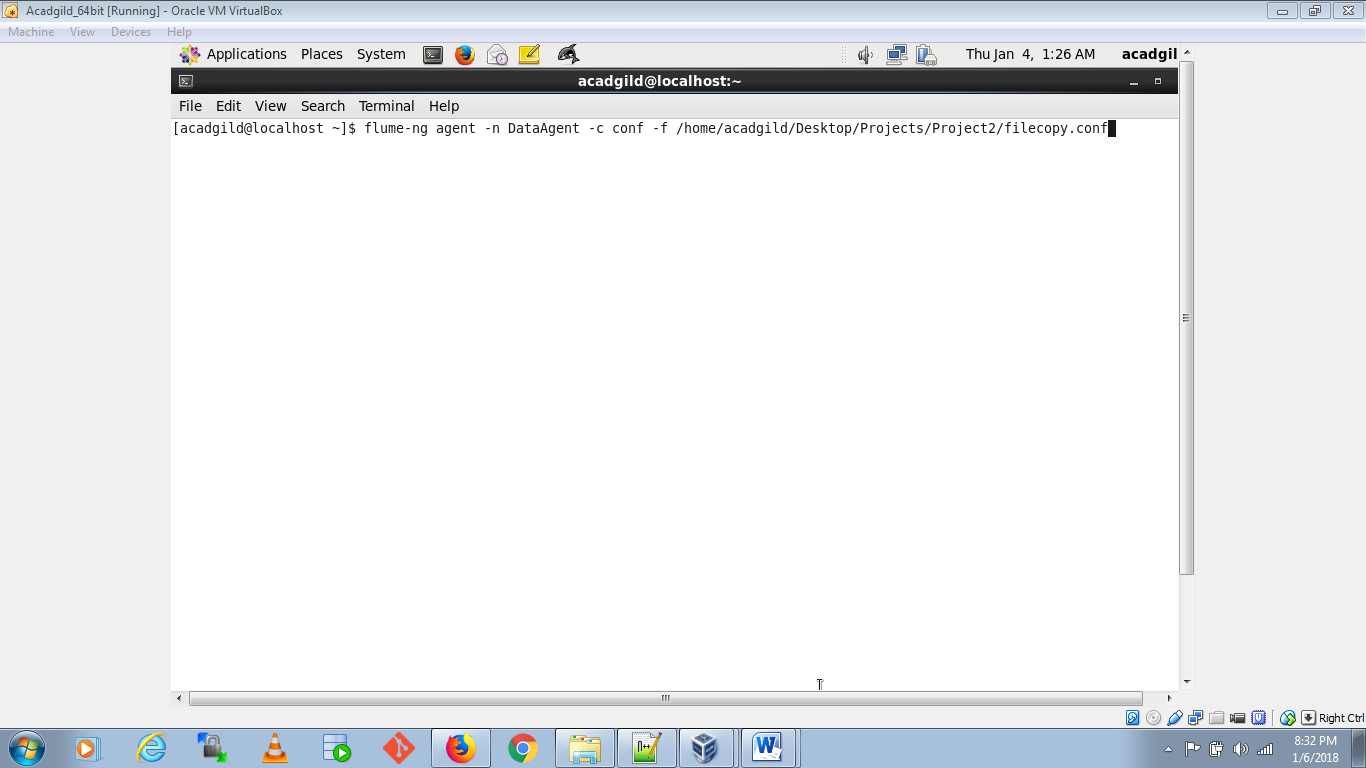
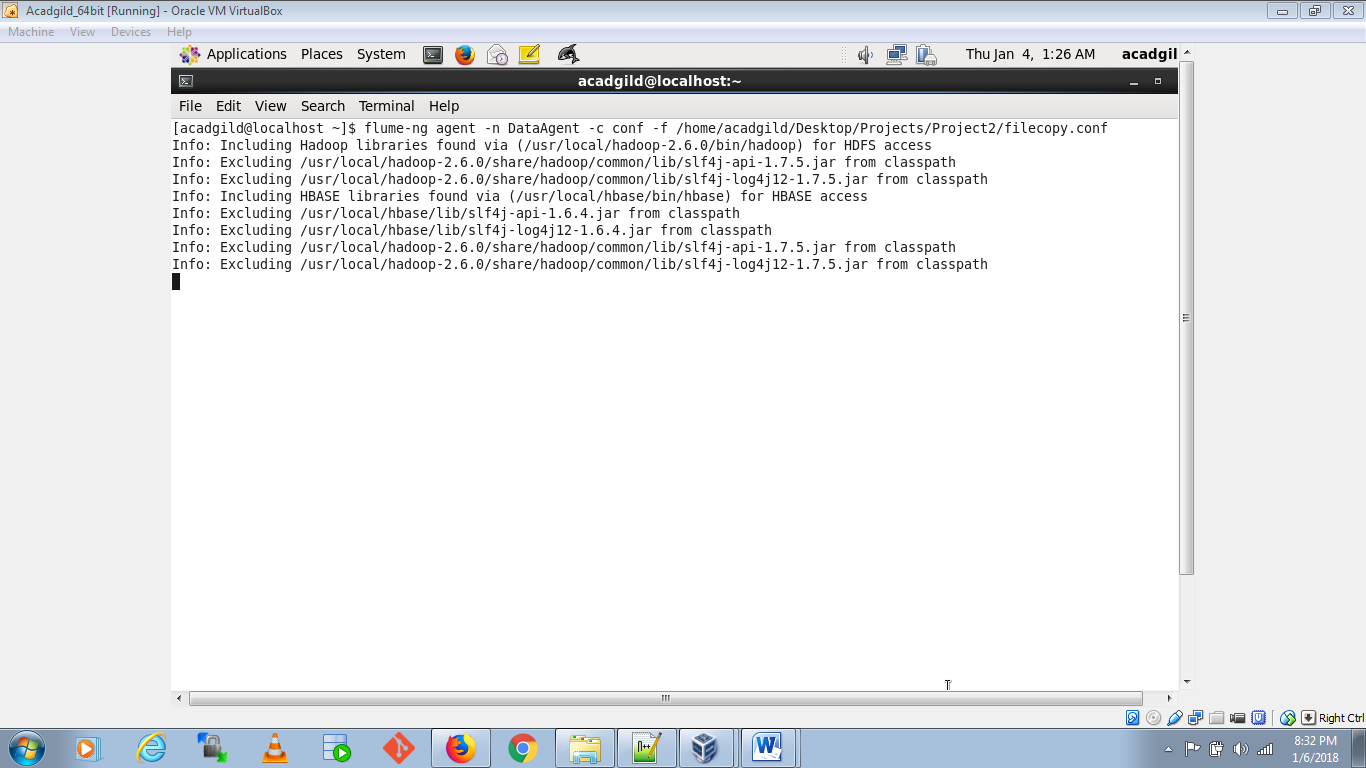
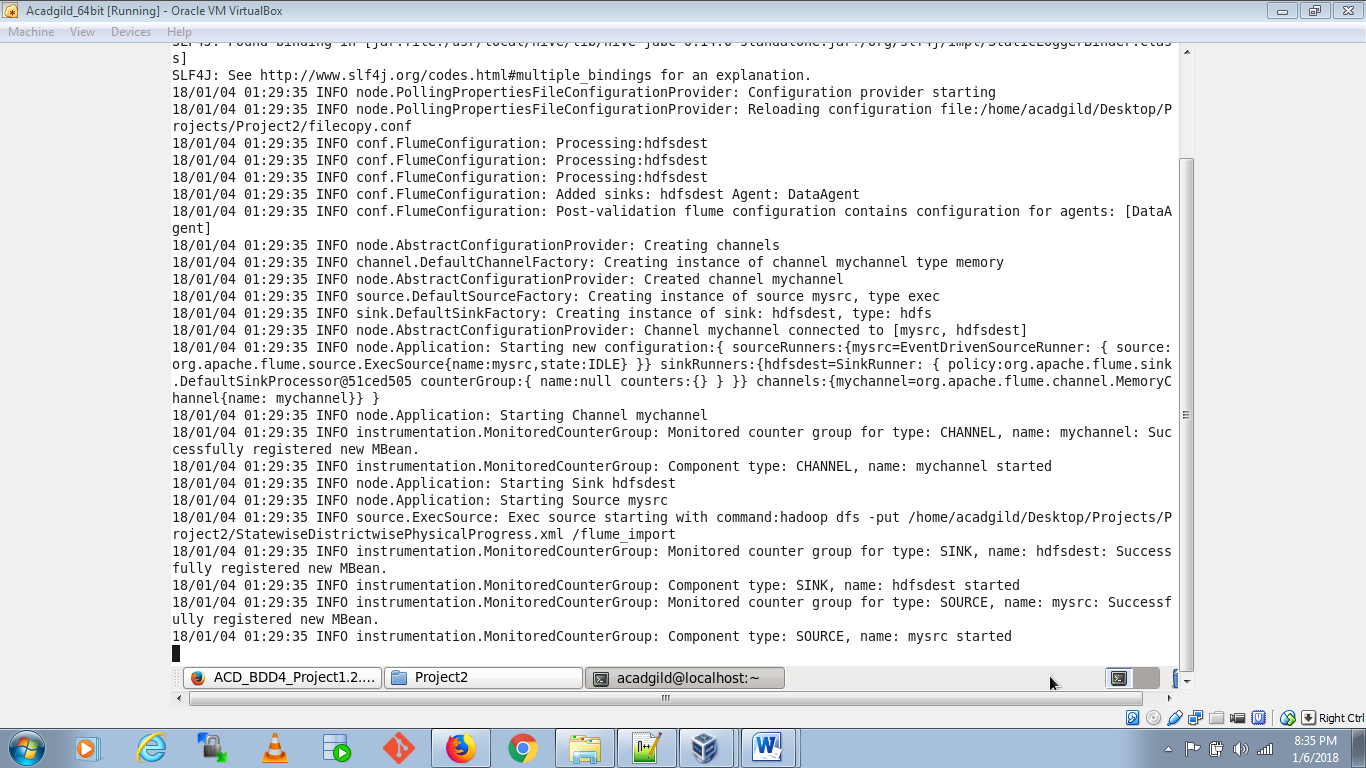
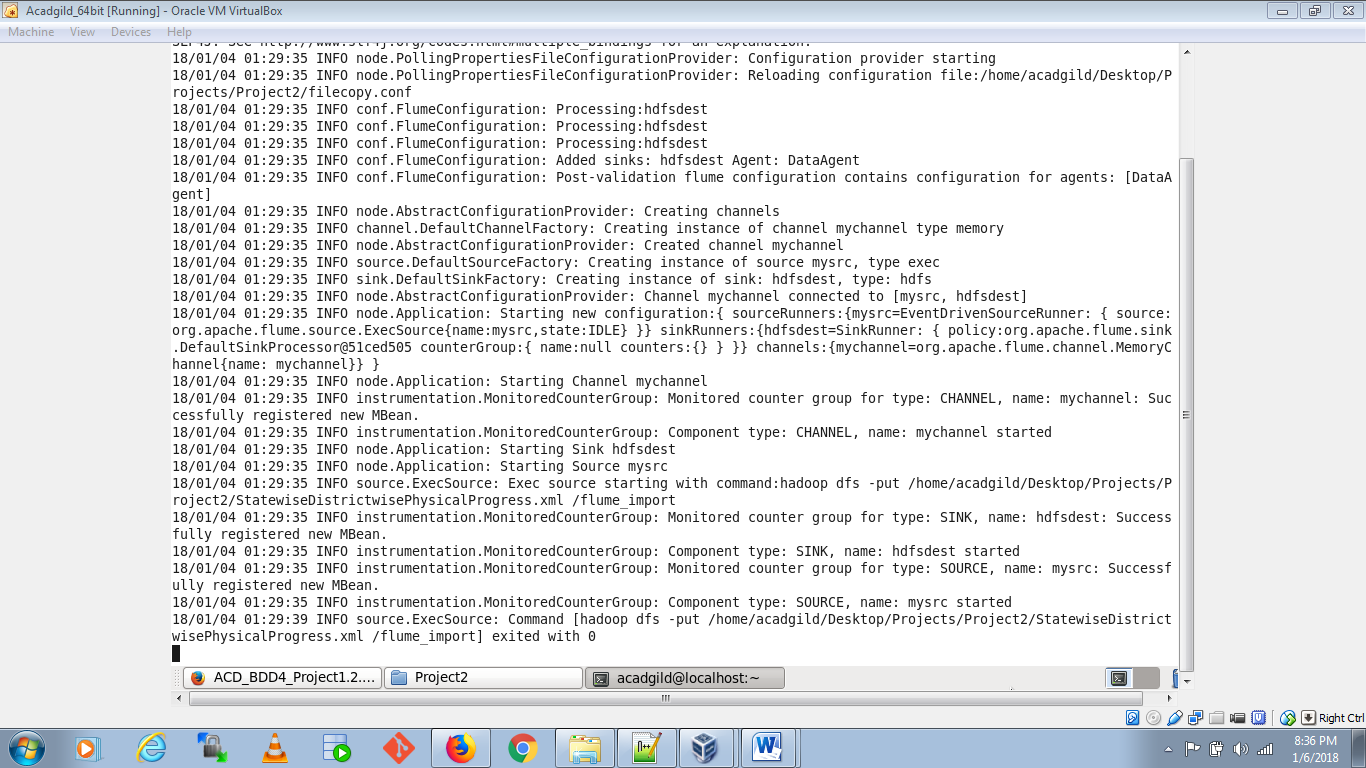
**STEP1:**

Copy dataset from local file system to HDFS using flume:

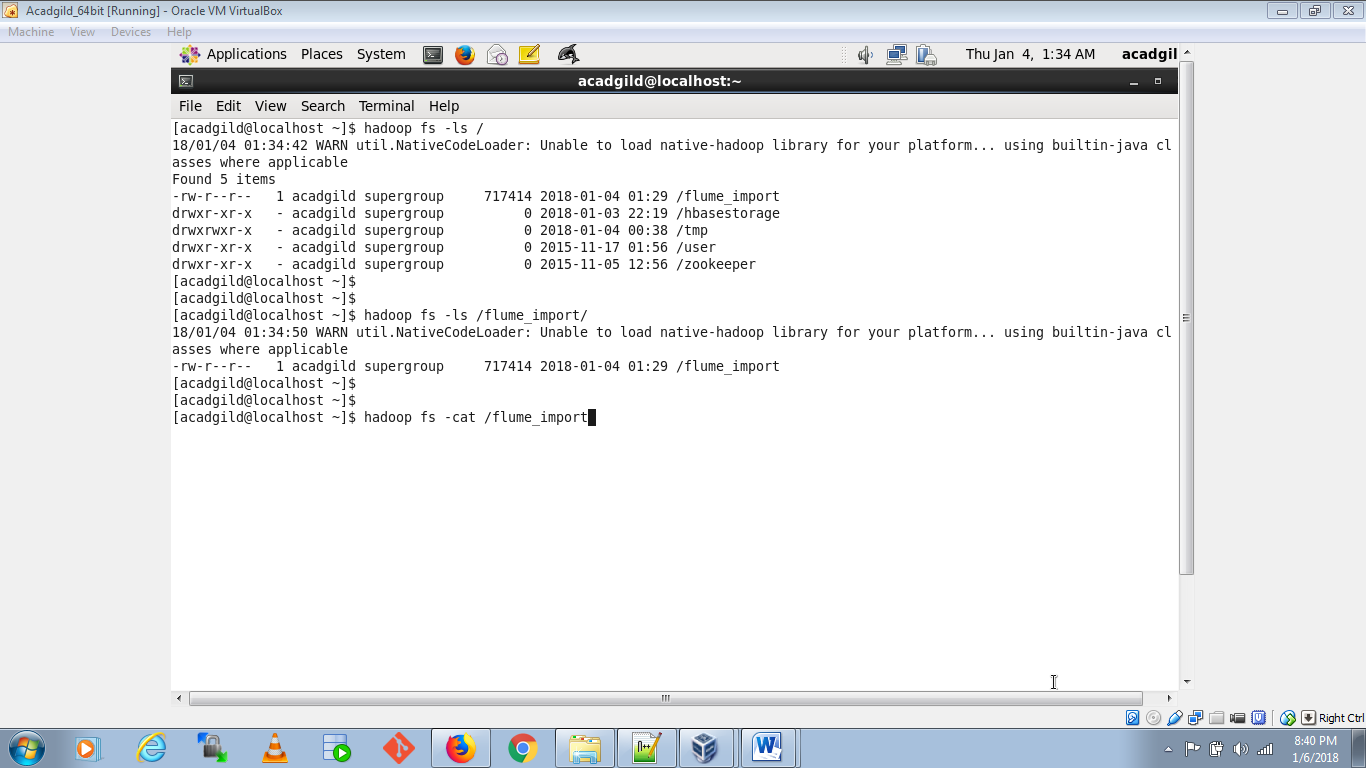


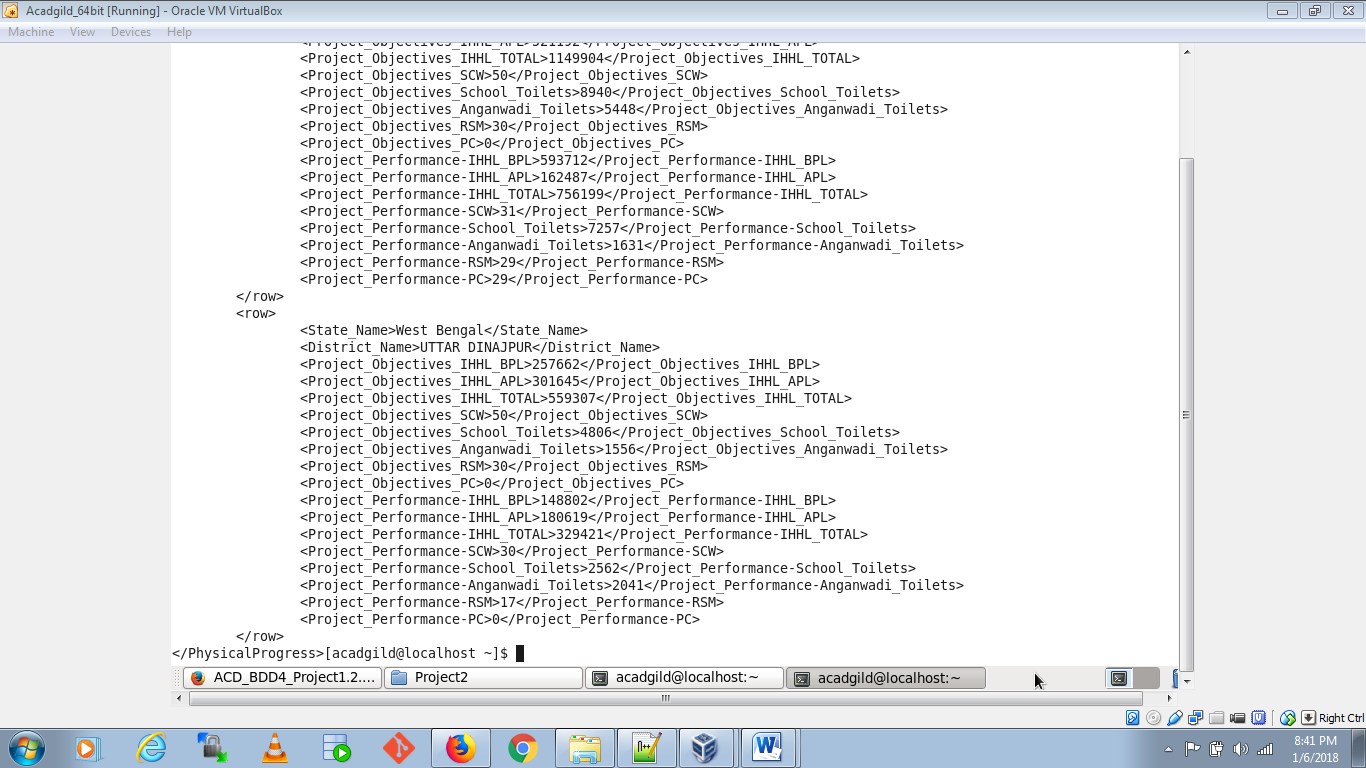






Data imported into HDFS:





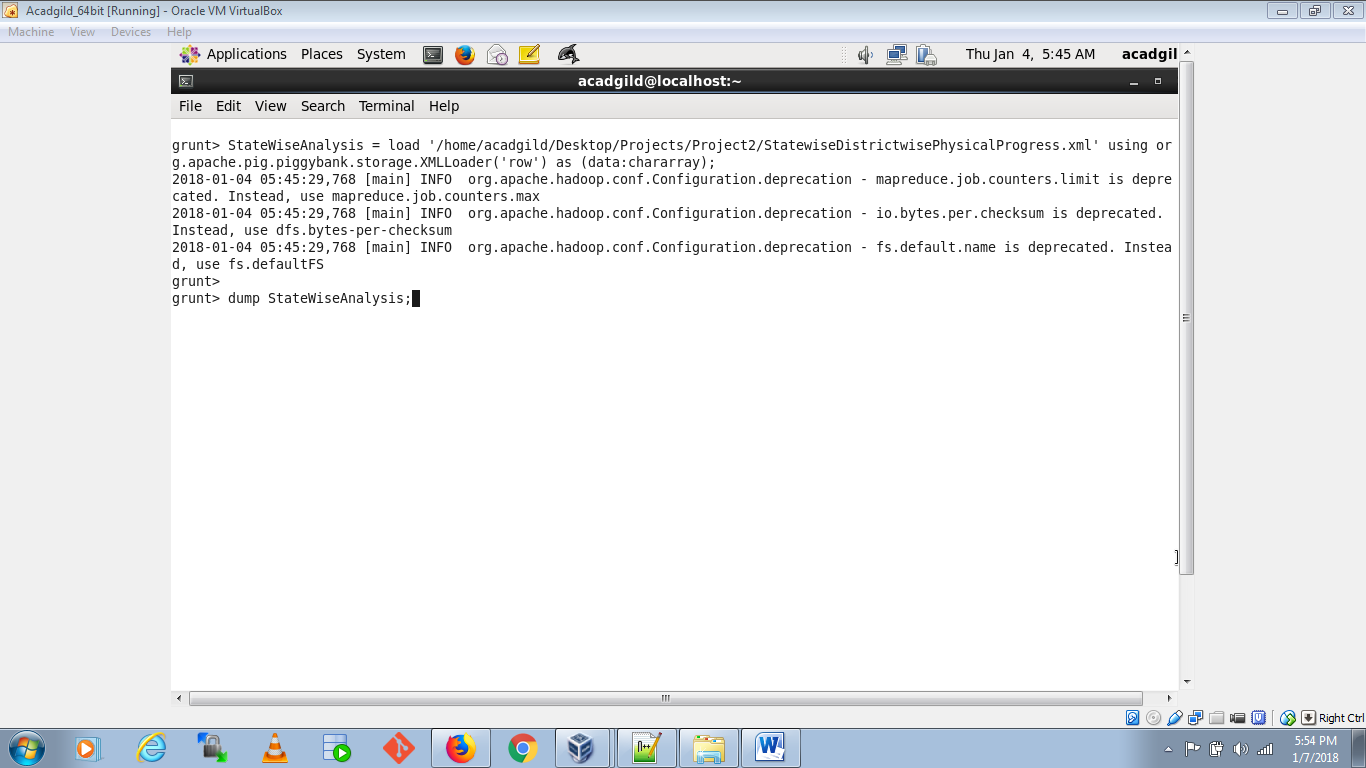
**STEP2:**

Register Piggybank.jar in grunt shell:



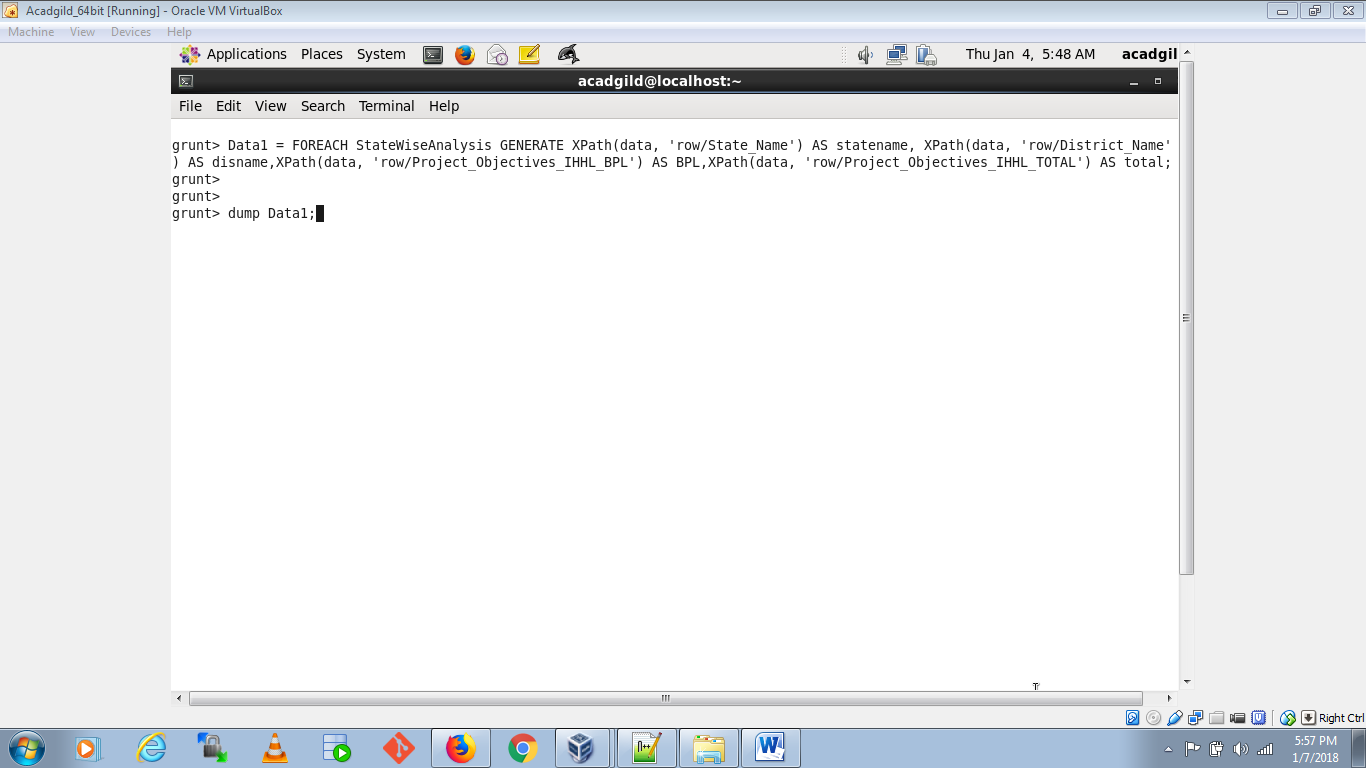
1. **Find out the districts who achieved 100 percent objective in BPL cards**

