

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> sbt 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年10月25日 1:25:19
[0]
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年10月25日 1:25:24
[0]
PS C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld> 
```

### – part1-2

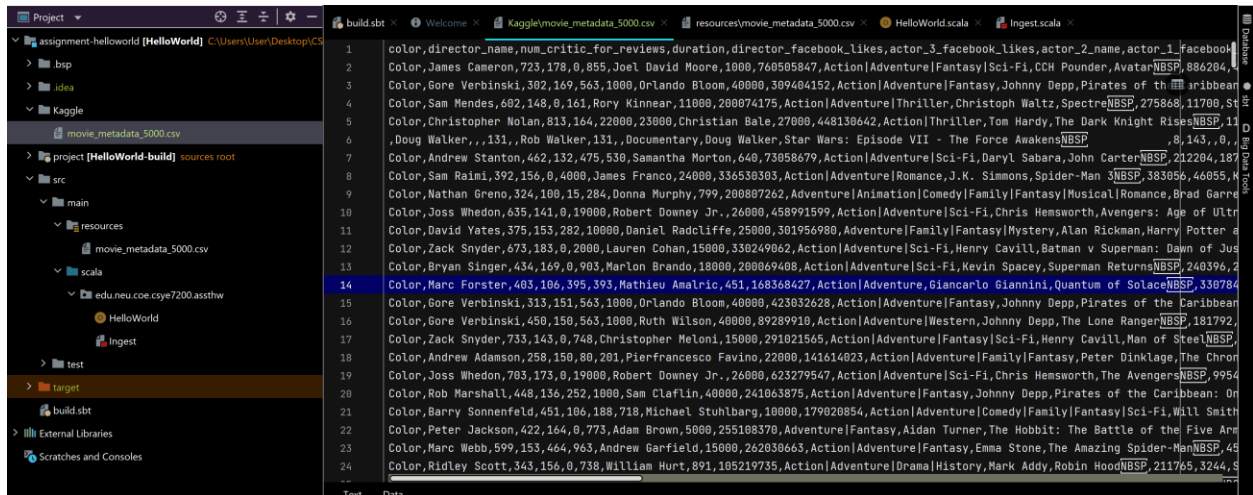
```
PS C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld> sbt run
[info] welcome to sbt 1.4.0 (Oracle Corporation Java 11.0.12)
[info] loading global plugins from C:\Users\User\.sbt\1.0\plugins
[info] loading project definition from C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld\project
[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/)

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: 13 s, completed 2021年10月25日 1:28:28
[0]
PS C:\Users\User\Desktop\CSYE7200-fall\assignment-helloworld> 
```

### – part2-1

Download csv file and put it into Kaggle folder



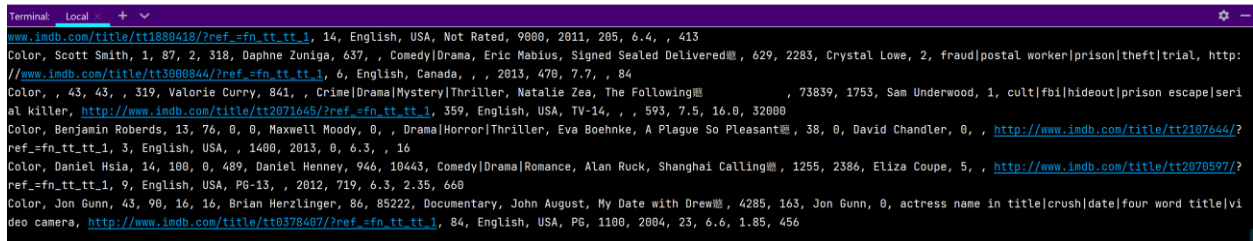
## - part2-2

```
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: 1

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



## -part2-3

```
36 }
37
38 // for (m <- ingester(source)) println(m.properties.mkString(", "))
39 val kiwiMovies = for (m <- ingester(source); if (m.properties(20)=="New Zealand")) yield m
40 println(kiwiMovies.size)
41 source.close()
42
43 // Please note that an alternative to the definition of source above would be as in the following comment:
44 // val source = Source.fromFile(if (args.length>0) args.head else "movie_metadata_5000.csv")
45
46
```

Process finished with exit code 0

sbt:HelloWorld> clean  
[success] Total time: 0 s, completed 2021年?键?0熾?消 始 2:03:06

Outcome : 15

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

final house keeping