United Airlines: New York City Flight Delays

Introduction

This analysis aims to identify the key factors contributing to departure delays for United Airlines departing from New York City for year 2013. Flight delays are a persistent issue that affect airline operations, passenger experience, and overall efficiency.

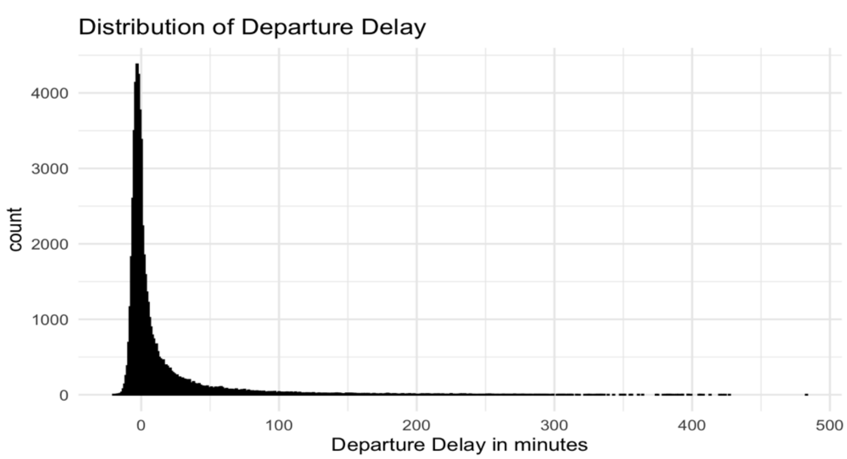
This study examines how the variables such as time of day, time of year, temperature, wind speed, precipitation and visibility influence the departure delays.

Data Methodology

The data package (nycflights13) we are using in our study contains information about all flights that departed from New York City.

The flights dataset contains on-time data for all flights that departed NYC (i.e. JFK, LGA or EWR) in 2013.The weather dataset contains information about hourly meteorological data for LGA, JFK and EWR. The flights and the weather dataset are combined using unique variables such as year, month, day, hour and origin. Here we focus only on data related to United Airlines. We have created additional 2 variables in the dataset. i.e. late (when Departure delay >0) and very late (when departure delay >30). This combined dataset is using for further analysis. The dataset has null values which we have ignored in our analysis.

United Airlines Flight Delays



The figure shows the distribution of departure delays for United Airlines. As we can see that most of the flights have very small or no departure delays and only few flights have a very large flight delay.

The summary statistics for the UA flights are given in the below table.

	Minimum	Mean	Median	Maximum
minutes	-20	12.09	0	483

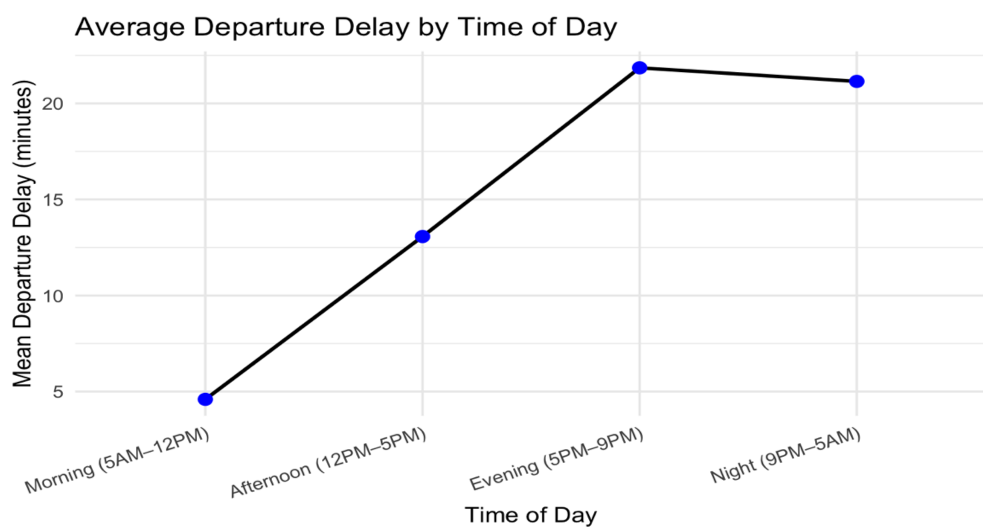
As we can see from the above statistics that few flights have departed early (-ve value). The average delay is 12.09 minutes. The median is 0 which is a positive thing for the airlines and the longest delay shows 483 minutes (8.05 hours).

