**Project Report**

**On**

**Qlik Analysis of Road Safety and Accident Patterns In India**

**Batch: May-june 2024**

**Submitted by**

**PUJITHA GOLLAMUDI**

**Skill Wallet ID : SWUID20240009651**

**Using**

**Qlik Platform**

****

**CONTENTS**

1.Introduction

2. Define Problem / Problem Understanding

* 1. Specify the business problem
  2. Business requirements
  3. Literature Survey
  4. Social or Business Impact

3. Data Collection

* 1. Collect the dataset

3.2 Connect Data with Qlik Sense

4. Data Preparation

4.1 Prepare the Data for Visualization

5. Data Visualizations

5.1 Number of Unique Visualizations

5.1.1 Bar chart

5.1.2 Combo chart

5.1.3 Line chart

5.1.4 Stacked bar chart

5.1.6 Scatter plot

5.1.7 Pie chart

5.1.8 Waterfall chart

6. Dashboard

6.1 Responsive and Design of Dashboard

7. Story Telling

7.1 Design of Story

8. Performance Testing

* 1. Amount of Data Rendered to DB
  2. Utilization of Data Filters

8.3 Number of Calculation Fields/Master Items

**INTRODUCTION**

**1.Introduction**

**Qlik Analysis Of Road Safety And Accident Patterns In India**

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.

scenario 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.

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.

**CHAPTER 2**

**Define Problem / Problem Understanding**

**2.Define Problem / Problem Understanding**

**2.1 Specify the business problem**

Technological advancements in transportation have significantly reduced distances and travel times, facilitating economic growth and personal mobility. However, these advancements have also brought about an increase in road traffic accidents, leading to a substantial loss of lives and injuries each year. To address this issue, a comprehensive study using Qlik Sense, a powerful data analytics platform, aims to analyze road safety and accident trends in India. This study involves examining various data points related to road incidents to uncover patterns and factors contributing to road safety or risks. The insights generated from this analysis will 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.

* 1. **Literature Survey**

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

* 1. **Social or Business Impact**

Social Impact Analysis:  
• Create visualizations to display the demographic distribution of accidents across the country.

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

• Explore any correlation between speeding, weather, and total accidents.

• Identify the leading causes of accidents.

• Examine the distribution of age groups and gender of the victims.

• Investigate the contribution of diverse types of vehicles to the total number of accidents.

**CHAPTER 3**

**DATA COLLECTION**

**3. Data Collection**

* 1. **Collect the dataset**

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

Data set has been collected from Kaggle

Kaggle is the world’s largest data science community with powerful tools and resources to help you achieve your data science goals...

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

About the Dataset

We have given a total of nine data sets where each dataset has the collected information about the accidents and their patterns in India and the datasets describes the following:

* 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 various data sources in Qlik Sense, integrate the data, and prepare it for analysis to study road safety and accident trends in India.

Step 1: Identify and Prepare Data Sources

Data that has been collected from kagggle

Step 2: Connect to Data Sources in Qlik Sense

Open Qlik Sense: Create a new app.

Add Data: Click “Add Data”.

File Upload: For CSV/Excel files.

Database Connection: For SQL Server, Oracle, etc.

API Connection: Using REST connectors for APIs.

Step 3: Integrate and Transform Data

Load Data: Use “Data Manager” or “Data Load Editor”.

Clean Data: Remove duplicates, handle missing values.

Transform Data: Merge tables, create calculated fields.

Associate Data: Define relationships using Qlik Sense’s associative model.

Step 4: Verify Data Integration

Preview Data: Ensure data is correctly loaded.

Validate Data: Cross-reference with original sources to ensure accuracy.

Step 5: Save and Document

Save App: With relevant name and description.

Document Process: Steps taken, sources used, and issues encountered.

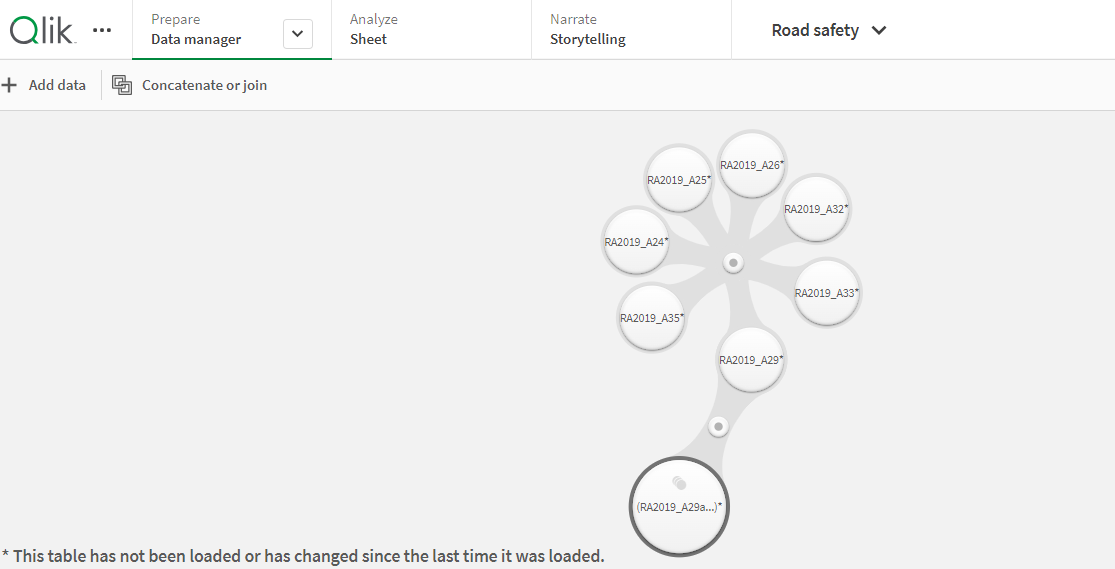
**CHAPTER 4**

**DATA PREPARATION**

**4. Data Preparation**

**4.1 Prepare the Data for Visualization**

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring that the data is accurate and complete. This process helps to make data easily understandable and ready for creating visualizations to gain insights.



**CHAPTER 5**

**DATA VISUALIZATIONS**

**5. Data Visualizations**

Data visualization is the process of creating graphical representations of data to help people understand information. 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.

**5.1 Number of Unique Visualizations**

The number of unique visualizations that can be created with a given dataset. Some common types of visualizations that can be used to analyse include bar charts, line charts, heat maps, scatter plots, pie charts, maps etc. These visualizations can be used to compare, track changes over time, show distribution, relationships between variables, breakdown of one category and much more.

5.1.1 Bar Chart:

Step 1: Load Your Data

Ensure your data related to road safety and accident patterns is loaded into Qlik Sense. This includes data on accident types, locations, causes, and other relevant factors.

Step 2: Create a Bar Chart

Drag the “Bar Chart” icon from the left panel onto the sheet.

Select Data:

Dimensions: Choose the dimension you want to represent on the x-axis. For instance, this could be the type of accident, location, or time period.

Measures: Choose the measure you want to represent on the y-axis. This could be the number of accidents, fatalities, injuries, etc.

Step 3: Customize the Bar Chart

Adding Dimensions, Measures, Adjusting appearance, colours, Labels, Legends etc.

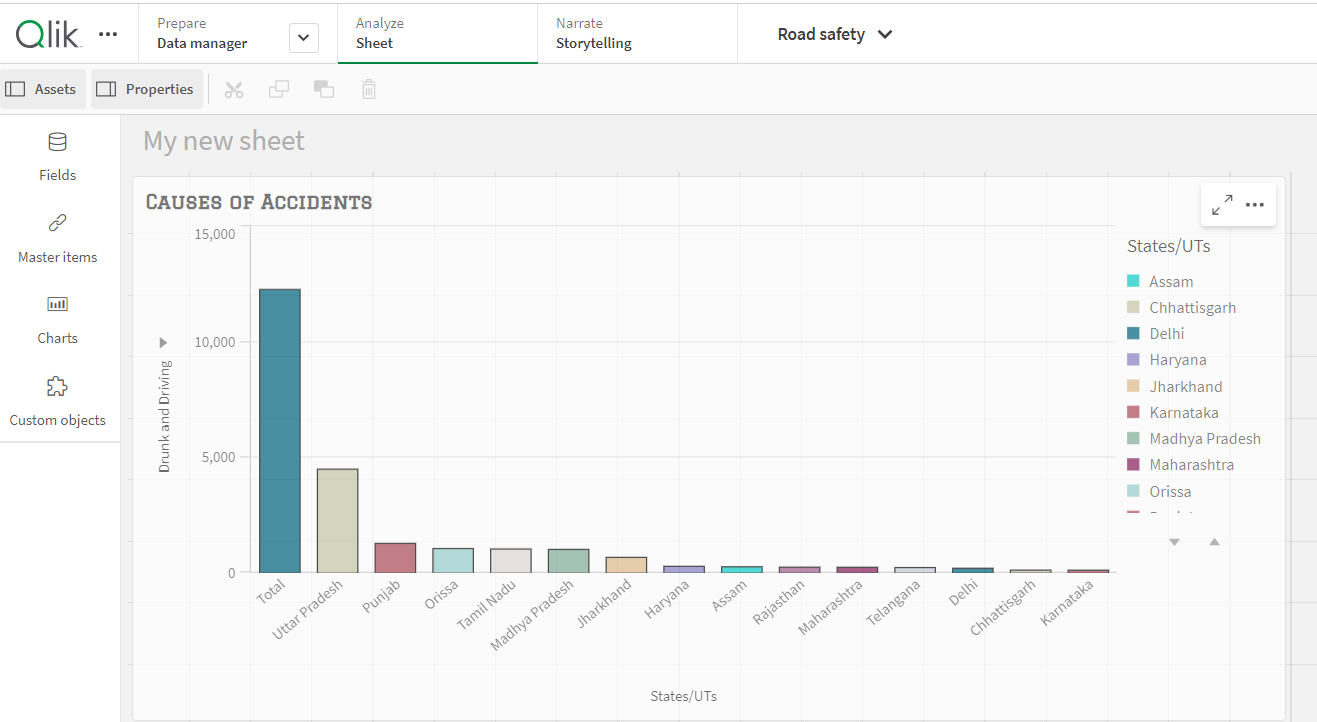
Step 4: Analyze and Interpret

Interact with the Chart:

Use Qlik Sense’s interactive features to explore the data. You can filter, drill down, and hover over bars to get detailed information.

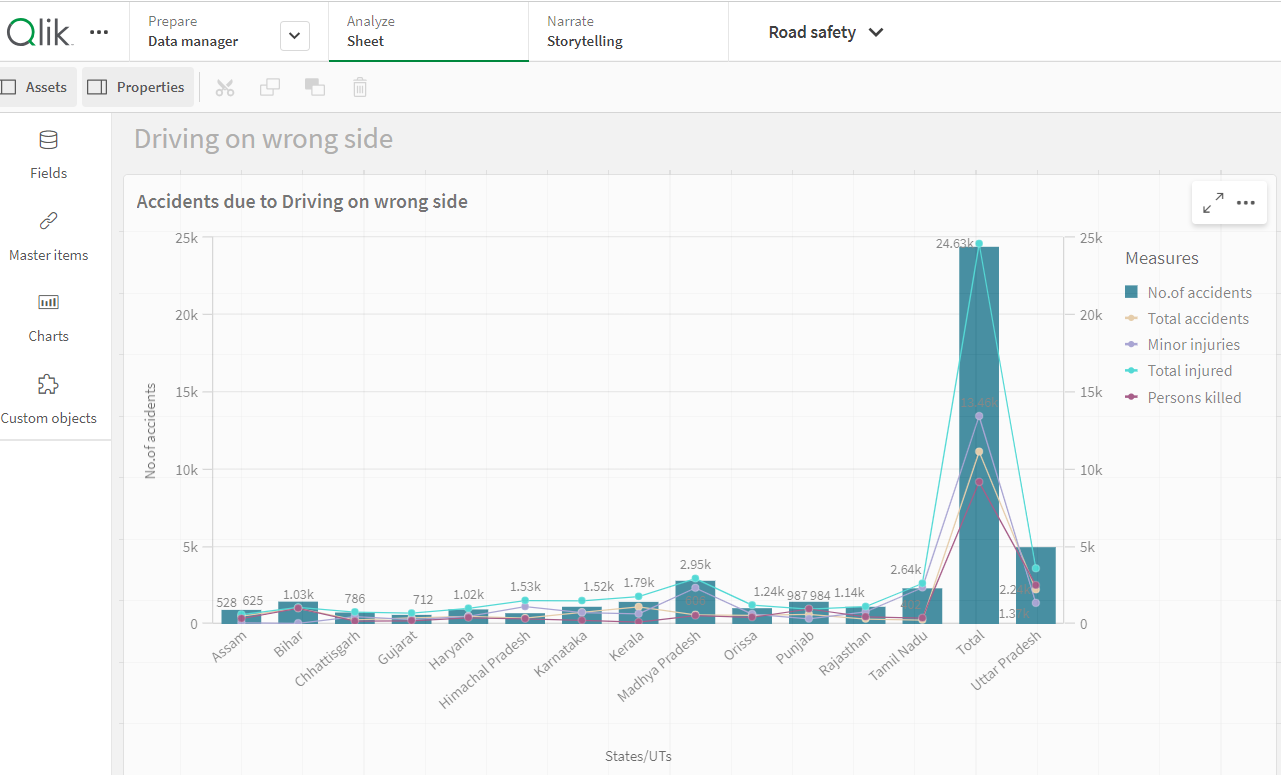
Derive Insights:

Analyze the bar chart to identify trends, patterns, and outliers. For instance, you might notice that certain types of accidents are more frequent or that specific locations are hotspots for accidents.

****

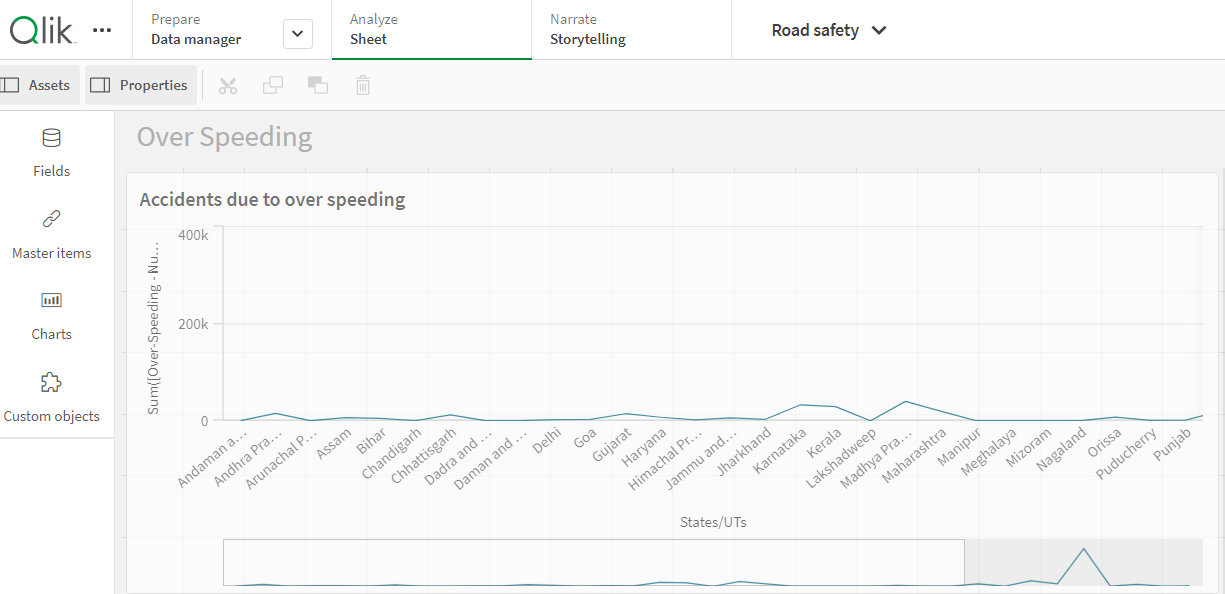
5.1.2 Combo Chart:

Showing accidents due to driving on wrong side in various states in India using combo chart



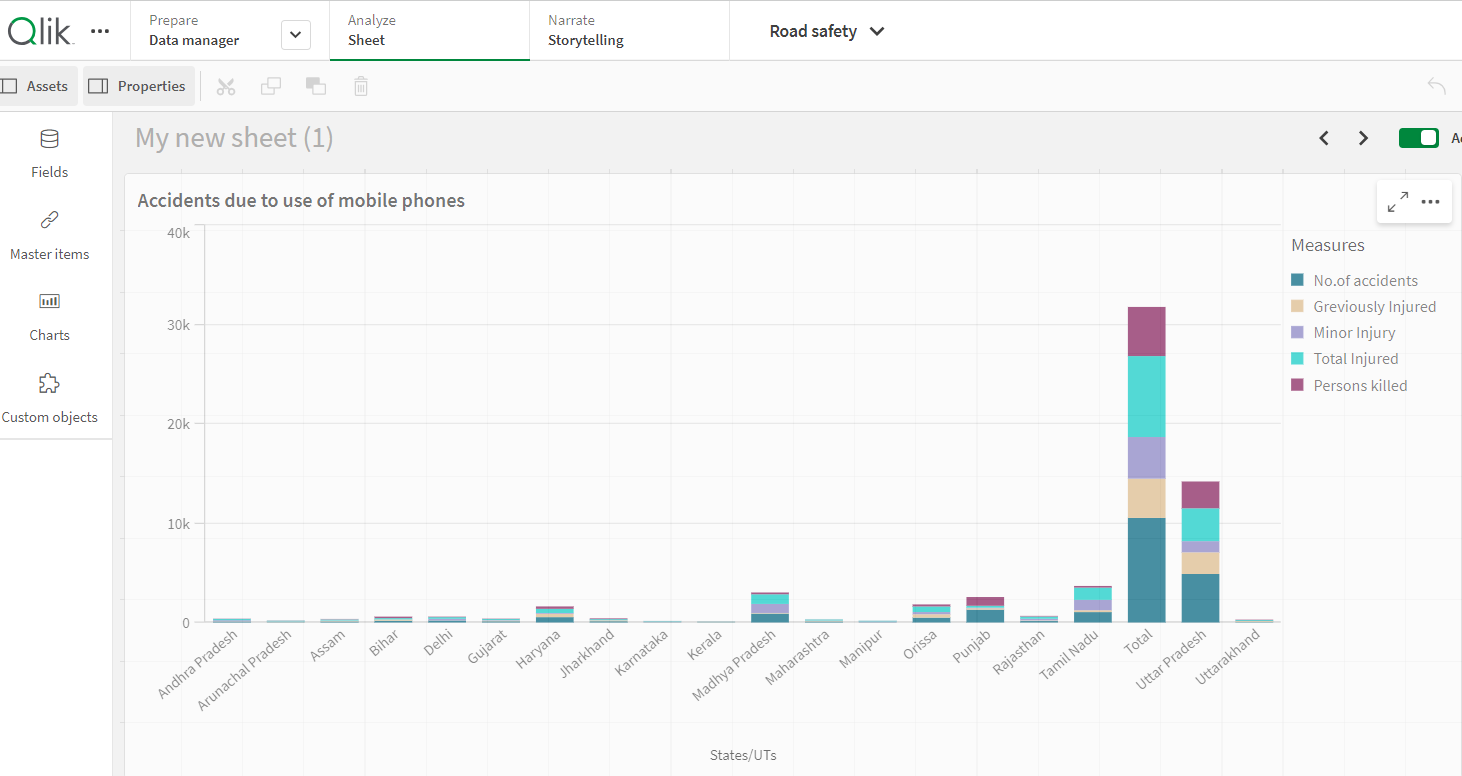
5.1.3 Line Chart

Showing accidents due to over speeding in various states using Line chart



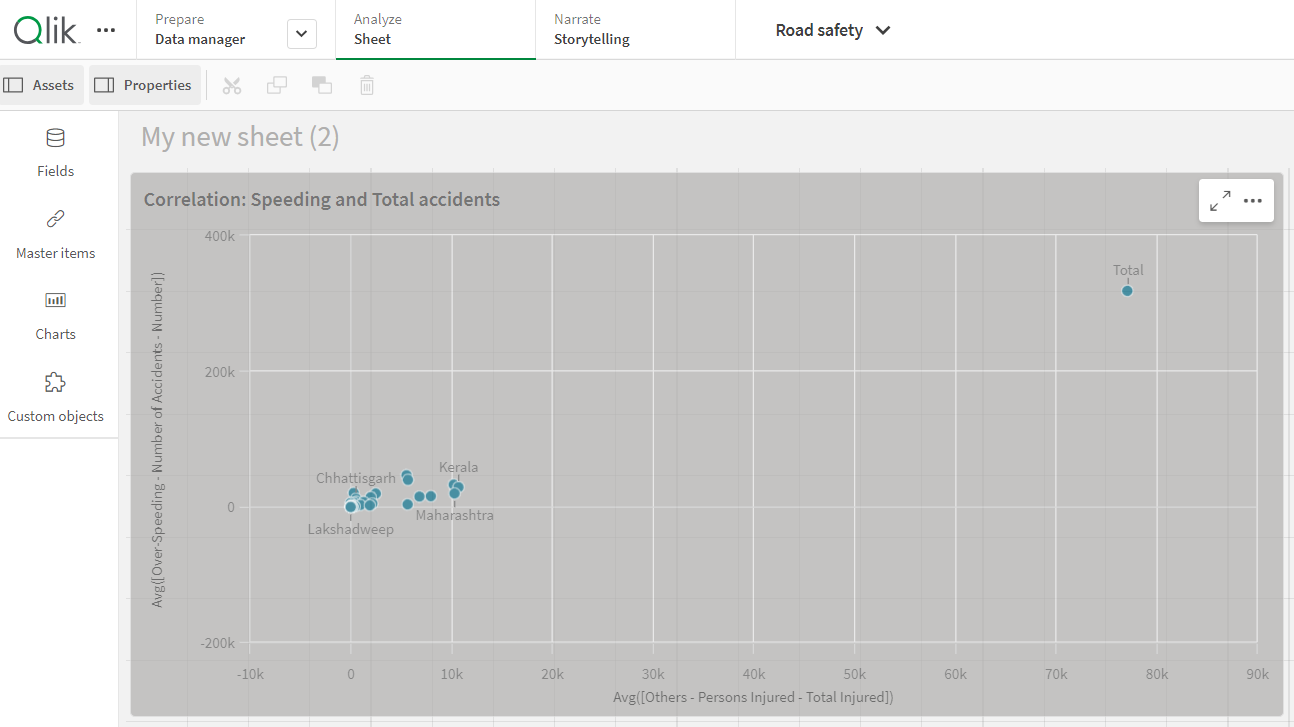
5.1.4 Stacked Bar Charts:

Showing Accidents occurred due to usage of mobile phones in various states of India



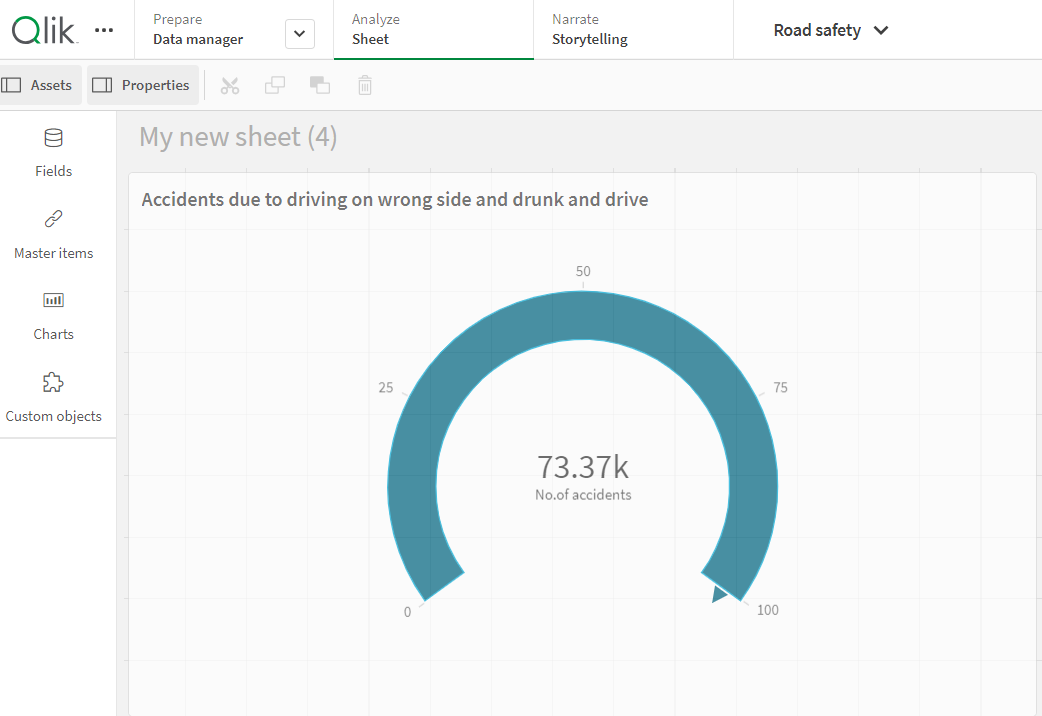
5.1.5 Scatter Plot:

Showing the correlation between speeding and occurrence of accidents



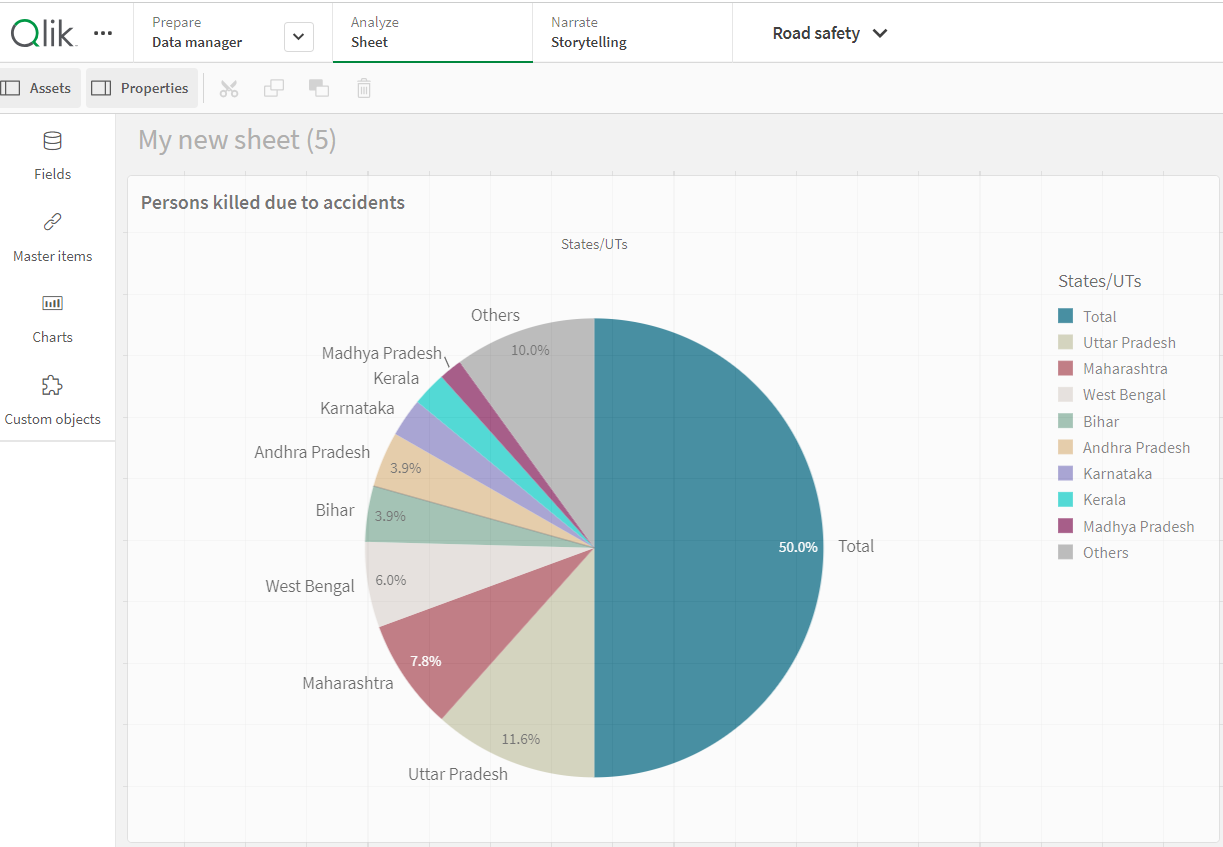
5.1.6 Guage:

Showing occurrence of accidents due to the combination of driving on wrong side and consumption of alcohol using a guage



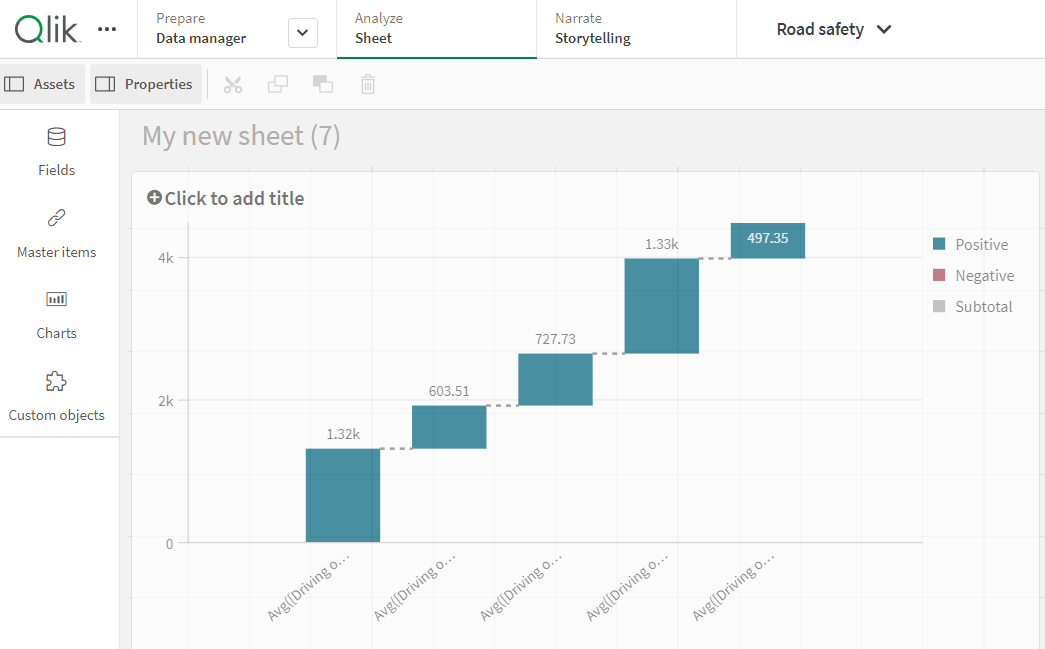
5.1.7 Pie chart

Showing the percentage each state occupied out of hundered based on persons killed due to accidents



5.1.8 Waterfall Chart

Showing the waterfall model of occurrence of accidents due to driving on wrong side



