



Predicting Cancellations

Beatriz Pereira
Mariana Domingues
Ana Luísa Mestre
Nadine Aldesouky



BUSINESS OBJECTIVE

Forecast net demand by predicting
future cancellations



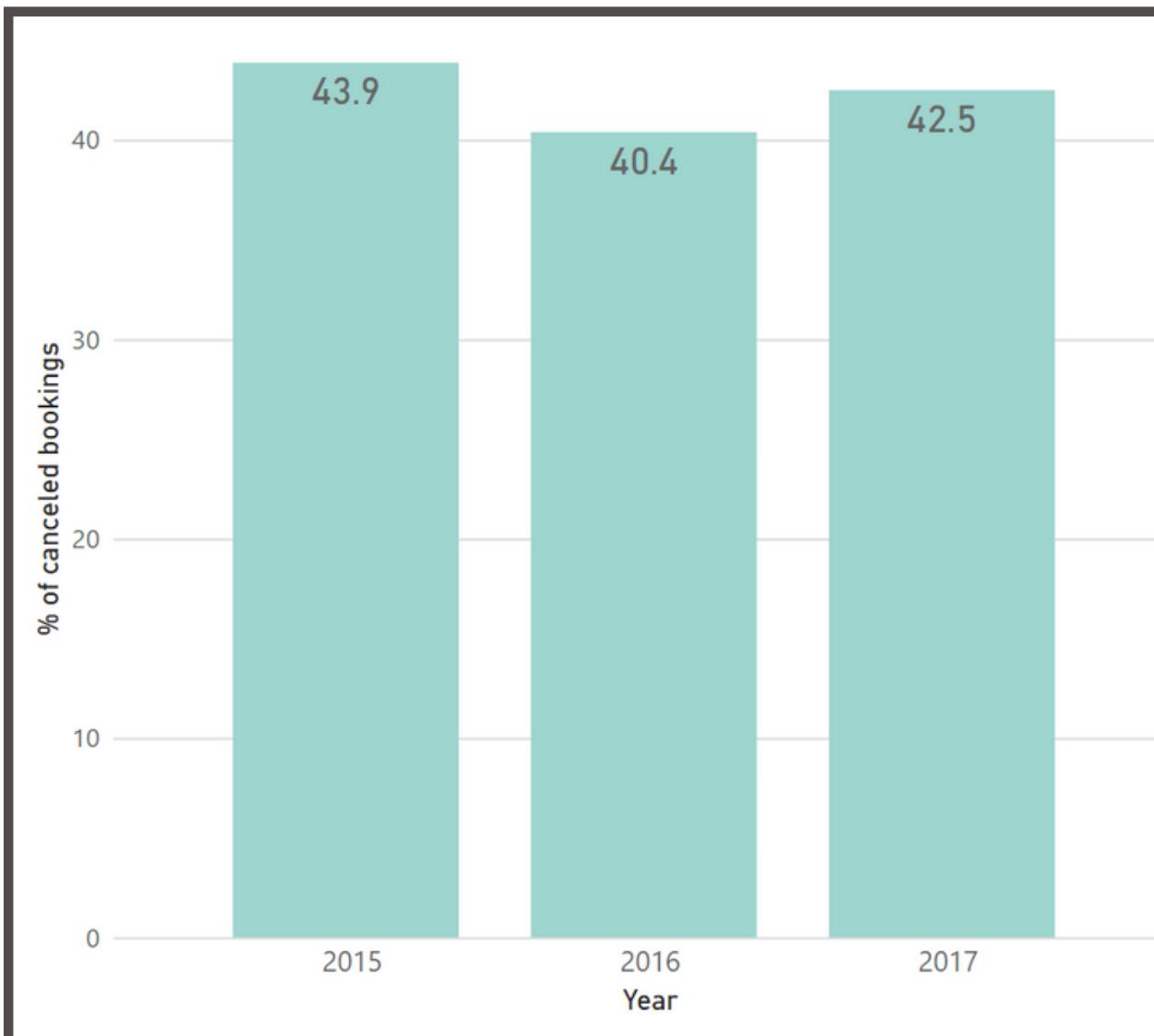
ASPECTS OF DATASET

- From Lisbon city hotel H2
 - From July 2015 to August 2017
 - 79330 bookings
 - 31 features
-

KNOWING YOUR DATA

- 33076 cancellations (42% of total bookings)

