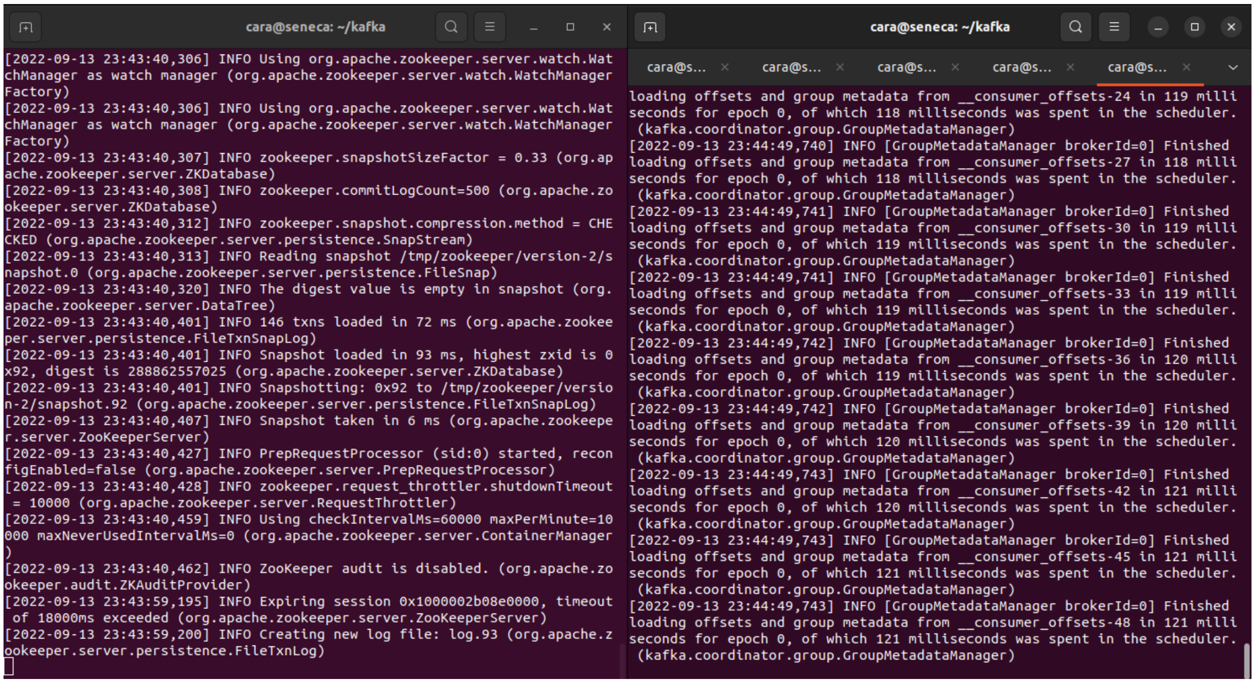
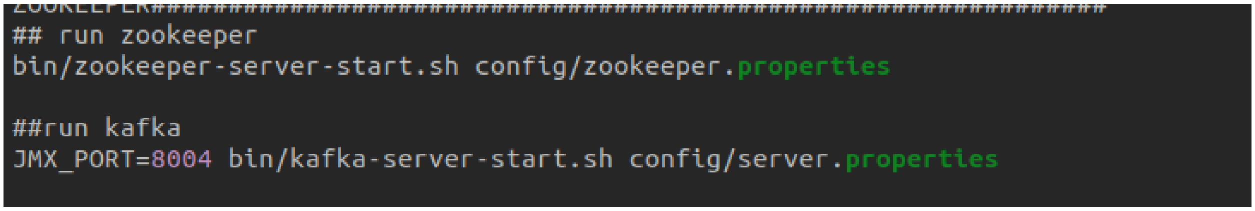
Big Data Streaming Pipeline and Integration Platform Project



```
[2022-09-13 23:43:40,306] INFO Using org.apache.zookeeper.server.watch.Wat
chManager as watch manager (org.apache.zookeeper.server.watch.WatchManager
Factory)
[2022-09-13 23:43:40,306] INFO Using org.apache.zookeeper.server.watch.Wat
chManager as watch manager (org.apache.zookeeper.server.watch.WatchManager
Factory)
[2022-09-13 23:43:40,307] INFO zookeeper.snapshotSizeFactor = 0.33 (org.ap
ache.zookeeper.server.ZKDatabase)
[2022-09-13 23:43:40,308] INFO zookeeper.commitLogCount=500 (org.apache.zo
okeeper.server.ZKDatabase)
[2022-09-13 23:43:40,312] INFO zookeeper.snapshot.compression.method = CHE
CKED (org.apache.zookeeper.server.persistence.SnapStream)
[2022-09-13 23:43:40,313] INFO Reading snapshot /tmp/zookeeper/version-2/s
napshot.0 (org.apache.zookeeper.server.persistence.FileSnap)
[2022-09-13 23:43:40,320] INFO The digest value is empty in snapshot (org.
apache.zookeeper.server.DataTree)
[2022-09-13 23:43:40,401] INFO 146 txns loaded in 72 ms (org.apache.zookee
per.server.persistence.FileTxnSnapLog)
[2022-09-13 23:43:40,401] INFO Snapshot loaded in 93 ms, highest zxid is 0
x92, digest is 288862557025 (org.apache.zookeeper.server.ZKDatabase)
[2022-09-13 23:43:40,401] INFO Snapshotting: 0x92 to /tmp/zookeeper/versio
n-2/snapshot.92 (org.apache.zookeeper.server.persistence.FileTxnSnapLog)
[2022-09-13 23:43:40,407] INFO Snapshot taken in 6 ms (org.apache.zookeepe
r.server.ZooKeeperServer)
[2022-09-13 23:43:40,427] INFO PrepRequestProcessor (sid:0) started, recon
figEnabled=false (org.apache.zookeeper.server.PrepRequestProcessor)
[2022-09-13 23:43:40,428] INFO zookeeper.request_throttler.shutdownTimeout
 = 10000 (org.apache.zookeeper.server.RequestThrottler)
[2022-09-13 23:43:40,459] INFO Using checkIntervalMs=60000 maxPerMinute=10
000 maxNeverUsedIntervalMs=0 (org.apache.zookeeper.server.ContainerManager
)
[2022-09-13 23:43:40,462] INFO ZooKeeper audit is disabled. (org.apache.zo
okeeper.audit.ZKAuditProvider)
[2022-09-13 23:43:59,195] INFO Expiring session 0x1000002b08e0000, timeout
 of 18000ms exceeded (org.apache.zookeeper.server.ZooKeeperServer)
[2022-09-13 23:43:59,200] INFO Creating new log file: log.93 (org.apache.z
ookeeper.server.persistence.FileTxnLog)
```

```
loading offsets and group metadata from __consumer_offsets-24 in 119 milli
seconds for epoch 0, of which 118 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,740] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-27 in 118 milli
seconds for epoch 0, of which 118 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,741] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-30 in 119 milli
seconds for epoch 0, of which 119 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,741] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-33 in 119 milli
seconds for epoch 0, of which 119 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,742] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-36 in 120 milli
seconds for epoch 0, of which 119 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,742] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-39 in 120 milli
seconds for epoch 0, of which 120 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,743] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-42 in 121 milli
seconds for epoch 0, of which 120 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,743] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-45 in 121 milli
seconds for epoch 0, of which 121 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
[2022-09-13 23:44:49,743] INFO [GroupMetadataManager brokerId=0] Finished
loading offsets and group metadata from __consumer_offsets-48 in 121 milli
seconds for epoch 0, of which 121 milliseconds was spent in the scheduler.
 (kafka.coordinator.group.GroupMetadataManager)
```

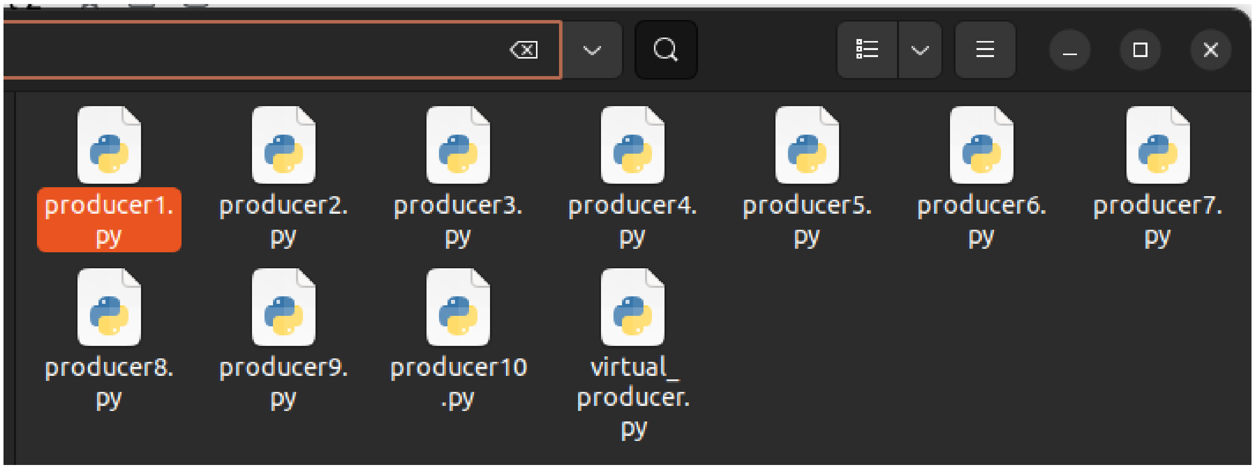
Zookeeper and Kakfa Running



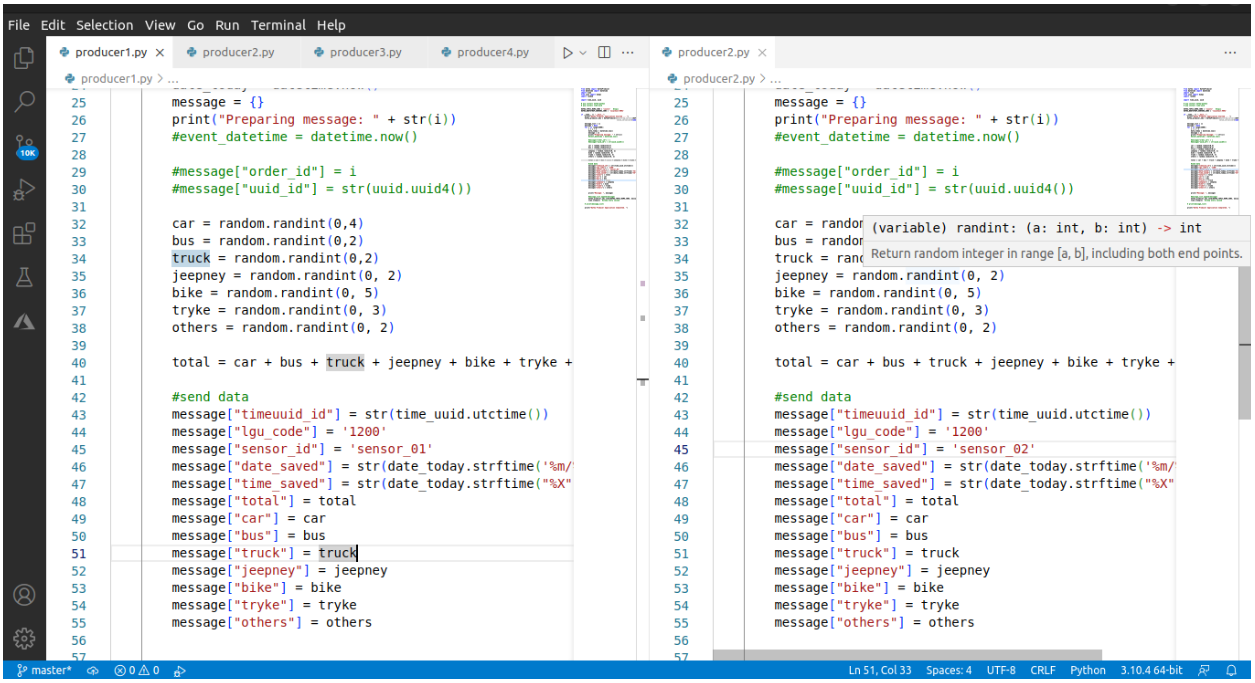
```
ZOOKEEPER##############################################################
## run zookeeper
bin/zookeeper-server-start.sh config/zookeeper.properties

##run kafka
JMX_PORT=8004 bin/kafka-server-start.sh config/server.properties
```

Commands used to run zookeeper and kafka.

Part 1: 10 snippets of producer's code



10 instances of producers, each one has unique sensor_id from 1 -10

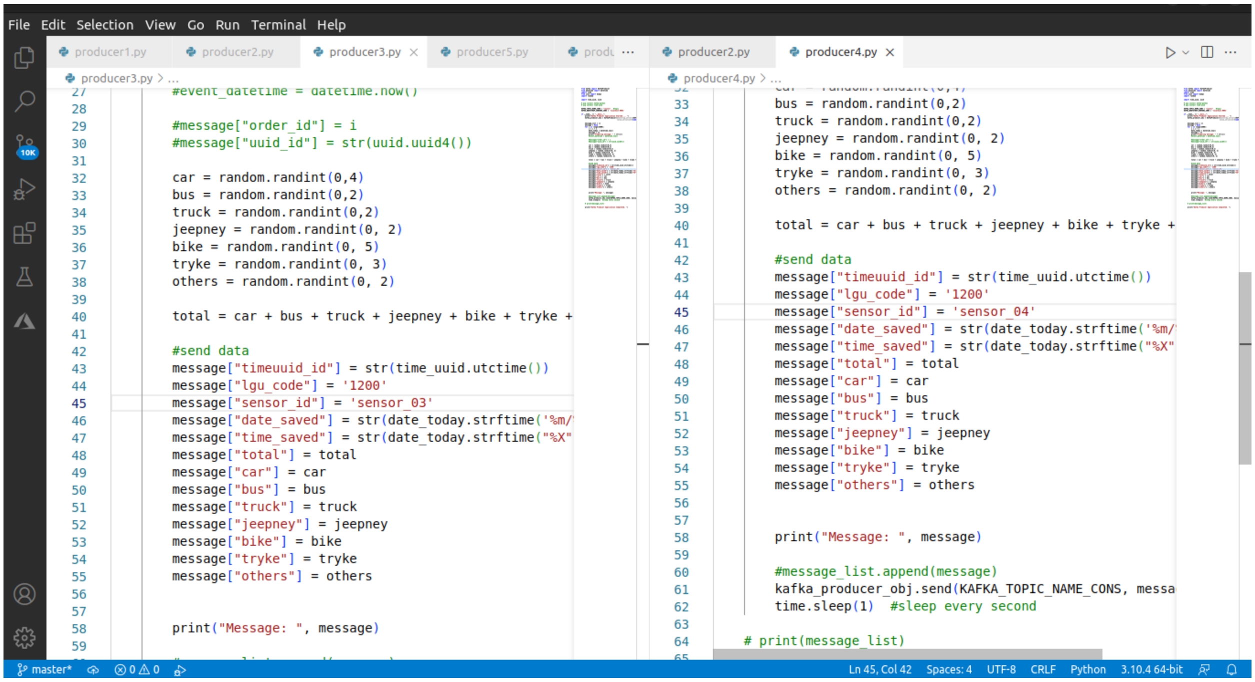


```python
        message = {}
        print("Preparing message: " + str(i))
        #event_datetime = datetime.now()

        #message["order_id"] = i
        #message["uuid_id"] = str(uuid.uuid4())

        car = random.randint(0,4)
        bus = random.randint(0,2)
        truck = random.randint(0,2)
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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
```

```python
        message = {}
        print("Preparing message: " + str(i))
        #event_datetime = datetime.now()

        #message["order_id"] = i
        #message["uuid_id"] = str(uuid.uuid4())

        car = randor
        bus = randor
        truck = ranc
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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
```

```
(variable) randint: (a: int, b: int) -> int
Return random integer in range [a, b], including both end points.
```

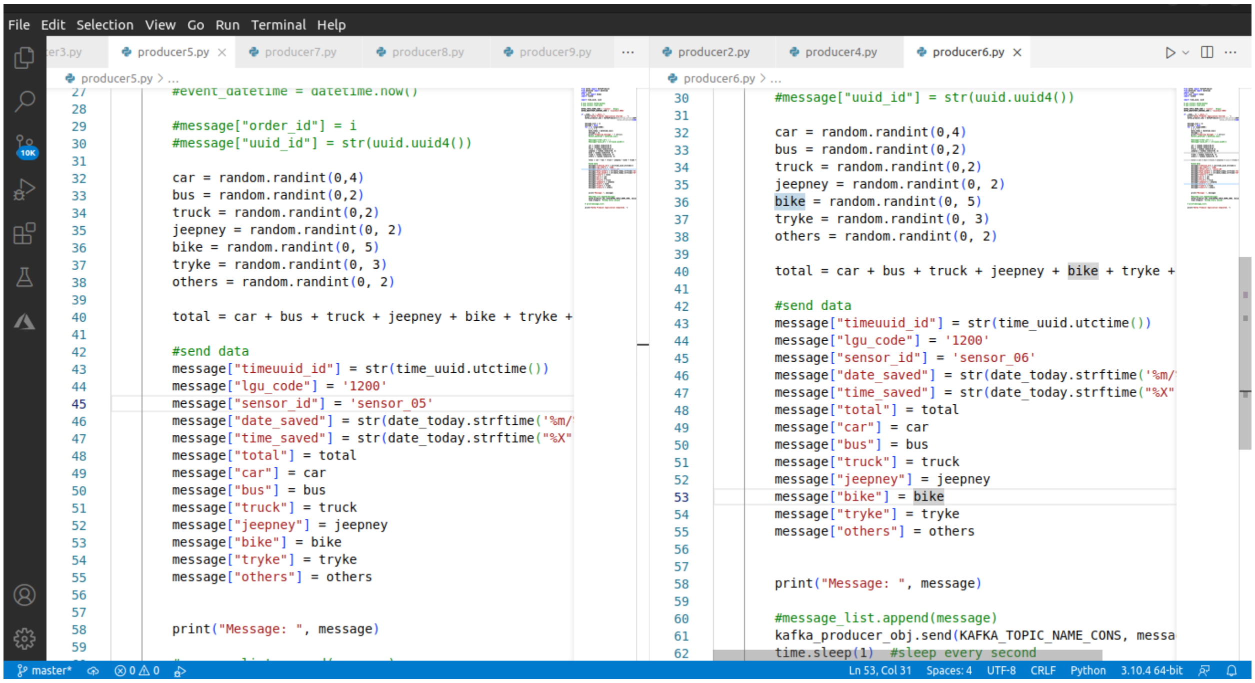
Producer 1 and 2



```python
        #event_datetime = datetime.now()

        #message["order_id"] = i
        #message["uuid_id"] = str(uuid.uuid4())

        car = random.randint(0,4)
        bus = random.randint(0,2)
        truck = random.randint(0,2)
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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


        print("Message: ", message)
```

```python
        bus = random.randint(0,2)
        truck = random.randint(0,2)
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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


        print("Message: ", message)

        #message_list.append(message)
        kafka_producer_obj.send(KAFKA_TOPIC_NAME_CONS, messa
        time.sleep(1)  #sleep every second

    # print(message_list)
```

