Predicting Reservation Cancellations for Hilton's Resorts

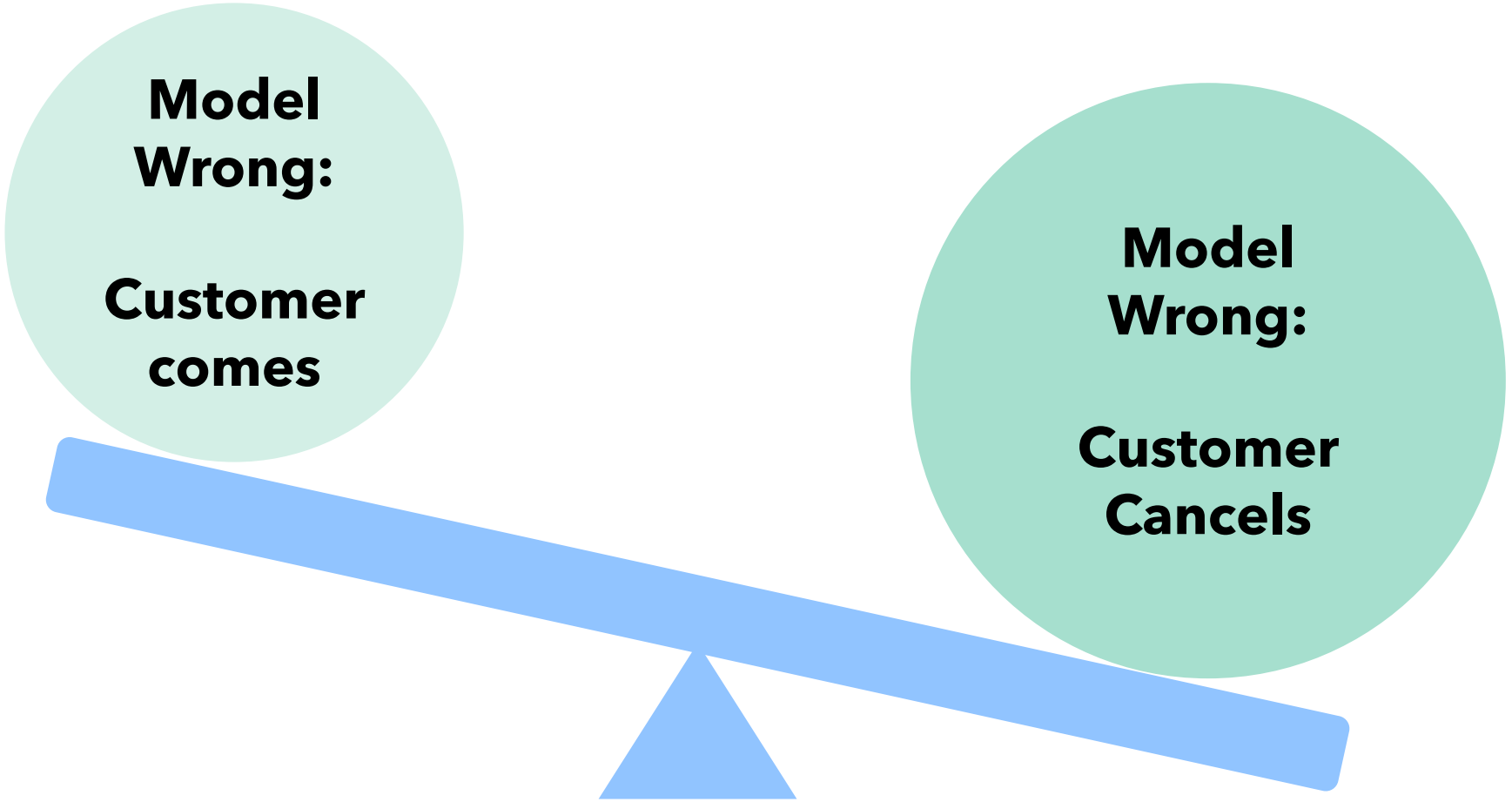
Jae Heon Kim

Outline

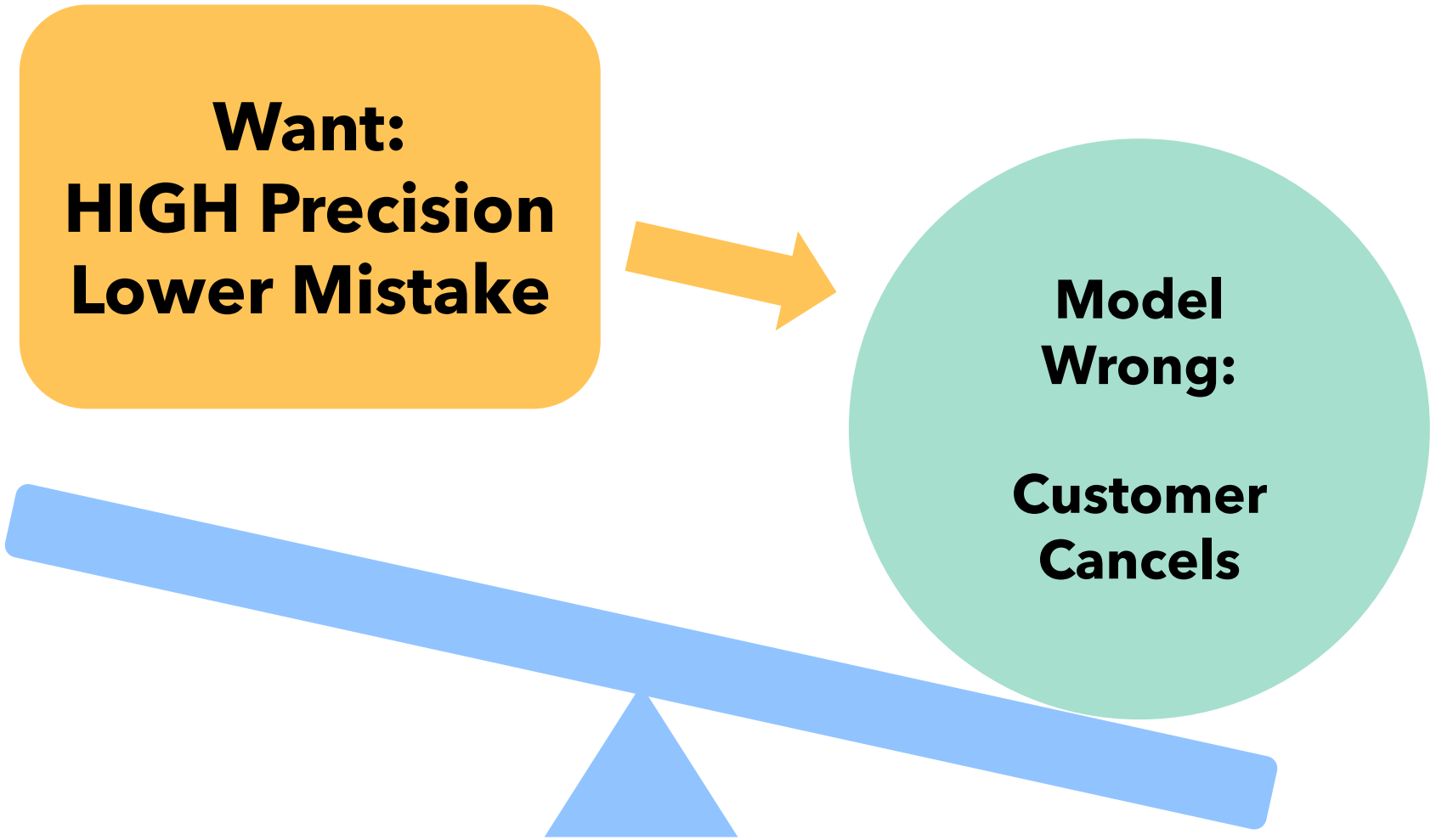
Problem	Data	Modeling	Evaluation	Conclusion
- Introduce and understand problem	- Explore data - Optimize data	- Build different models - Further tune models that work well	- Select the final model - Evaluate the final model in depths	- Make meaningful suggestions

Business Problem

Two Types of Wrong Predictions



Definition: Precision



Data Introduction

Source

- Kaggle Data Sets
- A resort hotel in Algarve, Portugal
- Booking information in 2017 and 2018

Features

- 19 features
- Information on :
  - lead time
  - meal plan type
  - repeated visits
  - room type
  - number of adults
  - number of children, etc..

Size

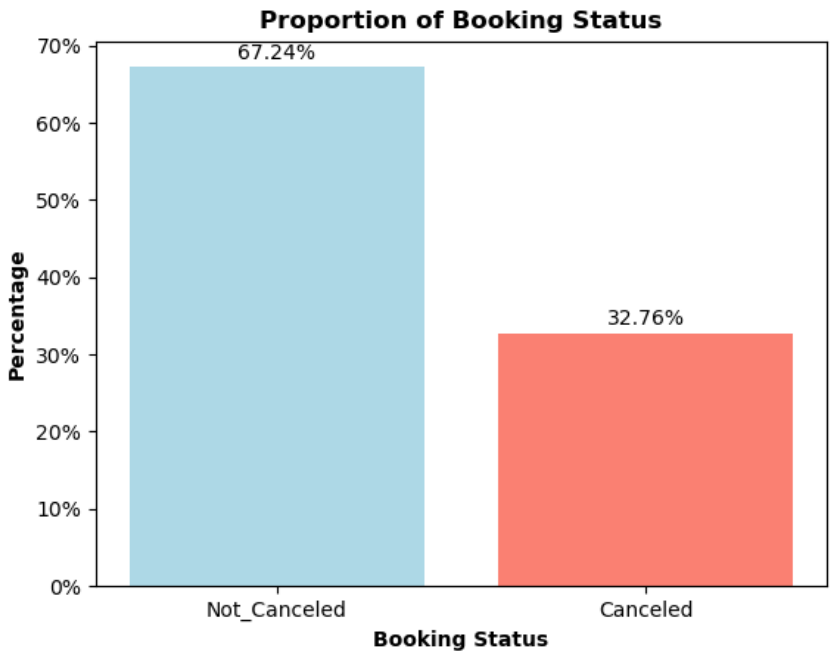
- 36,275 observations
- Very little missing values!

Working with Data

Preparing the Data

- Studied each of 19 features
- **Target feature: Booking Cancellation Status**
- Excluded irrelevant and redundant predictor features
- Tailored each feature to be fed into the machine learning models

How many bookings are cancelled?



Modeling Process

Step 1

- Try Different Models

Step 2

- Choose Good Models

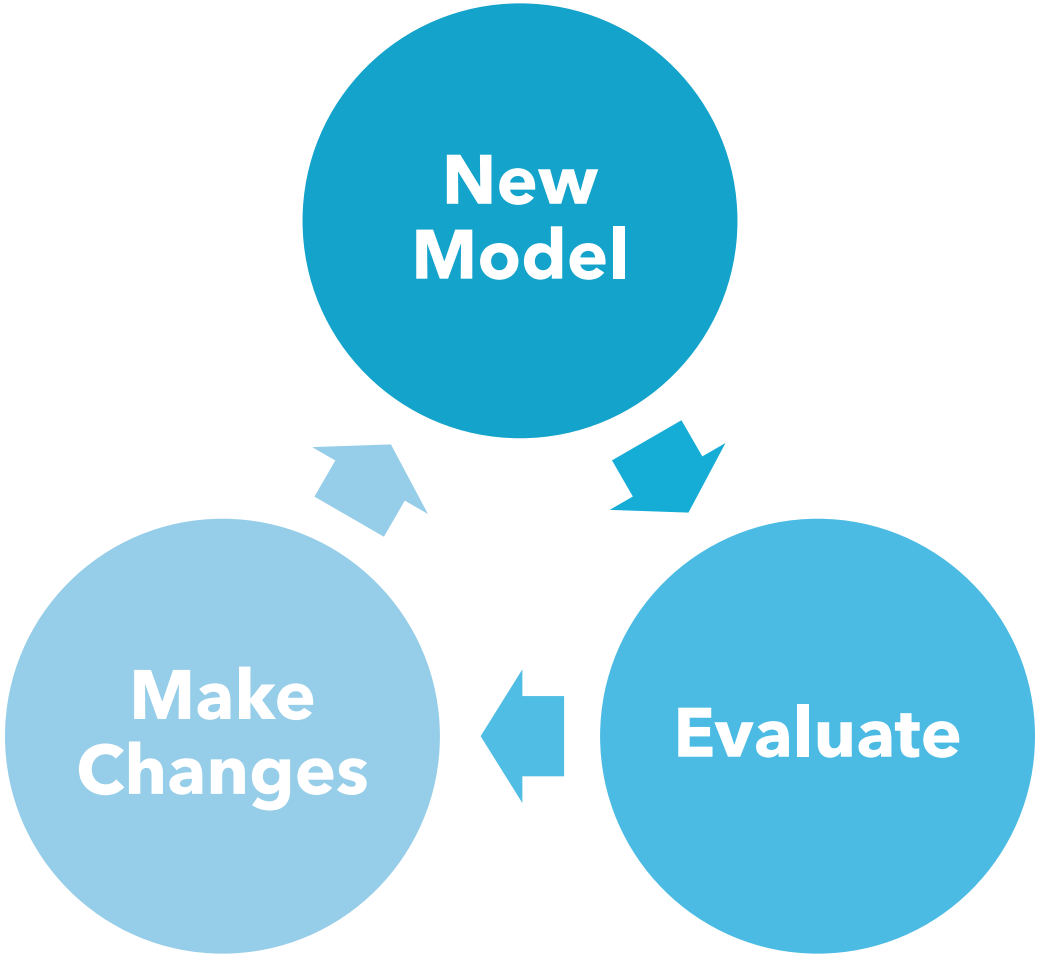
Step 3

- Combine Good Models to Make Better Models

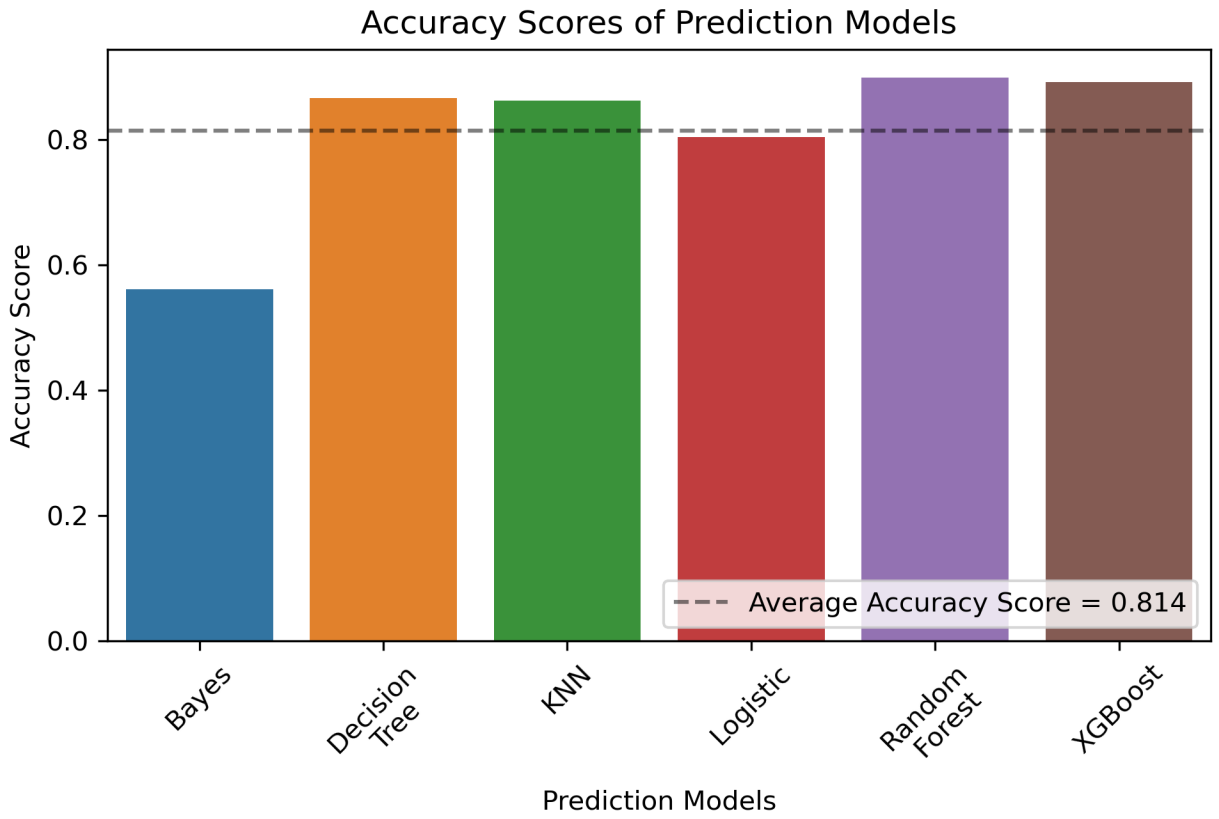
Step 4

- Compare Better Models and Choose the BEST MODEL

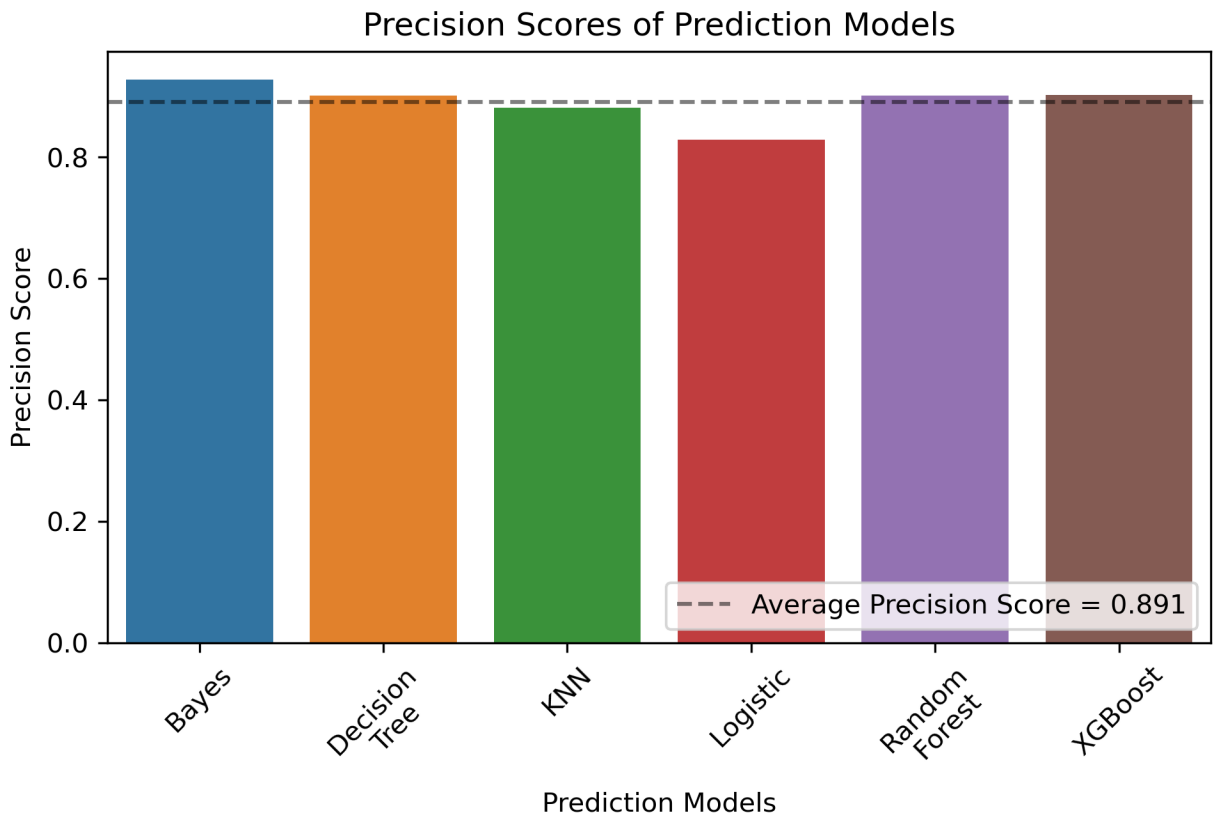
Modeling Process – Step 1



Comparing Models – Accuracy



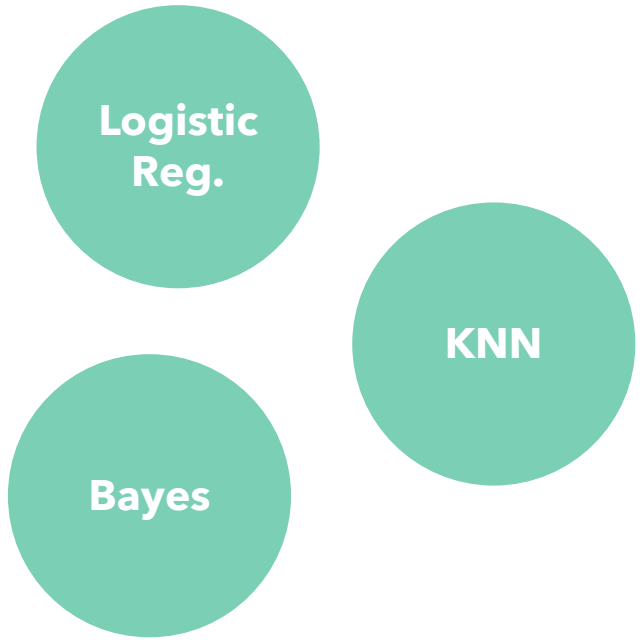
Comparing Models – Precision



Modeling Process – Step 2

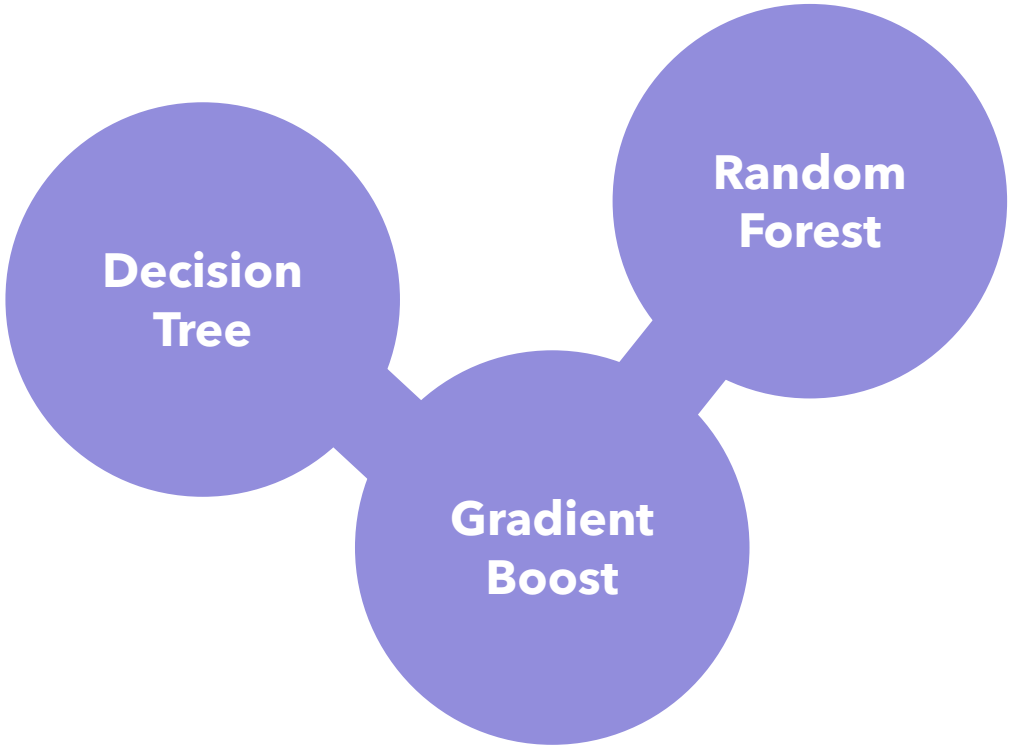
Good Models

Bad Models

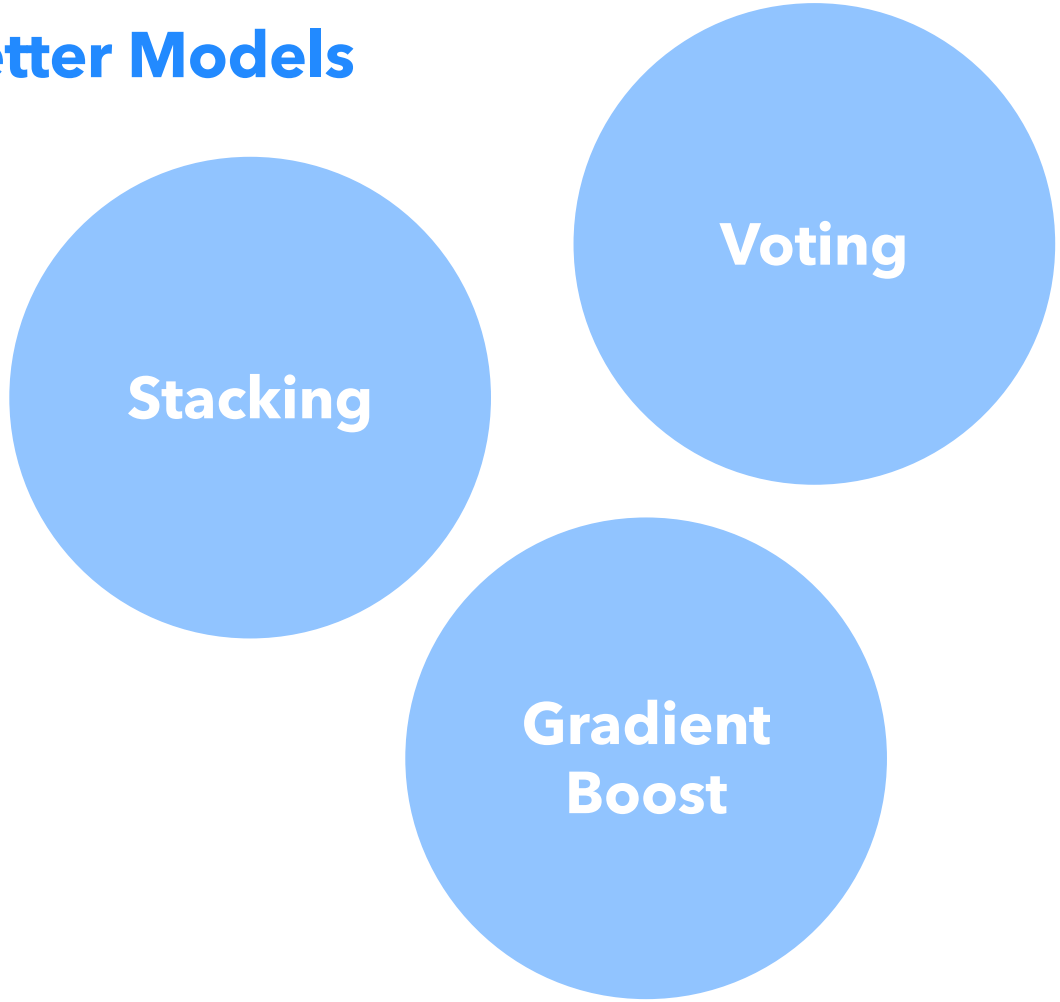


Modeling Process – Step 3

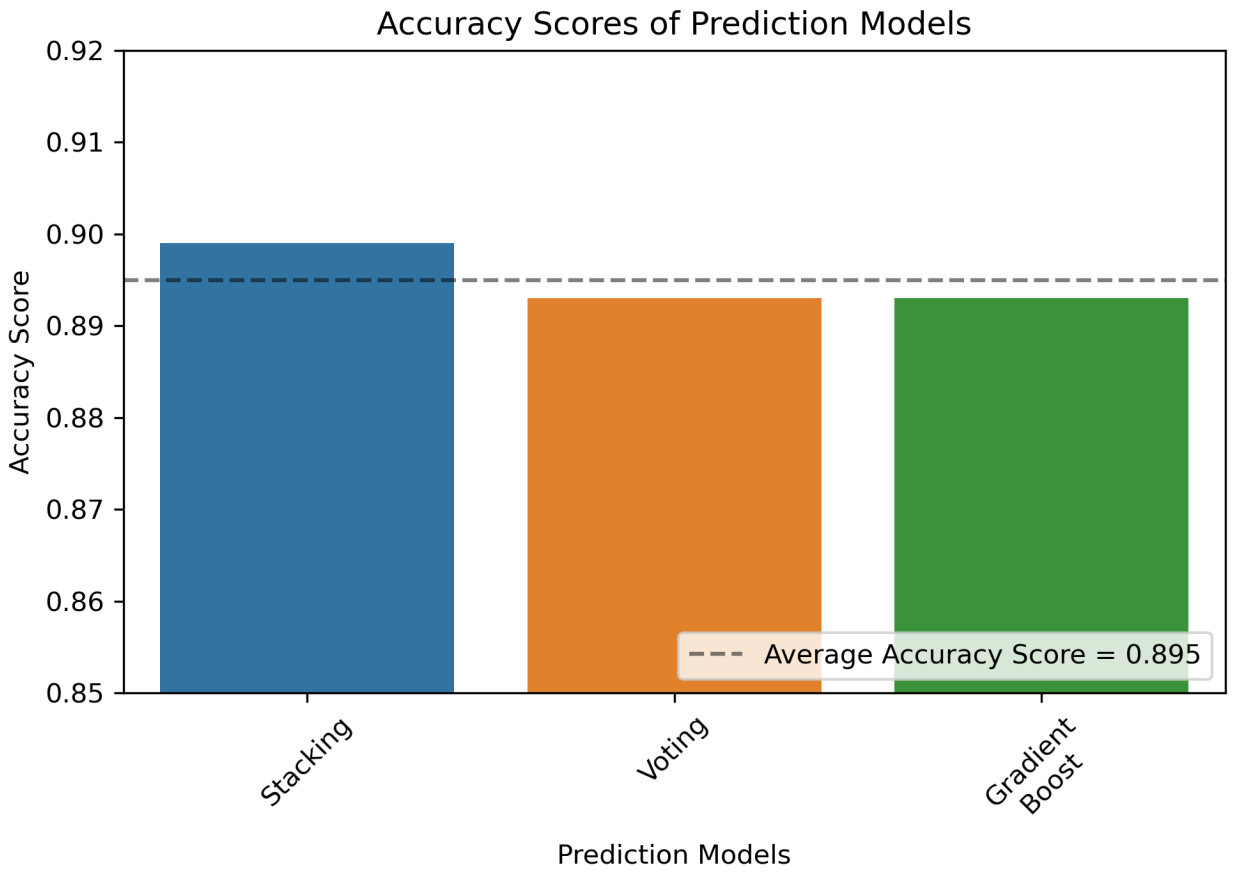
Good Models



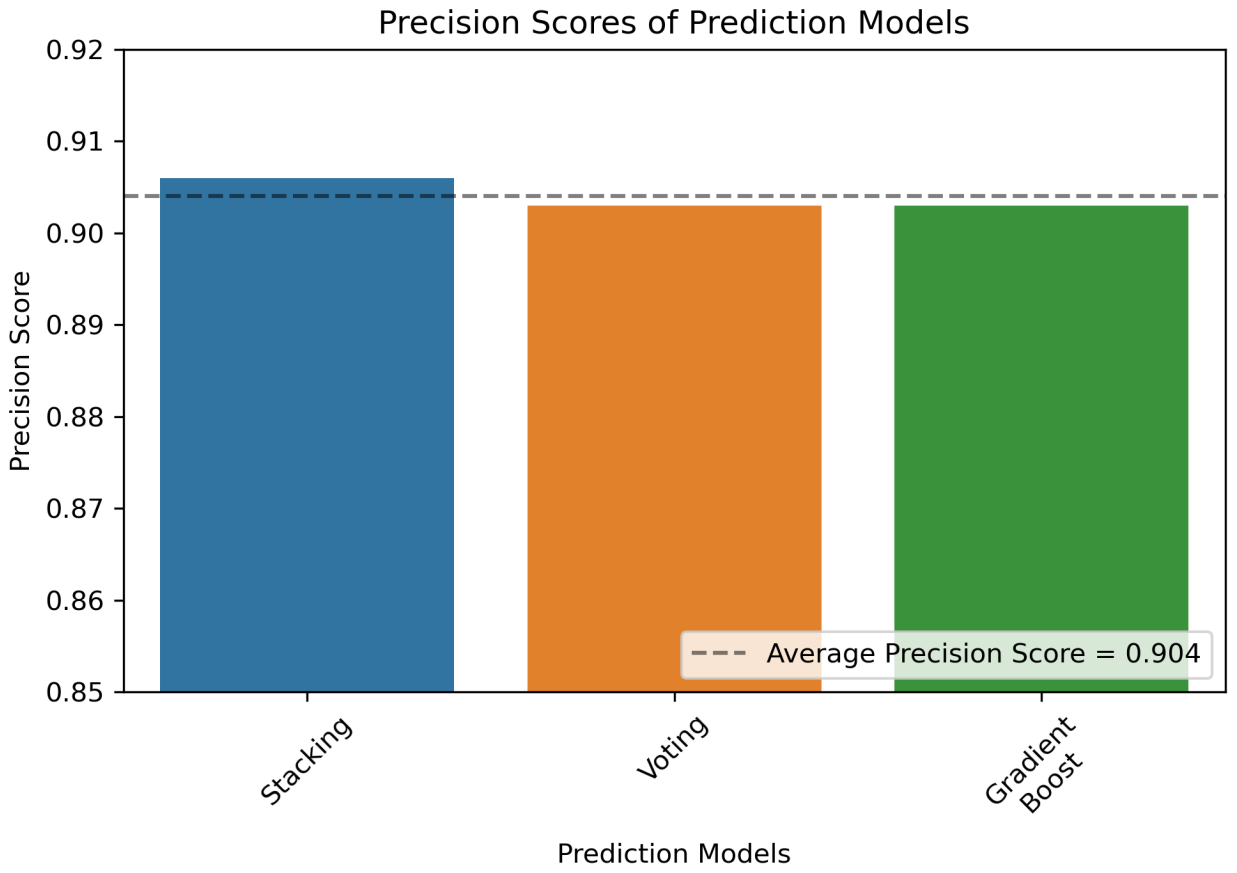
Better Models



Comparing Models



Comparing Models



Modeling Process – Step 4

- ***Stacking Classifier*** **is the** ***Final Model***
  - **90% Accuracy**
  - **91% Prediction**
    - **When the model predicts that customer will honor the reservation, it's correct 91% of the time**

