DEPI Final project

UK Railway Report

Data Analysis Project

By : Yousef Ahmed

Yomna Saad

Introduction

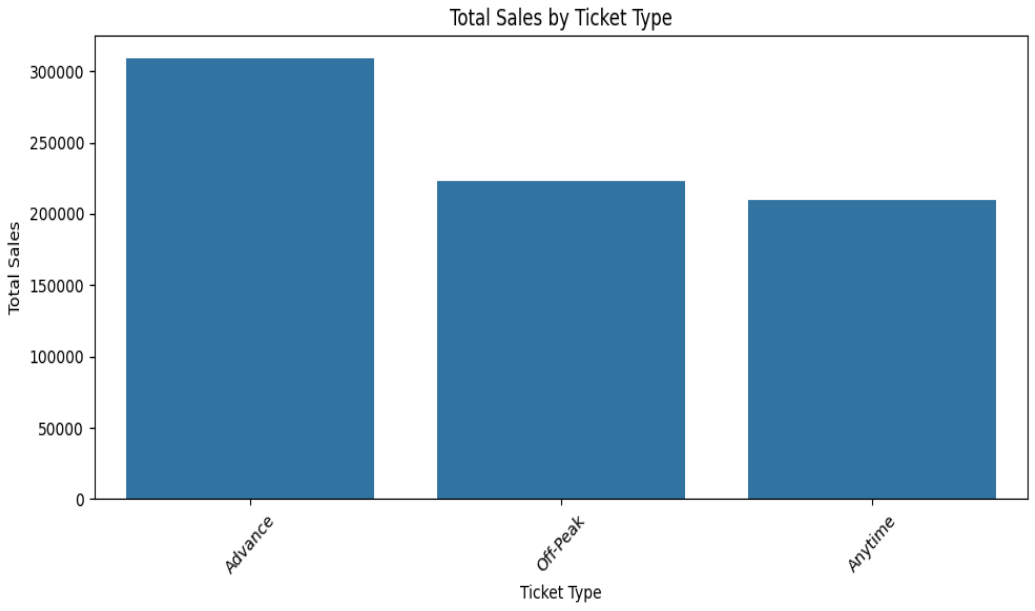
This report provides a comprehensive analysis of UK railway operations, focusing on ticket purchasing patterns, journey performance, and customer experiences. Utilizing a dataset of 31,653 transactions from December 2023 to April 2024, sourced from the "Railway_clean.csv" file, we explore key metrics such as ticket types, pricing, journey statuses, and delay patterns across major stations like London Paddington, Liverpool Lime Street, and Manchester Piccadilly. The dataset includes detailed information on purchase types (online vs. station), payment methods, railcard usage, and journey outcomes, offering a robust foundation for understanding operational efficiencies and customer behavior. Through data-driven insights, this report aims to identify trends, highlight areas for improvement, and provide actionable recommendations for enhancing railway services in the UK.

Data Visualization

1- Sales

1-1- Sales based on ticket type :

