PUBLIC TRANSPORTATION ANALYSIS

TEAM MEMBER

DHIVYA SOBIYA.M

HARINI.M

MADHUSRI.M

MERLIN.I

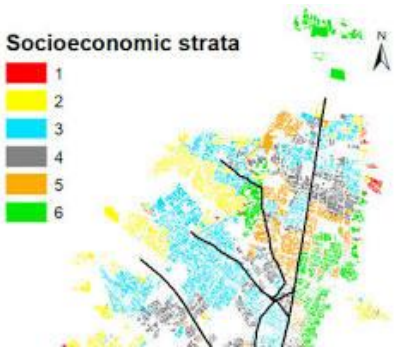
SAMYUKDHA.D

Phase 5 submission document

Project Title : Public Transportation Analysis

Phase 5 : Project Documentation & Submission

Topic: In this section we will document the complete project and prepare it for submission



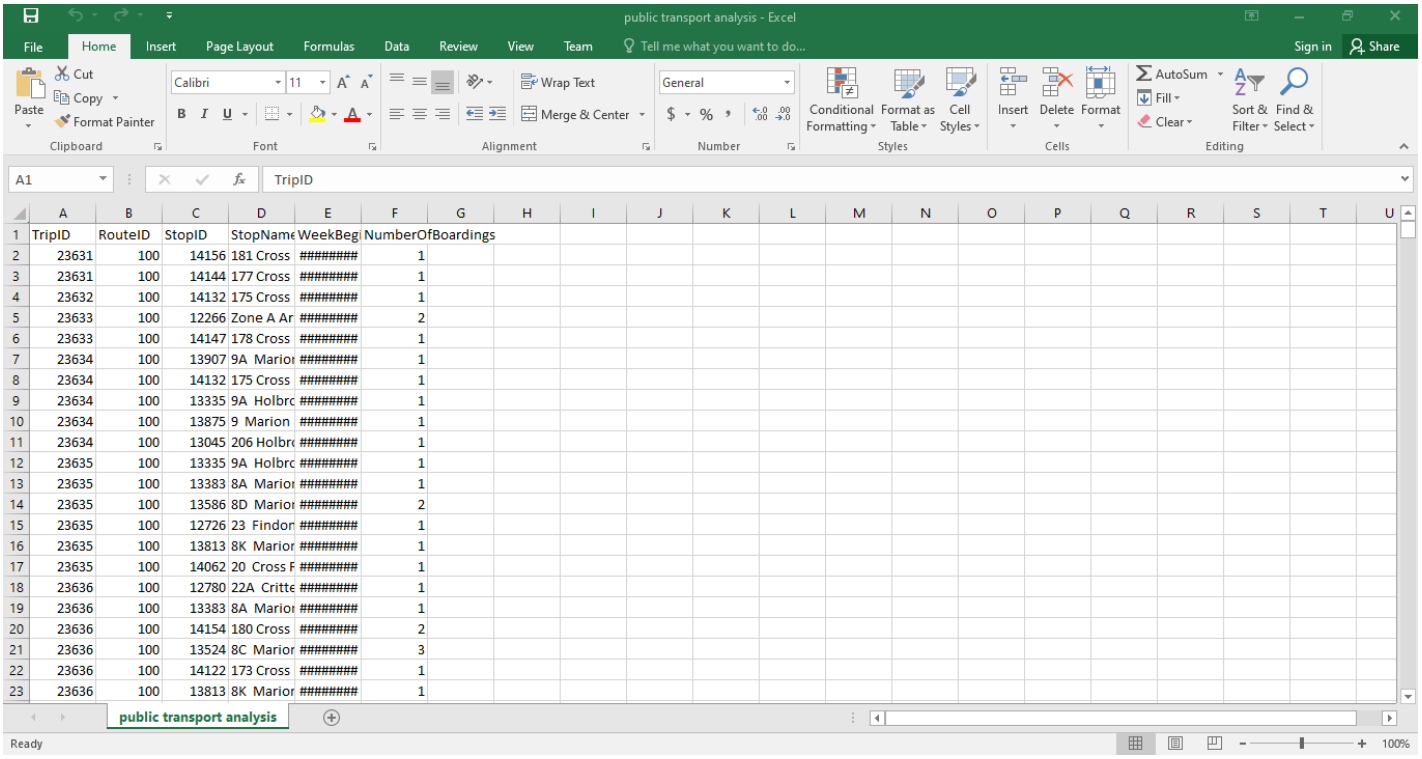
Public Transportation Analysis

Introduction :

- Public transportation analysis is the study of public transportation systems and their performance. It involves collecting, analyzing, and interpreting data on public transportation ridership, service levels, costs, and other factors.
- The goal of public transportation analysis is to identify ways to improve the efficiency, effectiveness, and equity of public transportation systems. It plays a vital role in modern urban environments, offering a cost-effective, eco-friendly, and efficient means of commuting for millions of people worldwide. The purpose of this analysis is to delve into the various aspects of public transportation, evaluating its performance, accessibility, and sustainability.
- This study considers both the challenges and opportunities associated with public transit systems, with a specific focus on improving the overall quality and accessibility of these services. This focuses on examining the efficiency, accessibility, and sustainability of public transportation systems in urban areas. The study employs a multidisciplinary approach, considering factors such as ridership, infrastructure, environmental impact, and technological advancements.
- Through the analysis of real-world case studies and data-driven insights, the research aims to provide a comprehensive overview of the current state of public transportation, highlighting its strengths and identifying areas for improvement.
- The findings of this analysis are crucial for urban planners, policymakers, and stakeholders to make informed decisions for enhancing public transport systems and promoting sustainable urban development. This public transport analysis focuses on examining the efficiency, accessibility, and sustainability of public transportation systems in urban areas.
- The study employs a multidisciplinary approach, considering factors such as ridership, infrastructure, environmental impact, and technological advancements. Through the analysis of real-world case studies and data-driven insights, the research aims to provide a comprehensive overview of the current state of public transportation, highlighting its strengths and identifying areas for improvement.

Dataset Link:(https://www.kaggle.com/datasets/rednivrug/unisys?select=20140711.CSV)

Given Dataset :



public transport analysis - Excel

TripID	RouteID	StopID	StopName	WeekBegi	NumberOfBoardings
23631	100	14156	181 Cross	########	1
23631	100	14144	177 Cross	########	1
23632	100	14132	175 Cross	########	1
23633	100	12266	Zone A Ar	########	2
23633	100	14147	178 Cross	########	1
23634	100	13907	9A Mario	########	1
23634	100	14132	175 Cross	########	1
23634	100	13335	9A Holbrc	########	1
23634	100	13875	9 Marion	########	1
23634	100	13045	206 Holbr	########	1
23635	100	13335	9A Holbrc	########	1
23635	100	13383	8A Mario	########	1
23635	100	13586	8D Mario	########	2
23635	100	12726	23 Findor	########	1
23635	100	13813	8K Mario	########	1
23635	100	14062	20 Cross F	########	1
23636	100	12780	22A Critte	########	1
23636	100	13383	8A Mario	########	1
23636	100	14154	180 Cross	########	2
23636	100	13524	8C Mario	########	3
23636	100	14122	173 Cross	########	1
23636	100	13813	8K Mario	########	1

Objectives :

- There must be an allocation of tasks between public and private transport;
- public transit services must be adapted to changing structures and needs
- attractiveness of public transit must be improved to draw motorists from their automobile need
- economic position of public transit must be guaranteed
- collaboration between transit enterprises must be increased in the passengers' interest
- cooperation between all transit operators must be achieved to optimize solutions to mutual problems.

Here's a list of tools and software commonly used in the process:

1. **Data Collection:**

   - ✓ Data are collected at a few locations taken to represent transport activity, travel movement and traffic flow across the study area or a sample of individual travellers.
   - ✓ The dataset may include information on various transport.

2. **Data Storage:**

   - ✞ **On-premises databases**: On-premises databases are installed and maintained on local servers. On-premises databases offer good performance and security, but they can be expensive to set up and maintain.
   - ✞ **Cloud databases**: Cloud databases are hosted by a third-party provider. Cloud databases are often more affordable and easier to maintain than on-premises databases, but they may have lower performance and security.
   - ✞ **Data lakes**: Data lakes are central repositories for storing all types of data, including raw data, processed data, and structured and unstructured data. Data lakes are often used for big data analytics, but they can also be used to store public transportation data.

3. **Data Visualization Tools:**

   - ✞ Tools like Esri ArcGIS, QGIS, Tableau, Power BI and D3.js are essential for data exploration and visualization.

4. **Data Preprocessing Tools:**

   - ✞ Libraries like pandas help with data cleaning, data transformation, data integration and manipulation.

5. Public Transportation Analysis Tools:

- Open Trip Planner (OTP), Mapzen, Google Transit Planner, Rome2rio, World Bank Transport for Communities (TTC) and Urban Mobility Analytics Platform (UMAP)

6. Additional Tools and Software:

- Programming languages such as Python and R
- Statistical software such as SPSS and Stata
- Machine learning software such as Tensor Flow and Py Torch.

7. External data sources:

External data sources can be used to complement and enhance public transportation analysis. These data sources can provide insights into a variety of factors that influence public transportation ridership, such as:

- **Land use and transportation**: This data can be used to understand how the built environment affects the demand for public transportation. For example, data on land use zoning, street networks, and pedestrian and bicycle infrastructure can be used to identify areas where public transportation service is likely to be most beneficial.
- **Traffic conditions**: This data can be used to understand how traffic congestion affects public transportation performance. For example, data on traffic speeds and volumes can be used to identify corridors where public transportation delays are likely to be high.
- **Weather conditions**: This data can be used to understand how weather conditions affect public transportation ridership and performance. For example, data on temperature, precipitation, and wind speed can be used to identify days when there is likely to be a decrease in ridership or an increase in delays.
- **Economic conditions**: This data can be used to understand how economic conditions affect public transportation ridership. For example, data on unemployment rates and income levels can be used to identify periods when ridership is likely to be higher or lower.

DESGIN THINKING AND PRESENT IN FORM OF DOCUMENT

Design Thinking Approach

Applying design thinking principles to the analysis and improvement of public Transportation can lead to innovative solutions that better meet the needs of Users and communities:

1) **Empathize:**

Conduct in-depth research to understand the needs, behaviors, and pain Points of public transportation users. This may involve surveys, interviews, Observations, and journey mapping.

- ❖ Engage with various stakeholders, including commuters, transit operators , Urban planners, and government officials, to gather diverse perspectives.

2) **Define:**
   - ❖ Clearly define the problem based on the insights gathered during the empathize stage. For example, the problem could be inadequate accessibility for people with disabilities or long wait times for buses.
   - ❖ Formulate a problem statement that focuses on the user experience, such as “How might we improve the daily commute for residents of this city.

3) **Ideate:**
   - ❖ Encourage creative thinking sessions with multidisciplinary teams to generate a wide range of potential solutions.
   - ❖ Use techniques like brainstorming, mind mapping, or “Crazy 8s” to explore

   innovative ideas without constraints.

4) **Prototype:**
   - ❖ Create low-fidelity prototypes of the most promising ideas. Prototypes can range from physical models of bus stops to digital simulations of transit apps.
   - ❖ Continuously refine and iterate on prototypes based on user feedback and feasibility assessments.

5) **Test:**
   - ❖ Testing can help identify strengths and weaknesses of proposed solutions.

- Make necessary adjustments to prototypes, which may involve several rounds of testing and refinement.

6) Implement:

- The most successful solutions and monitor their impact.
- This could involve making changes to existing policies and procedures, developing new technologies, or even investing in new infrastructure.

7) Evaluate:

- Design thinking is a valuable tool for improving public transportation analysis.
- It can help public transportation agencies to better understand the needs of users, develop more effective solutions to public transportation problems, and be more creative and innovative in their approach to public transportation planning and operations.

8) Iterate:

- Based on the feedback from users and the data collected from the testing phase, refine and improve the prototypes.
- The prototypes can then be tested again on a larger scale.

9) Educate and Train:

- Education and training in public transportation analysis is essential for developing the skills and knowledge necessary to improve public transportation systems.
- Public transportation agencies should invest in educating and training their employees in public transportation analysis, and individuals who are interested in a career in public transportation should consider pursuing an education in transportation engineering, planning, or urban planning.

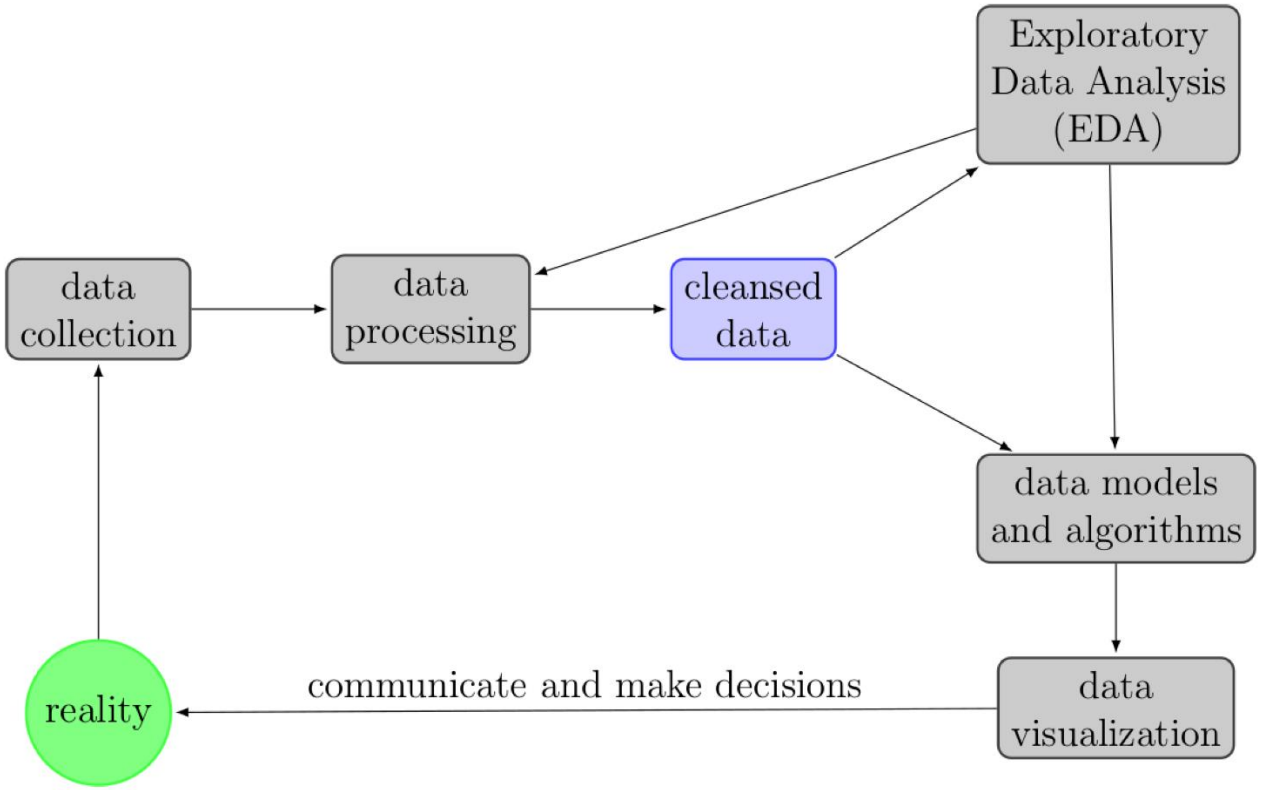
DESIGN INTO INNOVATION

A. Data Collection:

- ✓ In this methods can help to address these challenges. By using new technologies and approaches, public transportation agencies can collect more data, more efficiently, and at a lower cost.
- ✓ This data can then be used to improve the accuracy and completeness of public transportation analysis, leading to better decision-making and improved service for riders.

B. Data Preprocessing:

- ✓ Data preprocessing is an essential step in public transportation analysis, as it helps to ensure that the data is clean, consistent, and ready for analysis.
- ✓ Traditional data preprocessing techniques can be time-consuming and labor-intensive, which can limit the scope and frequency of public transportation analysis.



PYTHON PROGRAM:

```
import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

data = pd.read_csv(r"C:\Users\IT\Desktop\Samyukdha\public transport analysis.csv")

data.head()

correlation=data.corr()

correlation["TripID"] .sort_values

import numpy as np # linear algebra

import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

import matplotlib.pyplot as plt

import datetime

import os

from math import sqrt

import warnings

from IPython.core.interactiveshell import InteractiveShell

InteractiveShell.ast_node_interactivity = "all"

warnings.filterwarnings('ignore')

data = pd.read_csv(r"C:\Users\IT\Desktop\Samyukdha\public transport analysis.csv")

data.shape

data.head(10)

out_geo = pd.read_csv(r"C:\Users\IT\Desktop\Samyukdha\public transport analysis.csv")
```

```
out_geo.shape

out_geo.head()

fig,axrr=plt.subplots(2,2,figsize=(15,15))

ax=axrr[0][0]

ax.set_title("No of Boardings")

data['NumberOfBoardings'].value_counts().sort_index().head(20).plot.bar(ax=axrr[
0][0])

ax=axrr[0][1]

ax.set_title("WeekBeginning")

data['WeekBeginning'].value_counts().plot.area(ax=axrr[0][1])

ax=axrr[1][0]

ax.set_title("most Busiest Route")

data['RouteID'].value_counts().head(10).plot.bar(ax=axrr[1][0])

ax=axrr[1][1]

ax.set_title("least Busiest Route")

data['RouteID'].value_counts().tail(10).plot.bar(ax=axrr[1][1])
```

OUTPUT:

	TripID	RouteID	StopID	StopName	WeekBeginning	No.Of Boardings
0	23631	100	14156	181 Cross Rd	6/30/2013 0:00	1
1	23631	100	14144	177 Cross Rd	6/30/2013 0:00	1
2	23632	100	14132	175 Cross Rd	6/30/2013 0:00	1
3	23633	100	12266	Zone A Arndale Interchange	6/30/2013 0:00	2
4	23633	100	14147	178 Cross Rd	6/30/2013 0:00	1

<bound method Series.sort_values of TripID 1.000000

StopID 0.017946

NumberOfBoardings 0.005864

Name: TripID, dtype: float64>

(1048575, 6)

	TripID	RouteID	StopID	StopName	WeekBeginning	No.Of Boardings
0	23631	100	14156	181 Cross Rd	6/30/2013 0:00	1
1	23631	100	14144	177 Cross Rd	6/30/2013 0:00	1
2	23632	100	14132	175 Cross Rd	6/30/2013 0:00	1
3	23633	100	12266	Zone A Arndale Interchange	6/30/2013 0:00	2
4	23633	100	14147	178 Cross Rd	6/30/2013 0:00	1
5	23634	100	13907	9A Marion Rd	6/30/2013 0:00	1
6	23634	100	14132	175 Cross Rd	6/30/2013 0:00	1
7	23634	100	13335	9A Holbrooks Rd	6/30/2013 0:00	1
8	23634	100	13875	9 Marion Rd	6/30/2013 0:00	1
9	23634	100	13045	206 Holbrooks Rd	6/30/2013 0:00	1

(1048575, 6)

	TripID	RouteID	StopID	StopName	WeekBeginning	No.Of Boardings
0	23631	100	14156	181 Cross Rd	6/30/2013 0:00	1
1	23631	100	14144	177 Cross Rd	6/30/2013 0:00	1
2	23632	100	14132	175 Cross Rd	6/30/2013 0:00	1
3	23633	100	12266	Zone A Arndale Interchange	6/30/2013 0:00	2
4	23633	100	14147	178 Cross Rd	6/30/2013 0:00	1

Text(0.5, 1.0, 'No of Boardings')

<AxesSubplot:title={'center':'No of Boardings'}>

Text(0.5, 1.0, 'WeekBeginning')

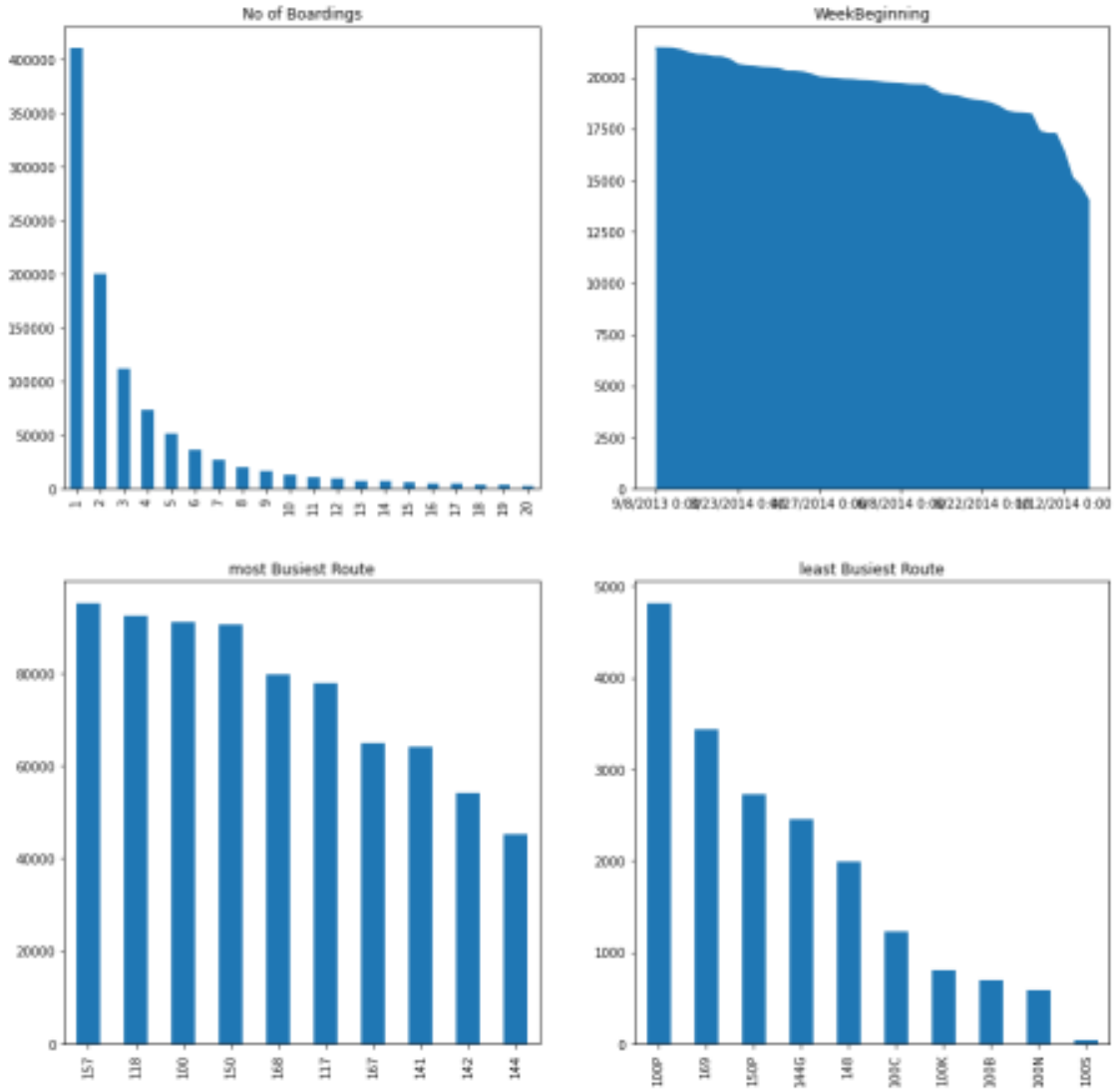
<AxesSubplot:title={'center':'WeekBeginning'}>

Text(0.5, 1.0, 'most Busiest Route')

<AxesSubplot:title={'center':'most Busiest Route'}>

Text(0.5, 1.0, 'least Busiest Route')

<AxesSubplot:title={'center':'least Busiest Route'}>



PERFORMING DIFFERENT ACTIVITIES LIKE FEATURE ENGINEERING, MODEL TRAINING, EVALUATION etc.,

I. Feature Engineering:

- The process of transforming and combining raw data into features that are more informative and useful for machine learning and statistical models.
- Feature engineering is an important step in public transportation analysis, as it can help to improve the accuracy and performance of models that are used to predict ridership, travel times, and other important metrics.

II. Data Preprocessing and Visualization

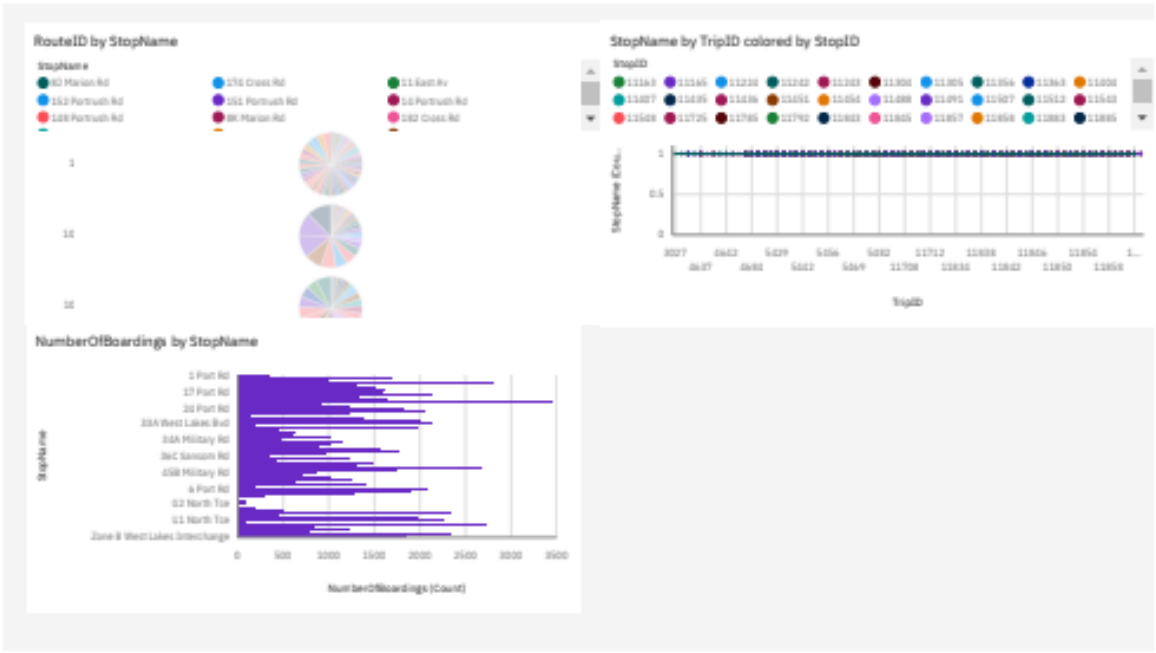
Data preprocessing: The process of cleaning, transforming, and organizing data so that it is ready for analysis. This may involve tasks such as:

- Removing errors and inconsistencies from the data
- Filling in missing values
- Converting the data to a consistent format
- Aggregating the data to a suitable level of granularity

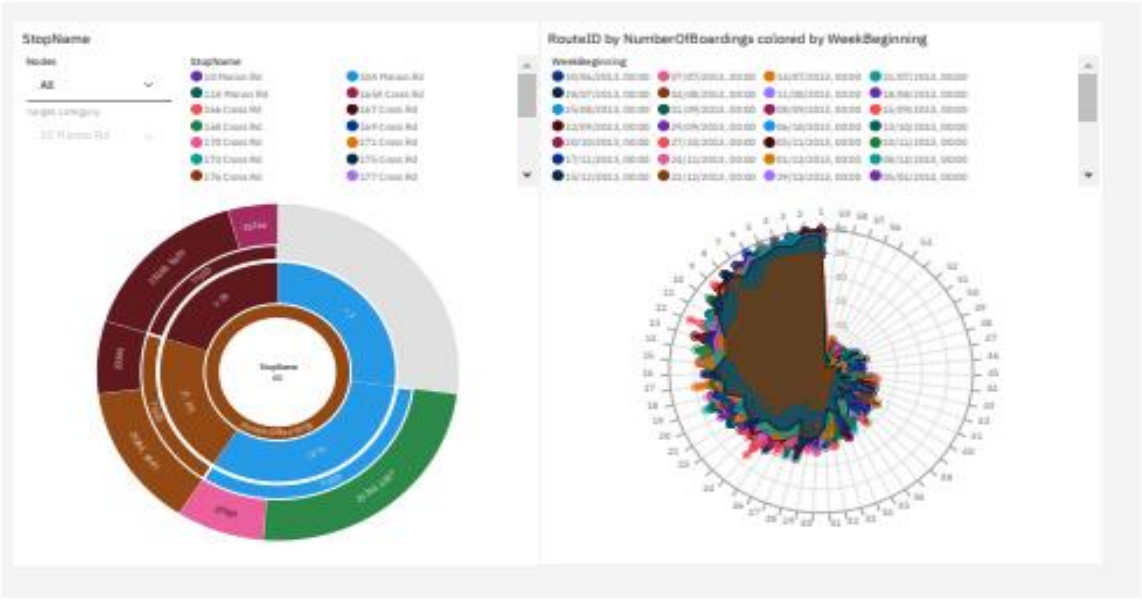
Data visualization: The process of creating graphical representations of data to help people understand and interpret it. This may involve tasks such as:

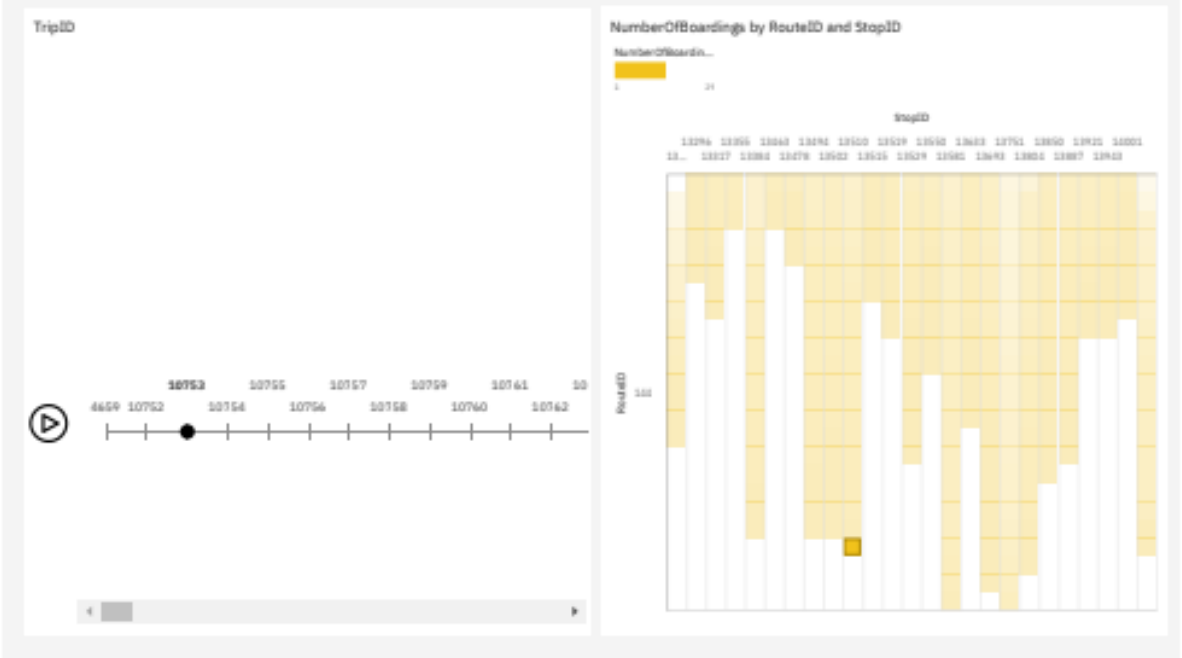
- Creating charts and graphs
- Mapping data to geographic locations
- Developing interactive visualizations

Tab 1



10/25/23, 8:55 PM New dashboard


StopName
RouteID by NumberOfBoardings colored by WeekBeginning


TripID
NumberOfBoardings by RouteID and StopID
StopID
RouteID

ADVANTAGES:

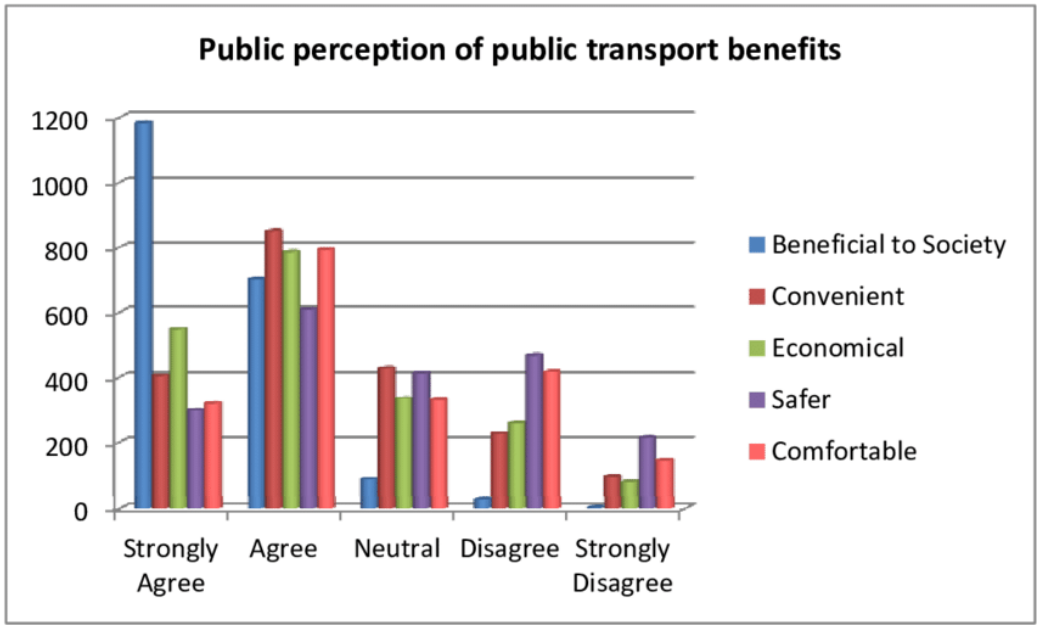
- Public Transport leads to less air pollution as more people commute via single vehicle eliminating the need to travel by different modes.
- Many people who cannot afford to buy their vehicles and cars generally travel by public transport. It is a very cost-effective mode of commuting for them.
- Public transport is a very fuel-efficient mode of transport per capita or person; fuel consumption is less.
- Talking about its economics, the government spends a lot to develop public transport infrastructure in the country. The revenue generated from public transport is a lot which helps in its maintenance. This is because it is the most used transport and very cost-effective.
- It is safer as they have high safety standards and protocols built-up.

DISADVANTAGES:

- Public delivery may be slower than visiting the usage of your non-public automobile because of strongly prescribed pace limits.
- The constant schedules that public delivery follows might not mesh exactly with the tour needs of a few users.
- Public transportation might not prevent an appropriate cope when a passenger visits, compelling them to search for extra means.
- Privacy is compromised a terrific deal. Travelers are crowded, and private areas might not be available.
- The breakdown of public delivery cars can cause a massive lack of time for many human beings.

BENIFITS:

- you can enjoy a less stressful journey by letting someone else do the driving
- you don’t have to worry about finding a parking space
- it reduces congestion in towns and cities using public transport is cheaper than owning and operating a car
- no more sitting in traffic jams in rush hour thanks to bus lanes and other bus priority measures
- it reduces your carbon footprint



CONCLUSION:

- Public transportation analysis is essential for improving the efficiency, effectiveness, and accessibility of public transportation systems. By analyzing public transportation data, researchers can gain valuable insights into ridership patterns, performance trends, and the impact of service changes.

- This information can then be used to develop and implement evidence-based policies and programs to improve public transportation for all users.In recent years, there has been a growing interest in using design thinking to innovate data preprocessing and visualization techniques for public transportation analysis.

- Design thinking is a human-centered approach to problem-solving that can help to ensure that the needs of users are central to the development of new solutions. By using design thinking to develop new data preprocessing and visualization tools and interfaces, researchers can make it easier and more efficient to analyze public transportation data.

- This can lead to new insights into public transportation ridership, performance, and trends. This information can then be used to improve public transportation planning and operations. Overall, public transportation analysis is a powerful tool for improving the quality of public transportation services.

- By using design thinking to innovate data preprocessing and visualization techniques, researchers can make public transportation analysis more efficient, effective, and accessible. This can lead to new insights and solutions that improve public transportation for all users.

LINK FOR PHASE SUBMISSION:

phase1: https://github.com/SamyukdhaDhamotharan/PublicTransportationAnalysis/blob/d91f11e1d5b8262e63b2b5f164221d755045e5ce/DAC_Phase1%20.pdf

phase2 :https://github.com/SamyukdhaDhamotharan/PublicTransportationAnalysis/blob/d91f11e1d5b8262e63b2b5f164221d755045e5ce/DAC_Phase2%20(1).pdf

phase 3:

https://github.com/SamyukdhaDhamotharan/PublicTransportationAnalysis/blob/d91f11e1d5b8262e63b2b5f164221d755045e5ce/DAC_Phase3.pdf

phase 4:

https://github.com/SamyukdhaDhamotharan/PublicTransportationAnalysis/blob/d91f11e1d5b8262e63b2b5f164221d755045e5ce/DAC_Phase4.pdf