

# Technical support related to sustainable urban mobility indicators (SUMI)

MOVE/B4/2017-358

## Harmonisation Guideline

Final (web) version

28 August 2020



# Table of contents

- TABLE OF CONTENTS..... 2**
- LIST OF ABBREVIATIONS ..... 3**
- 1. EXECUTIVE SUMMARY ..... 4**
- 2. INTRODUCTION..... 4**
  - 2.1 Background..... 4
  - 2.2 Goal of this document ..... 5
- 3. THE SUMI INDICATORS..... 6**
- 4. DEFINITIONS ..... 8**
- 5. HARMONISATION OF DATA.....17**
  - 5.1 Data format and categories ..... 18
  - 5.2 Data collection frequency ..... 23
    - 5.2.1 Non-survey data – frequency ..... 23
    - 5.2.2 Survey data – sampling ..... 23
    - 5.2.3 Alternative data sources versus surveys ..... 25

## List of abbreviations

CNG	Compressed Natural Gas
E.V.	Electric Vehicle
EC	European Commission
EU	European Union
GHG	Greenhouse Gas
GPS	Global Positioning System
ITF	International Transport Forum
LPG	Liquefied Petroleum Gas
NTS	National Travel Survey
NUTS	Nomenclature of Territorial Units for Statistics
OECD	The Organisation for Economic Co-operation and Development
PT	Public Transport
UNECE	United Nations Economic Commission for Europe

# 1. Executive summary

This document is the final version of the data harmonisation guidelines, which is also made available on the SUMI website. As it should be possible to use SUMI indicators to discuss an urban areas performance over time and with respect to other urban areas, it is important that the data used to calculate the indicators is consistent.

A first important element for data harmonisation is to have common definitions. The most important ones are described in this report. At the same time SUMI's indicator calculation spreadsheets provide most of the relevant data definitions in the respective "user guides" sheet. In addition, the data format and categories were discussed starting from a list of parameters which urban areas should track regularly.

The last section focusses on the frequency of data collection and briefly discussed – for surveys – on sampling and alternative sources. Overall, it is recommended to collect data yearly – with a maximum gap of five years.

## 2. Introduction

### 2.1 Background

A consortium led by Rupprecht Consult – Forschung & Beratung GmbH (Germany) and composed of TRT Trasporti e Territorio (Italy), Transport & Mobility Leuven (Belgium), Polis (Belgium), Eurocities (Belgium) and UITP, Union Internationale des Transports Publics (Belgium), has been selected by the European Commission – DG MOVE to support the testing of Sustainable Urban Mobility Indicators (SUMI) within the **"Service Contract: Technical support related to sustainable urban mobility indicators" (MOVE/B4/2017-358)**.

The partners were supported by eight subcontractors acting as additional "Urban Area Coaches": CERTH, The Centre for Research & Technology, (Greece), UIRS, the Urban Planning Institute of the Republic of Slovenia (Croatia), Žilina University (Slovakia), TRIVECTOR (Sweden), STRAFICA (Finland), Mobilissimus (Hungary), University of Gdansk (Poland) and TIS, Consultores em Transportes Inovação e Sistemas (Portugal), as well as by the European Cyclists' Federation (Belgium) to support on safety of non-motorised transport modes.

The starting point for the SUMI project were the **"SMP2.0 Sustainable Mobility Indicators"**<sup>1</sup> developed by WBCSD, the World Business Council for Sustainable Development. These have subsequently been revised by the SUMI consortium for use by European cities.

The common development and use of a methodologically sound, practically feasible and **harmonised** indicator set on sustainable urban mobility is fundamental for European urban areas in order to analyse progress towards their goals and policy objectives as well as to identify deficiency areas where additional action may be required. Moreover, urban areas need a system of indicators that is widely accepted and used in Europe, irrespective of city size and characteristics of the mobility system.

---

<sup>1</sup> <http://www.wbcd.org/Projects/SIMPLify/Resources/SMP2.0-Sustainable-Mobility-Indicators-2nd-Edition>

This important challenge needs a strong direct involvement of urban areas and the provision of technical support on the appraisal of and data gathering for such indicators, which had been provided to urban areas during the lifetime of the SUMI project.

## 2.2 Goal of this document

Given that the data is coming from different urban areas and even different countries, harmonisation of data has been of particular relevance for this project. The goal of this document is to proactively handle and prevent possible data harmonisation problems when collecting data. A first version of this document has served as a guide for Indicator Mentors and Urban Area Coaches to ensure a common approach on harmonisation and to build a knowledge database for the long-term deployment of good harmonisation practices.

This final version of the guidelines is an update of the first version, which builds heavily on Eurostat (2016) EU Transport Statistics – Eurostat guidelines on Passenger Mobility Statistics but has been tailored and updated to the needs of the SUMI indicators for urban areas.

Eurostat developed these guidelines in 2016 as passenger mobility statistics are not part of the regulated European Statistical System and no regular and harmonised data collection exist in the field of road passenger transport. The existing data collection usually responds to national/regional/local data needs. Hence, there is a need to have harmonisation guidelines for collecting data on urban areas.

In addition, this final version of the harmonisation guidelines incorporates the lessons learned as a result of the interactions between the Data and Statistics Experts, Indicator Mentors, Urban Area Coaches and the representatives from the urban areas/cities that cooperated with the SUMI project.

This document is organised in the following chapters:

- A short description of SUMI's core indicators
- A set of definitions as a common understanding is the first step towards harmonisation
- Methods of post-harmonisation of national/regional/local survey results
- Lessons learned during the SUMI project

### 3. The SUMI indicators

The table below provides an overview of the SUMI indicators, indicating whether they are a core indicator or not. Within the SUMI project, the cooperating urban areas were requested to gather all necessary data to calculate at least the core indicators. The calculation of the non-core indicators was voluntary.

