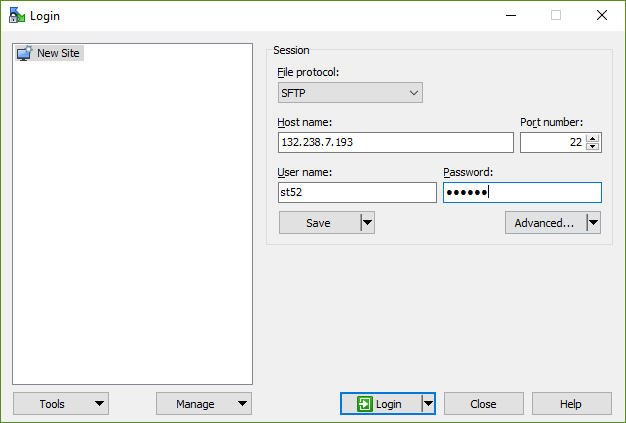
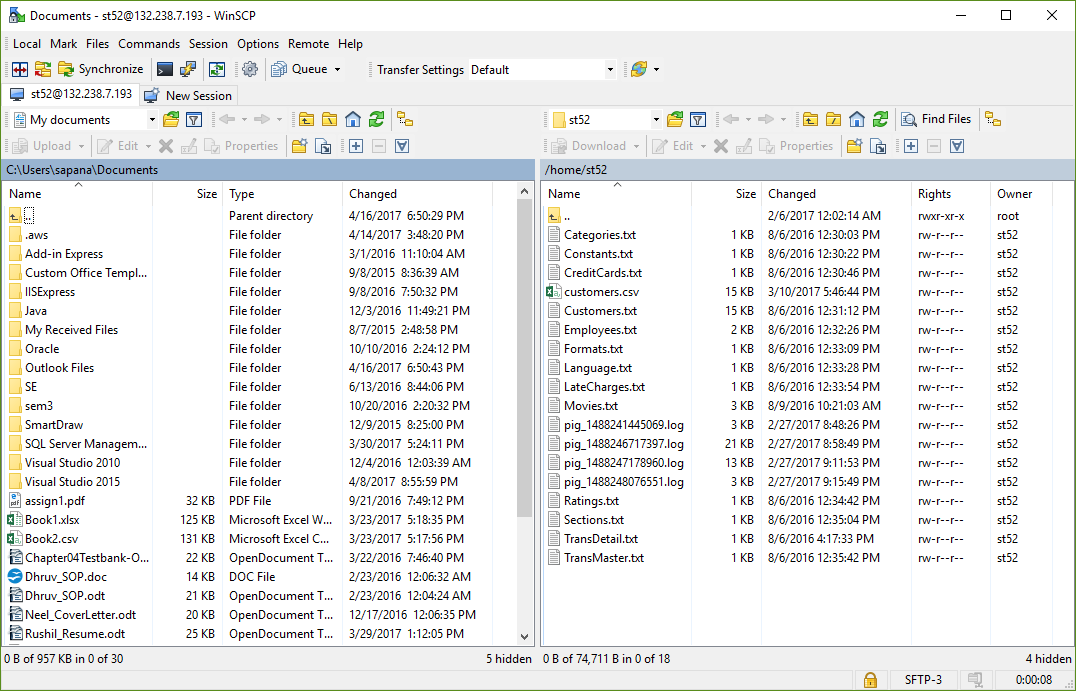
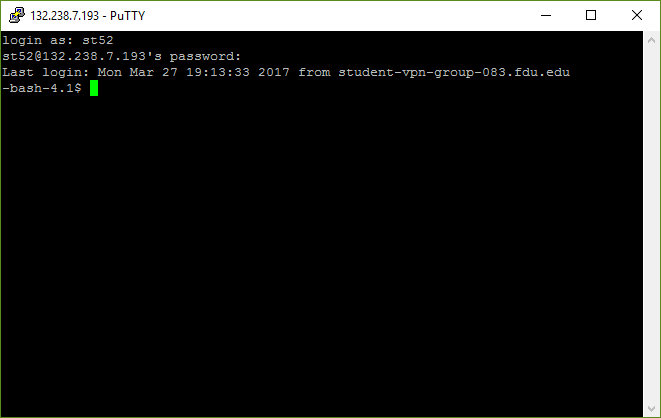
Project

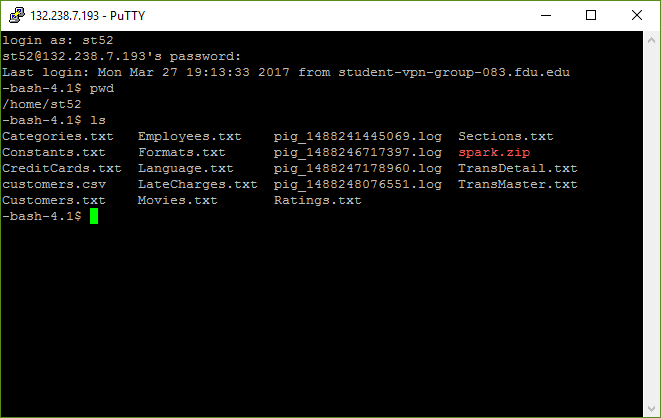
* Connect to WinSCP to Transfer file Spark.zip from my Computer to the fdu cluster.



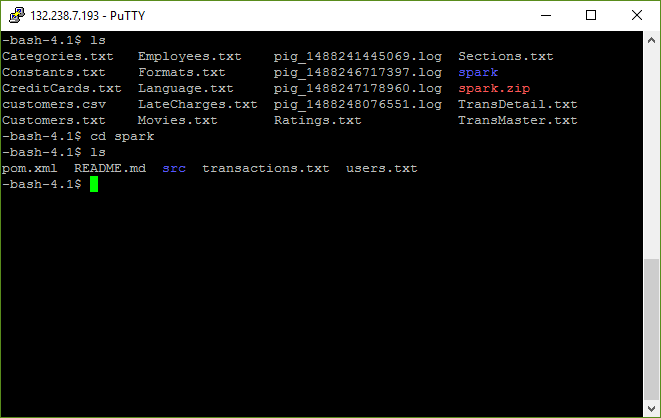


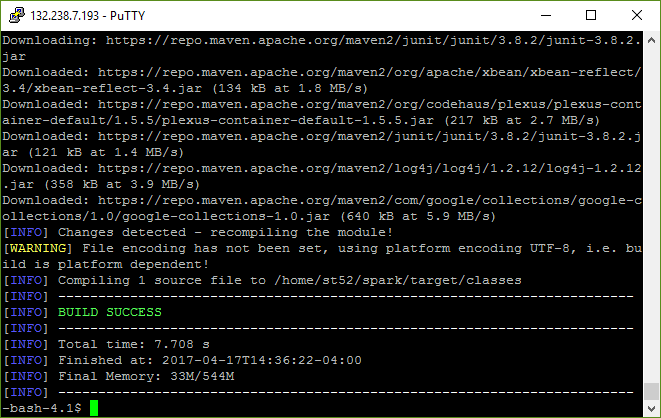
* Open PUTTY.

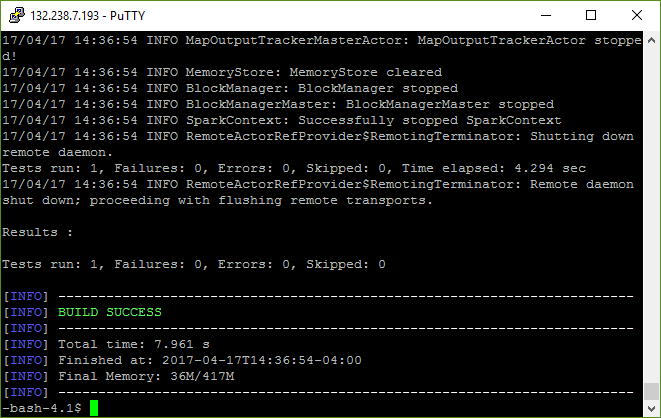


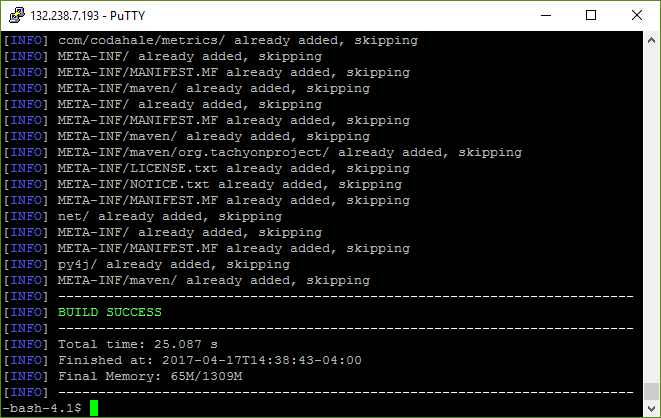


* Unzip the Spark.zip and Change directory to spark by cd spark.

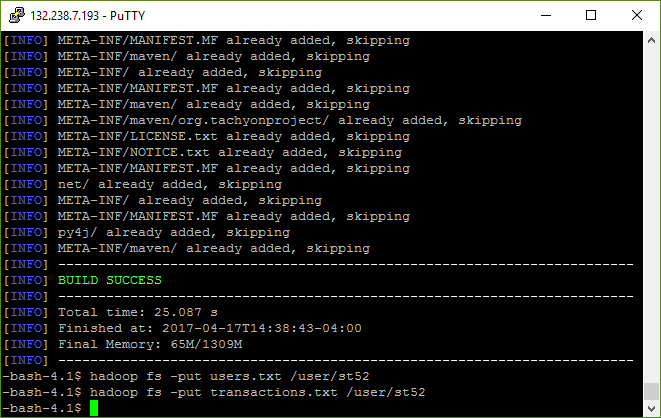


* Now Compile the Source Code using mvn compile by Apache Maven. Here is the Result.
* Lets Run mvn test command to perform task.

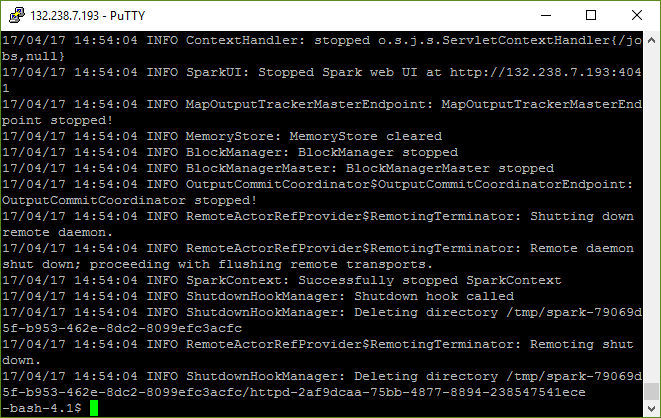




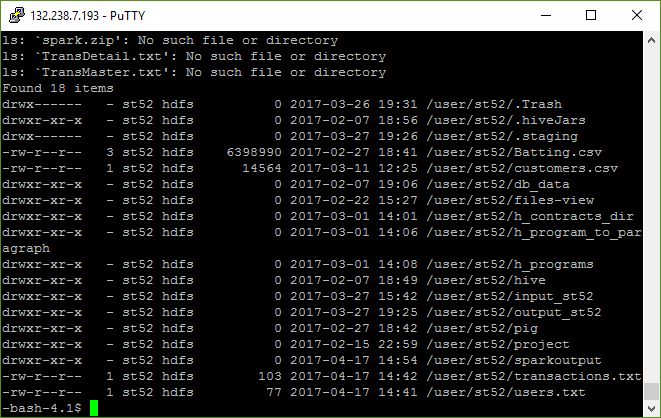
* Now copy users.txt and transactions.txt to my folder in HDFS.
* Hadoop fs –put users.txt /user/st52.

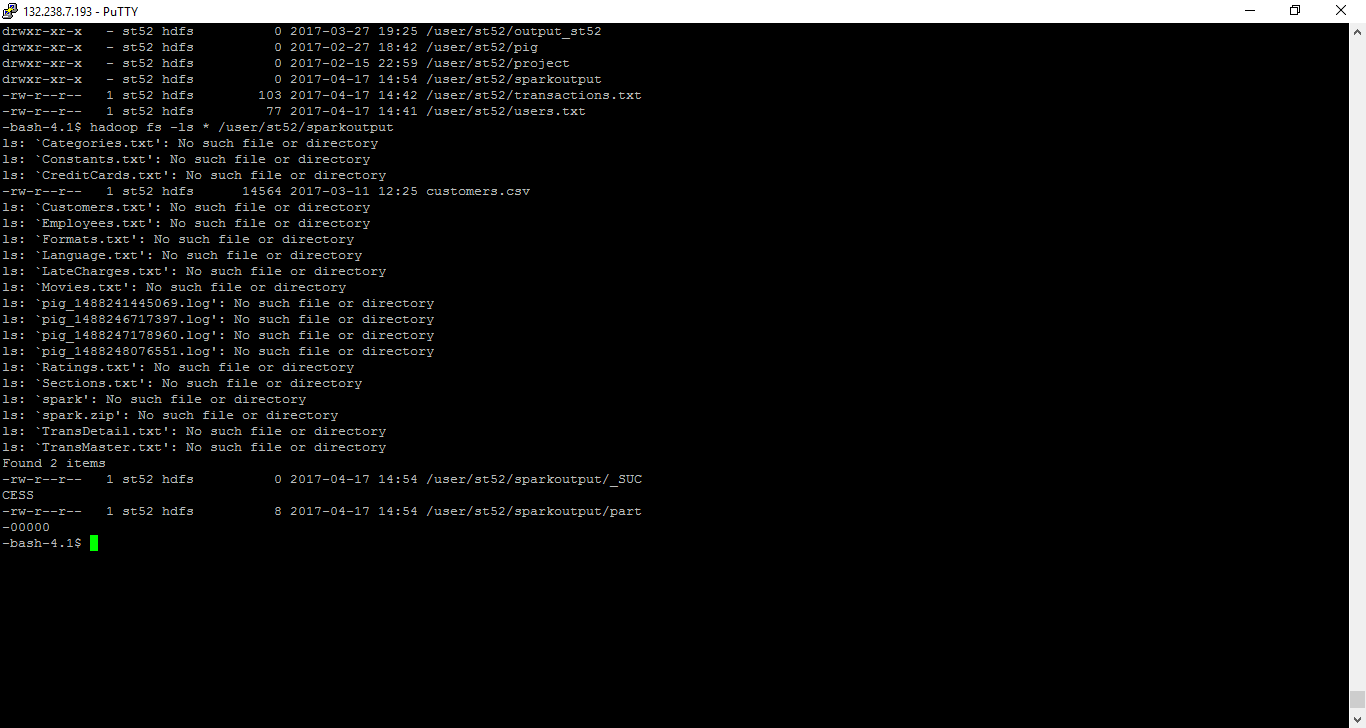


* Now use following command
* /usr/bin/spark-submit --class main.java.com.matthewrathbone.sparktest.ExampleJob -- master local ./target/spark-example-1.0-SNAPSHOT-jar-with-dependencies.jar /user/st52/transactions.txt /user/st52/users.txt /user/st52/sparkoutput

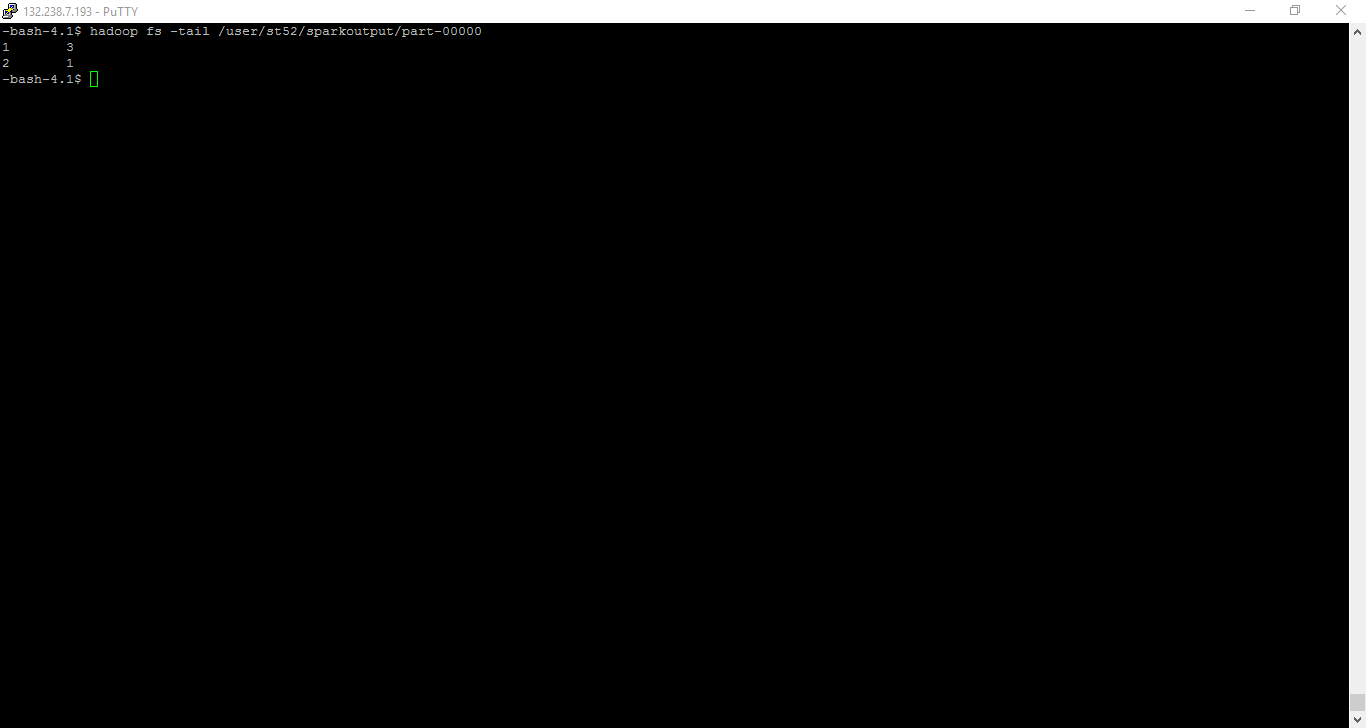


* There will be new directory created, sparkoutput. Check it out.





* Check output in sparkoutput folder.



**Install Tableau**

