Trend line

KPI safety belts and child restraint systems

Methodological guidelines

Version 2.2, January 4, 2024

Co-funded by the European Union

KPI safety belts and child restraint systems. Methodological guidelines

Please refer to this report as follows: Kšicová E., Forsman Å., Areal A., Laiou A., Valentová V., Temmerman P., Boets, S. (2023). KPI safety belts and child restraint systems. 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 Research for Road Safety Henri Faasdreef 312 – 2492 JP The Hague, The Netherlands
	Co-funded by the European Union

Report details:	
Version: **Dissemination level:** **Date:**	2.2 Public 04/01/2024

Report Author(s):

Kšicová, E., (CDV), Czech Republic
Forsman, Å., (VTI) Sweden
Areal, A., (PRP), Portugal
Laiou, A., (NTUA) Greece
Valentová, V., (CDV), Czech Republic
Temmerman, P., (Vias institute), Belgium
Boets, S. (Vias institute), Belgium

Revision history

Date	Version	Description
31/03/2023	1.0	Version from BASELINE project for the first meeting
16/04/2023	1.1	First draft version - modifications according to the discussions and conclusions from the BASELINE project
12/05/2023	2.0	Second draft
21/11/2023	2.1	Revision with clarifications
04/01/2024	2.2	Version published at the website

Legal Disclaimer

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

Passangeer car
Motor vehicle with 3 or 4 wheels, mainly used to trasnsport people, seating forno more than 8 occupants (excluding the driver).

Light goods vehicle
Goods vehicle under 3,5 t maximum gross weight: Smaller motor vehicle used only for the transport of goods.

Heavy goods vehicle
Goods vehicles over 3,5 t maximum gross weight. Larger motor vehicle used only for the transport of goods.

Driver and passenger
The driver is the one who controls the vehicle, the passenger just rides.

Motorways (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.

Urban roads (or road inside urban areas)
Public roads inside urban boundary signs.

Rural roads
Public roads outside urban boundary signs, excluding motorways and expressways.

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

Weekend – daytime
Saturday to Sunday 6.00 a.m. to 9.59 p.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 implementation 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 methodological guidelines for the KPI safety belts and child restraint systems. It describes the minimum methodological requirements to qualify for this KPI, defined as[1]:

Percentage of vehicle occupants

using the safety belt or child restraint system – part A and

using the safety belt or child restraint system correctly – part B

[1] This is an adaptation of the original KPI definition in the SWD (2019) "Percentage of vehicle occupants using the safety belt or child restraint system correctly" in order to distinguish the two different methods that are needed to collect the required data for this KPI.

The minimum requirements set by the EC for this KPI are described in the Commission Staff Working Document SWD (2019) 283 (see Appendix 1). These requirements are further quantified and specified in this document. This document is based on a review of the methodological guidelines that were developed within the Baseline project (Temmerman et al., 2021), expert consultation within the Trendline Key Expert Group and the recommendations from the Baseline KPI report on Safety Belt and Child Restraint Systems (Van den Broek et al., 2022).

Each section also includes optional supplementary methodological recommendations. Member States can decide whether to follow the minimuml requirements only or to extend (part of) their methodology, depending on available means and own research questions.

The target audience of this document are the persons in the participating Member States that will collect and/or analyse the data to deliver this KPI.

2. Scope

2.1. Two main methods

Given the complexity of determining the correct use of CRS, two types of observations are recommended:

1) Roadside observations: these allow collecting data on the use of safety belts and the number of children in the car and the presence of CRS can be observed for quantitative purposes – part A.

2) In-vehicle inspections; these allow collecting data on the correct use of CRS. This method requires the cooperation of the driver and is only possible during dedicated sessions in accessible locations – part B.

It is recommended to conduct both types of observations. However, since part B can be difficult to carry out in some countries (for example due to legislation), only part A is required and part B can be omitted. The reasons will need to be reported in the meta data file.

2.2. Definition of correct use, no use, and misuse

The objective is to estimate the percentage of vehicle occupants using a safety belt or child restraint system (CRS) through roadside observations (Part A) and using a child restraint system (CRS) correctly through in-vehicle inspections (Part B). The theoretical population refers to the total of all movements with the vehicles over the national territory. In other words, this reflects the total number of kilometres driven. Hence the percentage of vehicle occupants using a safety belt or child restraint system (Part A – roadside observations) and correct use of CRS (Part B – in-vehicle inspection) is related to the percentage of kilometers driven using a safety belt or child restraint system (Part A) and correct use of CRS (Part B).In some states, following the in-vehicle inspection (part B) is complicated due to GDPR and others, it is possible to omit it and perform only part A.

References for correct use should consist of:

- The national traffic legislation;
- The CRS's conformity and instruction label.

It is not required to take into account additional (national) recommendations for the optimal use of CRS (e.g., the Swedish recommendation to use a rearward facing CRS up to and including 4 years of age). Compliance with such recommendations could be included as optional information.As additional information in the metadata, indicate in abbreviated form the method of using the CRS for the given country.

Monitoring of safety belt and CRS parameters can be carried out in two ways; (A) road side observations (A); and (B) in-vehicle inspections. Correct use of CRS can only be assessed through in-vehicle inspection (B). In roadside observation studies only "use" or "no use" can be observed.

In method (A) this implies that observation of the "use" of safety belts may also include the incorrect use of safety belts, for example a belt under the arm, incorrect height adjustment of the upper safety belt guide, using "foreign objects" such as clothespins to deflect the safety belt or reduce its tension, etc.

In-vehicle inspection of CRS (B) should lead to one of the following statements (per child observed): "Correct use", "Misuse", or "No use".

Correct use is the complement of 'no use' and 'misuse'. As a result, both 'no use' and misuse must be detected. If there is no indication of no use or misuse, the usage is considered to be correct.

Possible ***misuses*** (non-exhaustive) of ***CRS*** can be grouped into 3 types:

- Inappropriate use
    - Child not in CRS while it should be (= no use),
    - Child in wrong group of CRS,
- Faulty fixation of CRS to vehicle.
    - Incorrect safety belt guidance around CRS,
    - Back tether or floor support (as complement to Isofix) not attached,
    - CRS wrongly orientated,
    - Frontal airbag not deactivated with rearward mounted CRS on place with frontal airbag.
- Faulty fixation of child to CRS
    - Belts too loose,
    - Wrong belt guidance.

Possible ***misuses*** of ***safety belts*** are (non-exhaustive):

- Belt behind the back,
- Belt under arm,
- Incorrect height setting of safety belt's top guidance,
- Use of 'foreign objects' such as clothespins to deviate the safety belt or reduce its tension.

2.3. Road types / vehicle types / occupants

SWD requires taking into account the following strata:

- Road type (**urban roads, rural roads, motorways including expressways if considered**),
- Week period (**weekday, weekend day**).

It is also recommended to collect:

- Place in vehicle (driver, front safety passengers, rear occupants),
- Age group (child 0-17, adults 18 and more),
- Sex (male, female).

Moreover, passenger cars are mandatory, other vecihle types are optional (goods and heavy goods vehicles).

2.4. Minimum sample size

For in-vehicle inspection of correct CRS use a minimum of 2000 observed vehicles with children among the occupants is more difficult to attain. The following sample sizes are requested:

Part A – roadside observations:

- **A minimum of 2000 observed vehicles overall for safety belt use, with a minimum of 500 observations per road type and per week period;**

Part B – in-vehicle inspection:

- **A minimum of 200 observed vehicles with children among the occupants for detailed in-vehicle inspection of child restraint system use, with a minimum of 50 observed vehicles per road type and per week period**.

Appendix 2 gives an overview of the argumentation behind the minimum sample size.

Beneficiaries not able to achieve the minimum requested number of observations need to justify this in detail.

