CSI 4142 Fundamentals of Data Science

Project Phase 3: OLAP Queries and BI Dashboard

University of Ottawa

School of Computer Science

Professor : Herna L Viktor

Student Name : Jiajie Xu
Student Number: 7881937

Student Name : Ali Khanafer
Student Number: 300010614

Student Name : Vidulash Rajaratnam
Student Number: 8190398

Queries

Drill Down

```
select D.month, D.day, count(*) as total_cases
from covid_19_tracking_fact_table as F, date_dimension as D
where F.reported_date_key = D.date_key
group by (D.month,D.day)
order by D.month, D.day
```

Roll Up

```
select P.province, P.city, sum(total_fatal) as total_fatal
from (select distinct * from covid_19_tracking_fact_table) F,
phu_location_dimension P
where F.patient_key in (select distinct patient_key from
covid_19_tracking_fact_table)
and F.phu_location_key = P.phu_location_key
group by rollup(P.province, P.city)
order by P.province, P.city
```

First Slice

```
select P.city, D.month, sum(F.total_fatal) as total_deaths,
sum(F.total_resolved) as total_resolved, sum(F.total_unresolved) as
total_unresolved
from (select distinct * from covid_19_tracking_fact_table) F,
phu_location_dimension P, date_dimension D
where F.patient_key in (select distinct patient_key from
covid_19_tracking_fact_table)
and F.phu_location_key = P.phu_location_key and F.reported_date_key =
D.date_key and P.city='Ottawa' and D.month != 3
group by (D.month, P.city)
order by D.month
```

Second Slice

```
select P.city, D.date, S.description, count(*) as total_cases
from covid_19_tracking_fact_table F, phu_location_dimension P,
date_dimension D, special_measures_dimension S
where F.phu_location_key = P.phu_location_key and F.reported_date_key =
D.date_key and F.special_measures_key = S.special_measures_key and D.month
!= 3 and S.special_measures_key = 1
group by (D.date, P.city, S.description)
order by (P.city,D.date)
```

First Dice

```
select D.month, P.city, sum(total_fatal) as total_deaths
from (select distinct * from covid_19_tracking_fact_table) F,date_dimension
D, phu_location_dimension P
where F.patient_key in (select distinct patient_key from
covid_19_tracking_fact_table) and
F.reported_date_key = D.date_key and F.phu_location_key =
P.phu_location_key and D.month in (10,11,12) and P.city in ('Ottawa',
'Toronto')
group by (D.month, P.city)
order by D.month, P.city
```

Second Dice

```
select D.date, P.city, M.retail_and_recreation, M.parks,count(*) as
total_cases
from covid_19_tracking_fact_table F, date_dimension D,
phu_location_dimension P, mobility_dimension M
where F.test_date_key = D.date_key and F.phu_location_key =
P.phu_location_key and F.mobility_key = M.mobility_key
and D.month in (10,11,12) and P.city in ('Ottawa', 'Toronto')
group by (D.date, P.city, M.retail_and_recreation, M.parks)
order by D.date, P.city
```

First Combination

```
select count(*) as Total_Cases, D.season, D.date
from covid_19_tracking_fact_table as F, date_dimension as D
where F.reported_date_key = D.date_key
group by (D.season, D.date)
```

Second Combination

```
select count(*) as Total_Cases, S.keyword1, D.date
from covid_19_tracking_fact_table as F, special_measures_dimension as S,
date_dimension D
where F.special_measures_key = S.special_measures_key and F.test_date_key =
D.date_key and S.keyword1 in
('Red-Control','Grey-Lockdown','Orange-Restrict')
group by (D.date, S.keyword1)
order by Total_Cases asc
```

Third Combination

```
select D.date, P.city, M.sub_region,  M.grocery_and_pharmarcy, M.parks,
M.residential, M.retail_and_recreation,M.transit_stations,M.workplaces
from covid_19_tracking_fact_table F, date_dimension D,
phu_location_dimension P, mobility_dimension M
where F.test_date_key = D.date_key and F.phu_location_key =
P.phu_location_key and F.mobility_key = M.mobility_key
and P.city in ('Ottawa', 'Newmarket') and D.month in (10, 11)
group by (D.date, P.city, M.sub_region, M.grocery_and_pharmarcy, M.parks,
M.residential, M.retail_and_recreation,M.transit_stations,M.workplaces)
order by D.date, P.city
```

Iceberg

```
select D.date, P.city, count(*) total_cases_reported
from covid_19_tracking_fact_table F, date_dimension D,
phu_location_dimension P
where F.reported_date_key = D.date_key and F.phu_location_key =
P.phu_location_key
group by D.date, P.city
order by count(*) desc limit 10
```

Window Clause

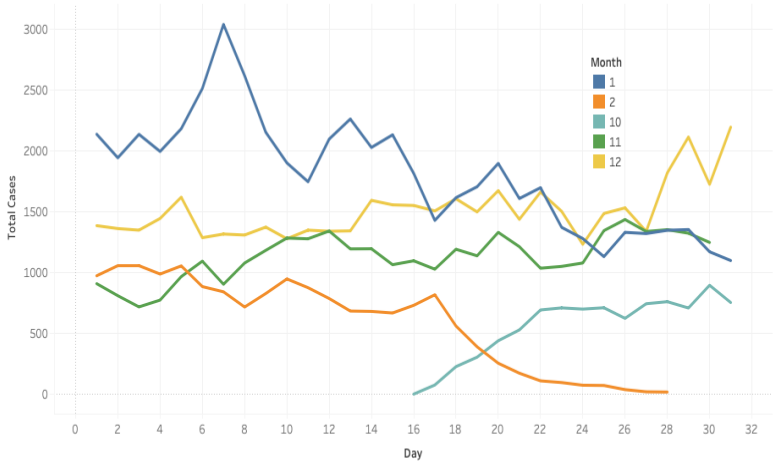
```
SELECT distinct P.age_group,L.city, sum(F.total_fatal) OVER W AS
Total_resolved
FROM (select distinct * from covid_19_tracking_fact_table) F,
phu_location_dimension L, patient_dimension P
WHERE F.phu_location_key = L.phu_location_key and F.patient_key =
P.patient_key and L.city = 'Toronto'
WINDOW W AS (PARTITION BY P.age_group
             ORDER BY L.city)
```

Windowing

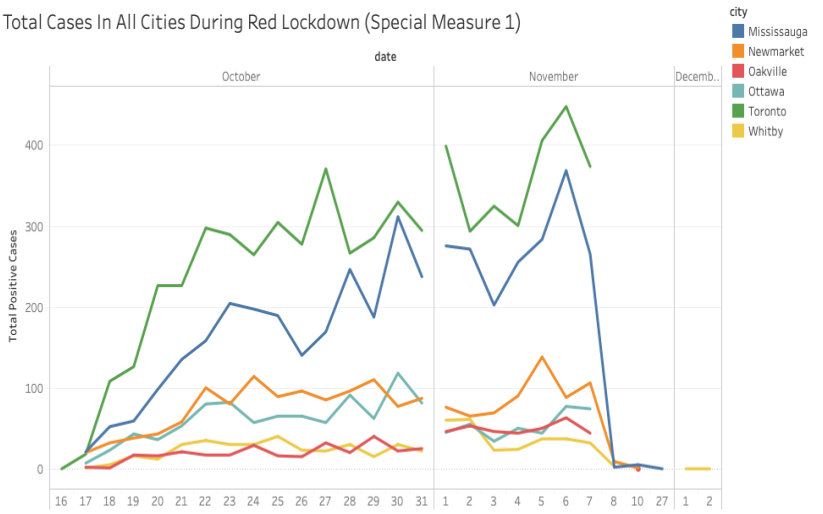
```
select tmp.month, tmp.phu_name, tmp.Total_cases,
RANK() over (partition by tmp.month order by tmp.Total_cases desc)
from (select b.month, c.phu_name , count(a.*) as Total_cases from
covid_19_tracking_fact_table a inner join date_dimension b on
a.reported_date_key = b.date_key
inner join phu_location_dimension c on a.phu_location_key =
c.phu_location_key
group by (b.month, c.phu_name)) tmp
```

Dashboard

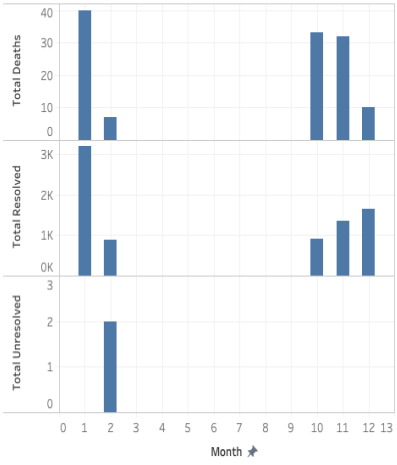
Daily total COVID-19 Cases Per Month



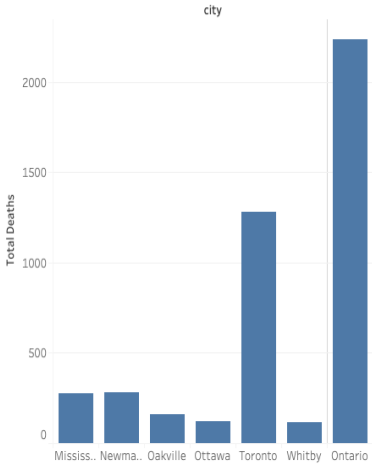
Total Cases In All Cities During Red Lockdown (Special Measure 1)



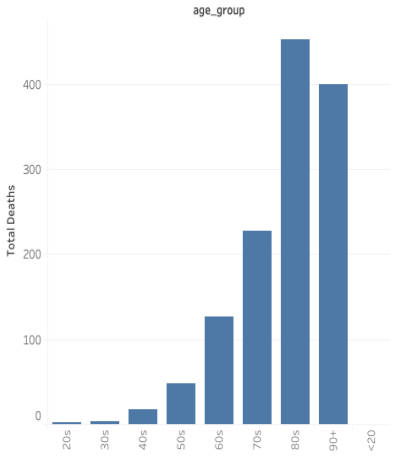
Outcomes Per Month in Ottawa



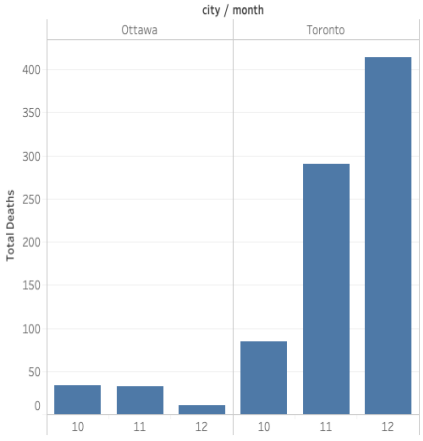
Total Deaths In City Compared to Province



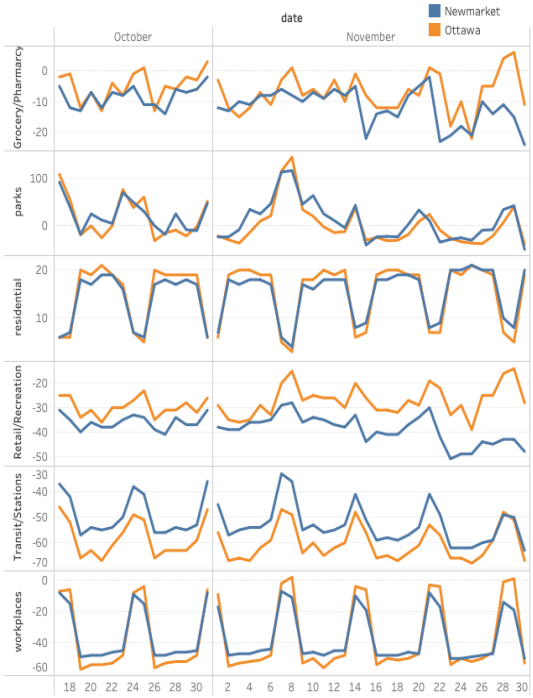
Number of Deaths By Age Group in Toronto



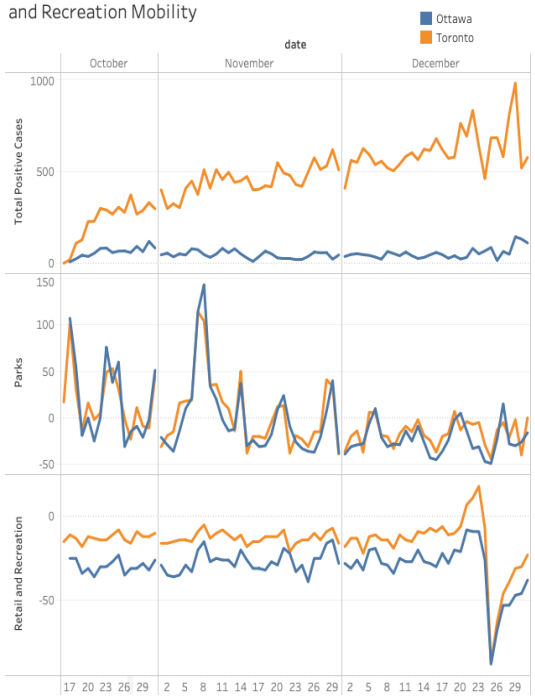
Total Deaths in Ocotber, November And December in Ottawa and Toronto



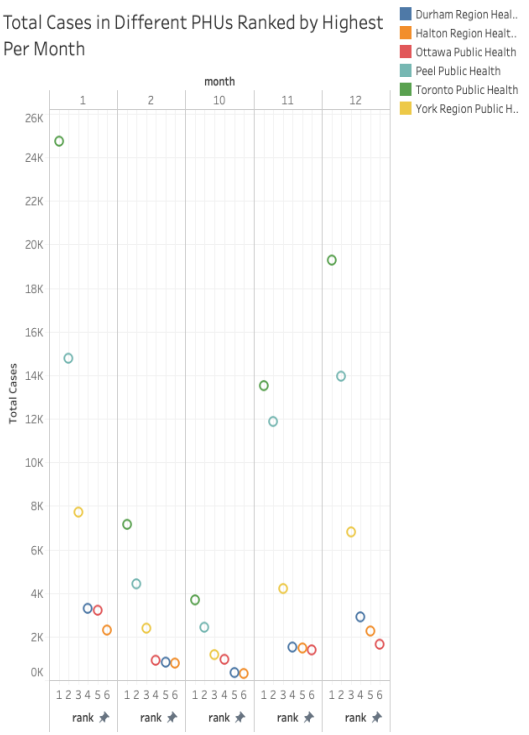
Comparing mobility data of Newmarket and Ottawa



Total Cases in Ottawa and Toronto Compared with Park, Retail and Recreation Mobility



Total Cases in Different PHUs Ranked by Highest Per Month



Top 10 Days With Largest Number of Cases

date									
January									December
2	5	6	7	8	9	12	15	19	29
921	947	991	1,614	1,152	940	987	948	973	991

Total Cases in Ontario During Autumn vs Winter

season	Year of date	Month of date	
Autumn	2020	October	8,880
		November	34,007
Winter	2021	January	56,077
		February	16,403
		March	61