Exercise: Discovering Open Data

The idea of this exercise is to help you gain a better understanding and be able to critically appraise the different types of open data publication practice on the web. While publishing data may just be about meeting targets, publishing usable data that can be consumed by others in a reliable fashion is entirely different. The aim of this exercise is to take a critical view on the ability to effectively consume open data and identify a number of key areas for improvement.

Throughout this exercise you should think exclusively as a consumer of open data. When analysing each dataset think about the ease of using the data, availability and opportunities to answer complex questions and perform data analysis. Questions may include:

- Do I understand what the dataset is? Does the title and description match the data itself?
- Am I able to access the data?
- Am I permitted to use the data?
- Do I understand the data itself?
- What questions can I answer using the data? Is the data too granular, too generic?
- Is the dataset supported long term?
- Is the data consistent?
- Is the data clean?
- How much effort would be required to make the data usable?
- Can I get support on the data and find what else it has been used for?

Once you have analysed the dataset as a consumer, think about how you would change the publication practices of the datasets in order to improve them. Do the best practice guidelines help here? Would you recommend this style of data publication to others?

Exercise Datasets

Dataset 1: Spend over £25,000 in the Foreign and Commonwealth Office
Location: http://data.gov.uk/dataset/financial-transactions-data-fco

Dataset 2: Tourism Malaysia, Facts and Figures
Location: http://data.gov.my/dataset.php?page=1&p=105&q=