Demographic and Geographic Heterogeneity between Subgroups of Cyclists

Jackson Tyler Lester

A thesis

Submitted in partial fulfillment of the

Requirements for the degree of

Master of Science in Civil Engineering

University of Washington

2016

Committee:

Don Mackenzie

Anne Vernez-Moudon

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

Demographic and Geographic Heterogeneity between Subgroups of Cyclists

Jackson Tyler Lester

Chair of the Supervisory Committee:

Assistant Professor Don Mackenzie

Civil and Environmental Engineering

This paper seeks to improve the understanding of heterogeneity of behavior, preferences, and satisfaction within a sample of cyclists and how that heterogeneity is connected with demographic and geographic factors. It develops a method for identifying subgroups of the sample of cyclists and discusses how those groups differ sociodemographically and geographically in their cycling behavior to provide policymakers with a tool to better understand how cycling infrastructure improvements will affect different groups within the sample. It describes a survey of 905 cyclists completed in 2015 in Malmö, Sweden. First it reduces the dimensionality of the information from the survey to eight components with Factor Analysis of Mixed Data, a principal component analysis method. Next it demonstrates a categorization with six groups of the respondents with heterogeneous cycling behavior, preferences, and satisfaction with hierarchical clustering. Then it details sociodemographic and geographic differences between those groups. This paper confirms the thesis that it is possible to recognize sociodemographic and geographic similarities within and differences between subgroups of cyclists.

Acknowledgements

Completing this thesis would not have been possible without the support and guidance of people from the University of Washington and the KTH Royal Institute of Technology in Stockholm.

Thank you to Professor Don Mackenzie for guiding me through this progress, and providing me with help in understanding the tools and concepts through Skype meetings with a nine-hour time difference throughout the year.

Thank you to Professor Joel Franklin and all of the wonderful people at the Center for Transport Studies at KTH for making me feel at home in Stockholm, providing me with work space, and guidance throughout the year.

Thank you to Mark Hallenbeck for encouraging me to go to Stockholm for the year and experience life and transportation research in Europe.

Thank you to Dayna Cole and Scott Rutherford from the Valle office for support throughout this process.

Thank you to Johan Irvenå of Malmö Stad for providing survey data and translation support.

This research was supported by the Valle Scholarship and Scandinavian Exchange Program at the University of Washington.

Table of Contents

1 Introduction 11

1.1 Rationale for promoting utilitarian cycling 11

1.1.1 Reducing the negative impacts of driving 12

1.1.2 Improving health of cyclists 13

1.2 Connections between infrastructure, policy, demographics, and cycling 14

1.2.1 Influences on cycling frequency 14

1.2.2 Influences on mode choice 16

1.2.3 Influences on route choice 18

1.3 Identifying and describing subgroups of cyclists 19

1.3.1 Segmenting populations of cyclists 20

1.3.2 Identifying demographic and geographic differences between subgroups of cyclists 22

1.4 Thesis Overview 24

2 Malmö 26

2.1 Cycling in Malmö 27

3 About the survey 30

3.1.1 Profile of survey respondents 34

3.1.2 Respondents' willingness to cycle in various conditions 36

3.1.3 Analysis of respondent satisfaction and preferences 38

4 Categorization of cyclists 45

4.1 Factor analysis of mixed data (FAMD) grouping of survey questions 46

4.2 Hierarchical clustering of cyclists 55

5 Sociodemographic aspects of clusters of cyclists 59

6 Spatial analysis of cycling behavior 65

7 Conclusions 78

8 Works Cited 82

Appendix A- Survey in original Swedish 91

Appendix B- Survey translated to English 99

Appendix C- Postal codes considered in the analysis 111

FIGURES

Figure 1- Images of neighborhoods not conducive (left) and conducive (right) to utilitarian cycling and/or walking (Design New Haven, 2009)... 17

Figure 2- Malmö's population from 1970 to 2015 (Statistics Sweden, 2016) 26

Figure 3 - Mode splits in Sweden’s 9 largest cities, year in parentheses, population below year (European Platform on Mobility Management, 2016)... 28

Figure 4- Malmö's bicycle routes and network of bicycle pumps (Malmö stad, 2016) 29

Figure 5- Scree plot of the portion of variance described by each successive component........... 50

Figure 6- Dendrogram of the hierarchical clustering using Ward's method with Manhattan distance, shown split into six clusters... 56

Figure 7- Values of significant (P>.95) components' within-cluster means................................ 58

Figure 8- Number of origins and destinations in central postal codes... 67

Figure 9- Trips in analysis by cluster... 69

EQUATIONS

Equation 1- Equation for Spearman rank correlation (Laerd Statistics, n.d.) 39

Equation 2- Scaling continuous responses to unit variance 48

Equation 3- Specific scaling of MCA for categorical variables 48

Equation 4- Manhattan distance between points P1 and P2 in N dimensions 55

Equation 5- Pearson's chi squared test statistic 59

TABLES

Table 1- Malmö climate information (belsoft, n.d.) 27

Table 2- Questions from survey used in study with responses and abbreviations used throughout paper... 31

Table 3- Survey respondents' cycling behavior in a variety of conditions, frequency, and perceptions of safety. (n=905) 36

Table 4- Demographic information about survey respondents and the area at large. (n=905)..... 37

Table 5 - Survey responses of important route attributes with correlations between the rankings of each attribute, n=644. On a scale from 1= least important to 10= most important. All respondents selected ranked importance, although not all respondents gave complete ranking. . 40

Table 6- Survey responses of satisfaction with aspects of Malmö as a city for bicycles. On a scale from 1= least important to 5= most important.(n=644) 42

Table 7- Survey responses of ratings of importance of aspects that make routes bicycle friendly. On a scale from 1= least important to 5= most important.(n=644) 43

Table 8 continued-Component coefficients and contributions for each variable from the original dataset 53

Table 9- Names and profiles of the eight components 54

Table 10- Names and descriptions of the six clusters of cyclists considered 57

Table 11- Contingency tables of each cluster's membership by gender, age, whether they have children, education, whether they have a partner, employment, and income 60

Table 12- Contingency table of each cluster's membership by driver's license status, car access, bicycle access, and how often they ride 61

Table 13- Demographic profiles of the clusters 62

Table 14- Origin-destination matrix for each of the three-digit postal codes with origins and/or destinations 66

Table 15- Area, population, density, and income of five-digit post codes grouped by zone of analysis (ESRI, 2016) 68

Table 16- Pie charts of each origin and destination zone by cluster with maps of each zone 70

Table 17- Heat maps of odds ratios of cluster membership within zones and cluster membership in the entire sample for all clusters 74

Table 18- Profiles of all clusters based on percentage of origins and destinations by zone 77

1 Introduction

Urban transportation has changed dramatically over the past two centuries, moving from domination by human and horse-powered modes, to domination by streetcars, to today's paradigm of domination increasingly by personal automobiles. As this automobile supremacy has taken hold, cities around the world have become more populated, congested, and polluted while their residents have become increasingly sedentary. To make room for more residents to live and travel within our cities, to reduce the pollution and noise created by automobiles, and to live healthier and more active lives, many cities are working to encourage active forms of transportation, namely walking and cycling.

Malmö, which draws much inspiration from neighboring Copenhagen, Denmark, is already a leader in Sweden and in the world for encouraging active transportation to its residents. Malmö has promoted cycling through an extensive network of cycle tracks, widespread bicycle parking, and campaigns to cut out unnecessary car trips. Because of these efforts, Malmö ranked 6th in the world in the 2015 Copenhagenize Index of bicycle friendly cities, which ranks cities based on their performance in 13 areas of bicycle friendliness. (Copenhagenize, 2015) In a 2013 survey of 12,400 residents of Malmö, 22 percent of total trips were taken by bicycle. (EPOMM, 2013)

1.1 Rationale for promoting utilitarian cycling

Because of increasing concern with the negative aspects of driving, the health of residents, and transit accessibility, cities have become interested in encouraging bicycling for utilitarian (non-recreational) trips. By convincing more residents to use bicycles, cities hope to make transportation more active, affordable, and convenient, while reducing congestion and pollution. Macmillan et al. (2014) used system dynamics modeling of commuter cycling under five policy scenarios in Auckland, New Zealand over the next 40 years. They estimate that upgrading urban roads to include separated cycle tracks on main roads and reduced speed limits on smaller roads, would yield a benefit/cost ratio of 10-25 with the largest benefits coming from increased physical activity, reduced fuel costs, and improved air quality. (Macmillan, et al., 2014) This demonstrates the potentially great value Auckland of investment in cycling infrastructure.

1.1.1 Reducing the negative impacts of driving

Reliance on automobiles has profound impacts on cities and how we experience them. Each automobile adds to the congestion, noise, and air pollution that affect all residents. When residents replace car trips with bicycle trips, they reduce the number of automobiles on the road and diminish the related negative impacts. Woodcock et al. (2009) forecasted 2030 transportation scenarios in London, England and Delhi, India with increases in active transportation. They estimated that policies to promote increases in active transportation will reduce the number of lost disability-adjusted-life-years per million population per year by 12,516 in Delhi and 7,332 in London. These reductions in lost disability-adjusted-life-years come from cyclists' increased physical activity, reduced air pollution, and a reduction in traffic crashes in Delhi. (Woodcock, et al., 2009) Rabl and de Nazelle (2012) argued that although reductions in CO_2 emissions are small, especially for smaller cities, the largest social gains from bicycling are the reduced noise, reduced congestion, and health gains for cyclists. They estimated the noise and congestion to benefit society by more than €1,700 per individual and the health gains to benefit society by more than €1,300 per individual who switches from driving to bicycling per year, based on estimates from Maibach et al. (2008) of the average damage costs due to congestion and noise per passenger car kilometer. (Rabl & de Nazelle, 2012) (Maibach, et al., 2008)

1.1.2 Improving health of cyclists

Using active modes of transportation in day-to-day life can greatly improve users' health outcomes. Increased physical activity has been demonstrated to reduce the risk of cardiovascular disease, depression, dementia, diabetes, obesity, and breast and colon cancer. Additionally, as the distance cycled in cities substantially increases, the absolute numbers of cyclists killed or seriously injured has been shown to decrease. (Woodcock, et al., 2009)

In studying cost benefit analysis of cycling infrastructure, Börjesson and Eliasson (2012) discussed whether the health benefits of cycling should be considered independently in a cost benefit analysis. They argued that while the health benefits of cycling are potentially very large and people tend to cycle more in response to infrastructure improvements, individuals take the health impacts of cycling into account when making their decision whether to cycle. Because of this decision level inclusion, they stated that including the health benefits in the cost benefit analysis would most likely be double counting them. (Börjesson & Eliasson, 2012) Based on this rationale, conceptually individuals should be eager to take up cycling, even more so if they are educated about the large health benefits.

The overwhelming consensus in prior research is that there are large public benefits to cycling both for the cyclists and society at large. Handy et al. (2014) discussed the challenges and research needs for promoting cycling into the future and provided a broad review of prior research on the social benefits of cycling. They argued that future transportation should focus on measuring how much cycling there is, who is doing it and why, to better understand what interventions may best and most cost effectively promote cycling. They additionally argued that the research must be accessible to planners and policy makers, sometimes at the expense of complexity, to ensure the research is used in practice. (Handy, et al., 2014)

1.2 Connections between infrastructure, policy, demographics, and cycling

Although past research shows utilitarian cycling to be a good thing for cyclists, motorists, and society at large, there are several important barriers that keep many people from cycling. These barriers can include poor physical fitness, safety concerns, long/hilly/busy commutes, bad weather, dropping off children, and lack of awareness of cycling options. (Bauman, et al., 2008) (Parkin, et al., 2007) Based on the evidence of health benefits, and reduced externalities from vehicle use, communities have enacted infrastructure and policy initiatives to address these barriers and encourage more people to commute by bike. Poorfakhraei and Rowangould (2015) studied the welfare change in terms of willingness to pay for improvements in cycling infrastructure including bicycle paths, bicycle lanes, and street lighting through a stated preference survey. They found that cycle tracks were valued most highly, followed by street lighting and then bicycle lanes. People with more cycling experience had lower willingness to pay for these improvements while older people had a higher willingness to pay. Predictably, higher income individuals and people living in higher income areas were more likely to value improvements more highly. (Poorfakhraei & Rowangould, 2015)

1.2.1 Influences on cycling frequency

Perception of cycling infrastructure has been shown to be an important factor in affecting cycling behavior. Ma et al. (2014) studied the perceived and objectively measured built environment and their impacts on cycling behavior in Portland, Oregon. They found that objectively good cycling infrastructure was only indirectly important to promote cycling by way of users' perception of good cycling infrastructure. They emphasized the importance of intervention programs to improve people's view of the built environment for cycling to increase cycling. (Ma, et al., 2014)

This demonstrates the importance of advocacy for cycling in addition to a conducive built environment to encourage more people to commute by bicycle. Grouping the factors that affect cycling behavior can help conceptualize proper interventions. Winters et al. (2011) surveyed 1402 current and potential cyclists in the Vancouver, Canada area. They asked about 73 motivators and deterrents of cycling, which they then grouped into 15 factors, analyzing average responses of cyclists in groups based on how often they ride. The factors they found to have the most influence on the likelihood of cycling included safety, ease of cycling, weather, route conditions, and interactions with motorized vehicles. (Winters, et al., 2011)

Advocating utilitarian cycling to recreational cyclists may be a convenient way to encourage people who have high utility of cycling to use it more for transportation. Sahlqvist and Heesch (2012) studied recreational cyclists in Queensland, Australia through an online survey of members of local cycling community and advocacy group. They found that 47% of respondents cycled for utilitarian purposes, and being male, young, full-time employed, and university educated made participants more likely to be utilitarian cyclists. They argued that because recreational cyclists already have a high utility in cycling and less than half of the recreational cyclists interviewed ride for utility, governments should take steps to promote utility cycling to recreational cyclists and their advocacy groups, especially to women and older cyclists. (Sahlqvist & Heesch, 2012) People who view cycling as a form of exercise have been shown to be more willing to cycle for transportation. Akar and Clifton (2009) conducted a web-based survey of cyclists on and around the College Park campus of the University of Maryland about bicycle infrastructure, policies, and programs. The study found that people are more time-sensitive for non-motorized forms of transportation and that women are less likely to commute by bicycle. Additionally it found that respondents who perceive walking and biking as a form of exercise and who find flexible departure time to be important are more likely to ride bicycle. Finally, Akar and Clifton found that most of the people driving to campus didn't perceive any other commute options and therefore never really considered cycling. (Akar & Clifton, 2009)

1.2.2 Influences on mode choice

People's surrounding land use can significantly impact their transportation mode choices. Cevero and Kockelman (1997) studied the aspects of the built environment, grouped into density, diversity, and design aspects, and their influence on travel demand in the San Francisco Bay area using 1990 travel diary and land use records. They found that density, diversity of land use, and pedestrian-centered design tend to promote non-vehicle travel and reduce total trip rates, although with elasticities between .06 and .18, the effect was not very strong. (Cervero & Kockelman, 1997) Handy et al. (2005) similarly examined the relationship between Northern California neighborhood design elements and resident travel preferences. Although prior studies had identified a relationship suggesting that residents drive less in dense, diverse, and transit accessible neighborhoods, the direction of the causal relationship has been the subject of much debate. They found that while travel behavior differences between suburban and traditional neighborhoods were largely explained by attitude, even when attitudes were controlled for they still found a relationship between neighborhood design and travel behavior, suggesting that travel behavior is influenced by neighborhood design. (Handy, et al., 2005) Figure 1 depicts a simplified depiction of a neighborhood designs that discourage and encourage active transportation.

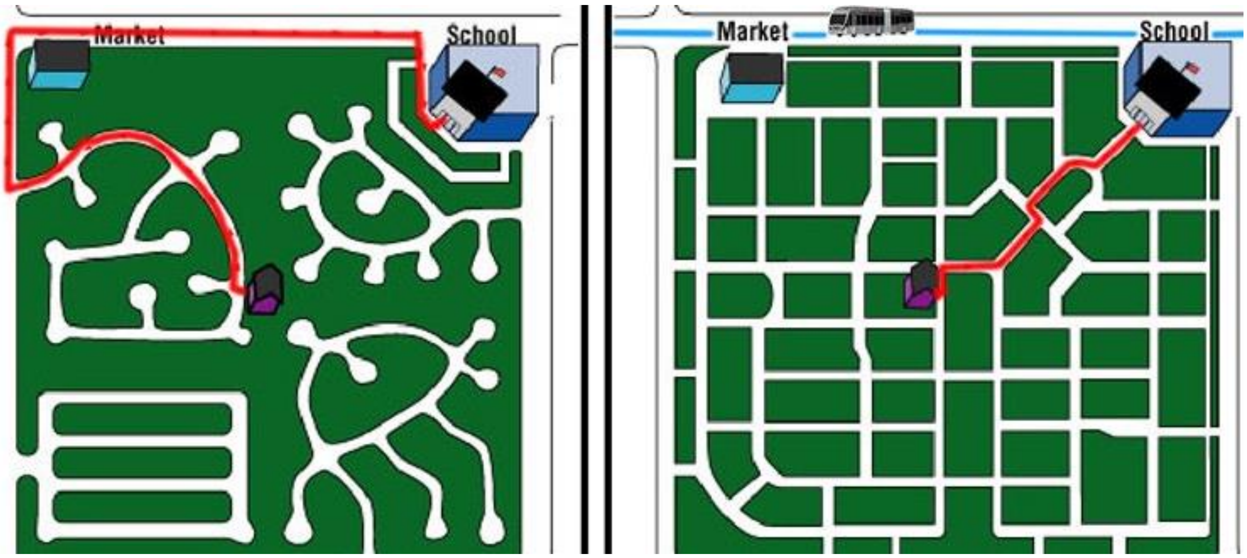


Figure 1- Images of neighborhoods not conducive (left) and conducive (right) to utilitarian cycling and/or walking (Design New Haven, 2009)

Quality cycling infrastructure has been demonstrated to influence a higher cycling mode share. Dill and Carr (2003) examined data from 43 large cities across the United States, finding that higher levels of cycling infrastructure have a positive and significant correlation with the rate of bicycle commuting. This supported claims from previous studies, but examined a wider variety of cities that were not exclusively college towns. (Dill & Carr, 2003) Titze et al. (2008) studied the commuter cycling mode share of residents through a phone survey of 1,000 residents of Graz, Austria. They found that cycling correlated positively with bicycle lane connectivity and cycling-related social support, and negatively correlated with people associating cycling with physical discomfort and an impractical mode of transportation. (Titze, et al., 2008)

While infrastructure improvements increase utilitarian cycling, policies such as direct payments to cyclists have been shown to also strongly encourage cycling. Wardman et al. (2007) studied the factors that affect individuals' propensity to cycle to work. They demonstrated through a variety of revealed and stated preference surveys that an improvement to completely segregated bicycle lanes for an entire route would only increase the total bike commute share by 4.2% of the baseline mode share for bikes, with less improvement for smaller-magnitude improvements. They demonstrated that direct payments to cyclists for commuting to work are an effective incentive and trip-end facilities such as showers and bike parking are shown to only shown to slightly increase bicycle commuting. Combining all three sets of strategies, their model predicted the largest shift to bicycle commuting. (Wardman, et al., 2007) Parkin et al. (2008) used census data to estimate the determinants of bicycle mode share for work commutes. They estimated a 43% saturation level for bicycle use. They found a smaller proportion of cyclists in areas with more females and car ownership. The most significant physical variable they found in affecting peoples' decision to bike to work was hilliness. They argued that infrastructure alone is insufficient to engender higher levels of cycling. (Parkin, et al., 2008)

Other cycling related policies can also have significant influences in cyclist mode split. Rietveld and Daniel (2004) analyzed the impacts of municipal policies on bicycle mode share by comparing policies and cycling behavior between cities around the world. They found that relevant policy-related variables include the number of stops cyclists need to make in their routes, obstacles to road use, and cyclist safety. They additionally found that the relative position of bikes compared with cars, meaning policies including car speeds and parking costs, affected cycling behavior. Finally they observed that cities' ethnic makeup affected cycling use, which is a variable typically not considered in similar studies. (Rietveld & Daniel, 2004)

1.2.3 Influences on route choice

Another method for observing how cyclists value cycling infrastructure and policies is to observe the characteristics of routes they ride. Tilahun et al. (2007) conducted a stated preference survey in which they offered respondents a choice between a superior facility with longer travel time and an inferior facility with reduced travel time. They found that respondents were willing to travel as much as 20 minutes longer to change from a road without a bike lane and with side parking to an off-road bike trail, with smaller changes connected with smaller improvements. (Tilahun, et al., 2007) Broach et al. (2012) observed the behavior of 164 cyclists with GPS units along with a log of each of the corresponding trips that the cyclists took. In generating a route choice model with the observed routes, they found that cyclists respond to distance, turn frequency, slope, signalization of intersections and traffic when choosing routes. They additionally found that cyclists placed particularly high value on off-street bike paths and “bicycle boulevards” with traffic calming features, but bike lanes on the street basically offset the negative impact of vehicle traffic. Finally they found that cyclists were more sensitive to distance and less sensitive to all other route attributes for commute trips compared with the other utilitarian trips considered in the study. (Broach, et al., 2012)

Not all cyclists experience route attributes in the same way, indicating potential for varying types and levels of infrastructure and policies to appease different people. Garrard et al. (2008) examined whether female commuter cyclists were more likely to use bicycle routes that provide separation from vehicles by counting cyclists observed during peak commute times in different types of locations in Melbourne, Australia. They found that female commuters more strongly preferred paths with more separation from vehicle traffic. (Garrard, et al., 2008)

1.3 Identifying and describing subgroups of cyclists

The previous sections have made the case for why policy makers and urban planners should encourage cycling, and which variables affect changes in cycling frequency, mode choice, and route choice. This brings into question who should be targeted by interventions and where areas of need exist. The following sections describe the literature on segmenting populations of cyclists into groups and identifying areas of cycling infrastructure need to develop a basis to target cycling infrastructure.

1.3.1 Segmenting populations of cyclists

Cyclists are a diverse segment of the general population, and are frequently separated into more homogenous subgroups for city planning purposes. This concept of cyclist segmentation is central to the purpose of this paper, which aims to tailor infrastructure to the people who are likely to change their behavior based on specific improvements in specific areas. Roger Geller (2009), Portland, Oregon's bicycle coordinator, published a paper called Four Types of Cyclists, which proposes a framework for grouping cyclist types from Portland's population at large, including non-cyclists. In it he estimated the population share of each of the groups and described the importance of mitigating fear as the catalyst for increasing the mode share of cycling compared with other modes of transportation. He developed the following categories of transportation cyclists and their needs to frame a conversation about bikeway treatments:

1. **Strong and fearless** less than 1% of the population that will travel by bicycle no matter the conditions
2. **Enthused and confiden**t around 7% of the population that will ride in the road with automobiles, but are more attracted by bicycle infrastructure
3. **Interested but concerned** around 60% of the population that likes riding bikes and would like to more often, but is largely limited by fear
4. **No way no how** around 33% of the population that cannot be convinced to ride bicycles

By reducing the fear of the people in groups three and four, Geller argued we can best encourage people to commute on bike. (Geller, 2009)

Despite giving conceptual arguments for his framework, Geller didn't back it up with any sort of rigorous evidence, which prompted Dill and McNeil (2013) to analyze Geller's categorization with a random phone survey of residents of Portland, Oregon. They found that nearly all of the respondents fit neatly into one of Geller's categories, and the majority (56%) were in the "interested but concerned" group. This was hypothesized to be the key group to help to increase the total bicycle share of commutes. Additionally, Dill and McNeil found that older adults and women were less likely to be a part of one of the more confident groups, and that decreased traffic speed and increased route separation from the road would best cater to the "interested but concerned" group. They found some strange results such as 34% of the strong and fearless, based on their stated level of comfort, had not cycled in the past 30 days. This was larger than the portion in groups two and three, which were 28% and 23% respectively. (Dill & McNeil, 2013) While Geller's framework provides a strong categorization for the population at large based on people's comfort cycling, it is not particularly useful for categorizing a sample of all cyclists, and limits the factors keeping people from cycling to only their confidence, when in reality the decision to cycle is more complex.

Damant-Sirois et al. (2014) proposed an alternate classification of distinct cyclist types that is meant to categorize people within a group of cyclists instead of the population at large. After trying testing configurations with two through eight groups of cyclists, using K-means clustering their best configuration included four groups. Their categorization of cyclists' motivators, deterrents, and infrastructure preferences, has four categories as follows:

1. **Dedicated cyclists** (24% of sample) are not strongly affected by weather and they choose to cycle largely for speed, flexibility, and predictability of travel time. They are less influenced by cycling infrastructure and sometimes even prefer riding on the road.
2. **Path-using cyclists** (36% of sample) are mainly motivated by convenience and fun. They tend to avoid cycling on streets and are highly influenced by a dedicated bicycle route that is separated by traffic.
3. **Fair weather utilitarians** (23% of sample) only cycle when it is convenient, and will choose another mode if the weather is bad or the destination is not easily accessible by bike. They tend to prefer bicycle paths and don't really consider themselves cyclists.

4. **Leisure cyclists** (17% of sample) cycle because they enjoy it and not because it is convenient or fast. They are heavily influenced by weather and are averse to riding in traffic. They tend to think of cycling as a hobby more than as a legitimate mode of transportation

Damant-Sirois et al. found significant differences in the four clusters in terms of demographics, infrastructure preferences, and behavior, which indicates that this typology can be a useful tool to predict how different interventions will affect these segments of the population of cyclists. (Damant-Sirois, et al., 2014)

Damant-Sirois and El-Geneidy (2015) used the same cyclist categorization as Damant-Sirois et al. (2014) to analyze factors that affected cycling frequency of 1707 cyclists in Montreal, Canada with an ordinal logistic regression model. After separating participants into the four groups, they found some significant differences in impact on cycling frequencies based on changes in the covariates between the four groups. After building a model that separated the cyclists by group membership, they found that different interventions affected members of each groups of cyclists differently, indicating the importance of segmenting cyclists to account for their heterogeneity. They found the most generally useful improvements in increasing cycling frequency to be making cyclists feel safe in all riding situations including roadways and stressing the affordability and convenience in cycling for transportation. Additionally they found that while men were more likely to commute to work, women were more likely to cycle for non-work utilitarian trips. (Damant-Sirois & El-Geneidy, 2015)

1.3.2 Identifying demographic and geographic differences between subgroups of cyclists

After establishing a rationale and some methods for segmenting the population of cyclists, the next goal is to establish a basis for identifying sociodemographic and geographic differences between those groups. Zahran et al. (2008) studied the spatial distribution of active modes of transportation by county across the entire contiguous US. They used GIS and linear regression analysis to study built, natural, socioeconomic, and civic environments. They found that the number of active commuters in a county increases with population density, natural amenities, education, wealth, and estimates of local civic concern. Alternatively, they found that the number of active commuters decreases with increases in pollution and average distance travelled by a work commuter. (Zahran, et al., 2008)

At a smaller scale, Shafizadeh and Niemeier (1997) examined the connections between geographic clustering and demographic characteristics of cyclists and their commuting travel durations by examining a survey distributed to cyclists. The survey was conducted in common weekday bicycle commute locations in Seattle, Washington at morning and afternoon peak commute times. They studied the varying origins and destinations of respondents' trips by categories of demographic variables, finding perceptible differences between income ranges, age groups, and genders. Using linear regression, they found that gender and income had opposite effects on travel time depending on whether the destinations were inside or outside the central business district. They additionally found that higher income and older participants tended to report longer bicycle commute times. (Shafizadeh & Niemeier, 1997)

Larsen et al. (2010) studied how far survey respondents were willing to walk or cycle to access different destinations in the Montreal, Canada area. They examined how distances traveled varied by location and individual sociodemographic and travel attributes. They found the median distance cycled to be around two kilometers with large variations by age, gender, and geographic area. Men had a larger median bicycle trip distance than women. They additionally found that respondents were willing to walk and cycle farther for work related trips than trips for other purposes. (Larsen, et al., 2010) This analysis will be pertinent to this paper's discussion of sociodemographic and geographic attributes that vary between groups of cyclists and distances traveled for reported trips.

1.4 Thesis Overview

This paper seeks to improve the understanding of heterogeneity within the population of cyclists. Developing a method for better understanding subgroups of cyclists by identifying categories of cyclists and recognizing demographic groups and locations where each of those groups ride to and will provide policymakers with a new tool to understand the impacts of cycling investment. Therefore it is hypothesized that,

> Through identification of subgroups based on cycling behavior, preferences, and satisfaction, it is possible to recognize sociodemographic and geographic similarities within and differences between subgroups.

The paper additionally aims to create a conceptually appealing framework for understanding how cyclists' preferences for route attributes differ and where each cyclist experiences relative comfort in commuting by bike. Chapter 2 describes Malmö and its background of cycling and cycling infrastructure, and describes the responses to the survey.

Chapter 3 describes the survey and gives an overview of survey responses. It first describes the demographic breakdown of respondents and of the regional population. Then it describes the correlations between the responses to questions within sets about importance of route attributes and satisfaction with Malmö as a cycling city.

Chapter 4 describes a framework and categorization of cyclists by cycling preferences, behavior, and satisfaction in Malmö. It begins by reducing the dimensionality of the data by using factor analysis of mixed data (FAMD), which is a principal component analysis (PCA) method meant for simultaneously analyzing continuous and categorical data. With FAMD the 46 variables considered are combined to form eight uncorrelated components that describe various aspects of cycling behavior, preference, and satisfaction of respondents. Those components are then each described based on the variables from the original data that influence them most heavily.

Based upon those eight components, the survey respondents are hierarchically clustered to form six clusters of cyclists that minimize within-cluster variation along the eight components created with FAMD. The attributes of those clusters are then described in terms of their within-cluster mean locations along the components to develop profiles of each cluster.

Chapter 5 examines the sociodemographic comparisons between the cyclists of each cluster. First, the clusters are cross-classified across categories of sociodemographic variables, and Chi-squared tests are performed to measure the significance of sociodemographic differences between cyclist clusters. Then the sociodemographic profiles of each of the clusters of cyclists are described to develop a better understanding of which types of people belong to each cluster of cyclists.

Chapter 6 examines the geospatial cycling patterns of each of the clusters developed in chapter 4. The clusters are first analyzed by the frequencies of origins and destinations within regions around Malmö, and chi-squared tests are performed to determine the significance of differences between clusters. Next, each cluster is described by a heat maps that describe the odds ratio of trip activity including origins and destinations occurring within different region. Finally geospatial profiles of origins and destinations for each of the clusters is described.

Finally, Chapter 7 presents discussion and implications of findings from the paper and considers limitations and future extensions of this research.

2 Malmö

Malmö is the 3rd largest city in Sweden, with over 320,000 residents, and is located across the Öresund Bridge from Copenhagen, Denmark. Malmö and the surrounding region are experiencing major post-industrial growth after a period of significant decline in the 1980’s and 1990’s due to the decay and subsequent closure of its formerly booming shipbuilding industry. In fact, Malmö’s population declined from around 265,000 in 1970 to around 230,000 in 1985, only to begin growing again in the 1990’s largely due to immigration of refugees from war-torn countries. (Baeten, 2011) Malmö’s population over time is shown in Figure 2. Malmö’s post-industrial development has included growth in the logistics, trade, and high-tech sectors, which have been especially exaggerated since the completion of the Öresund Bridge in 2000. Additionally Malmö University College, one of Sweden’s newest higher education institutions, began operating in the city in 1998, providing another resource to develop a highly educated workforce. (Malmö stad, n.d.)

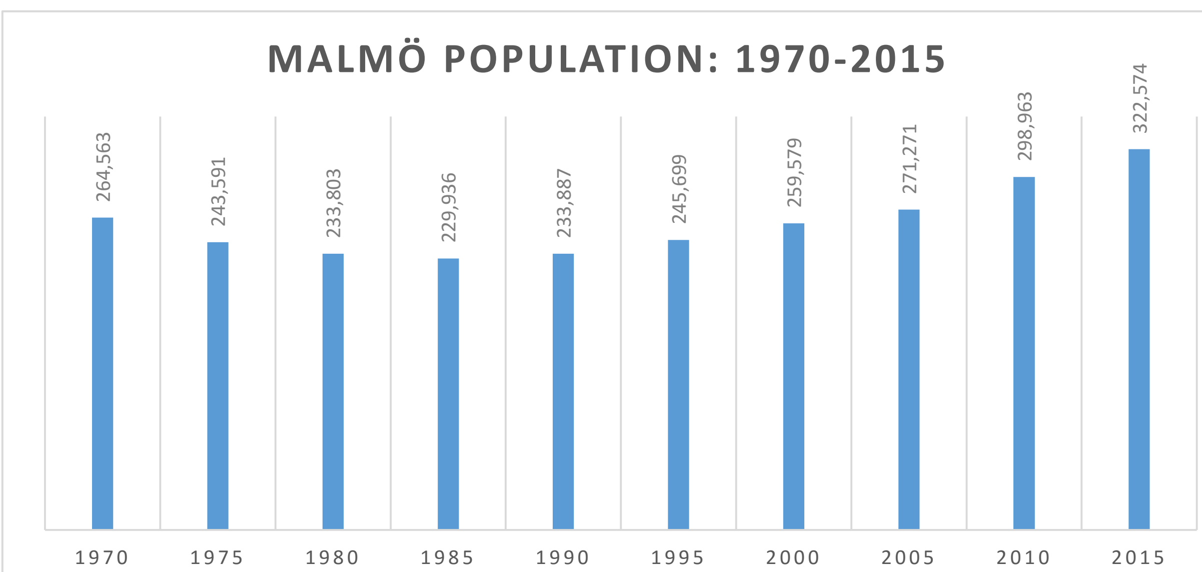


Figure 2- Malmö's population from 1970 to 2015 (Statistics Sweden, 2016)

Malmö has made sustainable planning and development a priority since the mid-1990s. The city has appropriated significant funds to sustainable business and housing developments, alternative fuel vehicles, active transportation infrastructure, and renewable energy resources. This trend has been especially apparent in areas such as the Western Harbor, where Malmö's former shipbuilding industry center has been turned into a model for sustainable development. In fact, Malmö is pursuing the goal of becoming carbon neutral by 2020, and by 2030 they propose to run the entire city on renewable energy. One of the major factors that will allow this transformation is a transition to clean forms of transportation. (Anderson, 2014)

2.1 Cycling in Malmö

Both Malmö and Copenhagen are widely considered to be among the most bicycle friendly cities in the world, and in both cities there is significant political will to improve sustainability by reducing reliance on non-renewable energy. One of the highly publicized methods of achieving this has been promoting travel by public transit and active modes of transportation. Compared with other cities in Sweden, Malmö is relatively flat, and due to its situation near the ocean and in Southern Sweden, it has a mild climate described in Table 1. The combination of these factors makes Malmö particularly well suited to travel by bicycle, allowing commuters to travel by bicycle all year in relative comfort.

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
Average high in °C	2	2	5	10	16	20	21	21	17	12	7	4
Average low in °C	-3	-3	-1	2	7	11	13	12	10	7	3	-1
Av. precipitation - mm	49	36	30	33	38	46	65	62	53	53	44	43
Days with precipitation	17	13	10	11	10	11	14	13	13	14	15	16
Hours of sunshine	39	66	125	177	269	263	253	215	172	84	35	20

Table 1- Malmö climate information (belsoft, n.d.)

Compared with other large Swedish cities, Malmö already has a significant portion of its residents commuting by bicycle. In fact as of 2013, 22 percent of commute trips and Malmö were taken by bicycle, which is a substantially larger portion than in Gothenburg and Stockholm, and is only surpassed by cities with larger portions of students among Sweden's nine largest cities, whose mode shares are displayed in Figure 3. Although Malmö is ideally situated for transportation by bicycle, the city has also achieved its high level of bicycle mode share through initiatives meant to do just that.

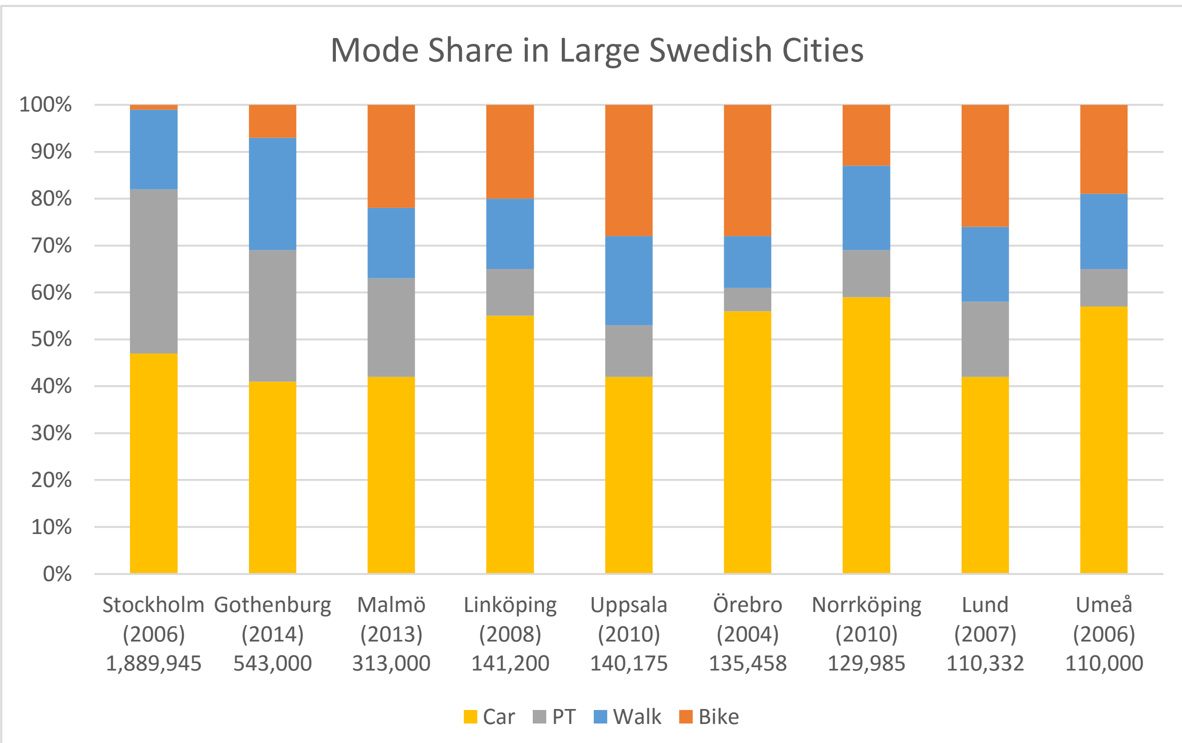


Figure 3 - Mode splits in Sweden's 9 largest cities, year in parentheses, population below year (European Platform on Mobility Management, 2016)

To encourage residents to commute by bicycle, Malmö has made cycling safer and more convenient. As of 2012, Malmö contained more than 470 km of bicycle paths, more than 60,000 bicycle parking spaces on municipal land, several air stations to fill tires dispersed around the city, and a network of cycling-specific traffic signals to accommodate cyclists crossing roads. Many of the traffic signals are now even capable of detecting cyclists with radar to trigger them to allow them through intersections on demand. (Bicycle Malmö 2012, 2012) A map of Malmö's bicycle path network and the distribution of free air stations is shown in Figure 4. To make cycling to and from public transportation more attractive, Malmö has created covered and secured bicycle parking garages near train stations to provide convenient and safe parking to cyclists and has encouraged cycling with bicycle counters that display how many cyclists have passed on a particular bicycle path each day. (Hu & Li, 2013) Although Malmö has taken significant steps to encourage bicycling, it is still interested in increasing the bicycle mode share to move closer to its sustainability goals. To achieve this objective and to understand how residents have responded to existing improvements, the city of Malmö conducts surveys of its residents about their cycling experience in Malmö.

Figure 4- Malmö's bicycle routes and network of bicycle pumps (Malmö stad, 2016)

3 About the survey

Between May and September of 2014 the traffic department of Malmö, Sweden conducted a survey of local cyclists online. The department encouraged residents to participate in the survey by distributing flyers targeted toward cyclists. In total, 905 cyclists responded to the survey. The survey consisted of 33 questions addressing the respondents' demographic information, cycling behavior and preferences, and policy priorities. After completing the survey participants were asked to demonstrate a typical route they ride on a map of Malmö by clicking their route on a map. Because the routes were simply straight lines between clicks on the map, and because Malmö has a high-density network of roadways where cyclists can ride, this part of the survey is not readily useful in determining cyclists' precise routes and therefore the particular route decisions the users made.

Cyclists indicated their preferences for route attributes, ranking the following attributes from most important to least important: distance, time, traffic, number of intersections, cycling network consistency, surface, safety, environment, and equipment. They also rated the importance of another set of route attributes on a scale from one to five including: safety at crossings, cycle lane width, snow removal, swept leaves, swept dirt, pedestrian-cyclist separation, removal of level differences, minimization of stops, smooth surface, control of plants, and police checks for proper use of cycling infrastructure. Finally they rated their satisfaction with attributes of Malmö's cycling infrastructure from one to five including: overall satisfaction, number of cycling paths, condition of cycling paths, path width, diversions at roadwork, and parking at home, shops, work or school, public places, train stations, and bus stops. The primary analysis in this paper focuses on the ranking and rating of these attributes along with the respondents' demographic characteristics and their behavior riding in a variety of conditions. The survey in its original Swedish and translated into English is included in Appendices A and B respectively. The questions from the survey analyzed in this thesis are described in Table 2 along with their responses and abbreviations for the questions and responses used throughout.

Table 2- Questions from survey used in study with responses and abbreviations used throughout paper

Question	Response	Abbreviation
Can you generally use a bicycle when you need to?		Bike access
	Yes, always	Always
	Yes, most of the time	Usually
	Yes, sometimes	Sometimes
	No, rarely	Rarely
Do you have driver's license?		Driver's license
	Yes	Yes
	No	No
Can you generally use a car when you need to?		Car access
	Yes, always	Always
	Yes, most of the time	Usually
	Yes, sometimes	Sometimes
	No, rarely	Rarely
	No, never	Never
How often do you ride normally?		How often do you ride?
	5-7 days per week	5-7 days per week
	3-4 days per week	3-4 days per week
	Once/few times per week	at least weekly
	Once/few times per month	at least monthly
	Once/few times a year	at least annually
During which seasons do you use a bicycle?		Season
	Spring	Spring
	Summer	Summer
	Winter	Winter
	Autumn	Fall

Do you avoid cycling in bad weather or conditions?		Riding conditions
	Yes, when it rains	Rain
	Yes, when it snows	Snow
	Yes, in strong winds	Wind
	Yes, when it's slippery	Slippery
	Yes, when it's cold	Cold
	Yes, when it’s poorly shoveled	Poorly shoveled
	Yes, when there is a lot of gravel/leaves	Gravel/leaves
	No, I bike whatever the weather	All weather
What do you think is the most important when choosing your bicycle route?		rank
	Distance	Distance
Rank them so that no one has the same value. (1 being the least important and 10 being the most important)	Time	Time
	Traffic safety	Traffic
	Few crossings	Intersections
	Separation from traffic	Separation
Feel safe: Well-lit roads, no concealing shrubbery, etc.	Consistent network	Consistency
	Smooth and solid surface	Surface
	Safety	Safety
Facilities: Toilets, bicycle pumps, shower / changing rooms, bicycle garage, water, etc	Surroundings / Environment	Environment
	Facilities	Facilities
How satisfied are you with the following in Malmö? (Scale 1-5)		Sat. w/
	Malmö as a cycling city	Satisfaction overall
	The amount of bike lanes	# of cycle paths
	Condition of bike lanes	Condition cycle paths
	Width of bike lanes	Path width
	Diversion at roadwork	Diversion @ roadwork
.(1 dissatisfied, 5 very satisfied)	Bicycle parking at home	Parking @ home
	Bicycle parking at the shops	Parking @ shops
	Bicycle parking at the workplace/school	Parking @ work/school
	Bicycle parking in public places	Parking public
	Bicycle parking at Malmö C, Triangeln and Hyllie	Parking @ stations
	Bicycle parking at bus stops	Parking @ bus stops
How important are the following measures you as a cyclist?		rating
	Increasing security/safety at crossings	Safety @ crossings
	Increase the width of cycle lanes	Cycle lane width
(1 not important, 5 very important)	Snow removal	Snow removal
	Sweep away leaves	Swept leaves
	Sweep away gravel	Swept gravel

	Measures to separate pedestrians and cyclists	Ped cyclist separation
	Remove level differences	Remove level differences
	Minimize the number of stops	Minimize stops
	Smooth surface	Smooth surface
	Keeping the shrubbery/trees	Control plants
	Police checks (eg check for lights, mopeds)	Police checks
What materials do you consider to be suitable as a surface on bike paths?		Surface preferences
	Asphalt	Asphalt
	Gravel	Gravel
(You can select multiple)	Hard packed gravel	Hard gravel
	Tiles	Tiles
	Cobble stone	Cobbles
How safe do you feel as a cyclist in Malmö?		Perception of safety
	Not safe	Not safe
	Partially safe	Partly safe
	Safe	Safe
Are you...?		Gender
	Female	Female
	Male	Male
	Prefer to not answer	No answer
How old are you?		Age
	Under 13 years	0-24
	13-17 years	0-24
	18-24 years	0-24
	25-39 years	25-39
	40-64 years	40-64
	65-75 years	65+
	76-84 years	65+
	85 or older	65+
What is your family situation?		Partner
	Partner I live with	Living with partner
	Single / not cohabitating	Not living with partner
	Living with parents or equivalent	Not living with partner
How many children do you have living at home?		Children
	None	No
	1 child	Yes
	2 children	Yes
	3 children	Yes
	4 or more children	Yes

What is your main occupation?		Employment
	Working	Works a job
	Studying	Student
	On sick leave	Neither
	Parental leave	Neither
	Seeking employment	Neither
	Pensioner	Neither
	Other	Neither
What is your highest completed education?		Education
	Attending primary school	Secondary or lower
	Primary school	Secondary or lower
	Secondary school	Secondary or lower
	Post-secondary education, other than college/university	post-secondary besides college
	College/university	university/college
Income per month:		Income
	Less than 10,000 SEK	<10,000 SEK
	10,000-14,999 SEK	10,000 SEK-14,999 SEK
	15,000-24,999 SEK	15,000 SEK-24,999 SEK
	25,000-34,999 SEK	25,000 SEK-34,999 SEK
	35,000-50,000 SEK	35,000 SEK-50,000 SEK
	More than 50,000 SEK	>50,000 SEK
	Do not wish to answer	No response

3.1.1 Profile of survey respondents

Table 4 shows demographic information about the surveyed sample and the population of the Malmö area at large. Within the survey, a relatively larger portion of the respondents were female than in the general population. 43.1% of survey respondents identified as male, 55.5% identified as female, and 1.4% chose not to disclose their gender. Within the 2015 general population in the Malmö region, 49.5% of residents were male and 50.5% were female.