Cancellations per Year



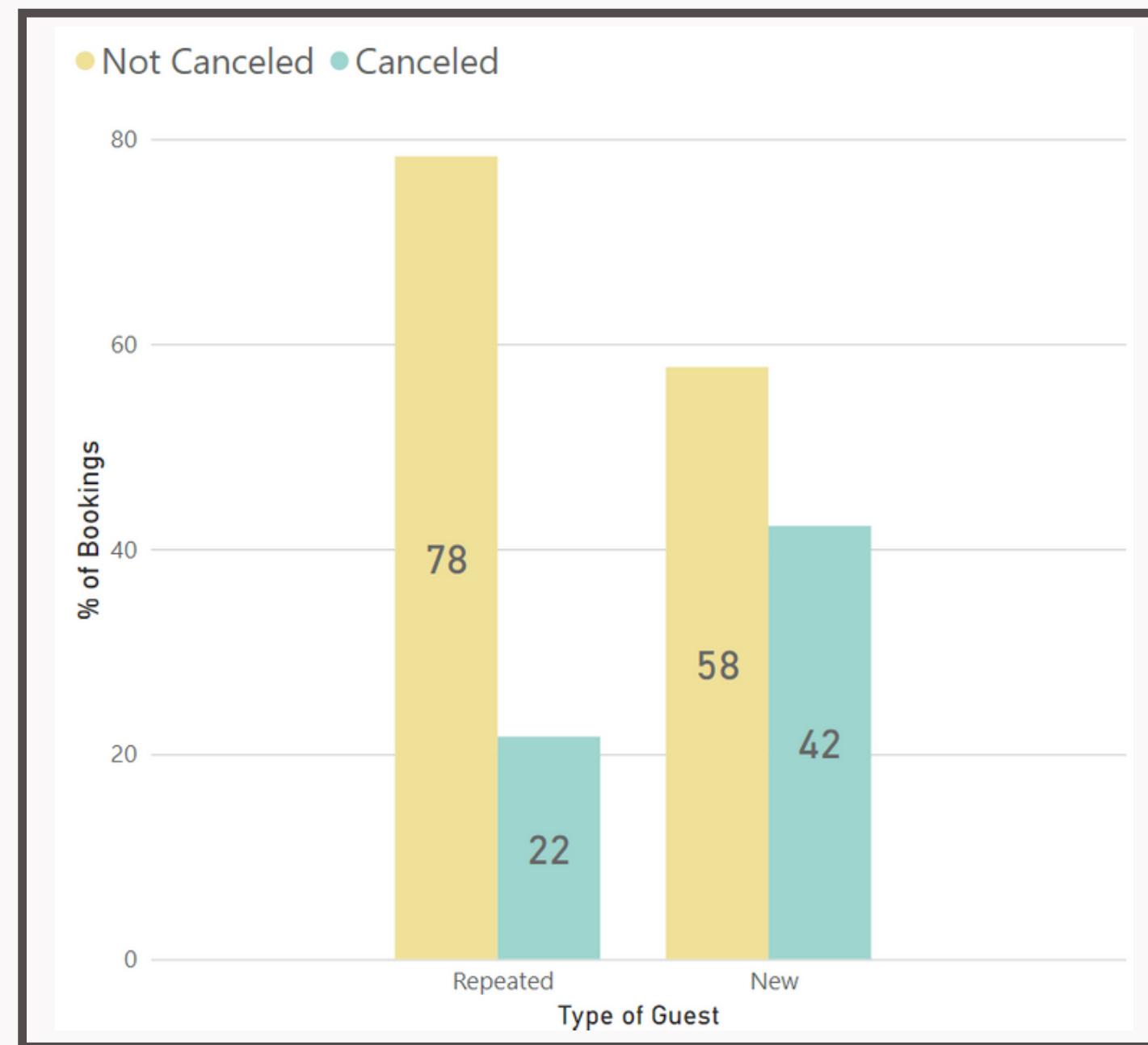
Cancellations per Month



KNOWING YOUR DATA

- Repeated Guests tend to like the hotel while New Guests have a higher risk of cancelling

