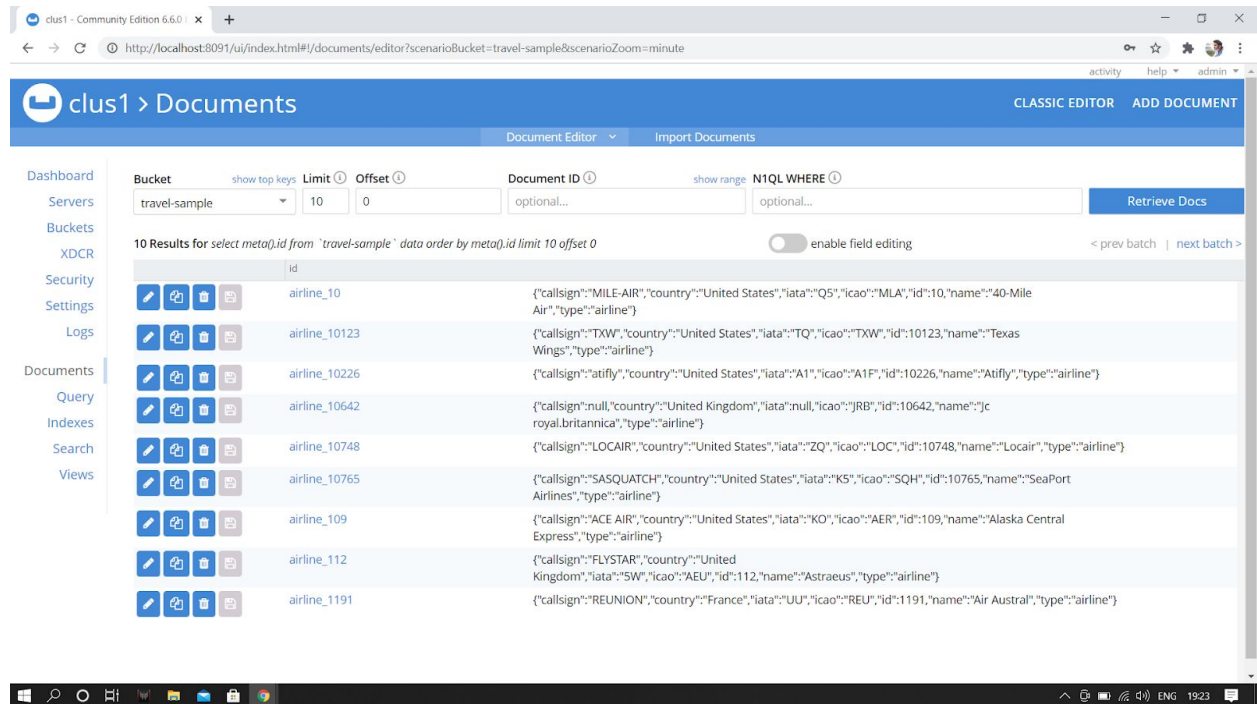


NoSQL Assignment

Maithreyan Kesavan

1. Import `travel-sample` bucket



2. Write a query to get the sum of all distances where type="route" for each airline id.

```
SELECT airlineid,SUM(distance)
FROM `travel-sample`
WHERE type="route"
GROUP BY airlineid
```

The screenshot shows the clus1 Query Workbench interface. The top navigation bar includes 'activity', 'help', and 'admin' links. The main header displays 'clus1 > Query' and 'Query Workbench'. The left sidebar contains navigation links for Dashboard, Servers, Buckets, XDCR, Security, Settings, Logs, Documents, Query, Indexes, Search, and Views.

The central 'Query Editor' displays the following SQL query:

```
1 SELECT airlineid,SUM(distance)
2 FROM `travel-sample`
3 WHERE type="route"
4 GROUP BY airlineid
```

Below the query editor, the 'Query Results' section shows the results in JSON format. The results are grouped by airlineid, showing the sum of distances for each route. The JSON output is as follows:

```
1 {
2   {
3     "$1": 436068.8215248117,
4     "airlineid": "airline_330"
5   },
6   {
7     "$1": 85923.3884873361,
8     "airlineid": "airline_2835"
9   },
10  {
11    "$1": 2802.2647632268013,
12    "airlineid": "airline_4822"
13  },
14  {
15    "$1": 578723.978014008,
16    "airlineid": "airline_35"
17  },
18  {
```

The right sidebar, 'Data Insights', provides a summary of the data. It shows the 'travel-sample' bucket with 1000 samples out of 31594. The 'stops' field is 0 for 'route' type, accounting for 79.7% of the data. The 'destinationairport' field is indexed and shows the following distribution:

- 'type' = "landmark": 12.6%
- 'type' = "hotel": 2.0%
- 'type' = "airline": 0.3%
- 'type' = "airport": 5.4%

The 'Indexes' section is also visible at the bottom of the Data Insights panel.

3. Write queries to join(LEFT,RIGHT,INNER) type="route" & "airline" and fetch the data whose sourceairport="SFO"

Inner Join

```
SELECT *
FROM `travel-sample` route
INNER JOIN `travel-sample` airline
ON route.airlineid = META(airline).id
WHERE route.type = "route"
AND route.destinationairport = "SFO";
```

The screenshot displays the clus1 Query Workbench interface. The top navigation bar includes the clus1 logo, the text 'Query', and buttons for 'IMPORT' and 'EXPORT'. Below this, the 'Query Editor' section contains a SQL query:

```
1 SELECT *
2 FROM `travel-sample` route
3 LEFT JOIN `travel-sample` airline
4 ON route.airlineid = META(airline).id
5 WHERE route.type = "route"
6 AND route.destinationairport = "SFO";
```

Buttons for 'Execute' and 'Explain' are visible, along with a status bar indicating 'success' and execution details. The 'Query Results' section shows the query output in JSON format:

```
1 {
2   {
3     "airline": {
4       "callsign": "JETBLUE",
5       "country": "United States",
6       "iata": "B6",
7       "icao": "JBU",
8       "id": "3029",
9       "name": "JetBlue Airways",
10      "type": "airline"
11    },
12    "route": {
13      "airline": "B6",
14      "airlineid": "airline_3029",
15      "destinationairport": "SFO",
16      "distance": 2416.0035377223094,
```

The 'Data Insights' section on the right provides a summary of the queryable buckets for the 'travel-sample' database, sampled from 1000 of 31594 records:

Bucket	Percentage
'type' = "route"	81.1%
'type' = "landmark"	11.0%
'type' = "hotel"	2.6%
'type' = "airline"	0.9%
'type' = "airport"	4.4%

A 'Refresh' button is located at the bottom right of the Data Insights section.

Left Join

```
SELECT *
FROM `travel-sample` route
LEFT JOIN `travel-sample` airline
ON route.airlineid = META(airline).id
WHERE route.type = "route"
AND route.destinationairport = "SFO";
```

clus1 - Community Edition 6.6.0 | x

http://localhost:8091/ui/index.html#/query/workbench?scenarioBucket=travel-sample&scenarioZoom=minute

clus1 > Query

Query Workbench

Dashboard
Servers
Buckets
XDCR
Security
Settings
Logs

Documents
Query
Indexes
Search
Views

Query Editor

```
1 SELECT *
2 FROM `travel-sample` route
3 LEFT JOIN `travel-sample` airline
4 ON route.airlineid = META(airline).id
5 WHERE route.type = "route"
6 AND route.destinationairport = "SFO";
```

Execute Explain External Query Advisor success just now | elapsed: 1.8s | execution: 1.8s | docs: 250 | size: 907733 bytes

Query Results

Table JSON Tree Plan Plan Text

```
1+ | [
2+ | {
3+ |   "route": {
4+ |     "airline": "AI",
5+ |     "airlineid": "airline_218",
6+ |     "destinationairport": "SFO",
7+ |     "distance": 11128.1820358009515,
8+ |     "equipment": "77W",
9+ |     "id": 10495,
10+ |    "schedule": [
11+ |      {
12+ |        "day": 0,
13+ |        "flight": "AI339",
14+ |        "utc": "23:05:00"
15+ |      },
16+ |      {
```

Data Insights

Queryable Buckets

travel-sample sampled 1000 of 31594

- ▶ 'type' = "route" 81.1%
- ▶ 'type' = "landmark" 11.0%
- ▶ 'type' = "hotel" 2.6%
- ▶ 'type' = "airline" 0.9%
- ▶ 'type' = "airport" 4.4%

Indexes

Refresh

Right Join

```
SELECT *
FROM `travel-sample` route
Right JOIN `travel-sample` airline
ON route.airlineid = META(airline).id
WHERE route.type = "route"
AND route.destinationairport = "SFO";
```

The screenshot shows the clus1 Query Workbench interface. The top navigation bar includes 'activity', 'help', and 'admin' links. The main interface is divided into three panels:

- Query Editor:** Contains a SQL query:


```
1 SELECT *
2 FROM `travel-sample` route
3 Right JOIN `travel-sample` airline
4 ON route.airlineid = META(airline).id
5 WHERE route.type = "route"
6 AND route.destinationairport = "SFO";
```

 Below the query are buttons for 'Execute', 'Explain', and 'format'.
- Query Results:** Displays the query results in JSON format. The first result is:


```
{
  "airline": {
    "callsign": "CITRUS",
    "country": "United States",
    "iata": "FL",
    "icao": "TRS",
    "id": 1316,
    "name": "AirTran Airways",
    "type": "airline"
  },
  "route": {
    "airline": "FL",
    "airlineid": "airline_1316",
    "destinationairport": "SFO",
    "distance": 3434.7108309317646,
  }
}
```
- Data Insights:**
 - travel-sample:** A bar chart showing the distribution of document types: 'route' (79.2%), 'landmark' (12.5%), 'hotel' (2.2%), 'airline' (0.4%), and 'airport' (5.7%).
 - MyNewBucket:** A bar chart showing the distribution of document types: 'route' (82.5%), 'landmark' (10.8%), 'hotel' (1.8%), 'airline' (0.4%), and 'airport' (4.5%).

4. Write a mapreduce to get the number of all documents based on entities(type).

Route

The screenshot shows the clus1 Query Workbench interface with a MapReduce job configuration. The top navigation bar includes 'activity', 'help', and 'admin' links. The main interface is divided into three panels:

- Documents:** Displays a document snippet:


```
10 "name": "Sentinel Dome",
11 "address": "Starts at the Sentinel Dome/Taft Point trailhead parking area, (6 mi / 9.6 km east of
12 Bridalveil creek Campground)",
13 "alt": "2.2 mi (3.5 km) round-trip",
14 "city": "Mariposa County",
```
- View Index Code:**
 - Map:**

```
1 function (doc, meta) {
2   if(doc.type == "route")
3     emit(meta.id, null);
4 }
```
 - Reduce (built in: _count, _sum, _stats):**

```
1 _count
```
- Results:**
 - Filter: ?limit=6&state=false&connection_timeout=60000&inclusive_end=true&skip=0&full_set=true
 - Development Time Subset: Full Cluster Data Set
 - Show Results button
 - Results table:

Key	Value
null	22950
undefined	

Landmarks

clus1 - Community Edition 6.6.0

http://localhost:8091/ui/index.html#/views/_design~2fdev_myview?scenarioBucket=MyNewBucket&scenarioZoom=minute&bucket=MyNewBucket&type=development&viewId=myview&isSpatial=0...

Documents

Query

Indexes

Search

Views

View Index Code

Map

```
1 function (doc, meta) {  
2   if(doc.type == "landmark")  
3     emit(meta.id, null);  
4 }
```

Reduce (built in: _count, _sum, _stats)

```
1 _count
```

Results filter: ?limit=6&state=false&connection_timeout=60000&inclusive_end=true&skip=0&full_set=true

Development Time Subset Full Cluster Data Set

Key	Value
null undefined	3422

Hotel

clus1 - Community Edition 6.6.0

http://localhost:8091/ui/index.html#/views/_design~2fdev_myview?scenarioBucket=MyNewBucket&scenarioZoom=minute&bucket=MyNewBucket&type=development&viewId=myview&isSpatial=0...

Documents

Query

Indexes

Search

Views

View Index Code

Map

```
1 function (doc, meta) {  
2   if(doc.type == "hotel")  
3     emit(meta.id, null);  
4 }
```

Reduce (built in: _count, _sum, _stats)

```
1 _count
```

Results filter: ?limit=6&state=false&connection_timeout=60000&inclusive_end=true&skip=0&full_set=true

