# Qlik Analysis of Road Safety and Accident Patterns in India

## CONTENTS

## 7 Report

7.1 Report Creation

## 8 Performance Testing

8.1 Amount of Data Rendered

8.2 Utilization of Data Filters

# 1 INTRODUCTION

## 1.1 Overview: A brief description about your project

This project aims to uncover trends, identify high-accident areas, and identify underlying causes of unfortunate or tragic road accidents or incidents through the integration of large amounts of data i.e traffic data, accident reports, weather, road data, population estimates and more. The objective of this project is to conduct a thorough analysis of accidents data, with the am of astracting valuable insights to anitance decision making, policy formulation, and operational efficiency within our country.

## 1.2 Purpose: The use of this project. What can be achieved using this.

This analysis is to improve road safety and reduce accidents in India by leveraging Qlik's data analytics platform. The analysis aims to:

**i) Improve Emergency Response:**

Optimize emergency response time by identifying accident-prone areas and potential hotspots in the future. Determine emergency services and treatment of areas of greatest need. Improve coordination between response teams through real-time data sharing and alerting.

**ii) To engage and educate the public:**

Provide datadriven findings to conduct public awareness campaigns focused on key risks and driving safety practices. Educate drivers, pedestrians and cyclists about common accident causes and prevention methods. Promote a culture of road safety through continuous participatory education programs informed by case studies.

**iii) Identify accident hot spots:**

Build specific areas or roads where accidents are common. Links accident data to factors such as traffic volume, road conditions and time of day to identify areas where accidents are more likely to occur. Provide critical feedback on targeted interventions such as enhanced traffic management, improved signage and speed limit reforms.

**iv) Development of predictive models:**

Uses predictive analytics to predict potential accidents based on realtime data input. Consider variables such as weather forecasts, traffic flow, and historical accident data. Provide early warning and suggest measures to prevent accidents. Enables authorities to use resources strategically and implement security priorities.

## Social Impact:

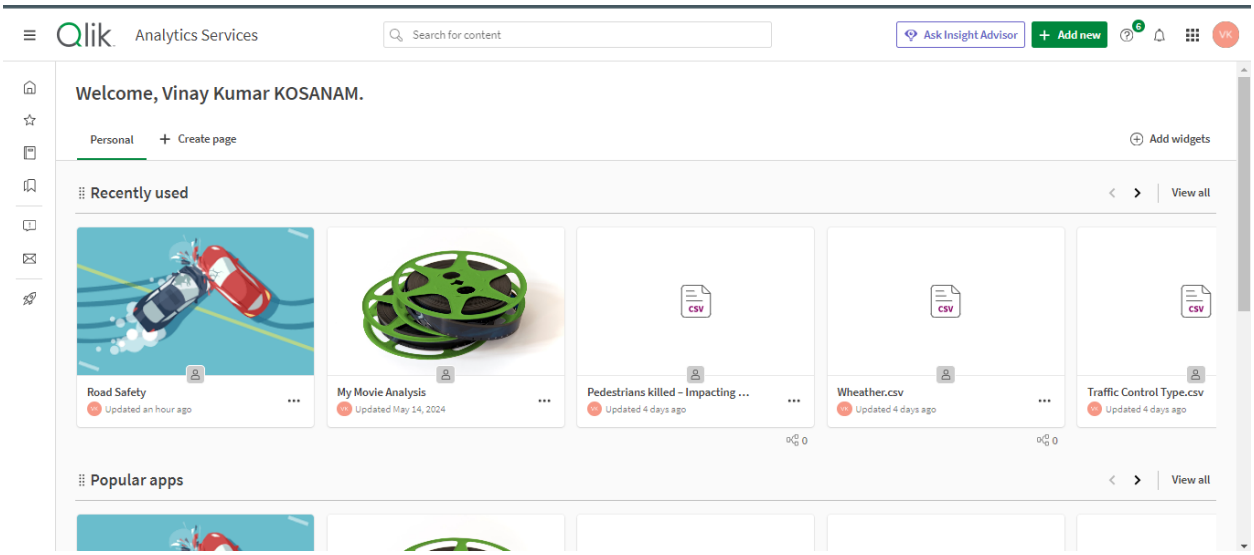
• Compare the severity of accidents in different areas of traffic control.
• Identify the leading causes of accidents.

## Business Model/Impact:

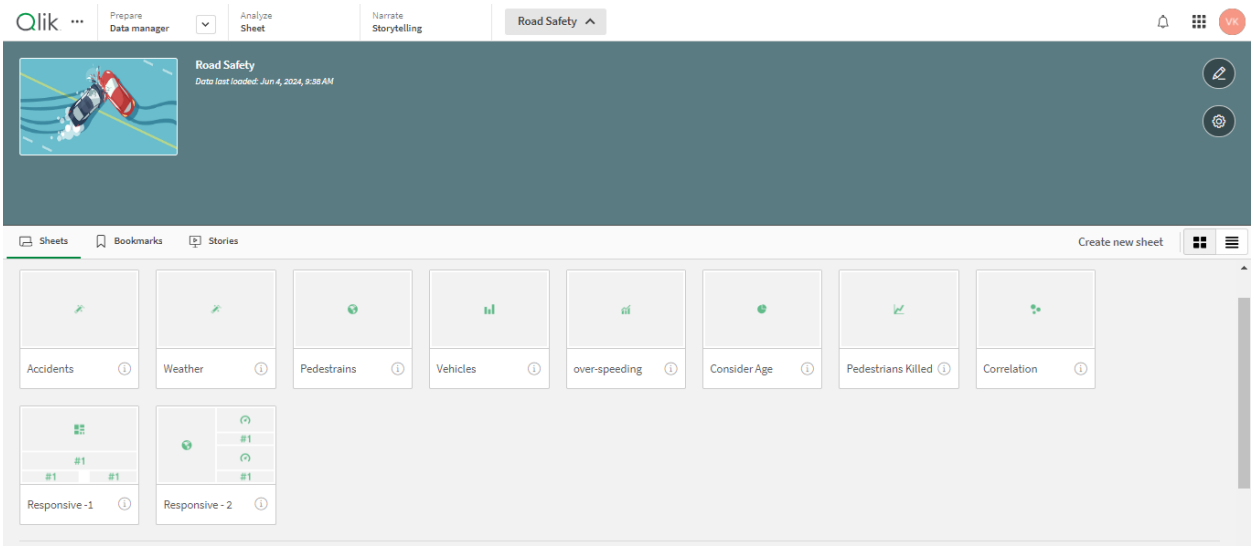
• Analysis of road accident data from 2019 can aid insurance companies, vehicle manufacturers. and transportation authorities developing strategies to mitigate risks, improve safety features, and enhance road infrastructure.

## 1.3 Technical Architecture

We used Qlik Cloud Analytics to create a visualization and for analytics . The Qlik engine supports self-service analytics, interactive dashboards, conversational analytics, custom and embedded analytics, mobile analytics, reporting, and alerting.




**Preview of Qlik Cloud Analytics Platform**




**Interface of Analytic App**

# 2 Define Problem/Problem Understanding

## 2.1 Specify the business problem

Hundreds of thousands of people lose their lives and millions of people are seriously injured every year in India due to accidents. There is an urgent need to analyze road safety and accident data to understand the root causes of incidents and the factors contributing to these incidents. The goal is to use Qlik Sense, a data analytics platform, to analyze data about road incidentsincluding accident types, locations, and causes to identify patterns and gain insights. These insights can then inform strategies and interventions aimed at improving road safety and reducing the number of deaths and injuries on Indian roads.

**Identify high-risk areas**: Identify specific locations and high-accident areas to prioritize safety measures.
**Understand seasonal and seasonal trends**: Analyze how accident rates change by season, time of day and day of week and adjust preventive measures accordingly.
**Methodological analysis**: Evaluate the impact of road conditions and infrastructure quality on accident rates to inform improvements and investments.
**Evaluate the effectiveness of current safety measures**: Review existing road safety measures and systems to determine their effectiveness and identify areas for improvement.
**Predicting future accident scenarios**: Use predictive modeling to predict potential future accidents based on historical data and real-time input, enabling proactive actions.
**Improve emergency response**: Understand accident planning and provide proper emergency services utilization to ensure timely delivery of appropriate assistance in high-risk areas.

**Promote public awareness and education**: Use data-driven insights to develop targeted awareness campaigns and education programs to enhance street culture

## 2.2 Business requirements

The project aims to analyze user demographics, accident patterns and problem areas through the creation of interactive and attractive dashboards These dashboards will provide valuable insights and support policy implementation and operational improvements to enable informed decision-making, enhanced security measures and regulatory compliance.

**Functional requirements**

- **Data Collection and Integration**: Collect data from multiple sources, including user profiles, incident reports, and the environment. Make sure the data is clean, accurate, and up-to-date. Integrate data into a centralized database for analysis.
- **Dashboard Enhancements**: Create interactive dashboards that allow users to visualize and explore data insights. Use a variety of visualization techniques (e.g., graphs, heat maps, charts) to effectively present data.Enable filtering and drill-down capabilities for in-depth analysis.
- **Analytical tools and methods**: Use statistical and machine learning models to analyze data and identify patterns. Use predictive analytics to anticipate potential problem areas and accidents. Ensure that tools are user-friendly and accessible to non-technical stakeholders.
- **Reporting and Analysis:** Produce regular reports summarizing key findings and insights from research. Provide actionable recommendations based on data insights. Ensure that reports are tailored to meet the needs of various stakeholders, including policy makers, operations managers, and compliance

personnel.

- **Security policies and compliance:** Insights must be developed to support the creation and improvement of security measures. Ensure all recommendations comply with legal requirements. Evaluate the impact of implemented changes on accident rates and user safety.

**Non-Functiona requirements**

- **Performance and Scalability:** Make sure the system can handle large amounts of data efficiently. Make the system scalable to fit future data developments.
- **Security and Privacy:** Implement strong security measures to protect sensitive user information. Ensure compliance with data protection laws and privacy policies.
- **Usability:** Create user-friendly interfaces for dashboards. Ensure that it is easy to use for participants with varying technical backgrounds.
- **Reliability :** Ensure that the system is highly reliable with minimal downtime. Provide users with access to dashboards and reports.

# 3 Data Collection

## 3.1 Collect the dataset

Kaggle is a well-known platform that offers a variety of datasets for data science and machine learning projects. Kaggle provides datasets across various domains, so it's crucial to select ones that align with your project's goals. Kaggle datasets often contain multiple types of data, such as numerical, categorical, and text data.

Choose datasets relevant to our analysis. Given link redirect to dataset in kaggle. There we can easily download the Dataset.
link - https://www.kaggle.com/datasets/aryakittukrishnasai/road-accidents-in-india

**The dataset contains different types of categorical data, which include :**

- State/UT-wise Pedestrians killed according to classification of age and sex during 2019
- State/UT-wise Pedestrians killed in Accidents Classified by the type of impacting vehicles during 2019
- State/UT-wise Accidents Classified according to Type of Traffic Control during 2019
- State/UT-wise Accidents classified according to Load Condition of Involved Vehicle during 2019
- State/UT-wise Two Wheelers killed in Accidents Classified by the type of impacting vehicles during 2019
- State/UT-wise Male and Female Persons Killed in Road Accidents in terms of Road User categories during 2019
- State/UT-wise Accidents Victims Classified according to Non-Use of Safety Device (Non Wearing of Helmet) during 2019 etc

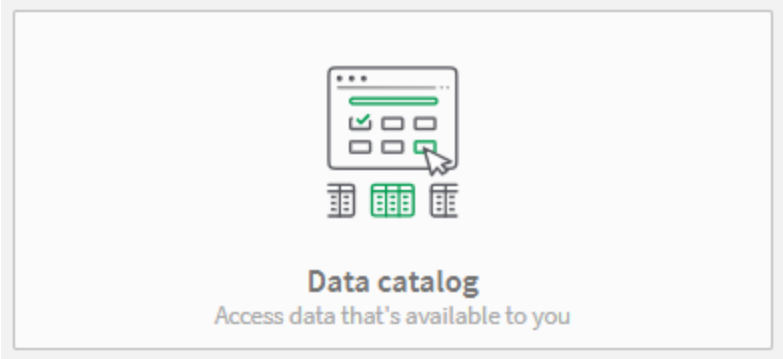
## 3.2 Connect Data with Qlik Sense

To connect data with Qlik sense we have Three Ways

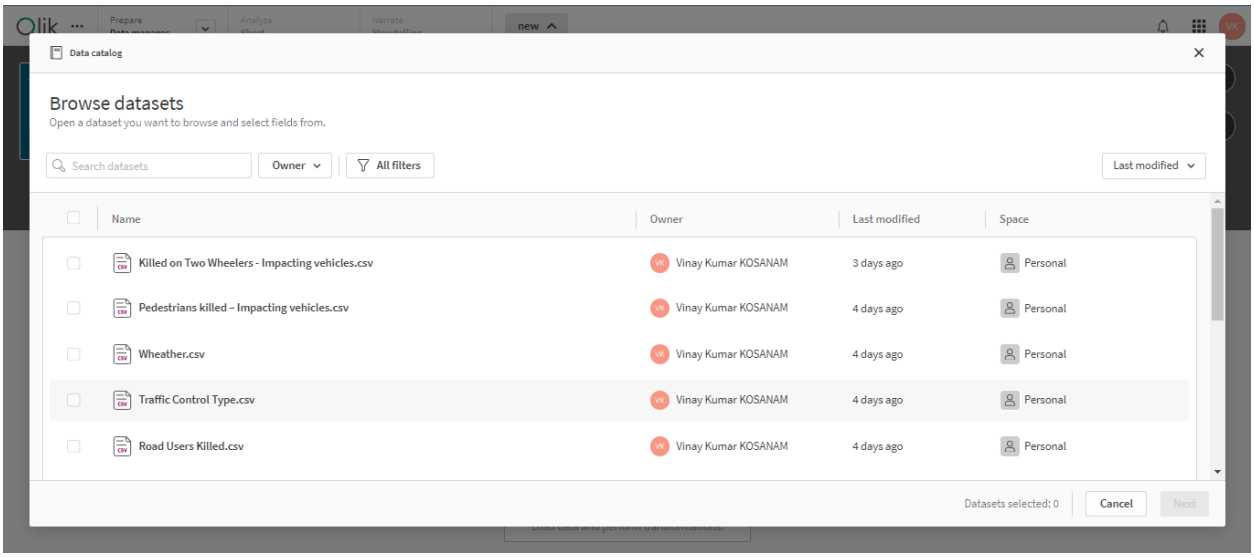
- Data Catalog
- Files and other sources
- Data load editor

