SDS 322E Project: Airport Delays

Team: Devin Xiao, Feifei Cao, Joy Richardson, Sofia Bautista, Vedika Shirtekar, Yasmeen Luna

Introduction

Our project analyzes the data of domestic flights in the United States. Our project goals were: (1) explore the relationship between airports and delay times and (2) explore which kind of delays cause longer departure delays.

In 2022 alone, U.S. airlines carried 853 million passengers (*Full Year*, 2023). Air travel is an extremely important mode of transportation in the modern age. According to the Air Travel Consumer Report, nearly a quarter of flights landed late in 2022 (Potter, 2023). On time arrival is critical not only for consumers but the airlines themselves. Delays are at best a minor inconvenience and at worst completely disastrous. Exploring relationships regarding delays is important for understanding underlying causes and development of better systems for dealing with delays.

Data

The Bureau of Transportation Statistics contains a database called Airline On-Time Performance Data. There are two datasets available but the one we used was the Reporting Carrier On-Time Performance. In this dataset, carriers report data for the flights operated, including their arrival/departure dates, carrier, origin, airport, causes of delay and cancellation, air time, and non-stop distance. The dataset spans from 1987 to the present.

Since we were interested in the relationship between flights and hurricane season, we decided to look at a span of years from 2020 to 2022. Notably, the 2020 Atlantic hurricane season was very active, having the most number of storms and second most number of hurricanes on record (*Record-breaking*, 2021). However, we must take into consideration that the entire aviation industry was greatly affected by COVID-19 in 2020, which may affect the generalizability of our findings. To control the size of our dataset, we focused on Texas flights, where either the origin, the destination, or both are Texas cities.

Additionally, we wished to combine our dataset with hurricane data to study the relationship between storms and flight delays, especially weather delays. We found information on dates, intensities, and states affected for storms and hurricanes by the National Oceanic and Atmospheric Administration. Then, we combined this data with the flight data by month, year, and states affected (*Continental United States Hurricane Impacts*, 2023). It may be more appropriate to combine the data by day rather than assuming the storm affects the entire month.

Data cleaning consisted of dropping missing data, selecting for variables of interest, and merging to simplify analysis. The data was filtered for delay types, origin and destination related variables (ex. Origin city), state, day, year, distance, and distance group. All delay types were grouped together under one variable to simplify analysis of delay type rather than several separate variables. Tidy long was performed when grouping delay types together, where "delay_minutes" was created to pair how late a flight was due to a particular delay. Origin and destination variables like "origin_city_name" and "dest_city_name" were split to separate the city from the state, and the distance group was converted to a categorical variable with levels as each numbered group was previously listed as a character. Then, we filtered for non-cancelled flights because these observations did not have information about delay times. We also mutated the scheduled departure and arrival times to consist of only the hour and discarded the minutes section. Lastly, we grouped by year and month and sampled 20000 observations from every month for a total of 720000 observations to make the dataset more manageable.

Exploratory Analysis

Airport Sizes and Delay Times

Our first idea was to explore the relationship between airport sizes and delay times. First, we created a new variable called number_flights that counted the number of flights from an origin airport as a measure for the size of the airport. Then, since our dataset had disproportionately more flights from Texas than any other state, we decided to consider only flights from Texas to study the relationship between airport size and delay times. Otherwise, we can imagine a large international airport having a smaller total number of departures than a moderately sized Texas airport in our dataset and thus being incorrectly classified as a smaller airport. To make a distinction between large and small airports, we used the cutoff of the top 6 airports. This number was chosen based on where there appeared to be a natural break or gap in the number of flights.

