NYC Flights

Emanuel Rodriguez Bedeman

7/12/2021

Flight Analysis from New York City (*NYC*) to Atlanta (*ATL*)

Curiosity has aroused me and I started investigatin the flights from **NYC** to **Atlanta**

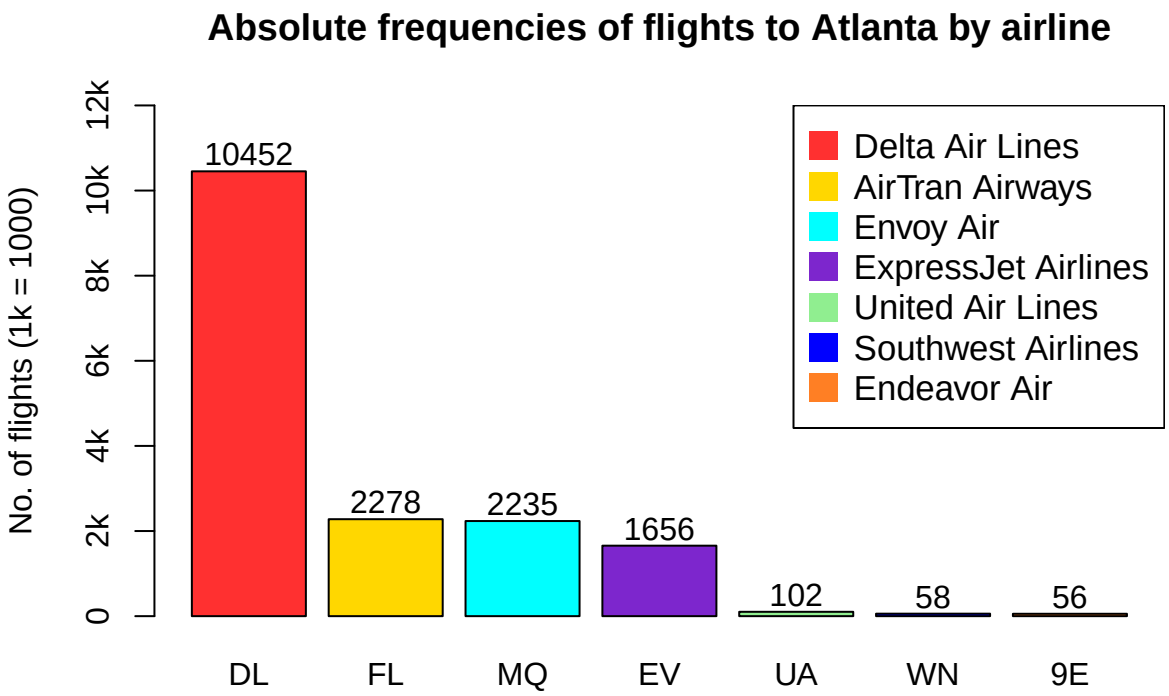
from the total flights (336776)**, only** 16837 **went to Atlanta.**

Which, in porcentage, is equal to:

$$\left(\frac{Total\,Flights}{Atlanta\,Flights}\right) \cdot 100 = 5\%$$

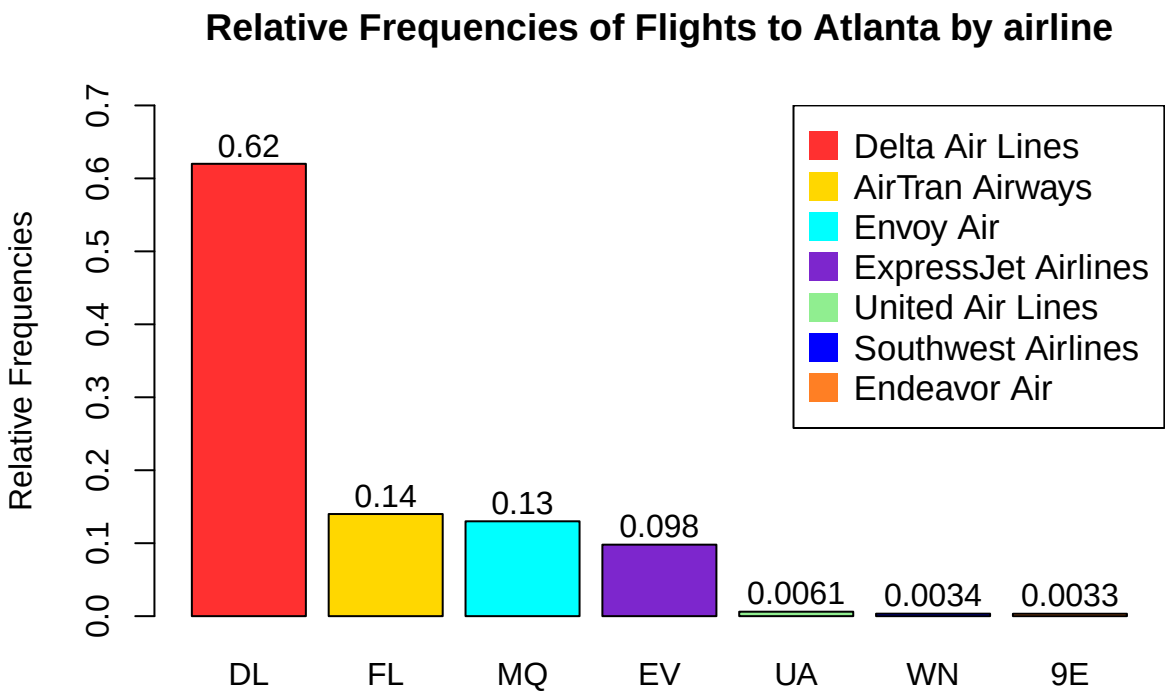
The available airlines to travel from NYC to ATL are:

Abbreviation	**Airline**
DL	Delta Air Lines
FL	AirTran Airways
MQ	Envoy Air
EV	ExpressJet Airlines
UA	United Air Lines
WN	Southwest Airlines
9E	Endeavor Air



In this graph of Absolute Frequencies we can see flights made in the year 2013, from *New York* to *Atlanta*, distributed by airlines:

- 10452 of flights made with the airline "***Delta Air Lines***".
- 2278 of flights made with the airline "***AirTran Airways***".
- 2235 of flights made with the airline "***Envoy Air***".
- 1656 of flights made with the airline "***ExpressJet Airlines***".
- 102 of flights made with the airline "***United Air Lines***".
- 58 of flights made with the airline "***Endeavor Air***".
- 56 of flights made with the airline "***Southwest Airlines***".

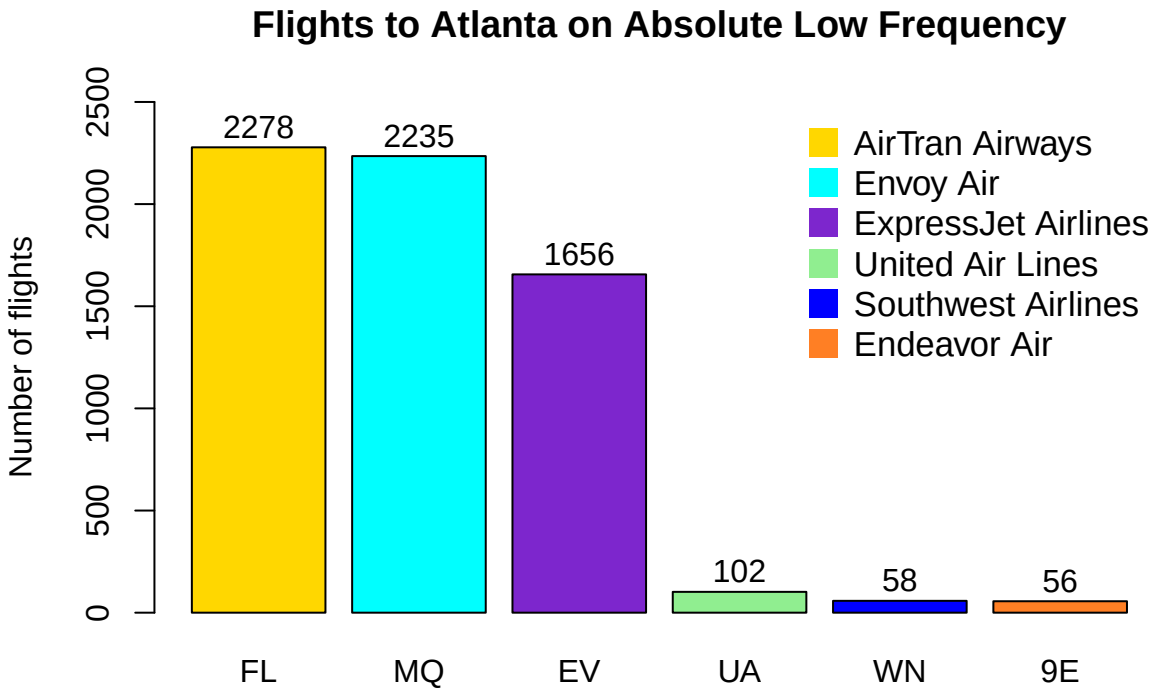


In this graph of *Relative Frequencies* of the flights to *Atlanta* by carriers , we can see how much the flights of each airline represent of the total flights to *Atlanta*, being that:

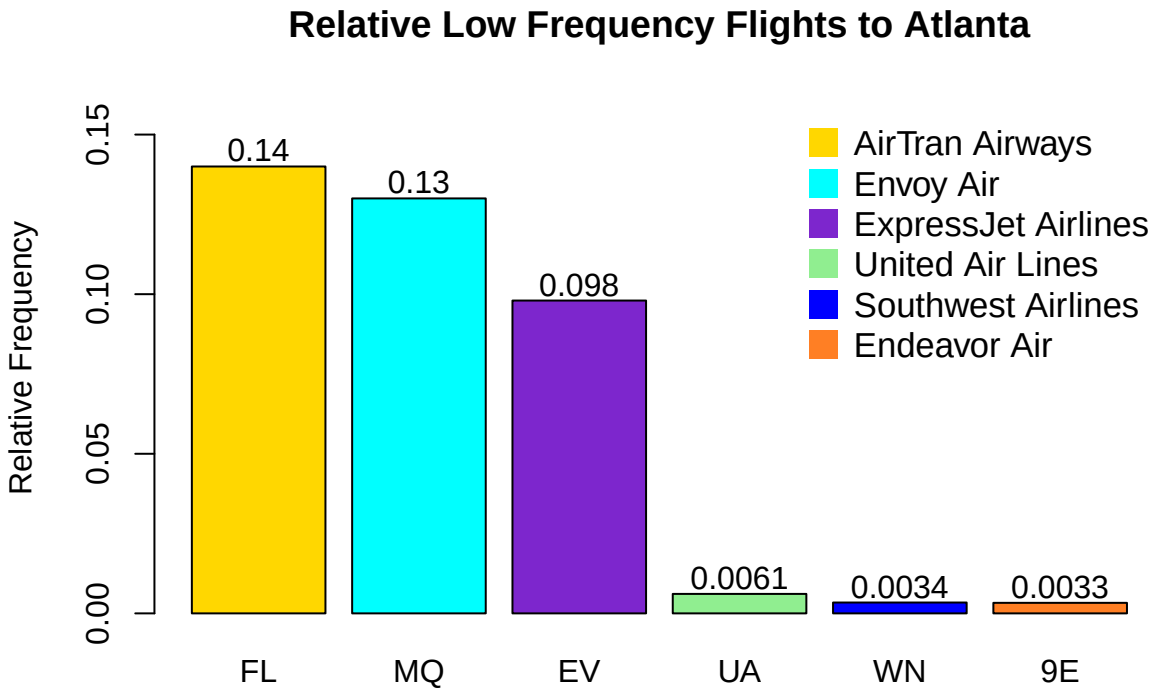
- 0.62 is the proportion of flights made by the airline "***Delta Air Lines***".
- 0.14 is the proportion of flights made by the airline "***AirTran Airways***".
- 0.13 is the proportion of flights made by the airline "***Envoy Air***".
- 0.098 is the proportion of flights made by the airline "***ExpressJet Airlines***".
- 0.0061 is the proportion of flights made by the airline "***United Air Lines***".
- 0.0034 is the proportion of flights made by the airline "***Endeavor Air***".
- 0.0033 is the proportion of flights made by the airline "***Southwest Airlines***".

Absolute low frequency airlines (<2,500 flights)

In the following graphs you can see in greater detail those airline flights considered (by me) as low frequency with respect to "***Delta Air Lines***" because this is, without a doubt, the one with the most trips.



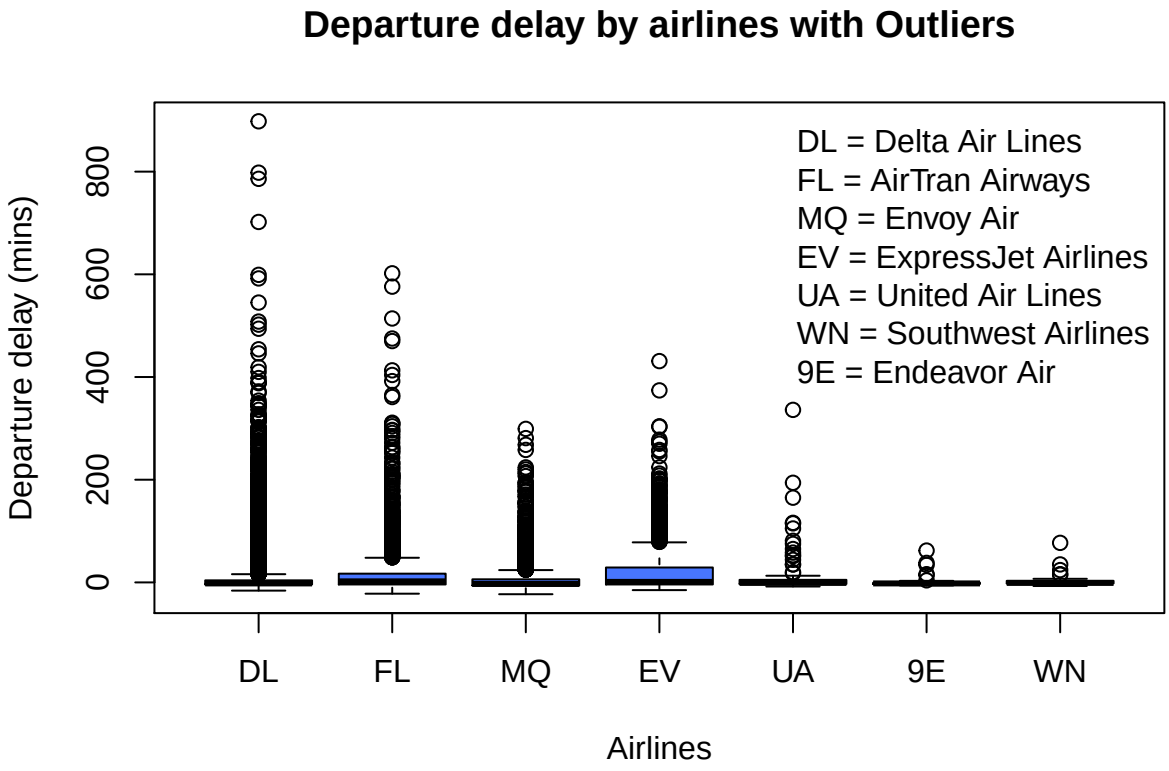
Relative low frequency airlines (<0.15)



As we can see in all these graphs, the carrier "*Delta Air Lines*" ($ DL $) is the most used airline (and by far) to go from *NYC* to *Atlanta*.

- Will it be the best airline to travel?
- Why is DL the most used carrier? -Is it the carrier with the least delays?
- Is it the person who compensates the most for the delays by arriving on time?
- How many Delta Air Lines company flights depart from each NYC airport to Atlanta?

Departure delays



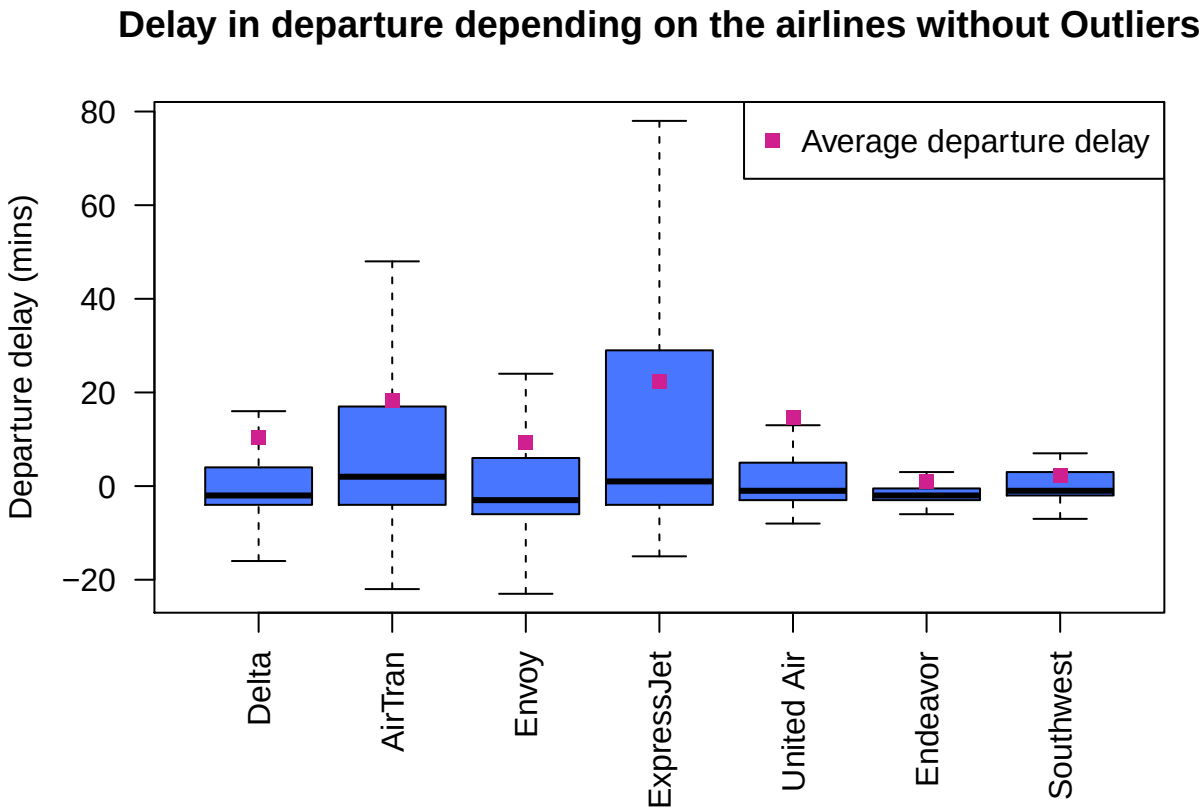
As we can see, although all airlines have delay problems, the company "***Delta Air Lines***" is the one that has the greatest problems to meet the departure time, with a maximum of: 898 *minutos* $\tilde{=}$ 15 *horas* of delay.

Then it follows, both in greater time of delay and greater problems to comply with the schedule:

- "***Airtran Airways***" with a maximum of 602 *minutos* $\tilde{=}$ 10 *horas* of delay.
- "***ExpressJet Airlines***" with a maximum of 431 *minutos* $\tilde{=}$ 7 *horas* of delay.
- "***United Air Lines***" with a maximum of 336 *minutos* $\tilde{=}$ 6 *horas* of delay.
- "***Envoy Air***" with a maximum of 299 *minutos* $\tilde{=}$ 5 *horas* of delay.
- "***Southwest Airlines***" with a maximum of 77 *minutos* $\tilde{=}$ 1,3 *horas* delay.
- "***Endeavor Air***" with a maximum of 62 *minutos* $\tilde{=}$ 1 *horas* of delay.

Airline	Minutes	Hours
Delta Airlines	898	15
Airtran Airways	602	10
ExpressJet Airlines	431	7
United Airlines	336	6
Envoy Air	299	5
Southwest Airlines	77	1,3
Endeavor Air	62	1

Let's look at this same graph but without the outliers:



Here we can see that the airline with the greatest distribution of departure delay time (**WITHOUT OUTLIERS**) is "*EXpressJet Airlines*", in second place the airline "*AirTran Airways*" and in third place "*Envoy Air*". It should be noted that these two airlines have the most cases in which the flight left earlier!

One thing to note in the boxplots of all these airlines is that the **arithmetic mean** (represented by the reddish-pink square) is greatly affected by the outliers because, as is the sum of values divided by the number of values, already having one or more values far outside the range of values generates very rude changes.

Average departure delay by airline:

Airlne	Average(mins)
ExpressJet Airlines	22.26
AirTran Airways	18.27
United Air Lines	14.67
Delta Air Lines	10.35
Envoy Air	9.33
Southwest Airlines	2.34
Endeavor Air	0.98

Absolute Median Deviation for the departure delay

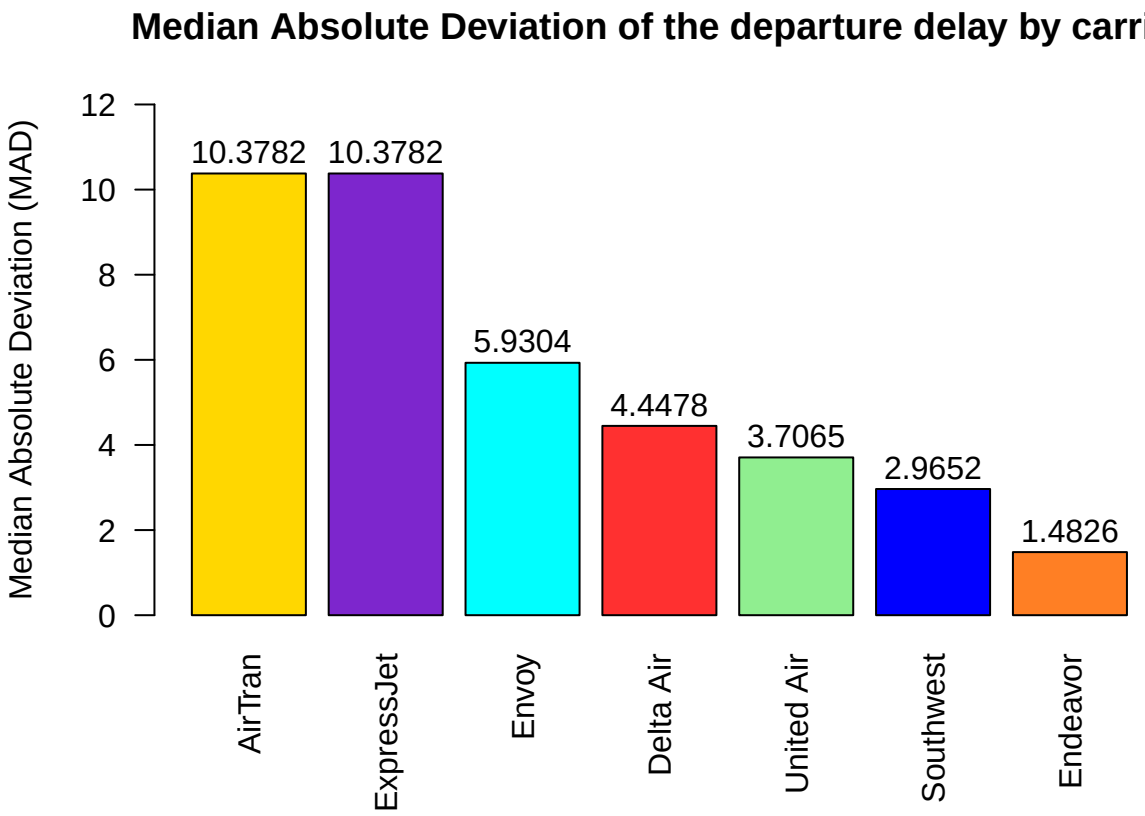
We can also calculate the ** absolute median deviation (DMA) ** so that those airlines whose mean is highly modified by very small or very large values, can be compared by their mean value (median), which is not so easy. to modify.

Calculation of the absolute median deviation:

Result = Absolute value of each value - the median of the values

... Using a statistical value $ k = 1.4826 $

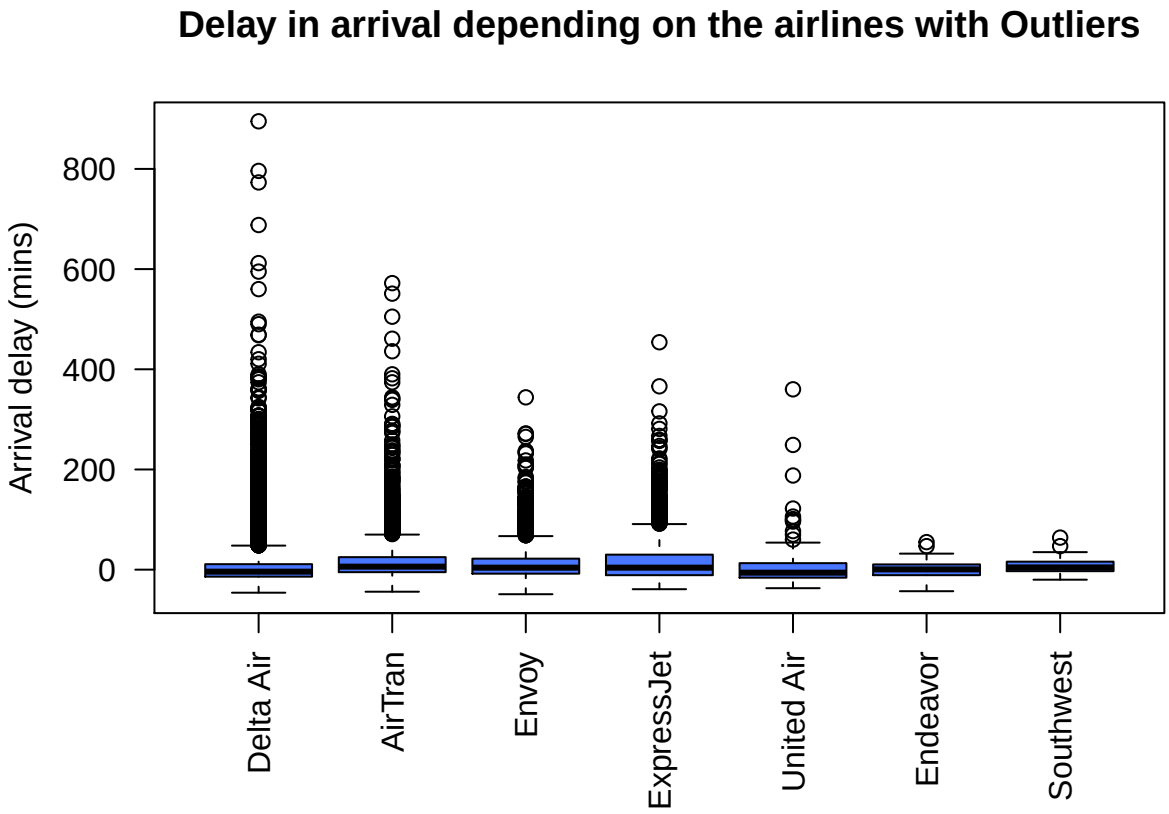
$(DMA) = median(Result) \cdot k$



Apparently, even though the company "*Delta Air Lines*" has many *outliers*, according to the MAD, the dispersion in terms of the delay of the flight departure is within all stable (it remains close to a central value).

The time to wait for the flight to depart is not so variable, compared to other airlines such as "*AirTran Airways*" and "*ExpressJet Airlines*" which has a very high dispersion, which I interpret as a longer waiting time for it to depart. a flight.

Arrival delays

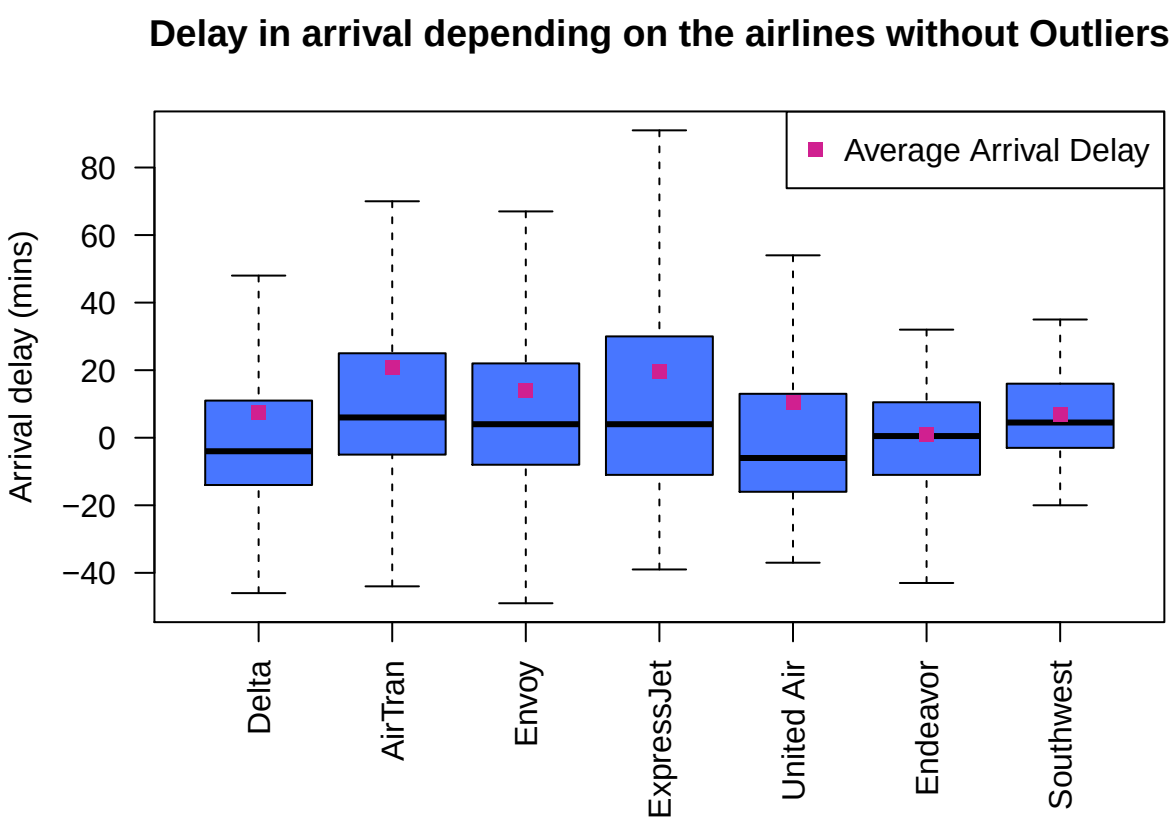


We see again that the company "***Delta Air Lines***" is the one with the greatest delay problems but in this case to comply with the arrival time, with a maximum of: 895 *minutos* $\tilde{=}$ 15 *horas* late.

Then it follows, both in greater time of delay and in greater problems to comply with the schedule:

- "***Airtran Airways***" with a maximum of 572 *minutos* $\tilde{=}$ 10 *horas* of delay.
- "***ExpressJet Airlines***" with a maximum of 454 *minutos* $\tilde{=}$ 8 *horas* of delay.
- "***United Air Lines***" with a maximum of 360 *minutos* $=$ 6 *horas* late.
- "***Envoy Air***" with a maximum of 344 *minutos* $\tilde{=}$ 5 *horas* of delay.
- "***Southwest Airlines***" with a maximum of 64 *minutos* $\tilde{=}$ 1 *horas* of delay.
- "***Endeavor Air***" with a maximum of 55 *minutos* $\tilde{=}$ 1 *horas* of delay.

Airline	Minutes	Hours
Delta Airlines	895	15
Airtran Airways	572	10
ExpressJet Airlines	454	8
United Airlines	360	6
Envoy Air	344	6
Southwest Airlines	64	1
Endeavor Air	55	1



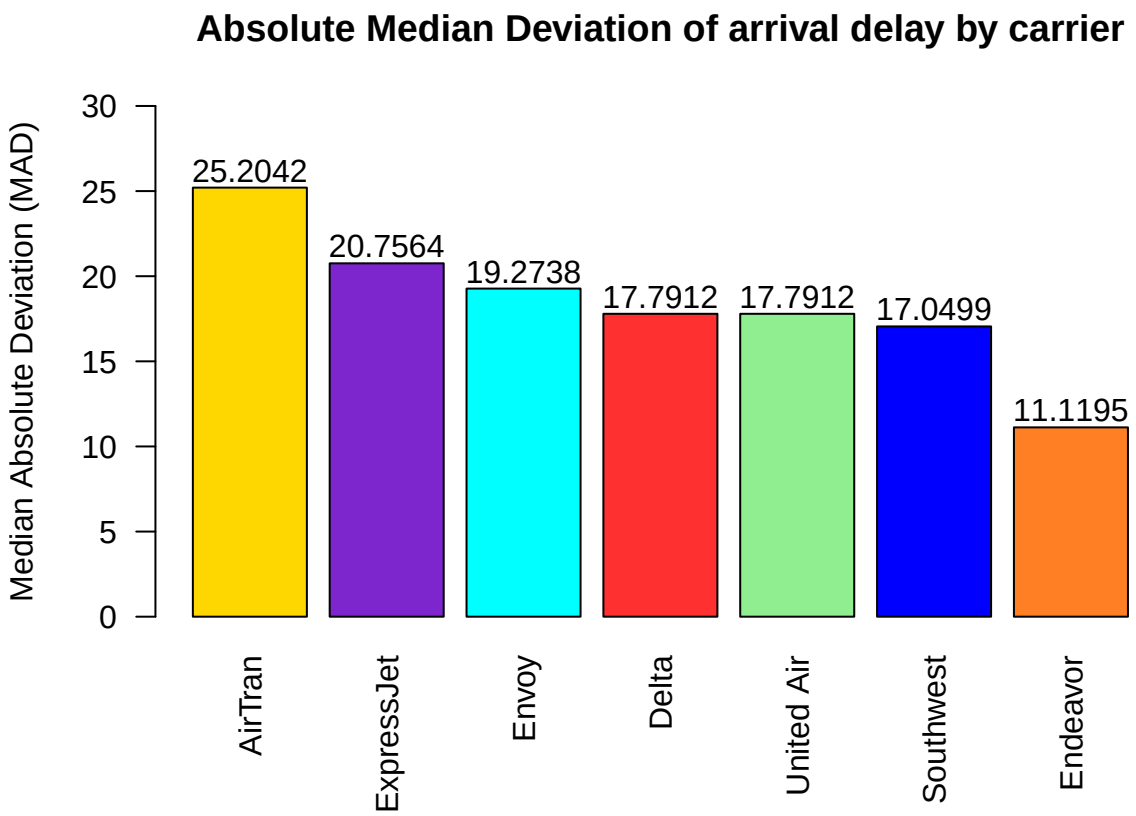
Average arrival delay by airline:

Airline	Average(Mins)
AirTran Airways	20.74
ExpressJet Airlines	19.64
Envoy Air	14.03
United Air Lines	10.50
Delta Air Lines	7.42
Southwest Airlines	6.90
Endeavor Air	0.86

As we can see, the carrier with the least dispersed arrival delays (**WITHOUT OUTLIERS**) is "*Southwest Airlines*" and the one with more dispersed delays is the airline "*ExpressJet Airlines*".

