

# **Qlik Analysis Of Road Safety And Accident**

**By Siddhant Bagga  
22BCE11280  
VIT Bhopal University**

- **INTRODUCTION**
  - 1.1 Overview: A brief description about your project**
  - 1.2 Purpose: The use of this project. What can be achieved using this.**
  - 1.3 Technical Architecture**
- **Define Problem / Problem Understanding**
  - 2.1 Specify the business problem**
  - 2.2 Business requirements**
  - 2.3 Literature Survey**
- **Data Collection**
  - 3.1 Collect dataset**
  - 3.2 Connect Data with Qlik Sense**
- **Data Preparation**
  - 4.1 Prepare the Data for Visualization**
- **Data Visualizations**
  - 5.1 Visualizations**
- **Dashboard**
  - 6.1 Responsive and Design of Dashboard**
- **Report**
  - 7.1 Report Creation**
- **Performance Testing**
  - 8.1 Amount of Data Rendered**
  - 8.2 Utilization of Data Filters**

# Patterns In India

## 1 Introduction

### 1.1 Overview/Project Description:

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

**scenario1** : 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.

**scenario 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.

**scenario 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**

### **Data Sources:**

1. **Excel Files:** installed from kaggle by linked provide in folder called 'Archive'

### **Data Flow:**

1. **Data Extraction:**
  - Import data from the Excel files into Qlik.
2. **Data Transformation:**
  - Clean and preprocess data to ensure consistency (e.g., handling missing values, standardizing state names).
  - Aggregate data as required (e.g., total accidents per state, total persons killed).
3. **Data Loading:**
  - Load the transformed data into Qlik's data model.
4. **Data Modeling:**

- Create associations between different datasets to enable comprehensive analysis.
- Define dimensions (e.g., States/UTs, weather conditions) and measures (e.g., total accidents, persons killed).

#### 5. **Data Visualization:**

- Develop various visualizations such as bar charts, line charts, and pie charts to represent key metrics.
- Design interactive dashboards to allow users to filter and drill down into the data.

### **Components:**

#### 1. **Data Source Layer:**

- Excel files (eg Book2.xlsx, Book3.xlsx, Book4.xlsx).

#### 2. **ETL (Extract, Transform, Load) Layer:**

- Qlik Data Load Editor to script the extraction and transformation of data.

#### 3. **Data Model Layer:**

- Qlik Associative Engine to create a data model linking different datasets.

#### 4. **Visualization Layer:**

- Qlik Sense/QlikView for creating interactive dashboards and reports.

#### 5. **User Interface Layer:**

- Dashboards and reports accessible to business users via Qlik Sense/QlikView.

| Excel Files | --> | Qlik Data Load | --> | Qlik Associative | -->  
 | (Book2,3,4.xlsx) | --> | Editor (ETL) | --> | Engine (Data | --> | Qlik  
 Sense/Qlik | --> | Modeling | --> | View (Dashboards | --> | Business  
 User |

# 2. Define Problem / Problem Understanding

## 2.1 Specify the Business Problem

The problem of road safety in India presents a significant challenge amid the backdrop of technological advancements that have minimized distances but concurrently heightened risks to human life. Despite the convenience afforded by modern transportation, the country witnesses a distressing annual toll of lives lost and injuries sustained due to road accidents. To address this pressing issue, a comprehensive study is needed to analyze road safety and accident trends in India. Leveraging the capabilities of Qlik Sense, a data analytics platform, this study aims to examine various parameters such as accident types, locations, causes, and contributing factors. By utilizing data-driven insights generated from this analysis, stakeholders can devise strategies and interventions aimed at mitigating risks and improving road safety outcomes across the nation. This problem statement underscores the urgency of leveraging technology and data analytics to confront the pervasive challenge of road safety in India and pave the way for a safer transportation environment for all.

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

Understanding the problem of road safety in India necessitates a comprehensive examination of the multifaceted factors contributing to the alarming rate of accidents and casualties on the nation's roads. Despite advancements in transportation technology, the persistent threat to human life remains a pressing concern. Analysis of historical data reveals patterns in accident types, locations, causes, and contributing factors, shedding light on the complexities of road safety challenges. Factors such as road infrastructure deficiencies, vehicle safety standards, human behavior, and enforcement of traffic regulations emerge as critical areas requiring attention. Moreover, socioeconomic disparities and urbanization trends amplify the risks faced by vulnerable road users, including pedestrians, cyclists, and motorcyclists. By leveraging data analytics tools like Qlik Sense, stakeholders gain a deeper understanding of these dynamics, enabling evidence-based decision-making to develop targeted interventions. This problem understanding underscores the need for a holistic approach to address the multifaceted nature of road safety challenges in India, encompassing infrastructure improvements, policy reforms, behavioral interventions, and public awareness campaigns to create safer road environments and reduce the toll of accidents and injuries on Indian roads.

## **2.2 Business Requirements**

The primary focus is on creating interactive and visually compelling

dashboards to support strategic planning and operational improvements. By examining the demographics of individuals involved in accidents, such as age, gender, and occupation, we can identify vulnerable groups and tailor safety campaigns accordingly. Analyzing the temporal and spatial patterns of accidents, including peak times, seasonal variations, and geographic hotspots, will help in deploying targeted interventions and resources. Identifying specific problem areas where accidents are frequent or severe, such as dangerous intersections, poorly lit roads, or areas with high pedestrian traffic, will assist in prioritizing infrastructural improvements and policy changes.

The creation of interactive and visually compelling dashboards using Qlik Sense is essential for effectively communicating these insights. These dashboards will provide a clear and intuitive overview of accident data, aiding long-term strategic planning and policy development. They will also offer valuable insights for operational improvements, such as optimizing traffic enforcement and emergency response strategies. By enabling stakeholders to make informed decisions based on real-time data and trends, these dashboards will enhance decision-making processes at various levels of governance and management.

The insights derived from this analysis will be instrumental in making informed decisions, implementing better safety protocols, and ensuring compliance with regulations. Data-driven insights will support the development of targeted prevention measures and identify best practices and areas for improvement in existing safety protocols to reduce accident rates. Additionally, the analysis will help ensure compliance with traffic regulations and standards by identifying gaps and areas for enhancement. Public awareness about road safety issues will also be raised, educating drivers, pedestrians, and other road users on safe practices. The ultimate goal is to reduce accidents, save lives, and create a safer transportation



environment through informed decision-making and effective safety measures.

## **2.3 Literature Survey**

### **Overview**

The literature survey aims to gather and review existing research, studies, articles, reports, and figures related to road safety and accident patterns. This involves exploring various sources, including academic databases and government publications, to understand the methodologies used, findings, and implications of previous studies. The goal is to build a robust foundation for our analysis by leveraging existing knowledge and identifying gaps that our study can address.

### **Methodologies for Analyzing Accident Data**

Previous studies on road safety and accident analysis have employed a variety of methodologies, each with its strengths and limitations. Common techniques include:

#### **1. Statistical Analysis:**

- Regression models to identify factors influencing accident frequency and severity.
- Time-series analysis to detect trends and seasonal variations in accident data.

#### **2. Machine Learning Models:**

- Classification algorithms (e.g., decision trees, random forests) to predict accident hotspots.
- Clustering techniques (e.g., k-means clustering) to group similar accident types and locations.

#### **3. Geographic Information Systems (GIS):**

- Spatial analysis to visualize accident locations and identify high-risk areas.

- Heat maps to highlight accident density and spatial distribution patterns.

## **Key Findings from Previous Studies**

Reviewing the results and conclusions of past research provides valuable benchmarks for our analysis. Notable findings include:

- **Demographic Insights:** Studies often reveal that certain demographic groups, such as young drivers and pedestrians, are at higher risk of accidents.
- **Temporal Patterns:** Research indicates that accidents tend to peak during certain times of the day and specific days of the week, often correlating with traffic volume and alcohol consumption.
- **Environmental Factors:** Poor road conditions, inadequate lighting, and adverse weather conditions are commonly cited as significant contributors to road accidents.

## **Government Reports and Publications**

Government documents are critical sources of comprehensive data and policy insights. Key reports and publications include:

- **National Crime Records Bureau (NCRB) Reports:** Annual statistics on road accidents, fatalities, and injuries across India.
- **Ministry of Road Transport and Highways (MoRTH) Publications:** Policies, safety guidelines, and strategic plans for reducing road accidents.
- **World Health Organization (WHO) Reports:** Global road safety statistics and recommendations for best practices in accident prevention.

## **Recommendations for Best Practices**

The literature survey also identifies best practices and successful

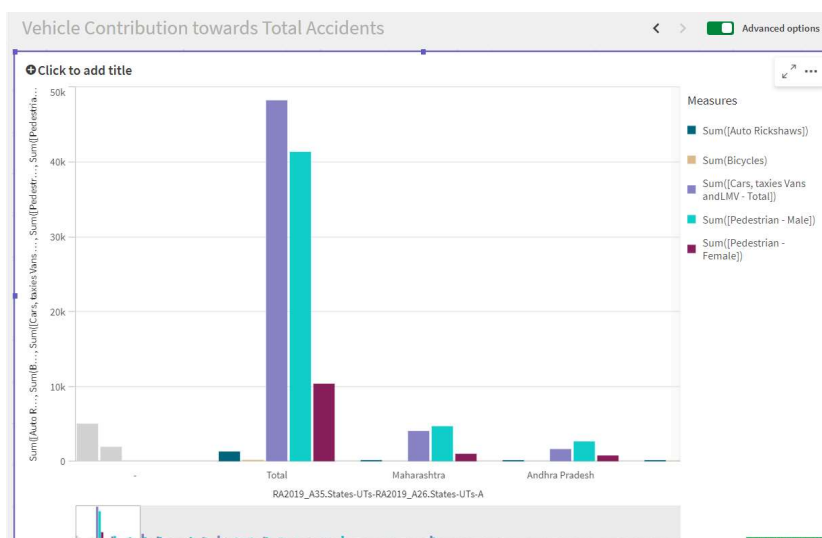
interventions implemented in various regions. These include:

- **Traffic Calming Measures:** Implementation of speed bumps, roundabouts, and pedestrian zones to reduce vehicle speeds and enhance safety.
- **Public Awareness Campaigns:** Educational programs aimed at promoting safe driving behaviors and increasing awareness of road safety issues.
- **Technology Integration:** Use of advanced technologies such as automated traffic enforcement systems and real-time traffic monitoring to prevent accidents.

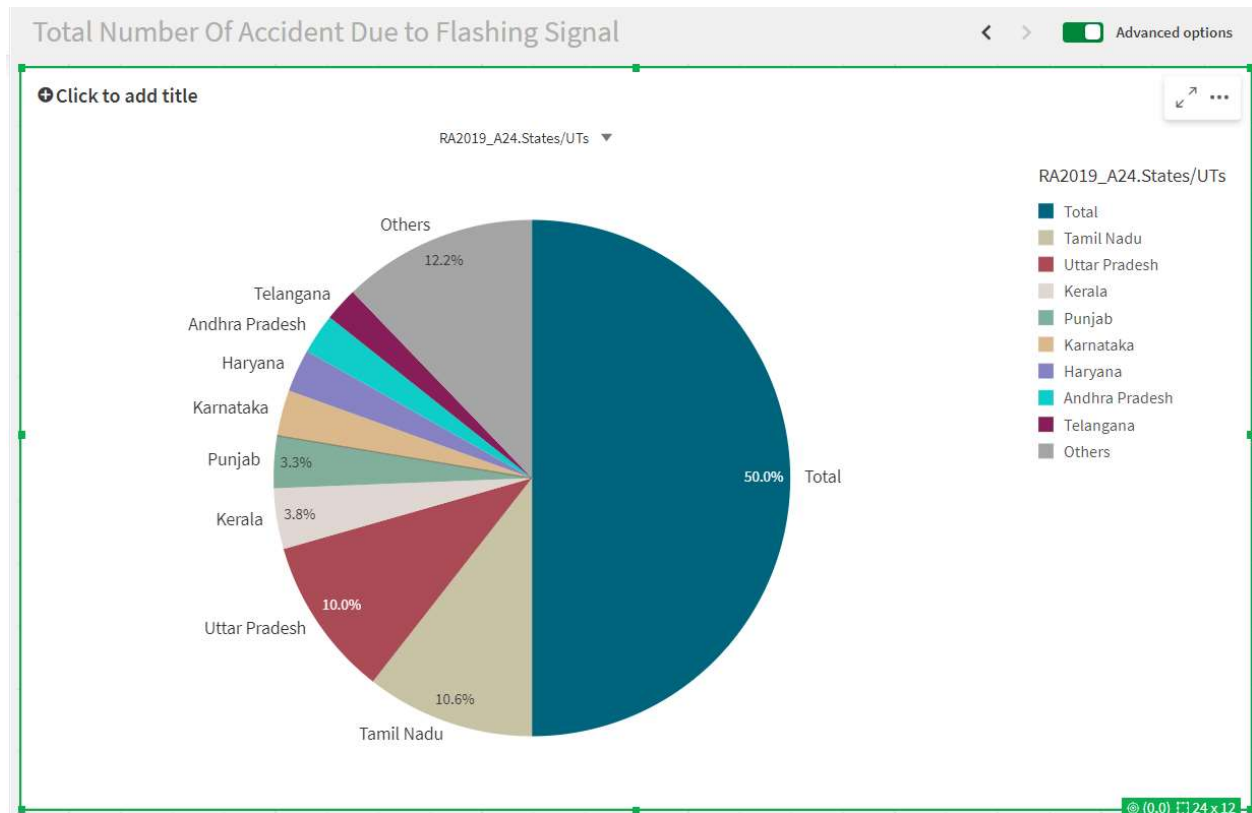
## Conclusion

The literature survey provides a comprehensive understanding of the existing body of knowledge on road safety and accident analysis. By synthesizing findings from various sources, including academic research and government reports, we establish a strong foundation for our study. This survey informs our methodological approach, highlights key factors influencing road safety, and guides the development of data-driven strategies for reducing accidents and improving road safety in India.

## 2.3 Social Impact



Analyzing the social impact of vehicle contributions towards accidents involves looking at how different types of vehicles contribute to road accidents and the resulting injuries or fatalities. This analysis can help in understanding how specific vehicle types influence road safety and what social measures can be taken to mitigate risks.



To analyze the social impact of accidents caused by flashing signals, we need to focus on data that specifically relates to accidents under such conditions. Since the datasets provided do not explicitly mention accidents due to flashing signals

## 3 Data collection

### 3.1 Collect the dataset

To collect the dataset from a Kaggle folder named 'Archive', first, navigate to the specific dataset page on Kaggle. Click the "Download" button to download the entire dataset as a ZIP file. Once downloaded, extract the contents of the ZIP file, which will typically create a folder named 'Archive' containing various CSV files, documentation, and possibly other data

formats. Use data analysis tools like Python (with pandas) or data visualization software to load and analyze these CSV files. For instance, in Python, you can use the pandas library to read the CSV files into DataFrames and begin your data processing and analysis workflow. Ensure to check for any README or documentation files within the 'Archive' folder that provide detailed information about the data structure and contents.

### **3.2 Connect Data with Qlik Sense**

To connect data with Qlik Sense, first ensure that your dataset from the 'Archive' folder on Kaggle is downloaded and extracted. Open Qlik Sense and create a new app. Within the app, click on "Add Data" and navigate to the folder containing your dataset. Select the appropriate CSV files or other data formats you wish to import. Qlik Sense will automatically profile the data and suggest associations between different tables. Review these associations to ensure accuracy, then load the data into your app. Once loaded, you can start creating visualizations and dashboards, leveraging Qlik Sense's powerful data analysis capabilities to gain insights from your dataset.

## **4 Data Preparation**

### **4.1 prepare the the data for visualization**

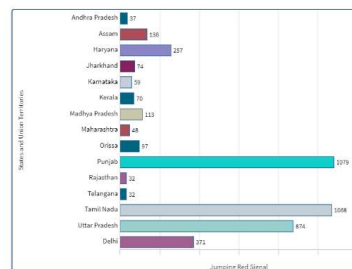
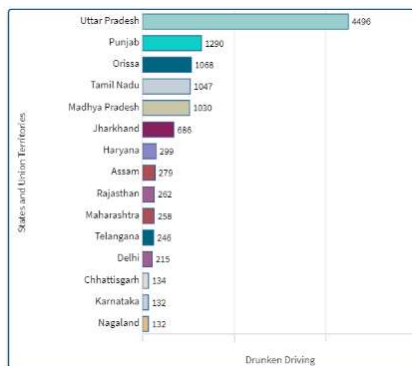
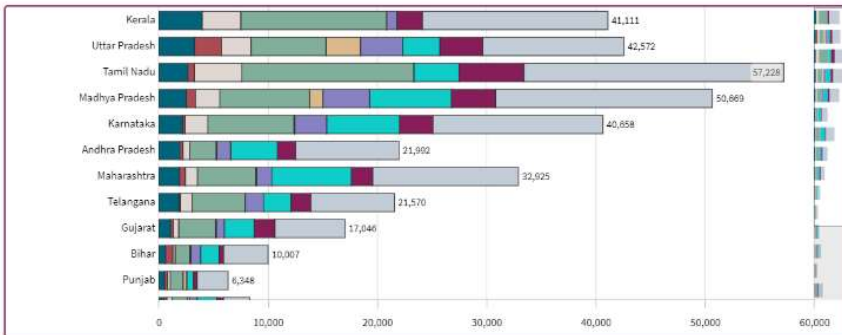
To prepare the data for visualization, start by cleaning the dataset to ensure accuracy and consistency. This involves handling missing values, correcting data types, and standardizing formats. Next, aggregate or transform the data as needed, creating new calculated fields or aggregating data points to provide meaningful insights. Ensure that the data is structured properly, with clear relationships between tables if working with multiple datasets. Finally, verify the integrity of the data and ensure it aligns with your visualization goals, making it ready for importing

into visualization tools like Qlik Sense.

## 5 Data Visualization

### 5.1 Visualization

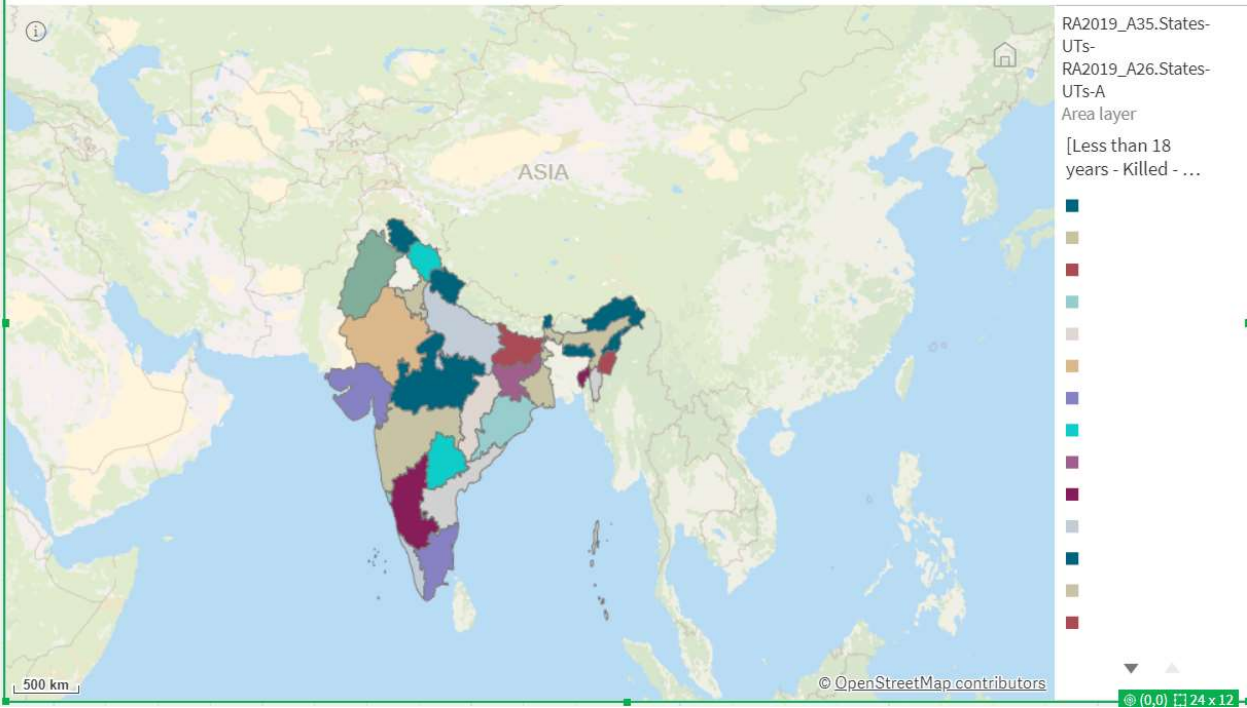
Here are few of the visualiztion to understand our business data



