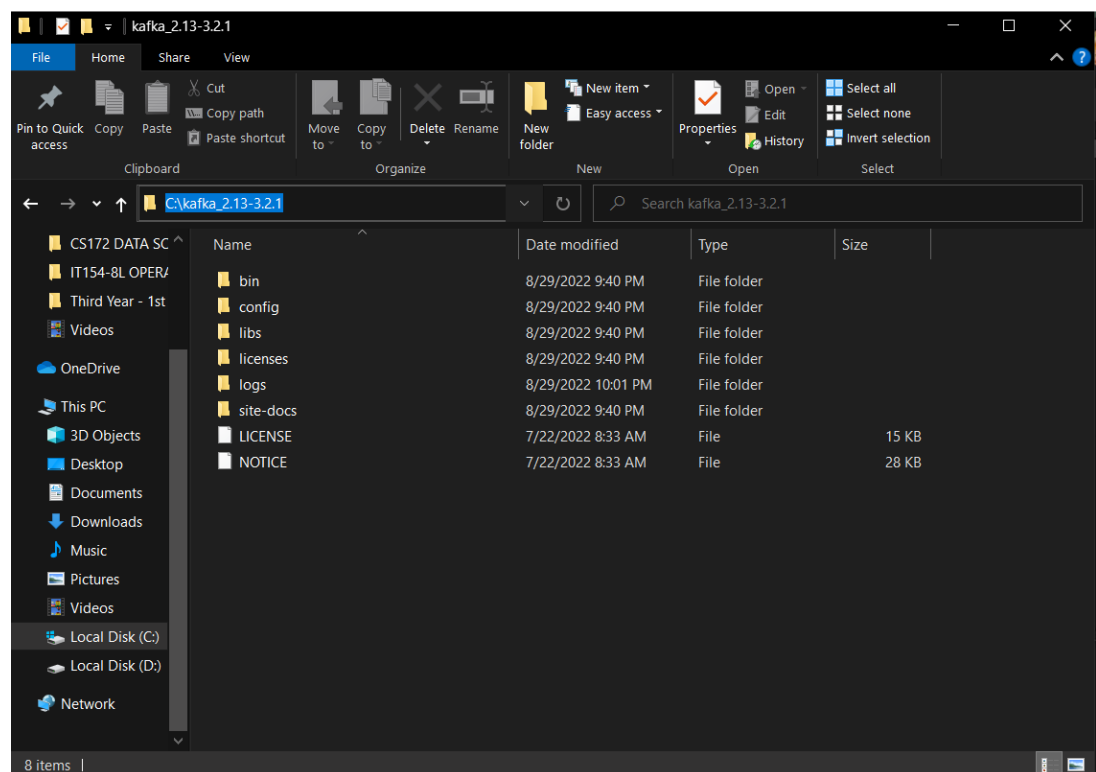


## **Module 1 Project – Part 2: Big Data Streaming Pipeline and Integration Platform**

### **I. Instantiating and Running Ten Instances of Producer through PyCharm and Displaying Them through Consumer using Apache Kafka**

PyCharm and Apache Kafka were both utilized to create instances of the producer and display the data sent by the different instances of the producer through the consumer. Since Apache Kafka had been previously installed, the Kafka environment was run by typing ‘cmd’ on the address bar of the Kafka download in the local disk (C:).



Running the Kafka environment required the following two services to be run before performing any commands in the environment itself: *Zookeeper* and *Kafka broker services*. (These steps have been discussed in further detail in a previous document, but they will be briefly discussed in this documentation process for the sake of instructional completeness.) The figures below indicate the commands and running processes of the prerequisite services of the Kafka environment.

```
.\bin\windows\zookeeper-server-start.bat
.\config\zookeeper.properties
```

[illegible]

## Zookeeper

In another command prompt terminal, which would be opened by going back to the Kafka download's location in the file explorer and typing 'cmd' in the address bar, as mentioned previously, the following command was entered to start the Kafka broker service:

```
.\bin\windows\kafka-server-start.bat
.\config\server.properties
```

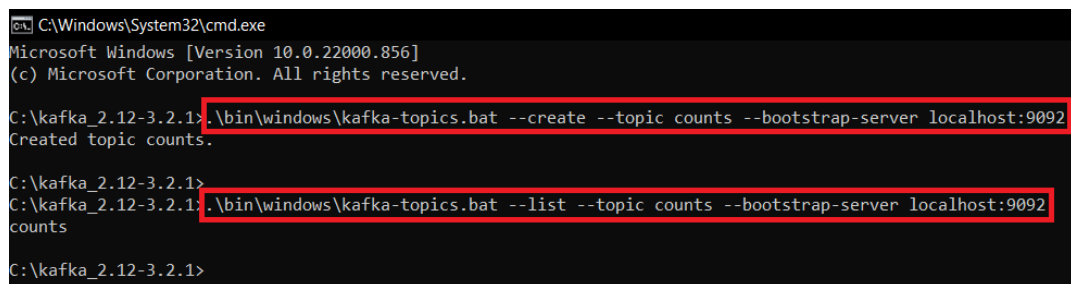
[illegible]

## Kafka broker services

Once both prerequisite services have been run, another command prompt terminal was opened to create the topic “counts.” (The topic was labeled “counts” because it is the default topic in the producer code.) The same topic was also included into a list, as indicated in the figure below. The commands used were the following:

```
.\bin\windows\kafka-topics.bat --create --topic counts --  
bootstrap-server localhost:9092
```

```
.\bin\windows\kafka-topics.bat --list --topic counts --  
bootstrap-server localhost:9092
```

A screenshot of a Windows command prompt window. The title bar reads 'C:\Windows\System32\cmd.exe'. The window content shows the following text: 'Microsoft Windows [Version 10.0.22000.856] (c) Microsoft Corporation. All rights reserved. C:\kafka\_2.12-3.2.1>.\bin\windows\kafka-topics.bat --create --topic counts --bootstrap-server localhost:9092 Created topic counts. C:\kafka\_2.12-3.2.1>C:\kafka\_2.12-3.2.1>.\bin\windows\kafka-topics.bat --list --topic counts --bootstrap-server localhost:9092 counts C:\kafka\_2.12-3.2.1>'. The two command lines are highlighted with red rectangular boxes.

Now that the necessary configurations for the Kafka environment have been made, the producer code should be ready to be run on PyCharm.

