REGULATORY AFFAIRS OF ROAD ACCIDENT DATA 2020

Road safety is a critical concern in urban areas, especially in cities with high traffic density. Analyzing accident data helps in identifying trends, causes, and outcomes of incidents, allowing authorities to implement effective safety measures. This project focuses on understanding the distribution of road accident outcomes, particularly minor injuries, across various cities. By analyzing this data, we can gain insights into which cities experience higher incidents of minor injuries and assess the impact of different cause categories on road safety.

-- by Priya Rajurkar

PROJECT OVERVIEW

- This project aims to analyze the distribution of incidents across various cities, focusing only on those where the count exceeds **15 cases.** By filtering and ranking cities based on their incident frequency, we can identify areas with a higher occurrence of minor injuries and assess the underlying factors contributing to these accidents. Understanding these patterns is essential for recognizing high-risk urban zones and addressing road safety concerns effectively.

- Using **SQL-based data analysis**, the project extracts meaningful insights from road accident records, categorizing incidents by city and cause. This structured approach allows us to pinpoint cities with frequent minor injury cases and examine which cause categories are most prevalent. The findings can help policymakers, traffic authorities, and urban planners develop targeted safety measures to reduce road accidents and improve overall traffic management.

TECHNOLOGY USED

This project primarily utilizes **SQL** for querying and analyzing road accident data. SQL is used to extract meaningful insights by filtering, grouping, and ranking cities based on the number of incidents. Various SQL functions such as COUNT(), GROUP BY, ORDER BY, and conditional filtering help in identifying trends related to minor injuries across different cities. Through these queries, we can determine which cities report the highest number of minor injury incidents and analyze the contributing factors. Additionally, **Excel** is used for an initial overview of the dataset, specifically to check the structure and available columns. This helps in understanding the data before writing SQL queries for analysis. Lastly, **PowerPoint** is utilized to present findings in a visually structured format, making it easier to communicate insights, trends, and recommendations effectively. By leveraging these technologies, the project ensures a systematic approach to data analysis and presentation.

QUIRES

1] Import & View Table

```
1 •   create database rc;
2 •   use rc;
3 •   SELECT * FROM`regulatory affairs of road accident data 2020 india`;
4
```

Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
Agra	Traffic Control	Flashing Signal/Blinker	Greviously Injured	0
Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0
Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0
Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0
Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0
Agra	Traffic Control	Others	Greviously Injured	175
Agra	Traffic Control	Others	Minor Injury	91
Agra	Traffic Control	Others	Persons Killed	373
Agra	Traffic Control	Others	Total Injured	266
Agra	Traffic Control	Others	Total number of Accidents	470
Agra	Traffic Control	Police Controlled	Greviously Injured	0
Agra	Traffic Control	Police Controlled	Minor Injury	0
Agra	Traffic Control	Police Controlled	Persons Killed	0

2] Rename The Table

```sql
17    # Rename Table name
18 •  rename table `regulatory affairs of road accident data 2020 india` to `road_accident`;
19 •  select * from road_accident;
20
```

3] Check Data-Types

```sql
9     # Chcek Data Types
10 •  describe road_accident;
11
```

Field	Type	Null	Key	Default	Extra
Million Plus Cities	text	YES		NULL	
Cause category	text	YES		NULL	
Cause Subcategory	text	YES		NULL	
Outcome of Incident	text	YES		NULL	
Count	double	YES		NULL	

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

4] Check Null Values

```
12      # Chcek Null values in columns
13 •    select * from road_accident
14      where `Million Plus Cities` is null or `Cause category` is null or `Cause Subcategory` or
15      `Outcome of Incident` is null or TRIM(`Count`) is null ;
16      -- OR
17 •    select * from road_accident
18      where `Million Plus Cities` = ' ' or `Cause category` = ' ' or `Cause Subcategory` or
19      `Outcome of Incident`= ' ' or `Count` = ' ' ;
20
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:

Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count

5] Top Cities with the Highest Incident Count

```
22      -- 1] Top Cities with the Highest Incident Count
23  •   select `Million Plus Cities`, count(*) as Highest_Incident_Count_cities
24      from road_accident
25      group by `Million Plus Cities`
26      order by  Highest_Incident_Count_cities desc;
27
28      -- Agra Incident Count
29  •   Select `Million Plus Cities`, count(*) AS incident_count
30      from road_accident
31      where `Million Plus Cities` = "Agra"
32      group by `Million Plus Cities`;
33
```

Million Plus 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

Result 7

Million Plus Cities	incident_count
Agra	191

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

```
35      -- 2] Distribution of "Minor Injury" Incident Outcomes Across Cities (Greater Than 15)
36  •   Select `Million Plus Cities`,`Outcome of Incident`,count(*) AS incident_count
37      from road_accident
38      where `Outcome of Incident`= "Minor Injury"
39      group by `Million Plus Cities`,`Outcome of Incident`
40      having incident_count > 15
41      order by `Million Plus Cities`,incident_count desc;
42
```

Million Plus Cities	Outcome of Incident	incident_count
Agra	Minor Injury	40
Ahmedabad	Minor Injury	40
Allahabad(Prayagraj)	Minor Injury	40
A Amritsar	Minor Injury	40
Asansol Durgapur	Minor Injury	40
Aurangabad	Minor Injury	40
Bengaluru	Minor Injury	40
Bhopal	Minor Injury	40

Result 8 ✕

7] Top Cities for Each Specific Cause Category

```sql
43     -- 3]Top Cities for Each Specific Cause Category
44  •⊝  with Rankcities as(
45     select`Million Plus Cities`, `Cause category`,count(*) as Cities_Each_Cause_Category,
46     row_number() over (partition by `Cause category` order by count(*) desc ) as rnk
47     from road_accident
48     group by `Million Plus Cities` ,  `Cause category` )
49     select `Million Plus Cities`, `Cause category`,Cities_Each_Cause_Category
50     from Rankcities
51     where rnk = 1;
```

Million Plus Cities	Cause 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

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

8] Most Common Outcome of Incidents Across All Cities

```
53      -- 4] Most Common Outcome of Incidents Across All Cities
54 •    select `Outcome of Incident` , count(*) as outcome_count
55      from road_accident
56      group by `Outcome of Incident`
57      order by outcome_count desc
58      limit 1 ;
59      # if i get multiple output then i will use rank function to get ony one outcome
60
```

Outcome of Incident	outcome_count
Total number of Accidents	2000

9] City-wise Distribution of Incident Causes

```
61    -- 5] City-wise Distribution of Incident Causes
62  • select `Million Plus Cities`, `Cause category` , count(*) as City_wise_Distribution_Incident
63    from road_accident
64    group by `Million Plus Cities`, `Cause category`
65    order by City_wise_Distribution_Incident desc;
66
```

Million Plus Cities	Cause 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

Result 11 ✕

10] Total Incident Count by Cause Category and Subcategory

```
67    -- 6] Total Incident Count by Cause Category and Subcategory
68 •  select `Cause category`, `Cause Subcategory` , count(*) as incident_count
69    from road_accident
70    group by `Cause category`,  `Cause Subcategory`
71    order by incident_count desc;
72
```

Cause category	Cause 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

Result 12 ✕

11] Percentage of Each Outcome Type in Total Incidents

```
73    -- 7] Percentage of Each Outcome Type in Total Incidents
74  •  select `Outcome of Incident`, count(*) as Total_count,
75    (count(*) * 100.0 / sum(count(*)) over()) as output_percentage
76    from road_accident |
77    group by `Outcome of Incident`
78    order by output_percentage desc;
79
```

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

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

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

```sql
80    -- 8] City with the Highest Incidents for a Specific Cause Category
81 •  select `Million Plus Cities`, count(*) as incident_count,
82    rank () over (order by count(*) desc) as city_rank
83    from road_accident
84    group by `Million Plus Cities`
85    order by city_rank;
86
```

Million Plus 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

Result 14 ✕

13] Identify Cities with the Lowest Number of Incidents

```sql
87    -- 9] Identify Cities with the Lowest Number of Incidents
88 •  select `Million Plus Cities`, count(*) as incident_count
89    from road_accident
90    group by `Million Plus Cities`
91    order by incident_count asc
92    limit 5;
93
```

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

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

```sql
95     -- 10] Find the Most Frequent Cause-Outcome Pair Across Cities
96  •  select `Cause category`, `Outcome of Incident`, count(*) as frequency
97     From road_accident
98     group by `Cause category`, `Outcome of Incident`
99     order by frequency desc
100    limit 1;
```

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

SUMMARY

•This project analyzed road accident data to examine the distribution of incidents across cities.

•Using **SQL**, we identified cities with more than **15 minor injury cases** and ranked them based on incident count.

•The study highlighted high-risk cities and provided insights into the most common causes of minor injuries.

•Findings from the analysis can help policymakers and traffic authorities implement **targeted road safety measures**.

•The results have been presented in a structured format to support **data-driven decision-making**.

- *Thank you for your time and attention!*
- *Feel free to ask any questions or share your thoughts.*
- *Looking forward to discussions and insights!*