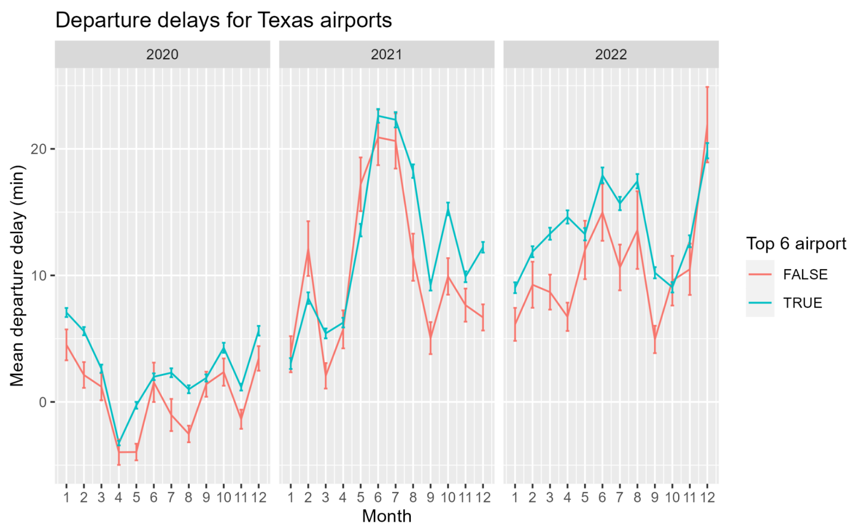
Figure 1. Number of Flights by Texas Airport

ORIGIN <chr>	n <int>
DFW	166315
IAH	84852
AUS	38017
DAL	37331
HOU	30088
SAT	17992
ELP	8944
MAF	4819
LBB	3652
CRP	2835

We can see that larger and smaller airports follow a similar monthly trend. Larger airports have longer mean departure delays compared to smaller ones. One explanation could be that a larger volume of flights is simply more difficult to manage. Otherwise, it could be that the mean distance for flights from larger airports is greater than the mean distance for flights from smaller airports. For instance, the longer the distance, the more probable that a problem will occur during planning or flight. It could be the case that smaller Texas airports have Texas destinations make up a larger proportion of their total flights, and flights from and to Texas are less likely to experience long delays.

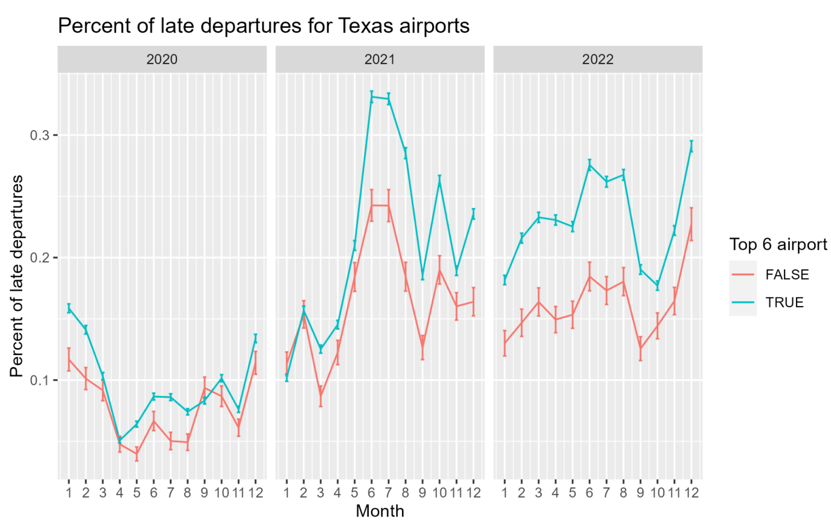
We can also see a significant decrease in delay times for the entirety of 2020, which had significantly less flights overall than the next two years. In fact, this supports the idea that the difference between delays for large and small airports is because more traffic is more difficult to manage. Still, we see a difference even when flights are more manageable for larger airports.

Figure 2. Mean Departure Delay Times for Texas Airports



Now, looking at the percent of flights that departed late, we find more separation between the larger and smaller airports. For instance, compare August 2022 in **Figure 2** and **Figure 3**. This seems to imply that smaller airports are better at preventing delays but have slightly less control over the actual delay times. Here, we also notice a smaller gap in percent between larger and smaller airports in 2020 compared to the next two years.

Figure 3. Percent of Late Departures for Texas Airports



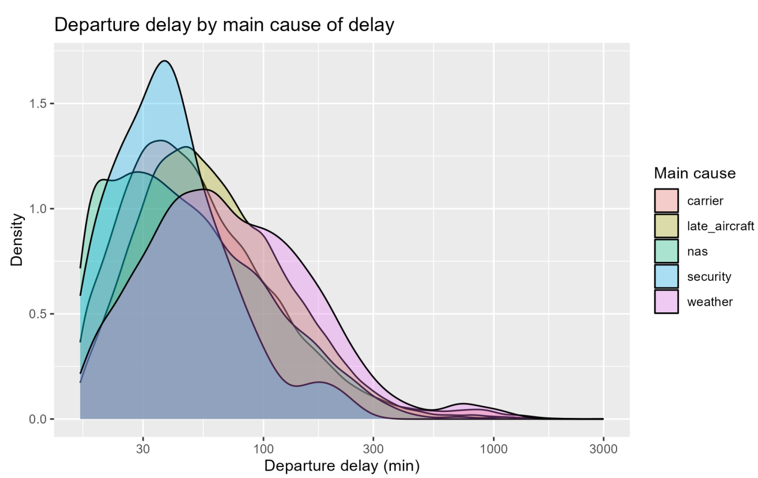
Causes of Delays

We then explored which kinds of delays cause the longest departure times. First, we filtered the dataset for late departures, so we know that the delay types have some effect on the departure and not entirely on the arrival. We created a density plot with all of the causes of departure delays as shown in **Figure 4**. Here, we used a logarithmic scale to visualize the delay times more clearly. We can see that weather is one of the main causes of long departure delay times as the density covers more area.

Here, we made the assumption that the main cause of delay contributes to the departure delay. While we do not know which parts of the delay types contribute to the departure or arrival, we know that arrival delay is affected by departure delay. That is, if a flight departs late, it is likely to arrive late as well. Thus, if we have that a certain type of delay is the main cause of delay for a flight and it has affected departure, then it will also be a major factor in arrival delay. Additionally, we do not consider the percentage contribution of each cause of delay. Oftentimes, there is more than one cause that contributes to the overall delay. By focusing on the main delay, we are discarding information on the details to gain a bigger picture understanding.

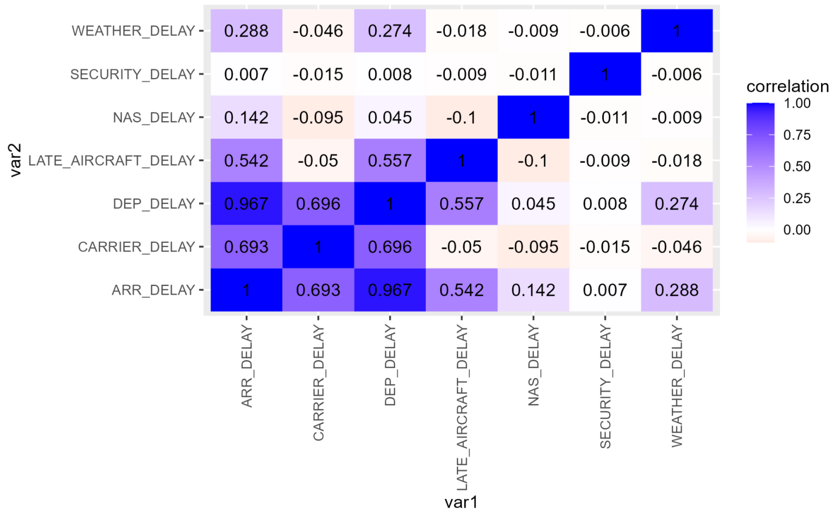
We hypothesize that weather delay is the most important predictor of delay times based on its density covering the longest delay times. This could be because weather conditions are uncontrollable. For instance, while we have some control over operations and security, the best we can do about bad weather is to wait it out.

Figure 4. Density Plot of Causes of Departure Delays



Looking at the correlation matrix of delay types, we can confirm that arrival delay and departure delay are highly correlated to each other. Additionally, it seems that carrier delay followed by late aircraft delay are the leading causes of delay times. Interestingly, while National Airspace System delay is not very correlated with either departure or arrival delay, it is significantly more correlated with arrival delay than departure delay. Lastly, unlike what we predicted, it seems that weather delay is only weakly correlated with arrival and departure delay. To create the correlation matrix, we did not make any changes to the relevant variables. This means that the matrix only considers late flights, since on-time flights have NAs for weather, security, NAS, late aircraft, and carrier delay.

