

# STREAM DATA INGESTION AND PROCESSING WORKSHOP

## TEAM MEMBERS

1. JAYARAJA S.K. – [jayaraja19041@cse.ssn.edu.in](mailto:jayaraja19041@cse.ssn.edu.in)
2. PRAVINKRISHNAN K – [pravinkrishnan19082@cse.ssn.edu.in](mailto:pravinkrishnan19082@cse.ssn.edu.in)
3. SAMYUKTHA G – [samyuktha19095@cse.ssn.edu.in](mailto:samyuktha19095@cse.ssn.edu.in)

## PROGRAMMING LANGUAGES USED

1. JAVA
2. SQL

## TECHNOLOGIES USED

1. SPARK
2. KAFKA

## GITHUB REPOSITORY LINK FOR THE PROJECT

<https://github.com/Jayaraja-SK/Stream-Data-Ingestion-and-Processing-Workshop>

## USE CASE

A Car Rental company which rents their electric cars of different variants to people and runs business in more than 100 countries

## FIELDS

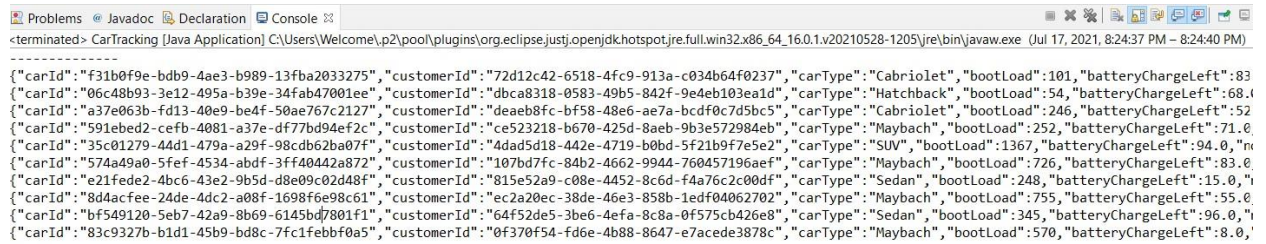
1. carId
2. customerId
3. carType – Randomly chosen from the following:
  - a. SUV
  - b. Sedan
  - c. Hatchback
  - d. Cabriolet
  - e. Maybach
  - f. Van
4. Speed
5. BootLoad
6. noOfPassengers – No. of passengers travelling
7. batteryChargeLeft – Charge left in the car
8. Time – The time at which the data is recorded
9. Location – In terms of latitude and longitude

## FILES

1. CarInfo – Contains the fields described above and their respective getter-setter methods
2. Functions – Contains the static functions to generate random data and is imported by all main functions
3. CarTracking – Main java function to generate random data and print the same to the console
4. CarTrackingToFile – Main java function to generate random data and save the same to different .txt files under a directory called as CarDataJson
5. CarTrackingSingleBroker – Main java function to generate random data by producer and send the same to a single consumer/broker
6. CarTrackingMultiBroker – Main java function to generate random data by producer and send the same to multiple consumer's/broker's
7. CarSparkStreaming – Main java function that gets data from Kafka Consumer's and sends the same to Spark UI
8. CarSparkStreamingSQL – Main java function that gets data from Kafka Consumer's and sends the same to Spark UI where the data is stored in a directory called CarDataSpark and could be later queried in Spark-Shell

## OUTPUT

### 1. PRINTING RANDOM DATA TO CONSOLE



The screenshot shows an IDE console window with the following content:

```
Problems Javadoc Declaration Console
terminated> CarTracking [Java Application] C:\Users\Welcome\p2\pool\plugins\org.eclipse.justj.openjdk.hotspot.jre.full.win32.x86_64_16.0.1.v20210528-1205\jre\bin\javaw.exe (Jul 17, 2021, 8:24:37 PM - 8:24:40 PM)
-----
{"carId":"f31b0f9e-bdb9-4ae3-b989-13fba2033275","customerId":"72d12c42-6518-4fc9-913a-c034b64f0237","carType":"Cabriolet","bootLoad":101,"batteryChargeLeft":83}
{"carId":"06c48b93-3e12-495a-b39e-34fab47001ee","customerId":"dbca8318-0583-49b5-842f-9e4eb103ea1d","carType":"Hatchback","bootLoad":54,"batteryChargeLeft":68.0}
{"carId":"a37e063b-fd13-40e9-be4f-50ae767c2127","customerId":"deaeb8fc-bf58-48e6-ae7a-bcdf0c7d5bc5","carType":"Cabriolet","bootLoad":246,"batteryChargeLeft":52}
{"carId":"591ebed2-cefb-4081-a37e-df77bd94ef2c","customerId":"ce523218-b670-425d-8aeb-9b3e572984eb","carType":"Maybach","bootLoad":252,"batteryChargeLeft":71.0}
{"carId":"35c01279-44d1-479a-a29f-98cdb62ba07f","customerId":"4dad5d18-442e-4719-b0bd-5f21b9f7e5e2","carType":"SUV","bootLoad":1367,"batteryChargeLeft":94.0,"n
{"carId":"574a49a0-5fef-4534-abdf-3ff40442a872","customerId":"107bd7fc-84b2-4662-994d-760457196aef","carType":"Maybach","bootLoad":726,"batteryChargeLeft":83.0}
{"carId":"e21fede2-4bc6-43e2-9b5d-d8e09c02d48f","customerId":"815e52a9-c08e-4452-8c6d-f4a76c2c00df","carType":"Sedan","bootLoad":248,"batteryChargeLeft":15.0,"i
{"carId":"8d4acfee-24de-4dc2-a08f-1698f6e98c61","customerId":"ec2a20ec-38de-46e3-858b-1edf04062702","carType":"Maybach","bootLoad":755,"batteryChargeLeft":55.0}
{"carId":"bf549120-5eb7-42a9-8b69-6145bd7801f1","customerId":"64f52de5-3be6-4efa-8c8a-0f575cb426e8","carType":"Sedan","bootLoad":345,"batteryChargeLeft":96.0,"i
{"carId":"83c9327b-b1d1-45b9-bd8c-7fc1febbf0a5","customerId":"0f370f54-fd6e-4b88-8647-e7acade3878c","carType":"Maybach","bootLoad":570,"batteryChargeLeft":8.0,'
-----
```

### 2. WORKING OF SINGLE BROKER SYSTEM

[https://drive.google.com/file/d/1AN-kGj\\_8DUqMLKAILGhD3Fkx6nM8lKeJ/view?usp=sharing](https://drive.google.com/file/d/1AN-kGj_8DUqMLKAILGhD3Fkx6nM8lKeJ/view?usp=sharing)

### 3. WORKING OF MULTI BROKER SYSTEM

[https://drive.google.com/file/d/19lilcBu3SWtBtx\\_7U0HbwbJMFLbYGI17/view?usp=sharing](https://drive.google.com/file/d/19lilcBu3SWtBtx_7U0HbwbJMFLbYGI17/view?usp=sharing)

### 4. WORKING OF STREAMING DATA TO SPARK AND LATER QUERIED USING SQL

[https://drive.google.com/file/d/1M-qdHXUsanUqodqGWkVGCZ6vAKzHA\\_9L/view?usp=sharing](https://drive.google.com/file/d/1M-qdHXUsanUqodqGWkVGCZ6vAKzHA_9L/view?usp=sharing)