

SQL PROJECT ON ROAD ACCIDENT DATA 2020

Presented by: Kirti Sharma

Presented by: Kirti sharma

Overview

1. Regulatory affairs of road accidentt data 2020

2. Project Overview

5. Summary

3. TechologyUsed

4. Quries



REGULATORY AFFAIRS OF ROAD ACCIDENT DATA 2020

Road safety remains a major concern in rapidly urbanizing regions, particularly in cities with high traffic density. A systematic analysis of accident data provides valuable insights into prevailing trends, key contributing factors, and the severity of outcomes. This project aims to examine the distribution of road-accident outcomes—specifically minor injuries—across multiple cities. By evaluating this data, we can identify cities with higher incidences of minor injuries and better understand the influence of various causal factors. These insights support authorities in designing targeted interventions to enhance overall road safety.

PROJECT OVERVIEW

This project focuses on analyzing the distribution of road-accident incidents across multiple cities, specifically considering only those with more than 15 reported cases. By filtering and ranking cities according to incident frequency, the study identifies locations with a higher prevalence of minor injuries and evaluates the key factors contributing to these occurrences. Understanding these patterns is crucial for recognizing high-risk urban zones and addressing road-safety challenges more effectively.

Using an SQL-driven analytical approach, the project extracts meaningful insights from structured accident records, categorizing incidents by both city and cause. This method enables a clear identification of cities with frequent minor-injury cases and highlights the most common cause categories involved. The insights generated from this analysis can support policymakers, traffic enforcement authorities, and urban planners in designing targeted safety interventions aimed at reducing accident rates and improving overall traffic management.

TECHNOLOGY USED

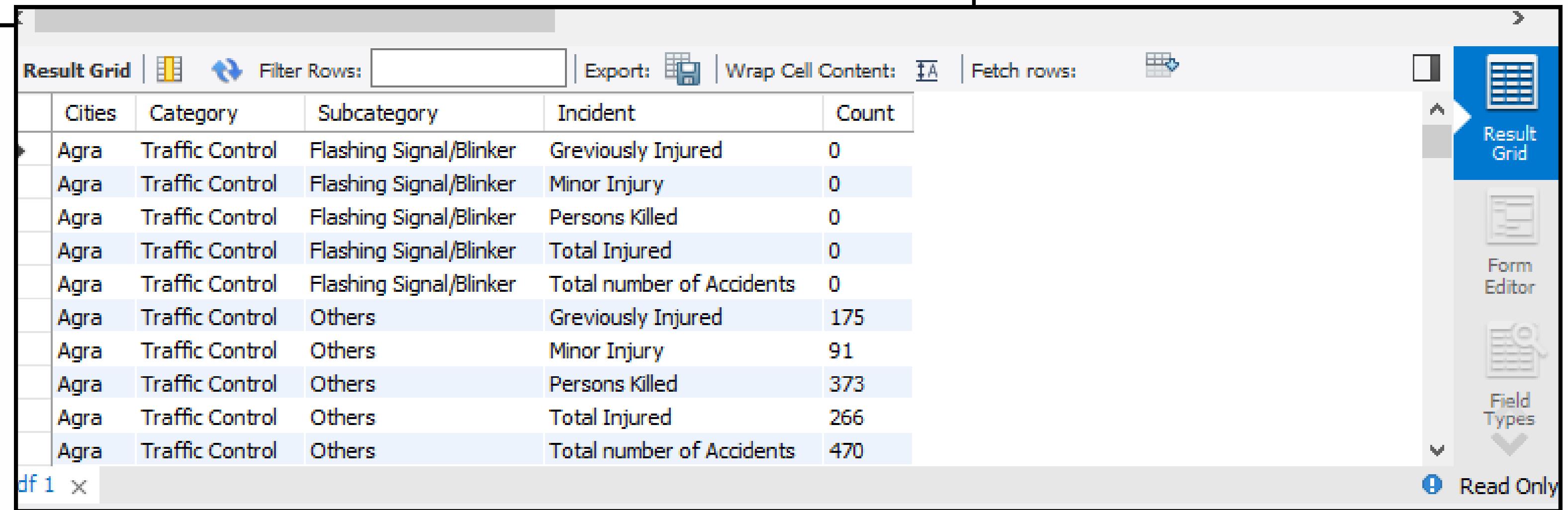
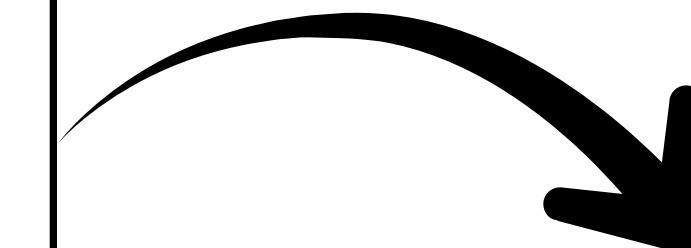
- This project leverages SQL as the primary tool for querying and analyzing road accident data. Through the use of SQL operations such as COUNT(), GROUP BY, ORDER BY, and conditional filters, the analysis extracts meaningful insights by identifying, comparing, and ranking cities based on the frequency of incident occurrences. These queries enable us to pinpoint cities with a high number of minor injury cases and explore potential factors contributing to these trends.
- In addition to SQL, Excel is utilized during the initial data-exploration phase to review the dataset's structure, verify column consistency, and understand the overall data layout. This preliminary assessment ensures accuracy and clarity before performing in-depth SQL analysis.
- Finally, PowerPoint is used to compile and present the analytical findings in a clear, visual, and organized format. Charts, tables, and summarized insights help effectively communicate key trends, high-risk areas, and recommendations.
- By integrating SQL for analysis, Excel for data understanding, and PowerPoint for impactful presentation, the project follows a systematic and professional approach to extracting valuable insights and presenting them with clarity and precision.

