Qlik Analysis Of Road Safety
And Accident Patterns
In India

Using Qlik Sense
Smart Internz

Mentor's Name : Anwar
Duration : 30 Days

Project by
Name: M Priyanka
B.TECH CSE (4TH YEAR)
AT
SRM Institute of Science
and Technology

1. Introduction

Road safety and accident trends in India will be studied using Qlik's data analytics technology. The project aims to discover trends, hotspots, and variables leading to road accidents by using diverse data sources such as meteorological conditions, demographic information, road infrastructure specifications, traffic statistics, and accident reports. Government officials, transportation agencies, and road safety groups may all benefit from this report by using it to inform data-driven choices that will increase road safety, decrease accident rates, and ultimately save lives.

1.1 Overview

This project intends to make use of the sophisticated data analytics platform that Qlik provides in order to investigate the patterns of road safety and accidents that occur throughout India. The purpose of the project is to discover insights and patterns that may be used to advise successful road safety actions.

This will be accomplished by combining a variety of data sources, such as traffic statistics, accident reports, meteorological conditions, roadway infrastructure details, and demographic information.

In order to minimize the number of accidents and save lives, the ultimate objective is to provide assistance to various stakeholders, such as government authorities, transportation agencies, and organizations concerned with road safety, in making choices based on data.

Objectives

- The purpose of this is to identify accident hotspots and give insights that may be put into action for targeted improvements.
- To conduct an analysis of patterns and recurrent variables that contribute to accidents on the road.
- For the purpose of developing predictive models for the purpose of projecting probable accident situations and enabling pre-emptive steps to be taken.
- The objective is to improve the general understanding of the dynamics of road safety in India in order to facilitate improved policymaking and implementation.
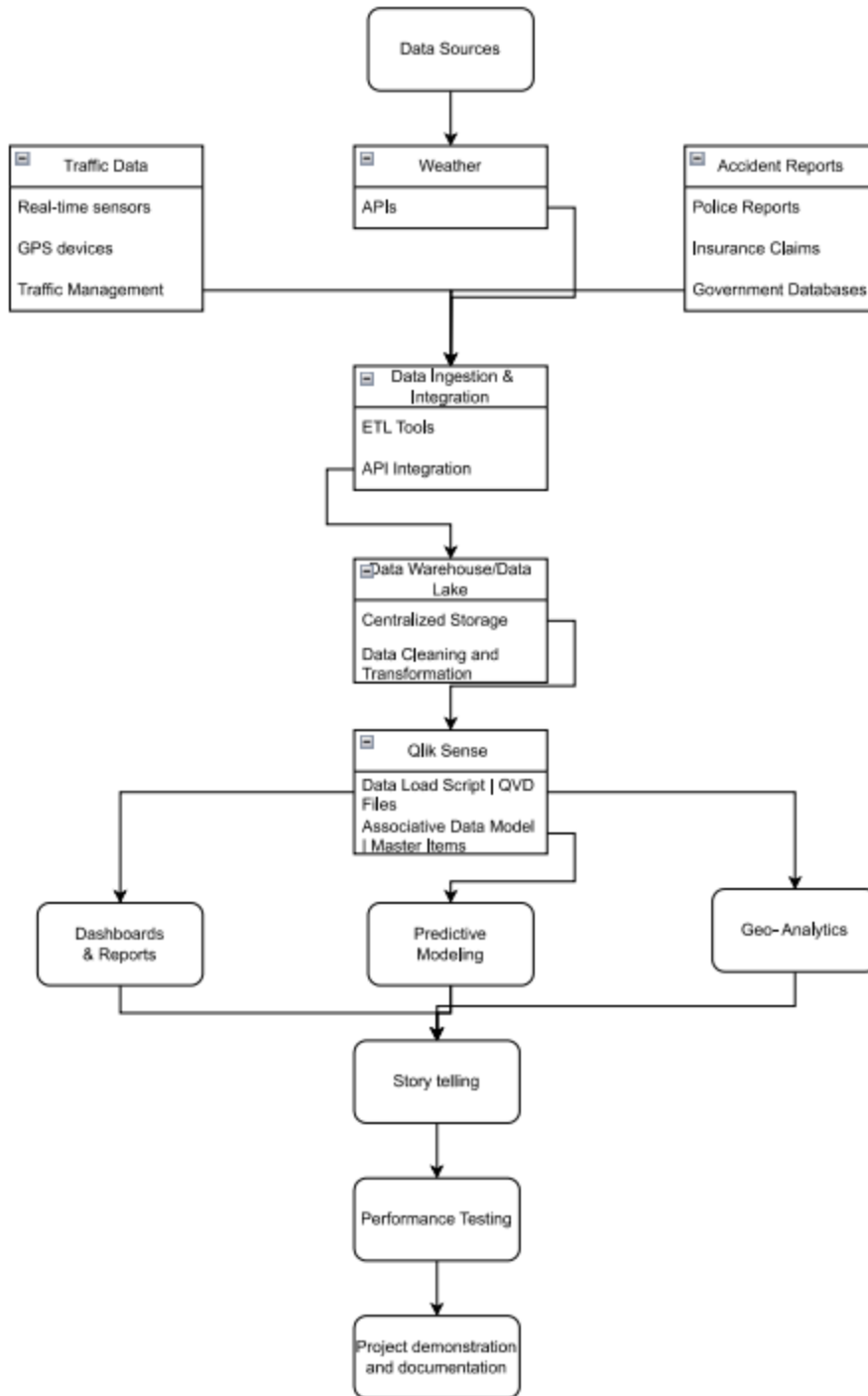
1.2 Purpose

There are some places or highways in India that have a high accident rate, and the analytics provided by Qlik can highlight these areas. It is possible for the platform to identify areas that are prone to accidents by combining accident data with other criteria such as the volume of traffic, the state of the roads, and the time of day. This information is essential for the implementation of targeted interventions such as increased traffic monitoring, improved signage, and modifications to speed limits.

Within the context of historical accident data, Qlik is able to undertake trend analysis in order to recognize trends and

elements that are a recurrent cause of accidents. One example of this is the examination of different sorts of accidents, such as collisions and pedestrian accidents, as well as seasonal fluctuations and driver behavior, such as speeding and distracted driving. For the purpose of addressing the underlying causes, the insights that are gathered can be used to direct awareness campaigns, driver training programs, and regulatory reforms.

With the use of predictive analytics, Qlik is able to make predictions about probable accident situations based on the data inputs that are real-time. The platform is able to give early warnings and preventative steps to prevent accidents by taking into consideration elements such as weather forecasts, patterns of traffic flow, and past accident trends. Authorities are given the flexibility to strategically deploy resources and take preventative safety measures as a result of this powerful predictive capacity. The purpose of this project is to create an organized pathway for studying road safety and accident trends in India using Qlik Sense. This will ensure that a complete and impactful strategy is used to improve road safety measures.

1.3 Technical Architecture

```
                          ┌─────────────────┐
                          │   Data Sources   │
                          └────────┬────────┘
                                   │
                                   ▼
┌───────────────────┐   ┌───────────────────┐   ┌───────────────────┐
│ ▭   Traffic Data   │   │ ▭     Weather      │   │ ▭  Accident Reports│
├───────────────────┤   ├───────────────────┤   ├───────────────────┤
│ Real-time sensors  │   │ APIs               │   │ Police Reports     │
│                    │   │                    │   │                    │
│ GPS devices        │   │                    │   │ Insurance Claims   │
│                    │   │                    │   │                    │
│ Traffic Management │   │                    │   │ Government Databases│
└───────────────────┘   └───────────────────┘   └───────────────────┘
                                   │
                                   ▼
                          ┌───────────────────┐
                          │ ▭ Data Ingestion & │
                          │     Integration    │
                          ├───────────────────┤
                          │ ETL Tools          │
                          │                    │
                          │ API Integration    │
                          └───────────────────┘
                                   │
                                   ▼
                          ┌───────────────────┐
                          │ ▭Data Warehouse/Data│
                          │       Lake          │
                          ├───────────────────┤
                          │ Centralized Storage │
                          │                     │
                          │ Data Cleaning and   │
                          │ Transformation      │
                          └───────────────────┘
                                   │
                                   ▼
                          ┌───────────────────┐
                          │ ▭     Qlik Sense   │
                          ├───────────────────┤
                          │ Data Load Script | QVD│
                          │ Files              │
                          │ Associative Data Model│
                          │ | Master Items     │
                          └───────────────────┘
```

| Dashboards & Reports | Predictive Modeling | Geo-Analytics |

```
                          ┌───────────────────┐
                          │    Story telling   │
                          └───────────────────┘
                                   │
                                   ▼
                          ┌───────────────────┐
                          │ Performance Testing │
                          └───────────────────┘
                                   │
                                   ▼
                          ┌───────────────────┐
                          │Project demonstration│
                          │ and documentation   │
                          └───────────────────┘
```

2. Problem Understanding

India is confronted with a huge issue in terms of road safety, since the country has high accident rates that result in a considerable number of fatalities and injuries each year. The primary challenge consists of locating accident hotspots, gaining a knowledge of demographic and behavioral characteristics, and evaluating the quality of the roads and the surroundings. Through the utilization of Qlik Sense, the overall objective of this project is to collect and integrate data from a wide variety of sources, carry out in-depth analysis, and provide interactive representations. The purpose of this project is to produce insights that may be put into action for the purpose of targeted interventions, informed decision-making, and enhanced road safety measures, with the ultimate objective of reducing accidents and improving public confidence.

2.1 Business problem

Both travel times and distances have been drastically cut down as a result of technological improvements in the transportation sector. It is estimated that each year in India, road accidents result in the loss of lakhs of lives and cause major injuries to crores of people. Due to the significant number of accidents that occur on the roads, there is a pressing need for a comprehensive investigation to determine the factors that are contributing to the problem and the patterns that are emerging.   .

 The first challenge for the company is to do a thorough investigation into the patterns of road safety and accidents that have occurred all across India in order to identify the causes that have contributed to these accidents and the locations that are considered to be high-risk. The purpose of this study is to provide insights that can be put into action, which may then be used to inform and guide plans for improving road safety, lowering the number of accidents reported.

2.2 Business requirements

The analysis aims to provide valuable insights into user demographics, accident patterns, and problem areas related to road safety in India using Qlik Sense. The primary focus is on creating interactive and visually compelling dashboards to support strategic planning and operational improvements. By integrating data from multiple sources, including traffic data, accident reports, weather conditions, road infrastructure, and demographic information, the project will enable a comprehensive understanding of the factors contributing to road accidents.

The analysis will identify accident hotspots, perform trend analysis, and develop predictive models to forecast potential accident scenarios. This data-driven approach will support stakeholders, including government authorities, transportation agencies, and road safety organizations, in making informed decisions, implementing better safety protocols, and ensuring compliance with regulations. The ultimate goal is to enhance road safety, reduce accidents, and save lives by providing actionable insights and facilitating targeted interventions based on the analysis.

2.3 Literature survey

There is a growing worry for public safety in India due to the rising number of road accidents, which calls for extensive research to be conducted in order to comprehend and address the factors that are contributing to the problem. The purpose of this literature review is to examine the research that has already been conducted on analysis of accidents and road safety, with a particular focus on the role that data analytics and visualization tools such as Qlik Sense play in solving this issue.

Speeding, driving while distracted, and driving under the influence of alcohol are the top behaviours that contribute to accidents, according to the World Health Organisation (2018).Road Conditions: According to Pathak (2016), the presence of poor road infrastructure, such as potholes and inadequate signs, drastically increases the likelihood of an accident occurring .In addition, the state of the vehicle and the kind of vehicle both play a significant part in the occurrence of accidents. Accidents are more likely to occur in cars that are both overloaded and poorly maintained, according to Kopits and Cropper (2005).Environmental factors: According to Brodsky and Hakkert (1988), the presence of weather conditions such as rain and fog is associated with an increased likelihood of accidents.

According to Haleem and Gan (2011), predictive analytics involves the utilization of historical data to forecast future accident hotspots and timings. This can be of assistance in the process of preventive planning and resource allocation. Identifying patterns and trends in accident data can help inspire policy changes and targeted actions (Zhang, 2015). Trend analysis is a technique that can be used to do this. Tools for Visualization: Platforms such as Qlik Sense make it possible to create interactive dashboards that make it easier to explore and make decisions on data in real time. The Swedish program known as "Vision Zero" makes use of data analytics in order to achieve large reductions in the number of deaths that occur on the roads. Furthermore, the initiative places an emphasis on data-driven interventions (Johansson, 2009).Improved Traffic Safety in New York City According to Weiss et al. (2017), the utilization of data analytics for the purpose of enhancing traffic safety in New York City has resulted in a striking reduction in the number of accidents that have occurred as a result of targeted enforcement and infrastructure enhancements.

The accuracy of analysis and predictions is hindered when the data is inconsistent and incomplete, according to Abdel-Aty and Haleem (2011). Data quality and availability are also important.The integration of data sources is a complicated process that needs sophisticated data integration techniques (Zhou et al., 2010). This is because combining data from several sources, such as traffic, weather, and demographic data, is a particularly difficult task. Scalability: It is essential to ensure that analytics systems are scalable in order to manage massive amounts of data in real time (Chen et al., 2014). This is especially important for making decisions in a timely manner.

3. Data collection

Data collection is the systematic process of gathering and measuring information from various sources to obtain a comprehensive and accurate dataset. This process is essential in research, analytics, and decision-making across numerous fields such as business, healthcare, social sciences, and environmental studies. Effective data collection ensures that the information gathered is relevant, reliable, and sufficient to address the specific research questions or business needs.

3.1 Collect the dataset

India is confronted with a huge issue in terms of road safety, since the country has high accident rates that result in a considerable number of fatalities and injuries each year. The primary challenge consists of locating accident hotspots, gaining a knowledge of demographic and behavioral characteristics, and evaluating the quality of the roads and the surroundings. Through the utilization of Qlik Sense, the overall objective of this project is to collect and integrate data from a wide variety of sources, carry out in-depth analysis, and provide interactive representations. The purpose of this project is to produce insights that may be put into action for the purpose of targeted interventions, informed decision-making, and enhanced road safety measures, with the ultimate objective of reducing accidents and improving public confidence.

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

3.2 Connect data with Qlik sense

- Log in to your Qlik Sense account.
- For Qlik Sense Desktop, open the application on your computer.
- For Qlik Sense Enterprise, log in through your web browser.
- Navigate to the Qlik Sense hub and click "Create a new app."
- Name the app (e.g., "Road Safety Analysis") and click "Create."
- Open the newly created app.
- Click on "Add Data" to start connecting your data sources.
- Click on "Data files" and upload files from local or cloud storage (e.g., Google Drive, Dropbox).
- Qlik Sense will display available tables and fields from the connected data source.
- Select the relevant tables and fields required for the analysis.
- Click "Load data" to import the data into Qlik Sense.

4. Data preparation

Making ensuring that the data that is put into Qlik Sense is clean, consistent, and suitable for analysis is an essential step that must be taken before the data can be loaded. Initially, the process begins with the cleaning of the data, which includes the elimination of duplicate entries, the management of missing information by either filling them in or omitting them, the standardization of data formats, and the correction of any entry mistakes. The next step is the transformation of the data, which includes the normalization of the data, the aggregated data at appropriate levels, the introduction of additional computed fields, and the merging of datasets in order to provide a comprehensive dataset. Creating a cohesive data model is an essential part of the process of structuring the data.

This is accomplished by defining relationships between the various tables and establishing critical fields for connecting. The Data Load Editor in Qlik Sense is where load scripts are created in order to load and convert data. This editor is responsible for implementing the transformations and preparations that are described above. Immediately following the execution of these scripts, the data is refreshed into Qlik Sense. In the subsequent step, the data is validated by doing spot checks and consistency tests in order to guarantee that it is accurate and consistent. Lastly, for the sake of future reference, comprehensive documentation of all conversions, cleaning methods, and metadata is provided.

By going through this extensive preparation procedure, you can be certain that the data that is utilized in Qlik Sense is accurate, trustworthy, and suitable for the generation of useful insights and visualizations.

Qlik ···
Prepare
Data manager

Analyze
Sheet

Narrate
Storytelling

Road Safety ∨

PM

+ Add data   Concatenate or join   ↶ ↷ ⚬ Associations ∨   ▷ Load data

Recommended associations   ? ✕

Total tables: 9
Unassociated tables: 0
Recommendations: 0

Preview all   Apply all

causes   Climate

safety   pedestrains killed

Accidents   Two wheelrs killed

Pedes invlo age & ge   ...es kill age&gender
Pedestrains Involved

To make associations manually, you can drag
one table onto another.

pedestrains killed 55.xlsx   Fields: 10

pedestrains killed.State/Uts	pedestrains ...	pedestrains ...	pedestrains ...	pedestrains ...	pedestrains ...	pedestrains ...	pedestrains ...	pedestrains ...	pedestrains ...
Andaman and Nicobar Island	0	3	0	1	1	0	1	0	6
Andhra Pradesh	0	534	164	322	288	126	0	289	1723
Arunachal Pradesh	8	0	0	0	0	0	0	0	8
Assam	0	269	43	212	75	47	2	70	718
Bihar	0	296	45	345	296	67	2	208	1259
Chandigarh	0	9	0	14	5	2	0	6	36

✎ ⟳ 🗑 ⛛ ···   ⌨ Hide data preview

5. Data visualization

Data visualization in Qlik Sense involves creating interactive and dynamic graphical representations of data to facilitate easier interpretation and insight generation. To begin, you prepare your dashboard by adding a new sheet and selecting appropriate visualization types, such as bar charts, line charts, pie charts, scatter plots, and maps, based on the nature of your data and the story you want to tell. Each visualization is configured by selecting relevant dimensions and measures, applying filters, and sorting data to highlight the most pertinent information. Customizing the appearance of visualizations with color schemes, labels, and tooltips enhances readability and context. Adding interactivity features like drill-downs, filters, and selections allows users to explore data in depth, and linking visualizations ensures that selections in one chart update other charts on the dashboard.

5.1
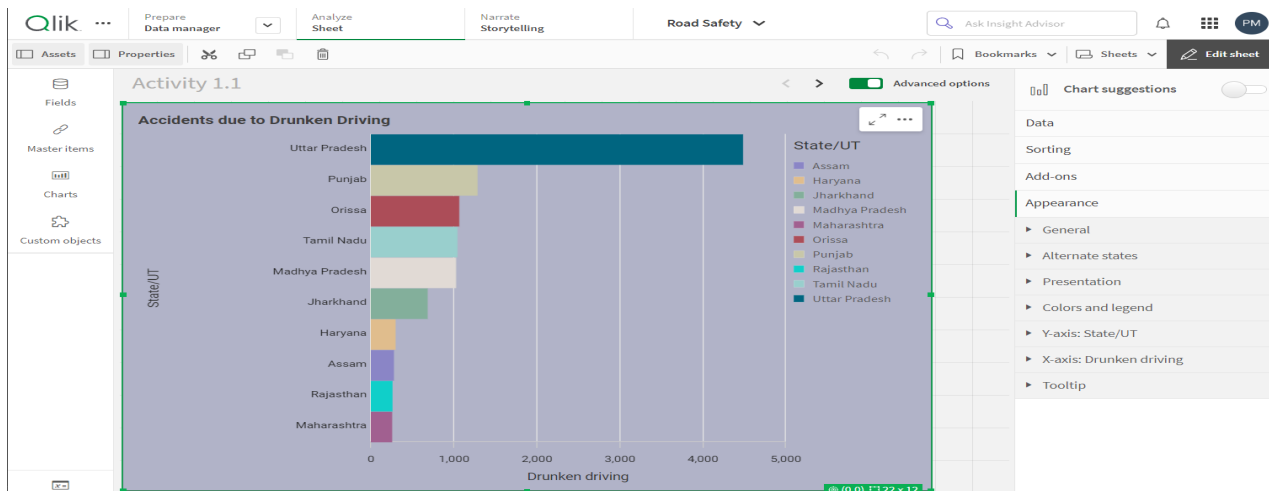
5.2
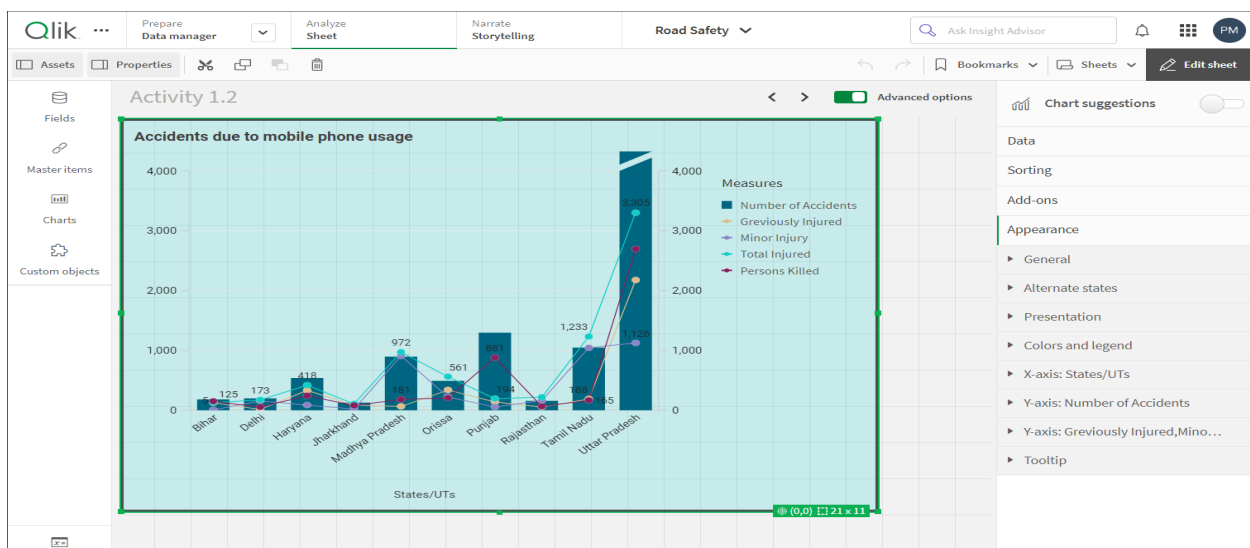
5.3
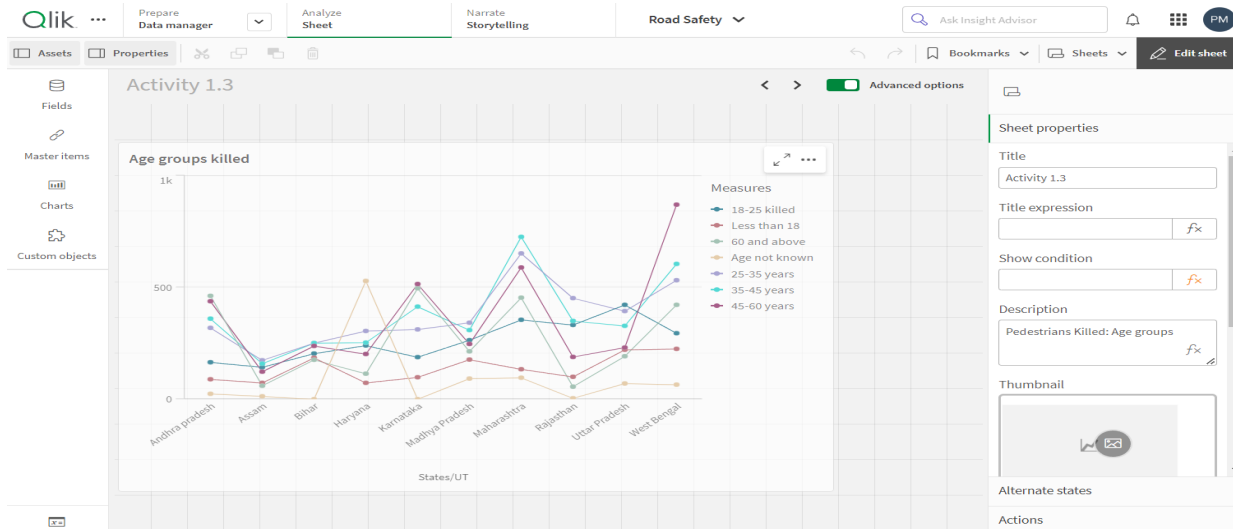
5.4
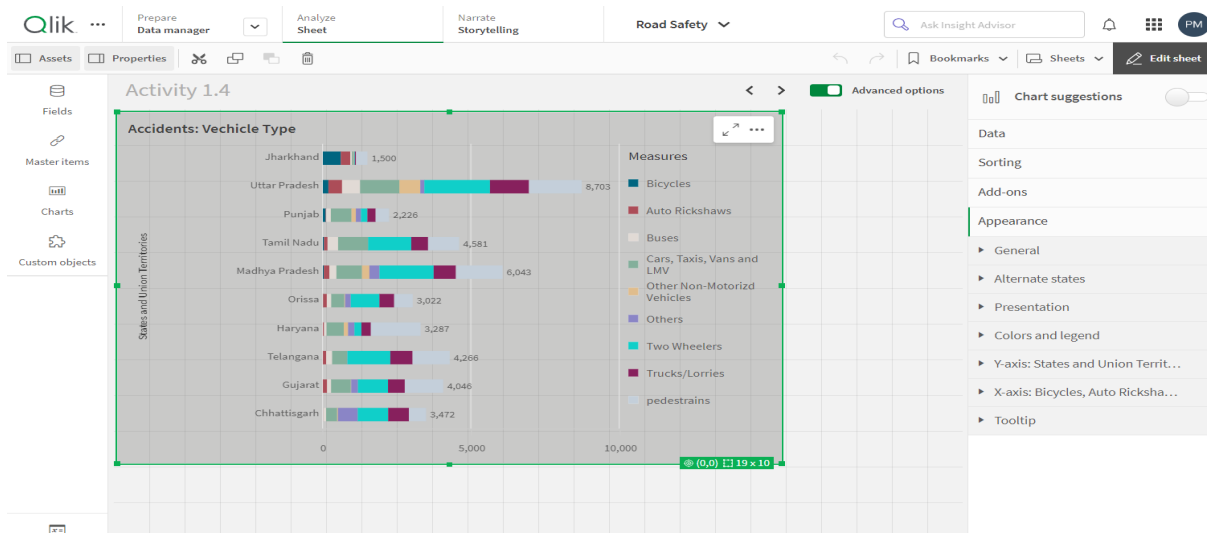
5.5
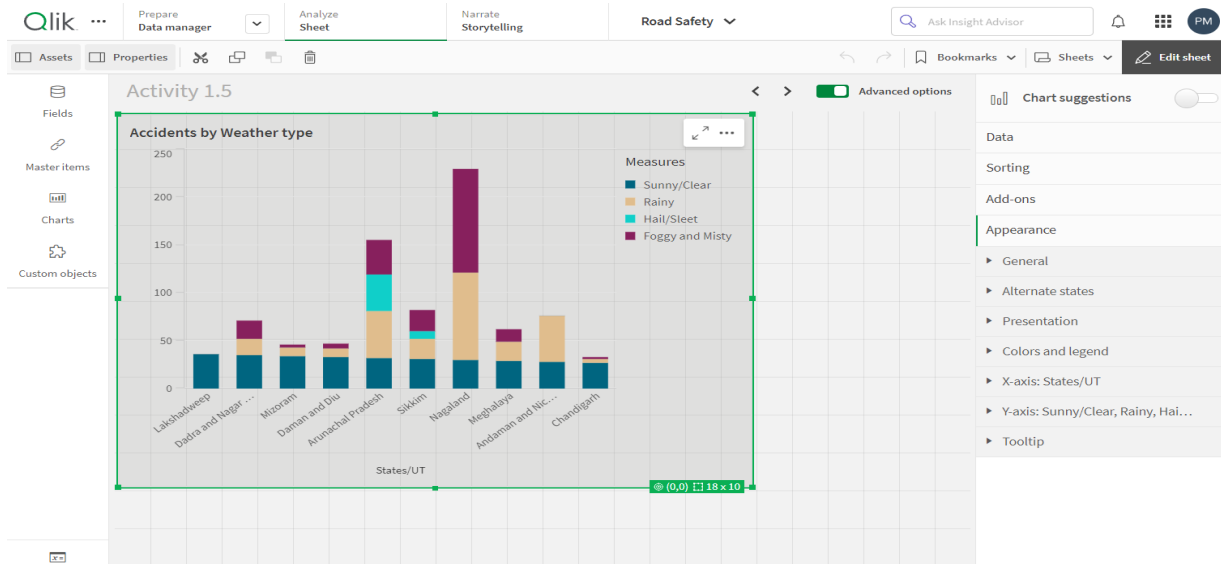
5.6
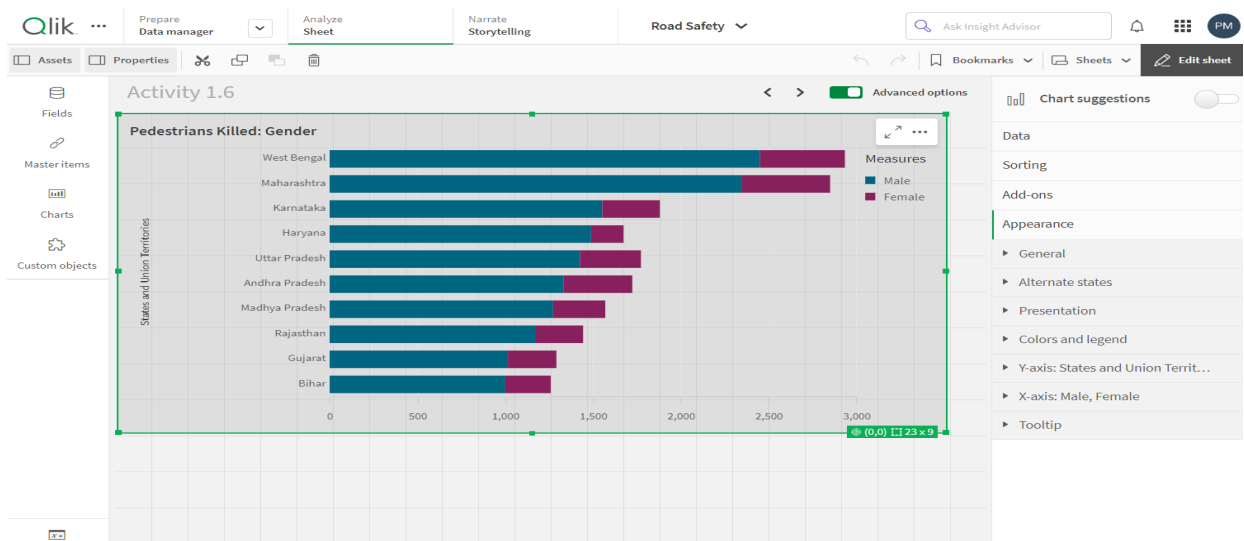
5.7
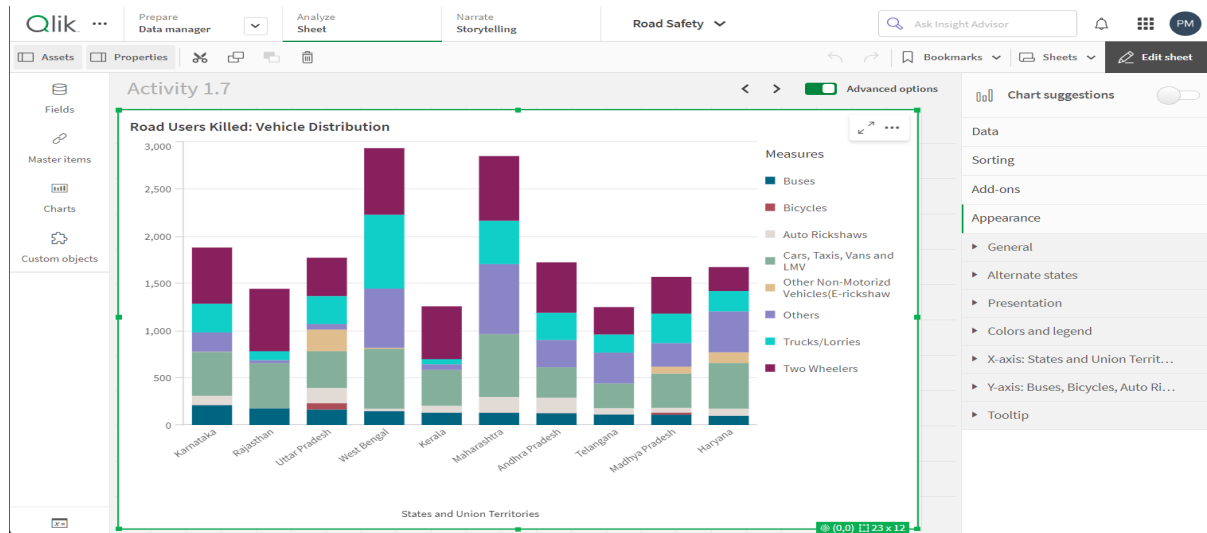
5.8
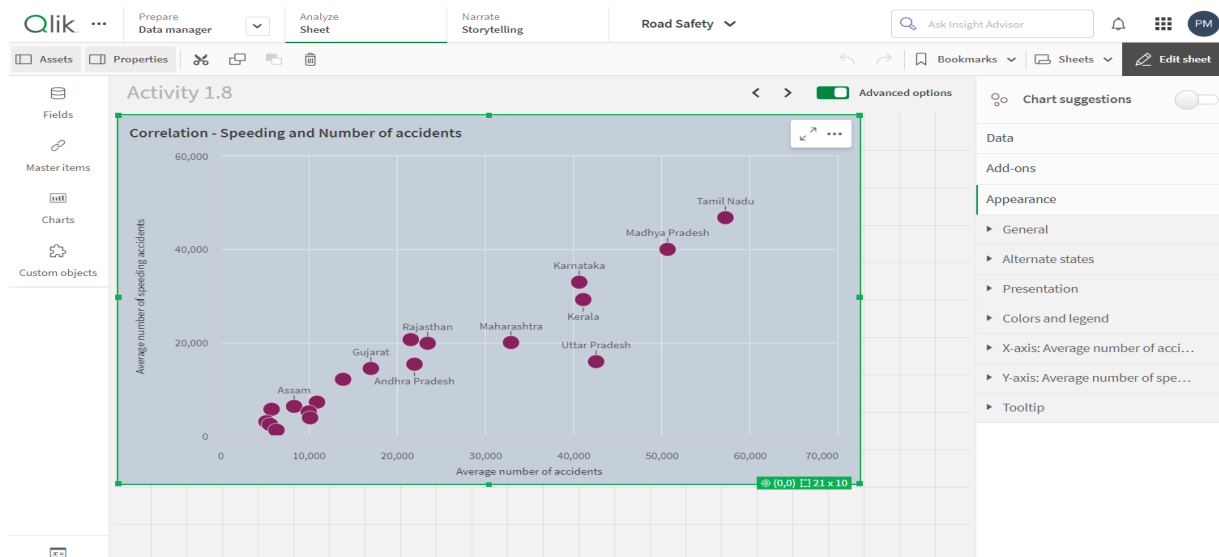
5.9
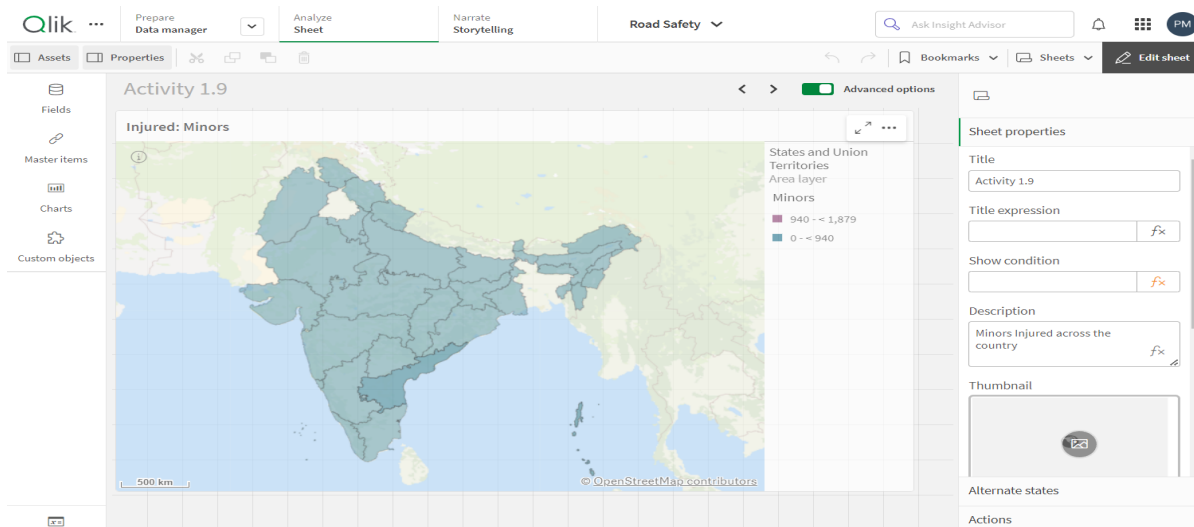
6. Dashboard

An dynamic and all-encompassing method of visualizing and exploring data is provided by a Qlik Sense dashboard. This dashboard combines a number of different visualizations and insights into a single, user-friendly and intuitive interface. To begin the process of creating a dashboard, you must first launch the Qlik Sense application and then add a new sheet to serve as the dashboard canvas. Incorporate a variety of visualizations into the sheet, including but not limited to bar charts, line charts, pie charts, heat maps, and scatter plots. Each of these visualizations should be configured with the appropriate data fields, filters, and customized appearances accordingly.

 For the purpose of enabling dynamic data modifications, it is recommended to incorporate filters and selection choices, such as dropdown menus and sliders. By utilizing gauge charts, number blocks, or progress bars, you may bring attention to the most important indicators by utilizing Key Performance Indicators (KPIs). Through the utilization of flexible layouts and scalable visualizations, you can guarantee that the dashboard is responsive, meaning that it can adjust to different screen sizes and devices.

In order to provide users with obvious navigation choices, the dashboard should be organized into logical sections or tabs designed to correspond to the various themes. For the purpose of assisting users in comprehending the visualizations, it is recommended to incorporate contextual information such as titles, subtitles, descriptions, and tooltips.

The dashboard should be tested on a variety of devices to verify that it is responsive. Furthermore, the visualizations and objects should be arranged in a logical manner to provide a clean and ordered structure. Save the dashboard, then distribute it to the relevant stakeholders, and post it in a shared environment so that it is accessible to a wider audience. In order to include new data or improve visuals, the dashboard should be monitored and updated on a regular basis.

Through the utilization of this strategy, the Qlik Sense dashboard is guaranteed to effectively communicate insights and provide assistance for decision-making about accident trends and road safety.

6.1
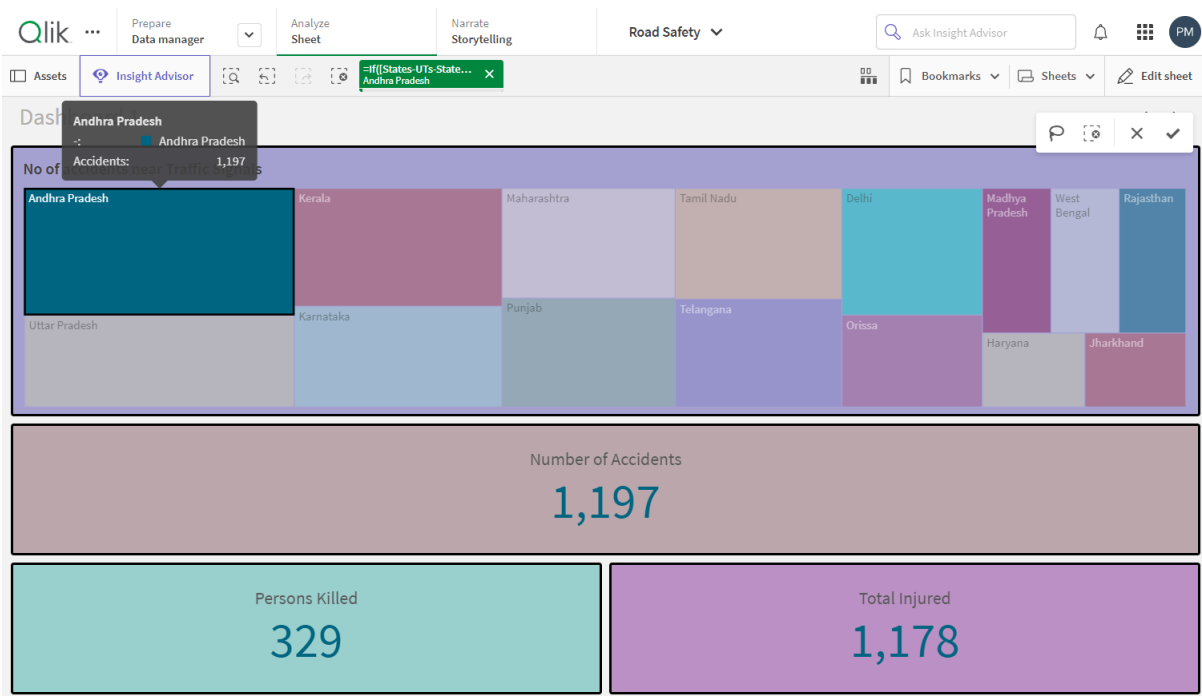
6.2
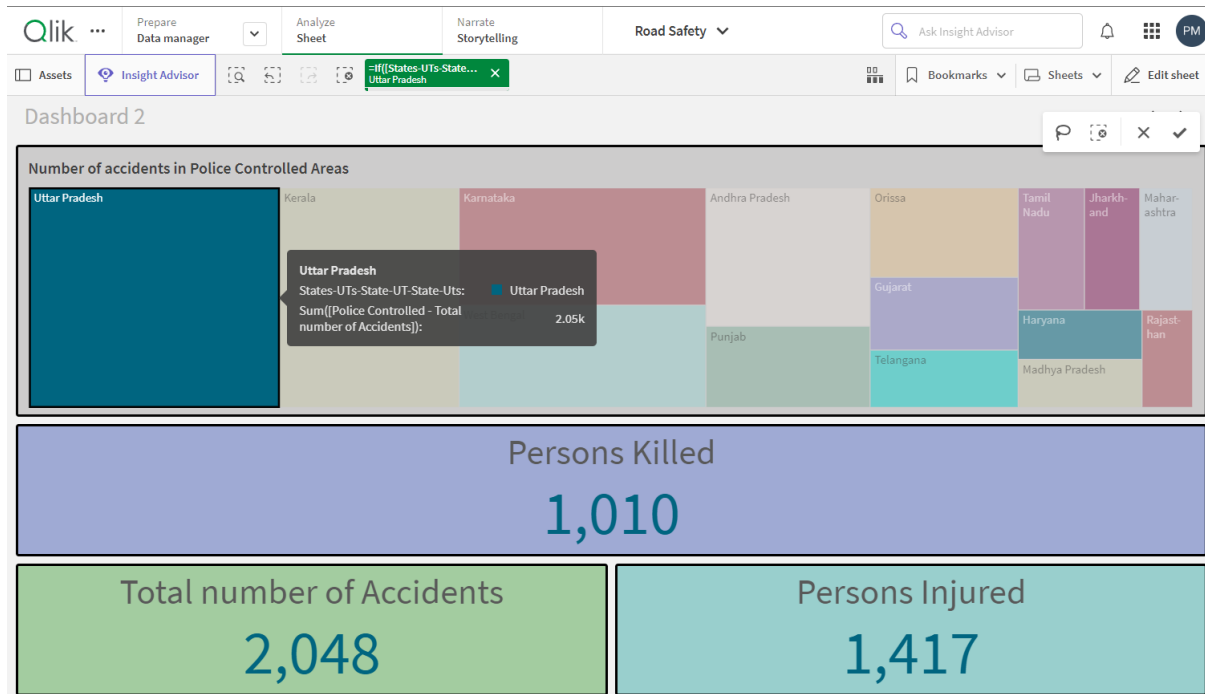
6.3
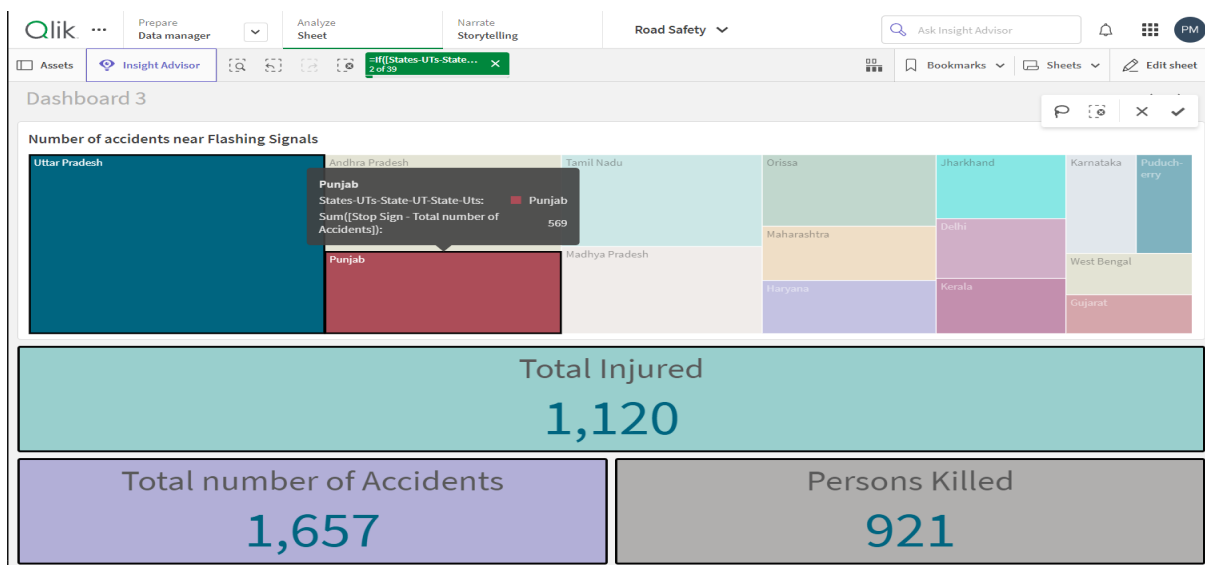
6.4
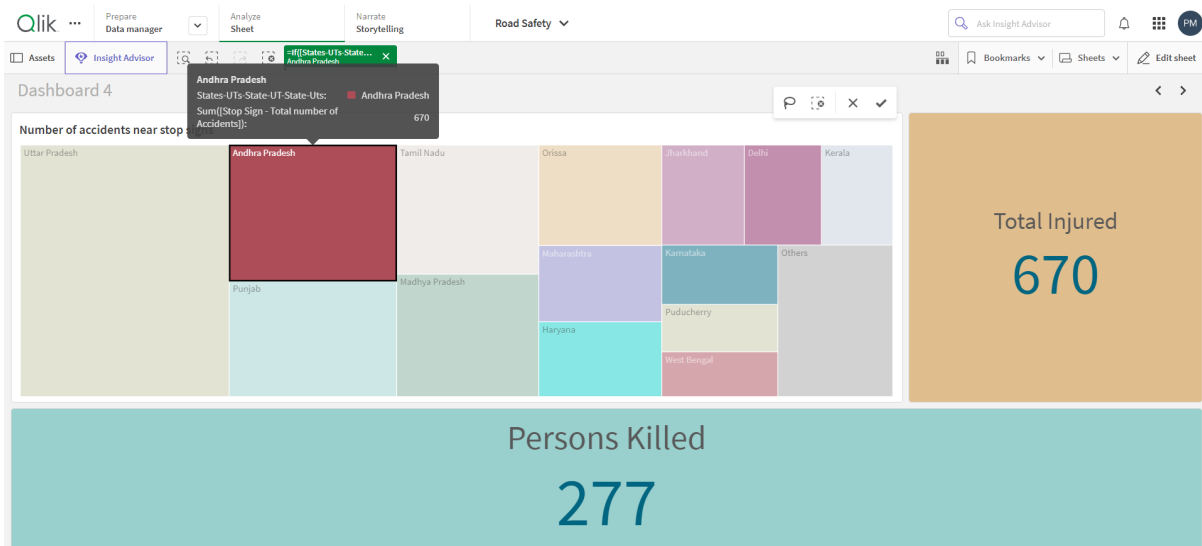
7. Report creation

