# SQL Execution

## 1) How many accidents happened in NYC , Austin , NYC

**By City**

select count(crash_id) from fact_crash  where source_system = 'New_York'

```
1    select count(crash_id) from fact_crash  where source_system = 'New_York'
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| count(crash_id) |
|---|
| 2075427 |

select count(crash_id) from fact_crash  where source_system = 'Chicago'

```
1    select count(crash_id) from fact_crash  where source_system = 'Chicago'
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| count(crash_id) |
|---|
| 817723 |

select count(crash_id) from fact_crash  where source_system = 'Austin'

```
1   select count(crash_id) from fact_crash  where source_system = 'Austin'
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| count(crash_id) |
|---|
| 147750 |

**OverAll**

Select count(crash_id) from fact_crash

```
1   select count(crash_id) from fact_crash
2
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| count(crash_id) |
|---|
| 3040900 |

## 2) Which areas in three cities has greatest number of accidents?

```
select l.latitude, l.longitude ,count(*)
from fact_crash fc
left join location_dim l on fc.loc_id = l.loc_id
group by l.latitude, l.longitude having l.latitude != 0
order by count(*) desc LIMIT 3
```

```
6
7   select l.latitude, l.longitude ,count(*)
8   from fact_crash fc
9   left join location_dim l on fc.loc_id = l.loc_id
10  group by l.latitude, l.longitude having l.latitude != 0
11  order by count(*) desc LIMIT 3
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

| latitude | longitude | count(*) |
|---|---|---|
| 30.000000000 | -98.000000000 | 145448 |
| 41.976201139 | -87.905309125 | 1305 |
| 41.900958919 | -87.619928174 | 759 |

3) How many accidents resulted in just injuries, at two levels 1) Overall 2) By City?

**OverAll**

select Just_Injured_Fl, count(*) from fact_crash group by Just_Injured_Fl

```
13
14      select Just_Injured_Fl, count(*) from fact_crash group by Just_Injured_Fl
15
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Just_Injured_Fl | count(*) |
|---|---|
| 1 | 650027 |
| 0 | 2390873 |

**By City**

select source_system, Just_Injured_Fl , count(*) from fact_crash group by source_system,Just_Injured_Fl

```
15
16      select source_system, Just_Injured_Fl , count(*) from fact_crash group by source_system,Just_Injured_Fl
17
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content:

| source_system | Just_Injured_Fl | count(*) |
|---|---|---|
| New_York | 1 | 473644 |
| New_York | 0 | 1601783 |
| Austin | 0 | 82980 |
| Austin | 1 | 64770 |
| Chicago | 1 | 111613 |
| Chicago | 0 | 706110 |

4) How often are the pedestrians involved in accidents at two levels 1) Overall 2) By City?

**OverAll**

```
19      select PEDESTRIAN_INVOLVED, count(*) from vehicle_int.fact_crash group by  PEDESTRIAN_INVOLVED
20
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| PEDESTRIAN_INVOLVED | count(*) |
|---|---|
| 0 | 2108316 |
| 1 | 114861 |
| NULL | 817723 |

**By City**

```
15
16  select source_system, Just_Injured_Fl , count(*) from fact_crash group by source_system,Just_Injured_Fl
17
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| source_system | PEDESTRIAN_INVOLVED | count(*) |
|---|---|---|
| New_York | 0 | 1961501 |
| New_York | 1 | 113926 |
| Austin | 0 | 146815 |
| Austin | 1 | 935 |
| Chicago | NULL | 817723 |

## 5) When do most accidents happen?

select t.Time_Of_Day,count(CRASH_ID) from fact_crash fc INNER JOIN time_dim t on fc.time_id = t.time_id group by t.Time_Of_Day

```
24
25  select t.Time_Of_Day,count(CRASH_ID) from fact_crash fc INNER JOIN time_dim t on fc.time_id = t.time_id group by t.Time_Of
26
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Time_Of_Day | count(CRASH_ID) |
|---|---|
| Night | 1060310 |
| Morning | 790917 |
| Afternoon | 1189673 |

Result Grid | Form Editor

## 6) How many motorist are injured or killed in accidents

- ➔ select count(*) from fact_crash order by Count_of_Motorist_involved desc
- ➔ select source_system,count(crash_id) from fact_crash group by crash_id, source_system order by count(Count_of_Motorist_involved) desc

```
30      select count(*) from fact_crash order by Count_of_Motorist_involved desc
31
32
33
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| count(*) |
|---|
| 3040900 |

## 7) Which top 5 areas in three cities have most fatal number of accidents?

```
select ld.latitude, ld.longitude ,count(fc.Total_death) from fact_crash fc INNER JOIN location_dim ld ON ld.loc_id = ld.loc_id group by ld.latitude, ld.longitude order by count(fc.Total_death) desc
```

```
72  21:01:09  select ld.latitude, ld.longitude ,count(fc.Total_death) from fact_crash fc INNER JOIN location_dim ld ON l...  Error Code: 2013. Lost connection to MySQL server during query
```

## 8) Time based analysis of accident?