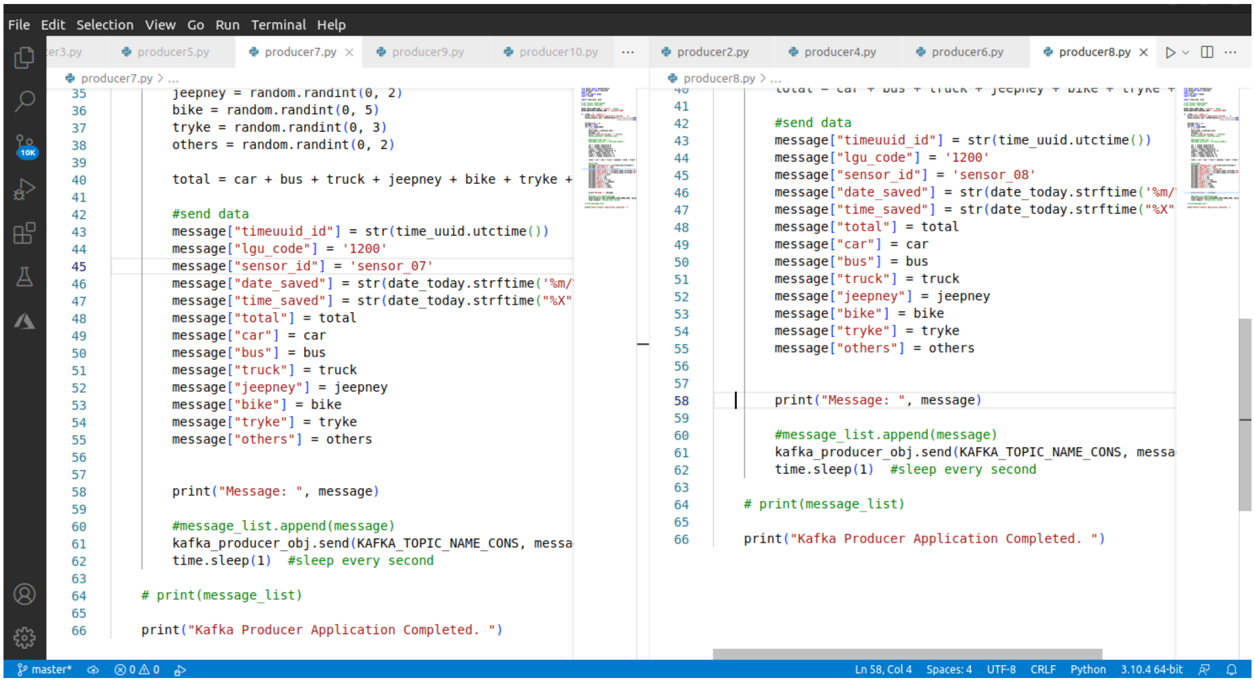
Producer 3 and 4



```python
        #event_datetime = datetime.now()

        #message["order_id"] = i
        #message["uuid_id"] = str(uuid.uuid4())

        car = random.randint(0,4)
        bus = random.randint(0,2)
        truck = random.randint(0,2)
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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


        print("Message: ", message)
```

```python
        #message["uuid_id"] = str(uuid.uuid4())

        car = random.randint(0,4)
        bus = random.randint(0,2)
        truck = random.randint(0,2)
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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


        print("Message: ", message)

        #message_list.append(message)
        kafka_producer_obj.send(KAFKA_TOPIC_NAME_CONS, messa
        time.sleep(1)  #sleep every second
```

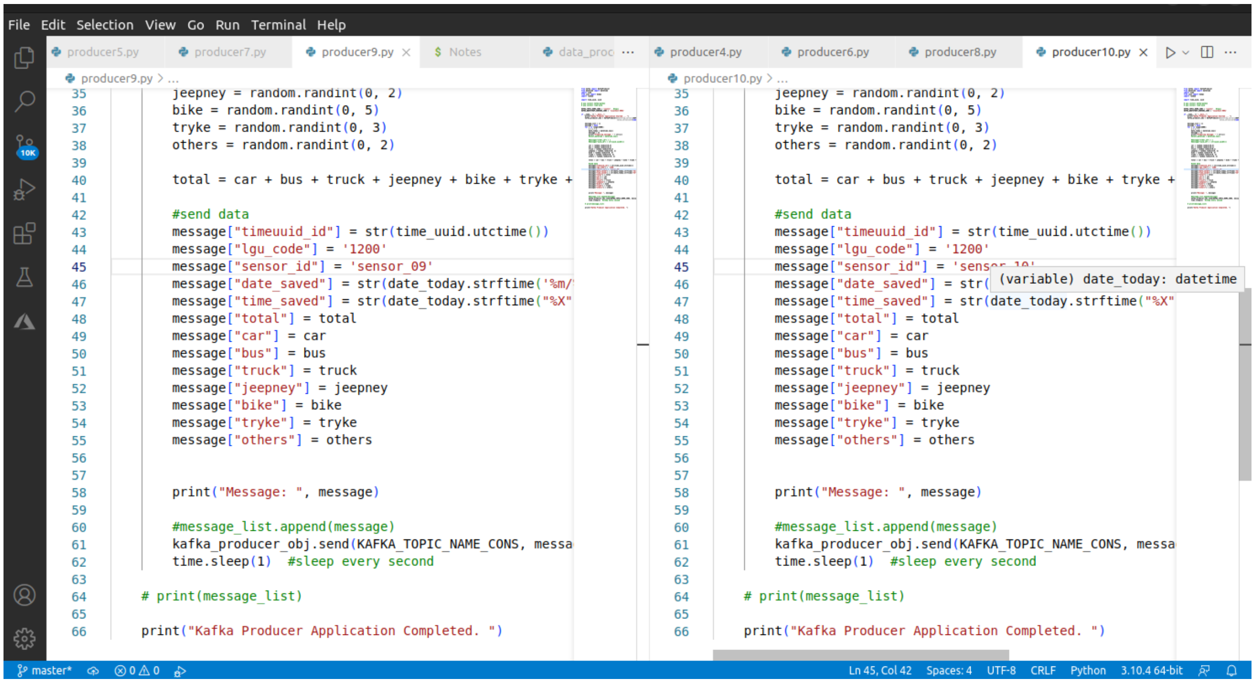
Producer 5 and 6



```python
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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


        print("Message: ", message)

        #message_list.append(message)
        kafka_producer_obj.send(KAFKA_TOPIC_NAME_CONS, messa
        time.sleep(1)  #sleep every second

    # print(message_list)

    print("Kafka Producer Application Completed. ")
```

```python
        total = car + bus + truck + jeepney + bike + tryke +

        #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/
        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


        print("Message: ", message)

        #message_list.append(message)
        kafka_producer_obj.send(KAFKA_TOPIC_NAME_CONS, messa
        time.sleep(1)  #sleep every second

    # print(message_list)

    print("Kafka Producer Application Completed. ")
```

Producer 7 and 8

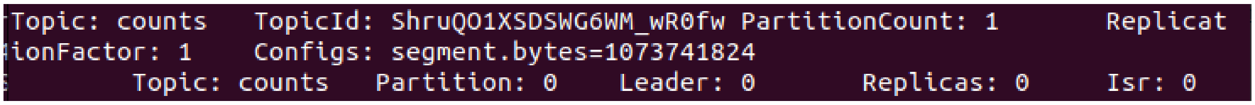


```python
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #send data
        message["timeuuid_id"] = str(time_uuid.utctime())
        message["lgu_code"] = '1200'
        message["sensor_id"] = 'sensor_09'
        message["date_saved"] = str(date_today.strftime('%m/
        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


        print("Message: ", message)

        #message_list.append(message)
        kafka_producer_obj.send(KAFKA_TOPIC_NAME_CONS, messa
        time.sleep(1)  #sleep every second

    # print(message_list)

    print("Kafka Producer Application Completed. ")
```

```python
        jeepney = random.randint(0, 2)
        bike = random.randint(0, 5)
        tryke = random.randint(0, 3)
        others = random.randint(0, 2)

        total = car + bus + truck + jeepney + bike + tryke +

        #send data
        message["timeuuid_id"] = str(time_uuid.utctime())
        message["lgu_code"] = '1200'
        message["sensor_id"] = 'sensor_10'
        message["date_saved"] = str(
        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


        print("Message: ", message)

        #message_list.append(message)
        kafka_producer_obj.send(KAFKA_TOPIC_NAME_CONS, messa
        time.sleep(1)  #sleep every second

    # print(message_list)

    print("Kafka Producer Application Completed. ")
```

