# TRAVEL CHOICE MODELLING (SET OF MODELS, CODE). FIRST RELEASE

## D3.2

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 955273

# Deliverable administrative information

| | |
|---|---|
| **Deliverable number** | D3.2 |
| **Deliverable title** | TRAVEL CHOICE MODELLING (SET OF MODELS, CODE). FIRST RELEASE |
| **Dissemination level** | Public |
| **Submission deadline** | 30/04/2023 |
| **Version number** | 1.0 |
| **Authors** | Christos Konstantinou (NTUA)<br>Marios Giouroukelis (NTUA)<br>Eleni Mantouka (NTUA)<br>Eleni Vlahogianni (NTUA) |
| **Internal reviewers** | Athina Tympakianaki (Aimsun)<br>Ynte Vanderhoydonc (imec) |
| **Document approval** | n/a |

## Legal Disclaimer



https://twitter.com/TANGENT_H2020

https://www.linkedin.com/company/tangent-project/

https://www.youtube.com/channel/UCjhz4kwEm_sTHj7fE4zXToA

For further information please visit http://www.tangent-h2020.eu/

# Executive summary

This document represents the second deliverable, namely D3.2 “Travel Choice Modelling” from Task 3.3, “Travel choice modelling” which takes input from the data collection process of Task 3.2 “Surveys Design and Data Collection” of WP3 “Travel Behaviour Modelling”. The objective of the deliverable is to define the process and methodology to develop travel choice models based on preference data elicited through a structured survey.

To collect data for creating and calibrating travel choice models, a travel behaviour survey was designed and disseminated through the CONEY toolkit, which employs a conversational approach to collect stated preference data by enhancing user experience. The scope of the survey was to determine how the introduction of new services, the new traffic management strategies or the alteration of strategic goals in a city may affect travellers’ choices. The survey targeted to collect information from 1,000 participants for each case study in all four cities, and it was disseminated online and on-site. Until M20, 1,859 questionnaires had been collected across all cities, with the collected sample being well-distributed and overall suitable for model development.

Based on the collected sample, parametric (Multinomial Logit) and machine learning (Random Forest) models have been developed to predict the mode choice in an aggregated way. The results indicated that the Random Forest models outperformed parametric models in predicting the travel choice of each user in all cities. Moreover, the study investigated the factors that influence the choice among the available travel modes for each user. The most important factors for all cities studied were found to be the trip cost and the total duration of the journey, while other significant factors included age, income, and the degree of satisfaction with the current level of service of public transport.

Overall, this study's findings aim to provide a basis for forecasting long-term demand shifts and provides valuable insights into travel mode choice which can inform future urban mobility planning efforts.

## Key words

Travel Behaviour Survey, Travel Choice Modelling, stated preference, CONEY, discrete choice analysis

# Table of contents

# List of figures



# List of tables

# List of abbreviations and acronyms

| Acronym | Meaning |
| --- | --- |
| API | Application Programming Interface |
| CAV | Connected and Autonomous Vehicle |
| COP | Common Operational Picture |
| DCP | Dynamic Congestion Pricing |
| ELSTAT | Hellenic Statistical Authority |
| EPDM | Personal Motorized Mobility Solutions |
| GA | Grant Agreement |
| INE | Instituto Nacional de Estatística - National Institute for Statistics (Portugal) |
| INSEE | Institut national de la statistique et des études économiques - National Institute of Statistics and Economic Studies (France) |
| M | Month |
| MLN | Multinomial Logit Model |
| OD | Origin-Destination |
| ONS | Office for National Statistics (UK) |
| PT | Public Transport |
| RF | Random Forest Model |
| TfGM | Transport for Greater Manchester |
| TMC | Traffic Management Centre |
| WP | Work Package |
| WTP | Willingness-to-pay |

# 1 Introduction

When planning a trip, individuals must make several decisions regarding their travel mode, route, and time of departure (Steed & Bhat, 2000). These decisions are mainly influenced by three types of factors: user-related, trip-related, and service-related. User-related factors include cognitive (such as gender and income) and affective (such as past experiences and the value of time) factors. Trip-related factors focus on the characteristics that define each trip, such as travel time and cost. Lastly, service-related factors pertain to the features that describe the services provided, such as safety and comfort (Gasparinatou et al., 2022). In this work, the interest is directed exclusively on the mode choice since decisions on users' main travel mode are directly affected by the implementation of network interventions and the introduction of new mobility alternatives (Gasparinatou et al., 2022).

## 1.1 Attainment of the objectives and explanation of deviations

As per the Grant Agreement (GA), there are two objectives of WP3 that should be addressed and are relevant for this deliverable:

*Objective B: Development of a set of models to describe travel choices and mobility shift in both an aggregated and in-depth level.*

This objective is addressed by the development of parametric and machine learning models for the description and quantification of travel choices and mobility shifts. The developed models have been trained on stated preference survey data from each of TANGENT's pilot cities. Activities and models related to the in-depth analysis of travel behaviour based on the revealed preferences of users will be the subject of Deliverable D3.3 "In-depth analysis of travel behaviour" (M28).

*Objective C: Identification of factors that affect behavioural shift under various traffic management measures (strategies) and unexpected events.*

These issues have been addressed in both model and survey design. Specifically, all traffic management strategies of TANGENT are approached by specific survey questions eliciting, among others, user acceptance and preferences with respect to each traffic management strategy, user behaviour in the case of unexpected events etc. Parametric and Machine Learning models are trained upon this information and are able to accurately predict mode choice when the transportation network is affected by the above eventualities.

Overall, the objectives related to this deliverable have been achieved in full and as scheduled.

## 1.2 Intended audience

This deliverable is public, thus accessible not only to TANGENT partners and project officers, but to anyone interested in travel behaviour analysis, mobility patterns and travel behaviour shifts that may occur when new traffic management measures are implemented in the road network.

## 1.3 Structure of the deliverable and links with other work packages/deliverables

This document outlines the process and methodology for creating travel choice models using stated preference data collected through a user questionnaire survey. Chapter 2 provides a thorough explanation of the questionnaire's structure, its integration with TANGENTS's services and use cases, its dissemination strategy, and insights on the collected sample. The second part of the chapter (Subsection 2.5) focuses on the development of mode choice models, discussing the similarities and differences between parametric and machine learning models and detailing the model development pipeline.

In Chapter 3, the models' findings are presented, and a comparative assessment of the models is provided.

Chapter 4 outlines the implementation of the models within the TANGENT solution framework, describing the inputs and outputs of the mode choice module. Finally, Chapter 5 includes the conclusions of this deliverable.

This deliverable is linked to several work packages and deliverables. For the development of travel choice models, input was required from the case studies (WP7) to further define which TANGENT services and traffic management strategies are relevant to each pilot, leading to the development of city-specific versions of the travel choice questionnaires. In addition, both the survey and models address the particularities (from a mode choice perspective) of the traffic management strategies developed within WP5, described in detail in Deliverable 5.2 “Optimization Models for Transport Network Management”.

In addition, this deliverable provides input for the following WPs:

- WP4: The probability of choosing each travel mode will serve as a key input for forecasting long term demand shifts
- WP6: The models will be integrated into the overall TANGENT solution to be developed.

# 2 Methodological approach

The main goal of this deliverable is to develop discrete choice models that can describe how travel mode decisions are made and then, to develop a travel choice module that can predict the probability of choosing each of the available travel modes for every population in each of the four city-pilots. In order to collect the necessary data for this task a thorough travel behaviour survey was designed and executed in Athens, Lisbon, Manchester and Rennes.

The survey builds on the principles of the Stated Preference technique, one of the most frequently utilized methods for gathering information on travel behaviour (Kroes & Sheldon, 1988). Stated preference surveys have the desirable property of combining a quick and cost-effective data collection with the ex-post examination of the impact of alternative configurations of the transport network after the introduction of new modes, infrastructure or policies. The elicited information encompasses sociodemographic traits, travel patterns, perceptions of system-level attributes, and preferences for different travel options as indicated by specific scenarios. More precisely, a site-specific questionnaire survey was designed meaning that a dedicated questionnaire version was developed for each pilot city to meet the mobility characteristics and network conditions of each city. The survey was conducted through the CONEY toolkit which enables a user-friendly data collection. A further description of CONEY is included in Section 2.4.

During the data collection process, a sample monitoring procedure was followed to ensure the effectiveness of data collection as well as the representativeness of the sample. At a further stage, corrective actions will be performed wherever needed to improve sample characteristics (see Section 2.4).

Further, a pre-processing phase took place where data were cleaned and codified to support the modelling process. Two types of models were developed corresponding to a more traditional yet interpretable one (Multinomial Logit Model) and a more advanced machine learning model, namely Random Forest. Finally, the models were compared in terms of accuracy on their predictions and significant results were drawn regarding mode choice decisions (Section 3).

The above-described methodology approach from survey design to model development is depicted in Figure 1.

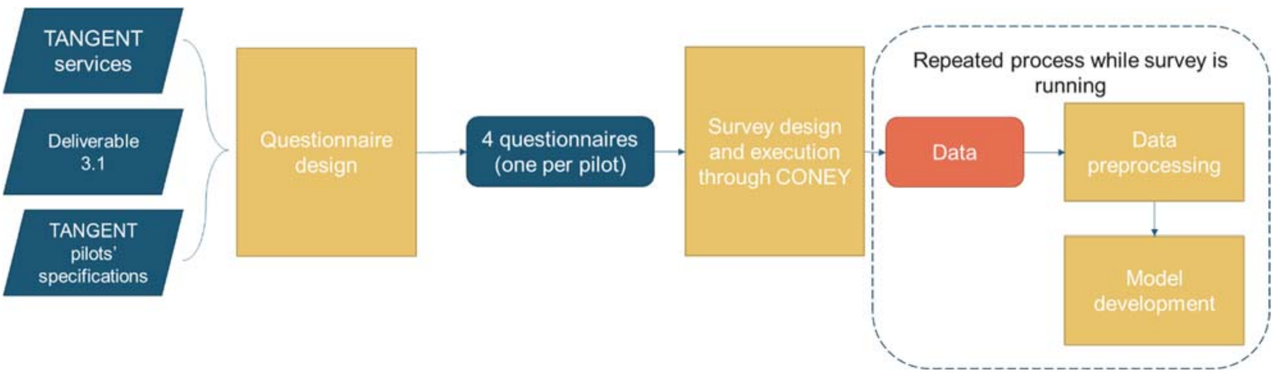


*Figure 1: Travel choice modelling pipeline*

## 2.1 Questionnaire development

The aim of this research is to understand commuters' travel patterns and perceptions in Athens, Lisbon, Manchester and Rennes. To achieve this, a questionnaire has been developed, consisting of four sections and around 30 questions, including stated preference scenarios. Each questionnaire includes questions regarding the mobility profile of the traveller, the travellers' perception on different system

specifications, their reactions to unexpected events, stated preference scenarios, as well as sociodemographic characteristics.
Four city-specific versions of the questionnaire have been developed and translated to the local language, as each city is interested only in specific functionalities and services of TANGENT. The transport network of each city also has specific characteristics, such as different topology and costs, which affect the stated preference scenarios.

Despite the variations between the city-specific versions, the overall questionnaire structure remains the same. All four versions of the questionnaire can be found in the Annex, while each section of the questionnaire is described briefly below.

**Section A: Mobility Profile**

This section aims at detecting the current mobility profile of each traveller. For that, each respondent is requested to provide information regarding their daily commute patterns such as their most frequent trip purpose and mode, whether they travel during peak hours or not, flexibility with respect to the arrival time on their destination, cost). They are also requested to provide an approximation of the number of weekly trips for different purposes (work, educational, personal, leisure etc.).

**Section B: Perception of users**

In this section, respondents are asked about their perception towards a variety of traffic management strategies and new services that aim at improving the network conditions. It is divided in 5 subsections:

*B1. Current Network Condition*
In this subsection, users express their satisfaction with the current transport network conditions by ranking selected public transport (PT) attributes from worst to best and by providing a quantitative perception of how much traffic congestion affects their everyday trips.

*B2. Pricing Schemes*
This part elicits respondents' willingness to accept the implementation of urban tolls. In addition, users are requested to provide input on how they will react in case such a system is implemented.

B3. *On-Demand Services*
Respondents state their willingness to use an on-demand service and proceed in selecting their preferred on-demand service business model from three available options: a first-last mile service to synchronise transfer between on-demand and regular public transit services, a fixed-stops service, and a door-to-door service.

*B4. Synchronization Of Traffic And PT*
Similar to the above, respondents express their willingness to use public transport means for their everyday trips if priority would be given to it through traffic lights, and rank various objectives related to the implementation of the proposed system.

*B5. Response To Unexpected Events*
Respondents state the likelihood of changing their usual travel mode in the case of various planned or unplanned system disruptions, by stating their maximum acceptable waiting time for critical and non-critical trips, their usual response to public transport service disruption, as well as their perception of the information channels that the PT service offers to passengers.

**Section C: Stated Preference Scenarios**

This part of the questionnaire formalises the stated preference scenarios for the estimation of travel choice probability. This technique presents hypothetical scenarios and asks respondents to choose their preferred transportation mode among the alternatives. The two distinct scenarios that have been developed are presented below.

The first scenario ("Synchronisation") presents users with the case of a typical trip from the suburbs to the city centre where the two available transportation modes are the private car and public transport. We are also considering that public transport modes are in synchronisation with traffic control (i.e. traffic lights), allowing for the improvement of the average speed of buses. This scenario is closely related to the *Synchronization of Public Transport and Traffic Control* optimization problem of Service 2 of TANGENT, presented in Section 2.2. In the case of Athens, the scenario description is as follows:

> ***In order to perform a typical trip from the suburbs to the centre of Athens, with length of approximately 10 km, during peak hours, the available modes are car and Public Transport. To improve traffic conditions traffic light management strategies that give priority to Public Transport are applied.***

The scenario presents respondents with a set of six hypothetical alternatives, for each of which users must select their preferred option from a choice of two. Each option is differentiated with respect to the mode (car or public transport), cost, and (in-vehicle and waiting) time. An example is shown in Figure 2.

| | |
|---|---|
| Option 1<br>Mode: Car<br>In-vehicle time: 65 minutes<br>Waiting time: -<br>Cost: 5€ | Option 2<br>Mode: Public Transport<br>In-vehicle time: 55 minutes<br>Waiting time: 15 minutes<br>Cost: 2€ |
| ○ | ○ |

*Figure 2: Available options for Stated Preference Scenario A "Synchronisation"*

In the second scenario ("Pricing"), respondents evaluate a similar case where the available transport modes are the car and public transport. In this case, we also consider that a congestion pricing scheme for the entrance in the city centre is also implemented, so commuters who choose the private car have the additional option of paying a toll for their entrance to the city centre or bypassing the tolled area altogether. This scenario is closely related to the *Dynamic Congestion Pricing* optimization problem of Service 2 of TANGENT, presented in Section 2.2. In the case of Athens, the scenario description is as follows:

> ***"For a typical trip from the suburbs to a central neighbourhood of Athens during morning peak hours, commuters can choose between using their private car or opting for the public transport system. To improve traffic conditions, a congestion pricing scheme is applied: car users who wish to pass through the city centre must pay a toll. Commuters who are not willing to pay the toll but nevertheless want to continue using their car can bypass the tolled area and use peripheral roads to reach their destination."***

The scenario presents respondents with a set of six hypothetical alternatives, for each of which users must select their preferred option from a choice of three. Each option is differentiated with respect to the mode (car with toll payment, car without toll payment or public transport), cost, distance, and time. An example is shown in Figure 3.

| | | |
|---|---|---|
| Option 1<br>Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 30 minutes<br>Cost: 3.5€ | Option 2<br>Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 35 minutes<br>Cost: 2€ | Option 3<br>Mode: Public Transport<br>Time: 45 minutes<br>Cost: 0.60€ |
| ○ | ○ | ○ |

*Figure 3: Available options for Stated Preference Scenario B "Pricing"*

**Section D: Demographic characteristics**.

In the final section of the questionnaire, the users are requested to provide personal information regarding their gender, age, annual personal income, occupation, car ownership and place of residence (within the city centre or not). In addition, they provided a self-categorization based on their usual transport mode choice (eco-traveller, time-saver, technology-enthusiast or car-dependent), a further explanation of which is included in the Annex. Respondents who wish to participate in the second phase of the survey, have the option to share their contact information (email) for future communication with the TANGENT team. The second phase of the survey is related to the elicitation of user revealed preferences and will provide the necessary data (trip trajectories from Google Timeline) for the training of in-depth travel behaviour models developed within Task 3.4 "In-depth analysis of travel behaviour". This part of the survey and the created models will be presented in Deliverable D3.4 "In-depth analysis of travel behaviour".

## 2.2 Survey to TANGENT functionalities mapping

The TANGENT project will implement an architecture that is functionally comprised of three distinct Services:

- **Service 1** - Enhanced Information Service for Multimodal Transport Management
- **Service 2** - Real-Time Traffic Management Services
- **Service 3** - Transport Network Optimisation for Transport Authorities

The three services are described in detail in Deliverable 1.2 'NTM needs assessment and system requirements' and their main functionalities are reproduced here:

- **Service 1** - Enhanced Information Service for Multimodal Transport Management
  - Visualization of the current and future (predicted) status of the network (including incidents, conditions, delays, etc.) in the TANGENT Dashboard
  - Collection and delivery of data from/to the TANGENT API regarding current and future (predicted) status of the network (including incidents, conditions, delays, etc.)
- **Service 2** - Real-Time Traffic Management Services
  - Cooperative Incident Management, supporting the:
    - Synchronization of On-Demand Transport and Public Transport
    - Synchronization of Traffic Control and Public Transport
  - Smart Network Load Balancing, supporting:
    - Adaptive traffic control and Connected and Automated Vehicles
    - Dynamic Congestion Pricing
    - Informing the transport passengers
- **Service 3** - Transport Network Optimisation for Transport Authorities
  - Common Operational Picture (COP) generation support
  - Response Plan generation support
  - Simulation of predefined scenarios

A further description of the Service 2 components that are related to the optimization problems that will be addressed within TANGENT is provided below. The interested reader is referred to Deliverable 5.1 “Analysis of Current Approaches in Optimization of Transport Network Management”.

- **Synchronization of Demand Responsive Traffic and Traffic Control**
  The objective of this problem is to find the optimal way to synchronize demand responsive traffic with public transport. In scenario where a user can initiate a travel request with feasible travel time windows for a given OD pair the objective of this problem is to find the nearest feasible PT stop for the origin and destination which adheres to the user given time windows and to schedule a pick up and drop service for the first mile (from user location to nearest feasible PT stop chosen) and last mile (from selected PT stop to the user destined destination) of the travel.

- **Synchronization of Public Transport and Traffic Control:** The objective of this section is to describe the formulation of the synchronization of public transport and traffic control problem. The main goal of this problem is to find the optimal traffic control plan and line-specific public transport plans (in terms of frequency and capacity) so that the user costs of the affected passengers and the operator costs of the optimized public transport lines both are reduced along with the systemwide pollution and congestion at selected axes of the network.

- **Dynamic Congestion Pricing**. In the DCP optimization problem the goal is to come up with an optimal pricing policy for the selected traffic zone so that the congestion at selected section of the network inside and outside the zone of interest should be reduced alongside the system wide emission. In the optimization problem our goal is also to maximize the toll revenue collection from the pricing scheme implementation.

- **Signalized Vehicle Couple Control with CAVs.** This section aims to formulate the signal vehicle coupled control problem. A fleet of fully Connected and Autonomous Vehicles (CAVs), which are connected with the Traffic Management Centre (TMC) and can be managed by the TMC (Centralized control). The origin and destination of each CAV is known by the TMC and between each origin and destination there are one or more pre-defined paths. In this problem, we have considered mixed traffic scenarios. The objective is to optimize the path assignment for the complete CAV fleet (for each OD-pair and path, define the percentage of CAVs that take that path) and the traffic control plan on specific arterial of the considered network according to different performance metrics that we will define below.

The survey and the corresponding models are conceived as an integral part of Service 2, by providing accurate mode choice predictions when the four traffic management strategies of Service 2 are applied. A mapping of Service 2 components to the survey’s sections is presented in Table 1.

| | | **Service 2 - Real-Time Traffic Management Services** | | | | |
|---|---|---|---|---|---|---|
| | | Cooperative Incident Management | | Smart Network Load Balancing | | |
| | | Synchronization of On-Demand Transport and Public Transport | Synchronization of Traffic Control and Public Transport | Adaptive traffic control and Connected and Automated Vehicles | Dynamic Congestion Pricing | Informing the transport passengers |
| **Section A: Mobility Profile** | | ✓ | ✓ | ✓ | ✓ | ✓ |
| | *B1. Current Network Conditions* | ✓ | ✓ | ✓ | ✓ | ✓ |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Section B: Perception of users** | *B2. Pricing Schemes* | | | | ✓ | |
| | B3. *On-Demand Services* | ✓ | | | | |
| | *B4. Synchronization Of Traffic And PT* | | ✓ | | | |
| | *B5. Response To Unexpected Events* | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Section C: Stated Preference Scenarios** | *Scenario A* | | ✓ | | | |
| | *Scenario B* | | | | ✓ | |
| **Section D: Demographic characteristics.** | | ✓ | ✓ | ✓ | ✓ | ✓ |

*Table 1: Survey to TANGENT functionalities mapping*

In addition to this mapping, the survey and the survey-trained models also support mode choice predictions during the occurrence of expected or unexpected events. Both cases are in the core of TANGENT's functionalities and scope and thus, they are relevant to all TANGENT Services.

## 2.3 Survey to TANGENT Use Cases Mapping

The questionnaire was site-specific, meaning that each of the pilot cities has their own version of the questionnaire where the content is modified accordingly based on the system characteristics as well as the type of TANGENT solutions that will be tested in the pilot. For each city hosting a pilot, the content and processes of the tests were defined in a testing protocol describing the Case Studies and modes to be tested. This work is described in detail in Derivable 7.1 "Definition of Case Studies, stakeholders' engagement and tests preparation" and Deliverable 7.2 "Operation of the tests: testing of subsystems". In this task, the pilot city partners assessed and selected the most suitable traffic management strategies for each case study. This process enabled customization of the questionnaire for each city by eliminating irrelevant questionnaire parts and thus, reducing the time required to complete the survey.

More specifically, partners from Lisbon, Manchester and Rennes have expressed interest in testing exclusively the *Synchronization of Public Transport and Traffic Control* traffic management strategy. This strategy is intrinsically linked to the particulates of the transport network in each of the three cities and aims to solve specific problems in each of them.

In more detail, Greater Manchester's bus services are currently provided by private operators (e.g. Stagecoach, Go North West etc.), thus making it not currently operationally feasible for Transport for Greater Manchester (TfGM) to facilitate this coordination. It is therefore of particular interest to TfGM to assess the potential benefits of this functionality through TANGENT, with this bus operator arrangement set to change over the next few years through Bus Franchising (TfGM is in the process of taking back control of bus services).

In Lisbon, synchronizing public transport and traffic control has a significant added value for the Lisbon case study, enabling not only the evaluation of the outcomes of distinct traffic management strategies, but also providing a visual and analytical tool to help foster stakeholder and decision maker engagement.

In Rennes, the traffic is controlled by PC Arthur, the Traffic Control centre, and the public transport services are operated by a private bus company, Keolis. Given this fragmentation, TANGENT can provide a significant improvement for the stakeholders, allowing them to implement joint and coordinated actions.

Based on the aforementioned considerations, we have eliminated survey sections that are only applicable to irrelevant traffic management strategies in all three cities (Table 2). On the other hand, given that the Athens use case is virtual and aims at the testing of the entire TANGENT solution, all four of the traffic management strategies have been addressed resulting in the most comprehensive questionnaire of the four.

| | | Athens | Lisbon | Manchester | Rennes |
|---|---|---|---|---|---|
| **Section A: Mobility Profile** | | ✓ | ✓ | ✓ | ✓ |
| **Section B: Perception of users** | *B1. Current Network Conditions* | ✓ | ✓ | ✓ | ✓ |
| | *B2. Pricing Schemes* | ✓ | | | |
| | B3. *On-Demand Services* | ✓ | | | ✓ |
| | *B4. Synchronization Of Traffic And PT* | ✓ | ✓ | ✓ | ✓ |
| | *B5. Response To Unexpected Events* | ✓ | ✓ | ✓ | ✓ |
| **Section C: Stated Preference Scenarios** | *Scenario A: Synchronization* | ✓ | ✓ | ✓ | ✓ |
| | *Scenario B: Pricing* | ✓ | | | |
| **Section D: Demographic characteristics.** | | ✓ | ✓ | ✓ | ✓ |

*Table 2: Survey to TANGENT Use Cases Mapping*

In addition to the above, each city's transport network characteristics have also been taken into account in the creation of the stated preference scenarios (Survey Part C). The transport network of each city shows different topology and service characteristics which affect the distance, travel time, waiting time and cost for a trip to the city centre. As such, city-specific values of the above values have been calculated and included in the questionnaire, aiming at presenting stakeholders with more realistic options and increasing the probability of eliciting realistic responses. Finally, the questionnaires have been translated into four languages: English, French, Greek and Portuguese. Respondents in each city were provided with the choice to answer in either the local language or in English.

## 2.4 Data collection

### 2.4.1 Dissemination strategy

The TANGENT survey was conducted as a comprehensive study aimed at gathering information from a required sample of 1000 participants for each case study in all four cities involved. The initial duration of the survey was set at four months, starting on 10 October 2022. To reach the predefined response target, the duration of the survey has been extended and it is expected to continue until October 2023.

The survey was executed using CONEY, an innovative toolkit designed by Cefriel to enhance the user experience in surveys completion. CONEY exploits a conversational approach: on the one hand, CONEY allows modelling a conversational survey with an intuitive graphical editor (Figure 4); on the other hand, it allows publishing and administering surveys through a chat interface. For a further description CONEY, the reader is referred to the CONEY toolkit website at https://coneytoolkit.cefriel.it/

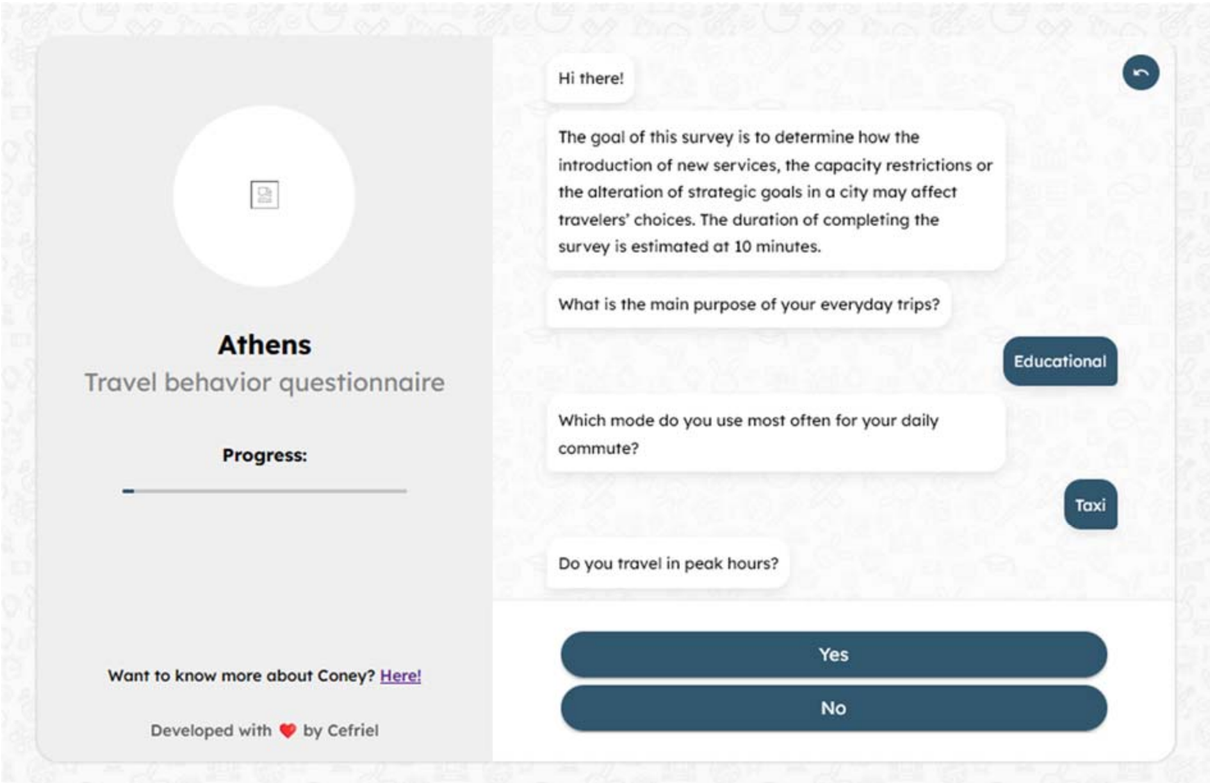


*Figure 4: The chat-based interface of CONEY*

To disseminate the survey, a combination of online and on-site approaches was utilized. The online methods involved the use of social media platforms, the TANGENT website, newsletter, mailing lists, and personal contacts, which enabled the project team to reach a broader audience and encourage participation in the survey. In certain cities, the acquisition of on-site data has been accomplished through the dissemination of the questionnaire on the field (e.g., in large multimodal exchange hubs or in frequented bus stops), or during scientific conferences where TANGENT was represented, such as the Transport Research Arena (TRA) 2022 conference, held in Lisbon.