Queries

Presented by:- Kirti Sharma

1) Import & View Table

- 1• **create database Report;**
- 2• **use Report;**
- 3• **select * from df;**
- 4



	Cities	Category	Subcategory	Incident	Count
1	Agra	Traffic Control	Flashing Signal/Blinker	Previously Injured	0
2	Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0
3	Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0
4	Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0
5	Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0
6	Agra	Traffic Control	Others	Previously Injured	175
7	Agra	Traffic Control	Others	Minor Injury	91
8	Agra	Traffic Control	Others	Persons Killed	373
9	Agra	Traffic Control	Others	Total Injured	266
10	Agra	Traffic Control	Others	Total number of Accidents	470

2) Rename The Table

```
6 #Rename the table  
7 • rename table df to road_accident ;  
8 • select * from road_accident ;  
9
```

3) Check Data-Types

```
10  
11 #Check Data Types  
12 • describe road_accident;  
13
```

The screenshot shows the MySQL Workbench interface with the 'Result Grid' tab selected. The results grid displays the following data:

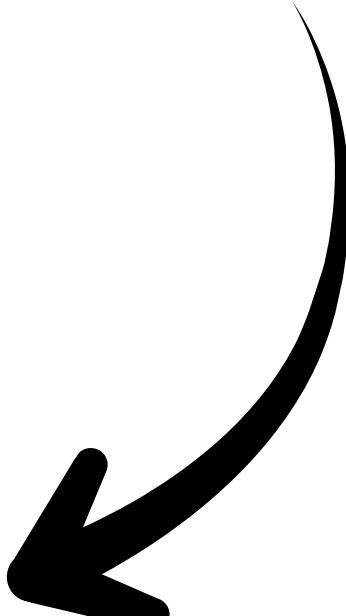
Field	Type	Null	Key	Default	Extra
Cities	text	YES		NULL	
Category	text	YES		NULL	
Subcategory	text	YES		NULL	
Incident	text	YES		NULL	
Count	int	YES		NULL	

Result 3 ×

Read Only

4) Check Null Values

```
9  
10    # Check Null Values in Columns  
11 • select * from road_accident  
12 where cities is null or category is null or subcategory or incident is null or TRIM(count) is null;  
13 -- OR  
14 • select * from road_accident  
15 where cities = '' or category = '' or subcategory or incident = '' or count= '' ;  
16
```



Result Grid				
	Cities	Category	Subcategory	Incident

5)Top Cities with the Highest Incident Count

```
16
17  #Top Cities With the Highest Incident Count
18 • select cities, count(*) as Highest_Incident_Count_cities
19   from road_accident
20   group by cities
21   order by Highest_Incident_Count_Cities desc;
22
23  #Agra Incident Count
24 • select cities, count(*) AS incident_count
25   from road_accident
26   where cities = "Agra"
27   group by cities;
```

The screenshot shows the MySQL Workbench interface with two result grids. The top grid, titled 'Result Grid', displays the top cities with the highest incident counts. The bottom grid, also titled 'Result Grid', shows the specific incident count for the city of Agra.

Top Grid (Result Grid):

cities	Highest_Incident_Count_cities
Agra	191
Ahmedabad	191
Allahabad(Prayagraj)	191
Amritsar	191
Asansol Durgapur	191
Aurangabad	191
Bengaluru	191
Bhopal	191
Chandigarh	191
Chennai	191
Coimbatore	191
Delhi	191
Dhanbad	191
Faridabad	191
Ghaziabad	191
Vizag	191
Hyderabad	191
Indore	191
Jabalpur	191

Bottom Grid (Result Grid):

cities	incident_count
Agra	191

6) Distribution of "Minor Injury" Incident Outcomes Across Cities (Greater Than 15)

```
28  
29  #Distributed of "Minor Injury" Incident Outcomes Across Cities (Greatest Than 15)  
30 • Select Cities, Incident, count(*) AS Incident_count  
31   from road_accident  
32   where Incident= "Minor Injury"  
33   group by Cities, Incident  
34   having Incident_count > 15  
35   order by Cities, Incident_count desc;  
36
```



Cities	Incident	Incident_count
Dhanbad	Minor Injury	40
Faridabad	Minor Injury	40
Ghaziabad	Minor Injury	40
Gwalior	Minor Injury	39
Hyderabad	Minor Injury	40
Indore	Minor Injury	40
Jabalpur	Minor Injury	40
Jaipur	Minor Injury	40
Jamshedpur	Minor Injury	40
Jodhpur	Minor Injury	40
Kannur	Minor Injury	40
Kanpur	Minor Injury	40
Khozikode	Minor Injury	40
Kochi	Minor Injury	40
Kolkata	Minor Injury	40
Kollam	Minor Injury	40
Kota	Minor Injury	40
Lucknow	Minor Injury	40
Ludhiana	Minor Injury	40

7) Top Cities for Each Specific Cause Category

```
37 #Top Cities for Each Specific Cause Category
38 • With RankCities as(
39   select Cities, category, count(*) as Cities_Each_Cause_Category,
40   row_number() over (partition by category order by count(*) desc
41   from road_accident
42   group by Cities , Category )
43   select Cities, Category, Cities_Each_Cause_Category
44   from Rankcities
45   where rnk = 1;
46
```

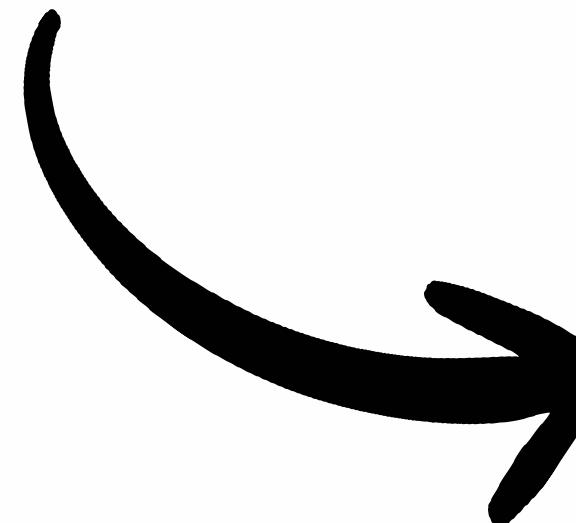


Result Grid | Filter Rows: [] | Export: | Wrap Cell Content: EA

	Cities	Category	Cities_Each_Cause_Category
▶	Agra	Impacting Vehicle/Object	36
	Khozikode	Junction	30
	Ahmedabad	Road Features	40
	Agra	Traffic Control	30
	Agra	Traffic Violation	30
	Agra	Weather	25

8) Most Common Outcome of Incident Across All Cities

```
47 #Most Common Outcome of Incident Across All Cities
48 • select Incident, count(*) as outcome_count
49 from road_accident
50 group by Incident
51 order by outcome_count desc
52 limit 1 ;
53 # if i get multiple output then i will use rank function to get only one outcome
```

