KPI Speed

Methodological guidelines

Version 1.4, 5 December, 2023

Co-funded by the European Union

Trend line

KPI Speed. Methodological guidelines

Please refer to this report as follows:
Laiou, A., Vadeby, A., Riguelle, F., Van Schagen, I., Cardoso, J., Tutka, P., Usami, DS., Wardenier, N. (2023). KPI Speed. Methdological Guidelines. Report produced as part of the Trendline project, supported by the European Union.

Project details:	
Project start date: **Duration:** **Project name:**	15/10/2022 36 months Trendline
Coordinator:	Wouter Van den Berghe & Agnieszka Stelling | SWOV Institute for Road Safety Research Henri Faasdreef 312 – 2492 JP The Hague, The Netherlands
	Co-funded by the European Union

Report details:	
Version: **Dissemination level:** **Date:**	1.4 Public 05/12/2023

Report Author(s):

Laiou, A. (NTUA), Greece
Vadeby, A. (VTI), Sweden
Riguelle, F. (AWSR), Belgium
Van Schagen, I. (SWOV Institute for Road Safety Research), the Netherlands
Cardoso, J. (LNEC), Portugal
Tutka, P. (WUT), Poland
Usami, DS. (CTL), Italy
Wardenier, N. (VIAS Institute), Belgium

Revision history

Date	Version	Description
15/03/2023	1.0	Version approved by the KEG Speed
05/04/2023	1.1	Version sent to project coordinator
10/05/2023	Final draft 1.1	Version sent to project coordinator
26/05/2023	Final draft 1.2	Version sent to project coordinator
04/10/2023	Final draft 1.3	Version sent to project coordinator
05/12/2023	1.4	Version published at the website

Legal Disclaimer

Trend line

Table of contents

About Trendline

Trendline brings together 29 European countries (25 EU Member States and 4 countries as observers) for data collection, data analysis, delivery of road safety KPIs and for using these within road safety policies. Trendline is co-funded by the European Union and builds on the experience gained in the Baseline project. KPIs – Key Performance Indicators – are indicators that provide information about factors that are associated with crash and injury risks. At the core of Trendline project are eight KPIs:

Indicator	Definition
Speed	Percentage of vehicles travelling within the speed limit
Safety belt	Percentage of vehicle occupants using the safety belt or child restraint system correctly
Protective equipment	Percentage of riders of powered two wheelers and bicycles wearing a protective helmet
Alcohol	Percentage of drivers driving within the legal limit for blood alcohol content (BAC)
Distraction	Percentage of drivers NOT using a handheld mobile device
Vehicle safety	Percentage of new passenger cars with a Euro NCAP safety rating equal or above a predefined threshold
Infrastructure	Percentage of distance driven over roads with a safety rating above an agreed threshold
Post-crash care	Time elapsed in minutes and seconds between the emergency call following a collision resulting in personal injury and the arrival at the scene of the collision of the emergency services

These 8 KPIs originate from the Commission Staff Working Document 'EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero" SWD (2019) 283 final.' In addition, some new experimental and complementary indicators will be tested within Trendline (provisional names):

- Driving under the influence of drugs
- Share of 30km/h road lane lengths in urban zones
- Red-light negations by road users
- Compliance with traffic rules at intersections
- Helmet wearing of PMD (Personal Mobility Devices) riders
- Self-reported risky behaviour
- Attitudes towards risky behaviour
- Use of lights by cyclists in the dark
- Enforcement of traffic regulations
- Alternative speeding indicators.

For each of the original eight KPIs and the experimental KPIs, a 'KPI Expert Group' (abbreviated as KEG) has been established. Their main role is to draft the common methodological guidelines, to give feedback on questions, and to review the report of the KPI which they are covering.

Website Trendline: https://www.trendlineproject.eu/

Terms and definitions

Vehicle type (European Commission, 2021):

Car or taxi: Motor vehicle with 3 or 4 wheels, mainly used to transport people, seating for no more than 8 occupants (excluding the driver). Motor vehicles with these characteristics used as taxis as well as motor caravans are also included.

Light goods vehicle: Goods vehicle under 3.5t maximum gross weight: Smaller motor vehicle used only for the transport of goods. (=also van for transport of equipment by workers such as electricians, plumbers...)

Heavy goods vehicle: includes road tractors and goods vehicle over 3.5t maximum gross weight. Road tractor: road motor vehicle designed, exclusively or primarily, to haul other road vehicles which are not power-driven (mainly semi-trailers). Goods vehicle over 3.5t mgw: larger motor vehicle used noly for the transport of goods.

Bus or coach: Bus: passenger-carrying vehicle, most commonly used for public transport, having more than 16 seats for passengers. Coach: passenger-carrying vehicle, having more than 16 seats for passengers. Most commonly used for interurban movements and touristic trips. To differentiate from other types of bus, a coach has a luggage hold separate from the passenger cabin.

Road type (European Commission, 2021):

Motorway (European Commission, 2019b):
A road, specially designed and built for motor traffic, which does not serve properties bordering on it and which meets the following criteria:
(a) it is provided, except at special points or temporarily, with separate carriageways for the two directions of traffic, separated from each other either by a dividing strip not intended for traffic or, exceptionally, by other means;
(b) it does not cross at level with any road, railway or tramway track, bicycle path or footpath;
(c) it is specifically designated as a motorway.

Expressway (European Commission, 2021):
Road specially built for motor traffic, which does not serve adjacent properties, and:
a) Is accessible only from interchanges or controlled junctions;
b) Is specially sign-posted as an express road and reserved for specific categories of road motor vehicles;
c) On which stopping and parking on the running carriageway are prohibited.
Entry and exit lanes are included irrespective of the location of the sign-posts.
Urban express roads are also included.

Rural road (European Commission, 2021):
Public road outside urban boundary signs, excluding motorways and expressways.

Urban road (or road inside urban areas) (European Commission, 2021):
Public road inside urban boundary signs.

Time of the week / day (European Commission, 2022):

Week – daytime

Monday to Friday 6.00 a.m. to 9.59 p.m.

Week – nighttime

Monday 10 p.m. to Tuesday 5.59 a.m., Tuesday 10 p.m. to Wednesday 5.59 a.m., Wednesday 10 p.m. to Thursday 5.59 a.m., Thursday 10 p.m. to Friday 5.59 a.m.

Weekend – daytime

Saturday to Sunday 6.00 a.m. to 9.59 p.m.

Weekend – nighttime

Friday 10 p.m. to Saturday 5.59 a.m., Saturday 10 p.m. to Sunday 5.59 a.m., Sunday 10 p.m. to Monday 5.59 a.m.

1. Introduction

1.1. Context

The Communication of the European Commission "Europe on the Move – Sustainable Mobility for Europe: safe, connected and clean" of the 13th May 2018 confirmed the EU's long-term goal of moving close to zero fatalities in road transport by 2050 and added that the same should be achieved for serious injuries. It also proposed new interim targets of reducing the number of road deaths by 50% between 2020 and 2030 as well as reducing the number of serious injuries by 50% in the same period. To measure progress, the most basic – and important – indicators are of course the result indicators on deaths and serious injuries.

In order to gain a much clearer understanding of the different issues that influence overall safety performance, the Commission has elaborated, in cooperation with Member State experts, a first set of key performance indicators (KPIs). The KPIs relate to main road safety challenges to be tackled, namely: (1) infrastructure safety, (2) vehicle safety, (3) safe road use including speed, alcohol, distraction and the use of protective equipment, and (4) emergency response. The aim of the KPIs is connected to EC target outcomes.

The Commission Implementing Decision C (2021)5763 final of 5.8.2021 concerning the adoption of the work programme for 2021-2023 and the financing decision for the imple-mentation of the CEF foresaw a technical assistance action for the collection of Key Performance Indicators for road safety in EU Member States. The action builds on a previous CEF support action in 2020-2022 which established the Baseline project to collect 8 road safety Key Performance Indicators (KPIs) in 18 EU Member States. On the 10th of August 2022, a call was published with reference "MOVE/C2/2022-54— Technical Assistance for the development and collection of Road safety Key Performance Indicators (KPI)". A consortium of 25 EU Member States proposed the "Trendline" project to continue and elaborate the work on key performance indicators.

