Patterns in human commuting behavior

Lars Rouvoet
VU University.
Amsterdam, The Netherlands
Lars.rouvoet@vu.nl

ABSTRACT

During daily life people behave differently depending on the situation they are in. Behavior of humans during a journey contains several patterns. Some of these patterns are related to their transport choice and their physical activity during the journey. Determining and studying these patterns for young adults resulted in advice which could be used during the development process of a lifestyle coaching application. Such applications aim at stimulating people to be more physical active. It is recommendable to increase the awareness of people about the diversity of transport options they could take. Their awareness is in most cases not optimal which therefore hold people back to choose for more active transport.

The coaching should be focused on short journeys done by car. These journeys do cover a distance which is currently already been covered by an average bike journey. The same occurs with most metro and bus journeys. People could be relatively effectively stimulated to take active travel instead due to the fact that the distance is not significantly different in these journeys. People take 35% less steps during days with relatively bad weather. These days have far more potential with respect to physical activity. Communicate with a weather API and help the user to predict if the weather is really an obstacle for using active transport.

1. INTRODUCTION

The number of kilometers travelled by residents of the Netherlands increased by 4 percent between 2000 and 2011 and therefore also car usage increases every year [1]. Also TNO stated about Dutch inhabitants that 49.5 % of the youth (4-17 years) and 30.3 % of the adults do not meet the physical activity norm [2], which influences human health negatively. This creates opportunities for researchers by combining the two domains: commuting behavior and human health.

Those two domains are studied by Active2Gether, a project which focuses on physical activity behavior of young adults. The goal of this project is to motivate young people to be more physically active while traveling and on specific locations. Active2Gether tries to achieve this goal by building a life-style coaching application which stimulates users via textual messages to be more physical active.

In order to motivate these people in an effective way, research concerning their activity behavior is required. This research can roughly be dived in two parts, namely a psychology part and a more technical part. This thesis focuses on the technical part of this research, which aims on whether we can use data from a mobile activity monitor and logs of GPS locations to get a better understanding about commuting behavior of young adults. More specifically, activity and position data is gathered and studied for commuting patterns.

A frequently common combination of specific factor values is called a pattern, see figure 1. The commuting behavior of a user can contain different patterns. These patterns consist of different daily life factors like time, weather or location.

Figure 1. Destination versus health active transport

The found behavior patterns will be used to create advice for developing a lifestyle coaching application. This advice will point out specific behavior of people with potential to improve their physical activity level. These improvement points will be the base for textual advice, sent to the user via a developed application.

The effectiveness of the application can be increased by making it more intelligent. By using intelligent methods, motivation techniques can be better adapted to the user. Finding user activity patterns are for example a way to do this. Patterns can be used to better understand and predict the user’s behavior and therefore be an important element of the advice given by the application. This thesis describes a study in which a research method of finding user patterns as well as the results of this method.

1.2 RELATED WORK

Research in the field of commuting behavior is also done by Oliver et al. (2010) In their paper GPS, GIS (Geographic Information System) and accelerometry data are combined to study travel behavior of adults. [3] Oliver et al.
mentioned the opportunity of combining commuting behavior with physical activity as follows: “TPA [Transport related Physical Activity] presents the opportunity to engage in physical activity by integrating regular, habitual activity into the daily routines of individuals, thereby overcoming the time constraints reported by many as a major barrier to physical activity”. About the size of this opportunity, Oliver et al. stated that it depends on built environment infrastructure, such as well-connected and maintained footpaths, cycle lanes and street networks. These appear to be important facilitators of TPA engagement, and are recognized as key correlates in combination with time and distance factors. This points out that the environment influence the commuting behavior of people. During a user study Oliver et al. measured physical activity with an accelerometer and GPS data. Assessing physical activity patterns using GPS units led to a significant amount of issues. Therefore it is recommendable to use (besides GPS data) other methods to measure physical activity. After the user study, Oliver et al. concluded that “relatively short distances found for motorized trips have identified the opportunity for intervention and promoting TPA for short commute trips”. This opportunity will be further discussed in this paper.

