**CSE375**

**PROJECT REPORT**

(Project Semester January–April 2025)

**AIRLINE PASSENGER SATISFACTION ANALYSIS**

Submitted by

Amandeep Singh

Registration No: 12303801

B.Tech CSE, Section: KM007

Course Code: CSE375

Under the Guidance of

**Dr. Mrinalini Rana, U.Id: 22138**

**Head of Department, Data Science**

**Discipline of CSE/IT**

**Lovely School of Computer Science and Engineering**

**Lovely Professional University, Phagwara**

**DECLARATION**

### I, Amandeep Singh, student of B.Tech Computer Science and Engineering, under the CSE/IT Discipline at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

### Date: Signature

### Registration No: 12303801 Amandeep Singh

**CERTIFICATE**

### This is to certify that Amandeep Singh, bearing Registration No: 12303801, has completed CSE375 project titled, “Airline Passenger Satisfaction Analysis” under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort and study.

### Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### Dr. Mrinalini Rana

### Head of Department – Data Science

### School of Computer Science and Engineering

### Lovely Professional University

### Phagwara, Punjab

### Date: \_\_\_\_\_\_\_\_\_\_\_

**ACKNOWLEDGEMENT**

I would like to express my sincere gratitude to all those who have supported me in completing this project titled **“Airline Passenger Satisfaction Analysis”**.

First and foremost, I extend my heartfelt thanks to **Lovely Professional University** for providing me with the opportunity to undertake this Data Science minor project.

I am especially grateful to my project mentor and guide, **Dr. Mrinalini Rana**, Head of Department, School of Computer Science and Engineering, for her invaluable support, guidance, and continuous encouragement throughout this project. Her knowledge and experience have been instrumental in shaping the direction of this work.

I would also like to acknowledge **Maven Analytics** for providing the dataset used in this project, and I appreciate the contributions of the open-source community whose libraries and resources enabled me to carry out the analysis effectively.

Lastly, I thank my family, friends, and peers for their constant motivation, feedback, and support during every phase of the project.

**TABLE OF CONTENT**

- Cover Page ....................................................................................................................... 1

- Declaration ....................................................................................................................... 2

- Certificate ........................................................................................................................ 3

- Acknowledgement ........................................................................................................... 4

1. [Introduction ....................................................................................................................](#_1._Introduction) 6

2. [Source of Dataset ...........................................................................................................](#_2._Source_of) 7

3. [EDA Process ..................................................................................................................](#_3._EDA_Process) 9

4. [Analysis on Dataset .......................................................................................................](#_4.1_Impact_of) 12

4.1 [Impact of Flight Distance on Service Experience and Satisfaction .........................](#_4.1_Impact_of) 12

4.2 [Satisfaction Rate by Class .......................................................................................](#_4.2_Satisfaction_Rate) 14

4.3 [Lowest Rated In-Flight Services .............................................................................](#_4.3_Lowest_Rated) 16

4.4 [Service Ratings by Age Group ...............................................................................](#_4.4_Service_Ratings) 18

4.5 [Delays vs Satisfaction .............................................................................................](#_4.5_Delays_vs) 21

5. [Conclusion ...................................................................................................................](#_5._Conclusion) 26

6. [Future Scope ................................................................................................................](#_6._Future_Scope) 29

7. [References ....................................................................................................................](#_7._References) 32

# 1. Introduction

In today’s highly competitive aviation industry, passenger satisfaction stands as a vital metric that directly influences customer loyalty, airline reputation, and long-term profitability. With increasing competition, rising customer expectations, and growing global air traffic, airlines are constantly seeking innovative ways to enhance service quality and deliver a seamless travel experience.

This project, titled **"Airline Passenger Satisfaction Analysis"**, aims to explore the various factors that influence how passengers perceive their journey. Using a real-world dataset sourced from Maven Analytics, which contains detailed feedback from over 100,000 airline passengers, the study investigates customer experiences related to seating comfort, on-board services, food and beverages, in-flight entertainment, cleanliness, delays, and more.

To conduct this analysis, we applied core data science techniques using Python libraries such as **Pandas**, **NumPy**, **Seaborn**, and **Matplotlib**. The methodology involved comprehensive **Exploratory Data Analysis (EDA)**, handling of missing values, **feature engineering** (such as the creation of a Total Service Score), and segmentation of passengers based on class, age group, and flight distance.

### The key objectives of this project include:

* Analyzing how **flight distance** impacts overall satisfaction and service experience.
* Measuring **satisfaction rates across different travel classes**.
* Identifying the **lowest-rated in-flight services** needing improvement.
* Exploring how **age groups perceive different service features**.
* Studying the effect of **arrival and departure delays** on customer satisfaction.

The insights derived from this analysis are visualized using bar plots, heatmaps, scatter plots, and line charts for better interpretability. Ultimately, this project seeks to offer **actionable, data-driven recommendations** that airlines can implement to improve satisfaction scores, enhance service delivery, and make strategic customer-centric decisions.

# 2. Source of Dataset

The dataset used for this analysis is titled **“Airline Passenger Satisfaction”** and was sourced from the **Maven Analytics Data Playground**—an educational platform offering real-world datasets for practicing data science and analytics. The dataset is publicly available and can be accessed via the following link:

<https://mavenanalytics.io/data-playground?page=9&pageSize=5>

This dataset simulates customer feedback collected by a commercial airline from its passengers following their flight experiences. It offers a balanced mix of demographic information, travel-related attributes, in-flight service ratings, and overall satisfaction levels, making it highly suitable for performing in-depth exploratory data analysis (EDA) and model-based diagnostics.

### ****Dataset Overview****

Each record in the dataset represents an individual passenger and includes **23 attributes**, categorized as follows:

#### ****1. Passenger Demographics****

* **ID**: A unique identifier for each passenger
* **Gender**: Gender of the passenger
* **Age**: Age in years

#### ****2. Travel Profile****

* **Customer Type**: Either ‘First-time’ or ‘Returning’
* **Type of Travel**: ‘Business’ or ‘Personal’
* **Class**: Travel class – Economy, Economy Plus, or Business
* **Flight Distance**: Flight distance in miles

#### ****3. Flight Punctuality****

* **Departure Delay** and **Arrival Delay** (in minutes): Captures flight timeliness

#### ****4. Service Ratings**** (Rated from 0 to 5)

Passengers rated the quality of services across different stages of their journey:

* Departure and Arrival Time Convenience
* Ease of Online Booking
* Check-in Service
* Online Boarding
* Gate Location
* On-board Service
* Seat Comfort
* Leg Room Service
* Cleanliness
* Food and Drink
* In-flight Service
* In-flight Wi-Fi Service
* In-flight Entertainment
* Baggage Handling

#### ****5. Overall Satisfaction****

* **Satisfaction**: The target variable indicating whether the passenger was **"Satisfied"** or **"Neutral or Dissatisfied"** with the flight experience.

### ****Suitability for Analysis****

This dataset is particularly well-suited for analytical tasks due to its structured design, balanced mix of numerical and categorical variables, and its relevance to real-world airline operations. It facilitated a comprehensive evaluation of:

* **Passenger satisfaction trends**
* **Service quality gaps**
* **Operational inefficiencies**
* And segmentation by **class**, **age**, and **flight distance**

Moreover, it provided an excellent platform to apply **statistical analysis**, **data visualization**, and **data science methodologies** in an industry-specific context.

By analyzing this dataset, the project aimed to extract actionable insights that can help airlines enhance customer experience, prioritize service improvements, and maintain a competitive edge in the aviation industry.

# 3. EDA Process

Exploratory Data Analysis (EDA) is an essential step in any data science project. It allows analysts to explore, clean, transform, and visualize data to uncover patterns and prepare it for meaningful analysis. In this project, EDA played a crucial role in preparing the airline passenger dataset for deeper insights into customer satisfaction.

The EDA process included the following key steps:

#### ****1. Data Loading and Setup****

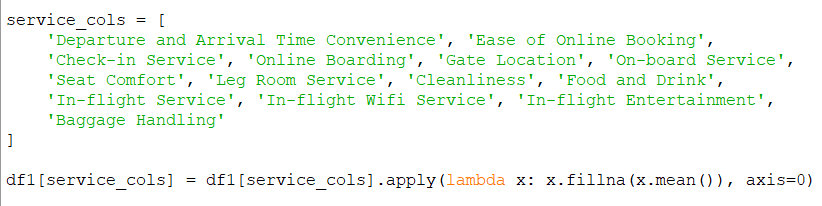
The dataset was loaded using the **Pandas** library into a DataFrame. A working copy was created to apply transformations and maintain the original structure. All further operations were performed on this processed copy.  


#### ****2. Data Cleaning****

Missing values were primarily found in service-related features such as:

* **In-flight WiFi Service**
* **Gate Location**
* **Leg Room Service**

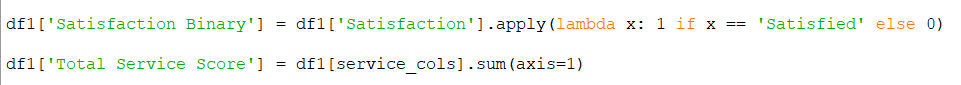
These missing entries were filled using the **mean value** of each respective column, ensuring consistency without distorting overall data distribution.



#### ****3. Feature Engineering****

Two new features were created to enhance analysis:

* **Satisfaction Binary**:  
  The original satisfaction column was encoded as binary:
  + 1 for “Satisfied”
  + 0 for “Neutral or Dissatisfied”
* **Total Service Score**:  
  This numeric feature was calculated by summing **14 service-related rating columns**, providing an overall quality indicator of each passenger’s experience.

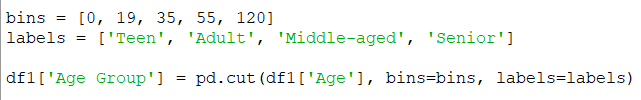


#### ****4. Age Group Segmentation****

Passengers were segmented into four age-based categories for demographic analysis:

* **Teen** (0–19 years)
* **Adult** (20–35 years)
* **Middle-aged** (36–55 years)
* **Senior** (56+ years)

This enabled a comparative view of how different age groups perceive airline services.

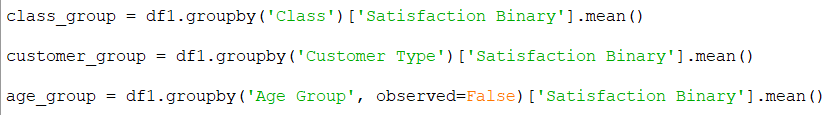


#### ****5. Grouping and Aggregation****

Grouping operations were performed using:

* **Class** (Economy, Economy Plus, Business)
* **Customer Type** (First-time, Returning)
* **Age Group**

These aggregations helped in identifying trends and variations in satisfaction and service expectations across different passenger segments.



#### ****6. Visual Explorations****

The data was explored visually using **Seaborn** and **Matplotlib** to support key insights:

* **Bar plots**: Satisfaction by travel class and distance group
* **Heatmaps**: Service rating differences across age groups
* **Line plots**: Trends in select services across demographic segments
* **Scatter plots**: Visualizing how delays affect satisfaction
* **Correlation matrix**: Understanding relationships between numeric variables like delays and satisfaction

This EDA process uncovered foundational patterns in the data and enabled the development of five targeted analysis objectives, which are presented in the following section.

**4. Analysis on Dataset**

## ****4.1 Impact of Flight Distance on Service Experience and Satisfaction****

### i. ****Introduction****

Flight distance is an important factor that can influence a passenger’s overall experience. On longer flights, passengers may place greater emphasis on comfort, entertainment, food, and service efficiency. In contrast, shorter flights may limit the scope for service delivery, making passengers more sensitive to delays or boarding processes. This objective aims to explore whether there is a significant difference in satisfaction and overall service perception based on the distance traveled.

### ii. ****General Description****

Passengers were segmented into three categories based on their flight distance:

* **Short-haul**: 0–1000 km
* **Medium-haul**: 1001–2000 km
* **Long-haul**: 2001+ km

For each group, we calculated:

* The **average satisfaction rate** (using the binary Satisfaction Binary column)
* The **Total Service Score**, a metric created earlier by summing up 14 service-related ratings

This comparison helps determine whether longer flights result in higher passenger satisfaction and better perceived service.

### iii. ****Specific Requirements, Functions and Formulas****

**Step 1: Segmenting Flight Distance and Grouping Averages**

### 

### iv. ****Analysis Results****

The grouping and aggregation revealed the following insights:

* **Long-haul passengers** had the **highest satisfaction rate** and **Total Service Score**
* **Medium-haul passengers** followed in both metrics
* **Short-haul passengers** reported the **lowest satisfaction**, potentially due to shorter duration limiting the ability to deliver enhanced services or recover from delays

This analysis suggests a **positive correlation between flight distance and satisfaction**, highlighting how service quality expectations shift based on travel time.

### v. ****Visualization****

#### Bar Chart: Satisfaction Rate by Distance Group

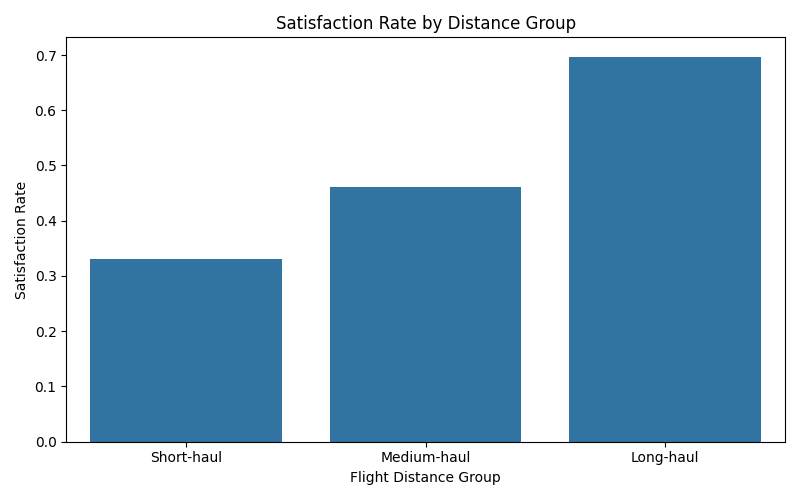
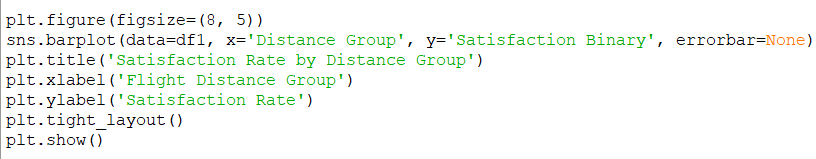


Figure 4.1(a): Satisfaction Rate by Distance Group



#### Line Chart: Total Service Score by Flight Distance

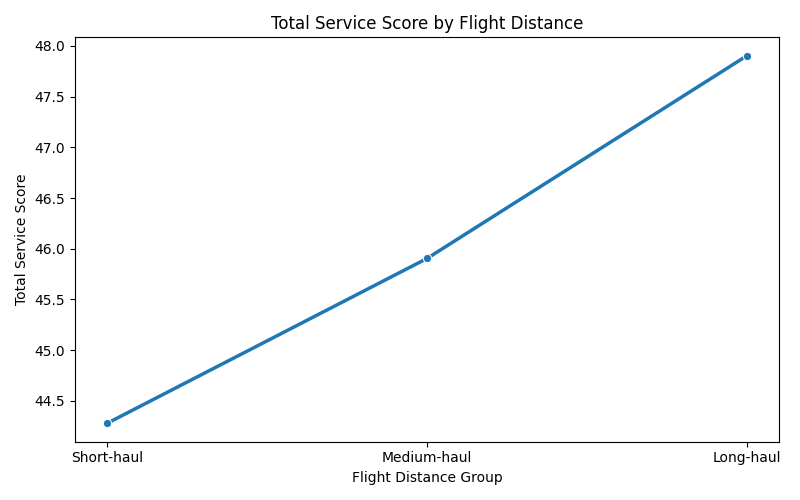
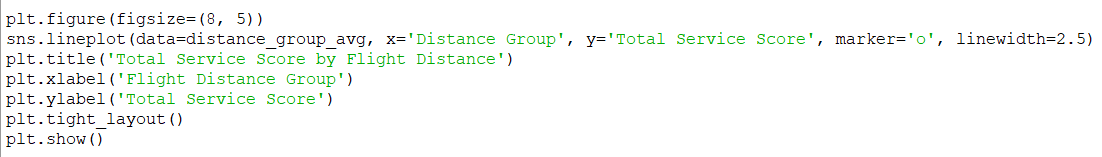


Figure 4.1(b): Total Service Score by Distance Group



These visuals clearly show that longer flights are generally associated with better service and higher satisfaction scores.