1.2. Purpose and basis of this document

This document presents the methodoligical guidelines for the KPI Speed. It describes the minimum methodological requirements to qualify for this KPI, defined as:

Percentage of vehicles travelling within the speed limit

The main target audience for this document are the persons in the participating countries that will collect and/or analyse the data to deliver the KPIs.

The minimal requirements set by the EC for this KPI are described in the Commission Staff Working Document SWD (2019) 283 (see also Annex 1), further referred to as "SWD". Most of those minimal requirements are incorporated in this guideline document. The requirements are quantified and specified for each of the parameters. This document is based on a review of the methodological guidelines that were developed within the Baseline project (Teuchies, 2021) and expert consultations within the (Trendline Key Expert Group Distraction).

The minimum requirement is to estimate **the percentage of vehicles travelling within the speed limit**. Besides this indicator, it is needed to also measure the **speed below which 85% of drivers are driving (V85)**, and the **average speed** (including standard error and standard deviation).

The theoretical population for this KPI refers to the percentage of vehicles travelling within the speed limit over the national territory on sections of roads that allow free flowing traffic. Optionally, when data is available on the total of all trips over the national territory (per vehicle type and road type), the KPI can also be presented as the total number of kilometres driven within the speed limit. For instance, the percentage of vehicles driving within thelegal speed limit then reflects the percentage of kilometers driven within the legal speed limit on sections of roads that allow free flowing traffic.

In addition to the specification of the minimum requirements to deliver the main KPI and the disaggregated indicators, this document also includes recommendations for optional additional activities. Member States can decide whether to follow the minimum requirements only or to extend (part of) their methodology, depending on available means and their own research questions.

2. Requirements for representative speed measurements

2.1. Free flowing traffic

The minimum requirement for the KPI on speeding is to only look at **free flowing traffic**. This means traffic conditions in which drivers can freely choose the speed they drive and are not restricted by traffic jams, infrastructure (e.g. speed bumps) or road works. To guarantee the observation of free-flowing traffic strict inclusion criteria are used for the measurement locations (see section 2.3.1). Next to selecting observation locations that allow free flowing traffic, there should also be enough headway between the vehicles of interest included in the analysis and the vehicle driving in front of it. Practically this means only including vehicles that have a headway similar to the distance travelled in 5 seconds at the current legal speed limit. This would for instance be 42, 69, and 167 meters for 30 km/h, 50 km/h, and 120 km/h respectively (Riguelle, 2008). It was found that using this 5 second headway is enough to guarantee free-flowing traffic (Global Road Safety Partnership, 2008).

2.2. Adequate observation equipment

2.2.1. Choice of measurement method

The SWD does not specify a required measurement method. However, the chosen method should allow the observation of momentaneous speed in free flowing traffic situations. The focus in these guidelines will be solely on tools that measure the momentaneous (or instantaneous) speed, thus producing spot speed data. Hakkert and Gitelman (2007) describe several methods to collect speed data, which will be explained in more detail in sections 2.2.5 through 2.2.7 of these guidelines.

Speed can also be measured over certain lengths of the road (e.g. with ANPR - Automatic Number Plate Recognition - systems), but this method is outside the scope of these guidelines, since for the KPI on speeding the requirement is to consider free flow traffic only and looking at speeds over certain lengths of road does not allow the analysis of free flow traffic. Similarly, floating car data can only be considered as a data source when it can be guaranteed that free flowing traffic can be analyzed and that the sample of drivers from which the data are obtained is representative of the broader population (e.g. are users of the smartphone application used to obtain the data representative of the Member State population).

2.2.2. Requirements for equipment

As mentioned above, the scope of these guidelines is limited to devices that measure instantaneous speed, or spotspeed at a particular location. The SWD does not specify the required equipment to do this. Because the speed measurement will usually be carried out over a short period of time (e.g. a couple of weeks) and ideally at a large number of locations, it is recommended to use equipment which can be installed quickly and flexibly. That is why portable systems such as radars or cameras that can be mounted quickly are preferred. It is possible to use permanent or semi-permanent systems such as loops as well, and therefore these are included in this section for the sake of completeness.

In general, three types of devices for collecting spot speed data can be distinguished (Knodler et al., 2005): out-of- road devices, in-road devices, and hand-held devices (although hand-held devices are not recommended but they are included for the sake of completeness). In this section, a basic overview of each of these devices is provided. Note that only the most commonly used devices are described.

The purpose of the measurements being to obtain the best possible measurement of the KPI (percentage of vehicles within limit), it is best not to take error margin of the equipment (measurement error) into account in defining the percentage of within versus over the limit. Taking measurement error into account like is often done in the use of radar measurements by the police for enforcement purposes, would come down to defining the speed limit of e.g. a 100 km/h speed zone at 103. This would cause a systematic bias in the results, reducing the percentage of vehicles over the limit artificially.

A typical 3% error margin of a measuring radar goes both ways (both + and -). Thus, the estimate of the KPI on percentage of vehicles within the speed limit can best be based on the measurement results of the radar as they are, not taking into account the error margin. Either way, treatment of measurement error should be included in the meta-data and will be taken into account in describing international differences in KPI.

2.2.3. Minimum requirements

Regardless of the device used, the equipment used should at least meet the following minimum requirements:

- be able to measure, store and deliver the instantaneous speed of individual vehicles;
- be able to measure the number of vehicles (traffic count);
- be able to measure the length of vehicles (in meters to one decimal place). The main reason to include length is that this variable is the most common means to determine vehicle type. When other means are available to determine vehicle type they can be used instead of length;
- be able to record the pass-by time of each vehicle (accurate to the second);
- be able to work uninterruptedly and store data for at least seven days;
- be able to collect data on at least 250,000 vehicles (either by internally storing the data in the device or by sending the data in real-time to an external server);
- be reasonably unobtrusive (not look like speed cameras);
- have a solid, stable installation. Also, the equipment should be calibrated and checked after installation to ensure correct data collection;
- be protected against theft and vandalism (optionally a small disclaimer can be added to the device explaining that the data are not used for law enforcement).
- be able to record speed measurement accuracy (defined by mean error and/or mean square error).

2.2.4. Unobtrusiveness of the equipment

Although this is not required by the SWD, to ensure the measurement of free flow traffic it is highly recommendedthat the equipment is as unobtrusive as possible. When drivers notice their speed is being measured, it will influence the speed they are driving, rendering the data less meaningful or unreliable. For this reason, hand-held devices are not recommended as it is hard to use these devices inconspicuously. Moreover, using hand-held devices limits the time window for the measurement, as drivers can report the measurement and warn other road users via various websites or smartphone applications.

2.2.5. Out-of-road devices

Doppler-based microwave radars are probably the most recommended method for speed measurements. These radars send a constant wave (24.5 GHz) which rebounds off the surface of the vehicle. From the modified frequency a number of variables can be deduced, typically including vehicle type (based on length), pas-by time, instantaneousspeed, and vehicle count. They are placed along the roadway using existing poles such as traffic signs or street lights. Microwave radars are relatively unaffected by weather conditions and are thus often preferred to other types of radars. Another similar technique is frequency modulated carrier wave radar (FMCW): however, the cost of theseradars is higher than Doppler radar but the performance is similar.

An advantage of radar systems is that they are relatively non-intrusive and there is usually no need to interrupt traffic to install them, although sometimes a road lane might have to be closed for a short amount of time to installthe radar. It must be noted that to ensure the safety of the people installing the radar equipment, this should be done by trained and experienced personnel in accordance with the safety regulations and traffic laws of the Member State. Both radar systems can be flexibly used and installed on a wide variety of locations ranging from city centers to motorways provided that poles or lampposts are available to install the radar equipment. This also offers flexibility in selecting locations, allowing a high-quality random locations sample.

A disadvantage is that with these devices, one can only obtain a coarse classification of vehicle types based on theirdimensions. Additionally, it is very important that this type of equipment is installed by experienced and trained personnel since poor installation can prejudice the data quality enormously. Also, the devices should be properly calibrated and checked after installation using another type of device (e.g. a speedgun) to ensure the installation issuccessful. This requirement for experience with the installation should be listed in the proposal if the installation of the equipment is subcontracted.

LIDAR devices (light detection and ranging) work similarly to radars but they use a different wavelength and a different type of wave. LIDARs use a laser wave and gather the reflected wave to obtain information on the detected objects. Their main field of application is enforcement because of their very high accuracy. The cost is significantly higher compared to Doppler-based radars.

Active infrared devices use the same principle as microwave radars but with infrared wavelengths. Smaller wavelengths make them more accurate than microwaves, which is especially helpful in distinguishing between vehicle types. This system too is more expensive compared to Doppler-based radars and is subject to errors in bad weather conditions.

Cameras can also be used when they are placed at a certain height above the roadway to film the passing vehicles.For a correct image setting two points of reference are used with a known distance between them. The device measures the time in which a vehicle drives from the first to the second point. The speed of the vehicle is calculatedby means of this time and the known distance. The vehicle length can be deduced as well using this method.

2.2.6. In-road devices

Some roads contain embedded devices that are capable of detecting vehicle speeds. These devices, such as loop detectors, are widely used for traffic surveillance purposes. They generally include a set of wires embedded into the roadway in a rectangular formation. Via the wires an electromagnetic field is created which can detect any vehicle that passes over the loop. Using these loops, data on traffic volumes can be derived directly. The main advantage isthat they are already in place, so data can be collected relatively easily.

An alternative in-road technology is the use of axle detectors. These detectors may be of different types: pneumatic, piezo-electric or quartz-electric. A rough classification of the vehicle can be detected provided that the headways between vehicles are not too small. As these devices do not count vehicles directly but count axles, a correction factor must be applied in order to establish correct traffic information. These correction factors are based on knowledge of the typical traffic characteristics of different road types but they must be adjusted depending on thespecificity of the road where measurements are carried out.

2.2.7. Hand-held devices

Radar guns and laser guns are portable instruments that are manually operated. The main advantage is their flexibility since they do not require any installation. The use of radar and laser guns would only be recommended onless-trafficked roads, as it is hard for the observer to monitor vehicle speed on roads with high traffic volumes. An advantage is that they can be used to distinguish particular types of vehicles which are not automatically detected by other systems such as vans, motorcycles, buses, ….

A major disadvantage of using radar or laser guns is their obtrusiveness, thereby possibly influencing the behaviour of the drivers. Another issue is that the overall cost of surveys with radar/laser guns is relatively high due to labour costs of the operators.

2.3. Adequate observation equipment

2.3.1. Choice of locations

Ideally the locations that are selected to obtain speed data should be representative for the whole network of roads in a Member State. Road design characteristics and the surrounding environment influence speeds at which drivers operate their vehicles, so not every location is suitable for free-flow speed measurements. Roads should meet some specific road design criteria in order to be suitable for free flowing traffic speed measurements. These specific requirements are described below.

The SWD specifies the following minimum requirements for the observation locations:

- The selection of the locations should be as random as possible with the objective of ensuring a representative sample for the national road network. However, roads where there are known or perceived speed problems are best omitted as these are not representative of the larger road network. The methodology for random sampling is not specified and is for the Member States to decide, but the method used for location selection should be described in the meta-data (sampling will be discussed in more detail in section 2.3.2).
- Measurements should not take place near speed cameras, neither fixed nor mobile.
- A minimum traffic flow of at least 10 vehicles passing per hour is required.

In order to ensure reliability and comparability of speed data, the locations at which speed measurements are carried out must be chosen carefully. All places where vehicles are likely to stop, accelerate or brake should be avoided, since at these locations free flowing traffic cannot be guaranteed. Each location should meet the following criteria as closely as possible:

- straight and uniform section of road (ideally there are no curves nearby that might influence the speed at the point of the measurement)
- section of road where it is possible to drive at a higher speed than the speed limit
- section with a small gradient (<5% on at least 500 meters preceding)
- away from junctions (>500 meters)
- away from any traffic calming device such as speed bumps or narrowing traffic lanes (> 500 meters)
- away from road works (> 500 meters)
- away from pedestrian crossings (> 500 meters)
- away from any speed limit change or sign (> 500 meters)
- away from sections where speed is enforced (e.g. traffic enforcement cameras).

If a location does not meet all the criteria listed above, it is recommended to mention this in the meta-data. It has to be noted that in all likelihood it will be hard to meet all of the above mentioned criteria, especially in built-up areas. As such, the criteria can be relaxed for locations in built-up areas. Still, it is recommended to select locations that meet these criteria as closely as possible under the circumstances.

A location that does not meet all the criteria listed above is considered the start of a non-free flow section which ends at the location where all criteria listed above are met again.

2.3.2. Sampling of locations

The SWD does not specify a required sampling method. Member States can define their own sampling methodology. It is important that the locations are representative for the national road network and cover the entire geographical area of the country. Ideally, over time it would be helpful for member states to work together with the European Commission to come up with common bases for sampling. In the meantime, sampling should be based on well-established statistical techniques aimed at achieving a properly representative result.

Selection of locations should be as random as possible, with a minimum requirement of covering the whole geographical area of the country. There are different options for random location selections: e.g. simple random, stratified random (e.g. random sampling in different regions). The basic process consists of three steps:

1. First the required number of locations is determined for the entire country or per region.
2. Next, these locations are randomly selected on a map using the entire area under consideration (e.g. countryor region), taking a sufficient geographical spread into account. The specific requirements for each location (e.g. road type or speed limit) do not have to be taken into account at this point. This step is just to ensure areasonable geographical spread of the randomly selected locations.
3. Finally, the exact locations that will be used for the observations are manually chosen in the area surroundingthe locations randomly selected in step two. At this point, the final selection must be based on the location requirements (different road types), inclusion/exclusion criteria (see section 2.3.1) and practical considerations. This final selection can be done using Google Street View for instance to search for observation locations near the randomly selected locations from step one that meet all the necessary criteria. The selected locations can then also be visited in

real life for a final check if needed. Pragmatic considerationsrelated to the observation locations can be taken into account (e.g. safety of observers or people installing measurement equipment should be guaranteed). Care should be taken to ensure that the different road types are also sufficiently geographically spread.

A convenient way of selecting locations randomly (step 2) is to use a GIS system (e.g. cartographic software like ARCView/ARCGIS) as such software can automatically randomly select location points within pre-defined areas. If Member States have no GIS software, step 2 can also be done manually using a national geographic map, e.g. Googlemaps/Google earth.

The sampled locations should be representative of the entire national territory. When stratification is used, results should be weighted according to traffic volumes by region. It is allowed to re-use the same sampled location for different times of day or days of week. In case such a crossed design is used, this should be indicated in the meta- data. The method used and rationale for choosing the locations of the measurements should be described in the method section of the study.

Ideally, the sampling procedure should comprise a selection from a database consisting of a list of uniform road segments, including their geographic coordinates and their characteristics such as:

- Road type (e.g. motorway, rural road, urban road...)
- Speed limit
- AADT (Annual Average Daily Traffic)
- Number of lanes (not including additional lanes at intersections)
- Length

Additional useful information is:

- Type of median provisions (median divided, flush median, no median)
- Surrounding environment (inner city, outer suburbs, extended shopping area)
- Road design characteristics (slope, curvature, etc.)

The basic characteristics of the locations should be recorded at the start of each observation: GPS coordinates, address or other geographical information, target lane or path and direction which is to be observed, traffic flow (should be free: no traffic jams, no road works). A code for the sampled location should be included in the database (at least as a qualitative code referring to the location).

In several countries, traffic counters have been placed on major roads with the general purpose of monitoring traffic flows on major roads of the road network. Since these counters can also produce speed data, the speed measurements in several countries are based on these traffic counters. In such cases, speed measurements are not based on a random sampling technique and will not be representative of the road network. For countries that already have permanent counters installed, it may be not feasible to change the system completely. If counters are installed only on main roads, an option would be to randomly sample fewer sites but to sample all of them on "non- main" roads. In this case a specific weighting procedure would be needed when calculating the speed indicators in order to take into account the respective share of main and "non-main" roads.

2.3.3. Minimum sample size

In order to ensure representative results for the entire road network**, the minimum required number of locations is 10 locations for each of the three road types** (urban, rural, motorway; see also section 2.4.1 on road types). The total minimum required number of **observed vehicles is 2000.** However, **for**

the first stratification level, a minimum of 500 observations per stratum is required (for the speeding KPI that means 500 observations per road type). Another minimum requirement is that **the proportion of observations at each of the three road types should be a minimum of 20%** (except if a certain road type, like motorways, is non-existent in a Member State).
Defining a minimum required sample size is by definition arbitrary since it depends on the level of accuracy that is considered adequate. With typical overall prevalence percentages in the range of 5 percent, accuracy in the order of range of 1 percent can be considered acceptable.

Accuracy for specific subgroups will by definition be lower. If higher accuracy levels are expected for particular subgroups (e.g. according to region), it is strongly recommended to increase the total sample size.

Since separate samples are taken for each road type and only straight segments of roads that fulfill certain requirements are considered (see section 2.3.1), the variance between locations should be quite small. If large variances are observed on a particular location in the sample, it is recommended to check whether that location fulfils all the requirements to be a good measuring location. If the location does not meet enough requirements, it is recommended to replace that location. If the location does meet enough requirements, it is recommended to increase the number of observations at that location.

Assuming a simple random sampling, the 95% confidence intervals for n=2000 and n=500 are, depending on prevalence (% of drivers within the speed limit) levels are shown in Table 2.1:

Table 2.1 95% confidence intervals for n=2000 and n=500 are, depending on prevalence

Prevalence	Lower bound, n=2000	Upper bound, n=2000	Lower bound, n=500	Upper bound, n=500
50%	47,8%	52,2%	45,5%	54,5%
75%	73,0%	76,9%	71,0%	78,7%
90%	88,6%	91,3%	87,0%	92,5%

To summarize, the minimum required sample sizes to provide the speeding KPI are:

- **min. 10 locations per road type = min. 30 locations in total**
- **min. 500 observations per road type**
- **min. 2000 observations in total**
- **the proportion of observations at each of the three road types should be at a minimum 20%**

For more information on random sampling of locations and for determining the minimum sample size, please referto the SafetyNet general recommendations for SPI (safety performance indicators):
http://www.dacota-project.eu/Links/erso/safetynet/fixed/WP3/sn_wp3_d3p8_spi_manual.pdf.
For the argumentation behind the minimum sample size see Appendix 2.

2.4. Stratifications and subpopulations

For speed measurements, **the minimum requirements** determined by the SWD should take into account **road type**(at a minimum urban, rural and motorways), **type of vehicle** (only cars are required, other types are optional), **time of day** (day is required, nights are optional), **day of the week** (weekdays are required, weekends are optional), and**the weather** (weather conditions must be good during the observations). In the sections below these minimum requirements will be discussed in more detail.

2.4.1. Road types

The SWD requires that the indicator should at a minimum cover **motorways, rural non-motorway roads** (defined as roads outside built-up areas)**, and urban roads** (defined as roads inside built-up areas**)**.

In Trendline three road types are generally considered:

- Motorways
- Urban roads: Public roads inside urban boundary signs (assuming exclusion of motorways)
- Rural roads: Public roads outside urban boundary signs, excluding motorways

For these road types the following definitions should be adopted:

Motorway

(definition according to Directive 2019/1936/EC)

A road, specially designed and built for motor traffic, which does not serve properties bordering on it and which meets the following criteria:

(a) it is provided, except at special points or temporarily, with separate carriageways for the two directions of traffic, separated from each other either by a dividing strip not intended for traffic or, exceptionally, by other means;

(b) it does not cross at level with any road, railway or tramway track, bicycle path or footpath;

(c) it is specifically designated as a motorway.

Expressway

Road specially built for motor traffic, which does not serve adjacent properties, and:

a) Is accessible only from interchanges or controlled junctions;

b) Is specially sign-posted as an express road and reserved for specific categories of road motor vehicles;

c) On which stopping and parking on the running carriageway are prohibited.

Entry and exit lanes are included irrespective of the location of the sign-posts.

Urban express roads are also included.

Rural road

Public road outside urban boundary signs, excluding motorways and expressways.

Urban road (or road inside urban areas)

Public road inside urban boundary signs.

Concerning motorways and especially for the KPI Speed, all Member States that collect data on motorways and on expressways are requested to report results **separately for motorways and expressways**.

- Expressways are roads with the same speed limit as motorways but with less strict rules for the occurrence of red-light crossings and access from other entrances than the official expressway entrances.

If another definition was applied in the Baseline results for specific countries, these countries might be requested to recalculate the Baseline results according to the present definition in order to evaluate the impact of changed definitions on the key estimates.

Ideally the locations that are selected to obtain speed data should be representative of the whole network of roads in a Member State.

In reality, road characteristics will vary between these different road types and therefore speed indicators should be computed separately for these three different road types. For countries where there is more than one speed limit per road type (for instance, rural roads with speed limits of 70 km/h and 90 km/h), it is recommended to compute the indicator either for each speed limit separately or for the most prevalent speed limit (it is not meaningful to aggregate data from roads with different speed limits).

For any given speed limit, it is not a minimum requirement to observe speed at both single and dual lane roads (if both exist). In Belgium, for instance, for most speed limits (50 km/h, 70/km/h and 90 km/h) there are both single and dual lane roads. It is, however, highly recommended to observe single and dual lane roads separately.

In the case of multi-lane roads, it is not necessary to measure speed on multiple lanes in the same direction simultaneously, but an option to do that is to use a radar per lane installed above the road (e.g. from a bridge). In Belgium, for instance, this is how measurements were performed (for three lane highways in one direction, three radars were used, one for each lane).

A second option might be to use a radar installed above the road focused on only one lane and to measure the left lane in 50% of the locations and the right lane in the other 50% of the locations (to be extended to 3 or more lanes if need be). Using a camera above the road allows to minimize interference in the measurements. Measurements using laser guns can also be considered, however, they are not recommended as they are hand-held devices, hard to use inconspicuously.

Aggregatingdata from single and dual lane roads with the same speed limit is not meaningful and is therefore not recommended. Should a Member State decide to look only at single lane roads or only at dual lane roads, it is recommended to choose the most prevalent type, thereby being more representative of the whole road network.

When communicating about the speeding indicators, some details should be provided about the design of the roadsincluded in the sample (e.g. number of lanes, type of division between opposite lanes, speed limit, ...).

2.4.2. Vehicle types

According to the SWD the **minimum requirement for the KPI is to observe the speed of passenger vehicles (cars)**.According to EuroStat, a passenger car is a road motor vehicle, other than a moped or a motorcycle, intended for the carriage of passengers and designed to seat no more than nine persons (including the driver). The term passenger car also cover microcars (small cars which, depending on individual Member State legislation, may need no permit to be driven and/ or benefit from lower vehicle taxation), taxis and other hired passenger cars, provided that they have fewer than 10 seats in total. This category may also include vans designed and used primarily for transport of passengers, as well as ambulances and motor homes. Excluded are light goods road vehicles, as well as motor coaches and buses and mini-buses/mini-coaches (https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Passenger_car).

This definition of a passenger car is similar to the UNECE definition of M1 vehicles: Vehicles used for carriage of passengers, comprising not more than eight seats in addition to the driver's = 9 seats total.

Optionally, motorcycles, vans, small trucks (between 6.00 meters – 12.00 meters) and trucks/ heavy goods vehicles(> 12 meters) can also be measured. When more vehicle types are considered, using the UNECE vehicle classification scheme is recommended.

Results should clearly define vehicle types included in the observations and should be presented separately for different vehicle types. Small vans might be hard to distinguish from person cars, and therefore, a certain percentage of the sample might contain small vans as well. This is hard to avoid and is acceptable, since in any event small vansare not that different from person cars in size and driving characteristics.

The way to distinguish between vehicle types depends on the measuring technique. With radar/laser guns, a humanobserver is present, allowing a more accurate categorization of vehicles. (It is recommended that an observer receives training to ensure that the classification is as accurate as possible). Most widespread automatic speed monitoring techniques (loops, tubes, radar classifiers) require that the classification of vehicles is obtained by indirect measurements:

- Roadside radars determine the lengths of vehicles on the basis of the time they stay in the beam of radar.
- Pneumatic tubes give information on vehicle lengths, number of axles and sometimes axle loads (based onthe pressure on the strips).
- Inductive loops use algorithms based on the expected vehicle distribution, the computed speeds and the occupancy rate of the loops to classify the vehicles. The determination of vehicle types becomes coarse when the traffic flow is heavy, usually resulting in an overestimation of the proportion of long vehicles.

