LUFTHANSA GROUP

Passenger Satisfaction Analysis

Wojtowicz K., Data Analysis and Visualisation, Faculty of Psychology and Cognitive Sciences, Adam Mickiewicz University, 2024/2025

All data used in this project is fictional.
Read readme.md file for additional information.

Introduction

This project analyzes key metrics related to airline passenger satisfaction using publicly available dataset. The objective is to validate hypotheses regarding airline customer satisfaction. The data was analyzed using a range of statistical methods tailored to the research questions, and the results were meticulously interpreted to derive meaningful insights.

The project simulates a real report. All contained data are fictitious.

Table of contents

Dataset overview

This dataset, sourced from Kaggle's Airline Passenger Satisfaction dataset, is a synthesized collection of data points reflecting common airline operations. It is specifically tailored for academic, research, and analytical purposes, offering a rich resource for exploring key drivers of passenger satisfaction and operational efficiency.

Source	Published on Kaggle by user teejmahal20.
Focus	Simulated airline operations and passenger satisfaction.
Dataset Size	103 904 rows and 24 columns.
Structure	Each row represents one passenger's feedback on a single flight.

#	id	Gender	Customer Type	Age	Type of Travel
0 104k	1 130k	Female 51% Male 49%	Loyal Customer 82% disloyal Customer 18%	7 85	Business travel 69 Personal Travel 31
0	70172	Male	Loyal Customer	13	Personal Travel
1	5047	Male	disloyal Customer	25	Business travel
2	110028	Female	Loyal Customer	26	Business travel
3	24026	Female	Loyal Customer	25	Business travel
4	119299	Male	Loyal Customer	61	Business travel
5	111157	Female	Loyal Customer	26	Personal Travel
6	82113	Male	Loyal Customer	47	Personal Travel

Dataset sample

Used columns

Column Name	Data description
Age	Passenger’s age in years (integer).
Type of Travel	Purpose of the travel (e.g., Business travel, Personal Travel).
Satisfaction	Overall satisfaction level (e.g., Satisfied, Neutral, Unsatisfied).
Gender	Passenger’s gender (e.g., Male, Female).
Seat comfort	Rating of seat comfort (1-5).
Ease of Online booking	Rating of online booking ease (1-5).
Leg room service	Rating of legroom space provided (1-5).
Flight distance	Distance of the flight in miles (integer).
Inflight service	Rating of overall in-flight service 1-5).
Class	Travel class (e.g., Business, Economy, Economy Plus).
Check-in service	Rating of check-in service 1-5)..
Customer Type	Type of customer (e.g., Loyal Customer, Disloyal Customer).

Research qustions

1. Are business travelers more likely to be satisfied?

This analysis aims to evaluate the effectiveness of the “Business Excellence Program” recently introduced to enhance the experience of business travelers. By comparing satisfaction levels between business and personal travels, we seek to assess the program’s impact and identify areas requiring further improvement.

Variables

Type of Travel: *Nominal*

Satisfaction: *Nominal*

Tests

Chi-squared test: *Group comparasion*

Tests results

Chi-squared test: $p < 0.05$

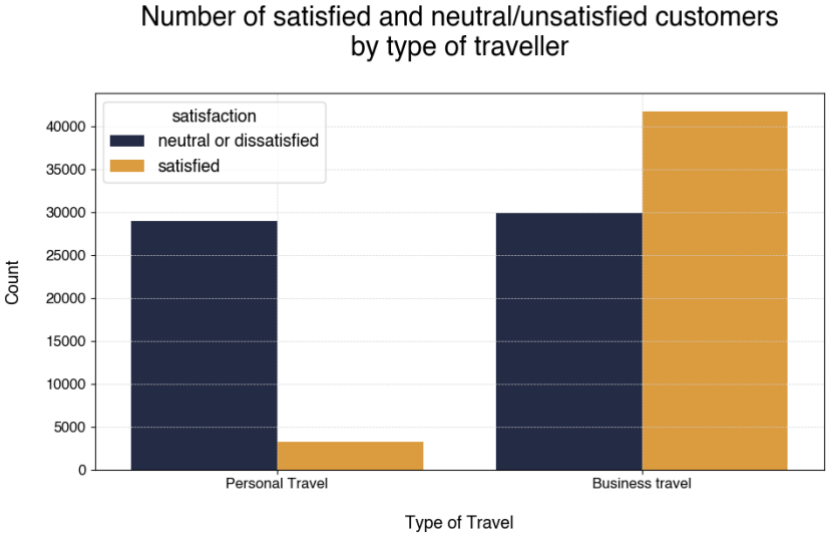
As the p value is below the significance level ($\alpha = 0.05$), we reject the null hypothesis (H0). This means that there is a statistically significant relationship between travel type and passenger satisfaction.

