

Reference metadata on methodology and quality



CARE - Road accidents database¹

A Data description (metadata)

Describe the main characteristics of the data set in an easily understandable manner, referring to the main data and indicators disseminated. The users should understand this short description immediately and easily.

Data description

CARE is a Community database on road accidents (commonly referred to as “crashes”) resulting in death or injury (no statistics on damage-only accidents). The legal basis for CARE is the Council Decision on the creation of a Community database on road accidents (93/704/EC , Oj No L329 of 30.12.1993, pp. 63-65). The Decision obliges Member States to provide annual data on road accidents but does not specify the content. The variables to be collected and disseminated were subsequently discussed and agreed upon by Member States and the Commission. In this regard, a CARE expert group composed of national representatives meets 1-2 times per year to discuss issues related to the database, the variables as well as the collection, processing and dissemination of data.

The major difference between CARE and most other existing international databases is the high level of disaggregation, i.e. CARE comprises detailed data on accidents, individuals and vehicles involved in the accident itself, as collected by the Member States. This structure allows for maximum flexibility and potential with regard to analysing the information contained in the system and opens up a whole set of new possibilities in the field of accident analysis. The main purpose is to provide evidence to identify and quantify road safety problems throughout the EU roads, evaluate the efficiency of road safety measures, determine the relevance of EU actions and facilitate the exchange of experience in this field.

Member States officials have access to the disaggregated CARE data through a dedicated reporting environment which makes available variables (to carry out their own queries and analysis. It does not however allow the possibility to retrieve single data on accidents, vehicles or persons.

Interested parties can also request DG MOVE to make queries for research purposes. DG MOVE regularly publishes on its website a wide range of aggregated data tables and analysis, under the umbrella of the European Road Safety Observatory.

Describe the target statistical population (one or more) which the data set refers to, i.e. the population about which information is to be sought.

Statistical population

The events in the scope of the database are road accidents resulting in death or injury. A traffic accident is defined as an accident involving at least one vehicle and happens on a public road. The damage-only accidents are excluded from the CARE database. The database is expected to cover all the events in the scope – census like data collection.

Glossary for transport statistics (Eurostat/UNECE/ITF) - <https://ec.europa.eu/eurostat/documents/3859598/10013293/KS-GQ-19-004-EN-N.pdf/b89e58d3-72ca-49e0-a353-b4ea0dc8988f?t=1568383761000>

Reference period

Statistical variables refer to specific time periods, which can be a specific day or a specific period (e.g. a month, a fiscal year, a calendar year or several calendar years). When there is a mismatch between the target and the actual reference period, for instance when data are not available for the target reference period, the difference should also be highlighted.

The date of the accident is used to aggregate and attribute indicators to a reference year. Fatally injured persons are those killed immediately or dying within 30 days as a result of an injury accident, excluding suicides.

Frequency of dissemination

The frequency with which the data is disseminated should be mentioned (e.g. monthly, quarterly, yearly). The frequency can also be expressed by using the codes released in the harmonised code list available for the European Statistical System.

The data are collected on an annual basis. Countries should deliver data for t-1 by end of September in year t but some countries have not finalised their data by this date. In addition, data need to be processed by DG MOVE which can lead to revised datasets being sent by the countries. When data are finalised for a country, they are made available in the CARE production database. In practice, this process of finalising data for t-1 runs from September in year t to spring in year t+1. Data for year t-1 are disseminated at EU level in April of year t+1 together with aggregate, preliminary data (number of fatalities per million population) for year t. The latter are gathered through national websites or from figures reported by countries

Time series for the main indicators start in 1999 while some historical series start in 1991 (limited set of variables).

Geographical

At European level: The geographical area covered by the data set disseminated (e.g. EU Members states, EU regions, USA, Japan, etc. as well as aggregates such as EU-27, EEA). At national level: the country, the regions and aggregates

¹ This document is based on a template provided by Eurostat and based on the [SIMS standard](#) for reporting on reference metadata and quality of statistical products. The content is the sole responsibility of the publisher.

