Code Logic - Retail Data Analysis

Below steps are taken-

1. Order System publish records to kafka topics.
2. Spark streaming jobs read data from kafka and apply logics and stored data in sinks.
3. 3 types of outputs-
4. Aggregated inputs ( 1 min)
5. Time based kpi (10 mins)
6. Country and time based kpi (10 mins)

Spark command-

spark-submit --packages org.apache.spark:spark-sql-kafka-0-10\_2.11:2.4.5 \ spark-streaming.py