		satisfaction	
		neutral or dissatisfied	**satisfied**
Type of Travel	**Business travel**	29909	41746
	Personal Travel	28970	3279

Satisfaction and Type of Travel Crosstab

Interpretation

Passengers travelling in different classes have different levels of satisfaction. The prevalence of satisfaction in Business travels is particularly evident, indicating that the Business Excellence Program is having its intended effect.



2. Is there an age difference between satisfied and neutral/unsatisfied customers?

Understanding whether satisfaction levels differ across age groups is crucial for tailoring services and marketing strategies. This analysis will provide insights into demographic trends and inform age-specific initiatives to enhance the passenger experience.

Variables

Age: *Quantitive*

Satisfaction: *Nominal*

Tests

Anderson–Darling test: *Normality verification*

Mann–Whitney U test: *Group comparasion*

Data preparation

The dataset was divided into two groups: Satisfied customers and neutral or unsatisfied customers

Tests results

1. Anderson–Darling test (for N > 5000):

Satisfied customers: statistic > critical_value (0.05)
Neutral or unsatisfied customers: statistic > critical_value (0.05)

For both age groups, the distributions are not normal (statistic > critical_value, H0 was rejected). Therefore, a Mann-Whitney test was performed.

2. Mann–Whitney U test: $p < 0.05$

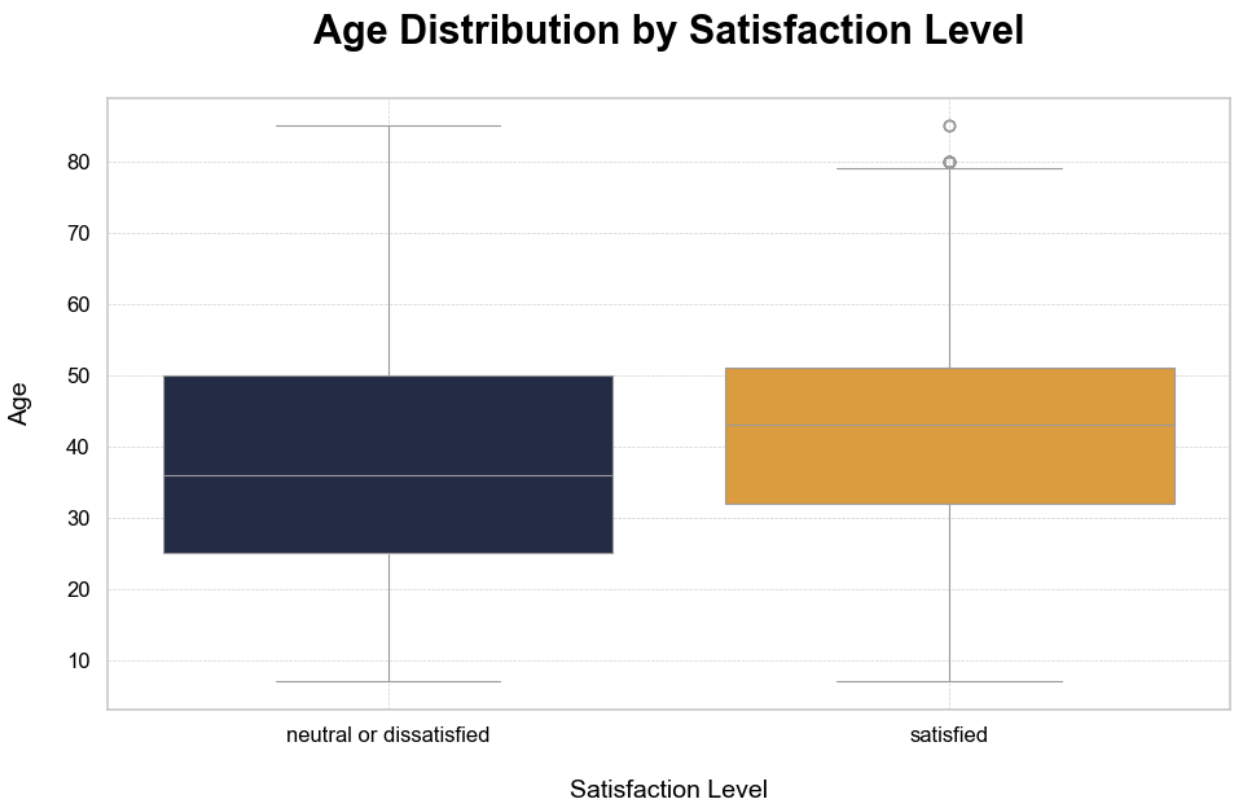
As the p-value is well below the significance level $\alpha = 0.05$, we reject the null hypothesis (H0). This means that there is a statistically significant difference in passenger age between the 'neutral or dissatisfied' and 'satisfied' groups.

