Final Report – Group 3
A Case Study of Swoop Airline

Members:
Priyanshi (ps4609), Moin (mk8793), Calvin (sh7787), Alex (tw2883)

TABLE OF CONTENT

The Motivation of Project

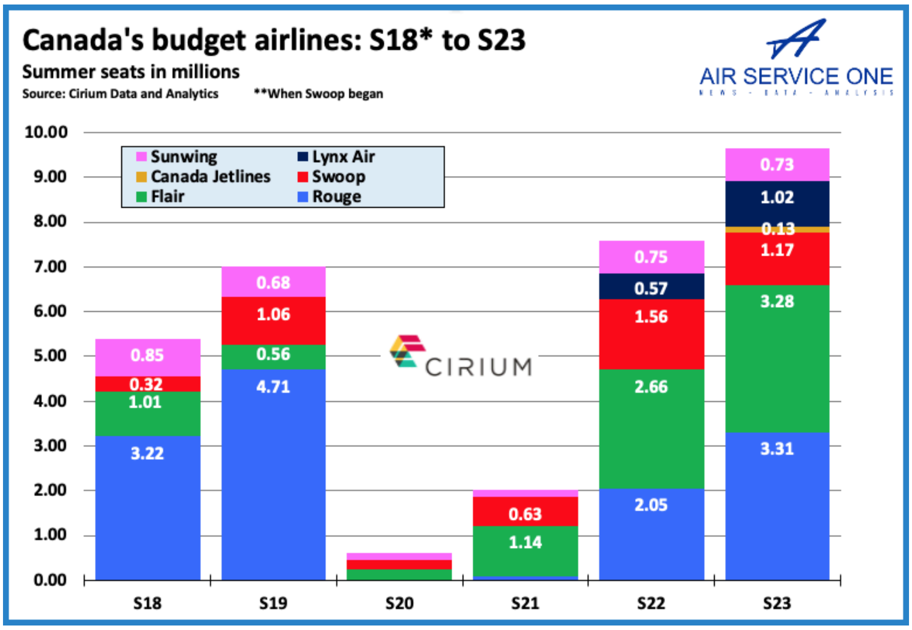
Swoop Airline, a Canadian low-cost carrier established in 2017, is set to cease operations in 2024, as announced by its parent company, WestJet. We will leverage a dataset encompassing customer ratings, which includes comprehensive reviews and ratings of airline flights, overall customer satisfaction, and recommendations. Our goal is to identify key predictors of positive recommendations to strategically enhance customer satisfaction. Additionally, we will conduct a comparative analysis of performance across various airlines to better understand our market position and refine our strategic approach by focusing on key competitors.

Business Understanding

Brief Introduction of Swoop Airline

Swoop Airlines, established as Canada's sixth-largest carrier in 2023, represents a strategic expansion in the ultra-low-cost aviation sector. Announced in September 2017 by WestJet, Canada's second-largest airline, Swoop commenced operations in June 2018, swiftly positioning itself as a key player in the industry. The airline employs over 400 individuals and manages a fleet of 16 aircraft, predominantly Boeing 737 and Boeing 737 Max models. While primarily servicing domestic routes, which account for 80% of its operations, Swoop also engages in international flights to the United States and Mexico, making up the remaining 20% of its service portfolio.

Business Challenges



As shown in the chart above, the entire market for Canada's budget airlines has been growing since 2018, except for the years 2020 and 2021 due to the impact of COVID-19. By 2022, the market had returned to pre-pandemic levels and experienced significant growth in 2023, increasing by 26% to reach a total of 9.5 million passengers.

Additionally, in terms of individual airline performance, the majority of carriers either improved or maintained their previous year's results. Despite the general upward trend, Swoop Airlines faced a decline, with its performance dropping by 25% in 2023, which translates to a reduction of approximately 0.39 million seats.

Data Understanding

Data Description

The dataset, titled "128K Airline Reviews" and sourced from Kaggle, comprises approximately 128 thousand data entries with a total size of 123 MB. It encompasses extensive reviews and ratings of airline flights, meticulously detailing aspects such as entertainment quality, food, seat comfort, overall customer satisfaction, and recommendations. This comprehensive collection serves as a valuable resource for in-depth analysis of consumer experiences and perceptions within the airline sector.

The following links are the dataset URL and the source code:

Dataset URL: https://www.kaggle.com/datasets/joelljungstrom/128k-airline-reviews
GitHub source code: https://github.com/alexwu0408/airline_reviews

Missing Values Handling

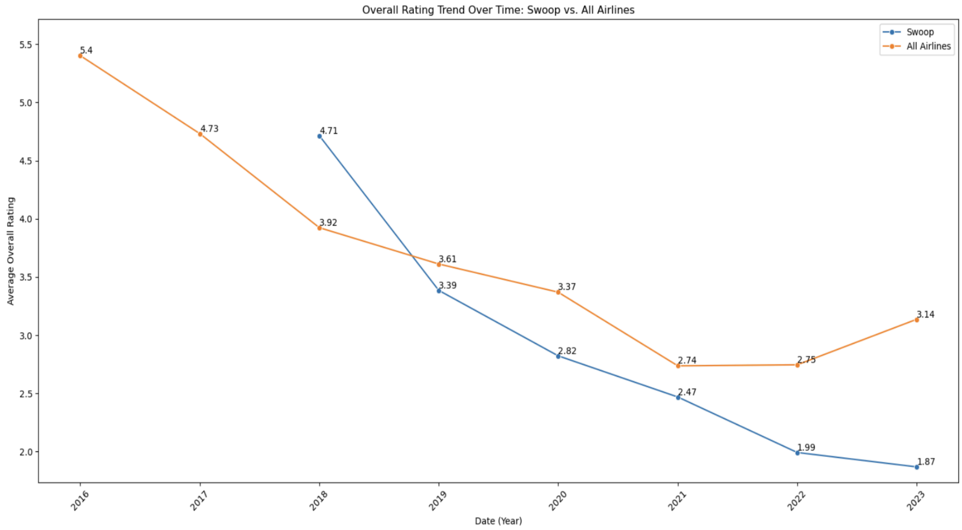
The raw data preprocessing involved standardizing formats, inferring missing values, and rectifying inconsistencies across various columns such as 'Aircraft', 'DatePubYearMonth', 'Dateflown', 'Review', 'TripVerified', and categorical columns. Missing values were filled appropriately, and textual comments were retained for analysis.

