VISUALIZATION OF THE BUSINESS REQUIREMENTS

TABLEAU VISUALIZATIONS

1.How many accidents occurred in NYC, Austin and Chicago?

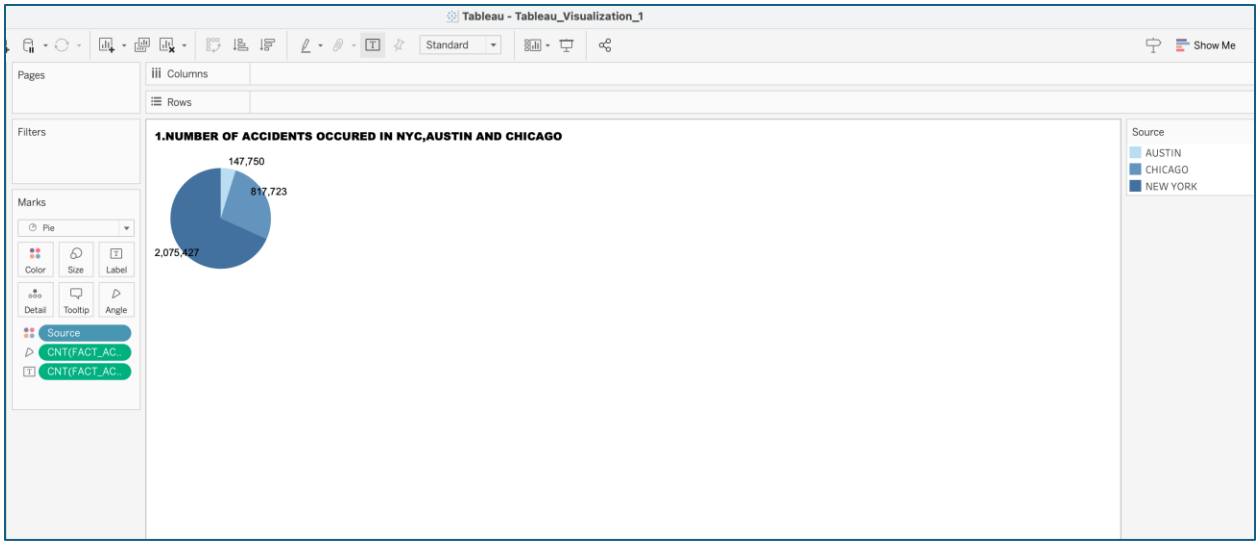
select count(*),s.`Source`
from FACT_ACCIDENTS a join DIM_Source s ON a.SK_Source=s.sk_source
group by a.SK_Source;

Run Cancel Disconnect Change Database: FinalProjectMVC

```
1  --1.How many accidents occurred in NYC, Austin and Chicago?
2  select count(*),s.`Source`
3  from FACT_ACCIDENTS a join DIM_Source s ON a.SK_Source=s.sk_source
4  group by a.SK_Source;
```

Results Messages

	count(*)	Source
1	2075427	NEW YORK
2	817723	CHICAGO
3	147750	AUSTIN



2.Which areas in the 3 cities had the greatest number of accidents?

```
WITH RankedAccidents AS (
  SELECT
    a.sk_location,
    s.SOURCE,
    l.street_name,
    COUNT(*) AS maxAccidents,
    DENSE_RANK() OVER (
      PARTITION BY a.sk_source
      ORDER BY COUNT(*) DESC
    ) AS SourceRank
  FROM FACT_ACCIDENTS a
  JOIN DIM_Location l ON a.sk_location = l.sk_location
  JOIN DIM_Source s ON a.sk_source = s.sk_source
  GROUP BY a.sk_location, a.sk_source, l.street_name
)
SELECT
  ra.sk_location,
  ra.SOURCE,
  ra.street_name,
  ra.maxAccidents,
  ra.SourceRank
FROM RankedAccidents ra
WHERE ra.SourceRank <= 3
ORDER BY ra.SOURCE, ra.SourceRank;
```

Welcome | SQLQuery_1 - localh... (root) | SQLQuery_3.sql - localh... (root) | SQLQuery_2 - localh... (root)

Run | Cancel | Disconnect | Change | Database: FinalProjectMVC

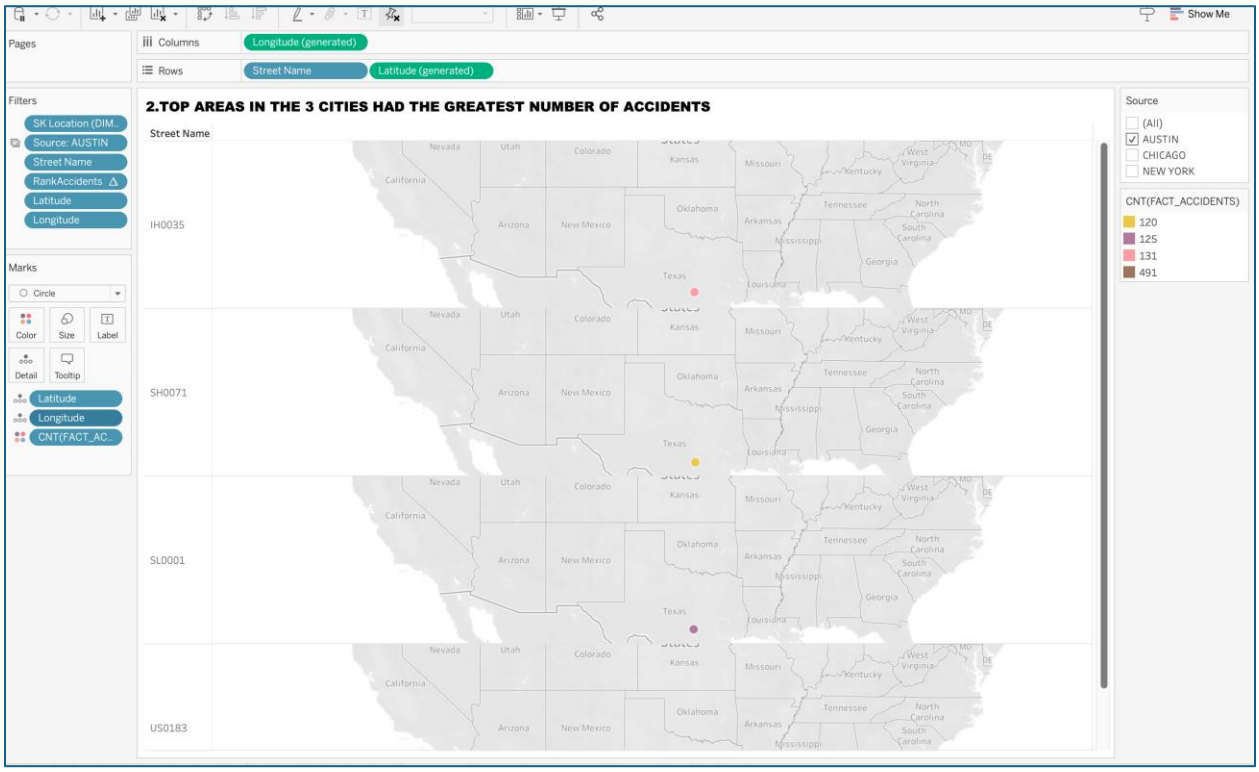
```
--2.Which areas in the 3 cities had the greatest number of accidents?
WITH RankedAccidents AS (
    SELECT
        a.sk_location,
```

Results | Messages

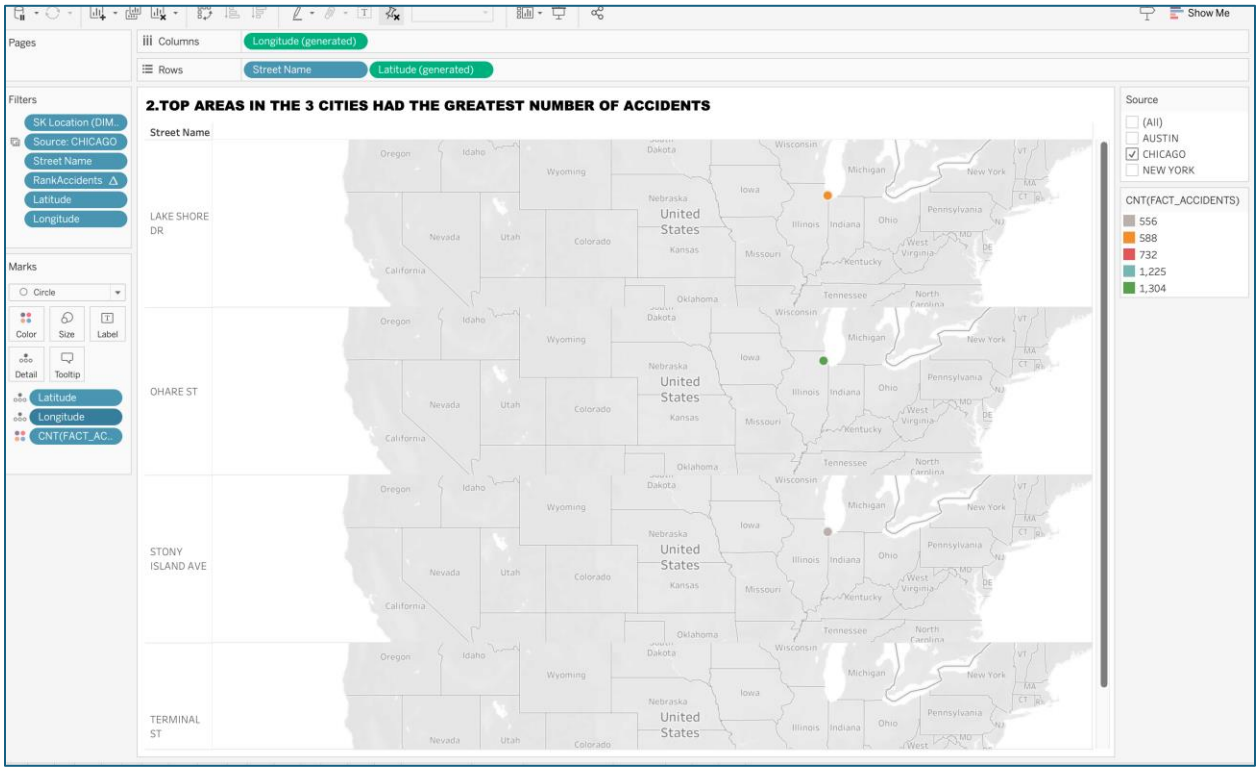
	sk_location	SOURCE	street_name	maxAccidents	SourceRank
1	104	AUSTIN	IH0035	491	1
2	1498	AUSTIN	IH0035	131	2
3	1153	AUSTIN	US0183	125	3
4	1164	AUSTIN	SL0001	125	3
5	4270	AUSTIN	SH0071	120	4
6	297	AUSTIN	SL0001	119	5
7	69432	CHICAGO	OHARE ST	1304	1
8	65672	CHICAGO	TERMINAL ST	1225	2
9	65669	CHICAGO	OHARE ST	732	3
10	73052	CHICAGO	LAKE SHORE DR	588	4
11	72262	CHICAGO	STONY ISLAND AVE	556	5
12	68497	NEW YORK	NA	56129	1
13	344519	NEW YORK	HORACE HARDING EXPRESSWAY	2665	2
14	342763	NEW YORK	FLATBUSH AVENUE	2081	3
15	342906	NEW YORK	VERRAZANO BRIDGE UPPER	1995	4
16	343514	NEW YORK	TRIBOROUGH BRIDGE	1863	5

Results grid

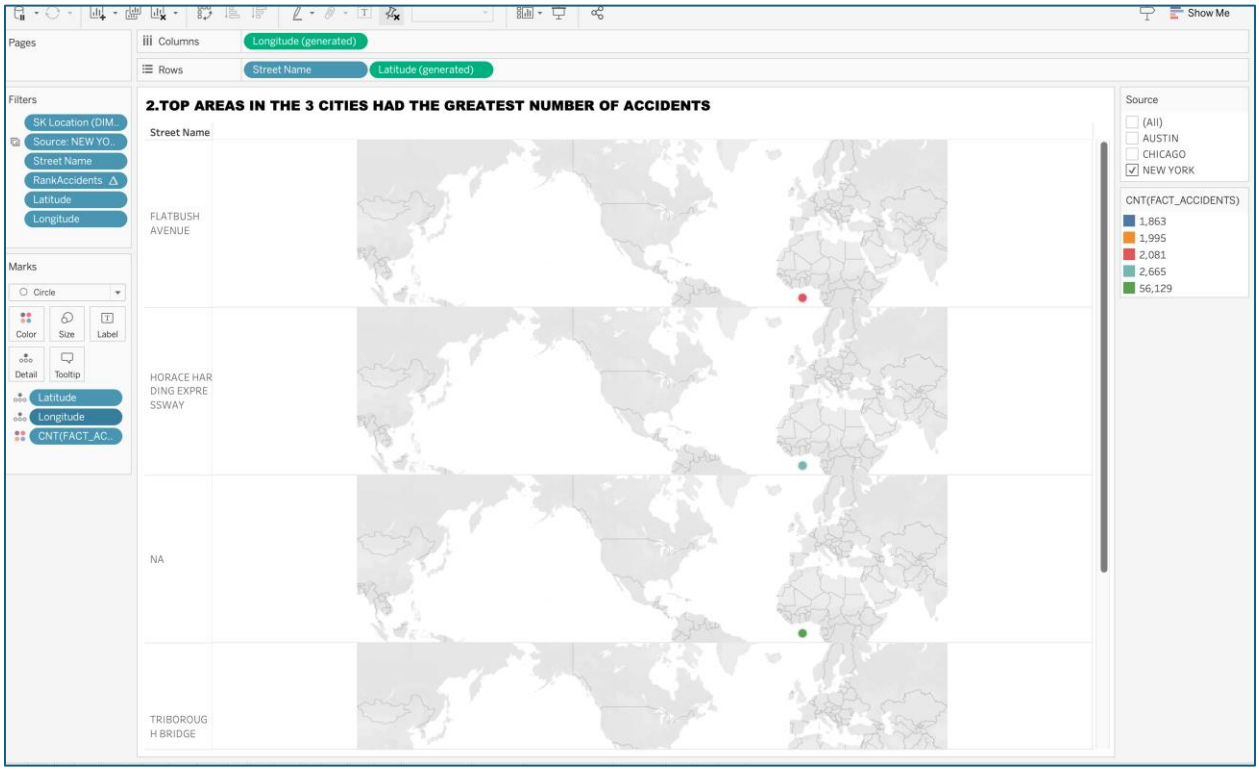
Austin:



CHICAGO:



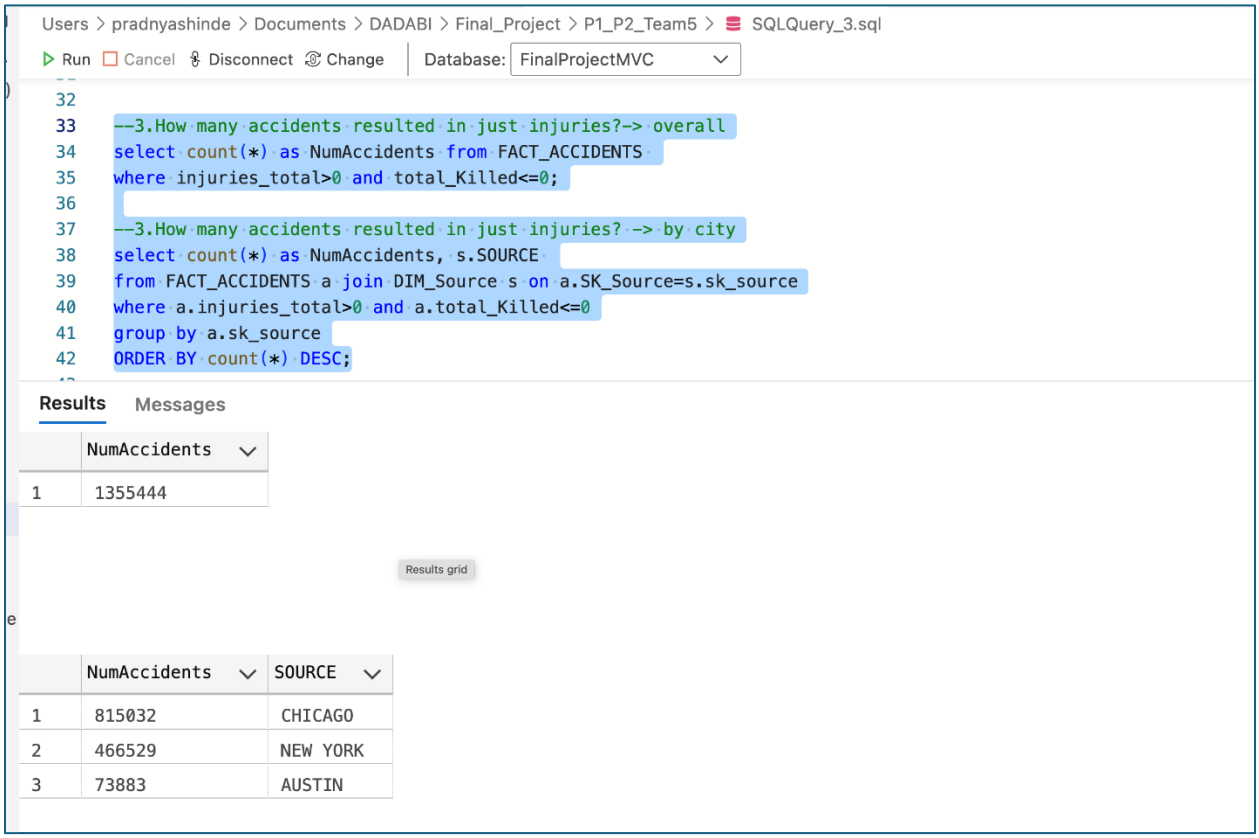
New York:

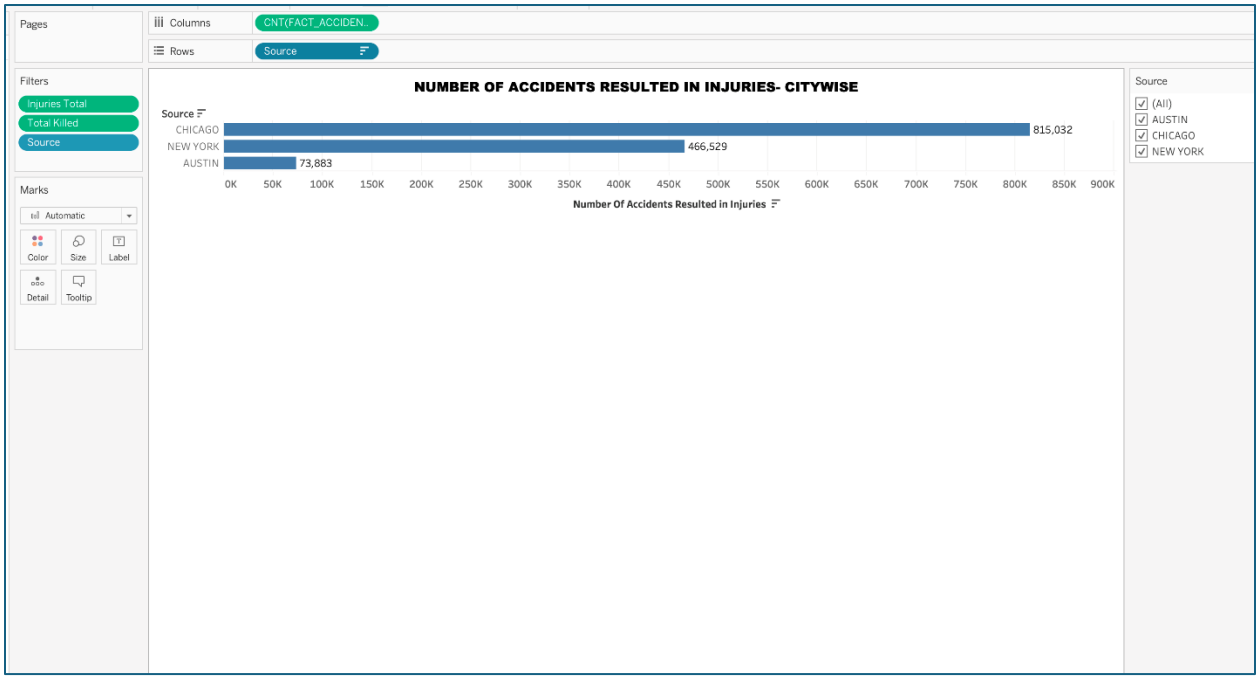


3.How many accidents resulted in just injuries?

```
select count(*) as NumAccidents, s.SOURCE
from FACT_ACCIDENTS a join DIM_Source s on a.SK_Source=s.sk_source
where a.injuries_total>0 and a.total_Killed<=0
group by a.sk_source
ORDER BY count(*) DESC;
```

```
select count(*) as NumAccidents from FACT_ACCIDENTS
where injuries_total>0 and total_Killed<=0;
```





4. How often are pedestrians involved in accidents?

```
select count(*) as NumAccidents from FACT_ACCIDENTS
where is_pedestrian=1;

select count(*) as NumAccidents, s.SOURCE
from FACT_ACCIDENTS a join DIM_Source s on a.SK_Source=s.sk_source
where a.is_pedestrian=1
group by a.sk_source;
```

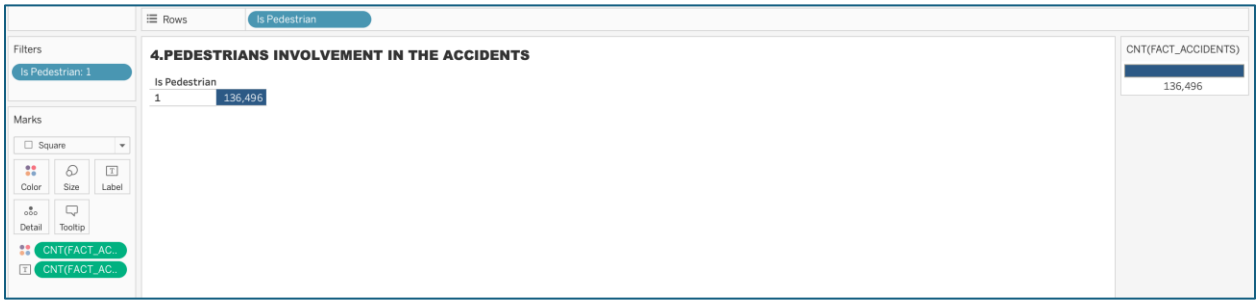
Run Cancel Disconnect Change | Database: FinalProjectMVC

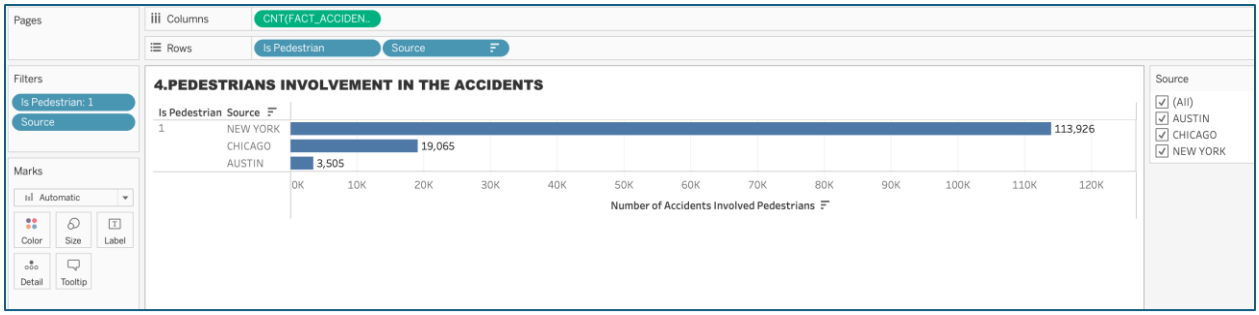
```
1   --4.How often are pedestrians involved in accidents?
2   select count(*) as NumAccidents from FACT_ACCIDENTS
3   where is_pedestrian=1;
4
5   select count(*) as NumAccidents, s.SOURCE
6   from FACT_ACCIDENTS a join DIM_Source s on a.SK_Source=s.sk_source
7   where a.is_pedestrian=1
```

Results Messages

	NumAccidents
1	136496

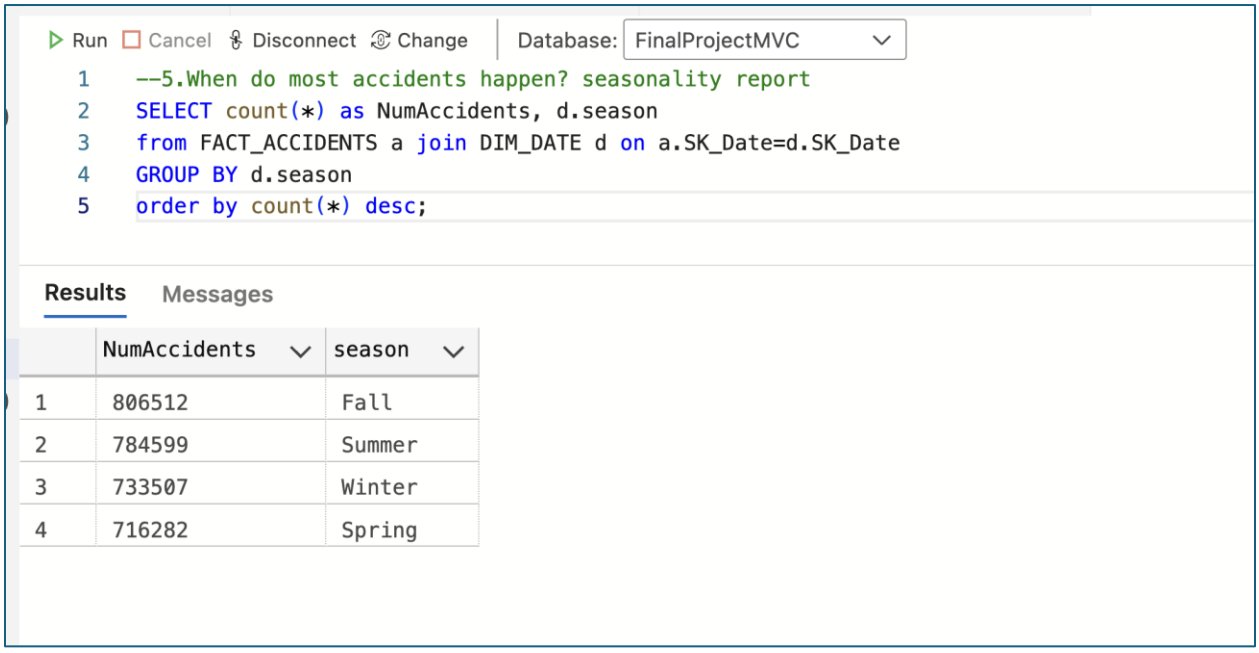
	NumAccidents	SOURCE
1	113926	NEW YORK
2	19065	CHICAGO
3	3505	AUSTIN



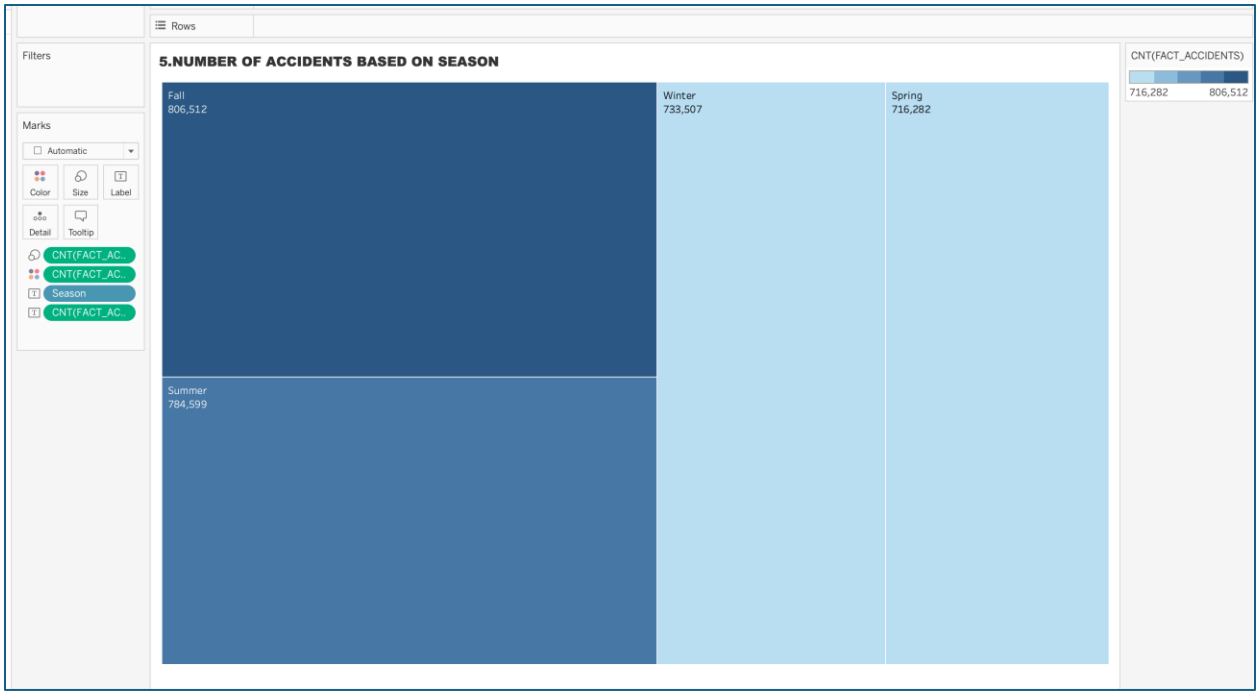


5.When do most accidents happen? seasonality report

SELECT count(*) as NumAccidents, d.season
from FACT_ACCIDENTS a join DIM_DATE d on a.SK_Date=d.SK_Date
GROUP BY d.season
order by count(*) desc;



	NumAccidents	season
1	806512	Fall
2	784599	Summer
3	733507	Winter
4	716282	Spring



6.How many motorists are injured or killed in accidents?

```
select sum(motorist_killed_count + motorist_injury_count) as NumMotorist
from FACT_ACCIDENTS
where motorist_killed_count >0 or motorist_injury_count> 0;
-- -> by city
select sum(a.motorist_killed_count + a.motorist_injury_count) as NumMotorist , s.SOURCE
from FACT_ACCIDENTS a join DIM_Source s on a.SK_Source=s.sk_source
where a.motorist_killed_count >0 or a.motorist_injury_count> 0
group by a.sk_source;
```

