GOPA
WORLDWIDE CONSULTANTS

Development, implementation and demonstration of a reference processing pipeline for the future production of official statistics based on multiple Mobile Network Operator data (TSS multi-MNO)

Service Contract Number – 2021.0400

Deliverable 2.3: Final version of technical documentation for scenarios, requirements, use cases and methods, and high-level architecture

Volume II – Use cases

In association with:

Istat

positium

NOMMON

Copyright @ 2024 European Union - Licensed under EUPL

Development, implementation and demonstration of a reference processing pipeline for the future production of official statistics based on multiple Mobile Network Operator data (TSS multi-MNO)

Service Contract Number – 2021.0400

Deliverable 2.3: Final version of technical documentation for scenarios, requirements, use cases and methods, and high-level architecture

Volume II – Use cases

Version: final

Date: 8 August 2025

The information and views set out in this publication are those of the authors and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Cover Photo by wd toromc from Pexels

Copyright @ 2024 European Union - Licensed under EUPL

\ ABSTRACT

The Multi-MNO project aimed to **develop, implement and demonstrate a proposal for a reference standard processing pipeline for the future production of official statistics in Europe based on MNO data from multiple operators**. If successful, the proposal developed by the project may be endorsed as European Statistical System (ESS) standard by the relevant ESS bodies. The term "processing pipeline" refers to the combination of a methodological framework and a reference open-source software adhering to such a framework. The processing pipeline developed in this project covers an initial set of use cases; nonetheless, it has been designed to be general enough to provide the flexibility and growth capability required to cover other future use cases. The pipeline has been demonstrated and evaluated on real data from multiple MNOs in various EU countries.

This report defines a set of use cases identifying relevant statistical outputs of interest for National Statistical Institutes (NSIs), evidencing at the same time the potentialities of MNO data. It complements the methodological framework proposed by the Multi-MNO project for the processing of multiple MNO data for official statistics (described in Volume I of this deliverable). This project deliverable focuses exclusively on the conceptual and methodological aspects. The technical specifications, the detailed architecture and the software design are defined in other project deliverables.

Authors: The report was prepared under the technical coordination of Tiziana Tuoto (Senior Researcher, ISTAT). Further contributions, expertise and research support were provided by: (in alphabetical order) Gabriele Ascari, Roberto Cenciotti, Erika Cerasti, Loredana Di Consiglio, Cristina Faricelli, Ricardo Herranz, Paolo Mattera, Miguel Picornell, Roberta Radini, Giorgia Simeoni, Margus Tiru, Villem Tonnison, Luca Valentino and Ivan Vasilyev. Specific contributions for the present population use case and the geolocation and continuous time segmentation methods (in Volumes II and III) were provided by: Edwin de Jonge, Jan van der Laan, Matthias Offermans and Martijn Tennekes.

Acknowledgments: The authors express their gratitude for the technical review and coordination ensured by Fabio Ricciato, responsible for steering the project on behalf of Eurostat.

This report was prepared in the context of the service contract ref. 2021.0400 awarded by Eurostat to the consortium led by GOPA (Germany), in collaboration with the industry partners NOMMON (Spain) and POSITIUM (Estonia), and the National Statistical Institutes ISTAT (Italy) and CBS (Netherlands).

On behalf of the contractor, project management is ensured by Florabela Carausu (GOPA).

Abbreviations

AB	Advisory Board
EC	European Commission
ESS	European Statistical System
EU	European Union
FUA	Functional Urban Area
GDPR	General Data Protection Regulation
LAU	Local Administrative Unit
MCC	Mobile Country Code
MNO	Mobile Network Operator
MS	Member State
NSI	National Statistical Institute
NUTS	Nomenclature of territorial units for statistics
ONA	Other National Authority
SDC	Statistical Disclosure Control
SDG	Sustainable Development Goals
SIM	Subscriber Identity Module
TFMNO	ESS Task Force on the Use of MNO data for Official Statistics
UC	Use Case
UE	Usual Environment
UM	Usual Mobility

Contents

1 Scope and document structure 1
2 MNO data and its potentialities to address statistical needs 2
3 UC 1.A: Present population estimation 4
4 UC 2.A: M-Usual Environment indicators 8
5 UC 2.B: M-Home Location indicators 11
6 UC 2.C: Access to services 15
7 UC 3.A: Usual mobility 19
8 UC 3.B: Commuting 25
9 UC 4: M-Functional Urban Areas and greater city 27
10 UC 5.A: Domestic touristic arrivals and nights spent 30
11 UC 5.B: Domestic same day visits 35
12 UC 5.C: Inbound tourism 39
13 UC 5.D: Outbound tourism 46
14 UC 6: Exposure to risks 51
15 UC 7: Internal migration 54

Index of Figures

Figure 1: Data processing steps: from daily movements between clusters of tiles to the definition of usual movements between administrative areas 20
Figure 2: Logic of the relationship between the input-output of the UE and the UM 21
Figure 3: Potential different international multi-roaming behaviours of a single device 40

Index of Tables

Table 1: UC 1.A – description and targeted statistical outputs 4
Table 2: UC 2.A – description and targeted statistical outputs 9
Table 3: UC 2.B – description and targeted statistical outputs 12
Table 4: UC 2.C – description and targeted statistical outputs 16
Table 5: UC 3.A – description and targeted statistical outputs 22
Table 6: From MNO data to UM outputs 24
Table 7: UC 3 – description and targeted statistical outputs 25
Table 8: UC 4 – description and targeted statistical outputs 27
Table 9: UC 5.A – description and targeted statistical outputs 30
Table 10: UC 5.B – description and targeted statistical outputs 35
Table 11: UC 5.C – description and targeted statistical outputs 42
Table 12: UC 5.D – description and targeted statistical outputs 46
Table 13: UC 6 – description and targeted statistical outputs 52
Table 14: UC 7 – description and targeted statistical outputs 55

1 SCOPE AND DOCUMENT STRUCTURE

This report is the second of a set of three separate volumes that form altogether Deliverable D2.3 (Final version of technical documentation for scenarios, requirements, use cases and methods, and high-level architecture) of the Multi-MNO project. For a better understanding of the content of this volume, we invite the readers to familiarise themselves with the content of Volume I – Detailed scope, requirements and methodological framework.

This document defines a set of use cases (UCs) considered by the project from the perspective of the expected statistical output. The focus is, therefore, on the potentialities which MNO data offer for addressing statistical needs. The description of the UCs focuses on conceptual and methodological aspects, having as basis the high-level description of the data processing flow detailed in Volume I of this deliverable; the software's technical specifications, architecture and design are defined in other project deliverables.

This document is organised as follows:

- **Chapter 2 'MNO data and its potentialities to address statistical needs'** introduces MNO data from the perspective of the purposes it can serve. The general description of MNO data narrows down to the identification of the UCs proposed for detailing and the interconnection between these.
- **Chapters 3 to 15** describe the UCs proposed by the Multi-MNO project. Each use case corresponds to a distinct statistical product from diverse statistical domains, such as population statistics, domestic commuting, spatial statistics, tourism statistics, etc. Most of the UCs described showcase the longitudinal perspective derived from MNO data. In most of the UCs the data processing can be envisaged and standardised at single MNO level and then be integrated in a multi-MNO perspective. On the other side, particularly the UCs on inbound and outbound tourism suggest the case for a more sophisticated integration of data from multiple MNOs.

2 MNO DATA AND ITS POTENTIALITIES TO ADDRESS STATISTICAL NEEDS

In this document, we describe a set of use cases (UCs) considered by the project from the perspective of the **expected statistical output**.

The UCs correspond to different sets of statistical products. We detail UCs related to population presence indicators (e.g. estimates of the number of people present in a certain area at a certain time), as well as to mobility and tourism. We cover statistical products relevant to the following domains:

- population (e.g. present population, resident population);
- spatial statistics (e.g. accessibility to services and exposure to risks);
- usual mobility, local commuting and functional urban areas;
- domestic tourism (i.e. arrivals, nights spent and same-day visits);
- inbound and outbound tourism;
- internal migration.

It is worth noting that in defining the UCs, we intentionally try to follow the **concepts and definitions** already adopted by the domain experts in the ESS and internationally. Thus, for instance, the concepts and definitions of 'tourism' and 'commuting' tend to mimic those already used in official statistics, as well as those of the resident population. Nevertheless, there are many well-known limitations and weaknesses in adapting the MNO data to the official statistical definitions, e.g. MNO data do not specifically provide information on travel **purposes** or intentions[1]. Hence, sometimes the possibility of deriving certain statistical outputs is conditioned by the ability to 'infer' a specific activity from the MNO data, as in the case of commuting, tourism and, to some degree, also for the assessment of the home location. Clearly, activity and purpose are **not directly observable** in the MNO data. It is, therefore, all the more important to properly evaluate the quality of the inference derived from the MNO data and to retain some degree of uncertainty related to the outputs for a proper assessment and usage of the results, according to the principle of *soft classification and rigorous uncertainty assessment*; see Section 4.1 – Fundamental Design Principles in Volume I.

In addition, the added value of MNO data for official statistics lies in the possibility of **exploring new and emerging phenomena** that might remain hidden in the current investigations, based on the current definitions and ongoing surveys and administrative data sources used. For these reasons, in addition to proposing UCs with outputs that are as close as possible to those currently covered by official statistics, we also propose 'variants' of potential outputs, often based on slightly different definitions. These 'new' statistical outputs, which can be derived from the MNO data, even if they are quite far from those currently adopted in the ESS, are intended to be complementary to the traditional ones. Indeed, while we cannot force the MNO data to accurately reproduce an official statistical concept for which they may not be suited, the aim is to use MNO data to investigate behaviour and habits that cannot be derived from other data sources. Based on MNO data, we expect that it will be possible to provide a uniform definition valid for the entire EU based on actual 'objective' behaviour (where a person actually spends the most of their time and is likely to have their social and economic connections in a given long-term

[1] See for example https://www.istat.it/it/files//2023/03/RSU-3-2022.pdf

reference period), rather than on intentions, declarations, registrations and subjective judgements on social connections.

There is no need to specify further that, according to the approach we defined in our reference and demonstrator scenarios (please refer to Volume I of this deliverable), all the proposed statistical products shall be privacy-safe, in accordance with the ongoing privacy guarantees of the MNOs and the Statistical Disclosure Control (SDC) procedures used by the NSIs.

In the development of the UCs, we exploit as much as possible the richness of the MNO data, in terms of temporal and spatial granularity. Hence, the relevant phenomena, like usual environment, commuting, tourism, are 'extracted' and analysed through the **longitudinal observation of the MNO data**, see Section 4.1 – Fundamental Design Principles in Volume I. For example, differently from other research on the topic, 'place of residence' and 'country of residence' are not assigned on the basis of subscriber contract information and of the MCC code of the SIM, respectively; rather, they are derived based on the longitudinal observation of the MNO data, that is in which places and in which country the devices spend the most of their time.

In line with the Pipeline fundamental design principles, a common aspect of the proposed UCs is that the estimation process is designed to produce outputs on a **spatial grid**, e.g. the INSPIRE grid. From these basic outputs, others can be derived at any (larger) territorial representation (e.g. local administrative units, municipalities, any NUTS level classification) by projecting the output from the finest grid level to the geographical unit systems relevant for the specific UC (see for further details Volume III – Methods and data objects). The use of the spatial grid system is a crucial step for the integration of data from multiple MNO and for the general standardisation purpose of this project.

There is no need to underline that both the temporal and spatial dimensions of the analysis are subject to uncertainty, and neither the finest granularity is able to guarantee we would 'infer' activities and purposes without errors. Indeed, there are different activities often concentrated in specific places (sports centres near shopping centres, which may also include a cinema and offices in the same building; universities near hospitals, etc.). In these cases, precautions are always recommended regarding the ability of the MNO data, even in combination with other data source(s), to correctly classify activities related to places according to any strong classification schema.

The target population of the proposed UCs is usually a population of individuals, and a population of areas in spatial statistics. It is never the case, to the best of our knowledge, that NSIs are interested in population of devices. The specification of the target population requires that in all the UCs, the MNO data are integrated with other data sources to be 'transformed' into the target population. The inference from mobile devices to target population might require additional data and potentially sophisticated integration methods. Both the identification of the most promising data sources to be integrated with MNO data and the development of integration methods are not the focus of this project, rather the topics of the parallel ESSnet project MNO-MINDS. Hence, in the description of the UCs, we will not explore them further.

Finally, it is worth to mention that the UCs proposed in this document and their description are intended as inspirational for NSIs. Other interesting statistical outputs or variants/modifications of the ones proposed can be envisaged and pursued. In addition, apart from the general recommendation and specification provided in this document, the concrete shape of the UC output can be cultural- or country-specific, hence requiring an adaptation of the several parametrisation settings that we have proposed. Indeed, the approach of parametrisation we rely on in all the phases of the project, from the methodological design to the code implementation, is meant to facilitate the adaptation and re-use of the methods and tools to fit different output needs. Flexibility in the parameters' setting will allow coping with national/local differences in the specification without compromising the methods' standardisation and the comparability of results.

3 UC 1.A: PRESENT POPULATION ESTIMATION

The objective of this UC is to estimate the number of people present in a specific geographical area at a particular fixed moment in time. This use case does not aim to further classify/label the present population into categories such as locals, visitors, tourists, or passers-by.

The statistical output are population estimates for the desired date and time in a regular spatial grid[2], and by census enumeration area or municipality (or other administrative geographical units), in a given country. The estimates represent the total number of individuals present in each specific area at that particular time, regardless of their residential status or temporary presence. The settings and parameters for the pipeline are described later in Chapter 14.1 of Volume III – Methods and data objects. Regarding the output of this UC, it is useful to note that in order to produce trends or density changes in the population, no further information is needed, e.g. the identification of the resident population (or a proxy of it) in the present population. However, if one is interested in absolute population counts, a classification of the present population into resident population could be needed for calibration purposes. **TABLE 1** provides an overview of the UC definition and output.

Table 1: UC 1.A – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	The count of people (regardless of whether they stay there for a while or they are just passing by) present in a geographical area (100x100 m and polygon) at a given fixed moment in time; e.g. people present on 2nd October 2022 at 10:00 PM by census enumeration area or by municipality in the whole country. See also the dashboards on Dutch population as an example. *Note: In this UC, the population estimates will not be further classified/labelled into locals, visitors, tourists, passing by, etc. Population segmentations will be introduced as an extension of this UC in a later stage.*	
Target statistical population	All persons (including children and elderly) present in a given area at a fixed timestamp.	

[2] as already in place for census-like estimates

USE CASE DEFINITION AND OUTPUTS		REMARKS
Spatial scope	The spatial scope will be chosen by the NSI; it could be an administrative geographical unit as well as a regular grid. In the *demonstrator scenario,* the computation of the output can be limited to some geographical areas or an example of them. Nevertheless, in the *reference scenario,* the methods and tools that we provide would support the whole country.	
Spatial resolution	The reference grid of the internal data processing will be the INSPIRE 100x100 m grid[3]. At the output level, the methods and tools will be flexible to allow the NSIs to choose any zoning system (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	The temporal resolution will be chosen by the NSI. In the *demonstrator scenario,* the output computation can be constrained to specific times and days, to determine the population at fixed intervals like 8:00 AM, 10:00 AM, 12:00 PM, 4:00 PM, etc. on dates like 2nd October 2022, 25th December 2022, and so forth. In the *reference scenario* the methods and tools that we provide would support the indicator to be produced for every single day and time. A set of pre-defined reference times will be considered (e.g. every 2 hours) while the possibility to select different times 'on demand' will be considered as an optional feature.	
Possible additional classification of the target statistical population	In this UC, the population estimates will not be further classified. The estimation of any possible classification is a potential extension/variation of this UC.	Additional possible classifications useful for enriching the value of the indicators produced can be: • Place of residence • Point of entry • Residents / commuters / tourists / other visitors • Type of activity performed in the area: home, work, transiting (i.e. people that transit Area i during the Time Window j, but do not stop there to perform any activity), tourist, other activities, etc. • Time spent in the area

[3] https://knowledge-base.inspire.ec.europa.eu/index_en

USE CASE DEFINITION AND OUTPUTS		REMARKS
		• Frequency of visits (i.e. what % of people visit the area at least X times per week, between X and Y times per week, etc.)
Possible extensions and variations	1. Present population in a given space and in a given time **interval**	
Target NSI departments	• Demography and population • Census • Social policy • Safety • National accounts (environmental accounts)	These estimates would provide insights into the overall population distribution and density at a particular moment, allowing for analysis and decision-making in various fields, including: 1. **Demography and population**: The data can be used for demographic analysis; e.g. understanding population growth rates, age structure, and internal migration patterns. This information is valuable for studying population trends and planning social services and policies. 2. **Census**: The population estimates can support census efforts by providing a snapshot of the population in a specific area at a particular time. This data can be compared with official census figures to assess accuracy and identify areas for further investigation. 3. **Social policy**: The information can be used to assess the impact of social policies and programs on the population. By understanding spatial population distribution and characteristics, policymakers can better design and target social policies to address specific needs and ensure equal access to services. 4. **Safety**: The population estimates can contribute to assessing safety and security needs. Understanding population density in different areas and at specific times can help to identify areas with higher demands for law enforcement, emergency services, and infrastructure to ensure public safety. 5. **National accounts (environmental accounts)**: The data on population distribution and density can be integrated into national and environmental accounts. It allows for a comprehensive understanding of the economic and environmental aspects influenced by population factors, such as resource consumption, environmental impact, and sustainable development. By incorporating these additional dimensions, the population estimates become even more valuable for policy formulation, resource allocation, and decision-making across various domains, promoting effective and sustainable development.
Target users of the new statistical product	• Urban planners • Environmental health specialists • Policy makers at local and national levels	1. **Urban planners**: The data assists urban planners in understanding population patterns and trends, aiding in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to accommodate population needs.

USE CASE DEFINITION AND OUTPUTS	REMARKS
• Safety and disaster management bodies	2. **Environmental health specialists**: The information supports environmental health specialists in assessing the impact of population density on environmental factors, such as air quality, noise pollution and waste management. It helps in identifying areas where interventions are needed to improve environmental health conditions. 3. **Policymakers at local and national levels**: The population estimates serve as a crucial input for policymakers in formulating strategies and policies related to housing, healthcare, education, and social services. They provide a basis for resource allocation and decision-making to address the specific needs of different population groups. 4. **Safety and disaster management bodies**: The data aids safety and disaster management professionals in understanding population distribution for effective emergency response planning. It assists in identifying areas with high population density that may require additional safety measures or evacuation plans in the event of natural disasters or emergencies.[4] 5. **Non-resident populations**: Census data may face challenges in accurately capturing non-resident populations, such as tourists, commuters, or temporary residents. MNO data can help to identify and estimate the presence of these populations based on their mobile phone activity, providing a more comprehensive picture of the population in a given area.

[4] See for example: https://www.easa.europa.eu/en/newsroom-and-events/events/workshop-population-density-services-uas-operations

4 UC 2.A: M-USUAL ENVIRONMENT INDICATORS

In the field of statistics, the term 'usual environment' (UE) refers to the specific location or setting where individuals typically spend the majority of their day or where they have their primary activities and interactions.[5] This UC aims to identify the main areas where each single person spends his/her time and performs his/her activities; i.e. an MNO data-based User Environment, in short M-User Environment. The statistical concept of 'usual environment' lends itself to the generation of a group of multiple UCs, based on the potential discrimination of places, activities, time, people involved and interaction among all these elements. For this reason, this UC is proposed as the first one, under the cluster of UCs that also collects the 'M-home location' and 'access to services' ones. In addition, the statistical outputs related to commuting and tourism are strictly dependent on the concept of 'usual environment', as we will specify and clarify in the sections devoted to them. Finally, also the UC on Exposure to risks can benefit from the definition of the M-Usual Environment indicators and its related concepts, as the M-Home Location indicators.

It is worth clarifying the potential ambiguity between the 'usual environment of an individual mobile device' and the 'set of aggregate indicators derived from the individual data elements'. The first, is a data element associated to an individual user; namely, an intermediate data element for the indicators that are derived from the downstream of the UC (i.e. the M-Usual Environment indicators). For example, one could count people by location and derive some M-FUA indicators in this way. However, it is also possible to use the (individual) UE to identify individual touristic trips (all trips falling outside the individual UEs) and, from there, derive other indicators related to tourism statistics. Therefore, the 'usual environment' as **intermediate data element** is then used to derive multiple UCs in different domains including the **genuine use case** M-Usual Environment indicators. This is the reasoning behind presenting the M-Usual Environment indicators UC as first in the list of the UCs, i.e. due to its link with the other UCs and its sufficiency to produce indicators specifically designed to describe the statistical concept of 'usual environment'.

Under this premise, it is important to note that the possibility of deriving certain statistical outputs from the UE concept is sometimes conditioned by the ability to infer the specific activity of interest from the MNO data, as in the case of commuting, tourism and, to some degree, also for the assessment of the home location. To this extent, we have already underlined how the **purpose** of 'presences' and 'movements' **is not directly observable** in the MNO data. It is therefore even more important to properly evaluate the quality of the inference derived from the MNO data and to retain some degree of uncertainty related to the outputs for a proper assessment and usage of the results.

Despite the due caution in classifying and distinguishing 'purposes' in the M-Usual Environment, the identification of an individual UE, without any further classification, is *per sè* an extremely valuable outcome for official statistics. This is due to the fact that it is complicated and expensive to derive such a measure from traditional data sources, and that many statistical outputs can be generated even without further classification of the individual UE. For instance, when the 'home location' can be derived from MNO data only with high uncertainty, it could be useful to rely on the individual UE, if it can be assessed with a lower uncertainty than the 'home location'. Clearly, in addition to the 'home location', any other breakdowns of the individual UE may be inaccurate, and in these

[5] See: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Usual_environment

instances keeping the wider concept of UE rather than the more detailed and more uncertain breakdown/classification (e.g. 'home location') would be advisable.

Table 2: UC 2.A – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	Collection of places/areas, such as home-place, work/study-place, and other places visited recurrently with or without a label assigned. The statistical output can be: 1. by regular square grid or municipality; for example, how many people have their M-UE completely contained in their home municipality? How many people gravitate in the area, even if it is not their home municipality? Classification of this M-UE 'other than home' by daily time, weekday, frequency, and purpose of the gravitation (e.g. work/study, second home, etc.); 2. the prevalent classification (tagging) of each area identified in a regular square grid or in a municipality (e.g. census enumeration area or other administrative territorial units); 3. how many people have their UE split into one, two, three or more municipalities and how far these usual places are from each other (e.g. within Rome there are places that are further away than two different municipalities or areas of other provinces) M-UE UCs may have some focus on big cities or other target areas, like municipalities on the coast where holiday houses are dominant, etc.	Location classifications, potentially identifiable with MNO data: 1. Home: this category represents the place where the individual usually resides and spends the majority of their time. 2. Second home: this category refers to individuals who have a secondary residence or vacation home in addition to their primary residence. 3. Work/Study: this category includes individuals' usual places of work or study, such as offices, factories, schools, universities, or other educational institutions. 4. 'Unspecified usual places': meaning places that are visited regularly by the mobile user and therefore are included in the UE, although we are not able to classify and associate these, with reasonable confidence, to a specific use. In principle, there are additional categories we might define: leisure/recreation (e.g. parks, sports facilities, entertainment venues, and shopping centres) or transportation (e.g. train stations, bus stops, airports, and main roadways, where people spend time during their daily commute), for example. For these additional categories, as well as for the basic ones listed in this section, we will explore the suitability of MNO data, in combination with other data sources, to assess them, along with the measures of uncertainty related to the classification itself, according to the principle of *Soft classification and rigorous uncertainty assessment*; see Section 4.1 – Fundamental Design Principles in Volume I.
Target statistical population	In the output of this UC, the statistical units under study are places. People counts and activities can be seen as attributes of places. Only people usually gravitating to the selected places are of interest to this UC. Different characterisations of places and related activities may produce extensions of this UC.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario,* the methods and tools that we provide would support the whole country. In the *demonstrator scenario*, we may limit the computation of the output to some specific areas. This point will be discussed jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	
Spatial resolution	The data processing will use the INSPIRE 100x100 m grid. The output will be any requested zoning system configurable by the NSI (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	The methods and tools shall be flexible to allow NSIs to choose any temporal scope and resolution. In the *reference scenario*, the methods and tools that we will provide would support the computation of the indicator at least considering the potential seasonality of the UE. The suggested option is 12 months. In the *demonstrator scenario*, we will discuss the temporal scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing)	Whatever the temporal scope and resolution of the output will be, the UE concept will require a longitudinal observation of the users to consider potential seasonality. This is the main reason for our recommendation of taking 12 months of observations. Apart from the general recommendation, since the concrete form of the UE concept can be cultural- or country-specific, our methods and tools will allow some flexibility in the parameters' setting, in order to cope with national differences.
Possible additional classification of the target statistical population	UE is already a kind of classification/tagging. No additional classification will be considered for this UC.	
Possible extensions and variations	1. Total time spent in target places 2. Average time spent in target places, e.g. in a day, in a week, etc. 3. UC - 2.C Access to services and UC – 6 Exposure to risks	
Target NSI departments	• Demography and population • Census • Social policy • Environment	
Target users of the new statistical product	• Urban planners • Policymakers at local and national levels	

5 UC 2.B: M-HOME LOCATION INDICATORS

The usual resident population refers to persons who have their 'usual residence' in a specific geographic area, which is the place where a person normally spends the daily period of rest, regardless of temporary absences for purposes of recreation, holidays, visits to friends and relatives, business, medical treatment or religious pilgrimage. The usual residence is defined as the place where a person has lived for a continuous period of at least 12 months before the reference date; or where the person arrived during the 12 months before the reference date with the intention of staying for at least one year.

The 'usually resident population' concept[6] takes into account the long-term residence and attachment to a specific geographic area, rather than focusing on temporary or short-term stays. It is a key concept used in official population statistics to measure the population of an area, in order to derive demographic and socio-economic insights.

Bearing in mind the official definition for 'usually resident population', this UC aims to provide an MNO-based approximation for the official statistics concept of resident population (i.e. M-Home Location indicators).

While we cannot force the MNO data to accurately reproduce an official statistical concept for which they may not be suited, the aim is to use MNO data to investigate behaviour and habits that cannot be derived from other data sources. Based on MNO data, we expect that it will be possible to provide a uniform definition valid for the entire EU, based on actual 'objective' behaviour (where a person actually spends the most of their time and is likely to have their social and economic connections in a given long-term reference period), rather than on intentions, declarations, registrations and subjective judgements on social connections.

Therefore, the aim of this UC is to produce estimated counts of people whose home place derived from the MNO data (i.e. the 'M-home location') is located in a given space at a given reference time, broken down by a regular square grid (e.g. 1km x 1km), as already done for census grid data, as well as by municipality or census enumeration areas.

To some extent, the M-home location concept, i.e. the residence derived based on MNO data, can be seen as the '*de facto*' population concept; that is, the MNO data provide the representation of the population where they are usually and regularly found, regardless of the official and recorded place of residence. On the other side, the usual resident population in use at NSIs and derived from population censuses, administrative registers, statistical population registers, can be seen as the '*de jure*' population concept; that is, the place where people declare to be located/registered, regardless of where they actually spend the majority of their time. The identification of two different concepts, the 'de facto' and 'de jure' population, help in clarifying that the corresponding measures will be inherently different. Only by chance the M-home location counts can reproduce the official statistics counts for the usual population. However, these numerical discrepancies are intrinsic to the different definitions and cannot be reconciled without removing the definition differences.

Finally, it is useful to note that the 'M-Home Location indicators' UC is not properly a use case *per sè*. Rather, it is one of the main outputs of the 'M-Usual Environment indicators' UC. Nevertheless, given the centrality of the resident population concept in many domains of official statistics, and the role that 'home location' plays as a proxy

[6] See: usually resident population (europa.eu)

for the resident population when dealing with MNO data, we decided to devote a dedicated effort to it, and highlight it in a separate section, as a specific product of the 'M-Usual Environment indicators' UC.

Table 3: UC 2.B – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	Estimated counts of individuals whose home location (a proxy for their place of residence) is situated within a specified area at a given reference time. These counts will be disaggregated by regular square grids, municipalities, census enumeration areas, or any other administrative or meaningful zoning system, provided that it is not finer than the INSPIRE 100x100 m grid. E.g. people with home location in Italy at the reference time 31st December 2022, by regular square grid 1km x 1km, by census enumeration area, or by municipality.	How to combine the information extracted from MNO data with census data or other non-MNO data to produce results for the total population is only tangent to this project. This project focuses on describing a process pipeline that encompasses this integration step. However, where state-of-the-art solutions do not provide satisfactory answers, the data workflow will be flexible and evolvable to implement more advanced solutions once they will become available. The integration of MNO data and non-MNO data sources is currently the main topic of the ESSnet MNO-MINDS research project. Therefore, in our demonstrator scenario we may apply a simplified method and leave more sophisticated methods to be assessed by the other project. Since NSIs might be interested in MNO data results before their integration with other data sources, in our demonstrator scenario we may allow the possibility to derive percentages and relative figures only based on MNO data.
Target statistical population	All persons with M-home location (proxy of the place of residence) in a country in a reference time.	It might be useful to consider some intermediate results on mobile devices as final outputs, before the inference to the whole population, for two main reasons: • counts on mobile users allow NSIs to perform some estimates on flows, in relative figures. As well, they can be used as auxiliary information in model-based approaches to improve the timeliness and the spatial granularity of official figures. • this is in line with a 'light' approach for the data workflow's 'Estimation' module detailed in Chapter 21 of Volume III. However, as often clarified, the inference from mobile devices to individuals might require additional more sophisticated methods and their investigation is not the focus of this project, rather the proposed pipeline will be ready to implement them once they will be developed; for example in the parallel ESSnet project MNO-MINDS.
Spatial scope	The spatial scope will be chosen by the NSI. In the *demonstrator scenario,* the output computation can be limited to some	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	geographical areas. We will discuss the spatial scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing). Nevertheless, in the *reference scenario*, the methods and tools provided would support the whole country.	
Spatial resolution	The spatial resolution of data processing will be the INSPIRE 100x100 m grid. The output spatial resolution can be any zoning system (e.g. square grid, administrative divisions, etc.), provided that it is not finer than the INSPIRE 100x100 m grid.	
Temporal scope and resolution	The methods and tools shall be flexible to let the NSIs choose any temporal resolution.	Regardless of the output temporal resolution, the 'M-home location' concept will require a longitudinal observation of the devices. Our recommendation is of at least 12 months (i.e. the length of observations will be 12 months). This recommended value is derived by similarity with the official definition of usual residence.
Possible additional classification of the target statistical population	Assigning a home location in the country is already a type of classification/tagging. In this UC, we will not consider any further classification other than the home location (as a proxy of the usual residence).	
Target NSI departments	• Demography and population • Census • Social policy • Environment	The fields of use for a proxy of resident population data are very diverse and extensive. However, in this study we will focus on its relevance to the census itself. More specific points are discussed below: • **Timeliness and frequency**: Census data is typically collected once every few years. Nonetheless, NSIs usually exploit administrative data and/or statistical models to produce annual demographic and migration statistics, as well as some infra-annual (bi-annual or quarterly) population products at national level exist. Since MNO data provide real-time or near real-time information, they may support and complement the production of more frequent updates on population movements and changes, allowing for timely and responsive planning and decision-making. • **Granularity and spatial accuracy**: MNO data can provide highly granular and geographically precise information about population

USE CASE DEFINITION AND OUTPUTS	REMARKS
	distribution. It can capture population movements at the individual level offering insights into population densities in specific areas. This level of detail helps to identify population hotspots, areas of high transit, or areas with rapidly changing populations. • **Behavioural insights**: MNO data can provide behavioural insights by analysing movement patterns, transport modes, and social interactions. This information can complement census data by offering a deeper understanding of how people use urban spaces, commute, and engage in various activities. It can inform urban planning, transport infrastructure development, and resource allocation.
Target users of the new statistical product	• Urban planners • Policymakers at local and national level

6 UC 2.C: ACCESS TO SERVICES

Access to services' refers to the measurement and analysis of the ability of individuals to reach and use essential services within a given geographical area. Principle 20 of the European Pillar of Social Rights[7] highlights that everyone has the right to access essential services of good quality that fulfil basic human needs and are key to well-being and social inclusion. Essential services that are key to active participation in the society and the labour market include early childhood education and care, education and training, healthcare, long-term care and social inclusion services. Access is a complex concept. It encompasses different aspects that need to be assessed, including financial, organisational and social or cultural factors, such as geographical distance, travel time to the nearest location of the service, adequate facilities and supply, sufficient delivery channels, provision of information in simple languages also for non-native speakers, convenient opening hours and reduced administrative burdens, affordability or the economic cost of purchasing the service relative to income. Clearly not all of these factors can be measured using MNO data. In this UC, we aim to measure the ability to access a service, taking into account the specific aspect of the geographical distance of the service from the home location of the person. In this UC, we apply the definition of MNO-home location as proposed in Chapter 5 UC 2.B: M-Home Location indicators and a binary evaluation, i.e. distance=0 if the service is placed within 1km from the MNO-home location, distance=1 if the service is placed further than 1km from the home location. It is clear that having a service 'close to home' does not 'guarantee' the access to it and its usage; at least, this proximity potentially "facilitates" service access and use.

It is worth noting that different indicators can be added (and compared), e.g. enlarging the observation of the distance from the MNO-home location to the Usual Environment of the person, as provided by the use case in Chapter 4 UC 2.A: M-Usual Environment indicators. Non-binary distance evaluation is a possibility as well. The outcome can also be enriched by taking into account the travel time, the calculation of which may vary in complexity, from the application of a set speed to the shortest path through a network, to the use of public transport timetables.

Regarding service classification, as mentioned above, they encompass several typologies, e.g. healthcare, education, transportation, culture and recreation. Some of them are easy to detect and locate on the territory. They are often concentrated in specific places (sports centres near shopping centres, which may also include a cinema and offices in the same building; universities near hospitals, etc.). In these cases, some precautions are recommended, particularly regarding the ability of the MNO data, in combination with other data source(s), to correctly classify places according to any classification schema.

In this UC, we will limit our attention to the binary distance of the home location from the nearest services, investigating health services and green areas as examples. However, we know that hospitals do not all provide the same treatments, as schools do not all provide the same quality of education. Potential different future choices can be represented by weighting the opportunities available and factoring in the travel time (and cost).

[7] See: https://ec.europa.eu/social/main.jsp?catId=1592&langId=en

Table 4: UC 2.C – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	Count of people potentially able to access a given service, e.g.: 1. count of population that have access to green areas within 1km from their M-home location; 2. count of population that have access to health services within 10km from their M-home location. Other measures of interest that can be potentially estimated but will not be investigated further during this project: 1. Proximity and Travel Time; 2. Service utilisation; 3. Disparities; i.e. the statistical analysis of access to services can also reveal inequalities or disparities in service provision across different population groups or geographic areas. It helps identify marginalised or underserved populations and can guide efforts to address these disparities through targeted interventions and policy initiatives.	Potential classification of services for which the MNO data can provide insights: 1. Healthcare services: this category includes access to hospitals, clinics, primary care centres, specialised medical facilities, pharmacies, and healthcare professionals. It encompasses services such as preventive care, diagnostics, treatment, and emergency medical services. 2. Education services: this category includes access to schools, colleges, universities, vocational training centres, and educational resources. It encompasses formal education as well as adult education and lifelong learning opportunities. 3. Transportation services: this category refers to access to transportation infrastructure and services, such as roads, public transit systems, railways, airports, and ports. It includes the availability of affordable and efficient transportation options for commuting, travel, and the movement of goods and services. 4. Utilities and infrastructure: this category includes access to basic utilities and infrastructure necessary for daily living, such as clean water and sanitation facilities, electricity, gas, and telecommunications services. It encompasses infrastructure for housing, waste management, and other essential public services. 5. Social services: this category includes access to social welfare programs, social assistance, childcare services, elderly care facilities, disability support services, and community centres. It encompasses services aimed at supporting vulnerable populations and promoting social well-being. 6. Financial services: this category refers to access to banking services, financial institutions, credit facilities, and insurance services. It includes the availability of financial resources and services for personal and business needs. 7. Cultural and recreational services: this category includes access to cultural institutions, libraries, museums, parks, recreational facilities, sports centres, and entertainment venues. It encompasses opportunities for leisure, cultural enrichment, and community engagement.

