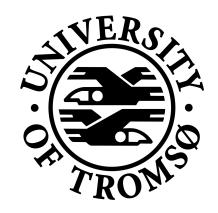UiT - Arctic University of Tromsø

Project Proposal:
Finding the best US based airline

Daniël ten Wolde

BED-2056: Introduction to Data Science

October 3, 2019

Project Plan

Initial idea

The initial idea for this project is to analyze and compare a number of United States based airlines. The aim is to highlight different areas of the airlines and try to answer the question: "Which is the best airline for different categories?". We will be comparing these airlines on a variety of topics. For now these topics include:

- Total revenue and expenses. This could be expanded into a more detailed comparison between airlines, depending on the amount of data found. It could be that a certain airline is more focused on a particular region. Which regions are more heavily contested?

- Size of fleet. Is the size proportional to the revenue?

- Number of employees.

- Amount of yearly / monthly passengers.

- Average Departure and Arrival Delay.

- Total number of flight movements per day / week / month.

- Total number of cancelled flights. Does a particular airline have more cancelled or diverted flights than others?

- Customer satisfaction. Is there any relation to be found between satisfaction and the size of an airline? Do smaller airlines (based on revenue) have achieve a higher customer satisfaction.

Data sources

Revenue and expenses data

One of the data sources we will be using is the Airline Data Project" [1] which has gathered data from Bureau Transportation Statistics [2]. We will use this to get data on the revenue, expenses and number of employees for different airlines. The data provides revenue and expenses split up into many different categories which can be compared.

Flight data

The Bureau Transportation Statistics also keeps track of Airline On-Time Statistics and Delay Causes, we will also make use of this data set. The data set can be found at the American Statistical Association [3]. The only downside is that the data set is no longer being updated and has only recorded data up to 2008. However there is still a large amount of data contained in each year, thus there should be enough to base research on, be it on old data.

Customer satisfaction data

The data to evaluate the passenger satisfaction will come from the American Customer Satisfaction Index [4] which brings out a yearly travel report on the customer satisfaction regarding different airlines. Another source we could use for this comes from J.D. Power [5]

Proposed Evaluation

We will be evaluating multiple airlines in different categories and look for airlines that score significantly higher than others. Since this involves multiple categories, we will evaluate every category in a different way.

Revenues and expenses

We will be looking to graph the total revenues and expenses for different airlines over the given time-period. This allows us to compare the potential rises and falls of different airlines and make comparisons. Another possibility is to graph the different regions of the revenues (International, Latin America, Pacific, Domestic and Atlantic) and see which airlines have more presence in each region.

Fleet size and number of employees

For this data we will also make a graph to compare the absolute values between the airlines. However it might also be interesting to graph the relative values combined with the total revenue. This allows us to make a fairer comparison between the airlines, since it is likely that an airline making having more revenue will also have a larger fleet and more employees. We will then see if there are any airlines that stand out.

Departure and Arrival Delay

We will calculate the average Departure and Arrival Delay that each airline has had up to 2008 and graph them. We can then compare the airlines and see if any stands out in either category. The same methodology can be done to analyze the amount of diverted or cancelled flights per airline. We can also analyze if the number of diverted or cancelled flights has reduced or increased over the years.

Flight movements

For this category we can simply calculate the amount of total flight movements over a certain period of time. Another possibility is to check if a certain airline has departed more from a particular airport than another airline. Lastly, we can also compare the average distance covered per flight per airline and see which airline focuses more on short- or long-distance flights.

Customer satisfaction

For this category it could be interesting to see which airline has a higher customer satisfaction. Do relatively smaller airlines have a higher customer satisfaction? Is it related to the amount of arrival or departure delay? Is the amount of cancelled or diverted flights of any influence?

Bibliography

[1] MIT, "Airline data project." Available at `http://web.mit.edu/airlinedata/www/default.html` (Accessed: 29-09-2019).

[2] B. of Transportation statistics, "Air carrier financial reports." Available at `https://www.transtats.bts.gov/Tables.asp?DB_ID=135` (Accessed: 29-09-2019).

[3] B. of Transportation statistics, "Airline on-time statistics and delay causes." Available at `http://stat-computing.org/dataexpo/2009/the-data.html` (Accessed: 29-09-2019).

[4] A. C. S. Index, "Acsi travel report 2018-2019." Available at `https://www.theacsi.org/news-and-resources/customer-satisfaction-reports/reports-2019/acsi-travel-report-2018-2019/acsi-travel-report-2018-2019-download` (Accessed: 29-09-2019).

[5] J. Power, "North america airline customer satisfaction improves for 7th consecutive year." Available at `https://www.jdpower.com/business/press-releases/jd-power-2018-north-america-airline-satisfaction-study` (Accessed: 29-09-2019).