# Airlines Passenger Satisfaction

Rama                                    Mikdad                                    Zehra

# Explaining the dataset

This dataset contains an airline passenger satisfaction survey.

What factors are highly correlated to a satisfied (or dissatisfied) passenger?

| | |
|---|---|
| id | |
| Gender | |
| age | |

| | |
|---|---|
| Customer Type | (Loyal customer, disloyal customer) |
| Type of Travel | (Personal Travel, Business Travel) |
| Class | (Business, Eco, Eco Plus) |
| Satisfaction | (Satisfied, neutral or dissatisfaction) |

Flight distance

Departure Delay in Minutes

Arrival Delay in Minutes

Inflight wifi service

Departure/Arrival time convenient

Ease of Online booking

Gate location

Food and drink:

Online boarding

Seat comfort:

Inflight entertainment:
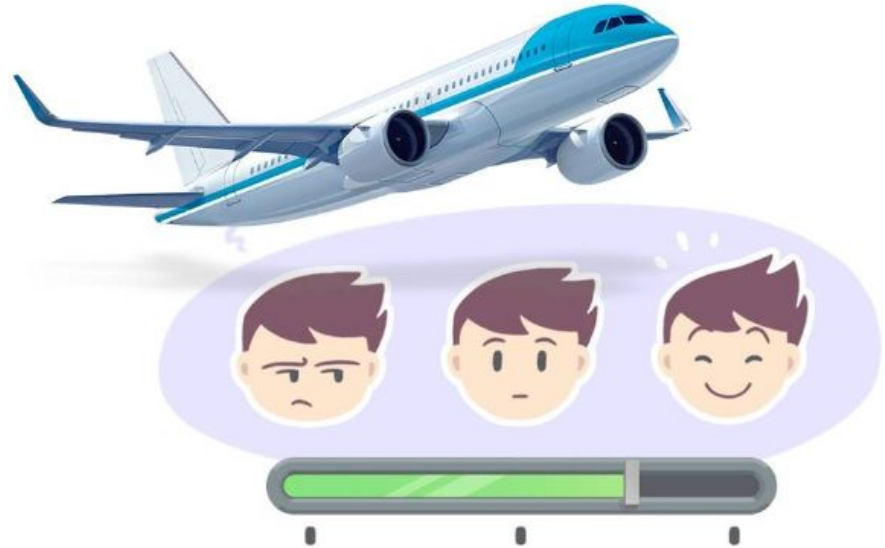
On-board service:

Leg room service:

Baggage handling:

Check-in service:

Inflight service:

Cleanliness

overall

# 1- Cleaning the data:



the percentage of missing value in
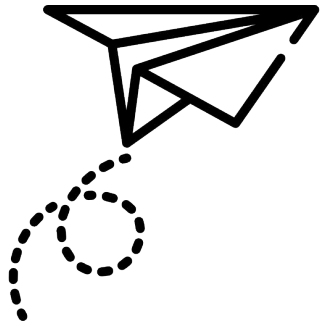comparison to the whole data:
0.003026


So we decided to drop them

```
df.isnull().sum()
```

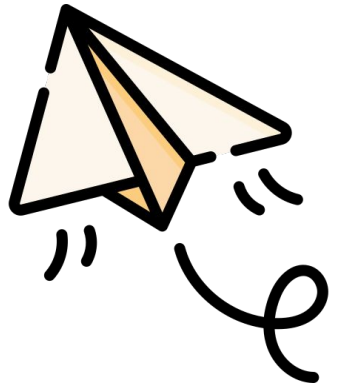| | |
|---|---|
| id | 0 |
| Gender | 0 |
| Customer Type | 0 |
| Age | 0 |
| Type of Travel | 0 |
| Class | 0 |
| Flight Distance | 0 |
| Inflight wifi service | 0 |
| Departure/Arrival time convenient | 0 |
| Ease of Online booking | 0 |
| Gate location | 0 |
| Food and drink | 0 |
| Online boarding | 0 |
| Seat comfort | 0 |
| Inflight entertainment | 0 |
| On-board service | 0 |
| Leg room service | 0 |
| Baggage handling | 0 |
| Checkin service | 0 |
| Inflight service | 0 |
| Cleanliness | 0 |
| Departure Delay in Minutes | 0 |
| Arrival Delay in Minutes | 393 |
| satisfaction | 0 |
| dtype: int64 | |

## 2- Ensuring that data is in the correct format (e.g. dates are in the correct format, categorical variables are properly encoded, etc.)

```
id                                [70172, 5047, 110028, 24026, 119299, 111157, 8...
Gender                                                            [Male, Female]
Customer Type                                     [Loyal Customer, disloyal Customer]
Age                               [13, 25, 26, 61, 47, 52, 41, 20, 24, 12, 53, 3...
Type of Travel                                   [Personal Travel, Business travel]
Class                                                 [Eco Plus, Business, Eco]
Flight Distance                   [460, 235, 1142, 562, 214, 1180, 1276, 2035, 8...
Inflight wifi service                                        [3, 2, 4, 1, 5, 0]
Departure/Arrival time convenient                            [4, 2, 5, 3, 1, 0]
Ease of Online booking                                       [3, 2, 5, 4, 1, 0]
Gate location                                                [1, 3, 2, 5, 4, 0]
Food and drink                                               [5, 1, 2, 4, 3, 0]
Online boarding                                              [3, 5, 2, 1, 4, 0]
Seat comfort                                                 [5, 1, 2, 3, 4, 0]
Inflight entertainment                                       [5, 1, 2, 3, 4, 0]
On-board service                                             [4, 1, 2, 3, 5, 0]
Leg room service                                             [3, 5, 4, 2, 1, 0]
Baggage handling                                                [4, 3, 5, 1, 2]
Checkin service                                              [4, 1, 3, 5, 2, 0]
Inflight service                                             [5, 4, 3, 1, 2, 0]
Cleanliness                                                  [5, 1, 2, 3, 4, 0]
Departure Delay in Minutes        [25, 1, 0, 11, 9, 4, 28, 43, 49, 7, 17, 52, 54...
Arrival Delay in Minutes          [18.0, 6.0, 0.0, 9.0, 23.0, 8.0, 35.0, 51.0, 1...
satisfaction                                       [neutral or dissatisfied, satisfied]
dtype: object
```
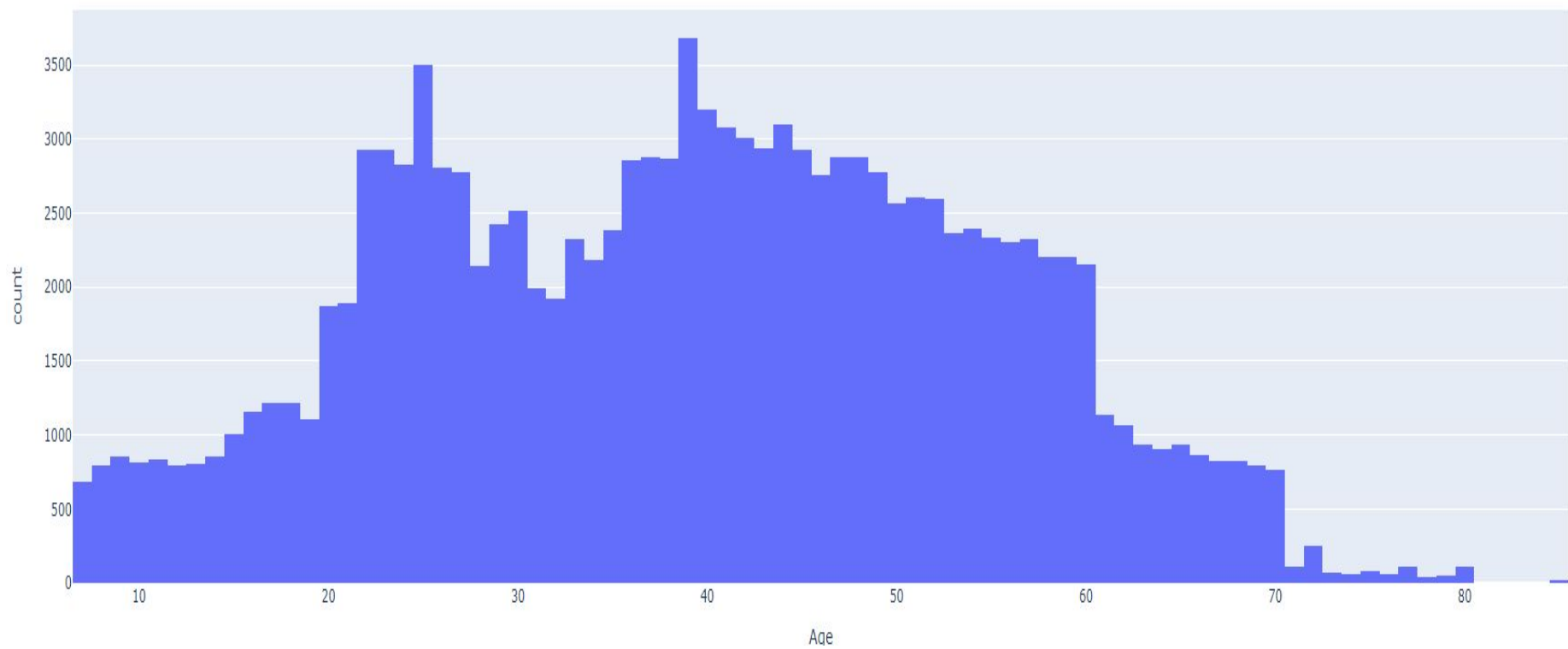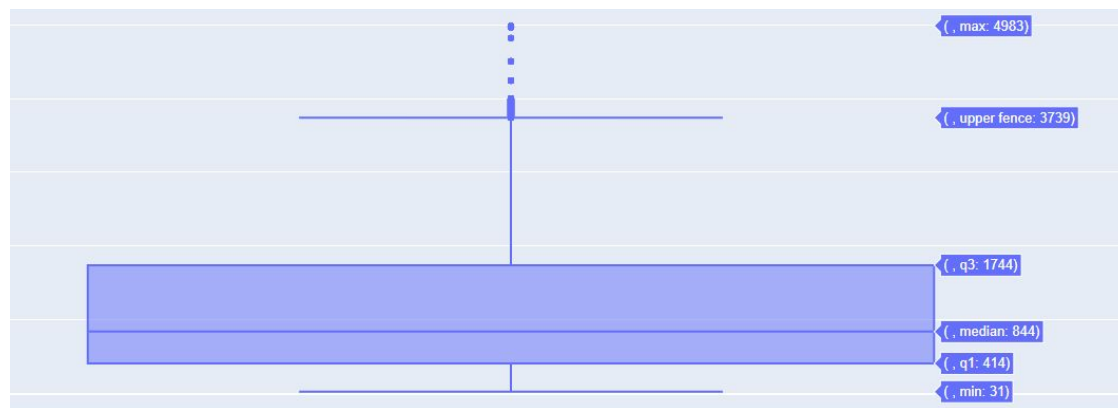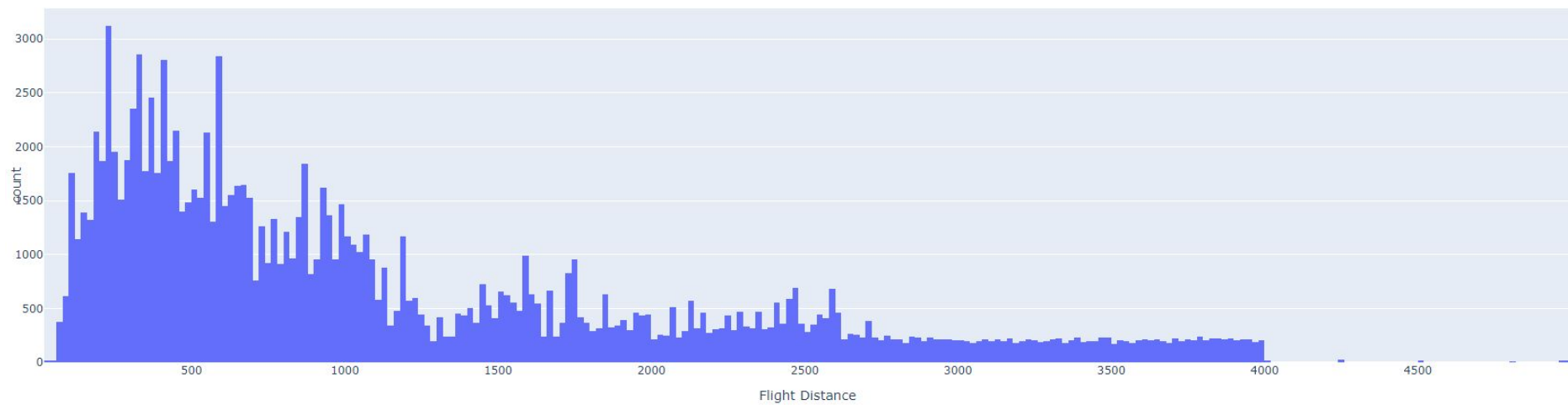
3- check if there are any duplicate rows

4- check for outliers

Flight Distance

( , max: 4983)

( , upper fence: 3739)

( , q3: 1744)

( , median: 844)

( , q1: 414)

( , min: 31)

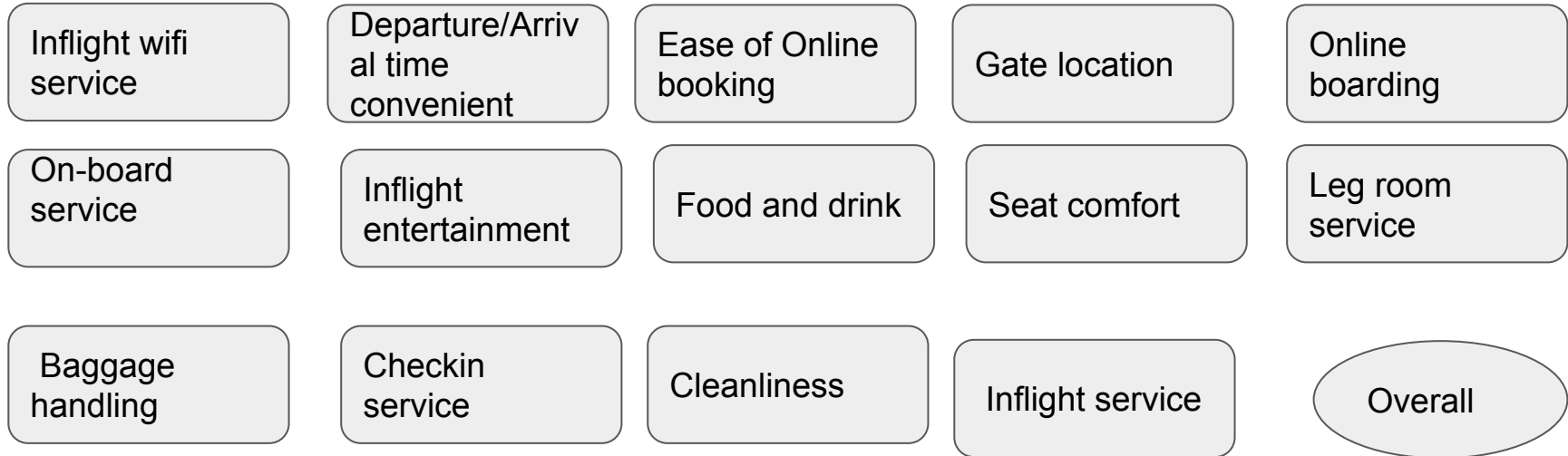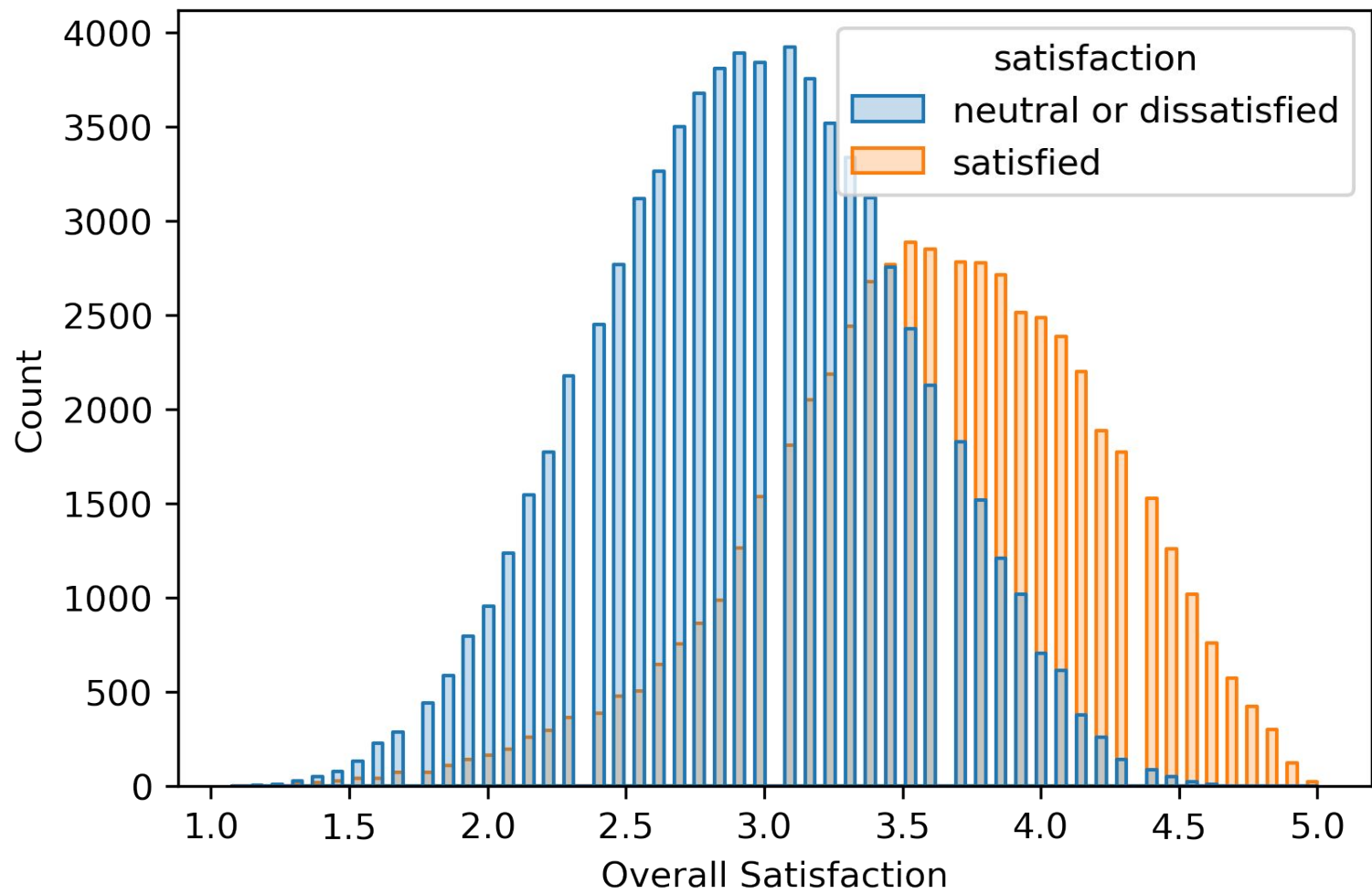# Mikdad's part in answering the research questions:
## Question 1 + 2

# Quantifying the 'Satisfaction'

We thought a continuous variable is beneficial to measure the 'overall satisfaction'

The 'overall satisfaction' is the mean of scores for each factor :

| | | | | |
|---|---|---|---|---|
| Inflight wifi service | Departure/Arrival time convenient | Ease of Online booking | Gate location | Online boarding |
| On-board service | Inflight entertainment | Food and drink | Seat comfort | Leg room service |
| Baggage handling | Checkin service | Cleanliness | Inflight service | Overall |

Distribution of the overall satisfaction for the two groups

"Density of Overall Satisfaction Scores for the two groups
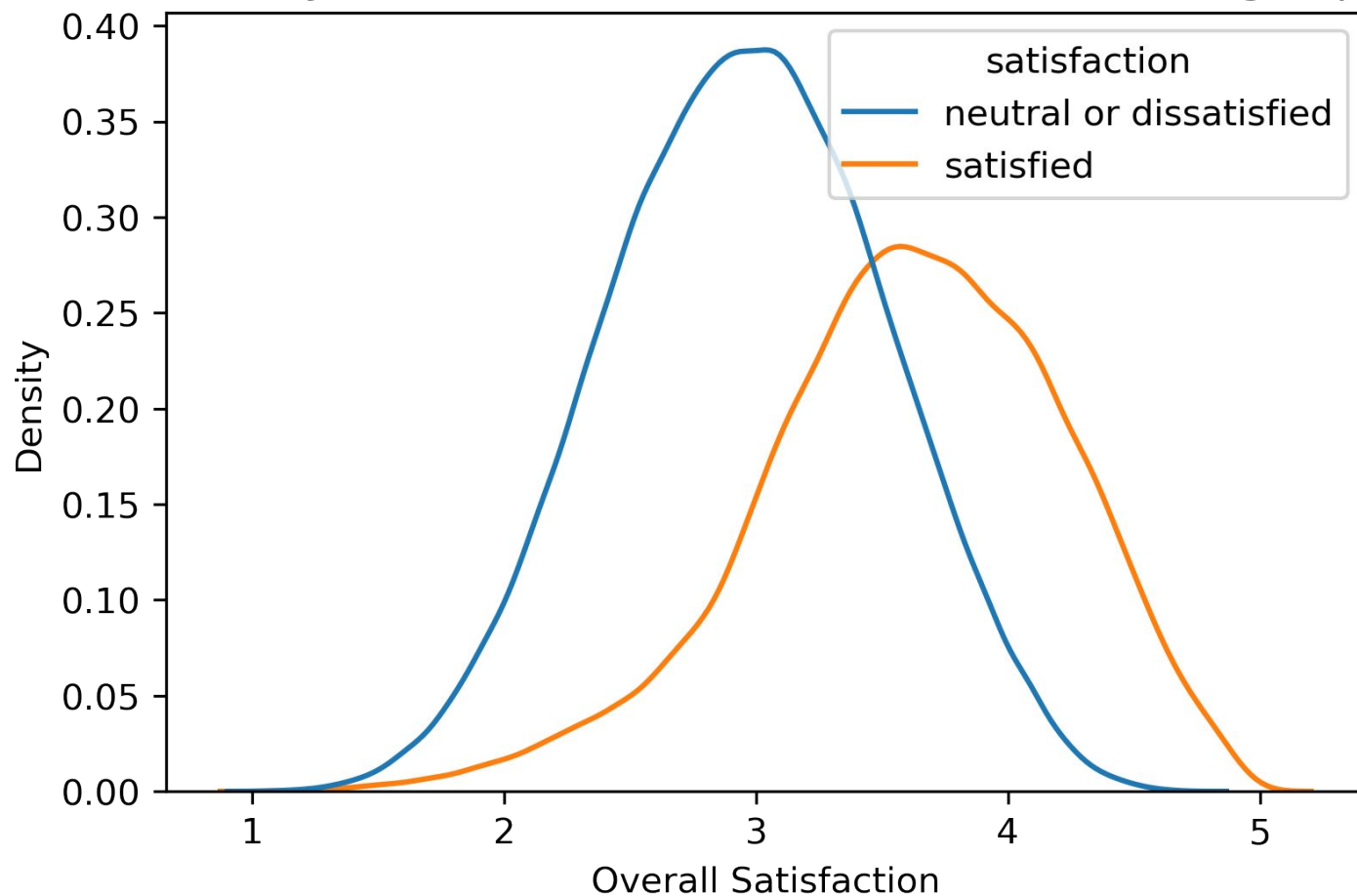
# How does the type of travel (personal vs business) affect passenger satisfaction?
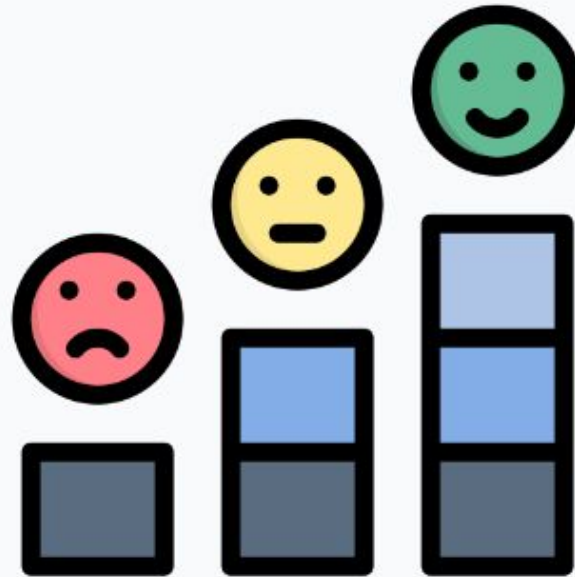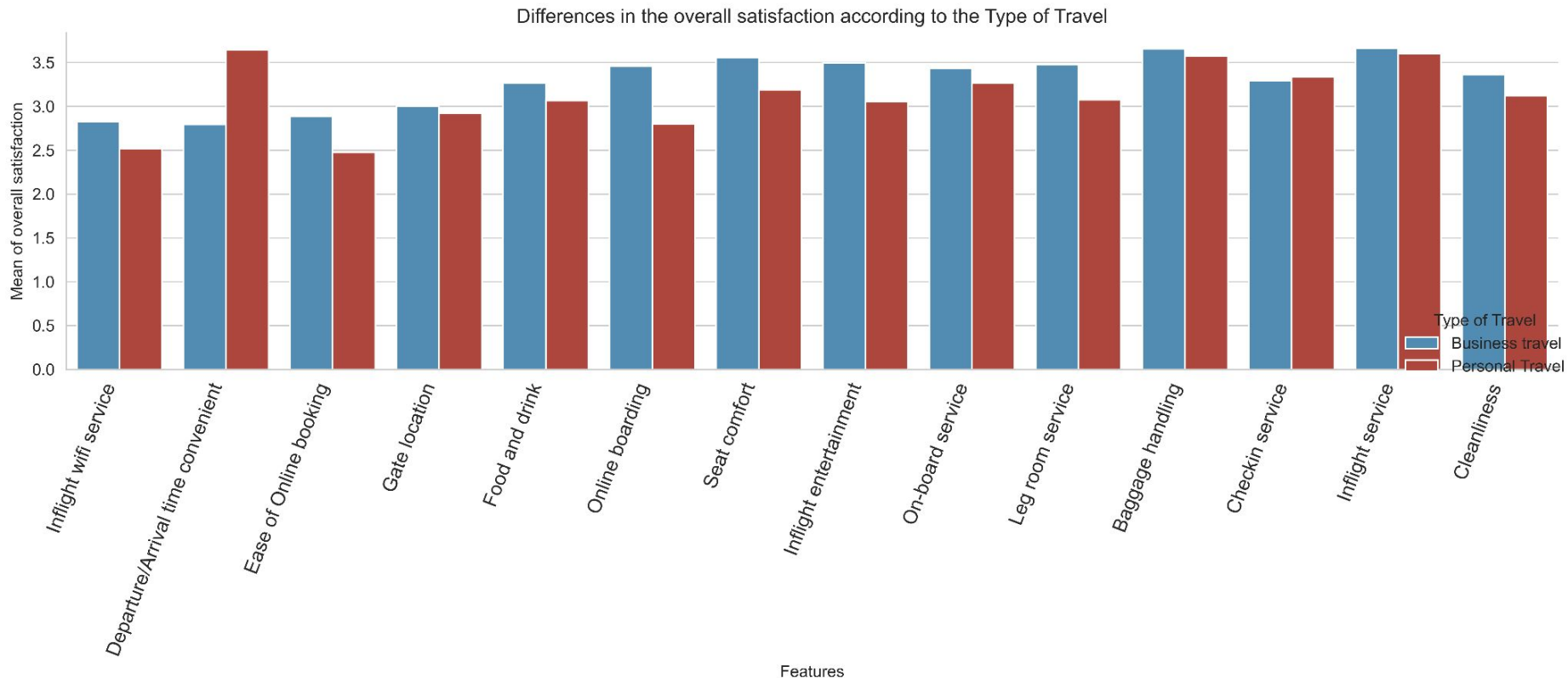
Type of Travel in differenct classes
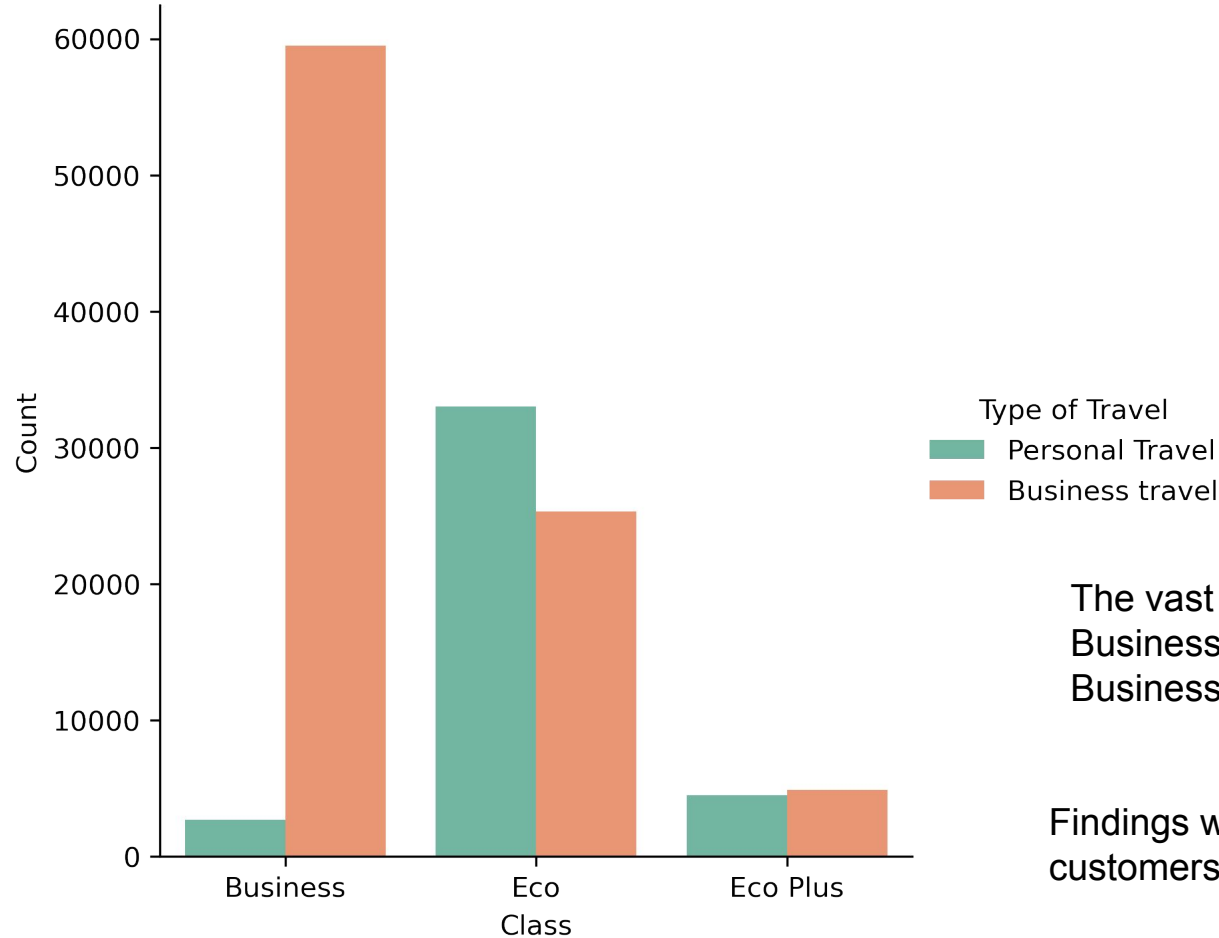
The vast majority of people using the Business class are traveling for Business purposes