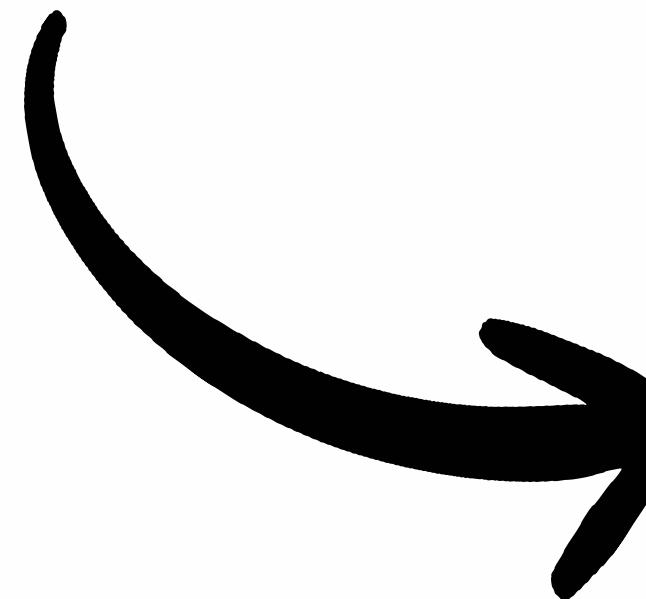


The screenshot shows a database query results grid. At the top, there are several buttons: 'Result Grid' (highlighted with a blue border), 'Filter Rows:', 'Export:', 'Wrap Cell Content:', and 'Fetch rows:'. Below the buttons is a table with two columns: 'Incident' and 'outcome_count'. The first row contains the header names, and the second row contains the data: 'Total number of Accidents' and '2000'.

	Incident	outcome_count
▶	Total number of Accidents	2000

9) City-Wise Distribution of Incident Causes

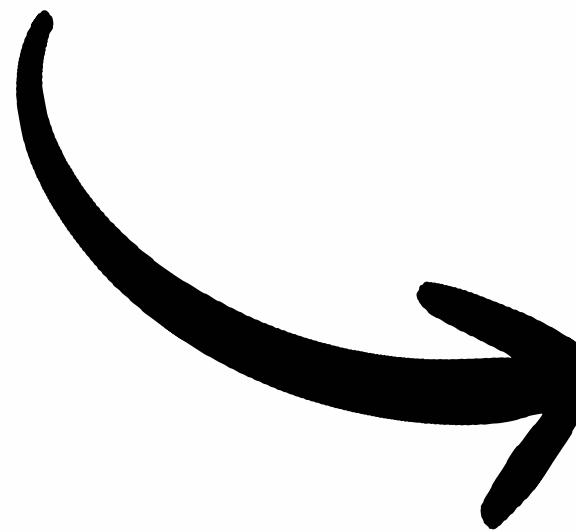
```
55 #City-wise Distributed of Incident Causes  
56 • select Cities, Category, count(*) as City_wise_Distribution_Incident  
57 from road_accident  
58 group by Cities, Category  
59 order by City_wise_Distribution_Incident desc;
```



	Cities	Category	City_wise_Distribution_Incident
▶	Agra	Road Features	40
	Ahmedabad	Road Features	40
	Allahabad(Prayagraj)	Road Features	40
	Amritsar	Road Features	40
	Asansol Durgapur	Road Features	40
	Aurangabad	Road Features	40
	Bengaluru	Road Features	40
	Bhopal	Road Features	40
	Chandigarh	Road Features	40
	Chennai	Road Features	40
	Coimbatore	Road Features	40
	Delhi	Road Features	40
	Dhanbad	Road Features	40

10) Total Incident Count by clause Category and Subcategory

```
61 #Total Incident Count by Cause Category and Subcategory
62 • select Category, Subcategory, count(*) as incident_count
63 from road_accident
64 group by Category, Subcategory
65 order by incident_count desc;
```

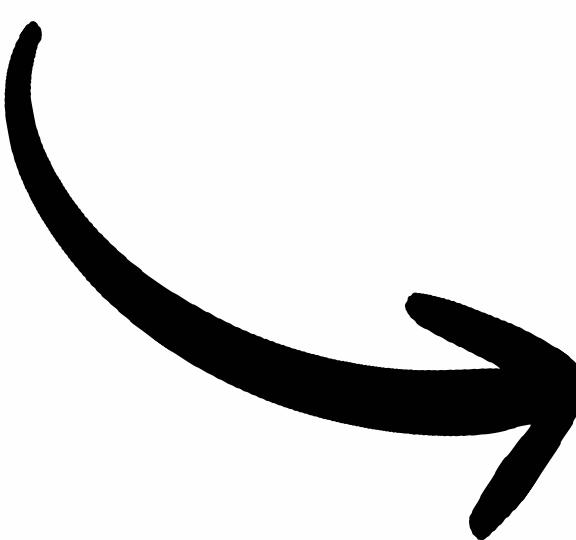


Result Grid | Filter Rows: Export: Wrap Cell Content:

	Category	Subcategory	incident_count
▶	Traffic Control	Flashing Signal/Blinker	250
	Traffic Control	Others	250
	Traffic Control	Police Controlled	250
	Traffic Control	Stop Sign	250
	Traffic Control	Traffic Light Signal	250
	Traffic Control	Uncontrolled	250
	Junction	Four arm Junction	250
	Junction	Others	250
	Junction	Round about Junction	250
	Junction	Staggered Junction	250

11) Percentage of Each Outcome Type in Total Incidents

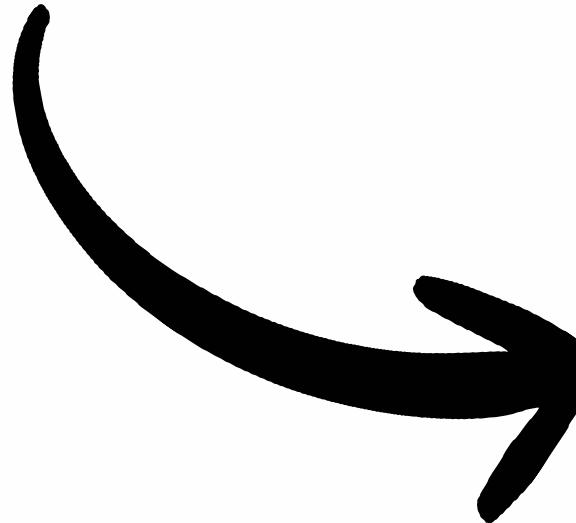
```
67  #Percentage of Each Outcome Type in Total Incidents
68 • select Incident, count(*) as Total_count,
69   (count(*) * 100.0 / sum(count(*)) over()) as output_percentage
70   from road_accident
71   group by Incident
72   order by output_percentage desc;
```



	Incident	Total_count	output_percentage
▶	Total number of Accidents	2000	20.94899
	Previously Injured	1999	20.93851
	Minor Injury	1999	20.93851
	Persons Killed	1999	20.93851
	Total Injured	1550	16.23547

12) City with the Highest Incidents for a Specific Cause Category

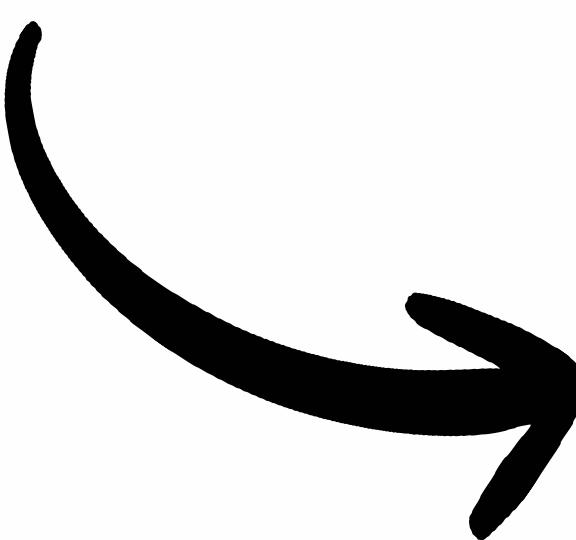
```
74  #City with the Highest Incidents for a specific Cause Category
75 • select Cities, Count(*) as incident_count,
76   rank() over (order by count(*) desc) as city_rank
77   from road_accident
78   group by Cities
79   order by city_rank;
```



	Cities	incident_count	city_rank
▶	Agra	191	1
	Ahmedabad	191	1
	Allahabad(Prayagraj)	191	1
	Amritsar	191	1
	Asansol Durgapur	191	1
	Aurangabad	191	1
	Bengaluru	191	1
	Bhopal	191	1
	Chandigarh	191	1
	Chennai	191	1
	Coimbatore	191	1

13) Identify Cities with the lowest Number of Incidents

```
81 # Identify Cities with the Lowest Number of Incidents  
82 • select Cities, count(*) as incident_count  
83 from road_accident  
84 group by Cities  
85 order by incident_count asc  
86 limit 5;
```



	Cities	incident_count
▶	Gwalior	188
▶	Agra	191
▶	Ahmedabad	191
▶	Allahabad(Prayagraj)	191
▶	Amritsar	191

14) Find the Most Frequent Cause-Outcome Pair Across Cities

```
89  #Find the Most Frequent Cause-Outcome Pair Across Cities
90 • select category, Incident, count(*) as frequency
91  from road_accident
92  group by category, Incident
93  order by Frequency desc
94  limit 1;
```



	category	Incident	frequency
▶	Impacting Vehicle/Object	Total number of Accidents	450

SUMMARY

This project conducted a detailed analysis of road accident data to evaluate the distribution of incidents across various cities. Using SQL-based querying, cities reporting more than 15 minor injury cases were identified and ranked according to incident frequency. The analysis highlighted high-risk urban areas and revealed key factors contributing to minor injuries. These insights can support policymakers and traffic authorities in developing targeted road safety strategies. All findings have been compiled and presented in a clear, structured format to facilitate effective, data-driven decision-making.



THANK YOU

Presented by: Kirti Sharma

