*Project 2 - USA Consumer Forum Data Analysis*

* Copy of the dataset into HDFS using Flume:

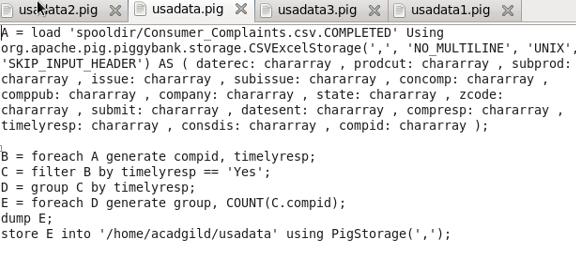


Output: Consumer\_Complaints.csv.COMPLETED



* Now Problem Part:

1. Pig script to find no of complaints which got timely response :



Output:

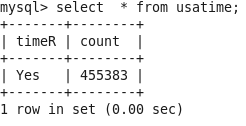


Output is stored and file part-r-00000 is generated:

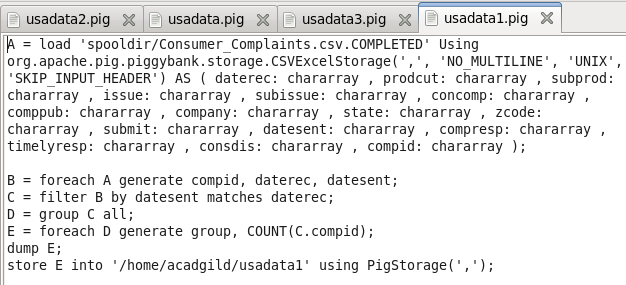


Storing the result into RDBMS using Sqoop:





1. Pig script to find no of complaints where consumer forum forwarded the complaint same day they received to respective company :



Output:

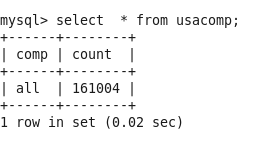


Output generated after storing:

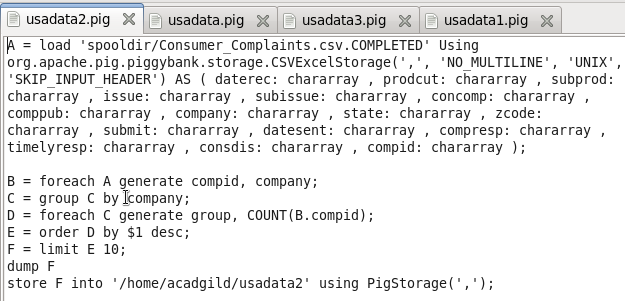


Storing the result into RDBMS using Sqoop:

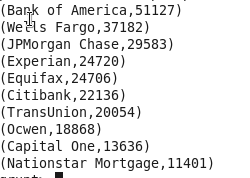




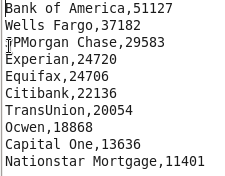
1. Pig script to find list of companies toping in complaint chart (companies with maximum number of complaints) :



Output:

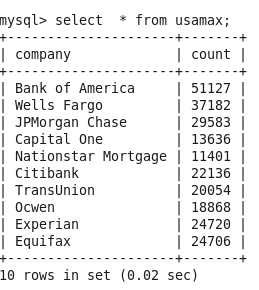


Output generated as part-r-00000 after storing:

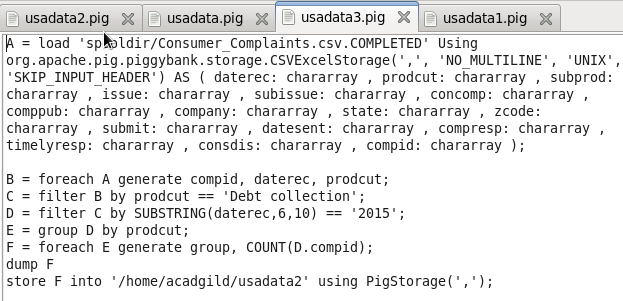


Storing the result into RDBMS using Sqoop:





1. Pig script to find no of complaints filed with product type has &quot;Debt collection&quot; for the year 2015 :



Output:



Output generated as part-r-00000 after storing:



Storing the result into RDBMS using Sqoop:



