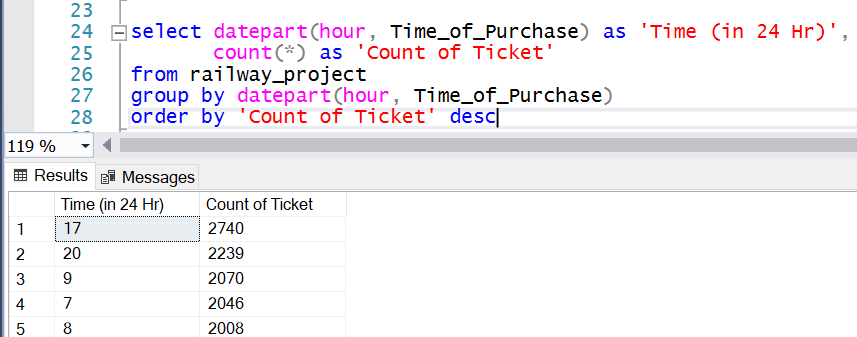
**Data Analytics Project**

**Part A-The SQL Analysis case study:**

**Query:1: Identify Peak Purchase Times and Their Impact on Delays:**

**Analysis:**



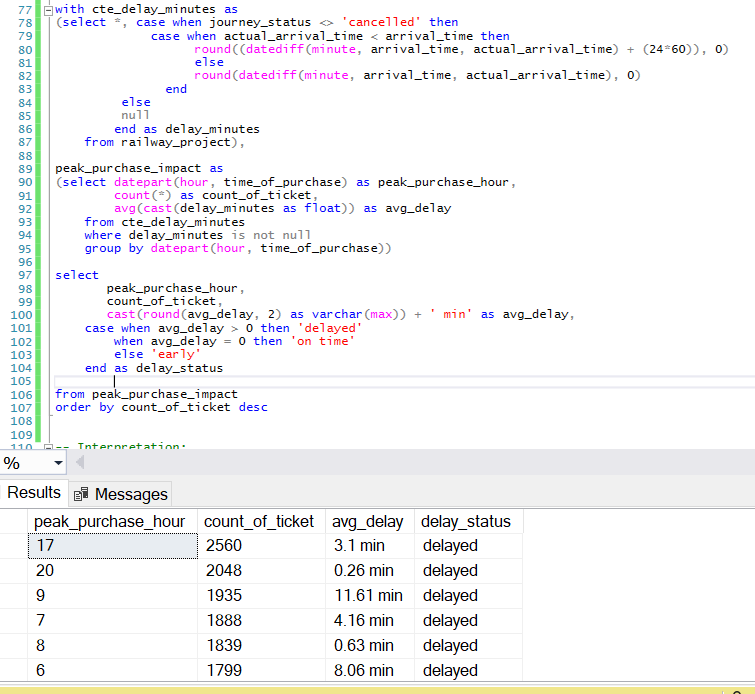
From the above interpretation, it indicates that ticket purchase times are highest in the morning and evening. This can be assumed to be due to the following reasons:

People are going to work in the morning, so ticket purchasing starts increasing from 8-9 AM, reaching its highest point during this period. Similarly, when people are returning home in the evening, ticket purchases also increase between 5 PM and 8 PM.

This increase in ticket purchases during these times can be attributed to people commuting to and from work, making these the peak times

for ticket purchases.

**Peak Time Impact on Delay and co-relation:**



From the above queries its indicate as I already understand from the previous query that the highest peak hours occur in the morning around 8-9 AM and in the evening around 5 PM and 8 PM, when the highest number of tickets are purchased by travellers, but when I tried to understand the corelation with delay its implies almost most of the peak times are associated with delay But most of Travelers purchasing tickets during these peak hours experience delays, with average delay times ranging from minimum 23 second to maximum 11.61 minutes.

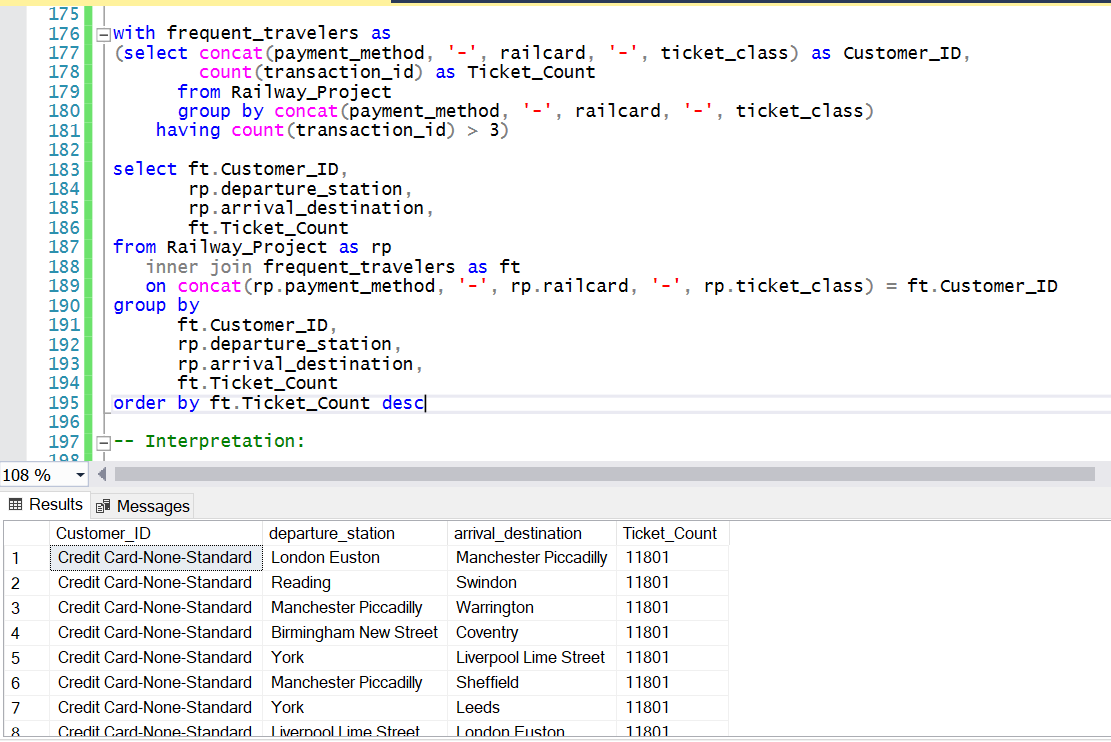
From the data, it can be assumed that delays are most significant during peak office hours, which include both morning and evening times, especially in the morning at 9 AM and in the evening at 6 PM. This indicates that delays are closely related to the start and end of working or office hours. Also, there is a noticeable impact during the late morning hours from 10 AM to 12 PM, immediately after the peak morning period.

If I interpret the correlation with peak hours, it generally shows moderate delays of 1 to 4 minutes, except for specific times like 9 AM, which have higher delays. Non-peak hours, particularly late at night, show no delays. We can assume that routes are busy until the end of the day due to traffic or operational issues.

I can be assumed from above peak purchase hours, ticket counts, and delays are interconnected. Travelers buying tickets during peak hours When ticket counts are high during peak hours, travellers are more likely to experience delays, it’s because of due with respect to working hour and returning home.

**Query:2 Analyze Journey Patterns of Frequent Travelers::**

**Analysis:**



I have identified due to not having sufficient data execution of analysis will

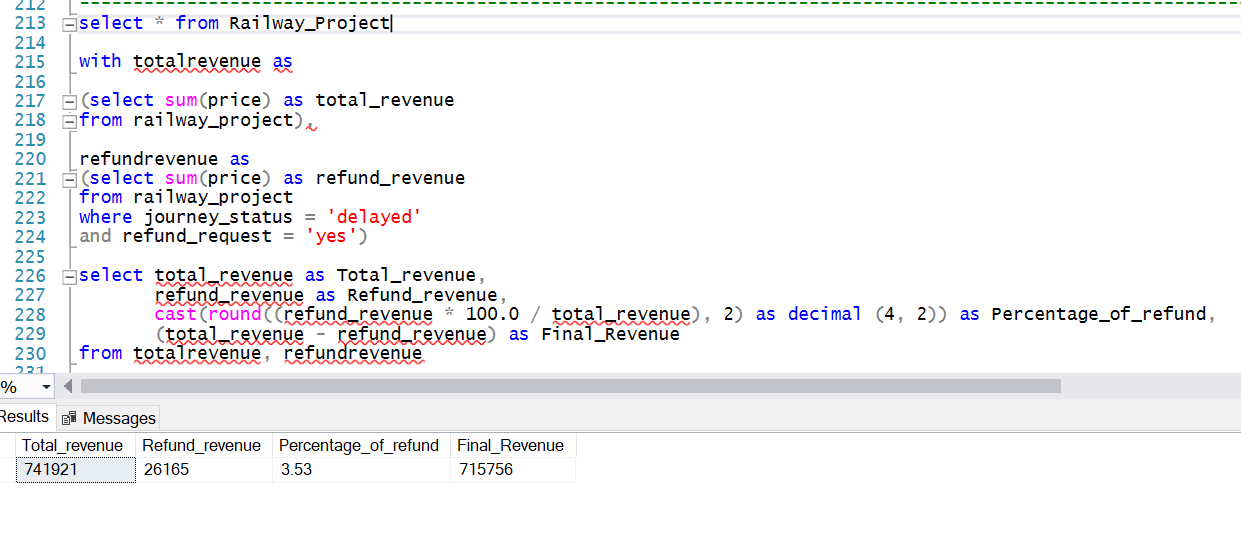
not be possible for which I have created a cust id by concating some columns to understand frequent travellers.

-- To understanding of frequent travellers and their travel behaviour of it can be interpret here:

The frequent travellers used routes are Manchester Piccadilly to London Euston, Liverpool Lime Street to Manchester Piccadilly and routes involving Birmingham New Street are the most frequently travelled and Credit Cards are the dominant payment method due to payment comfortability and followed by Contactless payments and Standard tickets, especially under without railcard holder category have the highest ticket counts, reflecting most traveller are frequent with economical journey with credit card, contactless payments method by without any railcard and adults category.

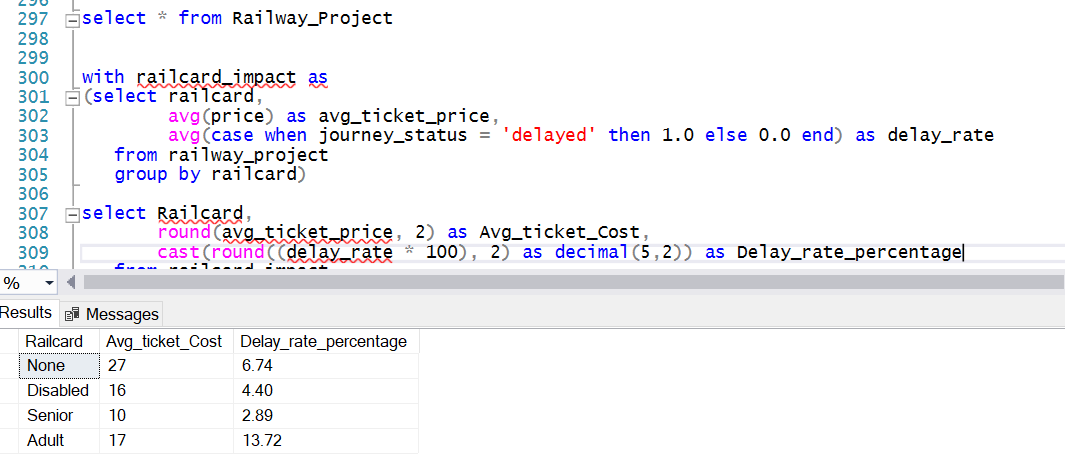
**Query:3 Revenue Loss Due to Delays with Refund Requests:**

**Analysis:**

****

From the above query it implies that total revenue of 741,921 from ticket sales. But because of delays its leading to refund requests due to train cancelled or technical issue or any reason where customer faced the issue they lost 26,165, which is about 3.53% of their total revenue. So, after refunds, railway collect the final revenue was 715,756.

**Query:4 Impact of Railcards on Ticket Prices and Journey Delays:**



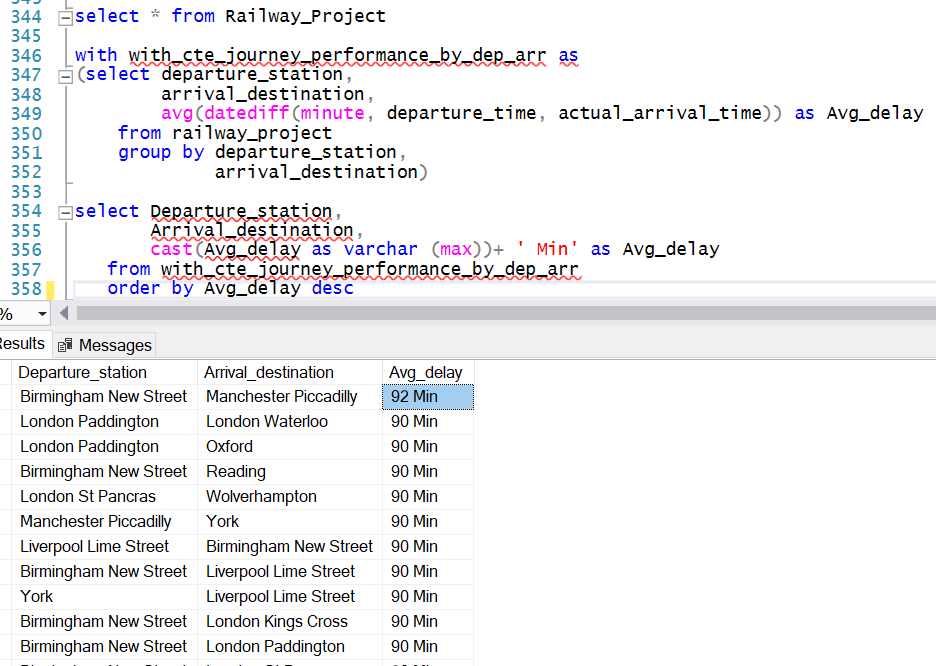
From the above query I can understand where Disabled and Senior railcard holders benefit from lower ticket costs and lower delay rates, making their journeys more economical and reliable and somehow punctuality of their journey.

Similarly, in case of adult holders are despite of paying moderately lower ticket prices compared to those without railcards experience the highest delay rates, it can be leads due to high no adult’s traveller and high rush.

None category railcard holder are pay the highest average ticket price but the delay percentage is 6.74%, indicating a moderate frequency of delays and I can assumed its quite satisfaction to their journey as compared to Adults holder.

**Query:5 Journey Performance by Departure and Arrival Stations:**

**Analysis:**



From the above interpretation I can assume most of the pair stations journey performances are extremely negative delay.

There is high in delay routes where it connects with Birmingham New Street and Liverpool Lime Street and London Paddington frequently experience high delays, indicating due to operational or traffic issues need to resolve better journey performance.

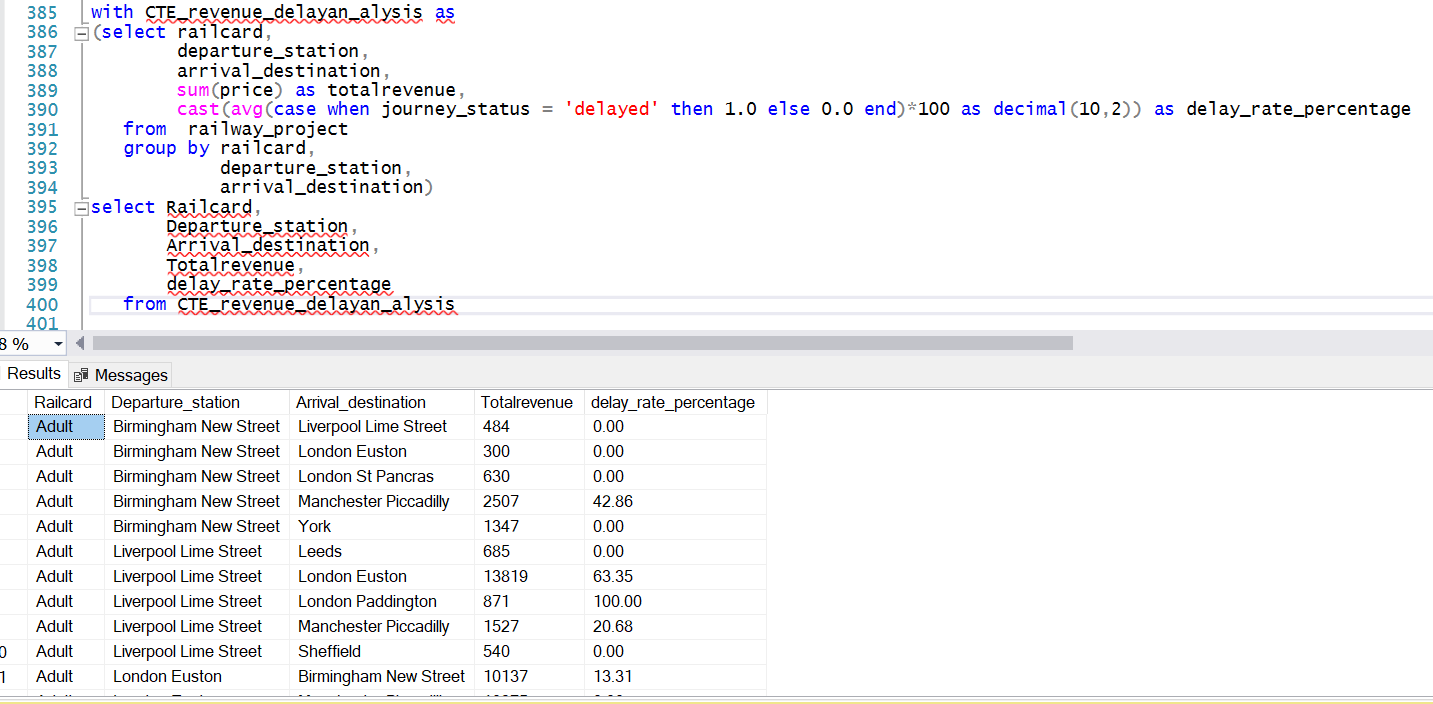
But in the case of the station from London Euston, Manchester Piccadilly, and York show a mix of early arrivals and moderate delays as compared to previous stations for better journey and customer performance need for better operational scheduling and traffic management.

Some routes like Reading, London Kings Cross, Bristol Temple Meads stations, and Coventry to Birmingham New Street manage to perform well with minimal delays or early arrivals. All these routes are indicating proper operational & traffic management. Also, I am assuming all these routes are performing delay management due to a mix of short and long distance but as a whole, all the routes are able to perform most low delay journey which is less than 40 minutes.

From this part of the interpretation, it's strongly recommended that railway authorities improve journey performance by addressing operational issues and traffic delays. Frequent delays can lead to cancellations or travellers opting for alternative transportation, negatively impacting railway revenue.

**Query:6 Revenue and Delay Analysis by Railcard and Station:**

**Analysis:**



From this interpretation I can come to some insights into how different railcards impact revenue and journey delays across various stations.

From Adult railcard: revenue contributions High revenue routes include Liverpool Lime Street to London Euston 13,819, London Euston to Birmingham New Street 10,137, and London Kings Cross to York 10,461, but some routes are high revenue but moderate delays on routes like Manchester Piccadilly to London Euston having 7,939 revenue with 44.86% delay rate. But in case of delay interpretation of adult railcard category, severe delays on routes such as Birmingham New Street to Manchester Piccadilly with (42.86%) and Liverpool Lime Street to London Euston with (63.35%) of delay rate but if I check complete delays on routes like Liverpool Lime Street to London Paddington with (100%) of delay rate and Manchester Piccadilly to Nottingham (100%) of delay.

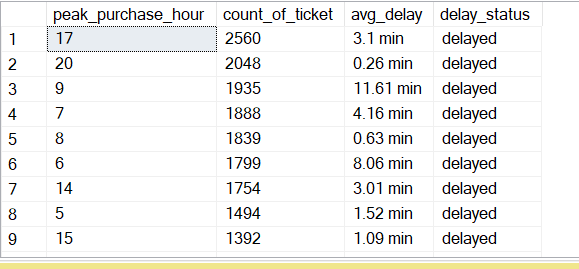
From disabled railcard: The high revenue routes include London Euston to Manchester Piccadilly with 10,829 of revenue and London Kings Cross to Edinburgh Waverley 6,084 and the moderate revenue from routes like Liverpool Lime Street to Manchester Piccadilly with 1,397 of revenue. But in case of delay interpretation, normally delay rates of this category are very low delay rates but specifically the routes from York to Wakefield with 100% of delay, rest of routes are trying to maintain as low rates of delay as possible which indicates efficient service for disabled railcard holders to be reliable and somehow punctuality for them.

From None railcard: For no holdings of any kinds of railcard contributing most highest revenue from London Kings Cross to York of 164,790 and Liverpool Lime Street to London Euston 97,376 of revenue but high revenue routes with significant delays include Liverpool Lime Street to London Euston of 71.74% of delay and Manchester Piccadilly to London Euston of 85.51% of delay. But interpretation of delay insights comes to the picture there are severe delays on certain high revenue routes such as London Euston to York with 100% and Manchester Piccadilly to London Euston of 85.51% delay, but some routes like London Paddington to Reading having minimal delays with 1.12% which is considerable.

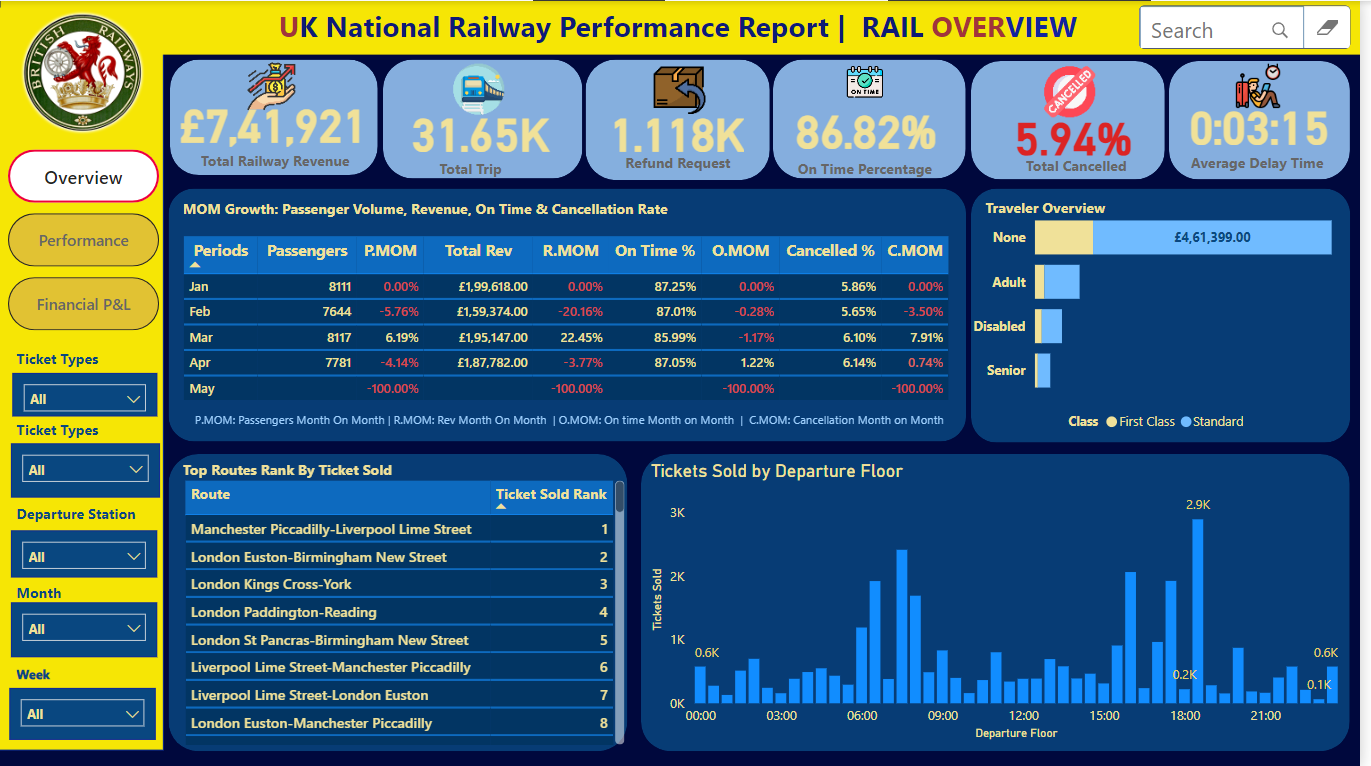
From Senior railcard: The senior category contributes the high revenue with no delays on routes like London St Pancras to Leicester with a revenue of 5,991 but on the other hand the moderate revenue from routes like Liverpool Lime Street to Crewe 1,563 and London Kings Cross to York 3,657. Now let's dive into the delay insights where significant delays on routes like Oxford to Bristol Temple Meads and Liverpool Lime Street to London Euston both are with 100% of delay, but most routes are low delays which indicate good management for the senior railcard holder to be reliable and punctuality and commitment for them.

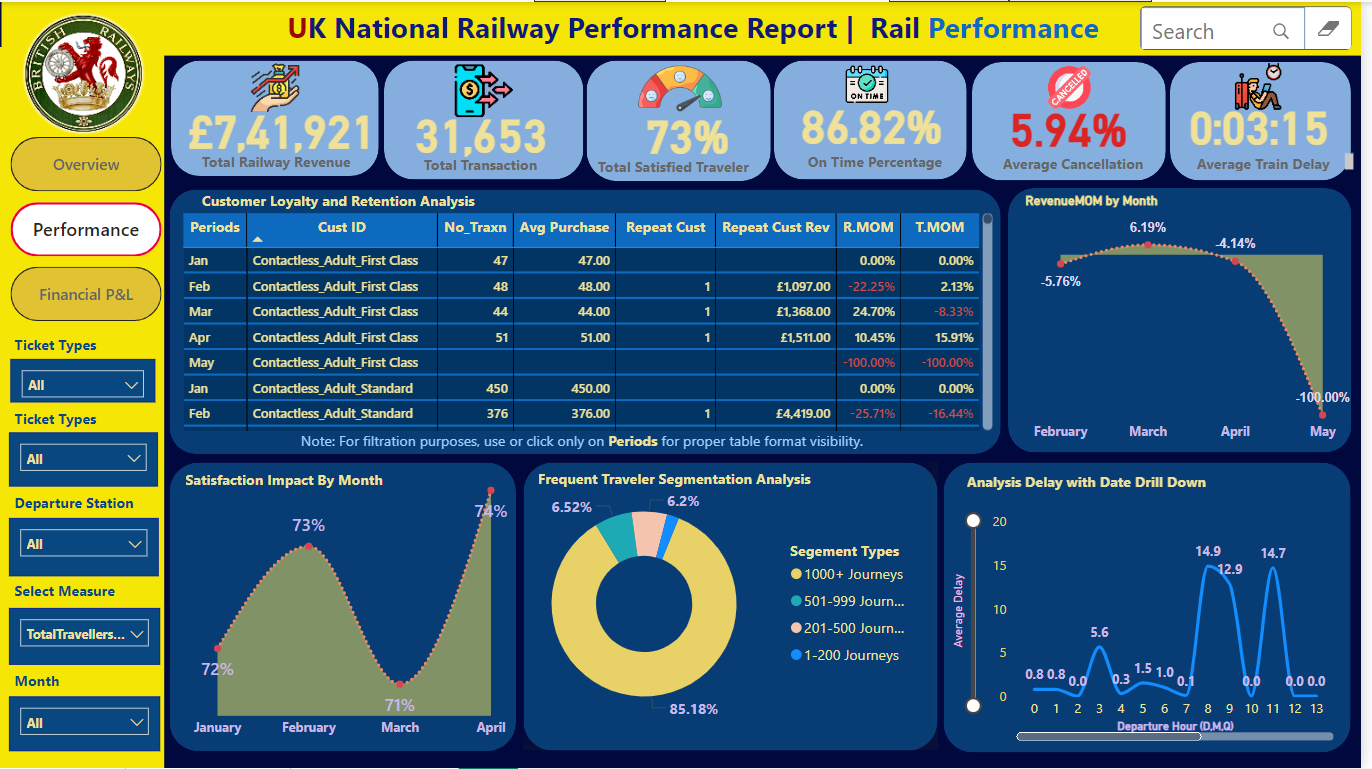
**Query:7 Journey Delay Impact Analysis by Hour of Day:**

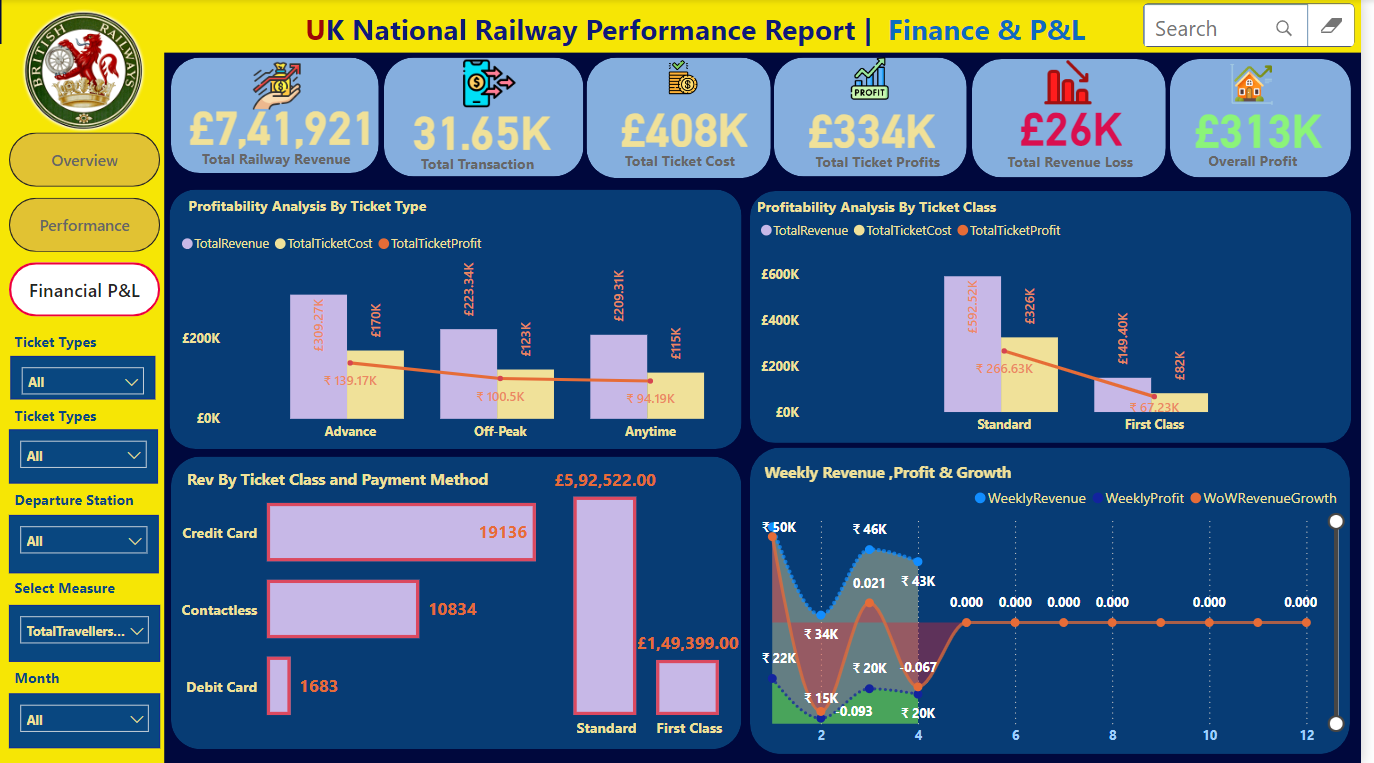
**Analysis:**

****

**Final Data Visualization Report:**







**Power Bi Live Dashboard Link: - (Use CTRL Click)**

[**https://app.powerbi.com/view?r=eyJrIjoiM2I0NjFlZDEtZGFlMi00MWE1LThiODEtZjc2OWMwZWZiNDUyIiwidCI6IjViNDYyNjIyLWI0ZDktNDk5OC05NGQ1LWNiMWJjMTljN2Y5NiJ9&pageName=31356b1948cb66f0c8e8**](https://app.powerbi.com/view?r=eyJrIjoiM2I0NjFlZDEtZGFlMi00MWE1LThiODEtZjc2OWMwZWZiNDUyIiwidCI6IjViNDYyNjIyLWI0ZDktNDk5OC05NGQ1LWNiMWJjMTljN2Y5NiJ9&pageName=31356b1948cb66f0c8e8)