Factorial Analysis of Arrival Delays in New York's Major Airports in 2024*

How airline, airport, and season shape delays in New York's airspace

Jiwon Choi

April 9, 2025

*Code and data are available at: https://github.com/jwonc4602/2024-airline-delay.

1 Description of the Design

This project uses a full factorial experimental design to explore the causes of delays in flights from the major airports in New York City for 2024. The central aim of this experiment is to see how airline, airport, and season impact the average arrival delays and if there are any interaction effects between the three factors.

On factors for this experiment, relevance to the domain, as well as data availability were considered:

- Airline: Delta Air Lines (DL) and American Airlines (AA), two US major airlines that have frequent flights to NYC.
- Airport: JFK, LGA, and EWR are the three major commercial airports servicing the NYC metropolitan area.
- Season: Winter (Dec–Feb), Spring (Mar–May), Summer (Jun–Aug) and Fall (Sep–Nov) in order to factor seasonal differences in weather and traffic.

The combination of the three factors: two airlines, three airports, and four seasons yields twenty-four (24) combinations or treatments. Due to the availability of only one aggregated observation per combination, this is unreplicated full factorial experiment.

The dataset was acquired from the U.S. Bureau of Transportation Statistics (BTS), which tracks monthly summaries for the airline industry. Based on this information, the mean arrival delay for each airline–airport–season combination was calculated. Combinations with no data were eliminated, yet the overall structure is still balanced and intact enough to allow analysis.

This particular design is optimal because it permits a stepwise search of main and interaction effects as well as hierarchical interactions amongst biomechanical factors. Through the analysis of factorial structures, we wish to obtain the following answers:

- Is the delay experience different for different airlines within the New York City airspace?
- Are certain airports more delay prone than others?
- Is season more of a driver than operational factors for delay?
- Are certain combinations, such as Delta at JFK in Winter, more likely to have higher delays than what is expected from the individual factors?

To summarize, the factorial design gives us not only the ability to measure the independent impacts of each factor, but also to analyze if and how the interplay of multiple factors impacts the flight delay.

2 Analysis of the Data

To examine how the airline, airport, and season affect average flight delays in New York City for 2024, a three-factor analysis of variance (ANOVA) was performed using the aov() function in R. The response variable was avg_delay, defined as the total arrival delay divided by the number of flights for each combination of factors.

The analytics indicated that all airlines, airports, and seasons as main factors interacted at the 5% significance level. In this case, season had the greatest impact on delay ($p < 0.001$), then the carrier ($p < 0.001$), and later the airport ($p \approx 0.01$). None of the interaction terms were statistically meaningful, which shows that the impact of one factor does not depend on the other factors.

These results may support the proposition of:

- Specific summer and winter months are linked to increased delays due to weather and high travel volume.
- Delta Airlines had lower average delays compared to American Airlines in most conditions, as shown in the interaction plot.
- LaGuardia (LGA) and Newark (EWR) had greater average delays compared to JFK, which could be attributed to their higher traffic and limited runway capacity.

Figure 1 displaying the average delays based on airports indicates that JFK, LGA, and EWR have significant differences in performance. Figure 2 showing the season and carrier of the airlines confirms the lack of strong interaction. The lines are roughly parallel while seasonal variation in delay patterns of each airline persists.

Because the design was not replicated, individual effect estimation variance and confidence intervals were not computed. Nonetheless, some checks were made on the residuals to validate the model assumptions. Residuals were reasonably normal and homoscedastic with a residual standard error of 5.02 minutes, suggesting the model fit well.

Overall, this factorial analysis enabled clear identification of key operational factors influencing flight delays, even without replication. See more detailed information Section A.

