Qlik Analysis Of Road Safety And Accident Patterns In India

Define Problem / Problem Understanding

Specify The Business Problem

Technological advancement in transportation has minimised the distances but has increased the risk to life. Every year, accidents result in loss of lakhs of lives and serious injuries to crores of people. A study to analyse road safety and accident trends in India is to be conducted using Qlik Sense, a data analytics platform. This study involves examining data related to road incidents, such as types of accidents, locations, causes, and potentially factors contributing to road safety or risks. The use of Qlik Sense is a data-driven approach, utilizing visualizations and insights generated from the analysis to understand patterns and potentially inform strategies for improving road safety in India.

Business Requirements

The evaluation targets to offer precious insights into person demographics, coincidence patterns, and trouble areas. The number one recognition is on developing interactive and visually compelling dashboards to aid strategic making plans and operational improvements. The insights derived from this evaluation can be instrumental in making knowledgeable decisions, enforcing higher protection protocols, and making sure compliance with regulations.

Literature Survey

A literature survey for the Road Safety and Accident Patterns evaluation could contain studying and reviewing preceding studies, articles, reviews and figures at the topic. This should encompass records at the strategies and strategies used for analysing injuries data, in addition to the consequences and conclusions of those studies. It is suggested to discover instructional databases consisting of PubMed, IEEE Xplore, Google Scholar, and institutional repositories. Additionally, authorities reviews and guides can offer insights into the ultra-modern developments.

Social 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.

DATA PREPROCESSING

Data preprocessing is a crucial step in preparing raw data for analysis. Qlik employs several techniques to enhance data quality, consistency, and suitability for modeling. Key aspects of Qlik's data preprocessing include:

Automatic Data Preparation and Transformation:

- Qlik's AutoML (Automatic Machine Learning) feature automatically preprocesses datasets before model training.
- Preprocessing steps include data preparation and transformation to improve data quality and model accuracy.
- Techniques such as imputing missing values, encoding categorical variables, and generating new features are applied.

Feature Classification and Handling:

- Qlik classifies columns in the dataset based on their feature type (categorical, numeric, date, or free text).
- It checks for sparsity, constants, and high cardinality (unique values) in columns.
- Features with too many null values or low predictive value are excluded.

Feature Scaling and Standardization:

- Qlik calculates summary statistics for each column to guide feature scaling.
- Standardization ensures that features have similar scales, preventing dominance by large values.

DATA VISUALIZATION

Data visualization transforms raw data into easily understandable graphical representations. This report explores the core principles and benefits of data visualization, emphasizing its role in maximizing information value. Key benefits include:

Enhanced Clarity:

- Visual representations simplify complex data sets, allowing users to grasp trends and patterns at a glance.

Improved Communication:

- Data visualizations effectively communicate insights and findings to a broader audience, fostering better understanding and collaboration.

Identification of Trends and Patterns:

- Graphical representations readily reveal trends and correlations that might be obscured in raw data.

Detection of Outliers:

- Data visualizations can highlight outliers, data points that deviate significantly from the norm, which may hold valuable insights.

Informed Decision Making:

- By providing a clear understanding of the data, data visualization empowers users to make informed and data-driven decisions.

DASHBOARD

A dashboard is a graphical user interface (GUI) that displays information and data in an organized and easy-to-read format. Dashboards provide real-time monitoring and analysis of data, typically designed for specific purposes or use cases. They can be used across various industries, including business, finance, manufacturing, and healthcare, to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

STORYTELLING

A data story presents data and analysis in a narrative format, making information more engaging and easier to understand. A typical data story includes:

- **Introduction**: Sets the stage and explains the context for the data.
- **Body**: Presents the data and analysis in a logical and systematic way.
- **Conclusion**: Summarizes key findings and highlights their implications.

Data stories can be conveyed through reports, presentations, interactive visualizations, and videos.

PERFORMANCE TESTING

Applications of Data Filters:

- Selections within data allow users to filter based on individual fields or dimensions, enabling focused analysis of relevant subsets.
- Users can choose specific values to include or exclude, set numeric and date ranges, and create complex filters using logical conditions like AND, OR, and NOT.
- Advanced features include dynamic, interactive filters such as dropdowns, sliders, and drill-down capabilities, as well as automated filters based on user behavior or predefined templates.
- These filtering capabilities are crucial for refining data analysis, improving decision-making, and enhancing operational efficiency across various applications like business intelligence, financial analysis, healthcare, and operational management.

Use of Master Items/Calculated Fields:

- Qlik Sense enhances data analysis by enabling the creation of reusable filter objects such as Master Items and Calculated Fields, streamlining the application of consistent filters across multiple visualizations and dashboards.
- Master Items allow users to define and save frequently used dimensions and measures, like "Region" or "Total Sales," ensuring uniform definitions and calculations throughout various analyses.
- These items can be easily applied across different charts and tables, maintaining consistency and saving time.
- Calculated Fields enable users to create custom calculations, such as "Profit Margin," that automatically update across all visualizations using these fields when underlying data changes.
- This approach simplifies the application of complex filters, ensuring consistent criteria and reducing errors.
- Shared resources like Master Items and Calculated Fields promote standardized data analysis within teams, enhancing collaboration and ensuring uniform reporting.
- Overall, Qlik Sense's reusable filter objects improve efficiency, accuracy, and consistency in data analysis

and reporting.

Number of Graphs/Visualizations:

- There are a total of 15 visualizations created in this project:
    a. Accidents due to weather conditions
    b. Mobile Injuries
    c. Accidents By Wrong Side Driving
    d. Minors Killed
    e. Correlation between Speeding and Accidents
    f. Road Users Killed: Vehicles Distribution
    g. Pedestrian Killed: Gender
    h. Vehicle Contribution Towards Accident
    i. Pedestrian Killed
    j. Flash Signal Accidents
    k. Mobile Usage
    l. Traffic Light Signal Accidents
    m. Police Controlled Accidents
    n. Stop Signed Killed
    o. Drunk Driving Accidents