HOTEL BOOKING

classifcation

T5 Bootcamp Data Science Project

Jawaher Alzahrani

OVERVIEW

- Data
- Project Goal
- Process
- Challenges
- Modeling

HOTEL BOOKING DEMAND

Kaggle

119390
Reservations

32
Features

(is_canceled)
Target variable

PROJECT GOAL

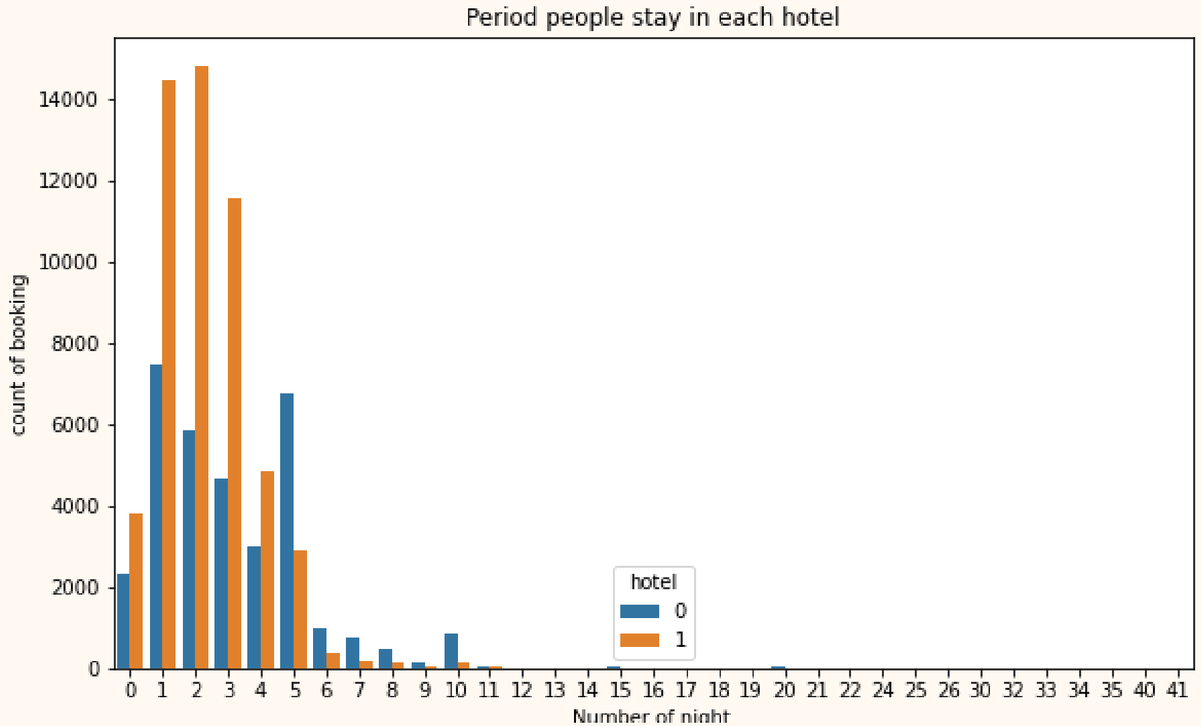
- The goal of this project was to use classification models to predict the hotel booking's cancellation to help improve operations and planning to arrange of these reservations

PROCESS

EDA → Data Cleaning → Modeling

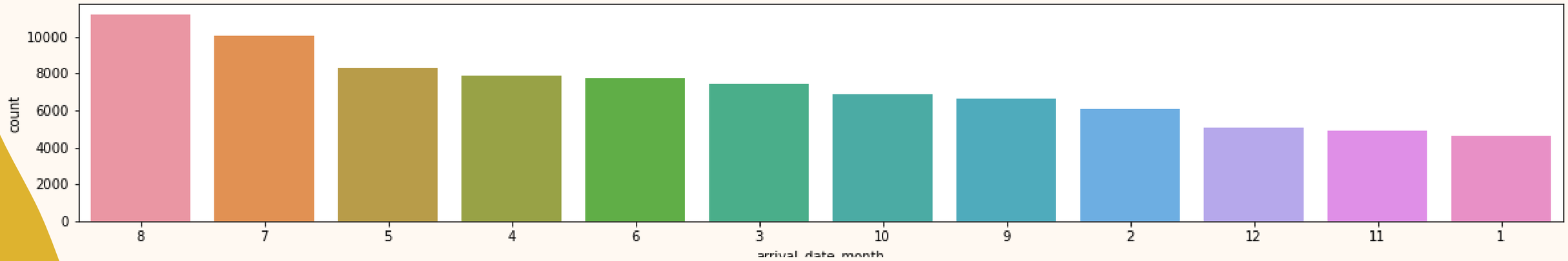
EDA AND VISUALIZATION

How Long People Stay in the hotel?



