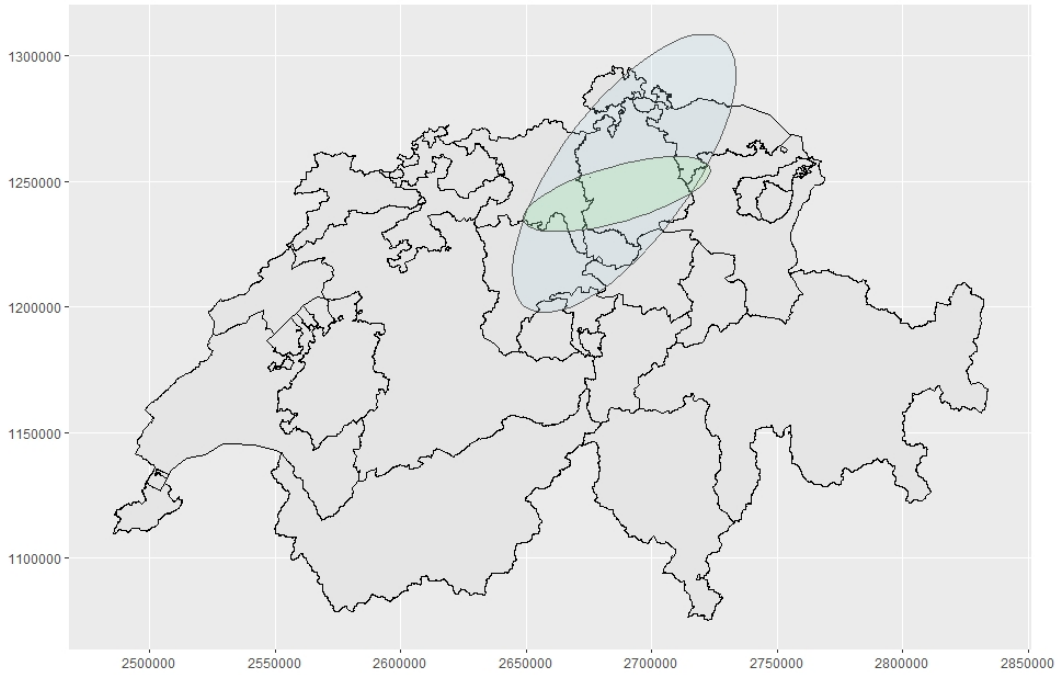


Activity spaces and behavioural innovation during COVID-19

Melanie Uhlmann

Supervisors:

Christopher Tchervenkov

Prof. Dr. Kay W. Axhausen

MSc Semester Project **June 2021**

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Activity spaces and behavioural innovation during COVID-19

Melanie Uhlmann

IVT
ETH Zürich
CH-8093 Zurich
muhlmann@student.ethz.ch

Supervisors:
Christopher Tchervenkov
Prof. Dr. Kay W. Axhausen
IVT
ETH Zurich
CH-8093 Zurich
christopher.tchervenkov@ivt.baug.ethz.ch
axhausen@ethz.ch

June 2021

Abstract

The goal of this report is to show the changes in people's mobility behaviour during COVID-19. It shows the distribution of activity spaces and innovation rates and sheds light on the socio-demographic aspects. User attributes such as age, income and household size are investigated and compared during the different phases of the COVID-19 pandemic. The results are based on the data from the MOBIS COVID-19 study. For the computation of the activity spaces and the innovation rates confidence ellipses and clustering methods were used. Further research on this topic is warranted to evaluate the results generated in this paper by more detail.

Keywords

Activity space, Innovation rate, 95%-confidence ellipses, socio-demographic attributes, MOBIS COVID study, COVID19

Suggested Citation

Uhlmann, M. (2021) Activity spaces and behavioural innovation during COVID-19, Semester project, Institut Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.

Activity spaces and behavioural innovation during COVID-19

Melanie Uhlmann

IVT
ETH Zürich
CH-8093 Zurich
muhlmann@student.ethz.ch

Supervisors:
Christopher Tchervenkov
Prof. Dr. Kay W. Axhausen
IVT
ETH Zurich
CH-8093 Zurich
christopher.tchervenkov@ivt.baug.ethz.ch
axhausen@ethz.ch

Juni 2021

Zusammenfassung

Dieser Bericht soll den Einfluss von COVID-19 auf das Mobilitätsverhalten von Personen aufzeigen. Es werden die Verteilungen der Aktivitätsräume sowie der Innovationsraten gezeigt und nach verschiedenen sozial demografischen Aspekte der Gesellschaft analysiert. Verschiedene Attribute wie Alter, Einkommen und Haushaltsgrösse wurden genauer betrachtet und über die verschiedenen Phasen der COVID-19 Pandemie verglichen. Die Resultat basieren auf den Daten der MOBIS COVID-19 Studie. Für die Berechnung der Aktivitätsräume und der Innovationsraten wurden Vertrauensellipsen und Cluster-Methoden verwendet. Für eine genauere Beurteilung der Resultate, welche in diesem Bericht generiert wurden, sind noch weitere Untersuchungen nötig.

Schlagworte

Activity space, Innovation rate, 95%-confidence ellipses, socio-demographic attributes, MOBIS COVID study, COVID19

Zitierungsvorschlag

Uhlmann, M. (2021) Activity spaces and behavioural innovation during COVID-19, Semester project, Institut Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.

Contents

List of Tables

List of Figures

1 Introduction

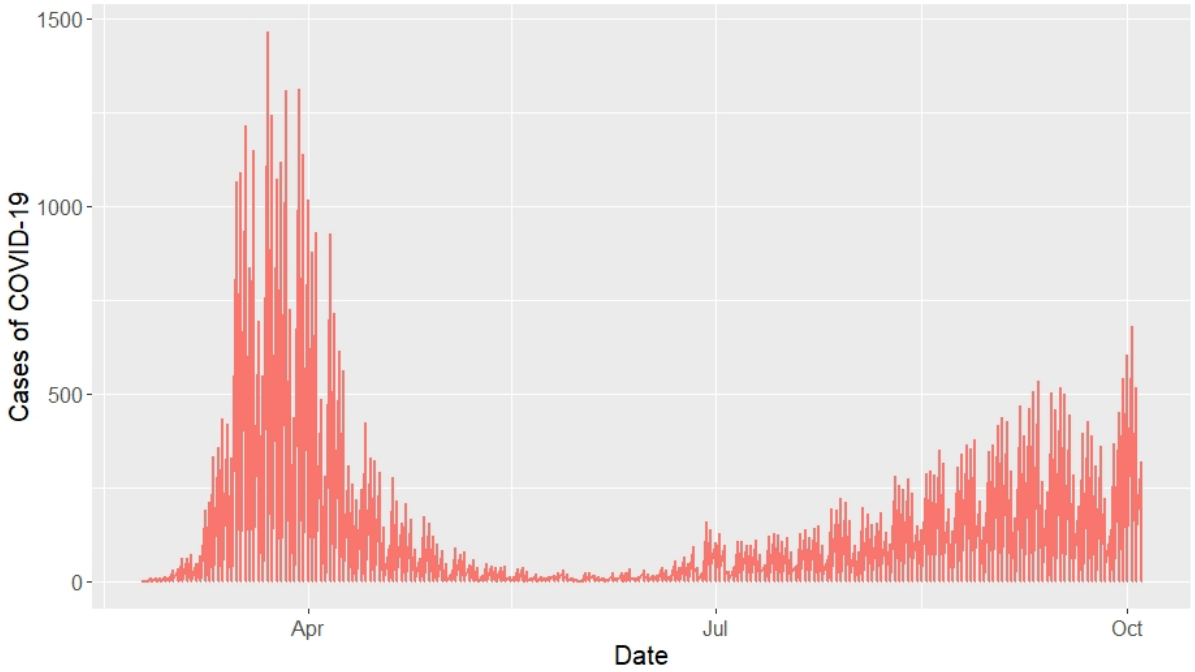
The COVID-19 pandemic has a large impact on the daily life of people all around the world. Countries took different measures to contain the virus. In a short time span borders were closed, shops and restaurants were not allowed to open and parents had to homeschool their children. Due to those measures the mobility behaviour of people was unrecognisable.
It is very challenging to understand the impact of the situation on mobility. Policies changed a lot and new policies were implemented. By plotting the activity spaces and approximating the innovation rate those changes in mobility behaviour are shown in this project.

1.1 COVID-19 timeline in Switzerland

On February 25, 2020 the first person who had COVID-19 in Switzerland was registered. The cases increased drastically and at the beginning of March over 400 people had been infected (Bühler, 2020).
On March 16, an extraordinary situation was declared by the Federal Council. Non-essential businesses as well as schools, recreational facilities and public parks were closed in order to contain the pandemic. Furthermore entry into Switzerland from other countries was only allowed for Swiss citizens, persons holding a residence permit for Switzerland and those who have to enter Switzerland for work-related reasons. Many people were working from home in order to avoid public transport. Where home-office was not possible the Federal Office of Public Health (FOPH) recommended to avoid rush hour travel. People who had symptoms of respiratory disease and people over 65 years of age were told not to use public transport. (Molloy *et al.*, 2021)
Although the Federal Council introduced many measures the number of COVID-19 cases increased rapidly. Hence new policies were introduced. Gatherings of more than 5 people were forbidden in public places and an inter-personal distance of 2 meters was mandated for groups with fewer than 5 people, but it was still allowed to go outside. In neighbouring countries as for example Austria, France and Italy people were allowed to leave the house only for limited purposes during the lockdown. On March 30, 2020 a peak of new COVID-19 cases was reached. 1,311 new infections were reported on that day and 58 people died in relation with COVID-19. After that the cases began to decrease. (Molloy *et al.*, 2021)
On April 27, 2020 hospitals were able to resume all medical procedures. As a result non-urgent procedures could be done again and also a first group of businesses was allowed

to re-open. The Federal Council had a strategy with three phases to emerge from the lockdown. The phases depended on the number of new infections, hospital admissions, deaths and on the hospital occupancy rates.
People who were feeling sick were advised to stay home in isolation. However in Switzerland there was no general obligation for healthy people to wear a mask. Keeping a distance of 2 meters and washing hands was seen as the most effective protective measures.(Molloy *et al.*, 2021)
On May 11, 2020 phase 2 was introduced by the Federal Council. In phase 2 most types of businesses were allowed to open as well as restaurants, museums, libraries and mandatory schools for students up to grade 9. Sport activities in small groups of up to 5 people were allowed but the rules on hygiene and social distancing still applied. The restrictions on entering Switzerland were relaxed and public transport was increasing its frequency. Furthermore the Federal Council introduced contact tracing for new infections with COVID-19. (Molloy *et al.*, 2021)
On June 6, 2020 phase 3 of the emergency plan was introduced. Events of up to 300 people were allowed as well as gatherings for up to 30 people. High schools and universities could resume. Leisure and entertainment businesses were able to reopen. At that point Switzerland reported a total number of cases of 30,988. 1,663 COVID-related deaths were reported and at that day there were only 16 new cases in Switzerland. A low level of cases, hospitalisations and deaths could be reached and stabilised. The extended power of the federal government expired on June 19, 2020.(Molloy *et al.*, 2021)
On June 22, 2020 most of the measures were lifted, only large events should not be allowed until at least the end of August. The minimum distance which must be kept between people was decreased from 2 to 1.5 meters, and where is was not possible to keep distance wearing a mask was strongly recommended.(Molloy *et al.*, 2021)
In mid-June the number of daily infections increased again and the Federal Council made it compulsory to wear a mask in public transport from July 6, 2020 (SRF News, 2020). The federal government also introduced a list with countries. When people visited a country on the list they had to quarantine for ten days upon their return.
An overview of the COVID-19 cases is shown on Figure 1 (Bundesamt für Gesundheit BAG, 2021). The time span shown is from February until the start of October 2020. The highest peak of this time span is in March. During June the cases are really low and they are increasing again in July.

Figure 1: COVID-19 cases in Switzerland from February - September 2020



Source: Bundesamt für Gesundheit BAG (2021)

1.2 Impacts of COVID-19 on mobility behaviour

Transportation is a key driver in the spread of infectious diseases (Baroyan and Rvachev 1967; Herrera-Valdez *et al.* 2011). Mobility plays a major role in a pandemic. This has been demonstrated for historical pandemics such as the Spanish flu in 1918 (Trilla *et al.*, 2008). Mobility is influenced by the measures to overcome pandemics. In the COVID-19 pandemic, especially during the lockdown, people tended to move less because stores, restaurants and other facilities were closed. As a result the daily life of the persons changed. The federal government suggested to minimise social contacts hence people did not meet much anymore. People were requested to stay within their household and therefore they spent more time at home.
There was a suggestion for home office. Therefore, persons who are working at a office mostly worked from home. This had a large impact on the mobility behaviour because the work place is a location which people go to most weekdays.

1.3 Activity spaces and innovation rates

To measure the mobility of the participants activity spaces and innovation rates are calculated and analysed. The activity space shows how far the participants moved whereas the innovation rates shown how many new places were visited during the observation period. Those values can be compared by grouping the participants. The participants are split by their age, by their income and by the household size they are living in.

2 Background

In this section the concept of the activity spaces and the innovation rates will be explained in order to get a better understanding.

2.1 Activity spaces

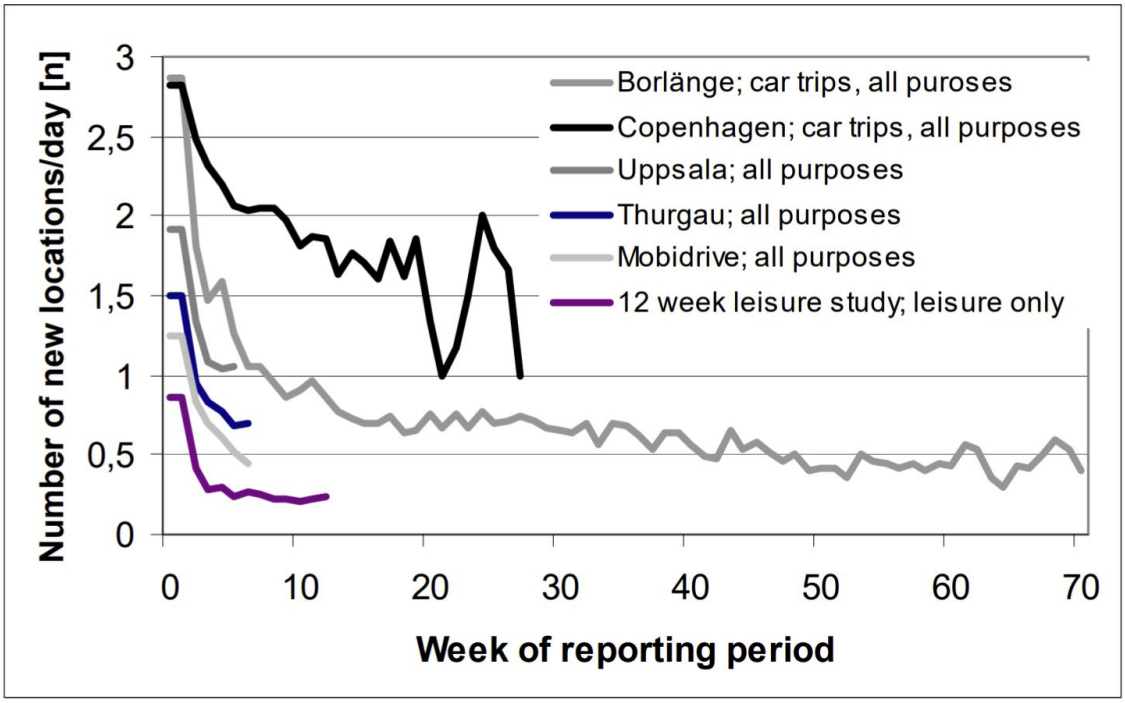
The first attempt to describe the space dimensions on human travel behaviour was made in the 1960s and 1970s and was called a traveller's personal world (Lynch *et al.*, 1960). This world of the participants was created based on a memory protocol which they created.
There were other attempts of graphically representing the space dimensions of human travel behaviour. People who created their own representation called their work differently and hence different terms were created. Those terms are awareness space (Brown and Moore, 1970), action space (Horton and Reynolds, 1971), space-time prisms (Lenntorp, 1977), mental maps (Muehrcke, 1978), perceptual space (Dürr, 1979) and activity repertoire or expectation space (Schönfelder and Axhausen, 2003). The term activity space was introduced by Golledge (1997). The three main determinants for this term are home, regular activities and the travel between and around these locations.
In order to estimate a person's activity space precise long-term observations are necessary. At that time the observations were handwritten diaries which the participants created. The same method was applied to the Mobidrive study (Axhausen *et al.*, 2002). To map a graphical representation, the locations which were visited by the participants were given a geocode.
With the introduction of GPS and Bluetooth it became possible to track people's movements automated and digitally. An early example of GPS-tracking was done in the study of Borlänge (Biding and Lind, 2002) (Schönfelder *et al.*, 2002). This study worked with geocode and hence it was possible to precisely estimate the activity spaces of the participants.
The activity spaces which are used today represent the distribution and allocation of different places. By recording a person's movement their travel pattern can be analysed. To improve the stability of the activity space the participants should be observed over a longer time frame.
With the new techniques available it was possible to calculate the activity spaces for single individuals, while previously only cross-sectional data for groups of respondents were depicted (Schönfelder and Axhausen, 2003).
Until the late 1990s, there were almost no studies which included the calculation and geographical representation of individual activity spaces. There was only a limited amount

of data in the Borlänge study meaning only car travel was included. Those findings show a variability on human activity spaces (Schönfelder and Samaga, 2003). In the Mobidrive study those results were investigated (Axhausen *et al.*, 2002) and additional analyses such as the clustering of activities and the interaction of activity density were performed. At that time it was already known how to compute the actual size of the activity spaces. There are three approaches to calculate the activity spaces of individuals namely a two-dimensional confidence ellipse, the kernel densities and a minimum spanning tree (network) or shortest-path network (Schönfelder and Axhausen, 2004).
In this project the activity spaces were approximated with the two-dimensional confidence ellipses and therefore the other two methods are not further considered after this point. The confidence ellipses were introduced in the 1970s (Zahavi, 1979) and were further investigated in the studies of Beckmann *et al.* (1983). While the concept of those is still used today, there were some changes and improvements which were made. The best centroid for the ellipses is not yet fully resolved. There are approaches taking the home location, the arithmetic mean of the geocoded data or merging the two main hubs within the data (Schönfelder and Axhausen, 2004). The calculation of the activity spaces is further explained in section 3.2.

2.2 Innovation rates

Innovation in relation with activity and locations is defined as new places. People visit new places on a regular basis. Whether to try a new restaurant or visit a museum one has not yet been, new places are added to a user's repertoire over time. So far few studies about innovation rates have been performed. This is due to a lack of available data. There are two studies, namely the Mobidrive data from 1999 and a study which was conducted in Thurgau in 2003 (Löchl *et al.*, 2005) which took a closer look into the innovation of the participants. Schlich *et al.* (2004) show that a regularity in travel behaviour over a observation period can be detected. Hence the assumption is that the innovation rate never drops to zero as the number of places which are to be visited is unlimited. With a higher number of trips the number of new locations increases (Schönfelder and Axhausen, 2004). The observed innovation rates for different studies is shown on Figure 2. The studies have shown that over a longer observation time the innovation rate decreases and reaches a plateau after some point.

Figure 2: Comparison of the innovation rates across studies



Source: Schönfelder and Axhausen (2004)

Research is showing that there are reasons for the assumption that the true innovation rate is lower than the values which are found in the studies. There are activities which are visited at a low frequency and hence might be detected as a new place even though they were already visited before. This could for example be a dentists office (Schönfelder and Axhausen, 2004). In the Thurgau studies (Löchl *et al.*, 2005) those places were taken into account by asking the participants whether they had visited the place before. This filtering led to a more realistic result of the innovation rates. However there are still some questions which come up by looking at the innovation rates in more detail. It still needs to be distinguished whether the locations are totally new to the participants or whether they have been visited at some point in the past. Also, the accuracy of the geodata needs to be improved by clustering them.

2.3 Overview of previous studies

There are several studies which have been conducted before in order to get a better insight in the spatial human behaviour (Schönfelder and Axhausen, 2010). The first survey which

was called Uppsala Household Travel Survey was done during 35 days, 144 participants were included and 23,000 trips were recorded. More studies followed which were conducted in different parts of the world. The most recent one was done in Atlanta and recorded 1,000,000 car trips. An overview of all the studies is given in Table 1.

Table 1: Summary of the conducted studies

Name of the survey*	**Year**	**Original focus**	**Locations(s)**	**Period**	**Resolution: geocoding**	**Resolution: purpses**	**Persons**	**Trips**
Uppsala Household Travel Survey	1971	Travel behaviour	Uppsala, Sweden	35 days	Building	All purposes	144	23'000
Mobidrive: Dynamics and routines of travel behaviour	1999	Stability of temporal patterns	Karlsruhe and Halle, Germany	42 days	Street block	All purposes	361	52'000
Borlänge GPS study (ISA Rätt Fart)	2000-2002	Speeding behaviour	Borlänge, Sweden	Up tp 80 weeks	Trip ends: GPS; unique locations: pre-defined clusters of trip ends	Unknown, potentially all	189 veh**	240'000 car trips
Leisure study (SVI Gesetzmässigkeiten des Wochenend-Freizeitverkehrs)	2002	Leisure travel behaviour and activities	Zürich, Switzerland	84 days	Post-code level	31 leisure purposes	75	9'900 leisure activities
Thurgau diary (SVI Study of the stability of transport behaviour)	2003	Stability of temporal patterns	Frauenfeld and villages in the Swiss canton of Thurgau	42 days	Building	All purposes	230	37'000
Copenhagen GPS study (AKTA Road Pricing Experiment in Copenhagen)	2001-2003	Route choice under road pricing	Copenhagen, Denmark	18-24 weeks	Trip ends: GPS; unique locations: pre-defined clusters of trip ends	Unknown, potentially all	500 veh.	250'000 car trips
Atlanta GPS study (Commute Atlanta Study)	2004-2006	Travel behaviour; test of policy measures such as pricing	Atlanta, USA	Up to two years	Trip ends: GPS; unique locations: pre-defined clusters of trip ends	Unknown, potentially all	Approx. 500 veh.	Approx. 1'000'000 car trips

*In the following, the data sets are simply titled Mobidrive, Thurgau, Uppsala, Borlänge, Copenhagen and Atlanta for better readability.
**Private cars only

Source: Rytz *et al.* (2020)

3 Methods

In this chapter a short summary of the MOBIS and the MOBIS COVID-19 studies is given and the methods which are used to approximate the activity spaces and the innovation rates are explained.

3.1 MOBIS and COVID-19 data

The MOBIS study (MObility Behaviour in Switzerland) started in September 2019. A sample of 5,375 persons who are living in Switzerland were recruited. With the help of an app their location was tracked and the trip stages and activities were identified. For a time span of 8 weeks the participants were tracked to investigate their response to a conceptual mobility pricing scheme. The last person finished in January and afterwards the participants could continue using the app. In March 2020 around 300 people of the MOBIS study were still using the app. For the MOBIS COVID-19 study people were recruted. There were 3,680 participants who completed the MOBIS study and around 1,600 of those volunteered to reactivate the tracking app. (Molloy *et al.*, 2021)
Since the start of the MOBIS COVID-19 study weekly reports have been produced (Molloy *et al.*, 2020). Those are available to the participants and the public in three languages (English, German and French). While using the app the participants also have access to a custom dashboard of their mobility behaviour. (Molloy *et al.*, 2021)

3.2 Calculation of activity spaces

The information in this section is based on the literature and implementations of Schönfelder and Axhausen (2004).

The activity space is a spatial representation of the space in which a certain participant moves. By taking all the activity geodata for one user a mean can be calculated. It is a conceptual point and is not necessarily a point the user visited. Before the activity spaces can be computed the activity coordinates need to be transformed to the Swiss reference system, projection code 2056. The method used to compute the confidence ellipse is the 95%-confidence ellipse (Zahavi, 1979).
The mean of all seen activity locations is represented by the centroid of the ellipse. It is weighted by the duration for each corresponding activity. In order to get the shapes of

the ellipses the covariance of the coordinates has to be determined. Those values can be calculated with the following formulas (Durandi and Künsch, 2015):

$$\mathrm{S} = \begin{pmatrix} s_{xx} & s_{xy} \\ s_{yx} & s_{yy} \end{pmatrix}$$

$$s_{xx} = \frac{1}{n-1} \sum_{i=1}^{n} w_i \cdot (x_i - \bar{x})^2$$

$$s_{yy} = \frac{1}{n-1} \sum_{i=1}^{n} w_i \cdot (y_i - \bar{y})^2$$

$$s_{xy} = s_{yx} = \frac{1}{n-1} \sum_{i=1}^{n} w_i \cdot (x_i - \bar{x})(y_i - \bar{y})$$

The weight of the locations is considered with the multiplication with w_i. The longer the participant stays at a certain location the higher is its presumed importance. By using the equations which are described above the confidence ellipse can be computed and therefore be plotted as well.
The size of the confidence ellipse can be calculated with the following formula:
$A = 6 \cdot \pi |S|^{1/2}$

3.3 Calculation of innovation rates

The innovation rate describes how many new places are visited during a specific time frame. The innovation of each participant and its rate must be found in order to apply a linear regression model. The computation of the innovation rates is based on the approach of Schönfelder and Axhausen (2004). The first step is to filter the participants. Those must have an observation time span of at least 52 days, because after the first few weeks a plateau rate is reached. In order to distinguish between the total innovation rate and this plateau rate a long-term observation time frame is needed.
The first step of the computation is to cluster the activities. This is important to detect locations and not only arbitrary points. Those clusters can be formed with the function DBSCAN (Density-based spatial clustering of applications with noise) in R. The points, which are defined, are core points, border points and outliers. Those clusters are computed within a radius of 50m. But this can be changed to any radius for which a cluster needs to be formed. The minimum number of points for a cluster is set to 3. This was chosen in order to prevent any location to be automatically treated as a new cluster.
After forming those clusters they can be arranged by date they had been visited and be assigned to an ID which links the date of the first visit to the location. The outlier

identification is the next step. Each new activity which is added starts as an outlier and over time it changes to a cluster when it is visited more often. Those outliers are assigned to an ID. The next step is to order the IDs chronologically. Hence it can be determined whether the IDs have been seen before or not. To find the new activities two sequential dates are compared. The logical value is set to 'TRUE' and an output is produces if a new location is detected. To compare the innovation rates of the participants the observation days must be standardised because the participants have started at different times. To compare the innovation rates of different groups of participants, the mean of each user can be calculated for a certain time span.

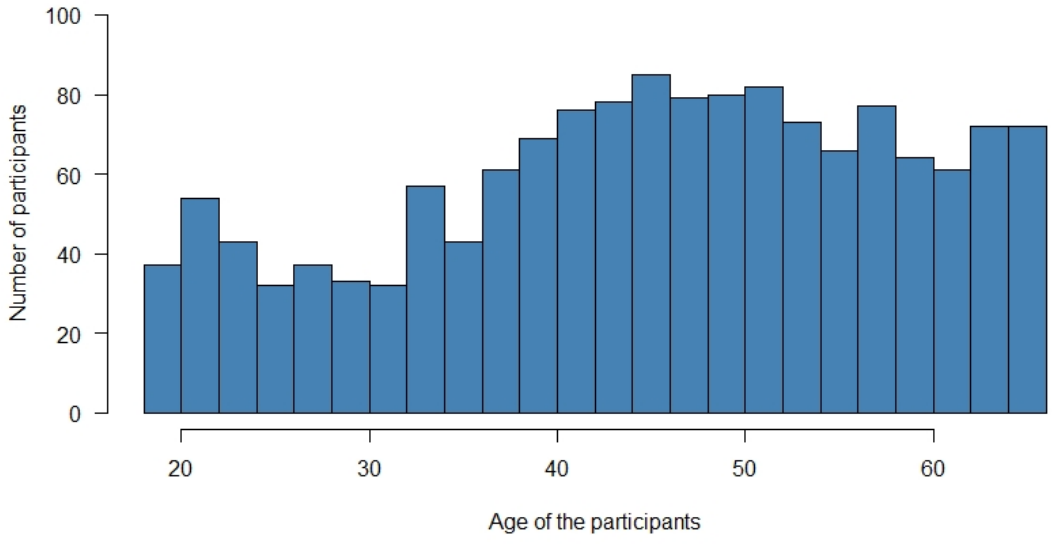
4 Results

In this reports not all results are shown. In the appendix the R-Code is provided to get a more detailed insight. Notice that not all outliers are shown in the box-plots in order to improve readability. The unit of the activity spaces are given in [km^2] and the innovation rates in [New activities/day].

4.1 Activity spaces

For the evaluation of the activity spaces the trips of 1,463 participants were analysed. Those participants have finished the intro survey and tracking can be analysed in more detail. The most important results of this intro survey are shown in this chapter.
Figure 3 shows the age distribution. The age of the participants varies from 19 - 66 years. There are fewer young people than people above 40.

Figure 3: Age distribution of the participants

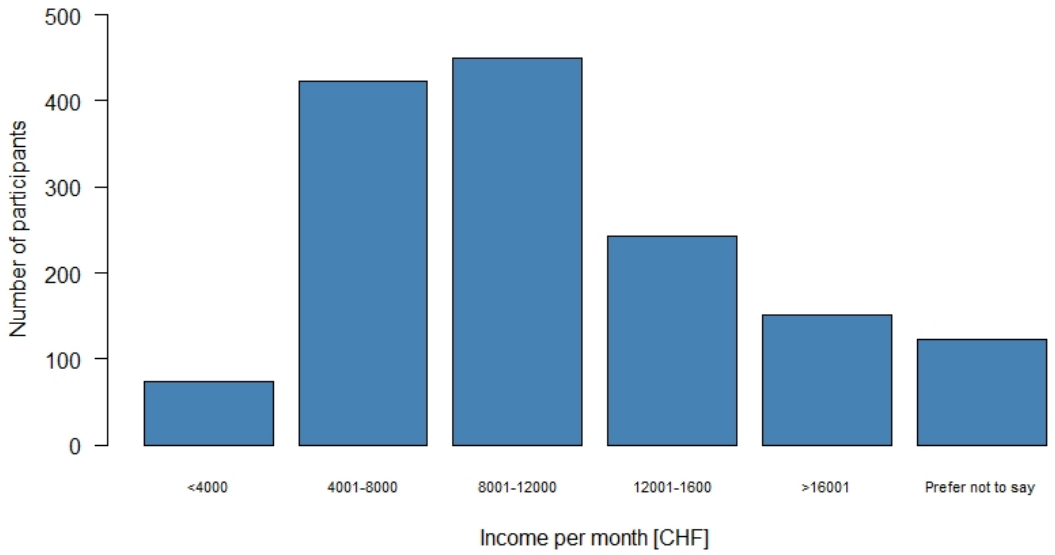


Source: MOBIS COVID-19 study

The income distribution of the participants is shown in Figure 4. Most participants earn between 4,000 and 12,000 CHF per month. There is also a group of participants who did

not want to tell their income in the survey.

Figure 4: Income of the participants



The household size in which the participants are living is shown in Figure 5. Most people live in a household with 2 persons and only a small number of participants are living in a large household with more than 5 persons.

Figure 5: Household size of the participants

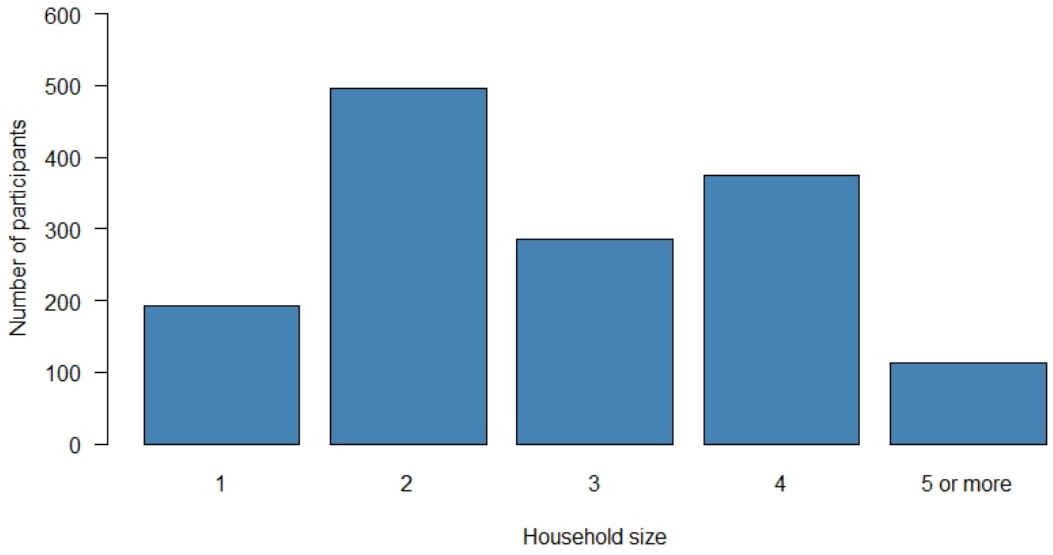
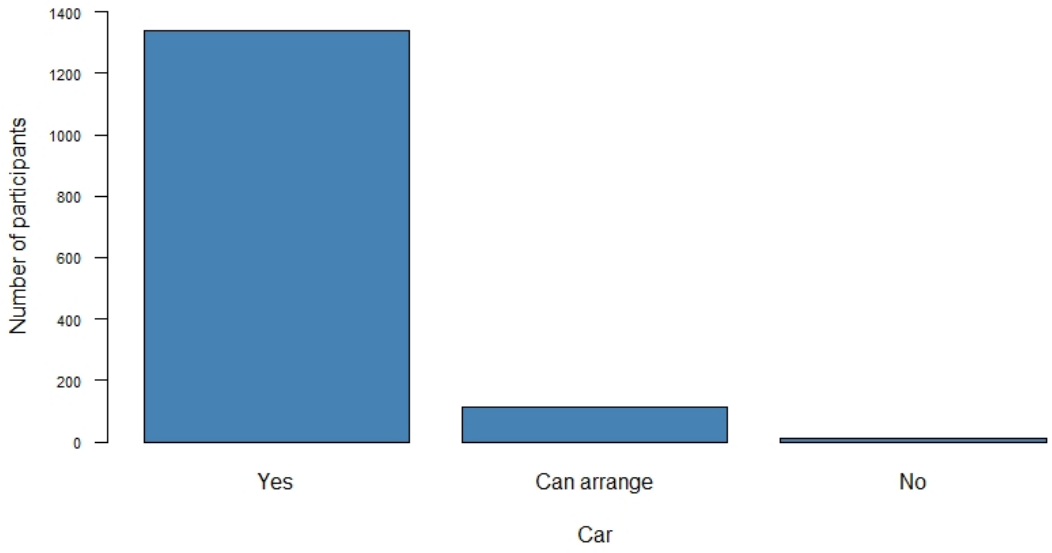


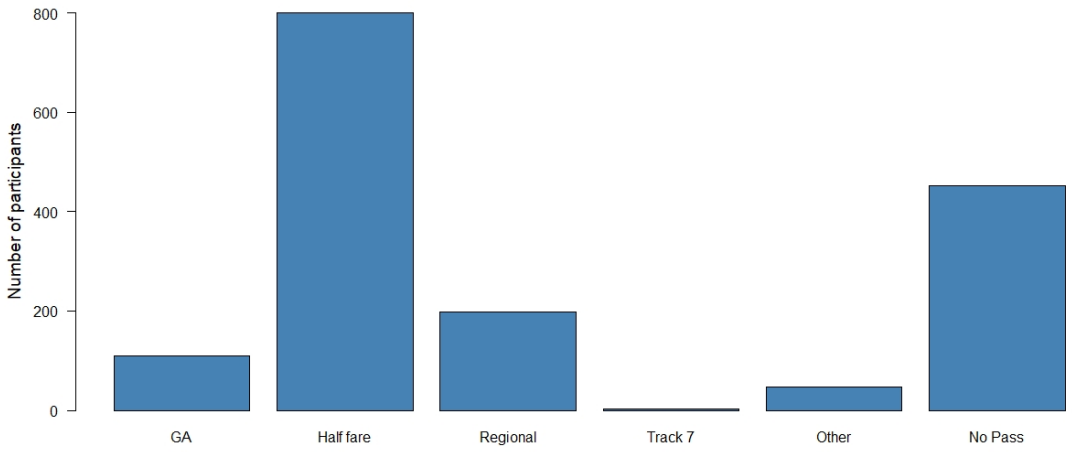
Figure 6 shows that most people who are participating in the study have a car or at least have access to a car. There is only an extremely low number of participants who do not have a car and no access to a car. It can be assumed that the study participants do most of their trips by car.

Figure 6: Car owners of the study



This assumption is supported by having a look at the travel passes which are owned by the participants. Figure 7 shows that most participants have either a half fare pass or no travel pass. With the half fare pass it is possible to travel for half the price on most public transport in Switzerland. This pass is worthwhile when making longer trips through Switzerland but not necessarily for travelling by public transport daily. A GA is only owned by a low number of participants. This travel pass allows to travel on most public transport in Switzerland for free.

Figure 7: Owners of travel passes



First the activity spaces were calculated for six phases which consist of six weeks with the first phase starting on February 3rd, 2020. Thus the first phase describes the situation before the COVID-19 outbreak. The second phase which starts on March 14 describes the behaviour of the participants during the lockdown in Switzerland. The third and the fourth phase last from end of April until mid of July. After the fourth phase masks became mandatory in public transport and this phase lasts until the end of August. The last phase describes the time from end of August until October 9, 2020. Table 2 shows a summary of those phases.

Table 2: Dates of the phases

	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Start date	01.02.20	14.03.20	25.04.20	06.06.20	18.07.20	29.08.20
End date	13.03.20	24.04.20	05.06.20	17.07.20	28.08.20	09.10.20

As the number of participants varies over the observation time, the numbers of participants have been calculated for the different phases. Table 3, 4 and 5 show the result of this analysis. During the first phase the number of participants is low for the whole study as the participants had to be recruited again. In the other phases the number of participants are higher and therefore more representative. The number varies a bit during the different phases but the differences are not as high as between the first two phases.

Table 3: Number of participants during the phases grouped by their age

	(10,20]	(20,30]	(30,40]	(40,50]	(50,60]	(60,Inf]
Phase 1	12	37	42	84	89	49
Phase 2	34	191	255	386	347	198
Phase 3	30	156	212	351	314	184
Phase 4	19	103	140	274	267	166
Phase 5	13	62	91	196	217	129
Phase 6	15	65	98	193	215	139

Table 4: Number of participants during the phases grouped by their income [CHF/month]

	< 4,000	4,000-8,000	8,001-12,000	12,001-16,000	>16,000	not said
Phase 1	14	95	84	51	41	28
Phase 2	72	404	437	233	148	117
Phase 3	59	353	382	209	137	107
Phase 4	42	271	292	170	114	80
Phase 5	34	196	218	110	93	57
Phase 6	33	202	236	114	95	58

Table 5: Number of participants during the phases grouped by their household size

	1	2	3	4	>5
Phase 1	34	105	68	81	25
Phase 2	188	480	277	356	110
Phase 3	166	426	241	314	100
Phase 4	126	341	173	251	78
Phase 5	97	264	128	170	49
Phase 6	103	263	133	174	52

For those different phases the activity spaces for all participants were calculated and are shown on Table 6. To analyse the trips only the one in Switzerland were taken into account. Otherwise large values would be created by people travelling to other countries. The mean is calculated by adding up all the calculated ellipse areas and dividing them by the number of summed up values while the median is calculated by putting the ellipse

areas in an ascending order and showing the value that is in the middle of the sample.

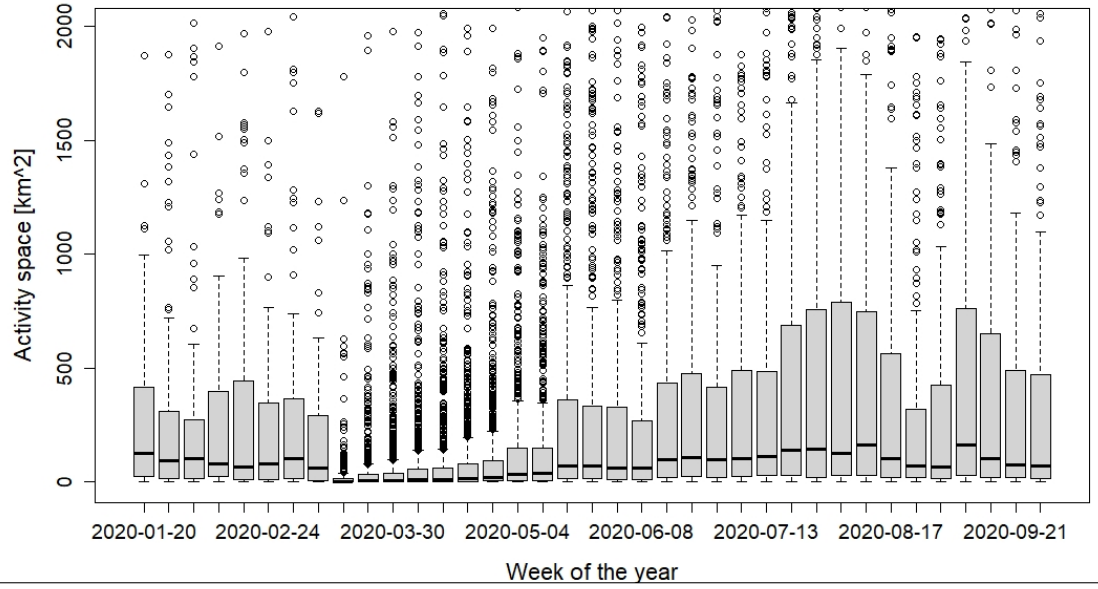
Table 6: Activity spaces for all participants

	Feb-Sep 20	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Mean	2'699	1'885	222	891	1'957	3'968	2'486
Median	1'164	374	29	170	357	796	855
Median/mean	0.43	0.20	0.13	0.19	0.18	0.20	0.34
Std. deviation	4'057	5'777	685	2'341	4'477	8'926	4'565
25%-quantile	370	76	6	41	96	175	227
75%-quantile	3'323	1'403	130	636	1'477	3'615	2'757

To analyse the activity spaces in more detail the rates were calculated weekly. With the weekly activity spaces the differences during the different phases of the pandemic can be seen in more detail.

Figure 8 shows the ellipse area over the whole study while the participants are not differentiated yet. The first COVID-19 case in Switzerland did not influence the participants a lot but as soon as the lockdown started the activity spaces were decreasing to a low value. After phase 2 of the emergency plan of the Federal Council the activity spaces increased again. There is a peak at the end of July and at the beginning of September.

Figure 8: Activity space over time for all the participants



Figures 9 shows the weighted mean of the activity spaces for the participants by age group. The weights are added for a better analysis between the different socio-demographic groups. For every age group the activity spaces decreases during the lockdown. Not every age group decreases their mobility at the same rate. After the lockdown the spaces are increasing but the age groups are adapting differently to the changes in policy which were introduced by the Federal Council.

Figure 9: Activity space over time for the different age groups

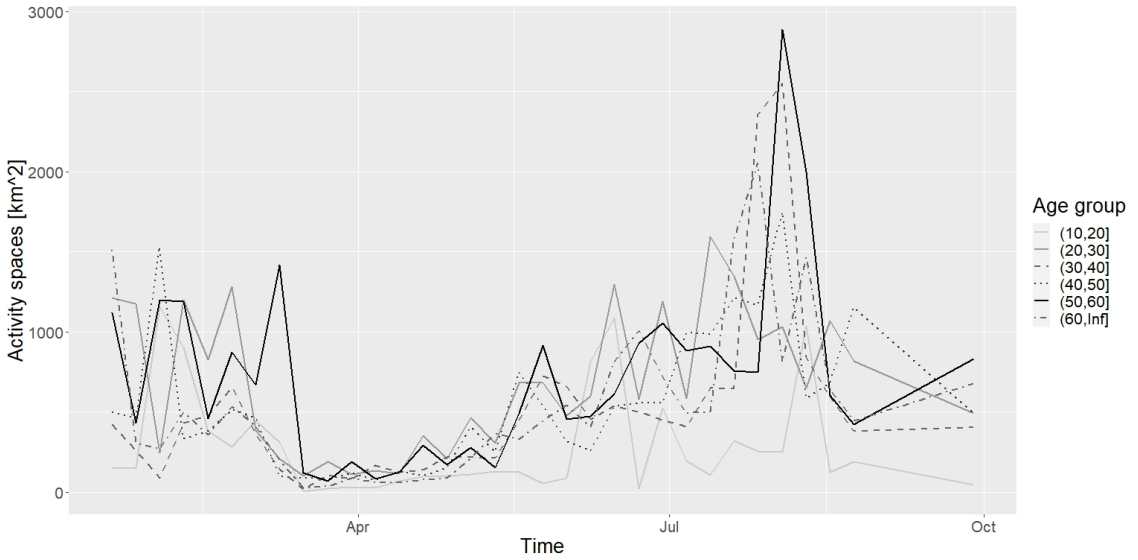


Figure 10 shows the weighted mean of the activity spaces for the different income groups. The activity space of people who earn between 8,001 and 12,000 are decreasing rapidly in March after having a high peak compared to the other income groups. People who preferred not to tell their income have two peaks which can be found rather at the end of the observation period.

Figure 10: Activity space over time for the different income groups

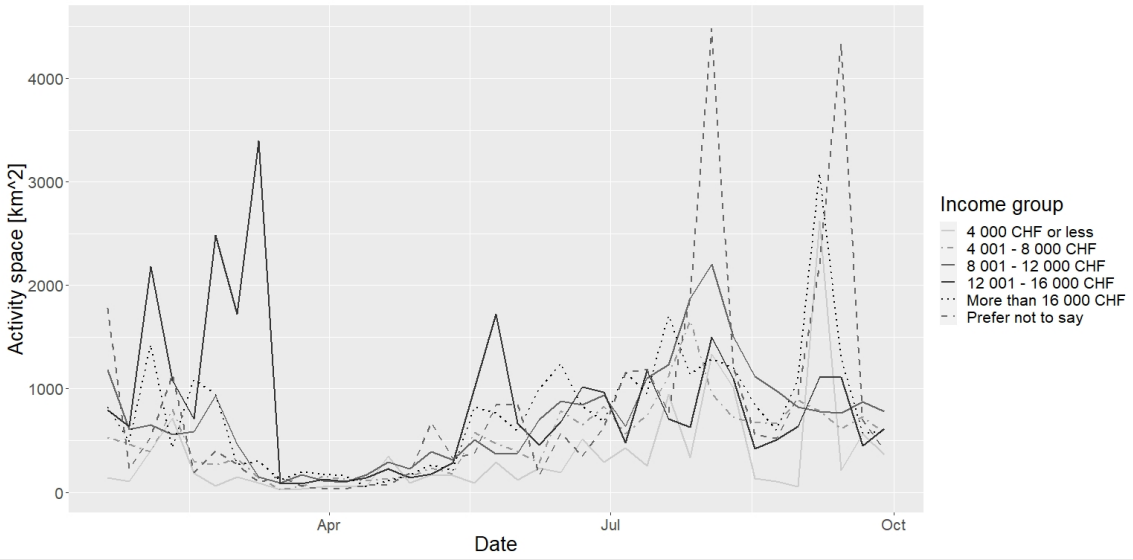
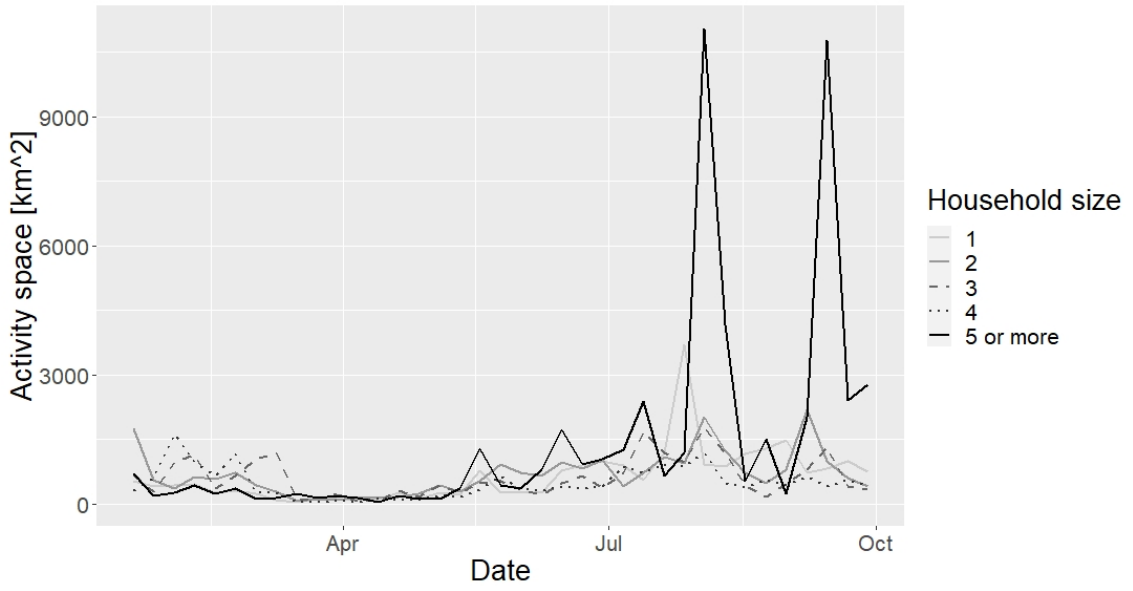


Figure 11 shows the activity spaces for the household sizes. The differences between the different household sizes are not that large compared to the other differentiations. There are two peaks for the largest households, one in August and one in September.

Figure 11: Activity space over time for the different household sizes



4.2 Innovation rate

The innovation rates were not analysed weekly but for the 6 weeks phases which are explained in chapter 4.1.
Figure 12 shows the innovation rate over the whole time span and for all participants. There are some high peaks of the innovation rate but also really low values.

Figure 12: Mean of the innovation rate over time for all participants

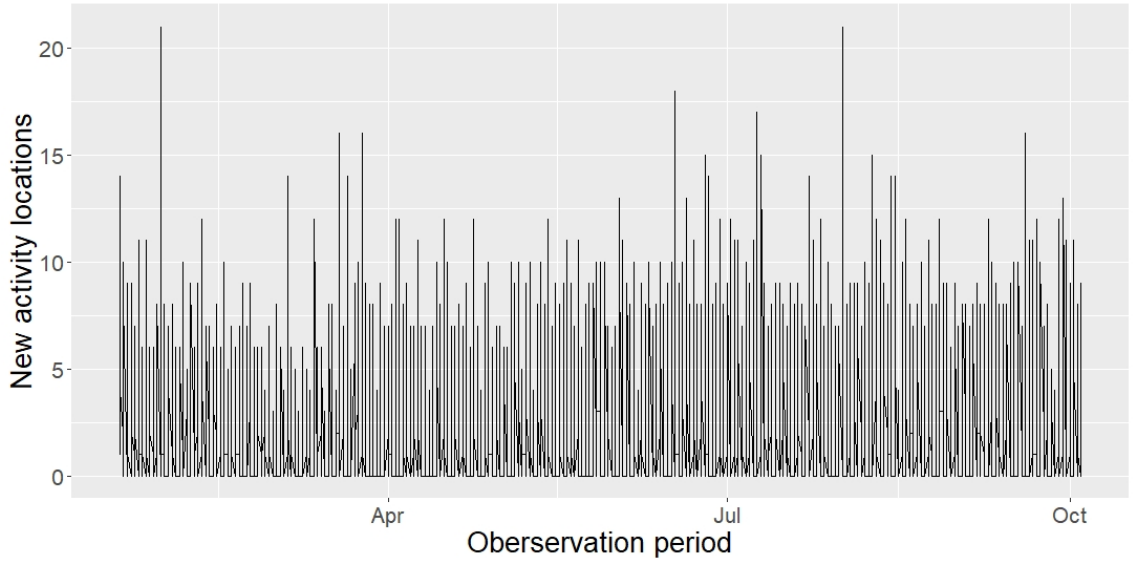
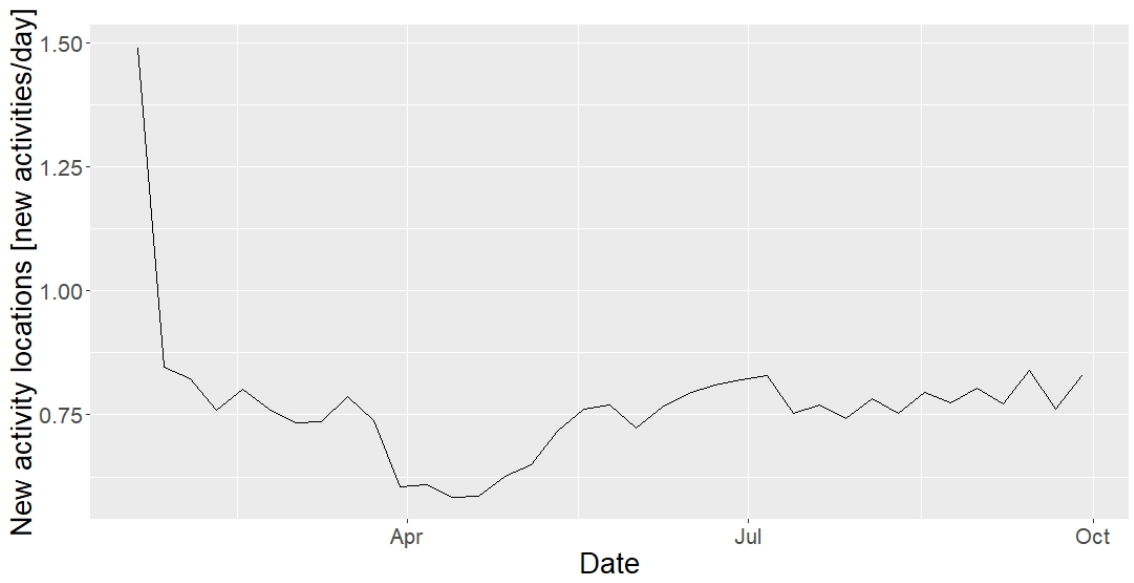


Figure 13 shows the weekly mean for innovation rate of all participants. Since there are a lot of zeros in the sample, it does not make sense to plot the median or to show a box plot for all the participants for the weekly values.

Figure 13: Mean of the innovation rate over time for all participants



In order to be able to calculate the innovation rates of the participants they need to record their trips over a certain amount of time, because first people are observed to see which places they usually visit. This is important to not track places they usually visit as new locations. Therefore the distributions of the participants were analysed separately from the distributions for the activity spaces.

Figure 14, 15 and 16 show the age distribution and the innovation rate for the different age groups. There are still more people above 35 years than young people. The innovation rate over the whole time does not show a huge difference for the different age groups. The age groups with more participants have more outliers. By looking at the innovation rate over the phases it is noticeable that the oldest age group has a innovation rate which is close to zero during the lockdown. Young people have a peak during phase 4 but the rate is decreasing again in phase 5. During phase 6 people aged 20 to 30 have the highest innovation rate.

Figure 14: Age distribution of the participants for the innovation rate

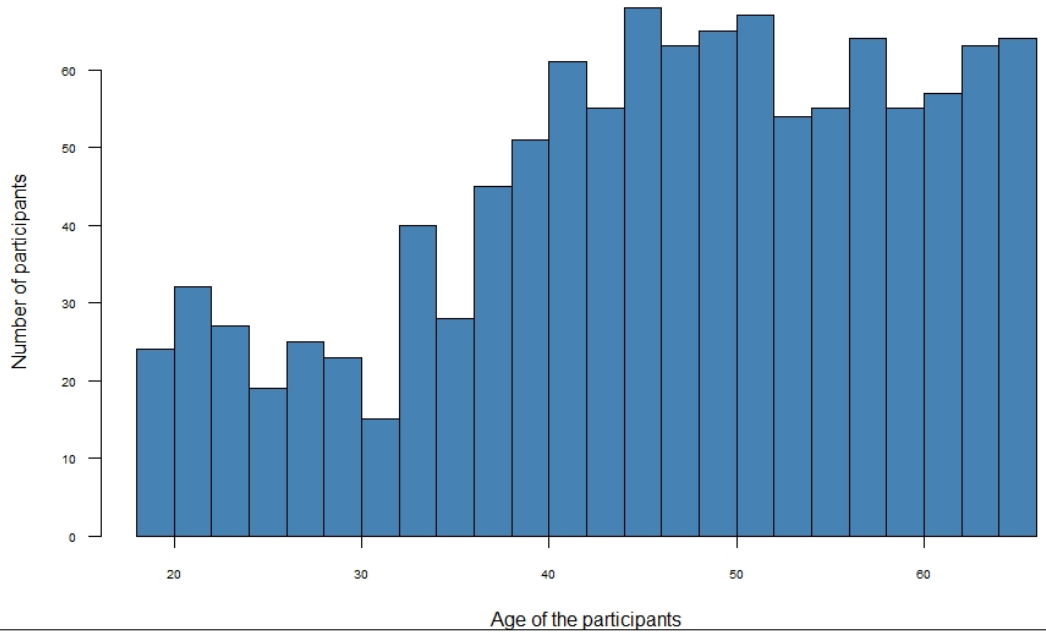


Figure 15: Innovation rate over the whole observation time for the different age groups

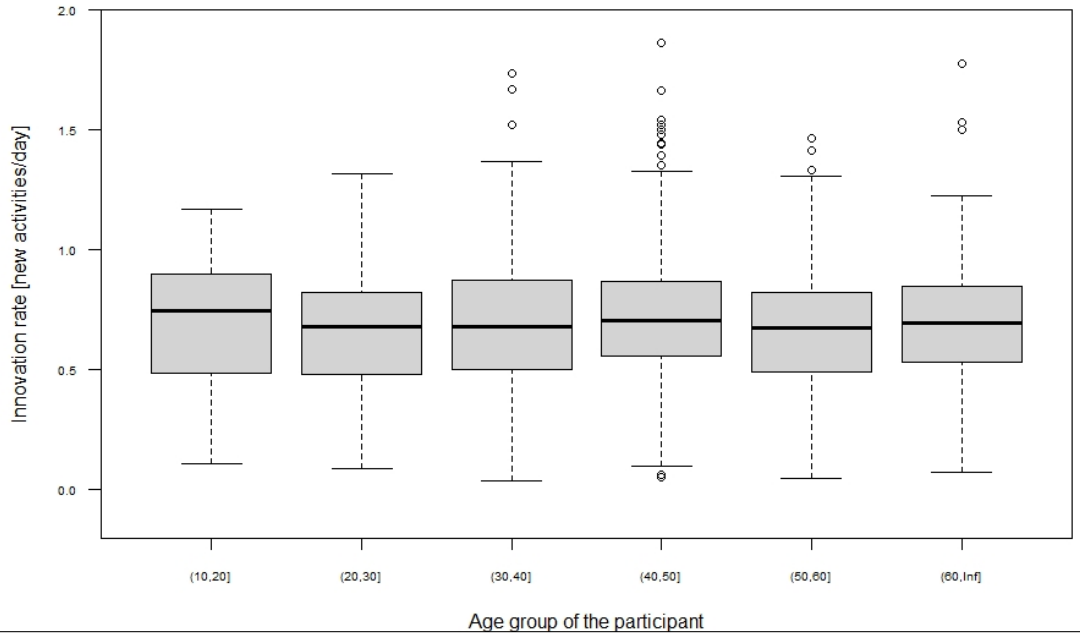
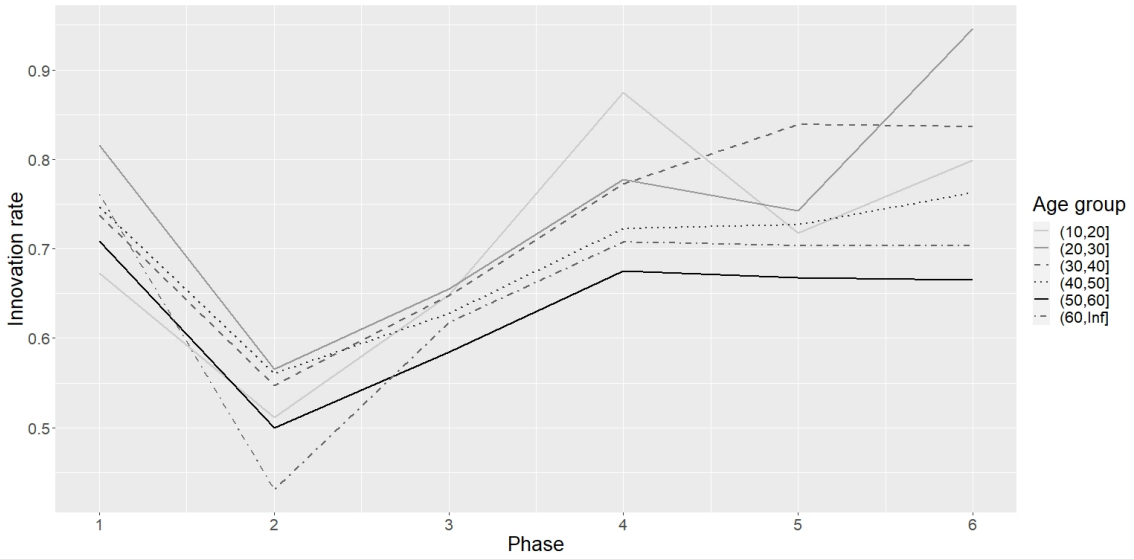


Figure 16: Weighted mean of the innovation rate for the age groups



The number of participants for the different income groups can be found below. The innovation rate over the whole time is shown in Figure 18 and for the different phases in Figure 19. The number of participants for the different income groups is not distributed evenly. There are many participants who earn between 4,000 and 12,000 CHF per month and less than 100 participants who earn less than 4,000 CHF.
Over the whole time all the income groups have a similar median, 25%- and 75%-quantile. Participants of the lowest income group have the highest innovation rate during all the phases of the observation period while the highest income group who earns more than 16,000 CHF per month has a rather low innovation rate.

Figure 17: Distribution of the participants for the different income groups

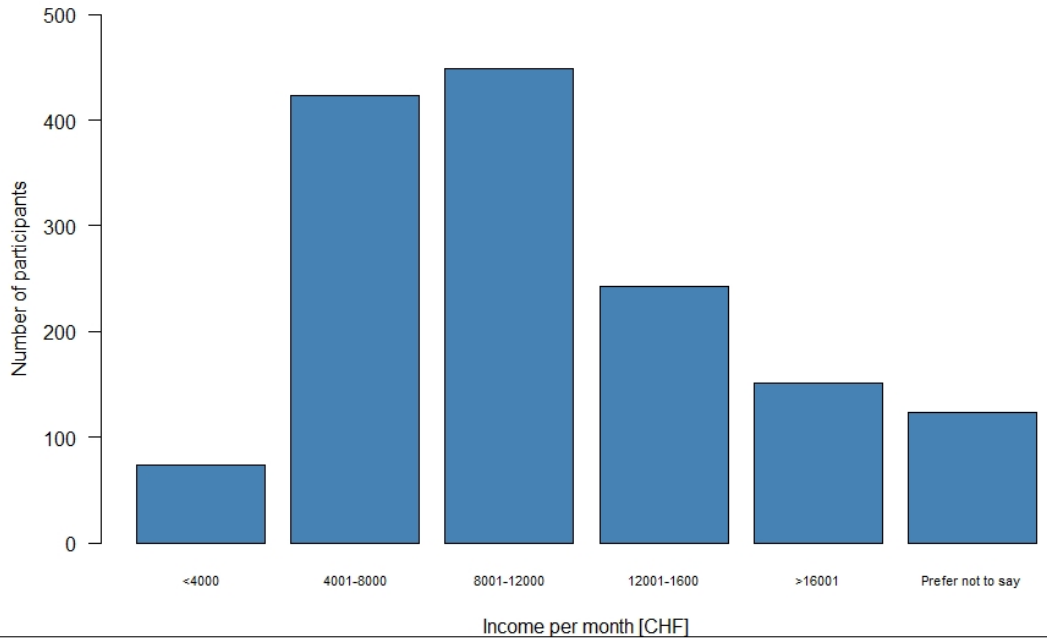


Figure 18: Innovation rate over the whole observation time for the different income groups

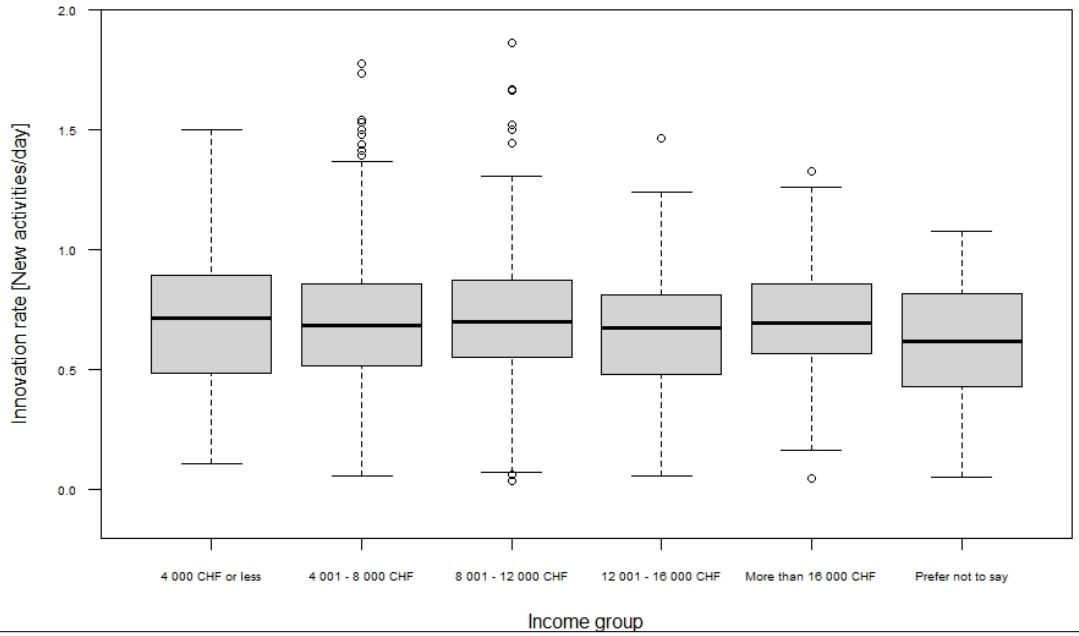


Figure 19: Weighted mean of the innovation rate for the income groups

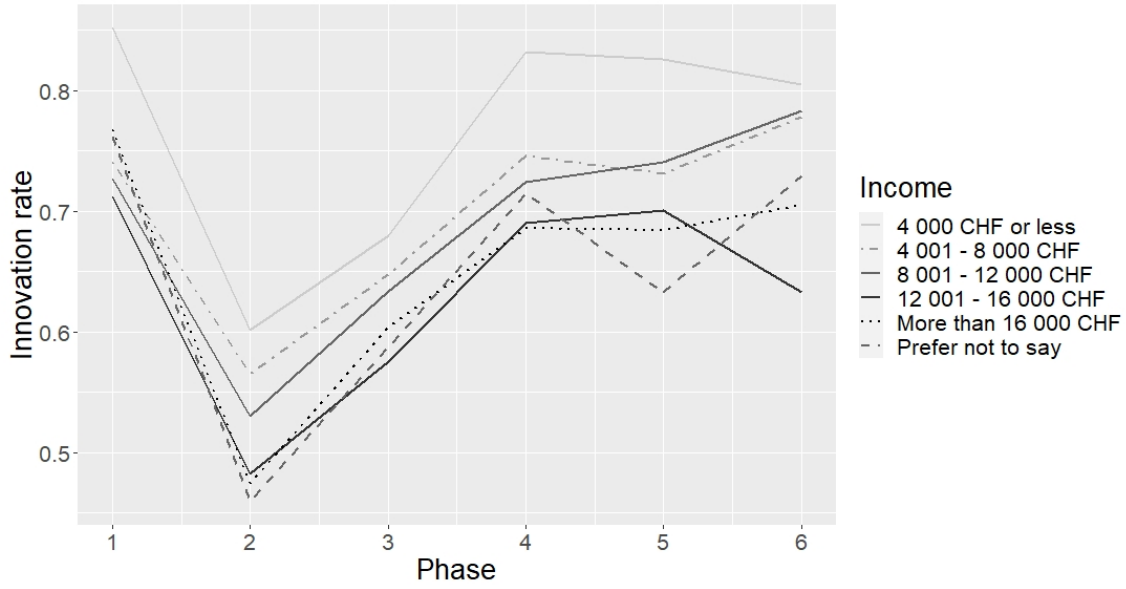


Figure 20 shows the distribution of the participants for the household sizes. The corresponding innovation rates can be found in Figure 21 and 22. Most people are living in a two-person household. There is only a small number of participants living in a household with 5 or more people. The box plots for the household sizes show similar values for all the sizes for the whole observation time. During the 6 phases households with two people have the lowest innovation rate while single people households have the highest innovation rate over almost the whole study time. Household with 5 or more persons have the highest peak during phase 4.

Figure 20: Distribution of the participants for the household sizes

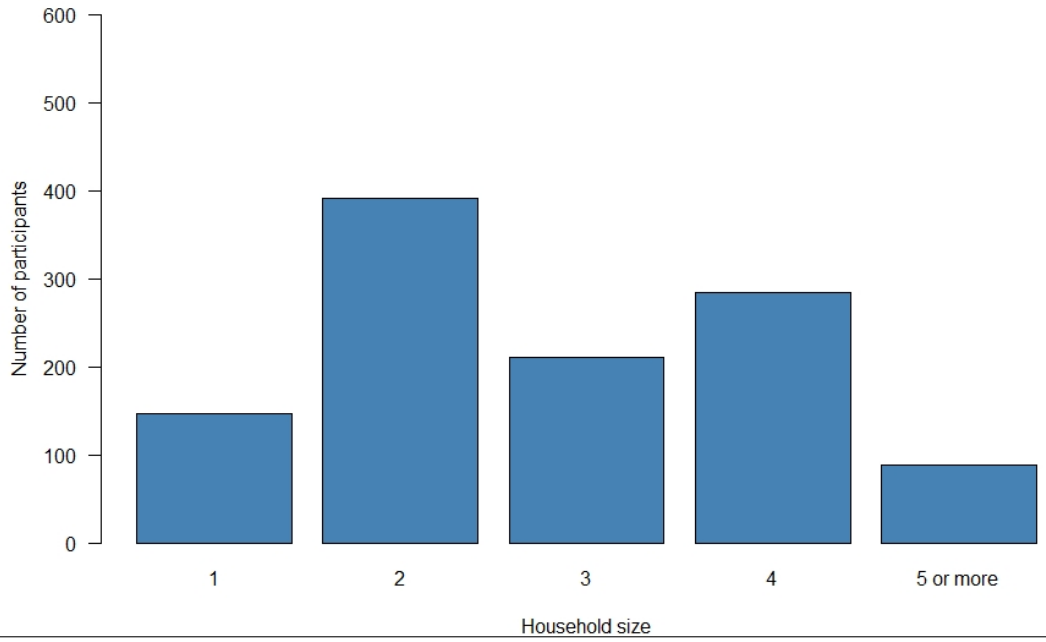


Figure 21: Innovation rate over the whole observation time for the household sizes

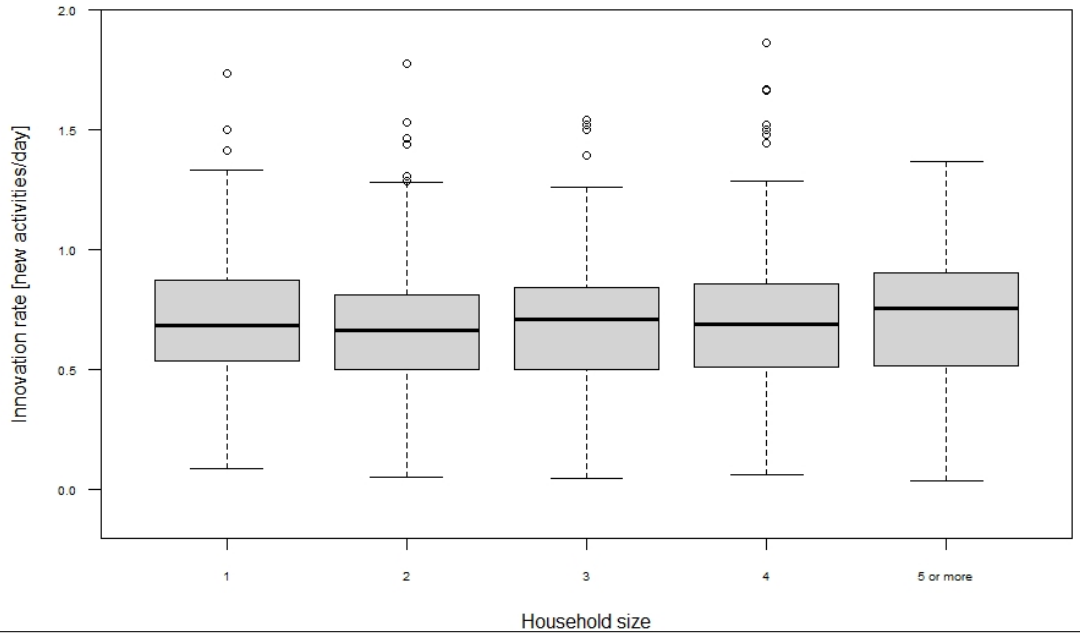
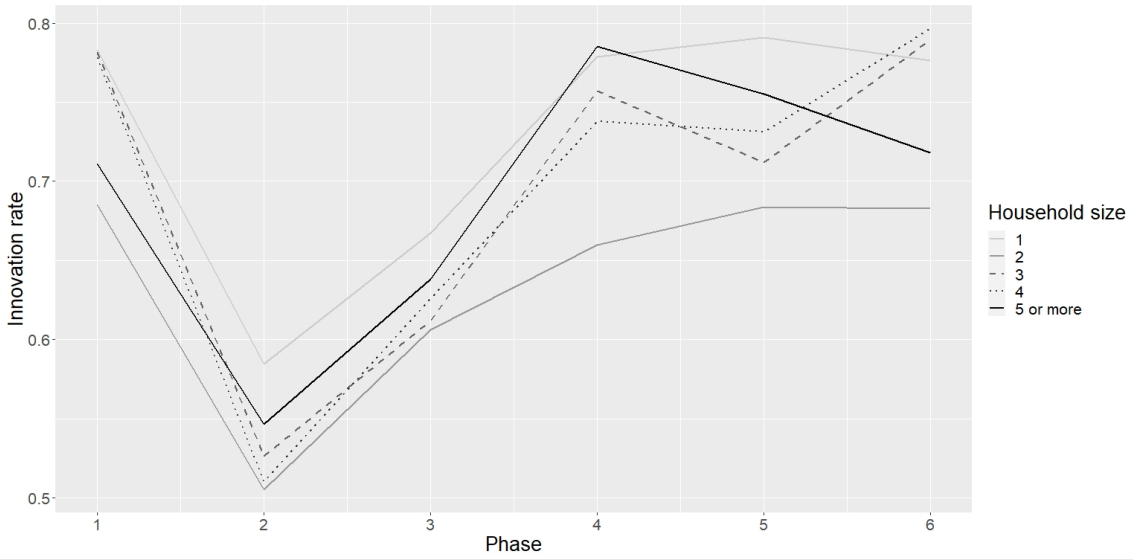


Figure 22: Weighted mean of the innovation rate for the household sizes



5 Discussion

In this section the results from chapter 4 are discussed in more detail and possible reasons for the behaviour of the people are explained.

5.1 Activity spaces

In general the activity spaces of all the people decreased during the lockdown. During this phase the median of the activity space as well as the 25% and the 75% quantiles are very low. In general people did not move a lot during this time hence there are a lot of outliers which are created by the people who were still travelling across Switzerland a lot. There is also a decrease in the activity space after mid August. At this time the COVID19-cases were slightly increasing again. This might be the reason why people did move less.
The ration between the median and the mean is the lowest during the lockdown. This means that the spread of the different activity spaces is not as large as in the other phases. The standard deviation also shows this fact. This means that the values are distributed more evenly than for the other phases where they have a lot of lower values but some really high values which increase the mean. The highest ratio is reached during phase 6 when the COVID-19 cases were increasing again. People behaved differently during this phase. Some people started to move a lot again whilst other people still tried to stay at home and avoid too much contact.

By differentiating the participants by their age, the different activities are visible. The youngest age group has the smallest activity space. The youngest age group consists of 37 participants over the whole observation time of which half are students who were no longer allowed to go to university and had their lectures online.
People who are older than 60 decreased their mobility in an early state. Those people were declared as a risk group at the beginning of the pandemic. People aged 50 to 60 have a high activity space until mid March but after this it decreases drastically. For this age group there is a peak in August. At that time the situation of the pandemic seemed to become better as the cases were only slightly increasing and therefore people started to increase their mobility.

The participants were also analysed according to their income. Most of the people gave information about their income but there was also a small group of participants who

preferred not to tell their monthly income. This group has very high peaks in autumn. However, as there is only a small number of participants in this group compared to the others a single participant has a higher influence on the weighted mean. Especially phase 5 and 6 which define the observation time in autumn have a low number of participants for this group. People who earn less than 4,000 CHF per month have the smallest activity space. Those people might not have much money to travel hence they do not move as far as people with a higher income.
The participants who are earning between 12,000 and 16,000 CHF per month did have a higher activity space during February and March than the other groups. People in that income group might have had some business meetings for which they needed to travel. But their activity space decreased rapidly during the lockdown and was never as high as in March again after the lockdown.
Participants with the highest income do not have a larger activity space than others. But notice that this group does not have as many participants as the groups with an income between 8,000 and 16,000 CHF. People with an income which is that high might not have as much time to travel than other participants.
The participant group who did not want to tell their income have two peeks at the end of the study period. There is one peak in August and one in September. Unfortunately this information does not tell much as their income is unknown. Also, one person who travels far can influence this value a lot because there are not many participants in this group.

The difference in the activity space of the different household sizes is not as large as in the other differentiations. Participants who live in a household with 5 or more people have two peaks: one in August and one in September. Those two peaks are really similar to the peaks in the income groups. People who preferred not to tell their income experience the same peaks as people living in a household with 5 or more people. There are 113 participants in total who live in a household with 5 or more people and only 11 of them did not want to tell their income. But the data shows that there is a participant with very high activity spaces who meets both criteria. This means it is possible that he is raising the mean of those groups. This probably happens in combination with other people who might have a large activity space. People living with 5 or more persons in a household live with their family or in a shared flat. Those might also make some trips together for which they travel across Switzerland.

5.2 Innovation rates

The innovation rate of all participants is the lowest in phase 2. This is due to the lockdown in Switzerland. In March the innovation rate decreases drastically. During this phase only grocery stores were allowed to stay open. Hence people did not have a chance to visit a lot of new places, but still people might change their supermarket because one is closer than the other or change it to one which is further away in order to be outside. Therefore the mean of the innovation rate is still not zero. The innovation rate stays low until May and then starts to increase again. After mid July the mean of the innovation rate is fluctuating because people behaved differently during that time. The cases were again increasing and people wanted to go out but the situation with COVID-19 was not clear. The innovation rate after the lockdown is quite high. However, the observation period starts at the begin of February and the lockdown was in March. This means that some places which the participants might usually visit are not recorded during the first weeks and hence are tracked as a new activity. This could for example be a cinema or a certain restaurant which are not new places for the participants but are still tracked as a new activity as they have not been seen before during the observation time.

In Figure 15 it is noticeable that the median of the innovation rate for the different age groups is very similar. But the innovation rate for the age group over time varies a lot and by splitting the rates for the different phases the innovation rate can be analysed in more detail.
By comparing the different age groups of the participants and the different phases it is noticeable that older people have a rather low innovation rate during the lockdown while the younger age groups have a high innovation rate. The elderly people were declared as a risk group by the Federal Council and were requested to stay at home. Therefore older people did not go to new places.
People who are between 30 and 40 years old have a low innovation rate during the lockdown compared to phases 5 and 6. Those people might be parents and they had to homeschool their children during the lockdown.
Young people are more likely to visit new places because they want to try new places even if it is just shopping at another supermarket than the usual one. The slope of the graph of the young people is steep. After the lockdown young people enjoyed to go out again. Also, most young people are not at risk when infected with COVID-19. The peak of the youngest participants is during phase 4. This can be explained by the summer break which most of them probably have. They have time to explore new locations during their break. After the steep slope of the innovation rate it reduces again. The innovation rate

decreases at the time of the introduction of the mask to public transport. Most of the young people of the study have an own car or at least access to a car, but this does not mean that they are doing all their trips with a car. The introduction of masks in public transport still can have an influence. At the beginning of the mask obligation people did not get used to it. It took some time for the people to adapt to this obligation.
After the lockdown people from 50 to 60 have the lowest innovation rate. This is possibly because they are the risk group. The Federal Council was still proposing for the old people to stay at home. The same reason applies to persons who are older than 60. They have a slightly higher innovation rate than the participants from 50 - 60 years but it is still lower than the rate of the other age groups.
During phase 6 the participants with an age from 20 until 30 years have the highest innovation rate. At that time events for up to 300 persons were allowed again. People at that age mostly like to go out and also try new places thus the mean of the innovation rate reaches nearly one.

Participants earning 4,000 CHF or less per month have the highest innovation rate over almost the whole time period. In this income group students are included and as they were no longer allowed to go to university they might have tried to go to new places. This income group also had the highest innovation rate before the pandemic. A possible reason is the age of the lowest income group there are a more young people than old people involved and young people tend to have a higher innovation rate.
People in the highest income group do not have a higher innovation rate than others. Those people often work a lot and therefore do not have as much time to explore new places as other participants might have.
People with a monthly income between 12,001 and 16,000 CHF have a decreasing innovation rate from phase 5 to phase 6. During this time meetings often were hold online again which might have an influence to this income group.

The innovation rate for the household sizes shows that people who are living in a single person household have the highest innovation rate. Those people often go out to see their friends and families and therefore they visit more new places than people from larger household sizes. Their innovation rate also decreased rapidly during the lockdown but increases again during phase 3 and 4. Participants who are living in a household with one other person have the lowest innovation rate. Those might be couples or close friends who enjoy their time at home and do not go out as much as a person who is living alone. The large household show a steep increase from phase 3 to phase 4. During the lockdown

they did not have the chance to explore new locations and they were requested to stay at home with their family or their roommates. They probably enjoy to go out again which might be the reason for the steep slope of the innovation rate.

6 Conclusion

In conclusion, the analysis of the activity spaces and the innovation rates shows that there are differences in mobility of the socio-demographic groups. Those can be observed and compared. It is also interesting that a low activity space does not always correspond to a low innovation rate. Especially young people and people with a low income show a lower activity space than other groups but still have a higher innovation rate. However, the innovation results should be analysed more thoroughly because the actual innovation rate can still be lower than the observed rate. Especially during the pandemic and the lockdown people did not always visit their usual places. Hence activities which were visited by the participants before the COVID-19 might be recorded as a new place in this project. It is also important to notice that there is a slight over representation of car owners in this study. This also has an influence of the travel behaviour of the participants. It is important to notice that during the first phase of the study the number of participants was lower than for the other phases. This mean that individuals have a higher influence on the mean of the activity spaces as well as for the innovation rates.

The research concerning human behaviour and spatial distribution of the individuals has just started and there are many possibilities to do further research on this topic. In order to be able mobility patterns of people as precisely as possible, more research in this area is warranted. This knowledge could play an important role in the future for the planning of public infrastructure.

Further research is also necessary to observe the behaviour of people during COVID-19. The time span of the study in this paper does only show the start of the COVID-19 timeline in Switzerland. The situation developed and therefore the mobility behaviour of people was still changing. To further analyse this travel behaviour is an important research topic.

7 Acknowledgement

I would like to express my gratitude to Prof. Dr. Kay W. Axhausen who was my supervisor for this project. He gave me the possibility to pursue this semester thesis. I also would like to express my thanks to Christopher Tchervenkov who was supporting me throughout this work. Both gave continuous inputs and feedback which were of great value for the project.
Furthermore, I would like to thank Thomas Schatzmann and Joseph Molloy who gave me support on the coding for this project.
Last but no least, I would like to express my thanks to my family and friends for their unconditional support. I specifically would like to thank Rico Schmuki and my flatmates who gave me their feedback on the manuscript and the presentation.

8 References

Axhausen, K. W., A. Zimmermann, S. Schönfelder, G. Rindsfüser and T. Haupt (2002) Observing the rhythms of daily life: A six-week travel diary, *Transportation*, **29** (2) 95–124.

Baroyan, O. and L. Rvachev (1967) Deterministic models of epidemics for a territory with a transport network, *Cybernetics*, **3** (3) 55–61.

Beckmann, M., T. Golob and Y. Zahavi (1983) Travel probability fields and urban spatial structure: 1. Theory, *Environment and Planning A*, **15** (5) 593–606.

Biding, T. and G. Lind (2002) Intelligent speed adaptation (isa): Results of large-scale trials in borlange, lidkoping, lund and umea during 1999-2002, *Technical Report*.

Brown, L. A. and E. G. Moore (1970) The intra-urban migration process: a perspective, *Geografiska Annaler: Series B, Human Geography*, **52** (1) 1–13.

Bundesamt für Gesundheit BAG (2021) Covid-19 Schweiz.

Bühler, C. (2020) Wie gefährlich ist das Coronavirus wirklich?, *Telebasel*.

Durandi, W. and H. R. Künsch (2015) Formeln, Tabellen, Begriffe: Mathematik-Physik-Chemie.

Dürr, H. (1979) *Planungsbezogene Aktionsraumforschung: theoretische Aspekte und eine empirische Pilotstudie*, Schroedel.

Golledge, R. G. (1997) *Spatial behavior: A geographic perspective*, Guilford Press.

Herrera-Valdez, M. A., M. Cruz-Aponte and C. Castillo-Chavez (2011) Multiple outbreaks for the same pandemic: Local transportation and social distancing explain the different" waves" of A-H1N1pdm cases observed in México during 2009, *Mathematical Biosciences & Engineering*, **8** (1) 21.

Horton, F. E. and D. R. Reynolds (1971) Effects of urban spatial structure on individual behavior, *Economic geography*, **47** (1) 36–48.

Lenntorp, B. (1977) Paths in space-time environments: A time-geographic study of movement possibilities of individuals, *Environment and Planning A*, **9** (8) 961–972.

Löchl, M., S. Schönfelder, R. Schlich, K. Axhausen, T. Buhl and P. Widmer (2005) Untersuchung der Stabilität des Verkehrsverhaltens.

Lynch, K. *et al.* (1960) *The image of the city*, vol. 11, MIT press.

Molloy, J., T. Schatzmann, B. Schoeman, C. Tchervenkov, B. Hintermann and K. W. Axhausen (2020) MOBIS: COVID-19, `https://ivtmobis.ethz.ch/mobis/covid19/`. Accessed: 2021-05-31.

Molloy, J., T. Schatzmann, B. Schoeman, C. Tchervenkov, B. Hintermann and K. W. Axhausen (2021) Observed impacts of the Covid-19 first wave on travel behaviour in Switzerland based on a large GPS panel, *Transport Policy*, **104**, 43–51.

Muehrcke, P. C. (1978) Maps in Minds: Reflections on Cognitive Mapping.

Rytz, S., J. Molloy and K. W. Axhausen (2020) Activity Spaces and Behavioral Innovation, Masterprojekt ETH Zürich.

Schlich, R., S. Schönfelder, S. Hanson and K. W. Axhausen (2004) Structures of leisure travel: temporal and spatial variability, *Transport Reviews*, **24** (2) 219–237.

Schönfelder, S. and K. W. Axhausen (2003) Activity spaces: measures of social exclusion?, *Transport policy*, **10** (4) 273–286.

Schönfelder, S. and K. W. Axhausen (2004) Structure and innovation of human activity spaces, *Arbeitsberichte Verkehrs-und Raumplanung*, **258**.

Schönfelder, S. and K. W. Axhausen (2010) *Urban rhythms and travel behaviour: spatial and temporal phenomena of daily travel*, Routledge.

Schönfelder, S., K. W. Axhausen, N. Antille and M. Bierlaire (2002) Exploring the potentials of automatically collected GPS data for travel behaviour analysis: A Swedish data source, *Arbeitsberichte Verkehrs-und Raumplanung*, **124**.

Schönfelder, S. and U. Samaga (2003) Where do you want to go today?: More observations on daily mobility, paper presented at the *STRC 2005 conference proceedings: 5th Swiss transport research conference: Monte Verità/Ascona, March 9-11, 2005*, vol. 179.

SRF News (2020) Maskenpflicht im ÖV: Der Bundesreat spricht ein Machtwort, *SRF*.

Trilla, A., G. Trilla and C. Daer (2008) The 1918 "Spanish flu" in Spain, *Clinical infectious diseases*, **47** (5) 668–673.

Zahavi, Y. (1979) The UMOT Project. A report prepared for the US Department of Transportation, Research and Special Programs Administration, and the Federal Republic of Germany Ministry of Transport, *Technical Report*, DOT-RSA–DPW-20-79-3.

9 Eigenständigkeitserklärung

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Eigenständigkeitserklärung

Die unterzeichnete Eigenständigkeitserklärung ist Bestandteil jeder während des Studiums verfassten Semester-, Bachelor- und Master-Arbeit oder anderen Abschlussarbeit (auch der jeweils elektronischen Version).

Die Dozentinnen und Dozenten können auch für andere bei ihnen verfasste schriftliche Arbeiten eine Eigenständigkeitserklärung verlangen.

Ich bestätige, die vorliegende Arbeit selbständig und in eigenen Worten verfasst zu haben. Davon ausgenommen sind sprachliche und inhaltliche Korrekturvorschläge durch die Betreuer und Betreuerinnen der Arbeit.

Titel der Arbeit (in Druckschrift):

Activity spaces and behavioural innovation during COVID-19

Verfasst von (in Druckschrift):
Bei Gruppenarbeiten sind die Namen aller Verfasserinnen und Verfasser erforderlich.

Name(n):	**Vorname(n):**
Uhlmann	Melanie

Ich bestätige mit meiner Unterschrift:

- Ich habe keine im Merkblatt „Zitier-Knigge" beschriebene Form des Plagiats begangen.
- Ich habe alle Methoden, Daten und Arbeitsabläufe wahrheitsgetreu dokumentiert.
- Ich habe keine Daten manipuliert.
- Ich habe alle Personen erwähnt, welche die Arbeit wesentlich unterstützt haben.

Ich nehme zur Kenntnis, dass die Arbeit mit elektronischen Hilfsmitteln auf Plagiate überprüft werden kann.

Ort, Datum	**Unterschrift(en)**
Zürich, 05.06.2021	M. Uhlmann

Bei Gruppenarbeiten sind die Namen aller Verfasserinnen und Verfasser erforderlich. Durch die Unterschriften bürgen sie gemeinsam für den gesamten Inhalt dieser schriftlichen Arbeit.

A R-Code

A.1 R-Code: Plot first page

```
#Plot map for the first page

library(dplyr)
library(sf)
library(car)
library(ggmap)
library(ggplot2)
library(lubridate)
library(sp)

setwd("C:/Users/melii/Desktop/Master Project Thesis/Data/Uhlmann")

#Import files of different phases
covid_activities <- read.csv("MOBIS_MDCEV/tracking/activities.csv", na="NA")


#combine the x and y value to one coordinate
activity_df_all <- st_as_sf(covid_activities, coords = c('x','y'),
                         crs = "+proj=longlat +datum=WGS84 +ellps=WGS84
                           +towgs84=0,0,0")
st_crs(activity_df_all)

#set the correct time for the data
activity_df_all$started_at <- as.POSIXct(activity_df_all$started_at
  ,tz=Sys.timezone())
activity_df_all$finished_at <- as.POSIXct(activity_df_all$finished_at
  ,tz=Sys.timezone())
activity_df_all$week <- format(activity_df_all$started_at, '%V')
activity_df_all$year <- format(activity_df_all$started_at, '%Y')
activity_df_all <- activity_df_all %>%
  filter(activity_df_all$year == '2020')
activity_df_all <- activity_df_all %>%
  filter(activity_df_all$week >= '04' & activity_df_all$week <= '40')

# ploting ellipses for a specific participant
activity_df_user <- activity_df_all %>%
  filter(user_id == 'AGHXU')
activity_df_user <- activity_df_user %>%
  filter(in_switzerland == 'TRUE')
covid_activities$started_at <- as.POSIXct(covid_activities$started_at
  ,tz=Sys.timezone())
```

```
covid_activities$week <- format(covid_activities$started_at, '%V')
covid_activities <- covid_activities %>%
  filter(covid_activities$week >= '04' & covid_activities$week <= '40')
covid_activities$year <- format(covid_activities$started_at, '%Y')
covid_activities <- covid_activities %>%
  filter(year == '2020')
coordinates <- covid_activities %>%
  filter(user_id == 'AGHXU') %>%
  filter(in_switzerland == 'TRUE')



coordinates$phase <- NULL
coordinates$user_id <- NULL
coordinates$activity_id <- NULL
coordinates$treatment <- NULL
coordinates$started_at <- NULL
coordinates$finished_at <- NULL
coordinates$duration <- NULL
coordinates$type <- NULL
coordinates$labelled_purpose <- NULL
coordinates$imputed_purpose <- NULL
coordinates$was_confirmed <- NULL
coordinates$in_switzerland <- NULL
coordinates$week <- NULL
coordinates$year<- NULL


#coordinates <- st_coordinates(coordinates$geometry) # for other user, change the
  user_id
centroid = colMeans(coordinates)
vcov_matrix = var(coordinates)

# or weighted by duration
weighted_cov = cov.wt(coordinates, activity_df_user$duration)
w_centroid <- weighted_cov$center
w_vcov_matrix <- weighted_cov$cov

# calculate the ellipse geometry
ellipse_geometry <- car::ellipse(center = centroid, shape = vcov_matrix,radius =
  sqrt(qchisq(.95, df=2)), draw=F)
w_ellipse_geometry <- car::ellipse(center = w_centroid, shape = w_vcov_matrix,radius
  = sqrt(qchisq(.95, df=2)), draw=F)

#convert this to an sf object so that we can plot it
el <- st_sfc(st_polygon(list(ellipse_geometry)), crs=st_crs(activity_df_user))
```

```
w_el <- st_sfc(st_polygon(list(w_ellipse_geometry)), crs=st_crs(activity_df_user))

kantons <-
  sf::st_read('BOUNDARIES_2021_04/DATEN/swissBOUNDARIES3D/SHAPEFILE_LV95_LN02/
swissBOUNDARIES3D_1_3_TLM_KANTONSGEBIET.shp') %>% st_transform(2056)


# plot the created confidence_ellipse (blue is unweighted, green is weighted)
ggplot(activity_df_user, datum=st_crs(4326)) +
  geom_sf(data = kantons, color = 'black') +
  geom_sf(data = activity_df_user$geometry, color = 'red', fill = 'red') +
  geom_sf(data=el, alpha=0.2, fill='lightblue') +
  geom_sf(data=w_el, alpha=0.2, fill='lightgreen') +
  coord_sf(crs=st_crs(2056), datum = sf::st_crs(2056))

ggplot() +
  geom_sf(data = kantons, color = 'black') +
  geom_sf(data = activity_df_user,
          inherit.aes = FALSE,
          colour = "#238443",
          fill = "#238443",
          alpha = 0.5,
          size = 6,
          shape = 21) +
  geom_sf(data = w_el, alpha = 0.2, fill = 'blue') +
  coord_sf(crs = st_crs(2056), datum = sf::st_crs(2056)) #important to display the
    map properly - try it without it.
```

A.2 R-Code: COVID-19 cases in Switzerland

```
  #Plot the COVID cases over time

library(readr)
library(dplyr)
library(ggmap)
library(ggplot2)
library(lubridate)

setwd("C:/Users/melii/Desktop/Master Project Thesis/Data/Uhlmann")

#read file
covid_cases <- read.csv('COVID19Cases_geoRegion.csv')
covid_cases$datum <- as.POSIXct(covid_cases$datum ,tz=Sys.timezone())
covid_cases$week <- format(covid_cases$datum, '%V')
covid_cases$year <- format(covid_cases$datum, '%Y')
covid_cases <- covid_cases %>%
  filter(covid_cases$year == '2020')
covid_cases <- covid_cases %>%
  filter(covid_cases$week >= '04' & covid_cases$week <= '40')
covid_cases$week_start <- floor_date(covid_cases$datum, unit = 'week', week_start
   =getOption('lubridate.week.start',1))

covid_cases_weekly <- covid_cases %>%
  group_by(week_start) %>%
  summarise(cases = sum(entries))


ggplot(data = covid_cases, aes(x=datum))+
  geom_line(aes(x=datum, y=entries, color ='1', linetype ='1'), size=1,
    show.legend=FALSE) +
  xlab('Date') +
  ylab('Cases of COVID-19') +
  theme(axis.title.x = element_text(size = 15), axis.text.x = element_text(size = 12),
        axis.title.y = element_text(size = 15), axis.text.y = element_text(size = 12))
```

A.3 R-Code: Participants

```
  #Participants
library(readr)
library(dplyr)
library(ggmap)
library(ggplot2)
library(lubridate)



#Import file of the survey
setwd("C:/Users/melii/Desktop/Master Project Thesis/Data/Uhlmann")
introSurvey_complete <-
  read.csv("csv/intro_survey_cleaned/introSurvey_complete.csv",na="NA")
study_participation_dates <- read.csv("study_participation_dates.csv",na="NA")
covid_activities <- read.csv("MOBIS_MDCEV/tracking/activities.csv", na="NA")
introSurvey_complete_mobis <-
  read.csv('MOBIS_MDCEV/surveys/mobis/intro_survey_cleaned/introSurvey_complete.csv')
introSurvey_complete_covid <-
  read.csv('MOBIS_MDCEV/surveys/covid/1_Initial/cleaned_mobis_covid_survey.csv')
introSurvey_complete <- bind_rows(introSurvey_complete_mobis,
  introSurvey_complete_covid)

#set the correct time for the data
covid_activities$started_at <- as.POSIXct(covid_activities$started_at
  ,tz=Sys.timezone())
covid_activities$finished_at <- as.POSIXct(covid_activities$finished_at
  ,tz=Sys.timezone())
covid_activities <- covid_activities %>% filter(covid_activities$started_at >=
  '2020-02-01' &
                                          covid_activities$finished_at <
                                            '2020-09-30')

#Filter participants who participated in COVID monitoring
corona_participants <- subset(study_participation_dates, coronavirus_registered_on !=
  'NULL')

#Filter the participants who have an activity space
participants_activities <- covid_activities %>%
  filter(user_id %in% unique(corona_participants$user_id)) %>%
  filter(in_switzerland == 'TRUE')

#Filter participants who completed the survey
introSurvey_completed <- subset(introSurvey_complete, intro_survey_finished == TRUE)

#Filter the participants in the intro survey based on the COVID study
```

```
participants <- introSurvey_completed %>%
  filter(participant_ID %in% unique(participants_activities$user_id))

#Plot the age distribution of the participants
participants$age <- as.numeric(as.character(participants$age))
hist(participants$age, breaks = 30, main = ' ', cex.main = 1, cex.axis = 1, cex.lab =
  1,
    xlab = 'Age of the participants', ylab = 'Number of participants', las=1,
      col='steelblue',
    ylim = c(0,100))

#Plot languages
language <- dplyr::count(participants, language)[2]
language_plot <- matrix(nrow=1,ncol=3)
language_plot[1] <- language[1,1]
language_plot[2] <- language[2,1]
language_plot[3] <- language[3,1]
language_plot <- as.numeric(language_plot)
barplot(language_plot, ylab='Number of participants', names = c('DE','EN','FR'),
  col='steelblue', las=1, cex.axis=0.7,
       ylim = c(0,1200))

#Plot the household size of the participants
householdsize <- dplyr::count(participants, household_size)[2]
householdsize_plot <- matrix(nrow=1,ncol=5)
householdsize_plot[1] <- householdsize[1,1]
householdsize_plot[2] <- householdsize[2,1]
householdsize_plot[3] <- householdsize[3,1]
householdsize_plot[4] <- householdsize[4,1]
householdsize_plot[5] <- householdsize[5,1]
householdsize_plot <- as.numeric(householdsize_plot)
barplot(householdsize_plot, names = c('1','2','3','4','5 or more'), las=1,
  col='steelblue', xlab='Household size',
       ylab='Number of participants', ylim=c(0,600))

#Plot the pulic transport tickets
pt_pass <- c(dplyr::count(participants, pt_pass_ga)[2,2],dplyr::count(participants,
  pt_pass_half_fare)[2,2],dplyr::count(participants,
  pt_pass_regional_pass)[2,2],dplyr::count(participants,
  pt_pass_track_7)[2,2],dplyr::count(participants,
  pt_pass_other)[2,2],dplyr::count(participants, pt_pass_no_pass)[2,2])
pt_pass <- as.numeric(pt_pass)
barplot(pt_pass, cex.main = 1, cex.axis = 1, cex.lab = 1.2, las = 1, col='steelblue',
  cex.names=1,
       names = c('GA', 'Half fare', 'Regional', 'Track 7', 'Other', 'No Pass'),
       ylab = 'Number of participants', ylim=c(0,800))
```

```
#Plot the income groups
income <- dplyr::count(participants, income)[2]
income_sorted <- matrix(nrow=1, ncol=6)
income_sorted[1] <- income[2,1]
income_sorted[2] <- income[3,1]
income_sorted[3] <- income[4,1]
income_sorted[4] <- income[1,1]
income_sorted[5] <- income[5,1]
income_sorted[6] <- income[6,1]
income_sorted <- as.numeric(income_sorted)
barplot(income_sorted, names =
  c('<4000','4001-8000','8001-12000','12001-1600','>16001', 'Prefer not to say'),
       col = 'steelblue', cex.names = 0.7, las=1, xlab = 'Income per month [CHF]',
         ylab = 'Number of participants',
       ylim=c(0,500))

#Plot education
education <- dplyr::count(participants, education)[2]
education_sorted <- matrix(nrow=1, ncol=3)
education_sorted[1] <- education[3,1]
education_sorted[2] <- education[2,1]
education_sorted[3] <- education[1,1]
education_sorted <- as.numeric(education_sorted)
barplot(education_sorted, names = c('Secondary','Mandatory', 'Higher education'),
  cex.names=0.7,col='steelblue',
       las=1, ylab='Number of participants', cex.axis=0.7, ylim=c(0,800))

#Plot work status
work_status <- c(dplyr::count(participants,
  work_status_employed)[2,2],dplyr::count(participants,
  work_status_self_employed)[2,2],dplyr::count(participants,
  work_status_unemployed)[2,2],dplyr::count(participants,
  work_status_apprentice)[2,2],dplyr::count(participants,
  work_status_student)[2,2],dplyr::count(participants,
  work_status_retired)[2,2],dplyr::count(participants, work_status_other)[2,2])
work_status <- as.numeric(work_status)
barplot(work_status, cex.axis = 0.7, cex.lab = 1, las = 1, col='steelblue',
  cex.names=0.4,
       names = c('Employed','Self
         employed','Unemployed','Apprentice','Student','Retired','Other'),
       ylab = 'Number of participants', ylim=c(0,1200))

#Plot workload main job
hist(participants$workload_jobs_main, ylab='Number of participants', xlab='Workload
  [%]',
```

```
    las=1, main='', cex.axis=0.7, col='steelblue', ylim= c(0,800))

#Plot the vehicle owners
car <- dplyr::count(participants, own_vehicles_car)[2]
car_sorted <- matrix(nrow=1, ncol=3)
car_sorted[1] <- car[3,1]
car_sorted[2] <- car[2,1]
car_sorted[3] <- car[1,1]
car_sorted <- as.numeric(car_sorted)
barplot(car_sorted, names = c('Yes','Can arrange', 'No'), col='steelblue', las=1,
       ylab='Number of participants', xlab='Car', cex.axis=0.7, ylim=c(0,1400))


#Plot the motorbike owners
motorbike <- dplyr::count(participants, own_vehicles_motorbike)[2]
motorbike_sorted <- matrix(nrow=1, ncol=3)
motorbike_sorted[1] <- motorbike[3,1]
motorbike_sorted[2] <- motorbike[2,1]
motorbike_sorted[3] <- motorbike[1,1]
motorbike_sorted <- as.numeric(motorbike_sorted)
barplot(motorbike_sorted, names = c('Yes','Can arrange', 'No'), col='steelblue',
  las=1,
       ylab='Number of participants', xlab='Motorbike', cex.axis=0.7, ylim=c(0,1400))

#Plot the bicycle owners
bicycle <- dplyr::count(participants, own_vehicles_bicycle)[2]
bicycle_sorted <- matrix(nrow=1, ncol=3)
bicycle_sorted[1] <- bicycle[3,1]
bicycle_sorted[2] <- bicycle[2,1]
bicycle_sorted[3] <- bicycle[1,1]
bicycle_sorted <- as.numeric(bicycle_sorted)
barplot(bicycle_sorted, names = c('Yes','Can arrange', 'No'), col='steelblue', las=1,
       ylab='Number of participants', xlab='Bicycle', cex.axis=0.7, ylim=c(0,1400))

#Plot the bike types
bike_type <- c(dplyr::count(participants,
  bike_type_regular)[2,2],dplyr::count(participants,
  bike_type_ebike_25)[2,2],dplyr::count(participants, bike_type_ebike_45)[2,2])
bike_type <- as.numeric(bike_type)
barplot(bike_type, cex.main = 1, cex.axis = 0.7, cex.lab = 1, las = 1,
  col='steelblue',
       names = c('Regular','E-Bike (25)', 'E-Bike (45)'), ylab = 'Number of
         participants',
       ylim=c(0,1400))
```