At least **ten instances** of the producer code were required, with each **sensor\_id** differing from sensor\_01 to sensor\_10:

```
# send data  
message["timeuuid_id"] = str(time_uuid.utctime())  
message["lgu_code"] = '1200'  
message["sensor_id"] = 'sensor_01'  
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))  
message["time_saved"] = str(date_today.strftime("%X"))  
message["total"] = total  
message["car"] = car  
message["bus"] = bus  
message["truck"] = truck  
message["jeepney"] = jeepney  
message["bike"] = bike  
message["tryke"] = tryke  
message["others"] = others
```

```
# send data
message["timeuuid_id"] = str(time_uuid.utctime())
message["lgu_code"] = '1200'
message["sensor_id"] = 'sensor_02'
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))
message["time_saved"] = str(date_today.strftime("%X"))
message["total"] = total
message["car"] = car
message["bus"] = bus
message["truck"] = truck
message["jeepney"] = jeepney
message["bike"] = bike
message["tryke"] = tryke
message["others"] = others
```

```
# send data
message["timeuuid_id"] = str(time_uuid.utctime())
message["lgu_code"] = '1200'
message["sensor_id"] = 'sensor_03'
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))
message["time_saved"] = str(date_today.strftime("%X"))
message["total"] = total
message["car"] = car
message["bus"] = bus
message["truck"] = truck
message["jeepney"] = jeepney
message["bike"] = bike
message["tryke"] = tryke
message["others"] = others
```

```
# send data
message["timeuuid_id"] = str(time_uuid.utctime())
message["lgu_code"] = '1200'
message["sensor_id"] = 'sensor_04'
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))
message["time_saved"] = str(date_today.strftime("%X"))
message["total"] = total
message["car"] = car
message["bus"] = bus
message["truck"] = truck
message["jeepney"] = jeepney
message["bike"] = bike
message["tryke"] = tryke
message["others"] = others
```

```
# send data
message["timeuuid_id"] = str(time_uuid.utctime())
```

```
message["lgu_code"] = '1200'  
message["sensor_id"] = 'sensor_05'  
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))  
message["time_saved"] = str(date_today.strftime("%X"))  
message["total"] = total  
message["car"] = car  
message["bus"] = bus  
message["truck"] = truck  
message["jeepney"] = jeepney  
message["bike"] = bike  
message["tryke"] = tryke  
message["others"] = others
```

```
# send data
```

```
message["timeuuid_id"] = str(time_uuid.utctime())  
message["lgu_code"] = '1200'  
message["sensor_id"] = 'sensor_06'  
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))  
message["time_saved"] = str(date_today.strftime("%X"))  
message["total"] = total  
message["car"] = car  
message["bus"] = bus  
message["truck"] = truck  
message["jeepney"] = jeepney  
message["bike"] = bike  
message["tryke"] = tryke  
message["others"] = others
```

```
# send data
```

```
message["timeuuid_id"] = str(time_uuid.utctime())  
message["lgu_code"] = '1200'  
message["sensor_id"] = 'sensor_07'  
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))  
message["time_saved"] = str(date_today.strftime("%X"))  
message["total"] = total  
message["car"] = car  
message["bus"] = bus  
message["truck"] = truck  
message["jeepney"] = jeepney  
message["bike"] = bike  
message["tryke"] = tryke  
message["others"] = others
```

```
# send data
```

```
message["timeuuid_id"] = str(time_uuid.utctime())  
message["lgu_code"] = '1200'  
message["sensor_id"] = 'sensor_08'
```

```

message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))
message["time_saved"] = str(date_today.strftime("%X"))
message["total"] = total
message["car"] = car
message["bus"] = bus
message["truck"] = truck
message["jeepney"] = jeepney
message["bike"] = bike
message["tryke"] = tryke
message["others"] = others

```

*# send data*

```

message["timeuuid_id"] = str(time_uuid.utctime())
message["lgc_code"] = '1200'
message["sensor_id"] = 'sensor_09'
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))
message["time_saved"] = str(date_today.strftime("%X"))
message["total"] = total
message["car"] = car
message["bus"] = bus
message["truck"] = truck
message["jeepney"] = jeepney
message["bike"] = bike
message["tryke"] = tryke
message["others"] = others

```

*# send data*

```

message["timeuuid_id"] = str(time_uuid.utctime())
message["lgc_code"] = '1200'
message["sensor_id"] = 'sensor_10'
message["date_saved"] = str(date_today.strftime('%m/%d/%Y'))
message["time_saved"] = str(date_today.strftime("%X"))
message["total"] = total
message["car"] = car
message["bus"] = bus
message["truck"] = truck
message["jeepney"] = jeepney
message["bike"] = bike
message["tryke"] = tryke
message["others"] = others

```

Each instance of the producer will be run through PyCharm so that the data produced will be sent to the counts topic in Kafka and then displayed through the consumer. If Kafka was configured and the topic created and labeled correctly, the consumer should display the data



that was sent by each producer when it is run simultaneously with the producer, as indicated below.

```
Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663107528.986746', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:48', 'total': 13, 'car': 1, 'bus': 2}
Preparing message: 2
Message: {'timeuuid_id': '1663107529.995914', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:49', 'total': 10, 'car': 1, 'bus': 1}
Preparing message: 3
Message: {'timeuuid_id': '1663107531.0037', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:51', 'total': 13, 'car': 2, 'bus': 2}
Preparing message: 4
Message: {'timeuuid_id': '1663107532.011985', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:52', 'total': 6, 'car': 1, 'bus': 1}
Preparing message: 5
Message: {'timeuuid_id': '1663107533.018433', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:53', 'total': 6, 'car': 3, 'bus': 0}
Preparing message: 6
Message: {'timeuuid_id': '1663107534.029639', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:54', 'total': 10, 'car': 0, 'bus': 1}
Preparing message: 7
Message: {'timeuuid_id': '1663107535.039989', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:55', 'total': 7, 'car': 0, 'bus': 0}
Preparing message: 8
Message: {'timeuuid_id': '1663107536.055275', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:56', 'total': 10, 'car': 1, 'bus': 2}
Preparing message: 9
Message: {'timeuuid_id': '1663107537.070778', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:57', 'total': 11, 'car': 3, 'bus': 1}
Preparing message: 10
Message: {'timeuuid_id': '1663107538.08569', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'time_saved': '06:18:58', 'total': 12, 'car': 4, 'bus': 0}

Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663108533.242835', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:33', 'total': 14, 'car': 3, 'bus': 0}
Preparing message: 2
Message: {'timeuuid_id': '1663108534.258181', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:34', 'total': 13, 'car': 4, 'bus': 0}
Preparing message: 3
Message: {'timeuuid_id': '1663108535.273544', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:35', 'total': 10, 'car': 1, 'bus': 0}
Preparing message: 4
Message: {'timeuuid_id': '1663108536.288965', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:36', 'total': 13, 'car': 4, 'bus': 0}
Preparing message: 5
Message: {'timeuuid_id': '1663108537.303727', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:37', 'total': 10, 'car': 4, 'bus': 0}
Preparing message: 6
Message: {'timeuuid_id': '1663108538.30657', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:38', 'total': 9, 'car': 1, 'bus': 0}
Preparing message: 7
Message: {'timeuuid_id': '1663108539.306729', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:39', 'total': 13, 'car': 3, 'bus': 0}
Preparing message: 8
Message: {'timeuuid_id': '1663108540.309405', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:40', 'total': 7, 'car': 0, 'bus': 0}
Preparing message: 9
Message: {'timeuuid_id': '1663108541.315883', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:41', 'total': 15, 'car': 4, 'bus': 0}
Preparing message: 10
Message: {'timeuuid_id': '1663108542.319062', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 'time_saved': '06:35:42', 'total': 5, 'car': 0, 'bus': 0}

Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663108670.34572', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:50', 'total': 10, 'car': 0, 'bus': 2}
Preparing message: 2
Message: {'timeuuid_id': '1663108671.361143', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:51', 'total': 8, 'car': 2, 'bus': 2}
Preparing message: 3
Message: {'timeuuid_id': '1663108672.361277', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:52', 'total': 4, 'car': 1, 'bus': 1}
Preparing message: 4
Message: {'timeuuid_id': '1663108673.376736', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:53', 'total': 15, 'car': 4, 'bus': 0}
Preparing message: 5
Message: {'timeuuid_id': '1663108674.377023', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:54', 'total': 13, 'car': 2, 'bus': 0}
Preparing message: 6
Message: {'timeuuid_id': '1663108675.383354', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:55', 'total': 9, 'car': 2, 'bus': 1}
Preparing message: 7
Message: {'timeuuid_id': '1663108676.398898', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:56', 'total': 11, 'car': 0, 'bus': 0}
Preparing message: 8
Message: {'timeuuid_id': '1663108677.405827', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:57', 'total': 11, 'car': 0, 'bus': 0}
Preparing message: 9
Message: {'timeuuid_id': '1663108678.412204', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:58', 'total': 6, 'car': 0, 'bus': 1}
Preparing message: 10
Message: {'timeuuid_id': '1663108679.424115', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 'time_saved': '06:37:59', 'total': 10, 'car': 1, 'bus': 0}

Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663108806.772574', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 'time_saved': '06:40:06', 'total': 6, 'car': 0, 'bus': 0}
Preparing message: 2
Message: {'timeuuid_id': '1663108807.783702', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 'time_saved': '06:40:07', 'total': 7, 'car': 2, 'bus': 0}
Preparing message: 3
Message: {'timeuuid_id': '1663108808.798885', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 'time_saved': '06:40:08', 'total': 10, 'car': 0, 'bus': 1}
Preparing message: 4
Message: {'timeuuid_id': '1663108809.808708', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 'time_saved': '06:40:09', 'total': 8, 'car': 0, 'bus': 2}

Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663108869.326692', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 'time_saved': '06:41:09', 'total': 8, 'car': 2, 'bus': 0}
Preparing message: 2
Message: {'timeuuid_id': '1663108870.336798', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 'time_saved': '06:41:10', 'total': 3, 'car': 0, 'bus': 0}
Preparing message: 3
Message: {'timeuuid_id': '1663108871.351573', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 'time_saved': '06:41:11', 'total': 11, 'car': 4, 'bus': 2}
Preparing message: 4
Message: {'timeuuid_id': '1663108872.362572', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 'time_saved': '06:41:12', 'total': 10, 'car': 2, 'bus': 1}

Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663108943.317013', 'lgu_code': '1200', 'sensor_id': 'sensor_06', 'date_saved': '09/14/2022', 'time_saved': '06:42:23', 'total': 15, 'car': 4, 'bus': 2}
Preparing message: 2
Message: {'timeuuid_id': '1663108944.326098', 'lgu_code': '1200', 'sensor_id': 'sensor_06', 'date_saved': '09/14/2022', 'time_saved': '06:42:24', 'total': 7, 'car': 1, 'bus': 1}
Preparing message: 3
Message: {'timeuuid_id': '1663108945.330967', 'lgu_code': '1200', 'sensor_id': 'sensor_06', 'date_saved': '09/14/2022', 'time_saved': '06:42:25', 'total': 12, 'car': 2, 'bus': 2}
Preparing message: 4
Message: {'timeuuid_id': '1663108946.342103', 'lgu_code': '1200', 'sensor_id': 'sensor_06', 'date_saved': '09/14/2022', 'time_saved': '06:42:26', 'total': 10, 'car': 4, 'bus': 2}
```

```
Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663109018.154734', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 'time_saved': '06:43:38', 'total': 7, 'car': 0, 'bus': 2,
Preparing message: 2
Message: {'timeuuid_id': '1663109019.170806', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 'time_saved': '06:43:39', 'total': 6, 'car': 0, 'bus': 1,
Preparing message: 3
Message: {'timeuuid_id': '1663109020.181974', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 'time_saved': '06:43:40', 'total': 13, 'car': 0, 'bus': 1
Preparing message: 4
Message: {'timeuuid_id': '1663109021.183327', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 'time_saved': '06:43:41', 'total': 10, 'car': 2, 'bus': 2
Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663109104.757234', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 'time_saved': '06:45:04', 'total': 10, 'car': 4, 'bus': 1
Preparing message: 2
Message: {'timeuuid_id': '1663109105.764428', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 'time_saved': '06:45:05', 'total': 13, 'car': 2, 'bus': 0
Preparing message: 3
Message: {'timeuuid_id': '1663109106.771872', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 'time_saved': '06:45:06', 'total': 10, 'car': 0, 'bus': 1
Preparing message: 4
Message: {'timeuuid_id': '1663109107.776004', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 'time_saved': '06:45:07', 'total': 8, 'car': 3, 'bus': 0,
Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663109176.854518', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 'time_saved': '06:46:16', 'total': 12, 'car': 4, 'bus': 1
Preparing message: 2
Message: {'timeuuid_id': '1663109177.869439', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 'time_saved': '06:46:17', 'total': 5, 'car': 2, 'bus': 0,
Preparing message: 3
Message: {'timeuuid_id': '1663109178.8852', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 'time_saved': '06:46:18', 'total': 6, 'car': 0, 'bus': 2,
Preparing message: 4
Message: {'timeuuid_id': '1663109179.899545', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 'time_saved': '06:46:19', 'total': 7, 'car': 1, 'bus': 1,
Kafka Producer Application Started ...
Preparing message: 1
Message: {'timeuuid_id': '1663109243.053172', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 'time_saved': '06:47:23', 'total': 19, 'car': 4, 'bus': 2
Preparing message: 2
Message: {'timeuuid_id': '1663109244.069559', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 'time_saved': '06:47:24', 'total': 9, 'car': 0, 'bus': 0,
Preparing message: 3
Message: {'timeuuid_id': '1663109245.073896', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 'time_saved': '06:47:25', 'total': 12, 'car': 2, 'bus': 1
Preparing message: 4
Message: {'timeuuid_id': '1663109246.076631', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 'time_saved': '06:47:26', 'total': 11, 'car': 2, 'bus': 2
```

## Producer

```
C:\Windows\System32\cmd.exe
{"timeuuid_id": "1663107535.039989", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:18:55", "total": 7, "car": 0, "bus": 0, "truck": 1, "jeepney": 2, "bike": 1, "tryk": 2, "others": 1}
{"timeuuid_id": "1663107536.055275", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:18:56", "total": 10, "car": 1, "bus": 2, "truck": 1, "jeepney": 0, "bike": 2, "tryk": 3, "others": 1}
{"timeuuid_id": "1663107537.070778", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:18:57", "total": 11, "car": 3, "bus": 1, "truck": 1, "jeepney": 0, "bike": 2, "tryk": 2, "others": 2}
{"timeuuid_id": "1663107538.08569", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:18:58", "total": 12, "car": 4, "bus": 0, "truck": 2, "jeepney": 2, "bike": 0, "tryk": 3, "others": 1}
{"timeuuid_id": "1663107539.101464", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:18:59", "total": 9, "car": 1, "bus": 2, "truck": 0, "jeepney": 2, "bike": 2, "tryk": 2, "others": 0}
{"timeuuid_id": "1663107540.101726", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:00", "total": 13, "car": 1, "bus": 0, "truck": 2, "jeepney": 2, "bike": 5, "tryk": 4, "others": 2}
{"timeuuid_id": "1663107541.11749", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:01", "total": 10, "car": 1, "bus": 1, "truck": 0, "jeepney": 2, "bike": 3, "tryk": 1, "others": 2}
{"timeuuid_id": "1663107542.132971", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:02", "total": 14, "car": 3, "bus": 2, "truck": 2, "jeepney": 0, "bike": 3, "tryk": 3, "others": 1}
{"timeuuid_id": "1663107543.147886", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:03", "total": 10, "car": 4, "bus": 0, "truck": 2, "jeepney": 0, "bike": 1, "tryk": 2, "others": 1}
{"timeuuid_id": "1663107544.148133", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:04", "total": 11, "car": 1, "bus": 2, "truck": 0, "jeepney": 1, "bike": 5, "tryk": 3, "others": 1}
{"timeuuid_id": "1663107545.161215", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:05", "total": 12, "car": 4, "bus": 1, "truck": 2, "jeepney": 1, "bike": 0, "tryk": 3, "others": 1}
{"timeuuid_id": "1663107546.174342", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:06", "total": 14, "car": 1, "bus": 2, "truck": 2, "jeepney": 2, "bike": 5, "tryk": 1, "others": 1}
{"timeuuid_id": "1663107547.187653", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:07", "total": 4, "car": 2, "bus": 0, "truck": 1, "jeepney": 0, "bike": 0, "tryk": 0, "others": 0}
{"timeuuid_id": "1663107548.202316", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:08", "total": 16, "car": 4, "bus": 2, "truck": 0, "jeepney": 2, "bike": 5, "tryk": 2, "others": 2}
{"timeuuid_id": "1663107549.20382", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:09", "total": 10, "car": 3, "bus": 1, "truck": 1, "jeepney": 0, "bike": 1, "tryk": 0, "others": 2}
{"timeuuid_id": "1663107550.219282", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:10", "total": 13, "car": 4, "bus": 0, "truck": 1, "jeepney": 2, "bike": 4, "tryk": 1, "others": 1}
{"timeuuid_id": "1663107551.22138", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:11", "total": 10, "car": 2, "bus": 1, "truck": 2, "jeepney": 2, "bike": 1, "tryk": 0, "others": 2}
{"timeuuid_id": "1663107552.23555", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:12", "total": 8, "car": 0, "bus": 0, "truck": 1, "jeepney": 0, "bike": 4, "tryk": 3, "others": 0}
{"timeuuid_id": "1663107553.238276", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:13", "total": 5, "car": 2, "bus": 0, "truck": 0, "jeepney": 0, "bike": 1, "tryk": 1, "others": 1}
{"timeuuid_id": "1663107554.243183", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:14", "total": 13, "car": 3, "bus": 0, "truck": 2, "jeepney": 1, "bike": 4, "tryk": 3, "others": 0}
{"timeuuid_id": "1663107555.25204", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:15", "total": 7, "car": 1, "bus": 1, "truck": 2, "jeepney": 0, "bike": 3, "tryk": 0, "others": 0}
{"timeuuid_id": "1663107556.259174", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:16", "total": 9, "car": 1, "bus": 0, "truck": 0, "jeepney": 0, "bike": 5, "tryk": 2, "others": 1}
{"timeuuid_id": "1663107557.263214", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:17", "total": 13, "car": 0, "bus": 2, "truck": 2, "jeepney": 1, "bike": 4, "tryk": 0, "others": 1}
{"timeuuid_id": "1663107558.270584", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:18", "total": 10, "car": 1, "bus": 1, "truck": 0, "jeepney": 2, "bike": 5, "tryk": 0, "others": 1}
{"timeuuid_id": "1663107559.272886", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:19", "total": 12, "car": 3, "bus": 2, "truck": 2, "jeepney": 1, "bike": 2, "tryk": 0, "others": 2}
{"timeuuid_id": "1663107560.277281", "lgu_code": "1200", "sensor_id": "sensor_01", "date_saved": "09/14/2022", "time_saved": "06:19:20", "total": 8, "car": 1, "bus": 1, "truck": 0, "jeepney": 1, "bike": 2, "tryk": 1, "others": 2}
```



