

---

# ***Annex 8 - Evolution of the database for the REGIO Rail Unit Cost Benchmarking Tool***

The benchmark analysis performed by the REGIO Rail Unit Cost Benchmarking Tool (RRUCBT) is based on the comparison of the investment cost of a new project with the cost of similar rail infrastructure projects carried out in the EU. Therefore, it was then necessary to gather and organise detailed information on a number of rail infrastructure projects to consider in the analysis.

In line with the ToR and in the view of building a significant dataset, all the rail infrastructure projects supported by EU funds over the 2000 – 2015 period have been considered to build the database. The sources of information on projects were the databases provided by DG REGIO and INEA. 867 projects were inserted and classified. Nonetheless, some of the projects included in the starting database were not relevant to the study because they were either out of scope or lack a minimum set of information to be analysed. Therefore, an accurate data cleaning process has been undertaken on the original database, which is outlined in details in the following paragraphs.

## **Cleaning the database from out of scope observations**

Out-of-scope projects were removed in two different phases:

- firstly, projects not related to the deployment of a rail infrastructure were excluded from the database;
- secondly, the remaining projects have been accurately assessed to determine which type of works they referred to and projects not concerning works on railway lines were excluded from the analysis.

Rail projects supported by EU but not related to the infrastructure include:

- projects related to studies (the selection did not take into consideration projects with both works and studies, in case the exact cost of the latter could be identified);
- projects related to rolling stock and/or on-board equipment.

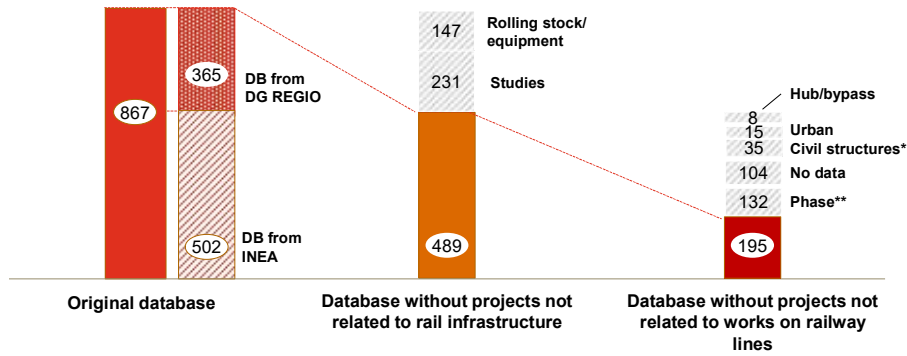
This first elaboration of the database reduced the number of projects to 489. Following, projects were clustered based on the type of work they concern. This way it was possible to identify projects not related to the deployment of railway line e.g. projects related to the upgrade of a station.

Specifically, five categories of works were not included in the analysis of the unit cost ranges:

- **Civil structures:** projects related to the building of a specific tunnel, bridge or viaduct. This category includes projects referring exclusively to the construction of a specific major structure or projects which the majority of the construction cost (more than 50%) refer to civil structures;
- **Stations:** projects concerning exclusively works for the construction, reconstruction or rehabilitation of specific railway stations and projects which majority construction cost (more than 50%) refer to stations;
- **Urban:** related to the deployment of an inner-urban rail infrastructure, this category includes railways lines linking two stations within the same metropolitan area;
- **Hub/bypass:** construction works performed to upgrade, reconstruct or bypass a transport hub;
- **Phase:** projects relating to individual sections of larger projects.

Further, additional 104 projects were removed as it was not possible to gather sufficient information (i.e. total investment cost and total length).

**Figure 1 – Phases of the data cleaning**



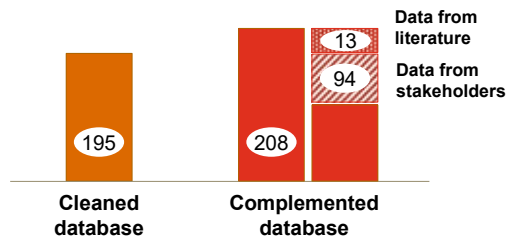
\*Projects concerning the construction of tunnels have been used to perform the analysis presented in the case study.

\*\* Whenever possible, the information on each phase of a project has been aggregated to include in the database the single larger project, which can be considered in the analysis.

**Complementing the information in the database**

To refine the quality of the database, an extensive desk research has been carried out to include in the analysis data retrieved from the relevant literature and from PwC databases as well as to complement the detailed information on the projects included in the database. Further, an EU-wide stakeholder consultation was performed to collect information directly from the IM.

**Figure 2 – Results of the data complementing**



**Assessment of the database**

The whole process enabled to create a large dataset, including over 200 rail infrastructure investments from all over Europe. A high level of detail on the information gathered was achieved as well, with almost 100 projects presenting complete, detailed information (i.e. information up to Tier 3).