How do people make decisions?
Human commuting behavior can be seen as a set of decisions. Cindy Dietrich (2010) stated that human decisions are based on certain factors, these are past experience, cognitive biases, age and individual differences, belief in personal relevance and an escalation of commitment. [4] Following Cindy Dietrich, understanding these factors that influence decision making process is import to understanding what decision are made. The factors that influence the process may impact the outcomes. Commuting patterns can be seen as an iterative set of decisions related to a certain commuting context. Influencing a decision via lifestyle coaching can be done by influencing decision factors. Understanding which patterns can effectively be changed and which not, relates with these decision factors. Understanding which patterns, and therefore which factors, effectively can be influenced by lifestyle coaching, is important for the development of the application.

1.3 OBJECTIVE
Commuting patterns could be used to find out on which point the user could improve and thus create personal and relevant advice for the user. The coaching advice can be adjusted to particular patterns and possibly aim at breaking through them. Finding relevant patterns is the main goal of this research and therefore the following main question will be answered:

Which patterns in human activity behavior can be derived from activity- and GPS data which can be used to create personal advice for people who want to be more physically active?

In the next sections the methodology of the research will be discussed, followed by the sections discussing the results and outcome of the research.

2. METHODOLOGY
A user study is done to create a better understanding of people’s activity behavior and to investigate if any patterns can be found. After the data is collected with the user study, it is preprocessed and analyzed.

2.1 USER STUDY
During the user study, data about physical activity behavior of participants is collected via a mobile device, called “Fitbit One”. It tracks the user’s steps, distance, calories burned and stairs climbed during the day. Besides the Fitbit, the location of the user was measured by a “GPS-Logger” application for Android phones. This application logs the GPS coordinates of the user’s mobile phone at specified intervals. This combination creates an opportunity to gain an insight in people’s activity behavior on specific locations and during journeys, which is necessary to answer the research question. It also limits the amount of issues, stated by Oliver et al., which occurs when measuring physical activity only by GPS. The participants of the experiment wore the Fitbit device for two weeks.

Following the target group of the Active2Gether project, the participants of the experiment were young people in the age range of 18 – 26 years. 24 People, living almost all in and nearby Amsterdam, participated in the experiment. The experiment lasted between the end of January until mid-March. More detailed information about the participants can be found in appendix A.

Questionnaires
An intake questionnaire for static information was held at the start of the experiment. It contained questions about general information of the participants, like age, sex, height and weight. It also contained questions about important places of the person, such as home, study, sport and work. The participants also answered a daily questionnaire during the experiment. Daily questionnaires contained questions about on what time the Fitbit was worn and taken off and if the participant has traveled that day. Information about departure location, destination, time and type of transport were asked per journey.

Both questionnaires gave more insight in the (daily) context of the user. The data from the questionnaires had a few purposes. One of the purposes was to validate if certain GPS coordinates from the GPS logger belongs to an important place of a particular person. An important place is a location which is visited more than average by a certain participant. Questionnaires could also validate journeys between these important places. Studying journeys
between important places were necessary to find potential commuting patterns. Due to the daily questionnaire data it was also possible to match the Fitbit data with related journeys.

2.2 PREPROCESSING THE DATA
After the experiment was finished, the activity data, GPS data and questionnaires were preprocessed before being analyzed for potential patterns. Before preprocessing the data, it had to be clear which factors were relevant for finding activity behavior patterns, because these factors influences the representation of the data and the data filtering process. For this study, commuting behavior is narrowed to type of transport taken, distance traveled and steps taken. Also weather data is compared against these factors. After selecting the data needed it was filtered, structured and analyzed with statistical computing software. During the preprocessing phase, Python [5] was used to gather the activity data using the REST API from Fitbit [6] and R [7] for statistical analyses and visualizations.

3. PATTERNS IN HUMAN COMMUTING BEHAVIOR
After preprocessing the data, it was ready to be analyzed on different commuting patterns. Potential patterns for this study were divided into three subcategories, related to the factors mentioned earlier in this paper. Research based on those factors led to the following sub questions:

- *Is there a relation between a particular place and the type of transport people prefer while travelling to that place?*
- *Is there a relation between the weather and commuting behavior?*
- *Is there a relation between distance traveled and the transport type taken?*

These three sub questions are discussed and answered in the upcoming sections.

