

A high-speed train, likely a TGV, is stopped at a train station platform. The train has a distinctive red and white livery. The background shows a European town with buildings and hills. A large, semi-transparent white box covers the middle portion of the image, containing the text.

UK Train Rids

First Quarter 2024 Operational Review

TABLE OF CONTENTS

1

**DATA
COLLECTION**

2

**Explore
Data
with Python**

3

**DATA
CLEANING**

4

**Data
visualization**

TEAM MEMBERS

AHMED TAHA

AHMED SHERIF

ZIAD MOHAMED

REEM KHALED

ASEM KHALED

YOUSSEF ELBEDAWY

DATA COLLECTION



1

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Transaction ID	Date of Purchase	Time of Purchase	Purchase Type	Payment Method	Railcard	Ticket Class	Ticket Type	Price	Departure Station	Arrival Destination	Date of Journey	Departure Time	Arrival Time	Actual Arrival Time	Journey Status	Reason for Delay	Refund Requested
da8e6ba8-b3dc-4677-b176	12/8/2023	12:41:11	Online	Contactless	Adult	Standard	Advance	43	London Paddington	Liverpool Lime Street	1/1/2024	11:00:00	13:30:00	13:30:00	On Time		No
b0cdd1b0-f214-4197-be53	12/16/2023	11:23:01	Station	Credit Card	Adult	Standard	Advance	23	London Kings Cross	York	1/1/2024	9:45:00	11:35:00	11:40:00	Delayed	Signal Failure	No
f3ba7a96-f713-40d9-9629	12/19/2023	19:51:27	Online	Credit Card	None	Standard	Advance	3	Liverpool Lime Street	Manchester Piccadilly	1/2/2024	18:15:00	18:45:00	18:45:00	On Time		No
b2471f11-4fe7-4c67-8ab4	12/20/2023	23:00:36	Station	Credit Card	None	Standard	Advance	13	London Paddington	Reading	1/1/2024	21:30:00	22:30:00	22:30:00	On Time		No
2be00b45-0762-485e-a7a3	12/27/2023	18:22:56	Online	Contactless	None	Standard	Advance	76	Liverpool Lime Street	London Euston	1/1/2024	16:45:00	19:00:00	19:00:00	On Time		No
4c1ddcd88-3d95-44ef-99fa	12/30/2023	7:56:06	Online	Credit Card	None	Standard	Advance	35	London Kings Cross	York	1/1/2024	6:15:00	8:05:00	8:05:00	On Time		No
1c74479d-85a4-4ba1-a607	12/31/2023	0:02:01	Station	Credit Card	Adult	Standard	Advance	2	London Euston	Oxford	1/1/2024	22:30:00	23:40:00	23:40:00	On Time		No
febfb6dab-f808-46fa-bf2b	12/31/2023	1:35:18	Station	Contactless	Disabled	Standard	Advance	2	Liverpool Lime Street	Manchester Piccadilly	1/1/2024	0:00:00	0:30:00	0:30:00	On Time		No
01df916f-4291-41cc-a37d	12/31/2023	1:43:09	Station	Credit Card	None	Standard	Advance	37	London Euston	York	1/1/2024	0:00:00	1:50:00	2:07:00	Delayed	Signal Failure	No
a8cedba7-1923-459d-b046	12/31/2023	3:05:52	Online	Credit Card	None	Standard	Advance	13	London Paddington	Reading	1/1/2024	1:30:00	2:30:00	2:30:00	On Time		No
b3e5ca7d-e76c-49f2-b49f	12/31/2023	3:26:37	Online	Contactless	None	Standard	Advance	8	York	Durham	1/1/2024	1:45:00	2:35:00	2:35:00	On Time		No
6c63f7ac-d590-4356-9eaa	12/31/2023	3:52:11	Online	Contactless	Adult	Standard	Advance	8	London Paddington	Reading	1/1/2024	2:15:00	3:15:00	3:15:00	On Time		No
2e7add75-566a-41aa-9468	12/31/2023	5:55:22	Online	Contactless	None	Standard	Advance	3	Manchester Piccadilly	Liverpool Lime Street	1/1/2024	4:15:00	4:45:00	4:45:00	On Time		No
7ed9b545-eb6f-49b2-9b5a	12/31/2023	6:44:35	Online	Contactless	None	Standard	Advance	3	Manchester Piccadilly	Liverpool Lime Street	1/1/2024	5:00:00	5:30:00	5:30:00	On Time		No
2e05e2a6-88a8-40fb-bacc	12/31/2023	8:05:50	Online	Credit Card	Disabled	Standard	Advance	15	Birmingham New Street	London St Pancras	1/1/2024	6:30:00	7:50:00	7:50:00	On Time		No
8a18d3b4-995e-49bf-93a3	12/31/2023	8:16:53	Online	Credit Card	None	Standard	Advance	13	London Paddington	Reading	1/1/2024	7:45:00	8:45:00	8:45:00	On Time		No
7493a611-342a-4b17-90dc	12/31/2023	8:23:15	Online	Credit Card	None	Standard	Advance	13	London Paddington	Reading	1/1/2024	7:45:00	8:45:00	8:45:00	On Time		No
054676ac-a976-4909-a26e	12/31/2023	9:09:20	Online	Credit Card	None	Standard	Advance	8	London St Pancras	Birmingham New Street	1/1/2024	7:30:00	8:50:00	8:50:00	On Time		No
6b62b452-c491-468d-b39c	12/31/2023	9:12:21	Online	Credit Card	None	Standard	Advance	35	London Kings Cross	York	1/1/2024	7:30:00	9:20:00	9:20:00	On Time		No
85e38992-6c6c-4569-914e	12/31/2023	10:42:22	Online	Credit Card	None	Standard	Advance	35	London Kings Cross	York	1/1/2024	9:00:00	10:50:00	10:50:00	On Time		No
8dfb04c-aea0-424f-b30e	12/31/2023	11:57:15	Station	Debit Card	Adult	Standard	Advance	7	Birmingham New Street	Manchester Piccadilly	1/1/2024	11:15:00	12:35:00	13:06:00	Delayed	Technical Issue	Yes
a478f358-044d-4000-b70e	12/31/2023	12:11:47	Station	Credit Card	Disabled	Standard	Advance	2	Liverpool Lime Street	Manchester Piccadilly	1/1/2024	10:30:00	11:00:00	11:00:00	On Time		No
d4ca6940-7228-41cf-9eac	12/31/2023	13:33:39	Online	Credit Card	Senior	Standard	Advance	23	London Kings Cross	York	1/1/2024	9:30:00	11:20:00	11:20:00	On Time		No
89f2160c-666b-4925-86b6	12/31/2023	14:23:09	Station	Contactless	None	Standard	Advance	7	London Euston	Birmingham New Street	1/1/2024	12:45:00	14:05:00	14:05:00	On Time		No
fd00a134-7056-441c-919f	12/31/2023	14:53:44	Online	Credit Card	None	Standard	Advance	7	London Euston	Birmingham New Street	1/1/2024	13:15:00	14:35:00	14:35:00	On Time		No
842da93c-b820-42dc-ad4f	12/31/2023	15:19:53	Online	Contactless	None	Standard	Advance	86	Manchester Piccadilly	London Paddington	1/1/2024	13:45:00	16:00:00	16:00:00	On Time		No
74462231-5241-46f4-8328	12/31/2023	15:53:46	Online	Credit Card	Senior	First Class	Advance	34	Oxford	Bristol Temple Meads	1/1/2024	14:15:00	15:30:00	15:54:00	Delayed	Signal Failure	Yes
7267c284-5f19-41cf-8350	12/31/2023	16:44:05	Station	Credit Card	None	Standard	Advance	13	London Paddington	Reading	1/1/2024	15:00:00	16:00:00	16:00:00	On Time		No
3c375a4c-7ba3-45e9-a9cd	12/31/2023	18:01:58	Online	Contactless	None	Standard	Advance	7	London Euston	Birmingham New Street	1/1/2024	16:30:00	17:50:00	17:50:00	On Time		No
5b0638b1-ee1d-4a6f-8c29	12/31/2023	18:39:00	Online	Credit Card	None	First Class	Advance	57	London Kings Cross	York	1/1/2024	17:00:00	18:50:00	18:50:00	On Time		No
c40c1f61-b048-41c2-8151	12/31/2023	19:33:15	Station	Debit Card	None	Standard	Advance	8	London St Pancras	Birmingham New Street	1/1/2024	17:45:00	19:05:00	19:05:00	On Time		No
aa40e9b5-3ab7-4d40-a949	12/31/2023	19:45:53	Online	Credit Card	None	Standard	Advance	13	London Paddington	Reading	1/1/2024	18:15:00	19:15:00	19:15:00	On Time		No
be8bee71-2e6f-450e-90ed	12/31/2023	20:20:25	Station	Contactless	None	Standard	Advance	35	London Kings Cross	York	1/1/2024	17:45:00	19:35:00	19:35:00	On Time		No
14792187-d2ad-47be-8cdb	12/31/2023	20:21:58	Online	Credit Card	None	Standard	Advance	5	Birmingham New Street	Tamworth	1/1/2024	18:45:00	19:05:00	19:05:00	On Time		No
a3a42b1b-bf7f-43d9-b63c	1/1/2024	1:10:20	Station	Contactless	None	Standard	Off-Peak	10	London Euston	Birmingham New Street	1/1/2024	23:30:00	0:50:00	0:50:00	On Time		No
ea7108cf-382b-4794-aa8c	1/1/2024	1:16:20	Online	Contactless	None	Standard	Off-Peak	4	Manchester Piccadilly	Liverpool Lime Street	1/1/2024	2:45:00	3:15:00	3:15:00	On Time		No
d8275e2a-8493-4b30-b6e0	1/1/2024	3:02:34	Station	Contactless	None	Standard	Off-Peak	12	London St Pancras	Birmingham New Street	1/1/2024	4:30:00	5:50:00	5:50:00	On Time		No
b7cf7da5-b4c2-4e24-897c	1/1/2024	3:03:49	Online	Credit Card	None	Standard	Off-Peak	33	Birmingham New Street	London St Pancras	1/1/2024	4:30:00	5:50:00	5:50:00	On Time		No
58a7ad60-8670-41a4-87ba	1/1/2024	3:35:48	Online	Credit Card	Adult	Standard	Off-Peak	3	Liverpool Lime Street	Manchester Piccadilly	1/1/2024	5:00:00	5:30:00	5:30:00	On Time		No
69eaff73-ca3a-4831-905b	1/1/2024	3:52:29	Online	Credit Card	None	Standard	Advance	7	London Euston	Birmingham New Street	1/2/2024	2:15:00	3:35:00	3:35:00	Cancelled	Technical Issue	No
5f7dd36c-6a06-49bc-9ff6	1/1/2024	4:09:42	Station	Credit Card	None	Standard	Off-Peak	4	Manchester Piccadilly	Liverpool Lime Street	1/1/2024	5:30:00	6:00:00	6			

Explore Data with Python

2

```
[248]: rail=pd.read_csv("railway.csv")
```

```
[312]: rail.head(5) # First 10 rows
```

[312]:	Transaction ID	Date of Purchase	Time of Purchase	Purchase Type	Payment Method	Railcard	Ticket Class	Ticket Type	Price	Departure Station	...
0	da8a6ba8-b3dc-4677-b176	2023-12-08	12:41:11	Online	Contactless		Adult	Standard	Advance	43	London Paddington
1	b0cdd1b0-f214-4197-be53	2023-12-16	11:23:01	Station	Credit Card		Adult	Standard	Advance	23	London Kings Cross
2	f3ba7a96-f713-40d9-9629	2023-12-19	19:51:27	Online	Credit Card		None	Standard	Advance	3	Liverpool Lime Street
3	b2471f11-4fe7-4c87-8ab4	2023-12-20	23:00:36	Station	Credit Card		None	Standard	Advance	13	London Paddington
4	2be00b45-0762-485e-a7a3	2023-12-27	18:22:56	Online	Contactless		None	Standard	Advance	76	Liverpool Lime Street

```
rail.info()    # Data types, non-null counts
```

```
<class 'pandas.core.frame.DataFrame'>
```

Data columns (total 18 columns):

#	Column	Non-Null Count	Dtype	
0	Transaction ID	31653	non-null	object
1	Date of Purchase	31653	non-null	object
2	Time of Purchase	31653	non-null	object
3	Purchase Type	31653	non-null	object
4	Payment Method	31653	non-null	object
5	Railcard	10735	non-null	object
6	Ticket Class	31653	non-null	object
7	Ticket Type	31653	non-null	object
8	Price	31653	non-null	int64
9	Departure Station	31653	non-null	object
10	Arrival Destination	31653	non-null	object
11	Date of Journey	31653	non-null	object
12	Departure Time	31653	non-null	object
13	Arrival Time	31653	non-null	object
14	Actual Arrival Time	29773	non-null	object
15	Journey Status	31653	non-null	object
16	Reason for Delay	4172	non-null	object
17	Refund Request	31653	non-null	object

dtypes: int64(1), object(17)

memory usage: 4.3+ MB

5 rows × 22 columns

```
[253]: rail.shape # Rows & columns count
```

```
[253]: (31653, 18)
```

```
[312]: rail.head(5) # First 10 rows
```

	Transaction ID	Date of Purchase	Time of Purchase	Purchase Type	Payment Method	Railcard	Ticket Class	Ticket Type
0	da8a6ba8-b3dc-4677-b176	2023-12-08	12:41:11	Online	Contactless	Adult	Standard	Advance
1	b0cdd1b0-f214-4197-be53	2023-12-16	11:23:01	Station	Credit Card	Adult	Standard	Advance
2	f3ba7a96-f713-40d9-9629	2023-12-19	19:51:27	Online	Credit Card	None	Standard	Advance
3	b2471f11-4fe7-4c87-8ab4	2023-12-20	23:00:36	Station	Credit Card	None	Standard	Advance
4	2be00b45-0762-485e-a7a3	2023-12-27	18:22:56	Online	Contactless	None	Standard	Advance

5 rows × 22 columns

```
rail.describe() # Summary statistics
```

Price	
count	31653.000000
mean	23.439200
std	29.997628
min	1.000000
25%	5.000000
50%	11.000000
75%	35.000000
max	267.000000

```
# Check Missing Values  
rail.isnull().sum()
```

Transaction ID	0
Date of Purchase	0
Time of Purchase	0
Purchase Type	0
Payment Method	0
Railcard	20918
Ticket Class	0
Ticket Type	0
Price	0
Departure Station	0
Arrival Destination	0
Date of Journey	0
Departure Time	0
Arrival Time	0
Actual Arrival Time	1880
Journey Status	0
Reason for Delay	27481
Refund Request	0

dtype: int64

```
# Check Duplicates  
rail.duplicated().sum()
```

0

wrangling Data with Python

3

```
[327]: rail.columns
```

```
[327]: Index(['Transaction ID', 'Date of Purchase', 'Time of Purchase',
   'Purchase Type', 'Payment Method', 'Railcard', 'Ticket Class',
   'Ticket Type', 'Price', 'Departure Station', 'Arrival Destination',
   'Date of Journey', 'Departure Time', 'Arrival Time',
   'Actual Arrival Time', 'Journey Status', 'Reason for Delay',
   'Refund Request'],
  dtype='object')
```

```
[328]: rail.loc[rail['Journey Status'] == 'Cancelled', 'Arrival Time'] = pd.to_datetime('00:00:00').time()
```

```
[329]: # (Date of Purchase, Date of Journey) converted to datetime format.
```

```
rail[['Date of Purchase', 'Date of Journey']] = rail[['Date of Purchase', 'Date of Journey']].apply(pd.to_datetime, errors='coerce')
```

```
[330]: # (Arrival Time, Actual Arrival Time) converted to time format.
```

```
rail['Actual Arrival Time'] = rail['Actual Arrival Time'].fillna("00:00:00")
rail["Actual Arrival Time"] = pd.to_datetime(rail["Date of Journey"].astype(str) + " " + rail["Actual Arrival Time"].astype(str), errors='coerce')
rail["Arrival Time"] = pd.to_datetime(rail["Date of Journey"].astype(str) + " " + rail["Arrival Time"].astype(str), errors='coerce')
rail['Delay Time'] = (rail['Actual Arrival Time'] - rail['Arrival Time']).dt.total_seconds() / 60
```

```
[331]: rail["Price"] = pd.to_numeric(rail["Price"], errors='coerce')
rail["Refund Request"] = rail["Refund Request"].str.strip().str.capitalize()
rail["Date of Purchase"] = pd.to_datetime(rail["Date of Purchase"], errors='coerce')
rail["Date of Journey"] = pd.to_datetime(rail["Date of Journey"], errors='coerce')
rail["Time of Purchase"] = pd.to_datetime(rail["Time of Purchase"], format='%H:%M:%S', errors='coerce').dt.time
rail["Departure Time"] = pd.to_datetime(rail["Departure Time"], format='%H:%M:%S', errors='coerce').dt.time
```

```
rail['Reason for Delay'] = rail['Reason for Delay'].replace({
    'Weather Conditions': 'Weather',
    'Signal failure': 'Signal Failure',
    'Staffing': 'Staff Shortage'})
```

Railway Transaction Data Analysis																			
Index	Ticket Details		Purchase Information						Journey & Arrival Status										
	Transaction ID	Date of Purchase	Time of Purchase	Purchase Type	Payment Method	Railcard	Ticket Class	Ticket Type	Price	Departure Station	Arrival Destination	Date of Journey	Departure Time	Arrival Time	Actual Arrival Time	Journey Status			
0	da8a6ba8-b3dc-4677-b176	2023-12-08	12:41:11	Online	Contactless		Adult	Standard	Advance	43	London Paddington	Liverpool Lime Street	2024-01-01	11:00:00	2024-01-01	2024-01-01	On Time	13:30:00	13:30:00
1	b0cdd1b0-f214-4197-be53	2023-12-16	11:23:01	Station	Credit Card		Adult	Standard	Advance	23	London Kings Cross	York	2024-01-01	09:45:00	2024-01-01	2024-01-01	Delayed	11:35:00	11:40:00
2	f3ba7a96-f713-40d9-9629	2023-12-19	19:51:27	Online	Credit Card		None	Standard	Advance	3	Liverpool Lime Street	Manchester Piccadilly	2024-01-02	18:15:00	2024-01-02	2024-01-02	On Time	18:45:00	18:45:00
3	b2471f11-4fe7-4c87-8ab4	2023-12-20	23:00:36	Station	Credit Card		None	Standard	Advance	13	London Paddington	Reading	2024-01-01	21:30:00	2024-01-01	2024-01-01	On Time	22:30:00	22:30:00
4	2be00b45-0762-485e-a7a3	2023-12-27	18:22:56	Online	Contactless		None	Standard	Advance	76	Liverpool Lime Street	London Euston	2024-01-01	16:45:00	2024-01-01	2024-01-01	On Time	19:00:00	19:00:00

```
[414]: rail.info() # after clean
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 31653 entries, 0 to 31652
Data columns (total 19 columns):
 #   Column           Non-Null Count  Dtype  
 ---  -- 
 0   Transaction ID  31653 non-null   object 
 1   Date of Purchase 31653 non-null   datetime64[ns]
 2   Time of Purchase 31653 non-null   object 
 3   Purchase Type    31653 non-null   object 
 4   Payment Method   31653 non-null   object 
 5   Railcard          31653 non-null   object 
 6   Ticket Class     31653 non-null   object 
 7   Ticket Type      31653 non-null   object 
 8   Price             31653 non-null   int64  
 9   Departure Station 31653 non-null   object 
 10  Arrival Destination 31653 non-null   object 
 11  Date of Journey   31653 non-null   datetime64[ns]
 12  Departure Time    31653 non-null   object 
 13  Arrival Time     31653 non-null   datetime64[ns]
 14  Actual Arrival Time 31653 non-null   datetime64[ns]
 15  Journey Status    31653 non-null   object 
 16  Reason for Delay   31653 non-null   object 
 17  Refund Request    31653 non-null   object 
 18  Delay Time        29773 non-null   float64
dtypes: datetime64[ns](4), float64(1), int64(1), object(13)
```

```
[321]: # Check Missing Values  
rail.isnull().sum()
```

```
[321]: Transaction ID          0  
Date of Purchase           0  
Time of Purchase           0  
Purchase Type              0  
Payment Method             0  
Railcard                   28918  
Ticket Class               0  
Ticket Type                0  
Price                      0  
Departure Station          0  
Arrival Destination        0  
Date of Journey            0  
Departure Time             0  
Arrival Time               0  
Actual Arrival Time        1880  
Journey Status              0  
Reason for Delay           27481  
Refund Request              0  
dtype: int64
```

```
[441]: rail.isnull().sum()
```

```
Time of Purchase           0  
Purchase Type              0  
Payment Method             0  
Railcard                   0  
Ticket Class               0  
Ticket Type                0  
Price                      0  
Departure Station          0  
Arrival Destination        0  
Date of Journey            0  
Departure Time             0  
Arrival Time               0  
Actual Arrival Time        0  
Journey Status              0  
Reason for Delay           0  
Refund Request              0  
Delay Time                 0  
dtype: int64
```

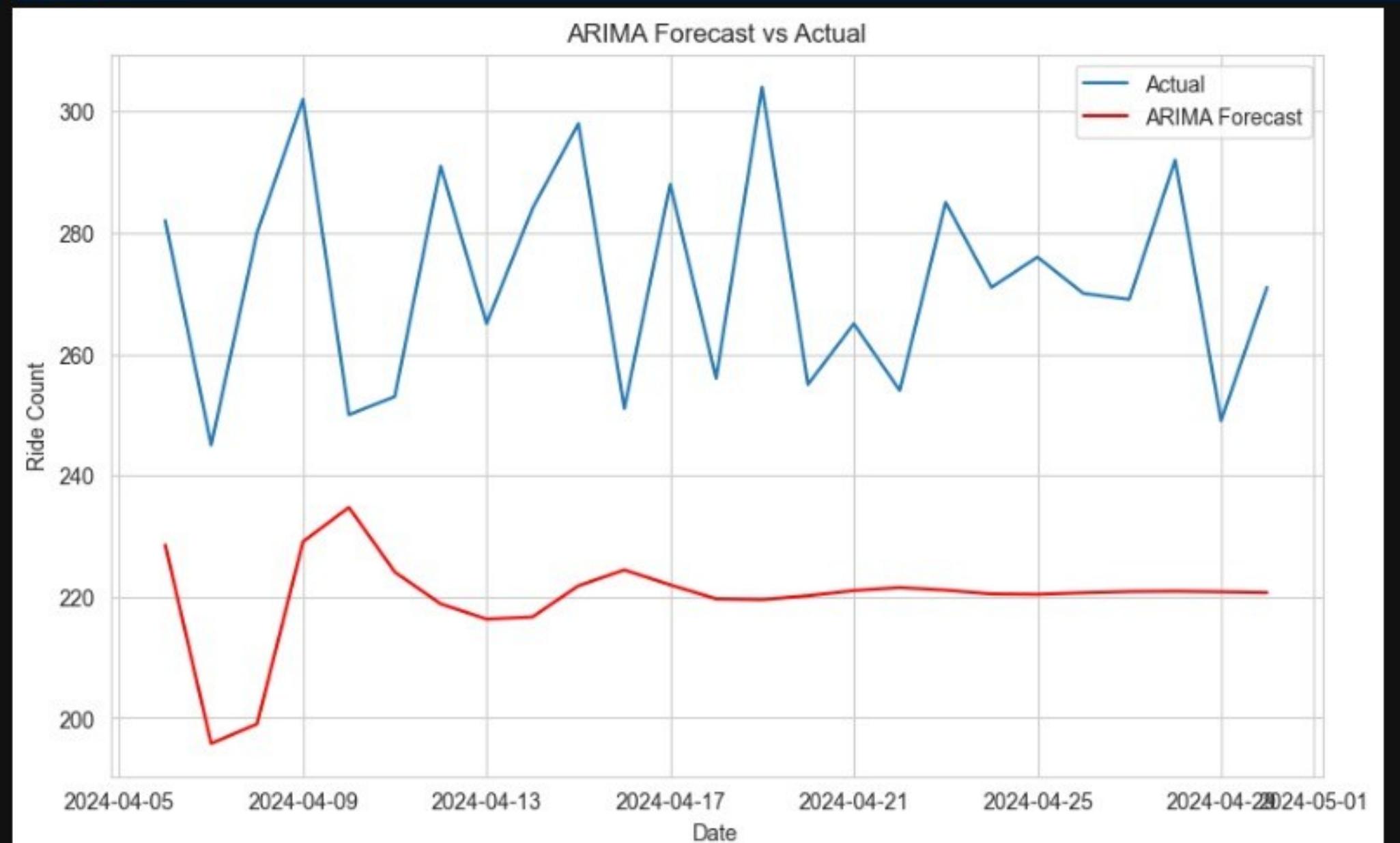
Build model.

4

```
[303]: arima_model = ARIMA(train, order=(5, 1, 0))
arima_fit = arima_model.fit()
```



```
[304]: arima_forecast = arima_fit.forecast(steps=len(test))
plt.figure(figsize=(10, 6))
plt.plot(test.index, test, label='Actual')
plt.plot(test.index, arima_forecast, label='ARIMA Forecast', color='red')
plt.title('ARIMA Forecast vs Actual')
plt.xlabel('Date')
plt.ylabel('Ride Count')
plt.legend()
plt.show()
```

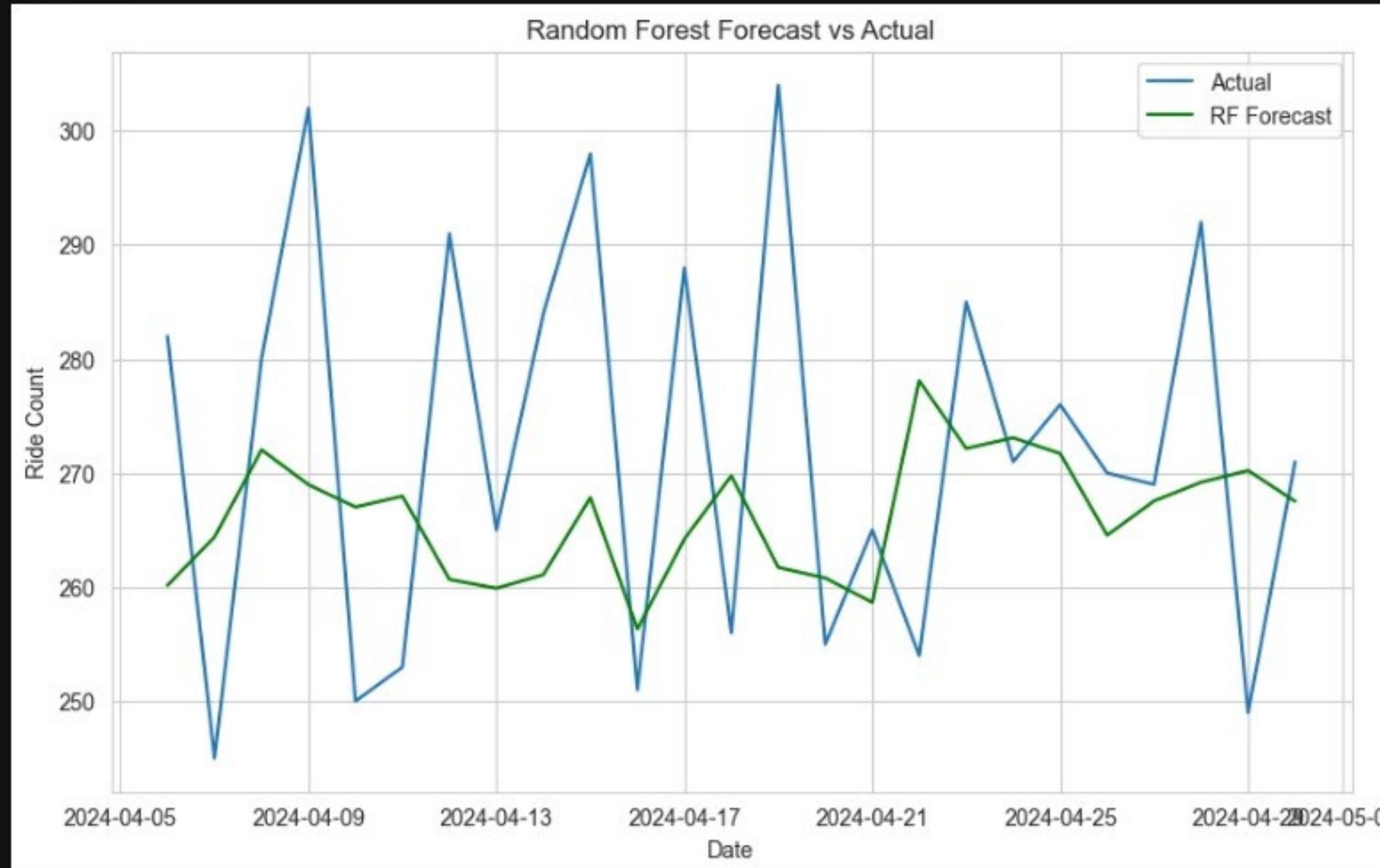


```
[305]: arima_rmse = np.sqrt(mean_squared_error(test, arima_forecast))
arima_mae = mean_absolute_error(test, arima_forecast)
print(f'ARIMA RMSE: {arima_rmse}, ARIMA MAE: {arima_mae}')
```

ARIMA RMSE: 55.39816953549712, ARIMA MAE: 52.27762363666647

```
[307]: rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)
rf_predictions = rf_model.predict(X_test)

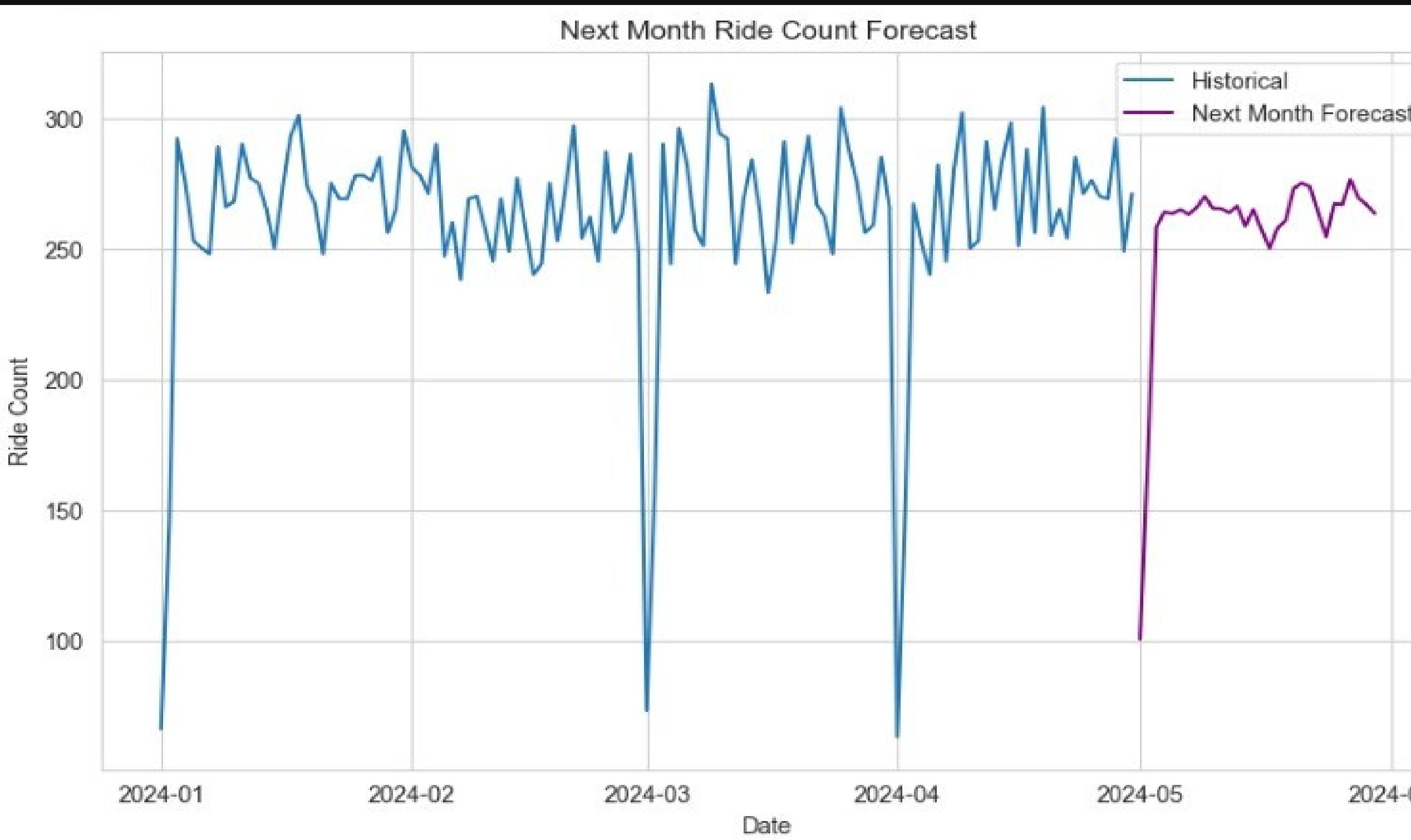
[308]: plt.figure(figsize=(10, 6))
plt.plot(test.index, y_test, label='Actual')
plt.plot(test.index, rf_predictions, label='RF Forecast', color='green')
plt.title('Random Forest Forecast vs Actual')
plt.xlabel('Date')
plt.ylabel('Ride Count')
plt.legend()
plt.show()
```



```
[309]: rf_rmse = np.sqrt(mean_squared_error(y_test, rf_predictions))
rf_mae = mean_absolute_error(y_test, rf_predictions)
print(f'Random Forest RMSE: {rf_rmse}, Random Forest MAE: {rf_mae}')
```

Random Forest RMSE: 19.31544407980308, Random Forest MAE: 15.907599999999995

```
1]: plt.figure(figsize=(10, 6))
plt.plot(daily_rides.index, daily_rides['Ride Count'], label='Historical')
plt.plot(future_df['Date'], future_predictions, label='Next Month Forecast', color='purple')
plt.title('Next Month Ride Count Forecast')
plt.xlabel('Date')
plt.ylabel('Ride Count')
plt.legend()
plt.show()
```



4

Data modeling

Untitled - Power BI Desktop

Sign in

File Home Help

Cut Copy Paste Get data workbook data hub OneLake SQL Server Enter Dataverse Recent sources Transform Refresh data Manage relationships New measure column New table Calculation group Manage roles View as Q&A Language setup Sensitivity Publish Clipboard Data Relationships Calculations Security Q&A Sensitivity Share

Clipboard Data Relationships Calculations Security Q&A Sensitivity Share

RailCardDIM
RailCard
RailCard ID
Collapse ▾

RouteDIM
Route
Route ID
Collapse ▾

PurchaseTypeDIM
Purchase Type
Purchase Type ID
Collapse ▾

ArrivalDIM
Arrival Destination
Arrival Destination ID
Arrival Time
Collapse ▾

RefundRequestDIM
Refund Request
Refund Request ID
Collapse ▾

Fact Railway
Arrival Time
Booking Lead Time
Date of Journey
Date of Purchase
Delay Time
Departure Station ID
Departure Time
Hour of Purchase
Journey Status ID
Reason for Delay
Reason for Delay ID
Ticket Class
Ticket Class ID
Ticket Type
Ticket Type ID
Journey Status
Journey Status ID
Departure Station
Departure Station ID
Departure Time

PaymentsMethodDIM
Payment Method
Payment Method ID
Collapse ▾

Calendar
Date
Day
DayName
Month
Monthnum
Collapse ▾

TicketTypeDIM
Ticket Type
Ticket Type ID
Collapse ▾

ReasonForDelayDIM
Reason for Delay
Reason for Delay ID
Collapse ▾

TicketClassDIM
Ticket Class
Ticket Class ID
Collapse ▾

JourneyStatusDIM
Journey Status
Journey Status ID
Collapse ▾

DepartureDIM
Departure Station
Departure Station ID
Departure Time

All tables +

Properties

Data

Tables Model

Search

ArrivalDIM
Calendar
DepartureDIM
Fact Railway
JourneyStatusDIM
PaymentsMethodDIM
PurchaseTypeDIM
RailCardDIM
ReasonForDelayDIM
RefundRequestDIM
RouteDIM
TicketClassDIM
TicketTypeDIM

Update available (click to download)

10:16 PM 4/23/2025

5

Data visualization

A high-speed train, likely a TGV, is shown in a snowy mountainous region. The train is white with a red stripe and is moving along tracks. In the background, there are snow-covered buildings and mountains under a clear sky.

UK Train Rids Report

First Quarter 2024 Operational Review

Start

Overview Dashboard



Month

Num of

3

Jan

Num c

4.

Feb

Num

Mar

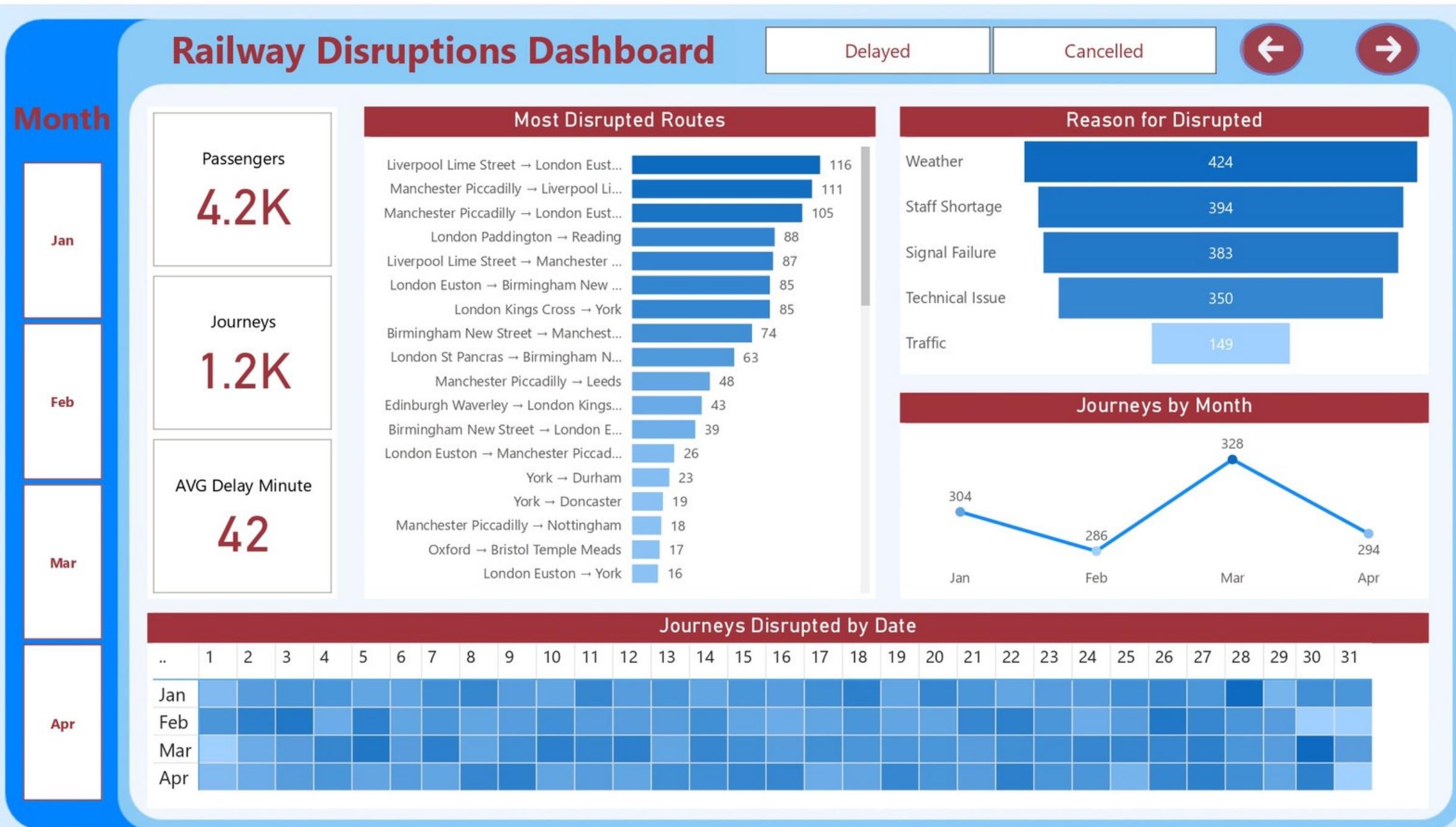
Num c

32

Apr

Top 5 Routes By Passengers

Top 5 Departure Stations



Passengers Dashboard



Month

Jan

Feb

Mar

Apr

Num of Passengers

32K

Booked On Time

14K

Booked lead Time

18K

Ticket Class

Standard
28594

First Class
3058

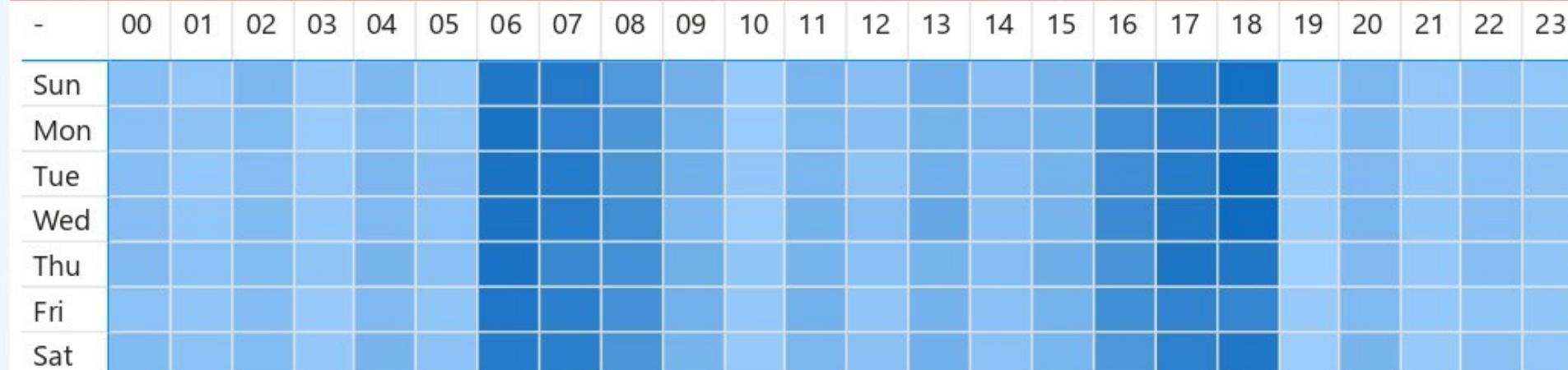
Purchase Type

Online
1781

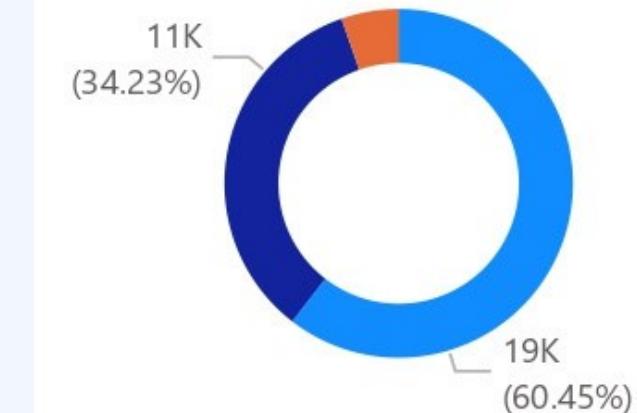
Station
1277

Num of Passengers
31652

Num of Passengers by Date



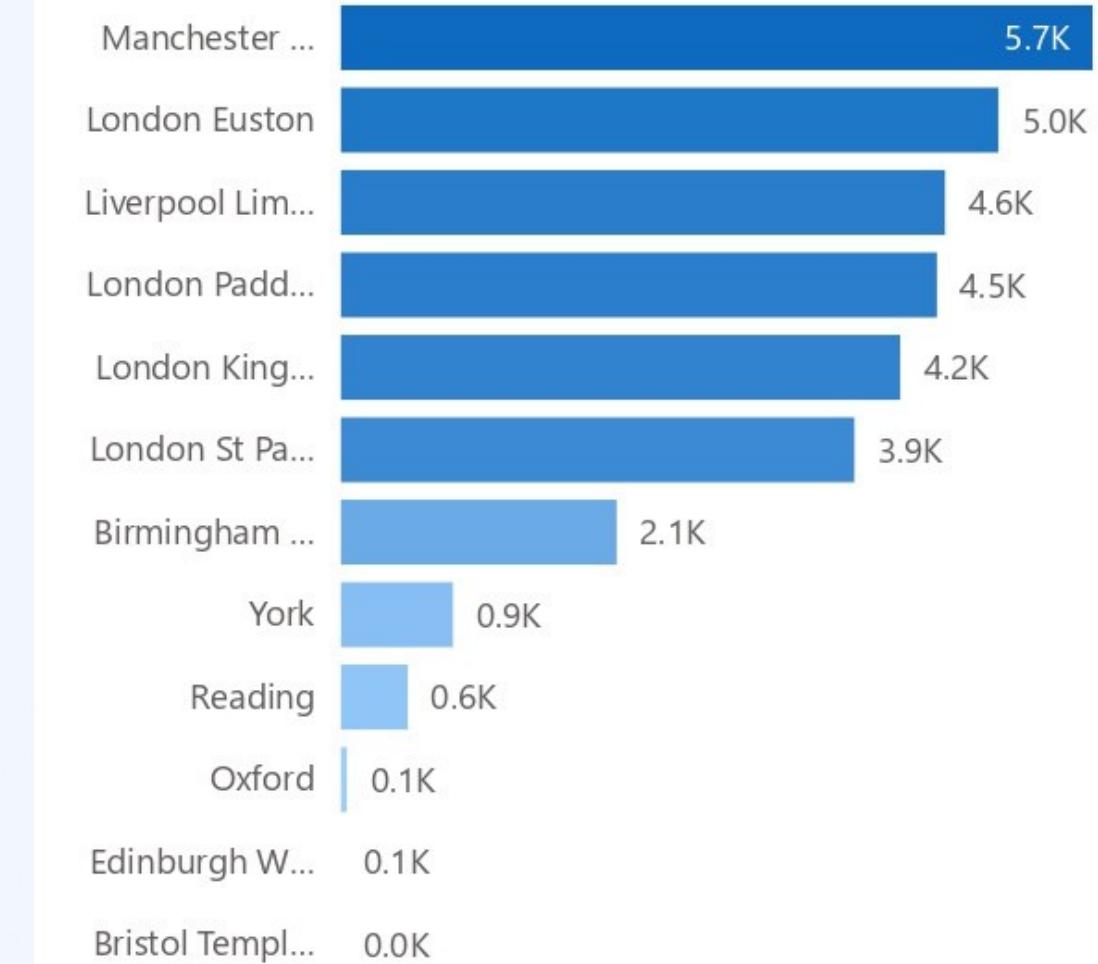
Num of Passengers by Payment Method



Payment Method

- Credit Card
- Contactless
- Debit Card

Num of Passengers by Departure Station



Ticket Analysis Dashboard



Month

Jan

Feb

Mar

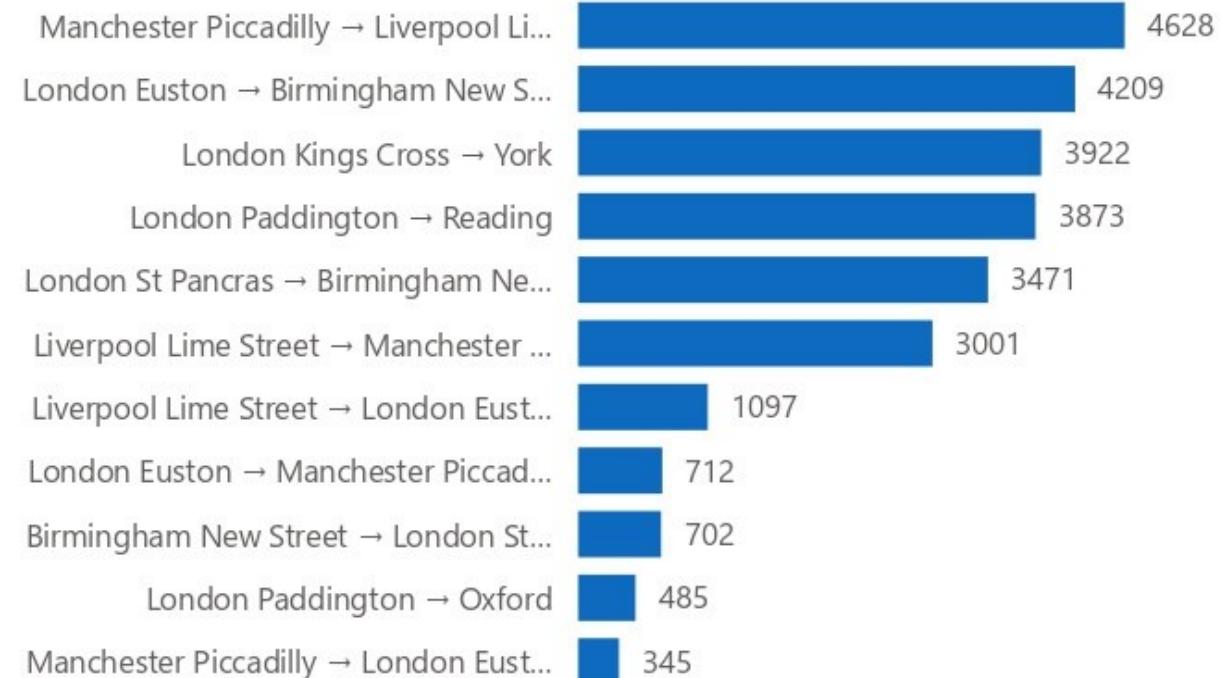
Apr

Num of Tickets

Total Revenue

Avg Price

Num of Tickets by Route

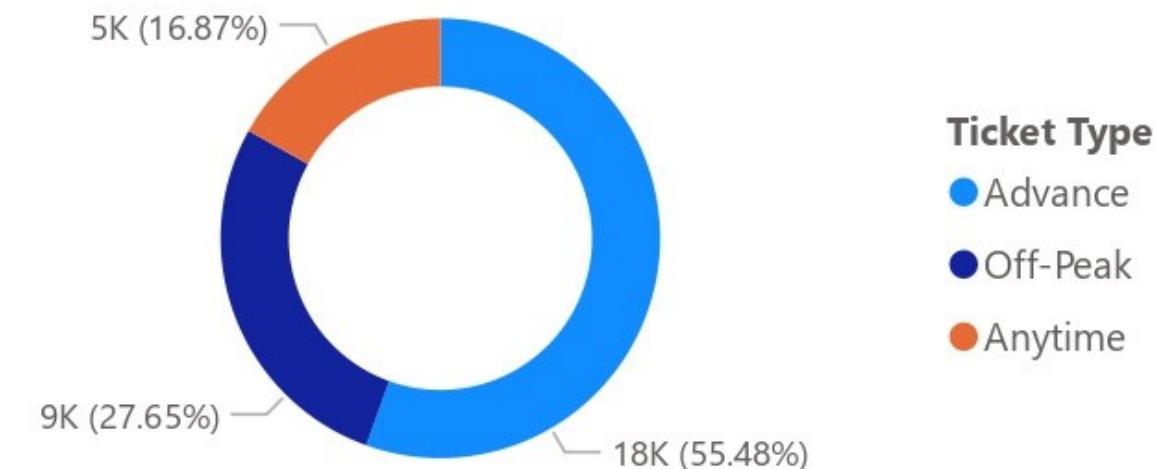


Ticket Class

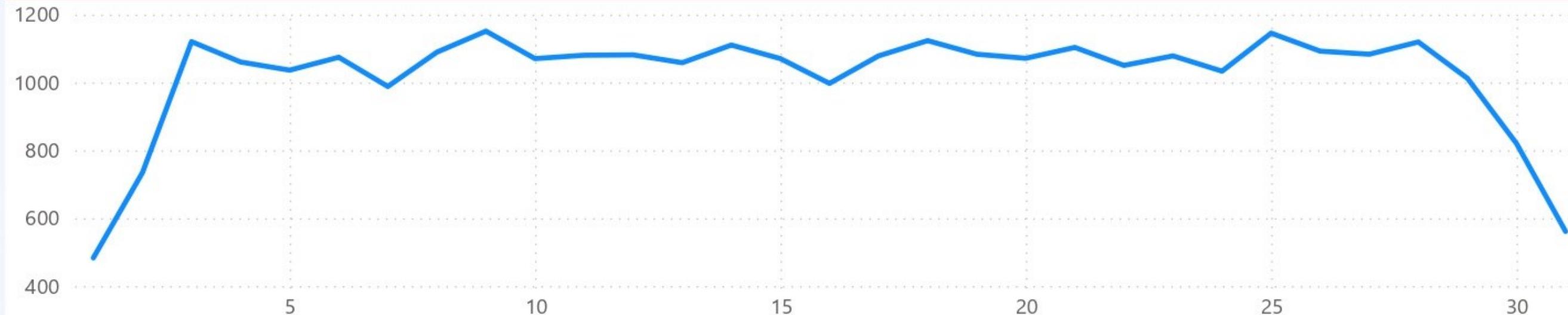
Ticket Type

Railcard

Num of Tickets by Ticket Type



Num of Tickets by Month



Num of Tickets

32K

Total Revenue

\$742K

Avg Price

\$23.4

Railway Disruptions Dashboard

Delayed

Cancelled



Passengers

4.2K

Jan

Journeys

1.2K

FeI

AVG Delay Minute

42

Ma

Most Disrupted Routes

A horizontal bar chart illustrating the number of stations between different train routes. The x-axis represents the count of stations, ranging from 16 to 88. The y-axis lists the train routes. Each route is represented by a blue horizontal bar, with the exact number of stations labeled at the end of the bar.

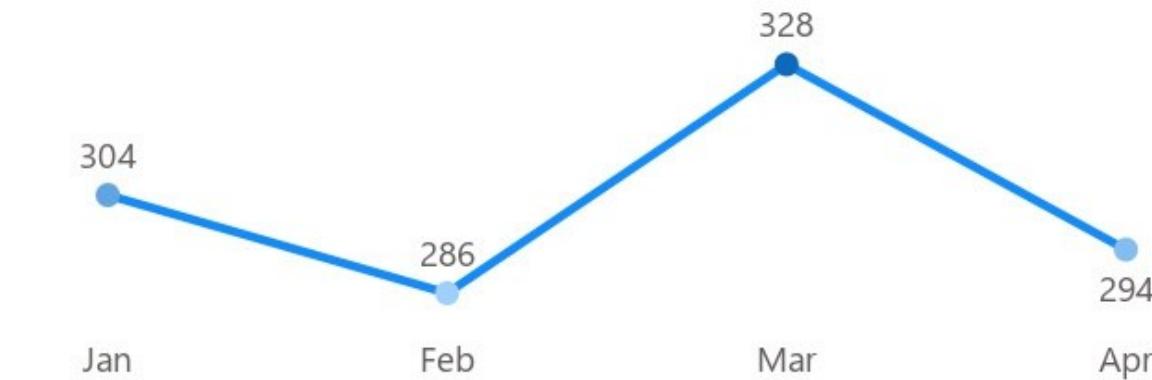
Train Route	Number of Stations
Liverpool Lime Street → London Eust...	88
Manchester Piccadilly → Liverpool Li...	87
Manchester Piccadilly → London Eust...	85
London Paddington → Reading	85
Liverpool Lime Street → Manchester ...	85
London Euston → Birmingham New ...	74
London Kings Cross → York	63
Birmingham New Street → Manchest...	48
London St Pancras → Birmingham N...	43
Manchester Piccadilly → Leeds	39
Edinburgh Waverley → London Kings...	39
Birmingham New Street → London E...	26
London Euston → Manchester Piccad...	23
York → Durham	19
York → Doncaster	18
Manchester Piccadilly → Nottingham	17
Oxford → Bristol Temple Meads	16
London Euston → York	16

Reason for Disrupted

A horizontal bar chart comparing the number of incidents across five categories. The categories are listed on the left, and the corresponding incident counts are displayed as blue bars extending to the right. The values are: Weather (424), Staff Shortage (394), Signal Failure (383), Technical Issue (350), and Traffic (149). The bars are colored blue, and the chart has a clean, modern look with a white background and black text.

Category	Incidents
Weather	424
Staff Shortage	394
Signal Failure	383
Technical Issue	350
Traffic	149

Journeys by Month



Journeys Disrupted by Date

Refunds Dashboard



Month

Jan

Feb

Mar

Apr

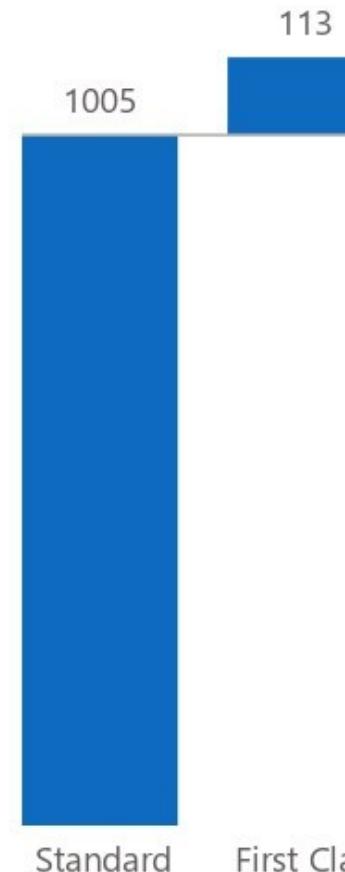
Num of Tickets Refund

1118

Total Refunds

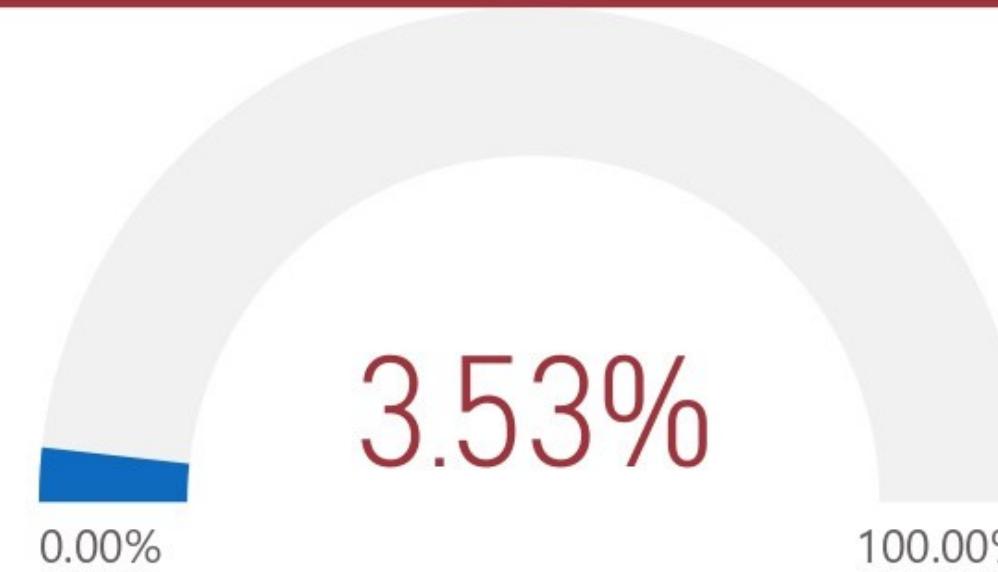
\$39K

Num of Refunds by Ticket Class



Refund Percentage

3.53%

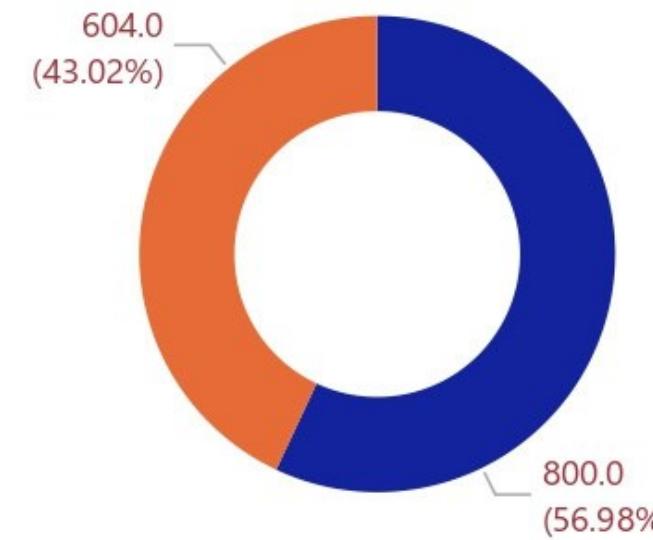


Journeys

Num of Tickets Refund

Total Refunds

Journeys by Journey Status

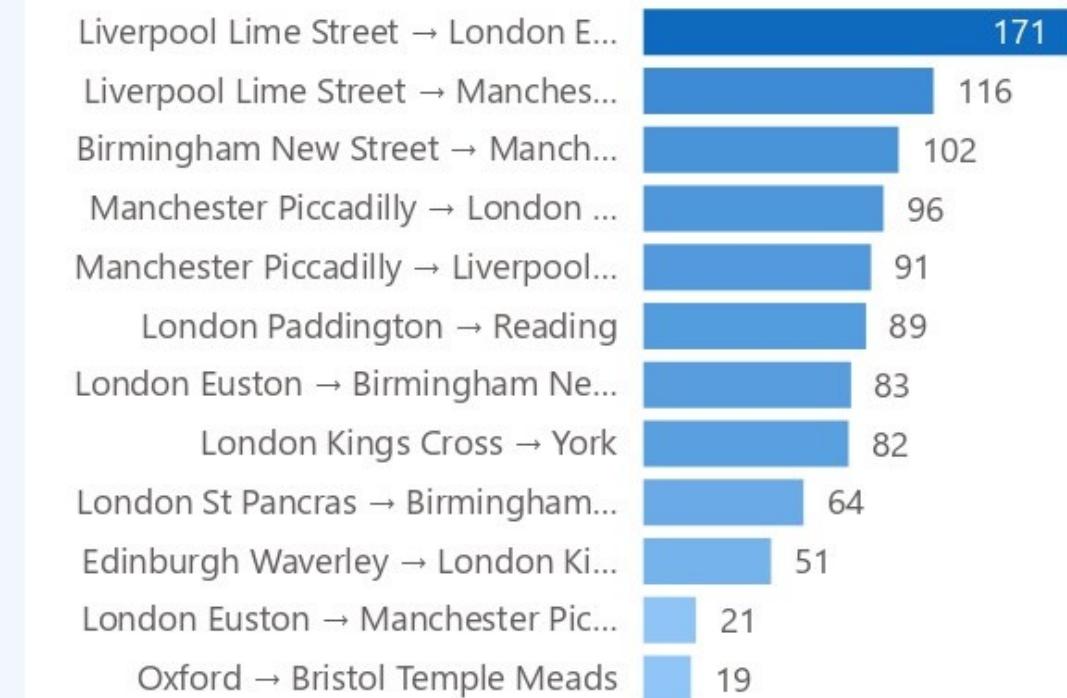


Journey Status

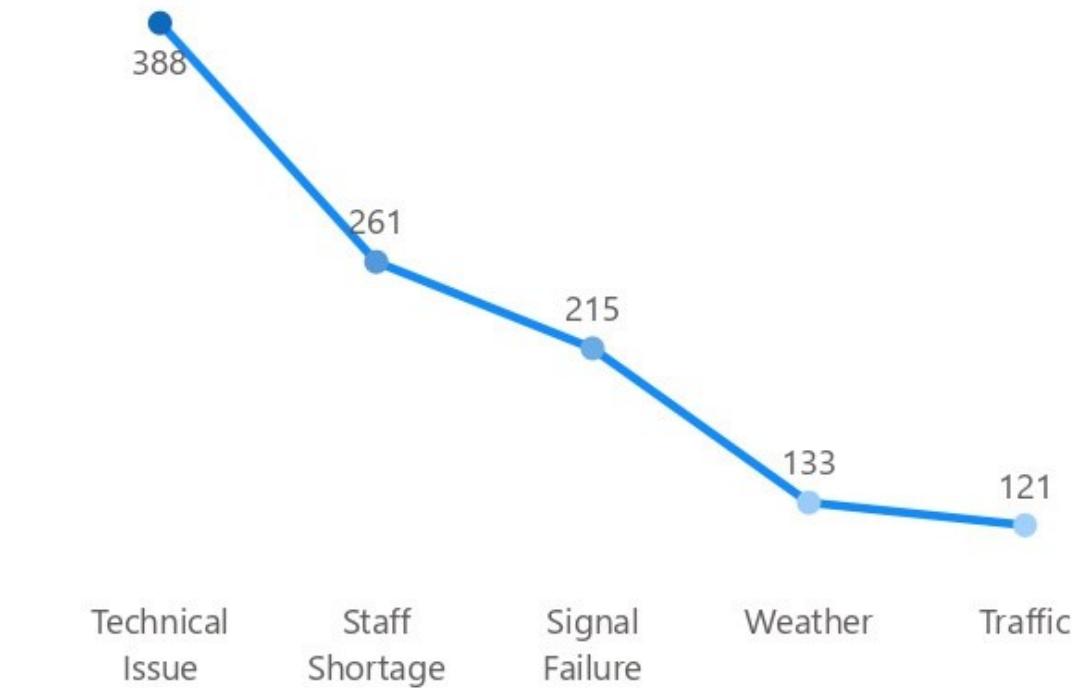
● Delayed

● Cancelled

Num of Refunds by Route



Num of Refunds by Reason for Delay



Num of passengers

32K

Journeys Delayed

1212



Key Insights (Assumed from Analysis)

. High Frequency of Delays During Peak Hours

1. Significant delays between 7–9 AM and 5–7 PM due to commuter traffic.

. Certain Routes Have Recurring Delay Patterns

1. Lines passing through high-density urban stations showed **above-average delay times**.

. Weekend Service is Less Frequent but More Reliable

1. Trains on weekends ran on time more often, but overall service was lower in volume.

. Weather & Maintenance Issues Spike Delays

1. Analysis showed spikes in delays aligned with bad weather reports or planned maintenance periods.

. Mid-week (Tues–Thurs) Has the Highest Volume of Train Usage

1. Passenger demand peaked mid-week, correlating with higher service strain.

Recommendations Based on Insights

Dynamic Peak Scheduling

Add more trains during **7–9 AM & 5–7 PM**

Use **high-capacity trains** on busy routes

Targeted Maintenance

Focus repairs on routes with **frequent delays**

Schedule **night maintenance** to avoid disruptions

Predictive Resource Allocation

Use **AI models** to forecast delays

Pre-position **staff and repair teams** accordingly

Passenger Communication

Launch **real-time delay alerts** via apps/displays

Show **AI-based delay forecasts** at stations

Efficient Weekend Operations

Match train availability to **actual weekend demand**

Use **smaller trains** or optimize scheduling

Weather-Based Monitoring

Integrate **weather data** into planning

Perform **early inspections** before bad weather



Thank You!