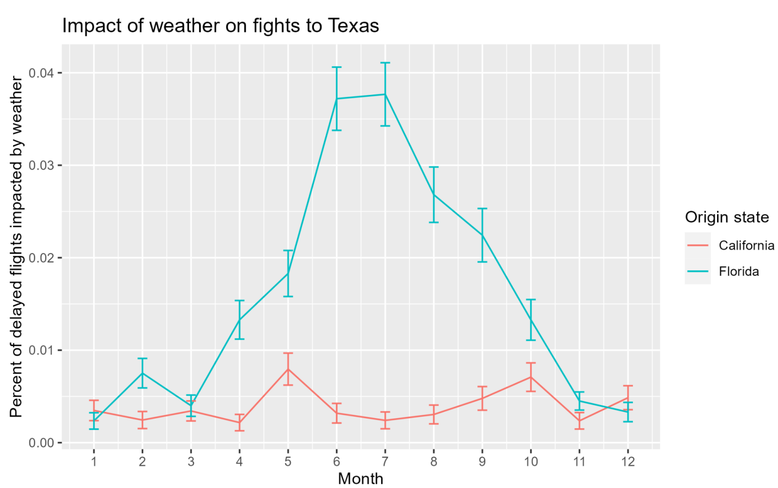
Figure 5. Correlation Matrix of Delay Types



Impact of Weather on Flights

Based on our findings in **Figure 4**, we explored if weather delays are specific to certain months of the year. For example, colder months like December - February might have higher departure delay times. In this case, we explored hurricane season. We defined hurricane season to be from June through November according to NESDIS (*2023 Atlantic Hurricane*, 2023). We chose to focus on Florida as it is known for hurricanes. We also chose to compare Florida and California as they had similar amounts of flights to Texas. We can see that in June through September of 2022, Florida had the highest percentages of delayed flights signifying to us that weather delays can be specific to certain months of the year. In fact, this period corresponds to hurricane season.

Figure 6. Impact of Weather on Flights to Texas



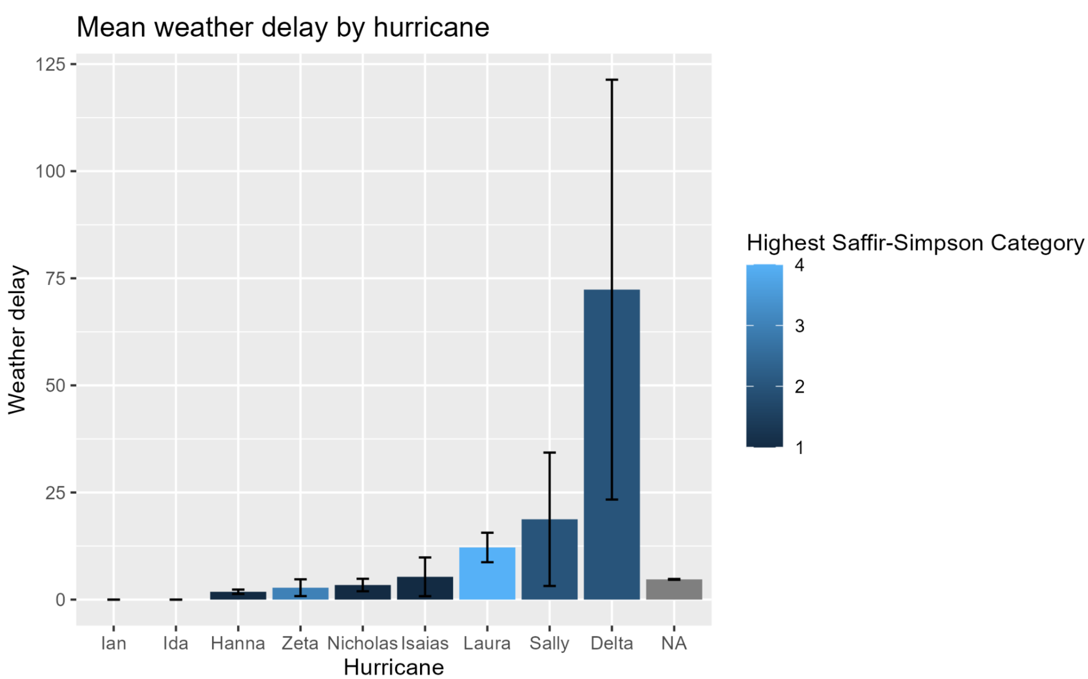
Another area of interest could be the effects of hurricanes on weather delays. First, since on-time flights do not have a weather delay, we decided to look at late flights only. However, another approach could have been to set weather delay to 0 for all non-delayed flights. Using the first approach, we can explore the contribution of hurricanes to delay times for delayed flights. Using the second approach, we can explore the impact of hurricanes on flight delays as a whole.

Previously, we established that a flight is hurricane-affected if it is in the same year and month as a hurricane and one of the states (origin or destination) was affected by the same hurricane. However, since this definition may be a little loose, we decided to look at all flights from 1 day before to 5 days after each hurricane landed within the same month.

The Saffir-Simpson Scale categorizes hurricanes based on their wind speed, with category 3 and above being considered major hurricanes. While we may expect more dangerous hurricanes to cause longer weather delays, we do not necessarily see that reflected in **Figure 7**. In fact, most of the hurricanes do not contribute at all to weather delay. This could be because rather than being delayed, flights scheduled during a hurricane are canceled. For instance, the large error bars for hurricanes Zeta, Delta, and Ida could be because there were not many flights scheduled during this selected time period.

Alternatively, we did not consider localized areas of impact or airports. For instance, we assumed that a hurricane that affected a state had the same impact across the entire state. Then, an airport whose location is far from the area of impact but still in the same state may not experience any special weather delay at all, but this is not accounted for.
https://www.weather.gov/mfl/saffirsimpson

Figure 7. Weather delay by hurricane



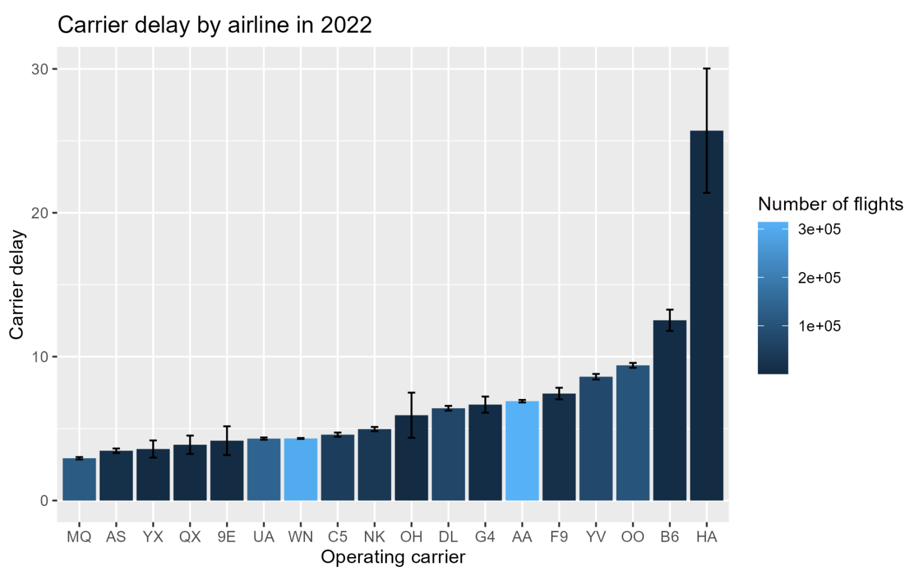
Relationship between airline (carrier) and carrier delay

Our last question revolves around the relationship between the operating airline and carrier delay. We were interested in whether certain carriers were better at preventing delays than others. In particular, there is a kind of delay called carrier delay, which concerns delays caused by an airline's operations and utilization of resources. If we find a significant difference in carrier delay between different carriers, then we may know which airlines are better organized or which airlines to avoid.

First, we looked at carrier delay. Since our original dataset did not have information about operating carriers, we needed to download a new dataset. One limitation of this new dataset is that it only contains information on flights from 2022. Since a flight does not have a carrier delay if it is not delayed, we set carrier delay to 0 for all non-delayed flights. However, it may also be interesting to remove NAs for carrier delay to see how different carriers handle operations when a flight is actually delayed. Then, since the size of a carrier may affect its organization and be a factor in delay times, we grouped by carrier and found the number of flights from each carrier.