The survey asked respondents to record their age range within ranges. Because the survey only included cyclists, the sample unsurprisingly included more people in the middle age groups.

Over 85% of the surveyed group were between 25 and 64 years old compared with only 52.4% of the regional population at large, however in both populations the median respondent was in the 25-39 age range. More than 55% of the residents of Malmö's region had children as of 2015, compared with only 35.4% of the surveyed sample.

The survey sample appears to be more educated, more employed, and more highly paid than the regional population at large. While within the survey sample only 8.8% of the respondents had a highest education completion of secondary (high school) or lower, 54.8% of the population at large does not have education beyond secondary. The general population of Malmö's region was 54.4% gainfully employed compared with 71.2% of the survey sample. This is certainly somewhat explained by the larger inclusion of pre and post-working-age individuals in the population at large, however it is important to note the higher employment of the sample from the survey. Accordingly, the median income of the survey sample is higher than the population at large over 20 years of age, and thus likely higher than the population at large of all ages. The median income of the sample group falls between 25,000 and 34,999 Swedish Kroner per month, which is higher than the median income of Sweden, the county of Skåne, Malmö, and Lund. In fact, the 90th percentile of income in Sweden, Skåne, and Malmö falls within that income bracket.

Within the survey sample, 84.5% of individuals had a driver's license, however only 51.9% had access to a car at least sometimes while 31.3% either had access to a car rarely or never. Conceptually, the group without frequent access to a car will be more likely to bike frequently due to their fewer alternatives. Considering that this is a survey of cyclists it is unsurprising that 64.9% always had access to a bike while 97.1% had access to a bike at least usually.

3.1.2 Respondents' willingness to cycle in various conditions

Table 3 depicts the cycling behavior of respondents from the survey in a variety of conditions. From the 905 survey respondents, a vast majority of 76.7% claimed to ride at least 5-7 days per week, while 13.7% claimed to ride 3-4 days per week, and only 9.6% claimed to ride less.

Suitability of surfaces	Yes	%	No	%
Asphalt	895	98.9%	10	1.1%
Gravel	58	6.4%	847	93.6%
Hard gravel	627	69.3%	278	30.7%
Tiles	234	25.9%	671	74.1%
Cobbles	72	8.0%	833	92.0%
Season	**Yes**	**%**	**No**	**%**
Spring	892	98.6%	13	1.4%
Summer	902	99.7%	3	0.3%
Fall	880	97.2%	25	2.8%
Winter	689	76.1%	216	23.9%
Riding conditions	**Yes**	**%**	**No**	**%**
Rain	739	81.7%	166	18.3%
Snow	573	63.3%	332	36.7%
Wind	704	77.8%	201	22.2%
Slippery	406	44.9%	499	55.1%
Cold	805	89.0%	100	11.0%
Poorly shoveled	431	47.6%	474	52.4%
Gravel/leaves	873	96.5%	32	3.5%
All weather	272	30.1%	633	69.9%
Frequency	**#**	**%**		
5-7 days per week	694	76.7%		
3-4 days per week	124	13.7%		
at least weekly	64	7.1%		
at least monthly	19	2.1%		
at least annually	4	0.4%		
Perception of safety	**#**	**%**		
Not safe	55	6.1%		
Partly safe	468	51.7%		
Safe	382	42.2%		

Table 3- Survey respondents' cycling behavior in a variety of conditions, frequency, and perceptions of safety. (n=905)

Descriptive Statistics of Survey Sample		
Gender	**#**	**%**
Male	390	43.1%
Female	502	55.5%
No answer	13	1.4%
Age	**#**	**%**
0-24	74	8.2%
25-39	386	42.7%
40-64	385	42.5%
65+	60	6.6%
Children	**#**	**%**
Yes	320	35.4%
No	583	64.6%
Highest Education	**#**	**%**
Secondary or lower	167	8.8%
post-secondary besides college	1105	58.1%
university/college	630	33.1%
Partner	**#**	**%**
Living with partner	568	63.0%
Not living with partner	333	37.0%
Employment	**#**	**%**
Works a job	642	71.2%
Student	131	14.5%
Neither	129	14.3%
Income	**#**	**%**
<10,000 SEK	107	11.8%
10,000 SEK - 14,999 SEK	93	10.3%
15,000 SEK - 24,999 SEK	108	11.9%
25,000 SEK - 34,999 SEK	320	35.4%
35,000 SEK - 50,000 SEK	142	15.7%
>50,000 SEK	43	4.8%
No response	91	10.1%
Driver's license	**#**	**%**
Yes	765	84.5%
No	140	15.5%
Car access	**#**	**%**
Always	214	23.9%
Usually	251	28.0%
Sometimes	150	16.8%
Rarely	160	17.9%
Never	120	13.4%
Bike access	**#**	**%**
Always	587	64.9%
Usually	291	32.2%
Sometimes	23	2.5%
Rarely	3	0.3%

Descriptive Statistics of General Population		
Gender (2015) *	**#**	**%**
Male	313,028	49.5%
Female	318,845	50.5%
Age (2015)*	**#**	**%**
0-24	193,524	30.6%
25-39	143,397	22.7%
40-64	187,715	29.7%
65+	107,237	17.0%
Children (2015)*	**#**	**%**
Yes	352,818	55.8%
No	260,413	41.2%
Unknown	18,642	3.0%
Highest Education (2015)*	**#**	**%**
Secondary or lower	279,178	54.8%
Post-secondary (< 3 years)	77,952	15.3%
Post-secondary (> 3 years)	137,497	27.0%
unknown	15,044	3.0%
Employment (2014)*	**#**	**%**
Gainfully employed	274,697	54.4%
Not gainfully employed	230,591	45.6%
Income (age>20, 2014)**	**Mean**	**10%**
Sweden	20,952	7,872
Skåne	19,850	6,884
Malmö	18,406	5,058
Lund	21,153	6,387
Income (age>20, 2014)**	**Median**	**90%**
Sweden	17,902	33,592
Skåne	17,001	32,674
Malmö	15,654	30,914
Lund	17,547	36,402

Table 4- Demographic information about survey respondents and the area at large. (n=905)
** Malmö/Lunds/Trelleborgs A-region*
*** Originally published as annual salary, divided by 12 to get monthly*

Accordingly, only 6.1% of the respondents answered that they feel unsafe riding a bicycle in Malmö, while 51.7% replied that they felt partly safe and 42.2% replied that they felt safe. This indicates that the survey sample was made up largely of frequent cyclists, however varying portions of the sample claimed to be deterred by a variety of riding conditions.

Survey respondents indicated differing levels of suitability of various surfaces. Asphalt was unsurprisingly shown to be the most favorable surface with 98.9% of respondents claiming to find it suitable. The survey group also favored hard gravel with 69.3% claiming to find it suitable. Alternatively gravel, cobblestones, and tiles were shown to be very unfavorable with 93.6%, 92.0%, and 74.1% of respondents not finding those surfaces suitable respectively.

While more than 97% of survey respondents rode in each spring, summer, and fall, 23.9% refrain from riding in the winter despite Malmö's relatively mild climate in Scandinavia. Only 30.1% of respondents claimed to ride their bicycle in all surveyed adverse conditions. Conditions that deter the most respondents from riding were slippery paths, poorly shoveled snowy paths, snow, and wind with 55.1%, 52.4%, 36.7%, and 22.2% refraining from riding in the respective conditions. Alternatively, only 3.5% of respondents claimed to avoid riding for gravel and/or leaves in the path and 11% for cold temperatures.

3.1.3 Analysis of respondent satisfaction and preferences

Table 5, Table 6, and Table 7 describe the correlations between the ranked route attributes the survey respondents found most important, their rated satisfaction with aspects of Malmö's cycling network, and their rated perception of importance of bicycle planning priorities. Because each of the sets is based on ranked or ordinal data, they are compared within sets of rankings or ratings by their spearman correlation described in Equation 1. In the correlation calculation, ρ represents the correlation between two items, d_i^2 represents the squared difference in rank between items in their respective sets, and n represents the number of entries being analyzed for correlations. Spearman rank correlation was used for this preliminary analysis because it is non-parametric and doesn't make any assumptions about the data's distribution. It is important to note that while means and standard deviations of ranks and ratings are included in the tables in addition to medians and modes, they are for preliminary analysis and should not be considered to have an equal interval between each level of their levels. In the initial set of 905 respondents, some didn't complete the rankings and ratings and were not included in further analysis. The following tables consider the 644 respondents who completed the ranking and rating tasks and gave at least three unique values in the ranking task shown in Table 5.

Equation 1- Equation for Spearman rank correlation (Laerd Statistics, n.d.)

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Table 5 shows the correlations between rankings of important cycling route attributes and the mean, median, mode, and standard deviation of the ranking for each attribute. Because the table consists of correlations between items' rankings and the rankings of the items they are being ranked against, some negative correlations may be coerced. This occurs because ranks are supposed to occupy all spots between one and ten, so when one item receives a higher or lower ranking, other items are more likely to receive the opposite. Because many respondents didn't view the task as a ranking in the survey and instead treated it as a rating, this effect is probably only partial. Many of the people treating the task as a rating were kept because only 387 observations could have been maintained observing the fully ranked responses.

	distance	time	traffic	intersections	separation	consistency	surface	safety	environment	facilities
Mean	6.15	6.50	7.81	5.59	6.91	6.27	6.20	6.54	5.15	4.23
Median	6	7	8	6	7	6	6	7	5	3
Mode	10	10	10	7	10	6	8	10	2	2
Std. dev.	2.90	2.76	2.31	2.48	2.51	2.55	2.42	2.66	2.59	2.73
Correlations	distance	time	traffic	intersections	separation	consistency	surface	safety	environment	equipment
distance	1.00									
time	0.71	1.00								
traffic	-0.21	-	1.00							
intersections	-0.12	-	0.25	1.00						
separation	-0.24	-	0.38	0.34	1.00					
consistency	-0.17	-	0.07	0.29	0.32	1.00				
surface	-0.13	-	-0.01	0.18	0.12	0.40	1.00			
safety	-0.20	-	0.33	0.05	0.22	0.09	0.19	1.00		
environment	-0.15	-	0.06	0.05	0.08	0.12	0.15	0.31	1.00	
facilities	-0.17	-	0.03	0.17	0.12	0.17	0.30	0.23	0.35	1.00

Table 5 - Survey responses of important route attributes with correlations between the rankings of each attribute, n=644. On a scale from 1= least important to 10= most important. All respondents selected ranked importance, although not all respondents gave complete ranking.

Despite the partial coercion of rankings, the correlations between rankings givers insight to the connection of route attributes' ranks to one another. In the ranking of route attributes, traffic appears to be the most important factor with the highest median and mean ranking, followed by separation, time, and safety. It is important to note that the most common response for importance of distance was the top rating of 10 with a median response of only six, and the highest standard deviation of any of the rankings considered. This indicates that some people value distance really highly while others don't value it nearly as highly as other attributes. Alternatively, separation, consistency, and safety each have similar median and mode responses indicating a more normally distributed set of preferences for those attributes with more similar numbers of people valuing them more and less than the mode.

There are a few discernible trends in the correlation between rankings of the route attributes. Distance and time have a strong correlation of .71, indicating that they tend be similarly important to individuals. All other route attributes are negatively correlated with distance and time, indicating that when distance and time are ranked highly other attributes tend to rank lowly and vice versa. This is certainly somewhat caused by the correlation of rankings forcing negative correlations. The strongest negative correlations with route distance and time are with separation from vehicle traffic, safety, and traffic, which may indicate a separation between cyclists who prefer fast routes and cyclists who prefer safe and comfortable routes. Additionally separation from vehicle traffic is unsurprisingly positively correlated with traffic and intersection as both attributes address cyclists' interactions with motor vehicles. Surface and consistency are also moderately positively correlated, which makes sense as they both address comfort and convenience of cycling. Finally, facilities and environment are moderately positively correlated, which may indicate some underlying preference for enjoyment of a route.

Table 6 describes the satisfaction section of the survey where respondents rated their satisfaction with different aspects of Malmö's bicycle network on a scale from 1 (dissatisfied) to 5 (very satisfied). The table gives the mean, median, mode, and standard deviation of cyclists' satisfaction ratings for each of the questions and then shows the correlations between them. Note that for each of the aspects of the cycling network being rated besides parking at home and path width, the median and mode response are the same, suggesting that the responses are dispersed at least somewhat evenly around the mode. Respondents were most satisfied with the cycling network overall, the number and condition of cycle paths, and parking at train stations and work/school. They were least satisfied with the parking at bus stops and diversions at road work.

	overall	# of cycle paths	condition cycle paths	path width	diversion @ roadwork	parking @ home	parking @ shops	parking @ work/ school	parking public	parking @ train stations	parking @ bus stops
Mean	4.02	3.67	3.6	3.38	2.62	3.4	2.94	3.54	3.1	3.61	2.58
Median	4	4	4	3	3	4	3	4	3	4	3
Mode	4	4	4	4	3	5	3	4	3	4	3
Std. dev.	0.78	0.86	0.91	1.02	1.08	1.30	1.10	1.19	1.04	1.17	1.03
Correlations	overall	# of cycle paths	condition cycle paths	path width	diversion @ roadwork	parking @ home	parking @ shops	parking @ work/ school	parking public	parking @ train stations	parking @ bus stops
overall	1.00										
# of cycle paths	0.64	1.00									
condition cycle paths	0.35	0.35	1.00								
path width	0.32	0.33	0.50	1.00							
diversion @ roadwork	0.30	0.26	0.36	0.32	1.00						
parking @ home	0.18	0.23	0.10	0.17	0.12	1.00					
parking @ shops	0.21	0.18	0.16	0.20	0.23	0.48	1.00				
parking @ work/ school	0.23	0.22	0.13	0.12	0.16	0.38	0.48	1.00			
parking public	0.21	0.21	0.22	0.26	0.19	0.21	0.48	0.47	1.00		
parking @ train stations	0.17	0.17	0.18	0.17	0.16	0.09	0.27	0.29	0.44	1.00	
parking @ bus stops	0.23	0.21	0.22	0.22	0.27	0.24	0.41	0.32	0.50	0.35	1.00

Table 6- Survey responses of satisfaction with aspects of Malmö as a city for bicycles. On a scale from 1= least important to 5= most important.(n=644)

Note that all of the correlations shown are positive, indicating that rating each aspect of the cycling network highly or lowly somewhat correlates with rating other aspects of the cycling network the same way. This may be partially because some individuals tend to give better or worse ratings at the same level of satisfaction, but it also may indicate a generally positive trending of attitudes between aspects of the bicycle network. Overall satisfaction is strongly correlated with the number of cycle paths, but is also moderately correlated with the condition and with of cycle paths. Satisfaction with the condition of cycling paths is well correlated with cycling path width and diversions at roadwork. Finally, satisfactions with parking at all of the locations mentioned are considerably correlated with one another with the exceptions of parking at home with parking at train stations.

	Safety @ crossings	Cycle lane width	Snow removal	Swept leaves	Swept gravel	Ped cyclist separation	Remove level differences	Minimize stops	Smooth surface	Control plants	Police checks
Mean	4.48	3.36	4.21	3.24	3.32	4.09	3.61	3.78	3.78	3.33	2.59
Median	5	3	5	3	3	4	4	4	4	3	3
Mode	5	3	5	3	3	5	3	4	4	3	3
Std. dev.	0.78	1.08	0.97	1.11	1.10	0.94	1.12	1.07	0.99	1.10	1.29
Correlations	Safety @ crossings	Cycle lane width	Snow removal	Swept leaves	Swept gravel	Pedestrian cyclist separation	Removal level differences	Minimize stops	Smooth surface	Control plants	Police checks
Safety @ crossings	1.00										
Cycle lane width	0.19	1.00									
Snow removal	0.14	0.08	1.00								
Swept leaves	0.08	0.22	0.39	1.00							
Swept gravel	0.09	0.21	0.30	0.72	1.00						
Ped cyclist separation	0.09	0.33	0.14	0.23	0.22	1.00					
Remove level differences	0.14	0.15	0.15	0.19	0.22	0.25	1.00				
Minimize stops	0.00	0.10	0.03	0.02	0.07	0.09	0.40	1.00			
Smooth surface	0.03	0.14	0.17	0.27	0.31	0.14	0.47	0.35	1.00		
Control plants	0.14	0.21	0.28	0.32	0.32	0.21	0.27	0.07	0.36	1.00	
Police checks	0.09	0.19	0.12	0.19	0.16	0.18	0.09	-0.09	0.11	0.30	1.00

Table 7- Survey responses of ratings of importance of aspects that make routes bicycle friendly. On a scale from 1= least important to 5= most important.(n=644)

Table 7 describes ratings on a scale from 1(not important) to 5 (very important) of survey respondents' attitudes about the importance of various cycling route attributes. The table includes the mean, median, mode, and standard deviation of responses for each attribute as well as the Spearman correlation between each of the attributes. Similarly to in Table 6, the median and mode response for most route attributes were the same with the exception of pedestrian-cyclist separation and removal of level differences, which each had a mode one point lower than their median. This again demonstrates that the responses were generally distributed at least somewhat evenly around the mode. The attributes with the highest rated importance were safety at crossings, snow removal, and pedestrian-cyclist separation. Alternatively the attribute with the lowest ratings of importance was police checks for proper use of cycling infrastructure.

Similarly to in Table 6, there were no strongly negative correlations included in the ratings, indicating that higher ratings for the importance of each attribute correlate with generally higher ratings for all other attributes. Predictably snow removal rating is correlated considerably with swept leaves, swept gravel, and controlled plants, as all of those attributes deal with clearing the path of obstructions. Similarly, swept leaves and swept gravel are substantially correlated with smooth surface and controlling plants. Additionally, removal of level differences, minimization stops, and smooth riding surface are all are all moderately correlated, which makes sense as they each address aspects of comfort and ease of riding, and thus could all be valued similarly.

After analyzing the survey responses and correlations of those responses within groups of questions, it is possible to develop a model that explains the connections between diverse individuals’ cycling behavior, route preferences, and satisfaction with cycling infrastructure in Malmö.

4 Categorization of cyclists

Categorization of riding behavior, route preferences, and satisfaction allows planners to develop a better understanding of cyclists' diverse preferences and travel patterns within more homogeneous groups. Within such sets, planners can target improvements to achieve a variety of goals that can be differently achieved for members within each latent group. These goals may include measures to better serve currently frequent cyclists, to encourage them to cycle more frequently, or to bring more cyclists into the fold. Section 1.3.1 describes the background of segmenting cyclists more in-depth, and this chapter will explain and implement a cyclist segmentation model. The model presented here consists of two major steps including the identification of principal components of cyclists based on 46 of the questions of the survey, and then the clustering of individuals from their coordinates on those components.

Similar methods have been used previously in transportation research. Krizek & El-Geneidy (2007) conducted surveys of public transit riders and non-riders separately, then completed a factor analysis based on responses to each of those surveys to find separate hierarchical clusters of both riders and non-riders. (Krizek & El-Geneidy, 2007) Gatersleben and Haddad (2007) conducted a survey of 244 cyclists and non-cyclists in England. They used principal component analysis (PCA) to separate respondents' perceptions of typical cyclists into four components including responsible, lifestyle, commuter, and hippy-go-lucky. They used ANOVA to determine that these component scores were different between cyclists and non-cyclists, then used regression analysis to explain how the principal components and past cycling behavior can explain respondents' future intention to travel by bicycle. They found that cycling in the past has a strong positive relationship with cycling in the future and that respondents who see typical cyclists as hippy-go-lucky or commuters were more likely to convey intention to cycle in the future. (Gatersleben & Haddad, 2010) Damant-Sirois et al. (2014) similarly used PCA to combine correlated survey questions into less correlated components that describe survey respondents' perceptions toward cycling. They ended up combining their 35 variables into 7 components, and then clustering respondents on those components into four groups. Their groups included dedicated cyclists, path-using cyclists, fairweather utilitarians, and leisure cyclists. (Damant-Sirois, et al., 2014) This method is the most similar to what this section of the paper will present to identify areas of sufficient and deficient infrastructure for each group within the heterogeneous population of cyclists.

This method for categorizing questions into components based on a correlation matrix of responses, then categorizing respondents from the distance between coordinates of their answers on the components, is effective because each step provides an explanatory framework to better understand how cyclists differ between groups. The grouping of questions through PCA and related methods develops links between responses to each of the questions, which helps understand connections between different attribute preferences, cycling behaviors and satisfactions with aspects of the bicycle network. Then the categorization of cyclists on those components explains the broad interests and preferences of each group.

4.1 Factor analysis of mixed data (FAMD) grouping of survey questions

PCA is a statistical method meant to extract meaningful information from complex and highly dimensional data in a way that generates a new set of principal components with fewer dimensions. The principal components are meant to explain as much of the variation of the dataset as possible while remaining orthogonal (and thus uncorrelated) with one another. The original data is converted into scores for each of the principal components through linear combinations with values of each of the original variables for each observation. (Abdi & Williams, 2010) Because the dataset from the survey considered in this paper includes a combination of binary variables, Likert Scale ratings, and rankings of the importance of route attributes, it was necessary to choose an analysis method that addresses correlation and scaling of the different types of data considered.

Factor analysis of mixed data (FAMD) enables the consideration of multiple data types by mixing elements of PCA for continuous data and multiple correspondence analysis (MCA) for categorical data. In this paper the FAMD function in the R Statistics Software package FactoMineR was used to reduce the number of dimensions in the data set. The FAMD function scales continuous data to unit variance while creating a disjunctive data table, which decomposes categorical variables into a separate variable for each category, with the categorical variables and then scaling them with MCA. This allows both types of data to be considered together to study the multi-variate similarities between individuals. (Husson, et al., 2016) To preserve the ordinal nature of all of the Likert and ranking data in this data set, all of the ordinally scaled data were treated as continuous, which makes the assumption that there is a latent underlying continuous variable that each of the scales lies upon. The following paragraphs describe the functionality of these scaling and dimension reduction techniques.

For data being treated as continuous, FAMD scales responses to unit variance to remove units from the variables and give each variable the same mean. Each individual standardized score x_{ik} is found by subtracting the mean $\bar{y}_k$ of each variable k from each individual i's value for that variable y_{ik}, then dividing by the standard deviation of that variable s_k. (Jackson, 1991) This is demonstrated in Equation 2.

Equation 2- Scaling continuous responses to unit variance

$$x_{ik} = (y_{ik} - \bar{y})/s_k$$

Alternatively for categorical variables, responses are put into a disjunctive data table then scaled using the specific scaling of MCA. This allows variables occurring less frequently to have larger weights and thus have a larger impact on the relationship matrix described in the next paragraph. Disjunctive data tables are made by decomposing each categorical column of responses into separate columns k for each of their responses. Those columns are filled in with a one for the column k corresponding with that individual i's value, which is represented by y_{ik} and a zero for the column(s) k corresponding with different values. In this data set, the disjunctive data table only has two columns for each categorical variable because they are all binary, asking whether respondents ride their bicycle in certain conditions. P_k represents the proportion of individuals with a one in column k and the scaled value for each individual's scaled score x_{ik} is demonstrated in Equation 3. (Pagès, 2014)

Equation 3- Specific scaling of MCA for categorical variables

$$x_{ik} = \frac{y_{ik}}{p_k} - 1$$

Using the scaled scores from Equation 2 for continuous variables and Equation 3 for categorical variables, a relationship matrix is built with different relationship coefficient calculations between two continuous variables, between two categorical variables, and between a continuous and a categorical variable. In the case of two continuous variables the squared Pearson's correlation coefficient, represented by r^2, is used. Pearson's correlation is the covariance between the two variables divided by the product of their standard deviations. In the case of a continuous and categorical variable, the squared correlation ratio is used. That is represented by η^2 and is a coefficient of non-linear association that compares the weighted variance of the mean of the continuous variable for each category of the categorical variable against the variance across all categories. Finally, when comparing two categorical variables, the statistic φ^2, which is equal to the chi-squared test for independence between the each of the variable's categories divided by the number of individuals, is used. (Pagès, 2014) With such a relationship matrix constructed it is possible to perform FAMD.