Fortunately, even the coarser classifications (by inductive loops or roadside radars) are satisfactory to distinguish light vehicles (such as passenger cars) from other vehicles, at least when the traffic flow is not too heavy.

A specific problem with heavy vehicles is that these often have different speed limits compared with cars or light duty vehicles. Furthermore, different types of vehicles are similar in length (buses, coaches, trucks) and may also have different speed limits. Devices that determine the vehicle type on the basis of vehicle length may thus, classify vehicles with different speed limits within the same category. Based on the national situation, computation of indicators for 'long vehicles' on the basis of this kind of equipment may thus, be less meaningful.

2.4.3. Time period (time of day, day of the week, month)

The SWD requires at a minimum Member States to carry out speed measurements **during daylight hours on weekdays**. Measurements at night and in the weekends are optional but highly recommended. Comparisons between day and night are especially recommended due to the difference in traffic conditions and in the composition of the population of drivers between the two periods. The results should be shown separately for day and night and weekdays and weekend days.

In order to harmonize definitions of week vs weekend and daytime vs night-time, the definitions adopted in the ERSO project (European Commission, 2022) should be used:

Week – daytime
Monday to Friday 6.00 a.m. to 9.59 p.m.

Week – nighttime
Monday 10 p.m. to Tuesday 5.59 a.m., Tuesday 10 p.m. to Wednesday 5.59 a.m., Wednesday 10 p.m. to Thursday 5.59 a.m., Thursday 10 p.m. to Friday 5.59 a.m.

Weekend – daytime
Saturday to Sunday 6.00 a.m. to 9.59 p.m.

Weekend – nighttime
Friday 10 p.m. to Saturday 5.59 a.m., Saturday 10 p.m. to Sunday 5.59 a.m., Sunday 10 p.m. to Monday 5.59 a.m.

Possible impact of applying this common definition on Baseline results is expected to be minimal, hence no recalculations of the Baseline results will be requested, even when in Baseline a different definition was adopted.

Ideally, measurements should be carried out in a month that is "neutral" as far as seasonal variation in traffic is concerned. This means avoiding both school and bank holiday periods (especially summer, as it has the longest holiday period) and the winter period (due to a risk of bad weather). It is thus, recommended to carry out the measurements during late spring or early autumn.

The number of periods of measurement and the length of time during which it is possible to measure might vary depending on the measuring technique that is used and on the available resources (e.g. handheld devices operatedby people versus roadside radars that can measure 24/7). The exact time periods covered by the measurements should be indicated in the meta-data.

2.4.4. Region

The SWD states that the indicator should be representative of the whole Member State territory. To obtain speed indicators at regional level, a stratified random sample of locations according to region (e.g. NUTS1 regions) can beconsidered. If there are exceptions (e.g. for islands), they should be precisely defined and communicated.

If Member States want to obtain meaningful speed indicators at regional level it is highly recommended to apply all the minimum requirements defined for the national level to the regional level. So, for instance, one should cover the three minimum required road types per region as well as the minimum required sample size (e.g. the 2000 observations and 10 locations per road type required at national level would then be recommended for the regionallevel).

When stratification by region is used, results should be weighted according to traffic volumes by region in order tocompute the KPI at national level (see also section 2.5.1 on Post stratification weights and statistical analysis).

2.4.5. Weather

Measurements should not be carried out in bad weather conditions (e.g. heavy rain, snow, ice, strong winds or fog). Member States are free to define the exclusion criteria and report them together with the data. The main reasons for wanting to avoid bad weather conditions such as heavy rain are that these conditions can affect both speed and radar measurements. It is recommended to consult the people installing the radar equipment on what amount of rain will have a negative impact on the data quality.

2.5. Data analysis

2.5.1. Post stratification weights and statistical analysis

Specifications on calculating weights and confidence intervals are provided in Annex 3 *Suggested approach for weighting sample data and calculation of statistics.*

2.5.2. Expected results, data delivery and metadata

The **minimum required speeding indicator is the percentage of vehicles driving within the speed limit** (at nationallevel).
In addition to this indicator, it is necessary to also report the following speed indicators:

- V85 (the speed below which 85% of drivers are driving, i.e. the 85th percentile of speed)
- average speed
- speed standard deviation and standard error

Results should also include the number of locations and the unweighted number of drivers the results are based on.

National speeding indicators should be reported separately according to the following minimum required parameters:

- Vehicle type (personal cars)
- Road type (motorways*, rural roads*, urban roads*)
- Time period (daytime on weekdays)

* It is recommended to also provide results separately for different speed limits. Aggregating data from roads withdifferent speed limits is not meaningful.

Optionally, data from non-free flow traffic can be analyzed and reported besides the required speeding indicators for free flow traffic.

Together with the above estimates, a report should be submitted that describes the methodology of the field workand the statistical techniques used to weight and analyze the results. Member States are free to determine the statistical analysis techniques and tools. In addition to this, all Member States are expected to provide **metadata** on the applied regulations and proceduresrelated to this KPI (e.g. legislation on speeding).

For the **data delivery** to the Trendline consortium (inclusion in the Trendline database), three possible levels of aggregation are possible (further instructions on dataset structure and variables will be provided later):

1) Minimum level requirement: point estimates for all the minimum required observation categories (speeding indicator for cars on 3 road types during daylight hours on weekdays).

2) Medium level: crossed-level matrix of all considered levels of disaggregation (crossed point estimates) + confidence intervals.
3) Ideal level: also, delivery of the raw cleaned data (not pure raw data). Data cleaning is the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete (only if the minimally required data is missing), irrelevant, duplicated, or improperly formatted. This data is usually not necessary or helpful when it comes to analyzing data because it may hinder the process or provide inaccurate results.

In any case, there should be no tolerance beyond the error margin of the measuring device, i.e. the values recorded should be those measured by the instrument.

Trendline

3. Summary of minimum requirements

	Minimum requirement	Optional
KPI	• Percentage of drivers within speed limit • V85 • Average speed (+ Standard Deviation and Standard Error/Confidence Interval) • Free-flow traffic	Non-free flow traffic data
Location	• Random selection • Representative of entire national roadnetwork • Covering the whole geographical area of the country • Measurements should not take place near speed cameras, either fixed or mobile • A minimum traffic flow of at least 10 vehiclespassing per hour is required	• Stratification by Regions
Road type	• Motorways • Rural roads (defined as roads outside built-up areas, but no motorways) • Urban roads (defined as roads inside built-up areas)	• Differentiate between single and dual lane roads for rural and urban roads • Differentiate between speed limits within rural and urban roads
Vehicle type	• Passenger cars	• Motorcycles • Vans and light trucks • Heavy trucks • Buses
Time period	• Weekdays • Daylight hours • Spring/autumn	• Weekend • Night-time hours
Weather	• Good conditions	
Sample size	• Min 2000 observations • Min 500 observations / road type • Min 10 locations / road type • The proportion of observations at each of the three road types should be at a minimum 20%	

References

European Commission (2019a). Commission staff working document EU road Safety Policy Framework 2021-2030 -Next steps towards "Vision Zero". SWD (2019) 283 final. Retrieved from https://ec.europa.eu/transport/sites/transport/files/legislation/swd20190283-roadsafety-vision-zero.pdf

European Commission (2019b). Directive (EU) 2019/1936 of the European Parliament and of the Council of 23 October 2019 amending Directive 2008/96/EC on road infrastructure safety management. 26.11.2019 EN Official Journal of the European Union L 305/1.