**CHAPTER 6**

**DASHBOARD**

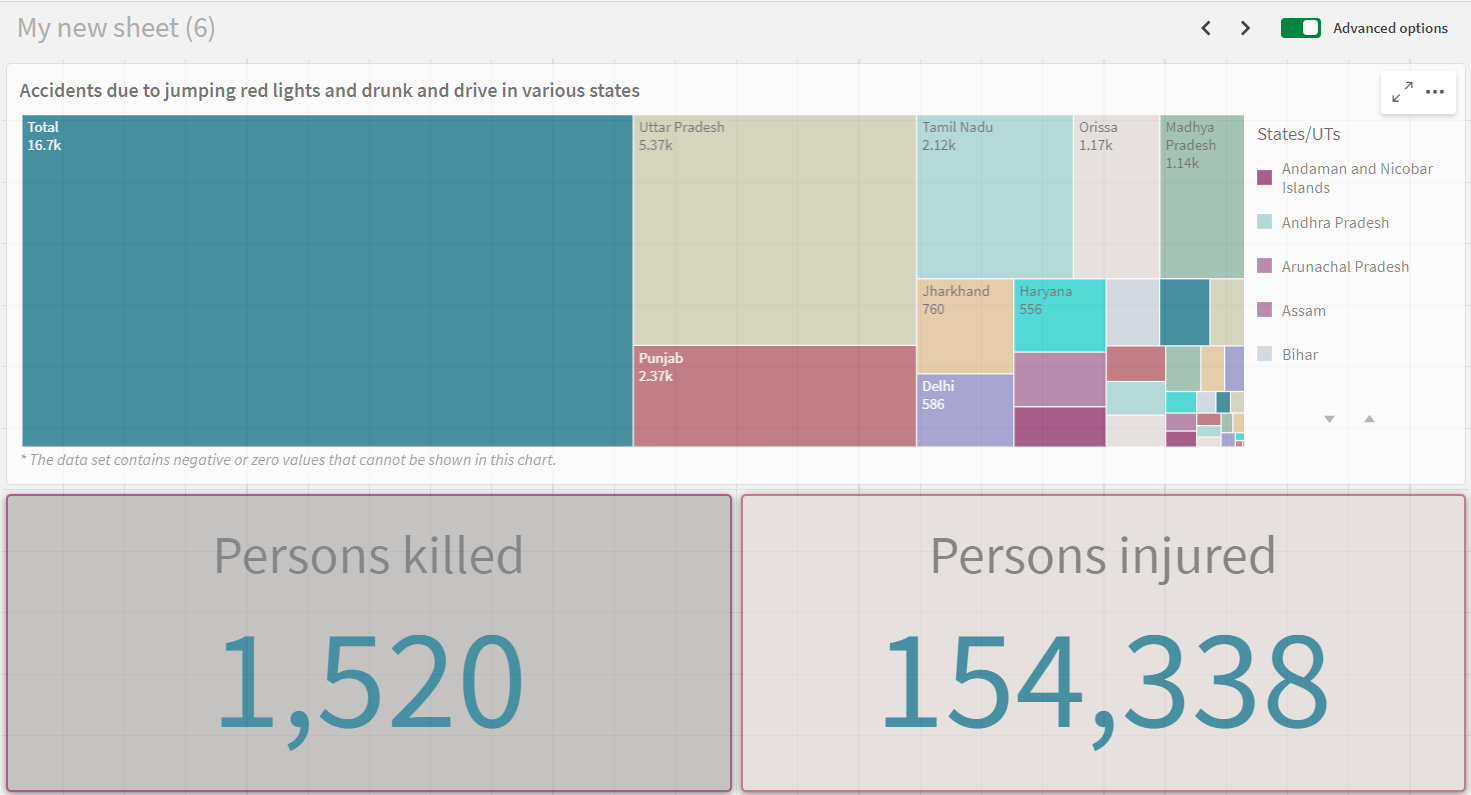
**6. Dashboard**

**6.1 Responsive and Design of Dashboard**

A dashboard is a graphical user interface (GUI) that displays information and data in anorganized 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.

Tree map and two Key Performance Indicators

Showing the accidents due to jumping red lights and drunk and drive in various states, Key performance indicators showing total number of persons killed and number of persons injured. By selecting particular state we can observe the change in number in KPIs



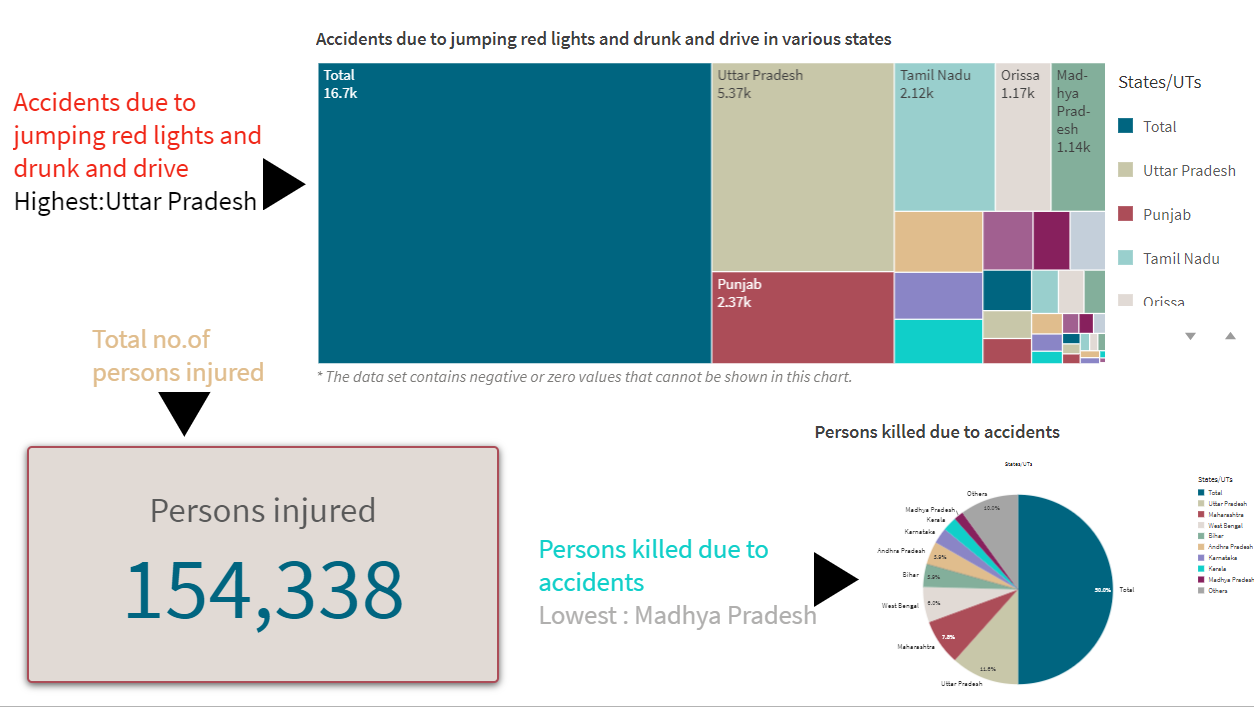
**CHAPTER 7**

**STORY TELLING**

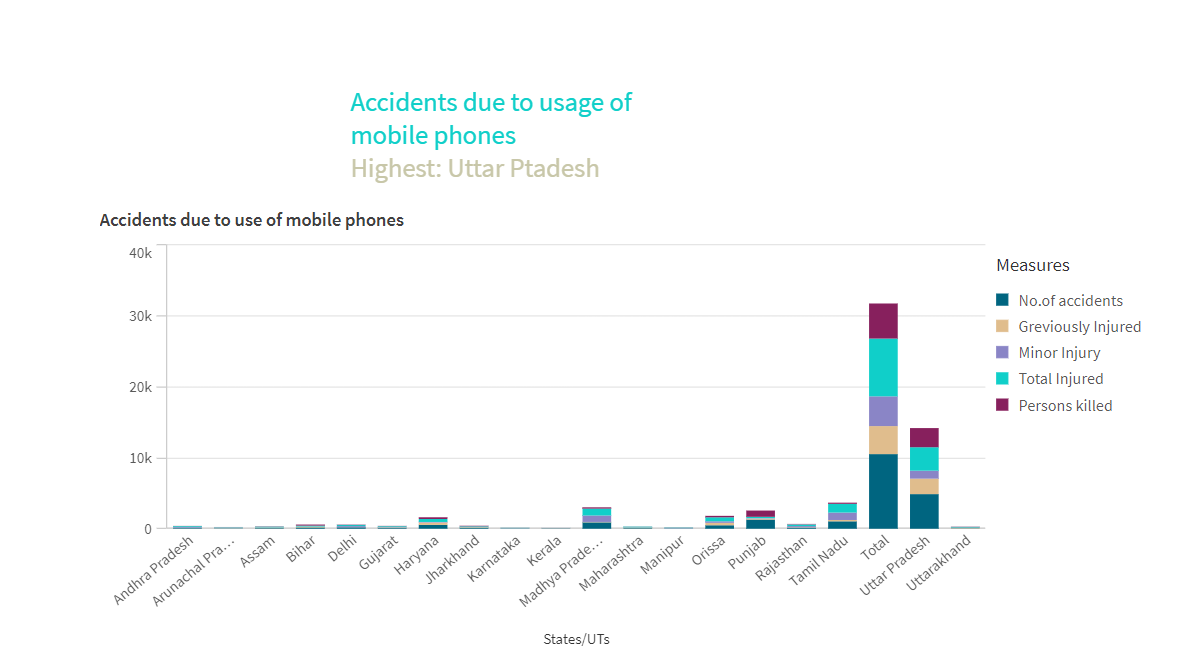
**7. Story Telling**

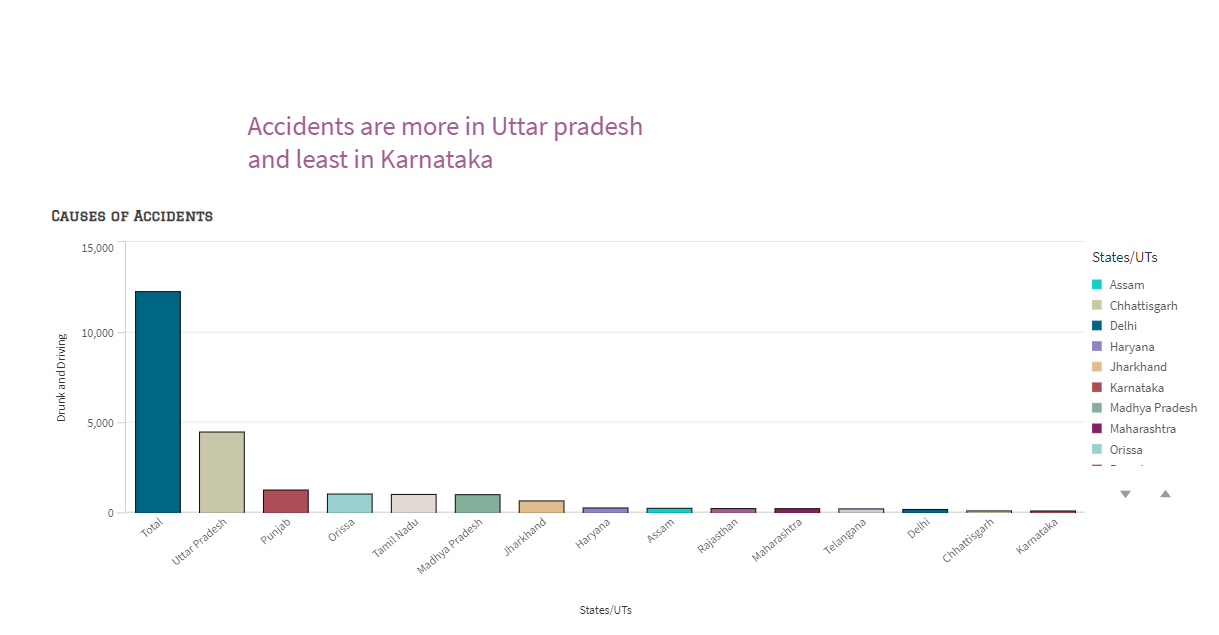
**7.1 Design of Story**

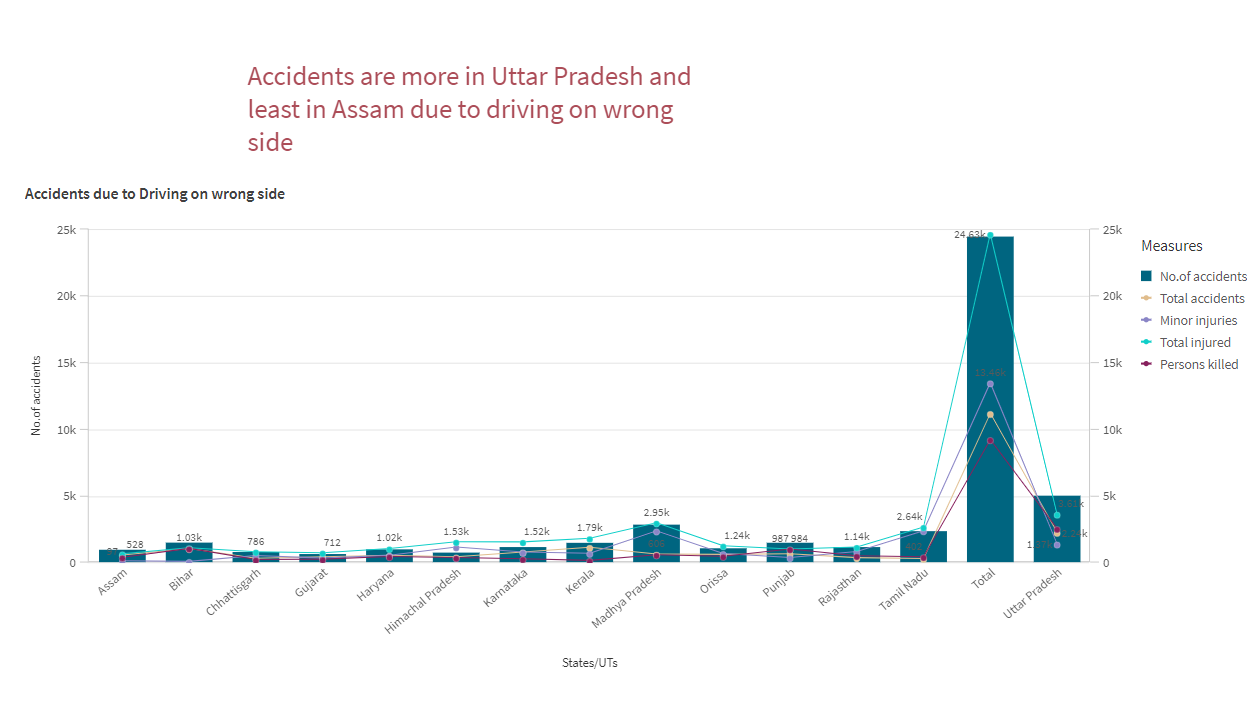
A data story is a way of presenting data and analysis in a narrative format, with the goal of making information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis 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.





****

****

****

**CHAPTER 8**

**PERFORMANCE TESTING**

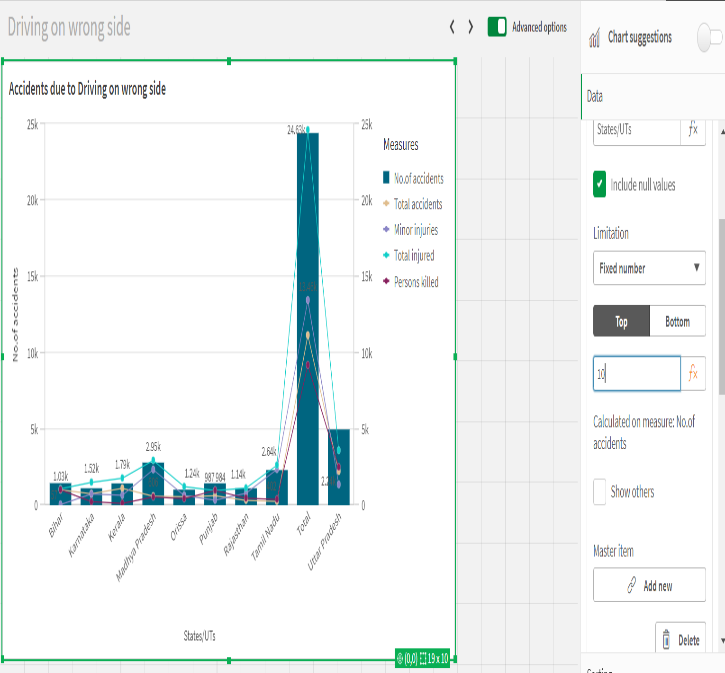
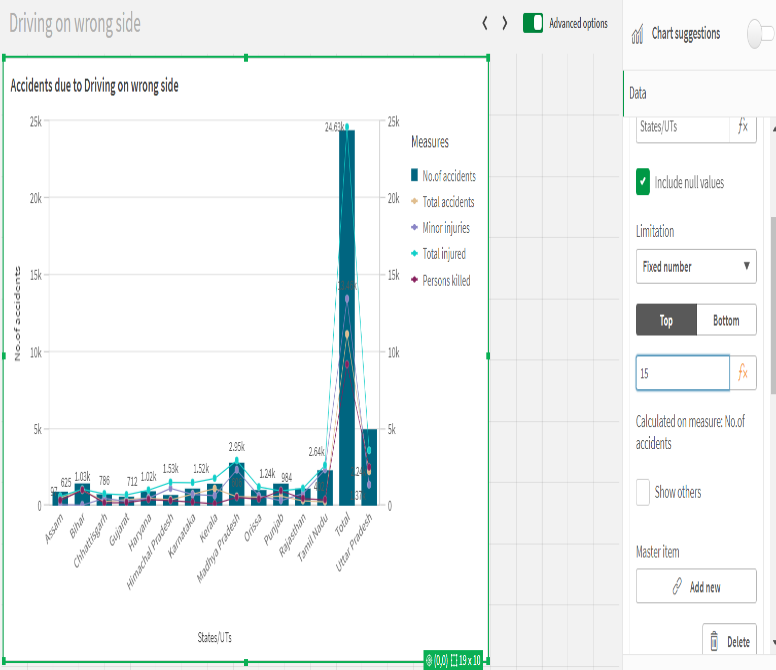
**8. Performance Testing**

**8.1 Amount of Data Rendered to DB**

Selections within the data allow 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**.**

**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.



**8.3 Number of Calculation Fields/Master Items**

Qlik Sense allows the creation of reusable filter objects like Master Items, Calculated Fields which can simplify the process of applying consistent filters across multiple visualizations and dashboards.

