

CREATING .py FILE TO WRITE THE CODE (ec2-user)  
vi spark-streaming.py

SUBMIT COMMAND USED TO EXECUTE THE spark-streaming.py CODE  
spark2-submit --jars spark-sql-kafka-0-10\_2.11-2.3.0.jar.1 spark-streaming.py "18.211.252.152"  
"9092" "real-time-project"

SERVER = "18.211.252.152"  
PORT NUMBER = "9092"  
TOPIC NAME = "real-time-project"

---

CHECKING IF THE JSON GOT REFLECTED INTO THE HDFS  
sudo -i  
su - hdfs

CHECKING TimeAndCountryBasedKPI1  
hadoop fs -ls /tmp/TimeAndCountryBasedKPI1

CHECKING TimeBasedKPI1  
hadoop fs -ls /tmp/TimeBasedKPI1

---FETCHING JSON FROM HDFS TO EC2---  
CREATING A DIRECTORY TO FETCH 'Country-and-timebased-KPI' JSON FILES  
mkdir Country-and-timebased-KPI

FETCHING JSON FILES TO EC2 DIRECTORY  
hadoop fs -get /tmp/TimeAndCountryBasedKPI1/ Country-and-timebased-KPI

CREATING A DIRECTORY TO FETCH 'Timebased-KPI' JSON FILES  
mkdir Timebased-KPI

FETCHING 'Timebased-KPI' JSON FILES TO EC2 DIRECTORY  
hadoop fs -get /tmp/TimeBasedKPI1/ Timebased-KPI