**Importing IMDB dataset into MySQL database**

Transferring IMDB dataset plain text files to MySQL server is an essential step before analyzing data mining tasks. I went through installing and running MySQL server on my machine. I am using OS X 10.11 64bit as the OS platform. In order to this long process to run, it requires the memory to have at least 4 GB of free space as a partition. We will also be using IMDbPy, a python based package. Of course, this library needs us to have Python installed on our machine as well as its development tools along with SQLObject. After installing required software, we will have to download all files from IMDb site: <ftp://ftp.fu-berlin.de/pub/misc/movies/database/>. Then, we can go ahead and create database and begin importing process.

**Steps for installing IMDbPy and SQLObject :**

1. Install required packages:

sudo easy\_install pip

sudo python setup.py install

sudo easy\_install MySQL-python

1. Download only .gz files from IMDb in the current directory:

wget -r --accept="\*.gz" --no-directories --no-host-directories --level 1 ftp://ftp.fu-berlin.de/pub/misc/movies/database/

1. Begin the import process:

imdbpy2sql.py -d /Users/Bader/Documents/Spring2016/Data-Mining/Project/ -u 'mysql://imdb:imdb@localhost/imdb'

We are aware that –d is the directory where .gz files are located and –u is the database connection string for the MySQL database. A screenshot of the output is given below:

