# Capstone

# Hotel Booking Prediction

Kristine Petrosyan
Flatiron School

# Problem Statement

- Reservations for resorts and hotels in Lisbon, Portugal are experiencing **37%** cancellations.

- Booking cancellations in the hospitality industry can result in:
  - Overbooking situations
  - Hotel's online social reputation
  - Revenue loss
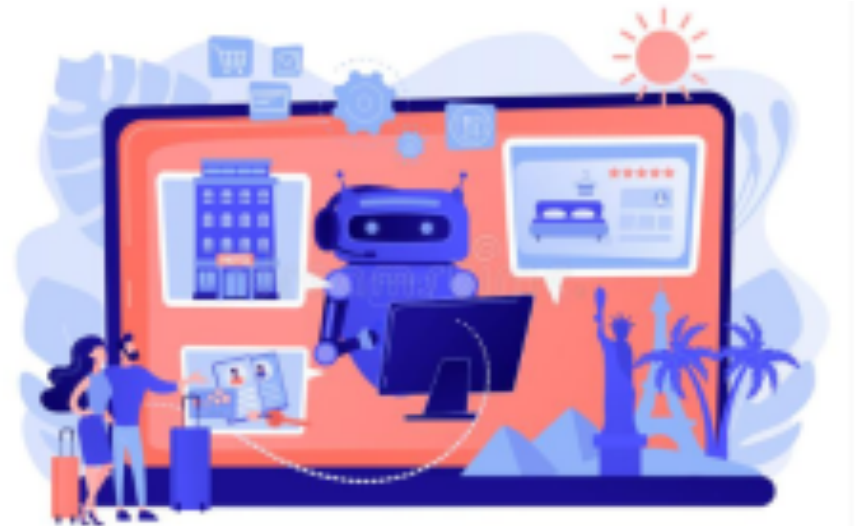  - Pricing, inventory and labor allocation decisions

# Business Value

- To predict hotel booking cancellations to decrease uncertainty and increase revenue.

- To explain how future cancelled reservations can be predicted in advance by machine learning methods.
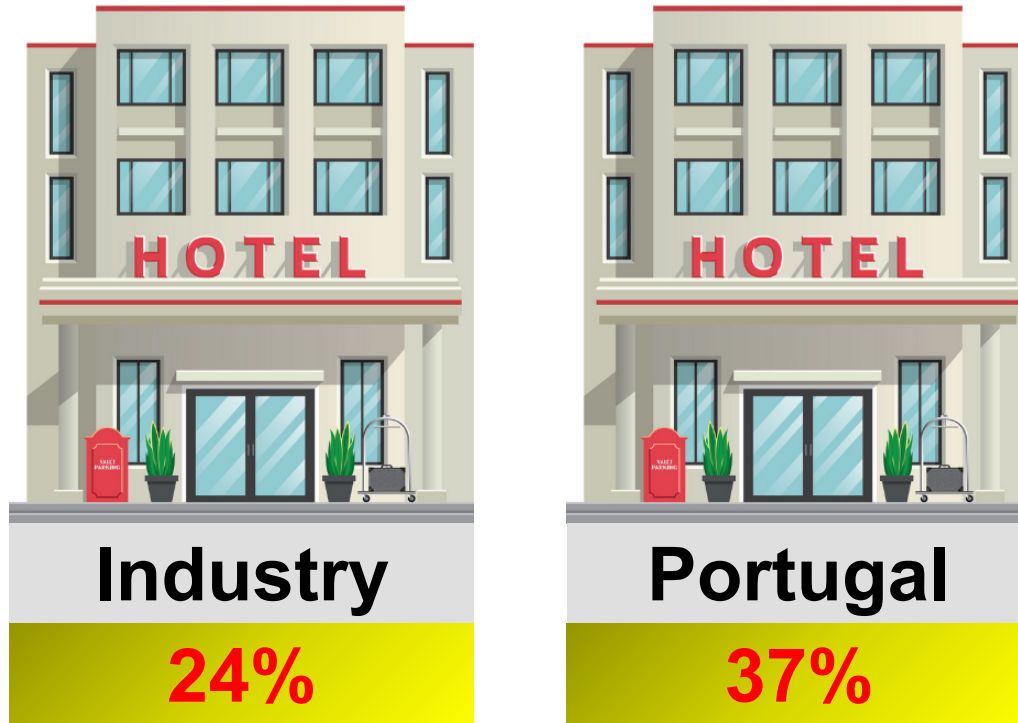
# **Methodology**

- Classification problem
- Data Science concepts used:
  - Machine Learning
  - Hypothesis Testing
  - Data Visualization
  - Time Series

# Industry Cancellation Rates



**Industry**

**24%**

**Portugal**

**37%**

- Cancellation rate is **13%** higher than the Industry in Portugal

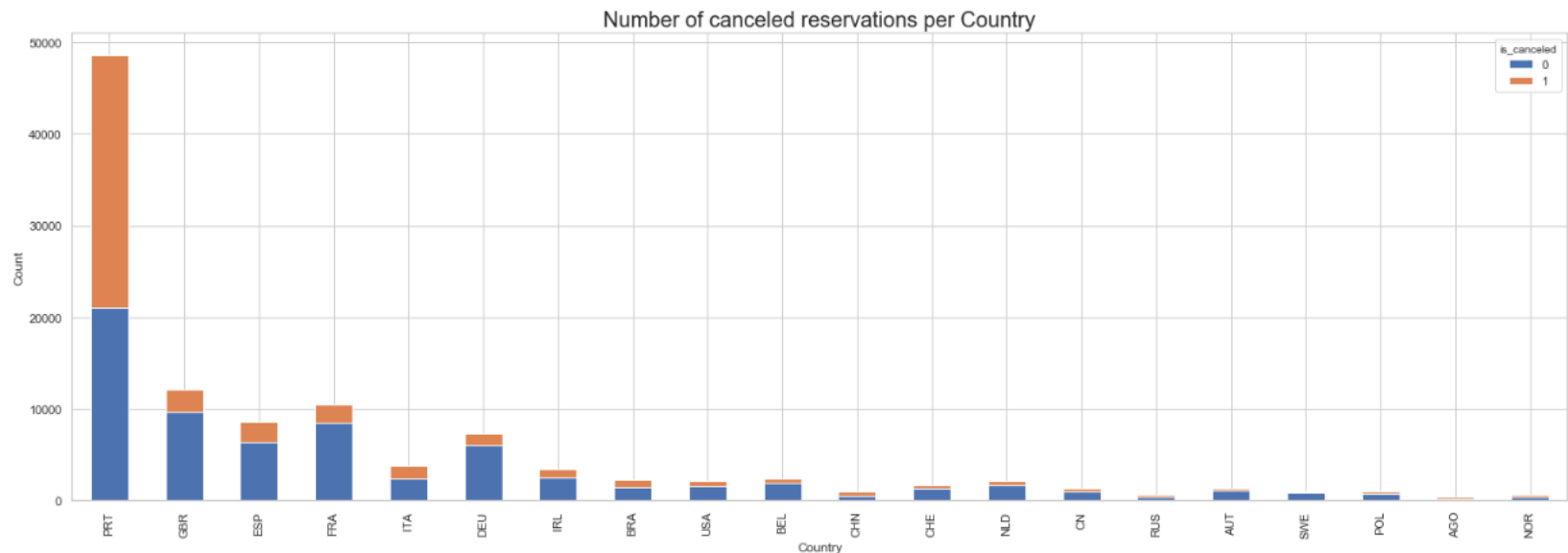# Loss Revenue

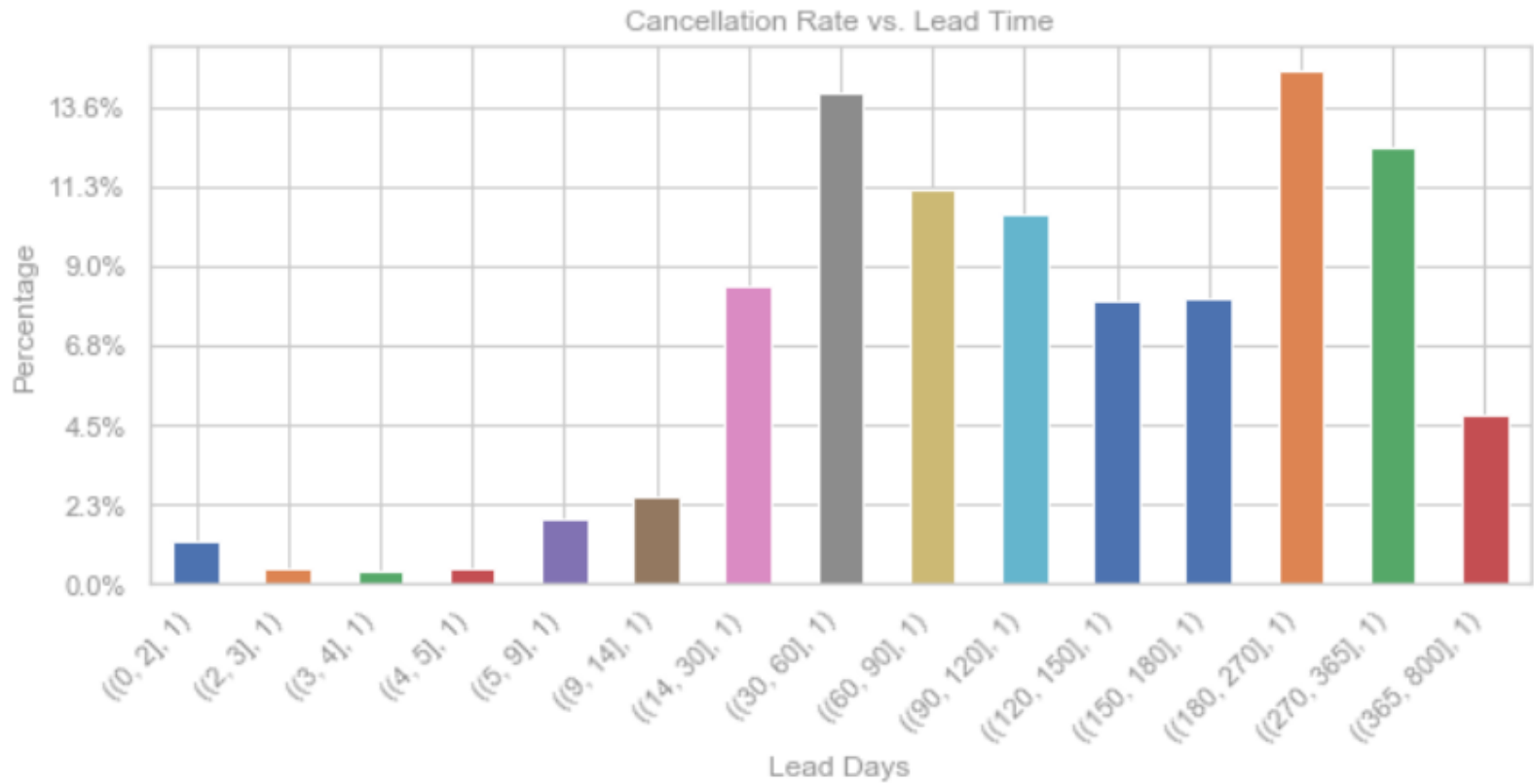- Average lost from cancellations per year

# Country

- Most of the cancellations are occurring in Portugal with 27,519

  Which is ten times more than the following:
  - GBR 2,453
  - ESP 2,177
  - FRA 1,934



Number of canceled reservations per Country

# Long Lead Times



Cancellation Rate vs. Lead Time

**85%** of cancellations occur past 30 days

# Findings - Most Predictive Features



- First predictor was '**Lead Time**'
- Others:
  - 'country',
  - 'market segment',
  - 'price',
  - 'customer type',
  - 'specific dates of reservation'

# Recommendations – Best Model for Predictions

- Random Forest Classification have the "highest" scoring metrics

| SCORING METRIC | VALUE |
|---|---|
| Test Accuracy Score | 88% |
| Precision | 87% |
| Test Error Rate | 12% |
| ROC/AUC | 87% |
| Sensitivity (Recall) | 81% |
| Specificity | 92% |

# Recommendations



- Hotels should analyze the growth rate of their respective marketing channels
- Allow hotel managers to act on bookings with high cancellation probability
- Improve overbooking and cancellation policies

# Future Work



- Perform analysis between the city hotel and resort separately.
- Build a neural network to increase accuracy from 88% to 90%.
- Create a GUI to convert high cancellation probabilities into revenue.
- Perform analysis to determine which OTA will serves hotel better.
- Analyze data from weather, holidays and online prices/inventory.

# Thank you!

*Contact: kristinelpetrosyan@gmail.com*