1. create a file named

vi datewise_bookings_aggregates_spark.py

2. execute the .py file using following command

spark2-submit --jars spark-sql-kafka-0-10_2.11-2.3.0.jar datewise_bookings_aggregates_spark.py

3. execute the .py file using following command

hdfs dfs -cp /home/datewise_aggreation /user/root/datewise_aggreation_new

4. see no of records and validate against validation doc

hadoop fs -cat
/user/root/datewise_aggreation_new/part-00000-9d97e6f7-e258-41eb-ad5c-1e5cc498064e-c00
0.csv | wc -l