Stopping at the company "*Delta Airlines*" we can see that it has a much more varied distribution of waiting times than the airline "*Sothwest Airlines*", besides that it is the one with the greatest maximum arrival delay (15hs) but still it is not as scattered as "*ExpressJet Airlines*".

Median Absolute Deviation



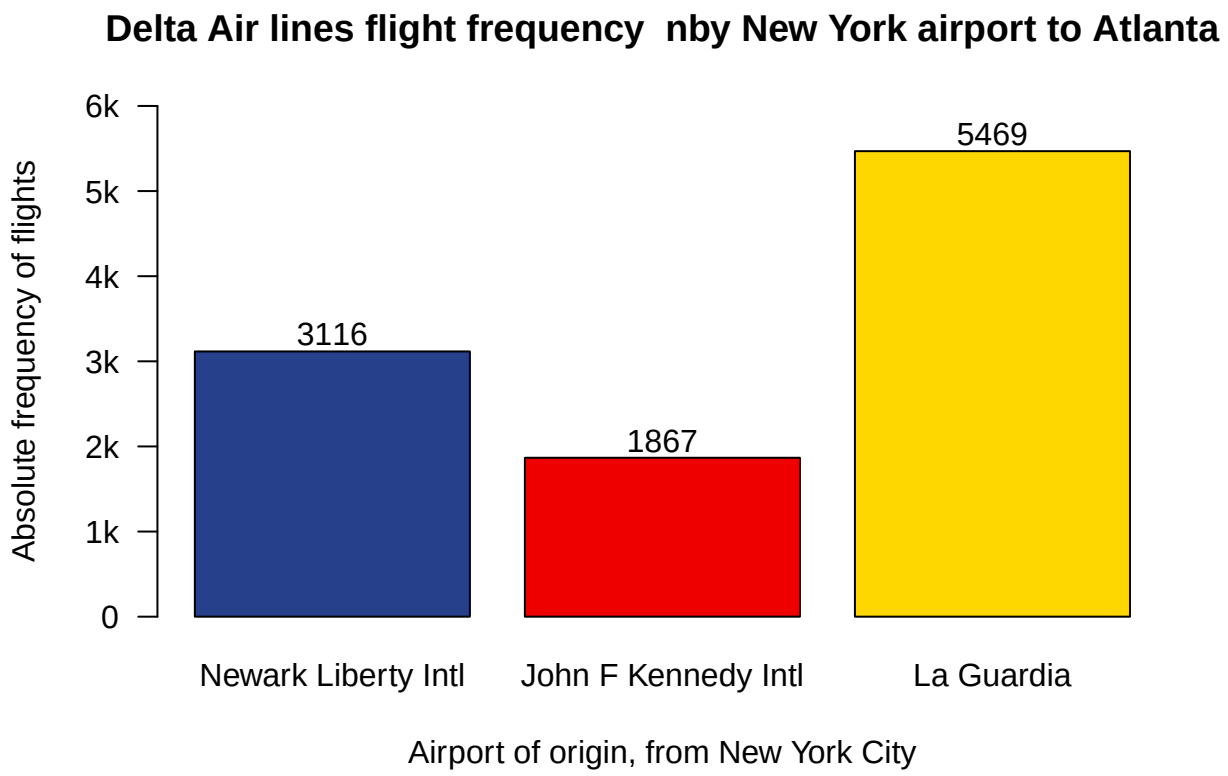
Now analyzing the MAD of the arrival delay by carrier, we can see that the company "*Endeavor Airlines*" is the one that has the least distributed arrival delay, being the airline "*AirTran Airlines*" the one that varies the most.

Focusing our gaze once more on the airline "*Delta Airlines*", we can see that although its delay values are quite distributed, they are not as distributed as other companies but it is not poorly distributed as other companies, but rather it is in a midpoint with respect to other airlines.

Unfortunately, the dataset does not include the price of the tickets of each airline to go to Atlanta, which would be very useful to know why the company "*Delta Airlines*" is the most used to travel, because seeing the delays, it is clear that "*Endeavor Air*" is the one with the least delays, this means that it is the one that complies with the schedules the most.

Flights through NYC airport to Atlanta by Delta Airlines

To conclude, we are going to see which airport is the most used by the airline "*Delta Air Lines*".



We can see that most of the flights departing from *New York* bound for *Atlanta* by the airline "*Delta Airlines*" depart from the airport "***La Guardia***".

Conclusion:

We could say that possibly the airline "*Delta Airlines*" is the one with the most flights, possibly due to its fame, although it is not the company with the least delays, but rather it is an intermediate one. Perhaps it will be due to its greater availability of flights on many dates and times. In addition to that perhaps it is the company with the cheapest tickets.

In this case, if we were in a hurry with the need to travel from New York to Atlanta and get there quickly, we would surely take the "*Endeavor Airlines*" flight because it is the one with the least departure and arrival delay. Perhaps it is not the most used because it has less availability of flights, fewer dates and times of flights or because it is very expensive. With the available data we cannot know.

End of analysis.