Results

Time of day

To make it easy to visualize in graphs we have divided the hours into 4 categories as below

Morning (5AM–12PM)
Afternoon (12PM–5PM)
Evening (5PM–9PM)
Night (9PM–5AM)



The analysis shows a clear relationship between the time of day and departure delays for United Airlines flights. Morning flights (typically between 5 AM – 12 PM) tend to have the lowest average delays and the highest on-time performance. As the day progresses, delays gradually increase, peaking during the late afternoon and evening hours (3 PM – 9 PM).

This pattern is likely due to cumulative delays from earlier flights that propagate throughout the day, as well as heavier air traffic and congestion during peak operational periods.

Night flights (after 9 PM) show a slight decline in delays, possibly because fewer flights operate during these hours, leading to reduced congestion. This time-of-day pattern highlights the importance of early scheduling and proactive delay management in the afternoon period to improve on-time performance.

The summary statistics for time of day is shown below

Time of day	Flights	Mean delay	Median delay	Max delay	Late %	Very late %
Morning (5AM–12PM)	23661	4.60	-2	420	33.26	6.15
Afternoon (12PM–5PM)	17512	13.07	1	483	51.07	13.15
Evening (5PM–9PM)	15333	21.85	5	406	62.33	22.85
Night (9PM–5AM)	1180	21.15	6	230	61.19	26.02

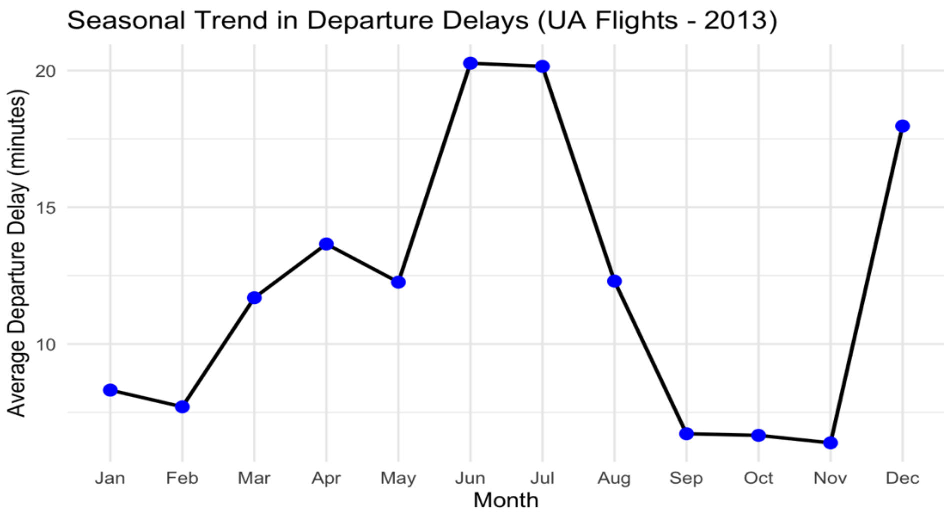
The results show a clear upward trend in departure delays as the day progresses. Morning flights (5 AM–12 PM) experience the lowest average delay of about 4.6 minutes, with only 33% of flights departing late. In contrast, afternoon flights (12 PM–5 PM) have a noticeably higher mean delay of 13.1 minutes, and over 51% of flights are late. Delays peak during the evening period (5 PM–9 PM), where the mean delay rises to 21.9 minutes and more than 62% of flights are delayed, indicating congestion and cascading effects from earlier departures.

Although night flights (9 PM–5 AM) have a slightly lower mean delay (21.2 minutes) compared to evening flights, they still show a high proportion of late departures (61%) and the highest rate of very late flights (26%).
These results suggest that operational congestion and accumulated delays throughout the day contribute significantly to later departures in the evening and nighttime hours.

We conducted permutation tests to check if there was significant difference in mean delays between the time of the day. We did a combination of 6 tests taking 2 categories at a time. The results for 5 pairs namely Morning & Afternoon, Morning & Evening, Morning & Night, Afternoon & Evening, Afternoon & Night we found that there were significant difference in the mean departure delays indicating that this could not have happened due to random chance. Only for one category i.e. Evening and Night the permutation result showed that there was no significant difference between the mean departure delays, and this could have happened based on random chance alone.

Hence, we can conclude that there is a relationship between Time of day and Departure Delay.

Time of Year



The chart shows noticeable seasonal variation in average departure delays. Delays are lowest during the winter months (January–February), with average delays below 10 minutes. They increase steadily through spring, peaking sharply in June and July, when the average delay exceeds 20 minutes. This summer spike likely reflects increased travel demand, airport congestion, and weather-related disruptions such as thunderstorms.

After July, delays decline sharply in late summer and fall (August–October), reaching their lowest levels around September and October. However, there is a rise again in December, which may be linked to holiday travel surges.
Overall, the data indicate that summer and holiday months experience the greatest delay challenges, while early fall and late winter tend to be the most punctual period.

Month	flights	Mean delay (minutes)	Median delay (minutes)	Max delay (minutes)	Late %	very_late %
Jan	4595	8.31	0	385	44.87	9.27
Feb	4157	7.70	-1	266	42.75	9.21
Mar	4926	11.69	0	408	46.04	13.01
Apr	4991	13.65	-1	427	44.80	15.03
May	4910	12.26	0	406	49.04	13.50
Jun	4910	20.27	2	420	56.07	20.18
Jul	4985	20.15	2	483	56.03	19.72
Aug	5064	12.30	0	424	49.13	13.39
Sep	4645	6.72	-2	422	33.39	7.51
Oct	5003	6.66	-1	292	37.56	8.41
Nov	4825	6.38	-1	351	38.13	8.00
Dec	4675	17.97	5	392	64.96	19.21

The summary statistics indicate that departure delays vary noticeably across months, reflecting clear seasonal trends. The lowest average delays occur during September to November, with mean delays of around 6–7 minutes and fewer than 40% of flights departing late, suggesting smoother operations in early fall. In contrast, summer months (June and July) experience the highest mean delays, averaging around 20 minutes, with more than 56% of flights leaving late, likely due to heavy travel demand, air traffic congestion, and weather-related disruptions such as thunderstorms.

A similar spike is observed in December, where both the mean delay and the percentage of late flights rise sharply, reflecting holiday travel surges and winter weather impacts. Overall, the data show that seasonal and weather-related factors strongly influence flight punctuality, with summer and winter months posing the greatest delay challenges for United Airlines.

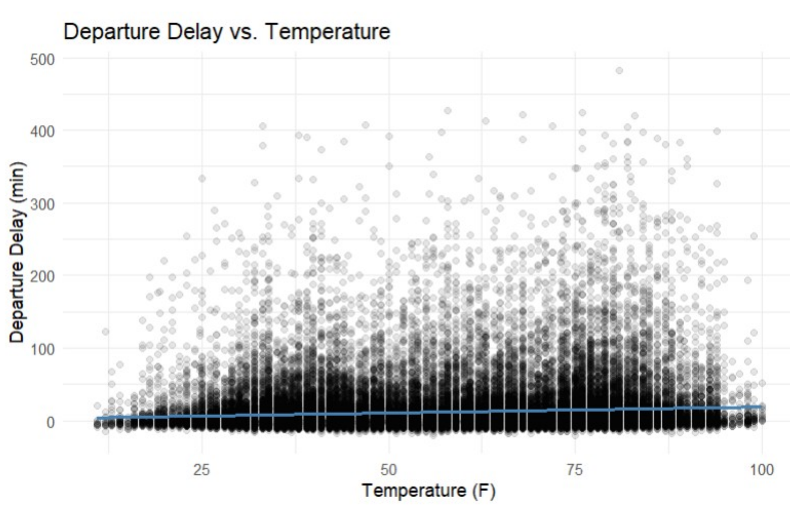
To make it convenient to conduct permutation tests we have divided the months into 4 categories as below

Fall	September, October, November
Winter	December, January, February
Spring	March, April, May
Summer	June, July, August

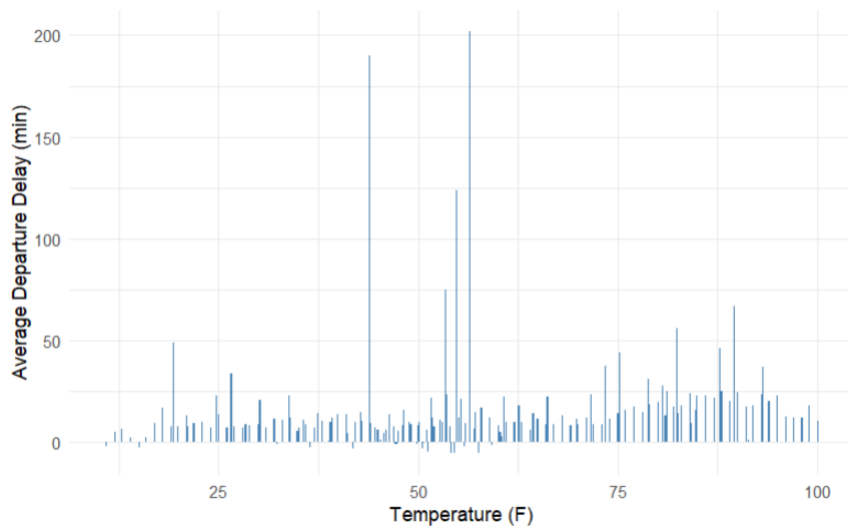
We conducted permutation tests across all seasons to analyze if there is significant difference in the mean departure delays and time of the year. Even here we conducted 6 permutation tests with a combination of 2 categories each. The results showed that the mean departure delay across all the seasons is statistically different and could not have occurred due to random chance.

Hence, we can infer that there is relationship between the mean departure delay and time of the year.

Temperature



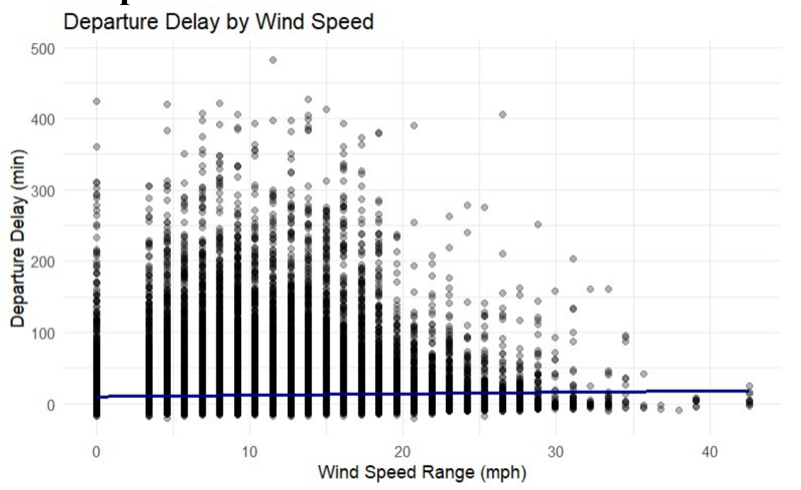
The plot shows a dense cloud of flight data points across the temperature range, with a red linear regression line that is nearly flat. This indicates that temperature has little to no linear effect on whether a flight is late or very late. Flights with delays greater than 0 minutes (late) and greater than 30 minutes (very late) are scattered throughout all temperatures from very cold to very hot. There's no visible clustering of severe delays at any specific temperature range. While moderate temperatures (around 50–75°F) host the highest volume of flights, this reflects operational frequency rather than delay severity. In short, the scatterplot suggests that temperature is not a strong predictor of delay status.