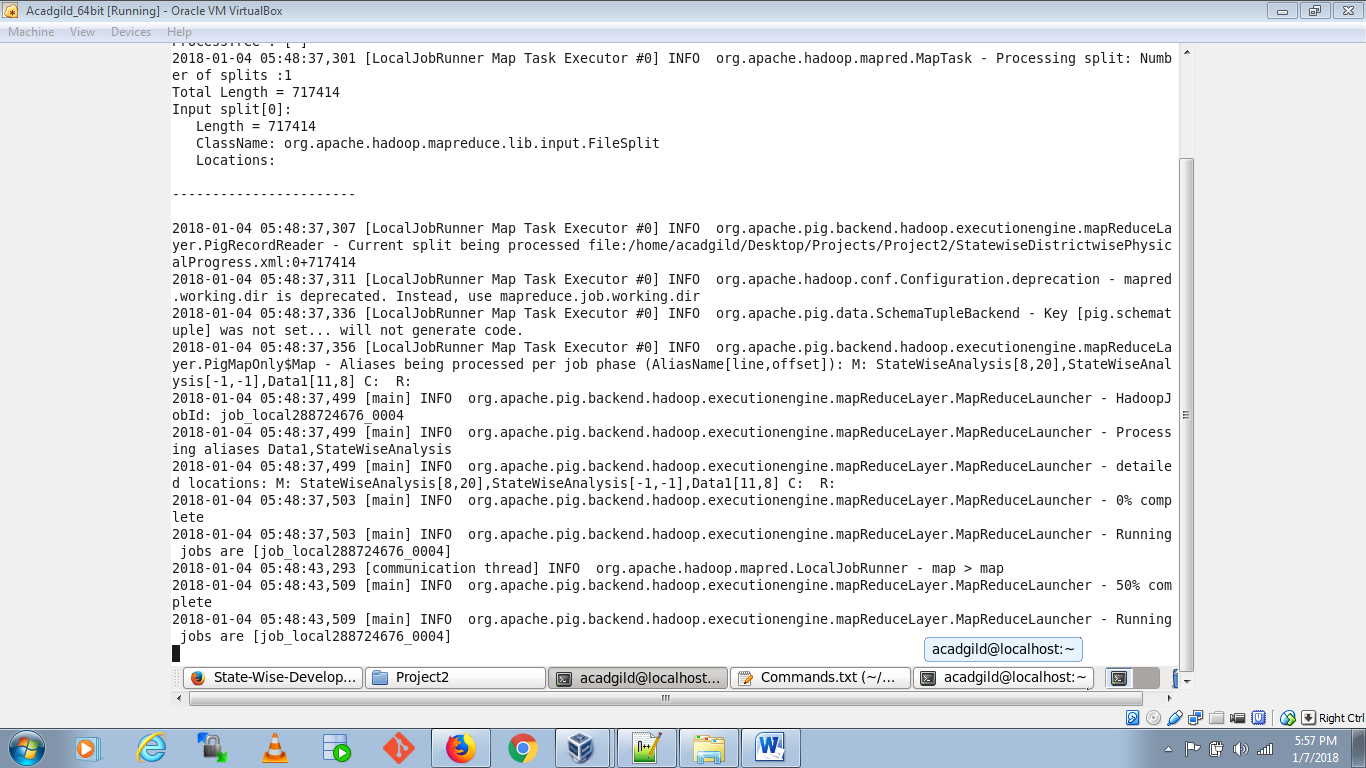
Loading HDFS xml file into PIG Grunt shell:





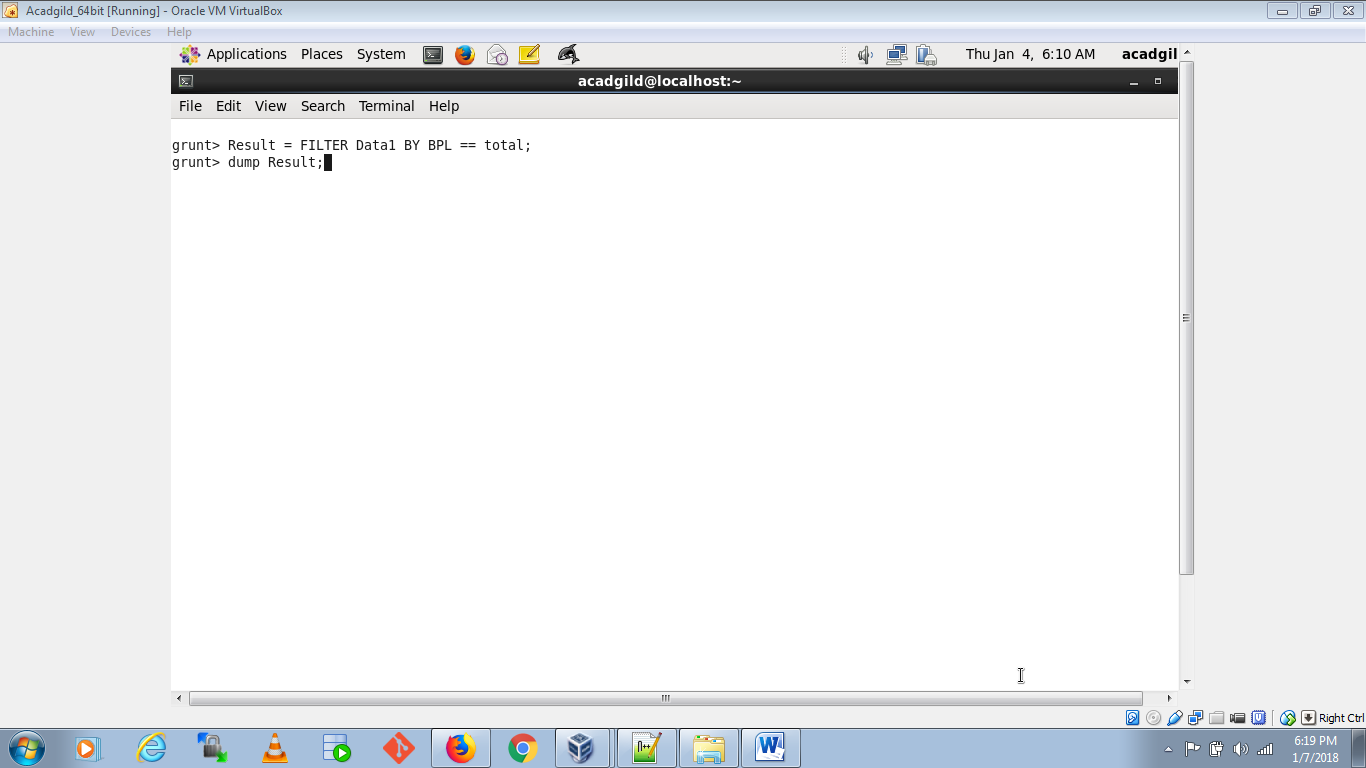
Parsing loaded XML data in PIG:

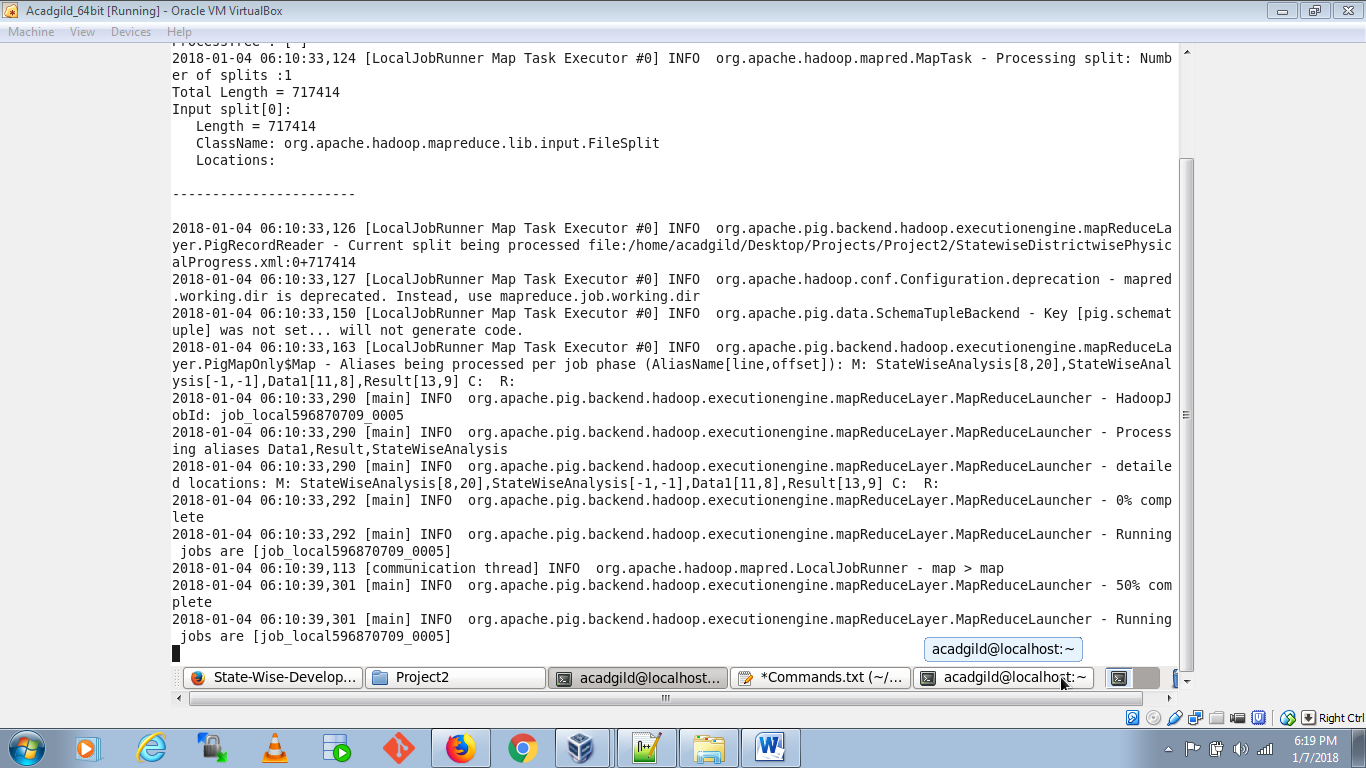


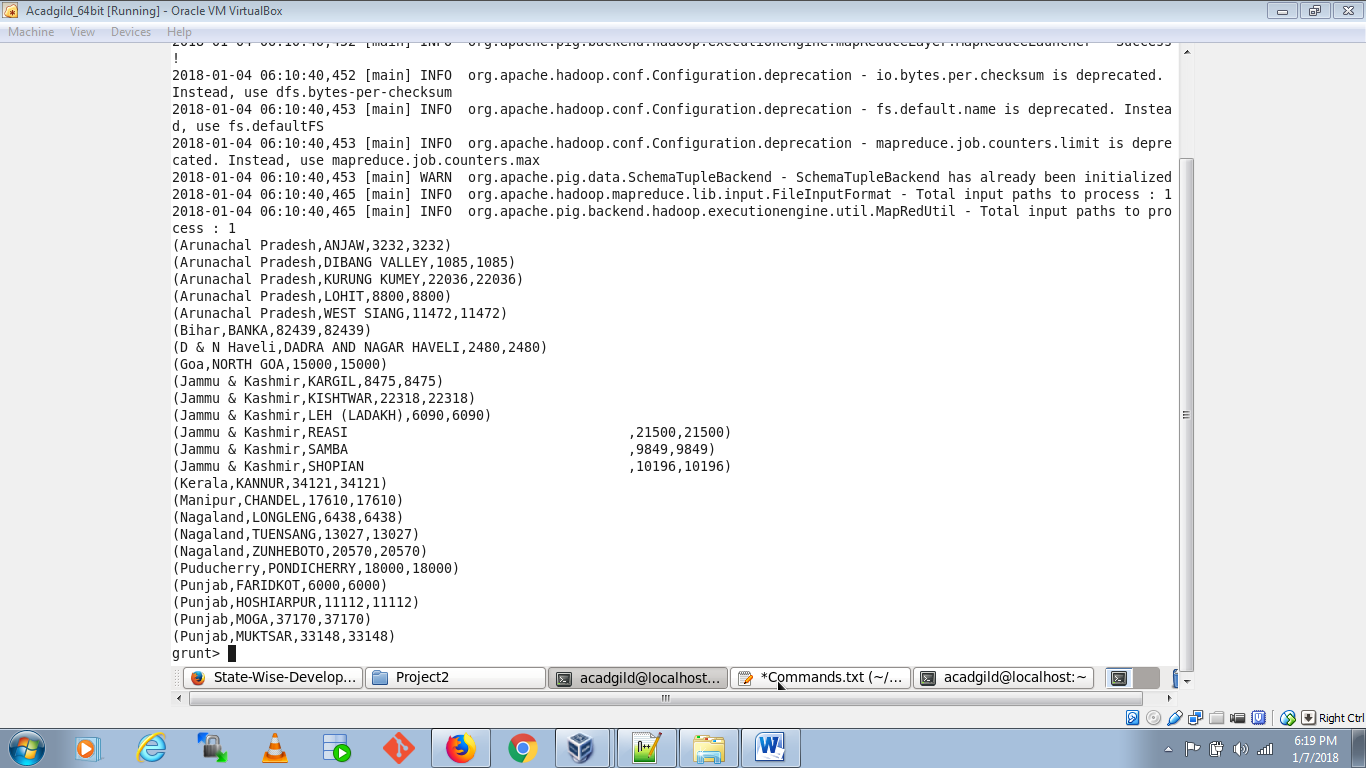




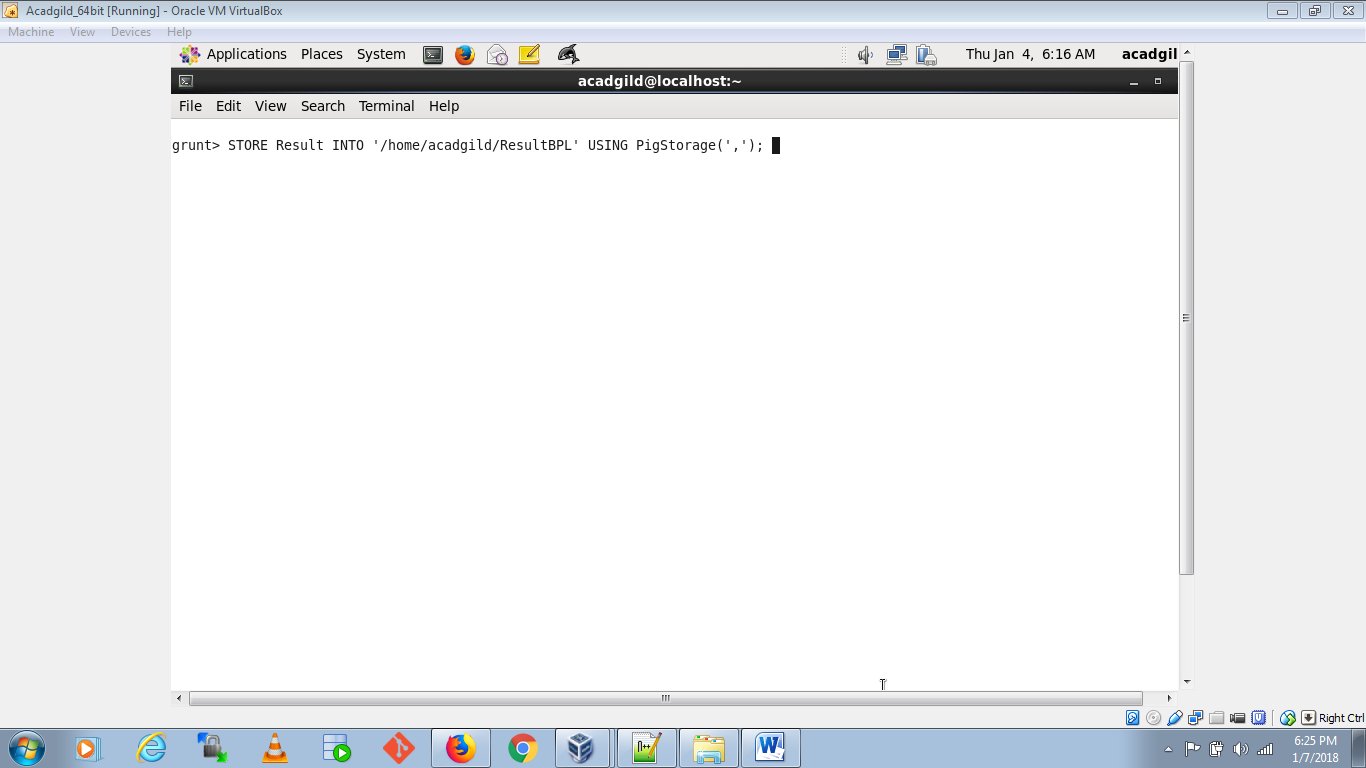
Find out the districts who achieved 100 percent objective in BPL cards



