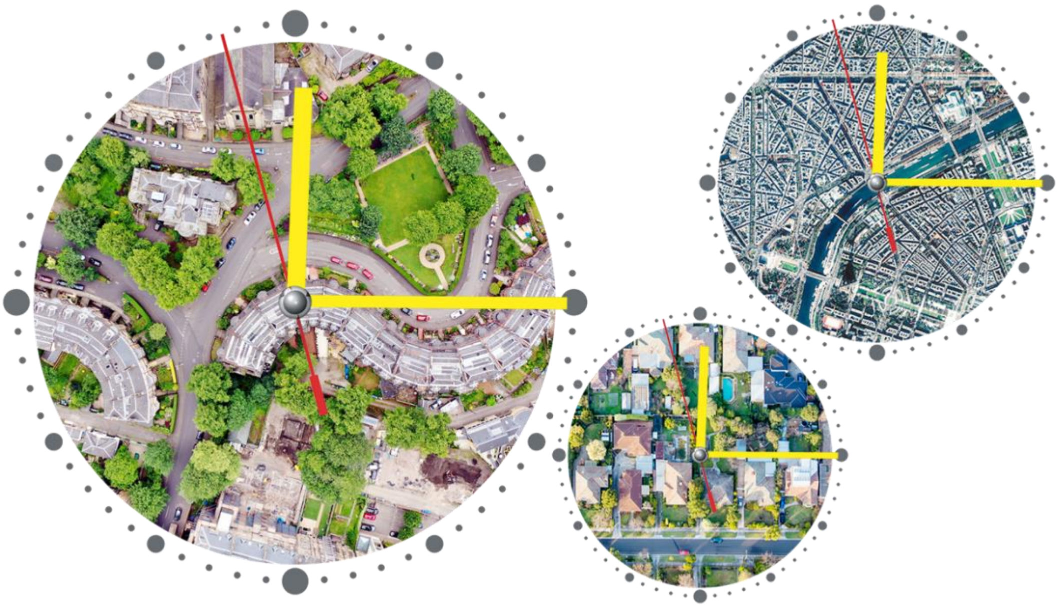
Evaluating Actual Travel Behaviour in Relation to 15-minute City Ambitions Through Mobility Data Analysis



© Shutterstock; Getty | Aerial views of Glasgow, Paris and Melbourne

Author:	Nicky de Theije
Student number:	2783854
Email:	n.de.theije@student.vu.nl
Program:	Earth, Economics & Sustainability
Faculty:	Faculty of Science
Supervisor:	dr. E. Koomen
Second Supervisor:	dr. J. Liu
Second Assessor:	dr. M. Lankhuizen
Date:	June 27, 2025

VRIJE UNIVERSITEIT AMSTERDAM

Abstract

The 15-minute city concept has gained traction as a planning concept that promotes compact, liveable neighbourhoods where residents can access essential services within a short walk or bike ride. While widely adopted in urban policy, questions remain about whether actual travel behaviour aligns with the intended spatial distribution of amenities in the 15-minute city concept. This study examines that alignment by analysing the relationship between cycling behaviour and accessibility of amenities in the municipality of Alkmaar, the Netherlands.

Using secondary data, this study combines socio-demographic statistics from Statistics Netherlands, accessibility scores previously calculated, and anonymised travel behaviour data collected by Mobidot. A series of Ordinary Least Squares (OLS) and Tobit regression models were used to analyse how the share of cycling trips per neighbourhood relates to amenity accessibility and socio-demographic variables.

The results showed that higher accessibility scores were generally associated with higher shares of short cycling trips. However, these associations only reached statistically significance in some models, and overall explanatory power remained modest. Socio-demographic variables contributed limited additional explanation, although the share of residents aged 25-45 and 65+ years emerged as consistent positive predictors in the OLS models. These results suggest that accessibility alone does not fully explain short-distance cycling behaviour, and that important behavioural or environmental factors are likely missing from the models.

To better understand and support sustainable travel patterns in line with the 15-minute city concept, future research should adopt a multi-method approach that incorporates behavioural insights, built environments, personal characteristics, and higher resolution mobility data.

Table of Contents

1. Introduction

In a globally connected world and increasingly smart cities, the demand for living in a physical neighbourhood where one can walk and cycle among familiar people and a variety of services is always alive (Abdelfattah et al., 2022). The 15-minute city concept responds to this demand by offering a planning approach that seeks to create compact, self-sufficient neighbourhoods in which residents can meet most of their daily needs within a short walking or cycling distance from their homes. By prioritising local access over long commutes, the concept marks a clear shift away from car-centric urban development and towards more liveable, inclusive, and environmentally friendly cities.

The 15-minute city concept emerged as a response to the overdependence on private cars (Manifesty & Park, 2022), and envisions urban environments where residents can enjoy a higher quality of life by effectively fulfilling six essential urban social functions (i.e., living, working, commerce, healthcare, education, and entertainment) within a 15-minute walk or bike ride (Moreno et al., 2016). It is a flexible concept that municipalities can tailor to their culture, circumstances and specific needs (C40 Cities Climate Leadership Group & C40 Knowledge Hub, 2020), taking various forms worldwide (from 'complete neighbourhoods' and '20-minute neighbourhoods' to a 'city of close proximities') (C40 Cities Climate Leadership Group & C40 Knowledge Hub, 2023). As a post-COVID urban recovery strategy, the 15-minute city not only seeks to decentralize urban functions and create self-sufficient neighbourhoods (Khavarian-Garmsir et al., 2023), but also integrates multiple policy goals related to climate, equity, health, and urban development, fostering engagement among residents, businesses and non-profits in shaping the future of their cities (C40 Cities Climate Leadership Group & C40 Knowledge Hub, 2021).

To move from vision to reality, cities around the world are implementing a range of spatial and policy interventions aimed at facilitating the 15-minute city. A commonly adopted strategy is transit-oriented development, which promotes denser, mixed-used development around public transport services, enabling a large-scale shift away from reliance on private vehicles (C40 Cities Climate Leadership Group & C40 Knowledge Hub, 2020). Moreover, improving infrastructure for active travel (i.e., walking and cycling) is fundamental to enable local access to essential services.

Equitable access to key amenities also plays a central role in the 15-minute concept, as it ensures that all residents, regardless of income or background, can benefit from nearby amenities and reducing social disparities. For instance, as part of its Green New Deal, Los Angeles set a 2035 goal for all low-income residents to live within half a mile of fresh food stores (C40 Cities Climate Leadership Group & C40 Knowledge Hub, 2020). In addition, cities are exploring the flexible use of existing buildings and public spaces to increase functionality and community value. Paris, for example, is transforming school playgrounds into green spaces and granting residents access outside school hours for recreation, community gardening and to escape the summer heat (C40 Cities Climate Leadership Group & C40 Knowledge Hub, 2020).

While such initiatives demonstrate the growing commitment of cities to realise the 15-minute city in practice, important questions remain about its actual effectiveness. Existing studies have primarily examined the concept from the perspective of accessibility and sustainability, often focusing on the development of indicators to assess the spatial and environmental performance of urban areas (Papadopoulos et al., 2023). These assessments typically measure whether neighbourhoods meet certain criteria, but offer limited insight into how residents interact with their urban environment and whether their travel patterns align with the goals of the 15-minute city concept (Papadopoulos et al., 2023).

Despite the concept's widespread appeal and theoretical promise, there are still gaps in understanding its practical application and impact, especially concerning the alignment of actual travel behaviour with the intended spatial distribution of amenities (Khavarian-Garmsir et al., 2023). A key challenge lies in the assumption that residents will automatically use nearby amenities when available. In reality, travel behaviour is shaped by various factors, such as personal preferences, habitual routines, socio-economic factors, or the perceived quality of local services. These behavioural influences may result in residents bypassing nearby amenities in favour of more distant options, thereby limiting the practical realisation of self-sufficient neighbourhoods envisioned by the 15-minute city concept.