### Data Catalog :

Data catalog is used to Acess data that's available to you in the Qilk Cloud

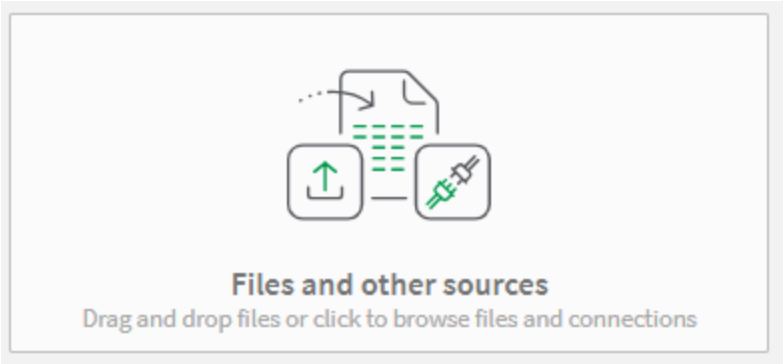



After clicking this we goto the Data catalog which contains data in different formats. Here we select the data which is necessary to our Analytics

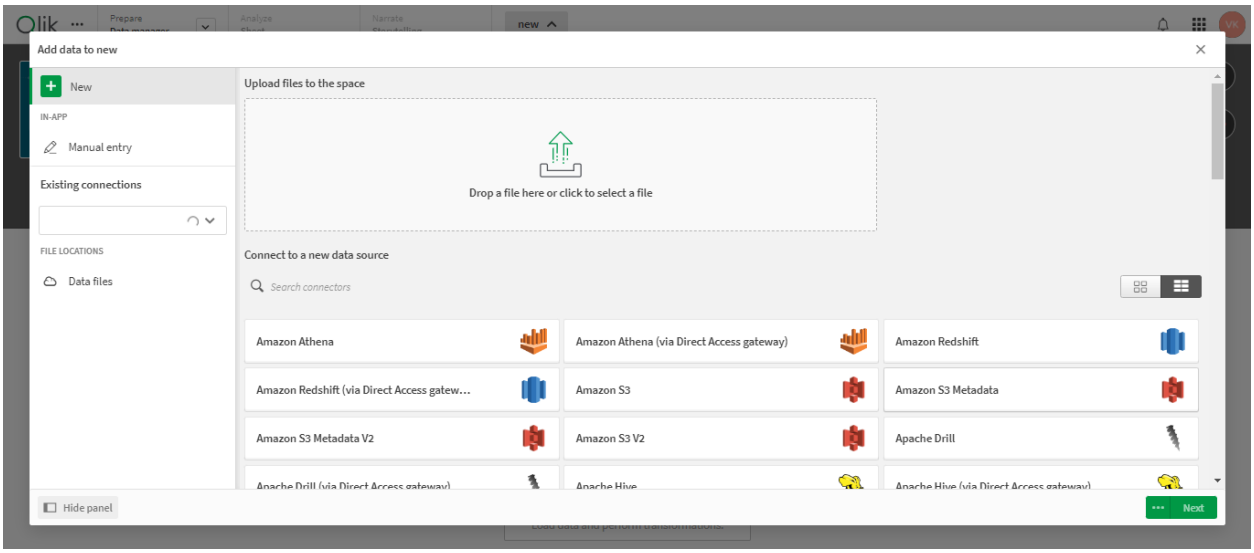


**Files and other sources :**

Files and other sources is used to upload files from External Sources.

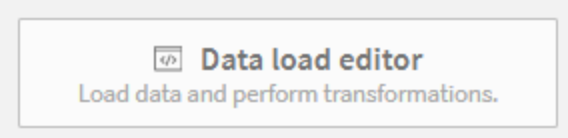



After clicking this we goto the Device Floders which contains data.
Here we select the data which is necessary to our Analytics.

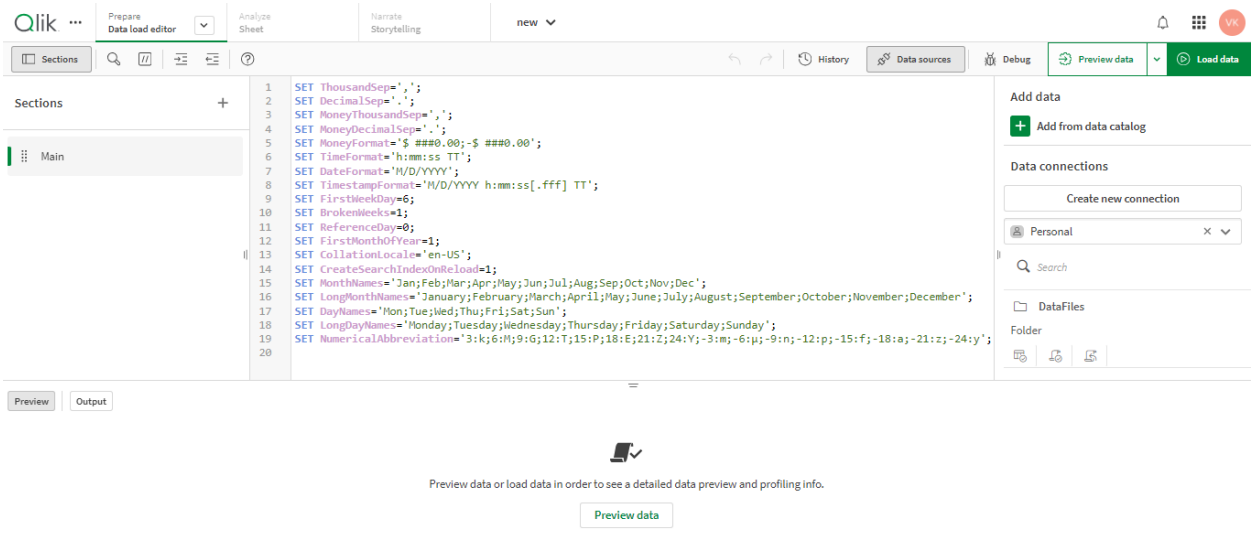



**Data load editor :**

Data load editor is used to load the data Manually and perform Transformations



After clicking this editor will open, here we Manually load the data and perform transformations with coding.