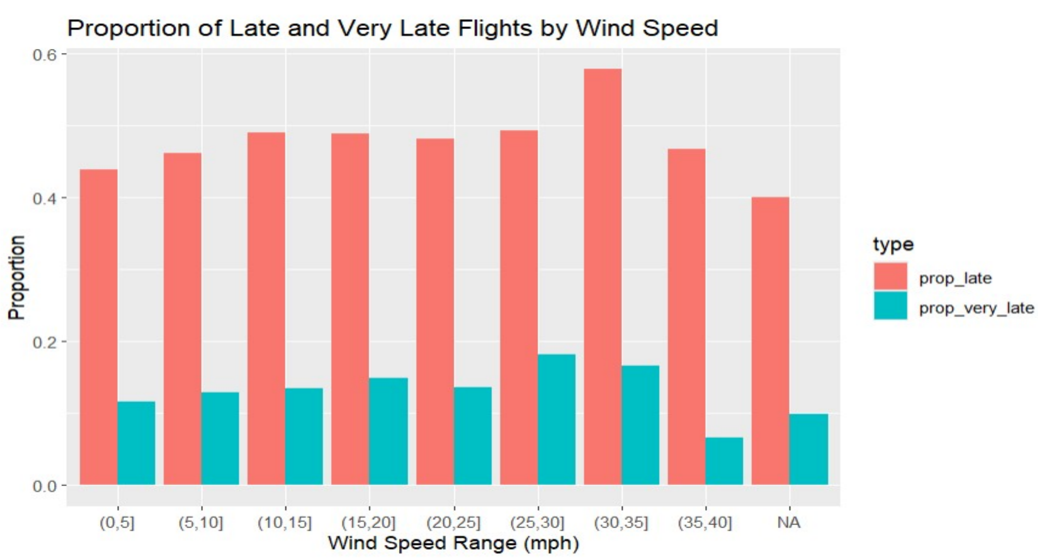
In addition, the bar chart above illustrates how average departure delay varies across temperature values, offering insight into whether certain temperatures are associated with longer delays. While the chart shows some spikes particularly around 30°F and 55°F, there is no consistent upward or downward trend across the temperature range. Most temperature values correspond to relatively low average delays, and the fluctuations appear irregular rather than systematic. This suggests that temperature alone is not a strong predictor of delay severity. The occasional spikes may reflect isolated weather events or operational disruptions rather than a generalizable pattern.

When we compared the temperature differences between departure delays using a statistical test, the results showed that this gap was very unlikely to be due to random chance. This suggests that temperature may play a meaningful role in whether a flight is delayed, even if the chart of average delays doesn't show a clear pattern. The occasional spikes around 30°F and 55°F might reflect isolated weather events or operational issues, but overall, the test points to temperature as a factor worth considering.

