EBA 2017 EU-wide transparency exercise dataset

Data user guide

For the 2017 EU-wide transparency exercise, the EBA published bank-by-bank data contained in 10 transparency templates (more than 4 000 data points). This exercise provides detailed data for 132 banks from 25 countries of the European Union (EU) and the European Economic Area (EEA). Data is also disclosed for the bucket 'All other banks', which includes aggregated values for the banks that are in the Risk Assessment Report sample but not in the transparency exercise, to enable reconciliation of the figures for the EU.

The EBA has developed a set of practical tools intended to clarify data use for the 2017 EU-wide transparency data. These include interactive maps, Excel aggregation tools and a complete dataset in CSV format, which can be imported into any analytical software for analysis purposes.

The transparency exercise dataset is stored in four CSV files. They include all the bank-by-bank data contained in the transparency templates, grouped into specific data categories to reflect the content of one or more transparency templates, as shown in the table below:

CSV file name	Transparency template(s)
Credit risk	Credit Risk_STA, Credit_Risk_IRB, NPE, Forborne Exposure
Market risk	Market Risk
Sovereign exposures	Sovereign
Other templates	Capital, Leverage, Risk Exposure Amount, P&L

With the CSV files, users will find the data dictionary table and the metadata table, which are helpful for understanding the file's database structure (as the four databases have different structures), and for setting up queries for data extraction and management.
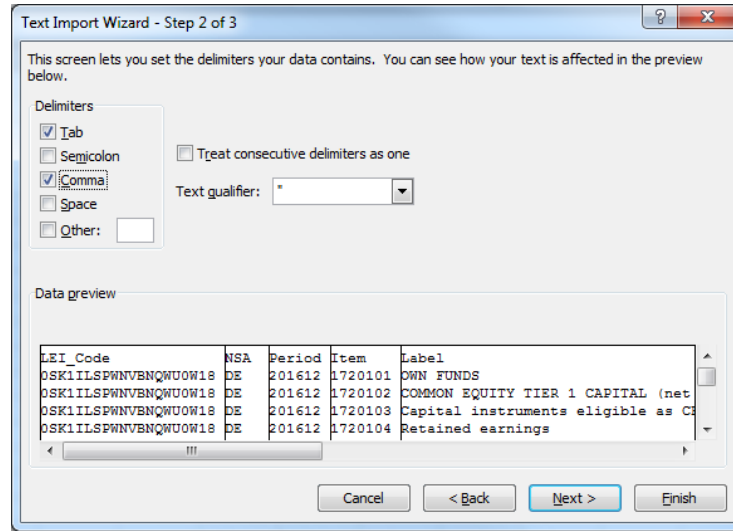
The example below shows how to use and query the EU-wide transparency exercise database.[1] The files are converted into spreadsheets, allowing the use of standard analytical tools embedded in Excel.

Capital: CET1 Ratio — fully loaded — for each bank by period using a pivot table
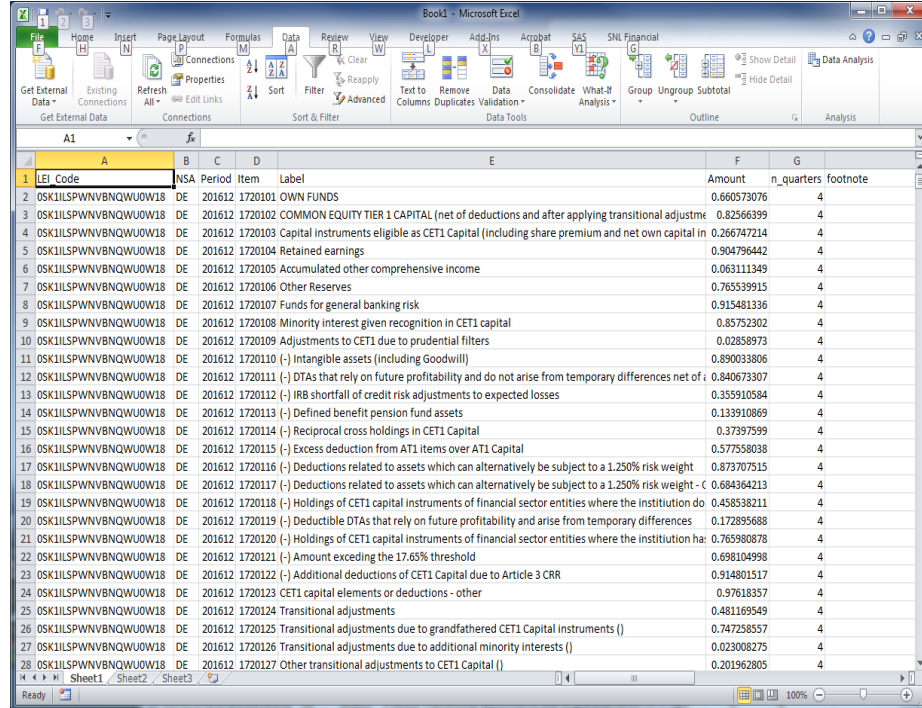
i)   Once you have downloaded the CSV file containing data on *Capital* (tr_oth.csv), import it into Excel using the Text Import Wizard:

---
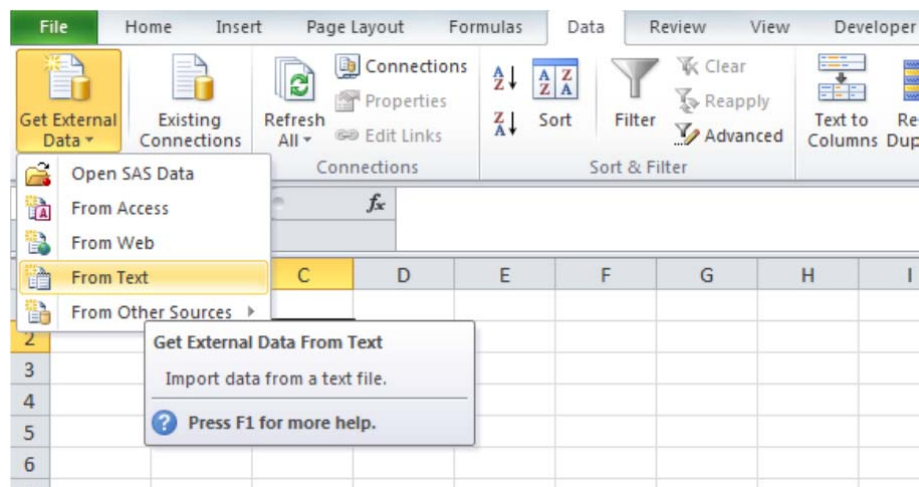
[1] Please bear in mind that this is only an example and that, therefore, **the figures show mock data**.

ii)   The database structure will appear as shown below:

iii) The database structure is explained in a metadata file, in which you will find a description of all the values that each column can assume. The dataset tr_oth has the following columns:
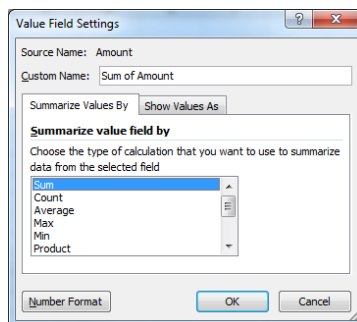
- *Lei_code*: a bank identifier;
- *NSA*: ISO code of the bank's country;
- *Period*: time period (201612 for December 2016 and 201706 for June 2017);
- *Item*: code of each variable;
- *Label*: decodification of the item;
- *Amount*: value that the variable assumes;
- *N_quarters*: the number of quarters to which P&L data (flow data) refer;
- *Footnote*: specific bank clarification as disclosed in the bank's PDF, added to all the items of the relevant templates.

