Faculdade de Engenharia da Universidade do Porto

Factors affecting transport mode choices in Porto Metropolitan Area: a discrete choice model

Hudyeron Hubner do Carmo Rocha

Dissertation submitted for the conclusion of the requisites of the degree
Master's In Civil Engineering

Supervisor: Sara Maria Pinho Ferreira (Phd, Professor)
Co-supervisor: António Manuel Cabral Vieira Lobo (PhD)

March 07th, 2022

“Every good gift and every perfect gift
is from above, coming down from the
Father of lights, with whom there is
no variation or shadow due to change.”

James, 1:17

Abstract

Sustainable urban mobility is a goal that supports the accomplishment of the post-carbon cities paradigm. Tackling congestion and improving air quality, accessibility and sustainability in most cities will require substantial changes in the mobility patterns of the population.

To understand these patterns and to influence people towards more sustainable behaviours, it is necessary to know the actual transport mode choices made by the population on a daily basis. The objective of this master's thesis is to analyse the transport mode choice in Porto Metropolitan Area using a discrete choice model, specifically the multinomial logit. The results allow for propose a better understanding of the behaviour of citizens in their travels, so that policymakers can promote strategies seeking to increase sustainable mobility, focus on the quality of public transports, and promote campaigns showing the benefits of active commuting.

To do that, the database generated by the 2017 Mobility Survey conducted by Statistics Portugal (INE) will be used. Several variables will be retrieved from this database, allowing the identification of the factors affecting the choice of transport mode and the estimation of its magnitude. After analysing the results of the model's application, it was found that the independent variables that had a greater explanatory significance for the model, was the "distance", "purpose of trip" and the "per capita income", being these decisive factors for the modal choices of individuals.

The advantage of applying the proposed model is the ability to identify the relationships between an individual's travel behaviour and his modal choice, thus allowing the influence of variables on the modal choice to be observed.

Acknowledgments

First, I would like to thank God for the strength, patience, and resilience He gave me every single day to get this far and never have abandoned me in times of need because I felt His presence by my side throughout my journey.

I want to thank my professors António José Fidalgo do Couto, José Pedro Maia Pimentel Tavares, to my supervisor Professor Sara Maria Pinho Ferreira and my co-supervisor António Manuel Cabral Vieira Lobo for helping me conduct my work with great dedication and mastery, thank you for your support and contribution, not only for my dissertation, but also to my growth and, especially, for your patience throughout the whole process.

I want to thank my whole family for always believing in me and not sparing efforts to provide all the necessary incentives to make this possible. Especially I thank my parents Geraldo and Shirley, also my grandparents Darli and Marlene for their unconditional love and support, for the affection and care they gave me throughout all my existence. You are the pillars of my development as a human being, and I am grateful for everything. Thank you for always believing in me.

I would also like to thank my friends Paulo, Daniel, Henrique, and Carolina for being great study partners, always looking for ways to obtain the best results, and in particular Suellen for the encouragement she gave to choose Territory Planning and Transports when I did not know which path to choose in my Master; and I thank my great friend Fabiola for her encouragement, her enormous understanding, and her precious advice, as well as for the critical and careful reading of the preliminary versions of the thesis, thus contributing to its improvement.

I would like to thank all the teachers and masters who contributed to my studies throughout my academic career and always encouraged me to look forward to knowledge.

And finally, a thank you to all the special people that crossed paths in my journey, always contributing to the realization of this dissertation.

This work is a result of project DynamiCITY: Fostering Dynamic Adaptation of Smart Cities to Cope with Crises and Disruptions, with reference NORTE-01-0145-FEDER-000073, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).

Contents

List of figures

List of tables

Abbreviations

AML	Metropolitan Areas of Lisbon
AMP	Metropolitan Areas of Porto
CAWI	Computer-assisted Web-based Interviewing
CAPI	Computer Assisted Personal Interviewing
EU	European Union
ICT	Information and Communication Technologies
INE	*Instituto Nacional de Estatística* in Portugal
IMob	Mobility Survey (*Inquérito à Mobilidade*)
MNL	Multinomial logistic regression
IBM	International Business Machines
PHV	Private hire vehicle
SPSS	Statistical Package for the Social Sciences
UN	United Nations

Chapter 1

Introduction

1.1 - Urban Mobility

In Population and Peace in the Pacific, Thompson (1946), foresaw that the contemporary problem facing all nations would be rapid population growth in urban areas. Modern transport systems and innovations in public health would reduce population mortality, but the affected countries and their rulers were not promoting industrialization and urbanization that plays an important role in increasing productivity and quality of life.

The great phenomenon of urbanization is characterized by "the movement of people from small communities concerned chiefly or solely with agriculture to other generally larger communities, whose activities are primarily centered in government, trade, manufacture or allied interest", Thompson (1935).

The process of urbanization in Portugal took place through the progressive concentration of people and activities along the coastline of the country, mainly in the two major cities: Lisbon and Porto. The polarization of the urban system around these two cities accentuated a historical trend for their development because historically the two largest national ports are at these cities, also both cities have in themselves the best natural conditions and accessibility of the country, so that the interior could not keep up with its growth rate.

In recent decades, due to the large migration of people to these cities, these trends have been exacerbated on a large scale, with the expansion of numerous suburbs and satellite towns around these two centers, forming an extensive metropolitan area around Lisbon and Porto.

According to the National Geographic Society (2011), "an urban area is the region surrounding a city. Most inhabitants of urban areas have non-agricultural jobs. Urban areas are very developed, meaning there is a density of human structures such as houses, commercial buildings, roads, bridges, and railways".

In 2014 a survey conducted by the European Union (EU) proved that 40.2% of people already lived in densely populated areas (urban centers), 27.8% in sparsely populated or considered rural areas, and 32.0% in intermediate areas (cities and suburbs). Through this survey, it was possible to verify that there are significant discrepancies between the EU Member States, where some of them have a predominantly urban population, while others have a predominantly rural population. In the world, according to the United Nations (UN, 2019), "55% of the world's population already lives in urban areas, a proportion expected to increase to 70% by 2050".

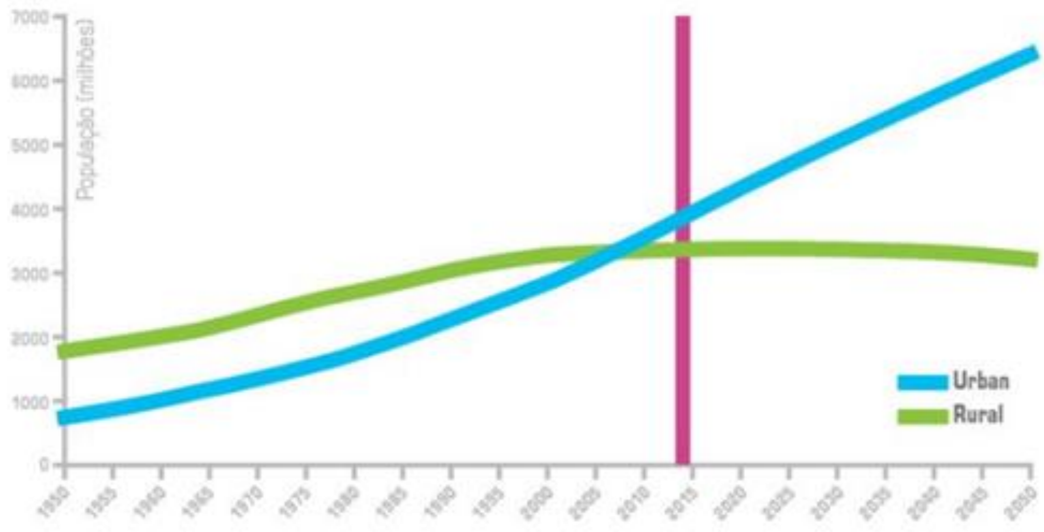

Figure 1 - Rural and urban population (in millions) in the world (WUP, 2019).

In Figure 1 is possible to see that 65.9% of Portugal's population already live in urban areas.

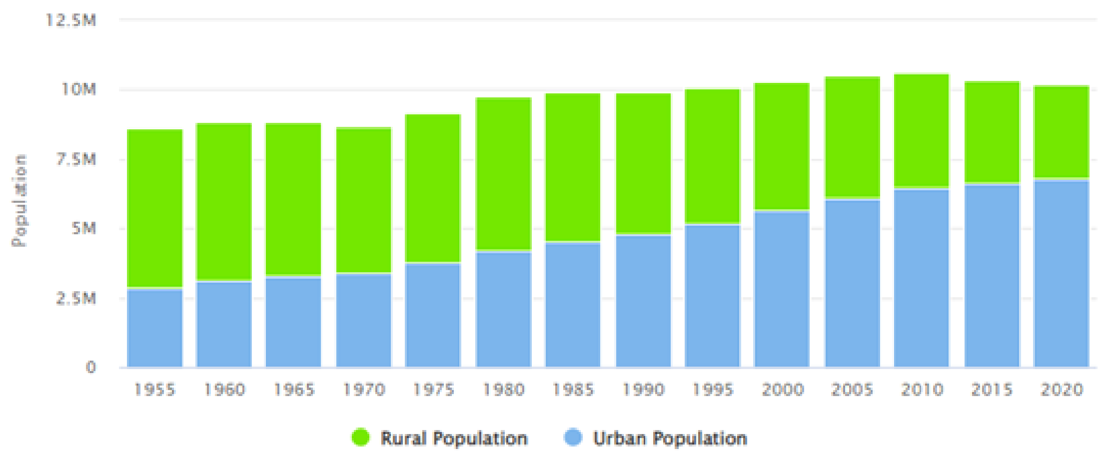

Figure 2 - Rural and urban population (in millions) in Portugal (Worldometers, 2020)

When a significant number of people migrate from rural areas to major urban centers in search of a better quality of life, they consequently end up directly affecting urban mobility.

Urban Mobility is the condition that allows people to move from one place to another in order to develop social and economic relations depending on the performance of the transportation system and the individual's characteristics (Tagore & Sikdar, 1995). (Thomson,1977) in his book "Great Cities and Their Traffic", provides an analysis of the seven reasons why people feel dissatisfied with the transportation systems in their cities, which are:

1. Movement and traffic congestion: Traffic congestion occurs when urban transport networks are no longer able to accommodate the volume of movements using them.

2. Crowding of Public Transport: The large number of people in public transport at peak times leads to borderline situations in social coexistence. A high number of trips of the day are made at peak hours, in which long queues happen at the bus stops, agglomeration at terminals, stairs, and ticket offices take place, as well as excessively long periods of travel with overcrowded vehicles.

3. Inadequacy of off-peak public transport: Many public transport companies may have sufficient vehicles to meet peak-hour demand, but in contrast with large work schedules of their employees, such companies may provide a limited quantity and large time difference between cars to adjust timetables.

4. Difficulties for pedestrians: Pedestrians are the biggest victims of traffic crashes and the attempts to improve their safety have generally failed to address the source of the problem (traffic volume and speed) and have instead focused on restricting movement on foot by making large areas as 'off limits' forcing pedestrians to use walkways and subways that are inadequately cleaned or policed. In addition, there is obstruction by parked cars and increasing pollution of the urban environment, with traffic noise and exhaust fumes affecting those on foot most directly.

5. Parking Difficulties: Many drivers who commute during peak hours search for parking spaces, and the slowness in finding them slows down urban traffic.

6. Environmental Impact: Besides pollution, traffic noise is also an environmental problem caused by traffic in urban areas. Traffic noise is irritating and can cause annoyance when walking or performing other activities in urban areas can become disturbed due to this factor, traffic noise can also penetrate inside buildings, making work difficult as it hinders concentration and conversation, it can also impair sleep and relaxation in domestic life.

7. Air Pollution: Smoke from motor vehicles is an extremely polluting activity. Although there are numerous other activities that cause environmental pollution as a result of the tremendous increase in the number of vehicles, it is only now that society is beginning to come to grips with the devastating and dangerous consequences that the excesses of motor vehicles use can cause.

As traffic volumes increase, so does air pollution. According to figure 3, no urban area is currently free from the effects of engine fumes that contribute to the increase in air pollution, especially those from poorly maintained diesel engines. In large cities, these gases are responsible for creating extremely unpleasant air pollution that is quite offensive to health.

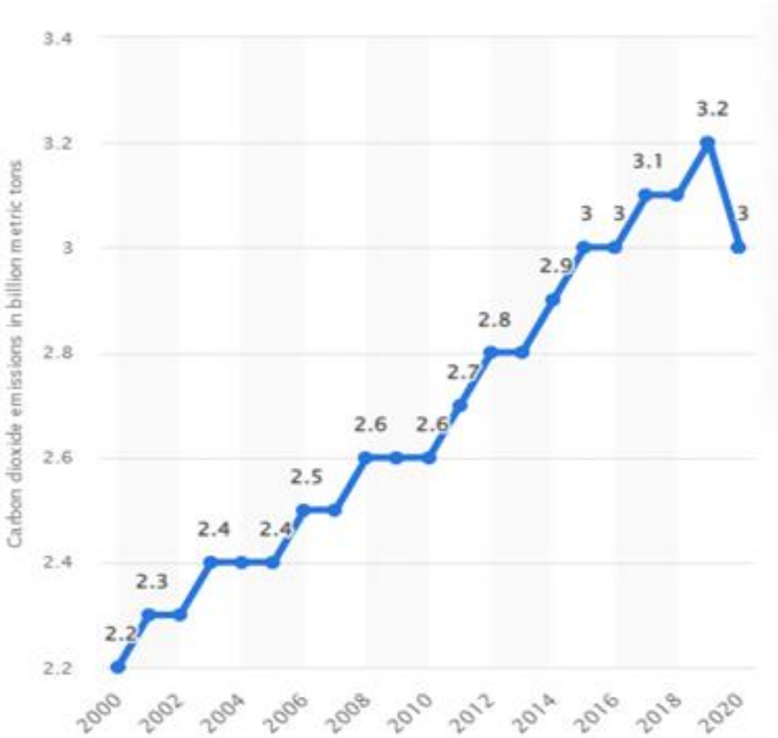

Figure 3 - Carbon Dioxide emission in billion tons worldwide from 2000 to 2020 (Statista, 2020)

When thinking about urban mobility, one must also take into consideration the power to offer citizens the right to enjoy the goods and services available in large urban centers, rather than just thinking about transport mode and traffic.

Ferreira et al., (2013), points out that sustainable mobility is promoted through integrated policies that can contribute directly or indirectly to the sustainable development of territories and society. According to Mackett (2003), encouraging alternative transports mode may be a more viable and more effective solution for urban mobility and Ramos (2001) defines the following objectives to increment sustainable mobility:

- Increasing non-motorized mobility;
- Increase the use of public transport;
- Reduce mobility by car;
- Reduce energy consumption;

- Promote intersectoral cooperation;
- Promote access to information and awareness among the population;
- Promote quality of life and general well-being.

Even though urban planning solutions end up prioritizing the private car over other modes of transportation (Souza, 2015) an increase in the number of private vehicles circulating in large centers ended up causing public transportation to lose space due to the great competition with the growing number of taxis or private hire vehicle (PHV), motorbikes and the private cars themselves, for being more comfortable alternatives. Thus, decreasing the performance of public transports, making their supply more expensive and no longer an attractive alternative for those who have other more comfortable transport options, entailing great difficulties for the development of the transport sector.

Considering the challenges of urbanization and the growing importance of promoting sustainable urban mobility, the creation of an effective and integrated urban mobility system starts with understanding consumer expectations by seeking to analyse what individuals want from urban mobility and to what extent transport systems are failing.

The adoption of Smart City initiatives in cities is being considered as a solution to help mitigate the impact of urbanization through the creation of solutions that allow effective and more efficient use of city resources.

The concept of a Smart City has several definitions and scope, according to Kourtit and Nijkamp (2012), Smart Cities are the result of creative and knowledge strategies aiming to improve socio-economic, environmental, logistical, and competitive performance of cities.

These cities are based on a set of:

- Human capital;
- Infrastructure;
- Share capital;
- Entrepreneurship capital.

Seeking to manage environmental, social, and resource economics sustainability, Smart Cities enable citizens and local government authorities to work together to launch initiatives and use available smart technologies to manage the assets and resources of the growing urban environment. A smart city should provide an urban environment with a high quality of life for residents as well as manage economic growth, providing a range of services to citizens while generating new revenue streams and operational efficiencies with the aim of saving the government and citizens money.

Smart Cities can or should use Information and Communication Technologies (ICT) such as sensors, monitoring systems, wireless networks, autonomous devices, and mobile applications, seeking to collect data to apply this data in urban projects in an easy way, taking advantage of the fact that information collection can occur in real-time.

With this idea in mind, in this century more than ever, cities need to build intelligent transport approaches with much greater consideration for future urban development, as only then can they suggest new desirable urban environments and prioritize the most critical and vital aspects for the future (Bratzel, 1999).

It is important to identify the reasons for travel that influence people's choice of mode of travel. The concept of intelligent transportation has been recognized by Smith et. all (2005), as a critical point to lead the community to a desirable local future through joint and shared actions focused on the citizen.

1.2 - Objective of the study

In order to monitor travel behavior in Portugal through mass evidence, the Instituto Nacional de Estatística in Portugal (INE), in partnership with the Statistical office of the European Union (Eurostat), which is the institution responsible for publishing high-quality statistics and indicators at a European level, allowing for comparison between countries and regions, conducted a mobility survey (IMob, Inquérito à Mobilidade) from October to December 2017, in the Metropolitan Areas of Lisbon (AML) and Porto (AMP). The data collected were related to people's place of residence, trip purposes, travel modes, travel distance and duration, among other indicators.

These surveys in AMP and AML were also conducted between April and June of the year 2000 with different metropolitan areas but with similar size sample. In 2012, Gonçalves developed a Ph.D. thesis named “Relevant factors for modal choice in urban areas” using data from the 2000’s survey for the AMP and this work will analyse some assumptions in the next chapters.

According to INE (2018), the main objective of this survey was to know the daily mobility patterns of the population of Lisbon and Porto, as well as their motivations when deciding the transport mode and their opinions about individual and collective transport.

This survey collected data on approximately 100,000 trips in each metropolitan area. According to the most recent estimates regarding the resident population in both areas, there are approximately 1.74 million people living in the AMP and 2.87 million in the AML. Together, these two areas account for 44.5% of the resident population in Portugal, almost half the population of mainland Portugal. Therefore, the sample of the survey corresponds to approximately 5.5% of the resident population in both areas together. In the case of AMP, which is the focus of this study, the sample corresponds to about 5.7% of the resident population.

According to the INE (2018), the car was the main used transport mode in the journeys made by residents in the AMP, accounting for 67.6% of the journeys, considering all days of the week. In the previous survey in the year 2000, the percentage of car usage was 50% (INE, 2002).

In the case of the AMP, the decision-makers face a major challenge in order to ensure a sustainable future for the resident population Indeed in recent years, there has been a growing change in mobility patterns, as a result of the intensification of motorization rates in areas with larger urban agglomerations. The idea that the quality of life revolved around the ease of travel by car, due to the comfort and agility of commuting, ended up causing too much disruption in large urban centers in terms of mobility.

A study carried out a company specialized in location technology TOMTOM Traffic Index throughout 2020, it analysed the traffic of 516 cities in 57 countries, in the results the city of Porto surpassed the city of Lisbon, becoming the most congested Portuguese city in 2020. Thereby, traffic congestions, noise, and air pollution are already some problems caused by the increase in motor vehicles, generating physical and mental health problems, which can bring a drop in the quality of life at urban centers. (Gärling and Steg, 2007).

Plus, following the guidelines of the European Commission's European Green Deal published in 2021, policymakers should adopt a strategy for a combination of measures to improve mobility by seeking to implement greater use of public transport and intermodal transport systems, investing in cycling and walking infrastructure, and developing communication campaigns that focus on the benefits of active mobility.

In order to identify the best policies and mobility measures to promote sustainable mobility, it is important to understand the characteristics of the people and the trips associated with the choice of each transport mode. In this context, this thesis study aims to identify the factors that interfere with individuals' choice of transport mode in their regular activities, mainly the sociodemographic characteristics and the characteristics of the trip as well as to understand the magnitude of each of them on the probability of choosing each transport mode.

For this analysis, a discrete choice model was applied, in particular, the multinomial logit model that considers as dependent variable the transport mode choice and, as independent variables, those that characterize the individual as well as the trip made.

The results obtained will allow us to better understand in which context a certain mode of transport is used, and it may constitute support for decision-making on public transport policies, that to be more effective, should be targeted to certain travel profiles and individuals. Moreover, considering the current challenges of cities, namely the creation and implementation of sustainable mobility plans, as well as the introduction of ICTs in the context of Smart Cities, it is essential to know the existing situation.

1.3 - Thesis structure

This work is organized into 4 chapters.

In the present chapter 1, an introduction to the theme of the work is made, with its framework, a review of the literature, and the objectives.

In chapter 2, the methodology used is presented along with the theories for modal choice modeling and the used model.

Then, chapter 3 refers to the characterization of the study area, the description, and organization of the variables used in the chosen model, such as the descriptive analyses and the results obtained.

Finally, chapter 4 includes the conclusions of the results obtained through the model application and plans for future research.

Chapter 2

Methodology

2.1 – Review of literature

According to Gonçalves (2012), in recent decades, discrete choice models have been the ones that have most potentiated evidence in studies on people's modal choices. These models, although they do not allow an analysis of the influence of the chain of the motives and the different trips throughout the day of the population, are presented as a powerful modeling tool, with a solid theoretical foundation and a wide application to different real situations.

The theoretical basis of discrete choice models makes them indispensable tools for the field of transport modeling, especially for studies that need to simulate the way individuals choose their modal options. Its most common models are the Mixed Logit, the Probit, and the multinomial Logit. This family of models presents a variety of formulations and can generate models based on simpler hypotheses and more complex models that can simulate possible scenarios and future environments.

According to van Wee (2002), there are still different patterns of travel within socio-economic and population groups demographically homogeneous and that is why some researchers have recently recommended the inclusion of more subjective variables, such as personal lifestyles, attitudes and preferences when building the model. As a result, subjective variables have been introduced in recent decades in empirical work on the relationship between the built environment and travel behavior (Handy et al., 2005; van Wee, 2009).

According to Salomon and Ben-Akiva (1983), these variables can refer to the individual's orientation to general topics such as leisure, family and work, thus managing to describe the individual in a broader context than those commonly used, such as income, age and family structure (Kitamura, 2009).

Therefore, it is always relevant to explore variables will have the greatest effects on travel behavior, besides looking only at the characteristics related to mobility or infrastructure. For example, Timmermans et al. (2002), after analysing the strengths and weaknesses of various models, applied the multinomial logit model and concluded that the work situation, age, the presence of children in the family, and availability of cars are related to the trip duration and to the mode choice behavior in daily activity.

To provide a case study, Ashalatha et al. (2013), analysed a study on modal preference in India using a multinomial logit model. In their analysis, it was found that the preference to use private vehicles increases with age, while the preference of using motorcycles tends to fall with age compared to the use of public transport. In addition, the results showed that an increase in time and cost also interferes with the modal choices of travelers, causing them to switch from public transport to private vehicles.

In another study conducted in Nanjing, China, to study modal choices regarding daily commuting and activities such as shopping and leisure, Feng et al. (2014), also using a multinomial logit regression model, found that "classical socio-demographic variables will have a larger explanatory power in models of mode choice in China than in other Western countries because of limited financial opportunities, a limited transport supply, and a stronger collectivist culture". This finding could be easily applied in various regions of the world, where there is a vast cultural and socio-economic difference between residents.

Lanzini and Khan (2017) analyzed the results of 58 primary studies on relevant psychological and behavioral factors to synthesize the evidence on what can determine an individual's modal choice, regarding behavioral intentions as well as actual travel behaviors. The results suggest that, in addition to travel intentions, habits represent the most relevant predictor, and environmental variables, on the other hand, play a relevant role in the formation of behavioral intentions, while their effect on actual behaviors showed little significance. For example, changes of city, job or any other event that can alter travel routines, can bring back the individual's conscious reasoning, causing their cognitive gears to reorient themselves back to deliberation, thus weakening the habit of choosing a private vehicle for these trips (Klöckner, 2004).

According to (Vovsha and Freedman, 2008) trip-based models have some limitations due to the consideration that each trip is independent of all others, thus making it difficult to analyze the chain of trips. Conventional models of modal choice define the basic unit of travel as the movement that a person makes from an origin to a destination (O-D) with a certain purpose through a certain mode of transport This concept of travel ends up being limited when trips are interdependent, and decisions made for one mode affect the others. An example of this is when the private car is passed over on the home-work trip, not being available on the return trip, a situation that in traditional modeling can be difficult to assess (Gonçalves, 2012).

Bowman and Ben-Akiva (1999), said that An important aspect of this structure is that the O-D concept was transformed from the characteristics of more intra-household incorporation, since 'joint' activities (those with the participation of two or more family members) and 'allocated' activities (those that derive from family needs rather than individual adaptations) are modeled at the whole family level, making it difficult to preserve a single model or a fully integrated pattern concept containing all trips throughout the day (Bowman and Ben-Akiva, 2001).

(Davidson et al., 2007) noted that more recent models have evolved to consider the concept of trip or trip chain to the basic unit of travel, where the modal choice is made considering the entire daily journey, so that the mode used on the return is the same as the mode used for the outward journey. Defining a journey as a closed chain of journeys, that is, where the origin of the first journey is the destination of the last journey, can be seen as an important step in the development of new operational models. This definition allows the journey to have only one focal point, the origin, with one or more potential destinations. In most journeys, the points of origin are the residence or the place of work and the fact that it is considered closed allows it to have properties that make it an extremely convenient modeling unit (Vovsha and Freedman, 2008)

A tour-based structure is one that is considered to be a complete trip structure as opposed to a model based only on an elementary trip. This structure preserves consistency between trips included in the same tour (Davidson, 2008).

So, to analyze socioeconomic variables and travel in a totalitarian way, the IMob survey by INE (2018) will allow us to generate a model in a tour-based concept because it contains data obtained through household interviews and will serve to estimate a model based on various activities. In this thesis we chose to use the Generation-Attraction (G-A) method instead of Origin-Destination (O-D) pairs.

Further, studies in other contexts likely will provide different results, moreover, it is evident that the factors that influence modal choice vary with the spatial context (country, region, city) but also temporal context. Hence, the importance of periodically updating mobility surveys and consequently studies on factors that influence the modal choice, as is the case of the present study.

2.2 Modal choice models

The factors that affect the choice of a transport mode are of interest to transport and mobility policy researchers. For this, a substantial qualitative analysis needs to rely on certain characteristics that will be presented in Chapter 3 of this document.

The scientific research community has increasingly produced relevant content on transport models based on citizens' daily activity. Transport data usually contain observations on trips made by a sample of individuals, using household surveys, onboard surveys, cordon-counts, etc.

Ferreira (2010) points out that "a model is a mathematical representation of reality", and it is common to think that a mathematical model is composed of several extremely complex formulas. In the field of urban transport modeling, this association may be correct since it is difficult to understand and replicate the various interactions that are inherent in transport systems.

Individual utility maximization theory provides a complete model of individual choice (McFadden, 1996). However, within the unstable economic framework and the vast motivation possibility of travel, there will always be unobserved characteristics, such as perceptions and attributes not measured during a survey, and which also vary according to the geographical location of the population studied.

For the citizen, before he/she travels, a choice of transport mode is performed against a set of all available discrete alternatives known to him/her. This means, therefore, that different elements can be considered, such as contextual factors, choice alternatives, and subjective values. The cost of private vehicles, income, age, and distance from the origin of departure to the destination is the most widely studied factors and considered in research studies (Wang et. al, 2018; Paulssen et. al, 2014). Therefore, statistical models should represent the relationship between attributes and behaviors.

De Witte et al. (2013), proposed three ways in which we can understand the complexity of commuting mode choice: rationalist approach, socio-geographical approach, and socio-psychological approach.

The rationalist approach is based on the random utility theory, applies the individual's behavioral theory to the demand, and portrays those travelers make comparisons between all the alternatives they have to decide how they will move, taking into consideration factors such as travel cost and time. According to this theory, citizens rank alternatives in order of preference and choose the one that best suits their needs, considering their constraints (e.g. income).

The socio-geographical approach assumes that the choice of travel mode is based on their social activities, such as family activities, and introduces a spatial dimension in the choice process, giving an optic that the act of traveling is not the end, but the means to do something (e.g. going to work, school, shopping, etc.).

The socio-psychological approach addresses subjective factors in the choice of travel modes, such as citizens' intentions and habits. According to Scheiner and Holz-Rau (2007), lifestyle perception, leisure aspects, consumption behavior, and social networks play an important role in the choice of travel mode.

However, of the many theories of transport modeling addressed and studied now a days, Ortuzar and Willumsen (2011) stated that we are dependent on the theoretical reference that it is adopt in each study to find a mathematical model with a high-reliability rate, stating that "the theoretical framework will also lend some credence to the ability of the model to forecast future behavior". For example, according to Meyer and Miller (2001) to have a transport model, the studied area must be well delimited, and we can find with a higher reliability rate in small regions, since, with the increase of the area analysed it becomes difficult to consider more and more options of origins and destinations.

In this sense, it is important to understand what influence theory and practical sides can have on each other and balance it. Ortuzar and Willumsen (2011), refer that it is possible to notice that practical models used in transport analysis have traditionally had a guiding influence on subsequent theoretical studies. They also refer that there are two major classical approaches to the development of transport models:

- Inductive, where the study is initiated with data to suggest general concepts of the outcome;
- Deductive, where you can build a model and test the predictions against observations.

It is the inductive formulation of probabilities through a rationalistic approach in the multiple-choice selection that will provide a practical database for the empirical analysis carried out in this study.

2.3. Multinomial Logit Model

Regression models constitute one of the most important tools in the statistical analysis of data when one intends to model relationships between variables. The main objective of these models is to explore the relationship between one or more explanatory (or independent) variables and a response (or dependent) variable.

According to Kuhnimhof et al (2016), the multinomial logit model is a widely used predictive method in travel behavior research. The multinomial logit model helps to understand that the random terms of the utility function are independent of each other and follow the same double exponential distribution. In the practical analysis, the traveler chooses the most efficient transport alternative under N given alternatives, and the choice is influenced by the characteristics of the traveler and the characteristics of the mode of travel. If these factors are known, then traveler choice behavior can also be predictable. As will be shown in Chapter 3, the prediction accuracy of the multinomial logit model is 61.6% and represents a fit of the model as good enough.

As presented by Ferreira (2010), qualitative response models, i.e., that are represented by categories, may correspond to a binary response (e.g., yes - 0 and no - 1), an ordered response (e.g., 0 - little, 1 - medium and 2 - a lot, not necessarily having a relationship between 0 and 2) and categorized response (e.g., 0 - motorized private transport, 1 - soft modes, 2 - public transport, not following any order of specification).

In this work we will use the qualitative response model with multiple unordered (multinomial) categories, represented by:

$$S_{jn} = \beta_j X_n + \varepsilon_{jn} \tag{1}$$

Where β_j are parameters to be estimated, X_n are the independent variables in categories and ε_{jn} is the unobserved random error term, where j is the alternative choice and n are representing the commuters.

Logistic regression is a statistical technique that aims to model, from a set of observations between the "logistic" relationship of a dichotomous response variable and a series of numerical (continuous, discrete) and/or categorical explanatory variables. Unlike the continuous distribution, the logistic distribution decreases when the probability of "success" tends towards zero and increases when the probability of success tends towards a normal distribution. This model has been widely used due to the easy interpretation of its parameters. According to Greene (2008), this model is represented as follows:

$$P(Y = 1|x) = \frac{e^{x\beta}}{1 + e^{x\beta}} = \Lambda(\mathrm{x}\beta)$$

(2)

Where $\Lambda(\mathrm{x}\beta)$ identifies the cumulative logistic distribution function.

In the multinomial logit model, a Gumbel distribution (also known as extreme value distribution) can be considered, as represented below:

$$Pn(l) = \sum_{n=1}^{N} \sum_{j=1}^{J} y_{nj} \, ln \frac{\exp{(\beta_j X_n)}}{\sum_j exp(\beta_J X_n)}$$

(3)

where y_{nj} are dummy variables indicating whether the *j*th alternative is chosen by the individual "*n*", y_{nj} = 1 if category *j* is chosen by citizen *n* and y_{nj} = 0 otherwise.

In this thesis, the multinomial logit model was estimated using IBM SPSS Statistic 27 software.

Chapter 3

Data processing and analysis of results

3.1 - Mobility survey

From October to December of 2017, the last Mobility Survey (IMob) available for the AMP was carried out by the Instituto Nacional de Estatística in Portugal (INE). The intention of the data measurement is the characterization of the trips of the residents in this region, composed of 17 municipalities and targeting Individuals aged between 6 and 84 years old. The municipalities involved are:

- Arouca;
- Espinho;
- Gondomar;
- Maia;
- Matosinhos;
- Oliveira de Azeméis;
- Paredes;
- Porto;
- Póvoa do Varzim;
- Santa Maria da Feira;
- Santo Tirso;
- São João da Madeira;
- Trofa;
- Vale de Cambra;
- Valongo;
- Vila do Conde;
- Vila Nova de Gaia;

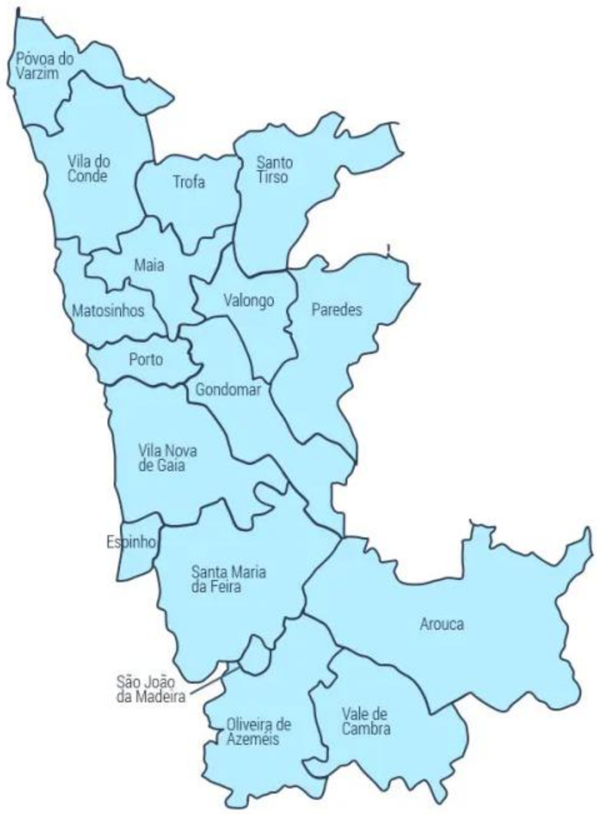

Figure 4 - Municipalities of the Porto Metropolitan Area (Portal AMP, 2018)

According to the data provided out by INE (2016), these municipalities together occupy an area of approximately 2,041 km2 and in 2016 there were already about 1.7 million people residing in the AMP, which corresponds to approximately 17% of the total resident population in Portugal.