A.4 R-Code: Activity spaces for the 6-week phases

```
#Activity Space for all Participants

library(dplyr)
library(sf)
library(car)
library(ggmap)
library(ggplot2)
library(lubridate)
library(sp)

#Import files of different phases
weight <- read.csv('enrichments/ipf_weights_sample.csv')
covid_activities <- read.csv("MOBIS_MDCEV/tracking/activities.csv", na="NA")
activity_df <- st_as_sf(covid_activities, coords = c('x','y'),
                        crs = "+proj=longlat +datum=WGS84 +ellps=WGS84
                          +towgs84=0,0,0")
st_crs(activity_df)
activity_df_all$started_at <- as.POSIXct(activity_df_all$started_at
  ,tz=Sys.timezone())
activity_df_all$finished_at <- as.POSIXct(activity_df_all$finished_at
  ,tz=Sys.timezone())



#separate the phases
activity_df_phase1 <- activity_df %>%
  filter(started_at > '2020-02-03') %>%
  filter(finished_at < '2020-03-14')
activity_df_phase2 <- activity_df %>%
  filter(started_at > '2020-03-14 00:00:00') %>%
  filter(finished_at < '2020-04-25 00:00:00')
activity_df_phase3 <- activity_df %>%
  filter(started_at > '2020-04-25 00:00:00') %>%
  filter(finished_at < '2020-06-06 00:00:00')
activity_df_phase4 <- activity_df %>%
  filter(started_at > '2020-06-06 00:00:00') %>%
  filter(finished_at < '2020-07-18 00:00:00')
activity_df_phase5 <- activity_df %>%
  filter(started_at > '2020-07-18 00:00:00') %>%
  filter(finished_at < '2020-08-29 00:00:00')
activity_df_phase6 <- activity_df %>%
  filter(started_at > '2020-08-29 00:00:00')
activity_df_baseline <- activity_df %>%
  filter(started_at > '2019-11-02 00:00:00') %>%
```

```
  filter(finished_at < '2019-12-14 00:00:00')

#Filter the participants and the trips in Switzerland
weight <- weight %>%
  filter(participant_ID %in% unique(participants$participant_ID))
activity_df <- activity_df %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_baseline <- activity_df_baseline %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_phase1 <- activity_df_phase1 %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_phase2 <- activity_df_phase2 %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_phase3 <- activity_df_phase3 %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_phase4 <- activity_df_phase4 %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_phase5 <- activity_df_phase5 %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_phase6 <- activity_df_phase6 %>%
  filter(user_id %in% unique(participants$participant_ID)) %>%
  filter(in_switzerland == 'TRUE')


#Activity spaces for all the participants
calculate_ellipse_area <- function(activity_df) {
  coordinates <- st_coordinates(activity_df$geometry)
  weighted_cov = cov.wt(coordinates, activity_df$duration)
  w_centroid <- weighted_cov$center
  w_vcov_matrix <- weighted_cov$cov

  p <- ifelse(
    anyNA(w_vcov_matrix), NA, prod(eigen(w_vcov_matrix)$values)
  )
  ellipse_radius = ifelse(is.na(p) || p<0, NA, sqrt(p))
  area <- pi * 5.991 * ellipse_radius / 1000^2

  el <- car::ellipse(center = w_centroid, shape = w_vcov_matrix, radius =
    sqrt(qchisq(.95,df=2)), draw = F)
```

```
  mode(el) <- "integer"
  aa <- list(el)
  el_geom <- st_sfc(st_polygon(list(el)), crs=st_crs(activity_df))
  return (data.frame(
    centroid_X = w_centroid['X'], centroid_Y = w_centroid['Y'],
    cov_XX = w_vcov_matrix[1,1], cov_XY = w_vcov_matrix[1,2], cov_YX =
      w_vcov_matrix[2,1],
    cov_YY = w_vcov_matrix[2,2], ellipse_area = area, stringASFactors=F, geometry =
      el_geom))

}
calculate_ellipse_area_baseline <- function(activity_df_baseline) {
  coordinates_baseline <- st_coordinates(activity_df_baseline$geometry)
  weighted_cov_baseline = cov.wt(coordinates_baseline, activity_df_baseline$duration)
  w_centroid_baseline <- weighted_cov_baseline$center
  w_vcov_matrix_baseline <- weighted_cov_baseline$cov

p_baseline <- ifelse(
  anyNA(w_vcov_matrix_baseline), NA, prod(eigen(w_vcov_matrix_baseline)$values)
)
ellipse_radius_baseline = ifelse(is.na(p_baseline) || p_baseline<0, NA,
   sqrt(p_baseline))
area_baseline <- pi * 5.991 * ellipse_radius_baseline / 1000^2

el_baseline <- car::ellipse(center = w_centroid_baseline, shape =
   w_vcov_matrix_baseline, radius = sqrt(qchisq(.95,df=2)), draw = F)
mode(el_baseline) <- "integer"
aa_baseline <- list(el_baseline)
el_geom_baseline <- st_sfc(st_polygon(list(el_baseline)),
   crs=st_crs(activity_df_baseline))
return (data.frame(
  centroid_X_baseline = w_centroid_baseline['X'], centroid_Y_baseline =
    w_centroid_baseline['Y'],
  cov_XX_baseline = w_vcov_matrix_baseline[1,1], cov_XY_baseline =
    w_vcov_matrix_baseline[1,2],
  cov_YX_baseline = w_vcov_matrix_baseline[2,1],
  cov_YY_baseline = w_vcov_matrix_baseline[2,2], ellipse_area_baseline =
    area_baseline,
  stringASFactors=F, geometry_baseline = el_geom_baseline))
}
calculate_ellipse_area_phase1 <- function(activity_df_phase1) {
  coordinates_phase1 <- st_coordinates(activity_df_phase1$geometry)
  weighted_cov_phase1 = cov.wt(coordinates_phase1, activity_df_phase1$duration)
  w_centroid_phase1 <- weighted_cov_phase1$center
  w_vcov_matrix_phase1 <- weighted_cov_phase1$cov
```

```
  p_phase1 <- ifelse(
    anyNA(w_vcov_matrix_phase1), NA, prod(eigen(w_vcov_matrix_phase1)$values)
  )
  ellipse_radius_phase1 = ifelse(is.na(p_phase1) || p_phase1<0, NA, sqrt(p_phase1))
  area_phase1 <- pi * 5.991 * ellipse_radius_phase1 / 1000^2

  el_phase1 <- car::ellipse(center = w_centroid_phase1, shape = w_vcov_matrix_phase1,
     radius = sqrt(qchisq(.95,df=2)), draw = F)
  mode(el_phase1) <- "integer"
  aa_phase1 <- list(el_phase1)
  el_geom_phase1 <- st_sfc(st_polygon(list(el_phase1)),
     crs=st_crs(activity_df_phase1))
  return (data.frame(
    centroid_X_phase1 = w_centroid_phase1['X'], centroid_Y_phase1 =
      w_centroid_phase1['Y'],
    cov_XX_phase1 = w_vcov_matrix_phase1[1,1], cov_XY_phase1 =
      w_vcov_matrix_phase1[1,2],
    cov_YX_phase1 = w_vcov_matrix_phase1[2,1], cov_YY_phase1 =
      w_vcov_matrix_phase1[2,2],
    ellipse_area_phase1 = area_phase1, stringASFactors=F, geometry_phase1 =
      el_geom_phase1))
}
calculate_ellipse_area_phase2 <- function(activity_df_phase2) {
  coordinates_phase2 <- st_coordinates(activity_df_phase2$geometry)
  weighted_cov_phase2 = cov.wt(coordinates_phase2, activity_df_phase2$duration)
  w_centroid_phase2 <- weighted_cov_phase2$center
  w_vcov_matrix_phase2 <- weighted_cov_phase2$cov

  p_phase2 <- ifelse(
    anyNA(w_vcov_matrix_phase2), NA, prod(eigen(w_vcov_matrix_phase2)$values)
  )
  ellipse_radius_phase2 = ifelse(is.na(p_phase2) || p_phase2<0, NA, sqrt(p_phase2))
  area_phase2 <- pi * 5.991 * ellipse_radius_phase2 / 1000^2

  el_phase2 <- car::ellipse(center = w_centroid_phase2, shape = w_vcov_matrix_phase2,
     radius = sqrt(qchisq(.95,df=2)), draw = F)
  mode(el_phase2) <- "integer"
  aa_phase2 <- list(el_phase2)
  el_geom_phase2 <- st_sfc(st_polygon(list(el_phase2)),
     crs=st_crs(activity_df_phase2))
  return (data.frame(
    centroid_X_phase2 = w_centroid_phase2['X'], centroid_Y_phase2 =
      w_centroid_phase2['Y'],
    cov_XX_phase2 = w_vcov_matrix_phase2[1,1], cov_XY_phase2 =
      w_vcov_matrix_phase2[1,2],
    cov_YX_phase2 = w_vcov_matrix_phase2[2,1], cov_YY_phase2 =
```

```
      w_vcov_matrix_phase2[2,2],
    ellipse_area_phase2 = area_phase2, stringASFactors=F, geometry_phase2 =
      el_geom_phase2))
}
calculate_ellipse_area_phase3 <- function(activity_df_phase3) {
  coordinates_phase3 <- st_coordinates(activity_df_phase3$geometry)
  weighted_cov_phase3 = cov.wt(coordinates_phase3, activity_df_phase3$duration)
  w_centroid_phase3 <- weighted_cov_phase3$center
  w_vcov_matrix_phase3 <- weighted_cov_phase3$cov

  p_phase3 <- ifelse(
    anyNA(w_vcov_matrix_phase3), NA, prod(eigen(w_vcov_matrix_phase3)$values)
  )
  ellipse_radius_phase3 = ifelse(is.na(p_phase3) || p_phase3<0, NA, sqrt(p_phase3))
  area_phase3 <- pi * 5.991 * ellipse_radius_phase3 / 1000^2

  el_phase3 <- car::ellipse(center = w_centroid_phase3, shape = w_vcov_matrix_phase3,
    radius = sqrt(qchisq(.95,df=2)), draw = F)
  mode(el_phase3) <- "integer"
  aa_phase3 <- list(el_phase3)
  el_geom_phase3 <- st_sfc(st_polygon(list(el_phase3)),
    crs=st_crs(activity_df_phase3))
  return (data.frame(
    centroid_X_phase3 = w_centroid_phase3['X'], centroid_Y_phase3 =
      w_centroid_phase3['Y'],
    cov_XX_phase3 = w_vcov_matrix_phase3[1,1], cov_XY_phase3 =
      w_vcov_matrix_phase3[1,2],
    cov_YX_phase3 = w_vcov_matrix_phase3[2,1], cov_YY_phase3 =
      w_vcov_matrix_phase3[2,2],
    ellipse_area_phase3 = area_phase3, stringASFactors=F, geometry_phase3 =
      el_geom_phase3))
}
calculate_ellipse_area_phase4 <- function(activity_df_phase4) {
  coordinates_phase4 <- st_coordinates(activity_df_phase4$geometry)
  weighted_cov_phase4 = cov.wt(coordinates_phase4, activity_df_phase4$duration)
  w_centroid_phase4 <- weighted_cov_phase4$center
  w_vcov_matrix_phase4 <- weighted_cov_phase4$cov

  p_phase4 <- ifelse(
    anyNA(w_vcov_matrix_phase4), NA, prod(eigen(w_vcov_matrix_phase4)$values)
  )
  ellipse_radius_phase4 = ifelse(is.na(p_phase4) || p_phase4<0, NA, sqrt(p_phase4))
  area_phase4 <- pi * 5.991 * ellipse_radius_phase4 / 1000^2

  el_phase4 <- car::ellipse(center = w_centroid_phase4, shape = w_vcov_matrix_phase4,
    radius = sqrt(qchisq(.95,df=2)), draw = F)
```

```
  mode(el_phase4) <- "integer"
  aa_phase4 <- list(el_phase4)
  el_geom_phase4 <- st_sfc(st_polygon(list(el_phase4)),
    crs=st_crs(activity_df_phase4))
  return (data.frame(
    centroid_X_phase4 = w_centroid_phase4['X'], centroid_Y_phase4 =
      w_centroid_phase4['Y'],
    cov_XX_phase4 = w_vcov_matrix_phase4[1,1], cov_XY_phase4 =
      w_vcov_matrix_phase4[1,2],
    cov_YX_phase4 = w_vcov_matrix_phase4[2,1], cov_YY_phase4 =
      w_vcov_matrix_phase4[2,2],
    ellipse_area_phase4 = area_phase4, stringASFactors=F, geometry_phase4 =
      el_geom_phase4))
}
calculate_ellipse_area_phase5 <- function(activity_df_phase5) {
  coordinates_phase5 <- st_coordinates(activity_df_phase5$geometry)
  weighted_cov_phase5 = cov.wt(coordinates_phase5, activity_df_phase5$duration)
  w_centroid_phase5 <- weighted_cov_phase5$center
  w_vcov_matrix_phase5 <- weighted_cov_phase5$cov

  p_phase5 <- ifelse(
    anyNA(w_vcov_matrix_phase5), NA, prod(eigen(w_vcov_matrix_phase5)$values)
  )
  ellipse_radius_phase5 = ifelse(is.na(p_phase5) || p_phase5<0, NA, sqrt(p_phase5))
  area_phase5 <- pi * 5.991 * ellipse_radius_phase5 / 1000^2

  el_phase5 <- car::ellipse(center = w_centroid_phase5, shape = w_vcov_matrix_phase5,
    radius = sqrt(qchisq(.95,df=2)), draw = F)
  mode(el_phase5) <- "integer"
  aa_phase5 <- list(el_phase5)
  el_geom_phase5 <- st_sfc(st_polygon(list(el_phase5)),
    crs=st_crs(activity_df_phase5))
  return (data.frame(
    centroid_X_phase5 = w_centroid_phase5['X'], centroid_Y_phase5 =
      w_centroid_phase5['Y'],
    cov_XX_phase5 = w_vcov_matrix_phase5[1,1], cov_XY_phase5 =
      w_vcov_matrix_phase5[1,2],
    cov_YX_phase5 = w_vcov_matrix_phase5[2,1], cov_YY_phase5 =
      w_vcov_matrix_phase5[2,2],
    ellipse_area_phase5 = area_phase5, stringASFactors=F, geometry_phase5 =
      el_geom_phase5))
}
calculate_ellipse_area_phase6 <- function(activity_df_phase6) {
  coordinates_phase6 <- st_coordinates(activity_df_phase6$geometry)
  weighted_cov_phase6 = cov.wt(coordinates_phase6, activity_df_phase6$duration)
  w_centroid_phase6 <- weighted_cov_phase6$center
```

```
  w_vcov_matrix_phase6 <- weighted_cov_phase6$cov

  p_phase6 <- ifelse(
    anyNA(w_vcov_matrix_phase6), NA, prod(eigen(w_vcov_matrix_phase6)$values)
  )
  ellipse_radius_phase6 = ifelse(is.na(p_phase6) || p_phase6<0, NA, sqrt(p_phase6))
  area_phase6 <- pi * 5.991 * ellipse_radius_phase6 / 1000^2

  el_phase6 <- car::ellipse(center = w_centroid_phase6, shape = w_vcov_matrix_phase6,
    radius = sqrt(qchisq(.95,df=2)), draw = F)
  mode(el_phase6) <- "integer"
  aa_phase6 <- list(el_phase6)
  el_geom_phase6 <- st_sfc(st_polygon(list(el_phase6)),
    crs=st_crs(activity_df_phase6))
  return (data.frame(
    centroid_X_phase6 = w_centroid_phase6['X'], centroid_Y_phase6 =
      w_centroid_phase6['Y'],
    cov_XX_phase6 = w_vcov_matrix_phase6[1,1], cov_XY_phase6 =
      w_vcov_matrix_phase6[1,2],
    cov_YX_phase6 = w_vcov_matrix_phase6[2,1], cov_YY_phase6 =
      w_vcov_matrix_phase6[2,2],
    ellipse_area_phase6 = area_phase6, stringASFactors=F, geometry_phase6 =
      el_geom_phase6))
}

#calculate daily num activities and ellipse area for each person
activity_spaces_df <- activity_df %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
  group_modify(~ calculate_ellipse_area(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df)) %>%
  ungroup()
activity_spaces_df_baseline <- activity_df_baseline %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
17

  group_modify(~ calculate_ellipse_area_baseline(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_baseline)) %>%
  ungroup()
activity_spaces_df_phase1 <- activity_df_phase1 %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
```

```
  group_modify(~ calculate_ellipse_area_phase1(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_phase1)) %>%
  ungroup()
activity_spaces_df_phase2 <- activity_df_phase2 %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
  group_modify(~ calculate_ellipse_area_phase2(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_phase2)) %>%
  ungroup()
activity_spaces_df_phase3 <- activity_df_phase3 %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
  group_modify(~ calculate_ellipse_area_phase3(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_phase3)) %>%
  ungroup()
activity_spaces_df_phase4 <- activity_df_phase4 %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
  group_modify(~ calculate_ellipse_area_phase4(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_phase4)) %>%
  ungroup()
activity_spaces_df_phase5 <- activity_df_phase5 %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
  group_modify(~ calculate_ellipse_area_phase5(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_phase5)) %>%
  ungroup()
activity_spaces_df_phase6 <- activity_df_phase6 %>%
  group_by(user_id) %>%
  mutate(n=n()) %>%
  filter(n>=3) %>%
  group_modify(~ calculate_ellipse_area_phase6(.x), keep = TRUE) %>%
  st_as_sf(crs=st_crs(activity_df_phase6)) %>%
  ungroup()

#getting the same length of argument
participants_for_analysise <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df$user_id))
activity_spaces_df_for_analysise <- activity_spaces_df %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_baseline <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_baseline$user_id))
```

```
activity_spaces_df_for_analysise_baseline <- activity_spaces_df_baseline %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_phase1 <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_phase1$user_id))
activity_spaces_df_for_analysise_phase1 <- activity_spaces_df_phase1 %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_phase2 <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_phase2$user_id))
activity_spaces_df_for_analysise_phase2 <- activity_spaces_df_phase2 %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_phase3 <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_phase3$user_id))
activity_spaces_df_for_analysise_phase3 <- activity_spaces_df_phase3 %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_phase4 <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_phase4$user_id))
activity_spaces_df_for_analysise_phase4 <- activity_spaces_df_phase4 %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_phase5 <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_phase5$user_id))
activity_spaces_df_for_analysise_phase5 <- activity_spaces_df_phase5 %>%
  filter(user_id %in% unique(participants$participant_ID))
participants_for_analysise_phase6 <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df_phase6$user_id))
activity_spaces_df_for_analysise_phase6 <- activity_spaces_df_phase6 %>%
  filter(user_id %in% unique(participants$participant_ID))

#Save results for all participants, before plotting
participants_activity_df <- dplyr::bind_cols(participants_for_analysise,
                                        activity_spaces_df_for_analysise %>%
                                          select(-user_id))
participants_activity_df$AgeGroup <- cut(participants_activity_df$age, breaks =
  c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df, 'activity_spaces_results.csv')

participants_activity_df$nonGrouping <- cut(participants_activity_df$age, breaks =
  c(seq(10, 65, by=100), Inf))
plot2 <- boxplot(ellipse_area ~ nonGrouping, participants_activity_df,
               cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
               main = 'Activity space distribution',
               ylab = 'all participants', xlab = 'Activity space [km2]',
               ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df$ellipse_area, breaks=1000, main='', xlab = 'value of
  the variable')

hist(participants_activity_df$age, participants_activity_df$ellipse_area, xlim =
```

```
  c(16, 67))

ggplot(participants_activity_df, aes()) +
                                    geom_histogram(x = participants_activity_df$age, y
                                      =
                                     participants_activity_df$ellipse_area)+
                                    labs(title='', x='age group of the participants',
                                      y='Activity space area [km^2]',
                                        color ='') +
                                  scale_color_manual(labels = c('mean', 'median'),
                                    values = c('blue', 'red'))

participants_activity_df_baseline <-
  dplyr::bind_cols(participants_for_analysise_baseline,
                                        activity_spaces_df_for_analysise_baseline
                                          %>% select(-user_id))
participants_activity_df_baseline$AgeGroup <-
  cut(participants_activity_df_baseline$age, breaks = c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df_baseline, 'activity_spaces_results_baseline.csv')

participants_activity_df_baseline$nonGrouping <-
  cut(participants_activity_df_baseline$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_baseline <- boxplot(ellipse_area_baseline ~ nonGrouping,
  participants_activity_df_baseline,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (baseline)',
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_baseline$ellipse_area_baseline, breaks=1000, main='',
  xlab = 'value of the variable')

hist(participants_activity_df_baseline$age,
  participants_activity_df_baseline$ellipse_area_baseline, xlim = c(16, 67))

ggplot(participants_activity_df_baseline, aes()) +
  geom_histogram(x = participants_activity_df_baseline$age, y =
                participants_activity_df_baseline$ellipse_area_baseline)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

participants_activity_df_phase1 <- dplyr::bind_cols(participants_for_analysise_phase1,
                                        activity_spaces_df_for_analysise_phase1 %>%
                                          select(-user_id))
participants_activity_df_phase1$AgeGroup <- cut(participants_activity_df_phase1$age,
  breaks = c(seq(10, 65, by=10), Inf))
```

```
write.csv(participants_activity_df_phase1, 'activity_spaces_results_phase1.csv')

participants_activity_df_phase1$nonGrouping <-
  cut(participants_activity_df_phase1$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_phase1 <- boxplot(ellipse_area_phase1 ~ nonGrouping,
  participants_activity_df_phase1,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (phase 1)',
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_phase1$ellipse_area_phase1, breaks=1000, main='', xlab
  = 'value of the variable')

hist(participants_activity_df_phase1$age,
  participants_activity_df_phase1$ellipse_area_phase1, xlim = c(16, 67))

ggplot(participants_activity_df_phase1, aes()) +
  geom_histogram(x = participants_activity_df_phase1$age, y =
                   participants_activity_df_phase1$ellipse_area_phase1)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

participants_activity_df_phase2 <- dplyr::bind_cols(participants_for_analysise_phase2,
                                          activity_spaces_df_for_analysise_phase2 %>%
                                            select(-user_id))
participants_activity_df_phase2$AgeGroup <- cut(participants_activity_df_phase2$age,
  breaks = c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df_phase2, 'activity_spaces_results_phase2.csv')

participants_activity_df_phase2$nonGrouping <-
  cut(participants_activity_df_phase2$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_phase2 <- boxplot(ellipse_area_phase2 ~ nonGrouping,
  participants_activity_df_phase2,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (phase 2)',
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_phase2$ellipse_area_phase2, breaks=1000, main='', xlab
  = 'value of the variable')

hist(participants_activity_df_phase2$age,
  participants_activity_df_phase2$ellipse_area_phase2, xlim = c(16, 67))

ggplot(participants_activity_df_phase2, aes()) +
  geom_histogram(x = participants_activity_df_phase2$age, y =
```

```
                participants_activity_df_phase2$ellipse_area_phase2)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

participants_activity_df_phase3 <- dplyr::bind_cols(participants_for_analysise_phase3,
                                        activity_spaces_df_for_analysise_phase3 %>%
                                          select(-user_id))
participants_activity_df_phase3$AgeGroup <- cut(participants_activity_df_phase3$age,
  breaks = c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df_phase3, 'activity_spaces_results_phase3.csv')

participants_activity_df_phase3$nonGrouping <-
  cut(participants_activity_df_phase3$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_phase3 <- boxplot(ellipse_area_phase3 ~ nonGrouping,
  participants_activity_df_phase3,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (phase 3)',
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_phase3$ellipse_area_phase3, breaks=1000, main='', xlab
  = 'value of the variable')

hist(participants_activity_df_phase3$age,
  participants_activity_df_phase3$ellipse_area_phase3, xlim = c(16, 67))

ggplot(participants_activity_df_phase3, aes()) +
  geom_histogram(x = participants_activity_df_phase3$age, y =
                   participants_activity_df_phase3$ellipse_area_phase3)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

participants_activity_df_phase4 <- dplyr::bind_cols(participants_for_analysise_phase4,
                                        activity_spaces_df_for_analysise_phase4 %>%
                                          select(-user_id))
participants_activity_df_phase4$AgeGroup <- cut(participants_activity_df_phase4$age,
  breaks = c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df_phase4, 'activity_spaces_results_phase4.csv')

participants_activity_df_phase4$nonGrouping <-
  cut(participants_activity_df_phase4$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_phase4 <- boxplot(ellipse_area_phase4 ~ nonGrouping,
  participants_activity_df_phase4,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (phase 4)',
```

```
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_phase4$ellipse_area_phase4, breaks=1000, main='', xlab
  = 'value of the variable')

hist(participants_activity_df_phase4$age,
  participants_activity_df_phase4$ellipse_area_phase4, xlim = c(16, 67))

ggplot(participants_activity_df_phase4, aes()) +
  geom_histogram(x = participants_activity_df_phase4$age, y =
                   participants_activity_df_phase4$ellipse_area_phase4)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

participants_activity_df_phase5 <- dplyr::bind_cols(participants_for_analysise_phase5,
                                        activity_spaces_df_for_analysise_phase5 %>%
                                          select(-user_id))
participants_activity_df_phase5$AgeGroup <- cut(participants_activity_df_phase5$age,
  breaks = c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df_phase5, 'activity_spaces_results_phase5.csv')

participants_activity_df_phase5$nonGrouping <-
  cut(participants_activity_df_phase5$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_phase5 <- boxplot(ellipse_area_phase5 ~ nonGrouping,
  participants_activity_df_phase5,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (phase 5)',
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_phase5$ellipse_area_phase5, breaks=1000, main='', xlab
  = 'value of the variable')

hist(participants_activity_df_phase5$age,
  participants_activity_df_phase5$ellipse_area_phase5, xlim = c(16, 67))

ggplot(participants_activity_df_phase5, aes()) +
  geom_histogram(x = participants_activity_df_phase5$age, y =
                   participants_activity_df_phase5$ellipse_area_phase5)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

participants_activity_df_phase6 <- dplyr::bind_cols(participants_for_analysise_phase6,
                                        activity_spaces_df_for_analysise_phase6 %>%
                                          select(-user_id))
```

```
participants_activity_df_phase6$AgeGroup <- cut(participants_activity_df_phase6$age,
  breaks = c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df_phase6, 'activity_spaces_results_phase6.csv')

participants_activity_df_phase6$nonGrouping <-
  cut(participants_activity_df_phase6$age, breaks = c(seq(10, 65, by=100), Inf))
plot2_phase6 <- boxplot(ellipse_area_phase6 ~ nonGrouping,
  participants_activity_df_phase6,
              cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
              main = 'Activity space distribution (phase 6)',
              ylab = 'all participants', xlab = 'Activity space [km2]',
              ylim = c(0,8000), las = 1, horizontal = TRUE)
hist(participants_activity_df_phase6$ellipse_area_phase6, breaks=1000, main='', xlab
  = 'value of the variable')

hist(participants_activity_df_phase6$age,
  participants_activity_df_phase6$ellipse_area_phase6, xlim = c(16, 67))

ggplot(participants_activity_df_phase6, aes()) +
  geom_histogram(x = participants_activity_df_phase6$age, y =
                   participants_activity_df_phase6$ellipse_area_phase6)+
  labs(title='', x='age group of the participants', y='Activity space area [km^2]',
       color ='') +
  scale_color_manual(labels = c('mean', 'median'), values = c('blue', 'red'))

#histogram for ellipse area
ggplot(data=participants_activity_df, aes(x=ellipse_area))+
  geom_histogram(breaks=seq(0, 2800, by=10),
              alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))

ggplot(data=participants_activity_df_baseline, aes(x=ellipse_area_baseline))+
  geom_histogram(breaks=seq(0, 2800, by=10),
              alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))
```

```
ggplot(data=participants_activity_df_phase1, aes(x=ellipse_area_phase1))+
  geom_histogram(breaks=seq(0, 2800, by=10),
                 alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))

ggplot(data=participants_activity_df_phase2, aes(x=ellipse_area_phase2))+
  geom_histogram(breaks=seq(0, 2800, by=10),
                 alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))

ggplot(data=participants_activity_df_phase3, aes(x=ellipse_area_phase3))+
  geom_histogram(breaks=seq(0, 2800, by=10),
                 alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))

ggplot(data=participants_activity_df_phase4, aes(x=ellipse_area_phase4))+
  geom_histogram(breaks=seq(0, 2800, by=10),
                 alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))

ggplot(data=participants_activity_df_phase5, aes(x=ellipse_area_phase5))+
  geom_histogram(breaks=seq(0, 2800, by=10),
                 alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
```

```
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))

ggplot(data=participants_activity_df_phase6, aes(x=ellipse_area_phase6))+
  geom_histogram(breaks=seq(0, 2800, by=10),
                 alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
        ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))


#key figures all participants
length_all <- nrow(participants_activity_df)
mean_all <- mean(participants_activity_df$ellipse_area)
median_all <- median(participants_activity_df$ellipse_area)
sd_all <- sd(participants_activity_df$ellipse_area)
quantile1_all <- quantile(participants_activity_df$ellipse_area, 0.25)
quantile2_all <- quantile(participants_activity_df$ellipse_area, 0.75)

key_figures_all <- c(length_all, mean_all, median_all, sd_all,quantile1_all,
  quantile2_all)

length_all_baseline <- nrow(participants_activity_df_baseline)
mean_all_baseline <- mean(participants_activity_df_baseline$ellipse_area_baseline)
median_all_baseline <- median(participants_activity_df_baseline$ellipse_area_baseline)
sd_all_baseline <- sd(participants_activity_df_baseline$ellipse_area_baseline)
quantile1_all_baseline <-
  quantile(participants_activity_df_baseline$ellipse_area_baseline, 0.25)
quantile2_all_baseline <-
  quantile(participants_activity_df_baseline$ellipse_area_baseline, 0.75)

key_figures_all_baseline <- c(length_all_baseline, mean_all_baseline,
                              median_all_baseline, sd_all_baseline,
                              quantile1_all_baseline, quantile2_all_baseline)

length_all_phase1 <- nrow(participants_activity_df_phase1)
mean_all_phase1 <- mean(participants_activity_df_phase1$ellipse_area_phase1)
median_all_phase1 <- median(participants_activity_df_phase1$ellipse_area_phase1)
sd_all_phase1 <- sd(participants_activity_df_phase1$ellipse_area_phase1)
quantile1_all_phase1 <- quantile(participants_activity_df_phase1$ellipse_area_phase1,
  0.25)
quantile2_all_phase1 <- quantile(participants_activity_df_phase1$ellipse_area_phase1,
```

```
  0.75)

key_figures_all_phase1 <- c(length_all_phase1, mean_all_phase1, median_all_phase1,
                        sd_all_phase1,quantile1_all_phase1, quantile2_all_phase1)

length_all_phase2 <- nrow(participants_activity_df_phase2)
mean_all_phase2 <- mean(participants_activity_df_phase2$ellipse_area_phase2)
median_all_phase2 <- median(participants_activity_df_phase2$ellipse_area_phase2)
sd_all_phase2 <- sd(participants_activity_df_phase2$ellipse_area_phase2)
quantile1_all_phase2 <- quantile(participants_activity_df_phase2$ellipse_area_phase2,
  0.25)
quantile2_all_phase2 <- quantile(participants_activity_df_phase2$ellipse_area_phase2,
  0.75)

key_figures_all_phase2 <- c(length_all_phase2, mean_all_phase2, median_all_phase2,
                        sd_all_phase2,quantile1_all_phase2, quantile2_all_phase2)

length_all_phase3 <- nrow(participants_activity_df_phase3)
mean_all_phase3 <- mean(participants_activity_df_phase3$ellipse_area_phase3)
median_all_phase3 <- median(participants_activity_df_phase3$ellipse_area_phase3)
sd_all_phase3 <- sd(participants_activity_df_phase3$ellipse_area_phase3)
quantile1_all_phase3 <- quantile(participants_activity_df_phase3$ellipse_area_phase3,
  0.25)
quantile2_all_phase3 <- quantile(participants_activity_df_phase3$ellipse_area_phase3,
  0.75)

key_figures_all_phase3 <- c(length_all_phase3, mean_all_phase3, median_all_phase3,
                        sd_all_phase3,quantile1_all_phase3, quantile2_all_phase3)

length_all_phase4 <- nrow(participants_activity_df_phase4)
mean_all_phase4 <- mean(participants_activity_df_phase4$ellipse_area_phase4)
median_all_phase4 <- median(participants_activity_df_phase4$ellipse_area_phase4)
sd_all_phase4 <- sd(participants_activity_df_phase4$ellipse_area_phase4)
quantile1_all_phase4 <- quantile(participants_activity_df_phase4$ellipse_area_phase4,
  0.25)
quantile2_all_phase4 <- quantile(participants_activity_df_phase4$ellipse_area_phase4,
  0.75)

key_figures_all_phase4 <- c(length_all_phase4, mean_all_phase4, median_all_phase4,
                        sd_all_phase4,quantile1_all_phase4, quantile2_all_phase4)

length_all_phase5 <- nrow(participants_activity_df_phase5)
mean_all_phase5 <- mean(participants_activity_df_phase5$ellipse_area_phase5)
median_all_phase5 <- median(participants_activity_df_phase5$ellipse_area_phase5)
sd_all_phase5 <- sd(participants_activity_df_phase5$ellipse_area_phase5)
quantile1_all_phase5 <- quantile(participants_activity_df_phase5$ellipse_area_phase5,
```

```
  0.25)
quantile2_all_phase5 <- quantile(participants_activity_df_phase5$ellipse_area_phase5,
  0.75)

key_figures_all_phase5 <- c(length_all_phase5, mean_all_phase5, median_all_phase5,
                        sd_all_phase5,quantile1_all_phase5, quantile2_all_phase5)

length_all_phase6 <- nrow(participants_activity_df_phase6)
mean_all_phase6 <- mean(participants_activity_df_phase6$ellipse_area_phase6)
median_all_phase6 <- median(participants_activity_df_phase6$ellipse_area_phase6)
sd_all_phase6 <- sd(participants_activity_df_phase6$ellipse_area_phase6)
quantile1_all_phase6 <- quantile(participants_activity_df_phase6$ellipse_area_phase6,
  0.25)
quantile2_all_phase6 <- quantile(participants_activity_df_phase6$ellipse_area_phase6,
  0.75)

key_figures_all_phase6 <- c(length_all_phase6, mean_all_phase6, median_all_phase6,
                        sd_all_phase6,quantile1_all_phase6, quantile2_all_phase6)


#activity space vs age
participants_activity_df_age <- participants_activity_df %>%
  filter(!is.na(age))
length_age_all <- nrow(participants_activity_df_age)
mean_activity_space_age <- participants_activity_df_age %>%
  group_by(age) %>%
  summarise(mean_space_age = mean(ellipse_area))
median_activity_space_age <- participants_activity_df_age %>%
  group_by(age) %>%
  summarise(median_space_age = median(ellipse_area))
std_activity_space_age <- participants_activity_df_age %>%
  group_by(age) %>%
  summarise(std_space_age = sd(ellipse_area))
activity_space_age <- bind_cols(mean_activity_space_age, median_activity_space_age %>%
                             select(-age), std_activity_space_age %>% select(-age))

participants_activity_df_age_baseline <- participants_activity_df_baseline %>%
  filter(!is.na(age))
length_age_all_baseline <- nrow(participants_activity_df_age_baseline)
mean_activity_space_age_baseline<- participants_activity_df_age_baseline %>%
  group_by(age) %>%
  summarise(mean_space_age_baseline = mean(ellipse_area_baseline))
median_activity_space_age_baseline <- participants_activity_df_age_baseline %>%
  group_by(age) %>%
  summarise(median_space_age_baseline = median(ellipse_area_baseline))
std_activity_space_age_baseline <- participants_activity_df_age_baseline %>%
```

```
  group_by(age) %>%
  summarise(std_space_age_baseline = sd(ellipse_area_baseline))
activity_space_age_baseline <- bind_cols(mean_activity_space_age_baseline,
   median_activity_space_age_baseline %>%
                                select(-age), std_activity_space_age_baseline %>%
                                  select(-age))

participants_activity_df_age_phase1 <- participants_activity_df_phase1 %>%
  filter(!is.na(age))
length_age_all_phase1 <- nrow(participants_activity_df_age_phase1)
mean_activity_space_age_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(age) %>%
  summarise(mean_space_age_phase1 = mean(ellipse_area_phase1))
median_activity_space_age_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(age) %>%
  summarise(median_space_age_phase1 = median(ellipse_area_phase1))
std_activity_space_age_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(age) %>%
  summarise(std_space_age_phase1 = sd(ellipse_area_phase1))
activity_space_age_phase1 <- bind_cols(mean_activity_space_age_phase1,
   median_activity_space_age_phase1 %>%
                                select(-age), std_activity_space_age_phase1 %>%
                                  select(-age))

participants_activity_df_age_phase2 <- participants_activity_df_phase2 %>%
  filter(!is.na(age))
length_age_all_phase2 <- nrow(participants_activity_df_age_phase2)
mean_activity_space_age_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(age) %>%
  summarise(mean_space_age_phase2 = mean(ellipse_area_phase2))
median_activity_space_age_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(age) %>%
  summarise(median_space_age_phase2 = median(ellipse_area_phase2))
std_activity_space_age_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(age) %>%
  summarise(std_space_age_phase2 = sd(ellipse_area_phase2))
activity_space_age_phase2 <- bind_cols(mean_activity_space_age_phase2,
   median_activity_space_age_phase2 %>%
                                select(-age), std_activity_space_age_phase2 %>%
                                  select(-age))

participants_activity_df_age_phase3 <- participants_activity_df_phase3 %>%
  filter(!is.na(age))
length_age_all_phase3 <- nrow(participants_activity_df_age_phase3)
mean_activity_space_age_phase3 <- participants_activity_df_age_phase3 %>%
  group_by(age) %>%
```

```
  summarise(mean_space_age_phase3 = mean(ellipse_area_phase3))
median_activity_space_age_phase3 <- participants_activity_df_age_phase3 %>%
  group_by(age) %>%
  summarise(median_space_age_phase3 = median(ellipse_area_phase3))
std_activity_space_age_phase3 <- participants_activity_df_age_phase3 %>%
  group_by(age) %>%
  summarise(std_space_age_phase3 = sd(ellipse_area_phase3))
activity_space_age_phase3 <- bind_cols(mean_activity_space_age_phase3,
   median_activity_space_age_phase3 %>%
                                select(-age), std_activity_space_age_phase3 %>%
                                  select(-age))

participants_activity_df_age_phase4 <- participants_activity_df_phase4 %>%
  filter(!is.na(age))
length_age_all_phase4 <- nrow(participants_activity_df_age_phase4)
mean_activity_space_age_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(age) %>%
  summarise(mean_space_age_phase4 = mean(ellipse_area_phase4))
median_activity_space_age_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(age) %>%
  summarise(median_space_age_phase4 = median(ellipse_area_phase4))
std_activity_space_age_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(age) %>%
  summarise(std_space_age_phase4 = sd(ellipse_area_phase4))
activity_space_age_phase4 <- bind_cols(mean_activity_space_age_phase4,
   median_activity_space_age_phase4 %>%
                                select(-age), std_activity_space_age_phase4 %>%
                                  select(-age))

participants_activity_df_age_phase5 <- participants_activity_df_phase5 %>%
  filter(!is.na(age))
length_age_all_phase5 <- nrow(participants_activity_df_age_phase5)
mean_activity_space_age_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(age) %>%
  summarise(mean_space_age_phase5 = mean(ellipse_area_phase5))
median_activity_space_age_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(age) %>%
  summarise(median_space_age_phase5 = median(ellipse_area_phase5))
std_activity_space_age_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(age) %>%
  summarise(std_space_age_phase5 = sd(ellipse_area_phase5))
activity_space_age_phase5 <- bind_cols(mean_activity_space_age_phase5,
   median_activity_space_age_phase5 %>%
                                select(-age), std_activity_space_age_phase5 %>%
                                  select(-age))
```

```
participants_activity_df_age_phase6 <- participants_activity_df_phase6 %>%
  filter(!is.na(age))
length_age_all_phase6 <- nrow(participants_activity_df_age_phase6)
mean_activity_space_age_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(age) %>%
  summarise(mean_space_age_phase6 = mean(ellipse_area_phase6))
median_activity_space_age_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(age) %>%
  summarise(median_space_age_phase6 = median(ellipse_area_phase6))
std_activity_space_age_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(age) %>%
  summarise(std_space_age_phase6 = sd(ellipse_area_phase6))
activity_space_age_phase6 <- bind_cols(mean_activity_space_age_phase6,
   median_activity_space_age_phase6 %>%
                                select(-age), std_activity_space_age_phase6 %>%
                                  select(-age))

#activity space vs AgeGroup
participants_activity_df_age$AgeGroup <- cut(participants_activity_df_age$age, breaks
  =
                                           c(seq(10,65, by=10), Inf))
mean_activity_space_age <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age = mean(ellipse_area))
median_activity_space_age <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age = median(ellipse_area))
std_activity_space_age <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age = sd(ellipse_area))
length_AgeGroup <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age = max(rownumberAge))
activity_space_age <- bind_cols(mean_activity_space_age, median_activity_space_age
                              %>%select(-AgeGroup),
                              std_activity_space_age %>% select(-AgeGroup),
                              length_AgeGroup %>%select(-AgeGroup))

boxplot(ellipse_area ~AgeGroup, participants_activity_df_age,
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,12000), las = 1,
       par(mar = c(4,6,1,1)))
```

```
op <- par(mar = c(5,4,4,2) + 0.1)
par(op)

boxplot(ellipse_area ~ age, participants_activity_df_age,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main = 'Activity space distribution by age',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,10000), las=1)

mean_age_all <- mean(activity_space_age$mean_space_age)
median_age_all <- median(activity_space_age$median_space_age)
sd_mean_age_all <- sd(activity_space_age$mean_space_age)
sd_median_age_all <- sd(activity_space_age$median_space_age)


key_figures_age_all <- c(length_age_all, mean_age_all, sd_mean_age_all,
  median_age_all, sd_median_age_all)

participants_activity_df_age_baseline$AgeGroup <-
  cut(participants_activity_df_age_baseline$age, breaks =
                                        c(seq(10,65, by=10), Inf))
mean_activity_space_age_baseline <- participants_activity_df_age_baseline %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_baseline = mean(ellipse_area_baseline))
median_activity_space_age_baseline <- participants_activity_df_age_baseline %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_baseline = median(ellipse_area_baseline))
std_activity_space_age_baseline <- participants_activity_df_age_baseline %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_baseline = sd(ellipse_area_baseline))
length_AgeGroup_baseline <- participants_activity_df_age_baseline %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age_baseline = max(rownumberAge))
activity_space_age_baseline <- bind_cols(mean_activity_space_age_baseline,
                                median_activity_space_age_baseline
                            %>%select(-AgeGroup),
                            std_activity_space_age_baseline %>% select(-AgeGroup),
                            length_AgeGroup_baseline %>%select(-AgeGroup))

boxplot(ellipse_area_baseline ~AgeGroup, participants_activity_df_age_baseline,
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='Baseline',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
```

```
       ylim = c(-100,12000), las = 1,
       par(mar = c(4,6,1,1)))

op_baseline <- par(mar = c(5,4,4,2) + 0.1)
par(op_baseline)

boxplot(ellipse_area_baseline ~ age, participants_activity_df_age_baseline,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main = 'Activity space distribution by age',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,10000), las=1)

mean_age_all_baseline <- mean(activity_space_age_baseline$mean_space_age)
median_age_all_baseline <- median(activity_space_age_baseline$median_space_age)
sd_mean_age_all_baseline <- sd(activity_space_age_baseline$mean_space_age)
sd_median_age_all_baseline <- sd(activity_space_age_baseline$median_space_age)


key_figures_age_all_baseline <- c(length_age_all_baseline, mean_age_all_baseline,
                                  sd_mean_age_all_baseline, median_age_all_baseline,
                                  sd_median_age_all_baseline)

participants_activity_df_age_phase1$AgeGroup <-
   cut(participants_activity_df_age_phase1$age, breaks =
                                           c(seq(10,65, by=10), Inf))
mean_activity_space_age_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_phase1 = mean(ellipse_area_phase1))
median_activity_space_age_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_phase1 = median(ellipse_area_phase1))
std_activity_space_age_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_phase1 = sd(ellipse_area_phase1))
length_AgeGroup_phase1 <- participants_activity_df_age_phase1 %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age_phase1 = max(rownumberAge))
activity_space_age_phase1 <- bind_cols(mean_activity_space_age_phase1,
   median_activity_space_age_phase1
                              %>%select(-AgeGroup),
                              std_activity_space_age_phase1 %>% select(-AgeGroup),
                              length_AgeGroup_phase1 %>%select(-AgeGroup))

boxplot(ellipse_area_phase1 ~AgeGroup, participants_activity_df_age_phase1,
```

```
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='Phase 1',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,12000), las = 1,
       par(mar = c(4,6,1,1)))

op_phase1 <- par(mar = c(5,4,4,2) + 0.1)
par(op_phase1)

boxplot(ellipse_area_phase1 ~ age, participants_activity_df_age_phase1,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main = 'Activity space distribution by age',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,10000), las=1)

mean_age_all_phase1 <- mean(activity_space_age_phase1$mean_space_age)
median_age_all_phase1 <- median(activity_space_age_phase1$median_space_age)
sd_mean_age_all_phase1 <- sd(activity_space_age_phase1$mean_space_age)
sd_median_age_all_phase1 <- sd(activity_space_age_phase1$median_space_age)


key_figures_age_all_phase1 <- c(length_age_all_phase1, mean_age_all_phase1,
                            sd_mean_age_all_phase1, median_age_all_phase1,
                            sd_median_age_all_phase1)

participants_activity_df_age_phase2$AgeGroup <-
  cut(participants_activity_df_age_phase2$age, breaks =
                                        c(seq(10,65, by=10), Inf))
mean_activity_space_age_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_phase2 = mean(ellipse_area_phase2))
median_activity_space_age_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_phase2 = median(ellipse_area_phase2))
std_activity_space_age_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_phase2 = sd(ellipse_area_phase2))
length_AgeGroup_phase2 <- participants_activity_df_age_phase2 %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age_phase2 = max(rownumberAge))
activity_space_age_phase2 <- bind_cols(mean_activity_space_age_phase2,
  median_activity_space_age_phase2
                            %>%select(-AgeGroup),
                            std_activity_space_age_phase2 %>% select(-AgeGroup),
```

```
                            length_AgeGroup_phase2 %>%select(-AgeGroup))

boxplot(ellipse_area_phase2 ~AgeGroup, participants_activity_df_age_phase2,
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='Phase 2',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,12000), las = 1,
       par(mar = c(4,6,1,1)))

op_phase2 <- par(mar = c(5,4,4,2) + 0.1)
par(op_phase2)

boxplot(ellipse_area_phase2 ~ age, participants_activity_df_age_phase2,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main = 'Activity space distribution by age',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,10000), las=1)

mean_age_all_phase2 <- mean(activity_space_age_phase2$mean_space_age)
median_age_all_phase2 <- median(activity_space_age_phase2$median_space_age)
sd_mean_age_all_phase2 <- sd(activity_space_age_phase2$mean_space_age)
sd_median_age_all_phase2 <- sd(activity_space_age_phase2$median_space_age)


key_figures_age_all_phase2 <- c(length_age_all_phase2, mean_age_all_phase2,
                            sd_mean_age_all_phase2, median_age_all_phase2,
                            sd_median_age_all_phase2)

participants_activity_df_age_phase3$AgeGroup <-
  cut(participants_activity_df_age_phase3$age, breaks =
                                        c(seq(10,65, by=10), Inf))
mean_activity_space_age_phase3 <- participants_activity_df_age_phase3 %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_phase3 = mean(ellipse_area_phase3))
median_activity_space_age_phase3 <- participants_activity_df_age_phase3 %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_phase3 = median(ellipse_area_phase3))
std_activity_space_age_phase3 <- participants_activity_df_ag_phase3 %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_phase3 = sd(ellipse_area_phase3))
length_AgeGroup_phase3 <- participants_activity_df_age_phase3 %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age_phase3 = max(rownumberAge))
activity_space_age_phase3 <- bind_cols(mean_activity_space_age_phase3,
```

```
  median_activity_space_age_phase3
                            %>%select(-AgeGroup),
                            std_activity_space_age_phase3 %>% select(-AgeGroup),
                            length_AgeGroup_phase3 %>%select(-AgeGroup))

boxplot(ellipse_area_phase3 ~AgeGroup, participants_activity_df_age_phase3,
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='Phase 3',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,12000), las = 1,
       par(mar = c(4,6,1,1)))

op_phase3 <- par(mar = c(5,4,4,2) + 0.1)
par(op_phase3)

boxplot(ellipse_area_phase3 ~ age, participants_activity_df_age_phase3,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main = 'Activity space distribution by age',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,10000), las=1)

mean_age_all_phase3 <- mean(activity_space_age_phase3$mean_space_age)
median_age_all_phase3 <- median(activity_space_age_phase3$median_space_age)
sd_mean_age_all_phase3 <- sd(activity_space_age_phase3$mean_space_age)
sd_median_age_all_phase3 <- sd(activity_space_age_phase3$median_space_age)


key_figures_age_all_phase3 <- c(length_age_all_phase3, mean_age_all_phase3,
                            sd_mean_age_all_phase3, median_age_all_phase3,
                            sd_median_age_all_phase3)

participants_activity_df_age_phase4$AgeGroup <-
  cut(participants_activity_df_age_phase4$age, breaks =
                                        c(seq(10,65, by=10), Inf))
mean_activity_space_age_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_phase4 = mean(ellipse_area_phase4))
median_activity_space_age_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_phase4 = median(ellipse_area_phase4))
std_activity_space_age_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_phase4 = sd(ellipse_area_phase4))
length_AgeGroup_phase4 <- participants_activity_df_age_phase4 %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
```

```
  group_by(AgeGroup) %>%
  summarise(length_space_age_phase4 = max(rownumberAge))
activity_space_age_phase4 <- bind_cols(mean_activity_space_age_phase4,
   median_activity_space_age_phase4
                              %>%select(-AgeGroup),
                              std_activity_space_age_phase4 %>% select(-AgeGroup),
                              length_AgeGroup_phase4 %>%select(-AgeGroup))

boxplot(ellipse_area_phase4 ~AgeGroup, participants_activity_df_age_phase4,
        cex.main = 1, cex.axis = 1, cex.lab =1,
        main ='Phase 4',
        ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
        ylim = c(-100,12000), las = 1,
        par(mar = c(4,6,1,1)))

op_phase4 <- par(mar = c(5,4,4,2) + 0.1)
par(op_phase4)

boxplot(ellipse_area_phase4 ~ age, participants_activity_df_age_phase4,
        cex.main = 1, cex.axis = 1, cex.lab = 1,
        main = 'Activity space distribution by age',
        ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
        ylim = c(-100,10000), las=1)

mean_age_all_phase4 <- mean(activity_space_age_phase4$mean_space_age)
median_age_all_phase4 <- median(activity_space_age_phase4$median_space_age)
sd_mean_age_all_phase4 <- sd(activity_space_age_phase4$mean_space_age)
sd_median_age_all_phase4 <- sd(activity_space_age_phase4$median_space_age)


key_figures_age_all_phase4 <- c(length_age_all_phase4, mean_age_all_phase4,
                              sd_mean_age_all_phase4, median_age_all_phase4,
                              sd_median_age_all_phase4)

participants_activity_df_age_phase5$AgeGroup <-
   cut(participants_activity_df_age_phase5$age, breaks =
                                          c(seq(10,65, by=10), Inf))
mean_activity_space_age_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_phase5 = mean(ellipse_area_phase5))
median_activity_space_age_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_phase5 = median(ellipse_area_phase5))
std_activity_space_age_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_phase5 = sd(ellipse_area_phase5))
```

```
length_AgeGroup_phase5 <- participants_activity_df_age_phase5 %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age_phase5 = max(rownumberAge))
activity_space_age_phase5 <- bind_cols(mean_activity_space_age_phase5,
   median_activity_space_age_phase5
                              %>%select(-AgeGroup),
                              std_activity_space_age_phase5 %>% select(-AgeGroup),
                              length_AgeGroup_phase5 %>%select(-AgeGroup))

boxplot(ellipse_area_phase5 ~AgeGroup, participants_activity_df_age_phase5,
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='Phase 5',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,12000), las = 1,
       par(mar = c(4,6,1,1)))

op_phase5 <- par(mar = c(5,4,4,2) + 0.1)
par(op_phase5)

boxplot(ellipse_area_phase5 ~ age, participants_activity_df_age_phase5,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main = 'Activity space distribution by age',
       ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
       ylim = c(-100,10000), las=1)

mean_age_all_phase5 <- mean(activity_space_age_phase5$mean_space_age)
median_age_all_phase5 <- median(activity_space_age_phase5$median_space_age)
sd_mean_age_all_phase5 <- sd(activity_space_age_phase5$mean_space_age)
sd_median_age_all_phase5 <- sd(activity_space_age_phase5$median_space_age)


key_figures_age_all_phase5 <- c(length_age_all_phase5, mean_age_all_phase5,
                              sd_mean_age_all_phase5, median_age_all_phase5,
                              sd_median_age_all_phase5)

participants_activity_df_age_phase6$AgeGroup <-
   cut(participants_activity_df_age_phase6$age, breaks =
                                          c(seq(10,65, by=10), Inf))
mean_activity_space_age_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age_phase6 = mean(ellipse_area_phase6))
median_activity_space_age_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age_phase6 = median(ellipse_area_phase6))
```

```
std_activity_space_age_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age_phase6 = sd(ellipse_area_phase6))
length_AgeGroup_phase6 <- participants_activity_df_age_phase6 %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age_phase6 = max(rownumberAge))
activity_space_age_phase6 <- bind_cols(mean_activity_space_age_phase6,
   median_activity_space_age_phase6
                              %>%select(-AgeGroup),
                              std_activity_space_age_phase6 %>% select(-AgeGroup),
                              length_AgeGroup_phase6 %>%select(-AgeGroup))

boxplot(ellipse_area_phase6 ~AgeGroup, participants_activity_df_age_phase6,
        cex.main = 1, cex.axis = 1, cex.lab =1,
        main ='Phase 6',
        ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
        ylim = c(-100,12000), las = 1,
        par(mar = c(4,6,1,1)))

op_phase6 <- par(mar = c(5,4,4,2) + 0.1)
par(op_phase6)

boxplot(ellipse_area_phase6 ~ age, participants_activity_df_age_phase6,
        cex.main = 1, cex.axis = 1, cex.lab = 1,
        main = 'Activity space distribution by age',
        ylab = 'Activity space [km2]', xlab = 'Age group of the participants',
        ylim = c(-100,10000), las=1)

mean_age_all_phase6 <- mean(activity_space_age_phase6$mean_space_age)
median_age_all_phase6 <- median(activity_space_age_phase6$median_space_age)
sd_mean_age_all_phase6 <- sd(activity_space_age_phase6$mean_space_age)
sd_median_age_all_phase6 <- sd(activity_space_age_phase6$median_space_age)


key_figures_age_all_phase6 <- c(length_age_all_phase6, mean_age_all_phase6,
                              sd_mean_age_all_phase6, median_age_all_phase6,
                              sd_median_age_all_phase6)

ellipse_area_age1_baseline <- participants_activity_df_age_baseline %>%
  filter(AgeGroup == '(10,20]')
ellipse_area_age1_baseline <- bind_cols(ellipse_area_age1_baseline,
                                        'Baseline')
ellipse_area_age1_ph1 <- participants_activity_df_age_phase1 %>%
  filter(AgeGroup == '(10,20]')
```

```
ellipse_area_income1_ph1 <- bind_cols(ellipse_area_age1_ph1,
                                      'Phase 1')
ellipse_area_age1_ph2 <- participants_activity_df_age_phase2 %>%
  filter(AgeGroup == '(10,20]')
ellipse_area_age1_ph2 <- bind_cols(ellipse_area_age1_ph2,
                                      'Phase 2')
ellipse_area_age1_ph3 <- participants_activity_df_age_phase3 %>%
  filter(AgeGroup == '(10,20]')
ellipse_area_age1_ph3 <- bind_cols(ellipse_area_age1_ph3,
                                      'Phase 3')
ellipse_area_age1_ph4 <- participants_activity_df_age_phase4 %>%
  filter(AgeGroup == '(10,20]')
ellipse_area_age1_ph4 <- bind_cols(ellipse_area_age1_ph4,
                                      'Phase 4')
ellipse_area_age1_ph5 <- participants_activity_df_age_phase5 %>%
  filter(AgeGroup == '(10,20]')
ellipse_area_age1_ph5 <- bind_cols(ellipse_area_age1_ph5,
                                      'Phase 5')
ellipse_area_age1_ph6 <- participants_activity_df_age_phase6 %>%
  filter(AgeGroup == '(10,20]')
ellipse_area_age1_ph6 <- bind_cols(ellipse_area_age1_ph6,
                                      'Phase 6')

ellipse_area_age1_baseline <- ellipse_area_age1_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_age1_ph1 <- ellipse_area_age1_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_age1_ph2 <- ellipse_area_age1_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_age1_ph3 <- ellipse_area_age1_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_age1_ph4 <- ellipse_area_age1_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
```

```
    phase = 247
  )
ellipse_area_age1_ph5 <- ellipse_area_age1_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_age1_ph6 <- ellipse_area_age1_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_age1 <- bind_rows(ellipse_area_age1_baseline,
                               ellipse_area_age1_ph1, ellipse_area_age1_ph2,
                               ellipse_area_age1_ph3, ellipse_area_age1_ph4,
                               ellipse_area_age1_ph5, ellipse_area_age1_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_age1,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Age group (10,20]', ylab = 'Activity space [km2]',
       xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_age1_plot <- ellipse_area_age1 %>%
  group_by(phase) %>%
  summarise(ellipse_area_age_1 = median(ellipse_area))

ellipse_area_age2_baseline <- participants_activity_df_age_baseline %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_baseline <- bind_cols(ellipse_area_age2_baseline,
                                        'Baseline')
ellipse_area_age2_ph1 <- participants_activity_df_age_phase1 %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_ph1 <- bind_cols(ellipse_area_age2_ph1,
                                   'Phase 1')
ellipse_area_age2_ph2 <- participants_activity_df_age_phase2 %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_ph2 <- bind_cols(ellipse_area_age2_ph2,
                                   'Phase 2')
ellipse_area_age2_ph3 <- participants_activity_df_age_phase3 %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_ph3 <- bind_cols(ellipse_area_age2_ph3,
                                   'Phase 3')
ellipse_area_age2_ph4 <- participants_activity_df_age_phase4 %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_ph4 <- bind_cols(ellipse_area_age2_ph4,
```

```
                                  'Phase 4')
ellipse_area_age2_ph5 <- participants_activity_df_age_phase5 %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_ph5 <- bind_cols(ellipse_area_age2_ph5,
                                  'Phase 5')
ellipse_area_age2_ph6 <- participants_activity_df_age_phase6 %>%
  filter(AgeGroup == '(20,30]')
ellipse_area_age2_ph6 <- bind_cols(ellipse_area_age2_ph6,
                                  'Phase 6')

ellipse_area_age2_baseline <- ellipse_area_age2_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_age2_ph1 <- ellipse_area_age2_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_age2_ph2 <- ellipse_area_age2_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_age2_ph3 <- ellipse_area_age2_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_age2_ph4 <- ellipse_area_age2_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_age2_ph5 <- ellipse_area_age2_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_age2_ph6 <- ellipse_area_age2_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )
```

```
ellipse_area_age2 <- bind_rows(ellipse_area_age2_baseline,
                               ellipse_area_age2_ph1, ellipse_area_age2_ph2,
                               ellipse_area_age2_ph3, ellipse_area_age2_ph4,
                               ellipse_area_age2_ph5, ellipse_area_age2_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_age2,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Age group (20,30]', ylab = 'Activity space [km2]',
       xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_age2_plot <- ellipse_area_age2 %>%
  group_by(phase) %>%
  summarise(ellipse_area_age_2 = median(ellipse_area))

ellipse_area_age3_baseline <- participants_activity_df_age_baseline %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_baseline <- bind_cols(ellipse_area_age3_baseline,
                                        'Baseline')
ellipse_area_age3_ph1 <- participants_activity_df_age_phase1 %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_ph1 <- bind_cols(ellipse_area_age3_ph1,
                                   'Phase 1')
ellipse_area_age3_ph2 <- participants_activity_df_age_phase2 %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_ph2 <- bind_cols(ellipse_area_age3_ph2,
                                   'Phase 2')
ellipse_area_age3_ph3 <- participants_activity_df_age_phase3 %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_ph3 <- bind_cols(ellipse_area_age3_ph3,
                                   'Phase 3')
ellipse_area_age3_ph4 <- participants_activity_df_age_phase4 %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_ph4 <- bind_cols(ellipse_area_age3_ph4,
                                   'Phase 4')
ellipse_area_age3_ph5 <- participants_activity_df_age_phase5 %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_ph5 <- bind_cols(ellipse_area_age3_ph5,
                                   'Phase 5')
ellipse_area_age3_ph6 <- participants_activity_df_age_phase6 %>%
  filter(AgeGroup == '(30,40]')
ellipse_area_age3_ph6 <- bind_cols(ellipse_area_age3_ph6,
                                   'Phase 6')

ellipse_area_age3_baseline <- ellipse_area_age3_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
```

```
    phase = 247
  )
ellipse_area_age3_ph1 <- ellipse_area_age3_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_age3_ph2 <- ellipse_area_age3_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_age3_ph3 <- ellipse_area_age3_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_age3_ph4 <- ellipse_area_age3_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_age3_ph5 <- ellipse_area_age3_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_age3_ph6 <- ellipse_area_age3_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_age3 <- bind_rows(ellipse_area_age3_baseline,
                               ellipse_area_age3_ph1, ellipse_area_age3_ph2,
                               ellipse_area_age3_ph3, ellipse_area_age3_ph4,
                               ellipse_area_age3_ph5, ellipse_area_age3_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_age3,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Age group (30,40]', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_age3_plot <- ellipse_area_age3 %>%
  group_by(phase) %>%
  summarise(ellipse_area_age_3 = median(ellipse_area))
```

```
ellipse_area_age4_baseline <- participants_activity_df_age_baseline %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_baseline <- bind_cols(ellipse_area_age4_baseline,
                                        'Baseline')
ellipse_area_age4_ph1 <- participants_activity_df_age_phase1 %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_ph1 <- bind_cols(ellipse_area_age4_ph1,
                                   'Phase 1')
ellipse_area_age4_ph2 <- participants_activity_df_age_phase2 %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_ph2 <- bind_cols(ellipse_area_age4_ph2,
                                   'Phase 2')
ellipse_area_age4_ph3 <- participants_activity_df_age_phase3 %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_ph3 <- bind_cols(ellipse_area_age4_ph3,
                                   'Phase 3')
ellipse_area_age4_ph4 <- participants_activity_df_age_phase4 %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_ph4 <- bind_cols(ellipse_area_age4_ph4,
                                   'Phase 4')
ellipse_area_age4_ph5 <- participants_activity_df_age_phase5 %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_ph5 <- bind_cols(ellipse_area_age4_ph5,
                                   'Phase 5')
ellipse_area_age4_ph6 <- participants_activity_df_age_phase6 %>%
  filter(AgeGroup == '(40,50]')
ellipse_area_age4_ph6 <- bind_cols(ellipse_area_age4_ph6,
                                   'Phase 6')

ellipse_area_age4_baseline <- ellipse_area_age4_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_age4_ph1 <- ellipse_area_age4_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_age4_ph2 <- ellipse_area_age4_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_age4_ph3 <- ellipse_area_age4_ph3 %>%
```

```
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_age4_ph4 <- ellipse_area_age4_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_age4_ph5 <- ellipse_area_age4_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_age4_ph6 <- ellipse_area_age4_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_age4 <- bind_rows(ellipse_area_age4_baseline,
                               ellipse_area_age4_ph1, ellipse_area_age4_ph2,
                               ellipse_area_age4_ph3, ellipse_area_age4_ph4,
                               ellipse_area_age4_ph5, ellipse_area_age4_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_age4,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Age group (40,50]', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_age4_plot <- ellipse_area_age4 %>%
  group_by(phase) %>%
  summarise(ellipse_area_age_4 = median(ellipse_area))

ellipse_area_age5_baseline <- participants_activity_df_age_baseline %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_baseline <- bind_cols(ellipse_area_age5_baseline,
                                        'Baseline')
ellipse_area_age5_ph1 <- participants_activity_df_age_phase1 %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_ph1 <- bind_cols(ellipse_area_age5_ph1,
                                   'Phase 1')
ellipse_area_age5_ph2 <- participants_activity_df_age_phase2 %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_ph2 <- bind_cols(ellipse_area_age5_ph2,
                                   'Phase 2')
```

```
ellipse_area_age5_ph3 <- participants_activity_df_age_phase3 %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_ph3 <- bind_cols(ellipse_area_age5_ph3,
                                   'Phase 3')
ellipse_area_age5_ph4 <- participants_activity_df_age_phase4 %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_ph4 <- bind_cols(ellipse_area_age5_ph4,
                                   'Phase 4')
ellipse_area_age5_ph5 <- participants_activity_df_age_phase5 %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_ph5 <- bind_cols(ellipse_area_age5_ph5,
                                   'Phase 5')
ellipse_area_age5_ph6 <- participants_activity_df_age_phase6 %>%
  filter(AgeGroup == '(50,60]')
ellipse_area_age5_ph6 <- bind_cols(ellipse_area_age5_ph6,
                                   'Phase 6')

ellipse_area_age5_baseline <- ellipse_area_age5_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_age5_ph1 <- ellipse_area_age5_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_age5_ph2 <- ellipse_area_age5_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_age5_ph3 <- ellipse_area_age5_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_age5_ph4 <- ellipse_area_age5_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_age5_ph5 <- ellipse_area_age5_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
```

```
  )
ellipse_area_age5_ph6 <- ellipse_area_age5_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_age5 <- bind_rows(ellipse_area_age5_baseline,
                               ellipse_area_age5_ph1, ellipse_area_age5_ph2,
                               ellipse_area_age5_ph3, ellipse_area_age5_ph4,
                               ellipse_area_age5_ph5, ellipse_area_age5_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_age5,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Age group (50,60]', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_age5_plot <- ellipse_area_age5 %>%
  group_by(phase) %>%
  summarise(ellipse_area_age_5 = median(ellipse_area))



ellipse_area_age6_baseline <- participants_activity_df_age_baseline %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_baseline <- bind_cols(ellipse_area_age6_baseline,
                                        'Baseline')
ellipse_area_age6_ph1 <- participants_activity_df_age_phase1 %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_ph1 <- bind_cols(ellipse_area_age6_ph1,
                                   'Phase 1')
ellipse_area_age6_ph2 <- participants_activity_df_age_phase2 %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_ph2 <- bind_cols(ellipse_area_age6_ph2,
                                   'Phase 2')
ellipse_area_age6_ph3 <- participants_activity_df_age_phase3 %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_ph3 <- bind_cols(ellipse_area_age6_ph3,
                                   'Phase 3')
ellipse_area_age6_ph4 <- participants_activity_df_age_phase4 %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_ph4 <- bind_cols(ellipse_area_age6_ph4,
                                   'Phase 4')
ellipse_area_age6_ph5 <- participants_activity_df_age_phase5 %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_ph5 <- bind_cols(ellipse_area_age6_ph5,
                                   'Phase 5')
```

```
ellipse_area_age6_ph6 <- participants_activity_df_age_phase6 %>%
  filter(AgeGroup == '(60,Inf]')
ellipse_area_age6_ph6 <- bind_cols(ellipse_area_age6_ph6,
                                   'Phase 6')

ellipse_area_age6_baseline <- ellipse_area_age6_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_age6_ph1 <- ellipse_area_age6_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_age6_ph2 <- ellipse_area_age6_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_age6_ph3 <- ellipse_area_age6_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_age6_ph4 <- ellipse_area_age6_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_age6_ph5 <- ellipse_area_age6_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_age6_ph6 <- ellipse_area_age6_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_age6 <- bind_rows(ellipse_area_age6_baseline,
                               ellipse_area_age6_ph1, ellipse_area_age6_ph2,
                               ellipse_area_age6_ph3, ellipse_area_age6_ph4,
                               ellipse_area_age6_ph5, ellipse_area_age6_ph6)
```

```
boxplot(ellipse_area ~ phase, ellipse_area_age6,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Age group (60,Inf]', ylab = 'Activity space [km2]',
       xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_age6_plot <- ellipse_area_age6 %>%
  group_by(phase) %>%
  summarise(ellipse_area_age_6 = median(ellipse_area))

#household size
participants_activity_df_household <- participants_activity_df %>%
  filter(!is.na(household_size))
length_household_all <- nrow(participants_activity_df_household)

mean_activity_space_household <- participants_activity_df_household %>%
  group_by(household_size) %>%
  summarise(mean_space_household = mean(ellipse_area))
median_activity_space_household <- participants_activity_df_household %>%
  group_by(household_size) %>%
  summarise(median_space_household = median(ellipse_area))
std_activity_space_household <- participants_activity_df_household %>%
  group_by(household_size) %>%
  summarise(std_space_household = sd(ellipse_area))
length_household_seq <- participants_activity_df_household %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold =row_number()) %>%
  summarise(length_space_household = max(rownumberhousehold))
activity_space_household <- bind_cols(mean_activity_space_household,
  median_activity_space_household %>%
                                    select(-household_size),
                                      std_activity_space_household %>%
                                    select(-household_size), length_household_seq %>%
                                    select(-household_size))

boxplot(ellipse_area ~ household_size, participants_activity_df_household,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all <- mean(activity_space_household$mean_space_household)
median_household_all <- median(activity_space_household$median_space_household)
sd_mean_household_all <- sd(activity_space_household$mean_space_household)
sd_median_household_all <- sd(activity_space_household$median_space_household)

key_figures_household_all <- c(length_household_all, mean_household_all,
  sd_mean_household_all,
```

```
                          median_household_all, sd_median_household_all)


participants_activity_df_household_baseline <- participants_activity_df_baseline %>%
  filter(!is.na(household_size))
length_household_all_baseline <- nrow(participants_activity_df_household_baseline)

mean_activity_space_household_baseline <- participants_activity_df_household_baseline
  %>%
  group_by(household_size) %>%
  summarise(mean_space_household_baseline = mean(ellipse_area_baseline))
median_activity_space_household_baseline <-
   participants_activity_df_household_baseline %>%
  group_by(household_size) %>%
  summarise(median_space_household_baseline = median(ellipse_area_baseline))
std_activity_space_household_baseline <- participants_activity_df_household_baseline
  %>%
  group_by(household_size) %>%
  summarise(std_space_household_baseline = sd(ellipse_area_baseline))
length_household_seq_baseline <- participants_activity_df_household_baseline %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_baseline =row_number()) %>%
  summarise(length_space_household_baseline = max(rownumberhousehold_baseline))
activity_space_household_baseline <-
   bind_cols(mean_activity_space_household_baseline,
   median_activity_space_household_baseline %>%
                                    select(-household_size),
                                      std_activity_space_household_baseline %>%
                                    select(-household_size),
                                      length_household_seq_baseline %>%
                                    select(-household_size))

boxplot(ellipse_area_baseline ~ household_size,
   participants_activity_df_household_baseline,
        cex.main = 1, cex.axis = 1, cex.lab = 1,
        main ='Baseline', ylab = 'Activity space [km]', xlab = 'Household size',
        ylim = c(-100,14000), las=1)

mean_household_all_baseline <-
   mean(activity_space_household_baseline$mean_space_household)
median_household_all_baseline <-
   median(activity_space_household_baseline$median_space_household)
sd_mean_household_all_baseline <-
   sd(activity_space_household_baseline$mean_space_household)
sd_median_household_all_baseline <-
   sd(activity_space_household_baseline$median_space_household)
```

```
key_figures_household_all_baseline <- c(length_household_all_baseline,
  mean_household_all_baseline,
                                    sd_mean_household_all_baseline,
                                      median_household_all_baseline,
                                    sd_median_household_all_baseline)

participants_activity_df_household_phase1 <- participants_activity_df_phase1 %>%
  filter(!is.na(household_size))
length_household_all_phase1 <- nrow(participants_activity_df_household_phase1)

mean_activity_space_household_phase1 <- participants_activity_df_household_phase1 %>%
  group_by(household_size) %>%
  summarise(mean_space_household_phase1 = mean(ellipse_area_phase1))
median_activity_space_household_phase1 <- participants_activity_df_household_phase1
  %>%
  group_by(household_size) %>%
  summarise(median_space_household_phase1 = median(ellipse_area_phase1))
std_activity_space_household_phase1 <- participants_activity_df_household_phase1 %>%
  group_by(household_size) %>%
  summarise(std_space_household_phase1 = sd(ellipse_area_phase1))
length_household_seq_phase1 <- participants_activity_df_household_phase1 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_phase1 =row_number()) %>%
  summarise(length_space_household_phase1 = max(rownumberhousehold_phase1))
activity_space_household_phase1 <- bind_cols(mean_activity_space_household_phase1,
  median_activity_space_household_phase1 %>%
                                    select(-household_size),
                                      std_activity_space_household_phase1 %>%
                                    select(-household_size),
                                      length_household_seq_phase1 %>%
                                    select(-household_size))

boxplot(ellipse_area_phase1 ~ household_size,
  participants_activity_df_household_phase1,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='Phase 1', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all_phase1 <-
  mean(activity_space_household_phase1$mean_space_household)
median_household_all_phase1 <-
  median(activity_space_household_phase1$median_space_household)
sd_mean_household_all_phase1 <-
  sd(activity_space_household_phase1$mean_space_household)
sd_median_household_all_phase1 <-
```

```
  sd(activity_space_household_phase1$median_space_household)

key_figures_household_all_phase1 <- c(length_household_all_phase1,
  mean_household_all_phase1,
                                    sd_mean_household_all_phase1,
                                      median_household_all_phase1,
                                    sd_median_household_all_phase1)

participants_activity_df_household_phase2 <- participants_activity_df_phase2 %>%
  filter(!is.na(household_size))
length_household_all_phase2 <- nrow(participants_activity_df_household_phase2)

mean_activity_space_household_phase2 <- participants_activity_df_household_phase2 %>%
  group_by(household_size) %>%
  summarise(mean_space_household_phase2 = mean(ellipse_area_phase2))
median_activity_space_household_phase2 <- participants_activity_df_household_phase2
  %>%
  group_by(household_size) %>%
  summarise(median_space_household_phase2 = median(ellipse_area_phase2))
std_activity_space_household_phase2 <- participants_activity_df_household_phase2 %>%
  group_by(household_size) %>%
  summarise(std_space_household_phase2 = sd(ellipse_area_phase2))
length_household_seq_phase2 <- participants_activity_df_household_phase2 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_phase2 =row_number()) %>%
  summarise(length_space_household_phase2 = max(rownumberhousehold_phase2))
activity_space_household_phase2 <- bind_cols(mean_activity_space_household_phase2,
  median_activity_space_household_phase2 %>%
                                    select(-household_size),
                                      std_activity_space_household_phase2 %>%
                                    select(-household_size),
                                      length_household_seq_phase2 %>%
                                    select(-household_size))

boxplot(ellipse_area_phase2 ~ household_size,
  participants_activity_df_household_phase2,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all_phase2 <-
  mean(activity_space_household_phase2$mean_space_household)
median_household_all_phase2 <-
  median(activity_space_household_phase2$median_space_household)
sd_mean_household_all_phase2 <-
  sd(activity_space_household_phase2$mean_space_household)
```

```
sd_median_household_all_phase2 <-
  sd(activity_space_household_phase2$median_space_household)

key_figures_household_all_phase2 <- c(length_household_all_phase2,
  mean_household_all_phase2,
                                    sd_mean_household_all_phase2,
                                      median_household_all_phase2,
                                    sd_median_household_all_phase2)

participants_activity_df_household_phase3 <- participants_activity_df_phase3 %>%
  filter(!is.na(household_size))
length_household_all_phase3 <- nrow(participants_activity_df_household_phase3)

mean_activity_space_household_phase3 <- participants_activity_df_household_phase3 %>%
  group_by(household_size) %>%
  summarise(mean_space_household_phase3 = mean(ellipse_area_phase3))
median_activity_space_household_phase3 <- participants_activity_df_household_phase3
  %>%
  group_by(household_size) %>%
  summarise(median_space_household_phase3 = median(ellipse_area_phase3))
std_activity_space_household_phase3 <- participants_activity_df_household_phase3 %>%
  group_by(household_size) %>%
  summarise(std_space_household_phase3 = sd(ellipse_area_phase3))
length_household_seq_phase3 <- participants_activity_df_household_phase3 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_phase3 =row_number()) %>%
  summarise(length_space_household_phase3 = max(rownumberhousehold_phase3))
activity_space_household_phase3 <- bind_cols(mean_activity_space_household_phase3,
  median_activity_space_household_phase3 %>%
                                    select(-household_size),
                                      std_activity_space_household_phase3 %>%
                                    select(-household_size),
                                      length_household_seq_phase3 %>%
                                    select(-household_size))

boxplot(ellipse_area_phase3 ~ household_size,
  participants_activity_df_household_phase3,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='Phase 3', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all_phase3 <-
  mean(activity_space_household_phase3$mean_space_household)
median_household_all_phase3 <-
  median(activity_space_household_phase3$median_space_household)
sd_mean_household_all_phase3 <-
```

```
  sd(activity_space_household_phase3$mean_space_household)
sd_median_household_all_phase3 <-
  sd(activity_space_household_phase3$median_space_household)

key_figures_household_all_phase3 <- c(length_household_all_phase3,
  mean_household_all_phase3,
                                    sd_mean_household_all_phase3,
                                      median_household_all_phase3,
                                    sd_median_household_all_phase3)

participants_activity_df_household_phase4 <- participants_activity_df_phase4 %>%
  filter(!is.na(household_size))
length_household_all_phase4 <- nrow(participants_activity_df_household_phase4)

mean_activity_space_household_phase4 <- participants_activity_df_household_phase4 %>%
  group_by(household_size) %>%
  summarise(mean_space_household_phase4 = mean(ellipse_area_phase4))
median_activity_space_household_phase4 <- participants_activity_df_household_phase4
  %>%
  group_by(household_size) %>%
  summarise(median_space_household_phase4 = median(ellipse_area_phase4))
std_activity_space_household_phase4 <- participants_activity_df_household_phase4 %>%
  group_by(household_size) %>%
  summarise(std_space_household_phase4 = sd(ellipse_area_phase4))
length_household_seq_phase4 <- participants_activity_df_household_phase4 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_phase4 =row_number()) %>%
  summarise(length_space_household_phase4 = max(rownumberhousehold_phase4))
activity_space_household_phase4 <- bind_cols(mean_activity_space_household_phase4,
  median_activity_space_household_phase4 %>%
                                    select(-household_size),
                                      std_activity_space_household_phase4 %>%
                                    select(-household_size),
                                      length_household_seq_phase4 %>%
                                    select(-household_size))

boxplot(ellipse_area_phase4 ~ household_size,
  participants_activity_df_household_phase4,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='Phase 4', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all_phase4 <-
  mean(activity_space_household_phase4$mean_space_household)
median_household_all_phase4 <-
  median(activity_space_household_phase4$median_space_household)
```

```
sd_mean_household_all_phase4 <-
  sd(activity_space_household_phase4$mean_space_household)
sd_median_household_all_phase4 <-
  sd(activity_space_household_phase4$median_space_household)

key_figures_household_all_phase4 <- c(length_household_all_phase4,
  mean_household_all_phase4,
                                      sd_mean_household_all_phase4,
                                        median_household_all_phase4,
                                      sd_median_household_all_phase4)

participants_activity_df_household_phase5 <- participants_activity_df_phase5 %>%
  filter(!is.na(household_size))
length_household_all_phase5 <- nrow(participants_activity_df_household_phase5)

mean_activity_space_household_phase5 <- participants_activity_df_household_phase5 %>%
  group_by(household_size) %>%
  summarise(mean_space_household_phase5 = mean(ellipse_area_phase5))
median_activity_space_household_phase5 <- participants_activity_df_household_phase5
  %>%
  group_by(household_size) %>%
  summarise(median_space_household_phase5 = median(ellipse_area_phase5))
std_activity_space_household_phase5 <- participants_activity_df_household_phase5 %>%
  group_by(household_size) %>%
  summarise(std_space_household_phase5 = sd(ellipse_area_phase5))
length_household_seq_phase5 <- participants_activity_df_household_phase5 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_phase5 =row_number()) %>%
  summarise(length_space_household_phase5 = max(rownumberhousehold_phase5))
activity_space_household_phase5 <- bind_cols(mean_activity_space_household_phase5,
  median_activity_space_household_phase5 %>%
                                       select(-household_size),
                                         std_activity_space_household_phase5 %>%
                                       select(-household_size),
                                         length_household_seq_phase5 %>%
                                       select(-household_size))

boxplot(ellipse_area_phase5 ~ household_size,
  participants_activity_df_household_phase5,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='Phase 5', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all_phase5 <-
  mean(activity_space_household_phase5$mean_space_household)
median_household_all_phase5 <-
```

```
  median(activity_space_household_phase5$median_space_household)
sd_mean_household_all_phase5 <-
  sd(activity_space_household_phase5$mean_space_household)
sd_median_household_all_phase5 <-
  sd(activity_space_household_phase5$median_space_household)

key_figures_household_all_phase5 <- c(length_household_all_phase5,
  mean_household_all_phase5,
                                    sd_mean_household_all_phase5,
                                      median_household_all_phase5,
                                    sd_median_household_all_phase5)

participants_activity_df_household_phase6 <- participants_activity_df_phase6 %>%
  filter(!is.na(household_size))
length_household_all_phase6 <- nrow(participants_activity_df_household_phase6)

mean_activity_space_household_phase6 <- participants_activity_df_household_phase6 %>%
  group_by(household_size) %>%
  summarise(mean_space_household_phase6 = mean(ellipse_area_phase6))
median_activity_space_household_phase6 <- participants_activity_df_household_phase6
  %>%
  group_by(household_size) %>%
  summarise(median_space_household_phase6 = median(ellipse_area_phase6))
std_activity_space_household_phase6 <- participants_activity_df_household_phase6 %>%
  group_by(household_size) %>%
  summarise(std_space_household_phase6 = sd(ellipse_area_phase6))
length_household_seq_phase6 <- participants_activity_df_household_phase6 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold_phase6 =row_number()) %>%
  summarise(length_space_household_phase6 = max(rownumberhousehold_phase6))
activity_space_household_phase6 <- bind_cols(mean_activity_space_household_phase6,
  median_activity_space_household_phase6 %>%
                                    select(-household_size),
                                      std_activity_space_household_phase6 %>%
                                    select(-household_size),
                                      length_household_seq_phase6 %>%
                                    select(-household_size))

boxplot(ellipse_area_phase6 ~ household_size,
  participants_activity_df_household_phase6,
       cex.main = 1, cex.axis = 1, cex.lab = 1,
       main ='Phase 6', ylab = 'Activity space [km]', xlab = 'Household size',
       ylim = c(-100,14000), las=1)

mean_household_all_phase6 <-
  mean(activity_space_household_phase6$mean_space_household)
```

```
median_household_all_phase6 <-
  median(activity_space_household_phase6$median_space_household)
sd_mean_household_all_phase6 <-
  sd(activity_space_household_phase6$mean_space_household)
sd_median_household_all_phase6 <-
  sd(activity_space_household_phase6$median_space_household)

key_figures_household_all_phase6 <- c(length_household_all_phase6,
  mean_household_all_phase6,
                                    sd_mean_household_all_phase6,
                                      median_household_all_phase6,
                                    sd_median_household_all_phase6)

ellipse_area_household1_baseline <- participants_activity_df_household_baseline %>%
  filter(household_size == '1')
ellipse_area_household_baseline <- bind_cols(ellipse_area_household1_baseline,
                                         'Baseline')
ellipse_area_household1_ph1 <- participants_activity_df_household_phase1 %>%
  filter(household_size == '1')
ellipse_area_household1_ph1 <- bind_cols(ellipse_area_household1_ph1,
                                    'Phase 1')
ellipse_area_household1_ph2 <- participants_activity_df_household_phase2 %>%
  filter(household_size == '1')
ellipse_area_household1_ph2 <- bind_cols(ellipse_area_household1_ph2,
                                    'Phase 2')
ellipse_area_household1_ph3 <- participants_activity_df_household_phase3 %>%
  filter(household_size == '1')
ellipse_area_household1_ph3 <- bind_cols(ellipse_area_household1_ph3,
                                    'Phase 3')
ellipse_area_household1_ph4 <- participants_activity_df_household_phase4 %>%
  filter(household_size == '1')
ellipse_area_household1_ph4 <- bind_cols(ellipse_area_household1_ph4,
                                    'Phase 4')
ellipse_area_household1_ph5 <- participants_activity_df_household_phase5 %>%
  filter(household_size == '1')
ellipse_area_household1_ph5 <- bind_cols(ellipse_area_household1_ph5,
                                    'Phase 5')
ellipse_area_household1_ph6 <- participants_activity_df_household_phase6 %>%
  filter(household_size == '1')
ellipse_area_household1_ph6 <- bind_cols(ellipse_area_household1_ph6,
                                    'Phase 6')

ellipse_area_household1_baseline <- ellipse_area_household1_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
```

```
  )
ellipse_area_household1_ph1 <- ellipse_area_household1_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_household1_ph2 <- ellipse_area_household1_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_household1_ph3 <- ellipse_area_household1_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_household1_ph4 <- ellipse_area_household1_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_household1_ph5 <- ellipse_area_household1_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_household1_ph6 <- ellipse_area_household1_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_household1 <- bind_rows(ellipse_area_household1_baseline,
                                ellipse_area_household1_ph1,
                                  ellipse_area_household1_ph2,
                                ellipse_area_household1_ph3,
                                  ellipse_area_household1_ph4,
                                ellipse_area_household1_ph5,
                                  ellipse_area_household1_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_household1,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Household size 1', ylab = 'Activity space [km2]',
       xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_household1_plot <- ellipse_area_household1 %>%
```

```
  group_by(phase) %>%
  summarise(ellipse_area_household_1 = median(ellipse_area))

ellipse_area_household2_baseline <- participants_activity_df_household_baseline %>%
  filter(household_size == '2')
ellipse_area_household2_baseline <- bind_cols(ellipse_area_household2_baseline,
                                      'Baseline')
ellipse_area_household2_ph1 <- participants_activity_df_household_phase1 %>%
  filter(household_size == '2')
ellipse_area_household2_ph1 <- bind_cols(ellipse_area_household2_ph1,
                                   'Phase 1')
ellipse_area_household2_ph2 <- participants_activity_df_household_phase2 %>%
  filter(household_size == '2')
ellipse_area_household2_ph2 <- bind_cols(ellipse_area_household2_ph2,
                                   'Phase 2')
ellipse_area_household2_ph3 <- participants_activity_df_household_phase3 %>%
  filter(household_size == '2')
ellipse_area_household2_ph3 <- bind_cols(ellipse_area_household2_ph3,
                                   'Phase 3')
ellipse_area_household2_ph4 <- participants_activity_df_household_phase4 %>%
  filter(household_size == '2')
ellipse_area_household2_ph4 <- bind_cols(ellipse_area_household2_ph4,
                                   'Phase 4')
ellipse_area_household2_ph5 <- participants_activity_df_household_phase5 %>%
  filter(household_size == '2')
ellipse_area_household2_ph5 <- bind_cols(ellipse_area_household2_ph5,
                                   'Phase 5')
ellipse_area_household2_ph6 <- participants_activity_df_household_phase6 %>%
  filter(household_size == '2')
ellipse_area_household2_ph6 <- bind_cols(ellipse_area_household2_ph6,
                                   'Phase 6')

ellipse_area_household2_baseline <- ellipse_area_household2_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_household2_ph1 <- ellipse_area_household2_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_household2_ph2 <- ellipse_area_household2_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
```

```
  )
ellipse_area_household2_ph3 <- ellipse_area_household2_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_household2_ph4 <- ellipse_area_household2_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_household2_ph5 <- ellipse_area_household2_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_household2_ph6 <- ellipse_area_household2_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_household2 <- bind_rows(ellipse_area_household2_baseline,
                                ellipse_area_household2_ph1,
                                  ellipse_area_household2_ph2,
                                ellipse_area_household2_ph3,
                                  ellipse_area_household2_ph4,
                                ellipse_area_household2_ph5,
                                  ellipse_area_household2_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_household2,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Household size 2', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_household2_plot <- ellipse_area_household2 %>%
  group_by(phase) %>%
  summarise(ellipse_area_household_2 = median(ellipse_area))

ellipse_area_household3_baseline <- participants_activity_df_household_baseline %>%
  filter(household_size == '3')
ellipse_area_household3_baseline <- bind_cols(ellipse_area_household3_baseline,
                                      'Baseline')
ellipse_area_household3_ph1 <- participants_activity_df_household_phase1 %>%
  filter(household_size == '3')
ellipse_area_household3_ph1 <- bind_cols(ellipse_area_household3_ph1,
```

```
                                  'Phase 1')
ellipse_area_household3_ph2 <- participants_activity_df_household_phase2 %>%
  filter(household_size == '3')
ellipse_area_household3_ph2 <- bind_cols(ellipse_area_household3_ph2,
                                  'Phase 2')
ellipse_area_household3_ph3 <- participants_activity_df_household_phase3 %>%
  filter(household_size == '3')
ellipse_area_household3_ph3 <- bind_cols(ellipse_area_household3_ph3,
                                  'Phase 3')
ellipse_area_household3_ph4 <- participants_activity_df_household_phase4 %>%
  filter(household_size == '3')
ellipse_area_household3_ph4 <- bind_cols(ellipse_area_household3_ph4,
                                  'Phase 4')
ellipse_area_household3_ph5 <- participants_activity_df_household_phase5 %>%
  filter(household_size == '3')
ellipse_area_household3_ph5 <- bind_cols(ellipse_area_household3_ph5,
                                  'Phase 5')
ellipse_area_household3_ph6 <- participants_activity_df_household_phase6 %>%
  filter(household_size == '3')
ellipse_area_household3_ph6 <- bind_cols(ellipse_area_household3_ph6,
                                  'Phase 6')

ellipse_area_household3_baseline <- ellipse_area_household3_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_household3_ph1 <- ellipse_area_household3_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_household3_ph2 <- ellipse_area_household3_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_household3_ph3 <- ellipse_area_household3_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_household3_ph4 <- ellipse_area_household3_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
```

```
  )
ellipse_area_household3_ph5 <- ellipse_area_household3_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_household3_ph6 <- ellipse_area_household3_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_household3 <- bind_rows(ellipse_area_household3_baseline,
                                ellipse_area_household3_ph1,
                                  ellipse_area_household3_ph2,
                                ellipse_area_household3_ph3,
                                  ellipse_area_household3_ph4,
                                ellipse_area_household3_ph5,
                                  ellipse_area_household3_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_household3,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Household size 3', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_household3_plot <- ellipse_area_household3 %>%
  group_by(phase) %>%
  summarise(ellipse_area_household_3 = median(ellipse_area))

ellipse_area_household4_baseline <- participants_activity_df_household_baseline %>%
  filter(household_size == '4')
ellipse_area_household4_baseline <- bind_cols(ellipse_area_household4_baseline,
                                      'Baseline')
ellipse_area_household4_ph1 <- participants_activity_df_household_phase1 %>%
  filter(household_size == '4')
ellipse_area_household4_ph1 <- bind_cols(ellipse_area_household4_ph1,
                                    'Phase 1')
ellipse_area_household4_ph2 <- participants_activity_df_household_phase2 %>%
  filter(household_size == '4')
ellipse_area_household4_ph2 <- bind_cols(ellipse_area_household4_ph2,
                                    'Phase 2')
ellipse_area_household4_ph3 <- participants_activity_df_household_phase3 %>%
  filter(household_size == '4')
ellipse_area_household4_ph3 <- bind_cols(ellipse_area_household4_ph3,
                                    'Phase 3')
ellipse_area_household4_ph4 <- participants_activity_df_household_phase4 %>%
```

```
  filter(household_size == '4')
ellipse_area_household4_ph4 <- bind_cols(ellipse_area_household4_ph4,
                                  'Phase 4')
ellipse_area_household4_ph5 <- participants_activity_df_household_phase5 %>%
  filter(household_size == '4')
ellipse_area_household4_ph5 <- bind_cols(ellipse_area_household4_ph5,
                                  'Phase 5')
ellipse_area_household4_ph6 <- participants_activity_df_household_phase6 %>%
  filter(household_size == '4')
ellipse_area_household4_ph6 <- bind_cols(ellipse_area_household4_ph6,
                                  'Phase 6')

ellipse_area_household4_baseline <- ellipse_area_household4_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_household4_ph1 <- ellipse_area_household4_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_household4_ph2 <- ellipse_area_household4_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_household4_ph3 <- ellipse_area_household4_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_household4_ph4 <- ellipse_area_household4_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_household4_ph5 <- ellipse_area_household4_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_household4_ph6 <- ellipse_area_household4_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
```

```
  )

ellipse_area_household4 <- bind_rows(ellipse_area_household4_baseline,
                                ellipse_area_household4_ph1,
                                  ellipse_area_household4_ph2,
                                ellipse_area_household4_ph3,
                                  ellipse_area_household4_ph4,
                                ellipse_area_household4_ph5,
                                  ellipse_area_household4_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_household4,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Household size 4', ylab = 'Activity space [km2]',
       xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_household4_plot <- ellipse_area_household4 %>%
  group_by(phase) %>%
  summarise(ellipse_area_household_4 = median(ellipse_area))

ellipse_area_household5_baseline <- participants_activity_df_household_baseline %>%
  filter(household_size == '5 or more')
ellipse_area_household5_baseline <- bind_cols(ellipse_area_household5_baseline,
                                      'Baseline')
ellipse_area_household5_ph1 <- participants_activity_df_household_phase1 %>%
  filter(household_size == '5 or more')
ellipse_area_household5_ph1 <- bind_cols(ellipse_area_household5_ph1,
                                    'Phase 1')
ellipse_area_household5_ph2 <- participants_activity_df_household_phase2 %>%
  filter(household_size == '5 or more')
ellipse_area_household5_ph2 <- bind_cols(ellipse_area_household5_ph2,
                                    'Phase 2')
ellipse_area_household5_ph3 <- participants_activity_df_household_phase3 %>%
  filter(household_size == '5 or more')
ellipse_area_household5_ph3 <- bind_cols(ellipse_area_household5_ph3,
                                    'Phase 3')
ellipse_area_household5_ph4 <- participants_activity_df_household_phase4 %>%
  filter(household_size == '5 or more')
ellipse_area_household5_ph4 <- bind_cols(ellipse_area_household5_ph4,
                                    'Phase 4')
ellipse_area_household5_ph5 <- participants_activity_df_household_phase5 %>%
  filter(household_size == '5 or more')
ellipse_area_income5_ph5 <- bind_cols(ellipse_area_income5_ph5,
                                    'Phase 5')
ellipse_area_household5_ph6 <- participants_activity_df_household_phase6 %>%
  filter(household_size == '5 or more')
ellipse_area_household5_ph6 <- bind_cols(ellipse_area_household5_ph6,
```

```
                                    'Phase 6')

ellipse_area_household5_baseline <- ellipse_area_household5_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 247
  )
ellipse_area_household5_ph1 <- ellipse_area_household5_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 247
  )
ellipse_area_household5_ph2 <- ellipse_area_household5_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 247
  )
ellipse_area_household5_ph3 <- ellipse_area_household5_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 247
  )
ellipse_area_household5_ph4 <- ellipse_area_household5_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 247
  )
ellipse_area_household5_ph5 <- ellipse_area_household5_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 247
  )
ellipse_area_household5_ph6 <- ellipse_area_household5_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 247
  )

ellipse_area_household5 <- bind_rows(ellipse_area_household5_baseline,
                                     ellipse_area_household5_ph1,
                                       ellipse_area_household5_ph2,
                                     ellipse_area_household5_ph3,
                                       ellipse_area_household5_ph4,
                                     ellipse_area_household5_ph5,
                                       ellipse_area_household5_ph6)
```

```
boxplot(ellipse_area ~ phase, ellipse_area_household5,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Household size 5 or more', ylab = 'Activity space [km2]',
       xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_household5_plot <- ellipse_area_household5 %>%
  group_by(phase) %>%
  summarise(ellipse_area_household_5 = median(ellipse_area))



#activity space vs income
participants_activity_df_income <- participants_activity_df %>%
  filter(income!=99)
length_income_all <- nrow(participants_activity_df_income)

mean_activity_space_income <- participants_activity_df_income %>%
  group_by(income) %>%
  summarise(mean_space_income = mean(ellipse_area))
median_activity_space_income <- participants_activity_df_income %>%
  group_by(income) %>%
  summarise(median_space_income = median(ellipse_area))
std_activity_space_income <- participants_activity_df_income %>%
  group_by(income) %>%
  summarise(std_space_income = sd(ellipse_area))
length_income_seq <- participants_activity_df_income %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
activity_space_income_all <- bind_cols(mean_activity_space_income,
                                median_activity_space_income %>% select(-income),
                                std_activity_space_income %>% select(-income),
                                length_income_seq %>% select(-income))
length_income_seq %>% select(-income)

participants_activity_df_income$income <-
   factor(participants_activity_df_income$income
                                    , levels=c('4 000 CHF or less',
                                               '4 001 - 8 000 CHF','8 001
                                                 - 12 000 CHF',
                                               '12 001 - 16 000
                                                 CHF','More than 16 000
                                                 CHF',
                                               'Prefer not to say'))
boxplot(ellipse_area ~ income, participants_activity_df_income,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
```

```
       main = '', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all <- mean(activity_space_income_all$mean_space_income)
median_income_all <- median(activity_space_income_all$median_space_income)
sd_mean_income_all <- sd(activity_space_income_all$mean_space_income)
sd_median_income_all <- sd(activity_space_income_all$median_space_income)

key_figures_income_all <- c(length_income_all, mean_income_all, sd_mean_income_all,
                            median_income_all, sd_median_income_all)

participants_activity_df_income_baseline <- participants_activity_df_baseline %>%
  filter(income!=99)
length_income_all_baseline <- nrow(participants_activity_df_income_baseline)

mean_activity_space_income_baseline <- participants_activity_df_income_baseline %>%
  group_by(income) %>%
  summarise(mean_space_income_baseline = mean(ellipse_area_baseline))
median_activity_space_income_baseline <- participants_activity_df_income_baseline %>%
  group_by(income) %>%
  summarise(median_space_income_baseline = median(ellipse_area_baseline))
std_activity_space_income_baseline <- participants_activity_df_income_baseline %>%
  group_by(income) %>%
  summarise(std_space_income_baseline = sd(ellipse_area_baseline))
length_income_seq_baseline <- participants_activity_df_income_baseline %>%
  group_by(income) %>%
  mutate(rownumberincome_baseline = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_baseline = max(rownumberincome_baseline))
activity_space_income_all_baseline <- bind_cols(mean_activity_space_income_baseline,
                                    median_activity_space_income_baseline %>%
                                      select(-income),
                                    std_activity_space_income_baseline %>%
                                      select(-income),
                                    length_income_seq_baseline %>% select(-income))
length_income_seq_baseline %>% select(-income)

participants_activity_df_income_baseline$income <-
   factor(participants_activity_df_income_baseline$income
                                          , levels=c('4 000 CHF or less',
                                                     '4 001 - 8 000 CHF','8 001 - 12
                                                       000 CHF',
                                                     '12 001 - 16 000 CHF','More
                                                       than 16 000 CHF',
                                                     'Prefer not to say'))
boxplot(ellipse_area_baseline ~ income, participants_activity_df_income_baseline,
```

```
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Baseline', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_baseline <-
  mean(activity_space_income_all_baseline$mean_space_income_baseline)
median_income_all_baseline <-
  median(activity_space_income_all_baseline$median_space_income_baseline)
sd_mean_income_all_baseline <-
  sd(activity_space_income_all_baseline$mean_space_income_baseline)
sd_median_income_all_baseline <-
  sd(activity_space_income_all_baseline$median_space_income_baseline)

key_figures_income_all_baseline <- c(length_income_all_baseline,
  mean_income_all_baseline,
                                sd_mean_income_all_baseline,
                                  median_income_all_baseline,
                                sd_median_income_all_baseline)

participants_activity_df_income_phase1 <- participants_activity_df_phase1 %>%
  filter(income!=99)
length_income_all_phase1 <- nrow(participants_activity_df_income_phase1)

mean_activity_space_income_phase1 <- participants_activity_df_income_phase1 %>%
  group_by(income) %>%
  summarise(mean_space_income_phase1 = mean(ellipse_area_phase1))
median_activity_space_income_phase1 <- participants_activity_df_income_phase1 %>%
  group_by(income) %>%
  summarise(median_space_income_phase1 = median(ellipse_area_phase1))
std_activity_space_income_phase1 <- participants_activity_df_income_phase1 %>%
  group_by(income) %>%
  summarise(std_space_income_phase1 = sd(ellipse_area_phase1))
length_income_seq_phase1 <- participants_activity_df_income_phase1 %>%
  group_by(income) %>%
  mutate(rownumberincome_phase1 = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_phase1 = max(rownumberincome_phase1))
activity_space_income_all_phase1 <- bind_cols(mean_activity_space_income_phase1,
                                  median_activity_space_income_phase1 %>%
                                    select(-income),
                                  std_activity_space_income_phase1 %>%
                                    select(-income),
                                  length_income_seq_phase1 %>% select(-income))
length_income_seq_phase1 %>% select(-income)

participants_activity_df_income_phase1$income <-
```

```
  factor(participants_activity_df_income_phase1$income
                                          , levels=c('4 000 CHF or less',
                                                     '4 001 - 8 000 CHF','8 001 - 12
                                                       000 CHF',
                                                     '12 001 - 16 000 CHF','More
                                                       than 16 000 CHF',
                                                     'Prefer not to say'))
boxplot(ellipse_area_phase1 ~ income, participants_activity_df_income_phase1,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Phase 1', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_phase1 <-
  mean(activity_space_income_all_phase1$mean_space_income_phase1)
median_income_all_phase1 <-
  median(activity_space_income_all_phase1$median_space_income_phase1)
sd_mean_income_all_phase1 <-
  sd(activity_space_income_all_phase1$mean_space_income_phase1)
sd_median_income_all_phase1 <-
  sd(activity_space_income_all_phase1$median_space_income_phase1)

key_figures_income_all_phase1 <- c(length_income_all_phase1, mean_income_all_phase1,
                                   sd_mean_income_all_phase1, median_income_all_phase1,
                                   sd_median_income_all_phase1)

participants_activity_df_income_phase2 <- participants_activity_df_phase2 %>%
  filter(income!=99)
length_income_all_phase2 <- nrow(participants_activity_df_income_phase2)

mean_activity_space_income_phase2 <- participants_activity_df_income_phase2 %>%
  group_by(income) %>%
  summarise(mean_space_income_phase2 = mean(ellipse_area_phase2))
median_activity_space_income_phase2 <- participants_activity_df_income_phase2 %>%
  group_by(income) %>%
  summarise(median_space_income_phase2 = median(ellipse_area_phase2))
std_activity_space_income_phase2 <- participants_activity_df_income_phase2 %>%
  group_by(income) %>%
  summarise(std_space_income_phase2 = sd(ellipse_area_phase2))
length_income_seq_phase2 <- participants_activity_df_income_phase2 %>%
  group_by(income) %>%
  mutate(rownumberincome_phase2 = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_phase2 = max(rownumberincome_phase2))
activity_space_income_all_phase2 <- bind_cols(mean_activity_space_income_phase2,
                                      median_activity_space_income_phase2 %>%
                                        select(-income),
```

```
                                    std_activity_space_income_phase2 %>%
                                      select(-income),
                                    length_income_seq_phase2 %>% select(-income))
length_income_seq_phase2 %>% select(-income)

participants_activity_df_income_phase2$income <-
  factor(participants_activity_df_income_phase2$income
                                          , levels=c('4 000 CHF or less',
                                                     '4 001 - 8 000 CHF','8 001 - 12
                                                       000 CHF',
                                                     '12 001 - 16 000 CHF','More
                                                       than 16 000 CHF',
                                                     'Prefer not to say'))
boxplot(ellipse_area_phase2 ~ income, participants_activity_df_income_phase2,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Phase 2', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_phase2 <-
  mean(activity_space_income_all_phase2$mean_space_income_phase2)
median_income_all_phase2 <-
  median(activity_space_income_all_phase2$median_space_income_phase2)
sd_mean_income_all_phase2 <-
  sd(activity_space_income_all_phase2$mean_space_income_phase2)
sd_median_income_all_phase2 <-
  sd(activity_space_income_all_phase2$median_space_income_phase2)

key_figures_income_all_phase2 <- c(length_income_all_phase2, mean_income_all_phase2,
                                   sd_mean_income_all_phase2, median_income_all_phase2,
                                   sd_median_income_all_phase2)

participants_activity_df_income_phase3 <- participants_activity_df_phase3 %>%
  filter(income!=99)
length_income_all_phase3 <- nrow(participants_activity_df_income_phase3)

mean_activity_space_income_phase3 <- participants_activity_df_income_phase3 %>%
  group_by(income) %>%
  summarise(mean_space_income_phase3 = mean(ellipse_area_phase3))
median_activity_space_income_phase3 <- participants_activity_df_income_phase3 %>%
  group_by(income) %>%
  summarise(median_space_income_phase3 = median(ellipse_area_phase3))
std_activity_space_income_phase3 <- participants_activity_df_income_phase3 %>%
  group_by(income) %>%
  summarise(std_space_income_phase3 = sd(ellipse_area_phase3))
length_income_seq_phase3 <- participants_activity_df_income_phase3 %>%
  group_by(income) %>%
```

```
  mutate(rownumberincome_phase3 = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_phase3 = max(rownumberincome_phase3))
activity_space_income_all_phase3 <- bind_cols(mean_activity_space_income_phase3,
                                    median_activity_space_income_phase3 %>%
                                      select(-income),
                                    std_activity_space_income_phase3 %>%
                                      select(-income),
                                    length_income_seq_phase3 %>% select(-income))
length_income_seq_phase3 %>% select(-income)

participants_activity_df_income_phase3$income <-
  factor(participants_activity_df_income_phase3$income
                                         , levels=c('4 000 CHF or less',
                                                    '4 001 - 8 000 CHF','8 001 - 12
                                                      000 CHF',
                                                    '12 001 - 16 000 CHF','More
                                                      than 16 000 CHF',
                                                    'Prefer not to say'))
boxplot(ellipse_area_phase3 ~ income, participants_activity_df_income_phase3,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Phase 3', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_phase3 <-
  mean(activity_space_income_all_phase3$mean_space_income_phase3)
median_income_all_phase3 <-
  median(activity_space_income_all_phase3$median_space_income_phase3)
sd_mean_income_all_phase3 <-
  sd(activity_space_income_all_phase3$mean_space_income_phase3)
sd_median_income_all_phase3 <-
  sd(activity_space_income_all_phase3$median_space_income_phase3)

key_figures_income_all_phase3 <- c(length_income_all_phase3, mean_income_all_phase3,
                                   sd_mean_income_all_phase3, median_income_all_phase3,
                                   sd_median_income_all_phase3)

participants_activity_df_income_phase4 <- participants_activity_df_phase4 %>%
  filter(income!=99)
length_income_all_phase4 <- nrow(participants_activity_df_income_phase4)

mean_activity_space_income_phase4 <- participants_activity_df_income_phase4 %>%
  group_by(income) %>%
  summarise(mean_space_income_phase4 = mean(ellipse_area_phase4))
median_activity_space_income_phase4 <- participants_activity_df_income_phase4 %>%
  group_by(income) %>%
```

```
  summarise(median_space_income_phase4 = median(ellipse_area_phase4))
std_activity_space_income_phase4 <- participants_activity_df_income_phase4 %>%
  group_by(income) %>%
  summarise(std_space_income_phase4 = sd(ellipse_area_phase4))
length_income_seq_phase4 <- participants_activity_df_income_phase4 %>%
  group_by(income) %>%
  mutate(rownumberincome_phase4 = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_phase4 = max(rownumberincome_phase4))
activity_space_income_all_phase4 <- bind_cols(mean_activity_space_income_phase4,
                                    median_activity_space_income_phase4 %>%
                                      select(-income),
                                    std_activity_space_income_phase4 %>%
                                      select(-income),
                                    length_income_seq_phase4 %>% select(-income))
length_income_seq_phase4 %>% select(-income)

participants_activity_df_income_phase4$income <-
  factor(participants_activity_df_income_phase4$income
                                          , levels=c('4 000 CHF or less',
                                                     '4 001 - 8 000 CHF','8 001 - 12
                                                       000 CHF',
                                                     '12 001 - 16 000 CHF','More
                                                       than 16 000 CHF',
                                                     'Prefer not to say'))
boxplot(ellipse_area_phase4 ~ income, participants_activity_df_income_phase4,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Phase 4', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_phase4 <-
  mean(activity_space_income_all_phase4$mean_space_income_phase4)
median_income_all_phase4 <-
  median(activity_space_income_all_phase4$median_space_income_phase4)
sd_mean_income_all_phase4 <-
  sd(activity_space_income_all_phase4$mean_space_income_phase4)
sd_median_income_all_phase4 <-
  sd(activity_space_income_all_phase4$median_space_income_phase4)

key_figures_income_all_phase4 <- c(length_income_all_phase4, mean_income_all_phase4,
                                   sd_mean_income_all_phase4, median_income_all_phase4,
                                   sd_median_income_all_phase4)

participants_activity_df_income_phase5 <- participants_activity_df_phase5 %>%
  filter(income!=99)
length_income_all_phase5 <- nrow(participants_activity_df_income_phase5)
```

```
mean_activity_space_income_phase5 <- participants_activity_df_income_phase5 %>%
  group_by(income) %>%
  summarise(mean_space_income_phase5 = mean(ellipse_area_phase5))
median_activity_space_income_phase5 <- participants_activity_df_income_phase5 %>%
  group_by(income) %>%
  summarise(median_space_income_phase5 = median(ellipse_area_phase5))
std_activity_space_income_phase5 <- participants_activity_df_income_phase5 %>%
  group_by(income) %>%
  summarise(std_space_income_phase5 = sd(ellipse_area_phase5))
length_income_seq_phase5 <- participants_activity_df_income_phase5 %>%
  group_by(income) %>%
  mutate(rownumberincome_phase5 = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_phase5 = max(rownumberincome_phase5))
activity_space_income_all_phase5 <- bind_cols(mean_activity_space_income_phase5,
                                  median_activity_space_income_phase5 %>%
                                    select(-income),
                                  std_activity_space_income_phase5 %>%
                                    select(-income),
                                  length_income_seq_phase5 %>% select(-income))
length_income_seq_phase5 %>% select(-income)

participants_activity_df_income_phase5$income <-
  factor(participants_activity_df_income_phase5$income
                                        , levels=c('4 000 CHF or less',
                                                   '4 001 - 8 000 CHF','8 001 - 12
                                                     000 CHF',
                                                   '12 001 - 16 000 CHF','More
                                                     than 16 000 CHF',
                                                   'Prefer not to say'))
boxplot(ellipse_area_phase5 ~ income, participants_activity_df_income_phase5,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Phase 5', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_phase5 <-
  mean(activity_space_income_all_phase5$mean_space_income_phase5)
median_income_all_phase5 <-
  median(activity_space_income_all_phase5$median_space_income_phase5)
sd_mean_income_all_phase5 <-
  sd(activity_space_income_all_phase5$mean_space_income_phase5)
sd_median_income_all_phase5 <-
  sd(activity_space_income_all_phase5$median_space_income_phase5)

key_figures_income_all_phase5 <- c(length_income_all_phase5, mean_income_all_phase5,
```

```
                                sd_mean_income_all_phase5, median_income_all_phase5,
                                sd_median_income_all_phase5)

participants_activity_df_income_phase6 <- participants_activity_df_phase6 %>%
  filter(income!=99)
length_income_all_phase6 <- nrow(participants_activity_df_income_phase6)

mean_activity_space_income_phase6 <- participants_activity_df_income_phase6 %>%
  group_by(income) %>%
  summarise(mean_space_income_phase6 = mean(ellipse_area_phase6))
median_activity_space_income_phase6 <- participants_activity_df_income_phase6 %>%
  group_by(income) %>%
  summarise(median_space_income_phase6 = median(ellipse_area_phase6))
std_activity_space_income_phase6 <- participants_activity_df_income_phase6 %>%
  group_by(income) %>%
  summarise(std_space_income_phase6 = sd(ellipse_area_phase6))
length_income_seq_phase6 <- participants_activity_df_income_phase6 %>%
  group_by(income) %>%
  mutate(rownumberincome_phase6 = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income_phase6 = max(rownumberincome_phase6))
activity_space_income_all_phase6 <- bind_cols(mean_activity_space_income_phase6,
                                  median_activity_space_income_phase6 %>%
                                    select(-income),
                                  std_activity_space_income_phase6 %>%
                                    select(-income),
                                  length_income_seq_phase6 %>% select(-income))
length_income_seq_phase6 %>% select(-income)

participants_activity_df_income_phase6$income <-
  factor(participants_activity_df_income_phase6$income
                                        , levels=c('4 000 CHF or less',
                                                   '4 001 - 8 000 CHF','8 001 - 12
                                                     000 CHF',
                                                   '12 001 - 16 000 CHF','More
                                                     than 16 000 CHF',
                                                   'Prefer not to say'))
boxplot(ellipse_area_phase6 ~ income, participants_activity_df_income_phase6,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Phase 6', ylab = 'Activity space [km2]',
       xlab = 'Income group', ylim = c(-100,16000), las = 1)

mean_income_all_phase6 <-
  mean(activity_space_income_all_phase6$mean_space_income_phase6)
median_income_all_phase6 <-
  median(activity_space_income_all_phase6$median_space_income_phase6)
```

```
sd_mean_income_all_phase6 <-
  sd(activity_space_income_all_phase6$mean_space_income_phase6)
sd_median_income_all_phase6 <-
  sd(activity_space_income_all_phase6$median_space_income_phase6)

key_figures_income_all_phase6 <- c(length_income_all_phase6, mean_income_all_phase6,
                                   sd_mean_income_all_phase6, median_income_all_phase6,
                                   sd_median_income_all_phase6)

ellipse_area_income1_baseline <- participants_activity_df_income_baseline %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_baseline <- bind_cols(ellipse_area_income1_baseline,
                                           'Baseline')
ellipse_area_income1_ph1 <- participants_activity_df_income_phase1 %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_ph1 <- bind_cols(ellipse_area_income1_ph1,
                                      'Phase 1')
ellipse_area_income1_ph2 <- participants_activity_df_income_phase2 %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_ph2 <- bind_cols(ellipse_area_income1_ph2,
                                      'Phase 2')
ellipse_area_income1_ph3 <- participants_activity_df_income_phase3 %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_ph3 <- bind_cols(ellipse_area_income1_ph3,
                                      'Phase 3')
ellipse_area_income1_ph4 <- participants_activity_df_income_phase4 %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_ph4 <- bind_cols(ellipse_area_income1_ph4,
                                      'Phase 4')
ellipse_area_income1_ph5 <- participants_activity_df_income_phase5 %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_ph5 <- bind_cols(ellipse_area_income1_ph5,
                                      'Phase 5')
ellipse_area_income1_ph6 <- participants_activity_df_income_phase6 %>%
  filter(income == '4 000 CHF or less')
ellipse_area_income1_ph6 <- bind_cols(ellipse_area_income1_ph6,
                                      'Phase 6')

ellipse_area_income1_baseline <- ellipse_area_income1_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 136
  )
ellipse_area_income1_ph1 <- ellipse_area_income1_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
```

```
    phase = 136
  )
ellipse_area_income1_ph2 <- ellipse_area_income1_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 136
  )
ellipse_area_income1_ph3 <- ellipse_area_income1_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 136
  )
ellipse_area_income1_ph4 <- ellipse_area_income1_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 136
  )
ellipse_area_income1_ph5 <- ellipse_area_income1_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 136
  )
ellipse_area_income1_ph6 <- ellipse_area_income1_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 136
  )

ellipse_area_income1 <- bind_rows(ellipse_area_income1_baseline,
                                  ellipse_area_income1_ph1, ellipse_area_income1_ph2,
                                  ellipse_area_income1_ph3, ellipse_area_income1_ph4,
                                  ellipse_area_income1_ph5, ellipse_area_income1_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_income1,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Income group 4 000 CHF or less', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_income1_plot <- ellipse_area_income1 %>%
  group_by(phase) %>%
  summarise(ellipse_area_income_1 = median(ellipse_area))

ellipse_area_income2_baseline <- participants_activity_df_income_baseline %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_baseline <- bind_cols(ellipse_area_income2_baseline,
                                           'Baseline')
```

```
ellipse_area_income2_ph1 <- participants_activity_df_income_phase1 %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_ph1 <- bind_cols(ellipse_area_income2_ph1,
                                      'Phase 1')
ellipse_area_income2_ph2 <- participants_activity_df_income_phase2 %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_ph2 <- bind_cols(ellipse_area_income2_ph2,
                                      'Phase 2')
ellipse_area_income2_ph3 <- participants_activity_df_income_phase3 %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_ph3 <- bind_cols(ellipse_area_income2_ph3,
                                      'Phase 3')
ellipse_area_income2_ph4 <- participants_activity_df_income_phase4 %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_ph4 <- bind_cols(ellipse_area_income2_ph4,
                                      'Phase 4')
ellipse_area_income2_ph5 <- participants_activity_df_income_phase5 %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_ph5 <- bind_cols(ellipse_area_income2_ph5,
                                      'Phase 5')
ellipse_area_income2_ph6 <- participants_activity_df_income_phase6 %>%
  filter(income == '4 001 - 8 000 CHF')
ellipse_area_income2_ph6 <- bind_cols(ellipse_area_income2_ph6,
                                      'Phase 6')

ellipse_area_income2_baseline <- ellipse_area_income2_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 136
  )
ellipse_area_income2_ph1 <- ellipse_area_income2_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 136
  )
ellipse_area_income2_ph2 <- ellipse_area_income2_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 136
  )
ellipse_area_income2_ph3 <- ellipse_area_income2_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 136
  )
ellipse_area_income2_ph4 <- ellipse_area_income2_ph4 %>%
```

```
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 136
  )
ellipse_area_income2_ph5 <- ellipse_area_income2_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 136
  )
ellipse_area_income2_ph6 <- ellipse_area_income2_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 136
  )

ellipse_area_income2 <- bind_rows(ellipse_area_income2_baseline,
                                  ellipse_area_income2_ph1, ellipse_area_income2_ph2,
                                  ellipse_area_income2_ph3, ellipse_area_income2_ph4,
                                  ellipse_area_income2_ph5, ellipse_area_income2_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_income2,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Income group 4 001 - 8 000 CHF', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_income2_plot <- ellipse_area_income2 %>%
  group_by(phase) %>%
  summarise(ellipse_area_income_2 = median(ellipse_area))

ellipse_area_income3_baseline <- participants_activity_df_income_baseline %>%
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_baseline <- bind_cols(ellipse_area_income3_baseline,
                                           'Baseline')
ellipse_area_income3_ph1 <- participants_activity_df_income_phase1 %>%
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_ph1 <- bind_cols(ellipse_area_income3_ph1,
                                      'Phase 1')
ellipse_area_income3_ph2 <- participants_activity_df_income_phase2 %>%
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_ph2 <- bind_cols(ellipse_area_income3_ph2,
                                      'Phase 2')
ellipse_area_income3_ph3 <- participants_activity_df_income_phase3 %>%
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_ph3 <- bind_cols(ellipse_area_income3_ph3,
                                      'Phase 3')
ellipse_area_income3_ph4 <- participants_activity_df_income_phase4 %>%
```

```
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_ph4 <- bind_cols(ellipse_area_income3_ph4,
                                      'Phase 4')
ellipse_area_income3_ph5 <- participants_activity_df_income_phase5 %>%
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_ph5 <- bind_cols(ellipse_area_income3_ph5,
                                      'Phase 5')
ellipse_area_income3_ph6 <- participants_activity_df_income_phase6 %>%
  filter(income == '8 001 - 12 000 CHF')
ellipse_area_income3_ph6 <- bind_cols(ellipse_area_income3_ph6,
                                      'Phase 6')

ellipse_area_income3_baseline <- ellipse_area_income3_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 136
  )
ellipse_area_income3_ph1 <- ellipse_area_income3_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 136
  )
ellipse_area_income3_ph2 <- ellipse_area_income3_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 136
  )
ellipse_area_income3_ph3 <- ellipse_area_income3_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 136
  )
ellipse_area_income3_ph4 <- ellipse_area_income3_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 136
  )
ellipse_area_income3_ph5 <- ellipse_area_income3_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 136
  )
ellipse_area_income3_ph6 <- ellipse_area_income3_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 136
```

```
  )

ellipse_area_income3 <- bind_rows(ellipse_area_income3_baseline,
                                  ellipse_area_income3_ph1, ellipse_area_income3_ph2,
                                  ellipse_area_income3_ph3, ellipse_area_income3_ph4,
                                  ellipse_area_income3_ph5, ellipse_area_income3_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_income3,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Income group 8 001 - 12 000 CHF', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_income3_plot <- ellipse_area_income3 %>%
  group_by(phase) %>%
  summarise(ellipse_area_income_3 = median(ellipse_area))

ellipse_area_income4_baseline <- participants_activity_df_income_baseline %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_baseline <- bind_cols(ellipse_area_income4_baseline,
                                           'Baseline')
ellipse_area_income4_ph1 <- participants_activity_df_income_phase1 %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_ph1 <- bind_cols(ellipse_area_income4_ph1,
                                      'Phase 1')
ellipse_area_income4_ph2 <- participants_activity_df_income_phase2 %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_ph2 <- bind_cols(ellipse_area_income4_ph2,
                                      'Phase 2')
ellipse_area_income4_ph3 <- participants_activity_df_income_phase3 %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_ph3 <- bind_cols(ellipse_area_income4_ph3,
                                      'Phase 3')
ellipse_area_income4_ph4 <- participants_activity_df_income_phase4 %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_ph4 <- bind_cols(ellipse_area_income4_ph4,
                                      'Phase 4')
ellipse_area_income4_ph5 <- participants_activity_df_income_phase5 %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_ph5 <- bind_cols(ellipse_area_income4_ph5,
                                      'Phase 5')
ellipse_area_income4_ph6 <- participants_activity_df_income_phase6 %>%
  filter(income == '12 001 - 16 000 CHF')
ellipse_area_income4_ph6 <- bind_cols(ellipse_area_income4_ph6,
                                      'Phase 6')

ellipse_area_income4_baseline <- ellipse_area_income4_baseline %>%
```

```
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 136
  )
ellipse_area_income4_ph1 <- ellipse_area_income4_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 136
  )
ellipse_area_income4_ph2 <- ellipse_area_income4_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 136
  )
ellipse_area_income4_ph3 <- ellipse_area_income4_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 136
  )
ellipse_area_income4_ph4 <- ellipse_area_income4_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 136
  )
ellipse_area_income4_ph5 <- ellipse_area_income4_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 136
  )
ellipse_area_income4_ph6 <- ellipse_area_income4_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 136
  )

ellipse_area_income4 <- bind_rows(ellipse_area_income4_baseline,
                                  ellipse_area_income4_ph1, ellipse_area_income4_ph2,
                                  ellipse_area_income4_ph3, ellipse_area_income4_ph4,
                                  ellipse_area_income4_ph5, ellipse_area_income4_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_income4,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Income group 12 001 - 16 000 CHF', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_income4_plot <- ellipse_area_income4 %>%
```

```
  group_by(phase) %>%
  summarise(ellipse_area_income_4 = median(ellipse_area))

ellipse_area_income5_baseline <- participants_activity_df_income_baseline %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_baseline <- bind_cols(ellipse_area_income5_baseline,
                                           'Baseline')
ellipse_area_income5_ph1 <- participants_activity_df_income_phase1 %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_ph1 <- bind_cols(ellipse_area_income5_ph1,
                                      'Phase 1')
ellipse_area_income5_ph2 <- participants_activity_df_income_phase2 %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_ph2 <- bind_cols(ellipse_area_income5_ph2,
                                      'Phase 2')
ellipse_area_income5_ph3 <- participants_activity_df_income_phase3 %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_ph3 <- bind_cols(ellipse_area_income5_ph3,
                                      'Phase 3')
ellipse_area_income5_ph4 <- participants_activity_df_income_phase4 %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_ph4 <- bind_cols(ellipse_area_income5_ph4,
                                      'Phase 4')
ellipse_area_income5_ph5 <- participants_activity_df_income_phase5 %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_ph5 <- bind_cols(ellipse_area_income5_ph5,
                                      'Phase 5')
ellipse_area_income5_ph6 <- participants_activity_df_income_phase6 %>%
  filter(income == 'More than 16 000 CHF')
ellipse_area_income5_ph6 <- bind_cols(ellipse_area_income5_ph6,
                                      'Phase 6')

ellipse_area_income5_baseline <- ellipse_area_income5_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 136
  )
ellipse_area_income5_ph1 <- ellipse_area_income5_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 136
  )
ellipse_area_income5_ph2 <- ellipse_area_income5_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 136
```

```
  )
ellipse_area_income5_ph3 <- ellipse_area_income5_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 136
  )
ellipse_area_income5_ph4 <- ellipse_area_income5_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 136
  )
ellipse_area_income5_ph5 <- ellipse_area_income5_ph5 %>%
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 136
  )
ellipse_area_income5_ph6 <- ellipse_area_income5_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 136
  )

ellipse_area_income5 <- bind_rows(ellipse_area_income5_baseline,
                                  ellipse_area_income5_ph1, ellipse_area_income5_ph2,
                                  ellipse_area_income5_ph3, ellipse_area_income5_ph4,
                                  ellipse_area_income5_ph5, ellipse_area_income5_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_income5,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Income group more than 16 000 CHF', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_income5_plot <- ellipse_area_income5 %>%
  group_by(phase) %>%
  summarise(ellipse_area_income_5 = median(ellipse_area))



ellipse_area_income6_baseline <- participants_activity_df_income_baseline %>%
  filter(income == 'Prefer not to say')
ellipse_area_income6_baseline <- bind_cols(ellipse_area_income6_baseline,
                                           'Baseline')
ellipse_area_income6_ph1 <- participants_activity_df_income_phase1 %>%
  filter(income == 'Prefer not to say')
ellipse_area_income6_ph1 <- bind_cols(ellipse_area_income6_ph1,
                                      'Phase 1')
ellipse_area_income6_ph2 <- participants_activity_df_income_phase2 %>%
```

```
  filter(income == 'Prefer not to say')
ellipse_area_income6_ph2 <- bind_cols(ellipse_area_income6_ph2,
                                      'Phase 2')
ellipse_area_income6_ph3 <- participants_activity_df_income_phase3 %>%
  filter(income == 'Prefer not to say')
ellipse_area_income6_ph3 <- bind_cols(ellipse_area_income6_ph3,
                                      'Phase 3')
ellipse_area_income6_ph4 <- participants_activity_df_income_phase4 %>%
  filter(income == 'Prefer not to say')
ellipse_area_income6_ph4 <- bind_cols(ellipse_area_income6_ph4,
                                      'Phase 4')
ellipse_area_income6_ph5 <- participants_activity_df_income_phase5 %>%
  filter(income == 'Prefer not to say')
ellipse_area_income6_ph5 <- bind_cols(ellipse_area_income6_ph5,
                                      'Phase 5')
ellipse_area_income6_ph6 <- participants_activity_df_income_phase6 %>%
  filter(income == 'Prefer not to say')
ellipse_area_income6_ph6 <- bind_cols(ellipse_area_income6_ph6,
                                      'Phase 6')

ellipse_area_income6_baseline <- ellipse_area_income6_baseline %>%
  rename(
    ellipse_area = ellipse_area_baseline,
    phase = 136
  )
ellipse_area_income6_ph1 <- ellipse_area_income6_ph1 %>%
  rename(
    ellipse_area = ellipse_area_phase1,
    phase = 136
  )
ellipse_area_income6_ph2 <- ellipse_area_income6_ph2 %>%
  rename(
    ellipse_area = ellipse_area_phase2,
    phase = 136
  )
ellipse_area_income6_ph3 <- ellipse_area_income6_ph3 %>%
  rename(
    ellipse_area = ellipse_area_phase3,
    phase = 136
  )
ellipse_area_income6_ph4 <- ellipse_area_income6_ph4 %>%
  rename(
    ellipse_area = ellipse_area_phase4,
    phase = 136
  )
ellipse_area_income6_ph5 <- ellipse_area_income6_ph5 %>%
```

```
  rename(
    ellipse_area = ellipse_area_phase5,
    phase = 136
  )
ellipse_area_income6_ph6 <- ellipse_area_income6_ph6 %>%
  rename(
    ellipse_area = ellipse_area_phase6,
    phase = 136
  )

ellipse_area_income6 <- bind_rows(ellipse_area_income6_baseline,
                                  ellipse_area_income6_ph1, ellipse_area_income6_ph2,
                                  ellipse_area_income6_ph3, ellipse_area_income6_ph4,
                                  ellipse_area_income6_ph5, ellipse_area_income6_ph6)

boxplot(ellipse_area ~ phase, ellipse_area_income6,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = 'Income group Prefer not to say', ylab = 'Activity space [km2]',
        xlab = 'Phase', ylim = c(-100,14000), las = 1)

ellipse_area_income6_plot <- ellipse_area_income6 %>%
  group_by(phase) %>%
  summarise(ellipse_area_income_6 = median(ellipse_area))
```