Although the ambition is to ensure all essential services are located within a walking or cycling distance, it remains unclear to what extent the proximity and density of amenities actually influence residents' mobility patterns. This study will therefor focus on the relationship between travel behaviour and the local spatial context, while also examining the extent to which socio-demographic characteristics help to explain observed patterns.

Addressing this knowledge gap requires investigating the extent to which residents' actual travel behaviour aligns with the spatial distribution and accessibility of amenities. While most existing studies focus on large urban areas, little is known about how the 15-minute city concept applies in medium-sized municipality. Studying medium-sized cities is important because they often face different spatial, social, and infrastructure conditions compared to bigger urban areas. Their scale may offer both opportunities and challenges for implementing the 15-minute city, making it valuable to explore how the concept translates into such contexts. For this reason, the municipality of Alkmaar was selected as the case study. As a mid-sized Dutch Municipality, Alkmaar combines a dense, historic city centre with suburban neighbourhoods and surrounding rural areas. This spatial diversity makes it a suitable case for evaluating the relationship between accessibility and travel behaviour.

Therefore, the main research question guiding this study is: *"To what extent does actual travel behaviour align with the spatial distribution and accessibility of amenities, and how does this alignment vary across different socio-demographic characteristics?"*

The general motivation of this research arises from the need to better understand how well the theoretical concept of the 15-minute city translates into practice. From an Earth, Economic, and Sustainability (EES) perspective, this is crucial for several reasons. Environmentally, the 15-minute city offers a possible solution in urban planning to reduce energy use in cities and reduce emissions to mitigate climate change (Knap et al., 2023). Economically, it can support local businesses and reduce transportation costs for residents (Papadopoulos et al., 2023). For instance, Badawi et al. (2018) found that transportation costs of households in walkable districts are half of those living in car-dependent areas. Socially, it highlights important questions about equity and accessibility, as not all residents may benefit equally from the 15-minute city concept (Khavarian-Garmsir et al., 2023). The distribution of amenities may also be unequal across neighbourhoods, which could reinforce existing disparities in access to essential services. Ultimately, the insights gained from this research can support policymakers in creating more sustainable, accessible, and equitable urban environments.

2. Methods

This study uses a quantitative research approach to investigate how the spatial distribution of amenities relates to residents' active travel behaviour in the municipality of Alkmaar. The analysis is based entirely on secondary data collected through desk research. The main sources include socio-demographic statistics form Statistics Netherlands (CBS), accessibility scores from previous spatial analysis, and travel behaviour data from Mobidot.

2.1 Data Collection and Characteristics

To begin with, socio-demographic characteristics were obtained from the CBS dataset "Kerncijfers wijken en buurten" for the years 2022 and 2023. As the travel behaviour data cover the year 2023, this year was prioritised where available to ensure consistency across datasets. The CBS datasets provide extensive insights into neighbourhood-level population statistics, including the percentage of households with and without children, as well as the age distribution of residents across five categories: 0-15, 15-25, 25-45, 45-60, and 60 years and older. However, some relevant variables such as average income per income recipient and the distribution of educational attainment levels (categorised as low, medium, or high) were not available for 2023. Therefore, these variables were sourced from the 2022 dataset. These socio-demographic variables are used as control variables in the subsequent analysis.

Central to this research are accessibility scores, which quantify how many amenities are reachable within a 15-minute walk or bike ride from each neighbourhood. These scores were calculated in a prior analysis (see De Theije, 2025) and are reused in this study to examine their relationship with active travel behaviour. The accessibility scores were computed based on six amenity categories: education, entertainment and leisure, grocery and shopping, healthcare, service, and sports. To ensure these scores accurately reflect accessibility from where people actually live, population-weighted centroids were used as starting points (De Theije, 2025). This approach avoids the common pitfall of relying on geometric centroids, which can fall into uninhabited areas such as parks or bodies of water, especially in neighbourhoods with uneven population distribution. For each residential location, the number of amenities within a 15-minute walking or cycling range was determined (De Theije, 2025). The resulting value reflects how well a neighbourhood is connected to the facilities in the before mentioned categories; a higher score indicates better spatial access. Figure 1 illustrates the spatial variation in cycling accessibility scores across neighbourhoods.

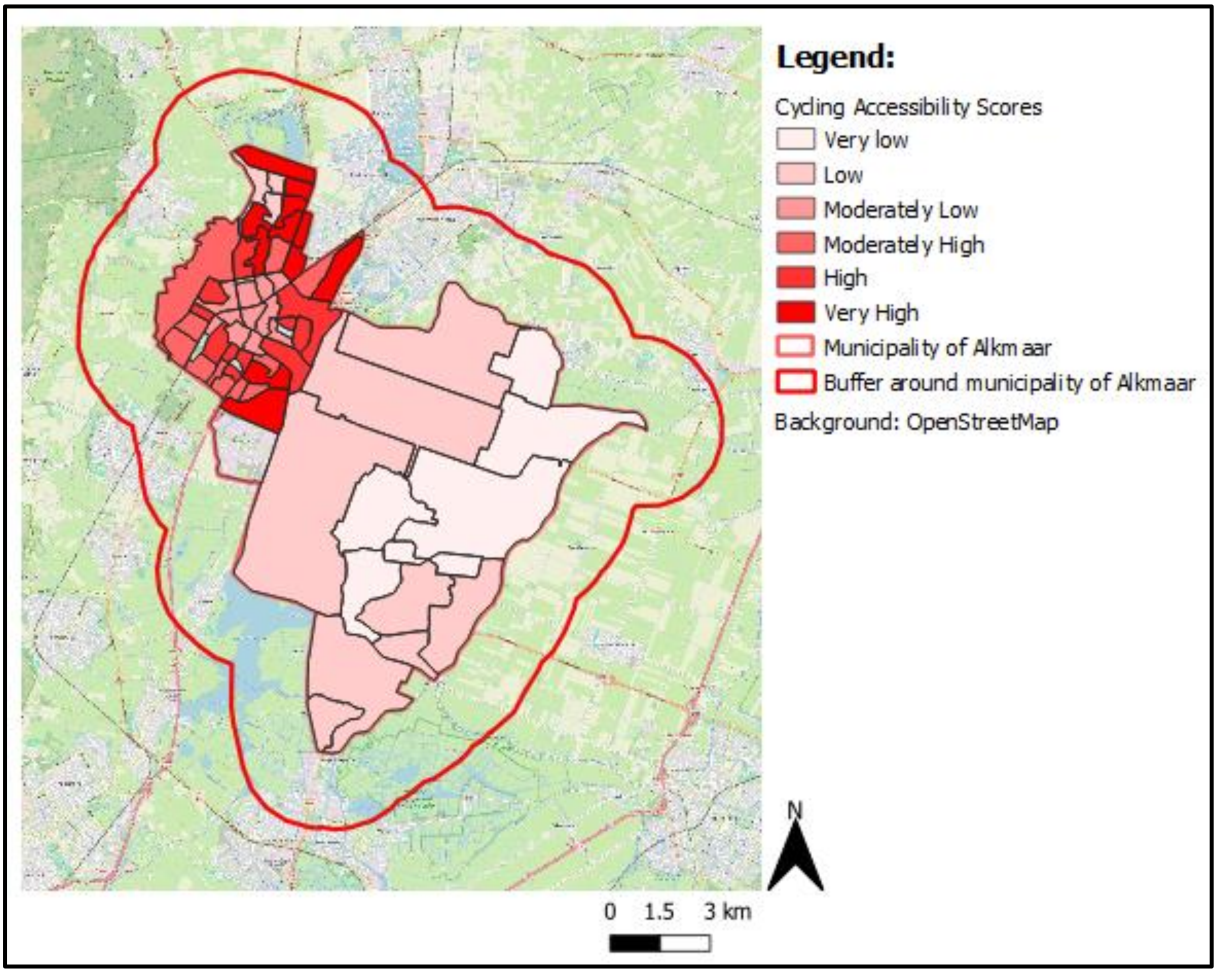


Figure 1 Cycling Accessibility Scores per neighbourhood (Own data VU, 2025)

In addition to the socio-demographic and spatial accessibility data, travel behaviour data were obtained from Mobidot. Due to privacy concerns, this study used aggregated data at the neighbourhood level, recording the number of journeys per neighbourhood by travel purpose over the period from January 1, 2023, to December 31, 2023. The data were collected via SESAMO, a smartphone app developed by Mobidot that passively tracks users' movements with their consent (*Mobidot, n.d.)*. Smartphone-based trip detection helps to reduce underreporting of trips, which is a common phenomenon in travel surveys (Thomas et al., 2018). Trips are classified both by mode of transport, including walking and cycling, and by travel motive. The motive classification includes categories such as visiting and/or staying, (grocery) shopping, other leisure activities, walking/strolling, professional, pickup and bring people, personal care services, commuting, sport/hobby, business visit and school/education. A trip is defined as a movement from a clearly identified origin point to a destination point. When multiple destinations occur within a single journey (for instance, stopping at several stores before returning home), the system splits the journey into separate trips for each leg.

By combining these diverse datasets, the study aims to explore how access to amenities from residential areas influences the proportion of walking and/or cycling trips in neighbourhoods of the municipality of Alkmaar.

2.2 Data Preparation

Before conducting the analysis, all datasets were cleaned and prepared for integration using Microsoft Excel. The socio-demographic data from CBS contained some missing or confidential values. These were systematically converted into NaN (Not a Number) values to ensure that they could be automatically excluded for maintaining the integrity and reliability of the analysis results.

Several variables required additional processing. For example, educational attainment was not provided in the CBS data as percentages, but as absolute numbers of individuals aged 15 to 75 years in each education category (low, medium, high). To allow for meaningful comparisons across neighbourhoods, these values were converted into percentages. This was done by dividing the number of individuals in each education category per neighbourhood by the total population aged 15+ years in that neighbourhood. The age group under 15 was excluded from the denominator, as individuals in that category are generally not expected to have completed any formal education and thus were not included in the educational attainment data. This transformation resulted in a more interpretable variable: the percentage of residents with low, medium, or high educational attainment.

The accessibility scores for walking and cycling, previously calculated for each neighbourhood, initially showed a skewed distribution, which could distort regression estimates and reduce interpretability. To correct for this and to normalize the distribution, a natural logarithmic transformation was applied to both scores. Since all scores were strictly positive and non-zero, no constant had to be added. The resulted variables represent the logarithmic transformation of the original accessibility scores and were used as the central independent variable.

In initial regression models, these continuous log-transformed accessibility scores were directly included. However, these models showed limited explanatory power and no significant effects. Therefore, the accessibility variables were recoded into categorical dummy variables to better capture potential non-linear relationships with travel behaviour.

To further explore these non-linear relationships, these log-transformed accessibility scores were also categorized into dummy variables (Figure 2). Based on visual inspection of scatterplots (see Section 3.1), thresholds were selected to create three categories: values below 5.5 (cycle dummy 1), values between 5.5 and 6.5 (cycle dummy 2), and values equal to or greater than 6.5 (cycle dummy 3).

Before running the regression models, all variables were converted to the appropriate data types (float or integers) to ensure compatibility with statistical software.

To ensure consistency with the accessibility scores, only trips associated with relevant purposes were retained in the travel behaviour data. Specifically, trips related to shopping, walking, leisure, school, care service and sports were included, while trips for commuting, professional activities, or business visits were excluded. This filtering ensured that the travel behaviour data were aligned with the types of amenities captured in the accessibility scores. Finally, to create the dependent variable for the regression models, the number of walking or cycling trips in each neighbourhood was divided by the total number of trips recorded there. This resulted in two proportional measures: the share of walking trips and the share of cycling trips (Figure 3).

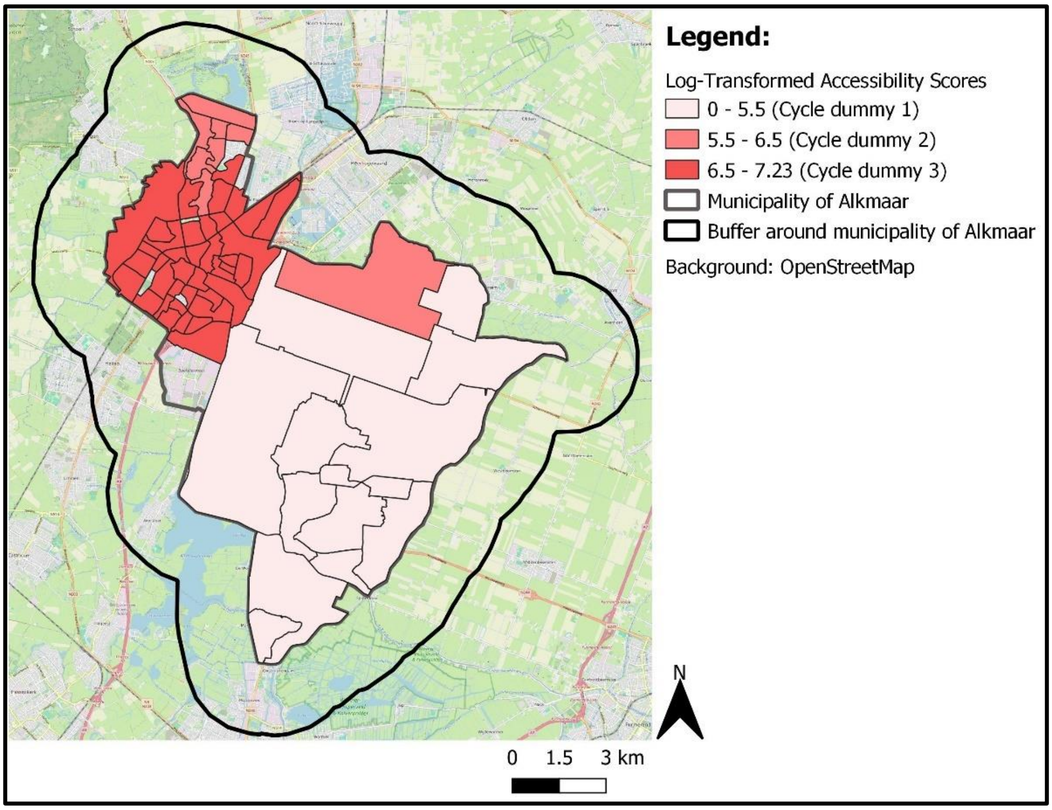


Figure 2 Log-Transformed Accessibility scores divided into dummy variables (Own data VU, 2025)

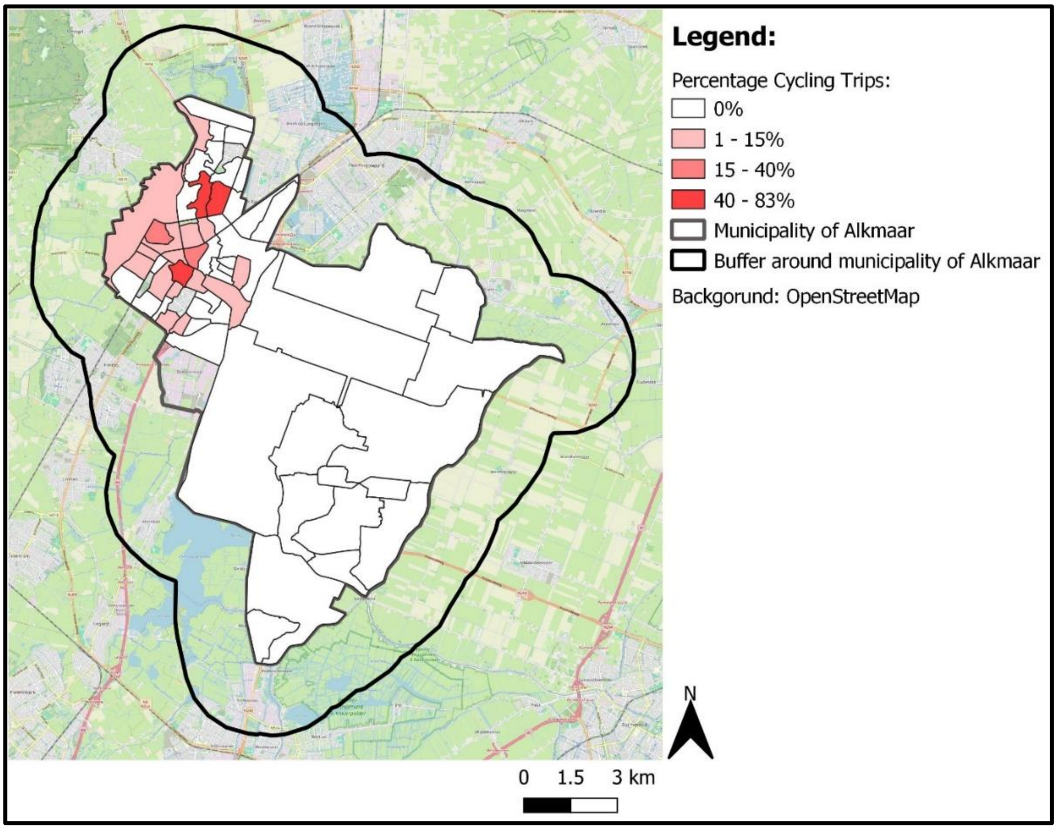


Figure 3 Percentage of cycling trips in comparison to the total amount of trips per neighbourhood (Mobidot, 2024; Own data VU, 2025)

2.3 Data Analysis

To investigate the relationship between spatial accessibility and active travel behaviour, a combination of exploratory and regression-based analyses was conducted. The dependent variable in this study is the share of cycling trips per neighbourhood, which was modelled against various socio-demographic and accessibility-related explanatory variables.

As a first step, exploratory scatterplots were created to visually examine the relationship between the dependent and key independent variables. These plots focused on the share of walking and cycling trips and the log-transformed accessibility scores for walking and cycling. This initial visual inspection served multiple purposes: it helped identify potential non-linear patterns and explore potential associations. Based on these patterns, dummy variables were created for cycling accessibility scores to better capture non-linear effects in the regression models and allow for a more nuanced understanding of how different levels of accessibility might relate to travel behaviour.

Before running the final regressions, initial OLS and Tobit models were estimated including the full dataset, including all neighbourhoods, to obtain a first impression of the results and identify potential issues such as multicollinearity or outliers. Subsequently, neighbourhoods with missing values were dropped from the analysis (n = 6), allowing the final regression models to be based on complete and consistent data. This step ensured the validity and reliability of the statistical findings.

To examine whether, and to what extent, the accessibility of amenities within a 15-minute walking or cycling distance influences residents' travel behaviour, two types of regression models were employed: an Ordinary Least Squares (OLS) regression model and a Tobit regression model.

An OLS regression is commonly used for estimating the linear relationship between one or more independent variables and a continuous dependent variable (Wooldridge, 2012). OLS estimates the regression coefficients by minimizing the sum of squared residuals, where each residual is the difference between the observed and predicted value (Wooldridge, 2012). While OLS provides intuitive and easily interpretable results, it assumes that the dependent variable is uncensored and normally distributed, which is not entirely appropriate in this case due to the prevalence of zeros in the dependent variable.

The Tobit model was therefore considered the primary model. The model accounts for censoring at both the lower and upper bounds, allowing for more accurate estimation when a significant number of observations are zero (Amemiya, 1984). So it is particularly suitable for this analysis, as many neighbourhoods recorded zero cycling travel trips within the selected categories.

A constant term (intercept) was included in the models to capture a non-zero baseline level of cycling travel when all predictors are set to zero. This prevents the regression line from being forced through the origin, which would likely worsen the model's fit and predictive accuracy.

To avoid issues of multicollinearity, particularly when dealing with categorical variables such as age groups and accessibility dummies, one category per variable was omitted to serve as the reference group. For example, in the cycling accessibility dummy variables, the lowest category (Cycle Dummy 1) was excluded and used as the baseline. This category primarily consists of neighbourhoods with low accessibility scores and is dominated by zeros in the dependent variable, making it an effective reference point. Similarly, among the age categories, the group aged 0-15 was excluded from the analysis, as this group does not appear in the Mobidot travel data and thus contributes no meaningful variation.

Overall, this analytical approach allowed for a nuanced investigation into how spatial accessibility to certain amenities shapes cycling travel behaviour.

3 Results

3.1 Descriptive analysis

Table 1 presents an overview of the variables used in this study. The dependent variable (share of cycling trips within 15 minutes) ranges from 0% to 83.33% across the 61 neighbourhoods included in the analysis. The log-transformed accessibility scores vary between 2.99 and 7.23.

Table 1 Overview of all the variables used in this study

Variables	Description	Unit	Min	Max
Share walking trips within 15 minutes	Proportion of walking trips (relative to all trips per neighbourhood) made within a 15-minute range, after excluding some motives	%	0	62.44
Share cycling trips within 15 minutes	Proportion of cycling trips (relative to all trips per neighbourhood) made within a 15-minute range, after excluding some motives	%	0	83.33
Walking Accessibility Score (Log)	Number of amenities within 15-minute cycling range (log-transformed)	Log-score	1.21	6.61
Cycling Accessibility Score (Log)	Number of amenities within 15-minute cycling range (log-transformed)	Log-score	2.99	7.23
Cycle dummy 2	Mid-level cycling accessibility	None	0	1
Cycle dummy 3	High-level cycling accessibility	None	0	1
% persons 15-25 years	Share of population aged 15-25	%	4	18
% persons 25-45 years	Share of population aged 25-45	%	12	44
% persons 45-65 years	Share of population aged 45-65	%	19	43
% persons 65 years and older	Share of population aged 65 and older	%	4	36
% households without children	Share of childless households	%	11	47
% education level high	Share of population with high educational attainment	%	9.1	50.3
Average income per income recipient	Average income per income recipient	€	24	69.9

3.2 Exploratory analysis

Figure 4 presents scatterplots that offered an initial insight into potential patterns and thresholds within the accessibility scores dataset. These visualizations guided the subsequent grouping of the accessibility scores into categorical dummy variables for the regression analysis.

The walking trip data contained a substantial number of zero values across the entire range, as observed in the scatterplot. Furthermore, the data indicated a consistent pattern whereby any neighbourhood with walking trip data was accompanied by cycling trip data. Due to these observations, from here on the analysis focused primarily on the cycling data, as it provided a more distinct and informative variable for the analysis.