Findings will be similar when stratifying customers according to **class** OR **travel type**

# Satisfied vs neutral or dissatisfied, according to Type of Travel



Customers traveling for **Business purposes tend to be more satisfied,** unlike those whose **type is 'Personal', who are almost always dissatisfied.**

Differences in the overall satisfaction according to the Type of Travel
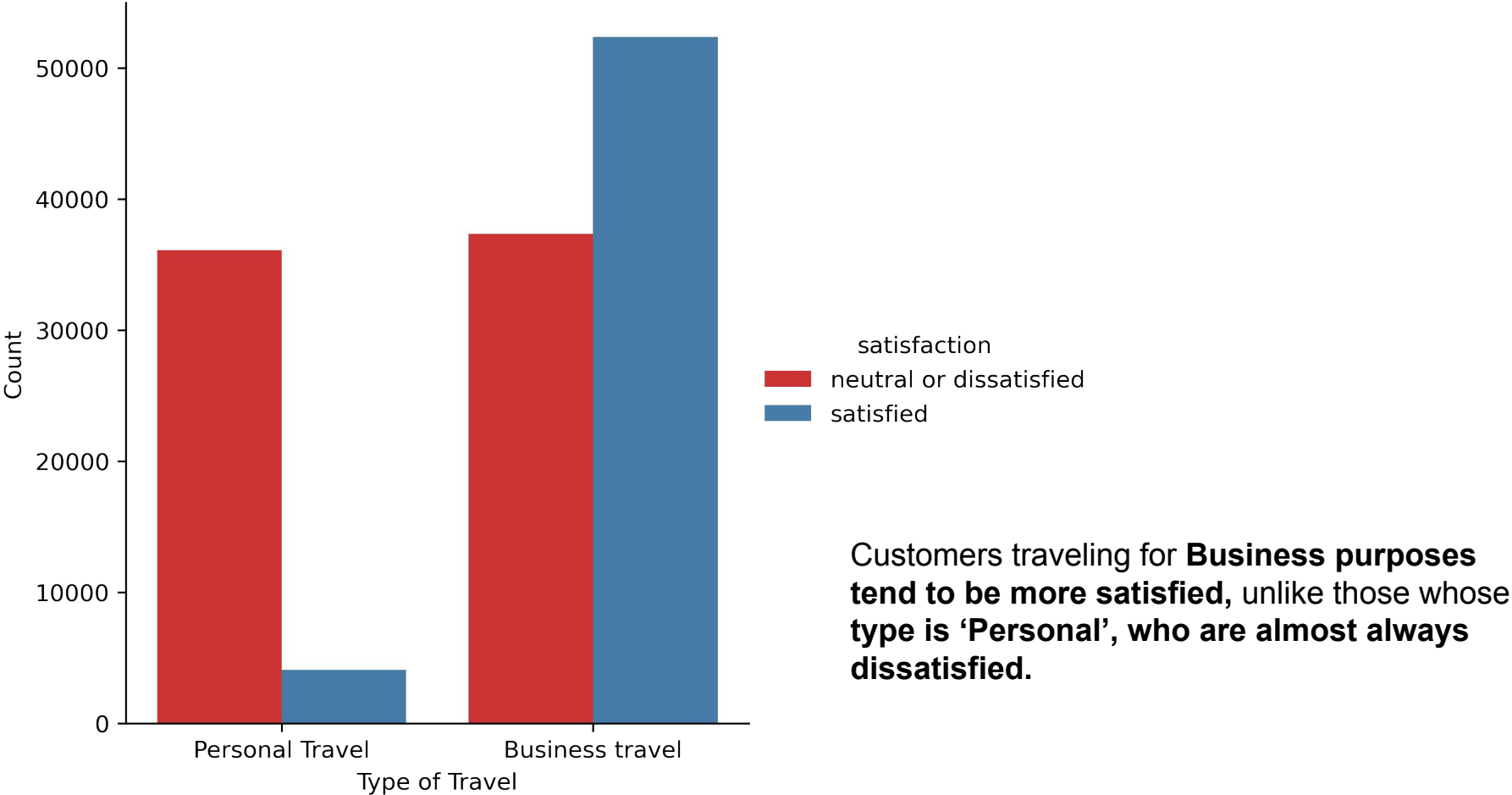
Except for the Departure/Arrival time convenient, customers who travel for business have almost always higher 'overall' satisfaction.
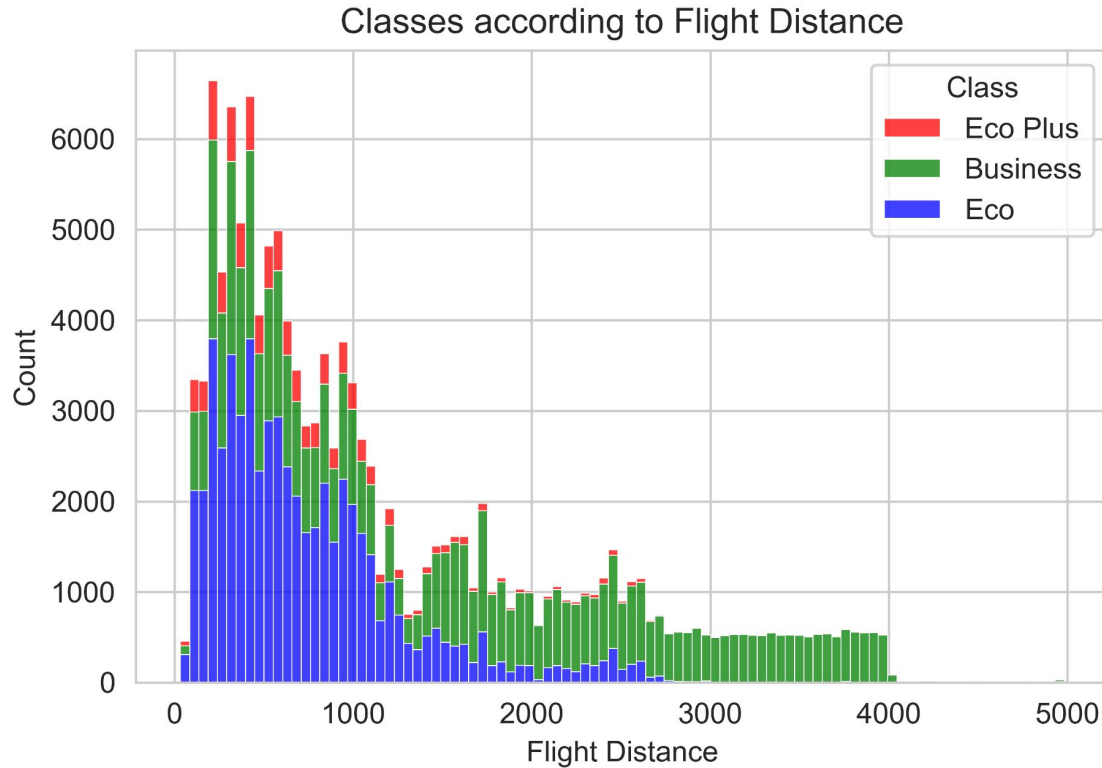
Relationship between **flight distance** and **passenger satisfaction**

Satisfaction according to Flight Distance
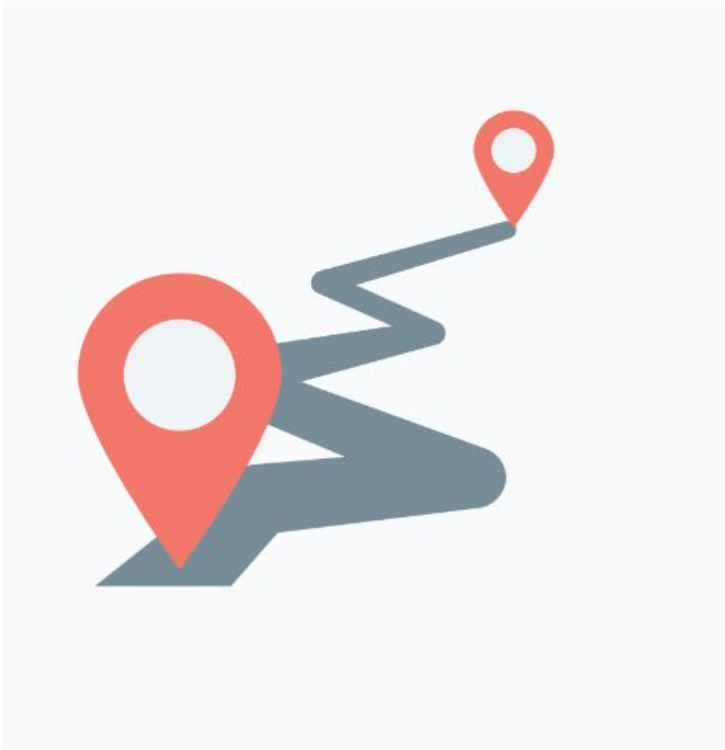
Satisfaction according to Flight Distance
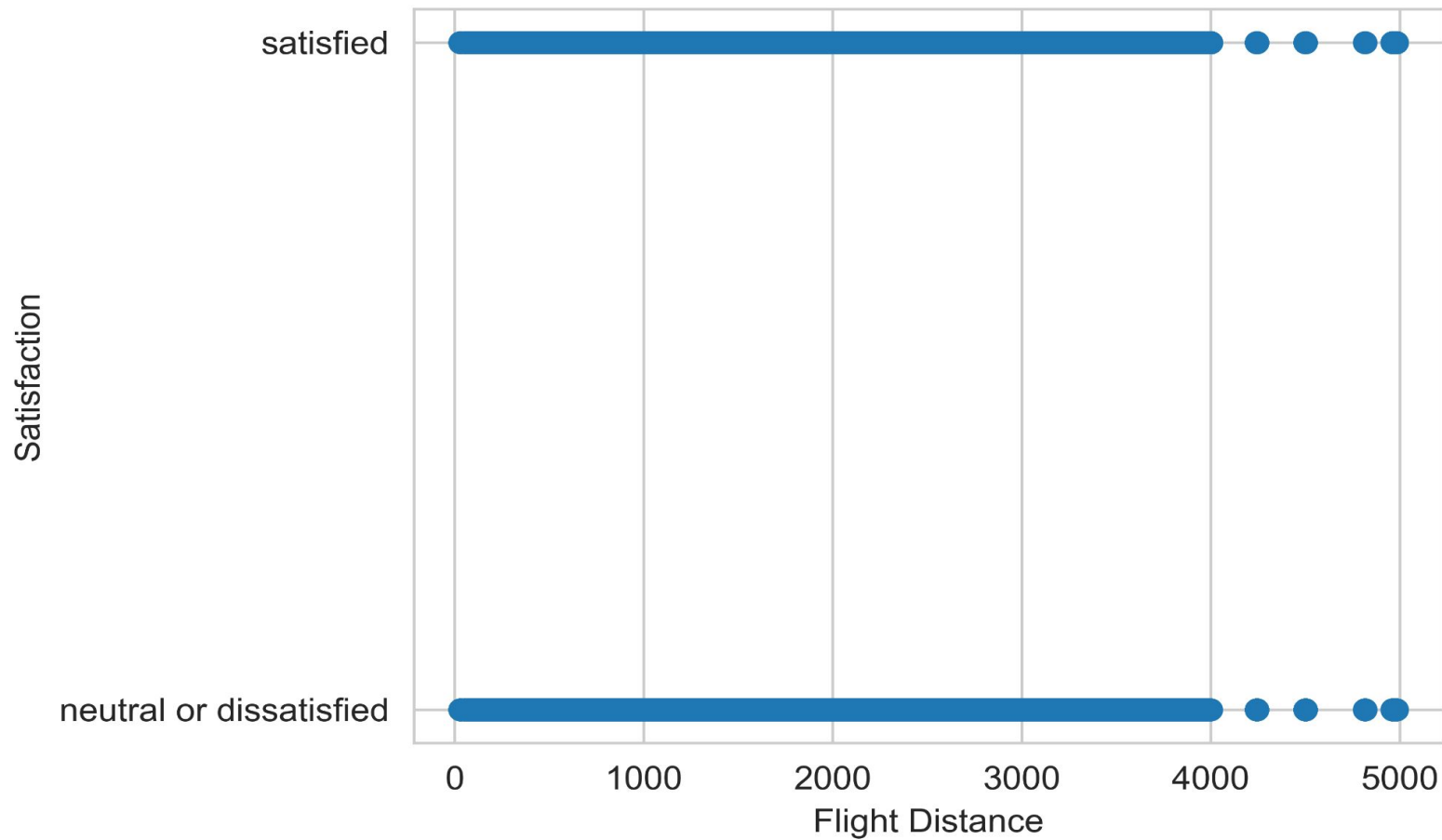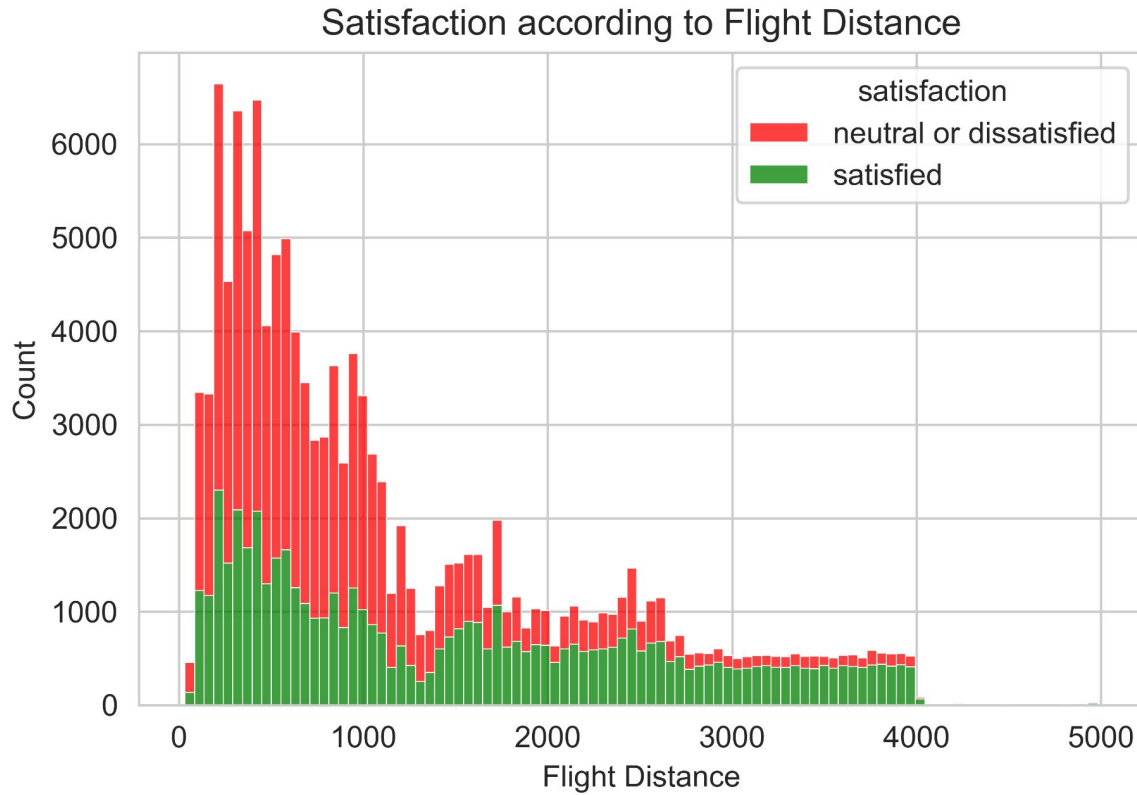
It is **tempting** to say that the longer the flight the more satisfied people are

Classes according to Flight Distance

It turned out that more ( almost all ) people are choosing for the business class for longer flights. This explains the higher satisfaction in longer flights.

We cannot conclude any **direct** relationship between Flight Distance and Satisfaction.
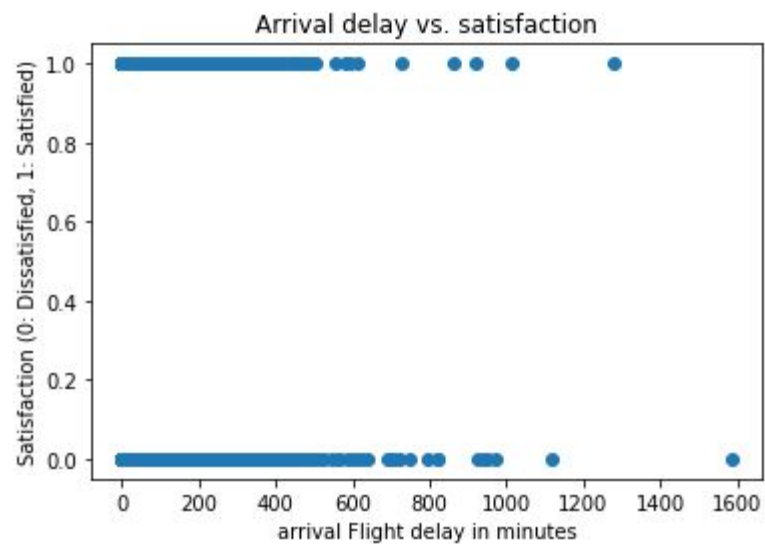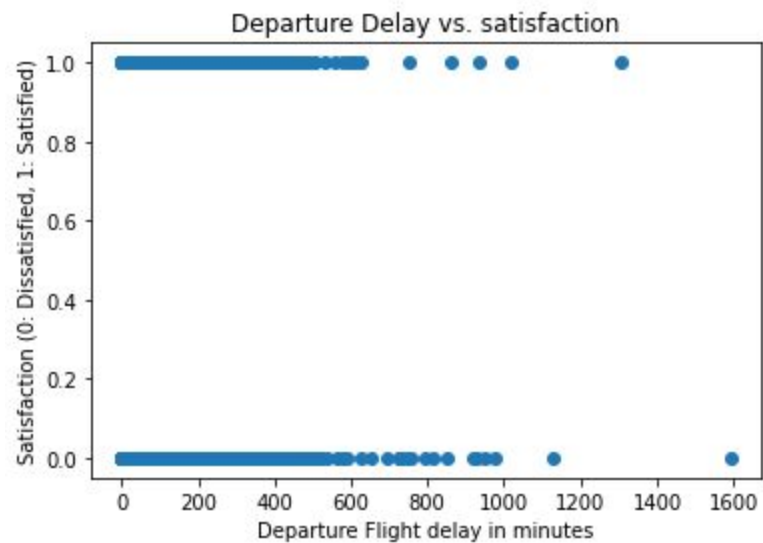
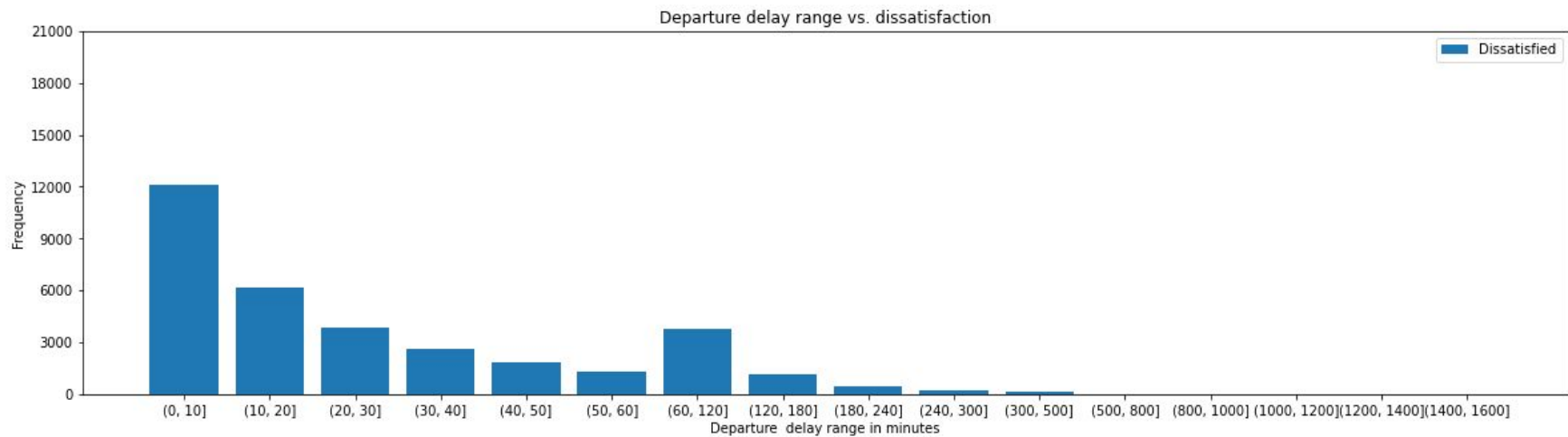# Rama part in answering the research questions
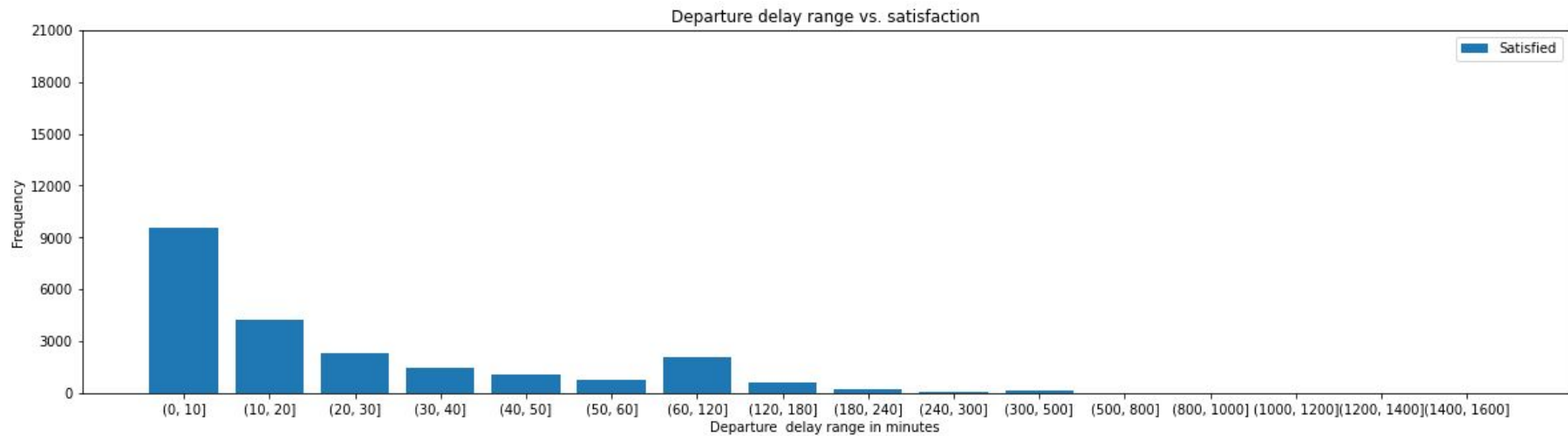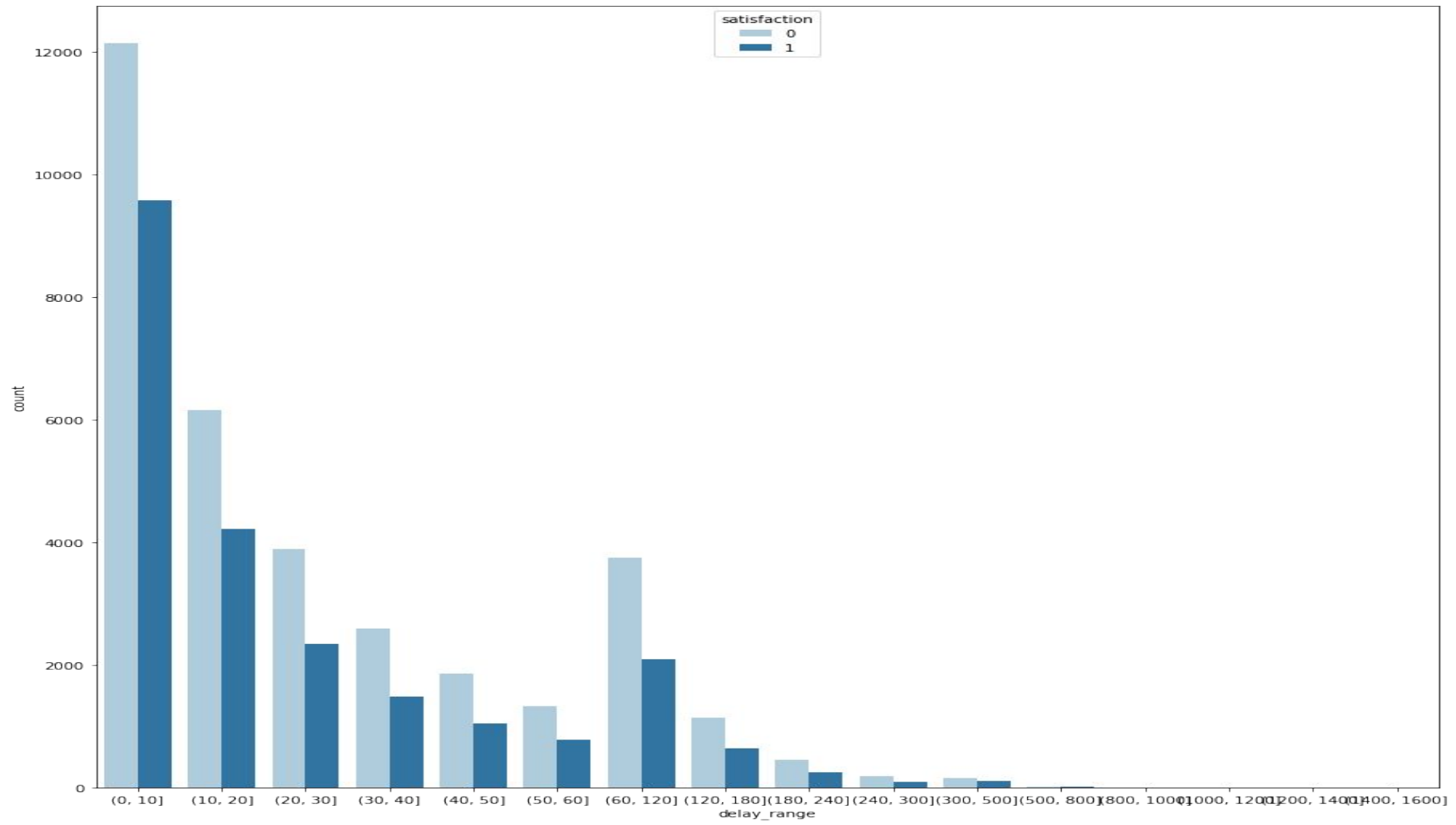## Question 3+4

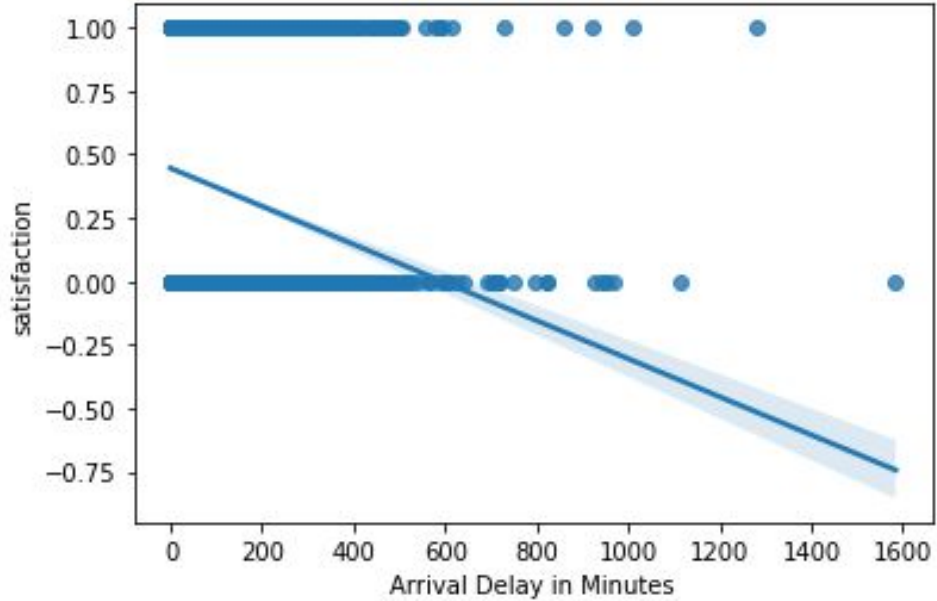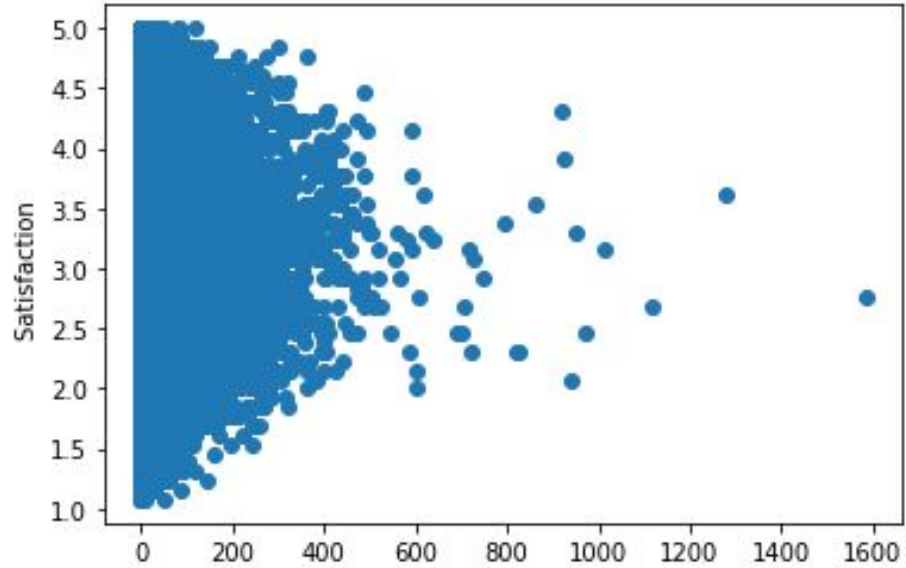## 3. What is the impact of flight delays on passenger satisfaction,

The institution: the more the delay, the more dissatisfied people would be in comparison to the number of satisfied people.

Departure Delay vs. satisfaction

Arrival delay vs. satisfaction

**Departure delay range vs. satisfaction**

Frequency vs. Departure delay range in minutes — Satisfied

**Departure delay range vs. dissatisfaction**

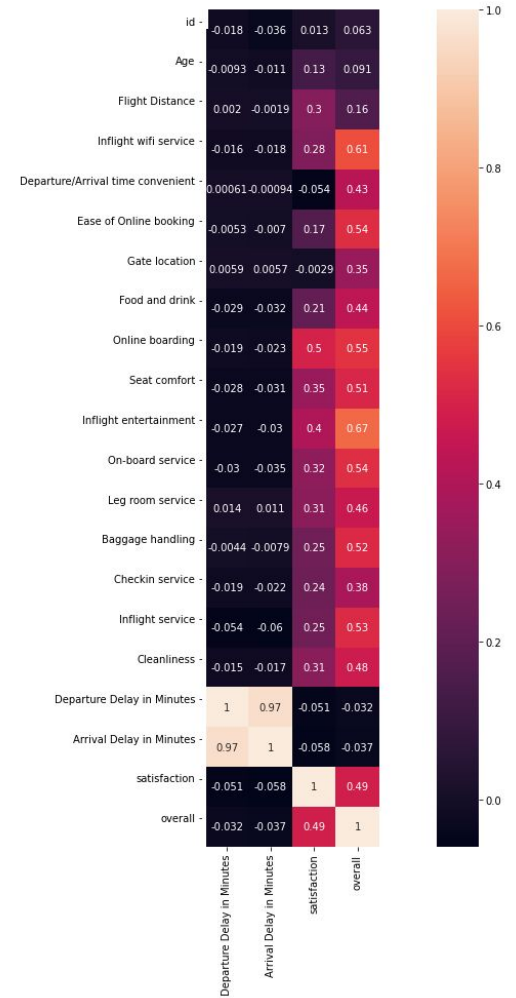Frequency vs. Departure delay range in minutes — Dissatisfied

x= Arrival Delay in Minutes

y= overall Satisfaction

# 3- which other factors are related to delays?



No other factors are related to delays as can be seen from the heat map

# 4.How do different factors affect overall passenger satisfaction?

the most important factors for the all the classes



| | overall |
|---|---|
| overall | 1 |
| Inflight entertainment | 0.67 |
| Inflight wifi service | 0.61 |
| Online boarding | 0.55 |
| On-board service | 0.54 |
| Ease of Online booking | 0.54 |
| Inflight service | 0.53 |
| Baggage handling | 0.52 |
| Seat comfort | 0.51 |
| satisfaction | 0.49 |
| Cleanliness | 0.48 |
| Leg room service | 0.46 |
| Food and drink | 0.44 |
| Departure/Arrival time convenient | 0.43 |
| Checkin service | 0.38 |
| Gate location | 0.35 |
| Flight Distance | 0.16 |
| Age | 0.091 |
| id | 0.063 |
| Departure Delay in Minutes | -0.032 |
| Arrival Delay in Minutes | -0.037 |

# 4.How do different factors affect overall passenger satisfaction?
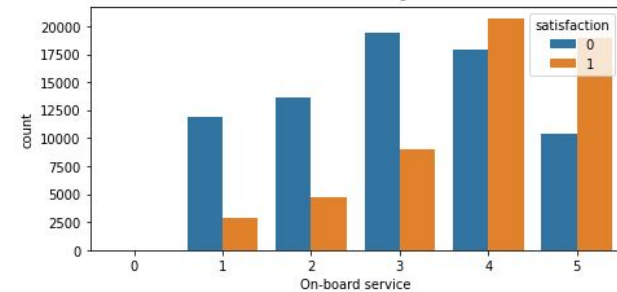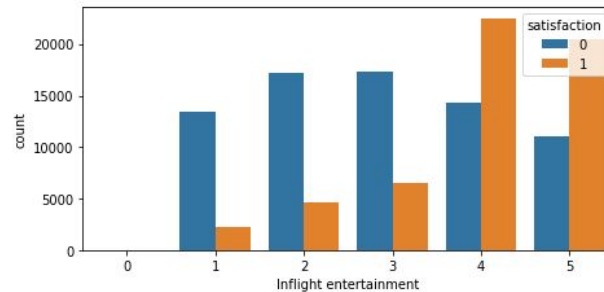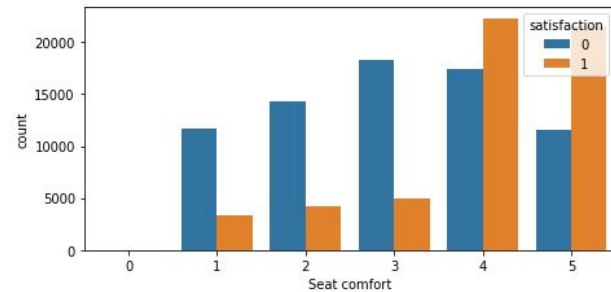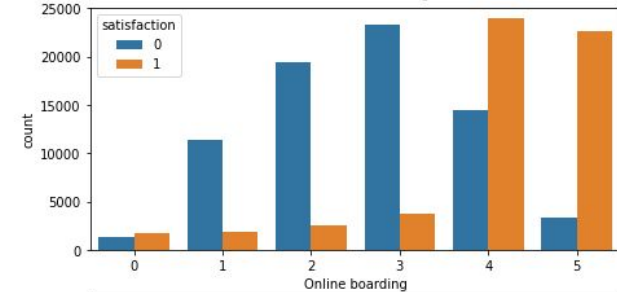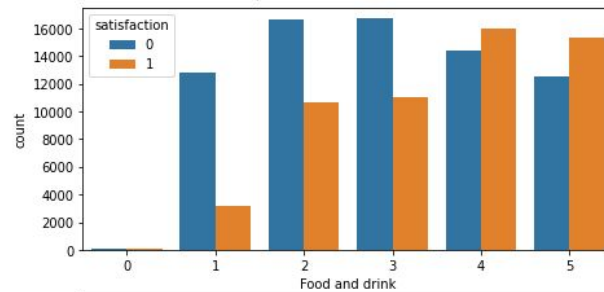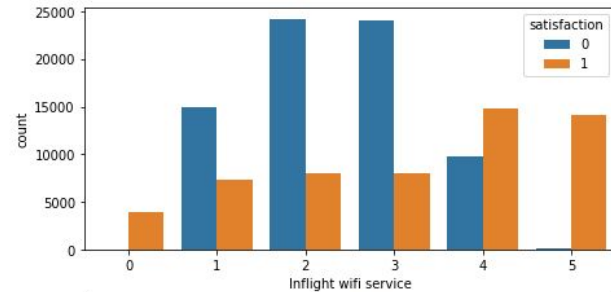
# 4.How do different factors affect overall passenger satisfaction?