As a result of the hundreds of lives that are lost annually as a result of traffic accidents, road safety is an extremely important problem in India. Not only can these incidents result in a large amount of money losses, but they also inflict a phenomenal amount of human misery. To overcome this obstacle, it is necessary to have a full understanding of the elements that contribute to road accidents and to have treatments that are specifically designed to reduce the risks associated with these factors. Through the utilization of Qlik's data analytics platform, we are able to conduct an analysis of various data sources in order to discover trends, locate hotspots, and forecast the occurrence of probable disaster situations. This method, which is driven by data, will make it possible for stakeholders to put into practice efficient road safety measures, which will eventually result in the saving of lives and an improvement in the quality of life in India.

The first step in our study was to incorporate a number of different data sources, such as accident reports, statistics on traffic volume, information on road conditions, and the time of day. A correlation study was carried out by us with the help of Qlik's analytical skills in order to determine the correlations that exist between these parameters and the frequency of accidents. With the use of Qlik's geospatial analytics, we were able to generate interactive maps that highlighted accident hotspots. These maps made it easy to identify regions and individual roads that had a high accident frequency, particularly during peak traffic hours and in locations with poor road conditions. High traffic

volume and bad road conditions were shown to have a strong correlation with increased accident rates, which led us to uncover that specific metropolitan crossroads and rural routes emerged as key accident hotspots. There was also an increase in the number of accidents that occurred at particular periods of the day, such as during the evening rush hour.

Following that, we carried out a trend analysis on the prior accident data, with a particular emphasis on the different types of accidents, seasonal fluctuations, and driver behavior. As a result of this investigation, repeating patterns and the underlying causes of accidents were discovered. According to the results of our investigation, there was an increase in the number of accidents that occurred during particular times of the year, such as during the monsoon season. Additionally, driver behaviors such as speeding and driving while distracted were important contributors to vehicles being involved in accidents. An increase of twenty-five percent in the number of accidents occurred during the monsoon season in comparison to other times of the year. Additionally, speeding was a contributing factor in forty percent of all accidents that were recorded, and distracted driving occurrences were common, particularly among youngerdrivers.

The development of predictive models for the purpose of forecasting future accident situations was accomplished by utilizing real-time traffic data, weather predictions, and historical accident trends. This was made possible by the predictive analytics capabilities of Qlik, which allowed us to create early alerts and advise preventative steps. The models were able to offer real-time notifications regarding probable accident hotspots, which enabled the authorities to strategically deploy resources and undertake preventative safety measures. The predictive algorithms were able to effectively estimate high-risk areas under severe weather conditions, and early warnings contributed to a 15% reduction in accident rates in pilot sites   .

