PHYTON FOR GOOGLE COLAB

# DATA SCIENCE

## HOTEL SATISFACTION DATA ANALYSIS

**DIGITAL SKILL FAIR 27.0**

14 - 24 MAY 2024


**by Sopia Ranty**

# OUR SERVICE

IMPORT DATA

HEAD DATA

INFO DATA

DESCRIPTIVE STATISTICS DATA

DATA ANALYSIS RESULTS

CONCLUSION

# IMPORT DATA

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
# Import data
df = pd.read_csv('Europe Hotel Booking Satisfaction Score.csv')
df
```

| | id | Gender | Age | purpose_of_travel | Type of Travel | Type Of Booking | Hotel wifi service | Departure/ Arrival convenience | Ease of Online booking | Hotel location | Food and drink | co |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 70172 | Male | 13 | aviation | Personal Travel | Not defined | 3 | 4 | 3 | 1 | 5 | |
| 1 | 5047 | Male | 25 | tourism | Group Travel | Group bookings | 3 | 2 | 3 | 3 | 1 | |
| 2 | 110028 | Female | 26 | tourism | Group Travel | Group bookings | 2 | 2 | 2 | 2 | 5 | |
| 3 | 24026 | Female | 25 | tourism | Group Travel | Group bookings | 2 | 5 | 5 | 5 | 2 | |
| 4 | 119299 | Male | 61 | aviation | Group Travel | Group bookings | 3 | 3 | 3 | 3 | 4 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 103899 | 94171 | Female | 23 | business | Group | Individual/ | 2 | 1 | 2 | 3 | 2 | |

## IMPORT LIBRARIES AND IMPORT DATA

The code imports the necessary libraries for data analysis and visualization, then imports the data from a CSV file into a DataFrame using pandas.

# DISPLAY HEAD DATA

The output of the print head displays the first five entries of the DataFrame. Each entry represents a hotel guest. The information presented includes details such as guest id, gender, age, purpose of travel, type of travel, type of booking, hotel service scores (such as wifi service, cleanliness, food and drink), and guest satisfaction level.

```
# Lihat sekilas data
print(df.head())
```

```
      id  Gender  Age purpose_of_travel   Type of Travel Type Of Booking  \
0  70172    Male   13          aviation  Personal Travel     Not defined
1   5047    Male   25           tourism     Group Travel  Group bookings
2 110028  Female   26           tourism     Group Travel  Group bookings
3  24026  Female   25           tourism     Group Travel  Group bookings
4 119299    Male   61          aviation     Group Travel  Group bookings

   Hotel wifi service  Departure/Arrival  convenience  Ease of Online booking  \
0                   3                                4                       3
1                   3                                2                       3
2                   2                                2                       2
3                   2                                5                       5
4                   3                                3                       3

   Hotel location  Food and drink  Stay comfort  Common Room entertainment  \
0               1               5             5                          5
1               3               1             1                          1
2               2               5             5                          5
3               5               2             2                          2
4               3               4             5                          3

   Checkin/Checkout service  Other service  Cleanliness  \
0                         4              5            5
1                         1              4            1
2                         4              4            5
3                         1              4            2
4                         3              3            3

             satisfaction
0  neutral or dissatisfied
1  neutral or dissatisfied
2                satisfied
3  neutral or dissatisfied
4                satisfied
```

# DISPLAY INFO DATA

The print info output gives a brief overview of the DataFrame. It shows the number of rows and columns, data types of each column, and memory usage. This summary indicates that there are 103,904 entries with 17 columns. The data types include integers and objects (strings), and there are no missing values in any column.

```
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 103904 entries, 0 to 103903
Data columns (total 17 columns):
 #   Column                          Non-Null Count   Dtype 
---  ------                          --------------   ----- 
 0   id                              103904 non-null  int64 
 1   Gender                          103904 non-null  object
 2   Age                             103904 non-null  int64 
 3   purpose_of_travel               103904 non-null  object
 4   Type of Travel                  103904 non-null  object
 5   Type Of Booking                 103904 non-null  object
 6   Hotel wifi service              103904 non-null  int64 
 7   Departure/Arrival  convenience  103904 non-null  int64 
 8   Ease of Online booking          103904 non-null  int64 
 9   Hotel location                  103904 non-null  int64 
 10  Food and drink                  103904 non-null  int64 
 11  Stay comfort                    103904 non-null  int64 
 12  Common Room entertainment       103904 non-null  int64 
 13  Checkin/Checkout service        103904 non-null  int64 
 14  Other service                   103904 non-null  int64 
 15  Cleanliness                     103904 non-null  int64 
 16  satisfaction                    103904 non-null  object
dtypes: int64(12), object(5)
memory usage: 13.5+ MB
None
```

# DESCRIPTIVE STATISTICS DATA

The output presents descriptive statistics for each numeric column in the DataFrame. These statistics provide insights into the distribution and spread of values within each feature. For instance, the average age of guests is approximately 39.38 years, with a standard deviation of around 15.11, indicating a significant variation in guest ages. The minimum age for guests is 7 years, while the maximum is 85 years. Quartiles provide information about the distribution of data within each column, with the second quartile (median) indicating the middle value of the data. Thus, this output gives an overview of the general profile of the observed data.

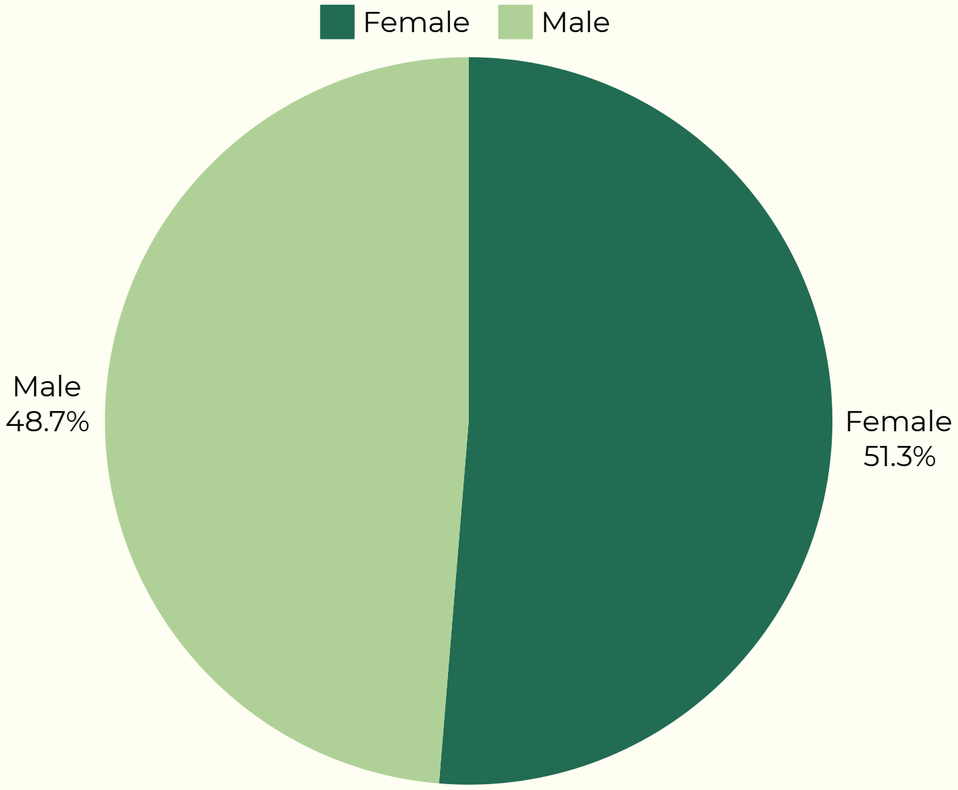
```
# Statistik deskriptif
print(df.describe())
```

```
                  id            Age  Hotel wifi service  \
count  103904.000000  103904.000000       103904.000000
mean    64924.210502      39.379706            2.729683
std     37463.812252      15.114964            1.327829
min         1.000000       7.000000            0.000000
25%     32533.750000      27.000000            2.000000
50%     64856.500000      40.000000            3.000000
75%     97368.250000      51.000000            4.000000
max    129880.000000      85.000000            5.000000

       Departure/Arrival  convenience  Ease of Online booking  Hotel location  \
count                     103904.000000           103904.000000   103904.000000
mean                           3.060296                2.756901        2.976883
std                            1.525075                1.398929        1.277621
min                            0.000000                0.000000        0.000000
25%                            2.000000                2.000000        2.000000
50%                            3.000000                3.000000        3.000000
75%                            4.000000                4.000000        4.000000
max                            5.000000                5.000000        5.000000

       Food and drink   Stay comfort  Common Room entertainment  \
count   103904.000000  103904.000000              103904.000000
mean         3.202129       3.439396                   3.358158
std          1.329533       1.319088                   1.332991
min          0.000000       0.000000                   0.000000
25%          2.000000       2.000000                   2.000000
50%          3.000000       4.000000                   4.000000
75%          4.000000       5.000000                   4.000000
max          5.000000       5.000000                   5.000000

       Checkin/Checkout service  Other service    Cleanliness
count             103904.000000  103904.000000  103904.000000
mean                   3.304290       3.640428       3.286351
std                    1.265396       1.175663       1.312273
min                    0.000000       0.000000       0.000000
25%                    3.000000       3.000000       2.000000
50%                    3.000000       4.000000       3.000000
75%                    4.000000       5.000000       4.000000
max                    5.000000       5.000000       5.000000
```

# DATA ANALYSIS RESULTS

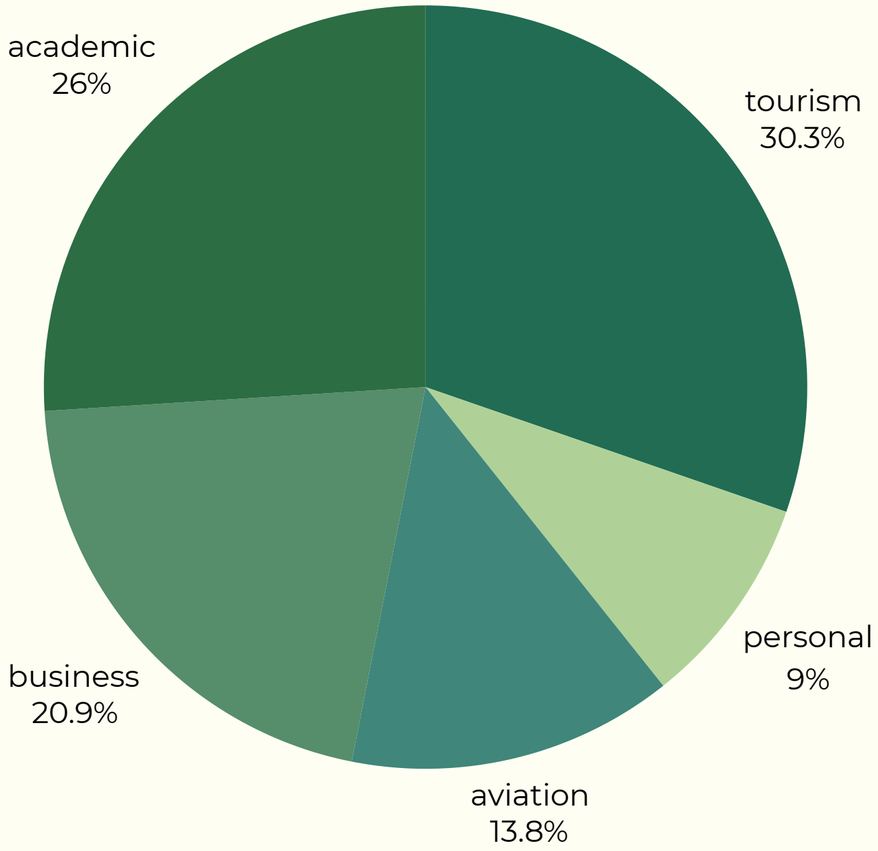
# CUSTOMER GENDER ANALYSIS

The analysis of hotel customer satisfaction data reveals that 51.3% of the customers are female, while male customers make up 48.7%. This suggests that the majority of hotel guests are female, with a small difference in percentage between the two gender groups. This information can help hotel management in creating service strategies that cater better to the needs and preferences of customers from different gender backgrounds.



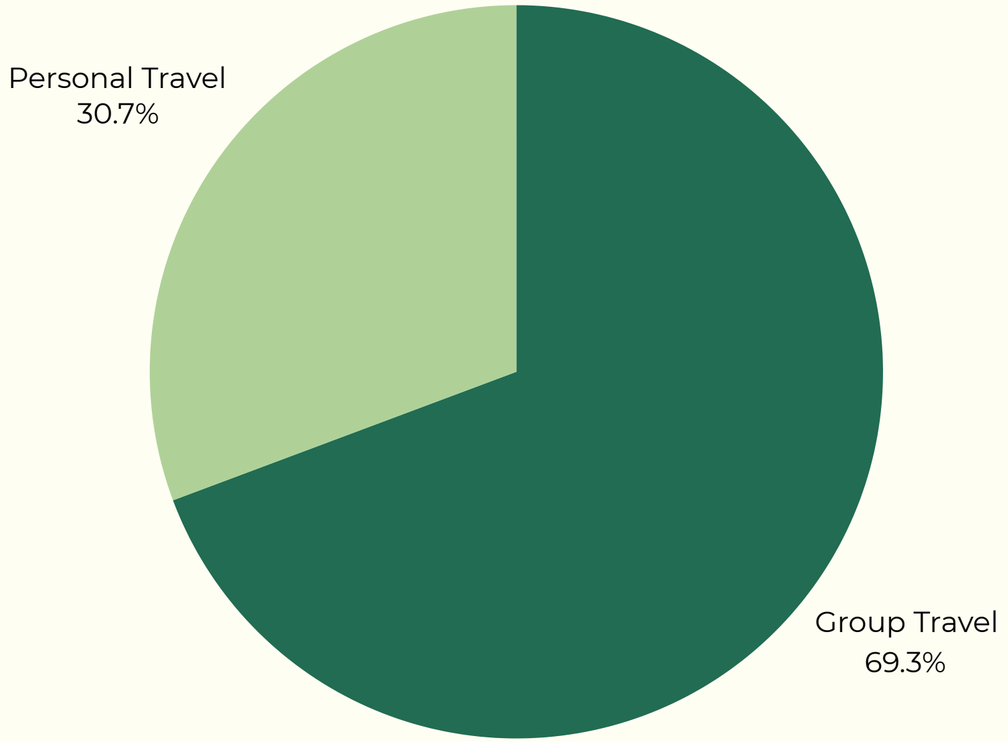
# PURPOSE OF TRAVEL ANALYSIS

Based on the analysis of customer satisfaction data regarding hotel services, it is evident that guests' travel purposes vary. The majority of guests travel for tourism (30.3%) and business (20.9%) purposes, emphasizing the importance of these sectors in contributing to the hotel's occupancy. Additionally, academic (26%) and aviation (13.8%) travels, along with personal journeys (9.0%), also significantly contribute to the guest travel profile. This conclusion underscores the importance of adapting hotel services to meet the diverse needs of guests' travel purposes.



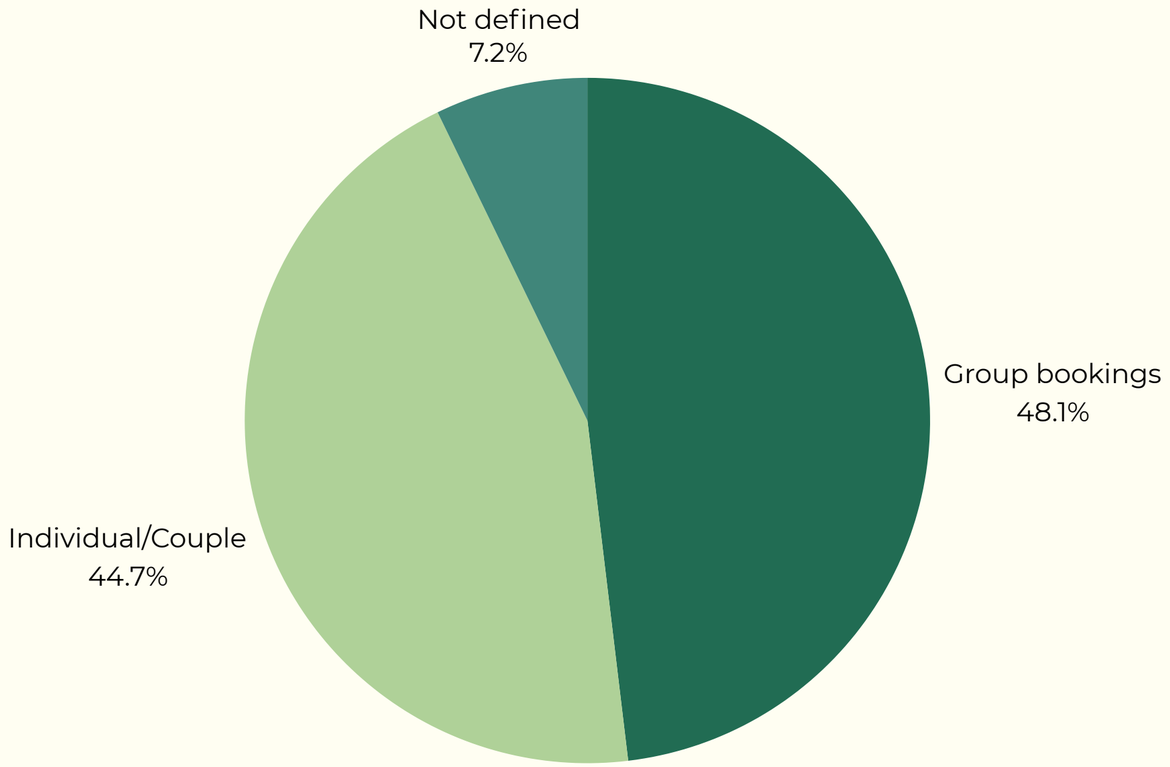
# TYPE OF TRAVEL ANALYSIS

From the analysis of travel types, it is evident that the majority of hotel guests travel in groups (Group Travel), constituting 69.3%, while personal travel (Personal Travel) accounts for 30.7%. The conclusion drawn from this data is that most guests tend to travel in groups, possibly for family vacations or group tours, while a smaller portion travels for personal reasons. This information can assist hotels in customizing their services to cater to the needs of different types of guests.



# TYPE OF BOOKING ANALYSIS

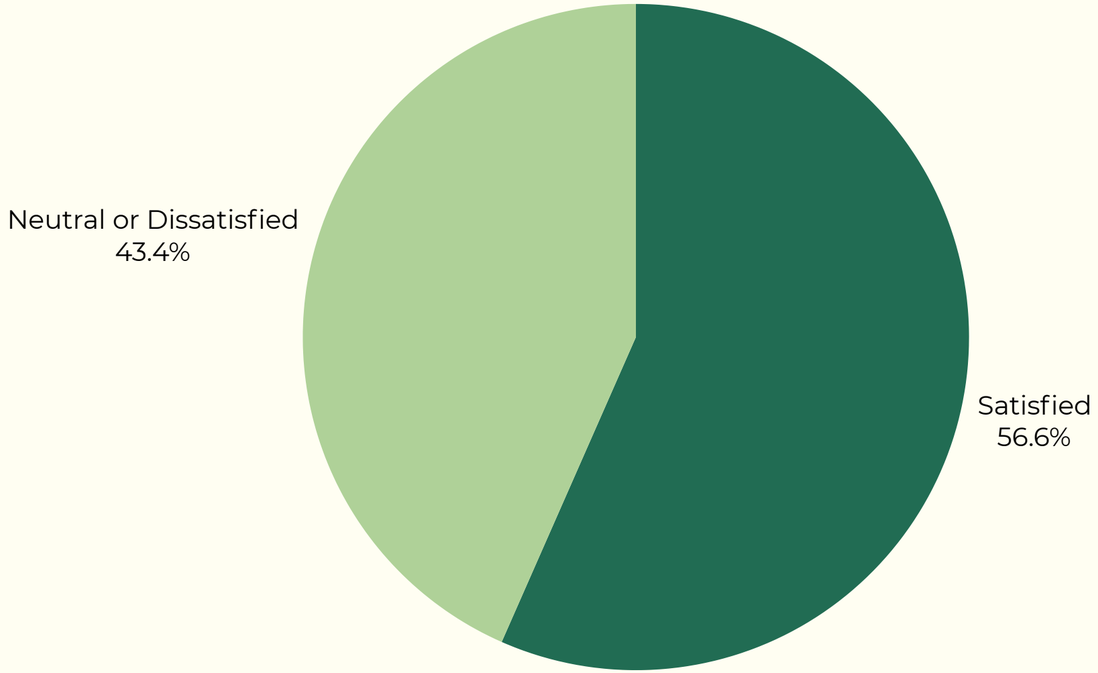
From the analysis of customer satisfaction data regarding hotel satisfaction, it is evident that there are three main types of bookings: Individual/Couple bookings contribute 44.7% of the total, while Undefined bookings represent 7.2%. However, the majority of bookings are Group bookings, contributing 48.1%. This indicates that most hotel guests make bookings as part of a group, while individual/couple bookings are also prevalent. The presence of Undefined bookings suggests the need for further clarity or categorization in the booking process. Overall, this data highlights the importance of providing services that cater to the needs of both group and individual/couple bookings to ensure customer satisfaction in the hotel industry.



# CONCLUSION OF HOTEL SATISFACTION DATA ANALYSIS

The analysis of customer satisfaction across various aspects of hotel services reveals that 43.4% of customers are satisfied, while 56.6% express neutrality or dissatisfaction. This suggests that the majority of customers feel neutral or dissatisfied with the hotel's services in the mentioned categories. Specifically, factors such as hotel WiFi service, Departure/Arrival convenience, Ease of Online booking, Hotel location, Food and drink, Stay comfort, Common Room entertainment, Check-in/Checkout service, Other service, and Cleanliness significantly influence customer satisfaction.

**Hence, to enhance customer satisfaction, it is imperative to evaluate and improve upon all the aforementioned service aspects, while also focusing on enhancing the quality and responsiveness of the services rendered.**



# THANK YOU