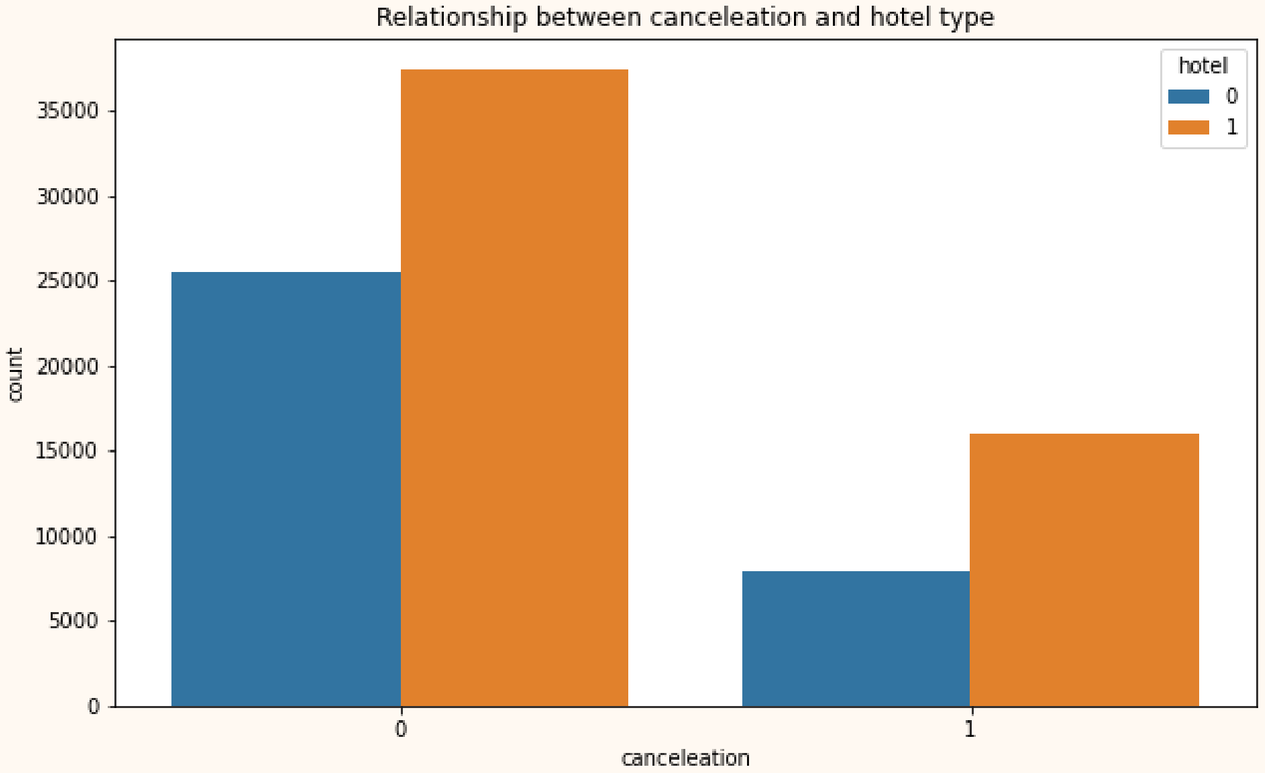
EDA AND VISUALIZATION

Which is the most busy month for hotel?



EDA AND VISUALIZATION

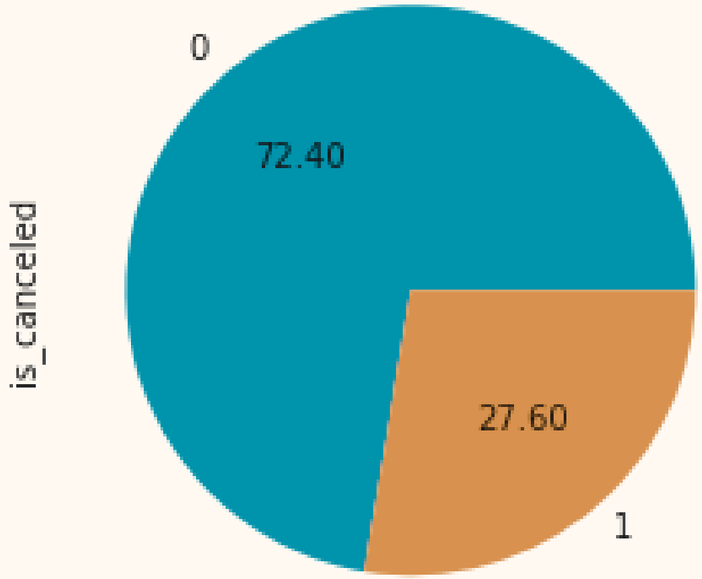
What is the proportion of cancelation between different hotel types ?



CHALLENGES

Imbalanced data

Score Models =100%



MODLING

LOGISTIC REGRESSION:

Accuracy: 97%

Precision: 94%

Recall: 95%

F1: 94.6%

KNN:

Accuracy: 97.4%

Precision: 94%

Recall: 96.8%

F1: 95%

RANDOM FOREST

Accuracy: 97.8%

Precision: 95.5%

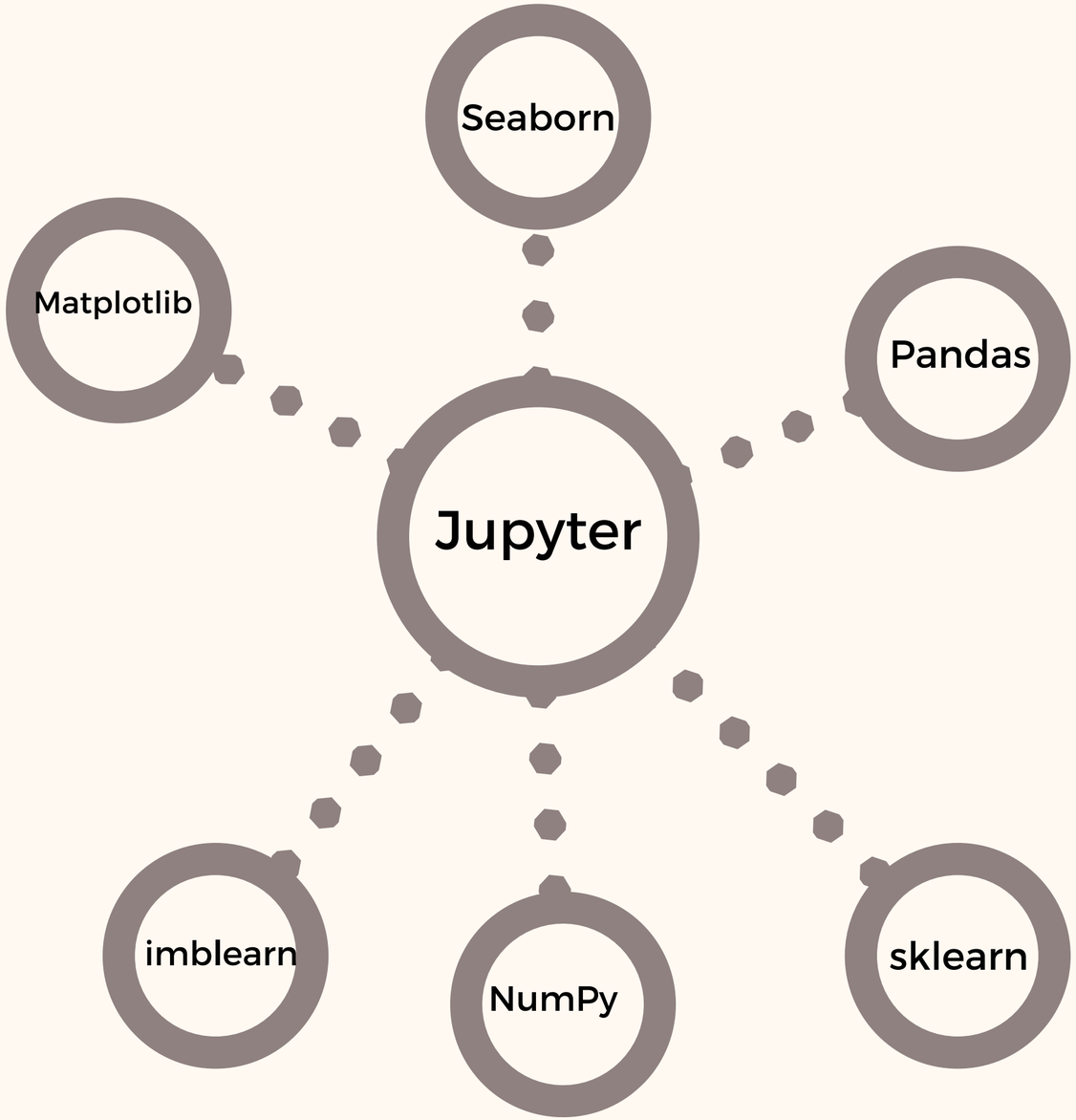
Recall: 96.6%

F1: 96%

CONCLUSION

To sum up, with this information hotels can, for example, contact clients that the model predicted will cancel in order to get a cancellation earlier - so they can have more time to resell the room. Or perhaps approach the client in a way to make them feel special and keep their reservation and therefore cancel the others he or she had made in other hotels in the same city.

TOOLS


Seaborn
Matplotlib
Pandas
Jupyter
imblearn
NumPy
sklearn

THANK YOU!