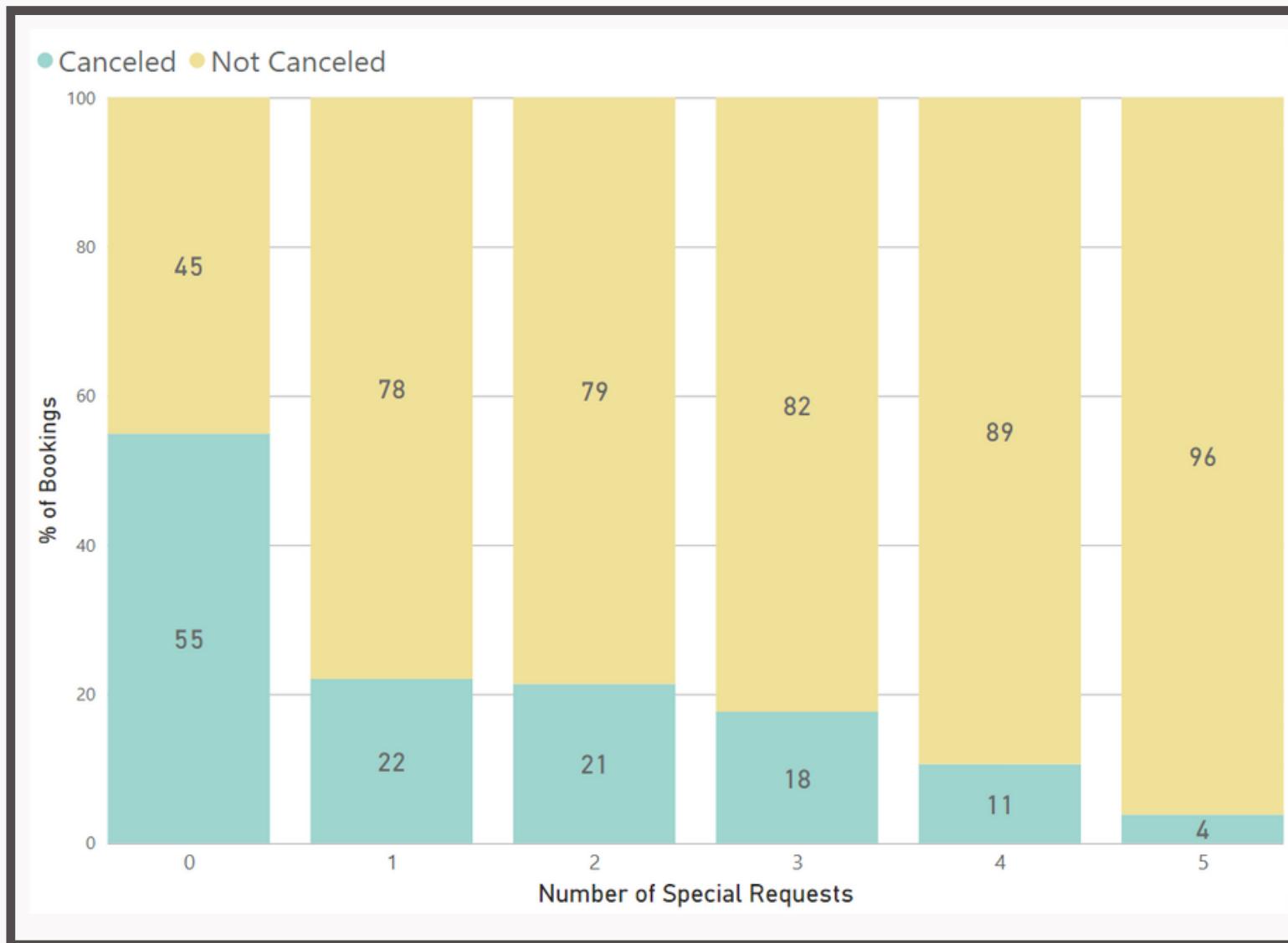
Repeated vs New Guests



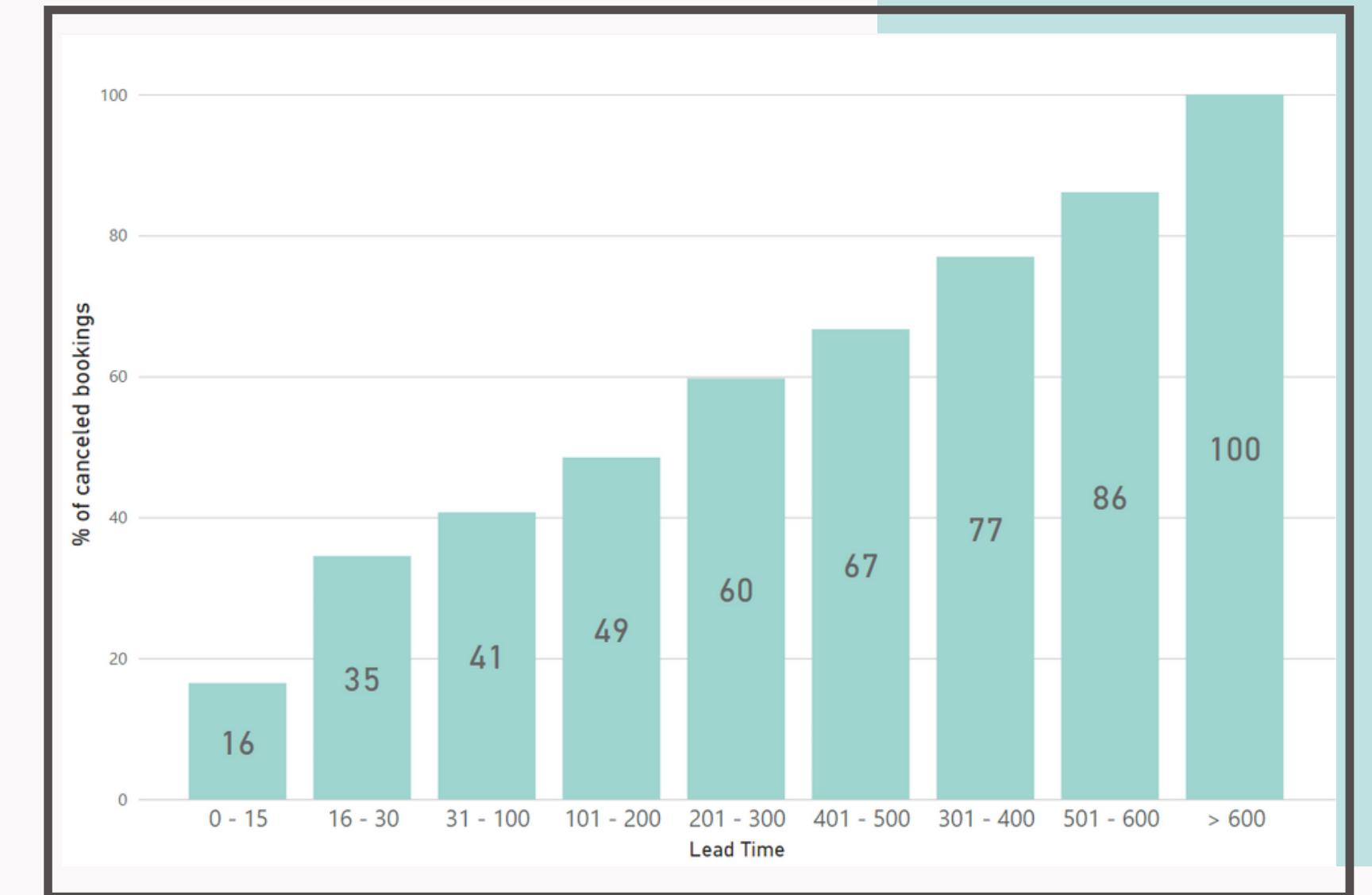
KNOWING YOUR DATA

- Lead time and total number of special requests have a large influence on cancelling a booking

Special Requests



Lead Time



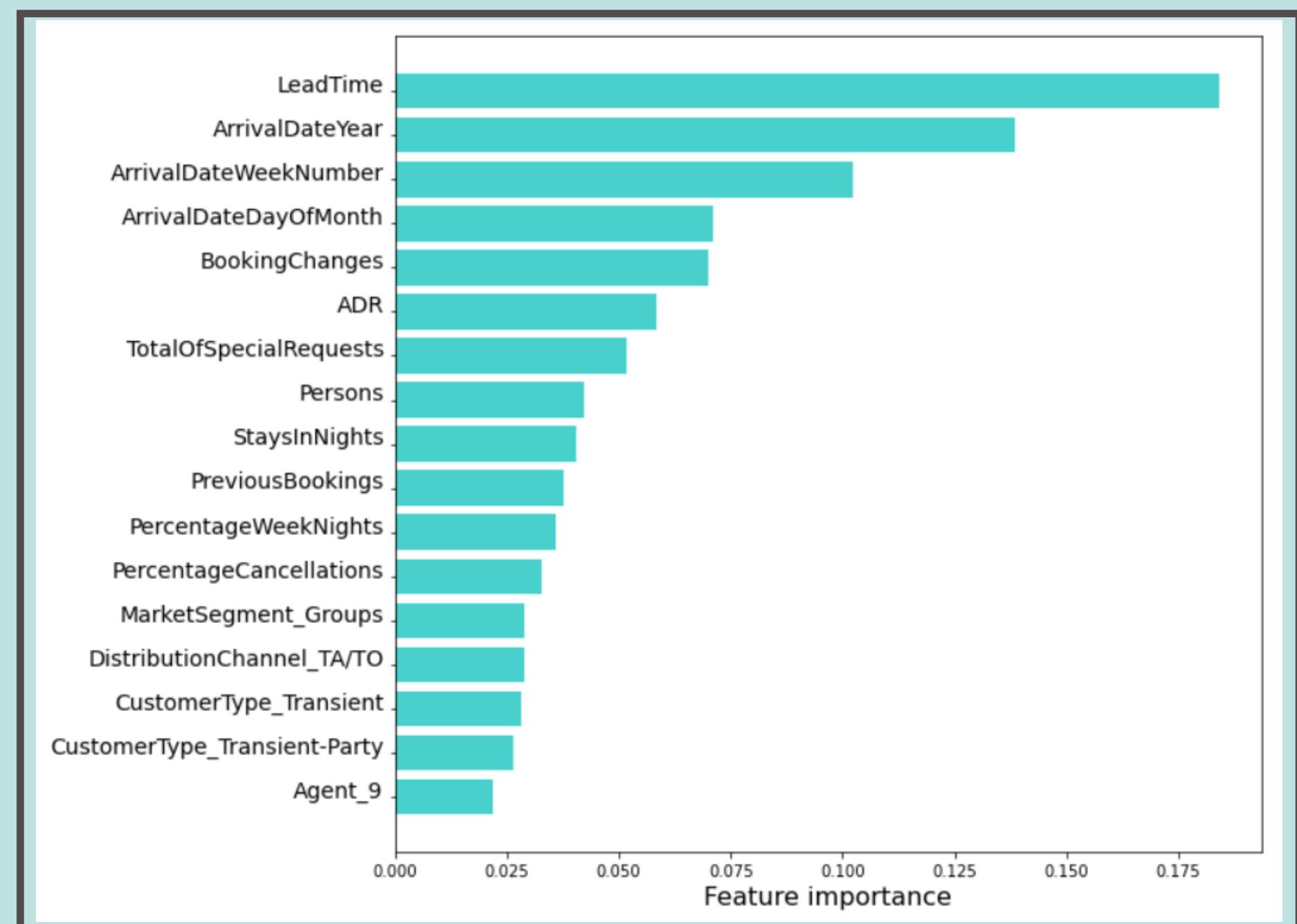
MODEL INPUT

Feature Selection

- New features created
- Recursive Feature Elimination



17 features



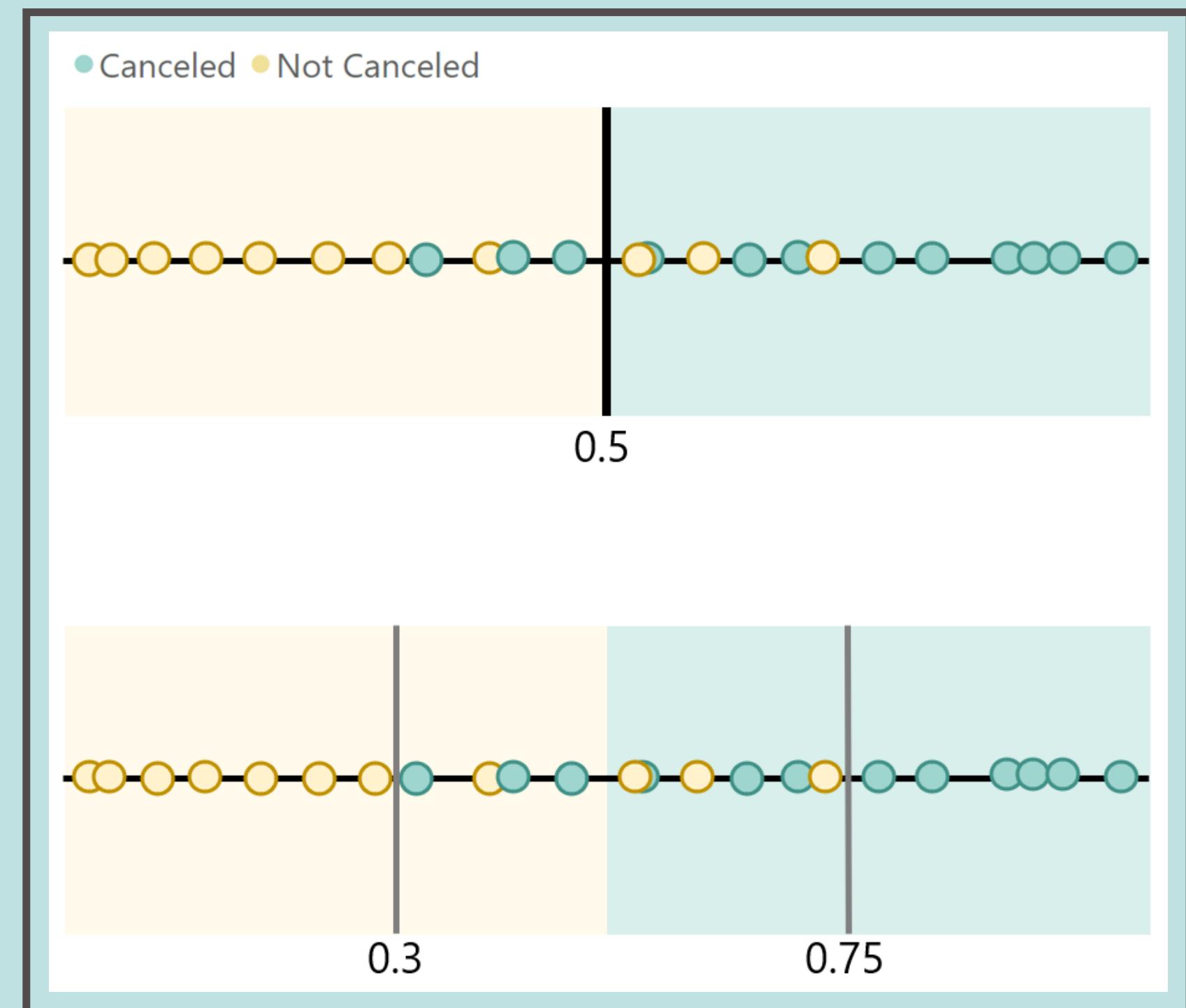
MODEL OUTPUT

Result

- The likelihood of cancelling (0 to 1)
- Create 2 classes using a threshold of 0.5

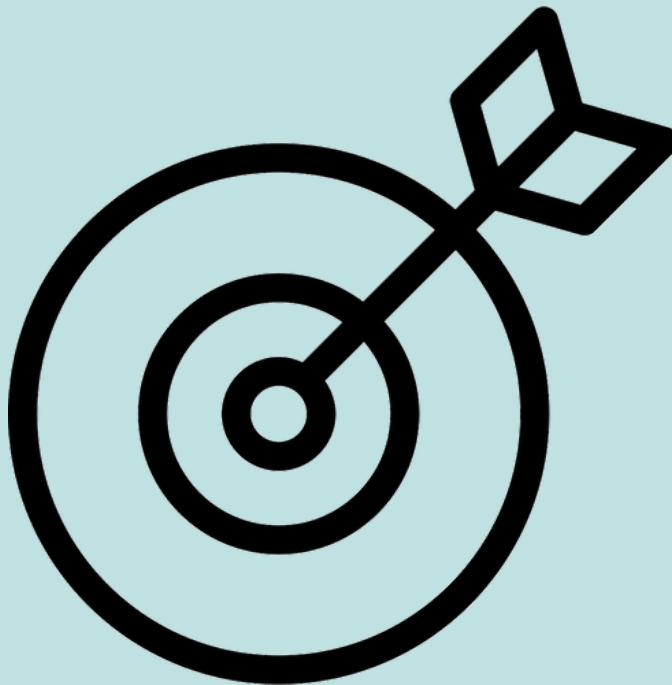
Suggestion for Application

- Create >2 classes using intervals to group cancellations
- Apply different policies per class



MODEL PERFORMANCE

Accuracy Score



86%

Error

		Predicted	
		True Non-cancellations	False Cancellations
Actual	True Non-cancellations	86%	14%
	False Non-cancellations	9.9%	4.1%

FNC: 9.9%

FC: 4.1%



Thank you!