No.	Indicator	Definition	Core indicator
1	Affordability of public transport for the poorest group	Share of the poorest quartile of the population's household budget required to hold public transport (PT) passes (unlimited monthly travel or equivalent) in the urban area of residence.	Yes
2	Accessibility of public transport for mobility-impaired groups	This indicator determines the accessibility of public transport services to persons with reduced mobility. Such vulnerability groups include those with visual and audial impairments and those with physical restrictions, such as pregnant women, users of wheelchairs and mobility devices, the elderly, parents and caregivers using buggies, and people with temporary injuries.	Yes
3	Air pollutant emissions	Air pollutant emissions of all passenger and freight transport modes (exhaust and non-exhaust for PM <sub>2.5</sub> ) in the urban area.	Yes
4	Noise hindrance	Hindrance of population by noise generated through urban transport.	Yes
5	Road deaths	Road deaths by all transport accidents in the urban area on a yearly basis.	Yes
6	Access to mobility services	Share of population with appropriate access to mobility services (public transport).	Yes
7	Greenhouse gas emissions (GHG)	Well-to-wheels GHG emissions by all urban area passenger and freight transport modes.	Yes
8	Congestion and delays	Delays in road traffic and in public transport during peak hours compared to off peak travel (private road traffic) and optimal public transport travel time (public transport).	Yes
9	Energy efficiency	Total energy use by urban transport per passenger km and tonne km (annual average over all modes).	Yes
10	Opportunity for active mobility	Infrastructure for active mobility, namely walking and cycling.	Yes
11	Multimodal integration	An interchange is any place where a traveller can switch from one mode of travel to another, with a minimum/ reasonable amount of walking or waiting. The more modes available at an interchange, the higher the level of multimodal integration.	Yes
12	Satisfaction with public transport	The perceived satisfaction of using public transport.	Yes

No.	Indicator	Definition	Core indicator
13	Traffic safety active modes	Fatalities of active modes users in traffic accidents in the city in relation to their exposure to traffic.	Yes
14	Quality of public spaces	The perceived satisfaction of public spaces.	No
15	Urban functional diversity	Functional diversity refers to a mix of spatial functions in an area, creating proximity of mutual interrelated activities	No
16	Commuting travel time	Duration of commute to and from work or an educational establishment, using any types of modes.	No
17	Mobility space usage	Proportion of land use, taken by all city transport modes, including direct and indirect uses.	No
18	Security	The perceived risk of crime and passenger security in urban transport.	No
	Modal split	<p>For passenger mobility:</p> <p>Modal split according to passenger kilometres ran</p> <p>Modal split according to vehicle kilometres ran</p> <p>Modal split according to the number of trips ran</p> <p>Modal split according to the number of vehicle kilometres per trip ran</p> <p>For freight:</p> <p>Modal split according to goods vehicles kilometres ran</p> <p>Modal split according to freight tonnes kilometres ran</p>	Yes

## 4. Definitions

This chapter contains recommendations on definitions for relevant terminology that is used in the methodology section and that are required to carry out a uniform analysis of the collected data. While Eurostat’s “EU Transport Statistics – Eurostat guidelines on Passenger Mobility Statistics” (2016) provide a wide range of definitions, the SUMI consortium has selected and updated the definitions of those parameters relevant to the needs of the SUMI indicators for urban areas. In addition, some definitions were also added in the SUMI indicator calculation spreadsheets.

Within this section, definitions are presented for different terminology used within the SUMI indicator set, but also in surveys (during data collection) and during the reporting of findings from surveys. The reason for keeping definitions related to surveys is that some of the indicators still rely on survey data.

The focus lies on those terminologies that require a specific definition in order to enhance comparability between urban areas. Hence not all parameters used to calculate the SUMI indicators are presented.

For each terminology, the following information is presented: (1) chosen definition, (2) possible definitions (if multiple definitions can/ could be found in the different input documents), and (3) an elaboration (including calculation if relevant).

### Reference population

#### *Definition*

The **reference population** for compilation of indicators provided should include **all residents of the urban area aged 15 to 84 inclusive**.

#### *Possible alternative definitions*

Urban area population

Different subgroups of population

#### *Elaboration*

The term ‘reference population’ refers to the population group for which the collected information is meant to be representative and, within the framework of SUMI, means the entire population of the urban area. The urban area population includes all inhabitants of the territory, including any non-nationals with a residence in the urban area. This means that both residents living in families and residents in institutions for e.g. elderly or disabled as well as students are included.

If surveys are carried out, typically age limits are introduced. The choice of using an age limits on respondents to be included in the sample and so also on the reference population reflects various data needs of national/regional/local transport policies and should be therefore left to the urban area. However, in order to maintain comparability between urban areas, it is recommended that all urban areas provide the requested indicators (linked to surveys) for the age group covering respondents from 15 years old up to 84 years old (inclusive).



## All days, weekday, working day, weekend day, average day

### *Definition*

**Weekdays** are defined as the five days Monday, Tuesday, Wednesday, Thursday or Friday (including any public holidays on those days)

A **working day** is defined as a Monday, Tuesday, Wednesday, Thursday or Friday (excluding any public holidays on those days)

**Non-working days** are Saturday or Sunday, or official holidays, including bank holidays and other official national or regional non-business days.

A **weekend day** is a Saturday or Sunday,

An **average day** is a weighted average of the working days and weekend days.

**All days** is all the days during the survey period weighted up to a year.

The minimum to provide is working days/non-working days.

### *Possible alternative definitions*

An alternative is to provide indicators for weekdays/weekend days independent of the holidays.

Another alternative is to include school holidays and local holidays in non-working trips.

### *Elaboration*

Countries and urban areas are using different methods when reporting data. However, most countries/urban areas are properly registering if a weekday is an official holiday. It is therefore preferable to use the definitions working days / non-working days for the indicators to be delivered. For countries which have not yet registered the official holidays and are not able to ex-post register it based on travelling dates it is acceptable to use the alternative definition for the indicators.

Holidays are defined as official holidays declared by law or similar or as an official agreement for the whole labour market. Official school holidays and the business summer holiday etc. are not included in the non-working days. The same is the case with weekdays during Christmas and after Ascension Day which many companies have declared a holiday. For some countries, for instance Germany, official holidays are at different dates in different regions. In case the Member State has decided to register these different dates for the sampled respondents, dependent on which region they live in these dates can be used as non-working days for the relevant part of the sample. In general, local holidays only relevant in one city or a few municipalities are not included in the non-working-days, but when collecting data at urban area level could be taken into account.

All indicators should be delivered per day weighted up to a year. For countries doing a survey all year round and including one travelling day no extra calculations are needed to find the indicator per day.

It is recommended in the survey to register the precise day of the week and not only weekday and weekend. Especially Saturdays and Sundays have quite different traffic patterns for e.g. shopping and leisure. Also, the weekdays are different; especially Friday is different from the first 4 days.

## Travel mode (or mode of transport)

### Definitions

Mode is defined as a vehicle or non-vehicle like walking used for travelling.

### Values for mode

The general definitions for ‘travel mode’ correspond to the main mode of transport used during a trip or journey. They could be grouped into four general categories: air, rail, road and water.

For the purpose of monitoring of the European transport policy and for reporting of Passenger Mobility Indicators described in these Guidelines, the following travel mode categories are distinguished:

Private car (as driver and/or as passenger): this category includes driving or being driven as a passenger in cars up to 8 passenger seats that are owned by the driver, leasing cars, company cars, rental cars, shared cars (a form of car rental where members subscribe to a service and can make use of the entire car fleet contained in the service), or borrowed cars, as well as informal carpooling systems or as a passenger on a remunerated ride (ride hailing such as Lyft, Uber,...). It does not include driving as a passenger in a taxi. Reported values should be split into as driver/as passenger.

Taxi (as passenger): the category includes all trips as a paying customer with a professional driver. This mode refers to travel as a passenger only. Professional trips of transport professionals (taxi/train/bus drivers) should not be included in the scope of the survey. However, all other trips of those professionals such as going to and from work, shopping, etc., using their professional vehicles should be included. Taxi driver using his taxi for other purposes than driving clients should report such trips as driver of a passenger car.

Motorcycle/moped/powerd two-wheelers: this category includes all powered two-wheelers, three-wheelers, and quads, as well as snow-scooters and similar.

Bus/coach: this category contains all vehicles of UNECE class M2 and M3 designed to seat more than nine persons (including the driver), used primarily for the transport of passengers, and also trolleybuses. Buses are vehicles of this type designed to carry both standing and sitting passengers. Coaches are vehicles designed to carry only sitting passengers. Trolleybuses are passenger road vehicles designed to seat more than nine persons (including the driver), connected to electric conductors but not rail-borne.

Metro: Metro is a form of rapid mass transit, with high frequency and stops generally no more than 1000m apart.

Tram/Light rail: Trams, tramways or light rail are vehicles running on tracks often integrated in the urban road system (as opposed to urban rail or metro systems which are generally separated from interactions with road vehicles). They can be both electric or diesel-powered, but are usually electric.

Train (split into total, High-speed Rail, Urban Rail, Regular train): three subcategories of train can be identified, but for SUMI they do not have to be reported separately:

- High-speed train: rail vehicles running mostly on both dedicated and upgraded high-speed railway lines, built to allow speeds of at least 250km/h and 200km/h respectively. Examples are TGV, Eurostar, ICE, etc.

- Urban rail: rapid transit with stops generally no more than 1200m apart – examples include RER, S-Bahn, etc.), but excluding metro and tram.
- Regular train: all other rail-borne vehicles not included in the first two sub-categories and excluding trams.

Cycling: this category covers all kinds of bicycles including electric bikes as well as motor-driven wheelchairs. Mail distribution and other professional duties are not included

Walking: this category includes running, roller skating, skating, manual wheelchairs and similar.

Waterways (total): includes both inland waterways (including rivers, lakes, and canals) and maritime transport and all types of water-borne vehicles including cruise boats, ferries, motorboats, sailboats, rowing boats, etc. Professional sailors are excluded.

Other: The use of this category should be limited only to travel modes which cannot be assimilated to any of the above-mentioned categories and respondents should always be asked for the precise travel mode to verify if this is the case; moreover, it is advisable to sub-divide this category into 'other motorised' and 'other non-motorised' to give it more meaning. Examples of travel modes grouped here can be horseback, horse carriage, dog sled, self-balancing scooters, one-wheeler (motorized), etc.

Van/lorry/light goods vehicles (<3.5 ton): this category includes all professional trips, in vehicles primarily intended for transport of freight (UNECE/EC classes N1)<sup>2</sup>.

Heavy goods vehicles (>3.5 ton): all professional trips in the transport of freight using heavy goods vehicles (UNECE/EC classes N2, N3)<sup>3</sup>.

Other freight modes include freight rail and inland waterways.

#### *Elaboration*

This list of modes is used for travel modes at both stages and trips. Whether mode information is collected as stages or at trip level is influencing the travel distance per mode and the comparability of data across countries/urban areas. Therefore, the recommendation is that, within a survey, the travel mode information should be collected for each stage of a trip.

## **Main travel mode**

### *Definitions*

Main travel mode is the travel mode within a trip which has been used for the **longest distance**.

### *Possible alternative definitions*

Main travel mode is considered the travel mode within a trip that is used for the longest time.

The main mode can also be defined from a priority list of travel modes. Examples of main mode prioritisation are ('a >b' indicates priority of 'a' over 'b'): plane > public transport > car > bike > walk or : other modes > rail > bus > car > ..., etc.