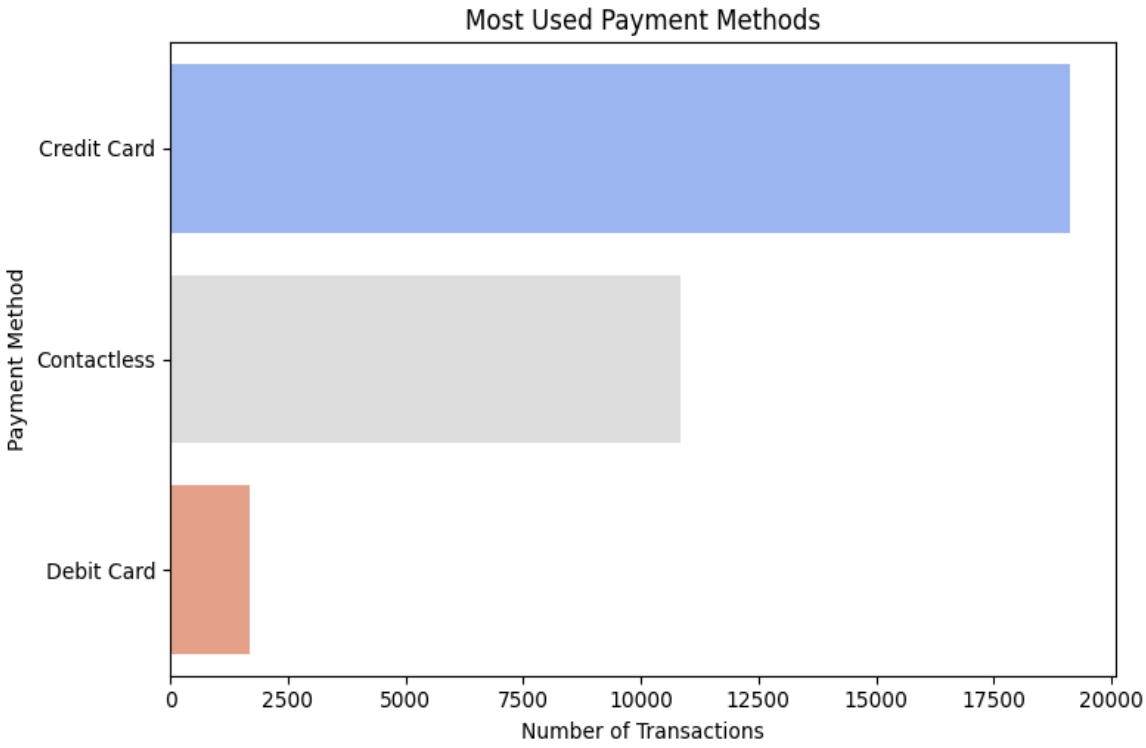
The railway dataset includes ticket types such as Advance and Off-Peak. The visualization below shows the distribution of ticket types purchased, highlighting passenger preferences.



As seen in the plot, Advance tickets are more common, likely due to their lower cost compared to Off-Peak tickets, which offer more flexibility.

1-2- Sales based on payment method :

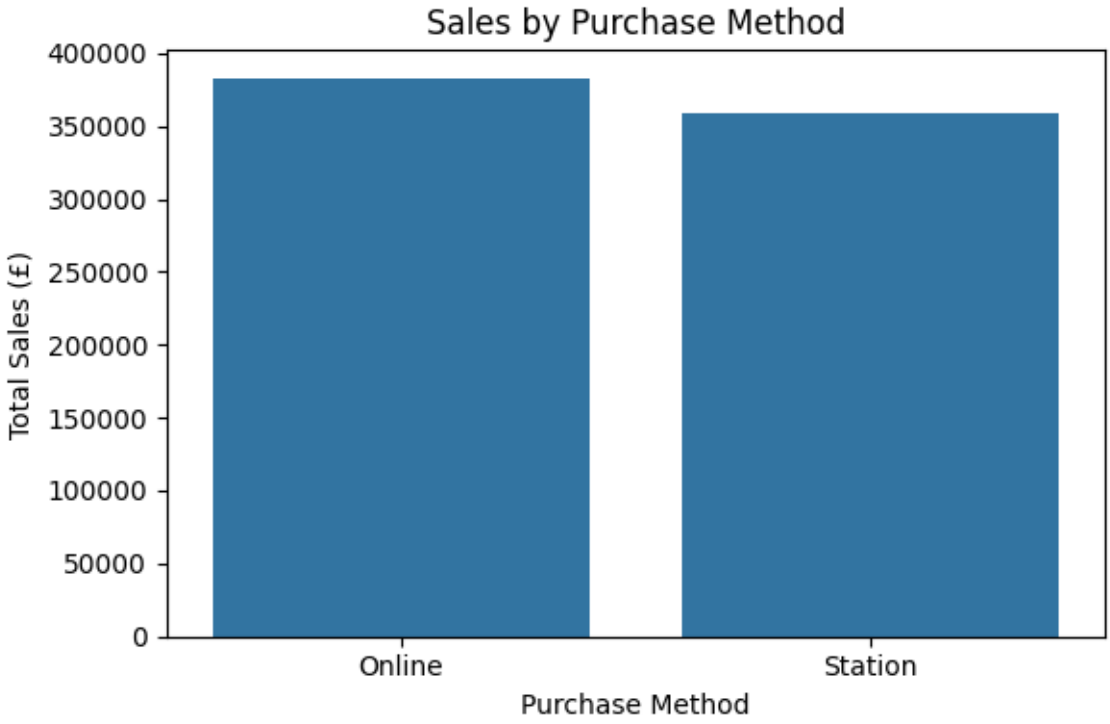
The dataset includes payment methods like Credit Card, Debit Card and Contactless. This visualization shows the usage of different payment methods across transactions.



