

```
1  STEPS FOR EXECUTION:
2  The execution of the program was done through Eclipse
3  1. The program needs 3 arguments for execution. We perform the check if the parameters
   have been provided.
4      a. Topic Name = stockData
5      b. Group Id = test_group
6      c. bootstrap server = 52.55.237.11:9092
7  These were provided in the Run As =>Run Configuration=> Under the Argument tab.
8
9  2. Then the spark context was setup followed by the JavaStreamingContext setting up
   the duration of poll to 1minute. This is the crux of Spark Streaming i.e. defining a
   micro batch.
10
11  3. Create instance for properties to access producer configs. Define a new HashMap for
   holding the Kafka information
12
13  4. Create direct Kafka stream with brokers and topics. LocationStrategy would allow
   partitions to be distributed consistently to the spark executors. ConsumerStrategy
   allows to subscribe to the Kafka topic.
14
15  5. Define the JavaInputDStream which would be a continuous input stream associated to
   the source. Create JavaInputDStream from kafka topic using kafkaUtils class.
16
17  6. Convert javaInputDStream into JavaDStream using appropriate map function.
18
19  7. Cache the output of previous point output.
20
21  8. Print the raw stock price batch coming from Kafka Topic.
22
23  9. Using the flapMaptoPair Function, transform the JavaDStream to
   JavaPairDStream<String, StockAverageTuple> where StockAverageTuple is java pojo of stock.
24
25  10. Apply reduceByKeyAndWindow function on output of previous input and calculate the
   sum of profit(closePrice- openPrice) , sum of tradingVolume, MovingAverage
```