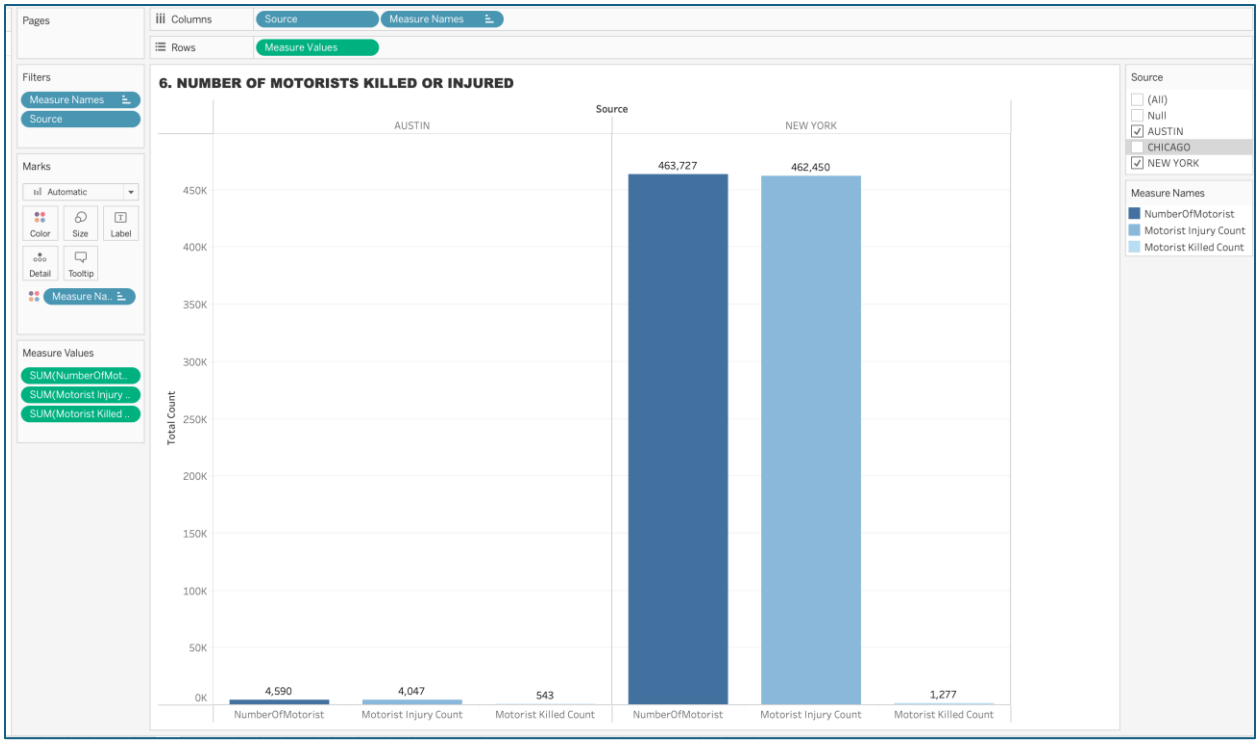
Run Cancel Disconnect Change Database: FinalProjectMVC

```
1  --6.How many motorists are injured or killed in accidents?
2  select sum(motorist_killed_count + motorist_injury_count) as NumMotorist
3  from FACT_ACCIDENTS
4  where motorist_killed_count >0 or motorist_injury_count> 0;
5  -- -> by city
6  select sum(a.motorist_killed_count + a.motorist_injury_count) as NumMotorist , s.SOURCE
7  from FACT_ACCIDENTS a join DIM_Source s on a.SK_Source=s.sk_source
8  where a.motorist_killed_count >0 or a.motorist_injury_count> 0
9  group by a.sk_source;
```

Results Messages

	NumMotorist
1	468317

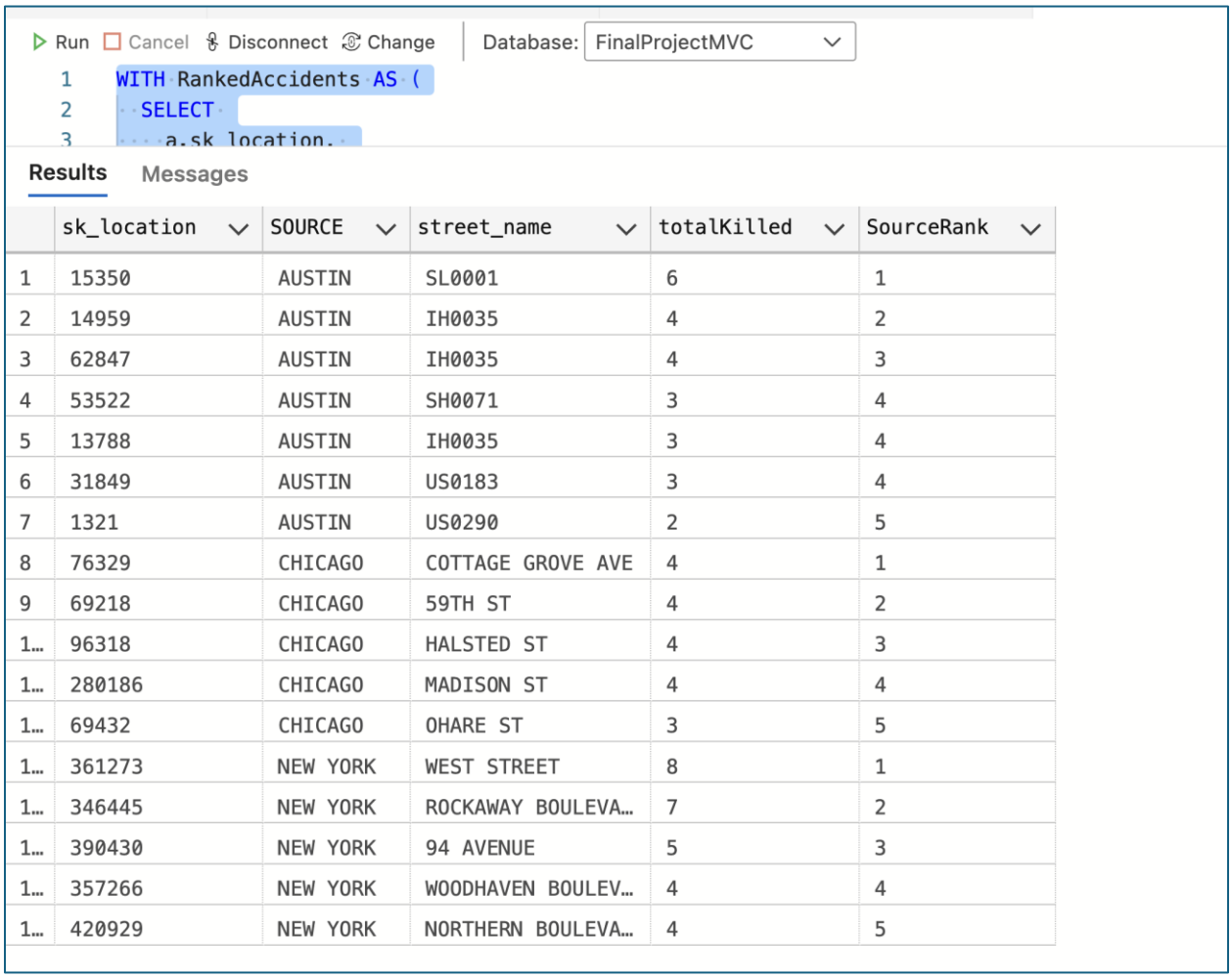
	NumMotorist	SOURCE
1	463727	NEW YORK
2	4590	AUSTIN



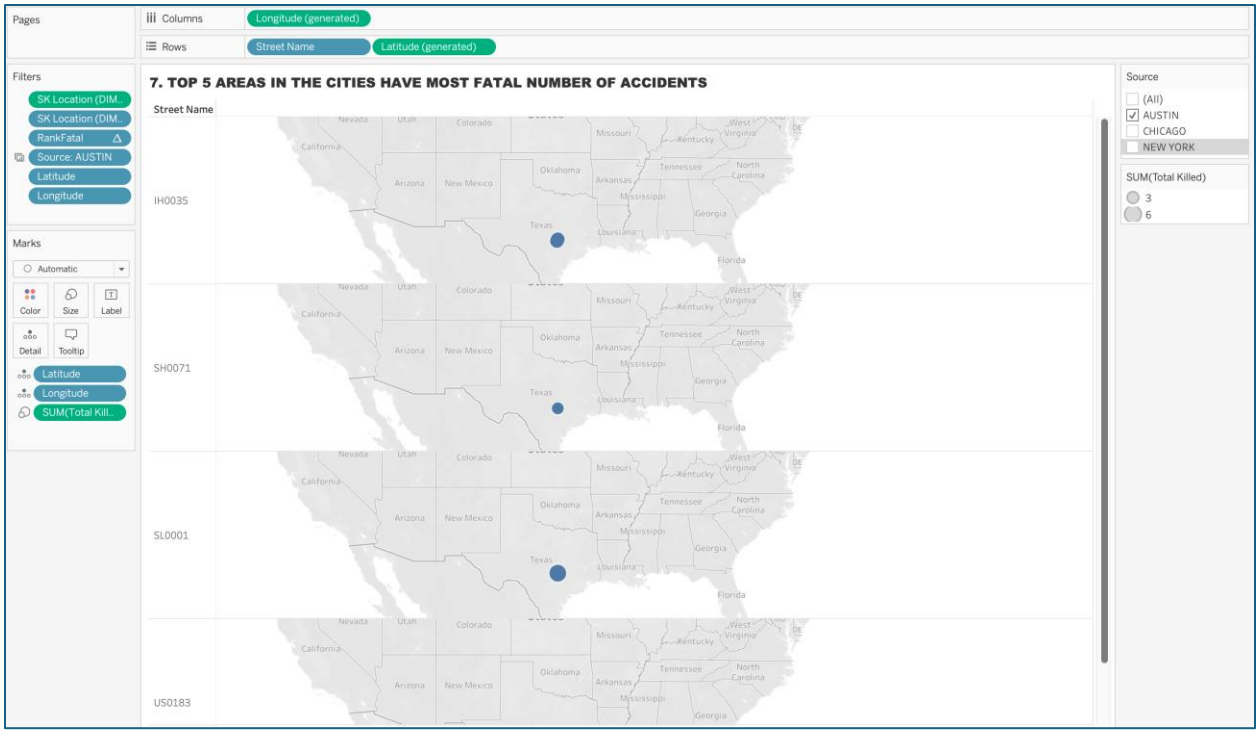
7. Which top 5 areas in 3 cities have the most fatal number of accidents?

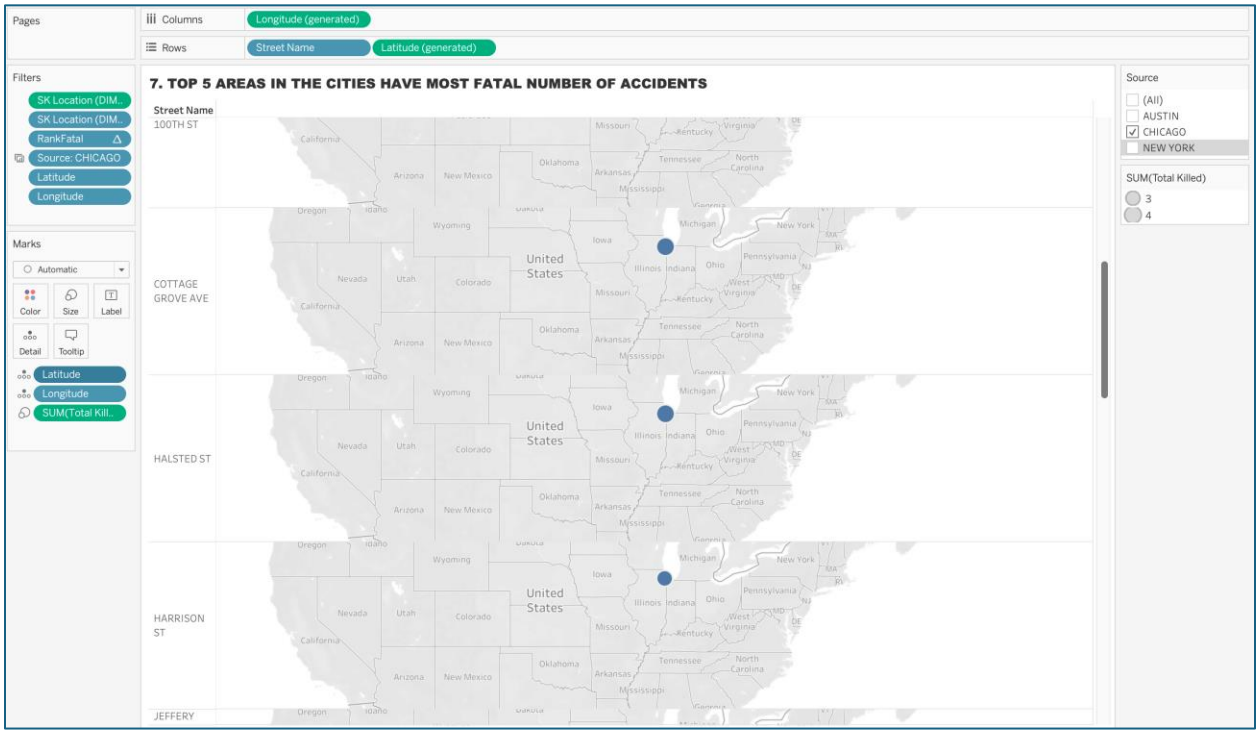
```
WITH RankedAccidents AS (
  SELECT
    a.sk_location,
    s.SOURCE,
    l.street_name,
    SUM(a.total_Killed) as totalKilled,
    DENSE_RANK() OVER (
      PARTITION BY s.SOURCE
      ORDER BY SUM(a.total_Killed) DESC, COUNT(*) DESC
    ) AS SourceRank
  FROM FACT_ACCIDENTS a
  JOIN DIM_Location l ON a.sk_location = l.sk_location
  JOIN DIM_Source s ON a.sk_source = s.sk_source
  WHERE (l.latitude <> -1 ) and ( l.longitude <> -1)
  GROUP BY a.sk_location, s.SOURCE, l.street_name
)
SELECT
  ra.sk_location,
  ra.SOURCE,
  ra.street_name,
  ra.totalKilled,
  ra.SourceRank
FROM RankedAccidents ra
```

```
WHERE ra.SourceRank <= 5
ORDER BY ra.SOURCE, ra.totalKilled DESC, ra.SourceRank;
```



	sk_location	SOURCE	street_name	totalKilled	SourceRank
1	15350	AUSTIN	SL0001	6	1
2	14959	AUSTIN	IH0035	4	2
3	62847	AUSTIN	IH0035	4	3
4	53522	AUSTIN	SH0071	3	4
5	13788	AUSTIN	IH0035	3	4
6	31849	AUSTIN	US0183	3	4
7	1321	AUSTIN	US0290	2	5
8	76329	CHICAGO	COTTAGE GROVE AVE	4	1
9	69218	CHICAGO	59TH ST	4	2
1…	96318	CHICAGO	HALSTED ST	4	3
1…	280186	CHICAGO	MADISON ST	4	4
1…	69432	CHICAGO	OHARE ST	3	5
1…	361273	NEW YORK	WEST STREET	8	1
1…	346445	NEW YORK	ROCKAWAY BOULEVA…	7	2
1…	390430	NEW YORK	94 AVENUE	5	3
1…	357266	NEW YORK	WOODHAVEN BOULEV…	4	4
1…	420929	NEW YORK	NORTHERN BOULEVA…	4	5


Pages
Columns
Longitude (generated)
Rows
Street Name
Latitude (generated)
Filters
SK Location (DIM..
SK Location (DIM..
RankFatal
Source: AUSTIN
Latitude
Longitude
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Latitude
Longitude
SUM(Total Kill..
7. TOP 5 AREAS IN THE CITIES HAVE MOST FATAL NUMBER OF ACCIDENTS
Street Name
IH0035
SH0071
SL0001
US0183
Source
(All)
AUSTIN
CHICAGO
NEW YORK
SUM(Total Killed)
3
6


Pages
Columns
Longitude (generated)
Rows
Street Name
Latitude (generated)
Filters
SK Location (DIM..
SK Location (DIM..
RankFatal
Source: CHICAGO
Latitude
Longitude
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Latitude
Longitude
SUM(Total Kill..
7. TOP 5 AREAS IN THE CITIES HAVE MOST FATAL NUMBER OF ACCIDENTS
Street Name
100TH ST
COTTAGE GROVE AVE
HALSTED ST
HARRISON ST
JEFFERY
United States
Source
(All)
AUSTIN
CHICAGO
NEW YORK
SUM(Total Killed)
3
4

8. Time based analysis of accidents

Time of the day, day of the week, weekdays or weekends.

