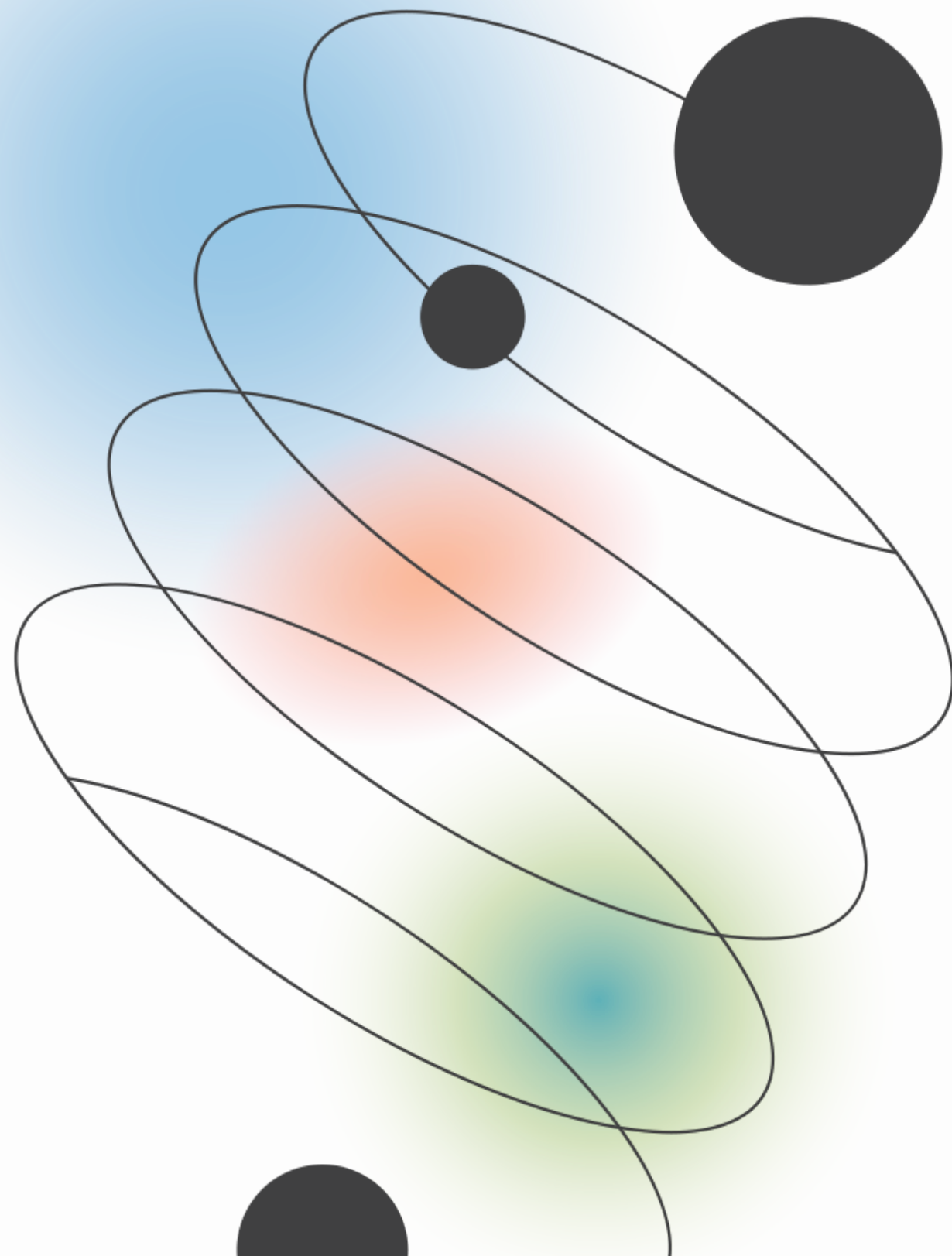
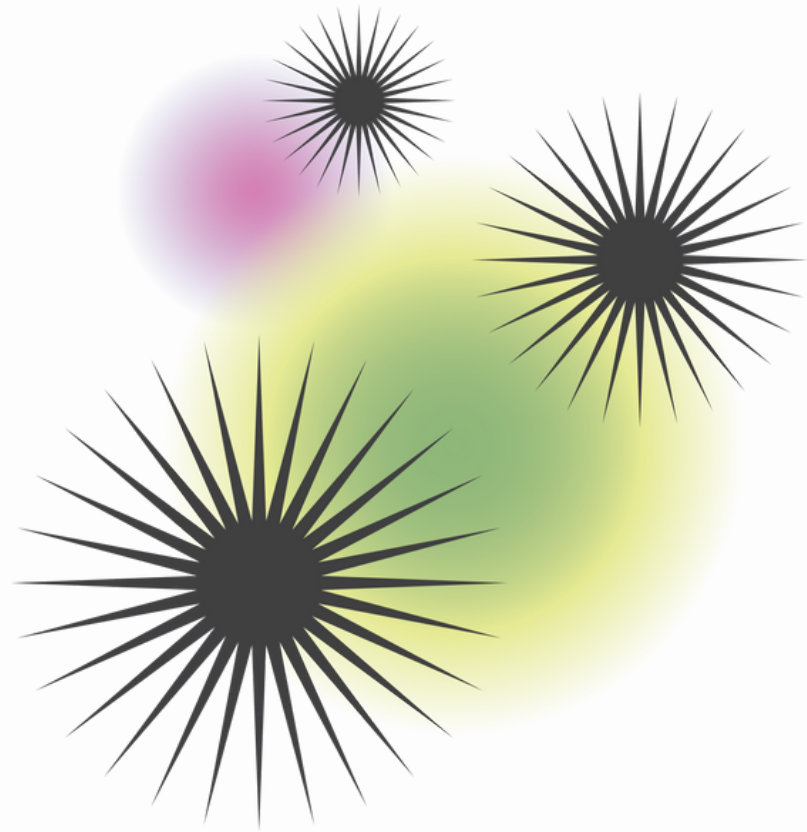


Hotel Reservations

Phạm Văn Chiến



Mục lục

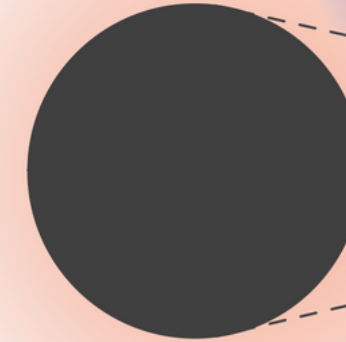


01

Giới thiệu về dataset

03

Mô hình



02

Xử lý và vẽ đồ thị

04

Giá trị thực tiễn

1.Dataset

```
RangeIndex: 36275 entries, 0 to 36274
```

```
Data columns (total 19 columns):
```

#	Column	Non-Null Count	Dtype
0	Booking_ID	36275 non-null	object
1	no_of_adults	36275 non-null	int64
2	no_of_children	36275 non-null	int64
3	no_of_weekend_nights	36275 non-null	int64
4	no_of_week_nights	36275 non-null	int64
5	type_of_meal_plan	36275 non-null	object
6	required_car_parking_space	36275 non-null	int64
7	room_type_reserved	36275 non-null	object
8	lead_time	36275 non-null	int64
9	arrival_year	36275 non-null	int64
10	arrival_month	36275 non-null	int64
11	arrival_date	36275 non-null	int64
12	market_segment_type	36275 non-null	object
13	repeated_guest	36275 non-null	int64
14	no_of_previous_cancellations	36275 non-null	int64
15	no_of_previous_bookings_not_canceled	36275 non-null	int64
16	avg_price_per_room	36275 non-null	float64
17	no_of_special_requests	36275 non-null	int64
18	booking_status	36275 non-null	object

```
dtypes: float64(1), int64(13), object(5)
```

```
memory usage: 5.3+ MB
```

Bộ data bao gồm 36275 dòng và 19 cột

Thể hiện thông tin của các đơn đặt phòng như: số người lớn, trẻ nhỏ, loại phòng, giá phòng, yêu cầu đặc biệt, tình trạng đơn đặt phòng...

Bài toán: từ bộ data có sẵn, xây dựng các mô hình dự đoán và tìm ra mô hình có khả năng dự đoán chính xác cao nhất

2. Biểu đồ

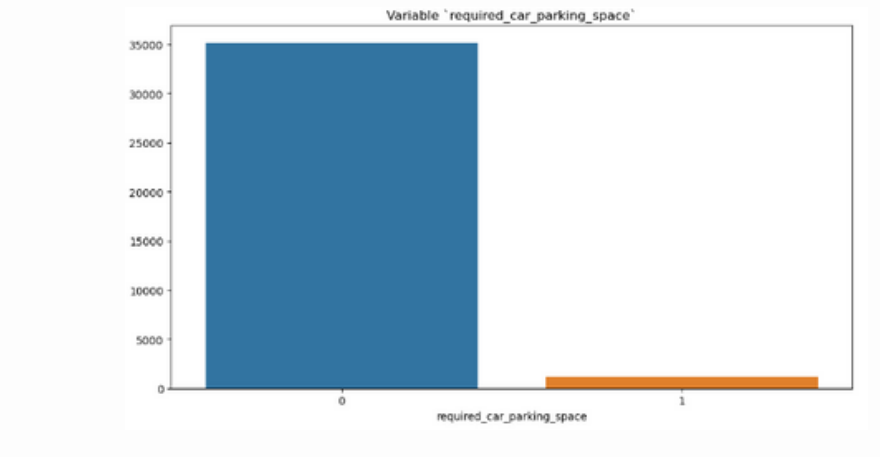
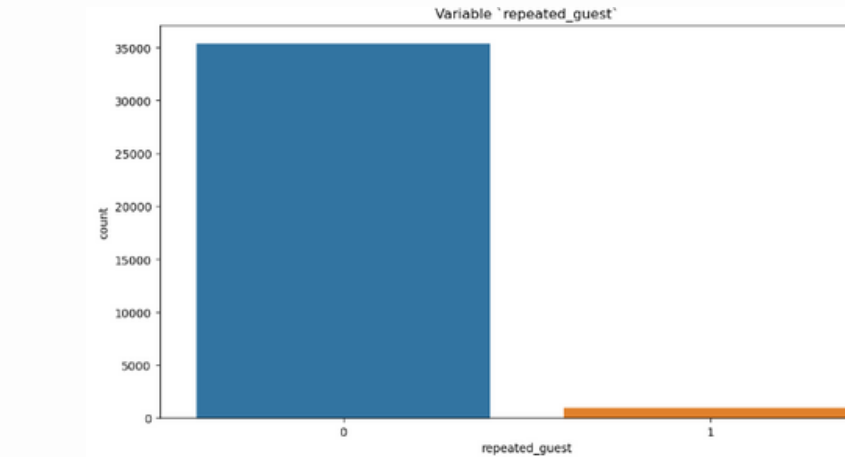
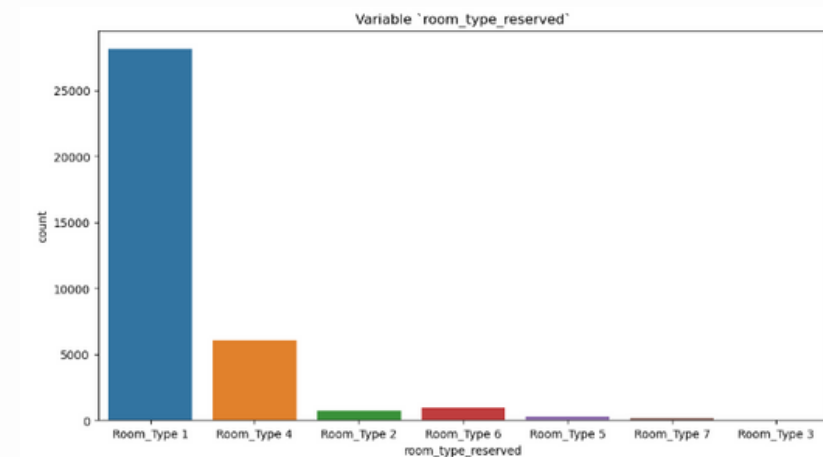
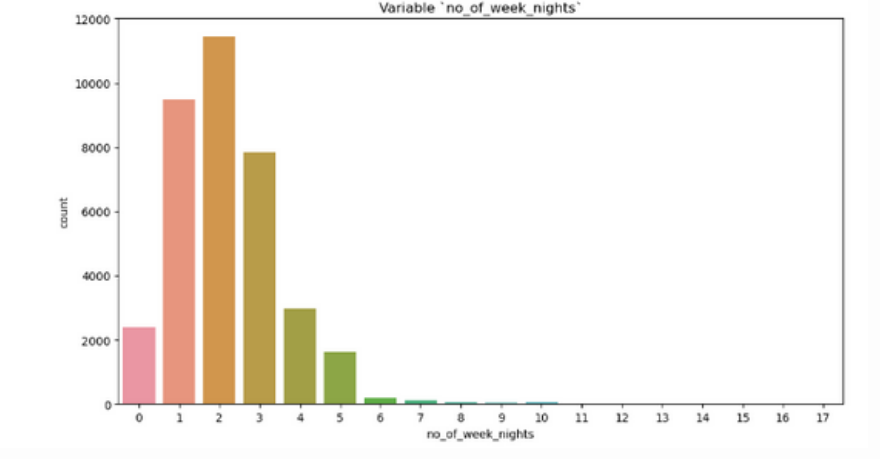
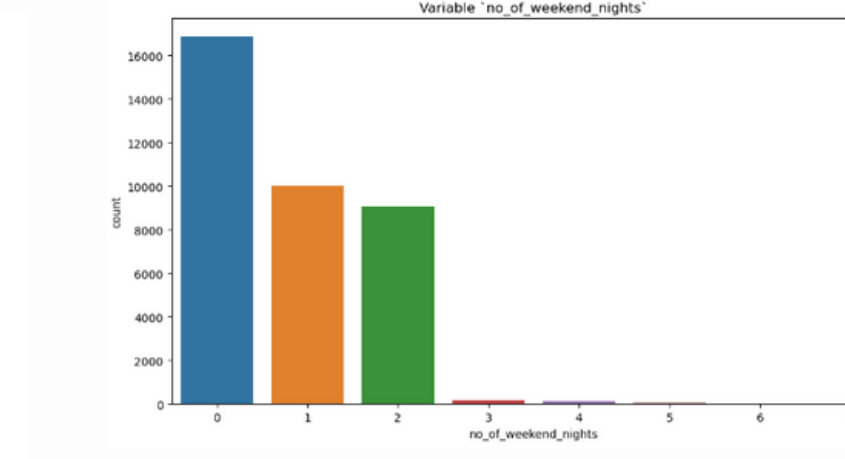
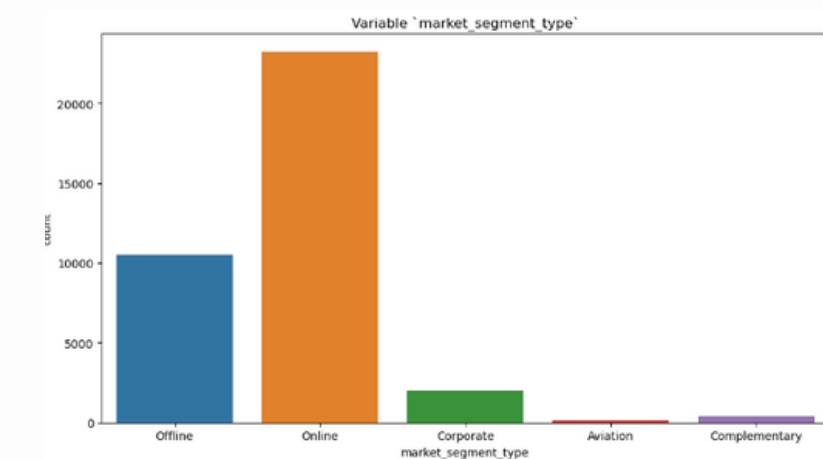
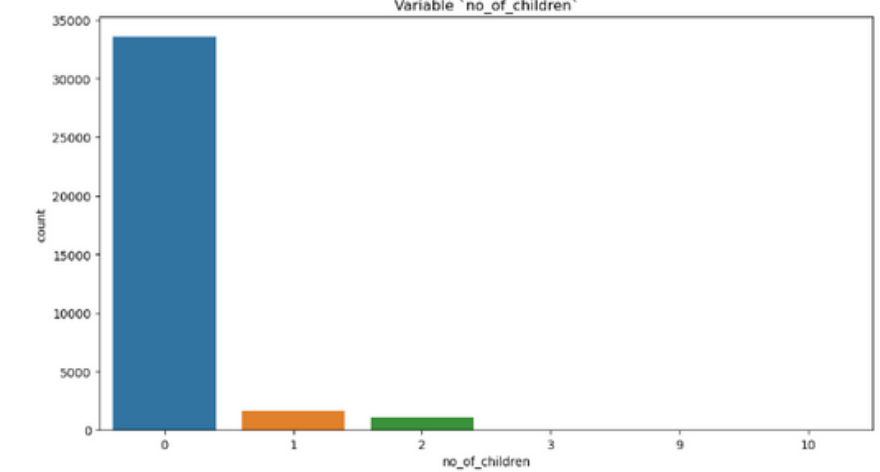
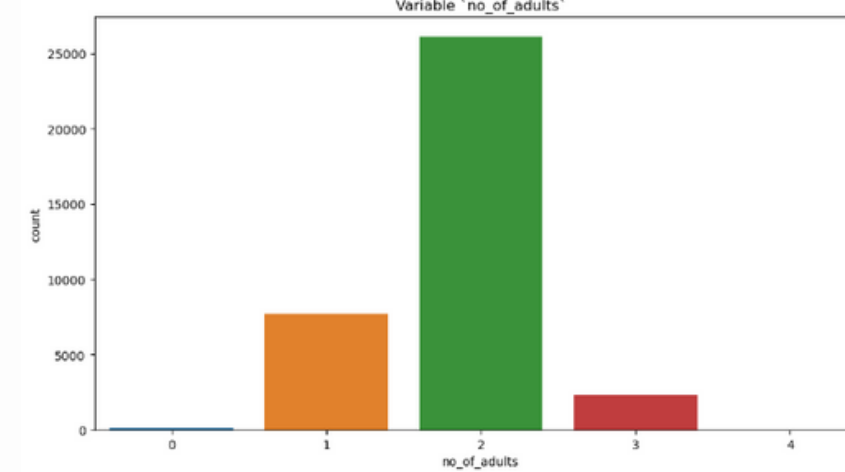
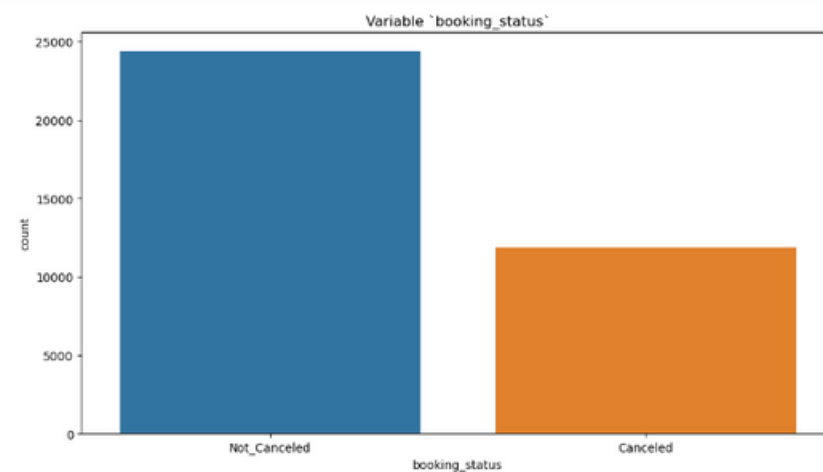
-Tình trạng huỷ phòng chiếm tới trên 30%

-Khách hàng chủ yếu đi 2 người

-Khách hàng của khách sạn chủ yếu ở trong tuần (2 ngày)

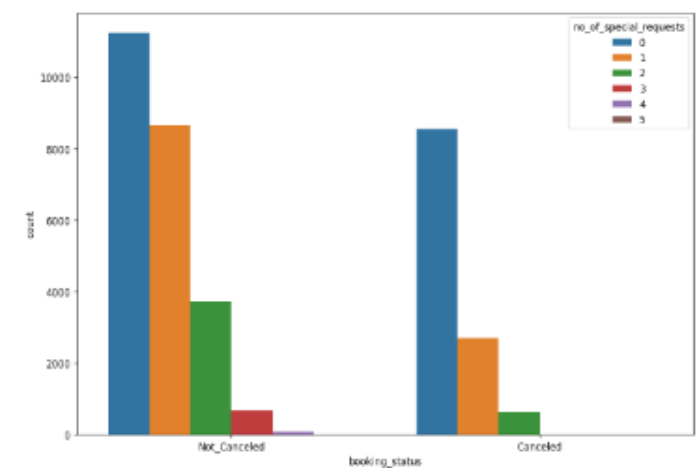
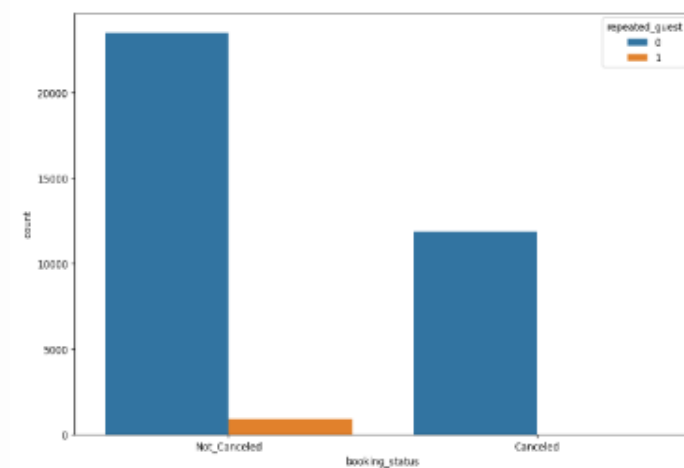
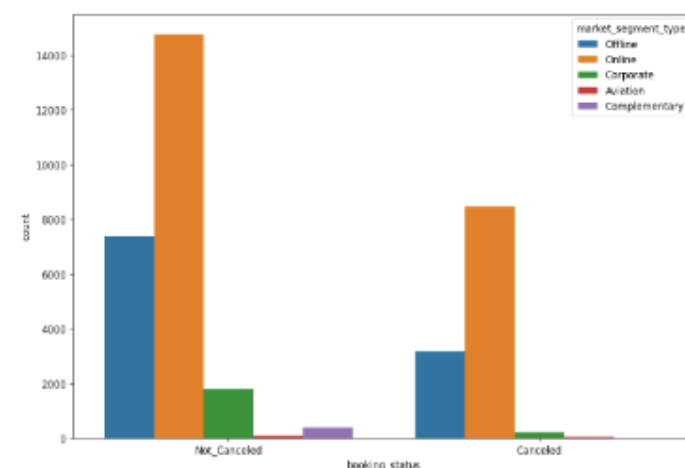
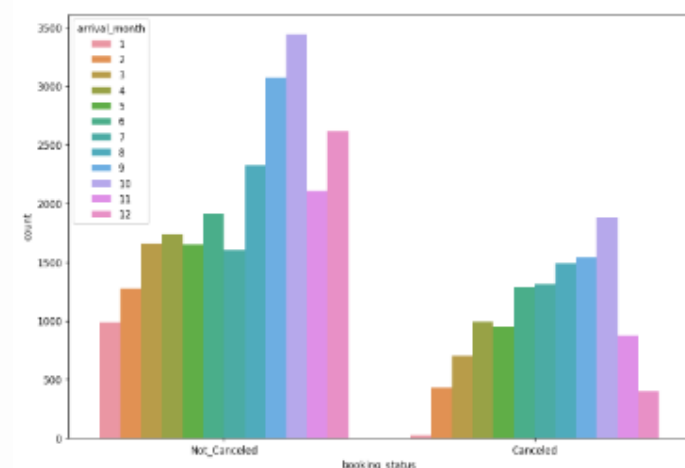
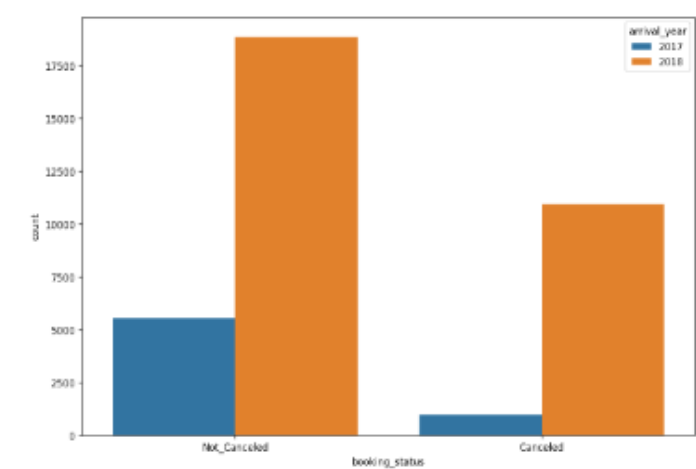
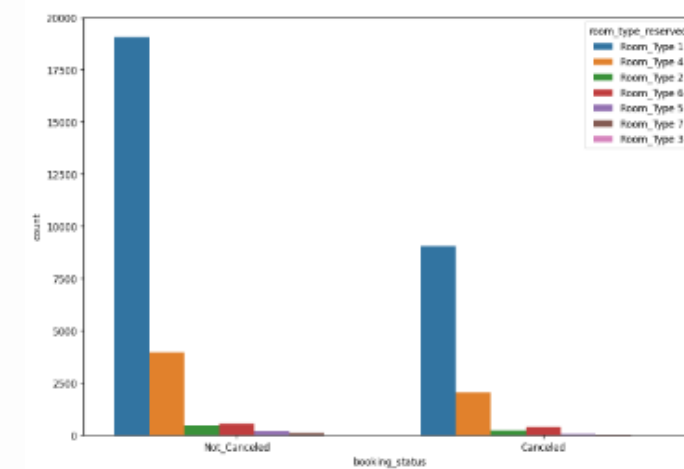
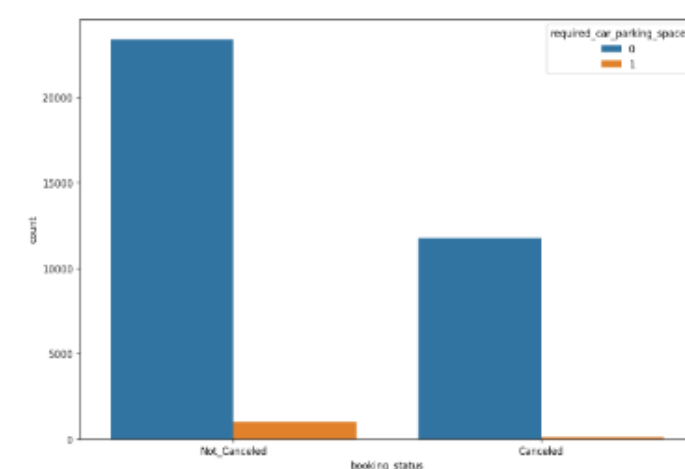
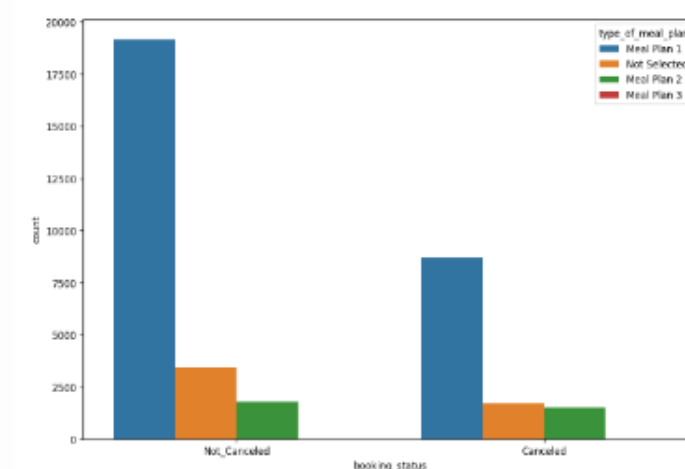
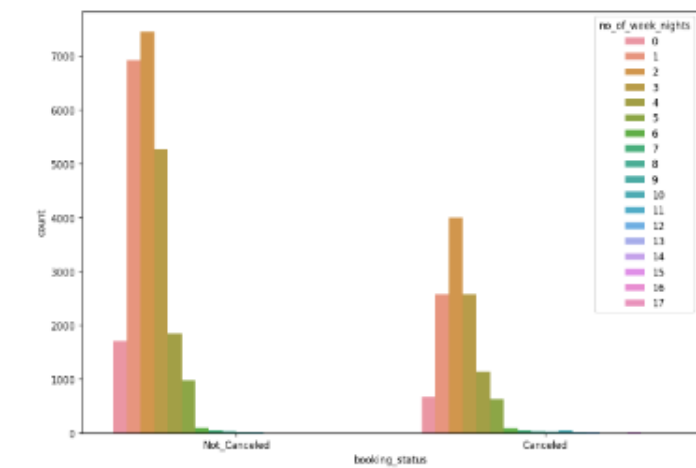
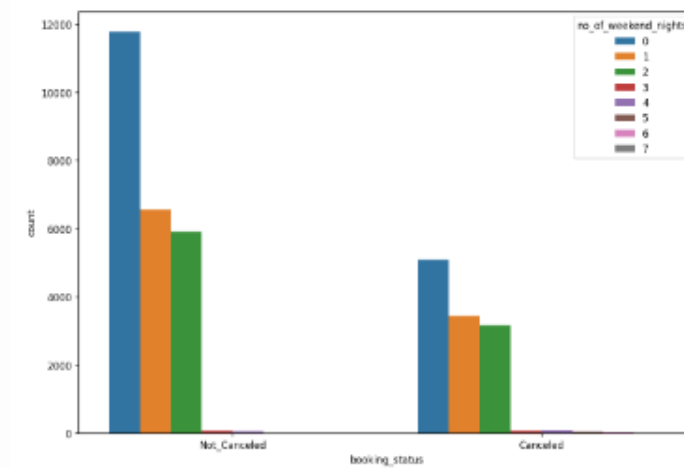
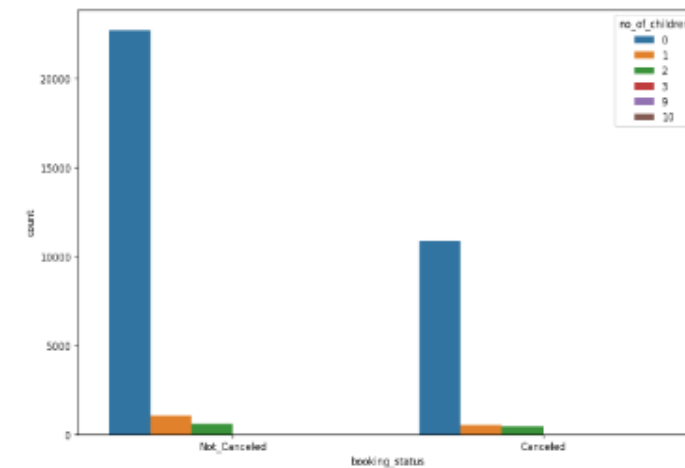
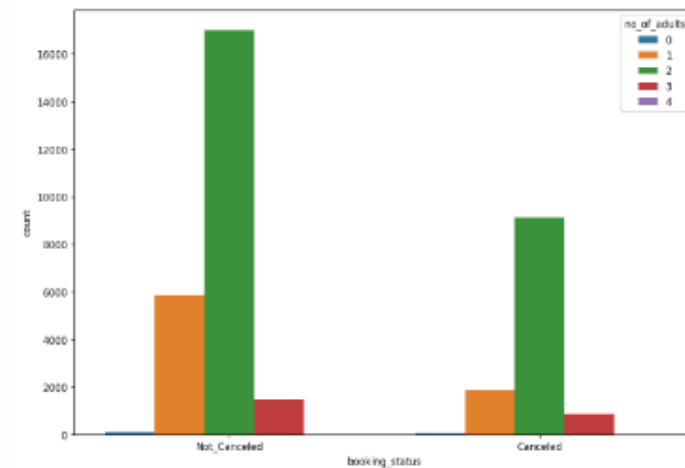
-Phòng loại 1 được ưa chuộng nhất

-Tỷ lệ khách hàng quay lại cực thấp



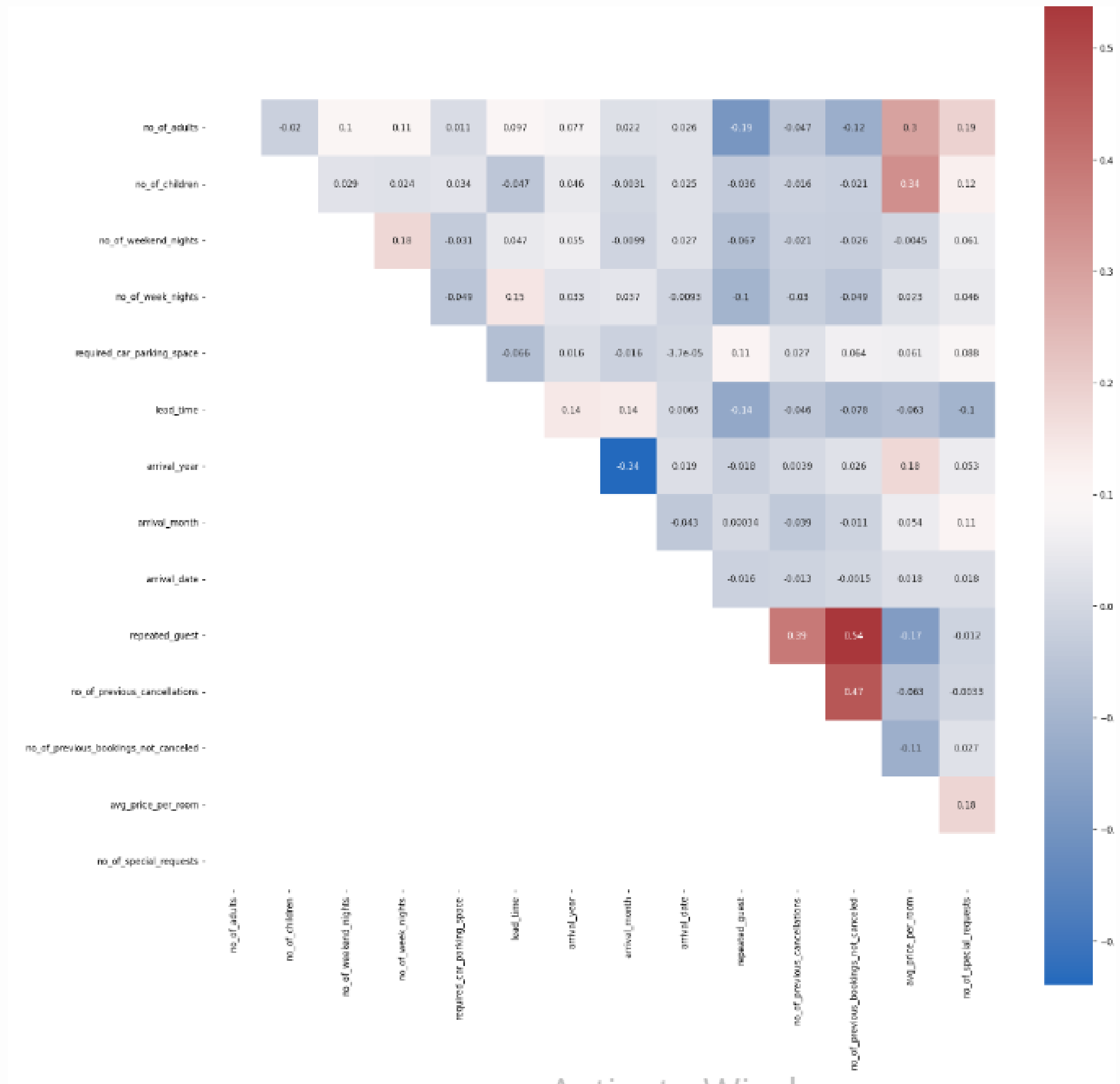
2. Biểu đồ

Biểu đồ thể hiện quan hệ giữa các yếu tố và tình trạng đơn đặt phòng



2. Biểu đồ

Biểu đồ tương quan giữa các biến



3. Mô hình

	Model	F1_Score	Accuracy
0	Random Forest	0.916299	0.884850
2	Decision Tree	0.904863	0.872186
7	GradientBoosting	0.890597	0.848296
4	SVC	0.859619	0.802077
1	KNN	0.859666	0.799963
3	Logistic Regression	0.846494	0.782780
5	SGD	0.834721	0.742810
6	XGB	0.919669	0.742810

Với điểm accuracy score:

- Random Forest cho kết quả cao nhất với 88,485%
- XGB và SGD cho kết quả thấp nhất với 74,281%

Hệ số F1_score:

- XGB cho kết quả cao nhất với 91,9669%
- SGD cho kết quả thấp nhất với 83,4721%

Confusion Matrix

True Positive: 6464

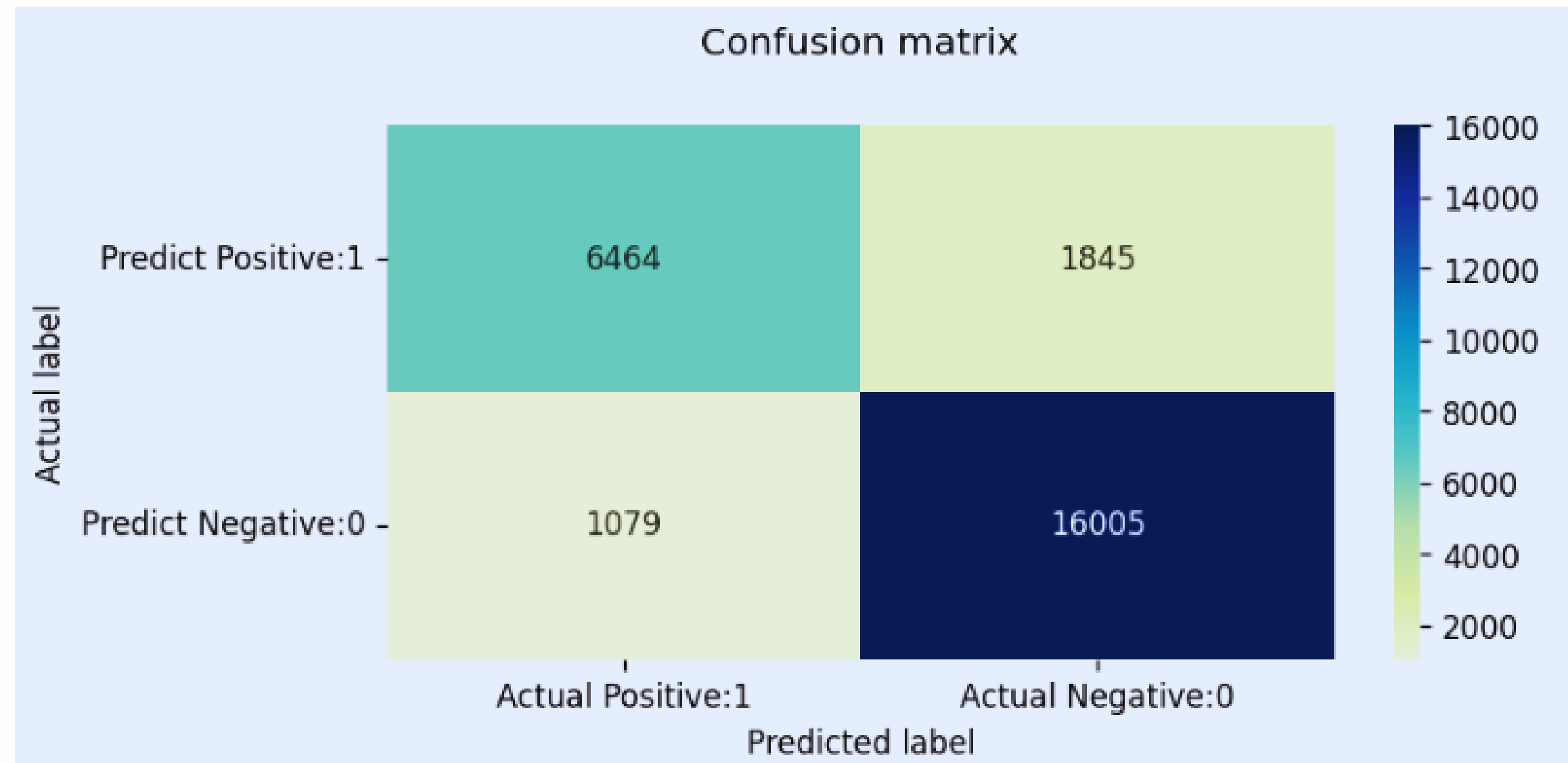
True Negative: 16005

False Positive: 1845

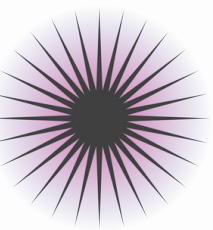
False Negative: 1079

Booking_status: 'Not Canceled' = 1

Booking_status: 'Canceled' = 0



Classification Report



	precision	recall	f1-score	support
0	0.86	0.78	0.82	8309
1	0.90	0.94	0.92	17084
accuracy			0.88	25393
macro avg	0.88	0.86	0.87	25393
weighted avg	0.88	0.88	0.88	25393

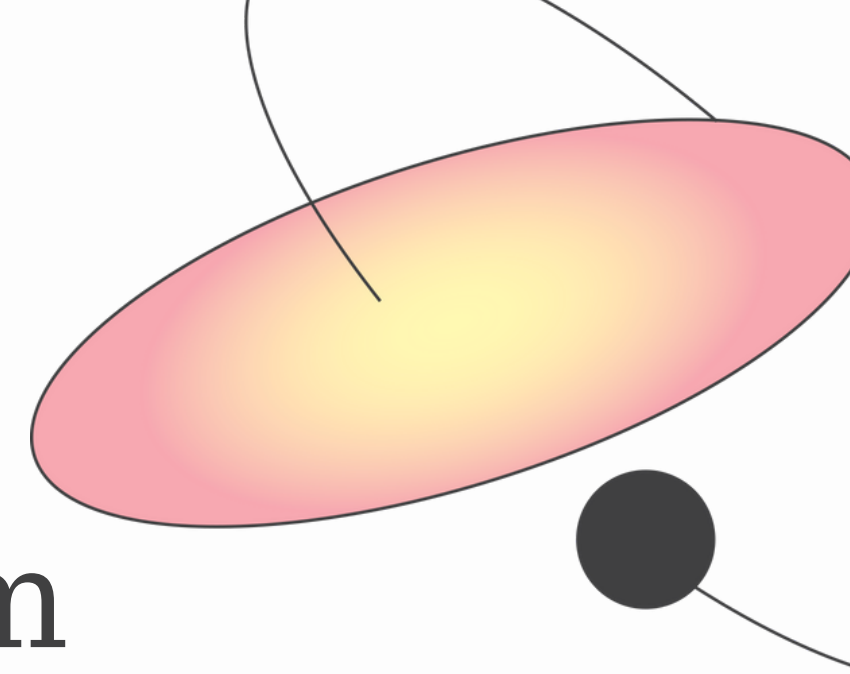
Trạng thái booking 0 và 1 có hệ số F1-score tương ứng là 0.82 và 0.92

=> Cao

Khả năng dự báo của model tốt



Thực tiễn



Ưu điểm

Tính toán tốt hơn về mặt nhân sự, chi phí (với khách sạn lớn)

- Mở thêm phòng trên nền tảng đặt phòng online => Tạo thêm doanh thu

- Thay đổi luật huỷ đặt phòng hợp lý

Nhược điểm

- Hậu quả của việc dự đoán sai là rất lớn.

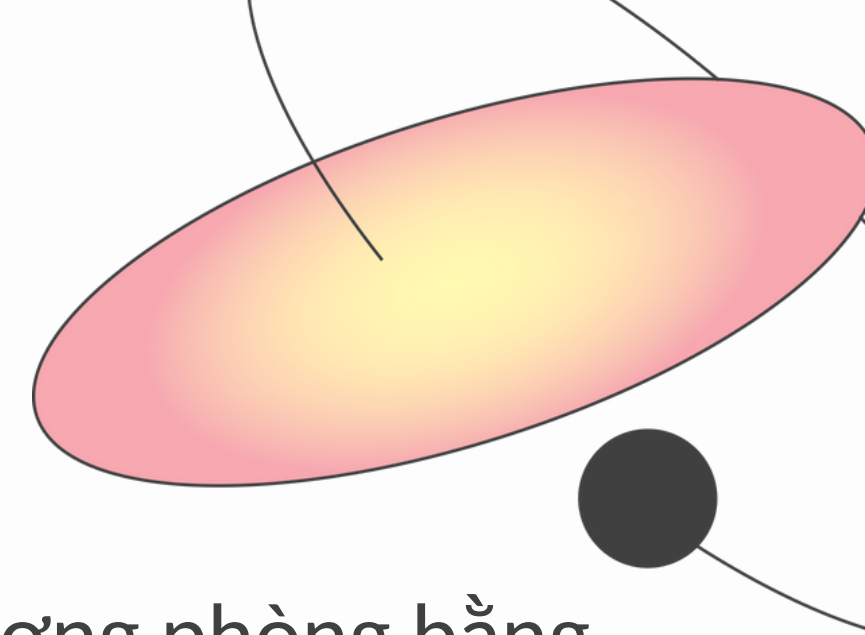
- Ví dụ, khách không huỷ

=> khách sạn thiếu phòng => phải book phòng khác cho khách và chịu toàn bộ chi phí

- Khách gây náo loạn khách sạn, gây ảnh hưởng đến các khách hàng khác

- Khách hàng review không tốt trên hệ thống, hạ uy tín của khách sạn (nặng nề nhất)

Giải pháp



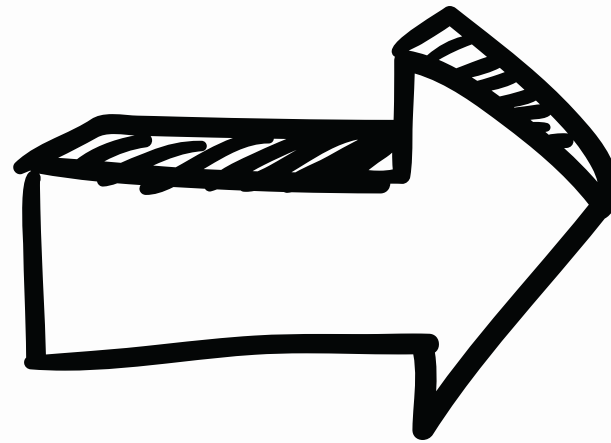
Với trường hợp của bài phân tích:

- Tỷ lệ huỷ phòng quá lớn

=> Ảnh hưởng doanh thu, có thể khách sạn rơi lòng luật huỷ phòng (thời gian cho phép quá ngắn)

- Tỷ lệ khách quay quá thấp

=> Có thể do chất lượng dịch vụ đang gặp vấn đề



- Mở thêm một lượng phòng bằng 80% số phòng sẽ bị huỷ theo mô hình

=> Tăng doanh thu và phòng ngừa rủi ro

- Siết luật huỷ phòng bằng cách nâng số ngày trước hạn được phép huỷ miễn phí

=> Tạo thêm rào cản huỷ phòng

- Cải thiện chất lượng dịch vụ



**Cảm ơn bạn
đã chú ý
lắng nghe!**