Table 1 - Description of the municipalities of the Porto Metropolitan Area (INE, 2016)

	Area	Population density	Total resident population	Resident population 6-84 years	Resident population 6-84 years (Men)	Resident population 6-84 years (Women)
	km^2	Hab./km^2	N.º			
Portugal	**92 225,60**	**111,8**	**10 309 573**	**9 499 839**	**4 521 972**	**4 977 867**
AMP	**2 041,31**	**842,1**	**1 719 021**	**1 596 802**	**755 563**	**841 239**
Arouca	329,11	64,4	21 211	19 533	9 339	10 194
Espinho	21,06	1 403,6	29 560	27 516	12 993	14 523
Gondomar	131,86	1 257,0	165 743	154 969	73 798	81 171
Maia	82,99	1 638,9	136 011	126 503	59 945	66 558
Matosinhos	62,42	2 777,0	173 339	161 086	75 820	85 266
Oliveira de Azeméis	161,10	412,8	66 496	62 302	29 839	32 463
Paredes	156,76	550,3	86 263	80 773	39 275	41 498
Porto	41,42	5 169,5	214 119	193 854	88 273	105 581
Póvoa de Varzim	82,21	758,4	62 344	58 160	27 451	30 709
Santa Maria da Feira	215,88	643,3	138 867	129 684	62 257	67 427
Santo Tirso	136,60	505,0	68 983	64 703	30 779	33 924
São João da Madeira	7,94	2 702,8	21 460	20 025	9 414	10 611
Trofa	72,02	530,5	38 210	35 757	17 061	18 696
Vale de Cambra	147,33	147,1	21 676	20 110	9 670	10 440
Valongo	75,12	1 270,1	95 411	88 941	42 159	46 782
Vila do Conde	149,03	532,3	79 327	73 581	35 341	38 240
Vila Nova de Gaia	168,46	1 780,8	300 001	279 305	132 149	147 156

According to INE (2018), the IMob survey aims to meet the needs of the Portuguese Metropolitan Area in terms of statistical information for transport and mobility, as well as the Statistical office of the European Union (Eurostat), so it has followed the Eurostat guidelines, seeking to provide harmonized and comparable statistics throughout the European territory.

Its main objective was to characterize the journeys made by the resident population, in this case at AMP, to understand the profile of the journeys made by the respondents, the users' opinion on individual or collective transport, and the main motivations that lead to their transport choices.

This present study analyse which variables influence citizens' choice of transport mode in their daily commuting. The data acquired by the survey provides detailed information on travel behaviour, including some socio-demographic variables. The dataset consists of the following variables:

- Residence information;
- Family size and household information (age, gender, education, occupation);
- Vehicles available;
- Characterization of household members' travel mobility (travel time, distance, reason for travel, destination, mode of transport);
- Transport costs (fuel cost, parking cost, transport cost, and tolls);
- Income information and respondents' opinions regarding transport.

QUESTIONÁRIO

LEGENDA:

A preto - questões colocadas ao respondente

A cinzento - variáveis calculadas na aplicação (não questionadas)

DF - SELEÇÃO DIA DE REFERÊNCIA

Inquérito à Mobilidade nas Áreas Metropolitanas do Porto e de Lisboa

Com este inquérito, pretende-se recolher informação sobre todos os residentes deste alojamento.

Contudo, sugere-se que seja preenchido por uma só pessoa - o **respondente**.

O respondente (18 ou mais anos) deve estar informado sobre as deslocações das demais pessoas, sendo desejável que as possa consultar, para uma resposta com a melhor qualidade possível.

O inquérito compõe-se pelas seguintes partes:

1 - Alojamento
2 - Identificação das pessoas residentes
3 - Veículos disponíveis
4 --Caracterização e condições de mobilidade das pessoas residentes
5 - Deslocações diárias (destinos, modos de transporte, motivos)
6 - Despesas relacionadas com transporte e rendimento (resposta em escalões)
7 - Opiniões

Para começar a responder pressione a seta para a direita, no canto inferior direito do ecrã.

Para consultar a informação registada pode utilizar o botão "JÁ RESPONDIDO" (do lado esquerdo do ecrã)

Figure 5 - Characterization of the survey variables (INE, 2018)

3.2 - AMP mobility patterns

According to the INE (2018), the mobile population is defined as "the set of people who have made at least one trip (starting on the reference day)". Based on this definition it is estimated that the number of trips per day in the AMP has amounted to the number of 3.4 million with a mobile population of 78.9% of its total resident population. If we consider only working days this value increases to 82.9% and on non-working days such as Saturdays, Sundays, and holidays, it falls to 71.2%.

The municipalities that origin the largest number of trips (by municipality of residence) were respectively Vila Nova de Gaia with around 635 thousand (17.6%) trips, Porto with 419 thousand (13.1%), and Matosinhos with 359 thousand (10.6%), contributing with a total of approximately 41% of the trips in all AMP (INE, 2018) per day.

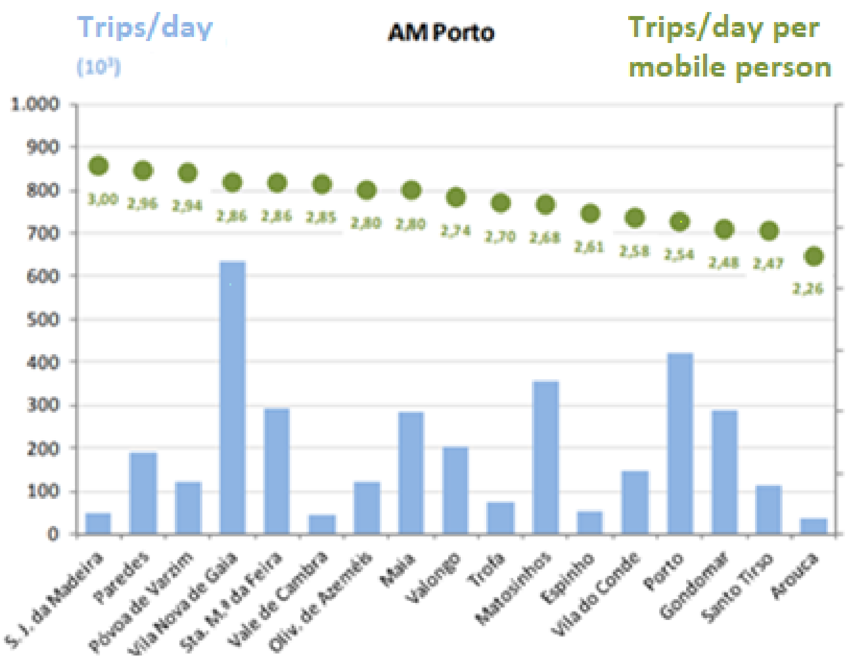

Figure 6 - Total number of trips/day and per mobile person, by the municipality of residence (INE, 2018)

To conduct the IMob (2017) survey, in a first phase of the survey, respondent information was collected through a self-completion form via the Internet (computer-assisted Web-based Interviewing - CAWI), and in a second phase, information was collected by face-to-face interviews (Computer Assisted Personal Interviewing - CAPI) at a subset of addresses selected from those residents who did not respond in the first phase.

From the data collected in the Metropolitan Area of Porto, the total sample size obtained a total of 18169 valid answers, contemplating a total of 40393 individuals and 80314 trips, thus making up a sample of approximately 2.4% in relation to the total number of trips presented by INE for the same year.

In the year 2000, a similar survey was also carried out in the AMP and AML, although at the time both metropolitan regions had a different demographic density than today, the sample size collected was similar to that of the present study. Gonçalves (2012), shows in his work that when working with the 2000 survey data, he acquired a final database with 95,426 trips made in AMP, which is an approximate value to the final database of this present study for AMP.

To analyze the changes that occurred in the modal choices of individuals from AMP over the years, an indirect comparison will be made of the results obtained by Gonçalves (2012), with those obtained in this study.

3.3 - Data processing

After a first analysis of the data raised by the survey to choose the variables to be used in the model, it was noticed that part of the variables included a very high number of categories, so it was necessary to make some changes to ensure representativeness in the results.

Of the original variables, the table below identifies how these were or were not considered in the present study (Used, Not used, and Changed):

Table 2 - Table of collected data

Variable Designation	Status	Variable Designation	Status
Accommodation address	Not used	Reason for the non-existence of trip	Not used
Cell phone	Not used	Place of origin of the first trip	Not used
Email	Not used	Place of travel destination	Not used
Gender	Used	Return home date	Not used
Age	Changed	Purpose of trip	Changed
Quantity of vehicles	Not used	Travel start time	Not used
Vehicle registration year	Not used	Transport mode	Changed
Vehicle Ownership/ Availability Type	Used	Which car is used for travel	Not used
Type of fuel	Not used	Passengers in a private car	Not used
No. of vehicle seats	Not used	What is the river crossing	Not used
Parking at home	Not used	No. of consecutive buses from the same operator	Not used
Relationship with the respondent	Not used	Public transport operator	Not used
Level of schooling	Changed	Entrance station	Not used
Condition towards employment	Changed	Exit station	Not used
Existence of physical limitation	Changed	Used transport title	Not used
Possession of a driving license	Used	Location data	Changed

Driving frequency	Changed	Waiting time	Not used
Possession of a transport pass	Used	Fuel expenses	Used*
Transport pass type	Not used	Parking expenses	Used*
Type of workplace	Used	Toll expenses	Used*
Location of the workplace	Not used	Public transport expenses	Used*
Parking at work	Not used	Monthly income	Changed
Location of the study site	Not used	Evaluation of the use of individual transport	Not used
Parking at the study site	Not used	Evaluation of the use of public transport	Not used
Existence of trip	Used	Public transport evaluation	Not used

*Used but removed from the final model

The variables identified in the table as altered were restructured, essentially by aggregating the categories, grouping those that could represent the same and/or did not generate relevant information for the study.

According to Ortuzar and Willumsen (2011), although there is no effective way to choose the number or type of categories needed to increase the reliability of a calculation, it is necessary that a detailed analysis of the categorization of each data presented, to find variables that define similar categories or categories that do not provide a substantial explanation of the variation in the data even variables that duplicate the explanation provided by other better variables.

Therefore, individual analysis of each unused or modified data category follows:

- Personal details:
  - Accommodation address;
  - Email;
  - Cell phone.

These variables were not available in the database used in this study due to the data protection of the respondents, thus seeking the privacy of the people, concerning the processing of personal data and the free movement of such data.

- Vehicle and passenger data:
  - Vehicle registration year;
  - Type of fuel;
  - Quantity of vehicles;
  - No. of vehicle seats;
  - Which car is used for travel;
  - Relationship with the respondent;
  - Passengers in a private car.

Regarding these variables, in the original survey, several conditions were questioned about the use of private vehicles. Also, the differentiation of drivers and passengers was made. Although this detailing may be relevant for other research, such variables will not differentiate or have no impact on the objective of this research, which is to know the type of transport used according to the specific conditions of the trip, so the variable used was "transport mode".

- Public transport data:
  - Transport pass type;
  - Used transport title;
  - Public transport operator;
  - No. of consecutive buses from the same operator.

These variables were discarded because it was not interesting for the study to know about the titles/types of transport pass used, neither about public transport operator, as they did not add benefits for this study. The relevant data to know is only if the respondent owned or not the transport pass, so it was selected for use only the field "Possession of a transport pass".

- Location data:
  - Entrance station;
  - Exit station;
  - Location of the workplace;
  - Location of the study site;
  - Return home date;
  - Place of origin of the first trip;
  - Place of travel destination.

The reason for not using these variables was due to the difficulty in confirming the correct place of origin and destination of the trips because they are coded in the database. On other hand, this study does not intend to know the locations, nor the dates mentioned. The study only considered the distance of the trip in km made by the individual on the day of the survey, through the variable “distance (Km)”.

- Temporal data:
  - Travel starts time;
  - Waiting time.

The variables were eliminated because what matters for the model is the total travel time, thus making no distinction between the time at the start of the journey and the expected time when making the journey, so the variable used was "Distance (Km)”.

- Non-existence data of the trip:
  - Reason for the non-existence of trip.

This variable was excluded because the final database only showed the individual trips that occurred; thus, it was not possible to know the causes of no trips.

- Data not relevant for AMP
    - River-crossing.

This variable does not apply to the AMP as it does not directly affect modal choices.

- Transport quality rating data:
    - Evaluation of the use of individual transport;
    - Evaluation of the use of public transport;
    - Public transport evaluation.

The fields referring to transport evaluations were not considered for modeling, as these are respondents' opinions regarding the transport mode in the AMP and as such, do not characterize their journey.

- Costs data:
    - Fuel expenses;
    - Parking expenses;
    - Toll expenses;
    - Public transport expenses;
    - Parking at home;
    - Parking at the study site;
    - Parking at work.

These variables were initially considered in the model, however, as the estimated coefficients showed low statistical significance they were removed from the final model, because despite being important variables, it would take a longer time to properly prepare these variables for use in the model, which was not possible in this work due to the short time for preparing the variables.

For better optimization of the results to be obtained by the chosen model, it was necessary to change the categories of some of the variables to obtain greater representativeness of the variables with interest to the objective of this study. With this, the data collected through the survey information were classified and coded into different groups that had similar characteristics.

After the new categorization of the variables, those chosen will be used together so that the final model does not lose explanatory value.

With this, the fields that suffered alterations were:

- Age:
  - < 25 years old;
  - 25 to 44 years old;
  - 45 to 64 years old;
  - ≥ 65 years old.

The variable "Age" was only changed in the classification below 25 years old, grouping all below this age since it was considered that they represent a group with less autonomy and power of decision in the modal choice. In the original research, it was possible to find the categories 0 to 14 years and 15 to 24 years. The remaining categories were maintained, as shown below:

- Monthly Income:

As the monthly income characterizes a potential economic level of the household, it was necessary to create a new variable called "per capita income", so that the average monthly individual income of the respondents could be better evaluated.

To create this new variable, the lowest and the highest income value of each group in the family grouping were selected, then this value was divided by the total number of people in that household.

Table 3 - Original categories of the variable "Monthly income"

BEFORE THE CHANGE

MONTHLY INCOME	
1	Less than 430 euros
2	From 430 to less than 600 euros
3	From 600 to less than 1000 euros
4	From 1000 to less than 1500 euros
5	From 1500 to less than 2600 euros
6	From 2600 to less than 3600 euros
7	From 3600 to less than 5700 euros
8	From 5700 to less than 7000 euros
9	7000 euros or more
10	Did not respond

For example, let's reclassify household income F1_1000 to income per capita, as per the original data below collected from the survey data table:

Table 4 - Example of calculation of the new variable "Per capita income"

BEFORE THE CHANGE				CALCULATIONS	
ID_accommodation	N_Individuals	Lower value of family income	Higher value of family income	Lower value per person	Higher value per person
F1_1000	4	1500	2600	375	650
F1_1002	2	600	1000	300	500
F1_1003	3	430	600	143.33	200
F1_1007	4	1000	1500	250	375
F1_1008	4	3600	5700	900	1425
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...

The household F1_1000 has a monthly income "From 1500 to less than 2600 euros" and is composed of 4 individuals. Thus, we apply the following formula:

$$PER\ CAPTA\ INCOME = \frac{Lower\ Value}{N\_Individual}\ and\ \frac{Higher\ Value}{N\_Individual} \tag{4}$$

Resulting in:

$$PER\ CAPTA\ INCOME\ F1_{1000} = LowerValue\,\frac{1500€}{4}\ and\ Higher\ Value\,\frac{2600€}{4}$$

$$PER\ CAPTA\ INCOME\ F1_{1000} = LowerValue\ 375€\ \ and\ Higher\ Value\ 650€$$

After performing these calculations for all households, the per capita income was considered, thus creating 5 new categories reported below:

- Up to 650 Euros per month;
- From 651 to 1000 Euros per month;
- From 1001 to 1500 Euros per month;
- From 1501 to 2000 Euros per month;
- Above 200 1Euros per month.

In the case of the example of household F1_1000, the per capita income obtained was corresponding to Category 1 (Up to 650 Euros).

- Transport mode:

As already explained at the beginning of this chapter, to make better use of the categorization, the 15 possible options of transport mode in the IMob survey were reduced to 5 categories, as shown in table 5:

Table 5 - Original categories of the variable "Transport mode"

BEFORE THE CHANGE

TRANSPORT MODE	
1	Passenger car - as a driver
2	Passenger car - as a passenger
3	Bus and coach - TE
4	Bus and coach - TP
5	Motorbike and Moped
6	Urban rail
7	Regular train
8	Walking
9	Cycling
10	Taxi as passenger
11	Van/Lorry/Tractor/Camper
12	Waterways
13	Aviation
14	Other
15	Unknow

Some of these means of travel were eliminated for having transport modes that, even when grouped, were used by less than 1.5% of the respondents, thus not presenting a significant sample for the analysis of modal choices in the modeling, totaling a value of 3813 trips eliminated. With this the data of the variable "Transport mode" eliminated were:

- Taxi as a passenger;
- Van/lorry/tractor/camper;
- Waterways;
- Aviation;
- Others;
- Unknow.

With the cut of these data, a new analysis of the database was carried out, thus finding a possibility of grouping the remaining categories, which had similar characteristics among themselves, thus further reducing the number of existing categories for a better analysis of the data.

The grouping of the categories that presented similar characteristics was done as follows:

- Passenger car - as a driver;
- Passenger car - as a passenger.

These categories were grouped becoming category 1 (Private vehicle), thus making no distinction between the passenger and the driver, but only the type of transport used for the trip.

- Bus and coach - TE (Electric public transport);
- Bus and coach - TP (ordinary public transport);
- Urban rail;
- Regular train.

These categories have been aggregated into category 2 (Public Transport) as they are non-individual modes of transport, and the use of a pass or ticket is also required to travel through them.

- Motorbike and moped.

These categories became category 3 (Motorcycle), which although it was not grouped with any other category, had its name simplified.

After the organization of the new categories, the variable suffered a reduction, thus remaining with only 5 options to choose from, as shown in table 6:

Table 6- New categorization of the variable "Transport mode"

AFTER CHANGE

TRANSPORT MODE	
1	Private vehicle
2	Public transport
3	Motorcycle
4	Walking
5	Bicycle

- Purpose of trip:

At first, this variable had 17 options to choose from, thus having some options of choices that had very similar characteristics, becoming the variable with the largest number of categories to be reduced:

Table 7 - Original categories of the variable "Purpose of trip"

BEFORE THE CHANGE

PURPOSE OF TRIP	
1	Go to work
2	Dealing with professional matters
3	Go to school or school activities
4	Take/pick up/accompany family or friends (children to school, etc.)
5	Visiting family or friends
6	Shopping (supermarket, grocery, utilities, etc.)
7	Go to consultation, treatments, medical exams, and the like
8	Dealing with personal matters (going to the bank, laundry, hairdressing, taking or picking up personal things, etc.)
9	Go to a restaurant, café, bar, disco, etc.
10	Practice outdoor activities (sports or leisure) or in a gym or pavilion
11	Attending sporting or cultural events (cinema, theatre, concert, football, etc.)
12	Carry out activities in groups or in a collective context (in associations, rallies, churches, volunteering, ...)
13	Take a pedestrian route (start and end in the same place), jogging, walking the dog, etc. (with at least 200 meters)
14	Other leisure activities, entertainment, or tourism
15	Return home
16	Other activity
17	Don't know

The aggregation of the variable categories was carried out as follows:

- Dealing with professional matters (option 2).

This purpose was added to option 1 (Go to work), as both are work-related trips.

- Take/pick up/accompany family or friends (children to school, etc.) (option 4).

It was associated with option 3 (Go to school or school activities), becoming option 2, as both were travel-related to school matters.

- Visiting family or friends (option 5);
- Shopping (supermarket, grocery, utilities, etc.) (option 6);
- Go to consultation, treatments, medical exams, and the like (option 7);
- Dealing with personal matters (going to the bank, laundry, hairdressing, taking or picking up personal things, etc.) (option 8).

These purposes were all grouped together because they were related to dealing with personal matters, thus becoming option 3 (Dealing with personal matters).

- Go to a restaurant, café, bar, disco, etc(option 9);
- Practice outdoor activities (sports or leisure) or in a gym or pavilion (option 10);
- Attending sporting or cultural events (cinema, theatre, concert, football, etc.) (option 11);

- Carry out activities in groups or in a collective context (in associations, rallies, churches, volunteering,) (option 12);
- Take a pedestrian route (start and end in the same place), jogging, walking the dog, etc. (with at least 200 meters) (option 13).

As they were all related to leisure activities, they were aggregated to become option 4 (Leisure activities).

- Other leisure activities, entertainment, or tourism (option 14);
- Other activity (option 16);
- Don't know (option 17).

Finally, the last reason to be added became option 5 (Other activities).

After the separation of the grouped categories, the variable suffered a reduction in its choice options, changing to 6 purposes of the trip, as shown in table 8:

Table 8 - Categories considered for the variable "Purpose of trip"

AFTER CHANGE

	PURPOSE OF TRIP
1	Go to work
2	Go to school or school activities
3	Dealing with personal matters
4	Leisure Activities
5	Other activity
6	Return home

In order to analyze the socio-economic variables presented, as said before, it was chosen to use the Generation-Attraction (G-A) method instead of Origin-Destination (O-D), where the destination zone (house) does not correspond to the attraction but rather the generation zone, with the attraction zone becoming the origin of the journey (Davidson, 2008).

- Level of schooling:

Because the categories "Not applicable", "Prefer not to answer" and "Do not know" had low number of observations and do not add relevant information for the model, they were grouped into a single category named "Prefer not to answer". Thus, the variables were altered as follows:

Table 9 - Original categories of the variable "Level of schooling"

BEFORE THE CHANGE

LEVEL OF SCHOOLING	
1	Not applicable
2	None or full 1st or 2nd or 3rd year
3	Basic Education (1st cycle, 2nd cycle or 3rd complete cycle)
4	Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)
5	Higher Education (Bachelor's Degree, Master's, Doctorate, Professional Higher Technical Course)
6	Prefer not to answer
7	Do not know

Table 10 - Categories considered for the variable "Level of schooling"

AFTER CHANGE

LEVEL OF SCHOOLING	
1	None or full 1st or 2nd or 3rd year
2	Basic Education (1st cycle, 2nd cycle or 3rd complete cycle)
3	Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)
4	Higher Education (Bachelor's Degree, Master's, Doctorate, Professional Higher Technical Course)
5	Prefer not to answer

- Condition towards employment:

In the Condition Towards Employment, the same categories' changes considered as the previous variable were applied.

Table 11 - Original categories of the variable "Condition towards employment"

BEFORE THE CHANGE

CONDITION TOWARDS EMPLOYMENT	
1	Not applicable
2	Employee
3	Unemployed
4	Student, Retired, Mainly occupied with household
5	Prefer not to answer
6	Do not know

Table 12 - Categories considered for the variable "Condition towards employment"

AFTER CHANGE

CONDITION TOWARDS EMPLOYMENT	
1	Employee
2	Unemployed
3	Student, Retired, Mainly occupied with household
4	Prefer not to answer

- Existence of physical limitation:

Similarly, to the previous ones the only categories aggregation applied in this variable was for the "Not applicable" and "Prefer not to respond" categories.

Table 13 - Original categories for the variable "Existence of physical limitation"

BEFORE THE CHANGE

EXISTENCE OF PHYSICAL LIMITATION	
1	Not applicable
2	Yes
3	No
4	Prefer not to respond

Table 14 - Categories considered for the variable "Existence of physical limitation"

AFTER CHANGE

EXISTENCE OF PHYSICAL LIMITATION	
1	Yes
2	No
3	Prefer not to respond

- Driving frequency:

Initially, the variable "Driving frequency" had 7 categories, as shown in the table 15:

Table 15 - Original categories of the variable "Driving frequency"

BEFORE THE CHANGE

DRIVE FREQUENCY	
1	Not applicable
2	Daily, or almost daily
3	1 to 3 days a week
4	From 1 to 3 days per month
5	Less than 1 day per month
6	Never, or hardly ever
7	Do not know

To simplify and reduce their choice the aggregation of the categories occurred as follows:

- Daily, or almost daily;
- 1 to 3 days a week.

Their data were aggregated according to the regularity with which individuals drive their vehicles, thus transformed into category 1 (Drives frequently).

- From 1 to 3 days per month;
- Never, or hardly ever.

The category 4 and 6 of table 15 were associated and had its name changed to category 2 (Drives sporadically) in table 16.

- Not applicable;
- Less than 1 day per month;
- Do not know;

Finally, these categories were grouped together either because they were not informative or because they had a very small number of observations thus making the category 3 (Unkown) in table 16.

After the changes made to the categories of "Drive frequency" variable, their choice of options changed from 7 to 3 as shown in table 16:

Table 16 - Categories considered for the variable "Driving frequency"

AFTER CHANGE

DRIVE FREQUENCY	
1	Drives frequently
2	Drives sporadically
3	Unknown

Posterior to separating the variables, 821 trips were eliminated (1%) of the 80314 trips identified in the original database, following the assumption of Gonçalves (2012), because such trips had a travel time of more than 120 minutes, not fitting into trips within the AMP territory.

After excluding these trips, the new database has a total of 75680 trips to be considered in this analysis.

3.3.1 - Classification of the variables

With the updated database, the variables were classified into 3 different groups:

- Individuals:
  - Gender;
  - Age;
  - Existence of physical limitation.

- Socioeconomics:
  - Level of schooling;
  - Condition towards employment;

- Type of workplace;
- Per capita income;
- Vehicle Ownership/Availability Type;
- Possession of a driving license;
- Driving frequency;
- Possession of a transport pass.

- Trips:

  - Purpose of the trip;
  - Transport mode;
  - Existence of trips.

Once the separation was concluded, some of the variables from each group were chosen to be used firstly in individual analyses in relation to the transport mode, thus trying to identify which would be their individual impact on modal choices during the journeys performed in the AMP and to easily understand the specific influence of the variables on each transport mode. At the end, a final model was ran with the entire set of variables relevant to the study.

3.4 - Statistical description of the variables

- Transport mode:

Since the variable "transport mode" is our dependent variable, for the choice of the reference category of the transport mode to be chosen in the model, a graph was generated in figure 7 to evaluate which mode of transport had the highest frequency in the individuals' choices:

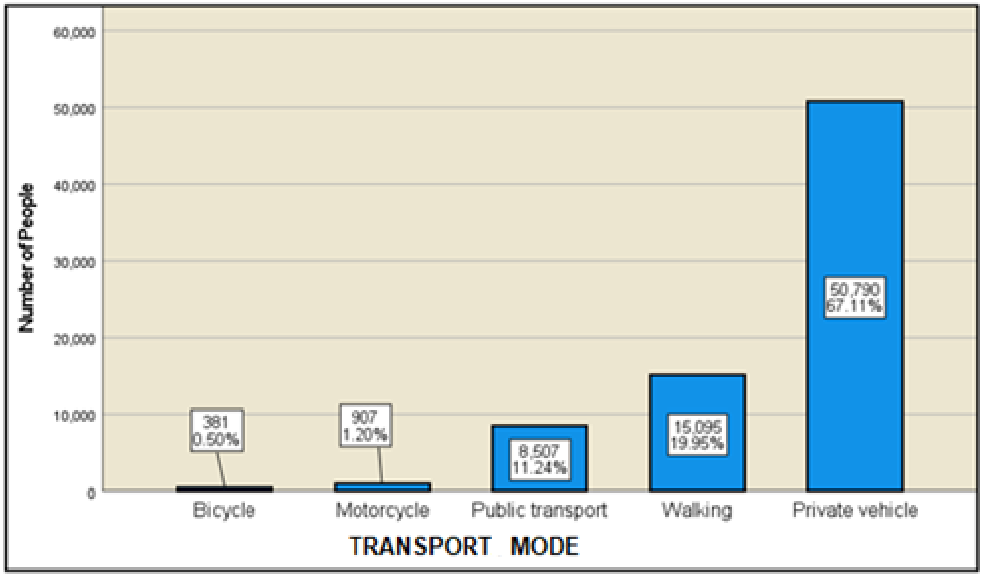

Figure 7 - Distribution of the variable "Transport mode" from SPSS software

After verifying that most of the trips were made by private vehicles, thus obtaining a total of about 67% observations of the sample, the type of transport "private vehicle" was selected as the reference variable for the analyses of the models generated.

- Gender:

The variable gender of the individual will assess the main differences between men and women in their main modal choices in their commutes.

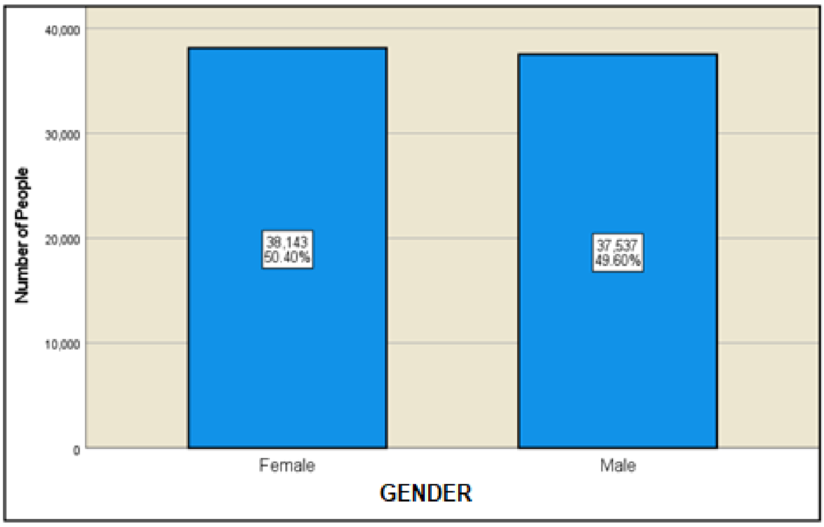

Figure 8 - Distribution of the variable "Gender" from SPSS software

Through the graph in figure 8, it was possible to see that the sample presents a balance in relation to the gender of the individuals in the sample.

Table 17 - Modal split according to the variable "Gender" from SPSS software

Observed and Predicted Frequencies			
Sex		Frequency Observed	Percentage Observed
Female	Bicycle	101	0.3%
	Motorcycle	177	0.5%
	Public transport	4944	13.0%
	Walking	8352	21.9%
	Private vehicle	24569	64.4%
Male	Bicycle	280	0.7%
	Motorcycle	730	1.9%
	Public transport	3563	9.5%
	Walking	6743	18.0%
	Private vehicle	26221	69.9%

According to the data presented by Gonçalves (2012), the male/female ratio is similar to that presented in this study. In the IMob database of the year 2000, 49.7% of respondents were male and 50.3% were female. Still, in table 18 frequencies about the modal choice preference for private vehicle and walking referring to each gender was also built, since they are the same categories presented in both studies despite the difference between the areas under study:

Table 18 - Comparison of the modal split for the variable "Gender"

	Walking				Private Vehicle			
	M (n)	M (%)	F (n)	F (%)	M (n)	M (%)	F (n)	F (%)
Gonçalves (2012)	4773	10.1	6947	14.5	29272	61.7	23854	49.7
Current data	6743	18	8352	21.9	26221	69.9	24569	64.4

- Age:

This variable allows deducing at which stage of the respondent's age group can influence their modal choices.

As per the age classification adopted by the IMob (2017) analysis, it was not possible to differentiate the school ages, nor to separate individuals under the age of 18 from other respondents, which would be interesting because this factor meets the minimum age for obtaining a driving license.

Figure 9 shows the distribution of the statistical data in the database used:

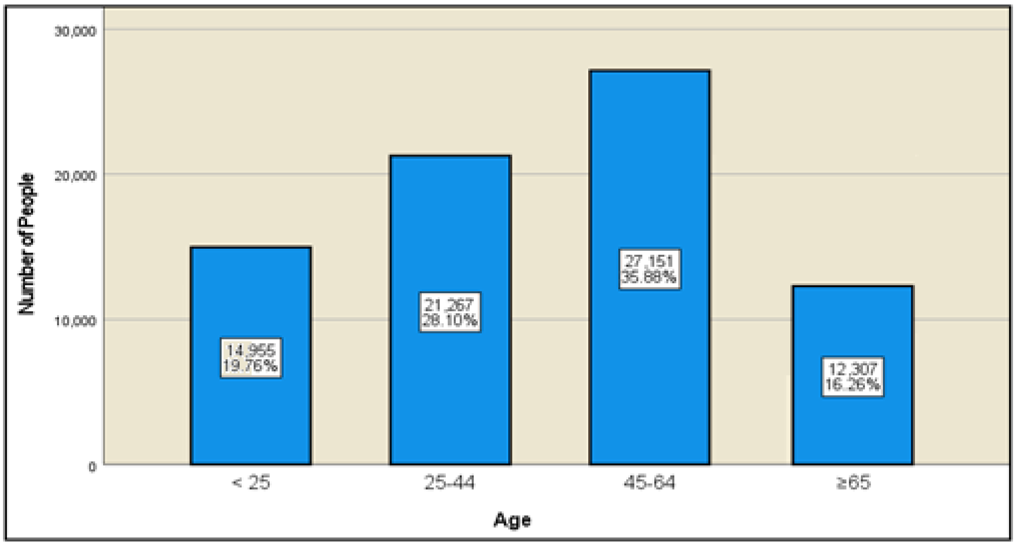

Figure 9 - Distribution of the variable "Age" from SPSS Software

It shows that the group with the highest number of observations in the sample are those aged between 45 and 64 years, with 35.9% of the observations.

Table 19 - Modal split according to the variable "Age" from SPSS software

Observed and Predicted Frequencies			
Age		Frequency Observed	Percentage Observed
< 25	Bicycle	62	0.4%
	Motorcycle	79	0.5%
	Public transport	3037	20.3%
	Walking	2670	17.9%
	Private vehicle	9107	60.9%
≥65	Bicycle	37	0.3%
	Motorcycle	111	0.9%
	Public transport	1381	11.2%
	Walking	4266	34.7%
	Private vehicle	6512	52.9%

Observed and Predicted Frequencies			
Age		Frequency Observed	Percentage Observed
25-44	Bicycle	129	0.6%
	Motorcycle	260	1.2%
	Public transport	1736	8.2%
	Walking	2816	13.2%
	Private vehicle	16326	76.8%
45-64	Bicycle	153	0.6%
	Motorcycle	457	1.7%
	Public transport	2353	8.7%
	Walking	5343	19.7%
	Private vehicle	18845	69.4%

According to the age groups defined, the categories < 25 years and ≥ 65 years were the ones that presented the lowest trends for the use of the private vehicle despite still being the highest cumulative frequency within the transport categories in these age groups. This trend could be due to the lack of driving qualification among those under 18 years old as they are in the category "< 25 years old" and in the case of the category "≥ 65 years old" is less likely to drive due to age-related loss of capabilities.

Under these conditions, it is expected that younger individuals have a greater propensity to use public transport since, in Portugal, people under 25 years old have discounts on the acquisition of public transport passes, and in the case of the walking mode when over 64 years old, due to the benefits that they may have when walking, such as the strengthening of the bones and joints, thus aiding in the disposition of the daily life. The age group that presents a greater tendency to use the private car is the one between the ages of 25 and 44 years old.

- Per capita income:

The graph below presents the distribution of statistical data in relation to the variable "Per capita income", to evaluate how these values can influence the decision about the modal choices of individuals:

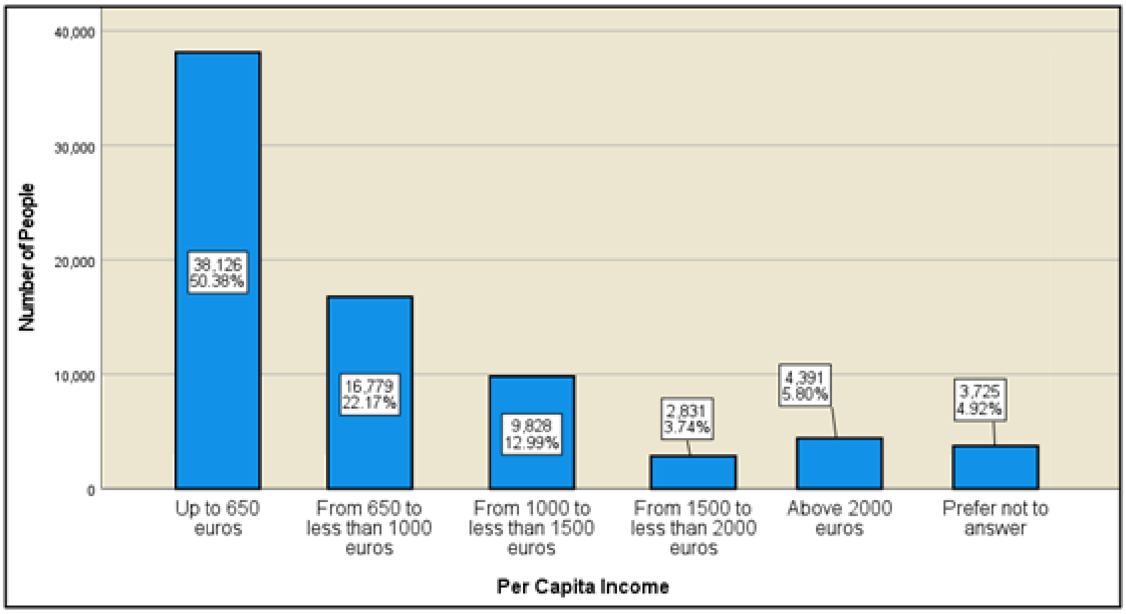

Figure 10 - Distribution of the variable "Per capita income" from SPSS software

Given the data in figure 10, it can be seen that 50.4% of the sample have a per capita income of up to €650 per month, so this will be the reference variable.

Table 20 - Modal split according to the variable "Per capita income" from SPSS software

Per Capita Income		Frequency Observed	Percentage Observed
Above 2000 euros	Bicycle	22	0.5%
	Motorcycle	51	1.2%
	Public transport	303	6.9%
	Walking	829	18.9%
	Private vehicle	3186	72.6%
From 1000 to less than 1500 euros	Bicycle	36	0.4%
	Motorcycle	121	1.2%
	Public transport	879	8.9%
	Walking	1703	17.3%
	Private vehicle	7089	72.1%
From 1500 to less than 2000 euros	Bicycle	19	0.7%
	Motorcycle	27	1.0%
	Public transport	166	5.9%
	Walking	477	16.8%
	Private vehicle	2142	75.7%

Per Capita Income		Frequency Observed	Percentage Observed
From 650 to less than 1000 euros	Bicycle	89	0.5%
	Motorcycle	157	0.9%
	Public transport	1680	10.0%
	Walking	2998	17.9%
	Private vehicle	11855	70.7%
Prefer not to answer	Bicycle	7	0.2%
	Motorcycle	41	1.1%
	Public transport	434	11.7%
	Walking	809	21.7%
	Private vehicle	2434	65.3%
Up to 650 euros	Bicycle	208	0.5%
	Motorcycle	510	1.3%
	Public transport	5045	13.2%
	Walking	8279	21.7%
	Private vehicle	24084	63.2%

After analysing the table, it can be seen that as the monthly income per capita increases, there is an increase in the use of the private vehicle in comparison with other transport modes, not only due to the probability of greater ease in acquiring and using it but also due to the dilution of the differences in the apprehension of the costs of its use in relation to other transport modes.

- Level of schooling:

With this variable we intend to characterize the potential influence of the level of education on modal choices:

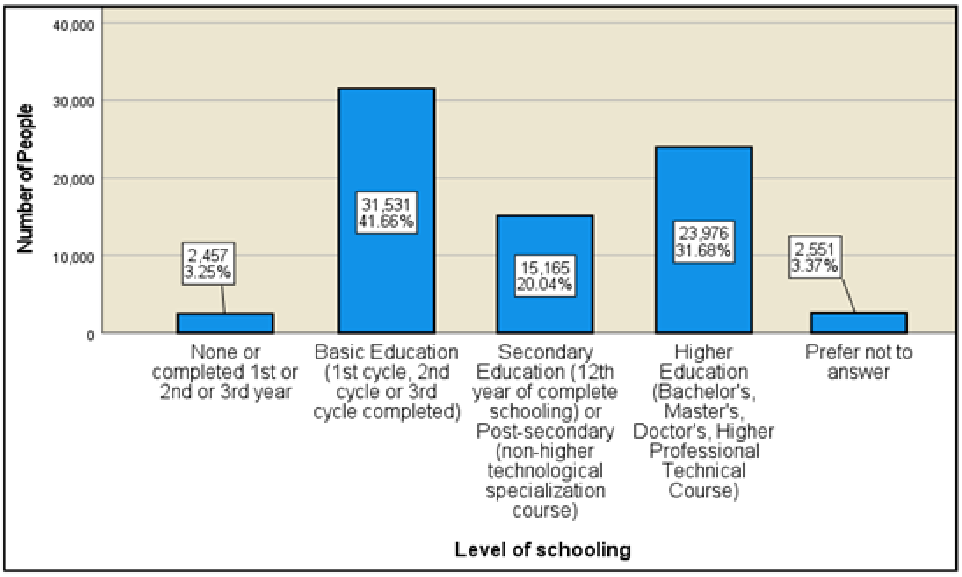

Figure 11 - Distribution of the variable "Level of schooling" from SPSS software

Considering the data in figure 11, it can be seen that most of the sample has only a basic level of education with 41.7%, this being the variable taken as reference.

Table 21 - Modal split according to the variable "Level of schooling" from SPSS software

Observed and Predicted Frequencies			
Level of schooling		Frequency Observed	Percentage Observed
Higher Education (Bachelor's, Master's, Doctor's, Higher Professional Technical Course)	Bicycle	125	0.5%
	Motorcycle	210	0.9%
	Public transport	1953	8.1%
	Walking	3704	15.4%
	Private vehicle	17984	75.0%
None or completed 1st or 2nd or 3rd year	Bicycle	11	0.4%
	Motorcycle	55	2.2%
	Public transport	332	13.5%
	Walking	906	36.9%
	Private vehicle	1153	46.9%
Prefer not to answer	Bicycle	9	0.4%
	Motorcycle	5	0.2%
	Public transport	208	8.2%
	Walking	460	18.0%
	Private vehicle	1869	73.3%

Observed and Predicted Frequencies			
Level of schooling		Frequency Observed	Percentage Observed
Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)	Bicycle	67	0.4%
	Motorcycle	134	0.9%
	Public transport	2146	14.2%
	Walking	2412	15.9%
	Private vehicle	10406	68.6%
Basic Education (1st cycle, 2nd cycle or 3rd cycle completed)	Bicycle	169	0.5%
	Motorcycle	503	1.6%
	Public transport	3868	12.3%
	Walking	7613	24.1%
	Private vehicle	19378	61.5%

After analysing table 21, considering the assumption that a higher level of education is equivalent to greater opportunities, it can be seen that with the increase in the level of education there is also a tendency for greater use of the private car in contrast to other transport options.

The category which has the lowest number of individuals using a private vehicle was category 2 (None or full 1st or 2nd or 3rd school year), the reason for which may be associated with the fact that to obtain a driving license in Portugal it is necessary to be able to read and write, besides having a minimum age of 18 years. Therefore, the lower the level of education, the greater the tendency to move by walking.

- Purpose of trip:

With this variable, we intend to characterize the influences that the reason for the trip will have on the choices of the type of transport of each individual in the AMP.

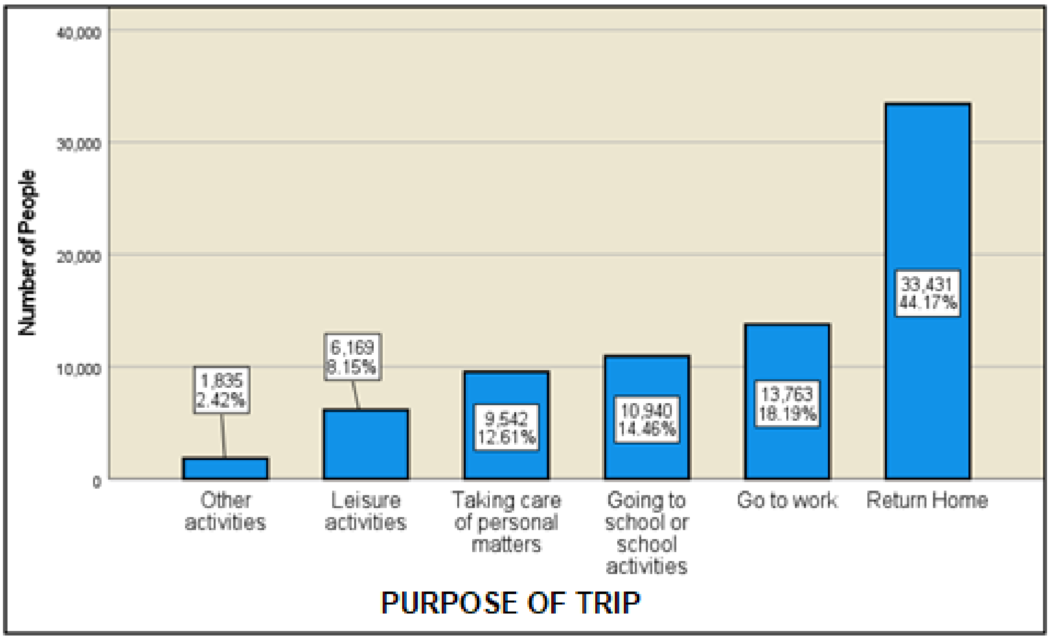

Figure 12 - Distribution of the variable "Purpose of trip" from SPSS software

Given the data in figure 12, it can be seen that the return home has 44.2% of the observations of the sample, so this will be the reference variable.

Table 22 - Modal split according to the variable "Purpose of trip" from SPSS software

Observed and Predicted Frequencies			
Reason for Displacement		Frequency Observed	Percentage Observed
Go to work	Bicycle	77	0.6%
	Motorcycle	278	2.0%
	Public transport	1758	12.8%
	Walking	1683	12.2%
	Private vehicle	9967	72.4%
Going to school or school activities	Bicycle	16	0.1%
	Motorcycle	41	0.4%
	Public transport	1523	13.9%
	Walking	1434	13.1%
	Private vehicle	7926	72.4%
Leisure activities	Bicycle	68	1.1%
	Motorcycle	73	1.2%
	Public transport	291	4.7%
	Walking	2178	35.3%
	Private vehicle	3559	57.7%

Observed and Predicted Frequencies			
Reason for Displacement		Frequency Observed	Percentage Observed
Other activities	Bicycle	10	0.5%
	Motorcycle	17	0.9%
	Public transport	233	12.7%
	Walking	454	24.7%
	Private vehicle	1121	61.1%
Taking care of personal matters	Bicycle	26	0.3%
	Motorcycle	76	0.8%
	Public transport	868	9.1%
	Walking	2322	24.3%
	Private vehicle	6250	65.5%
Return Home	Bicycle	184	0.6%
	Motorcycle	422	1.3%
	Public transport	3834	11.5%
	Walking	7024	21.0%
	Private vehicle	21967	65.7%

Analysing the modal split for each purpose of trip observed, the trips relating to "Go to work" and "Go to school or school activities", have a greater propensity to be carried out by private vehicles. These are also the types of commuting trips.

The other trips because they are associated with non-routine activities and without time pressure, have a higher propensity to be carried out by sustainable modes of transport, especially the "Leisure activities" motive, where walking had a higher percentage observed.

3.5 - Application and analysis of the final model

Considering the high number of variables, Sawalha and Sayed (2006), warn against building over-fitting models, containing many insignificant variables/categories that end up not being as beneficial as expected for the model. Taking this into account, a final model was created by adding all variables that showed the greatest potential to the model. Thus, the variables included in the final model were:

- Dependent variable:
    - Transport mode.

- Independent variables:
    - Gender;
    - Age;
    - Existence of physical limitation;
    - Level of schooling;
    - Condition towards employment;
    - Per Capita Income;
    - Workplace;
    - Workday;
    - Reason for Trip;
    - Distance (Km);
    - Has Vehicle;
    - Driving Frequency;
    - Transport Ticket.

After selecting the variables to be used in the model as described in 3.3, the analysis began with the application of the multinomial logit model considering the dependent variable "Transport mode". The following results were obtained thru statistical program (SPSS):

The "Model Fitting Information" in table 23 contains a chi-squared likelihood ratio test comparing the final model (i.e., containing all predictors) against a null model (intercept only). Statistical significance (Sig.) indicates that the final model of each data analysis has greater significant representativeness over the null model applied for our MNL model analysis since the confidence interval of the model is 95%, so:

- H_0: fit of the constructed model = fit of the null model $p > 0.05$;
- H_1: fit of the constructed model ≠ fit of the null model $p \leq 0.05$;

Table 23 - Evaluation of the Multinomial Regression Model

Model Fitting Information						
Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	138231.409	138268.346	138223.409			
Final	83428.333	84610.320	83172.333	55051.076	124	0.000

Since the p-value (Sig.) is 0.000, the H_0 (Intercept Only) is rejected, and the final model is accepted as more significant.

Table 24 shows the likelihood ratio tests of the overall contribution that each independent variable represents to the model. The likelihood ratio test shows that all variables included in the model explain the dependent variable since all of them have a p-value lower than 0.05.

