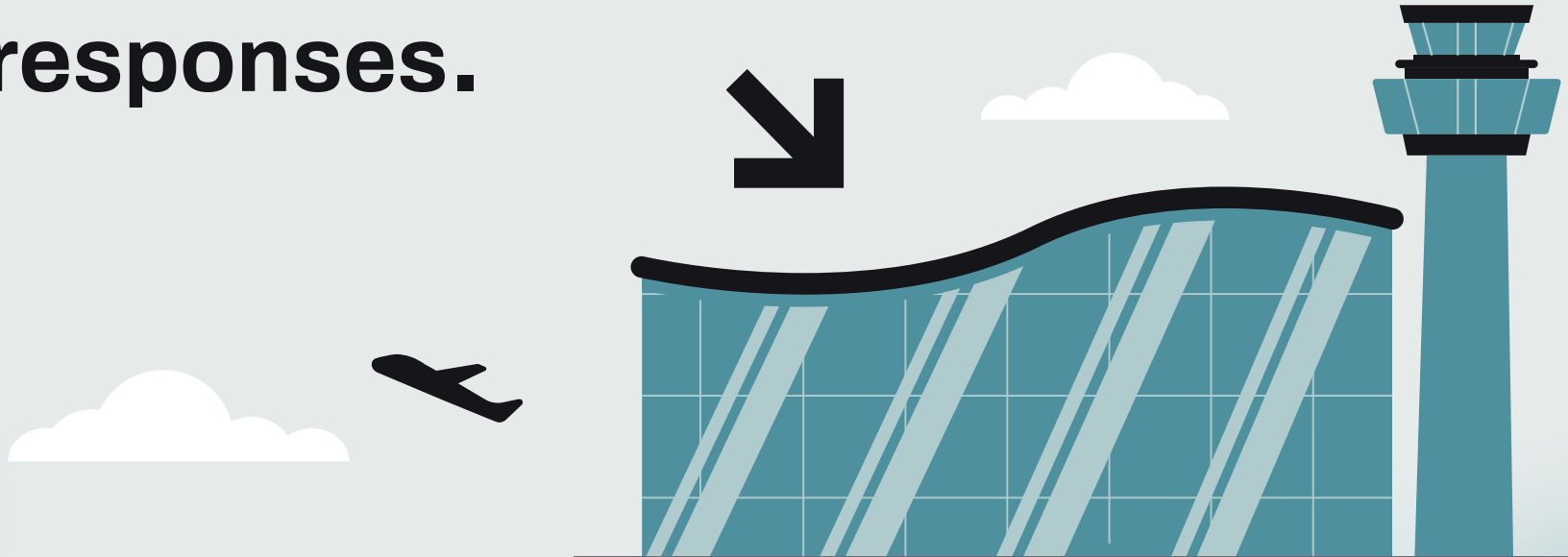


**Customer reviews of
several airlines' services
and their predicted
responses.**



BUSINESS UNDERSTANDING

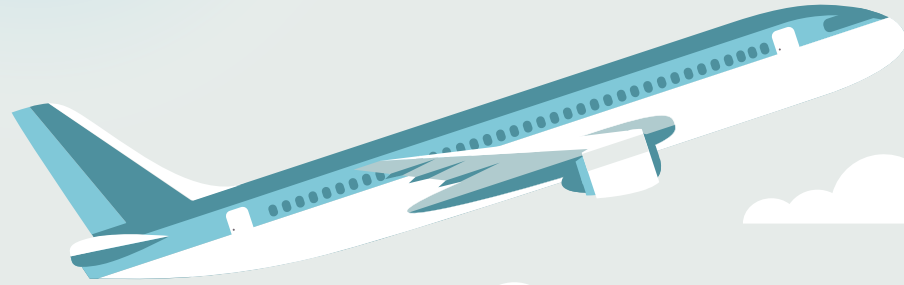
Building and Evaluating a Machine Learning Model Using Airline Reviews Data

Problem Description: The task is to develop a machine learning model that predicts whether a customer is likely to recommend an airline based on their reviews. This is a supervised learning problem, specifically a binary classification task, with the following output classes:

1: Yes (Customer is likely to recommend)

0: No (Customer is not likely to recommend)

1



**DATA
UNDERSTANDING**



Data Description: The dataset used for this task consists of airline reviews from various airlines from 2013-2024, comprising 6,216 records after data cleaning (equivalent to 6,216 reviews). Each review includes 17 relevant attributes:

- Airline: Airline name
- Type of Traveler: Type of traveler
- Month Flown: Month of flight
- Class: Ticket class
- Seat Comfort: Seat comfort rating
- Food and Beverages: Food and beverage rating
- Inflight Entertainment: Inflight entertainment rating
- Value For Money: Value for money rating
- Overall Rating: Overall rating
- Recommended: Whether the customer recommends the airline or not
- From: Departure point
- Via: Transit point
- To: Destination
- Name: Customer name
- Staff Service: Staff service rating
- Month: Month of flight
- Year: Year of flight

0	Name	6216 non-null	object
1	Airline	6216 non-null	object
2	Type of Traveller	6216 non-null	object
3	Month Flown	6216 non-null	datetime64[ns]
4	Class	6216 non-null	object
5	Seat Comfort	6216 non-null	int64
6	Staff Service	6216 non-null	int64
7	Food & Beverages	6216 non-null	int64
8	Inflight Entertainment	6216 non-null	int64
9	Value For Money	6216 non-null	int64
10	Overall Rating	6216 non-null	int64
11	Recommended	6216 non-null	object
12	Via	6216 non-null	object
13	From	6216 non-null	object
14	To	6216 non-null	object
15	Month	6216 non-null	datetime64[ns]
16	Year	6216 non-null	object

DATA PREPROCESSING

Data Description

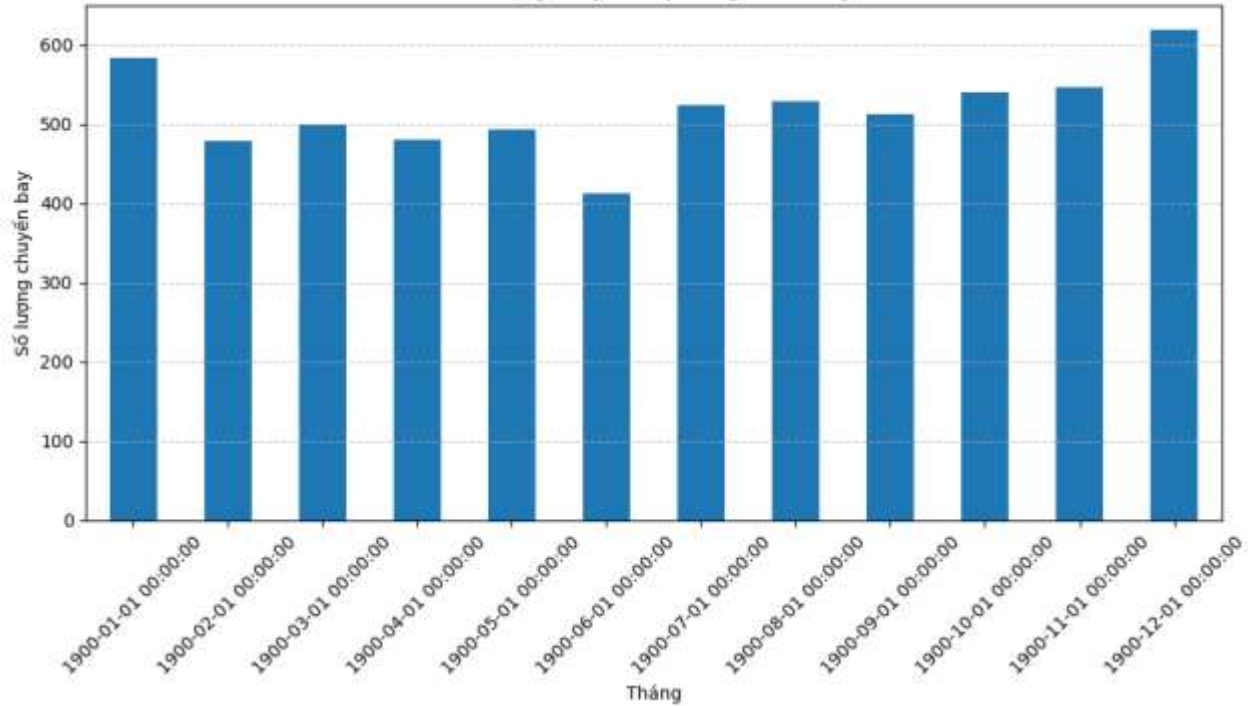
	Month	Flow	Seat Comfort	Staff Service	Food & Beverages	Inflight Entertainment	Value For Money	Overall Rating	Month
count		6216	6216.000000	6216.000000	6216.000000	6216.000000	6216.000000	6216.000000	6216
mean	2019-11-26 00:42:51.428571392		3.402671	3.568050	3.367117	3.635457	3.116634	5.573842	1900-06-20 14:24:19.459459328
min	2013-03-01 00:00:00		1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1900-01-01 00:00:00
25%	2018-03-01 00:00:00		2.000000	2.000000	2.000000	3.000000	1.000000	2.000000	1900-03-01 00:00:00
50%	2019-06-01 00:00:00		4.000000	4.000000	4.000000	4.000000	3.000000	6.000000	1900-07-01 00:00:00
75%	2022-02-01 00:00:00		5.000000	5.000000	5.000000	5.000000	5.000000	9.000000	1900-10-01 00:00:00
max	2024-03-01 00:00:00		5.000000	5.000000	5.000000	5.000000	5.000000	10.000000	1900-12-01 00:00:00
std		NaN	1.401018	1.527396	1.459909	1.310638	1.589307	3.535502	NaN



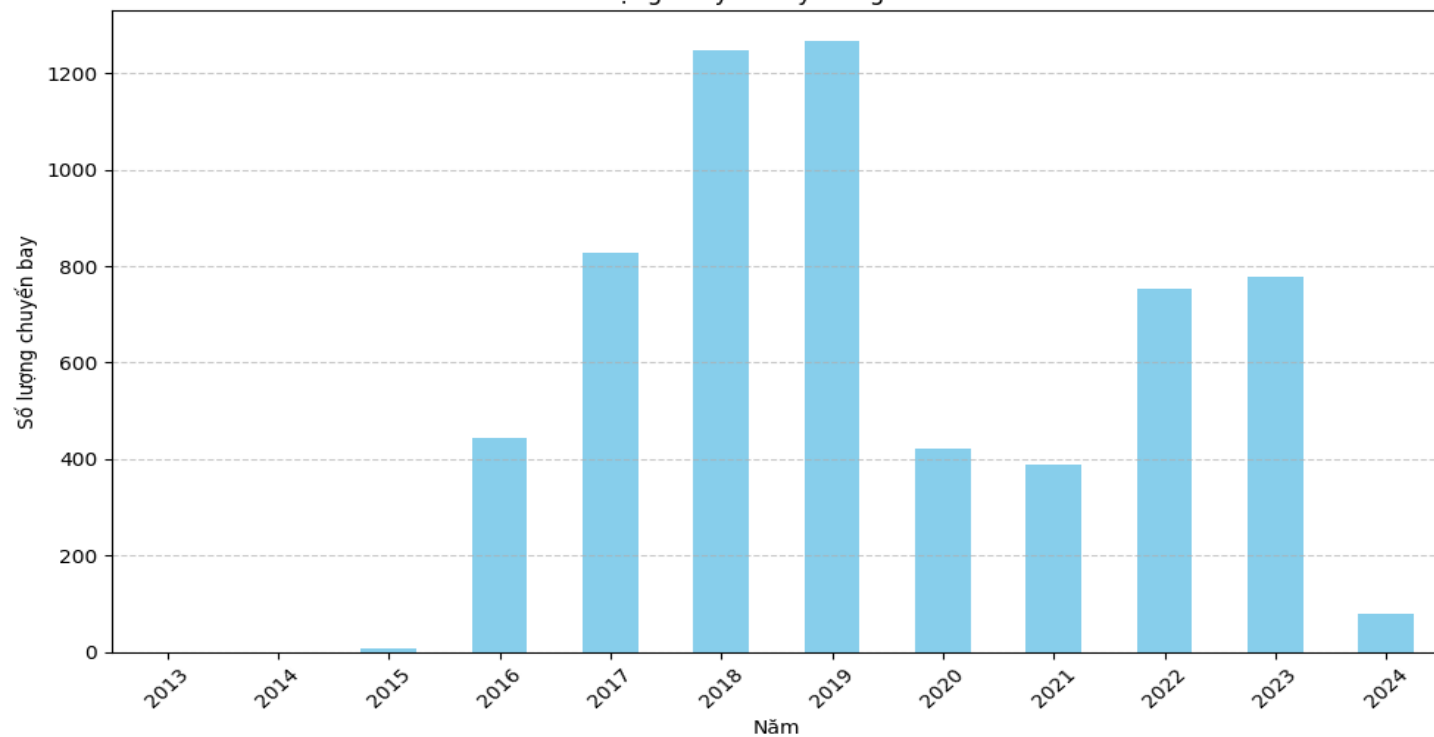
01

EDA

Số lượng chuyến bay trong các tháng

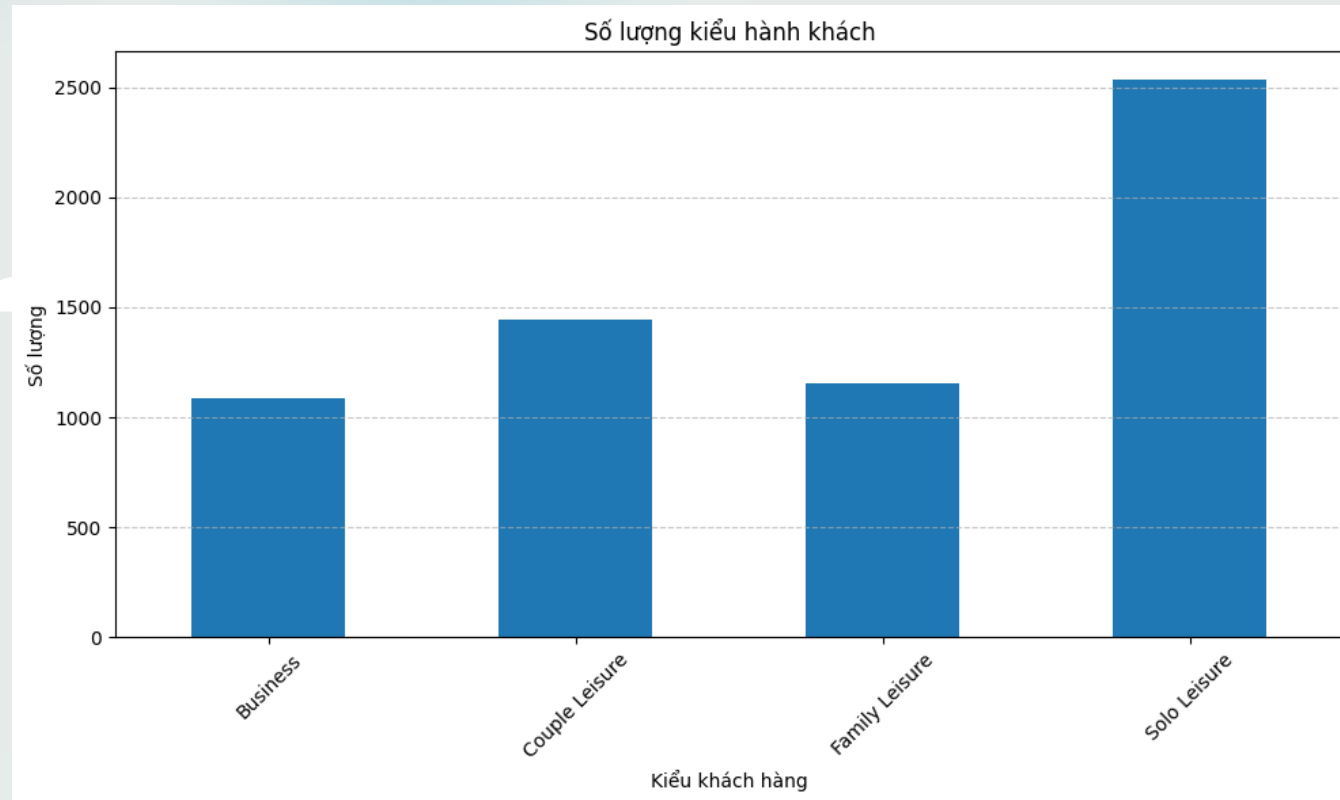


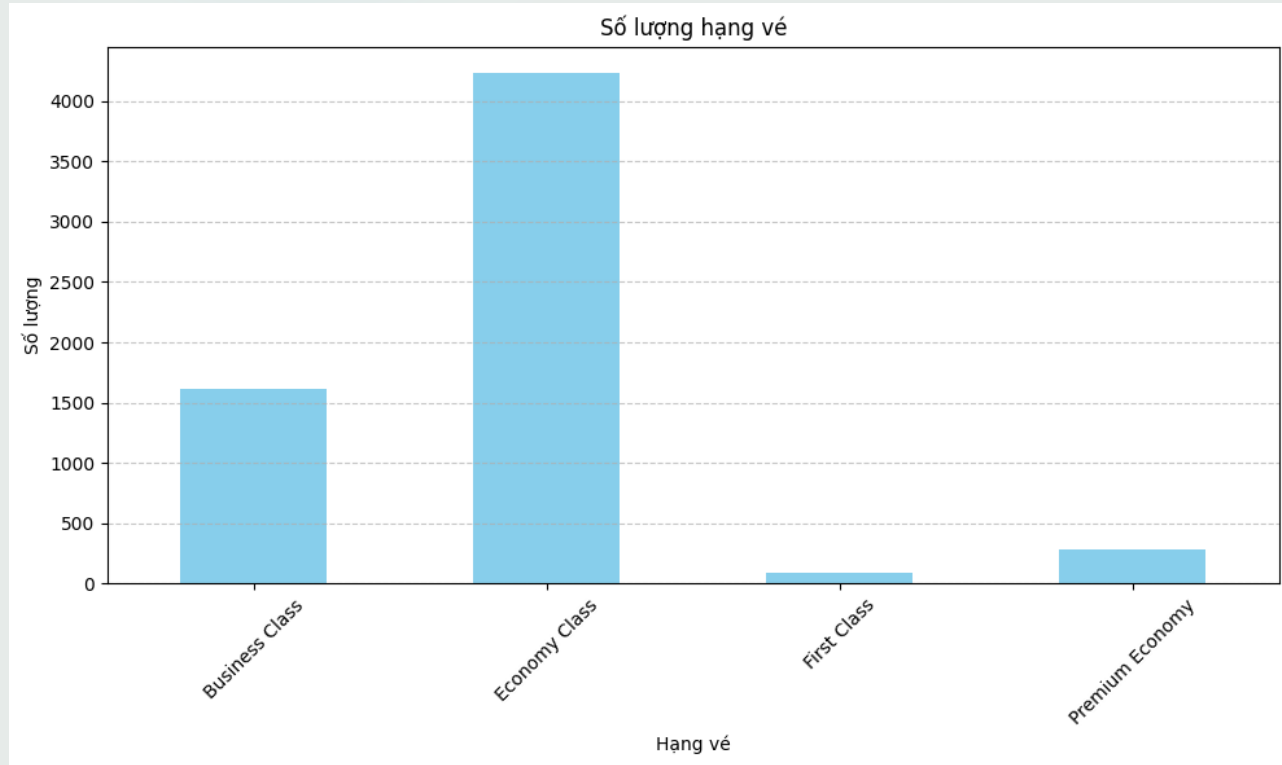
Số lượng chuyến bay trong các năm



The airlines began collecting reviews

Air France	2016
All Nippon Airways	2016
Cathay Pacific Airways	2015
EVA Air	2015
Emirates	2015
Japan Airlines	2015
Korean Air	2016
Qatar Airways	2016
Singapore Airlines	2013
Turkish Airlines	2015





