UK Train Rides Problem Definition and Research Questions

- Project Title: UK Train Rides

- Team Members: Mohamed Alaa, Abd El Rahman Maher, Ziad Mohamed, Aya Anwar , Amira Atef , Seif Marzouk

- Supervisor: Eng. Sherihan Ali

- Date: 12/02/2025

1-Introduction

The UK railway is essential for daily travel, linking cities and regions. But it faces challenges like changing passenger numbers, uneven income, and operational issues. To improve, it's important to understand when people travel, how they buy tickets, and how delays affect them. This helps optimize schedules, pricing, and the overall experience.

This project uses data to analyze ticket sales, travel patterns, and delays. The goal is to provide insights that help rail operators make better decisions and create a more reliable, passenger-friendly service. It's about making train travel smoother for everyone.

2-Business Problem

The UK train network faces multiple operational and financial challenges that impact both revenue generation and passenger experience. Some of the key issues include:

- **Unpredictable passenger demand, making capacity planning inefficient.**
- **Revenue fluctuations due to pricing variations and ticket class preferences.**
- **Frequent delays leading to refund requests and customer dissatisfaction.**
- **Lack of data-driven strategies to optimize scheduling and ticket pricing.**

By analyzing historical ride data, ticket sales, and journey patterns, we aim to uncover key insights that can help improve pricing strategies, forecast ridership trends, enhance service efficiency, and minimize revenue loss due to delays

3. Research Questions

To address the identified business challenges, the project focuses on answering the following key research questions:

1- Revenue & Ticket Sales Trends

- How do ticket prices, ticket types (Advance, Off-Peak, Anytime), and ticket classes (Standard, First Class) impact overall revenue?

- What is the relationship between time of purchase (peak vs. off-peak hours) and revenue generation?

- How do railcards (Adult, Disabled, Senior) affect revenue, and which railcard type contributes the most?

2- Journey Delays & Customer Satisfaction

- What are the most common reasons for delays, and how do they affect customer behavior (e.g., refund requests)?

- What is the financial impact of delayed or cancelled journeys on revenue?

- How does on-time performance impact customer satisfaction and repeat purchases?

3-Ticket Demand & Route Profitability

- Which routes and stations generate the highest revenue?

- Are there underperforming routes or stations that could be optimized for better profitability?

- How does customer segmentation (e.g., railcard holders, ticket types) affect ticket sales trends?

4- Pricing & Discount Strategies

- How do discounts (e.g., Advance ticket pricing) impact revenue and customer behavior?

- What is the optimal pricing strategy for different ticket types to maximize revenue?

- How does peak vs. off-peak pricing affect sales and demand?

5- Refunds & Customer Loyalty

- How do refund requests correlate with customer retention or repeat purchases?

- What are the most common reasons for refund requests, and how can they be reduced?

6- Operational Efficiency & Cost Reduction

- What are the most common causes of operational inefficiencies (e.g., delays, cancellations), and how can they be addressed?

- How can route optimization or scheduling improvements reduce costs and improve customer satisfaction?

4. Objectives

- Identify key drivers of revenue and ticket sales trends.

- Analyze the impact of ticket pricing, discounts, and purchase timing on revenue.

- Examine the relationship between journey delays and customer satisfaction.

- Assess the profitability of different routes and stations.

- Segment customers based on railcards, ticket types, and purchasing behavior.

- Evaluate the effect of refunds on customer loyalty and retention.

- Optimize operational efficiency by reducing delays and cancellations.

- Develop strategies to enhance cost reduction through route optimization.

5. Scope & Limitations

Scope:

This analysis will be conducted using historical UK train ride data, including information on refund requests, reasons for delays, journey status, actual vs. scheduled arrival times, departure details, ticket pricing, payment methods, and other transactional details.

The primary focus will be on:-

- Descriptive Analytics: Identifying trends in train delays, ticket purchases, passenger preferences, and refund requests.
- Diagnostic Analytics: Analyzing reasons for delays and refund patterns to understand their impact on customer experience and operational efficiency.