Table 24 - Likelihood ratio test of the Multinomial Regression model

Likelihood Ratio Tests						
Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	83428.333	84610.320	83172.333[a]	0.000	0	
Distance (Km)	104096.443	105241.492	103848.443	20676.110	4	0.000
Gender	83990.433	85135.482	83742.433	570.100	4	0.000
Existence of physical limitation	83499.579	84607.692	83259.579	87.246	8	0.000
Age	83614.169	84685.344	83382.169	209.836	12	0.000
Transport Ticket	93862.328	95007.378	93614.328	10441.995	4	0.000
Has Vehicle	85744.241	86889.290	85496.241	2323.908	4	0.000
Workday	84370.419	85515.468	84122.419	950.086	4	0.000
Level of schooling	83539.938	84574.176	83315.938	143.605	16	0.000
Condition towards employment	83768.379	84876.492	83528.379	356.046	8	0.000
Workplace	83540.075	84648.188	83300.075	127.742	8	0.000
Per Capita Income	83469.211	84466.513	83253.211	80.878	20	0.000
Purpose of trip	84290.447	85287.748	84074.447	902.114	20	0.000
Driving Frequency	86002.157	87110.269	85762.157	2589.824	8	0.000

The value of R-squared (coefficient of determination) is an indicator used to measure the quality of adjustment of a linear regression, it measures the proportion of the variance of the dependent variable that is explained by the independent variables. In the simple linear regression models, the dependent variable is assumed as being a continuous random variable. When using a Multinominal Logit model, i.e., when the dependent variable is qualitative being expressed by two or more categories, instead of the R-squared, the "pseudo-R-squared" is used to evaluate the adjustment quality of the model. It can be considered a modified version of R-squared, which is adjusted for the number of predictors in the qualitative model.

Although pseudo-R-square values applied to logistic regressions are available for analysis, it is worth remembering that Pseudo R-squared will always be smaller than R-squared because they are analogous values. As a result, Smith and McKenna (2013) when comparing the values of the Pseudo R-square, observed that the values of the Nagelkerke index were the closest to the values of the R-square. Therefore, this analysis is considering the Nagelkerke index as shown in table 25.

Table 25 - Pseudo R-square

Pseudo R-Square	
Cox and Snell	0.517
Nagelkerke	0.616
McFadden	0.398

The results of the final model are shown in Table 26. The analysis of these results is based on the column "Parameter Estimates" of the table, to compare the value estimated for the variable of each category of transport mode with the reference category (Private Vehicle). Therefore, the first interpretation will be based on the estimated coefficient presented here as Beta (coefficient of B or only B) to analyse whether the independent variable influences the probability of choosing a particular mode of transport in relation to the reference category (Private Vehicle). For this, the coefficient B must be statistically different from zero. When there are mathematical values close to zero, we can still use the Wald test to verify if the regression coefficient B is statistically different from zero.

- H_0: coefficient B = 0 if $p > 0.05$;
- H_1: coefficient B $\neq$ 0 if $p \leq 0.05$;

Exp(b) is regarded as the odds ratio and should be interpreted as the multiplicative adjustment of the odds of the outcome. If the unit of measurement is very small compared to the size of a significant change, the odds ratio will be close to one, thus one is the absence of effect:

- Exp (b) > 1 = positive relation;
- Exp (b) < 1 = inverse relation;

The summary of analysis of the variables is shown in figure 13 below:

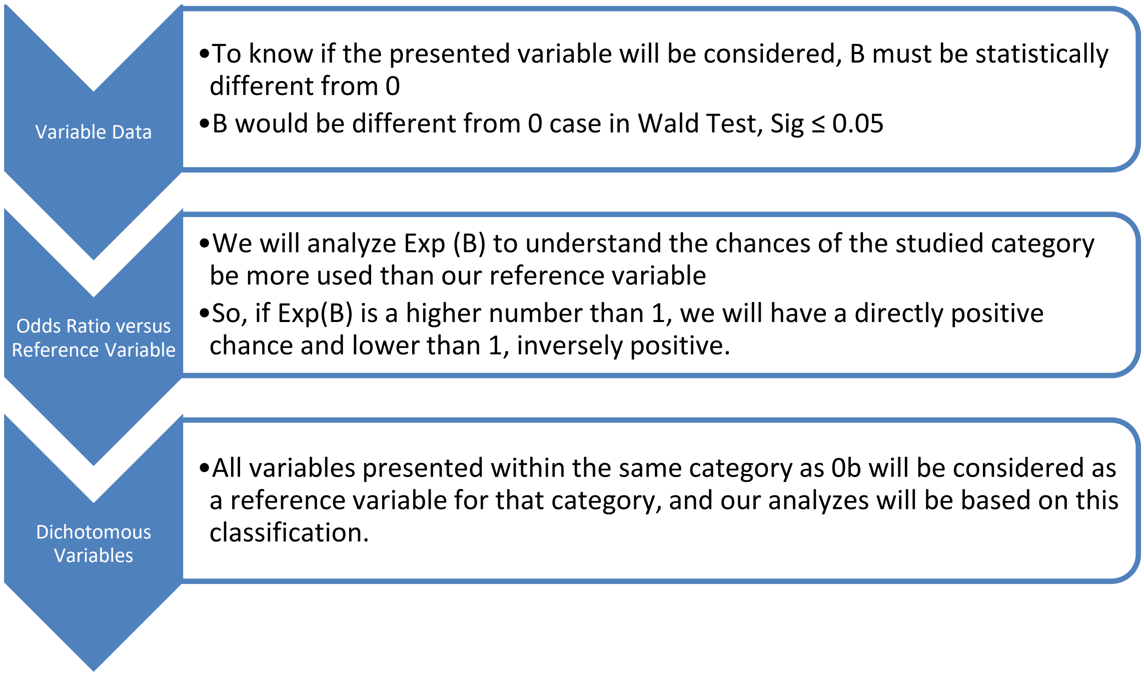

Figure 13 - Infographic explaining the analysis of variables' results

For better observation of the results, the table was divided according to the transport mode analysed, to facilitate the reading of the individual references to each transport mode. Note that all categories that show statistical significance are highlight in red.

- Bicycle

Table 26 - Parameter Estimation for the Transport Mode Bicycle from SPSS software

Parameter Estimates									
Transport mode[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	95% Confidence Interval for Exp(B) Upper Bound
Bicycle	Intercept	-5.028	0.265	361.073	1	0.000			
	Distance (Km)	-0.033	0.007	23.189	1	0.000	0.967	0.954	0.980
	Gender = Female	-1.373	0.123	125.503	1	0.000	0.253	0.199	0.322
	Gender = Male	0[b]			0				
	Existence of physical limitation = Prefer not to answer	-2.841	0.742	14.669	1	0.000	0.058	0.014	0.250
	Existence of physical limitation = Yes	-0.446	0.310	2.072	1	0.150	0.640	0.349	1.175
	Existence of physical limitation = No	0[b]			0				
	Age = 65	-0.885	0.227	15.257	1	0.000	0.413	0.265	0.643
	Age =< 25	-0.274	0.210	1.697	1	0.193	0.761	0.504	1.148
	Age = 25-44	0.164	0.128	1.651	1	0.199	1.178	0.917	1.513
	Age = 45-64	0[b]			0				
	Transport Ticket = No	0.371	0.205	3.275	1	0.070	1.449	0.970	2.166
	Transport Ticket = Yes	0[b]			0				
	Has Vehicle = No	1.564	0.155	101.330	1	0.000	4.777	3.523	6.478
	Has Vehicle = Yes	0[b]			0				
	Workday = No	0.169	0.115	2.153	1	0.142	1.184	0.945	1.484
	Workday = Yes	0[b]			0				
	Level of schooling = Higher Education (Bachelor's, Master's, Doctor's, Higher Professional Technical Course)	0.387	0.154	6.295	1	0.012	1.473	1.088	1.992
	Level of schooling = None or completed 1st or 2nd or 3rd year	-0.291	0.324	0.803	1	0.370	0.748	0.396	1.412
	Level of schooling = Prefer not to answer	1.733	0.434	15.969	1	0.000	5.658	2.418	13.237
	Level of schooling = Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)	-0.003	0.159	0.000	1	0.987	0.997	0.731	1.361
	Level of schooling = Basic Education (1st	0[b]			0				

	cycle, 2nd cycle, or 3rd cycle completed)							
	Condition towards employment = Prefer not to answer	-1.256	0.329	14.599	1	0.000	0.285	0.149	0.542
	Condition towards employment = Student, Retired, Mainly engaged in household chores, Permanently disabled or Other inactivity situation	0.063	0.185	0.114	1	0.735	1.065	0.741	1.530
	Condition towards employment = Unemployed	0.196	0.195	1.011	1	0.315	1.217	0.830	1.784
	Condition towards employment = Employee	0^b			0				
	Workplace = At home	0.931	0.285	10.694	1	0.001	2.538	1.452	4.434
	Workplace = Away from home with no fixed location (driver, salesperson, etc.)	-0.333	0.235	2.005	1	0.157	0.717	0.452	1.136
	Workplace = Prefer not to answer	0^b			0				
	Workplace = Away from home at a fixed location	0^b			0				
	Per Capita Income = Above 2000 euros	-0.007	0.244	0.001	1	0.978	0.993	0.615	1.604
	Per Capita Income = From 1000 to less than 1500 euros	-0.274	0.195	1.980	1	0.159	0.761	0.519	1.114
	Per Capita Income = From 1500 to less than 2000 euros	0.339	0.260	1.703	1	0.192	1.404	0.843	2.337
	Per Capita Income = From 650 to less than 1000 euros	0.030	0.135	0.048	1	0.827	1.030	0.791	1.341
	Per Capita Income = Prefer not to answer	-1.055	0.389	7.361	1	0.007	0.348	0.163	0.746
	Per Capita Income = Up to 650 euros	0^b			0				
	Purpose of trip = Go to work	0.043	0.145	0.089	1	0.766	1.044	0.786	1.387
	Purpose of trip = Going to school or school activities	-1.372	0.266	26.655	1	0.000	0.254	0.151	0.427
	Purpose of trip = Leisure activities	0.696	0.149	21.783	1	0.000	2.005	1.497	2.685
	Purpose of trip = Other activities	0.104	0.332	0.098	1	0.755	1.109	0.578	2.128
	Purpose of trip = Taking care of personal matters	-0.735	0.214	11.752	1	0.001	0.480	0.315	0.730
	Purpose of trip = Return Home	0^b			0				
	Driving Frequency = Drives sporadically	1.717	0.195	77. 887	1	0.000	5.566	3.802	8.150
	Driving Frequency = Unkown	1.796	0.162	122.637	1	0.000	6.025	4.384	8.279

	Driving Frequency = Drives frequently	0[b]			0				

a. The reference category is: Private vehicle

Regarding the choice of the bicycle compared to the private vehicle, it shows that for each kilometer of distance increased, there is a decrease of 3.3% in the choice of the bicycle as the main mode of transport.

The probability of men using the bicycle is higher than for women. Individuals over 64 years old have a 58.7% drop in the choice of bicycle for their trips versus the private vehicle when compared to individuals in the 45 to 64 age group. Not owning a vehicle is a crucial factor in opting for cycling, with 4.8 times more likely to travel by this mode compared to a private vehicle, even as a passenger.

Individuals with a high level of education are 47.3% more likely to travel by bicycle compared to those with a basic level of education. Individuals who work from home are 2.5 times more likely to commute by bicycle than those who work out of home.

For trips related to school activities and to deal with personal matters, there is a drop of 74.6% and 52%, respectively, in the choice of bicycle as the main mode of transport compared to the private vehicle, while in trips to Leisure activities there is an increase of 100% on the preference of the bicycle over the private vehicle. People who drive sporadically tend to commute 5.6 times more often by bicycle than those who drive frequently.

- Motorcycle

Table 27 - Parameter Estimation for the Transport Mode Motorcycle from SPSS software

Parameter Estimates									
Transport mode[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Motorcycle	Intercept	-3.963	0.231	293.743	1	0.000			
	Distance (Km)	-0.033	0.004	54.685	1	0.000	0.968	0.960	0.976
	Gender = Female	-1.646	0.090	338.103	1	0.000	0.193	0.162	0.230
	Gender = Male	0[b]			0				
	Existence of physical limitation = Prefer not to answer	-0.631	0.591	1.140	1	0.286	0.532	0.167	1.695
	Existence of physical limitation = Yes	-0.517	0.202	6.551	1	0.010	0.596	0.401	0.886
	Existence of physical limitation = No	0[b]			0				
	Age = 65	-0.497	0.146	11.507	1	0.001	0.609	0.457	0.811
	Age =< 25	-0.441	0.155	8.135	1	0.004	0.643	0.475	0.871
	Age = 25-44	-0.170	0.083	4.159	1	0.041	0.844	0.717	0.993
	Age = 45-64	0[b]			0				
	Transport Ticket = No	1.158	0.211	29.970	1	0.000	3.183	2.103	4.817
	Transport Ticket = Yes	0[b]			0				
	Has Vehicle = No	1.881	0.105	321.087	1	0.000	6.559	5.340	8.058
	Has Vehicle = Yes	0[b]			0				
	Workday = No	-0.401	0.087	21.273	1	0.000	0.670	0.565	0.794
	Workday = Yes	0[b]			0				
	Level of schooling = Higher Education (Bachelor's, Master's, Doctor's, Higher Professional Technical Course)	-0.479	0.103	21.594	1	0.000	0.620	0.506	0.758
	Level of schooling = None or completed 1st or 2nd or 3rd year	0.397	0.165	5.814	1	0.016	1.487	1.077	2.054
	Level of schooling = Prefer not to answer	-0.274	0.535	0.262	1	0.609	0.760	0.267	2.169
	Level of schooling = Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)	-0.530	0.105	25.448	1	0.000	0.588	0.479	0.723
	Level of schooling = Basic Education (1st	0[b]			0				

	cycle, 2nd cycle, or 3rd cycle completed)								
	Condition towards employment = Prefer not to answer	-3.133	0.347	81. 548	1	0.000	0.044	0.022	0.086
	Condition towards employment = Student, Retired, Mainly engaged in household chores, Permanently disabled or Other inactivity situation	-0.664	0.131	25.554	1	0.000	0.515	0.398	0.666
	Condition towards employment = Unemployed	-0.724	0.169	18.442	1	0.000	0.485	0.348	0.675
	Condition towards employment = Employee	0[b]			0				
	Workplace = At home	-0.764	0.386	3.927	1	0.048	0.466	0.219	0.992
	Workplace = Away from home with no fixed location (driver, salesperson, etc.)	-0.167	0.119	1.966	1	0.161	0.846	0.670	1.069
	Workplace = Prefer not to answer	0[b]			0				
	Workplace = Away from home at a fixed location	0[b]			0				
	PerCapitaIncome = Above 2000 euros	0.199	0.162	1.512	1	0.219	1.220	0.888	1.676
	PerCapitaIncome = From 1000 to less than 1500 euros	0.314	0.113	7.718	1	0.005	1.370	1.097	1.710
	Per Capita Income = From 1500 to less than 2000 euros	0.146	0.212	0.473	1	0.492	1.157	0.764	1.751
	Per Capita Income = From 650 to less than 1000 euros	-0.160	0.097	2.718	1	0.099	0.852	0.705	1.031
	Per Capita Income = Prefer not to answer	-0.001	0.168	0.000	1	0.994	0.999	0.719	1.388
	Per Capita Income = Up to 650 euros	0[b]			0				
	Purpose of trip = Go to work	0.235	0.085	7.718	1	0.005	1.265	1.072	1.493
	Purpose of trip = Going to school or school activities	-1.017	0.168	36.655	1	0.000	0.362	0.260	0.503
	Purpose of trip = Leisure activities	0.095	0.133	0.503	1	0.478	1.099	0.846	1.428
	Purpose of trip = Other activities	0.141	0.256	0.302	1	0.583	1.151	0.697	1.901

	Purpose of trip = Taking care of personal matters	-0.340	0.131	6.699	1	0.010	0.712	0.550	0.921
	Purpose of trip = Return Home	0[b]			0				
	Driving Frequency = Drives sporadically	-0.237	0.288	0.677	1	0.411	0.789	0.449	1.387
	Driving Frequency = Unkown	1.530	0.105	213.171	1	0.000	4.617	3.760	5.670
	Driving Frequency = Drives frequently	0[b]			0				

a. The reference category is: Private vehicle.

As well as the bicycle, there is a drop in the choice of motorcycle to travel long distances and women tend to travel by motorcycles 80.7% less than men. Regarding the age group, for any group age compared with the group reference, i.e., 45-64 years-old, the probability of travelling by motorcycle decreases compared with the private vehicle.

If an individual does not have car, it increases the chances to have a motorcycle by 6.6 times. If the individual does not have a transport ticket/pass, the chances of individuals have a motorcycle instead private vehicle is increased by 3.2 times compared with who have no transport ticket/pass. On non-working days, there is a 33% drop in trips by motorcycle.

Individuals with low or no education level are 1.5 times more likely to choose to travel by motorcycle compared to those with a basic level and those with a high educational level tend to prefer the private vehicle to make their journeys. “Student, Retired, mainly engaged in household chores, permanently disabled or Other inactivity situation” when compared to people Employed and working out of home have a higher tendency to use private vehicle in their trips rather than motorcycle.

Individuals with a per capita income between 1000 and 1500 euros and with the purpose of commuting to work tend to use a motorcycle over a private vehicle. Regarding the trips related to going to school and taking care of personal matters, the private vehicle has a higher probability of being chosen in relation to the motorcycle.

- Public Transport

Table 28 - Parameter Estimation for the Transport Mode Public transport from SPSS software

Parameter Estimates									
Transport mode[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Public transport	Intercept	-0.340	0.058	33.851	1	0.000			
	Distance (Km)	0.006	0.001	45.717	1	0. 000	1.00 6	1.005	1.008
	Gender = Female	-0.076	0.033	5.414	1	0.020	0.92 7	0.869	0.988
	Gender = Male	0[b]			0				
	Existence of physical limitation = Prefer not to answer	-0.375	0.179	4.374	1	0. 036	0.68 7	0.484	0.977
	Existence of physical limitation = Yes	-0.266	0.090	8.712	1	0.003	0.76 7	0.643	0.915
	Existence of physical limitation = No	0[b]			0				
	Age = 65	0.121	0.065	3.469	1	0.063	1.12 9	0.994	1.282
	Age =< 25	0.064	0.061	1.104	1	0.293	1.06 6	0.946	1.202
	Age = 25-44	-0.114	0.045	6.479	1	0.011	0.89 3	0.818	0.974
	Age = 45-64	0[b]			0				
	Transport Ticket = No	-3.242	0.035	8500.3 21	1	0.000	0.03 9	0.036	0.042
	Transport Ticket = Yes	0[b]			0				
	Has Vehicle = No	1.882	0.050	1424.8 06	1	0.000	6.56 9	5.958	7.244
	Has Vehicle = Yes	0[b]			0				
	Workday = No	-1.257	0.044	821.90 4	1	0.000	0.28 5	0.261	0.310
	Workday = Yes	0[b]			0				
	Level of schooling = Higher Education (Bachelor's, Master's, Doctor's, Higher Professional Technical Course)	0.107	0.048	4.883	1	0.027	1.11 3	1.012	1.224
	Level of schooling = None or completed 1st or 2nd or 3rd year	0.182	0.086	4.466	1	0.035	1.19 9	1.013	1.419

	Level of schooling = Prefer not to answer	0.588	0.173	11.533	1	0.001	1.801	1.282	2.528
	Level of schooling = Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)	0.127	0.045	7.971	1	0.005	1.136	1.040	1.241
	Level of schooling = Basic Education (1st cycle, 2nd cycle, or 3rd cycle completed)	0b			0				
	Condition towards employment = Prefer not to answer	-0.417	0.086	23.396	1	0.000	0.659	0.557	0.781
	Condition towards employment = Student, Retired, Mainly engaged in household chores, Permanently disabled or Other inactivity situation	0.283	0.058	23.684	1	0.000	1.327	1.184	1.487
	Condition towards employment = Unemployed	0.502	0.070	51.611	1	0.000	1.652	1.440	1.894
	Condition towards employment = Employee	0b			0				
	Workplace = At home	0.469	0.151	9.651	1	0.002	1.599	1.189	2.150
	Workplace = Away from home with no fixed location (driver, salesperson, etc.)	0.666	0.070	90.346	1	0.000	1.947	1.697	2.234
	Workplace = Prefer not to answer	0b			0				
	Workplace = Away from home at a fixed location	0b			0				
	Per Capita Income = Above 2000 euros	-0.180	0.081	4.995	1	0.025	0.835	0.713	0.978
	Per Capita Income = From 1000 to less than 1500 euros	-0.071	0.053	1.797	1	0.180	0.931	0.840	1.033
	Per Capita Income = From 1500 to less than 2000 euros	-0.291	0.100	8.412	1	0.004	0.747	0.614	0.910
	Per Capita Income = From 650 to less than 1000 euros	-0.198	0.041	23.009	1	0.000	0.821	0.757	0.890
	Per Capita Income = Prefer not to answer	0.107	0.072	2.229	1	0.135	1.113	0.967	1.282
	Per Capita Income = Up to 650 euros	0b			0				
	Purpose of trip = Go to work	0.369	0.045	68.236	1	0.000	1.447	1.325	1.579
	Purpose of trip = Going to school or school activities	-0.193	0.047	16.878	1	0.000	0.824	0.752	0.904

	Purpose of trip = Leisure activities	-0.614	0.078	62.462	1	0.000	0.541	0.465	0.630
	Purpose of trip = Other activities	0.258	0.100	6.664	1	0.010	1.295	1.064	1.575
	Purpose of trip = Taking care of personal matters	-0.254	0.054	22.471	1	0.000	0.776	0.698	0.862
	Purpose of trip = Return Home	0^b			0				
	Driving Frequency = Drives sporadically	1.437	0.057	636.072	1	0.000	4.209	3.764	4.707
	Driving Frequency = Unkown	1.241	0.048	680.427	1	0.000	3.460	3.152	3.798
	Driving Frequency = Drives frequently	0^b			0				

It was observed when the distance increases, the probability of individuals travelling by public transport increases even having a private vehicle at home. Also even having a private vehicle at home, men are more prone to use the public transport compared to women. Regarding the age group, individuals who are over 64 years old are more prone to travel by public transport than by private vehicle and in contrast, those who are between 25 and 44 years old are less prone to use public transport when compared with individuals in the age group of 45 to 64 years old.

There is a 96% drop in the choice of public transport by people who do not have a transport ticket compared to those who do. Additionally, those people who do not have their own vehicle tend to travel by public transport 6 times more than those who do have their own vehicle. Public transport is also more prone to be chosen than private vehicle on weekdays, with a drop of 71.5% on weekends.

Regarding the schooling level, it was observed that despite public transport being a mode with the higher probability to be used than private vehicles as shown by the estimated positive values for all categories, its more common for individuals who have no level of schooling or have an incomplete basic education the use of public transport. “Student, Retired, mainly engaged in household chores, Permanently disabled or Other inactivity situation” have 1.3 times more tendency to choose public transport instead of private vehicle, and unemployed have 1.7 times more tendency to use public transport than those who are employed. Individuals who work from home tend to use public transport 60% more when compared to those who work out of home at a fixed location; those who work out of home but without a fixed location compared to those who work at a fixed location tend to opt 94.7% more for public transport.

It can also be observed that individuals with a higher per capita income (more than 1500 euros per month) tend to use private vehicles over public transport, with 74.7% more chances when compared to individuals that have a per capita income up to 650 euros. Regarding the purpose of trip, it can be observed that for all trips there are a decrease in the probability of using public transport, except in the trips related to work, where there was an increase of 44.7% in the choice of public transport. Public transport is also an option for people who drive sporadically, being chosen at a value of 4.2 times more than those who drive frequently.

- Walking

Table 29 - Parameter Estimation for the Transport Mode Walking from SPSS software

Parameter Estimates									
Transport mode[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Walking	Intercept	0.762	0.056	183.281	1	0.000			
	Distance (Km)	-0.619	0.007	7513.349	1	0.000	0.538	0.531	0.546
	Gender = Female	-0.171	0.026	44.499	1	0.000	0.843	0.801	0.886
	Gender = Male	0[b]			0				
	Existence of physical limitation = Prefer not to answer	-0.539	0.148	13.300	1	0.000	0.583	0.437	0.779
	Existence of physical limitation = Yes	-0.502	0.069	53.421	1	0.000	0.606	0.529	0.693
	Existence of physical limitation = No	0[b]			0				
	Age = 65	-0.070	0.045	2.459	1	0.117	0.932	0.854	1.018
	Age =< 25	-0.545	0.052	108.026	1	0.000	0.580	0.523	0.643
	Age = 25-44	-0.268	0.034	62.955	1	0.000	0.765	0.715	0.817
	Age = 45-64	0[b]			0				
	Transport Ticket = No	-0.498	0.039	161.682	1	0.000	0.608	0.563	0.656
	Transport Ticket = Yes	0[b]			0				

	Has Vehicle = No	1.748	0.046	1435.50 9	1	0.00 0	5.744	5.247	6.288
	Has Vehicle = Yes	0b			0				
	Workday = No	-0.145	0.028	25.869	1	0.00 0	0.865	0.818	0.915
	Workday = Yes	0b			0				
	Level of schooling = Higher Education (Bachelor's, Master's, Doctor's, Higher Professional Technical Course)	0.242	0.037	43.742	1	0.00 0	1.274	1.186	1.369
	Level of schooling = None or completed 1st or 2nd or 3rd year	-0.195	0.066	8.826	1	0.00 3	0.823	0.723	0.936
	Level of schooling = Prefer not to answer	0.144	0.144	0.998	1	0.31 8	1.154	0.871	1.530
	Level of schooling = Secondary Education (12th year of complete schooling) or Post-secondary (non-higher technological specialization course)	0.115	0.037	9.650	1	0.00 2	1.122	1.043	1.206
	Level of schooling = Basic Education (1st cycle, 2nd cycle, or 3rd cycle completed)	0b			0				
	Condition towards employment = Prefer not to answer	-0. 346	0.069	24.845	1	0.00 0	0.707	0.617	0.810
	Condition towards employment = Student, Retired, Mainly engaged in household chores, Permanently disabled or Other inactivity situation	0.512	0.043	141.177	1	0.00 0	1.668	1.533	1.815
	Condition towards employment = Unemployed	0.596	0.050	143.550	1	0.00 0	1.815	1.646	2.000
	Condition towards employment = Employee	0b			0				
	Workplace = At home	0.385	0.096	16.050	1	0.00 0	1.469	1.217	1.774
	Workplace = Away from home with no fixed location (driver, salesperson, etc.)	-0.063	0.062	1.053	1	0.30 5	0.939	0.832	1.059
	Workplace = Prefer not to answer	0b			0				

	Workplace = Away from home at a fixed location	0[b]			0				
	PerCapitaIncome = Above 2000 euros	0.025	0.057	0.190	1	0.663	1.025	0.916	1.148
	PerCapitaIncome = From 1000 to less than 1500 euros	-0.041	0.041	1.014	1	0.314	0.960	0.886	1.040
	Per Capita Income = From 1500 to less than 2000 euros	-0.208	0.070	8.899	1	0.003	0.812	0.709	0.931
	Per Capita Income = From 650 to less than 1000 euros	-0.089	0.032	7.844	1	0.005	0.915	0.859	0.974
	Per Capita Income = Prefer not to answer	0.081	0.056	2.073	1	0.150	1.084	0.971	1.210
	Per Capita Income = Up to 650 euros	0[b]			0				
	Purpose of trip = Go to work	-0.038	0.039	0.927	1	0.336	0.963	0.891	1.040
	Purpose of trip = Going to school or school activities	-0.672	0.039	290.410	1	0.000	0.510	0.472	0.551
	Purpose of trip = Leisure activities	0.432	0.041	112.094	1	0.000	1.540	1.421	1.668
	Purpose of trip = Other activities	0.351	0.081	18.956	1	0.000	1.421	1.213	1.664
	Purpose of trip = Taking care of personal matters	-0.168	0.038	19.756	1	0.000	0.845	0.785	0.910
	Purpose of trip = Return Home	0[b]			0				
	Driving Frequency = Drives sporadically	1.559	0.049	999.921	1	0.000	4.756	4.318	5.238
	Driving Frequency = Unkown	1.394	0.039	1264.214	1	0.000	4.030	3.732	4.352
	Driving Frequency = Drives frequently	0[b]			0				

a. The reference category is: Private vehicle.

The choice for walking when compared to the private vehicle drops 46.2% as the distance increases. Regarding the age group, excluding individuals over 64 years of age, all the estimated values are negative denoting that the probability of walking decreases compared with those with the ages between 45-64 years old.

Individuals who do not have their own vehicle tend to choose to walk 5.7 times more than those who have their own vehicle. Those individuals who do not have a transport ticket/pass, a 39.2% drop could be observed in walking when compared to those who have a transport ticket/pass. Walking is preferable to the car on weekdays, with a 13.5% drop in citizens' choice of this transport mode on weekends.

In relation to the schooling level, when comparing all levels with the completed basic level, it is observed that the higher the educational level of individuals, more they tend to prefer travelling by walking instead of using a private vehicle. Students, retirees, and the unemployed, when compared to employees, also tend to travel more by walking than by private vehicle.

Regarding the per capita income, it was observed that the higher the individual's per capita income, the lower is the probability of he/she prefers to walk. For the purpose of trip, such as going to school and taking care of personal matters, the private vehicle tends to be more likely to be used than walking, while for Leisure activities, a 54% increase was observed in the preference of walking in relation to the use of a private vehicle. When comparing individuals who drive sporadically with those who drive frequently, it can be observed that those who drive sporadically tend to choose 4.8 times more walking than private vehicle for their trips.

Another detail observed in the analysis was that individuals with reduced mobility prefer to travel by private vehicle over any other mode of transport.

Chapter 4

Conclusion and future developments

Currently, sustainable urban mobility is a challenge for local and regional governments, cities seek to increase sustainability and implement new technologies. Furthermore, mobility plans are used to plan transport networks and services to improve urban mobility (Duarte, 2021).

Public policies can promote more efficient mobility if based on real evidence. Even academic researchers are increasing by the time, the level of knowledge about sustainable mobility or regarding Smart Cities need more developments and discussion with the population to become an important matter on daily basis actions and talks. According to the Smart Cities Index (INTELI, 2016), none of the 124 municipalities surveyed in Portugal has an integrated urban management platform in place, with only 6% providing a dashboard with critical urban indicators. According to the respondents, the most valuable information regarding this issue is only the real-time information regarding public transport timetables and traffic flow.

The low dissemination and lack of information on smart cities and smart mobility can be improved by creating additional efforts to promote citizen participation in the co-creation of solutions to urban problems and ideas to respond to the future challenges of the AMP, thus investing in the promotion of the solutions and in initiatives to raise citizens' awareness on these issues, through seminars and public workshops. With the large growth of the AMP in recent years, some public problems related to urban mobility have also started to appear, generating new obstacles for policymakers and urban planners in this region.

It is legitimate to consider that urban mobility, and in particular sustainable transport systems, should be part of the core structure of the political agenda, fulfilling a set of requirements concerning to ensure stability and adaptability to people's needs. It is necessary that public transport can meet the different demands of the population by ensuring interconnections and integration with different modes of transport, basing its operation on frequency, regularity, and reliability, and ensuring service levels suited to the needs.

The implementation of a smart city becomes primordial for the development of a more sustainable city, seeking to automate the development of AMP, based on technologies associated with the increasing availability of information and data, facilitating the availability, visualization, and use of these data, which can be handled by public managers or responsible companies, capable of connecting the public and private community, citizens, and entities.

Therefore, to understand the current situation of the AMP in relation to its urban mobility patterns, and to try to create public policies to influence people to adopt more sustainable behaviors, it is necessary to study more than one-day behavior of the population and follow a timeline of decisions in a real daily basis need.

Through a multinomial logit regression model to analyse the relationships between the modal choices of individuals according to their characteristics (age, income, the reason for travel, level of education, etc.), were examined the relations between the modal choices of individuals in their daily trips with their characteristics and daily activities, thus identifying those variables that had a greater effect for the citizen's choice of transport.

Several statistical tests were carried out in the quest to evaluate the variables collected in INE’s survey, thus the results found will be useful for transport modelers who would like to examine what factors may influence travelers’ decision-making regarding their modal choices.

After the analyses, it was observed that private vehicle was the main transport mode among the citizens of AMP. Through the report conducted by INE (2018), the main reason that led AMP residents to choose the private vehicle was the "speed", pointed out by 58.8% of respondents. It also pointed out that there is a small dissatisfaction with the public transport system in the AMP, regarding the "price/cost of public transport", its "accessibility for people with disabilities", and "reliability/punctuality" of the service along with the information provided to the public regarding schedules (period of operation and frequency of services).

As the main result of the application of the model, it was found that the independent variables that had a greater explanatory significance for the model, were the variables “distance”, “purpose of trip” and the “per capita income”, being these decisive factors for the modal choices of individuals. It was also observed that when it comes to the variable "purpose of trip", the individual when making a displacement that requires a degree of punctuality for their arrival at the destination place or for routine activities, such as taking care of personal matters or going to school activities, prefers the use of private vehicle compared to other modes of transport, possibly due to the speed, comfort of and flexibility associated to this transport mode.

Regarding the “schooling level”, it was observed that people who have a higher level of education tend to choose sustainable modes such as the use of public transport and walking when compared to those who have a basic level. About the “per capita income” Individuals with a lower per capita income have a greater tendency to have a public transport pass or not own a private vehicle due to the high costs of use and maintenance. Those who have a per capita income above 1500 euros per month or a higher level of education have a greater tendency to use private vehicles as opposed to public transport.

For the other socio-demographic variables of the model when comparing the modal preferences by gender, there was a certain homogeneity among the answers of the respondents, and thus there are not so many behavioral differences between genders regarding their modal choices. In relation to the age group, it was observed that the older the age the greater the tendency for the individual to select more sustainable modal choices. Therefore, according to the results of the model generated, it was possible to verify that the socio-demographic conditions of citizens have a high potential influence on their daily modal choices, also revealing that the owning a private vehicle may end up acting, as expected, as an important factor in the decision making of their modal choice.

Seeking also to make a comparison with the results currently found in AMP and the modal preferences of AML, a parallel was made between this study and a similar study, based on the same IMob data (2017), for AML.

Habibo (2020) also used the INE (2018) study to analyse AML region. The database for AML contained initial sample of 62,712 individuals and after applied several restrictions to obtain the appropriate working sample of his analysis, it was presented the final sample consisting of 7353 valid answers. Habibo (2020) used a multinomial regression model and he found the following results: as in AMP, the private vehicle is the main type of transport used by AML citizens, representing 60% of the trips; public transport represents around 24% in AML, which is still more than twice the value observed in AMP (11%); about the walking mode, AMP has practically the double of AML's value (around 10%), with this being approximately 20% of the individuals' preferred option; finally, the main reason for the difference between AML and AMP could be the distribution of the population in both regions, where Porto has a central area smaller than Lisbon, thus facilitating the trips by foot. It was also observed that both AML and AMP residents with higher education tend to use more private vehicles for their commuting and that people earning less than 1000 euros in AML tend to choose public transport over other transport modes. In the AMP, people with the same income tend to opt for walking on their journeys.

On the other hand, with the interest of observing whether the behaviors and modal choices of individuals undergo changes in the AMP over the years, the results found in the present study were indirect compared with the results already obtained in a similar study conducted in AMP.

In this indirect comparison with a similar study to the AMP region, it was observed that the preferences of both genders were stable as to its modal behavior. Gonçalves (2012), observed that young adult men and women with a higher level of education tend to use the private vehicle more than the older ones, values similar to what was found in this study.

Another interesting fact observed is that despite the private vehicle being the main modal choice of individuals in both studies, in the present study a drop of approximately 10% in the use of this mode of transport in individuals with a higher level of education was noted, a value which was transferred to the transport mode "walking". This may reveal a good idea about how public policies can promote changes on individuals' modal choices when provided to citizens initiatives for the use of soft transport modes, considering that, although the AMP still has few transport policies for citizens in this area, a small behavioral change can be observed when comparing the two studies. However, this effect may come from a global trend towards the use of these transportation modes.

A possible solution to reduce the use of the private vehicle among AMP residents would be to implement a greater collective awareness of the need to use the public transport network, to contribute to a cleaner planet, besides creating attractive conditions for its use, through better transport policies, promoting its use and offering appropriate alternatives such as better metro network, better trains, a greater supply of buses, and the creation of exclusive bus lanes in order to give priority to public transport (Afonso, 2015).

With the intention of generating some economic solutions for public planners, we can encourage some policies such as banning the access of private vehicles in the main centers of AMP that exceed the limits of pollutant emissions established by AMP and/or establishing rotations of private vehicles carrying only one occupant. Another possible solution is through parking, according to (Duarte, 2021) urban planners and the municipality can use the information on parking behaviors (existing parking spaces, parking availability, stops, loading and unloading bays, etc.) to study possible interventions on it, seeking to substantially increase the price by reducing available spaces, increasing the width of pavements and thus creating good cycling networks. Although these are considered unpopular measures, they would certainly contribute to the reduction of a few thousand private vehicles in urban centers, causing more people to seek more sustainable transportation mode or the sharing of the private vehicle in their trips (Rebordão, 2010).

One difficulty observed in the research based on surveys similar to the one used in this study is that these only use data relating to the one-day trip, not considering the variability in mode choice on other days.

As changes occur in AMP transport policies as well as in people's lifestyles namely with the increase of telecommuting, a phenomenon that had a considerable growth during the COVID-19 pandemic, future research should seek for an increased focus on the dynamic processes of travel behavior, such as learning and changing rhythms and routines of individuals, thus considering the intrapersonal variability in the studies. To this respect, according to (Schlich and Axhausen), previous studies have shown that travel behavior is not fully repetitive and not fully variable. Although many behaviors that represents a daily pattern are highly repetitive, the similarity between daily travel patterns on different days in an individual's longitudinal record is quite low (Huff and Hanson, 2015).

Based on other studies, it was observed that the oscillations of modal choices of individuals are related to the stages of life of each one. It was seen that young adolescents can be considered as multimodal people, due to the lack of a driving license and the difficulty of owning a private vehicle, causing those who cannot use the private vehicle as a passenger, to use alternative transportation modes, thus the percentage of multimodal people tends to decrease dramatically with the entry into professional life (Nobis, 2017). Regarding persons with a single marital status, these tend to be more multimodal than married couples. The car is often the best choice for families. However, public transport can be the best choice for single people in specific situations (Kuhnimhof, et.al).

Regarding age analysis, Pas and Sundar (1995) brings in their study that due to some restrictions regarding spatial mobility on elderly people, it was observed that modal variability is reduced when the person is older than 60 years, works full time, has low household income, has regular access to a private vehicle or bicycle, and does not have a public transport pass.

According to Heinen, et.al (2011), cyclists' decisions to commute by bicycle are more affected by positive weather conditions and cyclists have a lower tendency to use the bicycle when they need to ride in less favoring conditions, such as going against the wind speed, up hills and if they need to be in different locations with great distance from each other in the same day. This is an important factor in the AMP region and one that may be constraint to use this mode of transport.

On the other hand, the study carried out may contribute to the application of a 4 - step model, commonly used in traffic studies, consisting of the following steps: trip generation, trip distribution, modal choice, and trip allocation. To this end, some additional considerations will be necessary to better integrate the model developed in this type of traffic study, being, therefore, a possible future development of the present study.

Overall, the research provided promising insights into the modal choice behavior of the population in the AMP. However, some limitations resulting from the data should be mentioned, namely the fact that observations were discarded due to lack of information, and some difficulties in working with the data of origin and destination of the trips since they are codified. The exploration of other statistical models may bring additional insights, also improving the fit of the models to the database. This is the case of the mixed logit model, which by treating the heterogeneity of the observations may add specific information about some variables.

The variables related to the monthly costs of individuals about spending on fuel, tolls and public transport would be of extreme importance to further explore this subject, requiring special attention in relation to their aggregation and categorization.

It is expected that this study on modal choice with data from the 2017 mobility survey conducted in the AMP will be the basis for several future analyses contributing to the knowledge of mobility in this area and to the promotion of new policies and actions towards sustainable and smart mobility.

Bibliography

A European Green Deal. Available in https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/ (accessed on 19/01/2022).

Afonso, M. T. (2015); Transporte público e mobilidade mais sustentável: Análise e avaliação de boas práticas. Dissertação para obtenção do Grau de Mestre em Urbanismo e Ordenamento do Território - Tecnico de Lisboa.

Ashalatha, R., Manju, V.S. and Zacharia, A.B. (2013); Mode Choice Behavior of Commuters in Thiruvananthapuram City. J. Transp. Eng.; p. 494-502.

Bhat, C.R. and Pendyala, R.M. (2005); Modeling intra-household interactions and group decision-making. Transportation; vol. 32; p. 443-448.

Bratzel, S. (1999); Conditions of success in sustainable urban transport policy change in relatively successful European cities. Transp. Rev. 19; p. 177–190.

Banister, D. and Marshall, S. (2007); Land Use and Transport Planning - European Perspectives on Integrated Policies. London: Elsevier.

Bowman, J. and Ben-Akiva, M. (1999); The day activity schedule approach to travel demand analysis. Paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC.

Bowman, J. and Ben-Akiva, M. (2001); Activity-Based disaggregate travel demand model system with activity schedules. Transportation Research A 35A, 1–28.

Carbon Dioxide Emission. Statista, 2020. Available in https://www.statista.com/statistics/276629/global-co2-emissions/ (accessed on 20/01/2022).

Davidson, P. (2008); Introducing a 'new way'; Guest Editor; Traffic Engineering and Control; Vol. 49, p. 322-323.

Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J. and Picado, R. (2007); o-d. Transportation Research Part A: Policy and Practice. Vol. 41; p. 464-488.

De Witte, A., Hollevoet, J., Dobruszkes, F. and Hubert M. C. (2013); Linking modal choice to motility: A comprehensive review. Transportation Research Part A: Policy and Practice; vol. 49; p. 329-341.

Duarte, S. P. M. (2021); A participative service design framework for urban mobility in a socio-technical transition context thesis submitted to the faculty of engineering, university of porto for the doctoral program in transportation systems MIT portugal program.

Feng, J., Dijst, M. J., Wissink, B. and Prillwitz, J. (2014); Understanding mode choice in the Chinese context: the case of Nanjing metropolitan area Tijdschr. Econ. Soc. Geogr.; p. 315-330.

Ferreira, S. M. P. (2010); A Segurança Rodoviária no Processo de Planeamento de Redes de Transporte em Meio Urbano. Universidade do Porto, Dissertação para Doutoramento em Engenharia Civil. Porto, Portugal.

Ferreira, D., J. P. Silva e A. B. Silva (2013); Impactos dos Modos de Transporte Sustentáveis em Instituições de Ensino Superior – o Caso do Instituto Politécnico de Leiria. Anais 7° Congresso Rodoviário Português – Novos desafios para a atividade rodoviária, CRP, Lisboa.

Gärling, T and Steg, L. (2007); Threats to the quality of urban life from car traffic: problems, causes, and solutions, Elsevier, Amsterdam.

Gonçalves, J. H. G. (2012); Fatores relevantes para a escolha modal em áreas urbanas. Tese para obtenção do Grau de Doutor em Engenharia Civil pela Universidade Da Beira Interior.

Greene, W. H. (2008); Econometric Analysis. Pearson International Edition, United States of America.

Habibo, E. N. I. (2020); Exploring The Drivers Of Commuting Mode Choice In Lisbon´S Metropolitan Area. Master´s final work dissertation in economics and public policy: Universidade de Lisboa

Handy S.L., Cao X.Y. and Mokhtarian P.L. (2005); "Correlation or causality between the built environment and travel behavior? Evidence from Northern California", Transportation Research D, 10, pp. 427-444.

Heinen, E. and Chatterjee, K. (2015); The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. Transp. Res. Part A Policy Pract.; vol. 78; p. 266-282.

Heinen, E., Maat, K. and van Wee, B. (2011); Day-to-day choice to commute or not by bicycle. Transp. Res. Rec.; p. 9-18.

Huff, J.O. and Hanson, S. (1986); Repetition and variability in urban travel.

INE, I.P., Estimativas Anuais da População Residente (2016); MA- DGT, Carta Administrativa Oficial de Portugal

INE - Instituto Nacional de Estatística (2002); "Inquérito à Mobilidade da População Residente 2000"; Instituto Nacional de Estatística, Direção Regional do Norte; financ. Direção Geral de Transportes Terrestres. – Porto

Instituto Nacional de Estatística - Mobilidade e funcionalidade do território nas Áreas Metropolitanas do Porto e de Lisboa: 2017. Lisboa: INE, 2018. Available in www: url:https://www.ine.pt/xurl/pub/349495406 (accessed on 05/01/2022).

INTELI. Available in http://www.inteli.pt/pt/go/rener1 (accessed on 10/12/2021).

Kitamura, R. (2009); "Lifestyle and travel demand", Transportation, 36, pp. 679-710.

Klöckner, C. (2004). How single events change travel mode choice - a life span perspective.

Kourtit, K. and Nijkamp, P. (2012); Smart cities in the innovation age, Innovation: The European Journal of Social Science Research; p. 93-95.

Kuhnimhof, K., Chlond, B. and Von der Ruhren, S. (2006); Users of transport modes and multimodal travel behavior. Steps to understanding travelers' options and choices.

kuhnimhof, T., Kroger, L. and Trommer, S. (2016); Modelling the Impact of Automated Driving. Private AV Scenarios for Germany and the US.

Lanzini, P. and Khan, S.A. (2017); Shedding light on the psychological and behavioral determinants of travel mode choice: a meta-analysis Transport. Res. F Traffic Psychol. Behav.; vol. 48; p. 13-27.

Mackett, R. L. (2003); Why do people use their cars for short trips? Transportation, p. 329-349.

Mapa da AMP. Available in http://portal.amp.pt/pt/4/municipios/ (accessed on 10/12/2021).

Mcfadden, D. L. and Domencich, T. (1996); Urban Travel Demand: A Behavioral Analysis. North-Holland Publishing Co. Germany.

Meyer, M. D. and Miller, E. J. (2001); Urban Transportation Planning. McGraw Hill, United States of America.

Nobis, C. (2007); Multimodality-facets and causes of sustainable mobility behaviour. Transp. Res. Rec.

Ortuzar, J. D. and Willumsen, L. G. (2011); Modelling Transport 4th edition. John Wiley and Sons Ltd, United Kingdom.

Pas, E.I. and Sundar, S. (1995); Intrapersonal variability in daily urban travel behavior: Some additional evidence.

Paulssen, M., Temme, D., Vij, A. and Walker, J.L.. (2014); Values, attitudes and travel behavior: A hierarchical latent variable mixed logit model of travel mode choice. Transportation; vol. 41, p. 873-888.

Porto é a cidade com mais trânsito da Península Ibérica. Available in https://www.jn.pt/local/noticias/porto/porto/porto-e-a-cidade-com-mais-transito-da-peninsula-iberica-13224060.html (accessed on 05/03/2022).

Portugal Urban Population. Available in https://www.worldometers.info/demographics/portugal-demographics/#urb (accessed on 18/01/2022).

Ramos, P. (2001); Estruturação espacial, transportes e desenvolvimento sustentável da área metropolitana do Porto. Análise e interdependências e opiniões de política urbano. Dissertação (mestrado) Faculdade de engenharia da universidade do Porto, Porto,Portugal.

Rebordão, H.F.S.G. (2010); Mitigação dos problemas de estacionamento na área do Hospital Pêro da Covilhã. Dissertação para obtenção do Grau de Mestre em Engenharia Civil Universidade Da Beira Interior.

Rural and urban population. WUP, 2019. Available in https://population.un.org/wup/Country-Profiles/ (accessed on 14/12/2021).

Salomon, I. and Ben-Akiva, M. (1983); "The use of the life-style concept in travel demand models", Environment and Planning A, 15, pp. 623-638.

Scheiner, J. and Holz-Rau, C (2007); Travel mode choice: Affected by objective or subjective determinants?

Schlich, R. and Axhausen, K.W. (2003); Habitual travel behavior. Evidence from a six-week travel diary. Transportation; vol. 30; p. 16-36.

Smith, A., Stirling, A. and Berkhout, F. (2005); The governance of sustainable socio-technical transitions. Res. Policy.

Souza, A. C. S. (2015); Localização espacial de unidades de informação para o deslocamento de pedestres utilizando Isovistas. Dissertação de Mestrado em Transportes, Publicação T.DM-004A/2015, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, DF.

Tagore, M. and Sikdar, P. (1995); A new accessibility measure accounting mobility Parameters. Paper presented at 7th World Conference on Transport Research. The University of New South Wales: Sydney.

Timmermans, H., Arentze, T. and Joh, C.-H. (2002); Analyzing space time behaviour: New approaches to old problems. Prog. Hum. Geogr.; vol. 26; p. 175–190.

The level of knowledge about sustainable mobility. Available in https://smart-cities.pt/opiniao-entrevista/projectos-estruturados-integrados-catarina-selada2411/ (accessed on 20/02/2022).

The traffic of 516 cities in 57 countries. Available in https://www.tomtom.com/en_gb/traffic-index/porto-traffic/ (accessed on 20/02/2022).

Thompson, W. S. (1935); Urbanization. Encyclopedia of the Social Sciences, XV; p. 189.

Thompson, W. S. (1946); Population and peace in the Pacific. University of Chicago Press, 2nd ed.

Thomson, J. M. (1977); Great cities and their traffic; Gollancz. London: Victor Gollancz.

Urban Area. National Geographic, 2011. Available in https://www.nationalgeographic.org/encyclopedia/urban-area/ (accessed on 15/01/2022).

van Wee, B. (2002); "Land use and transport : Research and policy challenges", Journal of Transport Geography, 10, pp. 259-271.

van Wee, B. (2009); "Self-selection : A key to a better understanding of location choices, travel behaviour and transport externalities ?", *Transport Reviews*, *29*, pp. 279-292.

Vovsha, P. and Freedman, J. (2008); What is an Activity-Based Model and how does it work?. Traffic engineering and control. Vol. 49, p. 330-335.

World population projection to live in urban areas by 2050. UN, 2019. Available in https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html (accessed on 12/01/2022).

Wallace, B., Barnes, J. and Rutherford, G.S. (2000); Evaluating the Effect of traveler and trip characteristics on Trip chaining, with implications for Transportation Demand Management strategies.

Wang, F. and Ross, C.L. (2018); Machine Learning Travel Mode Choices: Comparing the Performance of an Extreme Gradient Boosting Model with a Multinomial Logit Model.

Yang, M., Wang, W., Chen, X., Wan, T. and Xu, R. (2007); Empirical analysis of commute trip chaining: Case study of Shangyu, China.