

Public transportation analysis

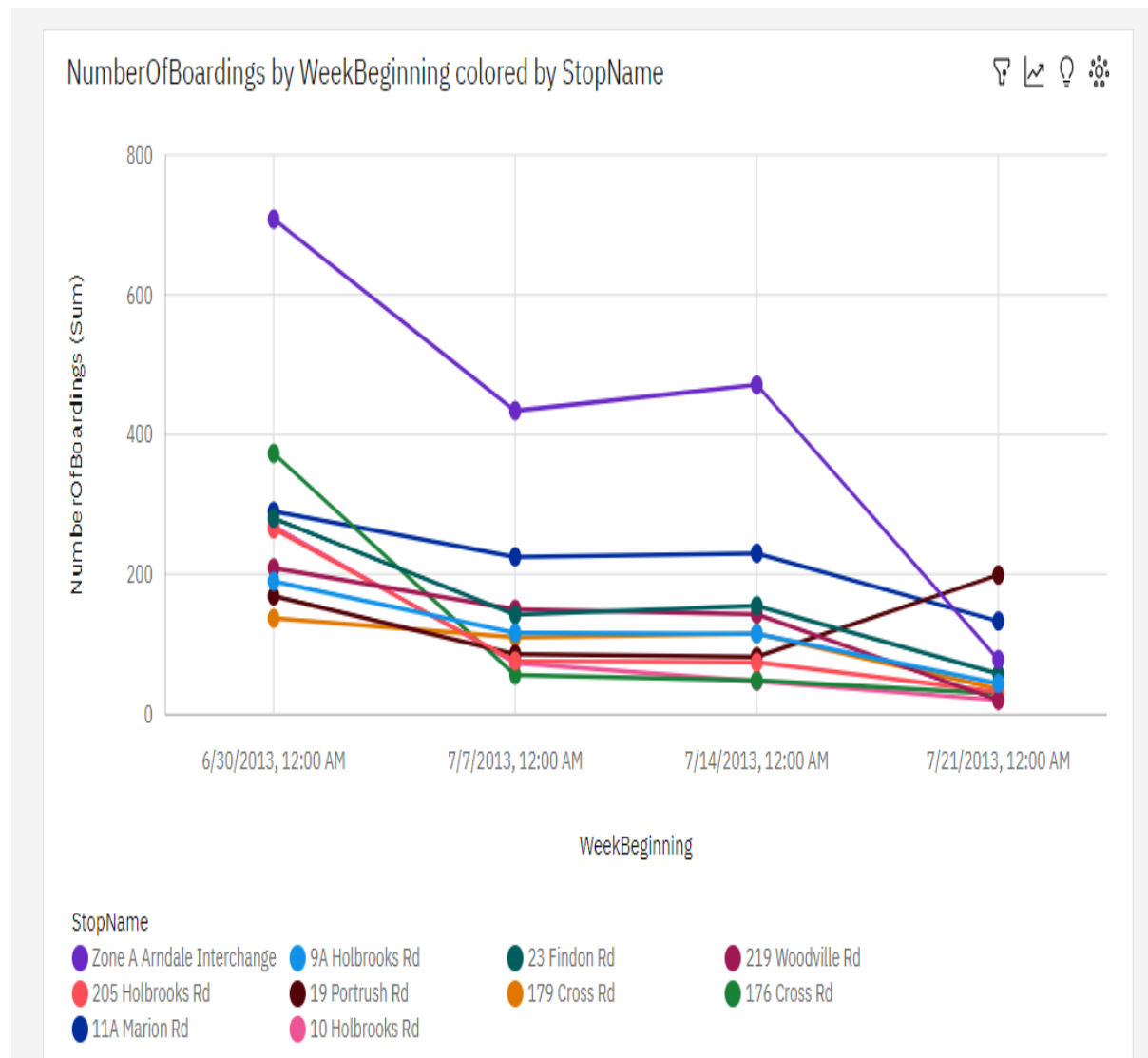
Phase 4: Development part 2



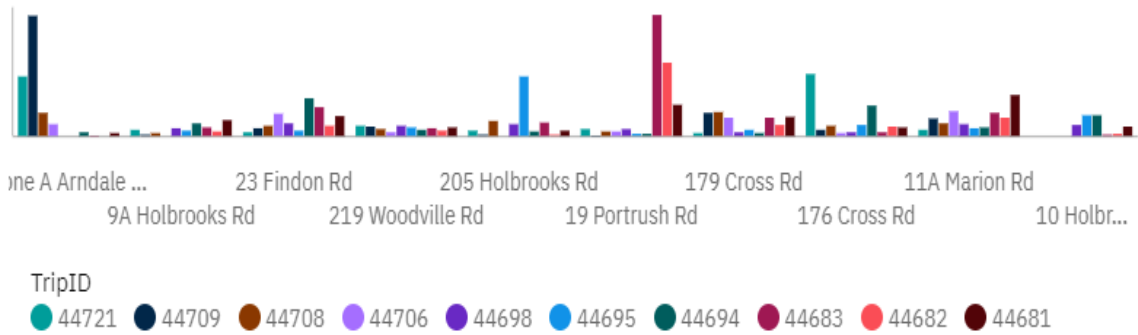
Visualization:

Visualization in data analytics is the process of representing data graphically to gain insights, identify patterns, and communicate findings effectively. Data visualization plays a crucial role in the data analysis process as it helps analysts and stakeholders better understand complex datasets, make data-driven decisions, and communicate their findings to others.

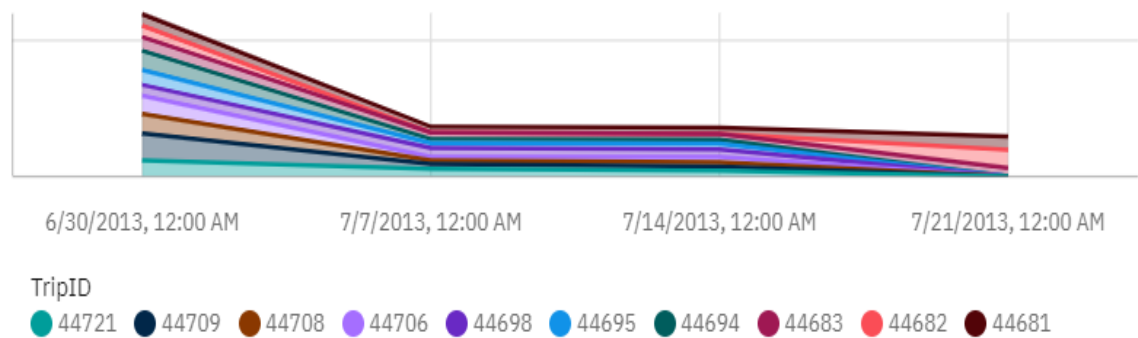
Using IBM Cognos for Visualization:



NumberOfBoardings by StopName colored by TripID



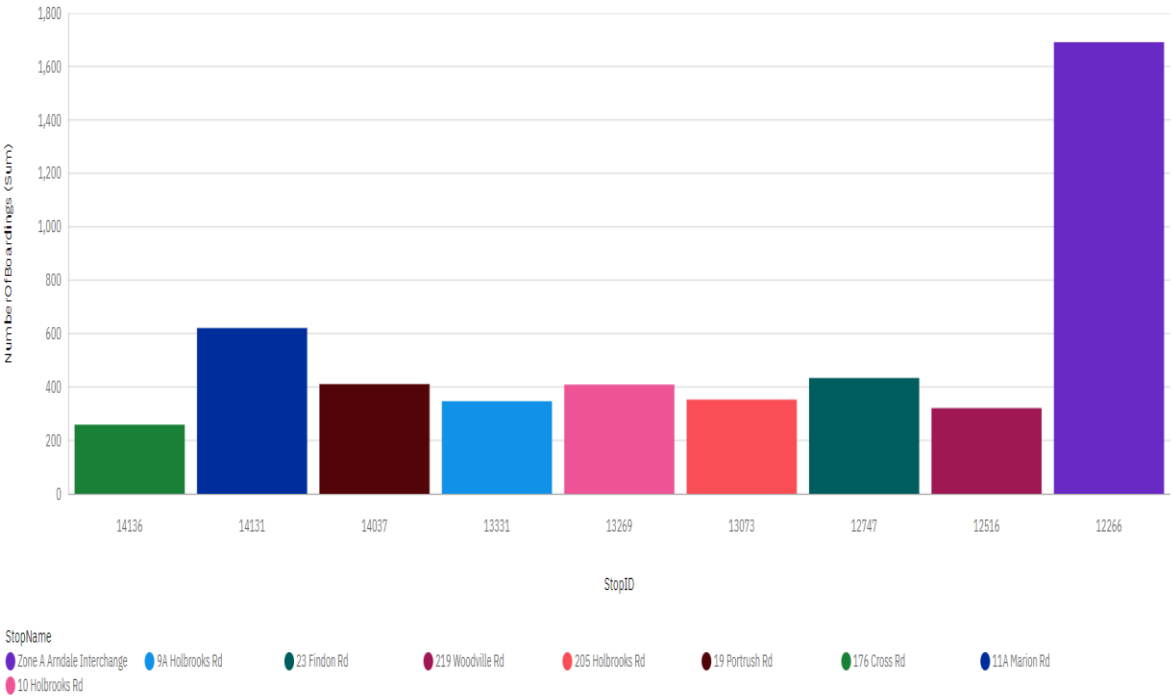
NumberOfBoardings by WeekBeginning colored by TripID



NumberOfBoardings



NumberOfBoardings by StopID colored by StopName



Report:

Public Transport Efficiency Analysis

visualization

NumberOfBoardings (Sum)

TripID

Report Description

This Report is all about the On-Time performance, passenger feedback, and service efficiency metrics.

Dataset Attributes

TripID	RouteID
23634	100
23638	
23645	
44676	
44681	
44687	
44694	
44700	
44701	
44722	
44724	
23632	
23639	
23641	
23644	
23647	
23650	
44673	
44677	
44678	

Code for advance Data Analysis:

```
import pandas as pd

# Load data from a CSV file (assuming the data is in a file
named 'data.csv')
data = pd.read_csv('project_dataset.csv')

# Calculate punctuality rate (assuming 'NumberOfBoardings'
represents successful boardings)
total_boardings = data['NumberOfBoardings'].sum()
total_trips = len(data['TripID'].unique())

punctuality_rate = (total_boardings / total_trips) * 100
print(f"Punctuality Rate: {punctuality_rate:.2f}%")
```