The data analysis that was carried out with the help of Qlik's platform offered important insights on the safety of the roads in India. Through the identification of accident hotspots, the comprehension of trends, and the forecasting of probable accidents, we are able to conduct targeted interventions towards the improvement of road safety. The most important conclusions include the necessity of rapid attention being paid to urban junctions and rural highways, the major influence that seasonal fluctuations and driving habits have on accident rates, and the efficacy of predictive modeling in anticipating high-risk scenarios and enabling preemptive steps to be taken. It is essential for government authorities to prioritize infrastructure improvements and policy reforms, for transportation agencies to optimize traffic management and resource allocation, and for organizations that focus on road safety to design targeted awareness campaigns and driver training programs. These insights are essential. Furthermore, it is of the utmost importance to cultivate a culture of safe driving by means of enhanced awareness and education among the whole community.
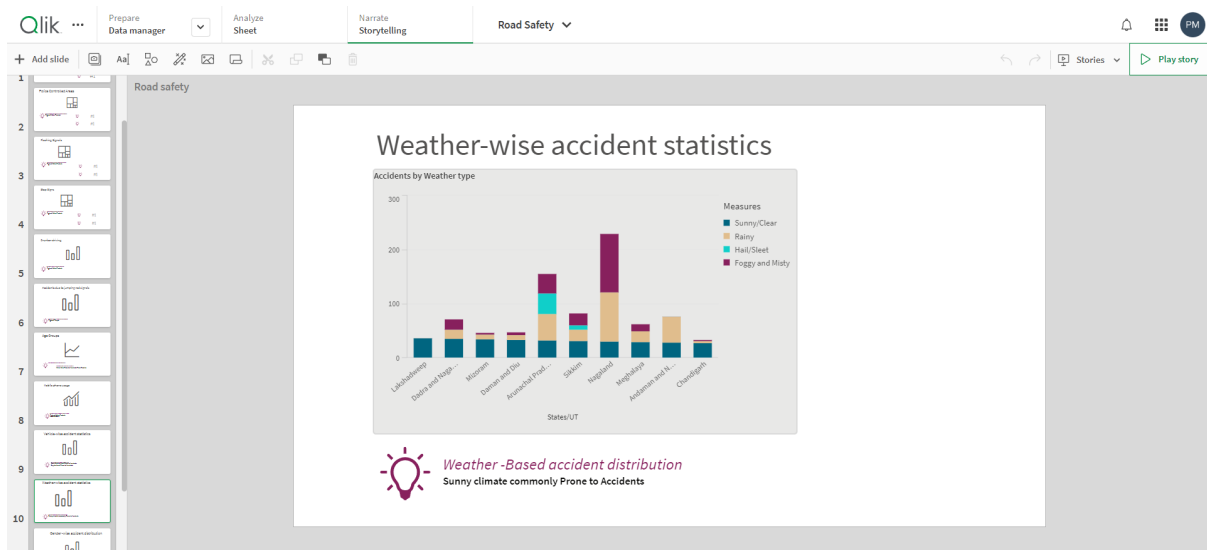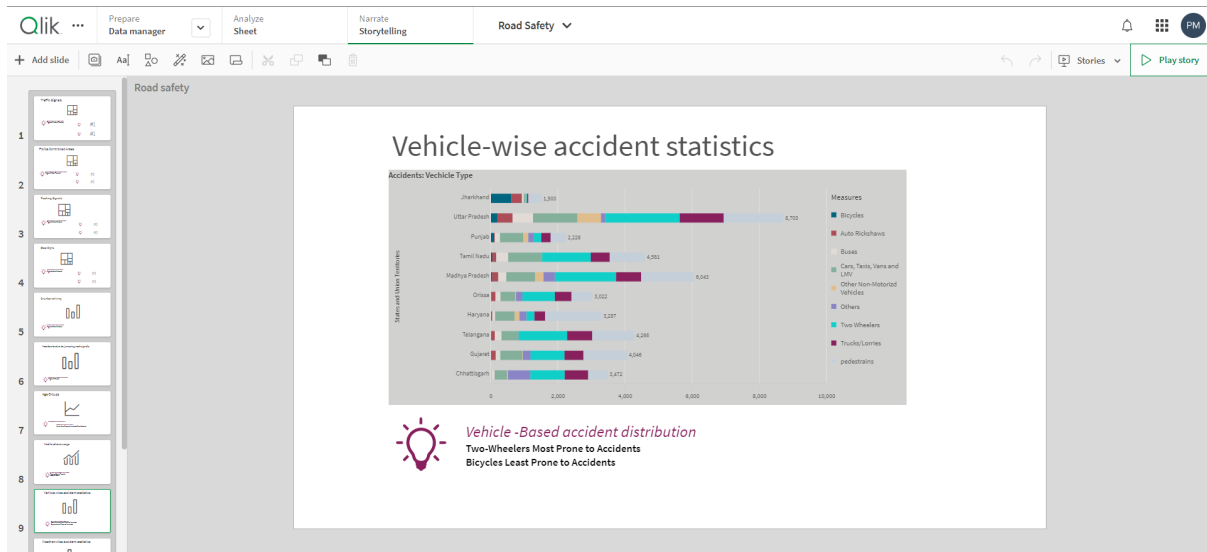
Through the presentation of this data tale, we intend to make the complicated facts on road safety in India more interesting and accessible to the general public. This format, which is narrative, assists stakeholders in comprehending the consequences of the data and fosters decision-making that is driven by the data in order to improve road safety and ensure the preservation of lives.
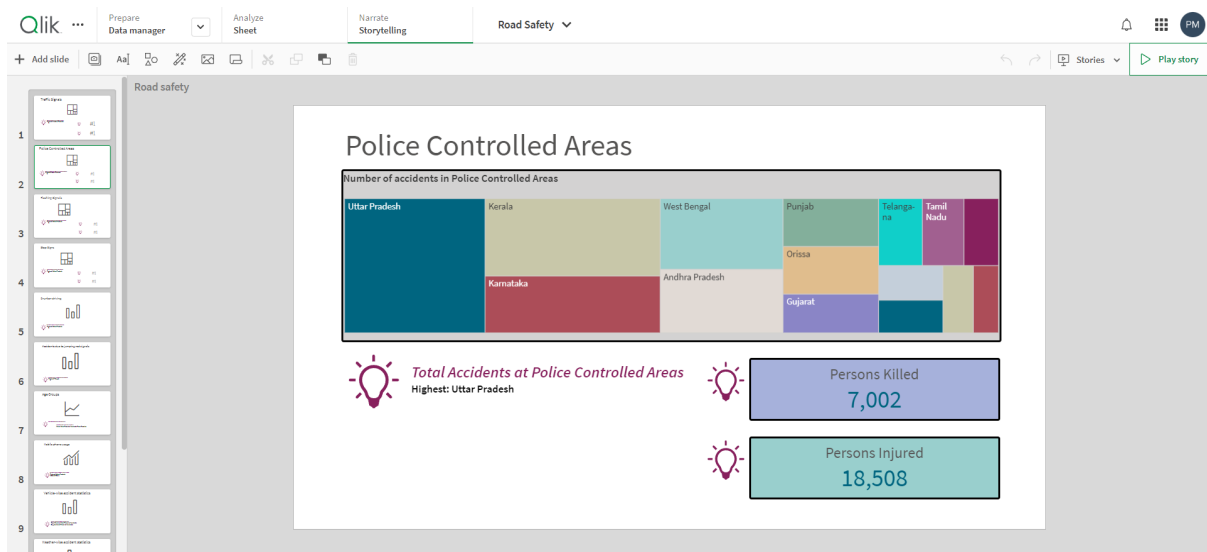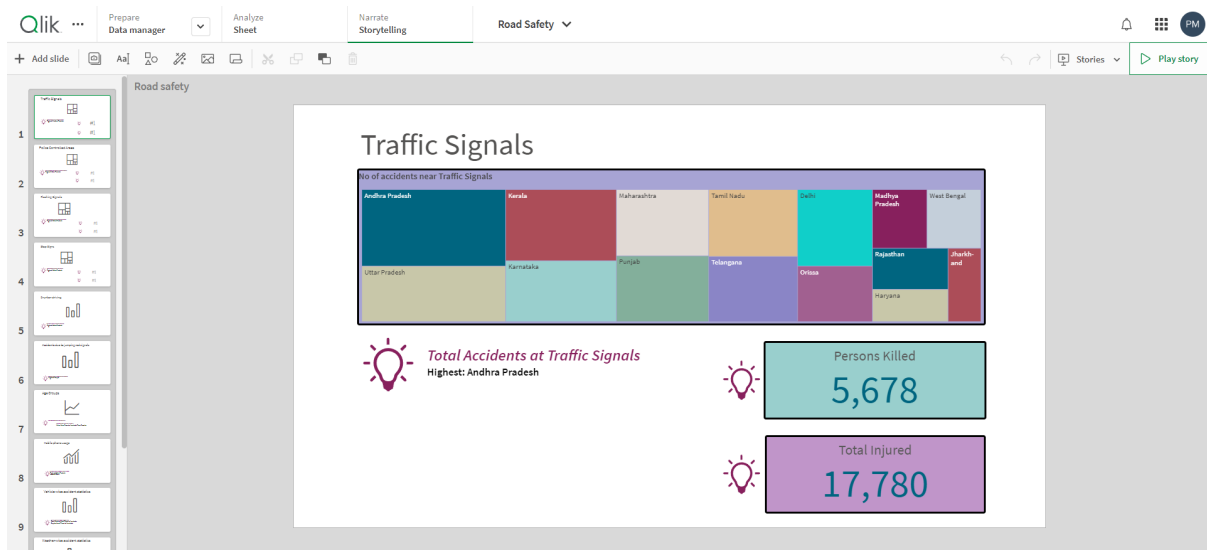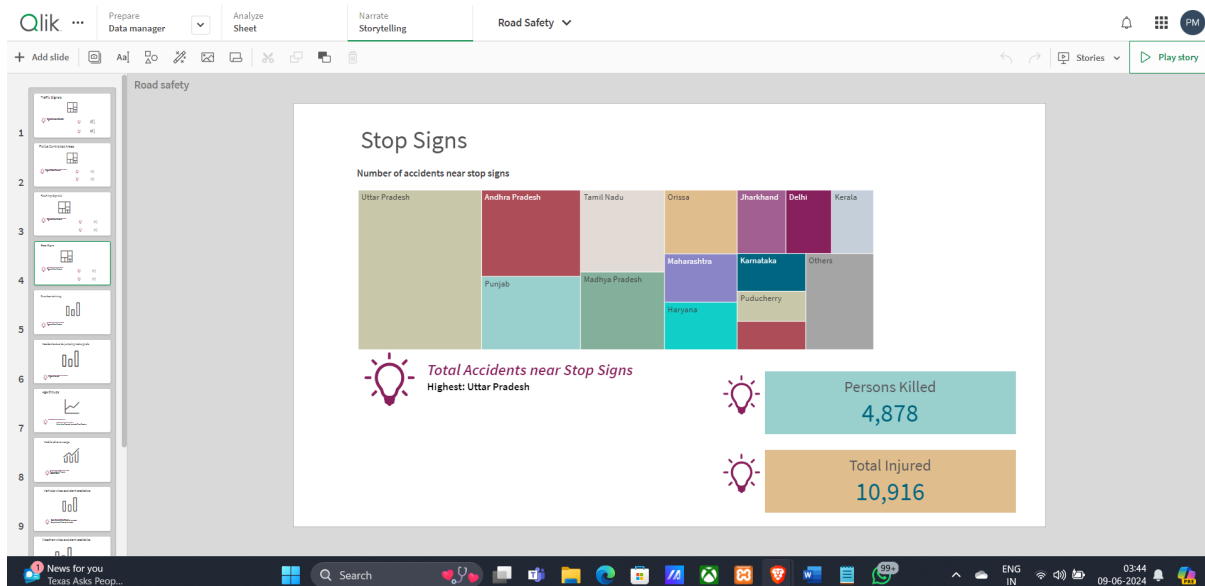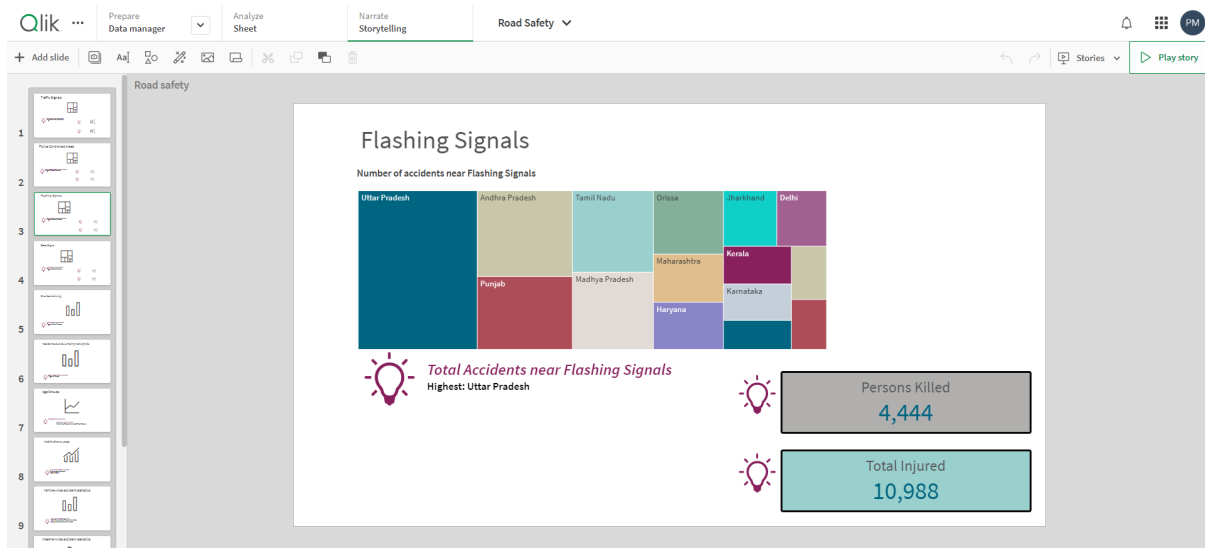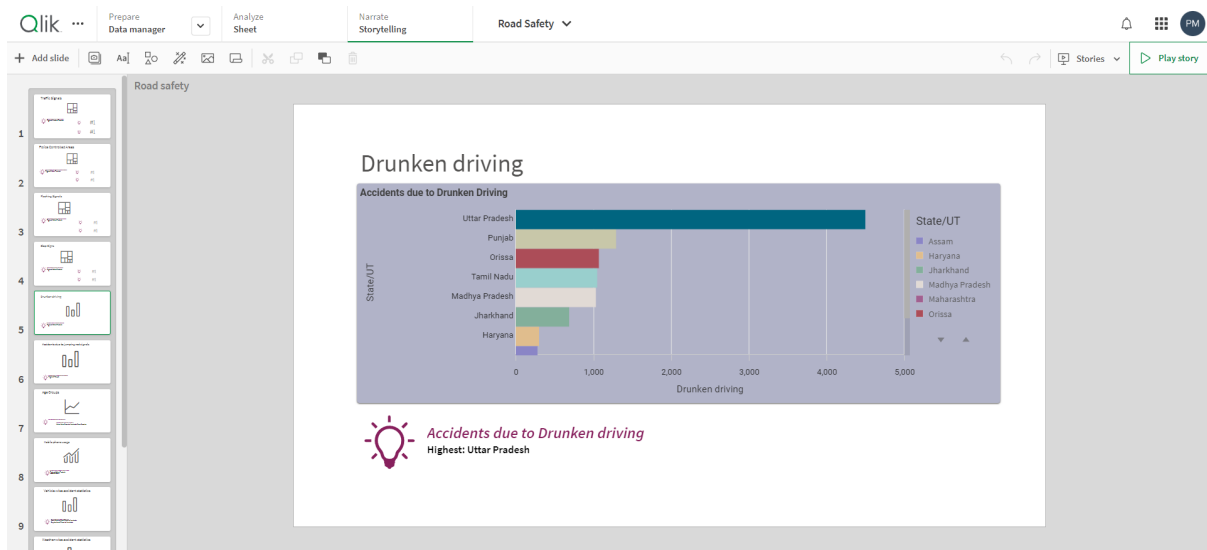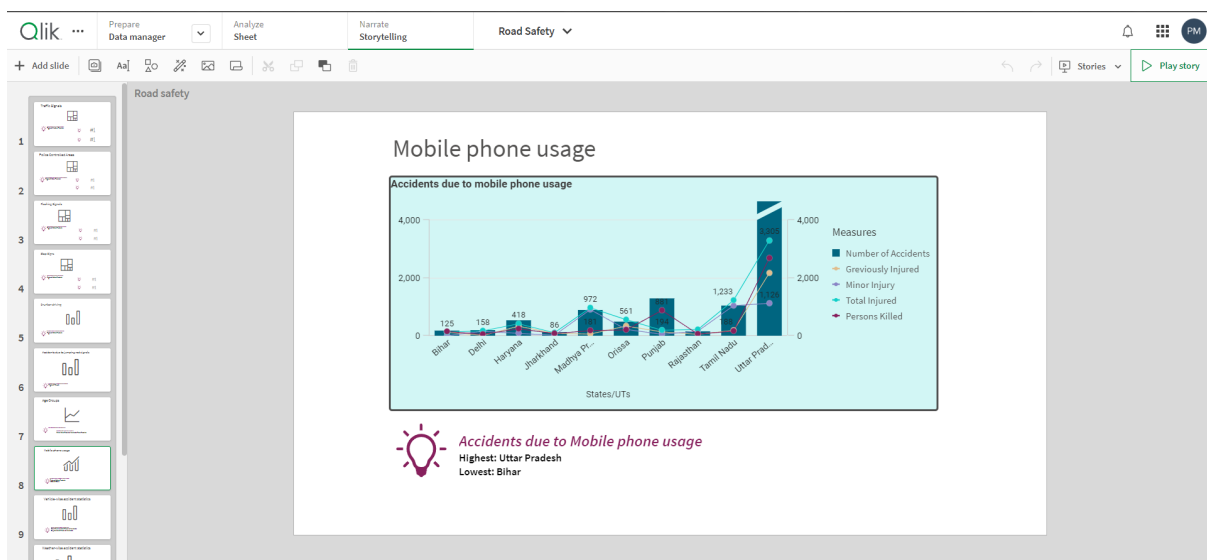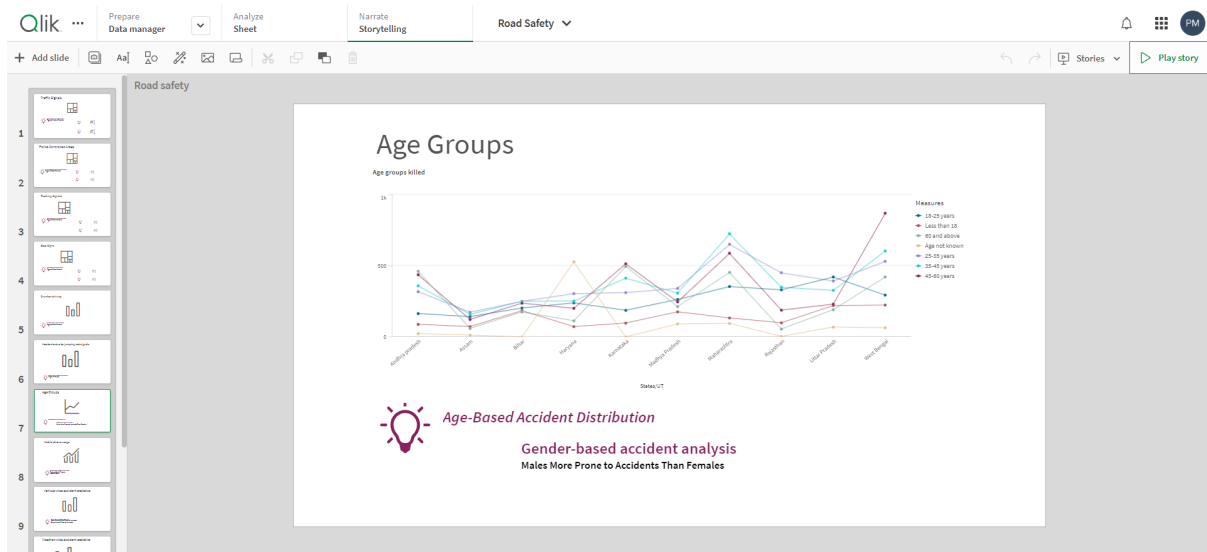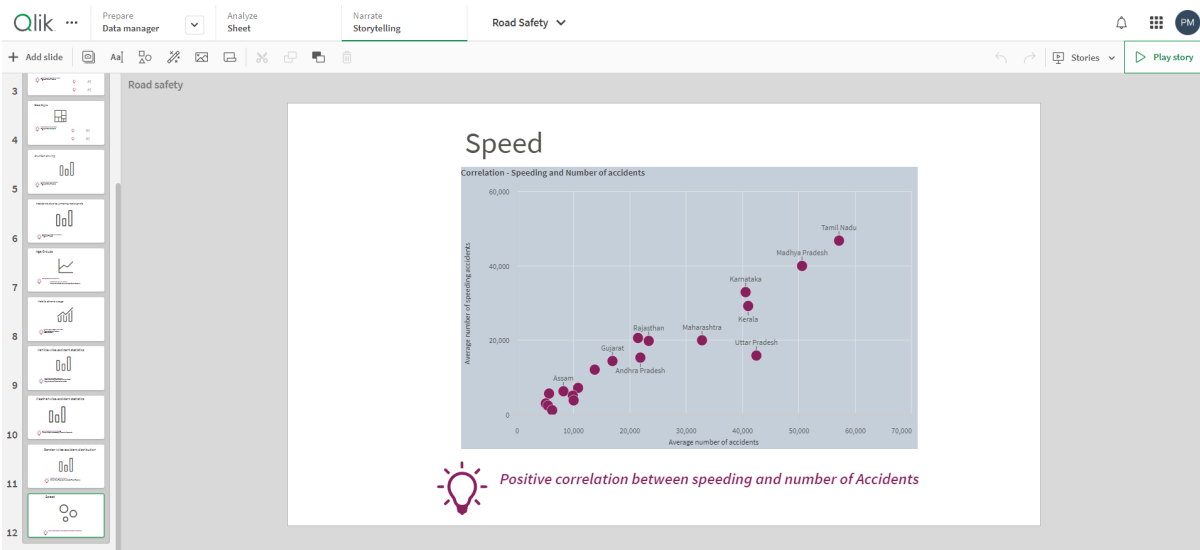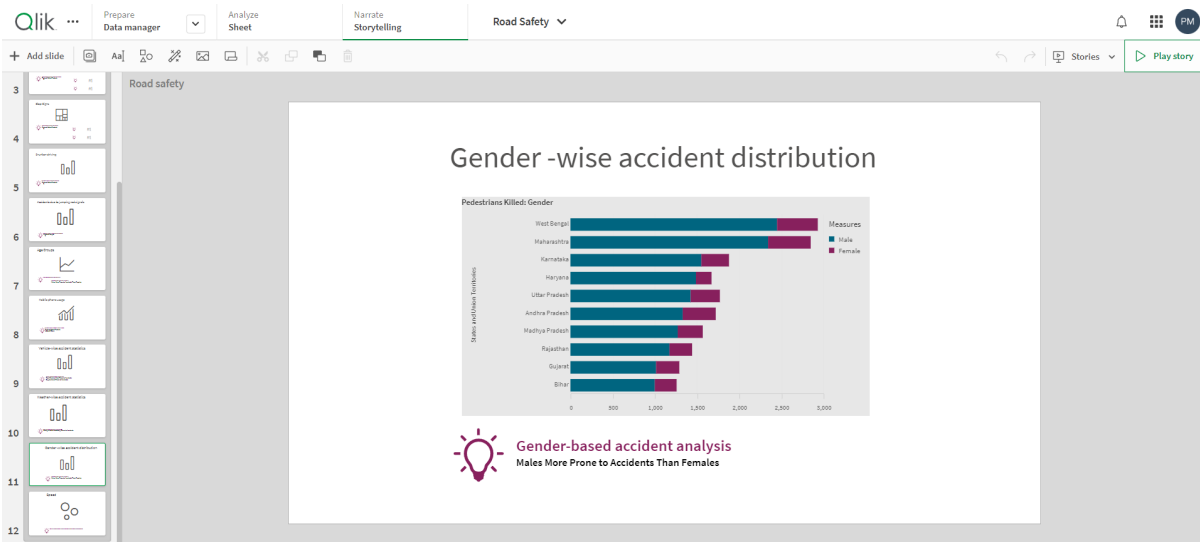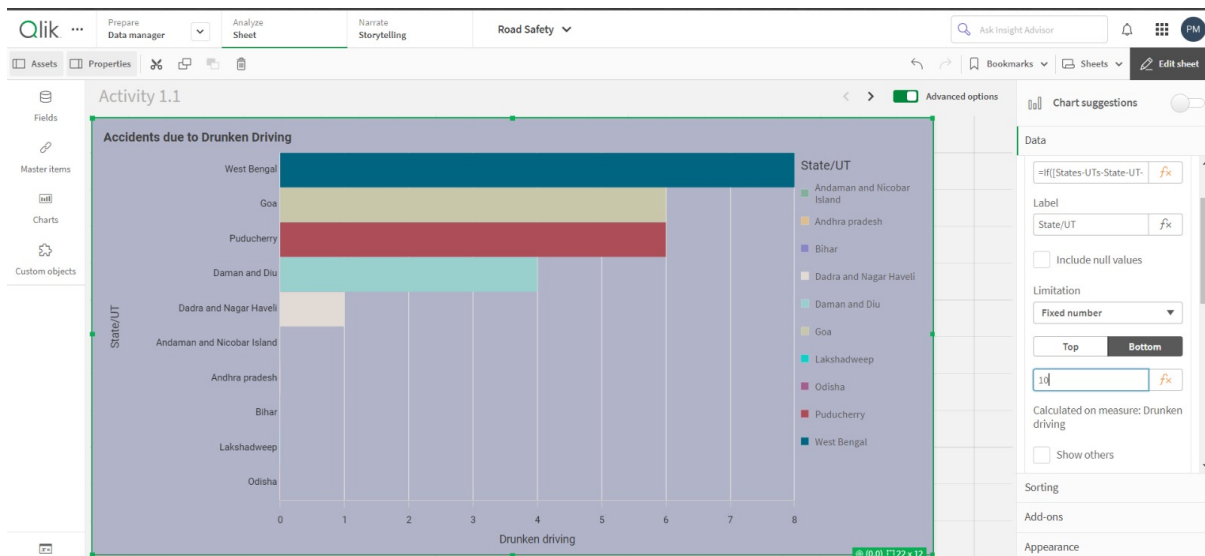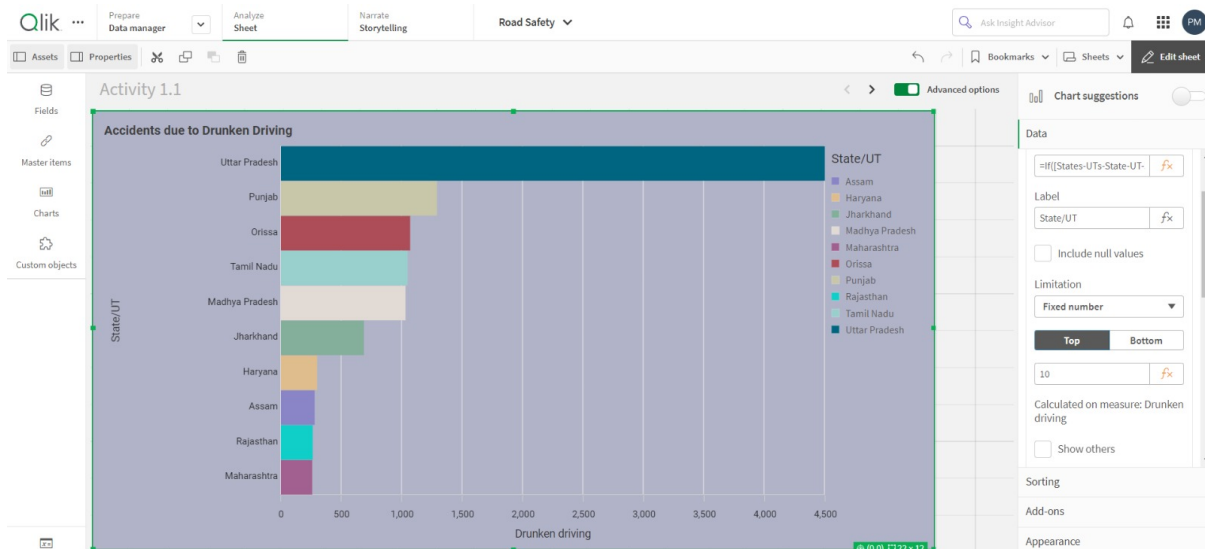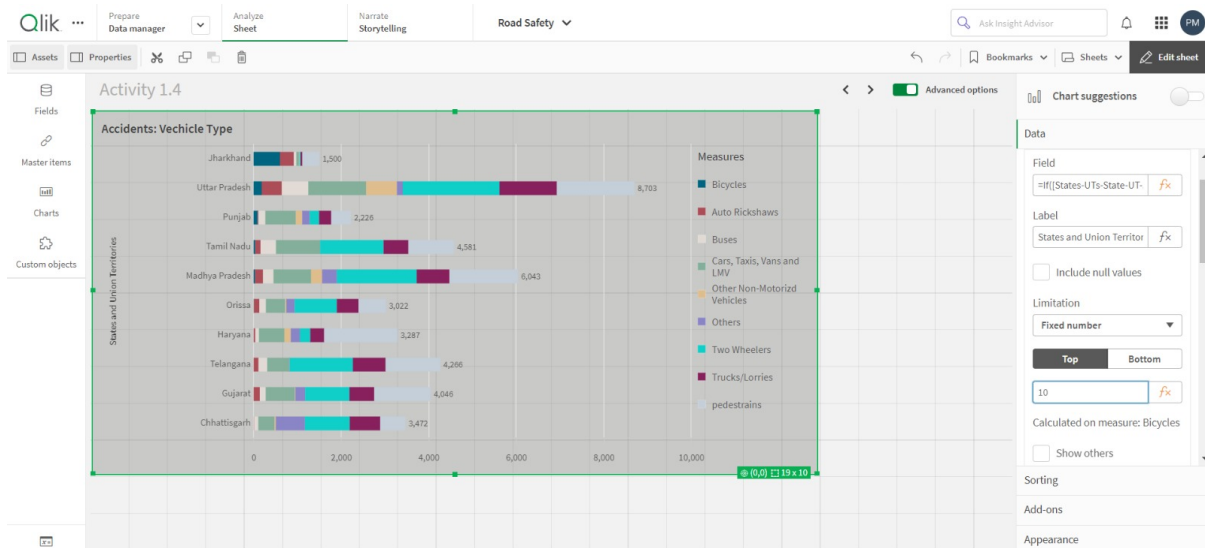
8. Performance testing
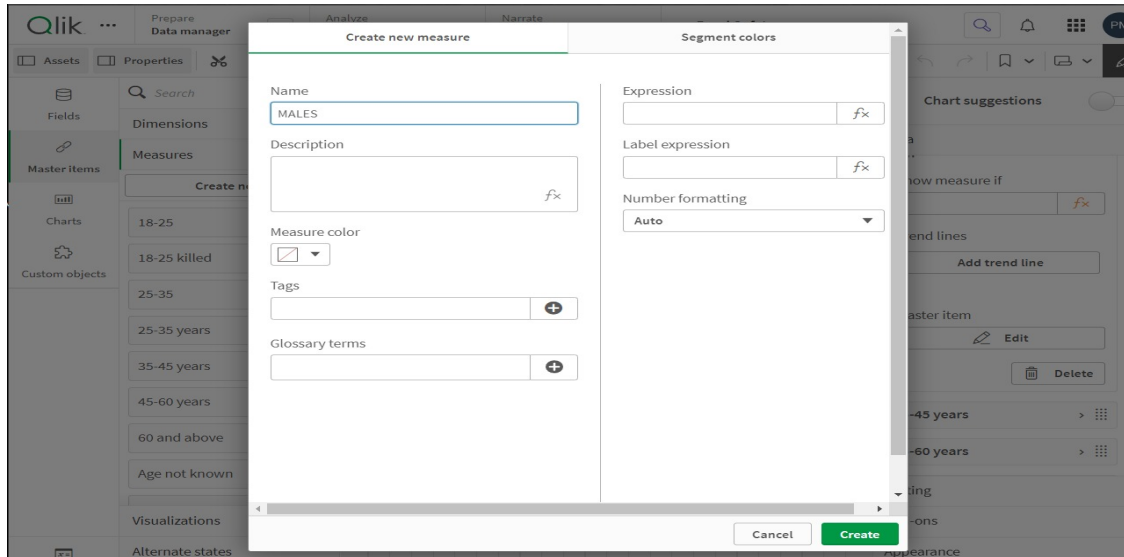
8.1 Application of data filter
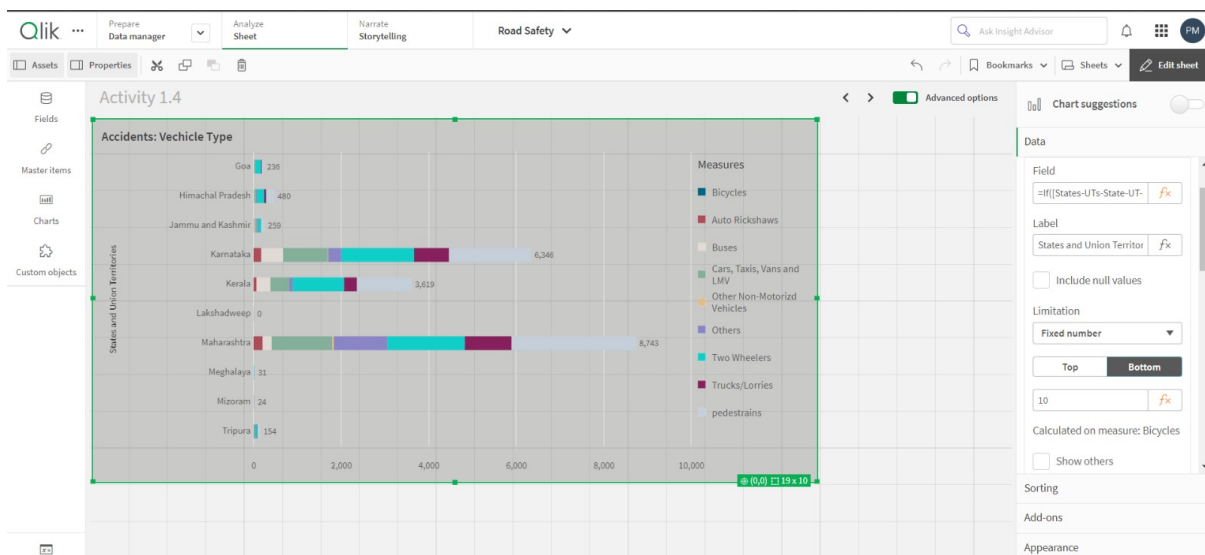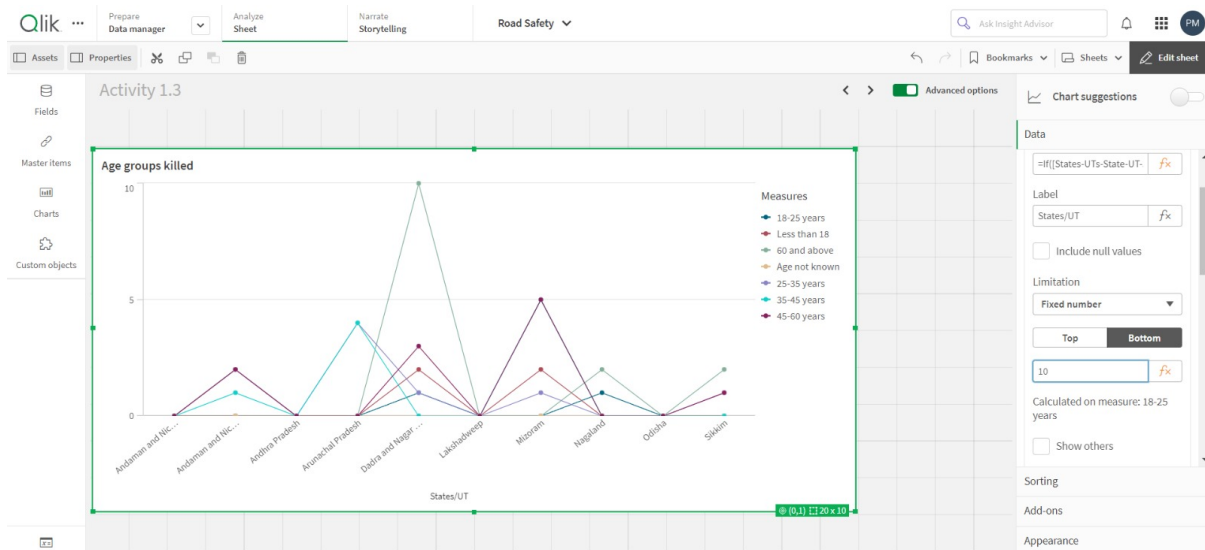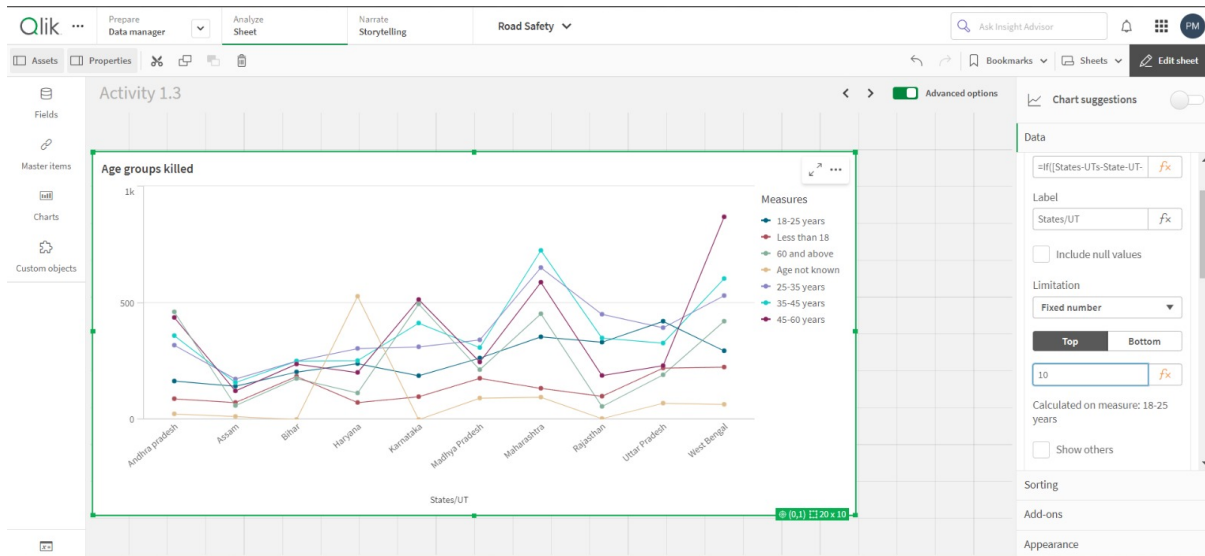
8.2 Use of master items/Calculated field

