LONDON RAILWAY ANALYSIS

IDENTIFYING HOW TO IMPROVE TICKET SALES AND PASSENGER EXPERIENCE

Emenike Kenechukwu A.

Data Analytics with Excel

4th April, 2025

ABSTRACT

This project presents an in-depth analysis of railway service transactions in London, focusing on operational performance, passenger behavior, and revenue optimization. Using a structured dataset containing detailed ticket purchases, journey timelines, payment methods, and passenger profiles, the analysis aims to uncover patterns in train punctuality, identify high-performing and delay-prone routes, assess the impact of railcards and ticket types on pricing, and evaluate refund behaviors. The insights derived provide actionable recommendations for improving customer experience, enhancing operational efficiency, and optimizing pricing strategies. Overall, the project offers data-driven solutions to support better decision-making within the railway transport sector.

INTRODUCTION

The London railway network serves as a critical transportation system for millions of passengers daily, facilitating seamless travel across the city and connecting major hubs. This project aims to analyze a comprehensive dataset of railway service transactions, with a focus on ticket sales, journey details, passenger behavior, and train performance metrics.

The dataset is a mock representation of National Rail ticket transactions in the UK, spanning from January to April 2024. It includes detailed attributes such as:

- Ticket type and class
- Date and time of purchase and journey
- Departure and arrival stations
- Ticket prices
- Payment methods and purchase channels
- Railcard usage
- Journey status (on-time, delayed, or canceled)
- Refund requests and reasons for delays

Scope of the Dataset

This rich dataset enables a multifaceted analysis that supports insights and includes key transactional and operational data, such as:

- **Transaction Details** – Unique ticket purchases, dates, and payment methods.
- **Journey Information** – Departure and arrival stations within London.
- **Train Performance** – Scheduled vs. actual arrival times, delays, and cancellations.
- **Passenger Preferences** – Ticket class, purchase type, and discount eligibility.

Objectives of the Analysis

By leveraging this dataset, the analysis aims to:

- **Assess Train Punctuality** – Identify delays and cancellation trends.
- **Analyze Passenger Behavior** – Understand ticket sales patterns and purchase preferences.
- **Evaluate Revenue & Pricing Strategies** – Examine the impact of ticket types and rail cards.
- **Enhance Customer Experience** – Investigate refund requests and service reliability.
- **Support Operational Improvements** – Provide insights for optimizing train schedules and resource allocation.

Expected Impact

The insights from this analysis will be valuable for railway operators, policymakers, and commuters, helping to improve efficiency, passenger satisfaction, and revenue optimization within London's railway network.

DATA COLLECTION & PREPARATION TECHNIQUE

The dataset used in this analysis was obtained through a **third-party source** from a trusted colleague. While the original source is not explicitly documented, the dataset appears to be structured and representative of railway service transactions in London. To ensure reliability, a preliminary assessment was conducted to validate its completeness and consistency.

After acquiring the dataset, it was imported into my workbook, which contained two worksheets:

- Data Metadata Sheet – This sheet provided a structured overview of the dataset, including column descriptions, data types, and definitions to ensure a clear understanding of the information.
- Dataset Sheet – The main dataset contained raw transactional and journey-related data, including ticket purchases, travel details, and train performance metrics.

Initial Data Assessment
To ensure the dataset was ready for analysis, the following preliminary checks were conducted:

- Data Completeness Check – Verified that no critical fields were missing.
- Data Consistency Review – Ensured uniform formatting across dates, times, and categorical values.
- Duplicate Detection – Checked for repeated transaction records that could affect accuracy using power query.
- Handling Missing Values – Assessed gaps in data and determined strategies for imputation or exclusion. Because in the reason for delay column, if the journey status is On Time, then there is no reason there. So, I had to fill in "No Delay".

DATA ANALYSIS FINDINGS

Based on my objectives of this analysis:

1. Assessing Train Punctuality: This objective was achieved by examining the percentage of trains that arrived on time, were delayed, or were canceled, along with identifying the top reasons for delays and average delay durations across routes.

    - *While delay durations vary across routes, the overall average delay across all journeys is 3.1 minutes.*
    - *The route with the highest cumulative delay is Manchester Piccadilly – Leeds, recording 64.8 minutes of delay, followed by London Euston – York with 36.4 minutes.*

On-Time Performance

	Cancelled	Delayed	On Time
■ Total	5.94%	7.24%	86.82%

Reason for Delay

	Signal Failure	Staff Shortage	Staffing	Technical Issue	Traffic	Weather	Weather Condi...
■ Total	970	399	410	707	314	995	377
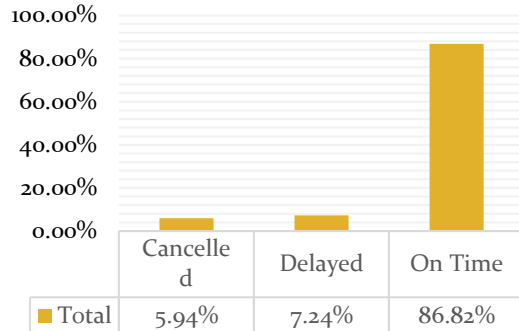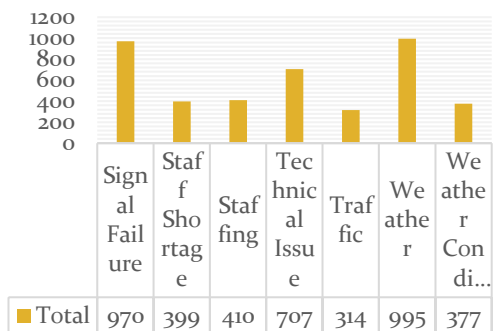
2. Analyzing Passenger Behavior: This involves understanding patterns in ticket sales (e.g., ticket types and class) and identifying preferred payment and purchase methods.