<sup>2</sup> UNECE, Consolidated Resolution on the Construction of Vehicles (R.E.3) Revision 2, 2011

<sup>3</sup> UNECE, Consolidated Resolution on the Construction of Vehicles (R.E.3) Revision 2, 2011

### *Elaboration*

Eurostat (2016) states that the definition based on distance is the most often used definition and possible for common use.

The definition based on longest time spent on modes may be considered as travelers often remember travel time better than distance. Although, this might result in a bias to slow modes like walking when this is combined with a bus or train ride.

The use of a definition based on a weighting procedure has the advantage of presenting a more pronounced effect when different modes are used, while at the same time reducing information on other modes. Such a definition may be preferred when particular policy options need to be analysed, but at the same time it restricts the more general use of information and reduces comparability over countries if different weighting procedures would be used.

## **Fuel type**

### *Definition*

The type of fuel (energy carrier) used in a vehicle: gasoline, diesel, gasoline-hybrid, diesel-hybrid, Electric Vehicle (E.V.), LPG, CNG, Ethanol, Bio-ethanol, bio diesel, Hydrogen.

### *Possible alternative definitions*

The type of fuel (energy carrier) used in a vehicle.

Different groupings can be suggested, depending on the level of detail required for policy analyses or national preferences.

Grouping option used within SUMI: gasoline, diesel, gasoline-hybrid, diesel-hybrid, Electric Vehicle (E.V.), LPG, CNG, Ethanol, Bio-ethanol, bio diesel, Hydrogen.

### *Elaboration*

In theory, a high number of grouping options can be identified where other or more detailed fuel types can be identified. For example: hydrogen fuel cells (FC) or flex-fuel.

As a general approach towards suggesting a definition for fuel types, a rationale can be followed where a balance is found between identifying common denominators across countries (allowing for grouping of more detailed information) and presenting a level of detail for which data collection is considered feasible.

Contrary to other indicators the grouping for this parameter may vary over time, depending on the availability and success of future energy carriers.

## **Travel Purpose (also trip purpose, travel motive, trip motive)**

### *Definition*

The main purpose of a trip without which the trip would not have occurred, with the distinction between work (commuting), other professional, education, shopping, escorting, professional, leisure, other personal business.

Travel purpose:	Inclusion of (more detailed or other) travel purposes listed in NTS by countries:
Work (commuting)	Daily/non-daily commute, travel to/from work, etc.
Other professional	Business, in course of work, trade, employee business, etc.
Education	Travel to/from school or educational institution, school field trips, etc.
Escorting	Picking up/accompanying/escorting people, taking children to school, etc.
Shopping	Shopping, food shopping, daily shopping, groceries, non-daily shopping, etc.
Leisure	Visiting friends/family, going out to eat or drink, touring/walking, Sport/Hobby, Recreation at water/beach/mountains, cultural activities, sightseeing, agro-tourism, visit/stay, holiday, sport, entertainment, voluntary work, other leisure, etc.
Other personal business	Services/personal care, health, treatment, personal reasons, religious, other, etc.

The category "escorting" should be used whenever respondent accompanies somebody else somewhere; it covers taking children to and from school, crèche, after-school activities, birthday parties etc., but also e.g. driving a spouse to and from work or taking someone to hospital, etc.

This means that driving to work, if it concerns the respondent directly, is "work" and dropping or picking up somebody else from work is "escorting".

Similarly, going to school or to attend training or a seminar is "education" if it concerns the respondent's education. However, taking or picking up children to/ from school is "escorting" and not "education".

The category "other personal business" should contain all purposes that are not professional and cannot be put in any of the other categories. It includes going to a doctor, hairdresser, post office, bank, etc.

#### *Possible alternative definitions*

The main purpose of a trip without which the trip would not have occurred (note: this definition focusses on individual trips).

The different purposes for which a person travelled (note: this definition groups different trip purposes from different successive trips).

Different groupings of purposes were identified as a result of information collected in different NTS:

- Grouping option 1: Work, education, shopping, business, leisure, other
- Grouping option 2: Personal, Professional
- Grouping option 3: Work (commuting), other professional, education, escorting, shopping, leisure, other personal business.

#### *Elaboration*

In theory, a high number of grouping options can be identified with many more purpose categories. Indeed, many countries use much more extensive lists, which foresee separate categories for activities such as: visiting family or friends, going out to eat or drink, medical visits, doing sports, food shopping, non-food shopping, etc.

As a general approach towards suggesting a definition for travel purposes, a rationale can be followed where a balance is found between identifying common denominators across countries (allowing for

grouping of more detailed information) and presenting a level of detail for which data collection is considered feasible. Another approach could have been to choose a very high level of aggregation and only make a split between personal and professional motives. The suggested definition (option 3 above) fits this general approach.

Some countries use a separate category for "home trip" or "return trip" or "return home" which is the last trip of the day or last trip of a set of trips (arriving at the home residence or similar). Other countries do not have such a separate category and attribute such return trips the same purpose as the trip before or yet the same as the main purpose of the series of trips.

In order to achieve comparability between countries Eurostat suggests that all countries attribute the "return" trips the same trip purpose as the one just before. For example: trip returning home from work should have purpose category "work", returning home from shopping should be attributed to "shopping", etc. This means that e.g. going for groceries after work, and then going home, should be considered "shopping" and not "work". This method has a disadvantage of over-representing the last travel purpose before returning home (often shopping or picking up children = escorting) but has the advantage of being simple to apply.

## Stage

### *Definition*

A stage is an uninterrupted movement making use of one transport mode, including any waiting time directly before or during the movement. A stage is defined by one single mode of transport. If a change of mode of transport takes place, this means that another stage is initiated upon the change of transport modes.

Example 1: changing from one bus to another bus is considered an interruption. Hence, this comprises two stages.

Example 2: biking from home to the station, taking the train to the next station and walking to work comprises three stages.

