In this homework we will build on top of the work from Homework 2
1. Use Apache Beam and Cloud Data Flow to do some data processing on the files in your cloud bucket. In particular you should find and print out the top 5 files with the most incoming links, and top 5 files with the most outgoing links (hint: the second should be much easier)
2. Run your code both on your local laptop and on the cloud dataflow engine and report on the respective runtimes.

What to turn in:

- The python code for your app
- A pdf file describing all the necessary steps to configure and run your app, the output of your code and an explanation of what your code does..