Satisfaction	Count	Mean	50%
neutral or dissatisfied	58879.0	37.566688	36.0
satisfied	45025.0	41.750583	43.0

Interpretation

Satisfied passengers have a higher median age compared to neutral or dissatisfied passengers. This means that older people are more likely to be satisfied with the airline's services compared to younger passengers.

Further analysis is requred to identify key pain points and create strategy amied at younger passengers satisfaction.



3. Do women find the seats more comfortable than men?

Following the implementation of redesigned seats, this study investigates whether perceptions of seat comfort vary between genders. The results will help determine whether the upgrades have effectively addressed the needs of all passengers.

Variables

Gender: *Nominal*

Seat comfort: *Quantitive*

Tests

Anderson–Darling test: *Normality verification*

Mann–Whitney U test: *Group comparasion*

Data preparation

The dataset was divided into two groups: Male customers and Female customers

Tests results

1. Anderson–Darling test (for N > 5000):

Male customers: statistic > critical_value (0.05)
Female customers: statistic > critical_value (0.05)

For both groups, the distributions are not normal (statistic > critical_value, H0 was rejected). Therefore, a Mann-Whitney test was performed.

2. Mann–Whitney U test: p < 0.05

As the p-value is well below the significance level α = 0.05, we reject the null hypothesis (H0). This means that there is a statistically significant difference in seat comfort rating between female and male customers groups.

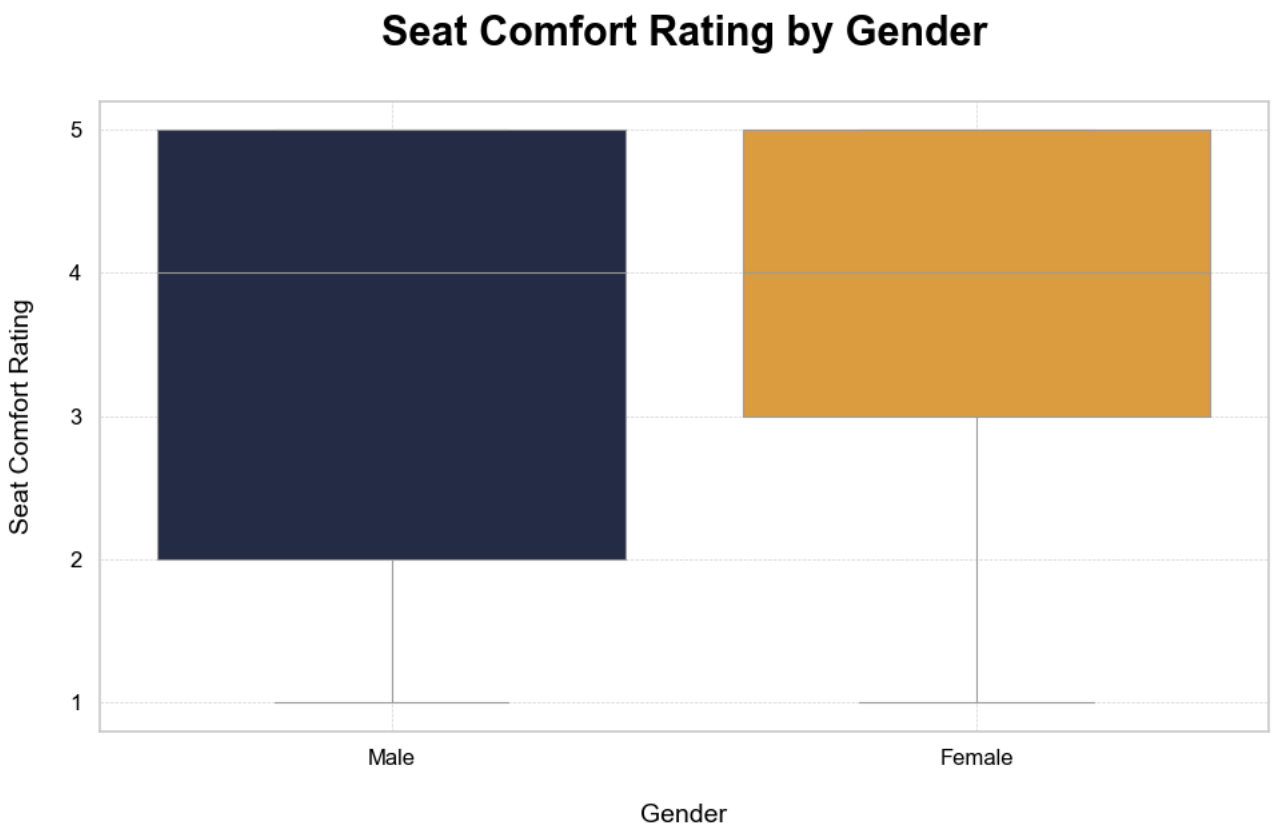
Descriptive statistics, central tendency

Gender	Count	Mean	50%
Female	52727.0	3.473837	4.0
Male	51177.0	3.403912	4.0

Interpretation

Women rate the comfort of the seats slightly higher than men, as suggested by the earlier descriptive statistics (mean). This reslt suggests that women are more likelly to rate seat comfort higher that men.

However, is worth considering additional research to understand each gender's specific expectations of seat comfort.



4. Does the rating of the ease of use of the online booking system decrease with the age of the passengers?

To evaluate the accessibility of the recently launched online booking platform, this analysis examines whether ease-of-use ratings decline with passenger age. The findings will inform potential adjustments to improve the platform's usability for all age groups.

Variables

Ease of Online booking: *Ordinal*

Age: *Quantitive*

Tests

Anderson–Darling test: *Normality verification*

Spearman's rank correlation: *Correlation*

Pearson correlation: *Correlation*

Tests results

1. Anderson–Darling test (for N > 5000):

Ease of Online booking: statistic > critical_value (0.05)
Age: statistic > critical_value (0.05)

For both variables, the distributions are not normal (statistic > critical_value, H0 was rejected). Therefore, a Spearman's rank correlation should be performed.

Additionally Pearson correlation was performed for comparasion.

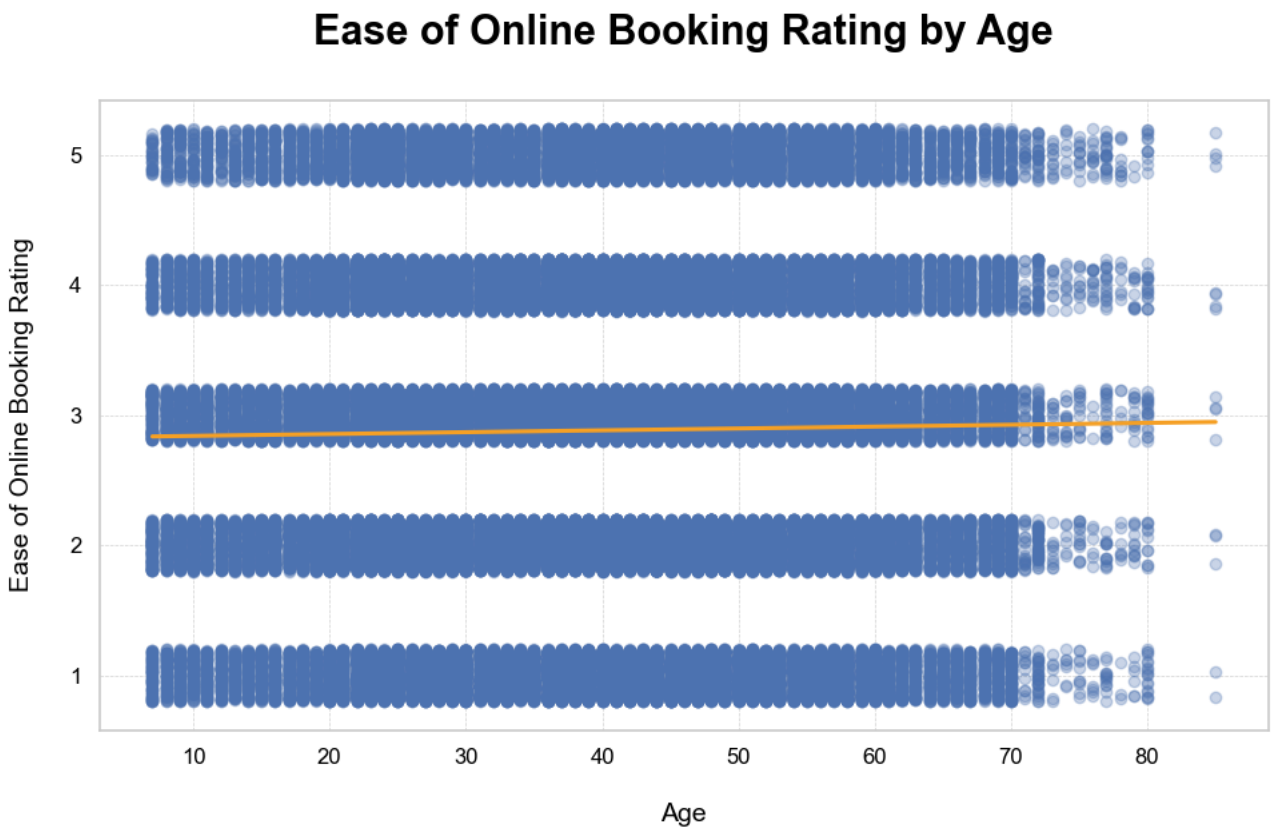
2. Spearman's rank correlation: $p < 0.05$
   Pearson correlation: $p < 0.05$

As the p-value is below the significance level $\alpha = 0.05$, we reject the null hypothesis (H0). This means that there is a statistically significant correlation between ease of booking and Age.

Interpretation

Although the result is statistically significant, the correlation value of r = 0.023 suggests that the relationship between the age of passengers and the rating of the ease of online booking is virtually extremly weak.

Further development of online booking system should involve proper accesibility features, but based on the currecnt data no additional action is required to increase usability for older people.



5. Does the leg room rating decrease as the flight distance increases?

This analysis explores whether satisfaction with legroom diminishes on longer flights. Identifying such trends will guide decisions regarding seat configurations on long-haul routes to enhance passenger comfort.

Variables

Flight distance: *Quantitive*

Leg room service: *Ordinal*

Tests

Anderson–Darling test: *Normality verification*

Spearman's rank correlation: *Correlation*

Tests results

1. Anderson–Darling test (for N > 5000):

Flight distance: statistic > critical_value (0.05)
Leg room service: statistic > critical_value (0.05)

For both variables, the distributions are not normal (statistic > critical_value, H0 was rejected). Therefore, a Spearman's rank correlation should be performed.

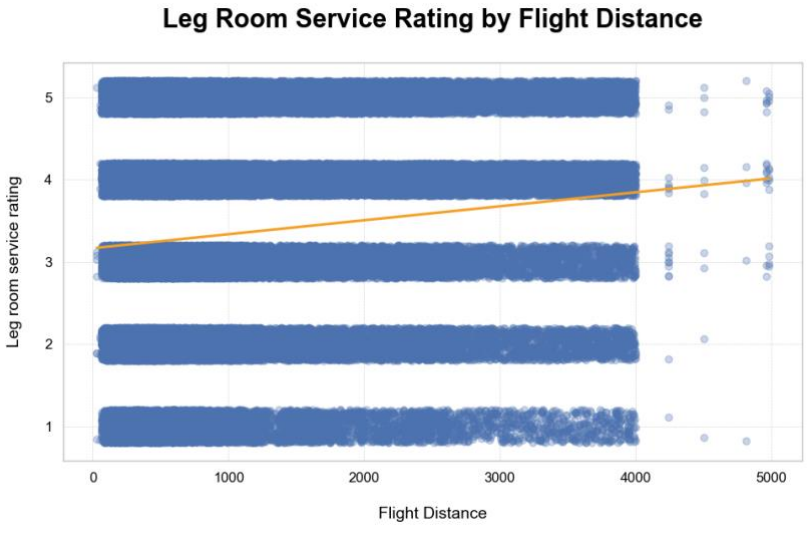
2. Spearman's rank correlation: $p < 0.05$

As the p-value is below the significance level $\alpha = 0.05$, we reject the null hypothesis (H0). This means that there is a statistically significant correlation between Flight distance and Age.

Interpretation

Although the result is statistically significant, the correlation value of $r = 0.1140$ suggests that the positive correlation between the flight distance and leg room service raiting is weak.

This means that the effect is almost not observable and no further adjustemnts of leg rooom service is required.



6. Are there significant differences in the rating of in-flight services between Business, Economy and Economy Plus class passengers?

To ensure consistency in service delivery, this study evaluates passenger ratings of in-flight services across travel classes. The results will highlight any disparities and inform initiatives to optimize the service experience for all passengers.

Variables	Tests
Inflight service: *Ordinal*	Anderson–Darling test: *Normality verification*
Class: *Nominal*	Kruskal-Wallis test: *Group comparasion*
	Post-Hoc Dunn Test: *Group comparasion*

Data preparation

The dataset was divided into three groups: Economy passengers and Economy Plus passengers and Business passengers

Tests results

1. Anderson–Darling test (for N > 5000):

Economy passengers: statistic > critical_value (0.05)
Economy Plus passengers: statistic > critical_value (0.05)
Business passengers: statistic > critical_value (0.05)

For all groups, the distributions are not normal (statistic > critical_value, H0 was rejected). Therefore, a Kruskal-Wallis test was performed.

2. Kruskal-Wallis test: $p < 0.05$

As the p-value is well below the significance level $\alpha = 0.05$, we reject the null hypothesis (H0). This means that there is a statistically significant difference in inflight service rating between Economy, Economy Plus Business passengers groups.

Descriptive statistics, central tendency

Class	Count	Mean	50%
Business	49662.0	3.844811	4.0
Economy Plus	7494.0	3.388444	4.0
Economy	46745.0	3.463921	4.0

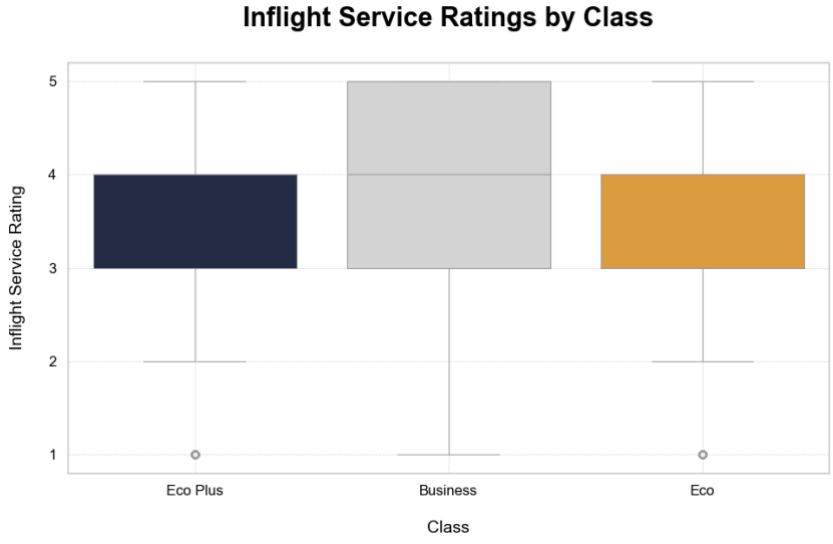
3. Post-Hoc Dunn Test:

Class 1	Class 2	P-value	Interpretation
Business	Economy	2.66e-08 < 0.05	Significant difference: Business > Economy.
Business	Economy Plus	1.28e-241 < 0.05	Significant difference: Business > Economy Plus.
Economy	Economy Plus	0.00e+00 < 0.05	Significant difference: Economy > Economy Plus.

Interpretation

Business class passengers rate the Inflight service slightly higher than Economy passengers, but Economy passengers rate the Inflight service slightly higher than Economy Plus passengers as suggested by the descriptive statistics (mean).

Further analysis is recommended to verify the reason of higher Inflight service rates in Economy class group compared to Economy Plus class group.



7. Does the check-in rating differ between loyal and disloyal customers travelling in business class?

As part of an evaluation of the loyalty program, this analysis compares check-in satisfaction ratings between loyal and non-loyal customers in business class. The findings will help identify any service gaps affecting customer retention and loyalty.

Variables	Tests
Check-in service: *Ordinal*	Anderson–Darling test: *Normality verification*
Customer Type: *Nominal*	Mann–Whitney U test: *Group comparasion*

Data preparation

The dataset was divided into two groups: Disloyal Business Customers and Loyal Business Customers.

Tests results

1. Anderson–Darling test (for N > 5000):

Disloyal Business Customers: statistic > critical_value (0.05)
Loyal Business Customers: statistic > critical_value (0.05)

For both groups, the distributions are not normal (statistic > critical_value, H0 was rejected). Therefore, a Mann-Whitney test was performed.

2. Mann–Whitney U test: $p < 0.05$

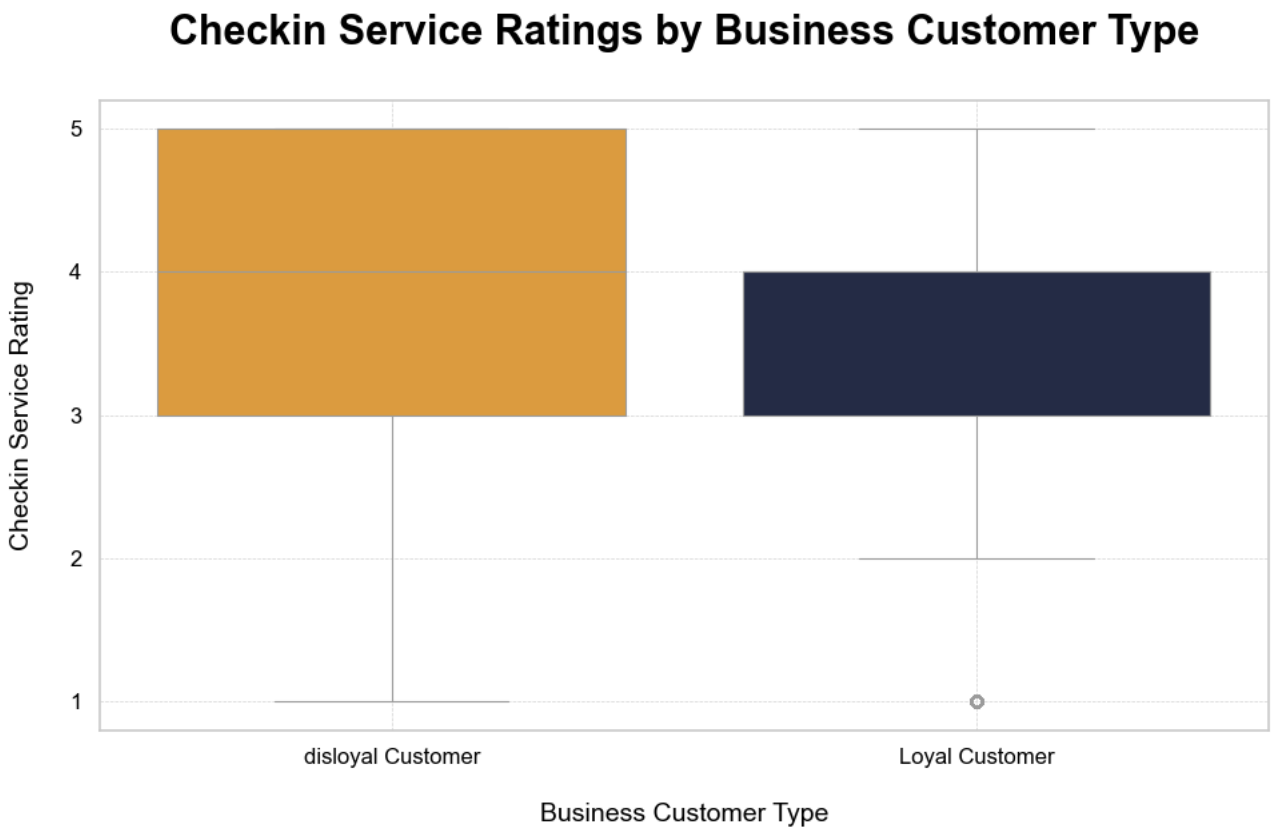
As the p-value is well below the significance level $\alpha = 0.05$, we reject the null hypothesis (H0). This means that there is a statistically significant difference in Check-in service rating between Disloyal Business Customers and Loyal Business Customers.

Descriptive statistics, central tendency

Class	Count	Mean	50%
Disloyal Business Customers	7356.0	3.714791	4.0
Loyal Business Customers	42308.0	3.485251	4.0

Interpretation

Disloyal Business Customers are on avarage more satisfied with Check-in service than Loyal Business Customers. This suggests that further analysis of Loyal Business Customers needs is required to identify areas for further improvement for these clients.



Discussion

The analysis highlights key insights into passenger satisfaction, including higher satisfaction among business travelers, older passengers, and female passengers for specific aspects like seat comfort. While the online booking system shows broad accessibility, periodic updates could enhance its usability further. Lower ratings in Economy Plus and unexpected satisfaction discrepancies among loyal customers warrant deeper investigation to address potential service gaps. Recommendations include refining loyalty programs, tailoring services for younger passengers, and improving in-flight experiences in Economy Plus. These actions will strengthen customer satisfaction and enhance the airline’s overall service quality.