QLIK ANALYSIS OF ROAD SAFETY AND ACCIDENT PATTERNS IN INDIA

ABSTRACT :

In India, road accidents have morphed into a critical public health concern, claiming countless lives and inflicting severe injuries. This research project strategically deploys data analytics and visualization techniques to illuminate the complexities of road safety and unveil recurring accident patterns across the nation. By leveraging the prowess of Qlik Sense, a powerful data analytics platform, the study meticulously examines various dimensions of road accidents. This includes a granular exploration of accident typologies, geographical distribution, causative factors, and the demographics of individuals impacted. The paramount objective is to extract actionable insights that can inform the formulation of data-driven strategies and policies. These interventions have the potential to significantly bolster road safety and mitigate the occurrence of accidents, ultimately safeguarding lives.

1 INTRODUCTION

1.1 OVERVIEW: Delving into the Epidemic of Road Accidents in India

India confronts a burgeoning crisis on its roads – a relentless surge in road accidents that translates into a devastating loss of life and life-altering injuries. To combat this escalating public health threat, this project undertakes a meticulous investigation into road safety and accident patterns across the nation. Leveraging the analytical prowess of Qlik Sense, a sophisticated data analytics and visualization platform, the project will dissect various facets of these accidents. This comprehensive examination will encompass the typologies of accidents, their geographical distribution, the root causes that precipitate them, and the demographics of the individuals impacted. The ultimate objective is to glean actionable insights that can illuminate the path towards data-driven strategies and interventions. These interventions have the potential to fortify road safety measures and significantly reduce the number of accidents, ultimately safeguarding lives on the roads of India.

1.2 PURPOSE: Unveiling Actionable Insights to Fortify Road Safety

The paramount objective of this project is to extract critical knowledge from the vast repository of road accident data in India. These gleaned insights will then be strategically utilized to formulate effective measures and policies that can significantly bolster road safety across the nation. Here's a comprehensive breakdown of the key goals:

Pattern Recognition: Illuminating Temporal and Spatial Trends- A meticulous analysis will be conducted to identify the distribution and frequency of accidents across various Indian states

and regions. This analysis will shed light on trends related to the time of day, location (urban vs. rural), and the severity of accidents. By unveiling these patterns, the project aims to illuminate the temporal and spatial landscapes of road safety challenges in India.

Risk Factor Exploration: Discerning the Precipitating Factors- The project will embark on a mission to pinpoint the leading contributors to accidents, such as speeding, driving under the influence of alcohol or drugs, or using mobile phones while driving. Additionally, it will explore potential correlations between various factors, including weather conditions, vehicle types (cars, motorcycles, trucks, etc.), and accident rates. By meticulously dissecting these risk factors, the project aims to illuminate the root causes of accidents and identify areas for targeted intervention.

Demographic Analysis: Identifying Vulnerable Populations- An in-depth investigation will be conducted into the demographics of accident victims, encompassing factors like age and gender. This analysis will be instrumental in identifying the most vulnerable groups on the road, such as young drivers, pedestrians, or cyclists. Furthermore, the project will assess the impact of accidents on different road users to gain a comprehensive understanding of the human cost associated with these incidents.

Data-Driven Decision Support: Empowering Policymakers and Law Enforcement- The project aims to equip policymakers, law enforcement agencies, and public safety officials with data-driven insights gleaned from the analysis. These insights can then be strategically utilized to formulate effective interventions and data-driven strategies that can bolster road safety and minimize the occurrence of accidents. By empowering these stakeholders with knowledge, the project strives to create a safer road environment for all.

Potential Outcomes:
Enhanced Road Safety: By gleaning valuable insights from the analysis, this project can pave the way for the development of improved safety protocols and regulations.
Targeted Interventions: Identifying high-risk areas and accident causes can aid in designing targeted interventions that address specific challenges.
Public Awareness Campaigns: The project's findings can inform public awareness campaigns designed to educate the public on road safety practices.
Data-Driven Policy Formulation: The project's robust data-driven evidence can support the development of effective policies aimed at reducing road accidents.

1.3 TECHNICAL ARCHITECTURE: The Cornerstones of Data Exploration
The technical architecture of this project is comprised of several crucial components that work in tandem to facilitate data analysis and visualization:

1. Data Acquisition: Sourcing the Fuel for Analysis
- **Raw Data Collection:** The project commences by gathering raw data on road accidents from trustworthy sources, potentially including a Kaggle dataset.
- **Data Integration:** If data is available from multiple sources, it will be meticulously integrated to ensure a comprehensive analysis.

2. Data Storage: Building a Secure Repository
- **Database Management:** The collected data will be housed in a structured format using a robust Database Management System (DBMS).
- **Data Cleaning and Preprocessing:** The data will undergo a meticulous cleaning process to eliminate inconsistencies, address missing values, and prepare it for insightful analysis.

3. Data Analysis and Processing: Unveiling Patterns
- **Qlik Sense as the Analytical Powerhouse:** Qlik Sense, a powerful data analytics platform, will be employed to scrutinize the data. Statistical methods will be applied to identify patterns and trends within the data.
- **Calculated Fields and Master Items:** Building Blocks for Efficiency: Calculated fields and master items will be created within Qlik Sense to streamline complex analyses and ensure consistency across visualizations.

4. Visualization and Dashboard Design: Transforming Data into Insights
- **Interactive Dashboards:** User-friendly and visually appealing interactive dashboards will be designed using Qlik Sense to present the analyzed data in a clear and compelling manner.
- **Tailored Visualizations:** A variety of visualization techniques, including bar charts, line charts, scatter plots, heat maps, and geographical maps, will be employed to effectively communicate the project's findings.

