

NYC Flights

Emanuel Rodriguez Bedeman

6/12/2021

Analisis vuelos desde New York City (*NYC*) a Atlanta (*ATL*)

Me ha despertado la curiosidad de investigar los vuelos que salen a **Atlanta** desde NYC.

Del total de vuelos (336776), solamente 16837 fueron hacia **Atlanta**.

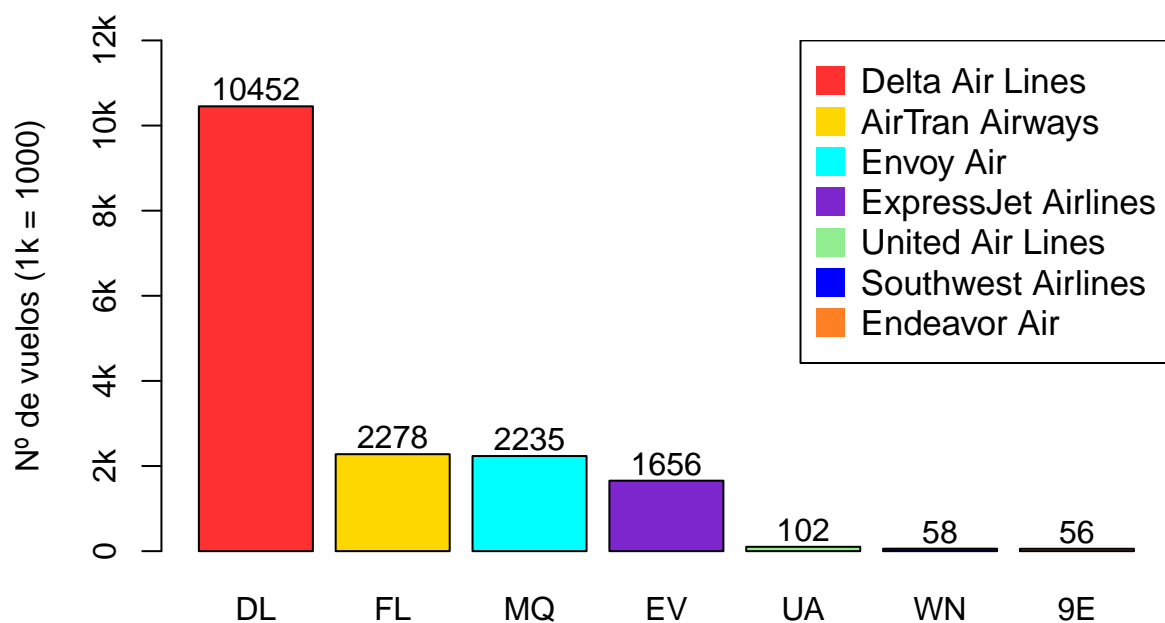
Lo que, dicho en porcentaje, es igual a:

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

Las aerolineas disponibles para viajar desde *New York City* a *Atlanta* son:

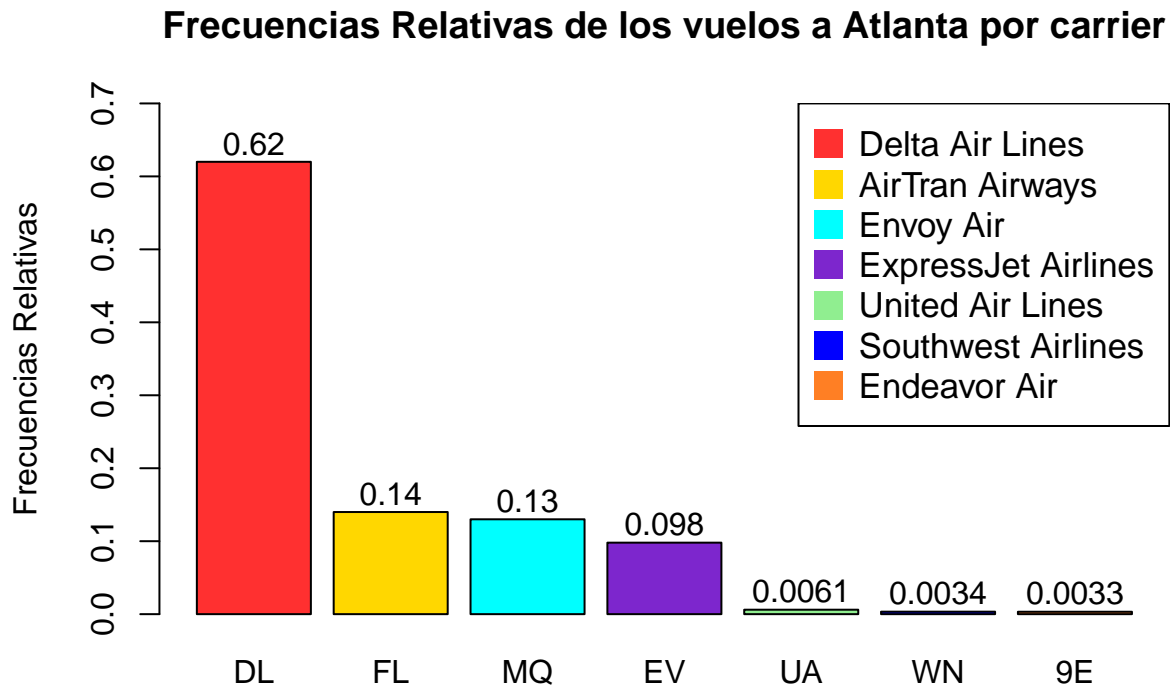
Abreviatura	Aerolinea
DL	Delta Air Lines
FL	AirTran Airways
MQ	Envoy Air
EV	ExpressJet Airlines
UA	United Air Lines
WN	Southwest Airlines
9E	Endeavor Air

Frecuencias Absolutas de los vuelos a Atlanta por carrier



En este gráfico de Frecuencias Absolutas podemos ver en valores absolutos vuelos realizados en el año 2013, desde *New York* a *Atlanta*, distribuidos por aerolíneas:

- 10452 de vuelos realizados con la aerolínea “*Delta Air Lines*”.
- 2278 de vuelos realizados con la aerolínea “*AirTran Airways*”.
- 2235 de vuelos realizados con la aerolínea “*Envoy Air*”.
- 1656 de vuelos realizados con la aerolínea “*ExpressJet Airlines*”.
- 102 de vuelos realizados con la aerolínea “*United Air Lines*”.
- 58 de vuelos realizados con la aerolínea “*Endeavor Air*”.
- 56 de vuelos realizados con la aerolínea “*Southwest Airlines*”.

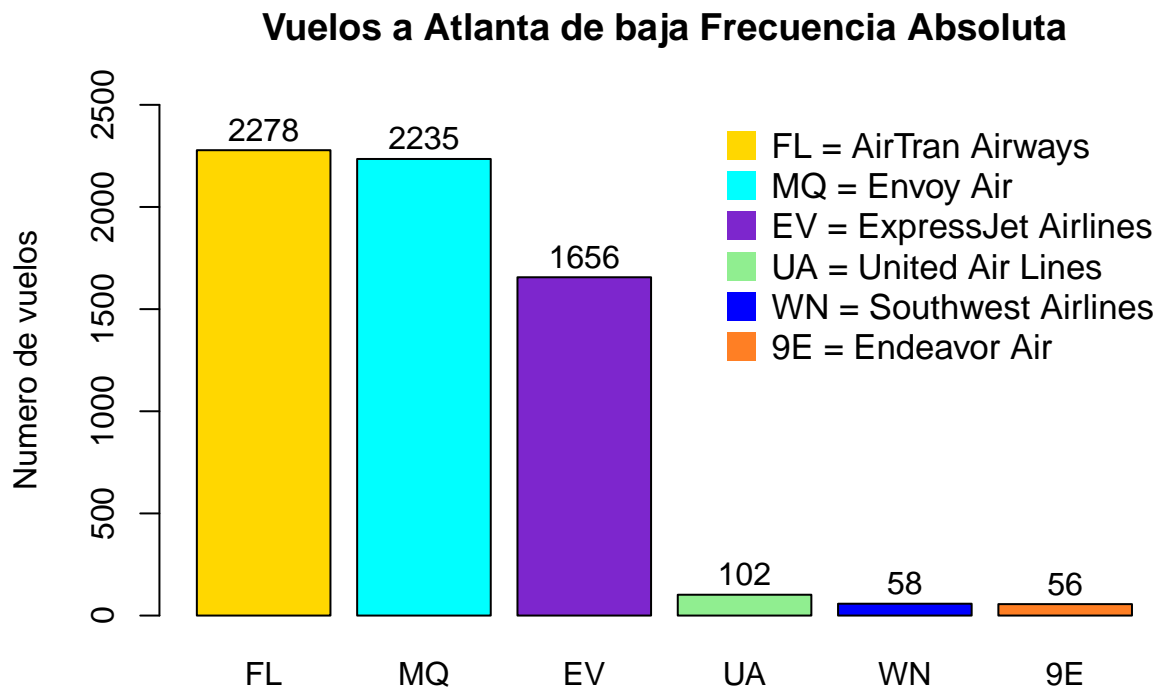


En este gráfico de *Frecuencias Relativas* de los vuelos a *Atlanta* por carriers, podemos observar cuanto representan los vuelos de cada aerolínea del total de vuelos a *Atlanta*, siendo que:

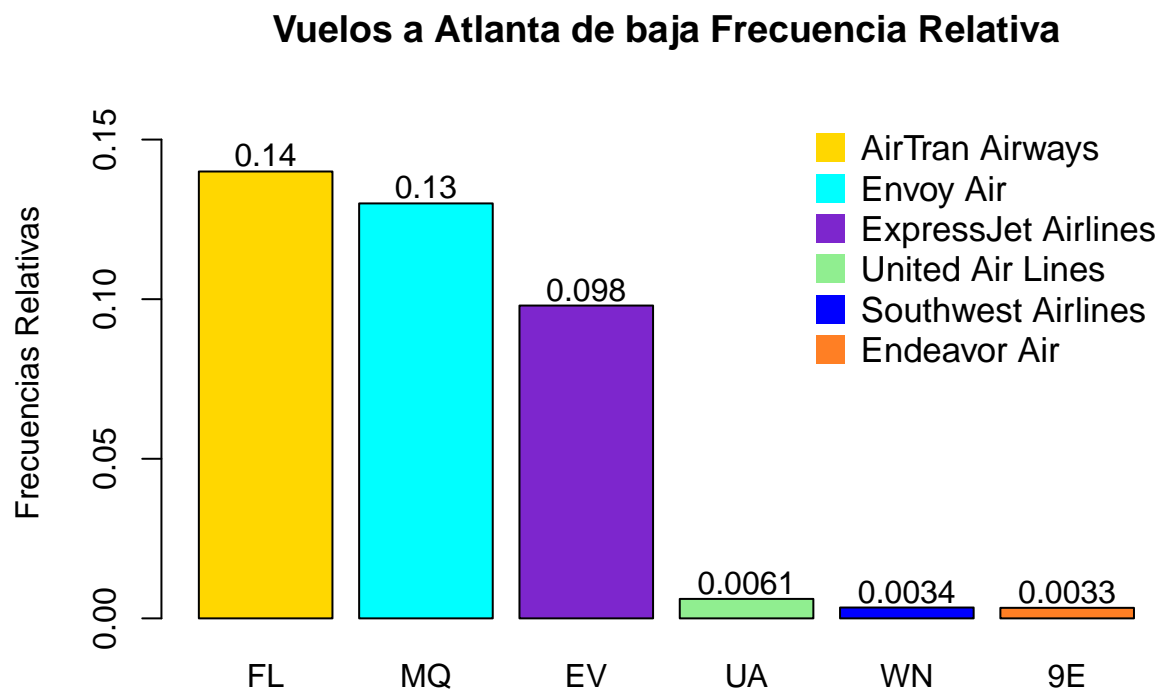
- 0.62 es la proporción de vuelos realizados por la aerolínea “*Delta Air Lines*”.
- 0.14 es la proporción de vuelos realizados por la aerolínea “*AirTran Airways*”.
- 0.13 es la proporción de vuelos realizados por la aerolínea “*Envoy Air*”.
- 0.098 es la proporción de vuelos realizados por la aerolínea “*ExpressJet Airlines*”.
- 0.0061 es la proporción de vuelos realizados por la aerolínea “*United Air Lines*”.
- 0.0034 es la proporción de vuelos realizados por la aerolínea “*Endeavor Air*”.
- 0.0033 es la proporción de vuelos realizados por la aerolínea “*Southwest Airlines*”.

Aerolineas de baja frecuencia absoluta (<2500 vuelos)

En los siguientes gráficos puede observarse en mayor detalle aquellos vuelos de aerolineas considerados (por mi) como de baja frecuencia con respecto a “*Delta Air Lines*” porque esta es, sin lugar a duda, la que más viajes tiene.



Aerolineas de baja frecuencia relativa (<0.15)

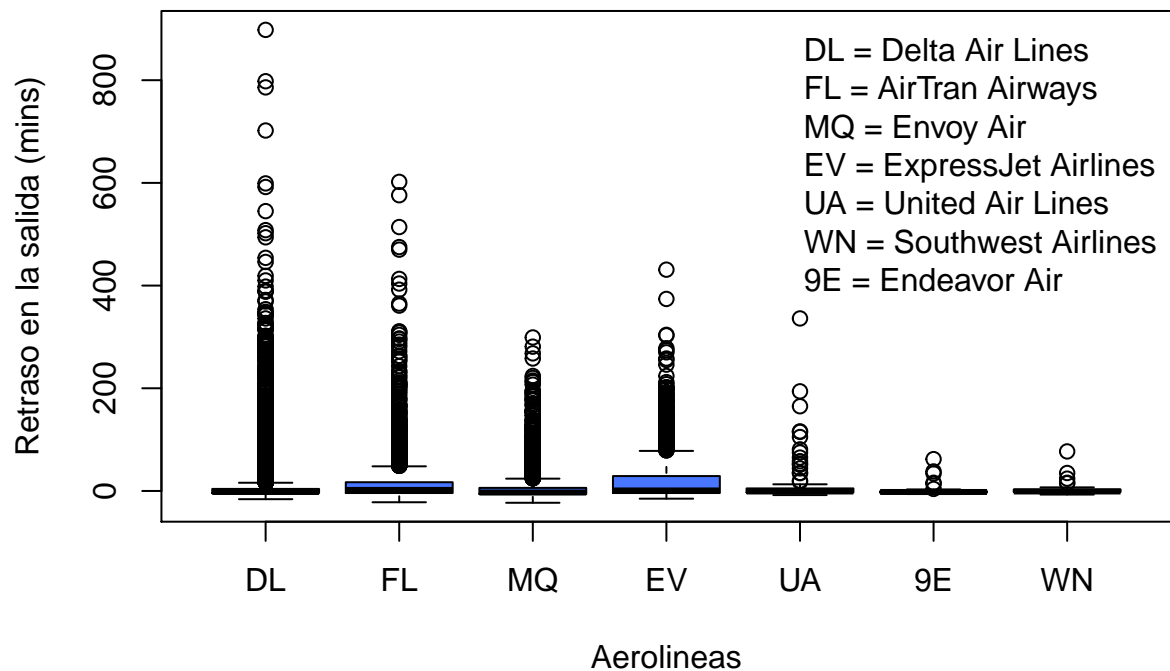


Como podemos observar en todos estos gráficos, el carrier “*Delta Air Lines*” (*DL*) es la aerolinea más usada (y por lejos) para ir desde *NYC* a *Atlanta*.

- ¿Será la mejor aerolinea para viajar?
- ¿Por qué es *DL* es el carrier más usado? -¿Es el carrier que menos retrasos tiene?
- ¿Es quien compensa más los retrasos llegando a tiempo?
- ¿Cuántos vuelos de la empresa Delta Air Lines salen de cada aeropuerto de NYC con destino a Atlanta?

Retrasos en salida

Retraso en salida en funcion de las aerolineas con Outliers



Como podemos observar, si bien todas las aerolineas tienen problemas de retrasos, la empresa " **Delta Air Lines** " es la que tiene mayores problemas para cumplir el horario de salida, con un maximo de: $898 \text{ minutos} \approx 15 \text{ horas}$ de retraso.

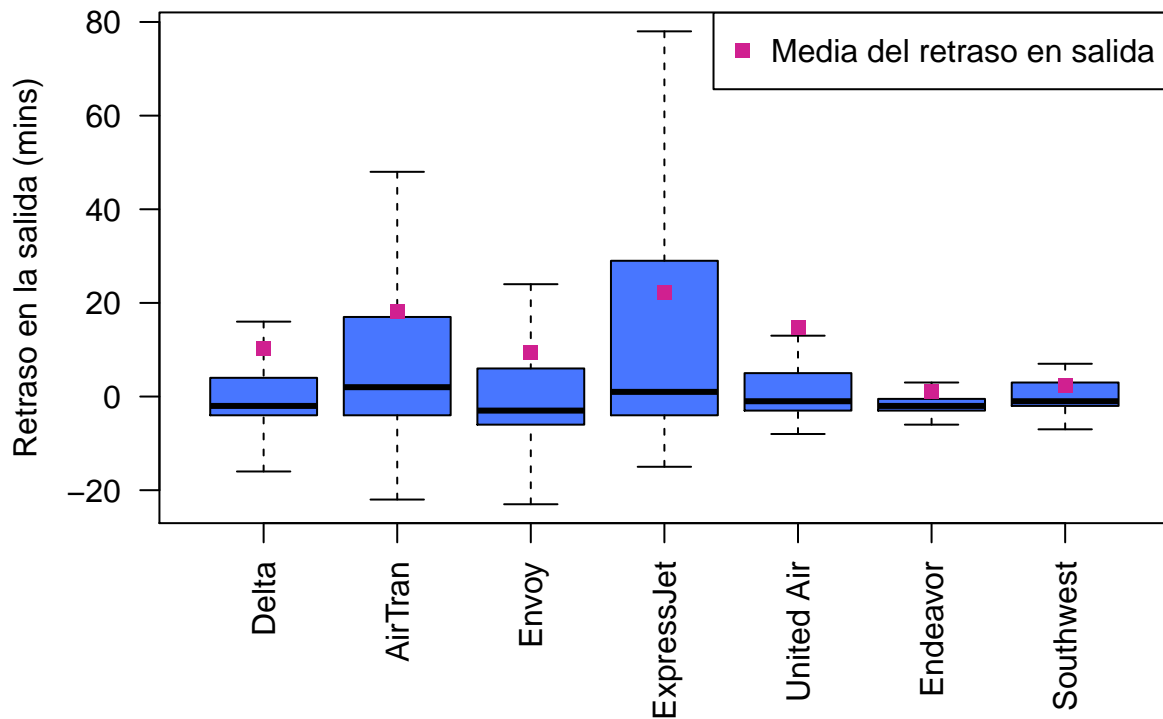
Luego le sigue, tanto en mayor tiempo de retraso como mayor problemas para cumplir el horario:

- " **Airtran Airways** " con un maximo de $602 \text{ minutos} \approx 10 \text{ horas}$ de retraso.
- " **ExpressJet Airlines** " con un maximo de $431 \text{ minutos} \approx 7 \text{ horas}$ de retraso.
- " **United Air Lines** " con un maximo de $336 \text{ minutos} \approx 6 \text{ horas}$ de retraso.
- " **Envoy Air** " con un maximo de $299 \text{ minutos} \approx 5 \text{ horas}$ de retraso.
- " **Southwest Airlines** " con un maximo de $77 \text{ minutos} \approx 1,3 \text{ horas}$ de retraso.
- " **Endeavor Air** " con un maximo de $62 \text{ minutos} \approx 1 \text{ horas}$ de retraso.

Aerolinea	Minutos	Horas
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

Observemos este mismo grafico pero sin los outliers:

Retraso en salida en funcion de las aerolineas sin Outliers



Aca podemos observar que la aerolinea que mayor distribucion de tiempo de retraso en salida (**SIN OUTLIERS**) es " *EXpressJet Airlines* ", en segundo lugar la aerolinea " *AirTran Airways* " y en tercer lugar " *Envoy Air* ". ¡ Cabe destacar que estas dos aerolineas son las que mas casos tienen en que el vuelo salio más temprano !

Una cosa a remarcar en los boxplots de todas estas aerolineas es que la **media aritmetica** (representada por el cuadrado color rosa rojiza) se ve muy afectada por los outliers porque, como es la suma de valores dividido el numero de valores, tener ya uno o varios valores muy por fuera del rango de valores, genera cambios muy groseros.

Media de retraso de salida por aerolinea:

Aerolinea	Media(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

Desviacion Mediana Absoluta para el retraso en salida

Tambien podemos calcular la **desviacion mediana absoluta (DMA)** para que, aquellas aerolineas que su media se encuentre muy modificada por valores muy pequeños o muy grandes, puedan ser comparadas por su valor del medio (mediana) que no es tan facil de modificar.

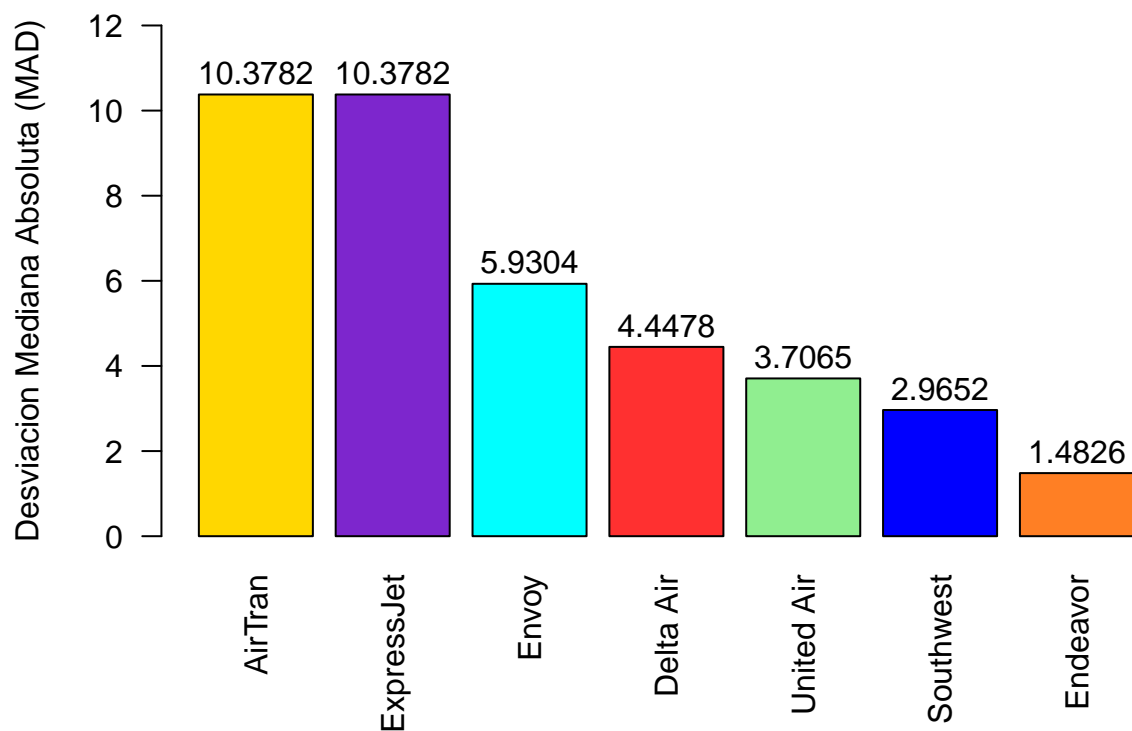
Calculo de la desviacion mediana absoluta:

Resultado = Valor absoluto de cada valor - la mediana de los valores

... Utilizando un valor estadistico $k = 1.4826$

(DMA) = mediana(Resultado) · k

Desviacion Mediana Absoluta del retraso en salida por carrier

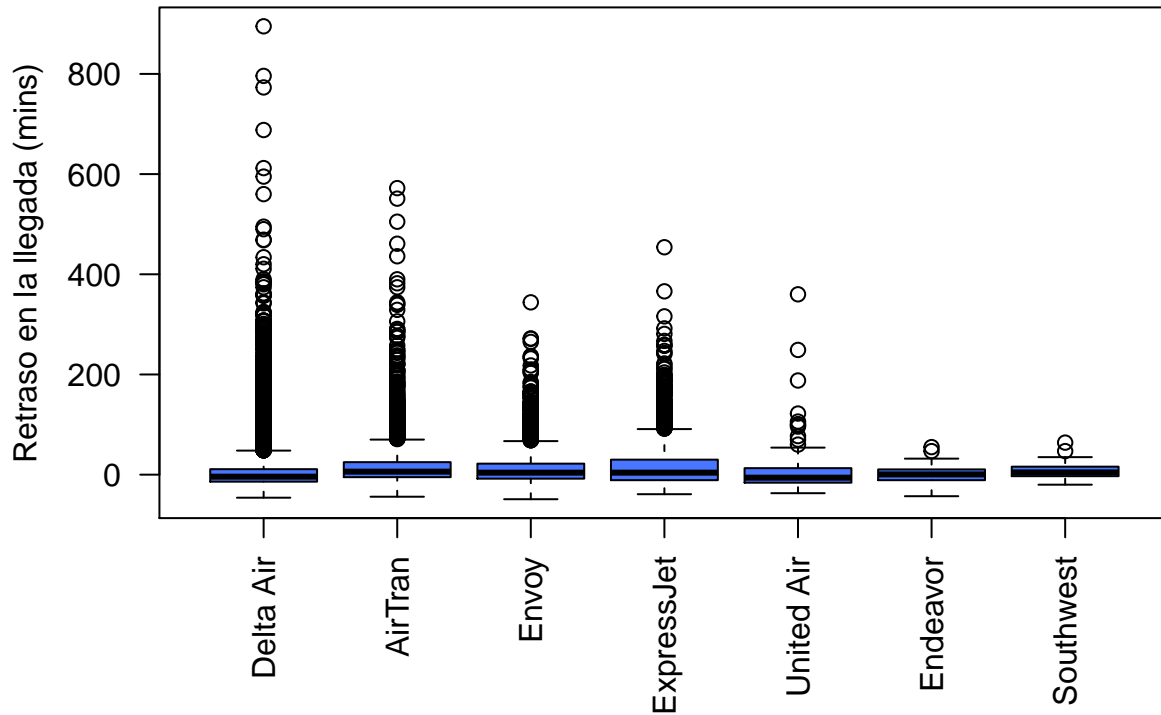


Por lo visto, por más de que tenga muchos *outliers* la empresa " *Delta Air Lines*", según el MAD, la dispersion en cuanto a el retraso de la salida del vuelo es dentro de todo estable (se mantiene cerca de un valor central).

No es tan variable el tiempo que hay que esperar para que salga el vuelo, en comparacion con otras aerolineas como " *AirTran Airways* " y " *ExpressJet Airlines* " que tiene una dispersion muy alta, lo cual interpreto como un mayor tiempo de espera para que salga un vuelo.

Retrasos en llegada

Retraso en llegada en funcion de las aerolineas con Outliers



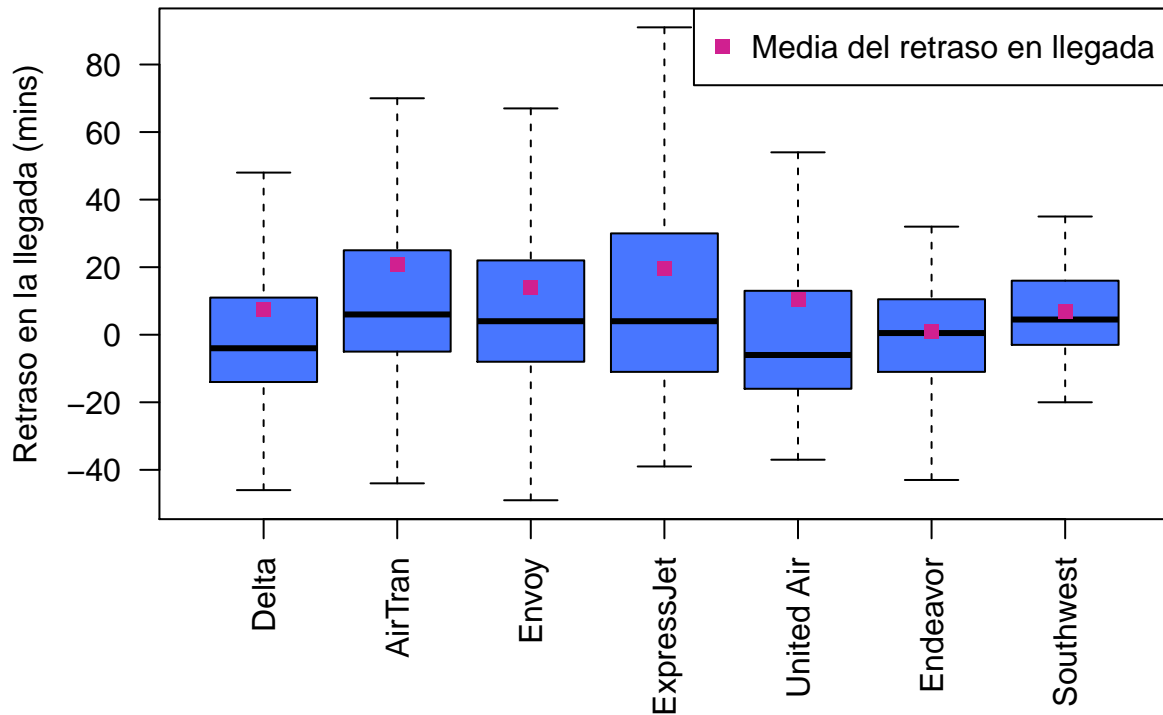
Vemos nuevamente que la empresa " *Delta Air Lines* " es la que tiene mayores problemas de retraso pero en este caso para cumplir con el horario de llegada, con un maximo de: $895 \text{ minutos} \approx 15 \text{ horas}$ de retraso.

