

TEAM: #READYTOWORK

TAN PHAN (Team leader)

- Email: vphan4@ur.rochester.com
- School/Major: University of Rochester - MS in Technical Entrepreneurship and Management (Data Science track)

WILLIAM LEE

- Email: slee287@ur.rochester.edu
- School/Major: University of Rochester - MS in Data Science

REVENUE MAXIMIZATION FOR GLOBAL HOTELS AND RESORTS (GHR)

EXECUTIVE SUMMARY

By leveraging classification and clustering models, we would like to provide following recommendations for revenue maximization at GHR:

1. CANCELATION:

- The canceled bookings contribute around **40%** of the total bookings. It is crucial that the hotel should predict if the booking is going to be canceled or not. If yes, the hotel can label the booking as potential cancellation, monitor and consider accepting additional bookings in case customers cancel their bookings.
 - Our model, **XGBoost** can predict whether a booking is going to be canceled with the **accuracy of 76.4%**
- Consider **increasing the threshold of last-minute cancel from 3 days to 5 or 7 days** to have more time to manage the bookings, as well as gain more revenue from penalty from cancellations.

2. FREE UPGRADE

- Free upgrade contributes **2.4%** of total normal bookings. However, this issue can be automatically resolved by effectively predicting the canceled bookings.

3. CUSTOMER SEGMENTATION

- Using K-means clustering, we can segment the customers into **4** different groups:
 - **Families:** Prioritize premium experiences with the highest daily rates, particularly during peak travel months from July to August, while frequently requesting amenities like parking and special accommodations
 - **Weekend chillers:** Contribute nearly half of revenue. They predominantly comprise couples or small groups of three, favor extended weekend stays from March to October.
 - **Business guests:** Prefer weekdays, stay briefly, and account for 83% of repeat guests, traveling year-round either alone or in pairs.
 - **Tour guests:** Have the longest lead time, a preference for half-board meals, and a tendency to travel in the summer.
- The hotel can leverage those insights of each segment to increase the number of bookings and spending per booking.

4. OCCUPATION RATE OPTIMIZATION

The occupancy rate for executive rooms and suites is still low: around **20-40%**. Improving the rate for this room type significantly helps increase the revenue.

DATA PREPARATION

1. BOOKING CANCELATION CLASSIFICATION:

- By leveraging **exploratory data analysis and feature importance technique**, we choose key variables that are the most effective to classify if a booking is going to be canceled or not: **Country, Lead Time, Deposit Type, Previous Cancellations, Total Of Special Requests, Average Daily Rate and Arrival Week Number**
- We used Arrival Date and Reservation Status Date to define the canceled vs. normal bookings.

2. FREE UPGRADE

- We consider normal bookings only (with check-out status). If **the reserved room is deluxe rooms and the assigned room is executive and suite room**, we define that booking as a free upgrade.

3. CUSTOMER SEGMENTATION

- Based on the data exploratory analysis, we choose the following variables for clustering model: **Number Of Adult, Is Family, Is Weekend, Total Nights', Market Segment, Customer Type, Revenue, Lead Time, Total Of Special Requests, Average Daily Rate**
- Only normal bookings are considered for the customer segmentation.

4. OCCUPATION RATE OPTIMIZATION

- Based on **the arrival date and check out date** (the reservation status date, with the booking status Check-out), we can calculate the actual occupation rate by room type for each day

DATA ANALYSIS & RECOMMENDATIONS

1. CANCELATION

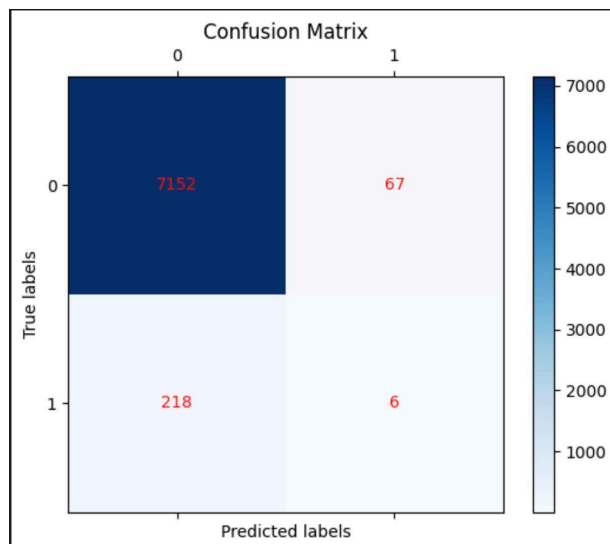
1.1. Cancellation Prediction Model

- Cancellation accounts for nearly **40% of total bookings**, which is a huge ratio. The high cancellation rate is driven by:
 - **Deposit type:** the hotel does not require customers to deposit to confirm their bookings, leading to the low commitment of the customers to their bookings.
 - **Bookings from Portugal** makes up the largest number of cancellations.
- **Therefore, it is valuable for the hotel if they can predict** whether a booking will be canceled or not. In case the booking is predicted to be canceled, they will consider accepting more booking requests.
- **Solution:** Build a classification model to predict whether a booking will be canceled or not.
- Applying 4 classification models to classify canceled bookings vs. normal bookings, the **XGBoost model outperforms in all performance metrics, with 76.4% of accuracy.**

Model	Accuracy	Kappa	AUC
Logistic Regression	0.7482	0.375	0.664
SVM	0.748	0.355	0.650
Random Forest	0.749	0.425	0.701
XGBoost	0.764	0.448	0.708

1.2. Early Canceled vs. Last-Minute Canceled Booking Prediction

- We try to take a deeper dive prediction that helps the hotel predict if the expected cancellation is early or last-minute. However, even though we handled the imbalance data (**only 2.69%** of canceled bookings are last-minute), the performance of classification is not reliable.



In the confusion matrix, the model can only correctly predict the early cancellation (**7152**), but **fails to predict the last-minute cancellation (218 last-minute canceled bookings are predicted as the early canceled bookings)**

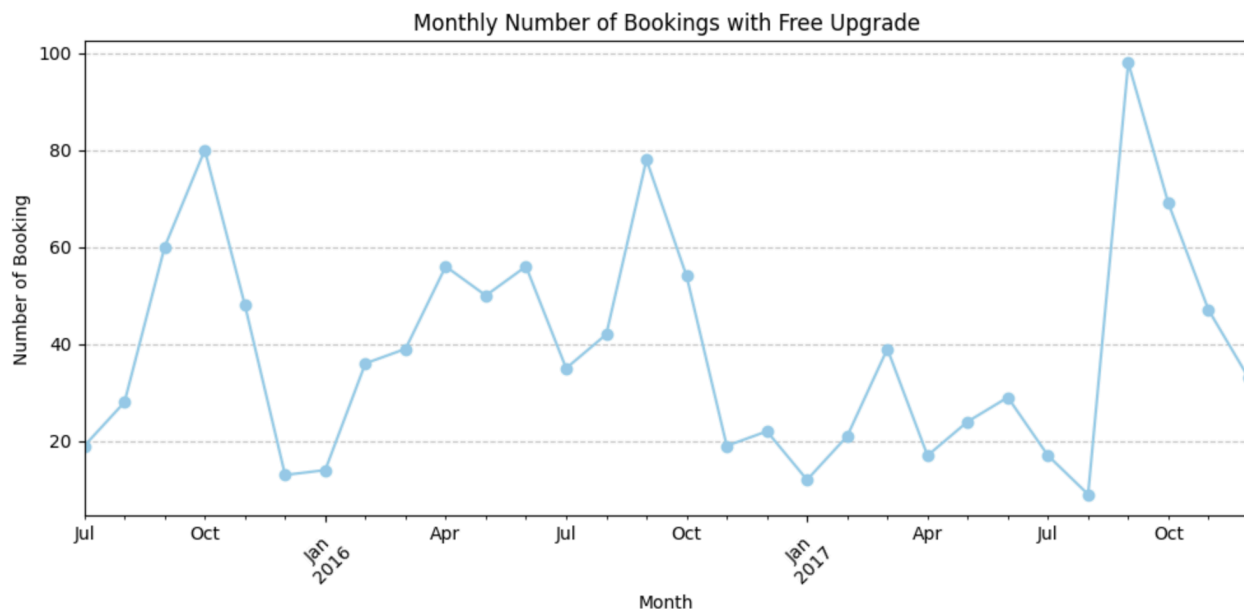
Therefore, we do not recommend building the model to predict the early canceled vs. last-minute canceled bookings. Instead, we recommend changing the threshold of 3 days to 5 or 7 days.

Threshold	Number of Late Cancellation	Revenue Contribution
3 days	934	0.57%
5 days	1481	0.90%
7 days	2051	1.24%

With the threshold of 3 days, more than **97%** of canceled bookings are early cancellations. Increasing the threshold to **5 or 7 days** does not only help the hotel increase from **0.3 to 0.6%** of total revenue, it also helps the hotel to have **more time** to find additional books for replacement.

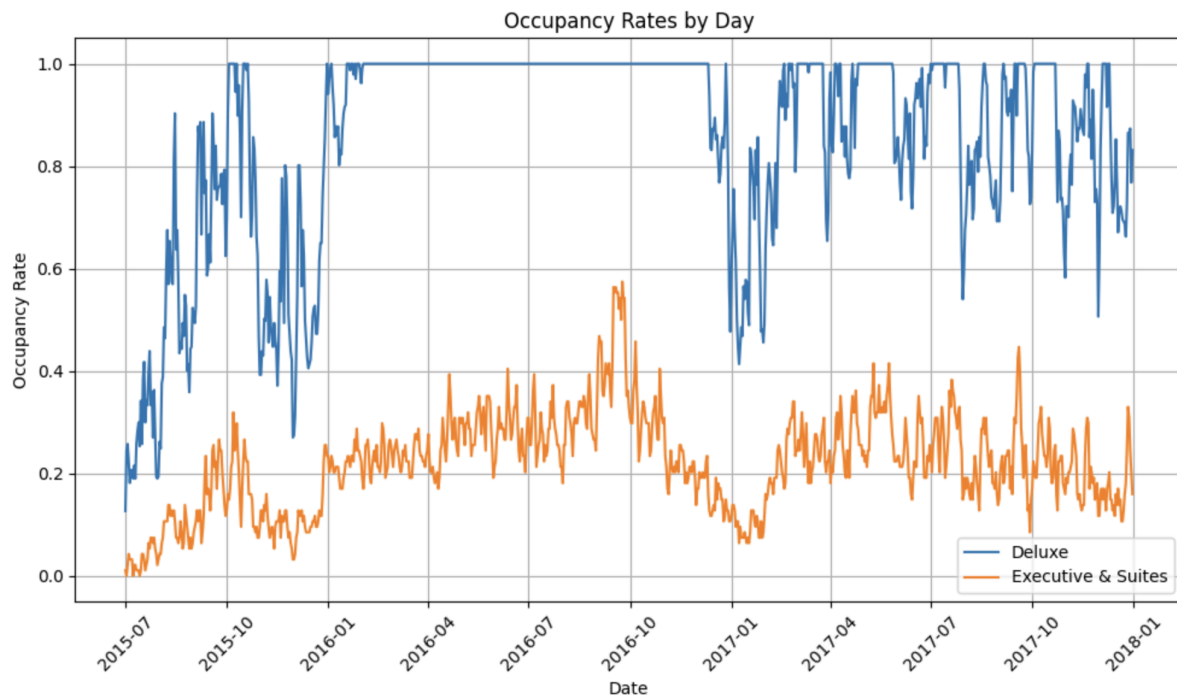
2. FREE-UPGRADE

It is interesting that out of all normal bookings (with valid check-outs), there are only **2.4%** of bookings that the hotel had to offer free room upgrades for customers.



We assume that **we just offer a free upgrade when all of the deluxe rooms are occupied**. In 2016, while the occupancy rate mostly reached 100%, it's reasonable that we need to offer them a free upgrade. However, looking at the occupancy rates by day and by room type in 2017, we see that the rate for deluxe rooms **dramatically fluctuates**, and we also have some free

upgrades during this year. If we can manage the cancellation prediction well, we do not have to provide the free upgrade while we still have deluxe rooms.



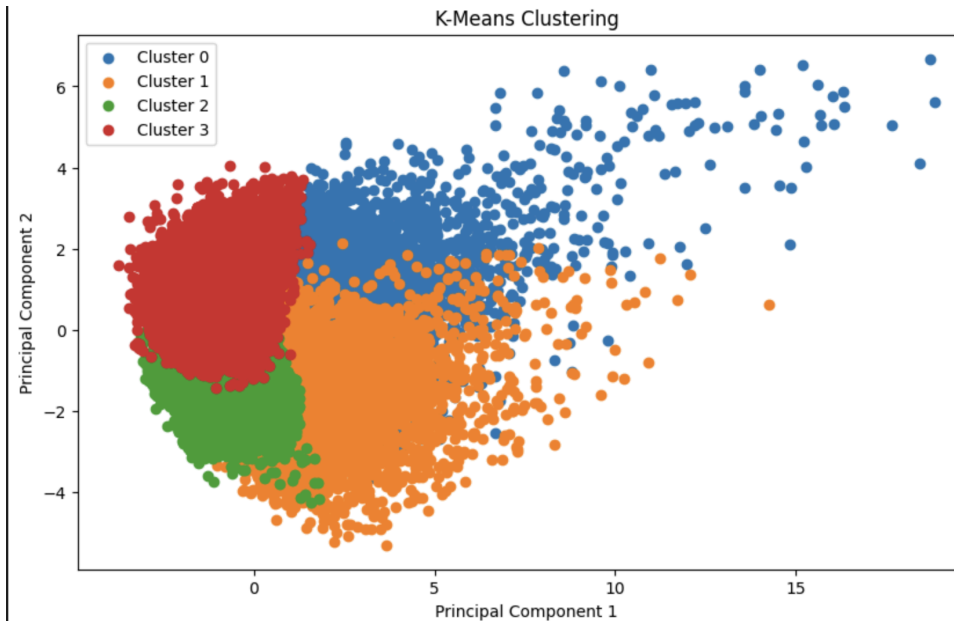
However, the number of bookings with free upgrades for each month is not high, except for December 2017. Therefore, if the hotel could solve the issues of cancellation prediction, the number of bookings with free upgrades would automatically reduce.

In short, from our team's perspective, **the free upgrade issue is not critical, and it can be handled by booking cancellation prediction.**

3. CUSTOMER SEGMENTATION

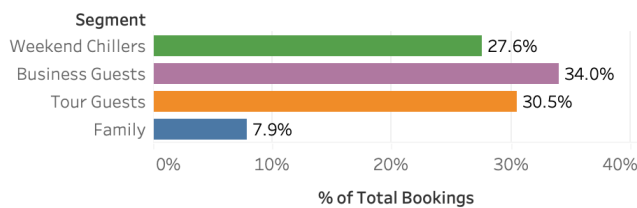
Based on the data exploratory analysis, we choose the following variables for clustering model: **Number Of Adult, Is Family, Is Weekend, Total Nights', Market Segment, Customer Type, Revenue, Lead Time, Total Of Special Requests, Average Daily Rate**

Note: Only normal bookings are considered for the customer segmentation.

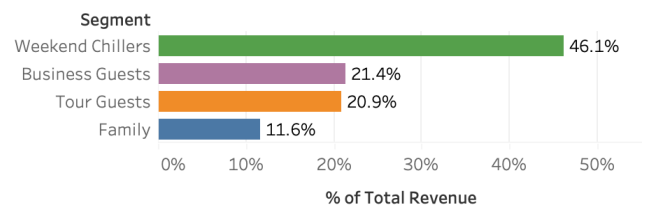


Visualization of 4 segments

Booking contribution (%)



Revenue contribution (%)



In general, we have 4 segments with the following key insights:

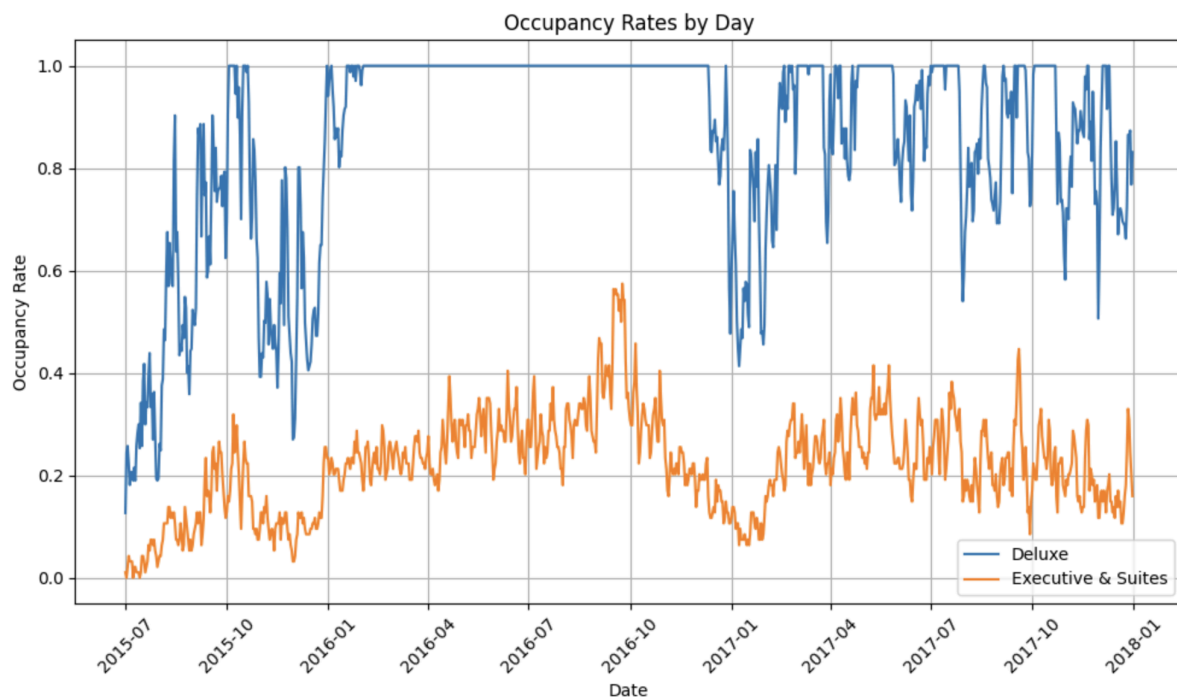
- **Families:** Prioritize **premium experiences** with the highest daily rates, particularly during peak travel months from July to August, while frequently requesting amenities like parking and special accommodations
- **Weekend chillers:** Contribute **nearly half of revenue**. They predominantly comprise couples or small groups of three, favor **extended weekend** stays from March to October.
- **Business guests:** Prefer **weekdays**, stay briefly, and account for 83% of repeat guests, traveling year-round either alone or in pairs.
- **Tour guests:** Have the **longest lead time**, a preference for half- board meals, and a tendency to travel in the summer.

The hotel can leverage behavioral insights of each segment to maximize revenue for hotel and resort.

- **Increase number of bookings:** Customize marketing strategies to target each segment more effectively to increase the number of bookings per segment.
- **Increase spending per booking:** Personalize services to match the needs of each segment to increase their spending per booking and enhance their experience. This

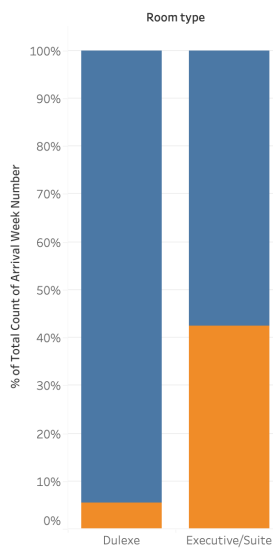
approach also encourages them to come back and recommend the services to their network.

4. OCCUPATION RATE OPTIMIZATION



Back to the occupancy rate by room type and by date, we see that the occupancy rate for executive rooms and suites is still low: around **20-40%**. Improving the rate for this room type significantly helps increase the revenue.

Family Bookings by Room Type



(Orange: Family)

We see that 40% of executive and suite rooms are booked by family. Therefore, we should suggest the option of upgrading to executive and suite rooms from families for more convenience and comfort.

Besides, we can try following strategies to increase the occupancy rate for this room type:

- **Spend more marketing budget to promote executive and suite rooms**
 - Increase the visibility of executive and suites rooms on your hotel's website, booking platforms, and other marketing channels.
 - Showcase professional photographs, virtual tours, and detailed descriptions of the rooms and their amenities to entice potential guests.
- **Promotions and Packages:**
 - Offer special promotions, packages, and discounts specifically for executive and suites rooms to attract guests and incentivize bookings.
 - Create attractive bundle deals that include additional perks or services, such as complimentary meals, spa treatments, or airport transfers, to enhance the value proposition for guests.
- **Personalized Upselling and Cross-Selling:**
 - Train your staff to actively promote executive and suites rooms during the booking process and at check-in, highlighting the benefits and advantages of upgrading.
 - Implement a targeted upselling strategy by offering guests the option to upgrade to an executive or suites room at a discounted rate when they book a standard room.

SUMMARY

By leveraging classification and clustering models, we would like to provide following recommendations for revenue maximization at GHR:

1. CANCELTATION:

- The canceled bookings contribute around **40%** of the total bookings. It is crucial that the hotel should predict if the booking is going to be canceled or not. If yes, the hotel can label the booking as potential cancellation, monitor and consider accepting additional bookings in case customers cancel their bookings.
 - Our model, **XGBoost** can predict whether a booking is going to be canceled with the **accuracy of 76.4%**
- Consider **increasing the threshold of last-minute cancel from 3 days to 5 or 7 days** to have more time to manage the bookings, as well as gain more revenue from penalty from cancellations.

2. FREE UPGRADE

- Free upgrade contributes **2.4%** of total normal bookings. However, this issue can be automatically resolved by effectiveling predicting the canceled bookings.

3. CUSTOMER SEGMENTATION

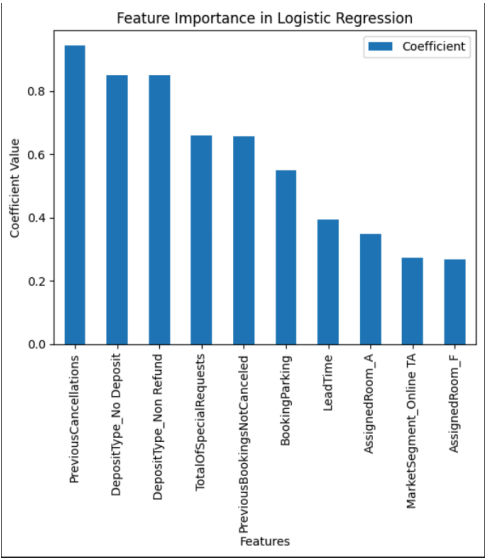
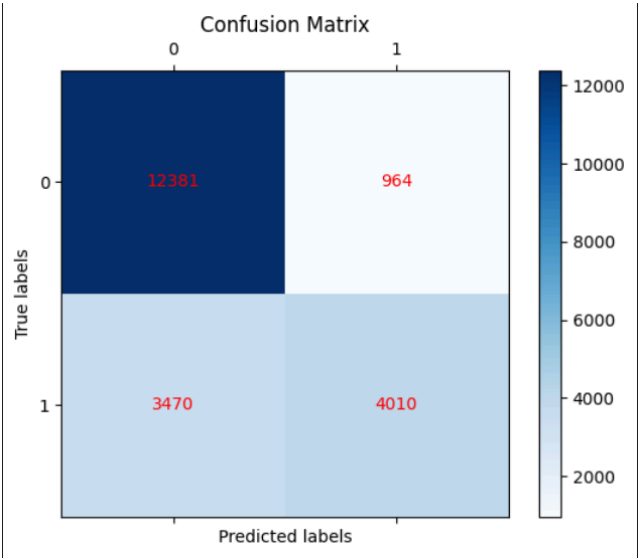
- Using K-means clustering, we can segment the customers into **4** different groups:
 - **Families:** Prioritize premium experiences with the highest daily rates, particularly during peak travel months from July to August, while frequently requesting amenities like parking and special accommodations
 - **Weekend chillers:** Contribute nearly half of revenue. They predominantly comprise couples or small groups of three, favor extended weekend stays from March to October.
 - **Business guests:** Prefer weekdays, stay briefly, and account for 83% of repeat guests, traveling year-round either alone or in pairs.
 - **Tour guests:** Have the longest lead time, a preference for half- board meals, and a tendency to travel in the summer.
- The hotel can leverage those insights of each segment to increase the number of bookings and spending per booking.

4. OCCUPATION RATE OPTIMIZATION

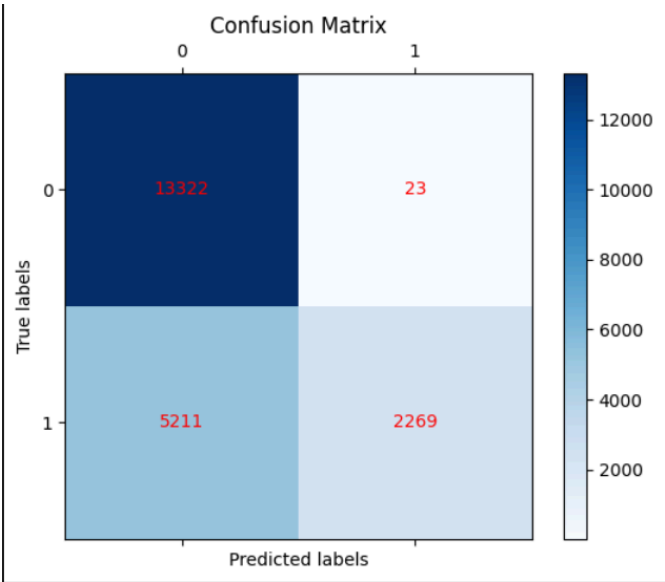
The occupancy rate for executive rooms and suites is still low: around **20-40%**. Improving the rate for this room type significantly helps increase the revenue.

APPENDIX

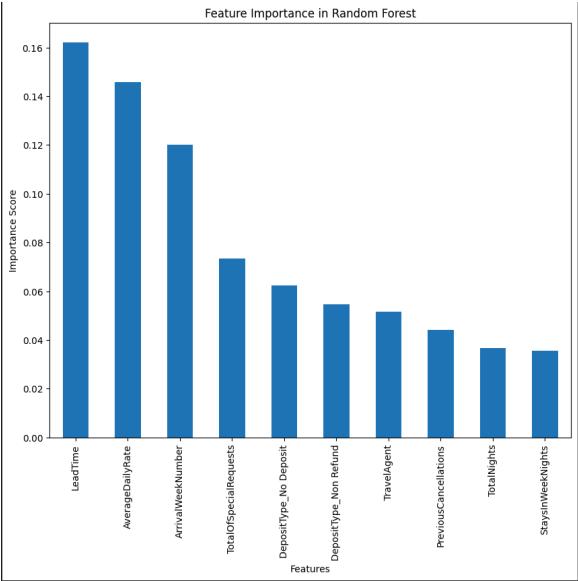
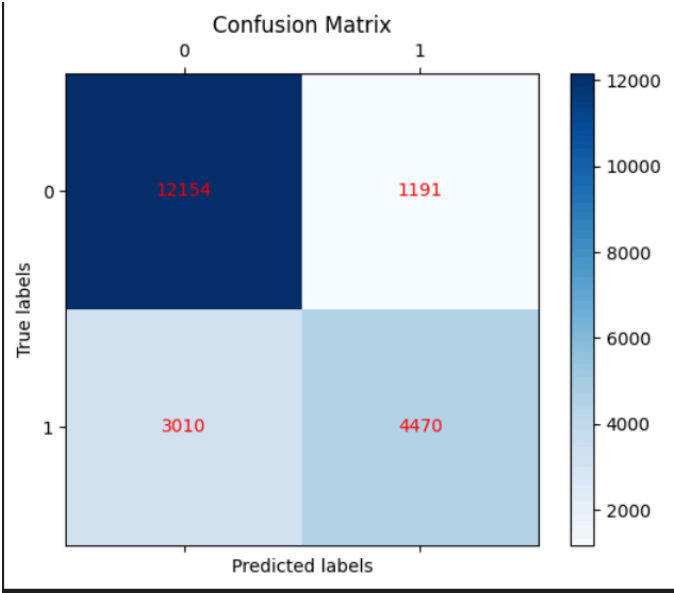
Non-Cancellation and Cancellation:
Logistic Regression



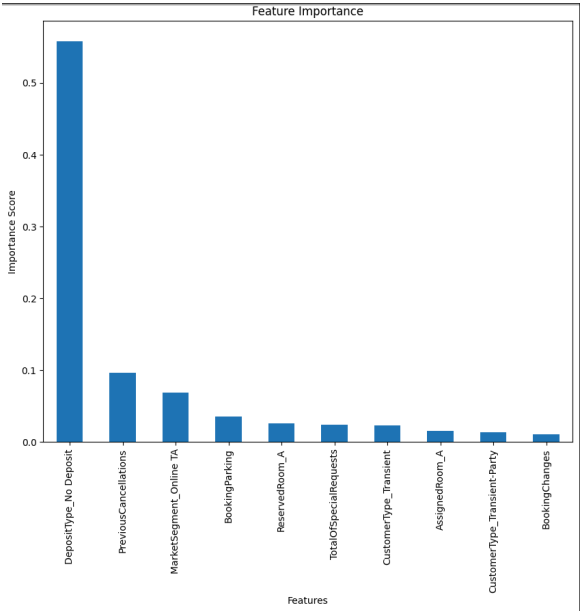
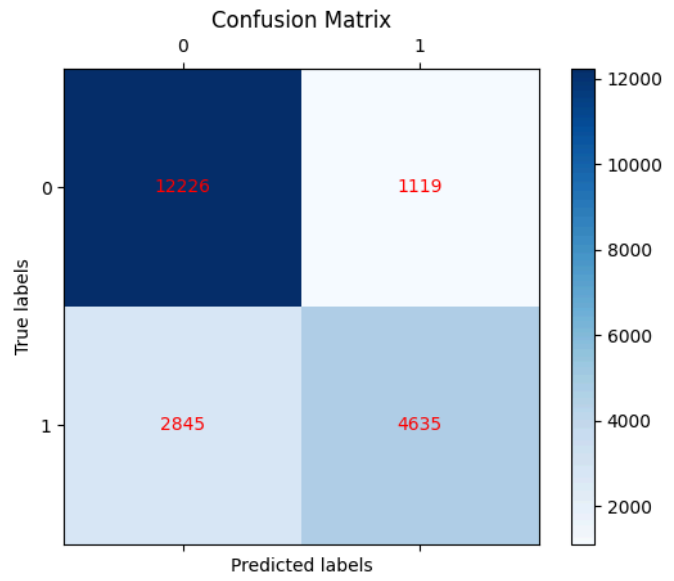
SVM



Random Forest

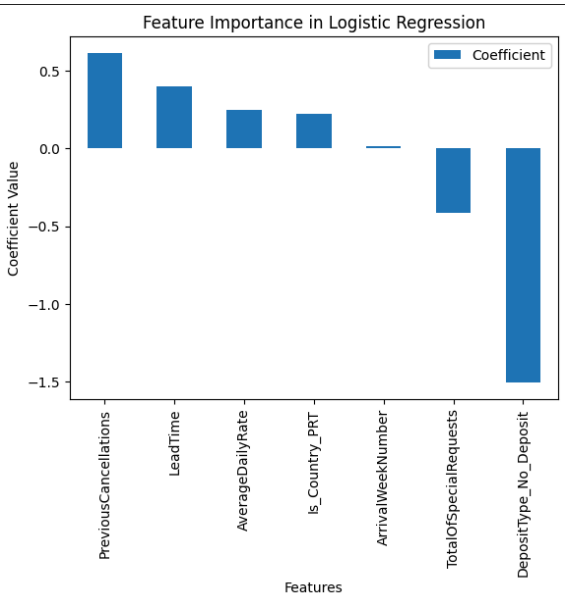
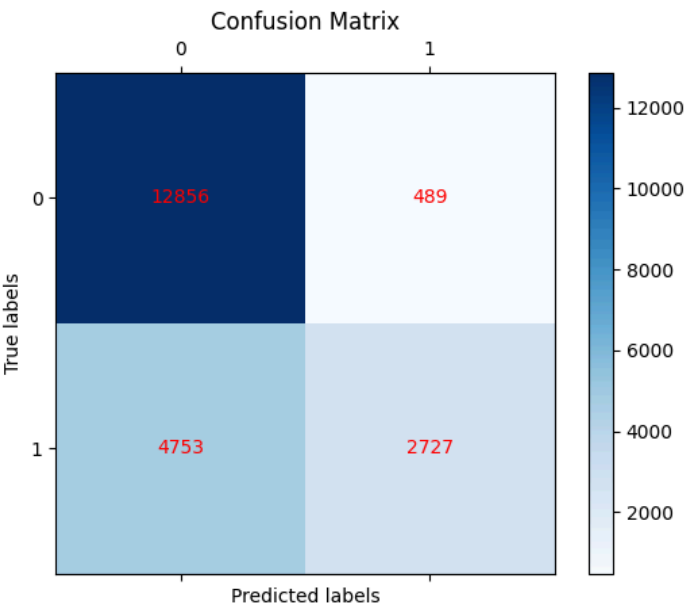


XGboost

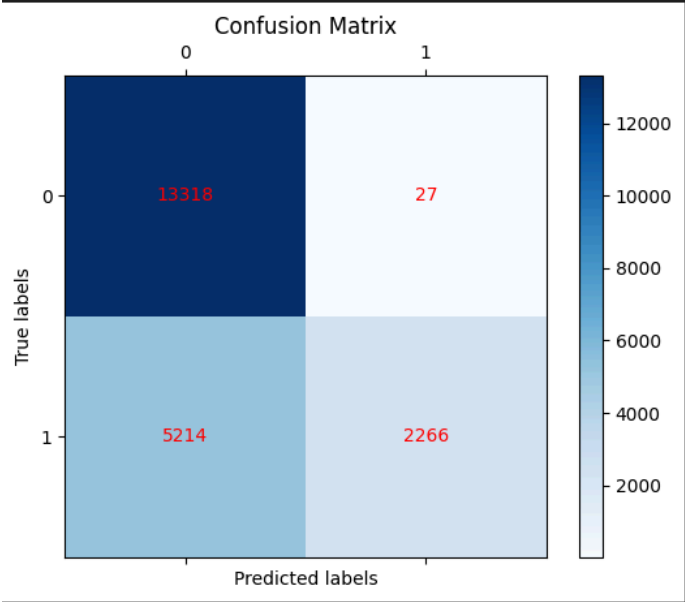


Feature Selection

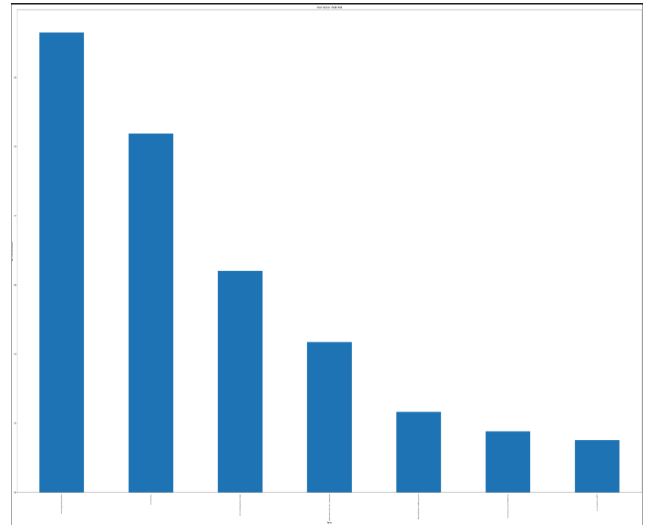
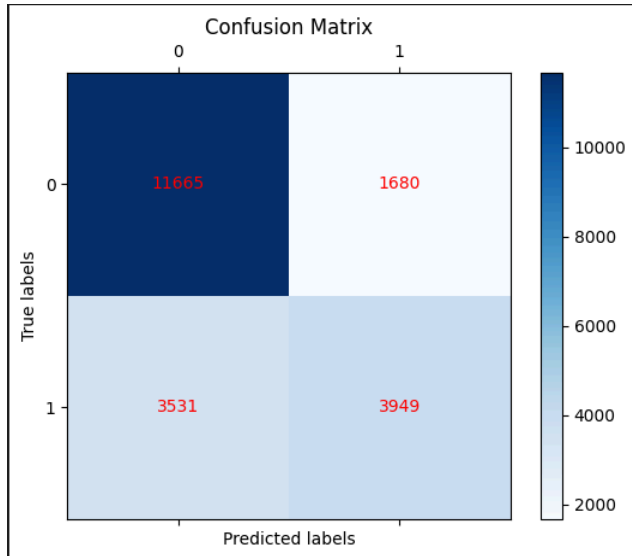
Logistic Regression



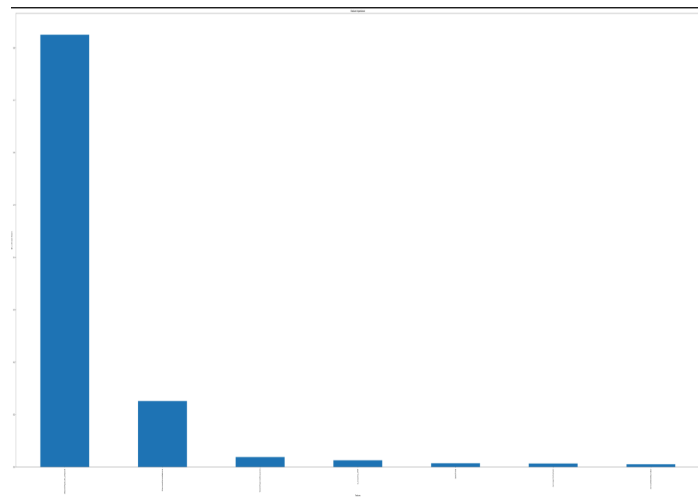
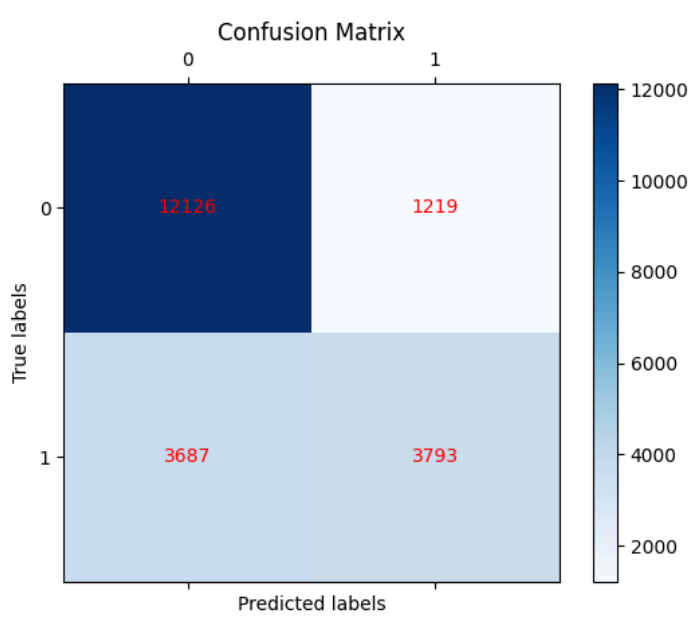
SVM



Random Forest

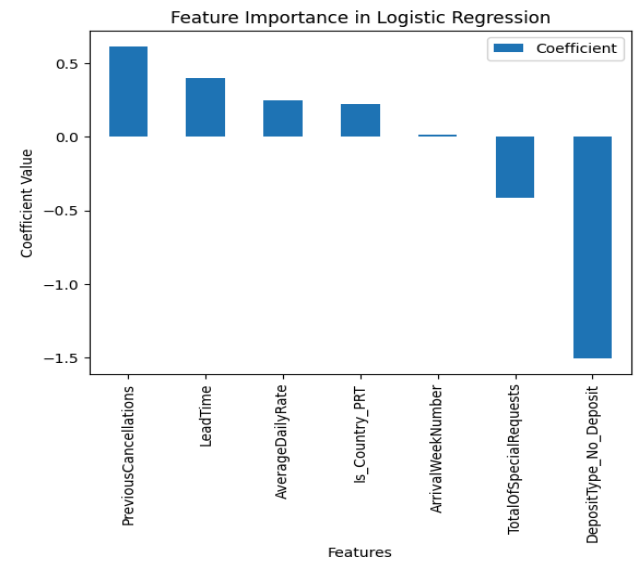
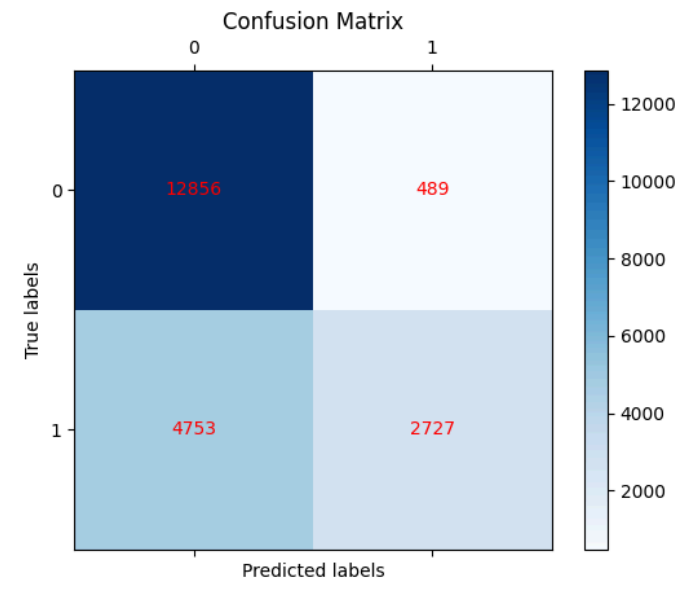


XGBoost

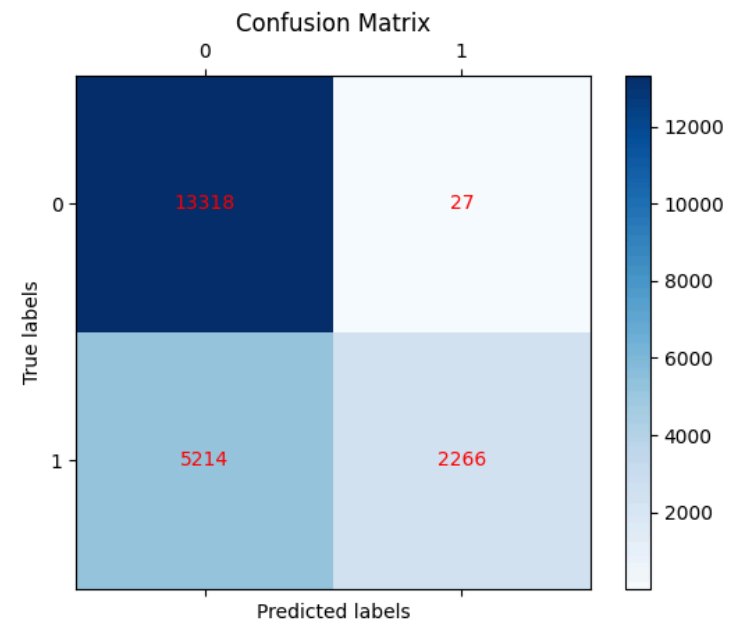


Feature Selection-Early cancellation and late cancellation:

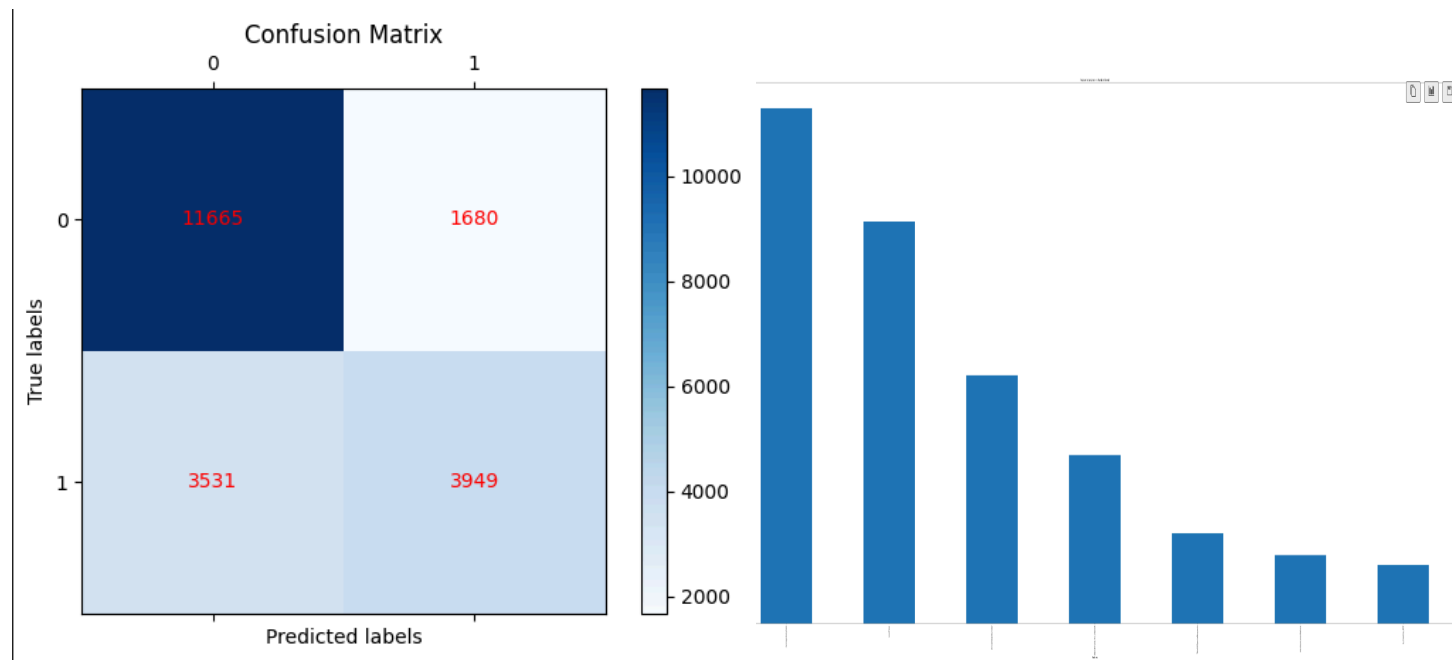
Logistic Regression



SVM



Random Forest



XGBoost

