Qlik Analysis Of Road Safety And Accident Patterns In India

Abstract

Road accidents have become a significant concern in India, resulting in substantial loss of life and injuries. This research project aims to leverage data analytics and visualization techniques to analyze road safety and accident patterns across the country. By utilizing Qlik Sense, a powerful data analytics platform, this study examines various aspects of road accidents, including their types, locations, causes, and the demographics of those involved. The primary objective is to derive actionable insights that can inform strategies and policies to improve road safety and reduce the occurrence of accidents.

1   Introduction

1.1 Overview

This project aims to leverage data analytics to analyze road safety and accident patterns in India. With the increasing number of road accidents causing significant loss of life and injuries, it is crucial to understand the underlying factors contributing to these incidents. By utilizing Qlik Sense, a powerful data analytics and visualization tool, this project will examine various aspects of road accidents, such as their types, locations, causes, and the demographics of those involved. The ultimate goal is to derive actionable insights that can inform strategies to improve road safety and reduce accidents.

1.2 Purpose

Purpose: The primary purpose of this project is to provide valuable insights into road accidents in India, which can be used to enhance road safety measures and policies. The key objectives include:

- **Understanding Accident Patterns**:

- Analyze the distribution and frequency of road accidents across different states and regions.
- Identify patterns related to the time, location, and severity of accidents.

- **Identifying Risk Factors**:
  - Determine the main causes of accidents, such as speeding, drunken driving, or mobile phone usage.
  - Explore correlations between various factors like weather conditions, vehicle types, and accident rates.

- **Demographic Analysis**:
  - Examine the demographics of victims, including age and gender, to identify the most vulnerable groups.
  - Assess the impact of accidents on different types of road users, such as pedestrians, cyclists, and vehicle occupants.

- **Supporting Decision Making**:
  - Provide data-driven insights to policymakers, law enforcement agencies, and public safety officials.
  - Assist in formulating effective strategies and interventions to enhance road safety and reduce accidents.

What Can Be Achieved:

- **Improved Road Safety**: Insights from the analysis can lead to better safety protocols and regulations.
- **Targeted Interventions**: Identifying high-risk areas and causes can help in designing specific interventions.
- **Public Awareness**: Findings can inform public awareness campaigns about road safety.
- **Policy Formulation**: Data-driven evidence can support the development of policies to reduce road accidents.

1.3 Technical Architecture

The technical architecture of this project involves several key components, each playing a vital role in the data analysis and visualization process:

1. Data Sources:

- **Raw Data**: Collect raw data on road accidents from reliable sources like the Kaggle dataset.
- **Data Integration**: Integrate data from multiple sources to ensure comprehensive

analysis.

2. Data Storage:

- **Database**: Store the collected data in a structured format using a database management system (DBMS).
- **Data Cleaning and Preprocessing**: Clean the data to remove inconsistencies, handle missing values, and prepare it for analysis.

3. Data Analysis and Processing:

- **Qlik Sense**: Utilize Qlik Sense for data analysis, applying statistical methods to identify patterns and trends.
- **Calculated Fields and Master Items**: Create calculated fields and master items in Qlik Sense to simplify complex analyses and ensure consistency across visualizations.

4. Visualization and Dashboard Design:

- **Interactive Dashboards**: Design interactive dashboards using Qlik Sense to present the analyzed data in a user-friendly and visually appealing manner.
- **Visualizations**: Employ various types of visualizations such as bar charts, line charts, scatter plots, heat maps, and geographical maps to convey insights effectively.

5. User Interaction:

- **Data Filters**: Implement data filters to allow users to dynamically interact with the data and explore specific aspects of road accidents.
- **Responsive Design**: Ensure the dashboards are responsive and accessible across different devices, providing a seamless user experience.

6. Performance Optimization:

- **Efficient Data Handling**: Optimize data handling and processing to ensure the dashboards perform well, even with large datasets.
- **Performance Testing**: Conduct performance testing to identify and address any bottlenecks or issues.

7. Documentation and Reporting:

- **Project Documentation**: Document the entire project development process, including data collection, preprocessing, analysis, and visualization methodologies.
- **Explanation Videos**: Record videos explaining the project's end-to-end solution, highlighting key findings and insights.