Looking at **Figure 8**, we can see a noticeable difference in carrier delays between different airlines, with Hawaiian Airlines having the largest mean delay by a large margin. However, it is difficult to see a trend between the carrier delay and the size of the airline. We can notice that smaller carriers have both the shortest and longest delay times, with larger carriers being in the middle. Interestingly, this result does not completely reflect our findings about airport sizes and delay times. This could be because smaller and larger airports have a similar proportion of flights operated by larger carriers and smaller carriers. That is, whether we look at a larger or smaller airport, most of the flights will be operated by Southwest Airlines (WN) or American Airlines (AA). Otherwise, long delay times for smaller carriers could be related to only flying to certain destinations, such as maintenance issues on a more limited number of airports.

Figure 8. Carrier Delay by Airline



Modeling

Predicting Late Arrivals

For this section of the project, we tried to predict if a flight will arrive late using random forest. The features used are listed below:

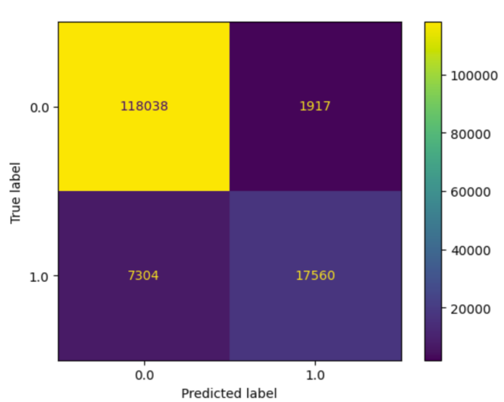
Figure 8. Features used to Classify Arrivals

```
YEAR
MONTH
DAY_OF_MONTH
DEP_DELAY
DISTANCE
CRS_DEP_HOUR
CRS_ARR_HOUR
AIRPORT_FLIGHTS
hurr_period
storm_period
```

Here, we have three new variables. Airport flights tells us information about the number of flights from an origin airport, including airports in and out of Texas. Like before, we must consider that there are disproportionately more flights from Texas in our dataset. Hurricane period is an indicator of whether a flight was scheduled in the period of 1 day before to 5 days after a hurricane landed in an affected state. Storm period is an indicator of whether a flight was scheduled in the period of 1 day before to 5 days after a tropical storm landed in an affected state. These periods only considered days in the same month as the hurricane. The idea is that the presence of a hurricane or storm in an area during a certain period of time may affect delay times.

We used a random forest model because we were interested in the importance of each feature to the arrival delay. Additionally, since the model considers the predictions of multiple decision trees, it is more robust than a simple decision tree. We then created a confusion matrix to evaluate our random forest classification model (**Figure 10)**. Our model incorrectly classified a large number of late arrivals as on-time.

Figure 10. Confusion Matrix of Late Arrival Classification



While our model has a high accuracy score, it has a relatively low recall score. Thus, an airport using this model may waste time and resources preparing for flights to arrive on time because many flights will arrive later than expected.

Figure 11: Classification Report for Late Arrivals

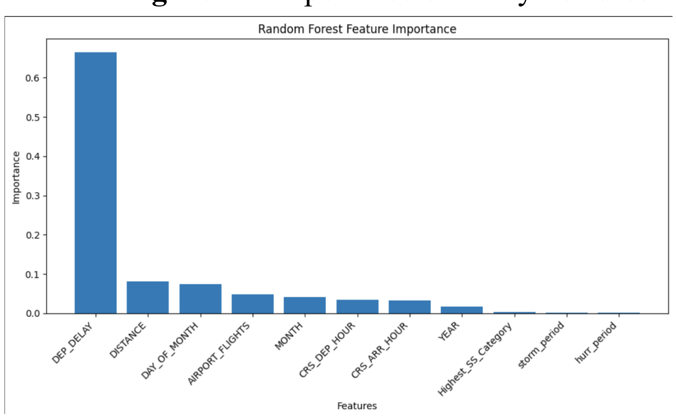
```
Accuracy: 0.94
Classification Report:
              precision    recall  f1-score   support

         0.0       0.94      0.98      0.96    119955
         1.0       0.90      0.71      0.79     24864

    accuracy                           0.94    144819
   macro avg       0.92      0.85      0.88    144819
weighted avg       0.93      0.94      0.93    144819
```

Looking at the feature importance chart, we immediately notice that departure delay is very important in predicting arrival delay. Additionally, it seems that the presence of a hurricane or storm in an area in the specified time period is not a good predictor of arrival delay. This may be because hurricanes and tropical storms are uncommon, so the indicators do not distinguish between many days. Additionally, flights scheduled during a hurricane or storm may be more likely to be canceled rather than delayed.