- **Voting Classifier – Quickest / Second-Best**

Evaluation of the Final Model

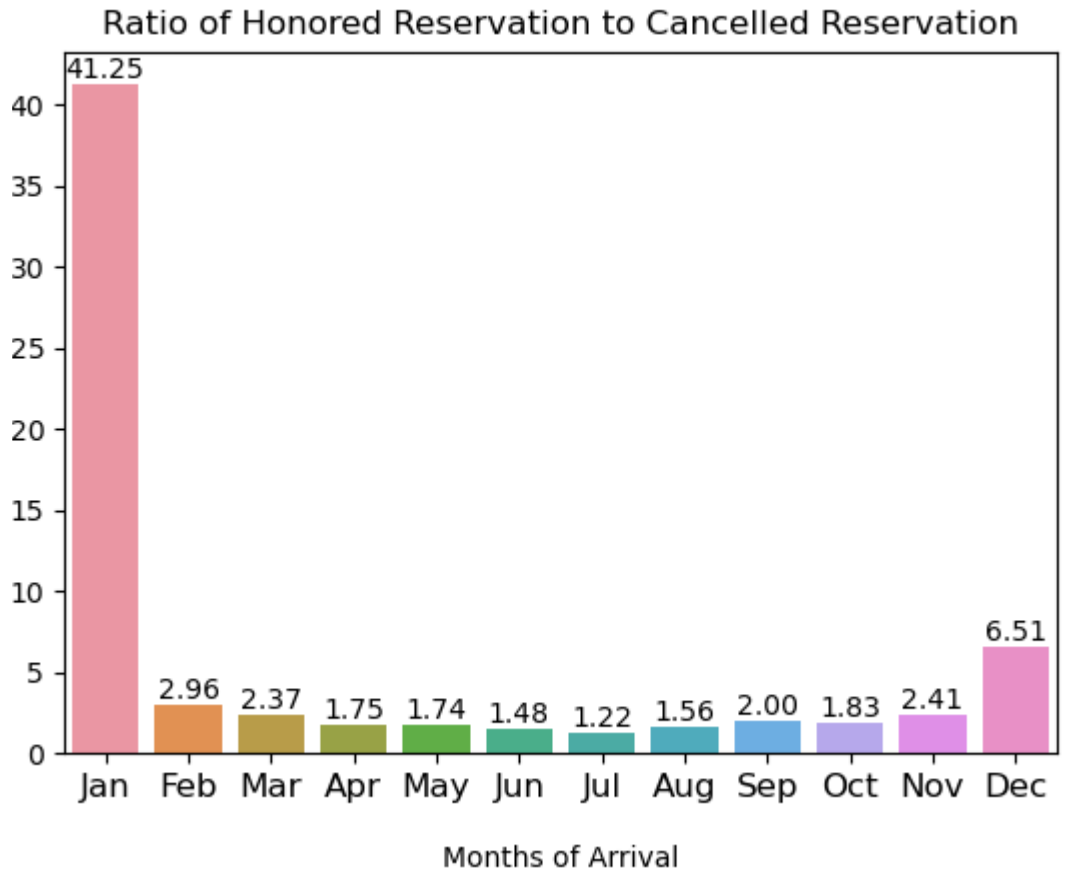
Most Important Features

- **Month of Arrival**
- **Reservation Segment Type**
- **Lead Time**
- **Average Price Per Night**
- **Number of Special Requests**