FAMD finds the underlying structures of the dataset by rotating the data to find each successive principal component that describes as much of the variation as possible while remaining orthogonal and thus uncorrelated with each of the other principal components. This is accomplished in the case of the first component by finding the axis around which there is the least possible orthogonal error, and thus the least amount of information in the dataset that the component does not describe. Successive components are generated by the same strategy with the condition that they are uncorrelated with each of the previous components. This process continues until there are as many principal components as in the original dataset and all of the variation in the dataset is described. Component scores for individuals can be described as linear combinations of the original variables, enabling descriptions of individual respondents on the principal components, with fewer variables that are less correlated than in the original data. These coefficients are multiplied by the scaled values of each of the variables to find these component scores or the individuals. Because the scores are scaled to have a mean of zero for each of the original variables, the mean of the component scores for the entire survey sample very close to zero.(Holland, 2008)

Figure 5 shows a scree plot of the percentage of variance from the original dataset described by each successive component through the fifteenth from FAMD. Although the described variance per component drops substantially after the fifth and sixth components, the eigenvalues for each of the components through the 14^{th} are greater than one, meaning they describe more variance than a single variable from the original dataset. After testing alignments with between six and 11 components, the eight-component alignment was chosen for further analysis. With eight components, variables from the original dataset strongly influencing each component were grouped in a way that most intuitively explains the underlying factors affecting cyclists' survey responses. Those eight components describe 46% of the variance from the 46 variables they refer to. Beyond the eighth component the coefficients for each variable and component do not group in a way that is useful for considering people's underlying preferences for and satisfaction with the cycling network.

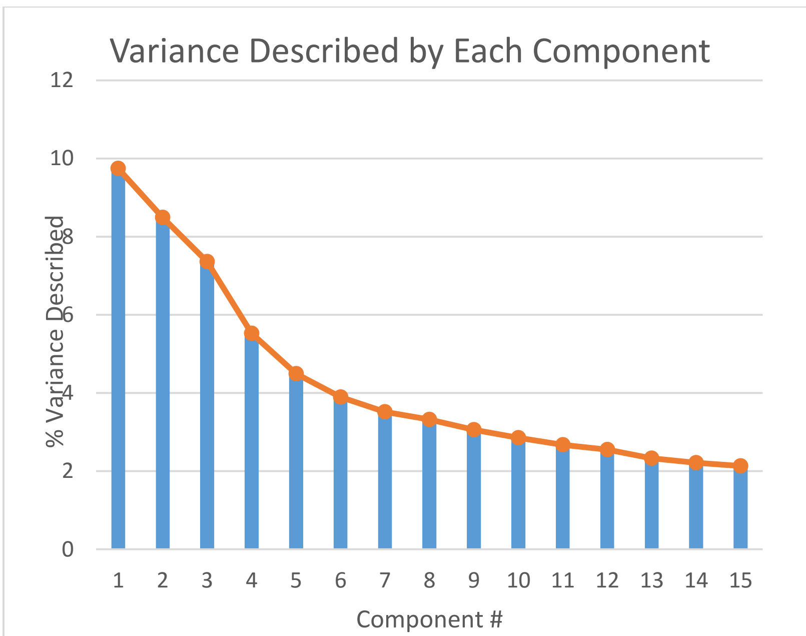


Figure 5- Scree plot of the portion of variance described by each successive component

Table 8 shows the coefficients for each of the variables from the original dataset used to find each survey respondent's score on each of the components as well as the contribution of each of the variables to those components in brackets. The variables are grouped by the components for which they have the largest contribution and ordered by the size of their contributions to those components. The contribution of each variable to each of the components indicates the variable's contribution in determining that component. It is a percentage, meaning that the sum of the contributions to each component is 100. The contributions are determined by dividing the squared cosine between a variable and a component by the summed squared cosines between the component and all variables, then multiplying by 100. Because the cosine approaches zero for large angles and approaches one for small angles, the more correlated the variable is to a component, the larger its contribution is determined to be. (Kassambara, 2015)

The coordinates of the variables on the principal components describe their coefficients in creating linear combinations of values to generate component scores for each of the individuals in the dataset. In other words, to generate the component scores for each individual, their scaled response value for each variable is multiplied by the corresponding coordinate for each component, and then summed across all original variables to find each component score.

Because the binary variables are each asking whether respondents ride under different conditions, they are decomposed into the condition (e.g. winter, hard gravel, slippery, etc.) and yes or no corresponding with each response. The eight components each correspond with identifiable aspects of the respondents' attributes relating to their preferences, cycling behavior, and satisfaction, and are broken down in Table 9.

Because so many of the variables presented address satisfaction and the avoidance of adverse conditions, those variables tend to describe the first and second components. While this indicates the clustering that in section 4.2 will substantially be based upon those responses, the other sources of variation are also suitably addressed and receive equal weight in the clustering as the first and second components. This is in effect applying weight to the lower components relative to the proportion of variance within the original dataset they describe, and thus removing some of the weight from the satisfaction and surface variables.

Satisfaction	C1	C2	C3	C4	C5	C6	C7	C8
sat. w/ path width	**0.6 [8]**	0.0 [0]	0.1 [0.3]	0.0 [0]	-0.2 [1.3]	0.1 [0.4]	-0.3 [4.6]	0.3 [6.1]
sat. w/ cycle tracks	**0.6 [7.5]**	0.1 [0.3]	0.1 [0.1]	-0.1 [0.3]	-0.2 [2.1]	0.2 [1.9]	0.0 [0]	0.3 [4.1]
satisfaction overall	**0.6 [7.1]**	0.0 [0]	0.3 [2.4]	0.0 [0]	-0.2 [2.8]	0.3 [4.4]	-0.2 [2.2]	0.2 [2.3]
sat. w/ parking public	**0.5 [6.5]**	0.1 [0.2]	0.4 [4]	0.1 [0.5]	0.3 [3.6]	-0.3 [3.6]	0.1 [1.1]	0.0 [0.1]
sat. w/ parking @ bus stops	**0.5 [6]**	0.1 [0.1]	0.3 [3.2]	0.0 [0]	0.2 [2.6]	-0.2 [2.1]	0.1 [0.9]	-0.1 [0.3]
sat. w/ parking @ shops	**0.5 [5.8]**	0.1 [0.2]	0.4 [4.5]	0.0 [0]	0.3 [4.3]	-0.3 [4.2]	0.1 [0.3]	-0.2 [2.9]
sat. w/ bike lanes	**0.5 [5.8]**	0.0 [0]	0.4 [3.8]	0.0 [0]	-0.3 [3.5]	0.3 [4]	-0.2 [2]	0.1 [1]
sat. w/ roadwork diversion	**0.5 [5]**	0.1 [0.5]	0.2 [1.4]	-0.1 [0.1]	-0.1 [0.3]	0.2 [2.3]	-0.1 [0.2]	0.0 [0]
sat. w/ parking @ stations	**0.4 [4]**	0.0 [0]	0.3 [1.9]	0.0 [0]	0.2 [2.3]	-0.2 [1.9]	0.2 [1.4]	0.1 [0.4]
Avoidance of adverse conditions	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
all weather yes	0.4 [0.3]	**-2.1 [9]**	-0.3 [0.2]	-0.5 [1.1]	0.4 [1.3]	-0.1 [0]	0.1 [0]	0.1 [0.1]
winter no	-0.6 [0.4]	**2.4 [8.9]**	0.3 [0.2]	0.5 [0.8]	0.3 [0.5]	0.0 [0]	0.1 [0]	0.0 [0]
snow no	-0.3 [0.2]	**1.9 [8.7]**	0.2 [0.1]	0.3 [0.5]	0.0 [0]	0.1 [0.1]	-0.2 [0.3]	0.1 [0.1]
cold no	-0.8 [0.3]	**3.2 [6.9]**	0.4 [0.2]	0.6 [0.6]	1.0 [2.6]	0.4 [0.4]	0.2 [0.2]	-0.1 [0]
wind no	-0.3 [0.1]	**2.1 [6.1]**	0.3 [0.2]	0.5 [0.9]	0.3 [0.5]	0.3 [0.6]	0.1 [0]	-0.1 [0.1]
slippery yes	0.4 [0.3]	**-1.4 [5.8]**	-0.2 [0.2]	-0.4 [0.9]	0.4 [1.8]	0.1 [0.1]	0.0 [0]	0.0 [0]
rain no	-0.5 [0.3]	**2.1 [5.3]**	0.1 [0]	0.7 [1.2]	0.8 [2.7]	0.5 [1.2]	-0.1 [0]	0.2 [0.4]
snow yes	0.2 [0.1]	**-1.1 [4.8]**	-0.1 [0]	-0.2 [0.3]	0.0 [0]	0.0 [0]	0.1 [0.2]	-0.1 [0.1]
slippery no	-0.3 [0.2]	**1.2 [4.8]**	0.2 [0.1]	0.3 [0.7]	-0.3 [1.5]	-0.1 [0.1]	0.0 [0]	0.0 [0]
all weather no	-0.2 [0.1]	**0.9 [4]**	0.1 [0.1]	0.2 [0.5]	-0.2 [0.6]	0.0 [0]	0.0 [0]	0.0 [0]
poorly shoveled yes	0.4 [0.3]	**-1.1 [4]**	-0.2 [0.2]	-0.3 [0.6]	0.4 [1.9]	0.0 [0]	0.1 [0.1]	0.0 [0]
poorly shoveled no	-0.4 [0.3]	**1.1 [3.8]**	0.2 [0.2]	0.3 [0.6]	-0.4 [1.8]	0.0 [0]	-0.1 [0.1]	0.0 [0]
winter yes	0.2 [0.1]	**-0.7 [2.6]**	-0.1 [0]	-0.1 [0.2]	-0.1 [0.1]	0.0 [0]	0.0 [0]	0.0 [0]
gravel/leaves no	-1.9 [0.6]	**3.2 [2.3]**	2.0 [1.2]	0.2 [0]	0.9 [0.6]	0.8 [0.7]	-0.1 [0]	-0.3 [0.1]
wind yes	0.1 [0]	**-0.6 [1.7]**	-0.1 [0]	-0.1 [0.2]	-0.1 [0.1]	-0.1 [0.2]	0.0 [0]	0.0 [0]
rain yes	0.1 [0.1]	**-0.5 [1.1]**	0.0 [0]	-0.1 [0.3]	-0.2 [0.6]	-0.1 [0.3]	0.0 [0]	-0.1 [0.1]
cold yes	0.1 [0]	**-0.4 [0.8]**	-0.1 [0]	-0.1 [0.1]	-0.1 [0.3]	0.0 [0.1]	0.0 [0]	0.0 [0]
gravel/leaves yes	0.1 [0]	**-0.1 [0.1]**	-0.1 [0]	0.0 [0]	0.0 [0]	0.0 [0]	0.0 [0]	0.0 [0]
Importance of orderly network	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
rating control plants	-0.3 [2.1]	-0.2 [0.6]	**0.5 [7.1]**	0.2 [1.2]	-0.2 [2.7]	0.0 [0.1]	0.0 [0]	-0.1 [0.2]
rating sweep leaves	-0.2 [1.1]	-0.2 [1.2]	**0.5 [6.9]**	0.3 [3.2]	-0.3 [4.6]	0.1 [0.8]	0.1 [1.3]	-0.1 [0.5]
rating sweep dirt	-0.2 [1.3]	-0.2 [1.3]	**0.5 [6.9]**	0.3 [3.9]	-0.2 [1.7]	0.1 [0.7]	0.1 [0.7]	-0.1 [0.2]
rating police checks	-0.2 [0.5]	0.0 [0]	**0.5 [5.9]**	-0.1 [0.1]	-0.2 [1.1]	-0.1 [0.3]	0.1 [0.8]	-0.1 [1]
rating smooth surface	-0.3 [1.7]	-0.3 [2.6]	**0.4 [4.9]**	0.3 [4.2]	0.2 [1.3]	0.2 [1.3]	-0.1 [0.9]	0.1 [0.1]
sat. w/ parking @ home	0.4 [2.8]	0.1 [0.2]	**0.3 [3.4]**	0.0 [0]	0.1 [1]	-0.2 [2.2]	-0.1 [0.4]	-0.2 [2.3]
rating ped/bike separation	-0.3 [2.4]	-0.2 [0.7]	**0.3 [2.9]**	0.0 [0.1]	-0.1 [0.3]	0.1 [0.5]	0.1 [1]	0.1 [0.2]

Table 8- Component coefficients and contributions for each variable from the original dataset

I. *Numbers outside brackets represent variable coordinates, numbers inside brackets represent variable contributions*

II. *Sat. w/ abbreviates satisfaction with referring to the questions described in* Table 6. *Rank refers to the ranked questions described in* Table 5. *Rating refers to the rating of route attributes described in* Table 7

III. *For binary variables yes/no indicate the score for the corresponding response*

Importance of speed	C1	C2	C3	C4	C5	C6	C7	C8
rank distance	0.2 [1]	-0.1 [0.1]	-0.2 [1.5]	**0.6 [13.4]**	0.0 [0]	0.0 [0]	0.3 [5]	0.2 [2.5]
rank time	0.2 [1]	-0.1 [0.2]	-0.2 [1.4]	**0.6 [12.6]**	0.0 [0]	0.0 [0]	0.3 [5.5]	0.2 [2.3]
rank safety	-0.2 [0.6]	0.1 [0.3]	0.4 [4.7]	**-0.4 [5.4]**	-0.2 [1.5]	-0.1 [0.2]	-0.1 [0.3]	-0.2 [1.7]
Importance of riding continuity	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
rating snow removal	0.0 [0]	-0.4 [3.5]	0.3 [2.6]	0.1 [0.8]	**-0.4 [8.4]**	-0.1 [0.1]	0.1 [0.2]	0.0 [0.1]
rating minimize stops	-0.2 [0.6]	-0.3 [2]	0.1 [0.2]	0.3 [2.8]	**0.4 [7.9]**	0.1 [0.2]	0.1 [0.4]	0.3 [7.7]
Importance of enjoyment	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
rank traffic	-0.2 [0.5]	0.1 [0.5]	0.2 [1.5]	-0.3 [4.4]	-0.3 [4.2]	**-0.4 [10.6]**	0.0 [0]	0.3 [4]
perception safety	0.6 [6.7]	-0.1 [0.2]	0.1 [0.3]	0.1 [0.4]	0.0 [0]	**0.4 [8.8]**	-0.1 [1.2]	0.1 [0.9]
rank environment	-0.1 [0.1]	0.1 [0.1]	0.2 [1.5]	-0.4 [5.5]	0.1 [0.2]	**0.4 [8.1]**	0.2 [3.4]	-0.3 [5]
rank equipment	0.0 [0]	0.0 [0]	0.3 [2.3]	-0.4 [4.7]	0.2 [2.8]	**0.4 [7.3]**	0.0 [0]	-0.3 [4]
sat. w/ parking work/school	0.5 [4.5]	0.1 [0.1]	0.4 [4.2]	0.1 [0.1]	0.2 [2.1]	**-0.4 [6.8]**	0.1 [0.8]	-0.1 [1.3]
rating safety @ crossings	-0.2 [0.7]	0.0 [0]	0.2 [1.5]	-0.1 [0.5]	-0.3 [3]	**-0.3 [6.4]**	0.1 [1]	0.3 [4.4]
rank surface	-0.3 [1.5]	-0.1 [0.5]	0.4 [3.9]	-0.1 [0.4]	0.3 [5]	**0.3 [6]**	-0.1 [0.8]	0.0 [0.1]
Suitability of surfaces	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
gravel yes	1.4 [0.6]	0.5 [0.1]	-0.3 [0]	-1.4 [2.1]	-0.6 [0.5]	0.9 [1.6]	**2.4 [14.4]**	0.4 [0.4]
cobblestones yes	1.1 [0.5]	0.7 [0.3]	-0.9 [0.5]	-0.7 [0.5]	-0.5 [0.5]	1.1 [2.9]	**2.0 [12.1]**	0.1 [0.1]
rating cycle lane width	-0.4 [3]	-0.1 [0.2]	0.3 [3.1]	0.0 [0]	0.0 [0]	0.0 [0]	**0.4 [10.5]**	-0.1 [0.2]
hard gravel yes	0.6 [0.5]	0.5 [0.5]	-0.4 [0.5]	-0.4 [0.9]	-0.4 [1]	0.2 [0.3]	**0.8 [8.8]**	0.2 [0.8]
tiles yes	0.8 [0.8]	0.3 [0.1]	-0.3 [0.2]	-0.3 [0.4]	-0.3 [0.4]	0.3 [0.5]	**0.8 [6.5]**	-0.2 [0.4]
hard gravel no	-0.3 [0.2]	-0.2 [0.2]	0.2 [0.2]	0.2 [0.4]	0.2 [0.5]	-0.1 [0.1]	**-0.4 [4.3]**	-0.1 [0.4]
tiles no	-0.3 [0.3]	-0.1 [0]	0.1 [0.1]	0.1 [0.2]	0.1 [0.1]	-0.1 [0.2]	**-0.3 [2.3]**	0.1 [0.2]
gravel no	-0.1 [0]	0.0 [0]	0.0 [0]	0.1 [0.2]	0.0 [0]	-0.1 [0.1]	**-0.2 [1.1]**	0.0 [0]
cobblestones no	-0.1 [0]	-0.1 [0]	0.1 [0]	0.1 [0]	0.0 [0]	-0.1 [0.2]	**-0.2 [1]**	0.0 [0]
Importance of safety	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
rank intersections	-0.2 [1.3]	0.1 [0.2]	0.2 [0.9]	-0.3 [3.7]	0.3 [3.1]	-0.1 [0.3]	0.1 [0.7]	**0.5 [14.5]**
rank separation	-0.2 [0.5]	0.2 [1]	0.3 [2]	-0.5 [8.9]	0.0 [0]	-0.2 [2.2]	-0.1 [0.3]	**0.4 [11.7]**
rank network consistency	-0.3 [1.7]	0.0 [0]	0.2 [0.9]	-0.3 [3.4]	0.3 [4.8]	0.2 [2.5]	0.1 [0.3]	**0.4 [9.2]**
rating remove level differences	-0.3 [1.8]	-0.3 [1.6]	0.3 [3.3]	0.4 [4.9]	0.2 [1.3]	0.0 [0]	0.0 [0]	**0.3 [5.2]**

Table 8 continued-Component coefficients and contributions for each variable from the original dataset

I. Numbers outside brackets represent variable coordinates, numbers inside brackets represent variable contributions

II. Sat. w/ abbreviates satisfaction with referring to the questions described in Table 6. Rank refers to the ranked questions described in Table 5. Rating refers to the rating of route attributes described in Table 7

III. For binary variables yes/no indicate the score for the corresponding response

Components	Description of component
1: Satisfaction 9.74% of variance	Higher for cyclists who are more satisfied with the cycling network and parking, perceive themselves to be safe, and ride in more conditions. Lower for individuals who value an orderly network.
2: Avoidance of adverse conditions 8.49% of variance	Higher for cyclists who do not ride in various weather and surface conditions, lower for those who do.
3: Importance of an orderly network 7.36% of variance	Higher for cyclists who value controlling bicycle path conditions including surrounding plants, leaves, and dirt, and for those who value police checks of path use, smooth surfaces, separation from vehicles, and parking at home. Higher for satisfied individuals and those who feel safe and lower for those who value speed.
4: Importance of speed 5.52% of variance	Higher for cyclists who value route distance, time, removal of level differences, surface, and an orderly bicycle network. Lower for those who value safety and recreational aspects of cycling.
5: Importance of network continuity 4.49% of variance	Higher for cyclists who value minimizing stops, network consistency and surface. Lower for cyclists who value snow removal and ride in the cold. Snow removal may not be important to this variable because respondents scoring highly in this component are less likely to ride in the cold.
6: Importance of enjoyment 3.89% of variance	Higher for cyclists who perceive cycling as safe and value route environment, facilities, and surface. Higher for those who ride on more surfaces. Lower for those who value traffic conditions and safety at crossings. Also lower for cyclists who are satisfied with parking at work or school.
7: Suitability of surfaces 3.52% of variance	Higher for cyclists who find all surfaces to be suitable, and value distance, time and cycle lane width. Lower for those who don't find the surfaces included suitable and those who are satisfied with cycling path width.
8: Importance of Safety 3.32% of variance	Higher for cyclists who value intersections, separation from vehicle traffic, bicycle network consistency, safety at crossings, traffic, distance, and time, and those who are satisfied with bicycle network attributes. Lower for those who are satisfied with parking and are concerned about the environment and equipment.

Table 9- Names and profiles of the eight components

4.2 Hierarchical clustering of cyclists

After using FAMD in section 4.1 to break down the dataset into comprehensible components that explain variation between individuals' survey responses, it is possible to cluster individuals on those components to identify more homogenous clusters than the total sample of cyclists. To create hierarchical clusters this paper uses Ward's method, which is an agglomerative clustering method. Agglomerative clustering starts with each member of a sample in their own individual clusters, combining clusters successively until they are all clustered together in a single tree. Ward's method uses a distance metric between points in a cluster and the cluster's centroid. The summation of the square of each of those distance is a metric for the variance within the cluster, and at each stage of ward's method the two clusters are joined that will result in the minimum total within-cluster variance for all clusters until the joining of the final two clusters. (Batagelj, 1988)

To determine distances between clusters and their centroids, Manhattan distance was used and is described in Equation 4. Manhattan distance is the aggregate distance between two variables on all of the axes considered individually. Equation 4 describes the distance between the points P1 and P2 which lie in an N dimensional plain with axes X to N using Manhattan distance. (Black, 2006)

Equation 4- Manhattan distance between points P1 and P2 in N dimensions

$$|X_1 - X_2| + \cdots + |N_1 - N_2|$$

Clustering the original observations of the 651 individuals using Ward's method and measured with Manhattan distance yields the dendrogram in Figure 6. The dendrogram is shown split into the six clusters considered in further analysis. A Six cluster alignment was chosen because it enables description of the most homogeneous clusters of cyclists while still maintaining sufficient group sizes in each cluster. The heights on the Y-axis of the dendrogram represent the sum-of-squares of the Manhattan distances between the points of a cluster and its centroid, showing the variation within each cluster where it is formed.

