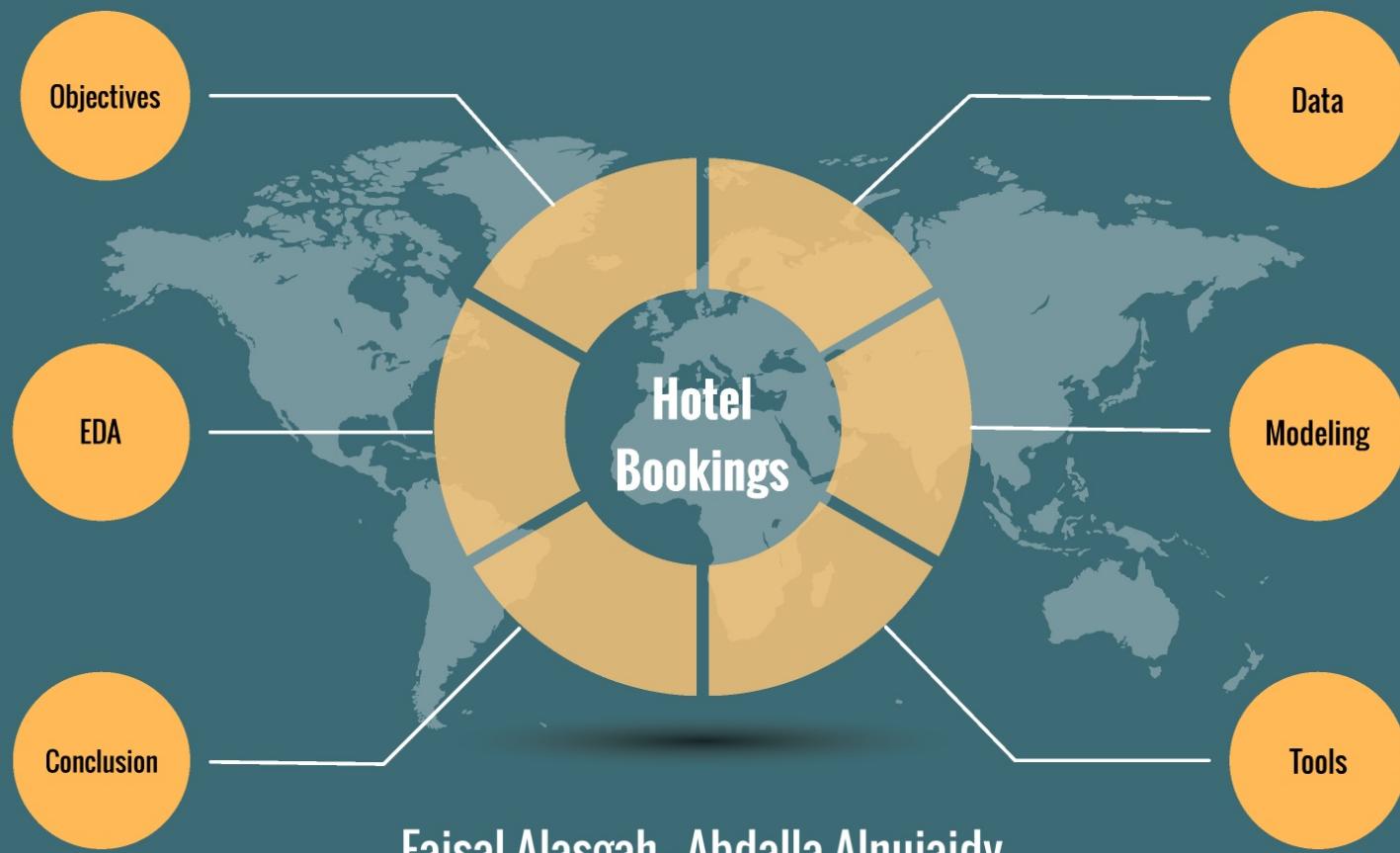


Faisal Alasgah - Abdalla Alnujaidy

- **Get the data**
- **EDA**
 - Data cleaning
 - Visualization and the relations between features.
 - Recommendations.
- **Modeling**
 - Data preparation
 - Apply different classification algorithms.
 - Compare between the results.



Faisal Alasgah - Abdalla Alnujaidy

Description of the Data

Features

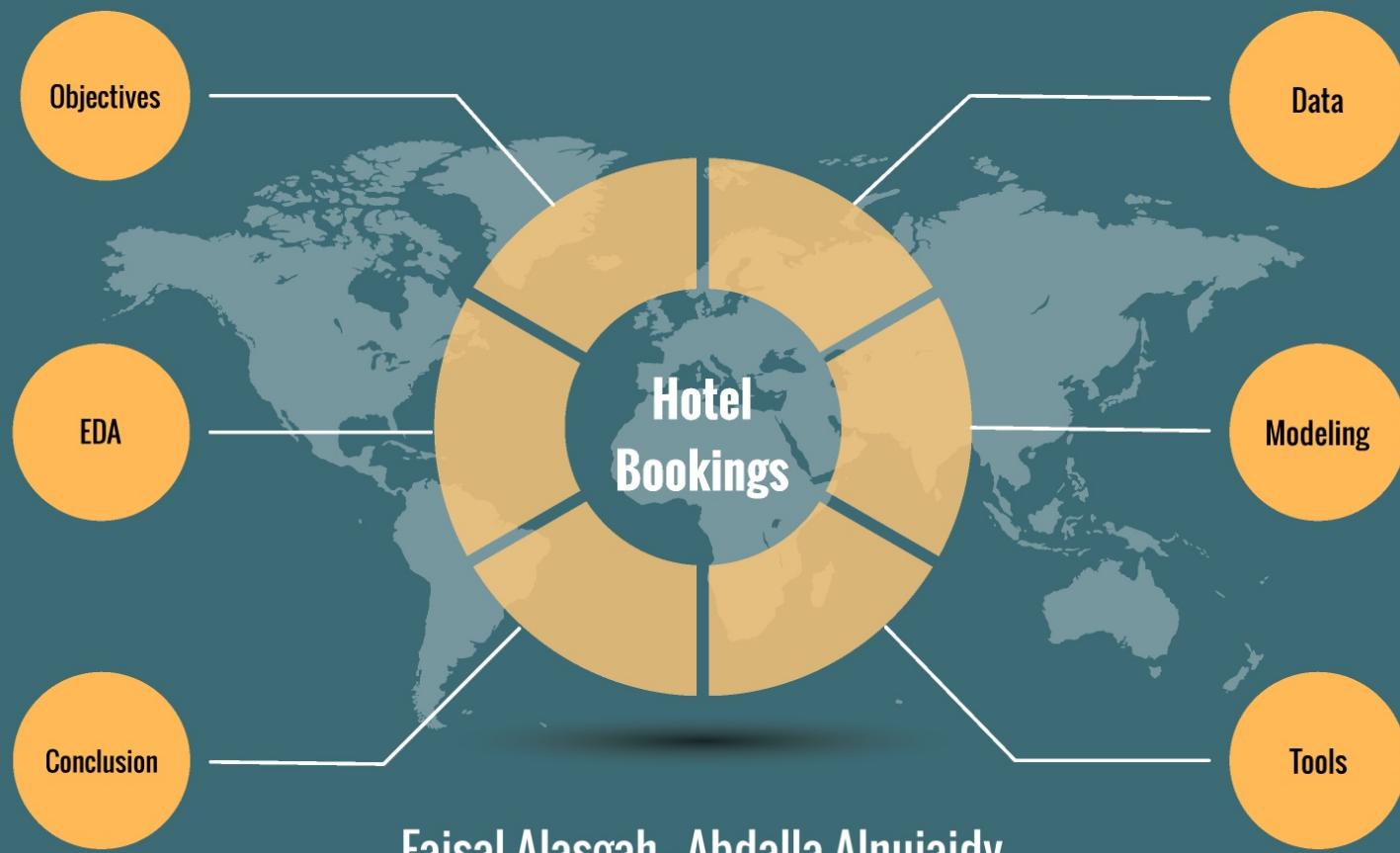
Source: Kaggle

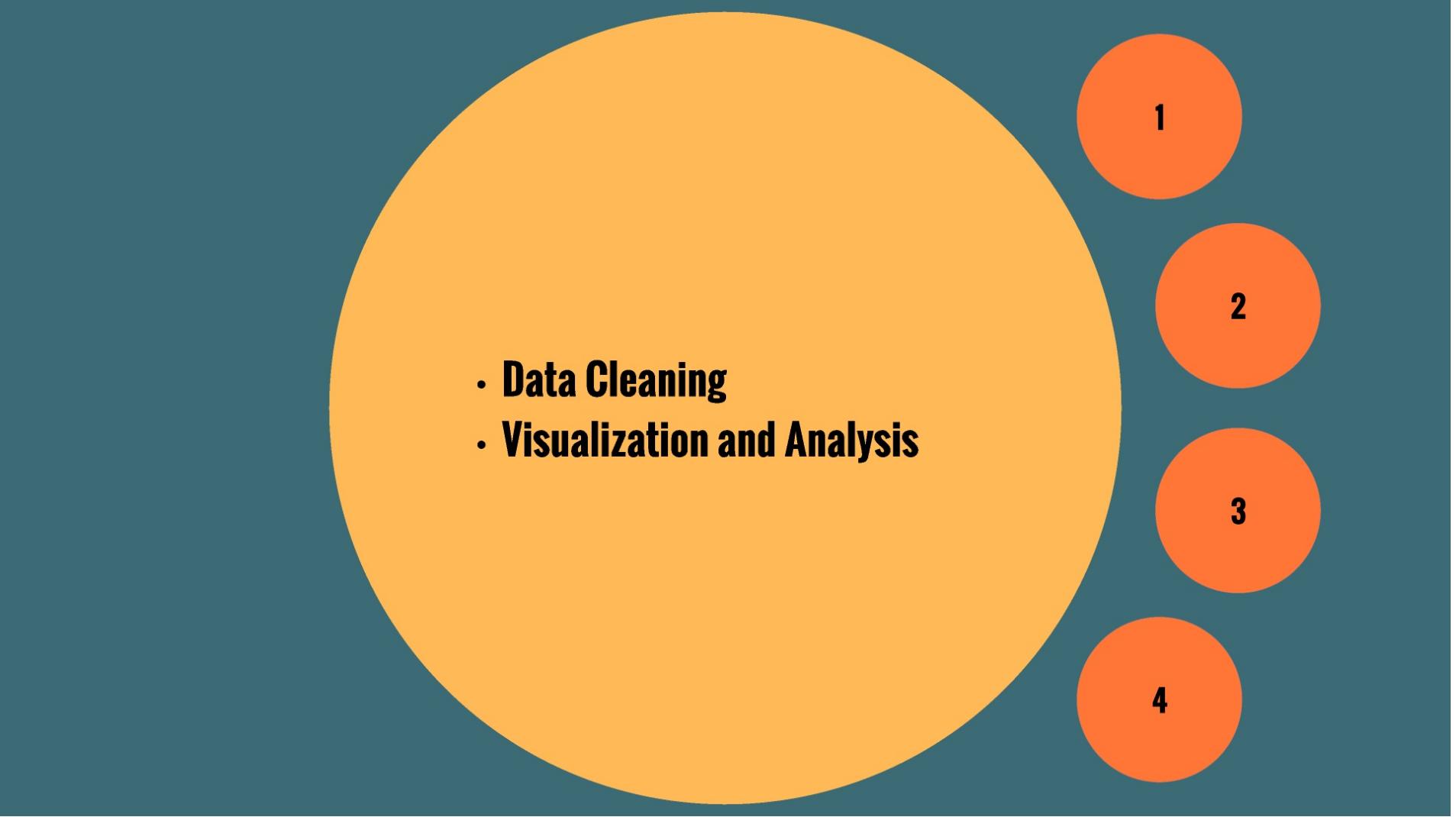
32 Features

Main Features:

- Type of Hotels
- Price
- Reservation status
- Number of guest
- Date
- Lead time
- Reservation canceled or not

119,000 Rows of reservations



- 
- Data Cleaning
 - Visualization and Analysis

1

2

3

4

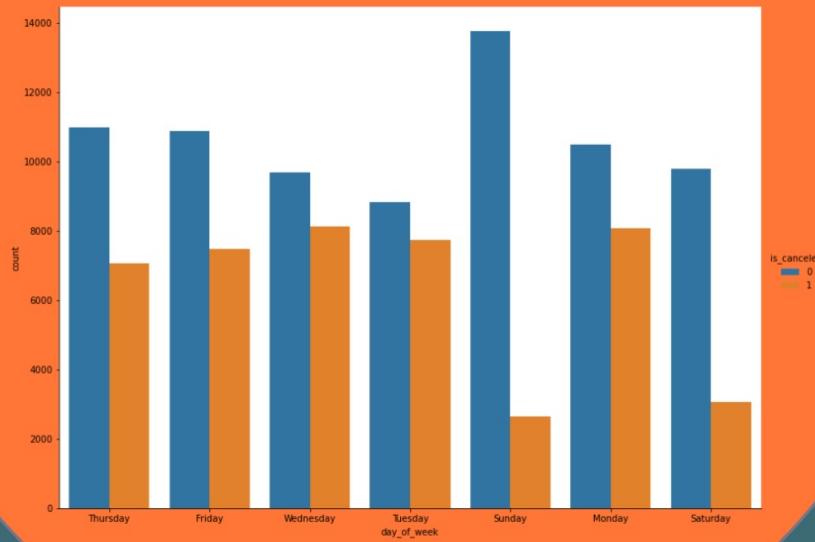
Which are the most busy months?



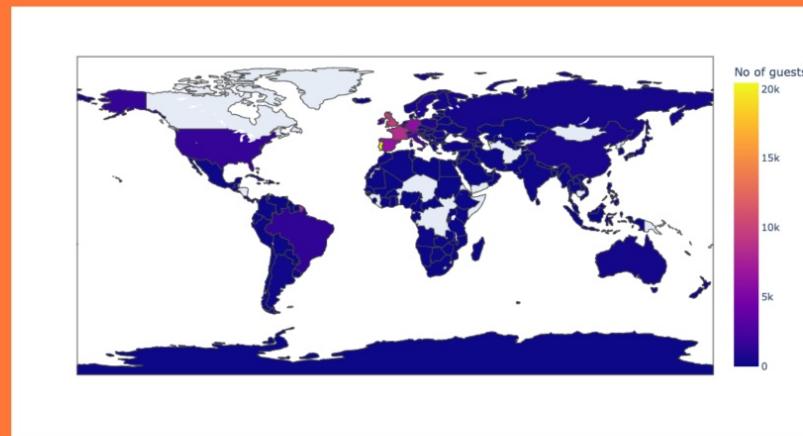
How does the price vary per night over the year?

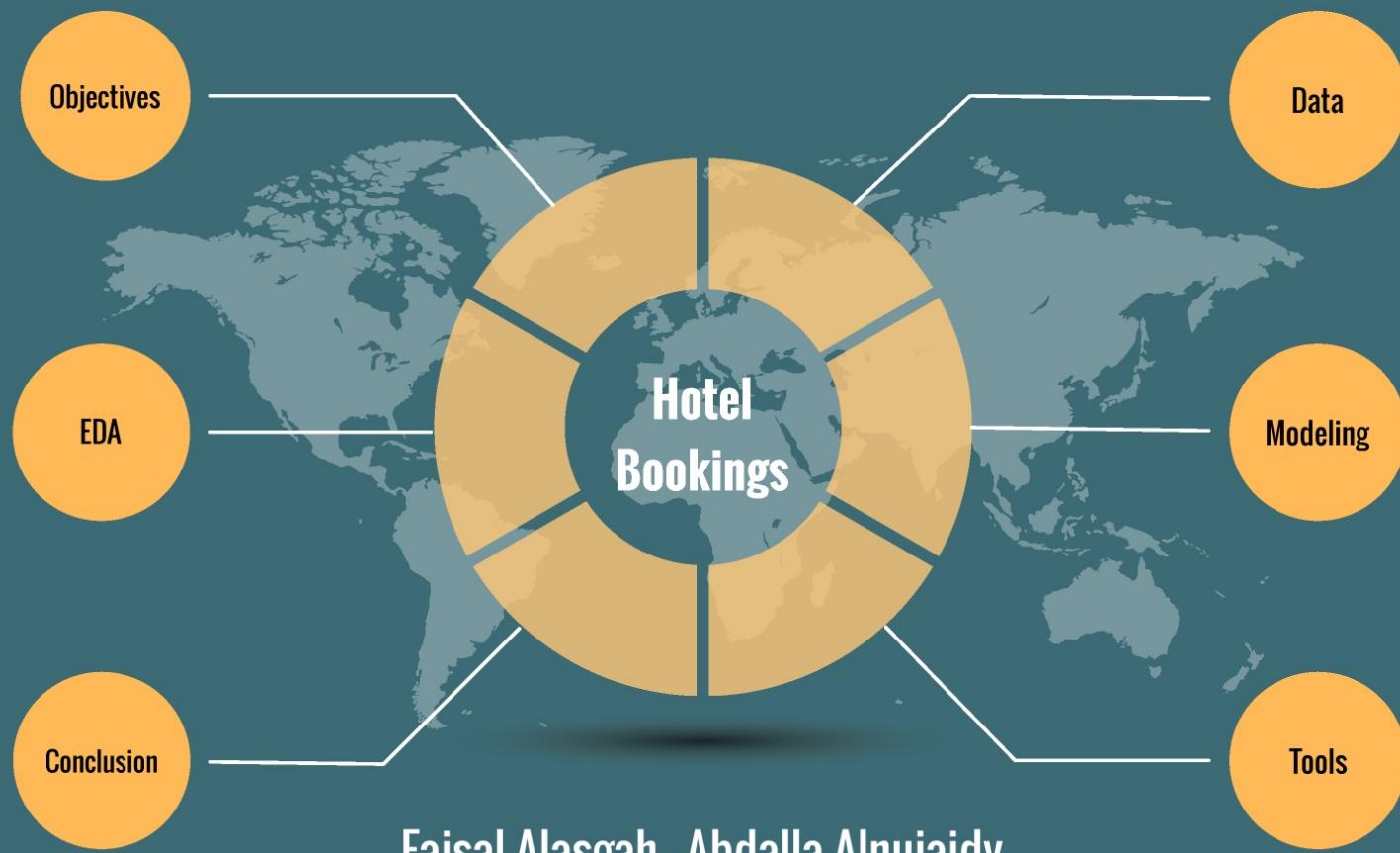


In which day people usually cancel their reservation?



From where the most guests are coming ?





Faisal Alasgah - Abdalla Alnujaidy

Modeling

- KNN
- Logistic Regression
- Decision Tree
- Random Forest

Data Prep

Models

Feature selecting and engineering

- Dropping unwanted columns.
- Split the reservation date into three columns.
- Maping some columns to binary variables.

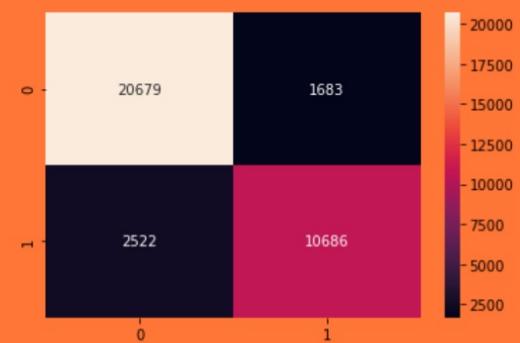
KNN_5

Accuracy
Precision
Recall
F1_score

86.404
88.688
72.653
79.873

KNN_1

88.178
86.393
80.906
83.559



Logistic regression

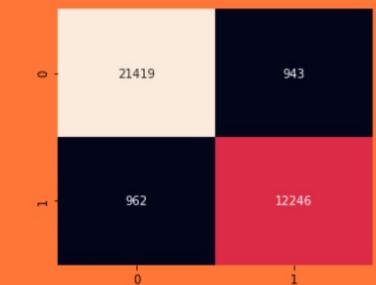
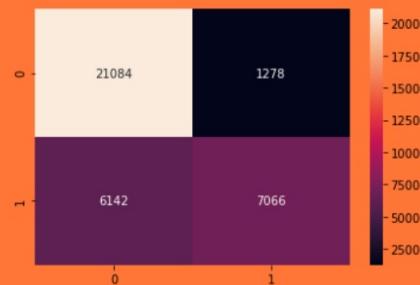
Accuracy	79.140
Precision	84.684
Recall	53.500
F1_score	65.572

Decision tree

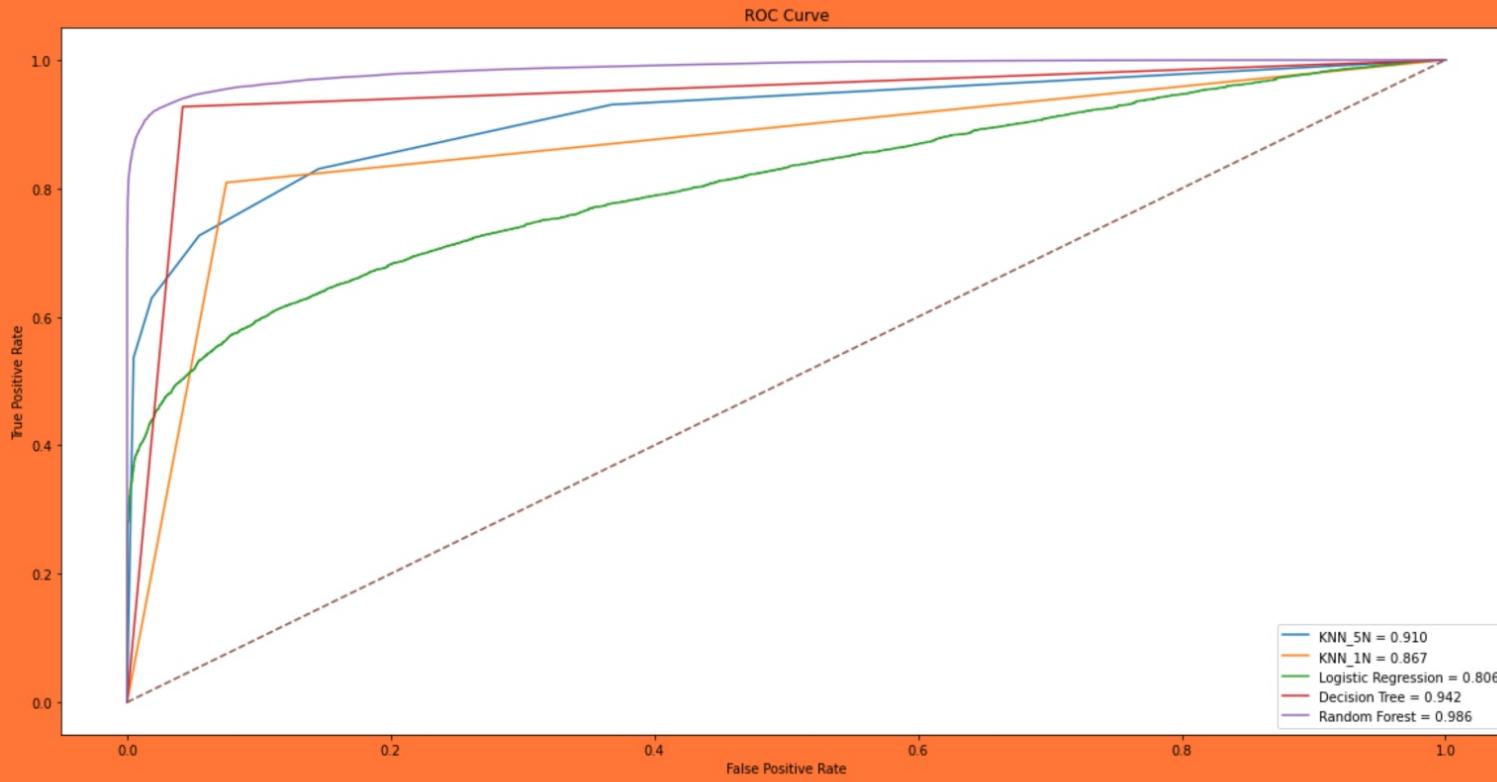
94.644
92.850
92.717
92.783

Random forest

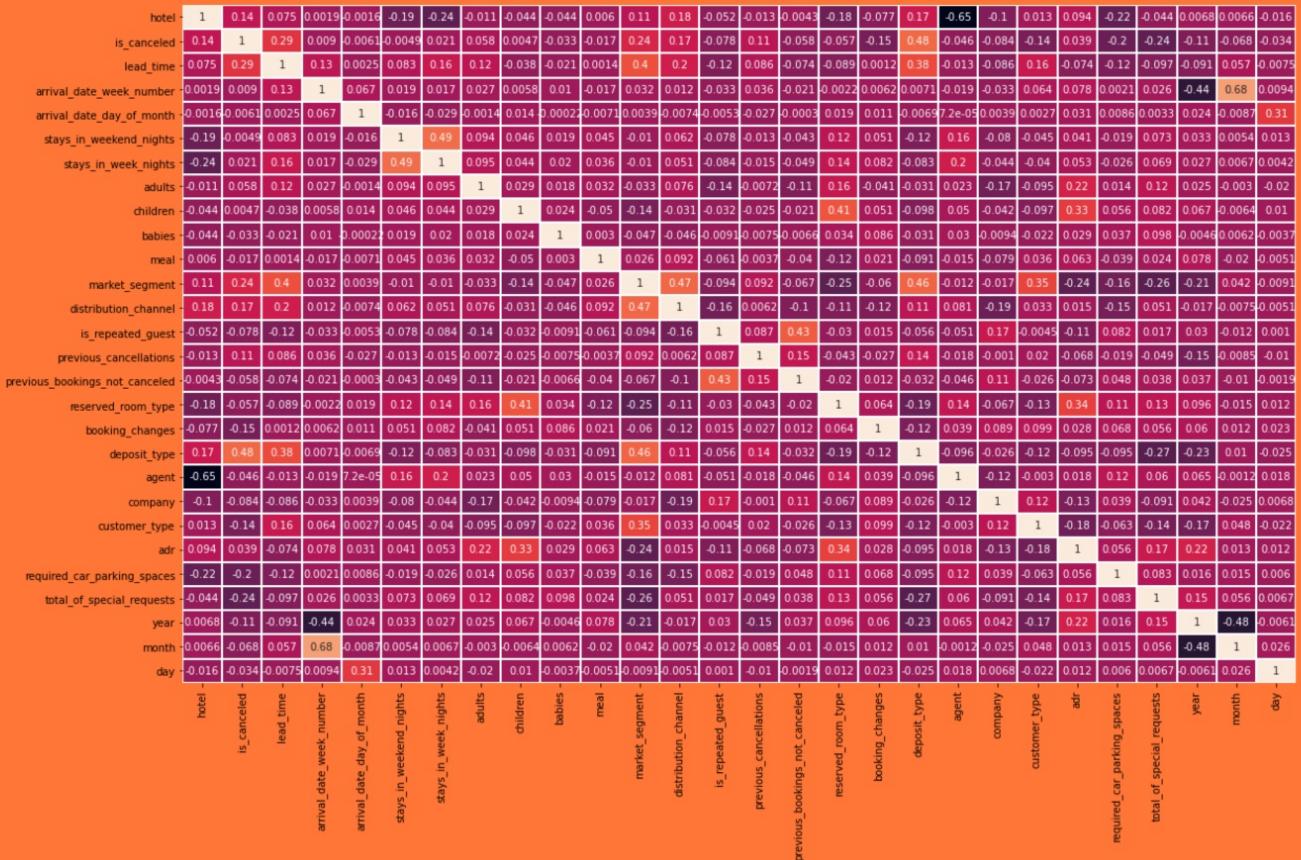
95.100
98.684
87.977
93.023

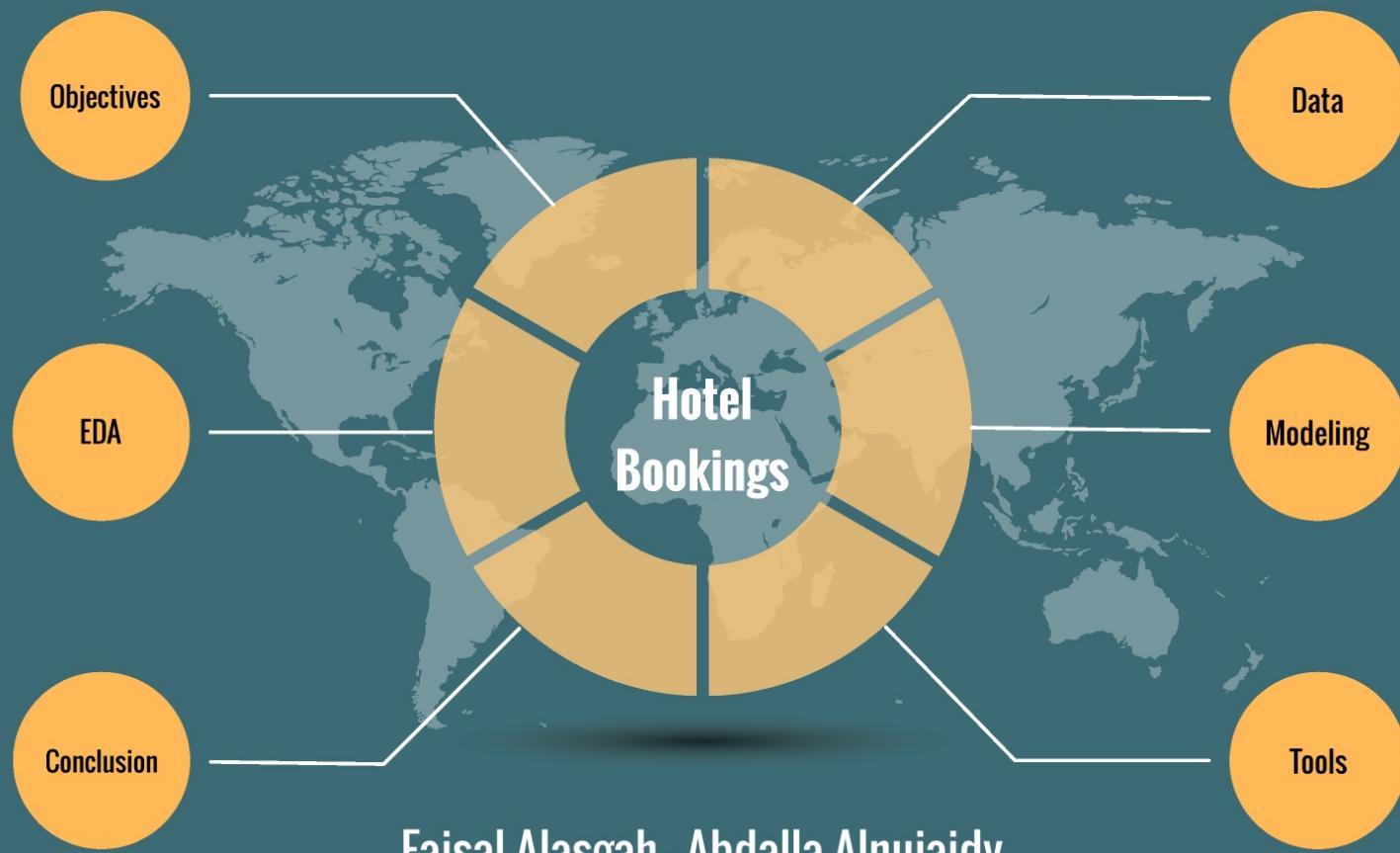


ROC



Results

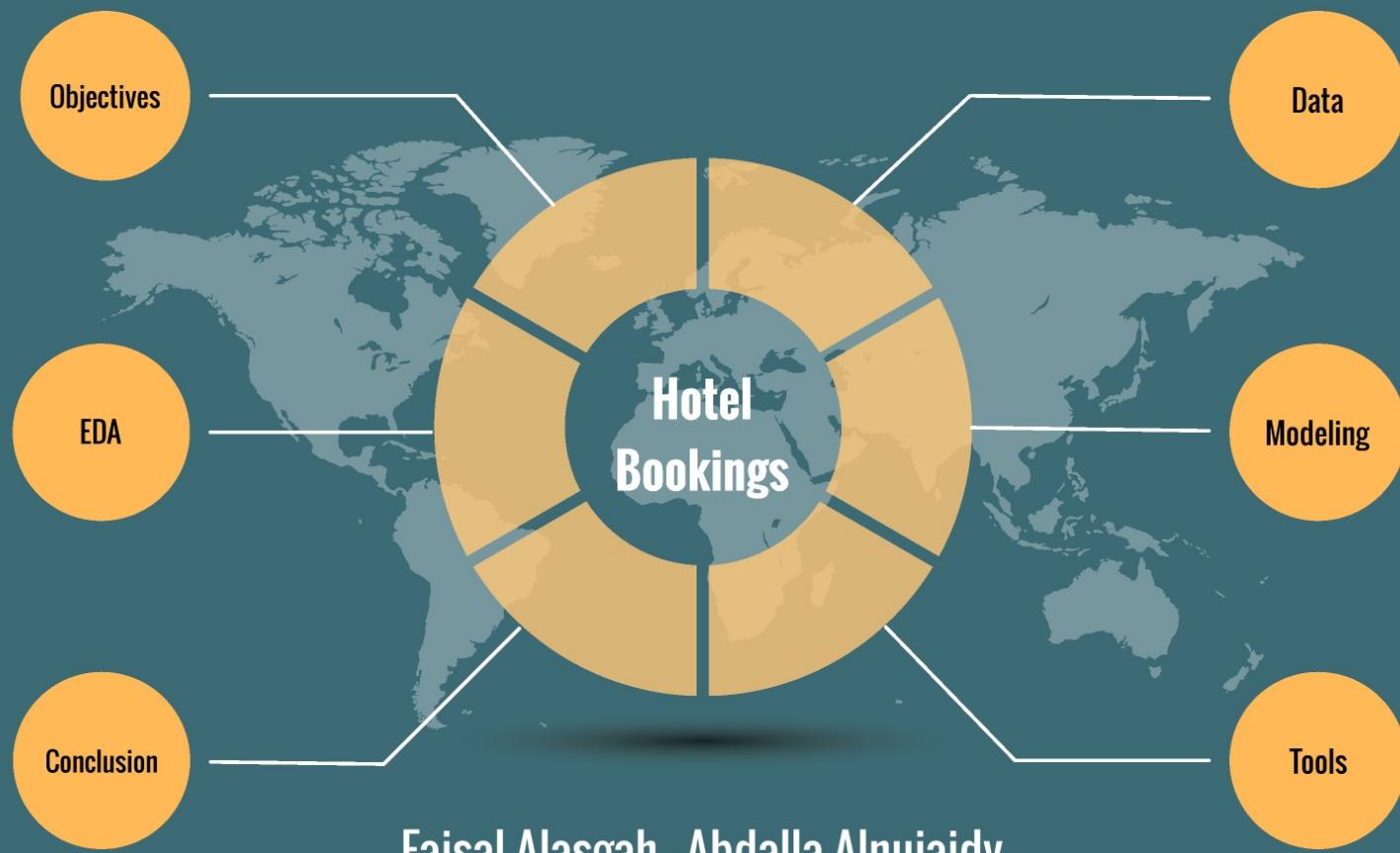




Faisal Alasgah - Abdalla Alnujaidy

Recommendations

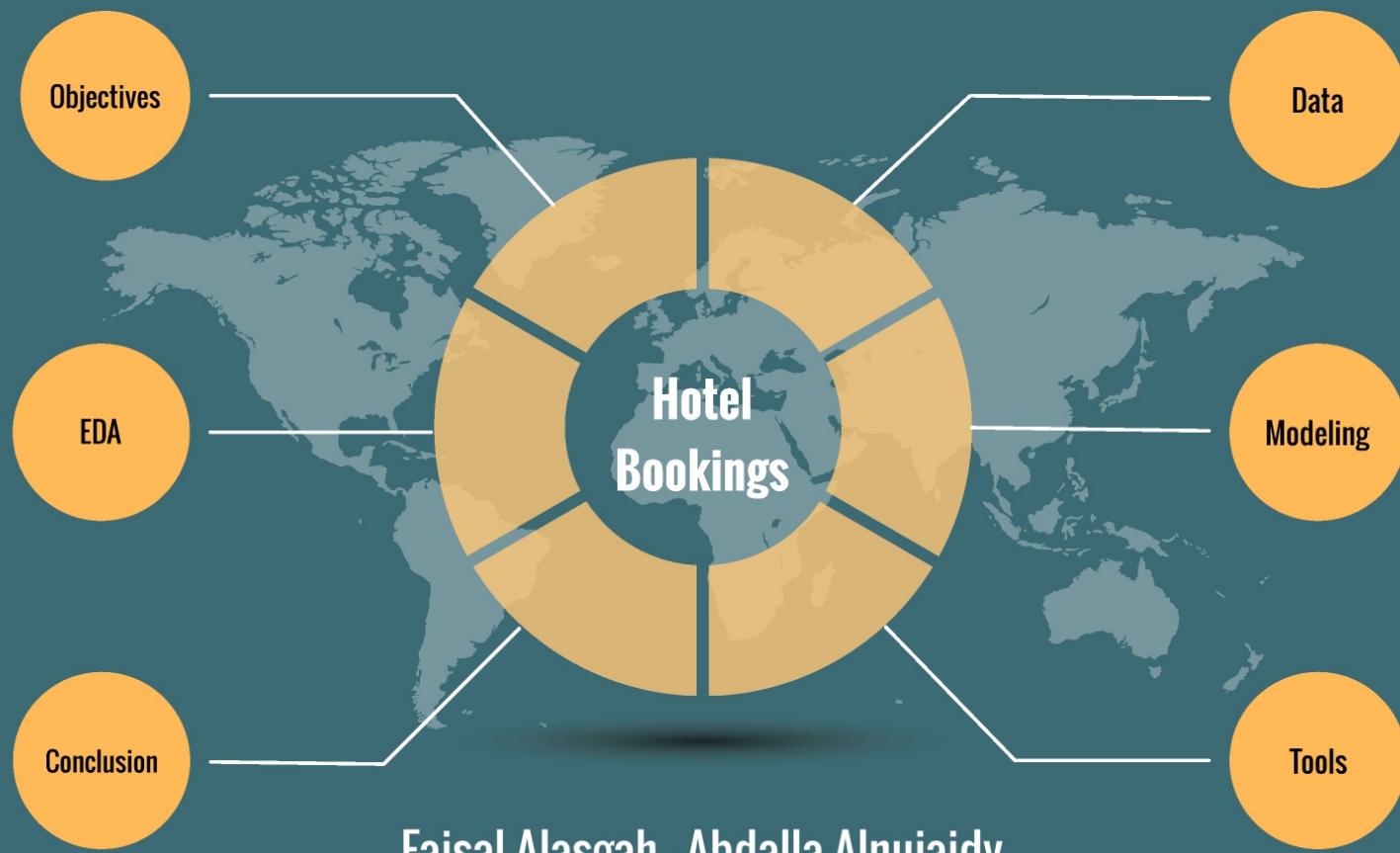
- Give offers on Monday and Wednesday where cancellation average is high.
- Keep the pierces as it is, because they are reasonable.
- Reduce the non-refundable deposit type.



Faisal Alasgah - Abdalla Alnujaidy

Libraries

- Pandas
- Numpy
- Seaborn
- Matplotlib
- Plotly
- Folium
- Sklearn



Faisal Alasgah - Abdalla Alnujaidy

Thank You ...

Any Questions ?