A.5 R-Code: Weekly activity spaces

```
#Activity Space for all Participants

library(dplyr)
library(sf)
library(car)
library(ggmap)
library(ggplot2)
library(lubridate)
library(sp)
library(tidyr)
library(Hmisc)
library(epade)

setwd("C:/Users/melii/Desktop/Master Project Thesis/Data/Uhlmann")

#Import files of different phases
weight <- read.csv('enrichments/ipf_weights_sample.csv')
covid_activities <- read.csv("MOBIS_MDCEV/tracking/activities.csv", na="NA")



#combine the x and y value to one coordinate
activity_df_all <- st_as_sf(covid_activities, coords = c('x','y'),
                            crs = "+proj=longlat +datum=WGS84 +ellps=WGS84
                              +towgs84=0,0,0")
st_crs(activity_df_all)

#set the correct time for the data
activity_df_all$started_at <- as.POSIXct(activity_df_all$started_at
  ,tz=Sys.timezone())
activity_df_all$finished_at <- as.POSIXct(activity_df_all$finished_at
  ,tz=Sys.timezone())
activity_df_all$week <- format(activity_df_all$started_at, '%V')
activity_df_all$year <- format(activity_df_all$started_at, '%Y')
activity_df_all <- activity_df_all %>%
  filter(activity_df_all$year == '2020')
activity_df_all <- activity_df_all %>%
  filter(activity_df_all$week >= '04' & activity_df_all$week <= '40')



#Filter the participants and the trips in Switzerland
weight <- weight %>%
  filter(participant_ID %in% unique(participants$participant_ID))
activity_df <- activity_df_all %>%
```

```
  filter(user_id %in% unique(participants$participant_ID))
activity_df <- activity_df %>%
  filter(activity_df$in_switzerland == 'TRUE')
activity_df <- activity_df %>%
  filter(activity_df$was_confirmed == 'TRUE')



#Activity spaces for all the participants
group_and_calculate_ellipse_area_weekly <- function(df1) {
  df1 %>%
    group_by(user_id, phase, week_start, week) %>%
    group_modify( ~calculate_ellipse_area(.), keep=T)

}
calculate_ellipse <- function(df1) {
  area = NA
  max_radius = NA
  try({
    coordinates <- st_coordinates(df1$geometry)
    weighted_cov=cov.wt(coordinates, df1$duration)
    w_centroid <- weighted_cov$center

    pyth = sweep(coordinates, 2, w_centroid)**2
    dist = sqrt(pyth[,1] + pyth[,2])
    max_dist = max(dist)

    w_vcov_matrix <- weighted_cov$cov

    p <- ifelse(
      anyNA(w_vcov_matrix), NA, prod(eigen(w_vcov_matrix)$values)
    )
    ellipse_radius = ifelse(is.na(p) || p < 0, NA, sqrt(p))
    area <- pi * 5.991* ellipse_radius / (1000*1000)
  }, silent = F)
  return (list(area=area, max_radius=max_dist))
}

calculate_ellipse_area <- function(df1, grouping_variables) {
  res <- calculate_ellipse(sf::st_as_sf(df1))
  n_activities = nrow(df1)
  return (data.frame( ellipse_area=res$area, max_radius=res$max_radius,
     n_activities=n_activities))
}

#Prepare data to calculate weekly activity spaces for the participants
```

```
activity_df$weekday <- format(activity_df$started_at, '%u')
activity_df$week <- as.integer(activity_df$week)
activity_df$day <- format(activity_df$started_at, '%j')
activity_df$day <- as.integer(activity_df$day)
activity_df$weekday <- as.integer(activity_df$weekday)
activity_df$week_start <- floor_date(activity_df$started_at, unit = 'week',
  week_start =getOption('lubridate.week.start',1))

#calculate daily num activities and ellipse area for each person
activity_spaces_df <- activity_df %>%
  group_modify(~ group_and_calculate_ellipse_area_weekly(.x), keep = TRUE) %>%
  ungroup()
activity_spaces_df <- activity_spaces_df[rowSums(is.na(activity_spaces_df)) == 0,]
write.csv(activity_space_df, 'WeeklyActivitySpaces.csv')


#getting the same length of argument
participants_for_analysise <- participants %>%
  filter(participant_ID %in% unique(activity_spaces_df$user_id))
activity_spaces_df_for_analysise <- activity_spaces_df %>%
  filter(user_id %in% unique(participants$participant_ID))

#add weights
weight$user_id <- weight$participant_ID
weight$week_start <- weight$week
weight$week_start <- as.POSIXct(weight$week_start)
weight <- weight %>%
  filter(user_id %in% unique(activity_spaces_df_for_analysise$user_id))
weighted_activity_spaces <- inner_join(activity_spaces_df_for_analysise, weight, by =
  c('user_id','week_start'))
weighted_activity_spaces$weighted_ellipse_area <-
  weighted_activity_spaces$ellipse_area * weighted_activity_spaces$weight
activity_spaces_df_for_analysise <- weighted_activity_spaces

#Save results for all participants, before plotting
participants_for_analysise$user_id <- participants_for_analysise$participant_ID
participants_activity_df <- left_join(participants_for_analysise,
  activity_spaces_df_for_analysise, by ='user_id')
participants_activity_df$AgeGroup <- cut(participants_activity_df$age, breaks =
  c(seq(10, 65, by=10), Inf))
write.csv(participants_activity_df, 'activity_spaces_results.csv')

participants_activity_df$nonGrouping <- cut(participants_activity_df$age, breaks =
  c(seq(10, 65, by=100), Inf))
plot2 <- boxplot(ellipse_area ~ nonGrouping, participants_activity_df,
               cex.main =1, cex.axis = 1.2, cex.lab = 1.2,
```

```
              main = 'Activity space distribution',
              ylab = 'all participants', xlab = 'Activity space [km^2]',
              ylim = c(0,1000), las = 1, horizontal = TRUE)


#histogram for ellipse area
ggplot(data=participants_activity_df, aes(x=ellipse_area))+
  geom_histogram(breaks=seq(0, 2800, by=10),
              alpha = .2, color = 'steelblue', fill = 'steelblue')+
  geom_boxplot()+
  coord_cartesian(xlim=c(0, 2800))+
  xlab('ellipse area [km^2]') +
  ylab('Frequency') +
  theme(axis.title.x = element_text(size = 20), axis.text.x = element_text(size = 15)
       ,axis.title.y = element_text(size = 20), axis.text.y = element_text(size = 15))


#key figures all participants
length_all <- nrow(participants_activity_df)
mean_all <- mean(participants_activity_df$ellipse_area, na.rm = TRUE)
median_all <- median(participants_activity_df$ellipse_area, na.rm = TRUE)
sd_all <- sd(participants_activity_df$ellipse_area, na.rm = TRUE)
quantile1_all <- quantile(participants_activity_df$ellipse_area, 0.25, na.rm = TRUE)
quantile2_all <- quantile(participants_activity_df$ellipse_area, 0.75, na.rm = TRUE)

key_figures_all <- c(length_all, mean_all, median_all, sd_all,quantile1_all,
   quantile2_all)


#activity space vs age
participants_activity_df_age <- participants_activity_df %>%
  filter(!is.na(age))
length_age_all <- nrow(participants_activity_df_age)
mean_activity_space_age <- participants_activity_df_age %>%
  group_by(age) %>%
  summarise(mean_space_age = mean(ellipse_area))
median_activity_space_age <- participants_activity_df_age %>%
  group_by(age) %>%
  summarise(median_space_age = median(ellipse_area))
std_activity_space_age <- participants_activity_df_age %>%
  group_by(age) %>%
  summarise(std_space_age = sd(ellipse_area))
activity_space_age <- bind_cols(mean_activity_space_age, median_activity_space_age %>%
                              select(-age), std_activity_space_age %>% select(-age))
```

```
#activity space vs AgeGroup
participants_activity_df_age$AgeGroup <- cut(participants_activity_df_age$age, breaks
  =
                                          c(seq(10,65, by=10), Inf))
mean_activity_space_age <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  summarise(mean_space_age = weighted.mean(ellipse_area))
median_activity_space_age <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  summarise(median_space_age = median(ellipse_area))
std_activity_space_age <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  summarise(std_space_age = sd(ellipse_area))
length_AgeGroup <- participants_activity_df_age %>%
  group_by(AgeGroup) %>%
  mutate(rownumberAge = row_number())%>%
  group_by(AgeGroup) %>%
  summarise(length_space_age = max(rownumberAge))
activity_space_age <- bind_cols(mean_activity_space_age, median_activity_space_age
                            %>%select(-AgeGroup),
                            std_activity_space_age %>% select(-AgeGroup),
                            length_AgeGroup %>%select(-AgeGroup))


boxplot(ellipse_area ~ week_start, w=weight, data = participants_activity_df, type =
  'boxplot',
       ylim = c(-10, 2000),
       cex.axis = 1.2, cex.lab =1.4,
       main = ' ',
       ylab = 'Activity space [km^2]', xlab = 'Week of the year')

boxplot(ellipse_area ~AgeGroup, participants_activity_df_age,
       cex.main = 1, cex.axis = 1, cex.lab =1,
       main ='',
       ylab = 'Activity space [km^2]', xlab = 'Age group of the participants',
       ylim = c(-10,1000), las = 1,
       par(mar = c(4,6,1,1)))


participants_activity_df_age1 <- participants_activity_df %>%
  filter(AgeGroup == '(10,20]')
participants_activity_df_age2 <- participants_activity_df %>%
  filter(AgeGroup == '(20,30]')
participants_activity_df_age3 <- participants_activity_df %>%
  filter(AgeGroup == '(30,40]')
participants_activity_df_age4 <- participants_activity_df %>%
```

```
  filter(AgeGroup == '(40,50]')
participants_activity_df_age5 <- participants_activity_df %>%
  filter(AgeGroup == '(50,60]')
participants_activity_df_age6 <- participants_activity_df %>%
  filter(AgeGroup == '(60,Inf]')

participants_activity_df_age1 <- participants_activity_df_age1 %>%
  group_by(week_start) %>%
  summarise(mean1=weighted.mean(ellipse_area, weight))
participants_activity_df_age2 <- participants_activity_df_age2 %>%
  group_by(week_start) %>%
  summarise(mean2=weighted.mean(ellipse_area, weight))
participants_activity_df_age3 <- participants_activity_df_age3 %>%
  group_by(week_start) %>%
  summarise(mean3=weighted.mean(ellipse_area, weight))
participants_activity_df_age4 <- participants_activity_df_age4 %>%
  group_by(week_start) %>%
  summarise(mean4=weighted.mean(ellipse_area, weight))
participants_activity_df_age5 <- participants_activity_df_age5 %>%
  group_by(week_start) %>%
  summarise(mean5=weighted.mean(ellipse_area, weight))
participants_activity_df_age6 <- participants_activity_df_age6 %>%
  group_by(week_start) %>%
  summarise(mean6=weighted.mean(ellipse_area, weight))

participants_activity_df_age_plot <- inner_join(participants_activity_df_age1,
  participants_activity_df_age2,
                                    by = 'week_start')
participants_activity_df_age_plot <- inner_join(participants_activity_df_age_plot,
  participants_activity_df_age3,
                                          by = 'week_start')
participants_activity_df_age_plot <- inner_join(participants_activity_df_age_plot,
  participants_activity_df_age4,
                                          by = 'week_start')
participants_activity_df_age_plot <- inner_join(participants_activity_df_age_plot,
  participants_activity_df_age5,
                                          by = 'week_start')
participants_activity_df_age_plot <- inner_join(participants_activity_df_age_plot,
  participants_activity_df_age6,
                                          by = 'week_start')

ggplot(data = participants_activity_df_age_plot) +
  geom_line(aes(x= week_start,y=mean1, color='1', linetype='1'), size=1) +
  geom_line(aes(x=week_start,y=mean2, color='2', linetype='2'), size=1) +
  geom_line(aes(x=week_start,y=mean3, color='3', linetype='3'), size=1) +
  geom_line(aes(x=week_start,y=mean4, color='4', linetype='4'), size=1) +
```

```
  geom_line(aes(x=week_start,y=mean5, color='5', linetype='5'), size=1) +
  geom_line(aes(x=week_start,y=mean6, color='6', linetype='6'), size=1) +
  labs(y='Activity spaces [km^2]', x= 'Date') +
  scale_color_manual(values = c('#CCCCCC', '#999999', '#666666', '#333333',
    '#000000', '#666666'),
                  name = 'Age group', labels = c('(10,20]','(20,30]', '(30,40]',
                                          '(40,50]', '(50,60]', '(60,Inf]'))+
  scale_linetype_manual(values=c(1,1,2,3,1,4), name = 'Age group',
                      labels = c('(10,20]','(20,30]', '(30,40]',
                                 '(40,50]', '(50,60]', '(60,Inf]'))+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))


op <- par(mar = c(5,4,4,2) + 0.1)
par(op)

boxplot(ellipse_area ~ age, participants_activity_df_age,
        cex.main = 1, cex.axis = 1, cex.lab = 1,
        main = 'Activity space distribution by age',
        ylab = 'Activity space [km^2]', xlab = 'Age group of the participants',
        ylim = c(-100,2000), las=1)

mean_age_all <- mean(activity_space_age$mean_space_age)
median_age_all <- median(activity_space_age$median_space_age)
sd_mean_age_all <- sd(activity_space_age$mean_space_age)
sd_median_age_all <- sd(activity_space_age$median_space_age)


key_figures_age_all <- c(length_age_all, mean_age_all, sd_mean_age_all,
   median_age_all, sd_median_age_all)


#household size
participants_activity_df_household <- participants_activity_df %>%
  filter(!is.na(household_size))
length_household_all <- nrow(participants_activity_df_household)

mean_activity_space_household <- participants_activity_df_household %>%
  group_by(household_size) %>%
  summarise(mean_space_household = mean(ellipse_area))
median_activity_space_household <- participants_activity_df_household %>%
  group_by(household_size) %>%
  summarise(median_space_household = median(ellipse_area))
std_activity_space_household <- participants_activity_df_household %>%
```

```
  group_by(household_size) %>%
  summarise(std_space_household = sd(ellipse_area))
length_household_seq <- participants_activity_df_household %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold =row_number()) %>%
  summarise(length_space_household = max(rownumberhousehold))
activity_space_household <- bind_cols(mean_activity_space_household,
   median_activity_space_household %>%
                                     select(-household_size),
                                       std_activity_space_household %>%
                                     select(-household_size), length_household_seq %>%
                                     select(-household_size))

boxplot(ellipse_area ~ household_size, participants_activity_df_household,
        cex.main = 1, cex.axis = 1, cex.lab = 1,
        main ='', ylab = 'Activity space [km^2]', xlab = 'Household size',
        ylim = c(-100,1000), las=1)

mean_household_all <- mean(activity_space_household$mean_space_household)
median_household_all <- median(activity_space_household$median_space_household)
sd_mean_household_all <- sd(activity_space_household$mean_space_household)
sd_median_household_all <- sd(activity_space_household$median_space_household)

key_figures_household_all <- c(length_household_all, mean_household_all,
   sd_mean_household_all,
                              median_household_all, sd_median_household_all)

participants_activity_df_household1 <- participants_activity_df %>%
  filter(household_size == '1')
participants_activity_df_household2 <- participants_activity_df %>%
  filter(household_size == '2')
participants_activity_df_household3 <- participants_activity_df %>%
  filter(household_size == '3')
participants_activity_df_household4 <- participants_activity_df %>%
  filter(household_size == '4')
participants_activity_df_household5 <- participants_activity_df %>%
  filter(household_size == '5 or more')

participants_activity_df_household1 <- participants_activity_df_household1 %>%
  group_by(week_start) %>%
  summarise(mean1=weighted.mean(ellipse_area, weight))
participants_activity_df_household2 <- participants_activity_df_household2 %>%
  group_by(week_start) %>%
  summarise(mean2=weighted.mean(ellipse_area, weight))
participants_activity_df_household3 <- participants_activity_df_household3 %>%
  group_by(week_start) %>%
```

```
  summarise(mean3=weighted.mean(ellipse_area, weight))
participants_activity_df_household4 <- participants_activity_df_household4 %>%
  group_by(week_start) %>%
  summarise(mean4=weighted.mean(ellipse_area, weight))
participants_activity_df_household5 <- participants_activity_df_household5 %>%
  group_by(week_start) %>%
  summarise(mean5=weighted.mean(ellipse_area, weight))


participants_activity_df_household_plot <-
   inner_join(participants_activity_df_household1, participants_activity_df_household2,
                                          by = 'week_start')
participants_activity_df_household_plot <-
   inner_join(participants_activity_df_household_plot,
   participants_activity_df_household3,
                                          by = 'week_start')
participants_activity_df_household_plot <-
   inner_join(participants_activity_df_household_plot,
   participants_activity_df_household4,
                                          by = 'week_start')
participants_activity_df_household_plot <-
   inner_join(participants_activity_df_household_plot,
   participants_activity_df_household5,
                                          by = 'week_start')


ggplot(data = participants_activity_df_household_plot) +
  geom_line(aes(x= week_start,y=mean1, color='1', linetype='1'), size=1) +
  geom_line(aes(x=week_start,y=mean2, color='2', linetype='2'), size=1) +
  geom_line(aes(x=week_start,y=mean3, color='3', linetype='3'), size=1) +
  geom_line(aes(x=week_start,y=mean4, color='4', linetype='4'), size=1) +
  geom_line(aes(x=week_start,y=mean5, color='5', linetype='5'), size=1) +
   labs(y='Activity space [km^2]', x= 'Date') +
  scale_color_manual(values = c('#CCCCCC', '#999999', '#666666', '#333333',
    '#000000', '#666666'),
                  name = 'Household size', labels = c('1','2', '3',
                                          '4', '5 or more'))+
  scale_linetype_manual(values=c(1,1,2,3,1,4), name = 'Household size',
                      labels = c('1','2', '3',
                                 '4', '5 or more'))+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))
```

```
#activity space vs income
participants_activity_df_income <- participants_activity_df %>%
  filter(income!=99)
length_income_all <- nrow(participants_activity_df_income)

mean_activity_space_income <- participants_activity_df_income %>%
  group_by(income) %>%
  summarise(mean_space_income = mean(ellipse_area))
median_activity_space_income <- participants_activity_df_income %>%
  group_by(income) %>%
  summarise(median_space_income = median(ellipse_area))
std_activity_space_income <- participants_activity_df_income %>%
  group_by(income) %>%
  summarise(std_space_income = sd(ellipse_area))
length_income_seq <- participants_activity_df_income %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
activity_space_income_all <- bind_cols(mean_activity_space_income,
                                       median_activity_space_income %>% select(-income),
                                       std_activity_space_income %>% select(-income),
                                       length_income_seq %>% select(-income))
length_income_seq %>% select(-income)

participants_activity_df_income$income <-
  factor(participants_activity_df_income$income
                                            , levels=c('4 000 CHF or less',
                                                       '4 001 - 8 000 CHF','8 001 - 12
                                                         000 CHF',
                                                       '12 001 - 16 000 CHF','More
                                                         than 16 000 CHF',
                                                       'Prefer not to say'))
boxplot(ellipse_area ~ income, participants_activity_df_income,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = '', ylab = 'Activity space [km^2]',
       xlab = 'Income group', ylim = c(-100,1000), las = 1)

mean_income_all <- mean(activity_space_income_all$mean_space_income)
median_income_all <- median(activity_space_income_all$median_space_income)
sd_mean_income_all <- sd(activity_space_income_all$mean_space_income)
sd_median_income_all <- sd(activity_space_income_all$median_space_income)

key_figures_income_all <- c(length_income_all, mean_income_all, sd_mean_income_all,
                            median_income_all, sd_median_income_all)
```

```
participants_activity_df_income1 <- participants_activity_df %>%
  filter(income == '4 000 CHF or less')
participants_activity_df_income2 <- participants_activity_df %>%
  filter(income == '4 001 - 8 000 CHF')
participants_activity_df_income3 <- participants_activity_df %>%
  filter(income == '8 001 - 12 000 CHF')
participants_activity_df_income4 <- participants_activity_df %>%
  filter(income == '12 001 - 16 000 CHF')
participants_activity_df_income5 <- participants_activity_df %>%
  filter(income == 'More than 16 000 CHF')
participants_activity_df_income6 <- participants_activity_df %>%
  filter(income == 'Prefer not to say')

participants_activity_df_income1 <- participants_activity_df_income1 %>%
  group_by(week_start) %>%
  summarise(mean1=weighted.mean(ellipse_area, weight))
participants_activity_df_income2 <- participants_activity_df_income2 %>%
  group_by(week_start) %>%
  summarise(mean2=weighted.mean(ellipse_area, weight))
participants_activity_df_income3 <- participants_activity_df_income3 %>%
  group_by(week_start) %>%
  summarise(mean3=weighted.mean(ellipse_area, weight))
participants_activity_df_income4 <- participants_activity_df_income4 %>%
  group_by(week_start) %>%
  summarise(mean4=weighted.mean(ellipse_area, weight))
participants_activity_df_income5 <- participants_activity_df_income5 %>%
  group_by(week_start) %>%
  summarise(mean5=weighted.mean(ellipse_area, weight))
participants_activity_df_income6 <- participants_activity_df_income6 %>%
  group_by(week_start) %>%
  summarise(mean6=weighted.mean(ellipse_area, weight))

participants_activity_df_income_plot <- inner_join(participants_activity_df_income1,
  participants_activity_df_income2,
                                        by = 'week_start')
participants_activity_df_income_plot <-
  inner_join(participants_activity_df_income_plot, participants_activity_df_income3,
                                        by = 'week_start')
participants_activity_df_income_plot <-
  inner_join(participants_activity_df_income_plot, participants_activity_df_income4,
                                        by = 'week_start')
participants_activity_df_income_plot <-
  inner_join(participants_activity_df_income_plot, participants_activity_df_income5,
                                        by = 'week_start')
participants_activity_df_income_plot <-
```

```
  inner_join(participants_activity_df_income_plot, participants_activity_df_income6,
                                          by = 'week_start')

ggplot(data = participants_activity_df_income_plot) +
  geom_line(aes(x= week_start,y=mean1, color='1', linetype='1'), size=1) +
  geom_line(aes(x=week_start,y=mean2, color='2', linetype='2'), size=1) +
  geom_line(aes(x=week_start,y=mean3, color='3', linetype='3'), size=1) +
  geom_line(aes(x=week_start,y=mean4, color='4', linetype='4'), size=1) +
  geom_line(aes(x=week_start,y=mean5, color='5', linetype='5'), size=1) +
  geom_line(aes(x=week_start,y=mean6, color='6', linetype='6'), size=1) +
  labs(y='Activity space [km^2]', x= 'Date') +
  scale_color_manual(values = c('#CCCCCC', '#999999', '#666666', '#333333',
    '#000000', '#666666'),
                     name = 'Income group', labels = c('4 000 CHF or less','4 001 - 8
                       000 CHF', '8 001 - 12 000 CHF',
                                          '12 001 - 16 000 CHF', 'More than 16 000
                                            CHF', 'Prefer not to say'))+
  scale_linetype_manual(values=c(1,4,1,1,3,2), name = 'Income group',
                        labels = c('4 000 CHF or less','4 001 - 8 000 CHF', '8 001 - 12
                          000 CHF',
                                   '12 001 - 16 000 CHF', 'More than 16 000 CHF',
                                     'Prefer not to say'))+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))



#Analysis of the participants for the discussion of the project
participants$AgeGroup <- cut(participants$age, breaks =
                                          c(seq(10,65, by=10), Inf))
participants_students <- participants %>%
  filter(main_employment == 'Student')
participants_young <- participants %>%
  filter(AgeGroup == '(10,20]')
participants_students_young <- participants %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(10,20]')
participants_students_young_lowincome <- participants %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(income == '4 000 CHF or less')
participants_students_young_notsaying <- participants %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(income == 'Prefer not to say')
```

```
participants_5 <- participants %>%
  filter(household_size == '5 or more')
participants_5_notsaying <- participants %>%
  filter(household_size == '5 or more') %>%
  filter(income == 'Prefer not to say')
participants_notsaying <- participants %>%
  filter(income == 'Prefer not to say')
participants_notsaying_5 <- participants %>%
  filter(income == 'Prefer not to say') %>%
  filter(household_size == '5 or more')
participants_students_2030 <- participants %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(20,30]')
participants_2030 <- participants %>%
  filter(AgeGroup == '(20,30]')
```

