Week 7: Neo4j tutorial

11.09.2018

Learning Objectives

In this week, we will go through the build-in tutorials of Neo4j Community edition to learn basic components of cypher query, such as patterns, clauses and sub-clauses.

Question 1: Neo4j and Cypher query language basics

Neo4j is installed locally on all lab PCs under the directory `C:\ProgramFiles\Neo4jCommunity`. To start the server, navigate to `C:\ProgramFiles\Neo4jCommunity\bin` and double click the executable file `neo4j-community`. A window will appear to allow you specify a database location. This is where neo4j stores its data. Create a new directory in your U drive, and type the full path of that directory in the neo4j startup window and click the `start` button at the bottom right corner. When the server finishes start up process, the Status box will display a message "Neo4j is ready. Browse to `http://localhost:7474/`". Click the link to open a browser and connect to the neo4j web interface. This is the preferred and primary user interface of Neo4j. It allows you to run Cypher commands to retrieve and update a graph; it also allows you to visualize the graph with customized styles. A detailed guide to the web interface can be found from `http://neo4j.com/developer/guide-neo4j-browser/`

When connecting to Neo4j, you will be prompted for `username` and `password`. If it is your first time connecting to Neo4j, you should use the default `username/password` pair "neo4j/neo4j". You will be asked to change the password after logging in. Please remember the password for later use.

At the very top of the browser window is a editor box for entering commands. All commands that have been executed along with their results will be displayed on the stream section in their own frames. You can easily discard any unwanted history command from the stream section. Figure 1 shows an example of browser window after connecting to the server and running one of the build in quick guides `:play concepts`. There are a few quick guides each has a link you can click and play with.

- **Intro** is a guided tour for the query browser; the command is `:play intro`
- **Concepts** contains basic concepts of graph database; the command is `:play concepts`
- **Cypher** introduces the Cypher query language; the command is `:play cypher`

- **The Movie Graph** is a comprehensive tutorial with commands to set up a mini graph model of movies and actors; it also contains several query use-cases, including simple query, aggregation and shortest path queries; the command is `:play movie graph`

- **The Northwind Database** is another tutorial showing how to import data from CSV files representing nodes and how to build relationships using Cypher commands; the command is `:play northwind graph`

You can follow the **movie graph tutorials** to get familiar with the graph database concept and the various constructs of Cypher query language. Please do not clear the database after you have finish the tutorial. We need the database for the next exercise. The **northwind database** tutorial will be useful for the project.
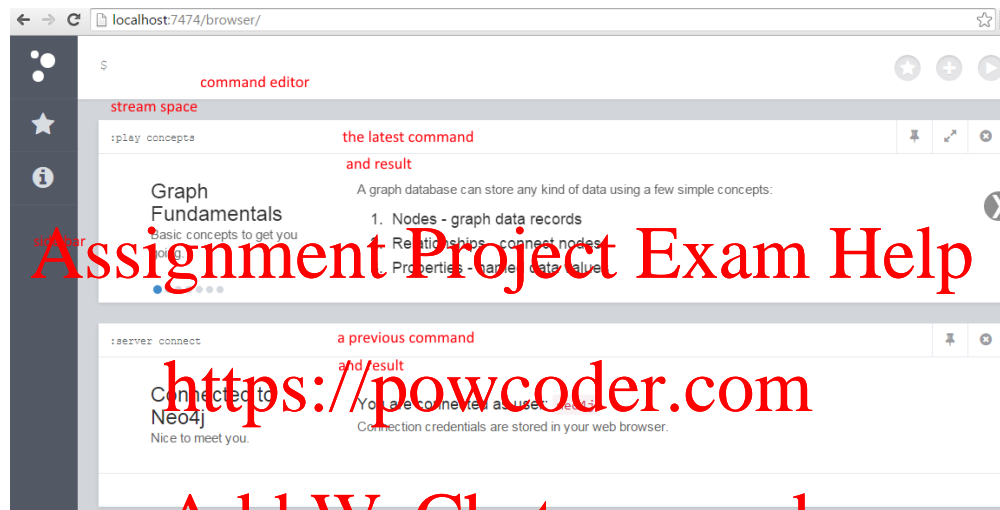

Figure 1: Neo4j Web Browser

Now try to answer the following questions using Movie Graph:

- Find out how many movies Tom Hanks acts in.

- Find out how many actors in the movie 'Cloud Atlas'.

- Find out the average rate of the movie 'The Replacements'.

- Find out the average age of actors in the movie 'Cloud Atlas'.

- Find out which movie has the highest average rate.

- Find out the top 3 movies based on the number of rating it received.