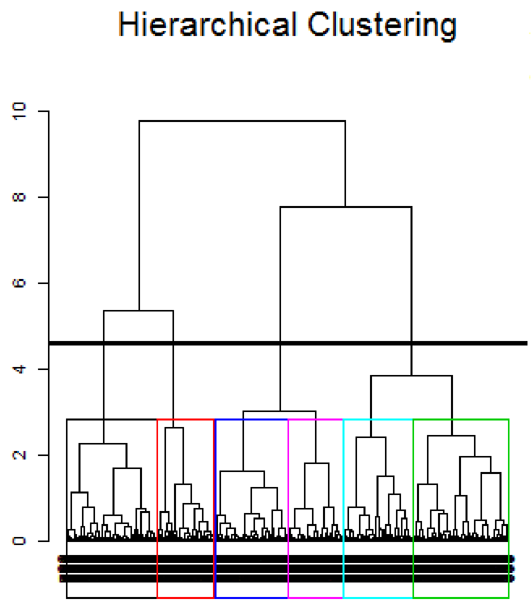


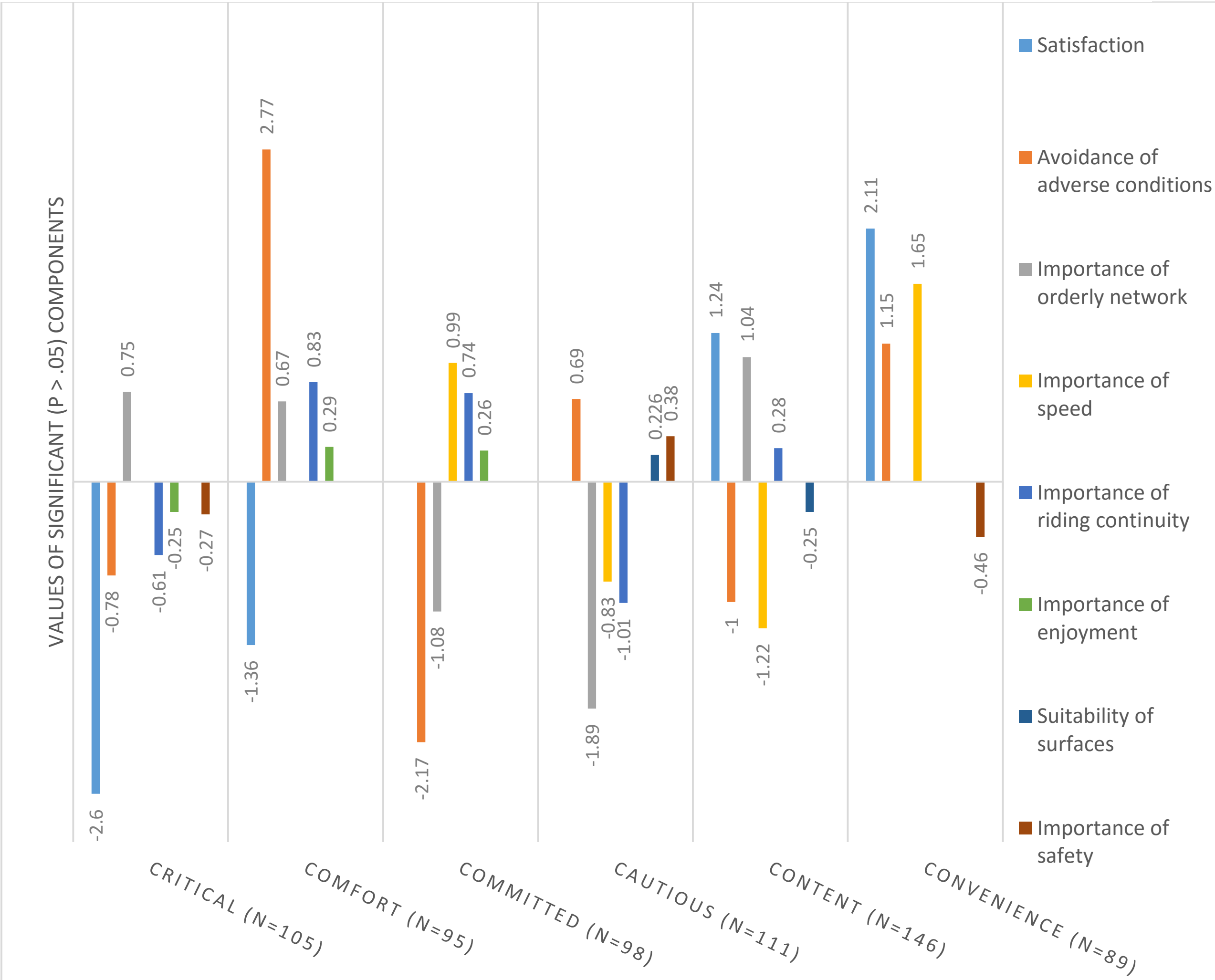
Figure 6- Dendrogram of the hierarchical clustering using Ward's method with Manhattan distance, shown split into six clusters

The six clusters of individuals that were chosen from the Ward's method clustering of the Manhattan distances between coordinates on the eight components are described in Table 10. The names are given to the clusters based on their within-cluster mean scores for the principal components that are shown in Figure 7. Each of the eight components from the previous section has a mean that is statistically significant at a 95% confidence level for at least two of the clusters based on a one-way analysis of variance test. This means that there is at least a 95% chance that the mean's value is due to differences between the cluster and overall sample and not due to random chance. Additionally each of the clusters is significantly described by at least five of the eight components, with both positive and negative relationships.

Cluster name	Cluster description
1: Critical cyclists	Critical cyclists value orderly network, are the least satisfied, and do not avoid adverse conditions. They have a low value of continuity, enjoyment, and importance of safety. They seem to ride because it is practical, but prefer to do so on the bicycle network.
2: Comfort cyclists	Comfort cyclists strongly avoid adverse conditions and value an orderly network, continuity, and recreation. Similarly to cluster 1, they have low satisfaction with conditions. They find the bicycle network to be important and only ride when the conditions are ideal.
3: Committed cyclists	Committed cyclists value speed, continuity, and enjoyment. They are the most willing to ride in adverse conditions and do not value an orderly network, however they do prefer continuity. They ride because it is practical but also value the recreational aspect and would likely do so with or without bicycle network improvements.
4: Cautious cyclists	Safety conscious cyclists avoid adverse conditions and traffic interaction. They are not worried about the speed of a route, which is inversely correlated with concern for safety. They are also not very concerned with an orderly network or continuity, and find a variety of surfaces to be suitable.
5: Content cyclists	Content cyclists are satisfied with the bicycle network. They do not avoid adverse conditions, but are also not at all concerned with speed. They are concerned with an orderly network and continuity but do not find diverse surfaces to be suitable.
6: Convenience cyclists	Convenience cyclists cycle because it is practical and fast, but only in ideal conditions. They are satisfied with the bicycle network, but they avoid adverse conditions. They are not very concerned with safety.

Table 10- Names and descriptions of the six clusters of cyclists considered

Figure 7- Values of significant (P>.95) components' within-cluster means



With this understanding of the clusters of cyclists, differentiated by preferences, behaviors, and satisfaction with bicycle infrastructure, it is possible to observe the similarities and differences between individuals in each cluster. This paper hypothesizes that these clusters will be substantially different demographically and geographically in their cycling behavior.

5 Sociodemographic aspects of clusters of cyclists

The survey respondents vary demographically with significant differences between clusters and categories of sociodemographic variables including gender, age, education, employment, income, car access, bicycle access, and how often they ride bicycles. The clusters and categories for each categorical demographic variable were tested with Pearson's chi-squared test to determine the likelihood that the differences between them occurred by chance. The chi squared test is described in Equation 5, where X^2 is the test statistic, O_i is the observed frequency for the ith level of the variable and E_i is the expected frequency of the ith variable. The P–value of the test is the likelihood of observing data at least as extreme as observed assuming the there is no difference between the categories. The chi-squared test is generally valid only in situations where the expected value of most or all of the categories within the cross classification is greater than five. (Franke, et al., 2012) For this reason, the demographics were grouped in a way to give each of the categories sufficient expected values. For the variable "Bike access", the P-value was taken considering only two groups, which were always and less than always. All groups were kept in the contingency table to better describe the sample and to remain parallel with the "car access" variable. P-values are included with the variables in Table 11 and Table 12 along with contingency tables between the categorical demographic variables and the clusters. Variables with a chi squared test P-value less than .05 are discussed following the contingency tables.

Equation 5- Pearson's chi squared test statistic

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

	OVERALL		Critical		Comfort		Committed		Cautious		Content		Convenience	
			#	%	#	%	#	%	#	%	#	%	#	%
Population			105	16.1%	95	14.6%	98	15.1%	111	17.1%	146	22.4%	89	13.7%
Gender *	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 3.353e-05	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Male	289	45.0%	35	33.3%	47	49.5%	61	63.5%	44	39.6%	73	50.0%	29	32.6%
Female	353	55.0%	70	66.7%	48	50.5%	35	36.5%	67	60.4%	73	50.0%	60	67.4%
Age *	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.001523	#	%	#	%	#	%	#	%	#	%	#	%	#	%
0-24	59	9.2%	6	5.7%	11	11.6%	8	8.3%	7	6.3%	14	9.6%	13	14.6%
25-39	292	45.5%	41	39.0%	34	35.8%	51	53.1%	67	60.4%	65	44.5%	34	38.2%
40-64	265	41.3%	55	52.4%	40	42.1%	36	37.5%	33	29.7%	62	42.5%	39	43.8%
65+	26	4.0%	3	2.9%	10	10.5%	1	1.0%	4	3.6%	5	3.4%	3	3.4%
Children	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.127	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Yes	238	37.1%	40	38.1%	23	24.2%	38	39.6%	47	42.3%	55	37.7%	35	39.3%
No	404	62.9%	65	61.9%	72	75.8%	58	60.4%	64	57.7%	91	62.3%	54	60.7%
Education *	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.03586	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Gymnasium or lower	115	17.9%	20	19.0%	21	22.1%	8	8.4%	14	12.6%	32	21.9%	20	22.5%
post-secondary not univ.	70	10.9%	11	10.5%	13	13.7%	7	7.4%	9	8.1%	21	14.4%	9	10.1%
university/college	456	71.1%	74	70.5%	61	64.2%	80	84.2%	88	79.3%	93	63.7%	60	67.4%
Partner	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.1569	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Living with partner	408	63.8%	56	53.8%	65	69.9%	58	60.4%	76	68.5%	93	63.7%	60	67.4%
Not living with partner	231	36.2%	48	46.2%	28	30.1%	38	39.6%	35	31.5%	53	36.3%	29	32.6%
Employment *	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.007099	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Works a job	469	73.2%	80	76.2%	59	62.1%	72	75.0%	86	77.5%	111	76.6%	61	68.5%
Student	103	16.1%	15	14.3%	14	14.7%	18	18.8%	14	12.6%	21	14.5%	21	23.6%
Neither	69	10.8%	10	9.5%	22	23.2%	6	6.3%	11	9.9%	13	9.0%	7	7.9%
Income *	**OVERALL**		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.009387	#	%	#	%	#	%	#	%	#	%	#	%	#	%
<10,000 SEK	84	13.1%	13	12.4%	16	16.8%	13	13.5%	9	8.1%	15	10.3%	18	20.2%
10,000 SEK - 14,999 SEK	62	9.7%	11	10.5%	11	11.6%	11	11.5%	11	9.9%	13	8.9%	5	5.6%
15,000 SEK - 24,999 SEK	72	11.2%	9	8.6%	9	9.5%	5	5.2%	12	10.8%	28	19.2%	9	10.1%
25,000 SEK - 34,999 SEK	231	36.0%	47	44.8%	24	25.3%	32	33.3%	45	40.5%	51	34.9%	32	36.0%
35,000 SEK - 50,000 SEK	105	16.4%	12	11.4%	15	15.8%	22	22.9%	20	18.0%	24	16.4%	12	13.5%
>50,000 SEK	31	4.8%	2	1.9%	4	4.2%	4	4.2%	6	5.4%	6	4.1%	9	10.1%
No response	57	8.9%	11	10.5%	16	16.8%	9	9.4%	8	7.2%	9	6.2%	4	4.5%

Table 11- Contingency tables of each cluster's membership by gender, age, whether they have children, education, whether they have a partner, employment, and income

			Critical		Comfort		Committed		Cautious		Content		Convenience	
Population			105	16.1%	95	14.6%	98	15.1%	111	17.1%	146	22.4%	89	13.7%
Driver's license	OVERALL		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.9542	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Yes	545	84.9%	88	83.8%	82	86.3%	82	85.4%	93	83.8%	122	83.6%	78	87.6%
No	97	15.1%	17	16.2%	13	13.7%	14	14.6%	18	16.2%	24	16.4%	11	12.4%
Car access *	OVERALL		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 0.01145	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Always	144	22.4%	20	19.0%	24	25.3%	22	22.9%	25	22.5%	31	21.2%	22	24.7%
Usually	184	28.7%	23	21.9%	34	35.8%	17	17.7%	39	35.1%	40	27.4%	31	34.8%
Sometimes	107	16.7%	18	17.1%	16	16.8%	20	20.8%	11	9.9%	24	16.4%	18	20.2%
Rarely	124	19.3%	23	21.9%	9	9.5%	23	24.0%	28	25.2%	28	19.2%	13	14.6%
Never	83	12.9%	21	20.0%	12	12.6%	14	14.6%	8	7.2%	23	15.8%	5	5.6%
Bike access *	OVERALL		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value = 3.091e-08	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Always	422	65.7%	67	63.8%	40	42.1%	68	70.8%	68	61.3%	120	82.2%	59	66.3%
Usually	206	32.1%	35	33.3%	46	48.4%	28	29.2%	42	37.8%	26	17.8%	29	32.6%
Sometimes	12	1.9%	2	1.9%	8	8.4%	0	0.0%	1	0.9%	0	0.0%	1	1.1%
Rarely	2	0.3%	1	1.0%	1	1.1%	0	0.0%	0	0.0%	0	0.0%	0	0.0%
How often do you ride? *	OVERALL		Critical		Comfort		Committed		Cautious		Content		Convenience	
p-value < 2.2e-16	#	%	#	%	#	%	#	%	#	%	#	%	#	%
5-7 days per week	500	77.9%	93	88.6%	43	45.3%	87	90.6%	90	81.1%	129	88.4%	58	65.2%
Less	142	22.1%	12	11.4%	52	54.7%	9	9.4%	21	18.9%	17	11.6%	31	34.8%

Table 12- Contingency table of each cluster's membership by driver's license status, car access, bicycle access, and how often they ride

Table 13 discusses notable and significant conclusions from the contingency tables in Table 11 and Table 12 to develop demographic profiles compared with the population at large for each of the clusters. The categories within the demographic variables that have substantially different distributions than the survey sample are noted. Demographic variables are considered that have chi-squared test p-values lower than .05, indicating a significant probability that their frequency distribution is due to differences between clusters and the categories of the demographic variables.

Table 13- Demographic profiles of the clusters

Cluster name	Demographic profile
1. Critical cyclists	66.7% female compared with the 55% female survey sample. Smaller percentage in the 0-24, 25-39, and 65+ age groups with a larger portion in the 40-64 age group. Higher percentage in the 25-35 thousand SEK/month income bracket with lower percentages in the 35-50 thousand and 50 thousand + SEK/month income brackets. Less likely to always or usually have car access, and the most likely of all clusters to never have car access. More likely than the survey sample to ride their bike 5-7 days per week.
2. Comfort cyclists	4.5% larger portion male than the survey sample. Of 26 people in the 65+ age group, 10 fall in the comfort cluster, while a substantially smaller portion fall in the 25-39 age group. A higher portion have a highest educational attainment of secondary or lower and a smaller portion have completed university or college than in the survey sample. A lower portion of comfort cyclists work a job and a higher portion are neither working nor in school. The largest portion of people who didn't respond to the income prompt are in the comfort cluster and there is a larger portion of the lowest income group. The smallest portion of people in the 25-35 thousand SEK/month income range are in the comfort cluster. A larger portion usually have car access, while a smaller portion never have car access. The smallest portion of comfort cyclists between all clusters always have bicycle access, and eight of the 12 who only sometimes have bicycle access fall in this cluster. The smallest portion, 45%, of comfort cyclists ride 5-7 days per week.
3. Committed cyclists	63.5% male compared with 45% in survey sample. Higher portion in the 25-39 age group, with lower portions in the older age groups compared with survey sample. The lowest portion of any cluster with highest

	education of secondary or lower and post-secondary besides university or college. The highest portion of any cluster with highest education of university or college. The lowest portion of any cluster that neither work nor attend school. Lowest portion of any cluster in the 15-25 thousand SEK/month income bracket with a higher portion than the survey sample in the 35-50 thousand SEK/month bracket. The lowest portion of all clusters to usually have access to a car, with higher portions rarely or never having car access. Higher portion to always have access to a bicycle, and the highest portion of all clusters to ride 5-7 days per week.
4. Cautious cyclists	5.4% larger portion female than the survey sample. Highest portion in the 25-39 age group and lowest portion in the 40-64 age group of all clusters. Lower portion in with highest education of secondary or lower and post-secondary besides university or college and higher portion having completed university or college than in the survey sample. Higher portion of workers and lower portion of students in cluster. Lower portion in the lowest income bracket and slightly higher portions in the three highest income brackets than in the survey sample. Higher portions usually and rarely have car access while lower portions sometimes and never have car access compared with the survey sample. A slightly smaller portion of the cluster always have bicycle access, and a slightly higher portion ride 5-7 days per week.
5. Content cyclists	5% larger portion male than in the survey sample. The lowest portion of any cluster with a highest educational attainment of university or college. A higher portion of content cyclists make 15-25 thousand SEK/month and a slightly lower portion make more and less than that in the survey sample. A somewhat larger portion of content cyclists never have access to a car while a substantially higher portion always have access to a bicycle. A larger portion in this cluster cycles 5-7 days per week than in the survey sample.

6. Convenience cyclists	The largest portion female of all clusters with the largest portion in the 0-24 age group. Smaller portion in the 25-39 age group than in the survey sample. Slightly higher portion with a highest education of secondary or lower potentially due to the large portion of young people. Larger portion of students and lower portion of workers or people who do neither than in the survey sample. Larger portions in both the lowest and highest income groups with equal or lower portions in all other income groups than in the survey overall. Of all clusters, the smallest portion of convenience cyclists ride 5-7 days per week and rarely or never have access to a car.

With an understanding of the sociodemographic similarities and differences between clusters of cyclists and the unique behaviors, preferences, and satisfaction levels of each cluster, this paper has developed specific profiles of each cluster. To further contribute to these profiles, Chapter 6 will examine the differences between origin and destinations of reported routes between clusters. Additionally it will discuss the route reporting task from the survey and the routes ridden by respondents.

6 Spatial analysis of cycling behavior

During the survey studied in this paper, shown in full in appendices A (Swedish) and B (English), respondents were asked to draw a map of a ride or multiple rides they complete in the city on a computer map interface. They were asked to enter starting and ending locations and then to adjust the route in between to make it as accurate as possible to indicate where they actually rode. Respondents were additionally asked to provide feedback about the route and provide a trip purpose for further analysis, although very few respondents completed this section of the task. The survey also had default settings for the route rating and purpose, which may have artificially skewed the responses toward those responses and made the responses not useful.

Respondents had varying degrees of success at making their reported route accurately mirror the transportation network. Because the route starts as a straight line between their origin and destination, and becomes more accurate as they add more points to the line, some respondents made it very clear where they rode, and some simply left their route as a perfectly straight line. Responses were filtered to retain only the somewhat reasonable responses, with a sufficient number respondents for analysis. Additionally circuitous routes that had more than three times the route distance to straight-line distance were filtered to concentrate on cycling routes used for transportation. Because matching routes to the transportation network would have been inaccurate due to the nature of the mapped routes, origin and destination locations were analyzed by postal codes and by clusters of cyclists.

Because some cyclists responded to this section of the survey with up to twelve routes, they could have skewed the data toward their own frequented postal codes if they were all considered. To avoid this effect, only the first route each of the respondents entered in the survey was considered for the analysis of this section. To verify differences between the clusters of cyclists and the origin and destination postal codes, chi-squared tests were performed on cross classification frequency tables. The p-value of the chi-squared test for the origins categorized by cluster was 0.058, indicating that the differences between origins and zones are nearly significant at a 95% confidence level. The p-value of the chi-squared test for the destination zones categorized by cluster was 0.039, indicating that the differences between those categories are significant at a 95% confidence level.

					Origins											
Destinations	211	212	213	214	215	216	217	218	222	223	230	232	234	235	238	245	Total
211	34	32	9	45	7	10	25	8	0	0	1	3	3	0	2	2	181
212	8	11	1	7	2	2	3	0	0	1	0	1	0	0	0	0	36
213	1	1	0	3	1	1	3	1	0	0	0	0	0	0	0	0	11
214	15	7	1	15	2	5	12	2	2	0	0	0	0	0	1	0	62
215	1	1	1	0	2	2	2	2	0	0	0	0	0	0	0	0	11
216	5	0	0	3	0	0	1	0	0	0	0	0	0	0	0	0	9
217	8	1	0	7	1	2	3	0	0	0	0	0	0	1	0	0	23
218	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	2
223	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	3
226	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
230	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
232	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2
237	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Total	73	54	12	85	15	22	52	15	2	1	1	4	3	1	3	2	345

Table 14- Origin-destination matrix for each of the three-digit postal codes with origins and/or destinations

Table 14 is an origin-destination matrix showing the number of trips going between each pair of zones housing origins and/or destinations. Figure 8 shows the number of origins and destinations in each three-digit postal code in the central areas considered. It demonstrates that while Postal Code 211 had more than twice as many destination points as origin points, all other zones with substantial activity had more origin points than destination points. Postal Code 211 had the largest number of destination while Postal Code 214 had the most origins. Additionally the most central postal codes had the largest numbers of origins and destinations. To ensure there were high enough expected frequencies for each of the categories, some postal codes were combined as demonstrated by the maps in Table 16.

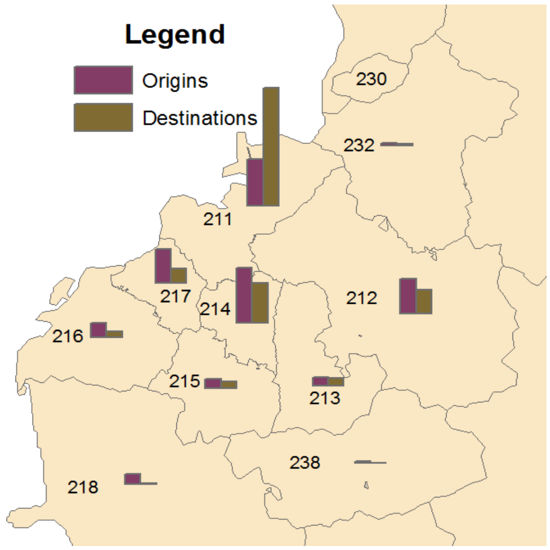


Figure 8- Number of origins and destinations in central postal codes

Table 15 shows the area, residential population and density, daytime population and density, and average income of each of the zones of analysis. Because the three-digit postal codes are larger areas that contain several smaller five-digit postal codes, the statistics do not represent the 3-digit zones as a whole. They instead describe the aggregate of the five-digit zones with origins and/or destinations in the dataset within each three-digit zone. This is meant to ensure the average income, and population densities describe the origins and destinations as accurately as possible. These areas of analysis within the post codes will henceforth be referred to as zones one through five.

Zone	Post Code(s)	Area (sq. km)	Population	Population (per sq. km)	Daytime Population	Daytime Population (per sq. km)	Average Income (SEK/year)
1	211	14	33,800	2,471	56,903	4,160	229,536
2	212-213	29	57,407	1,975	40,332	1,387	149,861
3	214	6	49,672	8,857	20,629	3,678	140,699
4	216-217	13	54,929	4,127	12,033	904	235,361
5	0thers	90	45,335	506	22,320	249	180,874

Table 15- Area, population, density, and income of five-digit post codes grouped by zone of analysis (ESRI, 2016)

Although Zone One had the third highest residential population density, it had the highest daytime population density and the second highest average income. Additionally it is the only zone with a higher daytime than residential population density, indicating that it is largely an employment center.