If Beneficiaries aim at distinguishing **regions** in the reported results (not required for Trendline), the minimum numbers of observations apply to each region. If Beneficiaries aim at distinguishing **vehicle types** in the reported results (not required for Trendline), the minimum numbers of observations apply to each vehicle type. If only passenger cars are considered or if there are insufficient observations of other vehicle types, the minimum numbers of observations apply to passenger cars (only minimum required vehicle type for Trendline).

Multiplying the minimum sample size can increase the accuracy of the estimates and allow delivery of additional disaggregated KPIs (e.g. for crossed strata like road type x week period). This is a decision up to the Beneficiary.

3. Observation method

3.1. Observation method

Part A – roadside observations:

SWD prescribes direct observation as the data collection method. Direct observation by trained observers should preferably be carried out alongside the road.

SWD also allows the use of cameras to collect data on safety belt use. In that case, it should be ensured that the cameras are installed on all road types to avoid selection bias. This technology could have advantages compared to using observers in terms of, for example, reliability, 24/7 observation, night-time use etc. Possible disadvantages should however be evaluated (e.g. lacking variables, visibility of rear occupants etc.). Its use should be tested and validated before deployment. For privacy reasons, faces and license plates should not be caught on camera. Each Member State will have to conform with national and international requirements regarding ethics, privacy, and data protection (GDPR).

Other methods are also possible, e.g. with observers in moving vehicles on the roads, if the accuracy of observation is sufficient.

Part B – in-vehicle inspection:

Determining the correct use of CRS requires detailed in-vehicle inspections. These sessions can take place in accessible locations such as parking lots, rest areas, etc. and require the driver's voluntary cooperation. Selection bias is inevitable in a survey based on voluntary participation. However, it is the only option for reliably detecting the correct use of CRS.

In one vehicle, it is possible to collect data on several children with a safety belt or in CRS. It is possible to include the monitoring of safety belt use in adults as well.

3.2. Coverage of road types

The indicator should cover **motorways (including expressways if these are considered), rural non-motorway roads (outside built-up areas), and urban roads (inside built-up areas)** (for definitions see *Terms and definitions*).

This is the minimum required categorisation. The results should be presented separately for these three different road types and also aggregated (after weighting) for the whole road network. A deviation from this requirement is only possible in the exceptional case that a specific road type is non-existent in a country (e.g., no motorways in Latvia and Malta).

When a Member State's road network does contain all required road types, but not all road types are included in the survey, the results for the remaining road types cannot be aggregated by the remaining road types and remain disaggregated for each remaining road type.

Countries that used another definition for the road types in Baseline might be asked to recalculate their Baseline KPIs according to the Trendline definition (if feasible) to evaluate the impact of changed definitions on the key estimates.

Roadside observations on motorways can be organised on the motorway but also at entrances or exits of motorways for feasibility reasons. This should be indicated in the meta-data.

Main characteristics of the included road types should be described in the meta-data (e.g., signs, speed regimes, number of lanes, lane separation, allowed vehicles) which allows to assess general correspondence of the road types between the countries (background/contextual information).

3.3. Selection of locations

Since SWD requires coverage of the three road types, the proportion of observations sampled at each of the three road types should be at least 20% to ensure a minimal number of observations for each stratum, even if this implies disproportionate sampling. Sample size could be allocated to the three road types according to traffic volume, assuming each of the three road types represents a share of traffic volume above 20% based on available national traffic volume data (e.g. estimates of vehicle kilometres driven per road type based on national traffic surveys). If such data is not available, a minimum number of 10 locations per road type should be selected for the national indicator.

The selection of locations should be as random as possible, covering the entire geographical area of the country. There are different options for random location selections: simple random, stratified random, cluster random etc. Cartographic software like ArcGIS can be used for selecting random points, e.g. https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/anhttps://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/an-introduction-to-sampling-monitoring-networks.htmintroduction-to-sampling-monitoring-networks.htm

Bigger countries may consider in a prior stage the selection of one or more regions/states which are considered to be representative for the country with regard to using restraints. This can add to the fieldwork feasibility. If this is done, it should be explained in the meta-data.

The basic procedure to randomly select locations consists of three steps:

1. The required number of locations (for the country or per region) is determined.
2. The number of locations is randomly selected on a map using the entire area in question (e.g., country or region), taking sufficient geographical spread into account. The specific requirements for each location (e.g., feasibility, visibility) do not have to be considered at this point yet. This step is to ensure a reasonable geographical spread of the randomly selected locations.
3. The final locations that will be used for the observations are manually chosen in the area surrounding the locations randomly selected in the previous step. At this point, the final selection must be based on the location requirements (different road types; for in-vehicle inspection: parking lots), inclusion/exclusion criteria (if applicable) and practical considerations (e.g., for roadside observations: locations where drivers must slow down). This final selection can be made using Google Street View or in cooperation with the police unit responsible for the respective location. Care should be taken to ensure that the locations for the different road types are sufficiently spread geographically.

The appropriate sample size should be estimated and used to determine the required number of locations or observational sessions, taking different vehicle types into account. For more information on random sampling of locations and for determination of the minimal sample size, reference can be made to 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.

Sample size calculators can be used to calculate the required minimal number of observations: e.g. https://samplehttps://sample-size.net/confidence-interval-proportion/size.net/confidence-interval-proportion/ (software determining the upper and lower bounds of the confidence interval for a proportion).

The rationale for choosing the observation locations should be documented. These include a minimum traffic flow (e.g. at least 10 relevant vehicles per hour) and a random selection of different regional locations. Ideally, a random sample of all possible locations within a designated area will be used. A random selection of locations will also include roads with low traffic volume. In that case, it is recommended to choose a nearby road with a higher traffic volume instead, if it is assumed that most drivers on the low-volume road drove or will drive on the high-volume road as well. Locations with less than 10 relevant vehicles passing per hour cannot be used. Member States can define a higher minimum.

In the next step data collection sessions are determined by attributing a week period (weekday or weekend day) to each selected location, balanced over the three road types and geographical spread. For in-vehicle inspection this allocation also depends on the type of parking (related type of activity/service of the parking, e.g., school, shopping). If locations per road type and week period are not sampled proportionally (which is the case when when the same number of locations is selected for each stratum, like the minimum 10), stratification weighting is needed in order to estimate nationally representative KPIs (specific guidelines are included as Appendix 3 of this document).

The required minimum number of different locations depends on the method.

Part A – roadside observations:

As an absolute minimum 10 different locations per level of stratification variable are required with the aim of getting sufficient data for the entire road network (three road types) and the two required week periods for meaningful KPI estimates:

- at least 10 locations on urban roads,
- at least 10 locations on rural roads,
- at least 10 locations on motorways (including expressways if considered,
- at least 10 locations on weekdays,
- at least 10 locations on weekend days.

The absolute minimum is thus 30 different locations. Each location can be used for different sessions (at different time intervals) or each location can be assigned (randomly) to a specific time interval.

To ensure a balanced sampling for each combination of road type (3) and week period (2), a minimum of 2 different locations for each combination of strata (i.e. 6 crossed strata) should be used:

- Urban roads x weekdays: minimum 2 locations,
- Urban roads x weekend days: minimum 2 locations,
- Rural roads x weekdays: minimum 2 locations,

- Rural roads x weekend days: minimum 2 locations,
- Motorways x weekdays: minimum 2 locations,
- Motorways x weekend days: minimum 2 locations.

The requirements concerning motorways (including expressways if considered) do not apply to countries with no motorways or where the network of motorways is very limited.

Ideally more than 10 locations for the different strata and more than 2 locations for the crossed strata are used for sampling. The recommendation is to boost the sample to allow a more accurate estimation of disaggregated indicators.

In the study's data basic characteristics of the locations should be documented - like the geographical coordinates (if possible), address or other geographical information, number of lanes, target lane and direction to be observed, and visibility of the traffic from the location.

Part B – in-vehicle inspection:

The minimum number of sites for in-vehicle inspections of correct CRS use is 2 for each combination of week period and road type (6 combinations if all road types are covered).