- **Column 'Aircraft':** The values in the 'Aircraft' column are quite chaotic, without a uniform format. For instance, the Airbus 320 model has variations such as 'A320', '320', '320 Airbus', and so on. To compound the issue, there is a high rate of missing values, at 71.86%. Fortunately, we found that each Aircraft Manufacturer typically follows a distinct naming convention, such as Boeing favoring names that begin with 7, and Airbus with 3. As a solution, we introduced a new column 'Aircraft Manufacturer' to infer the maker based on the model number. To tackle the problem of the 71.86% missing values, given their significant volume, we chose not to delete them but instead to randomly fill them in proportion to the non-missing values in the 'Aircraft Manufacturer' column.
- **Columns ' DatePubYearMonth ':** It indicates 'When the review was published,' and we have standardized its format to YYYY-MM-DD for easier further analysis.
- **Column 'Dateflown': '**Dateflown' signifies 'When the flight(s) occurred.' For the 29.71% of missing values, we are imputing this column with the year and month information from 'DatePubYearMonth'.
- **Column 'Review':** The 'Review' column contains customers' textual comments about their flight experience. As this column is neither categorical nor numeric and the missing values constitute only 0.64% of the data, we are opting to drop these missing entries directly.
- **Column 'Review':** The contents of the 'Review' column are presented in the format 'Departure to Destination'. Therefore, we utilized a split function to extract the departure and destination information, and stored them separately in two new columns named 'Route_Departure' and 'Route_Destination'.
- **Column 'TripVerified':** We are rectifying values that do not adhere to the format within the dataset. For example, 'Trip Verified, Trip Verified' is being corrected to 'Trip Verified', and 'Not Verified, Not Verified' to 'Not Verified'. Missing values are being filled randomly.
- **Columns 'CabinType', 'OriginCountry', 'Route', 'TravelType':** Since these columns are categorical and the missing values are a minor fraction of the overall data, we are filling them in with the mode (the most common value).
- **Column 'OverallScore':** With missing values comprising 3.35% of the data, we are imputing the median score.

EDA - Exploratory Data Analysis

GRAPH 1:

Analysis of Customer Satisfaction Trends in the Airline Industry with a Focus on Swoop Airlines



This report examines the customer satisfaction trends within the airline industry, highlighting the performance of Swoop Airlines in comparison to the industry average from 2016 to 2022. The data, sourced from industry-wide customer feedback, shows a notable divergence in satisfaction levels between Swoop Airlines and other airlines.

Industry-Wide Trends

From 2016 to 2021, the customer satisfaction ratings for all airlines (represented by the orange line in the graph) showed a downward trend. This period was characterized by increasing challenges in airline operations and a surge in travel demand, particularly stressing service capacities and affecting customer perceptions negatively. However, in 2022, there was a noticeable rebound in satisfaction levels, suggesting effective industry-wide adjustments to the operational challenges and improved customer experiences.

Swoop Airlines Performance

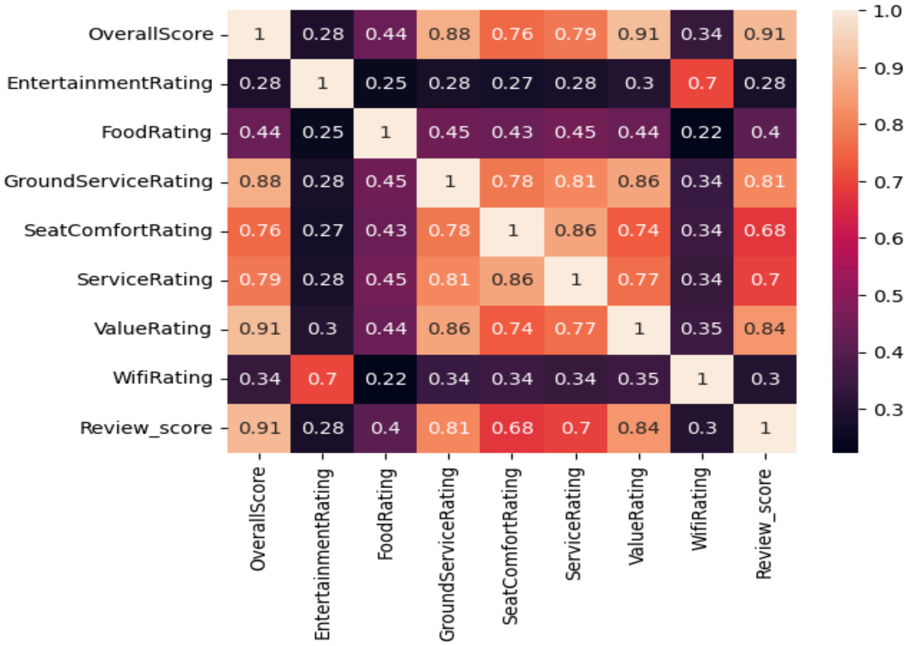
Contrasting sharply with the broader industry, Swoop Airlines, which began operations in 2018, is represented by the blue line in the graph. Unlike the industry trend, Swoop Airlines has not shown signs of recovery in customer satisfaction ratings. Since its inception, the airline has experienced a consistent decline in customer satisfaction, falling from an initial rating of 3.92 in 2018 to a significantly lower score of 1.87 by 2022. This decline is more pronounced and sustained than that observed in the broader industry.

Implications for Swoop Airlines

The data suggests that Swoop Airlines may be facing specific operational or service-related challenges that are not as prevalent in other airlines. The continuous decline in satisfaction scores, despite the industry's recovery in 2022, indicates deeper issues that may include customer service quality, reliability of service, pricing strategies, or passenger handling inefficiencies.

GRAPH 2:

Detailed Analysis of Correlation Patterns in Airline Service Factors



This section of the report presents a correlation analysis using a heatmap visualization, which explores the relationships between various service factors and overall customer satisfaction in the airline industry. The heatmap analysis provides a statistical basis to understand which aspects of airline service are most influential in determining overall customer satisfaction.

1. Value-Overall Satisfaction Correlation

The correlation analysis reveals a very strong positive correlation (0.91) between the perceived value (ValueRating) and the overall satisfaction score (OverallScore). This significant correlation underscores the critical impact of perceived value on customer satisfaction. Customers' perceptions of receiving good value for money are paramount in their overall satisfaction with the airline service.

2. Service Quality Correlations

Key service quality metrics including Ground Service (GroundServiceRating), Onboard Service (ServiceRating), and Seat Comfort (SeatComfortRating) are strongly correlated with the OverallScore, with correlation values of 0.88, 0.79, and 0.76, respectively. These factors are crucial in shaping the overall passenger experience, indicating that operational efficiencies in these areas are likely to enhance customer perceptions significantly.

3. Impact of Entertainment and Connectivity

Interestingly, Entertainment (EntertainmentRating) and WiFi services (WiFiRating), show weaker correlations with the OverallScore (0.28 and 0.34, respectively). This suggests that while these amenities contribute to the in-flight experience, they are less critical to overall customer satisfaction compared to core service factors.

4. Alignment between Customer Reviews and Overall Satisfaction

A very high correlation (0.91) between Review Scores (Review_score) and OverallScore was observed. This indicates that customer reviews are a direct and reliable reflection of overall satisfaction levels, emphasizing the importance of monitoring and managing online customer feedback as it closely aligns with broader satisfaction metrics.

Recommendations:

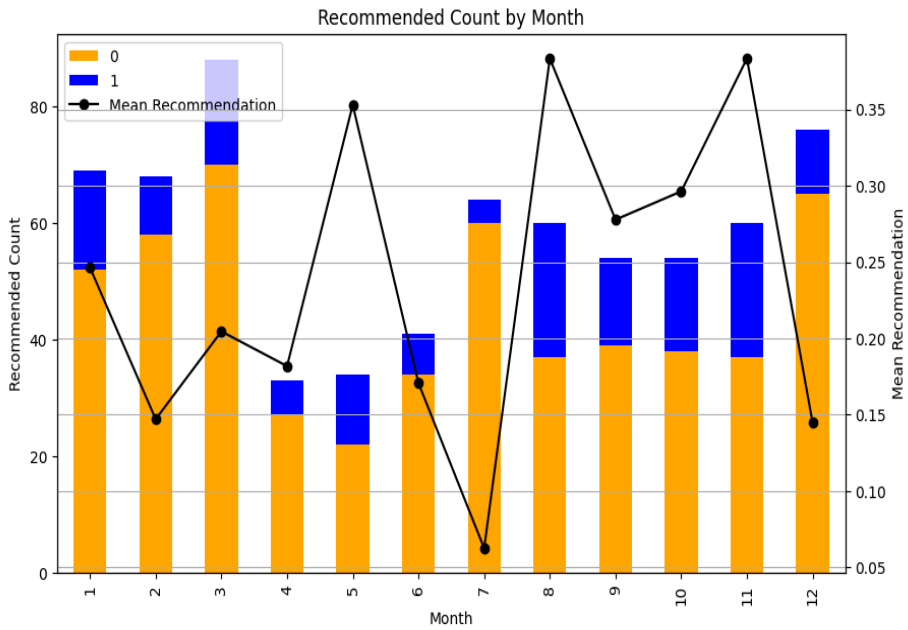
The heatmap analysis effectively highlights which aspects of airline service are most valued by customers. Airlines should focus on enhancing perceived value and core service quality areas such as ground service, onboard service, and seating comfort to boost overall customer satisfaction. Additionally, while entertainment and WiFi are important, they should not divert resources from more impactful areas. Lastly, the strong alignment between customer reviews and overall satisfaction scores highlights the need for airlines to engage actively with customer feedback platforms to monitor and respond to customer sentiments.

<u>GRAPH 3:</u>

Analyzing Trends in Customer Recommendations: Monthly and Yearly Distributions

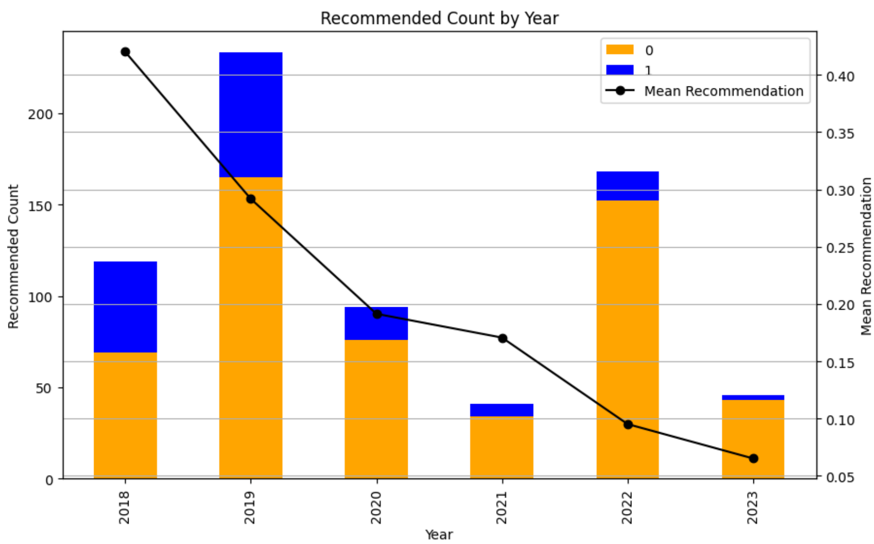
This report section provides a detailed examination of the distribution of customer recommendations on a monthly and annual basis, drawing insights from the data to inform strategic decisions.

Monthly Trends in Recommendations



The 'Recommended Count by Month' graph reveals significant variation in the number of recommendations (0 = not recommended, 1 = recommended) received each month. This is represented by the heights of the blue bars which denote positive recommendations. Notably, there are seasonal fluctuations that could be attributed to variations in customer experience or external factors influencing travel and service perceptions. The mean recommendation line, depicted in black, indicates the average level of recommendations per month, providing a quick visual assessment of over or underperformance relative to the monthly average.

Yearly Trends and Decline in Recommendations



The 'Recommended Count by Year' graph presents a more concerning trend. It is evident that there has been a consistent decline in the average number of recommendations since 2018. The yearly distribution shows a clear downward trajectory in both the total recommendations and the mean recommendation value. This decline suggests a potential erosion in customer satisfaction or changes in service quality, perceptions, or competitive dynamics over the years.

Implications and Strategic Considerations

The observed monthly variations suggest that while some months perform well above average, others significantly underperform, indicating potential opportunities for targeted improvements or promotional efforts to enhance customer experiences during lower-performing periods.

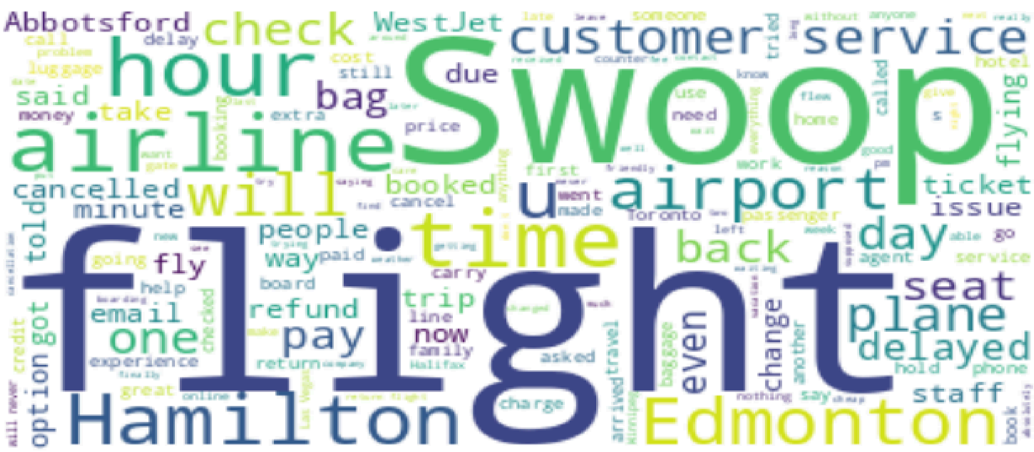
The yearly analysis, however, points to a more systemic issue that requires a thorough investigation. The consistent year-over-year decline in customer recommendations signals a need for a strategic review of operational practices, service offerings, and customer engagement strategies. Understanding the root causes driving this negative trend will be critical for implementing effective interventions.

Recommendations

1. Conduct a Root Cause Analysis: To understand the underlying factors contributing to the decline in customer recommendations, particularly analyzing changes in service delivery, customer expectations, and competitive pressures since 2018.
2. Enhance Customer Experience: Based on the insights from monthly data, develop strategies aimed at improving the customer experience during months with historically lower recommendation rates.
3. Monitor and Adapt: Regularly review customer feedback and recommendation trends to adapt services and offerings promptly to meet evolving customer needs and expectations.

<u>GRAPH 4:</u>

Insights from Word Cloud Analysis on Customer Feedback



This section of the report delves into the qualitative insights derived from a word cloud analysis, which synthesizes the main themes from customer feedback data for our airline services. The word cloud visually emphasizes the frequency and prominence of specific terms mentioned in customer reviews and feedback submissions.

Key Themes Identified:

1. **Customer Service Focus:**
   - The words "customer," "service," "airline," and "staff" appear prominently in the word cloud, indicating that these aspects are frequently mentioned in customer feedback. The size of these words suggests a high volume of comments centered around customer service

experiences. This underscores the critical importance of service quality in customer interactions and the significant impact it has on overall customer satisfaction.

2. **Operational Challenges:**
    - Terms related to operational efficiency such as "delay," "canceled," "time," "check," "line," and "refund" are notably visible. These words highlight common operational challenges faced by customers, including delays, cancellations, and issues with check-in processes. The frequent appearance of these terms points to areas where operational improvements are necessary.
3. **Specific Issues Highlighted:**
    - Other specific words like "booked," "cancel," "refund," and "ticket issue" reflect particular points of friction for customers. These terms often relate to the booking and ticketing process, suggesting areas where procedural enhancements or additional staff training may be required to improve the customer experience.

Recommendations:

1. Enhance Customer Service Training: Focus on training programs that empower staff to handle inquiries and complaints more effectively, ensuring that customer service representatives can address the concerns highlighted in the feedback.
2. Improve Operational Efficiency: Address common operational complaints by reviewing and enhancing procedures around scheduling, check-ins, and ticket handling to reduce delays and cancellations.
3. Streamline Booking and Refunds: Simplify the booking and refund processes to minimize customer frustration related to ticket issues and cancellations, possibly integrating more robust digital solutions to enhance user experience.

Modeling Appropriateness and Thoroughness

Establishing a Baseline Model for Predicting Customer Recommendations

In the process of building predictive models, it is essential to set a baseline to which more sophisticated models can be compared. This report outlines the creation and evaluation of such a baseline model using the Dummy Classifier provided by Scikit-learn, a popular machine-learning library.

Methodology

The baseline model is constructed by first transforming the 'Recommended' feature from the dataset into a binary variable. In this transformation, a response of 'yes' is encoded as 1 (indicative of a positive recommendation), and all other responses are encoded as 0. This binary variable serves as the target variable for the prediction model.

To set up the baseline, we employ the Dummy Classifier with the 'most_frequent' strategy. This strategy simplifies the model's prediction process by always selecting the most common label observed in the training data. This approach ensures that the baseline model predicts the majority class for every input.

Training and Testing the Model

The data is split into training and testing subsets to validate the effectiveness of the model. The Dummy Classifier is trained on the training data and subsequently used to make predictions on the test set.

Performance Evaluation

The performance of the baseline model is quantified using the accuracy metric, which measures the proportion of total correct predictions made by the model out of all predictions. The accuracy of this baseline model on the test set is approximately 76.8%. This indicates that the most frequent recommendation status, whether positive or negative, accounts for about 76.8% of the cases in the test data, providing a preliminary measure of model performance.

Implications

The established baseline model serves an essential function by providing a benchmark accuracy of 76.8%. Any advanced predictive model developed henceforth should aim to exceed this baseline performance to justify its complexity and deployment. Failure to significantly outperform the baseline model would suggest reevaluating the features used, the model chosen, or the data itself.

Analytical Approach and Model Performance Evaluation

This section of our report outlines the methodology and performance of three different predictive models—Logistic Regression, Decision Tree, and K-Nearest Neighbors (KNN)—used to analyze our dataset. The objective was to identify the most effective model for predicting customer recommendations based on a comprehensive set of data features.

Methodology

The models were applied as follows:

- **Logistic Regression:** A model used to estimate discrete values (such as binary values like 0 or 1) from a set of independent variables by using the logistic function. The optimal regularization strength parameter 'C' used was 10.
- **Decision Tree:** This model uses a tree-like graph of decisions and their possible consequences. It was optimized with a minimum of 70 samples required to split a leaf node.
- **K-Nearest Neighbors (KNN):** A method used for classifying cases based on a measure of distance (e.g., Euclidean) from the case to other cases with known categories. The model was configured with 141 neighbors for optimal results.

Results

The models' performances were primarily evaluated using the ROC AUC metric, which measures the area under the receiver operating characteristic curve. This metric is particularly useful for classification tasks with imbalanced classes. The results were as follows:

Model Performance Summary

Model	Optimal Parameter	ROC AUC (%)
Decision Tree	Leaf: 70	98.9
Logistic Regression	C: 10	99.338
K-Nearest Neighbors	Neighbors: 141	99.32

These scores indicate an excellent predictive capability across all models, with Logistic Regression slightly outperforming the others.

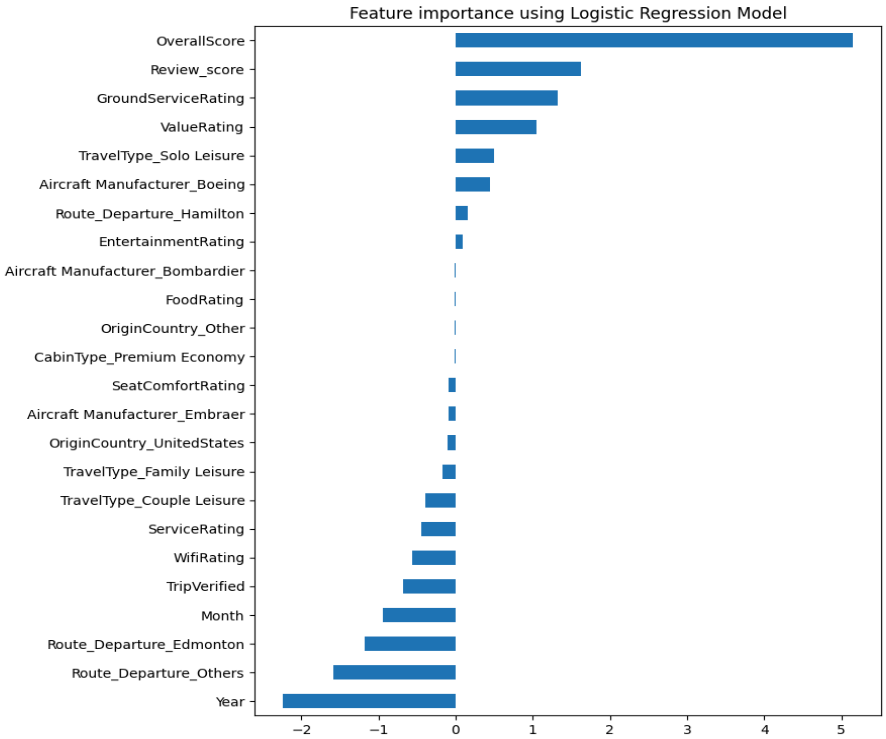
Implications

The high ROC AUC values demonstrate that all three models are highly effective in predicting customer recommendations using the available features. Logistic Regression, in particular, shows slightly higher performance, suggesting a better fit for handling binary classification in our dataset.