USE CASE DEFINITION AND OUTPUTS		REMARKS
		Cautionary remarks on the ability of the MNO data, in combination with other data source(s), to properly classify places according to the wide classification schema proposed in this section applies even more to this use case. Again, we will assess and eventually revise the proposed classification in Task 5 (testing) and as already stated we will accompany the use case results with measures of the uncertainty related to the classification itself, according to the principle of *Soft classification and rigorous uncertainty assessment*, see Section 4.1 – Fundamental Design Principles in Volume I.
Target statistical population	All persons with M-home location (proxy of the place of residence) in a country at a reference time. Different characterisations of places and related activities may produce extensions of this UC.	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	
Spatial resolution	The data processing will use the INSPIRE 100x100 m grid. The output will be flexible to let the NSIs choose any zoning system (e.g., square grid, administrative divisions, etc.)	
Temporal scope and resolution	The methods and tools shall be flexible to let the NSIs choose any temporal scope and resolution. In the *reference scenario*, the provided methods and tools would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	Whatever the temporal resolution output will be, the 'access to service' concept requires a proper longitudinal observation of the devices, to assess the individual's M-home location.
Possible additional classification of the	The classification of places in terms of the services they offer is crucial for this UC.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
target statistical population	This classification should be derived by combining MNO data with external / publicly available sources.	
Target NSI departments	• Demography and population • Census • Environment • Social policy (SDG, equitable and sustainable well-being indicators)	
Target users of the new statistical product	• Urban planners • Policymakers at local and national levels	

7 UC 3.A: USUAL MOBILITY

The UC on Usual Mobility (UM) encompasses all the movements of mobile devices within their Usual Environment. Hence, this UC is strongly correlated with the definitions in Chapter 4 UC 2.A: M-Usual Environment indicators. A fundamental concept in the scope of this UC is the 'path', defined as an ordered pair of stay places, namely 'origin' and 'destination' of a mobile device displacement. Notably, in the path definition only the origin and destination are considered, while the route covered to accomplish the displacement is disregarded.

The following criteria characterise a usual movement, defined as displacement with a valid path in the scope of the present UC:

- Both the origin and the destination defining the path must belong to the UE of the mobile device;
- The movement must occur within a single day.

For the **destination to be valid,** in the context of usual movements, the following additional criteria are considered, beyond its belonging to the UE of the device:

- Exclusion of reflexive paths, i.e. the destination does not coincide with the origin of the path.
- Length of stay: the mobile device remains at the destination for at least three consecutive hours during the day. This parameter can be tuned according to the activity to be analysed.
- Frequency of presence: the mobile device exceeds a threshold for the number of times the destination is visited in a given time window (e.g. 30 times in a time window of 9 weeks or 14 times in a time window of 5 weeks; to be decided after stakeholders feedback).

Once these conditions, for the path and the destination, are met, the path is considered a **usual movement**, characterised by the following **attributes**:

- the ordered pair of locations involved in the movement, each different from one another;
- the duration of the movement;
- the day of the week of occurrence and its classification (weekday, weekend);
- the time interval at the beginning of the movement over the day of occurrence (morning, afternoon, evening, night);
- the time interval at the end of the movement (morning, afternoon, evening, night).

The **inputs** for this UC are a set of places defined as UE for a mobile device, according to UC 2.A: M-Usual Environment indicators. Once received the input, the present UC analyses all the daily data observations for the mobile device to generate the daily sequence of tiles visited with a relevant permanence score.

Subsequently, all the daily paths that include locations with a high permanence score are **calculated**. At this stage, reflexive paths must be filtered out of the dataset.

Then, only the paths that meet the following criteria are **selected**: (i) exceed the frequency of the presence threshold and (ii) whose destinations exceed the length of stay threshold. No restrictions are applied for origin locations. The selection activity is applied daily, using the daily permanence score, and also on a long-term longitudinal observation among all the observed paths.

This selection can be implemented according to criteria defined by the specific mobility of interest using the frequency, the average stay time and the average travel time, stored as indicators for each path of the mobile

device and the related thresholds. The paths can be distinguished by the daily movement time interval of the origin and destination (morning, afternoon, evening, night) and the days of the week (weekday, weekend).

Once the usual movement paths are selected, the origin and destination are **transformed** from a group of tiles into an administrative territory (e.g. municipality, sub-municipal area, census area), or into a partition of the territory of interest for statistics (e.g. 1 km regular grids).

In the last step, the mobile devices are **aggregated** for each path with the same analysis dimensions defined (e.g. the origin-destination matrix during morning time over weekends).

The operations on locations and paths are illustrated in the figure below:

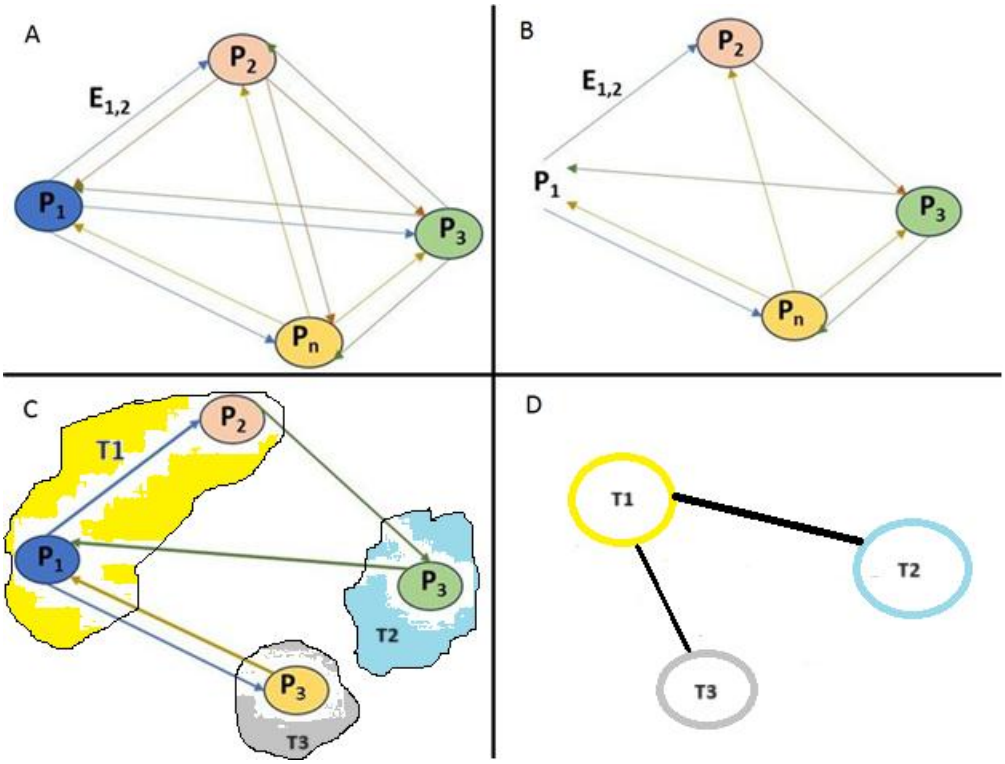

Figure 1: Data processing steps: from daily movements between clusters of tiles to the definition of usual movements between administrative areas

In panel A, all locations with high permanence score and the paths interconnecting these locations are **calculated**. In panel B, the paths that exceed the frequency of presence threshold are **selected**. In panel C, a different selection criterion on the destinations that exceed the length of stay threshold shows a different possible collection of paths, with less samples left. T Areas represent territorial **transformation**. Panel D shows an example of the **usual mobility between statistical areas** chosen to define the output.

The **output** will be an origin-destination matrix featuring the defined statistical indicators, as described in the table at the end of this chapter.

To clarify the relationship between the input-output of the UE and the UM, as well as the use by the UM of the intermediate products of the UE (such as the daily and long-term permanence score), we illustrate the logical flow of the data in **FIGURE 2**. The hypothesis illustrated shall be detailed and validated during the definition of the methods and the design of the software components for the pipeline's application to the UM UC.

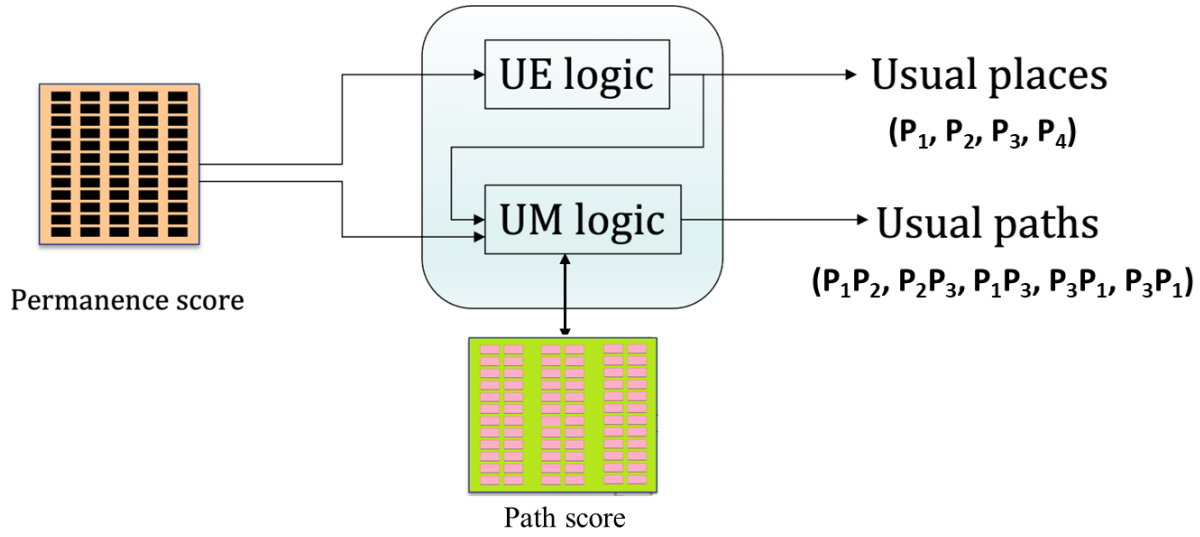

Figure 2: Logic of the relationship between the input-output of the UE and the UM

Table 5: UC 3.A – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	Indicators are calculated per each usual movement as: • number of mobile devices; • average time duration between origin to destination; • spatial distance between origin and destination. **Definitions**: • a path is defined as an ordered pair of stay places representing the origin and the destination of an individual's movement made within a single day. An usual movement is a path satisfying the following criteria: • the origin and destination places of the displacement are both inside the device UE; • The duration of a stop at an origin or destination must last at least 3 **consecutive** hours during the day; • the movement occurs within one day; • the movement needs to be recurring/frequent, i.e. usual. ***Note**: the regularity is detected setting a threshold for the number of times the location is visited in a given time window (e.g. at least 30 times in an observation time window of 9 weeks or 14 times in a time window of 5 week – to be discussed based on stakeholder feedback). The observation time window in weeks is a moving window.*	
Target statistical population	The statistical unit is the device resident in the considered country. The population under study includes all the devices having a UE in the country area. (see UC 2.A: M-Usual Environment indicators).	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	The spatial scope consists of LAUs/grid cells of specific length. In large cities, we can use the aggregate of the census area to define movements within the city.
Spatial resolution	The data processing will use the INSPIRE 100x100 m grid.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	The output will be flexible to let the NSIs choose any zoning system (e.g., square grid, administrative divisions, etc.).	
Temporal scope and resolution	The methods and tools shall be flexible to let the NSIs choose any temporal scope and resolution at output level. The default temporal scope is set to one month. The temporal resolution of the data processing is the one provided by the input MNO data. In the *reference scenario*, the methods and tools provided would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	Whatever the temporal resolution output will be, the 'usual mobility' concept requires a proper longitudinal observation of the devices, to assess the device Usual Environment.
Possible additional classification of the target statistical population	In this UC, usual movements are not further classified. The estimation of any possible usual movement classification can be seen as a potential extension/variation of this UC (to be discussed after stakeholder feedback).	Potential classifications, useful for enriching the value of the indicators produced, can be: • Breakdowns by weekend day or per sub-daily time intervals (morning, mid-day, afternoon, night). • Breakdowns by distance in km or by time duration of the movements (short, medium, long). • Breakdowns by degree of urbanisation of origin-destination areas or by other features of origin-destination areas (e.g. FUA, internal area, accessibility, degree of infrastructure development, etc.). • Breakdowns by purpose of the movement, subject to using additional data on the characterisation of areas or UE labels (e.g. school, university, workplace, leisure, unspecified, usual environment).
Target NSI departments	• Demography and population • Mobility • Environment and territory • Welfare	
Target users of the new statistical product	• Urban planners • Environmental specialists • Health specialists • Transportation planning • Managers of mobility services • Risk managers	

The table below clarifies further how a real-life scenario reflected in MNO data translates into outputs of this UC.