By following this technical architecture, the project aims to provide a robust and comprehensive analysis of road safety and accident patterns in India, offering valuable insights that can drive improvements in road safety measures and policies.

2 Problem Understanding

2.1 Problem Statement

Business Problem: Technological advancements in transportation have minimized distances but increased the risk to life due to road accidents. Every year, accidents result in significant loss of life and serious injuries in India. This project aims to analyze road safety and accident trends using Qlik Sense, examining data on road incidents to identify patterns and inform strategies to improve road safety.

2.2 Business Requirements

Business Requirements:
- **User Demographics Analysis**: Provide insights into the demographics of road accident victims.
- **Accident Patterns Identification**: Analyze patterns in road accidents across different regions and times.
- **Risk Factors Analysis**: Identify and analyze factors contributing to road accidents (e.g., speeding, weather conditions).
- **Visualization**: Create interactive and visually compelling dashboards to support strategic planning and operational improvements.
- **Decision Support**: Generate insights to help policymakers and safety officials make informed decisions and implement better safety protocols.
- **Compliance and Reporting**: Ensure compliance with regulations and facilitate reporting requirements through comprehensive data analysis.

2.3 Literature Survey

Literature Survey: A literature survey for this project involves researching and reviewing previous studies, articles, reports, and data on road safety and accident analysis.

A comprehensive literature review was conducted to gain a thorough understanding of the existing research in the field of road safety and accident analysis. Previous studies have explored various aspects of road accidents, including the role of human factors, vehicle conditions, road infrastructure, and environmental conditions [2, 3]. Several analytical techniques, such as regression analysis, spatial analysis, and data mining, have been employed to identify patterns and risk factors associated with road accidents [4, 5]. However, the literature review revealed a lack of comprehensive studies focused

specifically on India, considering the unique challenges and diverse geographical and socio-economic landscape of the country. Additionally, the integration of advanced data analytics and visualization tools, such as Qlik Sense, in the context of road safety analysis in India has been relatively unexplored.

By conducting a thorough literature survey, the project aims to build on existing knowledge and apply proven techniques to analyze road safety and accident patterns in India effectively. This will help in deriving actionable insights to improve road safety measures and reduce accidents.

3   Methodology

3.1 Data Collection

The first step in this research project involved collecting comprehensive data on road accidents in India from reliable sources. The data was obtained from government reports, datasets from transportation departments, and academic research studies. The dataset includes variables such as accident types, locations, causes, demographics of victims, vehicle types, and environmental conditions.

3.2 Data Preparation and Integration

The collected data underwent a rigorous cleaning and preprocessing process to ensure its accuracy and consistency. Missing values and outliers were handled appropriately, and the data was transformed into a format suitable for visualization and analysis. Qlik Sense's powerful data integration capabilities were utilized to connect and integrate data from multiple sources, facilitating a comprehensive analysis.

3.3 Data Analysis and Visualization

Qlik Sense's advanced analytical and visualization features were employed to analyze the road accident data. Statistical methods were applied to identify patterns, trends, and correlations between various factors contributing to road accidents. Calculated fields and master items were created within Qlik Sense to simplify complex analyses and ensure consistency across visualizations. Interactive dashboards were designed to present the analyzed data in a user-friendly and visually compelling manner. Various types of visualizations, such as bar charts, line charts, scatter plots, heat maps, and geographical maps, were utilized to convey insights effectively. Data filters were implemented to allow users to dynamically interact with the data and explore specific aspects of road accidents.

4 Data Collection

4.1 Collect the Data

Data Collection Process:
- Obtain comprehensive data on road accidents in India from reliable sources.
- Sources may include government reports, datasets from transportation departments, and research studies.
- Ensure the dataset includes variables such as accident types, locations, causes, and demographics of victims.

4.2 Connect the Data with Qlik Sense

Connecting Data with Qlik Sense:
- Import the collected dataset into Qlik Sense for analysis.
- Utilize Qlik Sense's data loading capabilities to connect to various data sources.
- Ensure proper data modeling to structure the data for analysis.
- Implement data transformations and cleaning processes as necessary to prepare the data for visualization.

