MASTER

Cycling performance in Dutch cities and the role of local cycling policies
A multilevel modelling approach

Gelissen, Martijn W.J.

Award date:
2026

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Built Environment

Cycling performance in Dutch cities and the role of local cycling policies: A multilevel modelling approach

by

M.W.J. Gelissen

MSC THESIS

Graduation		**Committee**	
Program:	Built Environment	Member 1 (chair):	dr. Q. Han
Date of defense:	December 17, 2025	Member 2:	ir. D.T.L. Andreoli
Student ID:	1360892	Member 3:	dr. ing. P.J.H.J. van der Waerden
Study load (ECTS):	45	Advisory member 1:	J. Kamminga
Track:	Urban Systems and Real Estate		

The research of this thesis has been carried out in collaboration with the Fietsersbond.

Summary

The importance of the bicycle in shaping urban sustainable mobility systems is often overlooked. In order to create healthier, cleaner, and more livable cities, a shift away from car dominance is needed, and subsequently, the bicycle should get a more prominent role in the mobility system. Cycling is associated with increased health benefits, reduces emissions, and requires less space than cars, an increasingly important factor with the space scarcity in increasingly dense areas in urban situations.

In the Netherlands, the bicycle already has a more prominent role in their mobility system as compared to other countries. It is a high-cycling country with cycling flavoring national policies over the last decades. Yet substantial differences are present in cycling performance between Dutch cities, explained by the differences in bicycle use. This thesis exposes the underlying determinants of these differences and dives into the role of local cycling policies involved. The research addresses a notable gap in the literature as existing literature often focuses on cross-country comparisons or single-city case studies, while providing an intra-country comparison analysis enables insights into location specific behavior within the same national contexts and national policies. It focuses thus on the local aspects of cycling, as cycling is a very location-specific phenomenon. Additionally, as the role of local cycling policy is underexposed, this study not only addresses this gap, but simultaneously provides societal relevance by providing municipalities with valuable insights into how they can increase their cycling performance with hands-on recommendations.

Following the insights from the literature review, it became clear that the determinants for cycling performance act on different levels, which are trip-, person-, environment- and city-level. This study conducts a multilevel modelling approach. This modelling technique allows for in-depth insights into not only the effect of all determinants acting on these levels, but also includes cross-effects between the levels. The study makes use of a mixed method technique in which existing quantitative data on travel behavior at the trip and person level (ODiN) has been combined with environmental level data from CBS and infrastructural data from the Fietsersbond. This has been combined with insights from 16 semi-structured interviews with experts on the role of local cycling policies in several city-specific contexts.

Key findings reveal that there is a crucial role of local authorities in shaping cycling performance, as several hardware and orgware measures appear to be of importance. Providing higher concentrations of separated cycling infrastructure, mixed-use urban forms, and increasing transit accessibility are all measures that increase cycling performance. Furthermore, regarding the role of local policy on cycling performance, one of the orgware measures included in the model seems particularly important, which is ensuring a stable organizational structure. This includes that municipalities should focus on formulating clear and ambitious policy goals and associated concrete implementation plans. Besides that, they should ensure structural financial resources. A key underlying condition of this is the presence of a consistent policy context, both in organizational stability and regarding politics. Besides the statistical model, the interviews showed that there is a mutual willingness for a constructive collaboration between municipalities and local departments of the Fietsersbond. The underlying condition to led these collaborations flourish is transparent communication.

The recommendations for municipalities are that they should invest in not only the above-mentioned hardware measures, but should simultaneously ensure a stable organizational structure with clear policy goals, concrete implementation plans, structural financial budgets, and a consistent policy environment. These measures can significantly improve cycling performance and contribute to a sustainable urban mobility system. In conclusion, this research demonstrates that while individual and trip-level factors dominate cycling behavior, the role of the municipality cannot be underestimated, as several hardware measures and a stable organizational structure play a crucial role in enabling more sustainable mobility systems. These insights give municipalities hands-on measures to work on, building further on creating healthier and sustainable cities.

Preface

This thesis marks the final step in my academic journey. During this period, I became increasingly interested in urban design, especially the domains linked to mobility. I've always been interested in how data can explain (mobility) patterns and how these insights can be used in order to influence systems and present dynamics. This graduation research gave me the chance to dive deeper into the domain in which data and policy meet each other in order to support more sustainable mobility systems. An urgent challenge, keeping in mind the sustainability concerns present today.

The results of my research show that local authorities have a clear role in shaping these sustainable mobility systems by giving the bicycle a more prominent role in their urban mobility system. This study shows that it not only requires investments in several hardware measures, but also states that a stable organizational structure is crucial for allowing cycling to flourish.

I would like to thank my graduation committee for their guidance throughout this research. First, I would like to thank Qi Han for being my chairwoman. Furthermore, I would like to thank my supervisors, Dennis Andreoli and Peter van der Waerden, for their valuable feedback and constructive advice throughout the process. Additionally, as this research was conducted in collaboration with the Fietsersbond, where I had an internship, I would like to thank the Fietsersbond for this opportunity. I am especially grateful to Jaap Kamminga for the excellent guidance and the opportunities given to work with all sorts of data from the Fietsersbond, even beyond this research scope. The direct exposure to a professional environment centered on cycling provided me with valuable insights and resulted in the outcome of this thesis. Finally, I would like to thank all the interview participants, both municipal and local department of the Fietsersbond representatives, as without their insights, this study could not have been conducted. The contributions were essential for understanding the dynamics creating by local cycling policies.

Contents

Chapter 1 - Introduction

1.1 Problem definition

Active modes of transportation like cycling are widely recognized as a more sustainable alternative to motorized vehicles and are also associated with health benefits (de Hartog et al., 2010; Mueller et al., 2015; Nijland, 2017; KiM, 2023). Multiple studies (e.g., Fishman et al., 2015; Götschi et al., 2015; de Hartog et al., 2010; Mueller et al., 2015; Nijland, 2017) emphasize that switching from car use to cycling for short trips, trips up to 7.5 kilometres, leads to substantial individual health benefits. These benefits lead to increased life expectancy and significantly outweigh the additional risks associated with cycling, including fatal traffic accidents and air pollution exposure. Besides the health benefits on the individual level, research points out the public benefits associated with cycling. A study by Fishman et al. (2015) showed that investments in cycling-friendly infrastructure result in long-term economic gains, as annual health benefits substantially outscore the annual investments made in infrastructure in the long term.

Additionally, cycling plays a crucial role in the urban mobility transition, being an active mode of transportation, for which a central prominent role is envisioned in for example the European transport policy (European Union, 2024). A key aspect of this transition is decarbonising transport by shifting away from motorized vehicles towards public transportation and active mobility modes such as walking and cycling in order to limit environmental burden (European Union, 2024). Cycling contributes significantly to this goal, with estimated emission savings of 133 g CO_2, 0.21 g NOx, and 0.02g PM_{10} per passenger kilometre travelled (KiM, 2023).

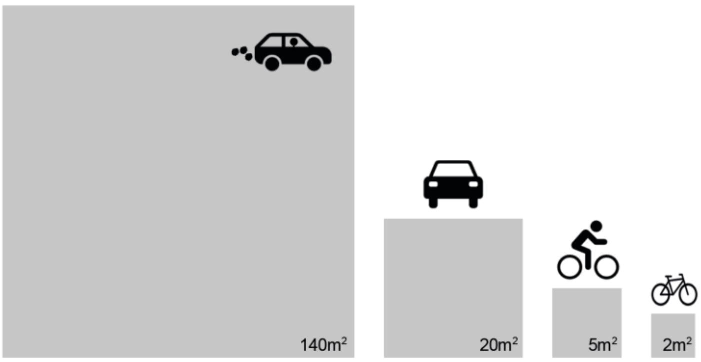

Figure 1 – Space utilization by cars and bicycle (moving and parked)

Furthermore, cycling requires substantially less urban space compared to cars, both in terms of road usage while driving and parking needs (Gemeente Amsterdam, 2017).Figure 1 illustrates that a moving car (at 50km/h) occupies approximately 140m^2 of space and around 20m^2 when parked, a bicycle at 15km/h requires only about 5m^2 while moving and only 2m^2 when parked. Given that cars are parked for approximately 95-97% of the time (Shoup, 1997; Bates & Leibling, 2012), a switch from the car to the bicycle could significantly free up valuable urban space in increasingly denser cities. This issue of spatial inefficiency of cars has long been recognized. Already in 1979, the "Ruimtegebruik" poster by the Eerste Enige Echte Nederlandse Wielrijdersbond (ENWB) gained international attention for visually demonstrating the contrast in space requirements between different transportation modes, especially exposing the out-of-proportion space consumption of the car (Slütter, 2020; Figure 2). This further illustrates the importance of facilitating a shift from car use to cycling. Considering the substantial benefits outlined above, related to individual health, sustainability, and spatial efficiency, the bicycle deserves a more central position within the current mobility system, in order to realise a more sustainable transportation system.

Figure 2: poster Ruimtegebruik ENWB (Slütter, 2020).

A country in which cycling already has a central role in the transport system is the Netherlands, where more than 25% of all trips are made by bicycle, which is the highest share anywhere in the world. This, in combination with having embraced cycling-oriented policies for several decades, has resulted in the Netherlands being internationally often seen as a cycling role model (Pucher and Buehler, 2008a; Pucher and Buehler, 2008b; Schepers et al., 2015). However, even within the Netherlands, there is still significant variation in cycling modal share between Dutch cities. While cities like Utrecht (42%) and Groningen (44%) achieve high cycling rates for inner-city trips, others, such as Rotterdam (26%), lag behind (KiM, 2023). Since national rules and regulations apply in all of these cities, the differences in cycling performance as explained above must be explained by local contexts.

Some literature included the local context as determinants for cycling performance explained as bike use (e.g., Heinen et al., 2009; Chen et al., 2017; Yang et al., 2019; Charreire et al., 2021; Zhou et al., 2023; Wu et al., 2024), by looking at aspects ranging from demographic variables and characteristics of the built environmental factors to infrastructural elements. Alternatively, other studies (e.g., Uijtdewilligen et al., 2024; Li et al., 2012; Berghoefer and Vollrath, 2023) used the same local factors but explained cycling performance as cycling preferences and (safety) perceptions. In this study, cycling performance is defined as the modal share of trips made by bicycle. This is the most commonly used indicator and gives the most comprehensive overview, as it illustrates the importance in the overall transport system (Rietveld & Daniel, 2004; Harms et al., 2015).

Despite these contributions, the largest body of this literature tends to focus on either cross-country or multinational regional comparisons (Pucher & Buehler, 2008a, Pucher & Buehler, 2008b; Charreire et al., 2021; Goel et al., 2021) or single case studies (Li et al., 2012; Aldred et al., 2015; Goel & Mohan, 2020; Wu et al., 2024). While these single-city case studies offer detailed insight into local conditions, they lack transferability due to the absence of geospatial comparative dimensions. On the other hand, international comparisons are useful for understanding differences between countries, they generalize effects on cycling on the country level, without accounting for the effects of the local context. Especially, a lack of research into the impact of local mobility policy is present, which is an important factor as differences in cycling performance appear at a local level suggesting need for local policy interventions (Harms et al., 2025). All of this together creates the need for an empirical intra-country comparison explaining the effect of cycling determinants, including local mobility policy effects.

1.2 research questions, objective & scope

Therefore, the goal of this study is to perform an extensive intra-country comparison of cycling performance across the ten largest cities in the Netherlands. This focus allows for an in-depth analysis while ensuring a manageable scope. By identifying key determinants, with special attention being paid to local policy effectiveness, this research aims to explain why some cities have higher bicycle use than others. As each city has its own local context and corresponding transportation policies, taking into consideration these local differences and policies is vital for making targeted and effective interventions. The findings will provide municipalities with insights into how cycling performance in their city can be improved in the most effective way. The research was conducted in collaboration with the Fietsersbond, where the author had an internship during the research period. This provided direct exposure to a professional environment centred on cycling.

This will be done on the basis of the following research questions:
"What factors explain the difference in cycling performance across Dutch cities, and to what extent do local cycling policies contribute to these differences?"

Subquestions:

- To what extent do trip-, personal-, and household-characteristics influence cycling performance?
- To what extent do features of the built environment and transport system affect cycling performance?
- How can the impact of local cycling policies on cycling performance be assessed, and what is their effect?

1.3 Relevance

Understanding the determinants of cycling performance and especially the role of local policy is essential for developing targeted and effective policies that promote cycling as a primary mode of urban transport, ultimately contributing to more sustainable and healthier cities. Given the substantial individual and societal benefits of cycling, including health improvements, reduced emissions, and more efficient use of urban space, insights from this research contribute to helping Dutch cities to optimize their cycling policies. This subsequently enables a more sustainable mobility system with a prominent role for the bicycle. Scientifically, this study contributes by addressing the lack of intra-country comparative research on cycling performance, in which the role of local cycling policies is largely underexplored.

1.4 Reading guide

Figure 3 presents the overall research design applied in this study. The research process begins with a literature review on travel behaviour, with special attention to travel mode choice. Furthermore, the determinants of bicycle use and the influence of policy measures are discussed. The insights derived from the literature review are the basis for the conceptual framework, which serves as a guideline for data collection. Based on this framework, variables corresponding to the research sub-questions are gathered and integrated into a comprehensive dataset. This dataset is subsequently analysed to identify patterns and relationships relevant to cycling performance. The analytical approach used for this research is explained in detail in Chapter 3. The results of this analysis form the basis for the conclusions presented in the final chapter.

This report is structured as follows. This chapter provides an introduction to the study and provides a problem definition with the associated societal and scientific relevance. Chapter 2 presents the literature review, providing existing literature on travel behaviour, the determinants for cycling performance, and the role of local cycling policies. Chapter 3 describes the methodology used for assessing the determinants and the role of local cycling policy, it also includes the analytical approach and data collection. Chapter 4 presents the empirical results, which are discussed in Chapter 5. Chapter 6 addresses the limitations of the study, while Chapter 7 offers recommendations for policymakers and directions for future research. Finally, Chapter 8 presents the main conclusions.

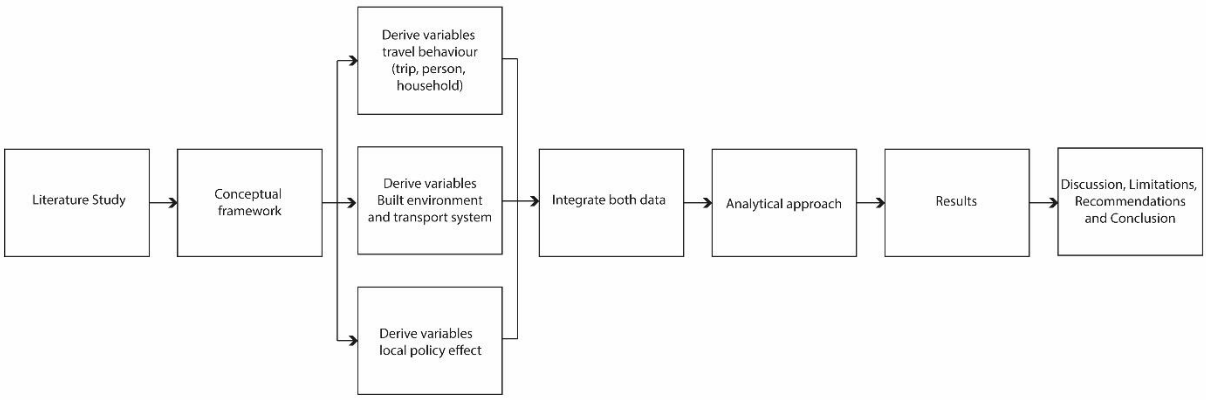

Figure 3 - Research design scheme

Chapter 2 – Literature review

The literature review begins with a general introduction to travel behaviour research, which is essential for understanding the dynamics creating the changes in cycling performance. It then explores the key determinants of cycling, with particular emphasis on the role of local policy in this, as effective policy interventions are crucial for improving cycling performance.

2.1 Travel behaviour research

In order to stimulate the urban mobility transition by enabling a shift away from motorized vehicles towards active modes of transportation such as cycling, it is essential to understand the underlying determinants of travel mode choices. This falls within the broader field of travel behaviour research, which is concerned with individuals' movement from their reference location (mostly home location) for any purpose (Axhausen, 2007). This research domain is well studied.

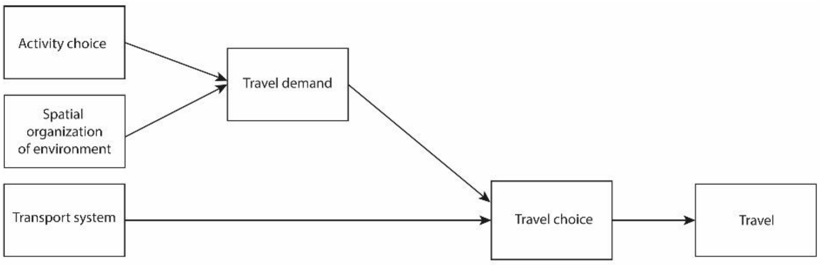

Figure 4: Determinants of travel model by Gärling (2005).

According to Gärling (2005), planning and executing trips involves a sequence of decisions that together shape individual travel behaviour. This means that different personal attitudes, demands, and needs are in place. Gärling (2005) argues that people make choices in chronological order on activity, destination, travel mode, and departure time (Figure 4). The most important element in this process model is the activity choice (Gärling 2005), because this triggers the need for a trip, as locations of activities are often spatially separated (Van Acker et al., 2010). The spatial organization of the environment thus determines where the next activity will take place and how far individuals must travel. In Garling's conceptual model, this is referred to as the travel demand for a certain trip. Gärling (2005) continues that the travel demand in combination with the current transportation system will determine the travel choice, in which the mode and time are chosen. These decisions are based on finding an optimal balance that suits the individual the best, taking into account factors such as travel time, costs, and personal preferences.

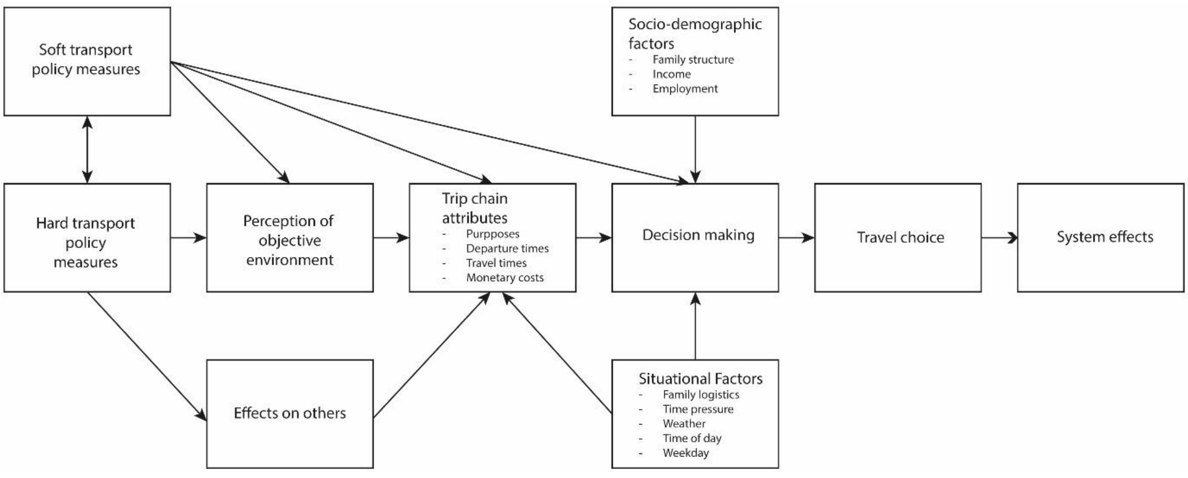

Figure 5: conceptual framework travel behaviour model Bamberg et al. (2011)

In a later conceptual model created by Bamberg et al. (2011), the decision-making process for travel choice is related to trip chain attributes, which is a result from the perception of the objective environment. This includes for example the spatial distribution of activities, the quality of activity facilitations, and the transport system (Figure 5). This perception influences so-called trip chain attributes, which correspond to the choice influences mentioned in the previous model of Gärling (2005). Thing links to the purposes departure and travel time and monetary costs. On top of that, (Bamberg et al., 2011) argue that socio-demographic factors such as income, family structure, and situational factors like family logistics, weather, and time pressure as well influence the decision0making process, leading to an interwoven and complex decision-making process.

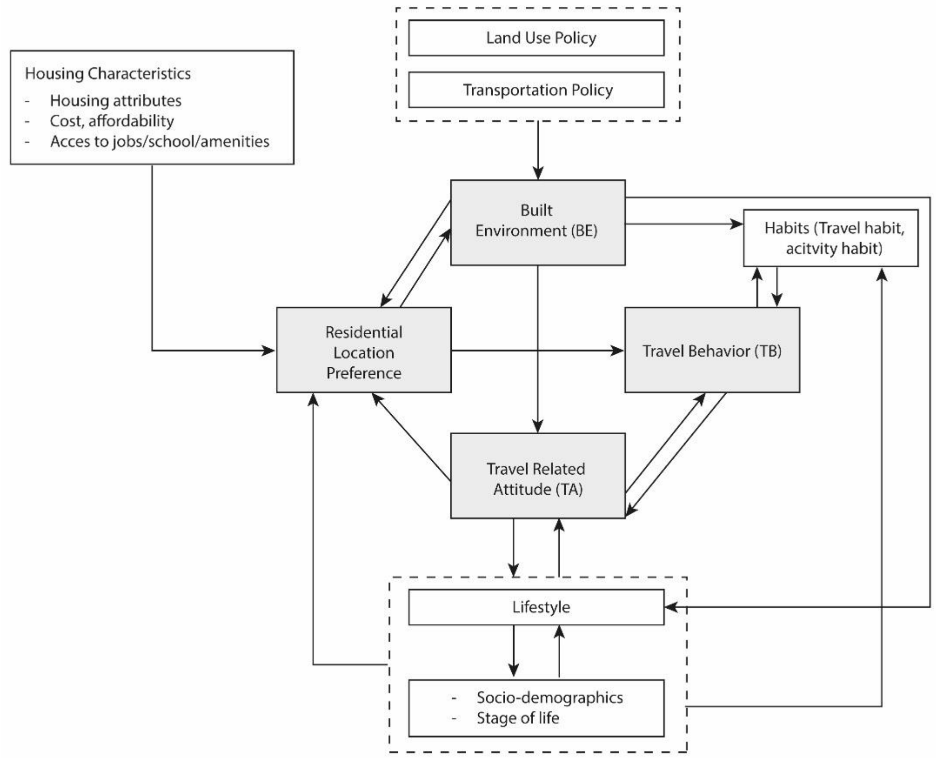

Figure 6: Conceptual model by Rahman and Sciara (2022) illustrating the relationships between built environment, travel behavior, travel attitudes and residential preference

This complexity is further enhanced by Rahman and Sciara (2022), who created a conceptual model (Figure 6) showing the complicated relationships between the built environment and travel attitudes and behaviour based on psychological theories. This model builds further on the previous models as it includes the effect of travel attitudes and travel habits. This changes the perception of a static process to a continuous process in which travel attitudes can change over time and thus also can lead to changes in travel behaviour through different psychological processes. Rahman and Sciara (2022) explain that travel attitudes are influenced by cognitive, normative, and behavioural triggers, of which especially the first one is important as gaining new information leads to attitude change, as explained by Anderson's Information Integration Theory (1971). Normative triggers include personal norms, moral beliefs and behavioral triggers which can influence travel attitude in such a way that positive experiences will be repeated. Individuals tend to minimize feelings of discomfort and repeat positive experiences (Festinger, 1957; Rahman and Sciara, 2022).

Besides the above-mentioned triggers leading to reasoned behaviour, habits create unreasoned behaviour. They are a result of routinized unconscious decisional-making. Automatization processes come into play where previous actions directly affects later travel behaviour. When habits are strong, the effect of attitude on behaviour becomes weak according to Triandis's Habit Theory (1977). Disrupting habits is studied in Habit Discontinuity Theory, in which changing the context can result in changes in habits. Multiple studies have shown that changing the built environment can act as a change in context as these studies assessed the effect of travel habits changes amongst movers positively (Bamberg, 2006; Verplanken et al, 2008; Walker et al, 2015; Haggar et al., 2019).

To conclude, explaining the dynamics of travel behaviour includes complex mechanisms that require an individual approach as psychological processes and perceptions come into play. The built environment, attitudes, habits, and travel behaviours are all in particular ways interacting with each other in such a way that targeted policies leading to particular changes in for example the built environment or promotional changes to attitudes could stimulate more sustainable travel behaviour.

2.1.1 Travel mode choice

A crucial element in travel behaviour research is the travel mode choice as it has a prominent role in all the previously discussed travel behaviour models. Furthermore, the travel mode choice is particularly relevant for understanding bicycle use, as it directly involves the selection of a transport mode, with the bicycle being one of the possible choices. Diving deeper into the literature on travel mode choices will lead to a comprehensive understanding of the factors that influence cycling rates.

The choice between transportation modes is in academic literature mostly explained by generating utilities for each transportation alternative in a choice set. Subsequently, the alternative with the highest utility is chosen (Zhang et al., 2023). Individuals choose the transportation mode that generates the highest utility. This utility depends on the contribution of attributes mentioned previously in the travel behaviour section. Additionally, research particularly focused on travel mode choice studies has tried to create additional conceptual models determining the factors that explain why individuals choose a certain transportation mode. Understanding the factors for bicycle usage is crucial for making informed policy decisions in order to promote cycling. This creates an explanatory model for bicycle usage. According to Stradling (2011), travel mode choice behaviour is linked to personal characteristics, household characteristics, environment characteristics, attitudes, journey purpose, and trip origin/destinations. Linking these determinants to current bicycle research creates a framework that explains bicycle usage.

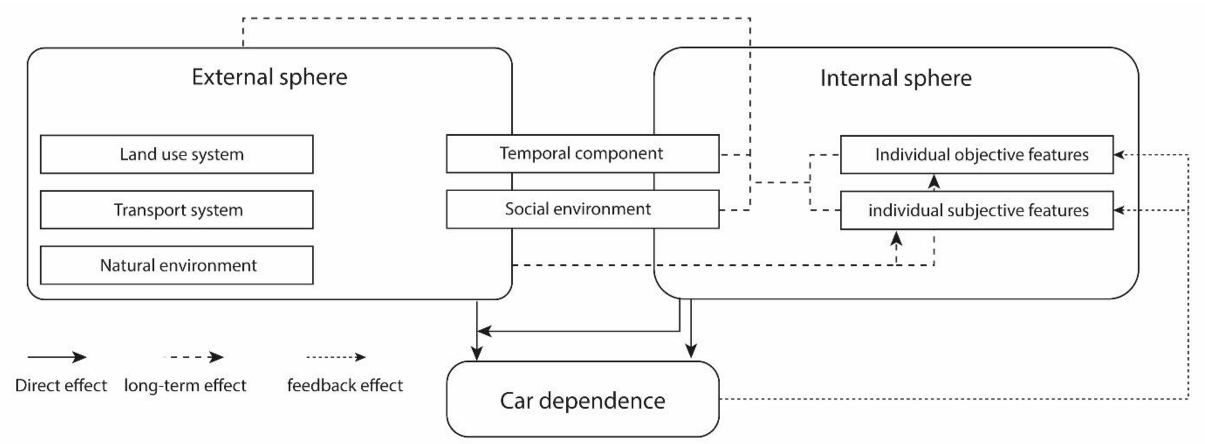

Figure 7: Conceptual model of car dependence by Cremer-Schulte et al. (2024)

Another notable approach is the conceptual model made by Cremer-Schulte et al. (2024), which explains current car dependence (Figure 7). This conceptual model has strong similarities with the determinants explained by Stradling (2011). However, Cremer-Schulte et al. (2024) distinguish it's determinants between an internal and external sphere. The internal sphere consists of individual objective features and individual subjective features, which show great similarities with the personal and household characteristics (objective) and attitudes (subjective) described in Strandling's (2011) conceptual model. The external sphere includes the land use system, transport system and natural environment. Similar to

Gärling (2005), the land use system and transport system are seen as separate components. The natural environment is explained as the "topological conditions, seasonal effects, and weather conditions" (Cremer-Schulte et al., 2024, p.182), referring to factors such as terrain features and climate aspects like rainfall. On top of that, a temporal component and a social environment component are added, acting between the external and internal spheres. While this conceptual model focuses on the dynamics of car dependence, elements from it can be integrated with the travel mode choice model by Stradling (2011).

2.1.2 Conclusion travel behaviour research models

In conclusion, both travel behaviour and travel mode choice research highlight the complex dynamics between individual characteristics, psychological processes, the social and physical environment in making travel decisions. In all discussed models, activity choice leads to a travel demand by individuals, who subsequently evaluate the utilities per transportation mode. This utility is an result of an evaluation of the environment (both social and physical) through both objective and subjective characteristics of the individual, such as socio-demographics, perceptions, and attitudes. This ultimately results in a travel mode choice, with the mode generating the highest utility being selected. Habitual behaviour can be influenced in new information, creating possible changes in attitudes. Here lies a key role for policies by affecting the utility of transportation modes experienced by individuals. This can be done through alterations to the built environment (hard measures) or by directly influencing attitudes and behaviours through soft measures. The remainder of the literature will dive into the different components which influence cycling performance, which are the trip-, person-, and environment characteristics. Furthermore, the tool to influence these characteristics, which is policy, is discussed subsequently.

2.2 Trip characteristics

First of all, what characterizes a trip taken by bicycle. In order to answer this question, it is crucial to keep in mind that local and national contexts significantly influence cycling performance (Pucher and Buehler 2008a; Heinen et al., 2009; Goel et al., 2021). In the Netherlands for example, cycling takes up a significant amount of trips of approximately 28%, of which roughly one third are made by the e-bike (KiM, 2023). On the other hand, cycling takes up only 2.1% of all trips being made in England (KiM, 2023). This highlights the bicycle's importance in some national transport systems, opposite to the negligence of the bicycle in transport systems elsewhere.

An important element that characterizes a bicycle trip and is an exception to the above statement is trip distance (Chen et al., 2017; Goel et al., 2021). Goel et al. (2021) showed that low cycling countries and cities showed the same distance distribution ratios as high cycling countries and cities, meaning that distance could be generalized across different international settings. The likelihood of cycling is especially high at lower distances (< 5km). The distance distribution itself shows that cycling is especially of importance at shorter distances (Pucher and Buehler, 2008a; Goel et al., 2021; KiM, 2023). This has to do with greater generalized travel costs being created by longer distances by the bicycle in comparison to other transportation modes such as motorized vehicles (Heinen et al., 2009). Differences in generalized travel costs between different types of bicycles are present as e-bike trips have an average 70% longer distance than bicycle trips made by the conventional bicycle (KiM, 2023). This can be explained by a higher comfort due to less physical effort needed for cycling on an e-bike (Huurman et al., 2024).

Additionally, when analysing the distribution of distance travelled by transport mode in the Netherlands, cycling accounts for 10% of the total distance travelled, which is a substantially smaller share than the share of trips made by bicycle (28%) in the Netherlands (KiM, 2023). Additionally, while active modes of transportation make up 50% of all trips, they only represent 14% of total distance travelled (KiM, 2023), reflecting again their suitability for relatively shorter journeys. In line with the findings of Goel et al. (2021) is the trip distance distribution in the Netherlands, as bicycle use peaks between 0.5 and 5 kilometers, taking up 32-44% of the trips made within these distance boundaries (KiM, 2023). Overall, the distribution of bicycle trips by distance has the shape of a right-skewed normal distribution, meaning that the influence of the bicycle gradually diminishes as trip distance increases. This pattern emphasizes the bicycle's potential as the main urban transport mode, while most distances in cities are shorter.

Another important element characterizing a bicycle trip is the purpose. Internationally, the bicycle is mostly used for commuting purposes, especially in urban areas. However, in high cycling countries, such as the Netherlands, the bicycle is more frequently used for other purposes such as leisure (Goel et al., 2021). Different trip purposes create different behaviours and preferences. Berghoefer and Vollrath (2022) for example showed that commuter cyclists are more concerned with deterrent aspects in route choices and prefer efficiency. Different purposes also result in different willingness to travel distances, as can be concluded from the Dutch cycling numbers in which average cycling distance per trip purpose heavily differentiates (KiM, 2023). The longest bicycle trips are made for leisure activities, whereas the shortest are linked to shopping. Furthermore, 28% of the Dutch commute trips are made by bicycle, with the bicycle especially dominating the shorter commute trips (<5km) with a share of 53% of the trips made (KiM, 2023). This illustrates the importance of compact urban areas in which residential and work locations are in close proximity, in order to ensure the bicycle as the main urban mode of transportation, as a concept by Pucher and Buehler (2008a) and Heinen et al. (2009). Again, as distance increases, the influence of the bicycle in the modal share decreases. However, due to the already mentioned lower generalized travel costs for the e-bike as compared to the conventional bicycle, leading to longer average distance travelled, could extend the influence of the bicycle.

All in all, travel distance and travel purpose are the most interesting trip characteristics to consider as trips made by bicycle are of relatively shorter distances and differ per trip purpose. Additionally, the travel purposes can influence the distances, as leisure purpose trips for example result on average in longer distances than other trip purposes. Knowing the trip characteristics of cycling can help make targeted policies for stimulating bicycle usage. Cities could for example aim for an urban design which provides locations of activities within the range of average distances travelled by bicycle by that particular activity purpose.

2.3 Person characteristics

As the previous section describes how the bicycle is especially suitable for short distances and how the trip purpose explains variation in trip distances and the modal share by bicycle, this part focuses on the characteristics of bicycle users. It explores who is riding the two-wheeler and what characterizes this individual, resulting in a user profile of the cyclist.

A key variable in the user profile is gender. On an international level, men are overrepresented among cyclists. However, in countries where cycling is more common such as the Netherlands, Denmark, and Germany, gender equality is more evident (Pucher & Buehler, 2008a; Pucher & Buehler, 2008b; Aldred et al., 2015; Goel et al., 2021). Goel et al. (2021) found that the variation in gender distribution among countries is related to the overall bicycle mode share. In places where the mode share is higher (>7%), gender equality or even overrepresentation of females on the bicycle can be observed, as illustrated in the case of Japan. In contrast, below this percentage gender inequality is more dominant, with ratios between males and females from 3:1 in the USA and UK and even 7:1 in certain cities in Argentina and Brazil (Goel et al., 2021). Contrary to these statements are the results from Aldred et al. (2015), who showed that biking rate increases in local contexts in the UK didn't result in less gender disparities in those particular regions, suggesting that there is more needed than only promoting a mode share increase. As alternations in established behaviour take time, policies should focus on promoting campaigns for underrepresented groups on the bicycle instead in order to accelerate the process of change (Aldred et al., 2015).

Studies have shown that women tend to make shorter trips and are less likely to cycle for commuting purposes than men (Krizek et al., 2005; Garrard et al., 2008). The same studies suggest that safer infrastructure could potentially increase the willingness of women to cycle longer distances and that they would consider taking the bicycle for commuting purposes more often. Therefore, policies should prioritize safer cycling infrastructure by for example providing separated cycling lanes and sufficient street lighting (Krizek et al., 2005; Garrard et al., 2008).

Besides gender, multiple studies have focused on the role of age in cycling behaviour (Pucher & Buehler, 2008a; Pucher & Buehler, 2008b; Charreire et al., 2021; Goel et al., 2021). Large international differences are identified again between countries and the relationship between age and bicycle usage is often found to be inconsistent (Charreire et al., 2021; Goel et al., 2021). Generally, in countries with a higher bicycle mode share, cycling among children (<18 years) tends to be more common (Goel et al., 2021). In the Netherlands, for example, nearly half of trips (48%) made by children (6-18 years) are made by bicycle (KiM, 2023). Furthermore, in the Netherlands, the use of bicycles is well represented across all ages (Pucher & Buehler, 2008a; Pucher & Buehler, 2008b), with a substantial usage rate drop after adolescence, but increasing again at the age of 60 years (KiM, 2023). A further increase is even observed after 65+ years (Pucher & Buehler, 2008a; Pucher & Buehler, 2008b, KiM 2023). This trend illustrates the importance of cycling for recreational purposes in the Netherlands as the average retirement age is 65.9 years (CBS, 2024a). This leads to a significant decrease in work-related trips after that age. This all implies that cities with an overrepresentation of underaged and older residents would imply higher cycling rates.

Another notable variable is the occupation of the individual, especially attending an educational institution. In other words, being a student has shown a positive relation to cycling in the past in the USA (Nelson & Allen, 1997), and in Edinburgh, Scotland (Ryley, 2006), and more recently in Sydney, Australia (Wu et al., 2024). This is in line with bicycle mode share for educational trips in the Netherlands, which accounts for than half of all educational trips made in 2023 (KiM, 2023). It is thus expected, that cities with relatively more students would produce higher cycling mode shares. Another result from Wu et al. (2024) included the mixed impact on commute cycling for the variable income. Similarly, Dill and Carr (2003) didn't find a significant relationship between income and bicycle usage for commute purposes. Heinen et al. (2009) concluded in their literature review that results were too diverse to make statements on. Building on that, Pucher & Buehler (2008a) and Pucher & Buehler (2008b) point out that cycling in the Netherlands, Denmark, and Germany is for everyone regardless of income level, as well as age as explained previously. As for the Dutch context, cycling is deeply rooted in daily life routines, and the bicycle can be seen as used by everyone. The ownership of 1.3 bicycles per person in the Netherlands (Fietsersbond, 2019), in combination with the significant modal share of the bicycle among different income groups, strengthens this statement (Pucher and Buehler, 2008b).

Additionally, multiple studies have established mixed results on the influence of immigration backgrounds and cultural backgrounds on cycling rates (Smart, 2010; Haustein, 2019; Faber et al., 2023; Wu et al., 2024). The variable of migration background touches upon behavioural attitudes by individuals in different societal settings, as they experience the same transport system as the native population but differ from travel habits created in their origin country (Haustein et al., 2020). Place-specific aspects will likely have an effect on the mobility patterns of individuals (Faber et al. 2023). Additionally, cultural backgrounds created behaviour and attitudes over generations, even for those already living in another country (Haustein et al., 2020). Yet the results are mixed. Where Smart (2010) argues that people with an immigration background in the USA, especially those with an eastern Asian background, cycle more than Americans without an immigration background. This is contrary to the findings from Wu et al. (2024), who argue a negative association between an eastern Asian background and cycling rates in Sydney, Australia. Even within the national context, there is debate, as Haustein (2019) argues that cycling rates in the Netherlands among individuals with an immigration background are lower than the travel rates of individuals without an migration background. Contrary to the results from Faber et al. (2023), who examined especially the differences between Dutch-born individuals and high-income immigrants, referred to as expats, as expats can solely base their modal choice on preferences rather than monetary feasibility. Their study showed a higher usage of the bicycle by the expats than by Dutch-born individuals.

The results from KiM (2023) reveal that the migration background type (Western vs non-Western) influences the statements made above, as Western background immigrants have somewhat the same cycling mode share as the Dutch population without a migration background (29% vs 28% respectively), and the non-Western background immigrants only make 23% of their trips by bicycle. Furthermore, it illustrates that between 2010 and 2022, the increase in cycling mode share among individuals with a

Western migration background was greater than that of individuals with a Dutch background and even surpassed the cycling mode share of the Dutch population at the end of this period. Spatially, this means that regions composed of different ethnic groups will result in different cycling rates between these places.

Another important aspect is the accessibility of the individual to transportation modes, especially the accessibility to the motorized vehicle as throughout time multiple studies showed that having access to a car has (strong) negative associations on bicycle usage (Kitamura et al., 1997; Buehler et al., 2011; Cervero and Duncan, 2003; Dill and Carr, 2003; Parkin et al., 2008; Heinen et al., 2010; Semenescu & Coca, 2022). Furthermore, individual cycling rates even drop by 10% if an individual is in possession of a driver's license (KiM, 2023). The quality and quantity level of public transport will also play a role in cycling performance. On the one hand, public transport can promote cycling as it can be used for first- and last-mile operations towards public transportation, especially train and metro usage (Pucher & Buehler, 2008a; Zhao & Li, 2017, KiM, 2023). On the other hand, it could be argued that a well-functioning public transport system could lead to a decrease in bicycle usage as cycling trips have the potential to be replaced by public transport trips.

In summary, the existing literature clearly indicates the influence of personal characteristics on cycling performance, although these effects vary across contexts. The main conclusion is that determinants differ among settings in which cycling is established in daily culture or not. The Netherlands is one of these high-cycling countries, and in these countries gender equality is observed. Furthermore, bicycle use is the highest among children and the elderly in these settings. The effect of income on remains mixed, and findings on origin and cultural background further enhance this complexity by also stating mixed results. Finally, access to alternative transport modes, especially linked to car ownership and driver's license ownership, strongly reduces bicycle use. On the other hand, increased public transport accessibility positively influences bicycle use. These findings underscore the integrated interplay between the bicycle and other modes of transport.

2.4 Environmental characteristics

This section discusses the environmental characteristics that influence bicycle usage. The section is separated into the natural and built environment. First, the natural environment will be discussed, which can be seen as the standard basic locational conditions with which the built environment, both the land use system and the transport system, has to deal with. These included landscape characteristics and weather and climate conditions, and differed across different spatial contexts. These factors have a stronger effect on cycling behaviour than on motorized vehicle transport behaviour (Heinen et al., 2009). Secondly, the factors effecting cycling behaviour from the built environment will be discussed. The built environment represents all the human-made buildings and infrastructure needed for human activities (Seyedrezaei et al, 2023). The three dimensions of density, diversity, and design, which explain the influence of the built environment on travel demand introduced by Cervero and Kockelman (1997), will be used to explain the effects of the built environment on cycling behaviour.

2.4.1 Natural environment

The environmental characteristics include the natural environment, which can be seen as the standard bare local conditions with which the built environment, both the land use system and the transport system, has to deal with. These include landscape characteristics and weather, climate conditions. These factors have a stronger effect on cycling behaviour than on motorized vehicle transport behaviour (Heinen et al., 2009).

First of all, the hilliness, or the slope, of the landscape has been found to be an important negative determinant of bicycle usage rates (Cervero & Duncan, 2003; Rodriguez & Joo, 2004; Parkin et al., 2008; Heinen et al., 2009; Chen et al., 2017; Zhou et al., 2023; Wu et al., 2024). The existence of a hilly terrain could result in discomfort among cyclists, due to the additional effort cyclists need to make (Rodriguez & Joo, 2004). Similar results are presented by Berghoefer and Vollrath (2023), resulting from their stated choice experiment, in which the highest impact on route choice influences turned out to be, next to the surface, the gradient of the route. Contrary to this are the results from Moudon et al. (2005), who have not found a significant effect of slopes on bicycle share. However, this study was performed on trips with a leisure purpose, potentially including cyclists who prefer additional physical activity during their leisure cycling trip.

Additionally, Gatersleben and Uzzel (2007) pointed out the importance of an attractive environment, and Blitz (2025) points out that an attractive local environment promotes regular bicycle use. Building on this, Zhou et al. (2023) proved that green and sky indexes were both positively correlated to cycling volumes, given a statistically significant explanation for what can be seen as an attractive environment for cyclists.

Furthermore, the effect of different seasons on cycling has widely been studied, resulting in lower cycling rates during the winter in comparison to other seasons (Bergström & Magnussen, 2003; Stinson & Bhat, 2004; Guo et al., 2007; Chen et al. 2017). Additionally, the extent to which the winters are fierce has an extra negative effect on cycling (Bergström and Magnussen, 2003; Heinen et al., 2009).

Going from seasons and the climate related effects towards to more day-to-day characteristics of the weather, aspects such as rainfall and winds seemed to be of interest for previous studies (Rietveld & Daniel, 2004; Heinen et al., 2009; Böcker et al., 2013). Rietveld & Daniel (2004) argue that wind has a stronger effect than rainfall. Furthermore, the results of the effect of rainfall are mixed (Heinen et al., 2009; Böcker et al., 2013), potentially deriving from the usage of different explanations of rainfall used in previous studies (Heinen et al., 2013).

Transferring all this information to the Netherlands, the scope of this study, the Netherlands forms an ideal natural environment with a relatively flat surface, mild sea climate, suitable for cycling (not too warm, not too cold), setting the base for high cycling rates. Yet, local differences in rainfall, winds (coastal areas vs more inland), and hilliness of terrain in cities can differ and potentially influence travel behaviour.

2.4.3 Infrastructure

The last dimension of the built environment is design and considers street network characteristics. It relates to all the elements of the infrastructure, such as street connectivity, pedestrian crossings and street widths (Ewing & Cervero, 2010). All of these elements from the transport system are of importance as specifically these features highly influence cycling behaviour. Especially the cycling infrastructure itself and its adjacent facilities are well studied. Research in this field generally follows two types of methodological approaches. The first consists of empirical research that investigates the impact of infrastructural elements on cycling performance using mostly modal share- or cycling rate data, census data or mobility survey data. These studies provide evidence on the effectiveness (or ineffectiveness) of specific infrastructural interventions. The second approach focuses more on individual perceptions and preferences, focusing on how people experience the infrastructural elements. These studies explore the effect of the infrastructural elements on aspects such as perceived safety, route or detour choice, and crowding (Vedel et al., 2017; Uijtdewilligen et al., 2024). This offers insights into how users objectively respond to infrastructural features. The results from both methodological approaches together form a more comprehensive view of the most important considerations to be taken into account for future infrastructural designs.

The most important aspect within the transport system is the cycling infrastructure itself, with especially the well proven high importance of separated bicycle lanes. Multiple studies have proven that separated lanes are associated with an increase in bicycle mode share (CPB, 2025; Dill & Carr, 2003). On top of that, they are also perceived as safer and more comfortable than other configurations, such as painted lanes or shared roads (Berghoefer & Vollrath, 2022; von Stülpnagel & Bining, 2022). These preferences hold across diverse populations and are particularly strong among risk-averse users, including women and less experienced cyclists.

Firstly, the effect of separated bicycle lanes on cycling numbers is well proven throughout the last decades and across different national contexts. Most recently, in the Netherlands, separated cycling lanes were found to create a 5 percentage point cycling modal share increase from 20% to 25% for commuting travel purposes (CPB, 2025). Similarly, Dill and Carr already proved in 2003 that even in car-centric environments like the United States, if cycling infrastructure (lanes and paths) is provided, commuter cycling rates will increase. On top of that, results from other studies also empirically prove the importance of the separated bicycle lanes in other contexts, such as in Australia and Europe (Berghoefer and Vollrath, 2008; Vedel et al., 2017; Wu et al., 2024).

The results from the studies focusing on user preferences and perceptions are much in line with the outcomes discussed above. Research shows that separated cycling infrastructure enhances perceived safety and is considered as one of the, if not the most important feature of the transport system (Parkin et al., 2007; Zhu et al., 2017; Chen et al., 2017; Berghoefer and Vollrath, 2023; Uijtdewilligen et al., 2024). Berghoefer and Vollrath (2023) found a clear hierarchy in bicycle road design for cyclists' preference in which the highest preference is for cycle paths with physical separation, followed by painted cycle lanes adjacent to the street, and then advisory lanes characterized by only dashed line marking. Their results showed that preferences not only influenced the willingness to cycle but also influenced the route choice. In making this route choice, next to motivational aspects, also deterrent aspects were found to be of importance, especially for commuting cyclists, as they prefer efficiency in their trip route.

However, the perceived benefits from the separated bicycle lines are less when they become overcrowded. Uijtdewilligen et al. (2024) and Li et al. (2012) proved that while cyclists do prefer separation from motorized vehicles, high crowding levels negatively affect this preference due to the difficulty of overtaking and reduced perceived safety due to an increased perception of the possibility of collision with other cyclists. This creates an important takeaway; stimulating cycling is not only about separating the cyclists from motorized vehicles, but also about taking into account the capacity of the biking infrastructure simultaneously (Uijtdewilligen et al., 2024).

Besides separation on street segment level, street connectivity in the network is also found to be an important aspect to positively affect cycling behaviour (Pucher & Buehler, 2008a; Yang et al., 2019). This touches upon making the cycling route as smooth as possible by a coherent network of for example separate bicycle lanes, priority bicycle traffic signals, and intersection modifications favouring cyclists (Pucher & Buehler, 2008a). Especially for commuting cyclists this efficiency is proven to be important in multiple studies (Yang et al., 2019; Uijtdewilligen et al., 2024). This positively influences the travel efficiency and accessibility by the bicycle in such a way that at a certain point it can compete with the car or even outperform the car in for example, in travel time. According to Pucher & Buehler (2008a), it is all about creating a low-stress environment, of which traffic calming measures such as speed reductions and bicycle priority streets play an important role.

Another infrastructural element found to be important is street lighting, which especially touches upon perceived safety. Krizek et al. (2005) found that during nighttime, street lighting is found to be important for especially women, who tend to be more sensitive to perceived safety during night. Additionally, the results from Krizek et al. (2005) showed that women put more emphasis on perceived safety related infrastructural elements than men do.

Besides the direct infrastructural elements along the route, the supporting facilities at the beginning or end points of the route are also of importance. Safe and convenient parking facilities at work or institutional locations for example have been found to be an important aspect in bicycle mode share choice (Pucher & Buehler, 2008a; Ricchetti et al., 2025). Besides parking facilities, a smooth transition to public transport can enhance cycling rates for first- and last-mile operations of the bicycle for public transport (Pucher & Buehler, 2008a; Harms et al., 2015; van Kuijk et al., 2022). Van Kuijk et al. (2022) argue that there is still potential for a better integration of shared bicycles and shared e-bikes for first- and last-mile operations to public transport.

2.4.4 Conclusion Environment

In summary, the environment can be divided into the natural and the built environment. Based on the literature, the Netherlands offers an ideal natural setting for cycling due to its relatively flat surface and mild sea climate (not too warm, not too cold), setting the base for high cycling rates all year round. Yet, local differences in rainfall, winds (coastal areas vs more inland), and hilliness of terrain in cities can differ and potentially influence travel behaviour.

Considering the built environment characteristics, from the density component, it can be concluded that higher population densities and compact urban forms usually support cycling by reducing average travel distances, especially in Western countries. Additionally, the main takeaway from the diversity component is to create mixed land-use patterns with multiple different facilities provided at shorter distances in order to increase cycling rates. This is an important conclusion to be taken into consideration on the governmental level, as political intervention by stimulating mixed land use development results in higher active modes of transportation.

Besides that, the infrastructure plays an important role too. However, specific infrastructural elements in the transport system have more effect on cycling behaviour than others do. Moreover, the results from the literature focusing more on objective indicators, such as cycling rates and mode shares, align with the results of the literature focusing on subjective matters, such as perceptions and preferences. It is clear that the separation of bicycle lanes is the most important element for cycling behaviour. Additionally, creating a smooth and low-stress environment by enhancing street connectivity (fewer intersections, but still ensuring high connectivity), traffic calming, and coherence networks boosts cyclists' travel behaviour. These aspects should be taken into account when performing the analysis. Additionally, these should be the main elements to be taken into consideration for policy interventions as they seem to have the biggest impact on cycling numbers and cyclists' satisfaction.

2.5 Policy

All previously discussed determinants represent the contextual factors that create a local specific framework that explains the cycling performance. Overarching these determinants is the policy domain, which operates at a higher level and can shape or adjust the local context. Through policy interventions, local authorities can influence determinants at the lower levels (the other determinants discussed in the previous sections), which enables them to indirectly affect cycling behaviour. Positively influencing this cycling performance requires a well-though location-specific policy plan. Academic literature widely confirms that policy interventions play a crucial role in changing cycling behaviour (Rietveld & Daniel, 2004; Pucher and Buehler, 2008a; Harms et al., 2015; Pawluk De-Toledo, 2022). To understand how policy can be assessed and how it impacts cycling performance, this part of the literature review dives deeper into the existing literature on the local policy domain and its effectiveness. Understanding the role of policy is key to understanding how actual travel behaviour can be changed towards more cycling in the most effective way.

2.4.2 Built environment

The first dimension from Cervero and Kockelman (1997) of the effect of the built environment on travel demand is density, which refers to the concentration of different elements such as the population, jobs, and houses. First of all, population density is proven to have a positive effect on cycling rates (Zahran et al., 2008; Heinen et al, 2017; Zhou et al., 2023). It is believed that the higher the concentration of residents, the lower the average distances travelled are as compared to a more sprawled out society. Contrary to this are the results from Goel and Mohan (2020), who proved that population density did not have a significant effect on cycling rates in India. Their research resulted in a positive result in population size, rather than population density. Goel and Mohan (2020) explain their result by arguing that cities in India are already of high population density and that these are low income cities, in which apparently other mechanisms are at place. This explains again the importance of taking into account the local context when performing bicycle related research. Overall, the population density does have a positive effect on cycling numbers in higher-income Western countries (Goel et al., 2021). On the other hand, the effect of household and residential density on cycling numbers is mixed as multiple studies do find a positive relationship (Parkin et al., 2008; Pucher and Buehler, 2006; Guo et al., 2007), while others did not find significant associations between the two (Chen et al., 2017; Rodriguez & Joo, 2004). All of this is highly related to distances that needs to be covered during trips (Pucher & Buehler, 2006; Heinen et al., 2009). Land-use concepts associated with designs focused on compact cities or 15-minute cities and realizing higher density levels produce higher shares of non-motorized travel (Heinen et al., 2009). Within such urban areas, distances become shorter, which are more suitable for active modes of transportation (Pucher and Buehler 2006; Pucher and Buehler 2008b; Heinen et al., 2009; KiM, 2023).

In line with the effects of the density dimension are the thoughts behind diversity. It is all about creating shorter distances between activities in order to make the bicycle more favourable to use, as the bicycle is used the most for relatively short distances (Cervero & Kockelman, 1997). Multiple researchers (Cervero & Duncan, 2003; Pucher & Bucher, 2006; Pucher & Bucher, 2008a; Hankey et al., 2012; Chen et al., 2017; Heinen et al., 2009) agree that by creating a more diverse urban landscape or a higher mixture of functions, shorter distances are created, which makes cycling more favourable. An example of this can be seen when studying the travel behaviour differences between people living in a city centre or in a suburb. Inhabitants of city centres tend to take the bicycle more than residents living in suburban areas, mostly due to the concentration of different functions in the city center (Witlox & Tindemans, 2004; Dill & Voros, 2007). As mixed land use environments have shorter distances to different functions deviating from homes, such as convenience stores, work, sports, leisure, and shopping, because land uses are mixed with living houses, cycling becomes more of an option. Shorter distances due to mixed land use developments result in higher cycling numbers. Pucher & Bucher (2006) show in their comparison between cities in the USA and Canada, the Canadian focus on mixed land use development resulted in higher cycling rates than comparable cities in the USA, that did not have governmental institutions that put focus on mixed land use development. Additionally, Wu et al. (2010) prove that accessibility to jobs in the proximity results to higher cycling levels and Chen et al. (2017) showed that areas with greater percentages of workplaces reflected higher cycling rates. The importance of mixed land use and especially by creating shorter home-work distances is well explained in current literature. All in all, it is about facilitating trip at distances favourable for cyclists, which can be done by urban designs promoting higher densities and mixed land uses (Heinen et al., 2009).

2.5.1 Policy mechanism

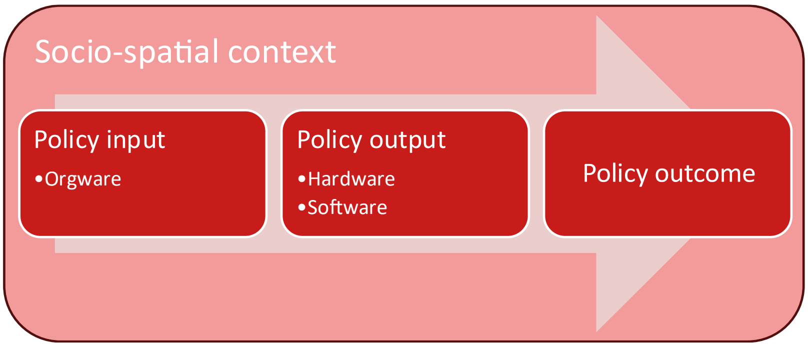

Figure 8: conceptual framework policy mechanism by Methorst et al. (2010).

The policy mechanism consists of several components as described by the conceptual framework from Methorst et al. (2010) and shown in Figure 8. It shows that the outcome of a policy is related to the policy input, the policy output, and socio-spatial contextual factors. The policy input refers to the 'orgware', the institutional conditions and framework in which the policy is made. The policy output is related to the measures taken, which can be divided into 'hardware' and 'software'. Finally, the context in which these measures will be taken has an influence on the effect of the policy (Methorst et al. (2010). These contextual factors relate to the determinants as explained previously. Taking this together, leads to the policy outcome, which represents the actual result of policies related in this study as bicycle mode share.

2.5.2 Orgware

The orgware is often overlooked in existing literature (Harms et al., 2015). It includes the organization and implementation of the policy, which is a crucial aspect in successfully implementing a policy. The limited available literature suggests that comprehensive planning with flexibility and adaptability is important for well working long-term cycling strategies. Additionally, debate exists on the involvement of various stakeholders, of which especially the involvement of citizens results in mixed recommendations. They could lead to inefficient planning and implementation and could create added costs (Harms et al., 2015). However, as seen in the Netherlands, the cycling advocacy organizations like the Dutch Cyclists' Union were a crucial pivot in promoting cycling (Pucher et al., 2011). A study by Harms et al. (2015) investigated the effectiveness of different municipal policies in Dutch cities and found that improving the 'orgware' positively affected the cycling rates. Crucial elements for an improved institutional framework and organizations seemed to be a measurable policy that can be monitored and has adaptability. It should furthermore create space for experimentation and include a high level of citizen involvement (Harms et al., 2015).

2.5.3 Hardware

The hardware component represents the direct influences, like the physical alterations to the transport system, regulations, and economic interventions (Steg, 2003). Examples of these measures include the construction of separated bicycle infrastructure, traffic calming networks, intersection treatments, and the implementation of speed limits (Schepers et al., 2017). All of these examples can be classified as 'pull' factors, which improve the attractiveness of cycling (Harms et al., 2015). Besides, improving the cycling performance in terms of cycling uptake in for example increased modal share, it also increases the safety (Reurings et al. 2012; Schepers et al. 2017). Besides improving the actual safety numbers, it also improves the perceived safety, a key determinant of cycling uptake as explained by multiple studies (Fishman et al., 2012; Berghoefer & Vollrath, 2023; Uijtdewilligen et al., 2024). Besides the physical changes to the infrastructure, regulations with direct financial consequences, such as pricing policies, are also considered as hard measures as they immediately result in behavioural changes (Steg, 2003). Such measures as increasing parking prices for cars or reducing car parking availability are classified as 'push' factors and make other modes, mostly focused on the car, less attractive and steer people towards the bicycle (Harms et al., 2015). All in all, the physical alterations of the infrastructure are highly related to the design component from the built environment discussed previously.

2.5.4 Software

On the other hand, soft measures aim to promote cycling through educational or promotional policy measures in order to change attitudes and perceptions, resulting in voluntary change in mode choices (Harms et al, 2015). Pucher and Buehler (2008a) mention that training, education, and promotional events are important policy focuses for promoting the bicycle. These communicational, educational, and training strategies raise awareness in order to create a behavioural change towards more sustainable transportation modes (Bamberg et al., 2011). This behavioral change is either focused on promoting these more sustainable transportation modes (Scheepers et al., 2014) or directly demotivating the use of private motorized vehicles (Graham-Rowe et al., 2011).

The importance of cycling education starting at a young age in order to form a cycling favourable behavioural routines at an older stage is found to be an important element (Pucher & Buehler, 2008; Hams et al., 2015). Dutch schools have incorporated compulsory traffic safety education since 1959 (Schepers et al., 2019) and municipalities actively encourage cycling through promotional efforts (Pucher and Buehler, 2008a). Although the direct impact of such educational soft measures may be modest (Reuring et al, 2012), their integration with infrastructural improvements can enhance cycling behaviour (Pucher and Buehler, 2008a). Besides education at schools, there is intensive motorist training and examination in the Netherlands (Pucher and Buehler, 2008a). Additionally, the Dutch law introduced at 1 January 1994 an important traffic-related law which made motorists held responsible for any collision with an injured child, cyclist, or pedestrian, which changed drivers' behaviour (Schepers et al., 2017). Another strategy discussed in the literature for promoting cycling is looking at individual life event changes, which seem to be the biggest driver of mode choice changes (Hams et al., 2015). On a broader scale, a National cycling skill program in Australia resulted in increased bicycle use, even after the program was finished (Rissel & Watkins, 2014). Yet, comparisons between cities on software policy outputs are limited and often lack data availability (Harms et al., 2015), which creates a significant literature and research gap.

2.5.5 Putting policy into practice

This section translates the theoretical policy mechanism into practice by presenting examples from both the Dutch and international contexts. This empirical evidences strongly suggest the effectiveness of especially the hard policy interventions of which some will be discussed below (Rietveld & Daniel, 2004; Harms et al, 2015; Pawluk De-Toledo, 2022). However, Harms et al. (2015) argue that most of these studies are location-specific and are thus not transferable to other contexts. Additionally, Harms et al. (2015) mention that the implementation of a control group is often overlooked, leading to possible misinterpretation of the actual effectiveness of the policy intervention.

The Netherlands experienced an 80% reduction in cyclist fatalities per billion kilometres cycled over the past 30 years (Schepers et al., 2017). According to Schepers et al. (2017), this success is due to an integrated road hierarchy with traffic-calmed areas, separated cycling paths, grade-separated crossings, and intersection treatments. The Dutch road hierarchy is mostly based on two main principles: homogeneity and functionality. Homogeneity refers to the desired state of having limited differences in speed, direction, and mass in order to create a safer transport system (Schepers et al., 2017). On the other hand, functionality refers to the road hierarchy, which makes differentiations between roads for handling mass traffic in order to ensure traffic flow, like distributor roads and through roads, and access roads for start and end locations of trips (Schepers et al., 2017). Combining the homogeneity and functionality together results in a safe hierarchical network with different speed limits, location of the cyclists, and functions. A study from Schepers et al. (2013) showed that one additional grade-separated intersection, like a cycling bridge or tunnel, per 10km at distributor roads (with car speed limits between 50 and 70km/h) and a 12% increase in traffic-calmed cycling kilometres results in a 24% lower fatality rate for bicycle-motor vehicle (BMV) crashes.

Besides the improved safety, the introduction of separated bicycle lanes also accounted for a higher modal share of cycling in the Netherlands, according to the CPB (2025). Their study showed that with the widespread introduction of separated cycling lanes, the share of cyclists for commute-related travel increased from 20% to 25%, largely driven by increased perceived safety and comfort.

Additionally, a higher modal share for cyclists increases their safety due to the Safety in Numbers (SiN) phenomenon (de Hartog et al., 2010). This phenomenon creates a positive feedback loop in which, if more people cycle, driver awareness increases and car traffic decreases as more people switch from the car to the bicycle. This together reduces the likelihood of crashes, which results in increased safety. This makes cycling even more attractive, resulting again in more people switching to the bicycle, further increasing driver awareness and reducing the likelihood of crashes.

This all illustrates the importance of a consistent and national policy favouring cycling to let the bicycle flourish. With this policy strategy, the Netherlands has become the country in which the bicycle plays the most important role in its transportation system. A national strategy can form a decent basis on which lower level governmental institutions such as provinces and municipalities can build. However, still differences between Dutch cities exist despite the fact that they all experienced the same cycling favourable national policy suggesting city-specific dynamics are at play.

Besides the above mentioned positive cycling related results from the Netherlands, similar results from policy interventions are seen internationally. Furthermore, multiple examples exist in which cities took manner into their own hands and with active policy positively changed cycling travel behaviour. Lowering speed limits for example has been widely empirically proven to improve cycling uptake and increase safety (Yannis, 2024; Kettle et al., 2017). Bologna, for example, serves as one of the first Italian cities to adapt to a speed limit of 30 km/h for almost their whole road network (90%) in 2023. As a result safety improved as a 13% reduction in traffic accidents is reported and the fatal traffic accidents nearly halved from 18 in 2023 to 10 in 2024. Furthermore, the cycling mode share increased by 10% and traffic-related air pollution decreased by 20% (Decisio, 2025). Similar results can be found in Wales, where lower speed limits around schools resulted in increased mode share in active transportation. Additionally, Helsinki has introduced lowered speed limits at multiple districts in the city over the last decades leading to an increased safety with no fatal traffic accidents amongst cyclists or pedestrians in 2019, when the last large scale alteration was made by giving two third of the city's road network a speed limit of 30km/h. On a broader scale, Yannis et al. (2024) investigated the effect of the introduction of the 30km/h speed limit in European cities and concluded that safety improved as substantial decreases were observed in crashes, fatalities, and injuries. Furthermore, the results indicated that the speed limit created multiple environmental benefits as emissions, noise pollution and fuel consumption all decreased by 18%, 2.5 dB, and 7% respectively.

Besides making rules and laws influencing travel behaviour, physical interventions to improve the infrastructure have widely proven to adjust travel patterns. Most of this is already discussed in the transport system literature section. These results underscore the importance for the physical implementation of cycling favourable infrastructure by governmental institutions. As literature on the effect of the software and the orgware is limited, this created a significant literature gap.

2.6 Conclusion literature

In summary, existing literature on travel behaviour highlights the complex dynamics behind mode choice, and thus bicycle use, explaining cycling performance. The choice is influenced by determinants operating across the various components at multiple levels. At the city level, local policies can adjust the built environment in such a way that creates circumstances that favour cycling. Furthermore, they can also directly influence individual behaviour by influencing attitudes and behaviours through soft measures. At the individual level, personal and household characteristics, attitudes, and habits determine how travel demand results in the choice for the bicycle. At the trip level, features such as distance, purpose, and timing further influence transport mode selection. Taken together, this illustrates a multilevel framework in which policy, environment, person, and trip factors are interconnected: higher levels influence lower ones, with policies shaping the environment and ultimately affecting individual habits and behaviour. This eventually leads to certain choices made at the trip level by individuals. This provides a hierarchical structured basis for analysing the determinants of cycling performance, in which the policy determinants can act as a tool for facilitating conditions for cycling.

Chapter 3 - Methodology

This chapter describes the methodology utilized to assess the determinants of cycling performance and the role of local cycling policies and local circumstances. Based on the literature review, it can be concluded that the cycling performance of cities is a result of a dynamic interplay of an extensive set of explanatory variables operating across multiple domains. Additionally, a significant knowledge gap regarding the effect of local cycling policies is present, which needs to be addressed. To address the research questions, the variables identified in the literature are subordinated into four main domains: trip characteristics, personal or household characteristics, environmental characteristics, and cycling policy characteristics. These domains reflect the structure used in the literature.

Examining the domains will be done by using a combination of analytical approaches. Quantitatively, existing travel behaviour data and geospatial data are used to describe the lower levels. In addition, semi-structured interviews are conducted in order to derive city-level policy indicators, which cannot be captured through existing databases as these are more of subjective manner. Gaining knowledge in this domain requires expert judgement. Combining both quantitative and interview-based approaches ensures a comprehensive understanding of both the behavioural and institutional dimensions. Given the nested structure of the data as a result of the literature review, a multilevel modelling approach is employed to account for variation at each scale level and allows for testing the influence of higher level variables on lower levels.

The remainder of this chapter outlines the methodological approach utilized in this research. First, the conceptual framework is introduced. Subsequently, the analytical techniques applied in the empirical analyses are described. The chapter then introduces the modelling approach, including a discussion of the models considered and the motivation for the final model used in the research. Finally, the procedures of data collection and refinement are outlined.

3.1 Conceptual framework

To address the research questions, the identified variables were grouped into four main domains, reflecting the structure used in the literature. The resulting conceptual framework can be found in Figure 9. The conceptual framework distinguishes four domains influencing cycling performance, which are trip characteristics, personal and household characteristics, environmental characteristics, and the policy domain. Simultaneously these domains correspond to different specific analytical scale level as described below and these all influence the cycling performance directly. The dependent variable, cycling performance, is operationalised at the trip level as a binary variable indicating whether the trip is made by a bicycle or not.

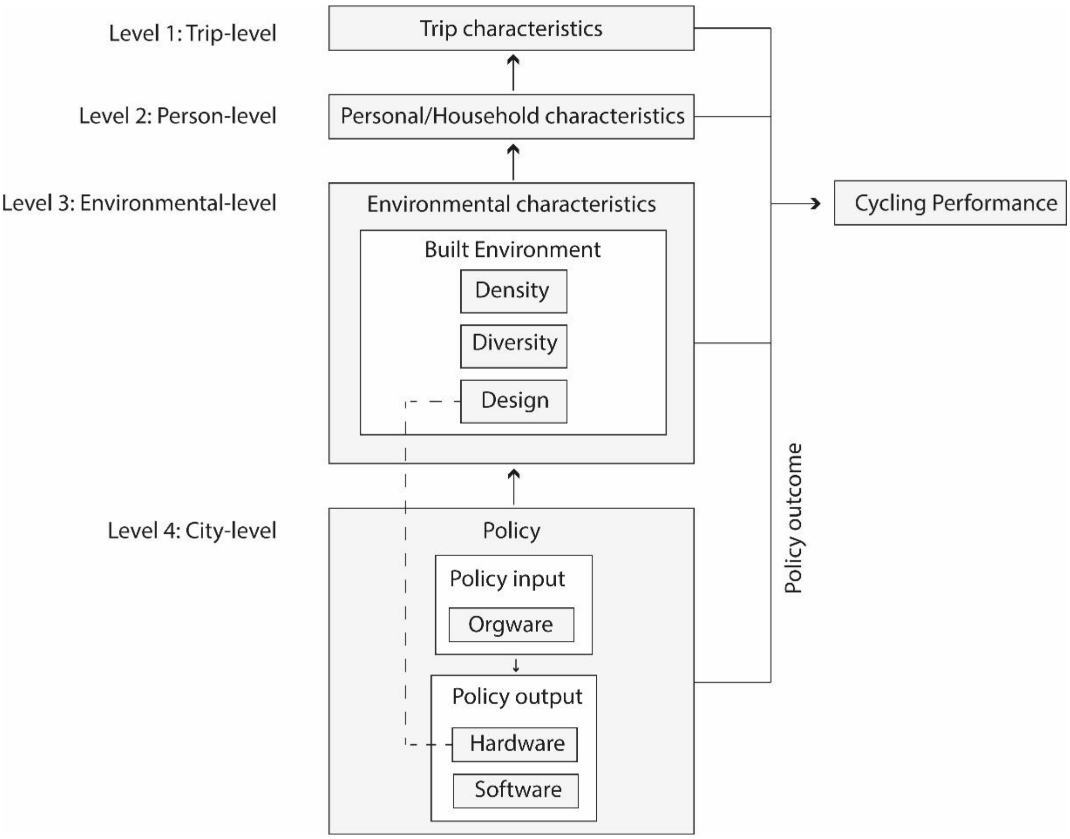

Figure 9 – Conceptual framework

Each domain thus contributes to cycling performance at a different level of analysis, ranging from the most detailed level, being trip-level to the broader city level. The trip characteristics act on the trip level (level 1), while trips are performed by persons (level 2), reflected by the personal and household characteristics. The persons travel within specific direct environments (level 3), reflecting the environmental characteristics expressed at the postal code level as will be discussed in section 3.3.1. These are in turn nested within cities (level 4), where policy factors are at play. This hierarchical structure suggests that higher-level domains can influence outcomes at lower levels, which is also shown in the conceptual model. This hierarchical nested data structure is illustrated in Figure 10.

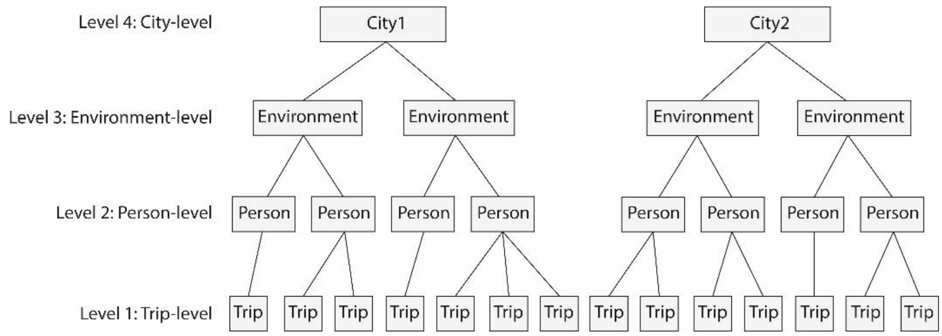

Figure 10 – Multi-level structure of the data

The conceptual framework implies that city-level policies may affect infrastructural components at the environmental level or both city-level and environmental level variables may affect individual- and trip characteristics. This possibility of cross-level effects should be taken into account when selecting the appropriate analytical method.

3.2 Model for empirical analysis

Building on the conceptual framework, this section outlines the analytical approach used to address the research questions. The framework shows that the determinants of cycling performance act on different levels which interact with each other in a hierarchical way from higher scale levels to lower levels. This requires a modelling strategy that accounts for these effects. Additionally, the dependent variable cycling performance is expressed as a binary indicator of bicycle use (bicycle taken for a trip or not). Given this, assessing the determinants of cycling performance requires a modelling approach that can handle both the hierarchical nested data structure and the binary outcome of the dependent variable. This shrinks the variety of different possible usable models substantially. The remainder of section 3.2 considers several modelling approaches possibly capable of handling both conditions.

3.2.1 Possible models for empirical analysis

Handling the binary outcome variable can be done with the basic standard binary logistic regression model, which estimates the outcome as the effect of the probability that a trip has been made by a bicycle (or not), influenced by the variables included in the equation. However, a main assumption of this regression is that all observations are independent. Given the hierarchical nested structure in which trips are made by persons who live in postal codes part of a city violates this assumption. Neglecting the clustering effect violates the independence, leading to biased standard error terms and overestimating p-values (Hox et al., 2017; Snijders & Bosker, 2012).

Another modelling technique suitable for binary outcomes is a fixed-effects model. This model controls for all unobserved heterogeneity at all levels. However, this model eliminates between-group variance, whereas comparisons between groups, especially on the postal code and city level, are one of the main aims of this study. This is done by introducing group-specific intercepts which represent all between-level variation, making higher-level predictors unidentifiable (Woolridge, 2010).

In summary, the standard binary logistic regression is not applicable because it ignores the hierarchical structure of the data. The fixed-effects model can indicate this hierarchical structure, but it eliminates the between-group variation on higher levels and can thus not make estimates on higher-level variables.

A model that can handle both hierarchical structure and between-group variation is a multilevel logistical regression model and is thus the most suitable analytical approach. It counteracts the limitations of the previously discussed models by implementing the hierarchical structure of the data by creating random intercepts at all levels. Furthermore, it allows for both within- and between-group effects. Additionally, it also tests how higher-level variables may influence lower levels (Rabe-Hesketh & Skrondal, 2012; Goldstein, 2011).

3.2.2 Multilevel logistic regression model

This subsection dives deeper into the multilevel logistic regression model. As described in the previous section, the model allows for variance between multiple levels and between groups within these levels. Additionally, it also tests the effect of influences of higher levels on lower level variables and accommodates correlations among observations within the same group (Goldstein, 2011; Hox et al., 2017; Rabe-Hesketh & Skrondal, 2012; Snijders & Bosker, 2012). This allows in-depth research possibilities for revealing the complex dynamics in cycling performance differences.

Mathematically, the multilevel logistic regression predicts the log-odds of the outcomes as a linear combination of the fixed and random effects as shown in Equation (1), in which $p_{i_1 i_2 \dots i_n} = \Pr(BicycleUse\ 1|\ level\ 1{:}\ i_1, level\ 2{:}\ i_2, \dots, level\ n{:}\ i_n)$, β_0 is the overall intercept and u_l is the random intercepts at the higher levels $l = 2, \dots, n$. If any intermediate level is included as a fixed effect, this will not result in a random intercept. The residual variation at the lowest level is captured by the model's error term ($\varepsilon_{i_1 i_2 \dots i_n}$).

$$logit(p_{i_1 i_2 \dots i_n}) = \eta_{i_1 i_2 \dots i_n} = \beta_0 + \sum_m (\beta_m X_{i_1 i_2 \dots i_n}) + \sum_{l=2}^{n} u_l + \varepsilon_{i_1 i_2 \dots i_n} \qquad \text{Eq. (1)}$$

By transforming the log-odds (logit) coefficient by Equation (2), the predicted probability can be assessed.

$$p_{i_1 i_2 \dots i_n} = \frac{e^{\eta_{i_1 i_2 \dots i_n}}}{1 + e^{\eta_{i_1 i_2 \dots i_n}}} \qquad \text{Eq. (2)}$$

This allows for the interpretation of the effect of each variable on the likelihood that the bicycle will be used for a trip, keeping in mind the hierarchical structure of the data (Hox et al., 2017).

The modelling procedure consists of several stages. First, the continuous variables on lower levels were grand-mean centered, which involved subtracting the overall sample mean from each observation. This results in the interpretation of intercepts relative to the mean. This allows for more stable regression coefficients and prevents comparison to a meaningless zero value in the continuous variables (Hox et al., 2017).

Secondly, all variables have been assessed for multicollinearity, because multicollinearity can bias estimates as independent variables possibly affect each other significantly, causing concerns about the interpretability of results. As multicollinearity is based on a fixed-effects design matrix instead of the hierarchical nested structure of the data, checking multicollinearity was done with an ordinary least-squares (OLS) model with all variables included at the same level (Hox et al., 2017; Snijders & Bosker, 2011). This is first checked using the Variance Inflation Factors (VIF), in which values above 5 indicate potential multicollinearity and values above 10 indicate serious correlation concerns (Hox et al., 2017; Kutner et al, 2004). O'Brien (2007) even indicates a more conservative VIF value of 4 for potential multicollinearity. As multiple categorical variables are present, these variables have multiple degrees of freedom and should be corrected for this using the generalized VIFs (GVIF) with $GVIF^{1/(2*df)}$ (Fox & Monette, 1992). Variables with a $GVIF^{1/(2*df)}$ value above 10 will be excluded from the research and variables with a $GVIF^{1/(2*df)}$ value between 4 and 10 will be further evaluated with an additional test, inspecting pairwise correlations.

Depending on the type of variable different correlation indicators will be used. For continuous variables the Pearson correlations will be conducted, in which a $|r| \geq 0.80$ is considered strong. For ordinal variables, the Spearman's rho will be used in which a $|\rho| \geq 0.80$ indicates a strong correlation between variables. Finally, if nominal variables indicate high VIFs, the Cramér's V will be conducted with a Chi-square test, in which $V \geq 0.50$ indicates a too strong correlation for this study (Cohen, 1988; Hox et al., 2017).

The next step is to create a so-called intercept-only (or null) model, which includes no predictors and explains the random intercepts for each level higher than the lowest level (Level 1). It estimates the proportion of the total variance which can be accounted for by each higher level, called the intraclass correlation coefficient ICC and acts as a justification test for performing a multilevel model (Hox et al., 2017). If the ICC has a certain threshold, it is considered that observations within the same group do not act independently from each other. Literature suggests a threshold of 0.05 for the ICC for this (Hox et al., 2017; Snijders & Bosker, 2012). Equation (3) shows the formula of the ICC at a level $l = 2$, in which $\pi^2/3$ represents the variance of the standard logistic distribution at the lowest level (level 1) (Hox et al., 2017; Snijders & Bosker, 2012).

$$ICC_l = \frac{\sigma_{l=2}{}^2}{\sigma_{l=2}{}^2 + \sigma_{l>2}{}^2 + \pi^2/3} \quad \text{(Eq. 3)}$$

An ICC of 0.05 for example at level $l = 2$ means that 5% of the total variance in the outcome can be linked to differences between groups in level $l = 2$, while the remaining part of the variation originates from other levels $l > 2$.

Finally, after justification of the multilevel logistic model by the intercept-only model results, the final model with all variables can be estimated and eventually be compared with the intercept-only model in order to assess the presence of improved performance. This will be done by assessing the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the -2Log Likelihood values (-2LL), in which lower values correspond to a higher model fit (Hoffman, 2013; Park & Park, 2022). estimates are in log-odds coefficients, meaning that they should be interpreted as odds ratios effects on the bicycle use dependent variable. All the steps made for the multilevel modelling approach are performed in the software R, Version 4.4.1 (Cran.R, 2024), which is an open-source programming language widely used for statistical computing.

3.3 Data

This section describes how the conceptual framework and the multi-level modelling approach are transferred into usable data for model input. The levels from the model can be linked to the domains of the conceptual model. A mixed-methods approach is used to include the extensive set of variables influencing cycling performance in Dutch cities. This study combines quantitative analysis of existing datasets with semi-structured interviews in order to derive city-level policy indicators. Both approaches are aligned with a multilevel conceptual framework, which distinguishes trip, person, environment, and city levels as mentioned in the previous sections. The trip and person-level variables originate from one data source. The variables on the environment level are coming from two data sources, and the variables at the city level are a result of the conducted semi-structured interviews. The analytical levels form the sequence of this section, in which for each analytical level, the data source, the filtering and extraction, and the operationalisation and preprocessing of the data are discussed.

3.3.1 Trip Level

Trip level information (and person-level and level division indicators) was obtained from the annually Dutch National Travel survey, 'Onderweg in Nederland' (ODiN) (CBS, 2024b), in which participants record their travel behaviour through a comprehensive multiple day travel diary enabling information on not only general trip- and personal-level characteristics, as well as more specific details on subjects ranging from vehicle information to deviations from participants' typical travel patterns. The extensiveness enables insights into daily mobility developments in the Netherlands. The 2023 edition is used for this research as it was the latest available version when starting the data collection process. The dataset contains 193,127 trip-level entries from 64,459 participants. The data is publicly available on request via Data Archiving and Networked Services (DANS). A repository from the Institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Dutch Research Council (NWO).

The trip-level variables gained from the ODiN dataset included the continuous variables reflecting the distance (AfstV, hereafter referred to as *Distance*, measured in hectometres) and the duration (Reisduur hereafter referred to as *Duration*, measured in minutes) of a trip, as well as a nominal variable indicating the trips motive (KMotiefV) in which 1= Home-work ,2= business and professional, 3= (personal)service, 4= shopping/grocery trips, 5= Education, 6= Visiting / staying over, 7= social/recreational , 8= Leisure walking trips, 9= other purposes). To ensure enough frequencies of each category in this analysis, the variable was recoded into a new variable (*Motive*) with five categories. In this variable the motives 1 and 2 were combined into “Work”, motives 3 and 4 into “Daily service/grocery”, motive 5 remained “Education”, motives 6, 7 and 8 were combined into “Recreational/Social”, and motive 9 remained “Other” purposes.

The dependent variable also originates from the ODiN dataset and also acts on the trip-level. It is based on the nominal trip-level variable indicating the main mode of transport (KHvm), in which 1= car - driver , 2= car - passenger, 3= train, 4= bus/tram/metro, 5= bicycle, 6= on foot, 7 = others. In order to investigate the factors associated with bicycle use, the variable was recoded into a new variable (*BicycleUse*), distinguishing trips in which the bicycle was the main mode of transport (coded as 1) or otherwise (coded as 0). Descriptive statistics of bicycle use will be shown in the results section and are based on the final dataset.

To extract the relevant data from the ODiN dataset, several filtering procedures were applied, as illustrated in Figure 11. These steps focus on converting the unit of analysis from stages into trips, removing irrelevant records, applying the spatial scope of the study, and removing outliers.

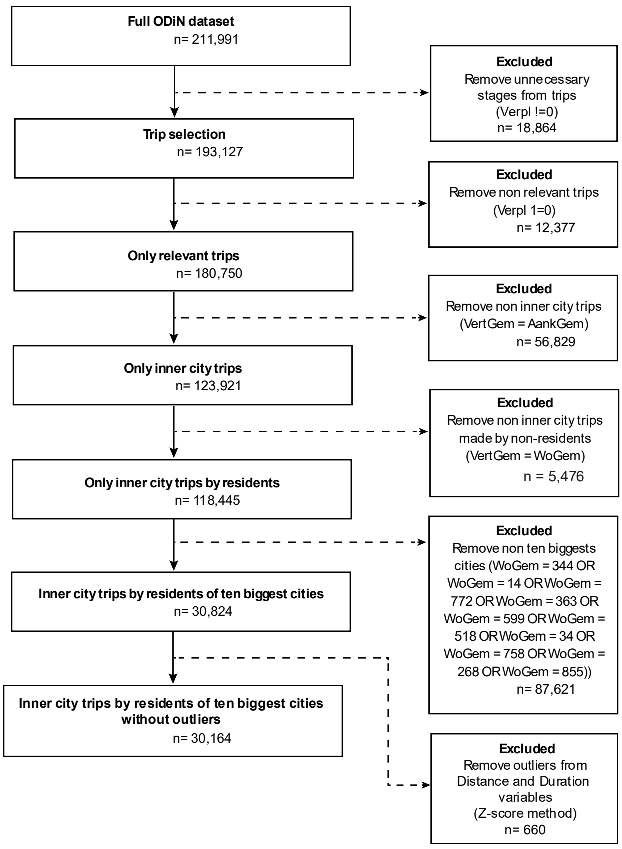

Figure 11 – Filtering process diagram

In ODiN, a trip is defined as a movement from an origin to a destination, which may consist of one or more stages. For example, a commute trip involving cycling to a station, taking the train, and walking to work includes three stages, but is treated as a single trip with the train as the main mode of transport. Trip characteristics are linked to the dependent variable representing the mode choice for the trip. As the mode is the same for all stages within a trip, duplicates at the stage level entries are removed from the dataset. This allows for trips as rows instead of stages in the final dataset. Additionally, extraordinary trips, such as work-related freight transport movements and trip series, are excluded from the analysis. This step scales the data down from 211,991 to 180,750 entries.

Next, the geospatial scope of the research was applied. As the primary goals of this study are to analyse differences in cycling performance between Dutch cites, only trips made within the ten selected cities were retained. To link postal code level data from other sources to the trips and person level variables, the residential postal code of each individual was used rather than the departure or arrival postal codes. This avoids mismatches caused by round-trip trips and better reflects the built environment in which individuals conduct most of their daily travel. This approach is consistent with previous research using trips over stages for travel behavioural research (Axhausen, 2007; Gärling, 2005)

This all leads to a filtering code which selects the trips made within the cities included in the research, by respondents living in that particular city, and can be found in Appendix A. This means a significant case reduction to 30,824 trips made by 10,395 persons in 431 postal codes in 10 cities.

To ensure that only meaningful observations are included in the analysis, outliers in the continuous variables *Duration* and *Distance* are removed. Excluding these outliers is done by using the Z-score Method, in which cases exceeding +-3 standard deviations from the variable mean are considered extreme and are excluded (Osborne & Overbay, 2004). Based on this criterion, trips longer than approximately 27.4 km (Distance: M = 37.9, St. Dev. = 78.710) and 136.8 minutes (Duration: M = 23.57, St. Dev. = 37.743) were removed. These thresholds sound reasonable within the context of inner-city travel in relatively compact Dutch Cities. With this step, 660 cases were removed, resulting in a final sample of 30,164 trips performed by 10,062 persons. A final division of the levels is shown in Table 1.

Table 1 – lower level data entry distribution per city

Level 4 - City	Level 3 - Postal code	Level 2 - Person	Level 1 - Trip
Almere	41	423	1,293
Amsterdam	72	1,762	5,311
Breda	24	339	1,112
Eindhoven	31	486	1,361
Groningen	39	549	1,852
Nijmegen	23	411	1,257
Rotterdam	69	2,124	6,087
Den Haag	61	1,899	5,319
Tilburg	28	409	1,286
Utrecht	43	1,660	5,286
Total	431	10,062	30,164

3.3.2 person-level variables

The variables at the person-level are also derived from the ODiN data source. The following variables could be used from ODiN to create person-level variables: age, gender, origin, education level, student occupation, and driver's license ownership. The preprocessing steps are described below.

First of all, the continuous age in years variable (Leeftijd, hereafter referred to as *Age*) was used. The binary gender variable (Geslacht, hereafter referred to as *Gender*) was recoded from 1/2 to 0 = male and 1 = female. This choice is made upon the analysis method chosen later and explained when discussing the modelling method.

The variable indicating the respondent's country of origin variable (HerkLand) was simplified from three categories (1= Netherlands, 2= Europe (excluding Netherlands), 3= Outside Europe (including unknown)) to two categories: 0 = Netherlands and 1 = Outside Netherlands.

Highest completed education (*Opleiding*) was used to create a five-level ordinal variable (*EducationLevel*) in which non-asked participants and 'Other' responses were combined with 'No education'. The highest completed education variable (Opleiding) was used to create a five category ordinal variable (EducationLevel) in which non-asked participants and 'Other' responses were combined with 'No education'.

Furthermore, the eight categorical nominal social participation variable (MaatsPart) was used to create a binary *Student* variable in which the occupation of being a student is 1 and all others 0. The binary driver's license variable (OPRijbewijsAu) was renamed to *DriverLicense.* Most of the variables are scaled down in fewer categories to improve interpretability, reduce computational times, and ensure enough entries per category.

To ensure external validity, a representativeness test will be conducted to assess whether the data sample corresponds with the actual population it should reflect. For this, the representativeness of the sample distribution across the ten cities is assessed. This is done by comparing the sample proportions to the actual population distribution of these cities. Population sizes at 1-1-2023 of the included cities from CBS were used (CBS, 2023). To evaluate whether the sample distribution deviates meaningfully from the population distribution, a chi-square goodness-of-fit test is conducted. This test examines whether observed categorical frequencies, in this case cities, differ from expected population frequencies in those cities (UCLA Institute for Digital Research and Education, 2024).

The chi-square test was statistically significant ($p<.001$), which is mostly due to the large sample size, as chi-square tests become highly sensitive when n is large, often producing significance even for small effect sizes (Serdar et al., 2021). Therefore, effect sizes were additionally calculated with Cohen's w (Eq. (4)) and Cramér's V (Eq. (5)), which provide information about the effect sizes (Ben-Shachar et al., 2023). The Cohen's w was 0.33, indicating a medium effect size, while Cramér's V was 0.11, indicating a small effect size (Ben-Shachar et al., 2023).

$$Cohen's\ \omega = \sqrt{\frac{\chi^2}{N}} \qquad \text{Eq. (4)}$$

$$Cramér's\ V = \sqrt{\frac{\chi^2}{n*(k-1)}} \qquad \text{Eq. (5)}$$

Additionally, a check with a comparison of percentage deviations between sample and population distributions shown in Table 2 indicates that most cities fall well within a ±5% range, except for Amsterdam (-6.56% difference) and Utrecht (+6.85% difference). The deviation in Utrecht can partly be explained by additional sampling effort conducted in Utrecht (CBS, 2024c). Overall, it is concluded that the sample distribution lies within acceptable terms to use in the analysis.

Table 2 – sample distribution difference in comparison to real population distribution

City	Difference
Almere	-1.64%
Amsterdam	-6.56%
Breda	-1.52%
Eindhoven	-1.56%
Groningen	-0.79%
Nijmegen	-0.70%
Rotterdam	+3.70%
s-Gravenhage	+4.12%
Tilburg	-1.91%
Utrecht	+6.85%

Additionally, tests have been made to check if the sample size is sufficiently large enough. The margin of error was calculated using Equation (6) (Ahmed, 2024; Taherdoost, 2017). The proportion used in this calculation (p = 41.3%) represents the share of respondents who used a bicycle at least once in their trips. This resulted in a margin of error of 0.00967 (0.97%), indicating a high degree of precision. With the same formula, the minimum required size for a 5% confidence interval can be calculated, which is 378.65. With a sample size of 10,062 persons, this is well above the minimum required sample size. Appendix B documents the chi-square results and the calculations for the above mentioned numbers. All in all, the sample is sufficiently large and is significantly representative for analysis.

$$E = \sqrt{\frac{Z^2 * p * (1-p)}{n}} \quad \text{Eq. (6)}$$

In which

- E = margin of error
- Z = Z-score from the standard normal distribution (1.96 for 5% confidence interval)
- p = proportion
- n = sample size

3.3.3 Built environment level

The built environment level information is derived from geospatial data sources. The analytical level at which the built environment domain will be explained is on the postal code level. For the density and diversity factors in the built environment as described by Cervero & Kockelman (1997), the postal code data from the Centraal Bureau voor de Statistiek (CBS) has been used, which provides detailed information on various subjects such as demographic composition, household structures, dwelling characteristics, the availability to local facilities, and various indicators of the urban form (CBS, 2025a). In the Netherlands the postal code is defined on different scale levels ranging from 4,070 numeric four postal codes to 464,138 four-digit and two-letter postal codes. For data completeness and the insurance of enough trip- and person-level entries for each postal code, the numeric four postal code is used. Mainly to align with the ODiN data, the data available in 2023 has been used (CBS, 2025b).

Density variables originate from the CBS postal code data source (CBS, 2025b) and are expressed using variables that represent either absolute counts or concentration on certain subjects. Absolute density indicators used in the research are the number of inhabitants (aantal_inwoners), households (aantal_part_huishoudens), and dwellings (aantal_woningen) within each postal code area, hereafter referred to as *Inhabitants*, *Households*, and *Houses* respectively.

In addition, a variable explaining the concentration of addresses within each postal code is included. This variable expresses address density using the '*omgevingsadressendichtheid'* (oad), a measure introduced in 1992 (den Dulk, van de Stadt, & Vliegen, 1992). The address density is defined as the number of addresses within a 1 kilometre radius of an address divided by the area of that circle. The average address density across all addresses within a postal code area is used to express the *AddressDensity* variable.

The variables explaining the diversity represent the demographic and urban composition and are also from the CBS postal code data source (CBS, 2025b). Two variables reflecting the population diversity are included in the analysis. The first variable explains the diversity in the population origin (percentage_geb_nederland_herkomst_nederland, hereafter referred to as *PerBornNL*). Same as for the variable on person-level, a variable is chosen indicating if the person has a Dutch origin and in all other cases a non-Dutch origin. However, this variable is expressed in the percentage of people of Dutch origin in that particular postal code. The second variable is the average household size (gemiddelde_huishoudensgrootte, hereafter referred to as *AverageHHSize*), which provides an indication of the composition of people living in the dwelling in the postal code area.

Two additional diversity variables are used to explain the characteristics of the dwellings. The first reflects the share of owner-occupied dwellings (percentage_koopwoningen, hereafter referred to as *PerOwnerOccupied*) as opposed to rental dwellings. The other variable captures the average housing value (gemiddelde_woz_waarde_woning, hereafter referred to as *AverageHousingValue*). With both variables, the socioeconomic differences between the postal code areas can be revealed.

Finally, an extended set of variables represents the shortest distance to certain facilities in the nearby surroundings. These indicators help indicating how diverse the postal code is, where shorter distances to facilities are expected to support the concept of the 15-minute city in which it is expected to have more pedestrian and bicycle trips due to the shorter distance (Cervero & Duncan, 2003; Pucher & Bucher 2006; Pucher & Bucher, 2008a; Hankey et al., 2012; Chen et al., 2017; Heinen et al., 2009). Conversely, if larger distances are present, the likelihood of active modes of transport will reduce. The included variables are linked to daily activities such as going to the supermarket or bringing your child to the child care center. Additionally, distances to different types of schools, to health facilities, and distances to facilities for other modes of transport have been included, of which the latter could indicate recommendations for push or pull factors for other modes of transport in order to increase cycling performance. An overview of all these variables can be found in the last section of this chapter, in which the grand overview of all variables used during this study is shown.

The design factor of the built environment is linked to the hardware component in the policy domain and covers the infrastructural components of the built environment. This data is derived from the Fietsersbond, which is a Dutch association committed to everything related to cycling. The organisation has its main office in Utrecht, the Netherlands, and 168 local departments spread throughout the country with more than 1,500 active volunteers. This enables influence on both higher and lower level regarding cycling policies.

Additionally, the Fietsersbond has a highly detailed route planner database, which enables the extraction of hardware measures regarding the cycling infrastructure from this source. This database consists out of sub-datasets on different subjects such as road segments, traffic lights, and street connectivity nodes. The route planner database is continuously updated, and the version used in this research has at least the update from 2 September 2025, with some sub-datasets in the database having even more recent updated data. The highly detailed and frequently updated information enables performing in-depth analysis of the infrastructure present in the postal codes. As the data from the route planner database is spatial data, it can be made visually interpretable in for example Geographical Information Systems (GIS) software such as the publicly available software QGIS. For this research, the QGIS 3.45.5 version has been used (QGIS, 2025). The individual sub-datasets can be loaded into QGIS as different layers. The sub dataset, which is the most data rich, is the *links* dataset (a line layer), which represents all the road segments in the Netherlands. It features 1,776,170 road segments with each segment having information on 43 different subjects such as length, location, quality, bicycle access, road authority, and lightning presence. Furthermore, another subdataset of interest for this research extracted from the route planner database is the *nodes* subdataset, representing the 1,244,785 network nodes where the links interact with each other. This represents the street connectivity.

The infrastructure is represented by multiple variables. First two types of roads are evaluated per postal code. The first one considers the road segments accessible to cyclists and the second type represents the road segments that physically separate cyclists from motorized traffic. Both types are expressed as relative proportions as well as densities, resulting in four variables in the model. The information originates from the *links* subdataset of the Fietsersbond route planner database.

The relative proportion of the bicycle accessible infrastructure is expressed as the percentage of bicycle accessible infrastructure length over the total infrastructure length within a postal code (*PerCyclingInfra*). Similarly, the relative proportion of separated cycling infrastructure is expressed as the percentage of separated cycling infrastructure over all bicycle accessible infrastructure (*PerSepCyclingInfra*). The densities are expressed as infrastructure type length in meters per square kilometre (*CyclingInfraKm2* & *SepCyclingInfraKm2*).

From the same *links* sub-dataset, the proportion of good quality cycling infrastructure (*PerQuality*) is calculated. This variable expresses how much of the bicycle accessible network is of good quality relative to the total bicycle accessible infrastructure per postal code. After consultation with the data owner, the decision was made to define good quality using only the asphalt/concrete-, vowels-, and tile roads assigned with good quality. The main reason for this is that the other surface road types represent less desirable cycling conditions such as unpaved paths. Additionally, as the percentage includes bicycle accessible infrastructure, only including asphalt or concrete roads would underestimate the true share of good quality cycling infrastructure. Only including the good quality of asphalt and concrete roads could be applicable if investigating the quality of the separated cycling infrastructure.

Furthermore, the street connectivity is expressed using the *nodes* (point layer) from the Fietsersbond route planner database, which represents all points at which road segments (from the *links* layer) intersect. To align with the cycling focus of the study, only the nodes of the cycling network are included. This still includes the places where cyclists and motorized vehicles interact with each other. This variable describes whether the connectivity of a cycling network is in such a way of high interaction with other roads is desirable or not.

As the CBS data and the Fietsersbond route planner data discussed in the previously are both interpreted at the postal code level and are of a spatial order, the data is combined spatially by using QGIS. To ensure consistency and avoid manual errors, a fully automated workflow was developed using the QGIS Model Designer option (See Appendix C). Automation also allows for smoother processing of large geospatial datasets and ensures reproducibility.

The first part of the Model Designer filters the applicable postal codes located within the selected cities. This ensures that all variables gathered in the process correspond to the spatial scope of this study. The second part extracts the correct information for the representation of the length of three types of infrastructure per postal code, which are the total infrastructure, the bicycle accessible infrastructure, and the separated bicycle infrastructure. It furthermore also calculates the length of good quality bicycle infrastructure. These attributes are the basis for constructing several indicators of (separated) cycling infrastructure and its quality.

The next part preprocesses the street connectivity by only including the nodes from the cycling network (including the places where these interact with motorized vehicles). The data is derived from the *nodes* subdataset. As this data set consists out of point geometries, the number of occurrences in each postal code can be computed.

The final part of the model designer transforms the raw infrastructural data into analytical variables, which can be used in the multi-level regression model. Because absolute lengths on the different infrastructure variables cannot be compared across postal codes, due to different postal code sizes, the variables of relative proportion and concentration, as described in the previous section, are calculated. The quality of the infrastructure will be presented in percentages. Furthermore, the street connectivity is expressed by the bicycle infrastructure length per occurrence of a node.

All the variables on the postal code are combined simultaneously in the above described process, leading to a final dataset including all variables explaining the built environment components. An extended explanation for reproducibility purposes can be found in Appendix D. This dataset is subsequently merged with the preprocessed ODiN dataset by connecting the CBS postal codes to the individual residential postal codes in the ODiN database as described earlier. An overview of all variables can be found in section 3.4 (Final variable overview).

3.3.4 City level

As the policy domain at city level is hard to capture through existing quantitative data, semi-structured interviews have been conducted to get insights on the role of local cycling policies. This type of information implies subjective perspectives, possible own-city biases and subjects that cannot be directly linked to cycling performance with quantitative measures.

In order to gain in-depth insights into the role of local cycling policies on cycling performance, semi-structured interviews were conducted. To obtain a robust overview of each city, interviews were conducted with both municipal representatives responsible for local cycling policies and with representatives of the local departments of the Fietsersbond. The contacts were mainly laid by the use of the network of the Fietsersbond. Meeting the representatives in person was preferred as asking follow-up questions became more applicable, resulting in more in-depth knowledge of the dynamics of local cycling policies.

Although the intended number of twenty interviews (two for each city) was not fully achieved, the sixteen completed interviews provided meaningful insights into the dynamics of local cycling policy measures and their effect on bicycle use. The missing interviews are with representatives from the municipalities of Amsterdam, Almere, and Den Haag, and with a representative from the local department of the Fietsersbond in Rotterdam. However, as each city was represented by at least one interviewee from either the municipality or the Fietsersbond, data on the city level for each city is ensured. In Tilburg, there is no local department of the Fietsersbond. Instead, Tilburg has a Fietsforum, which is a local cycling advocacy group that operates independently from the Fietsersbond and has a tight connection with the municipality of Tilburg.

Meeting the representatives in person was preferred as asking follow-up questions became more applicable, resulting in more in-depth knowledge of the dynamics of local cycling policies. However, from the sixteen interviews, eleven have been conducted in an offline setting. The remaining five interviews have been conducted via a video call due to the preference of the interviewee or the extensive travel distance for the author, which was not possible as multiple interviews occurred on the same day. Each interview was scheduled to be around approximately sixty minutes, of which some lasted forty-five minutes and others lasted almost ninety minutes. An overview of the presence of an interview, the interview setting, and the duration is shown in Table 3.

The interviews were conducted in a semi-structured manner, ensuring that specific predefined topics related to the software and orgware measures were discussed. The semi-structured format also left room for the interviewees to discuss and elaborate on city specific elements not mentioned in the predefined subjects. The first part of the interview is an open introduction to the city and its mobility system, serving as a base line needed for understanding the role of the bicycle in the current mobility system as this position is heavily influenced by historical developments and city specific characteristics. After this introduction, the predefined topics were introduced by predefined statements that should be scored on a five point scale. Interviewees were given the space to elaborate on the scoring, which gives in-depth insights for each topic. The statements used during the interviews originate from research conducted by Harms et al. (2015) in which medium sized cities in the Netherlands were evaluated on cycling performance.

Table 3 – overview interview characteristics

City	Representative of	Interview Setting	Duration (in minutes
Amsterdam	Fietsersbond	offline	71
	Municipality	MISSING	MISSING
Rotterdam	Fietsersbond	MISSING	MISSING
	Municipality	online	76
Den Haag	Fietsersbond	online	60
	Municipality	MISSING	MISSING
Utrecht	Fietsersbond	online	94
	Municipality	offline	60
Eindhoven	Fietsersbond	offline	50
	Municipality	offline	51
Groningen	Fietsersbond	online	42
	Municipality	online	60
Tilburg	FietsForum	offline	75
	Municipality	offline	68
Almere	Fietsersbond	offline	56
	Municipality	MISSING	MISSING
Breda	Fietsersbond	offline	65
	Municipality	offline	90
Nijmegen	Fietsersbond	offline	71
	Municipality	offline	65

First, four statements on the software have been taken as guidelines for this research and cover topics on education for children (1) and adults (2) and marketing campaigns with (3) and without incentives (4) in order to stimulate cycling behaviour.

Additionally, seven statements regarding the orgware are also originating from Harms et al. (2015). For this research the statements can be divided into two main topics of which the first one is the organizational structure and the second one is the topic of collaboration. For the organizational structure statements on the formulation of policy goals (5), the implementation of policy measures (6), the financial resources of municipalities for cycling policies and the policy consistency and adaptability have been taken. For the latter two, some changes have been made in the formulation of scoring these statements. For financial resources, the score indication is changed from how much the part is into how much effort is needed to get financial resources, think of ensuring structural budgets, subsidies by higher levels of governance and making accessible budgets through politics (7). The policy consistency and adaptability is changed into a policy consistency variable especially focussing on the consistency of local politics favouring cycling policies. The statement is thus double sided as it covers consistency in local policy leading college and if this college is consistently favouring cycling policies (8).

Three statements from Harms et al. (2015) are extracted to represent the collaboration factors of municipalities in order to increase cycling performance. The statements used are related to the topics of involvement of actors outside the policy area (9), the relationship between actors inside and outside the policy area (10), and the level of citizen participation (11). Additionally, an own made variable is added which covers the level of interaction with the Fietsersbond (12) with scoring a five is very constructive collaboration with for example including the Fietsersbond at the forefront of projects, during the projects and making cycling policies together to a 1 score indicating there is no to minimal contact and more in a conflict way than in a constructive collaboration manner. An overview from all twelve used statements with additional explanation is shown in Appendix E.

By combining quantitative data with insights from the semi-structured interviews, the study captures both existing measurable factors of cycling performance and the influence of more conceptual indicators on city-level policies and governance. This mixed-methods approach creates a comprehensive understanding of the dynamics of urban cycling. After all data was cleaned and restructured, a merged multilevel dataset was created capturing a four level nested structure with trips in persons, persons in postal codes, and postal codes in cities. The statistical software SPSS 28.0 from IBM (IBM, 2022) will be used for data processing and descriptive analysis.

3.4 Final variable overview

Figure 12 illustrates the completed conceptual framework with all variables included in the research. It demonstrates an overview and shows the complexity of urban cycling dynamics immediately. A table overview of all variables included in the research can be found in Appendix F. In total, forty-four variables will be used to explain cycling performance.

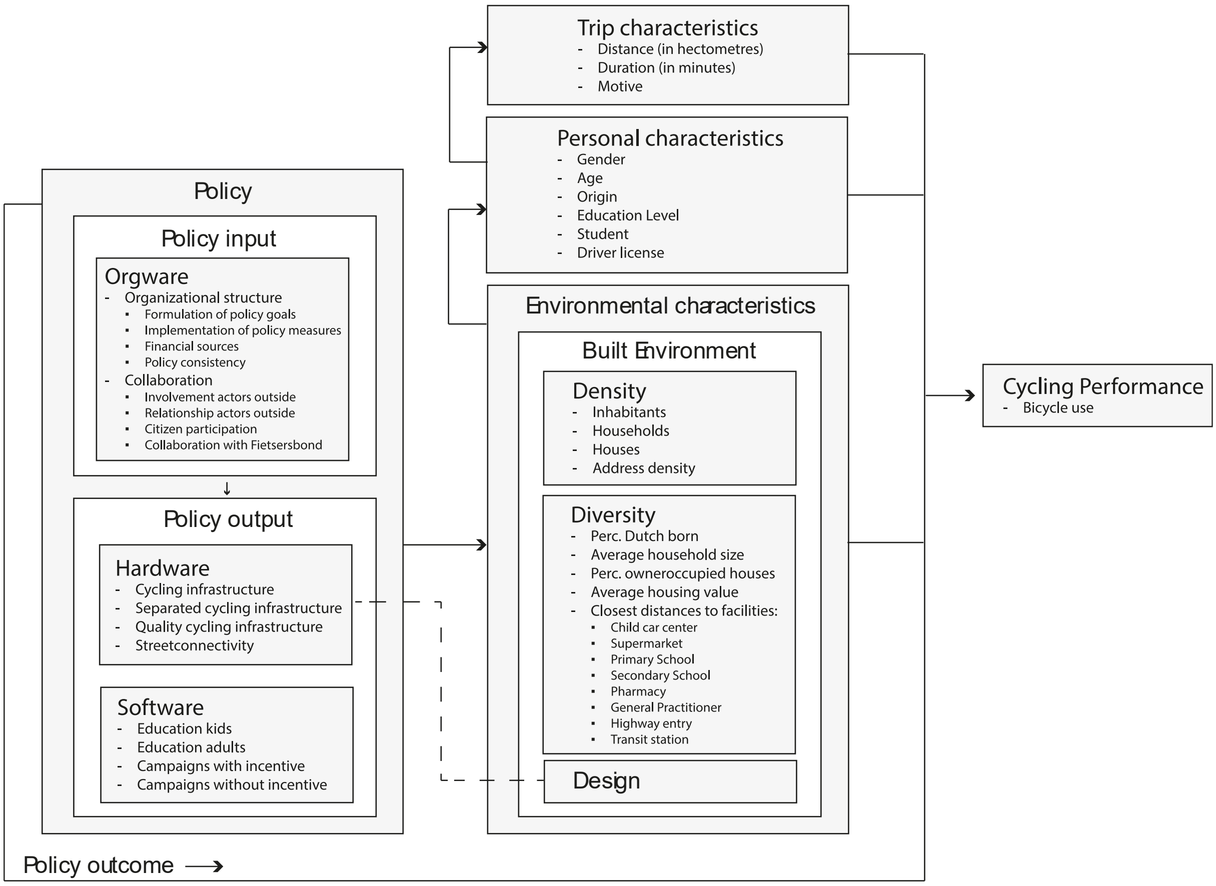

Figure 12 – Conceptual framework

Chapter 4 - Results

This chapter documents the results of the study. The chapter starts with preliminary results by explaining the descriptive statistics of each analytical level. Furthermore, this section discusses the in-depth and value insights gained from the performed semi-structured interviews. The second section deals with multicollinearity, after which the third section introduces the intercept-only model. Finally, this chapter ends with the documentation of the final multilevel logistic regression model.

4.1 Descriptive analysis

This section of the report describes the descriptive statistics of every variable included in the multilevel logistic regression model. This allows insights into the frequencies and the distribution of the variables, allowing a detailed understanding of the data that has been worked with. Additionally, it indicates missing values, outliers, or skewed distributions, all facets important to take into consideration as they could affect the model performance of the multilevel logistic regression(Hox et al., 2017). Furthermore, it operates as a final check to justify the data before usage in the model. The descriptive statistics will be discussed per analytical level. Moreover, additional scale construction steps at the city level are also present in this section. Finally, additional in-depth and valuable insights obtained through the interviews are also discussed in this section. Detailed tables of descriptive statistics can be found in Appendix G.

4.1.1 Trip level descriptive statistics

There are four variables at trip level of which two are nominal. The distribution of the binary nominal dependent variable *BicycleUse* is shown in Figure 13. Of the 30,164 trips included in the analysis, 37.1% were made by bicycle, while the other part (62.9%) was not made by bicycle.

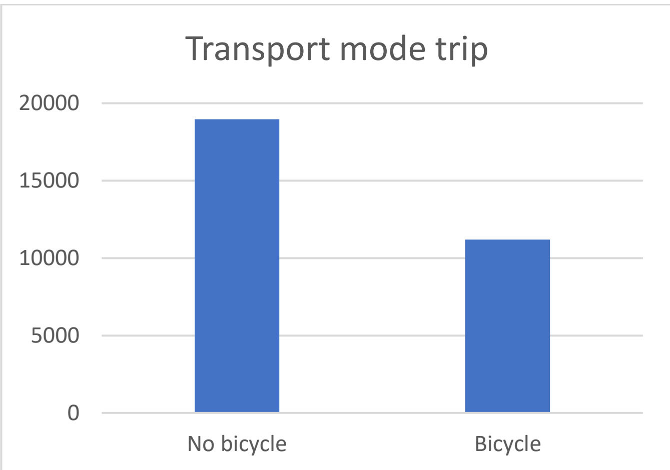

Figure 13 –Bicycle use distribution

The distribution of the other nominal variable, *Motive*, is shown in Figure14. From this, it can be concluded that most trips were related to recreational or social activities (35.8%), followed by daily service or grocery trips (31.0%). Trips related to work accounted for 13.4% and trip purposes defined as 'Others' accounted for 11.8%. The smallest motive for a trip in this sample is education related trips (7.9%).

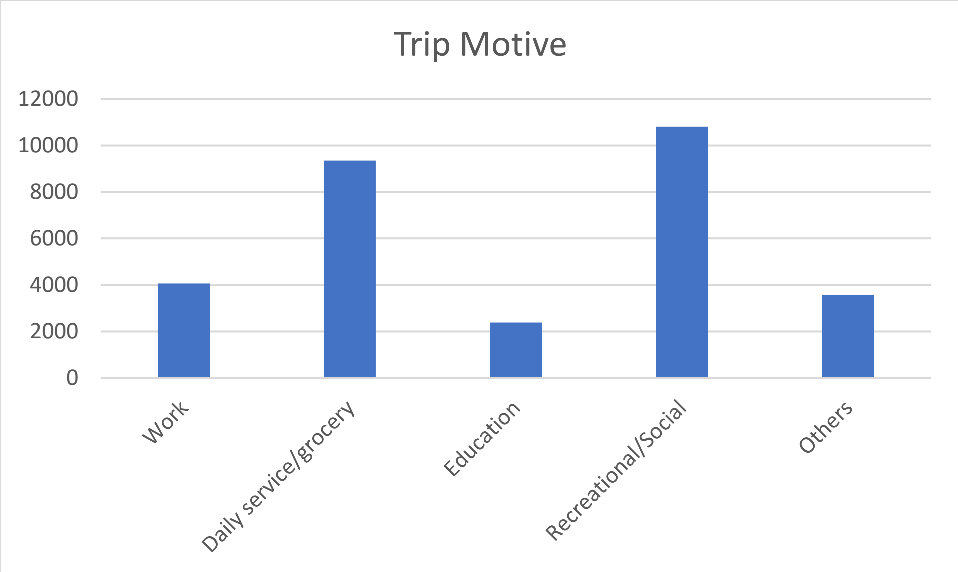

Figure 14 – Trip motive distribution

The last two variables in the trip level are continuous variables explaining the *Distance* (in hectometers) and *Duration* (in minutes) and are summarized in Table 4. Distances ranged from 1 to 270 hectometres with a mean of 32.52 (Std. Dev. = 34.40). Trip durations ranged from 1 to 136 minutes, with an average trip duration of 19.51 minutes (Std. Dev. = 18.95). As can be seen from the comparison of the mean and median, as well as the first quartile (Q1), median (Q2), and third quartile (Q3) assessment, both variables have a right-skewed distribution. This indicates shorter-than-average trips occur more frequently than longer-than-average trips. Outliers were already excluded using the Z-score method (*Distance* & *Duration* variables), as described in the data collection and refinement section.

Table 4 - Descriptive statistics distance and duration variables

Variable	n	Min	Max	Mean	Std. Deviation	Q1	Median	Q3
Distance	30,164	1	270	32.52	34.404	4	20	74
Duration	30,164	1	136	19.51	18.945	5	15	40

All variables on the trip-level have n = 30,164 indicating that there are no missing values are present. The descriptive statistics on trip level variables show that both distance and duration variables follow a right skewed distribution. Furthermore, the daily service/grocery and recreational/social motives for trips are by far the most common trip motives.

4.1.2 Person level descriptive statistics

The 30,164 trips in the dataset have been made by 10,062 unique individuals of whom their characteristics have been explained by six variables. Four of these variables are binary (*Gender, Origin, Student, Driverlicense),* one is ordinal (*Educationlevel*) and the latter is continuous (*Age*). The distributions of the nominal and ordinal variables are illustrated in Figure 15 (binary variables) and in Figure 16 (ordinal variables).

The distribution between men (49%) and women (51%) is almost evenly split. Individuals of Dutch origin are slightly overrepresented (58.3%). Furthermore, the sample contains substantially more non students (78.7%) than students (21.3%) and a majority has a driver's license (66.3%). Overall, the distribution of these variables is consistent with the expectations. Regarding education, there are more individuals in the highest education level category (51.6%) than all the other education level categories combined.

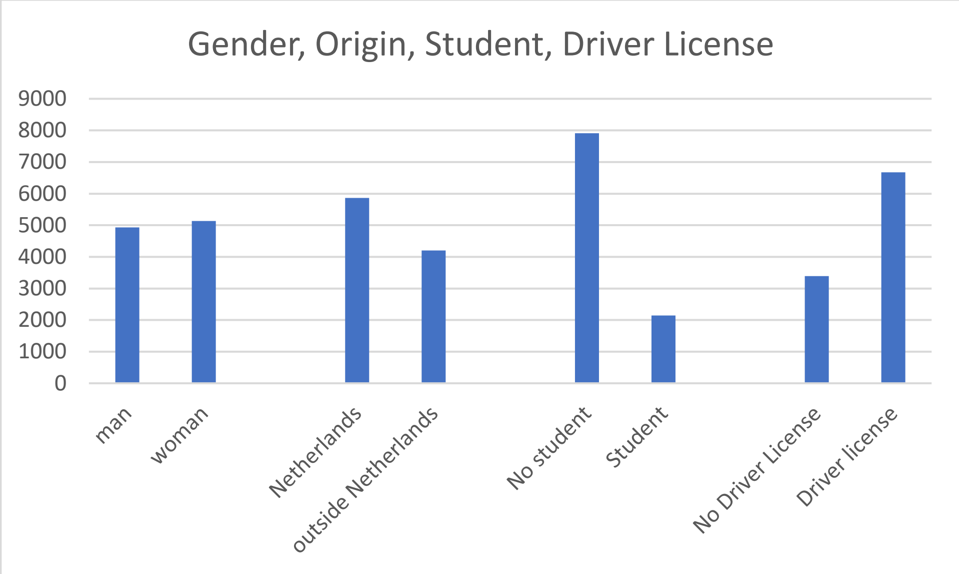

Figure 15 – Distribution of binary variables gender, origin, student and driver license

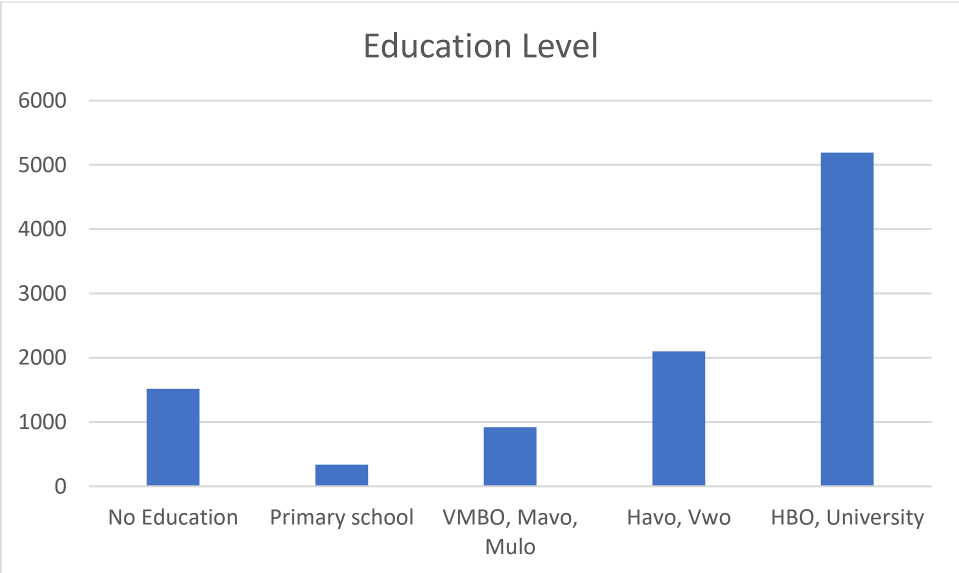

Figure 16 – bar chart education level variable

The continuous variable Age ranges from 6 to 99 years, with a mean of 40.64 years (Std. Dev. = 20.30). The distribution of the age variable is slightly right-skewed, indicating a higher occurrence of younger individuals relative to the mean. This can be concluded as the median is smaller than the mean and is closer to Q1 than to Q3 as can be seen in Table 5. Again, for all variables no missing values are present as for all variables n = 10,062.

Table 5 – Descriptive statistics Age variable

Variable	n	Min	Max	Mean	Std. Deviation	Q1	Median	Q3
Age	10,062	6	99	40.64	20.30	25	37	56

4.1.3 Postal code descriptive statistics

In total, 431 unique postal codes are present on the postal-code level scale. The variation between postal codes is explained by 23 continuous variables. For three variables (*Houses*, *PerOwnerOccupied,* and *AverageHousingValue*), two postal codes (1101 and 3041) have missing values, due to no data availability from the CBS postal code database (CBS, 2025b). The postal codes will automatically be excluded from the multilevel regression by R. This means that on lower levels 9 trips (from the 30,164 trips) and 2 individuals (from the total of 10,062 individuals) will be removed. All other variables on the postal code level are not reporting missing values.

A detailed overview of the descriptive statistics for the continuous variables is shown in Table G4 in Appendix G. Overall, substantial heterogeneity across postal codes can be concluded. The number of inhabitants ranges from 50 to 26,245, with a mean of 8,832.69 (Std. Dev. = 4933.83), and the number of households ranges from 35 to 14,140 households with a mean of 4,579.94 (Std Dev. = 2715.24), both indicating a substantial diversity of density numbers in population. Additionally, the address density descriptives are also underscoring the differences in density indicators between postal codes with a minimum address density of 41, a maximum of 11,760, and a mean of 3,465.92 with a standard deviation of 2,411.38.

Contrary is the relatively constant average household size with a minimum of 1.2 and a maximum of 3.2 with a mean of 1.99 (Std. Dev. = 0.36). The relatively constant average household size can mostly be indicated with the Q1, Median, and Q3 values, which don't deviate that much from each other. The distribution of the percentage of the population born in the Netherlands (mean 56.06 and Std. Dev. = 19.1) indicates a significant variation between postal codes.

Housing economicrelated variables indicate substantial diversity as well. The percentage of dwellings that are owner-occupied ranges between 0% and 100%, with a mean of 45.97% and a standard deviation of 21.52%. Furthermore, the housing value ranges from €189,000 to €1,267,000 with a mean of €407,680 and a standard deviation of €158,880. These numbers both indicate wide dispersion among postal codes.

The descriptive statistics on the accessibility variables, the variables reporting the shortest distance to certain facilities, show both similarities and dispersion. Some facilities are concentrated at smaller distances, such as Supermarkets, child care centres, and primary schools, which all have medians concentrated between 0.5 and 0.8. Secondary schools are located, in general, at a further distance, with a median of 1.1. Facilities with a longer distance include mostly accessibility to other transport possibilities such as highway entries (Median =2.1) and transit stations (Median = 4.1). Additionally, General practitioner centres are also located at further distances (Mean of 3.9).

The different cycling infrastructure variables show dispersion among postal codes, as minimum and maximum values are widely spread across the variables and relatively widespread values of the Q1 and Q3 quartiles. The same accounts for the quality of the infrastructure variable, which has a mean of 55.38 and a 38.70-72.51 Q1-Q3 range with a median of 52.25. Additionally, the street connectivity variable shows a wide range between minimum and maximum values (66.61-955.27). However, the Q1, median, and Q3 values do not differ on such a large range (99.11, 110.94, 133.91). This indicates that the majority has somewhat the same level of street connectivity.