## Consumer

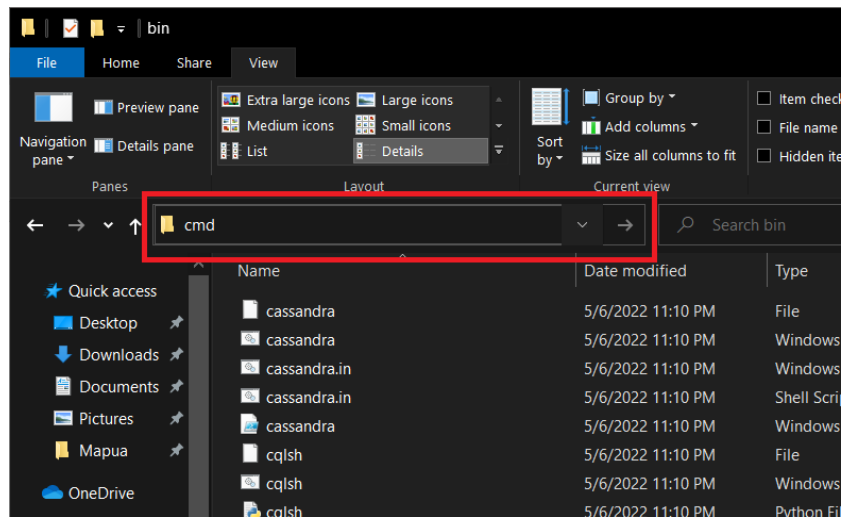
Note: Any mode through which the consumer results will be displayed is accepted, as long as the results show that the consumer is receiving the data being sent by the producer. For this instance, consumer results were displayed on a command prompt terminal opened through Kafka, instead of the provided consumer code. If the consumer code were run through PyCharm or a command prompt opened to where the consumer code was located, the same results still would have been gained.

```
C:\Windows\System32\cmd.exe - python virtual_consumer.py
id": "1663254469.54", "bike": 0, "truck": 1, "tryke": 1, "bus": 1, "date_saved": "09/15/2022", "total": 4}
{"sensor_id": "sensor_09", "time_saved": "23:07:49", "car": 4, "others": 0, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254469.54", "bike": 0, "truck": 2, "tryke": 1, "bus": 2, "date_saved": "09/15/2022", "total": 9}
{"sensor_id": "sensor_04", "time_saved": "23:07:49", "car": 0, "others": 0, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254469.57", "bike": 3, "truck": 2, "tryke": 0, "bus": 1, "date_saved": "09/15/2022", "total": 6}
{"sensor_id": "sensor_08", "time_saved": "23:07:49", "car": 0, "others": 1, "jeepney": 1, "lgu_code": "1200", "timeuuid_
id": "1663254469.57", "bike": 0, "truck": 2, "tryke": 3, "bus": 0, "date_saved": "09/15/2022", "total": 7}
{"sensor_id": "sensor_05", "time_saved": "23:07:49", "car": 0, "others": 0, "jeepney": 2, "lgu_code": "1200", "timeuuid_
id": "1663254469.63", "bike": 0, "truck": 0, "tryke": 3, "bus": 0, "date_saved": "09/15/2022", "total": 5}
{"sensor_id": "sensor_03", "time_saved": "23:07:50", "car": 2, "others": 0, "jeepney": 1, "lgu_code": "1200", "timeuuid_
id": "1663254470.43", "bike": 0, "truck": 0, "tryke": 1, "bus": 0, "date_saved": "09/15/2022", "total": 4}
{"sensor_id": "sensor_01", "time_saved": "23:07:50", "car": 2, "others": 2, "jeepney": 1, "lgu_code": "1200", "timeuuid_
id": "1663254470.51", "bike": 3, "truck": 1, "tryke": 0, "bus": 2, "date_saved": "09/15/2022", "total": 11}
{"sensor_id": "sensor_02", "time_saved": "23:07:50", "car": 3, "others": 2, "jeepney": 1, "lgu_code": "1200", "timeuuid_
id": "1663254470.51", "bike": 3, "truck": 1, "tryke": 3, "bus": 0, "date_saved": "09/15/2022", "total": 13}
{"sensor_id": "sensor_07", "time_saved": "23:07:50", "car": 3, "others": 2, "jeepney": 2, "lgu_code": "1200", "timeuuid_
id": "1663254470.53", "bike": 5, "truck": 2, "tryke": 2, "bus": 2, "date_saved": "09/15/2022", "total": 18}
{"sensor_id": "sensor_06", "time_saved": "23:07:50", "car": 0, "others": 0, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254470.53", "bike": 0, "truck": 0, "tryke": 1, "bus": 0, "date_saved": "09/15/2022", "total": 1}
{"sensor_id": "sensor_09", "time_saved": "23:07:50", "car": 0, "others": 0, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254470.56", "bike": 0, "truck": 0, "tryke": 3, "bus": 2, "date_saved": "09/15/2022", "total": 5}
{"sensor_id": "sensor_10", "time_saved": "23:07:50", "car": 1, "others": 1, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254470.56", "bike": 2, "truck": 1, "tryke": 0, "bus": 0, "date_saved": "09/15/2022", "total": 5}
{"sensor_id": "sensor_04", "time_saved": "23:07:50", "car": 4, "others": 1, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254470.59", "bike": 1, "truck": 0, "tryke": 2, "bus": 2, "date_saved": "09/15/2022", "total": 10}
{"sensor_id": "sensor_08", "time_saved": "23:07:50", "car": 0, "others": 0, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254470.59", "bike": 0, "truck": 0, "tryke": 1, "bus": 0, "date_saved": "09/15/2022", "total": 1}
{"sensor_id": "sensor_05", "time_saved": "23:07:50", "car": 2, "others": 2, "jeepney": 0, "lgu_code": "1200", "timeuuid_
id": "1663254470.65", "bike": 3, "truck": 1, "tryke": 0, "bus": 1, "date_saved": "09/15/2022", "total": 9}
```

## II. Keyspace (Database) and Table Creation in Cassandra (Producers)

A keyspace in Cassandra is similar to a schema/database in relational database management systems (RDMBS). In a Cassandra cluster, a keyspace is a data container that determines how data replicates on nodes.

To use Apache Cassandra, navigate to the bin folder within the Apache Cassandra folder. We can start the Windows Command Prompt directly from within the bin folder by typing 'cmd' in the address bar and pressing Enter.

