

US commercial flights analysis

Davide Masticci



Dataset

Research and Innovation Technology Administration

- 22 years of data [1987 - 2008]
- 123 M of observations
- 29 different variables

Data Management

Before

After

22 CSV

1 SQLite table

Hard to read, join and filter data

Easy to read, join and filter data

Same reading and retrieving data speed

Faster reading and retrieving data speed on
some field (Indices)

First Insights

Distance (miles)

Origin

Dest

6

Spokane Intl
(Washington)

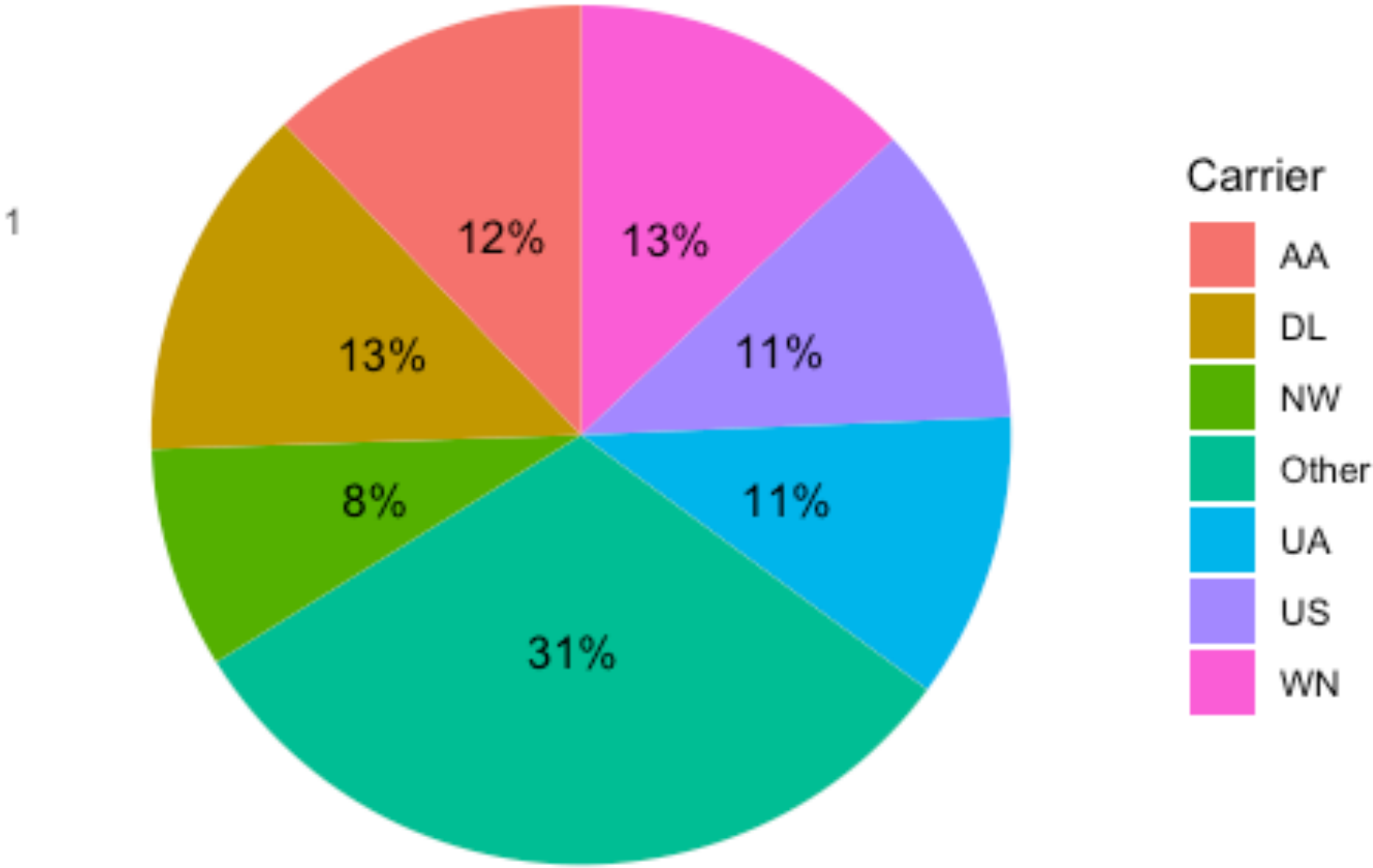
Fairchild AFB
(Washington)

