**Programming Assignment**

**1. Goal:** Make a basic information retrieval system with tf-idf weighting.

**2. Data Set**

- The cranfield data set is used (load the data set from the python package **ir-datasets**).

- You can get the answer set using the *get\_answer\_set(dataset, qid)* function.

**3. Detail**

- See the jupyter notebook file.

- Fill in “Your Code” parts.

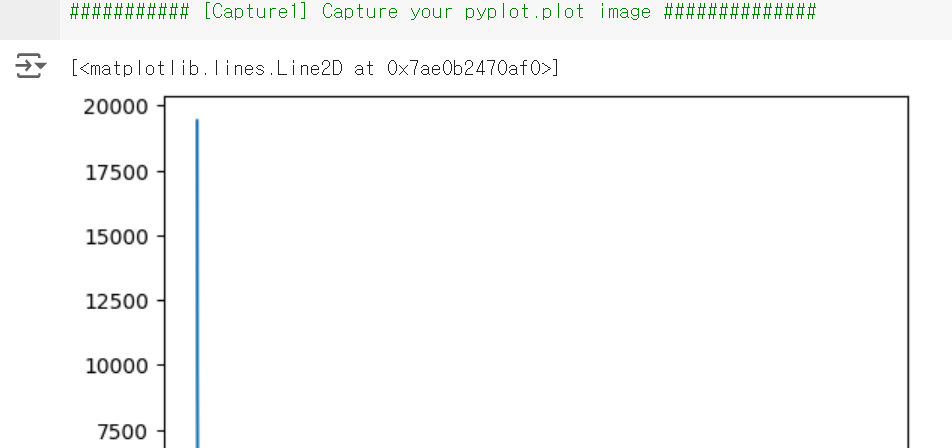
**4. Due date:** June 9. (Tentative)

**5. Submission Files**

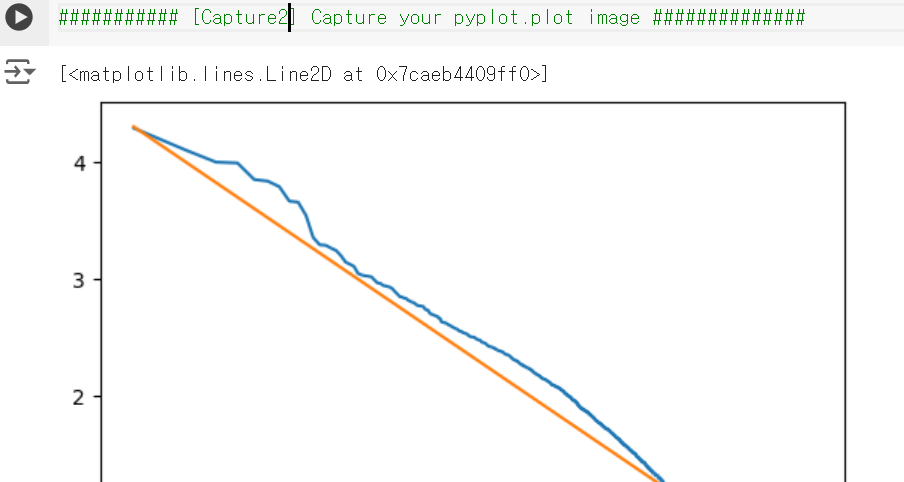
- your jupyter notebook file

- Captured Results

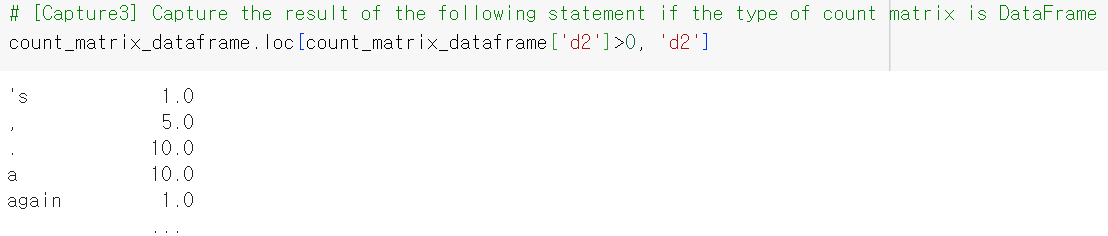
1) Zipf’s law graph

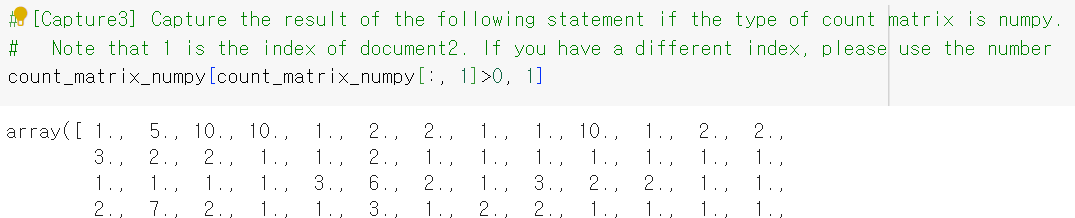


2) Zipf’s law graph with log scale



3) count matrix





4) Capture the results with mode=’nnn.nnn’ and mode=’lnc.ltn’ for query2.

(Sample Image for query 1)

