

# Day18 - Day19 Assignments

Bhargav Jangam

## Exercise 1

```
User Details Dataset:
24/11/29 22:19:20 INFO CodeGenerator: Code generated in 4.039542 ms
24/11/29 22:19:20 INFO CodeGenerator: Code generated in 3.26775 ms
+-----+-----+
|user_id|user_name|
+-----+-----+
|      1|   Alice|
|      2|    Bob|
|      3|  Charlie|
|      4|   David|
|      5|    Eve|
|      6|   Frank|
|      7|   Grace|
|      8|   Helen|
|      9|    Ivy|
|     10|   Jack|
+-----+-----+
```

```
Transaction Logs Dataset:
24/11/29 22:19:20 INFO CodeGenerator: Code generated in 4.052792 ms
24/11/29 22:19:20 INFO CodeGenerator: Code generated in 4.9725 ms
+-----+-----+-----+
|user_id|transaction_date|amount|
+-----+-----+-----+
|      1|    2024-01-1|    0.0|
|      2|    2024-01-2|   100.0|
|      3|    2024-01-3|   200.0|
|      4|    2024-01-4|   300.0|
|      5|    2024-01-5|   400.0|
|      6|    2024-01-6|   500.0|
|      7|    2024-01-7|   600.0|
|      8|    2024-01-8|   700.0|
|      9|    2024-01-9|   800.0|
|     10|    2024-01-10|  900.0|
|     11|    2024-01-11| 1000.0|
|     12|    2024-01-12| 1100.0|
|     13|    2024-01-13| 1200.0|
|     14|    2024-01-14| 1300.0|
|     15|    2024-01-15| 1400.0|
|     16|    2024-01-16| 1500.0|
|     17|    2024-01-17| 1600.0|
|     18|    2024-01-18| 1700.0|
|     19|    2024-01-19| 1800.0|
|     20|    2024-01-20| 1900.0|
+-----+-----+-----+
only showing top 20 rows
```

```
Time taken for join without broadcasting: 4.153107958 seconds
```

```
Time taken for join with broadcasting: 2.901073958 seconds
```

We observed that the time taken with broadcasting is significantly less compared to without broadcasting. While this difference is noticeable with just a few hundred rows, the impact becomes substantial when dealing with real-time data involving millions of rows.

## Exercise 2

```
Sales Dataset:
```

```
24/11/29 22:43:32 INFO CodeGenerator: Code generated in 4.385125 ms
```

```
24/11/29 22:43:32 INFO CodeGenerator: Code generated in 4.981875 ms
```

```
+-----+-----+-----+-----+
|sale_id| product|amount|      date|
+-----+-----+-----+-----+
|      1|ProductA| 100.0|2024-01-01|
|      2|ProductB| 200.0|2024-01-02|
|      3|ProductA| 150.0|2024-01-03|
|      4|ProductC| 300.0|2024-01-04|
|      5|ProductB| 250.0|2024-01-05|
|      6|ProductA| 120.0|2024-01-06|
|      7|ProductC| 350.0|2024-01-07|
+-----+-----+-----+-----+
```

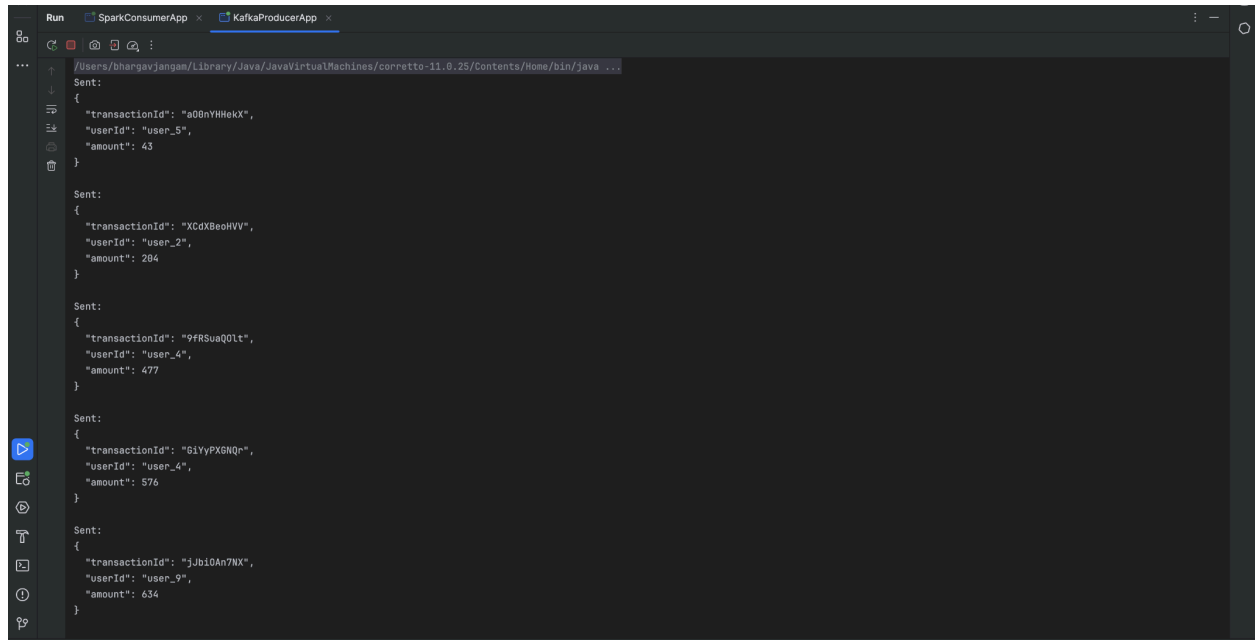
```
Time taken without caching: 0.586936208 seconds
```

```
Time taken with caching: 0.024308375 seconds
```

We observed that the time taken with caching is significantly reduced compared to without caching. While this difference is noticeable with smaller datasets, the impact becomes especially substantial when dealing with large datasets or multiple stages in a pipeline.

# Exercise 3

## KafkaProducerApp



```
Run SparkConsumerApp x KafkaProducerApp x
/Users/bhargavjangan/Library/Java/JavaVirtualMachines/corretto-11.0.25/Contents/Home/bin/java ...
Sent:
{
  "transactionId": "a08nYHheK",
  "userId": "user_5",
  "amount": 43
}

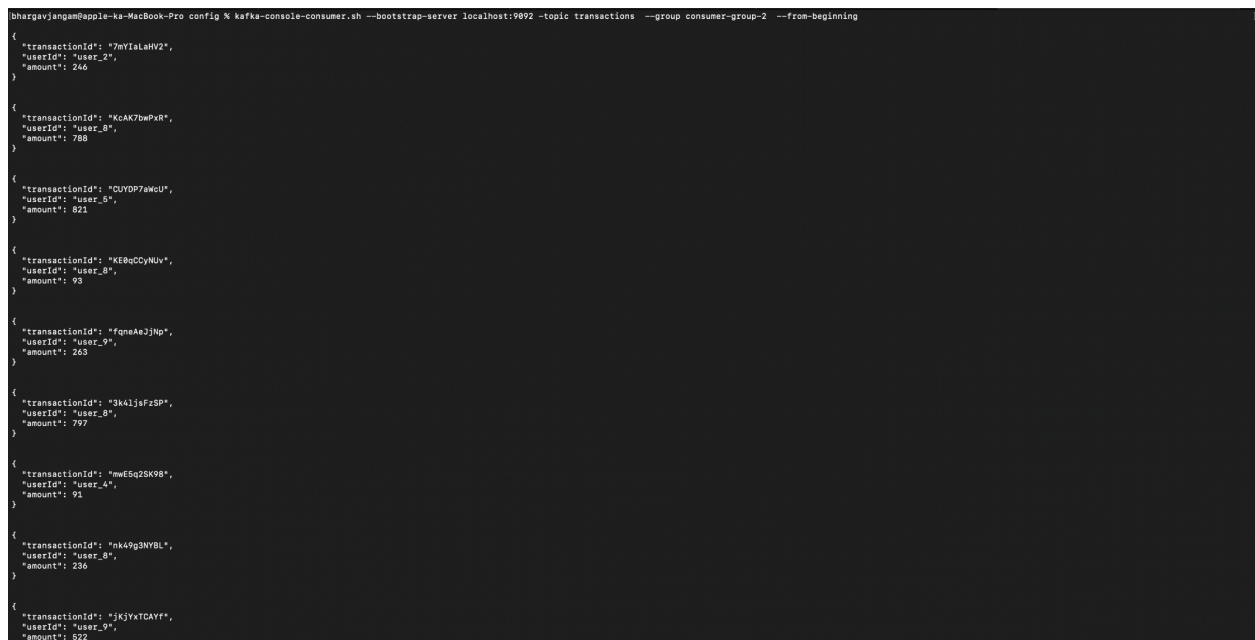
Sent:
{
  "transactionId": "XCdX8eohVV",
  "userId": "user_2",
  "amount": 284
}

Sent:
{
  "transactionId": "9fRSuaQ0It",
  "userId": "user_4",
  "amount": 477
}

Sent:
{
  "transactionId": "GiyYPXGNQ",
  "userId": "user_4",
  "amount": 576
}

Sent:
{
  "transactionId": "jJbi0An7NX",
  "userId": "user_9",
  "amount": 634
}
```

## Messages in the console consumer



```
bhargavjangan@apple-ka-MacBook-Pro config % kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic transactions --group consumer-group-2 --from-beginning
{"transactionId": "7mYtAlaHV2",
 "userId": "user_2",
 "amount": 246
}

{"transactionId": "KcAK7bPxr",
 "userId": "user_8",
 "amount": 788
}

{"transactionId": "OUYDP7wMcU",
 "userId": "user_6",
 "amount": 821
}

{"transactionId": "KE@qCCyNUv",
 "userId": "user_8",
 "amount": 93
}

{"transactionId": "fQneAeJjNp",
 "userId": "user_9",
 "amount": 263
}

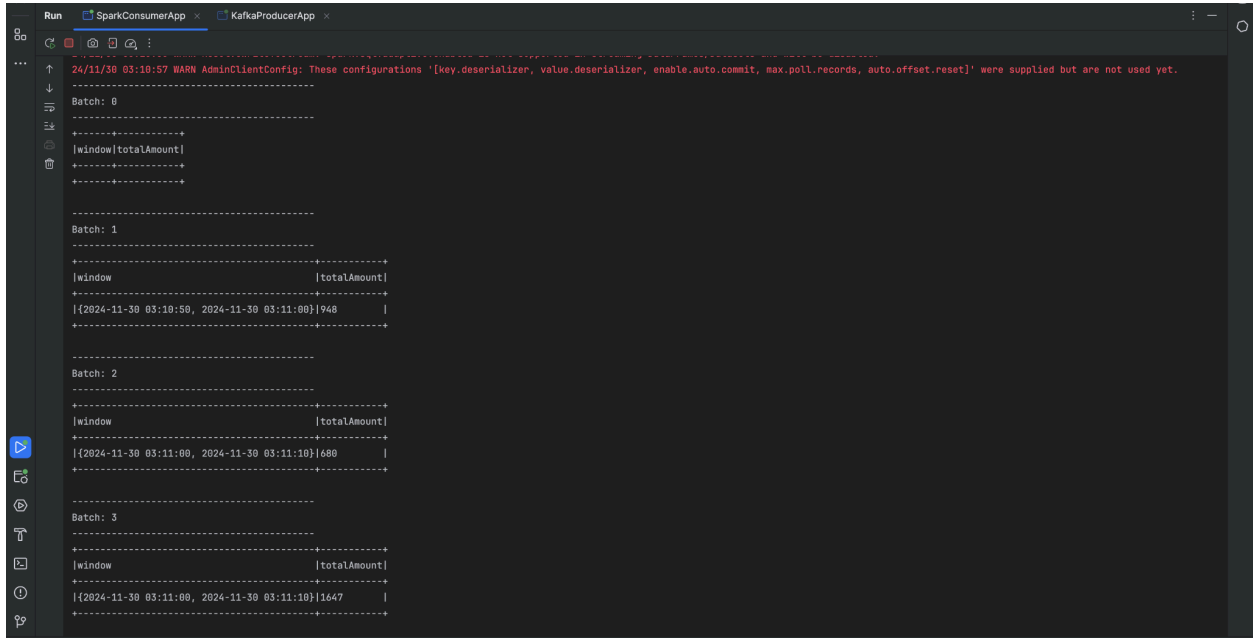
{"transactionId": "3k4jjsFzSP",
 "userId": "user_8",
 "amount": 797
}

{"transactionId": "mwEsq2SK9B",
 "userId": "user_4",
 "amount": 91
}

{"transactionId": "nk49g3MYBL",
 "userId": "user_8",
 "amount": 236
}

{"transactionId": "jKjYxTDAV*",
 "userId": "user_9",
 "amount": 622
}
```

# SparkConsumerApp



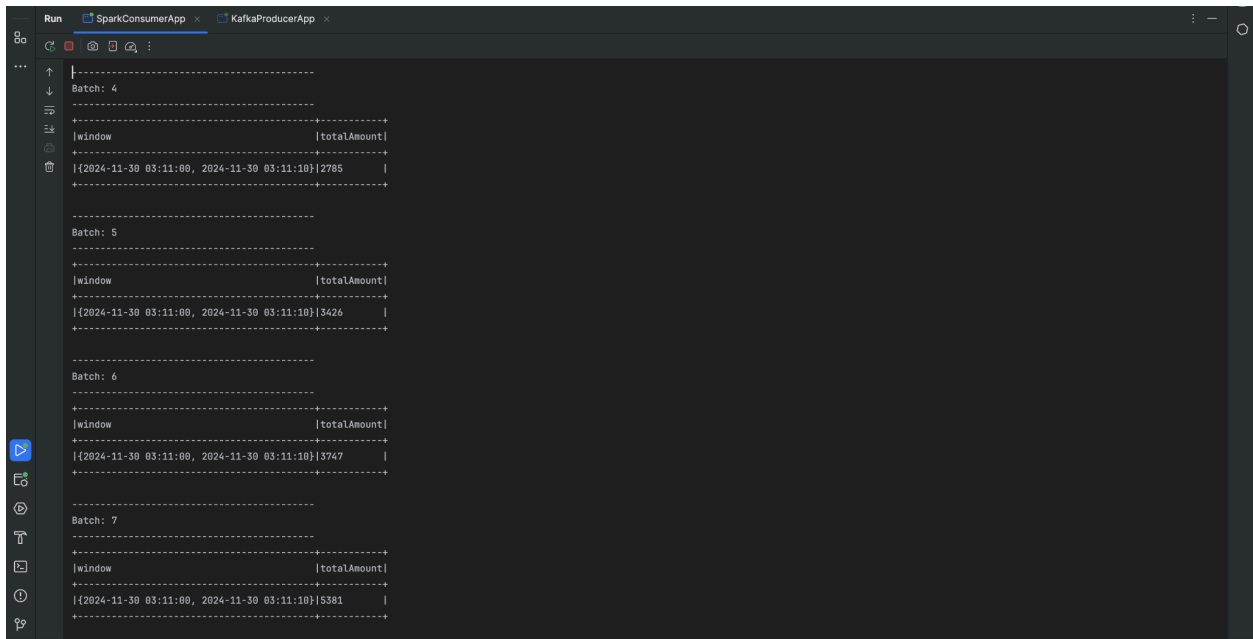
The screenshot shows the console output of the SparkConsumerApp. It displays four batches of data, each with a window of time and a total amount. The output is as follows:

```
Batch: 0
-----
|window|totalAmount|
-----
-----

Batch: 1
-----
|window|totalAmount|
-----
|{2024-11-30 03:10:50, 2024-11-30 03:11:00}|948|
-----

Batch: 2
-----
|window|totalAmount|
-----
|{2024-11-30 03:11:00, 2024-11-30 03:11:10}|680|
-----

Batch: 3
-----
|window|totalAmount|
-----
|{2024-11-30 03:11:00, 2024-11-30 03:11:10}|1647|
-----
```



The screenshot shows the console output of the SparkConsumerApp, continuing from the previous batch. It displays four more batches of data, each with a window of time and a total amount. The output is as follows:

```
Batch: 4
-----
|window|totalAmount|
-----
|{2024-11-30 03:11:00, 2024-11-30 03:11:10}|2785|
-----

Batch: 5
-----
|window|totalAmount|
-----
|{2024-11-30 03:11:00, 2024-11-30 03:11:10}|3426|
-----

Batch: 6
-----
|window|totalAmount|
-----
|{2024-11-30 03:11:00, 2024-11-30 03:11:10}|3747|
-----

Batch: 7
-----
|window|totalAmount|
-----
|{2024-11-30 03:11:00, 2024-11-30 03:11:10}|5381|
-----
```

## Exercise 4

Original Dataframe in the Parquet Format which is in the input folder of case\_study\_4 for Exercise 4 ,generated by GenerateParquetGcs.

```
+---+-----+-----+
| id|  name|   status|
+---+-----+-----+
| 46|  Tina|in_progress|
| 47|   Uma|  completed|
| 48| Victor|   pending|
| 49| Wendy|  completed|
| 50| Xavier|  completed|
| 21|   Uma|   pending|
| 22| Victor|  completed|
| 23| Wendy|in_progress|
| 24| Xavier|  completed|
| 25|  Yara|   pending|
|  1| Alice|  completed|
|  2|  Bob|   pending|
|  3| Charlie|  completed|
|  4| David|in_progress|
| 30| Diana|  completed|
| 31|  Eve|  completed|
| 32| Frank|in_progress|
| 33| Grace|   pending|
| 17| Quinn|in_progress|
| 18|  Rose|   pending|
+---+-----+-----+
only showing top 20 rows
```

| bhargav-assignments                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|--------------------------|--------------------------|------|------|------|---------|----------|-----|--------------------------|--------------------------|-----------------------------------|------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|-------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|---------|--------------------------|--------------------------|-----------------------------------|-------|--------------------------|--------------------------|-----------------------------------|------|--------------------------|--------------------------|
| Location                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Storage class | Public access            | Protection               |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| us (multiple regions in United States)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Standard      | Not public               | Soft Delete              |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| OBJECTS CONFIGURATION PERMISSIONS PROTECTION LIFECYCLE OBSERVABILITY INVENTORY REPORTS OPERATIONS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| Folder browser                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| Buckets > bhargav-assignments > Day18And19Task > case_study_4 > input                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| CREATE FOLDER UPLOAD TRANSFER DATA OTHER SERVICES                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| Filter by name prefix only Filter objects and folders Show Live objects only                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| <table><thead><tr><th>Name</th><th>Size</th><th>Type</th><th>Created</th></tr></thead><tbody><tr><td>_SUCCESS</td><td>0 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 2:00:54 AM</td></tr><tr><td>part-00000-9bdc18ba-829a-46c9-...</td><td>1 KB</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:45 AM</td></tr><tr><td>part-00001-9bdc18ba-829a-46c9-...</td><td>1,014 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:26 AM</td></tr><tr><td>part-00002-9bdc18ba-829a-46c9-...</td><td>1,011 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:33 AM</td></tr><tr><td>part-00003-9bdc18ba-829a-46c9-...</td><td>996 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 2:00:12 AM</td></tr><tr><td>part-00004-9bdc18ba-829a-46c9-...</td><td>1,018 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:59 AM</td></tr><tr><td>part-00005-9bdc18ba-829a-46c9-...</td><td>1 KB</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:08 AM</td></tr><tr><td>part-00006-9bdc18ba-829a-46c9-...</td><td>1,018 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 2:00:29 AM</td></tr><tr><td>part-00007-9bdc18ba-829a-46c9-...</td><td>1,023 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:17 AM</td></tr><tr><td>part-00008-9bdc18ba-829a-46c9-...</td><td>1,000 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 2:00:39 AM</td></tr><tr><td>part-00009-9bdc18ba-829a-46c9-...</td><td>1,015 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:58:57 AM</td></tr><tr><td>part-00010-9bdc18ba-829a-46c9-...</td><td>995 B</td><td>application/octet-stream</td><td>Nov 30, 2024, 1:59:02 AM</td></tr><tr><td>part-00011-9bdc18ba-829a-46c9-...</td><td>1 KB</td><td>application/octet-stream</td><td>Nov 30, 2024, 2:00:21 AM</td></tr></tbody></table> |               |                          |                          | Name | Size | Type | Created | _SUCCESS | 0 B | application/octet-stream | Nov 30, 2024, 2:00:54 AM | part-00000-9bdc18ba-829a-46c9-... | 1 KB | application/octet-stream | Nov 30, 2024, 1:59:45 AM | part-00001-9bdc18ba-829a-46c9-... | 1,014 B | application/octet-stream | Nov 30, 2024, 1:59:26 AM | part-00002-9bdc18ba-829a-46c9-... | 1,011 B | application/octet-stream | Nov 30, 2024, 1:59:33 AM | part-00003-9bdc18ba-829a-46c9-... | 996 B | application/octet-stream | Nov 30, 2024, 2:00:12 AM | part-00004-9bdc18ba-829a-46c9-... | 1,018 B | application/octet-stream | Nov 30, 2024, 1:59:59 AM | part-00005-9bdc18ba-829a-46c9-... | 1 KB | application/octet-stream | Nov 30, 2024, 1:59:08 AM | part-00006-9bdc18ba-829a-46c9-... | 1,018 B | application/octet-stream | Nov 30, 2024, 2:00:29 AM | part-00007-9bdc18ba-829a-46c9-... | 1,023 B | application/octet-stream | Nov 30, 2024, 1:59:17 AM | part-00008-9bdc18ba-829a-46c9-... | 1,000 B | application/octet-stream | Nov 30, 2024, 2:00:39 AM | part-00009-9bdc18ba-829a-46c9-... | 1,015 B | application/octet-stream | Nov 30, 2024, 1:58:57 AM | part-00010-9bdc18ba-829a-46c9-... | 995 B | application/octet-stream | Nov 30, 2024, 1:59:02 AM | part-00011-9bdc18ba-829a-46c9-... | 1 KB | application/octet-stream | Nov 30, 2024, 2:00:21 AM |
| Name                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Size          | Type                     | Created                  |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| _SUCCESS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 0 B           | application/octet-stream | Nov 30, 2024, 2:00:54 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00000-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1 KB          | application/octet-stream | Nov 30, 2024, 1:59:45 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00001-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,014 B       | application/octet-stream | Nov 30, 2024, 1:59:26 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00002-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,011 B       | application/octet-stream | Nov 30, 2024, 1:59:33 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00003-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 996 B         | application/octet-stream | Nov 30, 2024, 2:00:12 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00004-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,018 B       | application/octet-stream | Nov 30, 2024, 1:59:59 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00005-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1 KB          | application/octet-stream | Nov 30, 2024, 1:59:08 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00006-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,018 B       | application/octet-stream | Nov 30, 2024, 2:00:29 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00007-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,023 B       | application/octet-stream | Nov 30, 2024, 1:59:17 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00008-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,000 B       | application/octet-stream | Nov 30, 2024, 2:00:39 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00009-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1,015 B       | application/octet-stream | Nov 30, 2024, 1:58:57 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00010-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 995 B         | application/octet-stream | Nov 30, 2024, 1:59:02 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| part-00011-9bdc18ba-829a-46c9-...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1 KB          | application/octet-stream | Nov 30, 2024, 2:00:21 AM |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |
| Rows per page: 50 1 - 13 of 13                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |               |                          |                          |      |      |      |         |          |     |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |         |                          |                          |                                   |      |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |         |                          |                          |                                   |       |                          |                          |                                   |      |                          |                          |

Filtered Dataframe in the parquet Format which is in the output folder of case\_study\_4 for Exercise 4.

```

+---+-----+-----+
| id | name | status |
+---+-----+-----+
| 47 | Uma | completed |
| 49 | Wendy | completed |
| 50 | Xavier | completed |
| 22 | Victor | completed |
| 24 | Xavier | completed |
| 1 | Alice | completed |
| 3 | Charlie | completed |
| 30 | Diana | completed |
| 31 | Eve | completed |
| 19 | Sam | completed |
| 20 | Tina | completed |
| 26 | Zara | completed |
| 28 | Bella | completed |
| 39 | Mona | completed |
| 40 | Nick | completed |
| 5 | Eva | completed |
| 9 | Grace | completed |
| 7 | Ivy | completed |
| 11 | Karen | completed |
| 34 | Hank | completed |
+---+-----+-----+
only showing top 20 rows

```

**bhargav-assignments**

**Location**  
us (multiple regions in United States)

**Storage class**  
Standard

**Public access**  
Not public

**Protection**  
Soft Delete

**OBJECTS**    CONFIGURATION    PERMISSIONS    PROTECTION    LIFECYCLE    OBSERVABILITY    INVENTORY REPORTS    OPERATIONS

**Folder browser**

Buckets > bhargav-assignments > Day18And19Task > case\_study\_4 > output 📄

CREATE FOLDER    UPLOAD    TRANSFER DATA    OTHER SERVICES

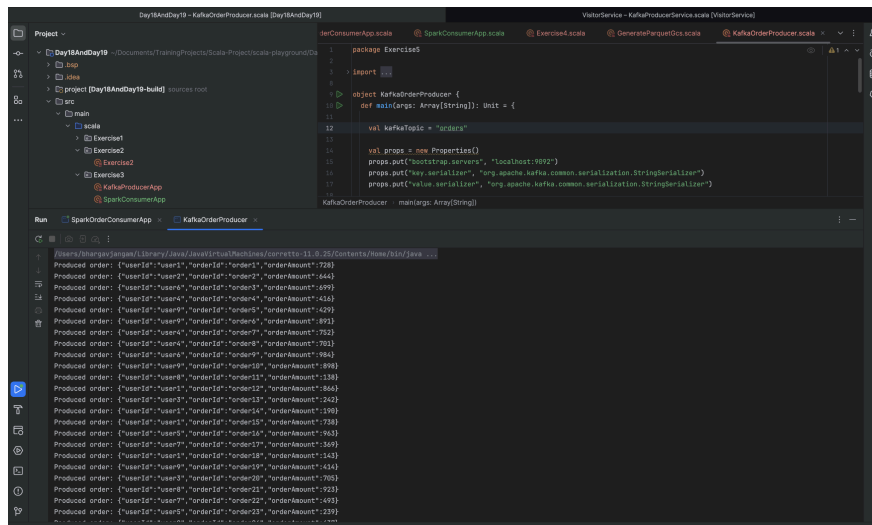
Filter by name prefix only ▾    Filter objects and folders    Show Live objects only ▾

| <input type="checkbox"/> | Name                           | Size    | Type                     | Created                  |            |
|--------------------------|--------------------------------|---------|--------------------------|--------------------------|------------|
| <input type="checkbox"/> | _SUCCESS                       | 0 B     | application/octet-stream | Nov 30, 2024, 2:07:30 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00000-d4c80971-c393-495e- | 1,010 B | application/octet-stream | Nov 30, 2024, 2:07:14 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00001-d4c80971-c393-495e- | 1,006 B | application/octet-stream | Nov 30, 2024, 2:07:09 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00002-d4c80971-c393-495e- | 1,007 B | application/octet-stream | Nov 30, 2024, 2:07:22 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00003-d4c80971-c393-495e- | 995 B   | application/octet-stream | Nov 30, 2024, 2:06:26 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00004-d4c80971-c393-495e- | 992 B   | application/octet-stream | Nov 30, 2024, 2:06:15 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00005-d4c80971-c393-495e- | 998 B   | application/octet-stream | Nov 30, 2024, 2:06:44 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00006-d4c80971-c393-495e- | 995 B   | application/octet-stream | Nov 30, 2024, 2:06:54 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00007-d4c80971-c393-495e- | 995 B   | application/octet-stream | Nov 30, 2024, 2:06:32 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00008-d4c80971-c393-495e- | 995 B   | application/octet-stream | Nov 30, 2024, 2:06:08 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00009-d4c80971-c393-495e- | 1,009 B | application/octet-stream | Nov 30, 2024, 2:06:38 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00010-d4c80971-c393-495e- | 1,005 B | application/octet-stream | Nov 30, 2024, 2:07:03 AM | 🔍 ⬇️ ⬆️ ⬇️ |
| <input type="checkbox"/> | part-00011-d4c80971-c393-495e- | 1,002 B | application/octet-stream | Nov 30, 2024, 2:07:18 AM | 🔍 ⬇️ ⬆️ ⬇️ |

Rows per page: 50 ▾    1 - 13 of 13    < >

# Exercise 5

## Kafka Producer Service for Order



```
package Exercise5

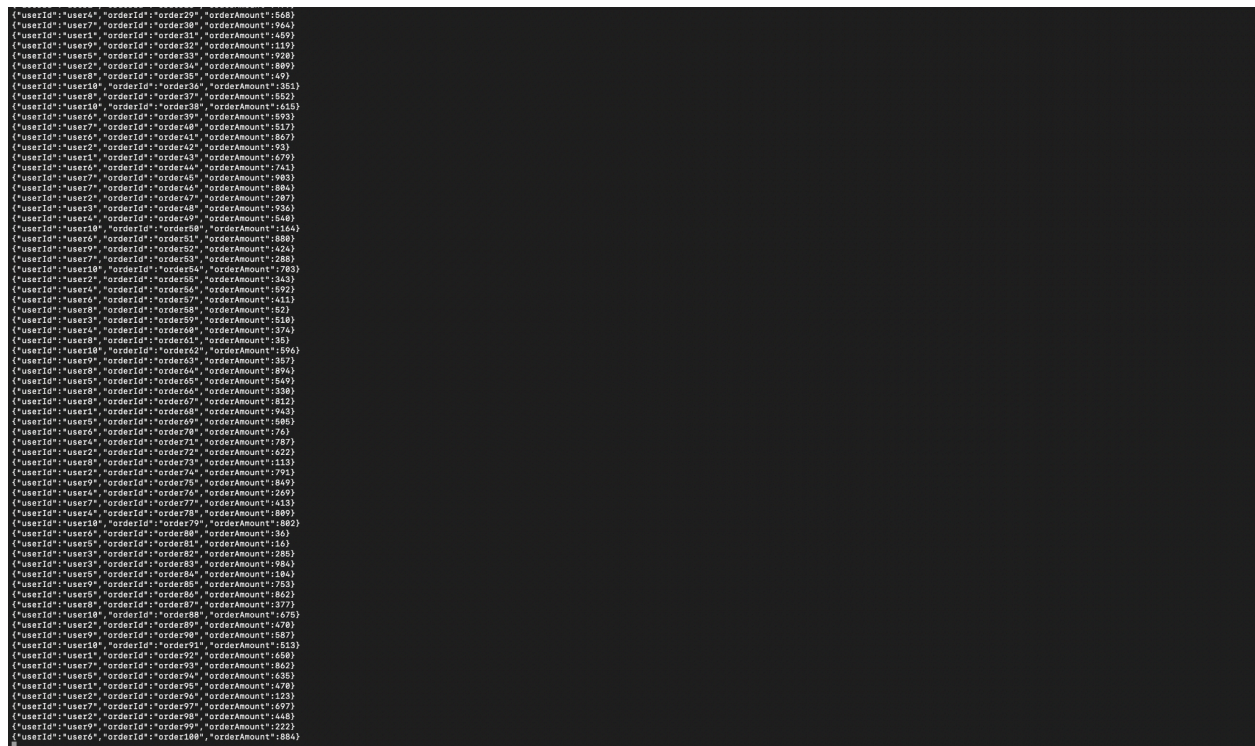
import org.apache.kafka.clients.producer.{KafkaProducer, ProducerRecord}
import org.apache.kafka.common.serialization.StringSerializer

object KafkaOrderProducer {
  def main(args: Array[String]): Unit = {
    val kafkaTopic = "orders"

    val props = new Properties()
    props.put("bootstrap.servers", "localhost:9092")
    props.put("key.serializer", "org.apache.kafka.common.serialization.StringSerializer")
    props.put("value.serializer", "org.apache.kafka.common.serialization.StringSerializer")

    KafkaOrderProducer mainArgs: Array[String])
  }
}
```

## Consumer Console



```
{
  "user": "user1",
  "order": "order1",
  "amount": 1000
}
{
  "user": "user2",
  "order": "order2",
  "amount": 1000
}
{
  "user": "user3",
  "order": "order3",
  "amount": 1000
}
{
  "user": "user4",
  "order": "order4",
  "amount": 1000
}
{
  "user": "user5",
  "order": "order5",
  "amount": 1000
}
{
  "user": "user6",
  "order": "order6",
  "amount": 1000
}
{
  "user": "user7",
  "order": "order7",
  "amount": 1000
}
{
  "user": "user8",
  "order": "order8",
  "amount": 1000
}
{
  "user": "user9",
  "order": "order9",
  "amount": 1000
}
{
  "user": "user10",
  "order": "order10",
  "amount": 1000
}
{
  "user": "user11",
  "order": "order11",
  "amount": 1000
}
{
  "user": "user12",
  "order": "order12",
  "amount": 1000
}
{
  "user": "user13",
  "order": "order13",
  "amount": 1000
}
{
  "user": "user14",
  "order": "order14",
  "amount": 1000
}
{
  "user": "user15",
  "order": "order15",
  "amount": 1000
}
{
  "user": "user16",
  "order": "order16",
  "amount": 1000
}
{
  "user": "user17",
  "order": "order17",
  "amount": 1000
}
{
  "user": "user18",
  "order": "order18",
  "amount": 1000
}
{
  "user": "user19",
  "order": "order19",
  "amount": 1000
}
{
  "user": "user20",
  "order": "order20",
  "amount": 1000
}
{
  "user": "user21",
  "order": "order21",
  "amount": 1000
}
{
  "user": "user22",
  "order": "order22",
  "amount": 1000
}
{
  "user": "user23",
  "order": "order23",
  "amount": 1000
}
{
  "user": "user24",
  "order": "order24",
  "amount": 1000
}
{
  "user": "user25",
  "order": "order25",
  "amount": 1000
}
{
  "user": "user26",
  "order": "order26",
  "amount": 1000
}
{
  "user": "user27",
  "order": "order27",
  "amount": 1000
}
{
  "user": "user28",
  "order": "order28",
  "amount": 1000
}
{
  "user": "user29",
  "order": "order29",
  "amount": 1000
}
{
  "user": "user30",
  "order": "order30",
  "amount": 1000
}
{
  "user": "user31",
  "order": "order31",
  "amount": 1000
}
{
  "user": "user32",
  "order": "order32",
  "amount": 1000
}
{
  "user": "user33",
  "order": "order33",
  "amount": 1000
}
{
  "user": "user34",
  "order": "order34",
  "amount": 1000
}
{
  "user": "user35",
  "order": "order35",
  "amount": 1000
}
{
  "user": "user36",
  "order": "order36",
  "amount": 1000
}
{
  "user": "user37",
  "order": "order37",
  "amount": 1000
}
{
  "user": "user38",
  "order": "order38",
  "amount": 1000
}
{
  "user": "user39",
  "order": "order39",
  "amount": 1000
}
{
  "user": "user40",
  "order": "order40",
  "amount": 1000
}
{
  "user": "user41",
  "order": "order41",
  "amount": 1000
}
{
  "user": "user42",
  "order": "order42",
  "amount": 1000
}
{
  "user": "user43",
  "order": "order43",
  "amount": 1000
}
{
  "user": "user44",
  "order": "order44",
  "amount": 1000
}
{
  "user": "user45",
  "order": "order45",
  "amount": 1000
}
{
  "user": "user46",
  "order": "order46",
  "amount": 1000
}
{
  "user": "user47",
  "order": "order47",
  "amount": 1000
}
{
  "user": "user48",
  "order": "order48",
  "amount": 1000
}
{
  "user": "user49",
  "order": "order49",
  "amount": 1000
}
{
  "user": "user50",
  "order": "order50",
  "amount": 1000
}
{
  "user": "user51",
  "order": "order51",
  "amount": 1000
}
{
  "user": "user52",
  "order": "order52",
  "amount": 1000
}
{
  "user": "user53",
  "order": "order53",
  "amount": 1000
}
{
  "user": "user54",
  "order": "order54",
  "amount": 1000
}
{
  "user": "user55",
  "order": "order55",
  "amount": 1000
}
{
  "user": "user56",
  "order": "order56",
  "amount": 1000
}
{
  "user": "user57",
  "order": "order57",
  "amount": 1000
}
{
  "user": "user58",
  "order": "order58",
  "amount": 1000
}
{
  "user": "user59",
  "order": "order59",
  "amount": 1000
}
{
  "user": "user60",
  "order": "order60",
  "amount": 1000
}
{
  "user": "user61",
  "order": "order61",
  "amount": 1000
}
{
  "user": "user62",
  "order": "order62",
  "amount": 1000
}
{
  "user": "user63",
  "order": "order63",
  "amount": 1000
}
{
  "user": "user64",
  "order": "order64",
  "amount": 1000
}
{
  "user": "user65",
  "order": "order65",
  "amount": 1000
}
{
  "user": "user66",
  "order": "order66",
  "amount": 1000
}
{
  "user": "user67",
  "order": "order67",
  "amount": 1000
}
{
  "user": "user68",
  "order": "order68",
  "amount": 1000
}
{
  "user": "user69",
  "order": "order69",
  "amount": 1000
}
{
  "user": "user70",
  "order": "order70",
  "amount": 1000
}
{
  "user": "user71",
  "order": "order71",
  "amount": 1000
}
{
  "user": "user72",
  "order": "order72",
  "amount": 1000
}
{
  "user": "user73",
  "order": "order73",
  "amount": 1000
}
{
  "user": "user74",
  "order": "order74",
  "amount": 1000
}
{
  "user": "user75",
  "order": "order75",
  "amount": 1000
}
{
  "user": "user76",
  "order": "order76",
  "amount": 1000
}
{
  "user": "user77",
  "order": "order77",
  "amount": 1000
}
{
  "user": "user78",
  "order": "order78",
  "amount": 1000
}
{
  "user": "user79",
  "order": "order79",
  "amount": 1000
}
{
  "user": "user80",
  "order": "order80",
  "amount": 1000
}
{
  "user": "user81",
  "order": "order81",
  "amount": 1000
}
{
  "user": "user82",
  "order": "order82",
  "amount": 1000
}
{
  "user": "user83",
  "order": "order83",
  "amount": 1000
}
{
  "user": "user84",
  "order": "order84",
  "amount": 1000
}
{
  "user": "user85",
  "order": "order85",
  "amount": 1000
}
{
  "user": "user86",
  "order": "order86",
  "amount": 1000
}
{
  "user": "user87",
  "order": "order87",
  "amount": 1000
}
{
  "user": "user88",
  "order": "order88",
  "amount": 1000
}
{
  "user": "user89",
  "order": "order89",
  "amount": 1000
}
{
  "user": "user90",
  "order": "order90",
  "amount": 1000
}
{
  "user": "user91",
  "order": "order91",
  "amount": 1000
}
{
  "user": "user92",
  "order": "order92",
  "amount": 1000
}
{
  "user": "user93",
  "order": "order93",
  "amount": 1000
}
{
  "user": "user94",
  "order": "order94",
  "amount": 1000
}
{
  "user": "user95",
  "order": "order95",
  "amount": 1000
}
{
  "user": "user96",
  "order": "order96",
  "amount": 1000
}
{
  "user": "user97",
  "order": "order97",
  "amount": 1000
}
{
  "user": "user98",
  "order": "order98",
  "amount": 1000
}
{
  "user": "user99",
  "order": "order99",
  "amount": 1000
}
{
  "user": "user100",
  "order": "order100",
  "amount": 1000
}
```

## Original DataFrame

```
Original DataFrame:
+-----+-----+-----+
|userId|      userName|      userEmail|
+-----+-----+-----+
| user1|    John Doe|john.doe@example.com|
| user2|   Jane Smith|jane.smith@exampl...|
| user3|  Bob Johnson|bob.johnson@examp...|
| user4|Alice Williams|alice.williams@ex...|
| user5| Charlie Brown|charlie.brown@exa...|
| user6|    David Lee|david.lee@example...|
| user7|   Emma Davis|emma.davis@exampl...|
| user8| Frank Harris|frank.harris@exam...|
| user9| Grace Wilson|grace.wilson@exam...|
|user10| Hannah Clark|hannah.clark@exam...|
+-----+-----+-----+
only showing top 10 rows
```

## Enriched DataFrame

```
-----
Batch: 29
-----
+-----+-----+-----+-----+-----+
| orderId|userId|orderAmount|      userName|      userEmail|
+-----+-----+-----+-----+-----+
|order675|user79|      62|    Zoe Perez|zoe.perez@example...|
|order846|user61|     267|Heather Robinson|heather.robinson@...|
|order566|user73|     641|   Todd Miller|todd.miller@examp...|
|order593|user64|     332|   Kevin Taylor|kevin.taylor@exam...|
|order450|user81|      36|Brendan Mitchell|brendan.mitchell@...|
| order62|user23|     736|   Vince Young|vince.young@examp...|
|order722|user82|     905|   Chloe Adams|chloe.adams@examp...|
|order752|user67|       2|   Nora Evans|nora.evans@exampl...|
|order164|user30|     515|Caitlyn Martinez|caitlyn.martinez@...|
|order944|user91|     190|   Liam Robinson|liam.robinson@exa...|
| order65|user60|     814|   George Allen|george.allen@exam...|
|order863|user95|     423|   Peter Wright|peter.wright@exam...|
|order677|user19|     873|Rebecca White|rebecca.white@exa...|
|order102|user53|     101|   Zane Scott|zane.scott@exampl...|
|order396|user57|     914|David Johnson|david.johnson@exa...|
|order487|user57|     535|David Johnson|david.johnson@exa...|
|order541|user73|     787|   Todd Miller|todd.miller@examp...|
|order213|user73|     787|   Todd Miller|todd.miller@examp...|
|order203|user78|     762|   Yara Moore|yara.moore@exampl...|
|order998| user8|     943|   Frank Harris|frank.harris@exam...|
+-----+-----+-----+-----+-----+
```



# Output-CheckPoints

Google Cloud

Bhargav1

buck

Search

Cloud Storage

Bucket details

GO TO PATH

REFRESH

LEARN

Overview

Buckets

Monitoring

Settings

bhargav-assignments

Location

us (multiple regions in United States)

Storage class

Standard

Public access

Not public

Protection

Soft Delete

OBJECTS

CONFIGURATION

PERMISSIONS

PROTECTION

LIFECYCLE

OBSERVABILITY

INVENTORY REPORTS

OPERATIONS

Folder browser

bhargav-assignments

Day16And17Task/

Day18And19Task/

case\_study\_4/

ClassWork/

Exercise\_5/

input/

output/

checkpoint/

commits/

offsets/

sources/

enriched\_orders/

Day18AndDay19Task/

CREATE FOLDER

UPLOAD

TRANSFER DATA

OTHER SERVICES

Filter by name prefix only

Filter

Filter objects and folders

Show

Live objects only

| Name     | Size | Type                     | Created                  |
|----------|------|--------------------------|--------------------------|
| commits/ | —    | Folder                   | —                        |
| metadata | 45 B | application/octet-stream | Nov 30, 2024, 5:20:46 AM |
| offsets/ | —    | Folder                   | —                        |
| sources/ | —    | Folder                   | —                        |

bhargav-assignments/Day18And19Task/ClassWork/ deleted

Get started with Cloud Storage

Getting bucket information

Uploading objects

Downloading objects

Use cases for Cloud Storage

Terraform samples

Architecture guides for storage

Prevent public access

# Output Json list

bhargav-assignments

Location

us (multiple regions in United States)

Storage class

Standard

Public access

Not public

Protection

Soft Delete

OBJECTS

CONFIGURATION

PERMISSIONS

PROTECTION

LIFECYCLE

OBSERVABILITY

INVENTORY REPORTS

OPERATIONS

Folder browser

bhargav-assignments

Day16And17Task/

Day18And19Task/

case\_study\_4/

Exercise\_5/

input/

output/

checkpoint/

MORE RESULTS

enriched\_orders/

Day18AndDay19Task/

CREATE FOLDER

UPLOAD

TRANSFER DATA

OTHER SERVICES

Filter by name prefix only

Filter

Filter objects and folders

Show

Live and noncurrent objects

| Name                          | Size  | Type             | Created                  | Storage  |
|-------------------------------|-------|------------------|--------------------------|----------|
| part-00000-1732928146101.json | 136 B | application/json | Nov 30, 2024, 6:32:48 AM | Standard |
| part-00001-1732928146102.json | 140 B | application/json | Nov 30, 2024, 6:32:49 AM | Standard |
| part-00002-1732928146103.json | 142 B | application/json | Nov 30, 2024, 6:32:48 AM | Standard |