Luego le sigue, tanto en mayor tiempo de retraso como en mayores problemas para cumplir el horario:

- " *Airtran Airways* " con un maximo de $572 \text{ minutos} \approx 10 \text{ horas}$ de retraso.
- " *ExpressJet Airlines* " con un maximo de $454 \text{ minutos} \approx 8 \text{ horas}$ de retraso.
- " *United Air Lines* " con un maximo de $360 \text{ minutos} = 6 \text{ horas}$ de retraso.
- " *Envoy Air* " con un maximo de $344 \text{ minutos} \approx 5 \text{ horas}$ de retraso.
- " *Southwest Airlines* " con un maximo de $64 \text{ minutos} \approx 1 \text{ horas}$ de retraso.
- " *Endeavor Air* " con un maximo de $55 \text{ minutos} \approx 1 \text{ horas}$ de retraso.

Aerolinea	Minutos	Horas
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

Retraso en llegada en funcion de las aerolineas sin Outliers



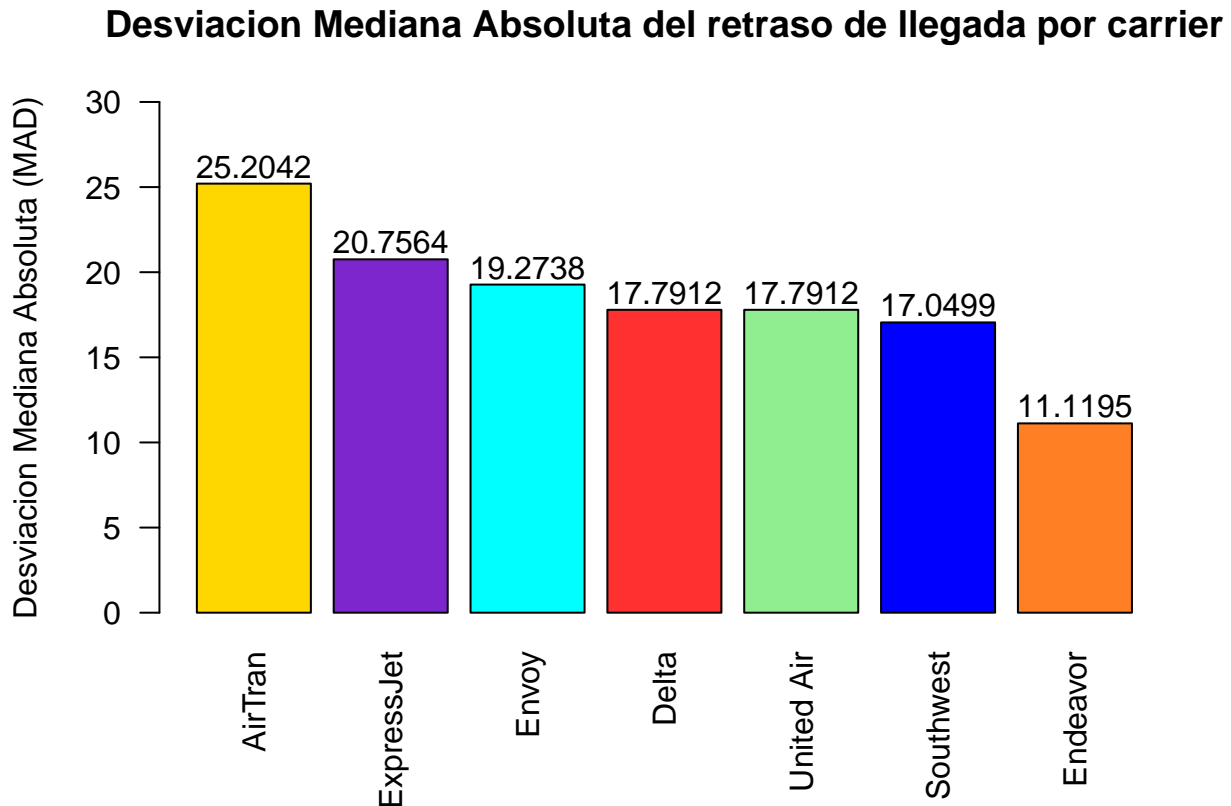
Media de retraso de llegada por aerolinea:

Aerolinea	Media(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

Como podemos observar, el carrier con retrasos de llegada menos dispersos (**SIN OUTLIERS**) es " *Southwest Airlines* " y el que tiene retrasos más dispersos es la aerolinea " *ExpressJet Airlines* ".

Deteniendonos en la empresa " *Delta Airlines* " podemos ver que tiene una distribucion de tiempos de espera mucho más variado que la aerolinea " *Southwest Airlines* ", además que es la que mayor retraso de llegada maximo tiene (15hs) pero aún así no es tan disperso como el de " *ExpressJet Airlines* ".