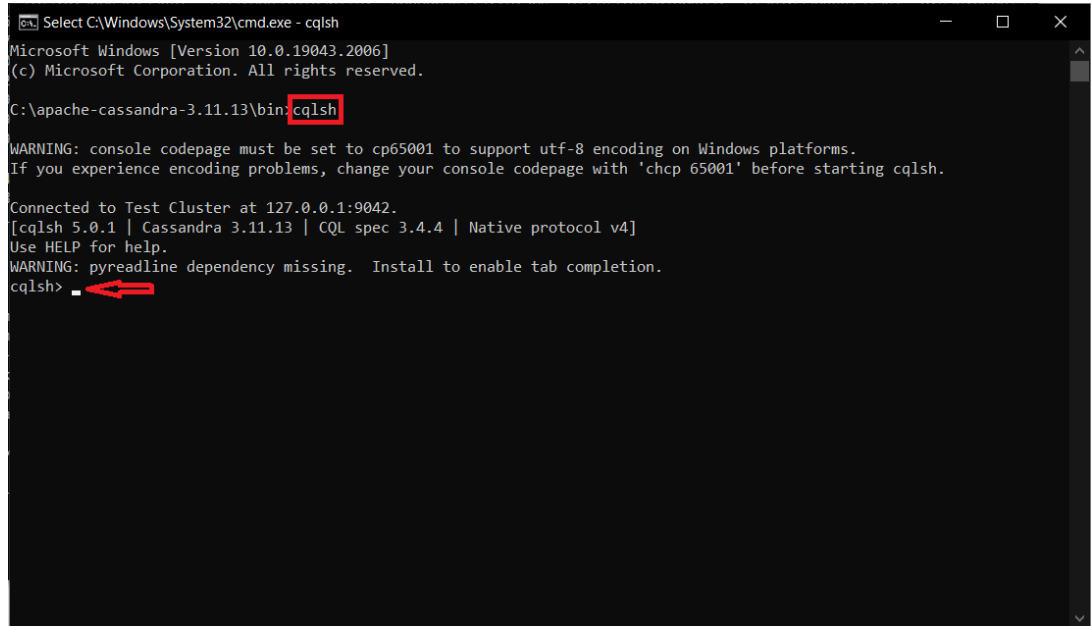


Once the command prompt is open, type in 'cassandra' to start the Cassandra server. **Do not close** this command prompt session.

```
C:\Windows\System32\cmd.exe - cassandra
Microsoft Windows [Version 10.0.19043.1889]
(c) Microsoft Corporation. All rights reserved.

C:\apache-cassandra-3.11.13\bin>cassandra
WARNING! Powershell script execution unavailable.
Please use 'powershell Set-ExecutionPolicy Unrestricted'
on this user-account to run cassandra with fully featured
functionality on this platform.
Starting with legacy startup options
Starting Cassandra Server
INFO [main] 2022-09-14 09:55:01,327 YamlConfigurationLoader.java:93 - Configuration location: file:/C:/apache-cassandra
-3.11.13/conf/cassandra.yaml
INFO [main] 2022-09-14 09:55:02,217 Config.java:555 - Node configuration:[allocate_tokens_for_keyspace=null; allow_extr
a_insecure_udfs=false; allow_insecure_udfs=false; authenticator=AllowAllAuthenticator; authorizer=AllowAllAuthorizer; au
to_bootstrap=true; auto_snapshot=true; back_pressure_enabled=false; back_pressure_strategy=org.apache.cassandra.net.Rate
BasedBackPressure[high_ratio=0.9, factor=5, flow=FAST]; batch_size_fail_threshold_in_kb=50; batch_size_warn_threshold_in
_kb=5; batchlog_replay_throttle_in_kb=1024; broadcast_address=null; broadcast_rpc_address=null; buffer_pool_use_heap_if_
exhausted=true; cache_load_timeout_seconds=30; cas_contention_timeout_in_ms=1000; cdc_enabled=false; cdc_free_space_chec
k_interval_ms=250; cdc_raw_directory=null; cdc_total_space_in_mb=0; check_for_duplicate_rows_during_compaction=true; che
ck_for_duplicate_rows_during_reads=true; client_encryption_options=<REDACTED>; cluster_name=Test Cluster; column_index_c
ache_size_in_kb=2; column_index_size_in_kb=64; commit_failure_policy=stop; commitlog_compression=null; commitlog_directo
ry=null; commitlog_max_compression_buffers_in_pool=3; commitlog_periodic_queue_size=-1; commitlog_segment_size_in_mb=32;
commitlog_sync=periodic; commitlog_sync_batch_window_in_ms=NaN; commitlog_sync_period_in_ms=10000; commitlog_total_spac
e_in_mb=null; compaction_large_partition_warning_threshold_mb=100; compaction_throughput_mb_per_sec=16; concurrent_compa
ctors=null; concurrent_counter_writes=32; concurrent_materialized_view_writes=32; concurrent_reads=32; concurrent_replic
ates=null; concurrent_writes=32; counter_cache_keys_to_save=2147483647; counter_cache_save_period=7200; counter_cache_si
ze_in_mb=null; counter_write_request_timeout_in_ms=5000; credentials_cache_max_entries=1000; credentials_update_interval
_in_ms=1; credentials_validity_in_ms=2000; cross_node_timeout=false; data_file_directories=[Ljava.lang.String;@34e9fd99
; disk_access_mode=auto; disk_failure_policy=stop; disk_optimization_estimate_percentile=0.95; disk_optimization_page_cr
oss_chance=0.1; disk_optimization_strategy=ssd; dynamic_snitch=true; dynamic_snitch_badness_threshold=0.1; dynamic_snitc
```

Navigate again to the bin folder within the Apache Cassandra folder. Open another command prompt directly by typing 'cmd' in the address bar and pressing Enter. This time, type 'cqlsh' to access the Cassandra **cqlsh** bash shell.



```
Select C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19043.2006]
(c) Microsoft Corporation. All rights reserved.

C:\apache-cassandra-3.11.13\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh>
```

In Apache Cassandra, the basic syntax for creating a keyspace with different replication settings is:

```
CREATE KEYSPACE keyspace_name WITH replication = {properties};
```

Following this basic syntax, we created a keyspace named **group23\_project** with **SimpleStrategy** and **replication\_factor 1**.

```
CREATE      KEYSPACE      group23_project      WITH      replication      =
{'class':'SimpleStrategy', 'replication_factor': 1};
```

We use SimpleStrategy since we do not intend to deploy a cluster to more than one data center. We also set the replication\_factor to 1 since we only have one node.

```
C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19043.2006]
(c) Microsoft Corporation. All rights reserved.

C:\apache-cassandra-3.11.13\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh> CREATE KEYSPACE group23_project WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
cqlsh>
```

Once we have created our keyspace, we can verify that the keyspace is on the list by using this command:

DESCRIBE KEYSPACES;                      or                      DESC KEYSPACES;

```
C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19043.2006]
(c) Microsoft Corporation. All rights reserved.

C:\apache-cassandra-3.11.13\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh> CREATE KEYSPACE group23_project WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
cqlsh> DESCRIBE KEYSPACES;

system_schema    system                      system_traces
system_auth      system_distributed    group23_project

cqlsh>
```

Now that we have created our keyspace/database, we can perform actions on it such as creating a table. Before we start creating our table, we must specify the keyspace where we want to create the table. There are two commands to do this:

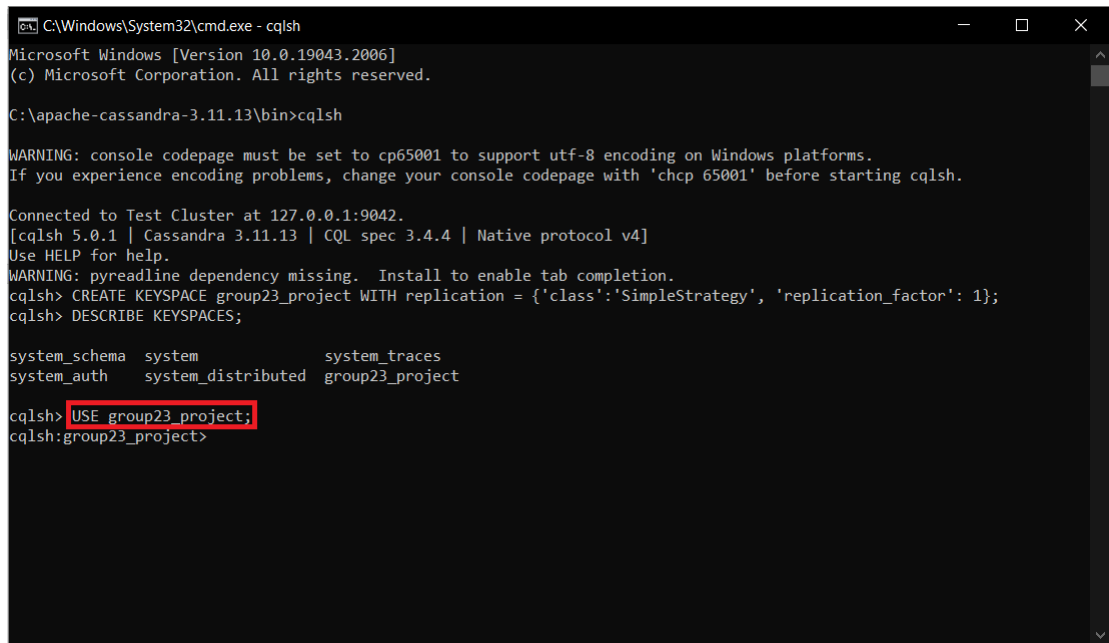
USE keyspace\_name;

or

```
CREATE TABLE keyspace_name.table_name
```

For this example, we will be using the USE command to select the keyspace. We selected the keyspace we made previously using the command:

```
USE group23_project;
```



```
C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19043.2006]
(c) Microsoft Corporation. All rights reserved.

C:\apache-cassandra-3.11.13\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh> CREATE KEYSPACE group23_project WITH replication = {'class':'SimpleStrategy', 'replication_factor': 1};
cqlsh> DESCRIBE KEYSPACES;

system_schema  system              system_traces
system_auth    system_distributed  group23_project

cqlsh> USE group23_project;
cqlsh:group23_project>
```

