

Week#8_Assignment

February 26, 2021

```
[1]: import json
import uuid

from kafka import KafkaProducer, KafkaAdminClient
from kafka.admin.new_topic import NewTopic
from kafka.errors import TopicAlreadyExistsError
```

0.0.1 Configuration Parameters

TODO: Change the configuration parameters to the appropriate values for your setup.

```
[2]: config = dict(
    bootstrap_servers=['kafka.kafka.svc.cluster.local:9092'],
    first_name='Sohan',
    last_name='Thakur'
)

config['client_id'] = '{}{}'.format(
    config['last_name'],
    config['first_name']
)
config['topic_prefix'] = '{}{}'.format(
    config['last_name'],
    config['first_name']
)

config
```

```
[2]: {'bootstrap_servers': ['kafka.kafka.svc.cluster.local:9092'],
      'first_name': 'Sohan',
      'last_name': 'Thakur',
      'client_id': 'ThakurSohan',
      'topic_prefix': 'ThakurSohan'}
```

0.0.2 Create Topic Utility Function

The `create_kafka_topic` helps create a Kafka topic based on your configuration settings. For instance, if your first name is *John* and your last name is *Doe*, `create_kafka_topic('locations')`

will create a topic with the name DoeJohn-locations. The function will not create the topic if it already exists.

```
[3]: def create_kafka_topic(topic_name, config=config, num_partitions=1,
    ↪ replication_factor=1):
    bootstrap_servers = config['bootstrap_servers']
    client_id = config['client_id']
    topic_prefix = config['topic_prefix']
    name = '{}-{}'.format(topic_prefix, topic_name)

    admin_client = KafkaAdminClient(
        bootstrap_servers=bootstrap_servers,
        client_id=client_id
    )

    topic = NewTopic(
        name=name,
        num_partitions=num_partitions,
        replication_factor=replication_factor
    )

    topic_list = [topic]
    try:
        admin_client.create_topics(new_topics=topic_list)
        print('Created topic "{}"'.format(name))
    except TopicAlreadyExistsError as e:
        print('Topic "{}" already exists'.format(name))

create_kafka_topic('locations')
```

Created topic "ThakurSohan-locations"

0.0.3 Kafka Producer

The following code creates a `KafkaProducer` object which you can use to send Python objects that are serialized as JSON.

Note: This producer serializes Python objects as JSON. This means that object must be JSON serializable. As an example, Python `DateTime` values are not JSON serializable and must be converted to a string (e.g. ISO 8601) or a numeric value (e.g. a Unix timestamp) before being sent.

```
[4]: producer = KafkaProducer(
    bootstrap_servers=config['bootstrap_servers'],
    value_serializer=lambda x: json.dumps(x).encode('utf-8')
)
```

0.0.4 Send Data Function

The `send_data` function sends a Python object to a Kafka topic. This function adds the `topic_prefix` to the topic so `send_data('locations', data)` sends a JSON serialized message to `DoeJohn-locations`. The function also registers callbacks to let you know if the message has been sent or if an error has occurred.

```
[5]: def on_send_success(record_metadata):
    print('Message sent:\n    Topic: "{}"\n    Partition: {}\n    Offset: {}'.
    ↪format(
        record_metadata.topic,
        record_metadata.partition,
        record_metadata.offset
    ))

def on_send_error(excp):
    print('I am an errback', exc_info=excp)
    # handle exception

def send_data(topic, data, config=config, producer=producer, msg_key=None):
    topic_prefix = config['topic_prefix']
    topic_name = '{}-{}'.format(topic_prefix, topic)

    if msg_key is not None:
        key = msg_key
    else:
        key = uuid.uuid4().hex

    producer.send(
        topic_name,
        value=data,
        key=key.encode('utf-8')
    ).add_callback(on_send_success).add_errback(on_send_error)
```

```
[6]: example_data = dict(
    key1='value1',
    key2='value2'
)

send_data('locations', example_data)
```

```
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 0
```

```
[25]: import os
import pandas as pd
```

```

os.getcwd()
import datetime
import threading
event = threading.Event()
print(datetime.datetime.now())
base_loc= '/home/jovyan/dsc650/data/processed/bdd/'
location_dir = base_loc+'locations/'
acceleration_dir = base_loc + 'accelerations/'

```

2021-02-26 12:00:42.529987

```

[73]: time_dir = os.listdir(location_dir)
time_list = list(set([float(x.split('=')[1]) for x in time_dir]))
time_dict = {}
for t in time_dir:
    time_dict[float(t.split("=")[1])] = t

```

```

[74]: def get_time_list_dict(folder_name):
    f_dir = base_loc+folder_name
    time_dir = os.listdir(f_dir)
    time_list = list(set([float(x.split('=')[1]) for x in time_dir]))
    time_dict = {}
    for t in time_dir:
        time_dict[float(t.split("=")[1])] = t
    return time_list,time_dict,f_dir

```

```

[75]: def get_json(directory):
    df = pd.read_parquet(final_dir)
    return df.to_json()

```

```

[78]: def send_file(file_type):
    i = 0.0
    time_list,time_dict,f_dir = get_time_list_dict(file_type)
    for t in time_list:
        final_dir = f_dir+'/' + time_dict[t] + '/'
        op_json = get_json(final_dir)
        event.wait(t - i)
        s_t = datetime.datetime.now()
        print("Sending file {}".format(final_dir))
        #print("write now time is {}, which is {} s from start time".
        ↪format(datetime.datetime.now(),datetime.datetime.now()-s_t))
        i = t
        send_data('locations', op_json)

```

```

[79]: send_file('locations')

```

Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=000.0/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 19
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=004.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 20
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=007.8/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 21
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=010.6/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 22
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=014.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 23
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=017.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 24
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=021.3/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 25
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=026.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 26
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=030.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 27
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=033.7/
Message sent:
Topic: "ThakurSohan-locations"

Partition: 0
Offset: 28
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=037.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 29
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=041.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 30
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=045.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 31
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=049.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 32
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=052.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 33
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=056.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 34
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=060.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 35
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=063.8/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 36
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=066.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 37
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=070.9/

Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 38
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=073.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 39
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=077.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 40
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=081.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 41
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=085.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 42
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=088.3/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 43
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=091.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 44
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=094.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 45
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=098.8/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 46
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=102.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0

```
Offset: 47
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=106.0/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 48
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=109.9/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 49
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=113.2/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 50
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=117.2/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 51
Sending file /home/jovyan/dsc650/data/processed/bdd/locations/t=121.4/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 52
```

```
[81]: send_file('accelerations')
```

```
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=000.0/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 56
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=004.5/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 57
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=007.8/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 58
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=010.6/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
```


Offset: 59
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=014.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 60
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=017.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 61
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=021.3/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 62
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=026.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 63
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=030.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 64
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=033.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 65
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=037.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 66
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=041.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 67
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=045.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 68
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=049.5/
Message sent:

Topic: "ThakurSohan-locations"
Partition: 0
Offset: 69
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=052.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 70
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=056.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 71
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=060.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 72
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=063.8/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 73
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=066.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 74
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=070.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 75
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=073.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 76
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=077.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 77
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=081.4/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 78

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=085.1/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 79

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=088.3/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 80

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=091.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 81

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=094.7/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 82

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=098.8/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 83

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=102.5/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 84

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=106.0/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 85

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=109.9/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 86

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=113.2/
Message sent:
Topic: "ThakurSohan-locations"
Partition: 0
Offset: 87

Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=117.2/
Message sent:
Topic: "ThakurSohan-locations"

```
Partition: 0
Offset: 88
Sending file /home/jovyan/dsc650/data/processed/bdd/accelerations/t=121.4/
Message sent:
  Topic: "ThakurSohan-locations"
  Partition: 0
  Offset: 89
```

[47]:

[]: