**Public Transportation Analysis**

**PROBLEM STATEMENT**

The project involves analyzing public transportation data to assess service efficiency, on time performance, and passenger feedback. The objective is to provide insights that support transportation improvement initiatives and enhance the overall public transportation experience. This project includes defining analysis objectives, collecting transportation data, designing relevant visualizations in IBM Cognos, and using code for data analysis.

**Analysis Objectives**

In this phase, we need to clearly define the specific objectives that will guide our analysis of Transportation Data. Our objectives will encompass several key aspects, including:

**Safety and Reliability Assessment:**

Analyze data related to accidents, breakdowns, or safety incidents to ensure the reliability and safety of public transportation services.

**Route Optimization:**

Optimize routes based on data analysis to reduce travel times, fuel consumption, and operational costs.

**Fare Pricing and Revenue Analysis:**

Analyze fare collection data to optimize pricing strategies, revenue generation, and fare evasion detection.

**Passenger Satisfaction Measurement:**

Gauge passenger satisfaction through surveys or feedback data. Determine specific satisfaction metrics, such as cleanliness, safety, courtesy of staff, and overall experience, aiming for a certain satisfaction score, like 80 out of 100.

**Data Collection**

Collecting transportation data encompasses diverse sources and methods crucial for managing efficient and passenger-friendly transit systems. Schedule data primarily originates from transit agencies and is disseminated through websites and mobile apps, offering passengers access to route timetables. Real-time updates, indispensable for commuters, rely on GPS tracking technology embedded in vehicles, enabling real-time tracking via mobile apps, public displays, and websites. Passenger feedback is gathered through surveys, customer service channels, and suggestion boxes, providing valuable insights into service quality and passenger satisfaction. Advanced sources include smart cards, WiFi, and beacons, while third-party data from app-based mobility services and traffic sensors can augment transit data. These multifaceted approaches to data collection form the foundation for optimizing transportation systems and enhancing the overall passenger experience.

**Visualization Strategy**

Planning to visualize insights using IBM Cognos involves several key steps. First, gather and organize the transportation data you've collected, including schedules, real-time updates, and passenger feedback, in a structured manner. Then, determine the specific insights and metrics you want to highlight in your dashboards and reports, such as on-time performance, passenger satisfaction, and service efficiency.

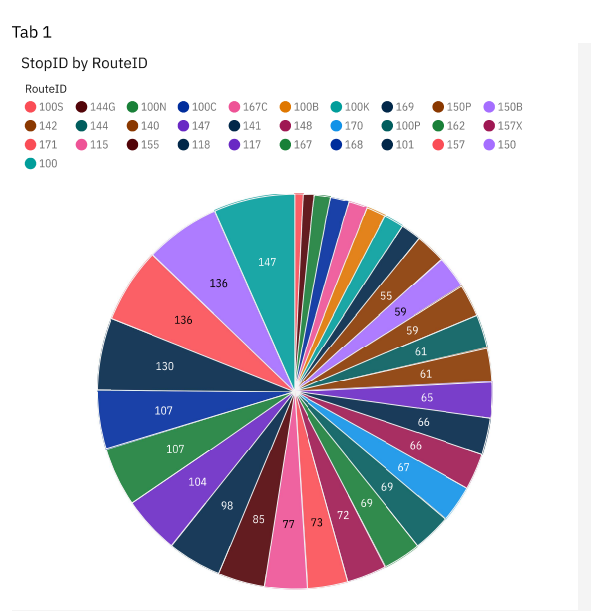
Next, within IBM Cognos, select the appropriate data visualization tools, charts, and widgets that best represent your data and insights. Utilize features like bar charts, line graphs, and heatmaps to effectively communicate trends and patterns. Ensure that the visualizations are easy to understand and user-friendly.

Consider the target audience for your dashboards and reports, whether it's transportation managers, stakeholders, or the general public. Tailor the design and content to meet their needs and expectations. For instance, create a real-time dashboard that displays live updates on vehicle locations and delays for commuters or a detailed performance report with historical data for management.

**Code Integration**

The use of code can significantly enhance various aspects of data analysis, particularly in transportation data analytics. Code is instrumental in automating repetitive tasks such as data cleaning, where it can swiftly identify and rectify missing values, outliers, and inconsistencies. Additionally, it empowers data transformation, allowing for the efficient aggregation of data, conversion of data types, and the creation of new variables, making the data more amenable to analysis. For statistical analysis, code can implement complex models and hypothesis tests, facilitating in-depth exploration of patterns and relationships within the data. Furthermore, machine learning algorithms can be harnessed through code to predict ridership, optimize routes, or detect anomalies. This integration of code not only streamlines the analysis process but also enables the handling of large datasets, real-time data streams, and customization of reports and visualizations, ultimately enhancing the depth and efficiency of transportation data analysis.

**Data Analyzed Using Cognos**

****