

```
!pip install kafka-python
```

```

Collecting kafka-python
  Downloading kafka_python-2.0.2-py2.py3-none-any.whl (246 kB)
    _____ 246.5/246.5 kB 4.0 MB/s eta 0
Installing collected packages: kafka-python
Successfully installed kafka-python-2.0.2

```

```
!pip install avro-python3
```

```

Collecting avro-python3
  Downloading avro-python3-1.10.2.tar.gz (38 kB)
  Preparing metadata (setup.py) ... done
Building wheels for collected packages: avro-python3
  Building wheel for avro-python3 (setup.py) ... done
  Created wheel for avro-python3: filename=avro_python3-1.10.2-py3-none-any
  Stored in directory: /root/.cache/pip/wheels/bc/85/62/6cdd81c56f923946b40
Successfully built avro-python3
Installing collected packages: avro-python3
Successfully installed avro-python3-1.10.2

```

```

# import required libraries
from kafka import KafkaConsumer

```

```

import avro.schema
import avro.io
import io

```

```
kafka_broker = 'lab.aimet.tech:9092'
```

```
!wget https://raw.githubusercontent.com/pvateekul/2110531_DSDE_2023s1/main/code/
```

```

--2023-10-23 10:59:12-- https://raw.githubusercontent.com/pvateekul/2110531_DSDE_2023s1/main/code/submit.avsc
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.108.133
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.108.133|:443
HTTP request sent, awaiting response... 200 OK
Length: 231 [text/plain]
Saving to: 'submit.avsc'

