Baseline

Considerations for sampling weights in Baseline

Version 1.3, December 6th, 2021

Baseline

Belgium | Austria | Bulgaria | Cyprus | Czech Republic | Finland | Germany | Greece | Ireland | Latvia | Lithuania | Luxembourg | Malta | Netherlands | PolandPortugal | Slovakia | Spain | Sweden

baseline.vias.be

Project:	This document has been prepared in the framework of the BASELINE project, for which a grant has been awarded by the European Commission. Information on this project can be found on the website www.baseline.vias.be
References:	(1) Grant agreement under the Connecting Europe Facility (CEF) No MOVE/C2/SUB/2019-558/CEF/PSA/SI2.835753 collection of Key Performance Indicators (KPIs) for road safety (2) Consortium agreement among the 19 partners of the Baseline project
Authors:	Silverans, P., & Boets, S.
Publisher:	Vias institute, Brussels, Belgium (www.vias.be)
Please refer to this document as follows:	Silverans, P., & Boets, S. (manuscript in preperation). Considerations for sampling weights. Baseline project, Brussels: Vias institute"

Any comments or feedback regarding these guidelines, should be sent to baseline@vias.be.

Version history

Version	Date	Changes
1.1	November 5th, 2021	Draft revised after discussion in KPI expert groups and technical committee.
1.2	November 15th, 2021	Tables and examples revised after review
1.3	December 6th, 2021	Correction of reference to inclusion probability in Tabeles 1 and 2 (higher corrected into lower)

Contents

1 Introduction

1.1 Objectives and importance of weighting observations

The aim of this document is to provide a guideline on data weighting for Member States participating in the Baseline Behavioural KPI measurements requiring assistance for calculating sampling weights.

As a general rule, weighing of data is the responsibility of the member states. The methods used for calculating sampling weights should be described in detail in the meta-data.

For countries seeking assistance in calculating sampling weights, a possible weight calculation method is described for Member States not using own validated weighting procedures.

The Baseline behavioural KPI's are requested in the format "Percentage of vehicles/road users … *travelling within the speed limit, using the safety belt or child restraint system correctly, wearing a protective helmet, driving within the legal limit for blood alcohol content, NOT using a handheld mobile device.* The overall aim though of Baseline is to estimate KPI's that are representative for the total of all kilometres driven (traffic volume) in a country[1], or more specifically, for the total of all kilometres driven in a country "by the vehicle/road user type(s) on the road types and within the time periods that are required in the KPI measurement".

The main aim of the estimates is to estimate the percentage of kilometres driven on the entire road network (over a period of time, which one could set to one year for instance) by vehicles respecting the legal limits and rules. In term of sampling this means that the population is the total traffic volume (hence 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 questions 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 confirming to the rules in the total population.

1.2 Principles to be considered

The actual vehicle/road user type(s), road types and time period(s) to be included as a minimum in the KPI measurements are defined in the Methodological Guidelines for each of the Baseline KPIs. These subgroups are referred to as strata, which are for most behavioural KPI's:

- Road type: 3 strata: motorways, rural roads, urban roads (except for KPI Helmet cyclists/mopeds: only urban and rural)
- Time period: 1 stratum (KPI distraction: daylight), 2 strata (KPI seatbelt and helmet: week day; weekend day) or 4 strata (KPI alcohol: week day; week night; weekend day; weekend night)
- Vehicle/road user type: from one to several modes per KPI; mostly passenger cars as required as a minimum (except for the KPI helmet and distraction).

For each KPI, minimum sampling requirements are determined per stratum, and if applicable also for combinations of strata (road type x time period), which refers to stratified sampling.

To make sure that a KPI measurement sample is representative for a country as a whole, it should be checked whether the presence of the specific subgroups or strata in the sample (sample strata proportions) are comparable to those in real life.

A theoretically possible strategy to estimate the overall national KPI would be to sample units (vehicles/road users) in all strata (if applicable in all combinations of strata) proportional to the actual traffic volume distributions of the strata, also called proportional stratified sampling, or to use other allocation methods. Such strategies require prior knowledge on the traffic volumes and variances in the different strata though, which most participating Member States do not have. Nationally spread traffic counters or representative mobility surveys can be used to estimate driven kilometre proportions, on the different road types and in the different time periods and ideally just for the vehicle/road user type(s) of interest.

---

[1] For some KPIs the definition deviates from this general approach. For the KPI on Speed, for instance, only vehicles which can freely choose their speed are included.

Often the method for collecting data for behavioural KPI measurements will be based on disproportional sampling, since a strategy of proportional stratified sampling could be detrimental for the accuracy of specific low-volume strata that are of interest.

