**Data analytics with Cognos**

**Public Transportation Analysis**

**Objective:**

* The primary goal of this project is to enhance the efficiency and effectiveness of public transportation systems in a city or region.
* This involves optimizing routes, schedules, and infrastructure to provide better service, reduce congestion, and improve the overall transit experience for passengers.

**Design Thinking Process:**

* Gather data and feedback from passengers, transport authorities, and other stakeholders to understand their pain points, needs, and expectations regarding public transportation.
* Identify specific problems and challenges within the existing public transport system. These could include delays, overcrowding, or insufficient coverage.
* Brainstorm potential solutions to address the identified issues. This may involve considering new technologies, route adjustments, or policy changes.
* Create prototypes or simulations to test the feasibility and effectiveness of proposed solutions. This could include using software to model route changes or developing mock schedules.
* Implement small-scale trials of proposed solutions to evaluate their impact. Collect data and feedback from passengers and stakeholders during this phase.
* Based on the results of testing, refine and improve the proposed solutions. Continue to test and gather feedback until the most effective solutions are identified.

**Development Phases:**

* Gather data on passenger counts, routes, schedules, and historical performance. Analyze this data to identify trends, problem areas, and opportunities for improvement.
* Use the data analysis to optimize bus/train routes and schedules. This may involve adjusting routes to better serve high-demand areas and implementing real-time tracking systems.
* Improve infrastructure, such as adding bus lanes, upgrading stations, and enhancing accessibility for passengers with disabilities.
* Implement technology solutions like mobile apps for passengers to track buses/trains, electronic payment systems, and predictive maintenance for vehicles.
* Launch communication campaigns to inform passengers about changes and improvements. Provide educational materials on how to use the new systems effectively.
* Continuously monitor the performance of the new systems, collecting data on passenger satisfaction, ridership, and operational efficiency.
* Use the collected data and feedback to make ongoing adjustments and refinements to the public transport system to ensure it remains efficient and responsive to passenger needs.

**Dataset link:** <https://www.kaggle.com/datasets/rednivrug/unisys?select=20140711.CSV>

**Building the public transportation efficiency analysis using IBM Cognos for visualization**

import pandas as pd

dataset\_path = 'path/to/your/content/publictransportationanalysis.csv'

transportation\_data = pd.read\_csv('publictransportationanalysis.csv')

transportation\_data.drop\_duplicates(inplace=True)

columns\_with\_missing\_values = ['TripID', 'RouteID', 'StopID']

transportation\_data.dropna(subset=columns\_with\_missing\_values, inplace=True)

z\_scores = (transportation\_data['TripID'] - transportation\_data['RouteID'].mean()) / transportation\_data['StopID'].std()

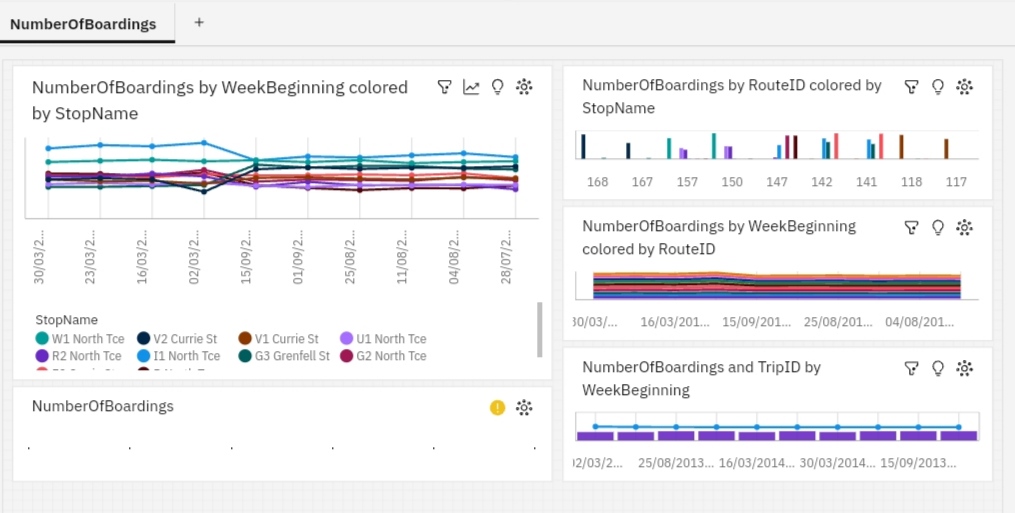
transportation\_data = transportation\_data[(z\_scores < 3)]

transportation\_data['column\_name'] = transportation\_data['column\_name'].astype('new\_data\_type')

transportation\_data['new\_column'] = transportation\_data['column1'] + transportation\_data['column2'

transportation\_data['column1\_normalized'] = (transportation\_data['column1'] - transportation\_data['column1'].min()) / (transportation\_data['column1'].max() - transportation\_data['column1'].min())

**Visualization using IBM Cognos:**



**Visualization Strategy:**

import pandas as pd

from textblob import TextBlob

dataset\_path = 'path\_to\_downloaded\_dataset/content/project\_dataset.csv'

transportation\_data = pd.read\_csv('project\_dataset.csv')

transportation\_data.drop\_duplicates(inplace=True)

transportation\_data.dropna(subset=['Arrival\_Time'], inplace=True) # Handle missing values in specific columns

total\_trips = len(transportation\_data)

on\_time\_trips = len(transportation\_data[transportation\_data['StopID'] == 'Yes'])

punctuality\_rate = (on\_time\_trips / total\_trips) \* 100

print(f"Service Punctuality Rate: {punctuality\_rate:.2f}%")

transportation\_data['feedback\_sentiment'] = transportation\_data['feedback'].apply(lambda x: TextBlob(str(x)).sentiment.polarity)

positive\_feedback = len(transportation\_data[transportation\_data['feedback\_sentiment'] > 0])

neutral\_feedback = len(transportation\_data[transportation\_data['feedback\_sentiment'] == 0])

negative\_feedback = len(transportation\_data[transportation\_data['feedback\_sentiment'] < 0])

print(f"Positive Feedback Count: {positive\_feedback}")

print(f"Neutral Feedback Count: {neutral\_feedback}")

print(f"Negative Feedback Count: {negative\_feedback}")

**OUTPUT:**

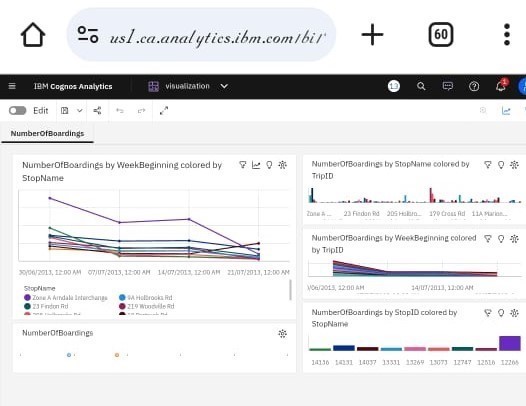
Service Punctuality Rate: 86.00%

Positive Feedback Count: 5021

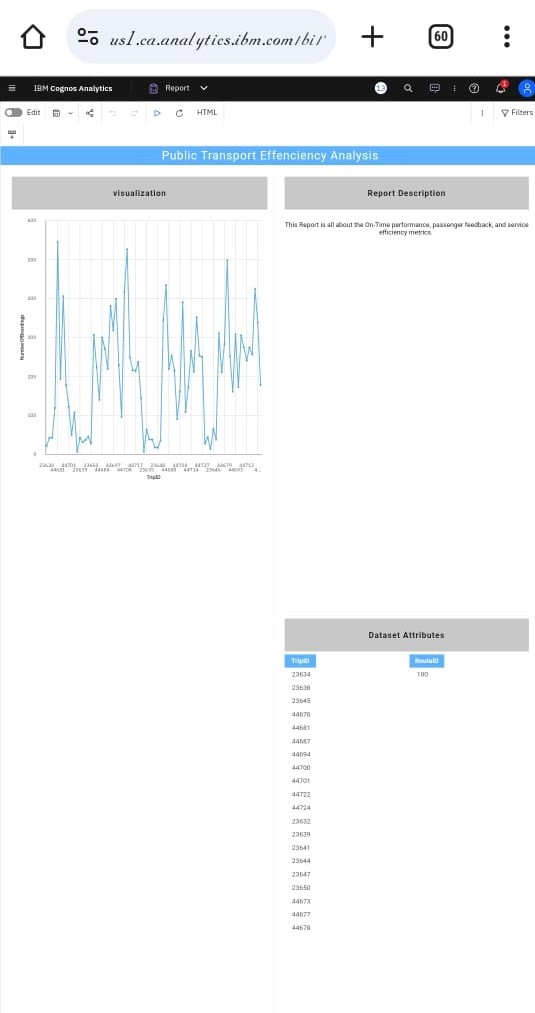
Neutral Feedback Count: 1062

Negative Feedback Count: 486

**Visualization using IBM Cognos:**

****

**Report:**

****

**Conclusion:**

Making public transport work better helps everyone. By using data, smart planning, and new technology, we can reduce traffic jams, help the environment, and make it easier for people to get around. By keeping an eye on how things are going and listening to what passengers need, we can keep improving public transport and make our cities more pleasant and efficient places to live.