Run Cancel Disconnect Change Database: FinalProjectMVC

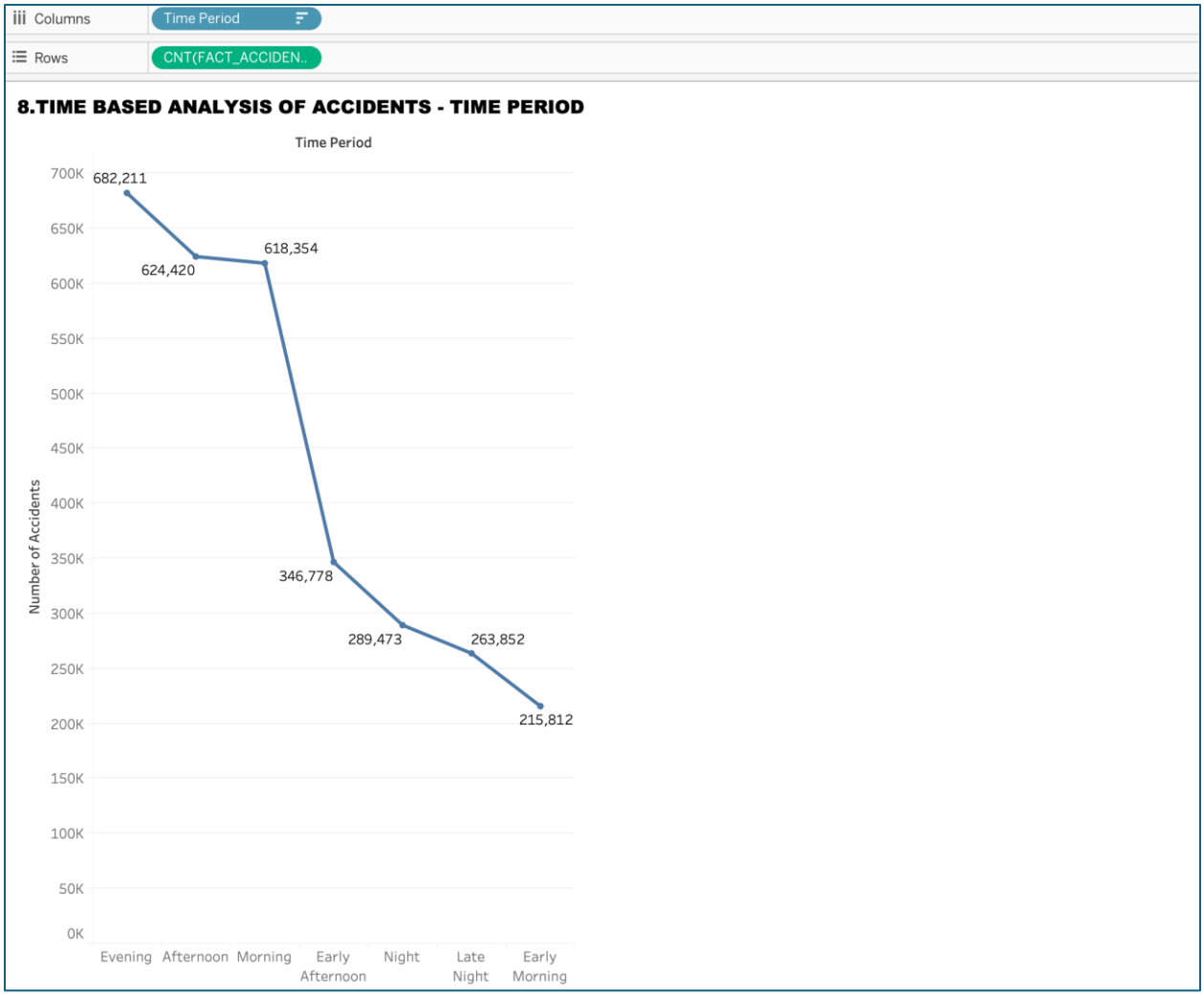
```
1  --8.Time based analysis of accidents
2  ------Time of the day, day of the week, weekdays or weekends.
3  SELECT count(*) as NumAccidents,t.time_period
4  from FACT_ACCIDENTS a join DIM_TIME t on a.SK_Time=t.SK_Time
5  GROUP BY t.time_period
```

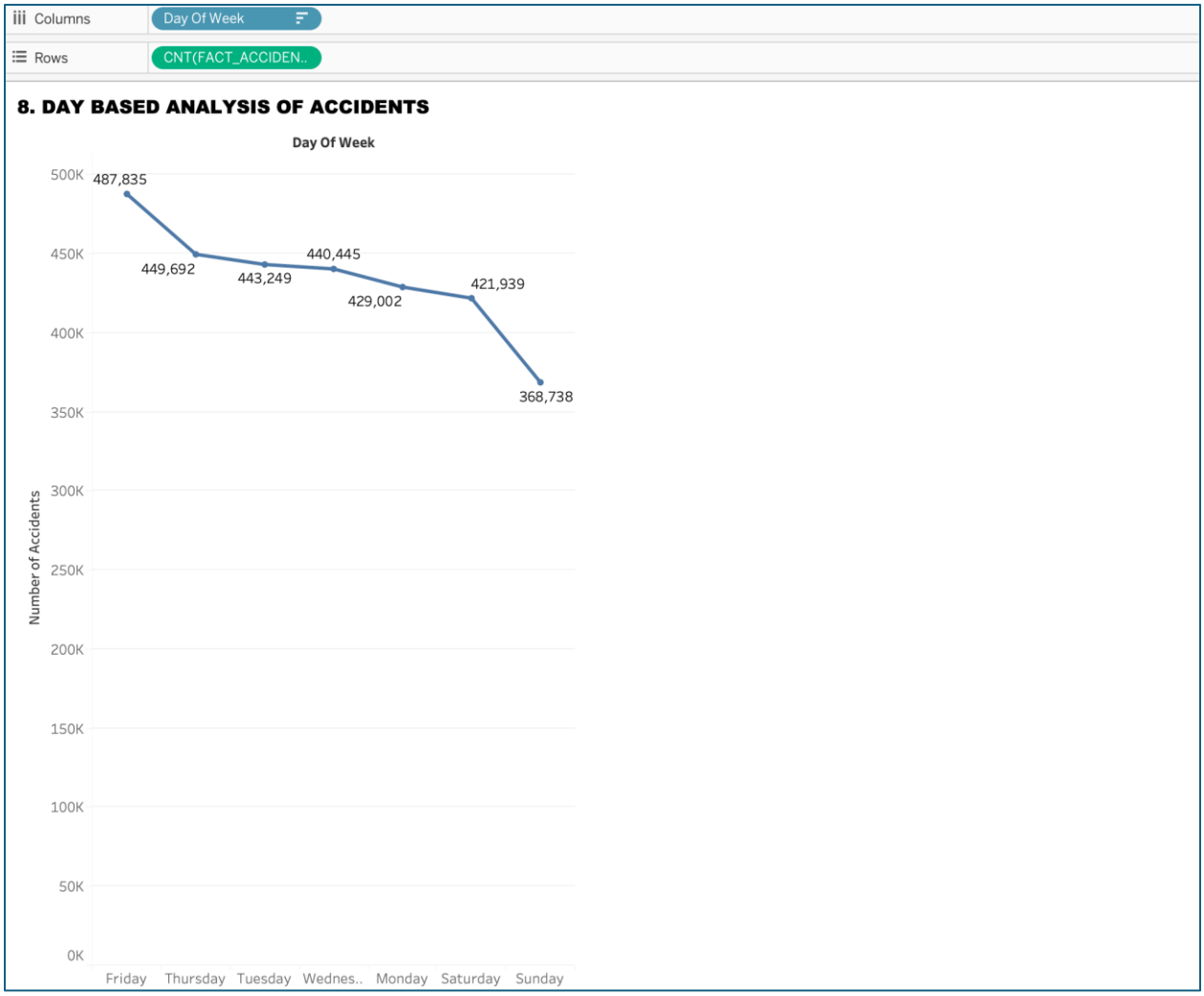
Results Messages

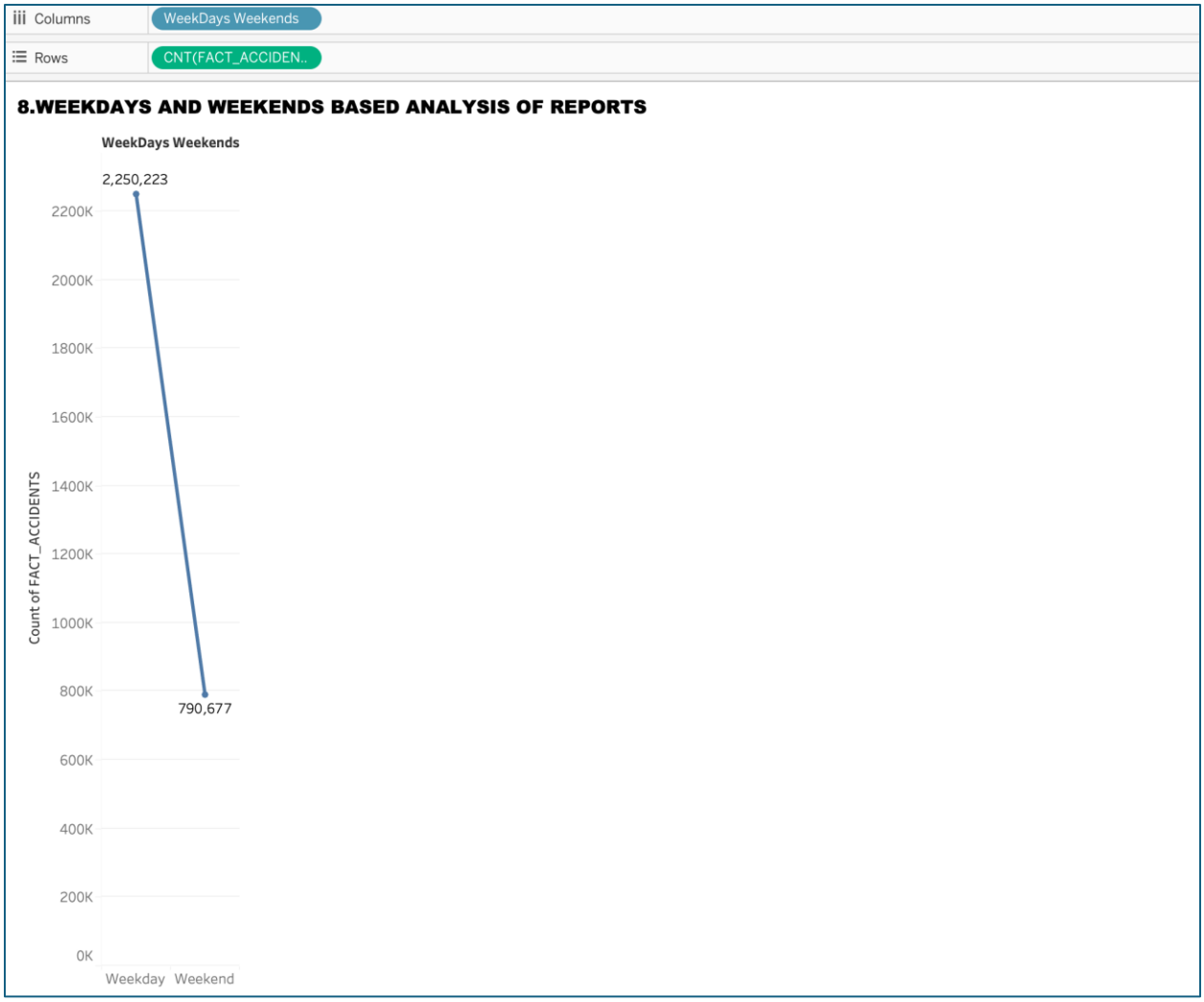
	NumAccidents	time_period
1	682211	Evening
2	624420	Afternoon
3	618354	Morning
4	346778	Early Afternoon
5	289473	Night
6	263852	Late Night
7	215812	Early Morning

Results grid

	NumAccidents	DayOfWeek
1	487835	Friday
2	449692	Thursday
3	443249	Tuesday
4	440445	Wednesday
5	429002	Monday
6	421939	Saturday
7	368738	Sunday


Columns
Time Period
Rows
CNT(FACT_ACCIDEN..
8.TIME BASED ANALYSIS OF ACCIDENTS - TIME PERIOD
Time Period
Number of Accidents
700K
650K
600K
550K
500K
450K
400K
350K
300K
250K
200K
150K
100K
50K
0K
682,211
624,420
618,354
346,778
289,473
263,852
215,812
Evening
Afternoon
Morning
Early Afternoon
Night
Late Night
Early Morning


Columns
Day Of Week
Rows
CNT(FACT_ACCIDEN..
8. DAY BASED ANALYSIS OF ACCIDENTS
Day Of Week
Number of Accidents
500K
450K
400K
350K
300K
250K
200K
150K
100K
50K
0K
487,835
449,692
443,249
440,445
429,002
421,939
368,738
Friday
Thursday
Tuesday
Wednes..
Monday
Saturday
Sunday



9. Fatality analysis - Are pedestrians killed more often than road users

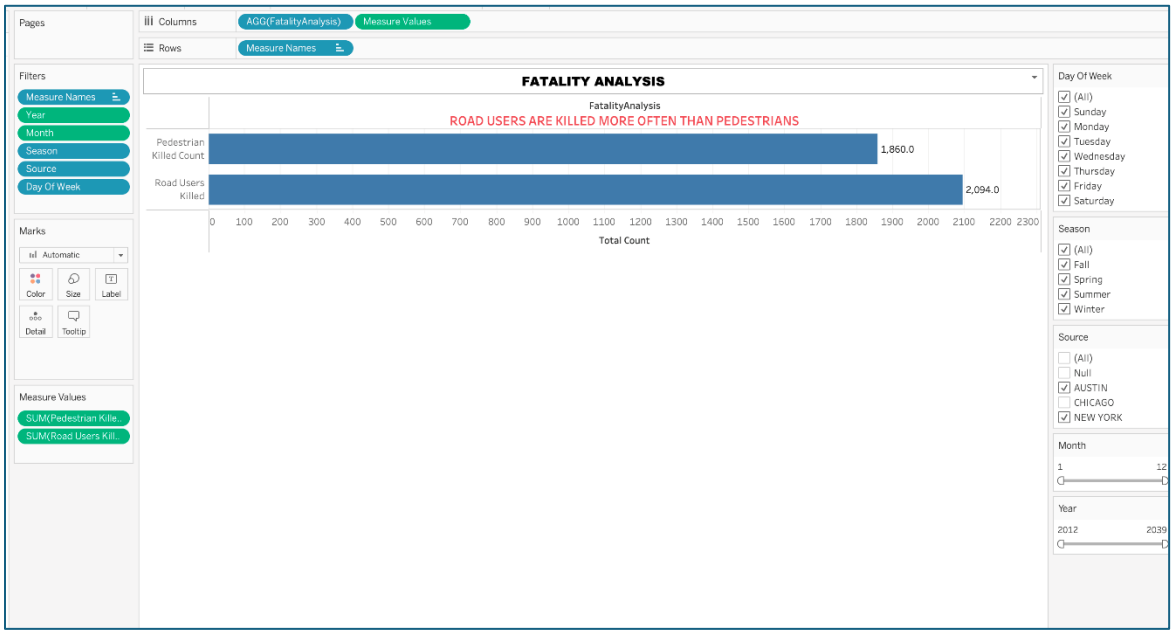
Users > pradnyashinde > Documents > DADABI > Final_Project > P1_P2_Team5 > SQLQuery_3.sql

Run | Cancel | Disconnect | Change | Database: FinalProjectMVC

```
114
115   --9.Fatality analysis
116   ------Are pedestrians killed more often than road users?
117   SELECT
118     (SELECT SUM(PEDESTRIAN_KILLED_COUNT) FROM FACT_ACCIDENTS) AS TotalPedestrianFatalities,
119     (SELECT SUM(road_users_killed) FROM FACT_ACCIDENTS) AS TotalMotoristFatalities,
120     CASE
121       WHEN (SELECT SUM(PEDESTRIAN_KILLED_COUNT) FROM FACT_ACCIDENTS) >
122            (SELECT SUM(road_users_killed) FROM FACT_ACCIDENTS)
123       THEN 'Pedestrians killed more often'
124       WHEN (SELECT SUM(PEDESTRIAN_KILLED_COUNT) FROM FACT_ACCIDENTS) <
125            (SELECT SUM(road_users_killed) FROM FACT_ACCIDENTS)
126       THEN 'Road Users killed more often'
127       ELSE 'Pedestrians and Road Users killed equally often'
128     END AS FatalityComparison
```

Results | Messages

	TotalPedestrianFatalities	TotalMotoristFatalities	FatalityComparison
1	1860	2094	Road Users killed more often



10.What are the most common factors involved in accidents?

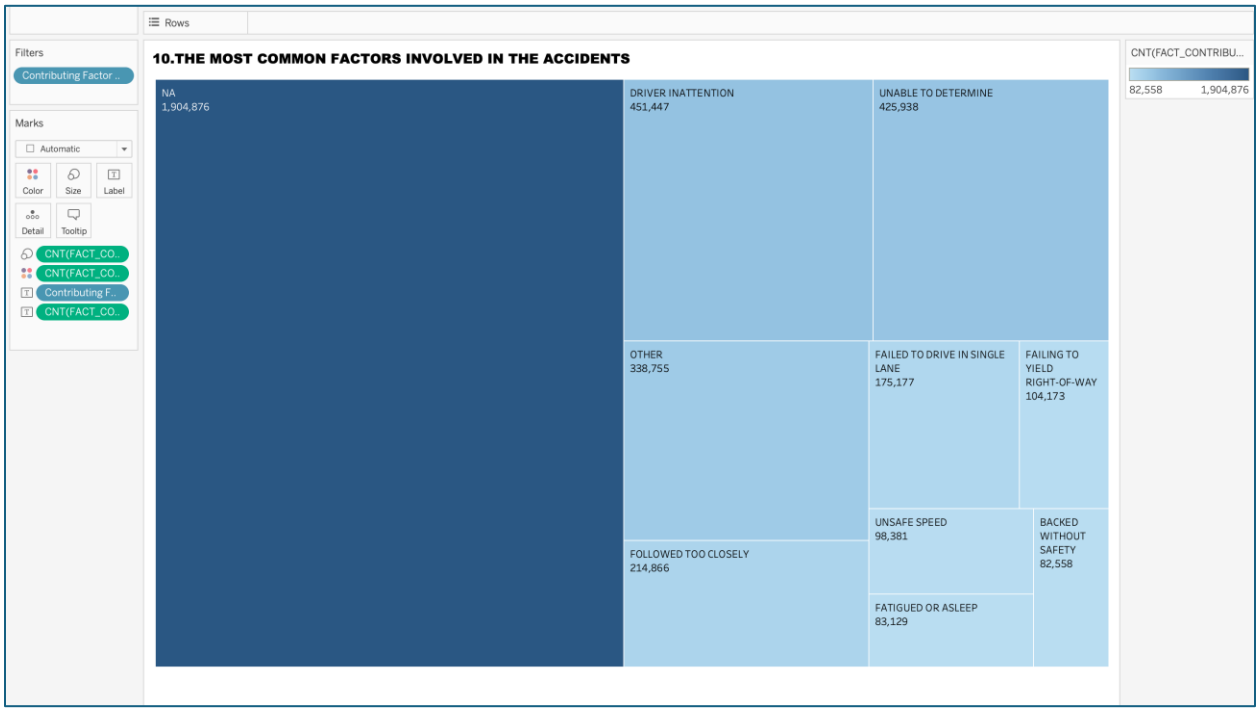
```
SELECT count(*) as
NumAccidents,c.contributing_factor_code,c.contributing_factor_description
from FACT_CONTRIBUTION f join Dim_Contribution c on
f.sk_contribution=c.sk_contribution
GROUP BY f.sk_contribution
order by count(*) desc
limit 10;
```

Run Cancel Disconnect Change Database: FinalProjectMVC

```
1  --10.What are the most common factors involved in accidents?
2  SELECT count(*) as NumAccidents,c.contributing_factor_code,c.contributing_factor_description
3  from FACT_CONTRIBUTION f join Dim_Contribution c on f.sk_contribution=c.sk_contribution
4  GROUP BY f.sk_contribution
5  order by count(*) desc
6  limit 10;
```

Results Messages

	NumAccidents	contributing_factor_code	contributing_factor_description
1	1904876	-1	NA
2	451447	20	DRIVER INATTENTION
3	425938	132	UNABLE TO DETERMINE
4	338755	101	OTHER
5	214866	44	FOLLOWED TOO CLOSELY
6	175177	23	FAILED TO DRIVE IN SINGLE LANE
7	104173	126	FAILING TO YIELD RIGHT-OF-WAY
8	98381	60	UNSAFE SPEED
9	83129	40	FATIGUED OR ASLEEP
10	82558	3	BACKED WITHOUT SAFETY



11.Store vehicle type/vehicles involved in accidents at its least graunality

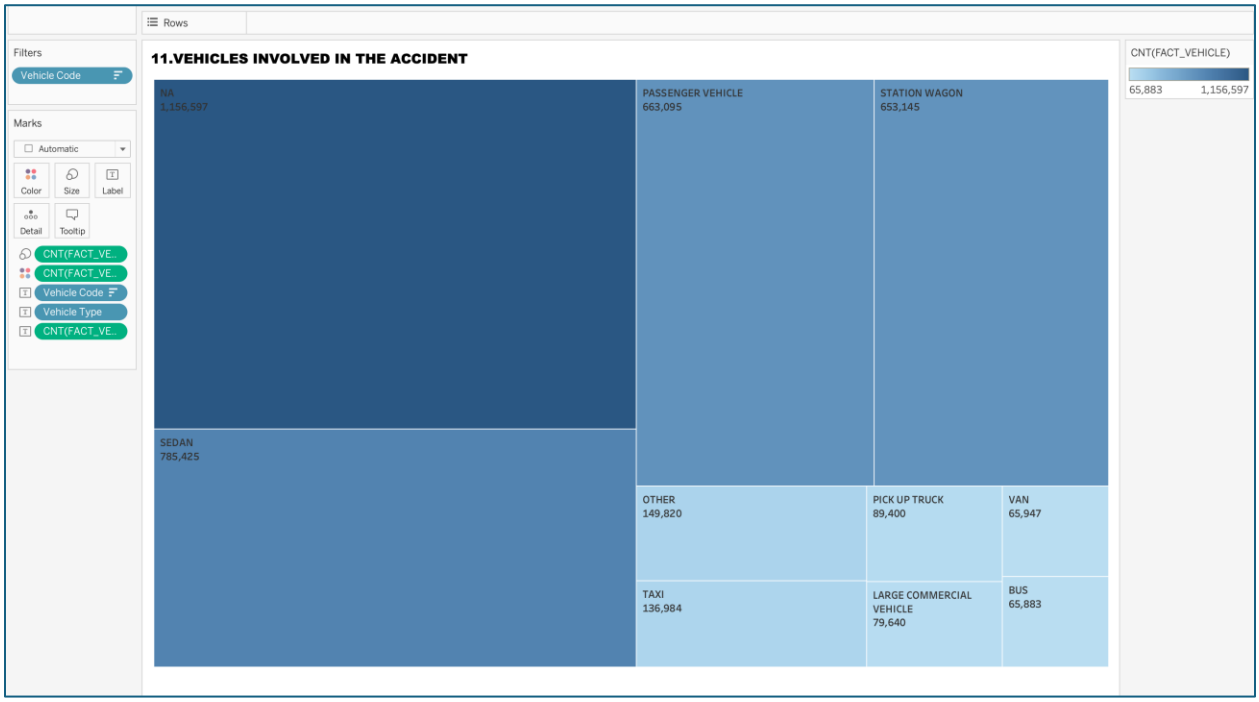
```
SELECT count(*) as NumAccidents,v.VEHICLE_CODE,v.VEHICLE_TYPE
from FACT_VEHICLE f join DIM_Vehicle_Type v on
f.SK_VEHICLE_TYPE=v.SK_VEHICLE_TYPE
GROUP BY f.SK_VEHICLE_TYPE
order by count(*) desc
limit 10;
```

Run Cancel Disconnect Change Database: FinalProjectMVC

```
1  --11.Store vehicle type/vehicles involved in accidents at its least graunality
2  SELECT count(*) as NumAccidents,v.VEHICLE_CODE,v.VEHICLE_TYPE
3  from FACT_VEHICLE f join DIM_Vehicle_Type v on f.SK_VEHICLE_TYPE=v.SK_VEHICLE_TYPE
4  GROUP BY f.SK_VEHICLE_TYPE
5  order by count(*) desc
6  limit 10;
```

Results Messages

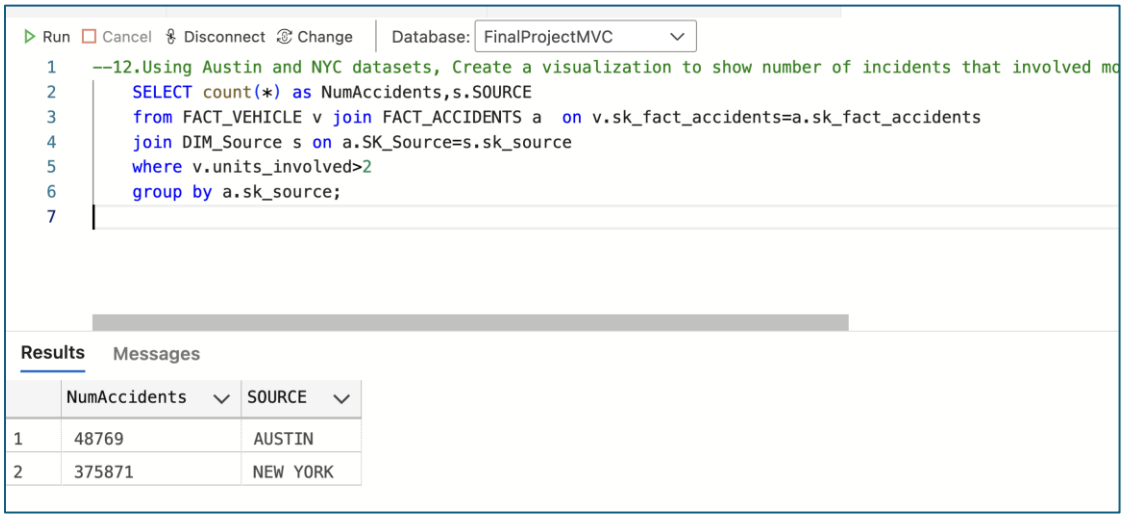
	NumAccidents	VEHICLE_CODE	VEHICLE_TYPE
1	1156597	219	NA
2	785425	168	SEDAN
3	663095	144	PASSENGER VEHICLE
4	653145	184	STATION WAGON
5	149820	142	OTHER
6	136984	188	TAXI
7	89400	147	PICK UP TRUCK
8	79640	109	LARGE COMMERCIAL VEHICLE
9	65947	205	VAN
10	65883	33	BUS

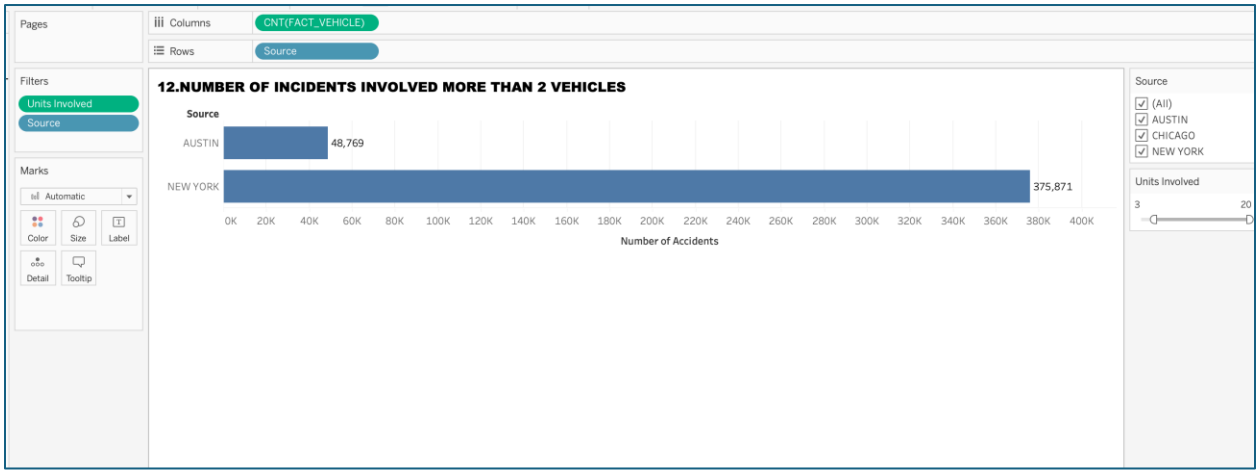


12.Using Austin and NYC datasets, Create a visualization to show number of incidents that involved more than 2 vehicles. Show this data as a comparision between these 2 cities.

```
    SELECT count(*) as NumAccidents,s.SOURCE
    from FACT_VEHICLE v join FACT_ACCIDENTS a  on
v.sk_fact_accidents=a.sk_fact_accidents
    join DIM_Source s on a.SK_Source=s.sk_source
    where v.units_involved>2
    group by a.sk_source;
```

```
SELECT
  count(*) as NumAccidents,
  s.SOURCE
FROM FACT_ACCIDENTS a
JOIN FACT_VEHICLE v ON v.sk_fact_accidents = a.sk_fact_accidents
JOIN DIM_Source s ON a.SK_Source = s.sk_source
WHERE v.units_involved > 2
GROUP BY a.sk_source, s.SOURCE;
```