Desviacion Mediana Absoluta



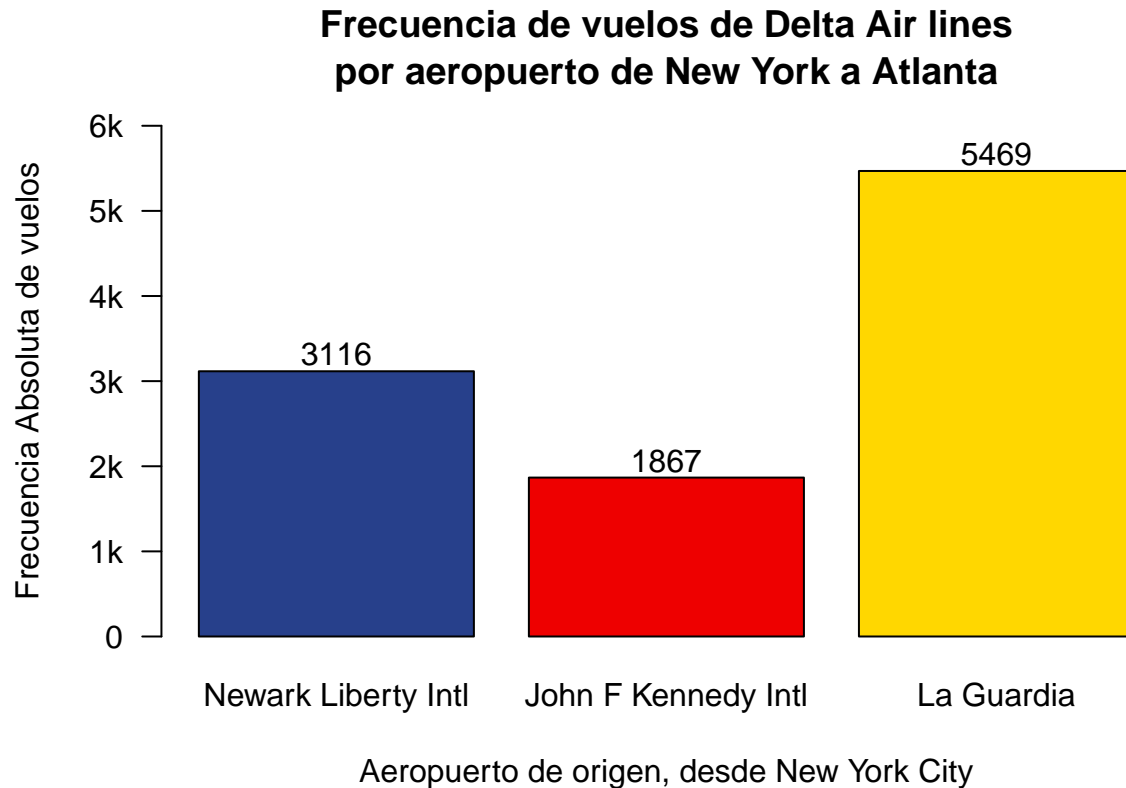
Analizando ahora el *MAD* del retraso de llegada por carrier, podemos ver que la empresa " *Endeavor Airlines* " es la que menos distribuidos tiene su retraso de llegada, siendo la que más varía la aerolínea " *AirTran Airlines* ".

Enfocando la mirada una vez más sobre la aerolínea " *Delta Airlines* ", podemos ver que si bien sus valores de retrasos están bastante distribuidos, no están tan distribuidos como otras empresas pero tampoco están pobremente distribuidos como otras empresas, sino que más bien se encuentra en un punto medio con respecto a otras aerolíneas.

Lamentablemente, el dataset no incluye el precio de los boletos de cada aerolínea para ir a Atlanta, lo cual sería muy útil para saber porque la empresa " *Delta Airlines* " es la más utilizada para viajar, porque viendo los retrasos, se ve bien que " *Endeavor Air* " es la que menos retrasos tiene, esto significa que es la que más cumple con los horarios.

Vuelos por aeropuerto de NYC con destino a Atlanta por Delta Airlines

Para concluir, vamos a ver que aeropuerto es el más usado por la aerolínea " *Delta Air Lines* ".



Podemos ver que la mayoría de los vuelos que salen de *New York* con destino a *Atlanta* por la aerolínea " *Delta Airlines* " salen del aeropuerto "***La Guardia***".

Conclusión:

Podríamos decir que posiblemente la aerolínea " *Delta Airlines* " es la que más vuelos tiene posiblemente por su fama, aunque no es la empresa que menos retrasos tenga, sino más bien es una intermedia. Tal vez será por su mayor disponibilidad de vuelos en muchas fechas y horarios. Además de que tal vez es la empresa con boletos más baratos.

En tal caso si estuviéramos apurados con la necesidad de viajar desde New York a Atlanta y llegar rápidamente, seguro tomaríamos el vuelo de " *Endeavor Airlines* " porque es la que menos retraso salida y de llegada tiene. Tal vez no es la mas usada porque tiene menos disponibilidad de vuelos, menos fechas y horarios de vuelos o porque es muy cara. Con los datos disponibles no podemos saberlo.