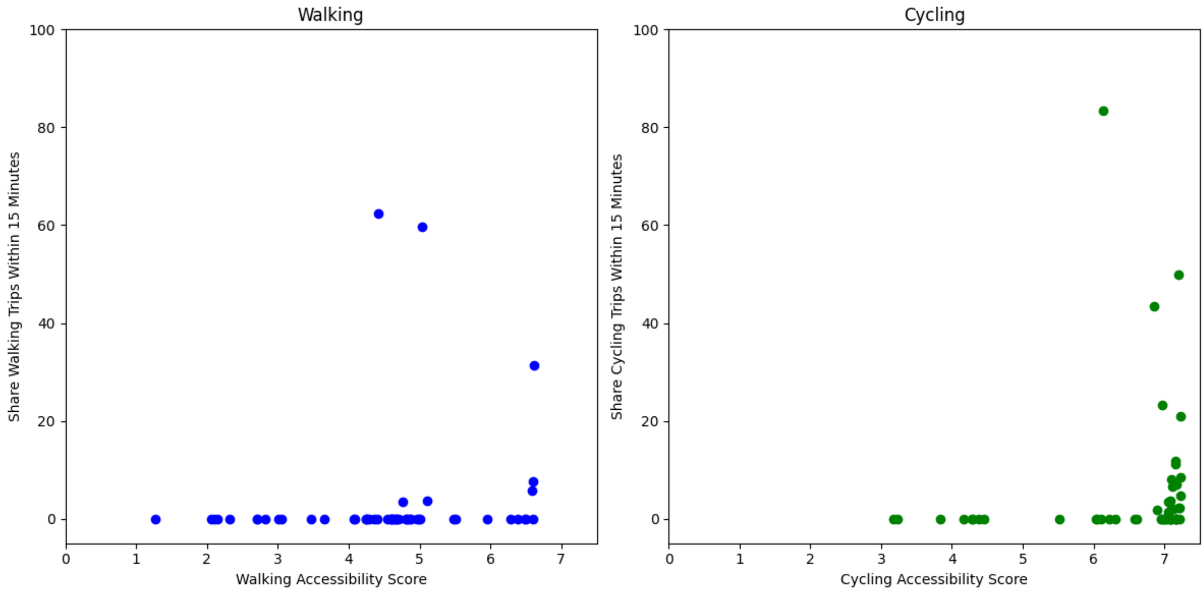


Figure 4 Scatterplots illustrating the relationship between the share of walking/cycling trips within 15 minutes and their respective accessibility scores (Mobidot, 2024; Own Data VU, 2025)

3.3 Regression analysis

3.2.1 OLS regression

To evaluate the relationship between travel behaviour and accessibility to amenities within a 15-minute cycling distance, a series of Ordinary Least Squares (OLS) regression models were performed. These models progressively incorporated additional control variables to better understand which factors might explain variation in the share of cycling trips across neighbourhoods.

Model 1a and Model 1b form the basis of the analysis. Model 1a included all available neighbourhoods (n=61), while model 1b excluded six neighbourhoods due to missing values (resulting n=55). Both models featured two categorical dummy variables representing mid-level and high-level accessibility scores (cycle dummy 2 and cycle dummy 3), using low accessibility (cycle dummy 1) as the reference category.

In model 1a, the coefficient for cycle dummy 2 is negative and statistically insignificant (coef = -1.275; $p > 0.1$), while cycle dummy 3 is positive and also insignificant (coef = 5.682; $p > 0.1$). The model explains 4.8% of the variance ($R^2 = 0.048$), with an AIC of 495.1.

Model 1b yields similar results, with cycle dummy 3 remaining positive but statistically insignificant (coef = 6.418; p> 0.1). with a slightly higher explained variance (R^2 = 0.054) and a lower AIC of 451.7, indicating a better model fit after removing incomplete observations.

In model 2, demographic control variables are introduced, specifically the proportion of residents in four age groups: 15-25, 25-45, 45-65, and 65 year or older. This addition aimed to determine whether age influenced travel behaviour. The coefficient for the age groups 25-45 years (coef = 1.236; p < 0.05) and 65 years and older (coef = 1.336; p < 0.05) are statistically significant and positive. Notably, the coefficient of cycle dummy 2 variable changed to a positive number and cycle dummy 3 changed to statistically significant (Table 2). The explained variance improves to 27.4% (R^2 = 0.274)

Model 3 introduces the percentage of households without children to the regression. The coefficient for this variable is negative and marginally significant (coef = -0.487; p < 0.1). The overall model fit decreases slightly (R^2 = 0.113).

Model 4 includes the share of highly educated residents. The coefficient for income is positive but statistically insignificant (coef = -0.154; p > 0.1). The R^2 drops to 5.4% (Table 2).

Model 5 adds average income per income recipient to assess whether economic factors play a role in cycling travel decisions. The coefficient for income is statistically insignificant (coef = -0.154; p > 0.1). The R^2 increases slightly to 5.9% (Table 2).

Finally model 6 again includes all variables combined. The coefficient for cycle dummy 3 is positive and statistically significant (coef = 12.194; p < 0.05). Age groups 25-45 years and 65 years and older remain statistically significant and positive (Table 2). The percentage of households without children is still negatively associated (coef = -0.840; p < 0.1), while educational attainment and income remain statistically insignificant. The explained variance increases to 34.3% (R^2 = 0.343).

Table 2 OLS regression results for cycling behaviour within 15 minutes (VU own data, 2025)

	Model 1a	Model 1b	Model 2	Model 3	Model 4	Model 5	Model 6
intercept	1.476 (3.792)	1.476 (3.970)	-68.815* (39.254)	15.529* (8.528)	1.016 (7.927)	7.513 (12.306)	-92.674** (41.646)
cycle dummy 2	-1.275 (5.929)	-1.218 (6.710)	0.249 (6.457)	-2.395 (6.589)	-1.096 (7.014)	-1.903 (6.885)	4.592 (7.200)
cycle dummy 3	5.682 (4.379)	6.418 (4.649)	10.161** (4.612)	6.535 (4.544)	6.455 (4.725)	6.405 (4.682)	12.194** (4.877)
% persons 15-25 years			1.737 (1.138)				1.921 (1.163)
% persons 25-45 years			1.236** (0.514)				1.202** (0.547)
% persons 45-65 years			-0.376 (0.544)				-0.085 (0.644)
% persons 65 years and older			1.336** (0.506)				1.819*** (0.563)
% households without children				-0.487* (0.263)			-0.840* (0.488)
% education level high					0.014 (0.208)		-0.021 (0.309)
average income per income recipient						-0.154 (0.297)	0.703 (0.481)
R^2	0.048	0.054	0.274	0.113	0.054	0.059	0.343
AIC	495.1	451.7	445.1	450.2	453.7	453.4	445.6
BIC	501.4	457.7	459.2	458.2	461.8	461.5	465.7

3.2.2 Tobit regression

Given the censored nature of the dependent variable, many neighbourhoods reported zero cycling trips for the selected purposes, a Tobit regression model was used as a more appropriate alternative to the OLS approach. As with the OLS models, model 1a and model 1b serves as the baseline Tobit models. Both included only two cycle accessibility dummies.

In model 1a, the coefficient for cycle dummy 2 is positive but not statistically significant (coef = 10.112; $p > 0.1$). The coefficient for cycle dummy 3 is also positive but statistically insignificant (coef = 9.525; $p > 0.1$). The model explains a very small portion of the variance with a pseudo R^2 of 0.009 and model fit statistics indicate an AIC of 203.8. (Table 3).

Model 1b excludes again six neighbourhoods due to missing values (n=55). The results remain similar to Model 1a. Cycle dummy 3 remains positive and statistically insignificant (coef = 8.137; $p > 0.1$), while cycle dummy 2 remains statistically insignificant as well (coef = 8.137; $p > 0.1$). Pseudo R^2 increases slightly to 0.010, with an improved model fit (AIC = 193.9).

In model 2, the age group variables are added to examine whether demographic composition contributes to explaining cycling behaviour. None of the age groups coefficients reach statistical significance (Table 3). The accessibility dummies remain positive, with cycle dummy 3 still insignificant (coef = 11.723; $p > 0.1$). The pseudo R^2 improves marginally to 0.013.

In model 3, the percentage of households without children was introduced. The coefficient was positive but insignificant (coef = 0.169; $p > 0.1$), which differs from the OLS results. This is different from the results from the OLS. The pseudo R^2 decreases slightly to 0.011 (Table 3).

In model 4, the share of highly educated residents is added. The coefficient for this variable is negative and statistically insignificant (coef = -0.395; $p > 0.1$). The accessibility dummies remain positive but not statistically significant. The pseudo R^2 improves to 0.019, with AIC = 194.2.

Model 5 introduces average income per income recipient. The coefficient for income remains statistically insignificant (coef = -0.045; $p > 0.1$). The Pseudo R^2 returns to 0.010, with AIC = 195.9.

Finally, model 6 includes all variables simultaneously. In this full model, nearly all variables remain statistical insignificance. Only education attainment reaches significance at the 10% level ($p < 0.1$). Although the pseudo R^2 increases to 0.033 (the highest among the Tobit models) it still indicates limited explanatory power. The AIC for this model is 203.5, reflecting the poorest model fit.

Table 3 Tobit regression results for cycling behaviour within 15 minutes (Own data VU, 2025)

	Model 1a	**Model 1b**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
intercept	13.739** (6.069)	13.886** (6.157)	-13.101 (77.687)	9.121 (15.076)	26.840** (12.019)	15.638 (21.641)	-40.198 (87.081)
cycle dummy 2	10.112 (11.505)	8.137 (12.068)	11.375 (13.195)	8.731 (12.333)	4.434 (12.111)	7.922 (12.323)	0.224 (13.825)
cycle dummy 3	9.525 (6.923)	10.101 (7.111)	11.723 (7.897)	10.404 (7.243)	9.820 (7.064)	10.058 (7.125)	9.306 (8.023)
% persons 15-25 years			0.056 (1.986)				0.343 (2.055)
% persons 25-45 years			0.074 (0.921)				0.406 (0.939)
% persons 45-65 years			0.517 (1.129)				0.258 (1.297)
% persons 65 years and older			0.427 (1.036)				-0.152 (1.222)
% households without children				0.169 (0.492)			0.231 (0.875)
% education level high					-0.395 (0.305)		-1.332* (0.740)
average income per income recipient						-0.045 (0.527)	1.917 (1.359)
Pseudo R^2	9	0.010	0.013	0.011	0.019	0.010	0.033
AIC	203.8	193.9	201.3	195.8	194.2	195.9	203.5
BIC	210.1	199.9	215.3	203.8	202.2	203.9	223.6

4 Discussion

The results of both the OLS and Tobit regression models provided limited statistical support for the hypothesised relationship between the accessibility of amenities and short cycling trips. While higher accessibility scores and certain socio-demographic factors showed some statistically significant associations in specific models, the overall explanatory power remained modest. This section reflects on these findings, exploring possible reasons for the low model performance, and considers broader limitations related to data quality, model specification, and theoretical framing.

4.1 Key findings and interpretation

Both the OLS and Tobit models included two categorical dummies for accessibility: cycle dummy 2 (mid-level accessibility) and cycle dummy 3 (high accessibility), with low accessibility as the reference. Across most models, the coefficient for cycle dummy 3 remained consistently positive, suggesting a possible association between higher accessibility and greater shares of cycling trips. However, in the initial models, these associations largely failed to reach statistical significance. Only after controlling for socio-demographic variables did cycle dummy 3 become significant in several OLS models. This suggests that simply increasing the proximity of amenities may not be sufficient to foster substantial changes in cycle travel patterns.

Such variation could indicate that spatial accessibility interacts with other factors beyond just proximity, such as personal preferences, cultural norms, or perceived safety. Previous studies have shown that the built environment influences active travel, but only in combination with attitudinal and social factors (Heinen et al., 2010; Levi & Baron-Epel, 2022). These findings underline that accessibility may be a necessary, but not sufficient, condition to encourage cycling behaviour.

In the full OLS model, neighbourhoods with a higher proportion of residents aged 25-45 years and 65+ years was significantly associated with an increased share of short cycling trips. These findings suggest that both younger adults and active seniors may engage more frequently in cycling trips within a 15-minute range. Interestingly, the share of households without children showed a marginally significant negative association in some models. In contrast, education level and income displayed no consistent or significant effects across most models, with the exception of one Tobit model where education showed a weak negative association.

While these results point to certain patterns, the overall explanatory power of both OLS (maximum R^2 = 0.343) and Tobit models (maximum pseudo R^2 = 0.033) remained low, indicating that substantial portions of variation in cycling behaviour remain unexplained.

4.2 Model limitations and missing variables

The consistently low explanatory power across all models suggests that key influencing factors of short-distance cycling are likely absent from the dataset. Although socio-demographic characteristics such as age distribution, household composition, education level, and income were included, these factors only partially capture the complex behavioural motivations behind cycling behaviour. In reality, the motivations for cycling are highly multifaceted, involving not only practical needs but also personal enjoyment, social influences, and environmental aspects (Levi & Baron-Epel, 2022). For instance, adolescents often perceive active travel as a source of personal time or an opportunity to spend time with friends (Levi & Baron-Epel, 2022). Moreover, pleasant surroundings, such as green spaces, can further enhance the appeal of active travel (Levi & Baron-Epel, 2022).

Furthermore, several key influencing factors may be missing from the model. The possible key factors could include attitudinal variables (e.g., environmental concerns, cycling experience), perceived safety, convenience of cycling infrastructure, weather conditions, and social norms (Heinen et al., 2010). Ultimately, active travel frequently reflects a personal choice shaped by a variety of contextual and subjective factors. Due to data limitations, these dimensions could not be incorporated into the current analysis, which likely contributes to the model's limited explanatory power.

4.3 Data quality and representativeness

Several limitations of the Mobidot dataset further constrain the findings. The dataset excludes trips with for example professional or commuting motives, which may have omitted relevant travel patterns, especially since such trips could still fall within 15-minutes ranges. For instance, individuals who cycle to work within a 15-minute distance may represent a significant share of active travel, yet they were removed from this analysis.

The walking data also showed a high proportion of zero observation which may suggest that walking is not a practical or an attractive option in certain neighbourhoods. This could reflect mismatches between residential preferences and the distribution of amenities, or broader structural barriers like infrastructure quality or perceived safety.

Moreover, selection bias may arise because the data only include participants who voluntarily used the SESAMO smartphone app. In some neighbourhoods, only a few residents participated, even if the area has thousands of residents. Consequently, using such sparse individual data to derive neighbourhood-level travel behaviour can lead to biased data or lack of representation. This issue is further complicated by the fact that socio-demographic averages were used to explain travel behaviour, while the actual individuals may not reflect the true composition of the neighbourhoods. As a result, there is a risk of mismatch between the few individuals and the aggregated statistics, which may weaken the explanatory power of the models and introduce additional noise.

This limited representativeness may also help explain the extremely high variation in cycling behaviour observed in the data, with some neighbourhoods showing 0% cycling trips and others exceeding 50%. However, beyond data limitations, this variation might also reflect a real divide between highly urbanized central neighbourhoods and more remote areas. Such spatial polarisation suggests that behavioural change may not occur gradually but rather depends on surpassing certain accessibility thresholds. In areas where the level of local amenities remains below this threshold, promoting active travel may prove challenging, whereas once sufficient facilities are present, active travel becomes a more realistic and attractive option for residents. This has important implications for policy, as efforts to stimulate behavioural change are probably unlikely to succeed unless a critical mass of accessible amenities is provided.

4.4 Challenges in defining '15-minute accessibility'

The strict 15-minute threshold used to define short-distance trips may also oversimplify actual mobility patterns. A trip lasting 16 or 17 minutes may still align with the conceptual principles of the 15-minute city, especially when considering personal mobility levels. As mentioned in the introduction, the 15-minute city concept is a fluid concept (C40 Cities Climate Leadership Group

& C40 Knowledge Hub, 2023). A more nuanced approach could involve analysing trip durations as a continuous variable or categorising them into layered thresholds (e.g., 10, 15, 20 minutes), allowing a more realistic representation of accessibility. Notably, very few of existing studies have applied different time thresholds for different types of urban amenities, despite the fact that it is questionable whether all services should fall within the same time range (Papadopoulos et al., 2023). The flexibility to define time thresholds specific to each amenity type is considered a valuable addition to measuring accessibility but should be guided by multiple contextual criteria (Papadopoulos et al., 2023). Moreover, within a 15-minute accessibility range, facilities may be located at opposite ends, meaning that trips covering all these facilities could easily exceed 15 minutes in actual travel time. These longer trips are not captured in the Mobidot data, which excludes trips exceeding 15 minutes, despite the resident still residing within the designated 15-minute range. This omission could lead to an underestimation of cycling travel patterns and misinterpretation of accessibility impacts.

4.5 Broader behavioural insights

The modest role of demographic variables also suggests that lifestyle or value-based factors may be more predictive of cycling travel behaviour than basic demographic characteristics. This points to the potential for more targeted segmentation in policy approaches, rather than one-size-fits-all planning.

Lastly, it is also important to note that spatial accessibility, even if objectively high, does not automatically lead to actual use. Residents may choose to travel beyond the 15-minute catchment area for preferred services due to loyalty habit, or perceived quality differences. These behavioural patterns means that accessibility, as measured objectively, may not reflect actual usage patterns. Integrating spatial analyses with qualitative or behavioural insights in future research, such as surveys or interviews, could enrich the understanding of how and why people choose certain travel behaviours despite local availability.

5 Conclusion

This study examined the extent to which actual cycle travel behaviour aligns with the spatial accessibility of amenities within the framework of the 15-minute city, using the Dutch municipality of Alkmaar as a case study. While the theoretical promise of the 15-minute city suggests that residents will use nearby services if they are available, the findings of this research point to a more complex reality.

The OLS and Tobit regression models showed that higher accessibility levels (cycle dummy 3) were generally associated with higher shares of short cycling trips. However, these associations only reached statistical significance in some models, and overall explanatory power remained modest. Socio-demographic variables contributed limited additional explanation, although the share of residents aged 25-45 and 65+ years emerges as consistent positive predictors in the OLS models. Educational attainment showed mixed effects: insignificant in most models but reaching significance in one Tobit model. Income and household type also showed no strong or consistent associations. These results suggest that accessibility alone does not fully explain short distance cycling behaviour, and that important behavioural or environmental factors are likely missing from the models.

Several limitations of the dataset further constrain interpretation. The Mobidot data is based on a small, self-selected sample of app users, potentially introducing selection bias. The exclusion of some trips also may have omitted an important share of short-distance cycling. Furthermore, the rigid application of a 15-minute threshold may oversimplify actual travel patterns, excluding trips that slightly exceed this limit but still align with the spirit of the 15-minute city concept. Together, these limitations may have led to underestimation of cycle travel behaviour and weakened the explanatory power of the statistical models.

In light of these findings, it becomes clear that relying solely on quantitative spatial models is insufficient to understand the reasons for making a sustainable trip. An multi-method approach that combines spatial models with qualitative data to capture residents' motivations would be preferred for a more valid answer. Future research should also incorporates behavioural insights, characteristics of the built environment, personal factors, and higher resolution mobility data. Additionally, a more flexible treatment of travel time thresholds is recommended; moving beyond strict binary definitions to layered distance thresholds that better reflect the nuances of sustainable travel.

From a policy perspective, the limited role of accessibility and socio-demographics variables point toward the need for more targeted segmentation in policy approaches, rather than one-size-fits-all planning. Assuming that spatial proximity and availability alone will shift behaviour will not work.

In Conclusion, this research highlights both the promise and the complexity of implementing the 15-minute city in practice. While spatial accessibility remains a vital foundation, actual travel behaviour is shaped by a broader set of factors. Addressing these elements more comprehensively will be key to turning the 15-minute city from a theoretical ideal into a lived urban reality.

6 References

Abdelfattah, L., Deponte, D., & Fossa, G. (2022). The 15-minute city: interpreting the model to bring out urban resiliencies. *Transportation research procedia*, *60*, 330-337.

Amemiya, T. (1984). Tobit models: A survey. *Journal of econometrics*, *24*(1-2), 3-61.

Badawi, Y., Maclean, F., & Mason, B. (2018). The economic case for investment in walking.

C40 Cities Climate Leadership Group & C40 Knowledge Hub. (2020, July). How to build back better with a 15-minute city. Accessed June 26, 2025. *https://www.c40knowledgehub.org/s/article/How-to-build-back-better-with-a-15-minute-city?language=en_U*

C40 Cities Climate Leadership Group & C40 Knowledge Hub. (2021, May). Why every city can benefit from a '15-minute city' vision. Accessed June 26, 2025. *https://www.c40knowledgehub.org/s/article/Why-every-city-can-benefit-from-a-15-minute-city-vision?language=en_US*

C40 Cities Climate Leadership Group, & C40 Knowledge Hub. (2023, March). *15-minute cities: Debunking the myths*. *Accessed June 26, 2025*. https://www.c40knowledgehub.org/s/article/15-minute-cities-Debunking-the-myths

Heinen, E., Van Wee, B., & Maat, K. (2010). Commuting by bicycle: an overview of the literature. *Transport reviews*, *30*(1), 59-96.

Khavarian-Garmsir, A. R., Sharifi, A., & Sadeghi, A. (2023). The 15-minute city: Urban planning and design efforts toward creating sustainable neighborhoods. *Cities*, *132*, 104101.

Knap, E., Ulak, M. B., Geurs, K. T., Mulders, A., & Van Der Drift, S. (2023). A composite X-minute city cycling accessibility metric and its role in assessing spatial and socioeconomic inequalities–A case study in Utrecht, the Netherlands. *Journal of Urban Mobility*, *3*, 100043.

Levi, S., & Baron-Epel, O. (2022). "It'sa kind of freedom": adolescents and parents speak about motivations for active travel and COVID-19. *International Journal of Qualitative Studies on Health and Well-being*, *17*(1), 2130508.

Manifesty, O. R., & Park, J. Y. (2022). A Case Study of a 15-Minute City Concept in Singapore's 2040 Land Transport Master Plan: 20-Minute Towns and a 45-Minute City. *International Journal of Sustainable Transportation Technology*, *5*(1), 1-11.

Mobidot – Create better multimodal insight & services with personal travel data. (n.d.). *https://mobidot.nl/*

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-Minute City": Sustainability, resilience and place identity in future post-pandemic cities. *Smart cities*, *4*(1), 93-111.

Papadopoulos, E., Sdoukopoulos, A., & Politis, I. (2023). Measuring compliance with the 15-minute city concept: State-of-the-art, major components and further requirements. *Sustainable Cities and Society*, *99*, 104875.

Thomas, T., Geurs, K. T., Koolwaaij, J., & Bijlsma, M. (2018). Automatic trip detection with the Dutch mobile mobility panel: towards reliable multiple-week trip registration for large samples. *Journal of Urban Technology*, *25*(2), 143-161.

Wooldridge, J. M. (2012). *Introductory Econometrics*: *A Modern Approach*. South-Western, Cengage Learning.

Appendix

A1. Explanatory analysis

In this script, two scatterplots were created to explore the relationship between the accessibility score for walking and cycling, on the one hand, and the percentage of trips within 15 minutes on foot or by bicycle, on the other.

```
# twee scatterplots in een plaatje
fig, axs = plt.subplots(1, 2, figsize=(12, 6))

# Eerste scatterplot: walking
axs[0].scatter(df_regression['Walk_Log_Score'], df_regression['Walk15_adjust_percentage'], color='blue')
axs[0].set_xlabel('Walking Accessibility Score')
axs[0].set_ylabel('Share Walking Trips Within 15 Minutes')
axs[0].set_title('Walking')

# Tweede scatterplot: cycling
axs[1].scatter(df_regression['Cycle_Log_Score'], df_regression['Cycle15_adjust_percentage'], color='green')
axs[1].set_xlabel('Cycling Accessibility Score')
axs[1].set_ylabel('Share Cycling Trips Within 15 Minutes')
axs[1].set_title('Cycling')

# Layout zelfde axis
for ax in axs:
    ax.set_xlim(0 , 7.5)
    ax.set_ylim(-5, 100)

plt.tight_layout()
plt.show()
```

A2. Cleaning and preparing the variables for regression

In this script, the dataset was cleaned and prepared for analysis. First, several columns containing percentages and income data were converted from text format (object) to numeric format (float). Next, rows with missing values only were removed, as well as rows containing some missing values.

```
df_regression['% Cycling trips within 15 minutes'] = df_regression['% Cycling trips within 15 minutes'].astype(str).str.replace(',', '.').astype(float)
df_regression['cycle dummy 1'] = df_regression['cycle dummy 1'].astype(str).str.replace(',', '.').astype(float)
df_regression['cycle dummy 2'] = df_regression['cycle dummy 2'].astype(str).str.replace(',', '.').astype(float)
df_regression['cycle dummy 3'] = df_regression['cycle dummy 3'].astype(str).str.replace(',', '.').astype(float)
df_regression['average income per income recipient'] = df_regression['average income per income recipient'].astype(str).str.replace(',', '.').astype(float)
df_regression['% education level high'] = df_regression['% education level high'].astype(str).str.replace(',', '.').astype(float)

df_regression = df_regression.dropna(how='all')

print(df_regression.shape)
print(df_regression.isna().sum())
df_regression.head()
```

Verborgen uitvoer tonen

```
#drop all rows with NaN of certain columns
df_regression = df_regression.dropna(subset=['% persons 0-15 years', '% persons 15-25 years',
       '% persons 25-45 years', '% persons 45-65 years',
       '% persons 65 years and older', '% households without children',
       '% households with children', '% education level high',
       'average income per income recipient'])

print(df_regression.shape)
print(df_regression.isna().sum())
```

A3. OLS regression models

In this script, several Ordinary Least Squares (OLS) regression models were specified and estimated to explain the share of cycling trips made within 15 minutes. The dependent variable was regressed on different combinations of explanatory variables. The script automatically extracted coefficients, standard errors, p-values, and calculated R^2, AIC and BIC scores for each model. A custom formatting function was used to present the results clearly and indicate statistical significance levels.

```python
import pandas as pd
import numpy as np
import statsmodels.api as sm
from IPython.display import display

# endog
endog = df_regression['% Cycling trips within 15 minutes']

# Models
model_specs = {
    "Model 1": ['cycle dummy 2', 'cycle dummy 3'],
    "Model 2": ['cycle dummy 2', 'cycle dummy 3', '% persons 15-25 years', '% persons 25-45 years', '% persons 45-65 years', '% persons 65 years and older'],
    "Model 3": ['cycle dummy 2', 'cycle dummy 3', '% households without children'],
    "Model 4": ['cycle dummy 2', 'cycle dummy 3', '% education level high'],
    "Model 5": ['cycle dummy 2', 'cycle dummy 3', 'average income per income recipient'],
    "Model 6": ['cycle dummy 2', 'cycle dummy 3', '% persons 15-25 years', '% persons 25-45 years', '% persons 45-65 years', '% persons 65 years and older',
            '% households without children', '% education level high', 'average income per income recipient'],
}

# Formatting p-value and coef
def format_coef(coef, stderr, pval):
    if pd.isna(coef) or pd.isna(stderr):
        return ""
    stars = ''
    if pval < 0.01:
        stars = '***'
    elif pval < 0.05:
        stars = '**'
    elif pval < 0.1:
        stars = '*'
    return f"{coef:.3f}{stars}\n({stderr:.3f})"

# Results gathering
results_table = {}

for model_name, variables in model_specs.items():
    # exog and intercept
    exog = df_regression[variables].copy()
    exog['intercept'] = 1
    exog = exog[['intercept'] + variables]

    # OLS regression
    model = sm.OLS(endog, exog)
    res = model.fit()

    # Results formatting
    result_column = {}
    for var in exog.columns:
        coef = res.params[var]
        pval = res.pvalues[var]
        stderr = res.bse[var]
        result_column[var] = format_coef(coef, stderr, pval)

    # R-squared
    result_column['R-squared'] = f"{res.rsquared:.3f}"

    results_table[model_name] = result_column
    #AIC
    results_table[model_name]['AIC'] = f"{res.aic:.3f}"
    #BIC
    results_table[model_name]['BIC'] = f"{res.bic:.3f}"
```

A4. Tobit regression models

In this script, several Tobit regression models were specified and estimated to explain the share of cycling trips made within 15 minutes. The dependent variable was regressed on different combinations of explanatory variables. The script automatically extracted coefficients, standard errors, p-values, and calculated pseudo R^2 , AIC and BIC scores for each model. A custom formatting function was used to present the results clearly and indicate statistical significance levels.

```
# Censuurcolumn
df_regression['censored'] = np.where(df_regression['% Cycling trips within 15 minutes'] <= 0, 1, 0)

# endog en cens
endog = df_regression['% Cycling trips within 15 minutes']
cens = df_regression['censored'].values

# Models
model_specs = {
    "Model 1": ['cycle dummy 2', 'cycle dummy 3'],
    "Model 2": ['cycle dummy 2', 'cycle dummy 3', '% persons 15-25 years', '% persons 25-45 years', '% persons 45-65 years', '% persons 65 years and older'],
    "Model 3": ['cycle dummy 2', 'cycle dummy 3', '% households without children'],
    "Model 4": ['cycle dummy 2', 'cycle dummy 3', '% education level high'],
    "Model 5": ['cycle dummy 2', 'cycle dummy 3', 'average income per income recipient'],
    "Model 6": ['cycle dummy 2', 'cycle dummy 3', '% persons 15-25 years', '% persons 25-45 years', '% persons 45-65 years',
                '% persons 65 years and older', '% households without children', '% education level high', 'average income per income recipient'],
}

# Formatting p-value and coef
def format_coef(coef, stderr, pval):
    if pd.isna(coef) or pd.isna(stderr):
        return ""
    stars = ''
    if pval < 0.01:
        stars = '***'
    elif pval < 0.05:
        stars = '**'
    elif pval < 0.1:
        stars = '*'
    return f"{coef:.3f}{stars}\n({stderr:.3f})"

# Results gathering
results_table = {}

for model_name, variables in model_specs.items():
    # Exog and intercept
    exog = df_regression[variables].copy()
    exog['intercept'] = 1
    exog = exog[['intercept'] + variables]

    # Tobit-regression
    model = tobit.Tobit(endog, exog, cens)
    res = model.fit(disp=False)

    # results per variable
    result_column = {}
    # Iterate using index to access numpy array results
    for idx, var in enumerate(exog.columns):
        # Access elements using integer index
        coef = res.params[idx]
        pval = res.pvalues[idx]
        stderr = res.bse[idx]
        result_column[var] = format_coef(coef, stderr, pval)

    results_table[model_name] = result_column

    #Pseudo R2 at the end
    results_table[model_name]['Pseudo R-squared'] = f"{res.prsquared:.3f}"
    # AIC
    results_table[model_name]['AIC'] = f"{res.aic:.3f}"
    # BIC
    results_table[model_name]['BIC'] = f"{res.bic:.3f}"
```