



# Economic and Social Council

Distr.: General  
31 March 2021

Original: English

---

## Economic Commission for Europe

### Inland Transport Committee

#### Working Party on Transport Statistics

##### Seventy-second session

Geneva, 9–11 June 2021

Item 6 of the provisional agenda

##### **Data collection, methodological development and harmonization of transport statistics**

## **An Update on the Eurostat/UNECE/ITF Joint Questionnaire**

### **Note by the secretariat**

#### *Summary*

This document describes the changes expected this year in how transport statistics are collected through the joint questionnaire of UNECE, Eurostat and the International Transport Forum.

## **I. Background**

1. Eurostat, the International Transport Forum (ITF) and the Economic Commission for Europe (ECE) have decided to modernise the collection and management systems of the Common Questionnaire data. The current IT system will be replaced by a new one, hosted by Eurostat, that will no longer be web-based. This system will be more user-friendly, and will take advantage of well-established tools and successful processes used by Eurostat for several other data collections. The new Common Questionnaire form for data collection will be an Excel file. Data will be stored in an Oracle database and its management and analysis will be performed by the software SAS (Statistical Analysis System). All members of the three organisations will transmit data to Eurostat, in a system that will give them easy access to all their files, allow revisions, create overview files and distribute data to the ITF and UNECE.

2. The design of the new system is ongoing and Eurostat, the ITF and UNECE are optimistic that the new tool will be used for the 2020 data collection, which will be launched in the summer of 2021. The components of new Common Questionnaire and the process of transmitting and managing files are described below.

## II. New data collection form for the Common Questionnaire

3. In the new Common Questionnaire, data will be transmitted using the Excel template shown in Figure 1. For each annual collection, one Excel file will be created per country and per domain – Road, Rail, Inland waterways, Oil pipelines, Buses and Coaches, and Road Transport (vehicle-kilometres). As shown in Figure 1, the year selection and the country filed will be fixed, whereas a drop-down list will enable users to choose the language in which they wish to complete the form (English, French, German or Russian). Figure below presents the form for the 2020 data collection, for Austria and for Railway transport, Austria has chosen to complete it in English.

Collection of data for the Common Questionnaire											
2020	Reference year				Drop-down list	Free text	Protected - cannot be changed				
AUT	Country				Can be empty	Can be empty					
ENG	Language				Can be empty						
RAILWAY TRANSPORT							Static Data				
DATABASE CODE	CHAPTER	LABEL	GLOSSARY	VALUE	FLAG	FOOTNOTE	2018	2017	2016	2015	2014
A-I-01-05-0-0-0-0	Infrastructure	Tracks/Length operated at 31.12 (km)/Total	<a href="#">Railway Transport Infrastructure</a>								
A-I-01-05-23.2-0-0	Infrastructure	Tracks/Length operated at 31.12 (km)/By type of traction/Electrified	<a href="#">Railway Transport Infrastructure</a>								
A-I-01-05-23.1-0-0	Infrastructure	Tracks/Length operated at 31.12 (km)/By type of traction/Non electrified	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-0-0-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/Total	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-11.1-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/By nature of Passenger only	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-11.2-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/By nature of Freight only	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-11.3-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/By nature of Passenger and freight	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-12.1-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/By number of tracks/Single track	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-12.2-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/By number of tracks/Double track or more	<a href="#">Railway Transport Infrastructure</a>								
A-I-02-05-15.1-0-0	Infrastructure	Lines/Length operated at 31.12 (km)/By track gauge/Standard gauge	<a href="#">Railway Transport Infrastructure</a>								

4. Each Common Questionnaire indicator is included in one row of the Excel table. Column A provides its unique code (Database code); column B, the relevant chapter of the theme (Chapter); column C provides the name of the indicator (Label); and column D, a link to the related chapter in the Glossary. Columns E, F and G will be completed by the data providers with the value, flag and a footnote for this indicator. All three cells can be empty. Columns H-L provide a static view of the values for the indicator that the country has provided for the 5 previous years. Data providers will not be able to modify these values which are presented only for informative purposes. To summarise, each form enables one country to send data in any of the four languages available; for all indicators in one of the five themes; only for one reference year. This year can be the latest one or an earlier one (revisions).

5. Every year, Eurostat will create a new form per country and per theme for the collection of data for the previous year (reference year). The form will be saved in a folder in CIRCABC (the European Commission central repository). Data providers will have access to CIRCABC to retrieve their empty data collection form, complete it and transmit it to Eurostat.

## II. Data transmission, storage, processing and dissemination

6. Data providers will transmit their Common Questionnaire forms via the EDAMIS Web Portal (at <https://webgate.ec.europa.eu/edamis4/>). EDAMIS implements Eurostat's "Single entry point" policy: data for all statistical domains should arrive securely at a common reception area in Eurostat, so they can be automatically monitored, checked and delivered into the target production environment.

7. Data providers for the Common Questionnaire will have to register as users in EDAMIS to transmit their Excel files and have access to their files saved in EDAMIS. In case some data providers are interested in automating their data transmission, they will have

the possibility to send data in more automation-friendly formats like comma separated variable (csv) formats.

8. Upon transmission of a file, it will go through a validation process and EDAMIS will send the data provider a validation report. If the file contains errors, the data provider will have to send it again after making the required corrections. If the file does not contain errors, it will be saved permanently in the “Received files” folder in EDAMIS; the data provider will receive a notification once the file is stored there. The data provider will be able to search the “Received files” folder by country, year and sending date in order to retrieve their input files any time they wish. Files for each country will be visible only by the data providers designated by the country, Eurostat, the ITF and ECE.

9. If a data provider wishes to revise data for reference year 2020 or after, they will retrieve the relevant file from EDAMIS, update its values and re-send it via EDAMIS. To revise data before year 2020, in the first phase of the migration, Eurostat plans to create and save in EDAMIS files per country, per theme and per year for the past 5 years (2015–2019). In this way, data revisions for these years will be similar to the revision of 2020 data. If a data provider needs to revise data prior to 2015, a tailor-made extraction of the appropriate questionnaire will be created for them.

10. In conclusion, EDAMIS will serve as an entry portal, a validation area, and a repository for all correct incoming files and for questionnaires created for reference years before 2020. Data that have succeeded the initial validation process in EDAMIS will be stored in an Oracle database and processed by the production system GSAST (Generic SAS Tool) that covers the entire data processing workflow. Further quality checks will be performed in GSAST and, eventually, fully validated data will be used for generating the Eurostat dissemination tables (Eurobase tables).

### **III. Data overview files**

11. Historical data will be stored in Excel files named “Overview” files. These files will be “static” – their content will be only for information and users will not be able to update them or send them to Eurostat for data revisions. For each country, Eurostat will create an Overview file per theme (Road, Rail, etc.) with all data for all years since the beginning of the Common Questionnaire collection exercise. These Overview files will be saved in CIRCABC and each country will only be able to see their own files. Each time a country revises their data via EDAMIS, the relevant Overview file will be updated automatically, and a new version will be stored in CIRCABC.

### **IV. Data access for the ITF and ECE**

12. The ITF and UNECE will need to receive the raw data sent by their member countries in order to store it in their own databases and disseminate it via their own channels. Eurostat will automatically provide the ITF and ECE with all relevant raw data via extractions from GSAST in agreed time intervals, e.g. three times per week during the data collection period. Extraction files will include data per country, per theme and per year in csv format. They will be stored in CIRCABC, the ITF and ECE will be able to download them from there.