Wind Speed



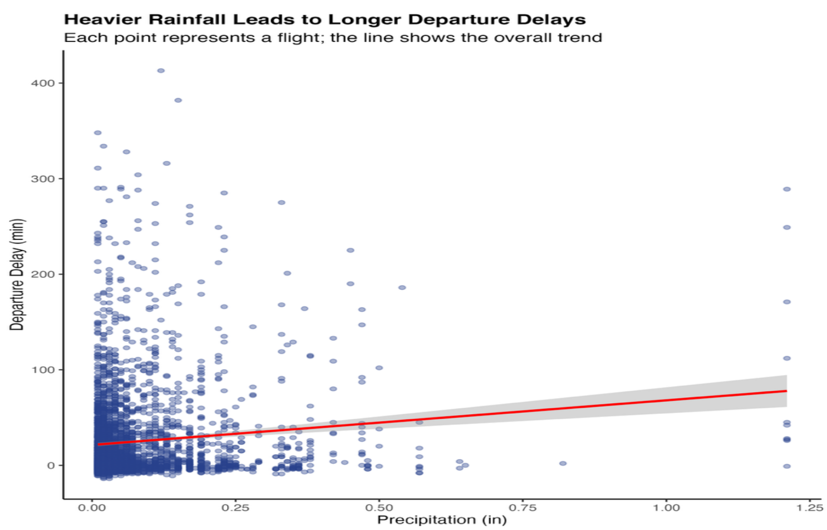
This chart shows how departure delays relate to wind speed. Each dot is a flight, and you can see that most delays happen when wind speeds are low. As wind speed increases, delays seem to get a little shorter, but the change is small. The blue line shows that overall trend, but the data is scattered, so it's not a strong or consistent pattern. In short, wind speed might have a slight effect on delays, but it's probably not the main factor.



The bar chart above compares the proportion of late (dep_delay > 0) and very late (dep_delay > 30) flights across wind speed bins, offering a clear view of how wind conditions might relate to delays. The proportions of late flights (red bars) are consistently higher than those of very late flights (blue bars) across all wind speed ranges. Interestingly, the highest proportion of late flights occurs in the (30,35] mph bin, which is relatively rare in the dataset. However, the overall pattern shows no strong or consistent trend, both late and very late proportions fluctuate slightly but remain stable across wind speeds.

However, a separate statistical test revealed that the difference in wind speed between late and on-time flights is very unlikely to be due to random chances. This suggests that wind speed may influence whether a flight is delayed, even if it doesn't strongly affect how long the delay lasts.

Precipitation



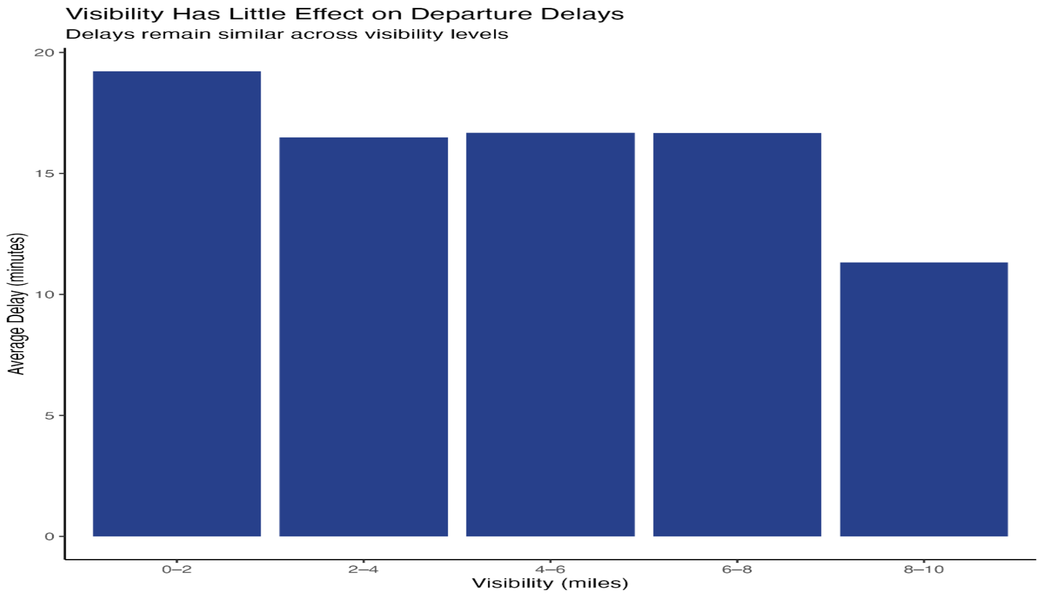
Departure delays were both more common and more severe when it rained.