4.1.4 City level descriptive statistics

The outcome of the interviews has been transformed into scores for each statement as explained in the methodology section. Eventually the final score on each statement for a city results from taking the average score from the representatives of the Fietsersbond and municipality per city and it illustrated in Table 6 (for software statements) and Table 7 (for orgware statements).

Table 6 – Final scores for software variables

City	Eduation Children	Education Adults	Marketing campaigns without incentive	Marketing campaigns without incentive
Amsterdam	4	5	4,5	3
Rotterdam	3	3	3	4
Den Haag	3	3	3	4
Utrecht	4	4	4,5	3.5
Eindhoven	3	3.5	3,5	3.5
Groningen	4	3.5	4	4
Tilburg	3.5	3	3	3.5
Almere	3	2	2	3
Breda	4	4	3	4.5
Nijmegen	3	3.5	3,5	2.5

Table 7 – Final scores per city

	Orgware - Organizational structure				Orgware - Collaboration			
City	Formulation policy goals	Implementation of policy measures	Financial sources	Policy consistency	Involvement actors outside policy area	Relationship between actors insides and outside policy area	Level of citizen participation	Level Fietsersbond
Amsterdam	4	4	4,5	5	4	4	4	5
Rotterdam	3	4	3,3	2	4	4	4	4
Den Haag	4	2	4	3	3	3	3	3
Utrecht	5	4	4	5	5	4	4	5
Eindhoven	4	3,5	5	4,5	3	4	3.5	4.5
Groningen	4	4,5	4,5	5	4,5	4	4.5	5
Tilburg	3,5	3,5	4	4	4	4	4	4.5
Almere	3	4	3	2	3	4	3	3
Breda	3	3	4	3	2	3	4	2
Nijmegen	4	3,5	4	5	2,5	3	4	2.5

The descriptive statistics of the city level variables show limited variations within each of the main domains (software measures, Orgware - organizational structure, and Orgware - collaboration). Additionally, variation within each variable is, in most cases, also limited. An overview of the descriptive statistics is shown in Table G5 in Appendix G.

The variables regarding software measures (education on kids and adults, and marketing campaigns with and without incentives), indicated very similar mean scores (Mean = 3.40-3.55) with relatively small standard deviations (0.497-0.798). The interquartile ranges (A1-A3) also lie within a limited one-point step on the Likert scale for all variables. This all indicates low dispersion within and between variables, indicating that municipalities either score the same on every individual variable or that respondents evaluate the distinct variables more or less the same, possibly due to the same underlying construct (software measures).

A similar pattern is indicated for the orgware variables explaining the organizational structure measures. Three of the four items show modest dispersion (Std. Dev. = 0.624-0.699) and small differences in interquartile ranges. This again suggests overlap in variables due to the same underlying construct. However, the variable explaining policy consistency by municipalities shows a substantially higher standard deviation (1.248) and a wider interquartile range (Q1-Q3 = 3-5), indicating more dispersion in policy consistency between cities.

The variables explaining collaboration measures made by the municipalities also indicated relatively low variability within and between variables. Although this variability is somewhat larger than the software and organizational structure measures. Most variables again show modest dispersion and interquartile ranges within one-point Likert scale intervals. However, the variable explaining the level of collaboration with the Fietsersbond noticeably differs with both a higher mean (3.85) and a larger standard deviation (1.132). Additionally, a broader interquartile range (Q1-Q3 = 3-5) is documented. This indicates that municipalities differ more strongly in their collaboration intensity with the Fietsersbond. Furthermore, the variable indicating the involvement outside the policy area of municipalities shows a slightly wider minimum and maximum range than the relationship with actors outside the policy area and citizenship. However, as it has a narrow Q1-Q3 range, this suggests that only a few cities differ significantly from the other cities.

Overall, the descriptive statistics show relatively low dispersion within and between the variables linked to the same constructs. This is likely due to the relatively small number of cities included (n = 10). Still, two variables (policy consistency and collaboration with Fietsersbond) indicate greater variations, indicating municipalities differ more substantially on these subjects.

Scale construction

Table 7 and Table G5 in Appendix G reveal low dispersion among cities and indicate little variation within variables on city-level. this suggests that these variables may represent different explanations for the same shared constructs described above (software measures, the organizational structure measures, and collaboration measures). To examine whether the variables could be meaningfully combined into these composite constructs, the inter-item correlations were examined.

Inter-item correlations show how strongly different variables (items) measure the same underlying construct (scale) (DeVellis, 2017). Desirable inter-item correlation typically falls between 0.15 and 0.50 (Clark & Watson, 1995). Correlations below 0.15 indicate weak conceptual coherence, and very high correlations (>0.80) could indicate redundancy or near-duplicates (DeVellis, 2017; Tavakol & Dennick, 2011).

Across the three overarching groups of variables, the inter-item correlations in general supported the assumption of underlying constructs of these overarching categories. Three of the four variables on software measures showed strong positive correlations ($r \geq 0.635$), with the other one (marketing campaigns without incentive) showing a weaker relationship with the other variables ($r = 0.006 – 0.289$). The inter-item correlations for the organizational structure measures follow the same pattern, in which again three variables indicating high shared variances ($r \geq 0.561$, but <0.8), and one variable (implementation of policy measures) showing small relationships ($r = 0.063 – 0.274$). Lastly, the variables linked to collaboration measures showed significant strong inter-item correlations ($r > 0.190$), indicating that these items indeed express closely related collaboration dynamics. Two correlations even exceeded 0.80, suggesting conceptual overlap, which is not in a problematic way given the distinct meaning of every variable.

Subsequently, internal consistency was also examined with a reliability test using Cronbach's Alpha. A Cronbach's Alpha threshold of 0.70 is widely accepted in literature in order to confirm combining variables into a new variable (scale) due to internal reliability (Nunnally & Bernstein, 1994; Gliem & Gliem, 2003; DeVellis, 2017). Given this, the scale capturing four software related measures showed an acceptable reliability (α = 0.720). Similarly, the four orgware measures related to organisational structure indicated acceptable internal consistency (α = 0.726). The four orgware variables linked to collaboration measures indicated an even stronger internal consistency (α = 0.831). An overview of the Cronbach's Apha results are shown in Table 8.

Table 8 - Cronbach's Alpha's

	Cronbach's Alpha	**N of items**
Software measures	0.720	4
Orgware - Organisational structure measures	0.726	4
Orgware – Collaboration measures	0.831	4

Given their conceptual coherence, acceptable inter-item correlations and reliability, the three groups of variables were each combined by subject into new variables explaining either software measures, organizational structural measures or collaborative measures, as shown in Table 8, by calculating their sum of scores. These combined variables provide more stable and interpretable measures of the overarching constructs and reduce unnecessary correlation between the individual variables if not combined. The three new variables are used as predictors in the multilevel regression analysis.

4.1.5 Complementary key insights from interviews on cycling policies

This section describes the additional insights captures during the conducted semi-structured interviews. This separated section is added to the study, as the quantitative results on city level do not fully capture the understanding of how local cycling policies act in practice. Leaving out the valuable in-depth findings gained during these interviews would omit essential information. Incorporating these insights allows for a more comprehensive view of urban cycling dynamics and the role of local cycling policies.

In general, the interviews revealed that the dynamics behind local cycling policies are complex and city specific. Although almost all cities have an ambitious and clearly stated cycling vision or an agenda, the mandate to implement them following a concrete action plan varies significantly between cities. In cities with lower bicycle use, this is often linked to an unstable political climate, which also does not explicitly prioritize cycling. A main conclusion is that all interviewees emphasized that a politically stable climate that favours cycling is a crucial underlying condition, which allows for long-term planning, structural budgets and an institutional organization that shows mandate for improvements in cycling-related projects and programs.

The interviews also indicated that there is interaction between the different policy components and that these interactions differ between cities. As for some cities ensuring stable and structural financial resources is a prominent problem (often linked to a non-cycling-oriented political climate), some cities seek better collaboration with organizations such as the Fietsersbond. Additionally, the mandate and getting the financial funds together differ between cities explaining the differences in cycling performance. These city-specific dynamics shape city based strategies, indicating that a one-size-fits-all approach is insufficient. Understanding this is essential for making a framework with effective interventions. Hence, an organisation that is structurally focussing on prioritizing cycling seems to also make more things possible at other policy components. Highlighting the need for a stable organizational structure and the existence of a real mandate again.

Furthermore, constructive collaboration between municipalities and local departments of the Fietsersbond is emphasized as highly valuable by all participants. They consistently indicate that this collaboration strengthens cycling initiatives, as the two organizations provide complementary perspectives. Municipalities acting with professional expertise and resources, and the Fietsersbond provides bottom-up insights and local knowledge. Additionally, several local Fietsersbond departments also have volunteers which are working in the mobility field, which results in constructive and on a professional level interactions between the municipalities and the Fietsersbond. In cities with high bicycle use, the collaboration is often transparent and very constructive, underscoring the value of the top-down and bottom-up perspective combination leading to cycling favouring policy implementations.

Participants, generally speaking, prefer implementing these policy measures with the use of a stable integrated policy agenda with long-term planning in which emphasis is placed on cycling. Hence, they also mention that so called incident politics can accelerate cycling projects due to political pressures. It is thus important to have plans and funds ready and come into action at the right moment.

Regarding the more softer policy measures, it can be concluded that municipalities do not put that much attention, with some exceptions, into these policy components as direct effects on bicycle use are mentioned to be not directly quantifiable. The real effect of these measures is lacking, as a result of which attention is generally given to other priorities. This suggests that more research on this topic could encourage municipalities to make more determination into these policy measures.

Overall, the semi-structured interviews provided highly valuable and city-specific insights into how cycling policies work on the local scale in the Netherlands. The main conclusions are the presence of city-specific policy implementation strategies due to location specific factors, the crucial role of political stability, and the importance of a constructive collaboration between municipalities and local departments of the Fietsersbond. These findings gave the author a deeper understanding of the dynamics behind the role of local cycling policies on cycling performance. This enabled more informed and grounded evaluations of the statement, which eventually led to the values for the variables on the city level.

4.2 Multicollinearity

As described in the methodology chapter, multicollinearity was assessed for all variables included in the multilevel logistic regression model to prevent increased standard errors and ensure stability and correct interpretability of regression coefficient estimates (Hox et al., 2017; O'Brien, 2007). An OLS regression including all independent variables was estimated, of which subsequently the Generalized Variance Inflation Factors (GVIFs) and adjusted $GVIF^{1/(2*df)}$ values were extracted in R. As two variables had multiple categories (*Motive* and *EducationLevel*), the adjusted $GVIF^{1/(2*df)}$ values were used for evaluation as recommended by (Fox & Monette, 1992). A complete overview of the GVIFs and adjusted $GVIF^{1/(2*df)}$ values are documented in Table H1 in Appendix H.

From this assessment, it can be concluded that no variables exceed the commonly used critical threshold of 10 for multicollinearity (Hox et al., 2017; O'Brien, 2007), and most values were well below the thresholds of (Hox et al., 2017) or even the more conservative thresholds of 4 by O'Brien (2007). Hence, the three postal code level variables *Inhabitants_gmc* (5.717), *Households_gmc* (8.479), and *Houses_gmc* (7.151235) fell within the intermediate range of 4/5 to 10, indicating the need for further examination by performing pairwise correlations to examine and support the interpretations of the VIF results.

As these variables are all continuous, Pearson correlations between the variables were calculated at the postal code level. As discussed in the methodology section, Pearson correlations above $|r| > 0.80$ indicate problematic collinearity. As can be seen in Table 9, the Pearson correlation coefficient r ranges between 0.948 and 0.988, which is well above the threshold and even significant at the 99% confidence level. As the three variables reflect the same underlying construct, the high VIFs and Pearson correlations are not surprisingly. To reduce multicollinearity, *Households_gmc* and *Houses_gmc* were removed from the model, as *Inhabitants_gmc* remained in the model, while it is the most direct interpretable variable for municipalities.

Table 9 – Pearson corelations *Houses, Households, Inhabitants*

Correlations				
		Houses	Households	Inhabitants
Houses	Pearson Correlation	1	.988**	.957**
	Sig. (2-tailed)		0.000	0.000
	N	430	430	430
Households	Pearson Correlation	.988**	1	.948**
	Sig. (2-tailed)	0.000		0.000
	N	430	431	431
Inhabitants	Pearson Correlation	.957**	.948**	1
	Sig. (2-tailed)	0.000	0.000	
	N	430	431	431
**. Correlation is significant at the 0.01 level (2-tailed).				

4.3 Intercept-only model

The first step of the multilevel modelling approach is to assess an intercept-only model. This model shows the variance on each analytical level and acts as a justification for using this modelling strategy as explained in section 3.2.2 (Hox et al., 2017).

4.3.1 Four level intercept-only model

Before making a multilevel logistic regression model with the explanatory variables discussed in Section 4.1, an intercept-only (or null) model will be constructed in order to assess the partition of the variance in bicycle use across trip (Level 1), person (Level 2), postal code (Level 3), and city level (Level 4). It will serve as a justification of performing a multilevel logistic regression model and the result of the intercept-only model follows Equation 7.

$$logit(p_{ijkl}) = \beta_0 + u_{0j} + v_{0ij} + w_{0ijk} \qquad \text{Eq. (7)}$$

In which: p_{ijkl} = probability that a person *k* in postal code *i* of city *j* uses a bicycle for a trip

β_0 = fixed intercept

u_{0j} = random intercept for city *j*

v_{0ij}= random intercept for postal code *i* in city *j*

w_{0ijk} = random intercept for person *k* within postal code *i* and city *j*

The results from the four-level intercept only model are presented in Table 10 and show that almost all variance will be at the person level (135.1). It furthermore shows that the variance on the postal code and city level are both negligible (both 0.0). An ICC calculation (Equation 8) shows that 97.6% of the variance is active at the person level. The remaining 2.4% act on the trip level and no variance happens at the city level. This intercept-only model furthermore indicates a relatively high fixed intercept, which eventually corresponds to the predicted probability of bicycle use in the intercept-only model of $p = \frac{e^{\beta_0}}{1+e^{\beta_0}} \frac{e^{-7.2067}}{1+e^{-7.2067}} \approx 0.0007 = 0.0741\%$. This is also a very low percentage considering that 37.1% percent of all trips are made by the bicycle.

$$ICC_{person} = \frac{\sigma^2_{person}}{\sigma^2_{Person} + \sigma^2_{PostalCode} + \sigma^2_{City} + \sigma^2_{Residual}} = \frac{135.10}{0.00+0.00+3.290} = 0.976 \qquad \text{Eq. (8)}$$

Table 10 – random and fixed effects four-level intercept-only model

Random effects for each level					
Level		Name	Variance		Std. Dev.
Person		w_{0ijk}	135.1		11.6
Postal Code		v_{0ij}	0.0		0.0
City		u_{0j}	0.0		0.0
Fixed effects					
	Estimate	Std. Error	z value	Pr(>|z|)	
β_0 (intercept)	-7.2067	0.1238	-58.23	<2e-16***	
Number of observations (Trips): 30,164, levels: Person, 10062; Postal Code, 431; City, 10					
sign ***: p<.001					

The fact that most variance is observed at the person level is likely due to the sparse clustering of trips per person (average of three trips per person). This scarce lower-level clustering can blow up the variance dominance on lower levels (person level in this case) and makes variation at the higher level uninterpretable (Maas & Hox, 2005). To solve this problem, the person level variables will be included as fixed effects rather than with a random intercept. They will act on the lowest level (trip level). This leads to a three-level model with trip, postal code and city level. This choice leads to variance on all included levels as will be discussed in section 4.3.2. The present variance on all levels justifies the use of the multilevel logistic regression modelling approach. Furthermore, the three-level model also ensures meaningful estimation of postal code and city-level effects.

4.3.2 Three level intercept only model

A new incept-only model was estimated to assess the partition of the variance in bicycle use across trip (Level 1), postal code (Level 2) and city level (Level 3). In this model, the person level variables are included as fixed effects, as explained in the previous model. Because of this, the person level is not included in the intercept-only model as this model still only evaluates the variation on different levels without including any variables. The three-level intercept-only model ensures that variation at the postal code and city levels can be assessed. The equation for this three-level model (Equation 9) is slightly different from the four-level intercept-only model equation (Equation 7) as it does not include the random intercept on person-level.

$$logit(p_{ij}) = \beta_0 + u_{0j} + v_{0ij} \qquad \text{Eq. (9)}$$

In which: p_{ij} = probability that a person in postal code *i* of city *j* uses a bicycle for a trip

β_0 = fixed intercept

u_{0j} = random intercept for city

v_{0ij}= random intercept for postal code

As can be seen from Table 11, the fixed-effect intercept was β_0= -0.586 (Std. Error = 0.112, p < .001), corresponding to an overall predicted probability of bicycle use of $p = \frac{e^{\beta_0}}{1+e^{\beta_0}} \frac{e^{-0.5859}}{1+e^{-0.5859}} \approx 0.358 = 35.8\%$. This seems reasonable as it does not differentiate much from the observed 37.1% bicycle use seen in the descriptive analysis. Table 11 also lists the random effect variances per analytical level of the intercept-only model. From this table, it can be concluded that the random effect variances for the postal code level ($\sigma^2_{PostalCode}$= 0.281, Std. Dev. = 0.530) and city level (σ^2_{City}= 0.132, Std. Dev. = 0.363) were moderate, indicating some sort of heterogeneity present across these levels.

Table 11 – Fixed effects intercept-only model

Random effect variances for each level				
Level		Name	Variance	Std. Dev.
Postal Code		v_{0ij}	0.2810	0.5301
City		u_{0j}	0.1315	0.3625
Fixed effects				
	Estimate	Std. Error	z value	Pr(>|z|)
β_0 (intercept)	-0.5859	0.1195	-4.902	9.47e-07***
Number of observations (Trips): 30,164, levels: Postal Code, 431; City, 10				
sign ***: <.001				

With these random effects, the intraclass correlation coefficients (ICCs) can be calculated (Eq. (10) and Eq. (11)). By doing so, the proportion of total variance accountable to each level can be assessed. As can be seen from the calculations below, the variance proportion at the postal code level is approximately 7.7% and 3.5% at city-level.

$$ICC_{City} = \frac{\sigma^2_{City}}{\sigma^2_{PostalCode}+\sigma^2_{City}+\sigma^2_{Residual}} = \frac{0.132}{0.281+0.132+3.290} = 0.0355 \qquad \text{Eq. (10)}$$

$$ICC_{PostalCode} = \frac{\sigma^2_{PostalCode}}{\sigma^2_{PostalCode}+\sigma^2_{City}+\sigma^2_{Residual}} = \frac{0.281}{0.281+0.132+3.290} = 0.0759 \qquad \text{Eq. (11)}$$

Although most of the variations are thus captured within trip and person level, these results still confirm that trips are not independent within postal codes and cities with even the ICC of postal codes being above the 0.05 threshold (Hox et al., 2017), justifying the use of a multilevel logistic model.

Additionally, the differences between cities can be indicated with some informative calculations using the different city-level random intercepts as shown in Table 12. For example, for Groningen (City = 14), the random intercept is 0.5063, meaning that the probability for cycling in Groningen can be calculated as follows:

$$p_{ij} = \frac{e^{\eta_{ij}}}{1+\ e^{\eta_{ij}}} = \frac{e^{\beta_0+u_{oj}}}{1+\ e^{\beta_0+u_{oj}}} = \frac{e^{-0.5859+\ 0.5063}}{1+\ e^{-0.5859+\ 0.5063}} \approx 0.4801 = 48.0\%\ \mathit{chance\ in\ bicycle\ use}$$

Doing the same for the city of Almere (City = 34) with a corresponding random intercept of -0.5903 shows the following change for a bicycle trip.

$$p_{ij} = \frac{e^{\eta_{ij}}}{1+\ e^{\eta_{ij}}} = \frac{e^{\beta_0+u_{oj}}}{1+\ e^{\beta_0+u_{oj}}} = \frac{e^{-0.5859-0.5903}}{1+\ e^{-0.5859-0.5903}} \approx 0.2356 = 23.6\%\ \mathit{in\ bicycle\ use}$$

Table 12 – Random intercepts per city and predicted bicycle use per city

City	City code	(Intercept) u_{0j}	Predicted probability p_{ij}
Groningen	14	0.5062	0.4801
Almere	34	-0.5903	0.2357
Nijmegen	268	0.3110	0.4317
Utrecht	344	0.4206	0.4588
Amsterdam	363	0.1282	0.3875
Den Haag	518	-0.2620	0.2999
Rotterdam	599	-0.4134	0.2691
Breda	758	0.0171	0.3615
Eindhoven	772	0.1194	0.3854
Tilburg	855	-0.1419	0.3257

The final conclusion of the intercept-only model is that following a four-level modelling approach would mean that the person level absorbs almost all variance in the model. This is due to the state of the data in which the persons are overrepresented in entries as compared to the higher levels postal code and city. Constructing a three-level modelling approach with trip-, postal code-, and city-level as analytical variables allowed for variance at higher levels. This is of importance as the resulting conclusions can inform municipalities on their role in influencing bicycle performance. Additionally, the model accounts for the real-world differences in bicycle use between Dutch cities, as reflected by the varying values of the random intercepts at the city level. All in all, the three-level model will provide a robust model which allows making conclusions and recommendations on environmental aspects and local mobility policies.

4.4 Final Multilevel Logistic Regression Model

The final three-level multilevel logistic regression model was estimated including all variables at the trip, person, postal code, and city level. The model also showed random intercepts for postal codes and cities. Comparing this model to the intercept-only model indicates a substantial improvement in the model fit, as can be seen from the model fit indicators in Table 13. The AIC and BIC values decreased substantially. Additionally, to test if the full three-level model including all predictors is statistically better than the intercept-only model a likelihood ratio test was conducted. The test indicated a significant improvement in model fit ($\chi^2(38)$ = 2,462.6, p < .001). This illustrates that the variables included in the model are significantly explaining the variance in bicycle use.

Table 13 – Model performance indicators

Model	AIC	BIC	logLik	-2*log(L)	df.resid
intercept-only model	38,319.0	38,343.9	-19,156.5	38,313.0	30,161
three level model	35,918.5	36,259.4	-17,918.2	35,836.5	30,114

Estimating the random effects in the final model showed reduced variance at the postal code level ($\sigma^2_{PostalCode}$= 0.168, Std. Dev. = 0.410) and even neglectable variance at the city level(σ^2_{City}= 1.165e-04, Std. Dev. = 0.008) (Table 14). This suggests that most of the heterogeneity on the postal code and city level is now captured in the model through the used variables. The intercept (β_0 = -1.062, Std. Error = 0.270, p< .001) corresponds to a baseline probability of bicycle use of 25.7% ($p = e^{-1.062}/(1 + e^{-1.062}) \approx 0.257$).

Table 14 – random effects three level logistic regression model

Random effects			
Groups	Name	Variance	Std. Dev.
Postal Code	(intercept)	0.168	0.410
City	(intercept)	1.165e-04	0.011
Number of obs (Trips): 30,155, groups: PostalCode, 429; City, 10			

A detailed overview of all fixed effect estimates is provided in Table 15. This overview shows that most variables on the trip level are significant at the 95% confidence level, except for the *Distance_gmc* (p=.507) variable, which is not significant. The variable Motive shows that all motives are leading to a decrease in probability of bicycle use as compared to the reference motive which is work related trips. Additionally, trip duration is a highly significant (p<.001) predictor of bicycle use in such a way that trips longer than the mean are associated with a lower probability of cycling, while shorter trips increase the probability.

All variables at the person level are significant at the 95% confidence level. Being female, non-Dutch, or having a driver's license is all associated with a lower probability of cycling, whereas being a student increases this probability. Additionally, higher age than the mean results in a lower probability of bicycle use. Most education levels are associated with higher cycling probability as compared to the reference category, which is no education. An exception to this is the highest category (HBO/University), which indicates slightly lower bicycle use probability.

Regarding the variables on the postal code level, several variables related to the diversity of the built environment significantly contribute to the probability of bicycle use. Higher percentages of residents born in the Netherlands (p<.001) and higher average housing values (p=.019) are associated with higher probabilities of cycling. Furthermore, increased shortest distances to transit stations significantly decrease the probability of bicycle use. Most other indicators regarding facility distances follow the expected direction of effect but are not statistically significant. An increase in the shortest distances to supermarkets and child care locations results in a lower probability for bicycle use, and if the shortest distance to a highway entry point increases, the probability for bicycle use decreases. However, none of these are statistically significant at the 95% confidence level.

For the hardware measures, the variables PerSepCyclingInfra_gmc (p=.012), CyclingInfraKm2_gmc (p=.004) and SepCyclingInfraKm2_gmc (p=.011) appear to be significant predictors of bicycle use probability. The results indicate that higher percentages of separated cycling infrastructure within a postal code, as well as greater concentration of cycling infrastructure, are associated with a lower probability of bicycle use. In contrast, a higher concentration of separated cycling infrastructure length increases the probability of bicycle use. Percentage-based measures produced patterns that at first appeared counterintuitive. For instance, higher shares of separated cycling infrastructure were associated with lower cycling probabilities. However, this is likely a compositional artefact of postal codes with very little absolute infrastructure length but high proportional values. When absolute infrastructure length was used instead, the results became clearer and more coherent. More total cycling infrastructure slightly reduced cycling probability, whereas more separated cycling infrastructure significantly increased it. This reinforces the importance of physically separated bicycle facilities and supports the argument that not all cycling infrastructure contributes equally to cycling behaviour. To avoid issues of multicollinearity, an additional check was conducted between these variables. The Pearson correlation coefficients ranged from r = -0.360 to r = 0.595, remaining well below the r = 0.800 threshold presented in Chapter 3. An overview of the Pearson correlation coefficients can be found in Table H2 in Appendix H. Given this, it is concluded that multicollinearity is not problematic in this case, and thus all variables remained in the model. It is important to note that all variables on the postal code level are grand mean centered, meaning that increases or decreases in these variables are interpreted relatively to their overall mean.

Table 15 – fixed effects three level logistic regression model

	Variable	Estimate	Std. Error	z value	p-value
	(Intercept)	-1.062E+00	2.698E-01	-3.938	<0.001***
Trip	Motive - Daily service/grocery	-1.034E+00	4.362E-02	-23.711	<0.001***
	Motive - Education	-1.862E-01	6.354E-02	-2.931	0.003**
	Motive - recreational/Social	-8.273E-01	4.209E-02	-19.656	<0.001***
	Motive - Others	-8.935E-01	5.208E-02	-17.156	<0.001***
	Distance_gmc	3.305E-04	4.983E-04	0.663	0.507
	Duration_gmc	-1.296E-02	9.923E-04	-13.055	<0.001***
Person	Gender - Woman	-5.747E-02	2.634E-02	-2.182	0.029*
	Age_gmc	-6.035E-03	9.456E-04	-6.382	<0.001***
	Origin - Outside Netherlands	-4.574E-01	2.965E-02	-15.427	<0.001***
	Student - Yes	3.412E-01	4.970E-02	6.865	<0.001***
	EducationLevel - Primary school	1.789E-01	4.457E-02	4.014	<0.001***
	EducationLevel - VMBO	3.410E-01	4.391E-02	7.768	<0.001***
	EducationLevel - HAVO/VWO	4.732E-01	5.308E-02	8.916	<0.001***
	EducationLevel - HBO/University	-5.166E-01	5.575E-02	-9.266	<0.001***
	Driverlicense - Yes	-4.268E-01	3.630E-02	-11.757	<0.001***
Environment	Inhabitants_gmc	-1.140E-05	6.708E-06	-1.699	0.089 .
	PerOwneroccupied_gmc	-1.802E-03	2.768E-03	-0.651	0.515
	PerBornNL_gmc	1.056E-02	2.510E-03	4.208	<0.001***
	AverageHHsize_gmc	-1.372E-01	1.436E-01	-0.956	0.339
	AverageHousingValue_gmc	5.835E-04	2.485E-04	2.348	0.019 *
	ClosestSupermarket_gmc	-8.242E-02	1.061E-01	-0.777	0.437
	ClostestChildCare_gmc	-3.077E-01	1.834E-01	-1.677	0.093 .
	ClosestHighwayEntry_gmc	4.114E-02	2.764E-02	1.488	0.137
	ClosestTransitStation_gmc	-5.113E-02	1.444E-02	-3.540	<0.001***
	ClosestPrimarySchool_gmc	1.020E-02	1.194E-01	0.085	0.932
	ClosestSecondarySchool_gmc	2.560E-02	4.316E-02	0.593	0.553
	ClosestPharmacy_gmc	1.457E-01	8.931E-02	1.631	0.103
	ClosestGPCenter_gmc	2.022E-03	1.659E-02	0.122	0.903
	AddressDensity_gmc	1.053E-05	2.052E-05	0.513	0.608
	PerCyclingInfra_gmc	-2.271E-03	4.008E-03	-0.566	0.571
	PerSepCyclingInfra_gmc	-1.870E-02	7.404E-03	-2.526	0.012 *
	PerQuality_gmc	-1.317E-03	1.597E-03	-0.825	0.409
	CyclingInfraKm2_gmc	-3.666E-02	1.260E-02	-2.909	0.004 **
	SepCyclingInfraKm2_gmc	9.762E-02	3.832E-02	2.548	0.011 *
	Streetconnectivityperm_gmc	-1.999E-03	1.177E-03	-1.698	0.090 .
City	Softwarescore	2.296E-02	2.494E-02	0.921	0.357
	Orgware - Organisation	6.953E-02	1.971E-02	3.527	<0.001***
	Orgware - Collaboration	4.125E-03	1.604E-02	0.257	0.797
	Signif. Codes: p<.001'***', p<.01'**', p<.05'*', p<.1'.'				

Finally, at the city level, the role of local policies is represented in three variables, of which the orgware - organisational structure variable is significantly positive (β= 0.070, p<.001) related to the bicycle use. An one unit increase in this variable corresponds to an odds ratio of 1.072, meaning that the odds of choosing the bicycle rather than another mode increase by 7.2%. Furthermore, the same one unit increase leads to a 1.3% increase in probability in bicycle use as compared to the baseline probability of 25.7%. The orgware organisational structure variable is constructed out of four different subjects related to the formulation of policy goals, the implementation of these policy measures, the associated financial resources and the policy consistency. This means that realizing higher scores on these subjects can effectively positively impact bicycle use. If for example a random city scores one point higher on all of these four subjects, the odds ratio corresponds to 1.320, which is a 32.0% increase in odds. In comparison to the baseline predicted bicycle use of 25.7%, this means that the predicted probability of cycling will increase to 31.3%. This is a notable increase to keep in mind for municipalities who strive to improve their cycling performance.

With the above listed fixed-effect estimates, more calculations regarding predicted probabilities of bicycle use can be made. For example, a trip 10 minutes shorter than the mean duration (19.51 minutes) corresponds to a log-odds increase of -0.1296 (10*-0.01296), which results in an $e^{0.1296}$ = 1.138 = 13.8% change in odds for cycling. This results in a predicted probability of bicycle use of approximately 28.2% ($e^{-1.062 + 0.130}$/(1+ $e^{-1.062+ 0.130}$ ≈ 0.282), compared to 25.7% baseline probability at the intercept.