Credit Card payments dominate, reflecting their widespread use, while Contactless payments are also popular, indicating a shift toward modern payment solutions.

1-2- Sales based on Purchase method :

Tickets can be purchased online or at the station. The following visualization illustrates the proportion of purchases made through each method.

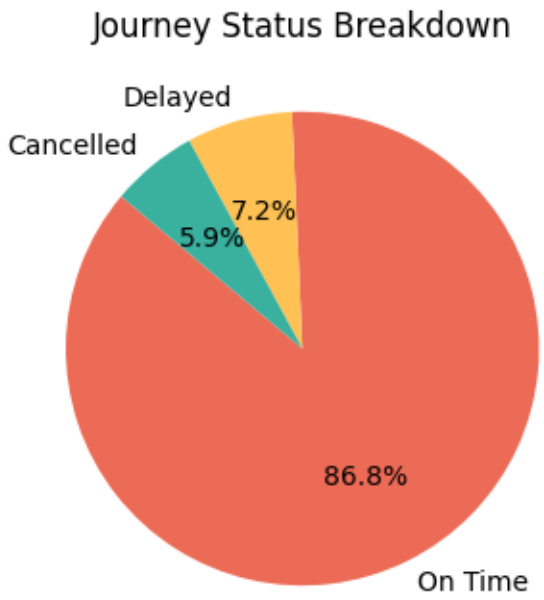


The data reveals a slight preference for online purchases, possibly due to convenience and accessibility, with station purchases still notable for last-minute travelers.

2- Journey performance

2.1. Journey Status Analysis

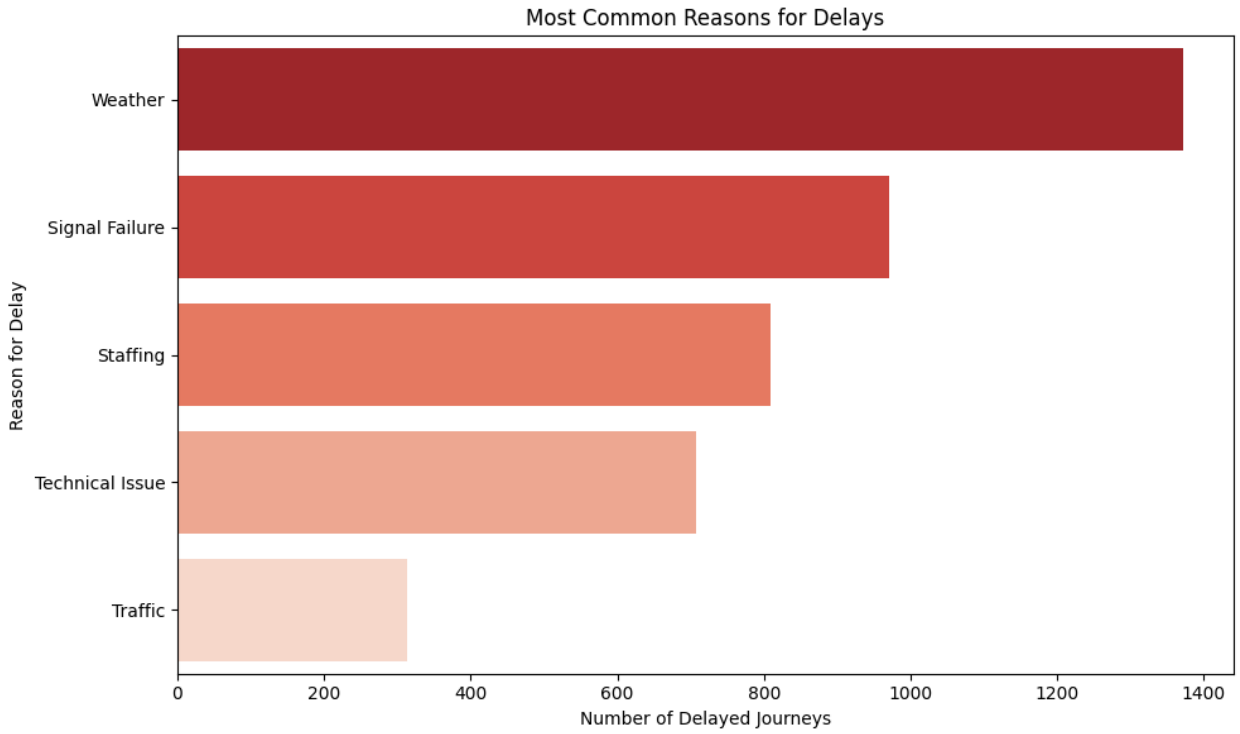
Journey status indicates whether trains were on time or delayed. The visualization below shows the proportion of on-time versus delayed journeys.