Zone Two covers a large area with the fourth highest residential population density and the third highest daytime population density of any of the zones. It has the second highest ratio of daytime to residential population density indicating it is also an area of much employment. Additionally it has the second lowest average income.

Zone Three is less than half the area of any of the zones with the lowest average income and by far the highest residential population density. It has the second highest daytime population density, although it is less than half its residential density. Zone Three is a dense urban population center that also has a lot of employment.

Zone Four is the second smallest zone with the highest average income. It is mostly residential with more than four times its daytime population in residential population. The two zones with the highest average income, Zone One and Zone Four are largely along the water.

Zone Five, which largely consists of post codes outside the city center predictably has the lowest residential and daytime population densities. With just over twice its daytime population in residential population it appears to be largely residential and has the third highest average income.

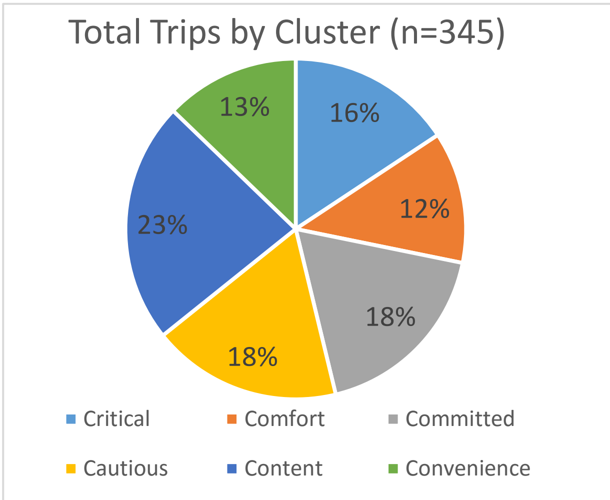
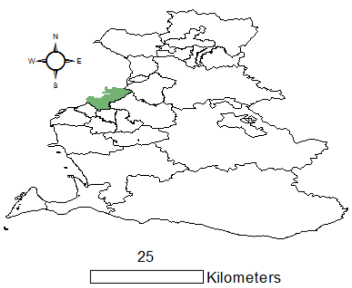
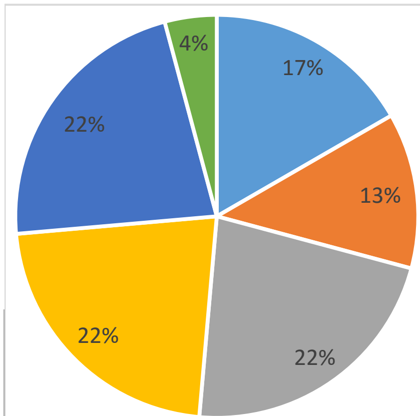
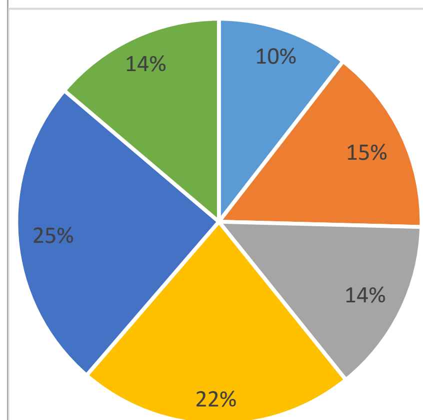
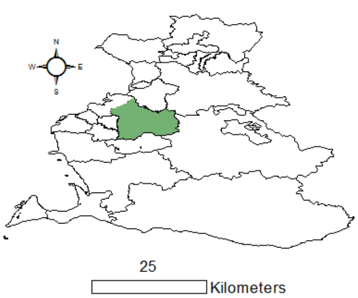
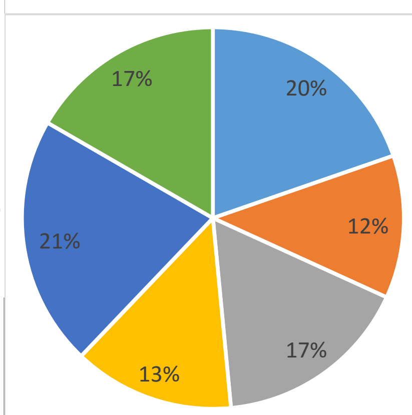
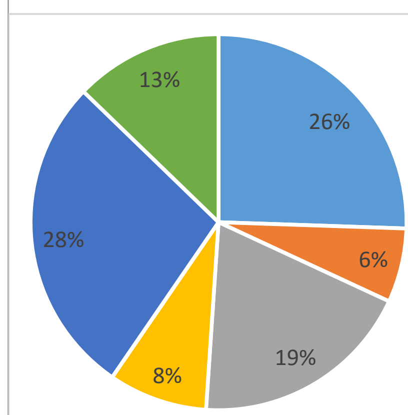


Figure 9- Trips in analysis by cluster

Table 16 shows the frequencies of each of the origin and destination zones for each of the clusters of cyclists along with shaded maps showing the geographic locations of the zones. A large labeled map of all of the zones is included in Appendix C. Figure 9 shows the total number of trips considered by members of each of the clusters.

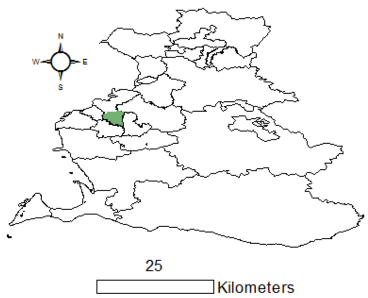
Table 17 shows heat maps of each of the cluster's odds ratio of origins plus destinations by three-digit post code (not the zones mentioned above, but the areas those zones are based upon) to demonstrate where each cluster has comparatively more and less activity. The odds ratios describe the ratio between the odds of cluster membership for origins and destinations within each three-digit post-code and the odds of cluster membership in the entire sample of origins and destinations. Some of the more outlying three-digit post codes have few or no origins and destinations, so the clusters that have origins and destinations there show up with very high odds ratios.

Table 16- Pie charts of each origin and destination zone by cluster with maps of each zone

Critical, Comfort, Committed, Cautious, Content, Convenience	P-value of X^2 test: 0.058	P-value of X^2 test: 0.039
 Zone 1	Origins (n = 73) 	Destinations (n = 181) 
 Zone 2	Origins (n = 66) 	Destinations (n = 47) 

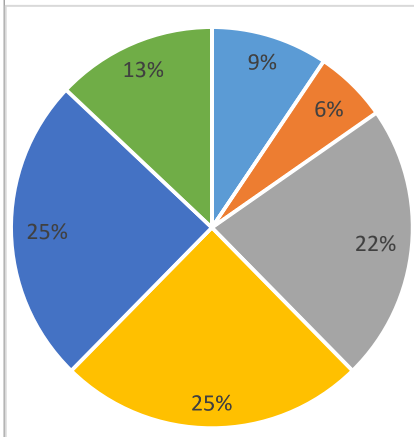
P-value of X^2 test: 0.058

P-value of X^2 test: 0.039

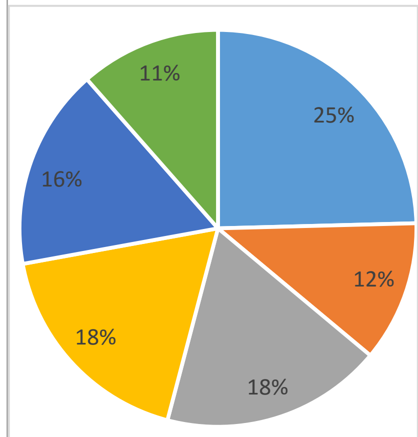


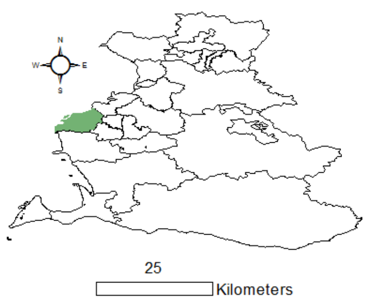
Zone 3

Origins (n = 85)



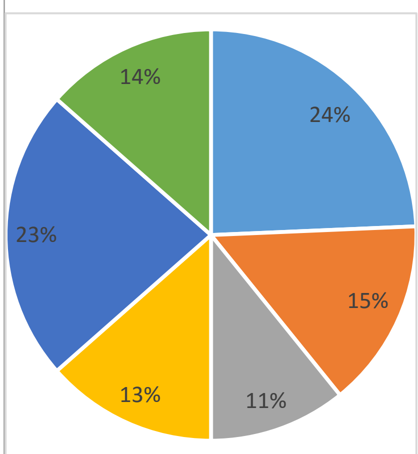
Destinations (n = 62)



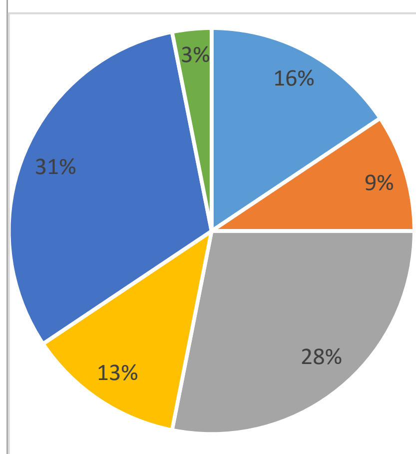


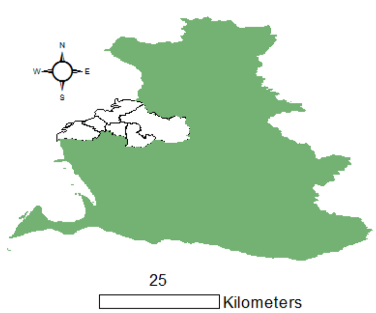
Zone 4

Origins (n = 37)



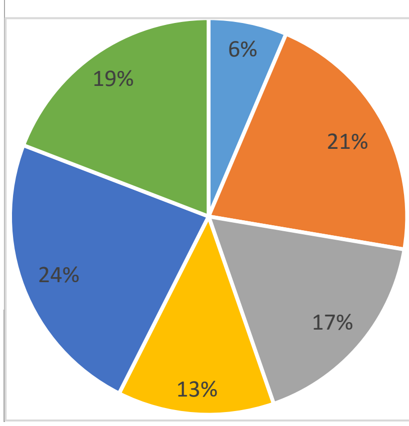
Destinations (n = 32)*



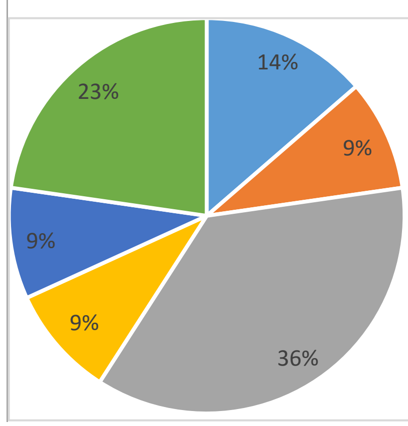


Zone 5

Origins (n = 47)

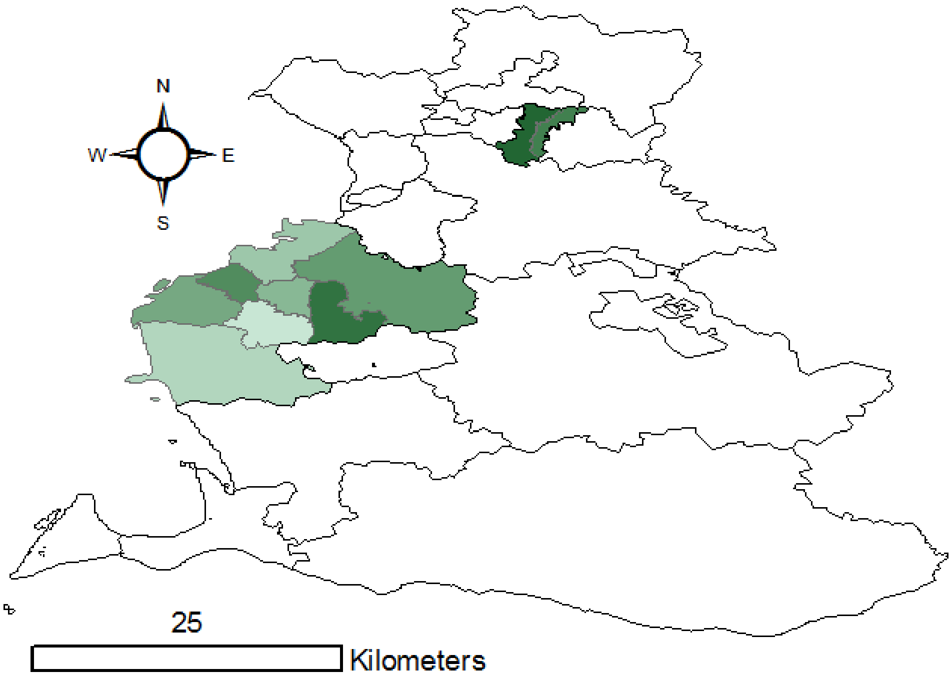


Destinations (n=23)*

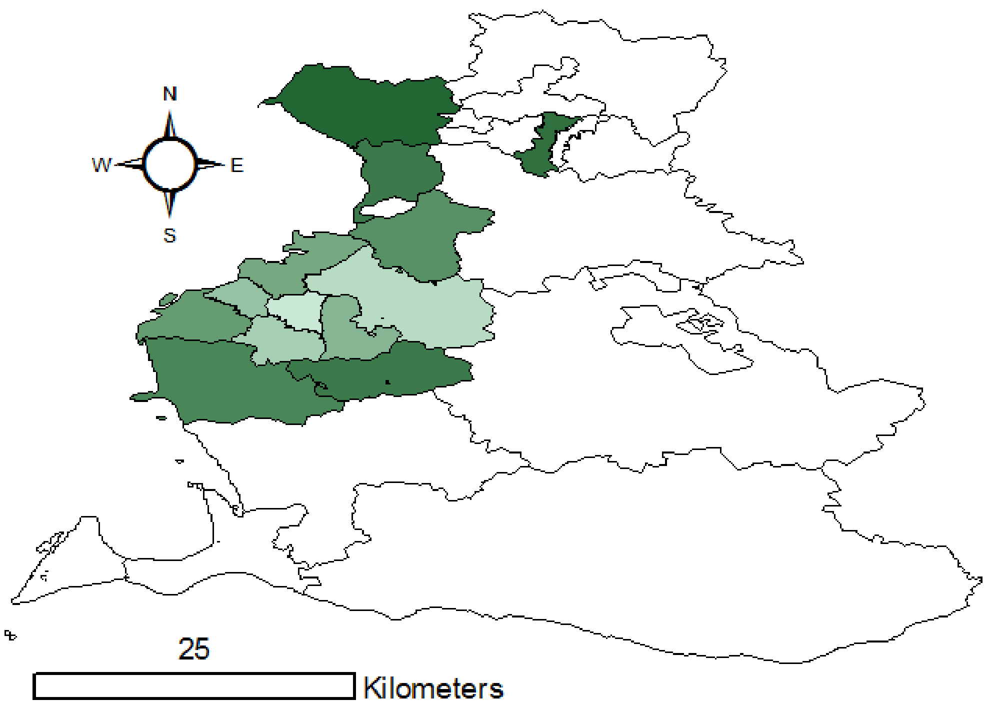


* Because of small number in category, percentages in each cluster should be considered lightly.

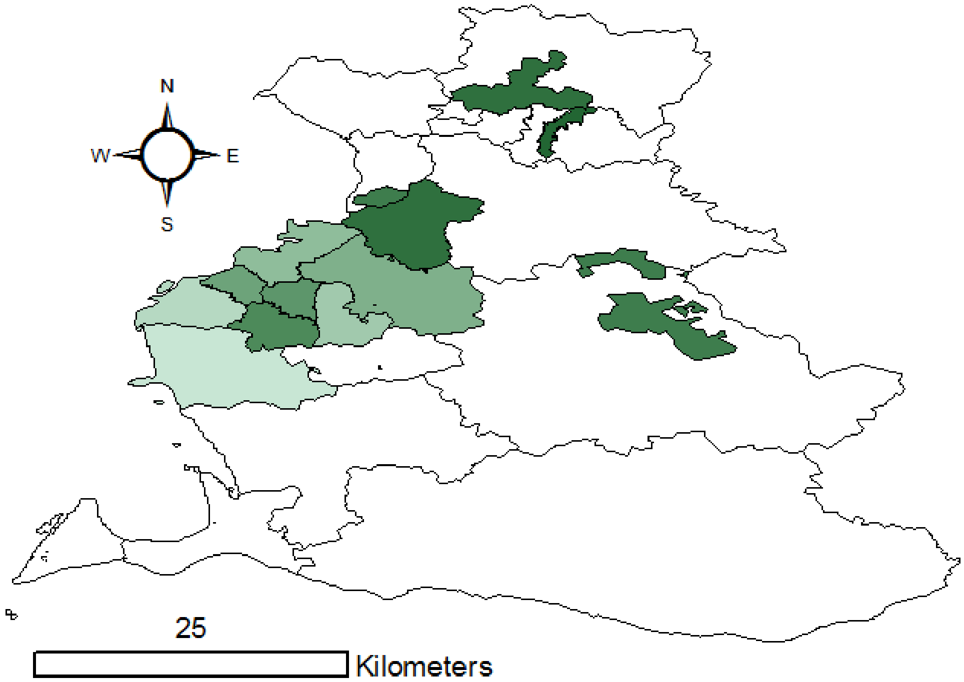
Cluster 1- odds ratios of critical cyclists by three-digit post code, darker green indicates higher odds ratio.



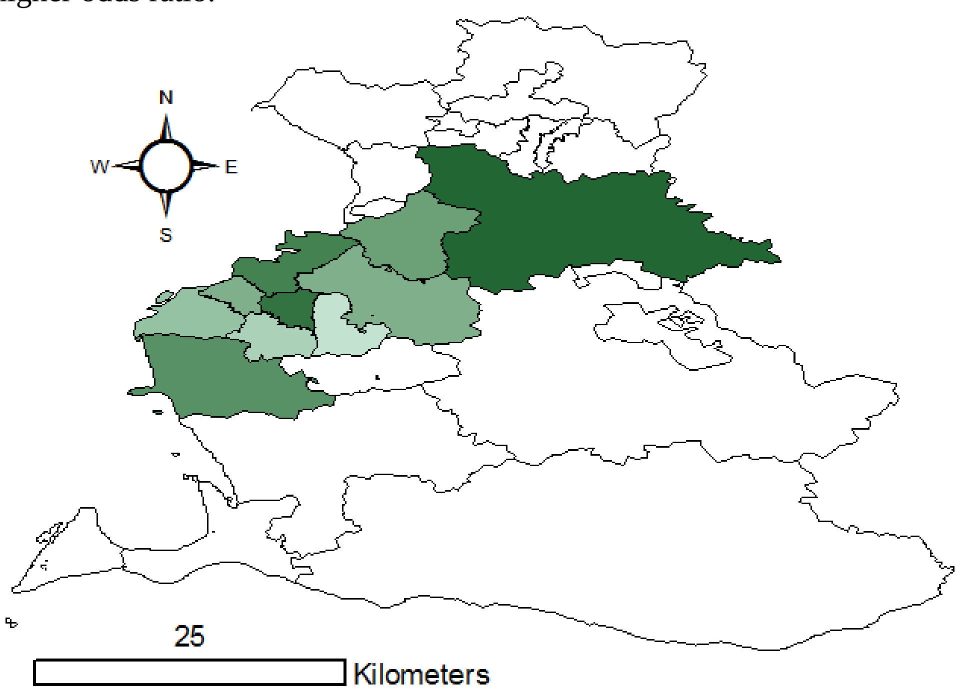
Cluster 2- odds ratio of comfort cyclists by three-digit post code, darker green indicates higher odds ratio.



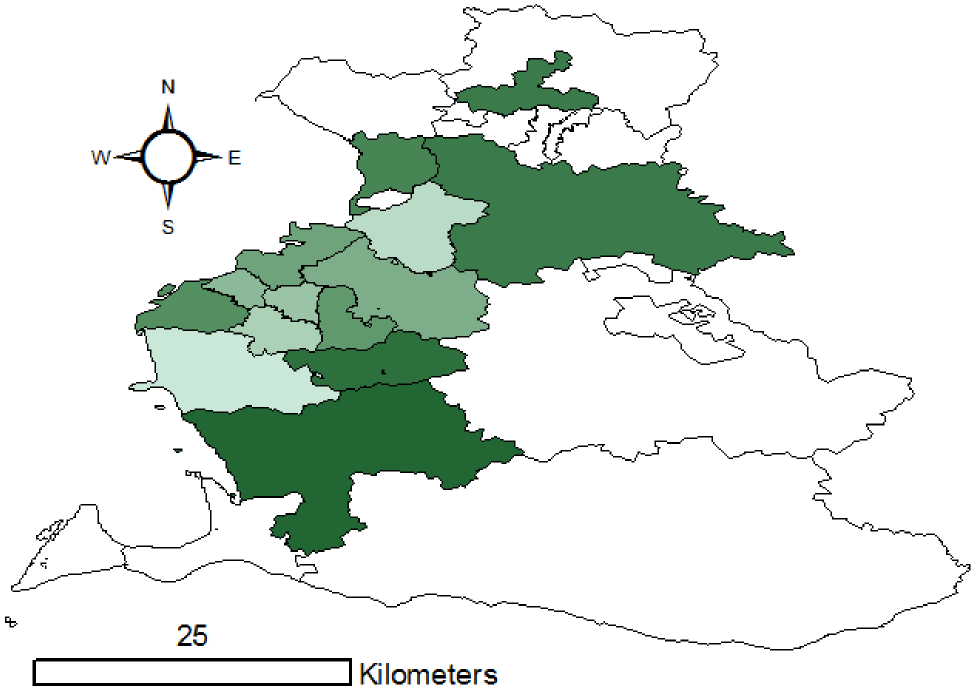
Cluster 3- odds ratio of committed cyclists by three-digit post code, darker green indicates higher odds ratio.



Cluster 4- odds ratio of cautious cyclists by three-digit post code, darker green indicates higher odds ratio.



Cluster 5- odds ratio of content cyclists by three-digit post code, darker green indicates higher odds ratio.



Cluster 6- Odds ratio of convenience cyclists by three-digit post code, darker green indicates higher odds ratio.