Least Important Features

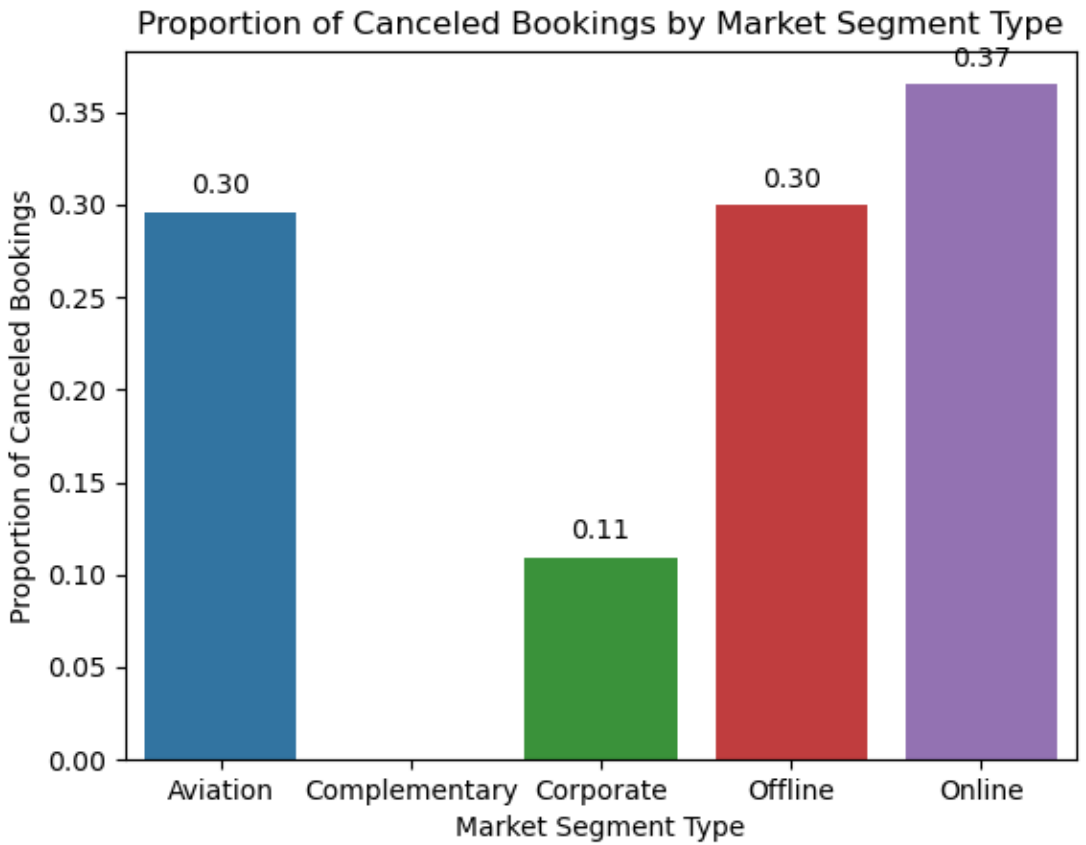
- **Meal Plan Types**
- **Number of Previous Cancellations**
- **Room Types**

Recommendation 1



Make strategies for dealing with Summer Reservations.

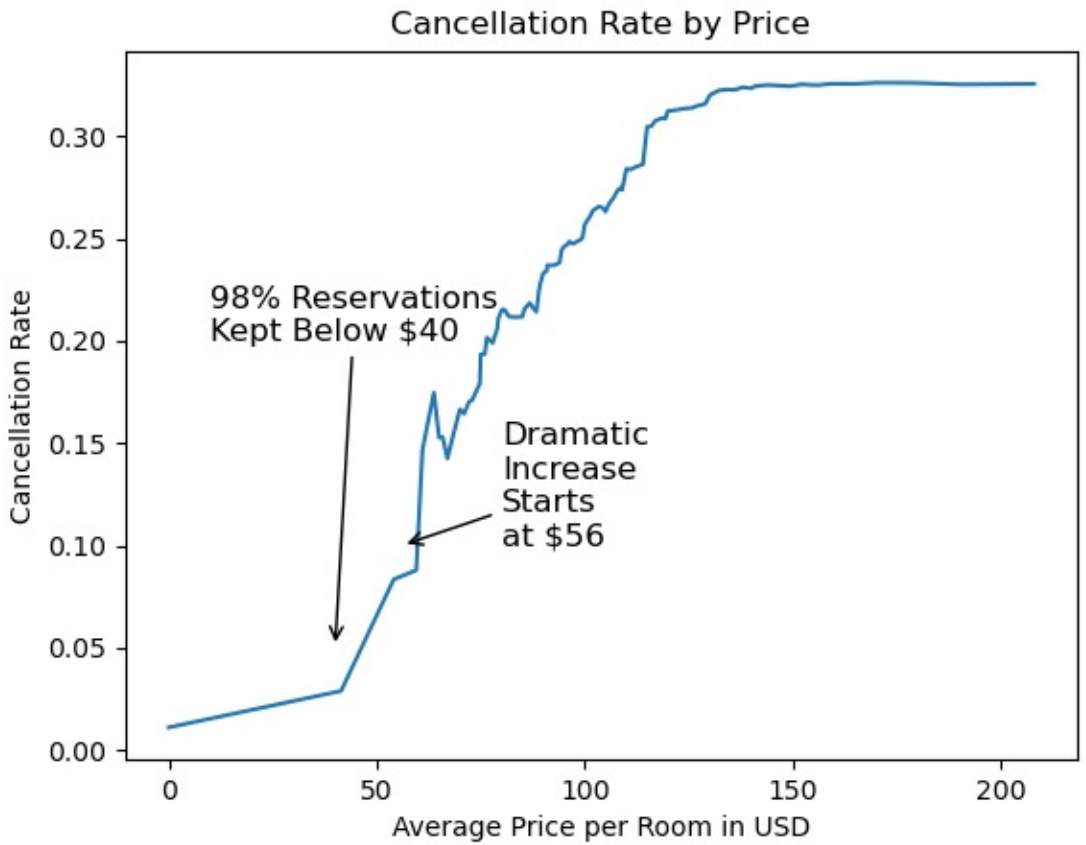
Recommendation 2



Make Business People Come.

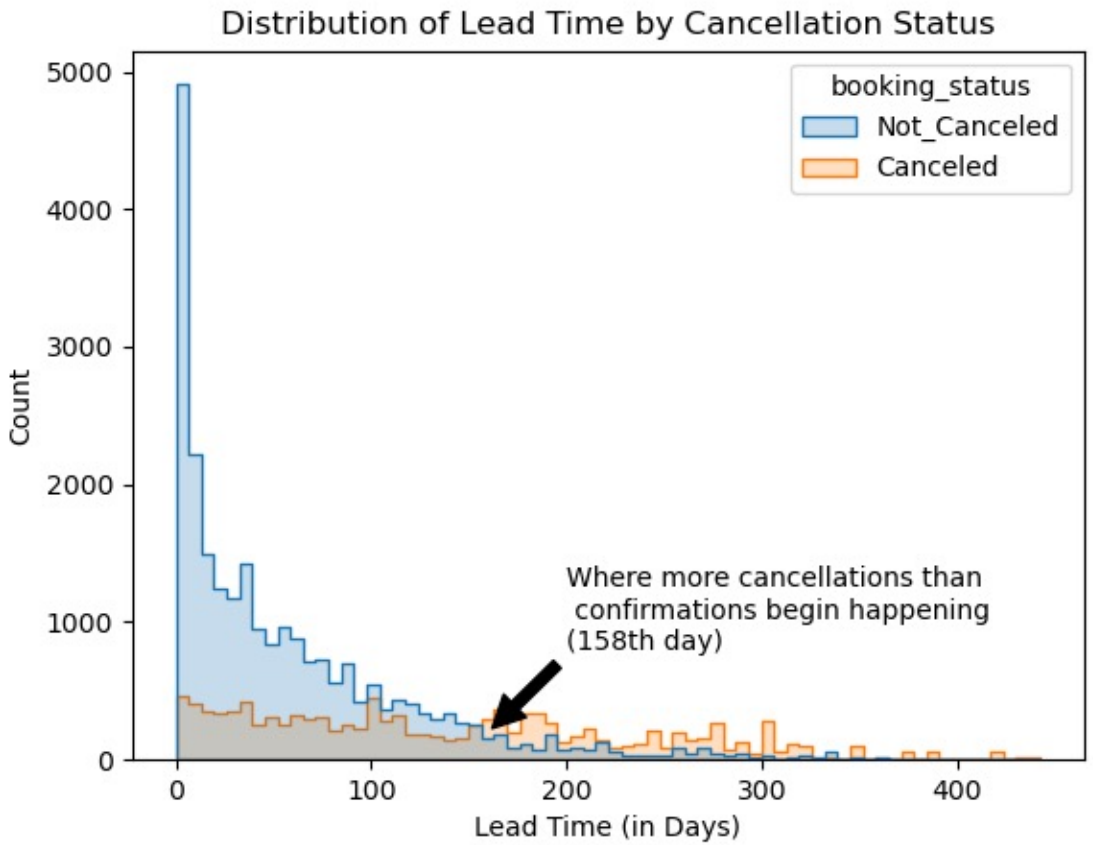
Enhance their experience.

Recommendation 3



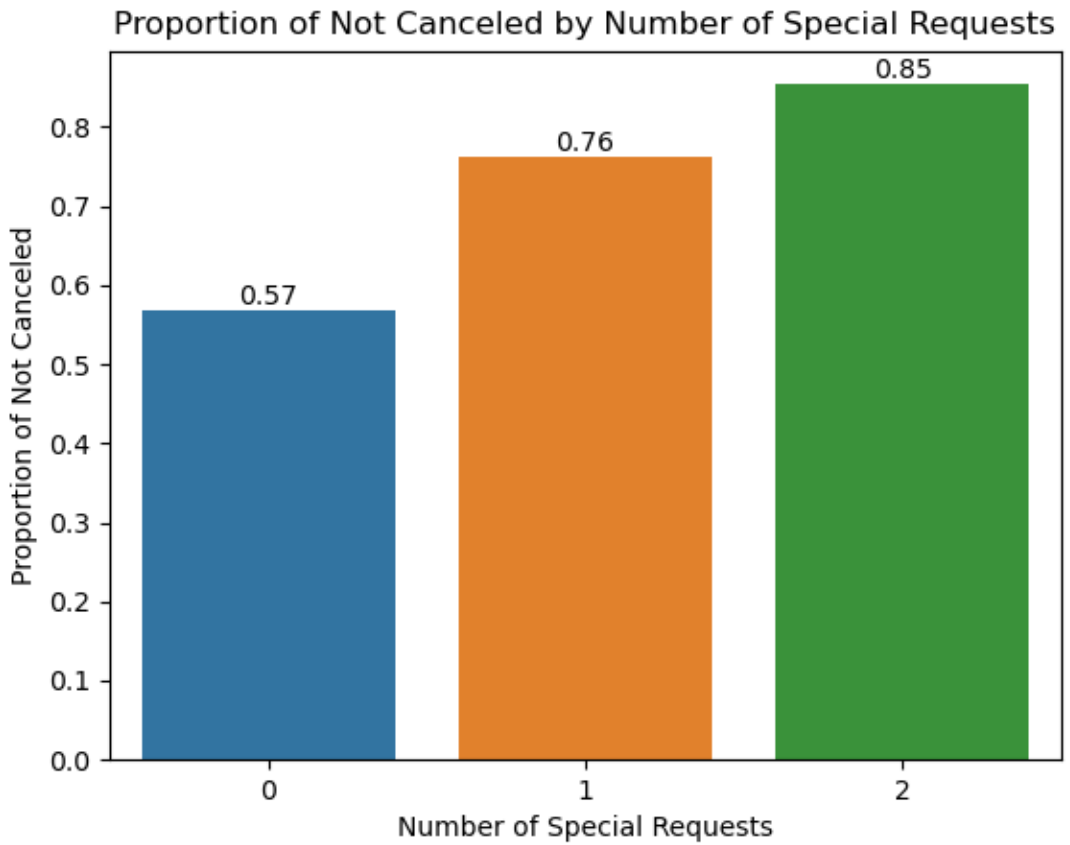
Price Your Rooms Just Before Steep Slope Increases.

Discovery 1



Reservation made 5 months prior are recipe for cancellation.

Discovery 2



More Special Requests are better.

Next Steps

For More Correctness

- Collect more personal information
  - Demographic
  - Loyalty Program
- Collect more geographic info.
  - Location at the time of making appointment
- Time it took for reservation. (for online reservations)

For More Usefulness

- Gather data from other hotels

Questions?

Thank you!

- **Please check out my other data science projects!**
- **Please click below for hyperlinks**