5. User Interaction: Empowering Exploration
- **Dynamic Data Exploration:** Data filters will be implemented to empower users to dynamically interact with the data and delve into specific aspects of road accidents.
- **Responsive Design for Accessibility:** The dashboards will be designed to be responsive and accessible across various devices, ensuring a seamless user experience.

6. Performance Optimization: Ensuring Smooth Operation
- **Efficient Data Handling:** Data handling and processing will be optimized to guarantee that the dashboards function flawlessly, even when dealing with large datasets.
- **Performance Testing:** Identifying and Addressing Bottlenecks: Performance testing will be conducted to pinpoint and address any potential bottlenecks or issues that might hinder performance.

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.

2 DEFINE PROBLEM / PROBLEM UNDERSTANDING

2.1 SPECIFY BUSINESS PROBLEM
Technological advancement in transportation has minimised the distances but has increased the risk to life. Every year, accidents result in loss of lakhs of lives and serious injuries to crores of people. A study to analyse road safety and accident trends in India is to be conducted using Qlik Sense, a data analytics platform. This study involves examining data related to road incidents, such as types of accidents, locations, causes, and potentially factors contributing to road safety or risks. The use of Qlik Sense is a data-driven approach, utilizing visualizations and insights generated from the analysis to understand patterns and potentially inform strategies for improving road safety in India.

Problem Statement: Road transport is an integral part of our lives. Everybody is a road user in one shape or form. The current transport system has minimized the distances but has increased the risk to life. In India, about eighty thousand people are killed in road crashes every year which is thirteen percent of the total fatality all over the world. The analysis aims to provide valuable insights into user demographics, accident patterns, and problem areas.

Objectives: The objective of this project is to conduct a thorough analysis of accidents data, with the aim of extracting valuable insights to enhance decision making, policy formulation, and operational efficiency within our country.

Social Impact:
- Compare the severity of accidents in different areas of traffic control.
- Identify the leading causes of accidents.

Business Model/Impact : Analysis of road accident data from 2019 can aid insurance companies, vehicle manufacturers, and transportation authorities in developing strategies to mitigate risks, improve safety features, and enhance road infrastructure.

2.2 BUSINESS REQUIREMENTS
This project prioritizes several key business requirements to effectively combat the challenge of road safety in India:
The analysis aims to provide valuable insights into user demographics, accident patterns, and

problem areas. The primary focus is on creating interactive and visually compelling dashboards to support strategic planning and operational improvements. The insights derived from this analysis will be instrumental in making informed decisions, implementing better safety protocols, and ensuring compliance with regulations.

2.3 LITERATURE SURVEY

A literature survey for the Road Safety and Accident Patterns analysis would involve researching and reviewing previous studies, articles, reports and figures on the topic. This could include information on the methods and techniques used for analysing accidents data, as well as the results and conclusions of these studies. It is recommended to explore academic databases such as PubMed, IEEE Xplore, Google Scholar, and institutional repositories. Additionally, government reports and publications can provide insights into the latest developments.

3 DATA COLLECTION

Data collection is the process of gathering and measuring information on variables of interest,in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

3.1 COLLECT THE DATA

To analyze road safety and accident patterns in India effectively, a comprehensive dataset is required. This dataset should include:

Accident Data: Number, types, and locations of accidents from sources like MoRTH, traffic police records, and state transport departments.
Vehicle Information: Details on type, model, age, and condition of vehicles involved.
Weather Conditions: Data from meteorological departments to assess weather impacts on accidents.
Road Conditions: Information on road surface type, maintenance status, and presence of road signs.
Demographic Data: Age, gender, and occupation of accident victims.
Time and Date: Timestamp data for analyzing patterns by time of day, week, and seasons.
Cause of Accidents: Primary causes such as speeding, drunk driving, or mechanical failures from police reports.

Steps for Data Collection:
Identify Sources: Government agencies, transportation departments, police records, and meteorological data.
Data Acquisition: Collect data through official requests, open data portals, and collaborations.
Data Validation: Ensure accuracy by cross-referencing with multiple sources.
Data Cleaning: Remove duplicates, correct errors, and handle missing values.

Data Integration: Combine data from different sources into a cohesive and consistent dataset.

3.2 CONNECT THE DATA WITH QLIK SENSE

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

Data Import: Import the cleaned dataset using Qlik Sense's data load editor, supporting formats like CSV, Excel, and database connections.

Data Modeling: Define relationships between data tables by setting primary and foreign keys.

Data Transformation: Refine the dataset with calculated fields, aggregations, and structural transformations.

Data Visualization: Design interactive dashboards using bar charts, line charts, maps, and scatter plots to display patterns and trends.

Interactivity and Filtering: Implement filters, drill-downs, and selections for dynamic data exploration.

Data Analysis: Conduct trend analysis, correlation analysis, and predictive modeling to identify key accident factors and preventive measures.

Collaboration and Reporting: Share dashboards and reports with stakeholders for informed decision-making.

These steps utilize Qlik Sense's capabilities to effectively analyze and visualize road safety and accident patterns in India, aiding in understanding and mitigating road safety issues.

4 DATA PREPARATION

4.1 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.
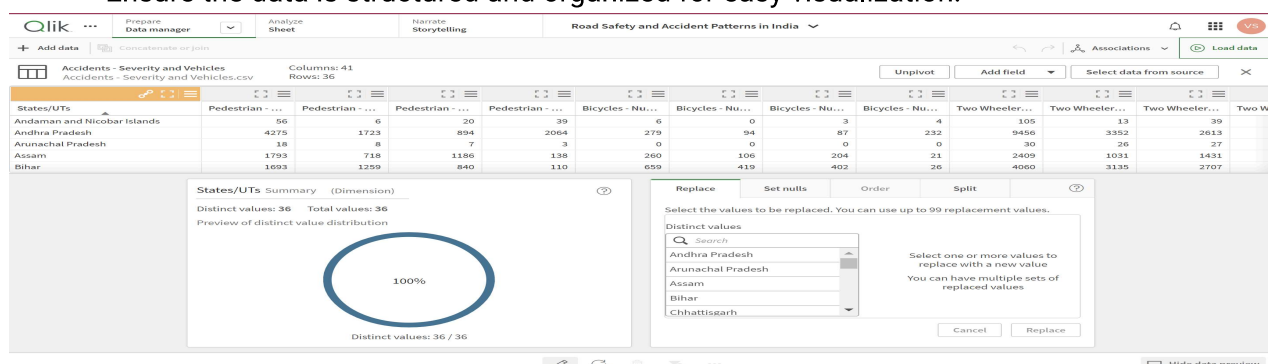

Fig 4.1: Data Preparation in Qlik Sense

5 DATA VISUALISATIONS

5.1 VISUALISATIONS
Step-by-Step Implementation:
- **Data Import:** Load the dataset into Qlik Sense.
- **Create Data Model:** Establish relationships between data tables.
- **Design Dashboards:** Utilize the various visualization techniques to create comprehensive dashboards.

  Accident Overview Dashboard: Use bar charts, pie charts, and KPIs to provide a high-level overview of accident statistics.

  Trends and Patterns Dashboard: Employ line charts and histograms to show temporal trends and distributions.

  Geographical Analysis Dashboard: Use maps and scatter plots to visualize accident hotspots and correlations.

  Detailed Analysis Dashboard: Include box plots, Sankey diagrams, and interactive filters for in-depth exploration of specific variables.
- **Enhance Interactivity:** Add filters and drill-down options to enable dynamic data exploration.
- **Collaborate and Share:** Publish the dashboards and share with stakeholders for insights and decision-making.

By employing these visualization techniques, the project will effectively analyze and present the road safety and accident patterns in India, facilitating a better understanding of the data and supporting informed decision-making.
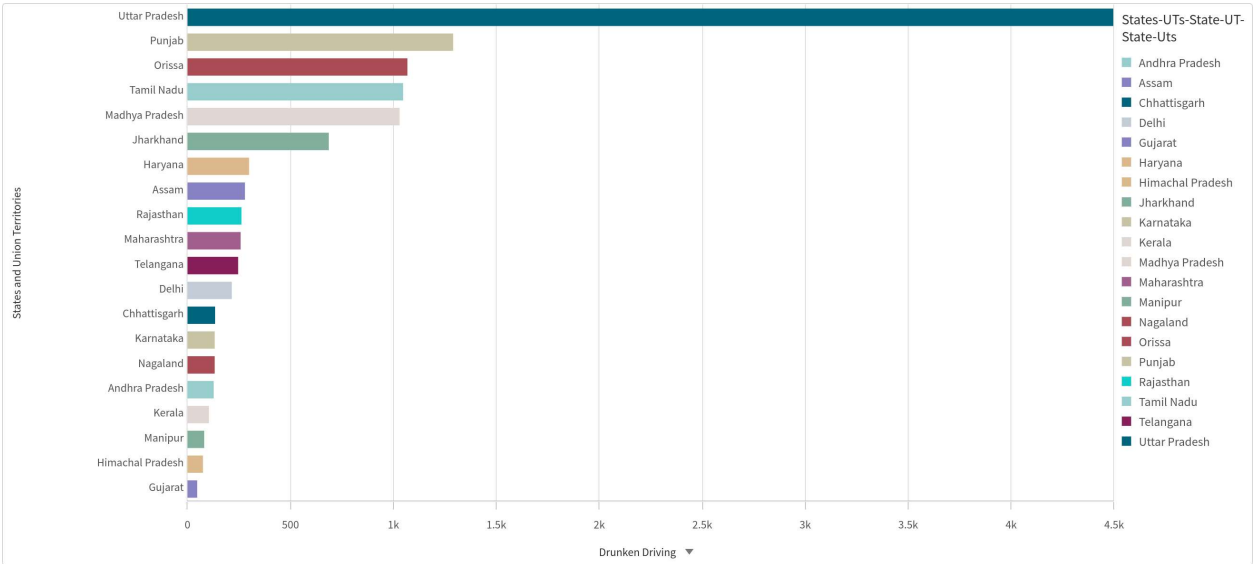
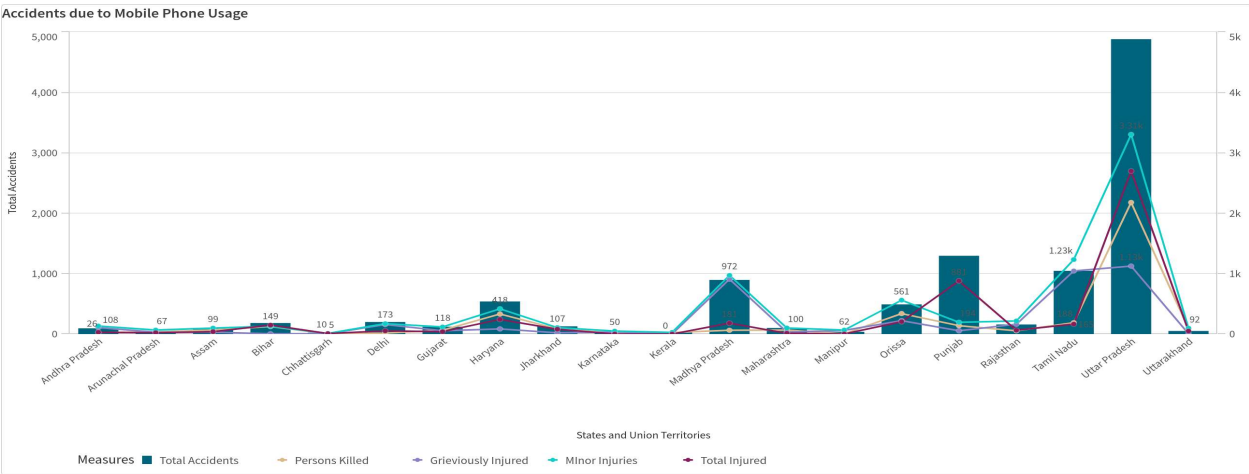Fig 5.1: Causes of Accidents in Drunken Driving
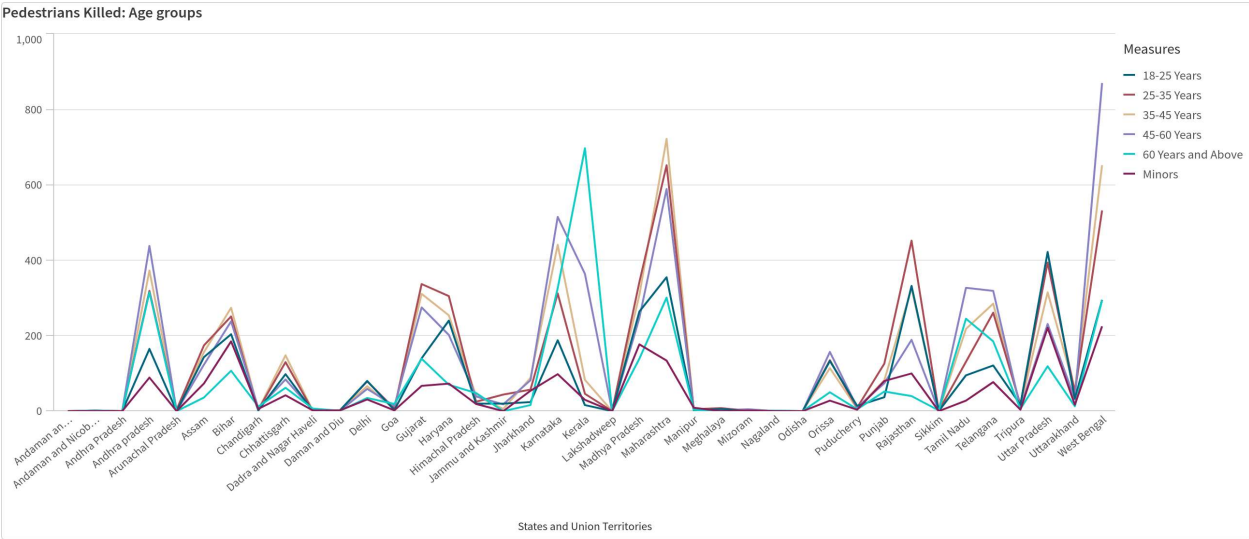
Fig 5.2: Accidents due to Mobile Phone Usage