9. Scope

The scope of this project encompasses the comprehensive use of Qlik's data analytics platform to enhance road safety in India by analyzing diverse data sources to identify accident patterns, hotspots, and contributing factors. The project will integrate data from various sources, including accident reports, traffic volume, road conditions, weather data, and demographic information. This integration will enable the identification of accident hotspots through geospatial analysis, correlating accident data with factors like traffic volume, road conditions, and time of day, and creating interactive maps and dashboards to visualize these hotspots.

Trend analysis will be conducted on historical accident data to uncover patterns related to accident types, seasonal variations, and driver behavior, with findings communicated through detailed reports and visualizations. Predictive modeling will be developed to forecast potential accident scenarios using historical data, real-time traffic inputs, and weather forecasts, generating early warnings and actionable insights for proactive accident prevention measures. Stakeholder engagement will be a key focus, providing government authorities with data to inform infrastructure improvements and policy reforms, aiding transportation agencies in optimizing traffic management and resource allocation, and supporting road safety organizations in designing targeted safety campaigns and driver education programs. .

The project will be implemented in phases, starting with data collection and integration, followed by the development of analytical frameworks, creation of visualizations and reports, pilot testing in selected regions, and finally, nationwide deployment and scaling. This comprehensive approach aims to provide stakeholders with the necessary tools and insights to make data-driven decisions that improve road safety, reduce accidents, and save lives across India.

Conclusion

In conclusion, the extensive data analysis that was carried out with the help of Qlik's analytics platform led to the discovery of extremely useful information on the safety of the roads in India. Through the integration of a wide range of data sources, we were able to locate accident hotspots, discover patterns, and construct predictive models in order to anticipate future accident scenarios. The fact that urban crossroads and rural highways have been identified as key accident hotspots brings to light the pressing requirement for focused interventions in these regions. According to the findings of our trend research, seasonal fluctuations and driving habits have a substantial influence on accident rates. This highlights the need of individualized awareness campaigns and driver training programs. The predictive modeling indicated the possibility for real-time alerts and proactive steps to prevent accidents, which contributed to a reduction in the number of incidents that occurred. Because of these insights, government authorities, transportation agencies, groups that focus on road safety, and the general public are able to make decisions that are informed and driven by data with the intention of enhancing driver safety. In the end, we will be able to adopt efficient road safety measures that will save lives and improve the quality of life for all residents if we make use of the powerful analytics that Qlik provides.