Limitations:

- External Factors Excluded: The study will not consider external factors such as weather conditions, strikes, maintenance issues, or government policies that may influence train delays and passenger behavior.
- No Predictive Modeling: The analysis will not include forecasting of future delays, passenger behavior, or sales trends.
- Limited Customer Insights: Since the dataset does not include detailed passenger demographics or feedback, customer behavior analysis will be limited to available ticketing and transaction data.
- Scope of Refund Analysis: Refund requests will be analyzed based on available reasons and transaction records but will not include customer service interactions or complaint resolution details.
- Real-Time Monitoring Not Included: This study focuses on historical data and will not involve real-time tracking of train journeys or ticket sales.

6. Key Performance Indicators (KPIs)

To measure the success of the analysis and track improvements, the following KPIs will be used:

Key Performance Indicators (KPIs) for Railway Operations

1. Sales & Revenue KPIs

Total Revenue ($)

- The total sum of all ticket sales revenue.
- Equation: Total Revenue = Sum of all Ticket Prices

Average Ticket Price ($)

- The mean price of all purchased tickets.
- Equation: Average Ticket Price = (Sum of all Ticket Prices) / (Total Number of Tickets Sold)

Sales Distribution by Purchase Type (%)

- The proportion of sales made online versus at the station.
- Equation: Sales Distribution = (Number of Transactions per Type) / (Total Transactions) * 100

Payment Method Distribution (%)

- The percentage breakdown of sales by payment method.
- Equation: Payment Method Distribution = (Number of Transactions per Payment Method) / (Total Transactions) * 100

Sales per Ticket Class

- The total revenue generated by each ticket class.
- Equation: Sales per Class = Sum of all Ticket Prices per Class

2. Operational Efficiency KPIs

On-time Performance (%)

- The percentage of journeys that arrived on time.
- Equation: On-time Performance = (Number of On-time Journeys) / (Total Journeys) * 100

Average Delay Duration (mins)

- The average delay time for all delayed journeys.
- Equation: Average Delay = (Sum of (Actual Arrival Time - Scheduled Arrival Time) for Delayed Journeys) / (Total Delayed Journeys)

Total Delayed Journeys

- The count of trips that arrived later than scheduled.
- Equation: Total Delayed Journeys = Count of Journeys with Status 'Delayed'

Top Reasons for Delay

- The most frequent causes of delays are based on recorded reasons.
- Equation: Top Reasons for Delay = Mode (Most Common) of 'Reason for Delay' Column

Refund Requests Rate (%)

- The percentage of trips for which a refund was requested.
- Equation: Refund Requests Rate = (Number of Refund Requests) / (Total Transactions) * 100

3. Passenger Insights KPIs

Total Number of Tickets Sold

- The total count of all transactions.
- Equation: Total Tickets Sold = Sum of all Transactions

Most Popular Routes

- The routes with the highest number of tickets sold.
- Equation: Most Popular Routes = Top N Routes based on Ticket Count

Ticket Sales Trend Over Time

- The trend of ticket sales based on purchase date.
- Equation: Ticket Sales Trend = Transactions Grouped by Date of Purchase

Railcard Usage (%)

- The percentage of passengers using a railcard.
- Equation: Railcard Usage = (Number of Transactions with Railcard) / (Total Transactions) * 100

Peak Booking Hours

- The most common time slots are when tickets are purchased.
- Equation: Peak Booking Hours = Top N Hours with Highest Ticket Sales

7. Expected Outcomes

Clear insights into passenger behavior and ticket purchasing patterns

- Identify peak purchase times, most preferred ticket types, and class demand distribution.
- Understand how rail card usage affects ticket sales and revenue.

Recommendations for pricing and discount strategies

- Optimize pricing structures based on demand trends for standard and first-class tickets.
- Assess the impact of rail cards and discounts on overall revenue and customer retention.

Optimized capacity planning for better resource allocation

- Forecast demand for different stations, routes, and time periods to improve seat availability.
- Adjust train schedules and station resources based on demand trends.

Improved delay and refund management strategies

- Identify key reasons for train delays and their correlation with refund requests.
- Reduce refund claims by addressing frequent delay issues and optimizing policies.

Enhanced customer segmentation for targeted marketing

- Classify passengers based on ticket class, travel frequency, and rail card type.
- Develop targeted marketing campaigns to increase first-class sales and rail card adoption.

Data-driven insights for operational efficiency improvements

- Improve ticket inventory management by analyzing demand fluctuations.
- Optimize staffing and service allocation at key stations based on passenger flow data