## Major Injury Across Country

< > ☒ Advanced options

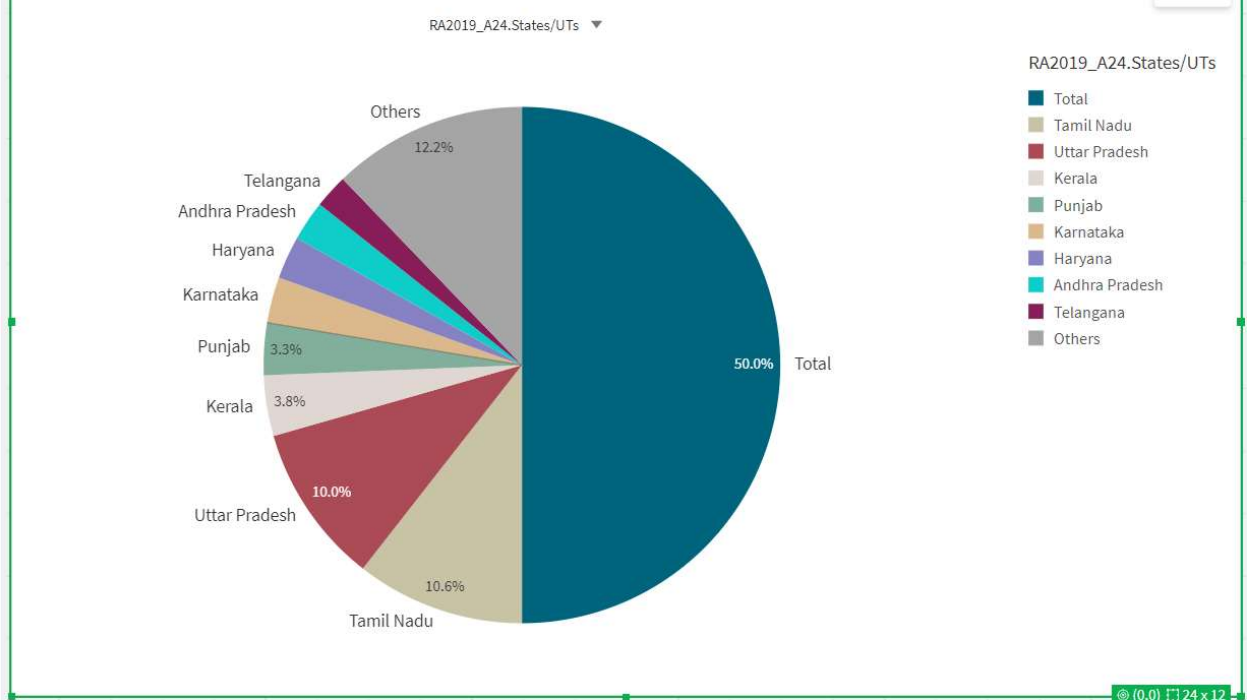
Click to add title



## Total Number Of Accident Due to Flashing Signal

< > ☒ Advanced options

Click to add title



< > ☒ Advanced options



< > ☒ Advanced options



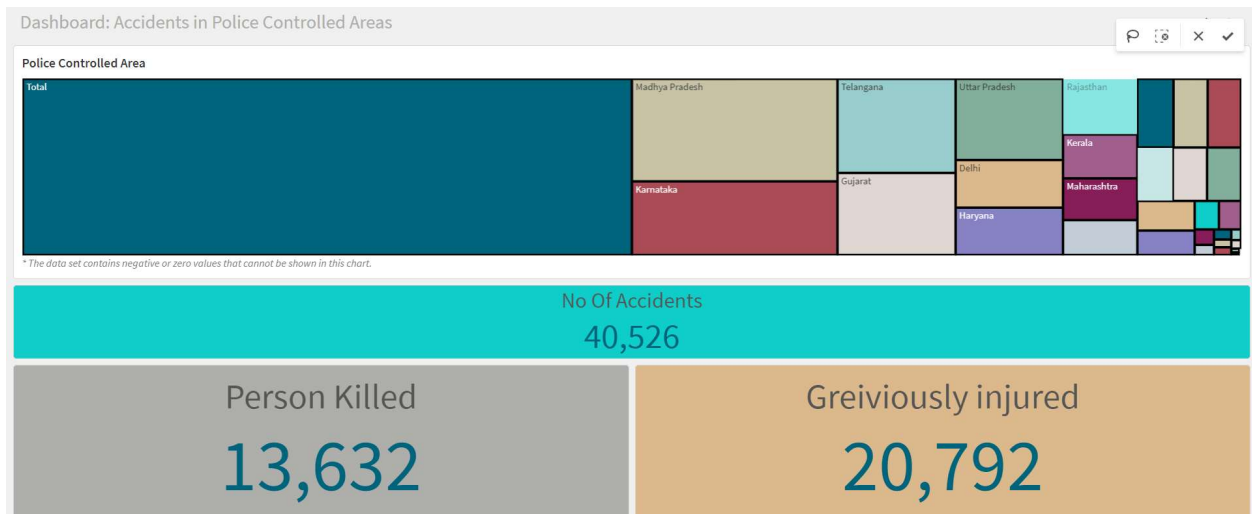


## 6 Dashboard

### 6.1 Responsive and design of Dashboard



A Responsive database of Accidents near Traffic Signal



A Responsive database of Accidents under Police controlled Areas

## 7 Report

## **7.1 Report Creation**

Report creation is essential for projects as it serves as a comprehensive documentation of progress, findings, and outcomes. It provides stakeholders with insights into project status, resource allocation, and potential challenges. Reports also facilitate communication, ensuring that all involved parties are informed and aligned with project goals.

Additionally, they serve as a valuable reference for future projects, enabling teams to learn from successes and failures, thus improving overall project management efficiency and effectiveness.

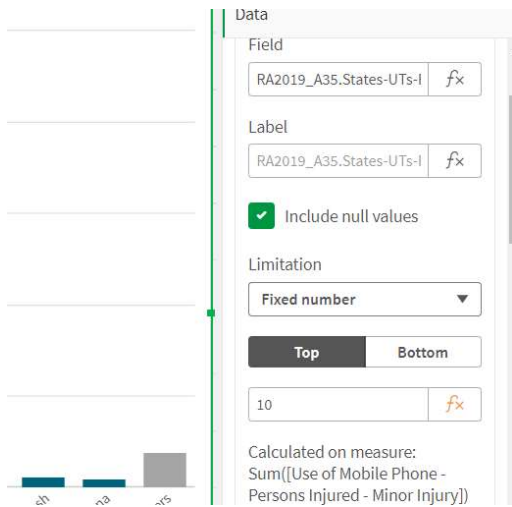
## **8 Performance Training**

### **8.1 Amount of Data Rendered**

The amount of data rendered depends on the complexity and size of the dataset being visualized. For large datasets, it's essential to optimize data processing to ensure smooth rendering and interaction. Techniques such as data aggregation, sampling, and efficient data modeling can help manage large volumes of data, ensuring that visualizations are responsive and informative without overwhelming the system or the user.

### **8.2 Utilization of Data Filter**

Utilization of data filters in visualization tools like Qlik Sense enhances data analysis by allowing users to focus on specific subsets of data. Filters enable users to drill down into the data, explore different dimensions, and uncover insights relevant to particular criteria such as time periods, geographic locations, or specific categories. This interactivity ensures that users can dynamically adjust their views, isolate trends, and make data-driven decisions based on precise and relevant information. Effective use of data filters improves the overall usability and functionality of visualizations, making them more powerful and insightful.



# Thank You