MASTER

Monitoring Cycleability
A Management Dashboard for Bike-Friendly Cities

Hordijk, Marjolein A.

Award date:
2021

Monitoring Cycleability:

A Management Dashboard for Bike-Friendly Cities

Graduation Thesis

Marjolein A. Hordijk

Submission date:	19-09-2021	
Submission year:	Academic year 2021/2022	
Publication date:	24-09-2021	
Author:	Marjolein A. Hordijk (0855820)	
E-mail:	m.a.hordijk@student.tue.nl // marjolein.hordijk@icloud.com	
Study load:	45 ECTS	
Faculties:	Build Environment (BE)	Industrial Engineering & Innovation Sciences (IEIS)
Programs:	MSc Architecture, Building and Planning	MSc Innovation Management
Tracks:	Urban Systems and Real Estate	Free
Course codes:	7Z45M0	1ZM96
Supervisors:	Gamze Dane Aloys Borgers	Néomie Raassens Sarah Gelper

Keywords: *cycling, GPS data, route choice, revealed choice, Path-Size Logit, Latent Class Analysis*

This Master's thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Integrity.

The digital repository for this thesis is available here.

ABSTRACT

This study analyzes GPS data from Utrecht's ikFiets mobile phone application to gain a better understanding of the relation between the urban environment and route choice behavior. The GPS data is enriched with open data to capture the characteristics of the routes. Choice sets with realistic alternative routes are generated using the routeplanner of De Fietsersbond. A Path Size Logit (PSL) is applied to quantify the preferences of cyclists. Interaction effects are introduced to account for varying preferences across departure times and trip purposes. Finally, a Latent Class Analysis is used to study preference heterogeneity. The results of the Latent Class Analysis (LCA) reveal two segments of cyclists with distinct preferences. The first segment seems to be particularly concerned with convenience and safety. The second segment has a much higher intention to cycle, is more likely to own a race bike and has a strong preference for green surroundings. Further, preferences for traffic lights are found to differ across on and off peak situations. To add, the attractiveness of green surroundings seems to depend on the trip purpose, but only for the first segment. A dashboard is developed to help policy developers understand these dynamics and plan effective interventions for each identified segment.

MANAGEMENT SUMMARY

Background: Active travel is argued to counteract a variety of challenges faced by Western societies, such as obesity, congestion and pollution. Consequently, planners and policy makers increasingly insist on the development of urban environments which facilitate and stimulate active transportation. The growing availability of crowd-generated GPS data presents an interesting opportunity to develop an understanding of how the urban environment influences cycling behavior. The current study leverages GPS data from Utrecht's ikFiets mobile phone application to answer the question; "How do built environmental and infrastructural characteristics influence route choices of cyclists in the municipality of Utrecht?". Further, it studies preference heterogeneity related to several personal characteristics, trip departure time and trip purpose. The results are presented in the form of a dashboard which could support Utrecht's policy makers to make certain infrastructure more appealing.

Methods: This study applies a Path Size Logit (PSL) model to study the route choices of the ikFiets sample. The GPS data is enriched with open data to capture the characteristics of the routes. Further, choice sets with various realistic alternative routes are generated using the

routeplanner of De Fietsersbond. Interaction effects are introduced to account for varying preferences across departure times and trip purposes. A Latent Class Analysis (LCA) is conducted to identify two segments of cyclists with distinct preferences.

Results: The results reveal two distinguishable segments of cyclists. The first group is characterized by their tendency to stick to the shortest route. They have relatively strong preferences when it comes to intersections, turns, speed limits and traffic lights. As such, it seems that this group is particularly concerned with convenience and safety. More specifically, they avoid traffic lights in general, but less so during peak hours, when signals may provide them with safe and efficient passage through heavy traffic. Further, they avoid agricultural surroundings during commutes, but not during leisure trips. They have a relatively low intention to cycle and are less likely to report to cycle because they enjoy it. Hence, they appear to consider a bike to be a mode of transport. The second group is willing to detour substantially more in comparison. These cyclists are more keen on green surroundings, regardless of their trip purpose. They appear to be the more advanced cyclists who are more likely to own a race bike and have a relatively high intention to cycle. To add, they report to cycle because they like being outside, it increases their physical and mental health and they simply enjoy it.

Conclusions: The results highlight that preferences of cyclists are not homogene. Moreover, preferences may differ across trip contexts. The developed dashboard helps policy developers understand these dynamics and plan effective interventions for each identified segment.

TABLE OF CONTENTS

1. Introduction

1.1 Context

Active travel is argued to counteract a variety of challenges faced by Western societies, such as obesity, congestion and pollution (Donnelly et al., 2009; Handy, Van Wee and Kroesen, 2014). In this regard, cycling poses a promising alternative to motorized traffic. That is, cycling is relatively fast and covers a larger radius when compared to walking. Meanwhile, it requires healthy exercise and produces no air pollution or noise disturbance. Consequently, planners and policy makers increasingly insist on the development of urban environments which facilitate and stimulate active transportation (Handy et al., 2014).

However, the development of effective policies to stimulate cycling is a complex process. As advocated by Handy et al. (2014), the assessment of cycling policies relies on close monitoring of the developments in cycling behavior. Particularly, it requires a clear understanding of the stimulants and deterrents of cycling. Handy et al. (2014) point to two types of studies required to guide policy makers in this process. First, *cross-sectional research* is necessary to identify aspects of the environment and infrastructure which influence cycling behavior. As becomes evident from the literature review and confirmed by Bernardi et al. (2018), this is often done through sending out stated preference surveys. Second, *longitudinal studies* are needed to evaluate the effectiveness of interventions. This requires extensive data collection at multiple points in time and is therefore inherently expensive and time consuming. Consequently, monitoring cycling behavior to aid policy development poses a heavy burden on municipalities.

Moreover, municipalities are often forced to conduct these studies themselves instead of relying on international theory and literature, as a consequence of the heterogeneity in cycling habits. To illustrate, active travel, including cycling, is much more popular in Europe compared to the U.S. (Donnelly et al., 2009). Further, Pucher and Buehler (2008) highlight the superiority of the Dutch, Danish and German cycling infrastructure in comparison to that of the U.K and U.S. To add, Mertens et al. (2017) shows substantial variation in utilitarian cycling rates across European countries. Moreover, variation in cycling habits is not only evident across continents and countries, but even between municipalities (Glaser and te Brömmelstroet, 2020). The literature indicates that these differences in cycling behavior, at least partially, originate from environmental and infrastructural variation across locations (Mertens et al., 2017). In the light of the above, it is difficult for municipalities to translate findings on stimulants and deterrents of cycling across contexts. Thus, they have to invest heavily in local studies.

Local governments could therefore benefit from a more efficient method to study cycling behavior. The growing availability of crowd-generated GPS data presents an interesting opportunity to tackle this challenge. That is, GPS data from mobile apps can be used to study which factors influence route choice behavior of citizens and to what extent they do so. Citizens can record their cycling movements using their own mobile phone. Hence, no surveys or GPS devices need to be distributed. Examples include the globally available apps Strava (n.d.) and BikeCitizens (n.d.), but also smaller local initiatives such as Moves (Pritchard et al., 2019), CycleTracks (Hood, Sall and Charlton, 2011; Melson, Duthie and Boyles, 2014) and ikFiets (Goedopweg, n.d.). The developers of these apps attract users by offering them insights in their own activities, the ability to share their activities on a social media platform and/or eligibility to receive promotions and prices. In some cases (e.g. Strava), municipalities can buy access to the GPS data. However, initiatives such as Moves, CycleTracks and ikFiets illustrate that local governments can also develop these applications themselves to gain direct access. In short, crowd-generated GPS data is becoming increasingly available to municipalities.

This novel approach, to use GPS data when studying cycling behavior, has successfully been applied in several studies, particularly in an American context. For example, Melson, Duthie and Boyles (2014) studied how the layout of bridges influence route choice behavior of cyclists in Texas, based on GPS data. Broach, Dill and Gliebe (2012) used GPS data to determine which infrastructural aspects attract and repel Oregon's cyclists to certain streets. The same has been done in California (Hood, Sall and Charlton, 2011), Ohio (Park and Akar, 2019) and Washington (Chen, Shen and Childress, 2018). Although some European studies exist (e.g. Menghini, Carrasco, Schüssler and Axhausen, 2010; Prato, Halldórsdóttir and Nielsen, 2018; Skov-Petersen, Barkow, Lundhede and Jacobsen, 2018), these remain exceptions. These studies indicate that GPS data can indeed supplant stated preference surveys.

Moreover, crowd-generated GPS data provides several benefits over the use of surveys. First, GPS datasets record the actual behavior of cyclists. It therefore circumvents the hypothetical bias documented for stated choice experiments (Murphy et al., 2005). That is, in some cases the hypothetical choices in a survey may sufficiently resemble real life decisions. For example, simple consumer products showcased together as if they were presented in a webshop may induce realistic choice strategies. In other situations, the alternatives must be experienced in reality to truly grasp the implications of their differences. Route choices fall in the latter category. Descriptions, illustrations, pictures or videos can be expected to fall short of capturing the true experience of a location and the context of the choice situation. In those cases, a stated preference survey will not be able to replicate a real decision. Hence, choices made in the survey can be expected to differ from choices made in real life, as Murphy et al. suggest (2005).

Moreover, GPS data does not rely on recall and is less sensitive to self-censoring (Larsen and El-Geneidy, 2011). It therefore provides a more accurate and complete view of someone's cycling habits. Further, availability of crowd-generated data has made GPS data a low cost alternative to surveys. In sum, the use of crowd-generated GPS data is an efficient way for municipalities to capture realistic choice behavior of cyclists.

1.2 The Municipality of Utrecht

A municipality that could benefit substantially from this approach is the Dutch city of Utrecht. According to “Actieplan Utrecht fietst!” (Gemeente Utrecht, 2015), a publication on the ambitions of the municipality, cycling plays a key role in maintaining a pleasant living environment in the city. Meanwhile, the city’s population is growing and an increasing number of visitors and tourists find their way to its historic center. This makes for a growing pressure on the cycling infrastructure, particularly during rush hours. Hence, the municipality is faced with the difficult task of maintaining the quality and efficiency of its infrastructure, whilst also striving for a growing user base. Monitoring the cycling infrastructure is therefore of the utmost importance to Utrecht.

However, the unique historical Dutch cycling culture makes it difficult for the municipality to leverage on findings of international studies. To illustrate, Pucher and Buehler (2008) rank The Netherlands among the leading countries when it comes to the quality of its cycling infrastructure. Further, considering its high density and flat topography, cycling often poses a suitable alternative to car travel. Indeed, the estimated number of bikes in The Netherlands exceeds the number of inhabitants (Statista, 2020), underlining the popularity of cycling among the Dutch. In comparison, less than one out of eight people in the United States cycles on a regular basis (Statista, 2021). Moreover, the city of Utrecht presents itself as the cycling city of the Netherlands. It sees the bike as “the symbol” for the city they want to be (Gemeente Utrecht, 2015, p.2). This strong focus on cycling in local policies makes Utrecht unique and difficult to compare even to other Dutch cities. The municipality of Utrecht therefore invests heavily in its own research departments and projects.

An example of Utrecht’s efforts to study cycling behavior is its collaboration with other cities in the region and the provincial government to develop the ikFiets app. This mobile application allows inhabitants of the province of Utrecht to record their cycling movements in exchange for rewards, insights and compelling challenges. The GPS data generated by these users provides Utrecht with a unique insight into the cycling habits of over a thousand of its citizens. Moreover, existing users keep generating data. Hence, the municipality continuously receives new data without much additional effort beyond the initial investment, other than some periodical campaigns to promote the application. Thus, Utrecht has continuous access to valuable crowd-generated GPS data from its inhabitants.

In short, Utrecht is in the perfect position to leverage upon the possibilities of crowd-generated GPS data to support the development of policies which stimulate cycling. Their strong focus on developing an attractive cycling infrastructure underlines the value of the insights that such analysis could provide to them. Moreover, they have access to a large stream of GPS data from the ikFiets app. It is therefore an ideal case to demonstrate how crowd-generated GPS data can be translated into valuable insights.

1.3 Research Design

Considering the above, the goal of the current study is twofold. First, it strives to identify and rank the aspects of the built environment which encourage and discourage cyclists in Utrecht. To achieve this, a route choice model is estimated based on GPS data. Specifically, a Path Size Logit (PSL) model is estimated on a total of 5091 regular trips made by 204 users of the ikFiets app. This model compares the attributes of each chosen route to a set of alternatives, which are generated using the routeplanner of the Dutch national cycling association (de Fietsersbond). Second, the study aims to illustrate how a municipality can leverage crowd-generated GPS data to support policy makers in their efforts to stimulate cycling. In this regard, a dashboard is developed based on the PSL estimates. The main performance indicator is a composite cycleability index, which can be broken down into more specific indices. The weights in this index are based on the relative importance of each attribute in the PSL model. This dashboard provides insight into the performance of the network, as well as possible causes of bottlenecks.

The goals introduced above translate into one main research question and two sub questions to be answered by the current study:

How do built environmental and infrastructural characteristics influence route choices of cyclists in the municipality of Utrecht?

A. How do these relations differ based on personal characteristics?
B. How do these relations differ based on trip context, in terms of departure time and trip purpose?

1.4 Academic Relevance

The literature review on indicators of cycleability illustrates the scarcity of Dutch studies on this topic. Several Dutch studies relate built environment aspects of home locations to the frequency or duration of cycling, usually in the context of active travel or mode choice (e.g. Gao et al., 2018; Noordzij et al., 2021; De Vries et al., 2010). However, only a handful of studies focus on route choice. An example of those studies that comes close to the current one is Bernardi et al. (2018), which analyzed 3,500 bike trips across the Netherlands using several extensions of the MNL model, including a Path Size Logit model. However, this study does not focus on just one city and therefore does not acknowledge the likelihood of intercity variation reported by Glaser and te Brömmelstroet (2020). Moreover, it does not distinguish between trips in rural and urbanized areas. Following the conclusions of Mertens et al. (2017), one would expect cycling habits to differ between those two distinct urbanization patterns. If the findings of the national study by Bernardi et al. (2018) and those of the current study differ substantially, the need for local studies is underscored. Thus, a route choice study in the context of the Dutch city of Utrecht poses a valuable contribution to the current literature.

Furthermore, the current study illustrates how crowd-generated data can be applied in route choice studies. Although this has been done before, stated preference surveys still remain the standard (Bernardi et al., 2018). Moreover, studies based on crowd-generated data are often conducted in the U.S. However, revealed choices from GPS data pose several benefits over stated preference studies. As discussed earlier, the latter are subject to a hypothetical bias, self-censoring and/or recall (Murphy et al., 2005; Larsen and El-Geneidy, 2011). They thus do not guarantee an accurate and complete impression of cycling habits. GPS data circumvents these issues. Particularly in the light of the growing availability of GPS data, demonstrating its usefulness in route choice studies is therefore another interesting contribution of the current study.

1.5 Practical Relevance

As advocated at the start of this introduction, municipalities have to invest substantially to gain insights into the stimulants and deterrents of cycling among their citizens. In this light, the municipality of Utrecht joined forces with other local governments to develop the ikFiets application. This mobile phone app generates a substantial amount of GPS data. The current study unlocks the potential of this data to support the development of an attractive cycling infrastructure in Utrecht. On the one hand, the developed dashboard provides insights into the performance of the infrastructure, bottlenecks and their potential causes. This formation can help Utrecht to reach its goal of becoming the cycling city of the world (Gemeente Utrecht, 2015). On the other hand, the study also demonstrates how the municipality could leverage the incoming stream of GPS data to monitor its cycling infrastructure in the future. That is, the dashboard can be considered as a static prototype for future efforts and could be extended to show the evolution of the cycling infrastructure over time. Moreover, the cycleability index could be used to evaluate the effects of future interventions. Hence, this first attempt to build a dashboard based on GPS data from the ikFiets app can be a valuable lesson for the future.

1.6 Reading Guide

The remainder of this report is structured as follows. First, the literature review (Chapter 2) summarizes the findings of over forty reviewed articles on bike route choice behavior. The main goal of this review is to provide an overview of the commonly studied indicators of cycleability in the literature. That is, it shows which factors are known to influence the route choices of cyclists and in what manner they do so. To add, indications of preference heterogeneity and the impact of trip context are discussed. Thereafter follows an overview of the current state of route choice modelling (Chapter 3). This chapter introduces important concepts related to route choice models and discusses multiple modelling techniques. Next, the methodology of the current study is outlined in Chapter 4. This includes the data collection and preparation process, choice set generation, model specification and the development of a dashboard. Chapter 5 provides a summary of the descriptive statistics regarding the demographic data of the sample, network and trip characteristics and the generated route alternatives. Next, Chapter 6 presents the results of the study. The chapter starts with a discussion of several correlation matrices, which guided the process of defining the final model. Further, it presents the results of several intermediate models and the final Path Size Logit model and Latent Class Analysis. As an elaboration on the latter, a comparison of the two identified classes is included. Finally, Chapter 7 summarizes the findings and discusses their practical and theoretical implications. To add, the limitations of the study and recommendations for future research efforts are discussed. Scripts which were used to prepare or analyse the data are included in the digital repository (see here).

2. Literature Review on Indicators of Cycleability

To gain insight into the current literature on stimulants and deterrents of cycling, a literature review is conducted. This chapter starts with briefly describing the review strategy. Next, an overview of the reviewed articles is provided in terms of research methods and origins. Thereafter, the conclusions of the reviewed articles regarding eight themes are summarized and contrasted. Each section on a theme is concluded with a table that summarizes the main take-aways (Tables 2.2 - 2.15).

2.1 Review Strategy

This literature study follows the guidelines on conducting a systematic literature review by Okoli and Schabram (2010). The complete review process is elaborated upon in Appendix I and summarized in Figure 2.1.

Planning	Selection	Extraction	Execution
Purpose of review: • Develop list of commonly studied indicators and their effects on cycleability. • Summarize variations among different types of cyclists. **Protocol:** • Define search criteria. • Define eligibility criteria. • Define extraction topics.	**Search criteria:** • No grey literature. • Dates from 1990 onwards. • Only published or in press. • Only English or Dutch. • SCimago Journal Rank Indicator Q1 or Q2 in transport category. **Eligibility criteria:** • Study should evaluate relationships between route characteristics and cycling behavior. • Infrastructural, built environment and natural characteristics. • Solely objective measures.	**Extraction strategy:** • Extract data on prespecified topics. • Cross-reference findings on route characteristics. • Study differences among demographic clusters / trip purposes. **Extraction Topics:** • Study design. • Study sample. • Route characteristics with significant effects on cycling behavior. • Nature of these effects and mechanisms behind them. • Differences among types of cyclists.	**Analysis of finding:** • Review studied indicators and their effects. • Review reported differences among types of cyclists. • Study differences & coherence among studies. **Writing the review** • Summarize findings. • Provide final list of indicators and expected effects. • Report potential differences among types of cyclists.

Figure 2.1 - Summary Literature Review Strategy

The goal of this literature study is twofold. First, it should produce a list of commonly studied indicators of cycleability, their reported relation to route choices and an indication of the magnitude of their effects. Second, the review should provide insight into the reported variations among different types of cyclists when it comes to the effects of the indicators. An understanding of these differences helps determine which personal characteristics should be considered during the current study. These goals are summarized in Table 2.1.

Table 2.1 - The Literature Review Goals	
1	**Develop a list of commonly studied indicators and their reported effects on cycleability.** ➔ Review the literature for the most commonly studied indicators. ➔ Summarize the nature of the relations.
2	**Summarize the variations in effects for different types of cyclists.** ➔ Review the literature for reported variations based on personal characteristics / trip context. ➔ Identify those personal characteristics that commonly capture distinct preferences and therefore warrant special attention.

2.2 Types of Studies

A tabular overview of the selected literature, including the research methods, type of sample and categorized findings of each article is included in the digital repository ('literatureSearch.xlsx'). As shown in Figure 2.2, most studies were conducted in the US, followed by Canada and Asian countries. Further, the strategy resulted in only one Dutch paper. Two studies compare two different countries.

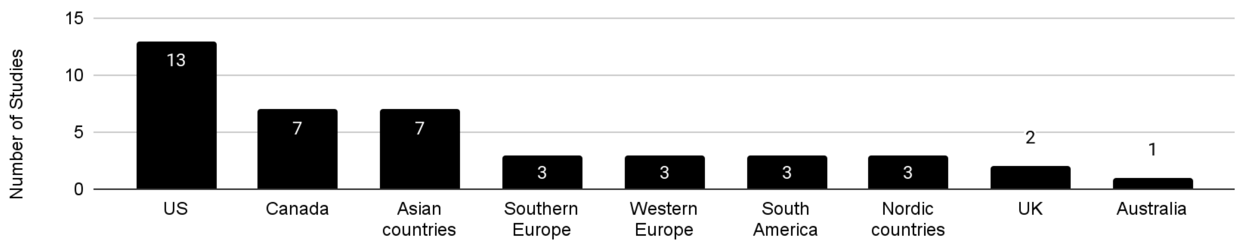


Figure 2.2 - Origin of Samples

As shown in Figure 2.3, Stated preference surveys appear to be most popular, followed by revealed preference experiments. These two approaches are sometimes combined with interviews or census data to validate the findings. One study combines stated and revealed preferences (Fitch and Handy, 2020), although not among the same sample. Most studies end with an overview of the preferences of the sample. Preference heterogeneity or context dependency are usually modelled as interactions in the models. Several articles translated the preferences to an index which scores the local infrastructure. This index is usually visualized on a static map, such as done by Arellana, Saltarín, Larrañaga, González and Henao (2020). Results are sometimes used for forecasting or traffic assignment models, as done by Arellana et al. (2020) and Duc-Nghiem, Hoang-Tung, Kojima and Kubota (2018).

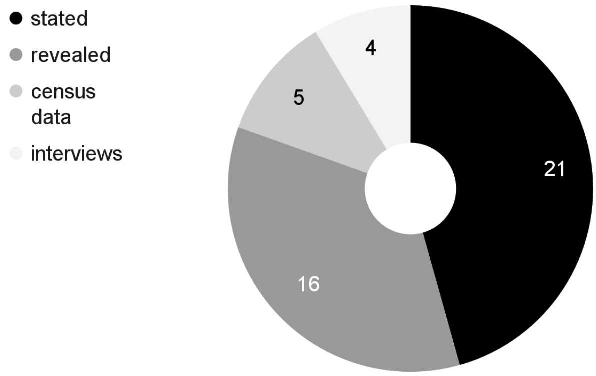


Figure 2.3 - Research methods

Figure 2.4 provides an overview of the themes that are covered by the selected articles. These themes correspond to those in the tabular format for data elicitation (see 'literatureSearch.xlsx' in the digital repository), where one can also find a list of subtopics that belong to each theme. Bike facilities, street layout, travel related concerns (e.g. travel time and distance) and nature and ambience are clearly recurring topics in many of the articles.

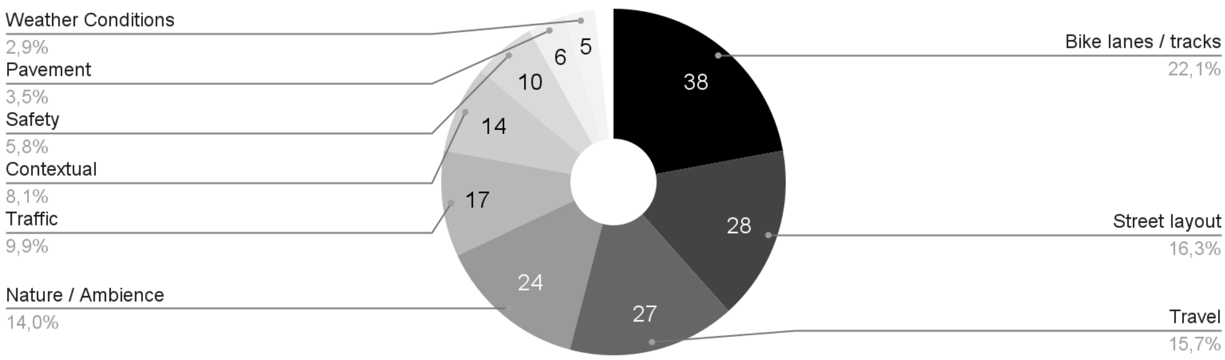


Figure 2.4 - Themes in Selected Studies

2.3 Cycling Facilities

A large body of literature underscores the appeal of dedicated bike facilities for most cyclists. For example, Winters, Davidson, Kao and Teschke (2011) ranked bike lanes, both off-street and on-street, among the top three motivators for cycling in Vancouver (Canada). To add, Manaugh, Boisjoly and El-Geneidy (2017) concluded that off-street bike paths strongly stimulate students and university staff members of a Canadian university to cycle to campus. Likewise, Parkin, Wardman and Page (2008) discovered that a high proportion of off-road bike facilities goes hand in hand with high levels of bike commuting in England and Wales. Orellana, Guerrero (2019) made the same discovery among cyclists in Ecuador. Moreover, Standen, Crane, Collins, Greaves and Rissel (2017) conclude that cyclists in Sydney (Australia) are likely to change their standard route

to incorporate a new bike track. The same appears to be true for Norwegian cyclists who responded to the introduction of a new two-way bike lane in Oslo (Pritchard, Bucher and Frøyen, 2019). Similarly, research shows that cyclists in Tel Aviv (Israël) (Ghanayim and Bekhor, 2018) and Columbus (America) (Park and Akar, 2019) are willing to detour substantially to include an existing bike track in their route. Furthermore, Chen, Shen and Childress (2018) found a strong preference for bike lanes over cycling in mixed traffic for cyclists in Seattle. The same holds for cycle tracks among commuting riders in Copenhagen (Denmark), according to Vedel, Jacobsen and Skov-Petersen (2017). Likewise, Lu, Scott and Dalumpines (2018) conclude that Canadian cyclists are attracted to safe on-street and off-street bikeways and clearly avoid busy streets without said facilities. In short, cyclists across the globe appear to favor routes with dedicated bike facilities.
However, as corroborated by the below, substantial preference heterogeneity seems to exist for the exact layouts of these facilities. Specifically, several socio-demographic clusters seem to have distinct wants and needs. Moreover, contextual factors, such as traffic volumes and speed limits, may also affect the preferences of riders. Thus, it is important to consider these nuances when studying preferences for cycling facilities.

2.3.1 On-Street Versus Off-Street Facilities

Reports in the literature on preferences for on-street versus off-street bike facilities are mixed. Several studies conclude that cyclists prefer off-street facilities over on-street ones, because the former evoke a sense of safety (Hopkinson and Wardman,1996; Krizek, El-Geneidy and Thompson, 2007; Parkin, Wardman and Page, 2008; Hunt and Abraham, 2007). Specifically, Hopkinson and Wardman (1996) found that U.K. cyclists are willing to trade shorter travel times for safe, segregated facilities. This preference for off-road facilities also seems to hold for commuters in both the U.K and Wales, as reported by Parkin, Wardman and Page (2008). Krizek, El-Geneidy and Thompson (2007) provide similar results for American riders and also conclude that cyclists are more tolerant of intersections when a route includes a separate bike path. Indeed, Hunt and Abraham (2007) drew similar conclusions based on their Canadian study. They argue that previous research confirms a cyclist's perceived risk of collision is reduced when cycling on a seperate bike path. Park and Akar (2019) confirm these findings for cyclists in Columbos and reason that off-street facilities not just offer safety, but also comfort. Consistent with the above, Skov-Petersen, Barkow, Lundhede, and Jacobsen (2018) report that separate bike tracks strongly influence the route choices of cyclists from Copenhagen. Further, Deenihan and Caulfield (2015) show that the need for off-street facilities is also evident among tourists in Ireland. The above indicates that the preference for off-street paths is present across a variety of nationalities.

To add, Melson, Duthie and Boyles (2014) concluded that cyclists in Texas are more strongly attracted to bridges that have physically separated facilities. The work of Broach, Dill and Gliebe (2012) also provides indications that Portland's cyclists are attracted to bridges with segregated facilities. Further, they report a strong preference for off-street, dedicated bike boulevards. Hence, it seems that the need for separate facilities extends beyond the default street infrastructure. All in all, these studies indicate a general preference for off-street facilities over on-street ones among cyclists.

However, this general consensus is not unanimous. To illustrate, a handful of studies claim that cyclists actually prefer on-street facilities. For example, Sener, Eluru and Bhat (2009) suggest that cyclists are more attracted to on-street facilities because these provide them with space to maneuver and psychological freedom. However, it is important to note that this study solely refers to commuters. Prato, Halldórsdóttir and Nielsen (2018) also focussed on commuters and found a strong preference for bike lanes in Copenhagen (Denmark), which surprisingly did not hold for seperate bike tracks. However, the authors explain that Copenhagen's bike lanes are highly available and provide sufficient separation from motorized traffic. The city's separate bike tracks, on the other hand, are commonly unpaved and therefore less attractive. Rossetti, Saud, and Hurtubia (2019) also report a general preference for on-street facilities among Santiago's (Spain) cyclists. However, they recorded substantial heterogeneity in their results. Hence, the effect might simply be a sum of the preferences of different types of cyclists. Moreover, some authors argue for the absence of any noteworthy preference for off-street or on-street facilities altogether (Buehler, Pucher, 2012; Fitch and Handy, 2020). In this case, the preferences of particular clusters might cancel eachother out. In short, understanding the preferences for off-street versus on-street bike facilities requires a closer inspection of specific target groups.

Interestingly, several authors studied the preferences for off-street and on-street in greater detail to indeed discover that particular target groups have specific needs. Furthermore, these preferences also appear to be context-dependent to some extent. This may explain the divergent findings of the aforementioned studies and the reported preference heterogeneity.

Frequent Versus Infrequent Cyclists

For example, frequent and infrequent cyclists display distinct preferences across multiple studies from various countries. To illustrate, Rossetti, Saud and Hurtubia (2011) discovered that infrequent cyclists in Santiago (Spain) strongly favor on-street facilities, possibly because their safety concerns are higher in comparison to experienced riders. In confirmation, Arellana, Saltarín, Larrañaga, González and Henao (2020) derived that infrequent Colombian cyclists seem particularly fond of a safe, separate bike infrastructure. Moreover, Rossetti et al. (2011) concluded that experienced cyclists are less sensitive to the absence of segregated bike facilities and are also less inclined to use sidewalks if said facilities are not available. They argue that experienced cyclists are more comfortable cycling near or between motorized traffic. Hunt and Abraham (2007) provide similar findings for Canadian cyclists. They reason that experienced cyclists might be in a better physical condition and can therefore more easily adapt to the speed of motorized traffic. In short, several studies underscore the importance of off-road facilities to inexperienced cyclists in particular.

Surprisingly, several studies report reverse effects. For example, Hood, Sall and Charlton (2011) report that frequent cyclists from San Francisco are more strongly attracted to streets with off-road facilities in comparison to infrequent cyclists. They argue that the stated preference approach of other studies might have clouded their results due to the overrepresentation of "vehicular cycling" promoters in their sample. Indeed, Rossetti et al. (2011) themselves highlight that their sample contains many experienced leisure cyclists, due to their promotion strategy. Due to their experience, these people are likely more comfortable cycling between traffic. The revealed preference method applied by Hood et al. (2011) is deemed to be less sensitive to this bias, because it studies actual route choice behavior instead of stated choices. Interestingly, the studies cited in the preceding paragraph are indeed based on stated choice experiments. Moreover, Melson, Duthie and Boyles (2014) drew a coherent conclusion in the context of Texan bridges, also based on revealed choices derived from GPS data. These authors reason that infrequent cyclists likely stick to a standard route and therefore do not deliberately go for a bridge with separate facilities. In sum, stated preference studies generally report a relatively strong preference for off-road facilities among infrequent versus frequent cyclists, whereas revealed choice studies conclude the opposite. Thus, there appears to be a lack of consensus regarding the preferences of (in)frequent cyclists for off-road versus on-road facilities. It seems that this inconsistency stems from the methodological differences.

Gender Specific Preferences

There are also indications in the literature for gender-specific preferences. For example, Standen, Crane, Collins, Greaves and Rissel (2017) concluded that Australian females have a stronger tendency to switch routes upon the introduction of a new off-road bike path. The authors attribute this finding to the fact that females are more risk averse. Therefore the introduction of a safe, separated cycleway is a stronger incentive to them in comparison to men (Garrard, Rose and Lo, 2008). The same seems to hold for Japanese cyclists, as suggested by the findings of Duc-Nghiem, Hoang-Tung, Kojima and Kubota (2018), who argue that men are less troubled by having to cycle in mixed traffic conditions. Likewise, Vedel, Jacobsen and Skov-Petersen (2017) conclude that female commuters from Denmark have a strong need for separate bike facilities, more so than the men in their sample. Deenihan and Caulfield (2015) also describe a male versus female distinction, specifically among tourists in Ireland. They found that female tourists are strongly discouraged by the absence of dedicated cycling facilities, whereas males were far less sensitive to this matter. However, female tourists appeared to have no specific preference for off-road or on-road facilities. In contrast, Rossetti, Saud and Hurtubia (2019) recorded that male cyclists in their sample have a stronger preference for separate facilities in comparison to females. However, they note that the women in their sample are mainly experienced cyclists, which is a likely cause of this divergent observation. All in all, it appears that female cyclists are generally more sensitive to the absence of a seperate, safe bike infrastructure, in comparison to men.

Impact of Traffic Volumes

Kang and Fricker (2013) identified an interesting nuance to the commonly observed preference for off-street facilities. Based on their human intercept survey at a university campus, they concluded that cyclists prefer on-street facilities in low traffic volume situations, whereas off-street paths become more popular along major arterials. They argue that streets around Purdue University (Indiana, America) generally feature high quality pavement and therefore allow cyclists to maintain high speeds and offer them a comfortable ride. Only in extreme cases, where high traffic volumes pose a serious risk to cyclists, did they move to the safety of a separate facility. This finding does not stand on its own. Broach, Dill and Gliebe (2012) also conclude that completely separate bike paths are particularly popular under high traffic volume conditions. More specifically, they report that the preference for off-road facilities among Portland's (Oregon, America) cyclists diminishes when traffic volumes are low. This indicates that separate bike paths simply offer cyclists protection against motorized traffic, but they are not more appealing than bike lanes in other aspects. In short, there are indications in the literature that preferences for off-road bike facilities are reduced under low traffic volume conditions.

Impact of Trip Purpose

Route choices for off-street and on-street bike facilities also seem to be affected by the purpose of the trip at hand. For example, Deenihan and Caulfield (2015) identified a preference for off-street cycle paths among tourist leisure cyclists. They discovered that tourists in Ireland are willing to take a detour of twice the original travel time to include segregated facilities in their rides. Moreover, tourists are willing to give up comfort in terms of steeper slopes in return for these facilities. In contrast, Duc-Nghiem, Hoang-Tung, Kojima, Kubota (2018) report that Japanese mountain bikers and race bike users are more likely to use on-street facilities in comparison to other cyclists. These sportive cyclists seemingly belong to a distinct breed of leisure cyclists. Further, bike commuters appear to be almost insensitive to the presence of separate bike facilities. As suggested by Arellana, Saltarín, Larrañaga, González, Henao (2020), these riders are more concerned with other factors such as safety, comfort and efficiency. These authors argue that commuters might be more inclined to stay on primary roads, because these provide efficient routing. Moreover, commuters generally face stronger time constraints in comparison to leisure cyclists. Hence, they can be expected not to deviate substantially from the shortest route to include an off-road facility in their route. Indeed several studies confirm that minimizing travel distance or time is very important to bike commuters (Sener, Eluru and Bhat, 2009; Parkin, Wardman and Page, 2008; Broach, Dill and Gliebe, 2012; Vedel, Jacobsen and Skov-Petersen, 2017; Anowar, Eluru and Hatzopoulou, 2017;) and less so for leisure cyclists (Chen, Shen and Childress, 2018; Bernardi, Geurs and Puello, 2018; Fitch and Handy 2020). This indicates that commuters do not necessarily dislike off-road facilities. That is, they simply put more value in time and efficiency. Hence, if the availability of off-road bike paths is low, they do not pose a suitable alternative to commuters looking for a direct connection. The logic for commuters seems transferable to utilitarian cyclists in general, as suggested by the findings of Bernardi, Geurs and Puello (2018). These authors concluded that main roadway links were more popular among their mainly utilitarian sample in comparison to separate facilities such as bike boulevards. In short, it appears that cyclists who are bound by stricter time constraints, such as commuters and utilitarian cyclists, are less likely to detour for seperate bike facilities in comparison to leisure cyclists.

2.3.2 Lane Width

Surprisingly, only a handful of studies describe preferences for lane widths in the context of cycling facilities. For example, Providelo and Da Penha Sanches (2011) concluded that Bazilian students and staff members consider wide lanes as an important factor in route choice. To add, a Spanish study among students by Rossetti, Saud and Hurtubia (2019) reports that lane width is only deemed relevant in an on-street scenario. However, several authors incorporate lane width indirectly. For example, Kang and Fricker (2013) and Kang and Fricker (2018) applied the bicycle compatibility index (BCI), originally developed by Harkey, Reinfurt, Knuiman, Stewart and Sorton (1998). This index measures perceived risk for cyclists and is, among other things, based on lane width. The wider a bike lane, the more safe it is considered to be. Moreover, lane width can be expected to covariate with several other factors such as the type of road (local vs. arterial) and traffic volumes. To add, there might be national regulations or habits in place which dictate a certain width, thus limiting the variation in width encountered by cyclists. Hence, it could be difficult to observe a preference for a particular lane width, particularly in revealed preference studies. Nevertheless, it can be expected that cyclists prefer wider lanes, because these provide them with sufficient room to move around when needed.

Table 2.2 - Preferences for Off vs. On-street facilities		**Reference**
General	Cyclists across the globe appear to favor routes with dedicated bike facilities. However, substantial preference heterogeneity seems to exist for the exact layouts of these facilities. These preferences appear to depend on both personal characteristics as well as trip context.	Hopkinson and Wardman (1996), Krizek, El-Geneidy and Thompson (2007), Parkin, Wardman and Page (2008), Hunt and Abraham (2007), Skov-Petersen, Barkow, Lundhede, and Jacobsen (2018)
Cycling frequency	There appears to be a lack of consensus regarding the preferences of (in)frequent cyclists for off-road versus on-road facilities. It seems that this inconsistency stems from methodological differences.	Rossetti, Saud and Hurtubia (2011), Larrañaga, González and Henao (2020), Hunt and Abraham (2007), Hood, Sall and Charlton (2011), Melson, Duthie and Boyles (2014)
Gender	It appears that female cyclists are generally more sensitive to the absence of a seperate, safe bike infrastructure, in comparison to men.	Standen, Crane, Collins, Greaves and Rissel (2017), Garrard, Rose and Lo (2008), Duc-Nghiem, Hoang-Tung, Kojima and Kubota (2018), Vedel, Jacobsen and Skov-Petersen (2017), Deenihan and Caulfield (2015)
Traffic volumes	There are indications in the literature that preferences for off-road bike facilities are reduced under low traffic volume conditions.	Kang and Fricker (2013), Broach, Dill and Gliebe (2012)
Trip purpose	Tourist leisure cyclists appear to have a strong preference for off-street facilities.	Deenihan and Caulfield (2015),
	Sporty cyclists on mountain bikes or race bikes are less concerned with using on-street facilities.	Duc-Nghiem, Hoang-Tung, Kojima, Kubota (2018)
	Commuters seem almost insensitive to the distinction between off and on-street facilities, because they are more concerned with efficiency.	Arellana, Saltarín, Larrañaga, González, Henao (2020), Bernardi, Geurs and Puello (2018)

Table 2.3 - Preferences for Lane Width		**References**
General	There are indications in the literature that cyclists do consider lane width in their route decisions.	Providelo and Da Penha Sanches (2011),
	Preferences for lane width are commonly studied in conjunction with other factors that relate to safety. Moreover, it might be difficult to measure distinct preferences for lane width due to limited variations and strong covariance with other factors.	Kang and Fricker (2013), Kang and Fricker (2018)
Off-street versus on-street	It appears that lane width is considered less important when it concerns an off-road facility in comparison to an on-street one.	Rossetti, Saud and Hurtubia (2019)

2.3.3 Sharing Facilities with Pedestrians

Several studies that were reviewed pay attention to the inclination of some cyclists to use facilities that are shared with pedestrians. Generally it appears that cyclists avoid these facilities (Hunt and Abraham, 2007; Skov-Petersen, Barkow, Lundhede and Jacobsen, 2018; Kang and Fricker, 2013). Hunt and Abraham (2007) argue that cyclists might be afraid to bump into pedestrians or be annoyed by having to adapt their speed. Kang and Fricker (2013) reason that the bike facilities around Purdue University, where they conducted a stated preference experiment, are of such good quality that walkways are not appealing to cyclists. However, in situations where bike facilities are lacking it could be possible that cyclists trade the street for a safe sidewalk. For example, Rossetti, Saud, Hurtubia (2019) report that inexperienced cyclists are particularly likely to use sidewalks if no dedicated bike facility is provided. Further, Kang and Fricker (2013) argue that wider sidewalks may encourage cyclists to move from the street. This behavior of cyclists to take over sidewalks has been studied more often (e.g. Kang, Fricker, 2016), but is out of the scope of this review. Overall, it seems that cyclists generally dislike cycling between pedestrians, although inexperienced cyclists may sometimes resort to the sidewalk in search of safety.

Table 2.4 - Preferences for Shared Facilities with Pedestrians		**References**
General	Cyclists generally appear to avoid facilities that are shared with pedestrians, possibly due to the difference in speed which causes dangerous and annoying situations.	Hunt and Abraham (2007), Skov-Petersen, Barkow, Lundhede and Jacobsen (2018), Kang and Fricker (2013)
Cycling experience	Inexperienced cyclists appear more inclined to use sidewalks if no dedicated bike facility is available to them, likely due to safety concerns.	Rossetti, Saud, Hurtubia (2019), Kang, Fricker (2016)

2.4 Street Layout and Pavement

The impact of infrastructure on route choices of cyclists extends beyond the characteristics of cycling facilities alone. That is, the literature indicates that the general layout of the street network also shapes the perceptions of cyclists. This includes parking, signing, street lights, bus stops, intersections and more. The overview provided below illustrates the broad relation between bike route choice and infrastructural aspects. Overall, the role of intersections and turns is the most extensively covered topic in the literature. One stream identifies intersections and turns as sources of irritation, delay and danger. These studies conclude that cyclists generally aim for a continuous route without interruptions. The other stream argues for connectedness and directedness. These studies argue that cyclists tend to stick to lower class roads, which form a dense network and offer a relatively short path to a destination. As a consequence of the road density, cyclists who stick to these roads will inevitably face a high number of intersections. Likewise, the findings for other infrastructural aspects can generally be explained by either a need for safety or for ease and speed. For example, some studies report a preference for traffic lights, because they offer a safe right of passage, whereas others argue against traffic lights, because they are a source of delay. To add, cyclists may avoid routes with on-street parking and bus stops, because these may cause interruptions. Further, some cyclists seem to avoid one-way streets, because cycling against traffic may cause dangerous situations, whereas others see them as a quick shortcut. All in all, speed versus safety seems to be an important trade-off in infrastructural preferences.

2.4.1 Intersections

The literature generally indicates that cyclists are tolerant of intersections. For example, Lu, Scott and Dalumpines (2018) conclude that chosen cycle routes, on average, contain more intersections compared to their shortest alternative. This might be explained by the fact that a highly connected infrastructure implies a larger number of intersections. Interestingly, the density of intersections along chosen routes is lower. Context may also shape people's perceptions of intersections. To illustrate, intersections along separate bike facilities (Krizek, El-Geneidy and Thompson, 2007) and those which feature good visibility (Providelo, da Penha Sanches, 2011) are reported to be experienced more positively. To add, Prato, Halldórsdóttir and Nielsen (2018) argue that experienced cyclists like roundabouts, because in Copenhagen cyclists have a right of way on them. In contrast, intersections with signs and traffic lights seem to discourage cyclists. This might be the case because those safety measures are often present at busy crossings, where motorized traffic poses a threat to cyclists (Kang and Fricker, 2013). The literature does not provide indications of preference heterogeneity regarding intersections, nor is there clear evidence against it. This topic therefore warrants further investigation.

Table 2.5 - Preferences for Intersections		**References**
General	In general, cyclists appear to be relatively tolerant of intersections.	Lu, Scott and Dalumpines (2018)
Bike Facilities	Intersections along separate bike facilities are experienced less negatively.	Krizek, El-Geneidy and Thompson (2007)
Visibility	Intersections with good visibility appear to bother cyclists less.	Providelo, da Penha Sanches (2011)
Signage and Signals	Intersections with signage and signals appear to be less appealing to cyclists, but it might be that these crossings are simply busier.	Kang and Fricker (2013)

2.4.2 Turns

Turns is a widely studied topic in route preference research. In general, cyclists seem to prefer simple routes with few turns (Providelo and da Penha Sanches, 2011; Hood, Sall and Charlton, 2011; Zimmermann, Mai and Frejinger, 2017; Ghanayim and Bekhor, 2018). Broach, Dill and Gliebe (2012) argue that turns delay cyclists and make it difficult for them to remember their route. The aversion appears to be particularly strong for left turns (Broach et al., 2012; Prato, Halldórsdóttir and Nielsen, 2018; Skov-Petersen, Barkow, Lundhede and Jacobsen, 2018). This can be explained by the fact that left turns require cyclists to cross oncoming traffic, introducing the risk of dangerous collisions. It is therefore not surprising that left turns are reported to be particularly discouraging in heavy traffic, when no bike facilities are available and when safety measures such as signs and traffic lights are lacking (Zimmermann et al., 2017; Broach et al., 2012). However cyclists appear more tolerant of turns at the start and end of their route (Skov-Petersen et al., 2018), indicating that wayfinding strategies change throughout a trip. Further, Sobhani, Aliabadi and Farooq (2019) conclude that cyclists in Toronto (Canada) choose routes with a relatively high number of turns. They attribute this finding to the city's dense network and the high number of one-way streets, which tend to be avoided by cyclists when going against traffic. Interestingly, Prato et al. (2018) mention that a specific group of cyclists has a particularly strong aversion towards both left and right turns. Unfortunately they do not specify a profile for this group. Thus, although there is an indication for preference heterogeneity regarding turns, the origin of this remains unclear.

Table 2.6 - Preferences for Turns		**References**
General	Cyclists seem to prefer simple routes with few turns. This finding is robust across a variety of studies.	Providelo and da Penha Sanches (2011), Hood, Sall and Charlton (2011), Zimmermann, Mai and Frejinger (2017), Ghanayim and Bekhor (2018)
	Left turns appear to be particularly bothersome to cyclists, likely because they require them to cross oncoming traffic.	Broach, Dill and Gliebe (2012), Prato, Halldórsdóttir and Nielsen (2018), Skov-Petersen, Barkow, Lundhede and Jacobsen (2018)
Heavy Traffic	Heavy traffic increases the danger of left turns, thus making them even less appealing.	Zimmermann, Mai and Frejinger (2017)
Bike Facilities	Dedicated bike facilities increase safety and therefore seem to make cyclists more tolerant of turns.	Broach, Dill and Gliebe (2012)
Signage and Signals	Cyclists seem more tolerant of turns on streets with signage and traffic signals.	Zimmermann, Mai and Frejinger (2017)

2.4.3 Traffic lights

Findings on preferences of traffic lights along cycling routes are mixed. As touched upon in the previous paragraph, traffic lights may offer safe passage and can therefore make intersections more appealing (Park and Akar, 2019). Particularly in situations with high traffic volumes, the safety benefits of traffic lights seem to outweigh the delay they cause (Broach, Dill and Gliebe, 2012). However, Skov-Petersen, Barkow, Lundhede and Jacobsen (2018) point out that local governments may simply grant popular cycling routes more traffic lights, thus complicating the entanglement of real preferences in revealed choice data. Further, another study in Zurich reports that cyclists actually avoid routes with a high number of traffic lights (Menghini, Carrasco, Schüssler and Axhausen, 2010). It is important to note that the data used in this study did not allow for modelling an interaction between traffic light preferences and traffic volumes, as Broach et al. (2012) did. Moreover, they did not test whether this finding might be clouded with a general aversion towards intersections. Thus, a study that distinguishes between the preference for traffic signals and intersections and which explores the interplay of those preferences with traffic volumes could provide clarification.

Table 2.7 - Preferences for Traffic Lights		**References**
General	Findings on preferences for traffic lights are mixed.	
High traffic volumes	Cyclists appear attracted to the safety and efficiency of traffic lights when traffic volumes are high.	Park and Akar, 2019; Broach, Dill and Gliebe, 2012

2.4.4 Car Parking

Cyclists seem to disfavor car parking across their route, potentially because it hinders sight and free movement and might cause dangerous situations (Hardinghaus and Papantoniou, 2020; Sener, Eluru and Bhat, 2009; Winters and Teschke, 2010). This includes on-street, angled and parallel parking (Sener et al., 2009). According to Sener et al. (2009), Male cyclists appear to be more bothered by parking compared to females, possibly because the former find it more important to keep a constant speed. The same seems to hold for long versus short commutes in their sample. Interestingly, they did not find cycling experience to affect preferences for car parking. Overall, the preference against parked cars seems to hold across various contexts and types of cyclists.

Table 2.8 - Preferences for Car Parking		References
General	Cyclists seem to disfavor car parking across their route, most likely because it can cause dangerous situations.	Hardinghaus and Papantoniou, 2020; Sener, Eluru and Bhat, 2009; Winters and Teschke, 2010
Gender & commute length	Males & long commuters appear to be more bothered by parking, possibly because they find it important to keep a constant speed.	Sener, Eluru and Bhat, 2009

2.4.5 Pavement Quality and Debris

In general, the literature provides indications for a strong preference for clean, smooth and high quality pavement among cyclists. To illustrate, Winters and Teschke (2010) report a general preference for paved roads over unpaved ones. The maintenance of paved roads also seems important. To illustrate, Parkin, Wardman and Page (2008) found that commuters are discouraged by poorly maintained pavement. They argue that poor maintenance not only decreases comfort, but also increases physical effort. Interestingly, Hardinghaus and Papantoniou (2020) report that frequent versus infrequent cyclists are more bothered by bad pavement, likely because they are exposed to it more often. Kang and Fricker (2013), in turn, discovered that cyclists are inclined to use the sidewalk in those situations, increasing the risk of collisions with pedestrians. This illustrates how badly cyclists want to avoid poorly maintained pavements. Further, cyclists seem to find it important that the road is free of glass and debris and does not become slippery when wet or icy (Winters, Davidson, Kao and Teschke K, 2011). Interestingly, Providelo, da Penha and Sanches (2011) report that cyclists in medium-sized Brazilian cities do not find pavement quality important. Their results indicate that in the context of these cities, safety related issues such as lane width, visibility, intersections and speed limits are considered more important than the comfort of high quality pavement. It thus appears that pavement quality is important to cyclists, under the condition that safety is ensured.

Table 2.9 - Preferences for Pavement & Debris		References
General	Cyclists prefer clean, smooth and high quality pavement.	Winters and Teschke, 2010; Winters, Davidson, Kao and Teschke K, 2011
	However, safety seems more important.	Providelo, da Penha and Sanches, 2011
Commuters	General findings hold for commuters.	Parkin, Wardman and Page, 2008
Cycling Frequency	Frequent cyclists are more bothered by bad pavement, likely because they are exposed to it more often.	Hardinghaus and Papantoniou, 2020

2.4.6 One-Way Streets

Although not extensively covered by the literature, there are indications that cyclists avoid cycling in the wrong direction down a one-way street (Hood, Sall and Charlton, 2011). Specifically, Prato, Halldórsdóttir and Nielsen (2018) conclude that cyclists perceive distances over twice as long when they cycle against the stream of motorized traffic. Sobhani, Aliabadi and Farooq (2019) argue that cyclists avoid one-way streets because they restrict their movement. Interestingly, Prato et al. (2018) do identify a specific group among cyclists which appears to prefer shortcuts that go against traffic. Thus, preference heterogeneity may exist.

2.4.7 Bridges

The literature indicates that bridges may also influence the behavior of cyclists. As one might expect, cyclists generally seem to avoid bridges (Zimmermann, Mai and Frejinger, 2017). However, bridges with separate bike facilities appear to be more appealing (Broach, Dill and Gliebe, 2012; Zimmermann et al., 2017). In particular, Prato, Halldórsdóttir and Nielsen (2018) report that dedicated bridges for cyclists appeal strongly to Copenhagen's cyclists. However, it might be that these facilities are simply granted to popular cycling routes or provide efficient routes across town. It would be interesting to evaluate if other highly bike friendly cities also show the appeal of bridges among cyclists.

Table 2.10 - Preferences for Bridges		**References**
General	Cyclists generally avoid bridges.	Zimmermann, Mai and Frejinger, 2017
Bike Facilities	Bike facilities make bridges more appealing.	Broach, Dill and Gliebe, 2012; Zimmermann, Mai and Frejinger, 2017; Prato, Halldórsdóttir and Nielsen, 2018

2.5 Nature and Topography

The literature indicates that preferences for surroundings come after those related to safety, comfort and efficiency (Bernardi, Geurs and Puello, 2018). Nevertheless, surroundings still do influence route choice. As discussed below, leisure cyclists in particular seem to have unique preferences when it comes to nature and topography.

2.5.1 Scenery and Green

Unique and green surroundings are known to attract cyclists (Hardinghaus, Papantoniou, 2020). For example, participants of a study in Vancouver (Canada) rated "beautiful scenery" among the top-3 motivators for cycling (Winters, Davidson, Kaoand Teschke, 2011). Ghanayim and Bekhor (2018) conclude that not only green, but also seashores can be attractive surroundings for cyclists. However, a preference for greenery is more commonly studied. This preference appears to be particularly strong among leisure cyclists (Chen, Shen and Childress, 2018) and also seems stronger among females compared to men (Vede, Jacobsen and Skov-Petersen, 2017). Park and Akar (2019) argue that greenery may also "serve as a buffer from other activities" (p.199). Following similar reasoning as Garrard, Rose and Lo (2008) provide for the strong preference of female cyclists for seperate bike facilities, the higher degree of risk aversion among females may also explain why they prefer a green buffer. Further, in areas with warm summers, such as Brazil (Providelo and da Penha Sanches, 2011) and parts of China (Liu, Yang, Timmermans, de Vries, 2020), trees may also be valued for the shade they provide. Interestingly, Park and Akar (2019) conclude that cyclists are only willing to detour for pleasurable surroundings when temperatures are above 5 degrees celsius. In sum, climate and weather may influence preferences for green surroundings. It is also important to note that preferences for greenery generally come after those regarding safety and comfort, especially for cyclists who do not detour substantially (Bernardi, Geurs and Puello, 2018). This conclusion is supported by Skov-Petersen, Barkow, Lundhede, and Jacobsen (2018), who report a disutility for green, which they ascribe to the fact that green areas are generally less safe and lack street lights.

2.5.2 Hilliness

Preferences regarding hilliness appear to differ substantially across different groups of cyclists and can also be related to trip purpose. In general cyclists are demotivated by steep hills (Chen, Shen and Childress, 2018; Parkin, Wardman, Page, 2008; Winters, Davidson, Kaoand Teschke, 2011; Sarjala, 2019). Specifically, Prato, Halldórsdóttir and Nielsen (2018) report the disutility of slopes increases as the gradient does. According to Zimmermann, Mai and Frejinger (2017), upslopes discourage cyclists starting at an angle of 4%. The preference for flat terrain appears particularly strong among commuters (Hood, Sall and Charlton, 2011; Anowar, Eluru and Hatzopoulou, 2017). Sobhani, Aliabadi and Farooq (2019) argue that commuters might not want to arrive at their meetings sweaty and out of breath. To add, regardless of trip purpose, female cyclists seem to be more bothered by steep hills compared to men (Sener, Eluru and Bhat, 2009; Hood, Sall, Charlton; Hood, Sall and Charlton, 2011). In this regard, Anowar et al. (2017) argue that slopes are hard and uncomfortable to climb, but can also be scary and dangerous to descend due to the high speed. They refer to other studies which highlight that women are less inclined to conduct physical exercise and are also more risk averse. The latter may explain why the effect is stronger for women in comparison to men. Overall, commuters and females seem to be particularly discouraged by slopes.

Interestingly, some cyclists appear to be more tolerant of slopes or even prefer slight hilliness. For example, tourists seem to be more tolerant of minor slopes, particularly if they are cycling on a segregated bike facility (Deenihan and Caulfield, 2015). To add, Lu, Scott and Dalumpines (2018) argue that bike-sharers might be more tolerant of minor slopes as the consequence of tradeoffs with other route characteristics. Further, Sener, Eluru and Bhat (2009) report a preference for some hilliness, particularly among leisure cyclists. They argue that this group prefers variation in the landscape and physical challenge. These needs might be tempered among commuters, due to their need for efficient transport. Further, Sener et al. (2009) also report that male cyclists prefer a hilly landscape, both during commute and leisure trips. In sum, not all cyclists are strongly discouraged by slopes and some even prefer them over flat terrain.

Some studies remain inconclusive regarding preferences for gradients. For example, Ghanayim and Bekhor (2018) did not find a significant effect for slope on route choice. They explain that in Tel Aviv the variation in gradients is limited, which makes it difficult to measure this preference based on GPS data. The same argument is used by Park and Akar (2019), who conducted a revealed preference study in Columbus (Ohio). This may also hold for Prato, Halldórsdóttir and Nielsen (2018), who did not observe the commonly reported differences across males and females in Copenhagen. To add, Menghini, Carrasco, Schüssler and Axhausen (2010) argue that the effects that they found could have been larger if the hills in the study area could be more easily avoided. In sum, the topography of a study area may influence the observed preferences for hilliness in revealed preference studies.

2.6 Traffic Volumes and Speed Limits

The effects of traffic volumes and speed limits on route choice are rarely studied together. That is, since the two can be expected to correlate strongly, it is difficult to separate these effects, particularly in revealed choice studies. It is therefore hard to tell what the individual effects of these route characteristics are. In that regard, Providelo and da Penha Sanches (2011) conducted a relatively unique research in which they used successive interval analysis with focus groups and attitude surveys. They conclude that speed limits are considered far more important compared to traffic volumes. More research is needed to confirm this finding.

Table 2.11 - Preferences for Nature and Topography		References
General	Cyclists are strongly discouraged by steep hills. Strong differences in preferences based on personal characteristics and trip purpose.	Chen, Shen and Childress, 2018; Parkin, Wardman, Page, 2008; Prato, Halldórsdóttir and Nielsen, 2018; Winters, Davidson, Kaoand Teschke, 2011; Sarjala, 2019; Zimmermann, Mai and Frejinger, 2017
	Cyclists are attracted to unique and green surroundings.	Hardinghaus, Papantoniou, 2020; Winters, Davidson, Kaoand Teschke, 2011, Ghanayim and Bekhor, 2018
	However, preferences for safety are more important.	Bernardi, Geurs and Puello, 2018; Skov-Petersen, Barkow, Lundhede, and Jacobsen, 2018
Commuters	Commuters seem to particularly avoid slopes.	Anowar, Eluru and Hatzopoulou, 2017; Hood, Sall and Charlton, 2011; Sobhani, Aliabadi and Farooq, 2019
Gender	Females have a stronger dislike for slopes.	Anowar, Eluru and Hatzopoulou, 2017, Sener, Eluru and Bhat, 2009; Hood, Sall, Charlton; Hood, Sall and Charlton, 2011
	Females appear to have a stronger preference for green surroundings, possibly because green can serve as a buffer.	Vede, Jacobsen and Skov-Petersen, 2017; Park and Akar, 2019
Bike-sharers / Leisure cyclists	Appear to be more tolerant of slopes, possibly because they like the challenge and changing landscape.	Lu, Scott and Dalumpines, 2018; Sener, Eluru and Bhat, 2009
Bike Facilities	Tourists are more tolerant of slopes if a seperate cyclist facility is available.	Deenihan and Caulfield, 2015
Topographical variation	Several studies find no significant or weak result. Possibly because the variation in the study area is too limited.	Ghanayim and Bekhor, 2018; Park and Akar, 2019; Menghini, Carrasco, Schüssler and Axhausen, 2010
Climate	Cyclists from warm climates may like trees because they provide shade.	Providelo and da Penha Sanches, 2011; Liu, Yang, Timmermans, de Vries, 2020

2.6.1 Traffic Volumes

There are strong indications in the literature that cyclists generally avoid streets with high traffic volumes (Ghanayim and Bekhor, 2018; Sener, Eluru and Bhat, 2009; Winters, Davidson, Kao and Teschke, 2011; Zimmermann, Mai and Frejinger, 2017). Anowar, Eluru and Hatzopoulou (2017) argue that this is related to safety concerns, because high traffic volumes implies a higher risk of collisions. Cyclists would therefore prefer streets with low traffic volumes, such as residential ones. To add, Parkin, Wardman and Page (2008) reason that areas with high traffic volumes have a strong focus on motorized traffic, hence, may not have been designed with cyclists in mind.

Melson, Duthie and Boyles (2014) show that this preference for low traffic volumes also holds for bridges. Interestingly, separate bike facilities seem to substantially reduce the negative effect of traffic volumes (Kang and Fricker, 2013; Broach, Dill and Gliebe, 2012; Park and Akar, 2019).

The findings of several studies also indicate that one should be careful in the definition of high and low traffic volumes. To illustrate, Parkin, Wardman and Page (2008) found no significant effect for the proportion of a route which is traffic free. To add, Zimmermann, Mai and Frejinger (2017) found no significant difference between medium and heavy traffic. Thus, there seems to be a certain threshold where traffic volumes become disturbing to cyclists. Therefore, a below versus above medium traffic volume measure might work better than using three categories or a continuous variable. This may partially explain why Hood, Sall and Charlton(2011), to their own surprise, did not find a significant result for traffic volumes.

Further, as Sener, Eluru and Bhat (2009) point out, preference heterogeneity regarding this preference is high. That is, some cyclists appear to be less bothered by traffic volumes. For example, Anowar, Eluru and Hatzopoulou (2017) report that females seem to be more drawn to low-traffic residential streets, due to their generally stronger risk aversion. In contrast, Sener, Eluru and Bhat (2009) conclude that men are more bothered by traffic, because they would find it more important to keep a constant speed. They also report that commuters are strongly discouraged by high traffic volumes for the same reason. Broach, Dill and Gliebe (2012) attribute this strong preference among cyclists to the simple fact that they are more exposed to peak hour traffic. Following similar reasoning, Arellana, Saltarín, Larrañaga, González and Henao (2020) argue that infrequent cyclists are less bothered by traffic volumes, since their exposure is relatively low. In short, personal characteristics and trip purpose may influence the strengths of preferences for traffic volumes.

Table 2.12 - Preferences for Traffic Volumes		**References**
General	Cyclists generally prefer low traffic volumes.	Anowar, Eluru and Hatzopoulou, 2017; Ghanayim and Bekhor 2018; Melson, Duthie and Boyles, 2014; Sener, Eluru and Bhat, 2009; Parkin, Wardman and Page, 2008; Winters, Davidson, Kao and Teschke, 2011; Zimmermann, Mai and Frejinger, 2017
Bike Facilities	Bike facilities can reduce the negative effect of high traffic volumes.	Kang and Fricker, 2013; Broach, Dill and Gliebe, 2012; Park and Akar, 2019
Gender	Some argue that female cyclists are less tolerant of high traffic volumes, possibly due to their stronger risk aversion.	Anowar, Eluru and Hatzopoulou, 2017
	Others conclude that men are more bothered by traffic, because they want to keep a constant speed.	Sener, Eluru and Bhat, 2009
Trip Purpose	Commuters appear to be more sensitive to high traffic volumes, possibly because they are exposed to them more often during peak hours.	Sener, Eluru and Bhat, 2009; Broach, Dill and Gliebe (2012)
Cycling Frequency	Infrequent cyclists are less bothered by traffic volumes, since their exposure is relatively low.	Arellana, Saltarín, Larrañaga, González and Henao, 2020

2.6.2 Speed Limits

The literature indicates that all cyclists prefer low speed limits over higher ones (Fitch and Handy, 2020; Chen, Shen and Childress, 2018; Providelo and da Penha Sanches, 2011). In particular, Winters, Davidson, Kao and Teschke (2011) observe that cyclists become substantially discouraged by speed limits above 50 km/hr. Interestingly, some cyclists seem more concerned with speed limits than others. For example, people cycling with children have a stronger preference for low speed limits (Hardinghaus and Papantoniou, 2020). Females also appear to be more careful and try to avoid high speed limits more often than their male counterparts (Fitch and Handy, 2020). The same seems to hold for inexperienced cyclists and short commuters (Handy et al., 2020; Sener, Eluru and Bhat, 2009). These groups are most likely extra concerned with safety or are less comfortable when cycling between high speed traffic. Indeed, Chen, Shen and Childress (2018) report that cyclists who find safety very important avoid roads with high speed limits. In contrast, leisure cyclists appear to be less bothered by speed limits, possibly due to their experience and agility. Intriguingly, Hardinghaus and Papantoniou (2020) report a very weak preference for low speed limits among German and in particular Greek cyclists. The authors argue that the effect of lower traffic speeds is limited for Greek cyclists because Greek drivers are less inclined to stick to traffic rules. Why the preference among German cyclists is weak compared to other studies remains unclear.

Table 2.13 - Preferences for Speed Limits		References
General	Cyclists generally prefer low speed limits.	Fitch and Handy, 2020; Chen, Shen and Childress, 2018; Providelo and da Penha Sanches, 2011; Winters, Davidson, Kao and Teschke, 2011
With Children	People cycling with children have a stronger preference for low speed limits.	Hardinghaus and Papantoniou, 2020
Gender	Females have a stronger preference for low speed limits.	Fitch and Handy, 2020
Cycling Frequency	Inexperienced cyclists have a stronger preference for low speed limits.	Handy et al., 2020; Sener, Eluru and Bhat, 2009
Safety	Cyclists who find safety important have a stronger preference for low speed limits.	Chen, Shen and Childress (2018)

2.7 Safety

The literature clearly shows that a safe environment is essential to get people on their bikes. To illustrate, Hopkinson and Wardman (1996) and Arellana, Saltarín, Larrañaga, González and Henao (2020) report that safety is among the top motivators to cycle. To add, Manaugh, Boisjoly and El-Geneidy (2013) observe that unsafe cycling infrastructure demotivates potential cyclists. Further, Buehler and Pucher (2012) report that cycling commute rates are higher in safe areas compared to unsafe ones. Thus, it appears that safety does influence cycling behavior.

It is important to note that safety does not only refer to minimizing the risk of collisions, it also encompasses a broader feeling of security. For example, darkness (Winters, Davidson, Kao and Teschke, 2006; Chen, Shen and Childress, 2018; Liu, Yang, Timmermans and De Vries, 2020; Majumdar and Mitra, 2017) and even scolding and crowded cycleways may unease cyclists (Vedel, Jacobsen and Skov-Petersen, 2017). Safety measures such as security cameras and traffic lights (Arellana, Saltarín, Larrañaga, González and Henao, 2020), reflective centerlines (Winters, Davidson, Kao and Teschke, 2006) and illuminated corridors (Majumdar and Mitra, 2017) can help cyclists to feel more safe.

Safety concerns may influence route choice among cyclists, although their relevance may differ across contexts. That is, Majumdar and Mitra (2017) report that Indian cyclists are strongly influenced by safety levels when it comes to their route choices. Likewise, Arellana, Saltarín, Larrañaga, González and Henao (2020) conclude that Colombian cyclists put a high emphasis on safety related issues, such as the presence of traffic control devices, (bike) traffic flows and speed, security cameras and street lighting. Further, Parkin, Wardman and Page (2008) argue that cycling is less common in low-income areas, possibly due to high crime rates. However, there are no clear indications that cyclists in Western countries have safety, in a broad sense, on top of their mind when they pick a route. As earlier sections indicate, these cyclists appear to be most concerned with the risk of collision. In contrast, they do not seem to consider other safety issues such as crime rates (Hood, Sall and Charlton, 2011). This could be related to the relatively safe situations in Western countries. Indeed, Kang and Fricker (2018) report that as the risk increases, safety becomes almost as important as distance in the route choices of commuting cyclists. Overall, it seems that Western-world cyclists are mostly concerned with traffic safety, whereas cyclists from other (more dangerous) areas may also consider other safety issues when selecting their route.

2.7.1 Accidents

Earlier sections already discuss literature which highlights the role of traffic safety concerns when selecting a route. For example, cyclists seem to prefer off-street facilities over on-street ones, because the former reduce the perceived risk of collisions (see Table 2.3). The same holds for the presence of traffic signals and good visibility at intersections (see Table 2.5). In sum, risk of collision is a returning element in the explanation of other preferences. As confirmed by Winters, Davidson, Kao and Teschke (2006), people are strongly discouraged to cycle in areas where they face the risk of injury from accidents with cars. They are therefore attracted to roads with safety measures such as traffic lights and off-street facilities. However, there are no indications in the reviewed literature that cyclists specifically avoid streets for the mere reason that they have a high number of accidents. It could be interesting to learn whether the possible avoidance of these streets stands separate from preferences for certain safety measures.

2.7.2 Street Lights and Visibility

The literature underpins the importance of street lights and visibility to perceived safety among cyclists. In general, people seem to prefer to cycle during daylight hours (Winters, Davidson, Kao and Teschke, 2006). As night falls, they value well lit roads (Winters et al., 2006; Chen, Shen and Childress, 2018; Liu, Yang, Timmermans and De Vries, 2020; Arellana, Saltarín, Larrañaga, González and Henao, 2020). Specifically, Liu et al. (2020) conclude that cyclists prefer street lights to be placed every fifteen to thirty meters. Female cyclists (Liu et al., 2020) seem particularly sensitive to badly lit roads. Interestingly, Chen, Shen and Childress (2018) report that cyclists who aim to minimize their trip length also have a relatively strong preference for a high street light density. Furthermore, cyclists also seem to like reflective centerlines, because they improve visibility (Winters et al., 2006). Overall, street lights and good visibility appear to be valued strongly by most cyclists.

Table 2.14 - Preferences for Safety		References
General	A safe environment is essential to get people on their bikes.	Hopkinson and Wardman, 1996; Arellana, Saltarín, Larrañaga, González and Henao, 2020; Manaugh, Boisjoly and El-Geneidy, 2013;Further, Buehler and Pucher, 2012
Accidents	There are no indications that cyclists avoid roads with a high number of accidents.	
Visibility	Cyclists prefer cycling during daylight hours.	Winters, Davidson, Kao and Teschke, 2006
	As night falls, they value well lit roads.	Winters et al., 2006; Chen, Shen and Childress, 2018; Liu, Yang, Timmermans and De Vries, 2020; Arellana, Saltarín, Larrañaga, González and Henao, 2020
	Cyclists like reflective centerlines.	Winters, Davidson, Kao and Teschke, 2006
Gender	Female cyclists seem particularly sensitive to badly lit roads	Liu, Yang, Timmermans and De Vries
Trip Length	Those who aim to minimize their trip length seem to find visibility particularly important.	Chen, Shen and Childress, 2018

2.8 Amenities

Few studies consider amenities as a potential factor in route choice behavior. It is known that cyclists are concerned with secure bike parking at the destination (Winters, Davidson, Kao and Teschke, 2006; Hunt and Abraham, 2007), but this does not influence how they get there. Findings by Chen and Chen (2013) indicate that amenities along a route might be particularly important to leisure cyclists. To illustrate, they report that recreational cyclists are generally attracted by routes that pass along attractions and offer facilities such as toilets, basic bike maintenance equipment and tourist information centers. Moreover, those who cycle a long distance appear to have a relatively strong preference for restaurants. Last, frequent leisure cyclists have a particularly strong preference for variation in amenities along their routes. Overall, there are indications that leisure cyclists consider amenities when selecting their route, but more research is needed to confirm this.

Table 2.15 - Preferences for Amenities		References
Trip Purpose	Amenities are particularly important to leisure cyclists. Examples include: toilets, basic bike maintenance equipment and tourist information centers.	Chen and Chen, 2013
Trip Length	On longer trips, cyclists prefer to pass restaurants.	Chen and Chen, 2013
Cycling Frequency	Frequent leisure cyclists like varying amenities.	Chen and Chen, 2013

2.9 Impact of Weather Conditions

Weather conditions are known to influence cycling behavior. For example, cyclists are discouraged by cold and snow, particularly in countries with harsh winters, such as Canada (Sobhani, Aliabadi and Farooq, 2019). In these countries, slippery and snowy pavements are likely to be an important deterrent of cycling (Winters, Davidson, Kao and Teschke, 2006). Further, rainfall and extreme temperatures can discourage commuters to travel by bike (Parkin, Wardman and Page, 2008). Weather conditions such as annual precipitation and annual hot or cold days are not reported to influence bike commute habits, according to Buehler and Pucher (2012). It might be that these annual based measurements are too broad to reveal preferences. Overall, it appears that weather influences when people decide to cycle.

Interestingly, the role of the interactions between weather conditions and route characteristics in route choices of cyclists are rarely reported among the reviewed articles. Deenihan and Caulfield (2015) report the very specific finding that tourists are tolerant of bad weather conditions if a segregated bike facility is available. To add, Prato, Halldórsdóttir and Nielsen (2018) conclude that weather conditions may impact the perception of a bike route. For example, cyclists appeared willing to detour for scenic areas only at temperatures above five degrees celsius. Hood, Sall and Charlton (2011) could not find significant interactions for hourly rainfall or daylight hours. However, their study was conducted in San Francisco, where weather conditions are generally mild and variation throughout the year is limited. Overall, much remains to be discovered regarding the impact of weather conditions on route choice.

2.10 General Willingness to Detour

There is consensus in the literature that cyclists have a general preference for short routes (Hood, Sall and Charlton, 2011; Broach, Dill and Gliebe, 2012; Manaugh, Boisjoly, El-Geneidy, 2013; Zimmermann, Mai and Frejinger, 2017; Ghanayim and Bekhor, 2018; Menghini, Carrasco, Schüssler and Axhausen, 2010). Specifically, most studies report an average degree of detour of about 11%, as shown in Table 2.16.

Table 2.16 - Reported Degree of Detour * = smartphone users only

Reference	Location	Data Elicitation	Study Population	Degree of Detour
Bernardi, Geurs and Puello (2018)	The Netherlands	smartphone GPS data (MoveSmarter)	current cyclists *	15%
Park and Akar (2019)	Columbus (US)	GPS app data (CycleTracks)	current cyclists *	13.%
Fitch and Handy (2020)	Davis (US)	online survey	students & faculty members, most likely commuters	5%
	San Francisco (US)	GPS app data (CycleTracks)	current cyclists *	12%
Lu, Scott and Dalumpines (2018)	Hamilton (Canada)	GPS-equipped shared bikes (SoBi)	bike-sharers	10%
Broach, Dill and Gliebe (2012)	Portland (US)	GPS trackers	commuters	11%
			non-commuters	12%
Average				**11%**

However, it appears that willingness to detour varies across contexts and groups of cyclists. For example, exposure to motorized traffic seems to decrease willingness to detour (Hunt and Abraham, 2007). Further, commuters appear to detour less (Sener, Eluru and Bhat, 2009), probably due to time constraints (Broach, Dill and Gliebe, 2012). This also seems to hold for utilitarian cyclists who, for example, go shopping (Chen, Shen and Childress, 2018). Further, women (Manaugh, Boisjoly, El-Geneidy, 2013; Anowar, Eluru and Hatzopoulou, 2017) and recreational cyclists (Melson, Duthie and Boyles, 2014) appear to have a relatively high willingness to detour, probably because they put more emphasis on comfort (Melson, Duthie and Boyles, 2014). To add, young (25-34 years old) commuters seem to be particularly sensitive to travel duration, possibly due to their fast lifestyles, as argued by Anowar, Eluru and Hatzopoulou (2017). In contrast, senior commuters (55+ years old) are less sensitive to travel time, possibly because they are less constrained than their younger counterparts, according to Anowar et al. (2017).

Interestingly, some studies use detour as a dependent variable, which allows them to translate preferences into willingness to detour for a particular route characteristic. Examples include Prato, Halldórsdóttir and Nielsen (2018), Vedel, Jacobsen and Skov-Petersen (2017) and Zimmermann, Mai and Frejinger (2017).

2.11 Conclusion

This chapter provides an overview of the findings of over forty articles on bike route choice from twenty one different countries. Considering these findings, it seems that preferences for many of the often studied route characteristics can be related to safety concerns. These include intersections, (left) turns, traffic lights, speed limits and on and off-street bicycle facilities. Several studies indicate that these preferences are relatively strong. To illustrate, pleasurable surroundings (Bernardi, Geurs and Puello, 2018; Skov-Petersen, Barkow, Lundhede, and Jacobsen, 2018) and pavement quality (Providelo, da Penha and Sanches, 2011) only seem to be important when safety is ensured. Moreover, there is consensus in the literature that cyclists have a general preference for short routes. Specifically, cyclists are unwilling to detour more than about 11% compared to the shortest route.

Most studies consider preference heterogeneity and relate this to personal characteristics and trip context. In particular, the distinctions between males and females are often evaluated. Overall, it seems that females are more risk averse and therefore have stronger preferences for safety aspects. For example, they are more sensitive to the absence of a separate bike facility and have a stronger preference for low speed limits. Moreover, they seem less willing to undergo physical effort, for example to climb a slope or take a detour. Further, studies often report clear differences between the preferences of commuters and leisure cyclists. The former are more concerned with efficiency and speed, since they are bound by stricter time constraints. They are therefore less concerned with the absence of dedicated facilities and particularly sensitive to steep slopes, which may slow them down. In contrast, leisure cyclists put more value in dedicated facilities and have a stronger preference for green scenery. Further, they are more tolerant of slopes, possibly because they prefer a varied landscape or like the physical challenge. Further, the findings indicate that frequent cyclists are less tolerant to some hindrances such as bad pavement and high traffic volumes, most likely because they are exposed to them on a regular basis. Moreover, several studies report interactions among route characteristics. Most importantly, the safety of separate bike facilities seem to make cyclists more tolerant of negatively experienced aspects such as (left) turns, bridges and slopes. Likewise, signage and traffic signals seem to make turns less attractive. Overall, the relations between route characteristics and route choice behavior turns out to be complex and differ across types of cyclists and context.

Interestingly, there are several topics on which the literature has not reached consensus. In particular, the preferences for traffic lights seem to vary greatly and it is unclear which cyclists prefer or avoid them. Further, Prato, Halldórsdóttir and Nielsen (2018) conclude that some cyclists have a particularly strong aversion towards turns, but they do not specify who these cyclists are. To add, there are indications that bridges might be appealing when they offer efficient connections (Prato, Halldórsdóttir and Nielsen, 2018). However, this finding has not been confirmed by other studies. In sum, some preferences remain not fully understood.

3. Route Choice Models

This chapter provides an overview of the current state of discrete choice modelling in the context of route choices. First the general concept of route choice modelling is outlined. The three paragraphs which follow introduce Multinomial Logit (MNL) modelling, Path Size Logit (PSL) modelling and Latent Class modelling techniques. Each of those paragraphs discusses the structures of the respective models and their pros and cons. Thereafter, stated and revealed choice modelling are contrasted. Finally, the process of choice set generation is elaborated upon.

3.1 Introduction to Route Choice Models

Route choice models serve to quantify the relations between a set of explanatory variables and route choice behavior (Schreckenberg & Selten, 2013). As such, they can be applied to evaluate, for example, which factors influence route choices of cyclists and to what extent they do so. The process of estimating a route choice model can be subdivided into two main steps (Schreckenberg & Selten, 2013). First, sets of alternative routes between origin-destination pairs have to be generated. That is, route choice models generally assume that people consider a finite set of route alternatives, which is referred to as a choice set. Each route in a choice set has distinct characteristics, which are recorded in the explanatory variables. The second step concerns the estimation of the likelihood that a participant choses a given route in each corresponding choice set. This is commonly done by means of a Multinomial Logit (MNL) model or its derivatives.

3.2 Multinomial Logit (MNL) Modelling

The Multinomial Logit (MNL) model is the most basic form of discrete choice modelling. It assumes that a decision maker attaches a certain degree of utility to each option in the choice set at hand. The alternative with the highest utility is expected to be selected, following the logic of utility-maximization (Train, 2009). The utility score can be quantified according to a utility function, which is based on the attributes of an alternative and contextual factors regarding the choice situation. Considering that choice behavior can never be completely understood, a portion of the utility remains unknown to the researcher. The total utility of alternative i for observation n is therefore split into the structural utility (V_{in}) and the random utility (ε_{in}) (Train, 2009).

The structural utility function of a MNL model can be defined as follows:

(1) $$V_{in} = \sum_{q} \beta_q \cdot x_{inq}$$

Where x_{inq} refers to the value for attribute q for alternative i of observation n. The parameters to be estimated (β_q) represent the relative contribution of attribute q to the total utility score, much like a weight. The random utility component ε_{in} captures the difference between the utility observed by the decision maker and the utility determined by the researcher according to the utility function. It is assumed to follow a standard Gumbel distribution.
The probability that observation n choses alternative i from choice set C_n is described by the following probability function:

(2) $$Pr(i|C_n) = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}$$

An important limitation of the MNL model is the independence from irrelevant alternatives (IIA) property (Train, 2009). This IIA property predicates that the odds of choosing one alternative over the other remain the same, indepent of the composition of the choice set. This makes a MNL model inappropriate to estimate choice behavior among similar alternatives (Bernardi et al., 2018), as discussed in the next paragraph.

3.3 Path Size Logit Models

As argued by Bernardi et al. (2018), the basic MNL model is generally unsuitable to apply to revealed route choices. That is, routes between the same origin-destination pair, with a realistically small degree of detour, can be expected to overlap. Moreover, deviations in the urban context might be minimal, particularly in the case of short routes which stay within a particular region. Hence, the attributes of routes within one choice set can be very similar. This can cause issues when estimating a MNL model, because the independence of irrelevant alternatives (IIA) property may be violated (Bernardi et al., 2018).

A Path Size Logit (PSL) model, as proposed first by Ben-Akiva and Bierlaire (1999), aims to overcome this issue by introducing a Path Size factor to account for the overlap between route alternatives. This factor ranges from zero, indicating a complete overlap between routes, to a maximum of one, meaning no overlap occurs within the choiceset. Several additional extensions of this model exist. For example, attempts have been made to account for excessively long routes (Ben-Akiva and Bierlaire, 1999) and to ensure that completely unique routes are not unnecessarily penalized (Ramming, 2002).

An extended Path Size Logit (PSL) model can be defined as follows (adapted from Bernardi, Geurs and Puello, 2018). The probability of choosing route i from the choice set C_n for observation n is defined as:

$$(3) \qquad Pr(i|C_n) = \frac{e^{V_{in} + ln(PS_{in})}}{\sum_{j \in C_n} e^{V_{jn} + ln(PS_{jn})}}$$

where V_{in} is the structural utility for route i, V_{jn} is the structural utility for alternative route j and PS_{in} and PS_{jn} are the Path Size factors for routes i and j respectively. In turn, the Path Size factor PS_{in} can be specified as (Bernardi, Geurs and Puello, 2018):

$$(4) \qquad PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \cdot \frac{1}{\sum_{j \in C_n} \frac{L^*_{C_n}}{L_j}^{\lambda} \cdot \delta_{aj}}$$

where Γ_i is the set of links that make up route i, L_i is the length of route i, L_a is the length of link a, C_n is the set of alternatives relevant to observation n, $L^*_{C_n}$ is the length of the shortest route among these alternatives, L_j is the length of route j, δ_{aj} is a dummy variable which indicates whether link a is part of route j ($\delta_{aj} = 1$) or not ($\delta_{aj} = 0$) and λ is a scaling parameter that reduces the contribution of relatively long routes to the path size factor of shorter routes.

The above specification of the Path Size factor includes two extensions to the original PSL mode proposed by Ben-Akiva and Bierlaire (1999). First, the weight term $\frac{L^*_{C_n}}{L_j}$ reduces the contribution of illogically long routes to the Path Size factor. The contribution of route j is weighted for the ratio of the length of route L_j and the length of the shortest route $L^*_{C_n}$. This extension was introduced by Ben-Akiva and Bierlaire (1999) themselves. Second, the scaling parameter λ was introduced by Ramming (2002) to ensure that completely unique routes are not inappropriately penalized. This specification was coined the Generalized Path Size Logit (GPSL) model.

The literature is indecisive regarding the added value of the Generalized Path Size Logit (GPSL) model proposed by Ramming (2002) over the original PSL model. There are indications that the GPSL model shows improved model fit (Duncan et al., 2020) and it has been successfully applied in earlier studies (e.g. Hoogendoorn-Lanser et al., 2005). Yet, other research suggests this specification can result in inapplicable corrections and implausible probabilities (Frejinger and Bierlaire, 2007). Furthermore, the need for these corrections will strongly depend on the nature of the data. That is, if excessive detours are uncommon, the estimation of additional parameters to account for them may be unnecessary (Broach, Dill and Gliebe, 2012). To add, the introduction of an additional scaling parameter may make the model overly complex and calls for larger datasets. Furthermore, the presence of illogical routes can, at least partially, be dealt with during the data cleaning process.

3.4 Latent Class Logit Modelling

Multinomial Logit (MNL) models can reveal the preferences of a sample. However, they do not take into account preference heterogeneity among participants. The results of these models can therefore be biased or misleading (Wen and Lai, 2010). A possible solution is to split the sample based on one or multiple characteristics of the subjects. However, the personal data may not suffice to capture the differences between segments. In those cases, the true preference segments may remain unapparent. Latent class models are able to identify segments in a sample based on preferences, rather than personal information. A latent class model includes a class membership model which estimates the chance that an individual belongs to a certain class. Further, it estimates a choice model for each class, with unique parameters or even distinct model specifications. As such, the probability of choosing an alternative in a latent class model depends on both the characteristics of the alternative as well as those of the individual. The probability of membership to class s for individual n can be defined as a logit function (Equation 5), similar to Equation 3.

$$(5) \qquad Pr(s|X_n) = \frac{e^{W_{sn}}}{\sum_{t \in S} e^{W_{tn}}}$$

An alternative to the latent class model is the Mixed Logit Model (MLM). Reports indicate that both MLM and latent class models can capture taste heterogeneity and outperform the basic Multinomial Logit model (Greene and Hensher, 2003; Hess, Ben-Akiva, Gopinath, and Walker, 2008). However, several authors argue that latent class models are more appealing. That is, latent class models can relate class membership to personal characteristics, which makes the results very insightful for policy makers (Hess et al., 2008). To do so using a MLM model requires parameterisation of the heterogeneity of the random distributions, which complicates the estimation (Greene, Hensher and Rose, 2006). As argued by Greene and Hensher (2003), a latent class model does not necessitate assumptions regarding distributions. In sum, the semiparametric latent class model might not be as flexible as the fully parametric MNL model, nevertheless it performs equally well and is less complex.

3.5 Stated Versus Revealed Route Choices

Route choice research methods can roughly be divided into stated choice and revealed choice approaches. Stated choice methods measure the route preferences of subjects based on hypothetical choice situations (Hensher, 1994). That is, the researcher selects a set of route attributes and specifies a limited number of levels for them. For example, the levels of the attribute "speed limit" might be defined as "30 km/h", "50 km/h", and so on. Next, scenarios are constructed based on combinations of these attribute levels. Participants of the study are presented with a number of scenarios in which they chose, score or rank the alternatives. The preferences of the participants are then derived based on the aggregate of their choices. Since different combinations of attribute levels are tested, the trade-offs made between them can be measured. The procedure described above is commonly applied by means of an on-paper or online survey, which makes it relatively easy to reach a large audience.

Revealed preference methods are based on real route choice behavior. That is, the researcher records a route chosen by a subject, for example through a GPS tracker (e.g. Broach, Dill, Gliebe, 2012) or a mobile phone application (e.g. Melson, Duthie and Boyles, 2014). Next, a set of alternatives is generated, which represent real life routes between the same origin and destination. The assumption is made that the alternatives in the set were under consideration at the moment the choice was made. The researcher then needs to collect data on the characteristics of these routes. Again, the choices made by the subjects reveal trade-offs made between route attributes.

Stated preference methods have several benefits over revealed preference ones (Broach, Dill, Gliebe, 2012). In particular, the data can be collected by means of a relatively simple and inexpensive survey method. Moreover, there is no need to generate real world alternative routes and collect data on the route characteristics. To add, the hypothetical scenarios allow researchers to study preferences for route attributes which are hard to observe or nonexistent in the real surroundings of the subject. As such, even preferences for futuristic interventions in the cycling infrastructure could be evaluated. These benefits are summarized in Table 3.1.

However, stated preference methods also have their downsides, as shown in Table 3.1. Most importantly, they study hypothetical choice situations. Consequently, they rely on the imagination of the subject, who has to make a choice based on the information provided by the researcher. Moreover, responses could be biased if subjects expect the results of the study to influence policy development (Broach, Dill, Gliebe, 2012). Although a revealed preference method might be more complex and costly compared to a stated one (Hood, Sall and Charlton, 2011), the former does deal with the issues discussed above, for it studies real choice behavior instead. Revealed methods are therefore able to capture more realistic relations between route characteristics and route choices (Chen, Shen and Childress, 2018). Further, the availability of GPS data has increased substantially (Hood, Sall and Charlton, 2011). That is, the growing popularity of activity apps has made large revealed preference datasets available to researchers at relatively low costs. Examples of these applications include "CycleTracks" (Chen, Shen and Childress, 2018; Melson, Duthie and Boyles, 2014; Hood, Sall and Charlton, 2011), "CycleLane" (Zimmermann, Mai and Frejinger, 2017) and "MoveSmarter" (Bernardi, Geurs and Puello, 2018). As touched upon earlier, a downside to the use of revealed preference data in route choice modelling is that routes in a choiceset may overlap. This violates the independence from irrelevant alternatives (IIA) property of the basic MNL model (Train, 2009). However, MNL extensions such as the Path Size Logit model discussed in §3.3 have been developed to tackle this issue. Nevertheless, it is still important to acknowledge that alternative routes in the same area may have similar characteristics. It is therefore important to generate a choice set that is both realistic and contains enough variety.

Table 3.1 - Stated Versus Revealed Route Choice Methods

	Stated Route Choice	**Revealed Route Choice**
Type of Data	Online or on-paper survey	Travel diary (in past), GPS trackers or mobile phone applications
Generation of Choice Sets	Based on hypothetical alternatives, generated following an experimental design.	Based on real world alternatives, assumed to be considered by the decision maker.
Advantages	• Data can be collected by means of a relatively simple and cheap survey method • Can include nonexistent or futuristic scenarios	• Large GPS datasets have become increasingly available to researchers • Studies real choice behavior
Disadvantages	• Relies on the imagination of a subject • Choice data may be biased • Studies hypothetical choices, which may not represent true behavior	• Need to generate realistic route alternatives • Need to collect data on characteristics of routes • Relies on recall in case of travel diary • IIA property of MNL model might be violated due to overlap or similarity between alternatives

3.6 Choice Set Generation in Revealed Route Preference Studies

As discussed in §3.5, revealed choice methods require researchers to generate alternative routes between the observed origin-destination pairs. There are several methods to do so. The most basic option at hand is to generate one alternative which minimizes travel distance. This can be done using Dijkstra's shortest-path algorithm (Dijkstra, 1959). Each link in the network is assigned a cost indicator, based on its length. The algorithm then tries to find the set of links between the origin and destination with the lowest total costs. Since this method provides only one alternative to each route, it does not generally result in a realistic choice set. Moreover, the literature review indicated that cyclists do not detour substantially (see §2.10). Thus, there is a risk that the generated route in the choice sets largely overlap with the chosen one or that their characteristics are very similar. The choisets may then not contain enough variation to capture preferences. Alternatively, a K-shortest path algorithm can be applied to generate multiple routes which minimize distance. However, the variation among these routes may still be minimal, because they are generated based on the same requirement. The "link elimination" technique can reduce this risk by successively eliminating links in the network after the generation of an alternative (Broach, Gliebe and Dill, 2010; Prato, 2009). Instead of link elimination, the costs of certain links may also be increased artificially to redirect routes. It is also possible to vary the link cost indicators in Dijkstra's algorithm, a technique called "labelling" (Prato, 2009). For example, routes can be generated which maximize exposure to green or minimize the number of intersections. Moreover, it is possible to combine different link costs in a weighted cost function. This allows the researcher to generate specific types of routes and compose a choice set which is both realistic and rich in variation.

The methods discussed above are referred to as deterministic techniques. That is, they are fully based on predefined parameters, such as the weights in the cost function. Alternatively, the generation of alternatives can be randomized. Methods that do so are called stochastic. Randomizing link attributes or weights removes the risk of bias introduced by predefined parameters. Moreover, it can increase the number of generated alternatives, as argued by Hood, Sall and Charlton (2011).

Another option to generate choice sets is the use of empirical data. This is possible when a sufficient number of repeated trips between the same origin-destination pair are observed (Bernardi, Geurs and Puello, 2018). The researcher may then assume that the variation among these trips captures the choice situation of a participant. To reduce the risk that some observed trips are too similar, one may cluster trips based on their attributes. For example, Bernardi et al. (2018) group trips into four categories based on the degree of detour from the shortest alternative. Further, Lu, Scott and Dalumpines (2018) study repeated trips between network hubs across different cyclists to ensure a sufficient number of unique alternatives, even if the number of repeated trips for a participant is low.

The different types of methods, deterministic, stochastic and empirical, may also be combined. For example, Hood, Sall and Charlton (2011) apply a method coined as the “doubly stochastic” method by Bovy & Fiorenzo-Catalano (2007). This method combines stochastic randomization of both the link attributes and cost coefficients with a deterministic labelling technique. The distribution of the coefficients was predetermined based on the data of the network to reduce the risk of generating bias routes. A similar method is applied by Skov-Petersen, Barkow, Lundhede and Jacobsen (2018). Broach, Dill and Gliebe (2012) combine “multiple distance constraints” (p.1733) with labelling. They found that this approach resulted in more alternatives compared to the basic labelling technique and more behaviorally realistic results than K-shortest paths. To add, Bernardi, Geurs and Puello (2018) added the shortest path to their empirically obtained alternatives to ensure that at least one alternative is available for routes with none or too few repeated trips.

3.7 Conclusion

The paragraphs above provide a brief overview of the current state of route choice modelling. All three modelling techniques which are discussed, Multinomial Logit (MNL), Path Size Logit (PSL) and Latent Class Logit modelling, are applied in the current study. Specifically, each technique is applied successively to illustrate their added value. Further these techniques are applied to revealed route choice behavior, retrieved from a GPS database. The choice sets are generated using an advanced cycle route planner from De Fietsersbond, which applies a method similar to that of Broach, Dill and Gliebe (2012).

4. Methodology

This chapter starts with an overview of the methodological approach of this study. Thereafter, the separate parts of this approach are elaborated upon. First, the different data sources are discussed. Next, the procedure for generating route alternatives is outlined. Then, the definitions of the context variables are discussed. After this, the model estimation process is described. Finally, the dashboard set-up is elaborated upon.

4.1 Overview

Figure 4.1 provides an overview of the methodological approach of the current study. As explained in §3.6, Multinomial Logit (MNL) route choice models and their derivatives require a set of alternatives to be generated for each observed route. In this study, this is done using De Fietsersbond Routeplanner. The chosen and alternative routes are stored as GPX files and then transferred to a PostGIS database. GIS data is collected and prepared to capture the characteristics of all routes in the choice sets. Personal characteristics of the participants, needed to estimate a probability model in the Latent Class Analysis (LCA), come from a supplementary survey. The Multinomial Logit (MNL) and Path Size Logit (PSL) models will provide insights into the general preferences of cyclists in Utrecht. A PSL model with interaction terms is estimated to evaluate the role of context variables (departure time and trip purpose). The LCA is used to identify segments of cyclists with distinct preferences, to study preference heterogeneity.

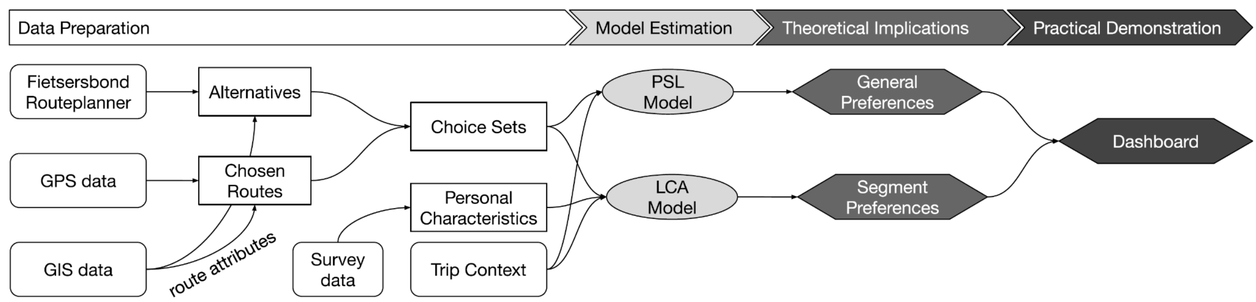


Figure 4.1 - Overview of Methodology

4.2 Data Collection and Preparation

4.2.1 GPS Data

The GPS data for this study has been collected using the ikFiets mobile phone application, developed on behalf of the Province of Utrecht. This app was launched to stimulate inhabitants of the region to cycle more often, both for leisure and transport. Users receive points for every kilometer that they cycle, which they can then exchange for promotions and prices. It also features special bonus challenges to gain extra points. The app is targeted at a wide audience, including e-bike users. Upon installation of the app, users are prompted to create an account. Their activities are linked to this account. Moreover, they can fill in a survey containing basic demographics and questions about their intention to cycle and their physical condition. Further, they are asked to rate the applicability of specific stimulants and deterrents of cycling to them. This survey is filled in on a voluntary basis. That is, the app can be used without finishing it. Hence, only a subset of the users supplies a survey.

Interestingly, the app registers all cycled routes automatically. That is, in contrast to some fitness apps like Strava, all cycle routes are recorded whenever cyclists have their phone on them. Therefore, data collection does not rely on users remembering to turn on the app. Moreover, self reporting bias is kept to a minimum, particularly because unrecorded routes do not generate any points. The modality is determined based on a Baysian network. The probability of each mode (e.g. bike) is calculated and the trip is assigned to the mode with the highest probability score. In sum, the data generated by this app provides a rather complete and unbiased image of the cycling behavior of users.

The trip data used in this study was collected between June 2020 and January 2021 using the ikFiets mobile phone application. During this period, 1107 users generated GPS data, of which 214 filled in the additional survey. Given the goal of this study, to understand the preferences of particular segments of cyclists, only users who filled in the survey are considered. Further analyses will focus on regular trips. These are fuzzy clusters of common trips with the same OD pair and therefore provide a good indicator of repeated travel behavior of users. A total of 205 users have at least one regular trip and filled in the survey. These users generated 5091 regular trips. Two users with unrealistic survey answers were removed from the analysis. In both cases the reported age surpassed 120 years. Further, all trips located outside the borders of the generated network, hence outside the province of Utrecht, were removed as well. To add, trips less than 500 meters long were removed, because generating a rich choice set of alternatives is not possible for routes that short. After this, 139 users remained who generated a total of 743 regular routes. The remaining users did not report owning an e-bike, thus it is assumed that their trips were conducted on a normal bike, without support.

4.2.2 Survey Data

A complementary survey is included in the ikFiets app, which focuses on several personal characteristics and stimulants and deterrents of cycling. This survey is not mandatory for app users, which is why only a subset of them supplied it. It contains questions on demographics including birth year, gender, educational level and household composition. Further, users are asked to indicate which type of bike they have available. To add, participants are asked to rate their own physical condition on a 0-100 scale and their intention to bike on a 0-5 scale. Additionally, users are asked to rate ten motivators and eight deterrents of cycling on a 7-point Likert scale, according to the degree to which those apply to them.

4.2.3 Spatial Data Sources

Table 4.1 provides an overview of the spatial data sources which are used to determine the route characteristics. These sources all have national or worldwide coverage. Thus, the methodology of this study can relatively easily be translated to a different study area in The Netherlands.

Table 4.1 - Overview of GIS Data Sources

Dataset	Coverage	Type	Source
Road Network • Road Type • Cycling Facilities • Speed Limits	World Wide	Lines	Geofabrik (2021)
Traffic Lights	World Wide	Points	
Shops	World Wide	Points	
Air Quality • Pm_{10} • No_x	Nationwide (NL)	Raster	Rivm (2021)
Accidents	Nationwide (NL)	Csv	Rijkswaterstaat (2020)

GeoFabrik

GeoFabrik is a community that collects data from OpenStreetMap and generates datasets for specific locations. Their website offers historic datasets at continental, national and regional levels. In the case of the province of Utrecht, the oldest dataset dates back to approximately two years ago. Since the GPX data of the bike routes was collected during the last half of 2020, the dataset from 2021-01-02 is used in this study.

Coordinate Reference System

Spatial data is unique in the sense that it relates to real locations on the earth. The location and form of this data can be captured by (a series of) coordinate pair(s). A Coordinate Reference System (CRS) is used to make the translation from a real world location to coordinates and back.

Some CRS's have been developed to cover the whole surface of the earth, whereas others are particularly accurate for specific regions only. Further, given the curve of the earth's surface, visualizing spatial data on a screen or paper requires a translation from the 3D world to a 2D space. A projected coordinate system, adapted to the region at stake, ensures that the distortions in this translation are kept to a minimum. In the case of the region of Utrecht, the Amersfoort / RD New projection (EPSG:28992), using the Bessel 1841 ellipsoid, is most applicable, offering an accuracy of one meter. This system is applied consistently across all analyses in this study.

4.2 Context Variables

Several context variables are determined for each trip. Although users of the ikFiets app are not asked to provide information regarding their (regular) trips, some conditions can be derived. For example, the average departure time for a regular trip is known. These are categorized into on-peak (07:00-09:00 and 17:00-19:00) and off-peak hours, representing distinct traffic conditions. Further, the type of trip is derived based on the type of origin and destination location. Specifically, trips between a home and work location, during peak hours, are considered commutes. Further, trips to or from a shopping location are considered shopping trips. Last, trips to or from a leisure location are categorized as leisure trips. Categorization of locations is done based on data from OpenStreetMap and according to the scheme included in Appendix II.

4.3 Generating Route Alternatives

As explained before (§3.6), modelling route choice behavior using a Multinomial Logit (MNL), Path Size Logit (PSL) or Latent Class model requires the generation of choice sets. These are sets of routes between the same origin and destination pair, including the chosen route and at least one alternative. The assumption is made that a cyclist considered these candidate routes when planning the observed trip. It is therefore important that the choice set is a realistic representation of the routes that a cyclist may have considered. For example, routes should not be overly long or follow inaccessible roads. Furthermore, the choice set must show sufficient variation in the attributes under investigation to be able to provide significant results and capture preferences.

Techniques for the generation of alternatives are discussed in §3.6. For example, Dijkstra's algorithm can be used to find the shortest route. Alternatively, K-shortest path search can be used to find a set of alternatives which minimize distance. As discussed in §3.6, both will likely not provide behaviorally realistic choice sets. In contrast, the approach by Broach, Dill and Gliebe (2012) combines multiple distance constraints with a labelling technique to generate several types of routes which each maximize or minimize certain route attributes. They found that this approach resulted in more alternatives compared to the basic labelling technique and more behaviorally realistic results than K-shortest paths. The current study uses a similar approach, based on the route planner of De Fietsersbond, a Dutch cycling association.

The route planner of De Fietsersbond provides a variety of route alternatives, each targeted at a specific audience. Together, these alternatives form a realistic choice set for a variety of cyclists. Table 4.2 describes the route types that were considered in this study. Each route type prioritizes specific aspects. The generated choice sets contain routes which should appeal to cyclists who put efficiency, safety and convenience first (route types: 2, 3, 5 and 9), as well as those who enjoy cycling in green surroundings (route types: 4, 6, 7, and 8). As done by Bernardi, Geurs and Puello (2018), the shortest path is included as well (route type 1).

A Python script is developed to generate the alternative routes and download them from the website of De Fietsersbond as GPX files (See digital repository: GenerateAlternatives.py). Another Python script serves to move the GPX files to a PostGIS database for subsequent processing (See digital repository: Move2PostGIS.py).

Table 4.2 - Alternative Route Types		
Route Type		**Description**
1	Shortest	These routes minimize the number of kilometers to travel. It does not account for any kind of obstacles, inconvenience or delays.
2	Easy Cycling	These routes are focussed on convenient cycling. For example, they avoid traffic lights which may cause delays and require cyclists to get off their bike. As a result, they are relatively fast. They are also easy to navigate, because they follow cycleways along main roads.
3	Conscious Cycling	These routes are somewhat similar to the easy cycling route. However, they prefer traffic lights over roundabouts. They also try to avoid steep gradients and are more likely to select asphalt roads over paved ones.
4	Cycle Network	These routes follow the national cycle network consisting of recommended cycling routes and a network of nodes. The cycling routes get priority over the nodes. If this does not provide a connected route between origin and destination, the gaps are covered according to the “easy cycling” route definition.
5	Low-Traffic	These routes minimize the exposure to motorized traffic. They select separate cycling facilities over those along streets. If no separate facility is available, they will try to follow quiet roads instead of busy ones.
6	Recreational	These routes are targeted at recreational cyclists and follow aesthetic roads. This judgement is based on the presence of nature or other green and unicity of the surroundings. Due to this focus, they usually also have a low exposure to motorized traffic and tend to follow the national cycling routes.
7	Nature	These routes specifically avoid urbanized areas and maximize the exposure to nature.
8	Racing bike	These routes are targeted at racing bike users. Therefore, they avoid unpaved roads and preferably select broad cycleways with high quality asphalt. Other than that, the route selection is similar to that of recreational routes.
9	Winter	In The Netherlands it is common to use salt and sand to keep roads accessible during the cold season. These routes follow the roads where this is done as much as possible. Usually, these concern the main roads and popular cycling infrastructure.

4.4 Model Estimation

This study aims to estimate three types of route choice models, namely, a Multinomial Logit (MNL) model, a Path Size Logit (PSL) model and a Latent Class Analysis (LCA) model. As discussed in §3.6, the PSL model is expected to outperform the MNL model. The results of the PLS model should reveal the general preferences of the sample. Further, interaction effects are used to evaluate the role of the context variables (subquestion B). In turn, the LCA model is estimated to study preference heterogeneity, based on the personal characteristics (subquestion A). The dependent variable in all models is a dummy indicating if a route was selected (1) or not (0). The regular trips are weighted according to their corresponding number of instances, putting more emphasis on often repeated choices. The combinations of independent variables to be entered in the final models are determined on a trial-and-error basis, guided by several correlation matrices and the changes in model fit.

A bi-variate Pearson correlation coefficient (r) captures the strength of a linear relationship among two variables (Illowsk, Dean and Holmes, 2017). A positive coefficient indicates that if one variable rises, the other does too, whereas a negative coefficient is observed for reversed relations. The closer the coefficient is to 1 or -1, the stronger the relationship. Thus, correlation matrices provide early indications of bi-variate, linear relationships in the data and can be used as a guideline for further analysis. With this in mind, several correlation matrices are generated using IBM SPSS. First, a complete correlation matrix of all route characteristics is used to evaluate which of them correlate strongly, either positively or negatively. This is important to know because it might be difficult to enter these variables together in a model, given that they covariate. The same is done for the personal characteristics. Next, a correlation matrix is generated which relates the route characteristics to the route choice behavior. This gives an indication of which factors influence route choices, in what manner, and to what degree. Finally, the sample is split based on several personal characteristics and separate correlation matrices are generated to explore potential differences across groups of cyclists when it comes to the effects of the route characteristics on route choice behavior. This may provide indications of preference heterogeneity.

The Akaike information criterion (AIC/N) and Mc Fadden's rho squared statistic are used to evaluate the goodness of fit. The AIC/N criterion is comparable across different model specifications and penalizes model complexity and therefore helps reduce the risk of overfitting (Cavanaugh and Neath, 2019). The lower the AIC/N, the better. Mc Fadden's rho squared ($pseudo-R^2$) statistic ranges from zero to one and is based on the Log Likelihood ratio between the estimated model (LL) and the null model (LL_0). It can be calculated according to Equation 6. Since the range of this measure is static ([0-1]), it is possible to apply general rules of thumb from the literature to determine whether the model fit is acceptable. For example, according to Hensher, Rose and Greene (2015), a pseudo-R2 above 0.3 is acceptable in discrete choice modelling.

6) $$pseudo - R^2 = 1 - \frac{LL}{LL_0}$$

A Path Size correction factor is included in the PSL and LCA models to account for overlap among alternative routes. This factor is calculated according to Equation 4. As argued in §3.3, the literature is indecisive on the added value of the Generalized Path Size Logit (GPSL) model proposed by Ramming (2002). Since excessive detours are uncommon among the generated alternatives, accounting for them is deemed to make the models unnecessarily complex (Broach, Dill and Gliebe, 2012). In this light, the scaling parameter in the current study is initially set to $\lambda = 0$, yielding the original PLS model proposed by Ben-Akiva and Bierlaire (1999). The Python code developed to calculate this correction factor is included in the digital repository as 'PathSizeCorrection.py'.

The specifications of the final MNL and PSL models are tested for multicollinearity by entering the selected route attributes and, if applicable, their interactions with trip context into a linear regression model in IBM SPSS and generating diagnostics. When all VIF scores are below 4, multicollinearity is deemed unproblematic, following the suggestions by Miles and Shevlin (2001). By comparison, Kang and Fricker (2013) apply a slightly more lenient threshold of 5 in a stated route choice experiment.

4.5 Dashboard

The results of the The Path Size Logit (PSL) model and Latent Class Analysis (LCA) are translated to a dashboard as a demonstration of how the results could be used in practice. To do this, the street network of Utrecht is split up into segments. These segments are based on the street segments in OpenStreetMap (OSM), such that each segment has a unique OSM ID. The characteristics of each segment are determined as if they were routes. Based on this data, the utility of each segment can be calculated according to Equation 1 (see §3.2). The utility values are then normalized to a scale ranging from -1 to +1, according to Equation 7.

7) $$v'_i = -1 + \frac{(v_i - min(V)) \cdot 2}{max(V) - min(V)}$$

Where v'_i is the rescaled utility score for segment i, v_i is the utility for segment i, and V is the collection of utilities for all segments in the network. The values of $min(V)$ and $max(V)$ may change based on the selected class or context. That is, a score of zero always reflects the average utility across the network for the selected class under the selected conditions. It is important to note that the rescaled utility scores cannot be used to calculate the probability of choosing a road segment. Based on this rescaled utility score, the segments in the network are colored in red, orange or green, representing low, average and high scores respectively. A hover-over tooltip contains the values for the variables that are included in the final model, such that users of the dashboard can evaluate the underlying causes of a low or high score. The dashboard is developed in Tableau (Tableau Software, n.d.).

4.6 Conclusion

This chapter outlines the methodology of the current study. In short, the GPS data is enriched with open GIS data to capture the characteristics of the chosen routes. Further, context variables are derived based on departure times and the types of origins and destinations. To add, a Python script is applied to automatically scrape choice sets with various types of route alternatives from the route planner of De Fietsersbond. Thereafter a main effects Multinomial Logit (MNL) model, a main effects Path Size Logit (PSL) model and a Path Size Logit (PSL) model with interaction effects are estimated successively. Then, a Latent Class Analysis (LCA) is conducted, including a model-free comparison of the characteristics of the identified classes. The final results are presented in an interactive dashboard featuring the cycling infrastructure of the province of Utrecht.

5. Data Preparation and Descriptives

This chapter discusses how the collected data is prepared for further analysis and provides an overview of the descriptive statistics. First, the demographics of the sample are discussed. Next, the preparation of the GIS data is elaborated upon. This data is used to determine the route and network link attribute values. Next the characteristics of the cycling infrastructure in Utrecht are summarized. Thereafter, the descriptive statistics regarding the regular trips are presented. Finally, the generated alternatives are discussed.

5.1 Demographics

Of the 139 users which are included in further analysis, about 67% are females and 33% are males. Men have, on average, generated slightly more regular trips (5.8) compared to women (5.1) during the data collection period.

As shown in Figure 5.1, the sample has a reasonable age distribution between twenty and seventy years. About half of the sample is aged between thirty and fifty. Further, ages 65 and up make up about 8% of the sample. Only two participants are less than twenty years old. In sum, most participants are (young) adults and a small portion has reached retirement age.

Figure 5.1 - Age Groups

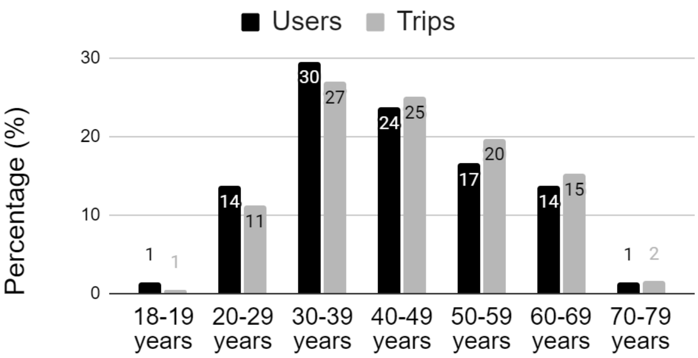
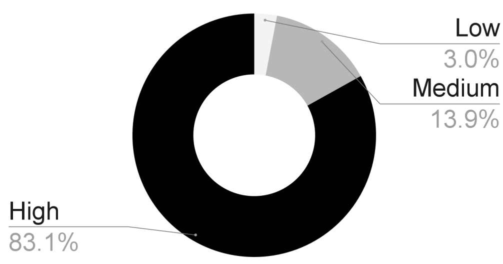


Figure 5.2 - Education

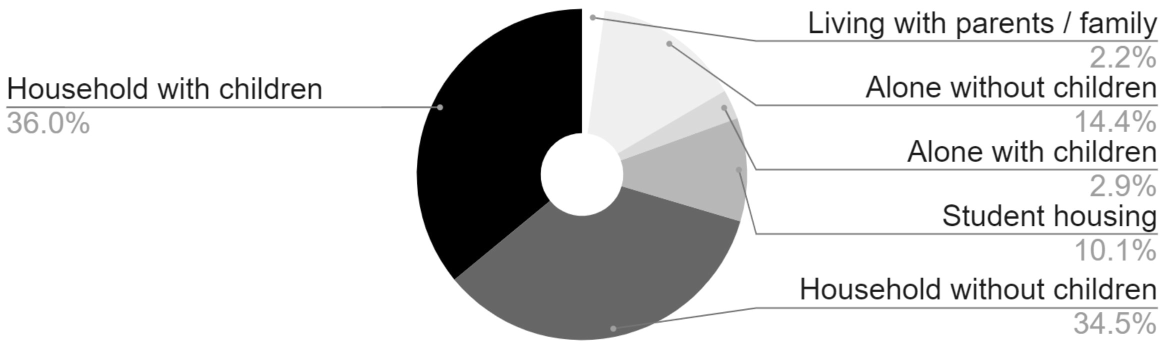


See Appendix III for coding.

As shown in Figure 5.2, The educational level of the sample is relatively high. That is, more than 80% of the sample has at least a Bachelor's degree, against 40% of the total Dutch population (CBS, 2021a). This notion has been observed in other studies regarding cycling behavior (e.g. Anowar, Eluru and Hatzopoulou, 2017 and Winters, Davidson, Kao and Teschke, 2011). It could be that the higher educated have a stronger will to join these kinds of studies, for they may understand their usefulness better. However, Anowar et al. (2017) argue that ridership is simply higher across the higher educated, particularly when it comes to commuting. Nevertheless, it should be acknowledged that the findings of this study might not be transferable to lower educated cyclists.

Figure 5.3 shows the frequencies for the household compositions. Most of the participants, about 70%, live in a two person household, of which approximately half has children. Students and those living in a single person household are also represented by at least 10% of the sample.

Figure 5.3 - Household Composition



The survey also includes questions on self-reported physical condition (100-point scale) and intention to cycle (7-point Likert-scale). The average physical condition score is 75 ($\sigma = 19$). About 7.2% of the sample rates their condition as insufficient (below 55). According to CBS (2021b), about 18.5% of the national population describes one's condition as “not good”. Thus, it appears that the people with a bad physical condition may be underrepresented in the sample. This could be expected since these people might also be less inclined to cycle in general. Further, the average intention to cycle score is 3.7 ($\sigma = 1.7$).

5.2 Preparation of GIS Data

In preparation for the model estimation process, the characteristics of each route have to be determined. Likewise, the attribute values for each network segment of the cycling infrastructure in Utrecht have to be established such that their utility scores can be calculated and displayed on the dashboard. This is done based on the GIS data from the sources discussed in §4.2.3. The data is processed using QGIS and PyQGIS. Several important considerations are elaborated upon below. Thereafter each attribute is discussed individually.

First of all, route characteristics should be comparable across routes of different lengths. It might therefore not always be applicable to use counts[1]. For example, the number of shops along

[1] Counts refer to the total number of occurrences along a route or network segment. For example, “shop count” refers to the total number of shops.

a long route might be high in comparison to a shorter one. However, the density of shops along both routes might be the same. It is important to acknowledge this distinction. Therefore, all characteristics that can be captured in a count are also translated into densities by dividing the count by the total length of the route. At a later stage, the best performing measurement is selected.

Further, special caution is required when dealing with the alternative routes, because these were generated against the road network of De Fietsersbond. This network deviates slightly from that of OpenStreetMap, to which the chosen routes have been mapped. Moreover, most data used to generate the route characteristics is mapped according to OpenStreetMap. Hence, some degree of tolerance is required when matching network links and other OpenStreetMap objects to the alternative routes. Upon visual inspection, the deviations between both networks are limited to about two meters. With this in mind, a tolerance of 2.5 meters is applied when matching OpenStreetMap data to the alternative routes. However, this tolerance may influence the matching process. For example, multiple traffic lights might be matched at an intersection. This may introduce a bias towards a higher number of traffic lights being recorded for alternative routes, in comparison to chosen routes. Therefore, the 2.5 meter tolerance is applied to both route types. Although not ideal, applying the correction to both route types helps to balance the errors.

5.2.1 Number of Traffic Signals

The locations of traffic signals are extracted from the Geofabrik OpenStreetMap dataset as points. These points are then snapped to the closest road network link using QGIS's "Snap Geometries to Layer" algorithm, with a tolerance of one meter. This ensures that traffic lights are precisely positioned on network links. Next, 2.5 meter buffers are drawn around the routes and network segments. The number of traffic lights within each buffer polygon is determined and the resulting counts are joined back to the original line features based on the unique IDs. For reasons discussed earlier, the density of traffic lights along each route is determined by dividing the count of traffic lights over the total length of the route.

5.2.2 Number of Intersections

An intersection is defined as a crossing of three or more street segments in the network. In order to determine the number of intersections in each route, a point layer is created containing all intersections in the network. First, the network links are dissolved based on the OpenStreetMap identifier, such that the segments from OpenStreetMap become uninterrupted lines. This ensures that streets which are not at the same level are not considered to be crossing. For example, a viaduct does not intersect with a street that passes under it. Next, the start and endpoints of each line are extracted. These are the candidate intersections. Thereafter, duplicate candidates are removed with a 0.1 meter tolerance to account for possible small gaps between network links. Next, a buffer of 0.1 meter is created around each remaining candidate and the number of network links crossing that buffer is determined using the “Join Attributes by Location (Summary)” algorithm. Candidates with less than two network links in their buffer are removed. As can be seen in Figure 5.4, these do not concern intersections, but turns along uninterrupted streets. The candidates that remain represent actual intersections of two or more streets, such as shown in Figure 5.5. Similar to the traffic signals, the number and density of intersections are determined for each route and network segment.

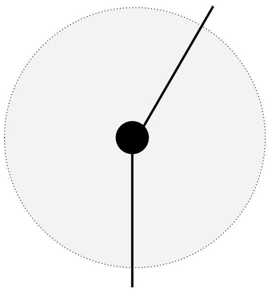


Figure 5.4 - No Intersection

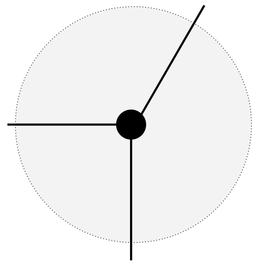


Figure 5.5 - Intersection

5.2.3 Number of Shops and Homes

Similar to the traffic signals, the shop and home locations are extracted from the Geofabrik OpenStreetMap dataset as points. Shops are identified by the OpenStreetMap key “`shop=*`”. Tags considered as homes are: “`building = house`”, “`building = detached`”, “`building = static_caravan`”, “`building = semidetached_house`”, “`building = bungalow`”, “`building = manor`”, “`building = villa`”, “`building = apartments`”, “`building = residential`”.

Again these points are snapped to the closest network link, as was done for the traffic lights. However, the tolerance has to be increased, given that the shop and home nodes in OpenStreetMap are positioned near the entrance of the store, which might be a couple of meters from the centerline of the street. Hence, a tolerance of fifteen meters is applied here. Thereafter, the routes are again buffered with a 2.5m radius. The number of shops and homes within each buffer are counted and joined back to the original route line features based on the unique IDs. Finally, the shop and home densities along each route are determined.

5.2.4 Number of Accidents

The number of accidents on each link is estimated based on data from Rijkswaterstaat. This data contains all accidents that have been administered by the local police forces. It is therefore limited to the more severe kinds of accidents which involve multiple parties, injuries and fatalities. The data on the accidents are stored in a comma-separated values (CSV) file with references to road features in a spatial database. Thus, the number of accidents is only available at a road-level. The total number of accidents in 2019 along each route and network segment is summed and the density of accidents is determined.

5.2.5 Proportion of Cycleway or Cycle Lane

Given the findings of the literature review, a distinction is made between separate and unseparate cycling facilities. The existence of an (un)separated facility is recorded in a boolean variable attached to each link in the network. Separated cycleways can be identified under the "`highway`" key in the OpenStreetMap data and have a "`cycleway=lane`" tag. The presence of a cycle lane along other types of links is recorded with the "`cycleway=lane`" tag. Based upon an inspection of the roads in the network, the tag "`cycleway=track`" is also considered indicative of a cyclelane. In contrast to the "`cycleway=lane`" tag, the "`cycleway=track`" tag officially refers to a cyclelane which is separated from a road by a physical barrier such as curbs, parking or vegetation. However, in reality the links that are marked as tracks in Utrecht turn out to have very minimal separation, usually in the form of a painted flat line or a small curb. Hence, these tracks are closer related to the unseparated cycle lanes compared to the fully separated cycleways.

The proportion of a route or network segment that follows a cycling facility is determined by summing the lengths of all network links covered if these are categorized as cycleway or cycle lane and dividing this by the total length. This is done by buffering all routes and network segments at one meter, filtering the network for the relevant cycling facility and then checking which links are contained by the buffers using the "Join Attributes by Location" algorithm. The lengths of the network links are then summed grouped by the unique route or network segment IDs and divided by the route or segment length. The final result is joined back to the line features based on the unique IDs.

It is important to note that this method ignores the first and last network link in the route. This happens because these links will not be completely contained by the route buffer. However, less restrictive predicates such as "cross" or "intersect" will result in intersecting roads to be included in the route. Several other methods were considered to overcome this issue. However, they all resulted in a significant increase in runtime. Fortunately, the impact of the exclusion of the first and last link of the network is minimal, given that these only represent a small section of the route. Moreover, the first and final link are the same among alternatives, given that the origin and destination are located along these links. Thus, the method described above is suitable to measure the proportion of a route that follows a cycling facility, despite the slight deviation from reality.

5.2.6 Cycling Facility Interruptions

To determine the number of interruptions in cycling facilities along each route or network segment, one first needs to determine which links in the network with cycling facilities are connected. This is done by merging adjacent network links with facilities into a single feature. To account for small imperfections on the OpenStreetMap network, a tolerance of ten centimeters is applied by buffering the network links. Next, the buffers are dissolved and split into unconnected single parts, which are then provided with a new unique id (uuid). Next, the route and segment features are buffered at 2.5m, again to account for the deviations between the OpenStreetMap network and that applied by De Fietsersbond. Next, the unconnected network parts with facilities within each buffer are counted. Finally, the resulting number of interruptions are joined back to the original line features. Figure 5.6 presents an example of a route and three identified facility interruptions.

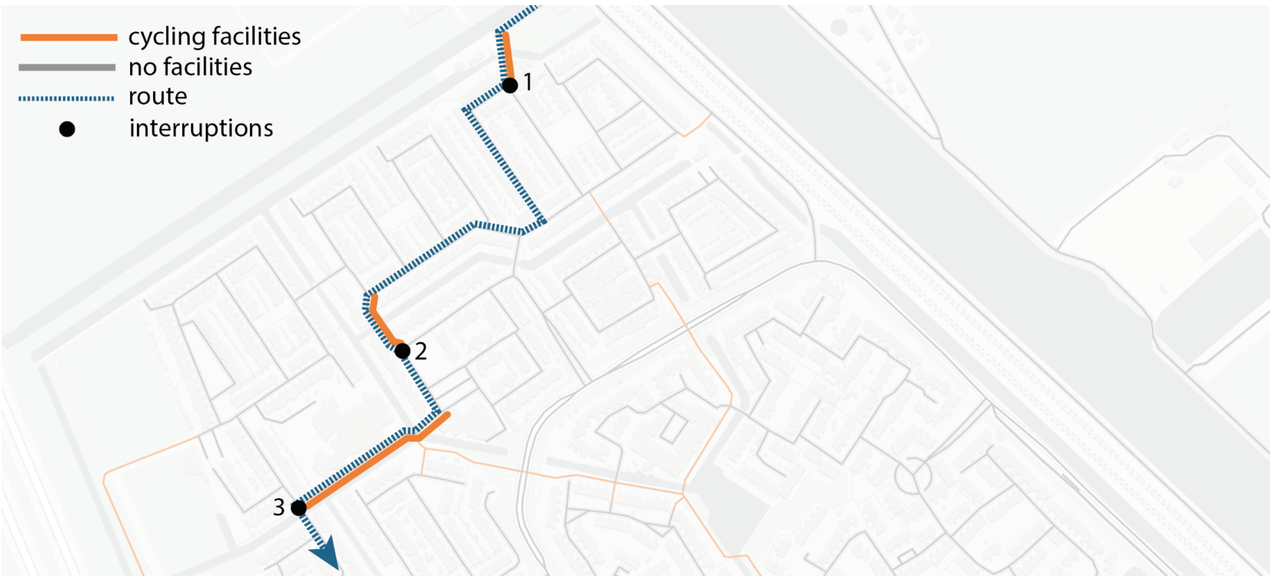


Figure 5.6 - Cycling Facility Interruptions

5.2.7 Air Quality

Few other studies have considered air quality as a factor which may influence route choice behavior of cyclists. Among the reviewed articles, only Anowar, Eluru and Hatzopoulou (2017) did so in the context of America and Canada. They conclude that some commuters, particularly experienced cyclists, have a tendency to avoid areas with high rates of pollution. This factor is therefore included in the current study, to possibly confirm this finding.

The air quality is measured in the form of PM_{10} and NO_x levels. These concentrations are indicative of traffic volumes and have a large impact on health (GGD and RIVM, 2014). It should be noted that these measurements are only available at a relatively low resolution of one by one kilometer. The concentration levels are related to the route and network segment features by creating buffers of one meter around each network link and taking the mean concentration within those buffers using the "Zonal Statistics" algorithm in QGIS.

5.2.8 Weighted Average Speed Limit

Unfortunately, data on the speed limits of roads in OpenStreetMap is often incomplete. However, the road type generally is available, based on which the speed limit can be inferred. Indeed, as can be seen in Figure 5.7, the number of network links without a known speed limit and road type is limited and generally pertains to links along squares or parking places. Therefore, an inferred speed limit is determined based on a set of rules (see pseudo code below) to replace null values in the OpenStreetMap data. The speed limits of cycleways, tracks, service roads and unclassified streets are set to 0 km/h. These rules are applied to the network link features using PyQGIS, following the logic shown below.

```
if osm_maxspeed is not NULL:
    speedlimit = osm_maxspeed
elif osm_roadtype is in (“residential”, “living street”):
    speedlimit = 30
elif osm_roadtype is in (“secondary”, “tertiary”):
    speedlimit = 50
else:
    #cycleways, tracks, service roads, unclassified and other
    speedlimit = 0
```

The contribution of each network link within a route or network segment to the average speed limit is weighted for the length of the link , according to the following equation:

8) $$WA = \frac{\sum_{l \in L_i} (speedlimit_l \cdot length_l)}{\sum_{l \in L_i} length_l}$$

Where L_i refers to the set of links included in route or network segment i

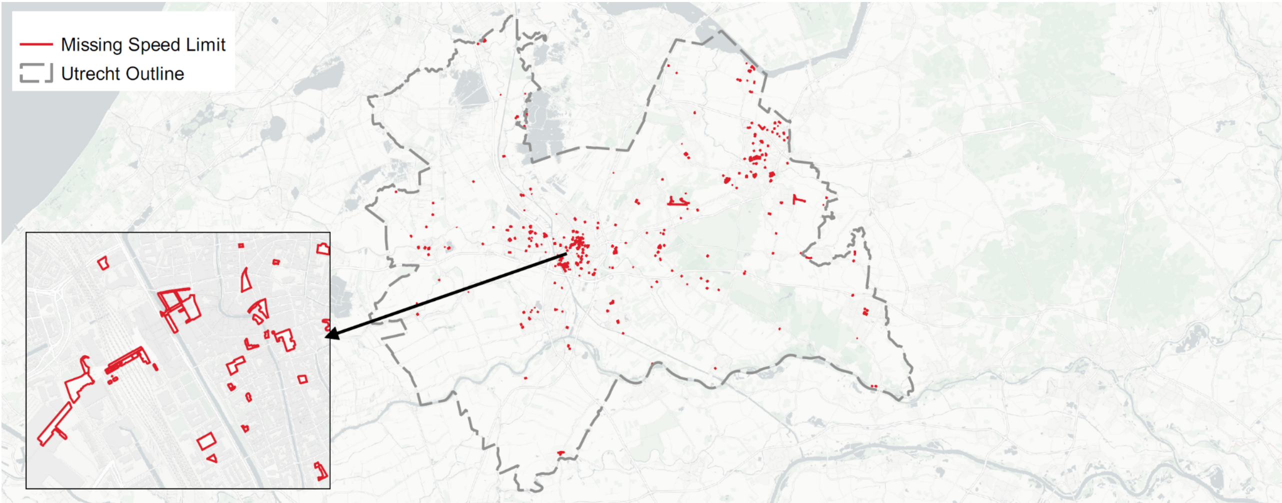


Figure 5.7 - Streets Without Speed Limit and Road Type

The weighted average speed limit of a route is determined both including and excluding the assumed speed limit of 0 km/h for unclassified roads. In the latter case, the denominator in Equation 8 concerns the sum of the lengths of those links within a route for which the speed limit is known. Equation 8 is applied to the network data in QGIS in the following manner. First, the route features are buffered at 2.5m, again to account for the deviations between the OpenStreetMap network and that applied by De Fietsersbond. A spatial index is created to boost the performance. Next, the QGIS "Join Attributes by Location" algorithm is used to join the applicable data of the network links to the route buffers which contain them. This is done on a one-to-many basis, meaning that the route buffers are duplicated for each matching network link. In the case of the average determined only for network links with a speed limit based on OpenStreetMap, a filter is applied to the resulting layer (`"NW_SpeedLimit" is not NULL AND "NW_SpeedLimit" != 9999`) such that links with an imputed speed limit are excluded. After filtering, the total length of the known network links in each route is determined using the field calculator (`sum("NW_length", group_by:="newid")`). For the calculation concerning all links, no filter is applied and the total length is determined based on the length of the route using the field calculator (`$length`). Next, the contribution of each, weighted for its length, is calculated based on Equation 8. Thereafter, these contributions are summed for each route or network segment (`sum("speed_weighted", group_by:="uuid")`), providing the weighted average speed. This result is then joined back to the original route or network segment features.

5.2.9 Landuse

To grasp the exposure to specific land use types, a PyQGIS algorithm is developed that takes a raw OpenStreetMap database file (.pbf format) and a QGIS line layer containing the network link or route features. This algorithm generates buffers of 25, 50 and 100 meter around each line feature and then determines the proportion of the area within each buffer categorized as a particular land use type. The land use categorization of OpenStreetMap is used as a basis, although some land use types are combined as a simplification, as seen in Table 5.1. Important to note is that a large number of meadows in the province of Utrecht are tagged as '`landuse=grass`' without an indication of an agricultural area, also not under other keys. The distinction between the "green - general" and "green - agriculture" category is therefore minimal. The process of the algorithm is summarized in Figure 5.8. The code is available in the digital repository ("landuses.py").

Table 5.1 - Land Use Categorization

Land Use Category	OpenStreetMap Land Use Types
residential	residential
green - general	allotments, animal_keeping, apiary, farm; grass, farmyard, farmyard;residential, framland, greenhouse_agricultural, greenhouse_horticulture, meadow, orchard, plant_nursery, vineyard, yard, grass, forest, forest;grass, garden, nature_conservation, nature_reserve, park, village_green
green - agriculture	allotments, animal_keeping, apiary, farm; grass, farmyard, farmyard;residential, framland, greenhouse_agricultural, greenhouse_horticulture, meadow, orchard, plant_nursery, vineyard, yard, grass
commercial	commercial
retail	retail
industrial	industrial, depot, landfill, salvage_yard

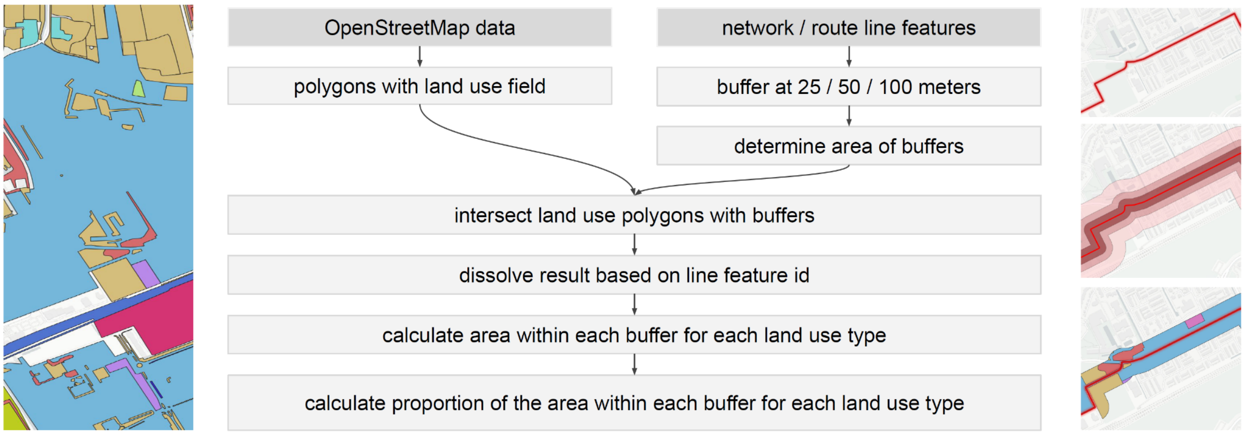


Figure 5.8 - Determining Land Use Proportions

5.3 Network Characteristics

Table 5.2 provides an overview of the descriptive statistics of the network links. As shown in this table, most links have a speed limit of 30 km/hr, as is standard for living streets in The Netherlands. It must be noted that streets with a speed limit above 80 km/hr have been excluded, since these are deemed unsuitable for cyclists. Although uncommon, some streets have a streetlimit below 30 km/hr. These are service roads, special residential roads or parking lots. The air quality measures do not deviate substantially. As shown in Figures 5.9 and 5.10, the air quality in the city centers is somewhat worse compared to the rest of the province. The links in the network are more or less straight, with only 0.34 turns on average (σ = 0.87). Further, they intersect with three other links (σ = 5) and pass 0.04 traffic lights on average. Traffic lights (Figure 5.11) and shops (Figure 5.12) are both mainly concentrated in strongly urbanized areas. Looking at the land use types, it can be concluded that most links are located in residential areas with some green. Other land use types are relatively rare.

Table 5.2 - Descriptives Network Links					
	Minimum	**Maximum**	**Mean**	**Std. Deviation**	**Median**
SpeedLimit	5	80	33	10	20
PM10 Level	15	21	17	0.4	17.4
NOx Level	9	44	25	3.9	26.3
Turns Count	0	39	0.34	0.87	0.0
Turns Density (/km)	0	0.51	0.002	0.007	0.0
Traffic Lights Count	0	28	0.04	0.33	0.0
Intersections Count	0	284	3	5	2.0
Homes Count	0	1372	16.17	43.63	0.0
Shops Count	0	96	0.09	1.24	0.0
Bridges Count	0	12	0.01	0.33	0.0
Land Use in 50m Buffer:					
- Agricultural Green	0%	100%	1.8%	8.5%	0.0%
- General Green	0%	100%	32.1%	32.6%	19.0%
- Commercial	0%	100%	2.6%	14.0%	0.0%
- Retail	0%	100%	0.7%	6.4%	0.0%
- Industrial	0%	100%	4.5%	19.0%	0.0%
- Residential	0%	100%	95%	19%	100%

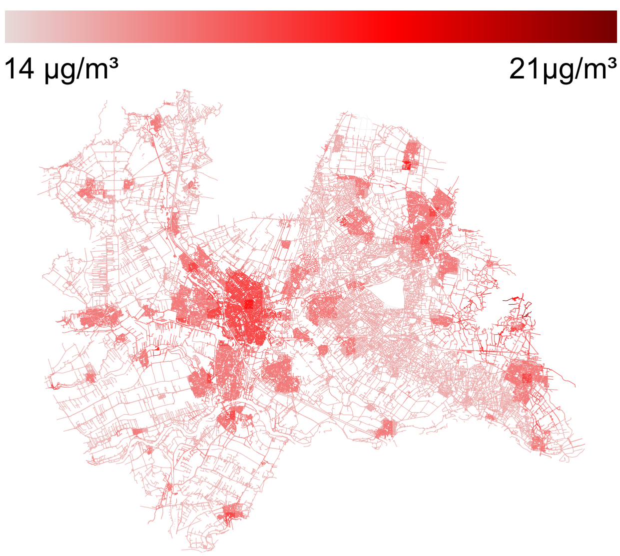


Figure 5.9
PM_{10} Concentrations At Network Link Level

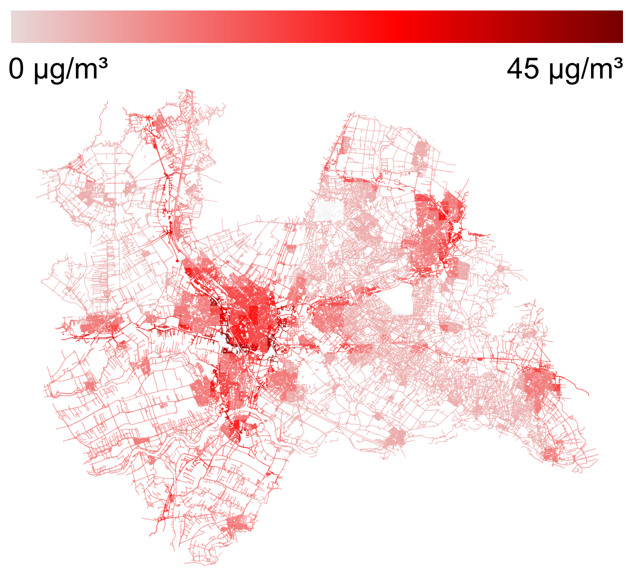


Figure 5.10
NO_x Concentrations At Network Link Level

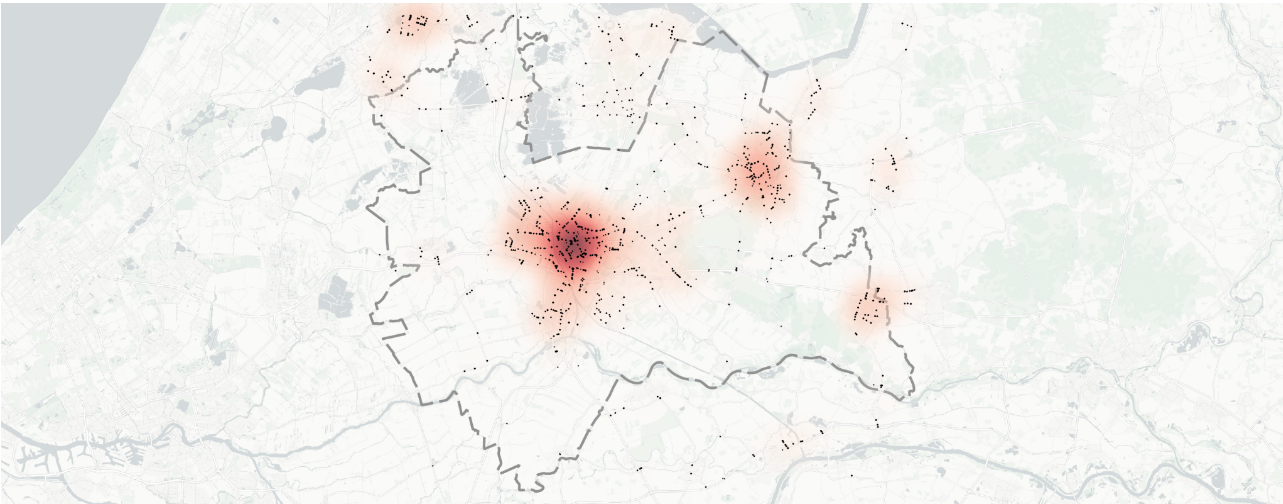


Figure 5.11 - Traffic Lights

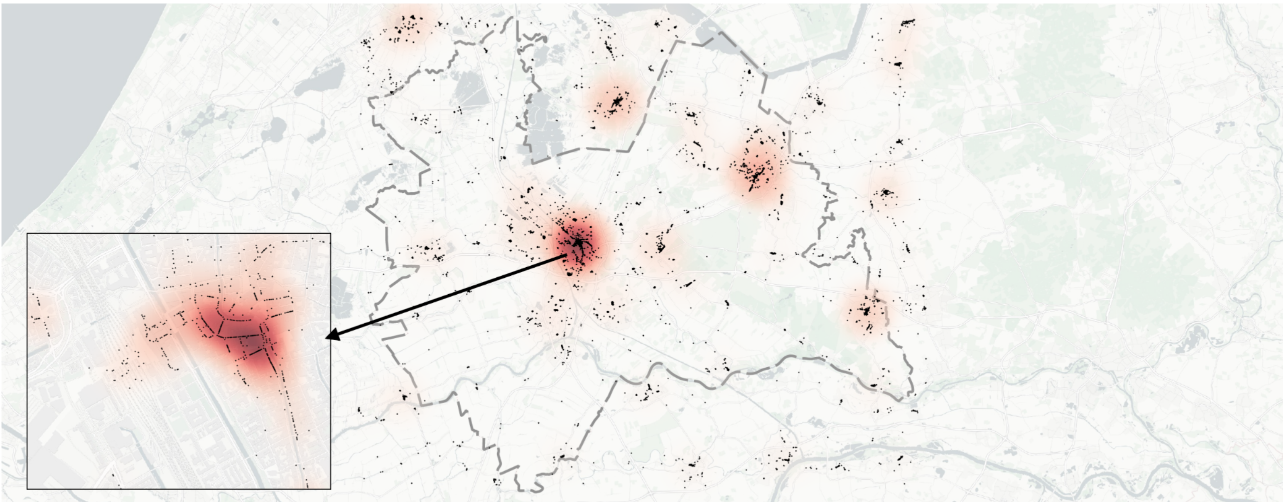


Figure 5.12 - Shops

5.4 Trip Characteristics

As can be seen in Figure 5.13, the starting times of the trips are spread out throughout the day. A small peak is visible around five ‘o clock, possibly due to cyclists returning home from work or other activities to have dinner. However a clear morning and evening peak are not evident. This is likely a consequence of the changes in travel behavior due to the Covid-19 pandemic, during which the data was collected.

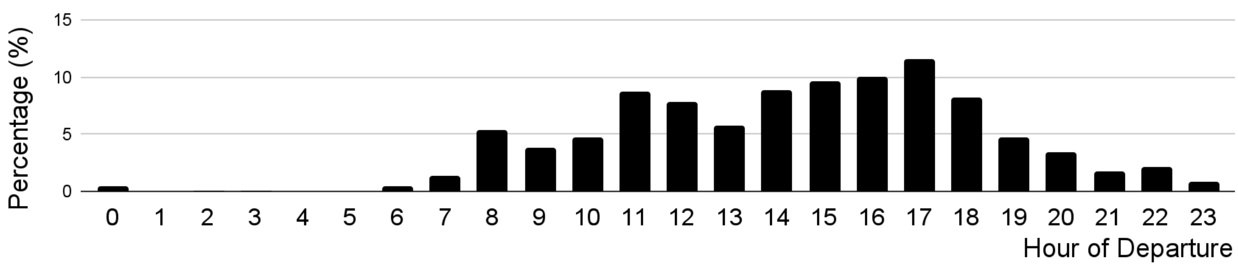


Figure 5.13 - Departure Time

As can be seen in Figure 5.14, most routes are between one and two kilometers long. The degree of detour seems to increase slightly for longer routes. The negative average degree of detour for routes less than one kilometer indicates that the shortest route generated by the Fietsersbond route planner is sometimes slightly longer compared to the chosen route. This may happen if someone took a shortcut, which is officially not accessible for cyclists. Further, there are minor deviations between the network used by the Fietsersbond route planner and the OpenStreetMap network to which the GPS data was mapped. The average degree of detour (mean = 5.5%, σ = 50%) is on the lower side of the values reported by reviewed articles (range 5% - 15%, see Table 2.16). Fitch and Handy (2020) observed a similar willingness to detour (5%) among students and staff members in Davis. They argue that most of the trips in their sample are likely to be commutes, which explains why the cyclists pick highly efficient routes. That is, commuters are known to be less willing to detour compared to others (Broach, Dill and Gliebe, 2012; Sener, Eluru and Bhat, 2009).

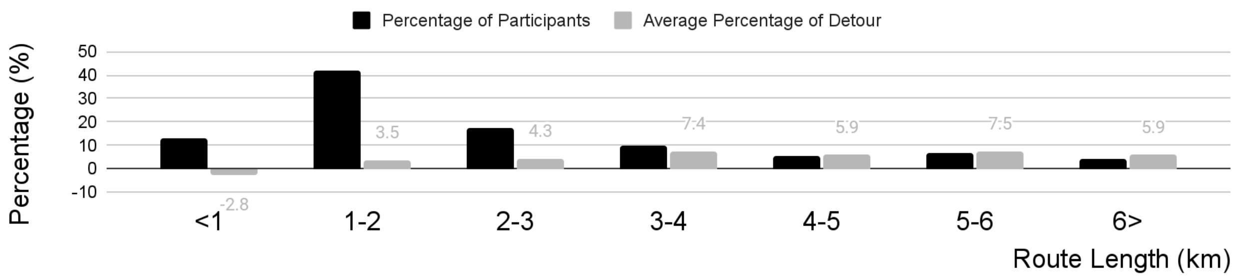


Figure 5.14 - Travel Distance

As discussed in the methodology, the origins and destinations of each trip are categorized as work, shopping, leisure or other, according to Appendix II. Trips between home and work locations, with a departure time between 07:00-09:00 and 17:00-19:00 hours, are considered on-peak commutes. As shown in Figure 5.15, a large number of routes are not classified. In those cases there was no clear origin or destination on which the categorization could be based or the type of the location was ambiguous. In comparison to other studies, the proportion of commute trips appears relatively low. In part, this can be attributed to the limitations of the derived categorization. However, this can also be a consequence of the large number of people working from home during the Covid-19 pandemic. In later analyses, both on-peak and general commute are considered, to account for potentially less regular start and end times of workers on account of the pandemic. The best performing measure is selected.

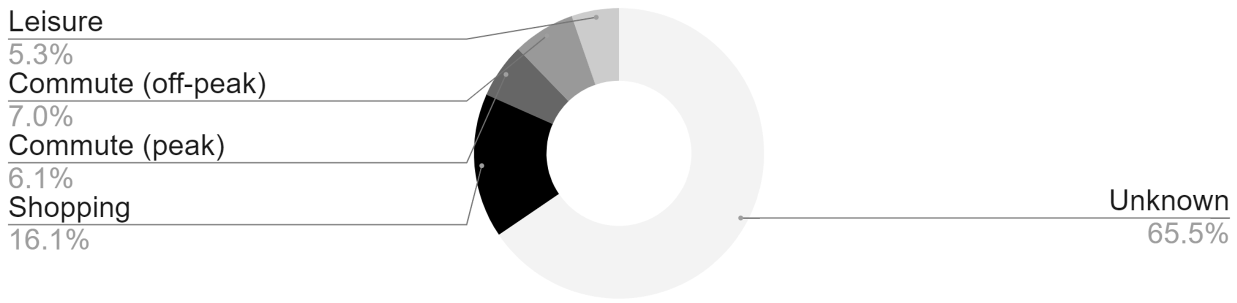


Figure 5.15 - Trip Purpose

As seen in the heatmap in Figures 5.16 and 5.17, most trips originate and end in the city center of Utrecht. It is important to note that this does not necessarily mean that cycling is more popular in the city compared to the whole province. The ikFiets app may simply be more promoted or popular among citizens of the city of Utrecht.

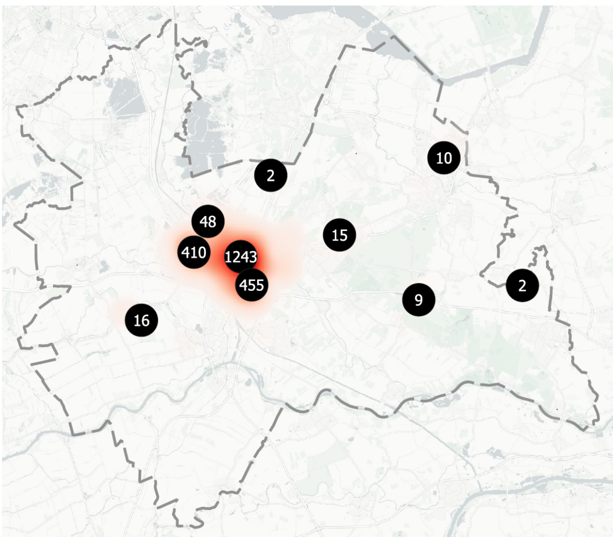


Figure 5.16 - Origins

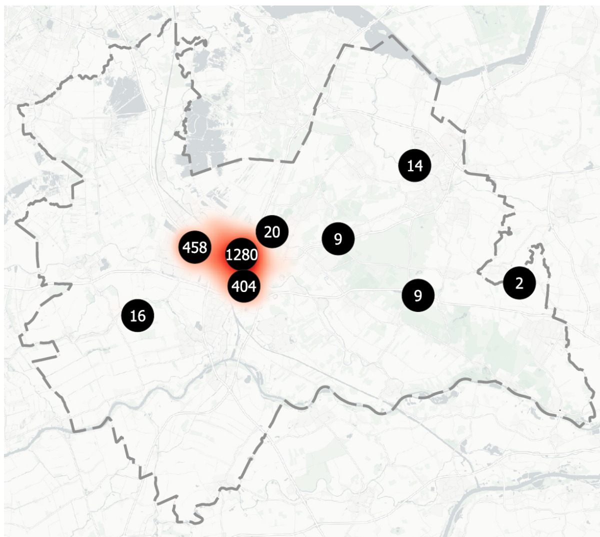


Figure 5.17 - Destinations

5.5 Alternative Routes

For each route in the sample, nine types of alternative routes were generated, according to the procedures in §4.3. Figure 5.18 provides an example of a chosen route and a corresponding set of alternatives. In this example, there are several clusters of routes which overlap. Some clusters follow a main road (e.g. the shortest, and conscious cycling routes), whereas others deviate further from the shortest route, into the rural areas (e.g. the nature, recreational and low-traffic routes).

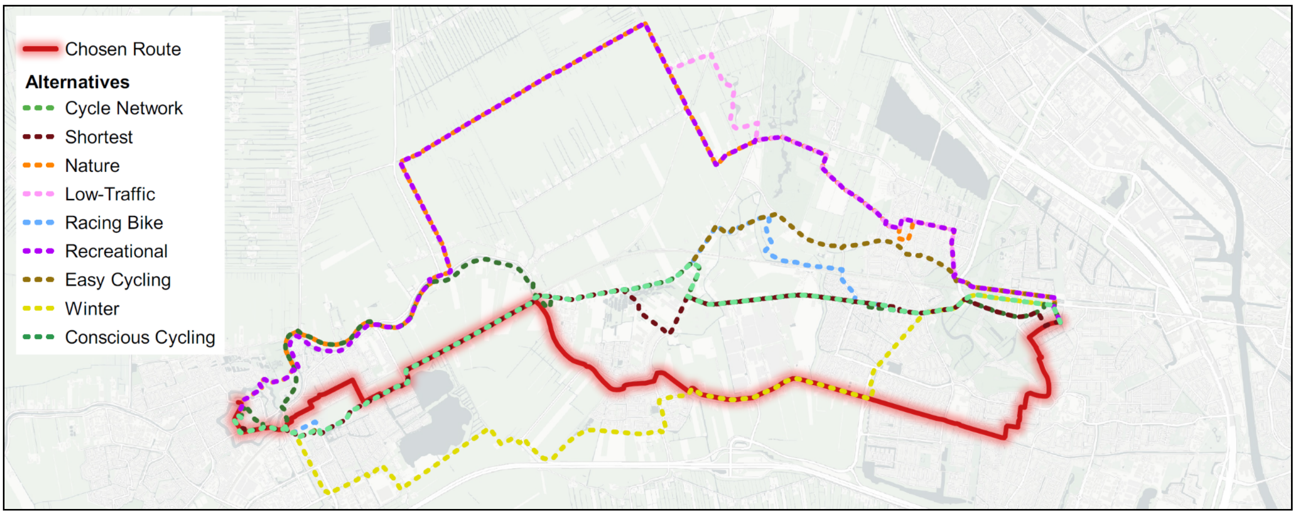


Figure 5.18 - Example Generated Alternatives

As shown in Table 5.3, the route types differ substantially in their degree of detour. The low-traffic, recreational, cycling network and nature routes show the highest degree of detour (15%+). These routes also tend to be more green, conforming to the descriptions provided in Table 4.2 (§4.3). The chosen routes overlap, on average, for about 46% with their alternatives. Further, five of the route types have an average degree of detour within the expected acceptable range according to the literature review (5% - 15%). All in all, the generated choice sets appear to contain several realistic alternatives in terms of detour and sufficient variation in terms of green versus urbanized areas.

Table 5.3 - Comparison Generated Alternatives

Route Type	Degree of Detour	Green (within 50m buffer)
Shortest	0%	16%
Easy Cycling	3%	18%
Conscious Cycling	5%	17%
Winter	7%	17%
Racing bike	9%	19%
Low-Traffic	17%	20%
Recreational	22%	21%
Cycle Network	25%	18%
Nature	33%	20%

5.6 Conclusion

This chapter elaborates on the data that is used to estimate the route choice models. Specifically it describes how the GPS data was enriched with open GIS data to determine the characteristics of chosen routes. Moreover it provides an overview of the network and (alternative) trip characteristics as well as the demographics of the sample. As such, it communicates the context to the findings presented in the next chapter.

Important takeaways include the following. First, the sample is predominantly highly educated, an issue that is common among similar studies. To add, females are slightly overrepresented. Further, it appears that the people with a bad physical condition are underrepresented in the sample. This could be expected since these people might also be less inclined to cycle in general. In terms of age and household composition, the distributions are comparable to national figures. The relations between these personal characteristics and route choice behavior are explored in the Latent Class Analysis (LCA). Regarding the network characteristics it is noteworthy that most links are located in residential areas with a speed limit of 30 km/h. Further, both shops and traffic lights are concentrated in urbanized areas. The most evident finding related to the chosen trips is the low degree of detour of no more than 5.5%. Further, trips are generally no longer than about 3km and both the origins and destinations concentrate around the city of Utrecht. For about 35% of the trips the purpose could be categorized according to the predefined rules. Specifically, 7% and 6% of the trips were categorized as off and on peak commute respectively, 16% as shopping and 5% as leisure. As discussed in the final paragraph, the choice sets contain routes which differ substantially in their degree of detour and exposure to green.

6. Results

This chapter presents the results of the study. First, the correlation matrices are discussed, which guide the model estimation process. Then, the main effects multinomial logit (MNL) model is discussed. Thereafter follows the main effects Path Size Logit (PSL) model, which contains a Path Size correction factor. Next, the PSL model with interaction terms is discussed. Finally, the latent class analysis (LCA) is elaborated upon, including a model-free comparison of the characteristics of the identified classes.

6.1 Correlations

The paragraphs below discuss the correlation matrices which were generated to support the model estimation process, as discussed in the methodology (§4.4). The correlations among route characteristics and those among personal characteristics suggest which combinations may cause issues of multicollinearity. The correlations between the route characteristics and route choice behavior helps identify important determinants of route choice behavior. Finally, the correlation matrices for specific subsamples indicate which personal characteristics relate to unique preferences.

6.1.1 Correlations Among Route Characteristics

Appendix V (see also digital repository) features the correlation matrix for the route characteristics. Several strong correlations in this matrix can be attributed to trivial relations between variables. For example, accidents are positively related to intersections. This makes sense because intersections increase the risks of collisions. To add, densities are derived based on counts and these are therefore always to some degree related. Further, some characteristics commonly coexist. To illustrate, traffic lights are generally located at intersections. Consequently, the correlation between the number of traffic lights and intersections is high. The same holds for bridges and water. Further, several route characteristics measured as occurrences along a route correlate strongly with the length of a route. This makes sense because the longer a route, the higher the chances of encountering some aspect. For example, longer routes can be expected to, on average, pass more traffic lights, bridges, turns and intersections. The same holds for routes with a longer total distance of cycleway. If one looks at the proportion of cycleway along a route instead, these correlations diminish.

However, some high correlations deserve extra attention. Specifically, some aspects seem to correlate strongly, because they concentrate in city centers. For example, the number of shops relates positively to the number of accidents and the presence of monuments. This can be explained by the fact that many of the stores and monuments are located in the busy city center of Utrecht, where accidents might be more likely to happen. Further, routes in urbanized areas also seem characterized by relatively high levels of PM10, which can be related to high traffic volumes and congestion. Among the land uses, the most notable observation in this regard is the negative correlation between green and residential. This makes sense because there is limited space for green in strongly urbanized areas. In sum, the distinction between urbanized and not urbanized areas seems to be an important source of high correlations among route attributes.

Further, several strong correlations may come as a surprise. For example, shops do not relate positively to retail areas. However, it is important to note that the city center of Utrecht is defined as a residential area, instead of retail, in OpenStreetMap. Further, the number of turns and intersections have a strong positive correlation, whereas their densities do not. To add, the number of turns and intersections appears to be higher among routes that also pass a large number of homes. This may be related to the high connectivity of residential streets. Furthermore, the weighted average speed limit does not relate strongly to any other route characteristics. Possibly, this is the case because most roads that were passed have relatively low speed limits (see Table 5.2), yielding limited variation. Overall, most of these bivariate relations can be logically explained. Nevertheless, they must be kept in mind when entering combinations of variables into the model, to avoid multicollinearity.

6.1.2 Correlations Among Personal Characteristics

Appendix VI (see also digital repository) shows the correlation matrix for the personal characteristics of the cyclists, the data obtained through the survey. Most noteworthy are the strong correlations among the motivators as well as the deterrents of cycling. To illustrate, the average correlation coefficient among the deterrents is *0.44*. Among the motivators there seem to be two clusters of variables with strong correlations. First, a group related to the convenience of cycling for transportation, including ease, security, speed and traffic costs. The other can be related to the positive experience of cycling, including benefits for physical and mental health, enjoyment and pleasure from being outside. Interestingly, the correlation matrix also indicates that females less often own a race or mountain bike, seem to enjoy cycling less and have a lower self-reported physical condition compared to their male counterparts. Further, enjoyment of cycling seems to increase with age ($r = 0.40$). The fact that these variables covariate can be expected to complicate the estimation of a model that includes multiple motivators or deterrents and personal characteristics. To gain a deeper understanding of the composition of the classes in the latent class model and what stimulates or discourages them may therefore require further analysis beyond the output of the latent class model.

6.1.3 Correlations Between Route Characteristics and Route Choice Behavior

Appendix VII provides the correlations between the route characteristics and route choice behavior, that is, the dummy variable for whether a route was chosen or not. As can be seen in the correlation matrix, these correlations are somewhat low. This could be an indication of conflicting preferences across specific segments, which may cancel eachother out. However, some relations do stand out and indicate differences between chosen routes and the alternatives. For example, the negative correlation for the degree of detour (*r = -0.06*) indicates that the chosen routes are generally shorter than the alternatives. Further, they appear to less commonly go through green areas, as indicated by the negative correlations for the proportion of green area within the 25m, 50m and 100m buffers. Moreover, the weighted average speed with imputations seems lower among chosen routes. The same goes for intersections and turns, both in terms of counts and densities. Interestingly, the bike facility interruption density appears to be higher among chosen routes. Overall, the degree of detour, presence of green, speed limits, turns and intersections and bike facility interruptions seem to have the strongest relations with route choice behavior.

6.1.4 Differences Based on Personal Characteristics

Appendix VIII (see also digital repository) features the correlations between the route characteristics and choice behavior for specific subsamples. These statistics were generated to gain insight into potential segments of cyclists with unique preferences. The subsample of cyclists with children stands out, showing much stronger correlations compared to the other groups. However, it must be noted that only four users reported to have children, which makes the results unreliable for generalization.

Some correlations appear to follow a similar pattern across the subsamples. Specifically, preference heterogeneity regarding turns and intersections appears to be minimal. However, interesting differences can be observed for other route characteristics. First, young adults appear to be particularly sensitive to the degree of detour (*r = -0.13*). In contrast, people who own a race bike seem to have a relatively high willingness to detour (*r = -0.01*). Interestingly, people who report not to cycle for enjoyment and young adults seem to be particularly discouraged by a high traffic light density (*r = -0.05 and r = -0.10*), whereas seniors do not seem to be bothered by traffic lights at all (*r = 0.00*). Further, those who do not seem to enjoy cycling appear to cycle in areas with shops, possibly in the city center. Surprisingly, most cyclists do not seem to choose green routes. However, this might be related to a high number of utilitarian trips in the sample. Hence it would be interesting to evaluate the interaction between a green landscape and trip purpose to study this relation in more detail. Interestingly, agricultural green appears to be particularly popular among cyclists who have reached their retirement age (*r = 0.13 for 50m buffer*) and low among adults aged 30-50 years (*r = -0.07 for 50m buffer*). Air quality, measured by PM10 and NOx levels, appears to be particularly bad along routes chosen by those who do not enjoy cycling, again indicating that the trips of those people may be restricted to the city centers. The same reasoning may apply to the relatively high number of monuments along their routes (*r = 0.14*). Further, considering the average imputed speed limit, most subsamples appear to choose routes with relatively low traffic speeds. However, young adults (*r = -0.17*) appear to be the most sensitive to high traffic speeds. In sum, these correlations do indicate that some preference heterogeneity exists.

6.2 Modelling Results

The set of independent variables to be entered in the final Path Size Logit (PSL) model and Latent Class Analysis (LCA) were selected on a trial-and-error basis, guided by the findings in §6.1. The Akaike information criterion (AIC/N) and Mc Fadden's rho squared were used to evaluate the goodness of fit. More than fifty model specifications were tested, varying in terms of route attributes, trip context variables and personal characteristics.[2] This process is summarized below. For all models, VIF scores were below the threshold of 4 (Miles and Shevlin, 2001; Kang and Fricker, 2013), indicative for the absence of problematic degrees of multicollinearity. Details regarding the specific models are discussed in the subsequent paragraphs.

First, a Multinomial Logit Model was estimated. The explanatory variables with the highest correlation with choice behavior were entered first (*r > 0.05*). Count and density measurements were alternated to evaluate which combination results in the best model fit. Those independent variables yielding insignificant results were omitted iteratively to evaluate the consequences for the remaining ones. Although the density of bike facility interruptions correlates relatively strongly with route choice behavior (*r = 0.10*), it was excluded due to its interpretability. That is, the positive correlation and MNL coefficient implies that cyclists prefer routes with a high density of bike facility interruptions, which is counterintuitive. This does not seem to be attributable to the notion that more bike facility interruptions also means more bike facilities, since the correlation between those two is negatieve (*r= -0.04*). Plausibly, the documentation of bike facilities in Utrecht by OpenStreetMap is lacking. As such, variables related to bike facilities were abandoned.

[2] Intermediate results are available upon request.

Thereafter, the Path Size correction factor was introduced, yielding a PSL model. The MNL model definition was taken as a starting point and again route attributes were included and excluded iteratively in an attempt to improve the model fit. The final PSL model definition remained the same as that of the MNL model, including the correction factor.

Next, several interaction effects with context variables were introduced to the PSL model. These context variables concern trip purpose (leisure, shopping, commute and other) and departure time (off- versus on-peak). Earlier excluded variables were reintroduced to test whether their effects might be context dependent.

Finally, a Latent Class Analysis was performed, based on the PSL model with interaction effects. The findings discussed in §6.1.4 were used to select additional explanatory variables which may show preference heterogeneity and personal characteristics which discriminate between the classes. The added independent variables were also included post hoc in earlier models to demonstrate that certain opposing preferences cancel eachother out.

6.2.1 Main Effects Multinomial Logit Model

Table 6.1 presents the results of the main effects Multinomial Logit Model. The McFadden Pseudo Rho square statistic indicates a moderate model fit (Hensher, Rose and Greene, 2015). Further, all VIF scores are below the threshold 4, the highest being 3.5 (see Appendix IX). The model mainly contains route attributes which are related to traffic safety, efficiency and convenience. An alternative model which considers only the non imputed speed limits was also tested (see Appendix IX). The performance of this model is similar to the one that includes imputed speed limits. However, the coefficient for the non imputed speed limit is positive, indicating that cyclists would select routes with higher traffic speeds compared to their alternatives. Since this seems implausible, the imputed version is used in subsequent analyses.

Table 6.1 - Results Main Effects Multinomial Logit Model			**McFadden Pseudo Rho square** = 0.24, **AIC/N** = 28.6				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Degree of Detour	**-0.07**	**	0.00	-34.51	0.00	-0.07	-0.07
Turn Density (/km)	**-0.39**	**	0.01	-28.64	0.00	-0.42	-0.37
Number of Intersections	**-0.10**	**	0.00	-26.30	0.00	-0.11	-0.10
Number of Traffic Lights	**-0.11**	**	0.01	-13.47	0.00	-0.13	-0.09
Speed Limit (Imputed)	**-0.08**	**	0.01	-16.53	0.00	-0.09	-0.07
Agriculture (50m Buffer)	**0.01**	*	0.01	1.83	0.07	0.00	0.03
Number of Bridges	**0.46**	**	0.02	23.98	0.00	0.42	0.50

** significant at 5% level, ** significant at 1% level*

Several interesting observations can be made regarding the results in Table 6.1. Cyclists appear to be sensitive to the degree of detour (*β = -0.07, p = 0.00*). This indicates that cyclists prefer routes which do not deviate substantially from their shortest alternative. The other results do indicate that cyclists are willing to detour to satisfy specific preferences. For example, cyclists seem to select routes with a low density of turns (*β = -0.39, p = 0.00*). That is, they seem to select simple routes over complex ones, possibly because the latter are harder to remember and may delay cyclists (Gliebe, 2012). Further, they avoid routes with a high number of intersections (*β = -0.10, p = 0.00*) and traffic lights (*β = -0.11, p = 0.01*). It must be noted that traffic lights may simply be positioned at busy intersections with a high risk of collision (Kang and Fricker, 2013), which could be the reason that these intersections are considered particularly unattractive. To add, cyclists seem to dislike routes with high speed limits (*β = -0.08, p = 0.00*). Interestingly, chosen routes seem to include a relatively high number of bridges (*β = 0.30, p = 0.00*). This could be related to the high number of bridges in the city center of Utrecht, which makes them hard to avoid. Moreover, these canal bridges are relatively flat and therefore do not require much effort to pass. To add, the strategic locations of these bridges may make them appealing connections between islands in the network, as is the case in Copenhagen (Prato, Halldórsdóttir and Nielsen, 2018). Finally, cyclists seem to be attracted to green surroundings (*β = 0.01, p = 0.07*), although this effect is less significant compared to the other attributes.

6.2.2 Main Effects Path Size Logit Model

Table 6.2 presents the results of the main effects Path Size Logit model. The significance of the Path Size correction factor highlights the importance of correcting for spatial overlap when dealing with revealed route choice data. The introduction of this parameter also severely improves model fit (*Δ McFadden Pseudo Rho square = 0.28, Δ AIC/N = -10.3*). Hence, the Path Size correction factor is included in all subsequent analyses. The McFadden Pseudo Rho square indicates that the model fit is beyond acceptable (Hensher, Rose and Greene, 2015). Further, all VIF scores are below the threshold 4, the highest being 3.9 (see Appendix IX).

Table 6.2 - Results Main Effects Path Size Logit Model			**McFadden Pseudo Rho square** = 0.52, **AIC/N** = 18.3				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Path Size Factor	**9.21**	**	0.13	69.35	0.00	8.95	9.47
Degree of Detour	**-0.11**	**	0.00	-46.35	0.00	-0.11	-0.10
Turn Density (/km)	**-0.49**	**	0.02	-32.54	0.00	-0.52	-0.46
Number of Intersections	**-0.13**	**	0.00	-30.60	0.00	-0.13	-0.12
Number of Traffic Lights	**-0.09**	**	0.01	-10.36	0.00	-0.11	-0.08
Speed Limit (Imputed)	**-0.10**	**	0.01	-18.85	0.00	-0.11	-0.09
Agriculture (50m Buffer)	-0.01		0.01	-1.43	0.15	-0.03	0.01
Number of Bridges	**0.31**	**	0.02	15.47	0.00	0.27	0.35

** significant at 5% level, ** significant at 1% level*

For almost all route attributes, the nature and significance of their effects correspond to the results of the main effects Multinomial Logit model. The change that stands out most concerns the effect of agricultural green, which has become insignificant. As discussed later, the preferences for green surroundings mixed, yielding an insignificant result for the general sample.

6.2.3 Path Size Logit Model with Interactions

The interaction model slightly outperforms the main effects Path Size Logit model (*Δ McFadden Pseudo Rho square = 0.01, Δ AIC/N = -0.5*). Again, all VIF scores are below the threshold 4, the highest being 3.9 (see Appendix IX). Further, it provides valuable insights into the role of trip purpose in route choice behavior. Specifically, the results in Table 6.3 indicate that preferences for agricultural green depend on whether someone is commuting or not. That is, cyclists appear to generally be attracted to farmland ($\beta = 0.06$, $p = 0.00$), but not when commuting ($\beta = -0.50$, $p = 0.00$). Most likely, commuters look for an efficient route and are not willing to trade speed, safety and comfort for pleasurable surroundings (Bernardi, Geurs and Puello, 2018). Therefore, they likely stick to urbanized areas, which offer a higher connectivity compared to agricultural landscapes. As discussed in the next paragraph, the interaction effect between traffic lights and peak hour departure time is only significant when the cyclists in the sample are segmentized.

Table 6.3 - Results Path Size Logit Model with Interactions			**McFadden Pseudo Rho square =** 0.53, **AIC/N =** 17.8				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Path Size Factor	**9.37**	**	0.13	69.45	0.00	9.11	9.64
Degree of Detour	**-0.10**	**	0.00	-45.20	0.00	-0.11	-0.10
Turn Density (/km)	**-0.45**	**	0.02	-29.49	0.00	-0.48	-0.42
Number of Intersections	**-0.14**	**	0.00	-31.32	0.00	-0.15	-0.13
Number of Traffic Lights	**-0.06**	**	0.01	-4.84	0.00	-0.08	-0.03
Number of Traffic Lights X Peak Hour	-0.03		0.02	-1.61	0.11	-0.07	0.01
Speed Limit (Imputed)	**-0.10**	**	0.01	-17.82	0.00	-0.11	-0.09
Agriculture (50m Buffer)	**0.06**	**	0.01	6.55	0.00	0.04	0.08
Agriculture X Commute	**-0.50**	**	0.03	-18.77	0.00	-0.55	-0.45
Agriculture X Leisure	-0.05		0.06	-0.82	0.41	-0.17	0.07
Number of Bridges	**0.30**	**	0.02	14.33	0.00	0.26	0.34

** significant at 5% level, ** significant at 1% level*

6.2.4 Latent Class Analysis

Table 6.4 presents the results of the Latent Class Analysis. The model specification corresponds to the Path Size Logit (PSL) model with interactions, including three personal characteristics. An attempt was made to estimate a three or four class model, without success. This could be expected due to the limited number of participants (*N=139*). The average class probabilities of class 1 and 2 are 0.75 and 0.25 respectively. The significant constant in the probability model (c =

3.73, p = 0.00) indicates that the preference heterogeneity cannot be fully explained by the personal characteristics in the model. This is not unexpected, given the limited demographic data that was available and the strong correlations across the motivators and deterrents of cycling (see §6.1.2). However, the model fit has increased substantially in comparison to the PSL model with interactions (*Δ McFadden Pseudo Rho square = 0.11, Δ AIC/N = -4.1*). This underscores the value of differentiating between the two identified classes of cyclists. Further, the results do provide some interesting insights into the compositions and preferences of these two classes.

Table 6.4 - Results Latent Class Model			**McFadden Pseudo Rho square** = 0.64, **AIC/N** = 13.7				
Variable	Coefficient	Sig.	Standard Error	z	Prob. $\lvert z\rvert>Z^*$	95% Confidence Interval	
Utility parameters in latent class 1		***Average class probability:*** 0.747					
Path Size Factor	**16.89**	**	0.37	45.79	0.00	16.17	17.62
Degree of Detour	**-0.28**	**	0.01	-42.13	0.00	-0.30	-0.27
Turn Density (/km)	**-0.93**	**	0.03	-30.88	0.00	-0.98	-0.87
Number of Intersections	**-0.12**	**	0.01	-22.32	0.00	-0.13	-0.11
Number of Traffic Lights	**-0.55**	**	0.02	-30.84	0.00	-0.59	-0.52
Number of Traffic Lights X Peak Hour	**0.20**	**	0.04	4.55	0.00	0.11	0.29
Speed Limit (Imputed)	**-0.58**	**	0.01	-56.06	0.00	-0.60	-0.56
Agriculture (50m Buffer)	**-0.16**	**	0.01	-14.22	0.00	-0.18	-0.14
Agriculture X Commute	**-1.11**	**	0.08	-13.89	0.00	-1.27	-0.95
Agriculture X Leisure	**0.27**	**	0.04	7.38	0.00	0.20	0.35
Number of Bridges	**0.71**	**	0.05	14.25	0.00	0.61	0.81
Utility parameters in latent class 2		***Average class probability:*** 0.253					
Path Size Factor	**6.87**	**	0.20	35.02	0.00	6.49	7.26
Degree of Detour	**-0.01**	**	0.00	-15.03	0.00	-0.01	-0.01
Turn Density (/km)	**-0.36**	**	0.03	-12.55	0.00	-0.41	-0.30
Number of Intersections	**-0.16**	**	0.01	-31.87	0.00	-0.17	-0.15
Number of Traffic Lights	0.01		0.03	0.45	0.65	-0.04	0.07
Number of Traffic Lights X Peak Hour	**-0.09**	*	0.05	-2.06	0.04	-0.18	0.00
Speed Limit (Imputed)	**-0.08**	**	0.01	-10.25	0.00	-0.09	-0.06
Agriculture (50m Buffer)	**0.16**	**	0.02	9.07	0.00	0.12	0.19
Agriculture X Commute	0.08		0.05	1.77	0.08	-0.01	0.17
Agriculture X Leisure	**0.53**	**	0.12	4.39	0.00	0.30	0.77
Number of Bridges	**-0.14**	**	0.03	-4.68	0.00	-0.20	-0.08
Probability model class 1	***(Parameters class 2 fixed to zero.)***						
Constant	**3.73**	**	0.66	5.67	0.00	2.44	5.02
Motivated by Enjoyment	**-0.34**	**	0.11	-3.21	0.00	-0.55	-0.13
Race Bike Ownership	**-1.44**	**	0.19	-7.68	0.00	-1.80	-1.07
Age 65+ (retired)	**-1.44**	**	0.37	-3.88	0.00	-2.16	-0.71

** significant at 5% level, ** significant at 1% level*

The results of the class membership model at the bottom of Table 6.4 give an impression of the compositions of classes 1 and 2. As indicated by the negative coefficients, cyclists who have not reached retirement (<65 years old), do not own a race bike and do not report to cycle for pleasure, are more likely to be a member of class 1.

Although some preferences appear to be similar in nature across the two classes, their magnitudes seem to differ. For example, cyclists in class 1 appear to be more sensitive to the degree of detour ($\beta = -0.28$, $p = 0.00$) compared to those in class 2 ($\beta = -0.01$, $p = 0.00$). In confirmation, the mean degree of detour among class 1 (M = 0.51%) is much lower compared to that of class 2 (M = 13%). Further, class 1 appears to be slightly less concerned with the number of intersections ($\beta = -0.12$, $p = 0.00$ *vs.* $\beta = -0.17$, $p = 0.00$) and more concerned with turn density ($\beta = -0.93$, $p = 0.00$ *vs.* $\beta = -0.34$, $p = 0.00$) and speed limits ($\beta = -0.58$, $p = 0.00$ *vs.* $\beta = -0.09$, $p = 0.00$).

Other route aspects show even stronger distinctions in preferences. First, the number of traffic lights does not appear to be relevant to cyclists in class 2 ($p = 0.65$), except during peak hour, in which case they tend to avoid them ($\beta = -0.09$, $p = 0.04$). In contrast, class 1 seems to select routes with fewer traffic signals in general ($\beta = -0.55$, $p = 0.00$), but appears to like them better during peak hours ($\beta = 0.20$, $p = 0.00$). These distinctive preferences explain why the 'Number of Traffic Lights X Peak Hour' interaction was insignificant in the earlier estimated PSL model with interactions. Further, class 2 seems to have a preference for agricultural landscapes ($\beta = 0.18$, $p = 0.00$), particularly when travelling to a leisure location ($\beta = 0.84$, $p = 0.00$). Class 1, on the other hand, appears to avoid farmland ($\beta = -0.16$, $p = 0.00$), especially during commute ($\beta = -1.11$, $p = 0.00$). However, class 1 does seem to choose routes with agricultural surroundings when on their way to a leisure location ($\beta = 0.27$, $p = 0.00$), albeit to a lesser extent compared to class 2 ($\beta = 0.84$, $p = 0.00$). Interestingly, cyclists in class 1 appear to be attracted to bridges ($\beta = 0.71$, $p = 0.00$), whereas those in class 2 seem to avoid them instead ($\beta = -0.13$, $p = 0.00$).

6.2.5 Class Comparison

As discussed in §6.1.2, there are strong correlations among the personal characteristics, in particular regarding the motivators and deterrents of cycling. Consequently, only a handful of them could be combined in the Latent Class Analysis (LCA). To gain a deeper understanding of the two groups, independent-samples T-tests (for continuous variables) and Chi-Square tests (for categorical variables) are performed. The participants are assigned to the class for which they have the highest posterior membership probability. Appendix X (see also the digital repository) provides the processed results of these tests.

Table 6.5 provides the mean statistics for those variables that show significant differences across the two identified segments of cyclists. These findings confirm the earlier conclusions regarding these two groups. That is, class 1 seems to consider the bike as a mode of transportation. This group is less motivated by physical or mental health or the mere enjoyment of cycling in itself. Moreover, they seem to enjoy being outside less and are more discouraged by distant destinations. They also report to be less motivated by traffic safety conditions compared to the other group. This could be an indication that these people experience the traffic safety of their surroundings to be suboptimal. Looking back at the results of the LCA, this experience may originate from the higher need for a safe infrastructure. That is, this group puts more emphasis on low speed limits and few crossings. Further, they have the tendency to seek the safety of signalised intersections during peak hours. When these aspects are lacking, cyclists belonging to class 1 will likely be less motivated to cycle. In sum, the above confirms the earlier reasoning that this class puts efficiency and safety above pleasurable surroundings.

In contrast, class 2 has a stronger intention to cycle and seems to do so out of pure enjoyment. They report being more motivated to cycle to benefit their physical and mental health and to like being outside more. Further, this group reports to be less discouraged by distant destinations, which corresponds to the earlier conclusion that this group is willing to tolerate a relatively high degree of detour. Moreover, race bikes are much more popular among these cyclists (41% versus 23%). Overall, it seems that this group consists of enthusiastic and more advanced riders. This may explain why these cyclists are comfortable cycling among motorized traffic at higher speeds (Sener, Eluru, Bhat, 2009). Further, their agility and confidence in traffic safety may also justify the higher tolerance for intersections and turns, as was observed in the LCA. Overall, these observations confirm the picture painted by the LCA.

Table 6.5 - Comparison Classes

		Class 1		**Class 2**	
		Mean	**Median**	**Mean**	**Median**
Motivators	*Physical Health (1-7)*	6.16	6.00	6.59	7.00
	Mental Health (1-7)	5.65	6.00	6.17	6.00
	Traffic Safety (1-7)	3.56	4.00	4.34	5.00
	Being Outside (1-7)	5.94	6.00	6.39	7.00
	Enjoyment (1-7)	5.52	6.00	6.17	6.00
Deterrents	*Distant Destination (1-7)*	5.22	5.50	4.71	5.00
Other	*Intention to bike (1-5)*	3.51	4.00	4.15	5.00
		Percentage		**Percentage**	
	Race Bike Ownership (0/1)	23%		41%	

6.3 Conclusion

The results reveal two distinguishable segments of cyclists. Table 6.6 summarizes the findings of the current study. The first group is characterized by their tendency to stick to the shortest route. They have relatively strong preferences when it comes to intersections, turns, speed limits and traffic lights. As such, it seems that this group is particularly concerned with convenience and safety. Indeed, they report that they are absolutely not motivated to cycle based on traffic safety. They have a relatively low intention to cycle and are less likely to report to cycle because they enjoy it. Hence, they appear to consider a bike to be a mode of transport. The second group is willing to detour substantially more in comparison. These cyclists are more keen on green surroundings, regardless of their trip purpose. They appear to be the more advanced cyclists who are more likely to own a race bike and have a relatively high intention to cycle. To add, they report to cycle because they like being outside, it increases their physical and mental health and they simply enjoy it. Therefore, these cyclists seem to consider the bike as more than a transport mode, they also cycle for pleasure.

Table 6.6 - Summary Latent Class Analysis		
Class 1: "Cycle for Transport"		***Class 2: "Cycle for Pleasure"***
Relatively low willingness to detour. Average detour: 0.51% Average route length: 2.7km		Relatively high willingness to detour. Average detour: 13% Average route length: 2.5km
Avoids crossings. ▲		Avoids crossings. ▼
Avoids traffic lights during off-peak hours, but finds them more appealing during on-peak hours. Passes 2.4 traffic lights on average.		Avoids traffic lights during on-peak hours. Passes 2.7 traffic lights on average.
Prefers low speed limits. ▲	30	Prefers low speed limits. ▼
Prefers straight roads. ▲		Prefers straight roads. ▼
Avoids farmland when cycling, particularly during commute. Is attracted to farmland when traveling to/from a leisure location.		Prefers cycling in agricultural surroundings, also during commute and particularly when going to/from a leisure location.
Less likely to own a race bike (23%).		More likely to own a race bike (41%).
Less likely to have reached retirement (4%).	***65+***	More likely to have reached retirement (10%).
Lower intention to bike (3.5 / 5).		Higher intention to bike (4.2 / 5).
Physical health ▼ Mental health ▼ Being outside ▼ Enjoyment ▼	***Motivators***	Physical health ▲ Mental health ▲ Being outside ▲ Enjoyment ▲ Traffic safety
Distant destination ▲ Traffic safety	***Deterrents***	Distant destination ▼

▼ = lower preference compared to other class, ▲ = higher preference compared to other class

7. Conclusions and Discussion

7.1 Summary of Findings

The results in Chapter 6 provide some interesting insights regarding the research questions defined in §1.3.1. First, the results of the Multinomial Logit (MNL) model (§6.3.2) indicate which built environmental and infrastructural characteristics influence route choices of cyclists in the municipality of Utrecht. As it turns out, aspects related to efficiency and safety are dominant. That is, all cyclists appear to be discouraged by roads with high speed limits and a large number of intersections and turns. Most likely, they perceive the risk of collision to be higher on those roads. To add, cyclists appear to be attracted to bridges, plausibly because these represent efficient connections between parts of the city. Further, in general, traffic lights seem to be avoided, possibly because they cause delays. Moreover, although cyclists seem attracted to agricultural green, they appear to avoid it during their commutes. Again, efficiency seems to play a role here. That is, cyclists seem to prefer strongly connected urbanized areas over losely connected farmland when commuting. Finally, the willingness to detour among Utrecht's cyclists is low, specifically, 5.5% on average. Overall, it seems that cyclists from Utrecht are mainly concerned with efficiency and safety.

The results also provide indications of preference heterogeneity among Utrecht's cyclists. In particular, two distinct segments are identified, as elaborated upon in §6.3.3 and §6.3.4. The first segment seems to consider the bike to be a mode of transport. They put more emphasis on efficiency, convenience and safety and are less willing to detour (0.51%). The second segment seems to cycle out of pure enjoyment. They have a higher intention to cycle, are more motivated to cycle to benefit their physical and mental health, like being outside more and are more likely to own a race bike. They also appear to be more attracted to agricultural landscapes, regardless of their trip purpose. Further, they are willing to detour considerably more (13%). Interestingly, opposing preferences for traffic lights can be observed for these two segments. That is, the first segment avoids them, but less so during peak hours, when signals may provide them with safe and efficient passage through heavy traffic. In contrast, the second segment avoids traffic lights during peak hours, possibly because they are willing to detour substantially to evade them. Table 6.6 (see §6.4) provides a complete summary of the differences between these segments.

Finally, there are also indications that context influences route choices. In particular, farmland appears to be more appealing to both segments when travelling to or from a leisure location. Possibly, cyclists are more concerned with efficiency during utilitarian trips, due to time constraints, and less so when cycling in their free time. Further, as discussed before, preferences for traffic lights turn out to differ across on and off-peak situations.

7.2 Theoretical Implications

The results of this study confirm several earlier findings discussed in the literature review (Chapter 2). In a broad sense, the results underscore the importance of traffic safety to cyclists, as discussed in §2.7. Further, several more specific findings are also replicated. For example, Kang and Fricker (2013) also conclude that intersections with traffic signals are generally less appealing to cyclists. They argue that these crossings might be more dangerous, which might be the underlying reason for this behavior. To add, the appeal of traffic lights during times of heavy traffic has also been observed by others (Park and Akar, 2019; Broach, Dill and Gliebe, 2012). As argued by Broach et al. (2012), the safety benefits of traffic lights seem to outweigh the delay they cause in those situations. Further, the aversion towards turns is broadly reported in the literature (Providelo and da Penha Sanches, 2011; Hood, Sall and Charlton, 2011; Zimmermann, Mai and Frejinger, 2017; Ghanayim and Bekhor, 2018; Broach, Dill and Gliebe, 2012; Prato, Halldórsdóttir and Nielsen, 2018; Skov-Petersen, Barkow, Lundhede and Jacobsen, 2018). Further, Skov-Petersen, Barkow, Lundhede, and Jacobsen (2018) also report a disutility for green areas among cyclists in Copenhagen. They argue that these areas are less safe and lack street lights. The current study indicates that this aversion is most evident among cyclists who put efficiency and safety first. To add, the results confirm that green surroundings are considered attractive in the context of leisure trips, as also concluded by Chen, Shen and Childress (2018). Further, the often reported preference for low speed limits is also evident in the current study (Anowar, Eluru and Hatzopoulou, 2017; Ghanayim and Bekhor, 2018; Melson, Duthie and Boyles, 2014; Parkin, Wardman and Page, 2008; Sener, Eluru and Bhat, 2009; Winters, Davidson, Kao and Teschke, 2011; Zimmermann, Mai and Frejinger, 2017). The insights discussed above are valuable, because only a handful of studies focus on the unique context of the Dutch cycling infrastructure. It is therefore interesting to see that some of the earlier observations from other countries apply to the cyclists in Utrecht as well.

However, some findings conflict with those of earlier studies. For example, Zimmermann, Mai and Frejinger (2017) report that cyclists avoid bridges. In contrast, cyclists in the current study generally seem to be attracted to routes with bridges. Prato, Halldórsdóttir and Nielsen (2018) make a similar observation for cyclists in Copenhagen. As they argue, bridges may provide efficient routes across town. It therefore seems important to consider urban layout when it comes to preferences for bridges. In some cases, bridges could potentially be avoided to some degree without the need to detour substantially. However, in other situations cyclists might be thankful for the quick shortcut they offer. Looking at the situation of the city center of Utrecht specifically, where a large portion of the routes is concentrated, bridges are often hard to avoid when selecting a short and efficient route. Moreover, the type of bridge may also be important to consider. For example, Copenhagen and Utrecht both know a lot of flat bridges with low traffic volumes (see examples in Appendix XI). In contrast, the bridges in Eugene, where Zimmermann et al. (2017) conducted their study, are generally high and part of major arteries. These may therefore be less appealing to cyclists, because of the effort to climb them and the exposure to motorized traffic. Another important contradiction between the findings of the current study and the reviewed literature is related to the preferences for intersections. Specifically, Lu, Scott and Dalumpines (2018) argue that cyclists are relatively tolerant of intersections, whereas the current study indicates the opposite. There are several plausible explanations for this difference. First, the sample of Lu et al. (2018) consists only of bikesharers, who may have distinct preferences from the general population of cyclists. Further, the traffic situations and infrastructural layout in Hamilton (Canada) might not be comparable to that of Utrecht. That is, the intersections passed by cyclists in Hamilton could be less dangerous or troublesome. To add, Lu et al. (2018) do not make a distinction between urbanized and agricultural areas. As observed in the current study, cyclists seem to prefer the strongly connected urbanized areas over losely connected agricultural ones. This preference for connectivity may have clouded the findings of Lu et al. (2018) regarding intersections. That is, when cyclists choose routes in well connected urbanized areas, they will simply be exposed to more intersections, but this does not mean they are attracted to them.

Further, several observations of the current study concern new contributions to the existing literature. For example, the results provide strong indications for preference heterogeneity regarding traffic lights. That is, it appears that some cyclists avoid them during peak hours, whereas others seem to be attracted to them at those times. This may explain why there is no clear consensus in the literature regarding this topic. Furthermore, recall that Prato, Halldórsdóttir and Nielsen (2018) observed a particularly strong aversion towards turns among a specific group of cyclists. Unfortunately, they did not discuss the personal characteristics of those cyclists. The current study made a similar observation for the first class in the Latent Class Analysis (LCA). The use of a LCA with a class membership model based on personal characteristics allows one to draw up a profile of both classes. Moreover, the supplementary analysis of the descriptives across both classes provides even more detail on who the members of these classes are. The current study is therefore able to conclude that those cyclists with a particularly strong aversion towards turns have a relatively low intention to cycle, are less likely to own a race bike, are more demotivated by distant destinations, and so on. In sum, the current study shows that it is important to consider trip context (e.g. departure time) and opposing preferences of different segments when studying route choice behavior.

7.3 Practical Implications

The findings of this study are valuable to the municipality of Utrecht, since they may guide future interventions to make the cycling infrastructure more attractive. Specifically, this study reveals two distinct segments of cyclists with unique preferences. First, the dominant group of cyclists in the sample find safety, convenience and efficiency most important. This group can be catered with reduced speed limits, and signalized intersections at busy crossings during peak hours. Possibly, the municipality could consider the installment of extra traffic lights which are only functional during peak hours along the routes which are popular among this group. Moreover, they have a preference for straight roads. Hence, new infrastructure targeted at these cyclists should be kept as straightforward and simple as possible. Further, it is important to realize that this group is very reluctant to detour. Thus, it will be very difficult to redirect these cyclists. Therefore, extremely attractive infrastructure is needed to guide them in a different direction. These cyclists are also more demotivated by distant destinations. Hence, a high facility density is needed to convince these people to get on their bike at all. Those in the second group have a higher intention to cycle and are willing to detour substantially. They are also more likely to own a race bike and to be seniors. These cyclists are less sensitive to aspects related to traffic safety, such as intersections, although these may still influence their route choices. In contrast, they put more emphasis on green surroundings. Hence, these cyclists would benefit from bike-friendly infrastructure in agricultural areas. Further, these cyclists appear to avoid traffic lights during peak hours, possibly because they delay them. Responsive traffic lights which prioritise cyclists over motorized traffic may make the inner city more appealing to this group during peak hours. In sum, the current study identifies two segments of cyclists with distinct preferences, which should be considered during future interventions by the municipality.

The findings of this study can also directly be translated to the cycling infrastructure of Utrecht. Specifically, the estimates from the Path Size Logit model with interactions and the Latent Class Analysis (LCA) can be used to score the links in Utrecht's network. The results of these calculations are translated to a dashboard, which can be accessed from the digital repository. The user is able to select a class from the LCA, a trip purpose, and a departure time (off or on peak). According to this input, an index score is calculated for each link in the network, following the procedures described in §4.5. The links are assigned a color according to their index score ranging from red (lowest) to green (highest). Thus, the user can intuitively identify segments which underperform. Moreover, it is possible to select one or more segments, which will trigger a refresh of the figures and charts. This allows the user to zoom in on a particular area of interest.

The dashboard makes the data generated in this study accessible to those unfamiliar with GIS software. This information is particularly valuable to policy makers. The dashboard allows them to intuitively compare the attractiveness of specific road segments and identify problematic situations. Moreover, the information in the tooltip and charts can help them derive why a segment is underperforming. Thus, the dashboard can support the municipality to develop successful interventions. For example, it may help them redirect cyclists away from overly crowded areas by improving certain aspects of underperforming alternative routes. Further, the dashboard could also be used by citizens or organizations such as De Fietsersbond, to give them more leverage when confronting the municipality with complaints or making suggestions to improve the cycling infrastructure. Appendix XII provides an overview of the most important functionality. A link to the dashboard is provided in the digital repository.

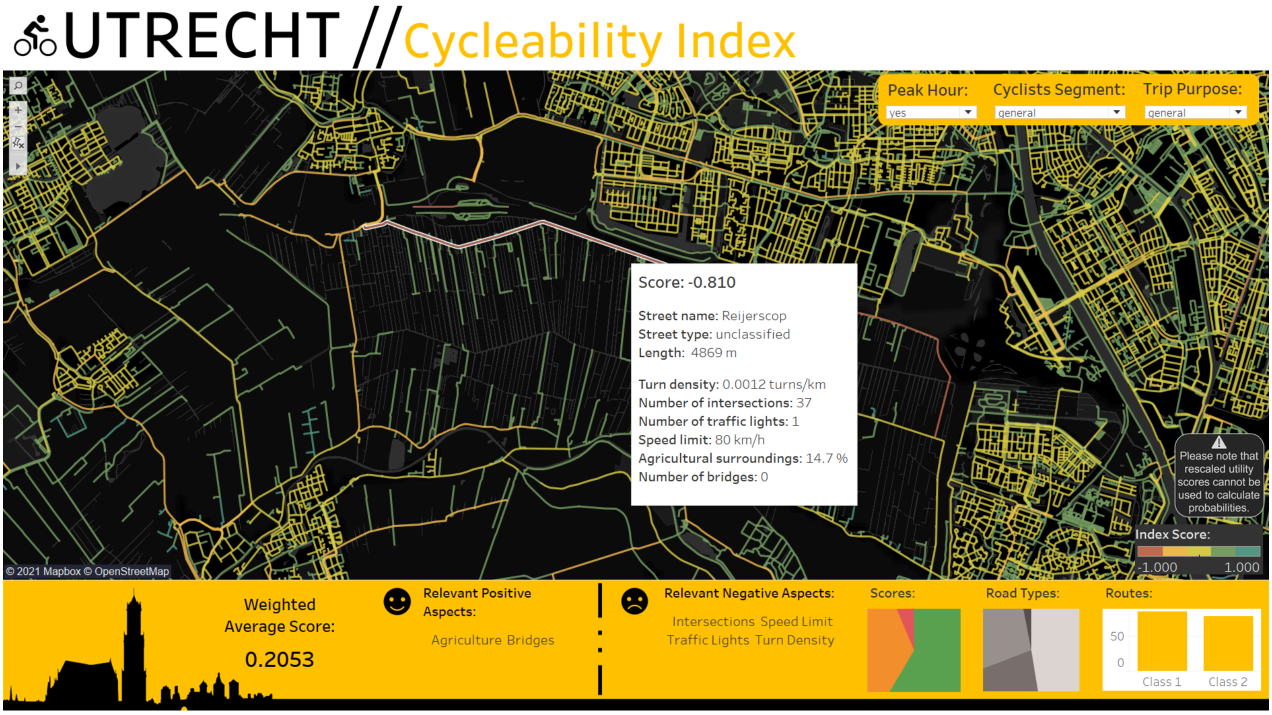


Figure 7.1 - Screenshot of Dashboard

7.4 Limitations and Future Research

Although the use of GPS data to study revealed preferences has proven to be a fruitful approach, it also imposes some limitations. First, strong correlations among route characteristics limit the number of attributes which can be combined in a route choice model. That is, some aspects simply do or do not coexist often in real life. In contrast, choice sets in stated preference experiments can be specifically designed to disentangle the effects of strongly related attributes.

This is possible because the alternatives are fictitious and their characteristics can be tweaked to the researcher's will. As the availability of GPS data increases, new methods to deal with correlated attributes in revealed route choice data could be very valuable. While these methods are still lacking, studies may benefit from a large research area to maximize the variation in choice situations. To add, it would be interesting to see more combinations of revealed and stated preference experiments. For example, Hensher (2008) pooled stated and revealed mode choice data to successfully "accommodate correlated observations" (p.23). In this regard, Ben-Akiva et al. (1994) discuss methods which combine revealed preference data with different types of stated preference data.

To add, revealed route choice studies rely on the available GIS data to determine the route characteristics. Some of this data is publically available at a worldwide scale. For example, the OpenStreetMap database documents speed limits and road types, along with many other aspects all over the world. However, it is important to acknowledge that this service runs on volunteers. The data is entered and reviewed by a community, but not by an official organization. It may therefore be error prone or incomplete. For example, in the current study the documentation of bike facilities, cycle lanes in particular, was clearly lacking. Specifically, only 230km of cycle lanes are documented in OpenStreetMap for the whole province, amounting to about 3% of the roads. Upon closer inspection, it turned out that many bike lanes were missing in the data. This made it difficult to evaluate preferences for bike facilities, as many other studies have done. This could be overcome by reaching out to local organizations such as De Fietsersbond to access professional databases, given that the budget is available. However, there is also a benefit to the use of public data. Since the current study relies only on public GIS data which is available at at least a national scale, the methodology could directly be applied to other Dutch cities without extra expenses.

Further, revealed route choice studies require the researcher to generate alternative routes. As discussed in §3.6, there are several methods to do so. The current study used the algorithm developed by De Fietsersbond to generate nine types of alternative routes. These nine distinct route types together compose a plausible set of alternatives that could be considered by cyclists. However, it must be noted that the generated alternatives concern recommended routes by De Fietsersbond, as such, they will not include extremely unattractive routes. In future research efforts, it could be valuable to include one or more seemingly unattractive routes to increase variation.

Further, there are several limitations of this study which can be related to the specific dataset which was used. For example, the data lacked confirmed trip purposes. Thus the potential purpose had to be derived based on the departure time and the type of origin and destination. This led to a relatively large number of trips which could not be classified. Therefore, it could be valuable to ask users of the ikFiets app or other apps alike to confirm their trip purpose afterwards in the future. Furthermore, weather conditions were not included as context variables in the analyses. This was a consequence of the decision to use regular trips, which are clusters of repeated trips made by one user. It was not possible to retrieve the weather conditions for these regular trips, because they are not assigned to a specific day. Future studies working with similar data may consider studying weather patterns of trip clusters, but this was outside the scope of the current study. Lastly, the data on the number of accidents turned out to be unsuitable for the analysis, because of the lack of detail. Unfortunately, the exact locations of the accidents were unknown. They could only be linked to complete roads. Moreover, only severe accidents which involve multiple parties, injuries and fatalities are included. An initiative among data specialists at the provincial level aims to generate a more complete view of the number of accidents, based on the administration of the local first aid departments. However, as for now, this data is only available for the main cycling infrastructure.

To add, the number of cyclists included in the analysis (N=139) is somewhat limited compared to other studies. A larger dataset might have allowed for more detailed segmentation in the Latent Class Analysis. Moreover, the cyclists participated in the survey at their own initiative. The sample may therefore be subject to a self selection bias. An observation that supports this proposition is the large proportion of highly educated participants compared to the national average (83.1% and 40% respectively). If highly educated cyclists have distinct preferences, this may have influenced the results of the study. Unfortunately, this is a common issue in these types of studies (see for example Anowar, Eluru and Hatzopoulou, 2017; Winters, Davidson, Kao and Teschke, 2011). Future studies could try to avoid this by specifically targeting unresponsive groups. Moreover, travel diary data from the Dutch Mobility Panel initiative could be replaced by GPS data in the near future, as suggested by Thomas, Geurs, Koolwaaij and Bijlsma (2015). It would be very interesting to see a similar methodology be applied to this data as it becomes available.

Finally, the current study has also sparked some suggestions for future research directions. For example, it remains unclear whether the layout, height and other characteristics of bridges influence route choice of cyclists. It would be valuable to see more studies such as Broach, Dill and Gliebe (2012), conducted in different countries, to better understand preferences for bridges. Moreover, it would be interesting to evaluate whether some relationships between route characteristics and route choice behavior could be nonlinear in nature. Further, the distinction between urbanized and agricultural areas in the current study suggest that connectivity plays a role in route choices of cyclists. A detailed study on connectivity could elaborate on this proposition. To add, it would be interesting to see a study which aims to capture route choice behavior of cyclists in an agent-based model. Such a model could be used to evaluate possible implications of interventions in the cycling infrastructure. The findings of the current study, and others alike, could be used to program the agents.

References

Abdullah, S., Markandya, A. And P.A.L.D. Nunes (2011) 'Introduction To Economic Valuation Methods' in Amit Batabyal & Peter Nijkamp (Eds.) Research Tools in Natural Resource and Environmental Economics, Chapter 5, pp. 143-187, World Scientific, US.

Anowar, S., Eluru, N., & Hatzopoulou, M. (2017). Quantifying the value of a clean ride: How far would you bicycle to avoid exposure to traffic-related air pollution? *Transportation Research Part A: Policy and Practice*, 105, 66–78. https://doi.org/10.1016/j.tra.2017.08.017

Arellana, J., Saltarín, M., Larrañaga, A. M., González, V. I., & Henao, C. A. (2020). Developing an urban bikeability index for different types of cyclists as a tool to prioritise bicycle infrastructure investments. *Transportation Research Part A: Policy and Practice*, 139, 310–334. https://doi.org/10.1016/j.tra.2020.07.010

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., & Rao, V. (1994). Combining revealed and stated preferences data. *Marketing Letters*, 5(4), 335–349. https://doi.org/10.1007/bf00999209

Bernardi, S., La Paix-Puello, L., & Geurs, K. (2018). Modelling route choice of Dutch cyclists using smartphone data. *Journal of Transport and Land Use*, 11(1), 883–900. https://doi.org/10.5198/jtlu.2018.1143

Bernardi, S., La Paix-Puello, L., & Geurs, K. (2018). Modelling route choice of Dutch cyclists using smartphone data. *Journal of Transport and Land Use*, 11(1). https://doi.org/10.5198/jtlu.2018.1143

Bike Citizens. (n.d.). Bike Citizens Homepage. Retrieved April 19, 2021, from https://www.bikecitizens.net/

Broach, J., Gliebe, J., & Dill, J. (2010). Calibrated labeling method for generating bicyclist route choice sets incorporating unbiased attribute variation. *Transportation Research Record: Journal of the Transportation Research Board*, 2197, 89–97.