MS aiming at getting more accurate estimations of correct CRS use are advised to boost the sample to at least the same number of locations as for the roadside observations (minimum 10 per stratum).

In the study's data basic characteristics of the locations should be documented - like the geographical coordinates (if possible), address or other geographical information, location type, related activity/service type (e.g. parking lot of school, shop, day-care etc.).

3.4. Observation sites and methods for different road types

For direct observations, strong wind, precipitation, and very low or high temperatures could negatively affect the observers' endurance and observation quality and recording quality (for camera recording). The road-side observations should be performed during reasonably good weather. The same applies to the in-vehicle inspections of correct CRS use.

Part A – roadside observations:

Observations of safety belt use on urban and rural roads should be carried out from a safe place along the road, preferably at locations where driving speed is reduced relative to the speed limit, such as entrances to cities/vilages or at intersections where vehicles slow down or stand in front of the traffic lights.

Observations on motorways are for example possible at:

- the last intersection before on-ramps to motorways,
- the first intersection after an off-ramp from a motorway,
- entrances and exists from,
- service stations,
- rest areas,
- toll stations etc.

Part B – in-vehicle inspection:

In-depth investigations of correct CRS use are carried out by observers that approach drivers that have just parked (before children are possibly unbuckled) or that are just ready for departing the parking spot (when children are expected to be just buckled up).

Observations of child restraint system use on urban and rural roads can be carried out at parking lots of places where children are expected (also considering the proper time slots), like day-care, schools, shops or places for leisure activities. In-depth investigation of child restraint system use on motorways is possible at parkings of service stations or rest areas.

3.5. Observation sessions

Each observation session should last at least 30 minutes, although a duration of 1 hour is advised. It should be kept in mind that this minimum session duration requirement does not include the time spent on traffic volume counting (see section 'Measuring traffic volume' below). Date and time (to the nearest hour) covered by the measurements should be indicated in the meta-data.

4. Other requirements and options to be considered

4.1. Vehicle types and occupants to be considered

Part A – roadside observations:

The road users to be observed are driveres, front passengers and rear occupants of at least passenger cars and optionally of goods vehicles as well (ideally separating light goods vehicles (LGV/vans) and heavy goods vehicles (HGV/lorries)).

At a minimum, separate test results for drivers, passenger car front occupants (wihtout drivers) and passenger car rear occupants are expected. If other vehicle categories are also included in the study, these results should be reported separately.

The different vehicle types and their specific categorization should be clearly defined and illustrated for the observers (training, briefing), e.g. some cars and LGVs share the same brand/model like Renault Kangoo (a passenger car has a backseat windows and passenger seats; a LGV has no backseat windows and no rear passenger seats).

Vehicles to be observed should be randomly selected from all the possible objects at the location where the observation is done. After coding one observation, the next passing target vehicle should be observed.

Vehicle occupants legally exempted from safety belt wearing should be excluded, e.g. postal delivery services, taxi drivers, emergency vehicles, etc. The most practical solution is to exclude the whole vehicle from the data collection. Because the legislation on (and exemptions from) safety belt use and on CRS use can vary between countries, it is requested that all countries document their legislation on safety belt use and CRS use and consequently document which vehicles were excluded from the observations.

Supplementary to safety belt usage, it might be valuable to include one or more of the following occupant characteristics for further analysis :

- Sex (observed),
- Age group (observed).

Age groups are divided as follows: child: 0-17, young: 18-24, medium: 25-64, senior: 65+.

Part B – in-vehicle inspection:

Since very few children are expected to be travelling in goods vehicles, it is recommended to only include passenger cars in the CRS observations.

During the in-vehicle inspections on correct CRS use, it might be valuable to include one or more of the following trip characteristics for further analysis:

- Trip purpose (question to driver),
- Trip length/duration (question to driver).

Supplementary to safety belt usage, it might be valuable to include one or more of the following occupant characteristics for further analysis (appropriate for in-vehicle inspections):

- Child sex (asked),
- Child age group (asked) – if the counry needs it.

Child age groups can be divided as follows: 0-10; 11-17 – but this is according to the regulation and interests of the country.

In in-vehicle inspections also the driver's age and sex can be estimated or asked to evaluate the link with CRS correct use – if countries are interested.

4.2. Optional breakdowns by region

Optionally, Member States can decide to distinguish different regions in the survey. In that case, countries can consider collecting data from each region or from a representative selection of regions. Member States wishing to have meaningful KPIs at regional level should take into account that the minimum sizes of the location sample and driver sample should ideally be applied in each region. If stratification in regions is used, results should be weighted according to traffic volumes by region.

5. Data analysis

5.1. Data to be recorded

Data to collect with regard to the **observation locations**:

- Unique location ID,
- Region (if applicable),
- Road type,
- Road number, address,
- Coordinates of exact observation spot (either here or in observation session details),
- Number of lanes,
- Target lane and direction to be observed (either here or in observation session details),
- Visibility of the traffic from the location (either here or in observation session details),
- In case of CSR inspection: related activity/service type (e.g. parking lot of school, shop, day-care etc.).

Data to collect with regard to the **observation sessions**:

- Unique session ID,
- Location (from which road type can be derived),
- Date (from which time period can be derived),
- Begin time of observations,
- End time of observations,
- Total duration of observation session (end time – begin time – count duration),
- Traffic count duration (not for (B) in-vehicle CRS inspection),
- Traffic count results per relevant vehicle type (not for (B) in-vehicle CRS inspection),
- Traffic count results per relevant vehicle type extrapolated to session duration (not for (B) in-vehicle CRS inspection),
- Short weather description.

Data to collect with regard to **roadside observations** (one data point = one observed vehicle):

- Vehicle type,
- Driver safety belt use (use / no use),
- Front passenger 1 safety belt use (use / no use),
- Front passenger 2 safety belt use (use / no use),
- Rear passenger 1 safety belt use (use / no use),
- Rear passenger 2 safety belt use (use / no use),
- …

"Use" or "no use" should be recorded. It is reminded that "use" can also include the incorrect use of safety belts.

Optionally, estimated age group, sex and other additional variables can be recorded per occupant as well.

Relevant data to collect with regard to **in-vehicle CRS inspections** – some variables depend on national regulation e.g. on child age and/or height (one data point = one observed vehicle):

Child characteristics	• Place of child in vehicle, • Frontal airbag on place of child (not present / activated / deactivated), • safety belt type on place of child (not present / 2-point / 3-point), • Isofix on place of CRS (correct / false), • Age of child, • Length of child, • Weight of child.
Safety belts / CRS characteristics	• Is the child fastened in the belt or in the seat (yes / no), • Is the seat, if any, fixed in/to the car? (yes/no), • Use of belt or CRS according to the age and/or length) of the child (correct / misuse), • Orientation of CRS (forward, rearward, sideways), • CRS group, • CRS homologation label.
Use of safety belts / CRS	• Safety belt guidance (correct / false / NA), • Safety belt tension (correct / too tight / too loose / NA), • CRS belts guidance (correct / false / NA), • CRS belts tension (correct / too tight / too loose / NA).

Requirements for the data delivery and data atrix for the Trendline dataset is provided in a separate document (results acquisition template).

In the case of observing adults in in-vehicle inspection, it is advisable to supplement data on the space in the vehicle (number of seats) and the use of safety belts (correct use / misuse / no use).

5.2. Stratification weights and statistical analysis

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

For method B (vehicle inspection), the number of 'observed vehicles with children among the occupants' should serve as a quantitative basis for the weighting of the qualitative data gathered with the in-vehicle inspections of CRS use.

5.3. Measuring traffic volume

For the roadside observations, traffic counts should be performed at each location and each observation session. This information is necessary to correctly calculate the confidence intervals and weighing factors. or the qualitative in-vehicle inspections of correct CRS use this is not necessary.

Traffic volumes should be estimated by traffic counts during the observation session: ideally either by counting all passing relevant vehicles (only the vehicle categories that are being observed) during the session, or by counting all passing relevant vehicles during a short interval in the middle, or partly before and partly after the measure. The counting should be done for the same vehicle categories at the same location and direction as the observations. The counting of all relevant vehicle categories should last at least 10 minutes. Optionally, an automatic counter can be used to determine traffic volume. Note that

in that case it might not be possible to exclude certain vehicle types. When observing at service stations or rest areas, the traffic volume to consider are the vehicles entering the service station or rest area.

These counts should then be extrapolated to the whole duration of the session. **It is required that the number of counted vehicles and the duration of the count are always coded together in order to be able to correctly calculate the number of passed vehicles per minute** (i.e. the traffic volume during the session) and so to avoid mistakes in that calculation. More information concerning measuring traffic volume can also be found in *Appendix 3*.

5.4. KPI values to provide

The main KPI to be provided is the percentage of vehicle occupants using restraints (A) and correctly using CRS (B) at all times and locations. At a minimum, the percentage of safety belt use by the driver, the front passenger in a passenger car, the percentage of safety belt use by the rear passenger, and the correct use of child restraint systems should be reported. The equivalent percentages in goods vehicles is optional. Results should also include the unweighted number of drivers the result is based on (as well as other variables, which will be included in the Trendline Datafiles).

Roadside observations:

As minimum requirement a point estimate and a corresponding 95% confidence interval is expected for each level of the following stratification variables:

- Nationally weighted aggregate indicator for correct safety belt use (all occupants, front occupants, rear occupants, drivers, front passengers),
- Nationally weighted aggregate indicator for CRS use,
- Indicators by road type (3 levels: motorways– including expressways if considered, rural non-motorway roads, and urban roads),
- Drivers vs front passengers vs rear occupant (in case of safety belt use in passenger car),
- Indicators by Week period (2 levels: weekdays vs weekend),
- Vehicle type (passenger cars).

Specific estimates for combinations hereof and for additional subgroups (e.g. goods vehicles) are recommended ifcountries have sufficient sample sizes for these.

For the data delivery to the Trendline coordinator, a specific results acquisition template is available.

Member States should provide the meta-data of their data collection and deliver this together with the dataset(s). Final info on this will be provided together with the Trendline Datafile info.

5.5. Temporal requirements

Observations should be timed as follows:

- late spring or early autumn. All months are allowed except for December, January, July and August. Holiday periods (bank / school holidays) and hard winter conditions should be avoided, as these disturb normal traffic patterns. In some Member States, the Winter or Summer holiday period could extend to other months as well, such as June, and in such cases these months should also be excluded.
- week days (excluding bank holidays) and weekend, observed and presented separately,
- daylight – observations should cover the whole daytime,
- reasonably good weather.

There should be a balance between all combinations of road types (3) and the different time factors above, to avoid a systematic sample bias.

Where Member States have historical series of measurements, it is recommended to use the same period(s) of the year as for the earlier measurements.

Member States willing to organise more than one roadside survey to deliver the KPIs (e.g. one in spring and one in autumn) can apply the minimal sample size requirements on the combination of both measurements. The data of both measures can be combined to deliver the main and disaggregate indicators.

It is furthermore recommended not to plan data collection in case the traffic situation and mobility patterns in a (large part of the) country are very different from the normal situation (low representativity) (e.g. COVID-19).

6. Summary of requirements and recommended options

SWD requirements are:

- Percentage using correctly safety belt and CRS,
- Method: observation,
- Road type: rural, urban, and motorways,
- Vehicles: passenger cars; goods vehicles optional,
- Front and rear seats,
- Child restraints vs safety belt,
- Location: random,
- Time: day,
- Day: week and weekend,
- Month: spring/autumn.

Within Trendline the additional or more specific requirements are:

- Percentage using safety belt and CRS (roadside observation method, part A)
  - Driver, front occupants, front seat passengers, rear occupants,
- Percentage using correctly safety belt and CRS (in-vehicle inspection, part B)
  - Only for children in CRS.

Part A – roadside observations:

- Direct observation : observers along the road or in moving vehicles or use of a camera,
- A minimum of 2000 observed vehicles overall for safety belt use, with a minimum of 500 observations per road type and per week period.

Part B – in-vehicle inspection:

- Min. 200 inspected vehicles with children that should legally be restraint with a CRS,
- a minimum of 50 observations per road type, and per week period,
- adult observation is also possible.

Recommended options:

- Also goods vehicles (ideally split LGV and HVG) – for Part A,
- Boost number of locations and sample size for more accurate indicators, also for more crossed strata and subgroups (Part A) and to allow a quantitative approach for estimating valid indicators of correct CRS use (Part B),
- Geographical coverage,
- Complete disaggregated data (if the sample size is sufficient).

References

European Commission (2017). Monitoring Road Safety in the EU: towards a comprehensive set of Safety Performance Indicators. European Commission, Directorate General for Transport https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/ersosynthesis2017-detailhttps://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/ersosynthesis2017-detail-performanceindicators15_en.pdfperformanceindicators15_en.pdf

European Commission (2019). Commission staff working document EU road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero". SWD (2019) 283 final. Retrieved from

https://transport.ec.europa.eu /system/files/2021-10/SWD2190283.pdf

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

Temmerman, P. et al. (2021). Methodological guidelines – KPI safety belts and child restraint

systems. Baseline project, Brussels: Vias institute

Van den Broek B., Aarts, L. & Silverans, P. (2022). Baseline report on the KPI Safety belt and Child

restraint systems. Baseline project, Brussels: Vias institute

Appendix 1 KPI for the use of safety belts and CRS

Ref: Commission Staff Working Document - EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero, SWD (2019) 238, https://transport.ec.europa.eu/system/files/2021-10/SWD2190283.pdf

Rationale

The use of the safety belt and child restraint systems is an essential element of passive safety. A significant proportion of fatally or seriously injured vehicle occupants have not used the safety belt or child restraint system correctly.

Definition of the KPI

Percentage of vehicle occupants using the safety belt or child restraint system correctly.

Minimum methodological requirements

Data collection method	Direct observation (if appropriate, using cameras).
Road type coverage	The indicator should cover motorways, non-urban roads and urban areas. The results could be presented separately for the three different road types if available.
Vehicle type	The indicator should include passenger cars as a minimum and goods vehicles (results shown separately) where possible.
Front and rear seats	For passenger cars the results should be presented separately for front and for rear seats.
Safety belts vs. child restraints systems	Safety belt and child restraint systems to be differentiated in the data collection.
Location	Random sample (methodology for Member States to decide).
Time of day	Observations to take place during daylight.
Day of week	Separate observations for week days and weekend and data to be shown separately.
Month	Late spring, early autumn.

The additional requirements established within Trendline are:

Data collection method	Part A: Direct observation on the road (if appropriate, using cameras, from moving vehicles...) for use of safety belts and of CRS. Part B: In-vehicle inspection for correct use of CRS.
Person in the vehicle	Part A: For passenger cars the results should be presented separately for drivers, front passengers, front occupants and rear occupants. Part B: Children in CRS, adult observation is also possible.

In the BASELINE methodology, safety belts were collected with a division into use / incorrect use / non-use. This division was changed based on experience with data inaccuracy - for part A - only collection will be carried out by roadside observation with a division into use/non-use. In case of detection of incorrect use (safety belt under the arm, etc.) observable by roadside observation, this behavior will be classified as "use".

Appendix 2 Rationale for 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 defined in 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 estimates for 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 and affordability, as a rule of thumb the minimum total number of observations was set (for part A – roadside observations) at 2,000, the minimum number of observations for different strata at 500. For part B – in-vehicle inspection - A minimum of 200 observed vehicles with children among the occupants for detailed in-vehicle inspection of child restraint system use, with a minimum of 50 observations per road type. 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 sampling of 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 mainly depends 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 ensure representativeness 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.

Appendix 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[2] 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[3]. 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.

[2] 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).

[3] 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[4]. 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:

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

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)*:

[4] 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.

(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:

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

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