Traffic and Capacity Analytics for Major Ports

Team ID : PNT2022TMID37693

Team Size:5

Team Leader : Uvagai.K.E

Team members :Anuja.G

Logeshwari.R,Monisha.B

Pavithra.V

T

Project Report Format

1. **INTRODUCTION**
    a. Project Objectives
    b. Project Flow
2. **LITERATURE SURVEY**
    a. References
    b. Problem Statement Definition
3. **IDEATION & PROPOSED SOLUTION**
    a. Empathy Map Canvas
    b. Ideation & Brainstorming
    c. Proposed Solution
    d. Problem Solution fit
4. **REQUIREMENT ANALYSIS**
    a. Functional requirement
    b. Non-Functional requirements
5. **PROJECT DESIGN**
    a. Data Flow Diagrams
    b. Solution & Technical Architecture
    c. User Stories

6. PROJECT PLANNING & SCHEDULING

a. Prepare Milestone and Activity List

b. Sprint Delivery Plan

7. TESTING

8. ADVANTAGES & DISADVANTAGES

9. CONCLUSION

10. APPENDIX

GitHub & Project Demo Link

INTRODUCTION

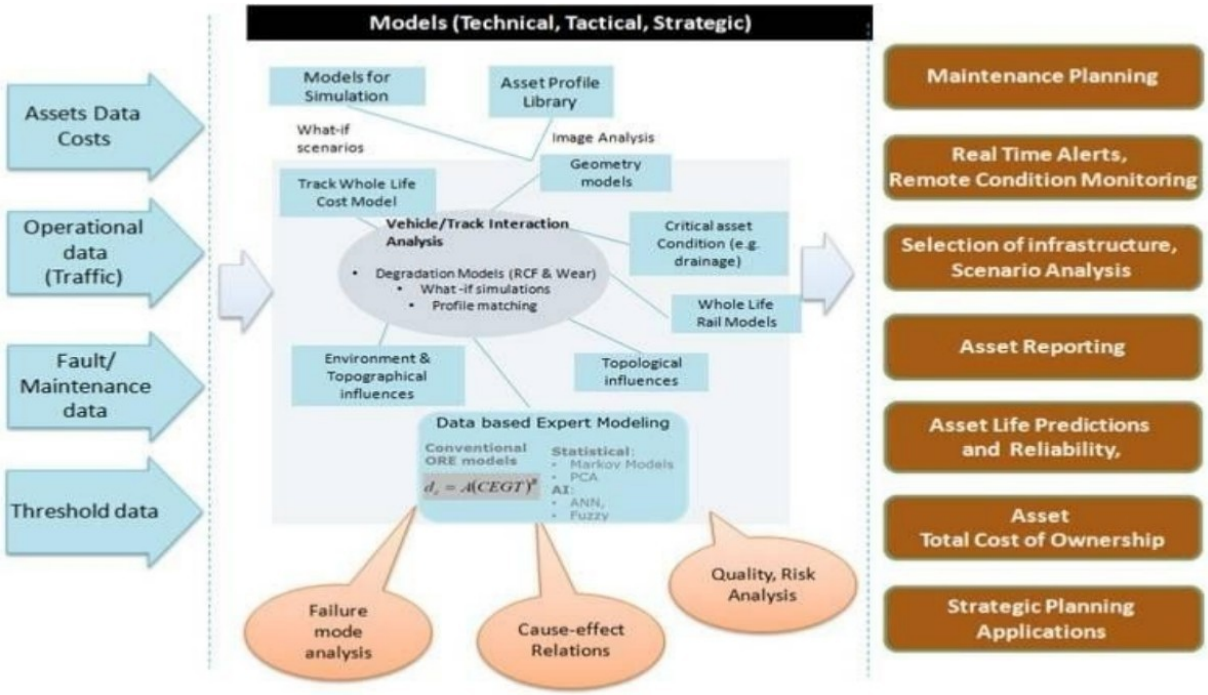
The Indian Railways has a capital base of about Rs. 100000 crores and is often referred to as the lifeline of the Indian economy because of its predominance in transportation of bulk freight and long distance passenger traffic. The network crisis-crosses the nation, binding it together by ferrying freight and passengers across the length and breadth of the country. As the Indian economy moves into a high growth trajectory the Railways have also stepped-up developmental efforts and are preparing themselves for an even bigger role in the future.

a. OBJECTIVES:

- Ports serves as an important link in global supply chain. The Indian Railways has a capital base of about Rs.100000 crores and is often referred to as the lifeline of the Indian economy because of its predominance in transportation of bulk freight and long distance passenger traffic. Data analytics can be used for analyzing the port performance.
- In this project, the port capacity topic was addressed through Cog nos analysis. Reducing the congestion on rail corridors and improving port-connectivity.

- Railways have also stepped-up developmental efforts and are preparing themselves for an even bigger role in the future. So, data analytics plays the major role in this project.

b. PROJECT FLOW:



LITERATURE SURVEY

a.References

Paper 1:

A systematic Analysis of Port Capacity Literature: Trends and Future Research Avenues Publication year:31 January ,2021 Author name: Cecil-Miguel Journal name: Journal of maritime transport & logistics Summary: The continuous growth in the world economy, technology, and the population still shapes the industrialization patterns. This massive progress has also shaped the international transportation requirements. Ports, as the one of the important infrastructure in international transportation and supply chains, have been pushed by these changes in terms of structuring their capacities to satisfy the demand. To do this, this study adopted a systematic literature review and content analysis together. The result of this study showed that the most attractive topics are service level and performance in main category.

Paper-2:

Performance analysis of major ports in India: A quantitative approach Publication year:January ,2016 Author name: Anindita-Man dal Journal name: International Journal of Business Performance Management Summary: The paper examines the performance of 13 major ports of India in respect of key operational performance indicators. Following rapid economic growth India's share in international trade is escalating. This puts increased pressure on these ports, which handle many of the trade to perform with optimal efficiency. The study presents a systematic analysis of different performance indicators for a 10 yr time period (2003 to 2013) using a variety of statistical methods and evaluates status of each port in different categories of performance.

Paper 3:

Analytics for Decision Making at Ports Publication year: October ,2015 Author name: Mrinal Markup Dupattas Journal name: publishing India Summary:

Ports serve as an important link in global supply chain. The Indian Union has endeavored to invest on major ports of the country to meet up to the global standards. The major ports lost its share to the minor ports under the state governments. This paper an attempt has been made to identify the dimensions of port performance and the causality between the dimensions. It chooses to take average turn round time (ATRT) as an indicator of port performance. The paper proposes an analytical framework to identify the causality that would aid the decision makers

Paper-4:

Towards Analytics-Enabled Efficiency Improvements in Maritime Transportation: A Case Study in a Mediterranean Port Publication year: 21 June ,2019 Author name: Pierluigi Zerbino Journal name: Department of Energy, Systems, Territory and Construction Engineering Summary: The current digitization trend, the increased attention towards sustainability, and the spread of the business analytics call for higher efficiency in port operations and for investigating the quantitative approaches for maritime logistics and freight transport systems. Process mining enabled enhancements in the overall export time length, which might improve the vessels' turnover

and reduce the corresponding operational costs, and supported the potential re-design of performance indicators in process control and monitoring.

Paper 5:

Dimensions of the Port Performance: A Review of Literature Publication year: 25 August ,2020 Author name: Bucak, U., Ba saran Journal name: Journal of ETA Maritime Science Summary: The port performance has frequently been studied in the academic literature, and the first studies on the subject are focused on financial or operational dimensions. However, today, port performance has become multi-dimensional due to the changing roles of the ports to its stakeholders, and the fact that local competition has been replaced by global competition through continuously developing routes, etc. Within this study, it is aimed to determine each dimension of the port performance concept which had been handled as a multidimensional process in recent years in literature. So, the concept of port performance had been divided into four basic dimensions which are operational, financial, sustainable, and logistics.

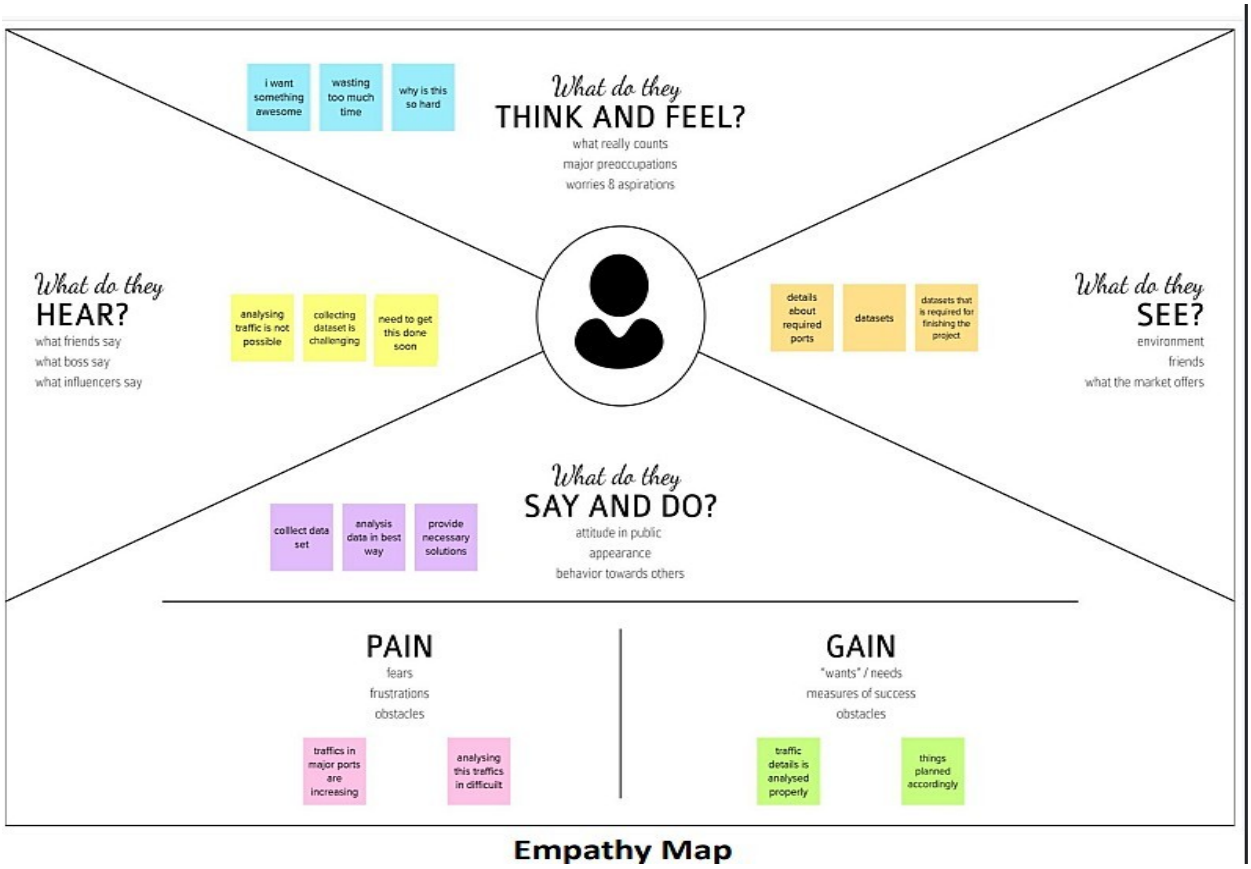
b.Problem Statement Definition

This aims at developing a Machine Learning Model for Traffic analysis for major ports. The Indian Railways has a capital base of about Rs. 100000 crores and is often referred to as the lifeline of the Indian economy because of its predominance in transportation of bulk freight and long distance passenger traffic. The network crisis-crosses the nation, binding it together by ferrying freight and passengers across the length and breadth of the country. As the Indian economy moves into a high growth trajectory the Railways have also stepped-up developmental efforts and are preparing themselves for an even bigger role in the future at the same time it became hard to analyze traffic in major ports and our project helps to overcome that problem.

QUESTION	DESCRIPTION
Who does the problem affect?	Indian railways
Why is it important?	As Indian railway play major role in Indian economy it is important to analyze the traffic in major ports
What are the benefits?	• AI along with ML model • Automatic Prediction • Data Analysis
How is it better than the others?	Faster Processing of data with higher accuracy and optimized model.
When to use?	Scenario where we want to analyze the traffic in major ports.

IDEATION & PROPOSED SOLUTION

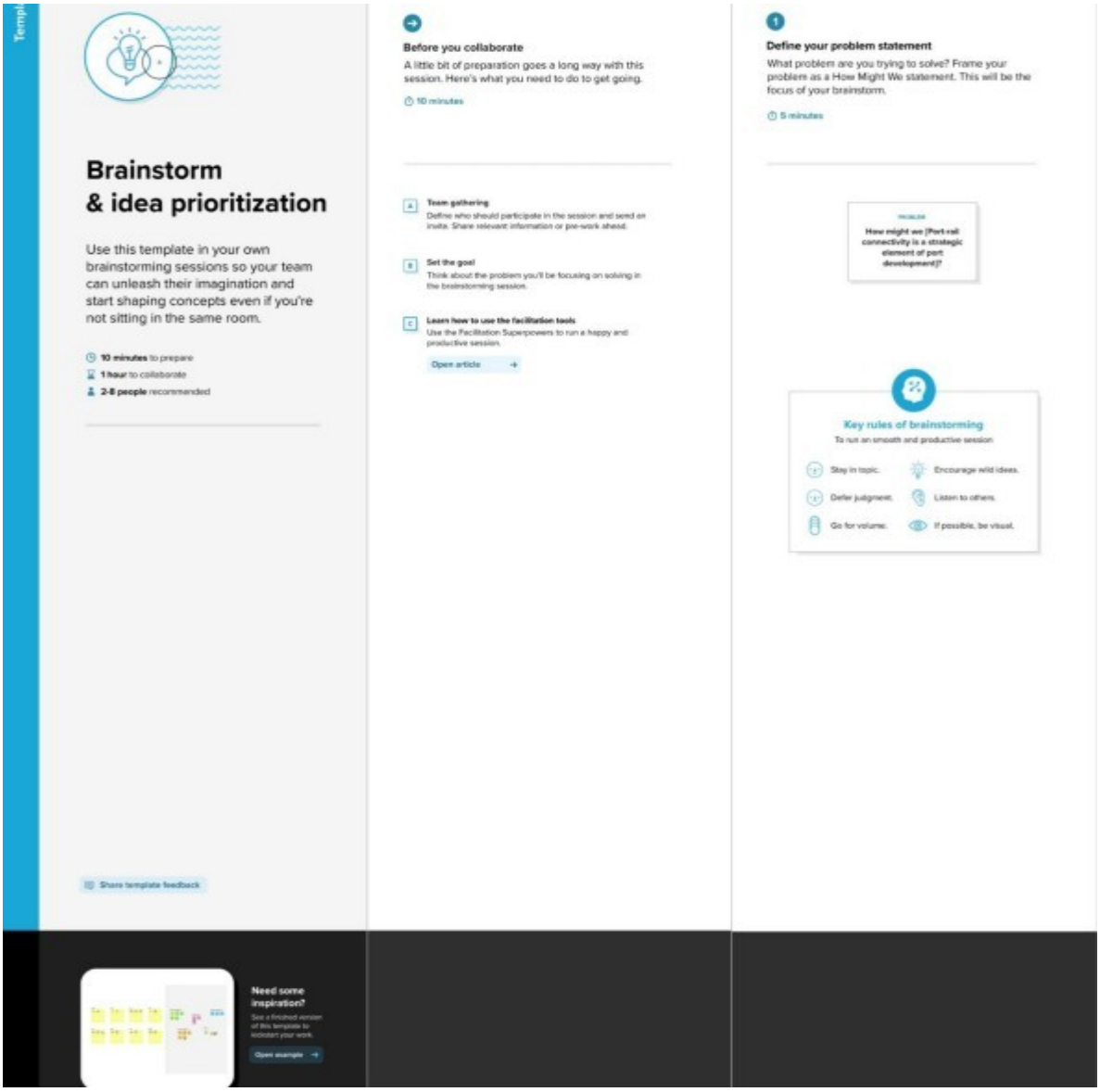
a.Empathy Map Canvas



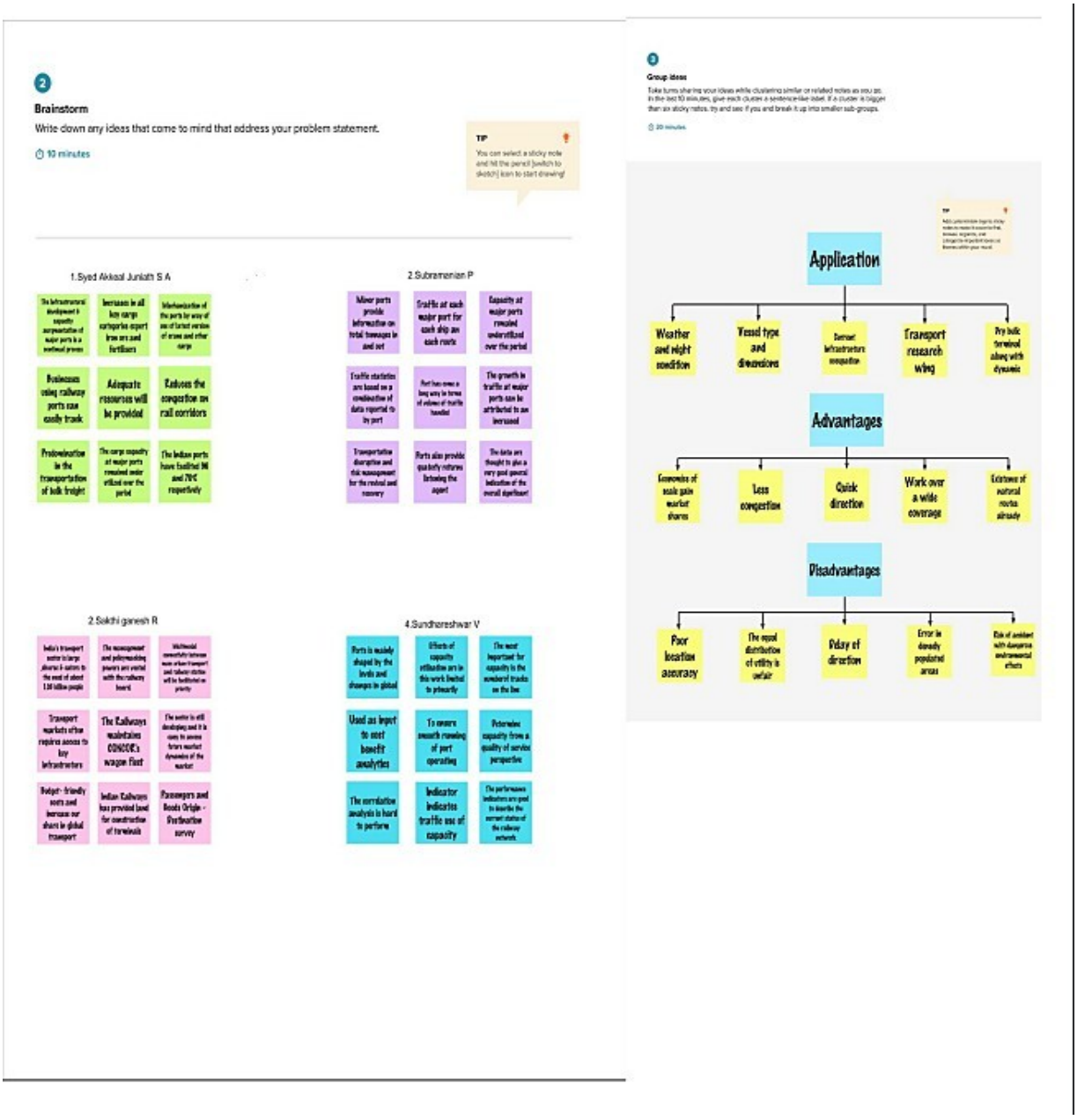
Empathy Map

b.Ideation & Brainstorming

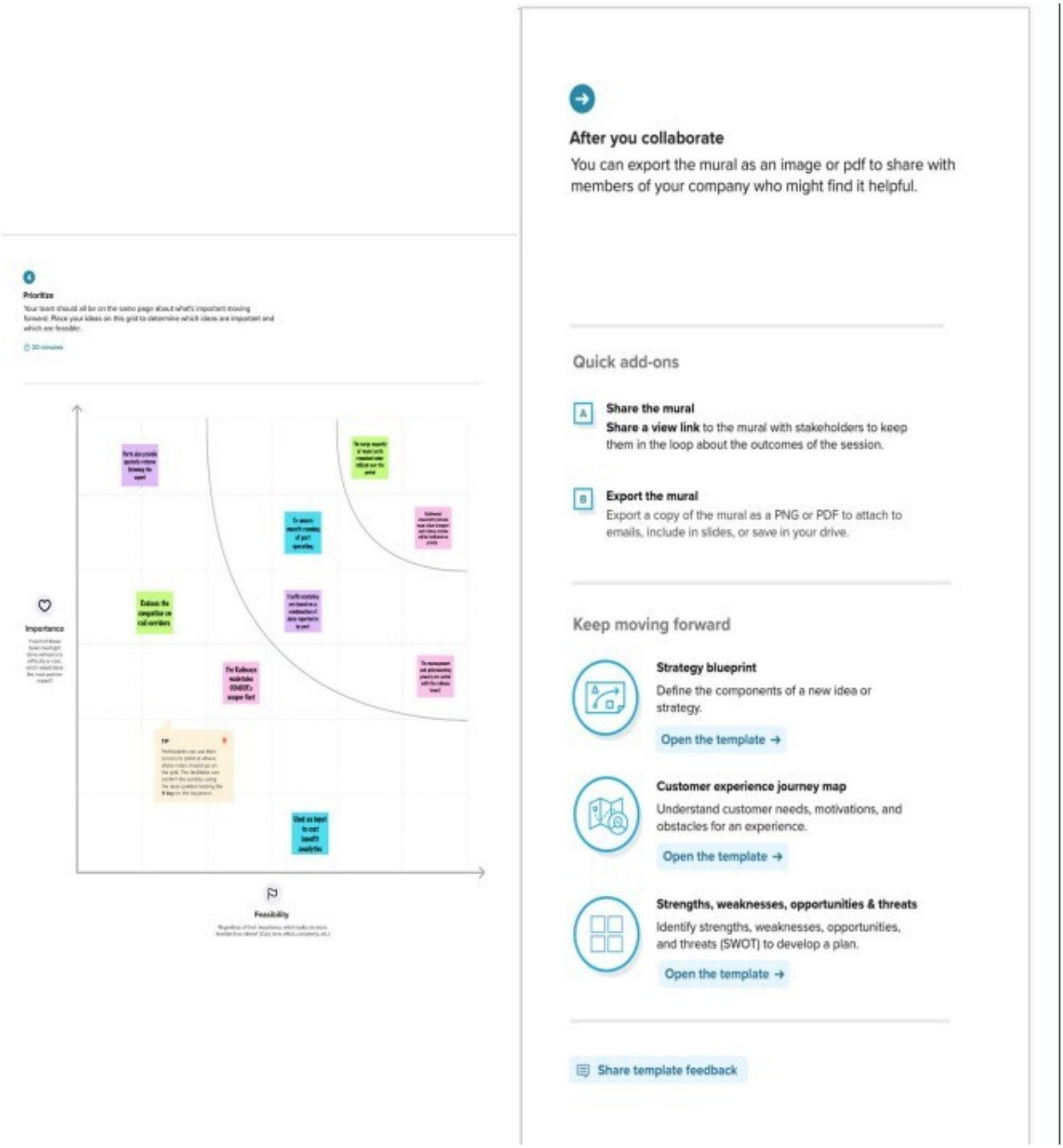
Step-1: Team Gathering, Collaboration and Select the Problem Statemen



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



c.Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	The Indian Railways has a capital base of about 1 lakhs crores and is often referred to as the lifeline of the Indian economy. As it includes transportation of bulk freight and long-distance passengers, traffic and congestion on rail corridors becomes a major challenge.
2.	Idea / Solution description	Data analytics can be applied to visualize freight transportation and congestion on rail corridors across major railway ports to get better insight of the working of port network and to improve the port connectivity.
3.	Novelty / Uniqueness	Can also predict the time at which the particular train will arrive and depart.
4.	Social Impact / Customer Satisfaction	Adequate resources will be provided for the customers regarding the arrival, departure and delay of the trains.
5.	Business Model (Revenue Model)	Businesses using railway ports can easily track the trains. Government can use data analytics dashboard to ensure less traffic on the ports.
6.	Scalability of the Solution	The solution can be used almost for all modes of transportation including the ships and so on. Thus it is scalable for almost all modes of transportation.

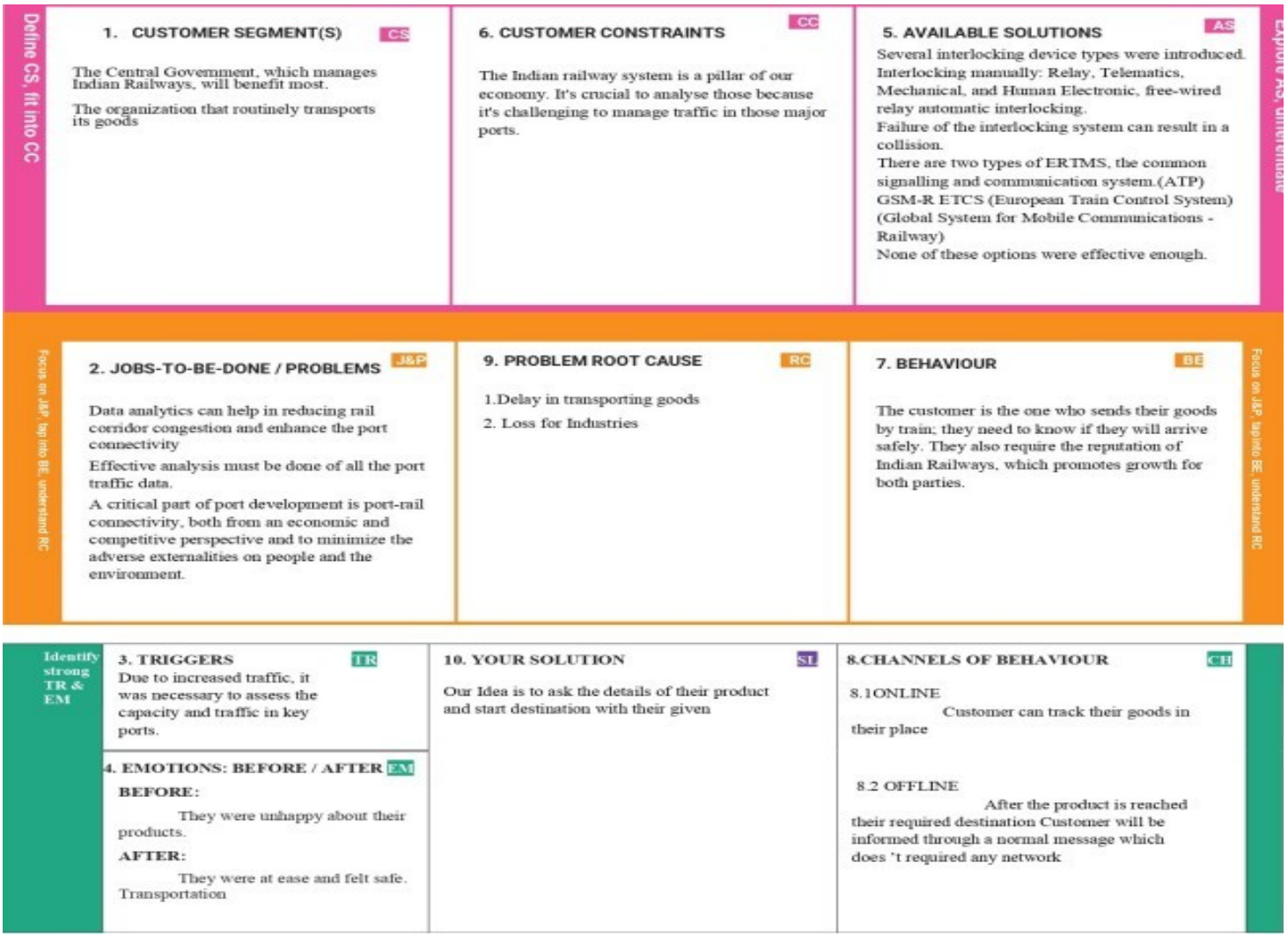
d.Problem – Solution Fit Template:

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem. It helps entrepreneurs, marketers and corporate innovators identify behavioral patterns and recognize what would work and why

Purpose:

a. Solve complex problems in a way that fits the state of your customers.
b. Succeed faster and increase your solution adoption by tapping into existing mediums and channels of behavior.
c. Sharpen your communication and marketing strategy with the right triggers and messaging.
d. Increase touch-points with your company by finding the right problembehavior fit and building trust by solving frequent annoyances, or urgent or costly problems.
e. Understand the existing situation in order to improve it for your target group.

Template:


Define CS, fit into CC
1. CUSTOMER SEGMENT(S) CS
The Central Government, which manages Indian Railways, will benefit most.
The organization that routinely transports its goods
6. CUSTOMER CONSTRAINTS CC
The Indian railway system is a pillar of our economy. It's crucial to analyse those because it's challenging to manage traffic in those major ports.
5. AVAILABLE SOLUTIONS AS
Several interlocking device types were introduced.
Interlocking manually: Relay, Telematics, Mechanical, and Human Electronic, free-wired relay automatic interlocking.
Failure of the interlocking system can result in a collision.
There are two types of ERTMS, the common signalling and communication system.(ATP)
GSM-R ETCS (European Train Control System) (Global System for Mobile Communications - Railway)
None of these options were effective enough.
Explore AS, differentiate
Focus on J&P, tap into BE, understand RC
2. JOBS-TO-BE-DONE / PROBLEMS J&P
Data analytics can help in reducing rail corridor congestion and enhance the port connectivity
Effective analysis must be done of all the port traffic data.
A critical part of port development is port-rail connectivity, both from an economic and competitive perspective and to minimize the adverse externalities on people and the environment.
9. PROBLEM ROOT CAUSE RC
1.Delay in transporting goods
2. Loss for Industries
7. BEHAVIOUR BE
The customer is the one who sends their goods by train; they need to know if they will arrive safely. They also require the reputation of Indian Railways, which promotes growth for both parties.
Focus on J&P, tap into BE, understand RC
Identify strong TR & EM
3. TRIGGERS TR
Due to increased traffic, it was necessary to assess the capacity and traffic in key ports.
4. EMOTIONS: BEFORE / AFTER EM
BEFORE:
They were unhappy about their products.
AFTER:
They were at ease and felt safe. Transportation
10. YOUR SOLUTION SL
Our Idea is to ask the details of their product and start destination with their given
8.CHANNELS OF BEHAVIOUR CH
8.1ONLINE
Customer can track their goods in their place
8.2 OFFLINE
After the product is reached their required destination Customer will be informed through a normal message which does 't required any network

REQUIREMENT ANALYSIS

A.Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	➢ Registers a new user through registration form or mail .
FR-2	User Confirmation	➢ Confirmation through Email or OTP.
FR-3	Data collection	➢ Data collection. Relevant data is gathered from operational systems, data warehouses, data lakes and other data sources. • Data discovery and profiling. • Data cleansing. • Data structuring. • Data transformation and enrichment. • Data validation and publishing.
FR-4	Data Pre-processing	➢ Data preprocessing a component of data preparation, describes any type of processing performed on raw data toprepare it for another data processing procedure. It has traditionally been an important preliminary step for the datamining process.
FR-5	Model Evaluation	➢ Model evaluation is the process of using different evaluationmetrics to understand a machine learning model's performance, as well as its strengths and weaknesses. Modelevaluation is important to assess the efficacy of a model during initial research phases, and it also plays a role in model monitoring

FR-6	Prediction Output	➢ Predictive analytics is the process of using data analytics to make predictions based on data. This process uses data alongwith analysis, statistics, and machine learning techniques to create a predictive model for forecasting future events.

Non-functional Requirements:

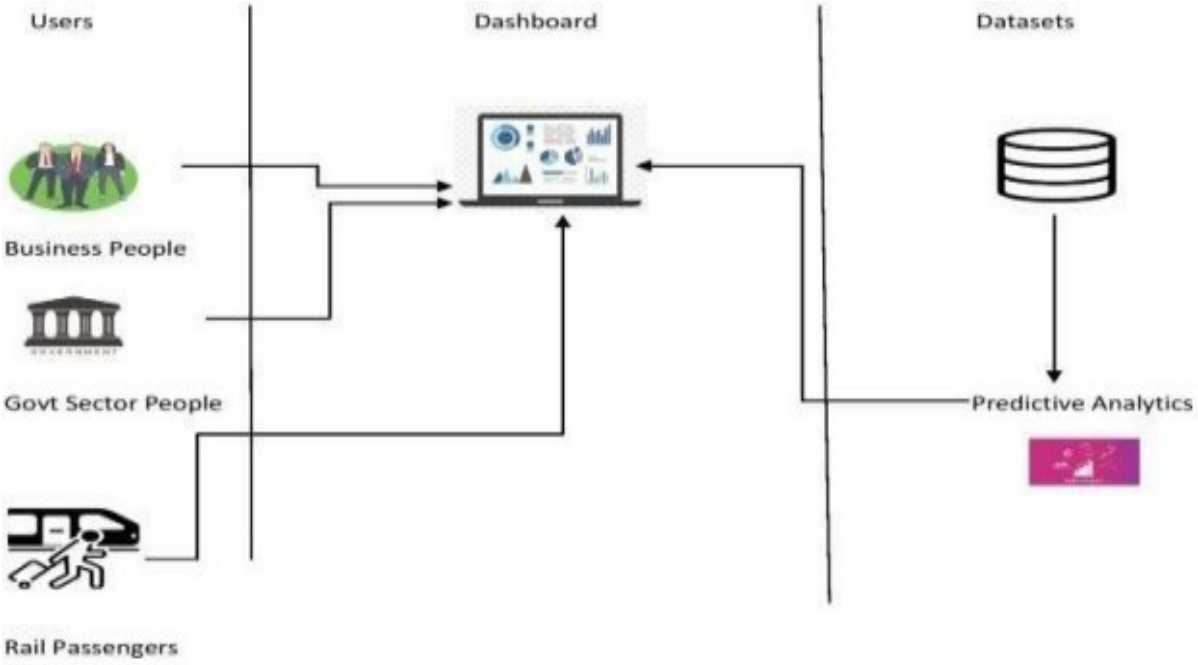
Following are the non-functional requirements of the proposed solution.

NFR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	➢ It helps the farmers to monitor the health of the crops in real time, create predictive analysis related to futureyield.
NFR-2	**Security**	➢ Data security functions to prevent data breaches, reduce risk of data exposure and ensure the ongoingsafe and secure use of private data by minimizing exposure risk.
NFR-3	**Reliability**	➢ The reliability of the data determines whether businesses can make good decisions or not. If the datais unreliable it is useless to the organizations
NFR-4	**Performance**	➢ Regularly evaluating the performance of the organization can help us to understand how much progress we're making towards our goal. A performance analysis is a tool you can use to check important metrics of crop yield for very month or yearand make plans for adjustment and improvement.
NFR-5	**Availability**	➢ Data should be available for access at anytime fromanywhere.
NFR-6	**Scalability**	➢ The software should be flexible and other developersmust be able to improve its capabilities.

PROJECT DESIGN:

a.Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and wheredata is stored.

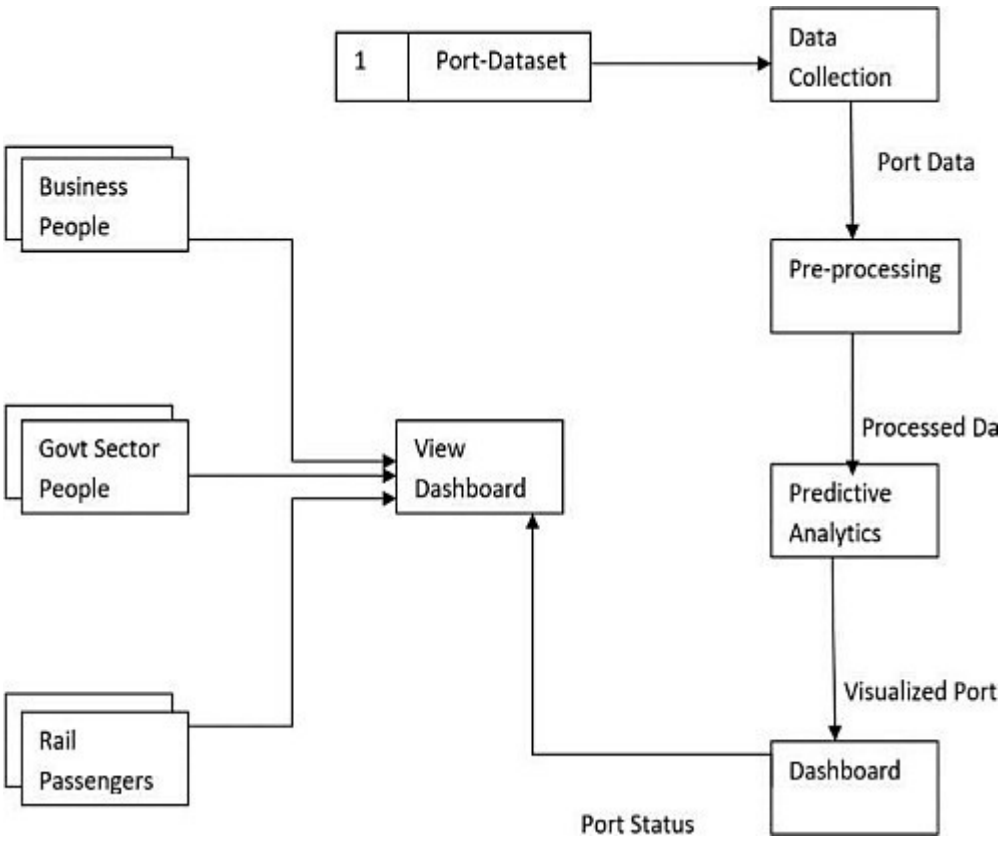


1.) Predictive analytics will be done from collected dataset and it will be updated in dashboard.

2.) Business People can able to view the dashboard to track their goods.

3.) Govt Sector People can able to predict the congestion in ports by viewing the dashboard and it helps to avoid congestion in future .

4.) Rail Passengers can able to track the correct time of rail in ports.



b.Solution Architecture:

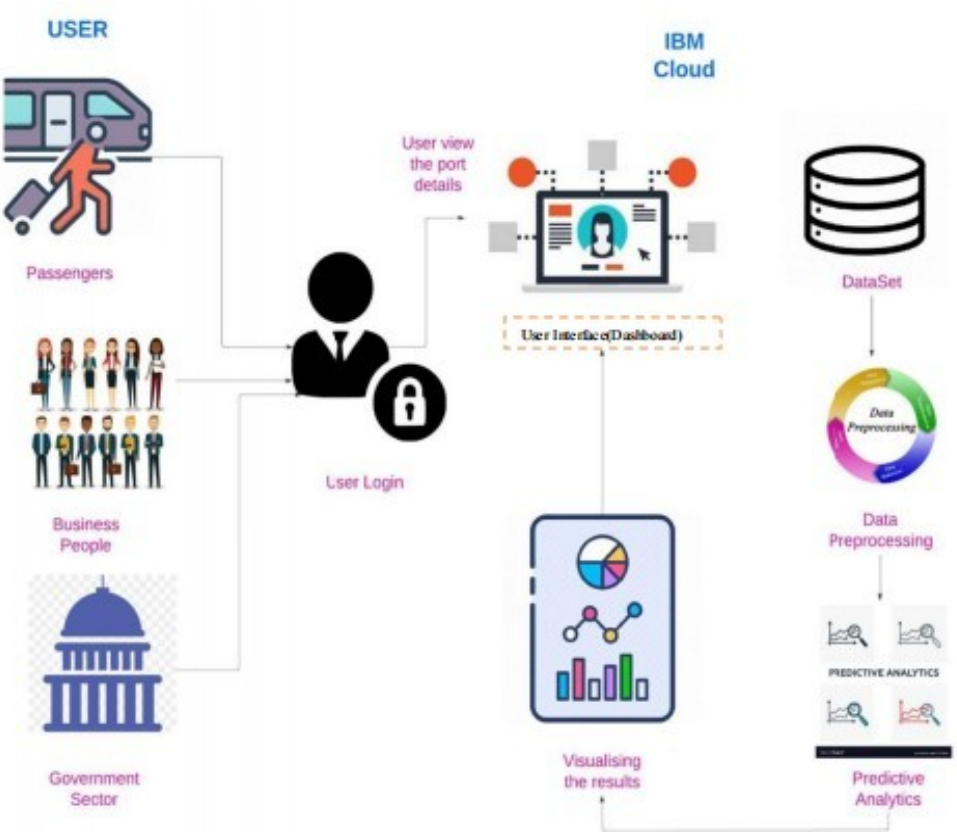
Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions.Its goals are to:

i. Find the best tech solution to solve existing business problems.

ii. Describe the structure, characteristics, behavior, and other aspects ofthe software to project stakeholders.

iii. Define features, development phases, and solution requirements.

iv. Provide specifications according to which the solution is defined, managed, and delivered.



c.User Stories:

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Business People	Monitoring	USN-1	As a user, I can view the dashboard to see the port status .	I can visualize the port status in dashboard.	High	Sprint-1
	Tracking	USN-2	As a user,I can track the goods.	I can track the goods by it's arrival/departure time	High	Sprint-1
Government Sector People	Viewing	USN-1	As a user,I can view the port status regularly	I can able to know the port status	Low	Sprint-2
	Predicting	USN-2	As a user,I will reduce the congestion in ports by predicting the port congestion through dashboard.	I can able to predict the congestion in future	High	Sprint-2
Passengers	Tracing	USN-1	As a user, I can trace the arrival/departure time of rail in ports.	I can able to track the correct time of rail.	High	Sprint-2

PROJECT PLANNING & SCHEDULING:

a.Prepare Milestone and Activity List

DURATION	MILESTONE	ASSIGNEE	STATUS	DESCRIPTION
1 WEEK	Data Collection-Download dataset	Uvagai.K.E,Anuja.G Monisha.B,Logeshwari.Pavithra.V	Completed	The dataset for Traffic and Capacity Analytics is to be collected.The datasetwhich is considered will have the port information
1 WEEK	Data Pre-processing 1.Renaming the coloumn names 2.Preparing calculations 3.Checking for NULL values 4.Checking for ouliers 5.Summarization of dataset 6.Label Encoding	Uvagai.K.E,Anuja.G Monisha.B Logeshwari.Pavithra.V	In-Progress	Preprocessing involves renaming the existing coloumn names into meaningfulone,preparing calculations such as calculating traffic percent,checking for NULL values in the dataset.
1 WEEK	Visualizing the dataset	Uvagai.K.E,Anuja.G Monisha.B Logeshwari Pavithra.V	completed	Visulaizing the dataset involves plotting thedataset using various plots and doing analysis on that.
1 WEEK	Model Building 1. Building the model using suitable machinelearning algorithm 2. Training and testing the model	Uvagai.K.E,Anuja.G Monisha.B Logeshwari Pavithra.V	In -Progress	Using certain algorithms to build the model.Those algorithms include 1.Linear regression

1 WEEK	Dashboard Creation	Uvagai. K.E,Anu ja.G Monisha.B Logeshwari Pavithra.V	completed	Dashboard for visualizing the port statuswill be developed.
1 WEEK	Ideation Phase 1. Litreature survey on the selected projectandinformation gathering. 2. Prepare the empathy map. 3. Ideation	Uvagai. K.E,Anu ja.G Monisha.B Logeshwari Pavithra.V	Completed	Start the ideation process
1 WEEK	Project Design Phase -1 1. Proposed solution. 2. Prepared fit solution 3. Solution Architecture	Uvagai. K.E,Anu ja.G Monisha.B Logeshwari Pavithra.V	Completed	Prepare the proposed solution document,which includes the novelty, feasibility ofidea, business model, social impact, scalability of solution, etc.
1 WEEK	Project Design Phase -2 1. Customer journey. 2. Functional requirements 3. Data flow diagram. 4. Technology architecture	Uvagai. K.E,Anu ja.G Monisha.B Logeshwari Pavithra.V	Completed	Prepare the customer journey maps to understand the user interactions & experiences with the application (entry to exit) ,Functional requirements and constructarchitecture
1 WEEK	Project Planning Phase 1.Milestone Activity List 2.Sprint Delivery plan	Uvagai. K.E,Anu ja.G Monisha.B Logeshwari Pavithra.V	Completed	Prepare milestone activity list and sprintdelivery plan for outline of work flow
1 WEEK	Project Development Phase1.Sprint -1 2 . S p r i n t - 2	Uvagai. K.E,Anu ja.G Monisha.B Logeshwari Pavithra.V	In-Progress	Plan of each task sprint to be developed.

	3. Sprint-3 4. Sprint-4			

b.Sprint Delivery Plan

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Project flow, Objectives	USN-1	Create a Project Flow	5	Medium	Whole Team
	IBM Account and Loading the dataset	USN-2	Create a IBM Account and dataset	5	Medium	Anuja
	Dataset and Calculations	USN-3	Prepare the Dataset and Prepare the Calculation	10	High	Pavithra
Sprint-2	Data Visualization Charts	USN-4	Port wise Traffic Distribution ,Traffic Vs Capacityand Prepare Line & Bar Chart ,Area Chart	20	High	Whole Team
Sprint-3	Literature Survey and Empathy Map , ideation	USN-5	Literature Survey On The Selected Project & Information Gathering and Prepare Empathy map,ideation	10	High	Whole Team
Sprint-3	Proposed Solution And Problem Solution fit	USN-6	To Prepare the Proposed Solution And Problem Solution fit	5	Medium	Whole Team
Sprint-3	Solution Architecture	USN-7	To Prepare the Solution Architecture	5	Medium	Whole Team
Sprint-4	Customer journey, functional Requirement	USN-8	To Create a Customer journey and functionalRequirement	10	high	Logeshwari, monisha,pavithra
Sprint-4	Data Flow, Technology Architecture	USN-9	To Prepare Data Flow, TechnologyArchitecture	5	medium	Uvagai

Sprint-4	Milestone &ActivityList	USN-10	To create Milestone and Activity List	5	mediu m	Uvagai

TESTING

```
import numpy as np
import pandas as pd

# Loading the dataset
df = pd.read_csv('D:/ibm/datafile_02.csv')
print(df.columns)
df.head()
```

```
Index(['Port', 'Traffic in Eleventh Plan (MT) (2011-12)Proj.',
       'Traffic in Eleventh Plan (MT) (2011-12) Ach.',
       'Traffic in Eleventh Plan (MT) (2011-12) %',
       'Total Capacity in Eleventh Plan (MT) (2011-12) Proj.',
       'Total Capacity  in Eleventh Plan (MT) (2011-12) Ach.',
       'Total Capacity in Eleventh Plan (MT) (2011-12) %'],
      dtype='object')
```

	Port	Traffic in Eleventh Plan (MT) (2011-12)Proj.	Traffic in Eleventh Plan (MT) (2011-12) Ach.	Traffic in Eleventh Plan (MT) (2011-12) %	Total Capacity in Eleventh Plan (MT) (2011-12) Proj.	Total Capacity in Eleventh Plan (MT) (2011-12) Ach.
0	Kolkata	1343	1223	9100	3145	1635
1	Haldia	4450	3101	7000	6340	5070
2	Paradeep	7640	5425	7100	10640	7650
3	Visakhapatnam	8220	6742	8200	10810	7293
4	Ennore	4700	1496	3200	6420	3100

```
   Total Capacity  in Eleventh Plan (MT) (2011-12) %
0                                                5100
1                                                7900
2                                                7100
3                                                6700
4                                                4800
```

```
# Preprocessing the dataset

# Renaming the columns
df.rename(columns = {'Traffic in Eleventh Plan (MT) (2011-
12)Proj.':'Traffic_Projected','Traffic in Eleventh Plan (MT) (2011-12)
Ach.':'Traffic_Achieved', 'Total Capacity in Eleventh Plan (MT)
(2011-12) Proj.':'Total_Capacity_Projected', 'Total Capacity in
Eleventh Plan (MT) (2011-12) Ach.':'Total_Capacity_Achieved'}, inplace
= True)
df
```

```
             Port  Traffic_Projected  Traffic_Achieved  \
0         Kolkata               1343              1223
1          Haldia               4450              3101
2        Paradeep               7640              5425
3   Visakhapatnam               8220              6742
4          Ennore               4700              1496
5         Chennai               5750              5571
6       Tuticorin               3172              2810
7          Cochin               3817              2010
8            NMPT               4881              3294
9        Mormugao               4455              3900
10         Mumbai               7105              5618
11           JNPT               6604              6575
12         Kandla               8672              8250

    Traffic in Eleventh Plan (MT) (2011-12) %
Total_Capacity_Projected  \
0                                          9100
3145
1                                          7000
6340
2                                          7100
10640
3                                          8200
10810
4                                          3200
6420
5                                          9700
7230
```

```
6                                        8900
6398
7                                        5300
5475
8                                        6800
6050
9                                        8800
6690
10                                       7900
9191
11                                      10000
9560
12                                       9500
12220

    Total_Capacity_Achieved Total Capacity in Eleventh Plan (MT)
(2011-12) %
0                       1635
5100
1                       5070
7900
2                       7650
7100
3                       7293
6700
4                       3100
4800
5                       7972
11000
6                       3334
5200
7                       4098
7400
8                       5097
8400
9                       4190
6200
10                      4453
4800
11                      6400
6600
12                      8691
7100
```

```
# Perparing the Calculations:

Traffic_Percent =
round((df.Traffic_Achieved/df.Traffic_Projected)*100,2)

Traffic_Percent
```

```
0     91.06
1     69.69
2     71.01
3     82.02
4     31.83
5     96.89
6     88.59
7     52.66
8     67.49
9     87.54
10    79.07
11    99.56
12    95.13
dtype: float64
```

```
Total_Percent =
round( (df.Total_Capacity_Achieved/df.Total_Capacity_Projected)*100,2)
Total_Percent
```

```
0      51.99
1      79.97
2      71.90
3      67.47
4      48.29
5     110.26
6      52.11
7      74.85
8      84.25
9      62.63
10     48.45
11     66.95
12     71.12
dtype: float64
```

```
# Replacing the existing columns with newly created columns
df.rename(columns = {'Traffic in Eleventh Plan (MT) (2011-12)
%':'Traffic_Percent','Total Capacity  in Eleventh Plan (MT) (2011-12)
%':'Total_Percent'}, inplace = True)
df.iloc[:,3:4] = Traffic_Percent
df.iloc[:,6:] = Total_Percent
df
```

```
             Port Traffic_Projected Traffic_Achieved
Traffic_Percent \
0         Kolkata              1343             1223
91.06
1          Haldia              4450             3101
69.69
2        Paradeep              7640             5425
71.01
3   Visakhapatnam              8220             6742
```

```
82.02
4          Ennore               4700              1496
31.83
5         Chennai               5750              5571
96.89
6       Tuticorin               3172              2810
88.59
7          Cochin               3817              2010
52.66
8            NMPT               4881              3294
67.49
9        Mormugao               4455              3900
87.54
10         Mumbai               7105              5618
79.07
11           JNPT               6604              6575
99.56
12         Kandla               8672              8250
95.13
```

	Total_Capacity_Projected	Total_Capacity_Achieved	Total_Percent
0	3145	1635	51.99
1	6340	5070	79.97
2	10640	7650	71.90
3	10810	7293	67.47
4	6420	3100	48.29
5	7230	7972	110.26
6	6398	3334	52.11
7	5475	4098	74.85
8	6050	5097	84.25
9	6690	4190	62.63
10	9191	4453	48.45
11	9560	6400	66.95
12	12220	8691	71.12

```
df.shape
```

```
(13, 7)
```

```
# Checking for null values
```

```
df.isnull().sum()
```

```
Port                        0
Traffic_Projected           0
Traffic_Achieved            0
Traffic_Percent             0
Total_Capacity_Projected    0
Total_Capacity_Achieved     0
Total_Percent               0
dtype: int64
```

```
# Summary of Dataset
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13 entries, 0 to 12
Data columns (total 7 columns):
 #   Column                    Non-Null Count  Dtype
---                            -------------  -------
 0   Port                      13 non-null     object
 1   Traffic_Projected         13 non-null     int64
 2   Traffic_Achieved          13 non-null     int64
 3   Traffic_Percent           13 non-null     float64
 4   Total_Capacity_Projected  13 non-null     int64
 5   Total_Capacity_Achieved   13 non-null     int64
 6   Total_Percent             13 non-null     float64
dtypes: float64(2), int64(4), object(1)
memory usage: 856.0+ bytes
```

```
df.describe()
```

```
       Traffic_Projected  Traffic_Achieved  Traffic_Percent  \
count          13.000000         13.000000        13.000000
mean         5446.846154       4308.846154        77.887692
std          2133.280019       2212.894855        19.382398
min          1343.000000       1223.000000        31.830000
25%          4450.000000       2810.000000        69.690000
50%          4881.000000       3900.000000        82.020000
75%          7105.000000       5618.000000        91.060000
max          8672.000000       8250.000000        99.560000

       Total_Capacity_Projected Total_Capacity_Achieved
Total_Percent
count                 13.000000               13.000000
13.000000
mean                7705.307692             5306.384615
68.480000
std                 2570.242673             2140.254796
17.252637
min                 3145.000000             1635.000000
48.290000
25%                 6340.000000             4098.000000
52.110000
50%                 6690.000000             5070.000000
67.470000
75%                 9560.000000             7293.000000
74.850000
max                12220.000000             8691.000000
110.260000
```

```
cor = df.corr
cor
```

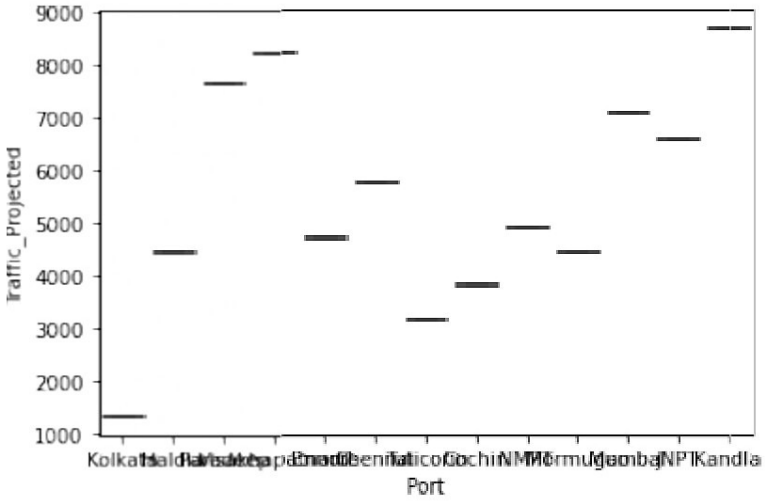
```
<bound method DataFrame.corr of              Port Traffic_Projected
Traffic_Achieved Traffic_Percent \
0         Kolkata               1343              1223
91.06
1          Haldia               4450              3101
69.69
2        Paradeep               7640              5425
71.01
3   Visakhapatnam               8220              6742
82.02
4          Ennore               4700              1496
31.83
5         Chennai               5750              5571
96.89
6       Tuticorin               3172              2810
88.59
7          Cochin               3817              2010
52.66
8            NMPT               4881              3294
67.49
9        Mormugao               4455              3900
87.54
10         Mumbai               7105              5618
79.07
11           JNPT               6604              6575
99.56
12         Kandla               8672              8250
95.13

    Total_Capacity_Projected  Total_Capacity_Achieved  Total_Percent
0                       3145                     1635          51.99
1                       6340                     5070          79.97
2                      10640                     7650          71.90
3                      10810                     7293          67.47
4                       6420                     3100          48.29
5                       7230                     7972         110.26
6                       6398                     3334          52.11
7                       5475                     4098          74.85
8                       6050                     5097          84.25
9                       6690                     4190          62.63
10                      9191                     4453          48.45
11                      9560                     6400          66.95
12                     12220                     8691          71.12
>
```

```
#Finding Outliers anr replacing the outliers
import matplotlib.pyplot as plt
import seaborn as sns

sns.boxplot(x='Port',y='Traffic_Projected',data=df)
```

```
plt.rcParams["figure.figsize"] = [17.50, 3.50]
plt.rcParams["figure.autolayout"] = True
```




```
# Check For Categorical Columns and do encoding


from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
print(df.Port.value_counts())

df.Port = le.fit_transform(df.Port)
print(df.Port.value_counts())
```

```
Kolkata          1
Haldia           1
Paradeep         1
Visakhapatnam    1
Ennore           1
Chennai          1
Tuticorin        1
Cochin           1
NMPT             1
Mormugao         1
Mumbai           1
JNPT             1
Kandla           1
Name: Port, dtype: int64
6     1
```

```
3     1
10    1
12    1
2     1
0     1
11    1
1     1
9     1
7     1
8     1
4     1
5     1
Name: Port, dtype: int64
```

```
# Classification

#y = df.Traffic_Percent
#print(y)

#df.drop(['Traffic_Percent'],axis=1)

df.head()
```

```
   Port  Traffic_Projected  Traffic_Achieved  Traffic_Percent  \
0     6               1343              1223            91.06
1     3               4450              3101            69.69
2    10               7640              5425            71.01
3    12               8220              6742            82.02
4     2               4700              1496            31.83

   Total_Capacity_Projected  Total_Capacity_Achieved  Total_Percent
0                      3145                     1635          51.99
1                      6340                     5070          79.97
2                     10640                     7650          71.90
3                     10810                     7293          67.47
4                      6420                     3100          48.29
```

```
ddf = df.drop(['Traffic_Percent'],axis=1)
ddf
```

```
    Port Traffic_Projected Traffic_Achieved
Total_Capacity_Projected \
0      6               1343              1223
3145
1      3               4450              3101
6340
2     10               7640              5425
10640
3     12               8220              6742
```

```
10810
4      2               4700              1496
6420
5      0               5750              5571
7230
6     11               3172              2810
6398
7      1               3817              2010
5475
8      9               4881              3294
6050
9      7               4455              3900
6690
10     8               7105              5618
9191
11     4               6604              6575
9560
12     5               8672              8250
12220
```

```
    Total_Capacity_Achieved  Total_Percent
0                      1635          51.99
1                      5070          79.97
2                      7650          71.90
3                      7293          67.47
4                      3100          48.29
5                      7972         110.26
6                      3334          52.11
7                      4098          74.85
8                      5097          84.25
9                      4190          62.63
10                     4453          48.45
11                     6400          66.95
12                     8691          71.12
```

```
x = ddf.iloc[:,1:]
print(x)
```

```
    Traffic_Projected  Traffic_Achieved  Total_Capacity_Projected  \
0                1343              1223                      3145
1                4450              3101                      6340
2                7640              5425                     10640
3                8220              6742                     10810
4                4700              1496                      6420
5                5750              5571                      7230
6                3172              2810                      6398
7                3817              2010                      5475
8                4881              3294                      6050
9                4455              3900                      6690
10               7105              5618                      9191
```

ADVANTAGES & DISADVANTAGES

The spatial distribution of the costs and benefits of port activity further complicate the tasks of attributing impacts, distributing mitigation and compensation, and identifying institutional actors with the willingness and ability to overcome the inevitable collective action problems. Strategies to internalise the externalities of port activity are vital, yet they are also limited by the spatial dynamics and other complexities that come with an activity that has multiple connections to the urban economy. In this context, governance frameworks that include all the relevant actors in the search for collaborative solutions to improve traffic management and planning, but that also have the institutional power to enforce them, are probably more important than any one single intervention. In that spirit, I will conclude with an observation about some common elements of the most promising strategies that are emerging in port-city-hinterland connectivity around the world. In the places where maritime and inland terminal operators, and the transport providers which provide the connective linkages between them, have come together to internalise the costs of some externalities, they have done so because they have been prompted by the political intervention of key stakeholders. For example, the PierPass system in Los Angeles and Long Beach really was a pre-emptive action by terminal operators to avoid even more stringent and

potentially unworkable regulation from the state agencies (Giuliano and Linder, 2013). These state agencies in turn were responding to pressures from locally elected representatives, who in turn were responding to the needs of their constituents (Hall, 2007). Likewise, the reservation system, and subsequent actions in Vancouver to compensate truckers for waiting time was the result of strike action by truckers. We have not yet seen the container terminal operating industry take proactive leadership alone in traffic issues beyond the gates, and perhaps this would be an unwelcome intrusion in an urban democracy. However, there are interesting and important examples of action by public authorities, often in partnership with private actors, of traffic planning that works for multiple interest.

CONCLUSION

Indian firms have focused on interconnected and lean supply chains to overcome the supply gaps in normal business operations. The COVID-19 pandemic has led to massive SCDs due to undiscovered supply chain vulnerabilities causedby government-imposed economic restrictions including transportation disruptions worldwide including India, which adversely impacted the normal functioning of the firms.

Many Indian firms have experienced severe disruptions in transportation and logistics services, including stronger impact on transportation and logistics data, time delays, and cargo cancellations due to drastically reduced freight capacity, limited mobility, ports shutdown, and problems in routine customs clearances. All this has also severely delayed the production of goods, transport consignments, and logistics services thereby caused massive delays and rerouting to final consumers. The suggested model of robust transport and ALS can be widely used by firms for speedier SCR in the context of economic crises like the COVID-19 pandemic. Over the period, the government has gradually removed most of the restrictions and the firms have made concerted efforts to speedily recover from SCDs, however, inadequate applications of robust TI and

ALS have delayed the SCR by the firms. This calls for reviewing current transport and ALS used by firms on priority for speedier SCR. Therefore, the suggested model can be widely applied to address the SCDs using robust intelligence transportation systems and ALS. The challenges and opportunities in operationalizing the suggested model along with optimization of transport and logistics resources should also be considered bythefirms.

APPENDIX

Github: https://github.com/IBM-EPBL/IBM-Project-15884-1664532451