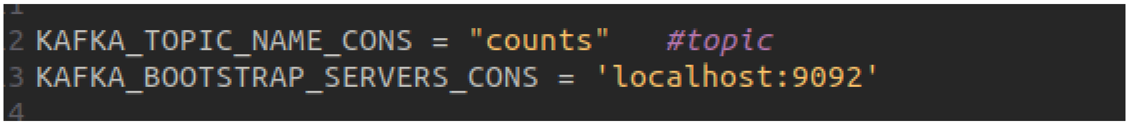
Producer 9 and 10

Part II. Create topic and run producers



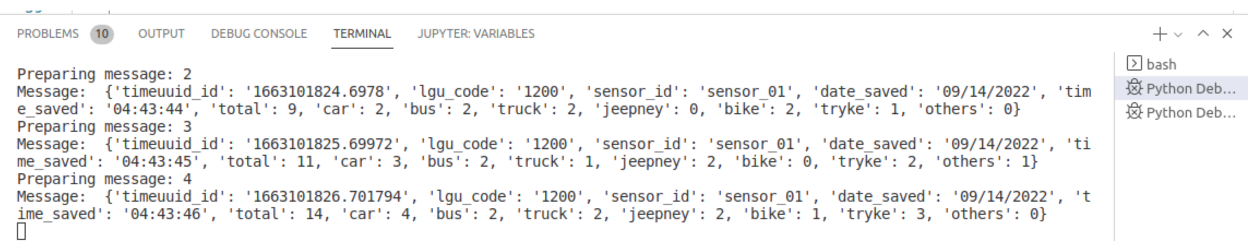
```
Topic: counts    TopicId: ShruQO1XSDSWG6WM_wR0fw PartitionCount: 1        Replicat
ionFactor: 1     Configs: segment.bytes=1073741824
         Topic: counts    Partition: 0     Leader: 0         Replicas: 0      Isr: 0
```

Created topic ‘counts’. Different producers



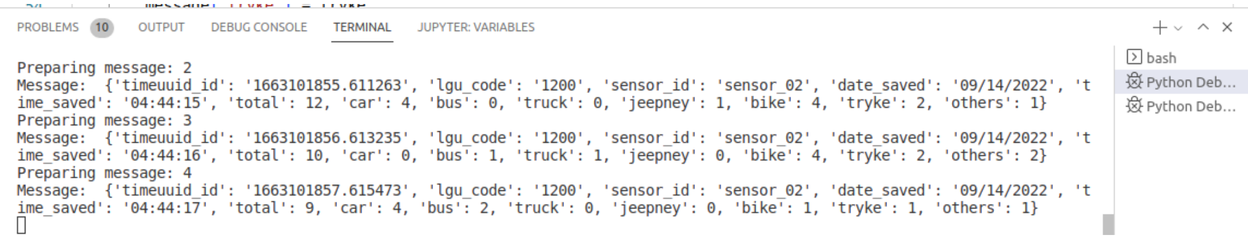
```
KAFKA_TOPIC_NAME_CONS = "counts"    #topic
KAFKA_BOOTSTRAP_SERVERS_CONS = 'localhost:9092'
```

Each producer are assigned to the ‘counts’ topic



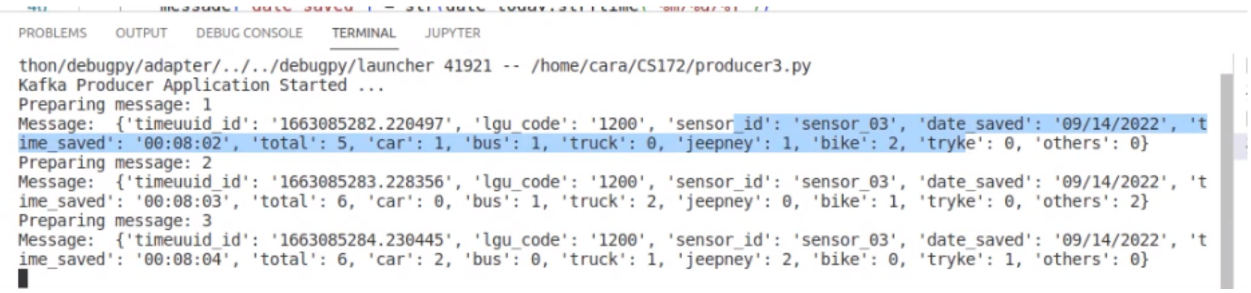
```
Preparing message: 2
Message:  {'timeuuid_id': '1663101824.6978', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'tim
e_saved': '04:43:44', 'total': 9, 'car': 2, 'bus': 2, 'truck': 2, 'jeepney': 0, 'bike': 2, 'tryke': 1, 'others': 0}
Preparing message: 3
Message:  {'timeuuid_id': '1663101825.69972', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 'ti
me_saved': '04:43:45', 'total': 11, 'car': 3, 'bus': 2, 'truck': 1, 'jeepney': 2, 'bike': 0, 'tryke': 2, 'others': 1}
Preparing message: 4
Message:  {'timeuuid_id': '1663101826.701794', 'lgu_code': '1200', 'sensor_id': 'sensor_01', 'date_saved': '09/14/2022', 't
ime_saved': '04:43:46', 'total': 14, 'car': 4, 'bus': 2, 'truck': 2, 'jeepney': 2, 'bike': 1, 'tryke': 3, 'others': 0}
```

Producer 1



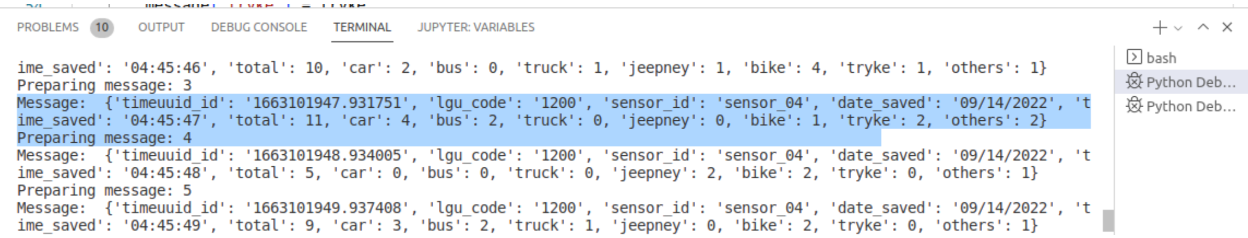
```
Preparing message: 2
Message:  {'timeuuid_id': '1663101855.611263', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 't
ime_saved': '04:44:15', 'total': 12, 'car': 4, 'bus': 0, 'truck': 0, 'jeepney': 1, 'bike': 4, 'tryke': 2, 'others': 1}
Preparing message: 3
Message:  {'timeuuid_id': '1663101856.613235', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 't
ime_saved': '04:44:16', 'total': 10, 'car': 0, 'bus': 1, 'truck': 1, 'jeepney': 0, 'bike': 4, 'tryke': 2, 'others': 2}
Preparing message: 4
Message:  {'timeuuid_id': '1663101857.615473', 'lgu_code': '1200', 'sensor_id': 'sensor_02', 'date_saved': '09/14/2022', 't
ime_saved': '04:44:17', 'total': 9, 'car': 4, 'bus': 2, 'truck': 0, 'jeepney': 0, 'bike': 1, 'tryke': 1, 'others': 1}
```

Producer 2



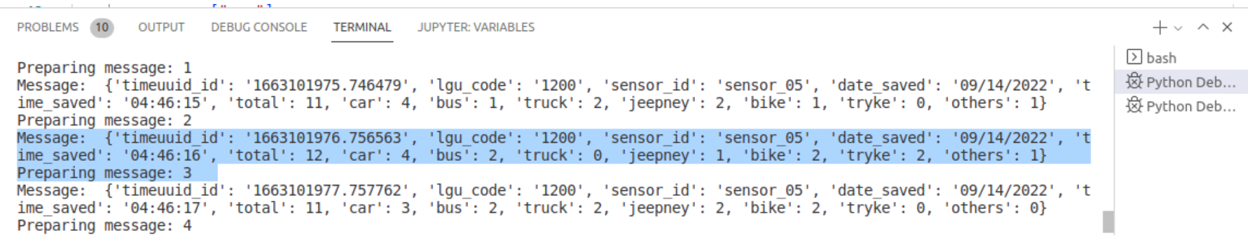
```
thon/debugpy/adapter/../../debugpy/launcher 41921 -- /home/cara/CS172/producer3.py
Kafka Producer Application Started ...
Preparing message: 1
Message:  {'timeuuid_id': '1663085282.220497', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 't
ime_saved': '00:08:02', 'total': 5, 'car': 1, 'bus': 1, 'truck': 0, 'jeepney': 1, 'bike': 2, 'tryke': 0, 'others': 0}
Preparing message: 2
Message:  {'timeuuid_id': '1663085283.228356', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 't
ime_saved': '00:08:03', 'total': 6, 'car': 0, 'bus': 1, 'truck': 2, 'jeepney': 0, 'bike': 1, 'tryke': 0, 'others': 2}
Preparing message: 3
Message:  {'timeuuid_id': '1663085284.230445', 'lgu_code': '1200', 'sensor_id': 'sensor_03', 'date_saved': '09/14/2022', 't
ime_saved': '00:08:04', 'total': 6, 'car': 2, 'bus': 0, 'truck': 1, 'jeepney': 2, 'bike': 0, 'tryke': 1, 'others': 0}
```

