Group VA24_13

Design and Implementation of Interactive Airline Ratings Visualizations

Exam Report

Monika Vedral

08.02.2025

Contents

1 Abstract

This paper dives into the design and development of an interactive visualization application for analyzing airline ratings. The system integrates three visualizations: a bar chart for category-wise ratings, a line chart for tracking average ratings over time, and a heatmap displaying the distribution of ratings. The paper focuses on the design rationale, technical implementation, and insights gained from the data. Additionally, related work in airline rating analysis and interactive data visualization techniques is reviewed.

2 Introduction

Airline reviews contain valuable information, including overall ratings, seat comfort, cabin staff service, in-flight entertainment, and other service-related aspects. Analyzing these ratings can help airlines improve their customer experience and allow travelers to make informed decisions.

This project aims to develop an interactive application that visualizes airline rating trends, enabling users to explore how airlines perform over time and in different service categories. The visual analytics component includes:

1. Ranking of airlines based on customer satisfaction: Providing an easy comparison across airlines.
2. Trends in ratings over time: Understanding how airline performance has evolved.
3. Clustering of airline ratings using dimensionality reduction (PCA): Identifying similarities between airlines.

The intended users of this system include airline companies looking to benchmark their services and travelers seeking insights into airline quality.

The data set[Bhojani and Pitroda(2024)] originally included a much larger set of airlines from around the world, but due to the restrictions of the AS index, I chose to focus on the 10 biggest airlines operating in Europe. Currently, the data set consists of 1010 tuples in 16 dimensions, leading to an AS index of 16160.

This paper describes the development of a web-based dashboard that presents airline rating data using three coordinated visualizations. The system was implemented using D3.js and JavaScript to enable interactivity and user-driven exploration.

3 Design And Implementation

3.1 Design Rationale

The design of the application was driven by the need for clear and interactive representations of the airline rating data. The data set consists of the following attributes: ID, Airline Name, Overall Rating, Review Date, Verified, Type of Traveller, Seat Type, Date

Flown, Seat Comfort, Cabin Staff Service, Food & Beverages, Ground Service, Inflight Entertainment, Wifi & Connectivity, Value For Money, and Recommendation Status.

Given the structure of the dataset, I designed three different types of visualizations:

1. Bar Chart: Displays category-wise ratings, allowing users to compare airlines based on specific criteria. This choice was based on the fact that multiple categorical rating attributes exist in the data set, which makes a comparative bar chart the best approach to visualize them. The user can also sort the x-axis alphabetically (enabled by default) or by rating score from highest to lowest.

2. Line Chart: Tracks changes in average ratings over time, enabling users to identify trends and performance improvements. Since the data set includes both a review date and a flight date, it allows for temporal analysis, making a line chart a logical choice. Also here the user can choose to only show data for a specific airline or aggregate data by displaying the average rating of all airlines.

3. Heatmap: Shows the distribution of ratings, highlighting common rating patterns and outliers. This visualization was chosen because the dataset contains multiple rating dimensions and I wanted to explore how frequently certain scores appear.

3.2 System Architecture

The system is built using JavaScript and D3.js, with a modular structure to handle different visualization components. The HTML layout consists of div elements designated for each visualization, and the JavaScript code dynamically populates these sections. Each chart is implemented as a separate module, making it easy to update or modify each component.

3.3 Implementation Details

The system consists of three primary JavaScript modules:

1. BarChart.js: Implements the bar chart visualization and provides interactive selection feature.

2. LineChart.js: Handles the line chart, allowing users to filter data by airline.

3. Heatmap.js: Generates the heatmap, mapping rating distributions to color intensities.

3.3.1 BarChart.js

This visualization incorporates interactive elements that enhance user engagement and exploratory data analysis. These features include:

1. Dropdown Functionality: Two dropdown menus enable users to select a rating category (e.g., "Seat Comfort") and sort airlines either alphabetically or by rating. This updates the chart by recalculating averages and reordering the x-axis.

2. Tooltips and User-Triggered Computations

   Chart Title Tooltip: A tooltip displays a brief description when hovering over the chart title.

   Bar Chart Tooltips: Bar tooltips show the selected category's average rating, the number of ratings, and the deviation from the overall mean.

3. Coordinated Interaction with the Line Chart

   Clicking on a bar triggers a coordinated update in the line chart. More specifically, the event listener calls a function that updates the line chart to reflect data corresponding to the selected airline. Clicking the same bar again resets the selection, restoring the default state in both the bar and line chart.

3.3.2 LineChart.js

The line chart visualization provides an interactive way to explore airline ratings over time. The chart is also designed to coordinate with the bar chart visualization.

1. Dropdown-Based Filtering: A dropdown menu allows users to select a specific airline, dynamically updating the displayed data. The default selection is "All", which aggregates the ratings across all airlines.

2. Interactive Tooltips:

   A tooltip is displayed when hovering over the chart title, providing a brief summary of the visualization.

   Data tooltips appear when hovering over a data point (circle) along the line graph, displaying the year of the rating, the mean rating for the selected airline in that year and the number of reviews contributing to the rating.

3. Coordinated Interaction with the Bar Chart: When an airline is selected from the dropdown, the event listener triggers updating the bar chart accordingly. If an airline is selected in the bar chart, the line chart is updated to reflect its historical trend. This ensures bidirectional interaction between the two visualizations.

3.3.3 Heatmap.js

The heatmap visualization provides a view into the distribution of airline ratings across different airlines. By utilizing Principal Component Analysis (PCA), the tool reduces the dimensionality of multiple review attributes. The heatmap uses color intensity to represent the number of reviews.

1. Data Preprocessing: The dataset is being preprocessed to convert and clean the data. Ratings for various categories (e.g., seat comfort, cabin staff service, etc.) are converted to numeric values, and rows with missing values are filtered out to ensure the integrity of the dataset.

2. Dimensionality Reduction via PCA: To capture the most significant patterns in the review data, PCA is applied to the numerical review attributes (e.g., seat comfort, food quality). This process reduces the dataset's dimensionality while preserving the key variance. The first principal component (PCA1) is extracted and assigned to each review.

3. Data Aggregation and Visualization: The data is then grouped by airline and rating, with the number of reviews counted for each combination. This aggregated data is used to generate the heatmap, where each cell represents a unique airline-rating combination, and the color intensity indicates the number of reviews.

4. Interactive Tooltips: A tooltip is displayed when hovering over the chart, offering brief details about the heatmap.

5. Color Legend: A color legend is included on the side of the heatmap to help users interpret the color scale, where lighter shades represent lower review counts, and darker shades indicate higher review frequencies.

4 Discovered Insights

Upon analyzing the dataset, several notable trends and insights emerged regarding airline ratings and performance.

First, for many airlines, it is difficult to draw definitive conclusions or identify clear trends. The data for certain airlines only spans a limited number of years, making it challenging to observe long-term patterns. With such restricted time frames, identifying a consistent trend in airline performance becomes unreliable, and decisions based on these limited data points would lack sufficient support.

Additionally, the overall ratings across airlines tend to be relatively low, with most ratings clustering in the lower range. High ratings are rare, making it clear that, generally, passengers' experiences are far from excellent. This suggests that many airlines are struggling to meet passengers' expectations.

Another key observation is that, over the years, there is a noticeable decline in ratings, which may indicate a deterioration in service quality, customer satisfaction, or other factors affecting the airline's overall performance.

Furthermore, some airlines, including Pegasus Airlines, Wizz Air, and easyJet, consistently appear as some of the poorer performers in most categories suggesting they receive lower ratings compared to competitors across multiple factors such as seat comfort, cabin staff service, and value for money.

These insights provide a deeper understanding of the airline industry's performance landscape and underscore the importance of continual improvement, especially for airlines lagging behind in customer satisfaction.

5 Related Work

Several studies and data sets provide valuable context for analyzing airline ratings and customer sentiment.

Masurah Mohamad[Mohamad(2023)] discusses how sentiment analysis can assess airline service quality by categorizing customer reviews as positive, negative, or neutral. This aligns with the project's focus on evaluating airline performance. Similarly, Eunil Park, Yeonju Jang, Jina Kim and Nam Jeong Jeong[Park et al.(2019)Park, Jang, Kim, and Jeong] investigate the factors influencing customer satisfaction in the airline industry, using sentiment analysis on a large dataset of feedback. Their work highlights key determinants of satisfaction, such as emotional responses and the distinction between low-cost and full-service carriers, which is relevant to our analysis.

Additionally, the Kaggle "Airline Reviews Dataset"[Bhojani and Pitroda(2024)] from airlinequality.com[skytrax(2024)] serves as the foundation for this project. It contains detailed airline reviews, enabling sentiment analysis, time series analysis, and trend prediction to uncover customer sentiment patterns.

6 Conclusion

By integrating a bar chart, line chart, and heatmap, the system allows users to explore trends in airline performance, ratings over time, and the distribution of reviews. The project incorporates a dimensionality reduction technique (PCA) to cluster airlines based on customer satisfaction.

The visualizations support users in making informed decisions, whether they are airline companies looking to benchmark their services or travelers evaluating potential flight options. Key insights from the analysis revealed that many airlines experience downward trends in their ratings, with most ratings being relatively low.

Overall, the interactive dashboard enhances the user experience by enabling dynamic exploration of airline data and facilitating data-driven insights into the quality of airline services. Future work could expand the dataset to include more global airlines and refine the system to incorporate additional predictive analytics.

References

[Bhojani and Pitroda(2024)] Juhi Bhojani and Khushi Pitroda. Airline reviews, February 2024. URL `https://www.kaggle.com/datasets/juhibhojani/airline-reviews/data`.

[Mohamad(2023)] Masurah Mohamad. A study on sentiment analysis on airline quality services: A conceptual paper. November 2023. URL https://www.researchgate.net/publication/375753151_A_Study_on_Sentiment_Analysis_on_Airline_Quality_Services_A_Conceptual_Paper.

[Park et al.(2019)Park, Jang, Kim, and Jeong] Eunil Park, Yeonju Jang, Jina Kim, and Nam Jeong Jeong. Determinants of customer satisfaction with airline services: An analysis of customer feedback big data. November 2019. URL https://www.researchgate.net/publication/336958586_Determinants_of_customer_satisfaction_with_airline_services_An_analysis_of_customer_feedback_big_data.

[skytrax(2024)] skytrax. airlinequality, February 2024. URL https://www.airlinequality.com/review-pages/a-z-airline-reviews/#mainMenu.