

---



# Hotel Analysis

Predicting booking cancellations

By Karina Basto Eyzaguirre

---



---

# TABLE OF CONTENTS

- 01• Business Understanding
  - 02• Data Understanding
  - 03• Data Preparation
  - 04• Modelling
  - 05• Evaluation
  - 06• Recommendations
  - 07• Next Steps
-



# Business

## •01• Understanding

*“Increased revenue loss due to  
last minute hotel booking  
cancellations”*







Data

•02• Understanding

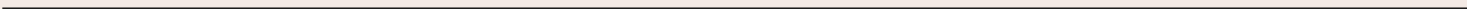
---

---

# Hotel Booking Cancellation Prediction Dataset



36,283 unique lines of information and 17 feature columns





---

# Data

## •03• Preparation

A scenic view of a coastal town, likely in the Canary Islands, featuring white buildings with red-tiled roofs, numerous palm trees, and a harbor with several boats. The town is built on a hillside overlooking a large body of water, with steep cliffs visible in the background. The text "Data" and "•03• Preparation" is overlaid on the image in a white serif font.

---

**·01·**

Data  
Exploration  
and Cleaning

**·02·**

Feature  
Engineering

**·03·**

Feature  
Selection

**·04·**

Data  
Transformation

---





---

# •04• Modelling

---



---

·1·

·Baseline Logistic  
Regression·

·2·

·Modified Logistic  
Regression·

statistically significant  
features

·3·

·Modified Logistic  
Regression·

Used SMOTE and Class  
Weights

---

---

·4·

·Baseline Decision  
Tree·

·5·

·Hyperparameter  
Tuning on Decision  
Tree·

Parameters: max\_depth,  
min\_samples\_split,  
min\_samples\_leaf,  
class\_weight, criterion

·6·

·Understanding  
Feature  
importance·

---



---

# •05•Evaluation



## Final Model: Tuned Decision Tree

---

max\_depth = 13  
min\_samples\_split = 2  
min\_samples\_leaf = 2  
class\_weight = None  
criterion = gini



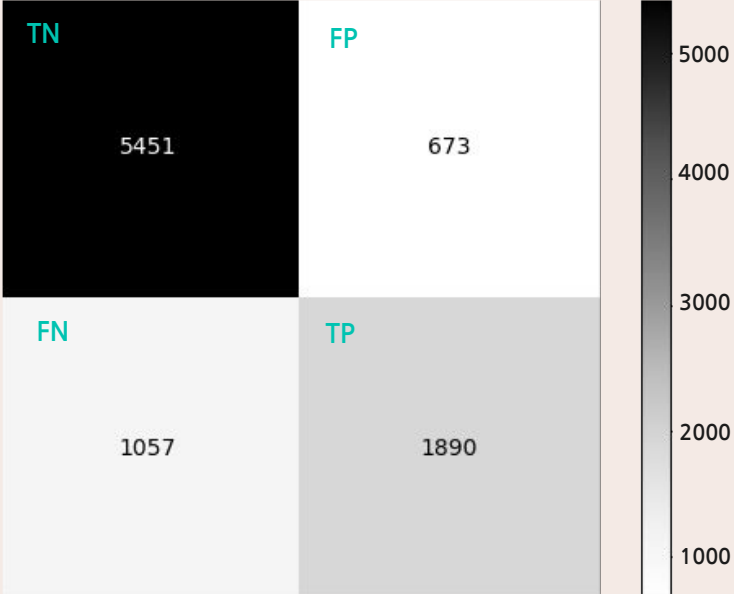


# Evaluation Metrics

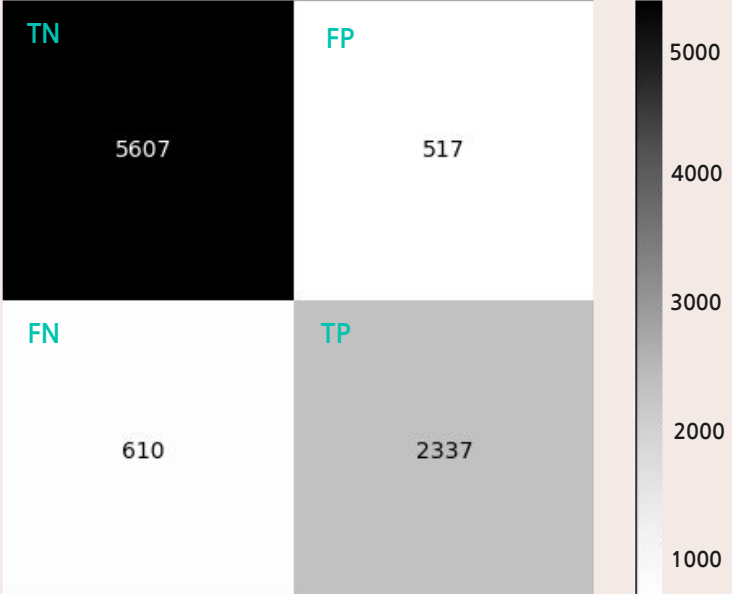
	Baseline Model	Final Model	Delta
Accuracy	81 %	88%	+7%
Precision	74%	82%	+8%
Recall	64%	80%	+16%
F1 score	69%	81%	+12%
AUC	87%	92%	+5%

# Confusion Matrix

Baseline Model

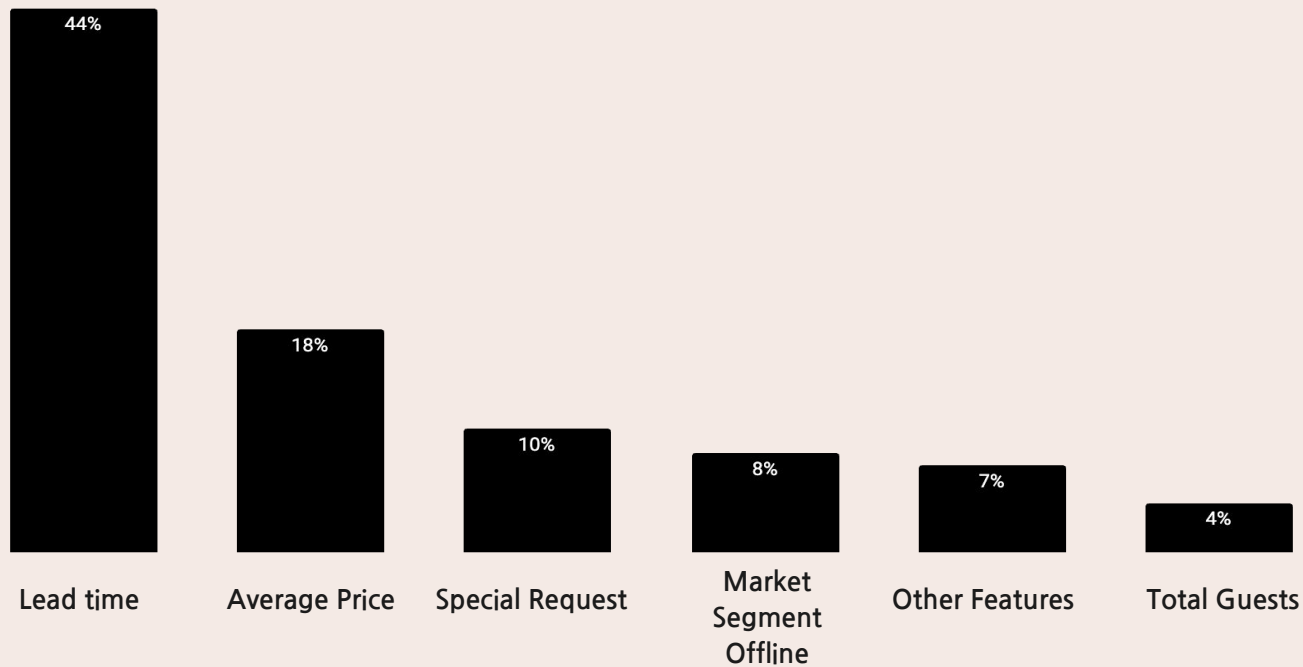


Final Model





# Feature Importance





---

# •06•Recommendations

---





**•01•**

Overselling strategy

**•02•**

features impacting cancellations

**•03•**

“Check with Guest” Program



# •07• Next Steps





Implement  
Recommendations



Investigate top five  
features



Hotel  
Datasets



Re-evaluate and  
experiment with model



# Thank you!

Author: Karina Basto-Eyzaguirre

 [karinabastoe@gmail.com](mailto:karinabastoe@gmail.com)