All partners have supported the dissemination of the survey while the pilot leaders are responsible for reaching the predefined target. To ensure a successful process, the WP3 team regularly provided partners with updated statistical characteristics of the sample collected so far and gave feedback on potential statistical imbalances and underrepresented user categories that may have been present in the sample. An overview of the dissemination characteristics of the survey is included in Table 3.

| | |
|---|---|
| **Required Sample** | 1000 per case study |
| **Cities involved** | All four case studies (Athens, Lisbon, Manchester, Rennes) |
| **Survey duration** | Initial: 4 months – continuing to receive responses until October 2023 |
| **Start date** | 10 October 2022 |
| **Dissemination methods** | On-line (social media, TANGENT website & newsletter, mailing lists, personal contacts) & on-site (fieldwork, stands in events) |
| **Survey form** | CONEY Toolkit |

*Table 3: Survey dissemination overview*

### 2.4.2 Sample identity

The following section presents a comprehensive summary of the data obtained from the questionnaire survey conducted across all four cities. Furthermore, an initial analysis is carried out to compare the responses of the sets of participants.

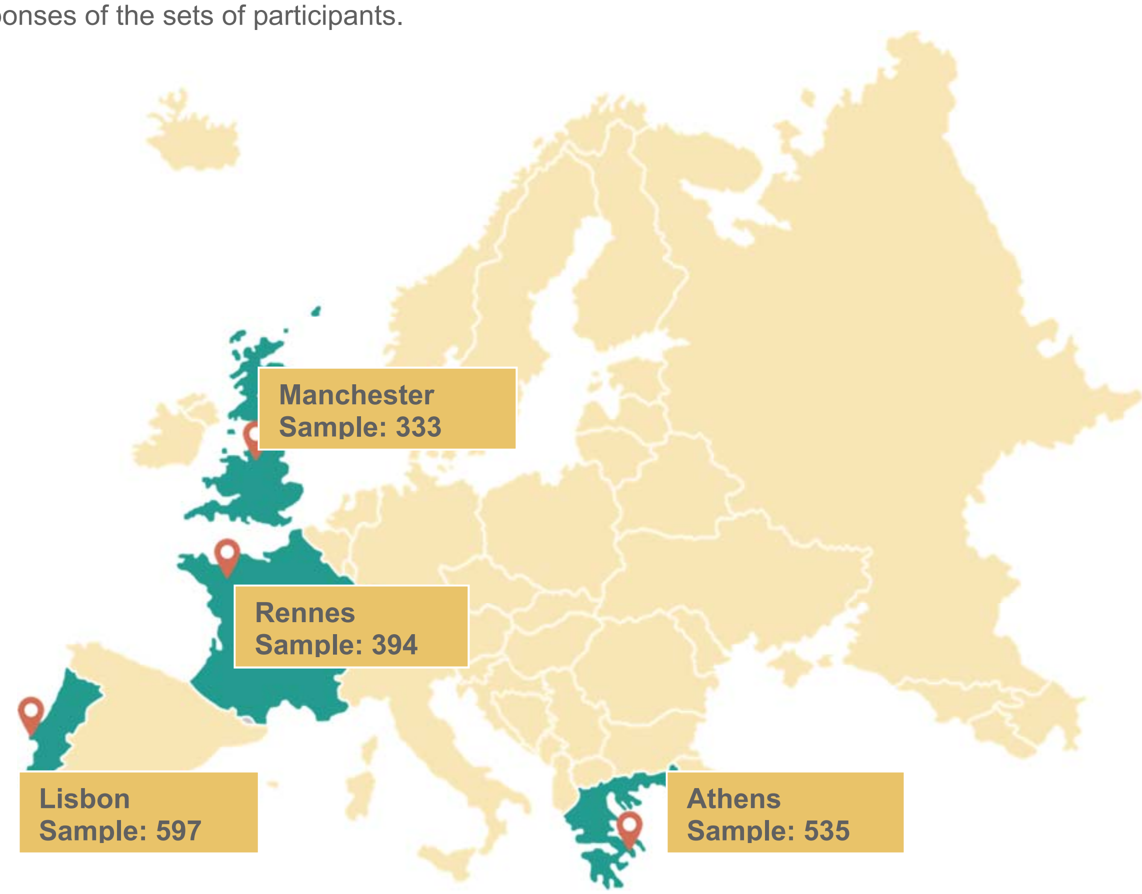


*Figure 5: Collected samples per city*

Up until M20, a total of 1859 valid (excluding incomplete and fault answers) responses were collected for the four pilot cities through the questionnaire survey prepared within the framework of WP3. Specifically, 535 valid responses were collected for the case of Athens, 597 for Lisbon, 333 for Manchester and 394 for Rennes.

The overall distribution of the collected sample with respect to gender and age is included in Table 4.

In the table, "Gender" refers to the distribution of participants in the sample according to their gender. The second variable presented in the table is "Age," which shows the distribution of participants across six age groups. For both variables, the number in the paretheses represents the actual population distribution with respect to gender and age, according to the latest population census data in each of the four cities.

| Sample Variables | | Athens | Lisbon | Manchester | Rennes |
|---|---|---|---|---|---|
| **Gender** | Male | 53.9% (48.2%) | 56.8% (46.3 %) | 55.0% (50.3%) | 40.6% (48.3%) |
| | Female | 46.1% (51.8%) | 43.2% (53.7%) | 45.0% (49.7%) | 59.4% (51.7%) |
| **Age** | 18-24 | 18.9% (9.5 %) | 22.1% (6.0%) | 24.9% (20.1%) | 22.6% (13.7%) |
| | 25-34 | 34.6% (18.7%) | 24.5% (16.7%) | 26.1% (23.9%) | 22.1% (19.1%) |
| | 35-44 | 15.5% (19.4%) | 22.8% (17.2%) | 23.1% (18.5%) | 19.5% (16.3%) |
| | 45-54 | 19.1% (17.0%) | 18.3% (16.6%) | 15.9% (14.6%) | 19.5% (16.3%) |
| | 55-64 | 10.8% (14.5%) | 9.0% (14.4%) | 7.8% (11.5%) | 11.4% (14.0%) |
| | 65+ | 1.1% (20.9%) | 3.4% (29.0%) | 2.1% (11.5%) | 4.8% (20.5%) |

*Table 4: Demographics of the survey sample together with the latest population census data for Rennes (INSEE, 2019), Lisbon (INE, 2021), Manchester (ONS, 2021) and Athens (ELSTAT, 2011).*

Based on the information provided in the table and expert judgement, the sample is well-distributed across the variables of gender and age in all cities. This is important because a well-distributed sample helps to ensure that the results of the survey are representative of the population being studied.
Regarding gender distribution, the table shows that there is almost a 50-50 split between men and women in all cities, with a slight overrepresentation of women in Rennes.

The age distribution in the sample is also well distributed, with a shift towards younger ages in all cities. We note that the population over 65 is underrepresented in all cities, which is a common issue in similar surveys as this population group tends to commute less frequently.

The overall distribution of the collected sample with respect to travel purpose and mode choice is included in Table 5**¡Error! No se encuentra el origen de la referencia.**. In this table, “Travel Purpose” refers to the distribution of participants in the sample according to their usual daily travel purpose, divided into four categories: professional, educational, leisure (visiting the gym or friends, other social reasons etc.) and personal (medical, family, housework reasons etc.). Similarly, “Mode Choice” refers to the distribution of users mode choices into car, public transport, walking/bicycling, motorcycle or taxi. It is noted that in Rennes, the Walking/Bicycling mode category was divided into two distinct options: Cycling and EPDMs (personal motorized mobility solutions) & Walking, which provides more detailed information on how people travel in Rennes and is compatible with past studies conducted in the city, such as the latest household travel survey of 2018[1].

| Sample Variables | Athens | Lisbon | Manchester | Rennes |
|---|---|---|---|---|
| Professional | 74.1% | 62.8% | 51.2% | 52.3% |

[1] Ménages Déplacements (EMD) en Ille-et-Vilaine (2018), available in https://tinyurl.com/3strk8y7

| | | | | | |
|---|---|---|---|---|---|
| **Travel purpose** | Educational | 11.4% | 20.9% | 20.6% | 24.4% |
| | Leisure | 8.4% | 8.2% | 17.2% | 9.6% |
| | Personal | 6.0% | 8.0% | 11.0% | 13.7% |
| **Mode choice** | Car | 49.2% | 21.1% | 35.6% | 34.3% |
| | Public transport | 36.0% | 71.5% | 42.0% | 38.3% |
| | Walking/Bicycling* | 7.9% | 5.5% | 21.5% | - |
| | Motorcycle | 5.6% | 1.7% | 0.3% | 1.5% |
| | Taxi | 1.3% | 0.2% | 0.6% | 0.8% |
| | Cycling and EPDMs (personal motorized mobility solutions) * | - | - | - | 14.0% |
| | Walking* | - | - | - | 11.2% |

*Table 5: Sample distribution with respect to travel purpose and mode choice for Athens, Lisbon, Manchester and Rennes.*

In terms of travel purpose, professional purposes dominate the samples in all cities, followed by educational and leisure/personal purposes. Athens has a slightly higher percentage of participants who reported that they mainly commute for professional purposes. Finally, the mode choice data shows that car and public transport dominate the observed modal split in all cities. In addition, public transport in Lisbon seems to be relatively overrepresented, as is walking/bicycling in Manchester.

Overall, the sample is well distributed, which should help to ensure that the results of the survey are representative of the populations being studied. The few issues that have been identified were communicated to the contact persons for each case study, and corrective measures will be evaluated to target more specific user categories in the rest of the sample collection, to be included in the second version of this deliverable.

## 2.5 Modelling Strategy

### 2.5.1 Model development

The objective of this study is to develop a travel choice model by comparing parametric methods and machine learning models. The aim is to identify the most reliable model for each scenario and city being considered. Developing the models includes 3 major steps: data pre-processing, statistical analysis and model development. The process in described schematically in Figure 6.

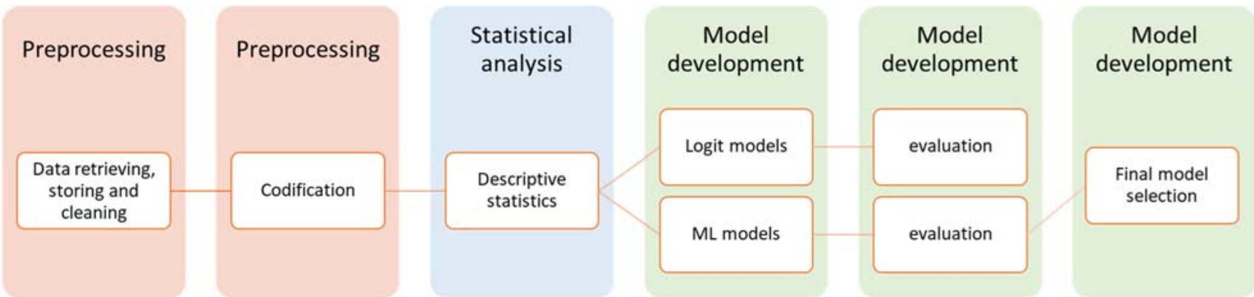


*Figure 6: Model development pipeline*

Data pre-processing is an essential step in the model development pipeline ensuring the quality and consistency of the data used for model training and testing. The first step is data retrieval, which involves exporting the data from CONEY for each of the case studies and storing it in a common database.

The next step is data cleaning, which refers to a set of techniques for identifying and dealing with various issues that may affect the quality of the data, such as missing values, duplicates, and outliers. Data cleaning also involves transforming the data into a standardized format, such as normalizing or scaling the values to ensure that they fall within a specific range. The appropriate implementation of these techniques can lead to the reduction of bias and improvement of accuracy in the developed models.

The final step in data pre-processing is data codification, which entails converting categorical data into a numerical format. Common methods of encoding categorical data include one-hot encoding, label encoding, and binary encoding. The choice of encoding method depends on the type of categorical data, the number of categories, and the analysis requirements.

The next step is to use descriptive statistic techniques in order to gain a better understanding of the data, monitor the sample and steer data collection efforts. By analysing the data collected so far, irregular patterns or anomalies in the data can be identified, which can help to detect potential errors or biases in the survey design or administration. This information can be used to adjust the survey design or steer data collection efforts by identifying specific respondent categories that require more attention. For example, if the analysis reveals that a particular demographic group has a low response rate, targeted outreach efforts can be implemented to increase their participation in the survey.

After this step, the collected data were utilized to train parametric (Multinomial Logistic Regression) and non-parametric models (Random Forest).

Traditional statistical techniques such as the Multinomial Logistic (Logit) regression are parametric models that assume a specific functional form for the relationship between the predictors and the response variable. These models have been extensively applied in mode choice modelling and are generally easy to interpret and provide estimates of the significance of each predictor. However, they have limited ability to capture complex and nonlinear relationships in the data.

On the other hand, Machine Learning techniques such as Random Forests, are non-parametric models that can capture complex and nonlinear relationships between predictors and response variables. These models are generally more accurate and can detect hidden patterns in the data. However, they may lack interpretability (often described as "black box" models), and it can be difficult to decipher how the model arrived at its predictions. Deliverable 3.1 provides an in-depth reference for the extensive application of both model families in predicting transport mode choice, for readers interested in further exploration of the topic.

The final step is the evaluation of the Logit and Machine Learning models. This will be carried out using a range of metrics, described in more detail in Section 3. Models that perform the best in this task will be selected as the optimal models and used for prediction purposes. The optimal models are expected to exhibit a high degree of generalizability to new data and demonstrate accurate prediction capabilities.

### 2.5.2 Multinomial Logit Model

The Multinomial Logit (MNL) model is a commonly used statistical model in social sciences, economics, and marketing research to analyse and predict the choices of individuals from a set of available options. The model is a type of regression model that predicts the probability of an individual choosing one option among several alternatives, based on a set of predictor variables. It assumes that the choices made by individuals are influenced by their preferences for the available options, as well as the characteristics of the individual and the options. The model is based on the idea that individuals choose the option that maximizes their utility, which is a function of the characteristics of the options and the individual's preferences. The Multinomial Logit model is a type of generalized linear model (GLM), which means that it assumes a linear relationship between the predictor variables and the log odds of the outcome. The model is called multinomial because it can handle more than two possible outcomes, each with its own set of predictors. It is typically used when the dependent variable is categorical and has three or more categories. Also, the model is estimated using maximum likelihood estimation, which involves finding the set of model parameters that maximize the likelihood of the observed data. The likelihood function is a function of the model parameters that describes how likely the observed data is to occur

given the model. Finally, the model assumes that the errors are independent and identically distributed, and that the variance of the errors is constant across all options, while the predictor variables are linearly related to the log odds of the outcome.
Assuming that we have a categorical dependent variable $Y$ with $J$ categories (options), $i$ predictor variables $Xi$, and $\beta$ being the set of $i$ unknown parameters to be estimated, the Multinomial Logit model assumes that the probability of observing category $j$ given the predictor variables is:

$$\Pr(yi|Xi) = \frac{\exp(Xi\beta j)}{\sum_{m=1}^{J} exp(Xi\beta m)} \qquad (1)$$

In this equation, the numerator represents the probability of observing category $j$ given the predictor variables, while the denominator is the sum of probabilities for all categories (including the reference category). Note that the denominator ensures that the probabilities for all categories add up to 1.

### 2.5.3 Random Forest Classifier

The Random Forest Classifier is a popular machine learning algorithm used for classification tasks. It belongs to the family of ensemble learning algorithms, which means it combines multiple decision trees to create a more robust and accurate model. The algorithm uses the bagging technique, which involves creating multiple decision trees using bootstrap samples of the training data. The basic idea behind the Random Forest Classifier algorithm is to create a large number of decision trees, each of which is trained on a random subset of the training data and a random subset of the predictor variables. This randomness helps to reduce overfitting and improve the accuracy of the model. The algorithm then uses these decision trees to predict the class of a new observation by taking a majority vote of the predictions made by the individual trees. The Random Forest Classifier algorithm can be used for both binary and multiclass classification problems. In the case of binary classification, the algorithm creates a single decision tree that predicts the probability of the positive class. In the case of multiclass classification, the algorithm creates multiple decision trees, each of which predicts the probability of one of the classes. The algorithm then selects the class with the highest probability as the predicted class.

# 3 Findings

Following the collection of the questionnaires, the codification of all the answers given by the participants was implemented in a Python environment. More specifically, the whole process of codification has been automated for every new data-answer input. The codification was implemented for the city of Athens, Lisbon, Manchester and Rennes.

Subsequently, feature engineering techniques were applied to both statistical and machine learning methods employed. Firstly, one-hot encoding technique was applied during codification by transforming the categorical variables to numerical. Also feature selection was applied by identifying the most relevant and informative variables that influence an individual's travel choices for every use case. As a result, a set of input variables were identified and selected, which when combined, yield the optimal prediction of travel choice selection of an individual for every city. The results of each model developed and significance of the variables for every use case are presented in the following chapters.

Table 6 describes the variables used to create the statistical and machine learning models for the cities considered.

| Input Variables | Description | Type | Unit |
|---|---|---|---|
| **IN_VEH_TIME_CAR** | In vehicle time of car | Integer | min |
| **IN_VEH_TIME_PT** | In vehicle time of Public Transport | Integer | min |
| **WAIT_TIME_PT** | Waiting time for Public Transport | Integer | min |
| **COST_CAR** | Cost of trip for Car | Integer | Euro |
| **COST_PT** | Cost of trip for Public Transport | Integer | Euro |
| **AGE** | Age of individual | Multiple choice [18-24, 25-34, 35-44, 45-54, 55-64, >65] | - |
| **INCOME** | Income of individual | 3-point scale [Low, Medium, High] | - |
| **PTN_SATIS** | Satisfaction with the level of service of Public Transport | 5-point scale [Not satisfied at all - Very satisfied] | - |
| **PEAK_HOUR** | Travelling on peak hours | Binary [Yes-No] | - |

*Table 6: Input variables for travel choice models (Prioritization to Public Transport)*

## 3.1 Travel choice model under prioritization to Public Transport

In this subsection the results of the models regarding to the travel choice under prioritization to Public Transport are presented. More specifically, a Random Forest classifier and a Multinomial Logit model were developed. For both models the evaluation metrics of precision, recall and F1-score are presented. In classification problems, precision, recall, and F1-score can be extended to assess the performance of a model. The metrics can be calculated for each output label separately, and then aggregated to obtain a single score for the model. The precision for a label is the proportion of true positive predictions for that label among all predicted positive instances for that label. The recall for a label is the proportion of true positive predictions for that label among all actual positive instances for that label. The F1-score for a label is the harmonic mean of precision and recall for that label and is calculated as the ratio of true positives to the sum of true positives, false positives, and false negatives.

Additionally, for each input variable to the Multinomial Logit model the coefficients, standard errors, z-values, and p-values are presented while for the Random Forest classifier the feature importance of each input variable is presented. All the above results are shown for the city of Athens, Lisbon and Manchester.

The results of the city of Rennes were not included as it was considered appropriate to create an updated version of the questionnaire in order to improve the representativeness of the different scenarios examined. Following the collection of new replies to the questionnaires, the results of Rennes' survey will also be presented.

### 3.1.1 City of Athens

In Table 7 the results of these metrics for every model are presented. MLN indicates the multinomial Logit model while RF the Random Forest classifier model.

| Model comparison for Athens | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Class** | **Precision** | | **Recall** | | **F1-Score** | | **Number of Samples** |
| **Model** | MLN | RF | MLN | RF | MLN | RF | |
| **CAR** | 0.620 | 0.706 | 0.492 | 0.598 | 0.549 | 0.648 | 189 |
| **Public Transport** | 0.717 | 0.769 | 0.810 | 0.843 | 0.761 | 0.804 | 300 |

*Table 7: Model comparison for the city of Athens (Prioritization to Public Transport)*

The results indicate that the Random Forest Classifier outperforms the Multinomial Logit algorithm in terms of F1-score for both classes, suggesting that it is better at achieving a balance between precision and recall.

Table 8 presents estimates of the coefficients, standard errors, z-values, and associated p-values for each input variable included in the Multinomial Logit model that was developed for the city of Athens. The coefficients are utilized to determine the direction and magnitude of the relationship between each input variable and the outcome variable. It should be noted that the standard errors, z-values, and p-values of the coefficients provide valuable information about the statistical significance of each input variable in the model. Specifically, the p-values allow for an assessment of the probability that the estimated coefficients are the result of random chance, with lower p-values indicating greater statistical significance.

| Feature | Coefficient | Std.Err. | z-val | P > z |
|---|---|---|---|---|
| **_intercept.PT** | 0.199 | 0.265 | 0.750 | 0.453 |
| **AGE.PT** | -0.130 | 0.039 | -3.308 | 0.000*** |
| **INCOME.PT** | -0.618 | 0.079 | -7.776 | 0.000*** |
| **PTN_SATIS.PT** | 0.318 | 0.049 | 6.484 | 0.000*** |
| **PEAK_HOUR.PT** | -0.277 | 0.138 | -2.011 | 0.044* |
| **IN_VEH_TIME_PT** | -0.070 | 0.007 | -10.209 | 0.000*** |
| **WAIT_TIME_PT** | -0.100 | 0.011 | -9.076 | 0.000*** |
| **COST_PT** | -0.248 | 0.044 | -5.659 | 0.000*** |
| Sign. 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | |

*Table 8: Multinomial Logit model results for the city of Athens (Prioritization to Public Transport)*

Similar with the significance of each variable computed for the statistical model, also the feature importance is calculated for the Random Forest classifier. Feature importance is a method used in machine learning to identify which input features are most relevant to the target variable and in our instance of Random Forest classifier it is calculated using the Mean Decrease Impurity (MDI) method which computes the reduction in impurity for each feature over all trees in the forest (Alicioglu & Sun, 2022). Table 9 shows the results of feature importance of the model developed for the city of Athens.

| Feature | Description | Importance Score |
|---|---|---|
| AGE | Age of individual | 0.292 |
| PTN_SATIS | Satisfaction with the level of service of Public Transport | 0.249 |
| INCOME | Income of individual | 0.145 |
| IN_VEH_TIME_PT | In vehicle time for Public Transport | 0.110 |
| PEAK_HOUR | Travelling on peak hours | 0.066 |
| COST_CAR | Trip cost with Car | 0.040 |
| IN_VEH_TIME_CAR | In vehicle time for Car | 0.038 |
| WAIT_TIME_PT | Waiting time for Public Transport | 0.032 |
| COST_PT | Trip cost with Public Transport | 0.023 |

*Table 9: Feature importance of Random Forest Classifier for the city of Athens (Prioritization to Public Transport)*

### 3.1.2 City of Lisbon

The outcomes of Multinomial Logit model and Random Forest classifier for Lisbon city in terms of evaluation metrics are presented in Table 10.

| Model comparison for Lisbon | | | | | | | |
|---|---|---|---|---|---|---|---|
| Class | Precision | | Recall | | F1-Score | | Number of Samples |
| Model | MLN | RF | MLN | RF | MLN | RF | |
| CAR | 0.706 | 0.655 | 0.089 | 0.267 | 0.158 | 0.379 | 135 |
| Public Transport | 0.820 | 0.846 | 0.991 | 0.966 | 0.897 | 0.902 | 564 |

*Table 10: Model comparison for the city of Lisbon (Prioritization to Public Transport)*

The results show that the F1-score for the CAR class is very low for both algorithms, indicating that they are not good at correctly identifying this class. However, for the Public Transport class, both algorithms achieve high precision and recall values, indicating that they are good at correctly identifying it. Specifically, the Random Forest Classifier achieves a slightly higher F1-score for both classes, suggesting that it is better at achieving a balance between precision and recall. Therefore, it can be concluded that the Random Forest Classifier might be a better choice for this classification task, as it shows superior performance in terms of the F1-score for both classes.

Table 11 demonstrates the coefficients, standard errors, z-values, and corresponding p-values for each input variable integrated into a Multinomial Logit model designed for Lisbon city.

| Feature | Coefficient | Std.Err. | z-val | P > z |
|---|---|---|---|---|
| _intercept.PT | 0.986 | 0.199 | 4.950 | 0.000*** |
| AGE.PT | -0.075 | 0.034 | -2.227 | 0.026 * |
| INCOME.PT | -0.588 | 0.062 | -9.514 | 0.000 *** |
| PTN_SATIS.PT | 0.205 | 0.034 | 5.982 | 0.000 *** |
| IN_VEH_TIME_PT | -0.036 | 0.006 | -5.968 | 0.000 *** |
| WAIT_TIME_PT | -0.065 | 0.011 | -5.774 | 0.000 *** |
| COST_PT | -0.329 | 0.046 | -7.133 | 0.000 *** |

Sign. 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

*Table 11: Multinomial Logit model results for the city of Lisbon (Prioritization to Public Transport)*

Table 12 shows the results of feature importance of the model developed for the city of Lisbon.

| Feature | Description | Importance Score |
|---|---|---|
| **AGE** | Age of individual | 0.364 |
| **PTN_SATIS** | Satisfaction with the level of service of Public Transport | 0.278 |
| **INCOME** | Income of individual | 0.189 |
| **IN_VEH_TIME_CAR** | In vehicle time for Car | 0.051 |
| **WAIT_TIME_PT** | Waiting time for Public Transport | 0.037 |
| **IN_VEH_TIME_PT** | In vehicle time for Public Transport | 0.032 |
| **COST_CAR** | Trip cost with Car | 0.028 |
| **COST_PT** | Trip cost with Public Transport | 0.018 |

*Table 12: Feature importance of Random Forest Classifier for the city of Lisbon (Prioritization to Public Transport)*

### 3.1.3 City of Manchester

Table 13 displays the evaluation metrics of both models developed for Manchester city.

| Metrics comparison for Manchester | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Class** | **Precision** | | **Recall** | | **F1-Score** | | **Number of Samples** |
| **Model** | MLN | RF | MLN | RF | MLN | RF | |
| **CAR** | 0.623 | 0.743 | 0.611 | 0.694 | 0.617 | 0.718 | 108 |
| **Public Transport** | 0.760 | 0.817 | 0.769 | 0.850 | 0.764 | 0.833 | 173 |

*Table 13: Model comparison for the city of Manchester (Prioritization to Public Transport)*

The results show that both algorithms achieve high precision and recall values for both classes, indicating that they are good at correctly identifying each class. However, the Random Forest Classifier outperforms the Multinomial Logit algorithm in terms of F1-score for both classes, suggesting that it is better at achieving a balance between precision and recall. Specifically, the Random Forest Classifier achieves an F1-score of 0.718 for the CAR class and 0.833 for the Public Transport class, compared to the Multinomial Logit algorithm's F1-scores of 0.617 and 0.764 for the respective classes. Therefore, it can be concluded that the Random Forest Classifier might be a better choice for this binary classification task, as it shows superior performance in terms of the F1-score.

In Table 14, the estimates of coefficients, standard errors, z-values, and corresponding p-values for every input variable, encompassed in a Multinomial Logit model developed for Manchester city, are presented.

| Feature | Coefficient | Std.Err. | z-val | P > z |
|---|---|---|---|---|
| **_intercept.PT** | -0.480 | 0.225 | 2.128 | 0.033* |
| **AGE.PT** | -0.169 | 0.059 | -2.860 | 0.004 ** |
| **INCOME.PT** | -0.880 | 0.108 | -8.181 | 0.000*** |
| **PTN_SATIS.PT** | 0.170 | 0.067 | 2.512 | 0.012 * |
| **IN_VEH_TIME_PT** | -0.082 | 0.007 | -11.829 | 0.000 *** |
| **COST_PT** | -0.056 | 0.069 | -0.813 | 0.416 |
| Sign. 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | |

*Table 14: Multinomial Logit model results for the city of Manchester (Prioritization to Public Transport)*

The outcomes of the feature importance of input features in the model created for Manchester city are illustrated in Table 15.

| Feature | Description | Importance Score |
|---|---|---|
| **AGE** | Age of individual | 0.337 |
| **PTN_SATIS** | Satisfaction with the level of service of Public Transport | 0.239 |
| **INCOME** | Income of individual | 0.152 |
| **IN_VEH_TIME_CAR** | In vehicle time with Car | 0.115 |
| **IN_VEH_TIME_PT** | In vehicle time with Public Transport | 0.093 |
| **COST_PT** | Trip cost with Public Transport | 0.041 |
| **COST_CAR** | Trip cost with Car | 0.020 |

*Table 15: Feature importance of Random Forest Classifier for the city of Manchester (Prioritization to Public Transport)*

## 3.2 Comparative assessment

### 3.2.1 Travel choice model under prioritization to Public Transport

In evaluating the ability of a Multinomial Logit model and a Random Forest classifier to predict travel choices under prioritization to Public Transport, their performance is assessed using mean accuracy from all classes and the weighted F1-score for every model and every city. Accuracy is the metric that measures the proportion of correct predictions made by the model out of the total number of predictions. The weighted average F1-score is a metric that takes into account both precision and recall for each class, providing a single score that summarizes the overall performance of the model. The weighted average F1-score takes the F1-score for each class and weighs it by the number of samples in that class, and then takes the average of the weighted F1-scores across all classes. A comparison of the two models' accuracy and weighted average F1-scores helps to determine the better-performing model for travel choice modelling purposes. In the context of the travel choice mode under prioritization to Public Transport, the results of the evaluation metrics are presented in Table 16.

| | Metric | Multinomial Logit | Random Forest Classifier |
|---|---|---|---|
| **Athens** | Accuracy (test-set 20%) | 0.69 | 0.75 |
| | Weighted average F1-score | 0.68 | 0.75 |
| **Lisbon** | Accuracy (test-set 20%) | 0.83 | 0.83 |
| | Weighted average F1-score | 0.75 | 0.80 |
| **Manchester** | Accuracy (test-set 20%) | 0.71 | 0.79 |
| | Weighted average F1-score | 0.71 | 0.79 |

*Table 16: Comparison of models for every city based on aggregative evaluation metrics for every city (Prioritization to Public Transport)*

Overall, the Random Forest Classifier seems to perform better than the Multinomial Logit algorithm on all three datasets, as it has higher accuracy and weighted F1-score values. However, the difference in performance between the two algorithms is more significant for some cities (e.g., Manchester) than others (e.g., Lisbon). These results indicate that the Random Forest Classifier is a more suitable algorithm for this particular classification task.

Table 17 shows the coefficients and significance levels of the factors that influence public transport selection in four different cities - Athens, Lisbon, Manchester, and Rennes. The coefficients indicate the direction and strength of the relationship between each factor and the likelihood of selecting public transport, while the significance levels indicate the probability of observing a coefficient as extreme as the one observed if there were no relationship between the factor and the outcome variable.

| Public Transport selection | Athens | | Lisbon | | Manchester | |
|---|---|---|---|---|---|---|
| **Factors** | Coef. | Sign. | Coef. | Sign. | Coef. | Sign. |
| COST_PT | - | *** | - | *** | - | |
| IN_VEHICLE_TIME_PT | - | *** | - | *** | - | *** |
| WAIT_TIME_PT | - | *** | - | *** | | |
| AGE | - | *** | - | *** | - | ** |
| INCOME | - | *** | - | *** | - | *** |
| PTN_SATIS | + | *** | + | *** | + | *** |
| PEAK_HOUR | - | * | | | | |
| Sign. 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | | | |

*Table 17: Coefficients and significance level of input variables for every city (Prioritization to Public Transport)*

The results show that the factors that influence Public Transport selection are consistent across the four cities, with COST_PT, IN_VEHICLE_TIME_PT, AGE, and INCOME all having negative coefficients and high significance levels, indicating that as these factors increase, the likelihood of selecting public transport decreases. On the other hand, PTN_SATIS has a positive coefficient and high significance level, indicating that higher levels of satisfaction with public transport increase the likelihood of selecting it. The factor PEAK_HOUR has a negative coefficient and low significance level, indicating that it has a weaker impact on public transport selection.

## 3.3 Travel choice model under Pricing Schemes

### 3.3.1 City of Athens

Table 18 presents the input variables of the models developed about travel mode choice in case a pricing scheme was implemented for the city of Athens.

| Input Variables | Description | Type | Unit |
|---|---|---|---|
| **TIME_CAR** | Total duration of trip with car | Integer | min |
| **COST_CAR** | Cost of trip with car | Integer | Euro |
| **TIME_CARTOLL** | Total duration of trip with car and toll | Integer | min |
| **COST_CARTOLL** | Cost of trip with car and toll | Integer | Euro |
| **TIME_PT** | Total duration of trip with Public Transport | Integer | min |
| **COST_PT** | Cost of trip with Public Transport | Integer | Euro |
| **PTN_SATIS** | Satisfaction with the level of service of Public Transport | 5-point scale [Not satisfied at all - Very satisfied] | - |
| **PEAK_HOUR** | Travelling on peak hours | Binary [Yes-No] | - |
| **INCOME** | Income of individual | 3-point scale [Low, Medium, High] | - |
| **OCCUP** | Occupation of individual | Multiple choice [Freelancer, Government employee, Private employee, Retired, Student, Unemployed] | - |

*Table 18: Input variables for travel choice models (Pricing schemes)*

Table 19 showcases the outcomes of both models (Multinomial Logit and Random Forest) designed for Athens, specifically for the pricing scheme mode choice model.

| Metrics comparison for Athens | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Class** | **Precision** | | **Recall** | | **F1-Score** | | **Number of Samples** |
| **Model** | MLN | RF | MLN | RF | MLN | RF | |
| **CAR** | 0.513 | 0.504 | 0.406 | 0.476 | 0.453 | 0.489 | 143 |
| **CAR and toll payment** | 0.478 | 0.430 | 0.413 | 0.419 | 0.443 | 0.425 | 155 |
| **Public Transport** | 0.541 | 0.581 | 0.686 | 0.618 | 0.605 | 0.599 | 191 |

*Table 19: Model comparison for the city of Athens (Pricing schemes)*

The results show that both algorithms have low precision and recall values for the ‘CAR’ and ‘CAR and toll payment’ classes, compared to the Public Transport class. This suggests that the algorithms are less effective at correctly identifying these classes. Furthermore, the F1-score for the CAR and CAR and toll payment classes is lower than that for the Public Transport class for both algorithms, indicating that they are less effective at achieving a balance between precision and recall for these classes.

Table 20 displays the estimates of coefficients, standard errors, z-values, and corresponding p-values for every input variable integrated into a Multinomial Logit model created under pricing schemes for Athens.

| Feature | Coefficient | Std.Err. | z-val | P > z |
|---|---|---|---|---|
| **_intercept.CARTOLL** | 1.241 | 240.596 | 0.005 | 0.996 |
| **_intercept.PT** | 0.402 | 781.939 | 0.001 | 1 |
| **PTN_SATIS.CARTOLL** | -0.101 | 0.055 | -1.836 | 0.066 |
| **PTN_SATIS.PT** | 0.190 | 0.055 | 3.411 | 0.001 *** |
| **PEAK_HOUR.CARTOLL** | -0.045 | 0.160 | -0.285 | 0.775 |
| **PEAK_HOUR.PT** | -0.325 | 0.154 | -2.108 | 0.035 * |
| **INCOME.CARTOLL** | -0.202 | 0.090 | -2.245 | 0.025 * |
| **INCOME.PT** | -0.572 | 0.091 | -6.276 | 0.000 *** |
| **OCCUP.CARTOLL** | -0.059 | 0.050 | -1.162 | 0.245 |
| **OCCUP.PT** | 0.146 | 0.049 | 2.954 | 0.003 ** |
| **DISTANCE** | -0.132 | 120.298 | -0.001 | 0.999 |
| **TIME** | -0.061 | 0.004 | -15.242 | 0.000*** |
| **COST** | -0.174 | 0.0498 | -3.504 | 0.000 *** |
| Sign. 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1 | | | | |

*Table 20: Multinomial Logit model results for the city of Athens (Pricing schemes)*

The results suggest that several factors significantly affect mode choice behaviour in both regions. Specifically, higher levels of public transportation satisfaction and income are positively associated with the likelihood of choosing public transportation, while peak hour and income have a negative effect on public transportation choice in Car with toll payment. Additionally, occupation has a positive effect on public transportation choice in Public Transport, while time and cost have a negative effect on the likelihood of travel mode choice.

Table 21 shows the results of feature importance of the model developed for the city of Athens under the implementation of the pricing schemes.

| Feature | Description | Importance Score |
|---|---|---|
| **PTN_SATIS** | Public Transport satisfaction | 0.306 |
| **OCCUP** | Occupation of individual | 0.283 |

| | | |
|---|---|---|
| **INCOME** | Income of individual | 0.157 |
| **PEAK_HOUR** | Travelling in peak hours | 0.093 |
| **TIME_ CAR** | Trip time with Car | 0.055 |
| **TIME_CARTOLL** | Trip time with Car and toll payment | 0.024 |
| **TIME_ PT** | Trip time with Public Transport | 0.024 |
| **COST_ CAR** | Cost of trip with Car | 0.021 |
| **COST_ PT** | Cost of trip with Public Transport | 0.016 |
| **COST_CARTOLL** | Cost of trip with Car and toll payment | 0.015 |

*Table 21: Feature importance of Random Forest Classifier for the city of Athens (Pricing schemes)*

Overall, these findings suggest that public transport satisfaction and occupation of the individual are the most important factors, followed by income and peak hour. Trip time and cost of travel with different modes of transportation have a lower impact on the travel mode choice.

We would like to highlight that the model findings shared thus far are provisional and depend on the responses obtained from the travel behaviour survey until M20. It's worth noting that the survey is still in progress, and as more data is gathered, the models will be retrained. The final outcomes will be integrated into Deliverable D3.4 "Travel choice modelling (set of models, code)" (M28).

# 4 Implementation

This deliverable results in the development of the travel choice module, which is designed as an analytical tool for modelling how individuals make travel choices under two scenarios:

1. Synchronization: For a typical trip from the suburbs to the city centre where public transport is synchronised with traffic control, commuters choose between a private car or the public transport system.
2. Pricing: For a trip from the suburbs to the city centre where a congestion charging scheme is enforced, commuters choose between driving a private car and paying the toll, driving a private car but bypassing the congestion charging area or using public transport.

The module will be a fundamental component of the final travel choice module which will be incorporated in the overall TANGENT solution.

At the input side, the module utilizes data collected from the travel behaviour survey described in Chapter 2, which is split into four databases, each corresponding to one of the case studies. From this data, city-specific choice models are created and calibrated. The output of the module will be either the estimation of acceptance of the proposed service or the distribution of the travellers across all the available travel modes. The process is described schematically in Figure 7.

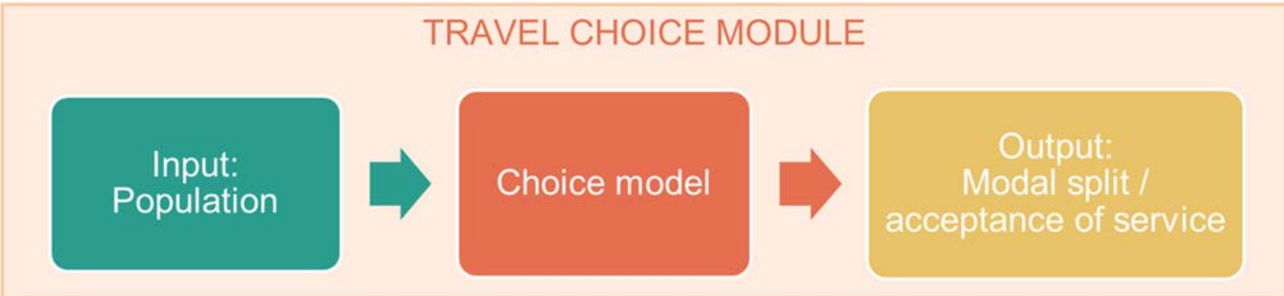


*Figure 7: Travel Choice Module*

By M20, eight models had been developed and calibrated based on stated preference scenarios, with two models (random forest and logit) created for each of Athens, Lisbon, and Manchester in the first scenario. For Rennes, the respective models will be recalibrated with the arrival of new survey data and presented in the second version of this deliverable at M28.

For the first scenario, two models (one random forest model and one logit model) have been created for each of Athens, Lisbon and Manchester. The model with the best accuracy from the two will be included in the final solution of TANGENT. Models for Rennes have not provided meaningful results at this stage and will be recalibrated based on new survey answers. Consequently, they are not included in this deliverable, and they will be presented in its second version in M28.