Development Time Subset Full Cluster Data Set

Key	Value
null undefined	672

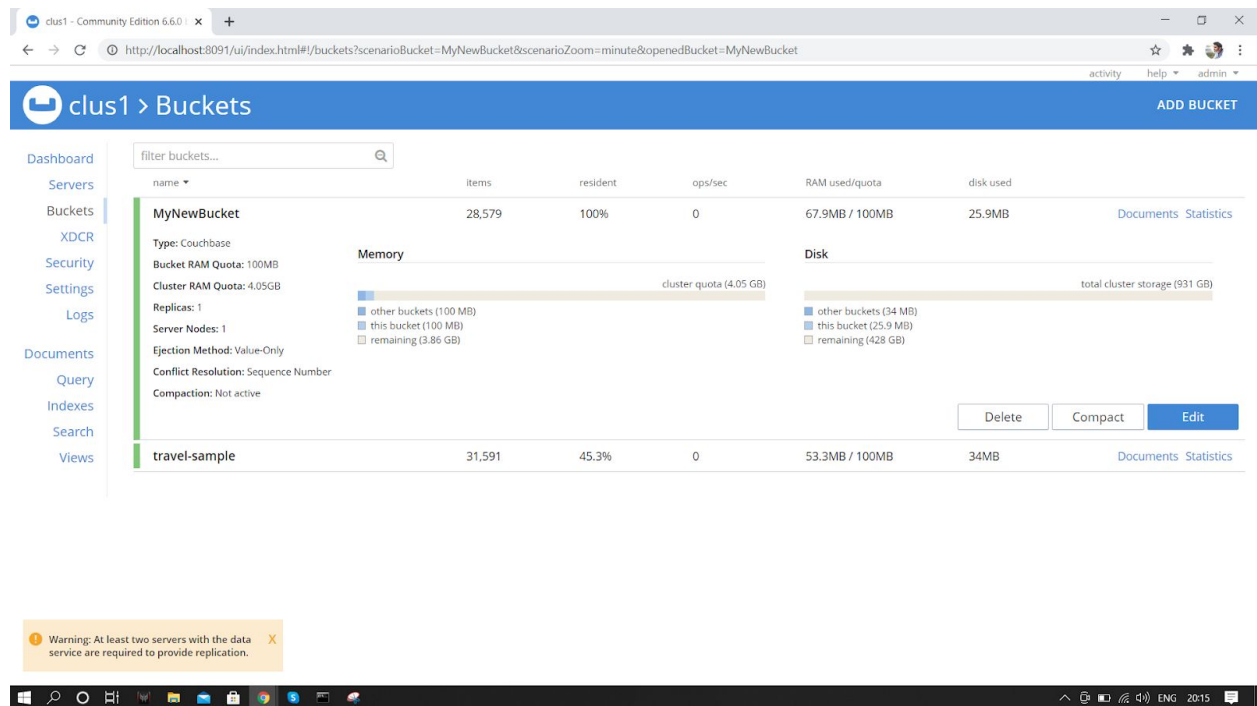
5. Refer CLI interface & try to export the travel sample data.

```
cbexport json -c couchbase://127.0.0.1 -u admin -p maithpass  
-b travel-sample -o D:/travelSample.json -f lines -t 4
```

```
cbimport json -c couchbase://127.0.0.1 -u admin -p maithpass -b MyNewBucket -f lines -d  
file://D:/travelSample.json -t 4 -g %id%
```

Exported and Imported the bucket

```
D:\softwares\Couchbase\Server\bin>cbexport json -c couchbase://127.0.0.1 -u admin -p maithpass -b travel-sample -o D:/travelSample.json -f lines -t 4  
JSON exported to 'D:/travelSample.json' successfully  
Documents exported: 31591 Documents skipped: 0  
  
D:\softwares\Couchbase\Server\bin>cbimport json -c couchbase://127.0.0.1 -u admin -p maithpass -b MyNewBucket -f lines -d file://D:/travelSample.json -t 4 -g %id%  
JSON 'file://D:/travelSample.json' imported to 'http://127.0.0.1:8091' successfully  
Documents imported: 31591 Documents failed: 0  
  
D:\softwares\Couchbase\Server\bin>
```



clus1 - Community Edition 6.6.0

http://localhost:8091/ui/index.html#/buckets?scenarioBucket=MyNewBucket&scenarioZoom=minute&openedBucket=MyNewBucket

clus1 > Buckets

filter buckets...

name	Items	resident	ops/sec	RAM used/quota	disk used	
MyNewBucket	28,579	100%	0	67.9MB / 100MB	25.9MB	Documents Statistics
Memory						
Type: Couchbase						
Bucket RAM Quota: 100MB						
Cluster RAM Quota: 4.05GB						
Replicas: 1						
Server Nodes: 1						
Ejection Method: Value-Only						
Conflict Resolution: Sequence Number						
Compaction: Not active						
Disk						
total cluster storage (931 GB)						
other buckets (34 MB)						
this bucket (25.9 MB)						
remaining (428 GB)						
Delete Compact Edit						
travel-sample	31,591	45.3%	0	53.3MB / 100MB	34MB	Documents Statistics

Warning: At least two servers with the data service are required to provide replication.