# 4 Data Preparation

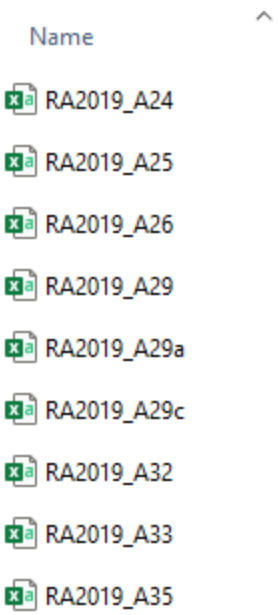
## 4.1 Prepare the Data for Visualization

Preparing data for visualization is a crucial step that involves several meticulous processes to ensure the final visualizations are insightful and accurate. This preparation includes:

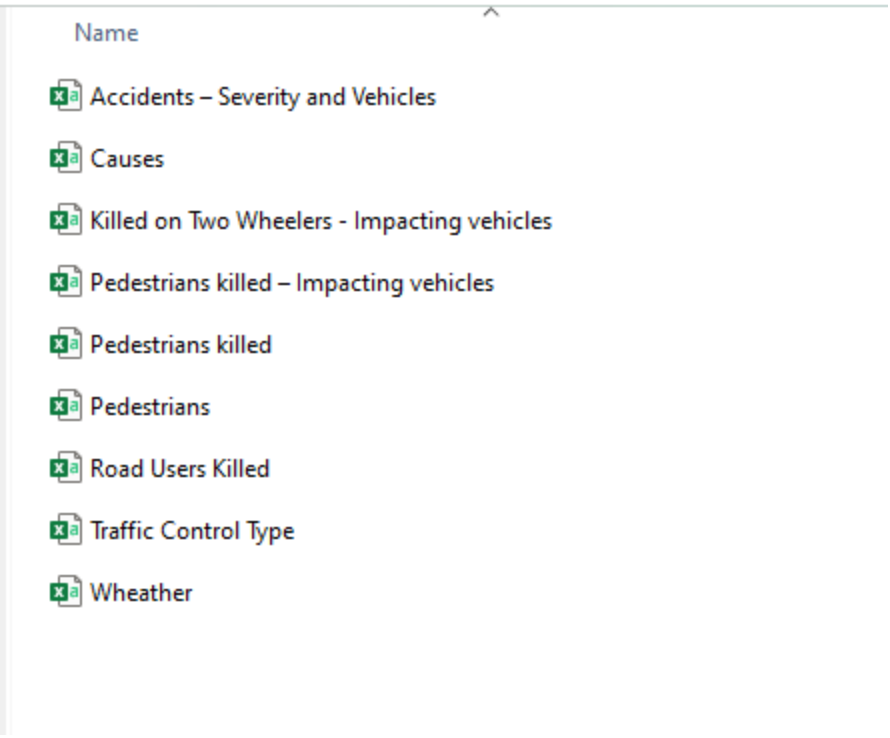
- Integrating Data Sources
- Cleaning the Data
- Transforming the Data
- Filtering the Data
- Calculating additional metrics
- Exploring the Data(EDA)

In files data is structured as Tables. Preparing is also called Preprocessing. After Preprocessing data , Each file mentioned with unqiue Name as shown in below image.

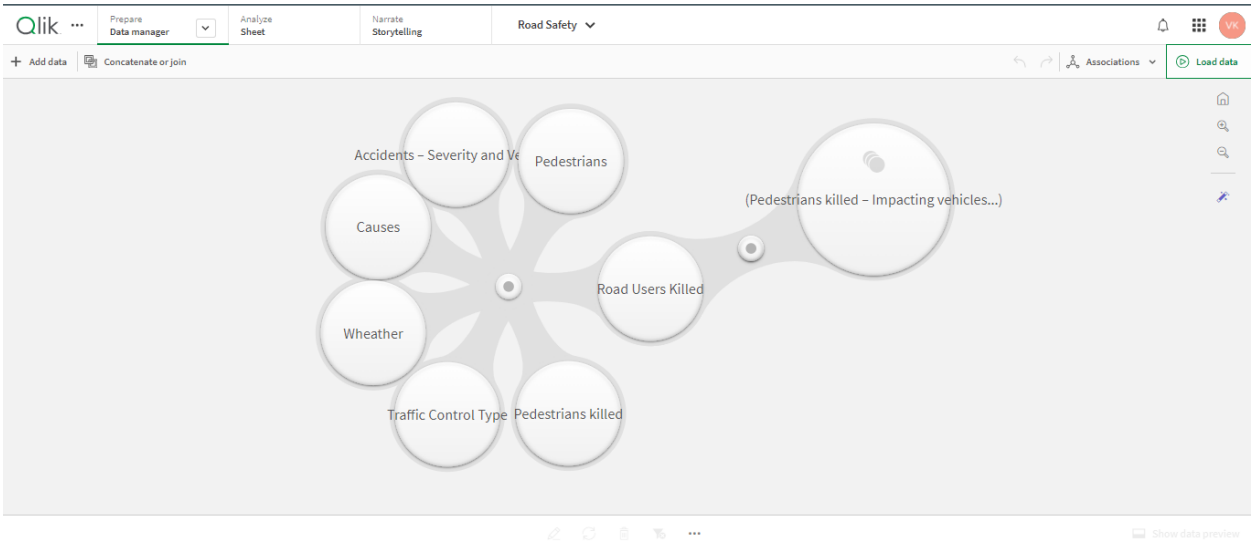
**Before Preprocessing**



**After Preprocessing :**




Processed data files are uploaded into anlaytic app i.e named as Road Satefy. After uploading we goto Data Manager, here tables represent as bubbles. Then we apply Associations between bubbles which have Similarites.



