**The impact of the COVID-19 pandemic on airlines’ passenger satisfaction**

<https://www.sciencedirect.com/science/article/pii/S0969699723000844#bib20>

Written reviews were analyzed using a text mining tool. The critical features to the flight experience were identified by evaluating the frequency at which each feature was mentioned. The most important attributes according to this paper are staff behavior, employee attitude towards dissatisfied customers, booking and cancellation policies, baggage handling, seat quality, boarding, check in, customer service, and food and drink.

A related study that is mentioned in this paper concluded that the most impacted passenger satisfaction was the queuing time, lounge comfort, cabin crew quality, and seat legroom. Of these features only seat legroom is included as a feature in our dataset. This could be a significant gap in our findings derived from our main dataset.

Other related studies have identified cabin crew quality as one of the most critical factors in passenger satisfaction.

This study concluded that the COVID-19 pandemic didn’t change which factors influence passenger satisfaction but did give an increased importance to refunds and cabin cleanliness. Our data was taken before the COVID-19 pandemic and thus may not reflect the heightened importance of cleanliness.

The data was taken from European airlines so there could be a cultural difference with our dataset since ours is taken from an American airline.

**Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews**

<https://www.sciencedirect.com/science/article/abs/pii/S0969699719302959>

The authors of this paper utilized Latent Dirichlet Allocation (LDA) to analyze over 55,000 written customer reviews to identify the topics that most influence customer satisfaction. The most important features identified were cabin staff, onboard service, value for money, and seats which is consistent with the findings in the other research papers.

The authors also found that certain segmentations can significantly change what customers find valuable. Some segmentations that showed significant differences were nationality, likely due to cultural reasons, and cabin flown. The practical implications to maximize customer satisfaction by cabin flown are: focus on customer service for first class passengers, comfort for premium economy passengers, and checking luggage and waiting time for economy class travelers.

This study used a database with over 55,000 online customer reviews, and covered over 400 airlines and passengers from over 170 countries. The authors claim that by having a much larger dataset that they have attained more reliable generalizations.

**An Optimized Deep Learning Approach for Improving Airline Services**

<https://www.techscience.com/cmc/v75n1/51460/pdf>

The author of this paper analyzed the same dataset that we are using with the purpose of creating the highest accuracy model possible. While the goals between our research and this author’s are different, the findings serve as an interesting comparison to our models.

An interesting approach that the author took for figuring out which features mattered most was to look at the correlation between satisfaction and the 24 features. The author concluded that the most important factors are online boarding, class, type of travel, and in-flight entertainment. Likewise the least important factors are age, gender, departure/arrival time convenience, departure delay, and arrival delay which the author decided to exclude from the model.

The techniques that the author tried were deep neural network with Adam (99.3% accuracy), artificial neural network with Adam (95% accuracy), support vector machine (95.19% accuracy), and random forest (95.90% accuracy).

**Research Summary**

Two of the papers that we investigated took a novel approach to customer satisfaction of analyzing publicly available written reviews and using natural language processing techniques to figure out which features airline customers care about. This seems like a more reliable approach because it doesn’t limit the possible factors that may turn out to be important. On the other hand the surveys such as the one from our main dataset assumes that the factors that matter are known and merely gives customers the options to rate those factors. This is a problem because it implies that our model could be missing a critical factor.

For instance cabin crew quality has been found to be a top contributor to customer satisfaction across multiple research papers on airline passenger satisfaction and it is not considered in our dataset. As a counter-example, seat legroom is also a top contributor to customer satisfaction across multiple research papers and it is included in our dataset.

Another concern with the approach taken by our dataset which is mentioned in multiple research papers is that there is a tendency for customers to not take questionnaires seriously. This means that many customers fill them out randomly which just creates noise in the data.