# 4.which are most important factors for different classes (business, economy, economy plus)?



the most important factors for the the Business class

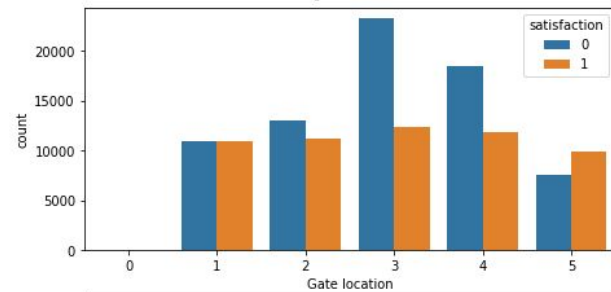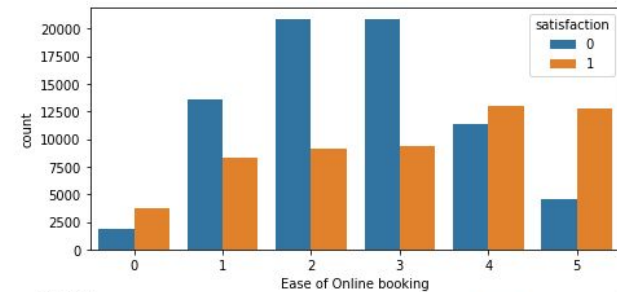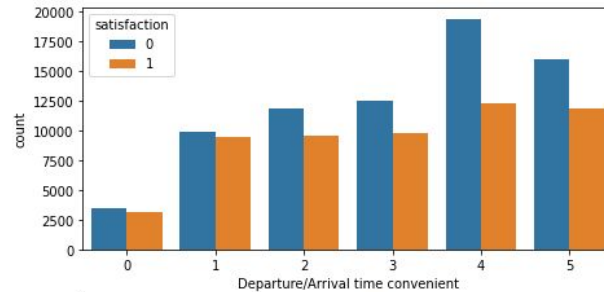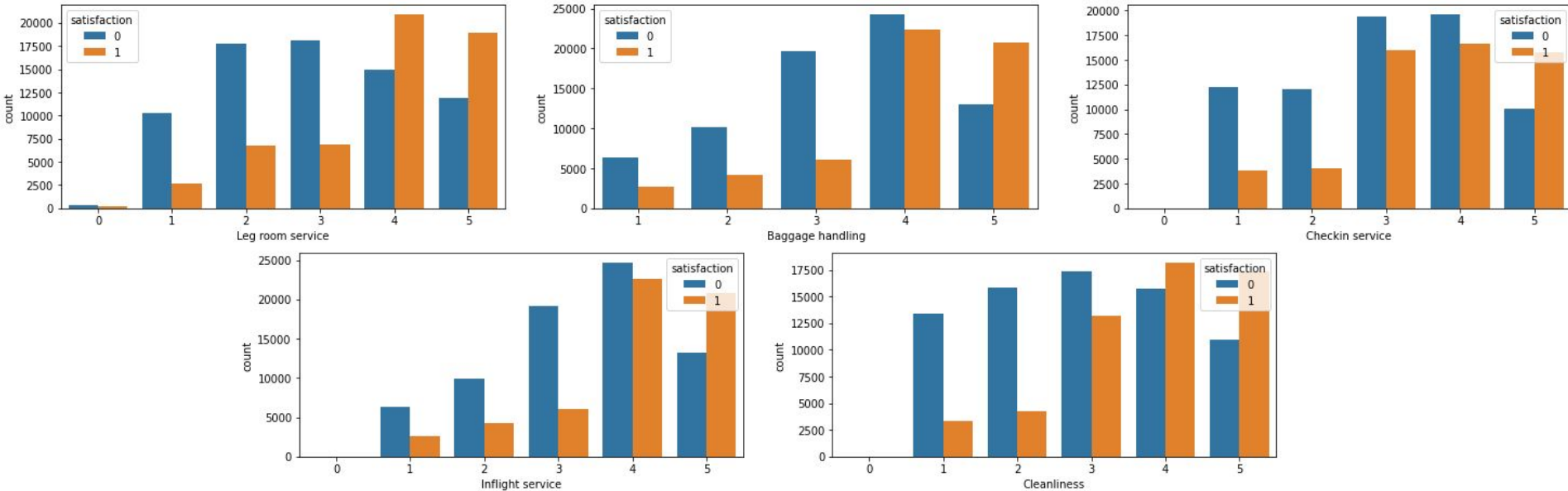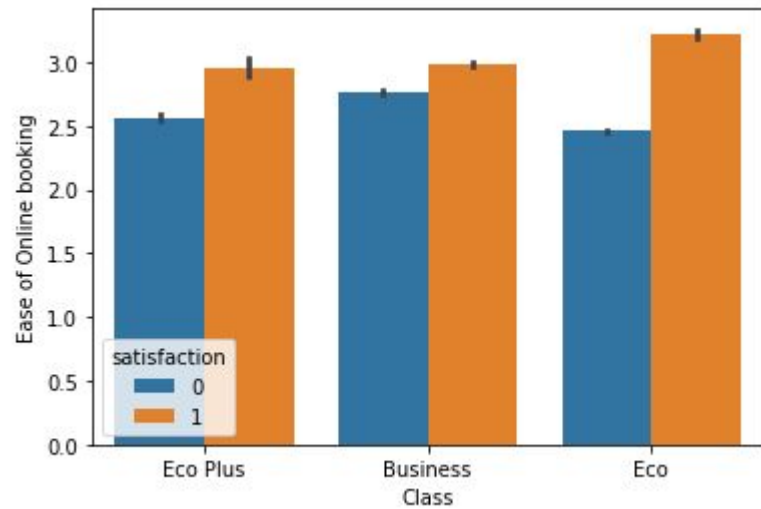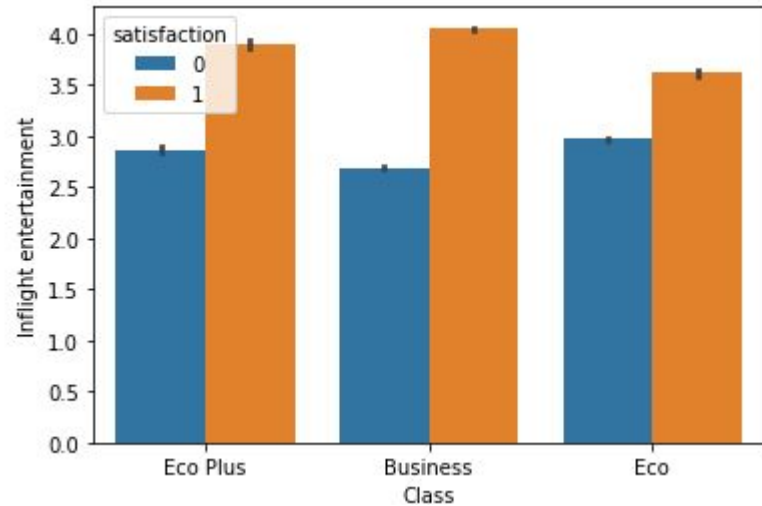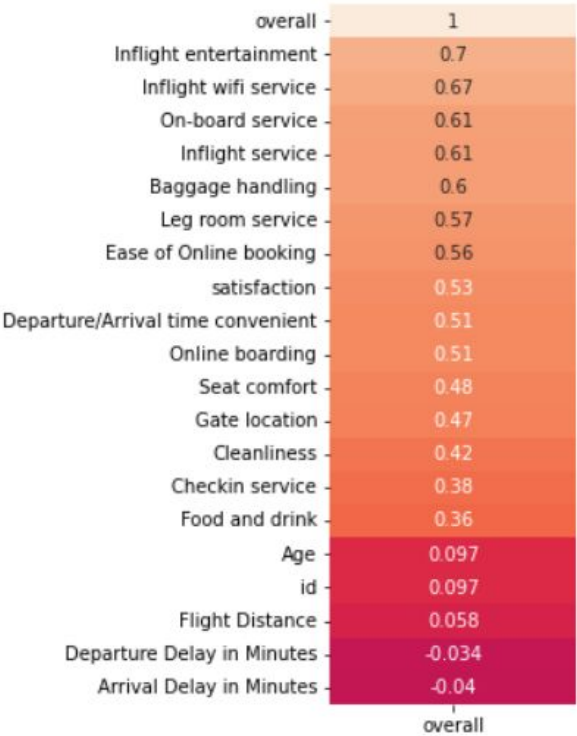| | overall |
|---|---|
| overall | 1 |
| Inflight entertainment | 0.7 |
| Inflight wifi service | 0.67 |
| On-board service | 0.61 |
| Inflight service | 0.61 |
| Baggage handling | 0.6 |
| Leg room service | 0.57 |
| Ease of Online booking | 0.56 |
| satisfaction | 0.53 |
| Departure/Arrival time convenient | 0.51 |
| Online boarding | 0.51 |
| Seat comfort | 0.48 |
| Gate location | 0.47 |
| Cleanliness | 0.42 |
| Checkin service | 0.38 |
| Food and drink | 0.36 |
| Age | 0.097 |
| id | 0.097 |
| Flight Distance | 0.058 |
| Departure Delay in Minutes | -0.034 |
| Arrival Delay in Minutes | -0.04 |

the most important factors for the the Eco class

| | overall |
|---|---|
| overall | 1 |
| Inflight entertainment | 0.6 |
| Inflight wifi service | 0.55 |
| Food and drink | 0.51 |
| Cleanliness | 0.5 |
| Ease of Online booking | 0.5 |
| Online boarding | 0.5 |
| Seat comfort | 0.49 |
| Departure/Arrival time convenient | 0.43 |
| On-board service | 0.42 |
| Inflight service | 0.42 |
| Baggage handling | 0.41 |
| Checkin service | 0.34 |
| Leg room service | 0.3 |
| satisfaction | 0.28 |
| Gate location | 0.21 |
| Age | 0.022 |
| Flight Distance | 0.002 |
| Departure Delay in Minutes | -0.024 |
| id | -0.026 |
| Arrival Delay in Minutes | -0.026 |

the most important factors for the the Eco plus class

| | overall |
|---|---|
| overall | 1 |
| Inflight entertainment | 0.65 |
| Inflight wifi service | 0.62 |
| Food and drink | 0.54 |
| Online boarding | 0.54 |
| Cleanliness | 0.54 |
| Seat comfort | 0.52 |
| Ease of Online booking | 0.49 |
| Departure/Arrival time convenient | 0.47 |
| On-board service | 0.4 |
| Baggage handling | 0.37 |
| Inflight service | 0.37 |
| Checkin service | 0.33 |
| satisfaction | 0.31 |
| Gate location | 0.3 |
| Leg room service | 0.3 |
| Age | 0.023 |
| Flight Distance | -0.034 |
| Departure Delay in Minutes | -0.037 |
| Arrival Delay in Minutes | -0.043 |
| id | -0.081 |

number of satisfied and unsatisfied people by Class Type
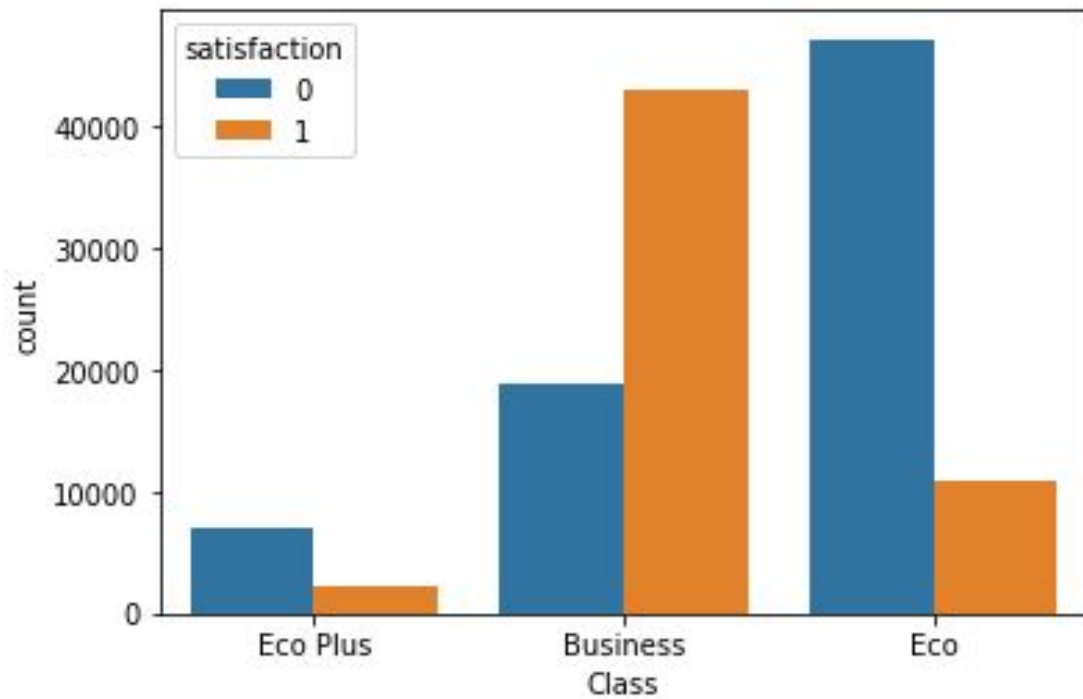
# Zehra's part in answering the research question

5. Are there differences in passenger satisfaction between loyal and disloyal customers, and if so, which factors are most important in driving these differences?

Barchart of the Number of the Customer Type

The number of loyal customers is about 4 times greater than the number of customers who are disloyal.
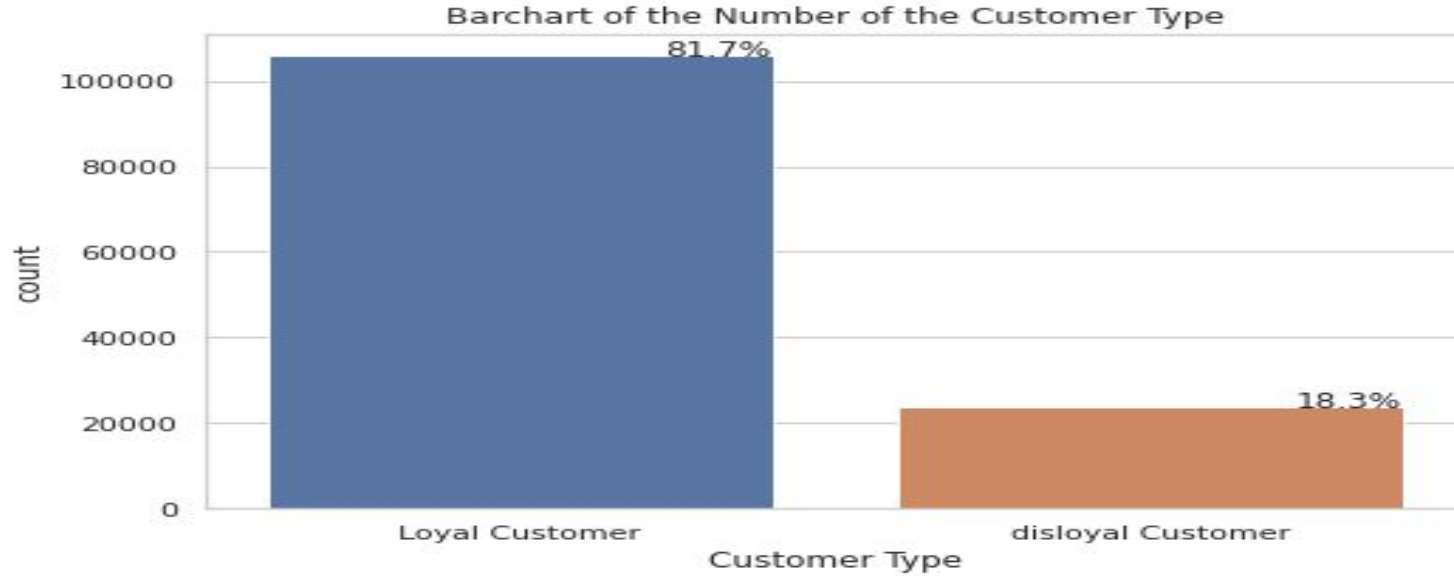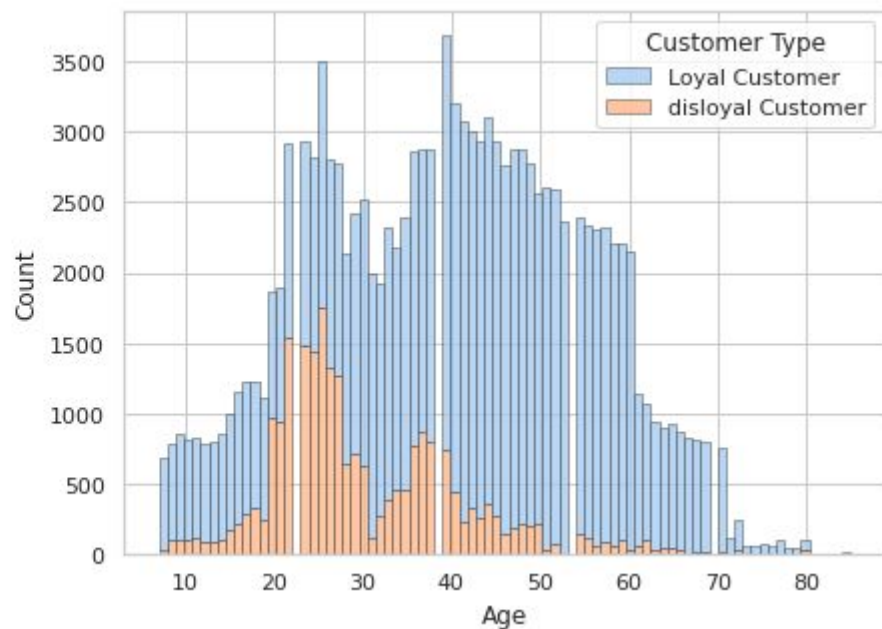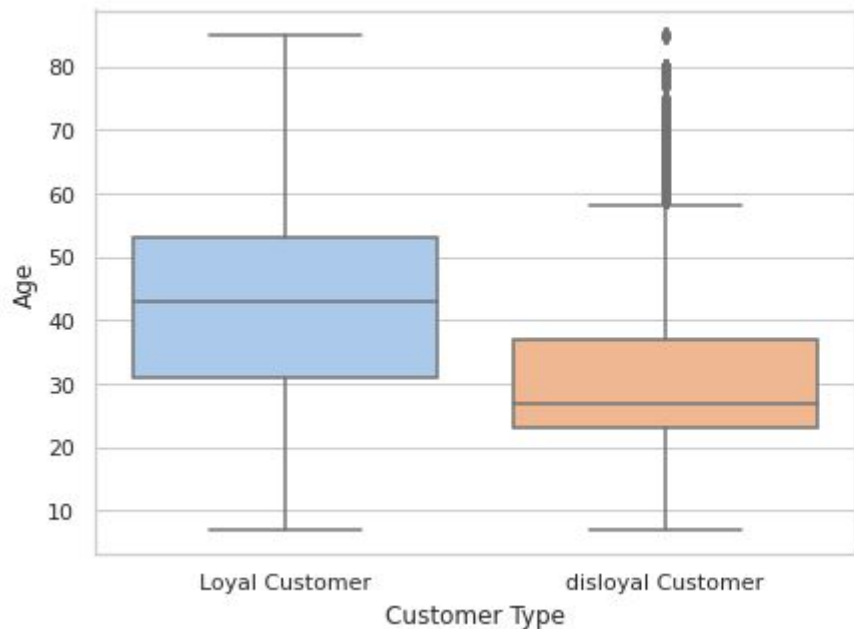
From this box diagram, we can conclude that most of the airline's loyal customers are between the ages of 30 and 50 (their average age is slightly over 40). The age range of disloyal customers is slightly smaller (from 25 to 40 years old, on average - a little less than 30).

Relation between Customer Type and Class
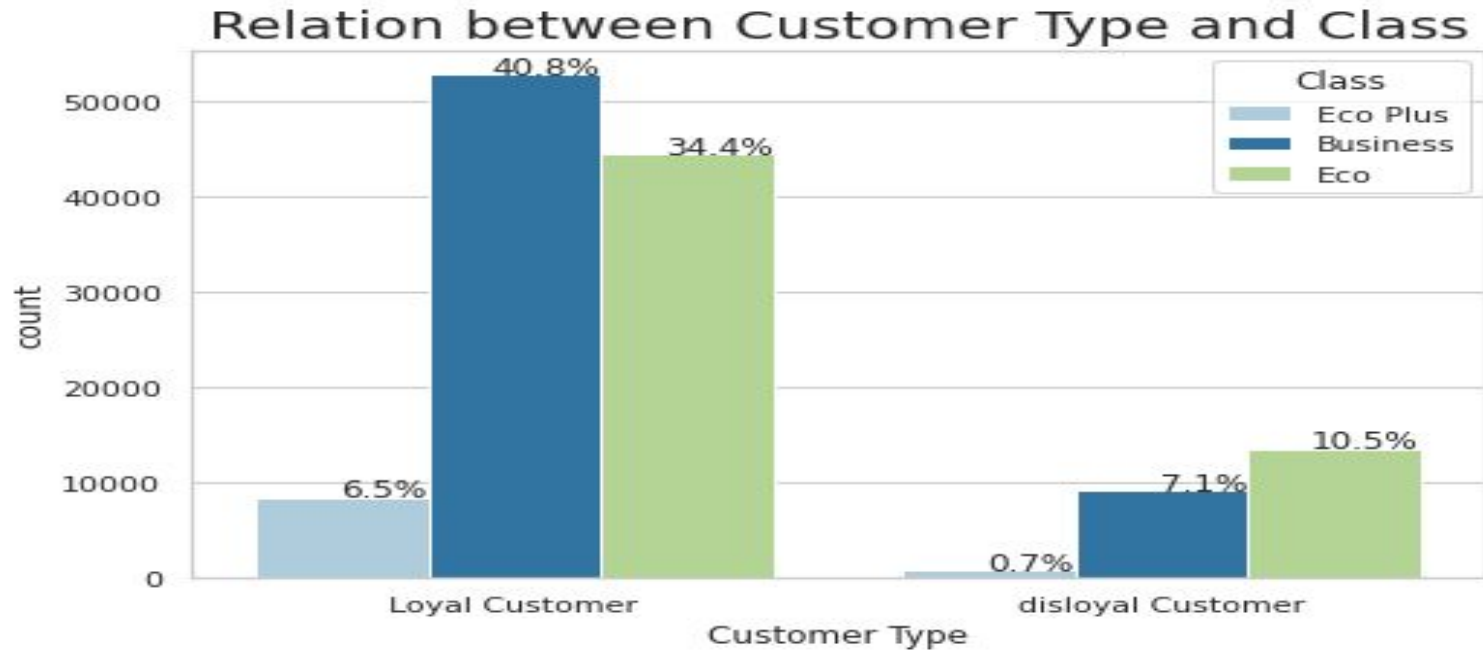
The most number of the loyal customers prefered the business class. But the most number of disloyal customer prefer eco class.

**Barchart of Customer Type by Customer Satisfaction**

42.6%

39.1%

13.9%

4.4%

satisfaction
- neutral or dissatisfied
- satisfied

count

50000

40000

30000

20000

10000

0

Loyal Customer

disloyal Customer
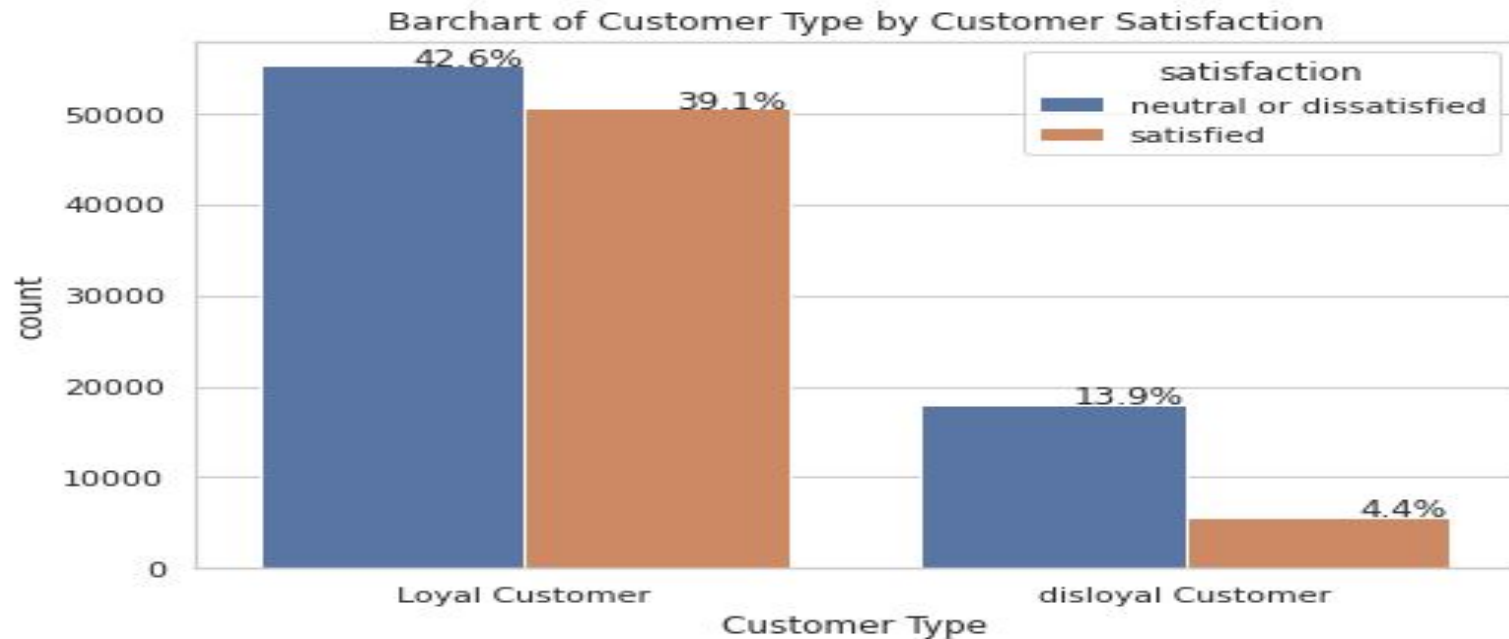
Customer Type

In the data set, there are a lot more loyal customers than disloyal customers. Through the normalized bar chart of satisfaction ratings by customer type, there is a higher proportion of neutral or dissatisfied customers among disloyal customers than there are among loyal customers. This chart provides us with the initial insight that perhaps customer loyalty may affect satisfaction ratings.

the most important factors for the Loyal Customer

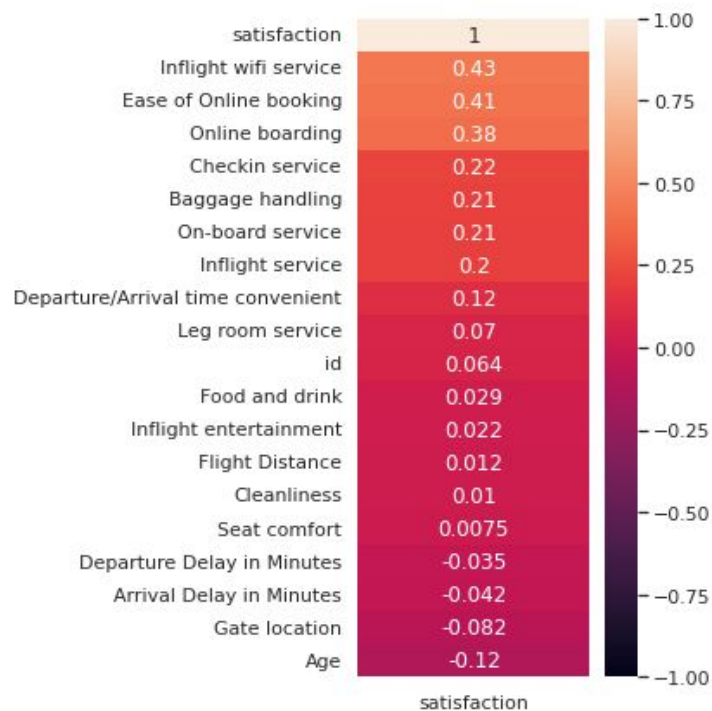| satisfaction | 1 |
|---|---|
| Online boarding | 0.5 |
| Inflight entertainment | 0.46 |
| Seat comfort | 0.4 |
| Cleanliness | 0.36 |
| Leg room service | 0.36 |
| On-board service | 0.34 |
| Flight Distance | 0.29 |
| Baggage handling | 0.27 |
| Inflight service | 0.26 |
| Inflight wifi service | 0.26 |
| Food and drink | 0.24 |
| Checkin service | 0.24 |
| Ease of Online booking | 0.12 |
| Age | 0.12 |
| Gate location | 0.011 |
| id | 0.0038 |
| Departure Delay in Minutes | -0.054 |
| Arrival Delay in Minutes | -0.062 |
| Departure/Arrival time convenient | -0.14 |

satisfaction

the most important factors for the Disloyal Customer

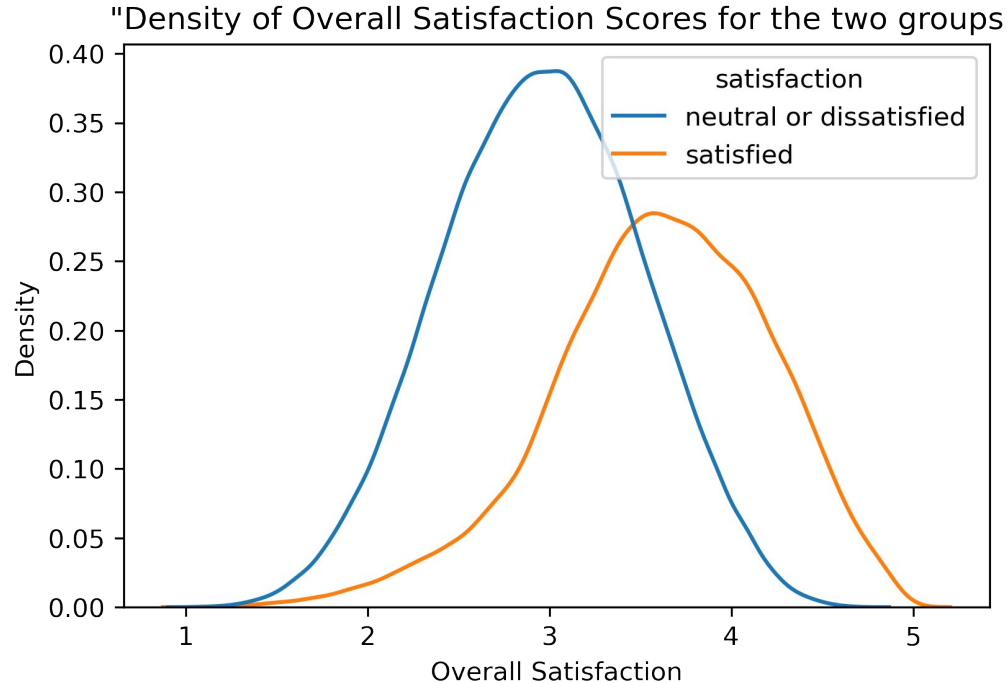| satisfaction | 1 |
|---|---|
| Inflight wifi service | 0.43 |
| Ease of Online booking | 0.41 |
| Online boarding | 0.38 |
| Checkin service | 0.22 |
| Baggage handling | 0.21 |
| On-board service | 0.21 |
| Inflight service | 0.2 |
| Departure/Arrival time convenient | 0.12 |
| Leg room service | 0.07 |
| id | 0.064 |
| Food and drink | 0.029 |
| Inflight entertainment | 0.022 |
| Flight Distance | 0.012 |
| Cleanliness | 0.01 |
| Seat comfort | 0.0075 |
| Departure Delay in Minutes | -0.035 |
| Arrival Delay in Minutes | -0.042 |
| Gate location | -0.082 |
| Age | -0.12 |

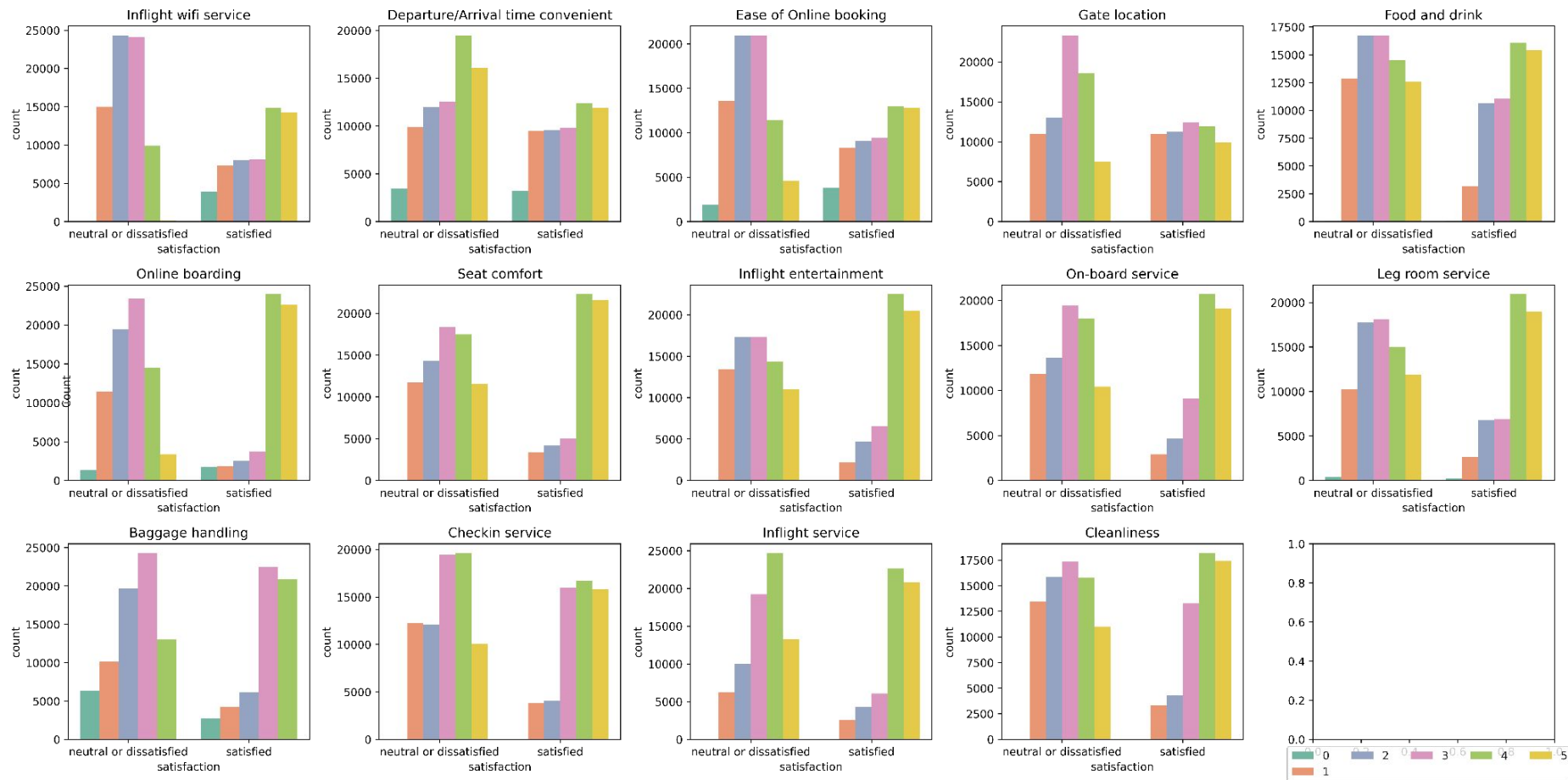satisfaction

# Strange findings:

1- score very low on overall satisfaction but still be satisfied
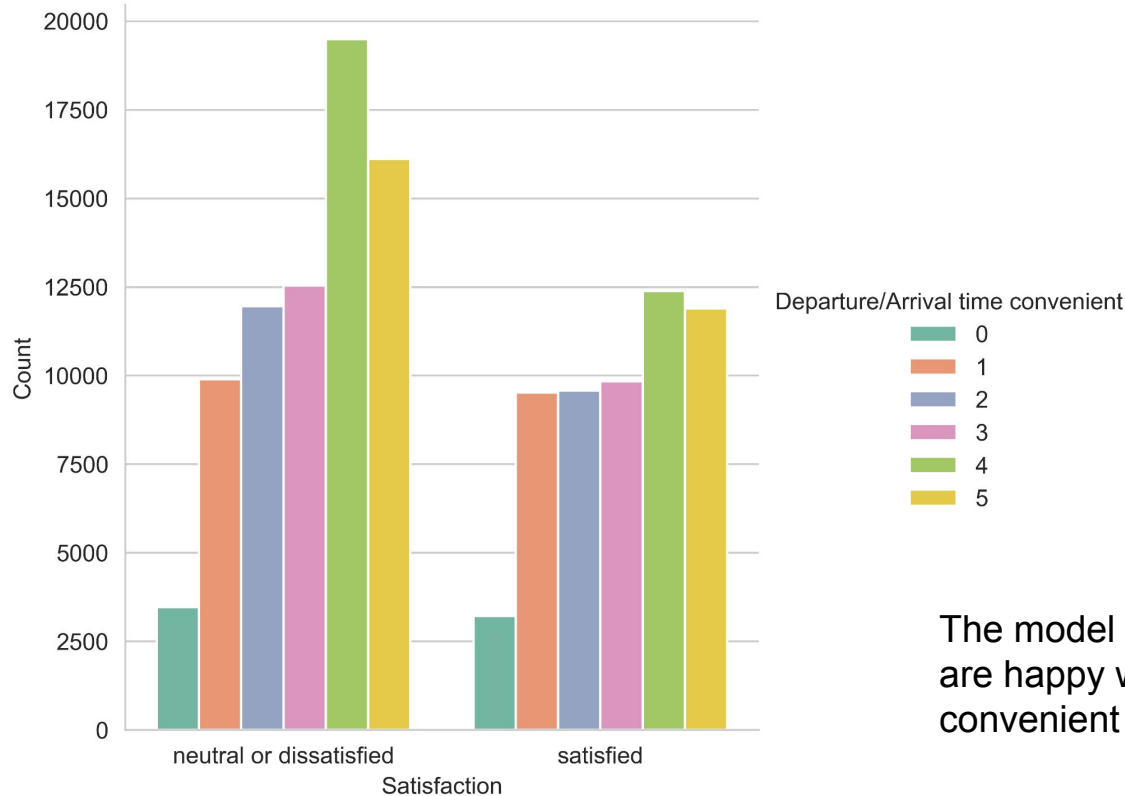
2- departure/arrival time convenient

- How can it be that a good percentages from the satisfied customers score very low in the overall score ? ( see the density graph)
- Same applies for people who score high on the overall satisfaction but still unsatisfied.



"Density of Overall Satisfaction Scores for the two groups

- How can it be that people who score higher in the Departure/Arrival time convenient tend to be less satisfied.
- If we want to **predict customers satisfaction** : the higher they score on Departure/Arrival time the higher the chance that they would be unsatisfied.



The model will classify you as unsatisfied if you are happy with the Departure/Arrival time convenient