## ****4.2 Satisfaction Rate by Class****

### i. ****Introduction****

Airlines typically offer multiple travel classes—such as **Economy**, **Economy Plus**, and **Business**—to cater to passengers with varying budgets and expectations. These classes differ significantly in terms of **seating comfort**, **meal quality**, **priority services**, and **overall in-flight experience**. This analysis aims to examine how satisfaction levels vary across these travel classes and determine which one consistently delivers the highest customer satisfaction.

### ii. ****General Description****

The dataset categorizes passengers based on the Class column into:

* Economy
* Economy Plus
* Business

To quantify satisfaction levels:

* The Satisfaction column was previously converted into a **binary format**:
  + 1 for “Satisfied”
  + 0 for “Neutral or Dissatisfied”

We then calculated the **average satisfaction rate per class** using the grouped mean of this binary column.

### iii. ****Specific Requirements, Functions and Formulas****

**Step 1: Binary Satisfaction Column**  
(Already handled during EDA)



**Step 2: Grouping Satisfaction by Travel Class**



This gives a numeric satisfaction rate for each travel class.

### iv. ****Analysis Results****

The class-wise satisfaction results revealed:

* **Business Class** had the **highest satisfaction**, with average values around **0.70**
* **Economy Plus** showed **moderate satisfaction**, around **0.26**
* **Economy Class** had the **lowest satisfaction**, at about **0.19**

These differences suggest that passengers in premium cabins are far more satisfied, likely due to **better service**, **more comfort**, and **priority treatment**. The gap between Economy and Business class highlights how crucial class-based features are to the overall passenger experience.

### v. ****Visualization****

#### Bar Chart: Satisfaction Rate by Class

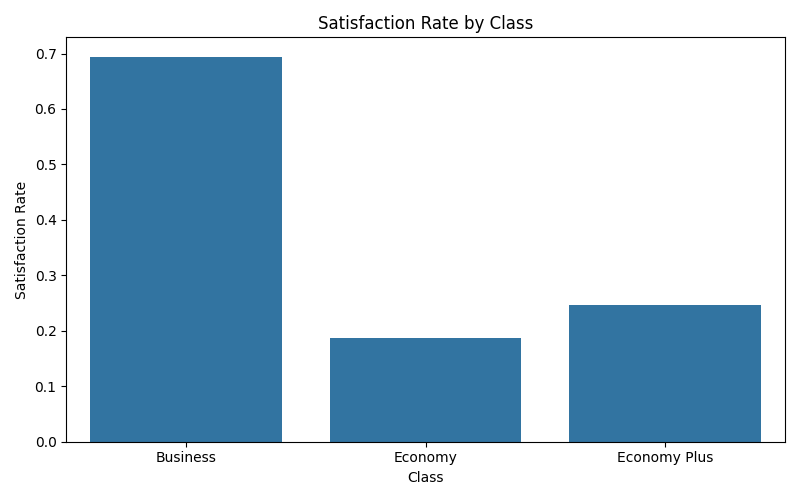
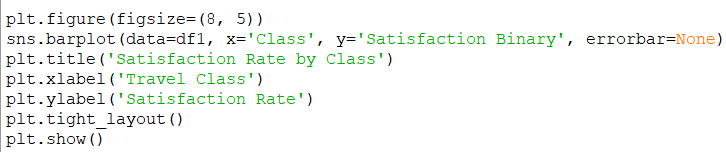


Figure 4.2: Satisfaction Rate by Travel Class



This visualization clearly reinforces the numeric analysis — Business class passengers report the highest satisfaction, followed by Economy Plus and then Economy.

## ****4.3 Lowest Rated In-Flight Services****

### i. ****Introduction****

While many passengers evaluate their flight experience based on overall satisfaction, it is equally important to examine **which specific services** are most frequently rated poorly. Identifying the **lowest-rated in-flight services** provides valuable insights into areas that require operational improvements. This objective helps airlines pinpoint **pain points** in the travel experience that negatively affect customer satisfaction.

### ii. ****General Description****

The dataset contains **14 service-related attributes**, each rated by passengers on a **scale of 0 to 5**. These include various aspects of the journey such as:

* In-flight Wi-Fi Service
* Food and Drink
* Gate Location
* Cleanliness
* Seat Comfort
* Online Boarding, etc.

To determine which services are underperforming, the **average score for each service** was calculated across all passengers. These scores were then **sorted in ascending order**, with the lowest-rated services highlighted.

### iii. ****Specific Requirements, Functions and Formulas****

**Step 1: Calculate Average Rating for Each Service**



This returns a sorted list of services from lowest to highest average rating.

### iv. ****Analysis Results****

The analysis revealed that the following services received the **lowest average ratings** from passengers:

* **In-flight Wi-Fi Service** (lowest score overall)
* **Ease of Online Booking**
* **Gate Location**

These findings suggest that **technological and logistical aspects** of the passenger journey are the most problematic. Poor Wi-Fi availability, unclear or inconvenient gate assignments, and online booking inefficiencies can significantly affect the customer experience — even more than traditional aspects like food or cleanliness.

Addressing these specific service gaps could lead to measurable improvements in **overall satisfaction**.

### v. ****Visualization****

#### Horizontal Bar Chart: Lowest Rated Services

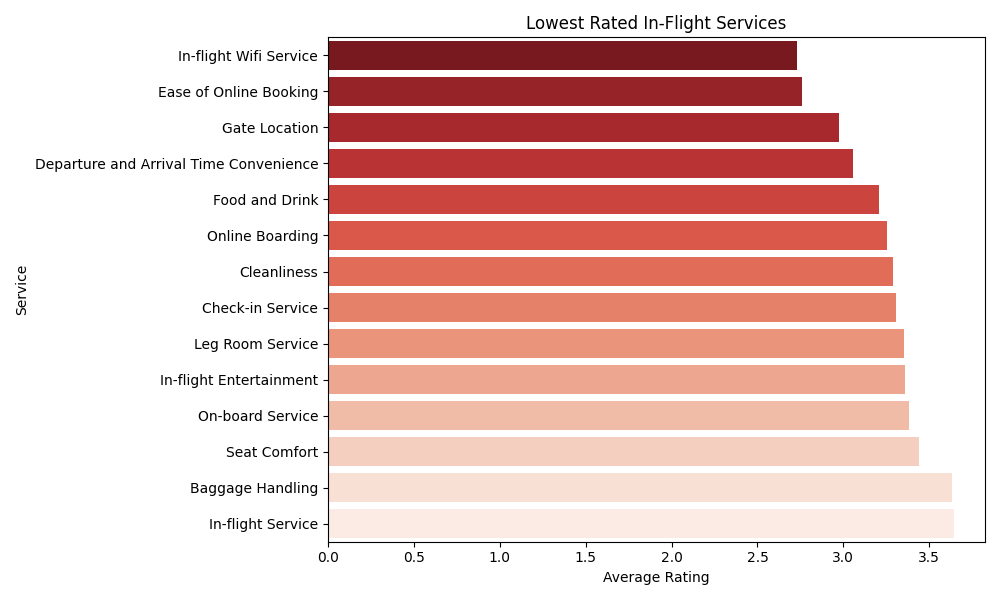
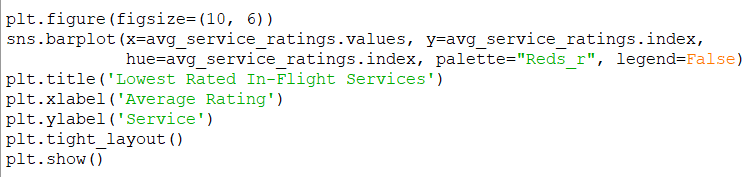


Figure 4.3: Lowest Rated In-Flight Services



This chart presents all services sorted by average rating, with those at the top being the **lowest performers**. The visual clarity allows quick identification of areas that need improvement.

## ****4.4 Service Ratings by Age Group****

### i. ****Introduction****

Passenger satisfaction is often influenced by demographic factors, particularly **age**. Different age groups may have varying expectations, comfort preferences, and service priorities. This objective aims to analyze how passengers of different **age categories** perceive various in-flight services, allowing airlines to tailor their services more effectively for different segments.

### ii. ****General Description****

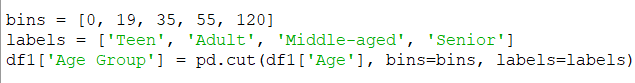
To perform this analysis, passengers were categorized into four distinct **age groups**:

* **Teen** (0–19 years)
* **Adult** (20–35 years)
* **Middle-aged** (36–55 years)
* **Senior** (56+ years)

Using these segments, we calculated the **average rating** for each of the 14 service-related columns. This allowed us to evaluate how each age group rated various aspects of their flight experience.

### iii. ****Specific Requirements, Functions and Formulas****

**Step 1: Create Age Groups**



**Step 2: Calculate Average Ratings by Age Group**



This returns a DataFrame of average ratings for each service, segmented by age group.

### iv. ****Analysis Results****

The results showed clear variations across age segments:

* **Seniors and Middle-aged passengers** consistently rated most services higher
* **Teens and Adults** gave relatively lower ratings to services like seat comfort, food, and in-flight entertainment
* Certain services such as **In-flight Wi-Fi** and **Gate Location** received low scores across all age groups, but particularly from younger passengers

These patterns suggest that **older travelers tend to be more satisfied** or may have different expectations, while **younger groups** might be more critical or digitally demanding.

This insight allows airlines to **customize services** like digital entertainment, food menus, or tech-based amenities based on passenger age groups.

### v. ****Visualization****

#### Heatmap: Average Service Ratings by Age Group

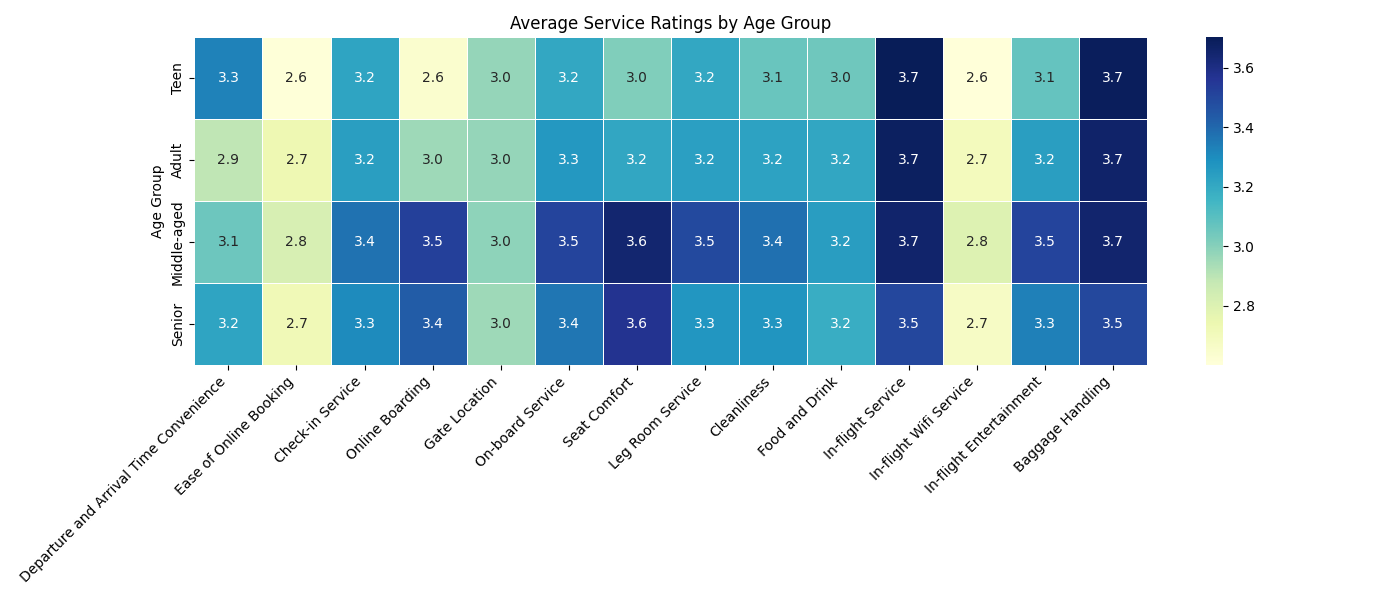
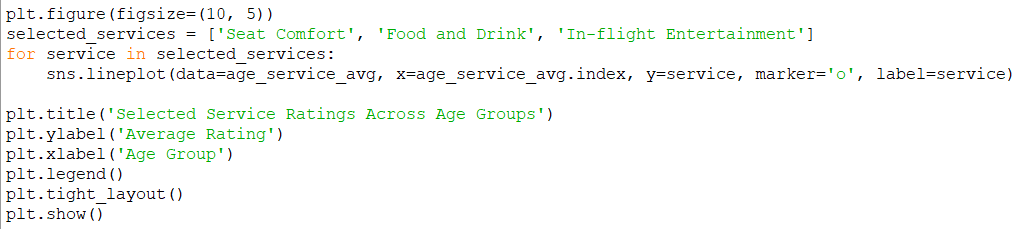


Figure 4.4(a): Heatmap of Average Service Ratings by Age Group

#### 

#### Line Plot: Selected Services by Age GroupC:\Users\HP\OneDrive\Desktop\Python for class\images\6.png

Figure 4.4(b): Line Chart of Selected Service Ratings Across Age Groups



These visuals highlight trends in service ratings across age groups — clearly showing that age influences expectations and satisfaction.

## ****4.5 Delays vs Satisfaction****

### i. ****Introduction****

Delays in flights — whether at **departure** or **arrival** — are a common source of passenger frustration. This analysis investigates how **punctuality affects passenger satisfaction**. By quantifying the relationship between delays and the satisfaction outcome, we can assess whether operational delays are a significant driver of dissatisfaction.

### ii. ****General Description****

Two delay-related columns were analyzed:

* **Departure Delay** (in minutes)
* **Arrival Delay** (in minutes)

These were compared against the **Satisfaction Binary** column (1 = Satisfied, 0 = Neutral or Dissatisfied). Using correlation and scatter plots, the strength and direction of the relationship between delays and satisfaction were visualized and interpreted.

### iii. ****Specific Requirements, Functions and Formulas****

**Step 1: Correlation Matrix**



**Step 2: Scatter Plots for Visual Relationship**

### 

### iv. ****Analysis Results****

The correlation analysis revealed a **slightly negative correlation** between both types of delays and passenger satisfaction. This means:

* As **departure or arrival delays increase**, the **likelihood of satisfaction decreases**
* The trend was not strongly linear, but **a noticeable drop** in satisfaction is visible as delays grow

These findings indicate that even moderate delays have a **measurable impact on how passengers rate their overall experience**. It confirms that **operational efficiency** is key to maintaining high satisfaction levels.

### v. ****Visualization****

#### Heatmap: Correlation Between Delays and Satisfaction

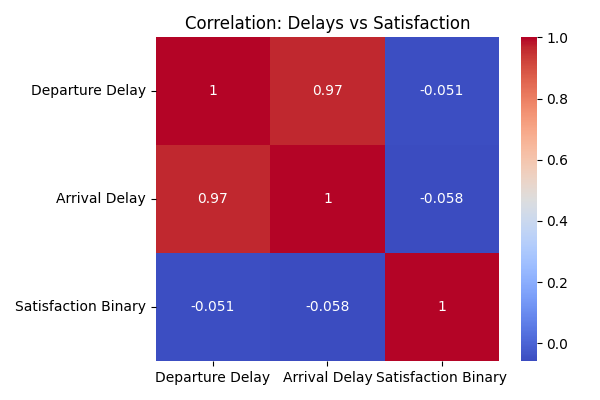
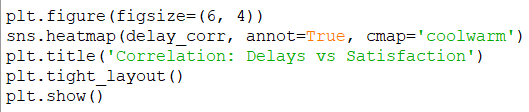


Figure 4.5(a): Correlation Heatmap – Delays vs Satisfaction



#### Scatter Plots: Delay Impact on Satisfaction

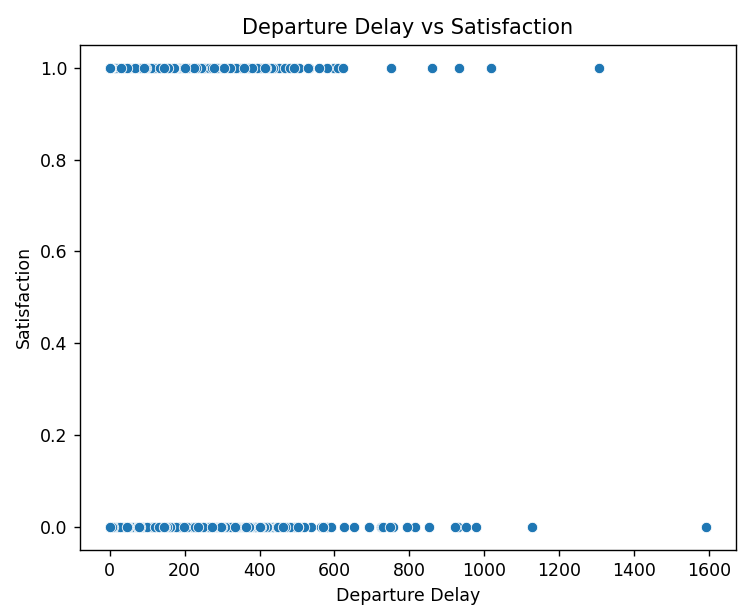


Figure 4.5(b): Scatter Plot – Departure Delay vs Satisfaction

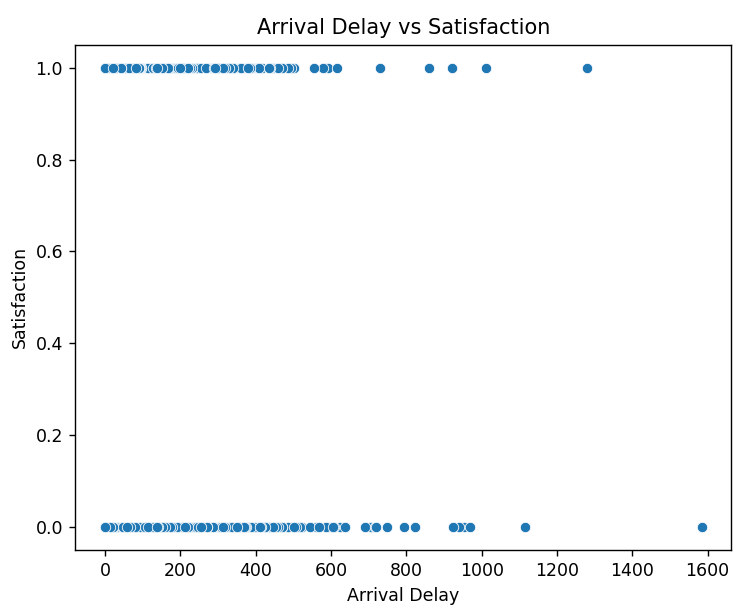
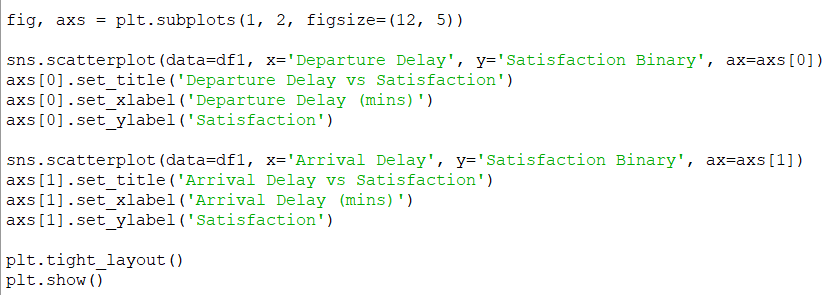


Figure 4.5(c): Scatter Plot – Arrival Delay vs Satisfaction



These plots visually confirm that **passenger satisfaction drops with longer delays**, especially when wait times exceed 50 minutes.

# 5. Conclusion

The **Airline Passenger Satisfaction Analysis** undertaken in this project served as a practical application of data science techniques to a real-world business problem—understanding and improving customer satisfaction in the airline industry. Leveraging powerful Python libraries like **Pandas**, **NumPy**, **Matplotlib**, and **Seaborn**, the analysis provided a multi-dimensional view of passenger sentiments, behaviors, and experiences.

The study adopted a data-driven approach to address critical questions around what makes passengers satisfied, which service areas require attention, and how factors like flight distance, age group, travel class, and punctuality influence perception.

### 1. Flight Distance as a Hidden Indicator of Satisfaction

Through custom segmentation, passengers were grouped into **Short-haul, Medium-haul, and Long-haul** categories. The analysis found that:

* **Long-haul travelers** consistently reported **higher satisfaction and better service ratings**.
* **Short-haul passengers** were the least satisfied, likely due to limited time for services or sensitivity to minor delays and boarding inefficiencies.

This suggests that **flight distance is not just a logistical variable—it’s a contextual driver of experience**. Airlines could adjust their service priorities based on route length to improve short-distance satisfaction and capitalize on long-haul expectations.

### 2. Travel Class Drives a Major Gap in Satisfaction

Unsurprisingly, **Business Class passengers** had significantly higher satisfaction scores compared to **Economy** and **Economy Plus**. This validates the premise that premium service offerings, such as extra legroom, priority boarding, and personalized meals, greatly enhance passenger perceptions.

Interestingly, the **gap between Economy and Economy Plus** was not as stark as the **gap between Economy and Business**, suggesting that Business Class offers the most value in terms of customer experience differentiation.

This finding highlights a strategic insight: **investing in selective premium services could improve mid-tier passenger satisfaction without requiring a full shift to Business Class levels.**

### 3. The Total Service Score as a Composite Satisfaction Metric

A new metric, **Total Service Score**, was engineered by aggregating 14 service-related features. This metric proved useful in capturing the **overall service perception** of passengers and served as a solid predictor of satisfaction. Services that consistently contributed to higher scores included:

* **Seat Comfort**
* **On-board Service**
* **Cleanliness**
* **Food and Drink**

The Total Service Score bridged the gap between subjective passenger responses and measurable service quality, providing a quantifiable lens through which to evaluate airline performance.

### 4. Consistently Low-Rated Service Areas

Not all service touchpoints met passenger expectations. Across classes and age groups, the following services were **consistently rated lowest**:

* **In-flight Wi-Fi Service**
* **Gate Location Convenience**
* **Online Booking**

These features, though sometimes overlooked by operations, are highly visible to tech-savvy and time-sensitive passengers. The results suggest that **modern expectations (e.g., reliable Wi-Fi and efficient boarding)** are unmet and offer an opportunity for competitive differentiation through improvement.

### 5. Age-Based Preferences Affect Service Perception

By categorizing passengers into **Teen, Adult, Middle-aged, and Senior** groups, it became evident that **age plays a key role** in service evaluation:

* **Senior and Middle-aged passengers** provided the most favorable ratings across most services.
* **Younger passengers**, especially Adults (20–35), were more critical—especially toward digital services and comfort-related features.

This reveals a clear opportunity for **demographic personalization**. Airlines should tailor marketing and service strategies by age segment—offering enhanced digital entertainment for younger flyers and seamless travel experiences for older travelers.

### 6. The Influence of Operational Delays on Satisfaction

Delay analysis, through correlation matrices and scatter plots, showed that:

* Both **arrival** and **departure delays** negatively impact satisfaction.
* While the correlation was weak, the **trend was clearly downward**—longer delays almost always correlated with lower satisfaction.

This underscores the operational importance of **punctuality**. Airlines should invest in **delay reduction**, **real-time communication**, and **efficient ground operations** to maintain satisfaction, especially for business travelers and short-haul routes where delays are less tolerable.

### 7. Overall Implications and Strategic Insights

This project validates the use of **exploratory data analysis (EDA)** and **feature engineering** to generate actionable insights in a service-heavy industry. Airlines can translate these insights into real-world improvements:

* **Personalize services by age group and flight duration**
* **Invest in underperforming service areas** (Wi-Fi, Gate Management)
* **Tailor offerings based on class preferences and expectations**
* **Proactively manage delays to mitigate dissatisfaction**
* **Use composite scores like Total Service Score to guide decisions**

Most importantly, the project demonstrates how **data science is not just about prediction—but about understanding and improving human experiences**. By listening to what the data says about passengers, airlines can create **more responsive, profitable, and passenger-friendly operations**.

# 6. Future Scope

While this project successfully uncovered key trends and insights through exploratory data analysis (EDA), it opens the door for more advanced, impactful applications. The foundation laid here can support strategic decision-making and pave the way for **predictive, personalized, and real-time data solutions** in the airline industry.

### ****1. Predictive Modeling****

Beyond descriptive analytics, this dataset can be used to **build classification models** that predict passenger satisfaction. Machine learning algorithms such as:

* Logistic Regression
* Random Forest
* XGBoost
* Support Vector Machines (SVM)

can be implemented to anticipate whether a passenger is likely to be satisfied based on their demographics, travel details, and service ratings. This would enable airlines to take **proactive measures** and mitigate dissatisfaction before it occurs.

### ****2. Sentiment Analysis on Textual Feedback****

If future datasets include customer reviews or open-ended comments, **Natural Language Processing (NLP)** techniques can be applied to perform **sentiment analysis**. This would enrich the analysis by revealing emotional tone and nuanced customer concerns.

Tools like:

* **VADER**, **TextBlob**, or transformer-based models like **BERT**

can help uncover sentiments, recurring themes, and hidden service issues that are not captured in numeric ratings.

### ****3. Time Series and Trend Analysis****

With access to time-stamped data, this project can be expanded to explore **temporal trends** in satisfaction, such as:

* Seasonal variations (e.g., holidays vs. off-season)
* Service quality differences by time of day or week
* Delay trends over months and years

Such insights could guide **staffing decisions**, **flight scheduling**, and **operational planning** more efficiently.

### ****4. Deep Demographic Segmentation****

Enriching the dataset with additional passenger details—such as:

* Frequent flyer status
* Ticket type and price
* Loyalty scores
* Country or region

would allow for **hyper-personalized analysis**. Airlines could then design experiences that better cater to different customer personas, enhancing satisfaction and loyalty.

### ****5. Integration with Business KPIs****

Another powerful extension would be to **link passenger satisfaction with financial and operational metrics**, including:

* Customer Lifetime Value (CLV)
* Repeat purchase behavior
* Net Promoter Score (NPS)
* Revenue per customer

This would transform satisfaction analysis from a support function to a **revenue-influencing strategy**, helping business teams understand the ROI of service upgrades.

### ****6. Real-Time Dashboards and Monitoring****

Building **interactive dashboards** using tools like:

* **Power BI** or **Tableau**, or
* Python-based apps with **Plotly Dash** or **Streamlit**

could enable real-time monitoring of passenger sentiment. Airline operations teams could act on dissatisfaction signals as they emerge—ensuring faster issue resolution and improved customer experience.

### ****Conclusion of Scope****

The future potential of this project lies in **scaling beyond exploration into prediction, personalization, and automation**. Integrating analytics with airline operations can help create a **customer-centric, data-informed environment** that supports loyalty, satisfaction, and operational excellence. As the airline industry grows more competitive, those who **listen to their data** will gain a lasting advantage.

# 7. References

[1] Maven Analytics, “Airline Passenger Satisfaction Dataset,” Maven Analytics Data Playground, [Online]. Available: <https://mavenanalytics.io/data-playground?page=9&pageSize=5>. [Accessed: Apr. 9, 2025].

[2] W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd ed. Sebastopol, CA, USA: O’Reilly Media, 2017.

[3] M. Waskom, “Seaborn: Statistical Data Visualization,” [Online]. Available: <https://seaborn.pydata.org/>. [Accessed: Apr. 9, 2025].

[4] J. D. Hunter, “Matplotlib: A 2D Graphics Environment,” Computing in Science & Engineering, vol. 9, no. 3, pp. 90–95, May–June 2007.

[5] The Pandas Development Team, “Pandas: Powerful Python Data Analysis Toolkit,” [Online]. Available: <https://pandas.pydata.org/>. [Accessed: Apr. 9, 2025].

[6] NumPy Developers, “NumPy,” [Online]. Available: <https://numpy.org/>. [Accessed: Apr. 9, 2025].