European Commission (2022) Facts and Figures Urban areas. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. https://road-safety.transport.ec.europa.eu/system/files/2022-07/ff_roads_inside_urban_areas_20220707.pdf

European Commission (2022) Facts and Figures Roads outside urban areas. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. https://road-safety.transport.ec.europa.eu/system/files/2022-07/ff_roads_outside_urban_areas_20220707.pdf

European Commission (2021) CARE DATABASE. CaDaS Common Accident Data Set. 04/06/2021 Version 3.8. Care Team. Directorate-General for Mobility and Transport. https://road-safety.transport.ec.europa.eu/system/files/2021-07/cadas_glossary_v_3_8.pdf

Global Road Safety Partnership (2008). Speed management: a road safety manual for decision-makers andpractitioners. Geneva, Switzerland: Global Road Safety Partnership

Hakkert, A.S and V. Gitelman (Eds.) (2007) Road Safety Performance Indicators: Manual. Deliverable D3.8 of theEU FP6 project SafetyNet. Retrieved from: http://www.dacota-project.eu/Links/erso/safetynet/fixed/WP3/sn_wp3_d3p8_spi_manual.pdf

Knodler, M.A., Rothenberg, H. and Benavente, M. (2005). Spot speed study pilot training. Manual and presentation. Report 04E-G019-001. Amherst, Massachusetts Traffic Safety Research Program (MassSAFE),University of Massachusetts.

Riguelle, F. (ed.) (2008). Safety Performance Indicators for Speed: Pilots in Belgium and Spain. Deliverable D3.10bof the EU FP6 project SafetyNet.

Trendline

Annex 1 SWD KPI Key Performance Indicator for speeding

Ref: COMMISSION STAFF WORKING DOCUMENT - EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero, SWD (2019)238, is https://transport.ec.europa.eu/system/files/2021-10/SWD2190283.pdf

Rationale: Speeding is very regularly cited as one of the most common collision causation factors and is related toboth collision occurrence and severity.

Definition: **Percentage of vehicles travelling within the speed limit**.

Methodological aspects

Aspect	Minimum methodological requirements
Road type coverage	The indicator should cover motorways, rural non-motorway roads, and urban roads. The results should be presented separately for the three different road types.
Vehicle type	The indicator should include at least passenger vehicles (cars). Buses and goods vehicles (light [less than 3.5t] and heavy [more than 3.5t]) and powered two-wheelers are optional in a first phase (results should be presented separately for each vehicle type if possible).
Location	Member States to decide on the locations of the measurements, but measurements should not take place near safety cameras whether fixed or mobile. The choice of locations should be based on random sampling if this is possible, and in any case made with the objective of ensuring a representative sample.
Time of day	All Member States should elaborate the indicator for day hours in free-flow traffic: the night indicator should be optional due to its higher cost. The results should be shown separately for day and night.
Day of the week	Measurements to be carried out on Tuesdays, Wednesdays or Thursdays. Weekend measurements also possible but optional, and again should be shown separately if carried out.
Month	Measurements to be carried out preferably in late spring and/or early autumn.
Weather	Measurements should not be taken in bad weather conditions (e.g. heavy rain, snow,ice, strong winds or fog). Member States will define the exclusion criteria and report them together with the data.
Tolerance	No tolerance (beyond the error margin of the measuring device), i.e. the values recorded should be those measured by the instrument.

Annex 2 Rationale behind the minimum sample requirements

The methodological guidelines for all KPIs are designed to ensure international comparability between KPI values while taking into account feasibility and affordability. To that end the methodological guidelines have been definedin such a way that accurate and representative results can be obtained for all parameters of interest at a reasonable cost.

Obviously, the larger the sample of observations and locations for observation, the more accurate the KPI estimatesfor the different strata will be (e.g. a KPI value for a particular type of road, or a particular part of the week). Increasing the number of observations and locations however implies increasing field work costs. Statistically, the required minimum sample size depends mainly on the desired accuracy of the final estimates, for which no absolute value can be determined *a priori*. Therefore, for the main KPI estimates a pragmatic evaluation was made of the expected confidence intervals at different sample sizes and population parameters. Giving priority to feasibility andaffordability, as a rule of thumb the minimum total number of observations was set at 2,000, the minimum numberof observations for different strata at 500. It was agreed that this should allow to identify statistically meaningful differences between countries at an affordable price. For some countries, this will imply disproportionate samplingof certain strata compared to the distribution of traffic volumes over different strata. This is however required to allow statistically meaningful international comparisons at the level of each of the strata at interest.

The same pragmatic logic was followed for determining the minimum number of 10 locations for observation for each of the required road types of interest. Once again, there is no statistical rationale for determining the required minimum number of locations to ensure representativeness of the observations for the entire country. This mainlydepends on the amount of variance between locations and within a country. Giving priority to affordability, a rule of thumb was also used to define the minimum number of locations at 10 per stratum. In order to ensurerepresentativeness for the entire country larger numbers of locations might be required for larger countries. Taking field work costs into account, it was however decided to only identify the minimum requirements and leave decisions on the final number of locations to the discretion of the member states. Equally importantly, in order to ensure representativeness of the measurement locations these should be randomly selected as far as possible.

The main objective in defining the minimum methodological requirements is to keep a balance between affordability of the field work and the requirements to make meaningful international and historical comparisons. Therefore, the emphasis is placed on the minimum requirements that can also be taken into account by smaller countries. It is however of interest to any member state to increase the accuracy of the KPI estimates by boosting the number of locations and the number of observations.

Annex 3 Suggested approach for weighting sample data and calculation of statistics

A. Introduction

Within Trendline, several of the "KPIs" (Key Performance Indicators) refer to the relative number of vehicles or road users that respect certain legal limits and rules. These are sometimes called the "behavioural" KPIs. They refer to speeding, driving under the influence of alcohol, use of protective equipment, wearing a seatbelt or distraction.

In general, it is impossible to measure the performance of all vehicles at all times. Therefore, the KPI values are actually estimates based on a sample of vehicles and/or road users observed or surveyed. The main aim of these estimates is to estimate the percentage of kilometres driven on the entire road network (over a period of time, which one could be set to one year for instance) by vehicles respecting the legal limits and rules.

In term of sampling this means that the statistical population to be considered is the total traffic volume (typically expressed in kilometres driven) of moving vehicles over a certain area (i.e. country or region) over a certain period of time (e.g. one year). Estimates are made by sampling individual vehicles (or road user) at particular locations and moments in time. Hence the question arises as to how each of these individual observations have to be weighed in order for the overall average or percentage to reflect the overall percentage of vehicles complying with the rules in the total population.

For many KPIs within the Trendline project, data is being collected during observations (e.g., for distraction by mobile phone) or surveys (e.g., for driving under the influence of alcohol) at different locations. For all behavioural KPIs sampling on three different road types is required (motorways, rural roads, urban roads). For some KPIs sampling of different time periods and/or vehicle types is also required (for other KPIs only one type is considered).

Sampling is done in 2 steps:

1) Random selection of locations. Most beneficiaries use a disproportionately stratified random sample of locations, e.g., a same amount of locations per considered road type.
2) Random selection of vehicles/road users (nested) in each session.

The minimum number of locations for observations or surveys in Trendline is 10 per road type. At a given location, there may be several observation sessions. If different time periods are required in the sampling, then time periods should be linked to locations in a balanced way and also a minimum of 10 locations per time period is required as well as minimum of 2 locations for each combination of road type and time period. These constitute the sessions.

The data collected during these sessions allows to calculate a KPI value for that session, and, if sufficient data are available, also for subcategories (e.g., male/female; position in the car, type of vehicle). Moreover, for every session at least the road type is coded:

- Motorways
- Urban roads
- Rural roads

These are the generally required minimum sampling strata for the behavioural KPIs.

For most behavioural KPIs also a time period is coded for the observation session, specifically:

- Weekday
- Weekend day

For drink driving, four time periods are considered (Weekday daytime, Weekday nighttime, Weekend daytime, Weekend nighttime). For some KPIs (e.g., distraction) only one time period is considered (weekday daytime).

Each combination of road time and time period should be considered as a separate stratum: a combination of 3 road types and 2 time periods would lead to 3x2 or 6 strata.

Calculating KPIs for crossed strata of road type x time period is generally not minimum required but recommended, in particular if these categories have been a part of a sampling strategy. For such strata to include sufficient and sufficiently reliable data, a minimum requirement is that for each stratum (combination of road type and time period) minimum 2 different locations are used (but more are recommended).

There is a need to **weight** the results at the observation locations **within** the stratum (to arrive at the best estimate for the KPI value within the stratum) but also **across** the strata (to obtain, for example, a value for all considered time periods or for all roads together).

