Ukrainian Catholic University

Faculty of Applied Sciences

Business Analytics & Computer Science Programmes

Econometrics
Final report

Beyond the clouds

Team:
Daryna Kuzyshyn
Olha Havryliuk
Anastasiia Dynia

26th of April 2024

1. Introduction

In today's world, what customers say online can really make or break a brand, and this is especially true for airlines. The airline business is all about providing great service and making sure passengers are happy. But as the world changes, especially with everyone using the internet more, airlines have to keep up not just with flying planes but also with what passengers expect and say about them online. Passengers' reviews and stories can now reach people all over the world, influencing others to choose or avoid an airline. Understanding why passengers are not satisfied is now key for airlines to get better and improve their services.

In this project, the authors are diving deep into what passengers say in their reviews to figure out what are the drawbacks of the services. It is essential to carefully look at the reviews to see what parts of the flight service stand out, either in a good or bad way. The authors are interested in everything from how comfortable the seats are to how friendly the flight attendants are. The authors are also willing to make their research as representative as they could.

Also, the authors will compare different airlines to see how they stack up according to passenger reviews. This will help to see what some airlines are doing right and where others might need to improve.

2. Literature review

Before doing the research and starting to formulate some hypotheses and assumptions, the authors analyzed the airline industry and gained some insights. First of all, since the dataset that is considered contains reviews from 2012 to 2023, one should analyze the impact of the COVID-19 on this type of industry. And, unfortunately, the pandemic had a huge influence on the airline services, resulting in a dramatic drop in demand and the change in the behaviour of passengers. The researches and articles that consider those topics are: "Pre- and post-COVID-19 condition performance", "The impact of the COVID-19 pandemic on airlines' passenger satisfaction". There is also another thing that need to be included: the duration, thus the distance of the flight really impact the overall score of the satisfaction of the passanger. So, the article about "short-haul, medium-haul and long-haul flights" was considered. During the work on the project, the authors also referred to "A guide to modern econometrics" to get the necessary theoretical information.

3. Data description and analysis

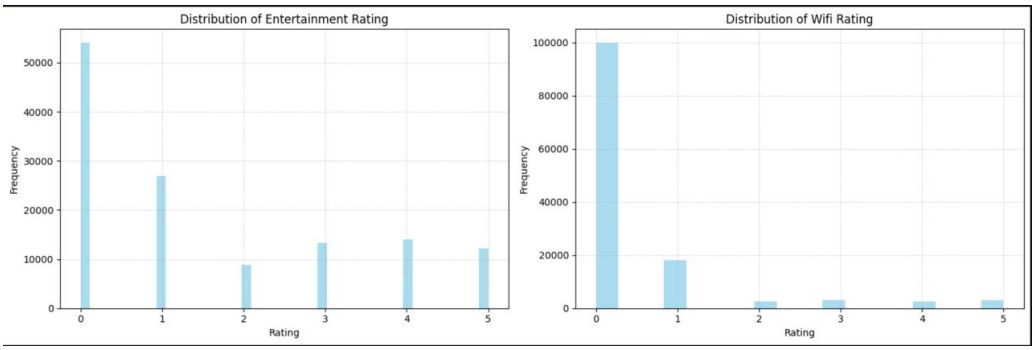
The project utilizes a dataset that is compiled from an online airline review platform - `https://www.airlinequality.com`, it provides a diverse range of passenger feedback on various airlines and flight experiences. This dataset has detailed reviews that are spanned from 2012 to 2023, offering an extensive view of passenger satisfaction and service quality in the airline industry during this period. It contains more than 129 thousands of rows and is **a time-series type of data**.