Most journeys are on time, which is encouraging for railway reliability, though delays still occur and warrant further investigation.

2.2. Reasons for Delays

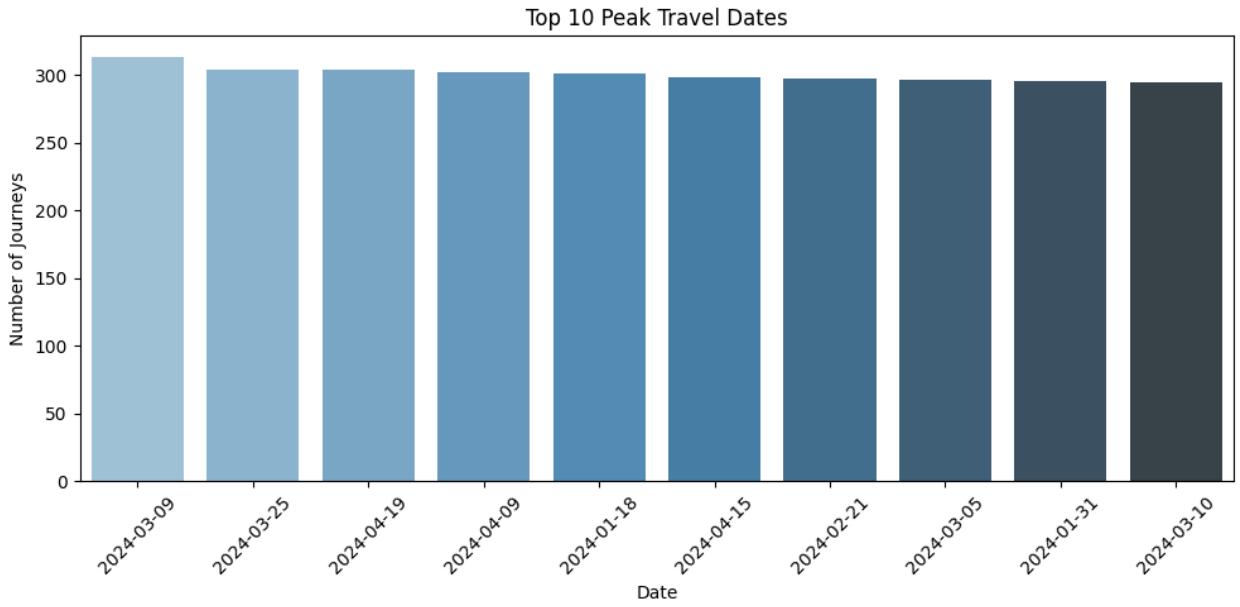
For delayed journeys, the dataset provides reasons such as Signal Failure. The following plot examines the frequency of different delay reasons.



Signal Failure appears as a common cause, suggesting areas for operational improvement to enhance service reliability.

2.3. Peak Travel Dates

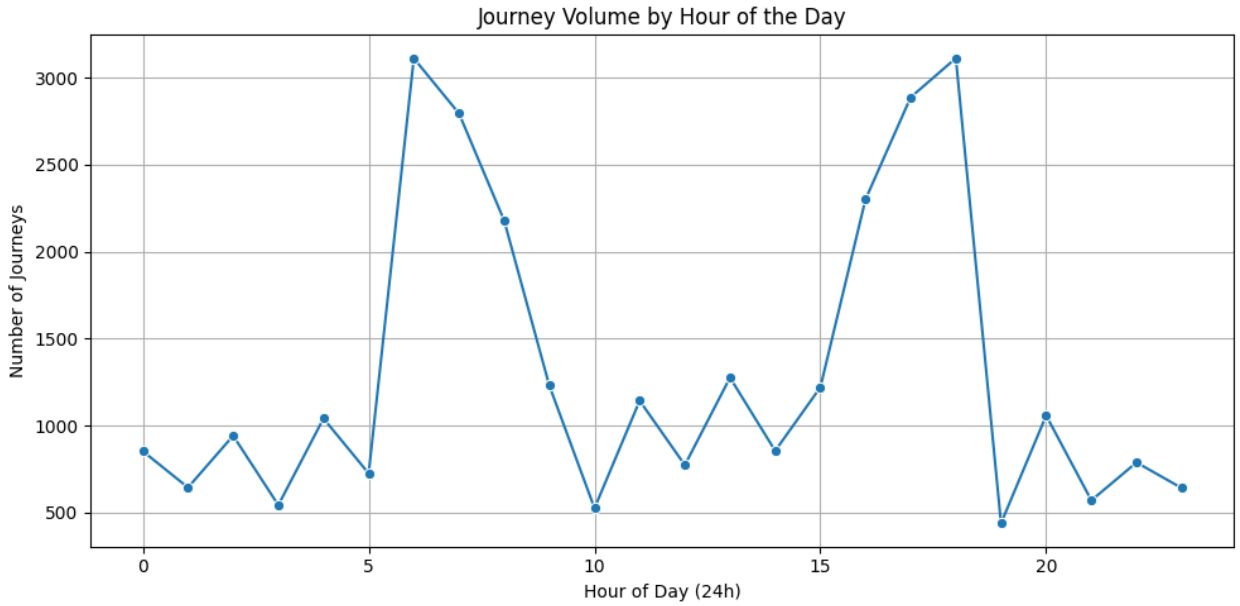
Understanding peak travel dates helps identify periods of high demand. The visualization below shows the top 10 dates with the highest number of journeys.



The plot highlights specific dates with unusually high travel volumes, such as early March and April, which may correspond to holidays or events driving increased railway usage.

2.4. Journey Volume by Hour of the Day

Understanding the distribution of journeys throughout the day can help identify peak travel hours. The visualization below shows the number of journeys by hour of the day (in a 24-hour format).



The plot reveals two distinct peaks: one around 6–8 AM and another around 5–6 PM, likely corresponding to morning and evening commutes. Journey volume is lowest in the early morning (around 3–4 AM), which is expected as fewer people travel at that time. This insight can help railway operators optimize schedules and manage capacity during peak hours.

3. Recommendations

Based on the analysis, we propose the following strategic recommendations:

1. **Enhance Signal Infrastructure**
    - Invest in modern signal systems to reduce the frequency of delays caused by signal failures.

- Implement predictive maintenance systems to anticipate and prevent faults.

2. **Promote Advance Booking Incentives**
   - Encourage passengers to buy Advance tickets through targeted discounts, increasing predictability in demand and load balancing.
3. **Optimize Staff and Train Allocation for Peak Hours**
   - Use hourly journey data to reallocate staff and train schedules, especially during morning and evening rush hours.
4. **Expand Digital Payment and Booking Options**
   - Enhance user experience on mobile apps and websites to further drive online and contactless transactions.
   - Introduce loyalty programs for frequent users to increase digital engagement.
5. **Plan Around Peak Travel Dates**
   - Prepare for demand spikes during holidays and special events by increasing service frequency and customer support capacity.
6. **Enhance Real-Time Communication**
   - Improve real-time updates and delay notifications to keep passengers informed and improve satisfaction during service disruptions.

---

4. Conclusion

This report presents actionable insights into UK railway operations. By addressing delay causes, enhancing customer experience, and optimizing resource allocation, railway services can significantly improve both operational efficiency and passenger satisfaction. The findings serve as a valuable guide for policy makers, transport planners, and railway operators.

Thank you