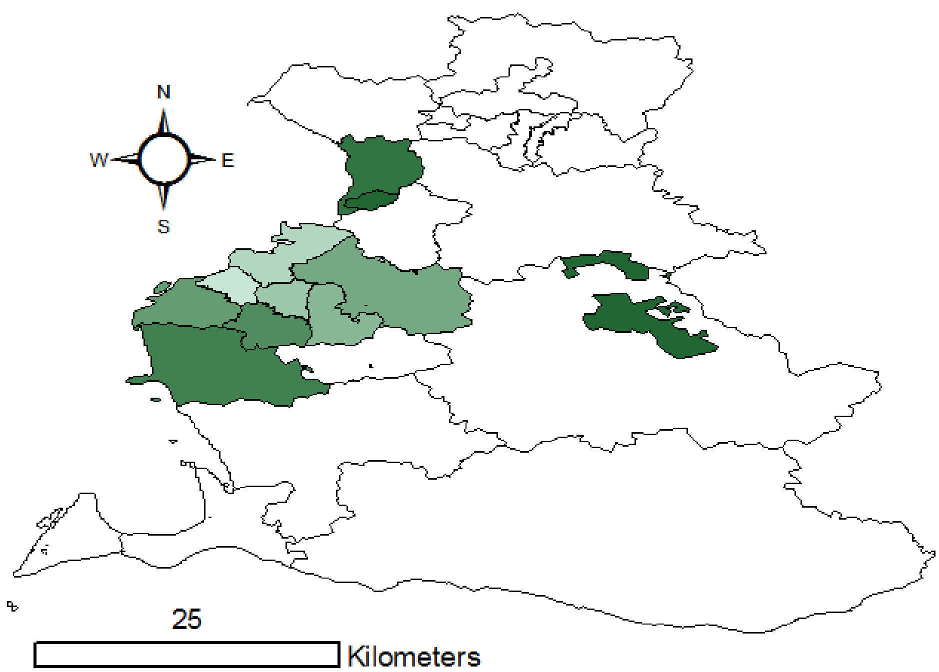


Table 17- Heat maps of odds ratios of cluster membership within zones and cluster membership in the entire sample for all clusters. Darker greens represent higher odds ratio and thus higher levels of activity. White post codes indicate no origins or destinations in the area. To see post code number labels see Appendix C.

Table 18 describes the profiles of each of the clusters of cyclists based on their portion of origins and destinations within each zone, indicating zones where each cluster's origins and/or destinations are at least 5% different than expected under the null hypothesis. The total trips by cluster described in Figure 9 are used for comparison with frequencies of origins and destinations by zone described in Table 16. It is important to note that there is no reliable information about trip purpose available for these trips, and for some of the respondents this represents only one of the several trips they marked on the map. Analysis of destinations in Zones 4 and 5 should be considered cautiously due to expected frequencies of fewer than 5 in some categories due to small sample sizes. Despite these limitations the information and insight provided are useful for analysis and significant at a 94% confidence level. Additionally the odds ratios of origins and destinations by three-digit-post code (not the same zones from table 16) compared with the whole survey for each cluster are shown in Table 17, and are used to further the discussion of cluster descriptions in Table 18.

Zone	Profiles of origins and destinations in zones by cluster
Critical (n=54)	Critical cyclists have 6% fewer than their expected trips to Zone One. They have 10% and 9% more trips than expected to Zones Two and Three Respectively, demonstrating that they travel more to the inland inner-city areas with lower income. They traveled 7% and 10% less than expected from Zones Three and Five respectively, and traveled 8% more from Zone Four which is the area with the highest average income. It is clear from the odds ratio map in Table 17 that critical cyclists have comparatively high levels of activity in post codes 213 and 217, which are parts Zone Two and Zone Four respectively. Alternatively, post code 215, which is a part of Zone Five, has a particularly low level of activity. Critical cyclists have activity generally condensed to the center with the exception of activity post codes 222 and 223.

Comfort (n=43)	Comfort cyclists traveled 6% less than expected to Zone Two and 6% less than expected from the dense Zone Three. Additionally they traveled 9% more than expected from Zone Five, which is the least dense zone with the highest income, including the outlying areas. Based on the second odds ratio map in Table 17, it is clear that comfort cyclists have a comparatively high level of activity in post codes 218, 222, 238, 232, 234, and 237, all of which are a part of Zone 5. Additionally, comfort cyclists have a particularly low level of activity in post code 214 which is the same area as the densest Zone Three. It is clear that comfort cyclists' activity level has a negative correlation with population density.
Committed (n=62)	Committed cyclists traveled 9% less than expected from Zone One and 10% more than expected to Zone Five. Additionally twice the expected trips to Zone Five were completed, although the sample size is too small for this to be trusted, it appears that these committed cyclists are more willing to ride to and from outlying and dense areas. According to the third odds ratio map from Table 17, it is clear that committed cyclists have high levels of activity in some of the outlying post codes that partially comprise Zone Five, while having a comparatively low level of activity in the central areas with the exceptions of post code 214, which is the densest area. Post codes 216 and 218 had especially low activity for committed cyclists.
Cautious (n=62)	Cautious cyclists traveled 5% and 10% less than expected from and to Zone Two respectively. They traveled 7% more than expected from the densely populated and low-income Zone Three. Additionally they traveled less than expected both from and to Zones Four and Five, which are the relatively high-income areas with low daytime population densities. Based on the fourth odds ratio map in Table 17, cautious cyclists have an especially high level of activity in post code 214, which wholly comprises Zone Three. Cautious cyclists had especially low levels of activity in post code 213, which is a part of Zone Two, and post code 215, which is a part of Zone Five.

Content (n=80)	Content cyclists traveled to the dense and low-income Zone Three 7% less than expected. Although destinations in Zones Four and Five do not have high enough frequencies for confidence in expectations, there were 8% more trips to high average income Zone Four and 14% fewer trips to low-population-density Zone Five than expected. Based on the fifth odds ratio map in Table 17, content cyclists had particularly low activity in post codes 215, 218, and 232, which are the constituents of Zone Five that had the most overall activity. Of the central post codes, content cyclists had the comparatively highest level of activity in post code 216, which is a part of Zone Four.
Convenience (n=44)	Convenience cyclists traveled 9% less from Zone One and 6% more from Zone Five than expected. They traveled 4% less to Zone Four and 10% more to Zone Five than expected, although those categories had too small quantities to properly test against expected values. Much like committed cyclists, convenience cyclists appear to be willing to travel more to and from the outlying areas. This makes sense as they seem to be more intense cyclists who are concerned with speed. According to the final odds ratio map from Table 17, convenience cyclists had a comparatively large amount of activity in the more central post codes of Zone Five. Alternatively they had comparatively low activity levels in the adjacent and expensive post codes 211 and 217.

Table 18- Profiles of all clusters based on percentage of origins and destinations by zone

At a 94% confidence level, the clusters of cyclists vary between origins and destinations in zones of consideration, despite the varying and unstated purposes of trips between individuals. It is clear that the convenience and committed cyclists are more willing to ride to and from the outlying areas, while the other clusters of cyclists seem to ride in the inner areas of the city more frequently. Despite that the clusters were formed by cyclists' behaviors, preferences, and satisfaction with cycling infrastructure, they vary significantly both demographically and geographically.

7 Conclusions

This thesis seeks to better the understanding of the heterogeneity within the population of cyclists by separating them into clusters and observing the similarities and differences between those clusters. Cyclists were clustered into categories based on components created through Factor Analysis of Mixed Data, which is a principal component analysis method for reducing the dimensionality of continuous and categorical variables together. (Pagès, 2014) The components were formed based on variables that describe survey respondents' preference for cycling route attributes, cycling behavior, and satisfaction with existing cycling infrastructure in Malmö. With scores from those components, respondents were grouped into six clusters of similar cyclists through hierarchical clustering. These clusters were then studied to determine whether they were also significantly different demographically and geographically based on origins and destinations of reported routes that they rode.

The analysis is interesting because it examines and categorizes a heterogeneous sample of cyclists based on one set of characteristics and then examines the relationship between those categories and other sets of characteristics. Understanding those relationships will allow planners to more holistically profile the population of cyclists, understanding better which clusters of people are riding in which locations. It is useful because these insights can be applied to planning cycling policies and infrastructure by targeting a group or groups of cyclists with which they are expected to have the desired outcomes.

Although this thesis intends to aid planners in increasing cycling, it is limited in that it exclusively includes current cyclists. The greatest opportunities to increase cycling may be in promoting cycling to potential cyclists who currently do not cycle or cycle infrequently, while more than 45% of the members of each group created in this research cycle 5-7 days per week. Additionally 77.9% of the total survey sample reported cycling 5-7 days per week. By considering and clustering more individuals who cycle less frequently, it could be more reasonable to draw conclusions about the differences between frequent cyclists, infrequent cyclists, and non-cyclists. Additionally it could help planners understand what measures could encourage the people who cycle less frequently to ride more often.

In Chapter Four, the 46 variables considered in the initial analysis are described by eight components that represent: 1.) Satisfaction, 2.) Avoidance of adverse conditions, 3.) Importance of an orderly bicycle network, 4.) Importance of speed, 5.) Importance of network continuity, 6.) Importance of enjoyment, 7.) Suitability of surfaces, and 8.) Importance of safety. Based on these components, groups of cyclists were created using hierarchical clustering. The resulting clusters of cyclists are: 1.) Critical cyclists, 2.) Comfort cyclists, 3.) Committed cyclists, 4.) Cautious cyclists, 5.) Content cyclists, and 6.) Convenience cyclists. Each of these clusters contains a more homogenous group of cyclists with significant variation between cyclists in different clusters.

Chapter Five and Chapter Six analyze the members of these clusters demographically and spatially based on their cycling behavior. Chapter Five finds significant differences between the clusters and categories of gender, age, education, employment, income, car access, bicycle access, and the frequency with which they ride bikes. Demographic profiles are developed based on this information for each of the clusters of cyclists. Chapter Six analyzes the origin and destination locations of reported trips by cyclists and finds substantial differences between clusters of cyclists and postal-code-based zones of origin and destination. It describes those zones demographically and develops profiles for each of the clusters of cyclists based on where their higher and lower than expected origin and destination frequencies arise. These analyses substantiate the thesis statement, demonstrating that cyclists clustered by cycling behavior, preference, and satisfaction are also substantially connected by their personal characteristics and their cycling behavior.

Future research should take clustering of cyclists farther by examining the difference of route choice behavior by the different clusters of cyclists to determine which route attributes different clusters seek out and avoid. Additionally it would be interesting to study how particular treatments affect cyclists' satisfaction by surveying the same groups before and after changes in infrastructure and policy. This would enable planners to understand how actual changes differently affect the subgroups within the population of cyclists. Finally future research aimed to increase cycling should examine a broader cross section of the community to better understand infrequent cyclists in addition to the generally frequent cyclists considered in this thesis.

8 Works Cited

Abdi, H. & Williams, L., 2010. Principal component analysis. In: *Wiley Interdisciplinary Reviews: Computational Statistics, Volume 2, Issue 4.* s.l.:John Wiley & Sons, Inc., pp. 433-459.

Akar, G. & Clifton, K., 2009. Influence of Individual Perceptions and Bicycle Infrastructure on Decision to Bike. *Transportation Research Record,* pp. 165-172.

Anderson, T., 2014. Malmo: A city in transition. *Cities,* Volume 39, pp. 10-20.

Baeten, G., 2011. Normalising Neoliberal Planning: The Case of Malmö, Sweden. In: T. Tasan-Kok & G. Baeten, eds. *Contradictions of Neoliberal Planning.* Dordrecht, The Netherlands: Springer, pp. 21-42.

Batagelj, V., 1988. Generalized Ward and Related Clustering Problems. *Classification and Related Methods of Data Analysis,* pp. 67-74.

Bauman, A. et al., 2008. Cycling: getting Australia moving - barriers, facilitators and interventions to get more Australians physically active through cycling. *Australian Transport Research Forum,* Volume 31, pp. 593-601.

belsoft, n.d. *Climate Malmo - Sweden.* [Online]
Available at: http://www.climatedata.eu/climate.php?loc=swxx0020&lang=en
[Accessed 21 04 2016].

Bicycle Malmö 2012. 2012. [Film] Directed by Martin Lang. Sweden: Malmö Stad.

Black, P., 2006. *Manhattan Distance.* [Online]

Available at: https://xlinux.nist.gov/dads//HTML/manhattanDistance.html

[Accessed 9 May 2016].

Börjesson, M. & Eliasson, J., 2012. The value of time and external benefits in bicycle appraisal. *Transportation Research Part A,* Volume 46, pp. 673-683.

Broach, J., Dill, J. & Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A 46*, pp. 1730-1740.

Cervero, R. & Kockelman, K., 1997. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D 2*, pp. 3: 199-219.

Copenhagenize, 2015. *Malmö, Sweden.* [Online]

Available at: http://copenhagenize.eu/index/06_malmo.html

[Accessed 18 APR 2016].

Damant-Sirois, G. & El-Geneidy, A., 2015. Who cycles more? Determining cycling frequency through a segmentation approach in Montreal, Canada. *Transportation Research Part A 77,* pp. 113-125.

Damant-Sirois, G., Grimsrud, M. & El-Geneidy, A., 2014. What's your type: a multidimensional cyclist typology. *Transportation*, pp. 41:1153-1169.

Design New Haven, 2009. *Mind the Gaps: Connected Street Networks Key to Economic Stimulus.* [Online]

Available at: http://downtownnewhaven.blogspot.se/2009/01/eliminating-gaps-connected-street.html

[Accessed 10 May 2016].

Dill, J. & Carr, T., 2003. Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them. *Transportation Research Reecord,* pp. 116-123.

Dill, J. & McNeil, N., 2013. Four Types of Cyclists? Examination of Typology for Better Understanding of Bicycling Behavior and Potential. *Transportation Research Record*, pp. 129-138.

EPOMM, 2013. *TEMS (The EPOMM Modwal Split Tool).* [Online]

Available at: http://www.epomm.eu/tems/result_city.phtml?city=46

[Accessed 18 APR 2016].

ESRI, 2016. *Open Data PostCode.* [Online]

Available at: http://opendata.arcgis.com/datasets/a65ed98be782459289d0dea41436dd60_0

[Accessed 20 May 2016].

European Platform on Mobility Management, 2016. *Modal Split Tool.* [Online]

Available at: http://www.epomm.eu/tems/result_cities.phtml?new=1

[Accessed 21 April 2016].

Franke, T. M., Ho, T. & Christie, C. A., 2012. The Chi-Square Test: Often Used and More Often Misinterpreted. *American Journal of Evaluation*, 33(3), pp. 448-458.

Garrard, J., Rose, G. & Lo, S. K., 2008. Promoting transportation cycling for women: The role of bicycle infrastructure. *Preventive Medicine 46*, pp. 55-59.

Gatersleben, B. & Haddad, H., 2010. Who is the typical bicyclist?. *Transportation Research Part F,* Volume 13, pp. 41-48.

Geller, R., 2009. *Four Types of Cyclists.* [Online]
Available at: https://www.portlandoregon.gov/transportation/article/237507

Google, 2016. *Translate.* [Online]
Available at: https://translate.google.com/
[Accessed 21 May 2016].

Handy, S., Cao, X. & Mokhtarian, P., 2005. Correlation or causality between the built environment and travel behavior?Evidence from Northern California. *Transportation Research Part D 10*, pp. 427-444.

Handy, S., van Wee, B. & Kroesen, M., 2014. Promoting Cycling for Transport: Research Needs and Challenges. *Transport Reviews,* 34(1), pp. 4-24.

Harkley, D. L. et al., 1998. *DEVELOPMENT OF THE BICYCLE COMPATIBILITY INDEX: A LEVEL OF SERVICE CONCEPT, FINAL REPORT,* s.l.: FHWA.

Holland, S., 2008. *Principal Component Analysis (PCA).* [Online]
Available at: http://strata.uga.edu/software/pdf/pcaTutorial.pdf
[Accessed 15 May 2016].

Husdson, F., Josse, J. & Pages, J., 2010. *Principal component methods - hierarchical clustering - partitional clustering: why would we need to choose for visualizing data?,* Rennes Cedex, France: Agrocampus Quest.

Husson, F., Josse, J., Le, S. & Mazet, J., 2016. *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining.* 1.32 ed. s.l.:s.n.

Hu, Y. & Li, X., 2013. Low-Carbon Transport System by Bicycle, in Malmö, Sweden. In: F. Chen, Y. Liu & G. Hua, eds. *LTLGB 2012.* Berlin: Springer Berlin Heidelberg, pp. 283-292.

Jackson, J. E., 1991. Chapter 3: Scaling of Data. In: *A User's Guide to Principal Components.* s.l.:John Wiley & Sons, Inc., pp. 63-79.

Jensen, S. U., 2007. Pedestrian and Bicyclist Level of Service on Roadway Segments. *Transportation Research Record,* Volume 2031, pp. 43-51.

Kassambara, A., 2015. *FactoMineR and factoextra : Principal Component Analysis Visualization - R software and data mining.* [Online]
Available at: http://www.sthda.com/english/wiki/factominer-and-factoextra-principal-component-analysis-visualization-r-software-and-data-mining#contributions-of-the-variables-to-the-principal-components
[Accessed 29 April 2015].

Klobucar, M. & Fricker, J., 2007. Network Evaluation Tool to Improve Real and Perceived Bicycle Safety. *Transportation Research Record,* Volume 2031, pp. 25-33.

Krizek, K. & El-Geneidy, A., 2007. Segmenting Preferences and Habits of Transit Users and Non-Users. *Journal of Public Transportation,* 10(3), pp. 71-94.

Laerd Statistics, n.d. *Spearman's Rank-Order Correlation.* [Online]
Available at: https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-

statistical-guide.php

[Accessed 21 May 2016].

Larsen, J., El-Geneidy, A. & Yasmin, F., 2010. Beyond the Quarter Mile: Examining Travel Distances by Walking and Cycling, Montreal, Canada. *Canadian Journal of Urban Research: Canadian Planning and Policy (supplement),* 19(1), pp. 70-88.

Larsen, J., Patterson, Z. & El-Geneidy, A., 2013. Build it. But where? The use of geographic information systems in identifying locations for new cycling infrastructure. *International Journal of Sustainable Transportation,* 7(4), pp. 299-317.

Lowry, M., Callister, D., Gresham, M. & Moore, B., 2012. Using Bicycle Level of service to Assess Community-Wide Bikeability. *Transportation Research Record*, Volume 2314, pp. 41-48.

Macmillan, A. et al., 2014. The Societal Costs and Benefits of Cummuter Bicycling: Simulating the Effects of Specific Policies Using System Dynamics Modeling. *Environmental Health Perspectives,* 122(4), pp. 335-344.

Maibach, M. et al., 2008. *Handbook on estimation of external costs in the transport sector Internalisation Measures and Policies for All external Cost of Transport (IMPACT)* , Delft: CE Delft.

Ma, L., Dill, J. & Mohr, C., 2014. The objective versus the perceived environment: what matters for bicycling?. *Transportation*, pp. 41:1135-1152.

Malmö stad, 2016. *kartor.* [Online]

Available at:

http://kartor.malmo.se/rest/leaf/1.0/?zoom=12&center=13,55.58&ol=adress,cykelpumpar&bl=stadskartan&config=../configs-1.0/malmo_atlas.js
[Accessed 25 April 2016].

Malmö stad, n.d. *Malmö – a city in transition.* [Online]
Available at: http://malmo.se/English.html
[Accessed 27 April 2016].

Pagès, . J., 2014. *Multiple Factor Analysis by Example Using R.* London: Chapman & Hall.

Parkin, J., Ryley, T. & Jones, T., 2007. Barriers to Cycling: An Exploration of Quantitative Analyses. In: D. Horton, P. Rosen & P. Cox, eds. *Cycling and Society.* Aldershot, England: Ashgate Publishing Limited, pp. 67-82.

Parkin, J., Wardman, M. & Page, M., 2008. Estimation of the determinants pf bicycle mode share for the journey to work using census data. *Transportation,* pp. 35:93-109.

Poorfakhraei, A. & Rowangould, G., 2015. Estimating Welfare Change Associated with Improvements in Urban Bicycling Facilities. *Journal of Transportation Engineering,* pp. 04015025:1-7.

Rabl, A. & de Nazelle, A., 2012. Benefits of shift from car to active transport. *Transport Policy,* Volume 19, pp. 121-131.

Rietveld, P. & Daniel, V., 2004. Determinants of bicycle use: do municipal policies matter?. *Transportation Research Part A 38,* pp. 531-550.

Sahlqvist, S. & Heesch, K., 2012. Characteristics of Utility Cyclists in Queensland, Australia: An Examination of the Associations Between Individual, Social, and Environmental Factors and Utility Cycling. *Journal of Physical Activity and Health,* pp. 9,818-828.

Shafizadeh, K. & Niemeier, D., 1997. Bicycle Journey-to-Work: Travel Behavior Characteristics and Spatial Attributes. *Transportation Research Record,* Volume 1578, pp. 84-90.

Statistics Sweden, 2016. *Population by region and year.* [Online]
Available at:
http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101A/BefolkningNy/table/tableViewLayout1/?rxid=8ed76d58-a3a9-4c71-bcbb-fe4bba923b15
[Accessed 28 April 2016].

Tilahun, N., Levinson, D. & Krizek, K., 2007. Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A 41,* pp. 287-301.

Titze, S., Stronegger, W., Janschitz, S. & Oja, P., 2008. Association of built-environment, social-environment and personal factors with bicycling as a mode of transportation among Austrian city dwellers. *Preventive Medicine 47,* pp. 252-259.

Transportation Research Board, 2011. *Highway Capacity Manual,* Washington, D.C.: s.n.

Wardman, M., Tight, M. & Page, M., 2007. Factors influencing the propensity to cycle to work. *Transportation Research Part A 41,* pp. 339-350.

Winters, M., Brauer, M., Setton, E. & Teschke, K., 2013. Mapping bikeability: a spatial tool to support sustainable travel. *Environment and Planning B: Planning and Design,* Volume 40, pp. 865-883.

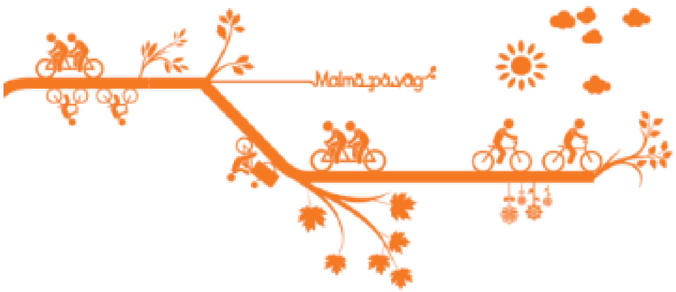
Winters, M., Davidson, G., Kao, D. & Teschke, K., 2011. Motivators and deterrents of bicycling: comparing influences on the decision to ride. *Transportation,* pp. 38:153-168.

Woodcock, J. et al., 2009. Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport. *The Lancet,* Volume 374, pp. 1930-1943.

Zahran, S. et al., 2008. Cycling and Walking: Explaining the spatial distribution of healthy modes of transportation in the United States. *Trandportation Research Part D,* Volume 13, pp. 462-470.

Appendix A- Survey in original Swedish

Malmö stad



Malmö har flera gånger utnämnts till Sveriges bästa cykelstad, senast förra året, och visst är det roligt med utmärkelser, men vi tycker att det är du som cyklar som gör Malmö till den bästa cykelstaden. Nu behöver vi din hjälp för att bli ännu bättre. Svara på några frågor om vad som är viktigast för dig och visa oss vilka vägar du väljer när du cyklar. Du kan även peka ut platser du tycker är bra eller dåliga ur cykelsynpunkt. Har du dessutom förslag på vad ett kommande hyrcykelsystem i Malmö skulle kunna heta, kan du vara med och tävla om en elcykel och två airbags för cyklister.

Kan du i allmänhet använda dig av cykel när du behöver?

- Ja, alltid
- Ja, för det mesta
- Ja, ibland
- Nej, sällan

Har du körkort för bil?

- Ja
- Nej

Kan du i allmänhet använda dig av bil när du behöver?

- Ja, alltid
- Ja, för det mesta
- Ja, ibland
- Nej, sällan
- Nej, aldrig

Sida 1/5

Malmö på väg

Malmö stad

Vilken typ av cykel och utrustning har du tillgång till?

