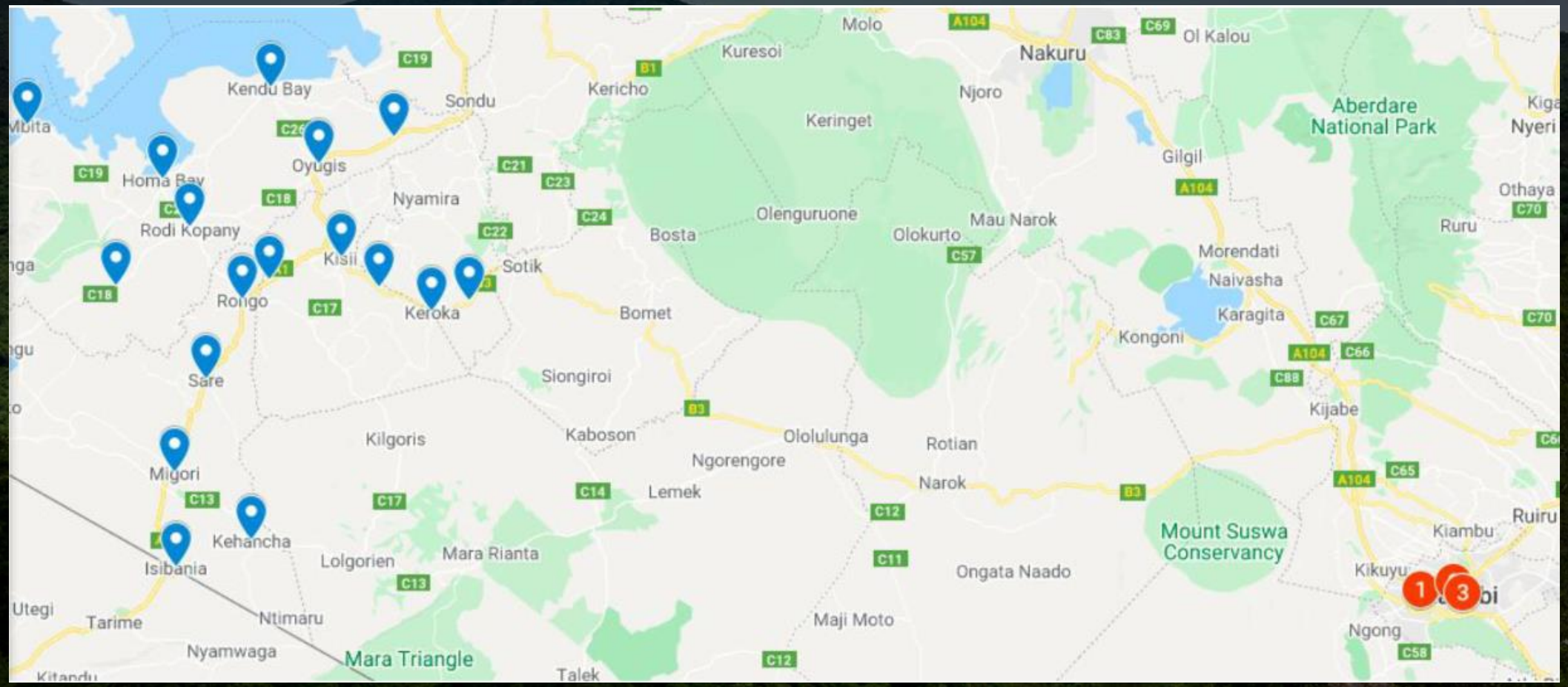
Nairobi Public Transportation Demand Prediction

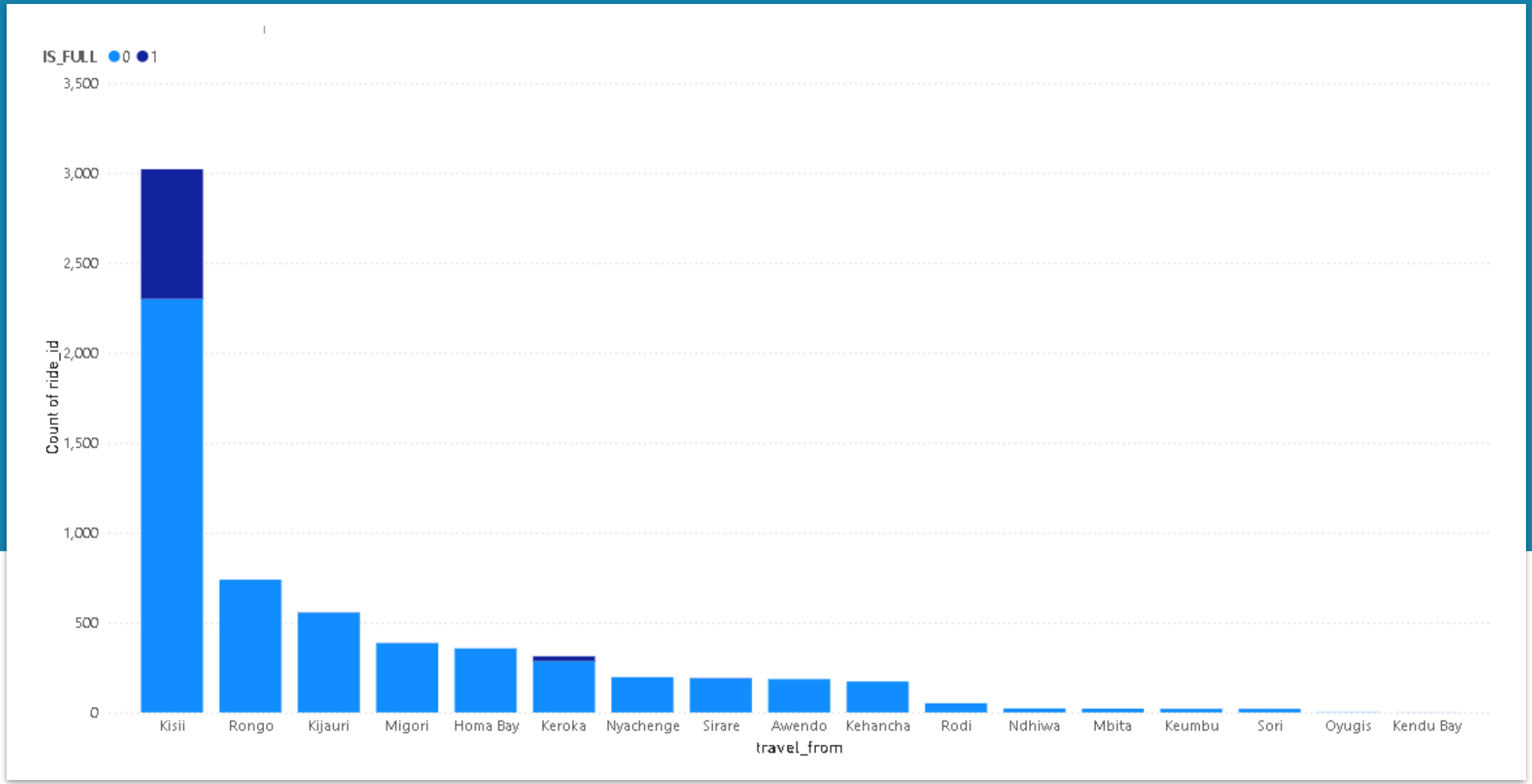
etisalat and

PROBLEM DEFINITION :
PREDICT THE NUMBER OF SEATS MOBITICKET CAN EXPECT TO SELL FOR EACH RIDE ON SPECIFIC ROUTES TO NAIROBI

Mobiticket

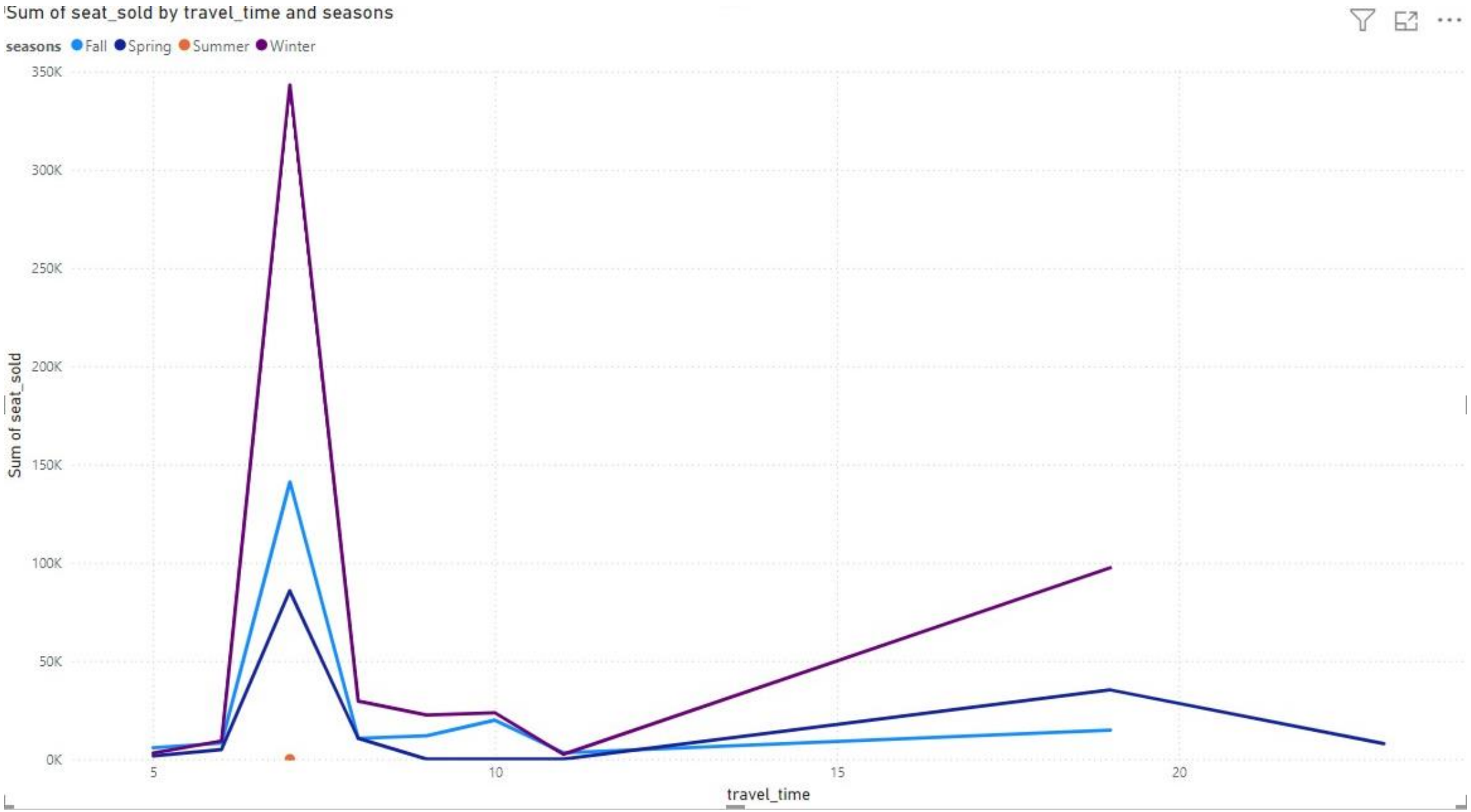
THE TOWNS FROM WHICH THESE ROUTES ORIGINATE ARE:




IS_FULL 0 1
3,500
3,000
2,500
2,000
1,500
1,000
500
0
Count of ride_id
Kisii
Rongo
Kijauri
Migori
Homa Bay
Keroka
Nyachenge
Sirare
Awendo
Kehancha
Rodi
Ndhiwa
Mbita
Keumbu
Sori
Oyugis
Kendu Bay
travel_from

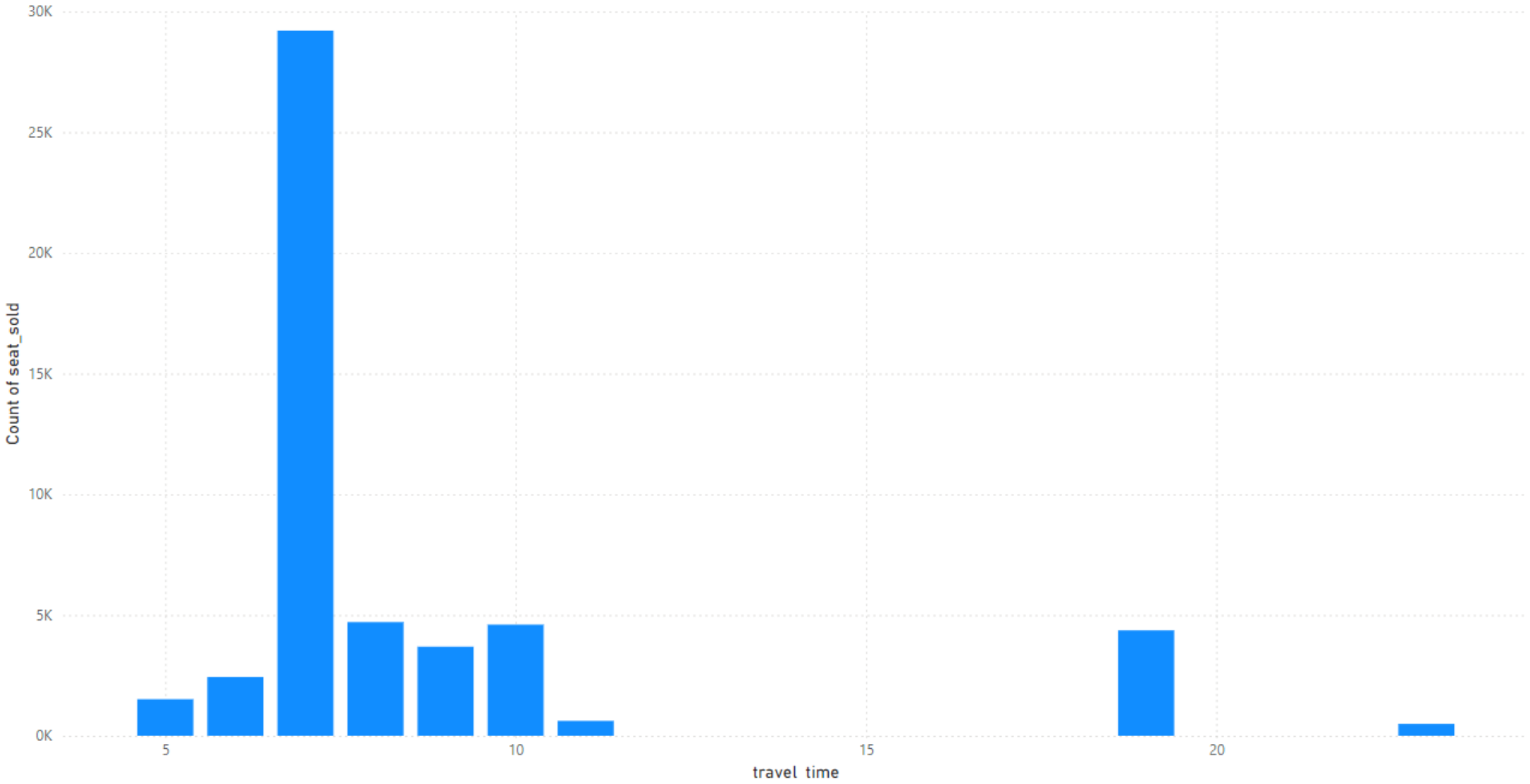
SEASONAL TRENDS IN TRANSPORTATION DEMAND

- In our analysis of seasonal seat sales, a notable trend has emerged: the percentage of seats sold during the summer is significantly lower compared to winter, fall and spring. This pattern suggests that transportation demand is closely linked to educational purposes.

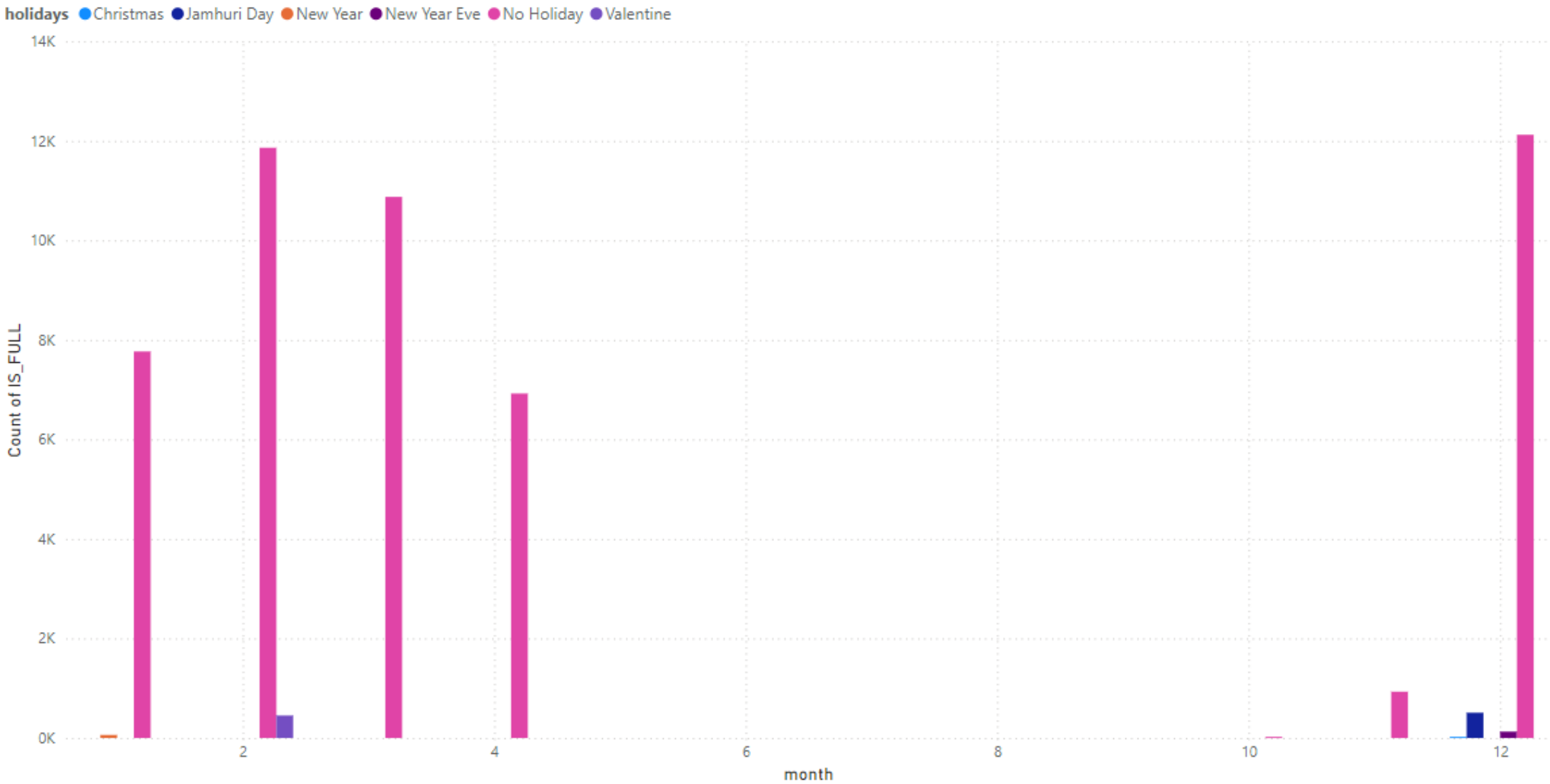

Sum of seat_sold by travel_time and seasons
seasons
Fall
Spring
Summer
Winter
350K
300K
250K
200K
150K
100K
50K
0K
Sum of seat_sold
5
10
15
20
travel_time

By examining the relationship between travel_date and seat_sold, we uncovered a seasonal pattern: number of tickets sold are much higher during the school year, suggesting that a large portion of our passengers are students

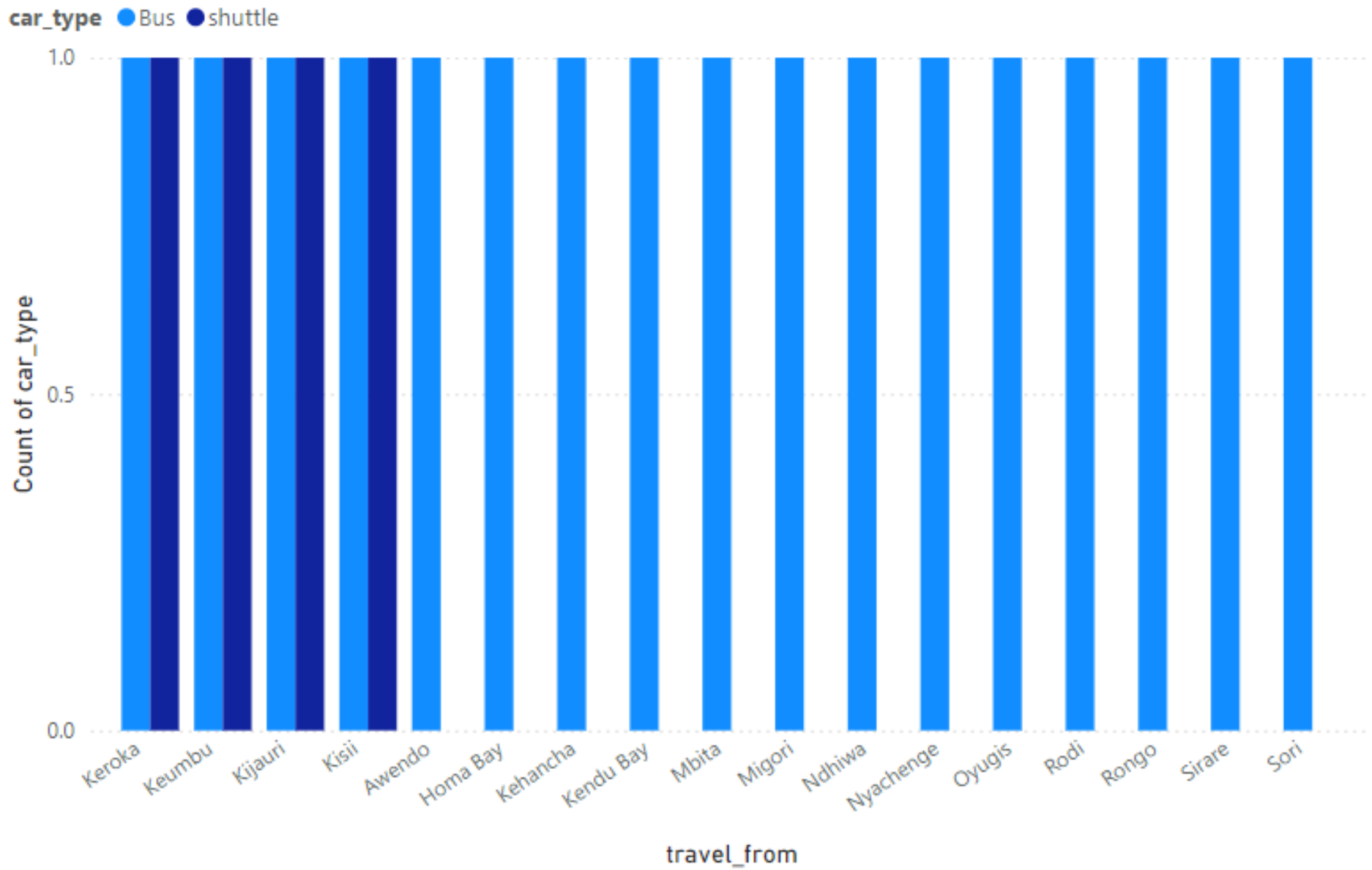
Insights to Determine When Rush Hour Occurs



The rides aren't affected by holidays

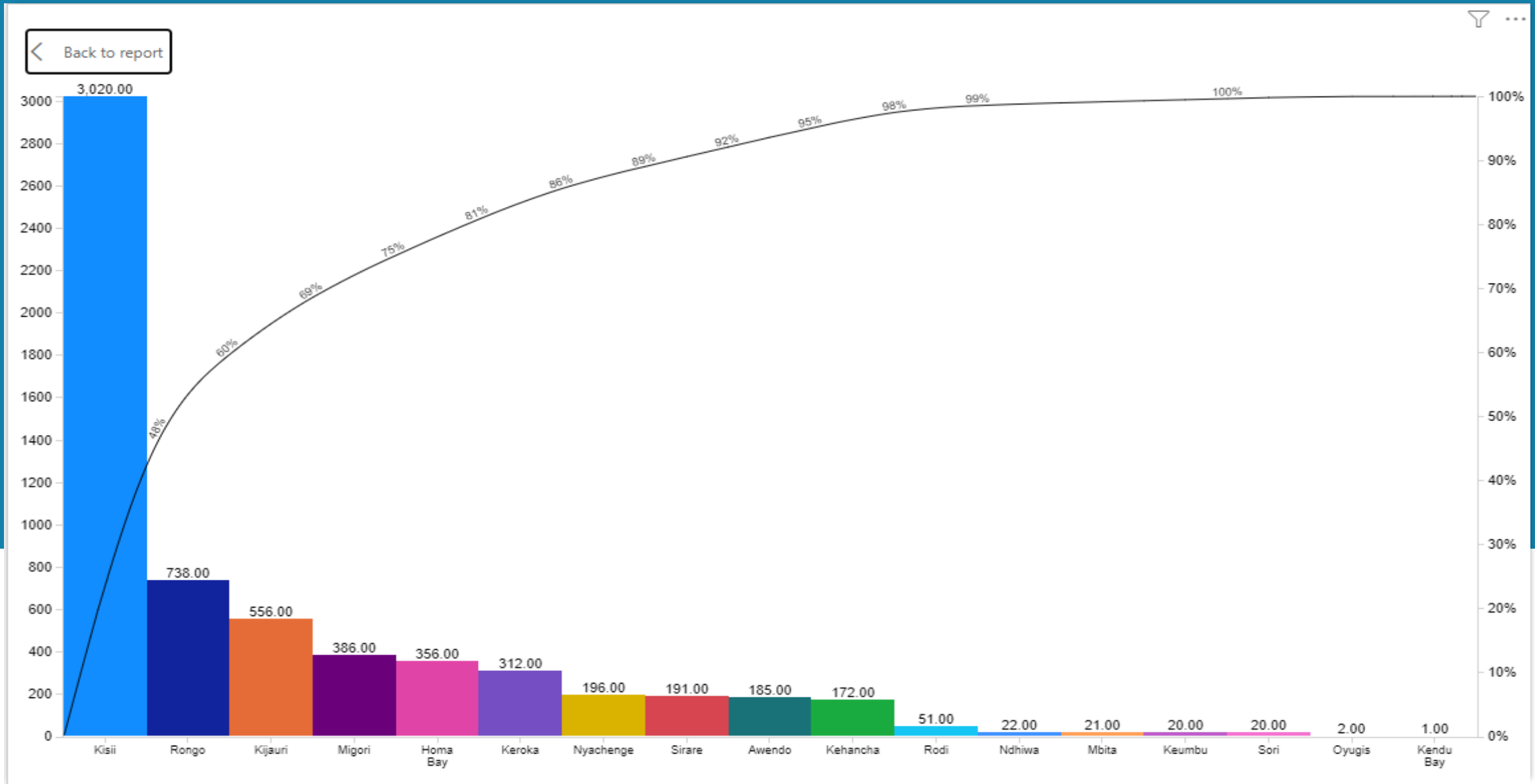


Only 4 towns use both shuttles and buses while, the others use buses only



ADDRESSING LOW TRAVEL DEMAND FROM CERTAIN TOWNS

In our analysis, we observed that the number of trips originating from certain towns, such as Sori, Keumbu, Oyugis, and Kendu Bay, is significantly lower compared to other locations like Kisii and Rongo. This discrepancy in travel demand presents a challenge, especially in optimizing routes and ensuring the profitability of services.


Back to report
3,020.00
738.00
556.00
386.00
356.00
312.00
196.00
191.00
185.00
172.00
51.00
22.00
21.00
20.00
20.00
2.00
1.00
48%
60%
69%
75%
81%
86%
89%
92%
95%
98%
99%
100%
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Kisii
Rongo
Kijauri
Migori
Homa Bay
Keroka
Nyachenge
Sirare
Awendo
Kehancha
Rodi
Ndhiwa
Mbita
Keumbu
Sori
Oyugis
Kendu Bay

RECOMMENDATIONS:

Cancellation: We can discontinue transportation services in these low-demand areas. This can help reduce unnecessary costs and allow resources to be reallocated to more profitable routes.

Reallocations : Shift resources, such as buses and staff, from these low-demand towns to cities with higher travel demand. This ensures that the transportation network remains focused on areas where it is most needed.

Shuttle-Only Service: Instead of operating large buses, convert the transportation service in these towns to a shuttle-only model. Shuttles are smaller, more flexible, and cost-effective, making them ideal for areas with low demand.

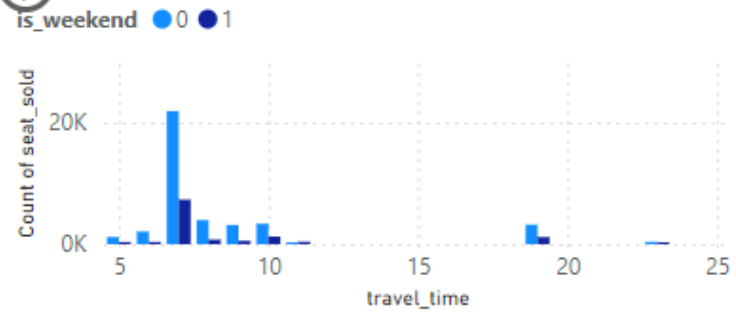
METHODS WE USED

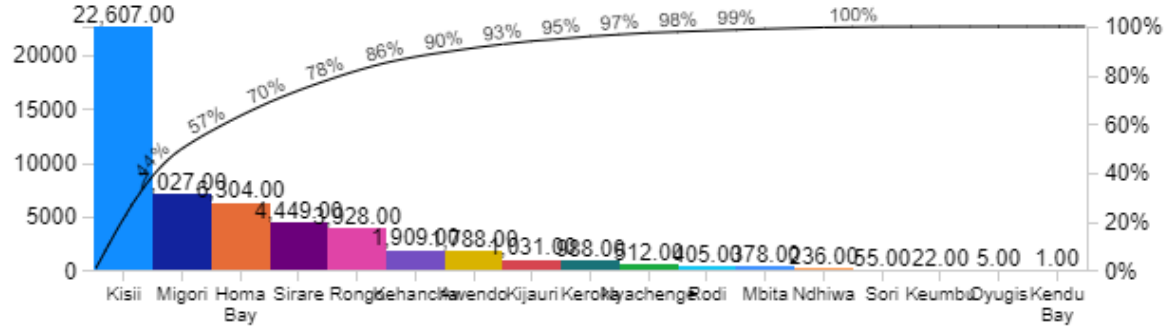
	XGBoost Regression	Random Forest
Mean Absolute Error(MAE)	3.318	2.94
R2 Score	0.81	0.832

We chose Mean Absolute Error (MAE) because it provides a simple, easy-to-understand measure of how accurate our predictions are. It tells us the average error in the number of tickets sold, which is useful for understanding how well our model is performing.

We chose XGBoost and Random Forest because they are strong at finding complex patterns in the data. These models are particularly good when dealing with multiple factors that influence ticket sales, like date, location, and car type

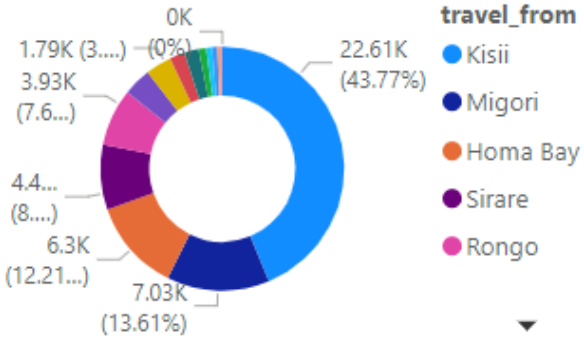
Count of seat_sold by travel_time and is_weekend



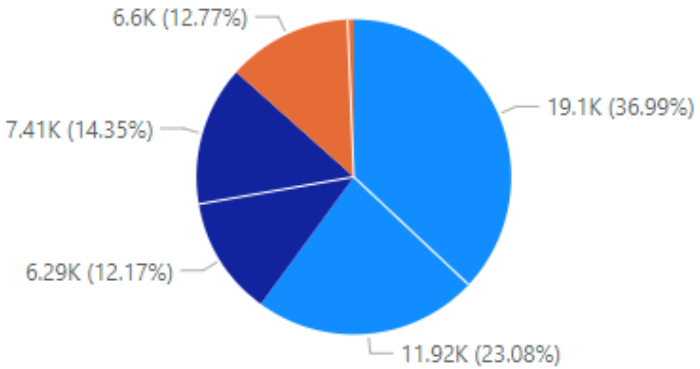


car_type	seat_sold	Count of ride_id
Bus	1	699
shuttle	1	759
Bus	2	343
shuttle	2	197
Total		**1998**

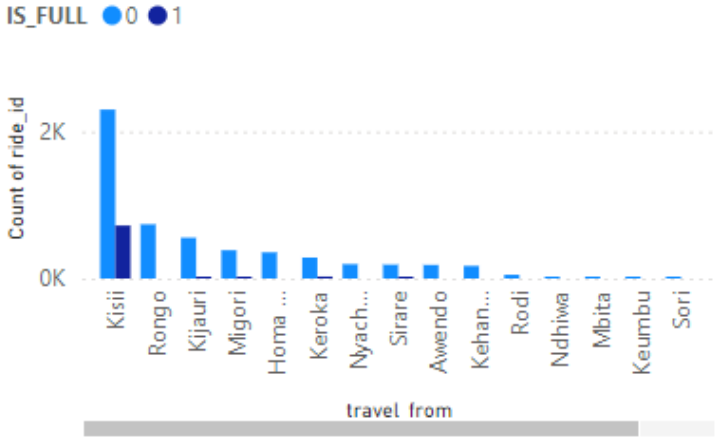
Count of IS_FULL by travel_from



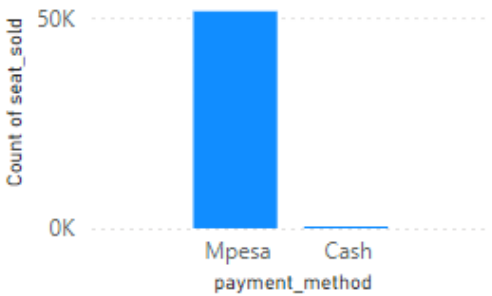
Count of seat_sold by seasons and car_type



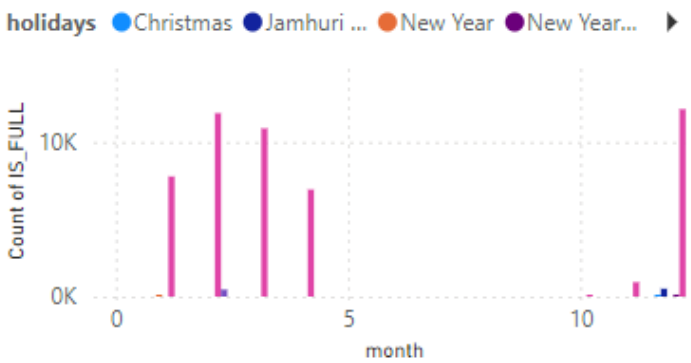
Count of ride_id by travel_from and IS_FULL



Count of seat_sold by payment_method



Count of IS_FULL by month and holidays



Count of seat_sold by Year