In the second scenario, which is linked to the Dynamic Congestion Pricing traffic management scheme and is relevant only for the Athens use case, two additional models were created and calibrated on Athens data.

It should be also noted that as the survey is still ongoing and anticipated to conclude in October 2023, all models will be recalibrated with the arrival of new survey data, and the updated results will be presented in the second version of this deliverable at M28.

The output of the module includes a modal split for each of the eight models, which provides the percentage breakdown of trips made by either car or public transport. This information serves as input for forecasting long-term demand shifts in WP4. The final users of the TANGENT solution may also use the travel choice model to simulate the potential impacts of new transportation services or policies, such as the introduction of a system the prioritizes public transit on signalized intersections (Scenario 1 models) or a dynamic congestion pricing scheme (Scenario 2 models for Athens). The travel choice module can estimate the acceptance of these new services by predicting how they will affect the modal split.

The presented results in the previous section showcase the performance of various models created for different cities, where the best performing model for each city was chosen based on various evaluation metrics. The primary objective of creating these models is to shift from a microscopic to a macroscopic view of the population's travel mode choices. Consequently, the final outcome is the percentage of individuals who prefer a specific mode of travel. To accomplish this, the framework utilizes a set of individuals and the variables that were considered during the creation of each model. Subsequently, it produces the percentage of individuals who selected each mode. The following results demonstrate the output for a randomized sample population, for each city. As already discussed in Section 3, the Random Forest classifier performed better in every city. The aggregative results on Figure 8 were created by utilizing the Random Forest classifier for the travel mode choice under prioritization to Public Transport for every city.

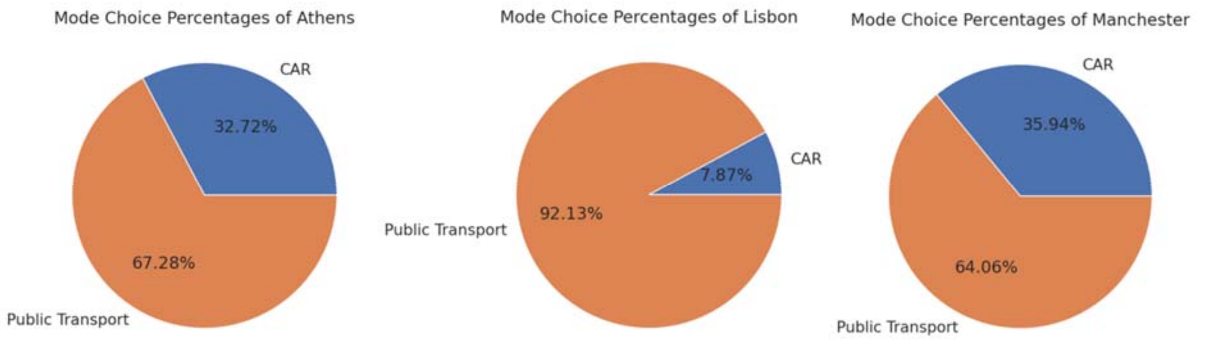


*Figure 8: Transport Mode choice percentages under Prioritization to Public Transport scenario for every city*

Also, the aggregative results for the travel mode choice under pricing schemes for the city of Athens are shown on Figure 9.

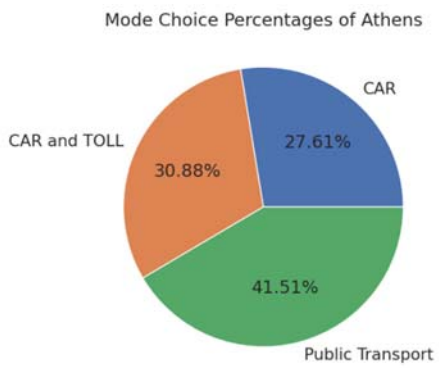


*Figure 9: Transport Mode choice percentages under pricing schemes for the city of Athens*

Same results as the above can be produced from the framework created, for every input of variables per model developed. The mode choice probability that is derived will be utilized as an input for forecasting long-term demand changes in WP4, as outlined in Deliverable 4.2 "Overview of the developed traffic supply forecasting approaches with benchmarks". The demand predictions will subsequently serve as an input for the optimization models in WP5 (see Deliverable 5.2 "Optimisation models for transport network management"), which will determine the most efficient network configuration while implementing the traffic management strategies of TANGENT Service 2.

# 5 Conclusions

The present study aimed to investigate the impact of changes in the traffic system and user choices on the transportation mode preferences of commuters in four different cities, namely Athens, Lisbon, Manchester, and Rennes.

To collect data to create and calibrate travel choice models, a travel behaviour survey has been conceived and launched. The respective questionnaire documents users' mobility profiles, perception of users towards new services and traffic management strategies, stated preference scenarios, and sociodemographic characteristics. The target of the survey is to collect information from 1000 participants for each case study in all four cities using the CONEY toolkit, which employs a conversational approach to enhance user experience. The survey was disseminated through online and on-site methods.

Until M20, 1859 questionnaires have been collected across all cities, while the collected sample is well distributed and overall suitable for model development. TANGENT partners will continue to distribute and monitor the quality of the sample until October 2023.

For the development of mode choice models, two distinct scenarios were examined for each city: (1) choice between car and Public Transport when prioritization is given to public transport over car, and (2) a choice between car, car with toll payment, and public transport based on the cost of the trip. To achieve this, the study employed statistical and machine learning models, namely Multinomial Logit and Random Forest models, to predict the travel choice of each user. The results indicated that the Random Forest machine learning model outperformed the Multinomial Logit model in predicting the travel choice of each user in all cities.

Moreover, the study investigated the factors that influence the choice of means of transport for each user. The most important factors for all cities studied were found to be the cost of travel and the total duration of the journey resulting from each mode choice. Other significant factors included age, income, and the degree of satisfaction with public transport. Additionally, for the city of Athens, it was observed that whether the user travels in peak hours is an essential parameter in their choice of transportation mode.

Closing, we note that accurate predictions of the modal split have significant potential to assist policymakers and transportation planners in promoting sustainable transportation modes. By estimating the impact of new services and traffic management strategies on travel demand for different modes, these predictions can help identify the factors that influence commuter behaviour and design effective strategies accordingly.

# 6 References

Alicioglu, G., & Sun, B. (2022). A survey of visual analytics for Explainable Artificial Intelligence methods. Computers & Graphics, 102, 502–520. https://doi.org/10.1016/j.cag.2021.09.002

Gasparinatou, C., Mantouka, E., Vlahogianni, E., David, A., & Serrano, L. (2022). Travel Behavior Shifts under Extreme System-Level Disruptions.

Kroes, E. P., & Sheldon, R. J. (1988). Stated preference methods: An introduction. Journal of Transport Economics and Policy, 11–25.

Steed, J. L., & Bhat, C. R. (2000). Modeling Departure Time Choice for Home-Based Non-Work Trips.

# Annex

The four distinct versions of the questionnaire tailored for Athens, Lisbon, Manchester, and Rennes can be found below. The CONEY version of the survey is accessible in: https://tangent.cefriel.com/survey/

# TRAVEL BEHAVIOR QUESTIONNAIRE

# (WP3)

## Athens version

# Section A: Mobility Profile

***This section aims at detecting the current mobility profile of each traveler and therefore, respondents should have in mind what they usually do during their daily trips when answering the questions.***

**A1. What is the main purpose of your everyday trips? (one answer)**
- ○ Professional
- ○ Educational
- ○ Personal (medical, family, housework)
- ○ Leisure (gym, friends, social, etc.)

**A2. Which mode do you use most often for your daily commute? (one answer)**
- ○ Car
- ○ Motorcycle
- ○ Public transport
- ○ Taxi
- ○ Walking/Bicycling

**A3. Do you travel in peak hours? (one answer)**
○ Yes ○ No

**A4. Do you have the flexibility to change the arrival time of your trips? (one answer)**
- ○ No
- ○ Yes, up to 30 minutes
- ○ Yes, up to 60 minutes
- ○ Yes, more than 60 minutes

**A5. What is the daily average cost of your trips (including commuting and parking)? (one answer)**○ 0€
- ○ 0-5€
- ○ 5-10€
- ○ >10

**A6.1.** ***Keeping in mind that a trip is considered a one-way movement from one point to another e.g., Home→Work, so if you travel daily from home to work and from work to home for five days per week you perform 10 trips*** **How many professional trips (e.g., going to the office, visiting a client etc.) do you perform on a weekly basis?__________**

**A6.2. How many educational trips (e.g., going to school or other educational establishments) do you perform on a weekly basis? __________**

**A6.3. How many personal trips do you perform on a weekly basis? (e.g., for medical, family or housework reasons) __________**

**A6.4. How many leisure trips do you perform on a weekly basis? (e.g., for visiting friends, a gym, other social activities etc.)__________**

# Section B: Perception of users

***The current transport network faces several challenges including traffic congestion and air pollution. In this section, respondents are asked about their perception towards a variety of traffic management strategies and new services that aim at improving the network conditions.***

**B1. CURRENT NETWORK CONDITIONS**

**1. How satisfied are you with the current Public Transport Network? (Only for respondents who replied "Public Transport" in A2)**

1(not satisfied at all)____2_____3_____4_____5(very satisfied)

**2. Please rank the PT attributes from the best to the worst. (Only for respondents who replied "Public Transport" in A2)**

○ Reliability
○ Frequency
○ Comfort
○ Cost
○ Availability
○ Safety
○ Information provision

**3. Regardless of the transit mode that you typically use for your daily commute, how much does traffic congestion affect your everyday trips?**

1(not at all)____2_____3_____4_____5(very much)

## B2. PRICING SCHEMES

***In order to relieve the city center from traffic congestion and minimize air pollution, an urban pricing scheme is applied, where private car users have to pay a form of toll for their entrance in the city center.***

**1. Under which conditions would you accept such a system and pay for your entrance in the city center? (multiple answers)**

□ The existence of reliable public transport
□ Significant reduction of traffic congestion
□ The charge is made according to the travelled distance
□ The charge is made according to the type of the vehicle (e.g., low vs high emissions vehicles)
□ Other, (define): _________
□ I am not willing to pay for my entrance to the city center

**2. In case a pricing scheme was applied in your city, which of the following would you choose to do? (one answer)**

○ Continue using my private car and pay the fee
○ Continue using my car, but change my route to avoid the charging area
○ Swift to Public Transport
○ Swift to an active travel mode (cycling, walking)
○ Other, please specify___________

## B3. ON-DEMAND SERVICES

***An on-demand service is a flexible, shared and pre-booked public transport service. On-Demand transport enables passengers to book their journey at a convenient time and to be picked up from an agreed location. Commuters can usually arrange the pickup and drop-off points through an application.***

**1. Would you be willing to use an on-demand service? (one answer)**

○ Yes
○ No

**2. Which of the following on-demand service's business model would you be willing to use? (one answer)**

○ **First-last mile** solution with synchronization with Public Transport - dynamically synchronize transfer between on-demand and regular public transit services to get you from your starting point to your final destination (low price)
○ An on-demand service with **fixed stops** - you can arrange to get in and out of the service at specific stops (medium price)
○ A **door-to-door** on-demand service where vehicles can move anywhere on the network (high price)

## B4. SYNCHRONIZATION OF TRAFFIC AND PT

***Transit signal priority (TSP) is a name for various techniques that facilitate the movement of public transport vehicles at intersections controlled by traffic signals.***

**1. Are you willing to use Public Transport for your everyday trips if priority would be given to it through traffic lights? (one answer)**

○ Yes
○ No

**2. Please rank the following objectives referring to the implementation of a Transit Signal Priority system at selected intersections, from the most to the least important:**

- Public transit schedule adherence and reliability are increased.
- Public transit travel time and delays are reduced.
- Impacts on normal traffic operations (e.g., network-wide congestion) are reasonable.
- Implementation costs are reasonable.
- Public transit GHG emissions are reduced.
- General traffic GHG emissions are reduced.
- Emergency vehicles can use the system to reduce response time.

