

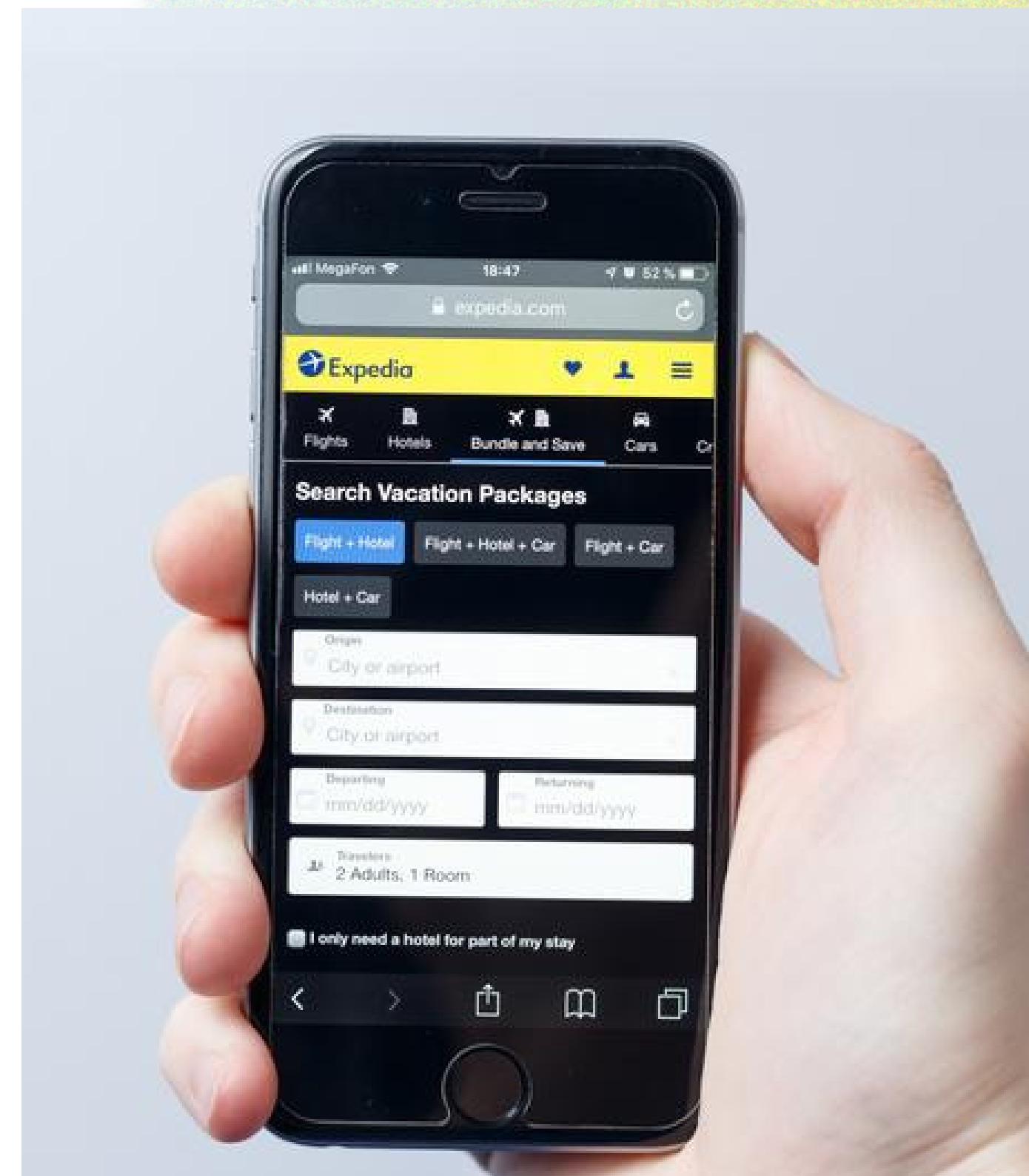
Expedia Hotel Recommendations

By Lais Amorim Menezes



Problem statement

- The website displays random Hotels to all the customer.
- Customer spend too much time on the website trying to find a hotel that matches their preferences.
- Expedia loses customer to other websites that are quick and easier to navigate.



Impact

Customer

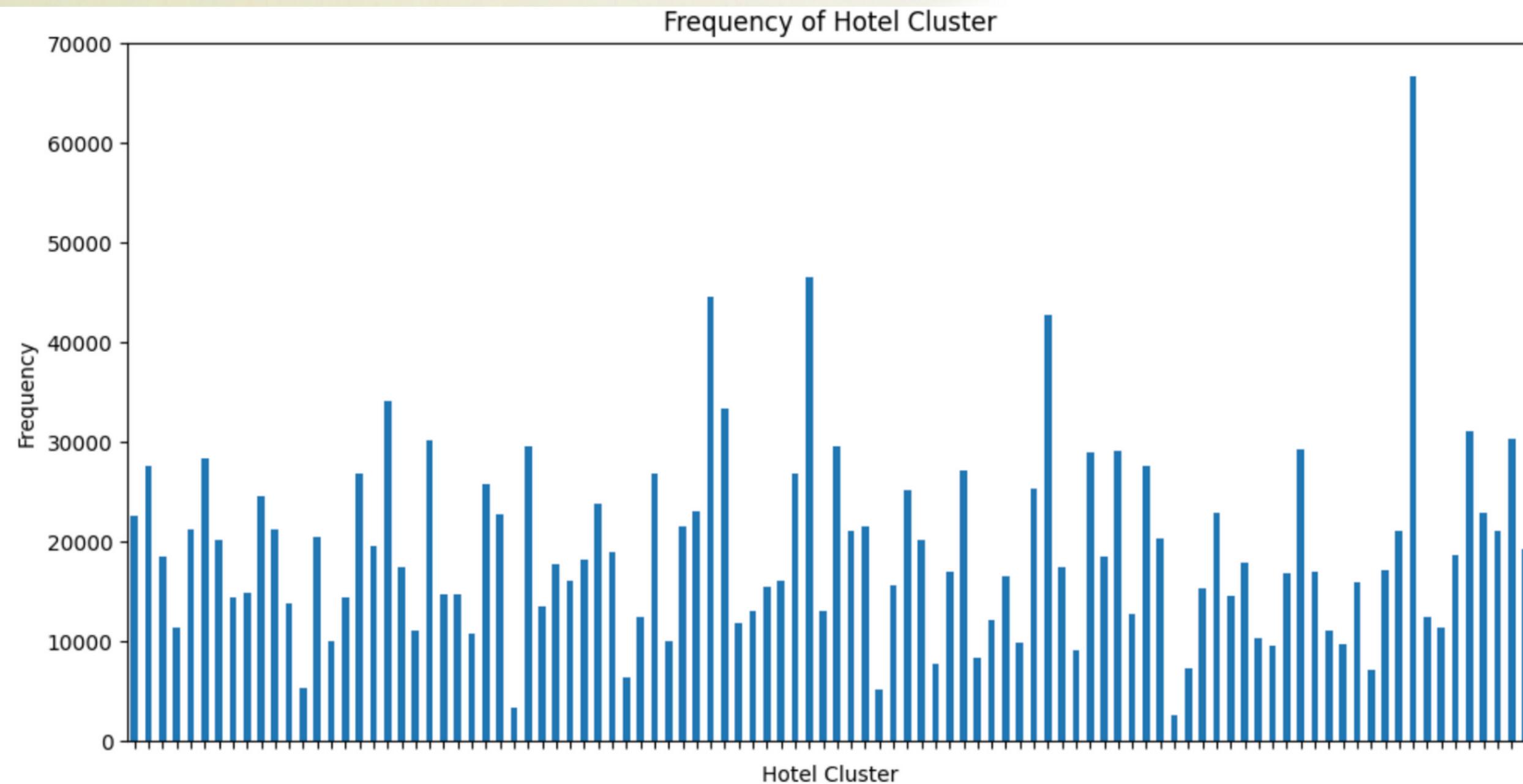


Company



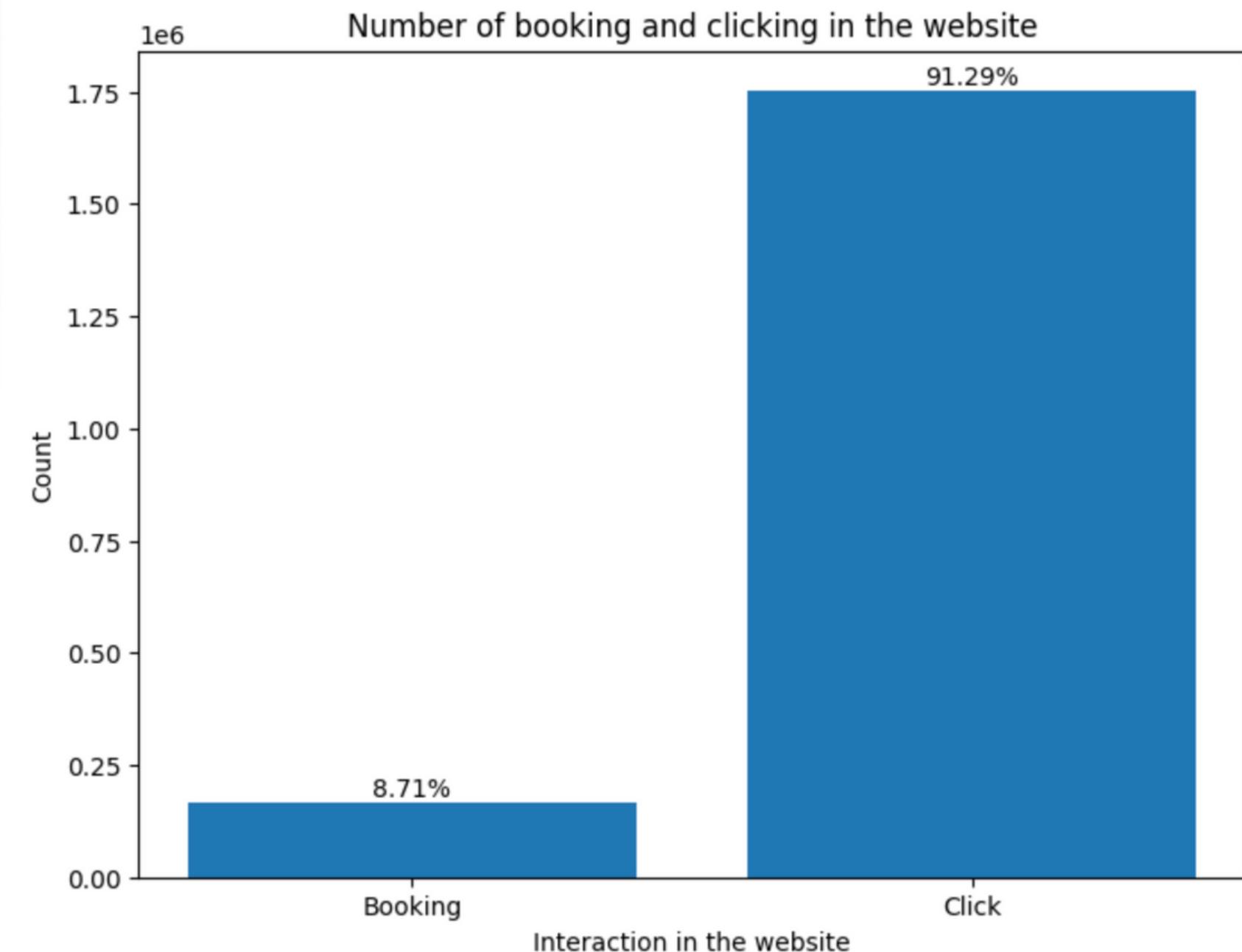
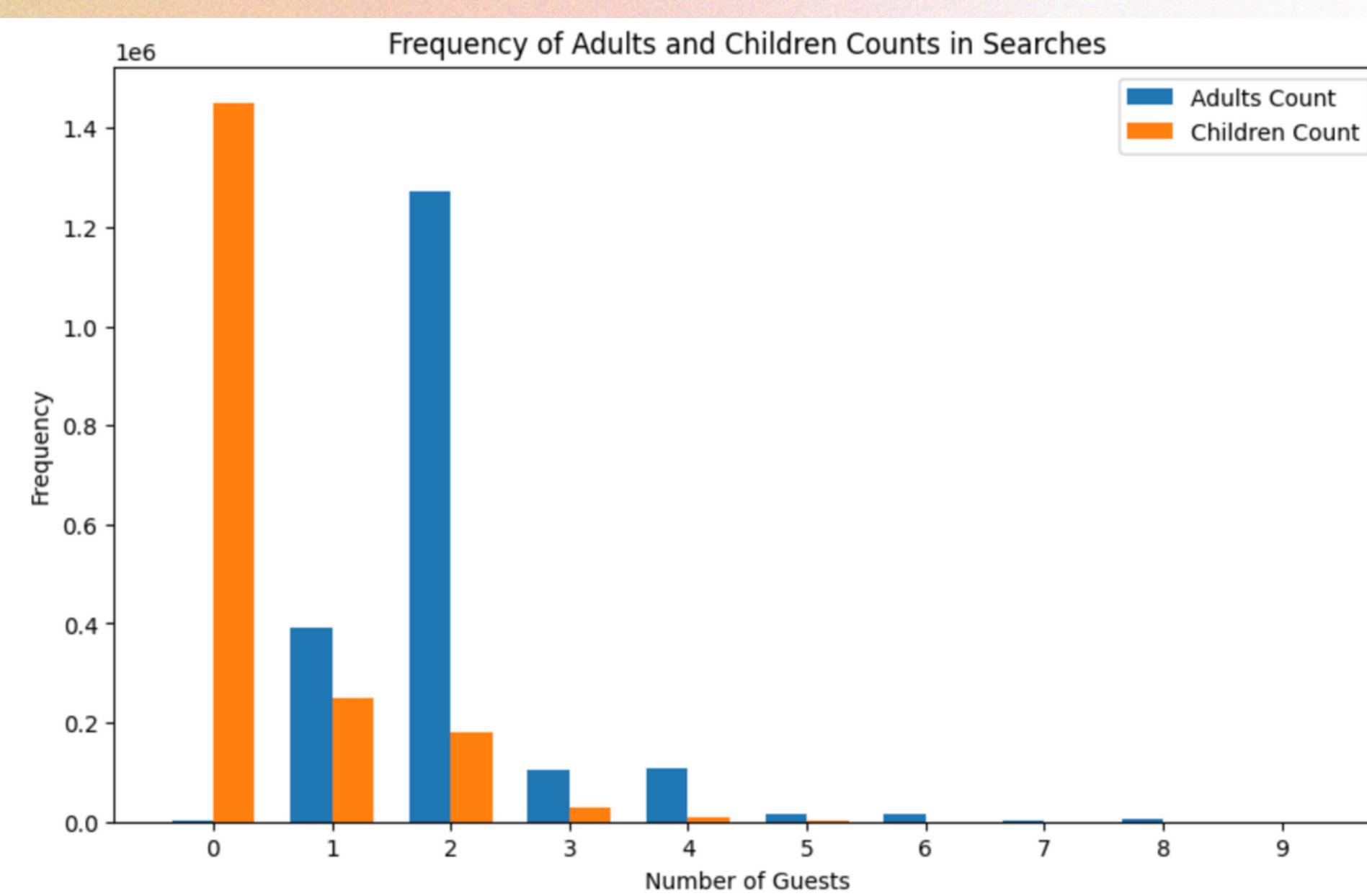
Dataset overview

- The train data from 2013 to July 2014.
- The test data from August to December 2014.
- The dataset doesn't contain categorical columns, instead , all columns are stored numerically.
- The target column will be hotel cluster, the distribution is multimodal with several peaks.

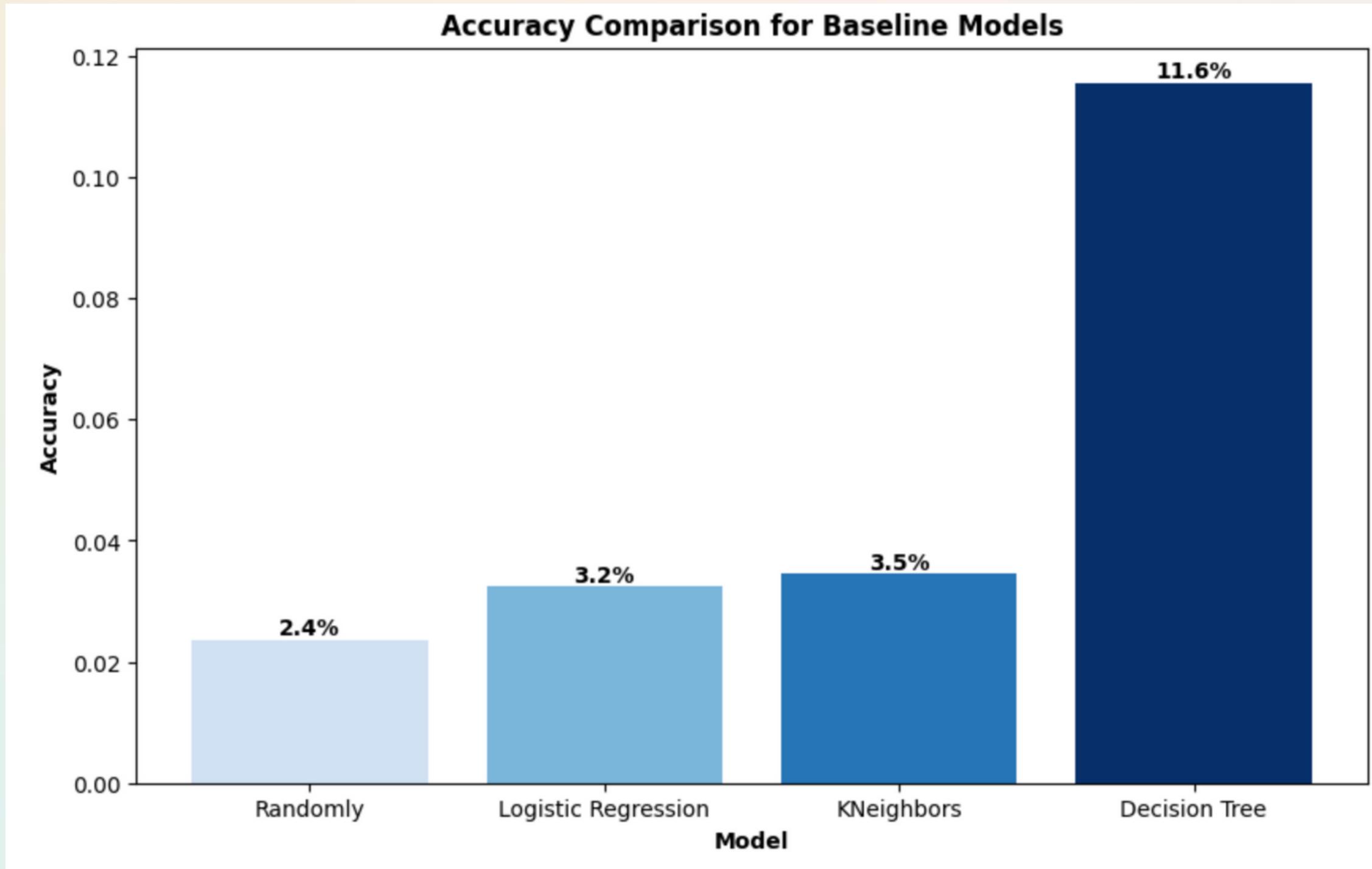


Data overview

- Most of the customers book a hotel for 1 day and for 2 persons without children.
- 91.29% of the data is just customers clicking.



Baseline Models and Evaluation Metrics



- Randomly
- Logistic Regression
- KNeighbors
- Decision Tree (The best one)

Best model: Decision Tree

	Base line Decision Tree	Hyperparameter Decision Tree
Accuracy	11.55%	15.49% 
Precision	11.59%	14.50%
Recall	11.55%	15.49%
F1 Score	11.55%	13.54%

Aplication

localhost:8501

Deploy :

Hotel Recommendation Expedia

In which country do you intend to travel?

Canada

What is the intended duration of your stay at the hotel?

5

Check-in Day

Month:

Day:

4 6

Number of adults

2

0 10

Number of children

0

0 10

Number of room

1

0 10

Recommend

The recommended hotel cluster is: 91



Thank you!