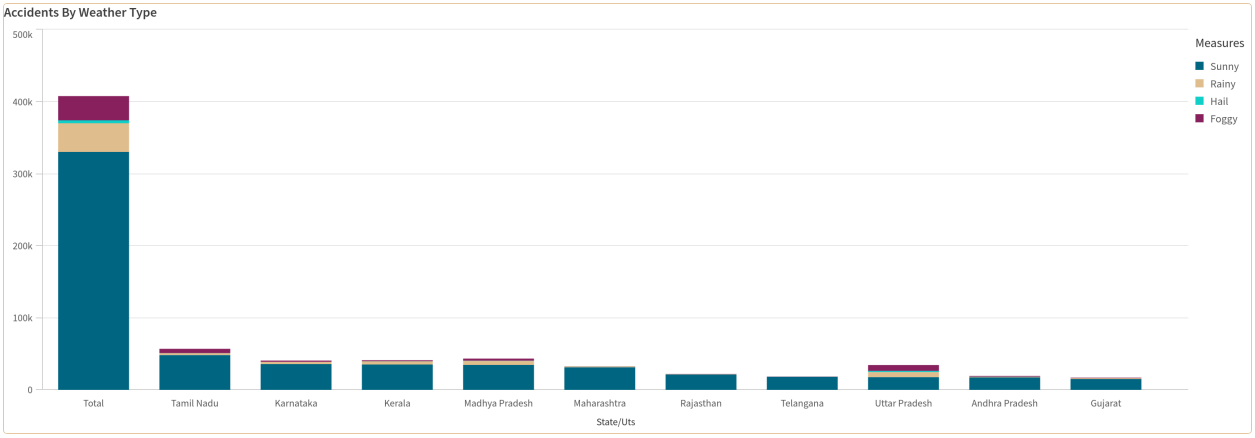
# 5 Data Visualizations

## 5.1 Visualizations

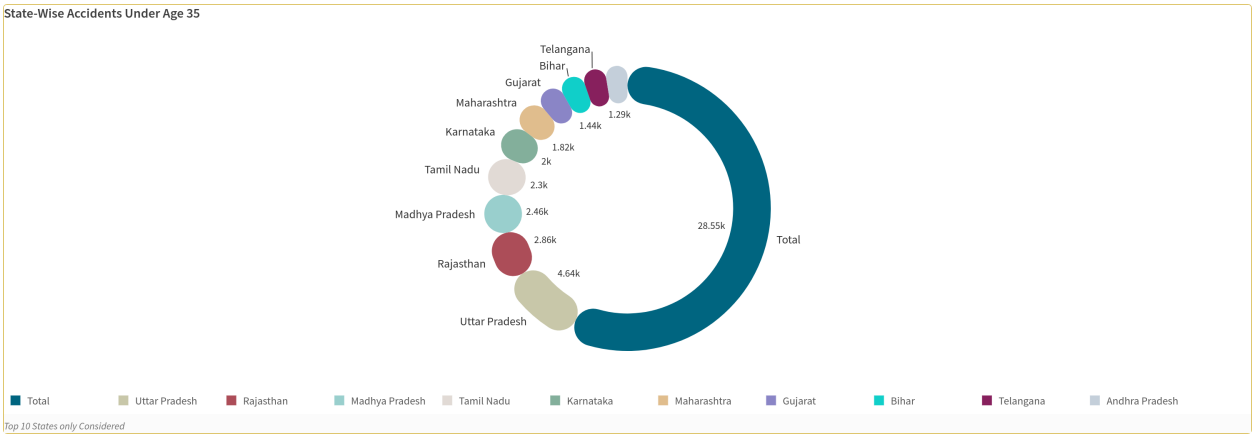
The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualization can help people identify patterns, trends, and outliers quickly in the data.

**Visualizations of Road Safety and Accident Patterns :**

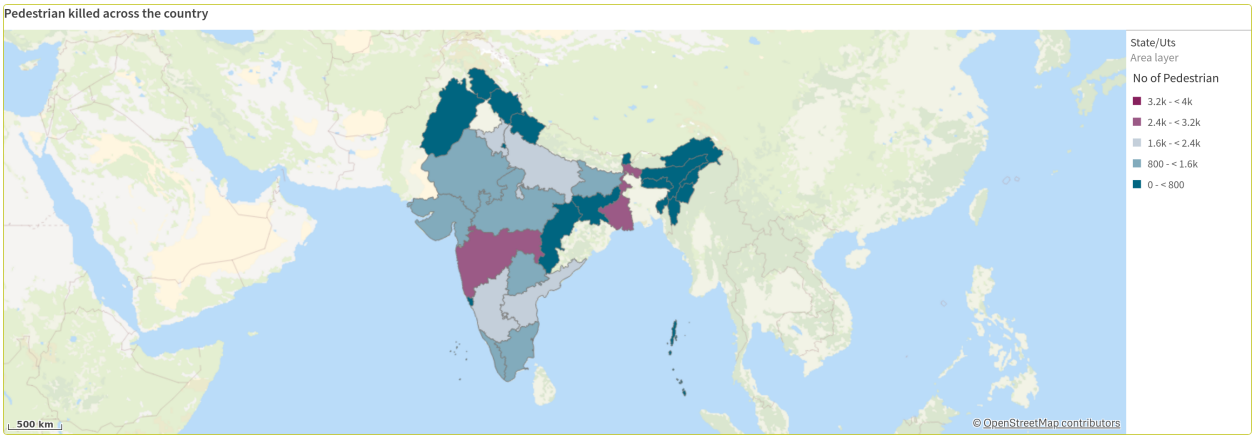
Activity 1.1: Accidents By Weather Type



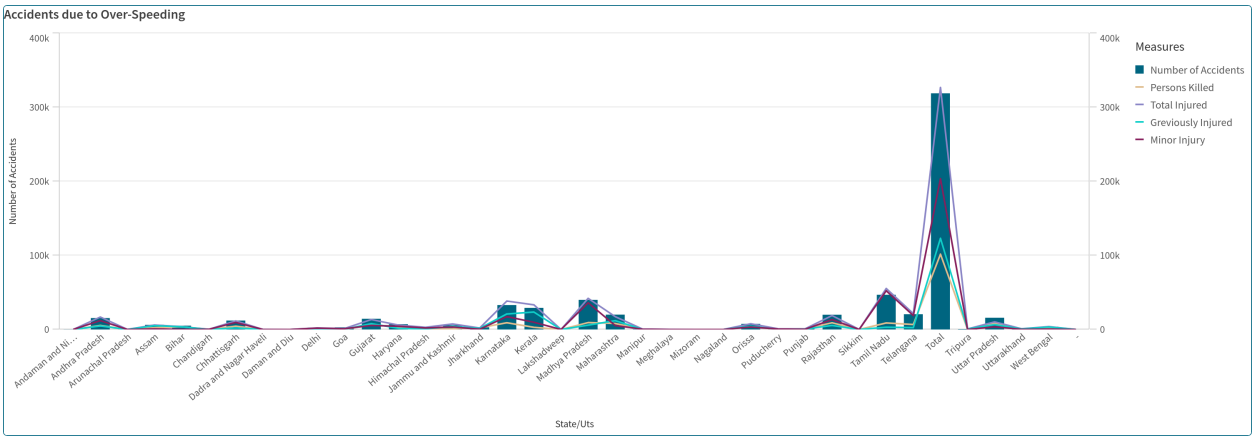

Activity 1.2: State-Wise Accidents Under Age 35



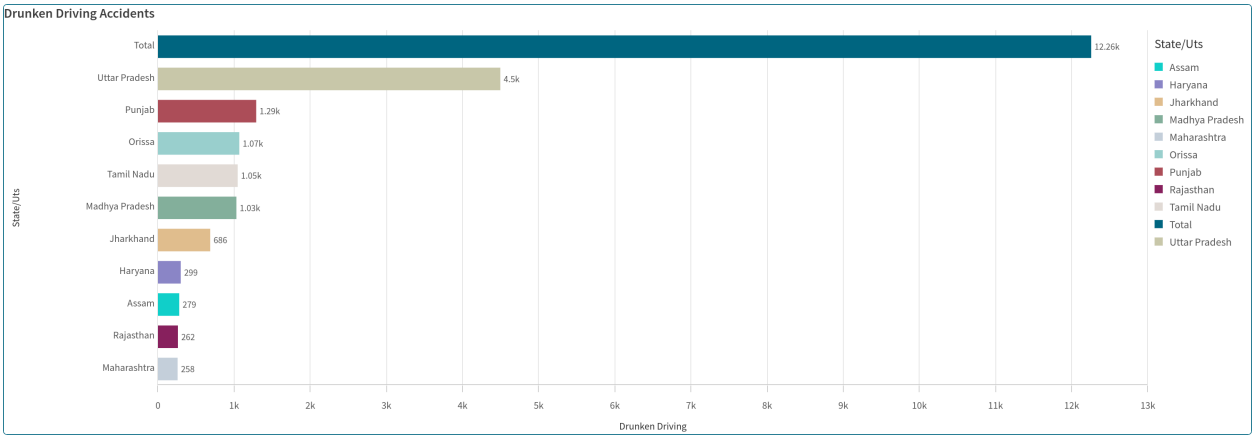
## Activity 1.3: Pedestrian Killed across the country



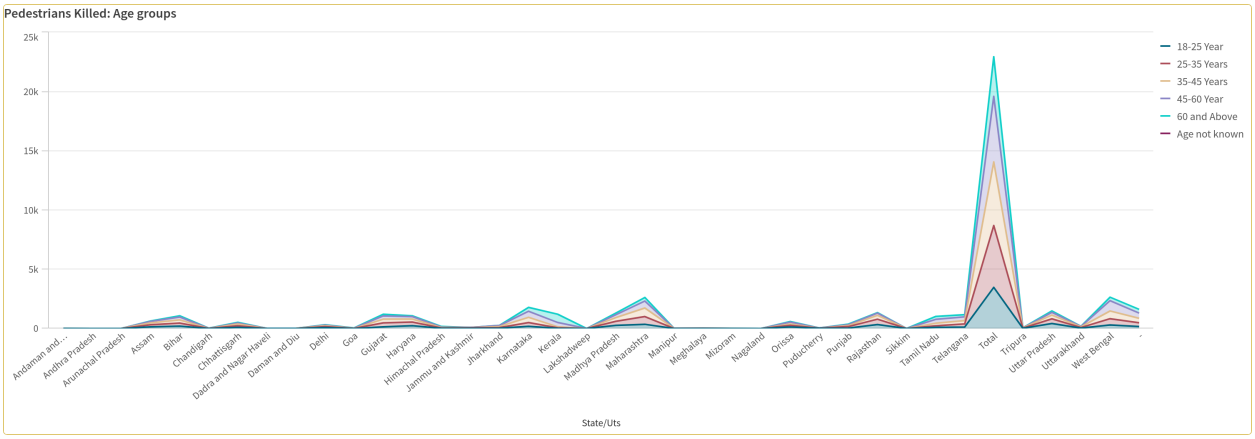

## Activity 1.4: Accidents due to Over-speeding



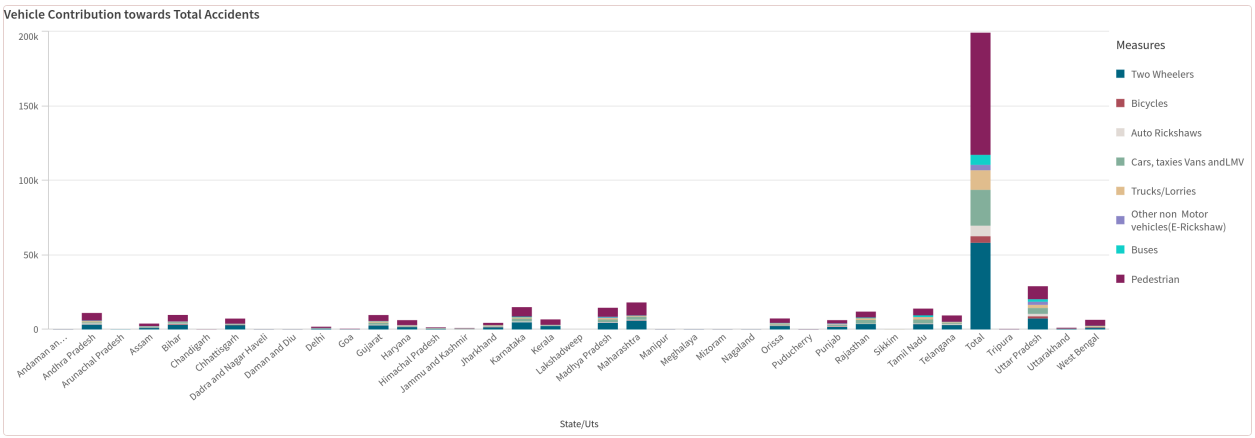

## Activity 1.5: Drunken Driving Accidents



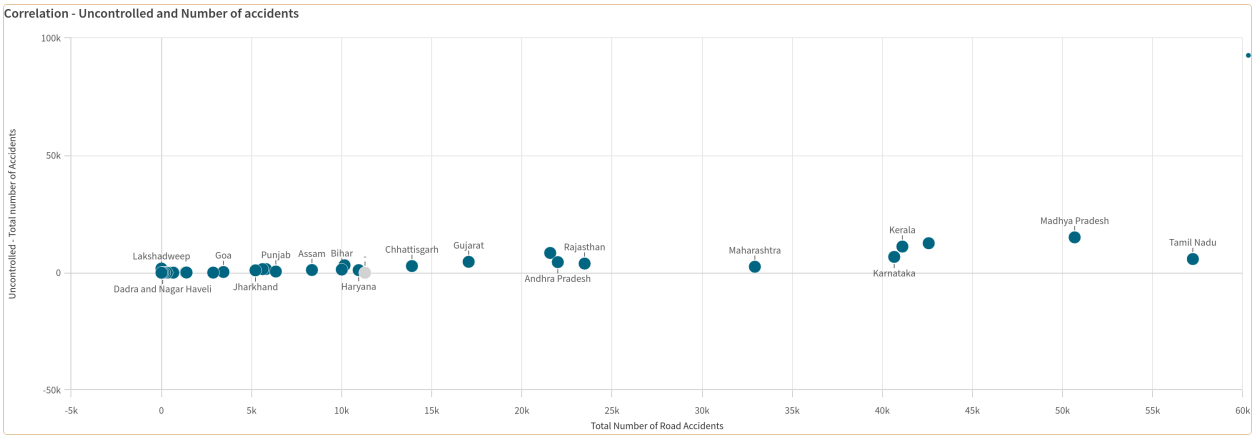
## Activity 1.6: Pedestrians Killed: Age Groups




## Activity 1.7: Vehicle Contribution towards Total Accidents




## Activity 1.8: Correlation - Uncontrolled and Number of Accidents

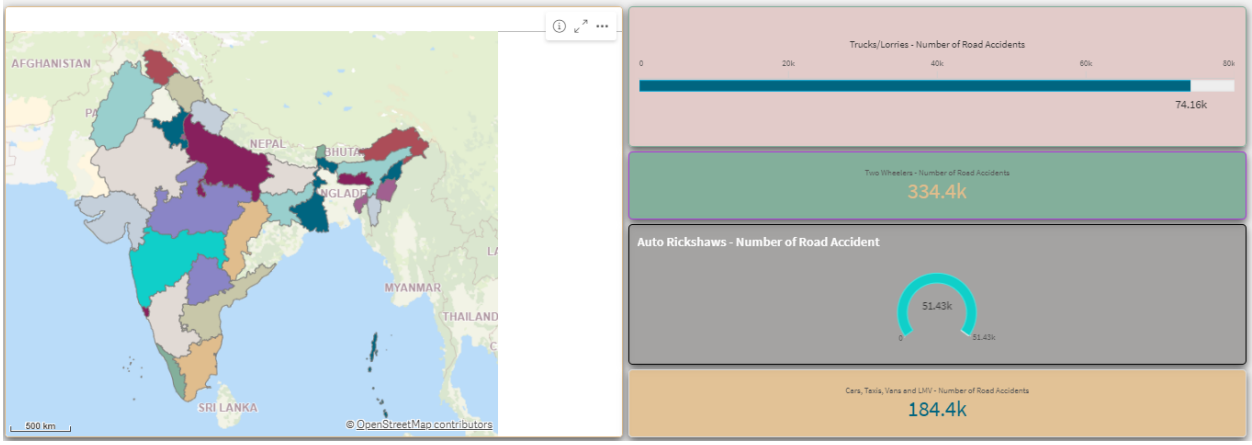


# 6 Dashboard

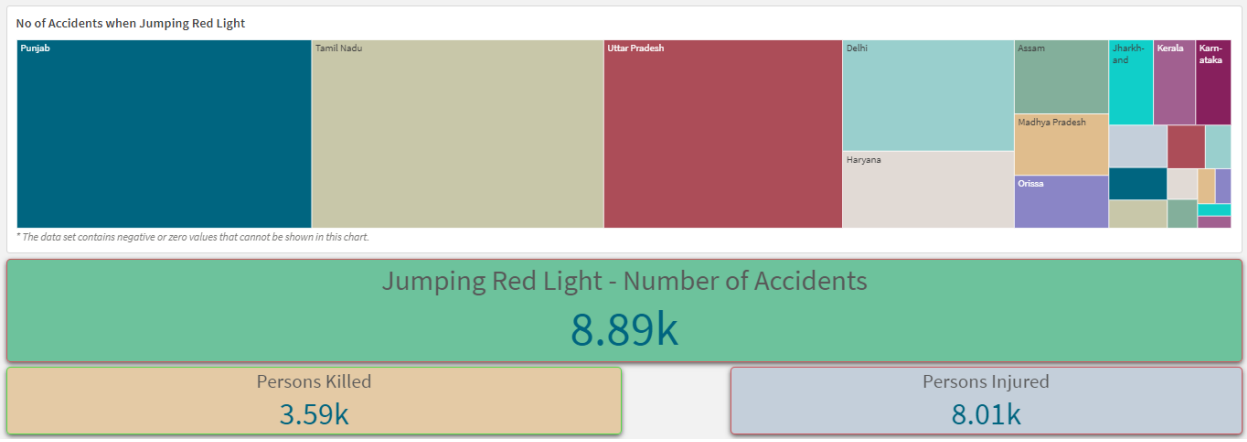
Dashboards are versatile tools that help in tracking, monitoring, and analyzing data across different industries. They provide a visual and interactive way to present information, making it easier to make informed decisions based on real-time data. Dashboards are powerful tools used for real-time monitoring and analysis of data. They are typically designed with a specific purpose or use case in mind.

## 6.1 Responsive and Design of Dashboard

Dashboard: Vehicle Impact-Related Accidents




Dashboard: Accidents when Jumping Red Light

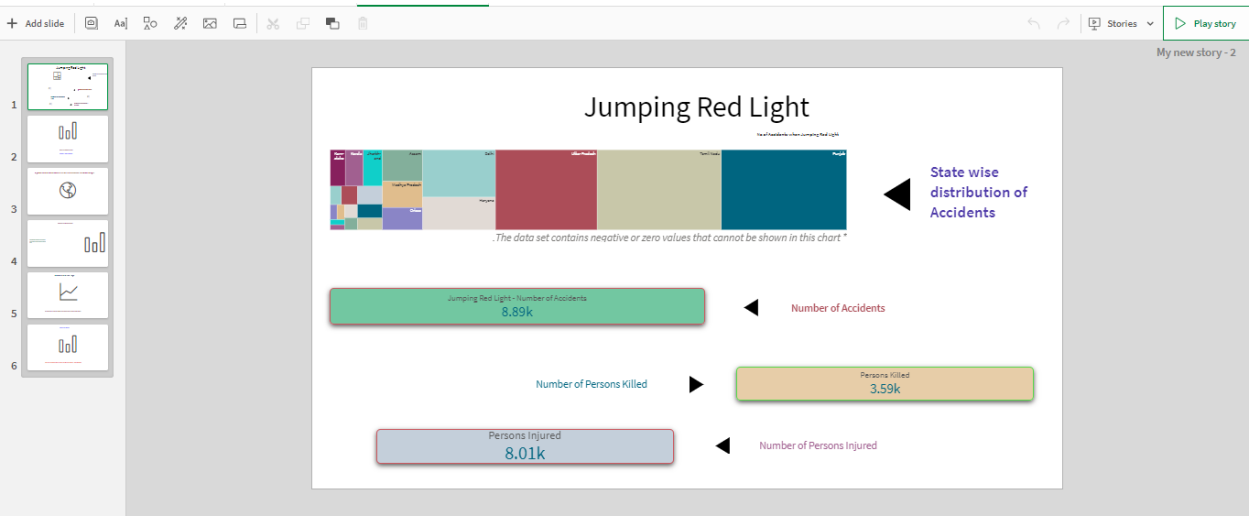


# 7 Report

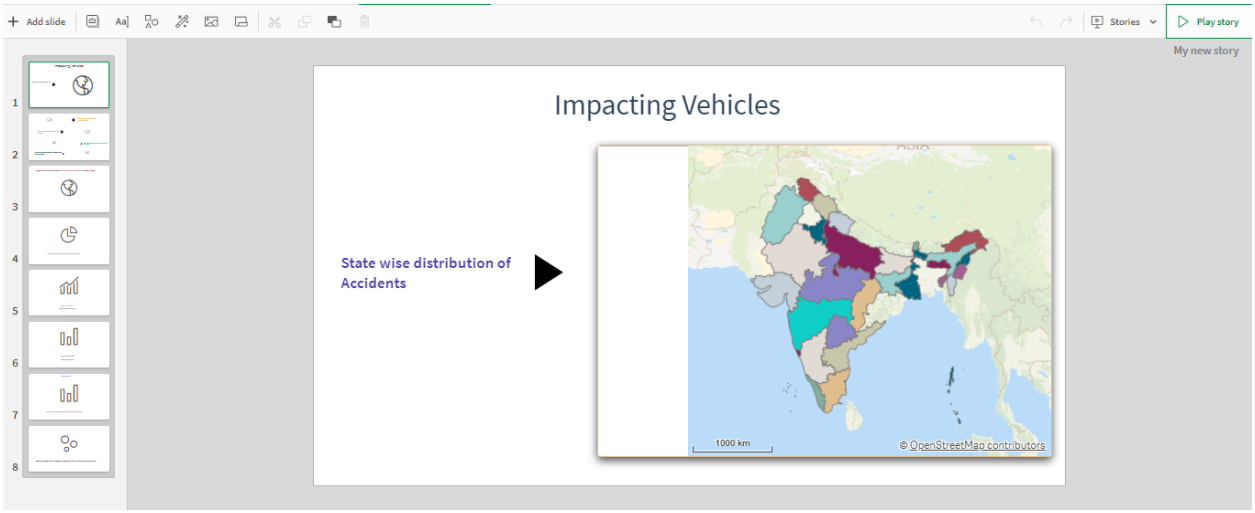
## 7.1 Report Creation

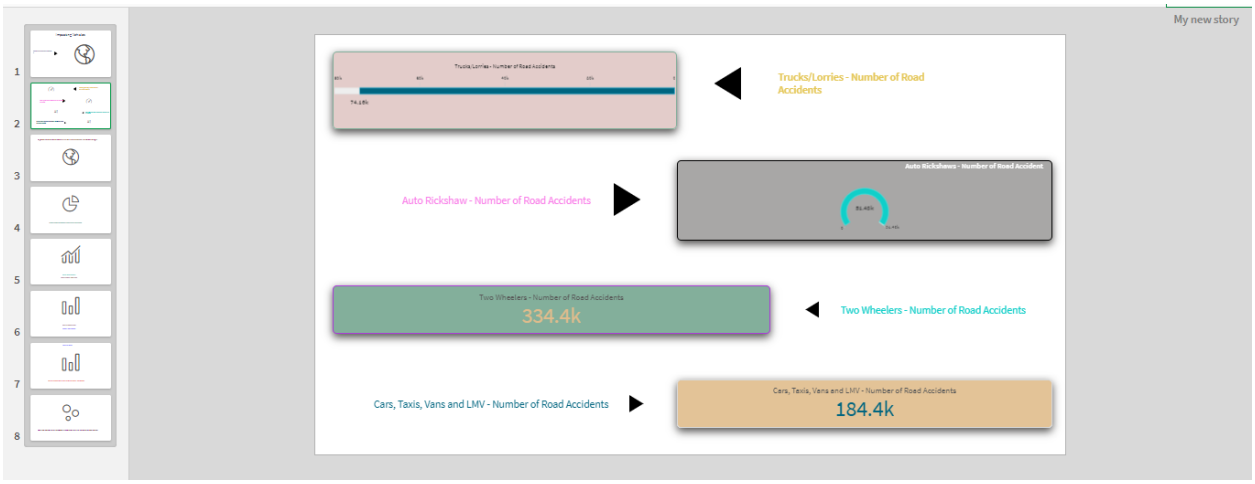
A report is a document or a statement that presents information in an organized format for a specific audience and purpose. A report is an account of information and evidence that has been applied to a particular product or matter and is usually organized and presented to a specific audience for a sole purpose.

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.

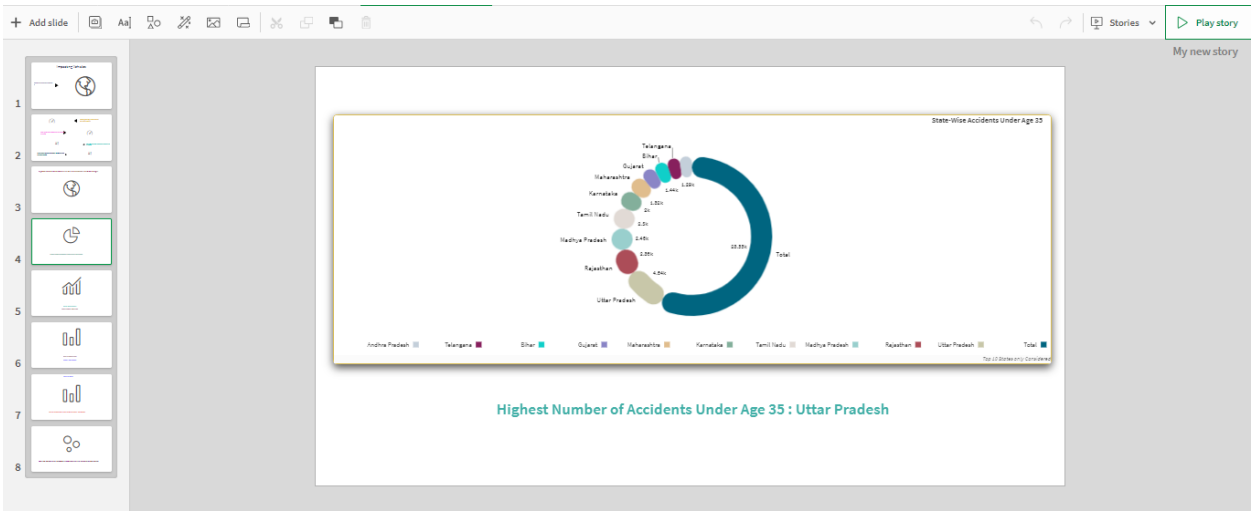



This shows how many persons killed, injured and no of Accidents because of Jumping Red light according to State-Wise.

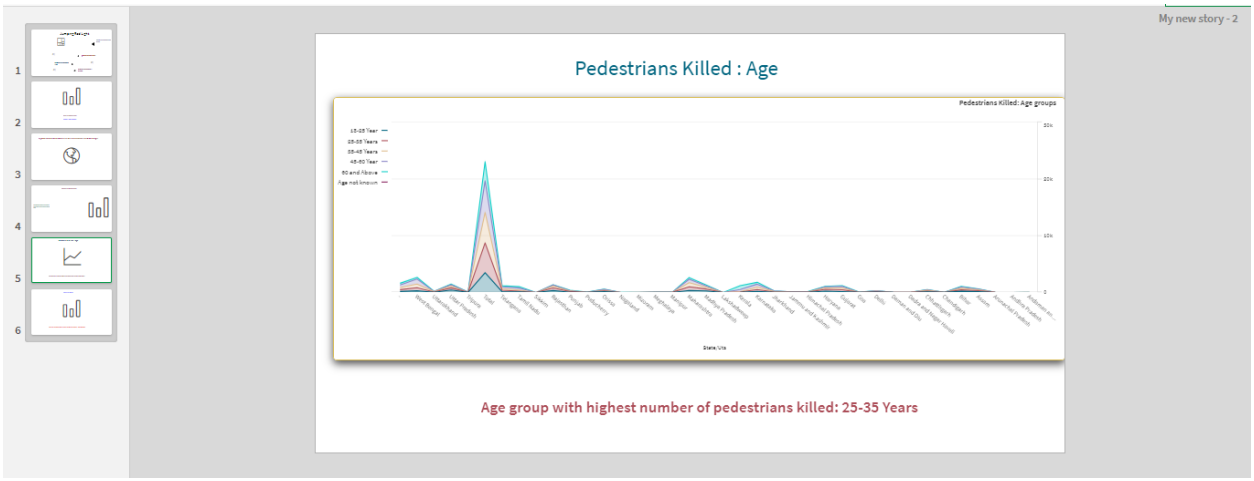



Above two slides shows that how many persons involved in Accidents with impacting Different types of Vehicles according to State - Wise.




This slide represents that Number of State-wise Accidents of Persons having Age under 35.




This slide represents that Number of Pedestrians killed categorized with Age Group

# 8 Performance Testing

## 8.1 Amount of Data Rendered

There are Nine different type of files contains 40 KiloBytes of data. Each file has different size as shown below :

- Accidents - Severity and Vehicles (8 KB)
- Causes  (6 KB)
- Killed on Two Wheelers - Impacting vehicles  (2 KB)
- Pedestrians killed - Impacting vehicles  (2 KB)
- Pedestrians killed  (3 KB)
- Pedestrians  (3 KB)
- Road Users Killed  (5 KB)
- Traffic Control  (6 KB)
- Type Wheather (5 KB)

**Number Of Graphs/ Visualizations**

1. Accidents By Weather Type
2. State-Wise Accidents Under Age 35
3. Pedestrian Killed across the country
4. Accidents due to Over-speeding
5. Drunken Driving Accidents
6. Pedestrians Killed: Age Groups
7. Vehicle Contribution towards Total Accidents
8. Correlation - Uncontrolled and Number of Accidents

## 8.2 Utilization 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.