Given these results, the Logistic Regression model can be recommended for operational use, although the KNN and Decision Tree models also provide valuable insights and could be used for further analysis or specific scenarios where interpretability (as provided by the Decision Tree) or the instance-based learning approach of KNN might be advantageous.

Evaluation of Feature Importance and Data Imbalance Issues

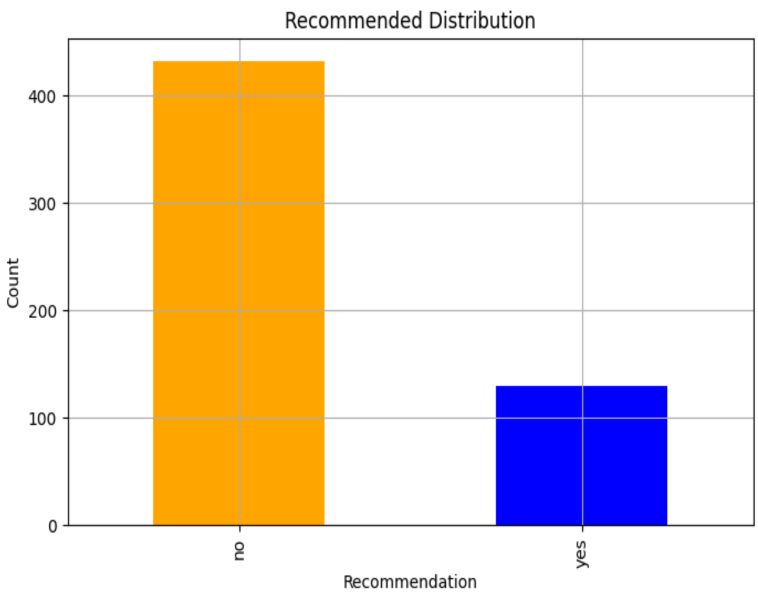
Feature Importance Analysis



The logistic regression model has highlighted several key features that significantly influence the prediction of customer recommendations. The OverallScore and Review_score are the most influential features, indicating that general customer satisfaction and specific feedback scores are critical in predicting whether a customer would recommend our services. Other important factors include GroundServiceRating and ValueRating, which suggest that the quality of ground services and perceived value for money are also crucial determinants.

Lesser but noteworthy influences come from specific travel types like TravelType_Solo_Leisure and aircraft manufacturers such as Boeing and Bombardier, which might reflect customer preferences or experiences related to different travel contexts and equipment used.

Data Imbalance and Mitigation



Initially, the dataset exhibited a significant imbalance in the distribution of the target variable 'Recommendation', with only 22% of the data representing positive recommendations (Yes) and a dominant 78% representing negative (No). Such an imbalance can bias a predictive model towards the majority class, potentially undermining its performance, especially in predicting the minority class.

To address this, sampling techniques were applied to balance the data within the training set, resulting in an equal representation of both classes (50% Yes, 50% No). This balance is crucial for training our models in a way that equally represents the outcomes, improving the model's ability to generalize and thus enhancing its predictive accuracy across unseen data.

Implications

The analysis of feature importance provides valuable insights into factors that most impact customer satisfaction and recommendation rates. Understanding these factors allows for targeted improvements in service areas that are most likely to influence customer recommendations positively.

Furthermore, the resolution of data imbalance through sampling techniques is expected to enhance the robustness of our predictive models, ensuring they perform well across different customer scenarios without bias towards the more frequently occurring class.

Evaluation: Metrics used and Interpretation

Evaluation of Predictive Models Using Balanced Data (Approach 2)

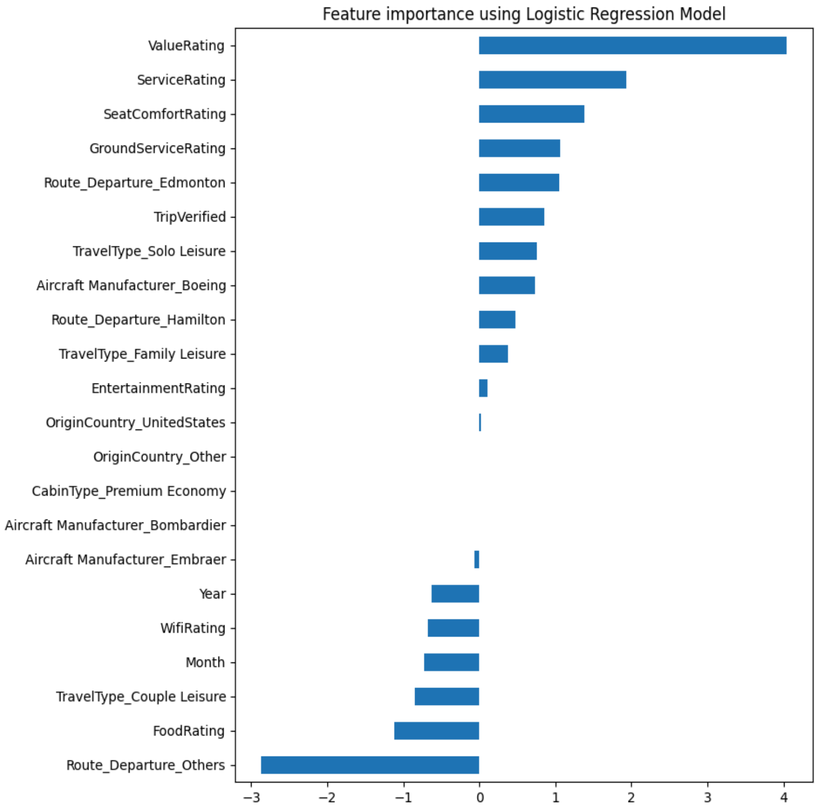
In this approach, we continue our analysis using balanced data to mitigate the bias introduced by imbalanced class distributions identified in previous analyses. This section details the performance of three predictive models: Logistic Regression, Decision Tree, and K-Nearest Neighbors (KNN), along with an analysis of feature importance as identified by the Logistic Regression model.

Model Performance

The models were evaluated on their ability to predict customer recommendations using the ROC AUC metric, a reliable indicator of classification performance. The results are as follows:

Model	Optimal Parameters	ROC_AUC
Logistic Regression	10	97.618
Decision Tree	10	98.974
KNN	121	99.571

Feature Importance Analysis



Using the Logistic Regression model, we identified several key features that have a significant impact on predictions. The most influential features include:

- **ValueRating:** The highest positive influence, indicating that customers' perception of value significantly affects their likelihood to recommend.
- **ServiceRating and SeatComfortRating:** Both show strong positive influences, underscoring the importance of service quality and comfort in customer satisfaction.
- **GroundServiceRating:** Also a significant positive contributor, highlighting the importance of the quality of ground services.

Less impactful but still notable features include various travel types and aircraft manufacturers, which suggest specific preferences or experiences might slightly affect customer recommendations.

Implications

The balanced approach to data has allowed for a more equitable representation of class labels, leading to high model performance across all three tested models. The insights from feature importance highlight critical areas for operational improvements and customer service enhancements.

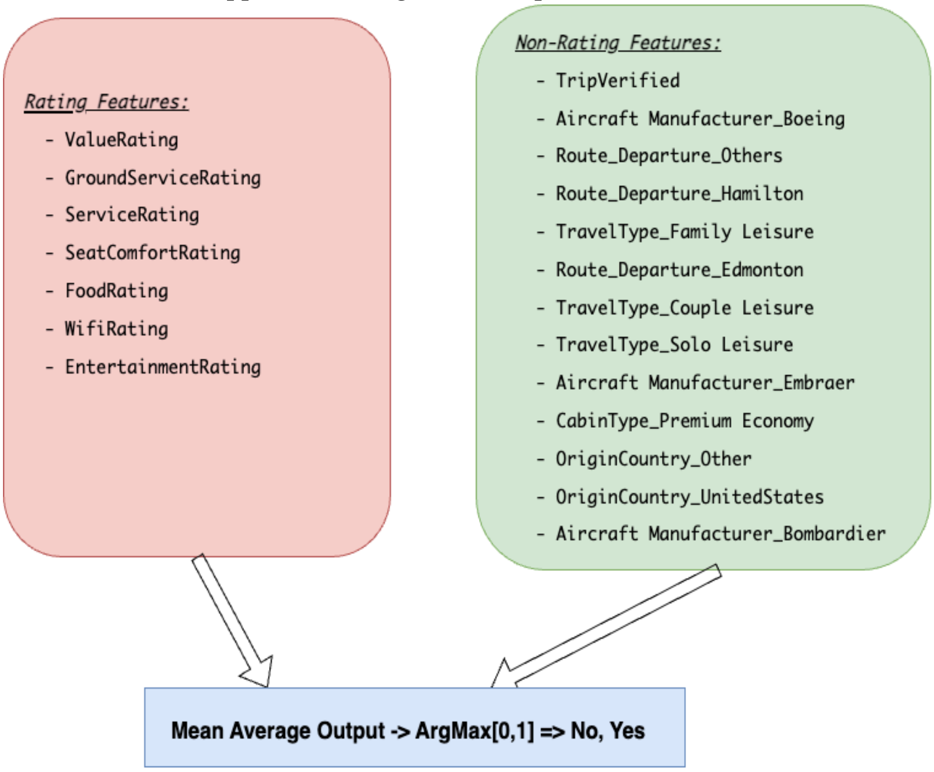
Given the superior performance of the KNN model and the significant features identified by the Logistic Regression model, these findings can guide strategic decisions to enhance customer satisfaction and recommendation rates. Furthermore, continuous monitoring and rebalancing of the dataset are recommended to maintain the efficacy of the predictive models.

Analysis of Ensemble Averaging Models in Predicting Customer Recommendations

In a strategic effort to enhance predictive accuracy, an ensemble model consisting of two distinct Random Forest models was employed. This section outlines the methodology, feature importance, and the impact of using balanced versus non-balanced data on model performance.

Model Configuration

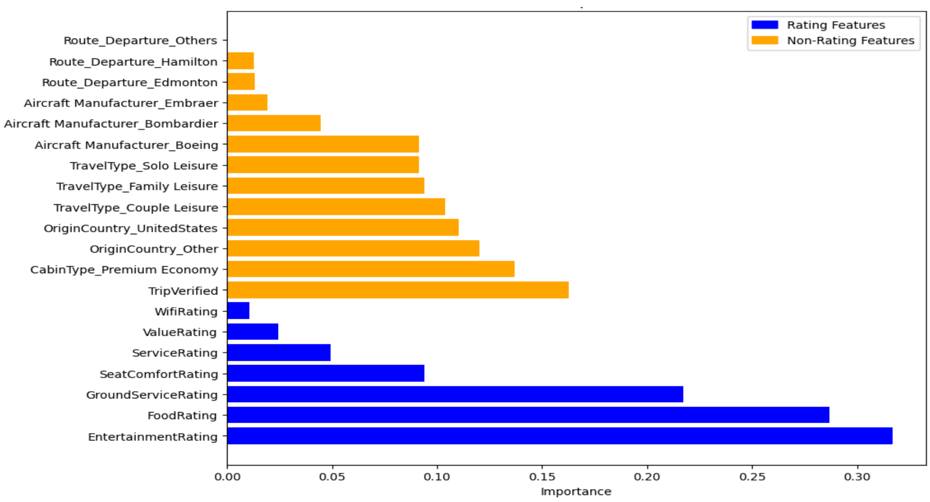
The ensemble approach integrated outputs from two Random Forest models:



- **Random Forest Model 1:** Focused on 'Rating Features' including ValueRating, GroundServiceRating, ServiceRating, SeatComfortRating, FoodRating, WiFiRating, and EntertainmentRating.
- **Random Forest Model 2:** Utilized a broader set of 'Non-Rating Features' such as TripVerified, aircraft manufacturer, route departure specifics, travel type, and other categorical variables like cabin type and origin country.

The outputs from these models were combined by averaging to predict customer recommendation, using a mechanism where the mean average output determines the final class (Recommend or Not Recommend).

Feature Importance Insights



The feature importance extracted from the ensemble model indicates a significant variance in the influence of different features:

- Among the 'Rating Features', ServiceRating, GroundServiceRating, and ValueRating were the most influential, highlighting their critical role in affecting customer satisfaction and subsequent recommendations.
- 'Non-Rating Features' such as Route_Departure_Edmonton and Aircraft Manufacturer_Embraer also showed substantial impact, suggesting that logistical and equipment-related aspects significantly influence customer perceptions.

Model Performance

The ensemble model demonstrated an initial test accuracy of 78% using non-balanced data. After applying techniques to balance the data, the test accuracy improved significantly to 87%. This improvement underscores the effectiveness of balancing the dataset in mitigating bias towards the majority class and enhancing the model's ability to generalize across varied scenarios.

Implications

The use of an ensemble averaging model leveraging both rating and non-rating features has proven effective in capturing a wide spectrum of factors influencing customer recommendations. The substantial improvement in accuracy with balanced data highlights the need for careful data preprocessing to ensure model reliability and validity.