### *Possible alternative definitions*

A stage is an uninterrupted movement making use of one transport mode.

A stage is defined by the use of one transport mode.

### *Elaboration*

The terminology 'stage' is commonly adapted in countries. One underlying definition for the different interpretation of the terminology 'stage' is that it is linked to a single mode of transport (or rather type of vehicle used).

Variations exist depending on whether it is a continuous mobility action or not, and whether waiting time is included in the stage or not.

Walking should always be reported as the starting stage (see model questionnaire in Chapter4), if someone does not have to walk (for example when using in-house parking space), then the starting stage is 0 meters long.

## Trip

### *Definition*

A trip is defined as change in location made in one or a series of stages. The change of location is not based on the start and end point being different, but on leaving the starting point for a certain amount of time.

### *Elaboration*

This concept is commonly in use in all countries.

A trip is usually defined as change in location made in one or a series of stages.

The change of location is not necessarily based on the start and end point being different, but on leaving the starting point for a certain amount of time. Trips consisting of a single stage and with the starting point the same as the end point are loops (e.g. jogging or cycling), which should also be included.

Some countries are also using a concept of a journey, especially in dedicated surveys on long distance mobility.

Journey is usually composed two trips: away trip and return trip.

However, the indicators to be provided to Eurostat should all be based on trips and not journeys.

Countries collecting information on journeys should transpose one journey into two trips.

## Travel time

### *Definition*

The time spent travelling from the moment of departure to the moment of arrival. The travel time includes the time spent waiting between two successive stages.

### *Possible alternative definitions*

The time spent travelling from the moment of departure to the moment of arrival.

The time spent travelling from the moment of departure to the moment of arrival, without the waiting time.

### *Elaboration*

The central element in the definition is the time between departure and arrival. The inclusion of waiting time could possibly cause a problem for the registration of data (insofar that non-automated mobility registration collection methods are concerned).

## Peak and off peak (travel time)

### *Definition*

Peak-hour: this corresponds to the beginning and end of the working day, when large numbers of people are travelling to or from work. The corresponding hours depend on citizens' habits and working legislation.

Off-peak: often, night-time traffic is measured and used for the estimation of "off-peak". However, this may not be the most relevant time period for such measurements, as vehicles also tend to speed more at night and this would therefore not represent a realistic measurement. It is therefore recommended to measure off-peak travel speed during the middle of the morning or afternoon.

## **Travel distance per person per day**

### *Definition*

Weighted travel distance in kilometres divided by weighted number of respondents at the actual day.

The weighted kilometres per person per day are calculated by summarising all respondents' kilometres multiplied by the "weight" of the respondent and divided by the weighted sum of respondents. This is done for all respondents and separate for respondents answering at working days respective non-working days. For a multi-day survey these results have to be divided by the number of days surveyed (for a 7 days survey the numbers are divided by 7). In case of a shorter survey period than a year the passenger kilometres have to be up scaled to a year.

## **Number of trips per person per day**

### *Definition*

Weighted number of trips divided by weighted number of respondents at the actual day.

Trips per person per day are calculated similar to the distance by replacing kilometres by 1 (for one trip).

## **Passenger kilometres for all reference population per year**

### *Definition*

Passenger kilometres are calculated as the total weighted travel distance in kilometre made by all persons within the reference population in a reference year.

## **Vehicle occupancy rate**

### *Definition*

Vehicle occupancy rate is the average number of persons travelling in a passenger vehicle on a reference year.

It is recommended to provide vehicle occupancy for passenger cars and taxis for all age groups and not only for 15-84 year.

### *Elaboration*

Calculation of vehicle occupancy rate for passenger cars and taxis is described in chapter 2.2. All calculations should be made on data weighted by the respondent. In order to obtain results for the entire population (all age 24 groups) it is necessary to complement the calculation above with additional information from all car drivers on number of trip companions to be asked for each stage of a trip. It is needed to make a distinction between age groups that are covered by the reference population and those that are not (e.g. below 15 years, between 15 and 84 years, and above 84 years). In case of a taxi trip all respondents are asked the same question.



As far as other travel modes are concerned (such as buses and coaches, trams, metro or trains) it is clear that the NTS cannot provide such information. It is therefore recommended to collect the information directly from public transport companies if possible.

## Urban Area

### *Definition*

Functional urban area consisting of a city and its commuting zone.

### *Possible alternative definitions*

Currently three possible definitions are used by a number of countries. They are drawn from practices that are used by Eurostat, OECD, the Directorates-General (DGs) for Regional and Urban Policy, Agriculture and Rural Development, ITF and UNECE.

- OECD-EC use the definition of a “larger urban zone” composed of a “city” and its “commuting zone”
- The EC classification of the “degree of urbanisation” (DEGURBA), distinguishing between three types of areas representing cities, towns and rural areas.
- “urban area” defined as an area within the administrative boundary or a set of administrative boundaries of a core city (settlement). This is the description given in the joint Eurostat/ITF/UNECE “Glossary on Transport Statistics”.

The three definitions have their advantages and disadvantages. The first definition was preferred by Eurostat (2016) as it would best fit the distinction between urban and non-urban trips.

### *Elaboration*

The definition of the "commuting zone" is mainly based on municipalities or similar administrative areas. However, for some countries, smaller administrative areas are used for definition of commuting zones which means that the border of the Functional Urban Area is not following the municipality borders completely.

The greater urban zones can be divided into 4 types of zones according to the overall population in the zone:

- Large Metropolitan area, population of 1.5 million people or more
- Metropolitan area, population between 500,000 and 1.5 million people
- Medium-sized urban area, population between 200,000 and 500,000 people
- Small urban area, population below 200,000 people

## 5. Harmonisation of data

Within this project, the information and data needed to calculate the indicators are collection by different people, in different urban areas and with different backgrounds. However, with the SUMI indicators one wants to

- follow the progress of an urban area over time with respect to the different indicators

- compare urban areas within countries and between countries; benchmarking urban areas with each other.

Hence it is important that the input data is comparable. Otherwise the results are not comparable.

In order to have comparable data, SUMI paid much attention to providing correct and concise definitions (see indicator spreadsheets and section 3 above). However, in practice differences might still occur. Hence, this section focusses on post-harmonisation. In this first version of the document, the focus lies on outlining the approach which will be taken. Next versions of the document will incorporate the lessons learned once the data is available. The final goal would be to propose a structure for harmonised data gathering at urban level.

A proposition for the content of a harmonised urban area data set is made in Section 4.1. This includes a listing of the set of parameters and variables for which data should be collected, but also the variables' type and suggested grouping. However, in itself, this information is not enough to provide data that is comparable between different countries. For all data, the data collection frequency is important. For data which relies on surveys, also information on sampling frequency and sample composition needs to be taken into account. A suggestion for these elements is presented in Section 4.2.

## 5.1 Data format and categories

Currently, a number of urban areas collect a fair share of the required data, even though the current data categorisation does not always match. Experience from these urban areas can provide a good starting point for a more harmonised data collection. However, in contrast to data collected at country level, there is not one institute such as, for example, Eurostat, collecting urban data and putting them into one database systematically. There are some databases such as UN Habitat Urban data (<http://urbandata.unhabitat.org/explore-data/>), European Commission JRC Urban data ([https://urban.jrc.ec.europa.eu/#/european-trends?level=CITY&year=2010&indicator=Population\\_pop&type=thematics](https://urban.jrc.ec.europa.eu/#/european-trends?level=CITY&year=2010&indicator=Population_pop&type=thematics)) and Eurostat – Regions, Metropolitan regions and cities (<https://ec.europa.eu/eurostat/web/metropolitan-regions/data/database>). However, these data sources usually capture higher level data than the data required to calculate the SUMI indicators.

Urban areas should be encouraged to collect the following parameters at a regular basis. With these data all SUMI indicators can be calculated at a regular basis.

**Table 1: Parameter listing for harmonised urban data collection – general information**

General information				
Parameter	Description	Unit	Quantitative/ qualitative	Recommended Grouping
GDP	Gross Domestic Product of the urban area	€/year	quantitative	
Surface	Surface of the urban area	km <sup>2</sup>	quantitative	
Direct land use of mobility	Direct land use of mobility in ha	ha	quantitative	roads/railways/inland ports and waterways/other direct land use (such as bus lanes, cycle lanes,...)
Indirect land use of mobility	Indirect land use of mobility in ha	ha	quantitative	public parking/private parking/service areas and petrol stations/storage and logistic centres/stations/other
Functional activities	Presence (yes/no) of functional activities in a zone (e.g. grid of 1km*1km)	0/1	quantitative	business/energy resources/hospital/services/school/commercial/sports recreation/residential/residence elderly/park and green

**Table 2: Parameter listing for harmonised urban data collection – demographic information**

Demographic information				
Parameter	Description	Unit	Quantitative/ qualitative	Recommended Grouping
Capita	Number of inhabitants of the urban area	#	quantitative	total number of inhabitants in each zone (e.g. grid of 1km*1km)
Gender	Number of males and females in the urban area	#	quantitative	female/male
Age	Number of inhabitants of the urban area in different age groups	#	quantitative	under 15, 15-17, 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75 and over
Employment	Number of inhabitants per employment status	#	quantitative	employed/unemployed/student/retired/other
Income	Average monthly income of the inhabitants of the urban area	€	quantitative	per quartile
Household size	Number of people in a household in the urban area	#	quantitative	

Table 3: Parameter listing for harmonised urban data collection – general mobility information

General mobility information				
Parameter	Description	Unit	Quantitative/ qualitative	Recommended Grouping
Length road network	Total length of urban area road network (excluding motorways)	km	quantitative	total/ with pavements (not in pedestrian zone)/with bike lanes (not in 30 km/h zone)/in 30 km/h zone/pedestrian zone
vkm/pkm/number of passenger trips/tonkm	vehicle - km, passenger-km, ton-km	vehicle - km, passenger-km, number of passenger trips, ton-km/year	quantitative	no grouping
vkm/pkm/number of passenger trips/tonkm per mode	vehicle - km, passenger-km, ton-km per transport mode	vehicle - km, passenger-km, number of passenger trips, ton-km/year	quantitative	air/road/rail/water
vkm/pkm/number of passenger trips/tonkm per transport type	vehicle - km, passenger-km, ton-km per transport type	vehicle - km, passenger-km, number of passenger trips, ton-km/year	quantitative	pedestrian/Cyclist/powered two-wheelers/motorised three-wheelers/own car/ taxi/ride-hailing/bus (M2;M3)/coach/metro/tram-lightrail/train/light goods vehicle (<3.5 ton: <1305;1305-1760;>1760) /heavy goods vehicle
vkm/number of trips shared mobility of vehicle	vkm/number of trips of shared mobility per transport type	vkm/number of trips/year	quantitative	car/bike/other
vkm/number of trips shared mobility of trip	vkm/number of trips	vkm/number of trips/year	quantitative	
Vehicle fleet per fuel type	Number of vehicles per engine/fuel type	number of vehicles per fuel type/engine	quantitative	car/bus (M2;M3)/coach/ PTW/ motorised three-wheeler/train/LGV (<1305 kg, 1305-1760,>1760)/HGV gasoline/diesel/CNG/LPG/ethanol/bio-ethanol/biodiesel/hydrogen/electricity/hybrid-gasoline/hybrid-diesel
Vehicle fleet per euronorm	Number of diesel and gasoline cars per euronorm	number of diesel and gasoline cars per euronorm	quantitative	car/bus (M2;M3)/coach/ PTW/ motorised three-wheeler/LGV (<1305 kg, 1305-1760,>1760)/HGV gasoline and diesel: pre Euro and Euro1/ Euro 2/Euro3/Euro4/Euro5/Euro 6 and post Euro6
Fuel consumption	Fuel or energy consumption	l/100 vkm	quantitative	car/bus (M2;M3)/coach/ PTW/ motorised three-wheeler/tram, metro/lightrail/train/inland waterways/LGV (<1305 kg, 1305-1760,>1760)/HGV gasoline/diesel/CNG/LPG/ethanol/bio-ethanol/biodiesel/hydrogen/electricity/hybrid-gasoline/hybrid-diesel
Fuel consumption per euronorm	Fuel or energy consumption per Euronorm	l/100 vkm	quantitative	three-wheeler/LGV (<1305 kg, 1305-1760,>1760)/HGV

Table 4: Parameter listing for harmonised urban data collection – transport quality

Transport quality				
Parameter	Description	Unit	Quantitative/ qualitative	Recommended Grouping
Modes (public transport and shared mobility)	Modes (public transport and shared mobility) available in city	yes/no	qualitative	long distance bus/railway/metro/LRT, tram/local bus/bike sharing stations/park&ride/reserved taxi rank/ferry total/ at interchanges
Monthly ticket price public transport	The price of a pass for one month travel using public transport	€	quantitative	per public transport company
No. of ticketing machines & offices	No. of ticketing machines & offices	#	quantitative	train/bus&trolley/tram/metro/cable car, funicular, racked railway/river shuttle, ferry
No. of vehicles	Number of vehicles	#	quantitative	train/bus&trolley/tram/metro/cable car, funicular, racked railway/river shuttle, ferry with on-board signage/with on-board audio announcement/with step free access/with designated space provision
No. of stops	Number of stops	#	quantitative	train/bus&trolley/tram/metro/cable car, funicular, racked railway/river shuttle, ferry with audio announcements/with step free access to station/with step free access within station/with step free access through lowered curb for ferry/water: with human assistance
Population living within each access typology	Number of people living in each access typology zone	#	quantitative	no access/low access/medium access/high access/very high access
Number of respondents survey general satisfaction public transport	Number of respondents survey general satisfaction public transport	#	quantitative	total/DK,NO/satisfied/rather satisfied/rather unsatisfied/ not at all satisfied with - public transport in your urban area
Number of respondents survey agreement quality public transport	Number of respondents survey agreement quality public transport	#	quantitative	total/DK,NO/strongly agree/ somewhat agree/somewhat disagree/strongly disagree with public transport in your urban area is - affordability - safe - easy to get - frequent - reliable

Table 5: Parameter listing for harmonised urban data collection – Effects on inhabitants

Effects on inhabitants (safety, noise,...)				
Parameter	Description	Unit	Quantitative/ qualitative	Recommended Grouping
Number of people exposed to different noise bands	Number of people exposed to different noise bands	#	quantitative	major roads/major railways/major airports Lden: 55-59/60-64/65-69/70-74/>75
Number of fatalities	Number of fatalities (30 days) per transport type	#	quantitative	pedestrian/bicycle/moped/motorcycle/car/LG V/HGV/bus/tram, lightrail,.../other/unknown
Number of respondents survey quality public areas	Number of respondents survey quality public areas	#	quantitative	total/DK,NO/satisfied/rather satisfied/rather unsatisfied/ not at all satisfied  with - green space such as parks and gardens - public space such as markets, squares, pedestrian areas
Number of respondents survey travel time	Number of respondents survey travel time	#	quantitative	/
Average distance work-home	The average distance from work to home for each respondent of the survey	km	quantitative	for each respondent of the survey
Main mode work -home	The main mode for going to work for each respondent of the survey	listing	qualitative	for each respondent of the survey
Average travel time going to work	The average time (in minutes) to go to work using the mode stipulated for each respondent of the survey	minutes	quantitative	for each respondent of the survey
Average travel time return	The average time (in minutes) to return from work using the mode stipulated for each respondent of the survey	minutes	quantitative	for each respondent of the survey
Number of car trips in the peak for 10 road corridors	Number of car trips in the peak for 10 road corridors	#	quantitative	for 10 road corridors
Car travel time in peak hour for 10 road corridors	Car travel time in peak hour for 10 road corridors	minutes	quantitative	for 10 road corridors
Car travel time off peak for 10 road corridors	Car travel time off peak for 10 road corridors	minutes	quantitative	for 10 road corridors
Number of public transport trips in the peak for 10 public transport corridors	Number of public transport trips in the peak for 10 public transport corridors	#	quantitative	for 10 public transport corridors
Public transport travel time in peak for 10 public transport corridors	Public transport travel time in peak for 10 public transport corridors	minutes	quantitative	for 10 public transport corridors
Optimal public transport travel time off peak for 10 public transport corridors	Optimal public transport travel time off peak for 10 public transport corridors	minutes	quantitative	for 10 public transport corridors
Average score of security	Average score of security question j related to mode i	score		for different security questions related to different modes (public transport/car/motorcycle/cyclist/walking)
Number of persons in survey related to security	Number of persons in survey related to security	#	quantitative	

## 5.2 Data collection frequency

### 5.2.1 Non-survey data – frequency

Within the context of policy support, a clear separation between an original situation and a situation after the implementation of a policy or technology is achieved resulting in clear “before implementation” and “after implementation” comparison. However, in practice such a situation is only seldom achieved. Projects leading to policy introductions in real life seldom include clear pre- and post-measurements of the mobility status. Furthermore, even if such a measurement is included, the influence of other factors that also influence mobility behaviour are only seldom measured and excluded from such an impact analysis. It could be argued that such a situation can be avoided by collecting data on a continuous basis. However, this often is impossible due to limited funding, limited project time or other reasons.