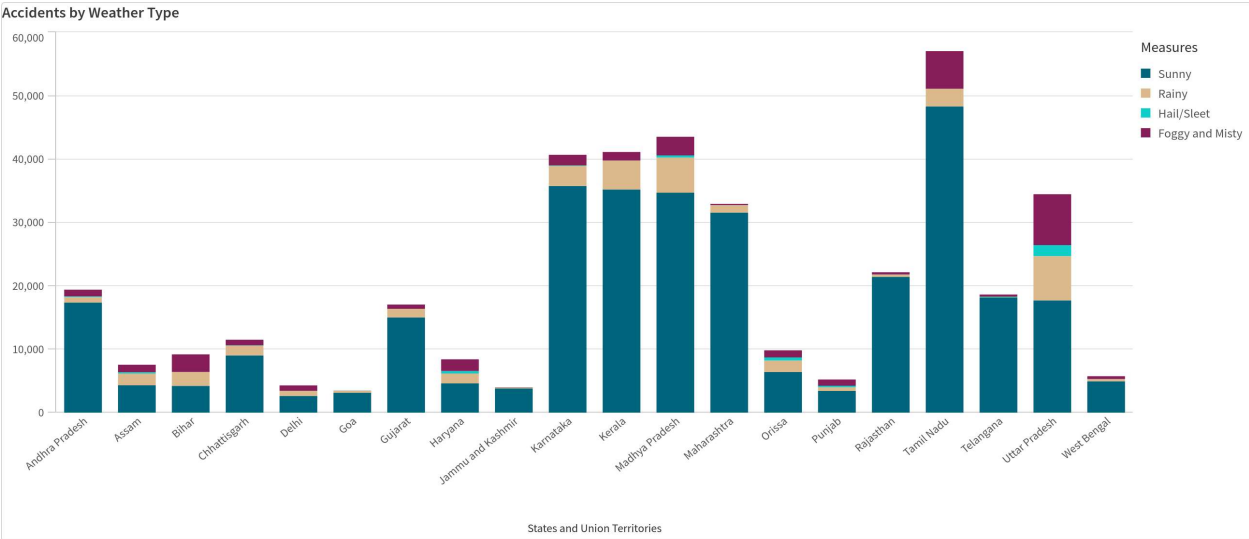Fig 5.3: Pedestrians Killed: Age groups

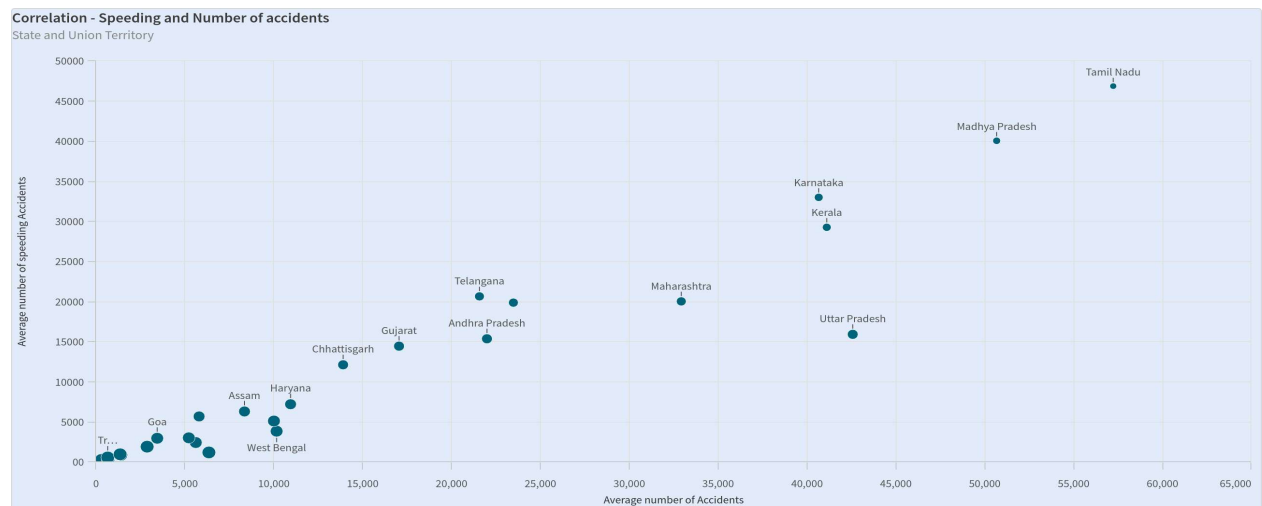

Fig 5.4: Accidents by Weather Type

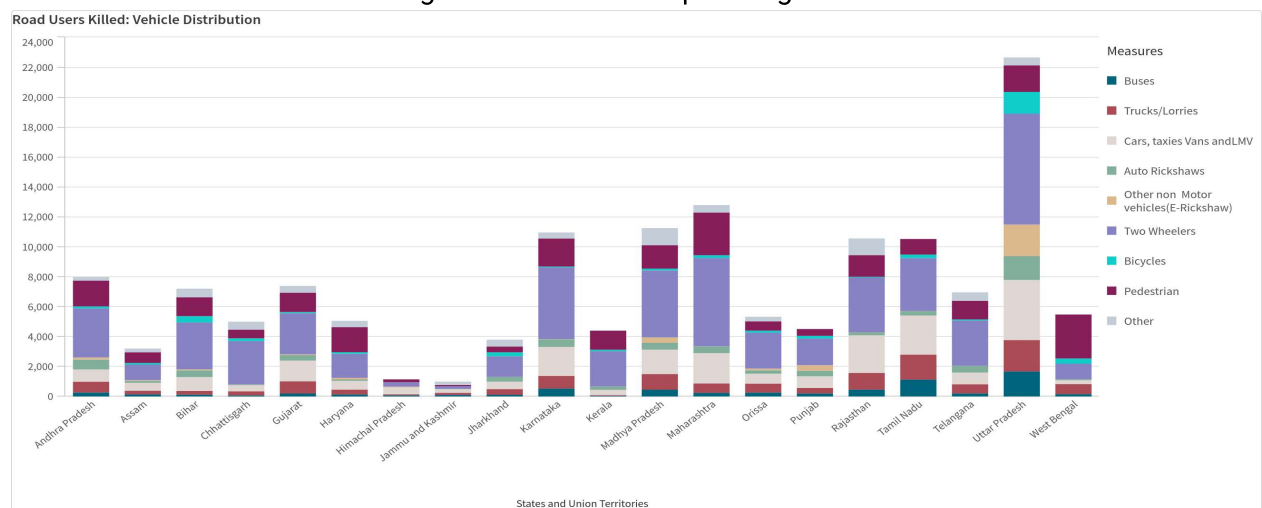Fig 5.5: Correlation - Speediing and Number of Accidents

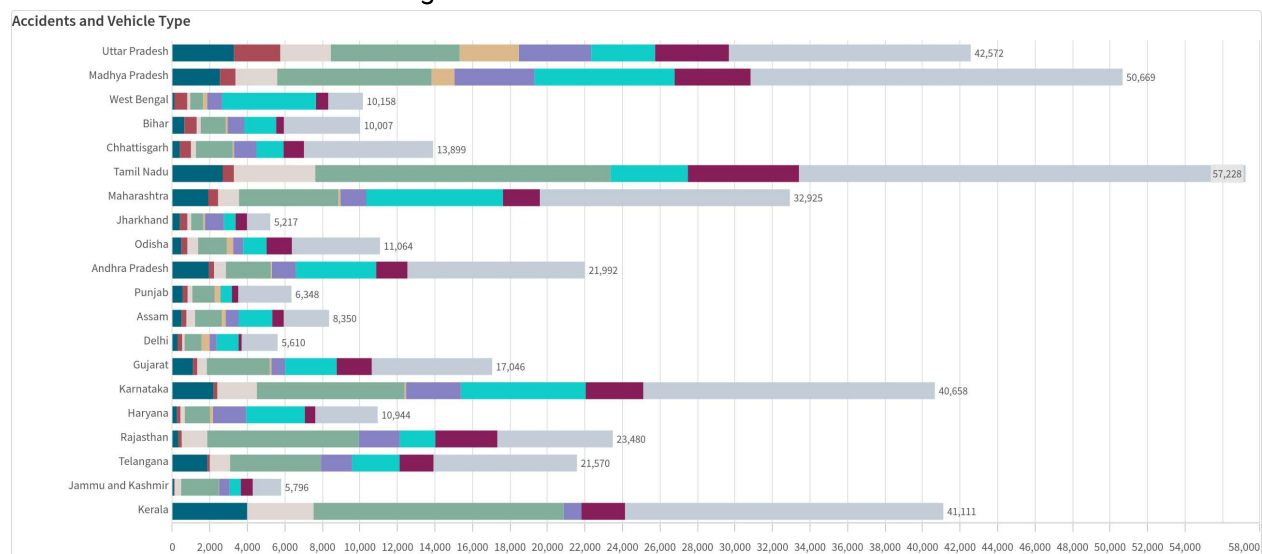Fig 5.6: Road Users Killed: Vehicle Distribution
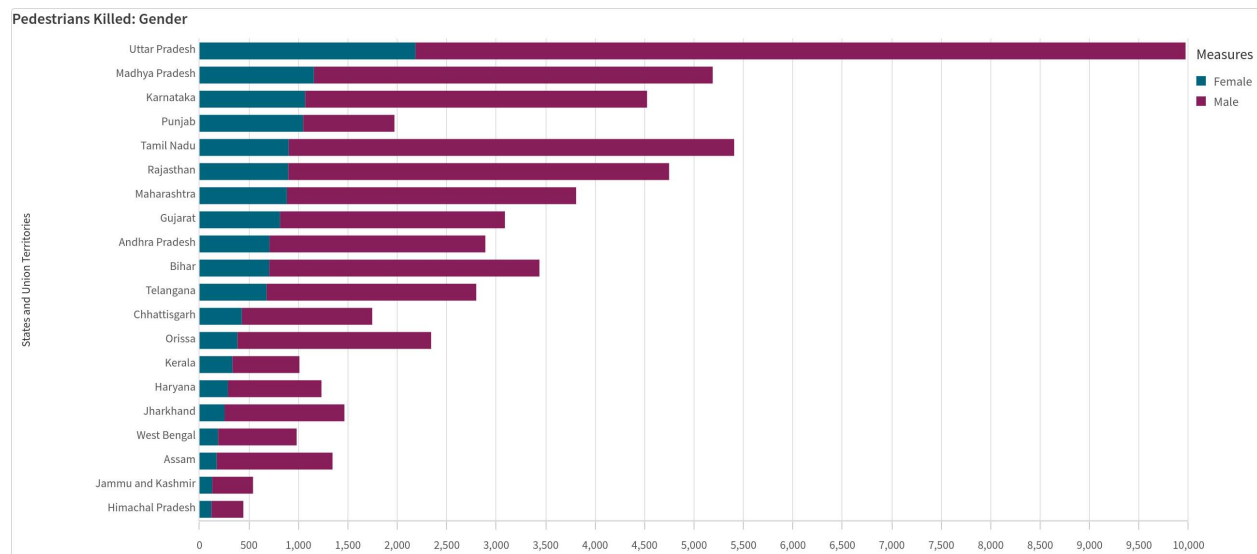

Fig 5.7: Accidents and Vehicle Type
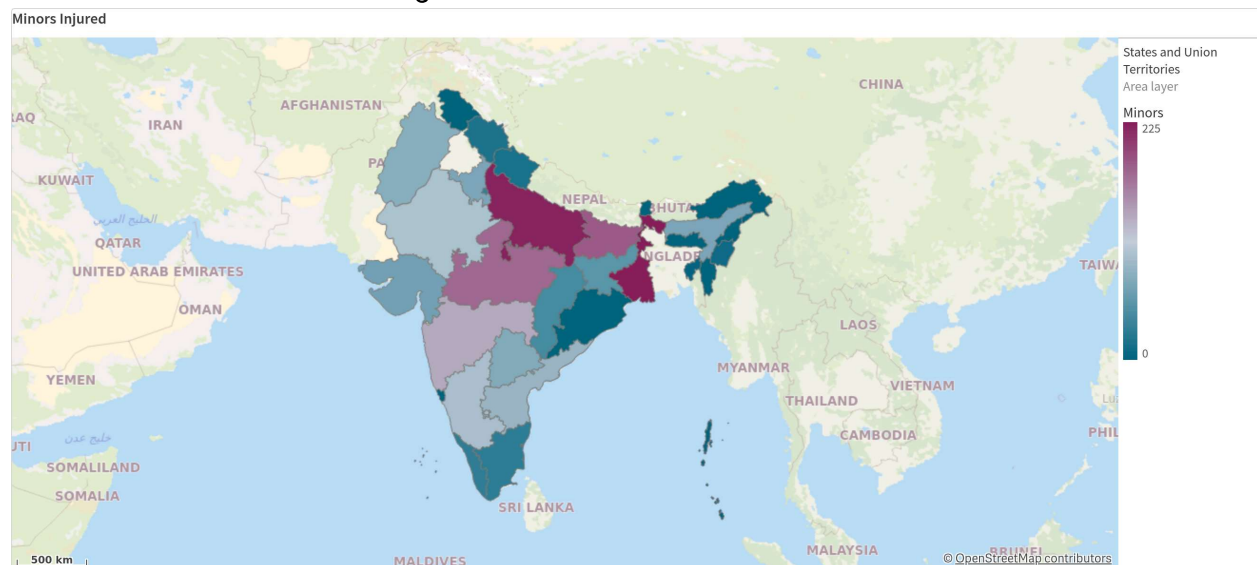
Fig 5.8: Pedestrians Killed: Gender

Fig 5.9: Minors Injured

6 DASHBOARD

A dashboard is a graphical user interface (GUI) that displays information and data in an organized and easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data. They are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

6.1 RESPONSIVE AND DESIGN OF DASHBOARD

Key Considerations:
- **User-Centric Design:** Cater to user needs with relevant metrics and visualizations.
- **Mobile/Desktop Compatibility:** Ensure accessibility on all devices with a responsive layout.
- **Intuitive Navigation:** Use clear navigation, consistent layouts, colors, and icons.
- **Performance Optimization:** Optimize for fast loading and smooth interactions.

Steps for Creating a Responsive Dashboard:
- **Plan the Layout:** Use a grid layout and organize with sections/tabs.
- **Design Visualizations:** Ensure consistent sizing, readable fonts, and a cohesive color scheme.
- **Add Interactive Elements:** Implement filters, selections, and drill-downs.
- **Test Responsiveness:** Ensure functionality on various devices and browsers.
- **Optimize Performance:** Use efficient data loading, aggregation, and dynamic data loading.
- **User Feedback and Iteration:** Collect feedback and continuously improve the dashboard.
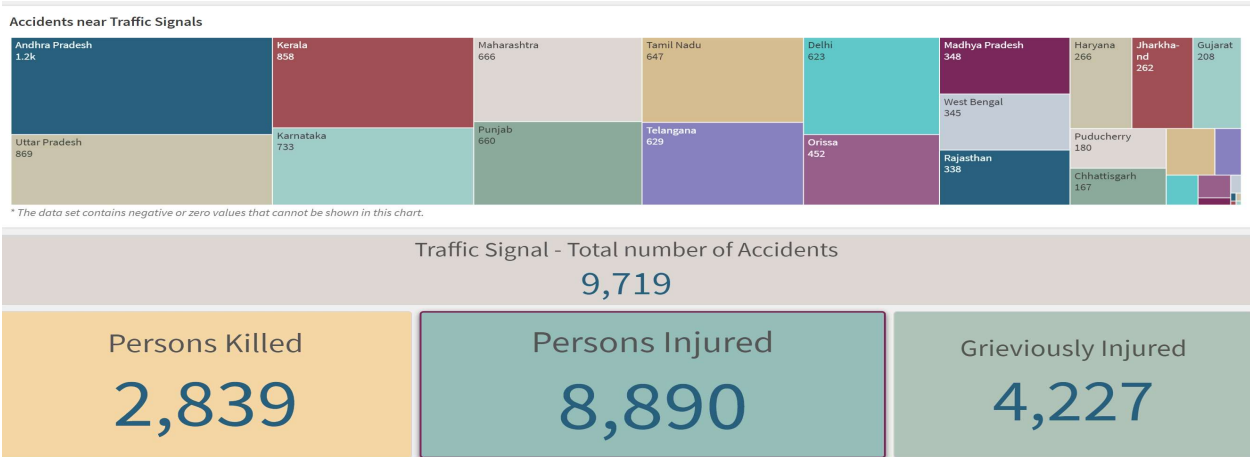

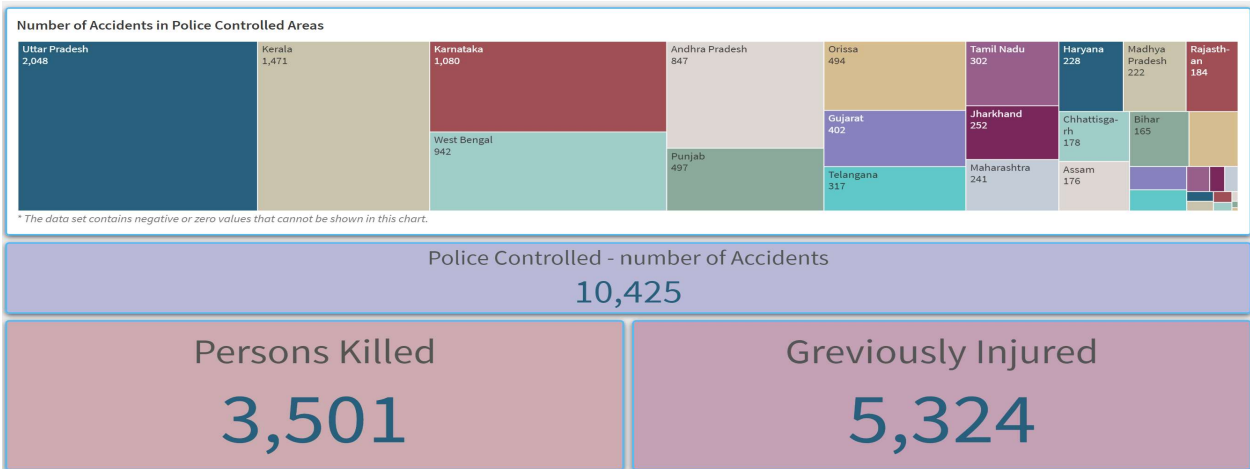Fig 6.1: Accidents near Traffic Signals

Fig 6.2: Number of Accidents in Police Controller Area

7 REPORT

7.1 REPORT CREATION
Clear Structure: Organize the report with a title page, table of contents, executive summary, introduction, and conclusion.
Data Collection: Detail sources and methods for collecting accident, vehicle, weather, and road condition data.
Data Preparation: Explain steps for cleaning and integrating data for analysis.
Effective Visualization: Use various visualizations (bar charts, maps) to highlight key findings and trends.
User-Friendly Dashboard: Design a responsive, intuitive dashboard with interactive elements like filters and drill-downs.
Comprehensive Analysis: Present findings on trends, causes, and demographics.
Actionable Recommendations: Offer clear, evidence-based recommendations for improving road safety.
Review and References: Proofread, incorporate feedback, and list all data sources and references.