Similarity: Competitors Navigation

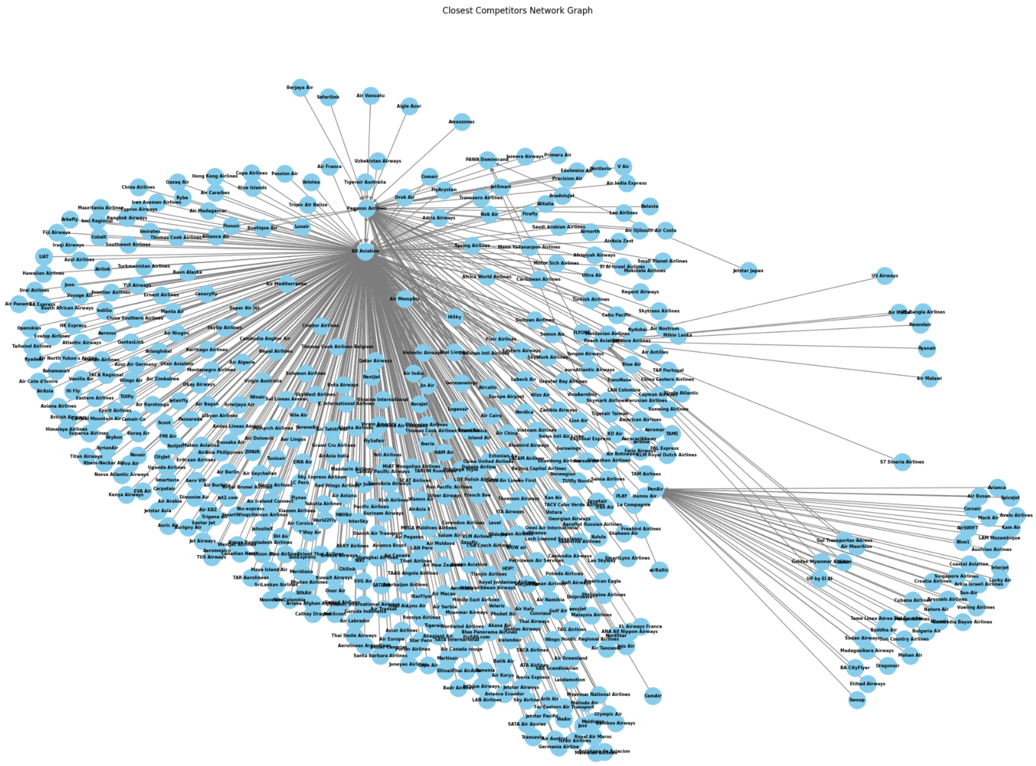
We also did airline performance benchmarking for the top 32 companies, to facilitate CEOs to make strategic decisions based on the moves of their top competitors!

Necessity

In our analysis of airline performance, we initially encountered a dataset featuring 547 unique companies. However, the distribution of flight records varied significantly across these companies. While some major players like "American Airlines" boasted 5456 entries, smaller airlines such as "Borajet" had only a single flight entry. To ensure our benchmarking efforts were impactful, we opted to focus solely on companies with a substantial presence in the dataset, specifically those with 1000 or more flight records. Despite this criterion, we made an exception for Swoop Airline, a key target for analysis, resulting in a final selection of 33 companies out of 547 for our benchmarking study.

Methodology

Our benchmarking analysis primarily revolves around key performance indicators (KPIs) related to passenger experience, including 'FoodRating', 'EntertainmentRating', 'GroundServiceRating', 'SeatComfortRating', 'ServiceRating', 'ValueRating', 'WifiRating', and 'OverallScore'. While acknowledging that other factors could influence performance, we focused solely on these ratings due to data availability.
To quantify performance, we calculated the mean ratings across these KPIs for each of the 33 selected companies. We then utilized cosine similarity to assess the similarity between companies based on these performance metrics. Additionally, we visualized the results through graph analysis, identifying clusters of interconnected nodes that may indicate regional markets or alliances within the airline industry.



Results

Our analysis revealed interesting insights into the performance landscape of the top airlines. **The closest two companies to Swoop Airline are" Penair" and "Pakistan Intl Airlines".** By focusing on passenger-centric KPIs, we were able to identify trends and patterns that can inform strategic decisions for airline CEOs. The benchmarking process highlighted areas of strength and weakness for each company, providing valuable insights into areas for improvement and potential competitive advantages.

It's important to acknowledge the limitations of our analysis, primarily stemming from the constraints of the dataset. While our current approach provides valuable insights, future iterations could benefit from additional data sources and more sophisticated analytical techniques. **If given more data, we can do a more comprehensive and dedicated performance benchmarking in the North American market,** in this way, we can analyze Swoop's competitors more thoroughly. Moving forward, we aim to expand our dataset and refine our methodology to ensure a more comprehensive and robust evaluation of airline performance.