Benefits of Connecting with Qlik Sense:
- Qlik Sense provides powerful data visualization and analytics capabilities.
- Connecting the data allows for real-time analysis and dynamic visualizations.
- Qlik Sense's intuitive interface facilitates data exploration and insights generation.

Data Integration Steps:
1. **Data Import**: Import the collected dataset into Qlik Sense.
2. **Data Connection**: Establish a connection between the dataset and Qlik Sense.
3. **Data Modeling**: Structure the data model within Qlik Sense to facilitate analysis.
4. **Data Transformation**: Clean and preprocess the data to ensure accuracy and consistency.
5. **Data Visualization**: Create visualizations and dashboards to present insights derived from the data.

By connecting the collected data with Qlik Sense, the project can leverage the platform's capabilities to analyze road safety and accident patterns effectively. This will enable the generation of actionable insights to improve road safety measures and reduce accidents in India.

5   Data Preparation

5.1 Prepare the Data for Visualization

Data Preparation Process:
- Clean the data to remove any inconsistencies or errors.
- Handle missing values and outliers appropriately.
- Transform the data into a format suitable for visualization.
- Perform any necessary aggregations or calculations to derive meaningful insights.
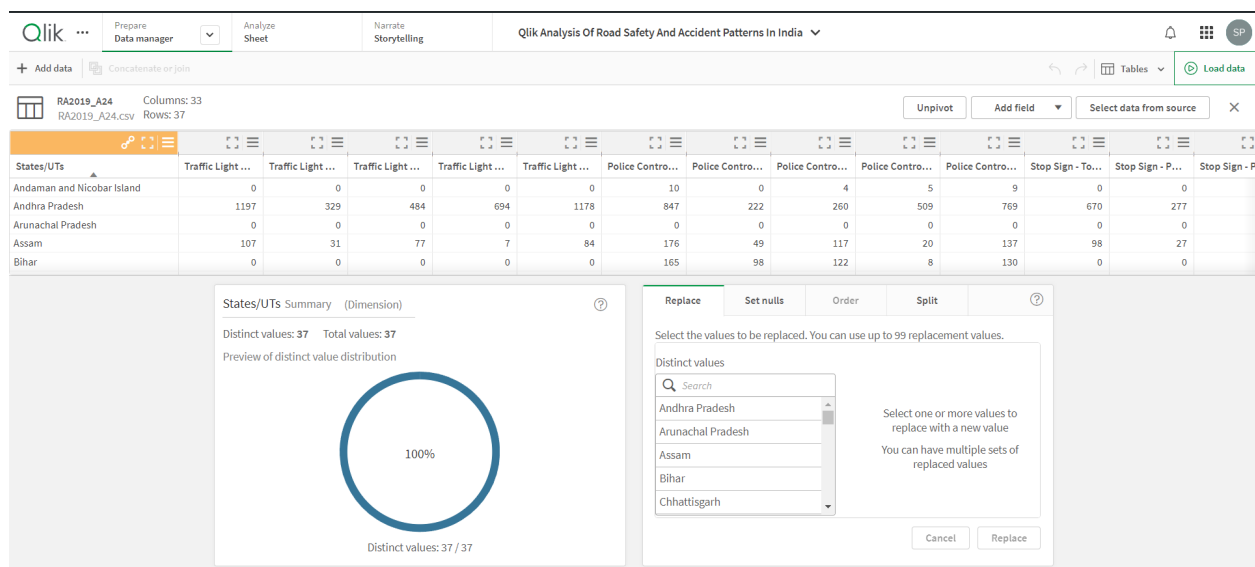- Ensure the data is structured and organized for easy visualization.