8 PERFORMANCE TESTING

8.1 AMOUNT OF DATA RENDERED
Data Volume Management: Pre-aggregate data to reduce processing loads.
Dynamic Loading: Load data dynamically based on user selections and filters.
Performance Optimization: Use Qlik Sense's data load editor for efficient data loading and fast rendering times.
Regular Testing: Test the dashboard with varying data volumes to identify and resolve performance issues.
Scalability: Ensure the dashboard can scale to handle increasing data volumes without compromising user experience.

8.2 UTILIZATION OF DATA FILTERS
Targeted Analysis: Isolate specific demographics (age, gender) or accident types to understand patterns affecting vulnerable populations or pinpoint recurring incidents.
Geographic Focus: Filter by region, state, or even specific roads to identify high-risk areas and pinpoint locations demanding heightened safety measures.
Root Cause Exploration: Focus on accidents influenced by speeding, adverse weather, or vehicle malfunctions to illuminate the factors contributing to these incidents.
Severity Assessment: Analyze trends by filtering accidents based on severity (fatal, injury, property damage), informing resource allocation and intervention strategies.
Comparative Insights: Compare accident patterns across demographics, locations, or

timeframes to uncover disparities and tailor interventions accordingly.

APPLICATION OF DATA FILTERS

Segmentation and Comparison: Data filters in Qlik Sense segment data by time, geography, or product categories to compare trends and identify outliers.
Drill-Down Analysis: Qlik Sense enables drill-down from yearly to daily sales with filters.
Performance Optimization: Qlik Sense ensures fast query response with efficient data filters.
Interactive Dashboards: Filters in Qlik Sense enhance dashboard interactivity for dynamic data views.
User-Driven Exploration: Qlik Sense empowers users with swift ad-hoc analysis using filters.

Selections within the data allows users to filter data based on individual fields or dimensions. Users 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.
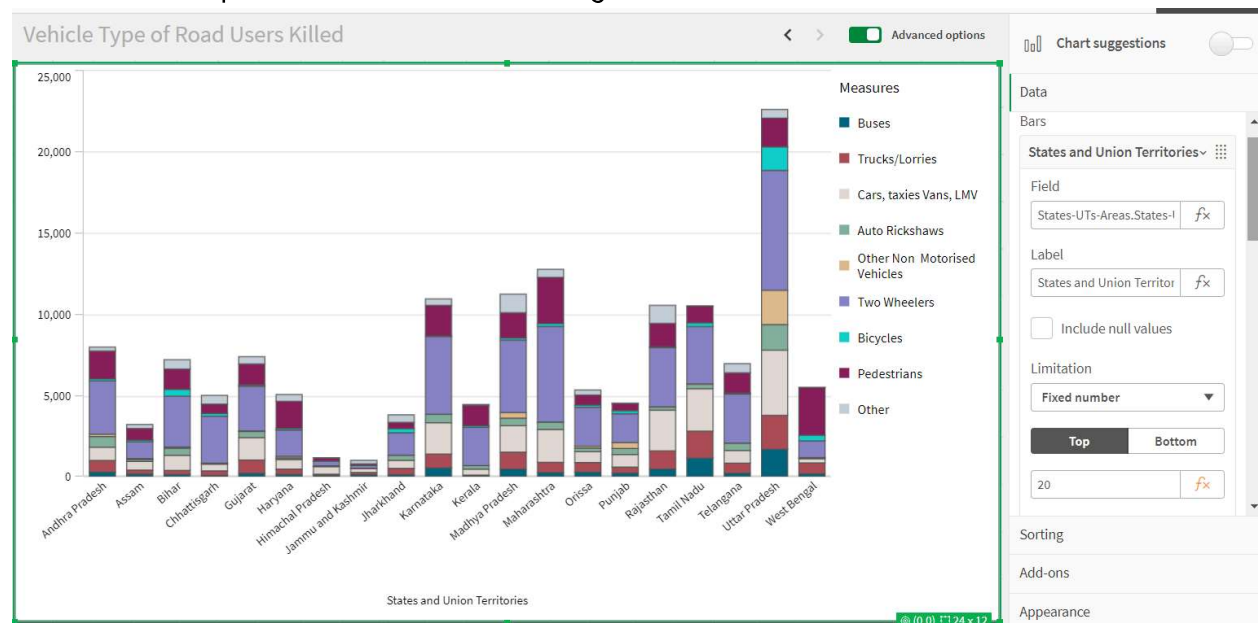

Fig 8.1: Vehicle Type of Road Users Killed

9 CONCLUSION

the Qlik analysis of road safety and accident patterns in India has provided valuable insights into the factors influencing traffic incidents across different regions and time periods. Utilizing Qlik Sense's robust data filtering capabilities, we were able to segment and compare accident data by various dimensions such as location, severity, and vehicle type, uncovering notable trends and patterns. The interactive dashboards facilitated user-driven exploration, enabling stakeholders to dynamically adjust data views and derive actionable insights swiftly. Performance optimization ensured that data filters operated efficiently, supporting fast query responses and enhancing overall analytical efficiency. This comprehensive analysis underscores

the importance of data-driven approaches in improving road safety strategies and decision-making processes in India.

10 REFERENCES

[1] https://morth.nic.in/road-accident-in-india

[2] https://www.validworth.com/road-safety-and-accident-prevention-strategies/

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

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

[5] 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.