Staffordshire University

DECISION ANALYTICS

Hotel Revenue Loss and Data Analysis Report

Author:
Mohamed Nausath Fathima Ilma

January 5, 2024

Abstract

Hotels are a fast-diving industry in the current economic world. However, one of the major problems the hotel industry faces is the revenue loss caused by booking cancellations and clients not showing up even after reserving a room. Therefore, the hotel is unable to schedule that specific reservation for another. However, when the customer doesn't cancel or cancel the booking, it causes a revenue loss to the hotel.

To solve the problem mentioned above, a data set is analyzed to analyze and implement strategies to find solutions to overcome the problem. The main tools used for data analysis of this project the Excel and Power-BI.

Based on the overall analysis of the revenue loss, more people in the South region and Online bookings had led to either cancellations or No-shows. However, there are more cancellations than no-shows in the data set used.

Therefore, by analyzing each important key metric, certain solutions in terms of suggestions are brought up. For example, the company can have strict policies for online bookings, like reservations cannot be made without a deposit, or a penalty fee can be charged for cancellations or No-show instances. Diving deep into audience requirements in the South region to reduce the occurrence of this situation and many more!

Acknowledgment

I would like to present my sincere gratitude and thanks to my Decision Analytics lecturer, Dr. Rajitha Nawaratne, who covered the theories and practicals related to the subject very clearly. His inspiration and continuous support helped to make this report an insightful and successful report. He made sure always to be available to solve our doubts and help us throughout the lecture time.

Contents

List of Figures

1 Introduction

The hotel industry is a part of the service industry. The main tasks of hotels are handling the guests' accommodations and lodging, preparing the rooms, and managing the overall activities, including the reservations. Generally, the hotel industry welcomes travelers from both locals and tourists.

One of the main problems the hotel industry faces is the revenue loss caused by booking cancellations and no-shows. When the customer cancels a reservation later, the hotel can't offer it to anyone who has requested the same reservation earlier (as it's late). This mainly gets tight during both the seasonal and off seasons. Also, another category of customers doesn't show up even after reserving, so they don't cancel the booking. Due to these unavoidable circumstances, it cause revenue loss.

In this data analysis, we will discuss each scenario in detail: the overall revenue loss, revenue loss caused by cancellations, the revenue loss caused by the show, and the solutions to overcome this problem.

2 Pre-Processing Work Performed

Before getting the insights, the following is pre-processing work performed on the data set.

2.1 Check For Duplicates And Null values

The data set was thoroughly checked for **duplicate** and **null** rows and columns. However, there were no duplicates, and null rows and columns were found.

2.2 Check For Data Accuracy

The data set was checked for data accuracy, and some values were less accurate.

The following are the accuracy issues found and solved in the given data set:

1. **Booking date was equal to check-out date:** There were rows where the 'Booking date' was equal to the 'check-out date' of the respective row. For example, the reservation ID of "45171160" had this problem where the booking date was **equal** to the check-out date. Such rows were filtered by creating a new column, namely "Check_BookingDate_And_CheckoutDate." The formula "=[Booking_date]=[Expected_checkout]" was applied. The columns with the "False" value were filtered.

2. **Check-in date was before the booking date:** For example, the reservation id "85985887" has the booking date as "6/2/2016" and the check-in date as "6/1/2016". Considering this scenario, a person can arrive at the hotel on the **same day** as the booking date or before the booking date. But, as per this example, the individual has arrived before the booking date, which is a data anomaly. Therefore, such rows were removed. To identify these columns, the formula "=[Booking_date]¡=[Expected_checkin]" and the columns with "True" values were removed.

The above methods were used to remove the Date outliers in the data set. Removing outliers helps clean the data set properly.

2.3 Data Filtering

As per the given question, the data should be filtered with only 15 months of data. Initially, the 'Booking_date' column was sorted from the 'Newest to Oldest.' Based on the results of finding the most recent date the data set has, the 'Booking_date' column is filtered to the dates between '6/1/2016' and '8/30/2017'. Meanwhile, the months between June 2016 and August 2017 include 15 months of data.

Therefore, after filtering the data according to the above guidelines, there were 6057 rows.

2.4 Maintaining Data Uniformity

The data headers are properly organized to maintain the standard and uniformity of data. For example, there were headers, namely "Room-rate" and "Checkin_Date." Therefore, to maintain uniformity, all the headers were changed according to the below format:

- **Capitalizing** every word (by maintaining Title Case). Example: Room_Rate and Checkin_Date.
- **Underscore** is maintained across all words as a good practice. Therefore, it is easy to read and understand.
- **Meaningful** header names are named. Therefore, it is easy to understand even if you review this data set after a couple of months or years.

2.5 Adding New Columns For Data Analysis

The following new columns were added to calculate the Revenue loss:

- **No_Of_People:** To calculate the number of people who would attend for the specific booking (adding the adults and children)
- **No_Of_Days_Booked:** To calculate the number of days booked (based on the expected check-in and check-out date)

- **No_Of_room:** Calculate the number of rooms booked for the specific booking (based on the number of people who attend)
- **Final_Roomrate:** Calculate the final room rate per booking (based on the number of rooms booked, no of days booked, room rate, and the discount offered)
- **Age_Group:** Group the given ages into specific age categories (convenient for data analysis based on ages)

3 Statistical Summary

The statistical summary involves the statistics data based on the **past 15 months of data**. The data used to generate the statistical summary is after pre-processing the data (based on the steps mentioned above).

These data include the 15 months of data, **being it a check-in, no-show, or canceled** (including all these 3 criteria).

The following figure represents the statistical summary of 15 months of data:

	Min	Max	SD	Count	Sum	Median	Mode	Q1	Q3
Reservation-Id	39530	99983431	28553986.9	6057	3E+11	50252661	55371044	25296167	75514365
Gender	-	-	-	0	-	-	F	-	-
Age	18	70	15.3183862	6057	267073	44	34	31	57
Ethnicity	-	-	-	6057	-	-	African American	-	-
Educational_Level	College	Grad	-	6057	-	-	College	-	-
Income	<25K	50K -- 100K	-	6057	-	-	50K -- 100K	-	-
Country_Region	-	-		6057	-	-	South	-	-
Hotel_Type	-	-		6057	-	-	Airport Hotels	-	-
Expected_Checkin	1/1/2015	9/6/2017		6057	-	-	-	-	-
Expected_Checkout	1/2/2015	9/9/2017		6057	-	-	-	-	-
Booking_Date	6/11/2014	8/30/2017		6057	-	-	-	-	-
Adults	1	5	1.19128133	6057	14209	2	2	2	3
Children	1	3	0.71936387	6057	10567	2	1	1	2
Babies	0	2	0.57668003	6057	2120	0	0	0	1
Meal_Type	-	-	-	6057	-	-	BB	-	-
Visted_Previously	-	-	-	6057	-	-	No	-	-
Previous_Cancellations	-	-	-	6057	-	-	No	-	-
Deposit_Type	No deposit	Non-Refundable		6057	-	-	No Deposit	-	-
Booking_Channel	-	-	-	6057	-	-	Online	-	-
Required_Car_Parking	-	-	-	6057	-	-	Yes	-	-
Reservation_Status	-	-	-	6057	-	-	Check-In	-	-
Use_Promotion	-	-	-	6057	-	-	Yes	-	-
Check_BookingDate_And_CheckinDate	-	-	-	6057	-	-	TRUE	-	-
Check_BookingDate_And_ExpectedCheckOutDate	-	-	-	6057	-	-	FALSE	-	-
Discount_Rate	0	40	11.227684	6057	76605	10	0	5	20
Room_Rate	100	250	44.3666269	6057	1058565	175	168	137	214
No_Of_People	2	8	1.38151724	6057	24776	4	4	3	5
No_Of_Days_Booked	1	4	0.97967557	6057	11045	1	1	1	2
No_Of_Rooms	1	2	0.36920157	6057	7043	1	1	1	1
Final_Roomrate	61.2	1944	239.665885	6057	1966389	240	216	160.2	417.6

Figure 1: Statistical summary of 15 months of data

3.1 Assessing The Statistical Summary

The **numeric value** columns in this data set are Reservation Id, Age, Adults, Children, Babies, Discount_Rate, Room_Rate, No_Of_People, No_Of_Days_Booked, No_Of_Rooms, and Final_Roomrate.

The **date column values** in this data set are Expected_Checkin, Expected_Checkout, and Booking_Date.

The **binary value** column in this data set is Gender (M/F), Visted Previously (Yes/No), Previous Cancellations (Yes/No), Required_Car Parking (Yes/No), Use Promotion (Yes/No), Checked BookingDate And CheckinDate (Yes/No), and Check BookingDate And ExpectedCheckOutDate (Yes/No).

The **ordinal value** columns in this data set are Educational_Level, Income, and Deposit_type.

The **nominal value** columns in this data set are Deposit_type, Booking_channel, Ethnicity, Country_region, Hotel_Type, Booking_channel, and Reservation_Status.

The statistical summary is represented under 9 categories as follows:

- **Min** (indicates the Minimum value of the respective data column): Since the same sheet has been used from the beginning for filtration, we cannot use the "Min" function. Therefore, the "SUBTOTAL(5, range)" function is used across all the numeric and date values columns. The number "5" indicates the "minimum value" for the filtered columns. Also, manually, the minimum value is considered for nominal value data types.

- **Max** (indicates the Maximum value of the respective data column): The "=SUBTOTAL(4, range)" function is used across all the numeric values columns and date values columns. The number "4" indicates the "Maximum value" for the filtered columns. Besides that, the minimum value is manually considered for nominal value data types.

- **SD** (indicates the Standard Deviation of the respective data column): The "=SUBTOTAL(7, range)" function is used across all the numeric values columns. The number "7" indicates the "Standard Deviation" for the filtered columns.

- **Count** (specify the total count of rows): The "=SUBTOTAL(2, range)" function is used here. The number "2" shows the "Count" for filtered columns.
- **Sum** (specify the sum of total rows): The "=SUBTOTAL(9, range)" function is used here for the numeric values columns. The number "9" shows the "Sum" for the filtered columns.
- **Median** (specify the Median of the respective column): The "=MEDIAN(range)" function is used here for numeric values columns.
- **Mode** (specify the Mode of the respective column): The "=MODE(range)" function is used here for the numeric values columns. The formula "=INDEX(range, MODE(MATCH(range: range)))" is helps to find the Mode for binary, ordinal, and nominal data types.
- **Q1** (specify the 1st Quartile of the respective column): The "=QUARTILE(range,1)" function is used across all the numeric values columns. The number "1" indicates the "1st Quartile" for the filtered columns.
- **Q3** (specify the 3rd Quartile of the respective column): The "=QUARTILE(range,3)" function is used across all the numeric values columns. The number "3" indicates the "3rd Quartile" for the filtered columns.

4 Answering Questions

4.1 What Is Revenue Loss Due To Booking Cancellations?

The revenue loss caused only due to the "Booking cancellations" is **"203988.85"**. The final sum value is based on the "Reservation_Status " filtered only for the "**Canceled** " column. This is because only these data types were properly canceled by the customer and recorded in the data set.

Type of Data	Value
Revenue loss due to cancellation	203988.85

Figure 2: Revenue loss due to booking cancellations

The overall revenue loss is taken considering four scenarios as follows:

- **Revenue caused due to "Cancellations" which have "No Deposit":** When there are Cancelled bookings, and the client has not deposited a pre-payment, the hotel bears a loss. It is because the hotel couldn't reserve the seat for another previously and had no monetary advantage.
- **Revenue caused due to "Cancellations" which are "Refundable":** When there are Cancelled bookings, and the client has paid under a Refundable condition, the hotel repays the pre-payment. After the booking gets canceled. However, the hotel bears a loss. The hotel couldn't reserve the seat for another previously and had to return the money.
- **Revenue caused due to "No-Show " which have "No Deposit":** When there are No-Show bookings, and the client has not deposited a pre-payment, the hotel bears a loss. The client didn't check in to the hotel or make any pre-payment. Therefore, the hotel couldn't decide whether the client would come and had to wait until the expected check-in and check-out date to reserve a seat for another previously.
- **Revenue caused due to "No-Show " which have "Refundable":** When there are No-Show bookings, they are paid under a Refundable condition. Thereby, after the booking is canceled, the hotel repays the pre-payment. And bears a loss. It is because the client didn't check in to the hotel and had to pay back the pre-payment. Therefore, the hotel couldn't decide whether the client would come and had to wait until the expected check-in and check-out date to reserve a seat for another previously.

The following diagram represents the revenue loss based on **"Cancellations" and "No-Show"**:

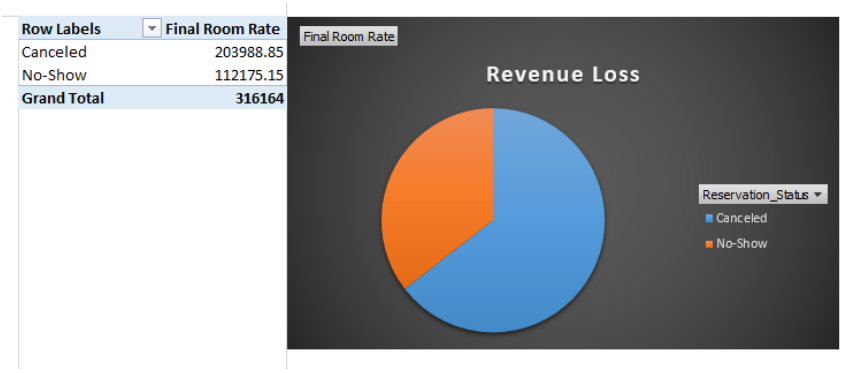

Figure 3: Revenue loss

The following diagram represents the revenue loss based on **"Cancellations" and "No-Show," including the "No Deposit" and "Refundable"**:

Row Labels	Final Room Rate
Canceled	**203988.85**
No Deposit	145039.65
Refundable	58949.2
No-Show	**112175.15**
No Deposit	82413.55
Refundable	29761.6
Grand Total	**316164**

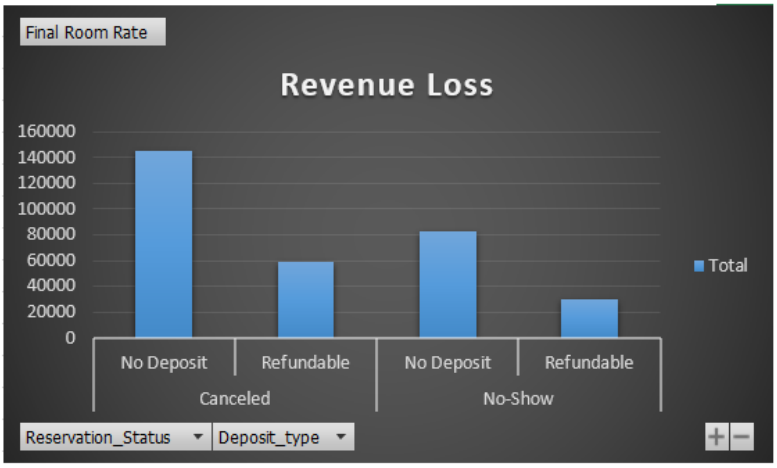

Figure 4: Revenue loss based on "No Deposit" and "Refundable"

4.2 What Is The Revenue Loss In Terms Of Different Breakdowns?

4.2.1 Gender

The following diagram represents the revenue loss based on "Gender":

Row Labels	Final Room Rate
F	157691.5
M	158472.5
Grand Total	**316164**

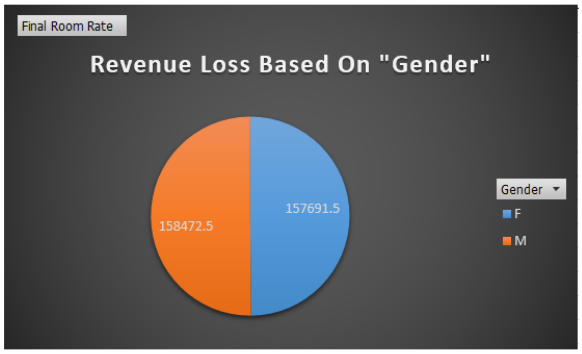

Figure 5: Revenue loss based on "Gender"

Insight: The above analysis shows that **more males** have been responsible for revenue loss than females (comparatively). Even though there is only a slight difference between the loss caused by Males and Females.

4.2.2 Age

The following diagram represents the revenue loss based on "Age" under several categories of age group:

Age Group	Final Room Rate
18-24	45377.85
25-34	56724.3
35-44	74907.05
45-54	56981
55-64	52108.65
65+	30065.15
Grand Total	**316164**

Figure 6: Revenue loss based on "Age"

Insight: Based on the above analysis, we can determine that the revenue loss caused between the age range of **35-44 is higher,** which is "74,907.05"

.

The following diagram represents the revenue loss based on the 'Count' on "Age":

Age Group	Count of Final_roomrate
18-24	120
25-34	180
35-44	198
45-54	183
55-64	175
65+	98
Grand Total	**954**

Figure 7: Revenue loss based on the 'Count' on "Age"

Insight: The number of people in the age group of 35-44 is 198. We can conclude that the age groups between **25 and 64** have mainly caused the revenue loss. However, if we sort the data in the descending order, it can be determined as follows:

- 35-44
- 45-54
- 25-34
- 55-64
- 18-24
- 65+

4.2.3 Ethnicity

The following diagram represents the overall revenue loss based on "Ethnicity":

Row Labels	Sum of Final_roomrate
African American	82753.35
Asian American	80629.9
caucasian	75838
Latino	76942.75
Grand Total	**316164**

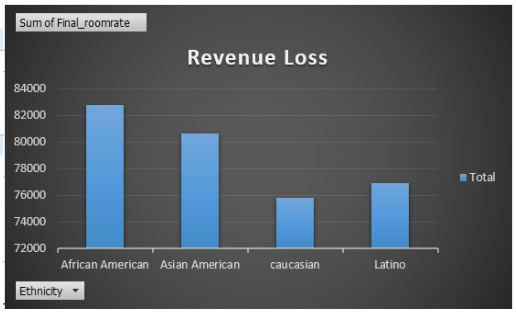

Figure 8: Revenue loss based on "Ethnicity"

Insight: African Americans and Latinos have been one of the main reasons causing the revenue loss comparatively to other Ethnic groups.

The above analysis is more detailed in the bar charts below.

The following diagram represents the revenue loss based on "Ethnicity" and "Canceled":

Row Labels	Final Room Rate
Canceled	**203988.85**
African American	45947
Asian American	59041.7
caucasian	50730.65
Latino	48269.5
Grand Total	**203988.85**

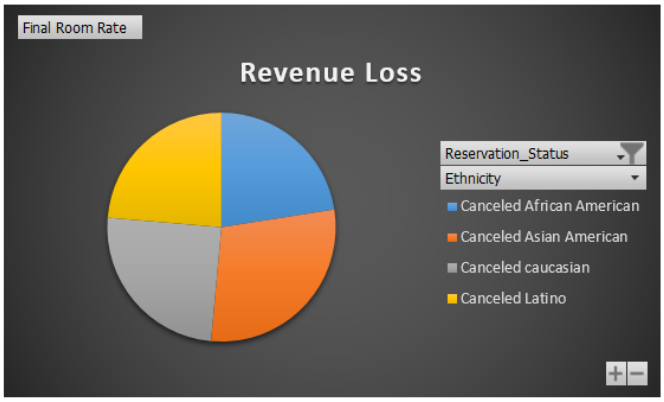

Figure 9: Revenue loss based on "Ethnicity" and "Cancelled"

The following diagram represents the revenue loss based on "Ethnicity" and "No-Show":

Row Labels	Final Room Rate
No-Show	**112175.15**
African American	36806.35
Asian American	21588.2
caucasian	25107.35
Latino	28673.25
Grand Total	**112175.15**

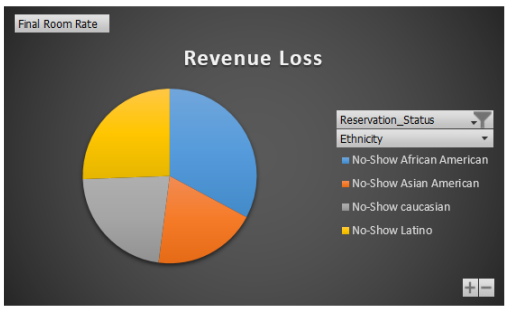

Figure 10: Revenue loss based on "Ethnicity" and "No-Show"

4.2.4 Educational Level

The following represents the overall revenue loss caused based on the "Education Level":

Row Labels	Sum of Final_roomrate
College	123909.7
High-School	76371.1
Mid-School	62037
Grad	53846.2
Grand Total	**316164**

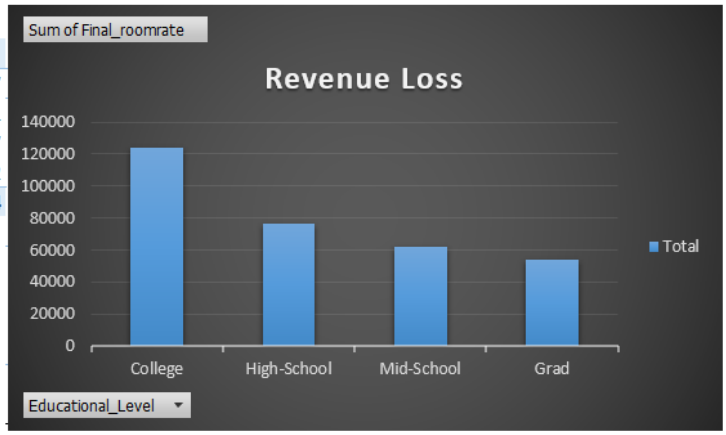

Figure 11: Revenue loss caused based on the "Education Level"

Insight: More **College** students have been the reason for revenue loss. Next comes the high-school students. One reason for this could be the "no deposit" and "refundable" options leading to irrelevant or careless bookings by these individuals. Stricting these policies can emphasize fewer booking cancellations.

The following diagram represents the revenue loss based on "Education Level" in terms of "Canceled" and "No-Show":

Row Labels	Sum of Final_roomrate	Count
Canceled	**203988.85**	**626**
College	84325.75	266
High-School	49305.65	137
Mid-School	36490.25	119
Grad	33867.2	104
No-Show	**112175.15**	**328**
College	39583.95	121
High-School	27065.45	71
Mid-School	25546.75	71
Grad	19979	65
Grand Total	**316164**	**954**

Figure 12: Revenue loss based on "Education Level" in terms of "Canceled" and "No-Show"

Insight: Even considering the above analysis, the **College** and **High School** students have been the primary reason for booking cancellations, and No show caused revenue loss.

4.2.5 Income

The following diagram represents the overall revenue loss based on "Income":

Row Labels	Sum of Final_roomrate
College	123909.7
High-School	76371.1
Mid-School	62037
Grad	53846.2
Grand Total	**316164**

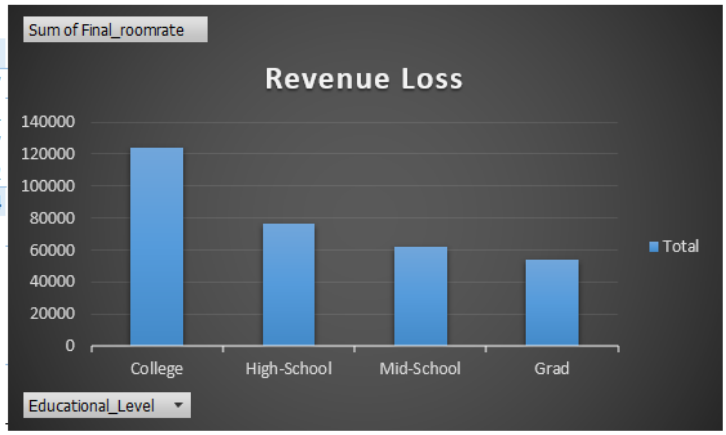

Figure 13: Revenue loss based on "Income"

Insight: The highest revenue loss has been caused by the income range of "**25K - 50K**" and "**50K-100K**". So, overall, people who earn between "25K and 100K" have caused a diverse area of revenue loss.

4.2.6 Country Region

The following diagram represents the revenue loss based on "Country Region":

Row Labels	Sum of Final_roomrate
East	63671.55
North	59099.1
South	129478.45
West	63914.9
Grand Total	**316164**

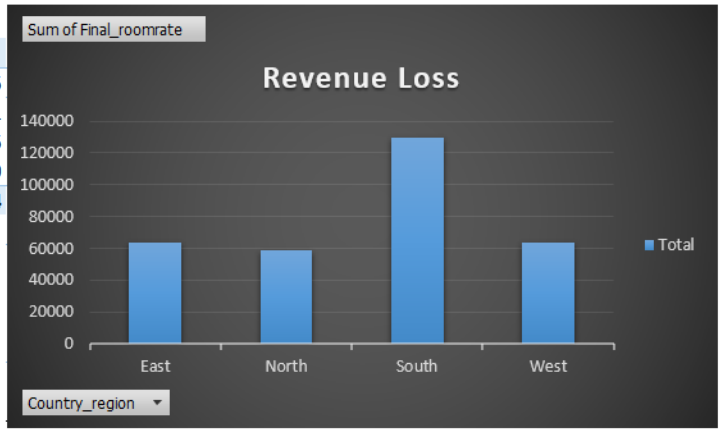

Figure 14: Revenue loss based on "Country Region"

Insight: Overall, the individuals in the **South** had been the primary cause of Revenue loss. Next comes the East (with comparatively less revenue loss). However, the South region should be given more attention and strategies to overcome the overall revenue loss.

4.2.7 Hotel Type

The following diagram represents the revenue loss based on "Hotel Type":

Row Labels	Sum of Final_roomrate
Airport Hotels	106199.3
City Hotel	102197.2
Resort	107767.5
Grand Total	**316164**

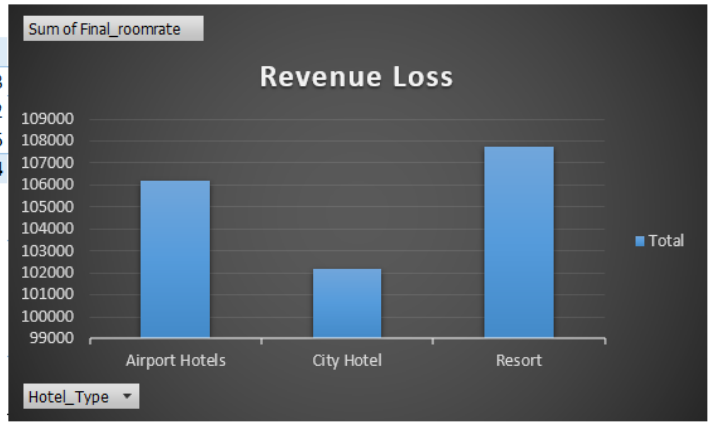

Figure 15: Revenue loss based on "Hotel Type"

Insight: Considering the insight, more people have **issues** with **Resorts** and **Airport hotels**.
Solution: Reasonable packages, effective promotions, feedback analysis, and target market activities can be implemented for Resorts and Airport hotels to overcome the revenue loss.

4.2.8 Meal Type

The following diagram represents the revenue loss based on "Meal Type":

Row Labels	Sum of Final_roomrate
BB	126635.1
FB	82349.5
HB	107179.4
Grand Total	**316164**

Figure 16: Revenue loss based on "Meal Type"

Insight: It seemed like most people **prefer the meal type "BB,"** and the second most is "HB." Therefore, effective **food-based promotional strategies** for these food types can be implemented.

4.2.9 Visited Previously

The following diagram represents the revenue loss based on "Visited Previously":

Row Labels	Sum of Final_roomrate
No	279250.3
Yes	36913.7
Grand Total	**316164**

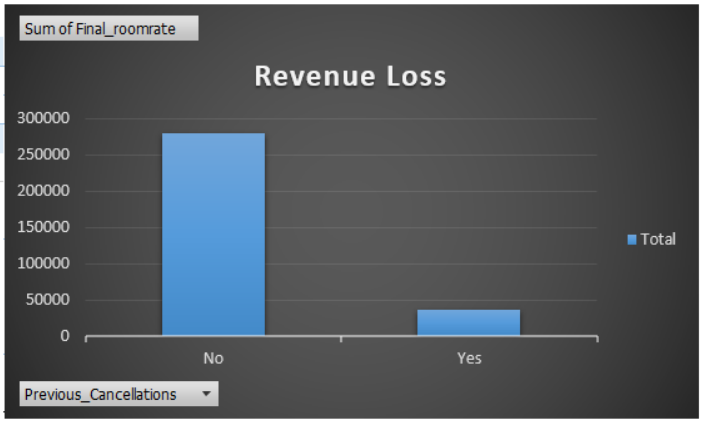

Figure 17: Revenue loss based on "Visited Previously"

Insight:

- Most of the people who had **not visited previously had canceled** the order. It means that they have less trust and confidence in the hotel. Therefore, effective **strategies** and **awareness** can be implemented to build **trust** in people. Therefore, it would help to decrease the revenue loss.
- Since some people who have visited have canceled the bookings, we can ask for **feedback** and the **reason** for cancellations in a polite manner. It helps to identify the problems faced by such individuals.

4.2.10 Previous Cancellations

The following diagram represents the revenue loss based on "Previous Cancellations":

Row Labels	Sum of Final_roomrate
No	279250.3
Yes	36913.7
Grand Total	**316164**

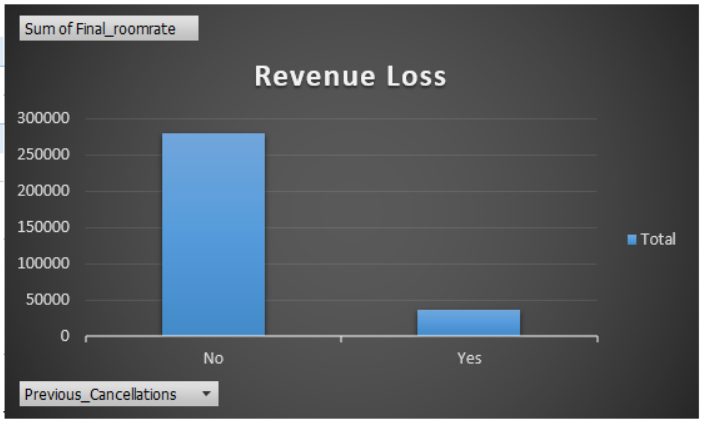

Figure 18: Revenue loss based on "Previous Cancellations"

Insight: More people who have **not previously canceled have canceled** or **didn't appear** for the hotel. Therefore, it is more likely that implementing proper **strategies** would work out for the hotel to reduce the revenue loss.

4.2.11 Deposit Type

The following diagram represents the revenue loss based on "Deposit Type":

Insight:

- The **absence of deposits** has caused more revenue loss. Therefore, the hotel should encourage this practice less.

Row Labels	Sum of Final_roomrate
No Deposit	227453.2
Refundable	88710.8
Grand Total	**316164**

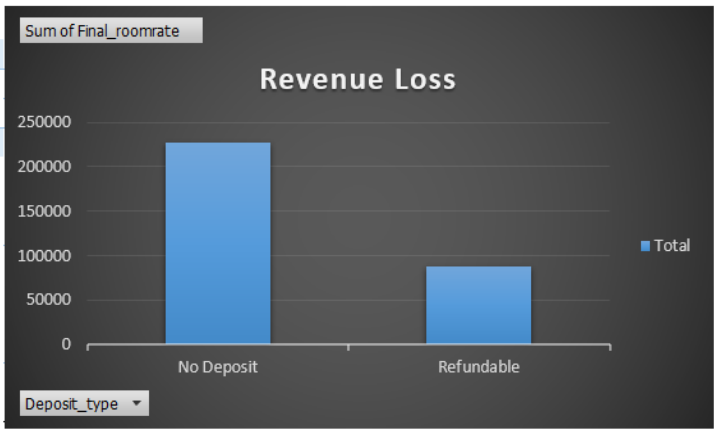

Figure 19: Revenue loss based on "Deposit Type"

- A considerable revenue loss is caused due to making deposits under a **Refundable** rule. Even though they are less likely to cancel or not show, hotels can implement the above strategies to reduce the overall revenue loss caused.

Solution:

- Hotels can implement strategies like previously, and people who have booked can re-book with no deposits.
- Hotels can provide more incentives or discounts for individuals who secure their reservation under a deposit.
- Therefore, they are less likely to cancel or not appear at the hotel. Also, if they refuse, you can retain more loyal customers.
- Introduce a penalty fee for Refundable deposits.

4.2.12 Booking Channel

The following diagram represents the revenue loss based on "Booking Channel":

Row Labels	Sum of Final_roomrate
Agent	46913.7
Direct	87120.75
Online	182129.55
Grand Total	**316164**

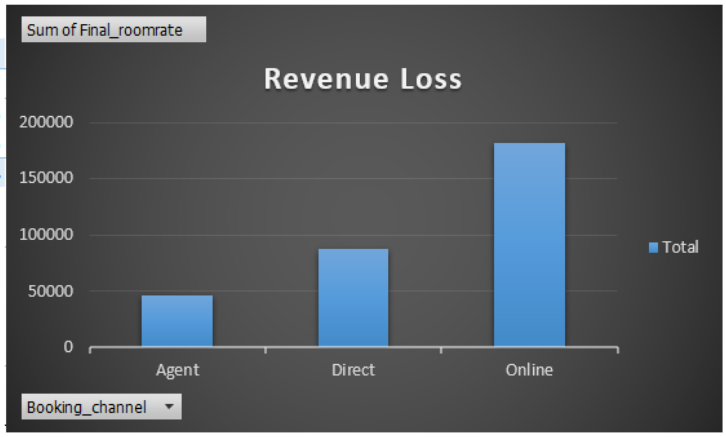

Figure 20: Revenue loss based on "Booking Channel"

Insight:

- More revenue loss has been caused due to **Online** booking as it is the easiest way to secure payments.
- Direct and agent bookings cause less revenue loss as they seem more trustworthy sources.

Solution:

- Implement measures to mitigate online booking cancellations, such as optimizing the website's user experience, providing clear cancellation policies, and
- Offering exclusive incentives for direct and agent bookings to encourage these channels.
- Provide more incentives and discounts for the bookings made through direct and agents. Also, diversify these two options to get more sales.

4.2.13 Required Car Parking

The following diagram represents the revenue loss based on "Required Car Parking":

Row Labels	Sum of Final_roomrate
No	86784.85
Yes	229379.15
Grand Total	**316164**

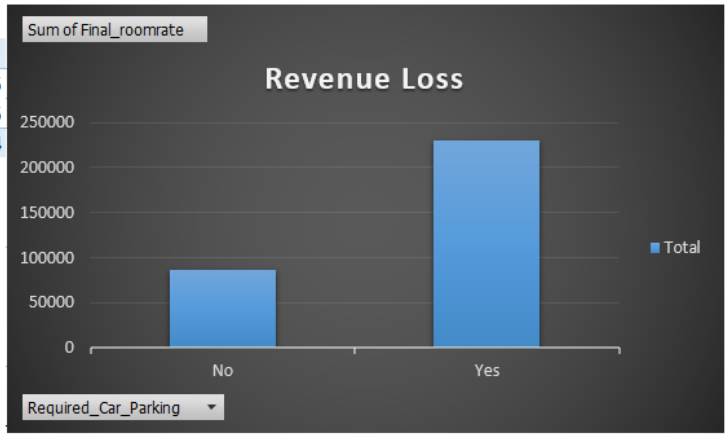

Figure 21: Revenue loss based on "Required Car Parking"

Insight:

- More revenue loss has been caused by the individuals **requesting car parks** than those not requiring car parking.

4.2.14 Reservation Status

The following diagram represents the revenue loss based on "Reservation Status":

Row Labels	Sum of Final_roomrate
Canceled	203988.85
No-Show	112175.15
Grand Total	**316164**

Figure 22: Revenue loss based on "Reservation Status"

Insight:

- More revenue loss has been caused by the people who have **canceled** the bookings than the No-show.

Solution:

- Enhance cancellation policies.
- Implement communication strategies to minimize cancellations, potentially introducing a nominal fee for cancellations made within a short time frame before the reservation date.
- Introduce a reminder system to remind the individuals about the bookings. Therefore, if they don't appear, they are more likely to remember and cancel the bookings.

4.2.15 Use Promotion

The following diagram represents the revenue loss based on "Use Promotion":

Row Labels	Sum of Final_roomrate
No	94338
Yes	221826
Grand Total	**316164**

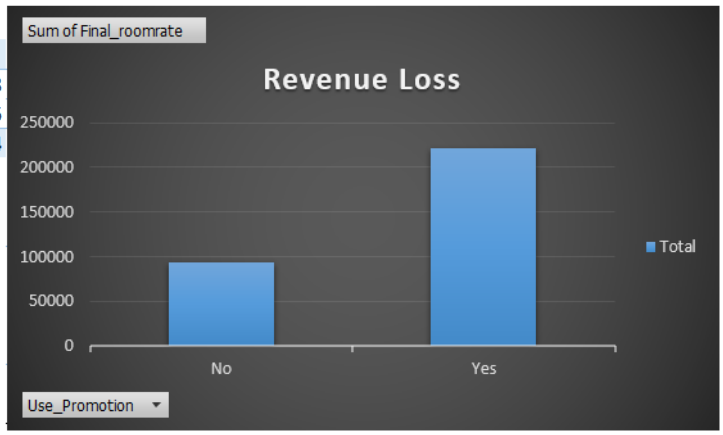

Figure 23: Revenue loss based on "Use Promotion"

Insight:

- More revenue loss has been caused by the reservations that appeared from **promotions**.

Solution:

- Refine promotional strategies by targeting a more qualified audience.
- Providing clearer terms and conditions.
- Analyze promotion performance to optimize for profitability.
- Strict the policies for bookings made through promotions by introducing compulsory deposits, penalty fees if canceled, and more!

5 Dashboard Analysis

5.1 Analysis 1

Here's a quick dashboard analysis made using Power BI to analyze the Revenue loss caused:

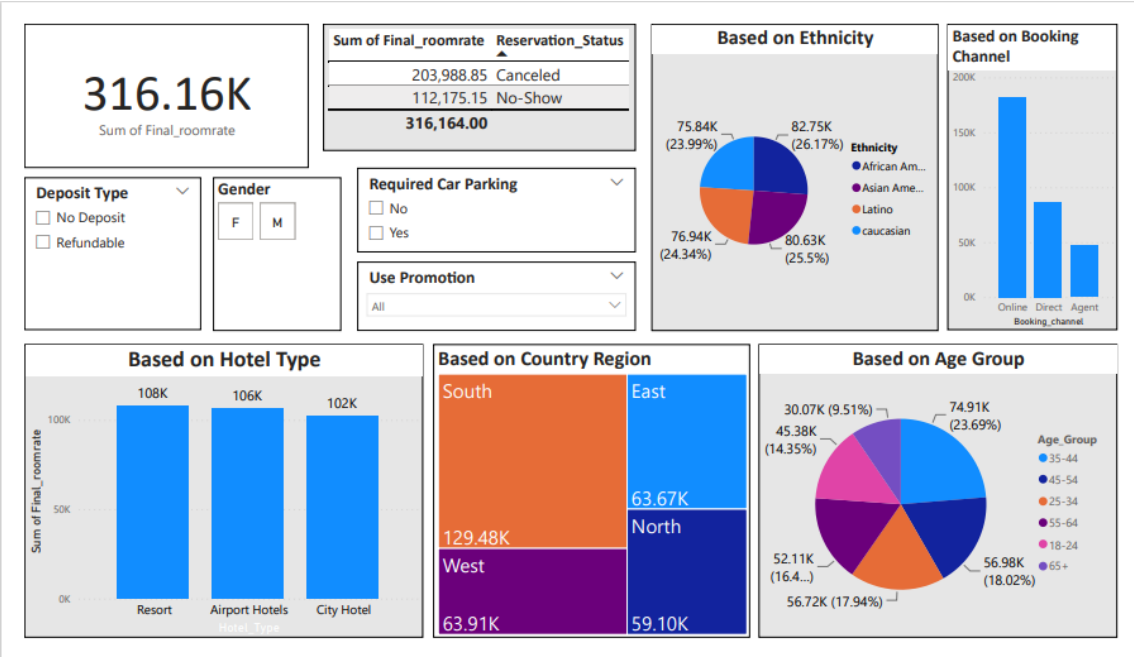

Figure 24: Analysis 1

The above diagram summarizes the information analyzed in the previous sections. However, the key metrics are only considered here, namely Deposit type, Gender, Reservation status, Ethnicity, Hotel type, Country region, and Age Group, by having the primary metric as the Final room rate.

Summary of suggested solutions based on the above insights:

- Refine promotional strategies by targeting a more qualified audience
- Optimize the policies related to "Cancellations" and "No-show."
- Analyzing promotion performance to optimize for profitability.
- The South region should be given more attention to overcome the overall revenue loss.
- More people have issues with Resorts and Airport hotels. Therefore, some reasonable packages, effective promotions, and target market activities can be implemented.
- Implement strategies to build trust in people.
- Since some visiting people have canceled the bookings, we can ask for feedback and the reason for cancellations.
- Hotels can implement strategies like previously, and people who have booked can re-book with no deposits.
- Hotels can provide more incentives or discounts for individuals who secure their reservation under a deposit and engage in Non-refundable deposits.
- Introduce a penalty fee for refunds and cancellations.
- Provide more incentives and discounts for the bookings made through direct and agents. Also, diversify these two options to get more sales.
- Implement communication strategies to minimize cancellations.
- Introducing a nominal fee for cancellations made within a short time before the expected check-in date.

- Introduce a reminder system to remind the individuals about the bookings.
- Strict the policies for bookings made through promotions by introducing compulsory deposits or penalty fees if canceled.

5.2 Analysis 2

Here's a quick dashboard analysis made to analyze the Revenue loss caused under different scenarios:

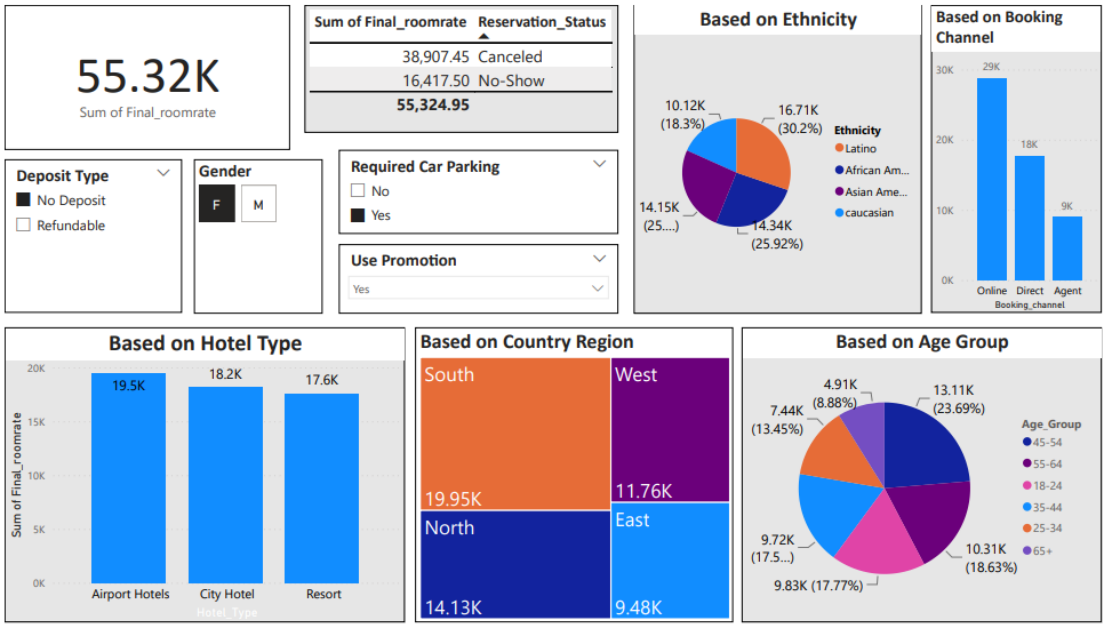

Figure 25: Analysis 2 based on Females

Insights from the above analysis:

- Revenue loss caused by females is $38,907.45$ who have booked (most people booked Online) and Cancelled by booking without depositing. Also, $16,417.50$ has not shown up as well.
- This had mainly happened between the age group of 45 to 54 booked under a Promotion.
- Also, they have mainly booked Airport hotels with the facility Car Park (People of Latinos in the South).
- Even though the promotion implemented during this season has been successful in convincing the age range of "45 - 55", there has been a considerable level of cancellations and no-shows.
- Conclusion: If the bookings had been done with a Non-refundable agreement or penalty fee, it is more likely that people would have visited the hotel.

Suggested **solutions** based on the above analysis:

- Implement a **non-refundable agreement** for bookings within the 45 to 54 age group under promotions, emphasizing the success of such agreements in increasing actual visits.
- Enhance **communication** about the benefits of non-refundable bookings and consider providing additional incentives for choosing this option.
- Analyze and adjust the promotion **strategy** for the Latino demographic in the South, considering specific preferences and concerns related to Airport hotels with car park facilities.
- Explore options to increase **direct** bookings or bookings through **agents**, emphasizing a more reliable source than online channels.

Here's a quick dashboard analysis made to analyze the Revenue loss caused under different scenarios:

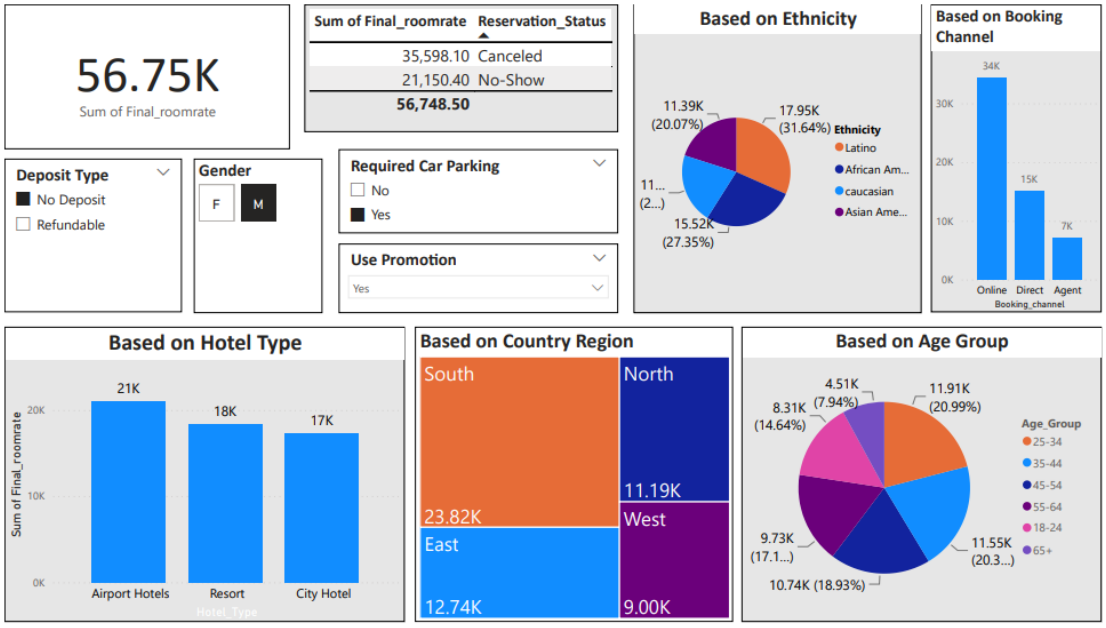

Figure 26: Analysis 2 based on Males

Considering the above **analysis** for the same scenario discussed above for "Males," the below insights are taken:

- The **Males** who booked with "No deposit," with promotion, requiring a car park, and have canceled have caused a revenue loss of $35,598.10$. Those who didn't show up caused a revenue loss of "$21,150.40$".
- But, they fall into the age group of "**25 to 34**, " a young age group. Even though the promotion has been successful for this age group of Latinos living in the South region, they still haven't shown up.
- Most of these **young** individuals have booked their reservations online.

Suggested **solutions**:

- Introduce a **nominal deposit** requirement for males aged 25 to 34 booking with the specified criteria and targeted communication on the benefits of honoring reservations to reduce cancellations and no-shows.
- Refine the **promotion** strategy for this **demographic**, considering additional **incentives** or engagement initiatives to increase **attendance** and mitigate revenue loss.

5.3 Analysis 3

Here's a quick dashboard **analysis** made to analyze the Revenue loss caused under different scenarios:

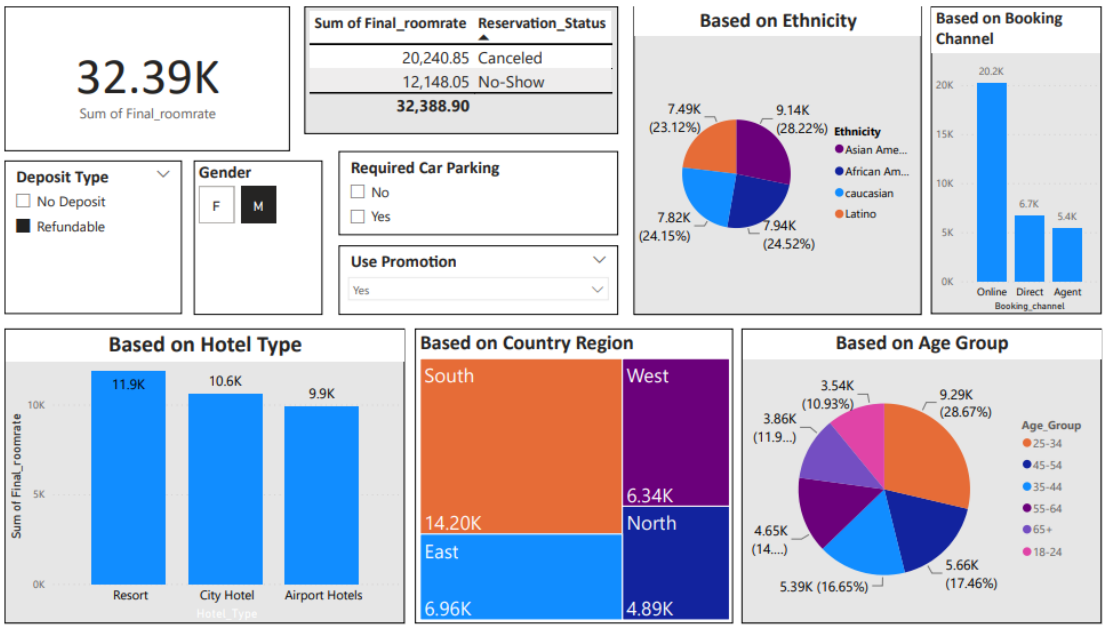

Figure 27: Analysis 3 based on "Males" and "Refundable"

Insights from the above analysis:

- Revenue loss caused by Males is $20,240.85$, who have canceled and requested a Refund. Also, $12,148.05$ has not shown up.
- This mainly happens between the age group of **25 to 34** booked under a Promotion.
- Also, they have mainly booked the **Resorts** (**Asian Americans** in the **South**).

Suggested **solutions**:

- Implement a stricter refund policy for the "**25 to 34**" age group booking under the promotion.
- Enhance communication regarding **cancellation policies**.
- Explore **targeted marketing** to understand and address specific preferences or concerns of the "**25 to 34**" age group.
- Evaluate the appeal of **Resorts** to the **Asian American demographic** in the South and consider diversifying offerings to cater to their preferences.

Here's a quick dashboard **analysis** made to analyze the Revenue loss caused under different scenarios:

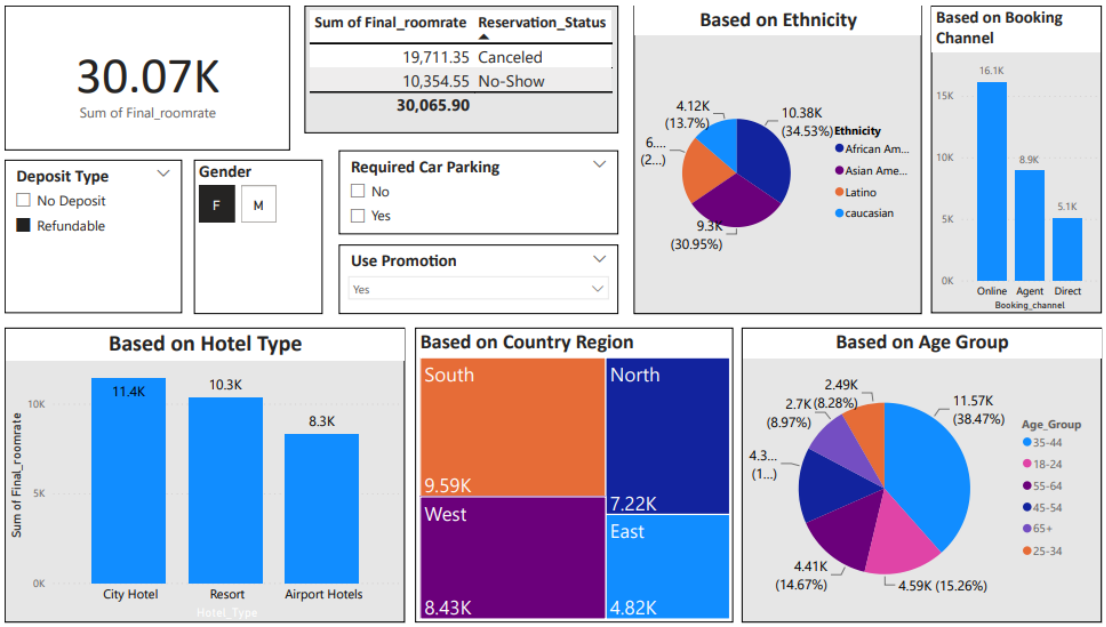

Figure 28: Analysis 3 based on "Females" and "Refundable"

Insights from the above analysis:

- Revenue loss caused by Females is $19,711.35$, who have canceled and requested a Refund. Also, $10,354.55$ has not shown up as well.
- This is mainly happened between the age group of 35-44 booked under a Promotion.
- Also, they have mainly booked the **City Hotels** (**African Americans in the South**).

Suggested **solutions**:

- Implement a more **stringent refund policy** for females aged **35-44** booking **under** promotions, emphasizing a commitment to reduce cancellations.
- Enhance **communication** about the importance of attendance and the benefits of the promotion to encourage this age group to honor their reservations.
- Evaluate the appeal of **City Hotels** to African Americans in the South, considering potential adjustments or additions to amenities and services that may better meet their preferences.

6 Additional Attributes That Hotel Management Should Collect

The following are the additional attributes the hotel management should collect:

- **Expected time of check-in and check-out:** Helps to narrow the decisions further.
- **Feedback/Satisfaction Score:** Include a column for customers to provide feedback or satisfaction scores after their stay. This information can be valuable for assessing overall customer experience.
- **Purpose of Visit:** Add a column specifying the purpose of the visit (e.g., business, leisure, family) to understand the different needs and behaviors of guests.
- **Member Status:** Include a column indicating the guest's membership or loyalty status if applicable. It can provide insights into the behavior of loyal customers.
- **Cancellation Reason:** Include a column for customers to specify the reason for cancellation, helping identify common issues and potential areas for improvement.
- **Promotion Details:** Add more to the "Use Promotion" column by including other information about the special offers visitors took part in. It lets for a more in-depth analysis for promotional effectiveness.
- **Room Preferences:** Get details about the room preference of the reservations such as bed type, if they want to smoke or not, bringing pets or special things they want.

7 Challenges Faced During The Data Analysis

The following are the main challenges faced in completing this report:

- Finding **outliers** of the data set.
- Finalizing the **additional** columns and data required to calculate the revenue loss.
- Using the tools **Excel**, **Power BI** for data analysis, and **Latex** for documenting played a major role. Since being less familiar, finding ways to solve the problems, the learning curve of these tools has been challenging.
- Designing and making the **charts** and **dashboards**.
- **Managing the time** to submit the report was challenging in the midst of the other assignments and study schedules.

8 Conclusion

In conclusion, revenue loss caused by cancellations and no-shows will always exist. Implementing proper and effective strategies based on the targetted strategies, marketing strategies based on the demographic regions, and understanding the target audience would help solve this problem. Therefore, the hotels can segment each region based on their requirements and analyze their feedback. Based on that, the hotel can implement effective strategies to suit each audience and market segment.