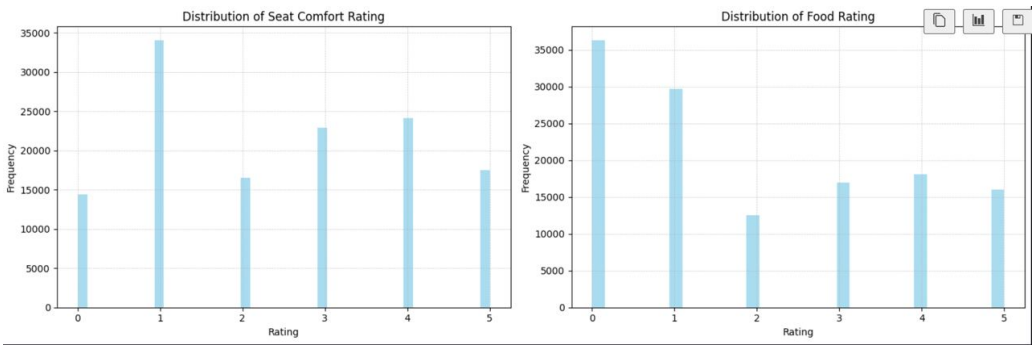
Key variables from the dataset:

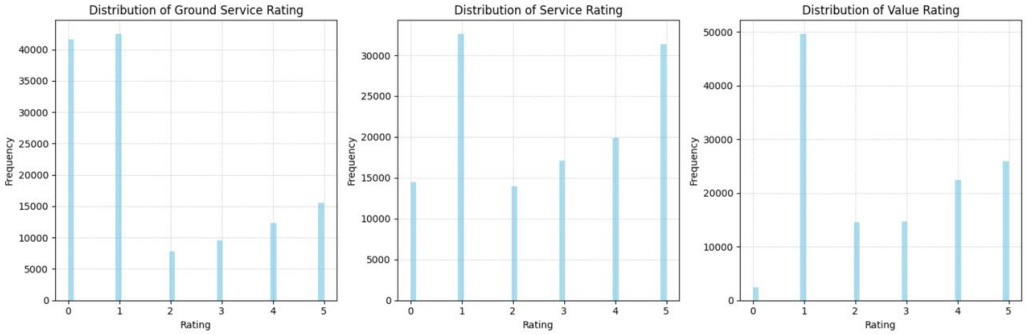
- **AirlineName:** the name of the airline being reviewed.

- **CabinType:** the class of service used by the reviewer (Economy, Business, First Class).
- **DateFlown:** the month and year of the flight.
- **EntertainmentRating, FoodRating, GroundServiceRating, SeatComfortRating, ServiceRating, WifiRating:** passenger ratings for each of these aspects on a scale from 1 to 5.
- **ValueRating:** this variable describes passengers' assessment of the value for money of their flight experience, rated on a scale from 0 to 5.
- **TripVerified:** a boolean variable indicating whether the review is verified.
- **OriginCountry:** the country from which the flight originated.
- **OverallScore:** an overall score for the flight experience on a scale from 1 to 10.
- **Recommended:** a boolean variable indicating whether a passenger recommends the airline ("yes" or "no").
- **Review:** textual review of the flight.
- **Route:** flight route.
- **TravelType:** the type of travel (e.g., Solo Leisure, Business). Different travel types can have distinct expectations and experiences.

Now the authors will proceed with exploratory data analyzing. At first they will visualize the general distribution of rating variables through all years.

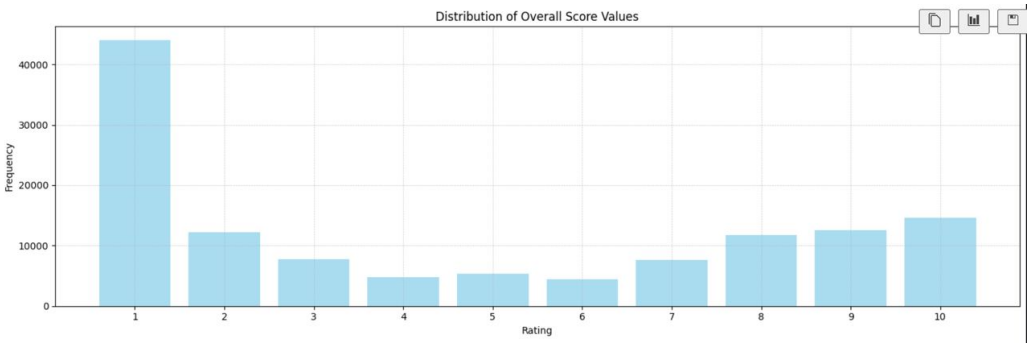






From the plots of rating columns it can be seen that 0 or 1 are the most frequent values. However, there is a big amount of high values, indicating that the passengers were satisfied with the provided service. The intermediate values such as 2 and 3 are in minority.

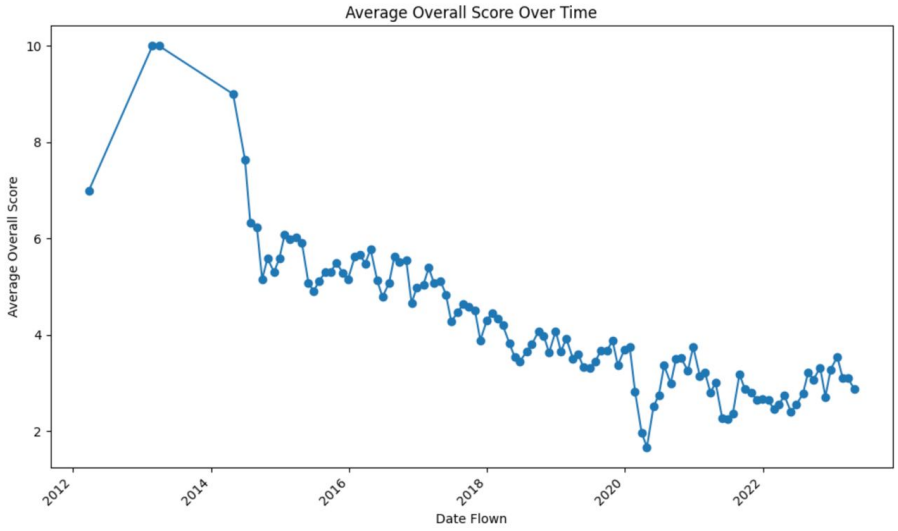
The authors also consider the overall score, as it is a comprehension measure of customer satisfaction.



From the distribution of the values in the overall score column, the similar pattern as in the distribution of values in the rating columns can be observed. The most frequent scores are values from 1 to 3. Also there are many high scores (values from 8 to 10), despite the fact that the intermediate scores(values from 4 to 7) are still in minority.

This can be explained by the fact that those people who left comments are either disappointed with their flight experience or vice versa: the flight conditions were very good and they were satisfied. Usually, when people receive good service, they do not want to share their feedback.

The next step is to visualise the overall score over a years in order to see whether there is a pattern in the data



From this plot the authors can conclude that the average value of the overall score decreases over time. It became especially low after a sharp drop in 2020. Based on related article to this topic, this can be explained by COVID-19 and the restrictions caused by it.

When conducting the initial study of the field in which the authors are currently working, they realized that COVID-19 restrictions have changed the usual way of air transportation. As the requirements to the flight have changed, it isn't appropriate to consider the data before COVID-19, because it isn't representative. The data from the 2021 to 2023 years will be taken for the research.

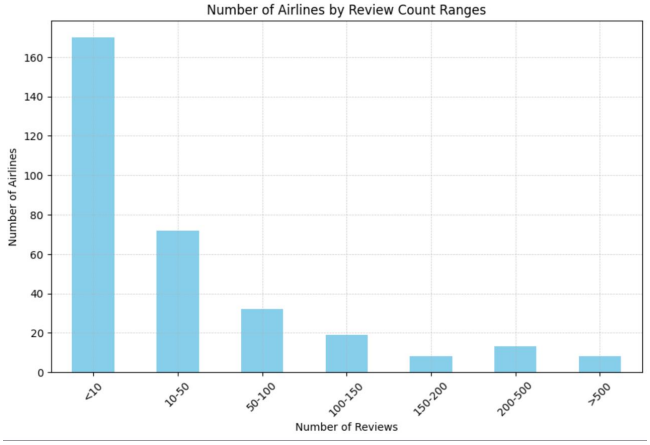
After cropping the dataset there were only 22 thousands rows left. As it was already mentioned, the dataset contains columns that represent different cabin type and travel type categories. Let's explore how they are related.

Count of people for each TravelType in each CabinType:

	TravelType	CabinType	Count
0	Business	Business Class	558
1	Business	Economy Class	2067
2	Business	First Class	95
3	Business	Premium Economy	137
4	Couple Leisure	Business Class	483
5	Couple Leisure	Economy Class	4848
6	Couple Leisure	First Class	101
7	Couple Leisure	Premium Economy	293
8	Family Leisure	Business Class	216
9	Family Leisure	Economy Class	5402
10	Family Leisure	First Class	49
11	Family Leisure	Premium Economy	204
12	Solo Leisure	Business Class	567
13	Solo Leisure	Economy Class	6809
14	Solo Leisure	First Class	123
15	Solo Leisure	Premium Economy	278

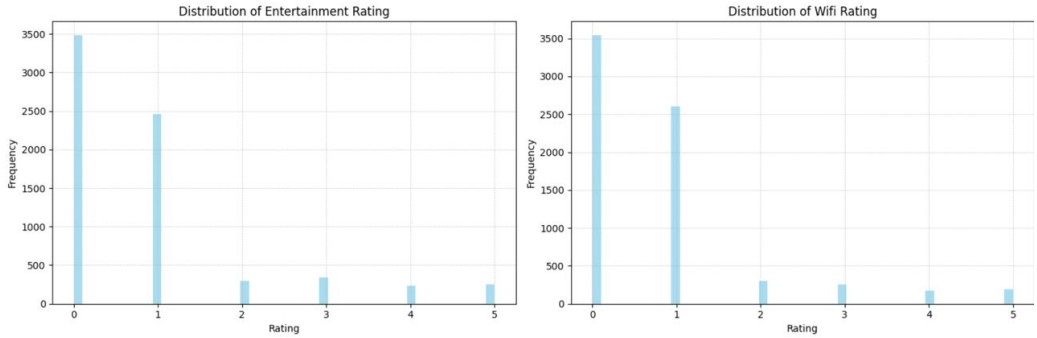
Despite the reasons why people travel, the vast majority choose economy class. Therefore in further research, the authors decided to study this class in more detail, as it is the most popular and may contain the most valuable information.

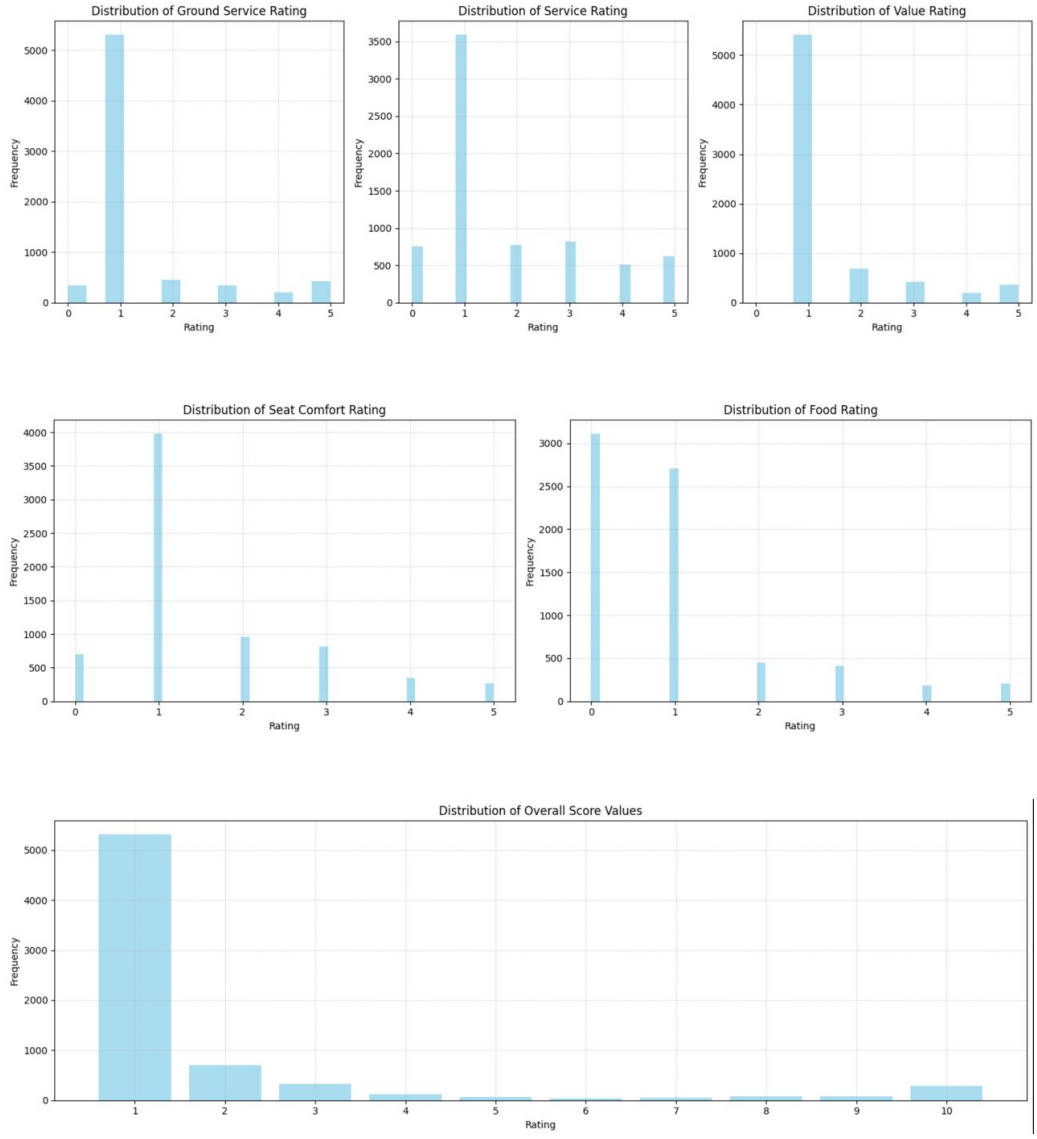
On the current stage the dataset consist of 322 unique airlines and the authors will explore how many unique reviews each airline has.



There are a lot of airlines with less than 10 reviews. This number of reviews isn't statistically significant in order to rely on them. Only 8 airlines have more than 500 reviews, making them the most popular and reliable among people. Among them are the following airlines: American Airlines, Spirit Airlines, Frontier Airlines, United Airlines, Delta Air Lines, Jetblue Airways, Allegiant Air, Southwest Airlines. It was found that all of these companies are American. That is why the authors decided to make the research based on information from reviewers on them.

After considering only economy class in those 8 American airlines the authors are left with only 7070 number of rows. One of the last step of dealing with data is to find distances of the flight and deleting those rows that contain None values in the column route because it makes them impossible to analyze. In final dataset we are left with 6800 values. Let's consider the distribution of rating columns of our final data in order to understand with what data the authors will work.





From the plots, it can be observed that in all rating columns the most common values are 0 or 1 while all other values together occupy a small percentage of reviews. This indicates that the services provided were of very poor quality. From the distribution of the values from overall score column, where again 1 is the most frequent value, the authors can conclude, that in the most cases people aren't satisfied with their flight experience.

4. Methodology explanation

After having the data visually represented, the next step is to analyze the main questions and to derive conclusions. While working on the project, the authors came up with the hypothesis that people from each type of travel may have different expectations about their flight, and their experience may differ. That is why their level of satisfaction may change, and the overall score they give for the trip may decrease. To test this hypothesis, the authors conducted an f-test.

- Null Hypothesis (H_0): There is no difference in the mean OverallScore across different TravelType groups.

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$$

- Alternative Hypothesis (H_1): At least one group has a mean OverallScore that is different, indicating varying levels of satisfaction.

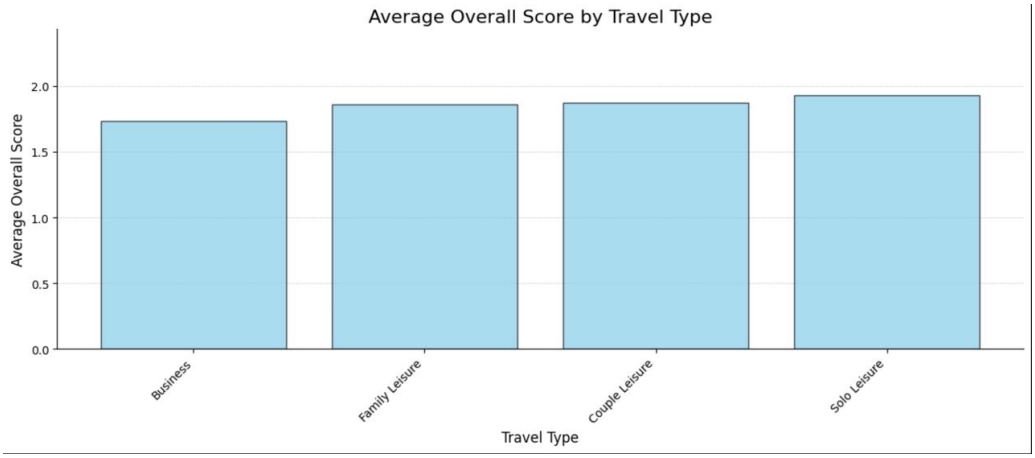
$$H_1 : \text{At least one } \mu_i \neq \mu_j \text{ for some } i \neq j$$

The results of this test are following:

- F-statistic: 1.6531394523006906;
- P-value: 0.17489958250191504;

The authors failed to reject the null hypothesis, there is no significant difference in overall score between people in different travel types.

This graph visually shows that the results of F-test are valid and we can't reject the null hypothesis. So people among all travel types have almost identical average overall score which is very low, indicating that they are all dissatisfied.



From the preliminary data analysis and the result of F-test the authors understood that all people generally are dissatisfied and in this stage they want to find out what factors increase the probability for people to be unsatisfied with their flight experience. In order to determine those values, the authors decided to run logit models.

For this, the binary column 'unsatisfied' was created, where 1 represents unsatisfied person(the person, whose overall score of the flight is in range between 1 and 3), and 0 - satisfied person(the person, whose overall score of the flight is in range between 4 and 10). As it was already mentioned, different travel types can have different requirements for the flight, that's why each travel type is considered separately. During the short-haul and long-haul flights, people may have different needs. That's why each category of the travel type was splitted according to the distance of the flight. So, to sum up, 8 logit models will be run with the dependent variable

ünsatisfied"that corresponds to each category. As independent variables, rating variables corresponding to each category were chosen: 'EntertainmentRating', 'FoodRating', 'ServiceRating', 'ValueRating', 'WifiRating', 'GroundServiceRating', 'SeatComfortRating'.

The general formula for the logit models is the following:

$$\text{unsatisfied} = \beta_0 + \beta_1 \cdot \text{EntertainmentRating} +$$

$$\beta_2 \cdot \text{FoodRating} + \beta_3 \cdot \text{ServiceRating} + \beta_4 \cdot \text{ValueRating} + \beta_5 \cdot \text{WifiRating} +$$

$$\beta_6 \cdot \text{GroundServiceRating} + \beta_7 \cdot \text{SeatComfortRating}$$

Let's consider the logit model for family leisure travel type on short-haul flights as an example:

Logit Regression Results

Dep. Variable:	unsatisfied	No. Observations:	1586
Model:	Logit	Df Residuals:	1578
Method:	MLE	Df Model:	7
Date:	Fri, 26 Apr 2024	Pseudo R-squ.:	0.7275
Time:	04:35:25	Log-Likelihood:	-128.40
converged:	True	LL-Null:	-471.12
Covariance Type:	nonrobust	LLR p-value:	9.544e-144

	coef	std err	z	P>|z|	[0.025	0.975]
const	8.4430	0.572	14.762	0.000	7.322	9.564
EntertainmentRating	-0.1174	0.152	-0.774	0.439	-0.415	0.180
FoodRating	-0.4656	0.141	-3.302	0.001	-0.742	-0.189
ServiceRating	-0.4818	0.137	-3.522	0.000	-0.750	-0.214
ValueRating	-1.1753	0.155	-7.568	0.000	-1.480	-0.871
WifiRating	0.1495	0.162	0.925	0.355	-0.167	0.466
GroundServiceRating	-0.5704	0.145	-3.928	0.000	-0.855	-0.286
SeatComfortRating	-0.1141	0.192	-0.595	0.552	-0.490	0.262

