Wednesday, March 24th, 2021
Sneha George: 300006801
Marissa Phul: 300022039
Toluwani Ogunsanya: 8677256

Deliverable D
Project Phase 3: OLAP Queries and BI Dashboard

Task Division
Part 1
Drill down and roll up - Sneha
Slice - Marissa
Dice - Tolu
Combination - all

Part 2
Iceberg - Marissa
Windowing queries -Sneha
Window clause - Tolu

Part 3
Each add our queries to tableau

Part 1. Standard OLAP operations – 9 queries in total

a. Drill down and roll up. – 2 queries

i For instance, explore the total number of positive cases in your data mart; drill down to a month (April 2020), and drill down to a specific day.

ii For instance, explore the total number of resolved cases in your data mart; drill down to a week (first week of April 2020), and drill down to a specific day.

iii For instance, consider all the unresolved cases in Toronto City, roll up to GTA, and roll up to all data in your data mart.

a) Drill down and roll up

i	SELECT COUNT(*) , D.month, D.day FROM fact_table_v2 as F, date as D WHERE F."Reported_Date_Key" = D.id AND D.month = 4 GROUP BY (D.month, D.day) ORDER BY D.month, D.day
ii	SELECT COUNT(*) , D.month, D.day FROM fact_table_v2 as F, date as D WHERE F."Reported_Date_Key" = D.id AND D.month = 4 AND F."Resolved" = 'yes' AND D.day IN (1,2,3,4,5,6,7) GROUP BY (D.month, D.day) ORDER BY D.month, D.day
iii	SELECT COUNT(*), P.City FROM fact_table_v2 as F, phu_location as P WHERE F."Resolved"= 'no' AND F."PHU_Key" = P.phu_location_id GROUP BY ROLLUP(P.City) ORDER BY P.City

b. **Slice,** where only one dimension is selected. **– 2 queries**
For instance, provide the (i) the number of cases in a specific PHU (resolved, unresolved and fatal), (ii) the number cases across the PHUs when a specific special measure was in place, (iii) the mobility levels in Ottawa, etc.

b) Slice

Slice

1. Number of Cases in a Specific PHU

Toronto:

```
1  SELECT COUNT(*)
2  FROM fact_table_v2 as F, phu_location as P
3  WHERE F."PHU_Key" = P.phu_location_id
4  AND P.City = 'Toronto'
```

Data Output | Explain | Messages | Notifications

	count bigint
1	13011

Ottawa:

```
1  SELECT COUNT(*)
2  FROM fact_table_v2 as F, phu_location as P
3  WHERE F."PHU_Key" = P.phu_location_id
4  AND P.City = 'Ottawa'
```

Data Output | Explain | Messages | Notifications

	count bigint
1	2245

All Cities:

```
1  SELECT COUNT(*), P.City
2  FROM fact_table_v2 as F, phu_location as P
3  WHERE F."PHU_Key" = P.phu_location_id
4  GROUP BY P.City
```

Data Output | Explain | Messages | Notifications

	count bigint	city character varying (80)
1	13011	Toronto
2	2245	Ottawa

2. Number of Cases in PHUs when a specific Special Measure was in place
Stage 3 Special Measure:

```
1  SELECT COUNT(*), S.status
2  FROM fact_table_v2 as F, special_measures_v2 as S
3  WHERE F."Special_Measure_Key" = S.special_measures_id
4  AND S.special_measures_id = 9
5  GROUP BY S.status
```

Data Output | Explain | Messages | Notifications

	count bigint	status text
1	19	stage 3

All Special Measures:

```
1  SELECT COUNT(*), S.status
2  FROM fact_table_v2 as F, special_measures_v2 as S
3  WHERE F."Special_Measure_Key" = S.special_measures_id
4  GROUP BY S.status
```

Data Output | Explain | Messages | Notifications

	count bigint	status text
1	1752	stage 2 extended to toronto
2	581	Parks Re-opened
3	16	stage 3 extended to toronto
4	141	Stage 2 of reopening.
5	19	stage 3
6	8961	First State of Emergency
7	3786	Restart

c. **Dice,** where one creates a sub-cube. **– 2 queries**
For instance, (i) provide the number of fatal cases during a period of two months, e.g., February and March, in Peel and Ottawa, (ii) provide the number of unresolved cases when contrasting two mobility locations, e.g., parks and transit, in Peel and Ottawa.

c) Tolu

DICE QUERY	
provide the number of fatal cases during a period of two Months in two PHU locations	SELECT P.city, COUNT(F."Fatal") FROM fact_table_v2 AS F, date AS D, phu_location as P WHERE F."Fatal"='no' AND D.month>=4 AND D.month <= 5 AND F."PHU_Key"=P."phu_location_id" AND P."phu_location_id" IN (2251, 3895) GROUP BY P.city
provide the number of unresolved cases between Stage 1&2 of restart	SELECT S.status, COUNT(F."Resolved") FROM fact_table_v2 AS F, special_measures_v2 AS S WHERE F."Resolved"='no' AND S."special_measures_id"=F."Special_Measure_Key" AND S."special_measures_id" IN (7, 8, 11) GROUP BY S.status

d. **Combining OLAP operations.** In these queries, we combine the above-mentioned operations. **– 3 queries**

For instance, we may aim to explore the number of cases i) during different periods of the year, ii) when certain types of measures are in place, iii) for different types of outcomes and weather conditions iv) contrasting mobility levels in Ottawa and Peel, v) comparing sunny versus rainy days, etc.

D.

i) num of deaths by city by month	SELECT D.month, P.city, COUNT(*), RANK() OVER(ORDER BY COUNT(*) DESC) FROM fact_table_v2 as F, date as D, phu_location AS P WHERE F."Reported_Date_Key"=D.id AND F."Fatal"='yes' AND P."phu_location_id"=F."PHU_Key" GROUP BY P.city, D.month LIMIT 10
ii) num of cases when weather	SELECT COUNT(*) as num_of_cases, CASE WHEN precipitation >= 20 THEN 'Very Rainy' WHEN precipitation < 20 AND precipitation > 0 THEN 'Rainy' WHEN precipitation <= 0 THEN 'Sunny' END precipitation_day FROM fact_table_v2 as F, weather as W WHERE F."Weather_Key" = W.weather_id GROUP BY precipitation_day
iii) num of cases when special measures	SELECT S.status, P.city, COUNT(*) FROM fact_table_v2 as F, special_measures_v2 as S, phu_location as P WHERE F."Special_Measure_Key" IN (2,9,10) AND S."special_measures_id"=F."Special_Measure_Key" AND F."PHU_Key"=P."phu_location_id" GROUP BY S.status, P.city

Part 2. Explorative operation – 3 queries

Identify general trends using advanced SQL operations. Give *one query* from each one of these categories.

a. **Iceberg queries.** For instance, find the five days with the highest numbers of resolved outcomes, find the location with the highest mobility in terms of visits to parks, to grocery stores and pharmacies, etc.
b. **Windowing queries.** For instance, show the ranking of the PHUs in terms of the number of cases per week, per outcome, per month, etc.
c. **Using the Window clause.** For instance, compare the number of resolved cases in Ottawa to that of the previous and next months, etc.

a. Iceberg Query

5 Days with the Highest Number of Resolved Outcomes

```
1  SELECT COUNT(F."Reported_Date_Key"),
2         TO_DATE(CONCAT(D.year::text,'0',D.month::text,D.day::text), 'YYYYMMDD') as myDate
3  FROM fact_table_v2 as F, date as D
4  WHERE F."Reported_Date_Key" = D."id" AND F."Resolved" = 'yes'
5  GROUP BY F."Reported_Date_Key", D.day, D.month, D.year
6  ORDER BY count DESC
7  LIMIT 5
```

Data Output | Explain | Messages | Notifications

	count bigint	mydate date
1	403	2020-05-29
2	332	2020-04-17
3	271	2020-04-20
4	247	2020-04-16
5	219	2020-04-15

B. Windowing queries

```
SELECT COUNT(*),P.City,  D.month, F."Resolved",
    RANK () OVER(ORDER BY COUNT(*) ASC) Rank
FROM fact_table_v2 as F, phu_location as P, date as D
GROUP BY ( P.City, D.month ,F."Resolved")
ORDER BY Rank;;
```

Query Editor | Query History

```
1  SELECT COUNT(*),P.City, D.month, F."Resolved",
2         RANK () OVER(ORDER BY COUNT(*) ASC) Rank
3  FROM fact_table_v2 as F, phu_location as P, date as D
4  GROUP BY ( P.City, D.month ,F."Resolved")
5  ORDER BY Rank;
```

Data Output | Explain | Messages | Notifications

	count bigint	city character varying (80)	month numeric	Resolved character varying	rank bigint
1	44220	Toronto	4	no	1
2	44220	Ottawa	6	no	1
3	44220	Ottawa	4	no	1
4	44220	Toronto	6	no	1
5	45694	Toronto	5	no	5
6	45694	Ottawa	5	no	5
7	45694	Ottawa	7	no	5
8	45694	Toronto	7	no	5
9	413460	Toronto	6	yes	9
10	413460	Toronto	4	yes	9
11	413460	Ottawa	4	yes	9
12	413460	Ottawa	6	yes	9
13	427242	Ottawa	7	yes	13
14	427242	Ottawa	5	yes	13
15	427242	Toronto	7	yes	13
16	427242	Toronto	5	yes	13

c) Window Clause

Query Editor | Query History

```
1  SELECT D.month, COUNT(F."Resolved") as resolvedCount, RANK()  OVER W
2  FROM fact_table_v2 as F, date as D
3  WHERE F."Resolved"='yes'
4  GROUP BY D.month
5  WINDOW W AS (ORDER BY COUNT(F."Resolved"))
6
```

Data Output | Explain | Messages | Notifications

	month numeric	resolvedcount bigint	rank bigint
1	6	413460	1
2	4	413460	1
3	5	427242	3
4	7	427242	3

Part 3. BI dashboard and Information Visualisation (40 marks)

Create a dashboard that allows the users to explore the data and to visualise trends. Your interface should include graphs and charts. You are encouraged to use Tableau, or any other dashboard tool of your choice.

Please see Tableau workbook attached.