3.1 RELATION BETWEEN A PARTICULAR PLACE AND THE TYPE OF TRANSPORT USED TO TRAVEL TO THAT PLACE
People in general prefer structure and consistency and therefore dislike behavior change. Therefore, it is plausible to think that people travel most of the time on the same way to a particular location. To test this hypothesis, the data is divided in three subcategories namely: journeys between home - sport, home - work and home – study. This is done to narrow the study and to focus on most common journeys for young adults. Only data from the questionnaires is used, due to the fact that this data contained information about which transport type was taken during a journey.

The routes are analyzed and studied on transport usage. A bar chart is made per destination which shows in how many cases which transport is taken. The results are shown in appendix B. Notice that a participant could take one or more different transport types while traveling to a certain location.

The choice for a specific type of transport varies heavily per destination, e.g. train is in 69% of the time used to travel to study while to work it is only 24%. The bar chart about transport to study contains two high peaks for bike and train transport, see figure 2.

![Figure 2. Transport to study](image)

Those two peaks represent the two subcategories in our sample group, namely people living in Amsterdam and outside Amsterdam. People outside Amsterdam preferred mostly the train as transport to study while people in Amsterdam mostly took the bike. The bar charts about work and sports in general contained relatively high diversity among the chosen transport type, which made it hard to conclude significant statements about a single transport type.

Transport types can also be combined and divided into two categories, active transport and non-active transport. Active transport means transport whereby people need to be physical active to move themselves, such as biking and walking. Non-active transport is contrary to active transport, people don’t need to be active in order to move. Public transport and traveling by car are examples of non-active transport.

A healthy score is calculated based on travels which are only done with active transport. The score represents how frequently people traveled actively (per foot/bike) to a certain location. This is only based on questionnaire data and not on activity data of the Fitbit device. Table 1 shows the healthy score and the average steps during a journey.

<table>
<thead>
<tr>
<th>Healthy Score</th>
<th>Study (%)</th>
<th>Work (%)</th>
<th>Sport (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average steps</td>
<td>1277</td>
<td>774</td>
<td>476</td>
</tr>
</tbody>
</table>

*Table 1. Destination versus active transport*

Table 1 shows a relatively low healthy score when people are traveling to work. Also notice that participants used about triple as much active transport to sport locations
compared to work. Keep in mind that the healthy score logically is related to the distance to a location. People tend to use active transport less frequently when they have to travel long distances. Due to the fact that the average distance to a study location is relatively far compared to sport locations, and therefore limiting the choice for active transport, could largely clarify the results. The average distance of study/work and sport is presented in figure 3.

Variation in transport
The participants had to fill in the different travel options to their important places during the intake questionnaire. Figure 4 shows the different transport choices people think they had when traveling to study, the results for work and sports can be found in appendix C.

The results show that the participants in general only filled in one way of travel to a particular destination. In other words, the participants mostly traveled the same way every time they traveled to a particular location. Participants were also possibly not fully aware of the different type of transport they could take to travel. This resulted in a small variation between different travel routes and transport to a particular location per participant.

3.2 Relation between weather and commuting behavior
The Pearson correlation method [8] is used to test the potential relationship between the weather and type of transport during a journey. The type of transport is tested for correlation with the following weather factors: rain, average temperature, average wind speed and a weather score. The weather score is for the greater part calculated according official guidelines [9]. According to these guidelines, the weather score is calculated on a scale of 1 to 10 and depends on cloud cover, fog and mist, precipitation and wind. Temperature has no influence on the weather score, due to the limited role it plays in experiencing weather quality. Some adjustments of the score calculation were made during this study to compensate for the limitations of the available weather data, downloaded from Weather Underground [10]. The downloaded weather data did not contain the exact amount of time fog and mist occurred but only if it occurred or not. Also the data did not state how long it has rained per day, but only the amount of millimeters. The weather score during the period of the experiment is presented in figure 5.

Figure 3. Average distance to important places

From figure 3 it can be concluded that the amount of steps walked while traveling (table 1) increases when the destination is more far away. It also indicates that work has a lower healthy score but is in average less far away than a study location. This can be partly due to that none of the participants ever took the car to study compared to 24% for work. If students don’t live at their parental home, they mostly live near their study location instead of near their (part time) work. This potentially resulted in more active transport and explains the difference in healthy score.

Figure 4. Transport routes to study

The results show a variation in weather score between a 5 and a 10, while the score is about 8 on average. The correlation between the weather score and single weather factors, like rain, temperature and wind speed is tested. According to these results, see appendix D, there are no strong linear relationships between the type of transport chosen and single weather factors on that particular day. This can be explained by the small variation of the transport type to a particular location of the participant.
They did not vary the way of transporting significantly during the user study.

Correlation weather score and amount of steps
Figure 6 shows a histogram which represents the variation in the amount of steps taken on a day. The participants walked on average 5462 steps per day. The box plot of this data can be found in appendix E.

![Figure 6. Histogram: amount of steps per day](image)

Unlike the lack of correlation between single weather factors and chosen transport type, the correlation test between the weather score and the average steps taken on that day resulted in a weak linear relationship of 0.3. Figure 7 shows the average amount of steps taken on a day and a graph line representing the relative weather score. Keep in mind that not every participant participated at the same time at the experiment. The last 10 days of figure 7 are less relevant due significant decrease in participants.

![Figure 7. Weather versus average steps](image)

It is hard to identify any relationship between the weather score and the amount of steps from figure 7. To get a better insight in the relation between steps taken and weather, each day was categorized in “good” or “bad”, representing good weather (weather score > 7) and bad weather. The average steps were calculated for each of those two day types. The result showed an increase of 1500 steps on a good day in comparison with bad days, see figure 8. A figure containing box plots for three weather subcategories (bad/normal/good) can be found in appendix F.

![Figure 8. Weather score versus average steps](image)

1500 Steps is a significant difference compared to the average amount of steps. It is an increase in steps of 35% compared to days with bad weather.

People do not prefer to be outside due to (precaution of) bad weather. If people are more aware about the weather forecast, the amount of steps taken on bad weather days could probably be increased. The activity level on days with bad weather has a lot of potential to be increased.

3.3 RELATION BETWEEN DISTANCE AND TRANSPORT TYPE

When people travel they can choose between active and non-active transport. Logically, when people have to travel relatively far, they prefer non-active transport because it takes less effort and time.

During the preprocessing phase, the distance traveled is calculated per transport type. Those results are presented in figure 9, 10 and 11. Each participant took in most cases the same route and transport to a particular location. The calculations discussed in this section are based on only unique journeys of each participant. This is done to prevent the results to be influenced by participants who travel more than average to a certain location. Figure 9 shows two box plots representing the distance traveled with non-active and active transport.
The box plot regarding the active transport shows the distance traveled lies between 0 and 5 kilometers. While the choice for non-active transport is spread between 0 and 39 kilometers. More interesting is the clear transition between active and non-active transport. The 1st quartile of the non-active box plot is about the same value as the maximum values of the active transport. Following figure 9, the threshold when people start taking non-active transport lies around 4 kilometers.

The data points in the non-active box plot are presented in two groups. One is more or less homogeneously spread between the median and the minimum value. The other group is widely spread in the top of the box plot. Those two groups are representing the different transport types of non-active transport. People prefer bus, metro and tram for relative small distances and train if they have to travel far. This is noticeable in figure 11, which shows box plots for every transport type representing the distance traveled.

The car box (figure 11) plot shows that the data points of the lowest 50% are relatively close to each other, while the top 50% are widely spread. This states that 50% of the car travels are relatively short journeys (2-7 km), while the other 50% is spread between 7 and 37 kilometer. Interesting to notice is that the 1st quartile of the car box plot is about the same value as the 3rd quartile of the bike box plot. This means that 25% of the car journeys cover the same distance as 75% of the bike journeys. People who use their car to travel such small distances could potentially effectively be stimulated to take active transport instead.

Figure 10 shows the enlarged active transport box plot. The lowest 50% of the data point are close positioned between 0.5 and 1.3 kilometer. This represents mostly the journeys done per foot. The bike is mostly taken to travel a larger distance between 1.3 and 4 kilometer. The bike and walk box plot are also represented in figure 11.

Notice that tram is excluded from figure 11. This was done because the tram data was based on only five travels in total and was as transport type far less popular than the others.

The median of the bike-, bus- and metro box plot are relatively close to each other, which indicate that distance is not a (single) decisive factor when choosing one of those transport types. When only focusing on the distance of a journey, it can be concluded (following figure 11) that the average journey with the metro and bus could be replaced by bike transport. Distance is not a significant obstacle for choosing bike transport instead.
4. DISCUSSION AND CONCLUSION
A lot of data is collected and analyzed during the research. The preprocessing phase resulted in structured and relevant data, ready for statistical analyses. Some results of these analyses, such as the lack of correlation between transport type and weather factors, were not expected beforehand. This was due to the small variation in chosen transport type.

Some commuting patterns found in the data can be potentially effectively changed to a more positive (active wise) direction. This is because distance (between two locations) was not always the decisive factor for choosing non-active transport. A commuting decision only based on the distance between two locations is hardly influenceable by extern sources (e.g. a mobile application). Therefore, a decision about how to travel will be more influenceable when distance is not a (significant) decision factor.

4.1 ADVICE
The analyses after the preprocessing phase yielded an advice which could be used during the development phase of a lifestyle coaching application. Following this advice will increase the effectiveness of lifestyle coaching of young adults and therefore pursue the goal of Active2Gether.

Increase awareness
Make people more aware about the diversity of travel options. People are not fully aware of the diversity in transport options they could take. The bar charts in appendix C shows the transport options people knew they could take to a certain location. People can’t choose between different ways of transporting themselves if they are not aware about the option diversity.

Focus on days with bad weather
People walk on average 35% less steps per day with bad weather compared to days with good weather, which is about 1500 steps. Focus the coaching advice on days with bad weather, which have relatively more potential to increase the activity. Obviously, people don’t prefer to use active transport when it is raining or during extreme weather. Helping people to predict if it is really going to rain (or not) while traveling, could stimulate active transport. With respect to the implementation of such feature, it is advisable to communicate with an API of a weather station.

Focus on short car travels
From the box plot about the different transport types in paragraph 3.3 (figure 11) can be concluded that the 1th quartile of the car box plot is about the same value as the 3th quartile of the bike box plot. This indicates that 25% of the car journey covers the same distance as 75% of the bike journeys. It is highly advisable to focus on these 25% of the car journeys. People who take these short journeys by car could be relatively more effectively stimulated to take the bike instead. This is due to the absence of the role travel distance normally plays while choosing certain transport types.

Focus on metro and bus travels
Paragraph 3.3 also concluded that metro and bus journeys largely cover the same distance as a bike journey. Alike short car journeys, distance is also not a relevant decision factor, which will make people more accessible for stimulation to use active transport for these journeys. Oliver et al. (2010) also mentioned the opportunity of relatively short distances for motorized trips. In according to the research of Oliver et al., short journeys found for non active transport have during this study identified the opportunity for intervention and promoting transport related physical activity for short commuting trips.

5. FUTURE WORK
The research discussed in this paper is done with a sample group, consisting of young adults who more or less all lives in or nearby Amsterdam. To validate the results in this paper, the research should be done a second time with participants living in another area. Amsterdam contains an advance public transport infrastructure which probably influenced the travel behavior of the participants, which is also mentioned by Oliver et al.(2010)

The experiment is done during the months February and March. Concluding statements using data about the weather during this period could be biased. The weather score during this period was almost continuously positive and did not drop below a 5. This way for example, activity behavior during “heavily bad weather” was not measurable. In case the experiment is done a second time, it is advisable to choose a period when the variation in weather is larger.

The participants were not fully aware of the diversity of transport options they could take to a certain location, which probably influenced their commuting behavior. Useful data could be gathered when the commuting behavior is measured while people do know all the transport options. This behavior could be compared against the behavior explained in paragraph 3.1, and the effect of awareness shortcoming can be analyzed and discussed.

REFERENCES


APPENDIX A
Static participant information.

<table>
<thead>
<tr>
<th>Answer</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>15</td>
<td>63%</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>38%</td>
</tr>
</tbody>
</table>

*Table 2: “What is your gender?”*

<table>
<thead>
<tr>
<th>Answer</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>8%</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>8%</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>25%</td>
</tr>
<tr>
<td>22</td>
<td>5</td>
<td>21%</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>24</td>
<td>4</td>
<td>17%</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>8%</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>4%</td>
</tr>
</tbody>
</table>

*Table 3: “What is your age?”*

<table>
<thead>
<tr>
<th>Answer</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I live at home</td>
<td>14</td>
<td>58%</td>
</tr>
<tr>
<td>I live by myself</td>
<td>10</td>
<td>42%</td>
</tr>
</tbody>
</table>

*Table 4: “Do you live at home or by yourself?”*

<table>
<thead>
<tr>
<th>Answer</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, fulltime</td>
<td>20</td>
<td>83%</td>
</tr>
<tr>
<td>Yes, part-time</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>No</td>
<td>4</td>
<td>17%</td>
</tr>
</tbody>
</table>

*Table 5: “Are you a student?”*

<table>
<thead>
<tr>
<th>Answer</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>4</td>
<td>17%</td>
</tr>
<tr>
<td>Yes, one</td>
<td>16</td>
<td>67%</td>
</tr>
<tr>
<td>Yes, more</td>
<td>4</td>
<td>17%</td>
</tr>
</tbody>
</table>

*Table 6: “Do you have a (part-time) job?”*

<table>
<thead>
<tr>
<th>Answer</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>2</td>
<td>8%</td>
</tr>
<tr>
<td>Yes, one</td>
<td>13</td>
<td>54%</td>
</tr>
<tr>
<td>Yes, more than one</td>
<td>9</td>
<td>38%</td>
</tr>
</tbody>
</table>

*Table 7: “Do you sport(s)?”*

APPENDIX B
Type of transport versus destination.

<table>
<thead>
<tr>
<th>Type of transport</th>
<th>Work(%)</th>
<th>Study(%)</th>
<th>Sport(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>35.79</td>
<td>19.39</td>
<td>15.31</td>
</tr>
<tr>
<td>Bike</td>
<td>36.84</td>
<td>68.37</td>
<td>46.94</td>
</tr>
<tr>
<td>Car</td>
<td>24.21</td>
<td>0</td>
<td>24.49</td>
</tr>
<tr>
<td>Tram</td>
<td>7.37</td>
<td>0</td>
<td>2.04</td>
</tr>
<tr>
<td>Metro</td>
<td>7.37</td>
<td>24.49</td>
<td>1.02</td>
</tr>
<tr>
<td>Train</td>
<td>25.26</td>
<td>69.39</td>
<td>3.06</td>
</tr>
<tr>
<td>Bus</td>
<td>41.05</td>
<td>26.53</td>
<td>7.14</td>
</tr>
<tr>
<td>Public Transport</td>
<td>52.63</td>
<td>69.39</td>
<td>9.18</td>
</tr>
</tbody>
</table>

*Table 8: Transport choice*

![Figure 12: Transport to sport](image)

![Figure 13: Transport to study](image)

![Figure 14: Transport to work](image)
APPENDIX C
Three bar charts representing amount of transport routes (including transport types) people were aware about.

Figure 15. Transport routes to sport

Figure 16. Transport routes to work

APPENDIX D
Correlation between weather and type of transport.

<table>
<thead>
<tr>
<th></th>
<th>Rain</th>
<th>Avg.Temp</th>
<th>Avg.WindSpd</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>0.02</td>
<td>-0.16</td>
<td>-0.16</td>
<td>-0.26</td>
</tr>
<tr>
<td>Bike</td>
<td>-0.03</td>
<td>-0.39</td>
<td>-0.13</td>
<td>-0.28</td>
</tr>
<tr>
<td>Tram</td>
<td>-0.03</td>
<td>-0.22</td>
<td>-0.06</td>
<td>-0.21</td>
</tr>
<tr>
<td>Metro</td>
<td>-0.02</td>
<td>-0.25</td>
<td>-0.24</td>
<td>-0.25</td>
</tr>
<tr>
<td>Train</td>
<td>0.05</td>
<td>-0.27</td>
<td>-0.14</td>
<td>-0.35</td>
</tr>
<tr>
<td>Bus</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-0.18</td>
<td>0.01</td>
</tr>
<tr>
<td>Public Transport</td>
<td>0.03</td>
<td>-0.23</td>
<td>-0.14</td>
<td>-0.27</td>
</tr>
</tbody>
</table>

Table 9: Correlation test for weather and transport

APPENDIX E
Box plot presenting the variation of the amount of steps walked per day.

Figure 18. Amount of steps

APPENDIX F
Box plot comparing weather versus amount of steps.

Figure 17. Amount of steps per weather category