In practice, and recognisable from the different data sources for which information was collected in the urban areas up to now, a repeat frequency of 1 year is suggested. This means that data is collected regularly on a yearly basis. Within the context of mobility policies, and the associated comparison of policy introduction effects, it is not recommended a time period of more than 5 years for comparison of data.

### 5.2.2 Survey data – sampling<sup>4</sup>

In this section, a proposal is made on two elements related to sampling: surveying frequency and sample composition. Both are relevant in order to obtain a certain level of comparability between data which relies on survey data. Often, data from successive surveys within one urban area are compared in order to analyse changes over time. As long as there is a simple quantitative comparison, a large time gap between surveys does not have to pose problems. However, when causal relations are introduced, this does become a problem. A multitude of factors influence mobility behaviour and change can happen rapidly. A frequency for repeating surveys is proposed in the following paragraphs.

At the same time, comparisons between urban areas are also often made. In these cases, the following elements are of importance: parameter definitions (does a figure in urban area A mean the same in urban area B? This is more likely if both countries belong to the same country. A solution is presented in Section 4.1.) and sample composition (is data collected from similar types of people? A sample selection is proposed in the section below). Both of these factors are required in order to consider the data collected from different surveys to be comparable.

#### *Sampling frequency*

Identifying a good frequency for repeating surveys depends on a number of elements related to practical feasibility on one hand and minimum requirements to guarantee data continuity and comparability over and across samples on the other hand.

---

<sup>4</sup> This section is based on the OPTISM project.

As with non-survey based data, a repeat frequency of 1 year is suggested. This means that data is collected regularly on a yearly basis. Within the context of mobility policies, and the associated comparison of policy introduction effects, it is not recommended a time period of more than 5 years for comparison of data.

### *Sample composition*

The second issue influencing the comparability of surveys is that the samples are composed in such a way that they are representative for the population they are meant to reflect upon. For example: it makes little sense to compare the mobility patterns of two urban areas from two different countries when no information is available on the type of population they represent (population composition, demographic information, economic situation, etc.). At the same time, an indication of the representativeness of each sample for the populations they represent needs to be foreseen.

As a starting point, it is suggested that population registries are used to compose population samples. They are, as such, less influenced by social differences that are introduced as a result of technology based sample bases (for example the use of telephone directories or email listings as sample base). Within the DATELINE project, other possible strategies are provided in case of unavailability of such registries<sup>5</sup>.

Often related to sample composition are issues in relation to sample size. For this, we refer to the working paper of Freedman<sup>6</sup>. In essence, it is impossible to set a specific percentage (of persons included in the sample, compared to the population size) that is representative for every population. Furthermore, response rates can be very different for different countries. This makes the net sample (the number of usable/valid returns at a unit level) important. Choosing a correct net sample depends on the following elements:

- Desired precision of results (error of margin): the difference between the real and the sampled population.
- Confidence level: the amount of risk that you are willing to take that the sample taken contains the populations' true values within the precision of results defined. A higher confidence interval requires a larger sample size.
- Degree of variability: the distribution of attributes or concepts that are measured in the questions in the total population. A homogenous population is easier to measure than a heterogeneous population. The degree of variability is often estimated based on prior information or expert information.
- Response rate: the amount of valid responses compared to the number of respondents that are tested, if possible over NUTS regions.

An example:

- A 5% error of margin means that the true value of a population is within 5% of the value that is found in the sample. If the value of a survey indicates that 20% of the people in the survey sample use public transport, and the error of margin is 5%, this means that the real population value lies between 15 and 25%.

---

<sup>5</sup> Socialdata (2000). DATELINE Deliverable 3: Sampling Methodology. FP5 DG TREN project DATELINE: Design and application of a travel survey for European long-distance trips based on an international network of expertise.

<sup>6</sup> <https://www.stat.berkeley.edu/~census/sample.pdf>



- A confidence level of 95%, given a 5% error or margin then means the following: if 100 samples were taken, 95 of these samples would have effectively have the real population value within the error of margin presented.
- A population with a 50%-50% division on an attribute is considered very heterogeneous. A population with a 80%-20% division homogenous.

In the case of urban surveys, it is suggested to aim at an error of margin that is not bigger than 5% and a confidence level of 95% or better.

### 5.2.3 Alternative data sources versus surveys

Although different technologies are becoming more available, they also come with specific problems. The use of GPS information, smartphones or other technologies is considered to be potentially valuable for specific types of information (route information, time information, travel times, modal choice, etc.). They can enhance data reliability, reduce data collection time, etc. However, some issues do exist with the use of these technologies: social and demographic biases, privacy issues but also the sheer quantity of data that can cause problems for analysts.

In the face of this, it can be considered that surveys remain valuable instruments that allow analysts to obtain more depth of analysis into the transport data. As well-known and accepted data sources, they allow for a degree of continuity in data collection, and as a result comparability over years of information. In particular cases they can even assist analysts in identifying underlying reasons for variations and changes in acquired data.

At the same time, the information carrier that is used to perform a survey can change over time, without the content of the survey in itself changing. The use of internet, smart phones, etc. can allow for a swifter completion of the questionnaires and a more regular flow of data from respondents to analysts. In some cases, they can assist the respondent in keeping daily trip diaries (i.e. pre-filled trip information offered for verification during the final interview). However, exactly this change could also lead to the exclusion of potentially interesting groups of respondents such as the poor, the elderly, or persons with disabilities. At the same time, it is not clear at the moment whether or not the introduction of new technologies is also more or less cost-efficient than current methodologies.

Because of this, it is currently recommended to not abandon the use of traditional instruments (pen and paper, telephone interview, etc.) for the collection of data through surveys. What can be considered is the parallel usage of traditional and new technologies (smartphone, etc.) for a certain time. This would allow for mutual data validation (checking for consistency), a swifter data flow, increased data usage and, equally important, the building of a working experience with such a new information carrier for surveys.