# PUBLIC TRANSPORTATION EFFICIENCY ANALYSIS

### Abstract:

Public transport has become one of the major transport options, especially when it comes to reducing motorized individual transport and achieving sustainability while reducing emissions, noise and so on. The use of public transport data has evolved and rapidly improved over the past decades. Indeed, the availability of data from different sources, coupled with advances in analytical and predictive approaches, has contributed to increased attention being paid to the exploitation of available data to improve public transport service. In this paper, we review the current state of the art of public transport data sources. More precisely, we summarize and analyse the potential and challenges of the main data sources. In addition, we show the complementary aspects of these data sources and how to merge them to broaden their contributions and face their challenges. This is complemented by an information management framework to enhance the use of data sources. Specifically, we seek to bridge the gap between traditional data sources and recent ones, present a unified overview of them and show how they can all leverage recent advances in data-driven methods and how they can help achieve a balance between transit service and passenger behaviour.

### Introduction:

Public transport provides an essential service whose relevance is increasingly recognized. Indeed, it helps to reduce road congestion, air pollution and energy as well as oil consumption. However, managing the public daily commute network is a difficult task, particularly today with rapid urbanization and the associated population increases, especially in developing countries. Therefore, public transportation systems need to adopt appropriate tools and take advantage of the available data to address these challenges. Indeed, it is shown in several studies that the acquisition of reliable information and data is crucial for the proper functioning of public transport information systems. We note, first, that, in general, a public transport (or transit) system is comprised of supply (presented through transit service) and demand (reflecting passenger behavior). The ultimate goal is then to achieve a balance in which transit services meet the needs of passengers, using the available information. On the one hand, classic data sources on transit service, which are generally provided by agencies, are based on transit schedules, stations and route information. However, these static data are not informative in terms of disruptions (e.g., delays, interruptions) as they are based entirely on schedules, which are expectations rather than observations of services. Moreover, we note that the actual observation of schedules which may or may not be published reveals that a schedule does not specify certain details, as is the case, e.g., when it is said that a bus or train runs every ten minutes, etc. Rather than using schedules, public transit services may also run on demand and other options may apply as well. On the other hand, traditional manual approaches to collect information on passenger demand and behaviour, such as household surveys, have shown limitations to cope with the current challenges alone. Therefore, in recent years, big data has emerged as an area to allow new perspectives on improving public transportation systems. That is, the availability of these data, coupled with advanced approaches to data analysis, has paved the way for the use of massive data sets in public transport. As a result, it is widely accepted today that the application of big data to public transport problems will offer new perspectives which were previously inaccessible by traditional transport data and analysis approaches. Therefore, the public transport community is increasingly emphasizing the importance of developing commonly collected data sources on public transport, as well as more powerful analytical tools. However, the massive increase in data availability poses many growing challenges in handling transit data, including their validation, in order to fully benefit from their potential.

# Enhancing Public Transportation Experience Through Data

### Analysis Project Description:

The project aims to leverage data analysis to improve the efficiency and overall quality of public transportation services. By assessing on-time performance, passenger satisfaction, and service efficiency, this project intends to provide valuable insights to support transportation improvement initiatives. Key components of this project include defining analysis objectives, collecting transportation data, designing informative visualizations using IBM Cognos, and integrating code for data analysis.

### Project Phases:

**Analysis Objectives:**

* Define specific objectives for analyzing public transportation data, focusing on on-time performance, passenger satisfaction, and service efficiency. This objectives will guide theentire project.

### Data Collection:

* Identify and gather transportation data from various sources, including schedules, real-timeupdates, and passenger fееdback. Ensure data quality and relevance to support meaningfulanalysis.

### Visualization Strategy:

* Plan the visualization strategy using IBM Cognos. Create informative dashboards and reports topresent the insights derived from the data analysis, making them accessible and actionable forstakeholders.

### Code Integration:

* Determine which aspects of the analysis can benefit from code integration, such as data

cleaning, transformation, and statistical analysis. This will enhance the accuracy and depth of theinsights generated.

### Project Goals:

1. Improve public transportation sеrvicеs by identifying areas for enhancement.
2. Enhance on-time performance to provide a more reliable service.
3. Increase passenger satisfaction by addressing their fееdback and concerns.
4. Optimize service efficiency to reduce operational costs and minimize delays.

### Expected Outcomes:

1. Actionable insights that can guide transportation improvement initiatives.
2. Informative visualizations that provide a clear understanding of the data analysis results.
3. Improved public transportation sеrvicеs, resulting in a better overall passenger experience.

### Project Team:

1. Data Analysts: Responsible for data collection, cleaning, and analysis.
2. Visualization Experts: Create informative dashboards and reports in IBM Cognos.
3. Programmers: Integrate code for data analysis and transformation.
4. Transportation Experts: Provide domain knowledge and insights.

# INNOVATION IN PUBLIC TRANSPORTATION EFFICIENCY ANALYSIS:

In recent years, there have been several innovations in public transportation efficiency analysis:

1. **Big Data Analytics**: Public transportation agencies are increasingly using big data to analyse vast amounts of information from various sources, including GPS tracking, passenger counts, and fare card transactions. This helps identify congestion points, optimize routes, and predict demand.
2. **Machine Learning Algorithms**: Machine learning is being used to develop predictive models for public transportation. These models can forecast ridership, optimize schedules, and even predict maintenance needs, leading to more efficient operations.
3. **Real-time Passenger Information Systems**: Mobile apps and digital signage at bus stops and train stations provide real- time information on arrivals and departures. This not only improves the passenger experience but also helps transportation agencies manage resources more effectively.
4. **Demand-Responsive Transportation**: On-demand services like ride-sharing and micro transit are being integrated into public transportation networks. Algorithms determine the best routes and pickup/drop-off points based on passenger demand, reducing inefficiencies.
5. **Sustainability Metrics**: With a growing focus on sustainability, public transportation efficiency analysis now includes metrics related to environmental impact. This includes evaluating the energy efficiency of vehicles and the reduction of greenhouse gas emissions.

These innovations aim to make public transportation more efficient, reliable, and attractive to passengers while reducing the environmental footprint and operational costs.

### BIG DATA ANALYSIS:

Big data analytics plays a pivotal role in enhancing public transportation efficiency analysis in several ways:

1. **Demand Forecasting**: By analysing historical and real-time data, transportation agencies can predict passenger demand accurately. This helps in optimizing routes, scheduling, and resource allocation, ensuring that services match actual needs.
2. **Traffic Management:** Big data analytics can monitor traffic patterns, congestion, and incidents in real-time. Transportation authorities can use this information to adjust routes dynamically, reduce delays, and improve overall efficiency.
3. **Predictive Maintenance**: Analysing data from sensors on vehicles and infrastructure allows for predictive maintenance. By identifying potential issues before they lead to breakdowns, agencies can minimize service disruptions and reduce maintenance costs.
4. **Rider Behaviour Analysis**: Understanding passenger behaviour through data analysis can lead to improved services. For instance, analysing fare card data can reveal popular routes and peak travel times, helping agencies adjust schedules accordingly.
5. **Route Optimization**: Big data can be used to optimize bus and train routes. By considering factors like traffic, weather, and special events, transportation agencies can find the most efficient paths for vehicles.
6. **Real-time Passenger Information**: Mobile apps and digital signage provide real-time updates to passengers. These systems rely on big data to relay accurate arrival times and service alerts, enhancing the passenger experience.
7. **Fare Pricing Strategies**: Data analysis can inform fare pricing strategies. Agencies can adjust fares based on demand, time of day, or other factors to encourage ridership during off-peak hours and maximize revenue.
8. **Environmental Impact Assessment**: Big data can quantify the environmental impact of public transportation systems, considering factors like fuel consumption and emissions. This information can guide efforts to reduce the carbon footprint.
9. **Security and Safety**: Analysing data from surveillance cameras and security systems can enhance passenger safety and deter criminal activity. Algorithms can also identify unusual patterns, potentially flagging security threats.
10. **Integration with Other Modes:** Public transportation often connects with other modes of transit, such as bike-sharing or ride-sharing services. Big data can facilitate seamless integration, making it easier for passengers to switch between modes.
11. **Budget Allocation**: Data-driven insights help agencies allocate their budgets more effectively. By identifying areas

where investments will have the most impact, agencies can make informed decisions about infrastructure improvements and service expansions.

Overall, big data analytics empowers public transportation agencies to make data-driven decisions that lead to more efficient, reliable, and passenger-friendly services while optimizing resource allocation and minimizing operational costs.

### MACHINE LEARNING ALGORITHM:

Machine learning algorithms have gained significance in public transportation efficiency analysis, offering various benefits and applications:

1. **Demand Prediction**: Machine learning models can forecast passenger demand based on historical data, weather conditions, and special events. This helps transportation agencies optimize schedules and allocate resources efficiently.
2. **Dynamic Route Optimization:** ML algorithms can adapt routes in real-time based on traffic congestion, accidents, or other disruptions. This dynamic adjustment minimizes delays and enhances efficiency.
3. **Predictive Maintenance**: Machine learning can predict when vehicles and infrastructure components require

maintenance. By analysing sensor data, agencies can schedule maintenance proactively, reducing unexpected breakdowns and associated service disruptions.

1. **Fare Pricing Optimization**: ML models analyse fare card data to optimize pricing strategies. Agencies can adjust fares dynamically to incentivize off-peak travel or balance demand on different routes.
2. **Passenger Flow Analysis**: Machine learning can analyse passenger flow within transportation hubs. This information helps optimize station layouts and facilitates efficient passenger movement.
3. **Anomaly Detection:** ML algorithms can identify anomalies in real-time data, such as unusual passenger behaviour or security threats. This enhances safety and security on public transportation systems.
4. **Predictive Delay Management**: Machine learning can predict delays caused by various factors like weather, accidents, or events. Agencies can proactively communicate delays to passengers and suggest alternative routes or transportation options.
5. **Energy Efficiency**: ML models analyse vehicle performance data to optimize energy usage. This can lead to more fuel- efficient driving patterns, reducing operational costs and environmental impact.
6. **Optimized Crew Scheduling**: Machine learning helps create efficient crew schedules, considering factors like driver availability, rest times, and route complexities. This minimizes labour costs while ensuring reliable service.
7. **Passenger Behaviour Analysis**: ML algorithms analyse passenger behaviour data to understand preferences, peak travel times, and route choices. This information guides service adjustments to better meet passenger needs.
8. **Safety Improvements**: ML-based predictive analytics can identify potential safety risks, such as intersections prone to accidents. Transportation agencies can prioritize safety enhancements in these areas.
9. **Integration with Other Modes**: Machine learning helps integrate various modes of transportation seamlessly. For example, ML algorithms can suggest optimal combinations of public transportation, bike-sharing, and ride-sharing for a passenger's journey.
10. **Capacity Planning**: ML models can predict peak passenger loads on different routes and vehicles. This information aids in planning for increased capacity during busy times.

Machine learning algorithms empower transportation agencies to make data-driven decisions, optimize operations, enhance the passenger experience, and ultimately achieve greater efficiency in public transportation systems.

### REAL TIME PASSENGER INFORMATION SYSTEM:

Real-time passenger information systems (RTPIS) play a crucial role in public transportation efficiency analysis by providing accurate and up-to-date information to passengers and transportation agencies. Here are some ways RTPIS contributes to efficiency analysis:

1. **Improved Passenger Experience**: RTPIS apps, digital signage, and announcements at stations and stops provide real-time updates on arrivals, departures, delays, and service disruptions. Passengers can make informed decisions, reducing

frustration and improving their overall experience.

1. **Enhanced Service Reliability**: By keeping passengers informed in real time, RTPIS helps transportation agencies manage service disruptions more effectively. Agencies can respond quickly to incidents, minimize delays, and maintain reliable operations.
2. **Rider Behaviour Analysis:** RTPIS systems can collect data on passenger behaviour, such as the stations they use and the times they travel. Analysing this data helps agencies understand ridership patterns, allowing for optimized scheduling and resource allocation.
3. **Data for Performance Metrics:** Real-time information feeds into key performance indicators (KPIs) used by

transportation agencies. Metrics such as on-time performance and schedule adherence can be calculated accurately, allowing agencies to track and improve their efficiency.

1. **Resource Allocation**: Transportation agencies can use RTPIS data to allocate resources dynamically. For example, if a particular station experiences higher passenger demand than usual, agencies can dispatch more vehicles or adjust schedules accordingly.
2. **Maintenance Alerts**: RTPIS systems can provide alerts related to maintenance issues or vehicle breakdowns in real time. This allows agencies to respond swiftly to ensure that maintenance tasks do not disrupt service.
3. **Demand Response:** By monitoring passenger flows in real time, agencies can identify overcrowding and respond by adding more vehicles or adjusting routes. This helps optimize service to match demand.
4. **Operational Efficiency**: Real-time data on vehicle locations and conditions can be used to optimize operations. For example, agencies can adjust the spacing between vehicles to maintain even service intervals.
5. **Feedback Mechanism:** RTPIS often includes mechanisms for passengers to provide feedback, such as reporting incidents or suggesting improvements. This feedback can be valuable for ongoing analysis and service improvements.
6. **Integration with Other Services:** RTPIS can be integrated with other transportation services, like ride-sharing or bike- sharing, to provide passengers with seamless multi-modal options. This integration can enhance the efficiency and convenience of public transportation networks.

In summary, real-time passenger information systems are a valuable tool for public transportation efficiency analysis. They not only improve the passenger experience but also provide critical data that transportation agencies can use to optimize operations, enhance reliability, and respond to changing demand patterns effectively.

### DEMAND RESPONSIVE TRANSPORTATION:

Demand-responsive transportation (DRT) is a flexible transportation service that adapts to passenger demand and can significantly contribute to public transportation efficiency analysis in the following ways:

1. **Optimized Resource Allocation**: DRT systems use real-time data and algorithms to allocate vehicles dynamically based on passenger demand. This optimizes the use of resources, reducing the number of empty or underutilized vehicles on the road.
2. **Reduced Operating Costs:** By aligning vehicle deployment with demand, DRT can reduce operational costs, including fuel, maintenance, and labour expenses. This leads to more cost-effective public transportation services.
3. **Minimized Congestion**: DRT systems can help alleviate congestion by efficiently routing vehicles to pick up and drop off passengers. This reduces traffic congestion and improves overall transportation system efficiency.
4. **First/Last-Mile Connectivity**: DRT often focuses on providing first/last-mile connectivity to traditional public

transportation services like buses and trains. This improves accessibility and encourages more people to use public transit, enhancing the efficiency of the entire network.

1. **Data-Driven Decision-Making:** DRT systems generate a wealth of data on passenger preferences, travel patterns, and service utilization. This data can be analysed to make data-driven decisions for optimizing routes, schedules, and vehicle deployment.
2. **Integration with Fixed Routes**: DRT can be integrated with fixed-route public transportation services. Analysing the integration's effectiveness helps transportation agencies balance fixed-route services with on-demand offerings to maximize efficiency.
3. **Service Customization:** DRT services can be tailored to meet specific community needs. Analysing passenger demand and preferences helps fine tune service parameters for different regions or times of day, increasing efficiency.
4. **Environmental Impact Analysis:** DRT services can be designed to reduce environmental impact by using electric or low- emission vehicles. Efficiency analysis can assess the environmental benefits and cost-effectiveness of these options.
5. **Improved Accessibility**: DRT can enhance accessibility for individuals with disabilities or those living in underserved areas. Analysing the impact on accessibility helps agencies meet regulatory requirements and serve diverse populations more efficiently.
6. **Real-Time Passenger Feedback**: DRT services often incorporate passenger feedback mechanisms. Analysing this

feedback helps transportation agencies identify areas for improvement and make real-time adjustments to service quality.

1. **Route Optimization Algorithms**: Advanced algorithms used in DRT systems can continually optimize routes based on real-time traffic conditions, ensuring that passengers reach their destinations efficiently.
2. **Reduced Empty Passenger Miles:** DRT aims to reduce empty passenger miles by matching passengers traveling in the same direction. This reduces unnecessary vehicle travel and enhances transportation efficiency.
3. **Peak Demand Management**: DRT can be used to address peak demand periods, such as rush hours or special events. Efficiency analysis helps agencies plan for these periods and allocate resources effectively.

Incorporating demand-responsive transportation into public transportation networks and conducting ongoing efficiency analysis can lead to more responsive, cost-effective, and passenger-friendly services while contributing to the overall efficiency of the transportation system.

Sustainability metrics are vital in public transportation efficiency analysis as they help assess the environmental, social, and economic impact of transit systems. Here are some key sustainability metrics used in this context:

### Sustainability Metrics:

1. **Greenhouse Gas Emissions:** Measuring the amount of greenhouse gases (e.g., CO2) produced per passenger mile or kilometre travelled provides insight into the environmental impact of public transportation. Lower emissions indicate a more sustainable system.
2. **Fuel Efficiency**: Analysing fuel consumption per vehicle or per passenger can help assess the efficiency of the transportation system. Efforts to reduce fuel usage contribute to sustainability.
3. **Energy Consumption**: Monitoring the energy consumption of public transit vehicles and infrastructure helps identify opportunities to reduce energy use and promote sustainable practices.
4. **Vehicle Emissions Standards**: Evaluating compliance with emissions standards, such as Euro VI for buses, ensures that public transportation fleets meet environmental regulations.
5. **Ridership and Mode Share:** Measuring the percentage of commuters who use public transit compared to other modes of transportation (e.g., private cars) provides an indication of the system's success in reducing individual vehicle trips.
6. **Air Quality Improvement**: Assessing the impact of public transportation on local air quality helps quantify the system's contribution to reducing pollutants that affect human health.
7. **Noise Pollution:** Analysing noise levels generated by public transit vehicles and infrastructure helps mitigate noise pollution in urban areas, contributing to a more sustainable and liveable environment.
8. **Accessibility Metrics**: Evaluating the accessibility of public transit services for people with disabilities and underserved communities ensures equitable access to sustainable transportation options.
9. **Land Use and Urban Development**: Assessing the impact of public transit on urban development, such as reducing urban sprawl and promoting transit-oriented development, supports sustainability goals.
10. **Waste Reduction**: Analysing waste generation and recycling efforts within public transportation systems helps reduce the environmental footprint.
11. **Economic Efficiency**: Sustainability metrics can include cost-effectiveness and return on investment assessments, ensuring that sustainability efforts align with budgetary constraints.
12. **Social Equity**: Evaluating how public transit systems serve diverse populations, including low-income communities, helps address social equity concerns and promote inclusive, sustainable transportation.
13. **Vehicle Fleet Electrification**: Tracking the adoption of electric or low-emission vehicles in public transit fleets measures progress toward sustainable transportation.
14. **Operational Efficiency**: Assessing the efficiency of operations, including routes, schedules, and maintenance practices, helps reduce resource waste and supports sustainability.
15. **Community Engagement**: Involving the community in sustainability initiatives and gathering feedback helps ensure that public transportation systems meet the needs and expectations of residents.

By analysing these sustainability metrics, public transportation agencies can make informed decisions, set sustainability goals, and implement strategies to reduce their environmental impact while providing efficient and equitable transit services to communities.

**DEVELOPMENT PART 1: LOADING AND PREPROCESSING THE DATASET**

**PROJECT DESCRIPTION**

In this part,building the project by loading and preproces ing the dataset.Start building the public transportation e ficiency analysis using IBM Cognos for visualization.

**DEFINE ANALYSIS OBJECTIVES**

1. Determine patterns of ridership, including peak hours, popular routes, and demographics of riders. This can help optimize service frequency and capacity.
2. Assess the punctuality of public transportation services. Define standards for on-time arrivals and departures, and analyze deviations from these standards.
3. Evaluate existing routes to identify underused or overcrowded segments. Optimize routes to reduce travel times and enhance overall system efficiency.
4. Examine service reliability by measuring factors such as the frequency of breakdowns, disruptions, and delays. This information can be used to improve maintenance and infrastructure.
5. Analyze the costs associated with operating public transportation services, including fuel, maintenance, labor, and infrastructure. Determine cost-effective measures to reduce expenses while maintaining or improving service quality.
6. Assess the environmental impact of public transportation, including emissions and fuel consumption. Identify opportunities for adopting more sustainable practices.
7. Measure customer satisfaction through surveys and feedback analysis. Understand the areas where passengers are most satisfied or dissatisfied to make targeted improvements.
8. Evaluate the accessibility of public transportation for all passengers, including those with disabilities. Identify areas where improvements are needed to make the system more inclusive.
9. Analyze safety-related incidents and accidents within the public transportation system. Develop strategies to minimize safety risks for passengers and employees.

Assess revenue collection methods and the efficiency of fare collection systems. Identify opportunities to reduce fare evasion and increase revenue.

**DATA COLLECTION**

Identify the sources of transportation data.This could include public transportation authorities government agenta or any other open portals.Here are some key sources of transportation data:

1. Cameras installed at key points on roadways to monitor tra fic flow, congestion, and incidents.

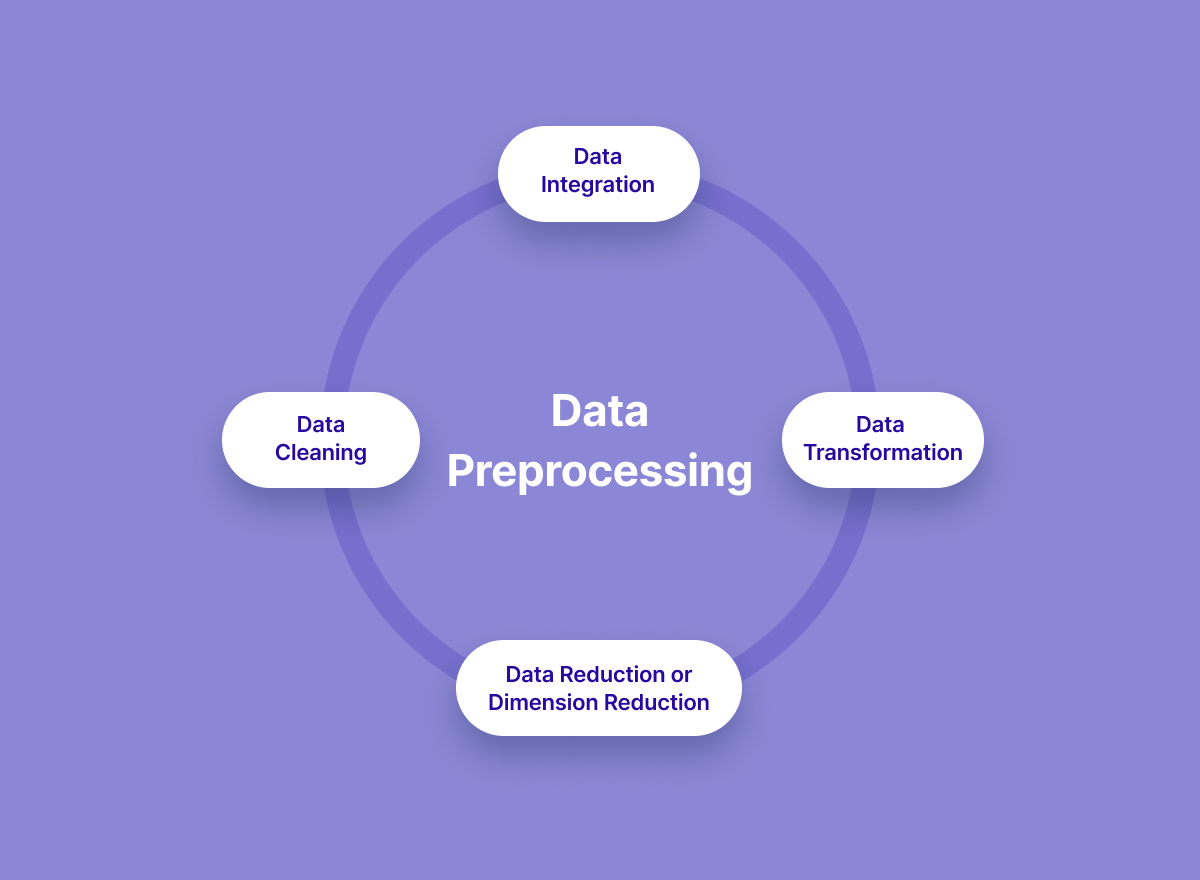
2. Data from GPS-enabled devices and navigation apps, which track the location and movement of vehicles.

3. Mobile apps, such as Waze and Google Maps, collect data from users to provide real-time traffic information.

4.Public transportation authorities collect data on routes, schedules, ridership, and service disruptions.

5.Toll systems capture data on vehicles passing through toll booths, including entry and exit times.

6.In-road or above-road sensors that monitor traffic speed, volume, and congestion.



**STEPS FOR DATA PREPROCESSING AND CLEANING**

Data processing and cleaning in transportation data using Python and Pandas typically involves several steps:

1. Data Import: Import the transportation data into a Pandas DataFrame. You can read data from various format like CSV, Excel, or databases.

import pandas as pd

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

2.Data Exploration: Examine the data to understand its structure, column names, and missing values.

df.head()  # View the first few rows

df.info()  # Get data types and non-null counts

3.Handling Missing Values: Deal with missing data by either removing rows with missing values or imputing values.

df.dropna()   # Remove rows with missing values

df.fillna(0)  # Replace missing values with 0

4.Data Cleaning:Clean and standardize data by removing duplicates, correcting data types, and renaming columns.

df.drop\_duplicates() # Remove duplicate rows

df['column\_name'] = df['column\_name'].astype(int) # Convert data types

df.rename(columns={'old\_name': 'new\_name'}, inplace=True)  # Rename columns

5.Data Transformation: Apply transformations to columns, such as datetime parsing or creating new features.

df['date'] = pd.to\_datetime(df['date\_column']) # Convert to datetime

df['hour'] = df['timestamp\_column'].dt.hour  # Extract hour from timestamp

6. Data Filtering:Filter data based on specific criteria.

df\_filtered = df[df['column\_name'] > 100]  # Filter rows where a column value is greater than 100

7.Data Aggregation: Aggregate data using groupby to gain insights.

mean\_values = df.groupby('category')['value\_column'].mean()  # Calculate mean values by category

8.Data Export:Save the cleaned data to a new file or database.

df.to\_csv('cleaned\_transportation\_data.csv', index=False)  # Save to a new CSV file

**PROGRAM**

#loading and preprocessing of dataset #import the required libraries import pandas as pd

import matplotlib.pyplot as plt # Step 1: Load the dataset

data = pd.read\_csv("transportdata.csv") # Step 2: Explore the data

# Display the first few rows to get an overview print(data.head()) # Step 3: Data Preprocessing

# Handle missing values (replace NaN with 0) data = data.fillna(0) # Data Cleaning (remove duplicates) data = data.drop\_duplicates() # Step 4: Exploratory Data Analysis (EDA)

# Calculate statistics and generate plots to explore the data # Example of how to calculate statistics (mean, median)

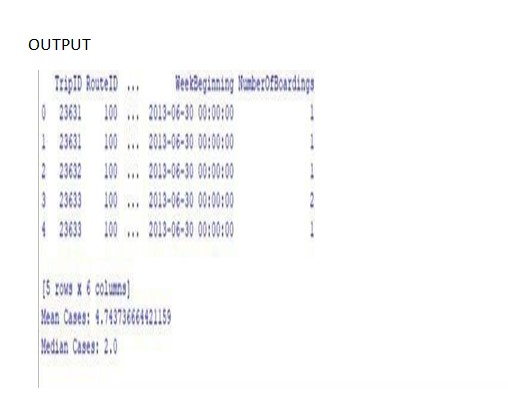
mean\_NumberOfBoardings = data['NumberOfBoardings'].mean() median\_NumberOfBoardings = data['NumberOfBoardings'].median() print("Mean Cases:", mean\_NumberOfBoardings)

print("Median Cases:", median\_NumberOfBoardings) # Step 5: Data Visualization (optional) column\_name = 'NumberOfBoardings' # Set the size of the plot plt.figure(figsize=(2,10))

plt.hist(data[column\_name], color='blue',alpha=0.5) plt.title('public transport analysis') plt.xlabel(column\_name) plt.ylabel('TripID') plt.xlim([2, 5])

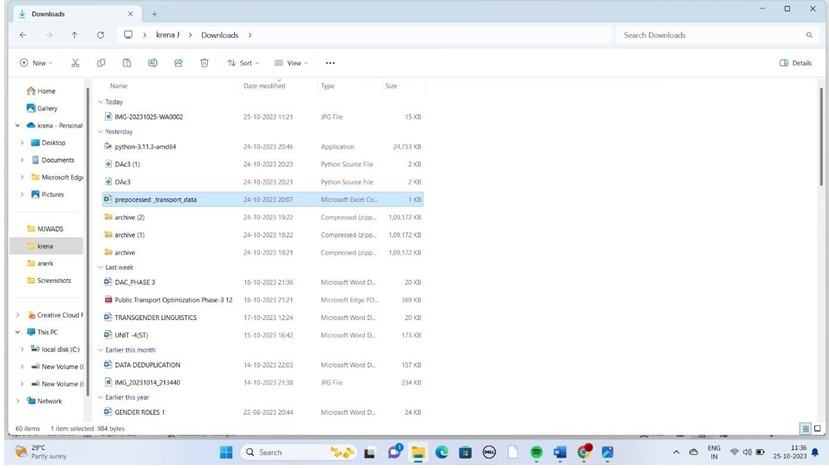
plt.ylim([0,1000])

plt.show()



**DATA PREPROCESSING**

It refers to the cleaning, transforming, and integrating of data in order to make it ready for analysis. The goal of data preprocessing is to improve the quality of the data and to make it more suitable for the specific data analysis task.

This may involve removing errors and inconsistencies,handling missing values,transforming the data into a consistent format and scaling the data to a suitable range. 

**VISUALIZATION STRATEGY: Plan how to visualize the insights using IBM Cognos to create informative dashboards and reports(Development part 2)**

**On-Time Performance:**

## Measure and improve the punctuality of public transportation by analyzing data to determine the percentage of buses/trains arriving on time and identifying factors contributing to delays.

* Create a real-time dashboard displaying the current status of buses/trains.

## Generate line charts showing historical on-time performance trends.

**Passenger Satisfaction:**

## Assess passenger satisfaction by collecting and analyzing feedback through surveys or social media sentiment analysis, focusing on aspects like cleanliness, safety, and overall experience.

* + Develop a sentiment analysis dashboard summarizing positive and negative feedback.
  + Create heatmaps or word maps to visualize common themes in passenger comments

### CODE:

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 inspecific 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}")

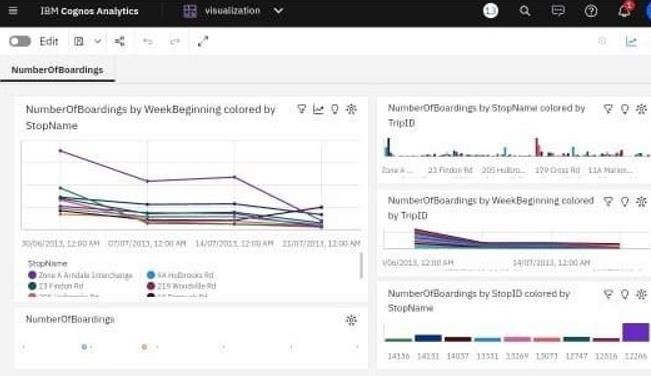
print(f"Neutral Feedback Count: {neutral\_feedback}")

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:



**Report:**