For example, for the KPI of DUI alcohol (four required time periods: week day, week night, weekend day, weekend night) for instance, night-time drivers, and more particularly weekend night drivers, are at a significantly higher risk for drink driving than weekday daytime drivers. Since traffic volumes during weekend nights are generally very low, strict proportionate sampling according to traffic volume data would lead to much wider confidence intervals (less accurate estimates) for weekend night drivers than for higher volume time periods. In disproportional stratified sampling, certain strata are thus over- and/or under-sampled compared to reality. Therefore it is required to "weight" the data collected in order to estimate the national aggregate KPI, which should reflect the real life strata proportions. MS that will strictly follow the Baseline minimum sample requirements (e.g. for KPI alcohol: min. 10 locations/500 sample units per road type, min. 10 locations/500 sample units per time period, and min. 2 locations per road x time period) use disproportional sampling.

Member States can use their own sampling and weighting techniques, taking the above mentioned strata and principles into account, and complying with the minimum sample requirements from the KPI guidelines. If this is not possible, the minimum sampling procedures described in the KPI guidelines and the weight procedure described in the next paragraph can be used.

All Baseline project participants should include the specificities of their sampling and weighting practices in the Datafile metadata and in the methodological report.

2 Proposed weight calculation method

This section describes a rather simple weight calculation method based on a minimalistic weight formula, which does not start from existing national traffic volume information per stratum but includes primarily data which is commonly available for Member States. At the end of this paragraph an extension of the formula is proposed in case reliable traffic volume data are available.

2.1 Probability of vehicles or road users to be included in observations

Every vehicle/road user in the population has a different probability of being included in the sample and that this should be accounted for in the analysis. In other words, it is likely that proportionally more observations will be available for certain subgroups in the population, meaning that other subgroups will be represented with less observations in the sample. If proportionally smaller samples represent more individuals in the population, the smaller observation numbers should be weighted more strongly in the final population estimate.

The selection of vehicles/road users during roadside observation studies is typically done in two steps:

1) Step 1 – Strata sampling: locations for the required road types are selected from all possible locations on the road network and required time periods are attributed from all possible week time periods.
2) Step 2 - Vehicle/road user sampling: in each session vehicles/road users are selected from all passing vehicles/road users.

Each of the two steps corresponds to the selection of a number of items from a "population". In the first step, the population in question is that of "all the locations and time periods for which it is theoretically possible to carry out behavioural measurements on the national territory". In the second step, the population is "all vehicles/road users who have passed during the selected measurement session".

With regard to Step 1, it is in general a question of "selecting" a given place on the road network and attribute a given time of the week. However, not all elements of this population have the same probability of inclusion in the sample. The probability of inclusion is higher for locations on a road type that constitutes the biggest part of the entire road network (e.g. a Member State can have a high proportion of rural roads and almost no motorways)[2].

[2] This will of course depend on how you allocate the number of locations per strata. But if you select the same number of locations per road type, then the inclusion probability will be smaller for locations on road types that constitutes the biggest part of the entire road network.

The probability of inclusion is also higher for sessions taking place during the day on weekdays (6am-22pm) as compared to during weekend days, as the proportion of weekdays is bigger than the proportion of weekend days in an entire week.

In Step 2 it may be possible to sample all (relevant) vehicles/road users in a session, e.g. when using radars for the KPI speed: all passing vehicles are measured. However, often a session sample will constitute only a part of all the relevant vehicles/road users that passed during the session. In this step the probability of inclusion is affected by, among other things, the traffic density during the session. The probability of inclusion of vehicles/road users is lower during sessions where traffic is dense (given the same measurement capacity). In order to take traffic density during the sampling into account, traffic counts of the relevant vehicles/road users are required during the sessions. The KPI guidelines indicate that, as a minimum, traffic density is based on a 10 minute count of all passing vehicles/road users during the measurement session (ideally separately for each relevant vehicle category). This count is then extrapolated to the duration of the session[3], e.g. a 1 hour session means that the 10 minute count result should be multiplied by 6.

To summarize, the inclusion probabilities in both sampling steps are :

1) Step 1 (stratum level: per road type or, if applicable, per combination of road type x time period): the probability that a particular type of location and moment in time is chosen
2) Step 2 (session level): the number of sampled vehicles/road users in the session divided by the total number of vehicles/road users that passed during the session (count extrapolated to session duration).

In order to get representative results, the measurements (observations) should be weighted inversely to the inclusion probabilities for each observation.

2.2 Weight formula if no traffic volume information is available

The proposed weight formula includes data which is generally available for Member States or can be estimated:

- an estimate of the actual proportion of each road type based on length on the total of all included road types in the national road network,
- if applicable (when more than one time period is required for a KPI), the proportion of each time period based on length (duration) on the total of all included week periods,
- data collected for the KPI measurements:
    - measurement duration per stratum (total of all session durations per stratum)
    - number of observed/tested vehicles/road users per session
    - traffic count per session (extrapolated to session duration),
    - measurement duration per session (if not all the same duration, the session observations should be weighted inversely to the duration of the session)

The tables below describe possible approaches to weighting taking the minimum requirements for specific KPIs into account. These can be used in case no more optimal weighting can be done.

2.2.1 KPIs with only one time period

We start with the simplest case of the KPI distraction, with as minimum required stratifications 3 road types and only 1 time period (daylight). The table below includes the parameters that determine the selection probability in each step of the sampling procedure and a formula to calculate the inclusion probability in each stage.

[3] As the KPI measures require specific vehicle/road user types, the theoretical ideal is to count only the relevant type, and if several types are included, to count them separately. If this is not feasible, it is allowed to count all passing vehicles on the same lane(s) and in the same direction(s) as the measurement.

Table 1. Parameters determining selection probabilities and formula to calculate selection probability for each sampling step for KPIs where only 1 stratum for time period is used, (e.g. KPI distraction)

Sampling step	**Step 1: selecting observation locations (e.g. 3 road types x 1 time period)**	**Step 2: Measured vehicles/road users in the session**
Description of inclusion process	Length of the road type in the total included national road network: the probability of inclusion is *lower* for locations on road types that constitute a larger part of the entire road network (As only one time period is considered, the inclusion probability is identical for the whole sample – a constant should not be considered in the formula)	Traffic density: the probability of inclusion of vehicles/road users is smaller during measurements with a higher traffic density
Parameters determining inclusion probability	M = relative size of road stratum = proportion of the road type in the total included road network m = number of measurement sessions in the road stratum	N = density of the traffic (traffic count result extrapolated to session duration: total number of vehicles/road users that passed during the session) n = number of vehicles/road users measured t = duration of the measurement session (e.g. in minutes)
Inverse sampling probability in sampling stage	M/m (One value per minimum road stratum)	N/(n*t) (One value per session)
Combined total sampling weight	Weight value = **[M/m] * [N/(n*t)]** (One weight value for each session, applied to all observations of the session)	

2.2.2 KPIs with several time period strata

In case there are several minimum strata for both road type and time period, then the formula should also include the inverse of the inclusion probability of the different time periods (e.g. KPI alcohol, seatbelt, helmet). The table below includes the parameters that determine the selection probability in each step of the sampling procedure and a formula to calculate the inclusion probability in each stage for this case.

Table 2. Parameters determining selection probabilities and formula to calculate selection probability for each sampling step for KPIs where several time periods are used (e.g. KPI alcohol, seatbelt, helmet)

Sampling step	**Step 1: selecting observation locations and time periods (e.g. 3 road types x 4 time periods)**	**Step 2: Measured vehicles/road users in the session**
Description of inclusion process	Length of the road type in the total included national road network: the probability of inclusion is *lower* for locations on road types that	Traffic density: the probability of inclusion of vehicles/road users is smaller during measurements with a higher traffic density

	constitute a larger part of the entire road network Length of the time period in the total included week span: the probability of inclusion is *lower* for measurements taking place during larger week periods (e.g. week days = 5 x 8h = 40h) vs. weekend days = 2 x 8h = 16h)	
Parameters determining inclusion probability	M = relative size of the combined stratum = proportion of the road type in the total included road network ***** proportion of the time period in the total included week span m = number of sessions in the combined stratum	N = density of the traffic (traffic count result extrapolated to session duration: total number of vehicles/road users***** that passed during the measurement session) n = number of vehicles/road users measured t = duration of the measurement session (e.g. in minutes)
Inverse selection probability per stage	M/m (one value per combination of strata)	N/(n*t) (one value per session)
Combined total sampling weight	Weight value: **[M/m] * [N/(n*t)]** (One weight value for each session, applied to all observations in the session)	

Specifications for possible special cases:

- M: If no official data on the length or proportion of the different road types is available, it is recommended to request estimates from experts from the relevant road administrations.
- N: In case all passing vehicles are included in the measurement N/n becomes 1 and irrelevant (e.g. KPI speed using radars).
- N: If not all passing vehicles/road users can be counted during the session, the minimum requirement is to count the passing traffic during a 10 minutes' count. The count result should then be extrapolated to the entire session duration (e.g. 1h session: count result multiplied by 6: 10 min * 6 = 60 min).
- N: If the minimum requirements consider different vehicle/road user types, the n and ideally also the N (traffic count) should be collected per relevant type, which allows calculating different weights per vehicle/road user type in each session. If traffic counts per vehicle/road user type are not feasible, traffic counts irrespective of type can be used as a proxy. As a minimum, all passing traffic on the observation lane(s) and direction(s) is counted for 10 minutes.
- **t:** duration of the measurement session. If all observation sessions last equally long, this is a constant and can be skipped. t is included since longer observation sessions increase the probability of inclusion of a passing vehicle. Put otherwise: doubling the amount of observation time at a specific location will double the number of vehicles observed and generally also the number of vehicles included in the traffic count N. Therefore, this factor is taken into account in calculating selection probabilities. This contrasts with for

instance doubling the number of observers (imagine police officers conducting alcohol checks), which will only lead to an increased n, but no change in N).

2.2.2.1 Example for calculating M

Fictious example: Calculation of the M values for 3 road type x 4 time period crossed strata

- Relative length of the road type:

	Km	Proportion
Rural roads	8,448	8,448/15,697=0.538
Urban roads	6,918	6,918/15,697=0.441
Motorways	331	331/1,5697=0.021
Total country	15,697	

- Relative length of the time period (e.g. 4 strata used for the KPI Alcohol)

	Week Day	**Week Night**	**Weekend Day**	**Weekend Night**
	Monday 6-22h	Monday 22-6h	Saturday 6-22h	Friday 22-6h
	Tuesday 6-22h	Tuesday 22-6h	Sunday 6-22h	Saturday 22-6h
	Wednesday 6-22h	Wednesday 22-6h		Sunday 22-6h
	Thursday 6-22h	Thursday 22-6h		
	Friday 6-22h			
Total	5*16h = 80h	4*8h= 32h	2*16h= 32h	3*8h=24h
Total week	**168h**			
Proportion week	80/168 = 0.476	32/168 = 0.190	32/168 = 0.190	24/168 = 0.143

- Values of M

	Rural roads	Urban roads	Motorways
Week Day	0.476*0.538 =0.256088	0.476*0.441 =0.209916	0.476*0.021 =0.009996
Week Night	0.190*0.538 =0.100222	0.190*0.441 =0.08379	0.190*0.021 =0.00399
Weekend Day	0.190*0.538 =0.100222	0.190*0.441 =0.08379	0.190*0.021 =0.00399
Weekend Night	0.143*0.538 =0.076934	0.143*0.441 =0.063063	0.143*0.021 =0.003003

2.2.2.2 Example for calculating combined sampling weights

Combining specific values of M with information on the other variables gives the final weight any observation included in the study is given.

Imagine for instance:

- 10 observation sessions for rural roads on weekdays: results in .256 / 10 as M/m (or .0256)

- 1000 vehicles driving by in an observation session in which 10 vehicles are tested for DUI in a 60 minute period: N/(n*t) = 1000/(10*60) = 10/6

- final weight for every observation in the session: M/m * N/(n*t) = .0256 * 1.666 = 0.042681333

Obviously, the final weight of every observation is a relative number, reflecting the combined impact of variations in the parameters determining the selection probability (which in turn reflects the part of the population the observation represents).

2.3 Extension when traffic volume data are available

In the formula above, traffic volumes are entirely estimated based on observed traffic volumes during the observations. Pre-existing independent traffic volume data are not yet taken into account in the formula.

The use of pre-existing traffic volume data should be evaluated carefully when considering to use this as a basis for correcting the observed data. Correcting traffic counts based on external estimates of traffic volumes should only be considered if the traffic volume estimates are based on better methods than using the traffic counts made during the observations.

To accommodate for that, the weight formula should be multiplied by the ratio of the percentual share in the total traffic volume of the stratum divided by the sum of the traffic count results during the measurement sessions per stratum. This corrects for any possible disproportionality between a priori traffic volume information from other sources and the traffic volume per stratum as observed during the observation sessions.

For instance, in case 20% of the traffic counts N are made in a particular stratum, while the "real" traffic share in the total number of vehicle kilometres driven is 30% in that stratum, this is corrected by multiplying the weight by a factor of 30/20.

In case no a priori external traffic volume information is available, the correction part of the formula falls away. Or put otherwise, if traffic volume is estimated based on the counts during the observation session, [driven/ride kms in stratum/traffic count in stratum] becomes [traffic count in stratum/traffic count in stratum] = 1.

The entire formula can hence be written as:

[M/m] * [N/(n*t)] * [TVS/TCS]

with:

- TVS : traffic volume share: share of traffic volume of the stratum based on a priori data (e.g. a proportion of 0.3 of vehicle kilometres are driven on motorways during weekdays)
- TCS : traffic count share: share of traffic counts in the stratum in the total traffic count over all sessions

---