```
select d.Season, count(CRASH_ID) from fact_crash fc INNER JOIN date_dim d on fc.date_id = d.date_id group by d.Season
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Season | count(CRASH_ID) |
|---|---|
| Summer | 550238 |
| Winter | 483262 |
| Spring | 497081 |
| Autumn | 547121 |

select count(CRASH_ID) from fact_crash fc INNER JOIN date_dim d on fc.date_id = d.date_id group by d.Weekday_Or_Weekend

```
51  select  d.Weekday_Or_Weekend, count(CRASH_ID) from fact_crash fc INNER JOIN date_dim d on fc.date_id = d.date_id group by
52
53
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Weekday_Or_Weekend | count(CRASH_ID) |
|---|---|
| Weekend | 529733 |
| Weekday | 1547969 |

select  d.Day_of_Week,count(CRASH_ID) from fact_crash fc INNER JOIN date_dim d on fc.date_id = d.date_id group by d.Day_of_Week

```
53  select  d.Day_of_Week,count(CRASH_ID) from fact_crash fc INNER JOIN date_dim d on fc.date_id = d.date_id group by d.Day_of
54
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Day_of_Week | count(CRASH_ID) |
|---|---|
| Saturday | 280061 |
| Wednesday | 303751 |
| Tuesday | 305805 |
| Monday | 296771 |
| Sunday | 249672 |
| Thursday | 310550 |
| Friday | 331092 |

Result Grid | Form Editor | Field Types

## 9) Fatality Analysis

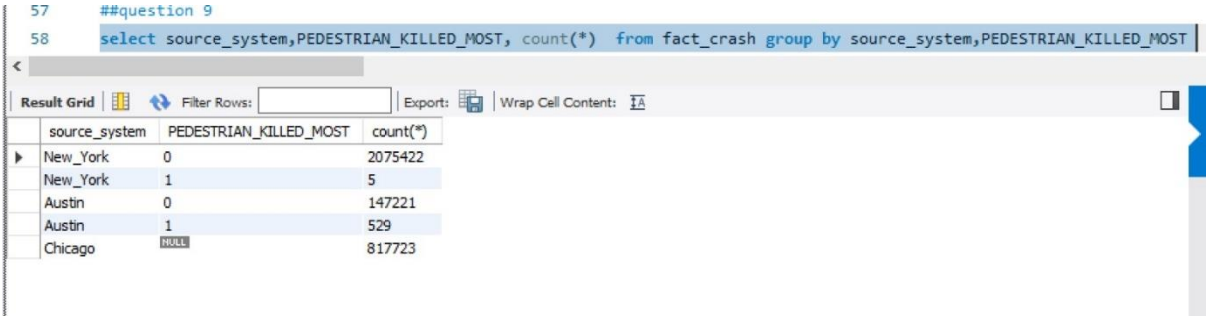
select PEDESTRIAN_KILLED_MOST, count(*)  from fact_crash group by PEDESTRIAN_KILLED_MOST

```
58  select PEDESTRIAN_KILLED_MOST, count(*)  from fact_crash group by PEDESTRIAN_KILLED_MOST
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

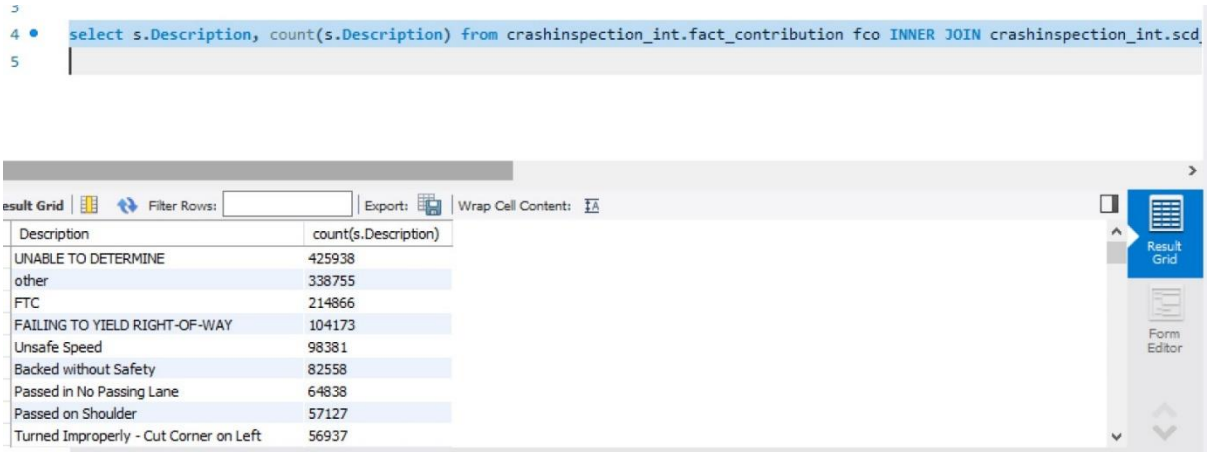
| PEDESTRIAN_KILLED_MOST | count(*) |
|---|---|
| 0 | 2222643 |
| 1 | 534 |
| NULL | 817723 |

select source_system,PEDESTRIAN_KILLED_MOST, count(*)  from fact_crash group by source_system,PEDESTRIAN_KILLED_MOST




## 10)What are most common factors involved in accidents?

select s.Description, count(s.Description) from vehicle_curation.fact_contribution fco INNER JOIN vehicle_curation.scd_contribution_dim s on fco.SK_Contri = s.SK_Contri group by s.Description order by count(s.Description) desc




## 11) 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 v.vehicle_type, count(v.vehicle_type) from fact_vehicle fv INNER JOIN vehicle_dim v on fv.vehicle_id = v.vehicle_id group by v.vehicle_type