4983

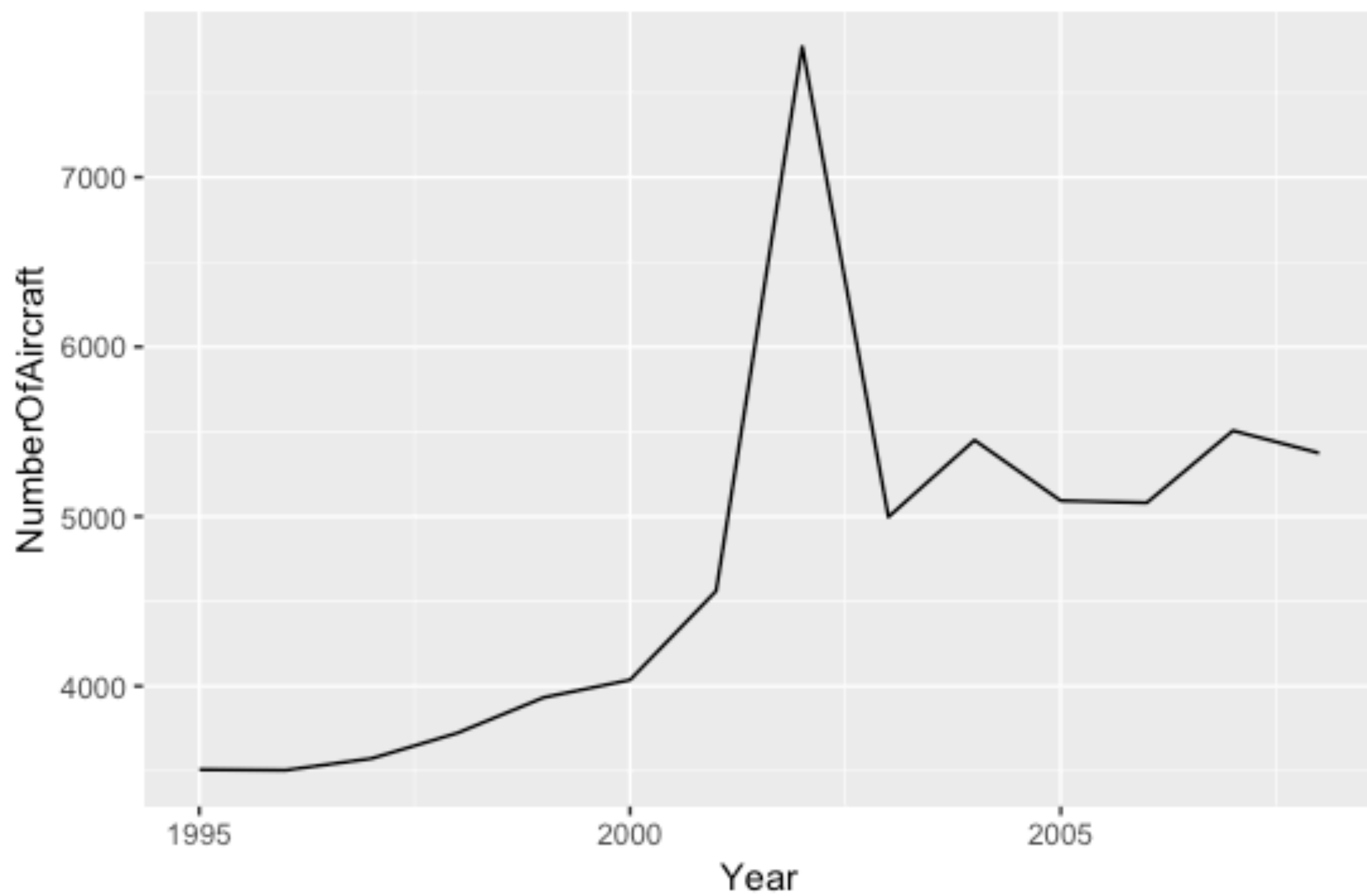
John F. Kennedy Intl.
(New York)

Honolulu Intl.
(Hawaii)

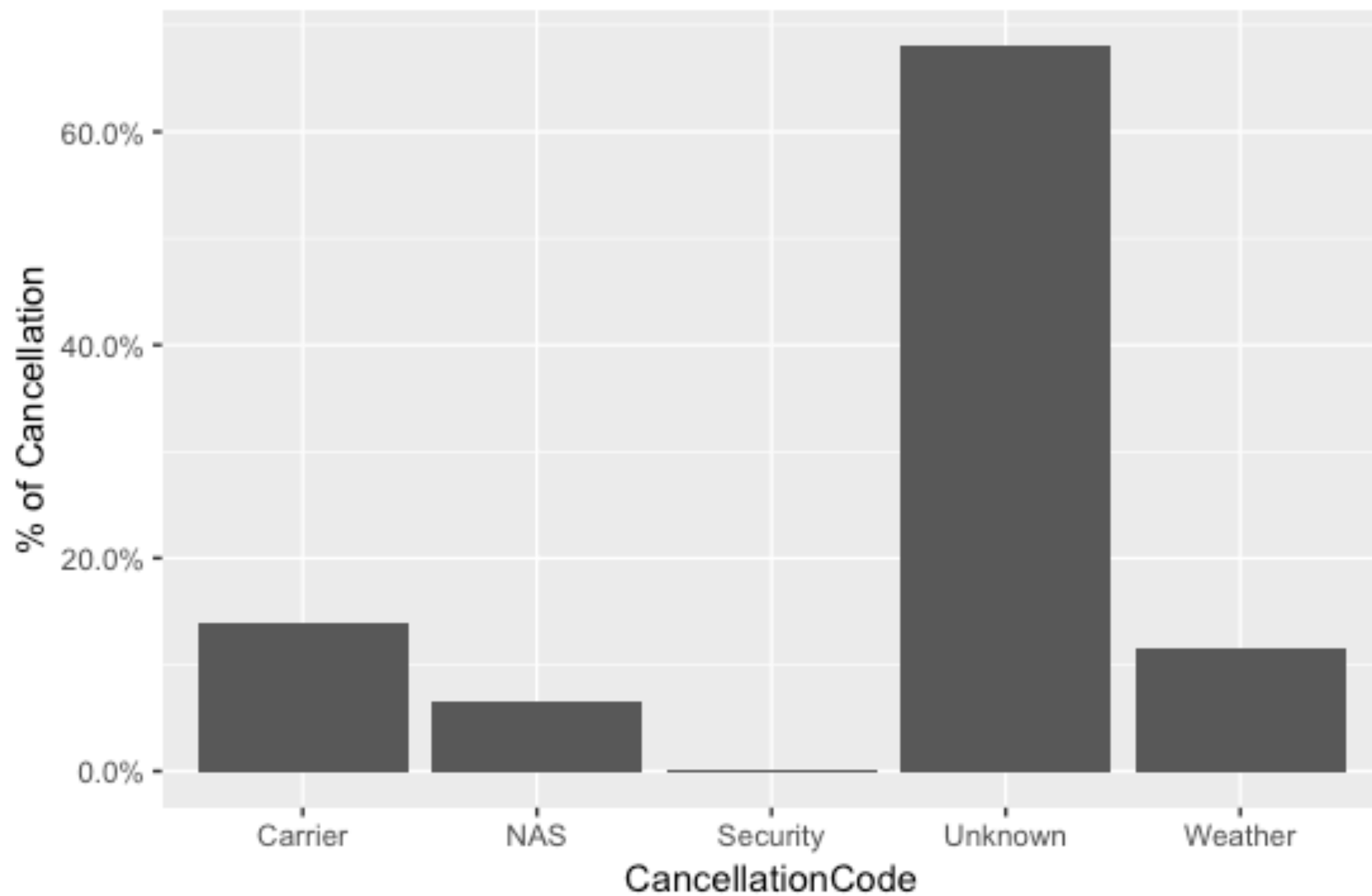
Carrier Distribution



Number of aircraft per year



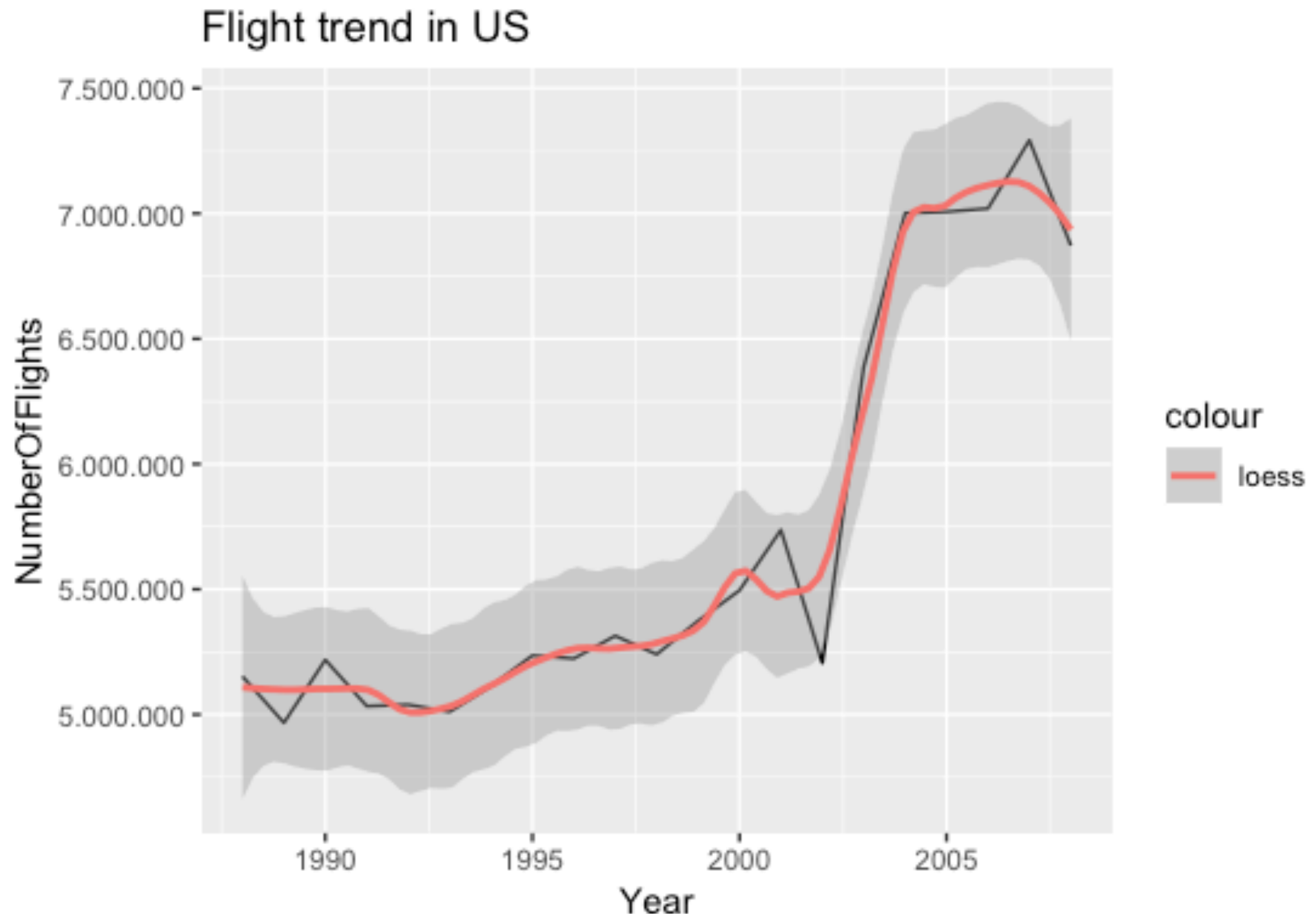
Causes of cancellations



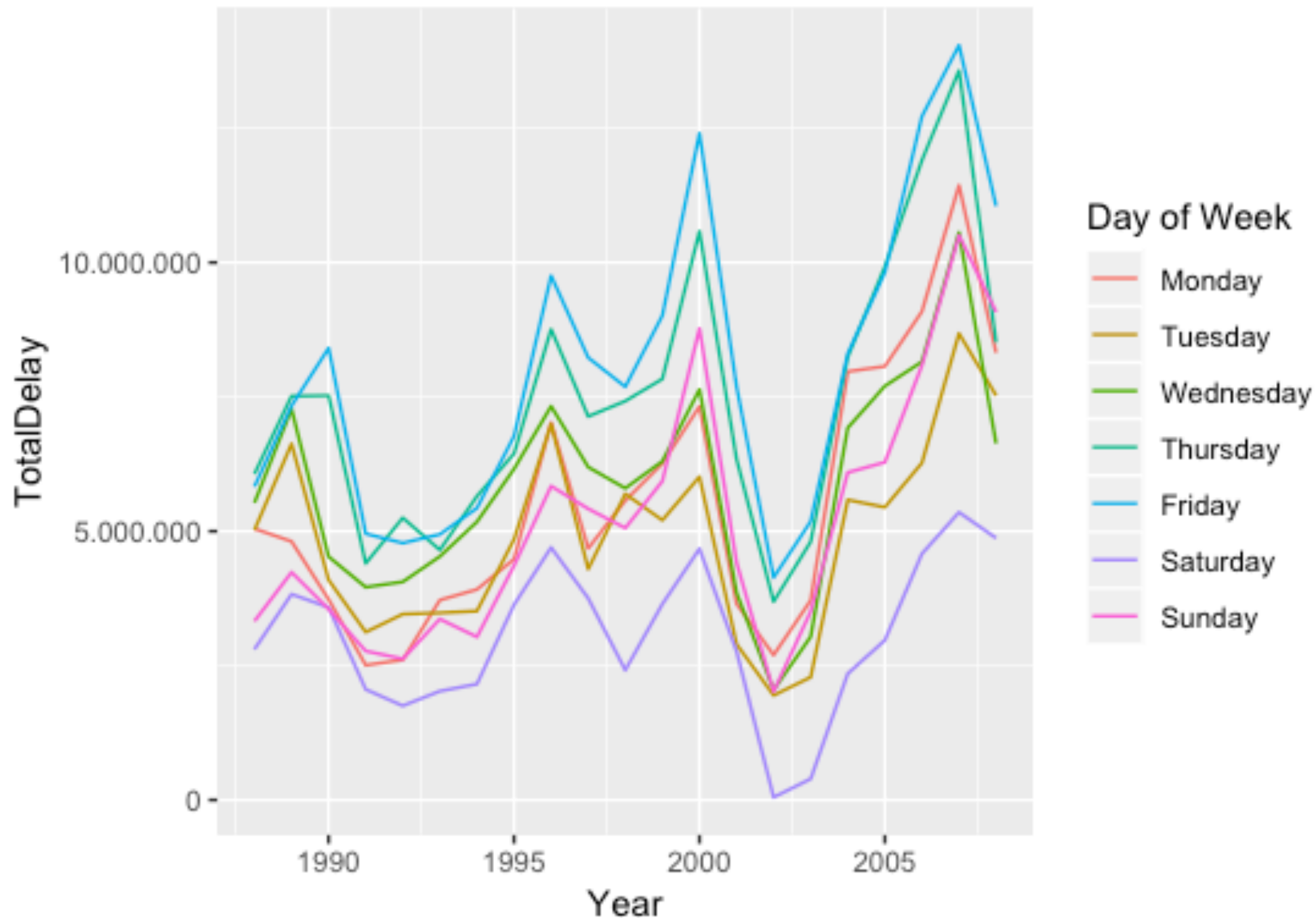
Analysis

1. How is the trend of commercial flights in the US?
2. Which is the best (worst) day of the week to take a flight?
3. Which is the best (worst) season to flight in terms of delay?

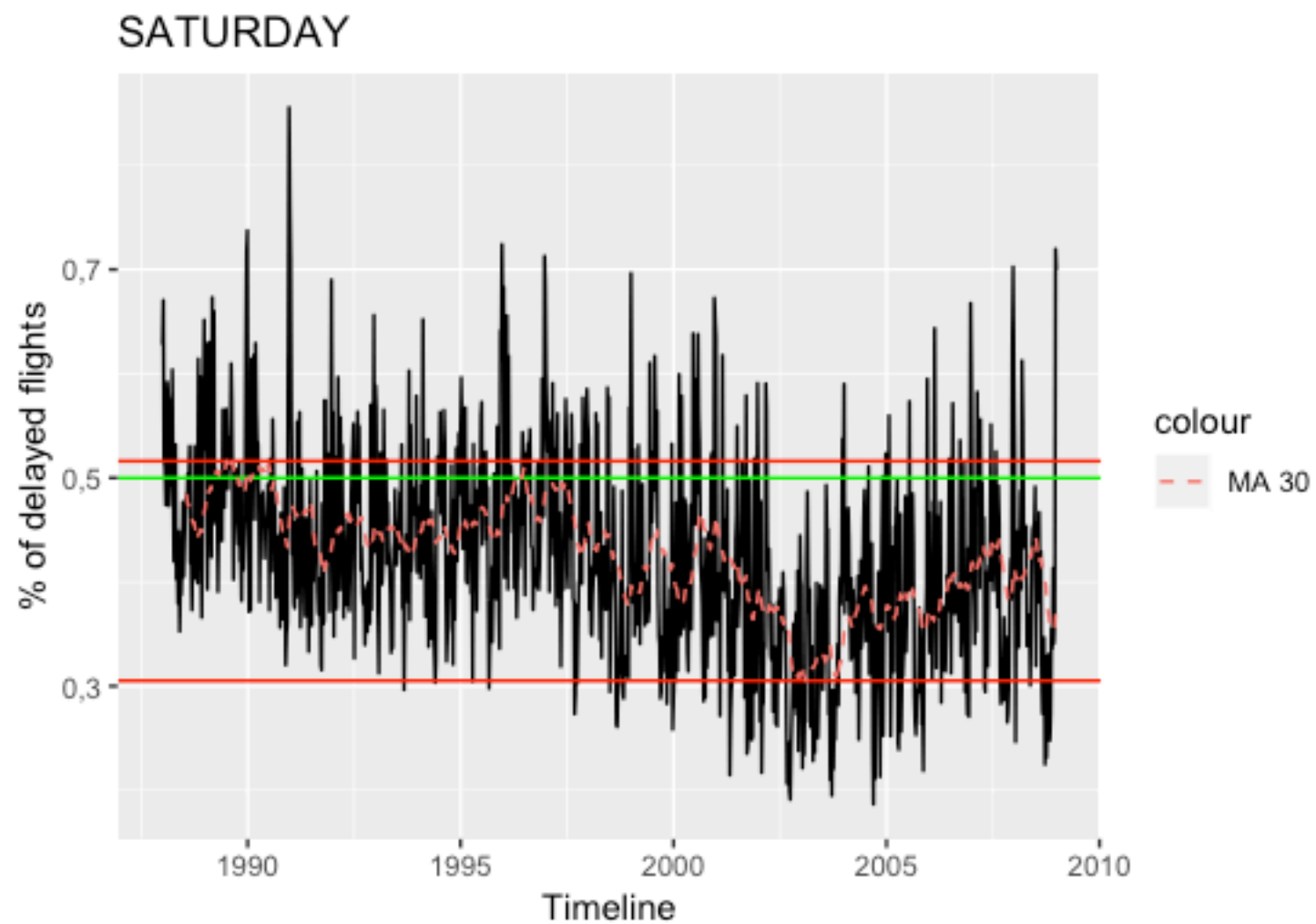
How is the trend of commercial flights in the US?



Which is the best (worst) day of the week to take a flight ?

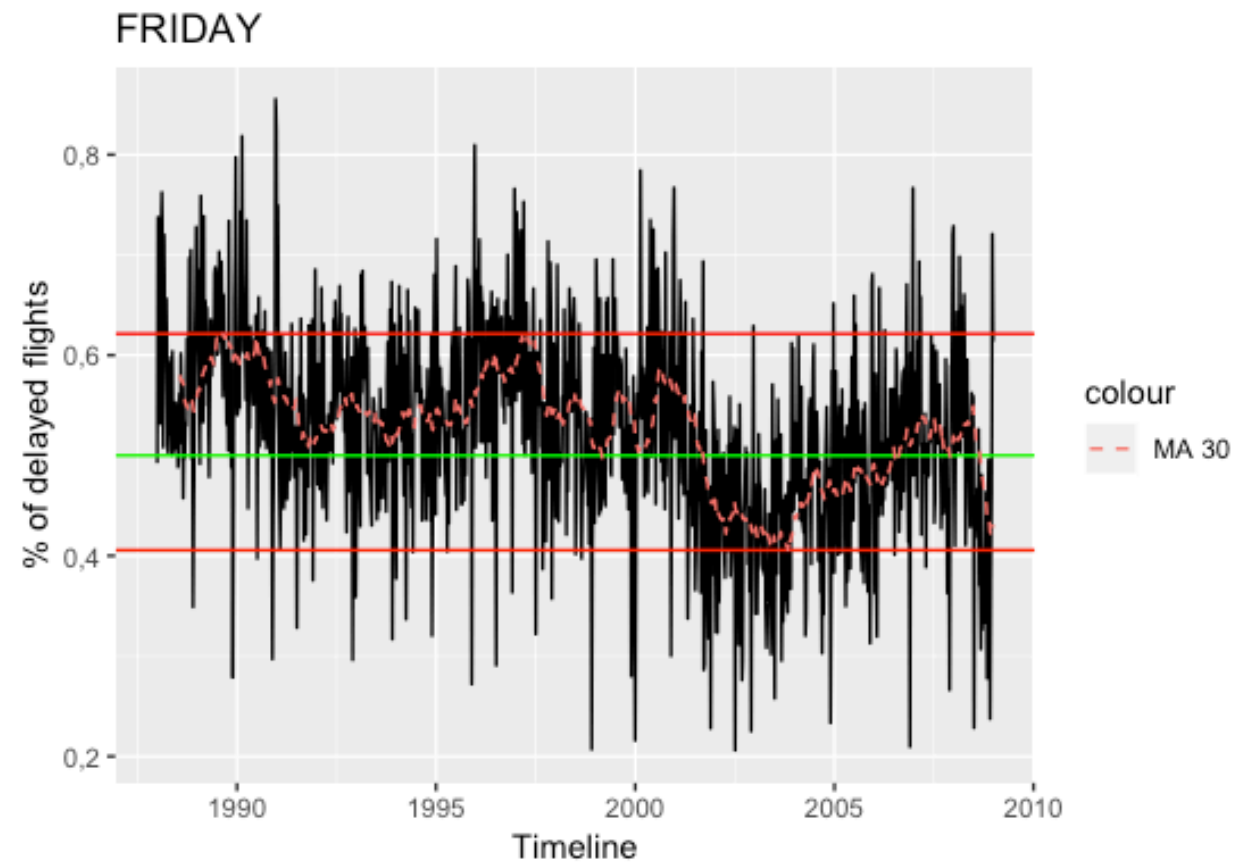


Best Day

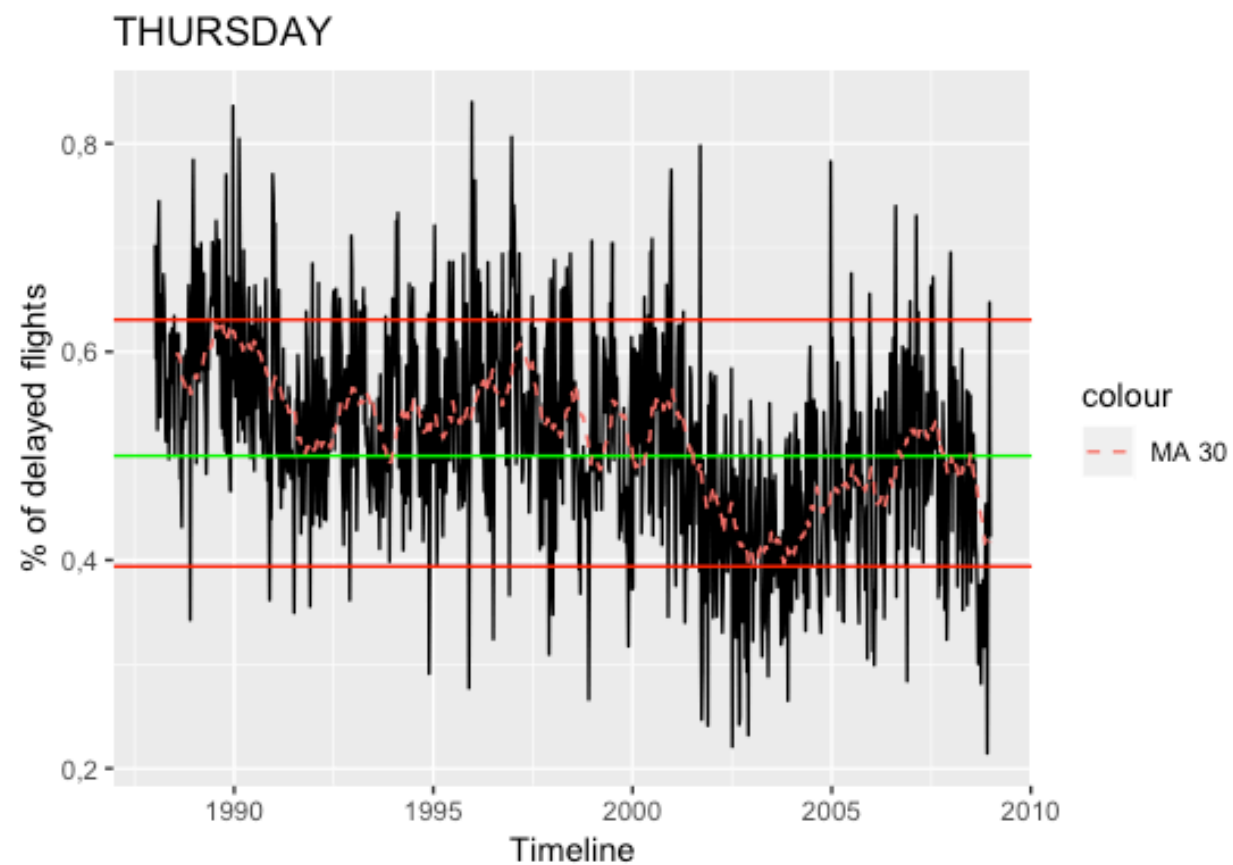


MA fall into [30% ; 52%]

Worst Days

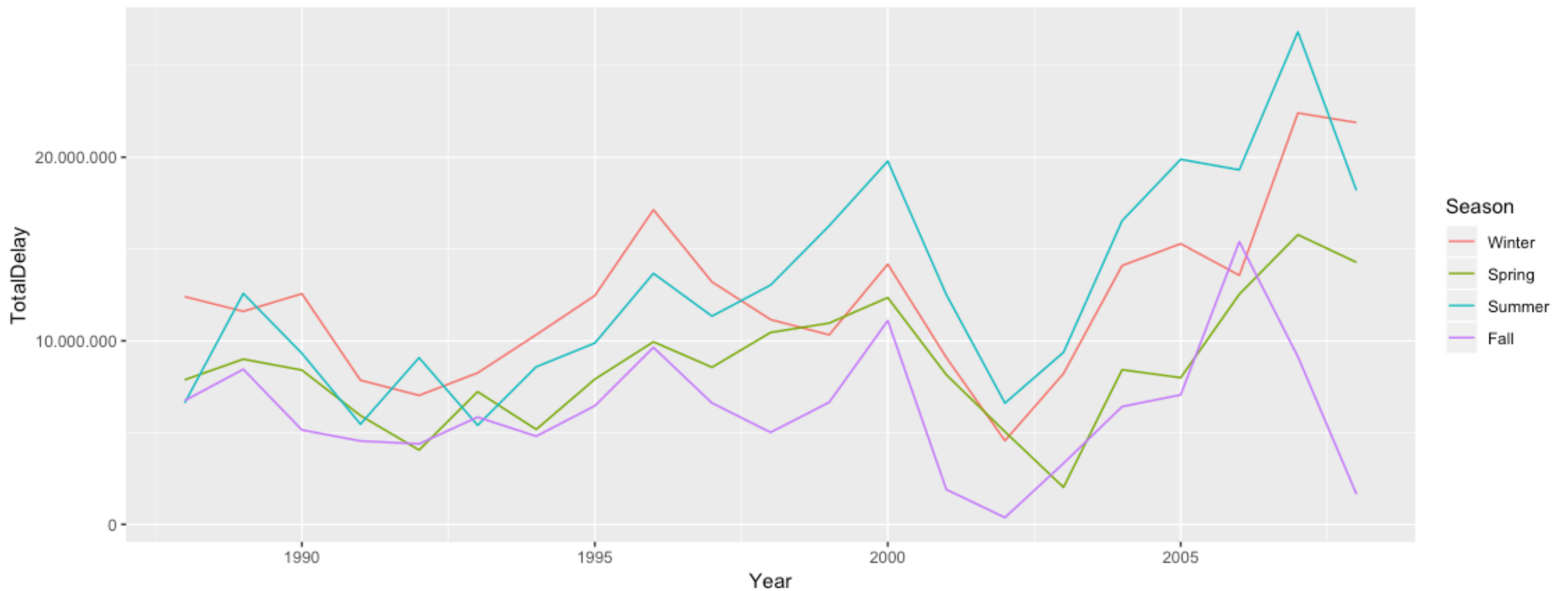


MA fall into [40% ; 63%]

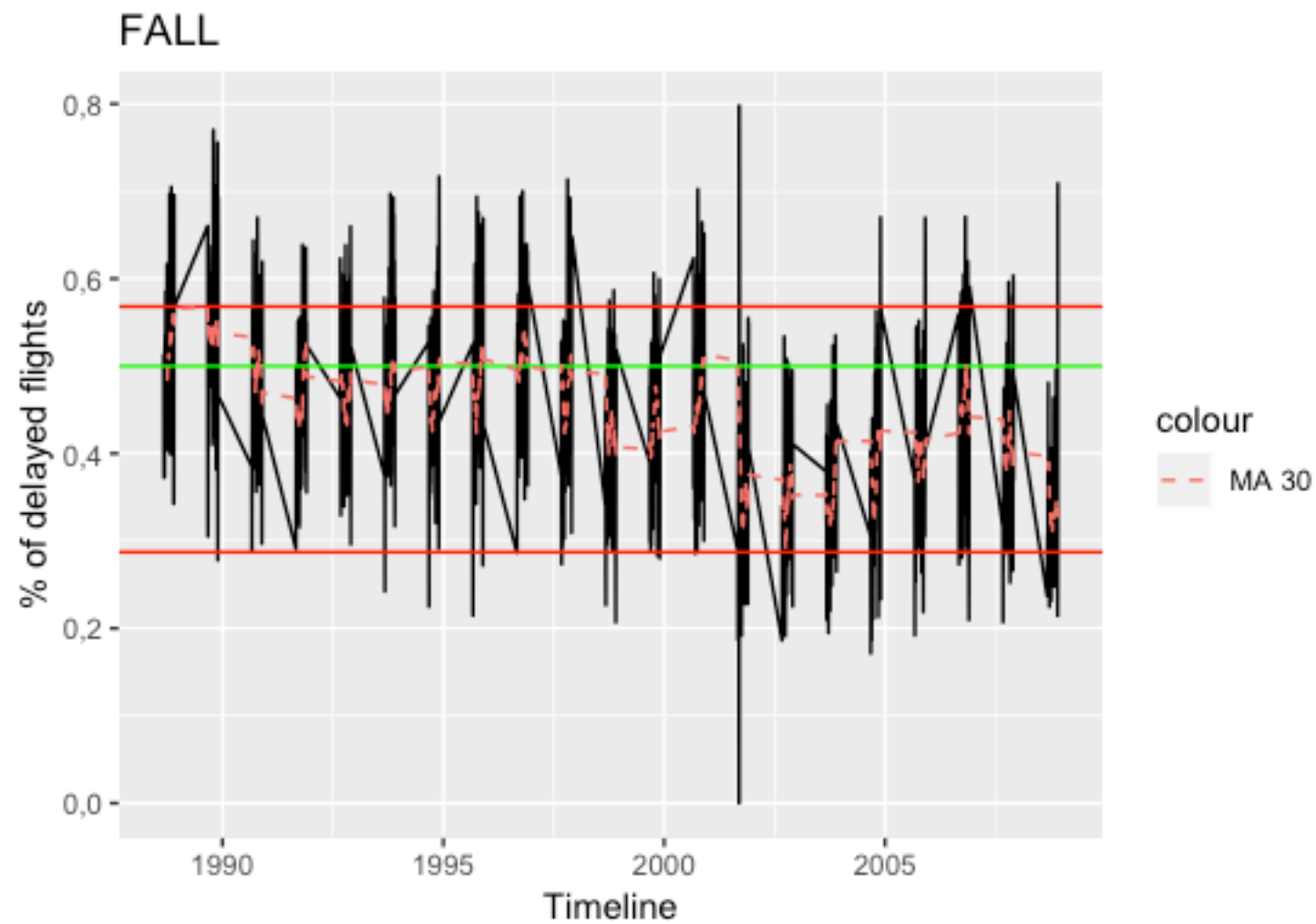


MA fall into [39% ; 63%]

Which is the best (worst) season to fly in terms of delay ?

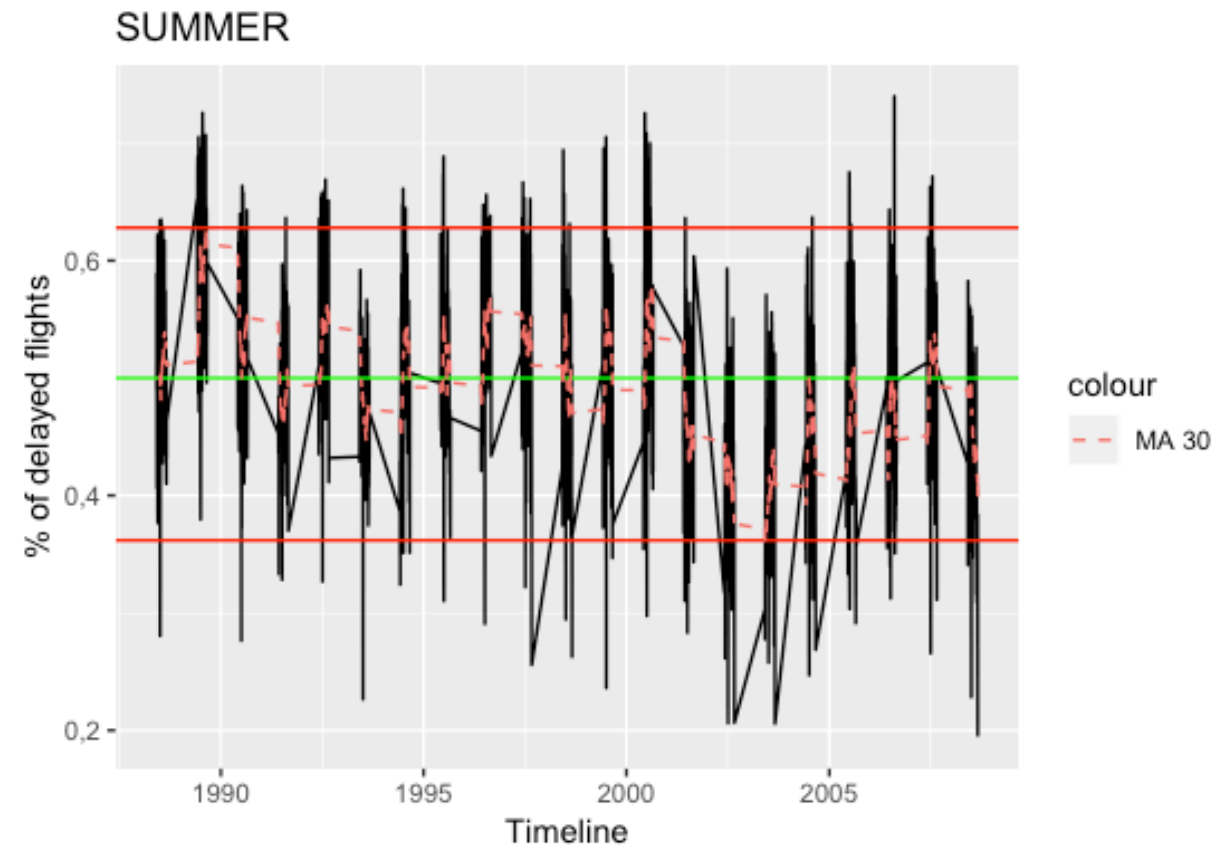


Best Season

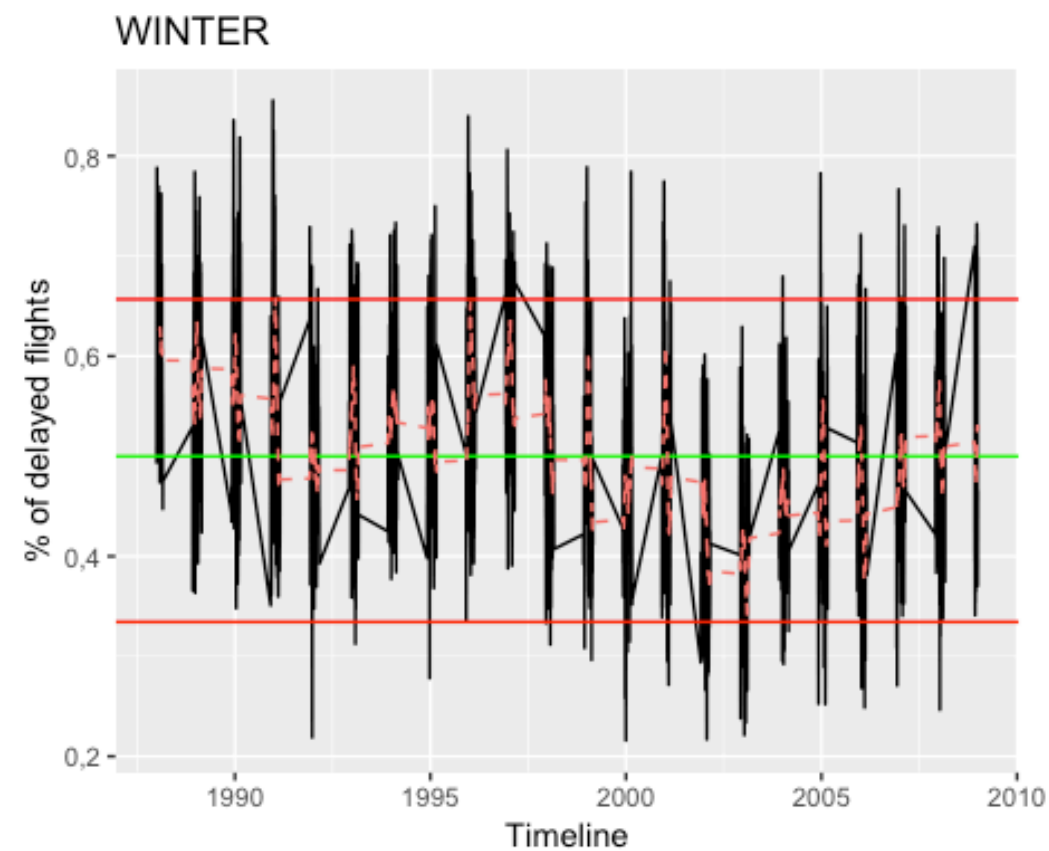


MA fall into [29% ; 57%]

Worst Seasons



MA fall into [36% ; 63%]



MA fall into [33% ; 66%]

Analisisys summary

- Good news for the carriers' wallet because the number of flights seem to increase over the year
- If you want to reduce the probability of having a delay, travel on Saturday and avoid Friday and Thursdays
- Do you want to have less problem during your holidays ? Plan them during the Fall and avoid Summer and Winter

“Viaggiare e analizzare i dati sono correlati dal fatto che alla fine si è sempre imparato qualcosa.”

–Davide Mastricci