Figure 1: The results of the logit model for family leisure travel type on short-haul flights

A higher food rating is associated with lower log-odds of being 'unsatisfied'. This is statistically significant, indicating that food quality is an important factor in overall satisfaction. Similarly, an improved service rating is associated with lower log-odds of being 'unsatisfied', suggesting that service quality significantly affects satisfaction. The authors also can state that higher ground service ratings are a considerable factor in overall satisfaction. The variable 'ValueRating' has strong negative relationship with being unsatisfied, which is intuitively understandable as money are a valuable resource for people and they do not want to waste them. All other variables are statistically insignificant. The results that are presented correspond to the intuition.

Now, after running all models, the authors are ready to share the obtained insights:

- For all people that are travelling on short-haul flights, except families, the variable food is considered as insignificant, meaning it is not the important factor in overall satisfaction. Basically, families is the only travel type for whom the food is significant both on short- and long-haul flights.

- For those, who are travelling on business purpose on the short-haul flights, the only thing that matters is service.
- There is also a contradiction to our intuition her because from the model the authors can conclude that the comfort of the seats is not always significant on both short- and long-haul flights.
- The variable Wifi is insignificant for all travel types, even for those, who are travelling on business purpose.
- It is also interesting that the ground service is crucial to all passengers, except people on business trips.

So, after gaining the results above, the authors decided to proceed with analyzing the ground service ratings, trying to find out for whom this factor matters more: for people that do or do not have the direct flight. For this the authors used the $t-test$ with the following hypotheses:

H_0 : Ground service is equally important for both groups with and without transfer.
H_1 : Ground service is more important for those with a transfer.

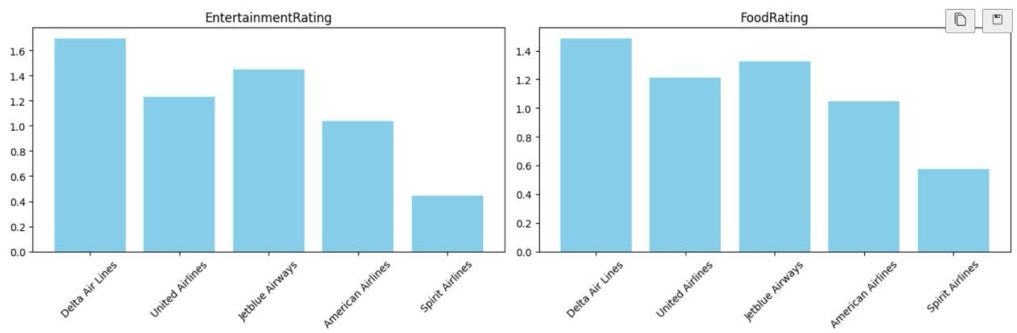
The results of this test are following:

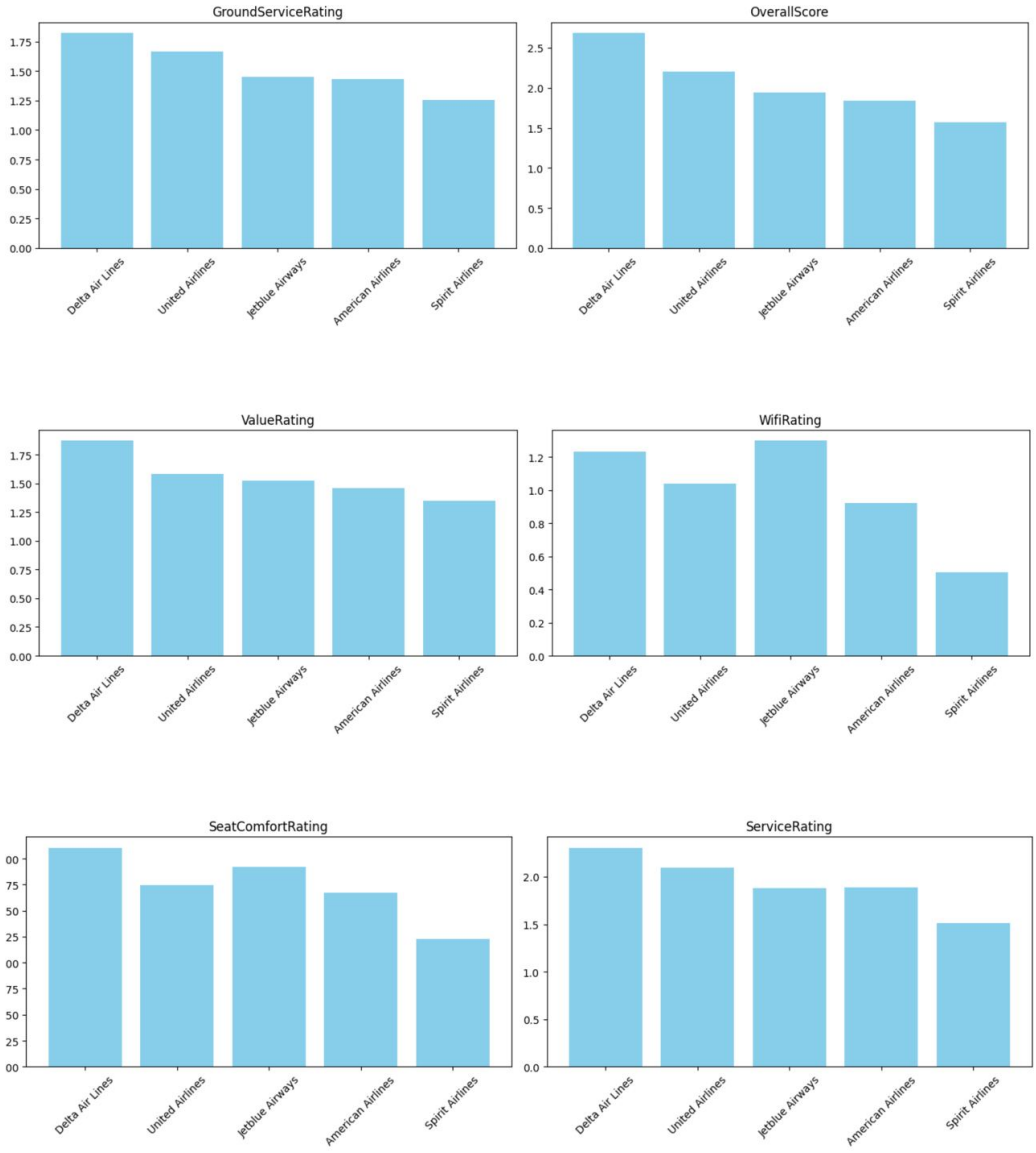
- T-statistic: 2.7752107474161916
- P-value: 0.0027889274379432048

From this, we can conclude that there is enough evidence to reject the null hypothesis. The authors also found out the top-5 airports with the highest average values of the ground service.

Philadelphia	1.645299
Tampa	1.512821
Phoenix	1.496914
Atlanta	1.480249
Houston	1.479310

Let's move on to the last question of this research: **comparing different airlines.** Let's take a closer look at different ratings of the airlines. Here the top-5 airlines with the biggest number of reviews is considered:





Overall, Delta Air Lines consistently ranks highest across all rating categories. Spirit Airlines generally has the lowest ratings, with its greatest weaknesses being entertainment, food, and Wifi. United Airlines and JetBlue Airways are often in close competition, usually ranking in the middle to upper range, and American Airlines typically falls between the top performers and Spirit Airlines.

5. Conclusions

Based on the findings from the report, the following recommendations can be made for both passengers and airlines in order to increase the satisfaction level of passanger:

1. Amenities like in-flight entertainment and Wifi are not significant for all. That's why airlines instead of investing in this areas should consider other areas to improve overall experience.

2. Since ground service has a significant impact on customer satisfaction, airlines should try to enhance the efficiency and friendliness of check-in, boarding, transfers, and baggage handling.

3. Airlines should consider improving meal options and catering because for family leisure travelers food quality is a significant factor.

By putting these recommendations into practice, it's possible for Airlines to elevate the quality of service to align more closely with customer expectations, thus creating a travel experience that is more satisfying for everyone involved.