A full overview of the code used in R for reproducibility purposes can be found in Appendix I. This includes the multicollinearity GVIFS, both intercept-only models, and the final model.

Chapter 5 – Discussion

The objective of this research was to identify the factors that explain differences in cycling performance between Dutch cities. The focus was especially put on assessing the extent to which local mobility policies contribute to these differences, as it is a significant current research gap. Cycling performance was expressed as the mode share of inner-city trips made by bicycle. The reason for this is that the mode share is the most commonly used indicator and gives a comprehensive overview, as it illustrates the importance of the overall transport system (Rietveld & Daniel, 2004; Harms et al., 2015). A multilevel logistic regression model was applied to account for the nested structure of the data and explain the variation on the different analytical levels. This approach allowed not only for the assessment of the effect of variables on different levels, but also allowed for explaining the interplay between the different levels and between groups at the same level. Due to the logistic manner of the model caused by the dependent variable, which indicates whether the bicycle is taken for a trip or not, the model predicts the probability that a trip is made by the bicycle under the circumstances included in the model.

The results show that variation in bicycle use exists at trip-, person-, postal code-, and city-levels, but is strongly dominated on the person level. The three-level intercept-only model showed moderate postal code and city-level variance. In the full model, postal code and city level variance became almost neglectable, indicating that most between-city and postal code differences can be explained by the variables included in the model. The remainder of this section follows the structure of the sub-questions stated in Chapter 1.

Sub-question 1: *To what extent do trip-, personal- and household-characteristics influence cycling performance?*

The first sub-research question was to define the extent to which trip and personal characteristics influence cycling performance. As explained above, much variation in the model happens on the trip and person level, directly answering a part of the question. Diving deeper into which characteristics are of importance, it can be concluded that on the trip level, especially the trip duration was found to be one of the most powerful predictors. In such a way that shorter trips (as supposed to the grand mean) substantially increased the likelihood of cycling, while longer trips significantly reduced this. This is in line with previous research which suggests that cycling is primarily used for relatively short trips (<5km) (Goel et al., 2021; KiM, 2023; Pucher and Buehler, 2008a). Furthermore, the motive of a trip also effects the bicycle use. In such a way that non-work related trips show lower cycling probabilities than work-related trips. Additionally, educational trips show the most comparison to work related trips. These results are somewhat contrary to prior literature, which indicates that recreational trips account for a large proportion of bicycle trips in the Netherlands (KiM, 2023; Goel et al., 2021). However, the findings are consistent with evidence showing that cycling remains an important mode for short commuting trips by KiM (2023).

The second part of the first sub-research question is related to the importance of person and household characteristics for cycling behaviour. It should be noted that household characteristics are included at the person level in the model, as with the privacy restrictions present in the data source, it was not possible to link persons to their households. Moreover, this nested structure would likely not occur given the random sampling procedure used in the data source used. It can be concluded that this effect is substantial as the most variance in the model happens at person level. Furthermore, all included variables on the person level showed a significant relationship with bicycle use probabilities. Being a female, of non-Dutch origin, older aged or holding a driver's licence reduced the probability of bicycle use, whereas students were more likely to take the bicycle. These effects largely align with existing literature, although some relationships appear weaker than in existing literature. The modest effect of gender indicated alignment with the widespread view in existing literature in which gender equality is frequently observed in high cycling countries (such as the Netherlands) (Pucher & Buehler, 2008a; Pucher & Buehler, 2008b;

Aldred et al., 2015; Goel et al., 2021). The negative effect of increased age corresponds with Goel et al (2021), who showed higher cycling rates among younger generations in high-cycling countries. Although the model is not capturing the U -U-shaped distribution of bicycle use over individual age, due to the generalized linear model structure, which is suggested by other literature (Pucher & Buehler, 2008a; Pucher & Buehler, 2008b). The positive effect of Dutch origin aligns with existing literature by Hausstein (2020). Furthermore, the positive effect of student occupation also matches the existing literature, which all report higher cycling levels among students (Nelson & Allen, 1997; Ryley, 2006; Wu et al., 2024). Subsequently, the effect of holding a driver's license showed expected results, although the results are slightly smaller than reported in existing literature. The model shows a 7.5% drop in bicycle use probabilities compared to a reported 10% drop in bicycle use, explained by KiM (2023). This could subsequently indicate that the role of the car diminishes in urban areas as this study focusses on the urban scale rather than the whole country (done by KiM, 2023).

Sub-question 2: *To what extent do features of the built environment and transport system affect cycling performance?*

The second sub-research question is related to the characteristics in the built environment and the transport system, which found to be of importance for cycling performance. Urban planners and policy makers should take this into account suggesting that integrated urban (mobility) planning is needed. However, the postal-code level contributed substantially less to overall variation in cycling performance as compared to the trip and person levels. Hence, several characteristics related to diversity and design components of the built environment still influence bicycle use. Increasing the values for indicators related to the socio-demographic diversity such as the percentage of residents born in the Netherlands and the average housing value indicated increased probability of bicycle use. Variables indicating the diversity of the urban composition by indicating the accessibility to different facilities most showed expected estimate effect directions, with shorter distances increasing cycling probability. This aligns with extensive existing literature explaining that more diverse landscapes with a higher mixture of functions at shorter distances positively influence bicycle uptake (Cervero & Duncan, 2003; Pucher & Bucher, 2006; Pucher & Bucher, 2008a; Hankey et al., 2012; Chen et al., 2017; Heinen et al., 2009). However, none of these variables were statistically significant at acceptable confidence intervals (95% or higher), limiting the explanatory power of this.

The accessibility indicators to other transport modes showed expected direction effects. The significantly positive relationship of higher transit accessibility resulting in higher probability of bicycle use suggests that the bicycle can play an important role in first- and last-mile operations for public transport, widely acknowledged in existing research (Pucher & Buehler, 2008a; Harms et al., 2015; van Kuijk et al., 2022). In contrast, but as expected, easier access to car infrastructure pulls users towards the car, resulting in less cycling. This is in line with existing literature describing that reduced car accessibility can lead to increases in cycling rates (Cervero and Duncan, 2003; Dill and Carr, 2003; Heinen et al., 2010). However, this variable is not statistically significant at acceptable confidence intervals, meaning no effect conclusions can be made on this variable.

Regarding the hardware measures, several infrastructure related variables showed significant but in first glance contrary associations with bicycle use. Contrary to existing literature (CPB, 2025; Dill & Carr, 2003; Berghoefer and Vollrath, 2008; Vedel et al., 2017; Wu et al., 2024), is that a higher percentage of separated cycling infrastructure was associated with a lower probability of bicycle use. The effect is likely to be overrepresented by postal codes with little absolute infrastructure length but with high proportion of separated bicycle infrastructure. Therefore, the absolute infrastructure variables provide a more interpretable outcome. These variables indicate that higher total cycling infrastructure concentration results in lower cycling probability, whereas higher separated cycling infrastructure concentration increases it. The focus should thus be on providing a substantial length of separated cycling infrastructure, a conclusion also widely accepted by existing literature (CPB, 2025; Dill & Carr, 2003; Berghoefer and Vollrath, 2008; Vedel et al., 2017; Wu et al., 2024). Furthermore, it shows that not all types of infrastructure contribute equally to cycling behaviour. This finding aligns strongly with existing literature which highlights the role of perceived safety in mode choice and the role of separated cycling infrastructure in this (Fishman et al., 2012; Berghoefer & Vollrath, 2023; Uijtdewilligen et al., 2024). The other infrastructural related variables representing street connectivity and infrastructure quality did not seem to have any significant effect on bicycle use.

Sub-question 3: *How can the impact of local cycling policies on cycling performance be assessed, and what is their effect?*

One of the key contributions of this research is the assessment of local mobility policies and is related to the third sub-research question. The impact of local cycling policies can be assessed by transforming insights gained from semi-structured interviews into quantifiable variables. The eventual outcome of the model revealed that on the local policy level, especially the organizational structure of municipalities has a significant positive effect on cycling performance. Each additional unit increase in this variable increased the odds of choosing the bicycle by approximately seven percent. This underscores the importance of a dedicated cycling-flavored organizational structure in municipalities in order to let cycling flourish. As the organizational structure is reflected by the formulation of policy goals, the implementation of these policy measures, the associated financial resources, and the policy consistency, these elements are hands-on subjects to improve. In contrast, software measures and collaboration indicators showed no significant effect in this model. This is not surprising, as capturing these effects is hard to quantify. This is in line with the qualitative results from the interviews, in which multiple interviewees mentioned that the direct effect of software and collaboration measures on cycling use remains generally undefinable.

Hence, the qualitative interviews revealed a more complex picture, which the model was not fully capable of capturing. Interviewees emphasized that a constructive collaboration between municipalities and local departments of the Fietsersbond is definitely highly valuable, as both organizations can strengthen one another through the complementary roles they fulfil. The municipalities have the professional expertise and resources, while the Fietsersbond provides bottom-up insights with local knowledge and an activism-driven pressure. Furthermore, current departments of the Fietsersbond often have volunteers who are operative in the professional mobility sector as well, leading to well-informed discussions with municipalities. Together, these top-down and bottom-up perspectives can reinforce each other, leading to an improvement in cycling performance, although not directly specified by the model. On top of that, in cities in which the collaboration between the two organizations has been assessed to be very positive from both sides, the underlying condition found to be very important was transparency.

Chapter 6 - Limitations

Despite the contributions of this research, several limitations should be acknowledged. First, one of the conclusions from the literature is that the effect on cycling performance on the environmental level can be divided into the natural environment and the built environment. The natural environment is not included in this model, as most of these factors are less critical in the Dutch context, because the geospatial size is rather small. This leads to almost universal weather conditions between cities. However, excluding these, according to the literature, relevant topics such as weather conditions could lead to possible omitted variable bias.

A second limitation of the study is related to the quality and completeness of the data, as part of the infrastructure data is derived from the Fietsersbond route planner dataset. This data source is partly based on volunteer contributions, leading to an objective interpretation of, for example, when a road segment is of good quality. This could introduce under- or overestimation of the infrastructure variables due to incomplete or incorrect reporting, outdated data, or overrepresentation due to targeted data collection in certain cities.

Several limitations also need to be taken into account for the multilevel model used in this study. Most variation is currently captured at the trip and person level, largely because these levels contain far more groups than the postal code or city level. According to Hox et al. (2017), multilevel models generally operate better with a large number of higher-level groups (n > 30), ensuring reliable estimation. The threshold recommended by Hox et al. (2017) is not met, as only ten cities are included. Still, the estimates provide helpful insights and form a basis to be expanded upon in further research. Additionally, the choice of using the residential postal code to define the analytical level for the built environment can be discussed as departure and arrival locations also play a role in travel behaviour (Axhausen, 2007).

A further limitation that needs to be addressed concerns the qualitative analysis component. The study aimed to conduct twenty interviews, but four interviews are missing. This results in insights on these cities from only one perspective, being either from the municipal perspective or from the local department of the Fietsersbond. Perspectives between these organisations can differ substantially, meaning that missing the second opinion may lead to an incorrect interpretation of local cycling policy measures. Additionally, own-city bias could be present in the interviewees. Given the small number of cities in the study, even small biases could affect the interpretation of the role of local cycling policies, especially if both interviews have not been conducted for a city.

Finally, the generalizability of the findings is limited to the Dutch national urban scale. The Netherlands is a high-cycling country with a long established cycling culture, which could influence cycling behaviour heavily (Haustein et al., 2020). Moreover, cycling in the Netherlands is currently undergoing rapid changes due to emerging cycling trends, such as changing speeds at the cycling infrastructure due to all types of e-bikes, such as the speed pedelec and fat bike. Additionally, in all cities, the capacity of the existing cycling infrastructure is under pressure, partly because of this. Furthermore, cities are implementing new mobility policies, which are also not incorporated in the model. A good example of this is the introduction of a universal speed limit of 30km/u for almost the whole city center of Amsterdam (Gemeente Amsterdam, 2025). As the data mainly originates from 2023, all these trends are not (fully) incorporated in the model.

Chapter 7 – Recommendations

This chapter presents recommendations based on the findings of this research. It is divided into two sections. The first section formulates recommendations for policymakers seeking to improve cycling performance in Dutch cities. These result from both the multilevel model outcome and the qualitative results from the interviews. The second section provides recommendations for future research based on analytical outcomes and limitations discussed in the previous chapter.

The findings of the study indicate that the orgware component plays a significant role in bicycle use, especially the organizational structure of local municipalities. Policymakers who seek to improve cycling performance in their city should focus on developing measurable and ambitious cycling policies that can be monitored and acted upon. Additionally, these policies should be enhanced by concrete action plans, structural financial resources, and a stable political climate acknowledging the pivotal role of the bicycle in the urban mobility system on both the short and long term. Political consistency acts as a crucial underlying condition for realising organisational improvements in the other organizational measures. Municipalities could increase cycling performance substantially if these conditions are in place and should thus invest in these conditions.

Additionally, policymakers should also focus on hardware measures by expanding the separated cycling facilities as a higher concentration of these will result in more bicycle uptake. Furthermore, urban development should prioritize compact and mixed-use neighbourhoods, which allow for short, bikeable distances to facilities. Moreover, local authorities should also take into consideration the push and pull effects regarding other modes of transportation. The bicycle could be complementary for public transport trips being an ideal mode for first and last mile operations. Contrary to this, easier access to highway entries decreases bicycle use probabilities, reflecting the push and pull factors of the car. Local policymakers should consider these dynamics with other modes in integrated mobility policies, creating push factors to the car, so these movements switch to other modes, such as active modes.

The main recommendation for further research is to include more cities in the research as more higher-level groups improve the multilevel model estimation reliability. This includes additional data gathering and conducting even more interviews, but gives more robust and valuable insights on urban cycling dynamics and the role of local policies on an even broader scale.

Additionally, further research is recommended to integrate the most up-to-date data due to the rapidly evolving cycling landscape, as described in the limitations chapter. Furthermore, additional hardware measures, such as the proportion of 30km/h roads in cities, could be included for example. Even more exploiting the multilevel model can be done by including cross effects in the model, explaining higher-level variable effects on lower-level effects. This allows for even more in-depth insights into the dynamics of urban cycling policies.

Another recommendation for further research could be considering another analysis approach, which could be for example a structural equation model. This method could provide deeper insights into the underlying relationships between different variables, as it examines both direct and indirect relationships between variables simultaneously. It can reveal complex causal pathways, offering deeper insights into the complex dynamics of this topic. All in all, this research could act as a basis for further research in this topic, in which expanding the geospatial scale is especially recommended in order to increase model performance and create more stable conclusions.

Chapter 8 – Conclusion

Overall, this research demonstrates that cycling performance in Dutch cities and the role of local cycling policies are shaped by a complex system of factors acting on interacting different scale levels. While trip- and person-characteristics account for much of the variation between cities, this study shows that a crucial role has been set aside for local authorities by showing that hardware and orgware measures significantly affect bicycle use. The findings related to policy measures at the city level expand the personal- and infrastructure-focused perspective in existing literature by proving that local governance structure is a key component for influencing cycling behaviour, even in high-cycling countries. In doing so, this research addresses a notable literature gap by providing an in-depth analysis of differences between urban areas within a high-cycling country, whereas most prior studies have focused on cross-country comparisons.

The answer to the main research question: *What factors explain the difference in cycling performance across Dutch cities, and to what extent do local cycling policies contribute to these differences?* Indicates that the differences can mainly be explained by characteristics at the trip and person level. However, the study also identified notable characteristics at the environmental and city level. From a hardware perspective, investments in separated cycling infrastructure and an integrated multimodal planning, particularly with high transit accessibility, yield the greatest gains. From an orgware perspective, a stable organisational structure is essential for municipalities aiming to improve their cycling performance. This requires clearly formulated policy goals and effective implementation of these policy goals through concrete action plans. It furthermore requires the assurance of structural financial resources and a consistent policy context in which municipalities can work. Furthermore, the insights from the interviews showed that there is a high willingness for a constructive and good collaboration from both the representatives from municipalities and the local department of the Fietsersbond with each other. In cities that indicate that there is already a good constructive collaboration present, indicate that the underlying condition of this is transparency. Cities that report having strong, constructive collaboration emphasize that transparency is the key underlying condition. This provides valuable insight for municipalities in which the collaboration can still be improved. Beyond its academic contribution, this research has a clear societal relevance. It provides actionable insights as listed above for policymakers who seek to strive for higher cycling performance in their cities. Such measures contribute to a more sustainable urban mobility system.

Chapter 9 – References

Ahmed, S. K. (2024). How to choose a sampling technique and determine sample size for research: A simplified guide for researchers. *Oral Oncology Reports, 12*, 100662. https://www.sciencedirect.com/science/article/pii/S2772906024005089

Aldred, R., Woodcock, J., & Goodman, A. (2015). Does More Cycling Mean More Diversity in Cycling? *Transport Reviews*, *36*(1), 28–44. https://doi.org/10.1080/01441647.2015.1014451

Axhausen, K. (2007). Concepts of travel behavior research. Retrieved from: https://www.researchgate.net/publication/237262766_Concepts_of_Travel_Behavior_Research

Bamberg, S., Fujii, S., Friman, M., & Gärling, T. (2011). Behaviour theory and soft transport policy measures. *Transport Policy, 18*(1), 228–235. https://doi.org/10.1016/j.tranpol.2010.08.006

Bates, J., & Leibling, D. (2012). Spaced out. Perspectives on parking policy. RAC foundation. Retrieved from: https://www.racfoundation.org/wp-content/uploads/2017/11/spaced_out-bates_leibling-jul12.pdf

Ben-Shachar, M.S., Patil, I., Thériault, R., Wiernik, B.M., & Lüdecke, D. (2023). Phi, Fei, Fo, Fum: Effect Sizes for Categorical Data That Use the Chi-Squared Statistic. *Mathematics* **2023**, *11*, 1982. https://doi.org/10.3390/math11091982

Berghoefer, F. L., & Vollrath, M. (2023). Motivational and deterrent effects of route attributes in cyclists' route choice. *Transportation Research Part F: Traffic Psychology and Behaviour, 95*, 343–354. https://doi.org/10.1016/j.trf.2023.04.003

Bergström, A., & Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. *Transportation Research Part A: Policy and Practice, 37*(8), 649–666. https://doi.org/10.1016/S0965-8564(03)00012-0

Blitz, A. (2021). How does the individual perception of local conditions affect cycling? An analysis of the impact of built and non-built environment factors on cycling behaviour and attitudes in an urban setting. *Travel Behaviour and Society. Volume 25. Pages 27-40.* https://doi.org/10.1016/j.tbs.2021.05.006.

Böcker, L., Dijst, M., & Prillwitz, J. (2013). Impact of everyday weather on individual daily travel behaviours in perspective: A literature review. *Transport Reviews, 33*, Article 747114. https://doi.org/10.1080/01441647.2012.747114

Centraal Bureau voor de Statistiek. (2023). Voorlopige bevolkingsaantallen per gemeente, 1-1-2023. Retrieved from: https://www.cbs.nl/nl-nl/maatwerk/2023/09/voorlopige-bevolkingsaantallen-per-gemeente-1-1-2023

Centraal Bureau voor de Statistiek. (2024a). Pensioenleeftijd werknemers nadert 66 jaar. Centraal Bureau voor de Statistiek. Retrieved from: https://www.cbs.nl/nl-nl/nieuws/2024/18/pensioenleeftijd-werknemers-nadert-66-jaar

Centraal Bureau voor de Statistiek. (2024b). Onderzoek onderweg in Nederland – ODiN 2023. [Dataset]. Data Archiving and Networked Services (DANS). https://doi.org/10.17026/SS/FNXJEU

Centraal Bureau voor de Statistiek. (2024c). Onderweg in Nederland (ODiN) 2023 – Onderzoeksbeschrijving. Retrieved from: https://www.cbs.nl/nl-nl/longread/rapportages/2024/onderweg-in-nederland--odin---2023-onderzoeksbeschrijving/4-steekproeftrekking-odin-2023

Centraal Bureau voor de Statistiek. (2025a). Kerncijfers per postcode. Retrieved from: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

Centraal Bureau voor de Statistiek. (2025b). 2025-cbs_pc4_2023_v2 [Dataset]. https://download.cbs.nl/postcode/2025-cbs_pc4_2023_v2.zip

Cervero, R., & Duncan, M. (2003). Walking, bicycling, and urban landscapes: Evidence from the San Francisco Bay Area. *American Journal of Public Health, 93*(9), 1478–1483. https://doi.org/10.2105/AJPH.93.9.1478

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment, 2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Charreire, H., Roda, C., Feuillet, T., Piombini, A., Bardos, H., Rutter, H., Compernolle, S., Mackenbach, J. D., Lakerveld, J., & Oppert, J. M. (2021). Walking, cycling, and public transport for commuting and non-commuting travels across 5 European urban regions: Modal choice correlates and motivations. *Journal of Transport Geography, 96*, 103196. https://doi.org/10.1016/j.jtrangeo.2021.103196

Chen, P., & Sun, F. (2017). Built environment determinants of bicycle volume: A longitudinal analysis. *Journal of Transport and Land Use, 10*, Article 892. https://doi.org/10.5198/jtlu.2017.892

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

CPB (2025). De effecten van fietsinfrastructuur op wonen, werken en reizen. Centraal Planbureau. Retrieved from: https://www.cpb.nl/de-effecten-van-fietsinfrastructuur-op-wonen-werken-en-reizen

Cran.R (2024). Software R 4.4.1. https://cran.r-project.org/bin/windows/base/old/4.4.1/

Cremer-Schulte, M., van Wee, B., & Heinen, E. (2024). Stuck in the driver's seat: a conceptualisation for understanding car dependence and its determinants. *Transport Reviews*, *45*(2), 173–193. https://doi.org/10.1080/01441647.2024.2430007

de Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the health benefits of cycling outweigh the risks? *Environmental Health Perspectives, 118*(8), 1109–1116. https://doi.org/10.1289/ehp.0901747

Decisio. (2025). 30 km per uur als norm. De maatschappelijke effecten op nationale schaal. Retrieved from: https://www.fietsberaad.nl/getmedia/75cd0178-0ebe-4633-bc5c-a04ef00bb756/Decisio-Rapport-30-km-per-uur-als-norm-Opdrachtgever-Fietsersbond.pdf.aspx

Den Dulk CJ, Van De Stadt H, Vliegen JM. Een nieuwe maatstaf voor stedelijkheid: de omgevingsadressendichtheid [A new measure for degree of urbanization: the address density of the surrounding area]. Maandstat Bevolking. 1992 Jul;40(7):14-27. Dutch. PMID: 12285285. Retrieved from: https://pubmed.ncbi.nlm.nih.gov/12285285/

DeVellis, R.F. (2017). Scale development : theory and applications (4th ed.). Retrieved from: https://tms.iau.ir/file/download/page/1635238305-develis-2017.pdf

Dill, J., & Carr, T. (2003). Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them. Transportation Research Record, 1828(1), 116-123. https://doi.org/10.3141/1828-14

Dill, J., & Voros, K. (2007). Factors affecting bicycling demand: Initial survey findings from the Portland, Oregon, region. *Transportation Research Record, 2031*(1), 9–17. https://doi.org/10.3141/2031-02

European Union. (2024). Europese verklaring over fietsen. Retrieved from: https://eur-lex.europa.eu/legal-content/NL/TXT/PDF/?uri=OJ:C_202402377

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Faber, K., Kingham, S., Conrow, L., & van Lierop, D. (2023). Differences in walking and cycling between professional immigrants and comparable ethnic Dutch: A quantitative analysis from the Netherlands. *Transportation Research Interdisciplinary Perspectives, 21*, 100915. https://doi.org/10.1016/j.trip.2023.100915

Fietsersbond. (2019). Nederland fietsland. Retrieved from: https://www.fietsersbond.nl/nieuws/hoe-hard-wij-fietsen-en-nog-veel-meer-interessante-fietsweetjes/

Fishman, E., Schepers, P., Barbara, C., & Kamphuis, M. (2015). Dutch cycling: Quantifying the health and related economic benefits. *American Journal of Public Health, 105*(8), e13–e15. https://doi.org/10.2105/AJPH.2015.302724

Fishman, E., Washington, S., & Haworth, N. (2012). Understanding the fear of bicycle riding in Australia. *Journal of the Australasian College of Road Safety, 23*, 19–27. Retrieved from: https://www.researchgate.net/publication/261983538_Understanding_the_fear_of_bicycle_riding_in_Australia

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association, 87*(417), 178–183. https://doi.org/10.2307/2290467

Gärling, T. (2005). Changes of private car use in response to travel demand management. In *Travel demand management and modal split.* Elsevier. https://doi.org/10.1016/B978-008044379-9/50200-4

Garrard, J., Rose, G., & Lo, S. K. (2008). Promoting transportation cycling for women: The role of bicycle infrastructure. *Preventive Medicine, 46*(1), 55–59. https://doi.org/10.1016/j.ypmed.2007.07.010

Gatersleben , B., & Appleton, K.M. (2007). Contemplating cycling to work: attitudes and perceptions in different stages of change . *Transportation Research Part A* , 41 (4) : 302 – 312 . https://doi.org/10.1016/j.tra.2006.09.002

Gemeente Amsterdam (2017). Meerjarenplan. Retrieved from: https://bicycleinfrastructuremanuals.com/manuals4/GemeenteAmsterdam-MeerjarenplanFiets_2017-2022_Nederlands.pdf.pdf

Gemeente Amsterdam. (2025). 30km/u in de stad, onderozeksrapport. Retrieved from: https://openresearch.amsterdam/nl/page/124453/onderzoeksrapport-30-km-u-in-de-stad

Gliem, J. A., & Gliem, R.R. (2003). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Paper presented at the 2003 Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education. Retrieved from: https://scholarworks.indianapolis.iu.edu/server/api/core/bitstreams/976cec6a-914f-4e49-84b2-f658d5b26ff9/content

Goel, R., & Mohan, D. (2020). Investigating the association between population density and travel patterns in Indian cities—An analysis of 2011 census data. *Cities, 100*, 102656. https://doi.org/10.1016/j.cities.2020.102656

Goel, R., Goodman, A., Aldred, R., Nakamura, R., Tatah, L., Garcia, L. M. T., & Woodcock, J. (2021). Cycling behaviour in 17 countries across 6 continents: levels of cycling, who cycles, for what purpose, and how far? *Transport Reviews*, *42*(1), 58–81. https://doi.org/10.1080/01441647.2021.1915898

Goldstein, H. (2011). *Multilevel statistical models*. John Wiley & Sons. Retrieved from: https://stats.oarc.ucla.edu/wp-content/uploads/2016/02/goldstein-1.pdf

Götschi, T., Garrard, J., & Giles-Corti, B. (2015). Cycling as a Part of Daily Life: A Review of Health Perspectives. *Transport Reviews*, *36*(1), 45–71. https://doi.org/10.1080/01441647.2015.1057877

Graham-Rowe, E., Skippon, S., Gardner, B., & Abraham, C. (2011). Can we reduce car use and, if so, how? A review of available evidence. *Transportation Research Part A: Policy and Practice, 45*(5), 401–418. https://doi.org/10.1016/j.tra.2011.02.001

Guo, J., Bhat, C., & Copperman, R. (2007). Effect of the built environment on motorized and nonmotorized trip making: Substitutive, complementary, or synergistic? *Transportation Research Record, 2010*(1), 1–11. https://doi.org/10.3141/2010-01

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., & Xu, Z. (2012). Estimating use of non-motorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN. *Landscape and Urban Planning, 107*(3), 307–316. https://doi.org/10.1016/j.landurbplan.2012.06.005

Harms, L., Bertolini, L., & Brömmelstroet, M. T. (2015). Performance of Municipal Cycling Policies in Medium-Sized Cities in the Netherlands since 2000. *Transport Reviews*, *36*(1), 134–162. https://doi.org/10.1080/01441647.2015.1059380

Haustein, S., Kroesen, M., & Mulalic, I. (2020). Cycling culture and socialisation: Modelling the effect of immigrant origin on cycling in Denmark and the Netherlands. *Transportation, 47*(4), 1689–1709. https://doi.org/10.1007/s11116-019-09978-6

Heinen, E., van Wee, B., & Maat, K. (2009). Commuting by bicycle: an overview of the literature. https://doi.org/10.1080/01441640903187001

Hoffman, L. (2013). *Applied multilevel models for longitudinal and clustered data* [PowerPoint slides]. QIPSR Workshop, University of Kentucky. Retrieved from: https://qipsr.as.uky.edu/sites/default/files/QIPSR_2013_Packet.pdf

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications* (3rd ed.). Routledge. https://doi.org/10.4324/9781315650982

IBM. (2022). Downloading IBM SPSS Statistics 28. Retrieved from: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-28

Kearney, M. (2017). Cramér's V. In *The SAGE Encyclopedia of Communication Research Methods* (Chapter 107). https://doi.org/10.4135/9781483381411.n107

Kettle, R., Crombie, H., & O'Rourke, D. (2017). *NG70 Air pollution: Outdoor air quality and health (NICE Guideline)*. https://doi.org/10.13140/RG.2.2.30955.66084

KiM Netherlands Institute for Transport Policy Analysis. (2023). Cycling facts 2023. Ministry of Infrastructure and Water Management. Retrieved from: https://english.kimnet.nl/publications/publications/2024/01/10/cycling-facts-2023

Kim, S., & Ulfarsson, G. (2008). Curbing automobile use for sustainable transportation: Analysis of mode choice on short home-based trips. *Transportation, 35*(6), 723–737. https://doi.org/10.1007/s11116-008-9177-5

Kitamura, R., Mokhtarian, P., & Daidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation, 24*(2), 125–158. https://doi.org/10.1023/A:1017959825565

Krizek, K., Johnson, P. J., & Tilahun, N. (2005). Gender differences in bicycling behavior and facility preferences. *Research on Women's Issues in Transportation*, 2, 31–40. Retrieved from: https://www.researchgate.net/publication/235909764_Gender_differences_in_bicycling_behavior_and_facility_preferences

Kutner, M. H., Nachtsheim, C., Neter, J., & Li, W. (2004). *Applied linear statistical models* (5th ed.). McGraw-Hill/Irwin. Retrieved from: https://users.stat.ufl.edu/~winner/sta4211/ALSM_5Ed_Kutner.pdf

Li, Z., Wang, W., Liu, P., & Ragland, D. R. (2012). Physical environments influencing bicyclists' perception of comfort on separated and on-street bicycle facilities. *Transportation Research Part D: Transport and Environment, 17*(3), 256–261. https://doi.org/10.1016/j.trd.2011.12.001

Methorst, R., Monterde-i-Bort, H., Risser, R., Sauter, D., Tight, M., & Walker, J. (2010). *PQN final report: Pedestrians' quality needs – Final report*. ISBN 978-0-9566903-0-2. https://swov.nl/en/publicatie/cost-358-pedestrians-quality-needs-pqn-final-report-part-c-executive-summary

