Qlik Analysis of Road Safety and Accident Patterns in India

Project Report

Submitted by

Vinitha G

Third Year

B.E. Computer Science Engineering

Thanthai Periyar Goverment Institute of Technology.

1. INTRODUCTION

1.1.OVERVIEW

The project aims to utilize Qlik's data analytics platform to analyze road safety and accident patterns in India. By leveraging various data sources such as traffic data, accident reports, weather conditions, road infrastructure details, and demographic information, the project seeks to identify trends, hotspots, and factors contributing to road accidents. This analysis can help stakeholders, including government authorities, transportation agencies, and road safety organizations, make data-driven decisions to improve road safety measures, reduce accidents, and save lives.

1.2.PURPOSE

1 : Hotspot Identification

Qlik's analytics can pinpoint regions or specific roads in India with a high frequency of accidents. By correlating accident data with factors like traffic volume, road conditions, and time of day, the platform can identify hotspots prone to accidents. This information is crucial for implementing targeted interventions such as enhanced traffic monitoring, improved signage, and speed limit adjustments.

2 : Trend Analysis

Qlik can perform trend analysis on historical accident data to identify patterns and recurring factors leading to accidents. This includes analyzing accident types (e.g., collisions, pedestrian accidents), seasonal variations, and driver behavior (e.g., speeding, distracted driving). Insights gained can guide awareness campaigns, driver training programs, and policy reforms aimed at addressing root causes.

3 : Predictive Modeling

Using predictive analytics, Qlik can forecast potential accident scenarios based on real-time data inputs. By considering variables like weather forecasts, traffic flow patterns, and historical accident trends, the platform can provide early warnings and proactive measures to prevent accidents. This predictive capability empowers authorities to deploy resources strategically and implement preemptive safety measures.

1.3. TECHNICAL ARCHITECTURE

1. Data Collection

- Collect the dataset
- Connect Data with Qlik Sense

2 Data Preparation

- Prepare the Data for Visualization

3. Data Visualizations

- Number of Unique Visualizations

4. Dashboard

- Responsive and Design of Dashboard

5. Story

- Story Creation

6. Performance Testing

- Amount of Data Rendered to DB
- Utilization of Data Filters
- Number of Calculation Fields/Master Items
- Number of Visualizations/Graphs

2.DEFINE PROBLEM / PROBLEM UNDERSTANDING

2.1 SPECIFY THE 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.

2.2 BUSINESS REQUIREMENTS

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

Road Safety and Accident Patterns

Road safety has been a critical concern globally, with various studies focusing on identifying factors contributing to accidents and developing strategies to mitigate them. According to the World Health Organization (WHO), road traffic

injuries are a leading cause of death and disability worldwide, especially in low-and middle-income countries, including India. Research has shown that factors such as road infrastructure, driver behavior, vehicle conditions, and environmental conditions significantly impact road safety (WHO, 2018).

Factors Contributing to Road Accidents

Several studies have identified key factors contributing to road accidents. These include:

Driver Behavior: Speeding, distracted driving, drunk driving, and non-compliance with traffic rules are major behavioral factors leading to accidents (Chen, 2009; Dingus et al., 2016).

Road Infrastructure: Poor road conditions, inadequate signage, lack of pedestrian facilities, and insufficient lighting are infrastructural issues that contribute to accidents (Elvik, 2008; Mohan, 2017).

Environmental Conditions: Weather conditions such as rain, fog, and extreme temperatures can adversely affect road safety (Edwards, 1998; Datla & Sharma, 2008)

Data Analytics in Road Safety

The advent of big data and advanced analytics has revolutionized the field of road safety analysis. Data analytics platforms like Qlik Sense offer powerful tools for visualizing and interpreting complex datasets, enabling stakeholders to make data-driven decisions.

3.DATA COLLECTION

3.1.COLLECT THE DATASET

For this project, data was sourced from Kaggle, a reputable platform for datasets. The dataset includes comprehensive information on road accidents across India, encompassing variables such as accident types, locations, times, weather conditions, and demographic details of those involved. The data was collected from multiple sources, ensuring a broad and representative sample for analysis. This dataset forms the foundation for our Qlik-based analysis, enabling a detailed examination of road safety and accident patterns in India.

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

3.2.CONNECT DATA WITH QLIK SENSE

After collecting the dataset from Kaggle, the next step involved integrating the data with Qlik Sense. This process included importing the dataset into Qlik Sense, ensuring compatibility and proper formatting for seamless analysis. Data connections were established, allowing Qlik Sense to efficiently access and process the information. This integration enables the use of Qlik Sense's powerful analytical tools and visualizations to explore the dataset in-depth. The successful connection of data with Qlik Sense is crucial for the subsequent stages of data preparation and analysis.
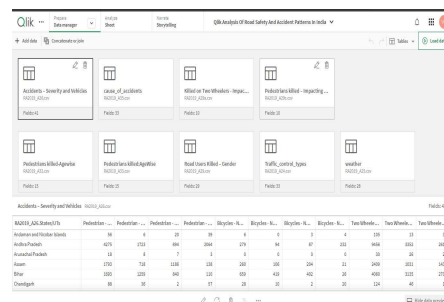
4.DATA PREPARATION

In the data preparation phase, several critical steps were undertaken to ensure the dataset's suitability for analysis. First, data cleaning was performed by removing null values and addressing any inconsistencies within the dataset. Next, the datasets were renamed for clarity and ease of use. Associations between different data fields were established to enable comprehensive analysis and meaningful insights. These preparatory steps were crucial for optimizing the dataset, ensuring it is ready for visualization and further analytical processes in Qlik Sense

4.1.PREPARING THE DATA FOR VISUALIZATION

To prepare the data for effective visualization, several key activities were performed:

Data Transformation: Adjusted the data structure to fit visualization requirements, including normalizing data ranges and aggregating data where necessary.

Field Creation: Created new calculated fields and metrics that are crucial for insightful visual analysis, such as accident rates per region or time-based trends.

Filtering and Segmentation: Applied filters and segmented the data to focus on specific areas of interest, such as high-accident zones or demographic-specific patterns.

Data Validation: Conducted validation checks to ensure that the prepared data accurately reflects the original dataset and is free from errors or biases.
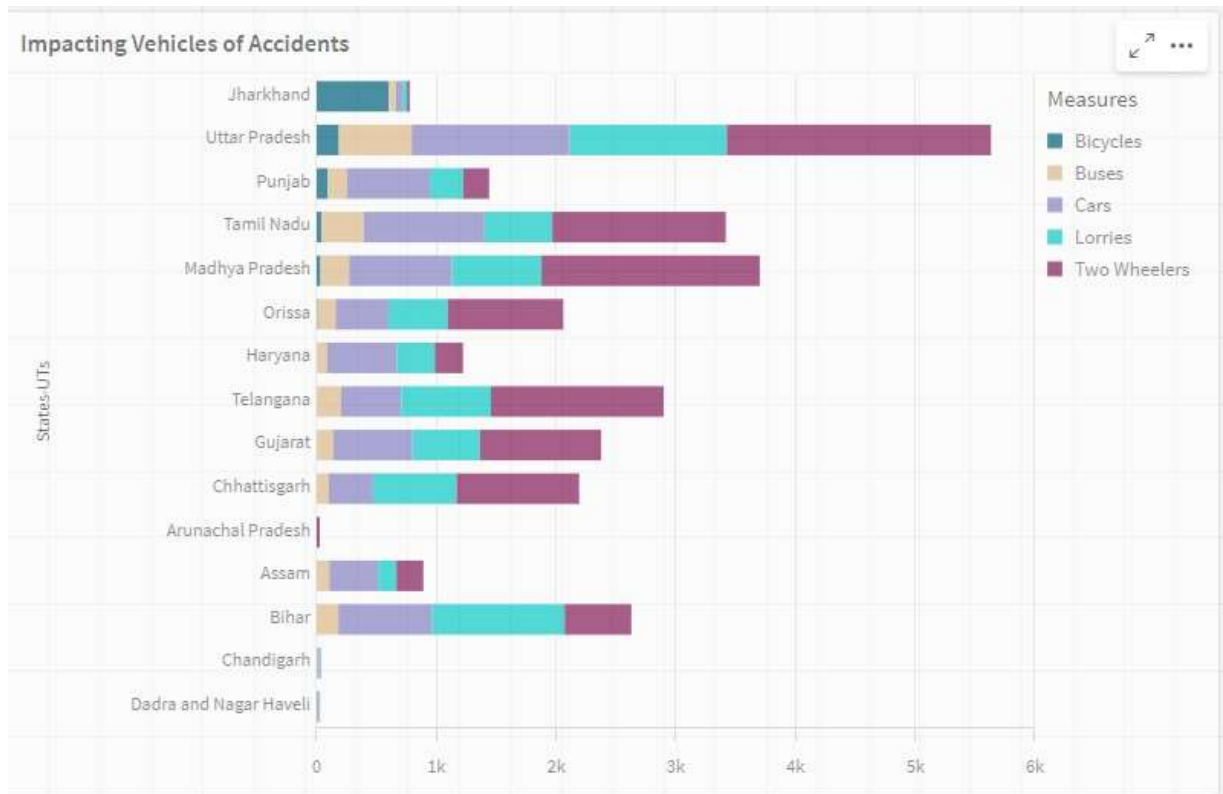
5.DATA VISUALIZATIONS

5.1. VISUALIZAIONS
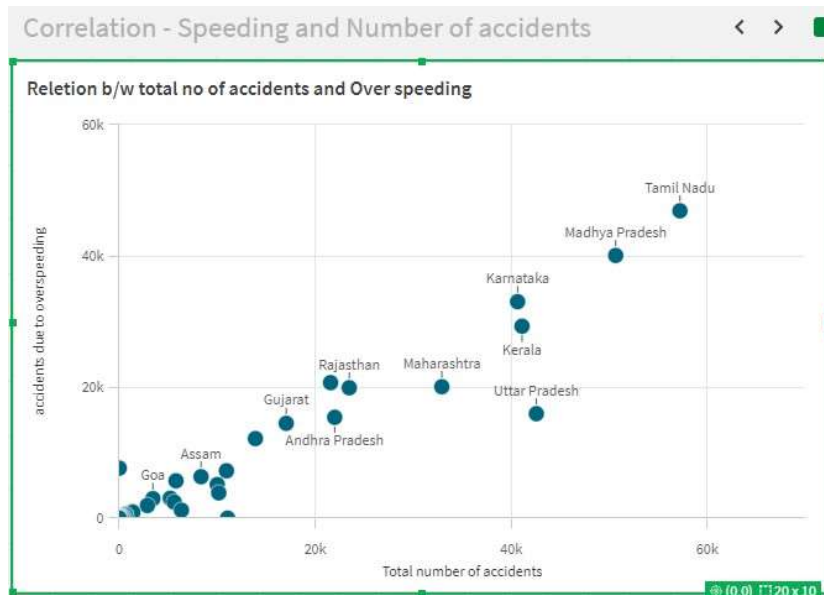
1. Accidents due to Drunken Driving
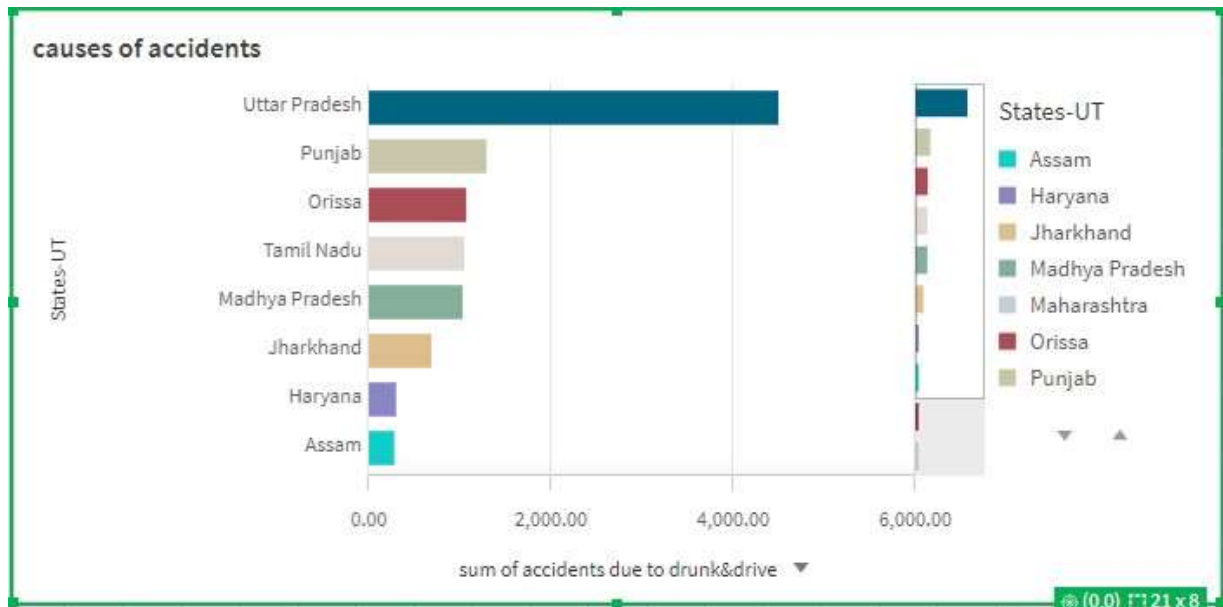
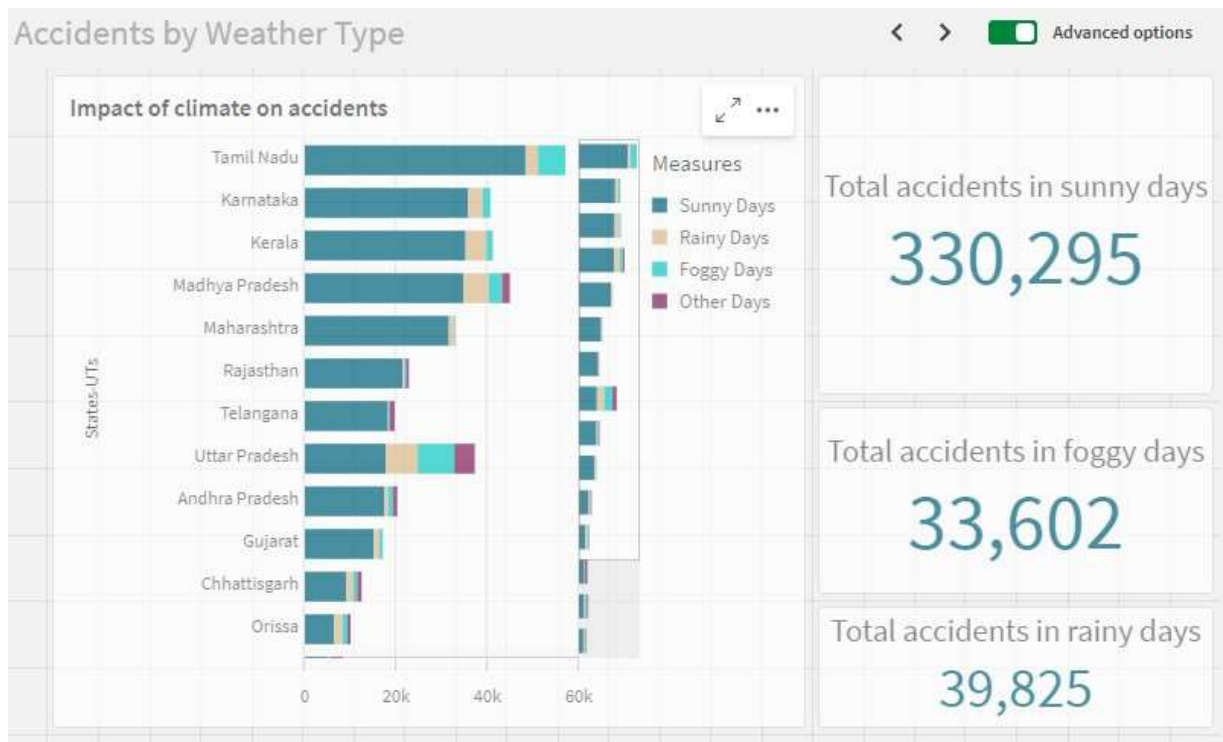### 2. State-wise Mobile Phone Usage
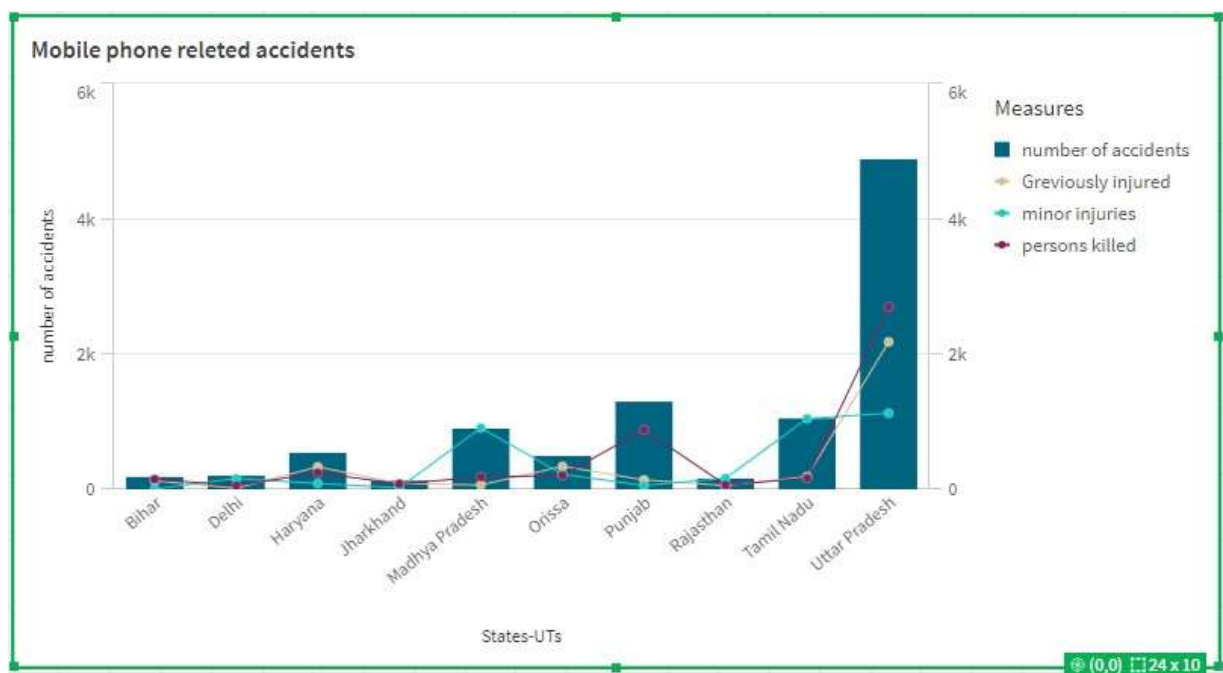
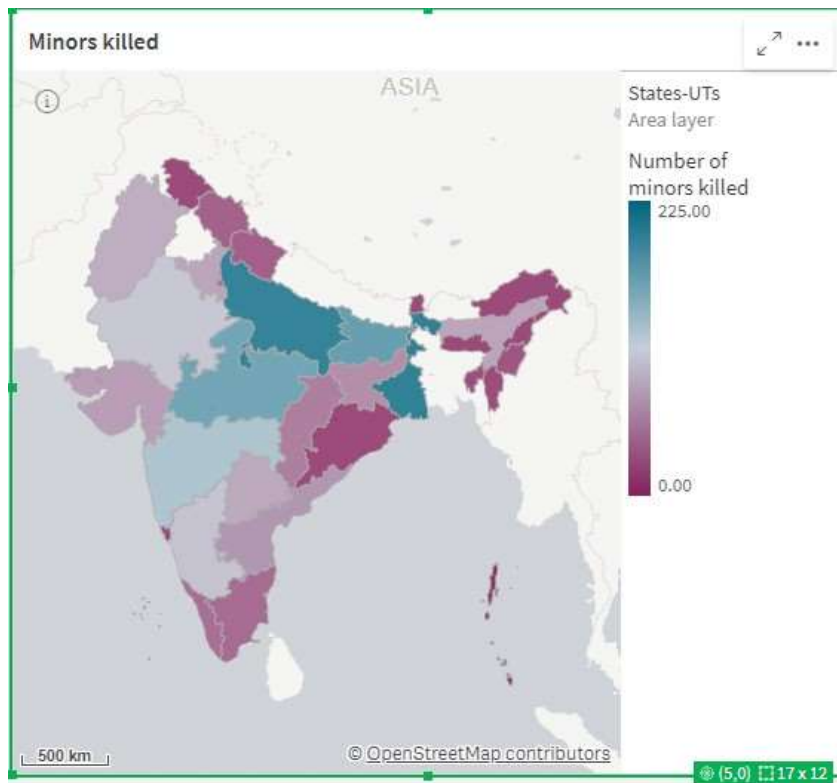## 3. Vehicle Contribution towards Total Accidents

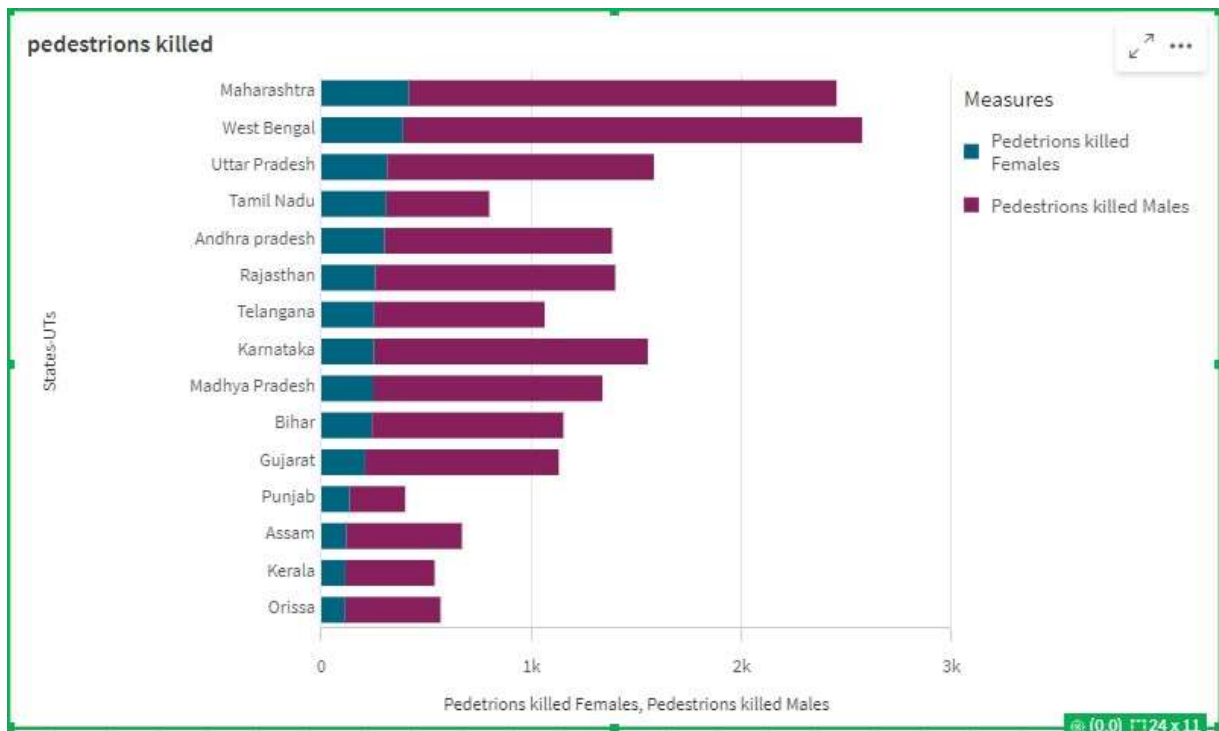## 4. Correlation - Speeding and Number of accidents
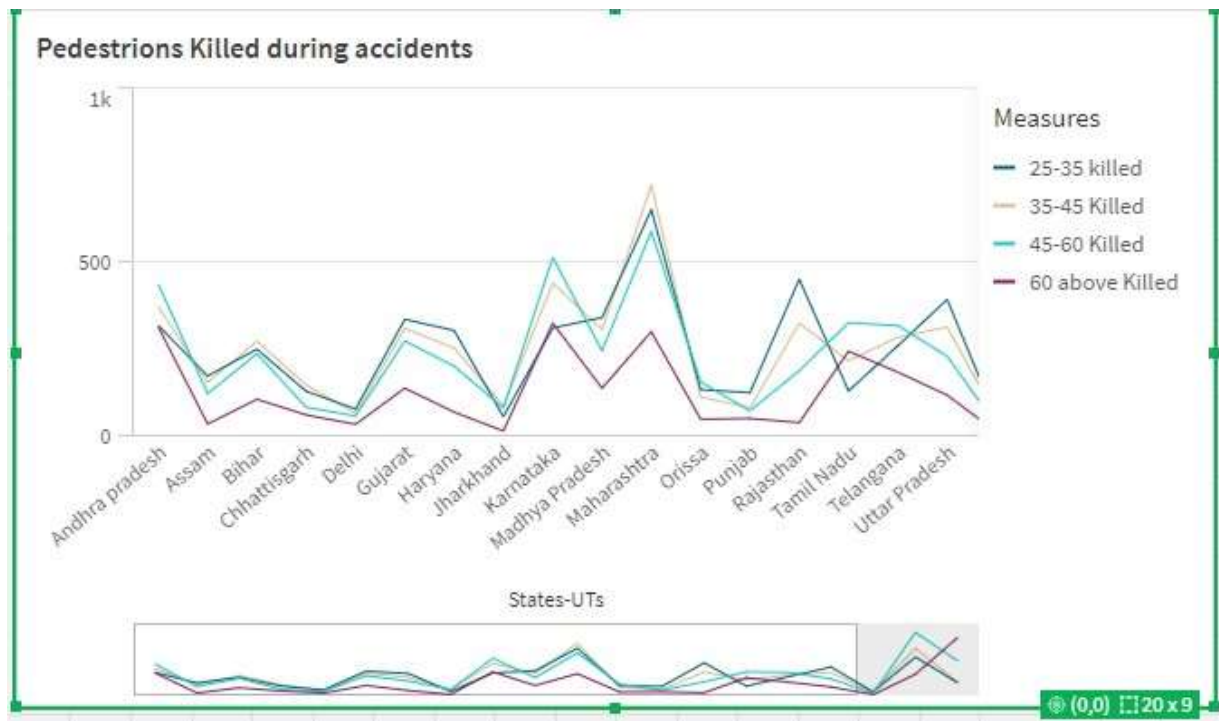
5. Accidents by Weather Type
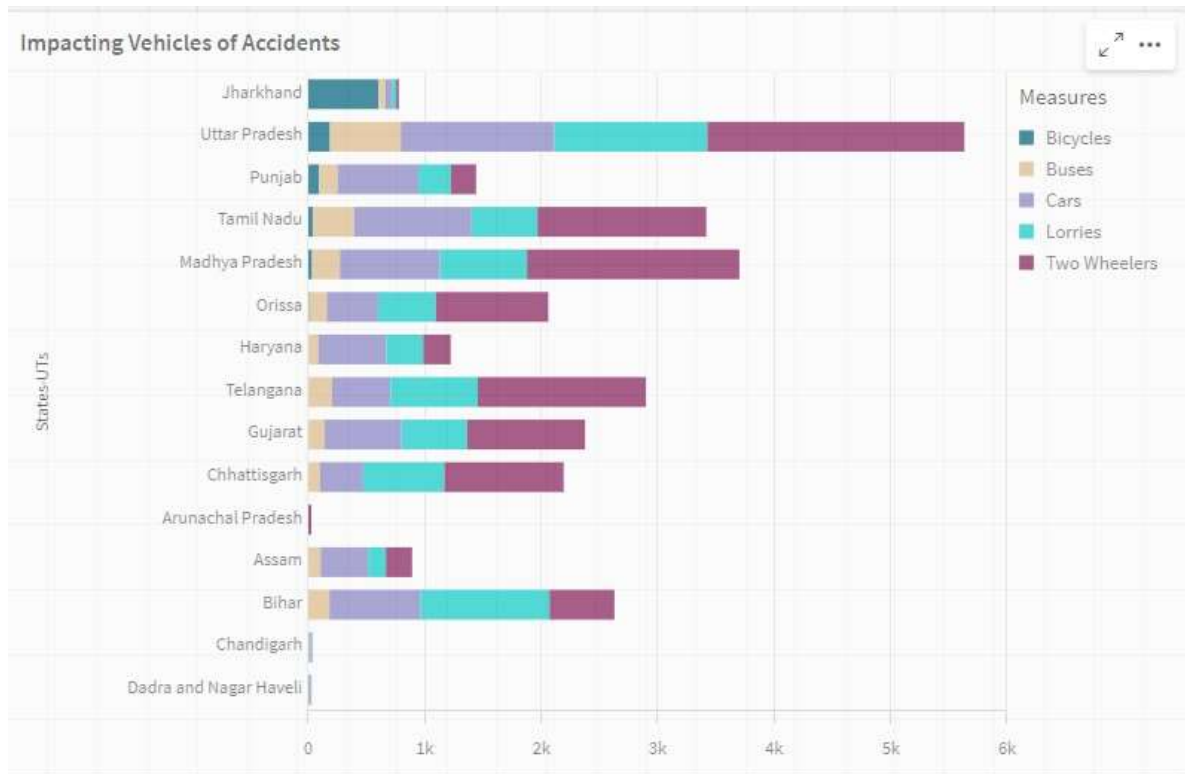

6. Minors Injured across the country

7. Pedestrians Killed: Gender

8. Pedestrians Killed: Age groups

9. Road Users Killed: Vehicle Distribution

Impacting Vehicles of Accidents

6.DASH BOARD
6.1 RESPONSIVE AND DESIGN OF 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.

Accidents near Traffic Signals
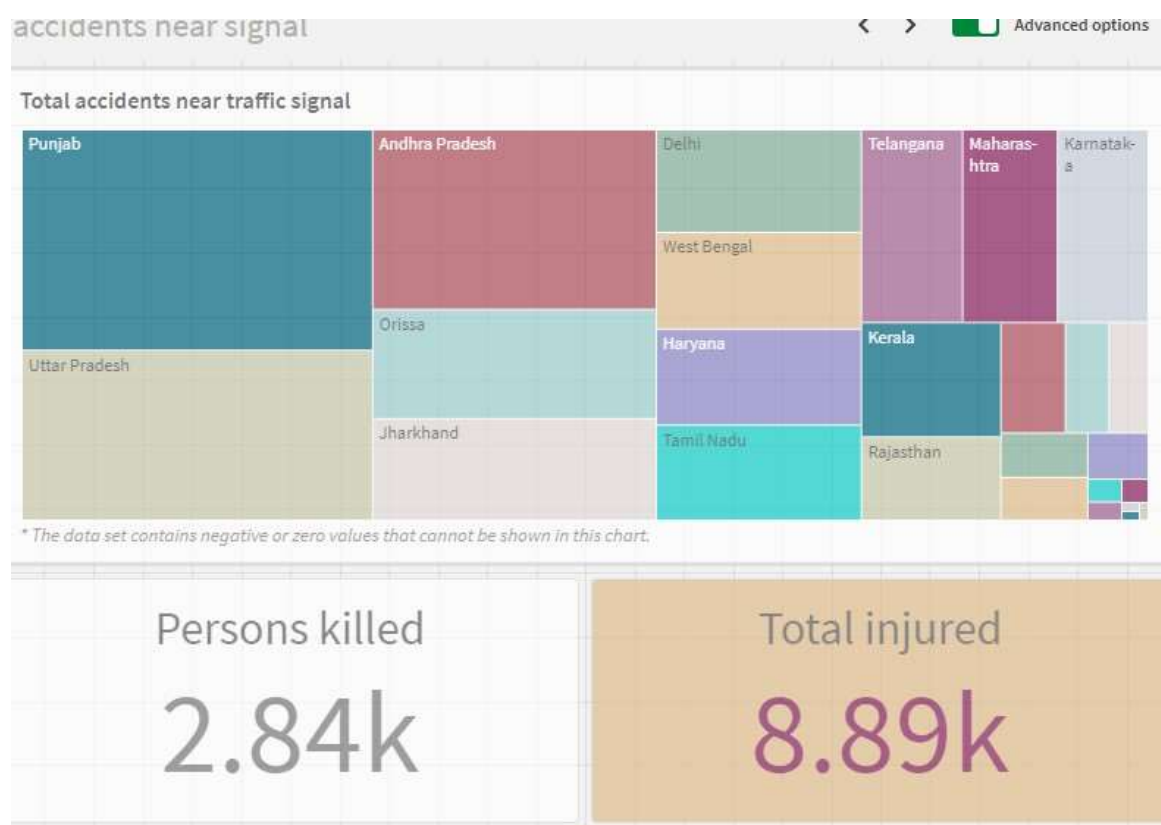
Accidents in Police Controlled Areas
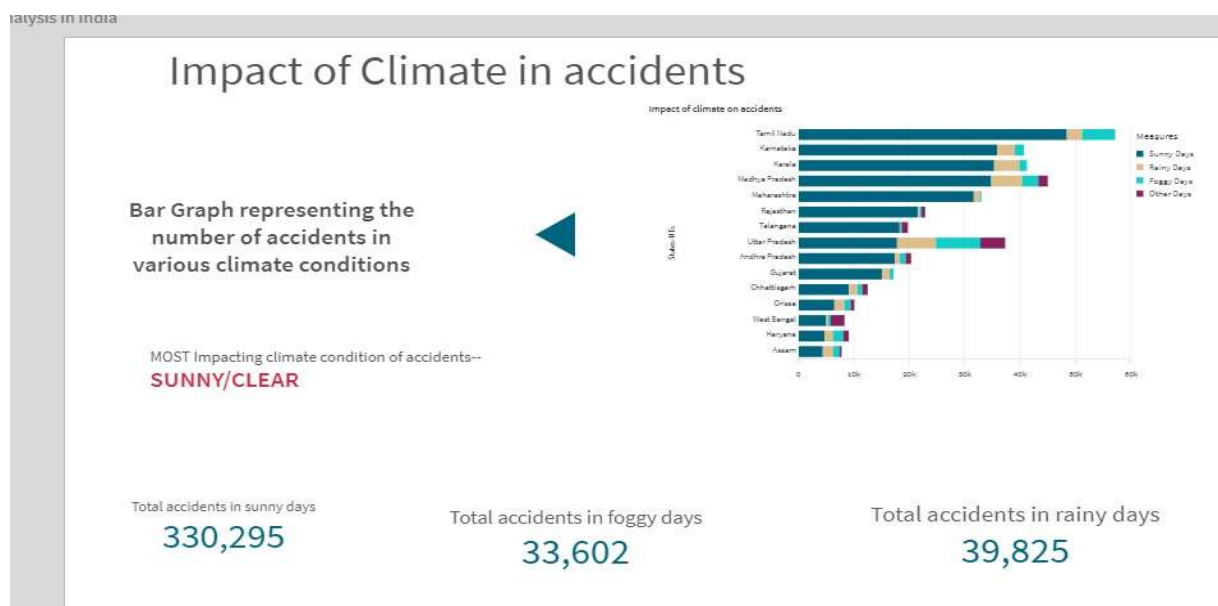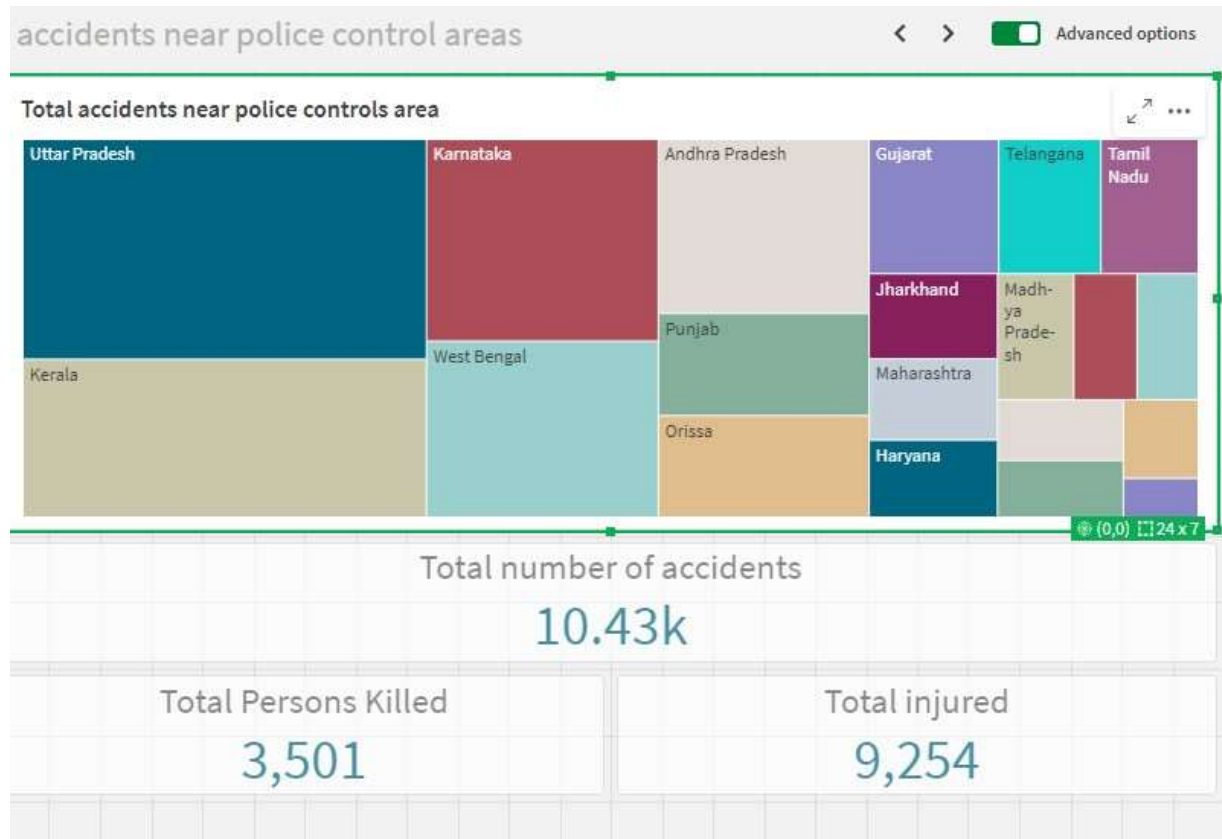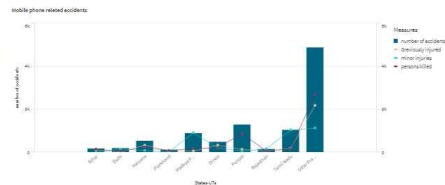
7.REPORT/STORY TELLING

7.1. Storytelling

A data story is a way of presenting data and analysis in a narrative format, with the goal of making information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of media, such as reports, presentations, interactive visualizations and videos.
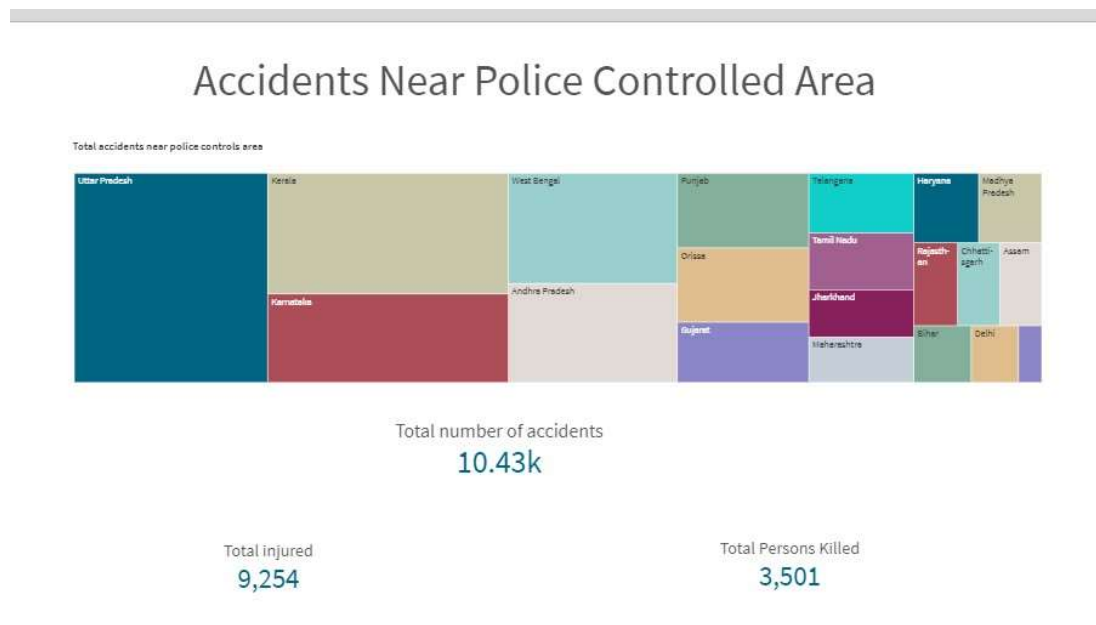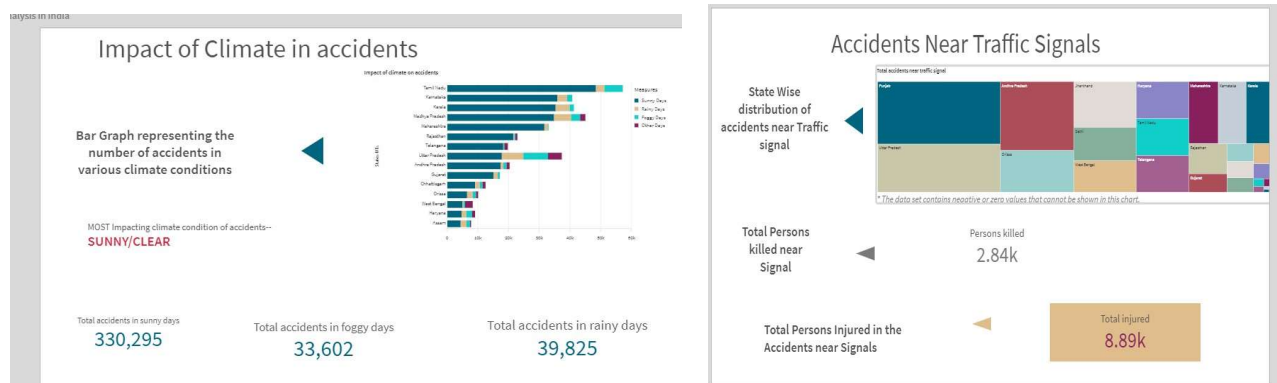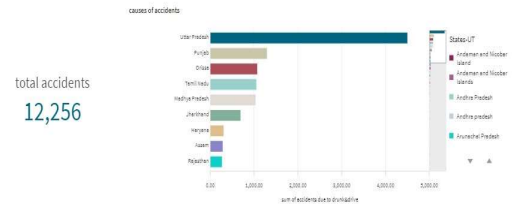
Mobile Phone related accidents

Cause Of Accident-
Use of mobile phone while driving

Mobile phone related accidents

Measures
- number of accidents
- previously injured
- minor injuries
- persons killed

States-UTs

Highest accidents - UTTAR PREDESH   **Second Highest - Punjab**

Accident due to drunk and drive

causes of accidents

total accidents
12,256

States-UT
- Andaman and Nicobar Island
- Andaman and Nicobar Islands
- Andhra Pradesh
- Andhra pradesh
- Arunachal Pradesh

sum of accidents due to drunk&drive

Highest accidents-UTTAR PREDESH   **Second Highest-PUNJAB**

Minors Injured during accidents across India

Minors killed

States-UTs
Area layer

Number of minors killed
225.00

0.00

© OpenStreetMap contributors

1000 km
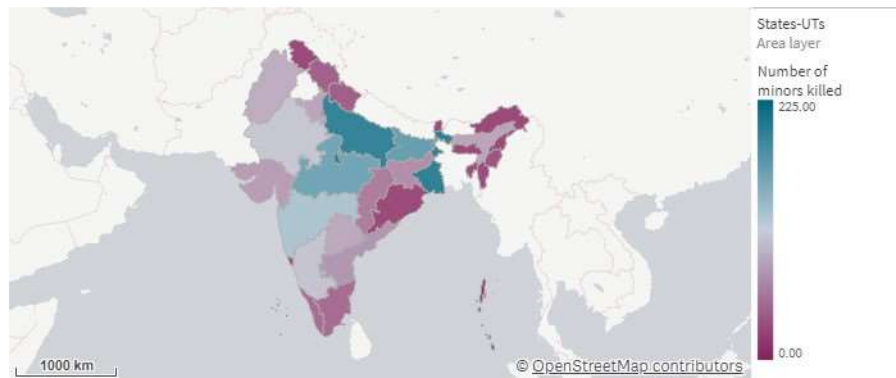
Highest -Uttar Predesh   **Second Highest - Bihar**

8.PERFORMANCE TESTING
8.1.UTILISATION OF DATA 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.