reference area	covered by the data set disseminated Countries of the EU-27 and EFTA and computation of EU aggregates.
Unit of measure	The units of measures used for the data set disseminated should be listed (units of measures are e.g. Euro, %, number of persons). Also the exact use of magnitude (e.g. thousand, million) should be added. Counting (and corresponding %) of accidents/victims (fatal or injured), by characteristics of the accident/victims. Counting per million inhabitants is also used in reporting.
Basic statistical concepts and definitions	Describe in short the main statistical variables provided. The definitions and types of variables provided should be listed. The data model used to register all the events in the scope of the database is based on four type of variables: Accident related variables, Road related variables, Traffic Units (vehicle and pedestrian), Person related variables. Data in the CARE database are collected according to the principles of a Common Accident Data Set (CADaS). Information on variables can be obtained through the CADaS reference guide, available through http://ec.europa.eu/transport/road_safety/sites/roadsafety/files/cadas_glossary_v_3_7.pdf The schema on page 10 summarize the available 4 types of variables and how they are linked. Countries are invited to provide their yearly datasets in a format referring to the CADAS model but they are not obliged to do so. When a dataset refers to a national format (other than CADAS) a conversion to CADAS is done,
Classifications used	List all classifications which are used for the data set produced (with their detailed names). Several variables include two distinct types of values, referring to different level of detail: 1. Detailed values: concern information at the highest level of detail. 2. Alternative values: concern information at a more aggregate level of detail, when more detailed values are not available. Alternative values do not differ from detailed values apart from their level of detail. These values are complementary and can be used when more detailed data are not available (for example concerning the "Traffic Unit type" variable, if a country does not collect the values "car" and "taxi" separately, it can provide this information through the "car or taxi" alternative value). In principle, the classifications used follow closely those used in the "Glossary for transport statistics (Eurostat/UNECE/ITF)" - https://ec.europa.eu/eurostat/documents/3859598/10013293/KS-GQ-19-004-EN-N.pdf/b89e58d3-72ca-49e0-a353-b4ea0dc8988f?t=1568383761000 For example, the typology of vehicle types is based on the categories identified under "Transport equipment (vehicles)" while the typology of road type (motorways/urban/rural) is based on "Category of road"
Statistical Confidentiality	Legislative measures or other formal procedures which prevent unauthorised disclosure of data that identify a person or economic entity either directly or indirectly. The data transmitted by countries to the CARE database do not contain personal or identity data. The microdata are accessible only to national officials working in the field of road safety, e.g., Statistical offices, Ministries of Transport or Interior, Road Safety Agencies, police departments. The querying interface (data mining service via SAP Business object BI4) allows users to consult a series of pre-defined tables or perform their own queries.

B Data quality

Relevance	Describe the degree to which statistical information meets current and potential needs of the users. Statistics on road accident fatalities in individual countries are generally available at aggregated level only. If more detailed data are made available at national level, they often cannot be compared with those of other countries due to methodological differences. Data for the CARE database are increasingly sent according to recommendations for a Common Accident Data Set (CADaS). This consists of a minimum set of standardised data elements, which allows for comparable road accident data to be available for the EU and EFTA countries. The CADaS can be implemented on a voluntary basis at the national accident collection systems and be gradually adopted by the EU countries. Thus, progressively, more and more common road accident data from the various countries will be available in a uniform format. In this way CARE, the European database with disaggregated data on road accidents, will gradually contain more and more compatible and comparable data, allowing for more reliable analyses and comparisons across the EU countries.
Timeliness	Indicate the length of time between data availability and the event or phenomenon they describe. The basic aggregated data are disseminated in April of year t+2 (preliminary data) and in November (final data). Disaggregated data in CARE are made available for each country as and when they are processed and validated by DG MOVE. See also "Frequency of dissemination", above.

Source data: Indicate if the data set is based on a survey or on administrative data sources. If sample surveys are used, some sample characteristics should also be given (e.g. gross and net sample size, type of sampling design, reporting domain etc.). If administrative registers are used, the description of registers should be given (source, year, primary purpose, potential deficiencies and solutions to address them, etc.)

Accuracy and reliability

Each Member State produces its own road accident statistics by following its own procedures according to national protocols and formats and by using road accident data collected by the police. In some countries, police data is complemented by information collected by hospitals (i.e. DK, NL, EL, SE, ES, SI) or by governmental organisations (BE, PT, HU). Where national concepts differ from European ones correction factors have been used to alleviate the problem. This applies mainly to the implementation of the 30-day period in countries where it differed. With a few exceptions (LV, MT, SI) all Member States systematically implement data quality cross-checking.

The majority of countries (22) follow the common CaDAS glossary and transmit figures to the CARE database directly. Some other countries transmit data according to their national data model and at EU level DG MOVE services apply the CAREPlus protocol to transform the data into the CaDAS model.

The recommendation for a Common Accident Data Set (CADaS) refers to the set of data to be voluntarily transmitted by each country to the EU, which should be derived from the national road accident data collection system. This means, that the EU countries will not be legally obliged to adopt the CADaS and can continue using their national systems, however, if they wish they can enhance them in order to be able to provide the CADaS data to the EU. In case the countries do not wish to adopt the CADaS they should continue transmitting national road accident data to the EU in the current format.

Based on Article 3(4) of the 1993 Council Decision on the creation of CARE, each Member State shall be responsible for the quality of the statistics it provides. However, DG MOVE carries out a series of quality checks when processing the annual data from each country (coherence with national publications and with previous years, etc).

Data collection and methods used: Describe the method used to gather data from respondents (e.g. postal survey, CAPI, on-line survey, etc.). Some additional information on questionnaire design and testing, interviewer training, methods used to monitor non-response etc. should be provided here.

Administrative data from multiple sources. Each country has its own national system. Data collection is the responsibility of each country's administration.

Publications: Regular or ad-hoc publications in which the data are made easily available to users.

Disaggregated CARE data and basic aggregated data on fatalities are published regularly on the DG MOVE website. The latest available data: https://ec.europa.eu/transport/road_safety/road-safety-facts-figures-1_en

The data are also used to carry out extensive analysis as part of the European Road Safety Observatory (https://ec.europa.eu/transport/road_safety/specialist/observatory_en) which contains a wide range of data and analysis: annual report, bi-annual monitoring report, a set of country profiles, thematic reports and Facts and Figures series.

Accessibility and clarity

Eurostat re-disseminate a limited number of indicators about the number of fatalities, as the definition of injuries is not entirely harmonised across the Member States. They are available in the [Road Transport Safety node \(tran_sf_road\)](#) of Eurostat database.

The DG MOVE database offers a data mining interface open to members of the Care expert group and to national officials upon request. The querying interface (data mining service via SAP Business object BI4) is presented in the following manual: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/care_-_bi_4_2_manual_20180404_v100.pdf

Quality documentation: Documentation on procedures applied for quality management and quality assessment.

The main characteristics are set out in the CaDAS glossary and in the following document: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/specialist/erso/pdf/methods/metadatabron_en.pdf

Geographical: Describe any problems of comparability between countries or regions. The reasons for the problems should be described and as well the order of magnitude of the effects of the main sources of errors.

Coherence and comparability

Please note that data referring to the French Départements d'Outre-Mer (overseas territories) are not available and hence excluded from the national total and the EU aggregates. Data for the Portuguese autonomous regions of Açores and Madeira are included as of 2018 for which there is therefore a break in the time series. They account for around 4% of total fatalities.

Data on fatalities are generally accepted as having a high level of comparability between countries. Data on injuries are less comparable between countries as they depend for example on national practices and customs (involvement of police in accidents, police recording of accidents with minor injuries and indeed on the severity of an injury). These data should however be reliable over time for any given country.

Over time: Provide information on the length of comparable time series, reference periods at which series breaks occur, the reasons for the breaks and treatments of them.

The EC recommended the use of the CADaS model for data provided after 2010. As of February 2021, 22 countries use the model.

Breaks in the time series, the use of coefficients, etc are included as metadata in the data mining service.
 Use of correction factors (when fatalities at 30 days were not available) are described in the referred CaDAS glossary.
 Completeness of data (with detailed breakdowns) is somewhat lacking for a number of countries for the early years of the CARE database. However, basic aggregate data are available since 1991.

C

Contact and update

	The name of the contact organisation for the data or metadata.
Contact organisation	DG MOVE Unit C2 - Road Safety Unit
	The name of the contact points for the data or metadata.
Contact name	As above
	E-mail address of the contact points for the data or metadata.
Contact email address	MOVE-CADAS@ec.europa.eu or for revisions, please contact ESTAT-EC-STATISTICAL-COORDINATION@ec.europa.eu
	The date on which a metadata element was inserted or modified in the database/on the website.
Metadata update	2/3/2021

Sources used to compile:

- inventory IOS fiche: <https://statisticstest.jrc.ec.eu.int/catalog/27/datasets/30>
- the « about » tab on the homepage : https://ec.europa.eu/transport/road_safety/specialist/statistics_en#
- the Eurostat ESMS fiche for the indicators extracted from CARE:
https://ec.europa.eu/eurostat/cache/metadata/en/tran_sf_road_esms.htm
- ESMS-IP: https://ec.europa.eu/eurostat/cache/metadata/en/sdg_11_40_esmsip2.htm
- the CaDAS manual: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/cadas_glossary_v_3_7.pdf
- note on quality:
https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/specialist/erso/pdf/methods/metadatabron_en.pdf
- all the CARE reports are available under the tab “CARE reports” of
https://ec.europa.eu/transport/road_safety/specialist/statistics_en#
- Glossary for transport statistics (Eurostat/UNECE/ITF) -
<https://ec.europa.eu/eurostat/documents/3859598/10013293/KS-GQ-19-004-EN-N.pdf/b89e58d3-72ca-49e0-a353-b4ea0dc8988f?t=1568383761000>
- The Council decided on 30 November 1993 the creation of a Community database on road accidents (Council Decision 93/704/EC , Oj No L329 of 30.12.1993, pp. 63-65).