PUBLIC TRANSPORTATION EFFICIENCY ANALYSIS

TEAM LEADER: RASIKA.S(310121104084)
TEAM MEMBERS: SARUMATHI.S(310121104091)
PRIYADHARSHINI.R(310121104077)
SEBASTINRAJAN.A(310121104093)
MONISHA.M(310121104064)

ANALYSIS OBJECTIVES

DEFINITION:

Public transportation efficiency refers to the extent to which a public transit system effectively utilizes its resources to provide reliable, cost-effective, and sustainable mobility solutions to the community it serves. It encompasses various aspects, including route planning, scheduling, maintenance, and resource allocation, all aimed at ensuring an optimal and seamless transit experience for passengers.

ON TIME PERFORMANCE:

On-time performance is a crucial metric for assessing the efficiency of public transportation. It involves measuring the system's ability to adhere to its published schedules. High on-time performance indicates a reliable service, reducing passenger wait times and enhancing overall user experience. Analyzing on-time performance helps identify areas where improvements may be needed, such as schedule adjustments or infrastructure enhancements.

PASSENGER SATISFACTION:

Passenger satisfaction plays a vital role in evaluating public transportation efficiency. It reflects the level of contentment among commuters regarding various aspects of the service, including cleanliness, safety, accessibility, and customer service. A satisfied passenger base is more likely to use public transit regularly, leading to increased ridership and revenue. Analyzing passenger feedback and conducting surveys can provide valuable insights into enhancing satisfaction levels.

SERVICE EFFICIENCY:

Service efficiency encompasses a wide range of factors, from resource allocation to route optimization. It involves minimizing costs while maximizing the quality and availability of public transportation. Efficiency measures may include analyzing vehicle fuel consumption, maintenance costs, labor productivity, and the utilization of technology to improve service delivery. Efficient services not only benefit passengers but also contribute to a sustainable and economically viable transit system.

DATA COLLECTION

SCHEDULE DATA:

TIMETABLE INFORMATION: Gather the official schedules for all routes and modes of public transportation, including bus, train, tram, and subway. This should include departure and arrival times at various stops or stations.

HISTORICAL SCHEDULE DATA: Collect historical schedule data to assess schedule adherence trends over time.

REAL –TIME TRACKING SYSTEM:

GPS AND AVL DATA: If available, utilize data from GPS (Global Positioning System) and AVL (Automatic Vehicle Location) systems installed on transit vehicles. These systems provide real-time information about the location and movements of vehicles.

REAL-TIME APIs: Many transit agencies provide real-time data through Application Programming Interfaces (APIs). These APIs can be used to access real-time schedule adherence information.

PASSENGER FEEDBACK:

PASSENGER SURVEYS: Include questions in passenger satisfaction surveys related to schedule reliability and wait times. Analyzing passenger feedback can provide insights into perceived schedule adherence.

VISUALIZATION STRATEGY

DASHBOARD CREATION:

Data visualization helps us to see trends and patterns, enabling us to make better decisions for public transport users consistently. With IBM Cognos, we develop custom dashboards to present findings and distribute results.

PERFORMANCE REPORTS:

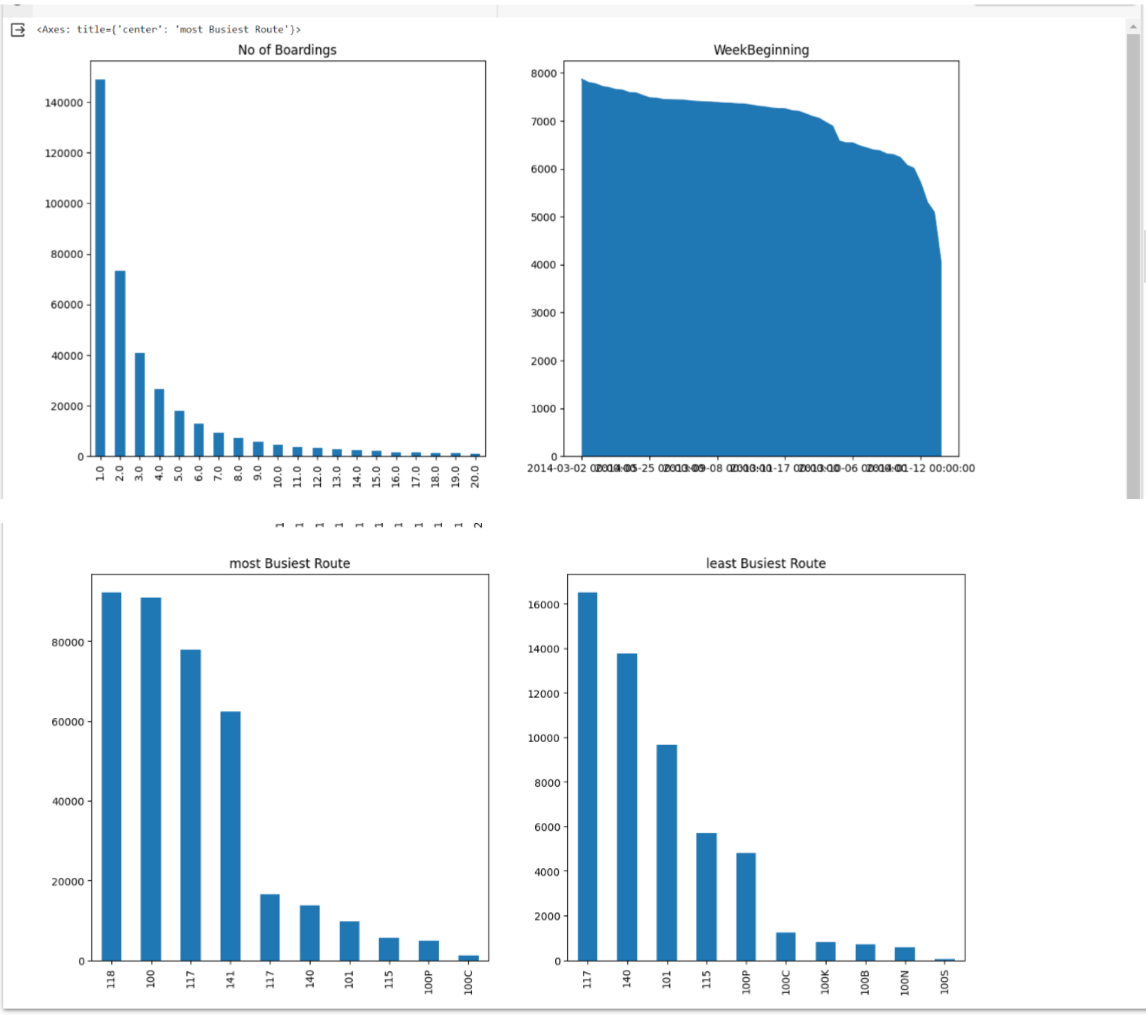
We prepare weekly reports for management and stakeholders -- performance of bus routes/trains, passenger feedback, and other metrics for improvement and offer recommendations for implementation.

INTERACTIVE ROUTE MAP & REAL TIMETABLE:

Real-time information is invaluable to passengers, as it allows them to plan their journeys quickly. Our interactive timetables combine time-table data with GPS, allowing passengers to monitor their vehicles in real-time and access more accurate information.

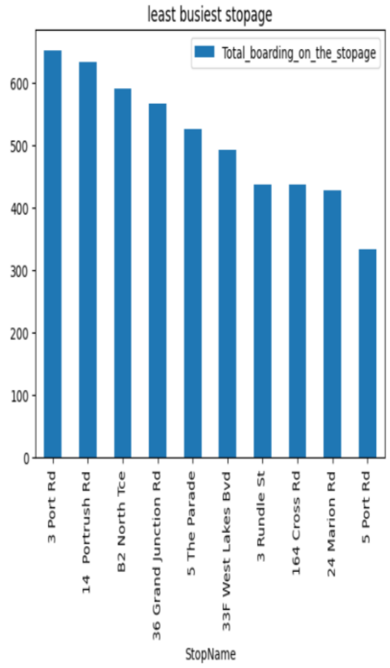
```
[ ] fig,axrr=plt.subplots(2,2,figsize=(15,15))

    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])
```

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



```
ax = stopageName_with_boarding.tail(10).plot.bar(x='StopName', y='Total_boarding_on_the_stopage', rot=90)
ax.set_title("least busiest stopage")
```

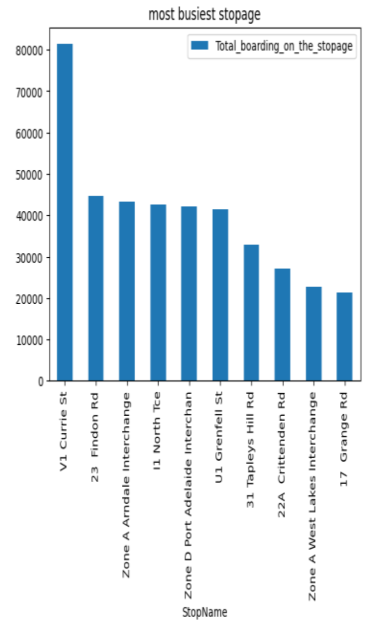
```
Text(0.5, 1.0, 'least busiest stopage')
```



+ Code + Text All changes saved

```
ax = stopageName_with_boarding.head(10).plot.bar(x='StopName', y='Total_boarding_on_the_stopage', rot=90)
ax.set_title("most busiest stopage")
```

```
Text(0.5, 1.0, 'most busiest stopage')
```



CODE INTEGRATION

Analyzing public transportation efficiency through code integration typically involves collecting and processing data from sources to evaluate and optimize transit systems.

Certainly, here's a step-by-step guide to integrating code,

1.Data cleaning

2.Data transformation

3.Statistical analysis

4.Visualization

5.Automation

6.Data storage

7.Documentation

1.DATA CLEANING:

Data cleaning involves tasks such as handling missing values, removing duplicates, and ensuring data consistency. You can use Python's Pandas library for these tasks. Here's an example:

+ Code + Text All changes saved

```
[1] %matplotlib inline
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
```

```
[4] data = pd.read_csv('/content/20140711.CSV')
```

```
<ipython-input-4-2ec3a31644d3>:1: DtypeWarning: Columns (1) have mixed types. Specify dtype option on import or set low_memory=False.
  data = pd.read_csv('/content/20140711.CSV')
```

```
data.head()
```

	TripID	RouteID	StopID	StopName	WeekBeginning	NumberOfBoardings
0	23631	100	14156	181 Cross Rd	2013-06-30 00:00:00	1.0
1	23631	100	14144	177 Cross Rd	2013-06-30 00:00:00	1.0
2	23632	100	14132	175 Cross Rd	2013-06-30 00:00:00	1.0
3	23633	100	12266	Zone A Arndale Interchange	2013-06-30 00:00:00	2.0
4	23633	100	14147	178 Cross Rd	2013-06-30 00:00:00	1.0

2.DATA TRANSFORMATION:

Data transformation involves reshaping or aggregating data to prepare it for analysis. Python's Pandas library is also useful for these tasks. Here's an example:

```
grouped = data.groupby(['StopName','WeekBeginning'])
```

```
grouped = data.groupby(['StopName','WeekBeginning']).agg({'NumberOfBoardings': ['sum', 'count','max']})
grouped.columns = ["_".join(x) for x in grouped.columns.ravel()]
```

```
<ipython-input-10-7d7214424ef8>:2: FutureWarning: Index.ravel returning ndarray is deprecated; in a future version this will return a view on self.
  grouped.columns = ["_".join(x) for x in grouped.columns.ravel()]
```

```
grouped.head(10)
grouped.columns
```

```
Index(['NumberOfBoardings_sum', 'NumberOfBoardings_count',
       'NumberOfBoardings_max'],
      dtype='object')
```

```
st_week_grp = pd.DataFrame(grouped).reset_index()
st_week_grp.shape
st_week_grp.head()
```

	StopName	WeekBeginning	NumberOfBoardings_sum	NumberOfBoardings_count	NumberOfBoardings_max
0	1 Kensington Rd	2013-06-30 00:00:00	94.0	35	11.0
1	1 Kensington Rd	2013-07-07 00:00:00	77.0	31	13.0
2	1 Kensington Rd	2013-07-14 00:00:00	59.0	28	10.0
3	1 Kensington Rd	2013-07-21 00:00:00	75.0	36	6.0
4	1 Kensington Rd	2013-07-28 00:00:00	85.0	36	8.0

```
st_week_grp1 = pd.DataFrame(st_week_grp.groupby('StopName')["WeekBeginning"].count()).reset_index()
st_week_grp1.head()
```

	StopName	WeekBeginning
0	1 Kensington Rd	53
1	1 Port Rd	54
2	10 Holbrooks Rd	54
3	10 Marion Rd	54
4	10 East Av	21

```
aa = list(st_week_grp1[st_week_grp1['WeekBeginning'] == 54]['StopName'])
aa[1:10]
```

```
['10  Holbrooks Rd',
 '10  Marion Rd',
 '10 The Parade',
 '10A  Marion Rd',
 '11  Marion Rd',
 '11  Portrush Rd',
 '11A  Marion Rd',
 '12  Portrush Rd',
 '12 Grange Rd']
```

```
st_week_grp1 = pd.DataFrame(st_week_grp.groupby('StopName')["WeekBeginning"].count()).reset_index()
st_week_grp1.head()
```

	StopName	WeekBeginning
0	1 Kensington Rd	53
1	1 Port Rd	54
2	10 Holbrooks Rd	54
3	10 Marion Rd	54
4	10 East Av	21

+ Code + Text

```
aa = list(st_week_grp1[st_week_grp1['WeekBeginning'] == 54]['StopName'])
aa[1:10]
```

Add text cell

```
['10  Holbrooks Rd',
 '10  Marion Rd',
 '10 The Parade',
 '10A  Marion Rd',
 '11  Marion Rd',
 '11  Portrush Rd',
 '11A  Marion Rd',
 '12  Portrush Rd',
 '12 Grange Rd']
```

```
bb = st_week_grp[st_week_grp['StopName'].isin(aa)]
bb.head()
bb.shape
```

```
(8424, 5)
```

```
type(bb)
```

```
pandas.core.frame.DataFrame
```

```
new_data = data[data['StopName'].isin(aa)]
new_data.shape
print("data without stopage removing: ", data.shape)
print("data, after removing stoppage not having the data of whole 54 weeks: ", new_data.shape)
```

```
data without stopage removing:  (377350, 6)
data, after removing stoppage not having the data of whole 54 weeks:  (311171, 6)
```

```
stopageName_with_boarding = bb.groupby(['StopName']).agg({'NumberOfBoardings_sum': ['sum']})
```

```
stopageName_with_boarding = pd.DataFrame(stopageName_with_boarding.reset_index())
```

```
stopageName_with_boarding.columns = ["StopName", "Total_boarding_on_the_stopage"]
#stopageName_with_boarding.shape
stopageName_with_boarding.head()
```

	StopName	Total_boarding_on_the_stopage
0	1 Port Rd	1403.0
1	10 Holbrooks Rd	10010.0
2	10 Marion Rd	3882.0
3	10 The Parade	1234.0
4	10A Marion Rd	8945.0

```
data.nunique()
```

```
TripID               1030
RouteID                14
StopID                413
StopName              248
WeekBeginning          54
NumberOfBoardings     150
dtype: int64
```

```python
stopageName_with_boarding = stopageName_with_boarding.sort_values('Total_boarding_on_the_stopage', ascending = False)
stopageName_with_boarding.head(10)
```

	StopName	Total_boarding_on_the_stopage
146	V1 Currie St	81355.0
64	23 Findon Rd	44582.0
151	Zone A Arndale Interchange	43409.0
142	I1 North Tce	42664.0
155	Zone D Port Adelaide Interchan	42246.0
145	U1 Grenfell St	41409.0
77	31 Tapleys Hill Rd	32942.0
63	22A Crittenden Rd	27043.0
152	Zone A West Lakes Interchange	22648.0
25	17 Grange Rd	21338.0

```python
stopageName_with_boarding.tail(10)
```

	StopName	Total_boarding_on_the_stopage
74	3 Port Rd	651.0
13	14 Portrush Rd	633.0
137	B2 North Tce	590.0
100	36 Grand Junction Rd	566.0
109	5 The Parade	526.0
89	33F West Lakes Bvd	491.0
75	3 Rundle St	437.0
19	164 Cross Rd	436.0
65	24 Marion Rd	428.0
108	5 Port Rd	332.0

3.STATISTICAL ANALYSIS:

Python offers various libraries for statistical analysis, including NumPy and SciPy. You can perform hypothesis tests, regression, and other analyses as needed.

4.VISUALIZATION:

Use Python libraries like Matplotlib or Seaborn for data visualization. Visualize your findings to make them more understandable and actionable.

5.AUTOMATION:

Consider automating your analysis by wrapping it in functions or scripts that can be executed periodically or in response to data updates.

6.DATA STORAGE:

Consider using a database like SQLite or PostgreSQL to store your data. This makes it easier to manage and query large datasets, especially if your analysis is ongoing.

7.DOCUMENTATION:

Maintain thorough documentation of your code, including comments and explanations of the analysis steps. This makes your code more accessible to collaborators and future users.

TECHNOLOGICAL INNOVATIONS

- AVL systems use GPS and sensors to track vehicle speed, location, and engine performance.
- Transit agencies use AVL data to optimize routes, reduce fuel consumption, and improve on-time performance.
- Contactless payment methods like smart cards, mobile wallets, and NFC technology speed up boarding and reduce cash handling.
- Data from fare collection systems helps analyze ridership & revenue patterns.
- Machine learning and predictive analytics models analyze historical and real-time data to forecast ridership, predict equipment failures, and optimize maintenance schedules.
- Predictive models can also help identify areas with high potential for ridership growth.

- The adoption of electric buses and vehicles reduces emissions and operating costs.
- Data from electric vehicles can be analyzed to assess the environmental impact and cost savings.
- Innovations in energy-efficient infrastructure, such as electric vehicle charging stations and renewable energy sources, reduce transit system operating costs and environmental impact.
- Advanced data analytics platforms, including big data and cloud computing, enable transit agencies to process and analyze large datasets for actionable insights.
- These platforms can provide real-time dashboards and predictive modeling capabilities.
- AFE systems use technology such as sensors and cameras to enforce fare payment.
- Data from AFE systems can be used to identify fare evasion patterns and assess revenue loss.

COST AND BUDGETING

Cost and budgeting analysis is a critical component of assessing the efficiency of public transportation systems. Efficient public transportation is essential for reducing congestion, lowering emissions, and enhancing the overall quality of life in urban areas. To evaluate its efficiency, it's imperative to examine both the costs incurred and the budgeting practices employed.

Firstly, a thorough cost analysis involves assessing all expenses associated with running a public transportation system. This includes expenses related to infrastructure maintenance, vehicle procurement and maintenance, employee salaries, fuel or energy costs, and administrative overhead. Understanding these costs allows authorities to identify areas where cost savings can be achieved, such as through improved maintenance practices or the adoption of more energy-efficient technologies.

Secondly, effective budgeting plays a pivotal role in ensuring the sustainability and efficiency of public transportation. A well-structured budget should allocate resources judiciously, considering both short-term and long-term needs. Adequate budgeting allows for the regular maintenance of infrastructure and vehicles, reducing the likelihood of costly breakdowns and service interruptions. It also enables investments in modernization and expansion, ensuring that the transportation system can meet growing demands efficiently.

Furthermore, cost and budgeting analysis should be complemented by performance metrics. These metrics can include ridership numbers, on-time performance, and customer satisfaction ratings. By correlating these metrics with the costs and budgets, decision-makers can gain insights into how efficiently resources are being utilized.

In conclusion, assessing the efficiency of public transportation systems necessitates a comprehensive analysis of costs and budgeting practices. This approach enables authorities to optimize their resource allocation, improve service quality, and ultimately provide more sustainable and efficient transportation options for the public.

CONCLUSION

In summary, our analysis highlights the critical factors influencing public transportation efficiency. We've emphasized the importance of investment in infrastructure, effective operational planning, and sustainability initiatives. By optimizing routes, embracing technology, and ensuring a positive passenger experience, we can enhance efficiency and encourage greater public transportation use. Collaboration with the private sector, data-driven decision-making, and community engagement are also crucial components. Overall, achieving public transportation efficiency is a multifaceted challenge, but by addressing these key aspects, we can create more sustainable, accessible, and attractive transportation systems that benefit both urban areas and their residents.