Producer 3



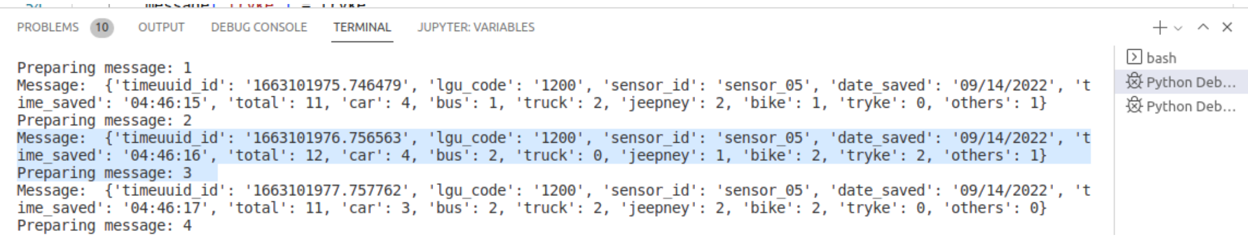
```
ime_saved': '04:45:46', 'total': 10, 'car': 2, 'bus': 0, 'truck': 1, 'jeepney': 1, 'bike': 4, 'tryke': 1, 'others': 1}
Preparing message: 3
Message:  {'timeuuid_id': '1663101947.931751', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 't
ime_saved': '04:45:47', 'total': 11, 'car': 4, 'bus': 2, 'truck': 0, 'jeepney': 0, 'bike': 1, 'tryke': 2, 'others': 2}
Preparing message: 4
Message:  {'timeuuid_id': '1663101948.934005', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 't
ime_saved': '04:45:48', 'total': 5, 'car': 0, 'bus': 0, 'truck': 0, 'jeepney': 2, 'bike': 2, 'tryke': 0, 'others': 1}
Preparing message: 5
Message:  {'timeuuid_id': '1663101949.937408', 'lgu_code': '1200', 'sensor_id': 'sensor_04', 'date_saved': '09/14/2022', 't
ime_saved': '04:45:49', 'total': 9, 'car': 3, 'bus': 2, 'truck': 1, 'jeepney': 0, 'bike': 2, 'tryke': 0, 'others': 1}
```

Producer 4



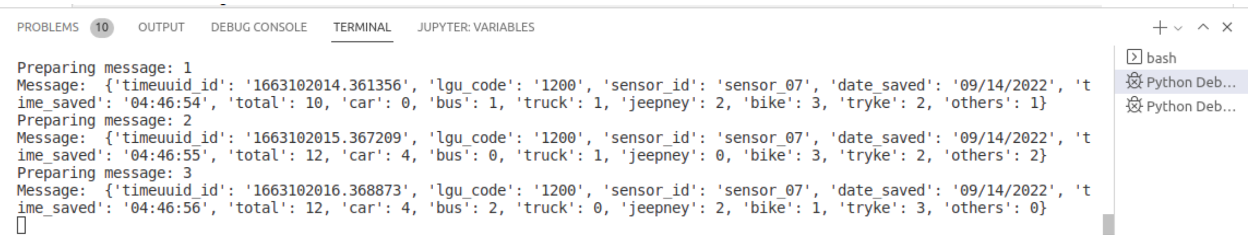
```
Preparing message: 1
Message:  {'timeuuid_id': '1663101975.746479', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:15', 'total': 11, 'car': 4, 'bus': 1, 'truck': 2, 'jeepney': 2, 'bike': 1, 'tryke': 0, 'others': 1}
Preparing message: 2
Message:  {'timeuuid_id': '1663101976.756563', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:16', 'total': 12, 'car': 4, 'bus': 2, 'truck': 0, 'jeepney': 1, 'bike': 2, 'tryke': 2, 'others': 1}
Preparing message: 3
Message:  {'timeuuid_id': '1663101977.757762', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:17', 'total': 11, 'car': 3, 'bus': 2, 'truck': 2, 'jeepney': 2, 'bike': 2, 'tryke': 0, 'others': 0}
Preparing message: 4
```

Producer 5



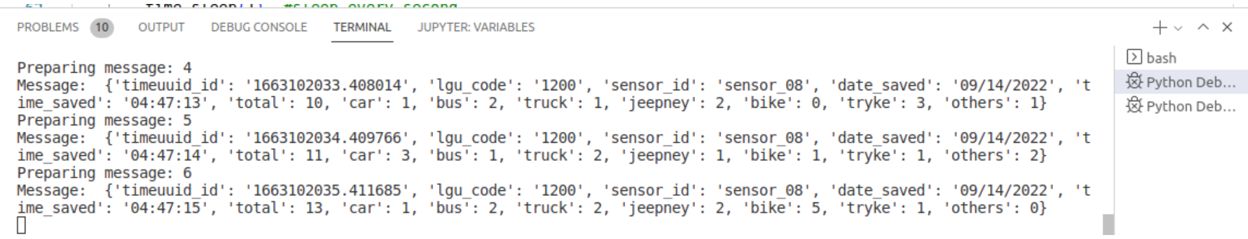
```
Preparing message: 1
Message:  {'timeuuid_id': '1663101975.746479', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:15', 'total': 11, 'car': 4, 'bus': 1, 'truck': 2, 'jeepney': 2, 'bike': 1, 'tryke': 0, 'others': 1}
Preparing message: 2
Message:  {'timeuuid_id': '1663101976.756563', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:16', 'total': 12, 'car': 4, 'bus': 2, 'truck': 0, 'jeepney': 1, 'bike': 2, 'tryke': 2, 'others': 1}
Preparing message: 3
Message:  {'timeuuid_id': '1663101977.757762', 'lgu_code': '1200', 'sensor_id': 'sensor_05', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:17', 'total': 11, 'car': 3, 'bus': 2, 'truck': 2, 'jeepney': 2, 'bike': 2, 'tryke': 0, 'others': 0}
Preparing message: 4
```

Producer 6



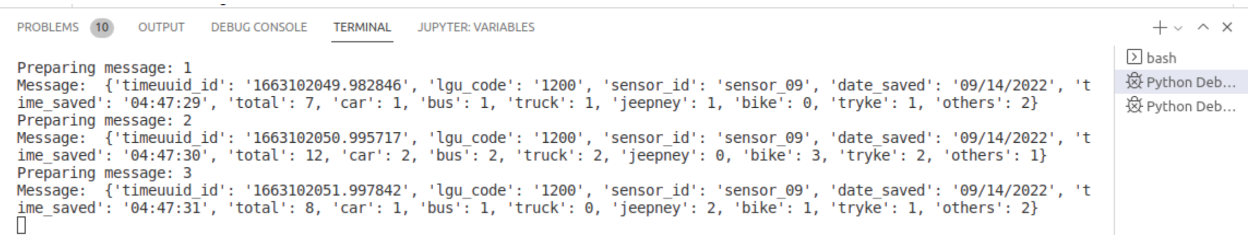
```
Preparing message: 1
Message:  {'timeuuid_id': '1663102014.361356', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:54', 'total': 10, 'car': 0, 'bus': 1, 'truck': 1, 'jeepney': 2, 'bike': 3, 'tryke': 2, 'others': 1}
Preparing message: 2
Message:  {'timeuuid_id': '1663102015.367209', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:55', 'total': 12, 'car': 4, 'bus': 0, 'truck': 1, 'jeepney': 0, 'bike': 3, 'tryke': 2, 'others': 2}
Preparing message: 3
Message:  {'timeuuid_id': '1663102016.368873', 'lgu_code': '1200', 'sensor_id': 'sensor_07', 'date_saved': '09/14/2022', 't
ime_saved': '04:46:56', 'total': 12, 'car': 4, 'bus': 2, 'truck': 0, 'jeepney': 2, 'bike': 1, 'tryke': 3, 'others': 0}
```

Producer 7



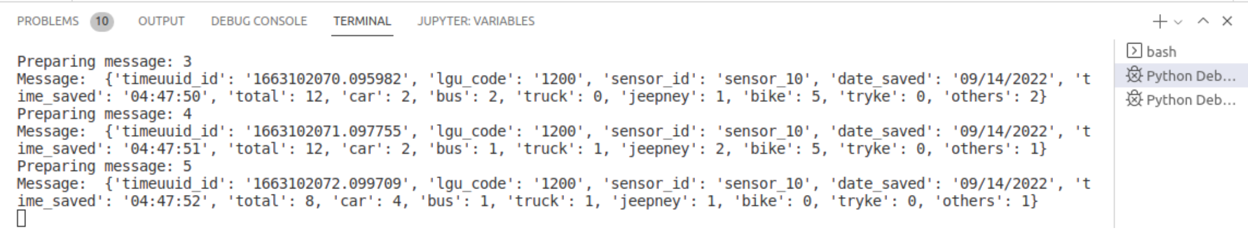
```
Preparing message: 4
Message:  {'timeuuid_id': '1663102033.408014', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:13', 'total': 10, 'car': 1, 'bus': 2, 'truck': 1, 'jeepney': 2, 'bike': 0, 'tryke': 3, 'others': 1}
Preparing message: 5
Message:  {'timeuuid_id': '1663102034.409766', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:14', 'total': 11, 'car': 3, 'bus': 1, 'truck': 2, 'jeepney': 1, 'bike': 1, 'tryke': 1, 'others': 2}
Preparing message: 6
Message:  {'timeuuid_id': '1663102035.411685', 'lgu_code': '1200', 'sensor_id': 'sensor_08', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:15', 'total': 13, 'car': 1, 'bus': 2, 'truck': 2, 'jeepney': 2, 'bike': 5, 'tryke': 1, 'others': 0}
```

Producer 8



```
Preparing message: 1
Message:  {'timeuuid_id': '1663102049.982846', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:29', 'total': 7, 'car': 1, 'bus': 1, 'truck': 1, 'jeepney': 1, 'bike': 0, 'tryke': 1, 'others': 2}
Preparing message: 2
Message:  {'timeuuid_id': '1663102050.995717', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:30', 'total': 12, 'car': 2, 'bus': 2, 'truck': 2, 'jeepney': 0, 'bike': 3, 'tryke': 2, 'others': 1}
Preparing message: 3
Message:  {'timeuuid_id': '1663102051.997842', 'lgu_code': '1200', 'sensor_id': 'sensor_09', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:31', 'total': 8, 'car': 1, 'bus': 1, 'truck': 0, 'jeepney': 2, 'bike': 1, 'tryke': 1, 'others': 2}
```

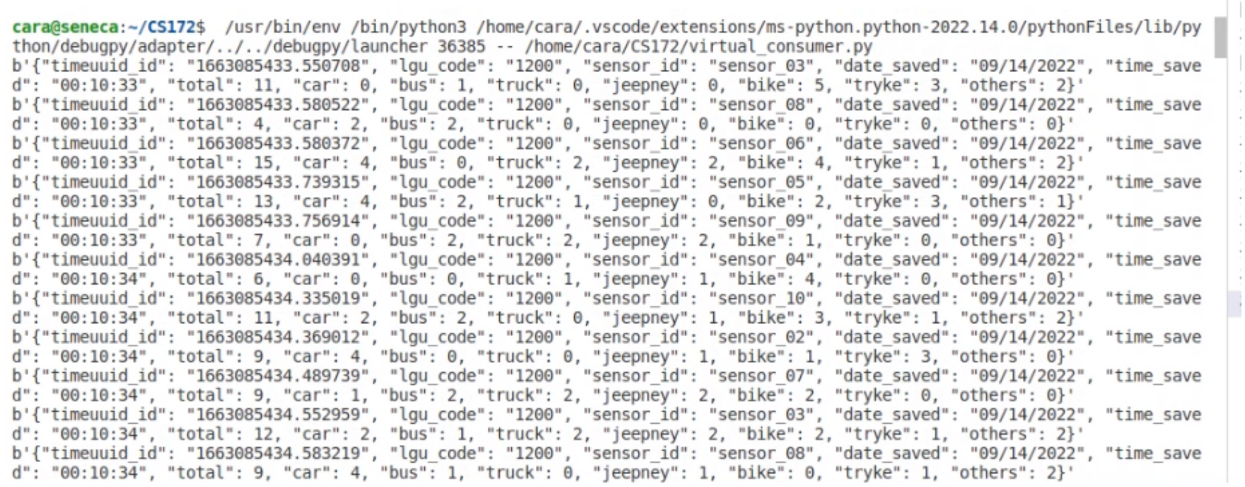
Producer 9



```
Preparing message: 3
Message:  {'timeuuid_id': '1663102070.095982', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:50', 'total': 12, 'car': 2, 'bus': 2, 'truck': 0, 'jeepney': 1, 'bike': 5, 'tryke': 0, 'others': 2}
Preparing message: 4
Message:  {'timeuuid_id': '1663102071.097755', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:51', 'total': 12, 'car': 2, 'bus': 1, 'truck': 1, 'jeepney': 2, 'bike': 5, 'tryke': 0, 'others': 1}
Preparing message: 5
Message:  {'timeuuid_id': '1663102072.099709', 'lgu_code': '1200', 'sensor_id': 'sensor_10', 'date_saved': '09/14/2022', 't
ime_saved': '04:47:52', 'total': 8, 'car': 4, 'bus': 1, 'truck': 1, 'jeepney': 1, 'bike': 0, 'tryke': 0, 'others': 1}
```

Producer 10

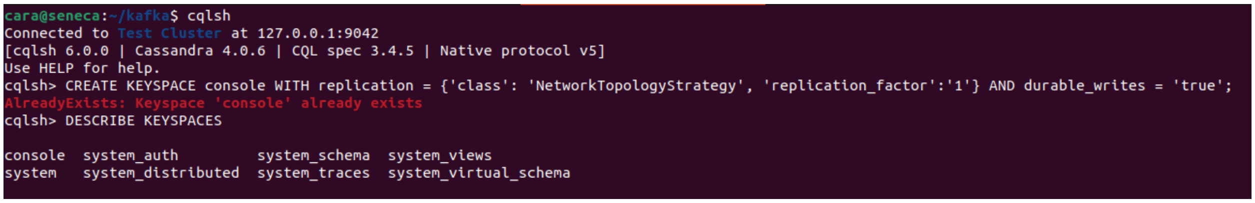
Part III. Data read by the consumer



```
cara@seneca:~/CS172$  /usr/bin/env /bin/python3 /home/cara/.vscode/extensions/ms-python.python-2022.14.0/pythonFiles/lib/py
thon/debugpy/adapter/../../debugpy/launcher 36385 -- /home/cara/CS172/virtual_consumer.py
b'{"timeuuid_id": "1663085433.550708", "lgu_code": "1200", "sensor_id": "sensor_03", "date_saved": "09/14/2022", "time_save
d": "00:10:33", "total": 11, "car": 0, "bus": 1, "truck": 0, "jeepney": 0, "bike": 5, "tryke": 3, "others": 2}'
b'{"timeuuid_id": "1663085433.580522", "lgu_code": "1200", "sensor_id": "sensor_08", "date_saved": "09/14/2022", "time_save
d": "00:10:33", "total": 4, "car": 2, "bus": 2, "truck": 0, "jeepney": 0, "bike": 0, "tryke": 0, "others": 0}'
b'{"timeuuid_id": "1663085433.580372", "lgu_code": "1200", "sensor_id": "sensor_06", "date_saved": "09/14/2022", "time_save
d": "00:10:33", "total": 15, "car": 4, "bus": 0, "truck": 2, "jeepney": 2, "bike": 4, "tryke": 1, "others": 2}'
b'{"timeuuid_id": "1663085433.739315", "lgu_code": "1200", "sensor_id": "sensor_05", "date_saved": "09/14/2022", "time_save
d": "00:10:33", "total": 13, "car": 4, "bus": 2, "truck": 1, "jeepney": 0, "bike": 2, "tryke": 3, "others": 1}'
b'{"timeuuid_id": "1663085433.756914", "lgu_code": "1200", "sensor_id": "sensor_09", "date_saved": "09/14/2022", "time_save
d": "00:10:33", "total": 7, "car": 0, "bus": 2, "truck": 2, "jeepney": 2, "bike": 1, "tryke": 0, "others": 0}'
b'{"timeuuid_id": "1663085434.040391", "lgu_code": "1200", "sensor_id": "sensor_04", "date_saved": "09/14/2022", "time_save
d": "00:10:34", "total": 6, "car": 0, "bus": 0, "truck": 1, "jeepney": 1, "bike": 4, "tryke": 0, "others": 0}'
b'{"timeuuid_id": "1663085434.335019", "lgu_code": "1200", "sensor_id": "sensor_10", "date_saved": "09/14/2022", "time_save
d": "00:10:34", "total": 11, "car": 2, "bus": 2, "truck": 0, "jeepney": 1, "bike": 3, "tryke": 1, "others": 2}'
b'{"timeuuid_id": "1663085434.369012", "lgu_code": "1200", "sensor_id": "sensor_02", "date_saved": "09/14/2022", "time_save
d": "00:10:34", "total": 9, "car": 4, "bus": 0, "truck": 0, "jeepney": 1, "bike": 1, "tryke": 3, "others": 0}'
b'{"timeuuid_id": "1663085434.489739", "lgu_code": "1200", "sensor_id": "sensor_07", "date_saved": "09/14/2022", "time_save
d": "00:10:34", "total": 9, "car": 1, "bus": 2, "truck": 2, "jeepney": 2, "bike": 2, "tryke": 0, "others": 0}'
b'{"timeuuid_id": "1663085434.552959", "lgu_code": "1200", "sensor_id": "sensor_03", "date_saved": "09/14/2022", "time_save
d": "00:10:34", "total": 12, "car": 2, "bus": 1, "truck": 2, "jeepney": 2, "bike": 2, "tryke": 1, "others": 2}'
b'{"timeuuid_id": "1663085434.583219", "lgu_code": "1200", "sensor_id": "sensor_08", "date_saved": "09/14/2022", "time_save
d": "00:10:34", "total": 9, "car": 4, "bus": 1, "truck": 0, "jeepney": 1, "bike": 0, "tryke": 1, "others": 2}'
```

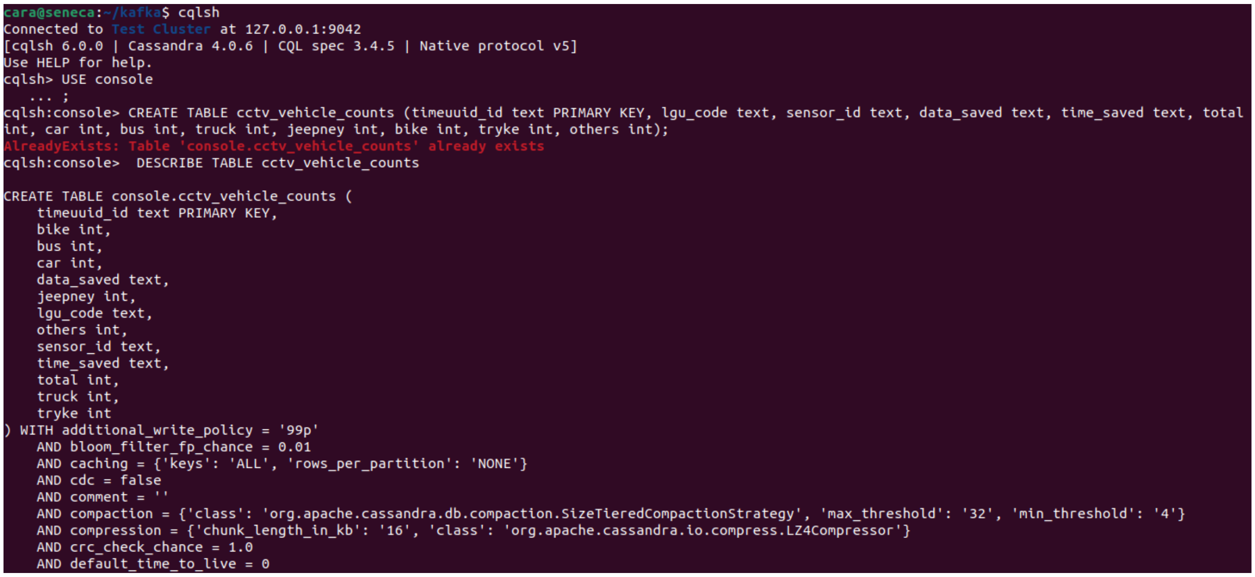
Data from different producers read by consumer

