

In this course we do coding in Python. You will need to have a working Python interpreter on your machine. If you need to install a new one, we recommend using Anaconda (<https://www.anaconda.com/products/individual>). If you do not have one already, we recommend installing a Python IDE, such as Pycharm (<https://www.jetbrains.com/pycharm/>).

You will need to install the following python packages: nltk, trec tools

Unzip the gov.zip file inside this assignment directory. The gov folder contains 3 sub directories:

- documents, the collection of documents in the corpus.
- topics, the list of queries to be run on this corpus.
- qrels, ground truth labels of which documents are relevant to which queries.

You are now ready to begin assignment 1.