A.6 R-Code: Innovation rate

```
#Innovation rate and new places visited

library(dplyr)
library(sf)
library(car)
library(ggmap)
library(ggplot2)
library(lubridate)
library(sp)
library(dbscan)
library(RColorBrewer)

setwd("C:/Users/melii/Desktop/Master Project Thesis/Data/Uhlmann")

#Import files of different phases
weight <- read.csv('enrichments/ipf_weights_sample.csv')
covid_activities <- read.csv("MOBIS_MDCEV/tracking/activities.csv", na="NA")



#combine the x and y value to one coordinate
activity_df_all <- st_as_sf(covid_activities, coords = c('x','y'),
                            crs = "+proj=longlat +datum=WGS84 +ellps=WGS84
                              +towgs84=0,0,0")
st_crs(activity_df_all)

#set the correct time for the data
activity_df_all$started_at <- as.POSIXct(activity_df_all$started_at
  ,tz=Sys.timezone())
activity_df_all$finished_at <- as.POSIXct(activity_df_all$finished_at
  ,tz=Sys.timezone())
activity_df_all$week <- format(activity_df_all$started_at, '%V')
activity_df_all$year <- format(activity_df_all$started_at, '%Y')
activity_df_all <- activity_df_all %>%
  filter(activity_df_all$year == '2020')
activity_df_all <- activity_df_all %>%
  filter(activity_df_all$week >= '04' & activity_df_all$week <= '40')

#Filter the participants and the trips in Switzerland
activity_df_innovation <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE')
activity_df_innovation_phase1 <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE') %>%
```

```
  filter(started_at >= '2020-02-03') %>%
  filter(finished_at <= '2020-03-15')
activity_df_innovation_phase2 <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE') %>%
  filter(started_at >= '2020-03-16') %>%
  filter(finished_at <= '2020-04-26')
activity_df_innovation_phase3 <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE') %>%
  filter(started_at >= '2020-04-27') %>%
  filter(finished_at <= '2020-06-07')
activity_df_innovation_phase4 <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE') %>%
  filter(started_at >= '2020-06-08') %>%
  filter(finished_at <= '2020-07-19')
activity_df_innovation_phase5 <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE') %>%
  filter(started_at >= '2020-07-20') %>%
  filter(finished_at <= '2020-08-30')
activity_df_innovation_phase6 <- activity_df_all %>%
  filter(user_id %in% unique(participants_activities$user_id)) %>%
  filter(in_switzerland == 'TRUE') %>%
  filter(started_at >= '2020-08-31') %>%
  filter(started_at <= '2020-10-11')


#get innovation for one person/user; cluster first and let the earliest date for
   location
#this will create cluster id's for eps values of 10, 20 and 50 meters
db <- function(x) {
  activity_coordinates <- x %>% st_coordinates()
  cluster_10 = dbscan::dbscan(activity_coordinates, eps = 10, minPts = 3)$cluster
  cluster_20 = dbscan::dbscan(activity_coordinates, eps = 20, minPts = 3)$cluster
  cluster_50 = dbscan::dbscan(activity_coordinates, eps = 50, minPts = 3)$cluster
  return(data.frame(cluster_10m=cluster_10,cluster_20m=cluster_20,
                    cluster_50m=cluster_50))
  #return(data.frame(cluster_50m=cluster_50))
  }


#innovation for all users
```

```
#use the dbscan algorithm above to cluster activities
activity_cluster_df <- activity_innovation %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()
activity_cluster_df_phase1 <- activity_innovation_phase1 %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()
activity_cluster_df_phase2 <- activity_innovation_phase2 %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()
activity_cluster_df_phase3 <- activity_innovation_phase3 %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()
activity_cluster_df_phase4 <- activity_innovation_phase4 %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()
activity_cluster_df_phase5 <- activity_innovation_phase5 %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()
activity_cluster_df_phase6 <- activity_innovation_phase6 %>%
  group_by(user_id) %>%
  group_modify(~db(.x)) %>% ungroup()

#join this back to the activities, don't duplicate user_id column
activity_df_with_clusters <- bind_cols(activity_innovation, activity_cluster_df %>%
                                       select(-user_id))
activity_df_with_clusters_phase1 <- bind_cols(activity_innovation_phase1,
                                            activity_cluster_df_phase1 %>%
                                       select(-user_id))
activity_df_with_clusters_phase2 <- bind_cols(activity_innovation_phase2,
                                            activity_cluster_df_phase2 %>%
                                       select(-user_id))
activity_df_with_clusters_phase3 <- bind_cols(activity_innovation_phase3,
                                            activity_cluster_df_phase3 %>%
                                       select(-user_id))
activity_df_with_clusters_phase4 <- bind_cols(activity_innovation_phase4,
                                            activity_cluster_df_phase4 %>%
                                       select(-user_id))
activity_df_with_clusters_phase5 <- bind_cols(activity_innovation_phase5,
                                            activity_cluster_df_phase5 %>%
                                       select(-user_id))
activity_df_with_clusters_phase6 <- bind_cols(activity_innovation_phase6,
                                            activity_cluster_df_phase6 %>%
                                       select(-user_id))

#this block gives id's to the outliers
```

```
non_outlier_activities <- activity_df_with_clusters %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)
non_outlier_activities_phase1 <- activity_df_with_clusters_phase1 %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)
non_outlier_activities_phase2 <- activity_df_with_clusters_phase2 %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)
non_outlier_activities_phase3 <- activity_df_with_clusters_phase3 %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)
non_outlier_activities_phase4 <- activity_df_with_clusters_phase4 %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)
non_outlier_activities_phase5 <- activity_df_with_clusters_phase5 %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)
non_outlier_activities_phase6 <- activity_df_with_clusters_phase6 %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_50m_no_outliers = cluster_50m)

#give outliers a cluster_id as well
outlier_activities <- activity_df_with_clusters %>%
  group_by(user_id) %>%
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df <- rbind(non_outlier_activities, outlier_activities)

outlier_activities_phase1 <- activity_df_with_clusters_phase1 %>%
  group_by(user_id) %>%
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df_phase1 <- rbind(non_outlier_activities_phase1,
   outlier_activities_phase1)

outlier_activities_phase2 <- activity_df_with_clusters_phase2 %>%
  group_by(user_id) %>%
```

```
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df_phase2 <- rbind(non_outlier_activities_phase2,
   outlier_activities_phase2)

outlier_activities_phase3 <- activity_df_with_clusters_phase3 %>%
  group_by(user_id) %>%
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df_phase3 <- rbind(non_outlier_activities_phase3,
   outlier_activities_phase3)

outlier_activities_phase4 <- activity_df_with_clusters_phase4 %>%
  group_by(user_id) %>%
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df_phase4 <- rbind(non_outlier_activities_phase4,
   outlier_activities_phase4)

outlier_activities_phase5 <- activity_df_with_clusters_phase5 %>%
  group_by(user_id) %>%
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df_phase5 <- rbind(non_outlier_activities_phase5,
   outlier_activities_phase5)

outlier_activities_phase6 <- activity_df_with_clusters_phase6 %>%
  group_by(user_id) %>%
  mutate(max_cluster_id = max(cluster_50m)) %>%
  filter(cluster_50m == 0) %>%
```

```
  arrange(started_at) %>%
  mutate(cluster_50m_no_outliers = max_cluster_id + row_number()) %>%
  select(-max_cluster_id)

re_clustered_activities_df_phase6 <- rbind(non_outlier_activities_phase6,
   outlier_activities_phase6)

#this block sets to true if the activity location wasn't seen before
re_clustered_activities_df1 <- re_clustered_activities_df %>%
  arrange(user_id, started_at) %>%
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                     unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))

re_clustered_activities_df1_phase1 <- re_clustered_activities_df_phase1 %>%
  arrange(user_id, started_at) %>%
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                     unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))

re_clustered_activities_df1_phase2 <- re_clustered_activities_df_phase2 %>%
  arrange(user_id, started_at) %>%
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                     unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))

re_clustered_activities_df1_phase3 <- re_clustered_activities_df_phase3 %>%
  arrange(user_id, started_at) %>%
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                     unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))

re_clustered_activities_df1_phase4 <- re_clustered_activities_df_phase4 %>%
  arrange(user_id, started_at) %>%
```

```
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                          unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))

re_clustered_activities_df1_phase5 <- re_clustered_activities_df_phase5 %>%
  arrange(user_id, started_at) %>%
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                          unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))

re_clustered_activities_df1_phase6 <- re_clustered_activities_df_phase6 %>%
  arrange(user_id, started_at) %>%
  group_by(user_id) %>%
  filter(cluster_50m > 0) %>%
  mutate(cluster_chronological_id = match(cluster_50m_no_outliers,
                                          unique(cluster_50m_no_outliers))) %>%
  group_by(user_id, day_g = floor_date(started_at, 'day')) %>%
  mutate(max_crono_id_day = max(cluster_chronological_id))


#get the max chronological id for the previous day
previous_max_crono_id_days_s <- re_clustered_activities_df1 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

previous_max_crono_id_days_s_phase1 <- re_clustered_activities_df1_phase1 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

previous_max_crono_id_days_s_phase2 <- re_clustered_activities_df1_phase2 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
```

```
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

previous_max_crono_id_days_s_phase3 <- re_clustered_activities_df1_phase3 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

previous_max_crono_id_days_s_phase4 <- re_clustered_activities_df1_phase4 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

previous_max_crono_id_days_s_phase5 <- re_clustered_activities_df1_phase5 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

previous_max_crono_id_days_s_phase6 <- re_clustered_activities_df1_phase6 %>%
  st_drop_geometry() %>%
  group_by(user_id, day_g) %>%
  summarise(max_crono_id_day = max(max_crono_id_day)) %>%
  group_by(user_id) %>%
  mutate(prev_max_crono_id_day = lag(max_crono_id_day, default=0))

#check each day to see if the chronological ids are greater than the max
#from the previous day, if so it is a new location
re_clustered_activites_df_with_new_activity_indicator <- re_clustered_activities_df1
  %>%
  left_join(previous_max_crono_id_days_s) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)

a <- re_clustered_activites_df_with_new_activity_indicator %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))

re_clustered_activites_df_with_new_activity_indicator_phase1 <-
```

```
  re_clustered_activities_df1_phase1 %>%
  left_join(previous_max_crono_id_days_s_phase1) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)

a_phase1 <- re_clustered_activites_df_with_new_activity_indicator_phase1 %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))

re_clustered_activites_df_with_new_activity_indicator_phase2 <-
  re_clustered_activities_df1_phase2 %>%
  left_join(previous_max_crono_id_days_s_phase2) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)

a_phase2 <- re_clustered_activites_df_with_new_activity_indicator_phase2 %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))

re_clustered_activites_df_with_new_activity_indicator_phase3 <-
  re_clustered_activities_df1_phase3 %>%
  left_join(previous_max_crono_id_days_s_phase3) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)

a_phase3 <- re_clustered_activites_df_with_new_activity_indicator_phase3 %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))

re_clustered_activites_df_with_new_activity_indicator_phase4 <-
  re_clustered_activities_df1_phase4 %>%
  left_join(previous_max_crono_id_days_s_phase4) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)
```

```
a_phase4 <- re_clustered_activites_df_with_new_activity_indicator_phase4 %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))

re_clustered_activites_df_with_new_activity_indicator_phase5 <-
   re_clustered_activities_df1_phase5 %>%
  left_join(previous_max_crono_id_days_s_phase5) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)

a_phase5 <- re_clustered_activites_df_with_new_activity_indicator_phase5 %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))

re_clustered_activites_df_with_new_activity_indicator_phase6 <-
   re_clustered_activities_df1_phase6 %>%
  left_join(previous_max_crono_id_days_s_phase6) %>%
  group_by(user_id) %>%
  mutate(is_new_activity_location_on_day = cluster_chronological_id >
    prev_max_crono_id_day) %>%
  select(-day_g, -max_crono_id_day, -prev_max_crono_id_day)

a_phase6 <- re_clustered_activites_df_with_new_activity_indicator_phase6 %>%
  st_drop_geometry() %>%
  group_by(user_id, day=floor_date(started_at, 'day')) %>%
  summarise(new_activities = sum(is_new_activity_location_on_day))


#get the days numbered in order to compare day 1 for all users no matter what date it
   was
a_num_date <- a %>%
  group_by(user_id) %>%
  mutate(n_day = row_number()) %>%
  ungroup()

a_num_date_phase1 <- a_phase1 %>%
  group_by(user_id) %>%
  mutate(n_day = row_number()) %>%
  ungroup()

a_num_date_phase2 <- a_phase2 %>%
  group_by(user_id) %>%
```

```
  mutate(n_day = row_number()) %>%
  ungroup()

a_num_date_phase3 <- a_phase3 %>%
  group_by(user_id) %>%
  mutate(n_day = row_number()) %>%
  ungroup()

a_num_date_phase4 <- a_phase4 %>%
  group_by(user_id) %>%
  mutate(n_day = row_number()) %>%
  ungroup()

a_num_date_phase5 <- a_phase5 %>%
  group_by(user_id) %>%
  mutate(n_day = row_number()) %>%
  ungroup()

a_num_date_phase6 <- a_phase6 %>%
  group_by(user_id) %>%
  mutate(n_day = row_number()) %>%
  ungroup()


#determine which participants need to take into considerations
obs_time <- a_num_date %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno <- obs_time %>%
  filter(num_obs_day >= 40)
write.csv(num_teilnehmer_inno, 'teilnehmer_innovation.csv')

obs_time_phase1 <- a_num_date_phase1 %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno_phase1 <- obs_time_phase1 %>%
  filter(num_obs_day >= 14)
write.csv(num_teilnehmer_inno_phase1, 'teilnehmer_innovation_phase1.csv')

obs_time_phase2 <- a_num_date_phase2 %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno_phase2 <- obs_time_phase2 %>%
```

```
  filter(num_obs_day >= 14)
write.csv(num_teilnehmer_inno_phase2, 'teilnehmer_innovation_phase2.csv')

obs_time_phase3 <- a_num_date_phase3 %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno_phase3 <- obs_time_phase3 %>%
  filter(num_obs_day >= 14)
write.csv(num_teilnehmer_inno_phase3, 'teilnehmer_innovation_phase3.csv')

obs_time_phase4 <- a_num_date_phase4 %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno_phase4 <- obs_time_phase4 %>%
  filter(num_obs_day >= 14)
write.csv(num_teilnehmer_inno_phase4, 'teilnehmer_innovation_phase4.csv')

obs_time_phase5 <- a_num_date_phase5 %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno_phase5 <- obs_time_phase5 %>%
  filter(num_obs_day >= 14)
write.csv(num_teilnehmer_inno_phase5, 'teilnehmer_innovation_phase5.csv')

obs_time_phase6 <- a_num_date_phase6 %>%
  group_by(user_id) %>%
  summarise(num_obs_day = max(n_day))

num_teilnehmer_inno_phase6 <- obs_time_phase6 %>%
  filter(num_obs_day >= 14)
write.csv(num_teilnehmer_inno_phase6, 'teilnehmer_innovation_phase6.csv')


#line plot for all new activities, these are not in order for days, rather in
   chronological sequence
ggplot(a) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = ' ', x = 'Oberservation period',
       y = 'New activity locations', color ='')+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))
```

```
ggplot(a_phase1) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = 'New activities phase 1', x = 'Oberservation period',
       y = 'New activity locations', color ='')

ggplot(a_phase2) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = 'New activities phase 2', x = 'Oberservation period',
       y = 'New activity locations', color ='')

ggplot(a_phase3) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = 'New activities phase 3', x = 'Oberservation period',
       y = 'New activity locations', color ='')

ggplot(a_phase4) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = 'New activities phase 4', x = 'Oberservation period',
       y = 'New activity locations', color ='')

ggplot(a_phase5) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = 'New activities phase 5', x = 'Oberservation period',
       y = 'New activity locations', color ='')

ggplot(a_phase6) +
  geom_line(aes(x=day, y=new_activities))+
  labs(title = 'New activities phase 6', x = 'Oberservation period',
       y = 'New activity locations', color ='')

#Plot the mean for the innovation rate over the observation time
a$week_start <- floor_date(a$day, unit = 'week', week_start
   =getOption('lubridate.week.start',1))
boxplot_inno <- a %>%
  group_by(week_start) %>%
  summarise(mean=mean(new_activities))
ggplot(boxplot_inno) +
  geom_line(aes(x=week_start, y=mean))+
  labs(title = ' ', x = 'Date',
       y = 'New activity locations [new activities/day]', color ='')+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))


#plot point to get the parameters for the ggplot
```

```
plot(a_num_date$n_day, a_num_date$new_activities, xlab = 'Day of observation',
    ylab = 'New activities', main = 'Number of new activities by day of observation')
regModel <- lm(new_activities ~ n_day, data = a_num_date)
abline(regModel$coef[1], regModel$coef[2], col= 'red', lwd = 3)
regModel_day_new_activities <- summary(regModel)


#group for id and get a rate for each user
innovation_rate_compare <- a_num_date %>%
  group_by(user_id)%>%
  summarise(rate = mean(new_activities))

innovation_rate <- mean(innovation_rate_compare$rate)

innovation_rate_compare_phase1 <- a_num_date_phase1 %>%
  group_by(user_id)%>%
  summarise(rate = mean(new_activities))

innovation_rate_phase1 <- mean(innovation_rate_compare_phase1$rate)

innovation_rate_compare_phase2 <- a_num_date_phase2 %>%
  group_by(user_id)%>%
  summarise(rate = mean(new_activities))

innovation_rate_phase2 <- mean(innovation_rate_compare_phase2$rate)

innovation_rate_compare_phase3 <- a_num_date_phase3 %>%
  group_by(user_id)%>%
  summarise(rate = mean(new_activities))

innovation_rate_phase3 <- mean(innovation_rate_compare_phase3$rate)

innovation_rate_compare_phase4 <- a_num_date_phase4 %>%
  group_by(user_id)%>%
  summarise(rate = mean(new_activities))

innovation_rate_phase4 <- mean(innovation_rate_compare_phase4$rate)

innovation_rate_compare_phase5 <- a_num_date_phase5 %>%
  group_by(user_id)%>%
  summarise(rate = mean(new_activities))

innovation_rate_phase5 <- mean(innovation_rate_compare_phase5$rate)

innovation_rate_compare_phase6 <- a_num_date_phase6 %>%
  group_by(user_id)%>%
```

```
  summarise(rate = mean(new_activities))

innovation_rate_phase6 <- mean(innovation_rate_compare_phase6$rate)


#add participants data to rate
innovation_participants_df_filtered <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare$user_id))
innovation_participants_df <- bind_cols(innovation_participants_df_filtered,
                                        innovation_rate_compare)

innovation_participants_df_filtered_phase1 <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare_phase1$user_id))
innovation_participants_df_phase1 <-
   bind_cols(innovation_participants_df_filtered_phase1,
                                        innovation_rate_compare_phase1)

innovation_participants_df_filtered_phase2 <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare_phase2$user_id))
innovation_participants_df_phase2 <-
   bind_cols(innovation_participants_df_filtered_phase2,
                                           innovation_rate_compare_phase2)

innovation_participants_df_filtered_phase3 <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare_phase3$user_id))
innovation_participants_df_phase3 <-
   bind_cols(innovation_participants_df_filtered_phase3,
                                           innovation_rate_compare_phase3)

innovation_participants_df_filtered_phase4 <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare_phase4$user_id))
innovation_participants_df_phase4 <-
   bind_cols(innovation_participants_df_filtered_phase4,
                                           innovation_rate_compare_phase4)

innovation_participants_df_filtered_phase5 <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare_phase5$user_id))
innovation_participants_df_phase5 <-
   bind_cols(innovation_participants_df_filtered_phase5,
                                           innovation_rate_compare_phase5)

innovation_participants_df_filtered_phase6 <- participants %>%
  filter(participant_ID %in% unique(innovation_rate_compare_phase6$user_id))
innovation_participants_df_phase6 <-
   bind_cols(innovation_participants_df_filtered_phase6,
                                           innovation_rate_compare_phase6)
```

```
#key figures for the participants_innovation data
label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno <- nrow(innovation_participants_df)
mean_inno <- mean(innovation_participants_df$rate)
median_inno <- median(innovation_participants_df$rate)
std_inno <- sd(innovation_participants_df$rate)
quantile1_inno <- quantile(innovation_participants_df$rate, 0.25)
quantile2_inno <- quantile(innovation_participants_df$rate, 0.75)
figures_innovation_total <- c(length_inno, mean_inno, median_inno,
                          std_inno, quantile1_inno, quantile2_inno)


innovation_participants_df$AgeGroup <- cut(innovation_participants_df$age, breaks =
                                     c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate over the whole observation time',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'Age group of the participant', ylim = c(-0.2, 2), las = 1)

label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno_phase1 <- nrow(innovation_participants_df_phase1)
mean_inno_phase1 <- mean(innovation_participants_df_phase1$rate)
median_inno_phase1 <- median(innovation_participants_df_phase1$rate)
std_inno_phase1 <- sd(innovation_participants_df_phase1$rate)
quantile1_inno_phase1 <- quantile(innovation_participants_df_phase1$rate, 0.25)
quantile2_inno_phase1 <- quantile(innovation_participants_df_phase1$rate, 0.75)
figures_innovation_total_phase1 <- c(length_inno_phase1, mean_inno_phase1,
  median_inno_phase1,
                          std_inno_phase1, quantile1_inno_phase1,
                            quantile2_inno_phase1)

innovation_participants_df_phase1$AgeGroup <-
  cut(innovation_participants_df_phase1$age, breaks =
                                     c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df_phase1,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate for phase 1',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'All participants', ylim = c(-0.2, 2), las = 1)

label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno_phase2 <- nrow(innovation_participants_df_phase2)
mean_inno_phase2 <- mean(innovation_participants_df_phase2$rate)
```

```
median_inno_phase2 <- median(innovation_participants_df_phase2$rate)
std_inno_phase2 <- sd(innovation_participants_df_phase2$rate)
quantile1_inno_phase2 <- quantile(innovation_participants_df_phase2$rate, 0.25)
quantile2_inno_phase2 <- quantile(innovation_participants_df_phase2$rate, 0.75)
figures_innovation_total_phase2 <- c(length_inno_phase2, mean_inno_phase,
  median_inno_phase2,
                                std_inno_phase2, quantile1_inno_phase2,
                                  quantile2_inno_phase2)

innovation_participants_df_phase2$AgeGroup <-
  cut(innovation_participants_df_phase2$age, breaks =
                                           c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df_phase2,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate for phase 2',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'All participants', ylim = c(-0.2, 2), las = 1)

label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno_phase3 <- nrow(innovation_participants_df_phase3)
mean_inno_phase3 <- mean(innovation_participants_df_phase3$rate)
median_inno_phase3 <- median(innovation_participants_df_phase3$rate)
std_inno_phase3 <- sd(innovation_participants_df_phase3$rate)
quantile1_inno_phase3 <- quantile(innovation_participants_df_phase3$rate, 0.25)
quantile2_inno_phase3 <- quantile(innovation_participants_df_phase3$rate, 0.75)
figures_innovation_total_phase3 <- c(length_inno_phase3, mean_inno_phase3,
  median_inno_phase3,
                                std_inno_phase3, quantile1_inno_phase3,
                                  quantile2_inno_phase3)

innovation_participants_df_phase3$AgeGroup <-
  cut(innovation_participants_df_phase3$age, breaks =
                                           c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df_phase3,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate for phase 3',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'All participants', ylim = c(-0.2, 2), las = 1)

label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno_phase4 <- nrow(innovation_participants_df_phase4)
mean_inno_phase4 <- mean(innovation_participants_df_phase4$rate)
median_inno_phase4 <- median(innovation_participants_df_phase4$rate)
std_inno_phase4 <- sd(innovation_participants_df_phase4$rate)
quantile1_inno_phase4 <- quantile(innovation_participants_df_phase4$rate, 0.25)
quantile2_inno_phase4 <- quantile(innovation_participants_df_phase4$rate, 0.75)
```

```
figures_innovation_total_phase4 <- c(length_inno_phase4, mean_inno_phase4,
  median_inno_phase4,
                                std_inno_phase4, quantile1_inno_phase4,
                                  quantile2_inno_phase4)

innovation_participants_df_phase4$AgeGroup <-
  cut(innovation_participants_df_phase4$age, breaks =
                                           c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df_phase4,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate for phase 4',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'All participants', ylim = c(-0.2, 2), las = 1)

label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno_phase5 <- nrow(innovation_participants_df_phase5)
mean_inno_phase5 <- mean(innovation_participants_df_phase5$rate)
median_inno_phase5 <- median(innovation_participants_df_phase5$rate)
std_inno_phase5 <- sd(innovation_participants_df_phase5$rate)
quantile1_inno_phase5 <- quantile(innovation_participants_df_phase5$rate, 0.25)
quantile2_inno_phase5 <- quantile(innovation_participants_df_phase5$rate, 0.75)
figures_innovation_total_phase5 <- c(length_inno_phase5, mean_inno_phase5,
  median_inno_phase5,
                                std_inno_phase5, quantile1_inno_phase5,
                                  quantile2_inno_phase5)

innovation_participants_df_phase5$AgeGroup <-
  cut(innovation_participants_df_phase5$age, breaks =
                                           c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df_phase5,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate for phase 5',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'All participants', ylim = c(-0.2, 2), las = 1)

label <- c('length','mean','std. deviation','25%-quantile','75%-quantile')
length_inno_phase6 <- nrow(innovation_participants_df_phase6)
mean_inno_phase6 <- mean(innovation_participants_df_phase6$rate)
median_inno_phase6 <- median(innovation_participants_df_phase6$rate)
std_inno_phase6 <- sd(innovation_participants_df_phase6$rate)
quantile1_inno_phase6 <- quantile(innovation_participants_df_phase6$rate, 0.25)
quantile2_inno_phase6 <- quantile(innovation_participants_df_phase6$rate, 0.75)
figures_innovation_total_phase6 <- c(length_inno_phase6, mean_inno_phase6,
  median_inno_phase6,
                                std_inno_phase6, quantile1_inno_phase6,
                                  quantile2_inno_phase6)
```

```
innovation_participants_df_phase6$AgeGroup <-
  cut(innovation_participants_df_phase6$age, breaks =
                                              c(seq(10, 65, by = 10), Inf))
boxplot(rate ~ AgeGroup, innovation_participants_df_phase6,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = 'Innovation rate for phase 6',
       ylab = 'Innovation rate [new activities/day]',
       xlab = 'All participants', ylim = c(-0.2, 2), las = 1)

innovation_participants_df_phase1$phase <- 1
innovation_participants_df_phase2$phase <- 2
innovation_participants_df_phase3$phase <- 3
innovation_participants_df_phase4$phase <- 4
innovation_participants_df_phase5$phase <- 5
innovation_participants_df_phase6$phase <- 6

innovation_plot_allphases <- bind_rows(innovation_participants_df_phase1,
  innovation_participants_df_phase2,
                                  innovation_participants_df_phase3,
                                  innovation_participants_df_phase4,
                                  innovation_participants_df_phase5,
                                  innovation_participants_df_phase6)

innovation_plot_agegroup1_phase1 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup1_phase2 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup1_phase3 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup1_phase4 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup1_phase5 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(10,20]') %>%
```

```
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup1_phase6 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup1 <-
   bind_rows(innovation_plot_agegroup1_phase1,innovation_plot_agegroup1_phase2,
                                  innovation_plot_agegroup1_phase3,
                                  innovation_plot_agegroup1_phase4,
                                  innovation_plot_agegroup1_phase5,
                                  innovation_plot_agegroup1_phase6)

innovation_plot_agegroup2_phase1 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(20,30]') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup2_phase2 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(20,30]') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup2_phase3 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(20,30]') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup2_phase4 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(20,30]') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup2_phase5 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(20,30]') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup2_phase6 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(20,30]') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup2 <-
```

```
  bind_rows(innovation_plot_agegroup2_phase1,innovation_plot_agegroup2_phase2,
                                  innovation_plot_agegroup2_phase3,
                                  innovation_plot_agegroup2_phase4,
                                  innovation_plot_agegroup2_phase5,
                                  innovation_plot_agegroup2_phase6)

innovation_plot_agegroup3_phase1 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(30,40]') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup3_phase2 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(30,40]') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup3_phase3 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(30,40]') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup3_phase4 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(30,40]') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup3_phase5 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(30,40]') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup3_phase6 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(30,40]') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup3 <-
  bind_rows(innovation_plot_agegroup3_phase1,innovation_plot_agegroup3_phase2,
                                  innovation_plot_agegroup3_phase3,
                                  innovation_plot_agegroup3_phase4,
                                  innovation_plot_agegroup3_phase5,
                                  innovation_plot_agegroup3_phase6)

innovation_plot_agegroup4_phase1 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(40,50]') %>%
  filter(phase == 1) %>%
```

```
  summarise(mean = mean(rate))

innovation_plot_agegroup4_phase2 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(40,50]') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup4_phase3 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(40,50]') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup4_phase4 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(40,50]') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup4_phase5 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(40,50]') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup4_phase6 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(40,50]') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup4 <-
   bind_rows(innovation_plot_agegroup4_phase1,innovation_plot_agegroup4_phase2,
                                    innovation_plot_agegroup4_phase3,
                                    innovation_plot_agegroup4_phase4,
                                    innovation_plot_agegroup4_phase5,
                                    innovation_plot_agegroup4_phase6)

innovation_plot_agegroup5_phase1 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(50,60]') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup5_phase2 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(50,60]') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup5_phase3 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(50,60]') %>%
```

```
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup5_phase4 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(50,60]') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup5_phase5 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(50,60]') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup5_phase6 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(50,60]') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup5 <-
   bind_rows(innovation_plot_agegroup5_phase1,innovation_plot_agegroup5_phase2,
                                  innovation_plot_agegroup5_phase3,
                                  innovation_plot_agegroup5_phase4,
                                  innovation_plot_agegroup5_phase5,
                                  innovation_plot_agegroup5_phase6)

innovation_plot_agegroup6_phase1 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(60,Inf]') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup6_phase2 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(60,Inf]') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup6_phase3 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(60,Inf]') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup6_phase4 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(60,Inf]') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup6_phase5 <- innovation_plot_allphases %>%
```

```
  filter(AgeGroup == '(60,Inf]') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup6_phase6 <- innovation_plot_allphases %>%
  filter(AgeGroup == '(60,Inf]') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_agegroup6 <-
   bind_rows(innovation_plot_agegroup6_phase1,innovation_plot_agegroup6_phase2,
                                    innovation_plot_agegroup6_phase3,
                                    innovation_plot_agegroup6_phase4,
                                    innovation_plot_agegroup6_phase5,
                                    innovation_plot_agegroup6_phase6)

innovation_plot_agegroup <- bind_cols(innovation_plot_agegroup1,
   innovation_plot_agegroup2,
                                  innovation_plot_agegroup3,
                                    innovation_plot_agegroup4,
                                  innovation_plot_agegroup5,
                                    innovation_plot_agegroup6)
innovation_plot_agegroup$phase <- c(1,2,3,4,5,6)

ggplot(data = innovation_plot_agegroup) +
  geom_line(aes(x= phase,y=mean...1, color='1', linetype='1'), size=1) +
  geom_line(aes(x=phase,y=mean...2, color='2', linetype='2'), size=1) +
  geom_line(aes(x=phase,y=mean...3, color='3', linetype='3'), size=1) +
  geom_line(aes(x=phase,y=mean...4, color='4', linetype='4'), size=1) +
  geom_line(aes(x=phase,y=mean...5, color='5', linetype='5'), size=1) +
  geom_line(aes(x=phase,y=mean...6, color='6', linetype='6'), size=1) +
  labs(y='Innovation rate', x = 'Phase') +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5, 6))+
  scale_color_manual(values = c('#CCCCCC', '#999999', '#666666', '#333333',
    '#000000', '#666666'),
                  name = 'Age group', labels = c('(10,20]','(20,30]', '(30,40]',
                                                 '(40,50]', '(50,60]', '(60,Inf]'))+
  scale_linetype_manual(values=c(1,1,2,3,1,4), name = 'Age group',
                      labels = c('(10,20]','(20,30]', '(30,40]',
                                 '(40,50]', '(50,60]', '(60,Inf]'))+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))


#innovation vs. income
```

```
innovation_participants_df_income <- innovation_participants_df %>%
  filter(income != 99)

innovation_participants_df_income_phase1 <- innovation_participants_df_phase1 %>%
  filter(income != 99)

innovation_participants_df_income_phase2 <- innovation_participants_df_phase2 %>%
  filter(income != 99)

innovation_participants_df_income_phase3 <- innovation_participants_df_phase3 %>%
  filter(income != 99)

innovation_participants_df_income_phase4 <- innovation_participants_df_phase4 %>%
  filter(income != 99)

innovation_participants_df_income_phase5 <- innovation_participants_df_phase5 %>%
  filter(income != 99)

innovation_participants_df_income_phase6 <- innovation_participants_df_phase6 %>%
  filter(income != 99)

innovation_participants_df_income$income <-
   factor(innovation_participants_df_income$income
                                          , levels=c('4 000 CHF or less',
                                                     '4 001 - 8 000 CHF','8 001 - 12
                                                       000 CHF',
                                                     '12 001 - 16 000 CHF','More
                                                       than 16 000 CHF',
                                                     'Prefer not to say'))

boxplot(rate ~income, innovation_participants_df_income,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main ='', ylab = 'Innovation rate [New activities/day]',
        xlab = 'Income group', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~income, innovation_participants_df_income_phase1,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main ='', ylab = 'Innovation rate [New activities/day]',
        xlab = 'Income group', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~income, innovation_participants_df_income_phase2,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main ='', ylab = 'Innovation rate [New activities/day]',
        xlab = 'Income group', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~income, innovation_participants_df_income_phase3,
```

```
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main ='', ylab = 'Innovation rate [New activities/day]',
       xlab = 'Income group', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~income, innovation_participants_df_income_phase4,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main ='', ylab = 'Innovation rate [New activities/day]',
       xlab = 'Income group', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~income, innovation_participants_df_income_phase5,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main ='', ylab = 'Innovation rate [New activities/day]',
       xlab = 'Income group', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~income, innovation_participants_df_income_phase6,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main ='', ylab = 'Innovation rate [New activities/day]',
       xlab = 'Income group', ylim = c(-0.2, 2), las = 1)


mean_innovation_income <- innovation_participants_df_income %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income <- innovation_participants_df_income %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income <- innovation_participants_df_income %>%
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income <- innovation_participants_df_income %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
innovation_income <- bind_cols(mean_innovation_income, median_innovation_income %>%
                                 select(-income), sd_innovation_income %>%
                                   select(-income),
                               length_innovation_income %>% select(-income))

mean_innovation_income_phase1 <- innovation_participants_df_income_phase1 %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income_phase1 <- innovation_participants_df_income_phase1 %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income_phase1 <- innovation_participants_df_income_phase1 %>%
```

```
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income_phase1 <- innovation_participants_df_income_phase1 %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
innovation_income_phase1 <- bind_cols(mean_innovation_income_phase1,
   median_innovation_income_phase1 %>%
                               select(-income), sd_innovation_income_phase1 %>%
                                 select(-income),
                              length_innovation_income_phase1 %>% select(-income))

mean_innovation_income_phase2 <- innovation_participants_df_income_phase2 %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income_phase2 <- innovation_participants_df_income_phase2 %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income_phase2 <- innovation_participants_df_income_phase2 %>%
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income_phase2 <- innovation_participants_df_income_phase2 %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
innovation_income_phase2 <- bind_cols(mean_innovation_income_phase2,
   median_innovation_income_phase2 %>%
                                    select(-income), sd_innovation_income_phase2 %>%
                                      select(-income),
                                   length_innovation_income_phase2 %>% select(-income))

mean_innovation_income_phase3 <- innovation_participants_df_income_phase3 %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income_phase3 <- innovation_participants_df_income_phase3 %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income_phase3 <- innovation_participants_df_income_phase3 %>%
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income_phase3 <- innovation_participants_df_income_phase3 %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
```

```
  summarise(length_space_income = max(rownumberincome))
innovation_income_phase3 <- bind_cols(mean_innovation_income_phase3,
   median_innovation_income_phase3 %>%
                                    select(-income), sd_innovation_income_phase3 %>%
                                      select(-income),
                                  length_innovation_income_phase3 %>% select(-income))

mean_innovation_income_phase4 <- innovation_participants_df_income_phase4 %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income_phase4 <- innovation_participants_df_income_phase4 %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income_phase4 <- innovation_participants_df_income_phase4 %>%
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income_phase4 <- innovation_participants_df_income_phase4 %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
innovation_income_phase4 <- bind_cols(mean_innovation_income_phase4,
   median_innovation_income_phase4 %>%
                                    select(-income), sd_innovation_income_phase4 %>%
                                      select(-income),
                                  length_innovation_income_phase4 %>% select(-income))

mean_innovation_income_phase5 <- innovation_participants_df_income_phase5 %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income_phase5 <- innovation_participants_df_income_phase5 %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income_phase5 <- innovation_participants_df_income_phase5 %>%
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income_phase5 <- innovation_participants_df_income_phase5 %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
innovation_income_phase5 <- bind_cols(mean_innovation_income_phase5,
   median_innovation_income_phase5 %>%
                                    select(-income), sd_innovation_income_phase5 %>%
                                      select(-income),
                                  length_innovation_income_phase5 %>% select(-income))
```

```
mean_innovation_income_phase6 <- innovation_participants_df_income_phase6 %>%
  group_by(income) %>%
  summarise(mean_inno_income = mean(rate))
median_innovation_income_phase6 <- innovation_participants_df_income_phase6 %>%
  group_by(income) %>%
  summarise(median_inno_income = median(rate))
sd_innovation_income_phase6 <- innovation_participants_df_income_phase6 %>%
  group_by(income) %>%
  summarise(sd_inno_income = sd(rate))
length_innovation_income_phase6 <- innovation_participants_df_income_phase6 %>%
  group_by(income) %>%
  mutate(rownumberincome = row_number()) %>%
  group_by(income) %>%
  summarise(length_space_income = max(rownumberincome))
innovation_income_phase6 <- bind_cols(mean_innovation_income_phase6,
   median_innovation_income_phase6 %>%
                                     select(-income), sd_innovation_income_phase6 %>%
                                       select(-income),
                                   length_innovation_income_phase6 %>% select(-income))



innovation_plot_income1_phase1 <- innovation_plot_allphases %>%
  filter(income == '4 000 CHF or less') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_income1_phase2 <- innovation_plot_allphases %>%
  filter(income == '4 000 CHF or less') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_income1_phase3 <- innovation_plot_allphases %>%
  filter(income == '4 000 CHF or less') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_income1_phase4 <- innovation_plot_allphases %>%
  filter(income == '4 000 CHF or less') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_income1_phase5 <- innovation_plot_allphases %>%
  filter(income == '4 000 CHF or less') %>%
  filter(phase == 5) %>%
```

```
  summarise(mean = mean(rate))

innovation_plot_income1_phase6 <- innovation_plot_allphases %>%
  filter(income == '4 000 CHF or less') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_income1 <-
   bind_rows(innovation_plot_income1_phase1,innovation_plot_income1_phase2,
                                    innovation_plot_income1_phase3,
                                    innovation_plot_income1_phase4,
                                    innovation_plot_income1_phase5,
                                    innovation_plot_income1_phase6)

innovation_plot_income2_phase1 <- innovation_plot_allphases %>%
  filter(income == '4 001 - 8 000 CHF') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_income2_phase2 <- innovation_plot_allphases %>%
  filter(income == '4 001 - 8 000 CHF') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_income2_phase3 <- innovation_plot_allphases %>%
  filter(income == '4 001 - 8 000 CHF') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_income2_phase4 <- innovation_plot_allphases %>%
  filter(income == '4 001 - 8 000 CHF') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_income2_phase5 <- innovation_plot_allphases %>%
  filter(income == '4 001 - 8 000 CHF') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_income2_phase6 <- innovation_plot_allphases %>%
  filter(income == '4 001 - 8 000 CHF') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_income2 <-
   bind_rows(innovation_plot_income2_phase1,innovation_plot_income2_phase2,
```

```
                                  innovation_plot_income2_phase3,
                                  innovation_plot_income2_phase4,
                                  innovation_plot_income2_phase5,
                                  innovation_plot_income2_phase6)

innovation_plot_income3_phase1 <- innovation_plot_allphases %>%
  filter(income == '8 001 - 12 000 CHF') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_income3_phase2 <- innovation_plot_allphases %>%
  filter(income == '8 001 - 12 000 CHF') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_income3_phase3 <- innovation_plot_allphases %>%
  filter(income == '8 001 - 12 000 CHF') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_income3_phase4 <- innovation_plot_allphases %>%
  filter(income == '8 001 - 12 000 CHF') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_income3_phase5 <- innovation_plot_allphases %>%
  filter(income == '8 001 - 12 000 CHF') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_income3_phase6 <- innovation_plot_allphases %>%
  filter(income == '8 001 - 12 000 CHF') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_income3 <-
   bind_rows(innovation_plot_income3_phase1,innovation_plot_income3_phase2,
                                  innovation_plot_income3_phase3,
                                  innovation_plot_income3_phase4,
                                  innovation_plot_income3_phase5,
                                  innovation_plot_income3_phase6)

innovation_plot_income4_phase1 <- innovation_plot_allphases %>%
  filter(income == '12 001 - 16 000 CHF') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))
```

```
innovation_plot_income4_phase2 <- innovation_plot_allphases %>%
  filter(income == '12 001 - 16 000 CHF') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_income4_phase3 <- innovation_plot_allphases %>%
  filter(income == '12 001 - 16 000 CHF') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_income4_phase4 <- innovation_plot_allphases %>%
  filter(income == '12 001 - 16 000 CHF') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_income4_phase5 <- innovation_plot_allphases %>%
  filter(income == '12 001 - 16 000 CHF') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_income4_phase6 <- innovation_plot_allphases %>%
  filter(income == '12 001 - 16 000 CHF') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_income4 <-
   bind_rows(innovation_plot_income4_phase1,innovation_plot_income4_phase2,
                                    innovation_plot_income4_phase3,
                                    innovation_plot_income4_phase4,
                                    innovation_plot_income4_phase5,
                                    innovation_plot_income4_phase6)

innovation_plot_income5_phase1 <- innovation_plot_allphases %>%
  filter(income == 'More than 16 000 CHF') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_income5_phase2 <- innovation_plot_allphases %>%
  filter(income == 'More than 16 000 CHF') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_income5_phase3 <- innovation_plot_allphases %>%
  filter(income == 'More than 16 000 CHF') %>%
  filter(phase == 3) %>%
```

```
  summarise(mean = mean(rate))

innovation_plot_income5_phase4 <- innovation_plot_allphases %>%
  filter(income == 'More than 16 000 CHF') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_income5_phase5 <- innovation_plot_allphases %>%
  filter(income == 'More than 16 000 CHF') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_income5_phase6 <- innovation_plot_allphases %>%
  filter(income == 'More than 16 000 CHF') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_income5 <-
   bind_rows(innovation_plot_income5_phase1,innovation_plot_income5_phase2,
                                    innovation_plot_income5_phase3,
                                    innovation_plot_income5_phase4,
                                    innovation_plot_income5_phase5,
                                    innovation_plot_income5_phase6)

innovation_plot_income6_phase1 <- innovation_plot_allphases %>%
  filter(income == 'Prefer not to say') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_income6_phase2 <- innovation_plot_allphases %>%
  filter(income == 'Prefer not to say') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_income6_phase3 <- innovation_plot_allphases %>%
  filter(income == 'Prefer not to say') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_income6_phase4 <- innovation_plot_allphases %>%
  filter(income == 'Prefer not to say') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_income6_phase5 <- innovation_plot_allphases %>%
  filter(income == 'Prefer not to say') %>%
```

```
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_income6_phase6 <- innovation_plot_allphases %>%
  filter(income == 'Prefer not to say') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_income6 <-
   bind_rows(innovation_plot_income6_phase1,innovation_plot_income6_phase2,
                                    innovation_plot_income6_phase3,
                                    innovation_plot_income6_phase4,
                                    innovation_plot_income6_phase5,
                                    innovation_plot_income6_phase6)

innovation_plot_income <- bind_cols(innovation_plot_income1, innovation_plot_income2,
                                  innovation_plot_income3, innovation_plot_income4,
                                innovation_plot_income5, innovation_plot_income6)
innovation_plot_income$phase <- c(1,2,3,4,5,6)

ggplot(data = innovation_plot_income) +
  geom_line(aes(x= phase,y=mean...1, color='1', linetype='1'), size=1) +
  geom_line(aes(x=phase,y=mean...2, color='2', linetype='2'), size=1) +
  geom_line(aes(x=phase,y=mean...3, color='3', linetype='3'), size=1) +
  geom_line(aes(x=phase,y=mean...4, color='4', linetype='4'), size=1) +
  geom_line(aes(x=phase,y=mean...5, color='5', linetype='5'), size=1) +
  geom_line(aes(x=phase,y=mean...6, color='6', linetype='6'), size=1) +
  labs(y='Innovation rate', x = 'Phase') +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5, 6))+
  scale_color_manual(values = c('#CCCCCC', '#999999', '#666666', '#333333',
    '#000000', '#666666'),
                     name = 'Income', labels = c('4 000 CHF or less','4 001 - 8 000
                       CHF', '8 001 - 12 000 CHF',
                                                 '12 001 - 16 000 CHF', 'More than 16
                                                   000 CHF', 'Prefer not to say'))+
  scale_linetype_manual(values=c(1,4,1,1,3,2), name = 'Income',
                        labels = c('4 000 CHF or less','4 001 - 8 000 CHF', '8 001 - 12
                          000 CHF',
                                   '12 001 - 16 000 CHF', 'More than 16 000 CHF',
                                     'Prefer not to say'))+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))
```

```
#innovation vs. household
innovation_participants_df_household <- innovation_participants_df %>%
  filter(!is.na(household_size))

innovation_participants_df_household_phase1 <- innovation_participants_df_phase1 %>%
  filter(!is.na(household_size))

innovation_participants_df_household_phase2 <- innovation_participants_df_phase2 %>%
  filter(!is.na(household_size))

innovation_participants_df_household_phase3 <- innovation_participants_df_phase3 %>%
  filter(!is.na(household_size))

innovation_participants_df_household_phase4 <- innovation_participants_df_phase4 %>%
  filter(!is.na(household_size))

innovation_participants_df_household_phase5 <- innovation_participants_df_phase5 %>%
  filter(!is.na(household_size))

innovation_participants_df_household_phase6 <- innovation_participants_df_phase6 %>%
  filter(!is.na(household_size))


boxplot(rate ~ household_size, innovation_participants_df_household,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = '', ylab = 'Innovation rate [new activities/day]',
        xlab = 'Household size', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~ household_size, innovation_participants_df_household_phase1,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = '', ylab = 'Innovation rate [new activities/day]',
        xlab = 'Household size', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~ household_size, innovation_participants_df_household_phase2,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = '', ylab = 'Innovation rate [new activities/day]',
        xlab = 'Household size', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~ household_size, innovation_participants_df_household_phase3,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = '', ylab = 'Innovation rate [new activities/day]',
        xlab = 'Household size', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~ household_size, innovation_participants_df_household_phase4,
        cex.main = 1, cex.axis =0.6, cex.lab = 1,
        main = '', ylab = 'Innovation rate [new activities/day]',
```

```
       xlab = 'Household size', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~ household_size, innovation_participants_df_household_phase5,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = '', ylab = 'Innovation rate [new activities/day]',
       xlab = 'Household size', ylim = c(-0.2, 2), las = 1)

boxplot(rate ~ household_size, innovation_participants_df_household_phase6,
       cex.main = 1, cex.axis =0.6, cex.lab = 1,
       main = '', ylab = 'Innovation rate [new activities/day]',
       xlab = 'Household size', ylim = c(-0.2, 2), las = 1)


mean_innovation_household <- innovation_participants_df_household %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household <- innovation_participants_df_household %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household <- innovation_participants_df_household %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household <- innovation_participants_df_household %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
innovation_household <- bind_cols(mean_innovation_household,
                                  median_innovation_household %>%
                                    select(-household_size),
                                  sd_innovation_household %>% select(-household_size),
                                  length_innovation_household %>%
                                    select(-household_size))

mean_innovation_household_phase1 <- innovation_participants_df_household_phase1 %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household_phase1 <- innovation_participants_df_household_phase1 %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household_phase1 <- innovation_participants_df_household_phase1 %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household_phase1 <- innovation_participants_df_household_phase1 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
```

```
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
innovation_household_phase1 <- bind_cols(mean_innovation_household_phase1,
                                median_innovation_household_phase1 %>%
                                  select(-household_size),
                                sd_innovation_household_phase1 %>%
                                  select(-household_size),
                                length_innovation_household_phase1 %>%
                                  select(-household_size))

mean_innovation_household_phase2 <- innovation_participants_df_household_phase2 %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household_phase2 <- innovation_participants_df_household_phase2 %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household_phase2 <- innovation_participants_df_household_phase2 %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household_phase2 <- innovation_participants_df_household_phase2 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
innovation_household_phase2 <- bind_cols(mean_innovation_household_phase2,
                                    median_innovation_household_phase2 %>%
                                      select(-household_size),
                                    sd_innovation_household_phase2 %>%
                                      select(-household_size),
                                    length_innovation_household_phase2 %>%
                                      select(-household_size))

mean_innovation_household_phase3 <- innovation_participants_df_household_phase3 %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household_phase3 <- innovation_participants_df_household_phase3 %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household_phase3 <- innovation_participants_df_household_phase3 %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household_phase3 <- innovation_participants_df_household_phase3 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
```

```
innovation_household_phase3 <- bind_cols(mean_innovation_household_phase3,
                                         median_innovation_household_phase3 %>%
                                           select(-household_size),
                                         sd_innovation_household_phase3 %>%
                                           select(-household_size),
                                         length_innovation_household_phase3 %>%
                                           select(-household_size))

mean_innovation_household_phase4 <- innovation_participants_df_household_phase4 %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household_phase4 <- innovation_participants_df_household_phase4 %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household_phase4 <- innovation_participants_df_household_phase4 %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household_phase4 <- innovation_participants_df_household_phase4 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
innovation_household_phase4 <- bind_cols(mean_innovation_household_phase4,
                                         median_innovation_household_phase4 %>%
                                           select(-household_size),
                                         sd_innovation_household_phase4 %>%
                                           select(-household_size),
                                         length_innovation_household_phase4 %>%
                                           select(-household_size))

mean_innovation_household_phase5 <- innovation_participants_df_household_phase5 %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household_phase5 <- innovation_participants_df_household_phase5 %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household_phase5 <- innovation_participants_df_household_phase5 %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household_phase5 <- innovation_participants_df_household_phase5 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
innovation_household_phase5 <- bind_cols(mean_innovation_household_phase5,
                                         median_innovation_household_phase5 %>%
```

```
                                    select(-household_size),
                                  sd_innovation_household_phase5 %>%
                                    select(-household_size),
                                  length_innovation_household_phase5 %>%
                                    select(-household_size))

mean_innovation_household_phase6 <- innovation_participants_df_household_phase6 %>%
  group_by(household_size) %>%
  summarise(mean_inno_household = mean(rate))
median_innovation_household_phase6 <- innovation_participants_df_household_phase6 %>%
  group_by(household_size) %>%
  summarise(median_inno_household = median(rate))
sd_innovation_household_phase6 <- innovation_participants_df_household_phase6 %>%
  group_by(household_size) %>%
  summarise(sd_inno_household = sd(rate))
length_innovation_household_phase6 <- innovation_participants_df_household_phase6 %>%
  group_by(household_size) %>%
  mutate(rownumberhousehold = row_number()) %>%
  group_by(household_size) %>%
  summarise(length_space_household = max(rownumberhousehold))
innovation_household_phase6 <- bind_cols(mean_innovation_household_phase6,
                                  median_innovation_household_phase6 %>%
                                    select(-household_size),
                                  sd_innovation_household_phase6 %>%
                                    select(-household_size),
                                  length_innovation_household_phase6 %>%
                                    select(-household_size))


innovation_plot_household_size1_phase1 <- innovation_plot_allphases %>%
  filter(household_size == '1') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size1_phase2 <- innovation_plot_allphases %>%
  filter(household_size == '1') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size1_phase3 <- innovation_plot_allphases %>%
  filter(household_size == '1') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size1_phase4 <- innovation_plot_allphases %>%
  filter(household_size == '1') %>%
```

```
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size1_phase5 <- innovation_plot_allphases %>%
  filter(household_size == '1') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size1_phase6 <- innovation_plot_allphases %>%
  filter(household_size == '1') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size1 <-
   bind_rows(innovation_plot_household_size1_phase1,innovation_plot_household_size1_phase2,
                                innovation_plot_household_size1_phase3,
                                innovation_plot_household_size1_phase4,
                                innovation_plot_household_size1_phase5,
                                innovation_plot_household_size1_phase6)

innovation_plot_household_size2_phase1 <- innovation_plot_allphases %>%
  filter(household_size == '2') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size2_phase2 <- innovation_plot_allphases %>%
  filter(household_size == '2') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size2_phase3 <- innovation_plot_allphases %>%
  filter(household_size == '2') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size2_phase4 <- innovation_plot_allphases %>%
  filter(household_size == '2') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size2_phase5 <- innovation_plot_allphases %>%
  filter(household_size == '2') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size2_phase6 <- innovation_plot_allphases %>%
```

```
  filter(household_size == '2') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size2 <-
   bind_rows(innovation_plot_household_size2_phase1,innovation_plot_household_size2_phase2,
                                innovation_plot_household_size2_phase3,
                                innovation_plot_household_size2_phase4,
                                innovation_plot_household_size2_phase5,
                                innovation_plot_household_size2_phase6)

innovation_plot_household_size3_phase1 <- innovation_plot_allphases %>%
  filter(household_size == '3') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size3_phase2 <- innovation_plot_allphases %>%
  filter(household_size == '3') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size3_phase3 <- innovation_plot_allphases %>%
  filter(household_size == '3') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size3_phase4 <- innovation_plot_allphases %>%
  filter(household_size == '3') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size3_phase5 <- innovation_plot_allphases %>%
  filter(household_size == '3') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size3_phase6 <- innovation_plot_allphases %>%
  filter(household_size == '3') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size3 <-
   bind_rows(innovation_plot_household_size3_phase1,innovation_plot_household_size3_phase2,
                                innovation_plot_household_size3_phase3,
                                innovation_plot_household_size3_phase4,
                                innovation_plot_household_size3_phase5,
```

```
                                innovation_plot_household_size3_phase6)

innovation_plot_household_size4_phase1 <- innovation_plot_allphases %>%
  filter(household_size == '4') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size4_phase2 <- innovation_plot_allphases %>%
  filter(household_size == '4') %>%
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size4_phase3 <- innovation_plot_allphases %>%
  filter(household_size == '4') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size4_phase4 <- innovation_plot_allphases %>%
  filter(household_size == '4') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size4_phase5 <- innovation_plot_allphases %>%
  filter(household_size == '4') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size4_phase6 <- innovation_plot_allphases %>%
  filter(household_size == '4') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size4 <-
   bind_rows(innovation_plot_household_size4_phase1,innovation_plot_household_size4_phase2,
                                innovation_plot_household_size4_phase3,
                                innovation_plot_household_size4_phase4,
                                innovation_plot_household_size4_phase5,
                                innovation_plot_household_size4_phase6)

innovation_plot_household_size5_phase1 <- innovation_plot_allphases %>%
  filter(household_size == '5 or more') %>%
  filter(phase == 1) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size5_phase2 <- innovation_plot_allphases %>%
  filter(household_size == '5 or more') %>%
```

```
  filter(phase == 2) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size5_phase3 <- innovation_plot_allphases %>%
  filter(household_size == '5 or more') %>%
  filter(phase == 3) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size5_phase4 <- innovation_plot_allphases %>%
  filter(household_size == '5 or more') %>%
  filter(phase == 4) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size5_phase5 <- innovation_plot_allphases %>%
  filter(household_size == '5 or more') %>%
  filter(phase == 5) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size5_phase6 <- innovation_plot_allphases %>%
  filter(household_size == '5 or more') %>%
  filter(phase == 6) %>%
  summarise(mean = mean(rate))

innovation_plot_household_size5 <-
   bind_rows(innovation_plot_household_size5_phase1,innovation_plot_household_size5_phase2,
                                  innovation_plot_household_size5_phase3,
                                  innovation_plot_household_size5_phase4,
                                  innovation_plot_household_size5_phase5,
                                  innovation_plot_household_size5_phase6)


innovation_plot_household_size <- bind_cols(innovation_plot_household_size1,
   innovation_plot_household_size2,
                                 innovation_plot_household_size3,
                                   innovation_plot_household_size4,
                                 innovation_plot_household_size5)
innovation_plot_household_size$phase <- c(1,2,3,4,5,6)

ggplot(data = innovation_plot_household_size) +
  geom_line(aes(x= phase,y=mean...1, color='1', linetype='1'), size=1) +
  geom_line(aes(x=phase,y=mean...2, color='2', linetype='2'), size=1) +
  geom_line(aes(x=phase,y=mean...3, color='3', linetype='3'), size=1) +
  geom_line(aes(x=phase,y=mean...4, color='4', linetype='4'), size=1) +
  geom_line(aes(x=phase,y=mean...5, color='5', linetype='5'), size=1) +
  labs(y='Innovation rate', x = 'Phase') +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5, 6))+
```

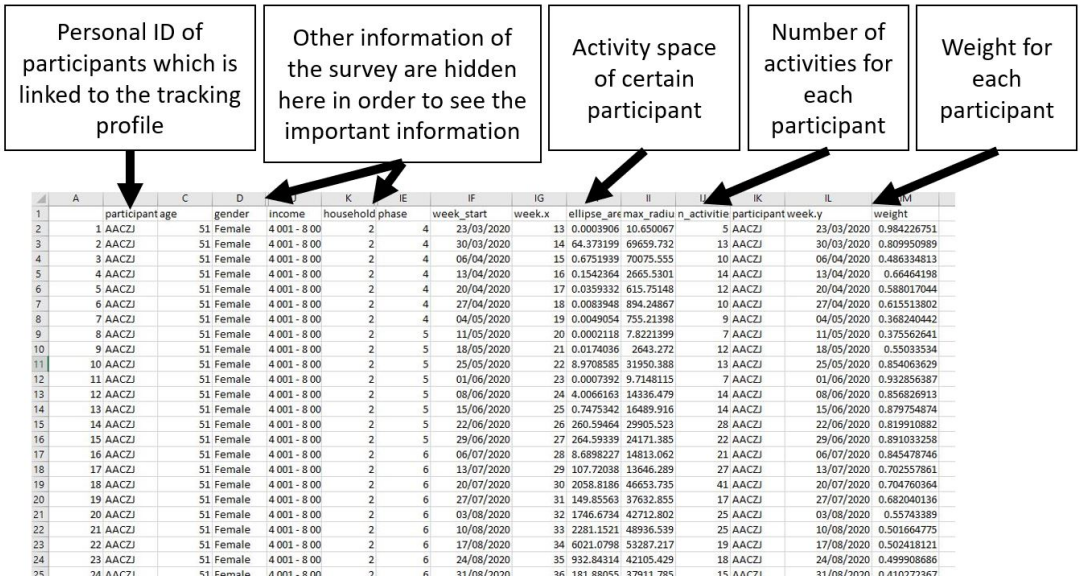
```
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_color_manual(values = c('#CCCCCC', '#999999', '#666666', '#333333',
    '#000000', '#666666'),
                     name = 'Household size', labels = c('1','2', '3',
                                                         '4', '5 or more'))+
  scale_linetype_manual(values=c(1,1,2,3,1,4), name = 'Household size',
                        labels = c('1','2', '3',
                                   '4', '5 or more'))+
  theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size=20),
        axis.text.x = element_text(size=16), axis.text.y = element_text(size=16),
        legend.title = element_text(size=20), legend.text = element_text(size=16))

#participants for innovation, plots
teilnehmer_innovation <- participants %>%
  filter(participant_ID %in% unique(num_teilnehmer_inno$user_id))


#age
hist(teilnehmer_innovation$age, breaks = 25, col = 'steelblue',
     cex.main = 1, cex.axis =0.6, cex.lab = 1,
     main = ' ',
     xlab = 'Age of the participants', ylab = 'Number of participants',
     las = 1)


#household size
householdsize_innovation <- dplyr::count(teilnehmer_innovation, household_size)[2]
householdsize_innovation_plot <- matrix(nrow=1,ncol=5)
householdsize_innovation_plot[1] <- householdsize_innovation[1,1]
householdsize_innovation_plot[2] <- householdsize_innovation[2,1]
householdsize_innovation_plot[3] <- householdsize_innovation[3,1]
householdsize_innovation_plot[4] <- householdsize_innovation[4,1]
householdsize_innovation_plot[5] <- householdsize_innovation[5,1]
householdsize_innovation_plot <- as.numeric(householdsize_innovation_plot)
barplot(householdsize_innovation_plot, names = c('1','2','3','4','5 or more'), las=1,
   col='steelblue', xlab='Household size',
        ylab='Number of participants', ylim=c(0,600))


#income
income_innovation <- dplyr::count(teilnehmer_innovation, income)[2]
income_innovation_plot <- matrix(nrow=1, ncol=6)
income_innovation_plot[1] <- income_innovation[2,1]
income_innovation_plot[2] <- income_innovation[3,1]
income_innovation_plot[3] <- income_innovation[4,1]
income_innovation_plot[4] <- income_innovation[1,1]
```

```
income_innovation_plot[5] <- income_innovation[5,1]
income_innovation_plot[6] <- income_innovation[6,1]
income_innovation_plot <- as.numeric(income_sorted)
barplot(income_innovation_plot, names =
  c('<4000','4001-8000','8001-12000','12001-1600','>16001', 'Prefer not to say'),
       col = 'steelblue', cex.names = 0.7, las=1, xlab = 'Income per month [CHF]',
         ylab = 'Number of participants',
       ylim=c(0,500))



# Save your results
## Output in a csv file
innovation_rate_results <- bind_cols(innovation_participants_df %>%select(-user_id),
  obs_time %>% select(-user_id))
write.csv(innovation_rate_results, 'innovation_rate_results.csv')

innovation_rate_results_phase1 <- bind_cols(innovation_participants_df_phase1
  %>%select(-user_id), obs_time_phase1 %>% select(-user_id))
write.csv(innovation_rate_results_phase1, 'innovation_rate_results_phase1.csv')

innovation_rate_results_phase2 <- bind_cols(innovation_participants_df_phase2
  %>%select(-user_id), obs_time_phase2 %>% select(-user_id))
write.csv(innovation_rate_results_phase2, 'innovation_rate_results_phase2.csv')

innovation_rate_results_phase3 <- bind_cols(innovation_participants_df_phase3
  %>%select(-user_id), obs_time_phase3 %>% select(-user_id))
write.csv(innovation_rate_results_phase3, 'innovation_rate_results_phase3.csv')

innovation_rate_results_phase4 <- bind_cols(innovation_participants_df_phase4
  %>%select(-user_id), obs_time_phase4 %>% select(-user_id))
write.csv(innovation_rate_results_phase4, 'innovation_rate_results_phase4.csv')

innovation_rate_results_phase5 <- bind_cols(innovation_participants_df_phase5
  %>%select(-user_id), obs_time_phase5 %>% select(-user_id))
write.csv(innovation_rate_results_phase5, 'innovation_rate_results_phase5.csv')

innovation_rate_results_phase6 <- bind_cols(innovation_participants_df_phase6
  %>%select(-user_id), obs_time_phase6 %>% select(-user_id))
write.csv(innovation_rate_results_phase6, 'innovation_rate_results_phase6.csv')



#Analysis of the participants for the discussion of the project
teilnehmer_innovation <- innovation_participants_df %>%
  filter(user_id %in% unique(teilnehmer_innovation$participant_ID))
innovation_participants_students <- teilnehmer_innovation %>%
```

```
  filter(main_employment == 'Student')
innovation_participants_young <- teilnehmer_innovation %>%
  filter(AgeGroup == '(10,20]')
innovation_participants_students_young <- teilnehmer_innovation %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(10,20]')
innovation_participants_students_young_lowincome <- teilnehmer_innovation %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(income == '4 000 CHF or less')
innovation_participants_students_young_notsaying <- teilnehmer_innovation %>%
  filter(main_employment == 'Student') %>%
  filter(AgeGroup == '(10,20]') %>%
  filter(income == 'Prefer not to say')
innovation_participants_5 <- teilnehmer_innovation %>%
  filter(household_size == '5 or more')
innovation_participants_5_notsaying <- teilnehmer_innovation %>%
  filter(household_size == '5 or more') %>%
  filter(income == 'Prefer not to say')
innovation_participants_notsaying <- teilnehmer_innovation %>%
  filter(income == 'Prefer not to say')
innovation_participants_notsaying_5 <- teilnehmer_innovation %>%
  filter(income == 'Prefer not to say') %>%
  filter(household_size == '5 or more')
```

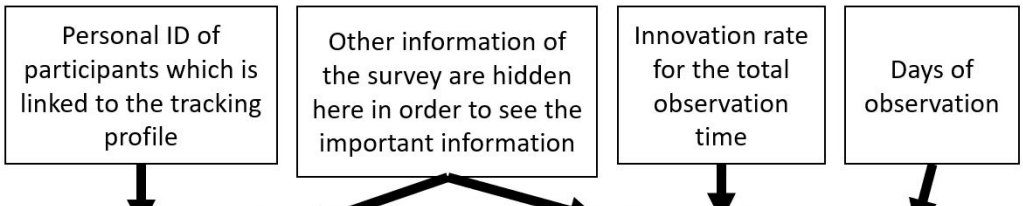
B Documentation of the results

B.1 Documentation of the results of the activity spaces



	participant	age	gender	income	household	phase	week_start	week.x	ellipse_are	max_radiu	n_activitie	participant	week.y	weight
1	AACZJ	51	Female	4 001 - 8 00	2	4	23/03/2020	13	0.0003906	10.650067	5	AACZJ	23/03/2020	0.984226751
2	AACZJ	51	Female	4 001 - 8 00	2	4	30/03/2020	14	64.373199	69659.732	13	AACZJ	30/03/2020	0.809950989
3	AACZJ	51	Female	4 001 - 8 00	2	4	06/04/2020	15	0.6751939	70075.555	10	AACZJ	06/04/2020	0.486334813
4	AACZJ	51	Female	4 001 - 8 00	2	4	13/04/2020	16	0.1542364	2665.5301	14	AACZJ	13/04/2020	0.66464198
5	AACZJ	51	Female	4 001 - 8 00	2	4	20/04/2020	17	0.0359332	615.75148	12	AACZJ	20/04/2020	0.588017044
6	AACZJ	51	Female	4 001 - 8 00	2	4	27/04/2020	18	0.0083948	894.24867	10	AACZJ	27/04/2020	0.615513802
7	AACZJ	51	Female	4 001 - 8 00	2	4	04/05/2020	19	0.0049054	755.21398	9	AACZJ	04/05/2020	0.368240442
8	AACZJ	51	Female	4 001 - 8 00	2	5	11/05/2020	20	0.0002118	7.8221399	7	AACZJ	11/05/2020	0.375562641
9	AACZJ	51	Female	4 001 - 8 00	2	5	18/05/2020	21	0.0174036	2643.272	12	AACZJ	18/05/2020	0.55033534
10	AACZJ	51	Female	4 001 - 8 00	2	5	25/05/2020	22	8.9708585	31950.388	13	AACZJ	25/05/2020	0.854063629
11	AACZJ	51	Female	4 001 - 8 00	2	5	01/06/2020	23	0.0007392	9.7148115	7	AACZJ	01/06/2020	0.932856387
12	AACZJ	51	Female	4 001 - 8 00	2	5	08/06/2020	24	4.0066163	14336.479	14	AACZJ	08/06/2020	0.856826913
13	AACZJ	51	Female	4 001 - 8 00	2	5	15/06/2020	25	0.7475342	16489.916	14	AACZJ	15/06/2020	0.879754874
14	AACZJ	51	Female	4 001 - 8 00	2	5	22/06/2020	26	260.59464	29905.523	28	AACZJ	22/06/2020	0.819910882
15	AACZJ	51	Female	4 001 - 8 00	2	5	29/06/2020	27	264.59339	24171.385	22	AACZJ	29/06/2020	0.891033258
16	AACZJ	51	Female	4 001 - 8 00	2	6	06/07/2020	28	8.6898227	14813.062	21	AACZJ	06/07/2020	0.845478746
17	AACZJ	51	Female	4 001 - 8 00	2	6	13/07/2020	29	107.72038	13646.289	27	AACZJ	13/07/2020	0.702557861
18	AACZJ	51	Female	4 001 - 8 00	2	6	20/07/2020	30	2058.8186	46653.735	41	AACZJ	20/07/2020	0.704760364
19	AACZJ	51	Female	4 001 - 8 00	2	6	27/07/2020	31	149.85563	37632.855	17	AACZJ	27/07/2020	0.682040136
20	AACZJ	51	Female	4 001 - 8 00	2	6	03/08/2020	32	1746.6734	42712.802	25	AACZJ	03/08/2020	0.55743389
21	AACZJ	51	Female	4 001 - 8 00	2	6	10/08/2020	33	2281.1521	48936.539	25	AACZJ	10/08/2020	0.501664775
22	AACZJ	51	Female	4 001 - 8 00	2	6	17/08/2020	34	6021.0798	53287.217	19	AACZJ	17/08/2020	0.502418121
23	AACZJ	51	Female	4 001 - 8 00	2	6	24/08/2020	35	932.84314	42105.429	18	AACZJ	24/08/2020	0.499908686
24	AACZJ	51	Female	4 001 - 8 00	2	6	31/08/2020	36	181.88055	37911.785	15	AACZJ	31/08/2020	0.410272367

B.2 Documentation of the results of the innovation rates



	participant_ID	age	income	household_size	AgeGroup	rate	num_obs_day
1	AACZJ	51	4 001 - 8 000 CHF	2	(50,60]	0.436170213	188
2	AARIH	53	12 001 - 16 000 CHF	2	(50,60]	0.285714286	28
3	ABASX	28	8 001 - 12 000 CHF	2	(20,30]	0.520833333	48
4	ACDFH	35	4 001 - 8 000 CHF	5 or more	(30,40]	1.368852459	122
5	ACKBM	48	More than 16 000 CHF	5 or more	(40,50]	0.876712329	146
6	ADARK	53	4 001 - 8 000 CHF	3	(50,60]	0.892857143	28
7	ADMTW	37	4 001 - 8 000 CHF	2	(30,40]	0.2	10
8	AEMRJ	20	Prefer not to say	5 or more	(10,20]	0.410852713	129
9	AETCS	42	8 001 - 12 000 CHF	1	(40,50]	0.946745562	169
10	AEWVK	30	4 001 - 8 000 CHF	1	(20,30]	0.525252525	198
11	AFEBJ	48	4 001 - 8 000 CHF	1	(40,50]	0.622857143	175
12	AGHXU	63	8 001 - 12 000 CHF	2	(60,Inf]	0.633986928	153
13	AGIIY	46	4 001 - 8 000 CHF	1	(40,50]	0.957446809	188
14	AGYOB	51	More than 16 000 CHF	5 or more	(50,60]	0.796407186	167
15	AHLNU	57	8 001 - 12 000 CHF	4	(50,60]	0.296296296	54
16	AHVGL	65	12 001 - 16 000 CHF	2	(60,Inf]	0.585365854	123
17	AHXQF	41	4 001 - 8 000 CHF	4	(40,50]	0.775280899	89
18	AIDVM	24	12 001 - 16 000 CHF	4	(20,30]	0.81875	160
19	AJDYU	63	12 001 - 16 000 CHF	2	(60,Inf]	0.821428571	28
20	AJSSC	41	8 001 - 12 000 CHF	1	(40,50]	0.885245902	61
21	AKKPT	48	More than 16 000 CHF	2	(40,50]	0.590909091	88