After selecting our keyspace, we can create a table using the basic syntax:

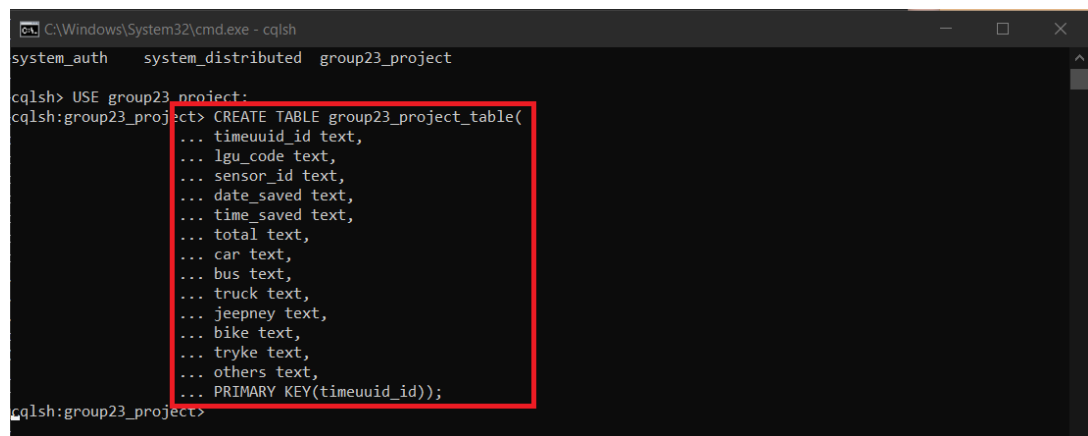
```
CREATE TABLE tableName (
    columnName1 dataType,
    columnName2 dataType,
    columnName2 datatype,
    PRIMARY KEY (columnName)
);
```

For our example, we will be creating a table with the name *group23\_project\_table* and with the following columns:

```
CREATE TABLE group23_project_table(
    timeuuid_id text,
    lgu_code text,
    sensor_id text,
    date_saved text,
```



```
time_saved text,  
total text,  
car text,  
bus text,  
truck text,  
jeepney text,  
bike text,  
tryke text,  
others text,  
PRIMARY KEY(timeuuid_id));
```

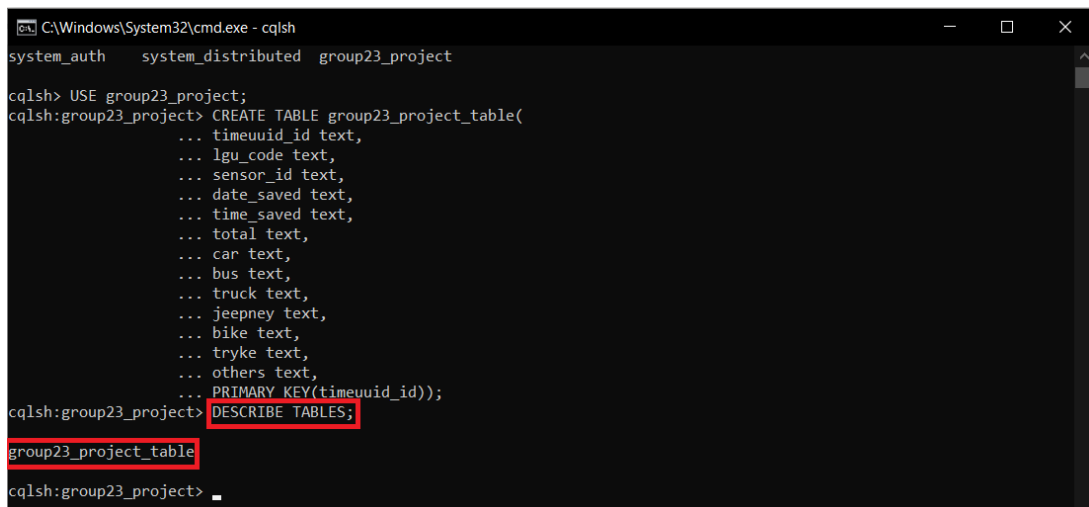


A screenshot of a CQLSH terminal window. The title bar shows the path 'C:\Windows\System32\cmd.exe - cqlsh'. The terminal content shows the user has selected the 'group23\_project' keyspace. The command 'CREATE TABLE group23\_project\_table(' is being entered, followed by a list of columns: 'timeuuid\_id text', 'lgu\_code text', 'sensor\_id text', 'date\_saved text', 'time\_saved text', 'total text', 'car text', 'bus text', 'truck text', 'jeepney text', 'bike text', 'tryke text', 'others text', and 'PRIMARY KEY(timeuuid\_id));'. The list of columns is highlighted with a red rectangle.

```
C:\Windows\System32\cmd.exe - cqlsh  
system_auth system_distributed group23_project  
cqlsh> USE group23_project;  
cqlsh:group23_project> CREATE TABLE group23_project_table(  
... timeuuid_id text,  
... lgu_code text,  
... sensor_id text,  
... date_saved text,  
... time_saved text,  
... total text,  
... car text,  
... bus text,  
... truck text,  
... jeepney text,  
... bike text,  
... tryke text,  
... others text,  
... PRIMARY KEY(timeuuid_id));  
cqlsh:group23_project>
```

To check if the table is in the keyspace, we can use the command:

```
DESCRIBE TABLES;
```



A screenshot of a CQLSH terminal window. The title bar shows the path 'C:\Windows\System32\cmd.exe - cqlsh'. The terminal content shows the user has selected the 'group23\_project' keyspace. The command 'DESCRIBE TABLES;' has been entered, and the output 'group23\_project\_table' is displayed. Both the command and the output are highlighted with red rectangles.

```
C:\Windows\System32\cmd.exe - cqlsh  
system_auth system_distributed group23_project  
cqlsh> USE group23_project;  
cqlsh:group23_project> CREATE TABLE group23_project_table(  
... timeuuid_id text,  
... lgu_code text,  
... sensor_id text,  
... date_saved text,  
... time_saved text,  
... total text,  
... car text,  
... bus text,  
... truck text,  
... jeepney text,  
... bike text,  
... tryke text,  
... others text,  
... PRIMARY KEY(timeuuid_id));  
cqlsh:group23_project> DESCRIBE TABLES;  
group23_project_table  
cqlsh:group23_project>
```

To show the contents of our table, we can use the command:

```
SELECT * FROM group23_project_table;
```

As we can see in the image below, it shows all the columns defined in the creation of our table. However, the table is empty since we have not inserted any data yet.

**Note:** Do not close this command prompt yet as we will still be using it later.



```
C:\Windows\System32\cmd.exe - cqlsh
cqlsh:group23_project> DESCRIBE TABLES;

group23_project_table
cqlsh:group23_project> SELECT * FROM group23_project_table;

timeuuid_id | bike | bus | car | date_saved | jeepney | lgu_code | others | sensor_id | time_saved | total | truck | tryke
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
(0 rows)
cqlsh:group23_project>
```

### III. Fetching data from Kafka using Python then Sending or Saving the Same Data to Cassandra

Now that the connection between the producer and consumer have been established and a table within a Cassandra keyspace created, data fetching from Apache Kafka to Cassandra may now be accomplished. To write the program that will fetch data from Kafka to Cassandra, the following libraries were initially imported to access Kafka and Cassandra from Python.

```
from future import print_function
from kafka import KafkaConsumer
from cassandra.cluster import Cluster
```

**Note:** Python 2.7 was used as the Python Interpreter as it is the latest version of Python that Kafka and Cassandra support.