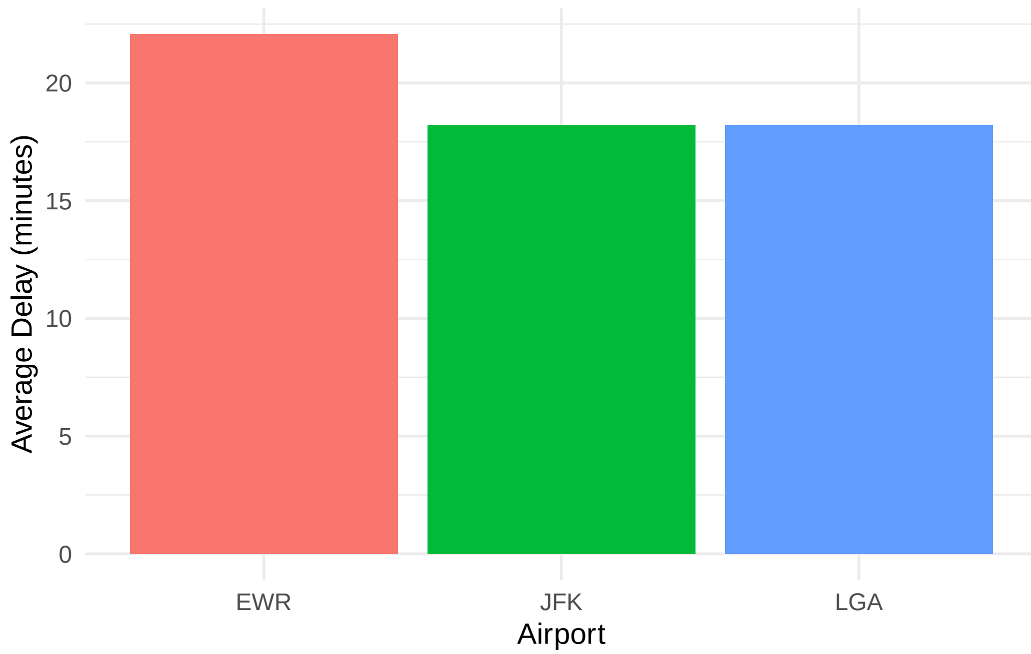


Figure 1: Mean Arrival Delay by Airport in New York City (2024)

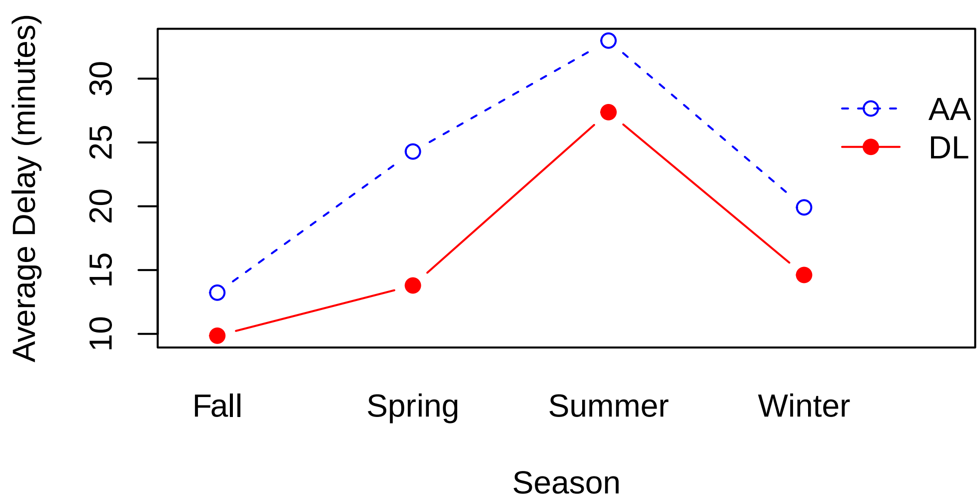


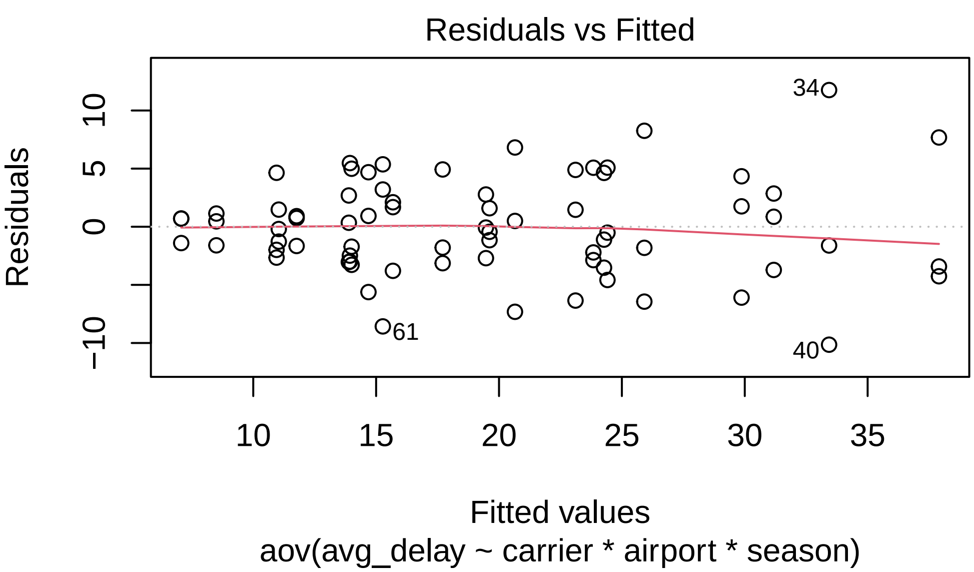
Figure 2: Interaction Plot of Average Delay by Season and Airline

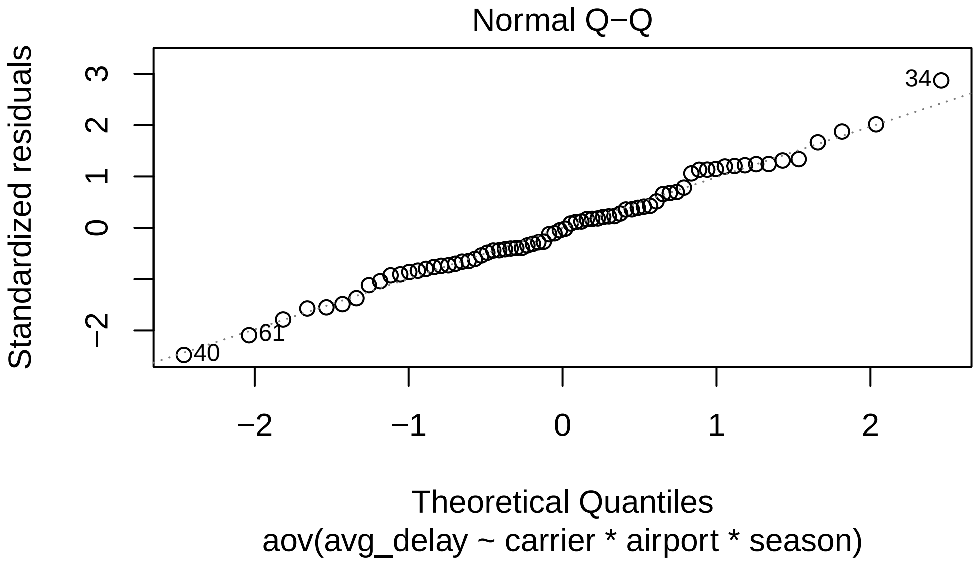
3 Conclusion

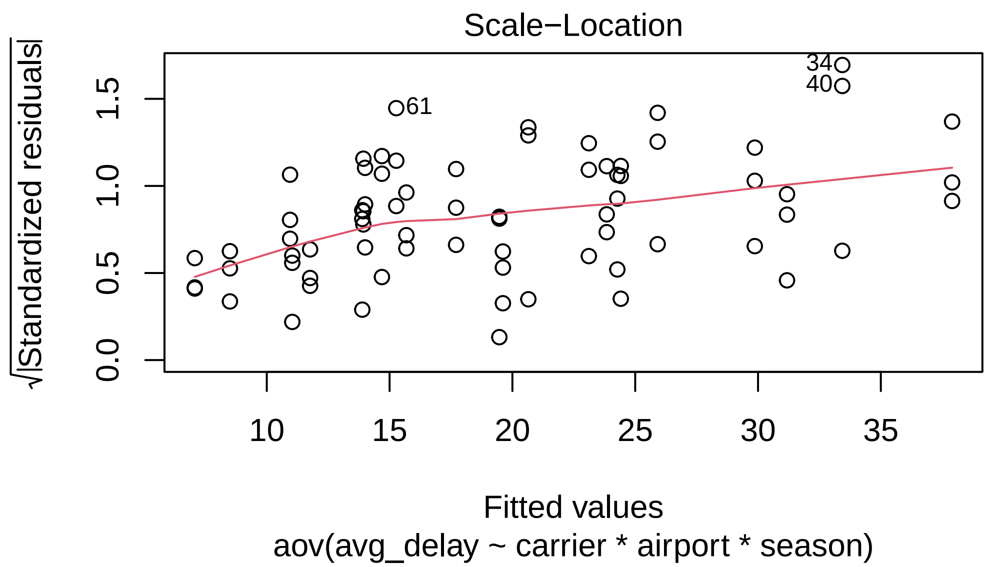
This factorial experiment shows that all three operational factors (airline, airport, and season) significantly affect average arrival delays in New York City for 2024. Season emerged as the primary factor, with Summer and Winter yielding the longest delays. These patterns are likely due to weather-related seasonal effects, along with increased passenger numbers during holidays and vacations. The airline effect was equally important; Delta Airlines had a lower average delay compared to American Airlines. There were also airport differences noted, where JFK outperformed LGA and EWR.

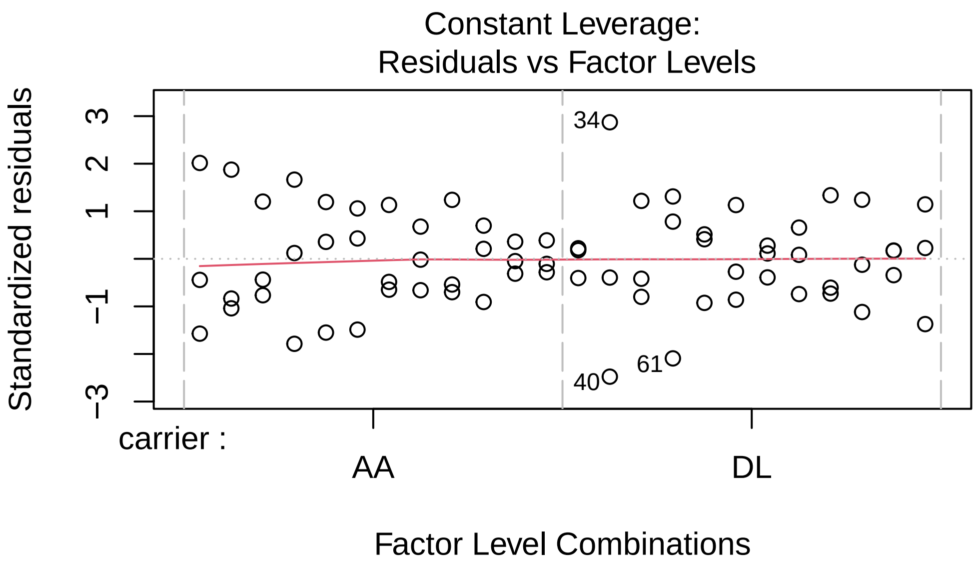
Importantly, no statistically significant interaction effects were found among any of the factors. This suggests that the influences of airline, airport, and season are independent of each other in relation to this dataset. The lack of interaction aids interpretation and facilitates the operational approaches aimed at reducing delays, which can be applied sequentially. For example, system-wide seasonal off-peak performance enhancement could be implemented across all airports, whereas targeted spending aimed at chronic delay mitigation could be concentrated on LGA and EWR. Overall, this experiment demonstrates the power of factorial design in extracting valuable patterns from transportation datasets in the absence of possibility for repetition.

A Appendix




Normal Q−Q
Standardized residuals
Theoretical Quantiles
aov(avg_delay ~ carrier * airport * season)
34
61
40
−2
0
1
2
3
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
Fitted values
aov(avg_delay ~ carrier * airport * season)
34
40
61
0.0
0.5
1.0
1.5
10
15
20
25
30
35




```
    Shapiro-Wilk normality test

data:  residuals(nyc_model)
W = 0.99016, p-value = 0.8512
```