subsidy stories.eu

Manual OpenSpending
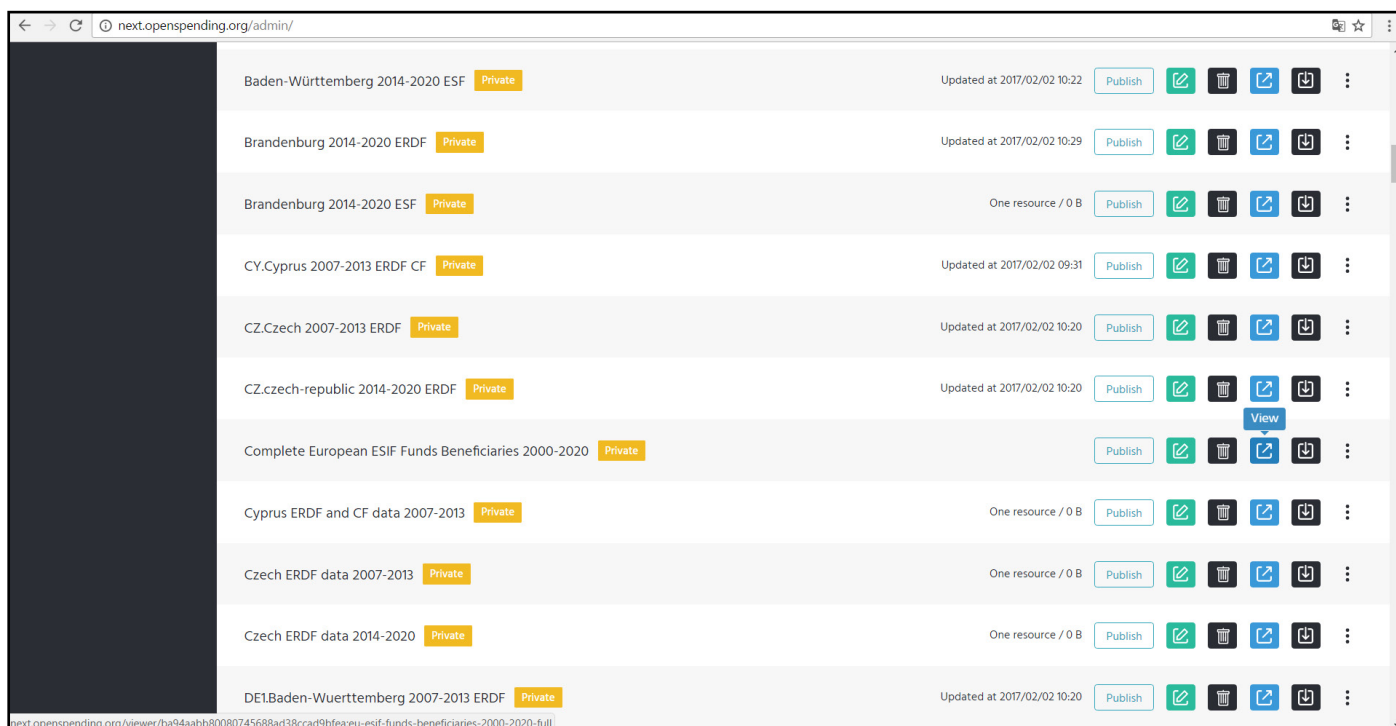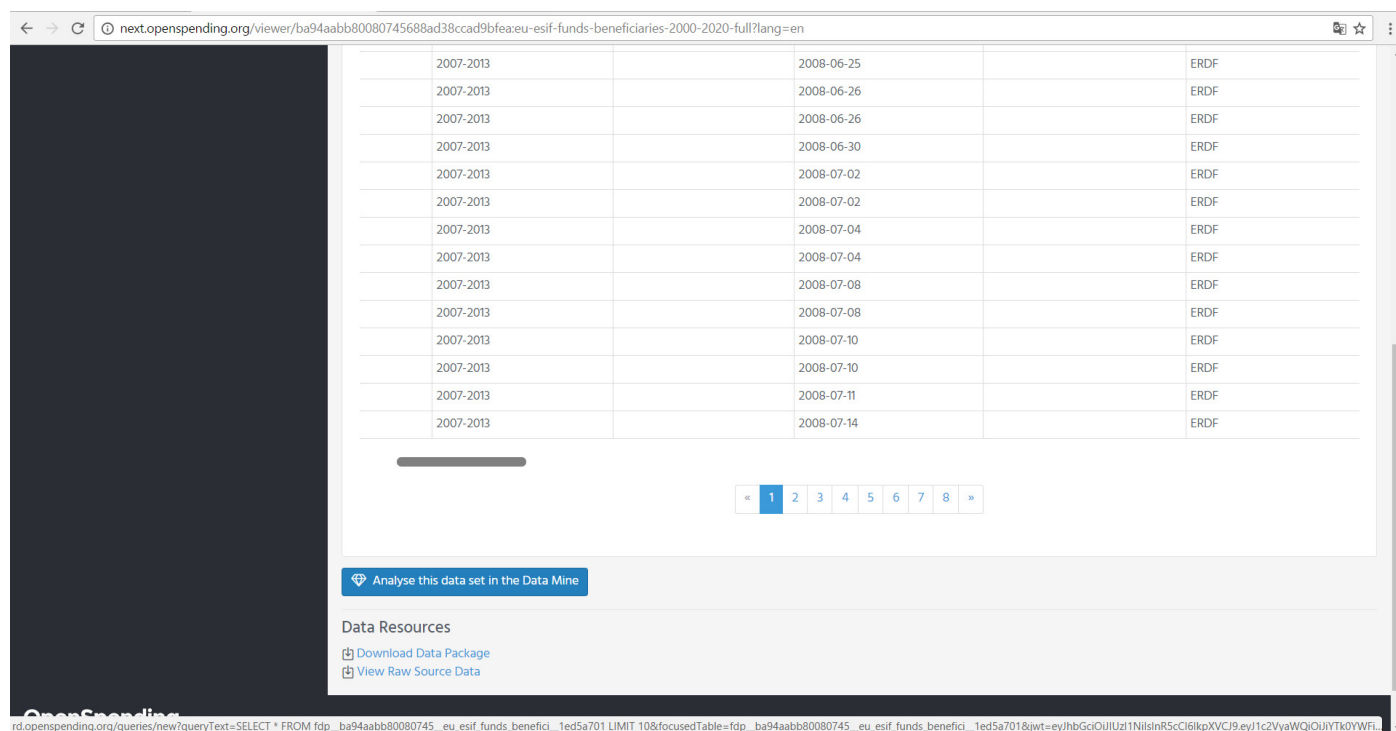
Datasets and how to use them

To access our datasets follow this link, which leads you to the OpenSpending viewer.

You are automatically logged in to our test account and have access to all our data. The joint dataset is called "Complete European ESIF Funds Beneficiaries 2000-2020" and includes all data. Furthermore, all single country/region datasets are accessible there too. To see an overview on what datasets are available go to the top right corner and click on OpenSpending Test (this is our test account) and then select "profile". This page offers an overview on the latest version of datasets available. From here on you have to decide how you wish to work with the data. You can view each dataset online by selecting "view" on the right side next to the dataset, which leads you to the next screen.
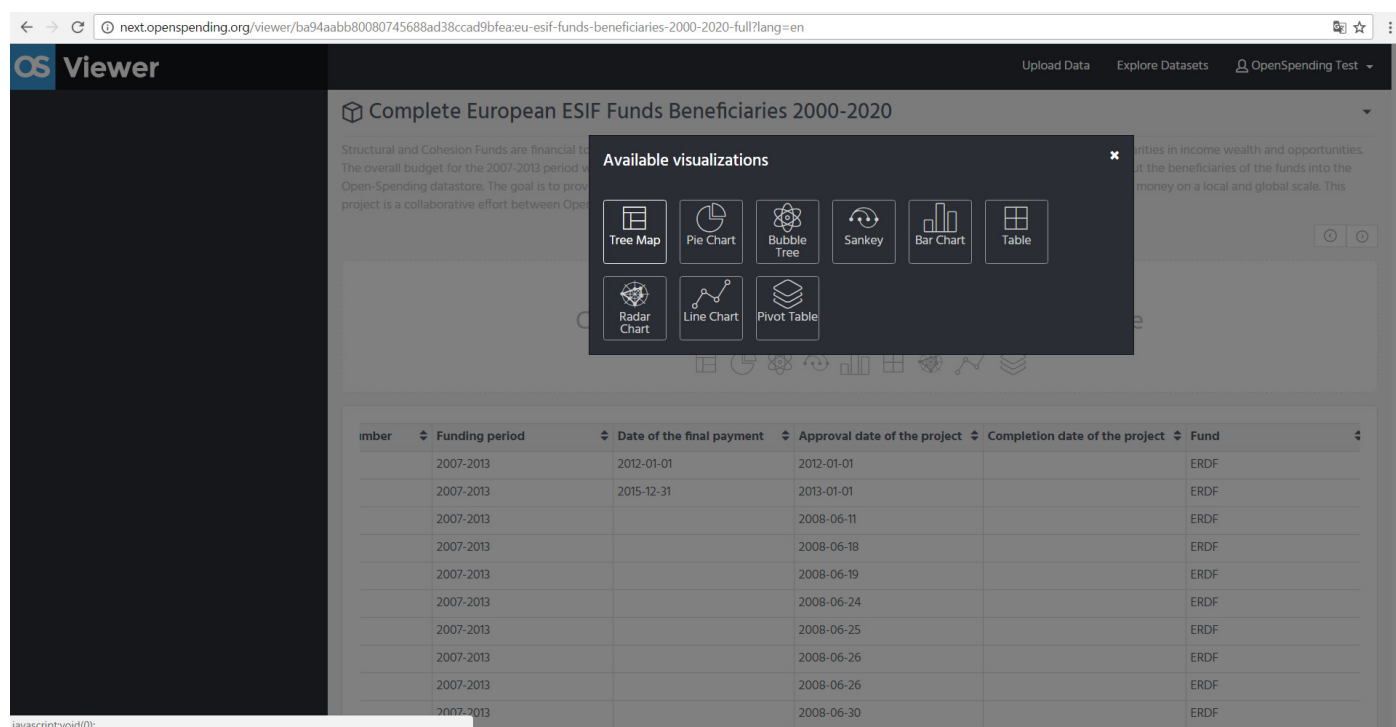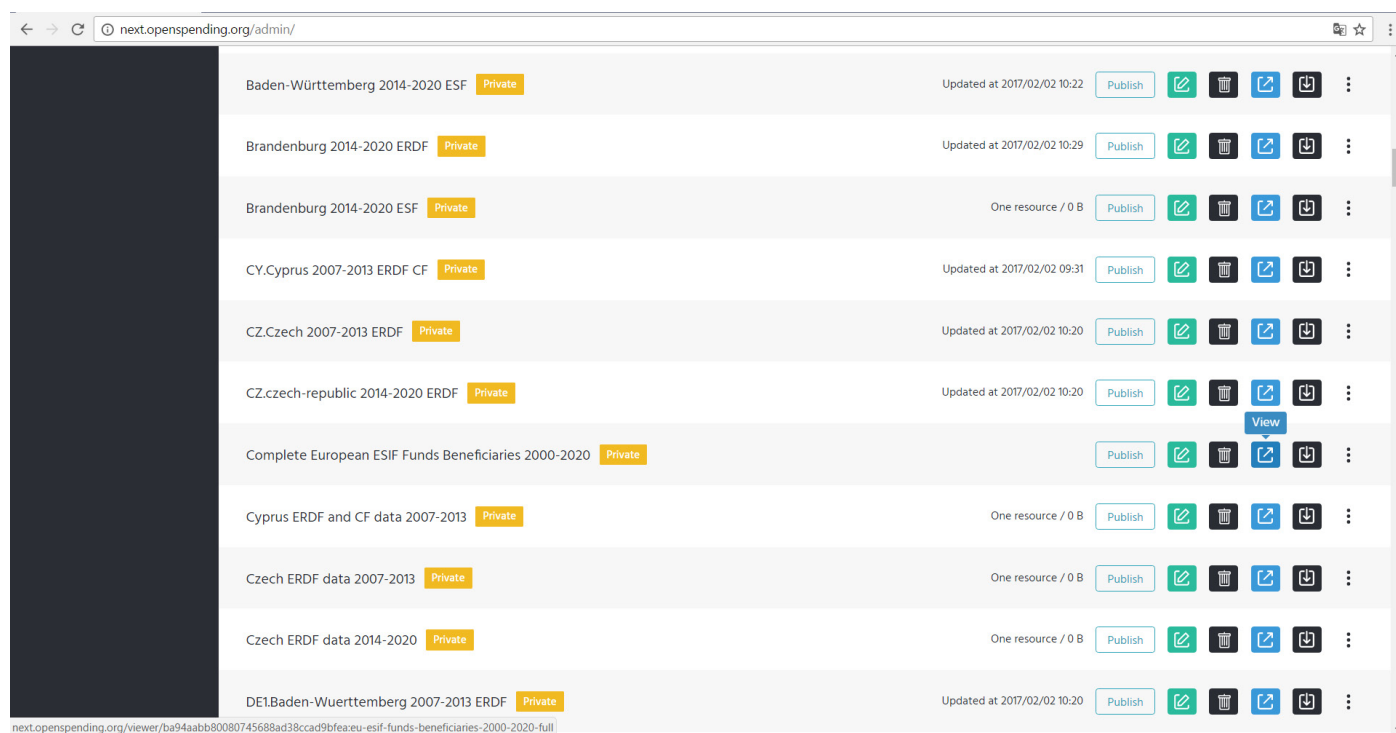
A new tab opens up showing the selected dataset - in the top you can choose to work with the visualizations OpenSpending has to offer such as Tree Maps, Bar Charts or Tables. If you scroll down (bottom left) and select to analyse the data in OpenSpending's data mine (a query tool based on SQL).

SQL queries are helpful to get first insights in large datasets e.g. basic statistics such as average amount per project (as depicted below). A good introductory video for using SQL can be found here.

However, for doing more in depth analysis downloading the dataset might be advisable. You can do this in the first overview by selecting "download" instead of "view". Choose "raw source data" and you will get a CSV containing all the data. You can open the data with Excel, OpenOffice, LibreOffice or Google Sheets and use the programs filtering and search options to navigate the data. Excel is known

to have some trouble with CSV files, but the other listed programs are free and enable you to navigate the data nicely.

Also consider Github was used for collecting all the data. By selecting "data" you will see an overview of all EU countries and can further select the country, region and time period you are looking for. In the "source.description.yaml" files there is often more information on where the data was originally found and especially how it was 'mapped'. Feel free to use any of this additional information especially if you are finding certain amounts questionable or think something was mapped incorrectly. If you cannot find any information on the source in the yaml-file - check the "Issues" section and just type in the name of the country you are looking for like "Slovenia" - just be aware that issues can be closed and open.