Figure:Dataset having zero null sets

6 Data Visualization

6.1 Visualization

Visualization Techniques:
- Utilize various visualization types such as bar charts, line charts, scatter plots, and maps to represent different aspects of the data.
- Choose visualization techniques that effectively communicate insights and trends in the data.
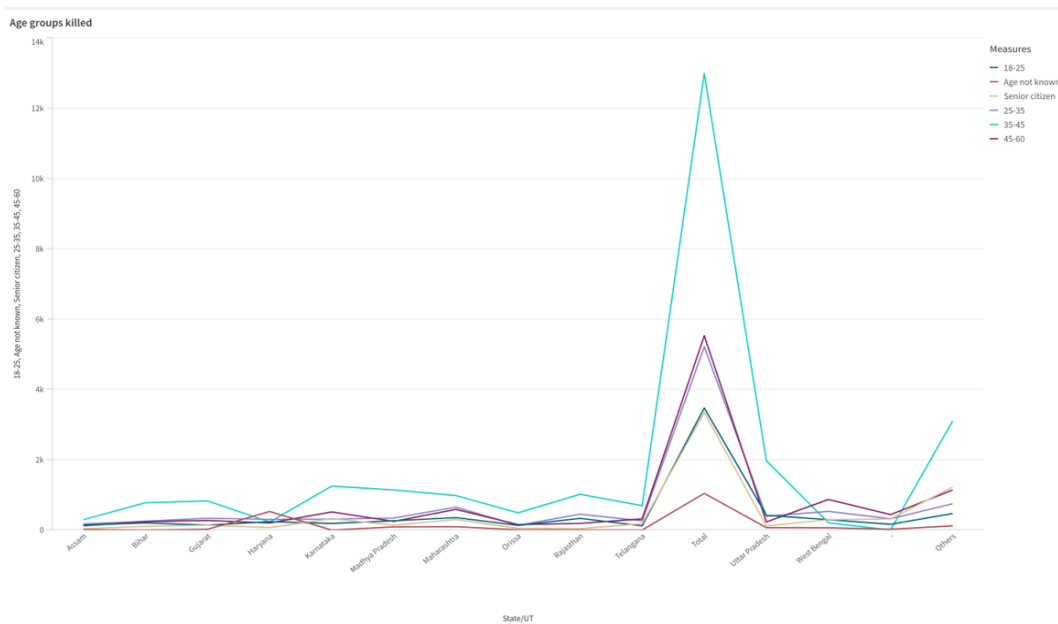- Incorporate interactivity to allow users to explore the data dynamically.

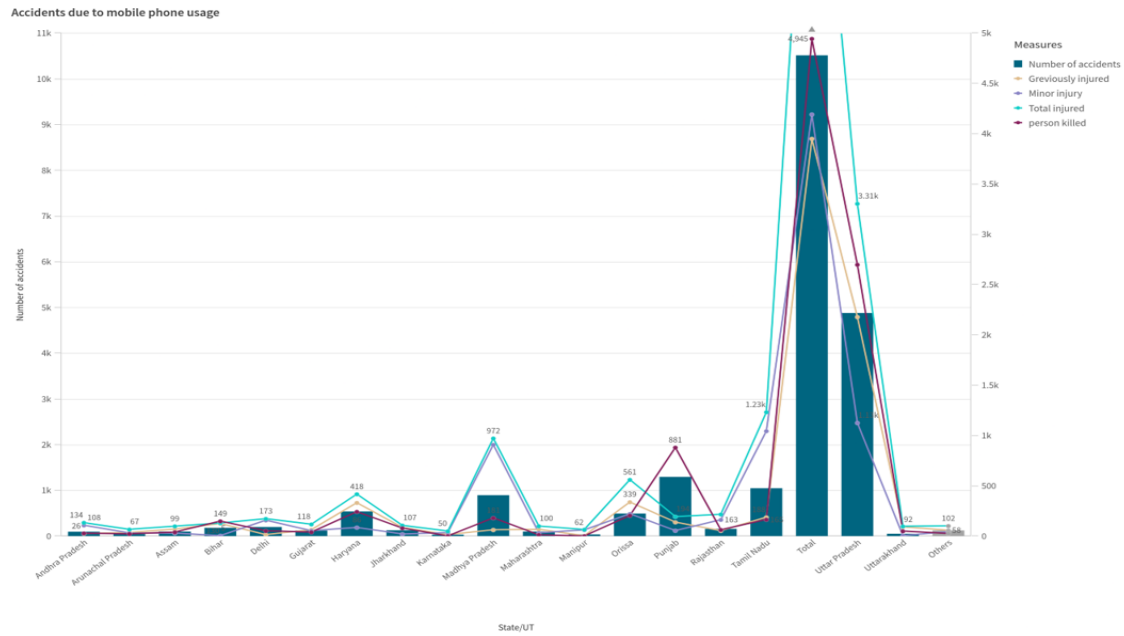*Figure: Different Age Group Killed in Accident*

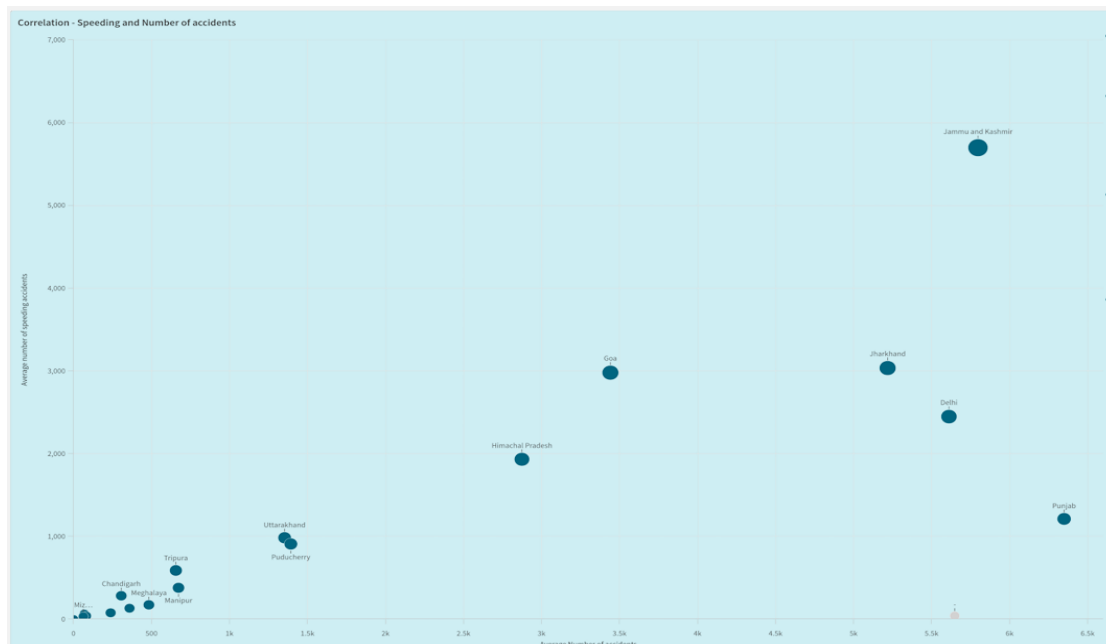*Figure:Accident due to mobile phone usages(Combine Graph)*

Figure:Speeding and Number of accident(Scatter graph)

7   Dashboard

7.1 Responsive and Design of Dashboard

Dashboard Design Principles:
- Design a visually appealing and intuitive dashboard layout.
- Ensure responsiveness across different devices and screen sizes.
- Organize visualizations logically to facilitate easy comprehension.
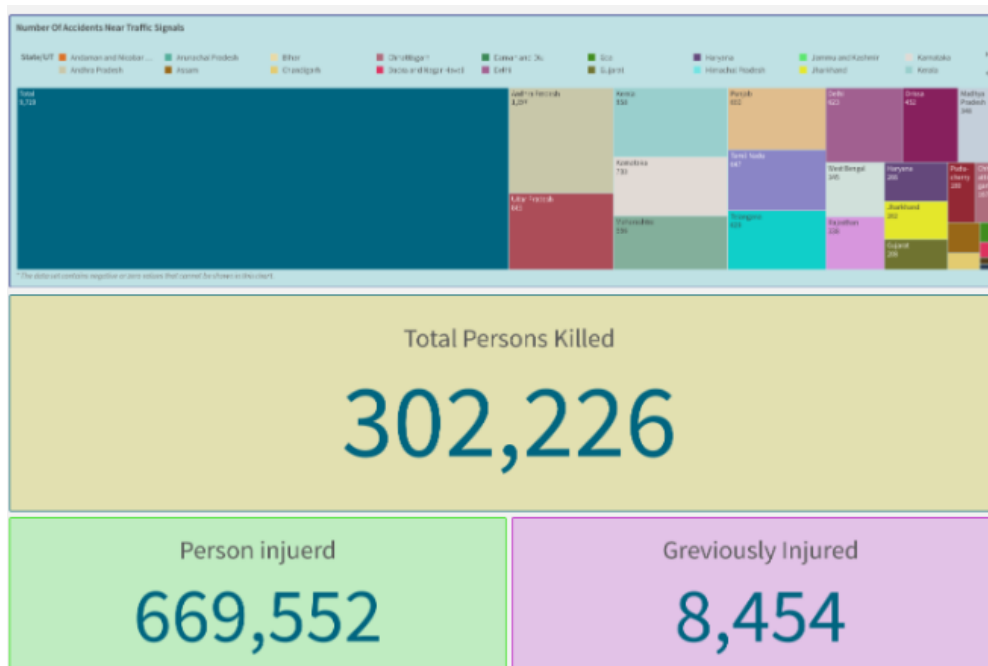- Incorporate user-friendly navigation and interaction elements.

Figure:Number Of Accident near Traffic Signals

8 Report

8.1 Report Creation

Report Components:
- Summarize key findings and insights derived from the data analysis.
- Provide visual representations of data trends and patterns.
- Include explanations and interpretations of the visualizations.
- Offer recommendations based on the analysis to improve road safety measures.

9 Performance Testing

9.18.1 Amount of Data Renders

Performance Testing Metrics:

- Measure the time taken to render different amounts of data in the dashboard.
- Assess the impact of data volume on dashboard performance.
- Optimize data rendering processes to ensure smooth and efficient performance.

9.2Utilization of Data Filters

Data Filter Utilization:

- Monitor the usage of data filters by users to understand their impact on dashboard performance.
- Evaluate the effectiveness of data filters in facilitating data exploration and analysis.
- Optimize data filter functionality to enhance user experience and performance.

10 Results and Discussion

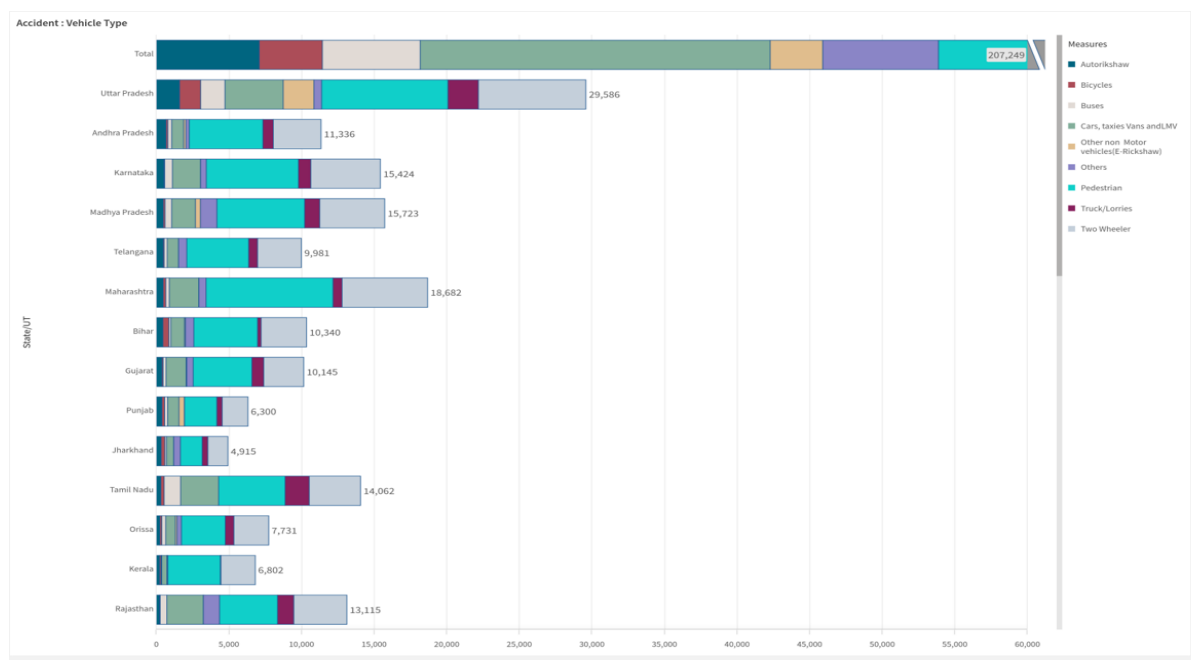The analysis of road accident data in India using Qlik Sense yielded several valuable insights and patterns.