    - *Less than 10% of passengers travel in First Class, indicating a clear preference for Standard Class.*

- *Advance Tickets account for 55% of total ticket purchases, suggesting that passengers favor discounted, pre-booked travel.*
- *Over 60% of transactions were made using credit cards, highlighting a strong preference for digital payment methods.*

Ticket Class

9.66%

90.34%

■ First Class   ■ Standard

Ticket Type

28%

17%

55%

■ Advance   ■ Anytime   ■ Off-Peak

Payment Type

5%

34%

61%

■ Contactless   ■ Credit Card   ■ Debit Card
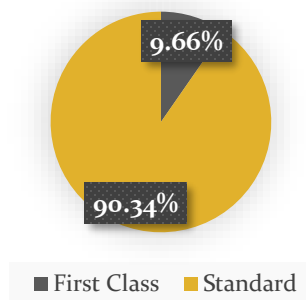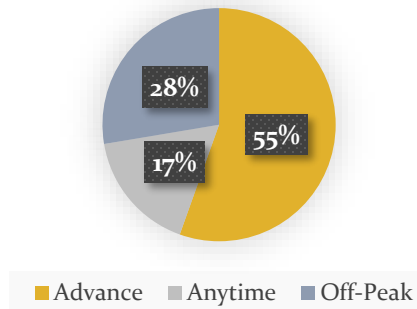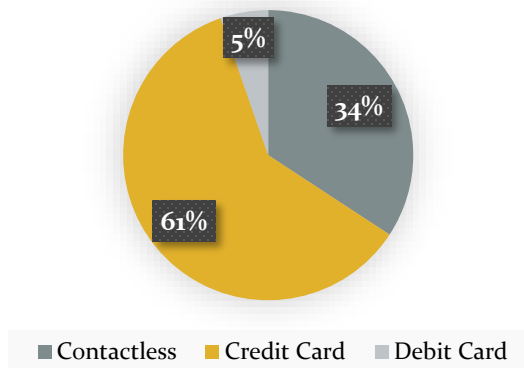
3. Evaluating Revenue & Pricing Strategies: This was approached by assessing the impact of rail cards on ticket prices and overall revenue.
   - *The analysis includes the total ticket prices and the average percentage difference in price between passengers with and without railcards.*
   - *This comparison highlights the discounts offered to various railcard holders (e.g., Adult, Senior, Disabled), demonstrating how railcards influence pricing and accessibility.*

Railcards Impact Discount
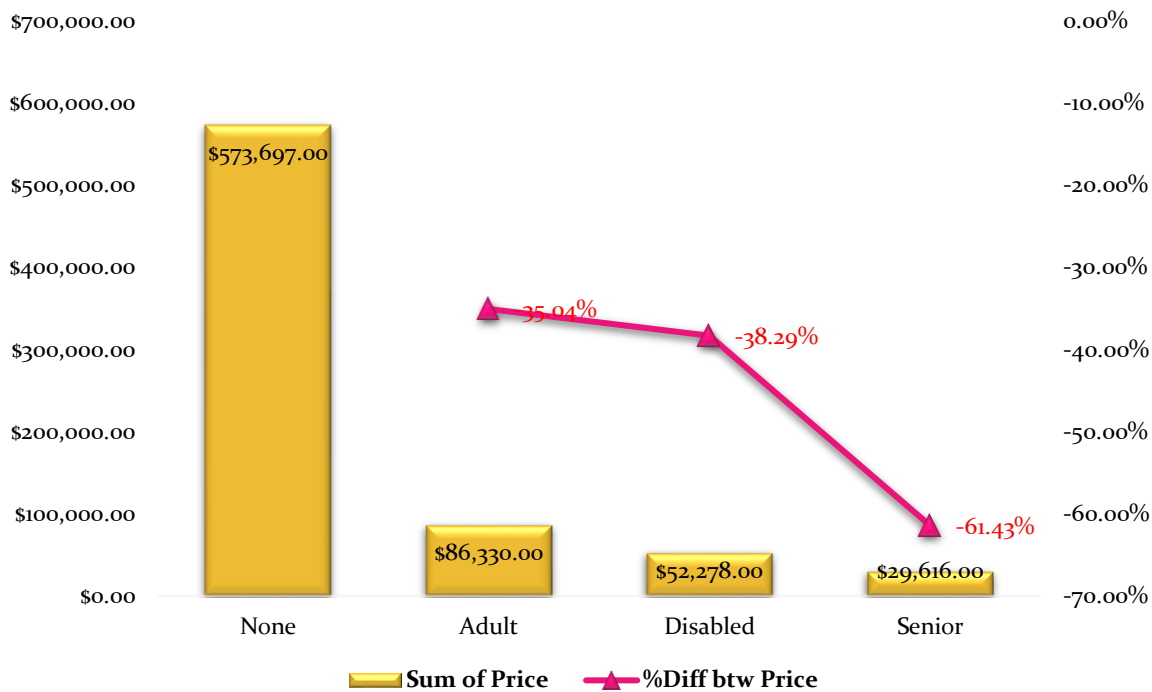
4. Enhance Customer Experience: To improve customer experience, refund patterns were analyzed—particularly after delays and cancellations—and routes with the highest refund requests were identified.
   - *Less than 31% of passengers on delayed or canceled trains requested a refund, meaning over 60% of affected travelers did not pursue compensation.*
   - *The Liverpool Lime Street – London Euston route had the highest number of refund requests and also recorded the most frequent delays in the dataset.*