POWER BI VISUALIZATIONS

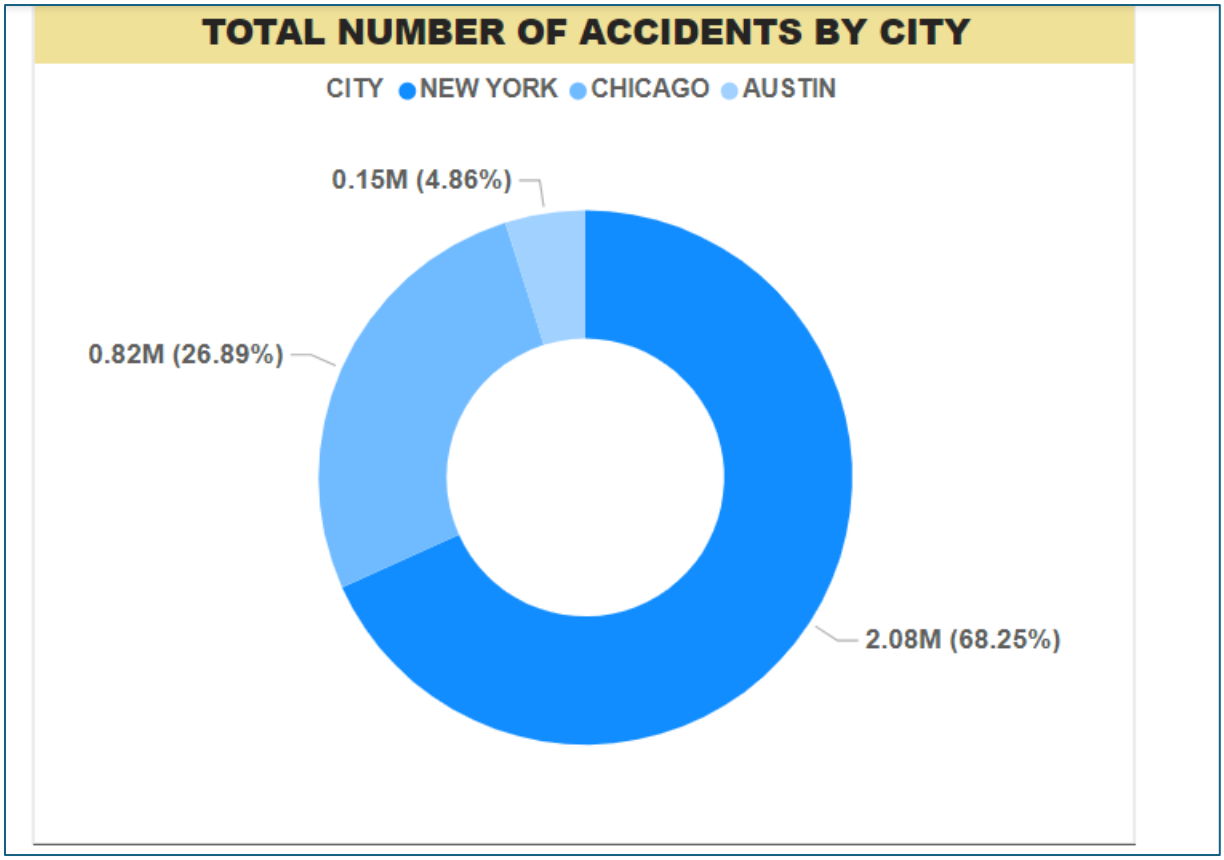
1.

```
--Q1. How many accidents occurred in NYC, Austin and Chicago

SELECT D.SOURCE,COUNT(SK_FACT_ACCIDENTS) AS TOTAL_NO_OF_ACCIDENTS FROM FACT_ACCIDENTS F
JOIN
    DIM_Source D ON D.SK_Source=F.SK_Source
GROUP BY
    D.SOURCE
```

110 %

Results | Messages

	SOURCE	TOTAL_NO_OF_ACCIDENTS
1	AUSTIN	147750
2	CHICAGO	817723
3	NEW YORK	2075427



2.

```
SELECT TOP 3
    a.sk_location,
    s.source,
    COUNT(*) AS maxAccidents,
    l.street_name
FROM FACT_ACCIDENTS a
JOIN DIM_Location l ON a.SK_Location = l.sk_location
JOIN DIM_Source s ON a.SK_Source = s.SK_Source
GROUP BY a.sk_location, s.source, l.street_name
ORDER BY COUNT(*) DESC;
```

110 %

Results | Messages

	sk_location	source	maxAccidents	street_name
1	303441	NEW YORK	56129	NA
2	398506	NEW YORK	2665	HORACE HARDING EXPRESSWAY
3	399251	NEW YORK	2081	FLATBUSH AVENUE

TOP 3 AREAS WITH HIGHEST ACCIDENT COUNTS

Count of Accidents	Latitude	Longitude	Street_Name	City
56129	-1.00	-1.00	NA	NEW YORK
2665	-1.00	-1.00	HORACE HARDING EXPRESSWAY	NEW YORK
2081	-1.00	-1.00	FLATBUSH AVENUE	NEW YORK
1	-1.00	-1.00	NA	CHICAGO
60876				

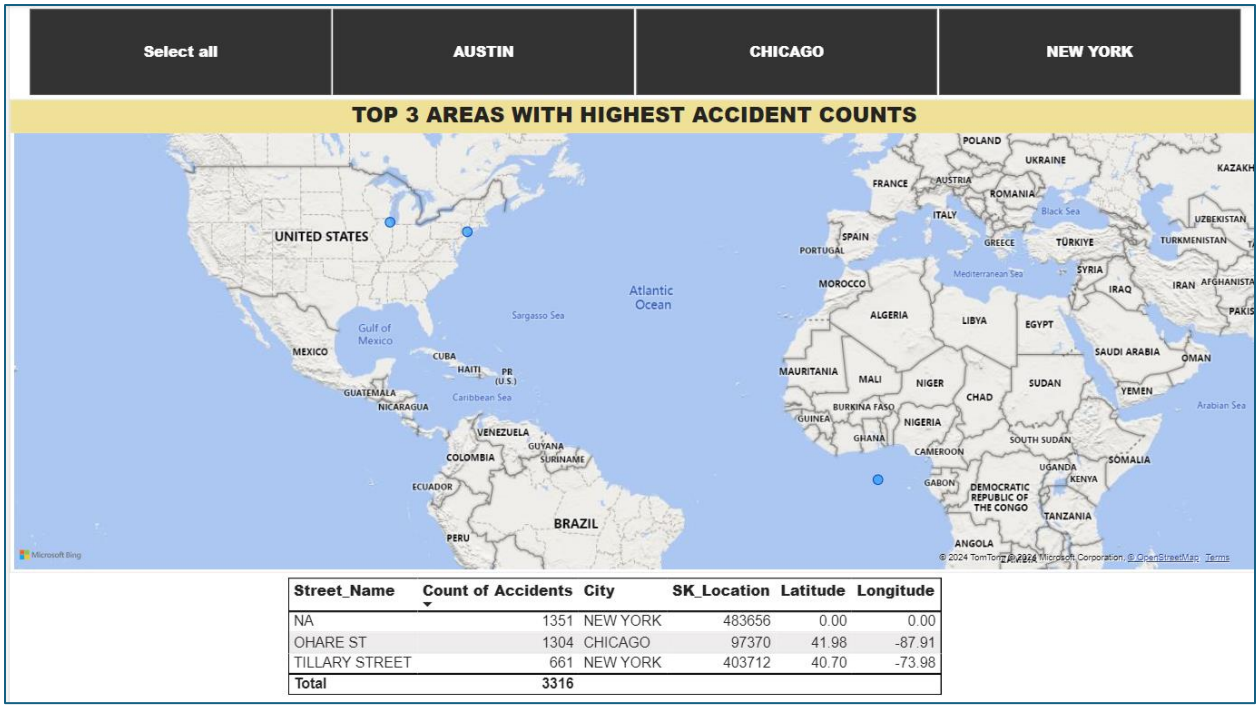
```sql
--Q2. Which areas in the 3 cities had the greatest number of accidents

--top areas within each city

WITH RankedLocations AS (
    SELECT
        a.sk_location,
        s.source,
        l.street_name,
        l.latitude,
        l.longitude,
        COUNT(*) AS maxAccidents,
        RANK() OVER (PARTITION BY a.sk_source ORDER BY COUNT(*) DESC) AS Rank
    FROM FACT_ACCIDENTS a
    JOIN DIM_Location l ON a.SK_Location = l.sk_location
    JOIN DIM_Source s ON a.SK_Source = s.SK_Source
    WHERE l.latitude != -1 AND l.longitude != -1
    GROUP BY a.sk_source, s.source, a.sk_location, l.street_name, l.latitude, l.longitude
)
SELECT
    sk_location,
    source,
    street_name,
    latitude,
    longitude,
    maxAccidents
FROM RankedLocations
WHERE Rank <= 3
ORDER BY source, maxAccidents DESC;
```

110 %

Results | Messages

	sk_location	source	street_name	latitude	longitude	maxAccidents
1	1219	AUSTIN	SL0001	30.24915	-97.8053	111
2	4413	AUSTIN	SH0071	30.20236	-97.63786	107
3	2032	AUSTIN	US0183	30.32557	-97.67332	97
4	97370	CHICAGO	OHARE ST	41.9762	-87.90531	1304
5	98298	CHICAGO	LAKE SHORE DR	41.79142	-87.58015	578
6	100230	CHICAGO	STONY ISLAND AVE	41.75146	-87.58598	556
7	483656	NEW YORK	NA	0	0	1351
8	403712	NEW YORK	TILLARY STREET	40.69603	-73.98453	661
9	404099	NEW YORK	FLATBUSH AVENUE EXTENSION	40.69603	-73.98453	568

Query executed successfully. VAISHVEER (16.0 RTM)



Street_Name	Count of Accidents	City	SK_Location	Latitude	Longitude
NA	1351	NEW YORK	483656	0.00	0.00
OHARE ST	1304	CHICAGO	97370	41.98	-87.91
TILLARY STREET	661	NEW YORK	403712	40.70	-73.98
Total	**3316**				

3.

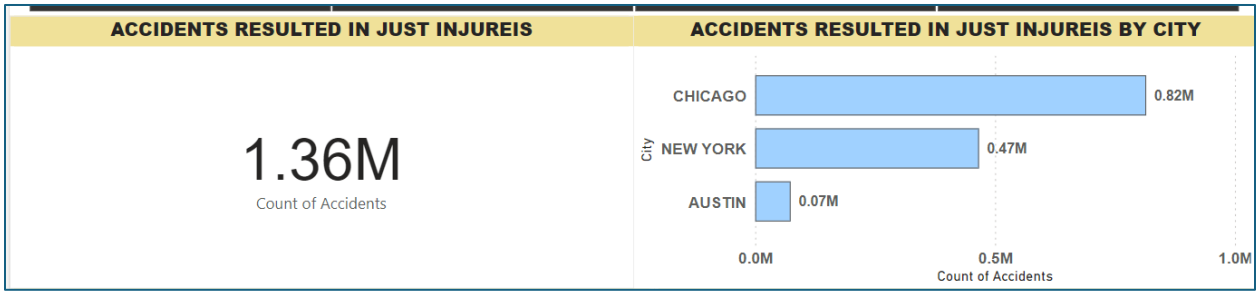
```
SELECT COUNT(*) AS TOTAL_ACCIDENTS,S.SOURCE
FROM FACT_ACCIDENTS f
join DIM_SOURCE S on S.SK_Source=F.SK_Source
WHERE f.INJURIES_TOTAL>0 and f.TOTAL_KILLED<=0
GROUP BY S.SOURCE
ORDER BY TOTAL_ACCIDENTS DESC

SELECT COUNT(*) AS TOTAL_ACCIDENTS
FROM FACT_ACCIDENTS
WHERE INJURIES_TOTAL>0 and TOTAL_KILLED<=0
```

110 %

Results | Messages

	TOTAL_ACCIDENTS	SOURCE
1	815032	CHICAGO
2	466529	NEW YORK
3	73883	AUSTIN

	TOTAL_ACCIDENTS
1	1355444



4.

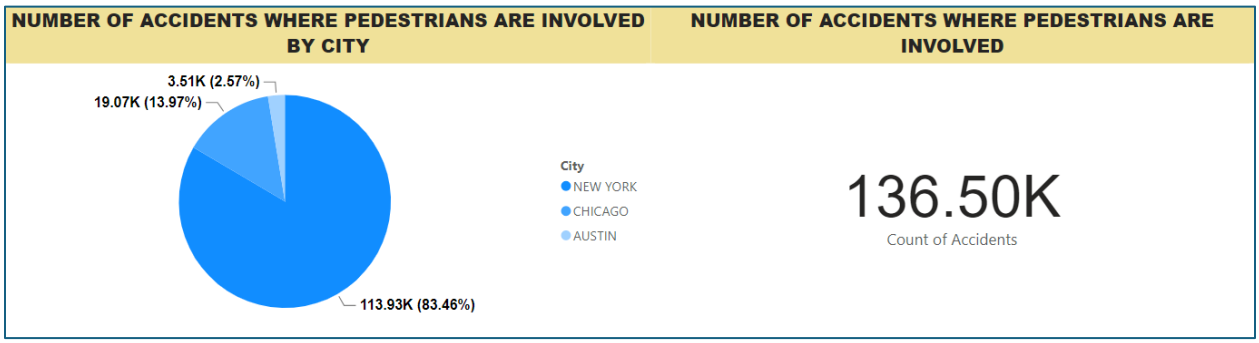
```
--Q4 How often are pedestrians involved in accidents

SELECT COUNT(SK_FACT_ACCIDENTS) AS TOTAL_NO_OF_ACCIDENTS_WHERE_PEDESTRIANS_INVOLVED
FROM FACT_ACCIDENTS
WHERE IS_PEDESTRIAN=1

SELECT COUNT(F.SK_FACT_ACCIDENTS) AS TOTAL_NO_OF_ACCIDENTS_WHERE_PEDESTRIANS_INVOLVED,S.Source
FROM FACT_ACCIDENTS F
JOIN DIM_Source S ON S.SK_Source=F.SK_Source
WHERE IS_PEDESTRIAN=1
GROUP BY S.Source
ORDER BY TOTAL_NO_OF_ACCIDENTS_WHERE_PEDESTRIANS_INVOLVED DESC
```