## B5. RESPONSE TO UNEXPECTED EVENTS

**1. How likely are you to change your usual travel mode in case of:**

- **Adverse weather**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Natural hazards (e.g., earthquake, flooding)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Intense traffic**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Unplanned PT disruption**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Planned events (e.g., roadworks, football match)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Hazardous events (e.g., COVID-19)**
  1(not likely at all)____2_____3_____4_____5(certain)

**2. Have you changed your trip patterns as a consequence of Public Transport disruptions' experience?**

| Never | Rarely | Sometimes | Often | Always |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

**3. What is the maximum acceptable waiting time for you before considering an alternative travel action?**

**For time-critical trips – need to arrive on time (e.g., work)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**For time-non-critical trips – no need to arrive on time (e.g., shopping)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**4. What kind of information does the PT Service in your city offer to passengers in case of a disruption? (select all that apply)**

○ Information about the incident (location, stations affected, etc.)
○ Real-time information about the recovery progress
○ Alternative routes
○ Alternative travel options (e.g., on-demand service)
○ None of the above

**5. How would you respond to PT service disruption that occurred while using it? (one answer)**

- ○ Wait at the stop
- ○ Use an alternative PT route proposed by the PT service
- ○ Use my own alternative PT route
- ○ Walk down to destination
- ○ Shift to private transport
- ○ Resign from travelling

# Section C: Stated Preference Scenarios

**C1.** ***In order to perform a typical trip from the suburbs to the center of Athens, with length of approximately 10 km, during peak hours, the available modes are car and Public Transport. To improve traffic conditions traffic light management strategies that give priority to Public Transport are applied.***

***For each of the following scenarios you are asked to indicate the alternative that best describes your preferences.***

**C1.1 Scenario 1**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 65 minutes<br>Waiting time: -<br>Cost: 7€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 10 minutes<br>Cost: 2€ |
| ○ | ○ |

**C1.2 Scenario 2**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 50 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 5 minutes<br>Cost: 1.5€ |
| ○ | ○ |

**C1.3 Scenario 3**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 60 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 55 minutes<br>Waiting time: 10 minutes<br>Cost: 2€ |
| ○ | ○ |

**C1.4 Scenario 4**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 65 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 55 minutes<br>Waiting time: 15 minutes<br>Cost: 2€ |
| ○ | ○ |

**C1.5 Scenario 5**

Which alternative would you choose?

| Alternative 1 | Alternative 2 |
|---|---|
| Mode: Car<br>In-vehicle time: 50 minutes<br>Waiting time: -<br>Cost: 7€ | Mode: Public Transport<br>In-vehicle time: 45 minutes<br>Waiting time: 5 minutes<br>Cost: 2€ |
| ◯ | ◯ |

### C1.6 Scenario 6

Which alternative would you choose?

| Alternative 1 | Alternative 2 |
|---|---|
| Mode: Car<br>In-vehicle time: 60 minutes<br>Waiting time: -<br>Cost: 7€ | Mode: Public Transport<br>In-vehicle time: 45 minutes<br>Waiting time: 15 minutes<br>Cost: 1.2€ |
| ◯ | ◯ |

**C2. *For a typical trip from the suburbs to a central neighborhood of Athens during morning peak hours, commuters can choose between using their private car or opting for the public transport system. To improve traffic conditions, a congestion pricing scheme is applied: car users who wish to pass through the city centre must pay a toll. Commuters who are not willing to pay the toll but nevertheless want to continue using their car can bypass the tolled area and use peripheral roads to reach their destination. For each of the following scenarios you are asked to indicate the alternative that best describes your preferences.***

### C2.1 Scenario 1

Which alternative would you choose?

| Alternative 1 | Alternative 2 | Alternative 3 |
|---|---|---|
| Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 30 minutes<br>Cost: 3.5€ | Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 35 minutes<br>Cost: 2€ | Mode: Public Transport<br>Time: 45 minutes<br>Cost: 0.60€ |
| ◯ | ◯ | ◯ |

### C2.2 Scenario 2

Which alternative would you choose?

| Alternative 1 | Alternative 2 | Alternative 3 |
|---|---|---|
| Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 25 minutes<br>Cost: 5.5€ | Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 45 minutes<br>Cost: 2.5€ | Mode: Public Transport<br>Time: 40 minutes<br>Cost: 0.60€ |
| ◯ | ◯ | ◯ |

### C2.3 Scenario 3

Which alternative would you choose?

| Alternative 1 | Alternative 2 | Alternative 3 |
|---|---|---|
| Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 20 minutes<br>Cost: 5 € | Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 45 minutes<br>Cost: 2€ | Mode: Public Transport<br>Time: 60 minutes<br>Cost: 1.2€ |
| ◯ | ◯ | ◯ |

### C2.4 Scenario 4

Which alternative would you choose?

| Alternative 1<br>Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 25 minutes<br>Cost: 5 € | Alternative 2<br>Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 35 minutes<br>Cost: 3€ | Alternative 3<br>Mode: Public Transport<br>Time: 60 minutes<br>Cost: 0.60€ |
|---|---|---|
| ◯ | ◯ | ◯ |

### C2.5 Scenario 5

Which alternative would you choose?

| Alternative 1<br>Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 30 minutes<br>Cost: 5.5 € | Alternative 2<br>Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 35 minutes<br>Cost: 2.5€ | Alternative 3<br>Mode: Public Transport<br>Time: 40 minutes<br>Cost: 1.2€ |
|---|---|---|
| ◯ | ◯ | ◯ |

### C2.6 Scenario 6

Which alternative would you choose?

| Alternative 1<br>Mode: Car, payment of toll<br>Distance: 6.5 km<br>Time: 20 minutes<br>Cost: 3.5 € | Alternative 2<br>Mode: Car, bypass the tolled area<br>Distance: 8.5 km<br>Time: 30 minutes<br>Cost: 2.5€ | Alternative 3<br>Mode: Public Transport<br>Time: 45 minutes<br>Cost: 1.2€ |
|---|---|---|
| ◯ | ◯ | ◯ |

# Section D: Demographic Characteristics

### D1. Gender:

○ Male ○ Female

### D2. Age:

○ 18-24
○ 25-34
○ 35-44
○ 45-54
○ 55-64
○ >65

### D3. Annual Personal Income:

○ Low (<10,000€)
○ Medium (10,000 – 25,000€)
○ High (>25,000€)

### D4. Occupation:

○ Government employee
○ Private employee
○ Freelancer
○ Unemployed
○ Retired
○ Student

### D5.1 Car ownership:

○ Yes○ No

### D5.2 If yes, how many cars do you own: ______

### D6. Number of household members: ______

**D7. Do you live in the city-center?**

○ Yes ○ No

**D8. How would you characterize yourself based on your travel habits?**

○ Eco-traveler, I always choose the most eco-friendly alternative for my trip.
○ Time-saver, I always choose the shortest path/itinerary for my trip.
○ Technology-enthusiast, I always use the most up-to-date mobility services (e.g., on-demand services, e-scooters, ride-sharing services).
○ Car-dependent, I go everywhere by car.

**D9. If you wish to take part in the second round of the survey, and get more details about the findings of the TANGENT project, please fill in your email address (By filling in your email address you declare that you understand that your personal data will be used according to the privacy policy and only for the purposes of this project) .**

# TRAVEL BEHAVIOR QUESTIONNAIRE

# (WP3)

## Lisbon version

# Section A: Mobility Profile

***This section aims at detecting the current mobility profile of each traveler and therefore, respondents should have in mind what they usually do during their daily trips when answering the questions.***

**A1. What is the main purpose of your everyday trips? (one answer)**
○ Professional
○ Educational
○ Personal (medical, family, housework)
○ Leisure (gym, friends, social, etc.)
**A2. Which mode do you use most often for your daily commute? (one answer)**○ Car
○ Motorcycle
○ Public transport
○ Taxi
○ Walking/Bicycling
**A3. Do you travel during peak hours? (one answer)**
○ Yes ○ No

**A4. Do you have the flexibility to change the arrival time of your trips? (one answer)**○ No
○ Yes, up to 30 minutes
○ Yes, up to 60 minutes
○ Yes, more than 60 minutes
**A5. What is the daily average cost of your trips (including commuting and parking)? (one answer)**○ 0€
○ 0-5€
○ 5-10€
○ >10
**A6.1. *Keeping in mind that a trip is considered a one-way movement from one point to another e.g., Home→Work.* How many professional trips (e.g., going to the office, visiting a client etc.) do you perform on a weekly basis?__________**

**A6.2. How many educational trips (e.g., going to school or other educational establishments) do you perform on a weekly basis? __________**

**A6.3. How many personal trips do you perform on a weekly basis? (e.g., for medical, family or housework reasons) __________**

**A6.4. How many leisure trips do you perform on a weekly basis? (e.g., for visiting friends, a gym, other social activities etc.)__________**

# Section B: Perception of users

***The current transport network faces several challenges including traffic congestion and air pollution. In this section, respondents are asked about their perception towards a variety of traffic management strategies and new services that aim at improving the network conditions.***

**B1. CURRENT NETWORK CONDITIONS**
**1. How satisfied are you with the current Public Transport Network?**
1(not satisfied at all)____2_____3_____4_____5(very satisfied)

**2. Please rank the PT attributes from the best to the worst.**
○ Reliability
○ Frequency
○ Comfort
○ Cost

○ Availability
○ Safety
○ Information provision

**3. Regardless of the transit mode that you typically use for your daily commute, how much does traffic congestion affect your everyday trips?**

1(not at all)____2_____3_____4_____5(very much)

**B2. SYNCHRONIZATION OF TRAFFIC AND PT**

***Transit signal priority (TSP) is a name for various techniques that facilitate the movement of public transport vehicles at intersections controlled by traffic signals.***

**1. Are you willing to use Public Transport for your everyday trips if priority would be given to it through traffic lights? (one answer)**

○ Yes
○ No

**2. Please rank the following objectives referring to the implementation of a Transit Signal Priority system at selected intersections, from the most to the least important:**

- Public transit schedule adherence and reliability are increased.
- Public transit travel time and delays are reduced.
- Impacts on normal traffic operations (e.g., network-wide congestion) are reasonable.
- Implementation costs are reasonable.
- Public transit GHG emissions are reduced.
- General traffic GHG emissions are reduced.
- Emergency vehicles can use the system to reduce response time.

**B3. RESPONSE TO UNEXPECTED EVENTS**

**1. How likely are you to change your usual travel mode in case of:**

- **Adverse weather**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Natural hazards (e.g., earthquake, flooding)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Intense traffic**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Unplanned PT disruption**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Planned events (e.g., roadworks, football match)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Hazardous events (e.g., COVID-19)**
  1(not likely at all)____2_____3_____4_____5(certain)

**2. Have you changed your trip patterns as a consequence of Public Transport disruptions' experience?**

| Never | Rarely | Sometimes | Often | Always |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

**3. What is the maximum acceptable waiting time for you before considering an alternative travel action?**

**For time-critical trips – need to arrive on time (e.g., work)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**For time-non-critical trips – no need to arrive on time (e.g., shopping)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**4. What kind of information does the PT Service in your city offer to passengers in case of a disruption? (select all that apply)**

- ○ Information about the incident (location, stations affected, etc.)
- ○ Real-time information about the recovery progress
- ○ Alternative routes
- ○ Alternative travel options (e.g., on-demand service)
- ○ None of the above

**5. How would you respond to PT service disruption that occurred while using it? (one answer)**

- ○ Wait at the stop
- ○ Use an alternative PT route proposed by the PT service
- ○ Use my own alternative PT route
- ○ Walk down to destination
- ○ Shift to private transport
- ○ Resign from travelling

# Section C: Stated Preference Scenarios

**C1.** ***In order to perform a typical trip from the suburbs to the center of Lisbon, with length of approximately 8 km, during peak hours, the available modes are car and Public Transport. To improve traffic conditions traffic light management strategies that give priority to Public Transport are applied.***

***For each of the following scenarios you are asked to indicate the alternative that best describes your preferences.***

**C1.1 Scenario 1**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 45 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 40 minutes<br>Waiting time: 5 minutes<br>Cost: 2.5€ |
| ◯ | ◯ |

**C1.2 Scenario 2**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 60 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 15 minutes<br>Cost: 1.5€ |
| ◯ | ◯ |

**C1.3 Scenario 3**

Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 60 minutes<br>Waiting time: -<br>Cost: 3€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 30 minutes<br>Waiting time: 5 minutes<br>Cost: 1.5€ |
| ◯ | ◯ |

### C1.4 Scenario 4
Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 55 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 10 minutes<br>Cost: 2.5€ |
| ◯ | ◯ |

### C1.5 Scenario 5
Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 55 minutes<br>Waiting time: -<br>Cost: 3€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 40 minutes<br>Waiting time: 10 minutes<br>Cost: 1.5€ |
| ◯ | ◯ |

### C1.6 Scenario 6
Which alternative would you choose?

| | |
|---|---|
| Alternative 1<br>Mode: Car<br>In-vehicle time: 45 minutes<br>Waiting time: -<br>Cost: 3€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 30 minutes<br>Waiting time: 15 minutes<br>Cost: 2.5€ |
| ◯ | ◯ |

# Section D: Demographic Characteristics

### E1. Gender:
○ Male ○ Female

### E2. Age:
○ 18-24
○ 25-34
○ 35-44
○ 45-54
○ 55-64
○ >65

### E3. Annual Personal Income:
○ Low (<10,000€)
○ Medium (10,000 – 25,000€)
○ High (>25,000€)

### E4. Occupation:
○ Government employee
○ Private employee
○ Freelancer
○ Unemployed
○ Retired
○ Student

### E5.1 Car ownership:
○ Yes○ No

**E5.2 If yes, how many cars do you own: ______**

**E6. Number of household members: ______**

**E7. Do you live in the city-center?**

○ Yes ○ No

**E8. How would you characterize yourself based on your travel habits?**

○ Eco-traveler, I always choose the most eco-friendly alternative for my trip.
○ Time-saver, I always choose the shortest path/itinerary for my trip.
○ Technology-enthusiast, I always use the most up-to-date mobility services (e.g., on-demand services, e-scooters, ride-sharing services).
○ Car-dependent, I go everywhere by car.

**E9. If you wish to take part in the second round of the survey, and get more details about the findings of the TANGENT project, please fill in your email address.**

# TRAVEL BEHAVIOR QUESTIONNAIRE

# (WP3)

## Manchester version

# Section A: Mobility Profile

***This section aims at detecting the current mobility profile of each traveler and therefore, respondents should have in mind what they usually do during their daily trips when answering the questions.***

**A1. What is the main purpose of your everyday trips? (one answer)**
- ○ Professional
- ○ Educational
- ○ Personal (medical, family, housework)
- ○ Leisure (gym, friends, social, etc.)

**A2. Which mode do you use most often for your daily commute? (one answer)**
- ○ Car
- ○ Motorcycle
- ○ Public transport
- ○ Taxi
- ○ Walking/Bicycling

**A3. Do you travel during peak hours? (one answer)**
- ○ Yes ○ No

**A4. Do you have the flexibility to change the arrival time of your trips? (one answer)**○ No
- ○ Yes, up to 30 minutes
- ○ Yes, up to 60 minutes
- ○ Yes, more than 60 minutes

**A5. What is the daily average cost of your trips (including commuting and parking)? (one answer)**○ 0£
- ○ 0£-4£
- ○ 4£-9£
- ○ >9 £

**A6.1. *Keeping in mind that a trip is considered a one-way movement from one point to another e.g., Home→Work.* How many professional trips (e.g., going to the office, visiting a client etc.) do you perform on a weekly basis?__________**

**A6.2. How many educational trips (e.g., going to school or other educational establishments) do you perform on a weekly basis? __________**

**A6.3. How many personal trips do you perform on a weekly basis? (e.g., for medical, family or housework reasons) __________**

**A6.4. How many leisure trips do you perform on a weekly basis? (e.g., for visiting friends, a gym, other social activities etc.)__________**

# Section B: Perception of users

***The current transport network faces several challenges including traffic congestion and air pollution. In this section, respondents are asked about their perception towards a variety of traffic management strategies and new services that aim at improving the network conditions.***

**B1. CURRENT NETWORK CONDITIONS**

**1. How satisfied are you with the current Public Transport Network?**

1(not satisfied at all)____2_____3_____4_____5(very satisfied)

**2. Please rank the PT attributes from the best to the worst.**○ Reliability
- ○ Frequency
- ○ Comfort
- ○ Cost

○ Availability
○ Safety
○ Information provision

**3. Regardless of the transit mode that you typically use for your daily commute, how much does traffic congestion affect your everyday trips?**

1(not at all)____2_____3_____4_____5(very much)

**B2. SYNCHRONIZATION OF TRAFFIC AND PT**

***Transit signal priority (TSP) is a name for various techniques that facilitate the movement of public transport vehicles at intersections controlled by traffic signals.***

**1. Are you willing to use Public Transport for your everyday trips if priority would be given to it through traffic lights? (one answer)**

○ Yes
○ No

**2. Please rank the following objectives referring to the implementation of a Transit Signal Priority system at selected intersections, from the most to the least important:**

- Public transit schedule adherence and reliability are increased.
- Public transit travel time and delays are reduced.
- Impacts on normal traffic operations (e.g., network-wide congestion) are reasonable.
- Implementation costs are reasonable.
- Public transit GHG emissions are reduced.
- General traffic GHG emissions are reduced.
- Emergency vehicles can use the system to reduce response time.

**B3. RESPONSE TO UNEXPECTED EVENTS**

**1. How likely are you to change your usual travel mode in case of:**

- **Adverse weather**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Natural hazards (e.g., earthquake, flooding)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Intense traffic**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Unplanned PT disruption**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Planned events (e.g., roadworks, football match)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Hazardous events (e.g., COVID-19)**
  1(not likely at all)____2_____3_____4_____5(certain)

**2. Have you changed your trip patterns as a consequence of Public Transport disruptions' experience?**

| Never | Rarely | Sometimes | Often | Always |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

**3. What is the maximum acceptable waiting time for you before considering an alternative travel action?**

**For time-critical trips – need to arrive on time (e.g., work)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**For time-non-critical trips – no need to arrive on time (e.g., shopping)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**4. What kind of information does the PT Service in your city offer to passengers in case of a disruption? (select all that apply)**

- ○ Information about the incident (location, stations affected, etc.)
- ○ Real-time information about the recovery progress
- ○ Alternative routes
- ○ Alternative travel options (e.g., on-demand service)
- ○ None of the above

**5. How would you respond to PT service disruption that occurred while using it? (one answer)**

- ○ Wait at the stop
- ○ Use an alternative PT route proposed by the PT service
- ○ Use my own alternative PT route
- ○ Walk down to destination
- ○ Shift to private transport
- ○ Resign from travelling

# Section C: Stated Preference Scenarios

**C1. *In order to perform a typical trip from the suburbs to the center of Manchester, with a length of approximately 7 km, during peak hours, the available modes are car and Public Transport. To improve traffic conditions traffic light management strategies that give priority to Public Transport are applied.***

***For each of the following scenarios you are asked to indicate the alternative that best describes your preferences.***

**C1.1 Scenario 1**

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 55 minutes<br>Waiting time: -<br>Cost: 4.5£ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 30 minutes<br>Waiting time: 5 minutes<br>Cost: 4£ |
|---|---|
| ○ | ○ |

**C1.2 Scenario 2**

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 45 minutes<br>Waiting time: -<br>Cost: 2.5£ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 50 minutes<br>Waiting time: 10 minutes<br>Cost: 4£ |
|---|---|
| ○ | ○ |

**C1.3 Scenario 3**

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 65 minutes<br>Waiting time: -<br>Cost: 4.5£ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 50 minutes<br>Waiting time: 5 minutes<br>Cost: 5£ |
|---|---|
| ○ | ○ |

### C1.4 Scenario 4

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 55 minutes<br>Waiting time: -<br>Cost: 2.5£ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 30 minutes<br>Waiting time: 2 minutes<br>Cost: 5£ |
|---|---|
| ◯ | ◯ |

### C1.5 Scenario 5

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 45 minutes<br>Waiting time: -<br>Cost: 4.5£ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 40 minutes<br>Waiting time: 10 minutes<br>Cost: 5£ |
|---|---|
| ◯ | ◯ |

### C1.6 Scenario 6

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 65 minutes<br>Waiting time: -<br>Cost: 2.5£ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 50 minutes<br>Waiting time: 2 minutes<br>Cost: 4£ |
|---|---|
| ◯ | ◯ |

# Section D: Demographic Characteristics

### E1. Gender:

○ Male ○ Female

### E2. Age:

- ○ 18-24
- ○ 25-34
- ○ 35-44
- ○ 45-54
- ○ 55-64
- ○ >65

### E3. Annual Personal Income:

- ○ Low (<£18,000)
- ○ Medium (<£45,000)
- ○ High (>£45,000)

### E4. Occupation:

- ○ Government employee
- ○ Private employee
- ○ Freelancer
- ○ Unemployed

○ Retired
○ Student

**E5.1 Car ownership:**
○ Yes○ No

**E5.2 If yes, how many cars do you own: ______**

**E6. Number of household members: ______**

**E7. Do you live in the city-center?**
○ Yes ○ No

**E8. How would you characterize yourself based on your travel habits?**
○ Eco-traveler, I always choose the most eco-friendly alternative for my trip.
○ Time-saver, I always choose the shortest path/itinerary for my trip.
○ Technology-enthusiast, I always use the most up-to-date mobility services (e.g., on-demand services, e-scooters, ride-sharing services).
○ Car-dependent, I go everywhere by car.

**E9. If you wish to take part in the second round of the survey, and get more details about the findings of the TANGENT project, please fill in your email address.**

# TRAVEL BEHAVIOR QUESTIONNAIRE

# (WP3)

## Rennes version

# Section A: Mobility Profile

***This section aims at detecting the current mobility profile of each traveler and therefore, respondents should have in mind what they usually do during their daily trips when answering the questions.***

**A1. What is the main purpose of your everyday trips? (one answer)**
- ○ Professional
- ○ Educational
- ○ Personal (medical, family, housework)
- ○ Leisure (gym, friends, social, etc.)

**A2. Which mode do you use most often for your daily commute? (one answer)**
- ○ Car
- ○ Motorcycle
- ○ Public transport
- ○ Taxi
- ○ Walking/Bicycling

**A3. Do you travel during peak hours? (one answer)**
- ○ Yes ○ No

**A4. Do you have the flexibility to change the arrival time of your trips? (one answer)**○ No
- ○ Yes, up to 30 minutes
- ○ Yes, up to 60 minutes
- ○ Yes, more than 60 minutes

**A5. What is the daily average cost of your trips (including commuting and parking)? (one answer)**○ 0€
- ○ 0-5€
- ○ 5-10€
- ○ >10

**A6.1. *Keeping in mind that a trip is considered a one-way movement from one point to another e.g., Home→Work.* How many professional trips (e.g., going to the office, visiting a client etc.) do you perform on a weekly basis?__________**

**A6.2. How many educational trips (e.g., going to school or other educational establishments) do you perform on a weekly basis? __________**

**A6.3. How many personal trips do you perform on a weekly basis? (e.g., for medical, family or housework reasons) __________**

**A6.4. How many leisure trips do you perform on a weekly basis? (e.g., for visiting friends, a gym, other social activities etc.)__________**

# Section B: Perception of users

***The current transport network faces several challenges including traffic congestion and air pollution. In this section, respondents are asked about their perception towards a variety of traffic management strategies and new services that aim at improving the network conditions.***

**B1. CURRENT NETWORK CONDITIONS**

**1. How satisfied are you with the current Public Transport Network?**

1(not satisfied at all)____2_____3_____4_____5(very satisfied)

**2. Please rank the PT attributes from the best to the worst.**
- ○ Reliability
- ○ Frequency
- ○ Comfort

○ Cost
○ Availability
○ Safety
○ Information provision

**3. Regardless of the transit mode that you typically use for your daily commute, how much does traffic congestion affect your everyday trips?**

1(not at all)____2_____3_____4_____5(very much)

**B2. ON-DEMAND SERVICES**

***An on-demand service is a flexible, shared and pre-booked public transport service. On-Demand transport enables passengers to book their journey at a convenient time and to be picked up from an agreed location. Commuters can usually arrange the pickup and drop-off points through an application.***

**1. Would you be willing to use an on-demand service? (one answer)**

○ Yes
○ No

**2. Which of the following on-demand service's business model would you be willing to use? (one answer)**

○ **First-last mile** solution with synchronization with Public Transport - dynamically synchronize transfer between on-demand and regular public transit services to get you from your starting point to your final destination (low price)

○ An on-demand service with **fixed stops** - you can arrange to get in and out of the service at specific stops (medium price)

○ A **door-to-door** on-demand service where vehicles can move anywhere on the network (high price)

**B3. SYNCHRONIZATION OF TRAFFIC AND PT**

***Transit signal priority (TSP) is a name for various techniques that facilitate the movement of public transport vehicles at intersections controlled by traffic signals.***

**1. Are you willing to use Public Transport for your everyday trips if priority would be given to it through traffic lights? (one answer)**

○ Yes
○ No

**2. Please rank the following objectives referring to the implementation of a Transit Signal Priority system at selected intersections, from the most to the least important:**

- Public transit schedule adherence and reliability are increased.
- Public transit travel time and delays are reduced.
- Impacts on normal traffic operations (e.g., network-wide congestion) are reasonable.
- Implementation costs are reasonable.
- Public transit GHG emissions are reduced.
- General traffic GHG emissions are reduced.
- Emergency vehicles can use the system to reduce response time.

**B4. RESPONSE TO UNEXPECTED EVENTS**

**1. How likely are you to change your usual travel mode in case of:**

- **Adverse weather**

  1(not likely at all)____2_____3_____4_____5(certain)

- **Natural hazards (e.g., earthquake, flooding)**

1(not likely at all)____2_____3_____4_____5(certain)

- **Intense traffic**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Unplanned PT disruption**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Planned events (e.g., roadworks, football match)**
  1(not likely at all)____2_____3_____4_____5(certain)
- **Hazardous events (e.g., COVID-19)**
  1(not likely at all)____2_____3_____4_____5(certain)

**2. Have you changed your trip patterns as a consequence of Public Transport disruptions' experience?**

| Never | Rarely | Sometimes | Often | Always |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

**3. What is the maximum acceptable waiting time for you before considering an alternative travel action?**

**For time-critical trips – need to arrive on time (e.g., work)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**For time-non-critical trips – no need to arrive on time (e.g., shopping)**

| 5 minutes | 10 minutes | 15 minutes | 20 minutes | 30 minutes | 60 minutes |
|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ |

**4. What kind of information does the PT Service in your city offer to passengers in case of a disruption? (select all that apply)**

○ Information about the incident (location, stations affected, etc.)
○ Real-time information about the recovery progress
○ Alternative routes
○ Alternative travel options (e.g., on-demand service)
○ None of the above

**5. How would you respond to PT service disruption that occurred while using it? (one answer)**

○ Wait at the stop
○ Use an alternative PT route proposed by the PT service
○ Use my own alternative PT route
○ Walk down to destination
○ Shift to private transport
○ Resign from travelling

# Section C: Stated Preference Scenarios

**C1.** ***In order to perform a typical trip from the suburbs to the center of Rennes, with a length of approximately 6 km, during peak hours, the available modes are car and Public Transport. To improve traffic conditions traffic light management strategies that give priority to Public Transport are applied.***

***For each of the following scenarios you are asked to indicate the alternative that best describes your preferences.***

**C1.1 Scenario 1**

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 30 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 25 minutes<br>Waiting time: 15 minutes<br>Cost: 2.5€ |
|---|---|
| ◯ | ◯ |

### C1.2 Scenario 2

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 30 minutes<br>Waiting time: -<br>Cost: 3€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 10 minutes<br>Cost: 1.5€ |
|---|---|
| ◯ | ◯ |

### C1.3 Scenario 3

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 45 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 20 minutes<br>Waiting time: 10 minutes<br>Cost: 2.5€ |
|---|---|
| ◯ | ◯ |

### C1.4 Scenario 4

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 45 minutes<br>Waiting time: -<br>Cost: 3€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 5 minutes<br>Cost: 1.5€ |
|---|---|
| ◯ | ◯ |

### C1.5 Scenario 5

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 25 minutes<br>Waiting time: -<br>Cost: 3€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 35 minutes<br>Waiting time: 15 minutes<br>Cost: 1.5€ |
|---|---|
| ◯ | ◯ |

### C1.6 Scenario 6

Which alternative would you choose?

| Alternative 1<br>Mode: Car<br>In-vehicle time: 25 minutes<br>Waiting time: -<br>Cost: 5€ | Alternative 2<br>Mode: Public Transport<br>In-vehicle time: 25 minutes<br>Waiting time: 15 minutes<br>Cost: 2.5€ |
|---|---|
| ◯ | ◯ |

# Section D: Demographic Characteristics

**E1. Gender:**

○ Male ○ Female

**E2. Age:**

○ 18-24
○ 25-34
○ 35-44
○ 45-54
○ 55-64
○ >65

**E3. Annual Personal Income:**

○ Low (<10,000€)
○ Medium (10,000 – 25,000€)
○ High (>25,000€)

**E4. Occupation:**

○ Government employee
○ Private employee
○ Freelancer
○ Unemployed
○ Retired
○ Student

**E5.1 Car ownership:**

○ Yes○ No

**E5.2 If yes, how many cars do you own:** ______

**E6. Number of household members:** ______

**E7. Do you live in the city-center?**

○ Yes ○ No

**E8. How would you characterize yourself based on your travel habits?**

○ Eco-traveler, I always choose the most eco-friendly alternative for my trip.
○ Time-saver, I always choose the shortest path/itinerary for my trip.
○ Technology-enthusiast, I always use the most up-to-date mobility services (e.g., on-demand services, e-scooters, ride-sharing services).
○ Car-dependent, I go everywhere by car.

**E9. If you wish to take part in the second round of the survey, and get more details about the findings of the TANGENT project, please fill in your email address.**