Refund Rate

No ■ Yes ■
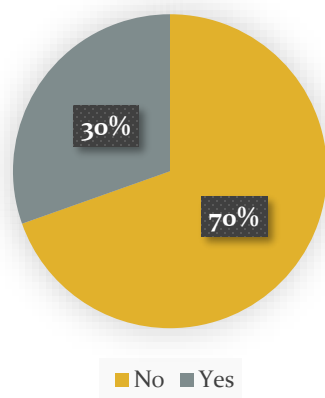
5. Operational Efficiency & Resource Allocation: To increase operational efficiency, we have to identify the stations with the highest passenger's volume, train routes with highest delay/cancellations and reason for the delay.

   - *The busiest route was Manchester Piccadilly – Liverpool Lime Street, with 4,628 individual transactions over four months.*
   - *This was followed by London Euston – Birmingham New Street, with 4,209 transactions.*
   - *Once again, the Liverpool Lime Street – London Euston route had the most frequent delays, with signal failure being the primary cause.*

4628   345

Count of Transaction ID
Value: 4628
Row: Manchester Piccadilly
Column: Liverpool Lime Street
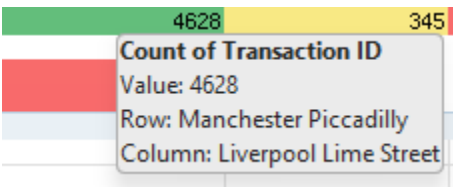
RECOMMENDATION

Below are my recommendations from the analysis

1. Improve Delay-Prone Routes: Focus on infrastructure enhancements to reduce delays on high-risk routes such as Manchester Piccadilly – Leeds and Liverpool Lime Street – London Euston. This includes addressing root causes like signal failures and implementing more reliable signaling systems. Additionally, constructing or upgrading weather-proof structures at key stations and platforms can help minimize the impact of adverse weather, ensuring smoother boarding and transit experiences for passengers during harsh conditions.

2. Boost Refund Awareness: Educate passengers on how to claim refunds efficiently after delays or cancellations to build trust and improve satisfaction. Many passengers may be unaware of their eligibility or find the process unclear, which results in a large percentage of unclaimed refunds. This lack of action could lead to frustration and the perception that the railway service is unresponsive or unfair, ultimately damaging customer loyalty.

3. Optimize Resource Allocation: Allocate more staff and support resources to high-traffic routes such as Manchester Piccadilly – Liverpool Lime Street. This ensures smoother operations, faster boarding times, and improved passenger assistance during peak hours or service disruptions.

4. Dynamic Pricing Strategy: Reevaluate ticket pricing and railcard discount policies to strike a balance between affordability and revenue. Promoting strategic use of railcards can help drive customer retention while maintaining financial sustainability.

5. Encourage Digital Payments: With over 60% of passengers opting for credit card payments, there's a clear preference for digital transactions. Enhancing the digital payment experience by making platforms more secure, faster, and user-

friendly will improve customer satisfaction. Additionally, offering incentives for using contactless or online payment methods can encourage more passengers to adopt them. This shift would also help reduce congestion at ticket sale stands, especially during peak hours, resulting in a more efficient and less stressful station environment.

6. First-Class Marketing Opportunities: Given that fewer than 10% of passengers choose First Class, there's room to attract more premium travelers. Targeted marketing campaigns, loyalty perks, and promotional offers could boost First Class usage and increase overall revenue.

CONCLUSION

In conclusion, the data revealed key insights into the operational performance and customer behavior within the London railway service. The average delay across all routes is relatively low at 3.1 minutes, but some routes experience significant delays, highlighting areas needing attention. Most passengers prefer Advance tickets and Standard class, with Credit Card being the dominant payment method. Despite many delayed or canceled trains, only a small portion of passengers request refunds, potentially due to lack of awareness or inconvenience.

The data also revealed specific routes that experience higher traffic and operational issues, such as the Liverpool Lime Street – London Euston line, which had both the highest refund rate and delay count. These insights can inform more efficient resource allocation and help improve customer satisfaction and service delivery.