Part IV. Create a database and a table in Cassandra



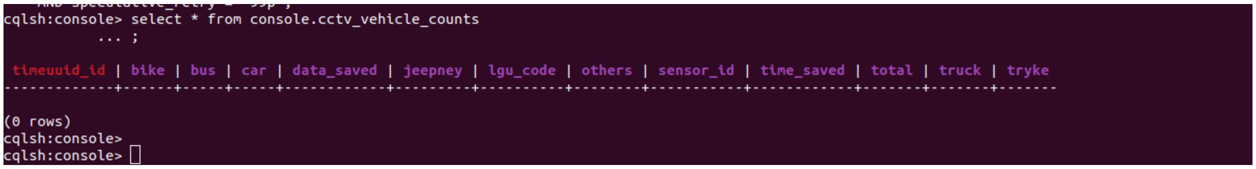
```
cara@seneca:~/kafka$ cqlsh
Connected to Test Cluster at 127.0.0.1:9042
[cqlsh 6.0.0 | Cassandra 4.0.6 | CQL spec 3.4.5 | Native protocol v5]
Use HELP for help.
cqlsh> CREATE KEYSPACE console WITH replication = {'class': 'NetworkTopologyStrategy', 'replication_factor':'1'} AND durable_writes = 'true';
AlreadyExists: Keyspace 'console' already exists
cqlsh> DESCRIBE KEYSPACES

console  system_auth         system_schema  system_views
system   system_distributed  system_traces  system_virtual_schema
```

Created Keyspace ‘console’.



```
cara@seneca:~/kafka$ cqlsh
Connected to Test Cluster at 127.0.0.1:9042
[cqlsh 6.0.0 | Cassandra 4.0.6 | CQL spec 3.4.5 | Native protocol v5]
Use HELP for help.
cqlsh> USE console
   ... ;
cqlsh:console> CREATE TABLE cctv_vehicle_counts (timeuuid_id text PRIMARY KEY, lgu_code text, sensor_id text, data_saved text, time_saved text, total
int, car int, bus int, truck int, jeepney int, bike int, tryke int, others int);
AlreadyExists: Table 'console.cctv_vehicle_counts' already exists
cqlsh:console>  DESCRIBE TABLE cctv_vehicle_counts

CREATE TABLE console.cctv_vehicle_counts (
    timeuuid_id text PRIMARY KEY,
    bike int,
    bus int,
    car int,
    data_saved text,
    jeepney int,
    lgu_code text,
    others int,
    sensor_id text,
    time_saved text,
    total int,
    truck int,
    tryke int
) WITH additional_write_policy = '99p'
    AND bloom_filter_fp_chance = 0.01
    AND caching = {'keys': 'ALL', 'rows_per_partition': 'NONE'}
    AND cdc = false
    AND comment = ''
    AND compaction = {'class': 'org.apache.cassandra.db.compaction.SizeTieredCompactionStrategy', 'max_threshold': '32', 'min_threshold': '4'}
    AND compression = {'chunk_length_in_kb': '16', 'class': 'org.apache.cassandra.io.compress.LZ4Compressor'}
    AND crc_check_chance = 1.0
    AND default_time_to_live = 0
```

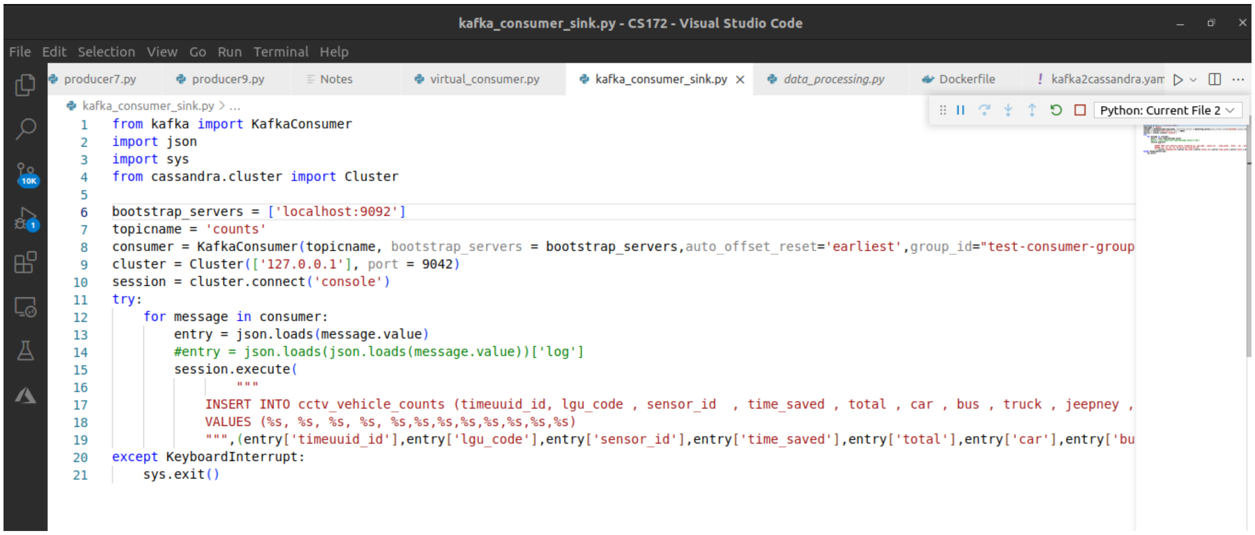
Created a table in the ‘console’ database named ‘cctv_vehicle_counts’



```
    AND speculative_retry = '99p';
cqlsh:console> select * from console.cctv_vehicle_counts
           ... ;

 timeuuid_id | bike | bus | car | data_saved | jeepney | lgu_code | others | sensor_id | time_saved | total | truck | tryke
-------------+------+-----+-----+------------+---------+----------+--------+-----------+------------+-------+-------+-------

(0 rows)
cqlsh:console>
cqlsh:console> 
```

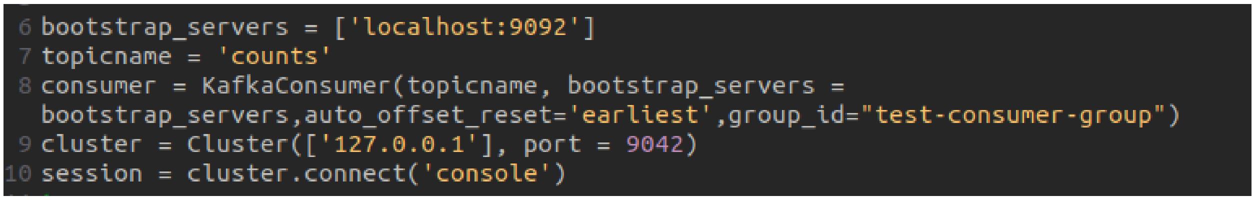
Check if table exists

Part V. Python program. Kafka to Cassandra



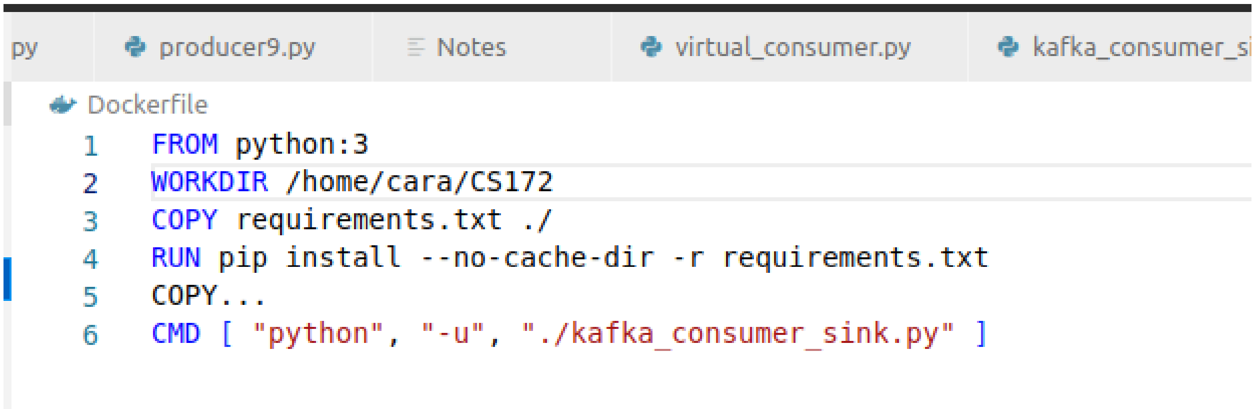
```python
from kafka import KafkaConsumer
import json
import sys
from cassandra.cluster import Cluster

bootstrap_servers = ['localhost:9092']
topicname = 'counts'
consumer = KafkaConsumer(topicname, bootstrap_servers = bootstrap_servers,auto_offset_reset='earliest',group_id="test-consumer-group
cluster = Cluster(['127.0.0.1'], port = 9042)
session = cluster.connect('console')
try:
    for message in consumer:
        entry = json.loads(message.value)
        #entry = json.loads(json.loads(message.value))['log']
        session.execute(
            """
            INSERT INTO cctv_vehicle_counts (timeuuid_id, lgu_code , sensor_id  , time_saved , total , car , bus , truck , jeepney ,
            VALUES (%s, %s, %s, %s, %s,%s,%s,%s,%s,%s,%s,%s)
            """,(entry['timeuuid_id'],entry['lgu_code'],entry['sensor_id'],entry['time_saved'],entry['total'],entry['car'],entry['bu
except KeyboardInterrupt:
    sys.exit()
```

Python code. Gets access to the cassandra database then loads values into it from kafka. Consumer sink is used here, a kafka connector.



