UK Train Rides Dashboard

I'll analyze this railway ticket sales dataset to extract meaningful insights that can be transformed into an impactful Power BI dashboard. Let's approach this systematically.

**## 1. Understanding the Dataset Structure**

The dataset contains 31653 Rows transactions and 18 Columns with the following fields:

- Transaction ID (31653)

- Date/Time of Purchase (12/8/2023) – (4/30/2024)

- Purchase Type ( Online = 18521 (85%) / Station = 13132 (41%) )

- Payment Method (Credit Card (19136) , Contactless (10834), Debit Card(1683)).

- Railcard type (Adult (4846), Disabled (3089), Senior (2800), None (20918))

- Ticket Class (Standard (28595) /First Class (3058))

- Ticket Type (Advance (17561)/Off-Peak (8752)/Anytime (5340) ) .

( 1/2 off a day prior to departure, 1/4 off used outside of peak hours (weekdays between 6-8am and 4-6pm), full price any time during the day).

- Price (Max = 267$) (Min = 1$)

- Departure (Max = Manchester Piccadilly – Min = Bristol Temple Meads) /Arrival Stations (Max = Birmingham New Street – Min = Warrington)

- Journey Date/Time

- Arrival Time

- Actual Arrival Time

- Journey Status (On Time (27481) / Delayed (2292) / Cancelled (1880))

- Reason for Delay (when applicable) (7 Main Reasons + 1 N/A)

- Refund Request (Yes ( 1118 ) / No (30535))

**## 2. Data Cleaning**

**1-Change column types**:

* Date of Purchase, Time of Purchase → Date/Time
* Date of Journey → Date
* Departure Time, Arrival Time, Actual Arrival Time → Time

**2-Add Columns and Calculating**:

* Journey Duration/hr and then change type
* Delay Duration.minute and then change type
* Route [Departure Station] → [Arrival Destination] and then change type

**2-Extracting**:

* Hour of Purchase and renamed column
* Day of purchase and renamed column

**3-Creating Categories**:

* Purchase Time Category and change type . Categorize Purchase Time /hr Column (>=5,<12 Morning/>=12,<17 Afternoon/>=17,<21 Evening)
* Ticket type Category and change type. Categorize Purchase Time /hr Column (>= 6, <= 8 Morning Peak/>= 16, <= 18 Evening Peak/ else "Off-Peak")
* Price Category and change type. Categorize Price Column (< 10 "Budget (<£10)"/ >= 10, < 50 "Standard (£10-50)"/ else "Premium (£50+)")

**4-Handling Missing Values**:

* Reason for Delay → Replace null values with N/A

**Measures for Key Metrics**

1. Total\_Transactions = 154.12k
2. Total Revenue = $154.12k
3. Total Reasons of Delay Count without N/A Values= 953
4. Refund Requests count (Yeas)= 244
5. Refund Rate = 3.33%
6. On Time Trips = 6384
7. On Time Rate = 87.01%
8. Delayed Trips = 537
9. Delayed Rate = 7.32%
10. Delay Reason Count = 537
11. Cancelled Trips = 416
12. Cancelled Rate =5.67%
13. Cancelled Reasons Count= 416
14. AVG Ticket Price= $21.01
15. Average Delay= 2.96 min

**\*\*Page 1: Summary\*\***

- KPI cards: Total Revenue, On-Time Performance, Refund Rate, Average Delay

- Line chart Total Revenue by Purchase /hr

- Map visualization of routes with ticket type bubbles

- Slicer All Departure Station

- Slicer All Years

- Slicer All Month Name

- Slicer Journey Station

- Slicer Day Name

**\*\*Page 2: Revenue Overview\*\***

- Slicer Total revenue = $154.12k

- Slicer Average ticket price: $21.01

- Line chart Total revenue by Price Category

- Area chart Total Transactions by Price

**\*\*Page 3: Tickets Revenue \*\***

* Pie chart Total Revenue by Ticket Class
* Stacked area chart Total Revenue by Purchase Type
* Clustered Bar Chart Total Revenue by Ticket Type
* Line clustered Revenue of Cancelled Trips and Delayed Trips by Reasons for Delay and cancelled
* Line chart Total Revenue by Ticket Type Category and Time Category
* Donut Total Revenue by Journey Status

**\*\*Page 4: Routes Revenue\*\***

* **Line chart Total Revenue and Count of Ticket Class by Departure Station**
* **Stacked chart Top 10 Route by Total Revenue**

**\*\*Page 5: Customer Insights\*\***

* **Tree map Total Transactions by Purchase Type**
* **Line chart Total transactions by PurchaseTime/hr**
* **Funnel Total Transactions by Payment Method**
* **Clustered chart Total Transactions by Railcard**

**\*\*Page 6: Reasons of Delayed Insights\*\***

* **Gauge Delay Reason Count**
* **Gauge Total Reason Count**
* **Gauge Cancelled Reason Count**
* **Line chart Total Transactions & AVG Delay.min by Reason for Delay**
* **Line clustered chart Cancelled Trips and Delayed Trips by Reasons for Delay and cancelled**

**\*\*Page 7: Operational Metrics \*\***

* **Gauge On Time %**
* **Gauge Delayed Rate %**
* **Gauge Cancelled Rate %**
* **Line Column Delayed Trips by Time Category**
* **Area Chart Cancelled Trips by Departure Time**
* **Clustered bar chart Cancelled Trips by Route**

**\*\*Page 8: Refund Insights \*\***

* **Slicer Refund Requests = 244 and Refund Rate = 3.33%**
* **Line Clustered Total Refund Request & Total Revenue by Journey Status**
* **Donut chart Total Refund Request by Journey Status**
* **Line Clustered Refund Requests by Reason for Delay and Journey Status**