Qlik Analysis of Road Safety and Accident Patterns in India.

1. DEFINE PROBLEM / PROBLEM UNDERSTANDING

1.1. Specify the Business Problem:

India faces a critical road safety challenge, with over 150,000 fatalities recorded annually due to road accidents. These accidents result in substantial social and economic costs, including loss of lives, medical expenses, and infrastructure damage. Despite efforts to improve traffic management and road infrastructure, accident rates remain high, and there is a lack of data-driven insights to guide effective interventions.

The key business problem is the need for an advanced data analytics solution to identify accident trends, high-risk areas, and contributing factors (such as vehicle types, weather, and traffic control systems). This solution would provide government authorities, transportation agencies, and road safety organizations with actionable insights to reduce accident rates, improve road infrastructure, and enhance traffic enforcement strategies, ultimately saving lives and minimizing economic losses

The project addresses the critical issue of road safety in India by analyzing accident patterns and identifying contributing factors. With a high frequency of road accidents leading to significant loss of life and property, there is an urgent need for data-driven solutions to enhance road safety measures.

1.2. Business Requirements:

The project will utilize Qlik's data analytics platform to achieve the following objectives:

- **Hotspot Identification:** Pinpoint regions or roads with a high incidence of accidents by correlating traffic data, road conditions, and time of day.
- **Trend Analysis:** Analyze historical accident data to uncover patterns related to accident types, seasonal variations, and driver behavior.
- **Predictive Modeling:** Forecast potential accident scenarios using real-time data inputs like weather conditions and traffic flow patterns, enabling proactive safety interventions.

1.3. Literature Survey:

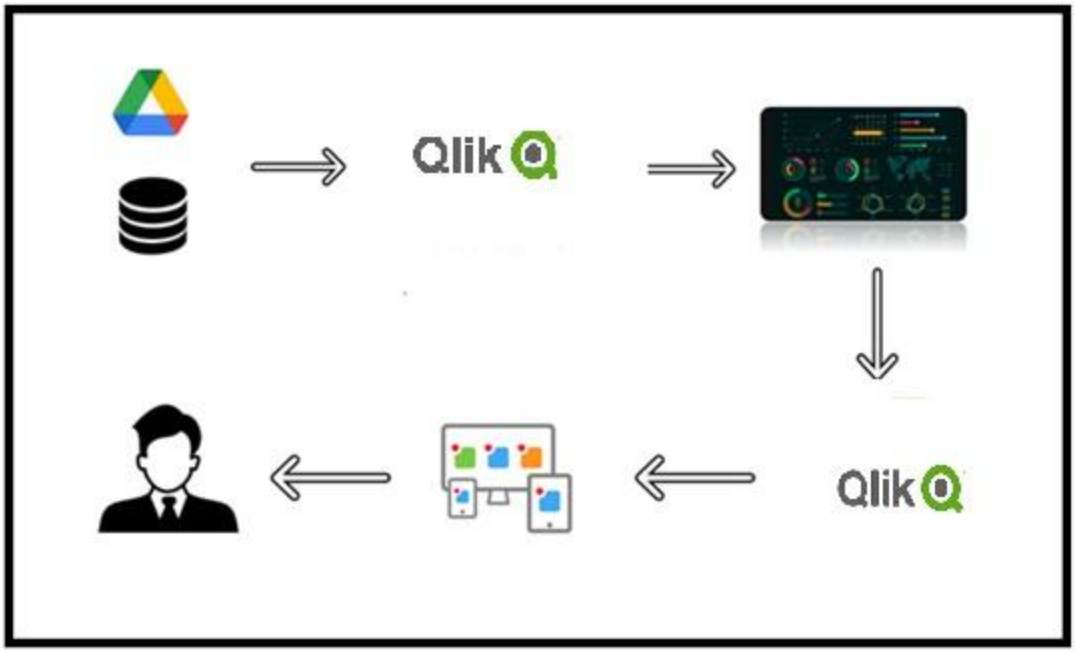
The literature review will encompass:

- **Previous Year Accident Data:** India's road safety has been a critical issue for years, with thousands of accidents occurring annually, resulting in significant loss of life and property. According to the National Crime Records Bureau (NCRB), there were over 150,000 fatalities in road accidents in 2021, making India one of the countries with the highest number of road accident deaths globally. The complexity of factors contributing to these accidents—ranging from vehicle types, weather conditions, and traffic controls to driver behavior—highlights the need for comprehensive analysis and intervention strategies.
- **Key Studies:** Several studies have utilized similar datasets to analyze road safety trends

and develop predictive models. For example, research has shown that over-speeding and distracted driving are leading causes of accidents, while improved road infrastructure and effective traffic management significantly reduce accident rates. Existing predictive models, such as those incorporating real-time traffic and weather data, have demonstrated the potential to forecast accident-prone areas and times, enabling preemptive actions to prevent accidents.

- **Existing Models:** A review of existing predictive models and analytical tools used in road safety, identifying their strengths and limitations to inform the development of the project's methodology.

2. TECHNICAL ARCHITECTURE



3. DATA COLLECTION

3.1. Collect the Dataset:

The To facilitate a detailed analysis of road safety, the project leverages the "ROAD ACCIDENTS IN INDIA" dataset from Kaggle. This dataset comprises nine Excel worksheets, each focusing on different aspects of road accidents in India during 2019. The data includes:

- **Pedestrians Involved and Killed:** This data is classified by age, gender, and type of impacting vehicles, providing insights into the most vulnerable groups and the types of vehicles most often involved in fatal accidents.

- **Traffic Control Types:** Analyzes accidents based on the presence and type of traffic controls, such as traffic lights, police controls, stop signs, and uncontrolled intersections, helping to assess the effectiveness of different traffic

management systems.

- **Weather Conditions:** Examines the impact of weather conditions like sunny, rainy, foggy, and others on the severity and frequency of accidents, revealing how environmental factors contribute to road safety.

- **Causes of Accidents:** Classifies accidents based on causes such as over-speeding, drunken driving, and the use of mobile phones while driving, offering a comprehensive view of the behavioral factors leading to accidents.

- **Severity by Vehicle Type:** Provides detailed information on the number of accidents, fatalities, and injuries by vehicle type, including pedestrians, bicycles, two-wheelers, auto rickshaws, and more

3.2. Connect Data with Qlik Sense

To analyze the dataset using Qlik Sense, follow these steps:

1. **Extract the Dataset:**
   - After downloading the dataset, extract the files to a specific location on your device.

2. **Create a New Qlik Sense App:**
   - Open Qlik Sense and create a new app named “Road Safety”.
   - Open the newly created app

3. **Add Data to Qlik Sense:**
   - Click on "Data Manager."
   - Click on "Add data" and select the dataset file from the location where it was extracted.

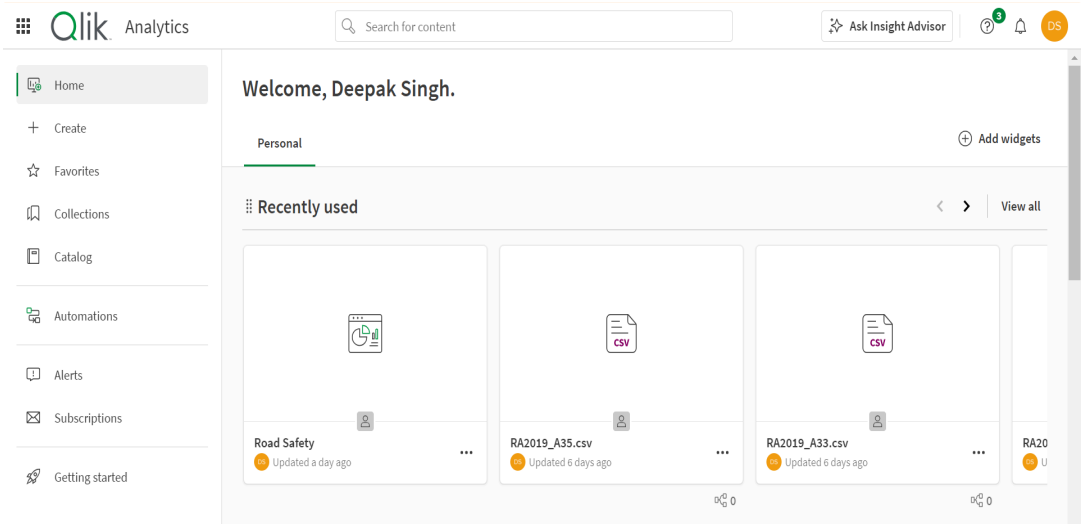
4. **Data Integration:**
   - Ensure that all relevant fields from the dataset are correctly mapped in Qlik Sense.
   - Check for any inconsistencies or missing values in the dataset and clean the data if necessary

5. **Data Mapping:**
   - Map fields such as State/UT to identify the geographical region or state in India where the accident data is recorded, ensuring it is correctly recognized by Qlik Sense for spatial analysis. Additionally, map age and gender categories, including fields like "Less than 18 years – Male," "Less than 18 years – Female," and "18-25 Years – Male," to categorize pedestrian involvement and fatalities accurately. For vehicle-related analysis, map fields such as "Bicycles," "Two Wheelers," "Auto Rickshaws," "Cars, Taxis, Vans and LMV," "Trucks/Lorries," and "Buses" to examine the types of vehicles involved in accidents, particularly

Qlik Analysis of Road Safety and Accident Patterns in India.

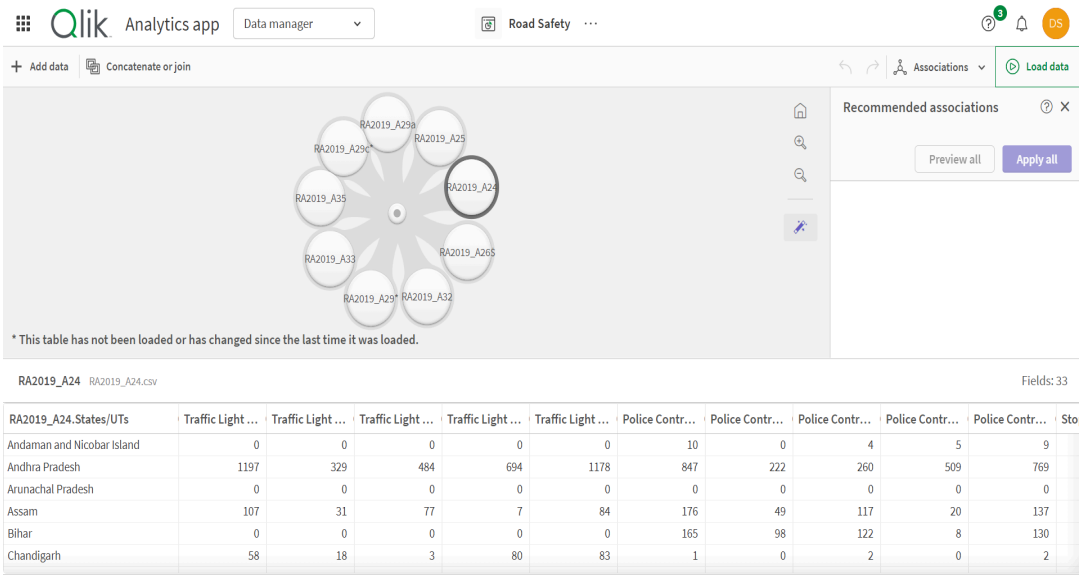
those resulting in fatalities. Traffic control types, including "Traffic Light Signal," "Police Controlled," "Stop Sign," "Flashing Signal/Blinker," and "Uncontrolled," should also be mapped to classify accidents based on traffic control measures, aiding in the evaluation of their effectiveness. Weather condition fields like "Sunny/Clear," "Rainy," "Foggy and Misty," and "Hail/Sleet" should be mapped to understand the influence of different weather conditions on accident frequency and severity. Furthermore, map causes of accidents fields, such as "Over-Speeding," "Drunken Driving/Consumption of Alcohol and Drugs," "Driving on Wrong Side," "Jumping Red Light," and "Use of Mobile Phone," to identify and analyze the root causes of road accidents. Severity by vehicle type should also be mapped, with fields like "Pedestrian - Number of Road Accidents," "Bicycles - Number of Road Accidents," and "Two Wheelers - Number of Road Accidents" to assess the severity of accidents based on the vehicle type involved. Lastly, map accident outcomes fields, including "Persons Killed," "Persons Injured - Grievously Injured," and "Persons Injured - Minor Injury," to facilitate the analysis of the human impact of road accidents.

By following these steps, you can successfully integrate and prepare your synthetic airline dataset in Qlik Sense for comprehensive analysis and visualization.



➤ Open Data Manager and we need to click on add data on left side top corner to add the data.

Qlik Analysis of Road Safety and Accident Patterns in India.



RA2019_A24.States/UTs	Traffic Light …	Traffic Light …	Traffic Light …	Traffic Light …	Traffic Light …	Police Contr…	Police Contr…	Police Contr…	Police Contr…	Police Contr…	Sto
Andaman and Nicobar Island	0	0	0	0	0	10	0	4	5	9	
Andhra Pradesh	1197	329	484	694	1178	847	222	260	509	769	
Arunachal Pradesh	0	0	0	0	0	0	0	0	0	0	
Assam	107	31	77	7	84	176	49	117	20	137	
Bihar	0	0	0	0	0	165	98	122	8	130	
Chandigarh	58	18	3	80	83	1	0	2	0	2	

4. DATA PREPARATION

4.1. Prepare the Data for Visualization:

➤ **Clean the Data**

i. **Remove Inconsistencies:**

- After downloading the dataset and converting it from CSV to Excel format, inspect the data for inconsistencies and anomalies.
- Rectify any discrepancies in the data entries to ensure uniformity.

ii. **Handle Missing Values:**

- Identify and address any missing values in the dataset.
- Fill in missing data points with appropriate values or remove records with substantial missing information.

➤ **Transform the Data**

- **Format for Analysis:**
  - Ensure the data is in a suitable format for analysis and visualization.
  - Check that dates, times, and numerical values are correctly formatted.
  - Ensure all fields are appropriately labeled.

4.2. Aggregate and Categorize Data

- **Remove Extra Columns:**
  - Identify and eliminate unnecessary columns such as 'others' and 'average' that are not relevant to the project's analysis.
  - During the data addition process in Qlik Sense, select only the columns required for analysis and discard extraneous ones.

Qlik Analysis of Road Safety and Accident Patterns in India.

- **Remove 'Total' Rows:**
  - Identify rows that contain "total" values, which are direct additions of each column.
  - Remove these rows from the dataset to ensure that aggregate data does not skew the analysis. In my case I have applied filter on it add excluded the total column from each visualization.

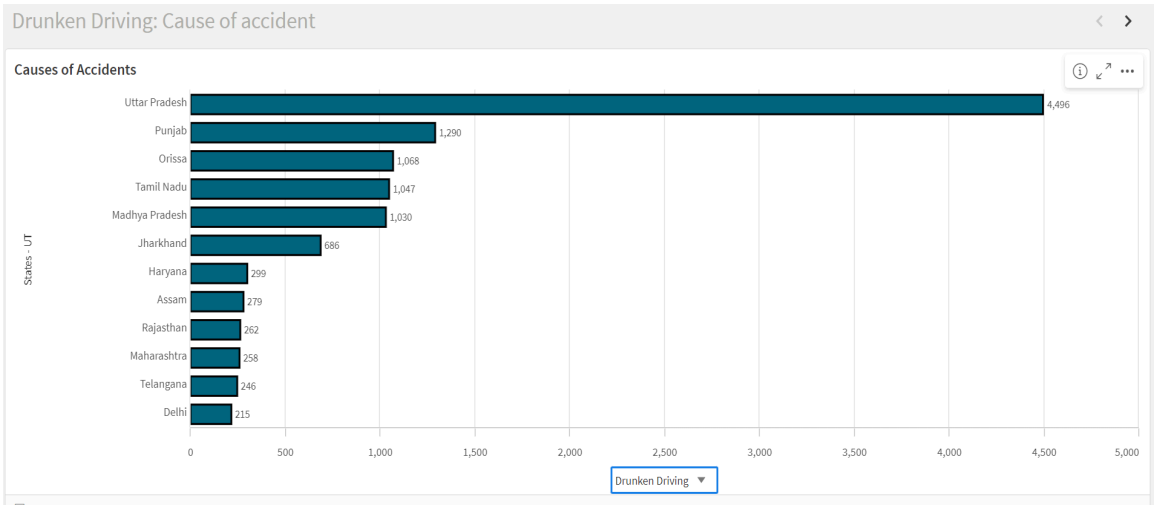
- **Re-upload Cleaned Data:**
  - After cleaning the dataset by removing unwanted columns and rows, re-upload the cleaned files to Qlik Sense.

4.3. Data Association

- **Qlik Sense Recommendations:**
  - Utilize Qlik Sense's recommendations for data associations to link related data fields across different tables.
  - Ensure that the data is properly connected and ready for comprehensive analysis.

5. DATA VISUALIZATIONS:

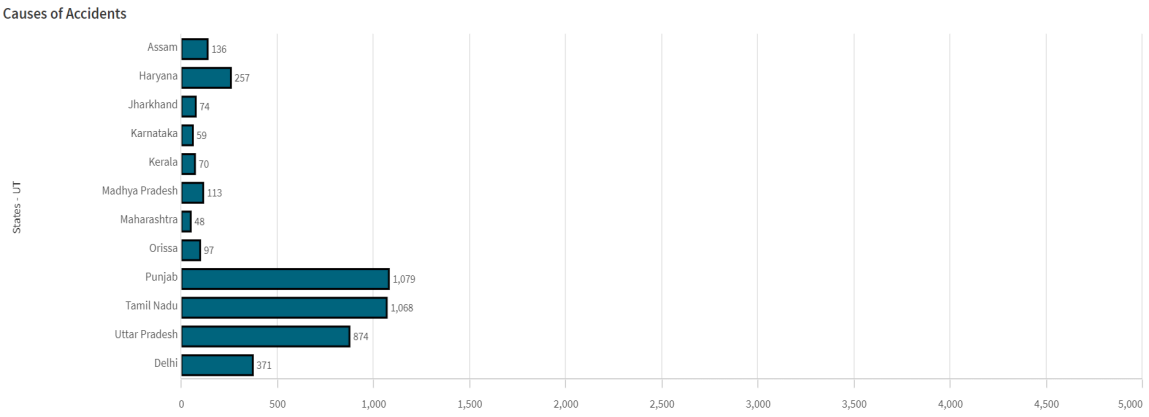
5.1. Bar Chart



Accidents due to Drunken Driving

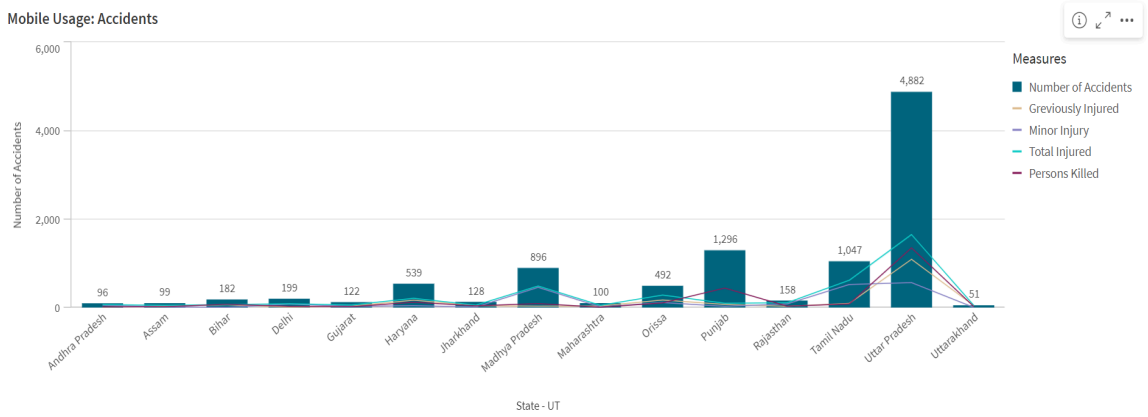
- The bar chart visualizes the frequency of accidents caused by drunken driving across various regions or states. The height of each bar represents the number of accidents, providing a clear comparison between different areas.
- **Insight**: The insight drawn from this chart indicates that certain regions have significantly higher incidents of drunken driving, suggesting a need for targeted awareness campaigns and stricter enforcement of driving regulations in these high-risk areas.

Qlik Analysis of Road Safety and Accident Patterns in India.



Accidents due to Jumping Red light

- The horizontal bar chart visualizes the count of accidents caused due to Jumping Reg light across various regions or states. The height of each bar represents the number of accidents, providing a clear comparison between different areas.
- **Insight**: The insight drawn from this chart indicates that certain regions have significantly higher incidents of red light, suggesting a need for targeted awareness campaigns and stricter enforcement of driving regulations in these high-risk areas.



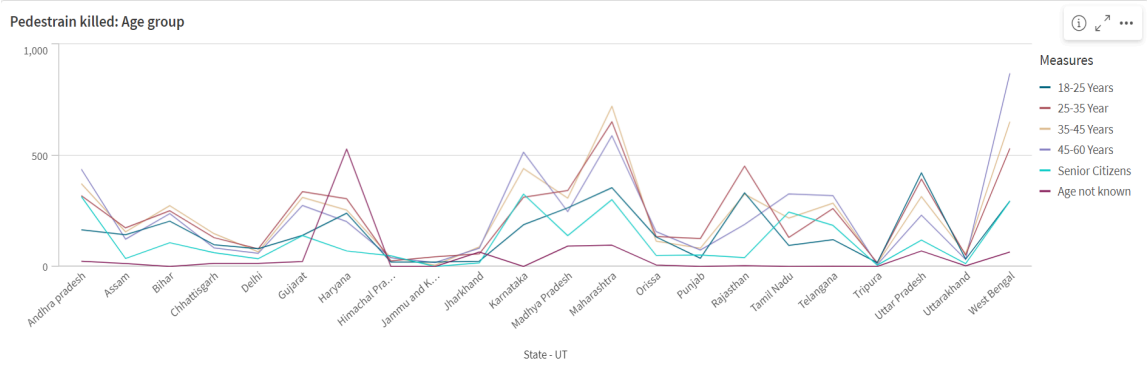
Accidents during Mobile usage

- This combo bar chart highlights the number of accidents associated with mobile phone usage while driving. The distribution across different states or regions shows how prevalent this risky behavior is and helps identify areas where mobile usage while driving is particularly problematic.
- The bar shows the number of Accidents and the line shows that what type of injuries

they had.

- **Insight**: The Insights suggest that regions with higher accidents due to mobile usage might benefit from campaigns focusing on the dangers of distracted driving and increased penalties for mobile usage behind the wheel.

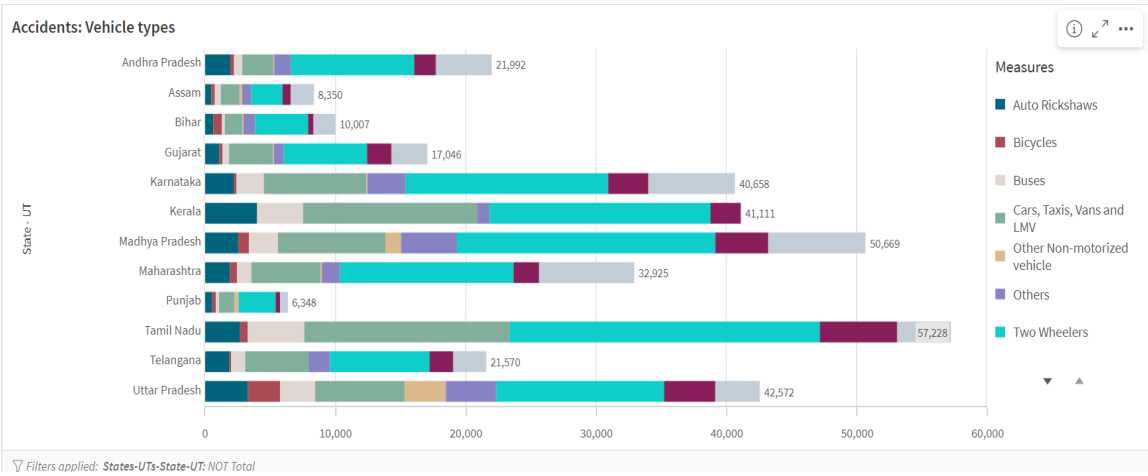
5.2. Line Chart



Pedestrians Killed by Age Group

- This line chart tracks the number of pedestrian fatalities across different age groups. The slope of the line indicates which age groups are more vulnerable to road accidents.
- **Insight**: The Insights from this chart reveal that certain age groups, such as the elderly or young adults, might be more susceptible to fatal accidents, highlighting the need for age-specific safety measures, such as better pedestrian crossings or targeted public safety messages.

5.3. Stacked BarChart



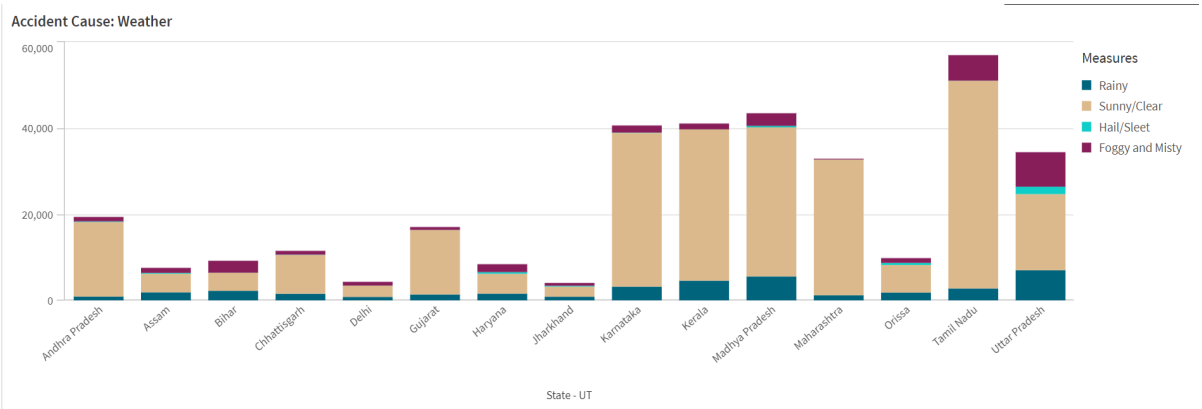
Accidents by Vehicle Types

- This stacked bar chart shows the proportion of accidents involving different types of

Qlik Analysis of Road Safety and Accident Patterns in India.

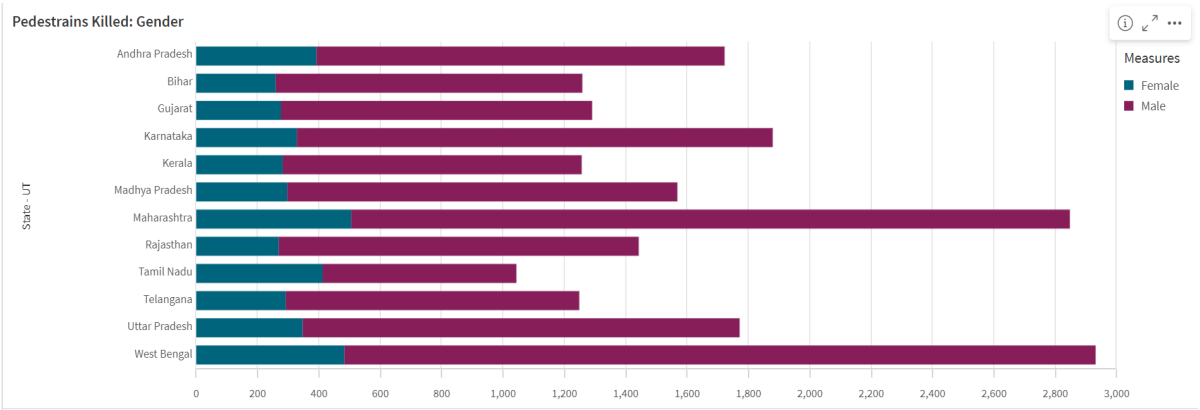
vehicles, such as two-wheelers, cars, and trucks. Each segment within the bar represents a different vehicle type, providing a visual representation of the distribution of accidents across vehicle categories.

- **Insight**: The Insights suggest that certain vehicle types, like two-wheelers, may be involved in a disproportionately high number of accidents, indicating a need for enhanced safety measures for these vehicles, such as better protective gear for riders or more stringent traffic regulations.



Accidents Based on Weather Conditions

- The stacked bar chart compares the number of accidents under various weather conditions, such as clear, rainy, and foggy days. Each segment of the bar corresponds to a specific weather condition, allowing for a comparison of how weather impacts road safety.
- **Insight**: The Insights drawn from this chart indicate that adverse weather conditions like rain or fog significantly contribute to the occurrence of accidents, suggesting the importance of improving road infrastructure, such as better drainage systems and clearer road markings, to mitigate these risks during bad weather.

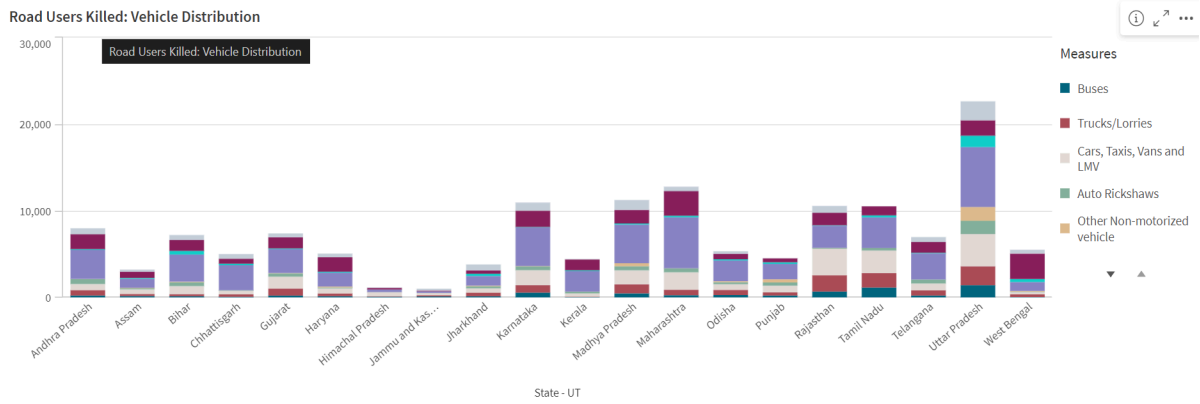


Overall Pedestrian Fatalities by Gender

- This visualization provides an overall view of pedestrian fatalities categorized by gender. It reveals the disparity between male and female pedestrian deaths, offering

insights into which gender is more vulnerable in road accidents.

- **Insight**: The insights suggest that targeted interventions might be required to address the specific safety needs of the more vulnerable gender group, such as enhancing street lighting and improving pedestrian facilities in areas where these fatalities are high.



Accidents Due to Vehicle Distribution

- This stacked bar chart visualizes the distribution of accidents based on different types of vehicles involved. Each segment within the bar represents a specific vehicle category, such as two-wheelers, cars, or trucks.
- **Insight**: The insights from this chart reveal which types of vehicles are most frequently involved in accidents, helping to identify high-risk vehicle categories. This information can guide safety interventions, such as targeted driver education programs or enhanced safety regulations for specific vehicle types.

5.4. Scatter Plot

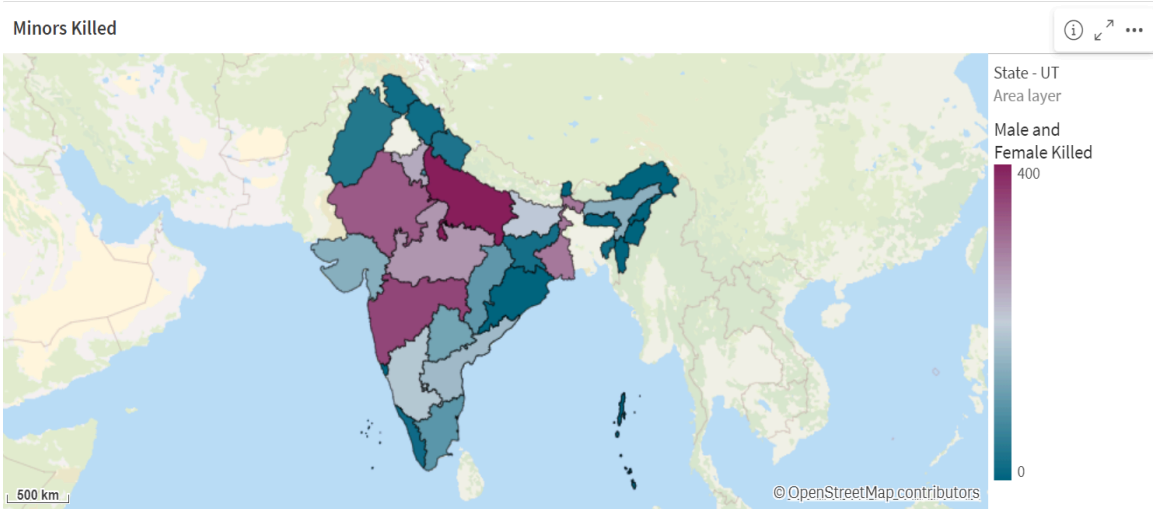
Relationship Between Average Number of Accidents in All States

- The scatter plot demonstrates a positive relationship between the average number of accidents across all states. Each point on the plot represents a state, with its position

indicating the frequency of accidents.

- **Insight**: The Insights from this visualization suggest that states with higher average accident rates may share common factors contributing to road accidents. This understanding can prompt further investigation into shared characteristics, such as infrastructure quality or traffic enforcement, to devise strategies that can be applied across multiple states.

5.5. Map Visualization

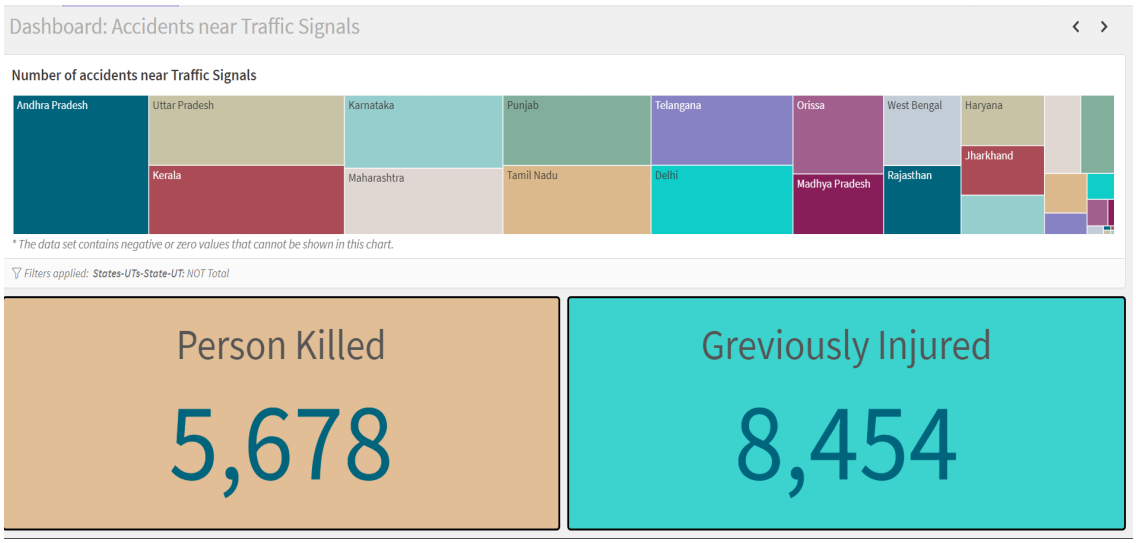


Accidents Due to Vehicle Distribution

- This map visualization highlights the states with the highest number of male and female fatalities in road accidents. The color intensity or markers on the map indicate the severity and location of these fatalities.
- **Insight**: The insights drawn from this map reveal regional patterns in gender-based fatalities, suggesting where gender-specific safety interventions might be necessary. For example, if certain states show a higher number of male fatalities, it could indicate risky driving behaviors prevalent among male drivers that need to be addressed.

6. DASHBOARD:

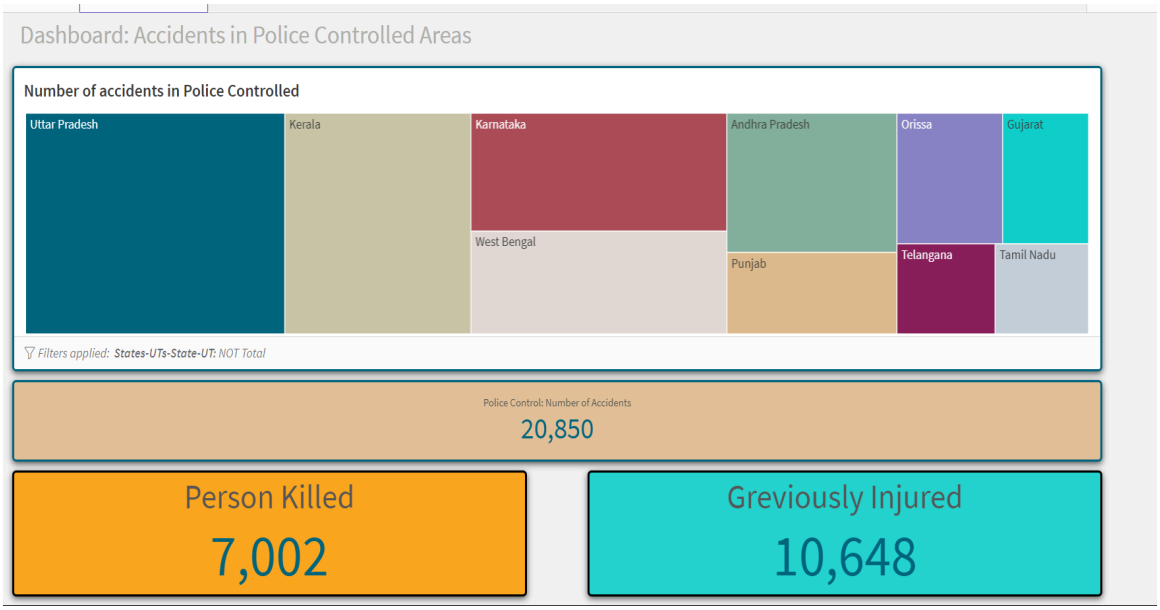
Qlik Analysis of Road Safety and Accident Patterns in India.



Tree Map and KPIs for Accident and Fatality Analysis

This dashboard utilizes a tree map to display the number of accidents across all states, with KPIs showing the total number of persons killed and grievously injured. The hierarchical structure of the tree chart allows for a detailed breakdown of accident data by state, while the Key Performance Indicator (KPIs) provide a quick overview of the most critical metrics. KPI show the number of persons killed that is 5.6k approximately and another KPI show the Injured person which are 8.4k approximately.

Insights from this dashboard suggest which states have the highest accident rates and the most severe outcomes, guiding resource allocation for emergency response and road safety improvements.



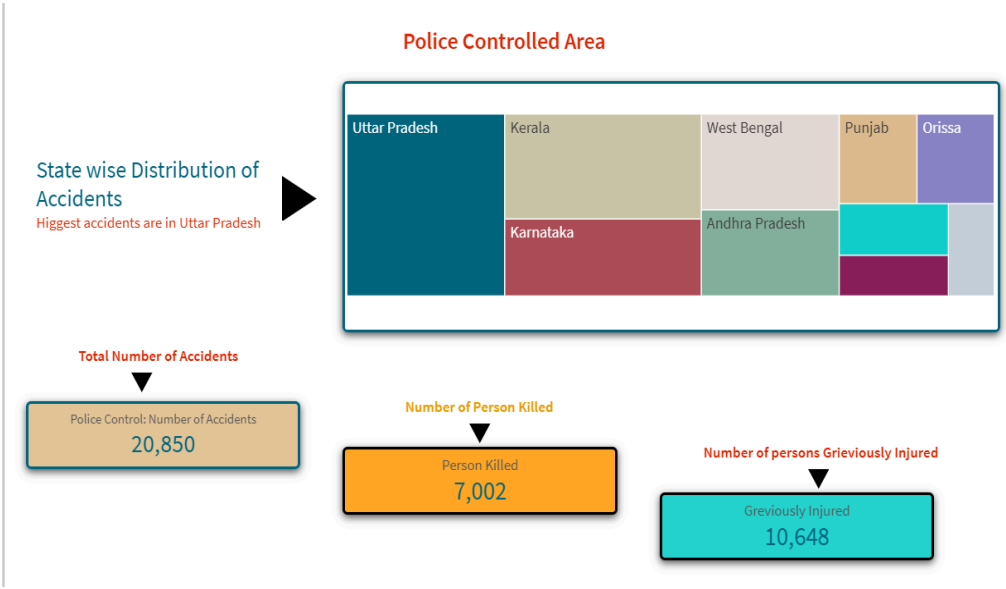
Qlik Analysis of Road Safety and Accident Patterns in India.

Tree Map and KPIs for Police-Controlled Area Accidents.

This dashboard focuses on accidents occurring in police-controlled areas, using a tree chart to show the number of accidents by state. The KPIs highlight the total number of accidents, persons killed, and grievously injured in these areas.

Insights from this dashboard can help assess the effectiveness of police-controlled traffic management and identify states where additional police resources or improved traffic control measures might be needed to reduce accidents and fatalities.

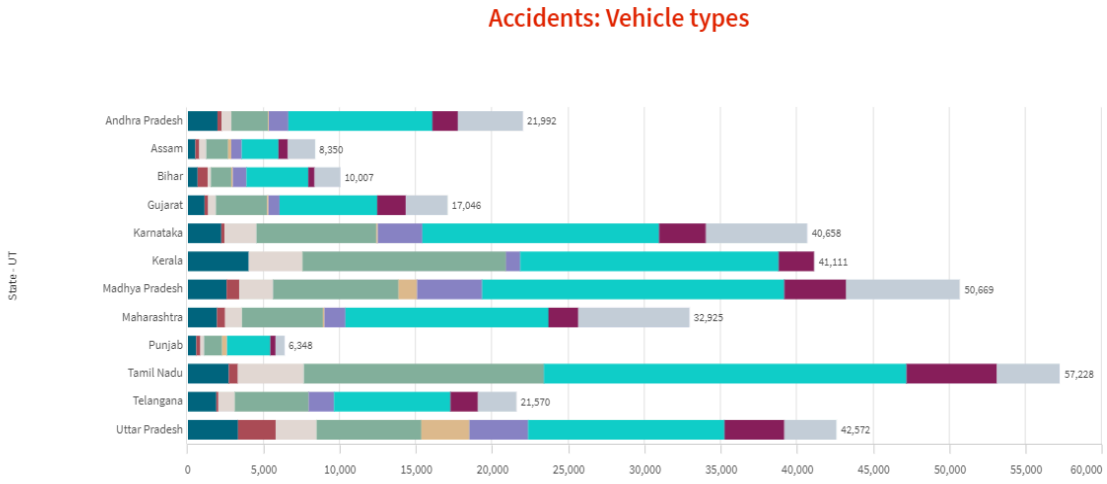
7. STORY



Tree Chart and KPIs for Accidents in Police-Controlled Areas

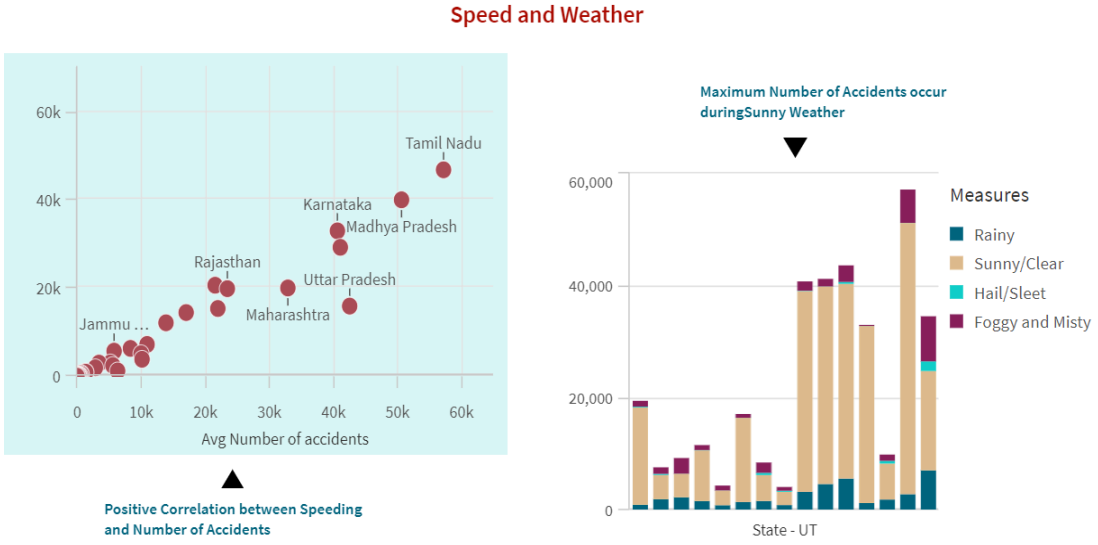
In this visualization we have a tree chart showing the number of accidents in all states, specifically focusing on police-controlled areas. Along with the number of accidents, the KPIs provide additional insights into the severity of these accidents, showing us how many people were killed or grievously injured. Despite police control, certain states still report a high number of serious accidents, indicating that additional measures—such as better enforcement, surveillance systems, or stricter penalties—are needed to improve safety in these areas.

Qlik Analysis of Road Safety and Accident Patterns in India.



Accident Hotspots by Vehicle Type

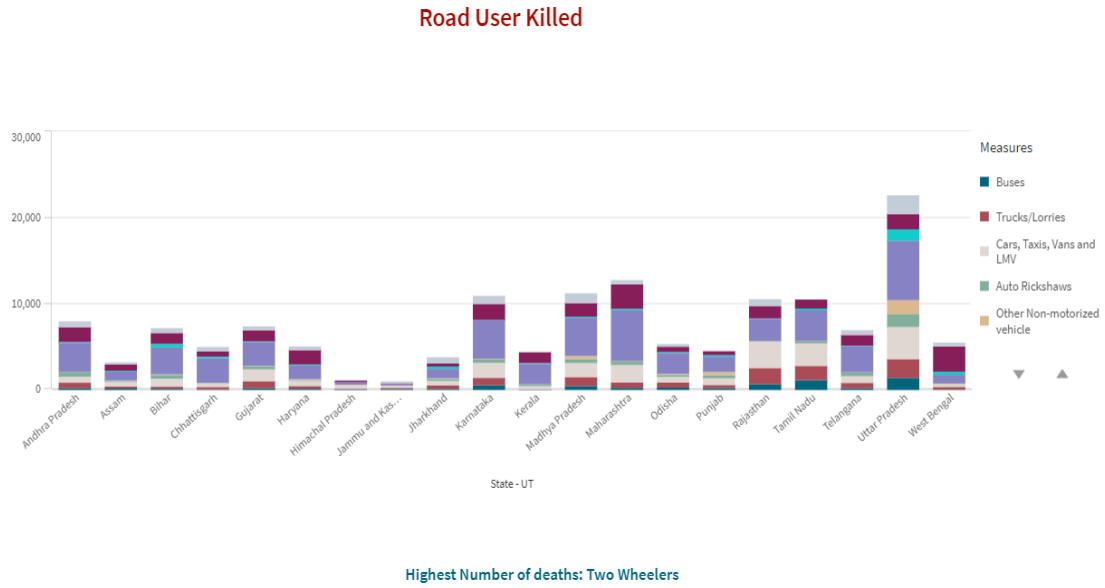
This visualization reveals the distribution of accidents by vehicle type across the country. As we can see, two-wheelers, cars, and trucks dominate the landscape of road accidents. Two-wheelers, in particular, represent a significant portion of these accidents, indicating a need for targeted safety interventions, such as stricter helmet enforcement and better rider training programs. By addressing these vehicle-specific risks, we can make a real impact in reducing accidents on the road.



Qlik Analysis of Road Safety and Accident Patterns in India.

Correlation Between Accidents and State-Wise Data

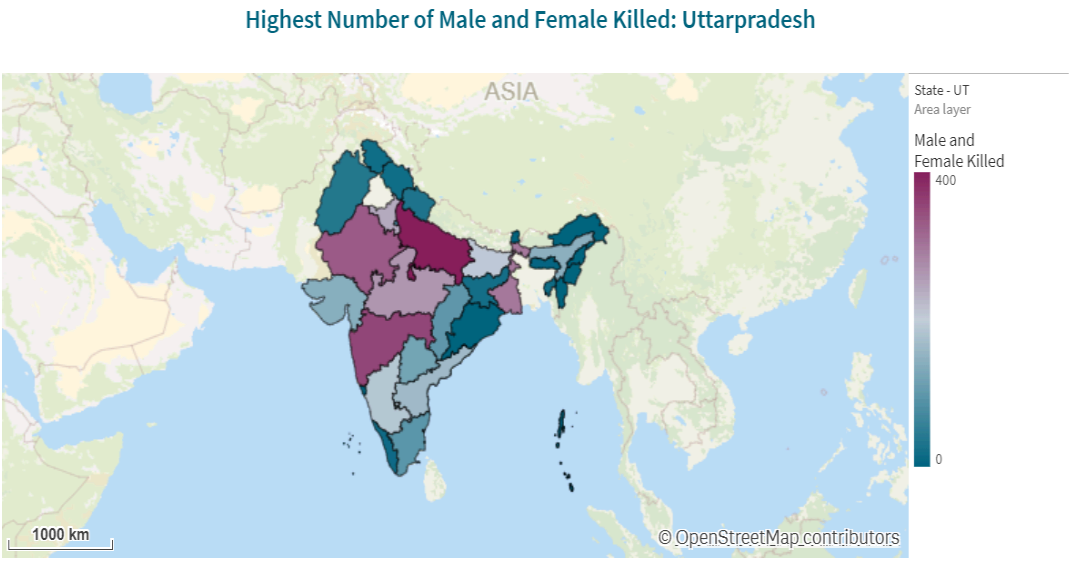
In this scatter plot, we can observe a positive correlation between the average number of accidents and the states. Each point represents a state, and it's clear that states with higher accident counts share certain characteristics, such as higher traffic volumes or poorer road infrastructure. This visualization can guide authorities to investigate these commonalities further and address state-specific risks with customized safety measures.



Accidents by Vehicle Distribution

A stacked bar chart, shows accidents by vehicle distribution. This chart clearly highlights the dominance of two-wheelers and cars in road accidents. As we can see, two-wheelers are particularly vulnerable, contributing significantly to the overall accident count. This insight calls for targeted interventions, such as improving helmet enforcement, educating riders on road safety, and enhancing road infrastructure to better accommodate two-wheeled vehicles.

Qlik Analysis of Road Safety and Accident Patterns in India.



Map Visualization of Male and Female Fatalities by State

Next, we look at a map visualization that shows the distribution of male and female fatalities by state. This map highlights the states where the number of fatalities is highest for both genders. The Uttarpradesh is the state which is having the highest number of deaths. What's particularly striking is the disparity between male and female fatalities, with males being disproportionately affected. This suggests a need for gender-sensitive road safety campaigns, focusing more on high-risk behaviors like speeding or reckless driving that are prevalent among male drivers.

These visualizations provide us with a data-driven roadmap for addressing road safety challenges. By focusing on key areas like vehicle safety, speeding, weather-related risks, and enforcement strategies, we can make meaningful improvements to reduce accidents and save lives on India's roads. The data is clear: we have the insights, and now it's time to take action

8. PERFORMANCE TESTING:

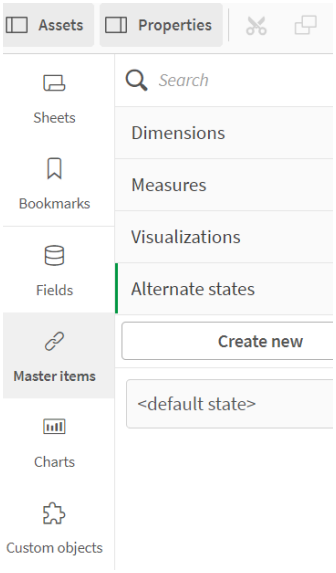
8.1. Application Of Data Filters:

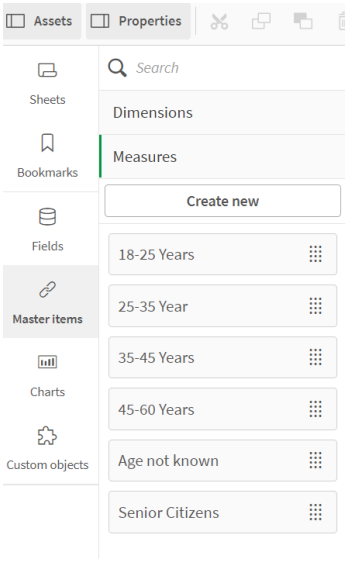
Qlik Analysis of Road Safety and Accident Patterns in India.

We have the option of selecting the data within the data which allows us to filter data based on individual fields or dimensions. We can choose specific values within a field to include or exclude from analysis. Complex filters based on predefined conditions and logic can also be created. In both the above charts we can see that we can see the data which we want it can be the top 20 states for what I want to see and I can go for 10 bottom states.

Qlik Analysis of Road Safety and Accident Patterns in India.

8.2. Calculated Fields:





We can use the Master items to create resuable filter objects, Calculated Fields which can simplify the process of applying consistent filters across multiple visualizations and dashboards.

8.3. Number of Visualizations

1. Accidents due to Drunken Driving
2. State-wise Mobile Phone Usage
3. Vehicle Contribution towards Total Accidents
4. Correlation - Speeding and Number of accidents
5. Accidents by Weather Type
6. Minors Injured across the country
7. Pedestrians Killed: Gender
8. Pedestrians Killed: Age groups
9. Road Users Killed: Vehicle Distribution

BY

DEEPAK SINGH