For certain KPIs other breakdowns are also possible (or even required), such as region, vehicle type/road user or sex. In such cases the number of strata that can be considered will be higher. However, in general strata with less than 500 data points should not be considered for calculating KPIs (unless specified differently in the minimum requirements of the methodological guidelines for the KPI), because the number of different observations and/or observation locations is too small and/or confidence intervals will be too wide. When strata with less than 500 observations are obtained and delivered to the Trendline coordination team, they will be treated differently in the tables and graphs of Trendline reports (e.g., shown in another colour or marked with an asterisk). However, such strata could be combined with or added to other strata to achieve this minimum. For instance, "weekday daytime" and "weekday nighttime" could be combined to "weekday".

B. First step: processing the data of each stratum individually

For each stratum (in the example above each of the 6) the following steps should be followed. Suppose you have *K* survey sessions in that stratum. For instance, you may have 6 observation sessions for observations on urban roads during weekdays. In that case, $K = 6$ for that stratum.

For each survey session *k* (with *k* varying between 1 and *K*) the traffic count(s) need to be determined. The traffic count obtained may concern all vehicles (or vehicles of a certain type) that passed by during the entire observation session, or for a fraction of the period (e.g. for 10 minutes in the middle of the session or for 5 minutes before and 5 minutes after the session). The duration of the counting is important. Please register both the actual count of the number of relevant vehicles and the time used to count. In case you have grounds to believe that the traffic density during the observation/survey session is quite different from the density during the counting session (e.g. because there was a sudden traffic jam causing much less vehicles to pass by during the observation, or because there was a bridge opening during the counting session), it is also useful to make an estimate of the number of relevant vehicles that passed by during the survey session. This estimate is somewhat redundant but would allow for unique unexpected situations.

Often it is planned that all observation or survey sessions have the same length of time (e.g., 60 minutes). This can be considered as the "standard duration" of a session. However, in practice, the duration of a session may deviate from the standard value, and this variation has to be accounted for when weighting the results.

So, for the session *k* in the stratum the following data is recorded:

Duration of the period used to count passing vehicles	$t_p(k)$
Number of passing (relevant) vehicles counted during the counting period	$N_p(k)$
Duration of the observation session	$T(k)$
Relative duration of the observation session = $\frac{T(k)}{Standard\ duration}$	$d(k)$
Estimated total number of (relevant) passing vehicles during the observation session, usually[1] this is equal to $N_p(k)\ x\ T(k)\ /\ t_p(k)$	$N(k)$
Number of (relevant) vehicles/individuals surveyed during the observation session	$n(k)$

It is important to have a good estimate of the total number of vehicles that passed this survey location during a session (this is $N(k)$). Otherwise, we do not know what share the individual survey sessions have within the stratum.

It is considered acceptable to assume that what is observed amongst the surveyed vehicles – $n(k)$ – is representative for all passing vehicles. Therefore, each surveyed vehicle represents $N(k)/n(k)$ vehicles in a session[2]. If the observation session took (a little) longer or shorter than the standard duration of the observation session (often the standard duration is 1 hour or 60 minutes), we can correct for that too (this is d(k)), yielding an observation weight for this vehicle type in this session in this stratum of :

$$Weight\ of\ observations\ in\ session\ k\ =\ W(k) = \frac{N(k)}{n(k) \times d(k)} \qquad (1)$$

When these weights are applied to all individual survey observations, the weights should add up to the number of vehicles that passed on all sessions in the stratum, had they been identical in duration.

[1] In exceptional cases where the traffic during the counting session is not representative for the traffic during the observation session, use the best estimate $N_h(k)$ (i.e. estimate of the total number of (relevant) passing vehicles 'per hour' during the observation session).

[2] If an observed vehicle represents 4 vehicles in the session, we have just one observation, not four, but it 'weights' for four vehicles

C. Calculate the KPI value per stratum

Now it is possible to create a database table or a spreadsheet with columns: this weight $W(k)$ and the actual observed values (surveyed vehicles – if required also vehicle type) and results noted as V(k), possibly augmented with administrative information (where, when, etc.) and further breakdowns (e.g., gender, position, ...) but keeping an eye on privacy of sensitive data. For instance, the observations of using a seatbelt in a survey could be ordered in the way as indicated in *Table 1* below (the other variables would concern the position of the person, whether he/she is driver or not, sex, ...).

Table 1. Data to be collected per observation

Date	Time	Location	Road type	Vehicle type	Time period	Within Stratum Weight W(k)	Seatbelt	Other variables
1-May-23	12:15	Site 51	Rural road	Passenger car	Weekend day	4	1	...
1-May-23	12:16	Site 51	Rural road	Passenger car	Weekend day	4	0	...
1-May-23	12:16	Site 51	Rural road	Truck	Weekend day	3	1	...
...								
2-May-23	12:15	Site 52	Urban road	Truck	Weekday	5	1	...
2-May-23	12:16	Site 52	Urban road	Passenger car	Weekday	3	1	...
2-May-23	12:16	Site 52	Urban road	Passenger car	Weekday	3	0	...
...								

Per session the KPI value V(k) can then be calculated as the average value of all observations. If a "positive" observation is given a score of 1 and a negative observation a score of 0, the average value is then a value between 0 and 1, which can be expressed as a percentage. We can then obtain a table with summary data on all the sessions. *Table 2* gives such information for the example of a stratum of passenger cars observed on weekdays on rural roads.

Table 2. Example of summary data of all sessions within a stratum

Session	Road type	Vehicle type	Time period	Observed vehicles n(k)	Within Stratum Weight W(k)	Seatbelt use V(k)
1	Rural	Passenger car	Week day	120	4.4	88.6%
2	Rural	Passenger car	Week day	110	3.8	92.7%
3	Rural	Passenger car	Week day	95	6.1	94.3%
4	Rural	Passenger car	Week day	130	2.6	78.6%
5	Rural	Passenger car	Week day	118	3.7	84.5%
6	Rural	Passenger car	Week day	84	4.1	94.3%
7	Rural	Passenger car	Week day	156	3.3	92.1%
8	Rural	Passenger car	Week day	124	4.0	86.2%
9	Rural	Passenger car	Week day	130	2.8	87.4%
10	Rural	Passenger car	Week day	145	2.7	88.1%

The formula for the KPI value of that stratum with K sessions is then:

$$KPI\ Value\ of\ the\ Stratum = \sum_{k=1}^{K} \frac{n(k)*W(k)*V(k)}{\sum_{1}^{k} n(k)*W(k)} \qquad (2)$$

For the example, the KPI value of the stratum would be 89%. For each different stratum, in general a different KPI value will be obtained.

D. The case of several vehicle types, road users or further breakdowns within the stratum

For some KPIs it is desirable or even required to make a distinction between several vehicle types and/or road users. This implies that each of these subgroups should be considered as a separate stratum; the logic discussed above should be applied to each considered vehicle or road user type.

However, this supposes that you can also count these different types during the traffic count in each session. If that is not possible, then you should assume that the distribution of vehicles passing by is the same as that of the vehicles observed/surveyed. This assumption is justified as the general rule during the fieldwork is to observe (or survey) the first arriving vehicle after coding the former one (random sampling - no deliberate over- or under-sampling of a specific vehicle/road user type).

This means that you have to adapt $N(k)$ above accordingly and use a value of *n(k)* per considered vehicle/road user type.

Other variables like age category and sex are generally no specified sampling strata in behavioural measurements on the road but collected variables of the surveyed road users[3]. If you, for instance, also want to make a distinction between male and female drivers, then the same assumption applies that the relative number of females in the set of the observed vehicles is the same in the set of the vehicles passing by.

E. Aggregation of the KPI results of different strata

From a policy perspective it can be useful to aggregate the data, for instance to arrive at a national indicator taking into account all road types, time periods and vehicle types. This is also desirable and often required within Trendline.

If two (or more) strata need to be aggregated, the relative importance of each stratum within the aggregation (sum) needs to be assessed. Within Trendline, the relative importance is based on the (estimated) volume of traffic in each of the strata. If the first stratum represents (or is representative for) 50% of traffic volume, the second represents 30% and the third 20%, the aggregated value is:

$$\text{Aggregated KPI value} = 0.5 \times KPI \text{ value stratum } 1 + 0.3 \times KPI \text{ value stratum } 2 + 0.2 \times KPI \text{ value stratum } 3.$$

[3] In questionnaire surveys age and sex are sampling strata - so there it makes sense to weight according to population statistics. But this is not the case in roadside surveys.