Users can find decoding information either in the metadata file (TR_Metadata.xlsx) and/or in the data dictionary file (TR Data dictionary.xlsx).
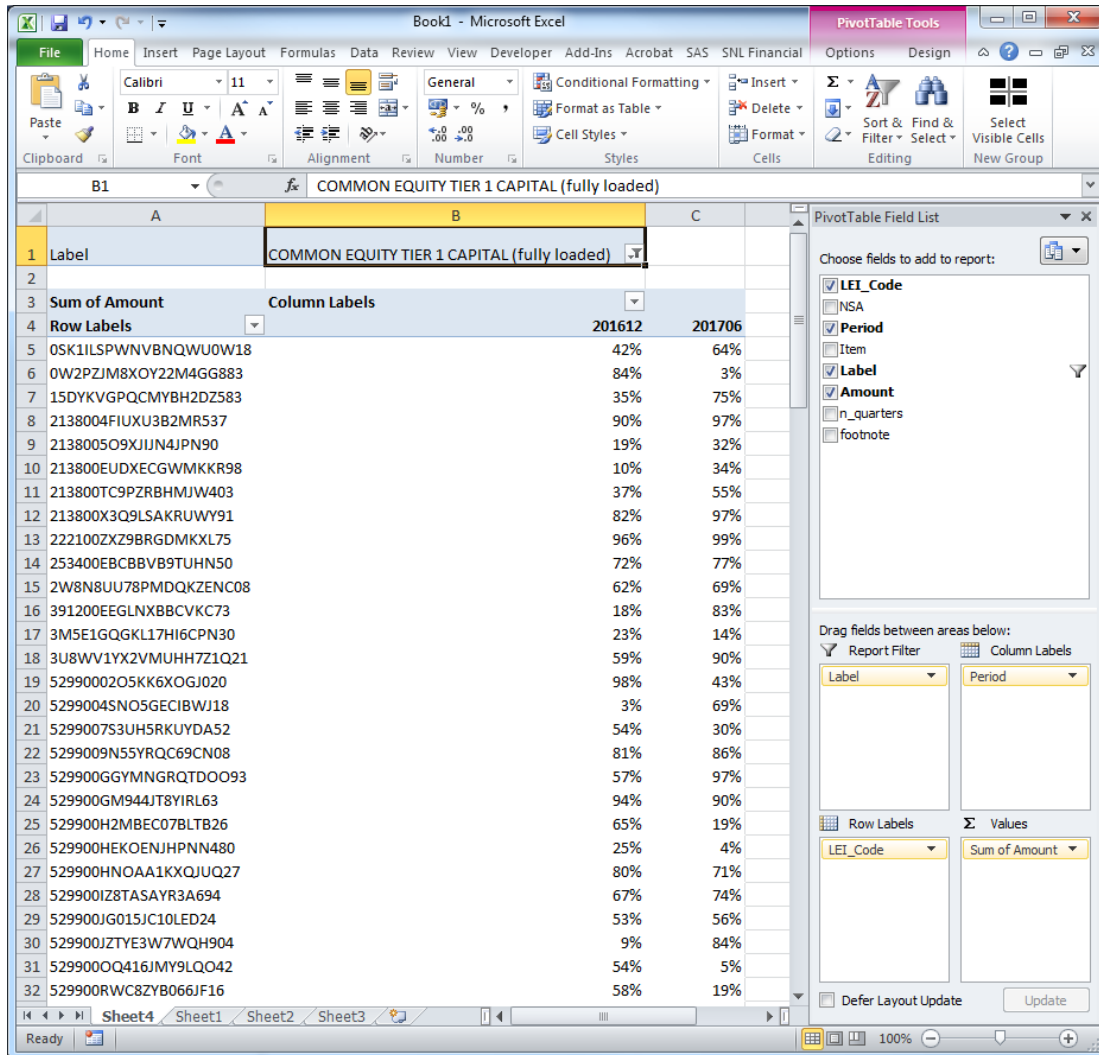
For instance, in the sheet 'Perf_status' of the metadata file, you can see that the dimension 'Perf_status' can assume values only from 0 to 4, and find the relevant explanation for this.

Perf_Status	Label
0	No breakdown by Perf_status
1	Performing
2	Non Performing
3	Performing but past due >30 days and <=90 days
4	Non Performing and Defaulted

iv) Now click on 'Pivot table' and select the entire dataset (or a subsample if you have already filtered the data you need) as the pivot table range. Set up the pivot table structure, dragging the variable 'LEI_code' into the box 'Row Labels' and the variable 'Period' into the box 'Column Labels'. Drag 'Label' into the box 'Report Filter' to select the item 'Common Equity Tier 1 Capital Ratio (fully loaded)' and show only the information for this item. Finally, you may drag in the box *Values* the variable *Amount*, where the variables' values are stored, and aggregate it by the sum.

v)   The final result should be as shown below:

In the 'Banks' sheet of the metadata file, you can see the name of the bank that the LEI code refers to, along with other properties of the bank (country of origin, financial year end, etc.).