DATA MINING PRESENTATION

Hawk Hotel Consultants

AGENDA

01. BACKGROUND INFORMATION

02. OUR GOALS

03. SOLUTION & RECOMMENDATIONS

04. LIMITATIONS

01.

BACKGROUND INFORMATION

PROBLEM DESCRIPTION

- Industry average was 12.8% in 2018
- Find strategies to keep hotel fully booked
- Target high cancellations & build customer loyalty

(Raza, 2023 & Revenue Hub, 2020)

BACKGROUND INFORMATION | OUR GOAL | SOLUTION | LIMITATIONS

02.

OUR GOALS

BUSINESS GOAL

- Keep hotel rooms fully booked to maximize profits
- Maintain customer satisfaction

BUSINESS GOAL

- Keep hotel rooms fully booked to maximize profits
- Maintain customer satisfaction

DATA MINING GOAL

- Determine most important factors that lead to cancellations
- Build model that predicts cancellations

03.

SOLUTIONS & RECOMMENDATIONS

SOLUTION #1

Training

Model	AUC	CA	F1	Precision	Recall	MCC
Neural Network	0.885	0.825	0.821	0.822	0.825	0.593
Decision Tree	0.943	0.877	0.876	0.875	0.877	0.718
Random Forest	0.911	0.857	0.853	0.856	0.857	0.668
Gradient Boosting	0.893	0.836	0.826	0.840	0.836	0.616

Testing

Model	AUC	CA	F1	Precision	Recall	MCC
Neural Network	0.878	0.819	0.815	0.815	0.819	0.573
Decision Tree	0.916	0.851	0.850	0.850	0.851	0.655
Random Forest	0.903	0.855	0.851	0.853	0.855	0.658
Gradient Boosting	0.888	0.833	0.823	0.835	0.833	0.602

Cross Validation (10 Fold)

Model	AUC	CA	F1	Precision	Recall	MCC
Neural Network	0.883	0.822	0.819	0.819	0.822	0.587
Decision Tree	0.920	0.858	0.857	0.856	0.858	0.675
Random Forest	0.906	0.852	0.847	0.852	0.852	0.656
Gradient Boosting	0.890	0.834	0.824	0.838	0.834	0.612

- Predictive model – classifies as “cancel” or “not cancel”
- Decision tree provides the highest AUC and does not overfit

SOLUTION #1

Decision Tree Model Factors

- Number of adults
- Number of children
- Number of weekend nights
- Number of week nights
- Meal plan
- Parking space
- Room type reserved
- Lead time
- Arrival month
- Market segment type
- Repeat guest
- # of previous cancellations
- # of bookings not canceled
- Average price per room
- Special requests

SOLUTION #1

Decision Tree Model Factors

- Number of adults
- Number of children
- Number of weekend nights
- Number of week nights
- Meal plan
- Parking space
- Room type reserved
- Lead time
- Arrival month
- Market segment type
- Repeat guest
- # of previous cancellations
- # of bookings not canceled
- Average price per room
- Special requests

Result

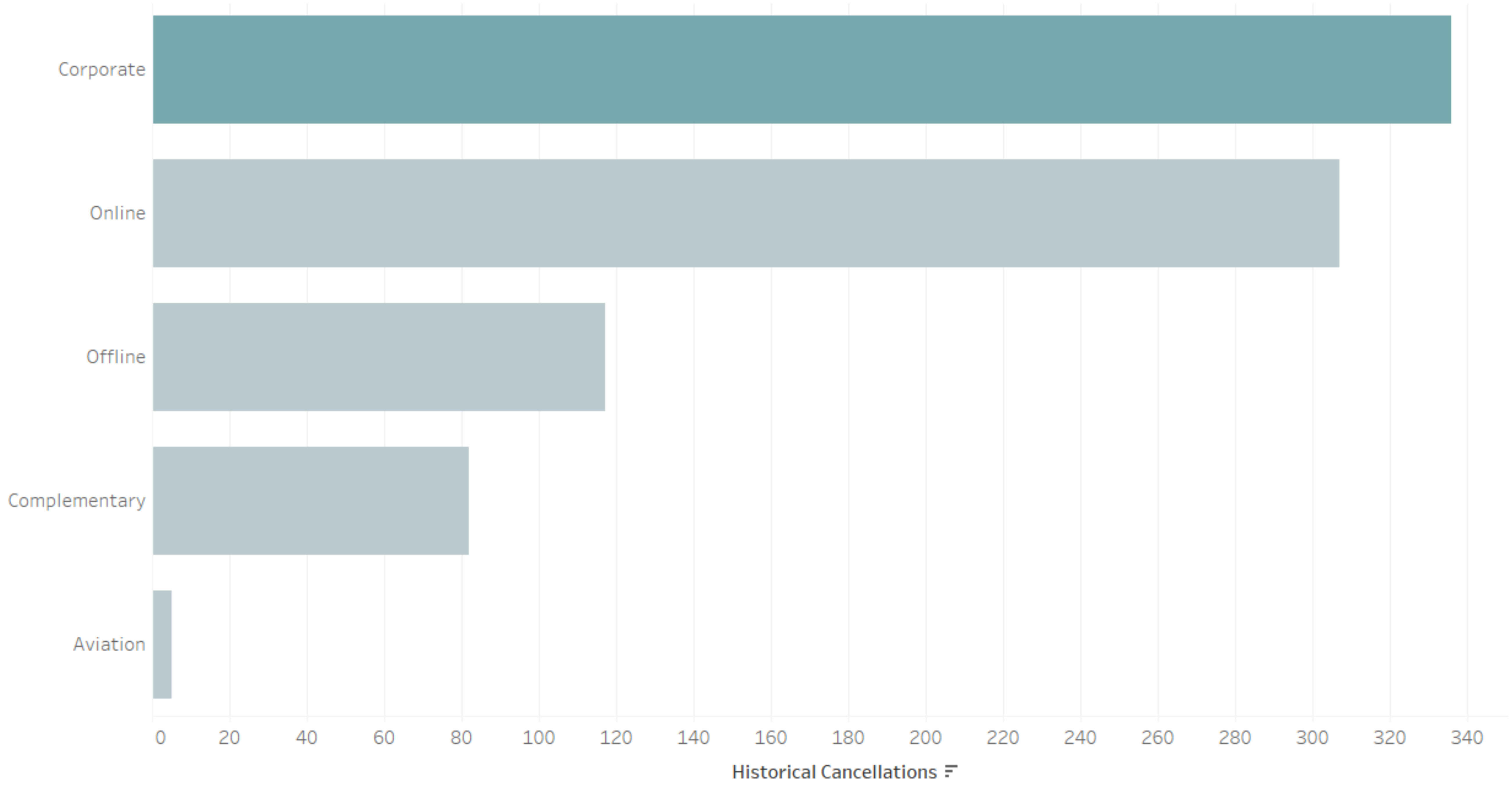
- Likely to cancel
- Unlikely to cancel

RECOMMENDATION #1

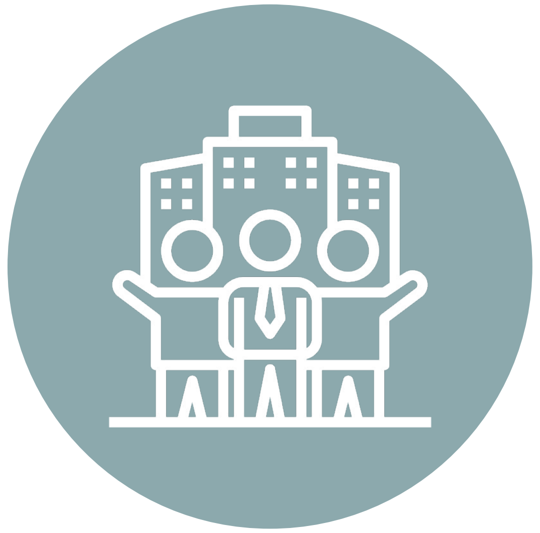
Implement this predictive model onto reservation website to determine the probability of reservation cancellation.

SOLUTION #2

Cancellations have come mostly from corporate groups.



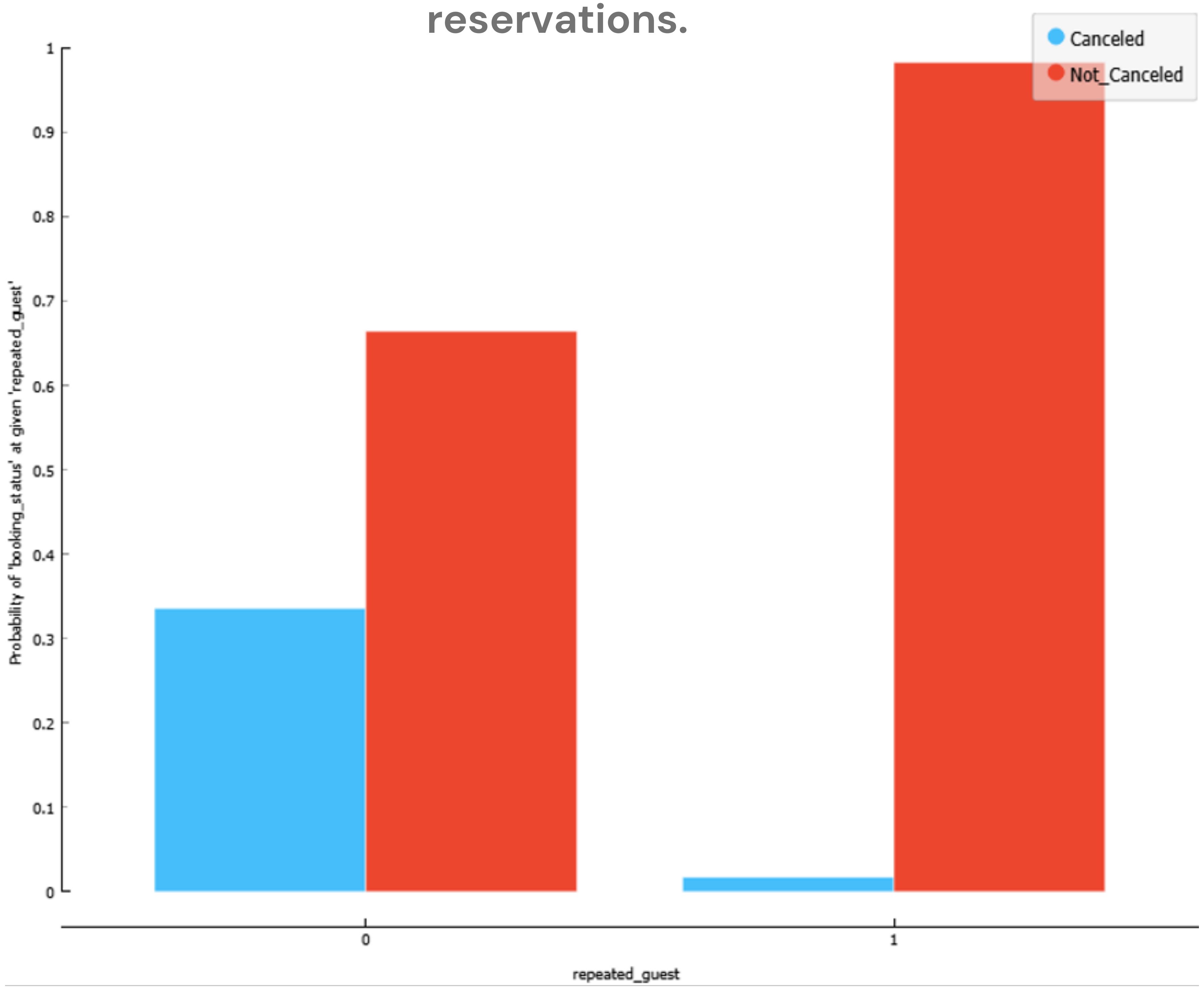
RECOMMENDATION #2



Target corporate groups by offering incentives and personalized services, or raise cancellation fees to decrease chances of cancellation.

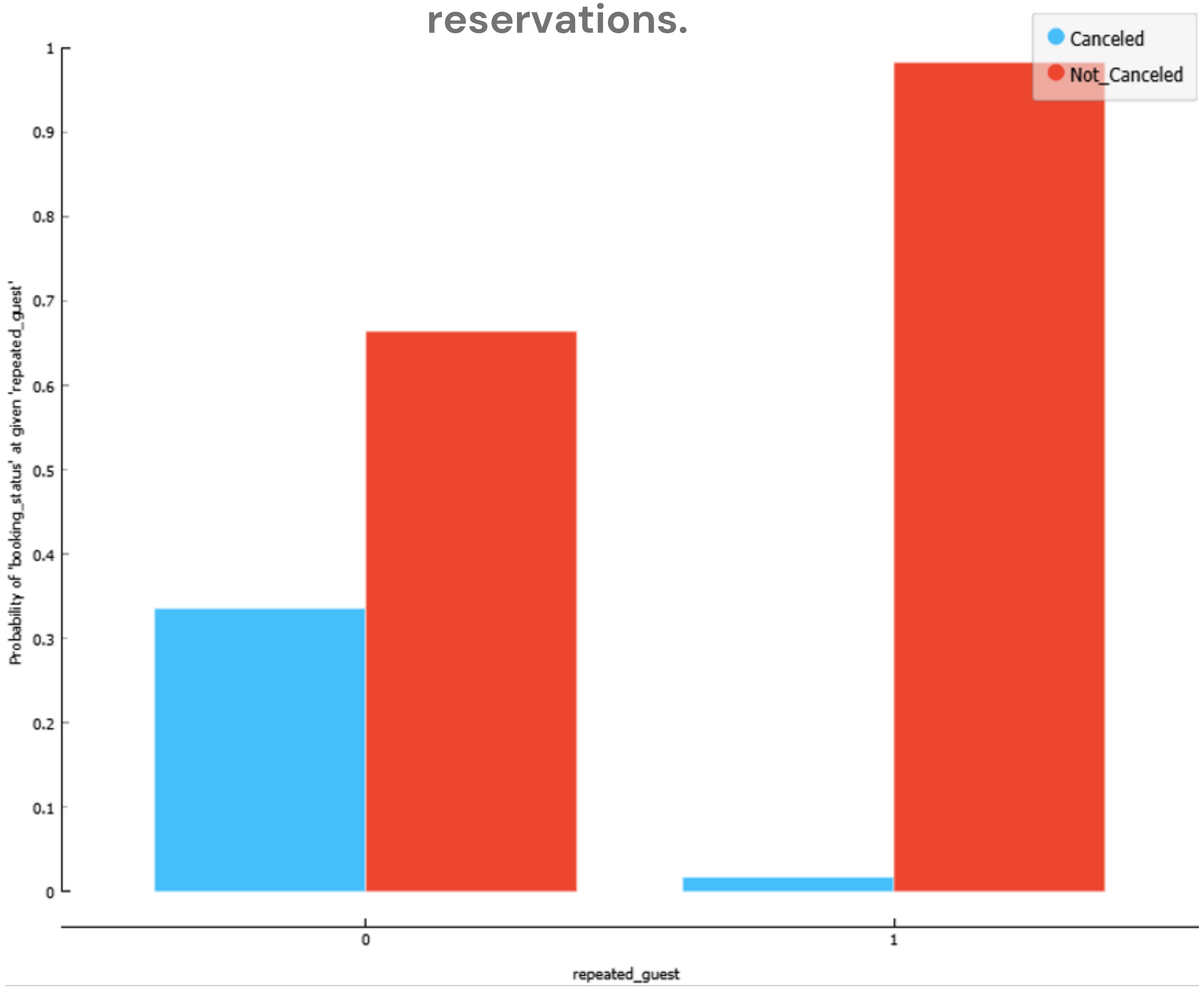
SOLUTION #3

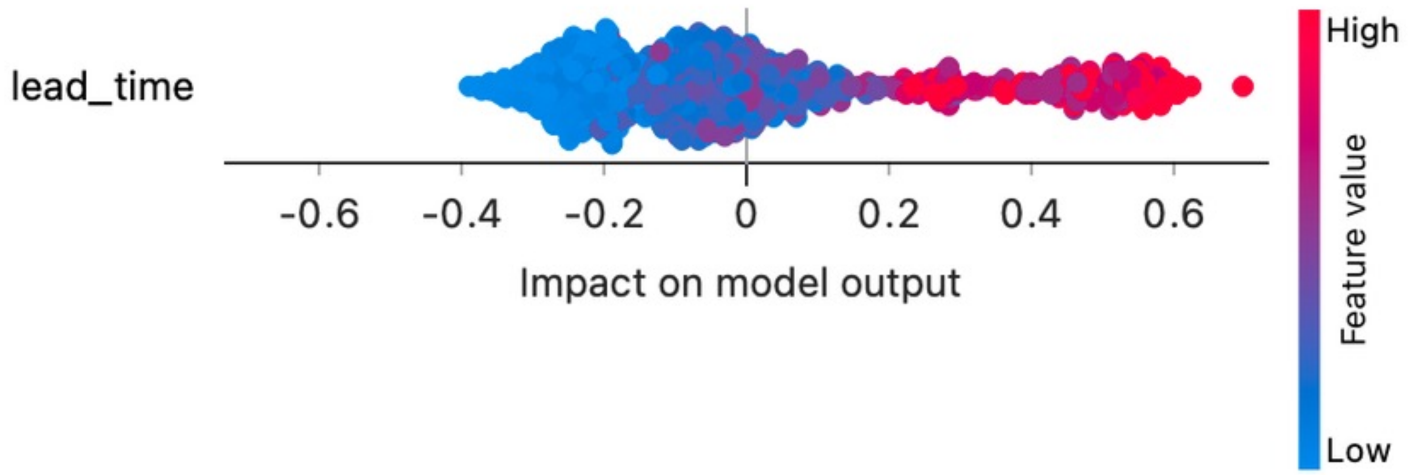
Repeat guests are less likely to cancel their hotel reservations.



SOLUTION #3

Repeat guests are less likely to cancel their hotel reservations.





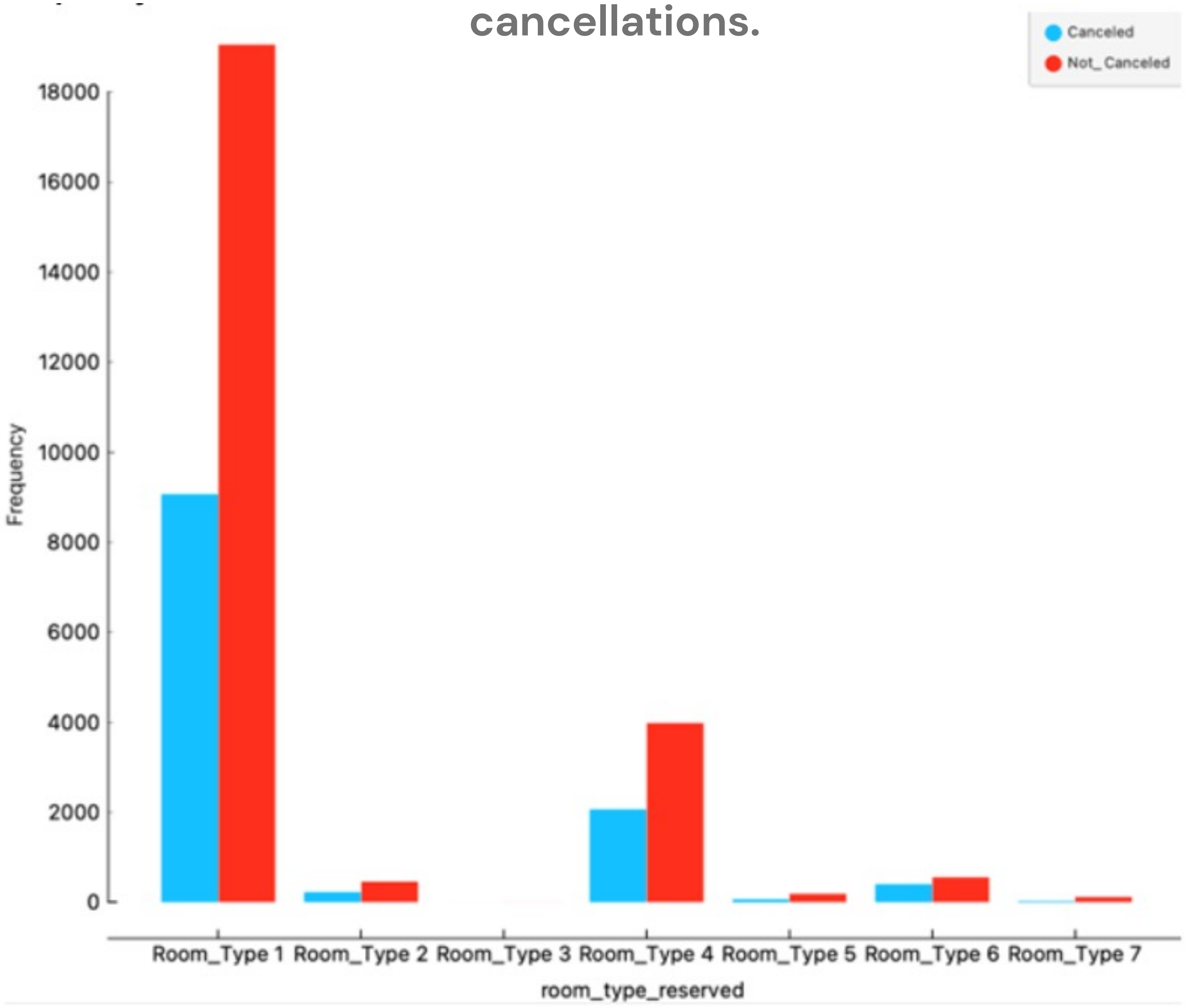
Lead time has the greatest impact on guest cancellations.

RECOMMENDATION #3

Create a loyalty program that prioritizes repeat guests in the booking process, allowing them to reserve rooms earlier.

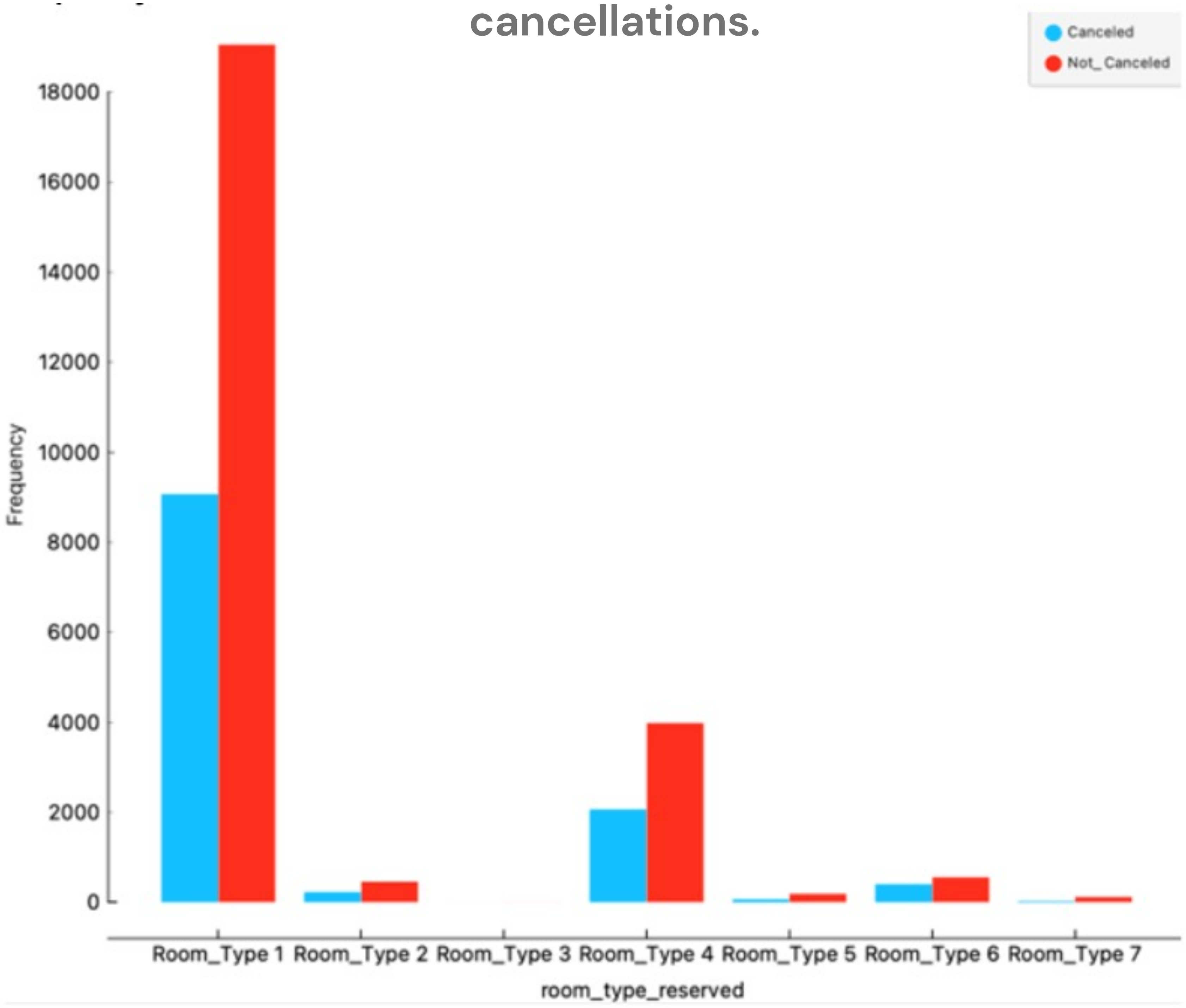
SOLUTION #4

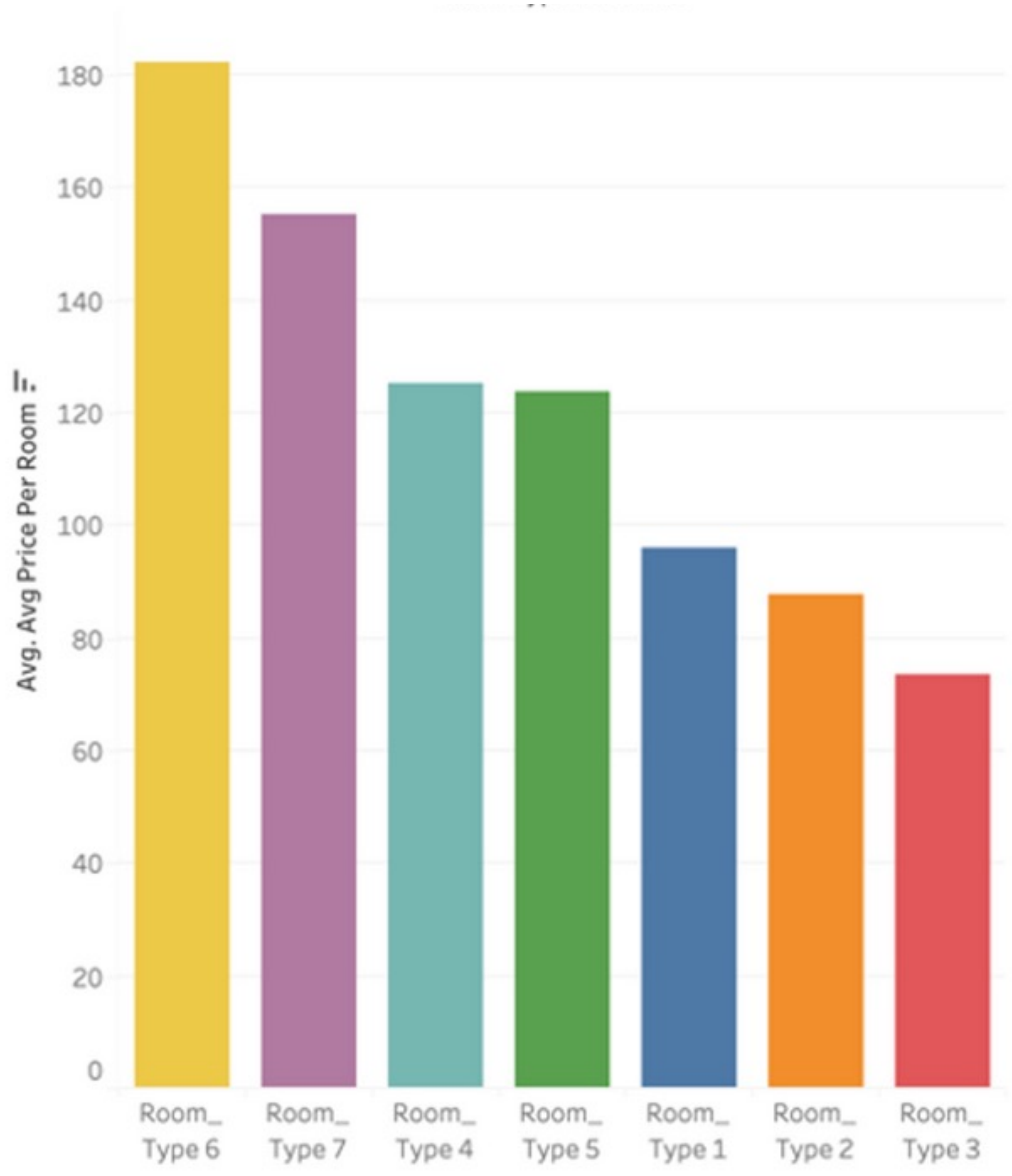
Room Type 1 has the most bookings and the most cancellations.



SOLUTION #4

Room Type 1 has the most bookings and the most cancellations.





Room Types 1 and 4 are moderately priced, relative to other rooms.

BACKGROUND INFORMATION
OUR GOAL
SOLUTION
LIMITATIONS

RECOMMENDATION #4

Restructure what Room Type 1 and Room Type 4 offer guests and how the value proposition is communicated to customers.

04.

LIMITATIONS

LIMITATIONS

Number of beds, bathroom configurations, and other amenities not provided in dataset

Pre-COVID Data (2017-2018)

Cancellations are circumstantial

05.

CONCLUSION

CONCLUSION

Implement predictive model on website to determine probability of reservation cancellation.

Target corporate groups by offering incentives and raise cancellation fees to decrease chances of cancellation.

Create a loyalty program that prioritizes repeat guests, allowing them to reserve rooms earlier.

Restructure what Room Type 1 and Type 4 offer and how the value proposition is communicated to guests.

06.

QUESTIONS?

07.

WORKS CITED

CITATIONS

Raza, A. (2023, January 4). Hotel Reservations Dataset. Kaggle.

https://www.kaggle.com/datasets/ahsan81/hotel-reservations-classification-dataset

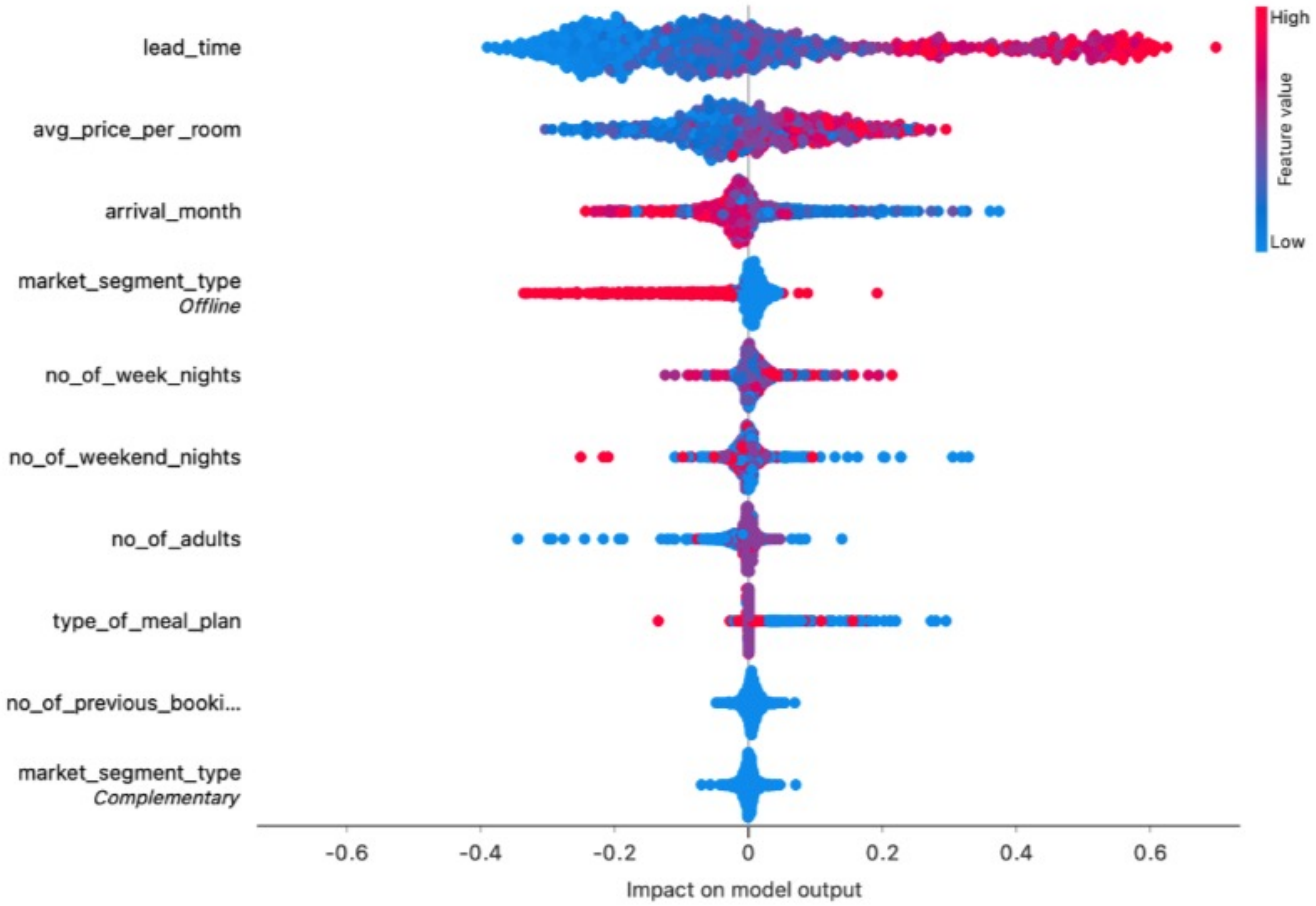
Revenue Hub. (2020). Cancellations Rates: Where do They Stand and How to Overcome Them?

https://revenue-hub.com/cancellations-rates-where-do-they-stand-and-how-to-overcome-them/

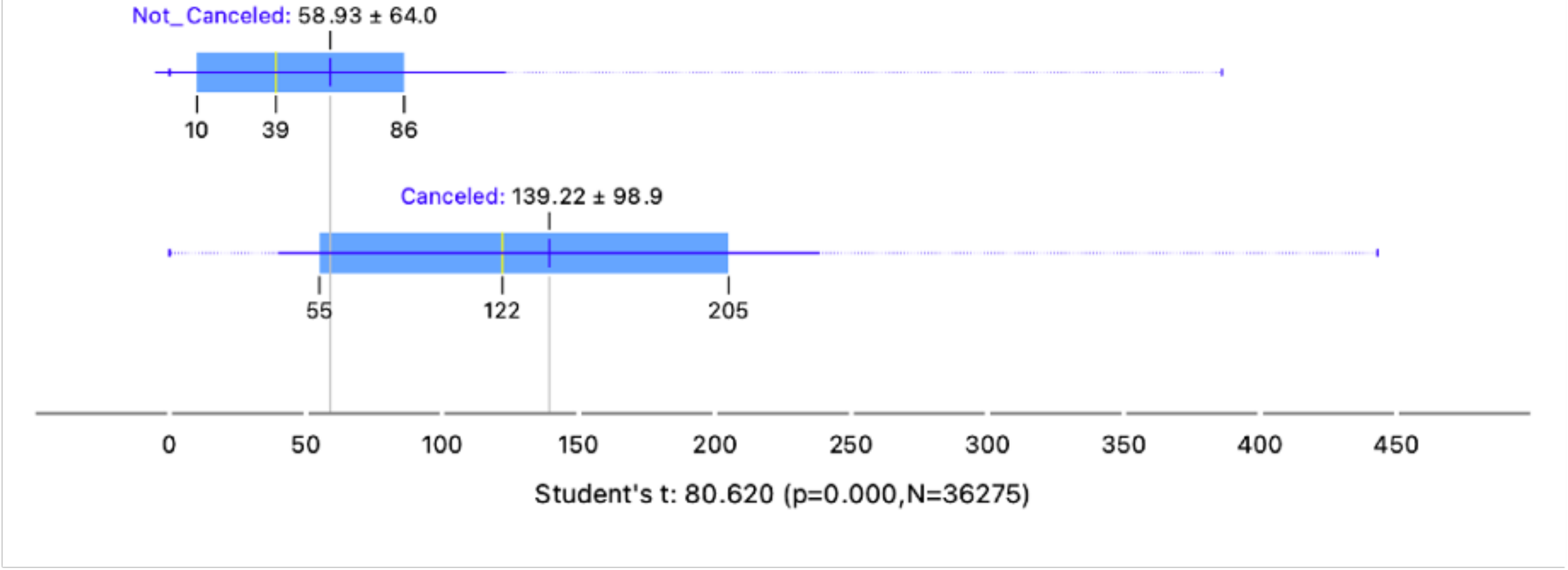
08.

APPENDIX

Lead time, price, and arrival month have the largest impact on whether a guest will cancel their room.



Hotel reservations booked closer to a guest's arrival date are less likely to be canceled.



Box plot for attribute 'lead_time' grouped by 'booking_status'

Decision Tree – Confusion Matrix

Actual \ Predicted	Canceled	Not_Canceled	Σ
Canceled	6598	1782	8380
Not_Canceled	1351	15662	17013
Σ	7949	17444	25393

Parameters for Models