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: Policy and Practice*, 46(10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Buehler, R., & Pucher, J. (2011). Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. *Transportation*, 39(2), 409–432. https://doi.org/10.1007/s11116-011-9355-8

Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. Wiley Interdisciplinary Reviews: Computational Statistics, 11(3), e1460. https://doi.org/10.1002/wics.1460

CBS. (2021a, May 18). Bevolking; onderwijsniveau; geslacht, leeftijd en migratieachtergrond [Dataset]. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82275NED/table?dl=535A6

CBS. (2021b, June 9). Gezondheid, leefstijl, zorggebruik en -aanbod, doodsoorzaken; kerncijfers [Dataset]. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81628NED/table?dl=55595

Chen, C. F., & Chen, P. C. (2013). Estimating recreational cyclists' preferences for bicycle routes – Evidence from Taiwan. *Transport Policy*, 26, 23–30. https://doi.org/10.1016/j.tranpol.2012.01.001

Chen, P., Shen, Q., & Childress, S. (2017). A GPS data-based analysis of built environment influences on bicyclist route preferences. *International Journal of Sustainable Transportation*, 12(3), 218–231. https://doi.org/10.1080/15568318.2017.1349222

De Vries, S. I., Hopman-Rock, M., Bakker, I., Hirasing, R. A., & Van Mechelen, W. (2010). Built Environmental Correlates of Walking and Cycling in Dutch Urban Children: Results from the SPACE Study. *International Journal of Environmental Research and Public Health*, 7(5), 2309–2324. https://doi.org/10.3390/ijerph7052309

Deenihan, G., & Caulfield, B. (2015). Do tourists value different levels of cycling infrastructure? *Tourism Management*, 46, 92–101. https://doi.org/10.1016/j.tourman.2014.06.012

Donnelly, J. E., Blair, S. N., Jakicic, J. M., Manore, M. M., Rankin, J. W., & Smith, B. K. (2009). Appropriate Physical Activity Intervention Strategies for Weight Loss and Prevention of Weight Regain for Adults. *Medicine & Science in Sports & Exercise*, 41(2), 459–471. https://doi.org/10.1249/mss.0b013e3181949333

Duc-Nghiem, N., Hoang-Tung, N., Kojima, A., & Kubota, H. (2018). Modeling cyclists' facility choice and its application in bike lane usage forecasting. *IATSS Research*, 42(2), 86–95. https://doi.org/10.1016/j.iatssr.2017.06.006

Fitch, D. T., & Handy, S. L. (2020). Road environments and bicyclist route choice: The cases of Davis and San Francisco, CA. *Journal of Transport Geography*, 85, 102705. https://doi.org/10.1016/j.jtrangeo.2020.102705

Gao, J., Kamphuis, C. B. M., Dijst, M., & Helbich, M. (2018). The role of the natural and built environment in cycling duration in the Netherlands. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1), 1–16. https://doi.org/10.1186/s12966-018-0715-z

Garrard, J., Rose, G., Lo, S.K., 2008. Promoting transportation cycling for women: the role of bicycle infrastructure. *Prev. Med*. 46, 55–59, http://dx.doi.org/10.1016/j.ypmed.2007.07.010

Gemeente Utrecht. (2015). Actieplan Utrecht fietst! [Press release]. https://www.utrecht.nl/wonen-en-leven/verkeer/fiets/actieplan-wereldfietsstad/

Geofabrik. (2021, March 23). OpenStreetMap data Utrecht [Dataset]. Geofabrik. http://download.geofabrik.de/europe/netherlands/utrecht.html

GGD & RIVM. (2014). Luchtkwaliteitsindex: Aanbevelingen voor de samenstelling en duiding (No. 2014–0050). https://www.rivm.nl/bibliotheek/rapporten/2014-0050.pdf

Ghanayim, M., & Bekhor, S. (2018). Modelling bicycle route choice using data from a GPS assisted household survey. *European Journal of Transport and Infrastructure Research*, 18(2), 158–177. https://doi.org/10.18757/ejtir.2018.18.2.3228

Glaser, M., & te Brömmelstroet, M. (2020). Unpacking policy transfer as a situated practice: blending social, spatial, and sensory learning at a conference. *Applied Mobilities*, 1–23. https://doi.org/10.1080/23800127.2020.1827559

Goedopweg. (n.d.). ikFiets Homepage. Ik fiets. Retrieved April 19, 2021, from https://www.ikfiets.nl/

GRASS Development Team. (n.d.). v.generalize. GRASS GIS Manual. Retrieved April 23, 2021, from https://grass.osgeo.org/grass76/manuals/v.generalize.html

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), 681–698. https://doi.org/10.1016/s0191-2615(02)00046-2

Greene, W. H., Hensher, D. A., Rose, J. M., 2006. Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. *Transportation Research Part B,* 40 (1), 75-92.

Handy, S., van Wee, B., & Kroesen, M. (2014). Promoting Cycling for Transport: Research Needs and Challenges. Transport Reviews, 34(1), 4–24. https://doi.org/10.1080/01441647.2013.860204

Hardinghaus, M., & Papantoniou, P. (2020). Evaluating Cyclists' Route Preferences with Respect to Infrastructure. Sustainability, 12(8), 3375. https://doi.org/10.3390/su12083375

Harkey, D.L., Reinfurt, D.W., Knuiman, M., Stewart, J.R., Sorton, A.: Development of the bicycle compatibility index: a level of service approach. Federal Highway Administration, Report FHWA-RD-98-072. Chapel Hill (1998)

Hensher, D. A. (1994). Stated preference analysis of travel choices: the state of practice. Transportation, 21(2), 107–133. https://doi.org/10.1007/bf01098788

Hensher, D. A. (2008). Empirical approaches to combining revealed and stated preference data: Some recent developments with reference to urban mode choice. *Research in Transportation Economics*, 23(1), 23–29. https://doi.org/10.1016/j.retrec.2008.10.003

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). Applied choice analysis. Cambridge, United Kingdom: Cambridge University Press. https://doi.org/10.1007/9781316136232

Hess, S., Ben-Akiva, M., Gopinath, D., Walker, J.: Advantages of latent class models over continuous mixture models in capturing heterogeneity. In: Proceedings of European Transport Conference (2008)

Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, 3(1), 63–75. https://doi.org/10.3328/tl.2011.03.01.63-75

Hopkinson, P., & Wardman, M. (1996). Evaluating the demand for new cycle facilities. *Transport Policy*, 3(4), 241–249. https://doi.org/10.1016/s0967-070x(96)00020-0

Hunt, J. D., & Abraham, J. E. (2006). Influences on bicycle use. *Transportation*, 34(4), 453–470. https://doi.org/10.1007/s11116-006-9109-1

Illowsk, B., Dean, S., Holmes, A. (2017). Introductory Business Statistics (1st ed.). OpenStax.

Kang, L., & Fricker, J. D. (2013). Bicyclist commuters' choice of on-street versus off-street route segments. *Transportation*, 40(5), 887–902. https://doi.org/10.1007/s11116-013-9453-x

Kang, L., & Fricker, J. D. (2016). Sharing urban sidewalks with bicyclists? An exploratory analysis of pedestrian perceptions and attitudes. *Transport Policy*, 49, 216–225. https://doi.org/10.1016/j.tranpol.2016.04.013

Kang, L., & Fricker, J. D. (2018). Bicycle-Route Choice Model Incorporating Distance and Perceived Risk. *Journal of Urban Planning and Development*, 144(4), 04018041. https://doi.org/10.1061/(asce)up.1943-5444.0000485

Kirner Providelo, J., & da Penha Sanches, S. (2011). Roadway and traffic characteristics for bicycling. *Transportation*, 38(5), 765–777. https://doi.org/10.1007/s11116-011-9353-x

Kreyszig, E. (2005). Advanced Engineering Mathematics. Wiley.

Krizek, K. J., El-Geneidy, A., & Thompson, K. (2007). A detailed analysis of how an urban trail system affects cyclists' travel. *Transportation*, 34(5), 611–624. https://doi.org/10.1007/s11116-007-9130-z

Liu, Y., Yang, D., Timmermans, H. J., & de Vries, B. (2020). Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: A stated choice experiment. *Journal of Transport Geography*, 82, 102570. https://doi.org/10.1016/j.jtrangeo.2019.102570

Lu, W., Scott, D. M., & Dalumpines, R. (2018). Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths. *Journal of Transport Geography*, 71, 172–181. https://doi.org/10.1016/j.jtrangeo.2018.07.012

Majumdar, B. B., & Mitra, S. (2017). Valuing Factors Influencing Bicycle Route Choice Using a Stated-Preference Survey. *Journal of Urban Planning and Development*, 143(3), 04017001. https://doi.org/10.1061/(asce)up.1943-5444.0000380

Manaugh, K., Boisjoly, G., & El-Geneidy, A. (2016). Overcoming barriers to cycling: understanding frequency of cycling in a University setting and the factors preventing commuters from cycling on a regular basis. *Transportation*, 44(4), 871–884. https://doi.org/10.1007/s11116-016-9682-x

Melson, C. L., Duthie, J. C., & Boyles, S. D. (2014). Influence of bridge facility attributes on bicycle travel behavior. *Transportation Letters*, 6(1), 46–54. https://doi.org/10.1179/1942787513y.0000000010

Melson, C. L., Duthie, J. C., & Boyles, S. D. (2014). Influence of bridge facility attributes on bicycle travel behavior. *Transportation Letters*, 6(1), 46–54. https://doi.org/10.1179/1942787513y.0000000010

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. (2010). Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice*, 44(9), 754–765. https://doi.org/10.1016/j.tra.2010.07.008

Mertens, L., Compernolle, S., Deforche, B., Mackenbach, J. D., Lakerveld, J., Brug, J., Roda, C., Feuillet, T., Oppert, J. M., Glonti, K., Rutter, H., Bardos, H., De Bourdeaudhuij, I., & Van Dyck, D. (2017). Built environmental correlates of cycling for transport across Europe. *Health & Place*, 44, 35–42. https://doi.org/10.1016/j.healthplace.2017.01.007

Miles, J., & Shevlin, M. (2001). Applying Regression and Correlation: A Guide for Students and Researchers (First ed.). SAGE Publications Ltd.

Move Smarter: Results from the Dutch Mobile Mobility Panel. Transportation Research Procedia, 11, 247–262.

Murphy, J. J., Allen, P. G., Stevens, T. H. and Weatherhead, D. 2005. A meta-analysis of hypothetical bias in stated preference valuation, *Environmental and Resource Economics*, 30, 313–325.

Noordzij, J. M., Beenackers, M. A., Groeniger, J. O., Timmermans, E. J., Motoc, I., Huisman, M., & van Lenthe, F. J. (2021). Land use mix and physical activity in middle-aged and older adults: a longitudinal study examining changes in land use mix in two Dutch cohorts. International *Journal of Behavioral Nutrition and Physical Activity*, 18(1), 1–10. https://doi.org/10.1186/s12966-021-01083-1

Okoli, C., Schabram, K. (2010). "A Guide to Conducting a Systematic Literature Review of Formation Systems Research," . Sprouts: Working Papers on Formation Systems, 10(26). http://sprouts.aisnet.org/10-26

Orellana, D., & Guerrero, M. L. (2019). Exploring the influence of road network structure on the spatial behaviour of cyclists using crowdsourced data. *Environment and Planning B: Urban Analytics and City Science*, 46(7), 1314–1330. https://doi.org/10.1177/2399808319863810

Park, Y., & Akar, G. (2019). Why do bicyclists take detours? A multilevel regression model using smartphone GPS data. *Journal of Transport Geography*, 74, 191–200. https://doi.org/10.1016/j.jtrangeo.2018.11.013

Parkin, J., Wardman, M., & Page, M. (2007). Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation*, 35(1), 93–109. https://doi.org/10.1007/s11116-007-9137-5

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. *Journal of Choice Modelling*, 2(1), 65–100. https://doi.org/10.1016/s1755-5345(13)70005-8

Prato, C. G., Halldórsdóttir, K., & Nielsen, O. A. (2018). Evaluation of land-use and transport network effects on cyclists' route choices in the Copenhagen Region in value-of-distance space. *International Journal of Sustainable Transportation*, 12(10), 770–781. https://doi.org/10.1080/15568318.2018.1437236

Pritchard, R., Bucher, D., & Frøyen, Y. (2019). Does new bicycle infrastructure result in new or rerouted bicyclists? A longitudinal GPS study in Oslo. *Journal of Transport Geography*, 77, 113–125. https://doi.org/10.1016/j.jtrangeo.2019.05.005

Pritchard, R., Bucher, D., & Frøyen, Y. (2019). Does new bicycle infrastructure result in new or rerouted bicyclists? A longitudinal GPS study in Oslo. *Journal of Transport Geography*, 77, 113–125. https://doi.org/10.1016/j.jtrangeo.2019.05.005

Pucher, J., & Buehler, R. (2008). Making Cycling Irresistible: Lessons from The Netherlands, Denmark and Germany. *Transport Reviews*, 28(4), 495–528. https://doi.org/10.1080/01441640701806612

Rijkswaterstaat. (2020). Bestand GeRegistreerde Ongevallen Nederland 2019 [Dataset]. https://www.rijkswaterstaat.nl/apps/geoservices/geodata/dmc/bron/01-01-2019_31-12-2019.zip

RIVM. (2021, March 19). GCN concentratiekaartbestanden achterliggende jaren [Dataset]. https://www.rivm.nl/gcn-gdn-kaarten/concentratiekaarten/cijfers-achter-concentratiekaarten/gcn-concentratiekaartbestanden-achterliggende-jaren

Rossetti, T., Saud, V., & Hurtubia, R. (2017). I want to ride it where I like: measuring design preferences in cycling infrastructure. *Transportation*, 46(3), 697–718. https://doi.org/10.1007/s11116-017-9830-y

Sarjala, S. (2019). Built environment determinants of pedestrians' and bicyclists' route choices on commute trips: Applying a new grid-based method for measuring the built environment along the route. *Journal of Transport Geography*, 78, 56–69. https://doi.org/10.1016/j.jtrangeo.2019.05.004

Schreckenberg, M., & Selten, R. (2013). Human Behaviour and Traffic Networks. Springer Publishing.

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, 36(5), 511–539. https://doi.org/10.1007/s11116-009-9201-4

Skov-Petersen, H., Barkow, B., Lundhede, T., & Jacobsen, J. B. (2018). How do cyclists make their way? - A GPS-based revealed preference study in Copenhagen. *International Journal of Geographical Information Science*, 32(7), 1469–1484. https://doi.org/10.1080/13658816.2018.1436713

Sobhani, A., Aliabadi, H. A., & Farooq, B. (2019). Metropolis-Hasting based Expanded Path Size Logit model for cyclists' route choice using GPS data. *International Journal of Transportation Science and Technology*, 8(2), 161–175. https://doi.org/10.1016/j.ijtst.2018.11.002

Standen, C., Crane, M., Collins, A., Greaves, S., & Rissel, C. (2017). Determinants of mode and route change following the opening of a new cycleway in Sydney, Australia. *Journal of Transport & Health*, 4, 255–266. https://doi.org/10.1016/j.jth.2016.10.004

Statista. (2020). Bicycle industry in the Netherlands (did-52703-1). https://www.statista.com/study/52703/bicycle-industry-in-the-netherlands/

Statista. (2021, March 4). Cycling - Statistics & Facts. https://www.statista.com/topics/1686/cycling/

Strava. (n.d.). Strava Homepage. Retrieved April 19, 2021, from https://www.strava.com/

Tableau Software. (n.d.). Tableau: Business Intelligence and Analytics Software. Tableau. Retrieved July 8, 2021, from https://www.tableau.com/

Thomas, T., Geurs, K. T., Koolwaaij, J., & Bijlsma, M. (2018). Automatic Trip Detection with the Dutch Mobile Mobility Panel: Towards Reliable Multiple-Week Trip Registration for Large Samples. *Journal of Urban Technology*, 25(2), 143–161. https://doi.org/10.1080/10630732.2018.1471874

Train, K. (2009). Discrete Choice Methods with Simulation. Cambridge University Press.

Vedel, S. E., Jacobsen, J. B., & Skov-Petersen, H. (2017). Bicyclists' preferences for route characteristics and crowding in Copenhagen – A choice experiment study of commuters. *Transportation Research Part A: Policy and Practice*, 100, 53–64. https://doi.org/10.1016/j.tra.2017.04.006

Wardman, M. (1988). A Comparison of Revealed Preference and Stated Preference Models of Travel Behaviour. *Journal of Transport Economics and Policy*, 22(1), 71–91. https://www.jstor.org/stable/20052836

Wen, C.H., Lai, S.C.: Latent class models of international air carrier choice. *Transp. Res. Part E Logist. Transp. Rev.* 46(2), 211–221 (2010)

Winters, M., & Teschke, K. (2010). Route Preferences among Adults in the Near Market for Bicycling: Findings of the Cycling in Cities Study. *American Journal of Health Promotion*, 25(1), 40–47. https://doi.org/10.4278/ajhp.081006-quan-236

Winters, M., Davidson, G., Kao, D., & Teschke, K. (2010). Motivators and deterrents of bicycling: comparing influences on decisions to ride. *Transportation*, 38(1), 153–168. https://doi.org/10.1007/s11116-010-9284-y

Zimmermann, M., Mai, T., & Frejinger, E. (2017). Bike route choice modeling using GPS data without choice sets of paths. *Transportation Research Part C: Emerging Technologies*, 75, 183–196. https://doi.org/10.1016/j.trc.2016.12.009

Appendix I - Literature Review Methodology

The literature study is executed following the guidelines by Okoli and Schabram (2010) on conducting a Systematic Literature Review. Accordingly, the process is initiated by explicitly specifying the purpose of the review. Given the notion that this review employed only the main author, no formal protocol is developed other than specifying the search and eligibility criteria and training of additional reviewers is unnecessary. However, the outline of the process is recorded and reviewed by a full professor and supervisor of the project. Next, the scope of the search is defined. This includes the databases which are to be queried as well as the search terms that are to be used. Comprehensiveness is ensured through combining both specific (e.g. cycleability) and broad (e.g. cycling) search terms employing Boolean operators. Furthermore, an online thesaurus is used to check for possible synonyms (e.g. cyclists, bicyclers, bicyclist, etc.). Once the documents are collected, a practical screen takes place to develop a reasonably comprehensive final list of publications, whilst accounting for the limitations of the reviewer. Documents that meet the eligibility criteria move on to the extraction phase. In this phase, the data on specific topics is extracted and cross-referenced. Predefined tabular formats assure that every study is reviewed thoroughly and for the same elements. Finally, the findings are summarized and contrasted to provide an overview of the literature on indicators of cycleability. To add, a list of commonly studied indicators and their reported effects is developed. Further, the variations among different types of cyclists (e.g. commuters vs. sportive cyclists) are summarized.

I.1 Purpose of the Review

This literature review is conducted as part of a larger study on cycleability. The goal of the review is twofold. First, it serves to generate a list of commonly studied indicators of cycleability and to summarize the reported nature and magnitude of their effects. This overview is a starting point for discussions with field experts on the relative importance and comprehensiveness of these indicators. Furthermore, it serves as a guide for selecting the indicators to be considered in a revealed preference experiment. Second, the review should provide insight into the reported variations among different types of cyclists when it comes to the effects of the indicators. An understanding of these differences helps determine which personal characteristics should be considered during the upcoming experiment. Moreover, the results of the revealed preference study can then be contrasted against earlier findings from the general literature. This may reveal important attitudinal differences between the Dutch cycling culture, which is the focus of the experiment, and the North-American culture, which is more commonly addressed in academic studies. All in all, the literature review should provide a strong foundation for the upcoming experiment.

I.2 Research Protocol

As discussed in detail later, a set of databases and qualified journals is selected to conduct the search. Further, a query is drawn up that fits the defined search criteria (see below). The resulting list of content is screened during a “practical screen” for the applicability to the review at hand. A set of eligibility criteria is defined to structure this process. These criteria can be found below. The search results, their inclusion verdict, and the extracted information are recorded digitally in tabular form. A predefined tabular format assures that every study is reviewed thoroughly and for the same elements. The results of the literature search are presented in this original format in the digital repository (‘literatureSearch.xlsx’).

Search Criteria:	*Eligibility Criteria:*
● No grey literature. ● Dates from 1990 onwards. ● Only published or in-press. No grey literature. ● Only English or Dutch. ● SCimago Journal Rank Indicator Q1 or Q2 in the “Transport” category.	An article should... ● evaluate relationships between route characteristics and cycling behavior. ● cover infrastructural, built environment and/or natural characteristics. ● contain predominantly objective measures. That is, it should not focus on attitudinal or perceptual aspects.

I.3 Literature Search

As discussed earlier, several search criteria have been drawn up to limit the scope of the search. The two most important ones, publication-quality and timeframe, are elaborated upon below. In addition, only English and Dutch language articles are considered, given the language capabilities of the reviewer. This should cover the general body of studies originating from Western societies. Further, the articles should be published or in press and therefore available in the databases accessible to the reviewer. Although Okoli and Schabram (2010) consider some of these issues part of the practical screen, these criteria can already be applied to the search query. Doing so reduces the burden of the practical screen.

I.3.1 Publication-Quality

The ScImago journal ranking is used to generate a list of journals to be queried for publications. This ranking is based on the SJR2 index, which was created to measure the “scientific prestige” of journals, based on the weighted number of citations (Guerrero-Bote & Moya-Anegón, 2012). The developers take into account varying citation customs, which makes the index comparable across research fields. Moreover, citations are weighted for the thematic relatedness of the two journals at hand. This adds more nuance to the indicator, as opposed to simpler measures such as the impact factor. The current ScImago journal ranking can be found on the ScimagoJR website (SCImago, n.d.). When filtering for a particular subject, the ranking table will also show whether a journal falls in the best quantile (Q1), second-best (Q2), and so on. The current literature review considers the ScImago journal ranking of 2019 given that this provides the most recent complete overview. Only journals from the first and second-best quantiles in the subject category "transportation" are considered. Further, some journals were excluded a posteriori upon a review of their descriptions, to limit the number of results. These concerned journals with a sole focus on public or maritime transport or logistics. The resulting journal list is provided in Table I.1.

Table I.1 - ScImago's (n.d.) Journal Ranking of 2019 in "Transportation" Category

Rank	Title	Quantile	Selected
1	Analytic Methods in Accident Research	Q1	✓
2	Transportation Research, Part C: Emerging Technologies	Q1	
3	Transport Reviews	Q1	✓
4	Tourism Management	Q1	✓
5	Journal of Travel Research	Q1	✓
6	Transportation Research Part B: Methodological	Q1	✓
7	Transportation Science	Q1	✓
8	International Journal of Physical Distribution and Logistics Management	Q1	
9	Transportation Research Part E: Logistics and Transportation Review	Q1	✓
10	Transportation Research Part A: Policy and Practice	Q1	✓
11	EURO Journal on Transportation and Logistics	Q1	
12	IEEE Transactions on Transportation Electrification	Q1	
13	Transportation	Q1	✓
14	Journal of Transport Geography	Q1	✓
15	Transportation Research Part D: Transport and Environment	Q1	✓
16	Transport Policy	Q1	✓
17	International Journal of Sustainable Transportation	Q1	✓
18	Travel Behaviour and Society	Q1	✓
19	Sustainable Cities and Society	Q1	✓
20	Maritime Policy and Management	Q1	
21	International Journal of Transportation Science and Technology	Q1	✓
22	Journal of Air Transport Management	Q1	
23	Economics of Transportation	Q1	✓

24	Transportation Research Part F: Traffic Psychology and Behaviour	Q1	✓
25	Transportation Geotechnics	Q1	✓
26	International Journal of Logistics Management	Q1	
27	Mobilization	Q1	✓
28	Transportmetrica B	Q2	✓
29	International Journal of Tourism Research	Q2	✓
30	Journal of Transport and Health	Q2	✓
31	Transportmetrica A: Transport Science	Q2	✓
32	Research in Transportation Business and Management	Q2	
33	IATSS Research	Q2	✓
34	Research in Transportation Economics	Q2	✓
35	Journal of Transport and Land Use	Q2	✓
36	Maritime Economics and Logistics	Q2	
37	Transportation Journal	Q2	✓
38	European Journal of Transport and Infrastructure Research	Q2	✓
39	International Journal of Rail Transportation	Q2	
40	Transportation Letters	Q2	✓
41	Journal of Public Transportation	Q2	
42	Urban Rail Transit	Q2	
43	European Transport Research Review	Q2	✓
44	Journal of Transport Economics and Policy	Q2	✓
45	Public Transport	Q2	
46	IET Intelligent Transport Systems	Q2	
47	Journal of Traffic and Transportation Engineering (English Edition)	Q2	✓
48	Case Studies on Transport Policy	Q2	✓
49	Archives of Transport	Q2	✓
50	Asian Journal of Shipping and Logistics	Q2	
51	Transportation Planning and Technology	Q2	✓
52	Journal of Transportation Safety and Security	Q2	✓
53	International Journal of Shipping and Transport Logistics	Q2	
54	Journal of Transportation Engineering	Q2	✓

I.3.2 Timeframe

Under the influence of cultural, economical, and societal changes, it can be expected that the behavior and preferences of cyclists have changed over time. It is therefore important to define a timeframe for the literature search that limits the scope to reasonably recent publications. This also makes the finding of the selected publications more comparable. An interesting development to consider is the growing accessibility to cars in Western countries. High levels of car ownership may mean that people are more selective in the use of a bike for particular trips, given they more

often have the car as an alternative. That is, car ownership influences mode choice behavior. Furthermore, the infrastructure has been adapted to accommodate the increasing stream of motorized traffic. This has changed the look and feel of the streets to cyclists. To add, an increase in traffic volume may also influence the cycling experience and could lead to more accidents. Hence, car ownership is argued to substantially influence multiple facets of cycling behavior.
In the Netherlands, one can observe a rapid growth in the number of cars per capita between 1970 and 1980 (see Figure I.1). This growth still continues today, but is slowly stagnating towards one car per two inhabitants. Taking this into consideration, it can be concluded that between the years 2000 and 2020, the level of accessibility to a car has been roughly the same. Therefore, the search is constrained to publications between 2000 and 2020.

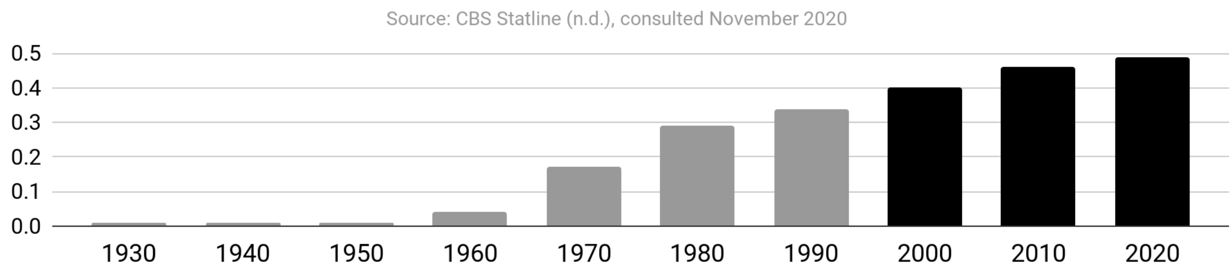


Figure I.1 - Cars Per Capita in The Netherlands

I.6 Practical Screen

The practical screen step is described by Okoli and Schabram (2010) as strongly subjective. Nevertheless, the choices of the reviewer must result in a reasonably comprehensive final list of publications, whilst accounting for the limitations of the reviewer. The guidelines for the practical screen are discussed below. These complement the search criteria discussed in the previous section.

I.6.1 Fit with Review Goals

The main inclusion criterion concerns the fit with the review goals. That is, a selected study should discuss infrastructural, built environment, and natural factors which influence cycling behavior. Preferably, a study should be related to preferences of cyclists or route choice behavior. However, studies on cycling habits (e.g. frequency), trends (e.g. counts), experiences or mode choice may also indirectly reveal preferences. For example, Snizek, Sick Nielsen and Skov-Petersen (2013) relate positive and negative experiences of cyclists to aspects of the cycling infrastructure. The aggregate of these perceptions captures the general attitude of cyclists towards particular aspects of the infrastructure. This attitude ultimately influences a cyclist's evaluation of a route and therefore the perceived cycleability. Hence, the literature review should not be limited to studies that explicitly discuss indicators of cycleability.

I.6.2 Research Design

Cycleability can be measured in a variety of ways. For example, subjects can be asked explicitly what boosts or hurts the cycleability of the infrastructure. Alternatively, discrete choice experiments can be conducted where participants are presented with sets of hypothetical situations and asked to choose, rate, or rank them. Their choices can then be studied to reveal their preferences for particular aspects. In contrast, revealed choice experiments derive preferences based on real-life actions. In the case of cycleability, these actions can be recorded as GPS data of routes chosen by cyclists. Comparing the characteristics of the selected route and the possible alternatives may again reveal the hidden preferences of the subject. Considering the above, the selection is not restricted to a particular research design.

I.6.3 Study Setting

As discussed, the results of this literature review will serve as a foundation for an empirical study on cycleability. This study will focus on the Dutch municipality "Utrecht". The Netherlands provides a unique case, given its historical cycling culture. Considering its densely populated areas and flat topography, cycling often poses a suitable alternative to car travel. It is important to realize that these characteristics may complicate the translation of research findings from other countries to the Dutch situation. It is therefore tempting to focus the review on the unique Dutch context only. However, it is expected that this will limit the number of selected studies substantially. Consequently, the review might miss out on some important indicators which have not been studied in the Netherlands. Therefore, studies conducted in other European and North-American countries are also included in this review. However, their results are explicitly contracted against the Dutch findings.

Appendix II - Coding Location Types

Retail / Eating Out

alcohol_shop, bakery_shop, bar, beauty_shop, bicycle_shop, cafe, charity_shop, chemist_shop, chocolate_shop, clothes_shop, coffee_and_tea_shop, commercial, confectionery_shop, convenience_shop, deli_shop, doityourself, electronics_shop, farm_shop, florist_shop, furniture_shop, garden_centre_shop, greengrocer_shop, hairdresser_shop, hearing_aids_shop, hifi_shop, houseware_shop, ice_cream, interior_design_office, internet_cafe, jewelry_shop, kitchen_utensils_shop, lighting_shop, locksmith_shop, mall, optician_shop, pastry_shop, perfumery_shop, pub, restaurant, retail, second_hand_shop, shop, soft_drugs_shop, sports_shop, supermarket, toys_shop, variety_store_shop, interior_decoration_shop, art_shop, books_shop, butcher_shop, shoes_shop, tattoo_shop, tobacco_shop

Work / School / Daycare

architect_office, educational_institution_office, events_office, industrial, kindergarten, lawyer_office, music_composer_office, newspaper_office, ngo_office, office, research_institute, school, tailor_school, university, interior_design_office, childcare, conference_centre, coworking_office

Services / Healthcare

bank, car_repair_shop, charging_station, dentist, doctors, estate_agent_shop, fuel, government_office, hospital, information, pharmacy, post_office, public_building, service, social_facility, townhall, veterinary, car_rental, police

Leisure Locations

arts_centre, artwork, athletics_pitch, attraction, boat_rental, camp_site, caravan_site, cinema, climbing_sports_centre, community_centre, cycling_sports_centre, equestrian_sports_centre, field_hockey_sports_centre, fitness_centre, frisbee_pitch, gambling, golf_course, gymnastics_sports_centre, hockey_sports_centre, ice_rink, library, museum, playground, recreation_ground, sauna, skiing_sports_centre, soccer_pitch, pitch, soccer_sports_centre, sports_centre, squash_sports_centre, stadium, swimming_pool, swimming_sports_centre, tennis_sports_centre, theatre, water_park, zoo, fort

Nature

forest, grass, nature_reserve, park, stream, water

Home / Visits

apartments, home, house, houseboat, neighbourhood, nursing_home, residence, residential, beach

Places to Stay

hostel, hotel, caravan_site

Other

area, bicycle_parking, bridge, bus, canal, cemetery, childcare, church, fence, fire_station, guest_house, greenfield, lock, mosque, parking, pedestrian area, picnic_table, place, place_of_worship, platform, rail, recycling, river, roof, shed, static_caravan, station, tower, tram, vending_machine, water_well, grave_yard, building, common, ferry_terminal, viewpoint

Appendix III - Coding Educational Level

Low: Primary School, LBO, LEAO, LHNO, LTS, MAVO, VMBO, MULO, MBO

Medium: HAVO, HBS, VWO, HBO, HTS, HEAO

High: University (BSc / MSc)

SPSS Syntax:

```
RECODE opleiding (1=1) (2=1) (3=1) (4=2) (5=3) (7=SYSMIS) (6=SYSMIS) INTO
Edu_cat.
```

Appendix IV - NLogit Syntax

Main Effects MNL Model:

```
Reset$
Read; File=path\to\data.csv $

create; Nalt=10 $

Nlogit
; lhs = chosen,Nalt
; rhs = detour, t_X46, t_X57, t_X3, t_X42, t_X27, t_X12
; pds = Nsets
; Parameters (Save posterior results)
; WTS = Ntrips
$
```

Main Effects PSL Model:

```
Reset$
Read; File=path\to\data.csv $

create; Nalt=10 $

Nlogit
; lhs = chosen,Nalt
; rhs = PSin, detour, t_X46, t_X57, t_X3, t_X42, t_X27, t_X12
; pds = Nsets
; Parameters (Save posterior results)
; WTS = Ntrips
$
```

PSL Model With Interaction Effects & LCA:

```
Reset$
Read; File=path\to\data.csv $

create; Nalt=10 $

Nlogit
; lhs = chosen,Nalt
; rhs = PSin, detour, t_X46, t_X57, t_X3, t3c15, t_X42, t_X27, t27c10, t27c9, t_X12
; lcm = r8, f1, u12
; Pts=2
; pds = Nsets
; Parameters (Save posterior results)
; WTS = Ntrips
$
```

Appendix V - Correlations Among Route Characteristics

	Route Total Length	Degree of Detour	Traffic Lights Count	Traffic Lights Density (/km)	Shops Count	Shops Denisty (/km)	Accidents Count	Accidents Denisty	Bridges Count	Bridges Denisty	Cycleway Total Length	Cycleway Proportion	Cycle Lane Total Length	Cycle Lane Proportion	Bike Facility Proportion	Landuse 25m buffer - Archiculture	Landuse 25m buffer - Green	Landuse 25m buffer - Retail	Landuse 25m buffer - Commercial	Landuse 25m buffer - Industrial
Route Total Length	1.00	0.30	0.48	-0.01	0.13	-0.12	0.41	-0.11	0.73	0.04	0.84	0.15	0.35	-0.02	0.15	0.26	0.31	-0.18	0.08	0.06
Degree of Detour	0.30	1.00	0.06	-0.06	0.09	-0.02	0.13	-0.03	0.22	0.01	0.17	-0.05	0.08	-0.01	-0.02	0.10	0.14	-0.07	0.00	-0.01
Traffic Lights Count	0.48	0.06	1.00	0.68	0.24	0.00	0.46	0.08	0.42	0.05	0.57	0.32	0.27	0.04	0.33	0.05	-0.03	-0.04	0.14	0.01
Traffic Lights Density (/km)	-0.01	-0.06	0.68	1.00	0.11	0.05	0.17	0.16	0.01	-0.02	0.11	0.29	0.05	0.02	0.29	-0.04	-0.16	0.17	0.08	-0.05
Shops Count	0.13	0.09	0.24	0.11	1.00	0.76	0.54	0.33	0.19	0.08	0.08	-0.04	0.28	0.20	0.06	-0.10	-0.28	-0.06	-0.04	-0.05
Shops Denisty (/km)	-0.12	-0.02	0.00	0.05	0.76	1.00	0.21	0.29	-0.06	0.02	-0.13	-0.10	0.11	0.24	0.00	-0.11	-0.32	0.00	-0.06	-0.07
Accidents Count	0.41	0.13	0.46	0.17	0.54	0.21	1.00	0.67	0.44	0.15	0.34	0.03	0.44	0.19	0.13	-0.05	-0.18	-0.08	0.01	-0.01
Accidents Denisty	-0.11	-0.03	0.08	0.16	0.33	0.29	0.67	1.00	0.01	0.14	-0.12	-0.09	0.15	0.21	0.01	-0.15	-0.33	0.09	-0.01	-0.04
Bridges Count	0.73	0.22	0.42	0.01	0.19	-0.06	0.44	0.01	1.00	0.55	0.63	0.09	0.34	0.06	0.11	0.13	0.17	-0.21	0.09	-0.01
Bridges Denisty	0.04	0.01	0.05	-0.02	0.08	0.02	0.15	0.14	0.55	1.00	0.01	-0.08	0.11	0.12	-0.03	-0.06	-0.06	-0.16	0.05	-0.09
Cycleway Total Length	0.84	0.17	0.57	0.11	0.08	-0.13	0.34	-0.12	0.63	0.01	1.00	0.54	0.19	-0.11	0.45	0.20	0.30	-0.12	0.10	0.09
Cycleway Proportion	0.15	-0.05	0.32	0.29	-0.04	-0.10	0.03	-0.09	0.09	-0.08	0.54	1.00	-0.13	-0.24	0.79	0.06	0.24	0.09	0.03	0.04
Cycle Lane Total Length	0.35	0.08	0.27	0.05	0.28	0.11	0.44	0.15	0.34	0.11	0.19	-0.13	1.00	0.71	0.20	-0.05	-0.12	-0.12	-0.03	0.05
Cycle Lane Proportion	-0.02	-0.01	0.04	0.02	0.20	0.24	0.19	0.21	0.06	0.12	-0.11	-0.24	0.71	1.00	0.20	-0.11	-0.22	-0.09	-0.04	0.00
Bike Facility Proportion	0.15	-0.02	0.33	0.29	0.06	0.00	0.13	0.01	0.11	-0.03	0.45	0.79	0.20	0.20	1.00	0.02	0.15	0.05	0.01	0.05
Landuse 25m buffer - Archiculture	0.26	0.10	0.05	-0.04	-0.10	-0.11	-0.05	-0.15	0.13	-0.06	0.20	0.06	-0.05	-0.11	0.02	1.00	0.44	0.02	0.08	0.04
Landuse 25m buffer - Green	0.31	0.14	-0.03	-0.16	-0.28	-0.32	-0.18	-0.33	0.17	-0.06	0.30	0.24	-0.12	-0.22	0.15	0.44	1.00	-0.03	-0.06	-0.09
Landuse 25m buffer - Retail	-0.18	-0.07	-0.04	0.17	-0.06	0.00	-0.08	0.09	-0.21	-0.16	-0.12	0.09	-0.12	-0.09	0.05	0.02	-0.03	1.00	-0.02	-0.02
Landuse 25m buffer - Commercial	0.08	0.00	0.14	0.08	-0.04	-0.06	0.01	-0.01	0.09	0.05	0.10	0.03	-0.03	-0.04	0.01	0.08	-0.06	-0.02	1.00	0.07
Landuse 25m buffer - Industrial	0.06	-0.01	0.01	-0.05	-0.05	-0.07	-0.01	-0.04	-0.01	-0.09	0.09	0.04	0.05	0.00	0.05	0.04	-0.09	-0.02	0.07	1.00
Landuse 25m buffer - Residential	-0.35	-0.08	-0.03	0.13	0.20	0.23	0.14	0.32	-0.16	0.11	-0.36	-0.25	0.05	0.16	-0.17	-0.47	-0.63	0.11	-0.07	-0.28
Landuse 50m buffer - Archiculture	0.23	0.10	0.07	0.01	-0.11	-0.12	-0.07	-0.15	0.08	-0.10	0.19	0.10	-0.05	-0.10	0.07	0.93	0.44	0.08	0.04	0.02
Landuse 50m buffer - Green	0.35	0.14	-0.01	-0.16	-0.26	-0.32	-0.18	-0.34	0.20	-0.05	0.33	0.23	-0.10	-0.21	0.14	0.47	0.97	-0.02	-0.05	-0.09
Landuse 50m buffer - Retail	-0.19	-0.08	-0.05	0.17	-0.08	-0.02	-0.09	0.08	-0.22	-0.17	-0.13	0.09	-0.13	-0.10	0.05	0.03	-0.02	0.99	-0.01	-0.03
Landuse 50m buffer - Commercial	0.09	-0.01	0.16	0.11	-0.04	-0.06	0.02	0.00	0.10	0.06	0.11	0.05	-0.04	-0.05	0.01	0.07	-0.06	-0.03	0.99	0.06
Landuse 50m buffer - Industrial	0.08	-0.01	0.02	-0.04	-0.06	-0.07	-0.01	-0.04	0.00	-0.09	0.10	0.06	0.05	-0.01	0.06	0.05	-0.08	-0.02	0.07	1.00
Landuse 50m buffer - Residential	-0.36	-0.09	-0.04	0.13	0.20	0.23	0.14	0.32	-0.17	0.10	-0.37	-0.23	0.05	0.16	-0.16	-0.48	-0.62	0.11	-0.08	-0.29
Landuse 100m buffer - Archiculture	0.24	0.10	0.09	0.01	-0.11	-0.12	-0.08	-0.18	0.09	-0.09	0.20	0.11	-0.01	-0.05	0.10	0.84	0.43	0.05	0.02	0.02
Landuse 100m buffer - Green	0.38	0.15	-0.01	-0.17	-0.25	-0.31	-0.17	-0.34	0.22	-0.05	0.34	0.20	-0.06	-0.18	0.12	0.47	0.90	-0.03	-0.05	-0.08
Landuse 100m buffer - Retail	-0.19	-0.08	-0.03	0.19	-0.10	-0.04	-0.09	0.08	-0.22	-0.18	-0.12	0.11	-0.13	-0.10	0.08	0.04	0.00	0.96	0.00	-0.03
Landuse 100m buffer - Commercial	0.09	-0.01	0.17	0.12	-0.04	-0.07	0.03	-0.01	0.10	0.06	0.12	0.06	-0.04	-0.06	0.02	0.06	-0.05	-0.04	0.94	0.05
Landuse 100m buffer - Industrial	0.09	-0.01	0.03	-0.03	-0.06	-0.08	-0.01	-0.05	0.01	-0.08	0.12	0.08	0.04	-0.02	0.07	0.06	-0.08	-0.03	0.08	0.98
Landuse 100m buffer - Residential	-0.37	-0.09	-0.03	0.15	0.20	0.24	0.14	0.32	-0.17	0.11	-0.37	-0.22	0.05	0.16	-0.14	-0.49	-0.62	0.12	-0.09	-0.30
PM10 Level	-0.23	-0.06	0.15	0.27	0.32	0.32	0.29	0.41	-0.02	0.19	-0.22	-0.15	0.14	0.24	-0.02	-0.37	-0.60	-0.04	0.02	-0.05
PM10 Maximum Level	0.13	0.03	0.33	0.24	0.37	0.27	0.45	0.35	0.25	0.19	0.08	-0.10	0.27	0.24	0.03	-0.27	-0.44	-0.11	0.03	-0.01
NOx Level	0.11	-0.01	0.23	0.17	0.13	0.07	0.21	0.18	0.19	0.14	0.10	-0.03	0.10	0.10	0.02	-0.09	-0.26	-0.13	0.25	0.04
NOx Maximum Level	0.42	0.10	0.29	0.03	0.06	-0.07	0.23	0.00	0.39	0.10	0.39	0.10	0.14	-0.01	0.09	0.11	0.06	-0.13	0.25	0.07
Weighted Average Speed - Imputed	0.16	0.10	-0.14	-0.21	0.05	0.10	0.06	0.08	0.10	0.04	-0.17	-0.59	0.16	0.19	-0.42	0.14	-0.01	-0.11	0.02	-0.03
Weighted Average Speed - Known Only	0.30	0.09	0.06	-0.02	-0.02	-0.02	0.04	-0.04	0.16	-0.04	0.19	-0.04	0.07	0.01	-0.03	0.19	0.14	0.04	0.07	0.13
Turn Count	0.85	0.31	0.43	0.00	0.11	-0.12	0.38	-0.10	0.74	0.12	0.66	0.02	0.30	-0.02	0.06	0.18	0.22	-0.15	0.08	0.04
Turn Density (/km)	-0.45	-0.10	-0.28	0.01	-0.11	0.06	-0.24	0.03	-0.27	0.11	-0.50	-0.39	-0.16	0.02	-0.31	-0.20	-0.29	0.20	-0.05	-0.07
Intersection Count	0.85	0.25	0.66	0.15	0.36	0.00	0.62	0.05	0.74	0.12	0.74	0.14	0.46	0.07	0.20	0.10	0.08	-0.17	0.07	0.02
Intersection Density (/km)	-0.35	-0.13	0.08	0.33	0.28	0.31	0.17	0.38	-0.16	0.14	-0.33	-0.14	0.07	0.19	-0.02	-0.29	-0.52	0.18	-0.05	-0.11
Bike Facility Interruption Count	0.76	0.28	0.34	-0.06	0.10	-0.10	0.33	-0.10	0.64	0.11	0.49	-0.09	0.30	-0.02	-0.08	0.13	0.16	-0.16	0.04	0.03
Bike Facility Interruption Denisity (/km)	-0.26	-0.03	-0.23	-0.12	-0.06	0.06	-0.16	-0.03	-0.13	0.16	-0.37	-0.39	-0.10	-0.01	-0.38	-0.14	-0.22	0.05	-0.06	-0.04
50m buffer - Water Area	0.77	0.27	0.25	-0.07	-0.01	-0.14	0.22	-0.11	0.63	0.10	0.58	0.05	0.29	0.00	0.07	0.19	0.20	-0.14	0.14	0.09
50m buffer - Water Area / km	0.25	0.15	0.04	-0.09	-0.11	-0.18	0.04	-0.08	0.34	0.29	0.16	-0.01	0.12	0.01	0.01	0.09	0.08	-0.12	0.13	0.05
Home Count	0.54	0.19	0.32	-0.06	0.31	0.08	0.44	0.01	0.49	0.06	0.37	-0.11	0.39	0.09	-0.04	-0.10	-0.13	-0.24	-0.07	0.01
Home Density (/km)	-0.34	-0.08	-0.26	-0.17	0.12	0.29	-0.11	0.04	-0.26	-0.04	-0.39	-0.39	-0.03	0.12	-0.31	-0.27	-0.44	-0.13	-0.15	-0.03
Monument Count	0.33	0.19	0.38	0.14	0.61	0.31	0.59	0.26	0.46	0.24	0.16	-0.15	0.47	0.27	0.00	-0.10	-0.23	-0.21	-0.04	-0.07
Monument Weighted Count	0.46	0.21	0.35	0.09	0.52	0.25	0.53	0.19	0.51	0.21	0.25	-0.14	0.49	0.23	0.00	-0.03	-0.15	-0.21	-0.05	-0.06

	Landuse 25m buffer - Residential	Landuse 50m buffer - Archiculture	Landuse 50m buffer - Green	Landuse 50m buffer - Retail	Landuse 50m buffer - Commercial	Landuse 50m buffer - Industrial	Landuse 50m buffer - Residential	Landuse 100m buffer - Archiculture	Landuse 100m buffer - Green	Landuse 100m buffer - Retail	Landuse 100m buffer - Commercial	Landuse 100m buffer - Industrial	Landuse 100m buffer - Residential	PM10 Level	PM10 Maximum Level	NOx Level	NOx Maximum Level	Weighted Average Speed - Imputed	Weighted Average Speed - Known Only	Turn Count
Route Total Length	-0.35	0.23	0.35	-0.19	0.09	0.08	-0.36	0.24	0.38	-0.19	0.09	0.09	-0.37	-0.23	0.13	0.11	0.42	0.16	0.30	0.85
Degree of Detour	-0.08	0.10	0.14	-0.08	-0.01	-0.01	-0.09	0.10	0.15	-0.08	-0.01	-0.01	-0.09	-0.06	0.03	-0.01	0.10	0.10	0.09	0.31
Traffic Lights Count	-0.03	0.07	-0.01	-0.05	0.16	0.02	-0.04	0.09	-0.01	-0.03	0.17	0.03	-0.03	0.15	0.33	0.23	0.29	-0.14	0.06	0.43
Traffic Lights Density (/km)	0.13	0.01	-0.16	0.17	0.11	-0.04	0.13	0.01	-0.17	0.19	0.12	-0.03	0.15	0.27	0.24	0.17	0.03	-0.21	-0.02	0.00
Shops Count	0.20	-0.11	-0.26	-0.08	-0.04	-0.06	0.20	-0.11	-0.25	-0.10	-0.04	-0.06	0.20	0.32	0.37	0.13	0.06	0.05	-0.02	0.11
Shops Denisty (/km)	0.23	-0.12	-0.32	-0.02	-0.06	-0.07	0.23	-0.12	-0.31	-0.04	-0.07	-0.08	0.24	0.32	0.27	0.07	-0.07	0.10	-0.02	-0.12
Accidents Count	0.14	-0.07	-0.18	-0.09	0.02	-0.01	0.14	-0.08	-0.17	-0.09	0.03	-0.01	0.14	0.29	0.45	0.21	0.23	0.06	0.04	0.38
Accidents Denisty	0.32	-0.15	-0.34	0.08	0.00	-0.04	0.32	-0.18	-0.34	0.08	-0.01	-0.05	0.32	0.41	0.35	0.18	0.00	0.08	-0.04	-0.10
Bridges Count	-0.16	0.08	0.20	-0.22	0.10	0.00	-0.17	0.09	0.22	-0.22	0.10	0.01	-0.17	-0.02	0.25	0.19	0.39	0.10	0.16	0.74
Bridges Denisty	0.11	-0.10	-0.05	-0.17	0.06	-0.09	0.10	-0.09	-0.05	-0.18	0.06	-0.08	0.11	0.19	0.19	0.14	0.10	0.04	-0.04	0.12
Cycleway Total Length	-0.36	0.19	0.33	-0.13	0.11	0.10	-0.37	0.20	0.34	-0.12	0.12	0.12	-0.37	-0.22	0.08	0.10	0.39	-0.17	0.19	0.66
Cycleway Proportion	-0.25	0.10	0.23	0.09	0.05	0.06	-0.23	0.11	0.20	0.11	0.06	0.08	-0.22	-0.15	-0.10	-0.03	0.10	-0.59	-0.04	0.02
Cycle Lane Total Length	0.05	-0.05	-0.10	-0.13	-0.04	0.05	0.05	-0.01	-0.06	-0.13	-0.04	0.04	0.05	0.14	0.27	0.10	0.14	0.16	0.07	0.30
Cycle Lane Proportion	0.16	-0.10	-0.21	-0.10	-0.05	-0.01	0.16	-0.05	-0.18	-0.10	-0.06	-0.02	0.16	0.24	0.24	0.10	-0.01	0.19	0.01	-0.02
Bike Facility Proportion	-0.17	0.07	0.14	0.05	0.01	0.06	-0.16	0.10	0.12	0.08	0.02	0.07	-0.14	-0.02	0.03	0.02	0.09	-0.42	-0.03	0.06
Landuse 25m buffer - Archiculture	-0.47	0.93	0.47	0.03	0.07	0.05	-0.48	0.84	0.47	0.04	0.06	0.06	-0.49	-0.37	-0.27	-0.09	0.11	0.14	0.19	0.18
Landuse 25m buffer - Green	-0.63	0.44	0.97	-0.02	-0.06	-0.08	-0.62	0.43	0.90	0.00	-0.05	-0.08	-0.62	-0.60	-0.44	-0.26	0.06	-0.01	0.14	0.22
Landuse 25m buffer - Retail	0.11	0.08	-0.02	0.99	-0.03	-0.02	0.11	0.05	-0.03	0.96	-0.04	-0.03	0.12	-0.04	-0.11	-0.13	-0.13	-0.11	0.04	-0.15
Landuse 25m buffer - Commercial	-0.07	0.04	-0.05	-0.01	0.99	0.07	-0.08	0.02	-0.05	0.00	0.94	0.08	-0.09	0.02	0.03	0.25	0.25	0.02	0.07	0.08
Landuse 25m buffer - Industrial	-0.28	0.02	-0.09	-0.03	0.06	1.00	-0.29	0.02	-0.08	-0.03	0.05	0.98	-0.30	-0.05	-0.01	0.04	0.07	-0.03	0.13	0.04
Landuse 25m buffer - Residential	1.00	-0.47	-0.66	0.11	-0.06	-0.29	1.00	-0.49	-0.67	0.09	-0.05	-0.29	0.98	0.57	0.39	0.05	-0.22	-0.01	-0.22	-0.20
Landuse 50m buffer - Archiculture	-0.47	1.00	0.47	0.08	0.03	0.03	-0.48	0.94	0.48	0.10	0.02	0.03	-0.49	-0.40	-0.32	-0.15	0.06	0.10	0.15	0.15
Landuse 50m buffer - Green	-0.66	0.47	1.00	-0.02	-0.05	-0.08	-0.66	0.48	0.97	0.00	-0.04	-0.08	-0.66	-0.61	-0.42	-0.23	0.08	0.04	0.18	0.25
Landuse 50m buffer - Retail	0.11	0.08	-0.02	1.00	-0.02	-0.03	0.11	0.06	-0.03	0.98	-0.03	-0.04	0.11	-0.05	-0.13	-0.14	-0.14	-0.11	0.02	-0.16
Landuse 50m buffer - Commercial	-0.06	0.03	-0.05	-0.02	1.00	0.06	-0.07	0.01	-0.05	-0.01	0.97	0.07	-0.07	0.03	0.05	0.26	0.26	0.00	0.07	0.08
Landuse 50m buffer - Industrial	-0.29	0.03	-0.08	-0.03	0.06	1.00	-0.30	0.03	-0.07	-0.03	0.05	0.99	-0.31	-0.05	-0.01	0.05	0.08	-0.03	0.14	0.05
Landuse 50m buffer - Residential	1.00	-0.48	-0.66	0.11	-0.07	-0.30	1.00	-0.51	-0.68	0.09	-0.06	-0.30	0.99	0.57	0.38	0.04	-0.24	-0.03	-0.23	-0.22
Landuse 100m buffer - Archiculture	-0.49	0.94	0.48	0.06	0.01	0.03	-0.51	1.00	0.51	0.07	-0.01	0.03	-0.52	-0.43	-0.32	-0.13	0.07	0.12	0.18	0.16
Landuse 100m buffer - Green	-0.67	0.48	0.97	-0.03	-0.05	-0.07	-0.68	0.51	1.00	-0.01	-0.05	-0.08	-0.68	-0.62	-0.40	-0.20	0.10	0.10	0.22	0.27
Landuse 100m buffer - Retail	0.09	0.10	0.00	0.98	-0.01	-0.03	0.09	0.07	-0.01	1.00	-0.02	-0.03	0.10	-0.06	-0.13	-0.14	-0.13	-0.13	0.02	-0.16
Landuse 100m buffer - Commercial	-0.05	0.02	-0.04	-0.03	0.97	0.05	-0.06	-0.01	-0.05	-0.02	1.00	0.06	-0.06	0.04	0.05	0.27	0.26	-0.02	0.05	0.07
Landuse 100m buffer - Industrial	-0.29	0.03	-0.08	-0.04	0.07	0.99	-0.30	0.03	-0.08	-0.03	0.06	1.00	-0.31	-0.04	0.00	0.06	0.09	-0.03	0.14	0.06
Landuse 100m buffer - Residential	0.98	-0.49	-0.66	0.11	-0.07	-0.31	0.99	-0.52	-0.68	0.10	-0.06	-0.31	1.00	0.59	0.39	0.04	-0.25	-0.05	-0.24	-0.23
PM10 Level	0.57	-0.40	-0.61	-0.05	0.03	-0.05	0.57	-0.43	-0.62	-0.06	0.04	-0.04	0.59	1.00	0.83	0.50	0.08	-0.01	-0.11	-0.07
PM10 Maximum Level	0.39	-0.32	-0.42	-0.13	0.05	-0.01	0.38	-0.32	-0.40	-0.13	0.05	0.00	0.39	0.83	1.00	0.56	0.33	0.10	0.07	0.23
NOx Level	0.05	-0.15	-0.23	-0.14	0.26	0.05	0.04	-0.13	-0.20	-0.14	0.27	0.06	0.04	0.50	0.56	1.00	0.73	0.13	0.15	0.16
NOx Maximum Level	-0.22	0.06	0.08	-0.14	0.26	0.08	-0.24	0.07	0.10	-0.13	0.26	0.09	-0.25	0.08	0.33	0.73	1.00	0.11	0.21	0.39
Weighted Average Speed - Imputed	-0.01	0.10	0.04	-0.11	0.00	-0.03	-0.03	0.12	0.10	-0.13	-0.02	-0.03	-0.05	-0.01	0.10	0.13	0.11	1.00	0.56	0.14
Weighted Average Speed - Known Only	-0.22	0.15	0.18	0.02	0.07	0.14	-0.23	0.18	0.22	0.02	0.05	0.14	-0.24	-0.11	0.07	0.15	0.21	0.56	1.00	0.20
Turn Count	-0.20	0.15	0.25	-0.16	0.08	0.05	-0.22	0.16	0.27	-0.16	0.07	0.06	-0.23	-0.07	0.23	0.16	0.39	0.14	0.20	1.00
Turn Density (/km)	0.34	-0.19	-0.30	0.21	-0.07	-0.09	0.34	-0.19	-0.32	0.21	-0.09	-0.10	0.34	0.26	0.04	-0.04	-0.27	0.04	-0.17	-0.09
Intersection Count	-0.10	0.08	0.11	-0.18	0.08	0.03	-0.11	0.09	0.14	-0.18	0.08	0.04	-0.11	0.05	0.36	0.18	0.39	0.08	0.18	0.83
Intersection Density (/km)	0.56	-0.26	-0.53	0.18	-0.05	-0.13	0.56	-0.27	-0.53	0.17	-0.06	-0.13	0.56	0.52	0.33	0.04	-0.23	-0.02	-0.14	-0.20
Bike Facility Interruption Count	-0.13	0.10	0.19	-0.17	0.04	0.04	-0.14	0.09	0.20	-0.17	0.04	0.04	-0.14	-0.06	0.21	0.10	0.33	0.18	0.17	0.79
Bike Facility Interruption Denisity (/km)	0.28	-0.16	-0.23	0.06	-0.06	-0.05	0.28	-0.19	-0.23	0.04	-0.06	-0.07	0.28	0.19	0.05	-0.06	-0.17	0.07	-0.14	-0.11
50m buffer - Water Area	-0.27	0.16	0.24	-0.15	0.15	0.10	-0.28	0.15	0.25	-0.14	0.16	0.11	-0.28	-0.15	0.07	0.09	0.29	0.17	0.26	0.66
50m buffer - Water Area / km	-0.10	0.07	0.10	-0.12	0.14	0.06	-0.10	0.04	0.08	-0.11	0.18	0.07	-0.10	0.00	0.05	0.08	0.14	0.04	0.07	0.23
Home Count	0.11	-0.12	-0.15	-0.25	-0.08	0.01	0.11	-0.11	-0.13	-0.26	-0.09	0.02	0.11	0.20	0.36	0.04	0.17	0.08	0.01	0.60
Home Density (/km)	0.39	-0.29	-0.48	-0.13	-0.17	-0.04	0.39	-0.29	-0.48	-0.16	-0.19	-0.04	0.39	0.35	0.18	-0.06	-0.23	0.12	-0.15	-0.26
Monument Count	0.20	-0.12	-0.21	-0.23	-0.03	-0.07	0.20	-0.12	-0.20	-0.24	-0.03	-0.08	0.21	0.37	0.48	0.19	0.13	0.11	-0.02	0.38
Monument Weighted Count	0.10	-0.05	-0.12	-0.22	-0.04	-0.05	0.10	-0.04	-0.09	-0.23	-0.04	-0.05	0.10	0.23	0.37	0.13	0.13	0.15	0.03	0.46

	Turn Density (/km)	Intersection Count	Intersection Density (/km)	Bike Facility Interruption Count	Bike Facility Interruption Denisity (/km)	50m buffer - Water Area	50m buffer - Water Area / km	Home Count	Home Density (/km)	Monument Count	Monument Weighted Count
Route Total Length	-0.45	**0.82**	-0.25	**0.76**	-0.26	**0.77**	0.25	**0.54**	-0.34	0.33	0.46
Degree of Detour	-0.10	0.27	-0.06	0.28	-0.03	0.27	0.15	0.19	-0.08	0.19	0.21
Traffic Lights Count	-0.28	**0.64**	0.12	0.34	-0.23	0.25	0.04	0.32	-0.26	0.38	0.35
Traffic Lights Density (/km)	0.01	0.16	0.33	-0.06	-0.12	-0.07	-0.09	-0.06	-0.17	0.14	0.09
Shops Count	-0.11	0.29	0.15	0.10	-0.06	-0.01	-0.11	0.31	0.12	**0.61**	**0.52**
Shops Denisty (/km)	0.06	-0.04	0.15	-0.10	0.06	-0.14	-0.18	0.08	0.29	0.31	0.25
Accidents Count	-0.24	**0.57**	0.13	0.33	-0.16	0.22	0.04	0.44	-0.11	**0.59**	**0.53**
Accidents Denisty	0.03	0.03	0.26	-0.10	-0.03	-0.11	-0.08	0.01	0.04	0.26	0.19
Bridges Count	-0.27	**0.70**	-0.12	**0.64**	-0.13	**0.63**	0.34	0.49	-0.26	0.46	**0.51**
Bridges Denisty	0.11	0.10	0.10	0.11	0.16	0.10	0.29	0.06	-0.04	0.24	0.21
Cycleway Total Length	-0.50	**0.71**	-0.21	0.49	-0.37	**0.58**	0.16	0.37	-0.39	0.16	0.25
Cycleway Proportion	-0.39	0.16	-0.03	-0.09	-0.39	0.05	-0.01	-0.11	-0.39	-0.15	-0.14
Cycle Lane Total Length	-0.16	0.37	-0.01	0.30	-0.10	0.29	0.12	0.39	-0.03	0.47	0.49
Cycle Lane Proportion	0.02	0.03	0.08	-0.02	-0.01	0.00	0.01	0.09	0.12	0.27	0.23
Bike Facility Proportion	-0.31	0.26	0.17	-0.08	-0.38	0.07	0.01	-0.04	-0.31	0.00	0.00
Landuse 25m buffer - Archiculture	-0.20	0.14	-0.17	0.13	-0.14	0.19	0.09	-0.10	-0.27	-0.10	-0.03
Landuse 25m buffer - Green	-0.29	0.13	-0.33	0.16	-0.22	0.20	0.08	-0.13	-0.44	-0.23	-0.15
Landuse 25m buffer - Retail	0.20	-0.15	0.14	-0.16	0.05	-0.14	-0.12	-0.24	-0.13	-0.21	-0.21
Landuse 25m buffer - Commercial	-0.05	0.07	-0.06	0.04	-0.06	0.14	0.13	-0.07	-0.15	-0.04	-0.05
Landuse 25m buffer - Industrial	-0.07	0.02	-0.10	0.03	-0.04	0.09	0.05	0.01	-0.03	-0.07	-0.06
Landuse 25m buffer - Residential	0.34	-0.14	0.38	-0.13	0.28	-0.27	-0.10	0.11	0.39	0.20	0.10
Landuse 50m buffer - Archiculture	-0.19	0.12	-0.13	0.10	-0.16	0.16	0.07	-0.12	-0.29	-0.12	-0.05
Landuse 50m buffer - Green	-0.30	0.15	-0.35	0.19	-0.23	0.24	0.10	-0.15	-0.48	-0.21	-0.12
Landuse 50m buffer - Retail	0.21	-0.16	0.15	-0.17	0.06	-0.15	-0.12	-0.25	-0.13	-0.23	-0.22
Landuse 50m buffer - Commercial	-0.07	0.08	-0.06	0.04	-0.06	0.15	0.14	-0.08	-0.17	-0.03	-0.04
Landuse 50m buffer - Industrial	-0.09	0.03	-0.11	0.04	-0.05	0.10	0.06	0.01	-0.04	-0.07	-0.05
Landuse 50m buffer - Residential	0.34	-0.15	0.38	-0.14	0.28	-0.28	-0.10	0.11	0.39	0.20	0.10
Landuse 100m buffer - Archiculture	-0.19	0.13	-0.13	0.09	-0.19	0.15	0.04	-0.11	-0.29	-0.12	-0.04
Landuse 100m buffer - Green	-0.32	0.17	-0.36	0.20	-0.23	0.25	0.08	-0.13	-0.48	-0.20	-0.09
Landuse 100m buffer - Retail	0.21	-0.15	0.15	-0.17	0.04	-0.14	-0.11	-0.26	-0.16	-0.24	-0.23
Landuse 100m buffer - Commercial	-0.09	0.08	-0.07	0.04	-0.06	0.16	0.18	-0.09	-0.19	-0.03	-0.04
Landuse 100m buffer - Industrial	-0.10	0.04	-0.11	0.04	-0.07	0.11	0.07	0.02	-0.04	-0.08	-0.05
Landuse 100m buffer - Residential	0.34	-0.15	0.38	-0.14	0.28	-0.28	-0.10	0.11	0.39	0.21	0.10
PM10 Level	0.26	0.00	0.32	-0.06	0.19	-0.15	0.00	0.20	0.35	0.37	0.23
PM10 Maximum Level	0.04	0.29	0.18	0.21	0.05	0.07	0.05	0.36	0.18	0.48	0.37
NOx Level	-0.04	0.16	0.01	0.10	-0.06	0.09	0.08	0.04	-0.06	0.19	0.13
NOx Maximum Level	-0.27	0.36	-0.18	0.33	-0.17	0.29	0.14	0.17	-0.23	0.13	0.13
Weighted Average Speed - Imputed	0.04	0.08	-0.02	0.18	0.07	0.17	0.04	0.08	0.12	0.11	0.15
Weighted Average Speed - Known Only	-0.17	0.14	-0.17	0.17	-0.14	0.26	0.07	0.01	-0.15	-0.02	0.03
Turn Count	-0.09	**0.81**	-0.11	**0.79**	-0.11	**0.66**	0.23	**0.60**	-0.26	0.38	0.46
Turn Density (/km)	1.00	-0.32	0.32	-0.21	0.44	-0.32	-0.16	-0.18	0.31	-0.10	-0.14
Intersection Count	-0.36	**0.94**	0.04	**0.74**	-0.19	**0.57**	0.15	**0.68**	-0.25	**0.57**	**0.59**
Intersection Density (/km)	0.41	0.00	**0.75**	-0.17	0.30	-0.34	-0.25	0.01	0.27	0.23	0.11
Bike Facility Interruption Count	-0.21	**0.68**	-0.14	1.00	0.26	**0.62**	0.26	**0.57**	-0.20	0.33	0.41
Bike Facility Interruption Denisity (/km)	0.44	-0.20	0.16	0.26	1.00	-0.15	0.01	-0.04	0.24	-0.06	-0.07
50m buffer - Water Area	-0.32	**0.55**	-0.24	**0.62**	-0.15	1.00	**0.65**	0.32	-0.29	0.26	0.41
50m buffer - Water Area / km	-0.16	0.15	-0.17	0.26	0.01	**0.65**	1.00	0.02	-0.26	0.15	0.19
Home Count	-0.18	**0.62**	-0.01	**0.57**	-0.04	0.32	0.02	1.00	0.37	0.47	0.48
Home Density (/km)	0.31	-0.25	0.16	-0.20	0.24	-0.29	-0.26	0.37	1.00	0.04	-0.02
Monument Count	-0.10	0.48	0.11	0.33	-0.06	0.26	0.15	0.47	0.04	1.00	**0.93**
Monument Weighted Count	-0.14	**0.51**	0.04	0.41	-0.07	0.41	0.19	0.48	-0.02	**0.93**	1.00

Appendix VI - Correlations Among Personal Characteristics

	Race Bike	Mountain Bike	Speed	Physical Health	Mental Health	Climate	Traffic Safety	Travel Costs	Being Outside	Enjoyment	Ease	Ensurance (e.g. no traffic jams)	Precipitation	mtb	Cold	Tired	Sweat	Distant Destination	Luggage	Fancy Clothing	Gender	Physical Condition	Intention to Cycle	Age	Living Alone	mtb	mtb	mtb	Lower Education Level	Medium Education Level	High Education Level	Age 65+
Race Bike	1.00	0.16	-0.12	-0.05	-0.02	-0.12	0.02	0.02	-0.01	0.06	-.194*	-0.13	0.15	0.10	0.09	0.14	0.14	-0.08	0.05	0.14	-.240**	0.09	0.04	-0.11	.192*	0.11	-.216*	-0.05	-0.01	-0.15	0.15	-0.01
Mountain Bike	0.16	1.00	-0.05	0.11	-0.02	0.16	0.03	-0.14	0.06	0.10	-0.06	-0.09	0.08	-0.03	0.07	0.03	0.16	0.10	0.05	0.13	-0.24	-0.02	0.08	-0.06	0.02	-0.02	0.02	-0.04	0.04	-0.06	0.04	-0.04
Speed	-0.12	-0.05	1.00	0.16	0.11	0.14	0.37	0.16	0.04	0.11	0.44	0.39	-0.18	-0.15	-0.11	-0.15	-0.27	-0.11	-0.16	-0.12	-0.02	0.05	0.08	0.10	-0.10	0.03	0.09	-0.09	-0.06	-0.01	0.04	0.00
Physical Health	-0.05	0.11	0.16	1.00	0.47	0.45	0.14	0.07	0.48	0.46	0.30	0.22	-0.11	-0.25	-0.20	-0.18	-0.10	-0.03	-0.20	-0.05	-0.02	0.08	0.11	0.26	0.03	0.04	0.13	-0.27	-0.04	-0.06	0.08	0.12
Mental Health	-0.02	-0.02	0.11	0.47	1.00	0.25	0.30	0.08	0.65	0.57	0.23	0.25	-0.19	-0.23	-0.14	-0.16	-0.21	-0.35	-0.10	0.00	-0.04	0.07	0.05	0.13	0.02	-0.10	0.23	-0.21	0.05	-0.03	0.00	-0.02
Climate	-0.12	0.16	0.14	0.45	0.25	1.00	0.23	0.04	0.38	0.27	0.37	0.28	-0.31	-0.36	-0.23	-0.27	-0.11	-0.10	-0.12	-0.03	0.07	0.09	0.10	0.24	-0.10	-0.03	0.16	-0.09	-0.07	0.01	0.02	0.17
Traffic Safety	0.02	0.03	0.37	0.14	0.30	0.23	1.00	0.15	0.27	0.38	0.27	0.24	-0.10	-0.16	-0.07	-0.06	-0.13	-0.20	-0.08	-0.02	-0.10	0.05	0.23	0.17	-0.04	-0.11	0.22	-0.15	0.12	0.02	-0.09	0.23
Travel Costs	0.02	-0.14	0.16	0.07	0.08	0.04	0.15	1.00	-0.05	-0.09	0.22	0.28	0.02	0.13	0.17	0.17	0.00	0.00	0.01	0.08	0.13	0.02	0.30	-0.18	-0.02	0.09	-0.15	0.08	0.01	0.11	-0.11	0.08
Being Outside	-0.01	0.06	0.04	0.48	0.65	0.38	0.27	-0.05	1.00	0.63	0.14	0.20	-0.18	-0.27	-0.23	-0.16	-0.15	-0.27	-0.01	0.00	-0.05	0.03	0.08	0.23	-0.05	-0.02	0.21	-0.22	0.03	0.00	-0.02	0.11
Enjoyment	0.06	0.10	0.11	0.46	0.57	0.27	0.38	-0.09	0.63	1.00	0.21	0.24	-0.13	-0.28	-0.26	-0.16	-0.17	-0.25	-0.15	-0.06	-0.24	0.10	0.09	0.40	0.05	0.04	0.16	-0.34	0.00	0.06	-0.06	0.23
Ease	-0.19	-0.06	0.44	0.30	0.23	0.37	0.27	0.22	0.14	0.21	1.00	0.39	-0.12	-0.03	-0.07	-0.10	-0.03	-0.08	-0.11	0.01	0.11	0.07	0.09	0.08	-0.15	0.08	0.06	-0.04	0.00	-0.01	0.01	0.04
Ensurance (e.g. no traffic jams)	-0.13	-0.09	0.39	0.22	0.25	0.28	0.24	0.28	0.20	0.24	0.39	1.00	-0.14	-0.15	-0.13	-0.11	-0.30	-0.14	0.04	-0.08	-0.06	0.19	0.15	0.05	-0.10	-0.02	0.20	-0.17	0.04	-0.03	0.01	0.06
Precipitation	0.15	0.08	-0.18	-0.11	-0.19	-0.31	-0.10	0.02	-0.18	-0.13	-0.12	-0.14	1.00	0.75	0.59	0.52	0.50	0.28	0.47	0.53	-0.01	-0.27	0.15	-0.22	0.10	0.12	-0.20	-0.01	-0.07	-0.07	0.11	-0.11
Wind	0.10	-0.03	-0.15	-0.25	-0.23	-0.36	-0.16	0.13	-0.27	-0.28	-0.03	-0.15	0.75	1.00	0.66	0.62	0.52	0.24	0.38	0.52	0.01	-0.18	0.18	-0.22	0.13	0.15	-0.27	0.04	-0.04	0.05	-0.02	-0.11
Cold	0.09	0.07	-0.11	-0.20	-0.14	-0.23	-0.07	0.17	-0.23	-0.26	-0.07	-0.13	0.59	0.66	1.00	0.57	0.49	0.29	0.36	0.43	0.01	-0.21	0.09	-0.24	0.14	0.06	-0.19	0.01	-0.10	0.04	0.02	-0.09
Tired	0.14	0.03	-0.15	-0.18	-0.16	-0.27	-0.06	0.17	-0.16	-0.16	-0.10	-0.11	0.52	0.62	0.57	1.00	0.49	0.33	0.36	0.42	0.07	-0.23	0.07	-0.11	0.10	0.02	-0.13	0.04	-0.06	0.06	-0.02	-0.03
Sweat	0.14	0.16	-0.27	-0.10	-0.21	-0.11	-0.13	0.00	-0.15	-0.17	-0.03	-0.30	0.50	0.52	0.49	0.49	1.00	0.27	0.34	0.49	-0.06	-0.29	-0.01	-0.08	0.15	-0.06	-0.09	0.05	-0.01	-0.11	0.11	-0.04
Distant Destination	-0.08	0.10	-0.11	-0.03	-0.35	-0.10	-0.20	0.00	-0.27	-0.25	-0.08	-0.14	0.28	0.24	0.29	0.33	0.27	1.00	0.38	0.23	0.07	-0.13	0.03	-0.09	-0.07	0.04	0.01	-0.03	-0.20	-0.06	0.17	-0.15
Luggage	0.05	0.05	-0.16	-0.20	-0.10	-0.12	-0.08	0.01	-0.01	-0.15	-0.11	0.04	0.47	0.38	0.36	0.36	0.34	0.38	1.00	0.35	-0.05	-0.25	0.05	-0.18	0.02	-0.01	0.02	-0.06	-0.17	-0.03	0.12	-0.23
Fancy Clothing	0.14	0.13	-0.12	-0.05	0.00	-0.03	-0.02	0.08	0.00	-0.06	0.01	-0.08	0.53	0.52	0.43	0.42	0.49	0.23	0.35	1.00	0.15	-0.21	0.15	-0.04	0.14	0.04	-0.12	-0.08	-0.10	0.00	0.06	-0.07
Gender	-0.24	-0.24	-0.02	-0.02	-0.04	0.07	-0.10	0.13	-0.05	-0.24	0.11	-0.06	-0.01	0.01	0.01	0.07	-0.06	0.07	-0.05	0.15	1.00	-0.23	-0.08	-0.22	0.04	-0.11	-0.08	0.24	-0.07	-0.10	0.14	-0.12
Physical Condition	0.09	-0.02	0.05	0.08	0.07	0.09	0.05	0.02	0.03	0.10	0.07	0.19	-0.27	-0.18	-0.21	-0.23	-0.29	-0.13	-0.25	-0.21	-0.23	1.00	-0.01	0.07	0.00	0.04	0.02	-0.12	0.11	-0.09	0.03	0.10
Intention to Cycle	0.04	0.08	0.08	0.11	0.05	0.10	0.23	0.30	0.08	0.09	0.09	0.15	0.15	0.18	0.09	0.07	-0.01	0.03	0.05	0.15	-0.08	-0.01	1.00	-0.03	-0.02	0.06	-0.04	0.01	0.13	0.04	-0.11	0.13
Age	-0.11	-0.06	0.10	0.26	0.13	0.24	0.17	-0.18	0.23	0.40	0.08	0.05	-0.22	-0.22	-0.24	-0.11	-0.08	-0.09	-0.18	-0.04	-0.22	0.07	-0.03	1.00	0.00	0.11	0.22	-0.47	-0.07	0.07	-0.03	0.55
Living Alone	0.19	0.02	-0.10	0.03	0.02	-0.10	-0.04	-0.02	-0.05	0.05	-0.15	-0.10	0.10	0.13	0.14	0.10	0.15	-0.07	0.02	0.14	0.04	0.00	-0.02	0.00	1.00	-0.30	-0.32	-0.15	-0.04	0.09	-0.07	0.11
Couple without children	0.11	-0.02	0.03	0.04	-0.10	-0.03	-0.11	0.09	-0.02	0.04	0.08	-0.02	0.12	0.15	0.06	0.02	-0.06	0.04	-0.01	0.04	-0.11	0.04	0.06	0.11	-0.30	1.00	-0.57	-0.27	-0.10	-0.02	0.07	0.24
Alone / couple, with children	-0.22	0.02	0.09	0.13	0.23	0.16	0.22	-0.15	0.21	0.16	0.06	0.20	-0.20	-0.27	-0.19	-0.13	-0.09	0.01	0.02	-0.12	-0.08	0.02	-0.04	0.22	-0.32	-0.57	1.00	-0.29	-0.06	-0.07	0.10	-0.23
Student housing, community or with family	-0.05	-0.04	-0.09	-0.27	-0.21	-0.09	-0.15	0.08	-0.22	-0.34	-0.04	-0.17	-0.01	0.04	0.01	0.04	0.05	-0.03	-0.06	-0.08	0.24	-0.12	0.01	-0.47	-0.15	-0.27	-0.29	1.00	0.20	0.05	-0.15	-0.11
Lower Education Level	-0.01	0.04	-0.06	-0.04	0.05	-0.07	0.12	0.01	0.03	0.00	0.00	0.04	-0.07	-0.04	-0.10	-0.06	-0.01	-0.20	-0.17	-0.10	-0.07	0.11	0.13	-0.07	-0.04	-0.10	-0.06	0.20	1.00	-0.23	-0.31	0.11
Medium Education Level	-0.15	-0.06	-0.01	-0.06	-0.03	0.01	0.02	0.11	0.00	0.06	-0.01	-0.03	-0.07	0.05	0.04	0.06	-0.11	-0.06	-0.03	0.00	-0.10	-0.09	0.04	0.07	0.09	-0.02	-0.07	0.05	-0.23	1.00	-0.85	0.04
High Education Level	0.15	0.04	0.04	0.08	0.00	0.02	-0.09	-0.11	-0.02	-0.06	0.01	0.01	0.11	-0.02	0.02	-0.02	0.11	0.17	0.12	0.06	0.14	0.03	-0.11	-0.03	-0.07	0.07	0.10	-0.15	-0.31	-0.85	1.00	-0.10
Age 65+	-0.01	-0.04	0.00	0.12	-0.02	0.17	0.23	0.08	0.11	0.23	0.04	0.06	-0.11	-0.11	-0.09	-0.03	-0.04	-0.15	-0.23	-0.07	-0.12	0.10	0.13	0.55	0.11	0.24	-0.23	-0.11	0.11	0.04	-0.10	1.00

Appendix VII - Correlations Route Characteristics and Behavior

Route Attribute	Correlation with Chosen Route Dummy
Route Total Length	-0.04
Degree of Detour	-0.06
Traffic Lights Count	-0.04
Traffic Lights Density (/km)	-0.03
Shops Count	0.02
Shops Denisty (/km)	0.03
Accidents Count	-0.01
Accidents Denisty	0.02
Bridges Count	0.00
Bridges Denisty	0.04
Cycleway Total Length	-0.03
Cycleway Proportion	-0.04
Cycle Lane Total Length	0.03
Cycle Lane Proportion	0.04
Bike Facility Proportion	0.01
Landuse 25m buffer - Agriculture	-0.02
Landuse 25m buffer - Green	-0.07
Landuse 25m buffer - Retail	-0.02
Landuse 25m buffer - Commercial	0.04
Landuse 25m buffer - Industrial	0.00
Landuse 25m buffer - Residential	0.02
Landuse 50m buffer - Agriculture	-0.03
Landuse 50m buffer - Green	-0.07
Landuse 50m buffer - Retail	-0.02
Landuse 50m buffer - Commercial	0.04
Landuse 50m buffer - Industrial	0.00
Landuse 50m buffer - Residential	0.02
Landuse 100m buffer - Agriculture	-0.04
Landuse 100m buffer - Green	-0.06
Landuse 100m buffer - Retail	-0.03
Landuse 100m buffer - Commercial	0.04
Landuse 100m buffer - Industrial	-0.01
Landuse 100m buffer - Residential	0.03
PM10 Level	0.01
PM10 Maximum Level	-0.01
NOx Level	0.00

Route Attribute	Correlation with Chosen Route Dummy
NOx Maximum Level	-0.02
Weighted Average Speed - Imputed	-0.10
Weighted Average Speed - Known Only	0.03
Turn Count	-0.11
Turn Density (/km)	-0.09
Intersection Count	-0.06
Intersection Density (/km)	-0.06
Bike Facility Interruption Count	0.02
Bike Facility Interruption Denisity (/km)	0.10
50m buffer - Water Area	-0.04
50m buffer - Water Area / km	-0.04
Home Count	-0.01
Home Density (/km)	0.03
Monument Count	-0.01
Monument Weighted Count	-0.01

Appendix VIII - Differences Based on Personal Characteristics

Correlation with Route Choice Dummy

	Overall Sample	Young Adults	Adults	Seniors	65-	65+	Females	Males	No Racebike	Racebike	Children	No Children	Low Enjoyment (1-3)	High Enjoyment (5-7)	min	max	range
Route Total Length	-0.04	-0.05	-0.03	-0.03	-0.04	-0.02	-0.04	-0.03	-0.04	-0.03	-0.01	-0.04	-0.03	-0.04	-0.05	-0.02	0.03
Degree of Detour	-0.06	-0.13	-0.03	-0.08	-0.06	-0.05	-0.06	-0.05	-0.09	-0.01	0.15	-0.08	-0.08	-0.04	-0.13	-0.01	0.12
Traffic Lights Count	-0.04	-0.05	-0.05	-0.02	-0.04	-0.01	-0.04	-0.04	-0.04	-0.05	-0.03	-0.03	-0.04	-0.04	-0.05	-0.01	0.04
Traffic Lights Density (/km)	-0.03	-0.05	-0.04	0.00	-0.03	-0.02	-0.03	-0.04	-0.02	-0.05	-0.06	-0.02	-0.10	-0.04	-0.10	0.00	0.10
Shops Count	0.02	0.04	0.03	0.01	0.02	0.03	0.01	0.04	0.03	0.02	0.09	0.03	0.13	0.02	0.01	0.13	0.13
Shops Denisty (/km)	0.03	0.06	0.02	0.03	0.03	0.05	0.02	0.04	0.03	0.02	0.21	0.03	0.04	0.02	0.02	0.06	0.04
Accidents Count	-0.01	-0.02	0.00	-0.02	-0.01	-0.01	-0.01	0.00	-0.01	-0.02	-0.01	-0.02	0.06	-0.02	-0.02	0.06	0.09
Accidents Density	0.02	0.01	0.02	0.01	0.02	0.00	0.01	0.02	0.02	0.01	-0.02	0.01	0.02	0.01	0.00	0.02	0.02
Bridges Count	0.00	-0.03	0.01	0.01	0.00	-0.01	-0.01	0.01	0.00	-0.01	0.04	-0.01	0.05	0.00	-0.03	0.05	0.08
Bridges Denisty	0.04	-0.01	0.04	0.07	0.04	0.05	0.03	0.05	0.05	0.02	0.05	0.03	0.07	0.05	-0.01	0.07	0.08
Cycleway Total Length	-0.03	-0.02	-0.03	-0.03	-0.03	-0.03	-0.03	-0.02	-0.03	-0.03	-0.06	-0.03	-0.03	-0.03	-0.03	-0.02	0.01
Cycleway Proportion	-0.04	-0.01	-0.03	-0.07	-0.04	-0.07	-0.05	-0.02	-0.04	-0.03	-0.08	-0.05	-0.10	-0.04	-0.10	-0.01	0.10
Cycle Lane Total Length	0.03	0.03	0.03	0.03	0.03	0.05	0.03	0.02	0.04	0.02	0.05	0.02	0.12	0.02	0.02	0.12	0.11
Cycle Lane Proportion	0.04	0.03	0.04	0.04	0.03	0.05	0.04	0.02	0.04	0.02	0.05	0.03	0.07	0.02	0.02	0.07	0.05
Bike Facility Proportion	0.01	0.03	0.01	-0.02	0.01	-0.07	0.00	0.01	0.00	0.01	-0.11	0.00	0.03	0.00	-0.07	0.03	0.10
Landuse 25m buffer - Archiculture	-0.02	0.00	-0.05	0.03	-0.03	0.10	-0.02	-0.01	-0.03	0.02	-0.04	0.02	0.04	-0.03	-0.05	0.10	0.15
Landuse 25m buffer - Green	-0.07	0.06	-0.12	-0.07	-0.07	-0.08	-0.09	-0.03	-0.09	-0.02	-0.28	-0.01	-0.04	-0.10	-0.12	0.06	0.18
Landuse 25m buffer - Retail	-0.02	-0.03	-0.03	0.02	-0.02	0.09	0.00	-0.03	-0.04	0.06	0.00	0.00	-0.03	-0.02	-0.04	0.09	0.13
Landuse 25m buffer - Commercial	0.04	-0.04	0.04	0.07	0.04	0.01	0.05	0.02	0.04	0.02	0.07	0.03	0.01	0.05	-0.04	0.07	0.11
Landuse 25m buffer - Industrial	0.00	-0.08	-0.02	0.09	-0.01	0.04	0.03	-0.03	0.05	-0.06	0.25	-0.04	-0.02	0.06	-0.08	0.09	0.17
Landuse 25m buffer - Residential	0.02	-0.06	0.06	0.01	0.01	0.08	0.03	0.00	0.03	0.01	0.13	0.00	0.05	0.03	-0.06	0.08	0.14
Landuse 50m buffer - Archiculture	-0.03	0.00	-0.07	0.04	-0.04	0.13	-0.03	-0.03	-0.04	0.00	-0.14	0.02	0.00	-0.04	-0.07	0.13	0.20
Landuse 50m buffer - Green	-0.07	0.03	-0.10	-0.09	-0.07	-0.10	-0.10	-0.02	-0.09	-0.03	-0.28	-0.03	-0.05	-0.10	-0.10	0.03	0.13
Landuse 50m buffer - Retail	-0.02	-0.03	-0.04	0.01	-0.03	0.08	-0.01	-0.03	-0.05	0.06	-0.02	0.00	-0.04	-0.03	-0.05	0.08	0.12
Landuse 50m buffer - Commercial	0.04	-0.04	0.05	0.08	0.05	0.02	0.06	0.03	0.05	0.03	0.10	0.04	-0.01	0.06	-0.04	0.08	0.12
Landuse 50m buffer - Industrial	0.00	-0.08	-0.02	0.08	-0.01	0.04	0.03	-0.04	0.05	-0.06	0.26	-0.04	-0.03	0.06	-0.08	0.08	0.16
Landuse 50m buffer - Residential	0.02	-0.06	0.06	0.00	0.02	0.07	0.03	0.01	0.02	0.02	0.11	0.01	0.06	0.03	-0.06	0.07	0.12
Landuse 100m buffer - Archiculture	-0.04	0.00	-0.08	0.02	-0.04	0.00	-0.04	-0.03	-0.05	0.00	-0.19	0.00	-0.02	-0.05	-0.08	0.02	0.09
Landuse 100m buffer - Green	-0.06	0.03	-0.08	-0.09	-0.06	-0.13	-0.09	0.00	-0.09	-0.02	-0.25	-0.03	-0.07	-0.09	-0.13	0.03	0.15
Landuse 100m buffer - Retail	-0.03	-0.04	-0.05	0.01	-0.04	0.08	-0.02	-0.03	-0.05	0.05	-0.05	-0.01	-0.03	-0.03	-0.05	0.08	0.13
Landuse 100m buffer - Commercial	0.04	-0.05	0.04	0.09	0.04	0.04	0.06	0.01	0.05	0.02	0.13	0.05	-0.03	0.06	-0.05	0.09	0.14
Landuse 100m buffer - Industrial	-0.01	-0.07	-0.02	0.08	-0.01	0.05	0.03	-0.04	0.04	-0.06	0.26	-0.04	-0.03	0.05	-0.07	0.08	0.15
Landuse 100m buffer - Residential	0.03	-0.05	0.07	0.00	0.02	0.07	0.03	0.01	0.02	0.02	0.11	0.01	0.07	0.03	-0.05	0.07	0.12
PM10 Level	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.01	0.02
PM10 Maximum Level	-0.01	0.01	0.00	-0.02	0.00	-0.02	-0.01	-0.01	-0.01	0.00	0.01	-0.01	0.06	-0.01	-0.02	0.06	0.08
NOx Level	0.00	0.01	0.00	-0.01	0.00	0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.06	0.00	-0.01	0.06	0.07
NOx Maximum Level	-0.02	-0.01	-0.01	-0.03	-0.02	0.01	-0.02	-0.02	-0.02	-0.01	0.04	-0.02	0.06	-0.02	-0.03	0.06	0.09
Weighted Average Speed - Imputed	-0.10	-0.17	-0.09	-0.08	-0.10	-0.12	-0.09	-0.12	-0.09	-0.14	-0.13	-0.11	-0.08	-0.10	-0.17	-0.08	0.10
Weighted Average Speed - Known Only	0.03	-0.03	0.03	0.05	0.03	0.02	0.03	0.02	0.02	0.05	-0.01	0.01	0.05	0.04	-0.03	0.05	0.08
Turn Count	-0.11	-0.12	-0.11	-0.10	-0.11	-0.08	-0.11	-0.11	-0.11	-0.11	-0.02	-0.11	-0.14	-0.11	-0.14	-0.08	0.06
Turn Density (/km)	-0.09	-0.08	-0.10	-0.07	-0.09	-0.05	-0.07	-0.11	-0.07	-0.12	-0.03	-0.10	-0.12	-0.08	-0.12	-0.05	0.07
Intersection Count	-0.06	-0.07	-0.06	-0.06	-0.06	-0.05	-0.06	-0.06	-0.06	-0.06	-0.03	-0.06	-0.04	-0.06	-0.07	-0.04	0.03
Intersection Density (/km)	-0.06	-0.05	-0.07	-0.04	-0.05	-0.08	-0.04	-0.08	-0.05	-0.08	-0.14	-0.06	0.01	-0.06	-0.08	0.01	0.10
Bike Facility Interruption Count	0.02	-0.01	0.02	0.03	0.02	0.03	0.02	0.02	0.02	0.03	0.09	0.02	0.07	0.03	-0.01	0.07	0.08
Bike Facility Interruption Denisity (/km)	0.10	0.06	0.10	0.13	0.10	0.12	0.12	0.08	0.11	0.10	0.08	0.10	0.14	0.11	0.06	0.14	0.08
50m buffer - Water Area	-0.04	-0.05	-0.04	-0.04	-0.04	-0.03	-0.04	-0.04	-0.03	-0.05	-0.03	-0.04	-0.04	-0.04	-0.05	-0.03	0.03
50m buffer - Water Area / km	-0.04	-0.03	-0.04	-0.04	-0.03	-0.06	-0.04	-0.03	-0.03	-0.05	-0.09	-0.03	-0.01	-0.04	-0.06	-0.01	0.05
Home Count	-0.01	-0.03	0.00	-0.01	-0.01	-0.04	-0.01	-0.01	-0.01	-0.01	0.03	-0.02	-0.02	-0.01	-0.04	0.00	0.04
Home Density (/km)	0.03	0.01	0.04	0.03	0.03	0.01	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.04	0.01	0.04	0.03
Monument Count	-0.01	0.01	-0.01	0.00	-0.01	0.00	-0.01	0.00	0.00	-0.01	0.02	-0.01	0.14	-0.01	-0.01	0.14	0.15
Monument Weighted Count	-0.01	0.00	-0.01	-0.01	-0.01	0.01	-0.01	0.00	-0.01	-0.01	0.05	-0.01	0.12	-0.01	-0.01	0.12	0.13

Appendix IX - Output Models

Main Effects MNL Model:

Results:

Results Main Effects Multinomial Logit Model			**McFadden Pseudo Rho square** = 0.24, **AIC/N** = 28.6				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Degree of Detour	**-0.07**	**	0.00	-34.51	0.00	-0.07	-0.07
Turn Density (/km)	**-0.39**	**	0.01	-28.64	0.00	-0.42	-0.37
Number of Intersections	**-0.10**	**	0.00	-26.30	0.00	-0.11	-0.10
Number of Traffic Lights	**-0.11**	**	0.01	-13.47	0.00	-0.13	-0.09
Speed Limit (Imputed)	**-0.08**	**	0.01	-16.53	0.00	-0.09	-0.07
Agriculture (50m Buffer)	**0.01**	*	0.01	1.83	0.07	0.00	0.03
Number of Bridges	**0.46**	**	0.02	23.98	0.00	0.42	0.50

* significant at 5% level, ** significant at 1% level

McFadden Pseudo Rho square:

LL = -10610.00634
LL0 = -14031.95356
pseudo-R2 = 1 - (-10610.00634 / -14031.95356) = 0.24

Collinearity Diagnostics:

VIF Scores	
Variable	**VIF**
Degree of Detour	1.012
Turn Density (/km)	1.196
Number of Intersections	3.545
Number of Traffic Lights	1.941
Speed Limit (Imputed)	1.095
Agriculture (50m Buffer)	1.055
Number of Bridges	2.288

Main Effects MNL Model With Non Imputed Speed Limit:

Results:

Results Main Effects Multinomial Logit Model			**McFadden Pseudo Rho square** = 0.24, **AIC/N** = 28.1				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Degree of Detour	**-0.07**	**	0.00	-33.65	0.00	-0.07	-0.06
Turn Density (/km)	**-0.41**	**	0.01	-29.65	0.00	-0.43	-0.38
Number of Intersections	**-0.12**	**	0.00	-31.16	0.00	-0.13	-0.11
Number of Traffic Lights	**-0.10**	**	0.01	-12.33	0.00	-0.11	-0.08
Speed Limit (Non Imputed)	**0.03**	**	0.00	10.37	0.00	0.03	0.04
Agriculture (50m Buffer)	0.01		0.01	1.20	0.23	-0.01	0.03
Number of Bridges	**0.45**	**	0.02	23.16	0.00	0.41	0.49

* significant at 5% level, ** significant at 1% level

McFadden Pseudo Rho square:

LL = -10697.04941
LL0 = -14031.95356
pseudo-R2 = 1 - (-10697.04941 / -14031.95356) = 0.24

Main Effects PSL Model:

Results:

Results Main Effects Multinomial Logit Model			McFadden Pseudo Rho square = 0.52, AIC/N = 18.3				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Path Size Factor	**9.21**	**	0.13	69.35	0.00	8.95	9.47
Degree of Detour	**-0.11**	**	0.00	-46.35	0.00	-0.11	-0.10
Turn Density (/km)	**-0.49**	**	0.02	-32.54	0.00	-0.52	-0.46
Number of Intersections	**-0.13**	**	0.00	-30.60	0.00	-0.13	-0.12
Number of Traffic Lights	**-0.09**	**	0.01	-10.36	0.00	-0.11	-0.08
Speed Limit (Imputed)	**-0.10**	**	0.01	-18.85	0.00	-0.11	-0.09
Agriculture (50m Buffer)	-0.01		0.01	-1.43	0.15	-0.03	0.01
Number of Bridges	**0.31**	**	0.02	15.47	0.00	0.27	0.35

* significant at 5% level, ** significant at 1% level

McFadden Pseudo Rho square:
LL = -6797.12911
LL0 = -14031.95356
pseudo-R2 = 1- (-6797.12911 / -14031.95356) = 0.52

Collinearity Diagnostics:

VIF Scores	
Variable	**VIF**
Path Size Factor	1.071
Degree of Detour	1.203
Turn Density (/km)	3.857
Number of Intersections	1.952
Number of Traffic Lights	1.095
Speed Limit (Imputed)	1.055
Agriculture (50m Buffer)	2.297
Number of Bridges	1.248

PSL Model With Interaction Effects:

Results:

Results Path Size Logit Model			McFadden Pseudo Rho square = 0.53, AIC/N = 17.8				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Path Size Factor	**9.37**	**	0.13	69.45	0.00	9.11	9.64
Degree of Detour	**-0.10**	**	0.00	-45.20	0.00	-0.11	-0.10
Turn Density (/km)	**-0.45**	**	0.02	-29.49	0.00	-0.48	-0.42
Number of Intersections	**-0.14**	**	0.00	-31.32	0.00	-0.15	-0.13
Number of Traffic Lights	**-0.06**	**	0.01	-4.84	0.00	-0.08	-0.03
Number of Traffic Lights X Peak Hour	-0.03		0.02	-1.61	0.11	-0.07	0.01
Speed Limit (Imputed)	**-0.10**	**	0.01	-17.82	0.00	-0.11	-0.09
Agriculture (50m Buffer)	**0.06**	**	0.01	6.55	0.00	0.04	0.08
Agriculture X Commute	**-0.50**	**	0.03	-18.77	0.00	-0.55	-0.45
Agriculture X Leisure	-0.05		0.06	-0.82	0.41	-0.17	0.07
Number of Bridges	**0.30**	**	0.02	14.33	0.00	0.26	0.34

* significant at 5% level, ** significant at 1% level

McFadden Pseudo Rho square:

LL = -6608.93184
LL0 = -14031.95356

pseudo-R2 = 1- (-6608.93184 / -14031.95356) = 0.53

Collinearity Diagnostics:

VIF Scores	
Variable	Coefficient
Path Size Factor	1.081
Degree of Detour	1.204
Turn Density (/km)	3.878
Number of Intersections	2.071
Number of Traffic Lights	1.201
Number of Traffic Lights X Peak Hour	1.103
Speed Limit (Imputed)	1.111
Agriculture (50m Buffer)	1.106
Agriculture X Commute	1.102
Agriculture X Leisure	2.314
Number of Bridges	1.248

Latent Class Analysis:

Remark: Both the AIC/N and Mc Fadden's rho squared statistics are calculated manually (including the value of the log likelihood function) in case of the Latent Class Analysis (LCA), since NLogit is not able to handle the weighted cases properly in that situation.

Results Latent Class Model			**McFadden Pseudo Rho square** = 0.64, **AIC/N** = 13.7				
Variable	Coefficient	Significance	Standard Error	z	Prob. |z|>Z*	95% Confidence Interval	
Random utility parameters in latent class 1			***Average class probability:*** 0.747				
Path Size Factor	**16.89**	**	0.37	45.79	0.00	16.17	17.62
Degree of Detour	**-0.28**	**	0.01	-42.13	0.00	-0.30	-0.27
Turn Density (/km)	**-0.93**	**	0.03	-30.88	0.00	-0.98	-0.87
Number of Intersections	**-0.12**	**	0.01	-22.32	0.00	-0.13	-0.11
Number of Traffic Lights	**-0.55**	**	0.02	-30.84	0.00	-0.59	-0.52
Number of Traffic Lights X Peak Hour	**0.20**	**	0.04	4.55	0.00	0.11	0.29
Speed Limit (Imputed)	**-0.58**	**	0.01	-56.06	0.00	-0.60	-0.56
Agriculture (50m Buffer)	**-0.16**	**	0.01	-14.22	0.00	-0.18	-0.14
Agriculture X Commute	**-1.11**	**	0.08	-13.89	0.00	-1.27	-0.95
Agriculture X Leisure	**0.27**	**	0.04	7.38	0.00	0.20	0.35
Number of Bridges	**0.71**	**	0.05	14.25	0.00	0.61	0.81
Random utility parameters in latent class 2			***Average class probability:*** 0.253				
Path Size Factor	**6.87**	**	0.20	35.02	0.00	6.49	7.26
Degree of Detour	**-0.01**	**	0.00	-15.03	0.00	-0.01	-0.01
Turn Density (/km)	**-0.36**	**	0.03	-12.55	0.00	-0.41	-0.30
Number of Intersections	**-0.16**	**	0.01	-31.87	0.00	-0.17	-0.15
Number of Traffic Lights	0.01		0.03	0.45	0.65	-0.04	0.07
Number of Traffic Lights X Peak Hour	**-0.09**	*	0.05	-2.06	0.04	-0.18	0.00
Speed Limit (Imputed)	**-0.08**	**	0.01	-10.25	0.00	-0.09	-0.06
Agriculture (50m Buffer)	**0.16**	**	0.02	9.07	0.00	0.12	0.19
Agriculture X Commute	0.08		0.05	1.77	0.08	-0.01	0.17
Agriculture X Leisure	**0.53**	**	0.12	4.39	0.00	0.30	0.77
Number of Bridges	**-0.14**	**	0.03	-4.68	0.00	-0.20	-0.08
Probability model class 1	***(Parameters class 2 fixed to zero.)***						
Constant	**3.73**	**	0.66	5.67	0.00	2.44	5.02
Motivated by Enjoyment	**-0.34**	**	0.11	-3.21	0.00	-0.55	-0.13
Race Bike Ownership	**-1.44**	**	0.19	-7.68	0.00	-1.80	-1.07
Age 65+ (retired)	**-1.44**	**	0.37	-3.88	0.00	-2.16	-0.71

** significant at 5% level, ** significant at 1% level*

McFadden Pseudo Rho square:
LL = -5075.024310
LL0 = -14031.95356

pseudo-R2 = 1- (-8082.96426 / -14031.95356) = 0.64

Akaike Information Criterion:
k = 11+11+ 4 = 26 (number of parameters)
N = 743 (number of choice sets)
LL = -5075.024310

$$\mathrm{AIC} = 2k - 2\ln(\hat{L})$$

AIC = 2(26) - 2(-5075.024310) = 10202.04
AIC/N = 10202.04/743 = 13.7

Appendix X - Class Comparison

Independent Samples Test

			Levene's Test for Equality of Variances		t-test for Equality of Means						
			F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
Motivators	Speed	Equal variances assumed	0.12	0.73	-1.13	137.00	0.26	-0.28	0.25	-0.78	0.22
		Equal variances not assumed			-1.13	75.26	0.26	-0.28	0.25	-0.79	0.22
	Physical Health	Equal variances assumed	0.20	0.66	-2.69	137.00	**0.01**	-0.42	0.16	-0.73	-0.11
		Equal variances not assumed			-2.96	94.08	0.00	-0.42	0.14	-0.71	-0.14
	Mental Health	Equal variances assumed	2.24	0.14	-2.47	137.00	**0.02**	-0.52	0.21	-0.93	-0.10
		Equal variances not assumed			-2.88	108.63	0.01	-0.52	0.18	-0.87	-0.16
	Climate / Global Warming	Equal variances assumed	0.00	0.96	-0.98	137.00	0.33	-0.21	0.21	-0.63	0.21
		Equal variances not assumed			-1.05	88.03	0.30	-0.21	0.20	-0.60	0.19
	Traffic Safety	Equal variances assumed	2.20	0.14	-2.72	137.00	**0.01**	-0.78	0.29	-1.35	-0.21
		Equal variances not assumed			-2.51	63.74	0.02	-0.78	0.31	-1.40	-0.16
	Travel Costs	Equal variances assumed	0.06	0.81	-0.70	137.00	0.49	-0.21	0.30	-0.80	0.38
		Equal variances not assumed			-0.68	72.64	0.50	-0.21	0.30	-0.81	0.40
	Being Outside	Equal variances assumed	0.39	0.53	-2.48	137.00	**0.01**	-0.45	0.18	-0.81	-0.09
		Equal variances not assumed			-2.87	106.57	0.01	-0.45	0.16	-0.76	-0.14
	Enjoyment	Equal variances assumed	4.36	0.04	-3.08	137.00	0.00	-0.65	0.21	-1.07	-0.23
		Equal variances not assumed			-3.66	113.56	**0.00**	-0.65	0.18	-1.00	-0.30
	Ease	Equal variances assumed	0.17	0.68	-1.12	137.00	0.27	-0.23	0.20	-0.63	0.17
		Equal variances not assumed			-1.13	76.93	0.26	-0.23	0.20	-0.63	0.17
	Security (e.g. no traffic jams)	Equal variances assumed	0.96	0.33	-1.19	137.00	0.24	-0.29	0.25	-0.78	0.19
		Equal variances not assumed			-1.19	74.88	0.24	-0.29	0.25	-0.79	0.20
Deterrents	Precipitation	Equal variances assumed	0.01	0.92	1.19	137.00	0.24	0.41	0.34	-0.27	1.09
		Equal variances not assumed			1.20	76.25	0.24	0.41	0.34	-0.27	1.09
	Wind	Equal variances assumed	1.80	0.18	0.94	137.00	0.35	0.29	0.31	-0.32	0.91
		Equal variances not assumed			0.90	67.70	0.37	0.29	0.33	-0.36	0.95
	Cold	Equal variances assumed	0.22	0.64	1.03	137.00	0.31	0.29	0.29	-0.27	0.86
		Equal variances not assumed			1.03	75.17	0.31	0.29	0.29	-0.28	0.86
	Tired	Equal variances assumed	1.10	0.30	0.20	137.00	0.84	0.06	0.30	-0.53	0.64
		Equal variances not assumed			0.19	68.75	0.85	0.06	0.31	-0.56	0.67
	Sweaty	Equal variances assumed	0.14	0.71	0.86	137.00	0.39	0.27	0.32	-0.36	0.91
		Equal variances not assumed			0.87	77.53	0.39	0.27	0.32	-0.35	0.90
	Distant Destination	Equal variances assumed	3.55	0.06	1.75	137.00	**0.08**	0.52	0.30	-0.07	1.10
		Equal variances not assumed			1.65	66.38	0.10	0.52	0.31	-0.11	1.14
	Luggage	Equal variances assumed	1.42	0.24	-0.08	137.00	0.94	-0.02	0.33	-0.67	0.62
		Equal variances not assumed			-0.07	70.00	0.94	-0.02	0.34	-0.70	0.65
	Fancy Clothing	Equal variances assumed	0.34	0.56	0.45	137.00	0.65	0.15	0.33	-0.50	0.80
		Equal variances not assumed			0.44	70.23	0.66	0.15	0.34	-0.53	0.83
	Physical Condition Rating	Equal variances assumed	0.09	0.77	-0.03	137.00	0.98	-0.08	2.64	-5.30	5.14
		Equal variances not assumed			-0.03	91.08	0.97	-0.08	2.43	-4.91	4.75
Other	**Intention to Bike**	Equal variances assumed	0.77	0.38	-2.01	137.00	0.05	-0.64	0.32	-1.26	-0.01
		Equal variances not assumed			-1.88	65.20	0.07	-0.64	0.34	-1.31	0.04
	Age	Equal variances assumed	0.00	0.98	-1.33	137.00	0.19	-3.28	2.47	-8.17	1.61
		Equal variances not assumed			-1.35	77.68	0.18	-3.28	2.43	-8.13	1.56

If the p-value of Levene's test is less than 0.05, the "Unequal variance" result is used.
Oterwise, the "Equal variance" result is used.

Descriptives for Variables in Independent Samples Test

	Descriptive Statistics - Class 1					
		Minimum	Maximum	Mean	Std. Deviation	Median
Motivators	Speed	1.00	7.00	5.08	1.36	5.00
	Physical Health	1.00	7.00	**6.16**	0.89	**6.00**
	Mental Health	1.00	7.00	**5.65**	1.23	**6.00**
	Climate / Global Warming	1.00	7.00	6.01	1.20	6.00
	Traffic Safety	1.00	6.00	**3.56**	1.44	**4.00**
	Travel Costs	1.00	7.00	5.06	1.58	5.00
	Being Outside	1.00	7.00	**5.94**	1.06	**6.00**
	Enjoyment	1.00	7.00	**5.52**	1.25	**6.00**
	Ease	1.00	7.00	6.04	1.10	6.00
	Security (e.g. no traffic jams)	1.00	7.00	5.63	1.33	6.00
Deterrents	Precipitation	1.00	7.00	3.90	1.86	4.00
	Wind	1.00	7.00	3.24	1.61	3.00
	Cold	1.00	6.00	2.81	1.54	2.00
	Tired	1.00	7.00	2.89	1.54	2.00
	Sweaty	1.00	7.00	3.59	1.74	3.00
	Distant Destination	1.00	7.00	**5.22**	1.52	**5.50**
	Luggage	1.00	7.00	4.44	1.72	5.00
	Fancy Clothing	1.00	7.00	3.49	1.73	3.00
Other	Physical Condition Rating	8.00	100.00	75.53	14.94	77.00
	Intention to Bike	1.00	7.00	**3.51**	1.61	**4.00**
	Age	18.00	72.00	42.13	13.43	40.00

	Descriptive Statistics - Class 2					
		Minimum	Maximum	Mean	Std. Deviation	Median
Motivators	Speed	2.00	7.00	5.37	1.36	6.00
	Physical Health	4.00	7.00	**6.59**	0.71	**7.00**
	Mental Health	4.00	7.00	**6.17**	0.83	**6.00**
	Climate / Global Warming	3.00	7.00	6.22	1.01	7.00
	Traffic Safety	1.00	7.00	**4.34**	1.76	**5.00**
	Travel Costs	1.00	7.00	5.27	1.64	6.00
	Being Outside	4.00	7.00	**6.39**	0.74	**7.00**
	Enjoyment	4.00	7.00	**6.17**	0.80	**6.00**
	Ease	2.00	7.00	6.27	1.07	7.00
	Security (e.g. no traffic jams)	1.00	7.00	5.93	1.33	6.00
Deterrents	Precipitation	1.00	7.00	3.49	1.83	3.00
	Wind	1.00	6.00	2.95	1.82	2.00
	Cold	1.00	6.00	2.51	1.53	2.00
	Tired	1.00	6.00	2.83	1.70	2.00
	Sweaty	1.00	6.00	3.32	1.68	3.00
	Distant Destination	1.00	7.00	**4.71**	1.75	**5.00**
	Luggage	1.00	7.00	4.46	1.86	5.00
	Fancy Clothing	1.00	7.00	3.34	1.87	3.00
Other	Physical Condition Rating	40.00	100.00	75.61	12.20	75.00
	Intention to Bike	1.00	7.00	**4.15**	1.91	**5.00**
	Age	27.00	68.00	45.41	12.94	43.00

Results Chi-Square Tests

	Pearson Chi-Square	df	Asymptotic Significance (2-sided)
Race Bike Ownership	4.57	1	**0.03**
MTB Ownership	0.11	1	0.74
Gender	1.24	1	0.27
Age Group	2.38	2	0.31
Aged 65+	1.46	1	0.23
Household Composition	1.43	3	0.70
Educational Level	0.17	2	0.92

Appendix XI - Examples Bridges

Snorrebro Bridge in Copenhagen

Van Asch van Wijckskade in Utrecht

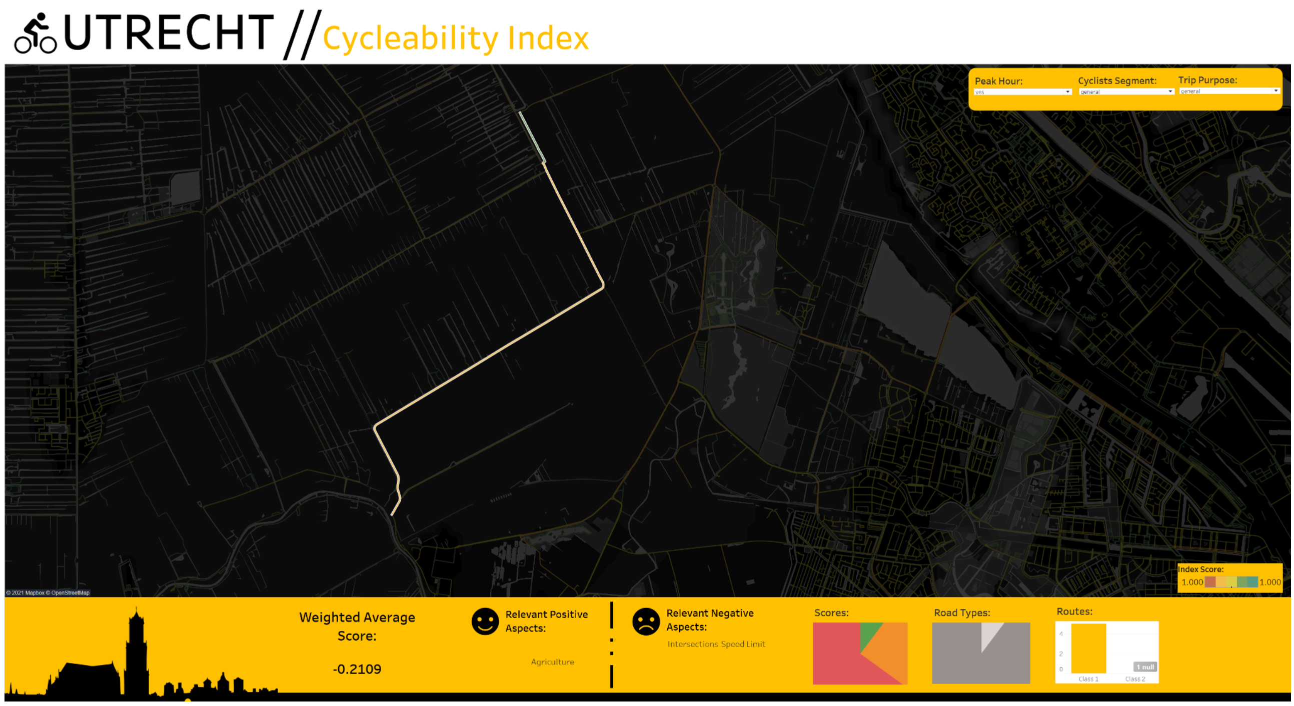
Ferry Street Bridge in Eugene

Source: Google Maps, Retrieved on July 1st 2021

Appendix XII - Functionality Dashboard

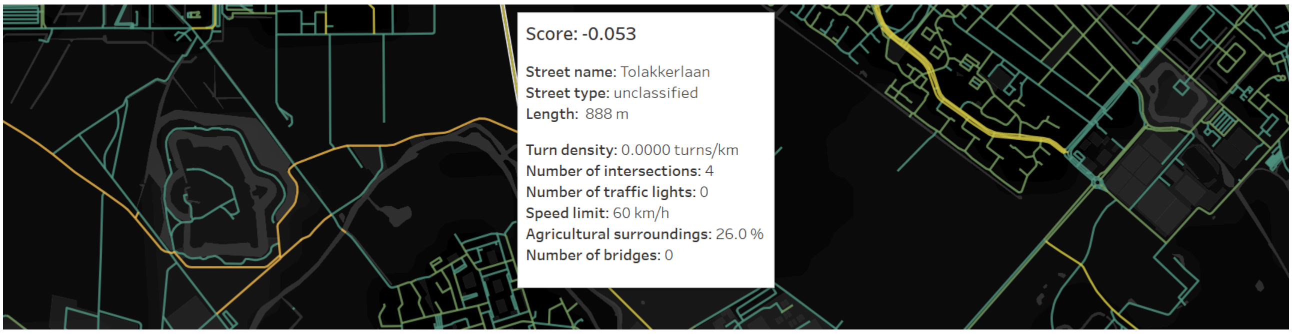
Selecting network segments, series of segments or areas.

One can select one segment to focus on by clicking on it in the map canvas. Multiple segments can be selected by holding the [control] button while clicking them. Further, the selection tool, which can be accessed from the top left map menu, can be used to select larger areas. The charts and figures on the dashboard will be refreshed automatically. The selection can be cancelled by clicking anywhere on the map and pressing [Esc] or making an empty selection with the selection tool.



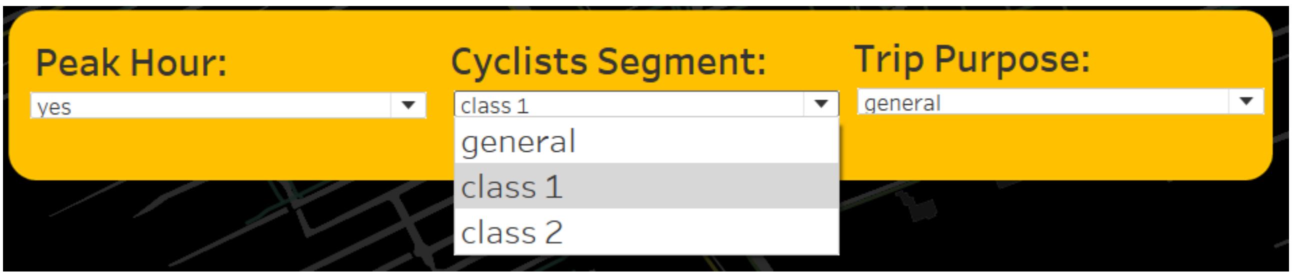
The bottom bar of the dashboard summarizes the information on the selected segment(s). One can see the average score of the segments (weighted for their length) and the relevant positive and negative aspects which contribute to this score. One can also see what proportion of the segments scores below, at or above average. Further, the total length of infrastructure for each road type is displayed in a pie chart. This allows the user to study the performance of the selected segments in detail. Moreover, the chart at the bottom right corner displays the number of routes of class 1 and class 2 which passed at least one of the segments in the selection. This provides an indication of which segments are (un)popular among each segment.

Information regarding a specific segment is viewed upon hover-over, which makes a tooltip appear. This tooltip provides data on all the relevant route characteristics, including the index score of the segment.



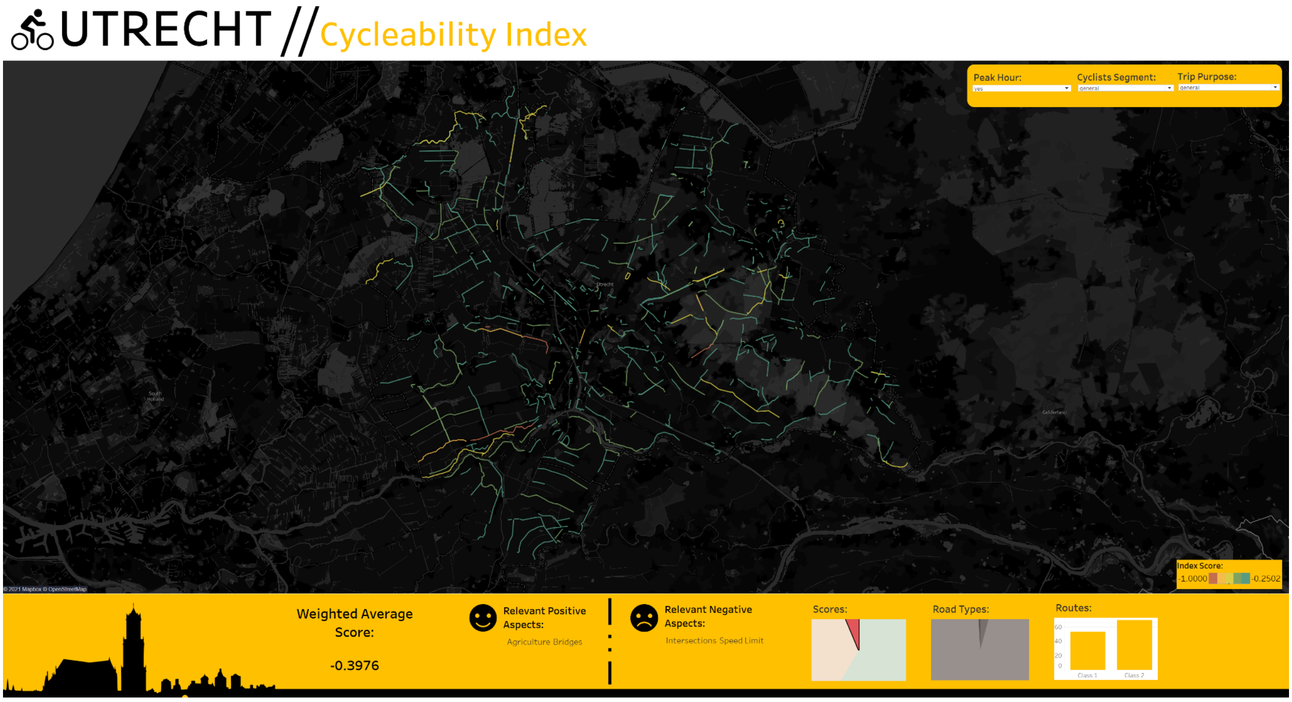
Select a class, trip purpose and departure time.

Based on the results in §6.2, one can conclude that trip purpose and departure time influence route preferences. These findings have been translated to the dashboard such that users can select a trip purpose and departure time. Further, two classes of cyclists were identified which display distinct preferences for specific route aspects (§6.2.4-§6.2.5). The dashboard therefore also allows the user to distinguish between the scores of class 1 and class 2. This can be done in the top-right menu. All charts and figures are refreshed upon making changes. This provides insight into the dynamics of the preferences.



Filter based on score or road type.

One can filter for the network segments which perform below, at or above average bij clicking the respective slice in the pie-chart in the bottom pane. The map will then only show those network segments and all charts and figures are recalculated. The same holds for the pie-chart for road types. Click the same slice to cancel the filter. Hold [ctrl] to select multiple slices.



Filter based on positive / negative aspects.

One can click on one of the positive or negative aspects in the bottom pane to display only those segments to which it is applicable. Click the same aspect again to cancel the filter.