Indian Railway Train Data Analysis Report

1. Objective

The goal of this project is to analyze Indian Railway train schedule data to uncover patterns, trends, and insights that can help optimize train operations and improve passenger service planning.

2. Dataset Overview

- File: Railway_info.csv
- Total Entries: 11,113 train records
- Columns:
- Train_No: Train number
- Train_Name: Name of the train
- Source_Station_Name: Starting station
- Destination_Station_Name: Ending station
- days: Day of operation

3. Data Preprocessing

Task 1.1: Load and Inspect

- Dataset loaded using pandas
- No missing values were found
- Data types: 1 integer (Train_No), 4 objects

Task 1.2: Basic Statistics

- Unique trains: 11,113
- Source stations: 921
- Destination stations: 924
- Most common station: CST-MUMBAI

Task 1.3: Cleaning

- Standardized all station names to uppercase

Indian Railway Train Data Analysis Report

- No null values were present

4. Data Analysis

Task 2.1: Filtering

- 1531 trains operate on Saturday

- 199 trains start from stations containing 'DELHI'

Task 2.2: Grouping & Aggregation

- Grouped by source station to get total and average trains per day

Added Column: Day Type

- Weekday trains: 7842

- Weekend trains: 3271

5. Visualizations

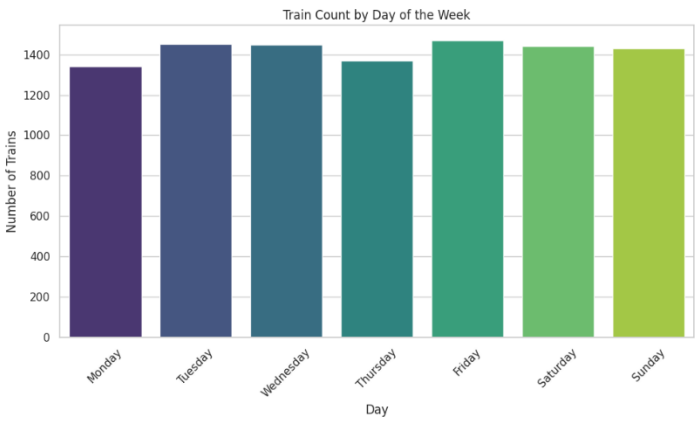
Visualizations were created using Matplotlib and Seaborn:

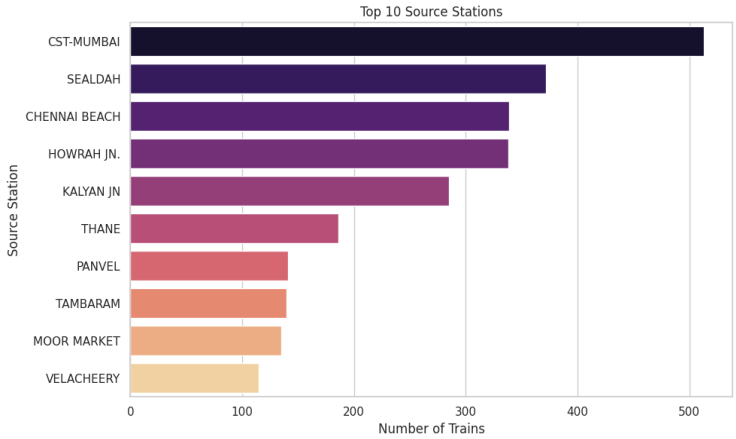
- Bar Chart: Trains per Day of the Week

- Bar Chart: Top 10 Source Stations

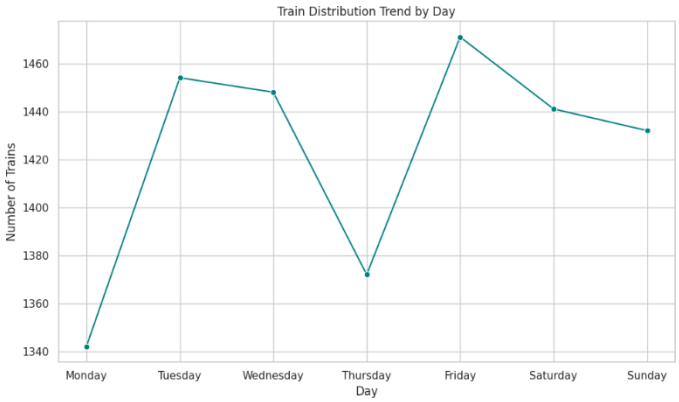
- Line Chart: Weekly Trend of Train Distribution

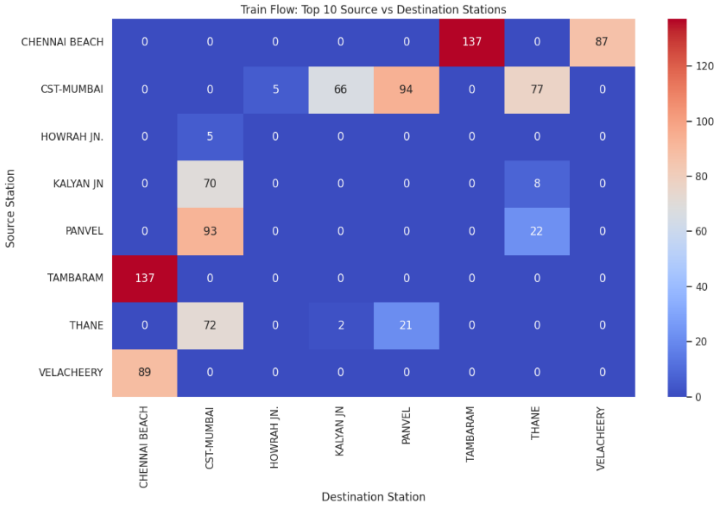
- Heatmap: Train Flow Between Top Source and Destination Stations





Indian Railway Train Data Analysis Report





6. Insights & Recommendations

Insights:

- Friday sees the highest number of trains
- CST-MUMBAI is the busiest source and destination- Weekdays have more train traffic compared to weekends

Recommendations:

- Increase weekend services on busy routes
- Optimize weekday schedules by evaluating underutilized trains
- Use station-wise data to improve local infrastructure and service planning

7. Conclusion

This analysis provides a foundational understanding of train scheduling patterns. The insights can help Indian Railways optimize service levels, improve passenger satisfaction, and plan better infrastructure allocation.

8. Appendix

- Python code (available on request)

Indian Railway Train Data Analysis Report

- CSV summary outputs

- Visualizations: Embedded in notebook or available as images