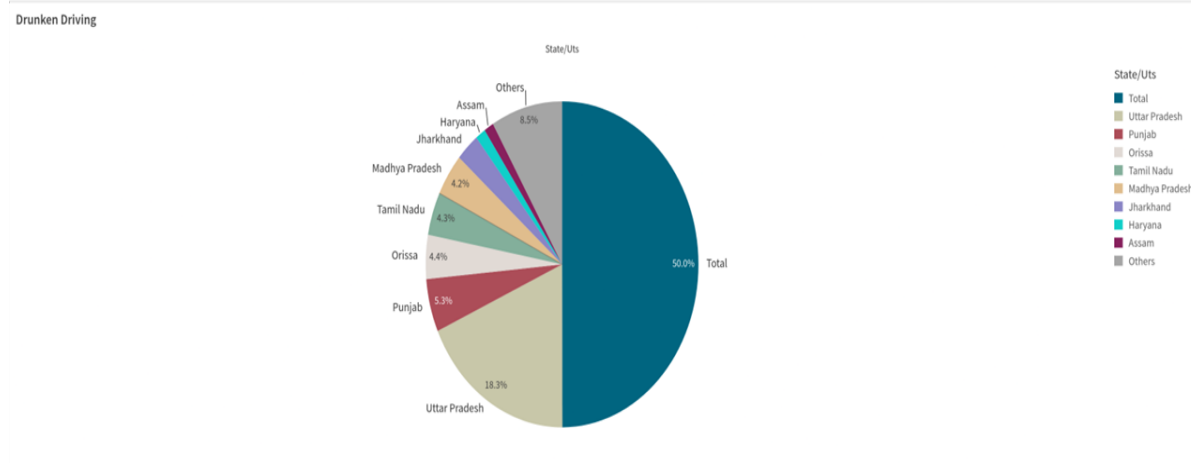*Figure: Accident caused due to various types of Vehicles*

Figure: Accident caused due to various types of Vehicles

11 Conclusion

This research project successfully leveraged the power of data analytics and visualization techniques, particularly Qlik Sense, to analyze road safety and accident patterns in India. The insights derived from this study can be instrumental in informing strategies and policies to enhance road safety measures and reduce the occurrence of accidents. By identifying high-risk areas, vulnerable groups, and contributing factors, targeted interventions and public awareness campaigns can be designed and implemented. Future research should focus on integrating additional data sources, such as real-time traffic data and weather information, to further enhance the analysis and provide more comprehensive insights. Additionally, exploring the application of advanced machine learning techniques for predictive modeling and risk assessment could potentially yield valuable contributions to the field of road safety.

12 References

[1] Ministry of Road Transport and Highways, Government of India. (2020). Road Accidents in India - 2019. Retrieved from [URL].

[2] Abdel-Aty, M., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. Accident Analysis & Prevention, 32(5), 633-642.

[3] Bener, A., Hussain, S. J., Lajunen, T., & Crundall, D. (2008). Road traffic crashes in the United Arab Emirates: Trends of morbidity and mortality during 1977–2006. Traffic Injury Prevention, 9(5), 461-467.

[4] Leden, L., Gårder, P., & Pulkkinen, U. (2000). An expert judgment model applied to estimating the safety effect of a bicycle facility. Accident Analysis & Prevention, 32(4), 589-599.

[5] Mitra, S., & Bhandari, S. (2019). A data mining approach to analyze road accident time series data. Journal of Big Data, 6(1), 1-19.