Since the data being retrieved is coming from the Kafka consumer, a KafkaConsumer function with the following parameters were initialized: Kafka topic name ('counts'), bootstrap\_servers (localhost: 9092), auto\_offset\_reset, and enable\_auto\_commit.

```
consumer = KafkaConsumer(
    'counts', # topic in Kafka
    bootstrap_servers = ['127.0.0.1:9092'],
    auto_offset_reset = 'latest',
    enable_auto_commit = True
)
```

A similar initializing of the local host and port number was fulfilled by the Cluster() function from the Cassandra.cluster.

```
cluster = Cluster(['127.0.0.1'], port = 9042)
```

Now that the necessary initializations have been established, the designated keyspace (group23\_project) in which the Cassandra table was created may now be connected to Python through the code below.

```
session = cluster.connect('group23_project')
```

To create the columns for the group23\_project\_table in Cassandra, the following code was written.

```
columns = ['timeuuid_id', 'lgu_code', 'sensor_id', 'date_saved',  
'time_saved', 'total', 'car', 'bus', 'truck', 'jeepney', 'bike',  
'tryke', 'others']
```

Since our data in the Kafka consumer is in string format, we want to separate each value by using message.replace() to replace the curly brackets and quotation marks into whitespace and message.split() to separate each word. We then arrange and store the data into an array based on its position in the string.

```
for message in consumer:  
    message = message.value  
    message = message.replace("{", "  
    message = message.replace("}", "  
    message = message.replace("'", "  
    message = message.split(", "  
  
    timeuuid_id = message[6].split(':')[1]  
    lgu_code = message[5].split(':')[1]  
    sensor_id = message[0].split(':')[1]  
    date_saved = message[11].split(':')[1]  
    time_saved = message[1].split(':')[1]  
    total = message[12].split(':')[1]  
    car = message[2].split(':')[1]  
    bus = message[10].split(':')[1]  
    truck = message[8].split(':')[1]  
    jeepney = message[4].split(':')[1]
```

```

bike = message[7].split(': ')[1]
tryke = message[9].split(': ')[1]
others = message[3].split(': ')[1]

```

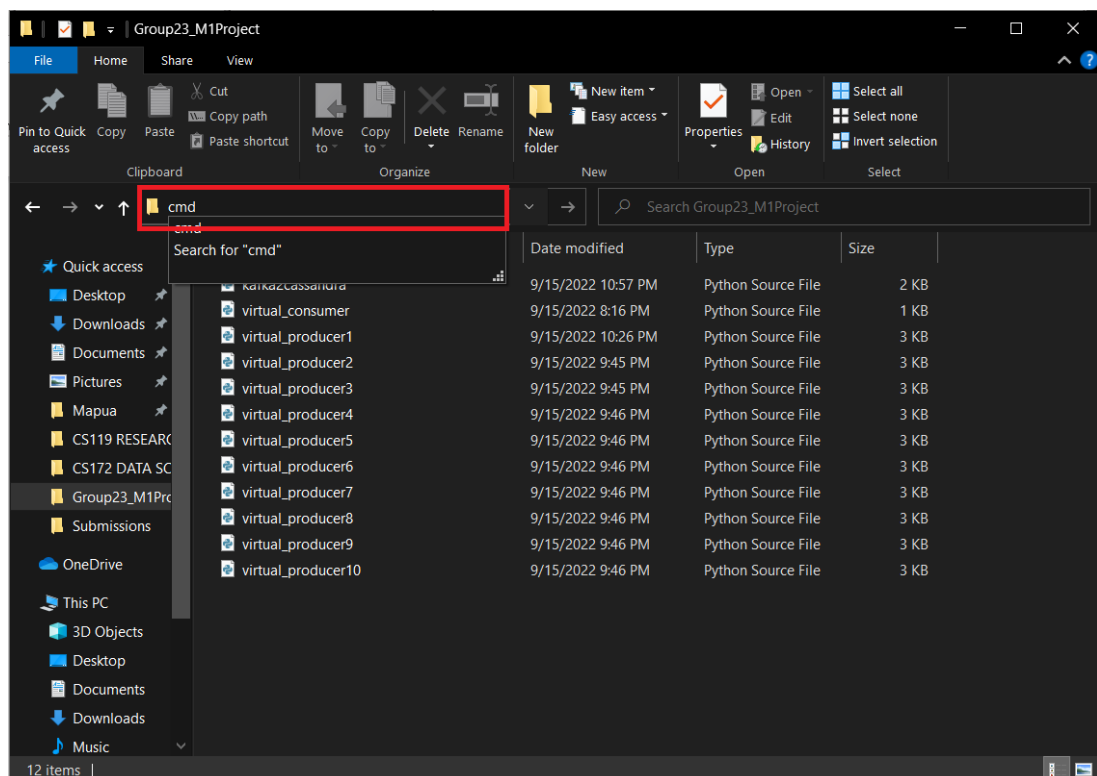
After making the necessary modifications to the consumer output, the following code was written to insert the values into the `group23_project_table` in Cassandra.

```

session.execute("INSERT INTO group23_project_table(timeuuid_id,
lgu_code, sensor_id, date_saved, time_saved, total, car, bus, truck,
jeepney, bike, tryke, others) VALUES('{}', '{}', '{}', '{}', '{}',
 '{}', '{}', '{}', '{}', '{}', '{}', '{}', '{}',
 '{}');".format(timeuuid_id, lgu_code, sensor_id, date_saved,
time_saved, total, car, bus, truck, jeepney, bike, tryke, others))

```

If the values were successfully inserted, an “Inserted...” prompt should appear when the `python kafka2cassandra.py` command is entered into a command prompt terminal opened in the project folder.





## REFERENCES

- DataMaking. (2019a, October 20). *Real-Time Apache Spark Project / Real-Time Data Analysis / Apache Kafka / Part 4 / DM / DataMaking* [Video]. YouTube. Retrieved September 14, 2022, from <https://www.youtube.com/watch?v=Tt7F5lZvO-E&list=PLelT0uBrDrfOYE8OwQvooPjmnP1zY3wFe&index=6>
- Jevtic, G. (2021, August 20). *How to Create Keyspace in Cassandra*. Knowledge Base by phoenixNAP. Retrieved September 15, 2022, from <https://phoenixnap.com/kb/cassandra-create-keyspace>
- Jevtic, G. (2021a, August 17). *How to Create, Drop, Alter, and Truncate Tables in Cassandra*. Knowledge Base by phoenixNAP. Retrieved September 15, 2022, from <https://phoenixnap.com/kb/create-drop-alter-and-truncate-tables-in-cassandra>