CS 1656 – Introduction to Data Science

Department of Computer Science – University of Pittsburgh
Instructor: Prof. Alexandros Labrinidis
Teaching Assistant: Phuong Pham – Additional Credits: Zuha Agha, Anatoli Shein

Linux Setup for Graph Databases: Neo4j & Cypher

Step 1: Install neo4j community edition 3.3.0 for your operating system from the link provided below.

https://neo4j.com/download/community-edition/

Step 2: Install official neo4j python driver which allows connecting to a neo4j graph database and run cypher queries from a python program. Type in the command below to install the driver.

```
pip3 install neo4j-driver
```

Other python community drivers are also available at the link below:

https://neo4j.com/developer/python/

Step 3: For this assignment, you will be using a graph database of movies found at the link below:

http://cs1656.org/data/cineasts_12k_movies_50k_actors.zip

Extract the zip file where ever you want, and copy the extracted directory **cineasts_12k_movies_50k_actors.db** to:

(For Linux)

/path/to/neo4j-community-3.3.0/data/databases/cineasts_12k_movies_50k_actors.db

(For Windows)

It can really go anywhere, but I put it next to the existing default.graphdb folder in:

C:\Users\aus\Documents\Neo4j\cineasts_12k_movies_50k_actors.db

Step 4: Edit the following configuration file:

/path/to/neo4j-community-3.3.0/conf/neo4j.conf

First we will switch the default database **graph.db** to our downloaded one **cineasts_12k_movies_50k_actors.db**. To do that, find the commented out the line:

#dbms.active_database=graph.db

And add the following line below it:

dbms.active_database=cineasts_12k_movies_50k_actors.db

Also, make sure that the following lines are **NOT** commented:

dbms.directories.import=import

dbms.security.auth_enabled=false

dbms.allow_upgrade=true

Step 4: Start the Neo4j server by running the following command in your terminal:

/path/to/neo4j-community-3.3.0/bin/neo4j console

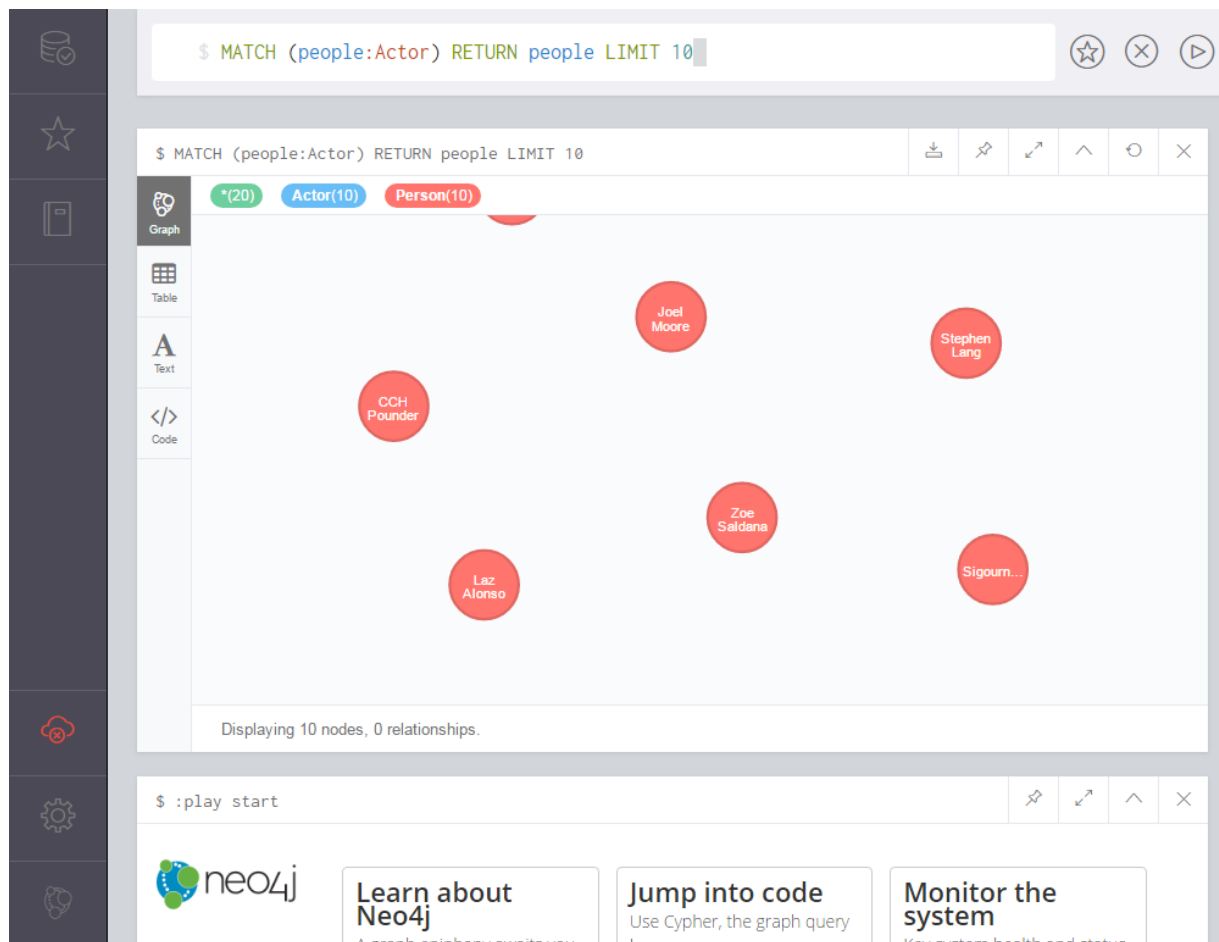Once the server is started, open your browser to:

http://localhost:7474/browser/

Now you can issue queries to the database from the input field at the top of this page.

For example try:

MATCH (people:Actor) RETURN people LIMIT 10

You should see the following screen if it was a success:

Step 5: You can also execute Cypher queries on this Neo4j movie database in python. The sample python program that uses neo4j-driver is provided in **recitation11.py**