(Du kan välja flera)

- ☐ Standard
- ☐ Racer
- ☐ Elcykel
- ☐ Lastcykel
- ☐ Liggcykel
- ☐ Kärra till cykel

Hur ofta cyklar du normalt?

- ○ 5-7 dagar per vecka
- ○ 3-4 dagar per vecka
- ○ Någon/några gånger per vecka
- ○ Någon/några gånger per månad
- ○ Någon/några gånger per år

Hur ofta reser du med cykel till följande?

	5 eller fler gånger per vecka	3-4 gånger per vecka	Någon gång per vecka	Några gånger per månad	Några gånger per år	Aldrig
Arbete/studier	○	○	○	○	○	○
Tjänsteresa	○	○	○	○	○	○
Inköp/ärende	○	○	○	○	○	○
Nöje/fritidsaktivitet	○	○	○	○	○	○
Besöka vänner/släkt	○	○	○	○	○	○
Hämta/lämna barn	○	○	○	○	○	○

Brukar du cykla till hållplats/station när du ska vidare med kollektivtrafik?

- ○ Ja, när jag ska åka tåg
- ○ Ja, när jag ska åka buss
- ○ Ja, både när jag ska åka tåg eller buss
- ○ Nej

Under vilka årstider använder du cykel?

(Du kan välja flera)

- ☐ Vår
- ☐ Sommar
- ☐ Höst
- ☐ Vinter

Undviker du att cykla vid dåligt väder eller underlag?

(Du kan välja flera)

- ☐ Ja, när det regnar
- ☐ Ja, när det snöar
- ☐ Ja, vid kraftig vind
- ☐ Ja, när det är halka
- ☐ Ja, när det är kallt
- ☐ Ja, när det är dåligt skottat
- ☐ Ja, när det ligger mycket grus/löv
- ☐ Nej, jag cyklar oavsett väder

Hur lång tid tar det i genomsnitt när du cyklar till följande, enkel väg?

	Ej aktuell	0-10 minuter	10-20 minuter	20-30 minuter	30-45 minuter	45-60 minuter	60-90 minuter	Mer än 90 minuter
Arbete/studier								
Tjänsteresa								
Inköp/ärende								
Nöje/fritidsaktivitet								
Besöka vänner/släkt								
Hämta/lämna barn								

Hur lång tid kan du som MAX tänka dig att cykla till följande, enkel väg?

	Ej aktuell	0-10 minuter	10-20 minuter	20-30 minuter	30-45 minuter	45-60 minuter	60-90 minuter	Mer än 90 minuter
Arbete/studier								
Tjänsteresa								
Inköp/ärende								
Nöje/fritidsaktivitet								
Besöka vänner/släkt								
Hämta/lämna barn								

Vad tycker du är viktigast vid val av rutt med cykel?

Rangordna dem så att ingen har samma värde.

(1 är minst viktig och 10 är mest viktig)

Se nedan för förklaringar

	1	2	3	4	5	6	7	8	9	10
Avstånd										
Tid										
Trafiksäker										
Få korsningar										
Separerat från biltrafik										
Sammanhängande nät										
Jämnt och bra underlag										
Kännas trygg										
Omgivning/miljö										
Utrustning										

Kännas trygg: Bra belysta vägar, ej skymmande buskage etc.

Utrustning: Toaletter, cykelpumpar, dusch/ombytesrum, cykelgarage, vatten etc.

Sida 2/5

Malmö på väg

Tillbaka Nästa

Malmö stad

På nästa sida kommer det komma upp en karta där vi vill att du ritar ut hur du cyklar i staden och markerar platser som du tycker fungerar bra eller dåligt ur cykelsynpunkt.

Du kan när som helst få upp information om hur du gör för att skapa färdväg och markering genom att trycka på frågetecknet längst upp till vänster på nästa sida.

Skapa färdväg

1. Skriv in adress på din startpunkt. Alternativt trycker du avbryt och letar upp startpunkt manuellt i kartan och vänsterklickar.
2. Leta upp din målpunkt och vänsterklicka. Nu har du skapat en start- och målpunkt. Det kommer upp en ruta där du kan välja beskrivning av färdvägen, vilken typ av ärende och sätta betyg på den.
3. Redigera färdvägen genom att dra i fyrkanterna på linjen, det skapas nya vart efter du justerar färdvägen. Försök vara så noggrann som möjligt. För att redigera beskrivning klickar du på själva linjen.
4. När du är klar med färdvägen kan du skapa fler färdvägar genom att trycka på pennan med ett plus.

Skapa markering

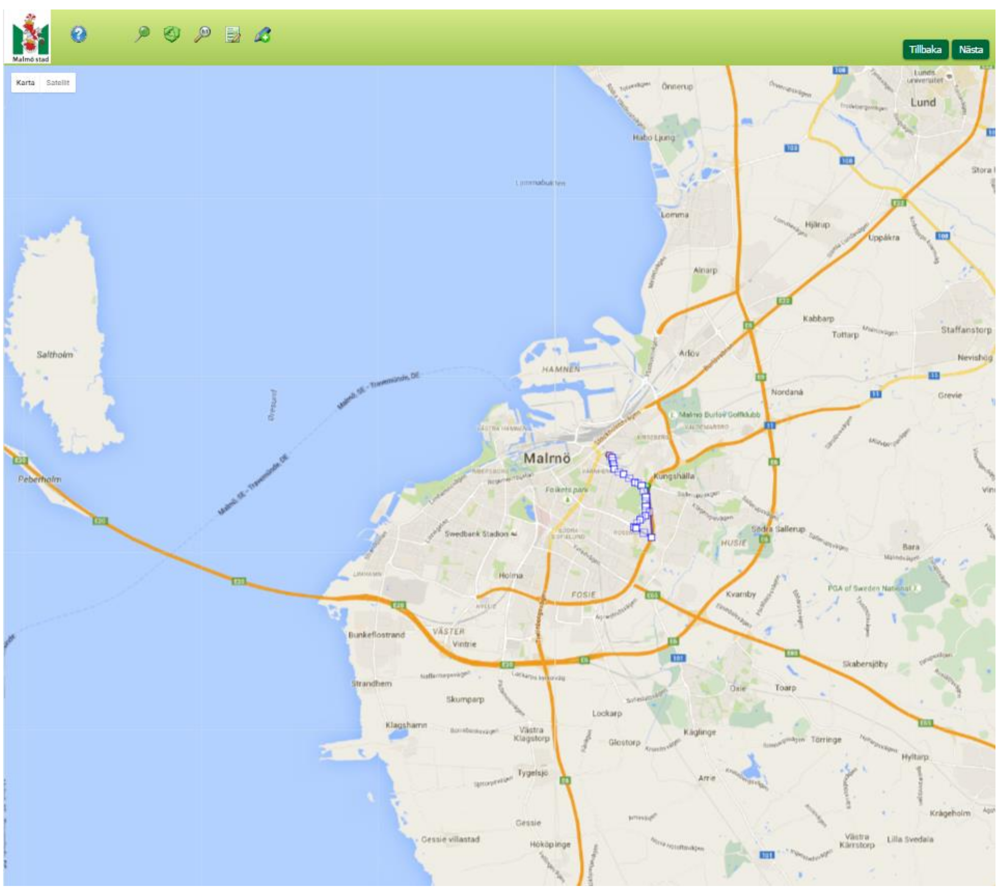
För att skapa en markering på kartan där du antingen tycker något är positivt, negativt eller annan synpunkt trycker du på knappnålen. Sedan zoomar du in och markerar platsen. Även här får du upp en ruta där du väljer vad markeringen avser och du kan även lämna kommentar.

Du kan även lämna förslag på förbättringsåtgärder och/eller välja översyn framkomlighet/trafiksäkerhet. Lämna då kommentar på vad.

Sida 3/5

Malmö på väg

Tillbaka Nästa



Hur nöjd är du med följande i Malmö? (Skala 1-5)

(1 missnöjd, 5 mycket nöjd)

	1	2	3	4	5
Malmö som cykelstad	○	○	○	○	○
Mängden cykelbanor	○	○	○	○	○
Cykelbanornas skick	○	○	○	○	○
Cykelbanornas bredd	○	○	○	○	○
Omledning vid vägarbeten	○	○	○	○	○
Cykelparkering vid bostad	○	○	○	○	○
Cykelparkering vid affärer	○	○	○	○	○
Cykelparkering vid arbetsplats/skola	○	○	○	○	○
Cykelparkering på offentliga platser	○	○	○	○	○
Cykelparkering vid Malmö C, Triangeln och Hyllie	○	○	○	○	○
Cykelparkering vid busshållplatser	○	○	○	○	○

Vilka faktorer är viktigast för dig när du parkerar din cykel för korta ärenden?

(Du kan MAX markera 3 stycken)

- ☐ Att parkeringen är övervakad (kameror/väktare)
- ☐ Att cykelparkeringen är låst
- ☐ Att parkeringen är upplyst
- ☐ Att det finns tak över cykeln
- ☐ Att jag kan låsa fast cykelns ram i stället
- ☐ Att parkeringen är lokaliserad nära där folk rör sig
- ☐ Att parkeringen är nära min målpunkt
- ☐ Att det är gott om plats för cykeln

Vilka faktorer är viktigast för att du ska känna dig trygg med att parkera din cykel en längre tid (en dag eller mer)?

(Du kan MAX markera 3 stycken)

- ☐ Att parkeringen är övervakad (kameror/väktare)
- ☐ Att cykelparkeringen är låst
- ☐ Att parkeringen är upplyst
- ☐ Att det finns tak över cykeln
- ☐ Att jag kan låsa fast cykelns ram i stället
- ☐ Att parkeringen är lokaliserad nära där folk rör sig
- ☐ Att parkeringen är nära min målpunkt
- ☐ Att det är gott om plats för cykeln

Kan du tänka dig att parkera lite längre ifrån din målpunkt om cykelparkeringen känns säker?

- ○ Ja, oftast
- ○ Ja, om cykeln ska stå parkerad en längre tid
- ○ Nej

Hur viktiga är följande åtgärder för dig som cyklist?

(1 ej viktig, 5 mycket viktig)

	1	2	3	4	5
Öka tryggheten/säkerheten i korsningar	○	○	○	○	○
Öka bredden på cykelbanor	○	○	○	○	○
Snöröjning	○	○	○	○	○
Sopa bort löv	○	○	○	○	○
Sopa bort grus	○	○	○	○	○
Åtgärder för att separera gående och cyklister	○	○	○	○	○
Ta bort nivåskillnader	○	○	○	○	○
Minimera antal stopp	○	○	○	○	○
Jämna till underlag	○	○	○	○	○
Hålla efter buskage/träd	○	○	○	○	○
Poliskontroller (t.ex. kolla belysning, mopedister)	○	○	○	○	○

Vilka material anser du vara lämpliga som underlag på cykelbanor?

(Du kan välja flera)

- ☐ Asfalt
- ☐ Grus
- ☐ Hårt packat grus
- ☐ Plattor
- ☐ Gatsten

Skulle du kunna tänka dig att samla in data åt oss med hjälp av gps, till exempel med en app i telefonen?

- ○ Ja
- ○ Nej

Hur trygg känner du dig som cyklist i Malmö?

Ej trygg	Delvis trygg	Trygg
○	○	○

Sida 4/5

Malmö på väg

Tillbaka Nästa

Malmö stad

Är du...?

- Kvinna
- Man
- Vill ej svara

Hur gammal är du?

- Under 13 år
- 13-17 år
- 18-24 år
- 25-39 år
- 40-64 år
- 65-75 år
- 76-84 år
- 85 år eller äldre

Hur ser din familjesituation ut?

- Partner som jag bor ihop med
- Ensamstående/särbo
- Bor med föräldrar eller motsvarande

Hur många barn har du som bor hemma?

- Inga
- 1 barn
- 2 barn
- 3 barn
- 4 barn eller fler

Vilken är din huvudsakliga sysselsättning?

- Förvärvsarbetar
- Studerar
- Sjukskriven
- Föräldraledig
- Arbetssökande
- Pensionär
- Annat

Vilken är din högsta avslutade utbildning?

- Går i grundskola
- Grundskola/folkskola
- Gymnasium/realskola
- Eftergymnasial utbildning, annan än högskola/universitet
- Utbildning på högskola/universitet

Inkomst per månad:

- Mindre än 10 000 kr
- 10 000 - 14 999 kr
- 15 000 - 24 999 kr
- 25 000 - 34 999 kr
- 35 000 - 50 000 kr
- Mer än 50 000 kr
- Vill ej svara

Tävlingsfråga

Göteborg har Styr & Ställ, Stockholm har sina City Bikes och Köpenhamn har bycykler.

Bidra med ditt kreativa förslag på vad ett hyrcykelsystem skulle kunna heta i Malmö. Det bästa förslaget prisas med en elcykel och ytterligare 2 stycken belönas med en airbag för cyklister.

(Vill du inte ha chansen att vinna kan du hoppa över att fylla i)

Ditt förslag:

Skriv in namn och kontakt-information, e-post och/eller telefonnummer, så att vi kan nå dig vid eventuell vinst.

(Vill du inte ha chansen att vinna kan du hoppa över att fylla i)

Namn:

E-post

Telefon:

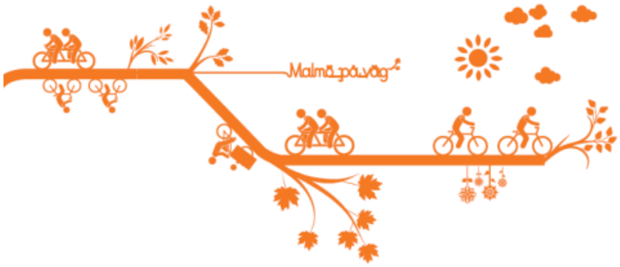
Sida 5/5

Malmö på väg

Tillbaka Skicka

Appendix B- Survey translated to English

Translated by Carl Wiman



Malmö has repeatedly been appointed the best cycling city of Sweden, most recently last year. Even though awards are important we think that it is you, the cyclist, who make Malmö the best cycling city. Now we need your help to get even better. Answer a few questions about what is important to you, and show us what paths you choose when you ride. You can also point out the places you think are good or bad from a cycling point of view. If you also have suggestions on what a future rental bike service in Malmö could be named, you could win an electric bicycle and two inflatable helmets.

Can you generally use a bicycle when you need to?

- Yes, always
- Yes, most of the time
- Yes, sometimes
- No, rarely

Do you have driver's license?

- Yes
- No

Can you generally use a car when you need to?

- Yes, always
- Yes, most of the time
- Yes, sometimes
- No, rarely
- No, never

Page 1/5

What kind of bicycle and equipment do you have access to?

(You can select multiple options)

- ☐ Standard
- ☐ Road bike
- ☐ Electric bike
- ☐ Cargo bike
- ☐ Recumbent
- ☐ Trailers for bicycles

How often do you ride normally?

- ○ 5-7 days per week
- ○ 3-4 days per week
- ○ Once/few times per week
- ○ Once/few times per month
- ○ Once/few times a year

How often do you travel by bicycle to the following?

	5 or more times per week	**3-4 times per week**	**Once / few times per week**	**A few times per month**	**Several times a year**	**Never**
Work/studies	○	○	○	○	○	○
Business travel	○	○	○	○	○	○
Purchase/errand	○	○	○	○	○	○
Entertainment/leisure	○	○	○	○	○	○
Visiting friends/family	○	○	○	○	○	○
Pickup/leave children	○	○	○	○	○	○

Do you cycle to the bus stop / train station to proceed with public transportation?

- ○ Yes, when I ride the train
- ○ Yes, when I ride the bus
- ○ Yes, both when I go by train and bus
- ○ No

During which seasons do you use a bicycle?

(You can select multiple options)

- ☐ Spring
- ☐ Summer
- ☐ Autumn
- ☐ Winter

Do you avoid cycling in bad weather or conditions?

(You can select multiple)

- ☐ Yes, when it rains
- ☐ Yes, when it snows
- ☐ Yes, in strong winds
- ☐ Yes, when it's slippery
- ☐ Yes, when it's cold
- ☐ Yes, when it’s poorly shoveled
- ☐ Yes, when there is a lot of gravel/leaves
- ☐ No, I bike whatever the weather

How long does it take, on average, when you cycle to the following one-way?

	Not relevant	0-10 minutes	10-20 minutes	20-30 minutes	30-45 minutes	45-60 minutes	60-90 minutes	More than 90 minutes
Work/studies	○	○	○	○	○	○	○	○
Business travel	○	○	○	○	○	○	○	○
Purchase/ errand	○	○	○	○	○	○	○	○
Entertainment/ leisure	○	○	○	○	○	○	○	○
Visiting friends/ family	○	○	○	○	○	○	○	○
Pick up/leave children	○	○	○	○	○	○	○	○

What is the maximum time you would be willing to spend to cycle to the following, one-way?

	Not relevant	0-10 minutes	10-20 minutes	20-30 minutes	30-45 minutes	45-60 minutes	60-90 minutes	More than 90 minutes
Work/studies	○	○	○	○	○	○	○	○
Business travel	○	○	○	○	○	○	○	○
Purchase/ errand	○	○	○	○	○	○	○	○
Entertainment/ leisure	○	○	○	○	○	○	○	○
Visiting friends/ family	○	○	○	○	○	○	○	○
Pick up/leave children	○	○	○	○	○	○	○	○

What do you think is the most important when choosing your bicycle route?

Rank them so that no one has the same value.

(1 being the least important and 10 being the most important)

Look below for explanations

	1	2	3	4	5	6	7	8	9	10
Distance	○	○	○	○	○	○	○	○	○	○
Time	○	○	○	○	○	○	○	○	○	○
Traffic safety	○	○	○	○	○	○	○	○	○	○
Few crossings	○	○	○	○	○	○	○	○	○	○
Separation from traffic	○	○	○	○	○	○	○	○	○	○
Consistent network	○	○	○	○	○	○	○	○	○	○
Smooth and solid surface	○	○	○	○	○	○	○	○	○	○
Safety	○	○	○	○	○	○	○	○	○	○
Surroundings / Environment	○	○	○	○	○	○	○	○	○	○
Facilities	○	○	○	○	○	○	○	○	○	○

Feel safe: Well-lit roads, no concealing shrubbery, etc.

Facilities: Toilets, bicycle pumps, shower / changing rooms, bicycle garage, water, etc.

Page 2/5

On the next page there will be a map where we want you to draw how you ride in the city and mark the places that you think works well or badly from a cycling point of view.

You can always get information on how to create and mark route selections by pressing the question mark at the top left on the next page.

Create route

1. Enter the address of your starting point. Alternatively, press cancel and find the starting point manually on the map and left-click.
2. Locate your target point and left click. Now you have created a start and destination point. A box will appear where you can describe the route, the type of errand and rate the route.
3. Edit the route by dragging the squares on the line, new squares will be created after each time you adjust the route. Try to be as accurate as possible. To edit the description, click on the line.
4. When you are finished with the route, you can create additional routes by pressing the pen with a plus.

Creating Selection
To create a mark on the map where you either think something is positive, negative, or otherwise, press the pin. Then zoom in and mark the location. A box appears where you can select what the marked route concerns and leave comments.

You can also make suggestions for improvements and/or review simplicity/road safety, then leave comment on which.

Page 3/5

How satisfied are you with the following in Malmö? (Scale 1-5)

(1 dissatisfied, 5 very satisfied)

	1	2	3	4	5
Malmö as a cycling city	○	○	○	○	○
The amount of bike lanes	○	○	○	○	○
Condition of bike lanes	○	○	○	○	○
Width of bike lanes	○	○	○	○	○
Diversion at roadwork	○	○	○	○	○
Bicycle parking at home	○	○	○	○	○
Bicycle parking at the shops	○	○	○	○	○
Bicycle parking at the workplace / school	○	○	○	○	○
Bicycle parking in public places	○	○	○	○	○
Bicycle parking at Malmö C, Triangeln and Hyllie	○	○	○	○	○
Bicycle parking at bus stops	○	○	○	○	○

What factors are most important when you park your bike for short errands?

(You can select three options)

- ☐ That the parking is monitored (cameras / guard)
- ☐ That the bicycle parking is locked
- ☐ That the parking is illuminated
- ☐ That there is a roof over the bicycle
- ☐ That I can lock the bicycle frame to a fixed point
- ☐ That the parking is located close to highly trafficked areas
- ☐ That the parking is near my destination point
- ☐ That there is plenty of room for the bicycle

What factors are most important for you to feel secure parking your bike for a longer time (a day or more)?

(You can select three options)

- ☐ That the parking is monitored (cameras / guard)
- ☐ That the bicycle parking is locked
- ☐ That the parking is illuminated
- ☐ That there is a roof over the bicycle
- ☐ That I can lock the bicycle frame to a fixed point
- ☐ That the parking is located close to highly trafficked areas
- ☐ That the parking is near my destination point
- ☐ That there is plenty of room for the bicycle

Can you imagine to park a little farther away from your destination point if the bicycle parking feels safe?

- ○ Yes, mostly
- ○ Yes, if the bike will be parked for a long time
- ○ No

How important are the following measures you as a cyclist?

(1 not important, 5 very important)

	1	2	3	4	5
Increasing security / safety at crossings	○	○	○	○	○
Increase the width of cycle lanes	○	○	○	○	○
Snow removal	○	○	○	○	○
Sweep away leaves	○	○	○	○	○
Sweep away gravel	○	○	○	○	○
Measures to separate pedestrians and cyclists	○	○	○	○	○
Remove level differences	○	○	○	○	○
Minimize the number of stops	○	○	○	○	○
Smooth surface	○	○	○	○	○
Keeping the shrubbery / trees	○	○	○	○	○
Police checks (eg check for lights, mopeds)	○	○	○	○	○

What materials do you consider to be suitable as a surface on bike paths?

(You can select multiple)

- ☐ Asphalt
- ☐ Gravel
- ☐ Hard packed gravel
- ☐ Tiles
- ☐ Cobble stone

Would you be willing to collect data for us using GPS, for example with an app on your phone?

- ○ Yes
- ○ No

How safe do you feel as a cyclist in Malmö?

Not safe	Partially safe	Safe
○	○	○

Page 4/5

Are you...?

- Female
- Male
- Prefer to not answer

How old are you?

- Under 13 years
- 13-17 years
- 18-24 years
- 25-39 years
- 40-64 years
- 65-75 years
- 76-84 years
- 85 or older

What is your family situation?

- Partner I live with
- Single / not cohabitating
- Living with parents or equivalent

How many children do you have living at home?

- None
- 1 child
- 2 children
- 3 children
- 4 or more children

What is your main occupation?

- Working
- Studying
- On sick leave
- Parental leave

- Seeking employment
- Pensioner
- Other

What is your highest completed education?

- Attending primary school
- Primary school
- High school
- Post-secondary education, other than college/university
- College/university

Income per month:

- Less than 10,000 SEK
- 10,000-14,999 SEK
- 15,000-24,999 SEK
- 25,000-34,999 SEK
- 35,000-50,000 SEK
- More than 50,000 SEK
- Do not wish to answer

Competition Question

Gothenburg has 'Styr & Ställ', Stockholm has its 'City Bikes' and Copenhagen has 'Bycykler'.

Contribute your creative suggestions on what a rental bike service could be named in Malmo. The best proposal will be rewarded with an electric bike and another two people with an inflatable helmet for cyclists.

(Do you not want the chance to win, you can skip this section)

Your suggestion:

Enter the name and contact information, e-mail and / or phone number so we can reach you if you win.

(Do you not want the chance to win, you can skip this section)

Name:

E-mail

Phone:

Page 5/5

Appendix C- Postal codes considered in the analysis