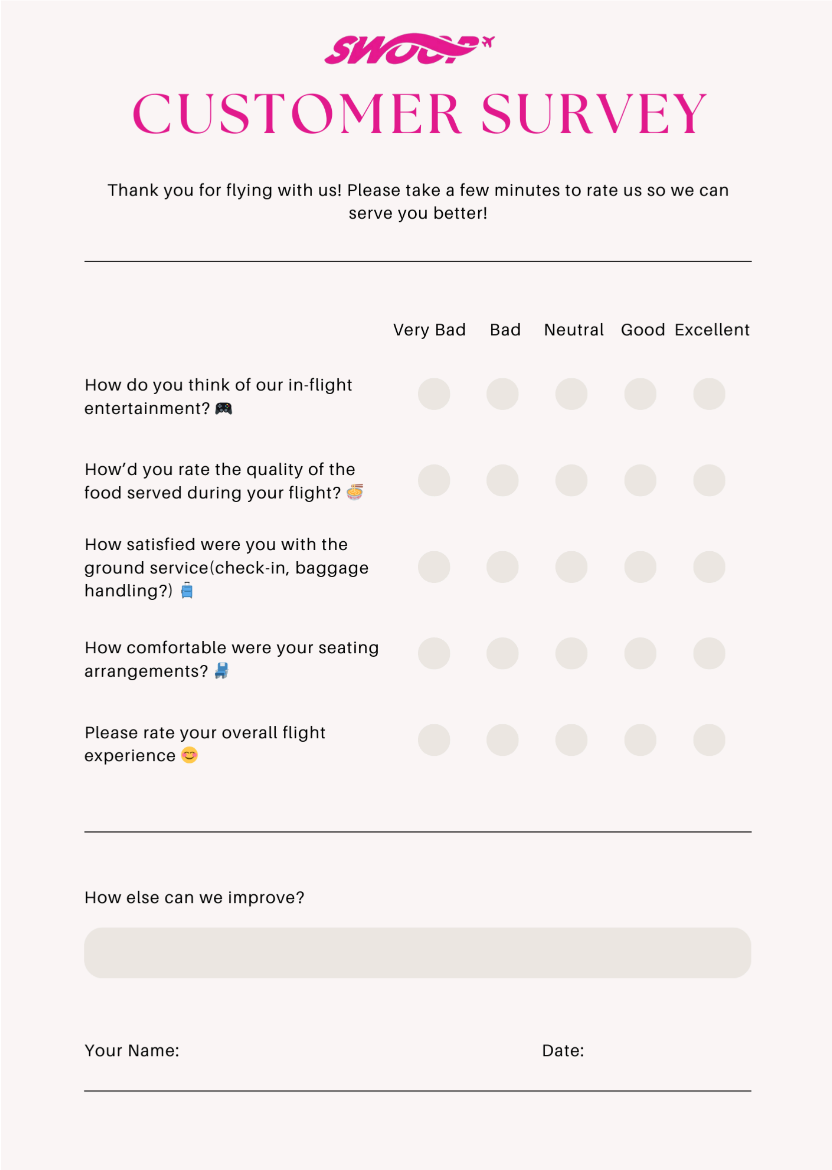
Business Strategies

Strategy I: Strategic Customer Surveys

In today's competitive airline industry, understanding and meeting customer needs are paramount for driving satisfaction and loyalty. To achieve this, airlines must design efficient customer surveys that focus on the most critical features. This strategy section outlines the importance of such surveys and their impact on customer satisfaction, loyalty, and market positioning.

Central to this endeavor is the design and implementation of efficient customer surveys. These surveys serve as invaluable tools for capturing the pulse of the customer base, providing airlines with actionable insights into the aspects of their service that matter most to passengers. By focusing on the most critical features through targeted survey design, airlines can effectively prioritize areas for improvement and innovation, thereby enhancing the overall passenger experience.

Based on the **“Feature importance insights”** section, we will design a survey like below:



SWOOP

CUSTOMER SURVEY

Thank you for flying with us! Please take a few minutes to rate us so we can serve you better!

	Very Bad	Bad	Neutral	Good	Excellent
How do you think of our in-flight entertainment? 🎮					
How’d you rate the quality of the food served during your flight? 🍜					
How satisfied were you with the ground service(check-in, baggage handling?) 🧳					
How comfortable were your seating arrangements? 💺					
Please rate your overall flight experience 😊					

How else can we improve?

Your Name: Date:

*(*It’s a survey demonstration, the survey will be conducted in the airline app in a real-life scenario.)*

- Designing surveys around the five most crucial features(entertainment, food, ground service, seat ratings) allows airlines to address customer needs effectively.
- Customers feel valued when their input is considered, leading to enhanced satisfaction and loyalty.
- Survey data informs decision-making processes, guiding resource allocation and service, and maximizing return on investment.
- Iterative use of survey data ensures continuous improvement and alignment with market expectations.

By identifying and prioritizing key features, airlines can tailor their services to meet customer expectations. On top of that, strategic surveys enable airlines to pinpoint areas for improvement, leading to enhanced satisfaction levels. Additionally, addressing customer needs directly contributes to increased loyalty and positive word-of-mouth.

Strategy II: Find the Edge - Market Positioning

In the fiercely competitive landscape of the airline industry in the post-COVID era, understanding how an airline stacks up against its closest competitors is crucial for strategic decision-making and long-term success.

By conducting a thorough analysis of competitive positioning, airlines can uncover their unique selling propositions (USPs) and identify areas where they excel compared to rivals. This knowledge not only informs marketing strategies but also empowers airlines to accentuate their strengths effectively in the eyes of consumers.

After achieving performance benchmarking, we put forward:

- Comparing the airline with its closest competitors provides insights into its unique selling propositions.
- Understanding competitive positioning guides marketing strategies to highlight the airline's strengths.
- Market positioning analysis helps differentiate the airline from competitors and target specific customer segments effectively.
- Insights gleaned from surveys enable airlines to refine service delivery and brand messaging.

Moreover, market positioning analysis serves as a compass for differentiation, allowing airlines to carve out a distinct identity amidst a sea of competitors and tailor their offerings to resonate with specific customer segments. Harnessing insights gleaned from customer surveys, airlines can refine their service delivery and brand messaging to better align with market expectations, ultimately driving customer satisfaction and loyalty. In this dynamic industry landscape, a deep understanding of competitive positioning is essential for airlines to thrive and maintain a competitive edge.

In conclusion, strategic customer surveys play a vital role in enhancing both customer satisfaction and market positioning for airlines. By focusing on key features, airlines can better understand customer needs, allocate resources efficiently, and differentiate themselves in a competitive market is also a ruler to help companies realize their role in the game. Continuous utilization of survey insights and attention to the moves of their competitors enables airlines to stay attuned to evolving customer preferences and maintain a competitive edge in the industry.

Future Endeavors

Deployment and Usage

- Integrate the predictive model into airline Crew Resource Management(CRM) systems.
- Enable real-time customer interaction and satisfaction tracking.
- Ensure continuous model performance monitoring and regular updates.

In the deployment and usage phase, our primary objective is to seamlessly integrate the predictive model into airline Crew Resource Management systems. This integration will facilitate real-time customer interaction and satisfaction tracking, empowering airline staff to respond promptly to passenger needs and preferences. Additionally, we prioritize continuous model performance monitoring and regular updates to ensure sustained relevance and accuracy. This iterative process allows us to adapt to evolving customer preferences and market dynamics, maximizing the utility of the predictive model in enhancing customer satisfaction and loyalty.

Data Challenges in the Wild

- Utilize ongoing customer feedback, flight data, and service interaction records.
- Address challenges like data drift and quality discrepancies.
- Ensure consistent data collection across all customer touchpoints.

Addressing data challenges in real-world applications requires a comprehensive strategy. We will leverage ongoing customer feedback, flight data, and service interaction records to mitigate issues such as data drift and quality discrepancies. Through meticulous data validation techniques and stringent quality control measures, we can maintain data integrity and reliability. Moreover, consistent data collection across all customer touchpoints is paramount to preserving the model's effectiveness. By implementing robust data management practices, we will ensure that the predictive model remains a dependable tool for accurately predicting customer needs and delivering personalized experiences.

Bias and Fairness Considerations

- Monitor for biases arising from non-representative training data.
- Adjust the model to accurately reflect diverse experiences and expectations.
- Promote fairness and inclusivity in the predictive modeling approach.

In our commitment to fairness and equity, we will prioritize the identification and mitigation of biases that may arise from non-representative training data. This involves rigorous monitoring for demographic or route-specific biases and proactive adjustments to the model to accurately reflect the diverse experiences and expectations of all passenger segments. By promoting fairness and inclusivity in our predictive modeling approach, we can ensure that our solutions resonate with all passengers, regardless of background or demographic factors!

Reference

- Dataset URL: https://www.kaggle.com/datasets/joelljungstrom/128k-airline-reviews
- GitHub code: https://github.com/alexwu0408/airline_reviews
- Presentation Slides: https://docs.google.com/presentation/d/