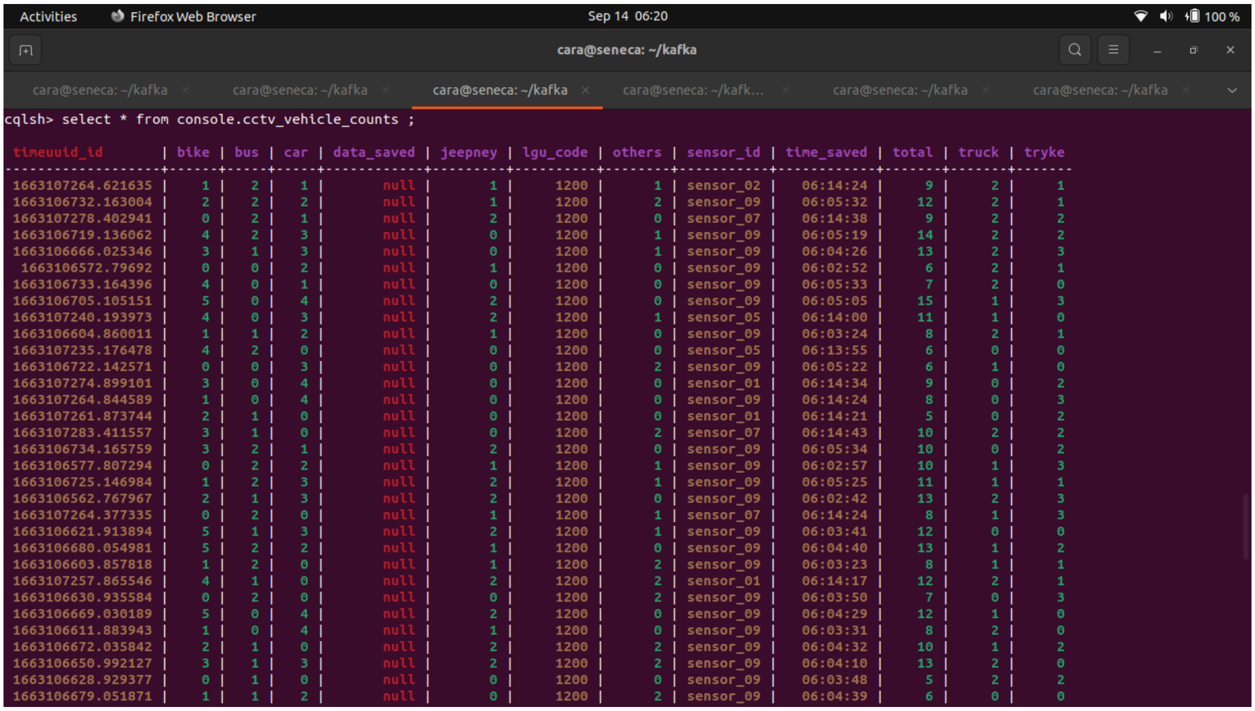
```
bootstrap_servers = ['localhost:9092']
topicname = 'counts'
consumer = KafkaConsumer(topicname, bootstrap_servers =
bootstrap_servers,auto_offset_reset='earliest',group_id="test-consumer-group")
cluster = Cluster(['127.0.0.1'], port = 9042)
session = cluster.connect('console')
```

This is the code to get access to the servers, topics, and cassandra.



```
FROM python:3
WORKDIR /home/cara/CS172
COPY requirements.txt ./
RUN pip install --no-cache-dir -r requirements.txt
COPY...
CMD [ "python", "-u", "./kafka_consumer_sink.py" ]
```

Dockerfile to run the program



```
cqlsh> select * from console.cctv_vehicle_counts ;

 timeuuid_id       | bike | bus | car | data_saved | jeepney | lgu_code | others | sensor_id | time_saved | total | truck | tryke
-------------------+------+-----+-----+------------+---------+----------+--------+-----------+------------+-------+-------+-------
 1663107264.621635 |    1 |   2 |   1 |       null |       1 |     1200 |      1 | sensor_02 |   06:14:24 |     9 |     2 |     1
 1663106732.163004 |    2 |   2 |   2 |       null |       1 |     1200 |      2 | sensor_09 |   06:05:32 |    12 |     2 |     1
 1663107278.402941 |    0 |   2 |   1 |       null |       2 |     1200 |      0 | sensor_07 |   06:14:38 |     9 |     2 |     2
 1663106719.136062 |    4 |   2 |   3 |       null |       0 |     1200 |      1 | sensor_09 |   06:05:19 |    14 |     2 |     2
 1663106666.025346 |    3 |   1 |   3 |       null |       0 |     1200 |      1 | sensor_09 |   06:04:26 |    13 |     2 |     3
  1663106572.79692 |    0 |   0 |   2 |       null |       1 |     1200 |      0 | sensor_09 |   06:02:52 |     6 |     2 |     1
 1663106733.164396 |    4 |   0 |   1 |       null |       0 |     1200 |      0 | sensor_09 |   06:05:33 |     7 |     2 |     0
 1663106705.105151 |    5 |   0 |   4 |       null |       2 |     1200 |      0 | sensor_09 |   06:05:05 |    15 |     1 |     3
 1663107240.193973 |    4 |   0 |   3 |       null |       2 |     1200 |      1 | sensor_05 |   06:14:00 |    11 |     1 |     0
 1663106604.860011 |    1 |   1 |   2 |       null |       1 |     1200 |      0 | sensor_09 |   06:03:24 |     8 |     2 |     1
 1663107235.176478 |    4 |   2 |   0 |       null |       0 |     1200 |      0 | sensor_05 |   06:13:55 |     6 |     0 |     0
 1663106722.142571 |    0 |   0 |   3 |       null |       0 |     1200 |      2 | sensor_09 |   06:05:22 |     6 |     1 |     0
 1663107274.899101 |    3 |   0 |   4 |       null |       0 |     1200 |      0 | sensor_01 |   06:14:34 |     9 |     0 |     2
 1663107264.844589 |    1 |   0 |   4 |       null |       0 |     1200 |      0 | sensor_09 |   06:14:24 |     8 |     0 |     3
 1663107261.873744 |    2 |   1 |   0 |       null |       0 |     1200 |      0 | sensor_01 |   06:14:21 |     5 |     0 |     2
 1663107283.411557 |    3 |   1 |   0 |       null |       0 |     1200 |      2 | sensor_07 |   06:14:43 |    10 |     2 |     2
 1663106734.165759 |    3 |   2 |   1 |       null |       2 |     1200 |      0 | sensor_09 |   06:05:34 |    10 |     0 |     2
 1663106577.807294 |    0 |   2 |   2 |       null |       1 |     1200 |      1 | sensor_09 |   06:02:57 |    10 |     1 |     3
 1663106725.146984 |    1 |   2 |   3 |       null |       2 |     1200 |      1 | sensor_09 |   06:05:25 |    11 |     1 |     1
 1663106562.767967 |    2 |   1 |   3 |       null |       2 |     1200 |      0 | sensor_09 |   06:02:42 |    13 |     2 |     3
 1663107264.377335 |    0 |   2 |   0 |       null |       1 |     1200 |      1 | sensor_07 |   06:14:24 |     8 |     1 |     3
 1663106621.913894 |    5 |   1 |   3 |       null |       2 |     1200 |      1 | sensor_09 |   06:03:41 |    12 |     0 |     0
 1663106680.054981 |    5 |   2 |   2 |       null |       1 |     1200 |      0 | sensor_09 |   06:04:40 |    13 |     1 |     2
 1663106603.857818 |    1 |   2 |   0 |       null |       1 |     1200 |      2 | sensor_09 |   06:03:23 |     8 |     1 |     1
 1663107257.865546 |    4 |   1 |   0 |       null |       2 |     1200 |      2 | sensor_01 |   06:14:17 |    12 |     2 |     1
 1663106630.935584 |    0 |   2 |   0 |       null |       0 |     1200 |      2 | sensor_09 |   06:03:50 |     7 |     0 |     3
 1663106669.030189 |    5 |   0 |   4 |       null |       2 |     1200 |      0 | sensor_09 |   06:04:29 |    12 |     1 |     0
 1663106611.883943 |    1 |   0 |   4 |       null |       1 |     1200 |      0 | sensor_09 |   06:03:31 |     8 |     2 |     0
 1663106672.035842 |    2 |   1 |   0 |       null |       2 |     1200 |      2 | sensor_09 |   06:04:32 |    10 |     1 |     2
 1663106650.992127 |    3 |   1 |   3 |       null |       2 |     1200 |      2 | sensor_09 |   06:04:10 |    13 |     2 |     0
 1663106628.929377 |    0 |   1 |   0 |       null |       0 |     1200 |      0 | sensor_09 |   06:03:48 |     5 |     2 |     2
 1663106679.051871 |    1 |   1 |   2 |       null |       0 |     1200 |      2 | sensor_09 |   06:04:39 |     6 |     0 |     0
```

After running different producers the ‘kafka_consumer_sink.py’ file will be run.
The data from kafka will show in the table ‘cctv_vehicle_counts’ in the ‘console’ keyspace. The image above outputs data from different producers using the *‘select *’* command.