Most flights took off on dry days, but when rain was present, flights tended to leave later and with greater variation between departure time. Some flights left only a few minutes late, while others were delayed for hours. Regardless, presence of rain still influences departure delay. This could be due to safety, runway congestion, or slower movement on the ground.

When the differences were compared with permutation tests, the results showed that this gap between rainy and dry flights would be extremely unlikely to happen by chance. Thus, there is strong evidence that rainy weather consistently leads to longer delays.

To validate this further, a separate test was conducted after converting precipitation into a binary variable, rain or no rain. The results were consistent with the continuous test, reinforcing the conclusion that precipitation has a meaningful effect on departure times. While most flights happened during dry weather, the ones that occurred in rainy conditions faced measurable disruptions in departure timing.

Visibility



Visibility was clustered around 10 miles, meaning most flights departed under clear conditions. When compared to departure delay on a scatterplot, the regression line was almost flat, implying just a small relationship between the visibility and departure delay.

When the differences were compared with permutation tests, strong p-values indicated a relationship between visibility and departure delay. The graphs show this; there is impact on departure delay and visibility, but the variance between delay time is slight.
Thus, visibility does not meaningfully influence flight departure delays.

Discussion

This study explored the key factors influencing departure delays for United Airlines flights departing from New York City in 2013, focusing on how time of day, time of year, and weather conditions such as temperature, wind speed, precipitation, and visibility affect flight punctuality. The most consistent and impactful finding was that time of day plays a significant role in shaping delays: morning flights experienced the shortest average delays and the lowest percentage of late departures, while delays increased steadily through the afternoon and peaked in the evening, likely due to cumulative operational congestion. Night flights showed slightly lower than average delays than evening flights but still had a high proportion of very late departures. Seasonal trends were also evident, with delays peaking in the summer months, particularly June and July and again in December, likely due to increased travel demand and weather disruptions, while fall months like September and October had the lowest delays.

In contrast, temperature and wind speed showed no strong or consistent relationship with delay severity; scatterplots and bar charts revealed high variability in delays across all temperature and wind ranges, with no clear upward or downward trends. Precipitation, however, emerged as a meaningful factor: flights during rainy conditions experienced significantly longer and more variable delays, a finding supported by permutation tests and reinforced by binary comparisons of rainy versus dry days. Visibility, while statistically significant in some tests, showed only a slight effect on delays, with most flights occurring under clear conditions and the regression line remaining nearly flat. Overall, the study concludes that while weather factors like rain can meaningfully disrupt flight schedules, operational timing, especially time of day and season, has a stronger and more consistent impact on departure delays.

The analysis helped uncover meaningful patterns in United Airlines departure delays, but it also opened new questions worth exploring. While we found strong links between delays and factors like time of day, season, and precipitation, it's clear that delays are shaped by a complex mix of operational and environmental conditions. Future work could dig deeper into how these variables interact, for instance, whether rain has a bigger impact during peak hours, or how delays build up over the course of the day. It may also be helpful to examine each airport individually, as JFK, LGA, and EWR likely operate under different conditions that shape their delay patterns. Another valuable direction is to explore how arrival delays influence later departures, particularly when aircraft are scheduled with minimal turnaround time. By understanding how these patterns build and interact, airlines could better anticipate points of congestion and adjust schedules more effectively. In the long run, integrating real-time weather and flight data into predictive systems could lead to more agile decision-making and help minimize disruptions for both passengers and operations.