

STAT 167 Final Project Report

Chloe Lavery, Shirley Lee, Zach Nguyen, Mihir Ranjan, Rhea Verma, Krishna Vijayendra

2025-06-08

Contents

Introduction	4
Methodology & Analysis Plan	5
Research Questions	5
Modeling Approach	5
Exploratory Data Analysis (EDA)	6
Arrival and Departure Delays by Airport	6
Setup	6
ANOVA Table	6
Tukey Test	6
Rank by mean delay	6
Conclusion	7
Distribution of Flight Times	7
Setup	7
ANOVA Table	7
Tukey Test	8
Distribution Visualization	8
Conclusion	8
Monthly Flight Volume by Airline/Airport	9
Conclusion	9
Data Exploration and Visualization	10
Question 1: What pattern(s) can be observed regarding flight punctuality in NYC? .	10
Setup	10
How do arrival and departure times vary throughout the day across NYC airports?	10
When during the day are flights most likely to leave and arrive on schedule/late?	11
What is the average delay time by airline?	12

Are weekend flights more punctual than weekday ones?	13
Are there certain months more punctual with flight departures/arrivals than others?	14
Conclusion	14
Question 2: Which of NYC's airports were most operationally efficient in 2013, and how has that changed to 2023? Which NYC airports have become more or less operationally efficient over the past decade?	15
Data Preparation	15
Delay Comparison	15
Flight Volume	16
Linear Model: Delays Over Time	16
Average Arrival/Departure Delay Comparisons by Airport	17
Boxplot of Arrival Delay Trends	18
Line Graph of Arrival Delay Trends By Month	19
Line Graph of Departure Delay Trends by Hour	20
Conclusion	20
Question 3: How does weather impact flight punctuality?	21
Setup	22
How does weather affect departure and arrival times?	22
Is there a monthly trend in weather and flight delays?	23
Are some airports more resilient to delays in severe weather than others?	24
How does weather predict flight delays? (linear regression)	24
Conclusion	25
Question 4: Are certain geographic areas or types of carriers linked with better or worse flight arrival times and operational efficiency? What geographic or airport characteristics are linked with consistently superior on-time performance?	26
Setup	27
Which of the three origin airports had the least average departure delays and which had the most?	27
Which of the three origin airports had the least average arrival delays and which had the most?	28
Which destination airport(s) had longer average departure delays and which had shorter?	29
Which destination airport(s) had longer average arrival delays and which had shorter?	30
Do average departure delays vary a lot between carriers? If so, which airline(s) had longer delays and which had shorter?	31
Do average arrival delays vary a lot between carriers? If so, which airline(s) had longer delays and which had shorter?	32
Conclusion	33

Question 5: Do air-traffic disruptions (higher average departure/arrival delays) predict short-term equity returns for the major carriers?	34
Setup	34
How do departure/arrival delays impact American airline stock returns?	34
Estimating prediction error using 10-fold cross-validation	35
Analysis	35
Conclusion	36
Conclusion & Implications	37
Real-World Impact	37
Limitations	37
Authors' contributions	38
Link(s) to Data/Code	38

Introduction

In this project, we wanted to look at several different influences and factors of air travel. These influences ranged from transport policy and finances to airport/airline efficiency and operations research. Once we identified these specific areas of interest, we drafted questions to further research these areas of study.

These questions are a mix of the topics of interest, but ultimately, all are aimed at making data-focused analyses, finding patterns, determining significance, and drawing conclusions. With strong data-focused results and analysis, it is possible to identify problems and find ideal solutions.

For example, if an airport did not have a sustainable budget, we could analyze its efficiency and find any external issues, such as airlines. Finding the problem and solving it can greatly improve the airport's sustainability and overall efficiency. There are nearly an infinite number of applications, such as the example given, by asking such questions and applying such research.

Methodology & Analysis Plan

Research Questions

1. What pattern(s) can be observed regarding flight punctuality in NYC?
2. Which of NYC's airports were most operationally efficient in 2013, and how has that changed to 2023? Which NYC airports have become more or less operationally efficient over the past decade?
3. How does weather impact flight punctuality?
4. Are certain geographic areas or types of carriers linked with better or worse flight arrival times and operational efficiency? What geographic or airport characteristics are linked with consistently superior on-time performance?
5. Do air-traffic disruptions (e.g. higher average departure/arrival delays) predict short-term equity returns for the major carriers?"

Modeling Approach

- Setup/Data Preparation
 - Extract any data that may be of use.
 - Clean data to prevent inaccurate calculations.
 - If needed, make sure to join different datasets.
- Data Analysis
 - Calculate variables such as mean and p-values.
 - Conduct any significance tests.
 - Make any comparisons in the context of the question.
- Visualization
- Use ggplot to provide visuals on the area of study to better understand.
- Conclusion/Results
 - Use plot visuals and important values and see if there are any patterns.
 - Conduct analysis.

Exploratory Data Analysis (EDA)

In our EDA, we had a few main goals. First, we wanted to ensure the data set was sufficient for our areas of study and whether it was necessary to add and cross-reference other data sets. We also aimed to clean up and quickly analyze the data and ensure there were not many outliers or missing values. Finally, we wanted to see if there were any specific patterns or trends in the data in some of the places of study. Under we have our three areas of focus.

1. The first was the arrival and departure delays of all three New York airports. Were they significant?
2. Second, is there variation in flight airtime by month?
3. Third, did the volume of flights at the three airports have any unusual patterns? Did some airlines also have notable patterns?

Arrival and Departure Delays by Airport

Setup

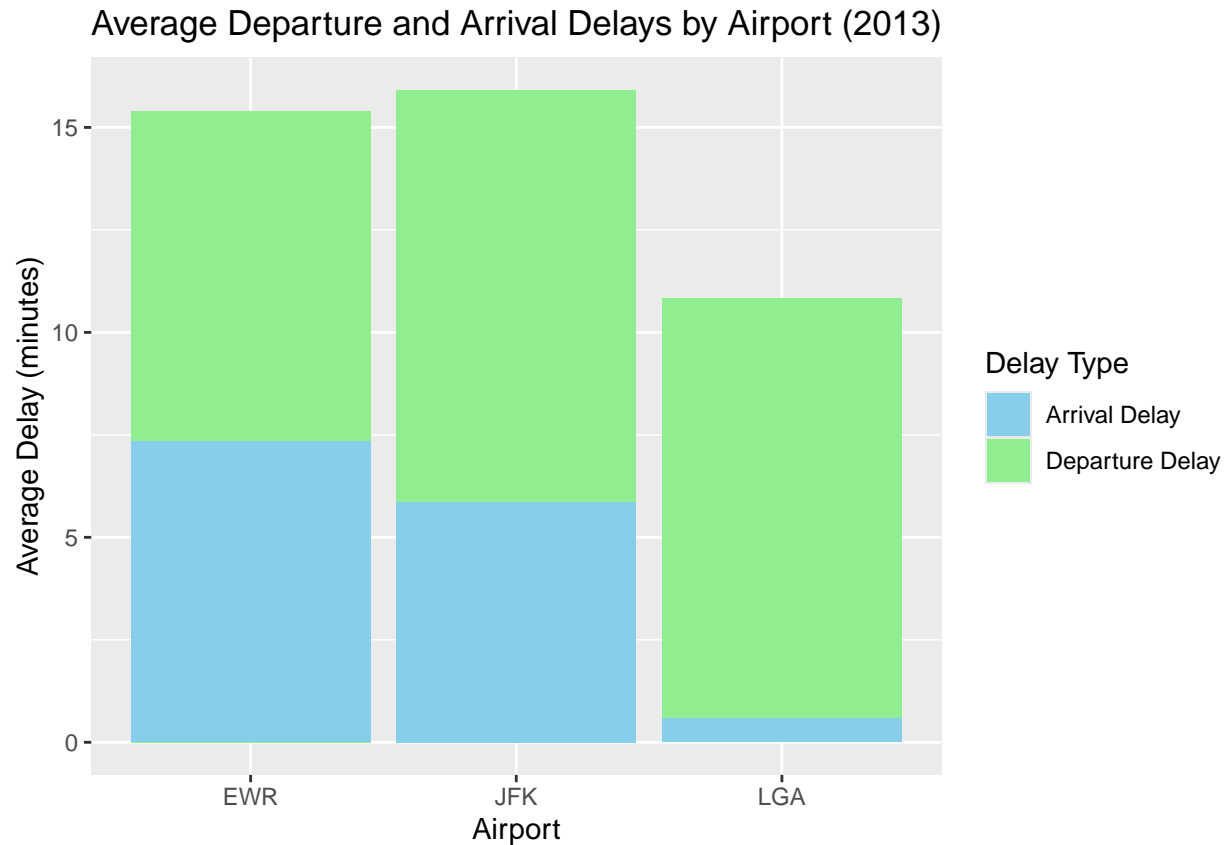
ANOVA Table

```
##              Df    Sum Sq Mean Sq F value    Pr(>F)
## origin          2 5.879e+06  2939560   935.62 < 2e-16 ***
## delay_type      1 1.909e+07 19090037 6076.11 < 2e-16 ***
## origin:delay_type 2 2.006e+05   100303   31.93 1.37e-14 ***
## Residuals      847426 2.662e+09     3142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tukey Test

Rank by mean delay

```
## # A tibble: 6 x 3
##   origin delay_type mean_delay
##   <chr>   <chr>         <dbl>
## 1 JFK    dep_delay        15.9
## 2 EWR    dep_delay        15.4
## 3 LGA    dep_delay        10.8
## 4 EWR    arr_delay         7.33
## 5 JFK    arr_delay         5.86
## 6 LGA    arr_delay         0.582
```



Conclusion

- With an ANOVA table and Tukey test we determined there was a significant difference in in both arrival and departure delays for all three airports at 95% confidence.
- It was noted that EWR had the highest departure and arrival delay.

Distribution of Flight Times

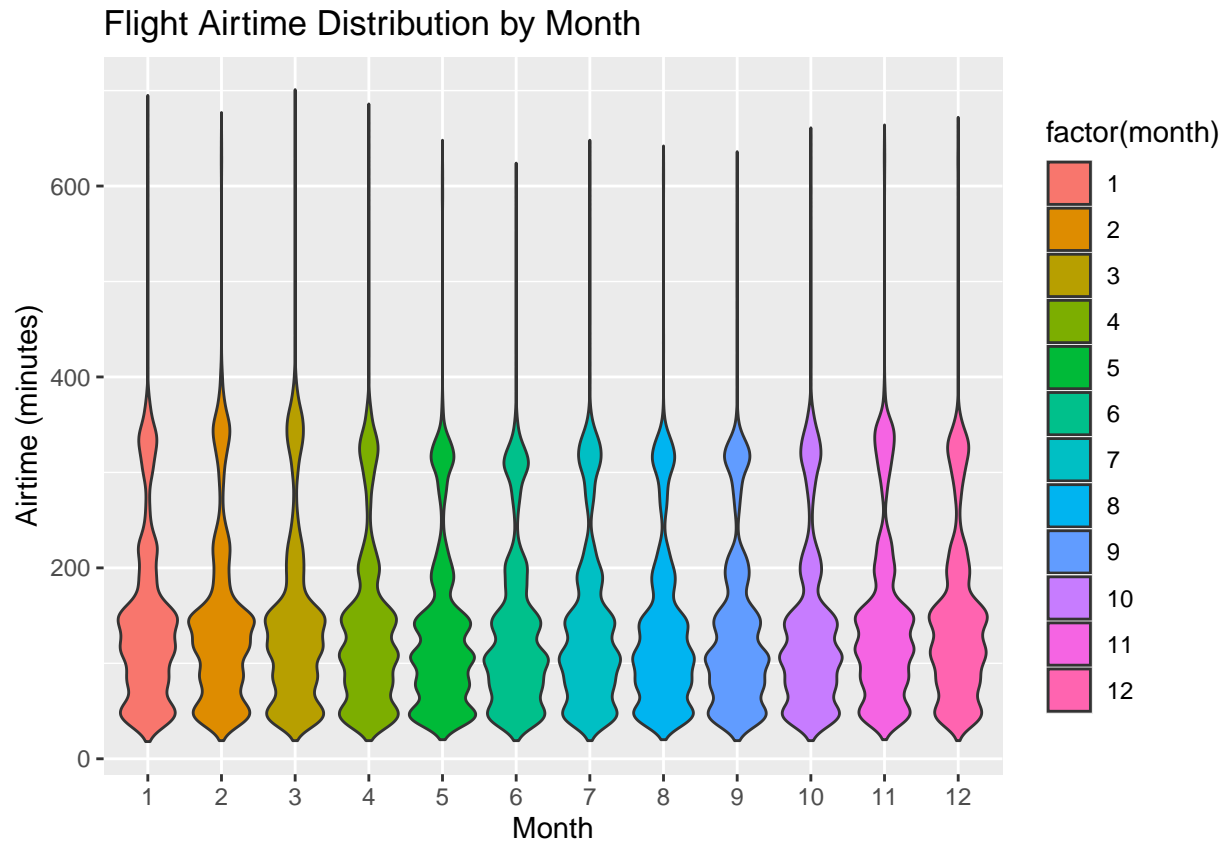
Setup

ANOVA Table

```
##           Df    Sum Sq Mean Sq F value Pr(>F)
## month      11 8.450e+06  768193   96.85 <2e-16 ***
## Residuals 422806 3.354e+09    7932
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tukey Test

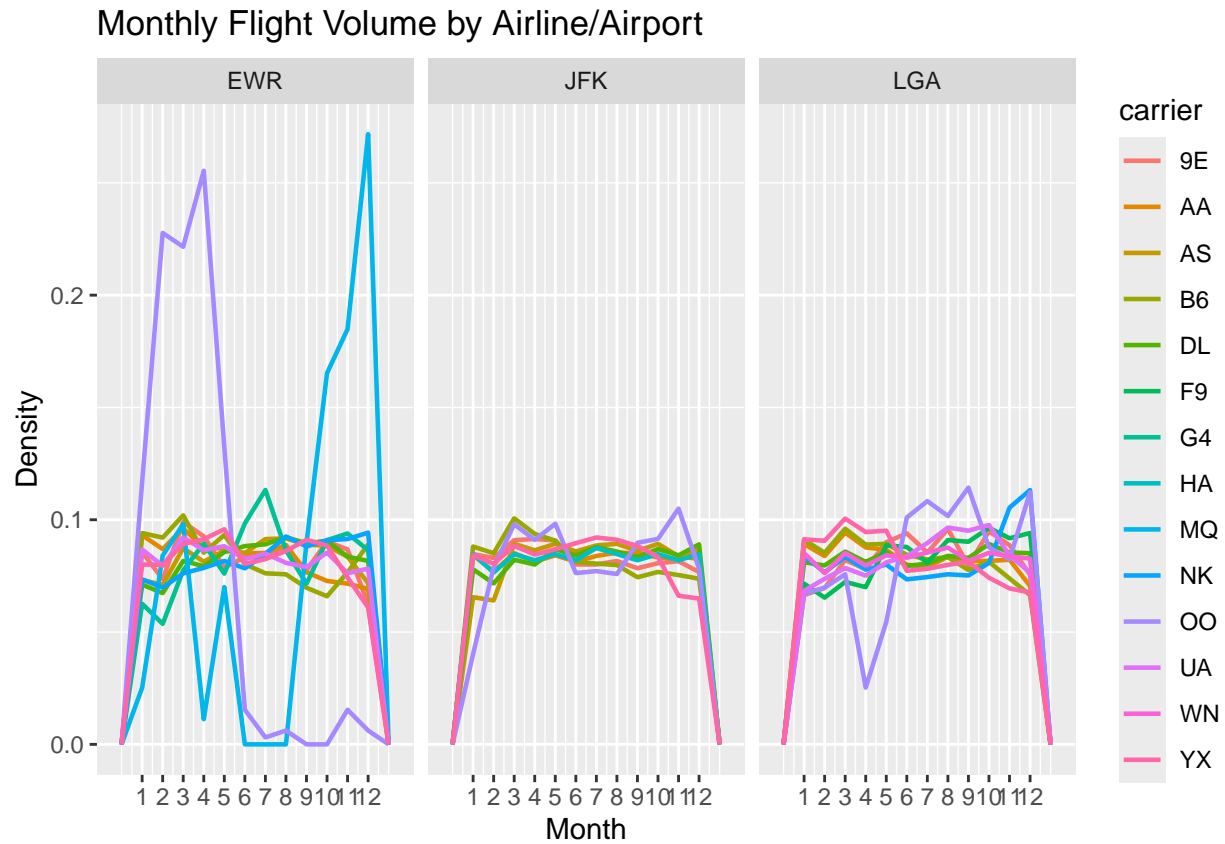
Distribution Visualization



Conclusion

- After conducting a Tukey test with a ANOVA table and determining the ranking, we noticed there was no significant difference between the months.
- It was noted there was higher airtime in month 12, possibly due to the holiday season, but not vastly different from any other month.