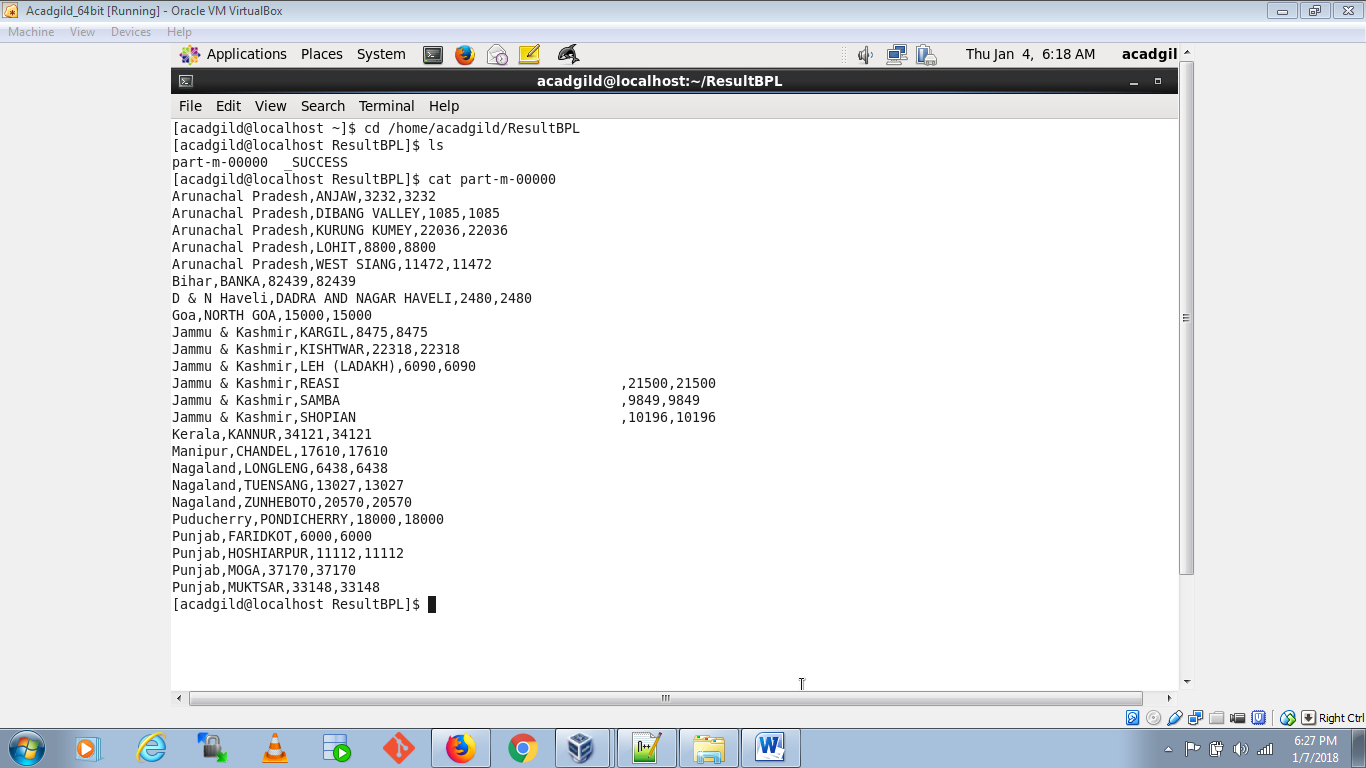




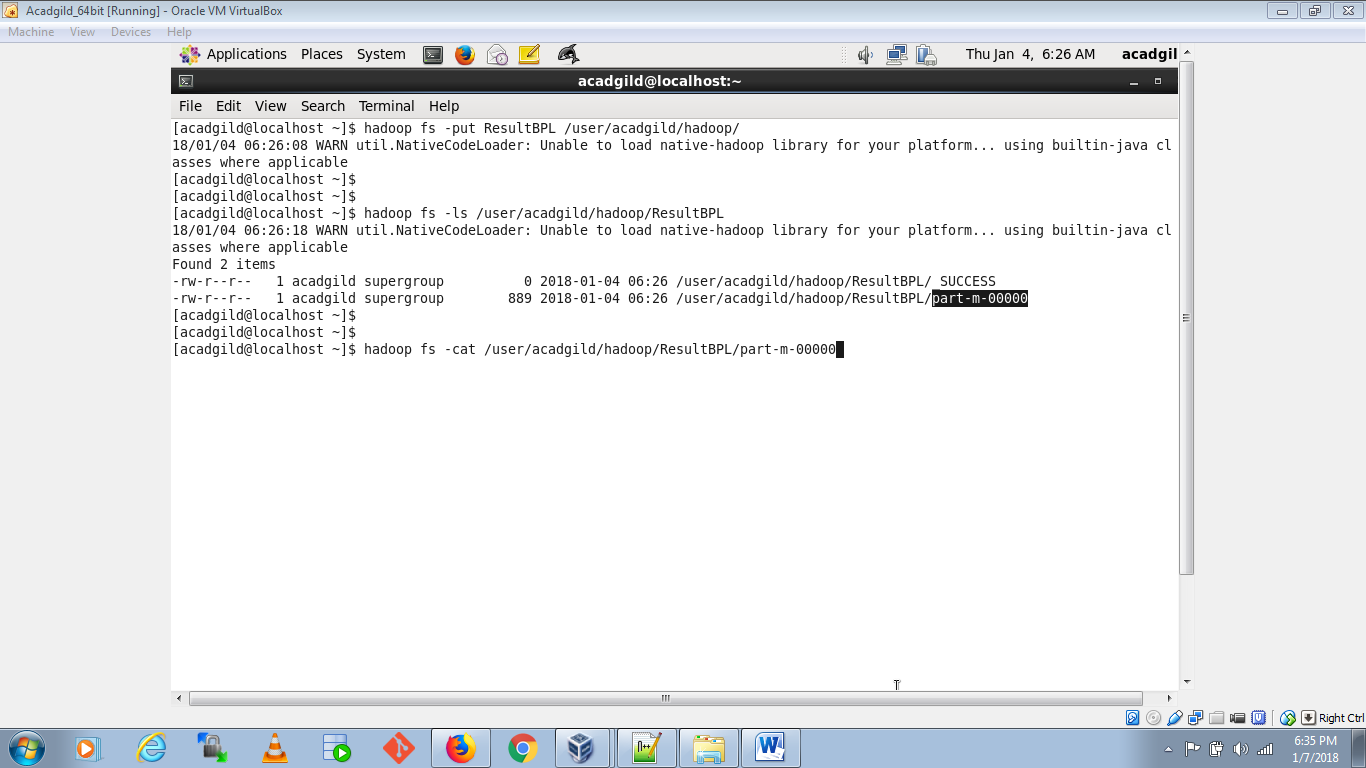
Export Result data from PIG shell to file system:

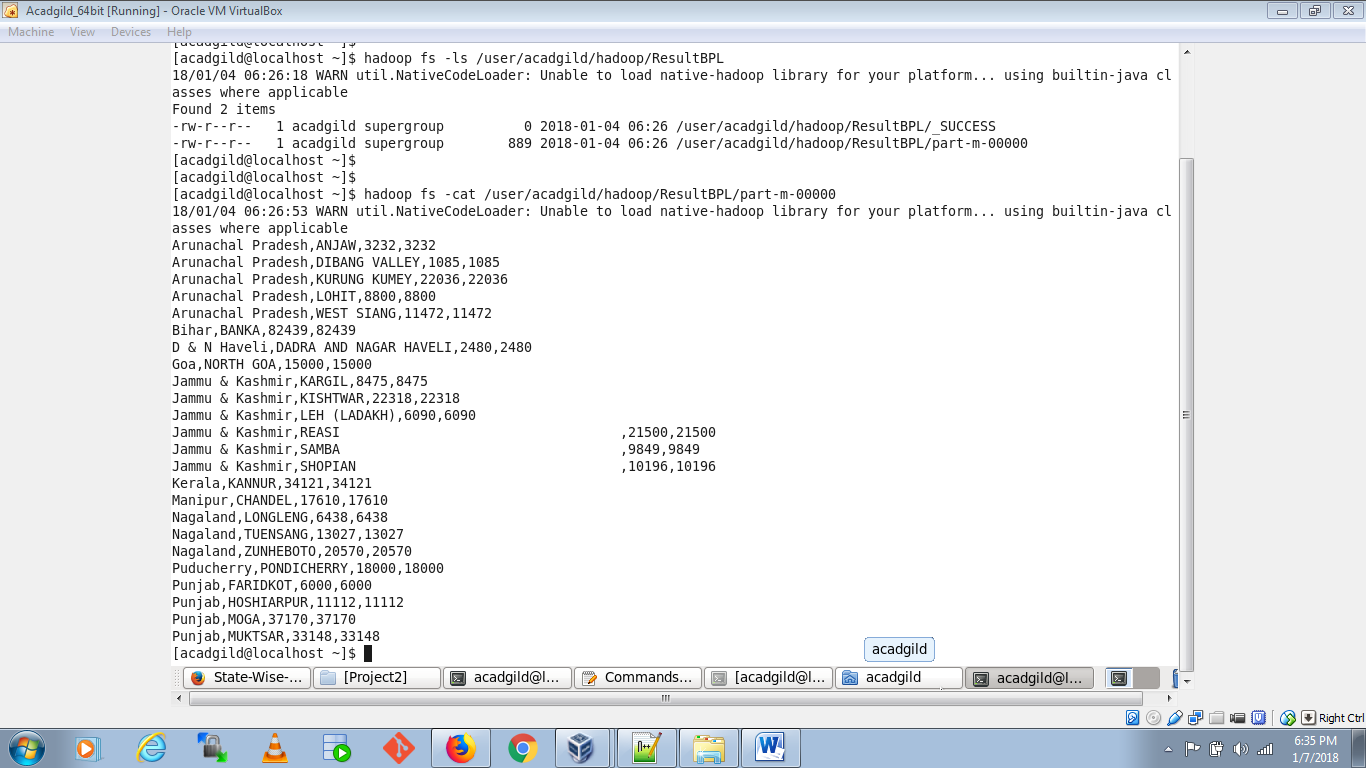




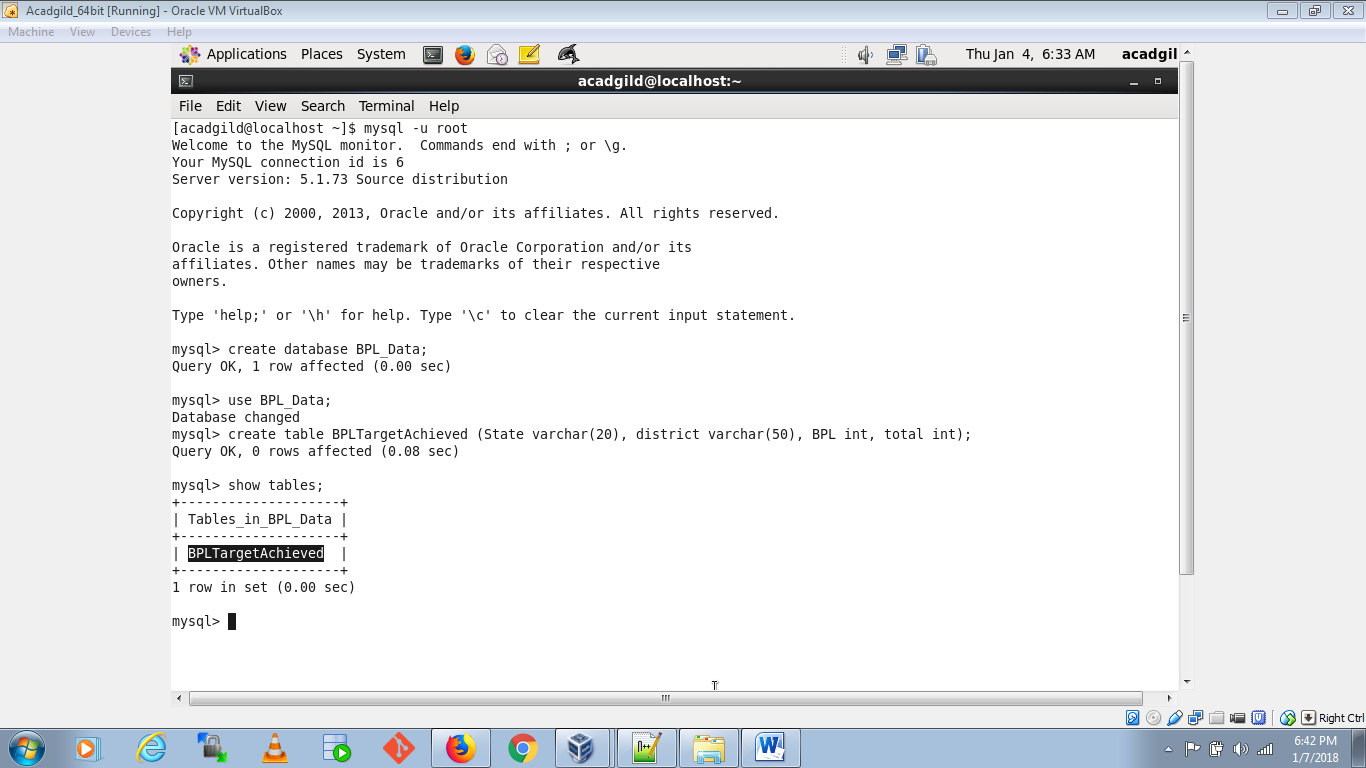


Copy Result data from file system to HDFS:

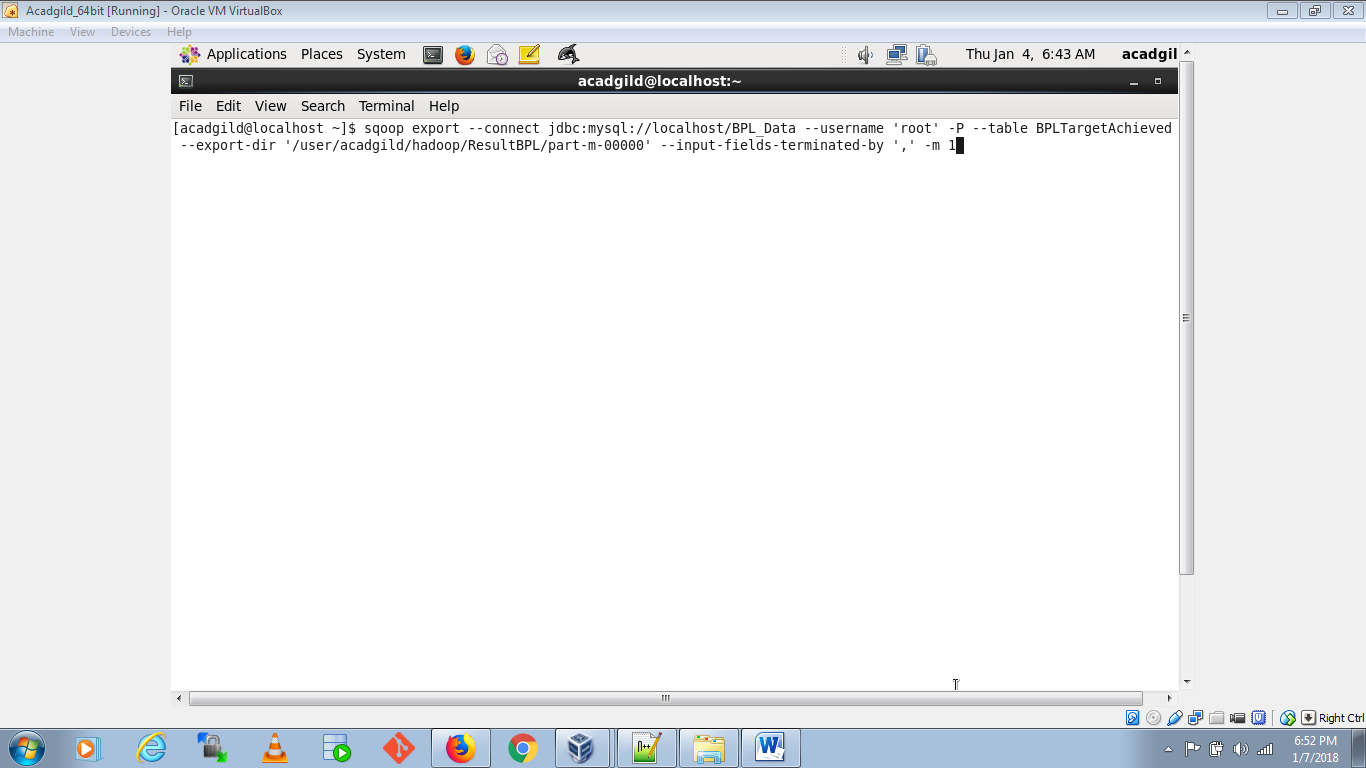


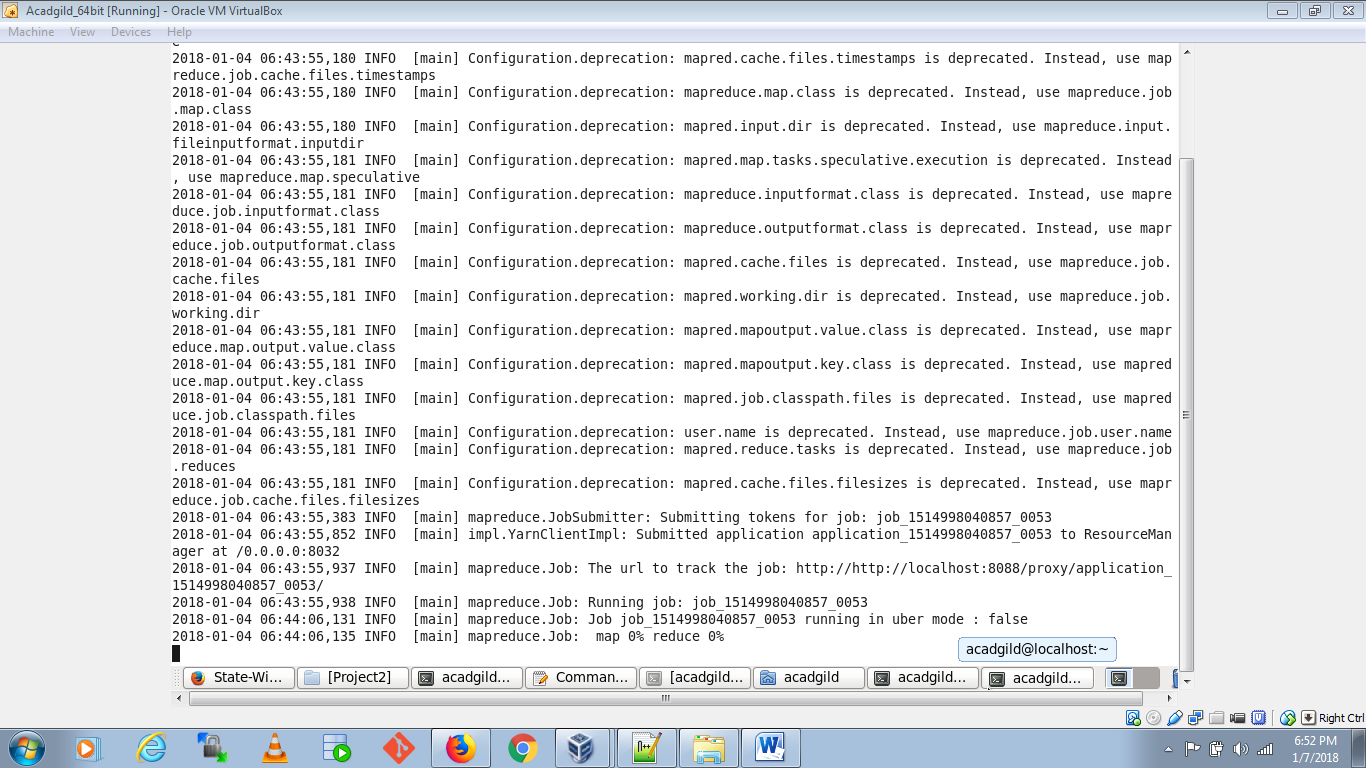


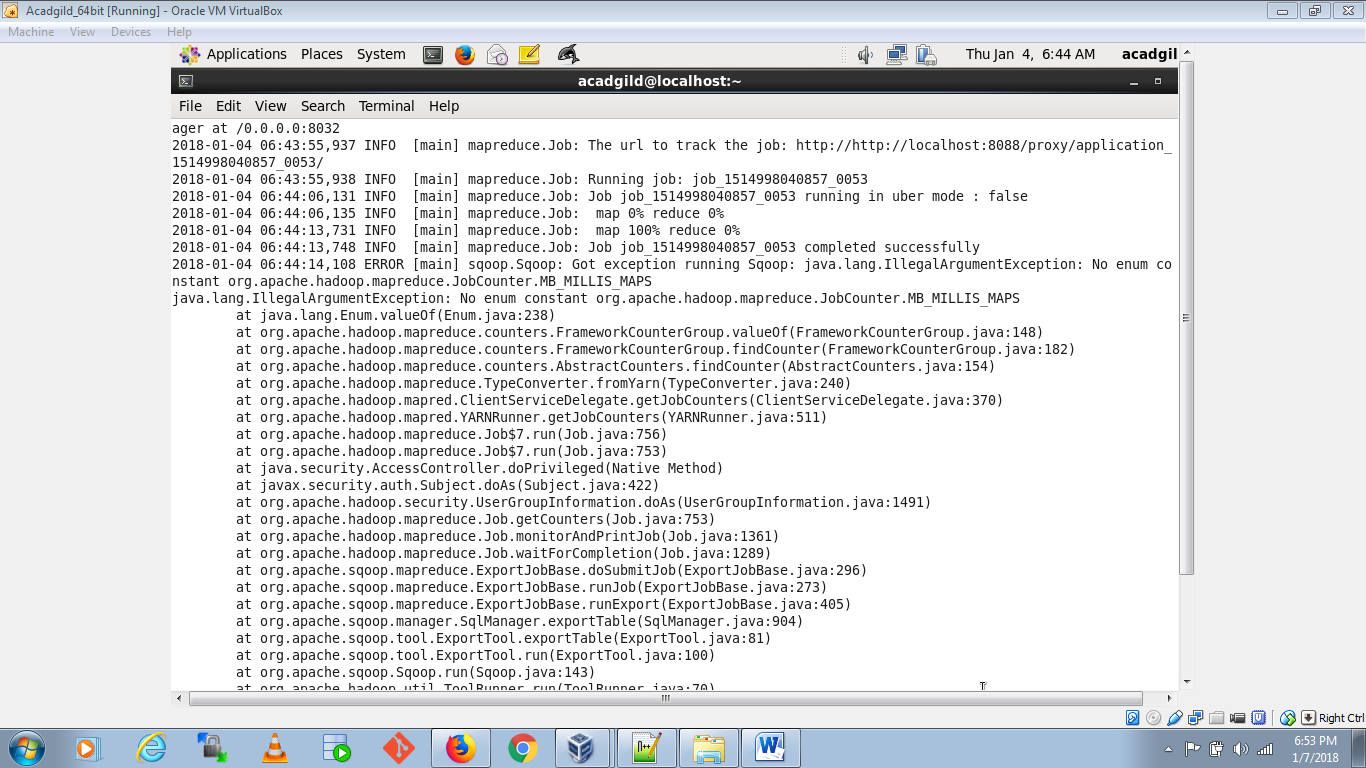
Create MySql table:



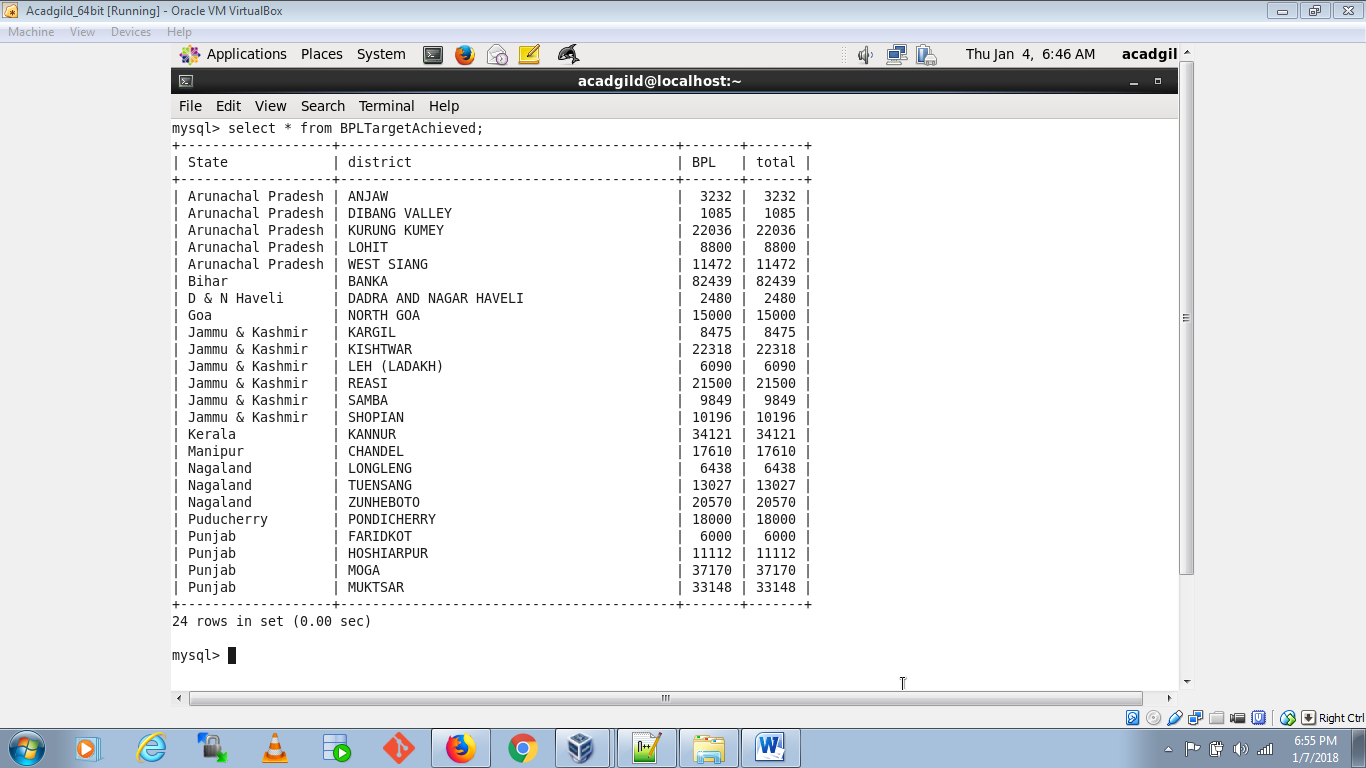
Export the results to mysql using sqoop



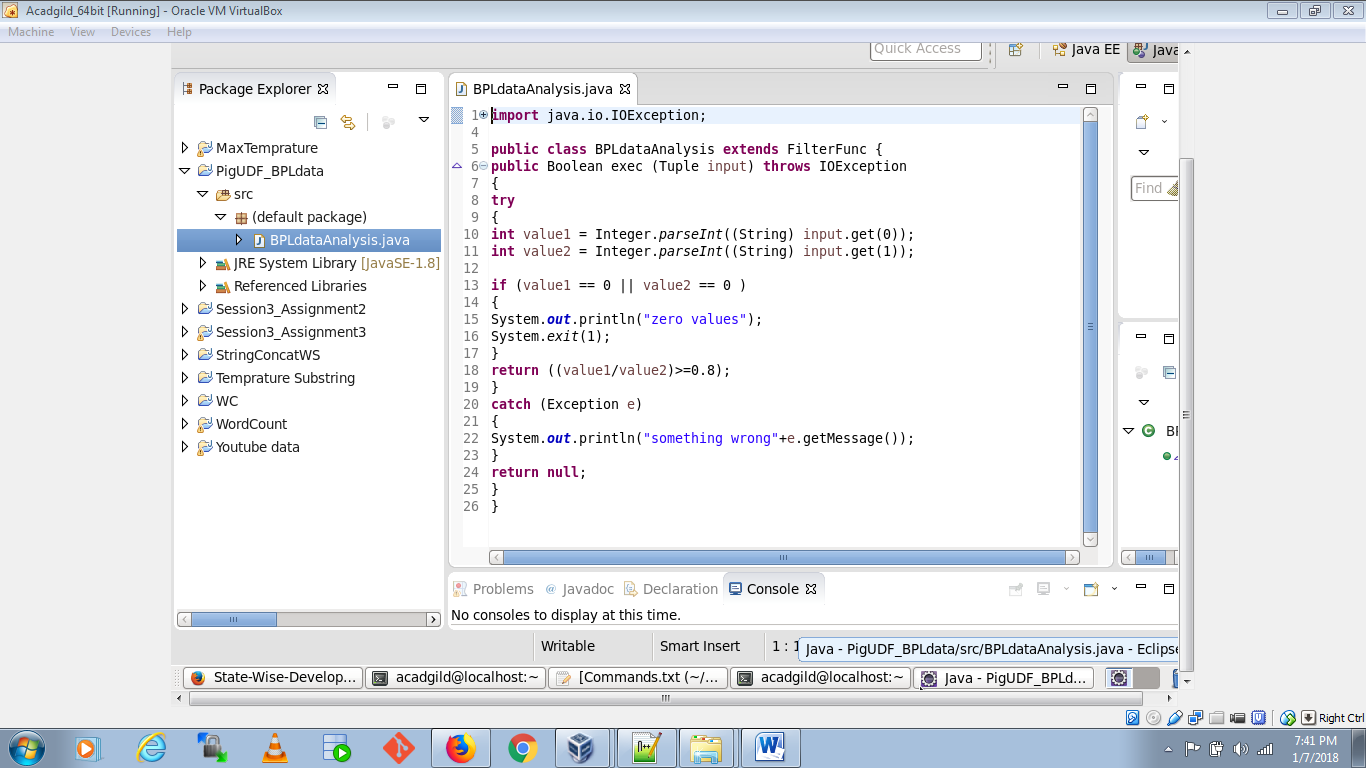




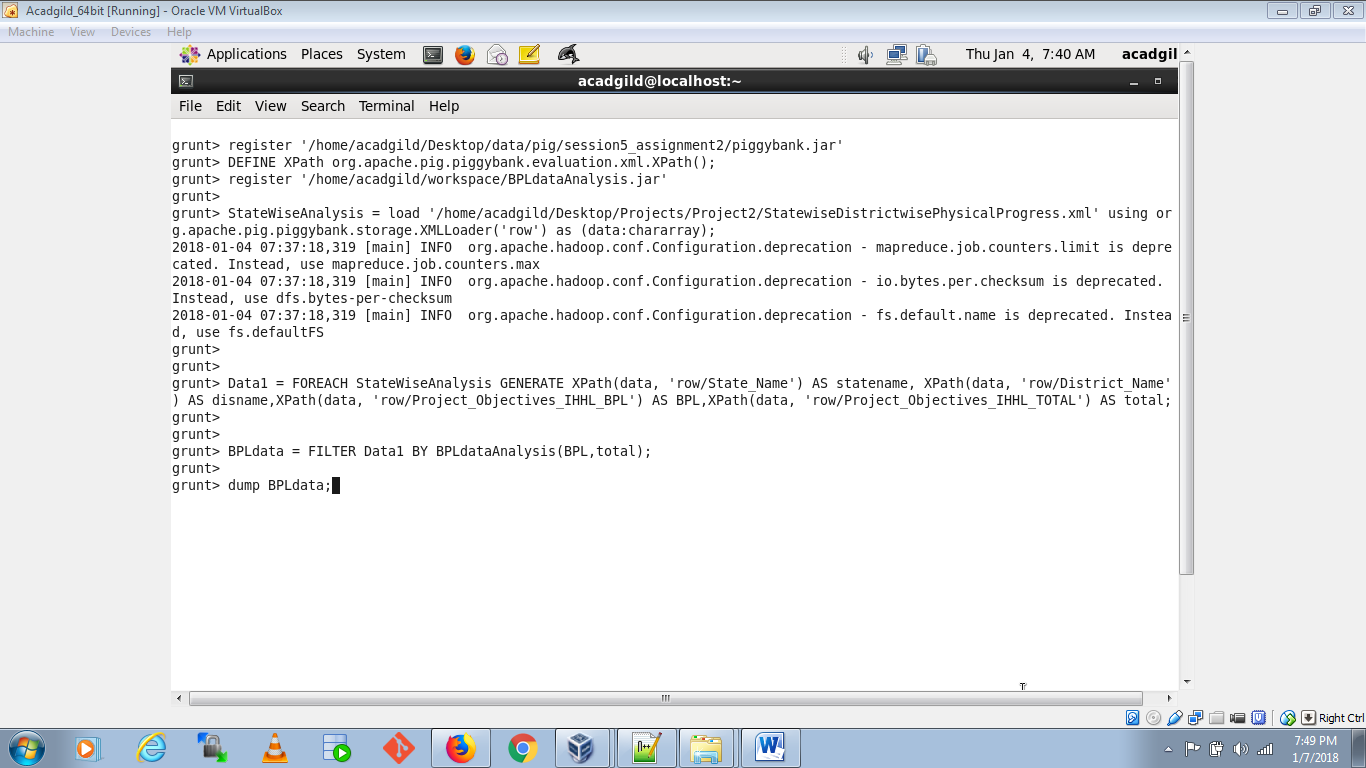
Verify data loaded into MySql table:

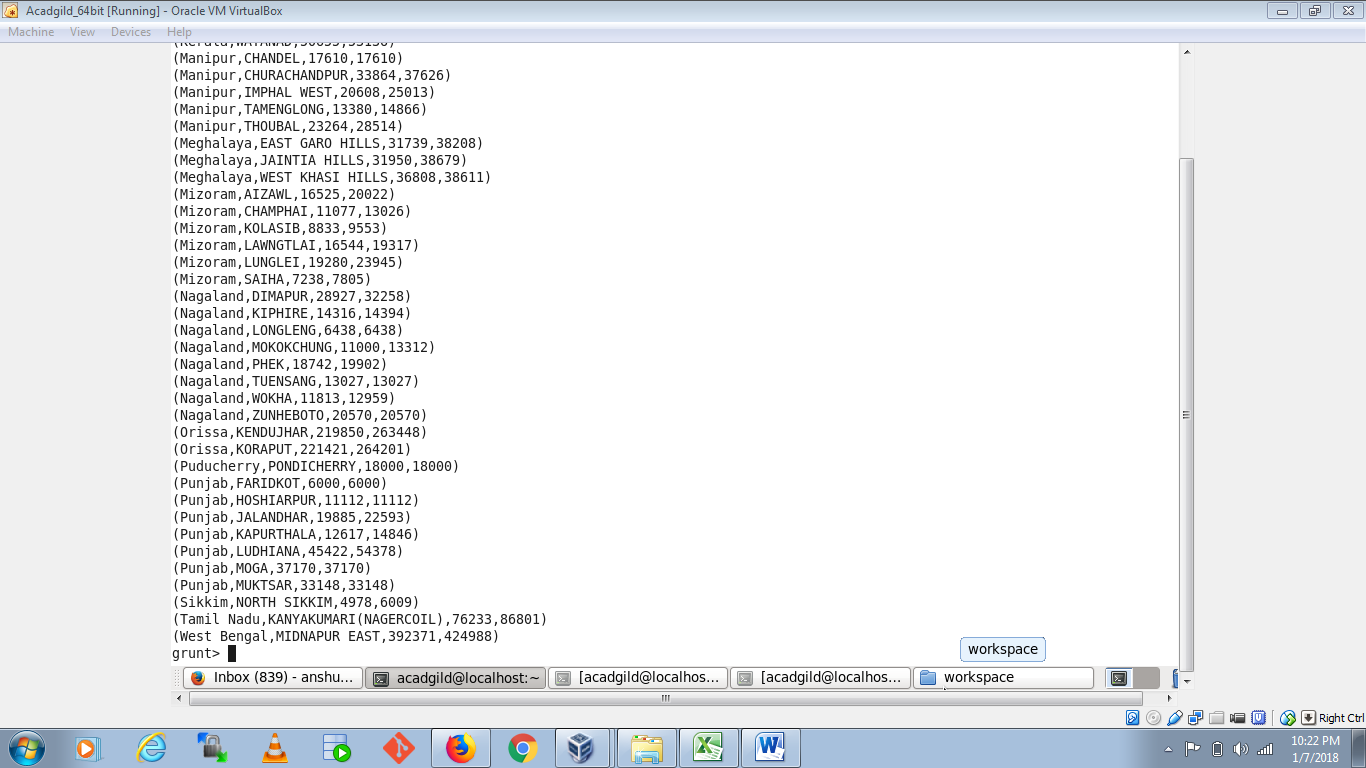


1. **Write a Pig UDF to filter the districts which have reached 80% of objectives of BPL cards.**

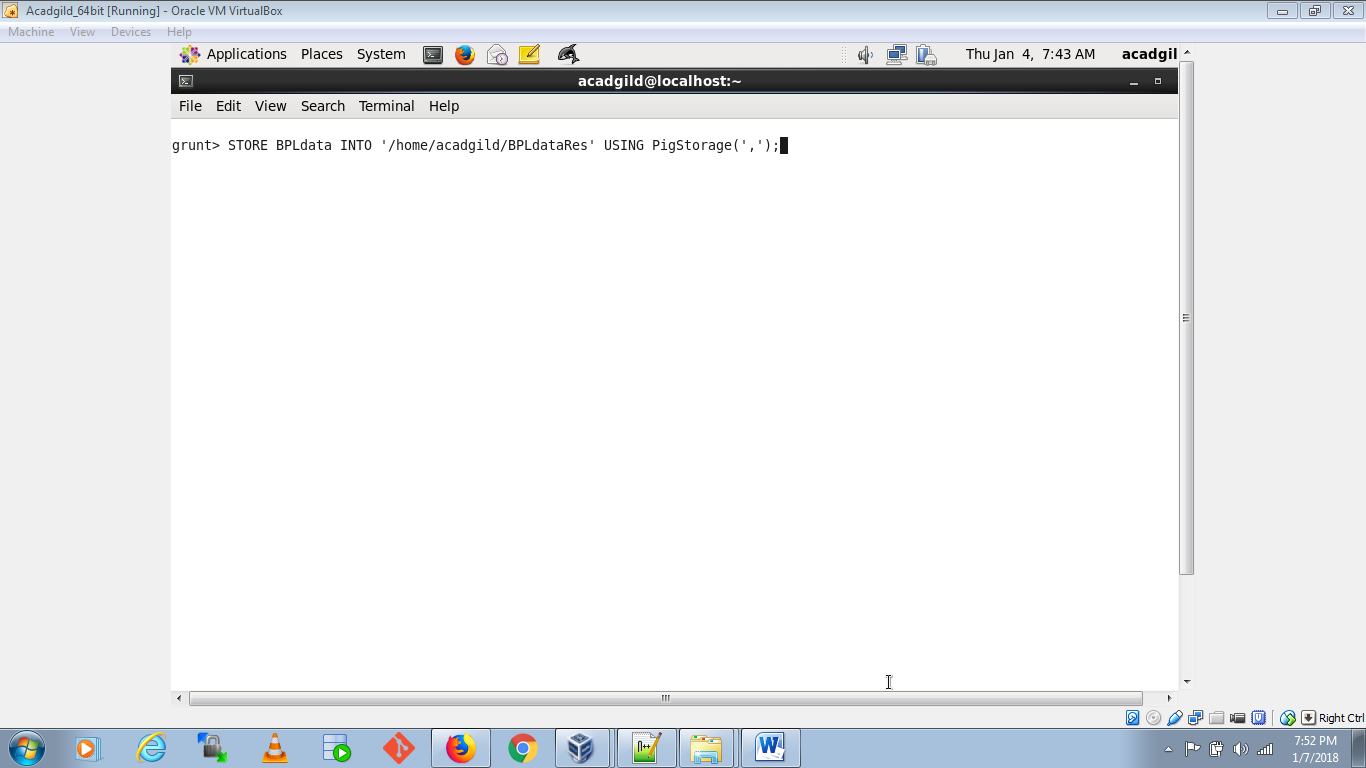


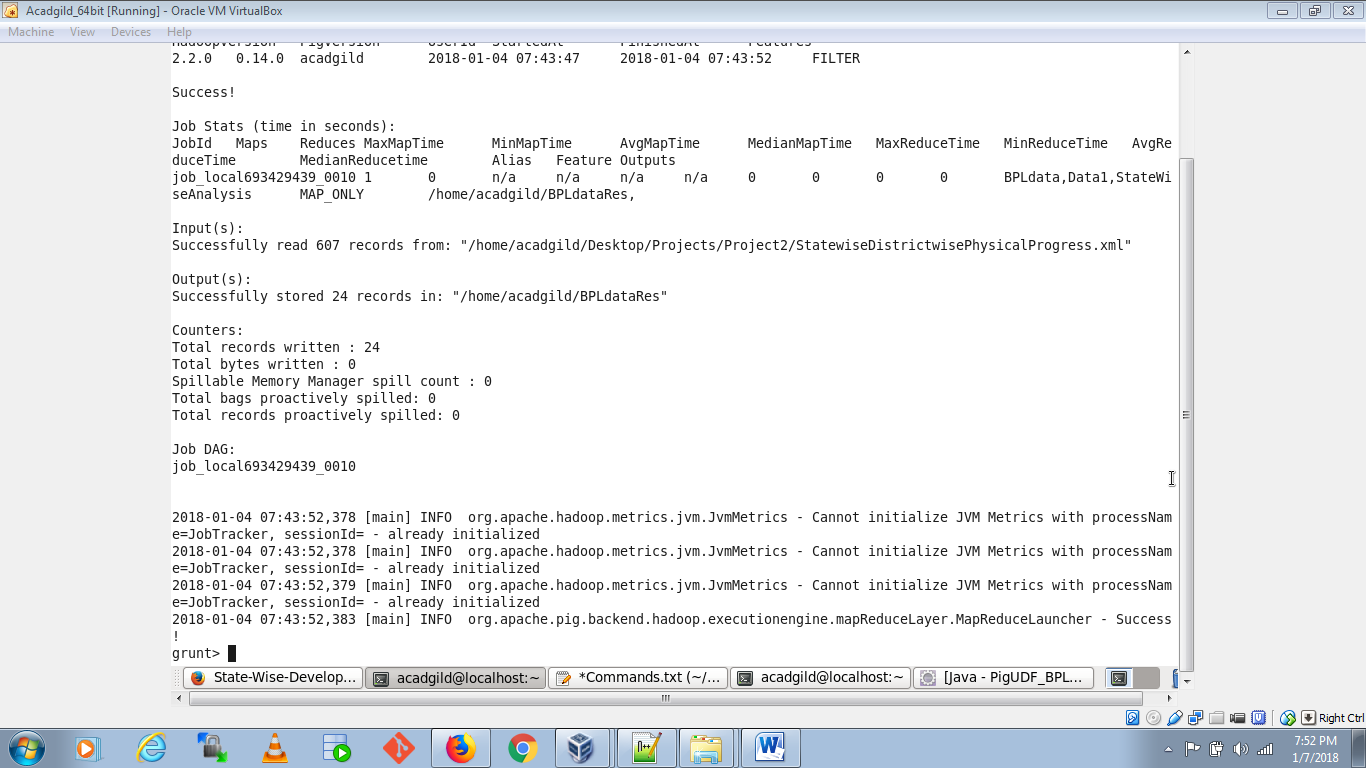
Export jar from above code. Register on grunt shell. And run filter on pig data:

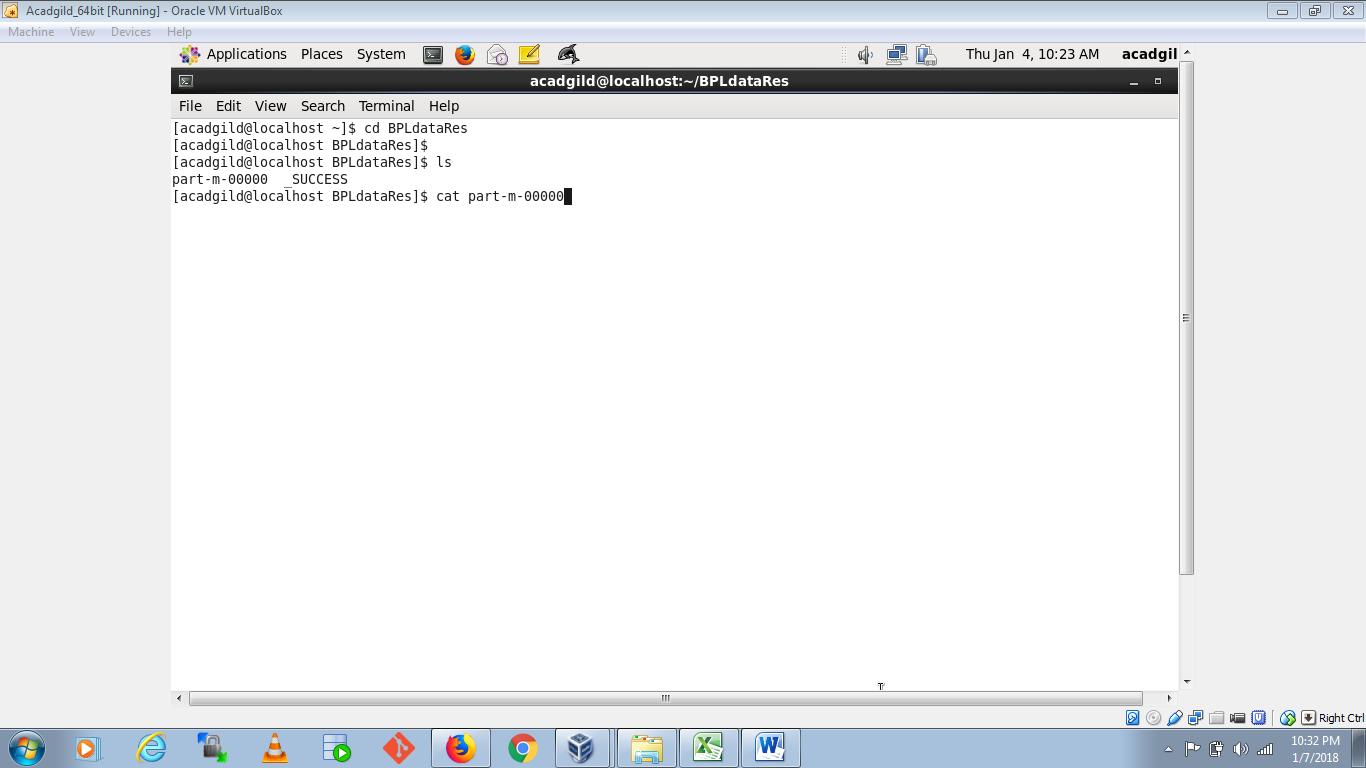


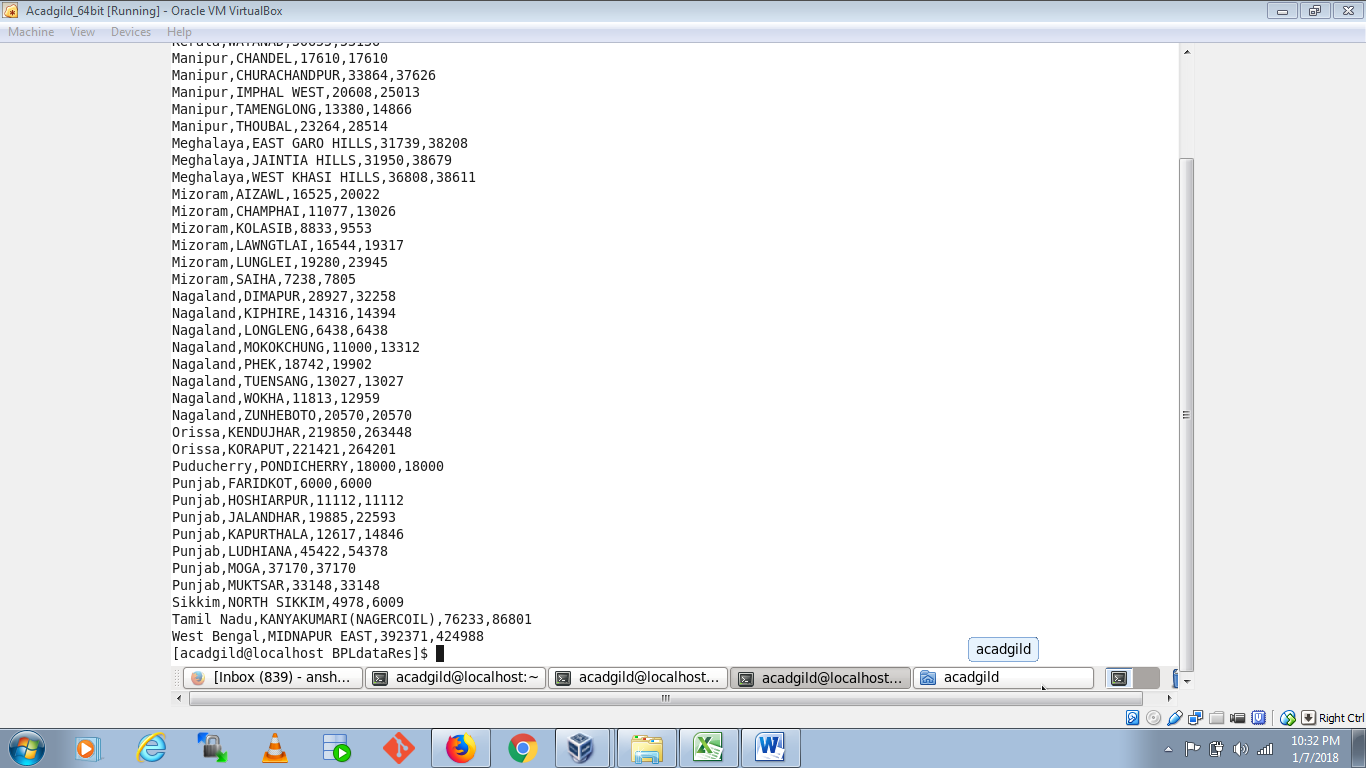


Export Result data from PIG shell to file system:

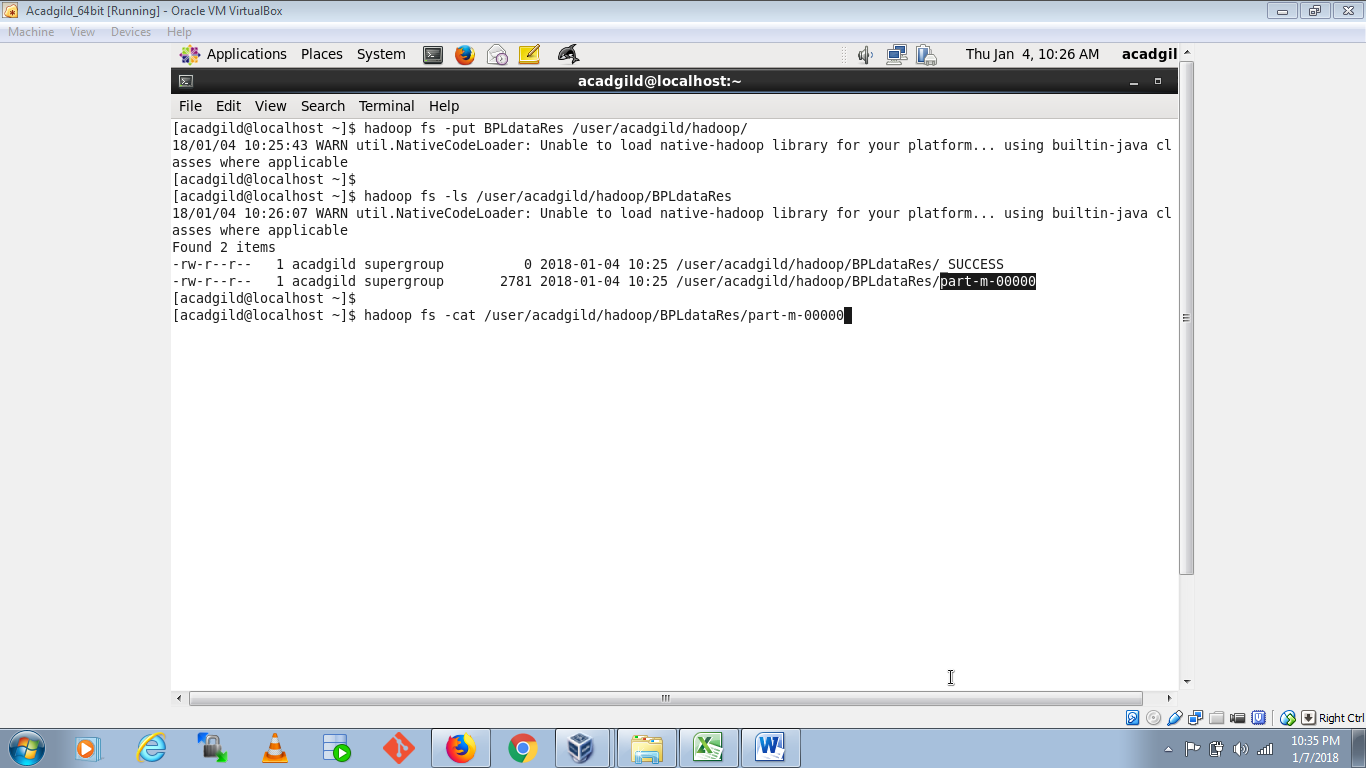


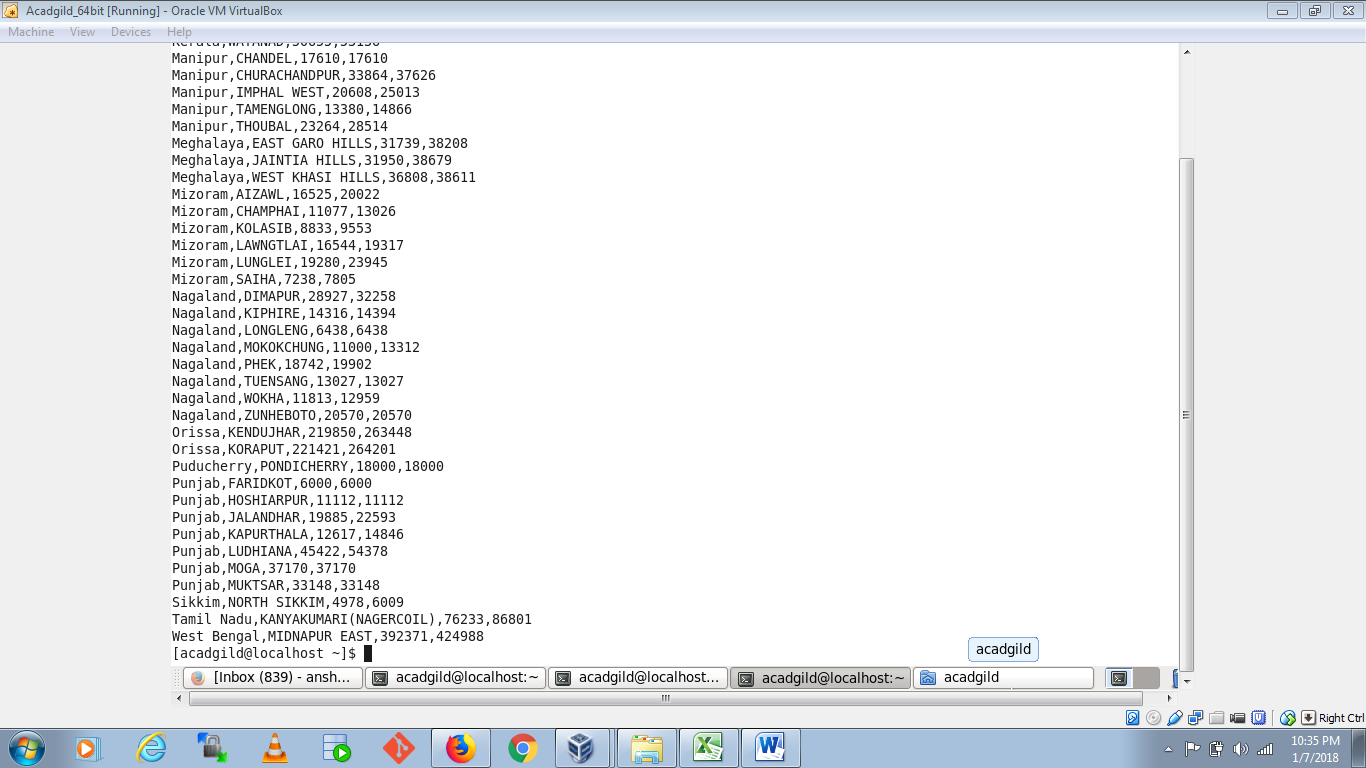




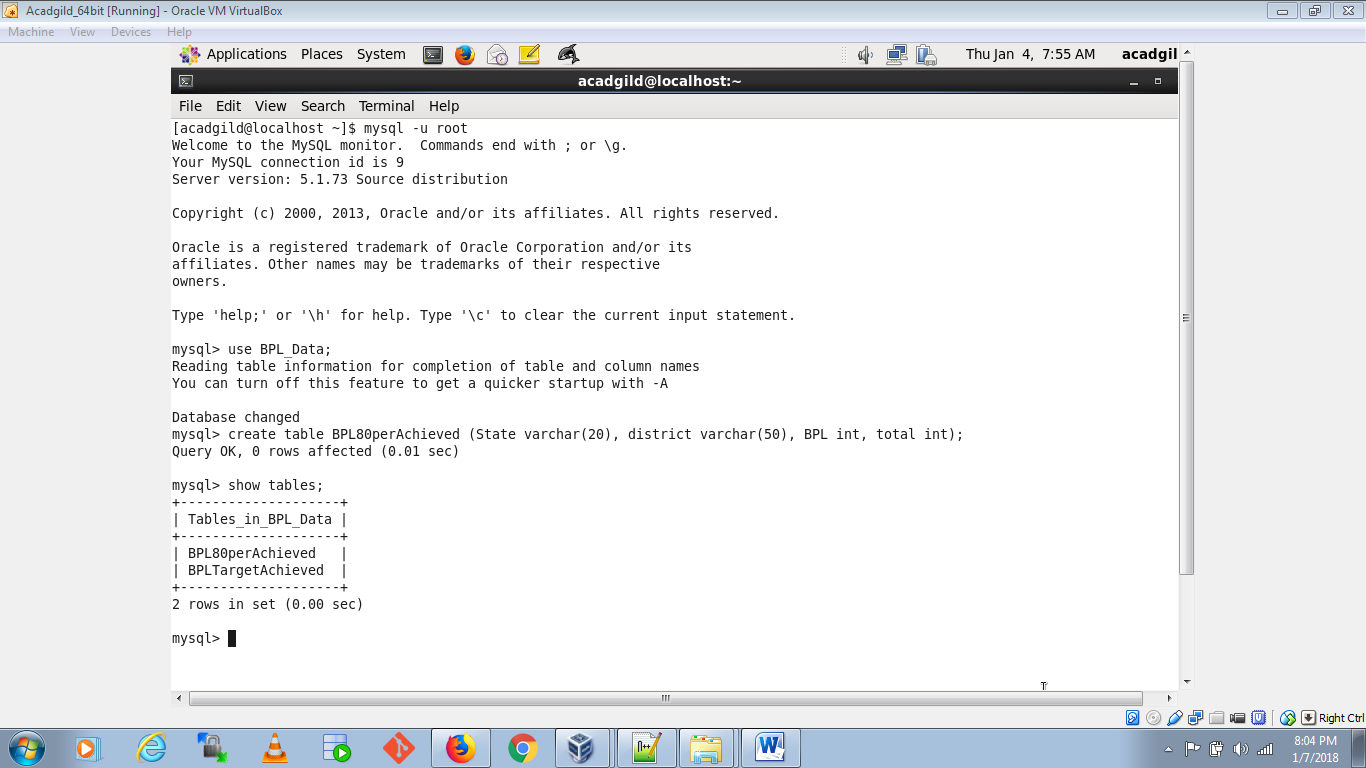


Copy Result data from file system to HDFS:

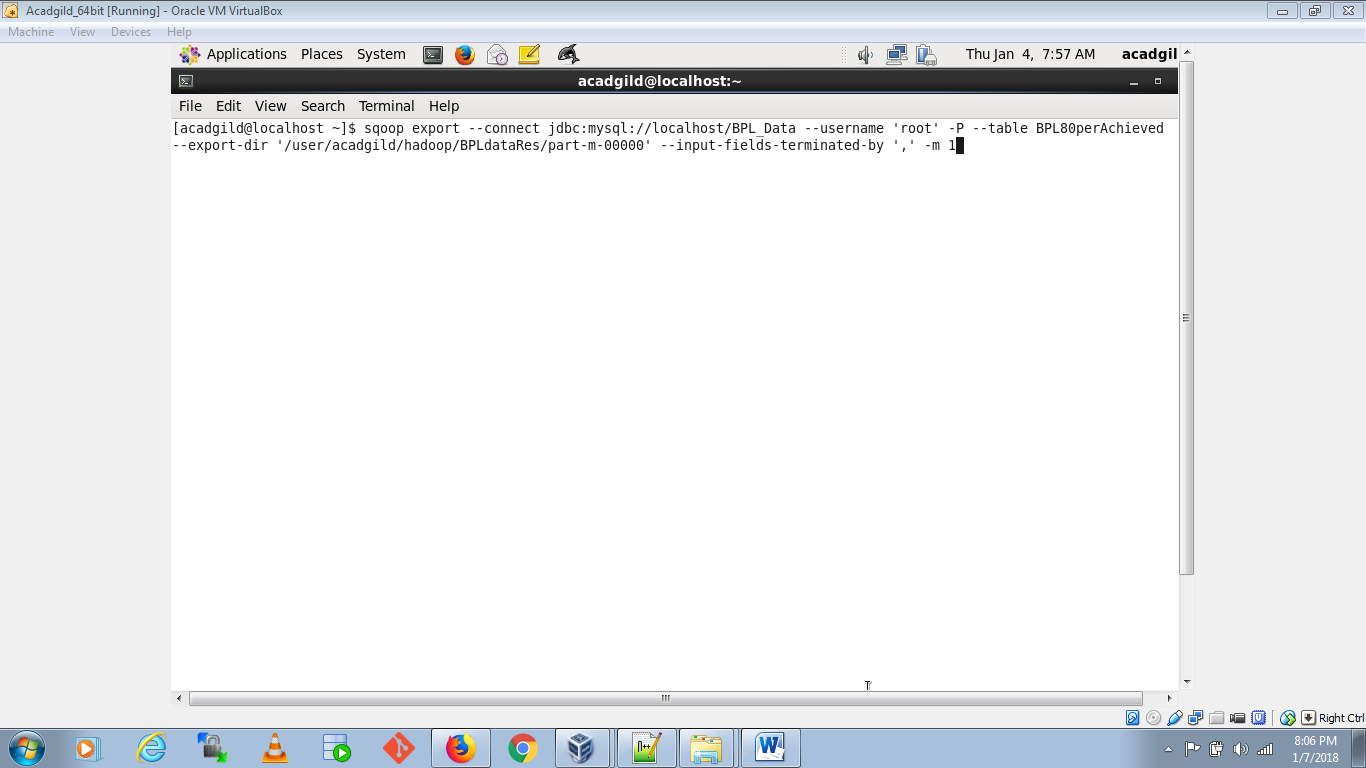


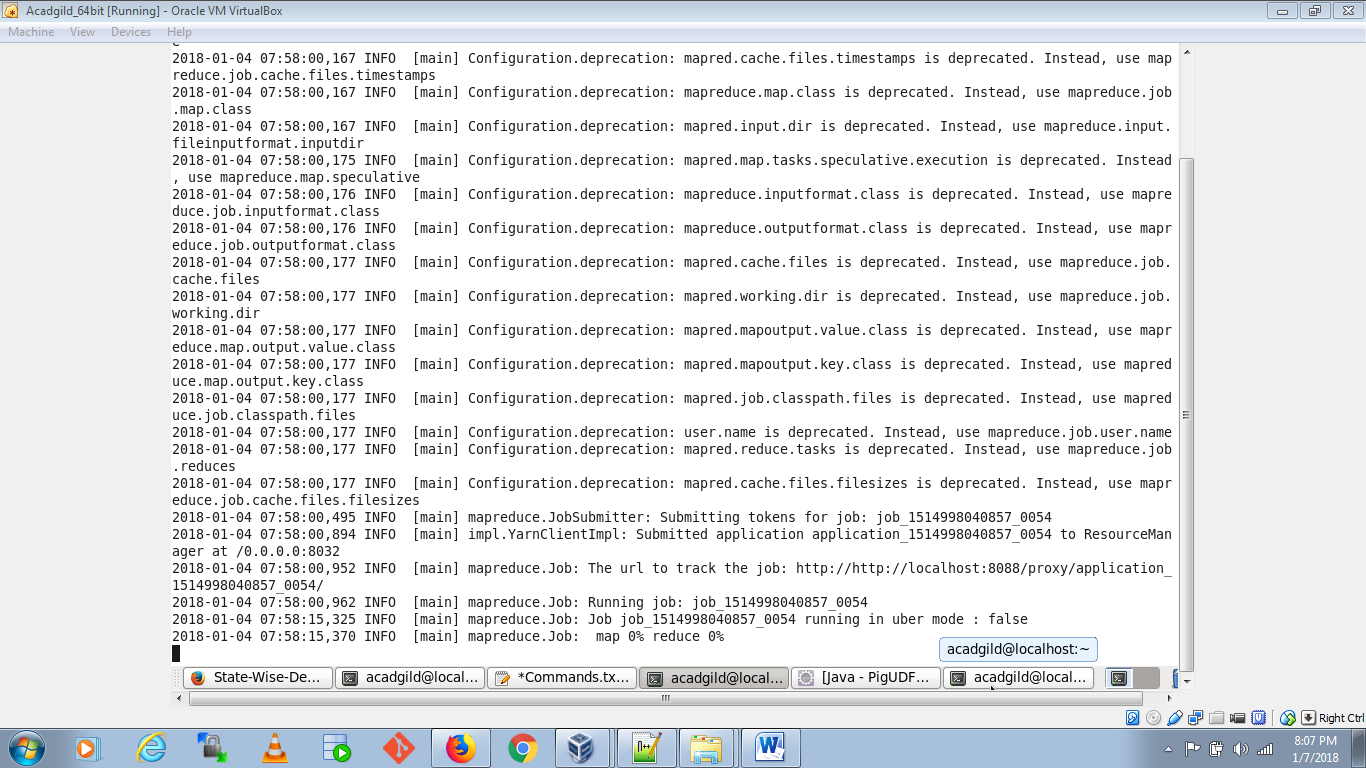


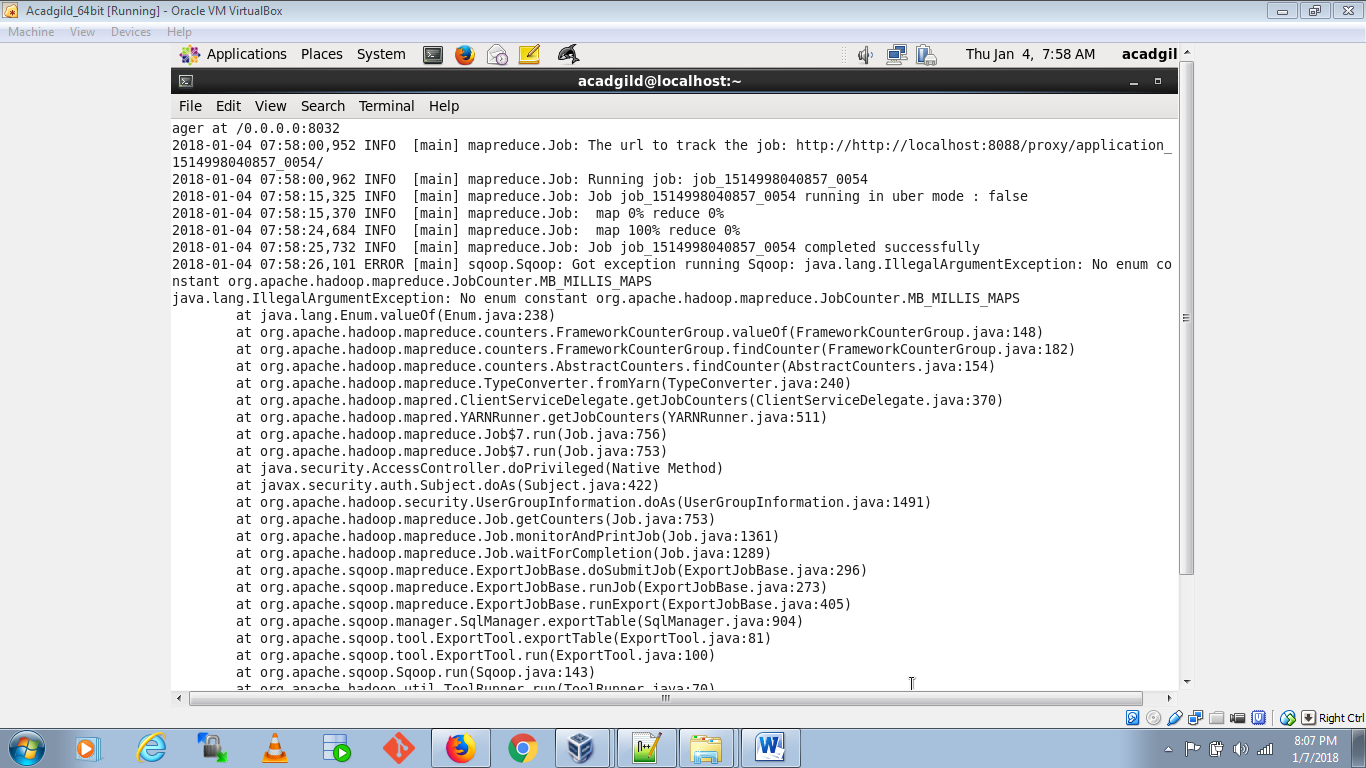
Create MySql table:



Export the results to mysql using sqoop







Verify data loaded into MySql table:

