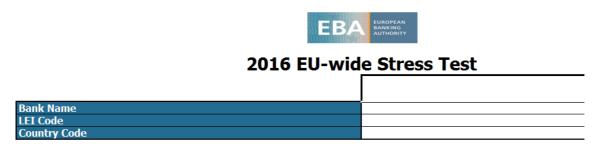


The EBA Stress Test data set

Guide for data exploitation

As a result of the 2016 EU-wide Stress Test Exercise, the EBA has published bank-by-bank data contained in 10 Transparency templates (up to 16,800 data points per bank for a sample of 51 banks).





The EBA has developed a range of practical tools that aim to facilitate the use of the stress test data. These include interactive maps and excel aggregation tools, as well as the complete stress test dataset in CSV format, which can be imported in any analytical software for analysis purposes.

The stress test dataset is stored in 3 different CSV files and includes all the bank-by-bank data contained in transparency templates. Each CSV file contains a specific stress test data category that reflects the content of one or more transparency templates as shown in the table below:

CSV Name	Stress Test category	Transparency Template
TRA_CR.csv	Credit risk	TRA_CR_STA TRA_CR_IRB TRA_CR_SEC TRA_NPE TRA_FORB
TRA_SOV.csv	Sovereign	TRA_MR_SOV
TRA_OTH.csv	Summary results, Capital, Risk exposure amount, P&L	TRA_SUM TRA_CAP TRA_P&L TRA_REA



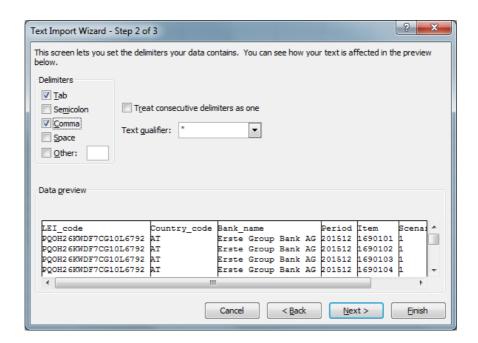
Along with the CSV, users will find the data dictionary table and the metadata table that are needed for understanding the database structure of each file (the tree databases have a different structure) as well as for setting up the queries to extract the data.

An example will be useful to understand how to use and query the EBA Stress test database (bear in mind that **the figures below show fake data**). In the example below, the files have been converted into excel files in order to use standard analytical tools embedded in excel.

<u>Capital</u>: <u>CET1 Ratio – fully loaded - for each bank by scenario using a pivot table</u>

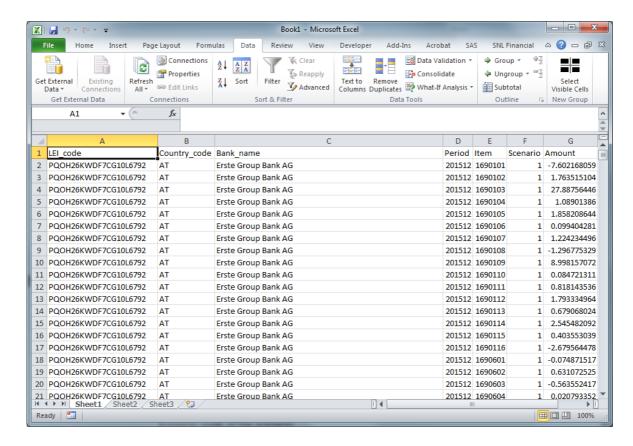
i) Once the CSV file containing data on *Capital* is downloaded (TRA_OTH.csv), we import it in excel using the text import wizard:

ii)



iii) The database structure turns to be the following:





- iv) The database structure is explained in a metadata file in which you one can find a description of all the values that each column can assume. For *Capital*, the database has 7 columns:
 - Lei: a bank identifier
 - Country_code: code of the country of the Bank
 - Bank Name: Name of the bank
 - Period: Time period
 - Item: Code of each variable
 - Scenario: code of the scenario
 - Amount: value that the variable assumes

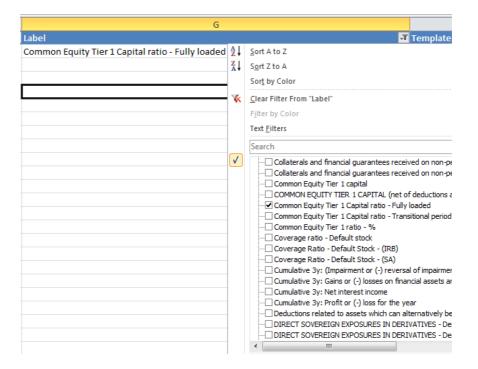
Users can find decoding information either in the metadata file (Metadata for the website.xlsx) and/or in the data dictionary file (TR Data dictionary for the website.xlsx).

For instance, in the sheet "Scenario" of the Metadata file, one can see that the dimension Scenario can only assume values equal to 0, 1, 2, 3 and find the corresponding explanation in it.



Code	Label
0	No breakdown by scenario
1	Actual figures
2	Baseline scenario
3	Adverse scenario

v) For identifying the item code associated with the financial concept "<u>CET1 Ratio – fully loaded"</u>, users can look for the name of the item in the column *Label* of the Data dictionary file and they will find that the item code is 1690850.



- vi) Now we click on "Pivot table" and select the entire dataset (or a subsample if you already filtered the data you need) as the pivot table range. We set up the pivot table structure, dragging in the box Row Label the variable Bank_name while in the columns we want the Period and the Scenario. Finally we drag in the box Values the variable Amount where the variables' values are stored and we aggregate them by sum.
- vii) Final result turns to be the following:



