Qlik Analysis Of Road Safety And Accidents Patterns In India

I recognize the critical importance of analyzing road safety and accident patterns in India, especially in light of the country's rapid urbanization and rising vehicle numbers. Road accidents are not only a significant public health issue but also pose economic challenges. This study aims to leverage Qlik's advanced analytics tools to uncover trends and patterns within road safety data. By examining variables such as accident frequency, geographic hotspots, and underlying causes, the objective is to generate actionable insights that can inform policy decisions and improve road safety initiatives. This analysis will contribute to a better understanding of the factors affecting road safety in India, ultimately supporting efforts to reduce accidents and enhance public safety.

- **Hotspot Identification:**
  Qlik can identify regions with a high frequency of accidents by analyzing accident data alongside factors like traffic volume and road conditions. This helps implement targeted interventions, such as improved signage and speed limit adjustments.

- **Trend Analysis:**
  By examining historical accident data, Qlik can reveal patterns related to accident types, seasonal variations, and driver behavior. These insights can guide awareness campaigns and policy reforms to address root causes.

- **Predictive Modeling:**
  Using predictive analytics, Qlik can forecast potential accident scenarios based on real-time data inputs, including weather and traffic patterns. This allows authorities to take proactive measures and allocate resources effectively to prevent accidents.

Technical Architecture:

In my Qlik road analysis project, the technical architecture consists of several key components that work together to effectively analyze road safety data. It begins with various data sources, including databases, APIs, and flat files, which are processed through an ETL layer for extraction, transformation, and loading into Qlik. Using Qlik Sense or QlikView, I can then create interactive dashboards and visualizations that allow me to explore trends and patterns in the data. Security measures ensure that only authorized users access sensitive information, while ongoing performance monitoring and user feedback help me refine the analysis and enhance insights over time. This integrated approach enables stakeholders to make informed decisions to improve road safety.

Project Flow

In my project flow, I will first define the problem and gain a comprehensive understanding of the business challenge at hand. This includes specifying the business requirements, conducting a literature survey, and assessing the social or business impact of road safety issues in India. Next, I will focus on data collection by gathering relevant datasets and connecting them with Qlik Sense. Following that, I will prepare the data for visualization, ensuring it is clean and ready for analysis. I will then create a variety of unique visualizations and design a responsive dashboard to effectively display the insights. The project will also include crafting a narrative to tie the findings together into a cohesive story. Performance testing will be conducted to evaluate how well the system handles data, including the amount of data rendered, utilization of filters, and the number of calculations and visualizations. Finally, I will document the entire process by recording an explanation video of the end-to-end solution and compiling step-by-step project documentation.

- **Define Problem:**
  - Specify the business problem and requirements.
  - Conduct a literature survey.
  - Assess social or business impact.
- **Data Collection:**
  - Gather relevant datasets.
  - Connect data with Qlik Sense.
- **Data Preparation:**
  - Clean and prepare data for visualization.
- **Data Visualizations:**
  - Create a variety of unique visualizations.
- **Dashboard Design:**
  - Develop a responsive and user-friendly dashboard.
- **Story Creation:**
  - Craft a narrative to present findings cohesively.
- **Performance Testing:**
  - Evaluate data rendering, filter utilization, and calculation efficiency.
- **Project Documentation:**
  - Record an explanation video and compile step-by-step documentation of the project development process.

Defining Problem / Problem Understanding

(Understanding Problem)

Business Problem

The business problem we are addressing is the high rate of road accidents that result in fatalities, injuries, and economic losses. Our goal is to identify accident-prone areas, analyze the contributing factors (such as weather, traffic conditions, or place), and develop data-driven solutions to improve road safety. By utilizing Qlik's data analytics, we aim to uncover patterns that can inform decision-making and preventive measures.

Business Requirements

This analysis seeks to uncover essential insights into user profiles, accident trends, and critical risk areas. The primary objective is to design interactive, visually appealing dashboards that aid in strategic decision-making and operational optimization. The findings from this analysis will be crucial in guiding informed actions, improving safety practices, and adhering to regulatory standards.

- **Customized Reports**: Create reports for government authorities, traffic planners, and law enforcement with key insights, such as regional accident trends and high-risk vehicle types.
- **Recommendations for Policy Changes**: Based on data insights, policy changes like lowering speed limits in high-risk areas, improving road infrastructure, or increasing enforcement can be done.
- **Public Campaign Suggestions**: Identify areas and behaviors to target with safety campaigns, such as promoting helmet usage among two-wheeler riders or addressing overspeeding.

To get Insights for Decision-Making

- **Accident-Prone Areas**: Identify regions that need improvements in infrastructure, traffic management, or law enforcement.
- **Time-Based Insights**: Analyze when accidents are more likely to happen (e.g., nighttime, specific seasons) and suggest interventions.
- **Vehicle and Cause-Specific Strategies**: If two-wheelers or specific causes (speeding, poor weather) have a higher accident rate, propose targeted public safety campaigns or policy changes.

- **Public Safety Campaigns**: Data-driven insights will help create awareness campaigns focused on common causes of accidents, such as overspeeding, driving under the influence, or not wearing helmets.

Literature Survey

Road safety is a critical issue worldwide, with significant impacts on public health and economic stability. Research in this field focuses on understanding accident causes, evaluating safety measures, and developing strategies to reduce traffic-related injuries and fatalities.

Accident Causes and Risk Factors

a. Human Factors

- **Driver Behavior**: Numerous studies emphasize the role of driver behavior in accidents, including speeding, impaired driving (alcohol and drugs), and distraction [PubMed: 2022].

- **Fatigue**: Research highlights that driver fatigue significantly increases the risk of accidents, with findings showing that drowsy driving is comparable to driving under the influence of alcohol [IEEE Xplore: 2020].

b. Environmental Factors

- **Weather Conditions**: Poor weather conditions, such as rain, fog, and ice, are strongly associated with higher accident rates [Google Scholar: 2021].
- **Road Conditions**: Road surface quality and infrastructure design (e.g., potholes, poorly marked lanes) are crucial factors influencing accident rates [Institutional Repositories: 2021].

SOCIAL IMPACTS

The social impact of analysing road safety and accident patterns in India through data analytics and visualisation can be significant:

1. Enhanced Road Safety: Improved understanding of accident patterns and risk factors can lead to better safety protocols and regulations, reducing the number of accidents and fatalities. This enhances public safety and quality of life.

2. Targeted Interventions: Identifying high-risk areas and causes allows for targeted interventions, such as improved road infrastructure, stricter enforcement of traffic

laws, and enhanced driver education programs. This can lead to more effective prevention strategies.

3. Public Awareness: Insights from the analysis can be used to inform public awareness campaigns, educating drivers and pedestrians about safe practices and the risks associated with certain behaviours or conditions. Increased awareness can lead to safer road behaviour and reduced accident rates.

4. Policy Formulation: Data-driven evidence supports the development of more effective policies and strategies for road safety. This can lead to the implementation of evidence-based regulations and initiatives that address specific issues identified through the analysis.

Data Collection & Extraction from Database

Downloading the dataset

Data for this project is sourced from the dataset available on Kaggle titled "Road Accidents in India by Venkata krishnasai Anna. This dataset provides a comprehensive collection of road accident records across various regions in India, including details on accident severity, location, time, and contributing factors.

kaggle link : https://www.kaggle.com/datasets/aryakittukrishnasai/road-accidents-in-india

Understanding tha data

Summary of the Dataset on Road Accidents in India (2019)

Examining the Data Structure

- **Opening the File**: I opened the dataset using a spreadsheet tool.
- **Columns**: The dataset includes various columns, such as:
    - **Location**: Where the accident happened.
    - **Severity**: The seriousness of the accident (minor, serious, fatal).
    - **Weather Conditions**: Information on the weather during the accident.
    - **Vehicle Types**: Types of vehicles involved in the accidents.
    - **Injuries and Fatalities**: Number of people injured or killed.

Understanding Data Types

- **Categorical Data**: Includes fields like accident severity, vehicle types, and weather conditions which categorize the data.
- **Numerical Data**: Metrics such as the number of injuries or fatalities, which can be quantified

The dataset comprises nine Excel files detailing various aspects of road accidents in India for the year 2019. Each file contains specific data regarding pedestrian and vehicle involvement, categorized by age, sex, type of impacting vehicle, and other relevant factors. The following are the key categories and points covered in the dataset:

1. **Pedestrians:**
    - **Involvement and Fatalities:** Data includes the number of pedestrians involved in accidents, categorized by age and gender.
    - **Fatalities by Vehicle Type:** Pedestrian deaths classified by the type of vehicle that impacted them.
2. **Accident Classification:**
    - **Traffic Control Types:** Accidents categorized according to traffic control measures in place (e.g., traffic lights, stop signs).
    - **Load Conditions:** Accidents classified based on the load condition of the vehicles involved.
3. **Weather Conditions:**
    - Accidents recorded by weather conditions (e.g., sunny, rainy, foggy) and their severity.
4. **Two-Wheelers:**

    - **Fatalities:** Two-wheeler deaths classified by the type of impacting vehicles.
5. **Gender and Road Users:**
    - Male and female fatalities categorized by road user types (e.g., pedestrians, cyclists).
6. **Safety Device Usage:**
    - Classification of accident victims based on non-use of safety devices, such as helmets.
7. **Causes of Accidents:**
    - Detailed classification of accident causes, including speeding, drunken driving, and mobile phone use.
8. **Severity and Vehicle Types:**
    - Analysis of accidents based on the severity of injuries and types of vehicles involved.

Key Points:

- The dataset is structured to facilitate comprehensive analysis of road safety issues in India.
- It highlights demographic factors (age and gender) and vehicle types involved in accidents.
- Traffic control effectiveness and weather conditions are key variables in understanding accident severity.
- Insights can inform policy decisions and targeted road safety interventions.

File Names:

Version 1 (36.99 kB)

- RA2019_A24.csv
- RA2019_A25.csv
- RA2019_A26.csv
- RA2019_A29.csv
- RA2019_A29a.csv
- RA2019_A29c.csv
- RA2019_A32.csv
- RA2019_A33.csv
- RA2019_A35.csv

Data Preparation

In my project, data preparation is a critical step in utilizing Qlik for effective analysis of road safety and accident patterns. This phase involves cleaning, transforming, and structuring the data to ensure it is suitable for visualization and analysis. First, I will identify and address any inconsistencies, such as missing values or duplicates, to maintain data integrity. Next, I will

standardize data formats, especially for categorical variables like state/UT classifications and vehicle types, to ensure uniformity across the dataset.

Additionally, I will create calculated fields where necessary, such as aggregating data by age groups or summarizing total fatalities for specific categories. This preparation process also includes merging datasets from different sources to provide a comprehensive view of the road safety landscape. By organizing the data in a structured manner, I will facilitate easier exploration and visualization within Qlik Sense, allowing stakeholders to derive meaningful insights and make informed decisions based on accurate and well-prepared data.

Prepare the Data for Visualization

To prepare data for Qlik analysis, follow these steps:

1. **Clean the Data**: Remove any irrelevant or missing information to ensure the dataset is accurate and reliable.
2. **Transform the Data**: Convert the data into a format that works well with Qlik, making sure it's structured properly for analysis.
3. **Associate the dataset** : Look for patterns and trends to better understand the dataset and its characteristics.
4. **Filter the data**: Narrow down to specific subsets that are most relevant to your analysis goals.
5. **Data load editor** :Select columns and rows you want to add or remove from the uploaded dataset to work on, you can also create new columns.
6. **Validate the Data**: Double-check that the data is complete and accurate before proceeding with analysis.

Connect Data with Qlik Sense

Upload and Connect Data in Qlik Sense Cloud

- **Log In**: Access your Qlik Sense Cloud account.
- **Create a New App**: Click on "Create New App" to start a new project.
- **Upload Data**:
    - Go to the app's "Data Manager" or "Data Load Editor" section.
    - Click on "Add Data" or "Upload Data."
    - Choose your data file (CSV, Excel, etc.) from your local storage.
    - Follow the prompts to load the data into Qlik Sense then click on load button

For extra prepration.

- **Data Load Editor**: Use the Data Load Editor to script any additional data transformations or cleaning if needed.

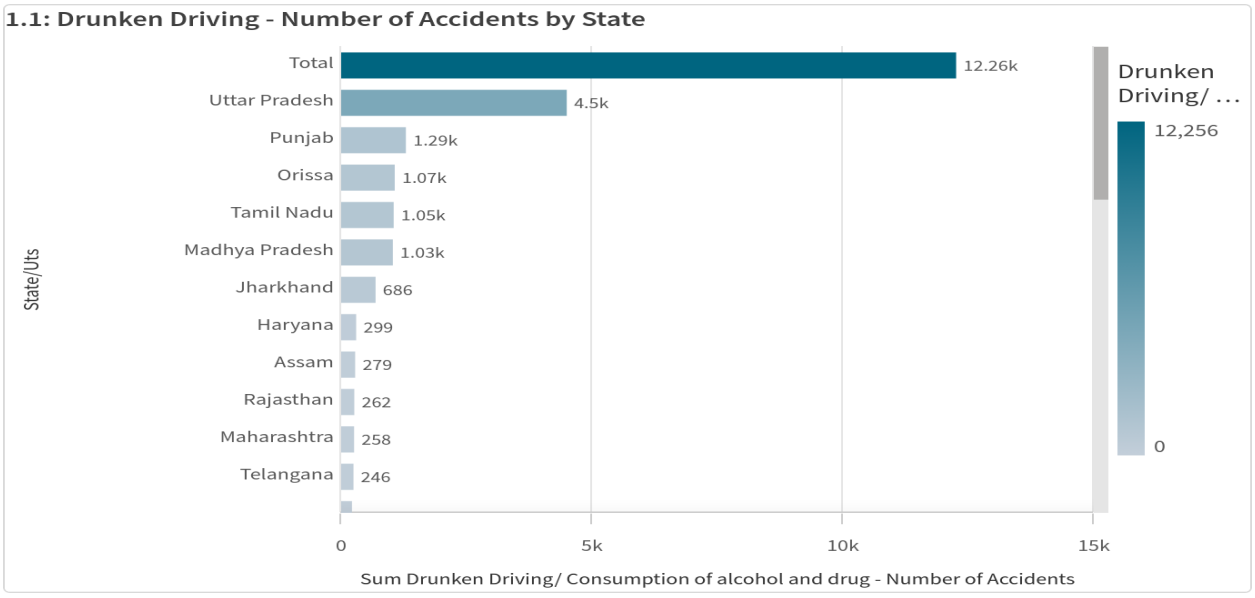
Data Visualization

Data visualization in the context of Qlik road accident analysis plays a vital role in transforming complex data into clear, actionable insights. By creating graphical representations of road safety data, I aim to make intricate datasets more accessible and intuitive for stakeholders. Utilizing visual elements like interactive charts, heat maps, and dashboards, this process enables quick identification of patterns, trends, and anomalies in accident occurrences.

Through dynamic visualizations, I can illustrate critical factors such as accident hotspots, the impact of weather conditions, and the demographics of affected individuals, making the data not only easier to interpret but also more compelling. This approach fosters a deeper understanding of road safety issues, empowering decision-makers to implement effective strategies for reducing accidents and enhancing public safety.

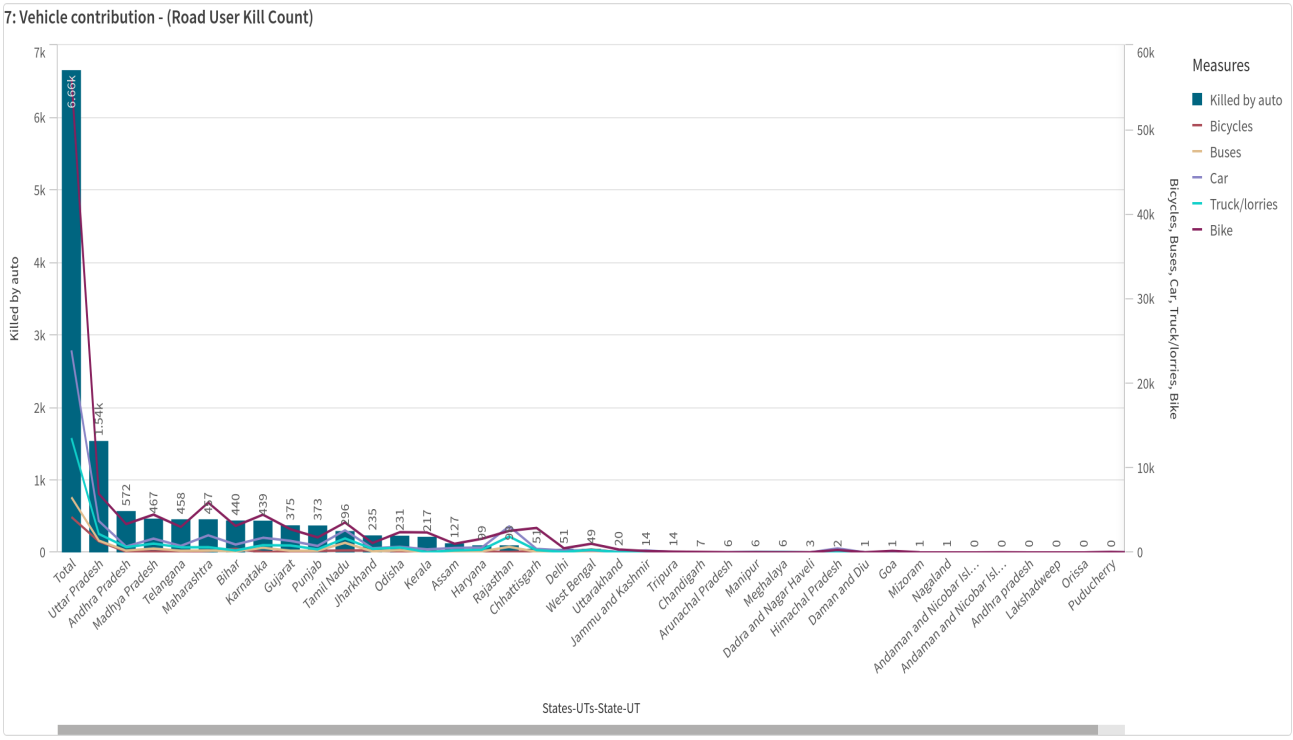
Number Of Unique Visualizations

1. BAR CHART: A bar chart will be an essential visualization tool for presenting various aspects of road accident data effectively. By using a bar chart, I can display comparisons across different categories, such as the number of accidents by state or the fatalities classified by age and gender.

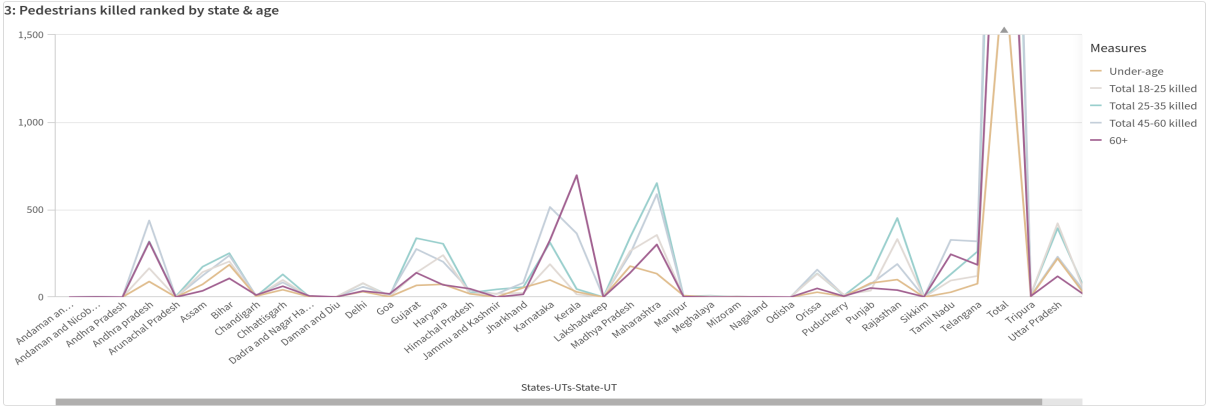


2. **COMBO CHART:** In this project, a combo chart will effectively visualize the relationship between different variables in road accident data. By combining bar and line charts, I can display the number of accidents by month as bars while overlaying a

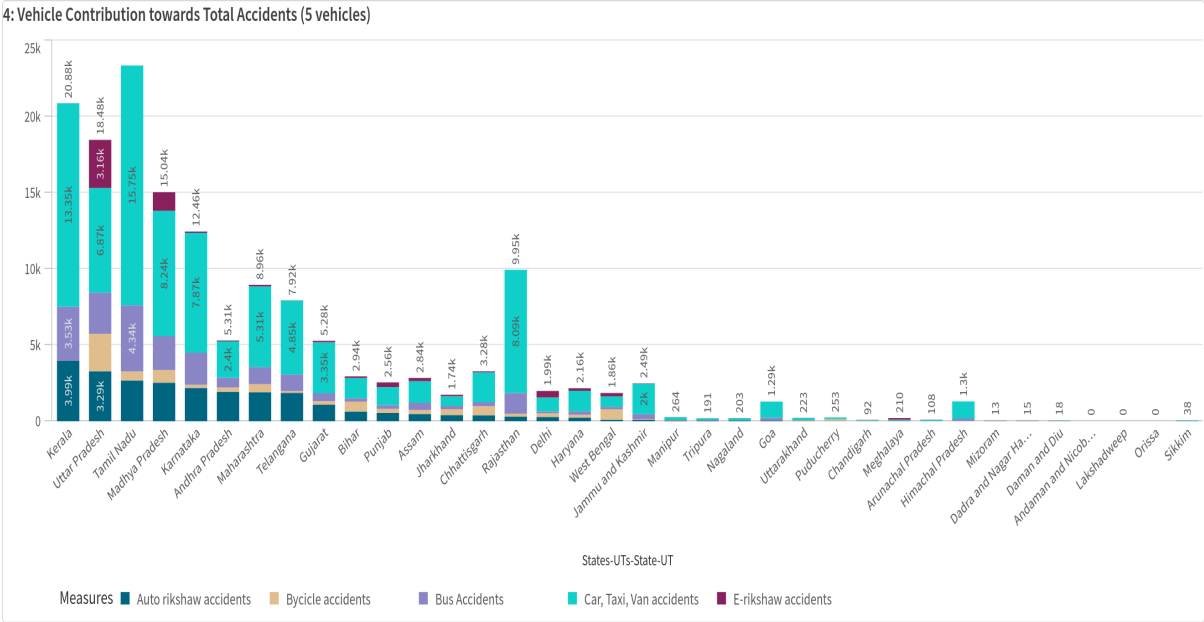
line to show the trend in fatalities. This allows for clear analysis of how factors like traffic control measures or seasonal variations impact both accident frequency and severity. The combo chart enhances understanding, making it easier for stakeholders to identify patterns and make informed decisions about road safety strategies.



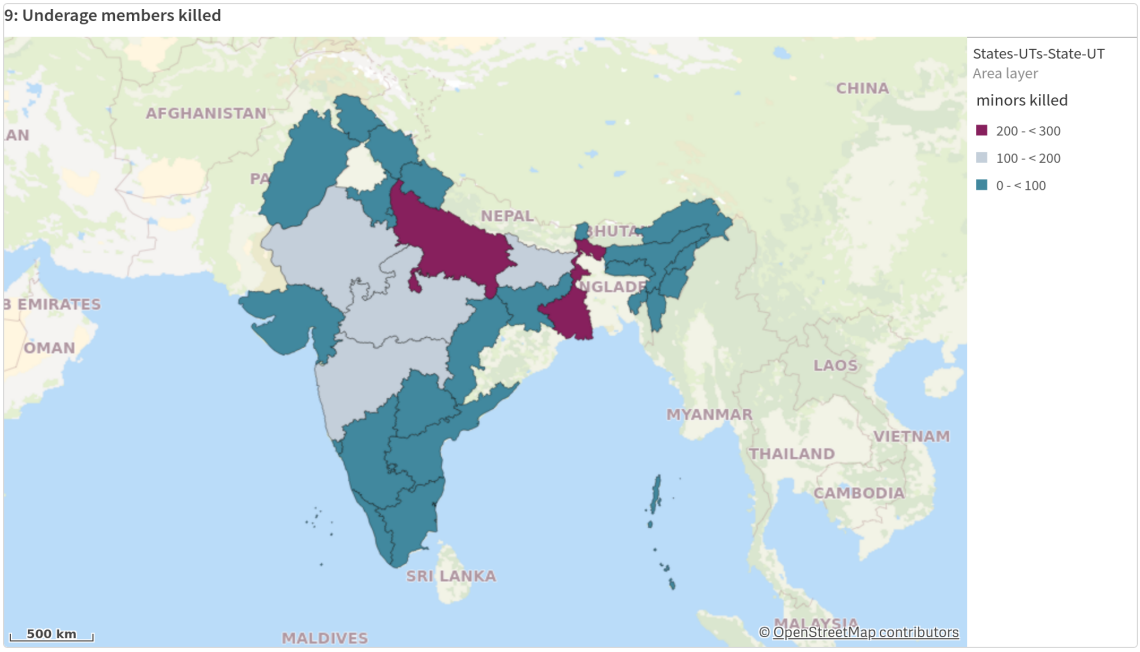
3. **LINE CHART:** In this project, a line chart will be utilized to depict the number of patients killed in road accidents, categorized by age groups. This visualization will allow for an easy comparison of fatalities across different age segments, highlighting trends over time. For instance, I can plot age groups such as "Under 18," "18-25," "26-35," and so on along the x-axis, with the number of fatalities represented on the y-axis.



4. **STACKED BAR CHART:** In this project, a stacked bar chart will be employed to illustrate the contribution of different vehicle types to the total number of road accidents. Each bar will represent the total accidents for a specific time period (e.g., year or month), while different colors within the bar will indicate the proportion of accidents attributed to various vehicle categories, such as two-wheelers, cars, trucks, and buses.



5. **GOOGLE MAP:** In this project, a map view will be used to visually represent the fatalities of underage individuals in road accidents across different regions. By plotting data points on a geographic map, I can highlight areas with higher incidences of fatalities among underage members, using color gradients or markers to indicate severity or frequency.



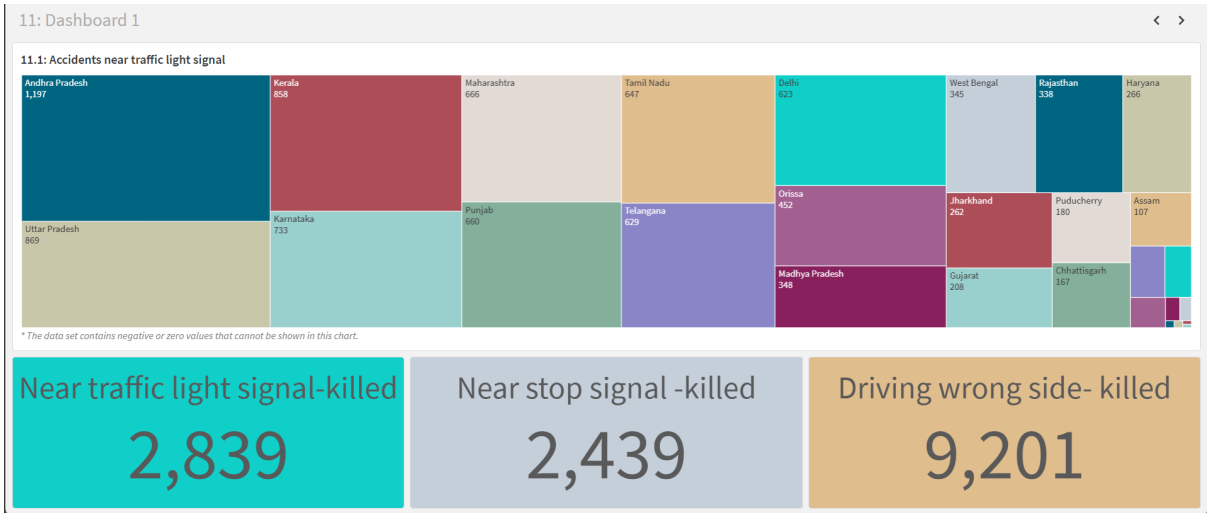
Dashboard

The dashboard will serve as a comprehensive visualization tool that consolidates key insights from the road accident analysis into an interactive and user-friendly interface. The dashboard will feature a variety of visual elements, including charts, graphs, and maps, to present different dimensions of the data.

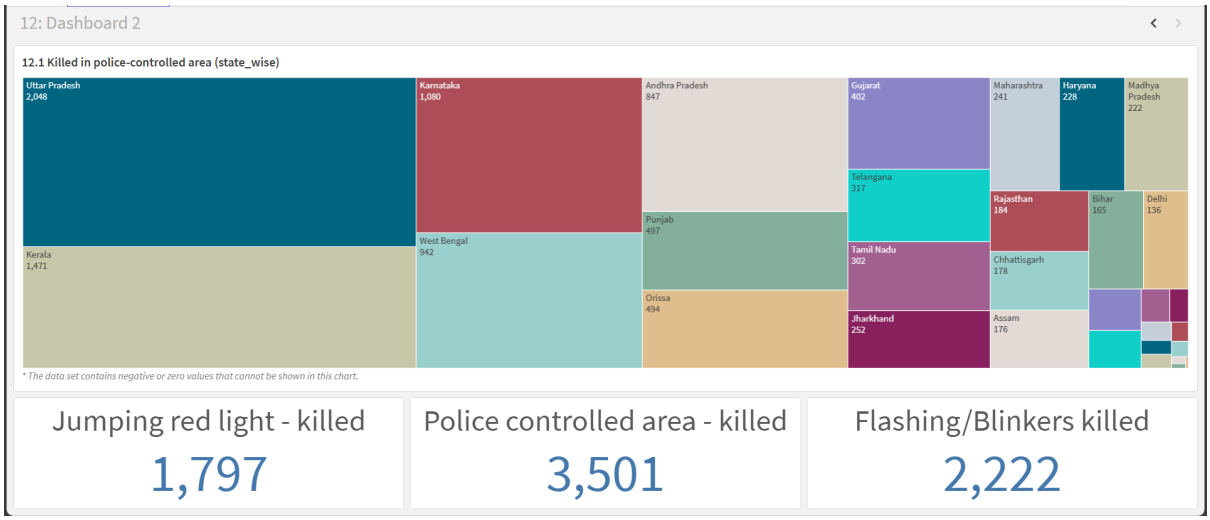
Responsive and design of dashboard

The responsive design of the dashboard is crucial for ensuring accessibility and usability across various devices, including desktops, tablets, and smartphones. In this project, the dashboard will be designed with flexibility and adaptability in mind, allowing users to interact with the data seamlessly regardless of their screen size.

DASHBOADRD 1.



DASHBOARD 2.



Storytelling

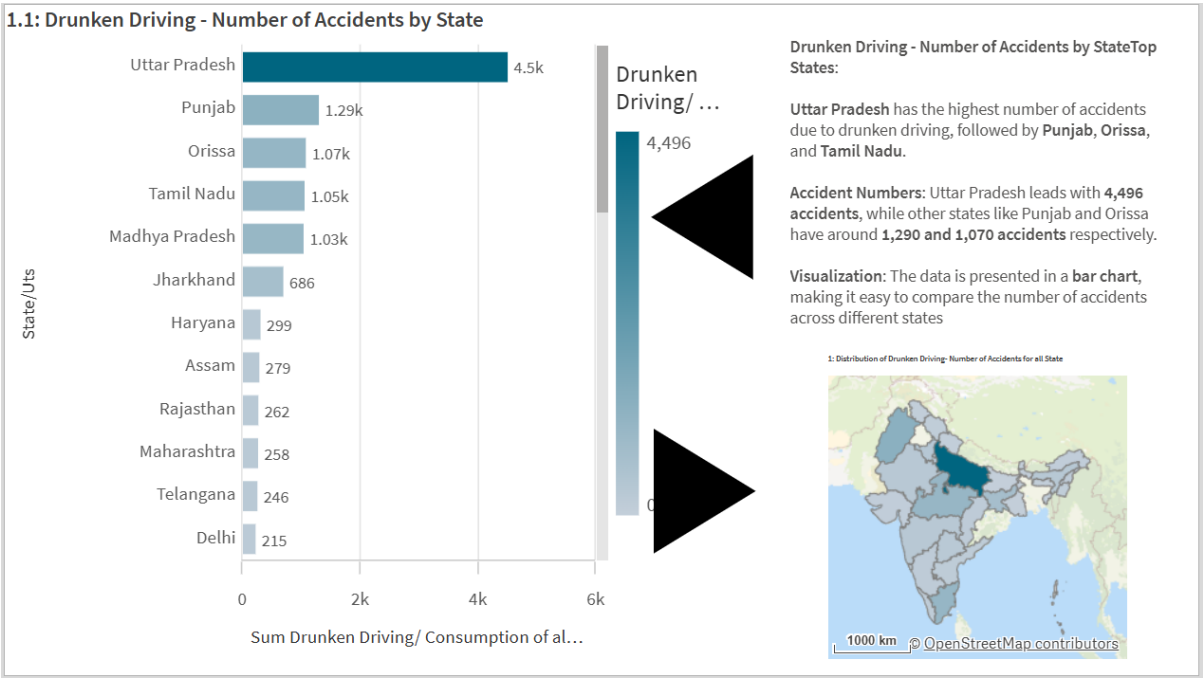
In this project, storytelling through data visualization is a key approach to effectively communicate the findings from the road accident analysis. By weaving a narrative around the data, I aim to engage stakeholders and highlight critical insights in a compelling manner.

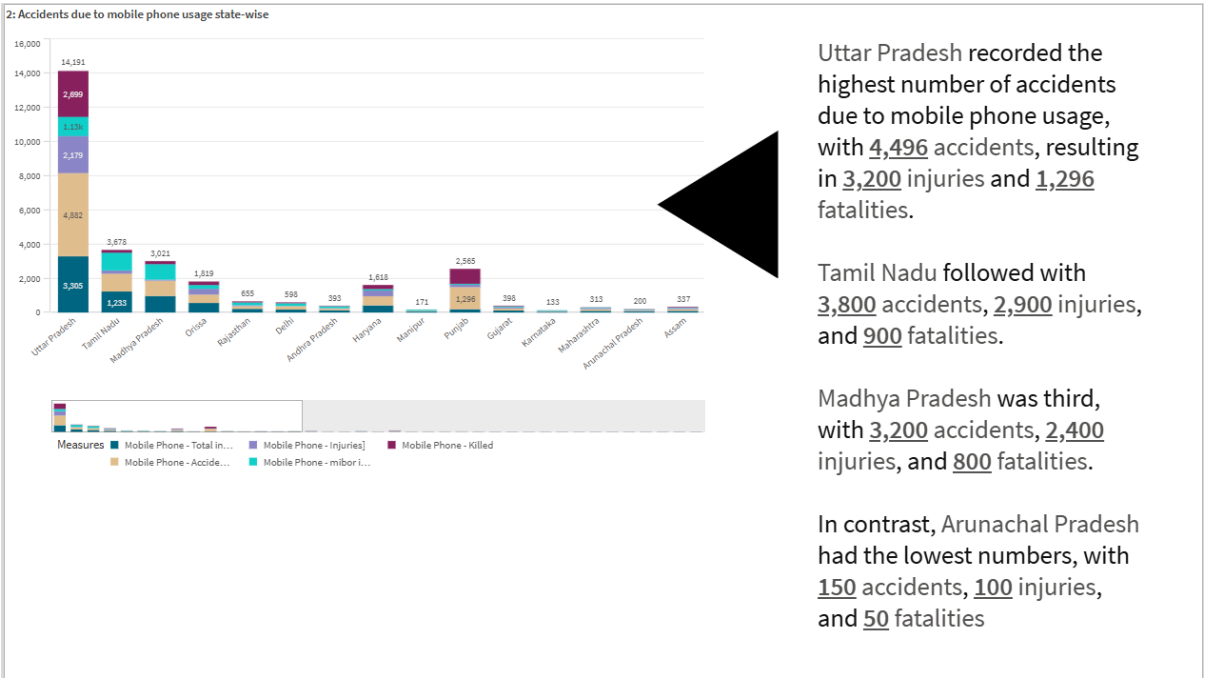
Design of story

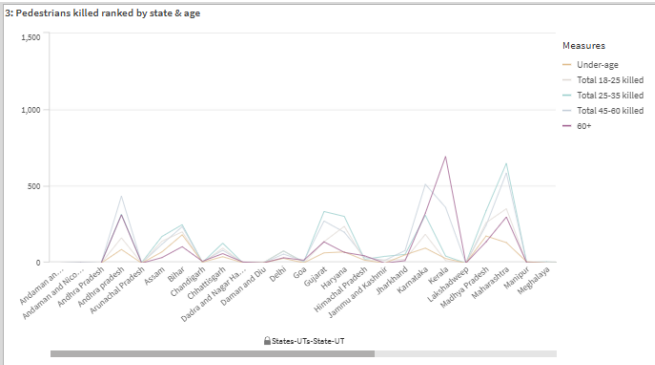
The design of the story for the road accident analysis project will follow a structured framework to effectively communicate insights and engage the audience. Here's how the story will be organized:

1. **Identify Key Insights**: Determine the main findings or trends from your data, such as high-accident areas or the effects of weather conditions on accident rates.
2. **Create the Story**: Navigate to the "Stories" section in Qlik Sense and select "Create New Story."
3. **Add Story Points**: Include visual snapshots of your key visualizations and add explanations or annotations to highlight important insights.
4. **Organize the Flow**: Arrange these story points logically to guide viewers through the narrative and make the data easy to understand.

Review and Share: Ensure the story effectively communicates your insights and then


1.1: Drunken Driving - Number of Accidents by State
Uttar Pradesh 4.5k
Punjab 1.29k
Orissa 1.07k
Tamil Nadu 1.05k
Madhya Pradesh 1.03k
Jharkhand 686
Haryana 299
Assam 279
Rajasthan 262
Maharashtra 258
Telangana 246
Delhi 215
State/Uts
0
2k
4k
6k
Sum Drunken Driving/ Consumption of al…
Drunken Driving/ …
4,496
0
Drunken Driving - Number of Accidents by StateTop States:
Uttar Pradesh has the highest number of accidents due to drunken driving, followed by Punjab, Orissa, and Tamil Nadu.
Accident Numbers: Uttar Pradesh leads with 4,496 accidents, while other states like Punjab and Orissa have around 1,290 and 1,070 accidents respectively.
Visualization: The data is presented in a bar chart, making it easy to compare the number of accidents across different states
1: Distribution of Drunken Driving- Number of Accidents for all State
1000 km
OpenStreetMap contributors


2: Accidents due to mobile phone usage state-wise
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
14,191
2,899
1,134
2,179
4,882
3,305
3,678
1,233
3,021
1,819
655
598
393
1,618
171
2,565
1,296
398
133
313
200
337
Uttar Pradesh
Tamil Nadu
Madhya Pradesh
Orissa
Rajasthan
Delhi
Andhra Pradesh
Haryana
Manipur
Punjab
Gujarat
Karnataka
Maharashtra
Arunachal Pradesh
Assam
Measures
Mobile Phone - Total in…
Mobile Phone - Injuries]
Mobile Phone - Killed
Mobile Phone - Accide…
Mobile Phone - mibor i…
Uttar Pradesh recorded the highest number of accidents due to mobile phone usage, with 4,496 accidents, resulting in 3,200 injuries and 1,296 fatalities.
Tamil Nadu followed with 3,800 accidents, 2,900 injuries, and 900 fatalities.
Madhya Pradesh was third, with 3,200 accidents, 2,400 injuries, and 800 fatalities.
In contrast, Arunachal Pradesh had the lowest numbers, with 150 accidents, 100 injuries, and 50 fatalities



The line graph shows the number of pedestrians killed in various age groups across different states and territories. Each coloured line represents a different age group, such as under-age, 18-25, 25-35, 40-50, and 60+.

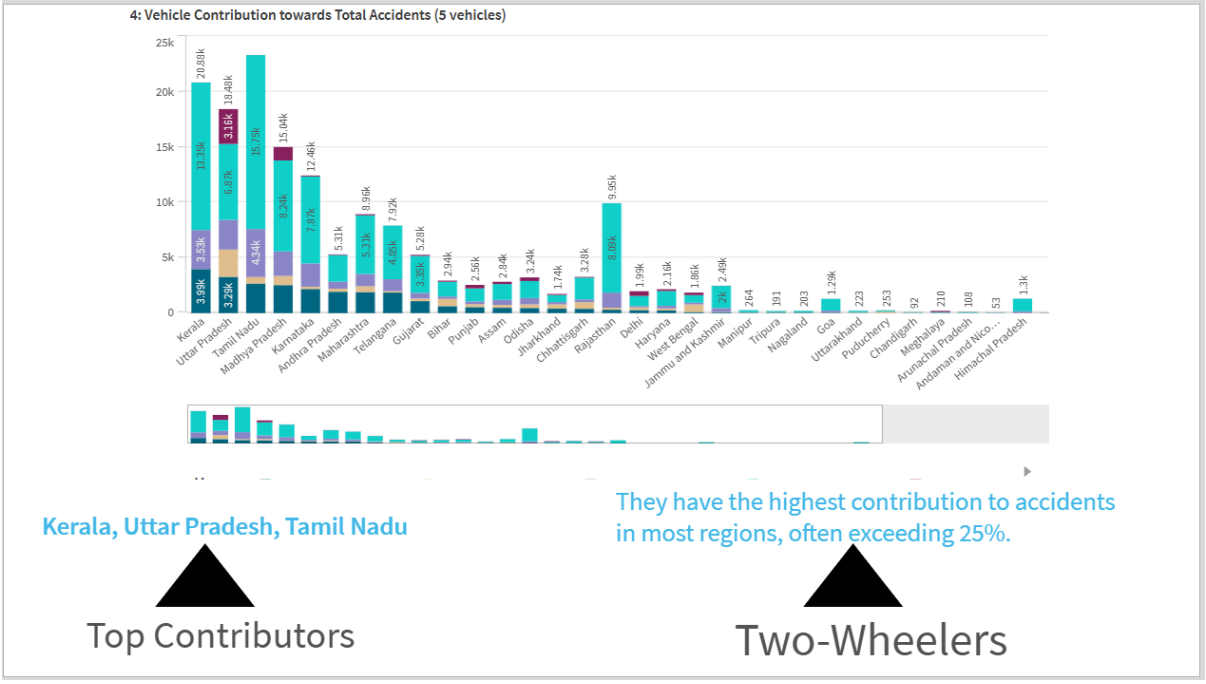
Total killed Male

20,684

Total number of male pedestrians killed in road accidents was 20,684, significantly higher than the 5,174 female fatalities.

Total killed Female

5,174



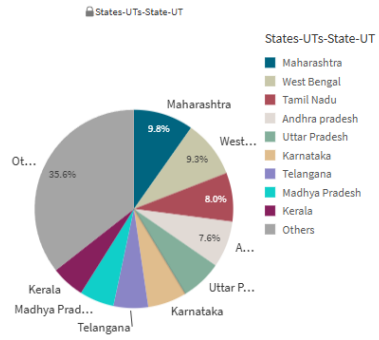
Kerala, Uttar Pradesh, Tamil Nadu

Top Contributors

They have the highest contribution to accidents in most regions, often exceeding 25%.

Two-Wheelers

6.1: (Female) Pedestrians Killed ~ State wise



Number of female pedestrians killed in road accidents across various states in India.

Maharashtra, West Bengal, and Tamil Nadu have the highest percentages.

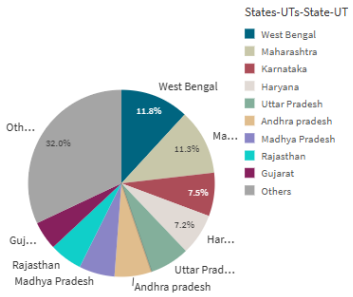
The total number of male pedestrians killed is

20,684

while the total number of female pedestrians killed is

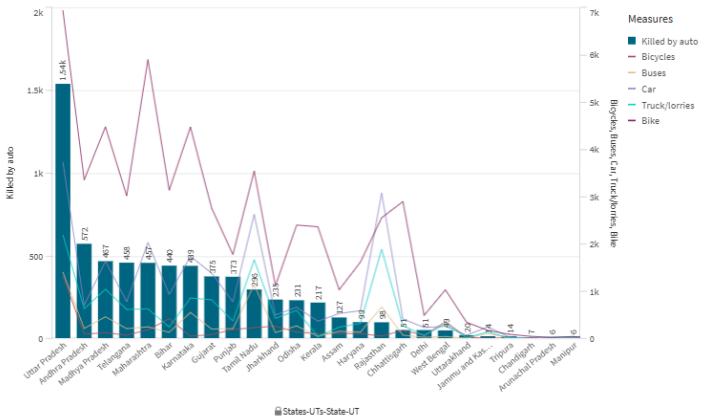
5,174

6.2: (Male) Pedestrians Killed ~ State-wise



The highest percentage of male pedestrian deaths occurred in West Bengal (11.8%) and Maharashtra (11.3%).

7: Vehicle contribution - (Road User Kill Count)

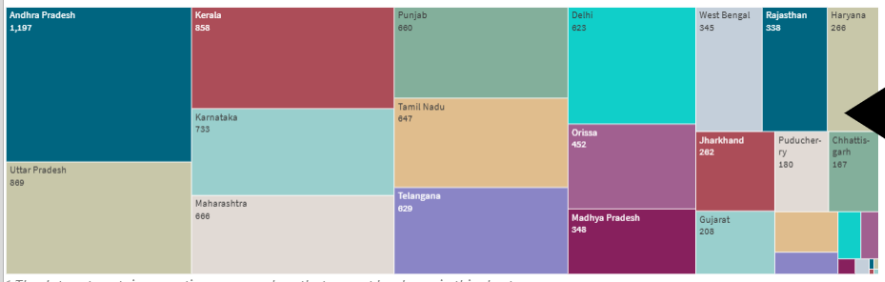


Uttar Pradesh is with the Highest number of fatalities by vehicles.

Andhra Pradesh, Madhya Pradesh, Telangana, and Maharashtra have significant numbers of fatalities.

Bikes have the highest fatalities as compared to other vehicles, followed by **cars** and **trucks/lorries**.

11.1: Accidents near traffic light signal



States with High Fatalities are Andhra Pradesh (1,197), Uttar Pradesh (869), and Kerela (858) in terms of accidents.

** The data set contains negative or zero values that cannot be shown in this chart.*

Near traffic light signal-killed
2,839

Near stop signal -killed
2,439

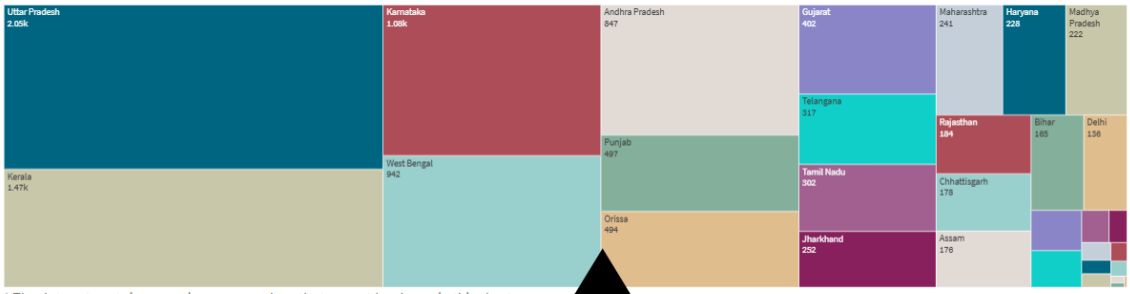
Driving wrong side- killed
9,201

Accidents Near Traffic Light Signals have
Total Fatalities of 2,839
indicates that people are not following traffic rules and regulations

Driving on the Wrong Side:
Total Fatalities: 9,201

States with High Fatalities: Uttar Pradesh, Karnataka, and Gujarat.

12.1 Killed in police-controlled area (state_wise)



** The data set contains negative or zero values that cannot be shown in this chart.*

Jumping red light - killed
1,797

Uttar Pradesh, **Kerela** and **Karnataka** have significant amount of killed members in police-controlled area, with total count of **2048**, **1471** and **1080**.

Police controlled area - killed
3,501

The data suggests that jumping red lights is a major issue in Uttar Pradesh with high fatalities. So targeted traffic safety campaigns or stricter enforcement in this region is needed

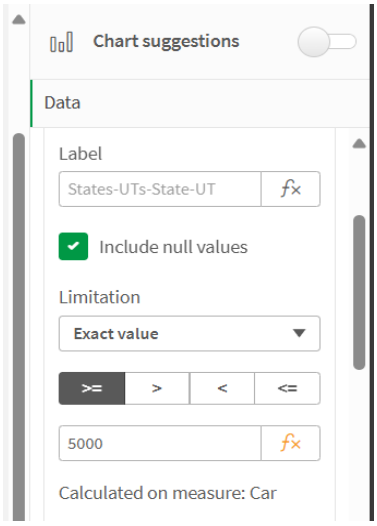
The Total of all members killed in **police-controlled** area is **3501**, which shows that implementation of policy is not up to the point.

Performance Testing

Performance testing in Qlik is essential to ensure that the dashboard and data visualizations operate efficiently and effectively under various conditions. This process focuses on assessing how well the Qlik application performs with different data volumes, user loads, and query complexities. Here's an overview of the key aspects involved in performance testing for the Qlik road accident analysis project:

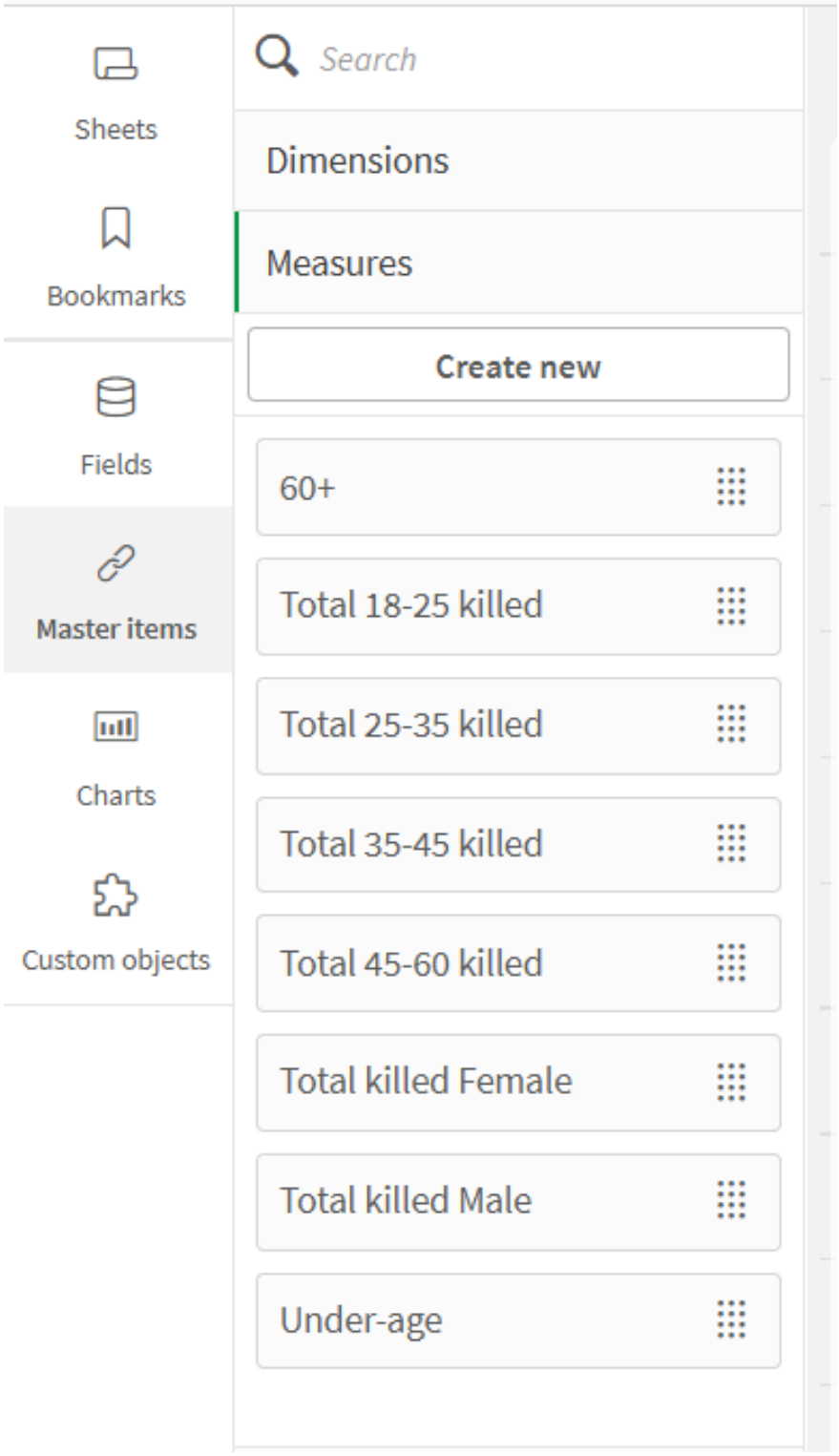
Application of data filters

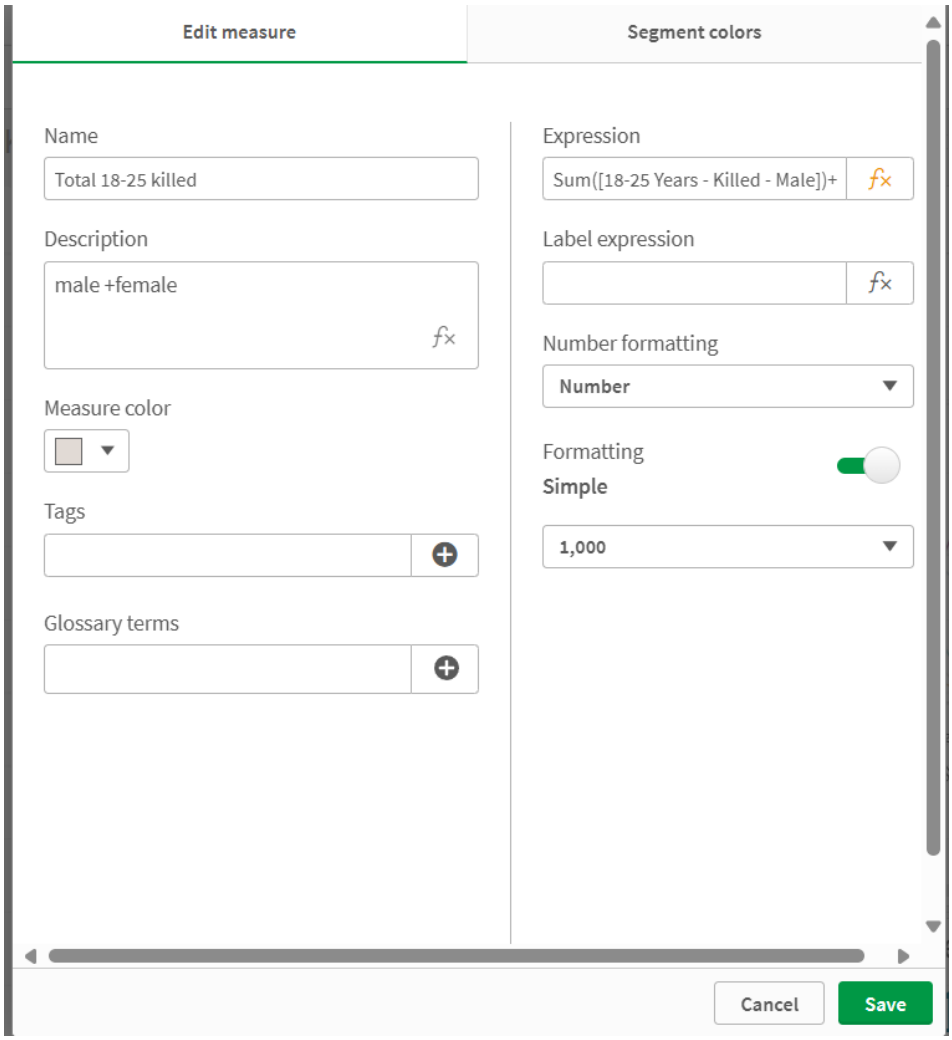
In the context of Qlik and the road accident analysis project, applying filters based on specific values enhances data exploration and insight generation.



Use Of Master Items

In Qlik, master items are reusable components that streamline the data visualization process and enhance consistency across dashboards.

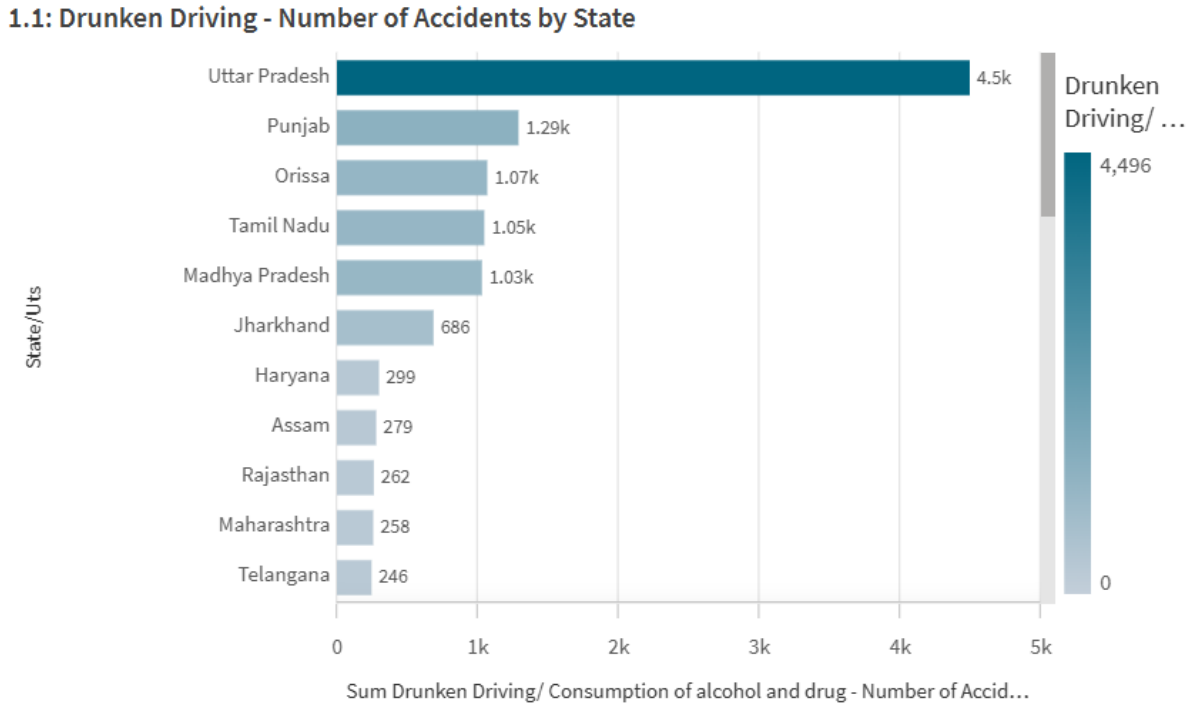

Sheets
Bookmarks
Fields
Master items
Charts
Custom objects
Search
Dimensions
Measures
Create new
60+
Total 18-25 killed
Total 25-35 killed
Total 35-45 killed
Total 45-60 killed
Total killed Female
Total killed Male
Under-age


Edit measure
Segment colors
Name
Total 18-25 killed
Description
male +female
Measure color
Tags
Glossary terms
Expression
Sum([18-25 Years - Killed - Male])+
Label expression
Number formatting
Number
Formatting
Simple
1,000
Cancel
Save

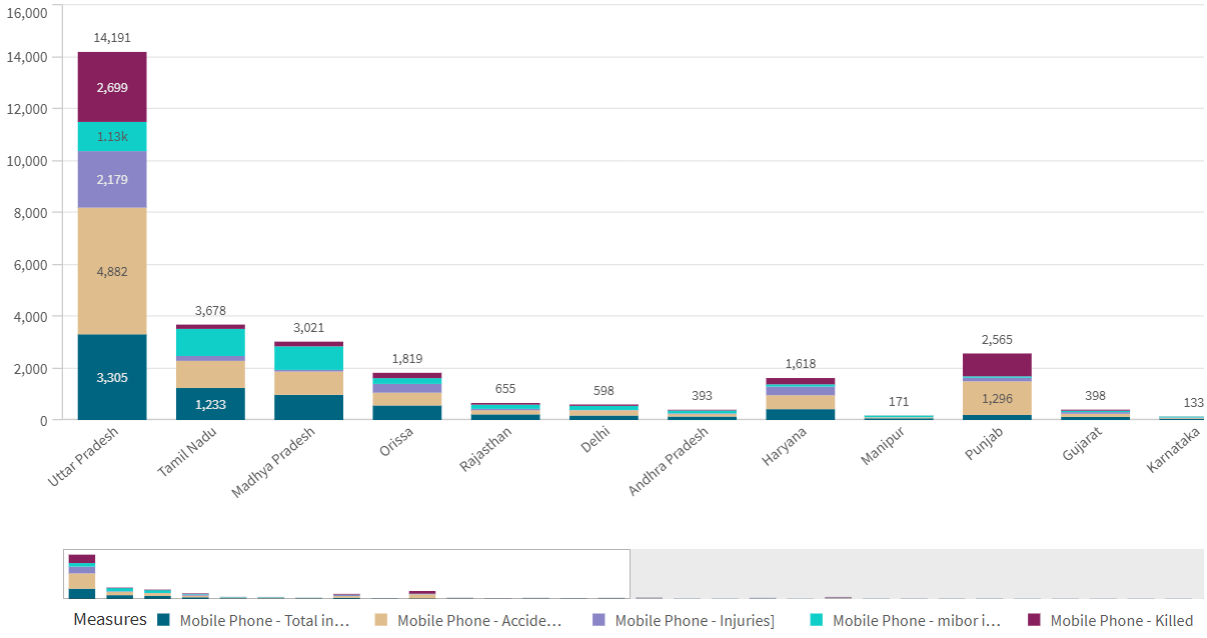

Edit expression
Sum([18-25 Years - Killed - Male])+Sum([18-25 Years - Killed - Female])
OK
Sum([18-25 Years - Killed - Male])+Sum([18-25 Years - Killed - Female])
Fields
Filter by table
All tables
Field
18-25 Years - Female
Aggregation functions
No aggregation
Set Expression
Distinct
Total
Insert
Expression generator
Functions
Set expressions
Variables
Cancel
Apply

Number Of Graphs

- Accidents due to Drunken Driving

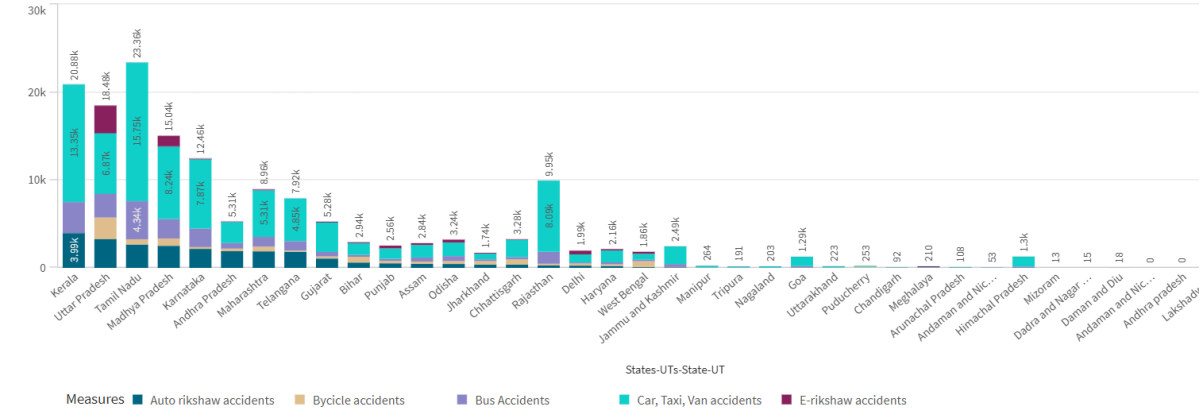


- State-wise Mobile Phone Usage



- Vehicle Contribution towards Total Accidents

4: Vehicle Contribution towards Total Accidents (5 vehicles)



Total Accidents vs total auto

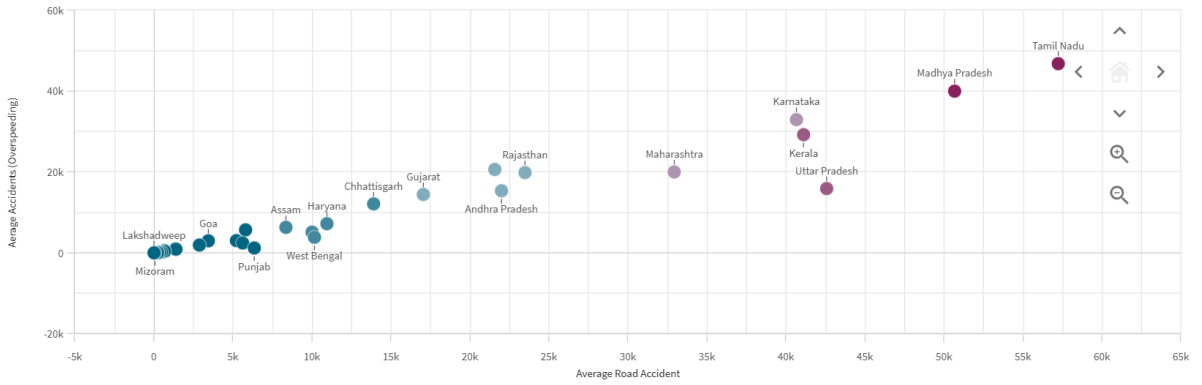
1,312 25.71k

Total truck vs person killed by truck

4,322 13.53k

- Correlation - Speeding and Number of accidents

8: Relation between Average Accidents and Over-speeding

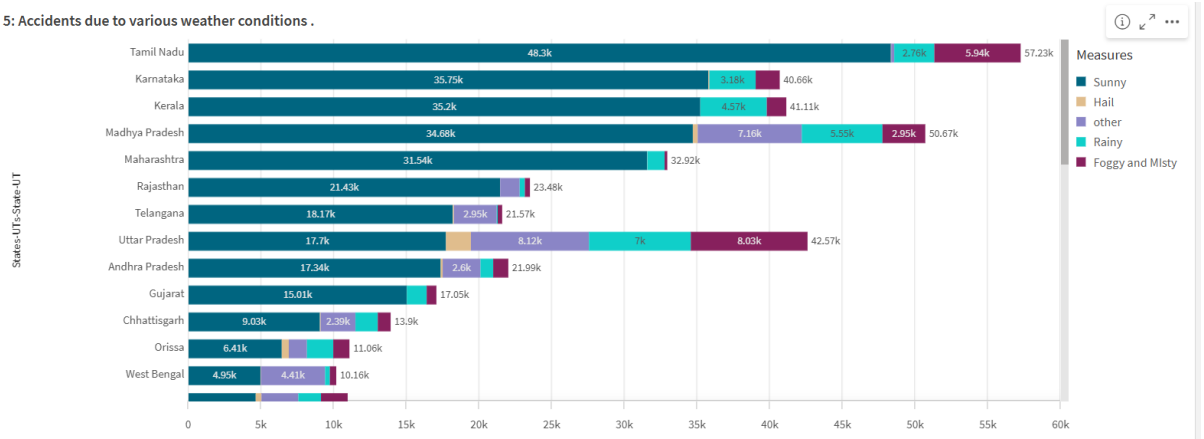


Mutual relation between accidents by overspeeding and total accidents

99.96%

- Accidents by Weather Type

5: Accidents due to various weather conditions .



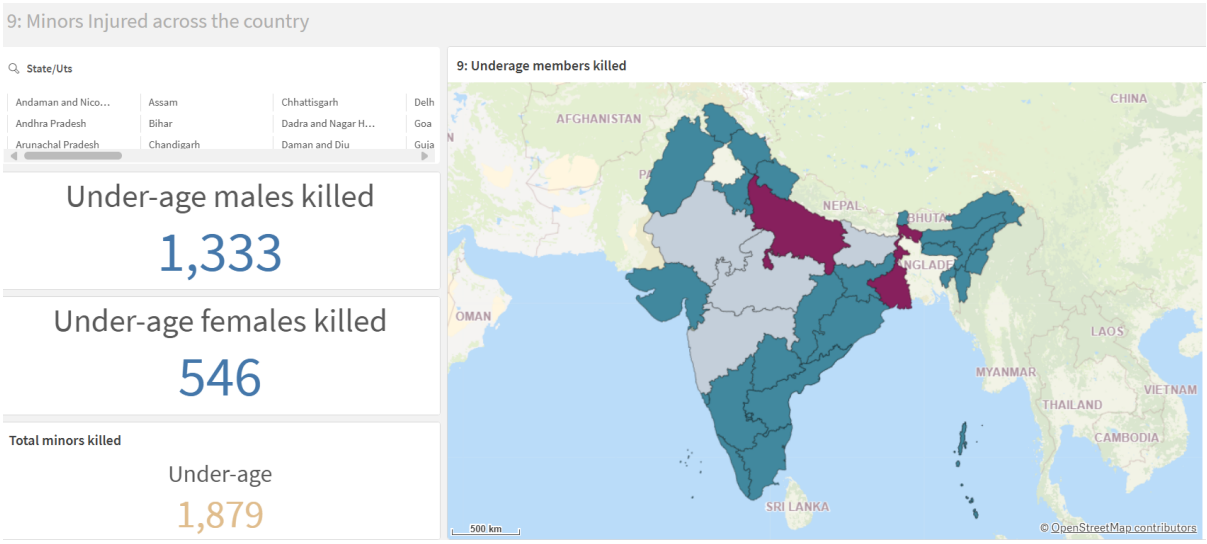
Death due to Foggy/Misty weather

13,405 33,602

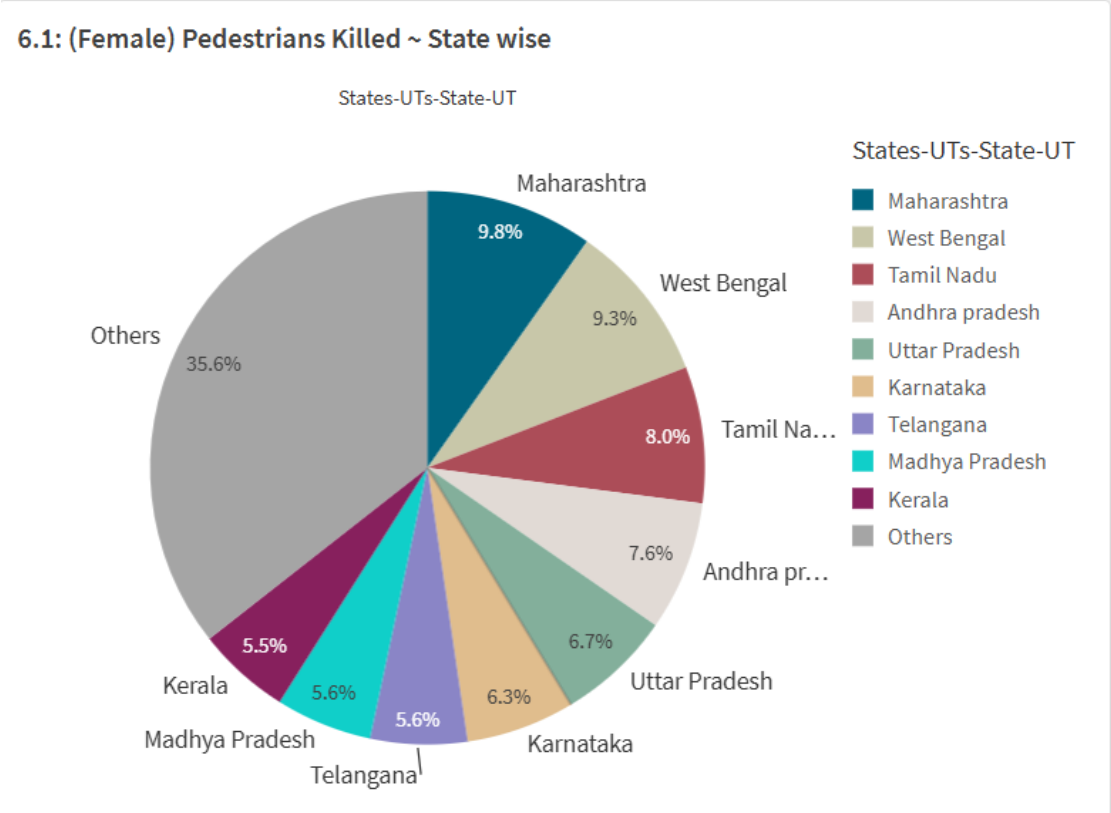
Death count due to Rainy weather

14,240 39,825

- Minors Injured across the country

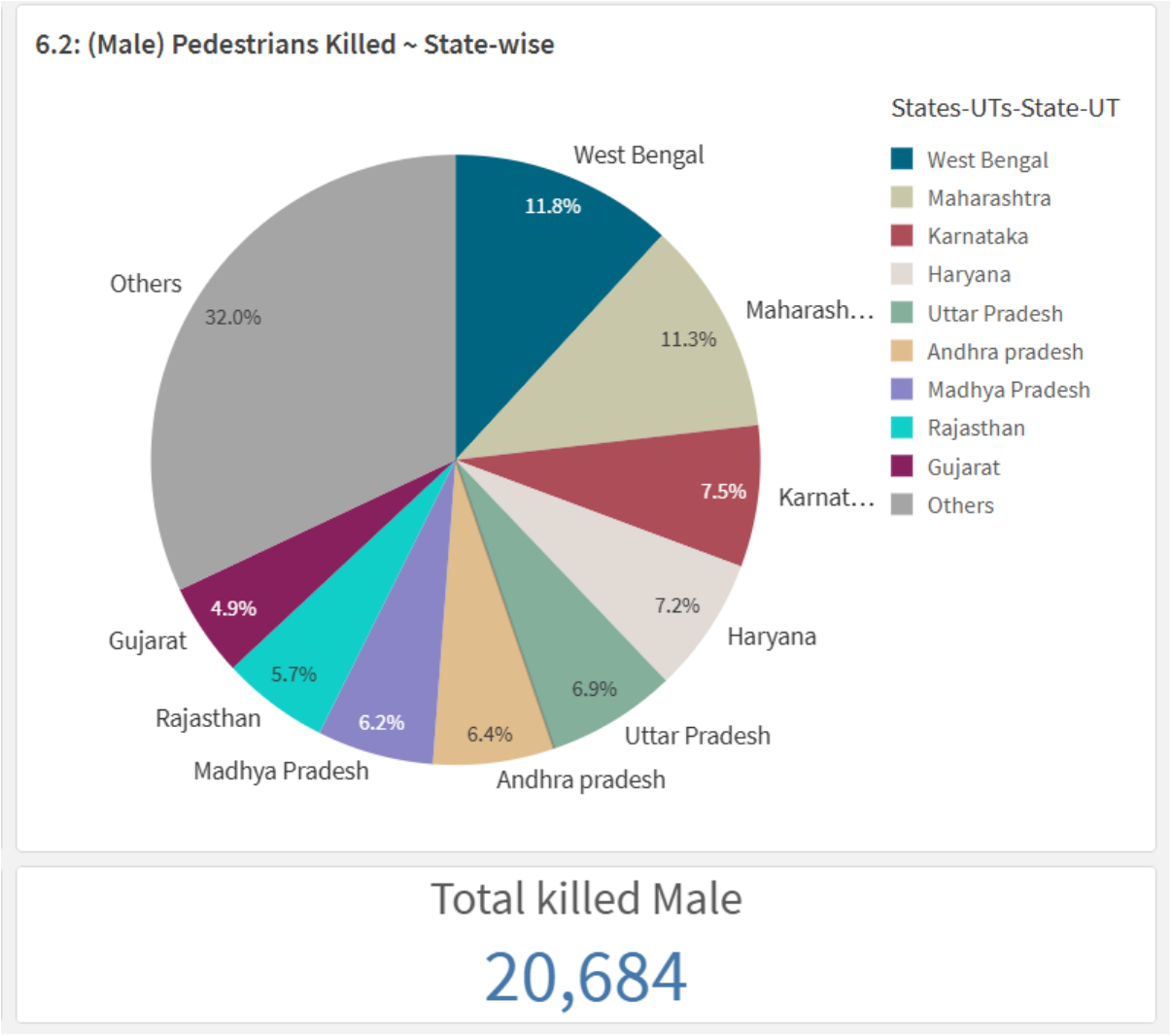


- Pedestrians Killed: Gender

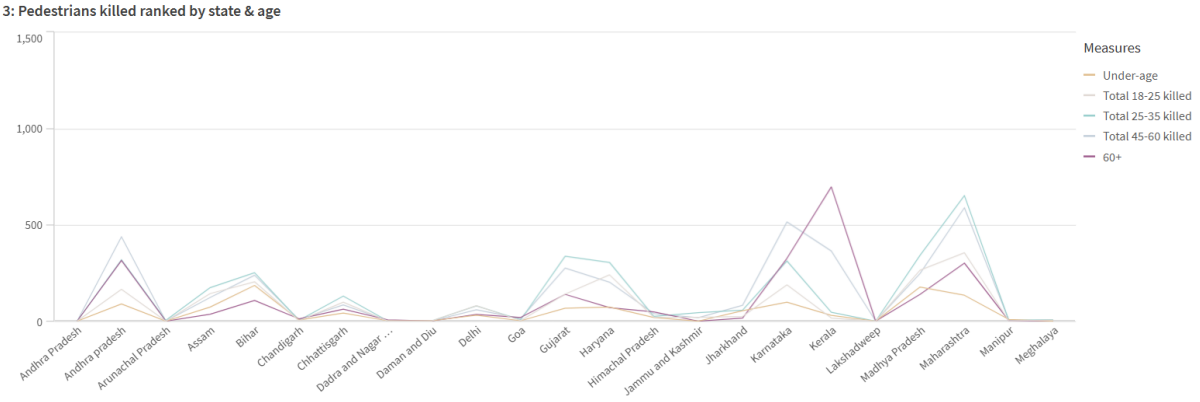


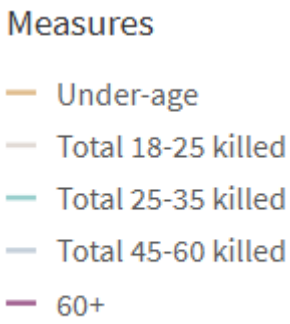
Total killed Female

5,174

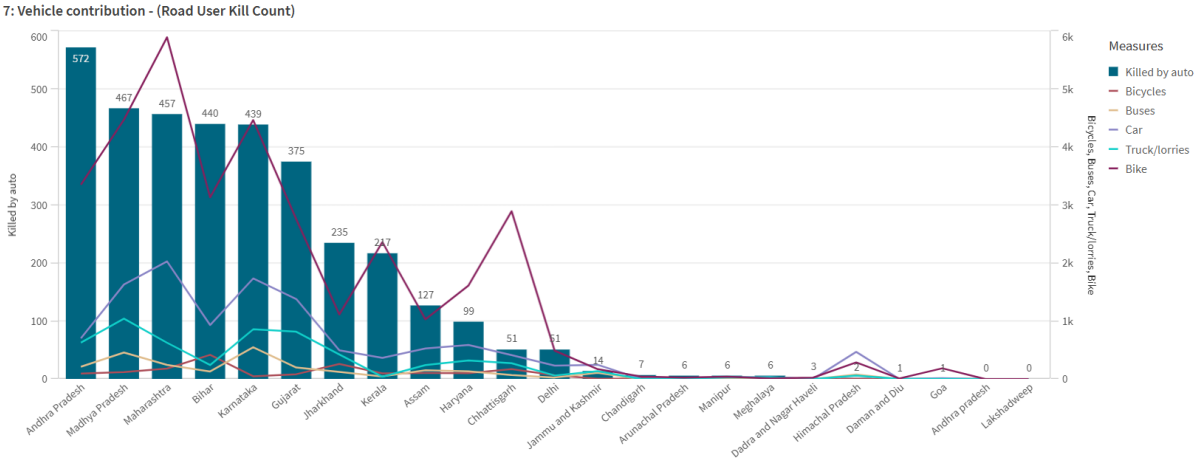


- Pedestrians Killed: Age groups





- Road Users Killed: Vehicle Distribution



Project Demonstration & Documentation

The project demonstration and documentation for the Qlik Road Accident Analysis will provide a comprehensive overview of the initiative's objectives and methodologies. The demonstration will begin with an introduction to the significance of road safety analysis in India, followed by a live walkthrough of the interactive dashboard. Key visualizations, such as heat maps for accident hotspots and age-wise fatality charts, will be showcased, allowing participants to explore the data through user-friendly filters. Major findings will highlight high-risk demographics, the impact of weather conditions, and common accident causes, fostering discussion on their implications for policy. Documentation will include a project summary, data sources, technical architecture, and a step-by-step development process, alongside a user guide for stakeholders. This thorough approach ensures that insights are accessible and actionable, paving the way for informed decision-making and future enhancements in road safety initiatives.

link: https://drive.google.com/file/d/1sSPAcL4u-XoV7bYTriAi245kDCMJlylk/view?usp=sharing

Conclusion

This research project effectively utilized data analytics and visualization tools, particularly Qlik Sense, to examine road safety and accident patterns in India. The insights gained from this study can significantly contribute to developing strategies and policies aimed at improving road safety measures and reducing accidents.
