Most importantly, add the entire drive folder to your drive using “Add to my drive”. This sets all the access and permissions required to run the files.

There are 3 important files with respect to running the code-

-buildgraph.ipynb

-movie\_metadata.csv

-queries.txt

movie\_metadata is a dataset created by IMdB movie site scraping. It contains tags such as actor\_1,actor\_2,actor\_3,genre, etc. This is in CSV format.

questions.txt is a text file that has a number of queries that we could pose to the graph written in natural language. We use NER to extract the supposed entities and relations and query the graph with the same.

buildgraph.ipynb is the runnable that builds our version of the knowledge graph from the movie\_metadata dataset. It leverages the below python packages-

-csv

-networkx

-matplotlib

-nltk

-pprint

-re

These must be installed before running the script.The following packages are downloaded as a part of the script. This is simple if running as ipynb.

-words

-averaged\_perceptron\_tagger

-punkt

-maxent\_ne\_chunker

The script then reads the queries file line by line and keeps track of how many of the queries have legitimate responses from our graph. The result is displayed.

To query the Google knowledge graph, we send requests to Google Knowledge Graph Search API. We use the Google Developers Console to create an API key at <https://developers.google.com/knowledge-graph/> and enable access to our application. Our key is provided in the script.