Table 6: From MNO data to UM outputs

REAL-LIFE SCENARIO REFLECTED IN DATA	WHAT IS MEASURED / WHAT IS OUTCOME IN THIS UC?
Home → outside of UE (<3hrs) → work (>3hrs) → Home	We observe 2 paths: Home → work, work → Home We calculate the indicators (described in Table 6 as output of the UC) for the above movements only if they are frequent
Home → outside of UE (>3hrs) → work (>3hrs) → Home	We observe 2 paths: Home → work, work → Home We calculate the indicators (described in Table 6 as output of the UC) for the above movements only if they are frequent
Home → outside of UE (>3hrs) → work (>3hrs) → outside of UE (for unspecified time, does not return home on that day)	We observe 1 path: Home → work We calculate the indicators (described in Table 6 as output of the UC) for the above movement only if it is frequent
Home → work (<3hrs) → Home	We don't observe any path

8 UC 3.B: COMMUTING

Commuters are workers and students *travelling from home to work or study places and back in a single day on a regular basis*. This is the definition adopted by NSIs. Additionally, some NSIs operationalise the commuter Census definition by requiring the travel to be done usually at *least 3 times a week*.

As already stated, from the MNO data it is not possible to explicitly define the purpose of the movement, as well as it is quite difficult to statistically infer it. Therefore, in this UC we consider a specific sub-case of the movements, i.e. the movements from the place defined 'home' in the UE to another place within the UE, with some specific characteristics in terms of temporal regularity such as time intervals in the day, duration, distribution over the days of the week.

The selected places of destination should have similar features to places where daily activities, such as work and study, are carried out. This movement involves returning home during the day. Thus, in this UC, the focus is on the identification of areas of long permanence and the subsequent analysis over a long period of time to identify possible patterns.

The UC can be expanded by considering the route as the sequence of paths travelled from origin to destination and vice versa. Considering the path travelled, the analyses can be expanded to estimate the speed and the type of vehicle used (i.e. by road, train, plane, ship, or simply motorised and not motorised vehicles). These expansions are recognised as complex ones to be evaluated through MNO data and might require additional resources compared to the ones allocated to this project. Even if these detailed analyses would be considered as lower priority for official statistics at the time of this project, it is important to highlight that the pipeline architecture is designed to support also these more complex UCs (e.g. by adding more specialised daily processing sub-modules). This enables the MNOs to enrich the reference pipeline with some additional modules in order to implement commercial analytics products alongside official statistics, based on the same pipeline.

Table 7: UC 3 – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS

Indicator(s) to be produced	Indicators are calculated per each origin-destination movement as:
	• number of people per origin-destination matrix of movements
	  • Commuting duration
	  • Commuting Spatial distance
	  • Usual 'time period' of departure
	  • Usual day of the week
		Definition of thresholds, patterns (how many days over the week, 'pseudo' commuters where the movement is done e.g. weekly), etc.
		In this UC we define:
		• Time-period - is a time slot of the day, which is labelled as, for example: morning, lunchtime, afternoon, evening, night, etc.
		• Days of the week - can be aggregated, for instance as weekdays and weekend.
		We will discuss the semantic vocabulary of 'time-period' and 'day of the week' with the ESS-members and interested stakeholders (e.g. the TFMNO, the Advisory Board) and will inform in the next release of this report about the outcome of the discussion on this choice.

USE CASE DEFINITION AND OUTPUTS		REMARKS
Target statistical population	Individuals / Anyone subject to mobility, e.g. students, workers.	
Spatial scope	The spatial scope will be chosen by the NSI. In the *demonstrator scenario,* the computation of the output can be limited to some geographical areas. We will discuss the spatial scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing). Nevertheless, in the *reference scenario*, the methods and tools provided would support the whole country.	
Spatial resolution	The spatial resolution of data processing will be the INSPIRE 100m x100 m grid. The output spatial resolution can be any zoning system (e.g. square grid, administrative divisions, etc.), provided that it is not finer than the INSPIRE 100x100 m grid.	
Temporal scope and resolution	The methods and tools shall be flexible to let the NSIs choose any temporal resolution. In principle, the UC measures phenomena with slow temporal variability; therefore, an annual observation can be adopted.	Regardless of the output temporal resolution, the 'commuting' concept will require a longitudinal observation of the devices for at least 12 months (i.e. the length of observation will be 12 months).
Possible additional classification of the target statistical population	Assigning a commuting label to a movement is already a type of classification/tagging. In this UC, we will not consider any further classification.	
Target NSI departments	• Demography and population • Environment and territory • Mobility • Welfare	
Target users of the new statistical product	• Urban planners • Environmental specialists • Health specialists • Transportation planning • Managers of mobility services • Risk managers • Train companies • Electric vehicles and car renting companies	

9 UC 4: M-FUNCTIONAL URBAN AREAS AND M-GREATER CITY

A functional urban area (FUA) consists of a city and its commuting zone. FUAs, therefore, consist of a densely inhabited city and a less densely populated commuting zone with a labour market that is highly integrated with the city (OECD, 2012). In the official definition[8], a city is a local administrative unit (LAU) where the majority of the population lives in an urban centre of at least 50 000 inhabitants and a commuting zone contains the surrounding travel-to-work areas of a city where at least 15% of employed residents work in the city. The FUA definition introduces a new concept, the 'Greater City', which is linked not only to population density but, above all, to the movement flows of its inhabitants between contiguous areas, regardless of administrative and, in some cases, even national borders.

This UC studies the possibility of measuring the dynamics of these areas in a more timely manner, analysing how the development of mobility infrastructures, production areas or environmental situations affect the distribution of the population on the territory, their regular movements and the mobility flow of workers.

Table 8: UC 4 – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	For each M-Greater City and M-FUA, the statistical output can be, e.g.: 1. the territorial variation over time; 2. the estimate of flow volumes by time (daily, by hours of the day); 3. the territorial classification: city, commuting zone, and urban centre.	The statistical output that can be obtained for each M-Greater City and M-FUA includes: 1. Territorial variation over time: this analysis examines the changes in population distribution, activities, and worker/student mobility within the FUAs and Greater Cities over time. It provides insights into how these areas have evolved, highlighting trends, growth patterns, and shifts in population and economic activities. The territorial variation can be measured yearly, comparing data from different years to observe changes and trends. 2. Estimate of flow volumes by time: this analysis focuses on understanding the mobility patterns and flow volumes of workers/students within the FUAs and Greater Cities. It provides information on the movement of workers/students throughout the day, allowing for a detailed understanding of peak hours, congestion patterns, and the intensity of commuting flows. Flow volumes can be estimated on a daily basis,

[8] See: Archive:European cities – the EU-OECD functional urban area definition - Statistics Explained (europa.eu);
Methodological manual on territorial typologies, 2018: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-008

USE CASE DEFINITION AND OUTPUTS		REMARKS
		broken down by hours of the day, to capture the temporal dynamics of worker/student mobility. 3. Territorial classification: this classification categorises the different areas within FUAs and Greater Cities based on their characteristics and functions. It includes the classification of city areas, commuting zones, and urban centres. City areas typically have a high population density and serve as the core administrative units, while commuting zones encompass the surrounding areas where a significant portion of residents commutes to the city for work/studying. Urban centres represent the densely populated and economically active areas within the FUAs and Greater Cities. By analysing these statistical outputs, policymakers, urban planners, and researchers can gain insights into the development and functioning of FUAs and Greater Cities. They can assess the impacts of various factors such as infrastructure investments, industrial developments, and environmental conditions on population distribution, economic activities, and worker/student mobility, enabling them to make informed decisions.
Target statistical population	The places that make up the M-FUA. Only people who are usually residents (with home location, according to what is possible to derive from MNO data) and usually gravitating to the selected places are of interest for this UC. Different characterisations of places and related activities may produce extensions of this UC.	
Spatial scope	The spatial scope is the LAU chosen by the NSI and its associated M-FUA. In the *reference scenario*, the indicator can be produced for all the Greater Cities in the country. In the *demonstrator scenario*, we can limit the computation of the output to only one Greater City. We will discuss the spatial scope of the demonstrator scenario jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	
Spatial resolution	The data processing will use the INSPIRE 100x100 m grid.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	The output will be any requested zoning system (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	In the *reference scenario*, the methods and tools shall be flexible to let the NSIs choose any temporal resolution. In the *demonstrator scenario*, we will discuss the temporal resolution jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	Whatever the output temporal resolution will be, the 'commuting' concept will require a longitudinal observation of the users for at least 12 months.
Possible additional classification of the target statistical population	FUA is already a kind of classification/tagging. No additional classification will be considered for this UC.	
Target NSI departments	• Demography and population • Census • Environment • Social policy	
Target users of the new statistical product	• Urban planners • Policymakers at local and national levels	

10 UC 5.A: DOMESTIC TOURISTIC ARRIVALS AND NIGHTS SPENT

The objective of this use case is to estimate the number of domestic touristic presences (i.e. overnight stays at a specific destination) and arrivals in a specific geographical area during a certain reference period. Domestic tourism refers to the activity of individuals or groups travelling within their country of residence for recreational, leisure or business purposes.[9] It is worth noting that, in the official definition of tourism, the purpose of the travel does not play a role and is only introduced later to characterise the travels. Therefore, from the perspective of this definition, the usual consideration that MNO data do not allow us to observe the purpose of the travel is not a limitation. This UC does not take into account touristic presences and arrivals generated by people coming from outside the country (inbound tourism) or by people leaving the country of residence (outbound tourism). Due to their peculiarities in terms of reference population and required analyses of MNO data, inbound and outbound tourism are the object of dedicated UCs, in Chapter 12 UC 5.C: Inbound tourism and Chapter 13 UC 5.D: Outbound tourism, respectively.

Statistical outputs of this UC are:

- the number of nights spent by tourists in a place at a given geographical level (regular square grid, municipality or other administrative territorial units), and
- the number of arrivals (counted as first night spent) of tourists in a place at a given geographical level (regular square grid, municipality or other administrative territorial units) per geographical area of tourists departure (regular square grid, municipality, or other administrative territorial units).

Data are processed at the single-night level and then longitudinally analysed to provide summaries at a monthly level. The output is provided for each month of the analysis period.

The definitions required for the implementation of the UC are chosen according to the definitions used in official tourism statistics as described in **TABLE 9**, showing an overview of the UC specifics and outputs.

It is worth noting that the statistical concept of tourism is closely related to that of UE, since in order to be considered a touristic trip, the travel must take the traveller outside his or her UE[10]. The UC on UE is provided in Chapter 4 UC 2.A: M-Usual Environment indicators and the description of the methods defined to derive the UE from MNO data according to the proposed pipeline is provided in Volume III of this deliverable.

Under the umbrella of domestic tourism statistics, we propose an additional UC, namely UC 5.B: Domestic same day visits dealing with same-day visits, which is described in the next section.

Table 9: UC 5.A – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	For this UC, the statistical output can be:	

[9] See: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Tourist
[10] Idem.

USE CASE DEFINITION AND OUTPUTS		REMARKS
	1. The number of touristic overnight stays, broken down by geographical area, month, and area of departure. 2. The number of touristic arrivals broken down by geographical area, month, and area of departure. For example, number of touristic overnight stays in August 2023 in the municipality of Rome spent by Italians; number of arrivals in the municipality of Rome in August 2023 from Milan province. The geographical area is identified by a regular square grid or by municipality or other administrative territorial units. For domestic tourism, only nights generated by resident people within the country are considered. *Definitions*: the definition of a geographical area as touristic depends on each individual. A place is defined as a touristic destination if it lies outside the geographical area corresponding to the individual's UE, as estimated in Chapter 4 UC 2.A: M-Usual Environment indicators. Hence, a touristic night is an overnight stay outside the geographical area corresponding to the UE. Touristic arrival is the first overnight stay in a given touristic destination.	
Target statistical population	The target statistical population is the resident population, and the focus of the analysis are the overnight stays outside their UE: • All overnight stays by people who are within the country, but outside their UE geographical area (i.e. in a domestic touristic place). • All nights of arrival (first overnight stay of a trip to a touristic place) spent by individuals who are within the country, but outside their UE geographical area (i.e. in a domestic touristic place).	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	
Spatial resolution	The spatial resolution of the data processing will be the INSPIRE 100m x100 m grid.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	At the output level, methods and tools will be flexible allowing the NSIs to choose any zoning system (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	The temporal scope will be chosen by the NSI, the default is set to the month. The methods and tools shall be flexible to let the NSIs choose any temporal resolution. In the *reference scenario*, the methods and tools provided would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	Whatever the output temporal resolution will be, the domestic tourism concept requires a proper longitudinal observation of the devices, to assess the individual UE.
Possible additional classification of the target statistical population	In this UC, overnight stays will not be further classified. The estimation of any possible classification can be seen as a potential extension/variation of this UC.	Possible future classifications useful for enriching the value of the produced indicators can be: • Number of nights in a month per weekend day or weekday. • Number of nights of arrivals in a month per weekend day or weekday.
Possible extensions and variations	Calculation of additional indicators, could include: • Duration of the touristic stay (average number of overnight stays per single trip); • Frequency of visits in a month/year (i.e. the % of people visiting the area at least X times per month/year); • Frequency of visits in a month/year to a place by area of departure (i.e. % of people visiting the place at least X times per month/year departing from a given area).	
Target NSI departments	• Statistics on tourism • Economic statistics • Social statistics • Environmental statistics	These estimates would provide insights into the overall touristic flow in a given month in terms of the number of overnight stays in a place and in terms of area of departure. They could be useful in official statistics in the following domains: 1. **Statistics on tourism:** output data can be used for measuring features of the touristic flow; for example, for: • understanding how touristic flow changes in an area depending on other variables or phenomena (climate change, infrastructure planning and building, social and marketing policies etc.).