Vernez Moudon, A., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., & Weather, R. D. (2005). Cycling and the built environment: A US perspective. *Transportation Research Part D: Transport and Environment, 10*(3), 245–261. https://doi.org/10.1016/j.trd.2005.04.001

Mueller, N., Rojas-Rueda, D., Cole-Hunter, T., de Nazelle, A., Dons, E., Gerike, R., Götschi, T., Int Panis, L., Kahlmeier, S., & Nieuwenhuijsen, M. (2015). Health impact assessment of active transportation: A systematic review. *Preventive Medicine, 76*, 103–114. https://doi.org/10.1016/j.ypmed.2015.04.010

Nijland, H. (2017). Fietsen leidt tot langer en gezond leven. PBL. Retrieved from: https://www.pbl.nl/publicaties/fietsen-leidt-tot-langer-en-gezond-leven

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Osborne, J., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation, 9*(6). https://doi.org/10.7275/b2q66q2t

Park, S., & Park, J. (2022). Multilevel Mixed-Effects Models to Identify Contributing Factors on Freight Vehicle Crash Severity. Sustainability, 14(19), 11804. https://doi.org/10.3390/su141911804

Parkin, J., Ryley, T., & Jones, T. (2007). Barriers to cycling: An exploration of quantitative analyses. In *Civil engineering: Book chapters*. Retrieved from: https://www.researchgate.net/publication/30502848_Barriers_to_Cycling_An_Exploration_of_Quantitative_Analyses

Parkin, J., Wardman, M., & Page, M. (2008). 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

Pawluk De-Toledo, K., O'Hern, S., & Koppel, S. (2022). Travel behaviour change research: A scientometric review and content analysis. *Travel Behaviour and Society, 28*, 141–154. https://doi.org/10.1016/j.tbs.2022.03.004

Pucher, J., & Buehler, R. (2006). Why Canadians cycle more than Americans: A comparative analysis of bicycling trends and policies. *Transport Policy, 13*(3), 265–279. https://doi.org/10.1016/j.tranpol.2005.11.001

Pucher, J., & Buehler, R. (2008a). Cycling for everyone: Lessons from Europe. *Transportation Research Record: Journal of the Transportation Research Board, 2074*, 58–65. https://doi.org/10.3141/2074-08

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

Pucher, J., Buehler, R., & Seinen, M. (2011). Bicycling renaissance in North America? An update and re-appraisal of cycling trends and policies. *Transportation Research Part A: Policy and Practice, 45*(6), 451–475. https://doi.org/10.1016/j.tra.2011.03.001

QGIS. (2025). Downloads. https://download.qgis.org/downloads/

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed., Vols. 1–2). Stata Press. Retrieved from: https://www.stata-press.com/books/preview/mlmus4-preview.pdf

Rahman, M., & Sciara, G.-C. (2022). Travel attitudes, the built environment and travel behavior relationships: Causal insights from social psychology theories. *Transport Policy, 123*, 44–54. https://doi.org/10.1016/j.tranpol.2022.04.012

Reurings, M., C., B., Vlakveld, W., P., Twisk, D., A., M., Dijkstra, A., & Wijnen, W. (2012). Van fietsongeval naar maatregelen: kennis en hiaten. *Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV*. Retrieved from: https://swov.nl/system/files/publication-downloads/r-2012-08.pdf

Ricchetti, C., Rotaris, L., & Scorrano, M. (2025). What drives university students to cycle? An investigation of their motivations. *International Journal of Sustainable Transportation, 19*(3), 211–226. https://doi.org/10.1080/15568318.2025.2455010

Rietveld, P., & Daniel, V. (2004). Determinants of bicycle use: Do municipal policies matter? *Transportation Research Part A: Policy and Practice, 38*(7), 531–550. https://doi.org/10.1016/j.tra.2004.05.003

Rissel, C., & Watkins, G. (2014). Impact on cycling behavior and weight loss of a national cycling skills program (AustCycle) in Australia 2010–2013. *Journal of Transport & Health, 1*(2), 134–140. https://doi.org/10.1016/j.jth.2014.01.002

Rodríguez, D., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment, 9*(2), 151–173. https://doi.org/10.1016/j.trd.2003.11.001

Ryley, T. (2006). Use of non-motorised modes and life stage in Edinburgh. *Journal of Transport Geography, 14*(5), 367–375. https://doi.org/10.1016/j.jtrangeo.2005.10.001

Scheepers, C. E., Wendel-Vos, G. C. W., den Broeder, J. M., van Kempen, E. E. M. M., van Wesemael, P. J. V., & Schuit, A. J. (2014). Shifting from car to active transport: A systematic review of the effectiveness of interventions. *Transportation Research Part A: Policy and Practice, 70*, 264–280. https://doi.org/10.1016/j.tra.2014.10.015

Schepers, P., Twisk, D., Fishman, E., Fyhri, A., & Jensen, A. (2017). The Dutch road to a high level of cycling safety. *Safety Science, 92*, 264–273. https://doi.org/10.1016/j.ssci.2015.06.005

Semenescu, A., & Coca, D. (2022). Why people fail to bike the talk: Car dependence as a barrier to cycling. *Transportation Research Part F: Traffic Psychology and Behaviour, 88*, 208–222. https://doi.org/10.1016/j.trf.2022.05.025

Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: Simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia Medica*, *31*(1), Article 010502. https://doi.org/10.11613/BM.2021.010502

Seyedrezaei, M., Becerik-Gerber, B., Awada, M., Contreras, S., & Boeing, G. (2023). Equity in the built environment: A systematic review. *Building and Environment, 245*, 110827. https://doi.org/10.1016/j.buildenv.2023.110827

Shoup, D. (1997). The high cost of free parking. *Journal of Planning and Education and Research, 17*(1), 3-20. https://doi.org/10.1177/0739456X9701700102

Slütter, M. (2020). Van wie is de ruimte?. Fietsersbond. Retrieved from: https://www.fietsersbond.nl/nieuws/van-wie-is-de-ruimte/

Smart, M. (2010). US immigrants and bicycling: Two-wheeled in Autopia. *Transport Policy, 17*(3), 153–159. https://doi.org/10.1016/j.tranpol.2010.01.002

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd ed.). Sage. Retrieved from: https://www.researchgate.net/profile/Tom-Snijders-2/publication/44827177_Multilevel_Analysis_An_Introduction_to_Basic_and_Advanced_Multilevel_M_ odeling/links/0c96051ffabd4ca210000000/Multilevel-Analysis-An-Introduction-to-Basic-and-Advanced-Multilevel-Modeling.pdf

Steg, L. (2003). Factors influencing the acceptability and effectiveness of transport pricing. In *Acceptability of transport pricing strategies* (pp. 187–202). Emerald Group Publishing. https://doi.org/10.1108/9781786359506-012

Stinson, M. A., & Bhat, C. R. (2004). Frequency of bicycle commuting: Internet-based survey analysis. *Transportation Research Record, 1878*(1), 122–130. https://doi.org/10.3141/1878-15

Stradling, S.,G. (2011). Travel Mode Choice. *Handbook of Traffic Psychology*, 485-502. https://doi.org/10.1016/B978-0-12-381984-0.10034-7

Taherdoost, H. (2017). *Determining sample size; How to calculate survey sample size*. Retrieved from: https://www.researchgate.net/publication/322887480_Determining_Sample_Size_How_to_Calculate_Survey_Sample_Size

Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011 Jun 27;2:53-55. doi: https://doi.org/10.5116/ijme.4dfb.8dfd

UCLA Institute for Digital Research and Education. (2024). What statistical analysis should I use? Statistical analyses using SAS. Retrieved from: https://stats.oarc.ucla.edu/sas/whatstat/what-statistical-analysis-should-i-usestatistical-analyses-using-sas/

Uijtdewilligen, T., Ulak, M. B., Wijlhuizen, G. J., & Geurs, K. T. (2024). Effects of crowding on route preferences and perceived safety of urban cyclists in the Netherlands. *Transportation Research Part A: Policy and Practice, 183*, 104030. https://doi.org/10.1016/j.tra.2024.104030

Van Acker, V., van Wee, B., & Witlox, F. (2010). When transport geography meets social psychology: Toward a conceptual model of travel behaviour. *Transport Reviews, 30*(2), 219–240. https://doi.org/10.1080/01441640902943453

van Kuijk, R. J., Correia, G. H. de A., van Oort, N., & van Arem, B. (2022). Preferences for first and last mile shared mobility between stops and activity locations: A case study of local public transport users in Utrecht, the Netherlands. *Transportation Research Part A: Policy and Practice, 166*, 285–306. https://doi.org/10.1016/j.tra.2022.10.008

von Stülpnagel, R., & Binnig, N. (2022). How safe do you feel? A large-scale survey concerning the subjective safety associated with different kinds of cycling lanes. *Accident Analysis & Prevention, 167*, 106577. https://doi.org/10.1016/j.aap.2022.106577

Witlox, F., & Tindemans, H. (2004). Evaluating bicycle–car transport mode competitiveness in an urban environment: An activity-based approach. *World Transport Policy and Practice, 8.* Retrieved from: https://www.researchgate.net/publication/235360117_Evaluating_bicycle-car_transport_mode_competitiveness_in_an_urban_environment_An_activity-based_approach

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press. Retrieved from: https://ipcid.org/evaluation/apoio/Wooldridge%20-%20Cross-section%20and%20Panel%20Data.pdf

Wu, H., Lee, J. (B.), & Pettit, C. (2024). What affects commute cycling in Sydney: Access, infrastructure, and demographics. *Transportation Research Interdisciplinary Perspectives, 24*, 101076. https://doi.org/10.1016/j.trip.2024.101076

Yang, Y., Wu, X., Zhou, P., Gou, Z., & Lu, Y. (2019). Towards a cycling-friendly city: An updated review of the associations between built environment and cycling behaviors (2007–2017). *Journal of Transport & Health, 14*, 100613. https://doi.org/10.1016/j.jth.2019.100613

Yannis, G., & Michelaraki, E. (2024). Review of City-Wide 30 km/h Speed Limit Benefits in Europe. *Sustainability*, *16*(11), 4382. https://doi.org/10.3390/su16114382

Zahran et al. (2008). Cycling and Walking: Explaining the spatial distribution of healthy modes of transportation in the United States. https://doi.org/10.1016/j.trd.2008.08.001

Zhang, H., Zhang, L., Liu, Y., & Zhang, L. (2023). Understanding Travel Mode Choice Behavior: Influencing Factors Analysis and Prediction with Machine Learning Method. Sustainability, 15(14), 11414. https://doi.org/10.3390/su151411414

Zhao, P., & Li, S. (2017). Bicycle-metro integration in a growing city: The determinants of cycling as a transfer mode in metro station areas in Beijing. *Transportation Research Part A: Policy and Practice, 99*, 46–60. https://doi.org/10.1016/j.tra.2017.03.003

Appendices

Appendix A – Filter code in SPSS for ODiN data

Select cases if:

Verpl = 1 AND VertGem=AankGem AND VertGem = WoGem AND (WoGem = 344 OR WoGem = 14 OR WoGem = 772 OR WoGem = 363 OR WoGem = 599 OR WoGem = 518 OR WoGem = 34 OR WoGem = 758 OR WoGem = 268 OR WoGem = 855)

Appendix B – Sample testing results

Table B1 and Table B2 - Chi-square goodness of fit test results (SPSS)

CityID			
	Observed N	Expected N	Residual
1	423	587.8	-164.8
2	1762	2422.1	-660.1
3	339	491.8	-152.8
4	486	643.0	-157.0
5	549	628.3	-79.3
6	411	481.4	-70.4
7	2124	1751.4	372.6
8	1899	1484.8	414.2
9	409	600.7	-191.7
10	1660	970.7	689.3
Total	10062		

Test Statistics	
	CityID
Chi-Square	1077.705[a]
df	9
Asymp. Sig.	0.000
a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 481.4.	

Equation B1 - Calculation for Cramér's V

$$Cramér's\ V = \sqrt{\frac{\chi^2}{n * (k-1)}} = \sqrt{\frac{1077.705}{10062 * (10-1)}} = 0.10909 \approx 0.11$$

Equation B2 -Calculation for Cohen's w

$$Cohen's\ \omega = \sqrt{\frac{\chi^2}{N}} = \sqrt{\frac{1077.705}{10062}} = 0.3272 \approx 0.33$$

Table B3 - Proportion of people using a bicycle or not (43.1 vs 56.9)

Bic_max					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5729	56.9	56.9	56.9
	1	4333	43.1	43.1	100.0
	Total	10062	100.0	100.0	

Equation B3 - Required sample size total population

$$n = \frac{Z^2 * p * (1-p)}{E^2} = \frac{1.96^2 * 0.431 * (1-0.431)}{0.05^2} \approx 376.84$$

Equation B4 - Margin of Error in current sample

$$E = \sqrt{\frac{Z^2 * p * (1-p)}{n}} = \sqrt{\frac{1.96^2 * 0.431 * (1-0.431)}{10062}} = 0.009676286 \approx 0.97\%$$

Appendix C – Diagram and python code workflow for automation process in QGIS

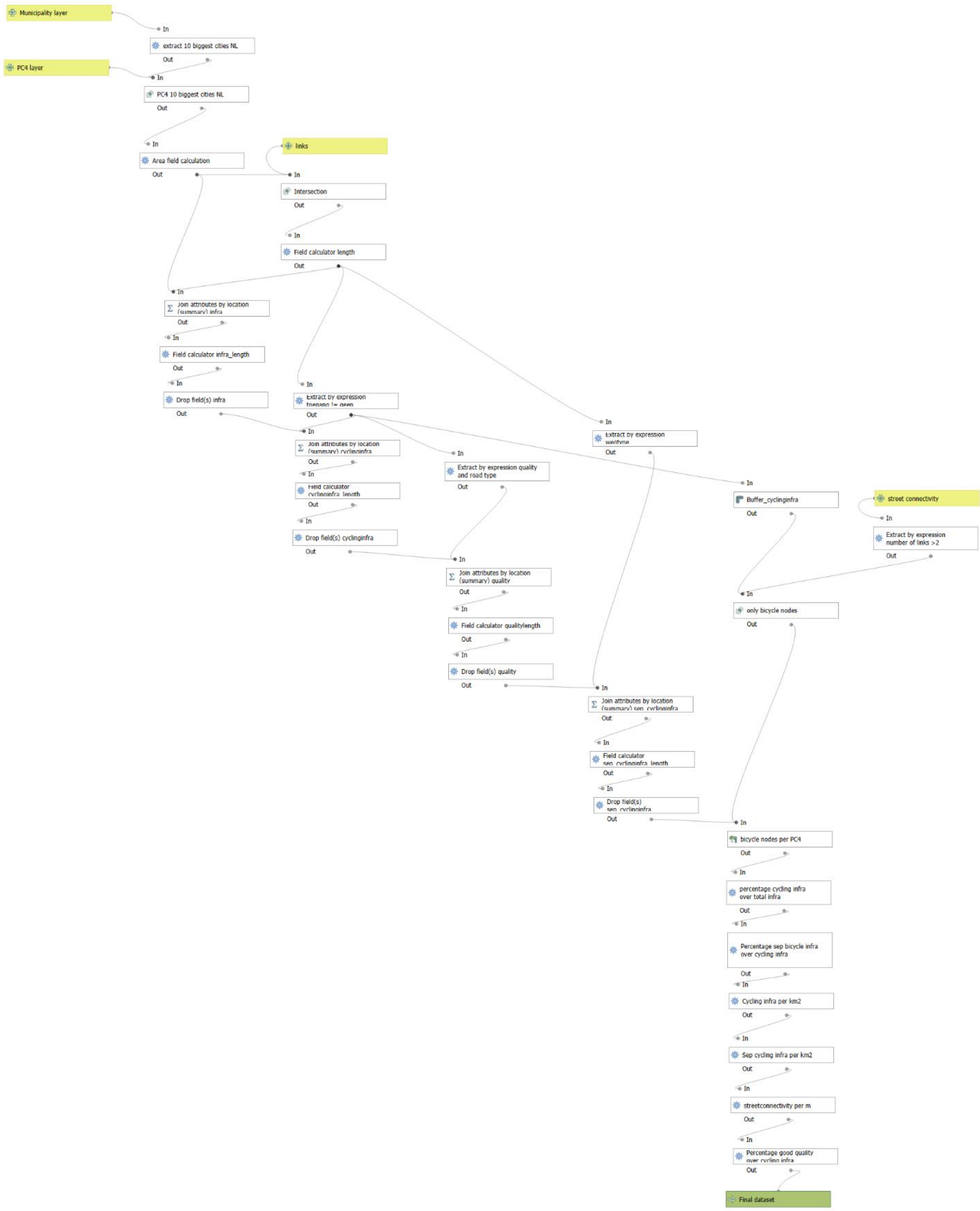

Figure C1 – Model Designer automatic workflow in QGIS for variables on postal code level

Code C1: Python Code for Model Designer illustrated in Figure C1

```
"""
Model exported as python.
Name : ModelDesignerPostalCodeLevel
Group :
With QGIS : 33405
"""

from qgis.core import QgsProcessing
from qgis.core import QgsProcessingAlgorithm
from qgis.core import QgsProcessingMultiStepFeedback
from qgis.core import QgsProcessingParameterVectorLayer
from qgis.core import QgsProcessingParameterFeatureSink
import processing

class Modeldesignerpostalcodelevel(QgsProcessingAlgorithm):

    def initAlgorithm(self, config=None):
        self.addParameter(QgsProcessingParameterVectorLayer('links', 'links', types=[QgsProcessing.TypeVectorLine],
defaultValue=None))
        self.addParameter(QgsProcessingParameterVectorLayer('municipality_layer', 'Municipality layer',
types=[QgsProcessing.TypeVectorPolygon], defaultValue=None))
        self.addParameter(QgsProcessingParameterVectorLayer('pc4_layer', 'PC4 layer',
types=[QgsProcessing.TypeVectorPolygon], defaultValue=None))
        self.addParameter(QgsProcessingParameterVectorLayer('street_connectivity', 'street connectivity',
types=[QgsProcessing.TypeVectorPoint], defaultValue=None))
        self.addParameter(QgsProcessingParameterFeatureSink('FinalDataset', 'Final dataset',
type=QgsProcessing.TypeVectorAnyGeometry, createByDefault=True, supportsAppend=True, defaultValue=None))

    def processAlgorithm(self, parameters, context, model_feedback):
        # Use a multi-step feedback, so that individual child algorithm progress reports are adjusted for the
        # overall progress through the model
        feedback = QgsProcessingMultiStepFeedback(30, model_feedback)
        results = {}
        outputs = {}

        # extract 10 biggest cities NL
        alg_params = {
            'EXPRESSION': '"naam" = \'Amsterdam\' OR "naam" = \'Rotterdam\' OR "naam" = \'Utrecht\' OR "naam" =
\'Eindhoven\'  OR "naam" = \'Tilburg\' OR "naam" = \'Groningen\' OR "naam" = \'Almere\' OR "naam"=\'Breda\' OR
"naam" = \'Nijmegen\' OR "code" = \'0518\'',
            'INPUT': parameters['municipality_layer'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['Extract10BiggestCitiesNl'] = processing.run('native:extractbyexpression', alg_params, context=context,
feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(1)
        if feedback.isCanceled():
            return {}

        # PC4 10 biggest cities NL
        alg_params = {
            'INPUT': parameters['pc4_layer'],
            'OVERLAY': outputs['Extract10BiggestCitiesNl']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
```

```
        outputs['Pc410BiggestCitiesNl'] = processing.run('native:clip', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(2)
        if feedback.isCanceled():
            return {}

        # Extract by expression number of links >2
        alg_params = {
            'EXPRESSION': '"NrOfLinks" > 2',
            'INPUT': parameters['street_connectivity'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['ExtractByExpressionNumberOfLinks2'] = processing.run('native:extractbyexpression', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(3)
        if feedback.isCanceled():
            return {}

        # Area field calculation
        alg_params = {
            'FIELD_LENGTH': 0,
            'FIELD_NAME': 'area',
            'FIELD_PRECISION': 0,
            'FIELD_TYPE': 1,  # Integer (32 bit)
            'FORMULA': '$area',
            'INPUT': outputs['Pc410BiggestCitiesNl']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['AreaFieldCalculation'] = processing.run('native:fieldcalculator', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(4)
        if feedback.isCanceled():
            return {}

        # Intersection
        alg_params = {
            'GRID_SIZE': None,
            'INPUT': parameters['links'],
            'INPUT_FIELDS': [''],
            'OVERLAY': outputs['AreaFieldCalculation']['OUTPUT'],
            'OVERLAY_FIELDS': [''],
            'OVERLAY_FIELDS_PREFIX': '',
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['Intersection'] = processing.run('native:intersection', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(5)
        if feedback.isCanceled():
            return {}

        # Field calculator length
        alg_params = {
            'FIELD_LENGTH': 10,
            'FIELD_NAME': 'length',
```

```
            'FIELD_PRECISION': 2,
            'FIELD_TYPE': 0,  # Decimal (double)
            'FORMULA': '$length\r\n',
            'INPUT': outputs['Intersection']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['FieldCalculatorLength'] = processing.run('native:fieldcalculator', alg_params, context=context,
feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(6)
        if feedback.isCanceled():
            return {}

        # Join attributes by location (summary) infra
        alg_params = {
            'DISCARD_NONMATCHING': False,
            'INPUT': outputs['AreaFieldCalculation']['OUTPUT'],
            'JOIN': outputs['FieldCalculatorLength']['OUTPUT'],
            'JOIN_FIELDS': ['length'],
            'PREDICATE': [0],  # intersect
            'SUMMARIES': [5],  # sum
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['JoinAttributesByLocationSummaryInfra'] = processing.run('native:joinbylocationsummary',
alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(7)
        if feedback.isCanceled():
            return {}

        # Extract by expression toegang != geen
        alg_params = {
            'EXPRESSION': '"toegang"  != \'geen\'',
            'INPUT': outputs['FieldCalculatorLength']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['ExtractByExpressionToegangGeen'] = processing.run('native:extractbyexpression', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(8)
        if feedback.isCanceled():
            return {}

        # Field calculator infra_length
        alg_params = {
            'FIELD_LENGTH': 0,
            'FIELD_NAME': 'infra_length',
            'FIELD_PRECISION': 2,
            'FIELD_TYPE': 0,  # Decimal (double)
            'FORMULA': '"length_sum"',
            'INPUT': outputs['JoinAttributesByLocationSummaryInfra']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['FieldCalculatorInfra_length'] = processing.run('native:fieldcalculator', alg_params, context=context,
feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(9)
        if feedback.isCanceled():
```

```
        return {}

    # Extract by expression wegtype
    alg_params = {
        'EXPRESSION': '"wegtype" = \'bromfietspad (langs weg)\' OR "wegtype" = \'solitair bromfietspad\' OR "wegtype" = \'solitair fietspad\' OR "wegtype" = \'fietspad (langs weg)\' OR "wegtype" = \'solitair onverplicht fietspad\' ',
        'INPUT': outputs['FieldCalculatorLength']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['ExtractByExpressionWegtype'] = processing.run('native:extractbyexpression', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(10)
    if feedback.isCanceled():
        return {}

    # Drop field(s) infra
    alg_params = {
        'COLUMN': ['length_sum'],
        'INPUT': outputs['FieldCalculatorInfra_length']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['DropFieldsInfra'] = processing.run('native:deletecolumn', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(11)
    if feedback.isCanceled():
        return {}

    # Buffer_cyclinginfra
    alg_params = {
        'DISSOLVE': True,
        'DISTANCE': 0.01,
        'END_CAP_STYLE': 0,  # Round
        'INPUT': outputs['ExtractByExpressionToegangGeen']['OUTPUT'],
        'JOIN_STYLE': 0,  # Round
        'MITER_LIMIT': 2,
        'SEGMENTS': 5,
        'SEPARATE_DISJOINT': False,
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['Buffer_cyclinginfra'] = processing.run('native:buffer', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(12)
    if feedback.isCanceled():
        return {}

    # Extract by expression quality and road type
    alg_params = {
        'EXPRESSION': '("wegdeksrt"  = \'asfalt/beton\' OR "wegdeksrt" = \'klinkers\' OR "wegdeksrt" = \'tegels\' ) AND "wegkwal" = \'goed\'',
        'INPUT': outputs['ExtractByExpressionToegangGeen']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['ExtractByExpressionQualityAndRoadType'] = processing.run('native:extractbyexpression', alg_params, context=context, feedback=feedback, is_child_algorithm=True)
```

```
        feedback.setCurrentStep(13)
        if feedback.isCanceled():
            return {}

        # Join attributes by location (summary) cyclinginfra
        alg_params = {
            'DISCARD_NONMATCHING': False,
            'INPUT': outputs['DropFieldsInfra']['OUTPUT'],
            'JOIN': outputs['ExtractByExpressionToegangGeen']['OUTPUT'],
            'JOIN_FIELDS': ['length'],
            'PREDICATE': [0],  # intersect
            'SUMMARIES': [5],  # sum
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['JoinAttributesByLocationSummaryCyclinginfra'] = processing.run('native:joinbylocationsummary', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(14)
        if feedback.isCanceled():
            return {}

        # Field calculator cyclinginfra_length
        alg_params = {
            'FIELD_LENGTH': 0,
            'FIELD_NAME': 'cyclinginfra_length',
            'FIELD_PRECISION': 2,
            'FIELD_TYPE': 0,  # Decimal (double)
            'FORMULA': '"length_sum"',
            'INPUT': outputs['JoinAttributesByLocationSummaryCyclinginfra']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['FieldCalculatorCyclinginfra_length'] = processing.run('native:fieldcalculator', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(15)
        if feedback.isCanceled():
            return {}

        # only bicycle nodes
        alg_params = {
            'INPUT': outputs['ExtractByExpressionNumberOfLinks2']['OUTPUT'],
            'OVERLAY': outputs['Buffer_cyclinginfra']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['OnlyBicycleNodes'] = processing.run('native:clip', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(16)
        if feedback.isCanceled():
            return {}

        # Drop field(s) cyclinginfra
        alg_params = {
            'COLUMN': ['length_sum'],
            'INPUT': outputs['FieldCalculatorCyclinginfra_length']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
```

```
    outputs['DropFieldsCyclinginfra'] = processing.run('native:deletecolumn', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(17)
    if feedback.isCanceled():
        return {}

    # Join attributes by location (summary) quality
    alg_params = {
        'DISCARD_NONMATCHING': False,
        'INPUT': outputs['DropFieldsCyclinginfra']['OUTPUT'],
        'JOIN': outputs['ExtractByExpressionQualityAndRoadType']['OUTPUT'],
        'JOIN_FIELDS': ['length'],
        'PREDICATE': [0],  # intersect
        'SUMMARIES': [5],  # sum
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['JoinAttributesByLocationSummaryQuality'] = processing.run('native:joinbylocationsummary', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(18)
    if feedback.isCanceled():
        return {}

    # Field calculator qualitylength
    alg_params = {
        'FIELD_LENGTH': 0,
        'FIELD_NAME': 'quality_length',
        'FIELD_PRECISION': 2,
        'FIELD_TYPE': 0,  # Decimal (double)
        'FORMULA': '"length_sum"',
        'INPUT': outputs['JoinAttributesByLocationSummaryQuality']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['FieldCalculatorQualitylength'] = processing.run('native:fieldcalculator', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(19)
    if feedback.isCanceled():
        return {}

    # Drop field(s) quality
    alg_params = {
        'COLUMN': ['length_sum'],
        'INPUT': outputs['FieldCalculatorQualitylength']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['DropFieldsQuality'] = processing.run('native:deletecolumn', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(20)
    if feedback.isCanceled():
        return {}

    # Join attributes by location (summary) sep_cyclinginfra
    alg_params = {
        'DISCARD_NONMATCHING': False,
        'INPUT': outputs['DropFieldsQuality']['OUTPUT'],
```

```
        'JOIN': outputs['ExtractByExpressionWegtype']['OUTPUT'],
        'JOIN_FIELDS': ['length'],
        'PREDICATE': [0],  # intersect
        'SUMMARIES': [5],  # sum
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['JoinAttributesByLocationSummarySep_cyclinginfra'] = processing.run('native:joinbylocationsummary', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(21)
    if feedback.isCanceled():
        return {}

    # Field calculator sep_cyclinginfra_length
    alg_params = {
        'FIELD_LENGTH': 0,
        'FIELD_NAME': 'sep_cycling_infra_length',
        'FIELD_PRECISION': 2,
        'FIELD_TYPE': 0,  # Decimal (double)
        'FORMULA': '"length_sum"',
        'INPUT': outputs['JoinAttributesByLocationSummarySep_cyclinginfra']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['FieldCalculatorSep_cyclinginfra_length'] = processing.run('native:fieldcalculator', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(22)
    if feedback.isCanceled():
        return {}

    # Drop field(s) sep_cyclinginfra
    alg_params = {
        'COLUMN': ['length_sum'],
        'INPUT': outputs['FieldCalculatorSep_cyclinginfra_length']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['DropFieldsSep_cyclinginfra'] = processing.run('native:deletecolumn', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(23)
    if feedback.isCanceled():
        return {}

    # bicycle nodes per PC4
    alg_params = {
        'CLASSFIELD': '',
        'FIELD': 'streetconnectivity',
        'POINTS': outputs['OnlyBicycleNodes']['OUTPUT'],
        'POLYGONS': outputs['DropFieldsSep_cyclinginfra']['OUTPUT'],
        'WEIGHT': '',
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['BicycleNodesPerPc4'] = processing.run('native:countpointsinpolygon', alg_params, context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(24)
    if feedback.isCanceled():
        return {}
```

```
    # percentage cycling infra over total infra
    alg_params = {
        'FIELD_LENGTH': 0,
        'FIELD_NAME': 'PerCyclingInfra',
        'FIELD_PRECISION': 3,
        'FIELD_TYPE': 0,  # Decimal (double)
        'FORMULA': '"cyclinginfra_length"/"infra_length"*100',
        'INPUT': outputs['BicycleNodesPerPc4']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['PercentageCyclingInfraOverTotalInfra'] = processing.run('native:fieldcalculator', alg_params, 
context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(25)
    if feedback.isCanceled():
        return {}

    # Percentage sep bicycle infra over cycling infra
    alg_params = {
        'FIELD_LENGTH': 0,
        'FIELD_NAME': 'PerSepCyclingInfra',
        'FIELD_PRECISION': 3,
        'FIELD_TYPE': 0,  # Decimal (double)
        'FORMULA': ' "sep_cycling_infra_length"/ "cyclinginfra_length" * 100',
        'INPUT': outputs['PercentageCyclingInfraOverTotalInfra']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['PercentageSepBicycleInfraOverCyclingInfra'] = processing.run('native:fieldcalculator', alg_params, 
context=context, feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(26)
    if feedback.isCanceled():
        return {}

    # Cycling infra per km2
    alg_params = {
        'FIELD_LENGTH': 0,
        'FIELD_NAME': 'CyclingInfraKm2',
        'FIELD_PRECISION': 3,
        'FIELD_TYPE': 0,  # Decimal (double)
        'FORMULA': '("cyclinginfra_length" * 1000) / "area"',
        'INPUT': outputs['PercentageSepBicycleInfraOverCyclingInfra']['OUTPUT'],
        'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
    }
    outputs['CyclingInfraPerKm2'] = processing.run('native:fieldcalculator', alg_params, context=context, 
feedback=feedback, is_child_algorithm=True)

    feedback.setCurrentStep(27)
    if feedback.isCanceled():
        return {}

    # Sep cycling infra per km2
    alg_params = {
        'FIELD_LENGTH': 0,
        'FIELD_NAME': 'SepCyclingInfraKm2',
        'FIELD_PRECISION': 3,
        'FIELD_TYPE': 0,  # Decimal (double)
```

```
            'FORMULA': '("sep_cycling_infra_length" * 1000)/"area"',
            'INPUT': outputs['CyclingInfraPerKm2']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['SepCyclingInfraPerKm2'] = processing.run('native:fieldcalculator', alg_params, context=context,
feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(28)
        if feedback.isCanceled():
            return {}

        # streetconnectivity per m
        alg_params = {
            'FIELD_LENGTH': 0,
            'FIELD_NAME': 'Streetconnectivityperm',
            'FIELD_PRECISION': 2,
            'FIELD_TYPE': 0,  # Decimal (double)
            'FORMULA': '"cyclinginfra_length"/"streetconnectivity"',
            'INPUT': outputs['SepCyclingInfraPerKm2']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['StreetconnectivityPerM'] = processing.run('native:fieldcalculator', alg_params, context=context,
feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(29)
        if feedback.isCanceled():
            return {}

        # Percentage good quality over cycling infra
        alg_params = {
            'FIELD_LENGTH': 10,
            'FIELD_NAME': 'Qualityperm',
            'FIELD_PRECISION': 2,
            'FIELD_TYPE': 0,  # Decimal (double)
            'FORMULA': '"quality_length"/"cyclinginfra_length"*100',
            'INPUT': outputs['StreetconnectivityPerM']['OUTPUT'],
            'OUTPUT': parameters['FinalDataset']
        }
        outputs['PercentageGoodQualityOverCyclingInfra'] = processing.run('native:fieldcalculator', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)
        results['FinalDataset'] = outputs['PercentageGoodQualityOverCyclingInfra']['OUTPUT']
        return results

    def name(self):
        return 'ModelDesignerPostalCodeLevel'
    def displayName(self):
        return 'ModelDesignerPostalCodeLevel'
    def group(self):
        return ''
    def groupId(self):
        return ''
    def createInstance(self):
        return Modeldesignerpostalcodelevel()
```

Appendix D – Extended description workflow QGIS

The data collection process in QGIS is done by using a model designer. This allows for an automatic workflow. For this particular model designer has six data entries. Firstly, a polygon layer with all municipalities is needed in order to extract the right geographical scope (the ten biggest cities in the Netherlands) from the CBS postal code data, which is the second required data entry and should also be a polygon. The other four data entries are extracted from the route planner and are the above-mentioned link line layers and the street connectivity, traffic light, and cycling route nodes point layers. Note that all layers should be projected to the same Coordinate Reference System (CRS) to ensure correct spatial alignment and enable calculations of distances and areas. The analysis uses the Dutch national projection *Amersfoort / RD New (EPSG:28992).*

The first part of the data processing includes getting only the applicable postal codes from the ten biggest cities in the Netherlands. For this step, downloading the PDOK Services Plugin in QGIS is recommended. This plugin provides access to authoritative Dutch geospatial web services, meaning that geographical information on municipality boundaries can easily be extracted from this by choosing the *gemeente_gegeneraliseerd* layer (type: WFS). The WFS type ensures that the vector features are still applicable instead of just showing a map image of the municipal boundaries. After extracting the ten biggest cities' boundaries, this can be combined with the CBS postal code layer (*Geoprocessing tools -> clip*). What remains are the postal codes, which are located in the ten biggest cities. An extra field representing the area is added, which is needed further in the process.

The next step is to preprocess the data needed from the links layer from the route planner. From this layer, the total infrastructural length (*infra_length*), the bicycle accessible infrastructural total length (*cyclinginfra_length*), the separated bicycle infrastructural length (*sep_cyclinginfra_length*), and the good quality bicycle accessible infrastructural length (*quality_length*) in meters will be calculated. First, the links not located within the ten biggest cities are excluded (geoprocessing tools -> intersection), after which the length for each remaining road segment is calculated. The method for calculating the length of each type of infrastructure within each postal code is similar, except for the filtering code used for each type of infrastructure. For the total infrastructural length, all links will be included. For the infrastructure accessible for the bicycle the road non-accessible for the bicycle should be removed (*toegang != 'geen'*) and the separated cycling lanes are filtered by selecting the road types associated with separated cycling lanes *("wegtype" = 'bromfietspad (langs weg)' OR "wegtype" = 'solitair bromfietspad' OR "wegtype" = 'solitair fietspad' OR "wegtype" = 'fietspad (langs weg)' OR "wegtype" = 'solitair onverplicht fietspad'*). Lastly, the good quality bicycle infrastructure is filtered (toegang != 'geen' AND ("wegdeksrt" = 'asfalt/beton' OR "wegdeksrt" = 'klinkers' OR "wegdeksrt" = 'tegels') AND "wegkwal" = 'goed').

The street connectivity needs some additional preprocessing. First, the layer will be filtered in such a way that only the nodes with more than 2 links connected to this node will be included in the research. This is done as the data includes nodes with fewer than three links, which do not represent a junction but a straight line. After this, only the nodes of the cycling network are included, and the non-cycling network nodes are excluded (with *geoprocessing tools Buffer* and *Intersection*). Finally, the amount of remaining nodes in the postal codes can be calculated with the *analysis tool count points in polygon* (*streetconnectivity*)

The final step in this process is to make calculations in order to make the variables interpretable and comparable between postal codes. For the infrastructural length percentages of the cycling infrastructure and separated cycling infrastructure have been calculated, respectively over the total infrastructural length and the cycling infrastructural length (*PerCyclingInfra* and *PerSepCyclingInfra*). Additionally, the cycling infrastructure and separated cycling infrastructure are both expressed as the amount of meters (separated) cycling infrastructure per square kilometres (*CyclingInfraKm2* and *SepCyclingInfraKm2*). The good quality of bicycle infrastructure is presented as a percentage relative to the total length of bicycle infrastructure. The street connectivity is expressed by the correct nodes per meter (*Streetconnectivityperm*, indicating how many meters of cycling infrastructure the indicator can be found.

Appendix E - statements used during semi-structured interviews

Software Measures

The software statements result directly from Harms et al. (2015).

Statement 1

Education children – Role of local government in learning or improving cycling skills and habits, and awareness of traffic rules and logic

1	No role
2	Very small role
3	Small role
4	Large role
5	Very large role

Statement 2

Education Adults – Role of local government in educating motorists and cyclists

1	No role
2	Very small role
3	Small role
4	Large role
5	Very large role

Statement 3

Marketing campaigns with incentive – Frequency of campaigns aiming to stimulate cycling use with incentive

1	No campaigns
2	One or few general campaigns
3	Many general campaigns
4	One or a few targeted (or individualized)
5	Many targeted (or individualized campaigns)

Statement 4

Marketing campaigns without incentive – Frequency of campaigns aiming to stimulate cycling use with incentive

1	No campaigns
2	One or few general campaigns
3	Many general campaigns
4	One or a few targeted (or individualized)
5	Many targeted (or individualized campaigns)

Orgware

The orgware statement result directly from Harms et al. (2015). However, statement 7 and statement 8 have been adjusted to meet the research interest better. Additionally statement 12 is added to the interview.

Statement 5
Formulation of policy goals - Whether or not cycling policy goals have been formulated which are measurable and have been monitored

1	No goals formulated
2	Goals formulated
3	Measurable goals formulated
4	Measurable goals formulated which are monitored
5	Measurable goals formulated which are monitored and which is acted upon

Statement 6
Implementation of policy measures - Whether or not cycling policy has been implemented

1	Almost nothing has been implemented
2	Few has been implemented
3	Roughly half has been implemented
4	Most have been implemented
5	All policy measures have been implemented

Statement 7
Financial possibilities regarding cycling policy – Availability of financial budget, the height of it and for cycling policies; subsidies (also from higher level governance), structural budgets, general infrastructural budgets, other

1	Little to no available budget, unpredictable funding
2	Small or unstable budget, incidential funding
3	Some structural budget exist, funding takes effort
4	Clear and stable budget available, structural funding
5	Very substantial and structural funding, long-term funding budgets

Statement 8
Policy consistency – Consistency in local politics favouring cycling policy

1	Very much adaptations
2	Many adaptations
3	Not much but also not few adaptations
4	Few adaptations
5	No adaptations

Statement 9

Involvement of actors outside policy area – Actors being involved with cycling policy: employees, schools and educational institutions; sport and recreational organizations; retailers; (public) transport organizations; cycling advocacy organizations; residents' groups, others

1	Not involved
2	Little involved
3	Not much but also not little involved
4	Much involved
5	Very much involved

Statement 10

Relationship between actors inside and outside policy area – communication between actors, collaborations, clarity of roles and tasks

1	None
2	Hardly
3	Average
4	Good/clear
5	Very good/clear

Statement 11

Level of citizen participation – Participation of civilians in policy formulation and implementation

1	No participation
2	Little participation
3	Sometimes participation
4	Often participation
5	Always particapation

Statement 12

Level of collaboration with Fietsersbond – How is the collaboration between the municipality and the Fietsersbond experienced, is the collaboration structural and constructive.

1	Low engagement: no structural meetings, little involvement
2	Litte engagement: structural meetings, further little communication
3	Medium engagement: structural meetings, sometimes collaboration at forefront of projects
4	High engagement: structural meetings, mostly collaboration at forefront of projects, Fietsersbond input for cycling policies
5	Very High engagement: structural meetings, collaboration at forefront of projects, writing cycling policies together

Appendix F – Variable overview

Table F1 - Level indicators

Variable name	Type	Value type	Explanation
TripID	Nominal	integer	Unique ID for every Trip
PersonID	Nominal	integer	Unique ID for every Person
PostalCode	Nominal	integer	Postal code of residential area Person
City	Nominal	integer	City in which Person lives and where trip is performed

Table F2 – Level 1: Trip-level variables

Variable name	**Type**	**Value type**	**Explanation**
BicycleUse	Nominal	0 = no bicycle 1 = bicycle	Bicycle used for trip (yes/no)
Distance	Continuous	integers	Distance (in hectometres)
Duration	Continuous	integers	Duration (in minutes)
Motive	Nominal	0 = Work 1 = Daily service/grocery 2 = Education 3= Recreational/Social 4= Others	Trip motive

Table F3 – Level 2: person-level variables

Variable name	**Type**	**Value type**	**Explanation**
Gender	Nominal	0 = Man 1 = Woman	Gender (woman/man)
Age	Continuous	integers	Age (in years)
Origin	Nominal	0 = Netherlands 1 = Outside Netherlands	Dutch origin (or not)
EducationLevel	Ordinal	0 = No education 1 = Primary school 2 = VMBO, Mavo, Mulo 3 = Havo, Vwo 4 = HBO/University	Highest finished education level
Student	Nominal	0 = No student 1 = Student	daily occupation is student (or not)
DriverLicense	Nominal	0 = No 1 = Yes	Driver license occupation (or not)

Table F4 – Level 3: postal code-level

Variable name	Type	Value type	Explanation
Inhabitants	Continuous	integer	amount of inhabitants
PerBornNL	Continuous	integer	Percentage born in Netherlands
Households	Continuous	integer	amount of households
AverageHHsize	Continuous	decimal	average household size
Houses	Continuous	integer	amount of houses
PerOwneroccupied	Continuous	integer	percentage dwellings owner occupied
AverageHousingValue	Continuous	decimal	average dwelling value
ClosestSupermarket	Continuous	decimal	Closest distance Supermarket in meters
ClostestChildCare	Continuous	decimal	Closest distance Child care in meters
ClosestHighwayEntry	Continuous	decimal	Closest distance Highway entry in meters
ClosestTransitStation	Continuous	decimal	Closest distance Transit station in meters
ClosestPrimarySchool	Continuous	decimal	Closest distance Primary school in meters
ClosestSecondarySchool	Continuous	decimal	Closest distance Secondary School in meters
ClosestPharmacy	Continuous	decimal	Closest distance Pharmacy in meters
ClosestGPCenter	Continuous	decimal	Closest distance General practitioner center in meters
AddressDensity	Continuous	decimal	Address Density
PerCyclingInfra	Continuous	decimal	Percentage bicycle infrastructure over all infrastructure
PerSepCyclingInfra	Continuous	decimal	Percentage separated cycling infrastructure over bicycle infrastructure
CyclingInfraKm2	Continuous	decimal	bicycle infrastructure length in metres per square kilometre
SepCyclingInfraKm2	Continuous	decimal	separated bicycle infrastructure length in metres per square kilometre
PerQuality	Continuous	Decimal	Percentage good quality bicycle infrastructure over bicycle infrastructure
Streetconnectivityperm	Continuous	decimal	Street connectivity explained in nodes per meters of bicycle infrastructure

Table F5 – Level 4: city-level

Variable name	Type	Value type	Explanation
EducationKids	Ordinal	five category Likert scale	Level of education for kids
EducationAdults	Ordinal	five category Likert scale	level of education for adults
MarketingCampaignsIncentive	Ordinal	five category Likert scale	level of marketing campaigns with incentive
Marketingcampainswithoutincentive	Ordinal	five category Likert scale	level of marketing campaigns withouth incentive
FormulationPolicyGoals	Ordinal	five category Likert scale	Level of formulation of policy goals
ImplementationPolicyMeasures	Ordinal	five category Likert scale	level of implementation of policy measures
FinancialSources	Ordinal	five category Likert scale	level of financial sources
PolicyConsistency	Ordinal	five category Likert scale	level of policy consistency
InvolvementOutsidePolicyArea	Ordinal	five category Likert scale	level of involvement outside policy area
RelationshipWithActorsOutside	Ordinal	five category Likert scale	level of relationship with actors outside policy area
CitizenParticipation	Ordinal	five category Likert scale	level of citizen participation
CollaborationFietsersbond	Ordinal	five category Likert scale	level of collaboration with Fietsersbond

Appendix G – Descriptive statistics

Trip level

Table G1 – Distribution of *BicycleUse*

Bicycle Use	n	Percentage (%)
No Bicycle	18,970	62.9
Bicycle	11,194	37.1

Table G2 – Motive descriptives

Motive	n	Percentage (%)
Work	4,057	13.4
Daily service/grocery	9,350	31.0
Education	2,386	7.9
Recreational/Social	10,806	35.8
Others	3,565	11.8

Person level

Table G3 – nominal variables on individual level (n = 10,062)

Variable	Value	n	Percentage (%)
Gender	Man	4,930	49.0
	Woman	5,132	51.0
Origin	Netherlands	5,862	58.3
	Outside Netherlands	4,200	41.7
Student	No student	7,914	78.7
	Student	2,148	21.3
Driver License	No Driver License	3,391	33.7
	Driver license	6,671	66.3
Education level	No Education	1,515	15.1
	Primary school	338	3.4
	VMBO, Mavo, Mulo	921	9.2
	Havo, Vwo	2,098	20.9
	HBO, University	5,190	51.6

Postal code level

Table G4 – descriptives continuos variables postal-code level

Variable	n	Min	Max	Mean	Std. Dev.	Q1	Median	Q3
Inhabitants	431	50.00	26245.00	8832.69	4933.83	5495	8485	11940
PerBornNL	431	10.00	100.00	56.06	19.11	40	60	70
Households	431	35.00	14140.00	4579.94	2715.24	2750	4270	6130
AverageHHsize	431	1.20	3.20	1.99	0.36	1.7	2	2.2
Houses	430	25.00	12790.00	4272.85	2501.82	2615	4030	5740
PerOwneroccupied	429	0.00	100.00	45.97	21.52	30	40	60
AverageHousingValue	430	189.00	1267.00	407.68	158.88	292	371	500
ClosestSupermarket	431	0.20	4.90	0.81	0.64	0.5	0.6	0.8
ClosestChildCare	431	0.20	2.90	0.56	0.37	0.4	0.5	0.6
ClosestHighwayEntry	431	0.40	6.10	2.29	1.06	1.5	2.1	3
ClosestTransitStation	431	0.60	20.30	2.80	4.61	2.7	4.1	5.9
ClosestPrimarySchool	431	0.30	5.10	0.75	0.52	0.5	0.6	0.8
ClosestSecondarySchool	431	0.40	11.10	1.49	1.19	0.8	1.1	1.7
ClosestPharmacy	431	0.30	7.10	1.02	0.82	0.6	0.8	1.1
ClosestGPCenter	431	0.60	17.10	4.29	2.29	2.6	3.9	5.4
AddressDensity	431	41.00	11,760.00	3,465.92	2,411.38	1,683.50	2,845.00	4,609.50
PerCyclingInfra	431	39.88	99.51	79.45	10.18	73.94	80.96	86.23
PerSepCyclingInfra	431	5.25	73.43	29.17	11.42	20.99	28.86	36.01
CyclingInfraKm2	431	1.89	30.84	17.00	6.88	11.76	18.08	22.24
SepCyclingInfraKm2	431	0.42	13.51	4.68	2.32	2.94	4.35	6.21
PerQuality	431	8.59	105.98	55.38	20.82	38.70	52.25	72.51
Streetconnectivityperm	431	66.61	955.27	130.17	73.37	99.11	110.94	133.91

City level

Table G5 - descriptive statistics city-level variables

Domains	Variable	n	Min	Max	Mean	Std. Dev.	Q1	Median	Q3
Software	Education Kids	10	3	4	3.45	0.497	3	3.25	4
	EducationAdults	10	2	5	3.45	0.798	3	3.5	4
	MarketingCampaignsIncentive	10	2	4.5	3.4	0.775	3	3.25	4
	Marketingcampainswithoutincentive	10	2.5	4.5	3.55	0.599	3	3.5	4
Orgware - organizational structure	formulationpolicygoals	10	3	5	3.75	0.635	3	4	4
	implementationPolicymeasures	10	2	4.5	3.6	0.699	3.5	3.75	4
	financialsources	10	3	5	4	0.624	4	4	4.5
	policyconsistency	10	2	5	3.85	1.248	3	4.25	5
Orgware - collaboration	Involvementoutsidepolicyarea	10	2	5	3.5	0.943	3	3.5	4
	relationshipwithactorsoutside	10	3	4	3.7	0.483	3	4	4
	Citizen Participation	10	3	4.5	3.8	0.483	3.5	4	4
	CollaborationFietsersbond	10	2	5	3.85	1.132	3	4.25	5

Appendix H – Multicollinearity results

Table H1 – Generalized Variance Inflation Factors (GVIF) and Adjusted GVIF$^{(1/2*df)}$ for all independent variables

Variable	GVIF	Df	GVIF^(1/(2*Df))
Motive	1.712	4	1.070
Distance_gmc	1.549	1	1.245
Duration_gmc	1.585	1	1.259
Gender	1.015	1	1.008
Age_gmc	2.019	1	1.421
Origin	1.199	1	1.095
Student	2.482	1	1.575
EducationLevel	2.090	4	1.097
Driverlicense	1.654	1	1.286
Inhabitants_gmc	32.680	1	5.717
PerOwneroccupied_gmc	4.802	1	2.191
PerBornNL_gmc	3.345	1	1.829
Households_gmc	71.901	1	8.479
Houses_gmc	51.135	1	7.151
AverageHHsize_gmc	9.387	1	3.064
AverageHousingValue_gmc	2.039	1	1.428
ClosestSupermarket_gmc	3.143	1	1.773
ClostestChildCare_gmc	2.110	1	1.453
ClosestHighwayEntry_gmc	1.381	1	1.175
ClosestTransitStation_gmc	2.192	1	1.481
ClosestPrimarySchool_gmc	2.514	1	1.586
ClosestSecondarySchool_gmc	1.879	1	1.371
ClosestPharmacy_gmc	3.257	1	1.805
ClosestGPCenter_gmc	1.681	1	1.296
AddressDensity_gmc	4.292	1	2.072
PerCyclingInfra_gmc	2.225	1	1.492
PerSepCyclingInfra_gmc	11.030	1	3.321
PerQuality_gmc	1.504	1	1.226
Streetconnectivityperm_gmc	2.851	1	1.688
CyclingInfraKm2_gmc	9.977	1	3.159
SepCyclingInfraKm2_gmc	13.226	1	3.637
Softwarescore	4.049	1	2.012
OrgwareOrganisation	4.308	1	2.075
OrgwareCollaboration	2.272	1	1.507

Table H2 – extra multicollinearity check on significant variables.

Correlations				
		PerSepCycling Infra	CyclingInfra Km2	SepCycling InfraKm2
PerSepCycling Infra	Pearson Correlation	1	-.360**	.458**
	Sig. (2-tailed)		0.000	0.000
	N	431	431	431
CyclingInfra Km2	Pearson Correlation	-.360**	1	.595**
	Sig. (2-tailed)	0.000		0.000
	N	431	431	431
SepCycling InfraKm2	Pearson Correlation	.458**	.595**	1
	Sig. (2-tailed)	0.000	0.000	
	N	431	431	431
**. Correlation is significant at the 0.01 level (2-tailed).				

Appendix I – R code used during research

Code I1 -R code for multicollinearity, intercept-only models and multilevel logistic regression model

```
install.packages("lme4")
install.packages("readr")
install.packages("dplyr")

library(lme4)
library(readr)
library(dplyr)

# Load and show csv file
df <- read.csv("FILEPATH", header = TRUE)
head(df)
str(df)

#checking missing values
sum(is.na(df$Houses))
sum(is.na(df$PerOwneroccupied))
sum(is.na(df$AverageHousingValue))

#remove -99997 (make missing values) -> check if true after
df$Houses[df$Houses == -99997] <- NA
df$PerOwneroccupied[df$PerOwneroccupied == -99997] <- NA
df$AverageHousingValue[df$AverageHousingValue == -99997] <- NA

#checking missing values
sum(is.na(df$Houses))
sum(is.na(df$PerOwneroccupied))
sum(is.na(df$AverageHousingValue))

#ensuring level indicators are factors
df$City <- as.factor(df$City)
df$PostalCode <- as.factor(df$PostalCode)
df$PersonID <- as.factor(df$PersonID)

#checking how many entries per level (extra check if data is complete)
n_trips <- nrow(df)
n_persons <- length(unique(df$PersonID))
n_postcodes <- length(unique(df$PostalCode))
n_cities <- length(unique(df$City))
cat("Trips:", n_trips, "\n",
   "Persons:", n_persons, "\n",
   "Postal codes:", n_postcodes, "\n",
   "Cities:", n_cities, "\n")

#making grand mean centered variables to make conclusions on general level
df$Distance_gmc <- df$Distance - mean(df$Distance, na.rm=TRUE)
df$Duration_gmc <- df$Duration - mean(df$Duration, na.rm=TRUE)
df$Age_gmc <- df$Age - mean(df$Age, na.rm=TRUE)
```

```
df$Inhabitants_gmc <- df$Inhabitants - mean(df$Inhabitants, na.rm=TRUE)
df$PerBornNL_gmc <- df$PerBornNL - mean(df$PerBornNL, na.rm=TRUE)
df$Households_gmc <- df$Households - mean(df$Households, na.rm=TRUE)
df$Houses_gmc <- df$Houses - mean(df$Houses, na.rm=TRUE)
df$AverageHHsize_gmc <- df$AverageHHsize - mean(df$AverageHHsize, na.rm=TRUE)
df$PerOwneroccupied_gmc <- df$PerOwneroccupied - mean(df$PerOwneroccupied, na.rm=TRUE)
df$AverageHousingValue_gmc <- df$AverageHousingValue - mean(df$AverageHousingValue,
na.rm=TRUE)

df$ClosestSupermarket_gmc <- df$ClosestSupermarket - mean(df$ClosestSupermarket, na.rm=TRUE)
df$ClostestChildCare_gmc <- df$ClostestChildCare - mean(df$ClostestChildCare, na.rm=TRUE)
df$ClosestHighwayEntry_gmc <- df$ClosestHighwayEntry - mean(df$ClosestHighwayEntry, na.rm=TRUE)
df$ClosestTransitStation_gmc <- df$ClosestTransitStation - mean(df$ClosestTransitStation, na.rm=TRUE)
df$ClosestPrimarySchool_gmc <- df$ClosestPrimarySchool - mean(df$ClosestPrimarySchool,
na.rm=TRUE)
df$ClosestSecondarySchool_gmc <- df$ClosestSecondarySchool - mean(df$ClosestSecondarySchool,
na.rm=TRUE)
df$ClosestPharmacy_gmc <- df$ClosestPharmacy - mean(df$ClosestPharmacy, na.rm=TRUE)
df$ClosestGPCenter_gmc <- df$ClosestGPCenter - mean(df$ClosestGPCenter, na.rm=TRUE)

df$AddressDensity_gmc <- df$AddressDensity - mean(df$AddressDensity, na.rm=TRUE)

df$PerCyclingInfra_gmc <- df$PerCyclingInfra - mean(df$PerCyclingInfra, na.rm=TRUE)
df$PerSepCyclingInfra_gmc <- df$PerSepCyclingInfra - mean(df$PerSepCyclingInfra, na.rm=TRUE)
df$CyclingInfraKm2_gmc <- df$CyclingInfraKm2 - mean(df$CyclingInfraKm2, na.rm=TRUE)
df$SepCyclingInfraKm2_gmc <- df$SepCyclingInfraKm2 - mean(df$SepCyclingInfraKm2, na.rm=TRUE)
df$StreetconnectivitypermV3_gmc <- df$StreetconnectivitypermV3 -
mean(df$StreetconnectivitypermV3, na.rm=TRUE)
df$PerQuality_gmc <- df$PerQuality - mean(df$PerQuality, na.rm=TRUE)
df$ClosestSuperDailyService_gmc <- df$ClosestSuperDailyService - mean(df$ClosestSuperDailyService,
na.rm=TRUE)

#Variables which should be factors, also ordered if ordinal
df$Motive <- as.factor(df$Motive)
df$Gender <- as.factor(df$Gender)
df$Origin <- as.factor(df$Origin)
df$Student <- as.factor(df$Student)
df$Driverlicense <- as.factor(df$Driverlicense)
df$EducationLevel <- factor(df$EducationLevel, ordered = TRUE,levels= c(0,1,2,3,4))

# Select all variables in model
vars_in_model <- c("BicycleUse", "Motive", "Distance_gmc", "Duration_gmc", "Gender", "Age_gmc",
        "Origin", "Student", "EducationLevel", "Driverlicense", "Inhabitants_gmc",
        "PerOwneroccupied_gmc", "PerBornNL_gmc", "Households_gmc", "Houses_gmc",
"AverageHHsize_gmc", "AverageHousingValue_gmc",
        "ClosestSupermarket_gmc", "ClostestChildCare_gmc",
        "ClosestHighwayEntry_gmc", "ClosestTransitStation_gmc",
        "ClosestPrimarySchool_gmc", "ClosestSecondarySchool_gmc", "ClosestPharmacy_gmc",
        "ClosestGPCenter_gmc","AddressDensity_gmc", "PerCyclingInfra_gmc",
        "PerSepCyclingInfra_gmc", "PerQuality_gmc", "CyclingInfraKm2_gmc",
"SepCyclingInfraKm2_gmc", "StreetconnectivitypermV3_gmc",
```

```
             "RouteNodes_gmc",
"TrafficLights_gmc","Softwarescore","OrgwareOrganisation","OrgwareCollaboration",
             "City", "PostalCode", "PersonID")

# count complete cases (again extra check)
sum(complete.cases(df[, vars_in_model]))

# instal for VIF
install.packages("car")
library(car)

# VIF model
model_vif <- lm(BicycleUse ~ Motive + Distance_gmc +  Duration_gmc +  Gender +  Age_gmc +
             Origin + Student + EducationLevel + Driverlicense + Inhabitants_gmc +
             PerOwneroccupied_gmc + PerBornNL_gmc + Households_gmc + Houses_gmc +
AverageHHsize_gmc + AverageHousingValue_gmc +
             ClosestSupermarket_gmc + ClostestChildCare_gmc +
             ClosestHighwayEntry_gmc + ClosestTransitStation_gmc +
             ClosestPrimarySchool_gmc + ClosestSecondarySchool_gmc +  ClosestPharmacy_gmc +
             ClosestGPCenter_gmc + AddressDensity_gmc + PerCyclingInfra_gmc +
             PerSepCyclingInfra_gmc + PerQuality_gmc + StreetconnectivitypermV3_gmc +
CyclingInfraKm2_gmc + SepCyclingInfraKm2_gmc +
              Softwarescore + OrgwareOrganisation +
OrgwareCollaboration
, data = df)

# Calculate VIF
vif(model_vif)

model_intercept4 <- glmer(
  BicycleUse ~ 1 +
    (1 | City) +
    (1 | City:PostalCode) +
    (1 | City:PostalCode:PersonID),
  data = df,
  family = binomial(link = "logit")
)
summary(model_intercept4)

# code for changing logodd to probability
plogis(fixef(model_intercept4)["(Intercept)"])

# deviation on city level for each city
ranef(model_intercept4)$City

model_intercept3 <- glmer(
  BicycleUse ~ 1 +
    (1 | City) +
    (1 | City:PostalCode),
  data = df,
  family = binomial(link = "logit")
)
```

```
summary(model_intercept3)

# code for changing logodd to probability
plogis(fixef(model_intercept3)["(Intercept)"])

# deviation on city level for each city
ranef(model_intercept3)$City

model_3level3 <- glmer(
  BicycleUse ~
    Motive + Distance_gmc + Duration_gmc + Gender + Age_gmc +
    Origin + Student + EducationLevel + Driverlicense + Inhabitants_gmc +
    PerOwneroccupied_gmc + PerBornNL_gmc + AverageHHsize_gmc + AverageHousingValue_gmc +
    ClosestSupermarket_gmc + ClostestChildCare_gmc +
    ClosestHighwayEntry_gmc + ClosestTransitStation_gmc +
    ClosestPrimarySchool_gmc + ClosestSecondarySchool_gmc + ClosestPharmacy_gmc +
    ClosestGPCenter_gmc + AddressDensity_gmc + PerCyclingInfra_gmc +
    PerSepCyclingInfra_gmc + PerQuality_gmc + CyclingInfraKm2_gmc + SepCyclingInfraKm2_gmc +
    StreetconnectivitypermV3_gmc + Softwarescore + OrgwareOrganisation +
    OrgwareCollaboration +
    (1 | City) + (1 | City:PostalCode),
  data = df,
  family = binomial(link = "logit")
)
summary(model_3level3)

#likelihood ratio-test
anova(model_intercept3, model_3level3, test = "Chisq")
```