Airport

From	
Singapore	311
London	256
Istanbul	235
Bangkok	224
Dubai	223

...	
India	1
Sochi	1
Saint Petersburg	1
Faisalabad	1
SEA	1

To	
Singapore	336
Bangkok	252
Istanbul	238
Doha	222
London	218

...	
Milano Malpensa	1
Mardin	1
Singapore Return	1
Samarkand	1
ULN	1

Via	
No	3012
Doha	807
Istanbul	674
Dubai	477
Singapore	208

...	
Harare	1
Dubia	1
Dubai, Bangkok, Kuala Lumpur	1
dubai	1
Arlanda Sweden	1

Airline	Food & Beverages
All Nippon Airways	4.015000
Qatar Airways	3.932876
EVA Air	3.845794
Japan Airlines	3.785276
Korean Air	3.537879
Singapore Airlines	3.536618
Cathay Pacific Airways	3.175221
Air France	3.094463
Turkish Airlines	2.972419
Emirates	2.950853

Airline	Seat Comfort
All Nippon Airways	4.140000
Qatar Airways	3.955759
EVA Air	3.920561
Japan Airlines	3.877301
Singapore Airlines	3.668442
Korean Air	3.636364
Cathay Pacific Airways	3.598230
Emirates	3.128385
Air France	2.920195
Turkish Airlines	2.747045

Airline	Staff Service
All Nippon Airways	4.515000
Qatar Airways	4.291381
EVA Air	4.191589
Japan Airlines	4.141104
Singapore Airlines	3.902796
Korean Air	3.825758
Cathay Pacific Airways	3.596460
Air France	3.298046
Emirates	2.939819
Turkish Airlines	2.879433

Airline	Value For Money
All Nippon Airways	4.085000
EVA Air	3.929907
Qatar Airways	3.711670
Japan Airlines	3.687117
Korean Air	3.469697
Singapore Airlines	3.452730
Cathay Pacific Airways	3.320354
Emirates	2.674022
Air France	2.672638
Turkish Airlines	2.375099

Airline	Overall Rating
All Nippon Airways	7.850000
EVA Air	7.224299
Qatar Airways	7.072464
Japan Airlines	6.858896
Singapore Airlines	6.549933
Korean Air	6.136364
Cathay Pacific Airways	6.099115
Air France	4.664495
Emirates	4.520562
Turkish Airlines	3.620961

Airline	Inflight Entertainment
Qatar Airways	4.130435
All Nippon Airways	3.890000
Singapore Airlines	3.882823
Cathay Pacific Airways	3.789381
EVA Air	3.785047
Emirates	3.670010
Japan Airlines	3.521472
Korean Air	3.272727
Air France	3.174267
Turkish Airlines	3.092199

	Airline	Flight	Recommended_rate
1	Qatar Airways	1311	0,71
2	Turkish Airlines	1269	0,29
3	Emirates	997	0,37
4	Singapore Airlines	751	0,65
5	Air France	614	0,41
6	Cathay Pacific Airways	565	0,58
7	EVA Air	214	0,74
8	All Nippon Airways	200	0,8
9	Japan Airlines	163	0,66
10	Korean Air	132	0,6



	Market_share ▾	AVG ▾
1	high	0,5
2	low	0,69
3	medium	0,55

	Airline	Seat_Comfort	Staff_Service	Food_Beverages	Inflight_Entertainment	Value_For_Money	Overall_Rating
1	EVA Air	3	4	3	3	3	7
2	Qatar Airways	3	4	3	4	3	7
3	All Nippon Airways	4	4	4	3	4	7
4	Cathay Pacific Airways	3	3	3	3	3	6
5	Japan Airlines	3	4	3	3	3	6
6	Korean Air	3	3	3	3	3	6
7	Singapore Airlines	3	3	3	3	3	6
8	Emirates	3	2	2	3	2	4
9	Air France	2	3	3	3	2	4
10	Turkish Airlines	2	2	2	3	2	3



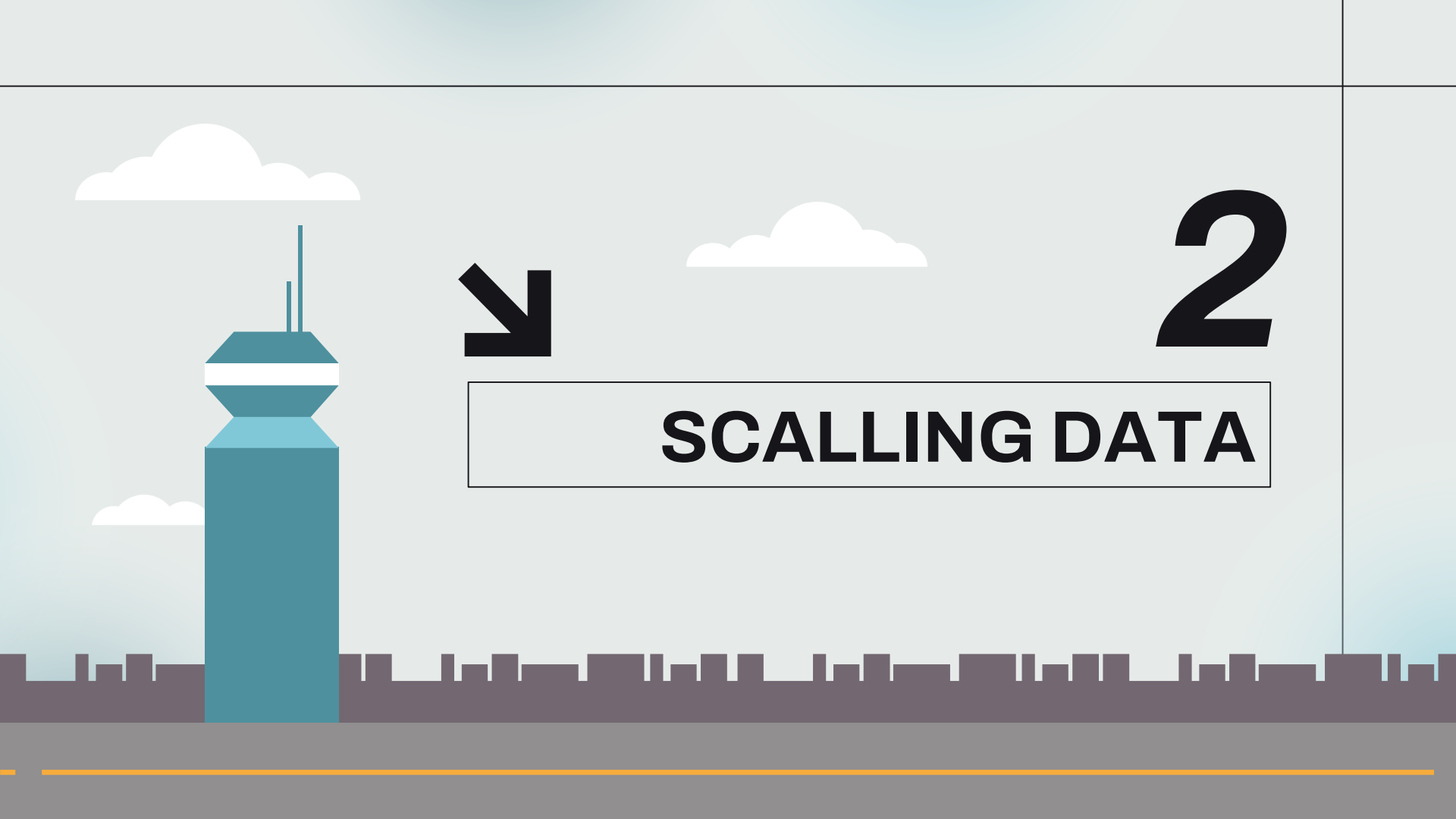
Class

- Dominant at Economy Class: Turkish Airlines, Korean Air, Emirates, Air France
- Dominant at Business Class: Qatar Airways
- Dominant at First Class: Emirates, All Nippon Airways
- Dominant at Premium Economy: EVA Air



Type of Traveler

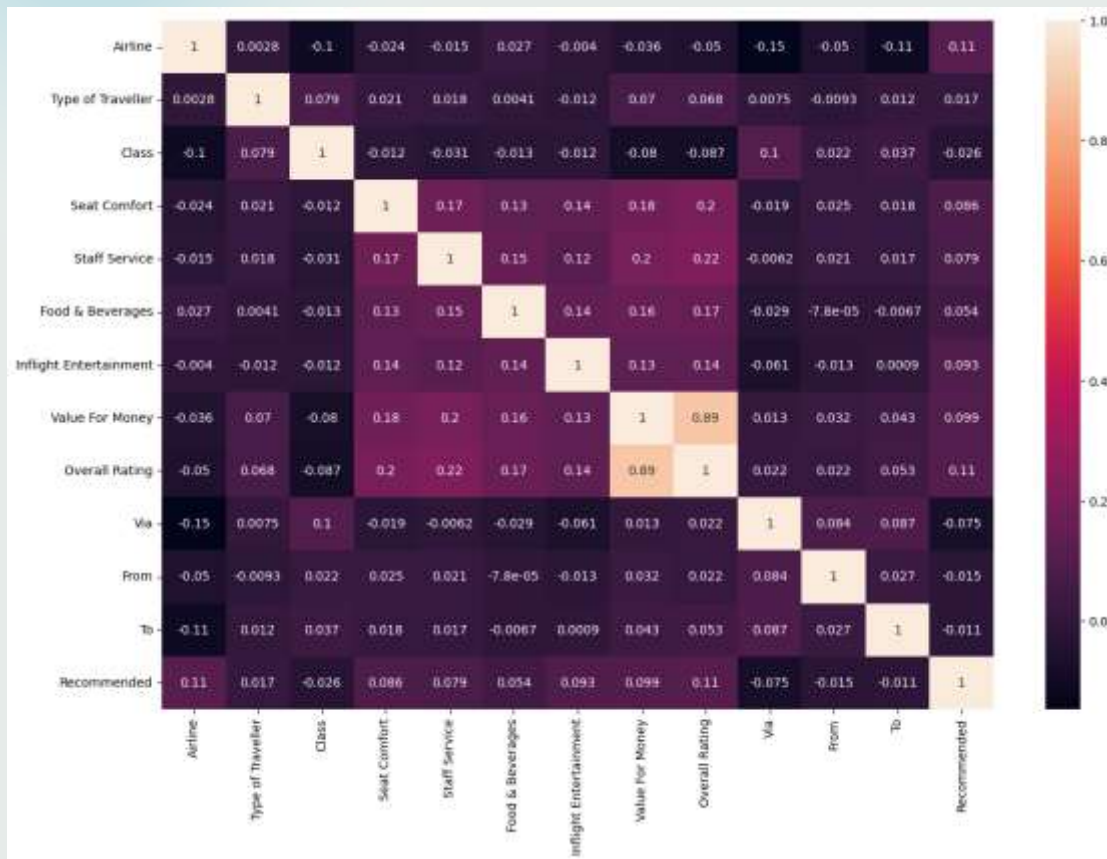
- Business Class: Business
- Economy Class: Family Leisure
- First Class: Solo Leisure
- Premium Economy: Couple Leisure



2

SCALLING DATA

CORRELATION



The key factors influencing customers' decisions to recommend an airline include:



Airline




Seat Comfort



Staff Service



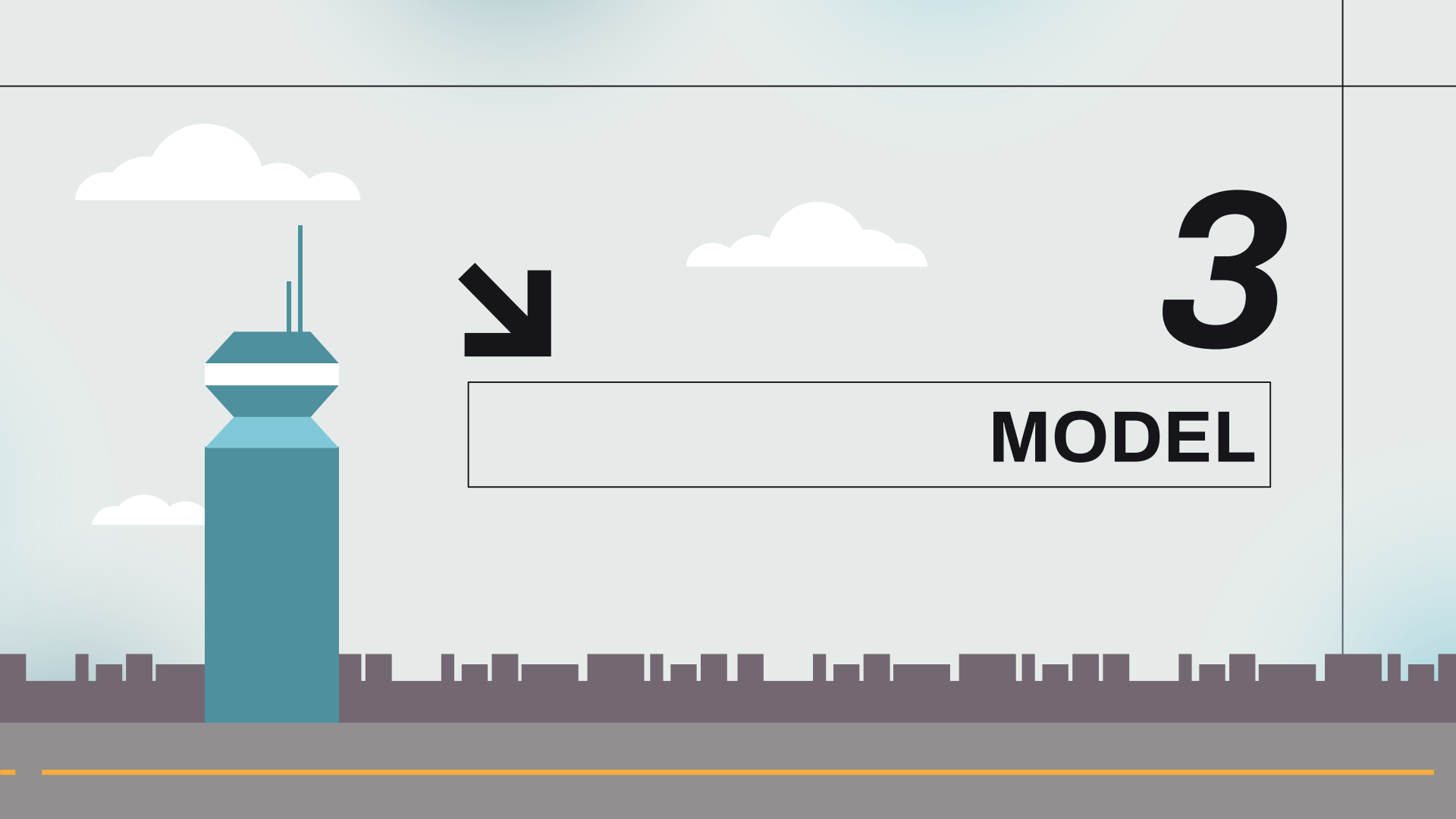
**Inflight
Entertainment**



**Value of
money**



**Transit
experience**



3

MODEL

Accuracy Score of Logistic Regression is : 0.6334405144694534

Confusion Matrix :

[[678 181]
[355 118]]

Classification Report :		recall	f1-score	support
precision				
0.0	0.66	0.87	0.75	779
1.0	0.52	0.24	0.33	465
accuracy			0.63	1244
macro avg	0.59	0.55	0.54	1244
weighted avg	0.61	0.63	0.59	1244

Accuracy Score of Random Forest is : 0.6438868167282572

Confusion Matrix :