USE CASE DEFINITION AND OUTPUTS		REMARKS
		• understanding if services and accommodations are adequate for the demand. 2. **Economic statistics**: output data can be used for the study of economic aspects and economic growth of places at several geographical levels. 3. **Social statistics**: the information can contribute to the assessment of quality-of-life aspects, for SDG and well-being statistics, and can be useful for assessing changes in social habits. 4. **Environmental statistics**: the information can be used for assessing the relation between touristic flow characteristics and environmental features.
Target users of the new statistical product	• Touristic services • Urban planners • Social policy • Safety and health managements • Environmental health and policy making	These estimates would provide insights into the overall touristic flow in each month in terms of number of nights spent in a place and in terms of area of departure, allowing for analyses that support decision-making in several domains, including: 1. **Touristic services**: output data can assist urban planners in understanding touristic patterns and needs, aiding in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to accommodate the needs of the population. 2. **Urban planners**: output data can assist urban planners in understanding touristic patterns and needs, aiding in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to accommodate the needs of the population. 3. **Social policy**: information can be used for assessing the impact of social policies and programs on tourism. 4. **Safety and health management**: output data can contribute to the assessment of safety and security needs in specific periods of time and places, according to touristic flows. They can help identify areas with higher demands of health care in different periods, emergency services

USE CASE DEFINITION AND OUTPUTS	REMARKS
	and infrastructure to ensure public safety. 5. **Environmental health and policymaking:** output data on touristic flows can support environmental specialists in assessing the impact of tourism on environmental factors (such as air quality, noise pollution, and waste management). It can help in identifying areas where interventions are needed to improve environmental health conditions and contributing to sustainable development solutions.

11 UC 5.B: DOMESTIC SAME DAY VISITS

The objective of this UC is to determine the number of residents visiting a place within the country, outside their UE, during the day (i.e. without an overnight stay).[11]

The detailed definitions required for the implementation of this UC are chosen according to the definitions in the official tourism statistics and described in **TABLE 10** which provides an overview of the UC specifics and outputs.

Table 10: UC 5.B – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	Number of same-day visits to a geographical area in a given period, by area of departure (e.g. number of same day-visits in August 2023 in the municipality of Rome from Naples province). *Definitions*: The definition of a geographical area as touristic depends on each individual. A geographical area is considered to be touristic if it lies outside the individual's UE (see UC 2.A: M-Usual Environment indicators). The number of same day visits is calculated as the number of daily domestic visitors to a touristic place in a given time per area of departure. For an individual, a place is defined as a touristic destination if it is outside the geographical area corresponding to their UE and if the individual departs from - and returns to - their usual residence (i.e. spend the night before and the night after the visit in their UE). In some European countries, the length of the same-day visit is also considered to operationalise the international definition in the household questionnaire used for the	

[11] See: https://www.unwto.org/glossary-tourism-terms#S

USE CASE DEFINITION AND OUTPUTS		REMARKS
	sample survey. Hence it is required for instance to count as same day visit those who last at least 3 hours. Even when dealing with MNO data, it could be helpful to introduce a length of the stay for the same-day visit to be able to identify the 'visits' and distinguish them from passing-by. We propose to consider as same day visit those lasting more than 3 hours. This will be implemented as a configurable parameter and its default value will be discussed with interested stakeholders (e.g. the TFMNO).	
Target statistical population	The target statistical population is the resident population of a country, as it could be estimated through MNO data, e.g. via the "home location" UC. Along with the 'home location' to estimate the residence through the MNO data, this UC also requires estimating the UE of the considered individuals. The latter is provided in the UC 'M-Usual Environment indicators'.	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	
Spatial resolution	The spatial resolution of the data processing will be the INSPIRE 100m x100m grid for the place of visit, as well as for the usual environment. At the output level, the methods and tools will be flexible, allowing the NSIs to choose any zoning system (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	The temporal scope will be chosen by the NSI, the default is set to the month.	Whatever the output temporal resolution will be, the same day visit concept requires a proper longitudinal

USE CASE DEFINITION AND OUTPUTS		REMARKS
	The methods and tools shall be flexible to let the NSIs choose any temporal resolution. In the *reference scenario*, the methods and tools provided here would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	observation of the devices, to assess the individual usual environment.
Possible additional classification of the target statistical population	In this UC, same day visits will not be further classified. The estimation of any possible classification can be seen as potential extensions/variations of this UC.	Possible future classifications, useful for enriching the value of the produced indicators, can be: • Number of same day visits in a month per weekend day or weekday. • Number of same day visits in a month per time slot within the day.
Possible extensions and variations	• Duration of same day visits (average number of hours) • Frequency of same day visits in a month/year per place (i.e. % of people visit the area at least X times per month/year) • Frequency of same day visits in a month/year in a visit place per area of departure (i.e. % of people visit the area at least X times per month/year departing from a given area) • Number of same day visits taken during a touristic trip (the visitor does not spend the night before and/or the night after the visit in his usual environment).	
Target NSI departments	• Statistics on tourism • Economic statistics • Social statistics • Environmental statistics	These estimates would provide insights into the overall flow of same day visits in a given month in terms of number of visits in a place and in terms of area of departure. They could be useful in official statistics in the following domains: 1. **Statistics on tourism:** output data can be used for measuring features of the flow of same day visits, for example: - understanding how the flow changes in an area depending on other variables or phenomena (climate change, infrastructure planning and building, social and marketing policies etc.);

USE CASE DEFINITION AND OUTPUTS		REMARKS
		- understanding if services and accommodations are adequate for the demand. 2. **Economic statistics**: output data can be used for the study of economic aspects and economic growth of places at several geographical levels. 3. **Social statistics:** the information can contribute to assessing quality-of-life aspects, for SDG and well-being statistics, and can be useful for assessing changes in social habits. 4. **Environmental statistics**: the information can be used to assess the relation between touristic flow characteristics and environmental features.
Target users of the new statistical product	• Touristic services • Urban planners • Social policy • Safety and health managements • Environmental health and policy making	These estimates would provide insights into the overall flow of same day visits in each period in terms of number of visits in a place and in terms of area of departure, allowing for analyses that support decision-making in several domains, including: 1. **Touristic services**: output data can assist urban planners in understanding touristic patterns and needs, helping in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to meet the needs of the population. 2. **Urban planners**: output data can assist urban planners in understanding excursion patterns and needs, helping in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to meet the population's needs. 3. **Social policy**: information can be used to assess the impact of social policies and programs on tourism. 4. **Safety and health managements**: output data can contribute to the assessment of safety and security needs in specific periods of time and places, according to visit flows. They can help to identify areas with higher health care demands in different periods, emergency services and infrastructure to ensure public safety. 5. **Environmental health and policymaking**: output data on the flow of excursionists can support environmental specialists in assessing the impact on environmental factors (such as air quality, noise pollution, and waste management). It can help to identify areas where interventions are needed to improve environmental health conditions and contribute to sustainable development solutions.

12 UC 5.C: INBOUND TOURISM

The objective of this use case is to estimate: the number of inbound touristic presences (i.e. overnight stays, at a specific destination within a country, by visitors who are not residents of that country) and arrivals in a specific geographical area during a certain reference period. Unless differently specified, we assume the reference period is one calendar month.

Statistical outputs of this UC are:

- the total number of nights spent by inbound tourists in a place at a given geographical level (regular square grid, municipality, or other administrative territorial units);
- the average number of nights spent by inbound tourists in a place at a given geographical level (regular square grid, municipality, or other administrative territorial units);
- the total number of arrivals (counted as first night spent) of inbound tourists in a place at a given geographical level (regular square grid, municipality or other administrative territorial units) per geographical area of tourists' departure (regular square grid, municipality, or other administrative territorial units);
- multi-destination trips: average number of destinations and average number of nights spent per destination by inbound tourists in the visited country. In multi-destination trips, the destination geographical level is the same geographical level as for the other outputs, say 1kmx1km grid or some administrative system (e.g. municipality, province, etc.).

To produce the statistical output, the data are processed on a daily basis and then longitudinally analysed to provide summaries at month level. The output is provided for each month of the period analysed. The statistical outputs listed above resemble official statistics production for the tourism statistical domain, nonetheless, being adapted to MNO as data sources. This adaptation will require the definition of algorithms and the adoption of conventions. Algorithms and conventions will be proposed in the methods section of this report and tested in the implementation stage. For instance, 'night spent' means the device spends a minimum number of hours (parameter configurable by the NSIs, e.g. 5-6 hours) in a certain place over a functional midnight (suggested 4 AM). When processing the data, if the night is spent travelling between different places, one can decide to consider this as one night spent within the country, but not at finer geographical disaggregation.

This and other conventional arrangements will be specified and discussed in the implementation stage of the use case. The definitions adopted for nights spent, arrivals and multi-destination trip resemble as much as possible those already in use in official statistics and guide us in the elaboration of the outputs. Again, for instance, trips that start in a given month and end in the following one will be counted according to the standard conventions already applied in official releases on tourism statistics. At European level, for instance, the adopted convention recommends accounting a trip in the reference period of its ending month. In addition to this convention, even the starting month is communicated, to properly evaluate the duration of the trips.

It is worth noting that the statistical concept of tourism is closely related to that of UE, since in order to be considered a touristic trip, the travel must take the traveler outside his or her UE.[12] When extracting information

[12] Idem.
Remarks on the proposed indicators:

on the inbound tourism from MNO data, a simple assumption is to consider inbound tourists those carrying on a foreign SIM. However, the concept of UE applies also to inbound tourism. A person with a foreign SIM present in the country of analysis for a period longer than a given threshold (usually three months) and/or on a regular basis, cannot be considered an inbound tourist. The continuity and the regularity of the observation of the foreign SIM can be derived using the algorithms already described for the UC 2.A: M-Usual Environment indicators. However, an additional complication arises with foreign SIM which deserve a specific discussion.

Foreign SIMs are served in the visited country through agreements of international roaming. During the international roaming a single SIM can be served by multiple MNOs and this phenomenon complicates considerably the analysis of the single device at each temporal scale (from daily to mid-term to long-term). To illustrate the problem, the figure below depicts a toy case.

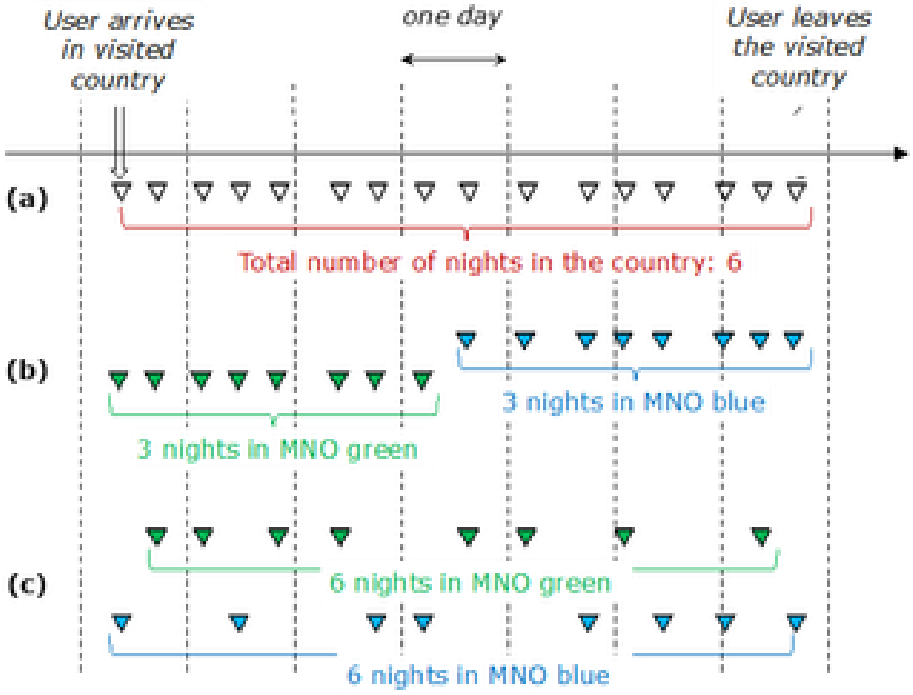

Figure 3: Potential different international multi-roaming behaviours of a single device

Complete view of a foreign SIM (inbound tourist) visiting a country for 6 days in red (a). Situations (b) and (c) represent two different international roaming across two visited MNOs, the blue and the green one

FIGURE 3 shows a device with a foreign SIM visiting the country of analysis for 6 days and generating MNO events represented by the triangles. Line (a) represents the complete view of them. In line (b) the same visitor is served by two visited MNOs, the green one in the first three days of his/her stay, the blue one in the last three days. Finally, in line (c) a different situation of international multi-MNO roaming is depicted, i.e. the device is served by the two

- It is worth noting that the aggregated number of visits are non-additive over space and time. Hence, in line with official statistics on tourism, in this use case we consider 'nights spent' rather than visitors, Nights spent can be aggregated over time and space, assuming only one night is assigned to any specific destination per each device.
- Another popular indicator in official tourism statistics is the number of 'same day' visit. According to the official definition a same-day visit starts and ends in the usual environment of the device. Other short visits like stop-over or short visit without a night spent during a long trip are not 'same-day visit' in the official definition. According to this definition, official 'same-day visit' are unusual for inbound and outbound, rather close to the usual environment and relevant for domestic tourism. For this reason, we do not consider same day visit related indicators in this use case, neither we consider stop-over or short visit without a night spent during a long trip to avoid ambiguity with official statistics indictors on 'same-day visit'.

visited MNOs in each single day of his/her stay. Clearly, **FIGURE 3** should be intended for illustrative purposes only and does not claim to represent all the possible multi-MNO international roaming behaviours.

Clearly, in both situations (b) and (c) the single MNO analysis is partial and will produce biased results in terms of the output statistical indicators described above. Similarly, the multi-MNO aggregations of single-MNO observations coming from situations (b) and (c) risk to produce seriously biased results.

In situations (b) and (c), as well as in other similar situations generated by multi-MNO international roaming, one would be able to reproduce and analyse the complete event generation only if inter-MNO linking at micro level is allowed, e.g. for each device it is possible to join all the events (or daily summaries thereof) generated by the visited MNOs. This may require the adoption of a secure computation environment based on advanced Privacy-Enhancing Technologies (as supplementary technical and organisational measures) to comply with GDPR. Otherwise, single-MNO observations as well as any aggregation of these might produce seriously biased results in terms of nights spent and arrivals. In principle, this bias could be reduced or even removed through sophisticated statistical methods and the analyses based on multi-MNO aggregations of single-MNO counts can be adjusted somewhat. However, at the time of writing this report, further investigations are needed to assess:

1. the severity of the phenomenon of multi-MNO international roaming, and how such severity varies across the visited countries and home operators;
2. the loss in accuracy when deriving the indicators from aggregated data based on partial views from single MNO processing, compared to the fusion of individual data.

An additional issue with inbound statistics based on MNO data is related to the interest of official statistics in breaking down the output indicator listed above by citizenship of the tourists. When dealing with MNO data, an approximation of the citizenship of the person using the device can be provided by the country of the home provider. However, especially extra-EU visitors can be expected to buy a new SIM in the first European country of their stay and then retain that SIM for the rest of the trip around Europe, since the European roaming agreements makes more convenient to use the European SIM than their original non-European one in all the different visited European countries.

Again, further investigations are required to assess the size of this behaviour and the way it affects the break down by citizenship of the previous indicators.

The definitions required for the implementation of the UC are described in the table below, showing an overview of the UC specifics and targeted statistical outputs.

Table 11: UC 5.C – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	For this UC, the statistical output can be: 1. The total number of nights spent by inbound tourists, broken down by geographical area, month, and country of departure (country of the home operator). 2. mean number of nights spent by inbound tourists, broken down by geographical area, month, and country of departure (country of the home operator). 3. the total number of inbound tourists arrivals broken down by geographical area, month, and country of departure (country of the home operator). 4. multi-destination trips: average number of destinations and average number of nights spent per destination by inbound tourists in the visited country, broken down by month and country of departure (country of the home operator). For example, number of inbound touristic overnight stays in August 2023 in the municipality of Rome spent by EU and non-EU visitors; number of arrivals in the municipality of Rome in August 2023 from the Netherlands. The geographical area is identified by the predefined standard Europe grid INSPIRE (from 1kmx1km to less finer resolutions) or by any zoning systems provided (e.g. municipality or other administrative territorial units). For inbound tourism, nights generated by foreign people (people carrying on a foreign SIM) excluding those with M-home location or M-usual environment within the country. Definitions: • The labelling of a geographical area as 'touristic' is related to the UE of the device and, therefore,	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	the same place may be labelled as 'touristic' destination for one device, but not for another. • A place is defined as a touristic destination if it lies outside the geographical area corresponding to the individual's UE, as estimated in the UC 2.A: M-Usual Environment indicators. • Hence, a touristic night is an overnight stay outside the UE's geographical area. • Touristic arrival is the first overnight stay in a given touristic destination. In the case of inbound tourism, the inbound touristic arrival is considered the first place where the foreign person spent the night in the country. When consecutive nights are spent in different places within the country, this generates a multi-destination trip.	
Target statistical population	The target statistical population are the foreign visitors of a given country, with a focus on the overnight stays outside their UE. It is worth noting that the population of foreign visitors is theoretically definable; nonetheless, in practice it is not easy to deal with such a reference population for the visited country perspective.	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	
Spatial resolution	The spatial resolution of the data processing will be the INSPIRE 100m x100m grid for the place of visit, as well as for the usual environment. At the output level, the methods and tools will be flexible, allowing the NSIs to choose any zoning system (e.g.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	square grid, administrative divisions, etc.).	
Temporal scope and resolution	The temporal scope will be chosen by the NSI, the default is set to the month. The methods and tools shall be flexible to let the NSIs choose any temporal resolution. In the *reference scenario*, the methods and tools provided here would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	Whatever the output temporal resolution will be, the inbound tourism concept requires a longitudinal observation of the devices, to assess the individual UE. This proper longitudinal observation can be jeopardised in some cases due to the international multi-MNO roaming, as explained in the text above. The detailed assessment of this phenomenon, as well as the consequential loss of accuracy in the results, fall outside of the scope of this project and are left as potential topic of study in future follow-up projects.
Possible additional classification of the target statistical population	In this UC, overnight stays will not be further classified. The estimation of any possible classification can be seen as a potential extension/variation of this UC.	Possible future classifications useful for enriching the value of the produced indicators can be: • Number of nights in a month per weekend day or weekday. • Number of arrivals in a month per weekend day or weekday.
Possible extensions and variations	Calculation of additional indicators could include: • Duration of the touristic stay (average number of overnight stays per single trip); • Frequency of visits in a month/year (i.e. the % of people visiting the area at least X times per month/year); • Frequency of visits in a month/year to a place by area of departure (i.e. % of people visiting the place at least X times per month/year departing from a given area).	
Target NSI departments	• Statistics on tourism • Economic statistics • Social statistics • Environmental statistics	These estimates would provide insights into the overall touristic flow in a given month in terms of the number of overnight stays in a place and in terms of area of departure. They could be useful in official statistics in the following domains: 1. **Statistics on tourism**: output data can be used for measuring features of the touristic flow; for example, for:

USE CASE DEFINITION AND OUTPUTS		REMARKS
		• understanding how touristic flow changes in an area depending on other variables or phenomena (climate change, infrastructure planning and building, social and marketing policies etc.). • understanding if services and accommodations are adequate for the demand. 2. **Economic statistics**: output data can be used for the study of economic aspects and economic growth of places at several geographical levels. 3. **Social statistics**: the information can contribute to the assessment of quality-of-life aspects, for SDG and well-being statistics, and can be useful for assessing changes in social habits. 4. **Environmental statistics**: the information can be used for assessing the relation between touristic flow characteristics and environmental features.
Target users of the new statistical product	• Touristic services • Urban planners • Social policy • Safety and health managements • Environmental health and policy making	These estimates would provide insights into the overall touristic flow in each month in terms of number of nights spent in a place and in terms of area of departure, allowing for analyses that support decision-making in several domains, including: 1. **Touristic services**: output data can assist urban planners in understanding touristic patterns and needs, helping in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to meet the needs of the population. 2. **Urban planners**: output data can assist urban planners in understanding touristic patterns and needs, helping in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to meet the population's needs. 3. **Social policy**: information can be used to assess the impact of social policies and programs on tourism. 4. **Safety and health managements**: output data can contribute to the assessment of safety and security needs in specific periods of time and places, according to touristic flows. They can help to identify areas with higher health care demands in different periods, emergency services and infrastructure to ensure public safety. 5. **Environmental health and policymaking**: output data on touristic flows can support environmental specialists in assessing the impact of tourism on environmental factors (such as air quality, noise pollution, and waste management). It can help to identify areas where interventions are needed to improve environmental health conditions and contribute to sustainable development solutions.

13 UC 5.D: OUTBOUND TOURISM

The objective of this use case is to estimate the number of outbound touristic presences (i.e. overnight stays in a country which is not the residence of the visitors) during a certain reference period.

The statistical output of this UC is:

- the total number of nights spent by tourists outside their country of residence, broken down by the number of visited countries in case of a multi-country trip.

To produce the statistical output, the data are processed daily at single-device level and then longitudinally analysed to provide summaries at a month level. The output is provided for each month of the period analysed.

Again, it is worth noting that the statistical concept of tourism is closely related to that of UE since, in order to be considered a touristic trip, the travel must take the traveler outside his or her UE.[13] When extracting information on the outbound tourism from MNO data, we should still apply the concept of UE and consider a person as outbound tourist in a foreign country only if that country is outside his/her UE: i.e. if the person is not present in that country for a period longer than a given threshold (usually three months) and/or on a regular basis. Indeed, in such cases the foreign country is to be taken as UE and the person should not be considered a tourist. The continuity and the regularity of the observation of the device in a foreign country can be derived using an approach similar to the one described for the UC 2.A: M-Usual Environment indicators.

Indeed, when the UE of a device extends to foreign countries, it should be noted that the MNO data will be able to identify this only at the country level, without additional detailed spatial information about the exact areas within the country. Hence, when a foreign country is labelled as UE for a device, we are not able to distinguish potentially different unusual destinations within the country. This lack of detailed spatial information in foreign countries can be relevant especially for nearby/border countries; in this case, it is likely that the nearby areas are included in the UE while more remote areas are not. To cope with this issue, we can identify separately and count the nights spent in outbound countries by devices that have their UE including the foreign countries.

The definitions required for the implementation of the UC are described in the table below, showing an overview of the UC specifics and statistical output.

Table 12: UC 5.D – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	For this UC, the statistical output can be:	

[13] Idem.
Additional note:

- The analysis of outbound tourism can help to investigate the international multi-MNO roaming, which is indeed not an issue for outbound tourism, though might seriously affect inbound tourism, as described in UC 5.C. In practice, the home MNO can retain the information on how many visited MNOs in the same foreign country a given device connects to, how long it is connected to each of them, and other characteristics of the multi-MNO international roaming useful to assess the size of the phenomenon and to elaborate some adjustment factors for the aggregates based on single-MNO counts.