Monthly Flight Volume by Airline/Airport



Conclusion

- By looking at this graph, we can see that JFK was relatively normalized compared to the other airports and had no significant spikes or dips.
- The other airports had spikes at unexpected times, notably at month 8 for the LGA airport.
- We are unsure what caused these spikes; some possibilities include the school season or holidays.

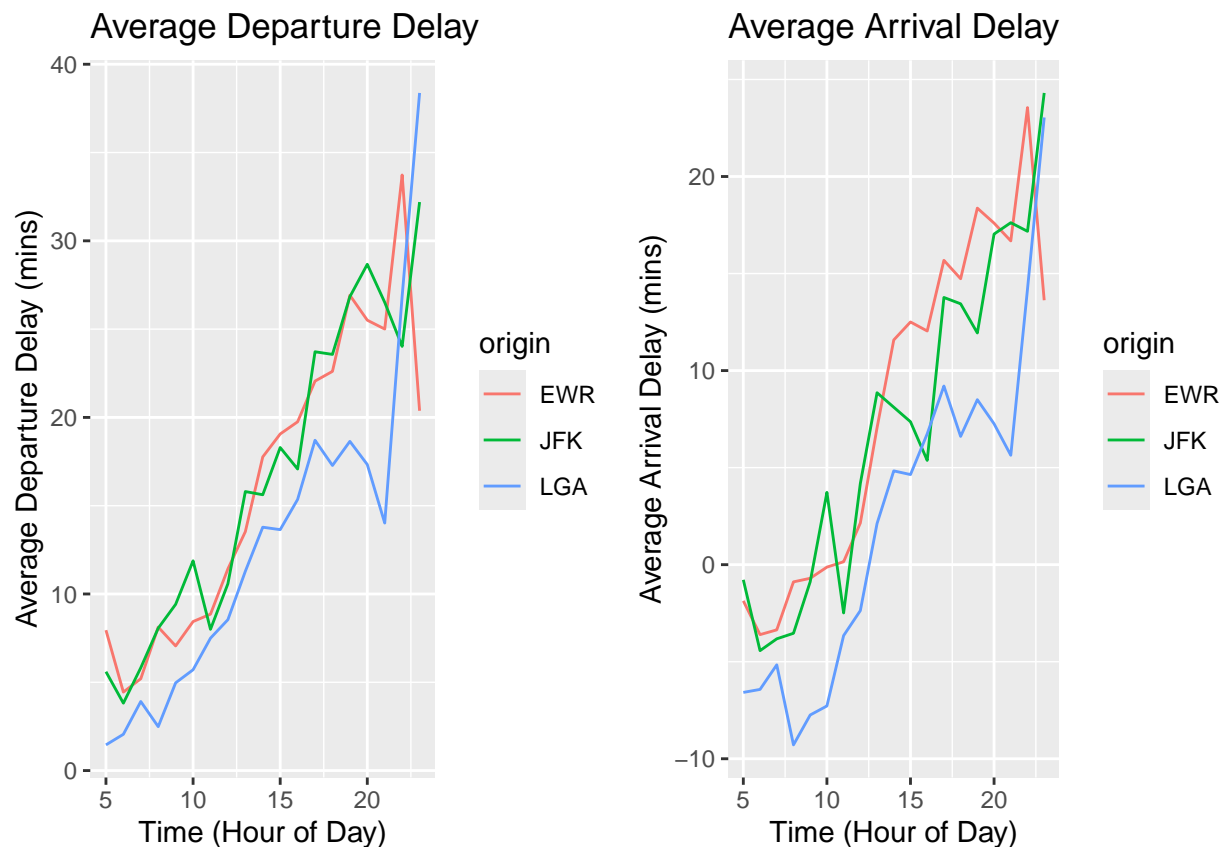
Data Exploration and Visualization

Question 1: What pattern(s) can be observed regarding flight punctuality in NYC?

- How do arrival and departure times vary throughout the day across NYC airports?
- When during the day are flights most likely to leave and arrive on schedule/late?
- What is the average delay time by airline?
- Are weekend flights more punctual than weekday ones?
- Are there certain months more punctual with flight departures/arrivals than others?

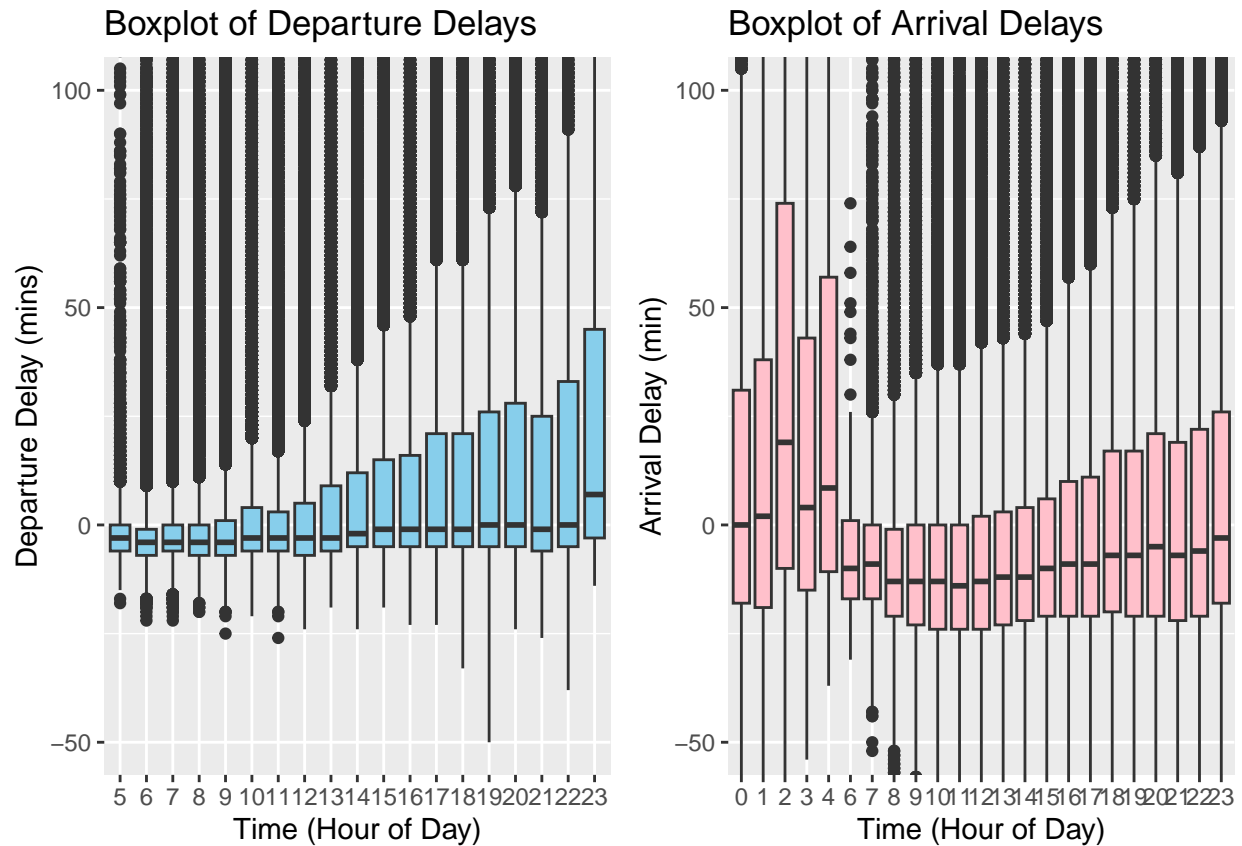
Setup

How do arrival and departure times vary throughout the day across NYC airports?



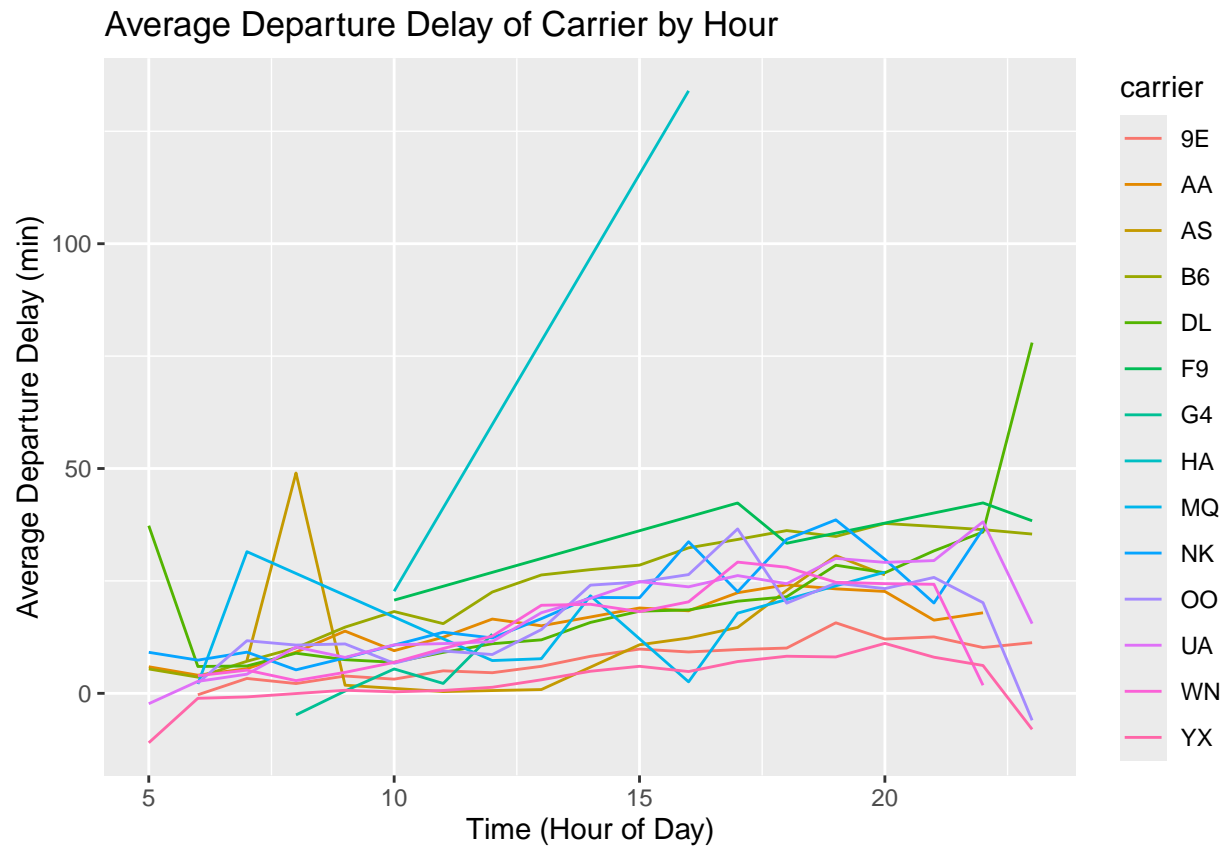
- Delays increase as the day progresses across all three airports (EWR, JFK, LGA).
- LGA (blue) has the lowest average departure delays most of the day.
- JFK (red) and EWR (green) show higher delays especially after 15:00 (hour 15).
- Arrival delays rise throughout the day, peaking after 18:00.
- JFK consistently has higher arrival delays in the evening.
- LGA again tends to have the lowest average arrival delays.
- Overall, flights are more delayed later in the day, with JFK generally experiencing the worst delays.

When during the day are flights most likely to leave and arrive on schedule/late?



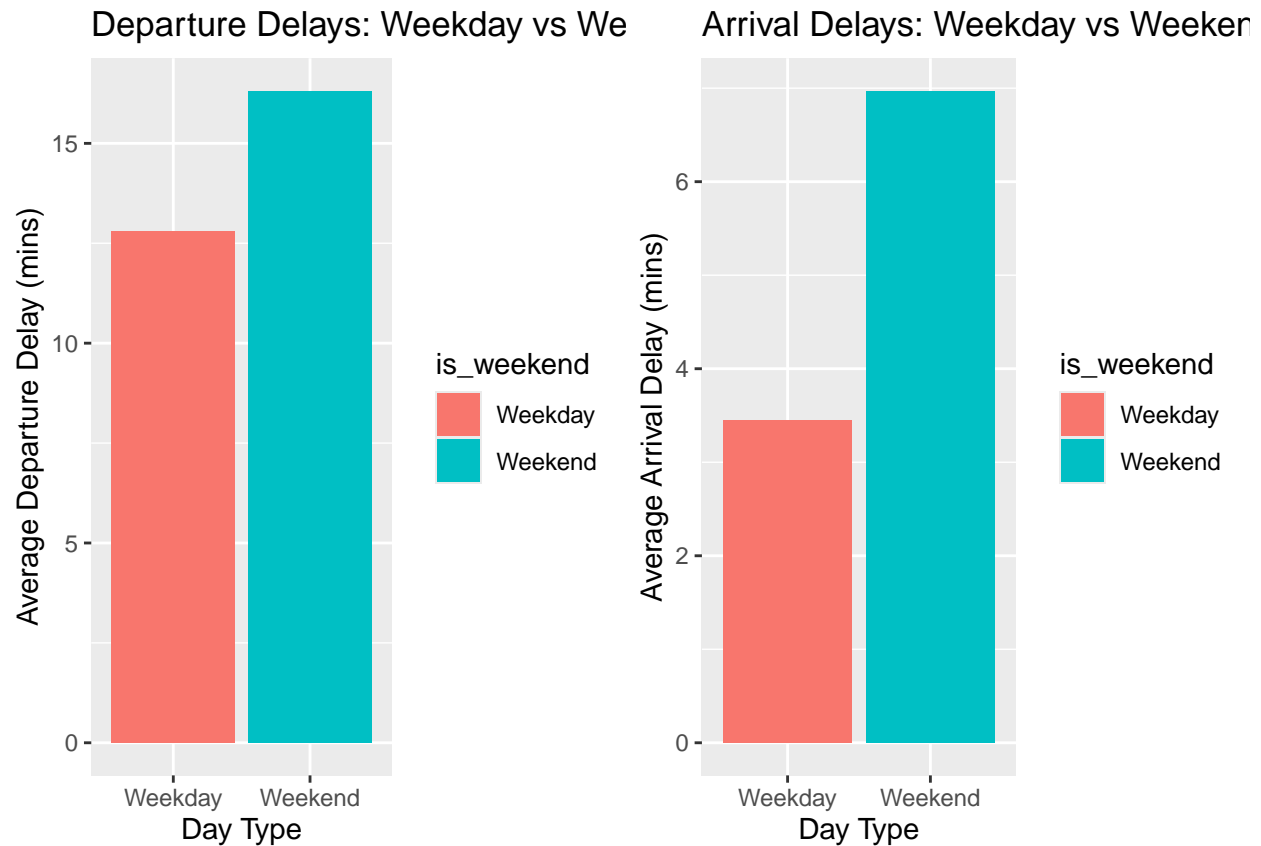
- Flights departing in the early morning hours (5am - 8am) have the lowest median delays and the least variability, making them the most likely to leave on time.
- Departure delays increase steadily throughout the day, with the worst delays occurring in the late evening between 8pm and 11pm
- Flights arriving between 6am and 11am show the lowest median delays and smallest interquartile ranges, indicating greater punctuality during these hours.
- Flights arriving in the early morning hours (1am-3am) and late evening tend to experience the highest and most variable delays.
- Overall, the most reliable time to fly in terms of both departure and arrival punctuality is early in the day, while delays become more frequent and severe as the day advances.