110 %

Results | Messages

	TOTAL_NO_OF_ACCIDENTS_WHERE_PEDESTRIANS_INVOLVED
1	136496

	TOTAL_NO_OF_ACCIDENTS_WHERE_PEDESTRIANS_INVOLVED	Source
1	113926	NEW YORK
2	19065	CHICAGO
3	3505	AUSTIN



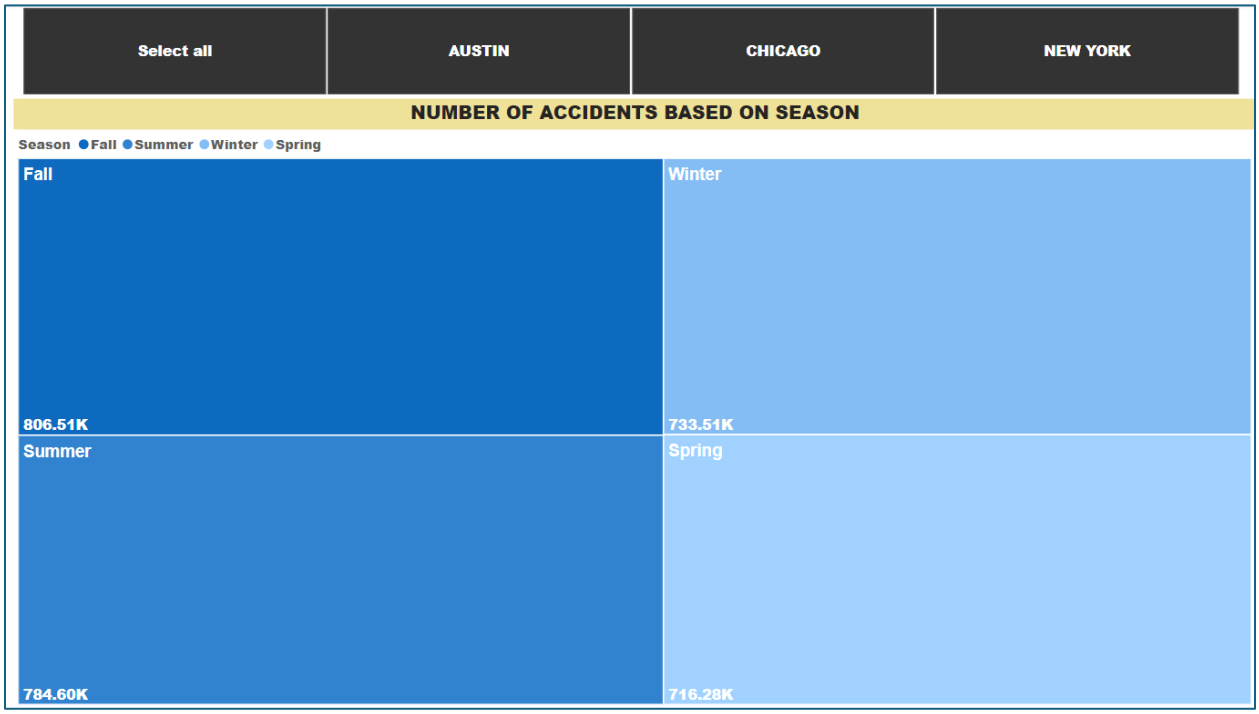
5.

```sql
--Q5 When do most accidents happen

SELECT COUNT(F.SK_FACT_ACCIDENTS) AS TOTAL_NO_OF_ACCIDENTS,D.Season
FROM FACT_ACCIDENTS F
JOIN DIM_DATE D ON D.SK_Date=F.SK_Date
GROUP BY D.Season
ORDER BY TOTAL_NO_OF_ACCIDENTS DESC
```

110 %

Results | Messages

	TOTAL_NO_OF_ACCIDENTS	Season
1	806512	Fall
2	784599	Summer
3	733507	Winter
4	716282	Spring



6.

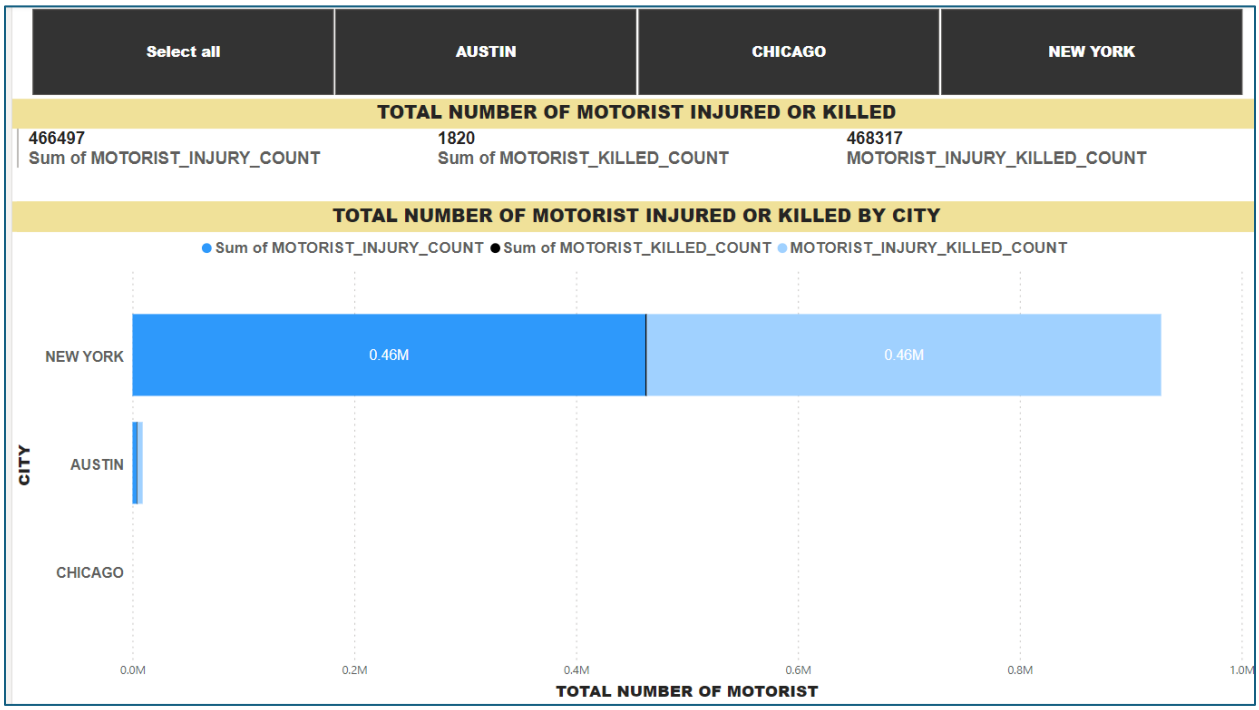
```sql
--Q6 How many motorists are injured or killed in accidents?

SELECT SUM(MOTORIST_INJURY_COUNT) AS TOTAL_MOTORIST_INJURED,
       SUM(MOTORIST_KILLED_COUNT) AS TOTAL_MOTORIST_KILLED,
       sum(motorist_killed_count + motorist_injury_count) as NumMotorist_Killed_Injured
FROM FACT_ACCIDENTS where motorist_killed_count >0 or motorist_injury_count> 0

SELECT SUM(MOTORIST_INJURY_COUNT) AS TOTAL_MOTORIST_INJURED,
       SUM(MOTORIST_KILLED_COUNT) AS TOTAL_MOTORIST_KILLED,
       S.SOURCE
FROM FACT_ACCIDENTS F
JOIN DIM_Source S ON F.SK_Source=S.SK_Source
GROUP BY S.SOURCE
ORDER BY S.Source ASC
```

110 %

Results | Messages

	TOTAL_MOTORIST_INJURED	TOTAL_MOTORIST_KILLED	NumMotorist_Killed_Injured
1	466497	1820	468317

	TOTAL_MOTORIST_INJURED	TOTAL_MOTORIST_KILLED	SOURCE
1	4047	543	AUSTIN
2	0	0	CHICAGO
3	462450	1277	NEW YORK



7.

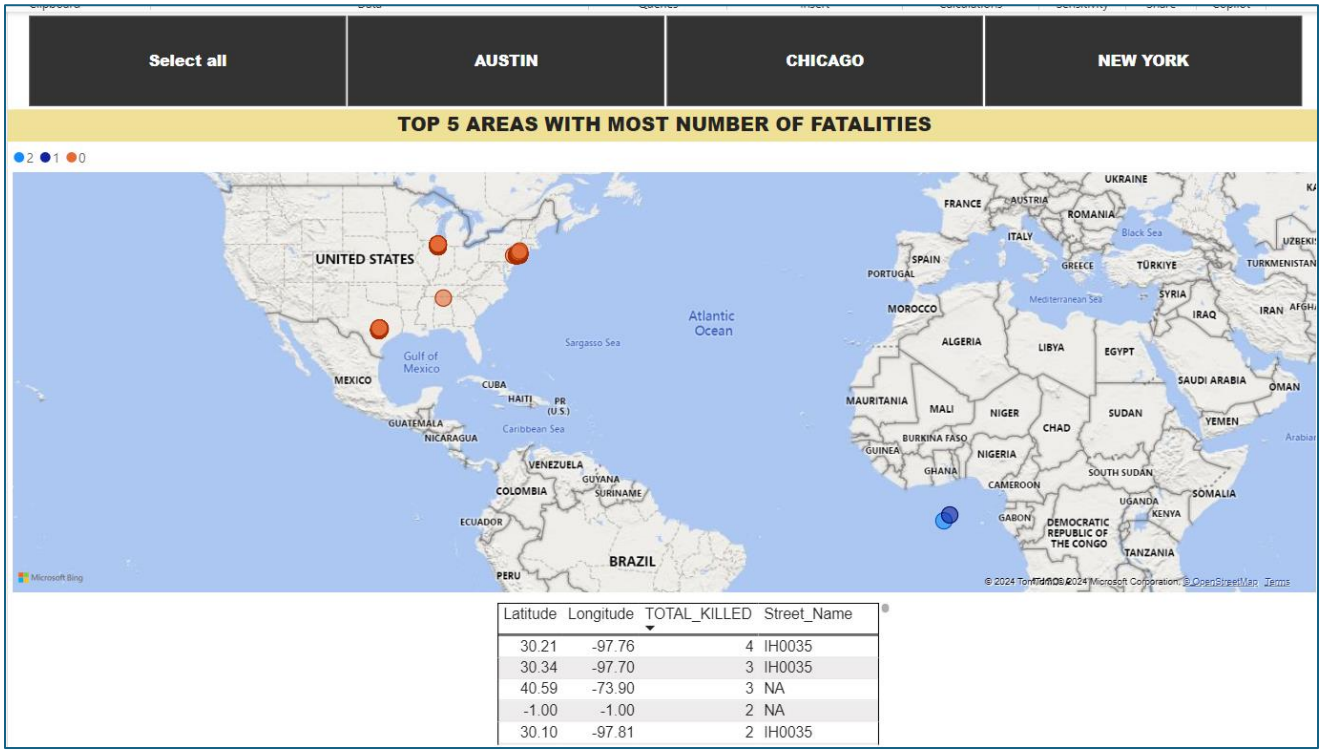
```sql
--Q7 Which top 5 areas in 3 cities have the most fatal number of accidents

WITH RankedAccidents AS (
  SELECT
    a.sk_location,
    s.SOURCE,
    l.street_name,
    l.latitude,
    l.longitude,
    SUM(a.total_Killed) AS totalKilled,
    COUNT(*) AS maxAccidents,
    DENSE_RANK() OVER (
      PARTITION BY s.SOURCE
      ORDER BY SUM(a.total_Killed) DESC, COUNT(*) DESC
    ) AS SourceRank
  FROM FACT_ACCIDENTS a
  JOIN DIM_Location l ON a.sk_location = l.sk_location
  JOIN DIM_Source s ON a.sk_source = s.sk_source
  WHERE l.latitude != -1 AND l.longitude != -1
  GROUP BY a.sk_location, s.SOURCE, l.street_name, l.latitude, l.longitude
)
SELECT
  sk_location
```

110 %

Results | Messages

	sk_location	SOURCE	street_name	latitude	longitude	maxAccidents	totalKilled	SourceRank
1	91862	AUSTIN	IH0035	30.20716	-97.75766	1	4	1
2	17062	AUSTIN	SL0001	30.38976	-97.73376	1	4	1
3	39169	AUSTIN	US0183	30.38449	-97.74249	1	3	2
4	74406	AUSTIN	SH0071	30.21491	-97.73246	1	3	2
5	15142	AUSTIN	IH0035	30.34393	-97.697	1	3	2
6	1379	AUSTIN	US0290	30.3405	-97.5911	29	2	3
7	16572	AUSTIN	IH0035	30.31618	-97.70763	17	2	4
8	21878	AUSTIN	E STASSNEY LN	30.19163	-97.73063	3	2	5
9	128133	CHICAGO	COTTAGE GROVE AVE	41.69274	-87.61002	46	4	1
10	106613	CHICAGO	59TH ST	41.78648	-87.67667	24	4	2
11	199633	CHICAGO	MADISON ST	41.88099	-87.69331	1	4	3
12	97370	CHICAGO	OHARE ST	41.9762	-87.90531	1304	3	4
13	99260	CHICAGO	LAKE SHORE DR NB	41.90096	-87.61993	331	3	5
14	563613	NEW YORK	WEST STREET	40.72905	-74.01073	45	8	1
15	400340	NEW YORK	ROCKAWAY BOULEVARD	40.65616	-73.76736	288	7	2
16	438531	NEW YORK	94 AVENUE	40.6972	-73.81333	44	5	3
17	603182	NEW YORK	NORTHERN BOULEVA...	40.76298	-73.83198	85	4	4
18	509881	NEW YORK	HENRY HUDSON PARK...	40.85061	-73.94437	11	4	5

Query executed successfully. VAISHVEER (16.0 RTM)



Latitude	Longitude	TOTAL_KILLED	Street_Name
30.21	-97.76	4	IH0035
30.34	-97.70	3	IH0035
40.59	-73.90	3	NA
-1.00	-1.00	2	NA
30.10	-97.81	2	IH0035

8.

```sql
--Q8 Time based analysis of accidents: Time of the day, day of the week, weekdays or weekends.
SELECT count(*) as NumAccidents,t.time_period
from FACT_ACCIDENTS a
join DIM_TIME t on a.SK_Time=t.SK_Time
GROUP BY t.time_period;

SELECT count(*) as NumAccidents,d.Day_Of_The_Week
from FACT_ACCIDENTS a join DIM_DATE d on a.SK_Date=d.SK_Date
GROUP BY d.Day_Of_The_Week
ORDER BY count(*) desc;

SELECT count(*) as NumAccidents,d.WeekDays_weekends
from FACT_ACCIDENTS a join DIM_DATE d on a.SK_Date=d.SK_Date
GROUP BY d.WeekDays_weekends
ORDER BY count(*) desc;

SELECT count(*) as NumAccidents,T.Hour_24
from FACT_ACCIDENTS a join DIM_TIME T on a.SK_Time=T.SK_Time
GROUP BY T.Hour_24
ORDER BY count(*) desc;
```