```

```
submit.avsc          100%[=====>]          231  --.-KB/s    in 0s
```

```
2023-10-23 10:59:12 (5.46 MB/s) - 'submit.avsc' saved [231/231]
```

```
!wget https://raw.githubusercontent.com/pvateekul/2110531_DSDE_2023s1/main/code/
```

```
--2023-10-23 11:22:49-- https://raw.githubusercontent.com/pvateekul/2110531_DSDE_2023s1/main/code/
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.108.153
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.108.153:
HTTP request sent, awaiting response... 200 OK
Length: 359 [text/plain]
Saving to: 'result.avsc'
```

```
result.avsc          100%[=====>]          359  --.-KB/s    in 0s
```

```
2023-10-23 11:22:49 (8.38 MB/s) - 'result.avsc' saved [359/359]
```

```
schema_file = '/content/result.avsc'
schema = avro.schema.parse(open(schema_file).read())
```

```
def deserialize(schema, raw_bytes):
    bytes_reader = io.BytesIO(raw_bytes)
    decoder = avro.io.BinaryDecoder(bytes_reader)
    reader = avro.io.DatumReader(schema)
    return reader.read(decoder)
```

```
# Connect to kafka broker running in your local host (docker). Change this to y
kafka_broker = 'lab.aimet.tech:9092'
```

```
consumer = KafkaConsumer(
    'result',
    bootstrap_servers=[kafka_broker],
    enable_auto_commit=True,
    value_deserializer=lambda x: deserialize(schema, x))
```

```
WARNING:kafka.coordinator.consumer:group_id is None: disabling auto-commit.
```

```
print('Running Consumer with AVRO')
for message in consumer:
    print(message.value)
```

```
{'timestamp': 1698061000, 'vid': 'V878613', 'txid': 'TX00335', 'checksum':
{'timestamp': 1698061002, 'vid': 'V878613', 'txid': 'TX09388', 'checksum':
{'timestamp': 1698061003, 'vid': 'V926090', 'txid': 'TX06293', 'checksum':
{'timestamp': 1698061004, 'vid': 'V878613', 'txid': 'TX00447', 'checksum':
{'timestamp': 1698061006, 'vid': 'V878613', 'txid': 'TX09776', 'checksum':
{'timestamp': 1698061008, 'vid': 'V878613', 'txid': 'TX07238', 'checksum':
{'timestamp': 1698061010, 'vid': 'V878613', 'txid': 'TX00709', 'checksum':
{'timestamp': 1698061012, 'vid': 'V878613', 'txid': 'TX03530', 'checksum':
```

```
{ 'timestamp': 1698061013, 'vid': 'V926090', 'txid': 'TX00598', 'checksum':
{ 'timestamp': 1698061014, 'vid': 'V878613', 'txid': 'TX05943', 'checksum':
{ 'timestamp': 1698061016, 'vid': 'V878613', 'txid': 'TX00259', 'checksum':
{ 'timestamp': 1698061018, 'vid': 'V878613', 'txid': 'TX00388', 'checksum':
{ 'timestamp': 1698061020, 'vid': 'V878613', 'txid': 'TX01822', 'checksum':
{ 'timestamp': 1698061021, 'vid': 'V926090', 'txid': 'TX09448', 'checksum':
{ 'timestamp': 1698061022, 'vid': 'V878613', 'txid': 'TX04966', 'checksum':
{ 'timestamp': 1698061024, 'vid': 'V878613', 'txid': 'TX07371', 'checksum':
{ 'timestamp': 1698061026, 'vid': 'V878613', 'txid': 'TX08013', 'checksum':
{ 'timestamp': 1698061029, 'vid': 'V878613', 'txid': 'TX08337', 'checksum':
{ 'timestamp': 1698061031, 'vid': 'V878613', 'txid': 'TX01597', 'checksum':
{ 'timestamp': 1698061032, 'vid': 'V926090', 'txid': 'TX03442', 'checksum':
{ 'timestamp': 1698061033, 'vid': 'V878613', 'txid': 'TX01797', 'checksum':
{ 'timestamp': 1698061035, 'vid': 'V878613', 'txid': 'TX02319', 'checksum':
{ 'timestamp': 1698061037, 'vid': 'V878613', 'txid': 'TX06854', 'checksum':
{ 'timestamp': 1698061038, 'vid': 'V926090', 'txid': 'TX09155', 'checksum':
{ 'timestamp': 1698061039, 'vid': 'V878613', 'txid': 'TX05696', 'checksum':
{ 'timestamp': 1698061041, 'vid': 'V878613', 'txid': 'TX06293', 'checksum':
{ 'timestamp': 1698061043, 'vid': 'V878613', 'txid': 'TX00598', 'checksum':
{ 'timestamp': 1698061045, 'vid': 'V878613', 'txid': 'TX09448', 'checksum':
{ 'timestamp': 1698061046, 'vid': 'V926090', 'txid': 'TX03174', 'checksum':
{ 'timestamp': 1698061047, 'vid': 'V878613', 'txid': 'TX03442', 'checksum':
{ 'timestamp': 1698061049, 'vid': 'V878613', 'txid': 'TX09155', 'checksum':
{ 'timestamp': 1698061051, 'vid': 'V878613', 'txid': 'TX03174', 'checksum':
{ 'timestamp': 1698061052, 'vid': 'V926090', 'txid': 'TX04784', 'checksum':
{ 'timestamp': 1698061053, 'vid': 'V878613', 'txid': 'TX04784', 'checksum':
{ 'timestamp': 1698061057, 'vid': 'V926090', 'txid': 'TX05957', 'checksum':
{ 'timestamp': 1698061058, 'vid': 'V878613', 'txid': 'TX05957', 'checksum':
{ 'timestamp': 1698061064, 'vid': 'V878613', 'txid': 'TX04069', 'checksum':
{ 'timestamp': 1698061065, 'vid': 'V926090', 'txid': 'TX04069', 'checksum':
{ 'timestamp': 1698061074, 'vid': 'V926090', 'txid': 'TX02753', 'checksum':
{ 'timestamp': 1698061083, 'vid': 'V926090', 'txid': 'TX01818', 'checksum':
{ 'timestamp': 1698061089, 'vid': 'V926090', 'txid': 'TX04837', 'checksum':
{ 'timestamp': 1698061094, 'vid': 'V926090', 'txid': 'TX08336', 'checksum':
{ 'timestamp': 1698061104, 'vid': 'V926090', 'txid': 'TX09556', 'checksum':
{ 'timestamp': 1698061110, 'vid': 'V926090', 'txid': 'TX06944', 'checksum':
{ 'timestamp': 1698061116, 'vid': 'V926090', 'txid': 'TX07687', 'checksum':
{ 'timestamp': 1698061126, 'vid': 'V926090', 'txid': 'TX02268', 'checksum':
{ 'timestamp': 1698061131, 'vid': 'V926090', 'txid': 'TX01416', 'checksum':
{ 'timestamp': 1698061137, 'vid': 'V926090', 'txid': 'TX00030', 'checksum':
{ 'timestamp': 1698061157, 'vid': 'V878613', 'txid': 'TX02753', 'checksum':
{ 'timestamp': 1698061159, 'vid': 'V878613', 'txid': 'TX01818', 'checksum':
{ 'timestamp': 1698061161, 'vid': 'V878613', 'txid': 'TX04837', 'checksum':
{ 'timestamp': 1698061163, 'vid': 'V878613', 'txid': 'TX08336', 'checksum':
{ 'timestamp': 1698061165, 'vid': 'V878613', 'txid': 'TX09556', 'checksum':
{ 'timestamp': 1698061167, 'vid': 'V878613', 'txid': 'TX06944', 'checksum':
{ 'timestamp': 1698061169, 'vid': 'V878613', 'txid': 'TX07687', 'checksum':
{ 'timestamp': 1698061171, 'vid': 'V878613', 'txid': 'TX02268', 'checksum':
{ 'timestamp': 1698061173, 'vid': 'V878613', 'txid': 'TX01416', 'checksum':
{ 'timestamp': 1698061175, 'vid': 'V878613', 'txid': 'TX00030', 'checksum':
{ 'timestamp': 1698061177, 'vid': 'V878613', 'txid': 'TX00801', 'checksum':
```

