Jiaqi Wang- (NUID 001023711)

# Big Data System Engineering with Scala Fall 2021 sec01 Assignment No. 1



#### Task

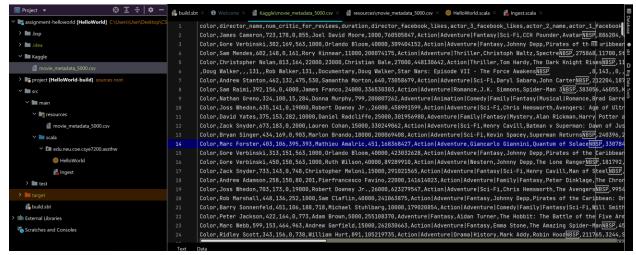
# part1-1

```
PS C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld> <mark>sbt</mark> test run
[info] welcome to sbt 1.4.0 (Oracle Corporation Java 11.0.12)
[info] loading global plugins from C:\Users\User\.sbt\1.0\plugins
[info] loading project definition from C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld\project
[info] loading settings for project assignment-helloworld from build.sbt ...
[info] set current project to HelloWorld (in build file:/C:/Users/User/Desktop/CSYE7200-fall/assignment-helloworld/)
[info] HelloWorldSpec
[info] HelloWorld
[info] - should get the correct greeting
[info] Ingest
[info] - should find the correct number of movies
[info] Run completed in 578 milliseconds.
[info] Total number of tests run: 2
[info] Suites: completed 1, aborted 0
[info] Tests: succeeded 2, failed 0, canceled 0, ignored 0, pending 0
[info] All tests passed.
 success] Total time: 2 s, completed 2021赛?鏈?0鏃?涓婂崃1:25:19
[0J
Multiple main classes detected. Select one to run:
[1] edu.neu.coe.csye7200.assthw.HelloWorld
[2] edu.neu.coe.csye7200.assthw.Ingest
Enter number: 1
[info] running edu.neu.coe.csye7200.assthw.HelloWorld
Hello World!
 SUCCESS] Total time: 4 s, completed 2021寒?鏈?0鏃?涓嫦崃1:25:24
PS C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld>
```

### part1-2

## part2-1

Download csv file and put it into Kaggle folder



#### - part2-2

```
PS C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld> sbt run Kaggle/movie_metadata_5000.csv
[info] welcome to sbt 1.4.0 (Oracle Corporation Java 11.0.12)
[info] loading global plugins from C:\Users\User\.sbt\1.0\plugins
[info] loading project definition from C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld\project
[info] loading settings for project assignment-helloworld from build.sbt ...
[info] set current project to HelloWorld (in build file:/C:/Users/User/Desktop/CSYE7200-fall/assignment-helloworld/)

Multiple main classes detected. Select one to run:
[1] edu.neu.coe.csye7200.assthw.HelloWorld
[2] edu.neu.coe.csye7200.assthw.Ingest

Enter number:

| => assignment-helloworld / Compile / selectMainClass 22s
```

```
Terminat Local + 

www.indb.com/title/tt1888418/?ref_=fn_tt_tt_1, 14, English, USA, Not Rated, 9808, 2011, 285, 6.4, , 413

Color, Scott Smith, 1, 87, 2, 318, Daphne Zuniga, 637, Conecylibrana, Eric Mabius, Signed Sealed Delivered®, 629, 2283, Crystal Lowe, 2, fraud|postal worker|prison|theft|trial, http://www.indb.com/title/tt3808844/?ref_=fn_tt_tt_1, 6, English, Lanada, , , 2013, 470, 7.7, 84

Color, , 43, 43, , 319, Valorie Curry, 841, Crime|Drama|Mystery|Thriller, Natalie Zea, The Following® , 73839, 1753, Sam Underwood, 1, cult|fbi|hideout|prison escape|serial killer, http://www.imdb.com/title/tt2071645/?ref_=fn_tt_tt_1, 359, English, USA, TV-14, , , 593, 7.5, 16.0, 32000

Color, Benjain Roberds, 15, 76, 0, Maxwell Moody, 0, Drama|Horror|Thriller, Eva Boehnke, A Plague So Pleasant®, 38, 0, David Chandler, 0, http://www.imdb.com/title/tt207644/?

ref_=fn_tt_tt_1, 3, English, USA, , 1400, 2013, 0, 6.3, , 16

Color, Daniel Hsia, 14, 100, 0, 489, Daniel Henney, 946, 10443, Comedy|Drama|Romance, Alan Ruck, Shanghai Calling®, 1255, 2386, Eliza Coupe, 5, http://www.imdb.com/title/tt2070597/?

ref_=fn_tt_tt_1, 9, English, USA, PG-13, , 2012, 719, 6.3, 2.35, 660

Color, Jon Gunn, 43, 90, 16, 16, Brian Herzlinger, 86, 85222, Documentary, John August, My Date with Drew®, 4285, 163, Jon Gunn, 0, actress name in title|crush|date|four word title|video camera, http://www.imdb.com/title/tt0378407/?ref_=fn_tt_tl_1, 84, English, USA, PG, 1100, 2004, 23, 6.6, 1.85, 456
```

#### -part2-3



Outcome: 15

github link:https://github.com/JiaqiWang1996/CSYE7200

# final house keeping

```
sbt:HelloWorld> clean
[success] Total time: 0 s, completed 2021寒?鏈?9鏃?涓嫦崍2:03:06
```