Thus, more general,

- if there are *M* strata to be aggregated
- let *TR(i)* represent the relative traffic volume of stratum *i* (*i* ranging from 1 to *M*)
- let *KPI(i)* be the KPI value of stratum *i*

Then:

$$Aggregated\ KPI\ Value = \sum_{i=1}^{M} TR(i) * KPI(i) \quad (3)$$

If crossed strata are considered, traffic information can come from different sources (e.g., national counts on roads for the proportions on the road types, and online representative mobility survey data for the relative proportions according to time period) which should be combined in a logical way to calculate a traffic volume % for each stratum (all summing up to 100%).

There are two possible ways to account for the relative importance of traffic volume and hence to determine or estimate *TR(i)*:

(1) National data on traffic volume (vehicle kilometres driven) by type of vehicle and type of road and time period. In the ideal situation national traffic volume data is available for all considered crossed strata but possibly this information has to come from combing different sources. It is also possible that no data is available for specific strata (e.g., no indication of national traffic volume according to the considered time periods).
Information on traffic volume can come from different sources such as national counts on roads for proportions on the road types. Representative online mobility survey data may be available for the relative proportions according to time period. If traffic volume data are available for each road type and information is available or can be estimated for the distribution of traffic volume over the time periods (e.g. 10 % of traffic at night, 20 % of traffic in the weekend), these proportions should be combined in a logical way to calculate a percentage of the traffic volume for each crossed stratum, all summing up to 100%.

(2) If no traffic volume information is available but a reliable estimate of the length of the roads of each road type is available, one could alternatively use the traffic counts from the sessions in the stratum to make an estimate of the hourly number of vehicles at the survey locations (= Nh(k)). If the locations are randomly selected, this average (time-standardized) vehicle count is an estimate of the average hourly vehicle count of all locations in the stratum. This value, multiplied by the estimate of the length of the roads in the stratum – and, if different time periods are considered, the number of hours in the time period considered – should give some estimate of the traffic volume in the stratum. These values could then be used to weight strata.

Let us develop this second approach which is based on road length:

- if there are *M* strata to be aggregated
- let *Ns(i)* be the average number of vehicles per hour (or any other duration standard) for stratum *i* (*i* ranging from 1 to *M*)
- let *Ps(i)* be the relative proportion of the time periods considered (e.g., 5/7 for weekdays, 2/7 for weekend days)
- let *RL(i)* be the total road length of stratum *i*
- *KPI(i)* be the KPI value of stratum *i*

Then:

$$Aggregated\ KPI\ Value = \sum_{i=1}^{M} \frac{Ns(i) * Ps\ (i) * RL(i) * KPI(i)}{\sum_{1}^{M} Ns(i) * Ps\ (i) * RL(i)} \quad (4)$$

Note that $Ns(i)$ is the average number of passing vehicles per hour on the road type (e.g., urban roads) and within the time period (e.g., weekdays) the stratum (i) represents. $Ns(i)$ is equal to the mean of all $N_h(k)$ in the stratum *i.*

As an example, consider the following data for six different strata:

Table 3. Example of data for different strata

i	Road type	Time period	Road length (km)	Ns(i)	PS(i)	KPI(i)
1	Urban	Weekday	10 000	100	5/7	87%
2	Urban	Weekend	10 000	80	2/7	92%
3	Rural	Weekday	25 000	50	5/7	82%
4	Rural	Weekend	25 000	30	2/7	79%
5	Motorway	Weekday	3 000	600	5/7	78%
6	Motorway	Weekend	3 000	350	2/7	74%

Application of formula (4) will then yield an aggregated KPI value of 81.4%.

In order to get an idea of how realistic this approach is (this analysis may lead to rejecting this approach rather than accepting it) it can be bootstrapped. The $Aggregated\ KPI\ Value$ value above depends on the average number of vehicles per hour value $Ns(i)$ which is calculated for each stratum. For each stratum $Ns(i)$ is calculated from the $N_h(k)$values obtained from the survey sessions. The purpose of this bootstrapping approach is to see what values the $Aggregated\ KPI\ Value$ could have attained if the $Ns(i)$ values were consistent with the $N_h(k)$ values, but reasonably different.

A way to do this is for each $Ns(i)$ collect the $N_h(k)$ for k = 1, ..., K. The "bootstrap" way would be of selecting L values (with L<K) from $N_h(k)$, k = 1, ..., *K* and calculate a new value for $Ns(i)$. Do this for each stratum *i* and equation (4) can be applied to obtain a new value of $Aggregated\ KPI\ Value$. When applying this step quite a number of times (with replacing the L values), one gets an idea of how well determined the $Aggregated\ KPI\ Value$ is.

The idea behind this approach is that both $RL(i)$ and $KPI(i)$ are quite accurately known compared to $Ns(i)$ within each stratum. Obviously, $RL(i)$ is constant within the stratum and we ***assume*** $KPI(i)$ is reasonably similar within the stratum (e.g., on motorways at night, you have this percentage of seatbelt use). Assuming this assumption holds, and we took another sample, we would have identical $RL(i)$ and quite similar $KPI(i)$ but only different $N_h(k)$, k = 1, ..., *K*. The best guess for the values the $N_h(k)$, *k* = 1, ..., *K* are the *K* values that were counted. Therefore, we sample with replacement *K* values from that set to get an estimate of $Ns(i)$. If the range of values for $Aggregated\ KPI\ Value$ obtained this way is too large to be useful (e.g., varying with more than 5%), the whole approach is probably not accurate enough. Unfortunately, if the range is too large to be useful, we still have the assumption that the $KPI(i)$ are reasonably similar within each stratum. This may not hold, so we cannot conclude, but we might tentatively assume the approach is not too bad.

Trendline beneficiaries should also report in their metadata whether bootstrapping has been applied.

Reporting:
When reporting results to the project coordinator, Trendline beneficiaries need to report, for each stratum used in the analysis, an estimate of the traffic volume (or at least percentual share of it), since this is a key element in assuring respect for minimal requirements for weighting and to assure internationally comparable results.

Important: if no vehicle counts or no road length information is available, or no otherwise obtained (actual or estimated) traffic volume information, one should only treat the strata separately, and defer from aggregation. In such cases, some of minimum required KPIs in Trendline cannot be delivered.

F. Calculation of confidence intervals (CI)

Calculation of confidence intervals for the data described above is far from trivial. The statistical reference works considered do not precisely cover the sampling problem considered and the methods discussed that appear to be feasible for implementation. Some Trendline beneficiaries appear to use gaussian approximations to statistics to aggregate over sample sessions within strata and aggregate over strata, although there are also some who are using statistical software taking the complex sampling design into account. In general, using gaussian approximations in the aggregation process is acceptable for the averages and percentages themselves but may cause serious problems determining confidence intervals thereof.

Weighting factors for observations within a stratum are given in formula (1) and weighting approaches for aggregation of different strata in formulas (3) and (4).

Trendline beneficiaries should use a method for calculating Confidence Intervals that takes the sampling design method into account, in particular the fact that observations are nested in sessions. Trendline beneficiaries need to indicate in the metadata how they calculated the CIs. Since approximations that assume simple random sampling clearly lead to unrealistically small confidence intervals, approximations using simple random sampling are not acceptable.

G. Using appropriate statistical software

It is advised to use dedicated survey software, as readily available in R and other software packages. Table 1 introduced above and all other variables needed for the weighting will serve as input to these procedures.

Packages that can be considered are:

- R Survey Package https://cran.r-project.org/web/packages/survey/index.html
- STATA Analysis of Complex Survey Data in Stata e.g. https://www.stata.com/meeting/mexico10/mex10sug_canette.pdf
- SPSS: https://www.ibm.com/products/spss-statistics/complex-samples
- SAS: https://support.sas.com/documentation/cdl/en/statug/63033/PDF/default/statug.pdf (hefty document including documentation of proc survey means)

Books considered:

Cochran, W. G. (1977). Sampling Techniques. Wiley
Thompson, S. K. (2012). Sampling. Wiley
Wu, C., Thompson, M. E. (2020). Sampling Theory and Practice. Springer International Publishing