

Producing the data (10%)

In []:

```
In [1]: # import statements
from time import sleep
from json import dumps
from kafka import KafkaProducer
import random
import datetime as dt
import csv
import pandas as pd

# A function to publish message to customer
def publish_message(producer_instance, topic_name, data):
    try:
        producer_instance.send(topic_name, data)
        #print('Message published successfully. Data: ' + str(data))
    except Exception as ex:
        print('Exception in publishing message.')
        print(str(ex))

# Connect to kafka
def connect_kafka_producer():
    _producer = None
    try:
        _producer = KafkaProducer(bootstrap_servers=['localhost:9092'],
                                   value_serializer=lambda x: dumps(x).encode('ascii'),
                                   api_version=(0, 10))
    except Exception as ex:
        print('Exception while connecting Kafka.')
        print(str(ex))
    finally:
        return _producer

# Main function
if __name__ == '__main__':

    topic_name = 'Assignment'

    print('Publishing records..')
    producer = connect_kafka_producer()

    # Using panda to read csv to a dataframe
    df = pd.read_csv("Streaming_Pedestrian_December_counts_per_hour.csv")

    # Using a for function to group by day
    for group in df.groupby(['Year', 'Month', 'Mdate']):
        print("Message published successfully. Date:", group[0][2])
        for data in group[1].iterrows():
            publish_message(producer, topic_name, data[1].to_dict()) # Convert data to a dictionary to publish

    # Sleep every 5 seconds
    sleep(5)
```

Publishing records..
Message published successfully. Date: 1
Message published successfully. Date: 2
Message published successfully. Date: 3
Message published successfully. Date: 4
Message published successfully. Date: 5
Message published successfully. Date: 6
Message published successfully. Date: 7
Message published successfully. Date: 8
Message published successfully. Date: 9
Message published successfully. Date: 10
Message published successfully. Date: 11
Message published successfully. Date: 12
Message published successfully. Date: 13
Message published successfully. Date: 14
Message published successfully. Date: 15
Message published successfully. Date: 16
Message published successfully. Date: 17
Message published successfully. Date: 18
Message published successfully. Date: 19
Message published successfully. Date: 20
Message published successfully. Date: 21
Message published successfully. Date: 22
Message published successfully. Date: 23
Message published successfully. Date: 24
Message published successfully. Date: 25
Message published successfully. Date: 26
Message published successfully. Date: 27
Message published successfully. Date: 28
Message published successfully. Date: 29
Message published successfully. Date: 30
Message published successfully. Date: 31

In []: