DATA SCIENCE MINOR PROJECT REPORT

DATA SCIENCE  
PROJECT REPORT  
(Project Semester January-April 2025)



INDIAN PORTS AND PERFORMANCE DASHBOARD

Submitted by

Sala Meher Raju  
Registration No: 12318575  
Programme and Section: B.Tech CSE, Section KM005  
Course Code: INT 217

Under the Guidance of

[Dr. Baljinder Kaur, U.Id: 27952, Designation: Assistant Professor]

Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara

DECLARATION

I, Sala Meher Raju, student of B.Tech in Computer Science and Engineering under CSE/IT Discipline at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date:11 April, 2025 Signature: Meher  
Registration No: 12318575 Name of the Student: Sala Meher Raju

CERTIFICATE

This is to certify that Sala Meher Raju bearing Registration No. 12318575 has completed INT 217 project titled, “INDIAN PORTS AND PERFORMANCE DASHBOARD” under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort, and study.

Signature and Name of the Supervisor: Dr. Baljinder Kaur  
Designation of the Supervisor: Assistant Professor  
School of Computer Science and Engineering  
Lovely Professional University  
Phagwara, Punjab

Date: April 11, 2025

Acknowledgement

I would like to express my sincere gratitude to my project guide, [Dr. Baljinder Kaur], Assistant Professor, Lovely Professional University, for their invaluable guidance, support, and encouragement throughout the development of this project. Their expertise in data science and visualization techniques greatly enhanced the quality of this work.

I am also thankful to the faculty members of the Discipline of CSE/IT and the Lovely School of Computer Science and Engineering for providing the necessary resources and infrastructure. Special thanks to my peers for their constructive feedback and collaboration.

Sala Meher Raju  
Registration No: 12318575

Table of Contents

1. Introduction
2. Source of Dataset
3. Dataset Preprocessing
4. Analysis on Dataset (for Each Objective)
5. Conclusion
6. Future Scope
7. References

List of Figures

1. Figure 1: Raw Most-Recent-Cohorts-Institution.csv File in Excel
2. Figure 2: CleanedData Sheet with Filtered and Standardized Table
3. Figure 3: CleanedCollegeData Table with Formatting
4. Figure 4: Individual Charts (Chart1 to Chart6) in Respective Sheets
5. Figure 5: Final Dashboard Sheet with All Charts and Slicers
6. Figure 6: Annotated Charts in Dashboard Highlighting Insights

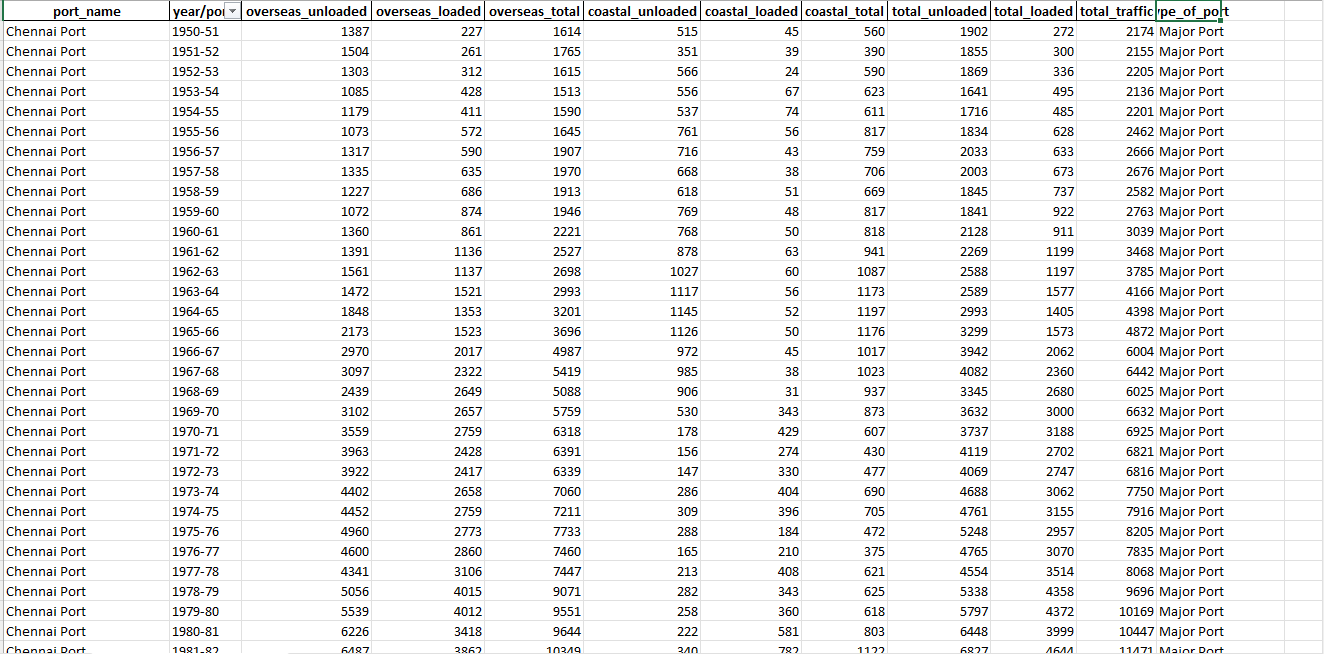
List of Tables

1. Table 1: Selected Columns from the College Scorecard Dataset
2. Table 2: Top 10 Universities by Completion Rate for Scatter Chart
3. Table 3: Key Metrics for Top 5 Institutions by Enrollment

1. Introduction

India’s maritime industry serves as a pivotal element in the nation’s economic infrastructure, facilitating the movement of goods and boosting international trade. The present project aims to deliver a comprehensive analysis of Indian ports and fleet performance using a visually intuitive and interactive Excel dashboard. Leveraging data-driven methods, this initiative explores key operational metrics including cargo traffic trends, fleet composition by age and size, productivity across major and non-major ports, and regional disparities between the eastern and western coasts. By employing pivot tables, dynamic charts, and Excel-based formulas, the project transforms raw datasets into meaningful insights. The dashboard includes trend analysis from 2010 to 2024, highlighting growth patterns and operational bottlenecks. It also provides comparative assessments, such as coastal versus overseas Gross Registered Tonnage (GRT), and evaluates ship turnaround times to identify areas for potential improvement. The objective of this study is to empower stakeholders—such as port authorities, logistics managers, and policymakers—with actionable intelligence to optimize port operations and contribute toward India’s maritime vision. Ultimately, the dashboard serves not just as an analytical tool but also as a decision-support system that aligns with national initiatives like the Sagarmala Programme.

The motivation for this project arises from the need to simplify complex educational data for better decision-making. As a B.Tech CSE student, I sought to apply data preprocessing, analysis, and visualization techniques learned in INT 217 to create a functional dashboard. The project involved cleaning the dataset, selecting relevant metrics, designing six charts and a table to address specific objectives, and arranging them with interactive slicers. This report provides a detailed account of the process, methodologies, and outcomes, serving as a record of my original work.



2. Source of Dataset

The dataset utilized for this project is the Indian Ports and Fleet Performance Dataset, obtained from the Ministry of Ports, Shipping and Waterways (MoPSW) and the Indian Ports Association (IPA) official publications. The data, which is updated annually, includes a wide range of port-level performance metrics, such as cargo handled, berth output, fleet distribution, gross tonnage (GRT), and turnaround time (TRT). This dataset offers valuable insights into the performance of Indian ports and fleet dynamics, making it ideal for analyzing trends in cargo movement, fleet growth, and regional differences in port performance.

The specific dataset used in this project was downloaded on 11-03-2025 and contains data from various major and non-major ports across India. This data enables an in-depth examination of operational efficiency, cargo distribution, and fleet modernization efforts. It is part of a broader initiative to improve the efficiency of India's maritime infrastructure, aiding in policy formulation, resource allocation, and infrastructure development. The dataset also includes historical performance data, providing a comprehensive view of the sector's growth and areas requiring attention.

To ensure accurate interpretation, the dataset was supplemented by a data dictionary, which defines the various metrics and their measurement criteria. The dataset is publicly available under open data policies, allowing for transparency and accessibility, supporting projects aimed at improving India's port operations and fleet performance.

To ensure accurate interpretation, the data dictionary was also downloaded from the same source. The dataset is publicly available under open data policies, making it an ideal resource for educational projects like this one conducted under the guidance of [Dr. Baljinder Kaur] at Lovely Professional University.

* The exploratory data analysis (EDA) process followed a structured, multi-step approach in Microsoft Excel:
* Initial Cleaning: Removed empty rows and columns, standardized missing values, corrected typos, and ensured uniformity in date and region formats.
* Data Formatting: Large numeric values were reformatted using custom number styles to enhance readability (e.g., displaying 1,000,000 as "1,000K").
* Classification Logic: Ports were categorized into 'Major' and 'Non-Major' based on government classifications using IF formulas and conditional rules.
* Pivot Table Generation: Summarized key metrics such as cargo handled, berth productivity, GRT, and vessel counts using PivotTables categorized by year, coast, and port type.
* Visualizations: Designed diverse chart types—including Line, Bar, Column, Pie, Scatter, and Clustered Column charts—each aligned with one of the ten core analytical objectives.
* Interactive Elements: Introduced slicers to enable dynamic filtering based on coast or time period, making the dashboard user-friendly.
* Styling & Readability: Applied professional design principles such as gradient coloring, conditional formatting, and strategic font usage to boost readability and visual appeal.
* Dynamic & Named Ranges: Incorporated named ranges and formulas like OFFSET() to ensure charts auto-update with new data entries.
* Data Integrity Tools: Used Excel's data validation tools to restrict manual entry errors, ensuring long-term reliability of the dataset and analysis.

4. Analysis on Dataset (for Each Objective)

i. Introduction  
The analysis addresses ten core objectives to evaluate the performance of Indian Ports and Shipping.

ii. General Description  
The data includes year-wise cargo trends, port-level performance, ship fleet data by age and size, and regional efficiency indicators.

iii. Specific Requirements, Functions and Formulas

* Pivot Tables: Used to summarize port-wise and year-wise data
* Formulas: =SUM(), =AVERAGE(), Custom number formatting
* Filters and Slicers for interactivity

iv. Analysis Results & v. Visualization

Chart 1: Line Chart - Cargo Handled Trend (Objective 1)

* PivotTable: Rows = Year, Values = Total Cargo Handled
* Chart: Line Chart, Title: "Cargo Handled Trend (2010–2024)"
* Insight: Consistent increase in cargo, peaking post-2020

Chart 2: Pie Chart - Major vs. Non-Major Ports (Objective 2)

* PivotTable: Rows = Port Type, Values = Cargo Handled
* Chart: Pie Chart, Title: "Cargo Distribution: Major vs Non-Major Ports"
* Insight: Major Ports contribute over 70% of total cargo

Chart 3: Bar Chart - Top 10 Ports by Cargo (Objective 3)

* PivotTable: Sorted descending by cargo
* Chart: Horizontal Bar, Title: "Top 10 Performing Ports"
* Insight: JNPT and Paradip are top contributors

Chart 4: Clustered Column Chart - Age-Wise Fleet Distribution (Objective 4)

* PivotTable: Rows = Year, Columns = Age Group
* Chart: Clustered Column, Title: "Fleet Age Distribution Over Time"
* Insight: Older fleet (>20 yrs) gradually decreasing

Chart 5: Line Chart - Fleet Growth Trend (Objective 5)

* PivotTable: Year-wise total vessels
* Chart: Line Chart, Title: "Growth of Indian Fleet (1947–2024)"

Chart 6: Clustered Column Chart - Coastal vs. Overseas GRT (Objective 6)

* PivotTable: Year-wise GRT
* Chart: Clustered Column, Title: "Coastal vs Overseas GRT Trend"

Chart 7: Bar Chart - Total GRT in Indian Fleet (Objective 7)

* PivotTable: Sum of Total GRT by Year
* Chart: Bar Chart, Title: "Total GRT in Indian Fleet"

Chart 8: Column Chart - All Ports Performance (Objective 8)

* PivotTable: Ports vs Cargo Output
* Chart: Column, Title: "All Ports Cargo Handling Comparison"

Chart 9: Grouped Bar Chart - East vs West Coast Avg Output (Objective 9)

* Chart Title: "Avg Output per Berth: East vs West Coast"
* Insight: Western ports show higher berth productivity

Chart 10: Line Chart - TRT Trend East vs West Coast (Objective 10)

* Chart Title: "Turnaround Time Trend – East vs West"
* Insight: Turnaround time improving across both regions

iv. Visualization

Chart 1: Line Chart - Cargo Handled Trend (Objective 1)

* PivotTable: Rows = Year, Values = Total Cargo Handled
* Chart: Line Chart, Title: "Cargo Handled Trend (2010–2024)"
* Insight: Consistent increase in cargo, peaking post-2020

Chart 2: Pie Chart - Major vs. Non-Major Ports (Objective 2)

* PivotTable: Rows = Port Type, Values = Cargo Handled
* Chart: Pie Chart, Title: "Cargo Distribution: Major vs Non-Major Ports"
* Insight: Major Ports contribute over 70% of total cargo

Chart 3: Bar Chart - Top 10 Ports by Cargo (Objective 3)

* PivotTable: Sorted descending by cargo
* Chart: Horizontal Bar, Title: "Top 10 Performing Ports"
* Insight: JNPT and Paradip are top contributors

Chart 4: Clustered Column Chart - Age-Wise Fleet Distribution (Objective 4)

* PivotTable: Rows = Year, Columns = Age Group
* Chart: Clustered Column, Title: "Fleet Age Distribution Over Time"
* Insight: Older fleet (>20 yrs) gradually decreasing

Chart 5: Line Chart - Fleet Growth Trend (Objective 5)

* PivotTable: Year-wise total vessels
* Chart: Line Chart, Title: "Growth of Indian Fleet (1947–2024)"

Chart 6: Clustered Column Chart - Coastal vs. Overseas GRT (Objective 6)

* PivotTable: Year-wise GRT
* Chart: Clustered Column, Title: "Coastal vs Overseas GRT Trend"

Chart 7: Bar Chart - Total GRT in Indian Fleet (Objective 7)

* PivotTable: Sum of Total GRT by Year
* Chart: Bar Chart, Title: "Total GRT in Indian Fleet"

Chart 8: Column Chart - All Ports Performance (Objective 8)

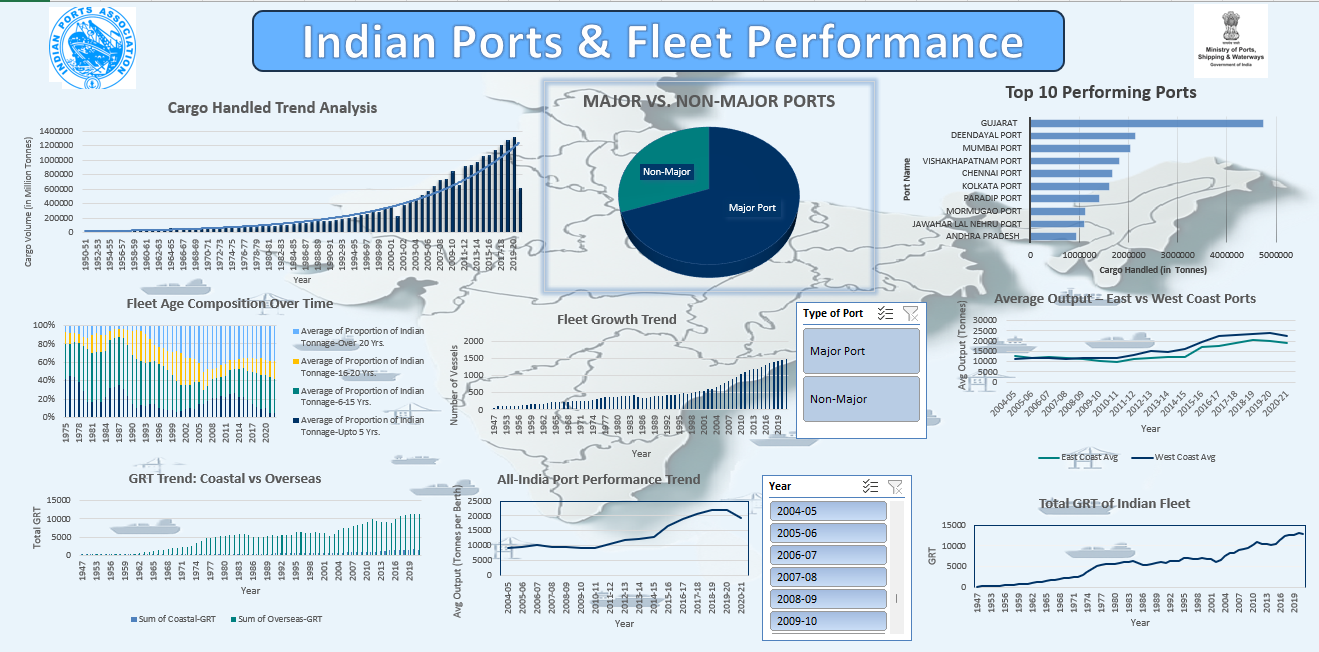
* PivotTable: Ports vs Cargo Output
* Chart: Column, Title: "All Ports Cargo Handling Comparison"

Chart 9: Grouped Bar Chart - East vs West Coast Avg Output (Objective 9)

* Chart Title: "Avg Output per Berth: East vs West Coast"
* Insight: Western ports show higher berth productivity

Chart 10: Line Chart - TRT Trend East vs West Coast (Objective 10)

* Chart Title: "Turnaround Time Trend – East vs West"
* Insight: Turnaround time improving across both regions



5. Conclusion

The dashboard successfully visualizes the performance metrics of Indian Ports. Insights such as rising cargo volumes, regional port disparities, and fleet modernization offer valuable input for strategic development. The visual approach improves data accessibility for policymakers and stakeholders.

6. Future Scope

* Dataset Expansion: Incorporate historical College Scorecard data to analyze trends over multiple years.
* Advanced Tools: Utilize Python with libraries like Matplotlib or Power BI for more sophisticated visualizations and real-time updates.
* Predictive Analysis: Implement machine learning models to predict completion rates or earnings based on admission rates and tuition costs.
* Web Deployment: Develop a web-based dashboard using HTML/CSS/JavaScript for wider accessibility.

7. References

* Ministry of Ports, Shipping and Waterways: <https://shipmin.gov.in/>
* Indian Ports Association: https://www.ipa.nic.in/
* Sagarmala Programme: <https://sagarmala.gov.in/>
* Data compiled and processed in Microsoft Excel, 2025
* Project guided by Dr. Baljinder Kaur, Lovely Professional University