USE CASE DEFINITION AND OUTPUTS		REMARKS
	1. The total number of nights spent by tourists outside their country of residence, broken down by visited country and region of origin. For example, number of outbound touristic overnight stays in August 2023 in the Netherlands by people resident in Italy, broken down by 'home location' in Italy, representing the region of origin of the devices. The geographical area is identified by a regular square grid or by municipality or other administrative territorial units. For outbound tourism, nights generated by resident people within their UE, if it contains the visited country, should not be considered. Definitions: • The definition of a geographical area as touristic depends on each individual. A place is defined as a touristic destination if it lies outside the geographical area corresponding to the individual's UE, as estimated in the UC 2.A: M-Usual Environment indicators. • Hence, a touristic night is an overnight stay outside the geographical area corresponding to the UE.	
Target statistical population	The target statistical population are the resident persons visiting a foreign country, which is outside their UE.	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	
Spatial resolution	For outbound tourism, the spatial resolution of the visited country is the country itself.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	For the breakdown of the origin area, the spatial resolution of the data processing will be the INSPIRE 100m x100 m grid. At the output level, methods and tools will be flexible, allowing the NSIs to choose any zoning system (e.g. square grid, administrative divisions, etc.) of the origin area.	
Temporal scope and resolution	The temporal scope will be chosen by the NSI, the default is set to the month. The methods and tools shall be flexible to let the NSIs choose any temporal resolution. In the *reference scenario*, the methods and tools provided would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with the MNO partners involved in the project and Eurostat in Task 5 (testing).	Whatever the output temporal resolution will be, the outbound tourism concept requires a proper longitudinal observation of the devices, to assess the individual UE.
Possible additional classification of the target statistical population	In this UC, overnight stays will not be further classified. The estimation of any possible classification can be seen as a potential extension/variation of this UC.	Possible future classifications, useful for enriching the value of the produced indicators, can be: • Number of nights in a month per weekend day or weekday.
Possible extensions and variations	Calculation of additional indicators, could include: • Duration of the touristic stay (average number of overnight stays per single trip); • Frequency of visits in a month/year (i.e. the % of people visiting the area at least X times per month/year); • Frequency of visits in a month/year to a place by area of departure (i.e. % of people visiting the place at least X times per month/year departing from a given area).	
Target NSI departments	• Statistics on tourism • Economic statistics • Social statistics	These estimates would provide insights into the overall touristic flow in a given month in terms of number of

USE CASE DEFINITION AND OUTPUTS		REMARKS
	• Environmental statistics	the number of overnight stays in a place and in terms of area of departure. They could be useful in official statistics in the following domains: 1. **Statistics on tourism:** output data can be used for measuring features of the touristic flow, for example: - understanding how touristic flow changes in an area depending on other variables or phenomena (climate change, infrastructure planning and building, social and marketing policies etc.); - understanding if services and accommodations are adequate for the demand. 2. **Economic statistics**: output data can be used for the study of economic aspects and economic growth of places at several geographical levels. 3. **Social statistics:** the information can contribute to assessing quality-of-life aspects, for SDG and well-being statistics, and can be useful for assessing changes in social habits. 4. **Environmental statistics**: the information can be used to assess the relation between touristic flow characteristics and environmental features.
Target users of the new statistical product	• Touristic services • Urban planners • Social policy • Safety and health managements • Environmental health and policy making	These estimates would provide insights into the overall touristic flow in each month in terms of number of nights spent in a place and in terms of area of departure, allowing for analyses that support decision-making in several domains, including: 1. **Touristic services**: output data can assist urban planners in understanding touristic patterns and needs, helping in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to meet the needs of the population. 2. **Urban planners**: output data can assist urban planners in understanding touristic patterns and needs, helping in the development of sustainable cities, efficient transportation systems, and appropriate infrastructure to meet the population's needs. 3. **Social policy**: information can be used to assess the impact of social policies and programs on tourism. 4. **Safety and health managements**: output data can contribute to the assessment of safety and security needs in specific periods of time and places, according to touristic flows. They can help to identify areas with higher health care demands in different periods, emergency services and infrastructure to ensure public safety.

USE CASE DEFINITION AND OUTPUTS	REMARKS
	5. **Environmental health and policymaking**: output data on touristic flows can support environmental specialists in assessing the impact of tourism on environmental factors (such as air quality, noise pollution, and waste management). It can help to identify areas where interventions are needed to improve environmental health conditions and contribute to sustainable development solutions.

14 UC 6: EXPOSURE TO RISKS

Satellites such as Copernicus Sentinel-5P, Sentinel-4 and Sentinel-5, along with instruments like Ozone Monitoring Instrument (OMI) and Global Ozone Monitoring Experiment -2 (GOME-2), collect data to monitor pollution in Europe. These satellites provide information on air quality, pollutant concentrations and pollution sources. This UC aims to estimate the number of people exposed to a given phenomenon or to a given risk. Currently, the main data source for pollution mapping is census data, which does not reflect the actual location of individuals. Combining the present population data from the use case in Chapter 3 UC 1.A: Present population estimation with satellite images makes it possible to obtain more timely and accurate estimates of the exposure of people to a specific environmental risk at a given time. On the other hand, if we are interested in people who are 'usually' exposed to a risk, rather than in a snapshot of people exposed at a precise reference time, we can combine satellite imagery data with the results provided by the UC in Chapter 5 UC 2.B: M-Home Location indicators. In this way, we can estimate the number of people who usually live in places exposed to some sources of pollutants. Similar indicators can be produced by replacing the results of the UC in Chapter 5 UC 2.B: M-Home Location indicators with the UC assessing the UE in Chapter 4 UC 2.A: M-Usual Environment indicators. This allows us to estimate the people exposed to pollution or some environmental risk, not only when they are at home, but in the set of places where they regularly spend their time (e.g. school, workplace, etc.).

In this project, we will concentrate on estimating the number of people who are *usually* exposed to some risks/pollutants, defining this 'usuality' by means of the UC in Chapter 4 UC 2.A: M-Usual Environment indicators. In practice, we will count as exposed to a risk those people for whom the places of their UE present a level of pollution that is higher than a given threshold, or who are exposed to an environmental risk (e.g. less than 1km from a volcano). Our indicators will be based on binary counts, meaning that a person will be considered as "exposed" regardless of the duration of the exposition or the actual exposure at a given time (e.g. at the time of the potential eruption of the volcano).

Further indicators can be produced by adopting a fractional approach, e.g. assigning a risk proportional to the duration of exposure to exposed people, characterised by some additional factors (e.g. exposed when at home, exposed when at work/school), and defining specific reference times (e.g. exposed at the 12 noon every Sunday). These additional outputs are natural and easy extensions of the basic outputs that we will explore during this project.

Table 13: UC 6 – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	The statistical output can be: 1. the size of population exposed to high levels of air pollution; 2. the size of population exposed to certain environmental risks (flooding, volcano eruption, etc.) This UC might require some focus for big cities or for target areas.	To calculate the size of the population exposed to a certain level of risk, we first binary associate the risk to the individual's UE and then we count the number of persons exposed to a certain level of risk by summing up.
Target statistical population	People, regardless their age and other characteristics; i.e. individuals of any age and demographic characteristics, encompassing the entire population present in an area for a specific time period.	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	
Spatial resolution	The data processing will use the INSPIRE 100x100 m grid. The output will be any requested zoning system (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	The methods and tools shall be flexible to let the NSIs choose any temporal scope and resolution. In the *reference scenario*, the methods and tools provided would support any temporal scope. In the *demonstrator scenario*, we will discuss the temporal scope jointly with Eurostat and the MNO partners involved in the project Task 5 (testing).	Whatever the output temporal resolution will be, the "exposure to risk" concept defined in this use case requires a proper longitudinal observation of the devices, to assess the individual usual environment.
Possible additional classification of the	Classification of places in terms of the risks they pose is crucial for this UC.	

USE CASE DEFINITION AND OUTPUTS		REMARKS
target statistical population	This classification will be derived from the integration of MNO data with additional external/publicly available sources.	
Target NSI departments	• Demography and population • Census • Environment • Social policy (SDG, equitable and sustainable well-being indicators) • Health	
Target users of the new statistical product	• Urban planners • Policymakers at local and national levels	

15 UC 7: INTERNAL MIGRATION

In official statistics, 'internal migration' is defined as 'the movement of a person within a state (internal migration) for more than one year, irrespective of the causes, voluntary or involuntary, and the means, regular or irregular, used to migrate.[14] The movement must cause a change in 'the place of usual residence' from an administrative area to another within the same country.[15]

In this UC, we mimic the 'internal migration' concept by considering stable changes in mobile devices' home location occurring in a time period of observation comparing two different six month consecutive periods (for the definition of 'home location', please refer to UC 2.B: M-Home Location indicators). As MNO data cannot accurately reproduce an official statistical concept, the aim is to use MNO data to investigate behaviours and habits that cannot be derived from other data sources (i.e. specifically for this UC 'current home location' vs 'official residence').

More precisely, for each device, this UC takes as **input** the 'home location' output of the M-UE indicators UC, where the home location label is assigned to the device (the default value of the time of reference in M-UE indicators UC is 6 months).

Taking 12 months as the observation period, we can compare the home locations assigned to the first and the second 6 months of the observation period and detect, for each device, if the home location changes. In this UC, we assume non-overlapping reference periods; however, the methods developed for this UC need to be adaptable to allow, in the future, also the implementation for month-to-month rolling reference periods.

Whereas the first setting reflects more closely the official statistics indicator, the second one with a series of overlapping 6-month intervals can be useful for additional analyses.

In conclusion, the aim of this UC is to produce estimated counts of people changing 'home location' in a given reference time period, broken down by municipality of origin and destination or other more or less granular geographical areas.

[14] https://home-affairs.ec.europa.eu/networks/european-migration-network-emn/emn-asylum-and-migration-glossary/glossary/migration_en

[15] In the following the word 'country' replaces 'state'.

Table 14: UC 7 – description and targeted statistical outputs

USE CASE DEFINITION AND OUTPUTS		REMARKS
Indicator(s) to be produced	Indicators are calculated as: number of individuals/devices whose home location (a proxy for the place of residence) changed from an administrative area to another, within the six-month period of reference. Indicators are calculated per pair of area of residence, in order to obtain indicators for all origin-destination pairs, relative to the considered period of reference. **Definitions**: the period of reference is 6 months by default. The area extension (spatial resolution of the output) is usually an administrative area. Default locations are the country's municipalities. **Example:** number of people having a home location in municipality A in the period July-December 2022 and a home location in municipality B in the period January-June 2023.	The method to combine the information extracted from MNO data with census data or other non-MNO data source, to produce results for the total population, is not the focus of this project. This project aims to define and describe a process pipeline that encompasses this integration step. However, where the state-of-the-art solutions do not provide satisfactory answers, the pipeline will be flexible and evolvable in order to implement more advanced solutions once they will be available. The integration of MNO data and non-MNO data sources is currently the main topic of the MNO-MINDS ESSnet project. It might be useful to consider intermediate results on mobile devices as final outputs, before the inference to the whole population, for two reasons: • counts of mobile devices allow NSIs to perform estimates on relative figures of flows, as auxiliary information for the modelling of official figures; • It is in line with the flexible approach for the pipeline's 'Estimation' module (detailed in Chapter 21 in Volume III), which may require, in the future, inserting additional, more sophisticated, methods. Since NSIs may be interested in results before the integration with other data sources, in the testing scenario we may allow the possibility to derive MNO based percentages/rates.
Target statistical population	The statistical unit is the individual/device resident in the spatial area under scope. The population under study includes all people having a 'home location' (defined by UC 2.B: M-Home Location indicators) in the considered geographical area in the first half of the observation period and a home location in a different geographical area in the second half the observation period (n.b. the focus is	

USE CASE DEFINITION AND OUTPUTS		REMARKS
	on the population flows from one area to another).	
Spatial scope	The spatial scope will be chosen by the NSI. In the *reference scenario*, the methods and tools provided would support the whole country. In the *demonstrator scenario*, we will discuss the spatial scope jointly with Eurostat and the MNO partners involved in the project in Task 5 (testing).	
Spatial resolution	The data processing will use the INSPIRE 100x100 m grid. The output will be any requested zoning system (e.g. square grid, administrative divisions, etc.).	
Temporal scope and resolution	The methods and tools shall be flexible to let the NSIs choose any temporal scope and resolution. In the *reference scenario*, the methods and tools provided would support the production of indicators for several years In the *demonstrator scenario*, the output computation will be limited to the selected specific 6-month period.	The 'home location' concept adopted in official statistics will require a longitudinal observation of the devices for at least 6 months. Accordingly, in the UE UC the 'home location' label is assigned for periods of 6 months as default period of reference.
Possible additional classification of the target statistical population	Assigning a home location in the country is already a type of classification/tagging. In this UC, we will not consider any further classification other than the home location.	
Target NSI departments	• Demography and population • Census • Social policy • Environment	Domains benefiting from a proxy of the resident population estimate are very diverse and extensive. However, in this study we will focus on its relevance to official population statistics. More specific points are discussed below: **Timeliness and frequency**: Official population statistics are often collected once every year (with full censuses only every few years) and with a considerable delay of the publications after the respective reference periods, while MNO data can provide real-time or near real-time information. By integrating MNO data, statistical authorities can obtain more frequent

USE CASE DEFINITION AND OUTPUTS		REMARKS
		updates on population movements and changes of residence within the country, allowing for more timely and responsive planning and decision-making. **Granularity and spatial accuracy**: MNO data can provide highly granular and geographically precise information about population distribution. They can capture population movements at the individual level offering insights into population densities in specific areas. This level of detail helps to identify population hotspots, areas of high transit, or areas with rapidly changing populations
Target users of the new statistical product	• Urban planners • Policymakers at local and national levels	