What is the average delay time by airline?



- G4 (Allegiant Air) stands out with extremely high average delays, especially from 10 AM to 5 PM, peaking around 135 minutes, which is significantly higher than other carriers. This suggests frequent or severe operational disruptions during this time block. It's a major outlier in delay performance and likely drags down overall punctuality averages.
- F9 (Frontier) and DL (Delta) also show above-average and increasing delays throughout the day, with spikes after 6 PM, hinting at cumulative delay effects.
- 9E (Endeavor) and WN (Southwest) consistently maintain low average departure delays, often staying below 15 minutes, suggesting better schedule adherence or buffer management.
- Most other airlines (e.g., AA, AS, B6, UA) cluster in the 10-30 minute delay range, with a mild upward trend throughout the day.
- Hour of Day effect: Across most airlines, delays increase slightly as the day progresses, but the degree of increase varies widely by carrier.

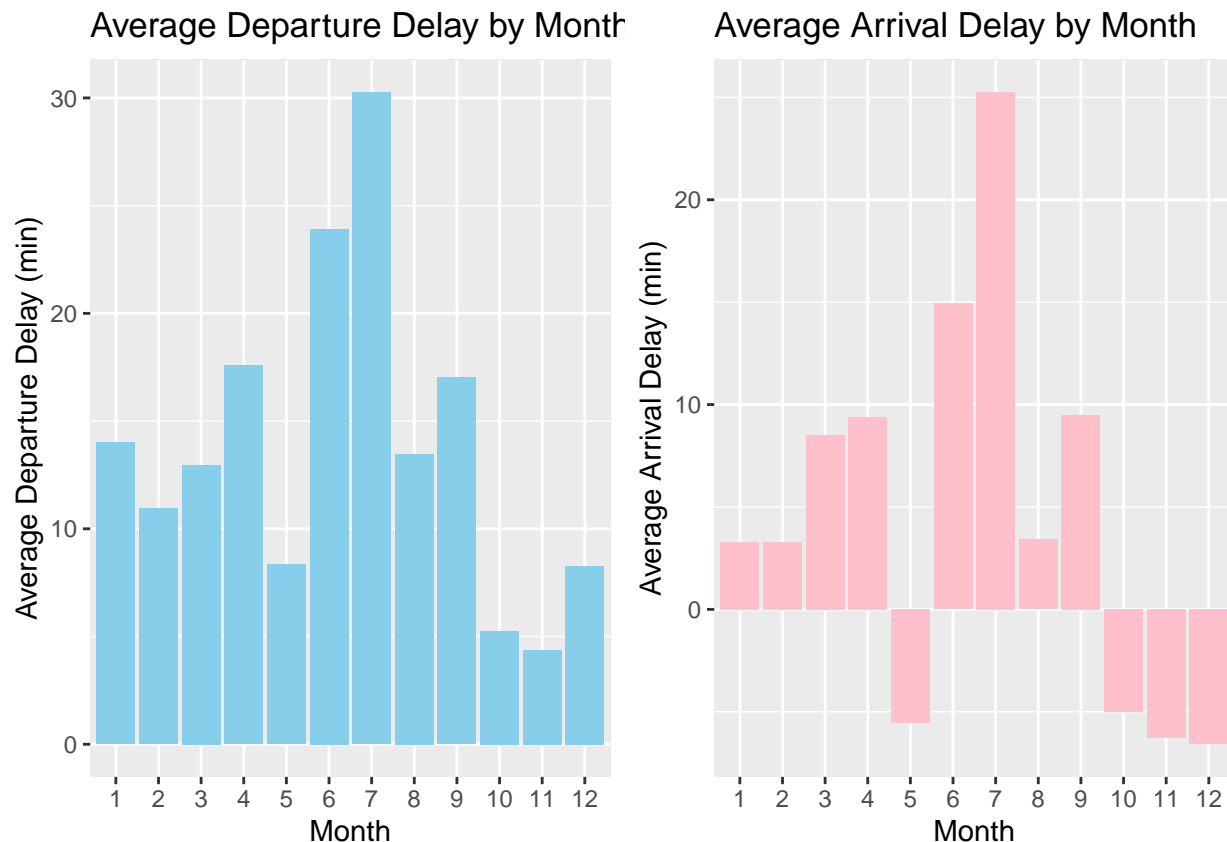
Are weekend flights more punctual than weekday ones?



```
## # A tibble: 2 x 3
##   is_weekend avg_dep_delay avg_arr_delay
##   <chr>      <dbl>      <dbl>
## 1 Weekday      12.8        3.44
## 2 Weekend      16.3        6.97
```

- Weekend flights show longer average delays than weekday ones.
- Departure delays: 12.8 (weekend) vs 16.31 min (weekday)
- Arrival delays: 3.44 min (weekend) vs 6.97 min (weekday)
- Average departure delays (14.55 mins) are greater in time than average arrival delays (5.21 mins)

Are there certain months more punctual with flight departures/arrivals than others?



- July has the highest delays for both departures and arrivals.
- October and November show the lowest departure delays.
- May, October, November and December report negative arrival delays.
- Summer months (June & July) are the worst for delays, while fall and winter months are more reliable.

Conclusion

The analyses of the questions above reveal clear patterns in flight punctuality across NYC airports. Flights are most likely to depart and arrive on time during early morning hours, while delays increase steadily throughout the day. LaGuardia generally experiences fewer delays than JFK and Newark. Some airlines consistently perform better than others, with Southwest (WN) and Endeavor (9E) typically showing lower average delays compared to carriers like Allegiant (G4) and Frontier (F9), which experience notably higher delays. Departure delays are greater than arrival delays. Weekend flights tend to be less punctual than weekday flights, and seasonal trends show greater delays in the summer, especially July, with more reliable performance in the fall.

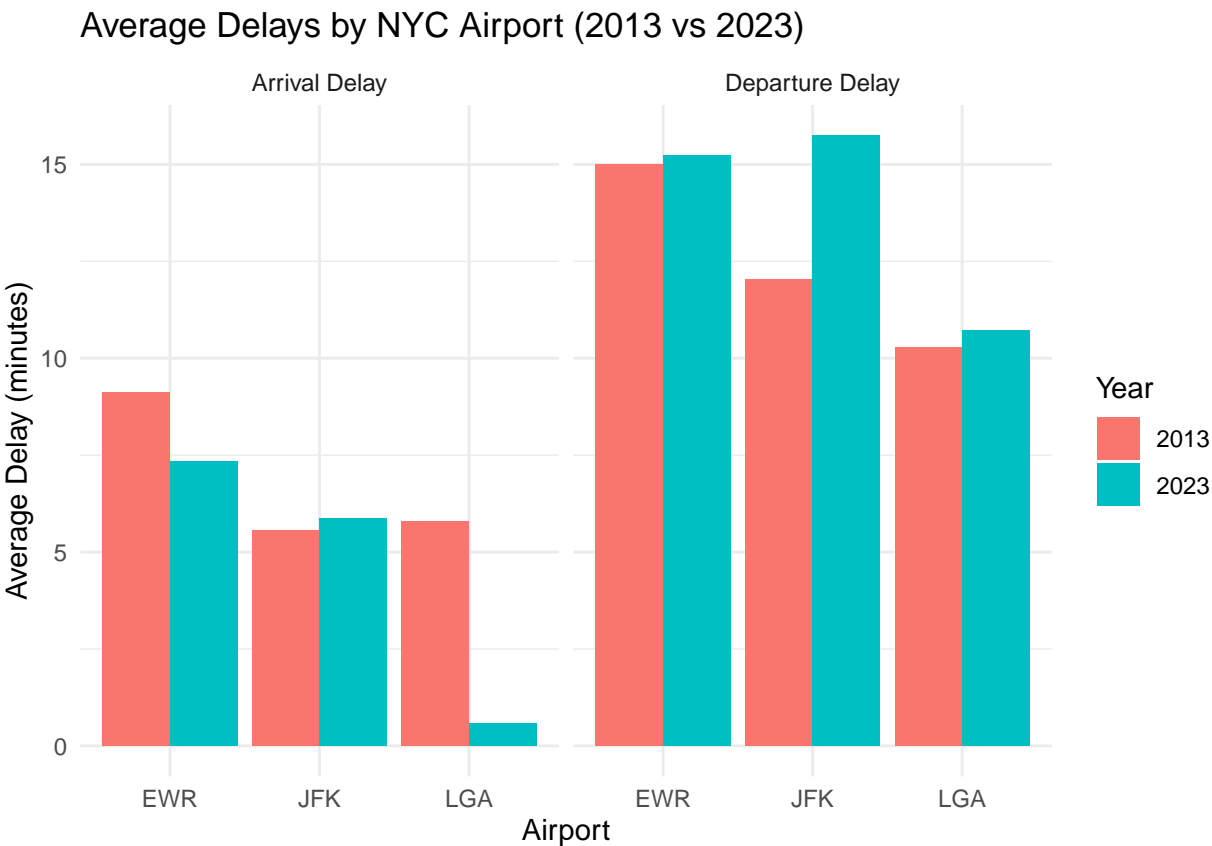
Question 2: Which of NYC’s airports were most operationally efficient in 2013, and how has that changed to 2023? Which NYC airports have become more or less operationally efficient over the past decade?

The following analysis compares the operational efficiency of New York City’s three major airports — JFK, LGA, and EWR — in the years 2013 and 2023. We define operational efficiency using: average arrival delay, average departure delay, and flight volume. We use data from the nycflights13 and nycflights23 packages to explore trends and assess changes over time.

Data Preparation

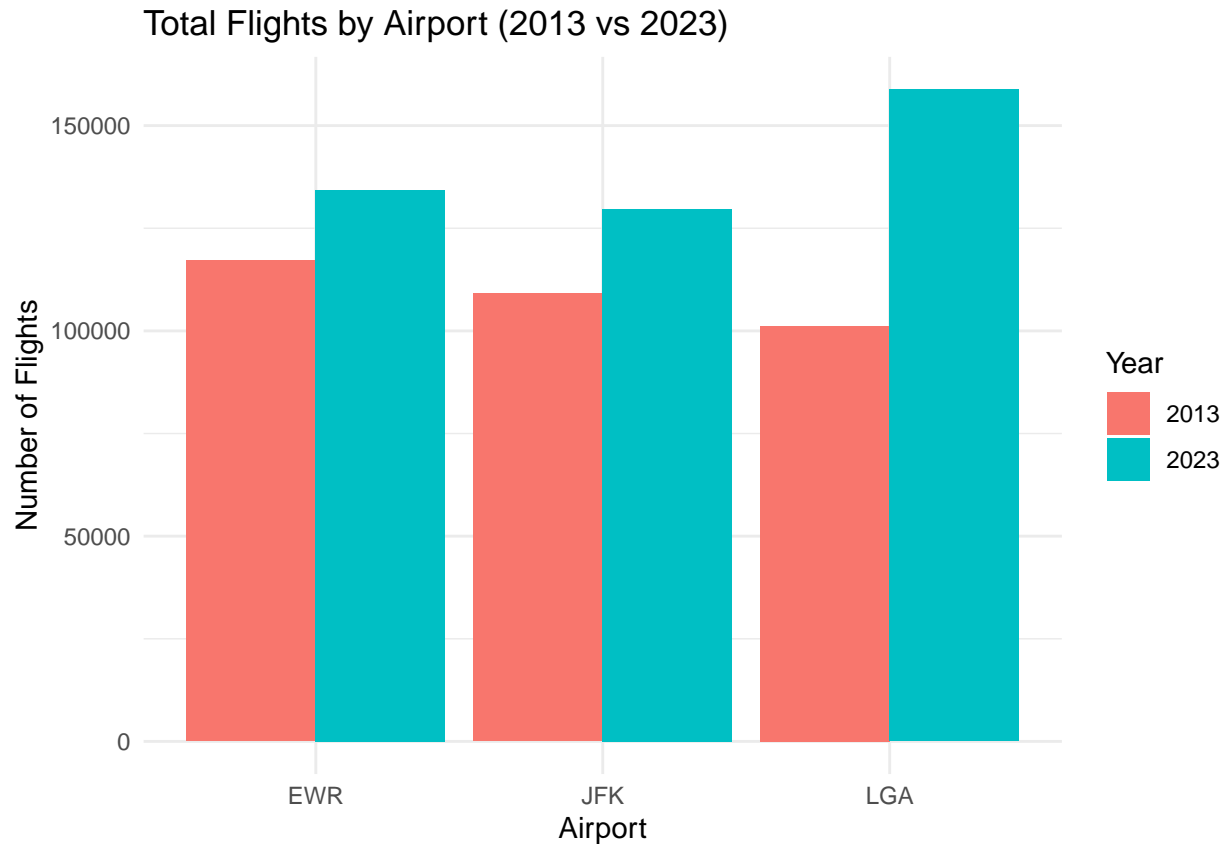
```
## # A tibble: 6 x 5
##   origin avg_arr_delay avg_dep_delay num_flights year
##   <chr>      <dbl>      <dbl>      <int> <dbl>
## 1 EWR          9.11         15.0    117127 2013
## 2 JFK          5.55         12.0    109079 2013
## 3 LGA          5.78         10.3    101140 2013
## 4 EWR          7.33         15.2    134398 2023
## 5 JFK          5.86         15.7    129620 2023
## 6 LGA          0.582        10.7     158800 2023
```

Delay Comparison



- LGA shows the greatest improvement in both arrival and departure delays, suggesting increased operational efficiency.
- JFK had moderate delays in both years, but 2023 delays increased slightly.
- EWR consistently had the highest average delays, indicating long-standing operational challenges.

Flight Volume



- Flight volume has remained high at all airports, though JFK and LGA saw small decreases, possibly reflecting post-COVID changes.
- Efficiency improvements at LGA are especially notable given its sustained flight volume.

Linear Model: Delays Over Time

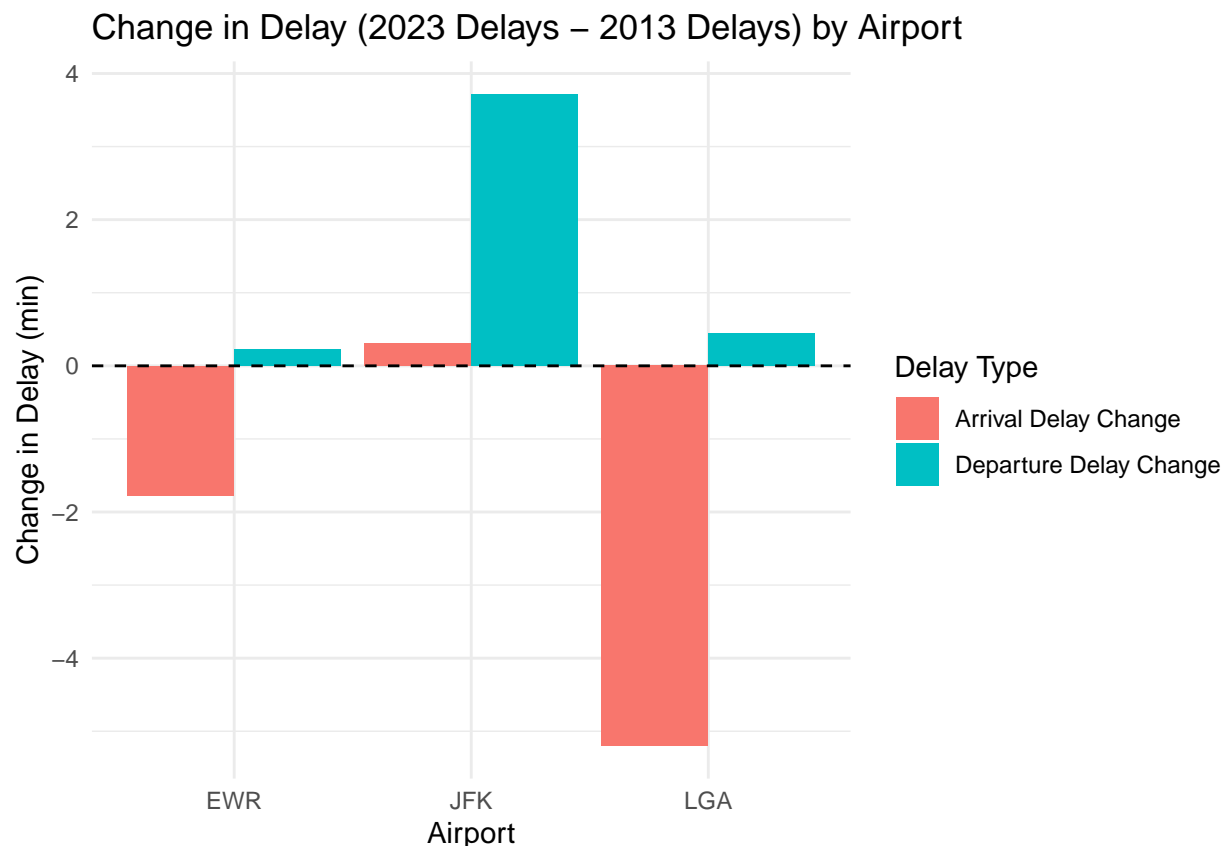
```
##
## Call:
## lm(formula = arr_delay ~ origin + year, data = combined_flights)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -101.69  -24.94  -12.98    5.89  1804.89
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  463.65750   24.70990   18.76  <2e-16 ***
```



```
## originJFK    -2.42171    0.14990   -16.16   <2e-16 ***
## originLGA    -5.38005    0.14702   -36.59   <2e-16 ***
## year         -0.22568    0.01224   -18.43   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 52.46 on 750160 degrees of freedom
## Multiple R-squared:  0.002368,    Adjusted R-squared:  0.002364
## F-statistic: 593.6 on 3 and 750160 DF,  p-value: < 2.2e-16
```

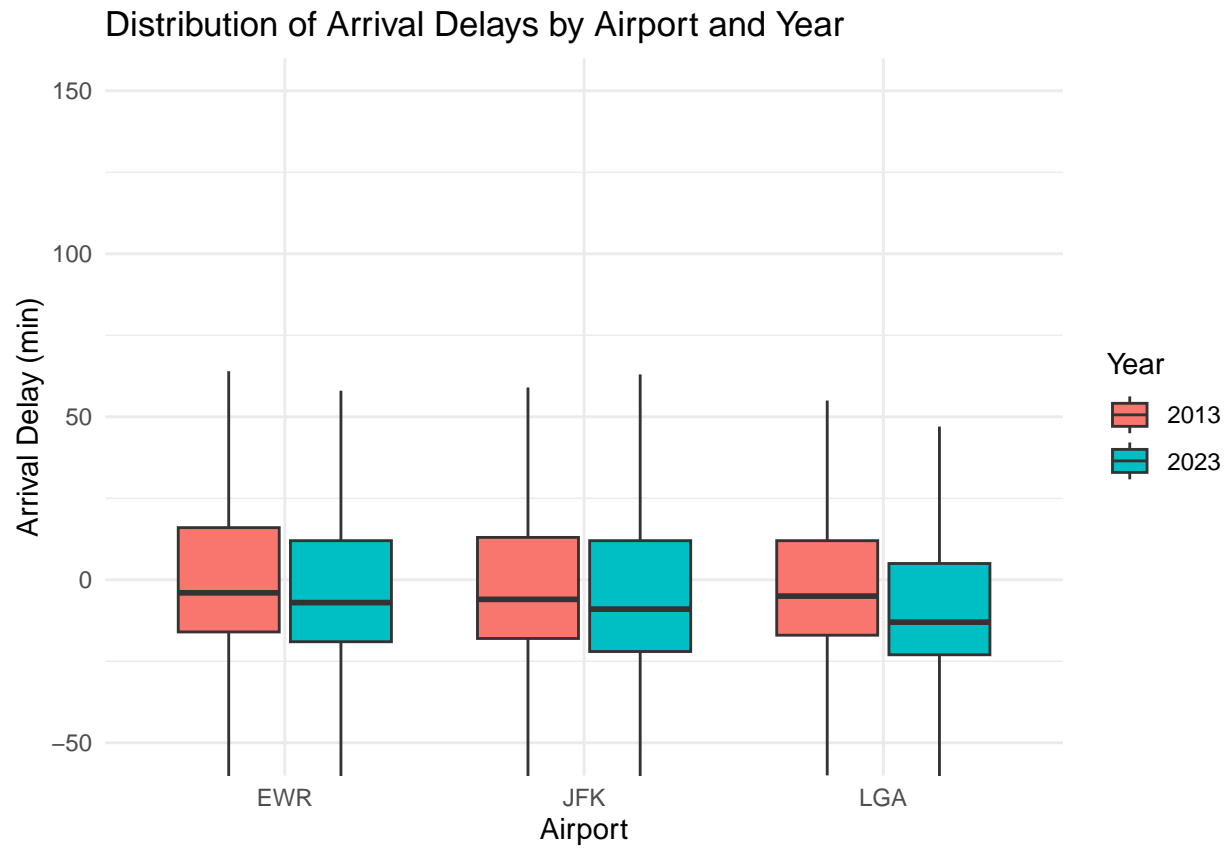
- Flights from JFK had, on average, 2.42 less minutes of arrival delay than EWR, while LGA had 5.38 less minutes of arrival delay than EWR on average.
- An adjusted R-squared value of 0.0024 suggests very low explanatory power as only 0.24% of the variance in arrival delays can be explained. A residual standard error of 52.46 indicates high variability in flight delays that aren't captured by airports.
- The model does suggest that arrival delays have decreased slightly over the decade, though not enough to warrant significance.

Average Arrival/Departure Delay Comparisons by Airport



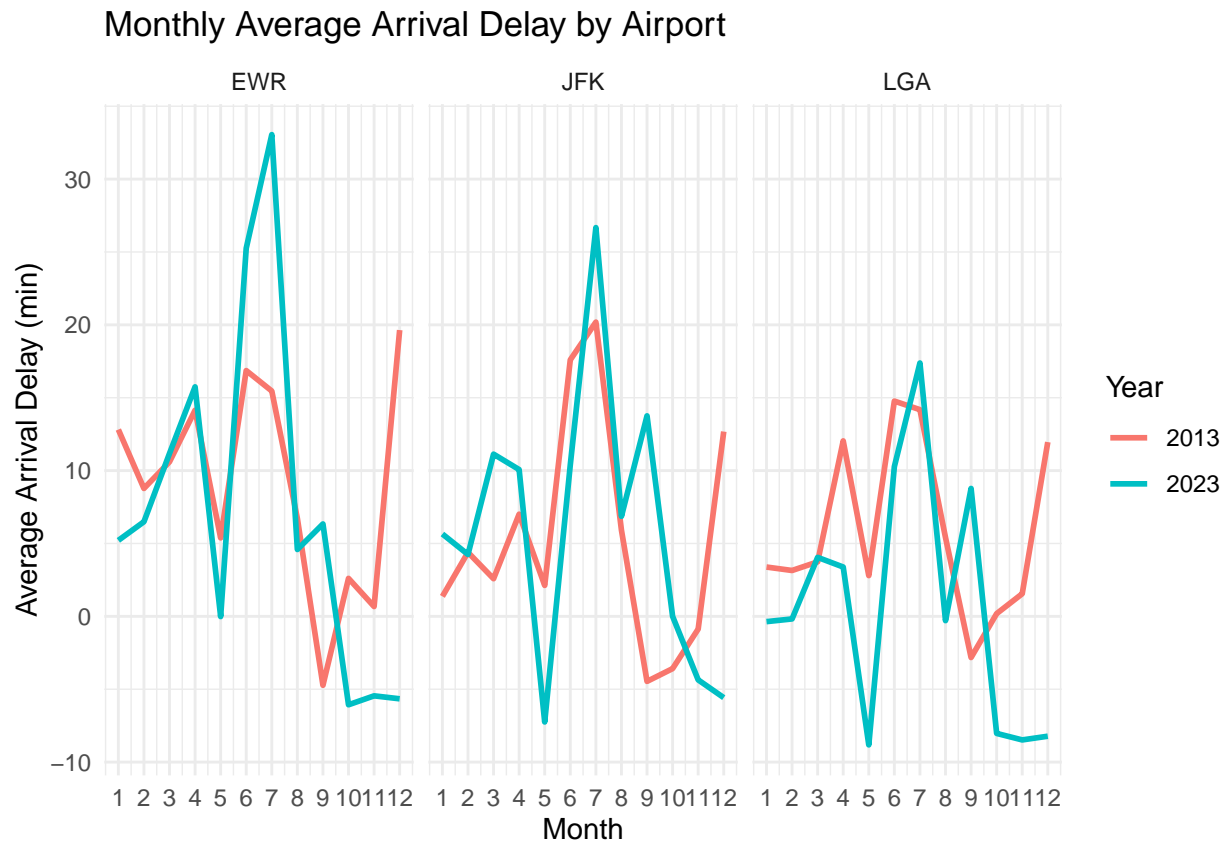
The graph shows a decrease in average arrival delay between 2013 and 2023 and an increase in arrival departure delay between 2013 and 2023. In particular, the arrival delay decrease in LGA and the departure delay increase in JFK is noteworthy.

Boxplot of Arrival Delay Trends



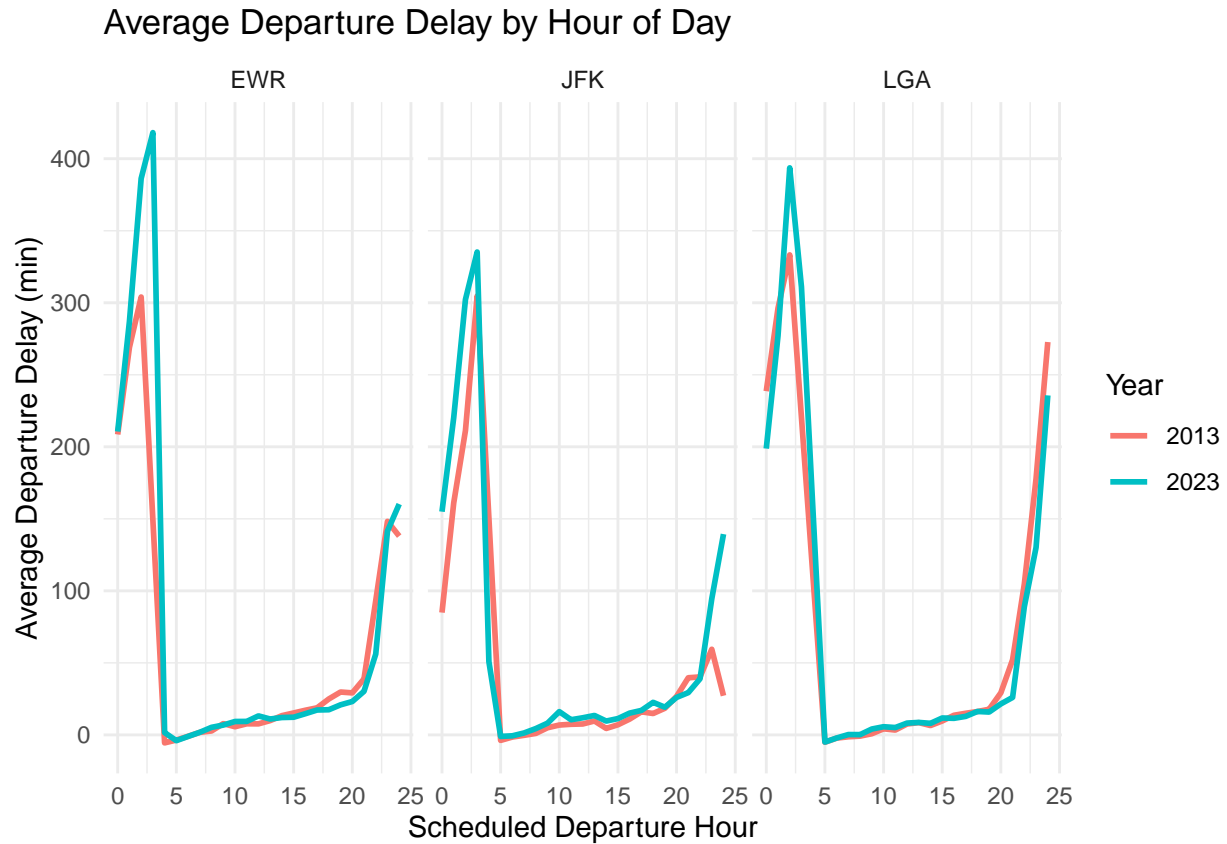
Boxplotting the arrival delay trends will give you a slight but consistent decrease throughout all three major New York airports. Specifically, the median, 1st quantile, and 3rd quantile values are all smaller for 2023 than for 2013 across the six different boxplots.

Line Graph of Arrival Delay Trends By Month



We see that the decrease in average arrival delay in the month of December and May is noteworthy across airports. There is a slight increase in average arrival delay in the month of September.

Line Graph of Departure Delay Trends by Hour



If we split the average departure delay by hour, we can see that the results are strikingly similar from 2013 to 2023. There is a notable increase in average delay at 12:00 AM in LGA and 1 AM in EWR, though the sample size of departed flights in those hours is to be questioned.

Conclusion

In conclusion, there does seem to be a slight increase in departure delay coinciding with a slight decrease in arrival delay across all three airports. Some results, like arrival delay in LGA or departure delay in JFK, were notable, but there isn't really a consistent trend that suggests that airports have improved their efficiency regarding delays after 10 years.

Question 3: How does weather impact flight punctuality?

With the following analysis, I will be investigating and answering the following questions:

- How does weather affect departure and arrival times?
- Is there a seasonal trend in weather and flight delays?
- Are some airports more resilient to weather-related delays than others?
- How does weather predict flight delays?

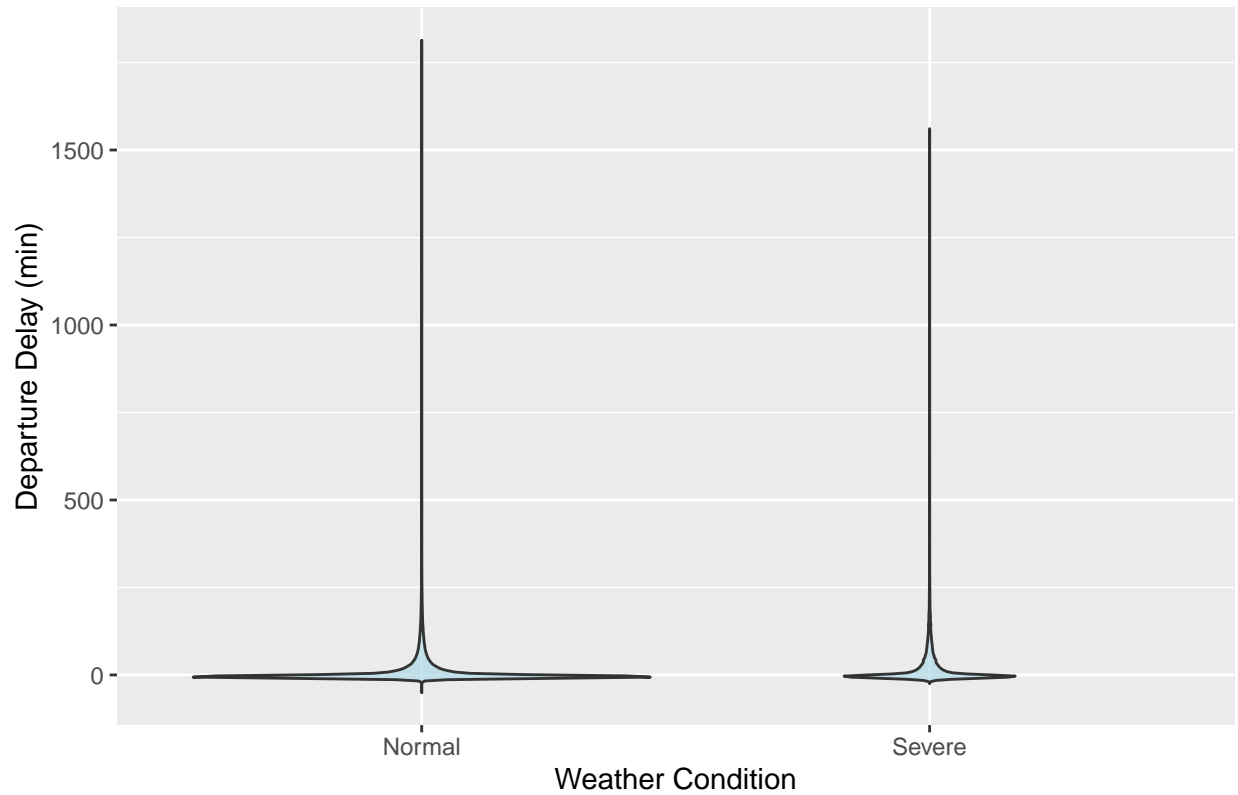
Hypotheses:

- Weather conditions impact the departure and arrival times of flights, with more severe weather (wind, rain, etc.) causing greater delays.
- There is a monthly trend in delays, correlating with seasonal weather.
- Certain airports are more resilient to delays in severe weather as they have less weather variability.
- Different weather conditions/variables (wind, rain, etc.) can be used to predict flight delays.

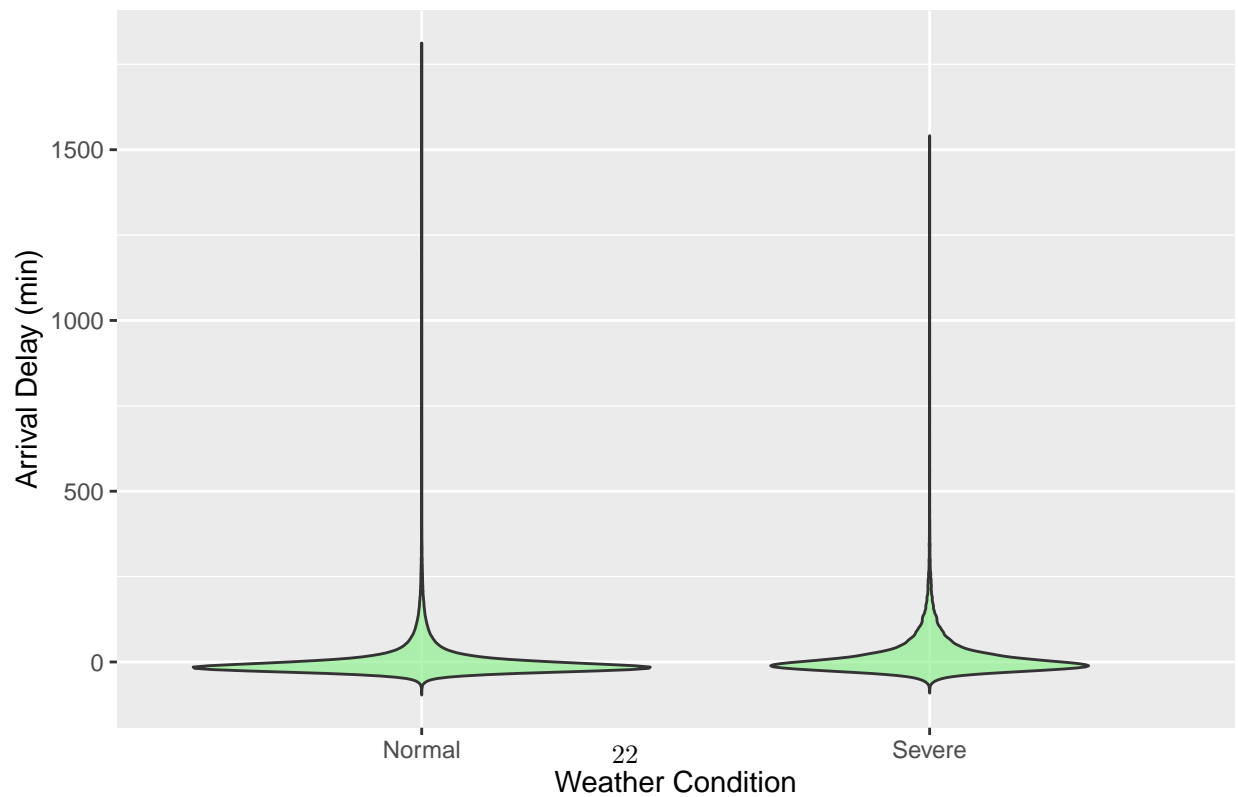
Setup

How does weather affect departure and arrival times?

Effect of Severe Weather on Departure Delays



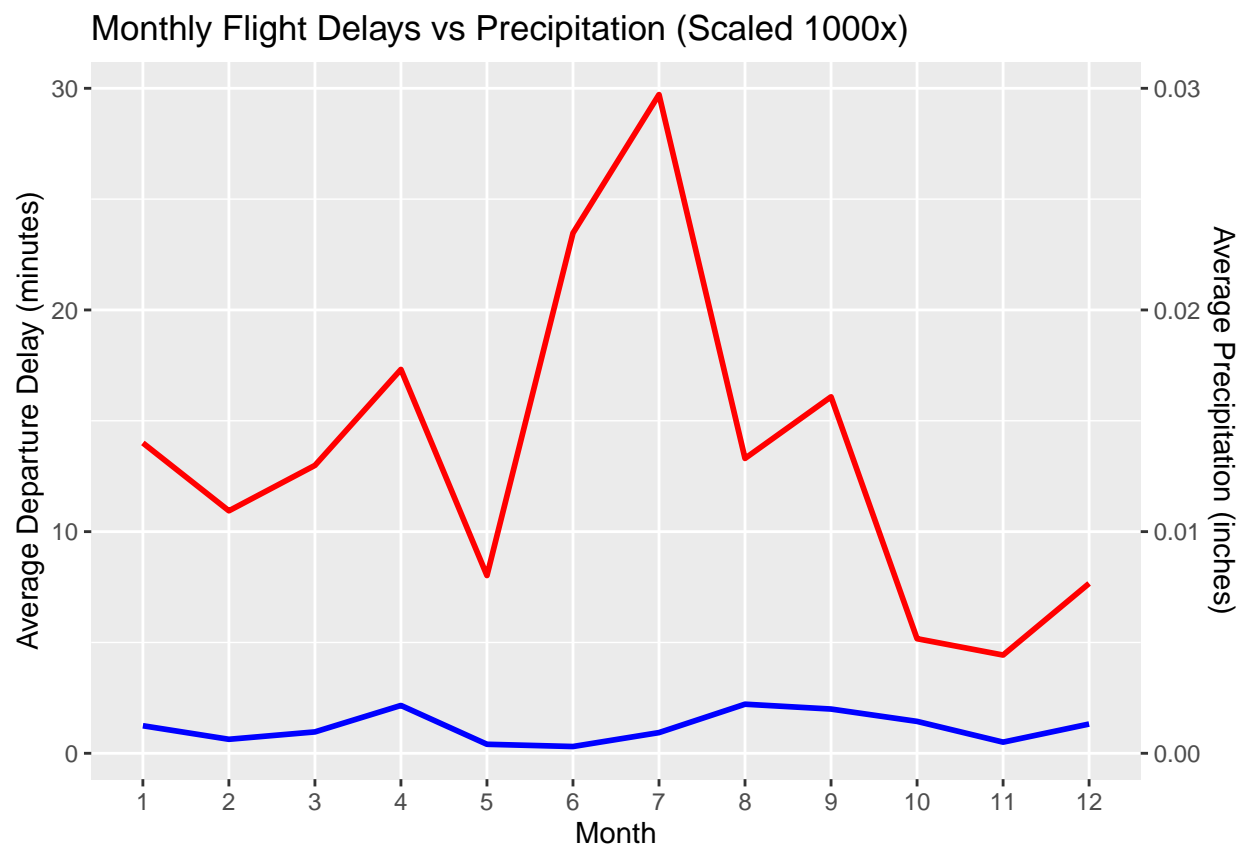
Effect of Severe Weather on Arrival Delays



Effect of Severe Weather on Departure Delays: It appears as if there is not a significant difference between departure delays in normal versus severe weather conditions. Perhaps the severe parameters can be altered to present a more pronounced difference. There is a larger density of flights taking off during normal conditions, as to be expected. The variability of normal weather condition departure delays is larger than severe's, with its maximum being greater than severe's. This is surprising, as one would expect the severe weather condition's plot to contain the maximum departure delay. This means that the delays can perhaps be attributed to other variables.

Effect of Severe Weather on Arrival Delays: Similarly to the Departure Delays, the Arrival Delays do not present much of a difference between the two categories. Once again, the normal weather condition's density is larger as well as its variance. The two plots (Departure vs Arrival) are very similar, with the only notable difference being more variability in the Arrival Delay plot.

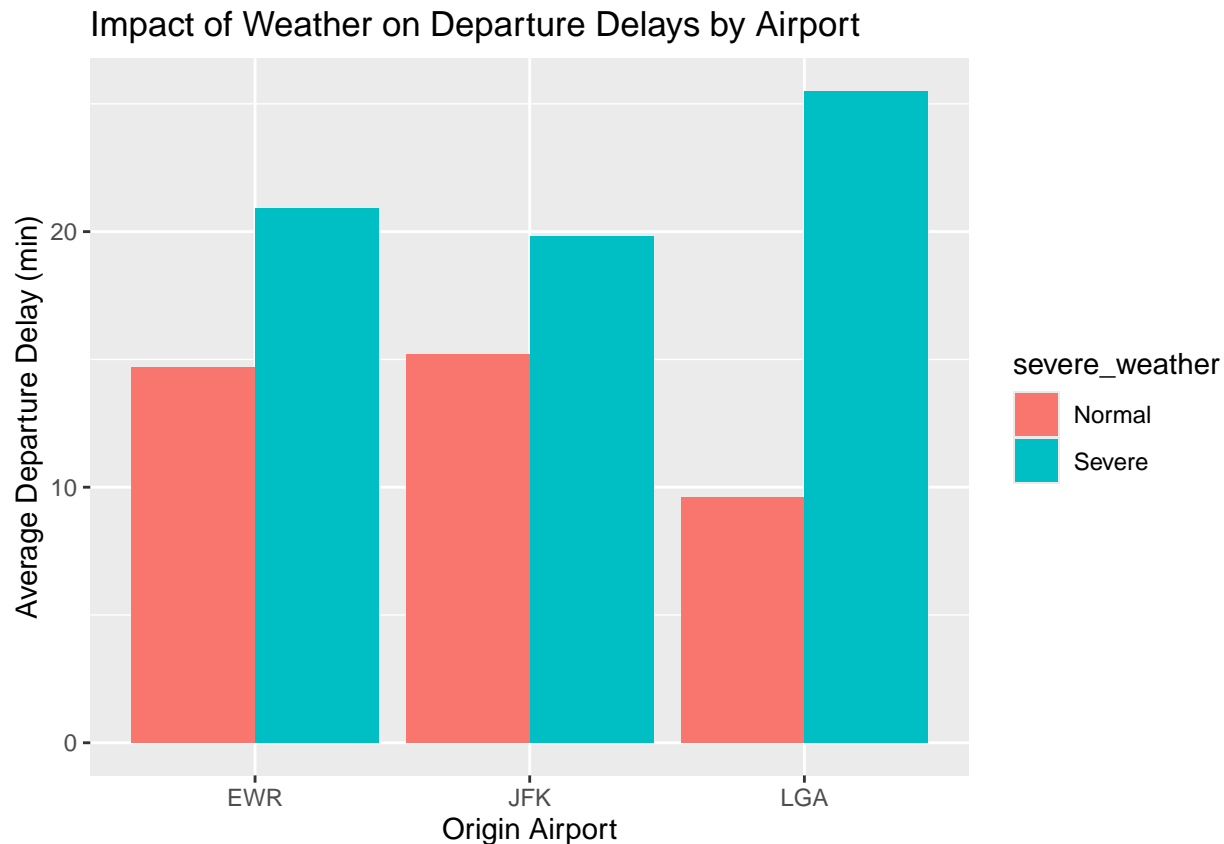
Is there a monthly trend in weather and flight delays?



Important Note: I scaled precipitation by 1000 for better visualization; this visually exaggerates precipitation's role in delays. My intention is to emphasize the mirrored trends of the two variables.

The average departure delays' trend mirrors the average precipitation per month. They both experience a large spike in June, however, July's average precipitation takes a large dip while the average delay increases to its max. I think it's important to consider other factors, such as increased travel volume during summer vacation being responsible for delays, rather than weather. The fall/winter months experience both a decrease in rain and a decrease in delays, both spiking in December then drastically declining in January. It seems as if there is some relationship between the two variables, based on the graph. *Once again, it's important to take note if the inflated precipitation graph as it is scaled by 1000 for better visualization.

Are some airports more resilient to delays in severe weather than others?



It appears that, of the three departure airports in NYC: EWR, JFK, and LGA, LGA is the most resilient to delays in severe weather as well as departure delays in general. Overall, EWR has the worst departure delays out of all three of the airports, both in severe weather and normal. Something to consider regarding each airports' delay times would be whether other factors have a bigger impact on these delays rather than weather. Perhaps EWR is the busiest airport in general, pricing has an impacts, etc.; correlation doesn't necessarily equal causation.

EWR: - In normal weather, EWR has an average departure delay of approximately 14 minutes. - In severe weather, it has an average departure delay of approximately 24 minutes.

JFK: - In normal weather, JFK has an average departure delay of approximately 12 minutes. - In severe weather, it has an average departure delay of approximately 19 minutes.

LGA: - In normal weather, LGA has an average departure delay of approximately 9 minutes. - In severe weather, it has an average departure delay of approximately 17 minutes.

How does weather predict flight delays? (linear regression)

```
##
## Call:
## lm(formula = dep_delay ~ precip + wind_speed + visib, data = flights_weather)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -85.74  -19.37  -14.96   -3.96 1802.04
```



```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  22.06605     0.46178   47.78  <2e-16 ***
## precip      109.23993    10.19331   10.72  <2e-16 ***
## wind_speed    0.40775     0.01589   25.66  <2e-16 ***
## visib       -1.34483     0.04651  -28.92  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 53.83 on 405041 degrees of freedom
## Multiple R-squared:  0.004708, Adjusted R-squared:  0.004701
## F-statistic: 638.7 on 3 and 405041 DF, p-value: < 2.2e-16
```

P-Value/Statistical Significance: - All variables have a p-value of $< 2e-16$, approximately equal to 0, meaning that all variables: precipitation, wind speed, and visibility, are statistically significant. This means that all three variables are significant predictors of departure delays.

Coefficients: - Intercept (22.56777 min): When all three weather factors are equal to 0, the expected departure delay is approx. ~22.57 minutes. - Precipitation (86.51568 min per inch): A 1 inch increase in precipitation leads to an average departure delay increase of ~86.52 minutes. This is the best predictor of departure delays in this model. - Wind Speed (0.36662 min per mph): A 1 mph increase in wind speed leads to an average departure delay increase of ~0.37 minutes. - Visibility (-1.55374 min per mile): A 1 mile increase in visibility leads to an average departure delay decrease of ~1.55 minutes.

R-squared Values: - Multiple R-squared: 0.0153 - This means the model explains only 1.53% of the variance in departure delays. - This low R-squared value suggests that there are other factors/variables that have a greater impact on departure delays.

Conclusion

In conclusion, weather conditions do have an impact on flight punctuality, however, other factors may have a greater influence, as delays during normal weather show more variability. While there is a monthly trend in delays that correlates with seasonal (“bad”) weather, other external factors, such as increased travel during summer months, may play a bigger role. Certain airports, such as LGA, are more resilient to delays in severe weather compared to others like EWR, which consistently has the highest delays. Once again, other factors such as operational differences may contribute to these trends. While weather variables are statistically significant predictors of departure delays, with precipitation having the strongest impact, the linear regression model explains only a small portion of the variance, indicating that additional factors beyond weather significantly influence flight delays.

Question 4: Are certain geographic areas or types of carriers linked with better or worse flight arrival times and operational efficiency? What geographic or airport characteristics are linked with consistently superior on-time performance?

Research question objectives:

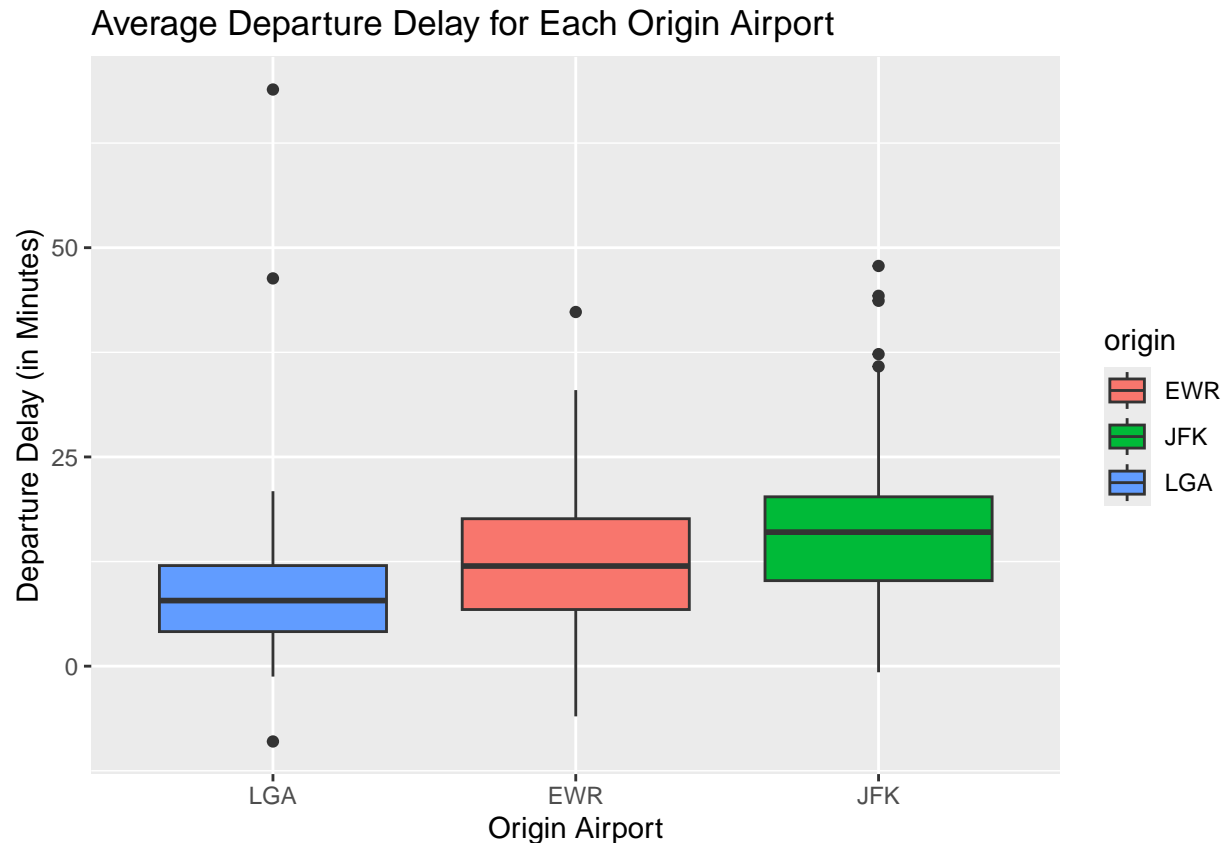
- Which of the three origin airports had the least average departure/arrival delays and which had the most?
- Which destination airport(s) had longer average departure/arrival delays and which had shorter?
- Do average departure/arrival delays vary a lot between carriers? If so, which airline(s) had longer delays and which had shorter?

Hypotheses:

- International airports will have longer departure delays than domestic airport due to busy air traffic.
- Airports on the West Coast will have longer arrival delays due to the higher unpredictability from longer flight distances.
- Arrival delays have little to no correlation with origin airports and departure delays have little to no correlation with destination airports.
- Budget airlines will have longer departure and arrival delays due to cost-cutting induced issues such as poor ground handling or flight scheduling.

Setup

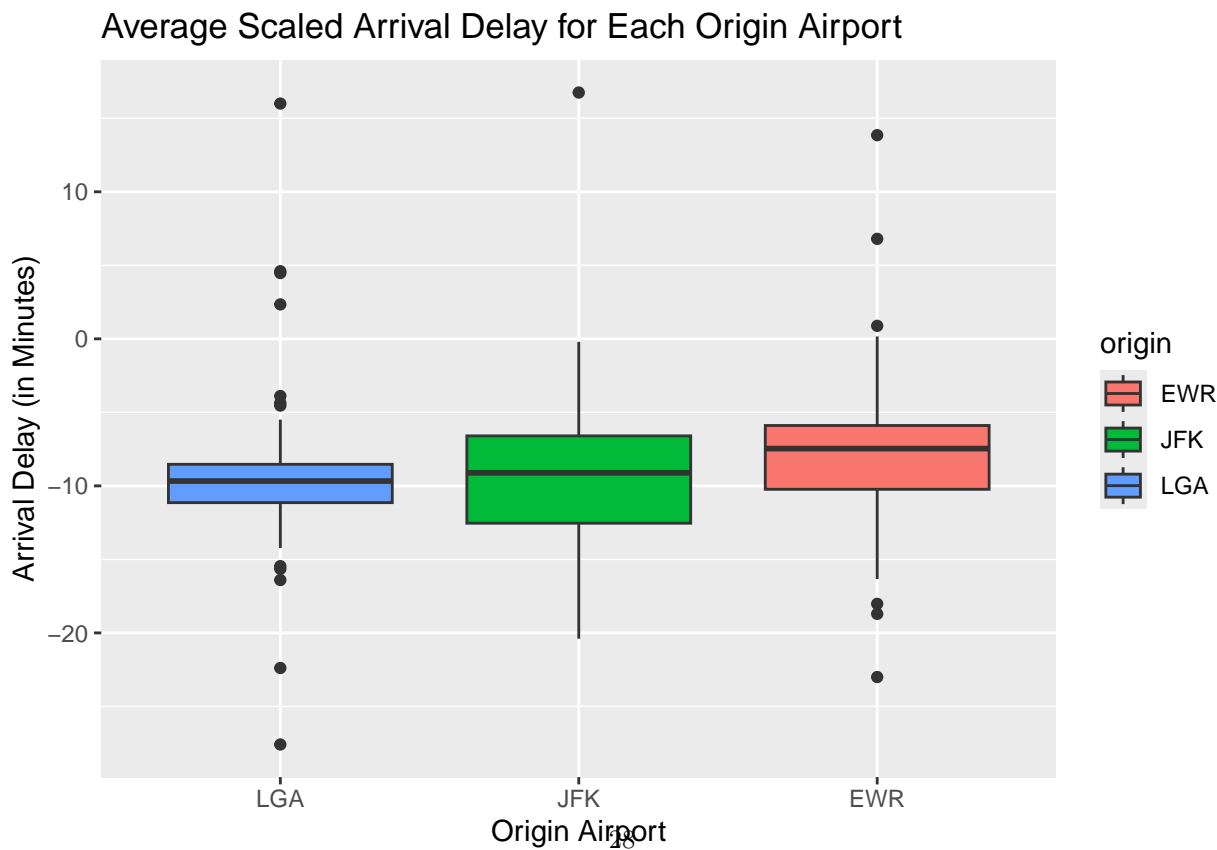
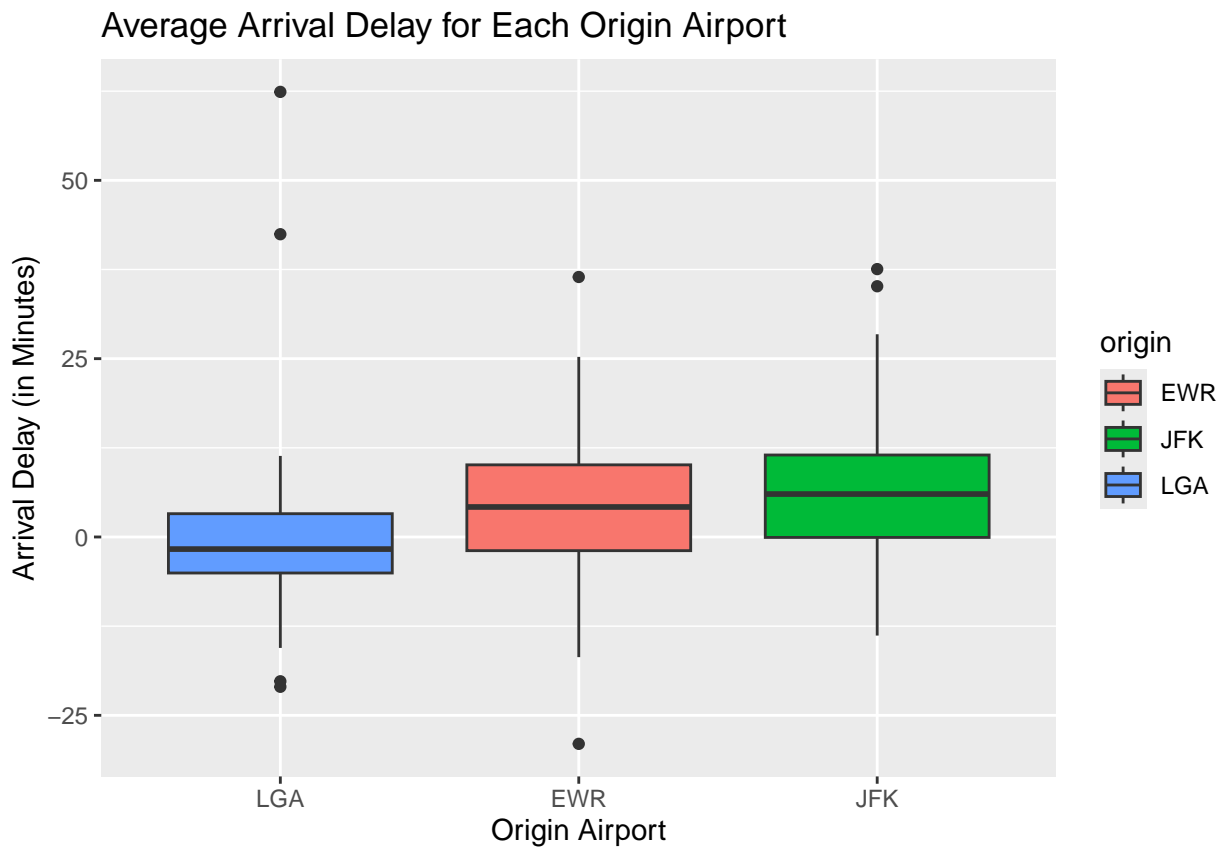
Which of the three origin airports had the least average departure delays and which had the most?



```
## # A tibble: 3 x 2
##   origin      n
##   <chr>   <int>
## 1 EWR    134398
## 2 JFK    129620
## 3 LGA    158800
```

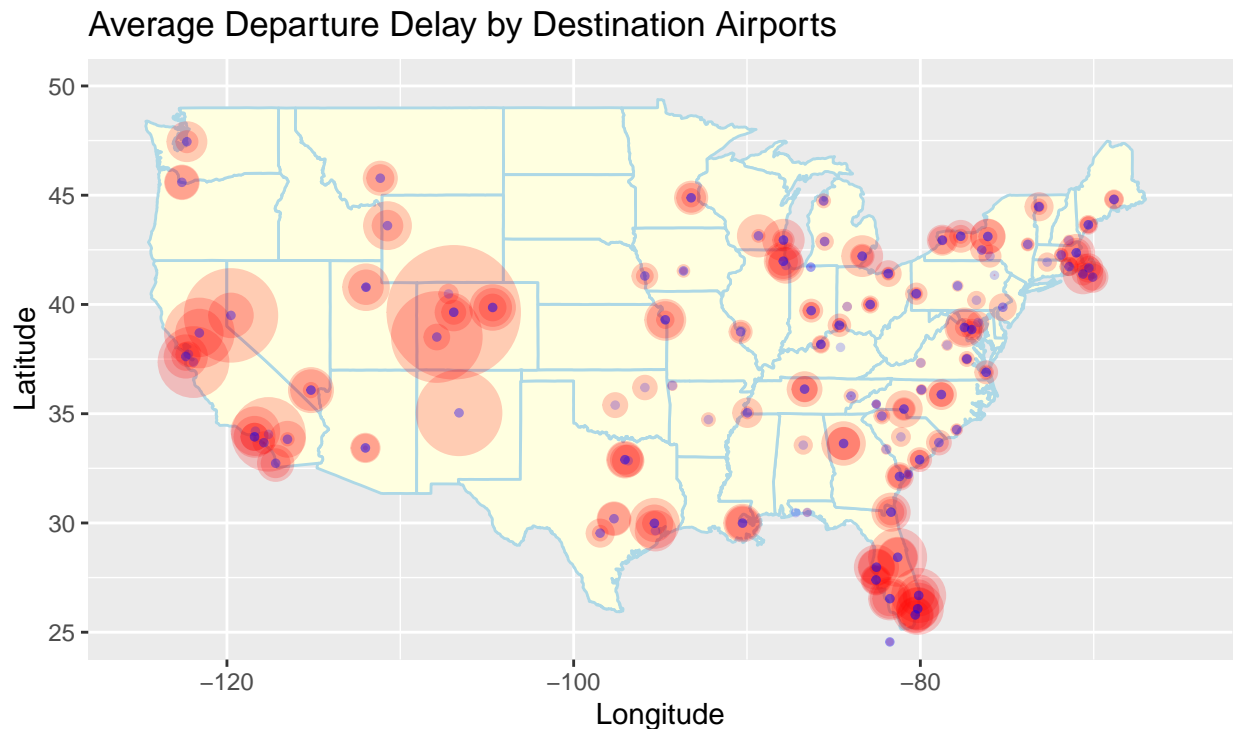
LGA (LaGuardia Airport) had the least average departure delays median, whereas JFK (John F. Kennedy International Airport) had the most. This could result from JFK being a busy airport with lots of international air traffic affecting airport efficiency and timeliness on domestic flights compared to LGA. The latter is a much smaller civil airport that flies mostly to domestic destinations with less or almost no interference from international flights. Another possible reason could be that since JFK is an international airport (whereas LGA is not), resulting in LGA's much larger number of domestic flights, which contributes to spreading out the average delay time. When looking at the number of domestic flights, LGA had the most, followed by EWR, then JFK, which is in the reverse order of average departure delays. However, since airport efficiency can be affected by many other different factors, a more complex, thorough investigation must be conducted to acquire statistically significant enough evidence to draw a strong conclusion about the cause for unsatisfactory flight punctuality.

Which of the three origin airports had the least average arrival delays and which had the most?



Same as departure delays, LGA had the least average arrival delays, and JFK had the most, which is reasonable as departure delays directly influence arrival delays. To acquire a more accurate vision of which airport had the worst arrival delays (not due to longer departure delays, but longer air time), the arrival delays are subtracted from the departure delays to normalize the effect of departure delays on arrival delays. After scaling out the departure delays factor, the order changed: LGA still had the least average arrival delays, but EWR now exceeded JFK to become the airport that had the most average arrival delays. JFK originally had its average arrival delays far more than LGA, but now comes close to it, suggesting that the airport only had the worst arrival delays due to its non-punctual departures. Another noticeable observation is that after subtracting the departure delays, all three origin airports had an average arrival departure of negative numbers. This means flights usually arrive earlier than the scheduled arrival time if departure delays are disregarded, implying an overall shorter air time than predicted.

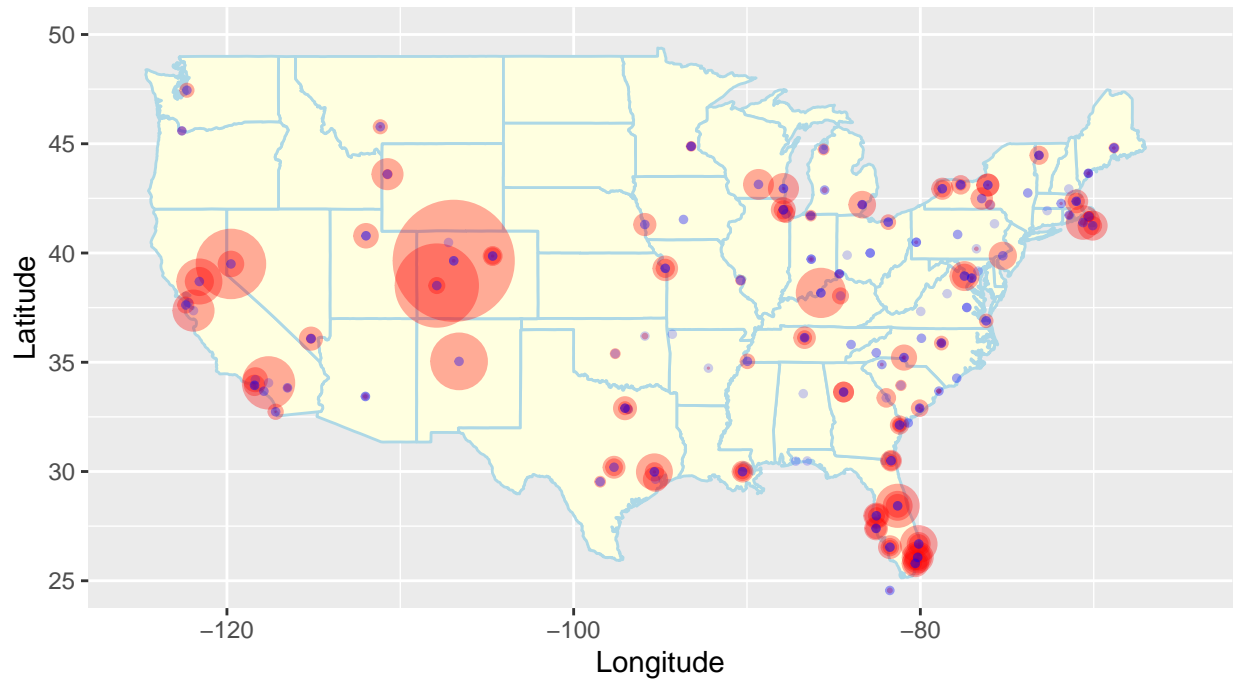
Which destination airport(s) had longer average departure delays and which had shorter?



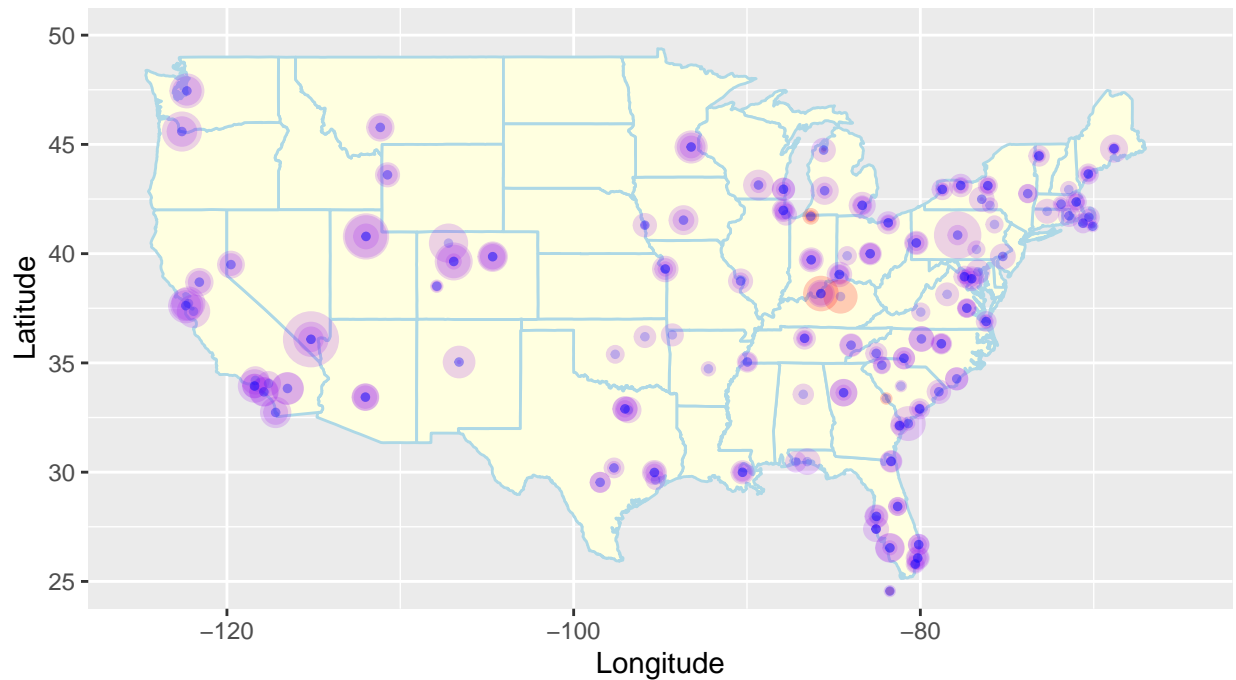
Airports on the West Coast and in the South/Southeastern regions tend to have more departure delay, but since departure delay shouldn't have much association with the airports at which the flight arrive, this could be a result of an origin airport that tends to have more delays have more flights to these destination airports.

Which destination airport(s) had longer average arrival delays and which had shorter?

Average Arrival Delay by Destination Airports

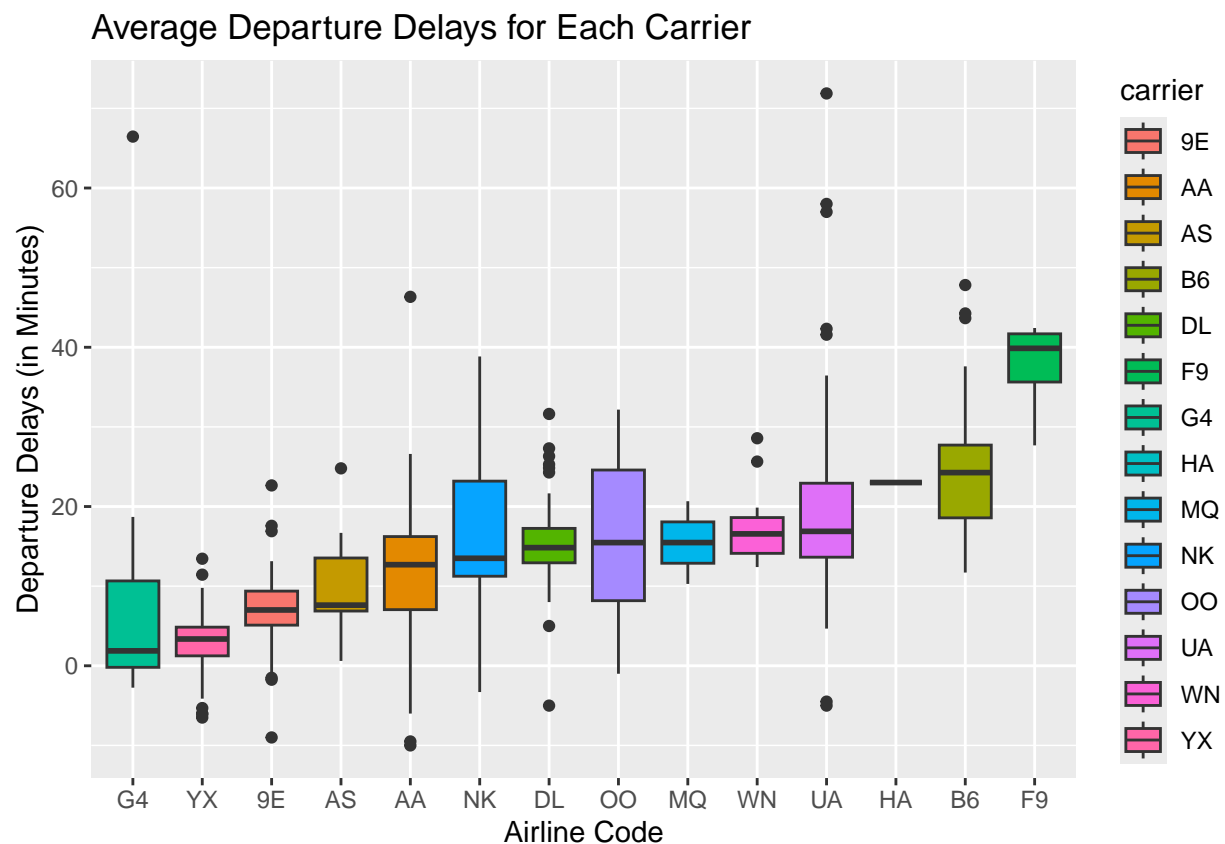


Average Scaled Arrival Delay by Destination Airports



The overall pattern suggests that flights of longer distances tend to cause longer average arrival delays, with a few exceptions on the East Coast. Since long-distance flights are generally more unpredictable than flights of shorter distances, it is reasonable that airports on the West Coast have longer average arrival delays. The rapid change in weather conditions, emergencies during the flight, landing conflicts due to less predictable air time, and many other factors could lead to this result. After normalizing the departure delays, arrival delays among destination airports had little variations. Most airports had negative delays (meaning the flights took less time than expected to arrive at the destination) with very few exceptions labeled on the map as red dots. This result suggests that there is not a clear relationship between destination airports and the magnitude of arrival delays.

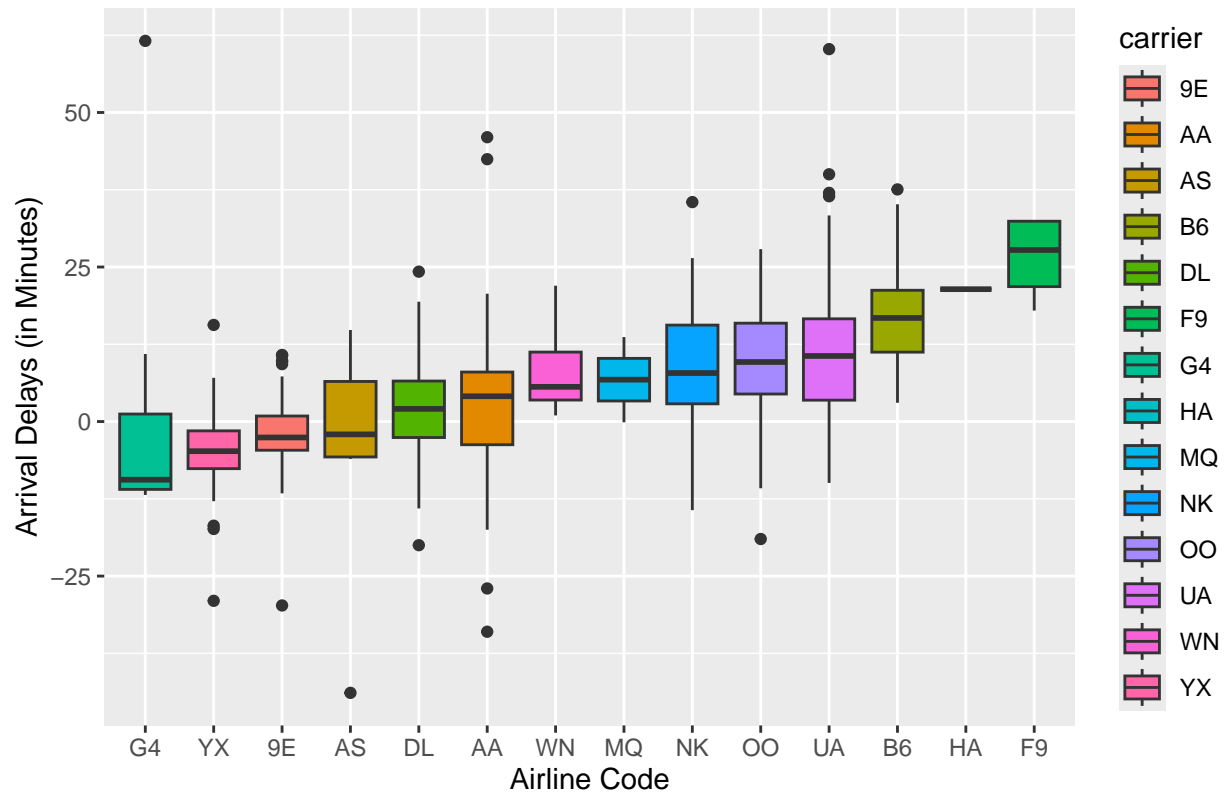
Do average departure delays vary a lot between carriers? If so, which airline(s) had longer delays and which had shorter?



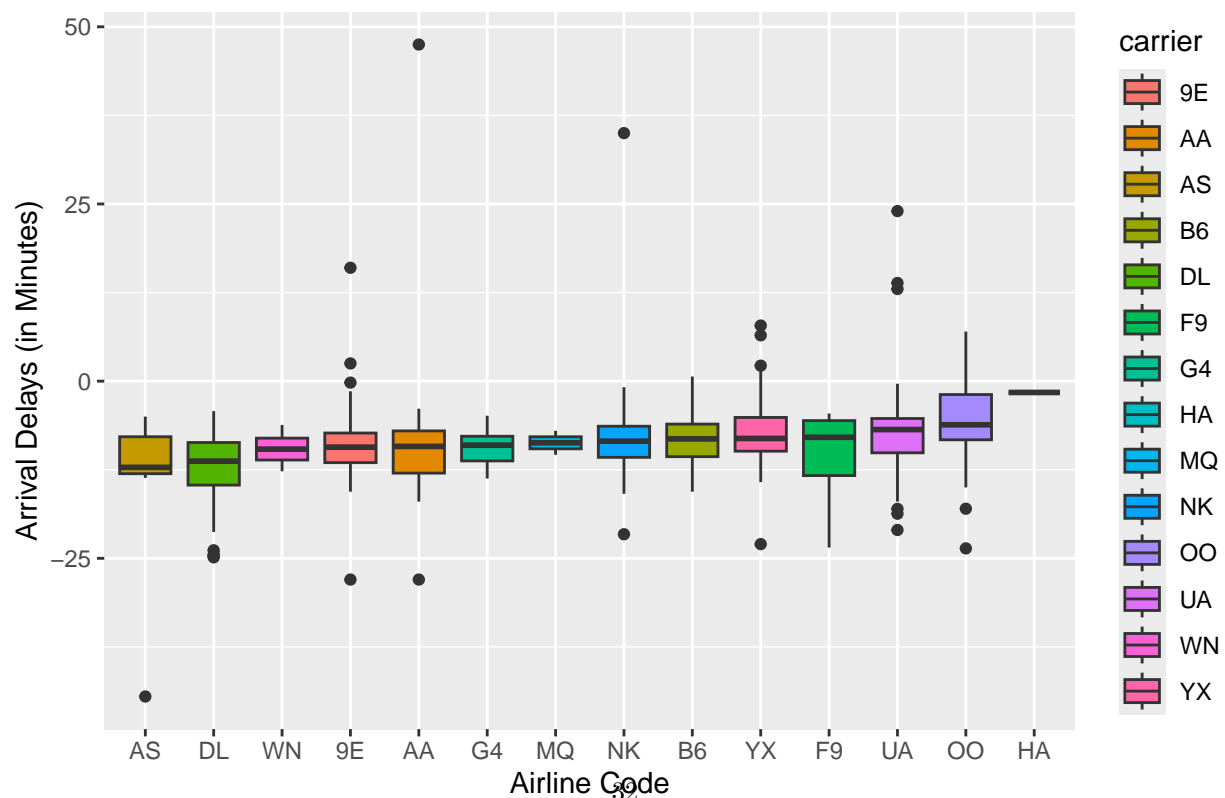
Departure delays do vary among airlines, with F9 (Frontier Airlines) leading the worst departure delays by a significant amount and G4 (Allegiant Air) having the least delays. Among all the airlines, G4, NK (Spirit Air Lines), and OO (SkyWest Airlines) have the longest “box,” suggesting a wider range of departure delays performance, which indicates the carriers’ instability on punctuality. On the other hand, HA (Hawaiian Airlines), DL (Delta Air Lines), YX (Republic Airline), and a few other carriers have relatively shorter “boxes,” demonstrating their more stable on-time performance. Even though G4, B6, and F9 were all budget airlines, G4 was able to achieve an average departure delay median of almost 0 minutes, whereas B6 and F9 both struggled to meet their scheduled departure time. The possible reason behind this interesting phenomenon could be that G4 had significantly better scheduling, ground handling, and staffing ability compared to B6 and F9, allowing punctual flights to be possible while minimizing utilization costs, which the other two budget airlines failed to do so. Though lacking an absolute trend as departure delays could be affected by many other factors, it is still believed that budget airlines, in order to suppress costs, tend to have longer departure delays.

Do average arrival delays vary a lot between carriers? If so, which airline(s) had longer delays and which had shorter?

Average Arrival Delays for Each Carrier



Average Scaled Arrival Delays for Each Carrier



The rank was roughly the same as that of departure delays, which was expected as departure delays would result in arrival delays. To acquire a more accurate rankings, the same normalization technique is applied (subtracting arrival delays by departure delays). After scaling, there was not a big difference in average arrival delays among carriers. Same as the result when analyzing the relationship between origin airports and arrival delays, all airlines had negative arrival delays, suggesting an overall shorter air time than predicted. Therefore, it can be concluded that carriers that had longer arrival delays before normalization were probably affected by their departure delays, not extended air time.

Conclusion

The analysis results support most of the hypotheses. JFK and EWR (both are international airports) had longer average departure delays, whereas LGA (serves primarily domestically) had the shortest average departure delays. The same conclusion can be applied to longer average arrival delays, which means there is still a correlation between origin airports and arrival delays (even after scaling out the departure delays), but the origin airports are probably not the cause to this phenomenon. Destination airports of further distance had longer average departure and arrival delays, with the latter showing stronger correlation as arrival delays are inextricably linked with flight distance. In regards to the exceptions on closer airports having long delays, since operational efficiency also impacts arrival delays, poor ground management induced landing delays might be the cause. Budget airlines do tend to have longer delays, with Allegiant Air being an exception for having the least average delays out of all carriers, possibly due to outstanding ground handling, operational performance, and flight scheduling. A bonus observation from the analysis is that after scaling out the departure delays, average arrival delays based on both airports and carriers are mostly negative, suggesting an overall shorter air time than predicted.

Question 5: Do air-traffic disruptions (higher average departure/arrival delays) predict short-term equity returns for the major carriers?

With the following analysis, I will be investigating and answering the following questions:

- How does departure delay affect airline carrier stock prices?
- How does arrival delay affect airline carrier stock price?
- What is the prediction error?

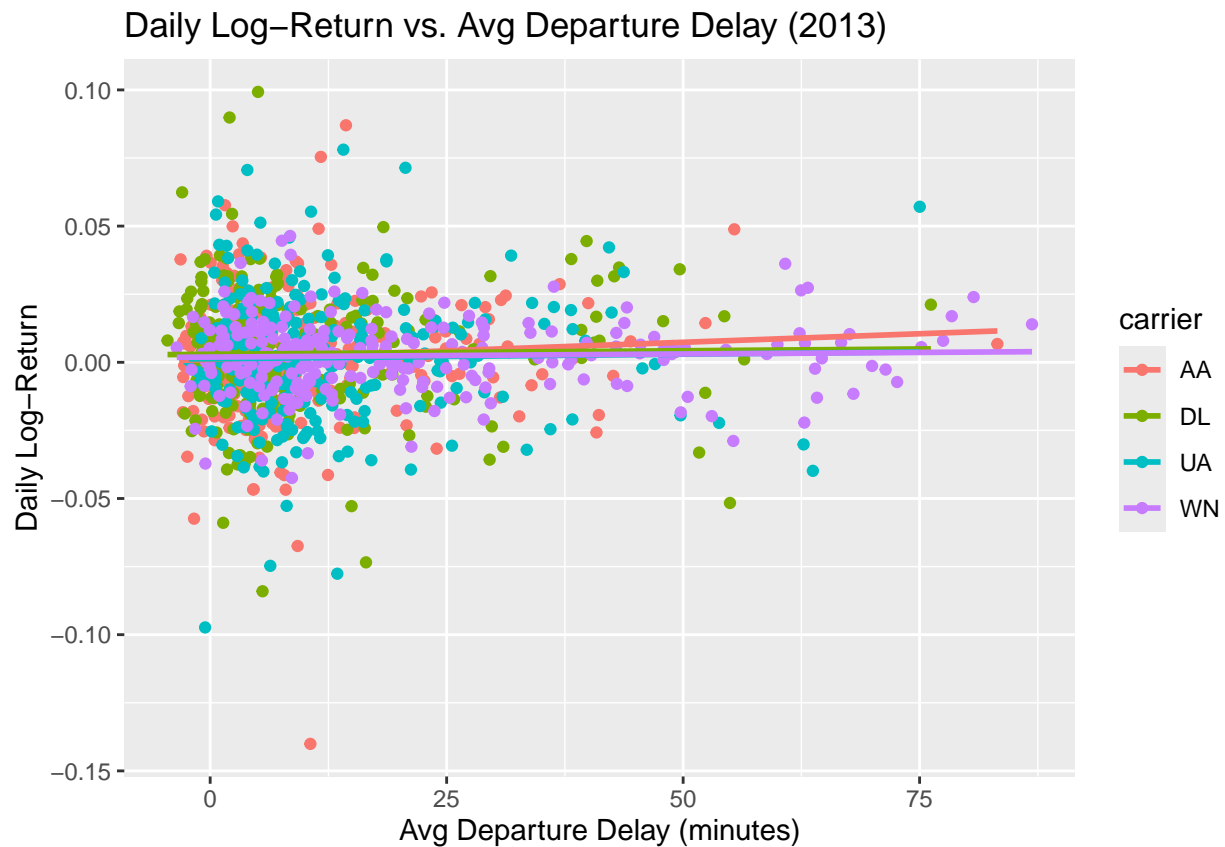
Hypotheses:

- Departure delay will have a negative affect on carrier stock prices.
- Arrival delay will have a negative affect on carrier stock prices.

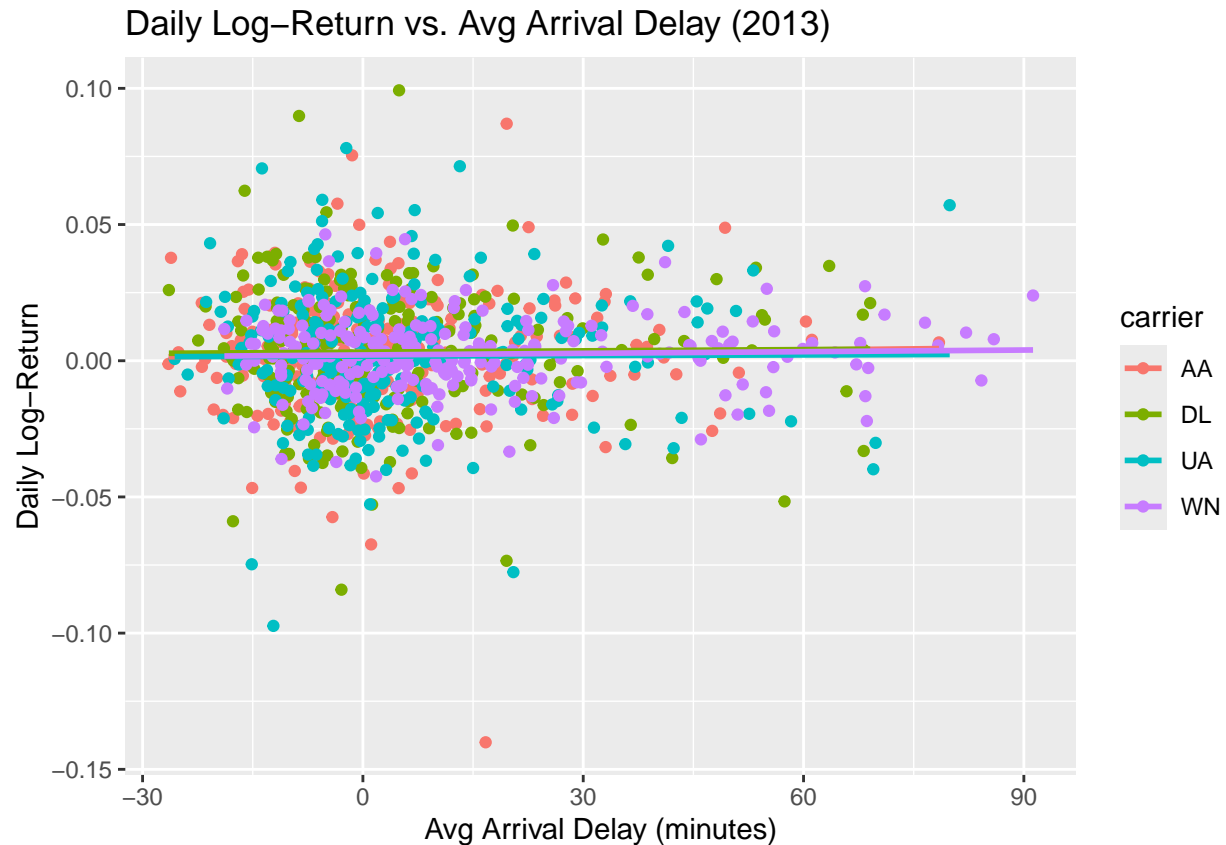
Setup

How do departure/arrival delays impact American airline stock returns?

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
## 'geom_smooth()' using formula = 'y ~ x'
```



Estimating prediction error using 10-fold cross-validation

```
## [1] 0.0004528666
```

Analysis

P-value/Statistical Significance: - Intercept ($p = 0.459$) is not significant at the 0.05, so we cannot conclude the baseline return (for AA when delays = 0) differs from 0 -avg_dep_delay ($p = 0.259$) is not significant, so there is no evidence that average departure delay predicts daily log-return. - avg_arr_delay ($p = 0.4$) is not significant, so there is no evidence that the average arrival delay predicts daily log-returns - factor(carrier)DL ($p = 0.622$), UA ($p = 0.690$), WN ($p = 0.797$): None of the carrier-dummy coefficients differ significantly from the AA baseline.

Thus, all six predictors have p-values greater than 0.05, so none are statistically significant predictors of daily stock returns using multiple regression model

Coefficients: - Intercept = 0.001202: When avg_dep_delay = avg_arr_delay = 0 and for carrier AA, the expected log-return is about 0.12%. - avg_dep_delay = 0.0001330, so each additional minute of average departure delay is associated with a 0.0133% increase in log-return, which is not significant). - avg_arr_delay = -0.00007639: Each additional minute of average arrival delay is associated with a 0.0076% decrease in change in log-return, which is not significant - DL vs AA = 0.0009331: After adjusting for delays, DL's average log-return is about 0.093% higher than AA's, which is not significant - UA vs AA = -0.0007567: UA's average log-return is about 0.076% lower than AA's, which is not significant - WN vs AA = -0.0005032: WN's average log-return is about 0.050% lower than AA's, which is not significant

R-squared: -Multiple R-squared is about 0.0022m so the multiple linear regression model only explains about 0.22% of the variability in daily log-return.

Cross-Validation Analysis: - $MSE = 0.0004529$, the average of the squared differences between the model's predicted log-returns and the observed log-returns across the 10 folds - $RMSE = 0.0213$, so on the log-return scale, our model's predictions are off by about 2.13% on average

Given that daily log-returns for airlines often range only a few percent, an average error of over 2% indicates the model has very limited predictive accuracy, which is no better than predicting the overall mean. Also our model did not over fit as the MSE and 10-fold MSE are about the same. However, given a very low R-square, it is likely the model is under fitting rather than over fitting.

Conclusion

With this extremely low R-squared and lack of statistical significance across predictor suggest that average flight delays and carrier identity, as measured in the code above, do not meaningfully predict daily stock returns. Other factors like macroeconomic news, broad market trends, fuel costs, etc. are likely far more important than delays.

Conclusion & Implications

Real-World Impact

Q1. Looking at the patterns regarding flight punctuality in NYC, airlines can enhance scheduling strategies, lessen operational bottlenecks, and better manage staffing by identifying consistent patterns in flight punctuality, such as delays increasing throughout the day and peaking during the summer. For travelers, this information supports smarter booking decisions, potentially leading to fewer missed connections and a more reliable travel experience.

Q2. Comparing operational efficiency from 2013 to 2023 is crucial to gauging how well airport operations have fared. Airports have implemented changes over the past decade, including new terminals, runaways, technological improvements, or gate management. Comparing the average delays can help whether these investments actually worked. A 10-year comparison also allows us to highlight whether delays have worsened or improved over time. It helps airport authorities and policymakers understand which operations are working and where strategic improvement is possibly needed. In this 10-year comparison, not enough evidence was shown to indicate an improvement in airport operations. However, they haven't worsened drastically, despite the flight volume increasing significantly across airports, which goes to show that the airport improvements that have been made within the 10 year span may only be significant enough to offset the increased volume.

Q3. The ability to use weather as a predictor for flight punctuality would be immensely helpful to airlines, potentially reducing flight delays and minimizing cancellations. While keeping customers happy is an added bonus, it could also increase safety as weather patterns become more predictable.

Q4. Analyzing the relationship between average departure/arrival delay and airports not only give passengers a general idea of airport efficiency, but it also builds a foundation for further study focusing on the deciding factors behind airport efficiency and provide insights on possible improvements. The same can be applied to carriers, with additional possible fields to study such as price-performance comparison among carriers.

Q5. From the results, airline delays do not have statistical significance on carrier stock prices. Thus, investors are not likely to anticipate routine delays, especially the daily swings in average delays, and these get priced in before market close. Macroeconomic factors such as oil prices, interest rates, earnings releases, industry-wide news (M&A deals, regulatory actions) and broad market sentiment typically reflect a relatively minor effect of one day's operational hiccups.

Limitations

One limitation we had was that a large amount of the analysis was based on a single dataset. Even with additional datasets to complement nycflights13, there was still a limit to how much information could be extracted.

Another possibility was potential interpretation bias. Despite the data analysis giving accurate results, everyone may have had some interpretation bias, which would skew how we each look at the results. For example, one person may determine that a spike on a graph is insignificant, while another may see it as notable, creating two different conclusions.

Another limitation may be sample size and failure to meet the normality/variance condition. Some of the data we extracted for the dataset was very limited in numbers, even when supplemented. As a result, some of the data in a few areas of study may not have been normalized. This can possibly have caused skewed and inaccurate results, weakening the overall analysis.

Authors' contributions

- Chloe Lavery: Data Exploration and Visualization (Question 3), Conclusion & Implications
- Shirley Lee: Data Exploration and Visualization (Question 4), Conclusion & Implications
- Zach Nguyen: Data Exploration and Visualization (Question 2), Conclusion & Implications
- Mihir Ranjan: Data Exploration and Visualization (Question 5), Conclusion & Implications
- Rhea Verma: Data Exploration and Visualization (Question 1), Conclusion & Implications
- Krishna Vijayendra: Introduction, Methodology & Analysis Plan, Conclusion & Implications

Link(s) to Data/Code

- <https://github.com/tidyverse/nycflights13>
- <https://github.com/moderndive/nycflights23/>