[[613 166]
[278 187]]

Classification Report :		recall	f1-score	support
precision				
0.0	0.69	0.79	0.73	779
1.0	0.53	0.48	0.46	465
accuracy			0.64	1244
macro avg	0.61	0.59	0.60	1244
weighted avg	0.63	0.64	0.63	1244

Accuracy Score of Decision Tree is : 0.5393890675241158

Confusion Matrix :

[[469 318]
[263 202]]

Classification Report :		recall	f1-score	support
precision				
0.0	0.64	0.68	0.62	779
1.0	0.39	0.43	0.41	465
accuracy			0.54	1244
macro avg	0.52	0.52	0.52	1244
weighted avg	0.55	0.54	0.54	1244

Accuracy Score of Ada Boost Classifier is : 0.6583681286173034

Confusion Matrix :

[[588 191]
[234 231]]

Classification Report :		recall	f1-score	support
precision				
0.0	0.72	0.75	0.73	779
1.0	0.55	0.58	0.52	465
accuracy			0.66	1244
macro avg	0.63	0.63	0.63	1244
weighted avg	0.65	0.66	0.65	1244

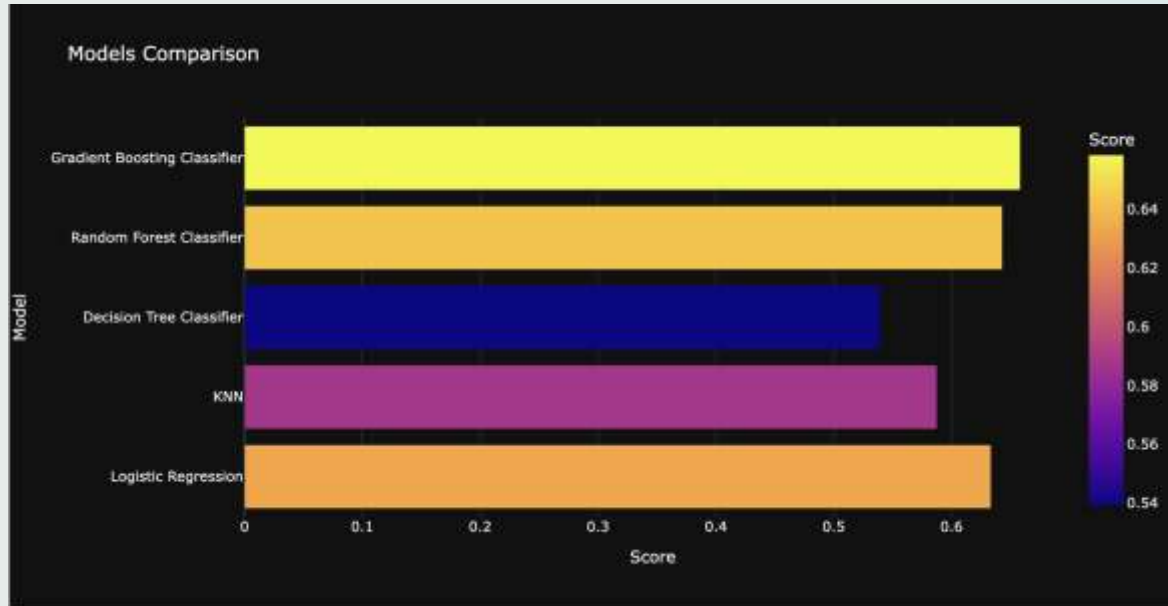
Accuracy Score of XGB is : 0.587628578778135

Confusion Matrix :

[[528 251]
[262 203]]

Classification Report :		recall	f1-score	support
precision				
0.0	0.67	0.68	0.67	779
1.0	0.45	0.44	0.44	465
accuracy			0.59	1244
macro avg	0.56	0.56	0.56	1244
weighted avg	0.59	0.59	0.59	1244

MODEL COMPARISON






4

CONCLUSION



Some general remarks about the data

- The majority of passengers opt for economy class for their flights. However, airlines primarily focused on this segment often lack highly positive reviews. Conversely, airlines dominating the higher fare segments tend to receive very good reviews.
 - In almost all evaluation criteria, All Nippon Airways consistently leads, except for Inflight Entertainment where Qatar Airways leads.
 - Qatar Airways, Turkish Airlines, and Emirates are considered the largest airlines, yet only Qatar Airways receives positive evaluations, while Turkish Airlines and Emirates have very low ratings compared to smaller airlines.
 - The airlines with the highest Recommended ratings are Qatar Airways and All Nippon Airways. The Recommended ratings of airlines are inversely proportional to their scale.
- 
-

The slide features a light blue background with a white city skyline silhouette at the bottom. Three stylized white clouds are positioned in the upper left, center, and lower left areas. A black-bordered box in the upper right contains the title.

Some directions for airlines

- Enhancing the economy class experience is crucial due to its potential large customer base. However, focusing on high-yield passenger segments remains essential for maximizing efficiency.
- Criteria for evaluating airline quality, such as ticket prices and staff service, should also be improved as much as possible.
- Based on this data, pricing strategies can be adjusted during special occasions, leveraging ticket purchase frequency to maximize profits. Additionally, promotions should be offered in regions with low flight volumes to stimulate demand.