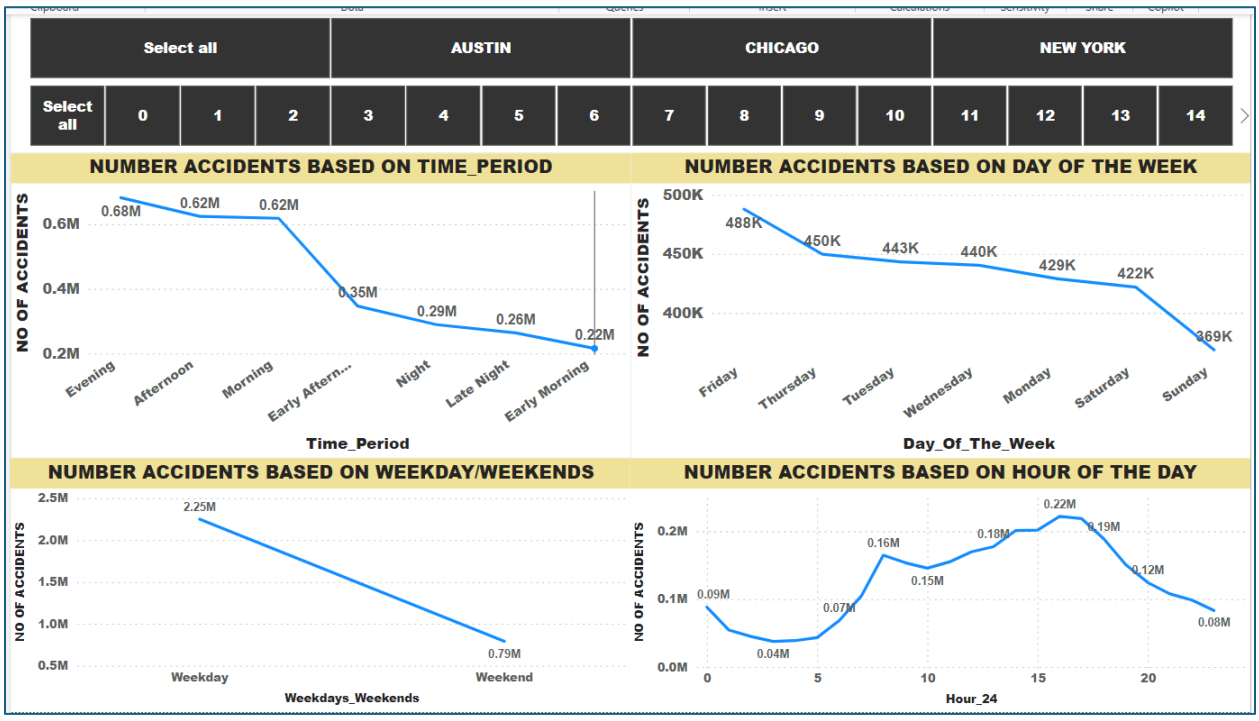
110 %

Results | Messages

	NumAccidents	time_period
1	624420	Afternoon
2	346778	Early Afternoon
3	215812	Early Morning
4	682211	Evening
5	263852	Late Night
6	618354	Morning
7	289473	Night

	NumAccidents	Day_Of_The_Week
1	487835	Friday
2	449692	Thursday
3	443249	Tuesday
4	440445	Wednesday
5	429002	Monday
6	421939	Saturday
7	368738	Sunday

	NumAccidents	WeekDays_weekends
1	2250223	Weekday
2	790677	Weekend



9.

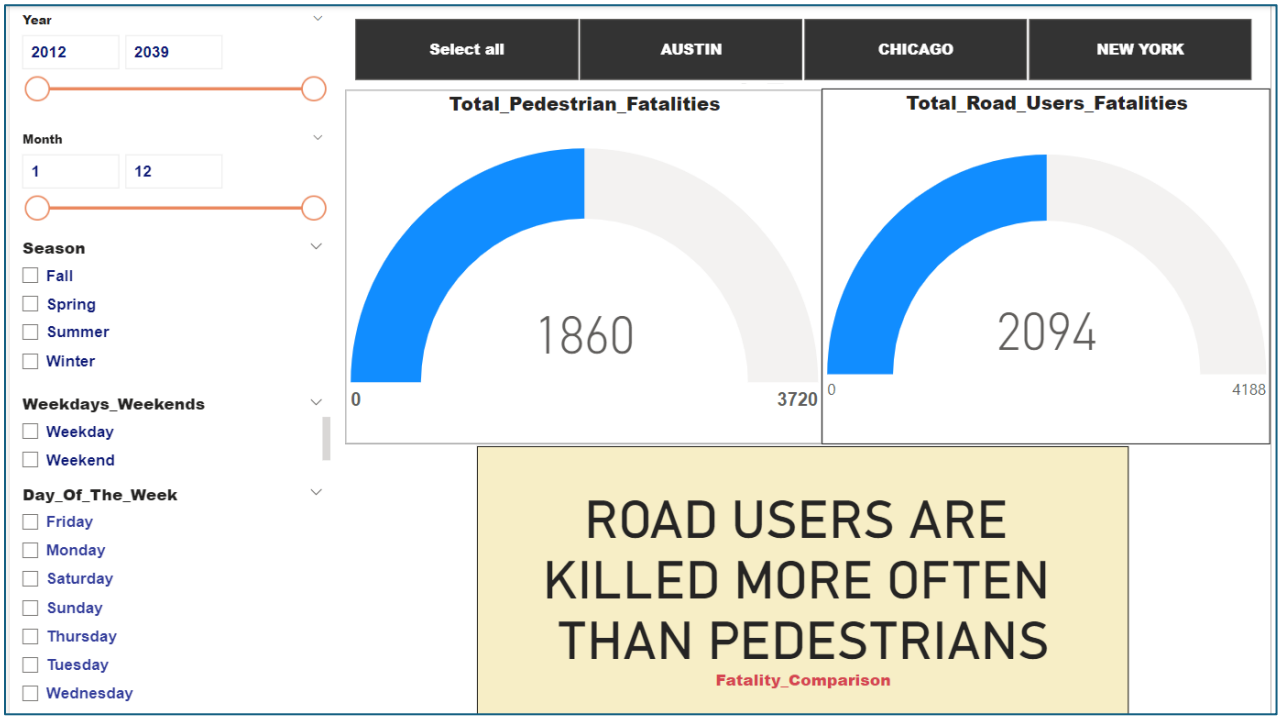
```sql
-- Q9 Fatality analysis: Are pedestrians killed more often than road users?

SELECT

(SELECT SUM(PEDESTRIAN_KILLED_COUNT) FROM FACT_ACCIDENTS) AS TotalPedestrianFatalities,

(SELECT SUM(road_users_killed) FROM FACT_ACCIDENTS) AS TotalMotoristFatalities,

CASE

WHEN (SELECT SUM(PEDESTRIAN_KILLED_COUNT) FROM FACT_ACCIDENTS) >

(SELECT SUM(road_users_killed) FROM FACT_ACCIDENTS)

THEN 'Pedestrians killed more often'

WHEN (SELECT SUM(PEDESTRIAN_KILLED_COUNT) FROM FACT_ACCIDENTS) <

(SELECT SUM(road_users_killed) FROM FACT_ACCIDENTS)

THEN 'Road Users killed more often'

ELSE 'Pedestrians and Road Users killed equally often'

END AS FatalityComparison
```

110 %

Results | Messages

	TotalPedestrianFatalities	TotalMotoristFatalities	FatalityComparison
1	1860	2094	Road Users killed more often

Query executed successfully. VAISHVEER (16.0 RTM)



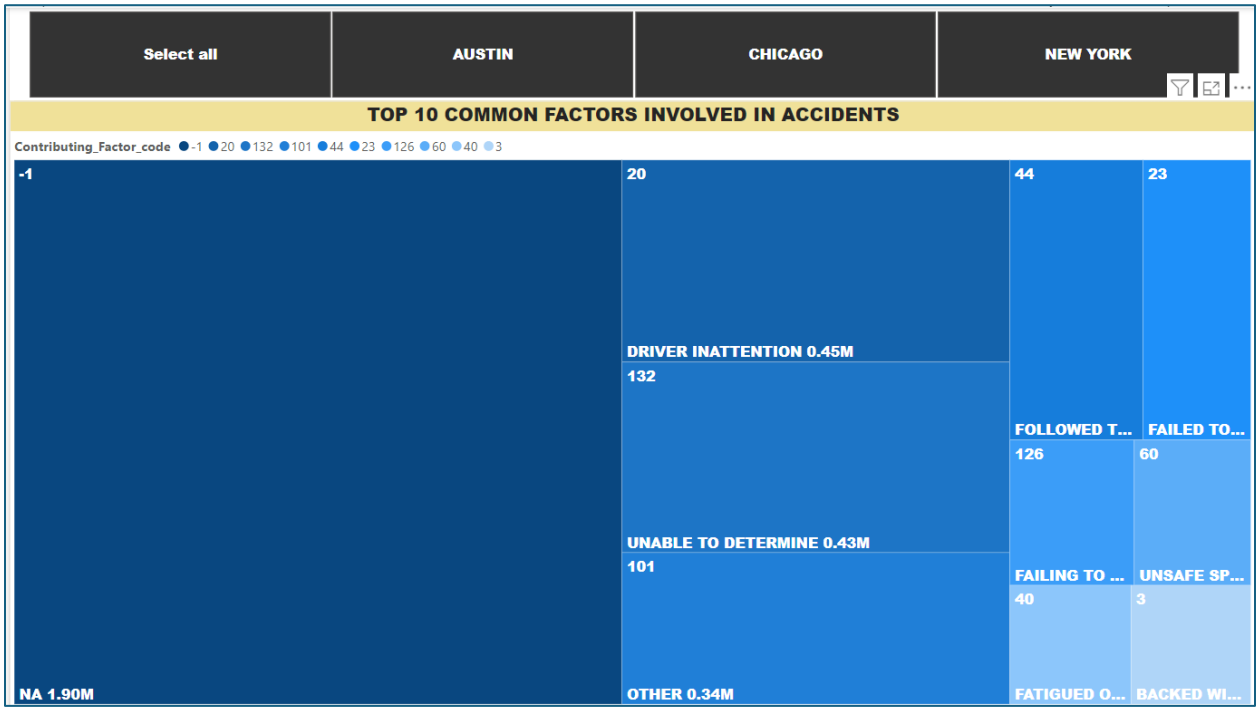
10.

```sql
--Q10 most common factors involved in accidents

SELECT TOP 10 Contributing_Factor_code,Contributing_Factor_DESCRIPTION,
COUNT(fa.SK_FACT_ACCIDENTS) AS COUNT_OF_ACCIDENTS
FROM Dim_Contribution dc
join FACT_CONTRIBUTION fc on dc.SK_Contribution=fc.SK_Contribution
join FACT_ACCIDENTS fa on fa.SK_FACT_ACCIDENTS=fc.SK_FACT_ACCIDENTS
GROUP BY Contributing_Factor_code,Contributing_Factor_DESCRIPTION
ORDER BY COUNT_OF_ACCIDENTS DESC
```

110 %

Results | Messages

	Contributing_Factor_code	Contributing_Factor_DESCRIPTION	COUNT_OF_ACCIDENTS
1	-1	NA	1904876
2	20	DRIVER INATTENTION	451447
3	132	UNABLE TO DETERMINE	425938
4	101	OTHER	338755
5	44	FOLLOWED TOO CLOSELY	214866
6	23	FAILED TO DRIVE IN SINGLE LANE	175177
7	126	FAILING TO YIELD RIGHT-OF-WAY	104173
8	60	UNSAFE SPEED	98381
9	40	FATIGUED OR ASLEEP	83129
10	3	BACKED WITHOUT SAFETY	82558



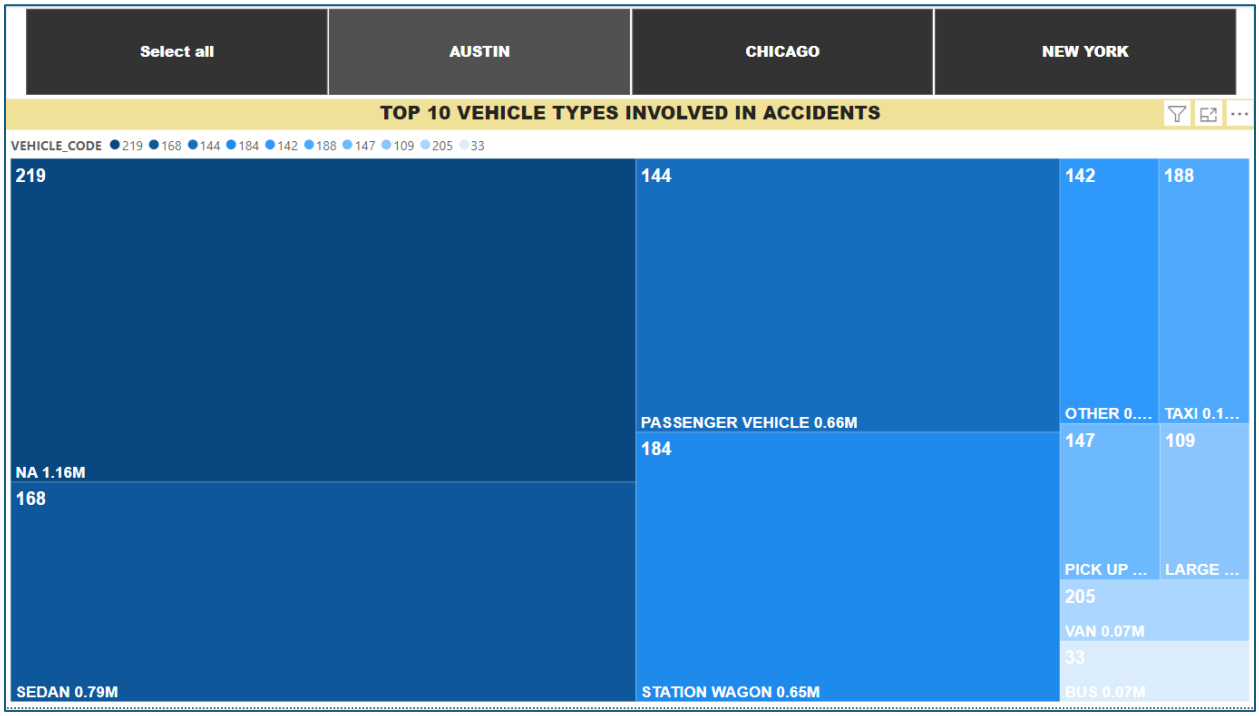
11.

```
--Q11 Store vehicle type/vehicles involved in accidents at its least graunality

SELECT TOP 10 count(*) as NumAccidents,v.VEHICLE_CODE,v.Vehicle_Type
from FACT_VEHICLE f join DIM_Vehicle_Type v on f.SK_VEHICLE_TYPE=v.SK_VEHICLE_TYPE
GROUP BY v.VEHICLE_CODE,v.Vehicle_Type
order by count(*) desc
```

110 %

Results | Messages

	NumAccidents	VEHICLE_CODE	Vehicle_Type
1	1156597	219	NA
2	785425	168	SEDAN
3	663095	144	PASSENGER VEHICLE
4	653145	184	STATION WAGON
5	149820	142	OTHER
6	136984	188	TAXI
7	89400	147	PICK UP TRUCK
8	79640	109	LARGE COMMERCIAL VEHICLE
9	65947	205	VAN
10	65883	33	BUS



12.

```
--12.Using Austin and NYC datasets, Create a visualization to show number of incidents that involved more than 2 vehicles. Show this data as a comp

SELECT count(*) as NumAccidents,s.SOURCE
from FACT_ACCIDENTS a join FACT_VEHICLE v on v.sk_fact_accidents=a.sk_fact_accidents
join DIM_Source s on a.SK_Source=s.sk_source
where v.units_involved>2
group by s.source;
```

110 %

Results | Messages

	NumAccidents	SOURCE
1	48769	AUSTIN
2	375871	NEW YORK