Figure 12. Importance of Delay Features

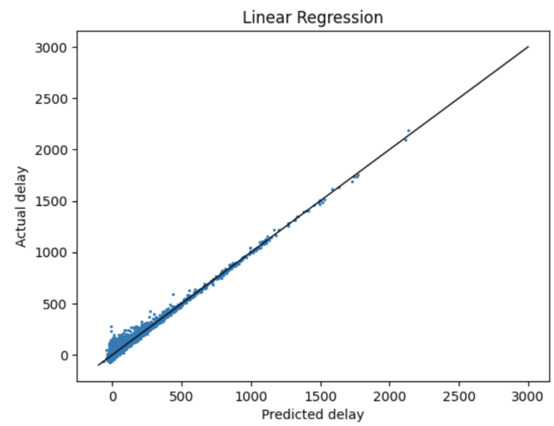


Predicting Arrival Delay Times

Next, we tried to predict the actual arrival delay using linear regression. We see that the model does a better job of predicting delays the longer the actual delay is. In those conditions, it is likely that the departure delay is exceptionally long, causing the arrival delay to also be exceptionally long. As we previously discussed, the model prioritizes the departure delay in predicting arrival delay, so the model works well in these cases.

On the other hand, the model does not predict shorter delays as well. We have a mean squared error of about 168, so there is a large difference between actual and predicted values. Here, it can often be the case that the major component of the delay is not during departure, but during the actual flight. This means that the model may not be able to rely on an exceptional departure delay to inform its prediction.

Figure 13. Actual by Predicted Arrival Delays



Predicting Hurricane-Prone States

Lastly, we wanted to see if flight information could predict which states are hurricane-prone. Based on records from 1851 to 2018, we decided to classify the top 8 states with the most hurricane impacts as hurricane-prone (Heil, n.d.). Depending on our classification of which states are hurricane-prone, the success of our predictive model will vary.

First, we had to make a few adjustments to our dataset. Since we needed to pick a variable to predict, we chose the origin state of the airport. Then, since a majority of our flights were from Texas and we considered Texas a hurricane-prone state, the prediction would not be interesting in these cases. Thus, we only considered flights from outside of Texas to Texas.

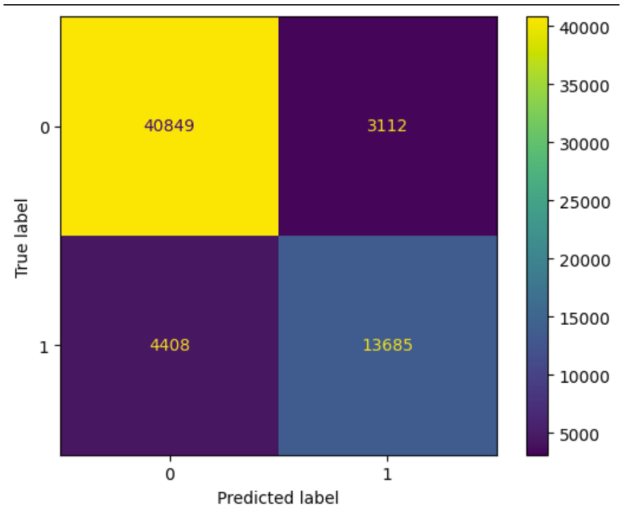
We used the same feature variables as before with the addition of weather delay, arrival delay, and distance group, which tells us information about the distance between two airports.

Figure 14. Features used to Classify Hurricane-Prone States

```
YEAR
MONTH
DAY_OF_MONTH
DEP_DELAY
ARR_DELAY
DISTANCE_GROUP
WEATHER_DELAY
CRS_DEP_HOUR
CRS_ARR_HOUR
AIRPORT_SIZE_GROUP
hurr_period
storm_period
```

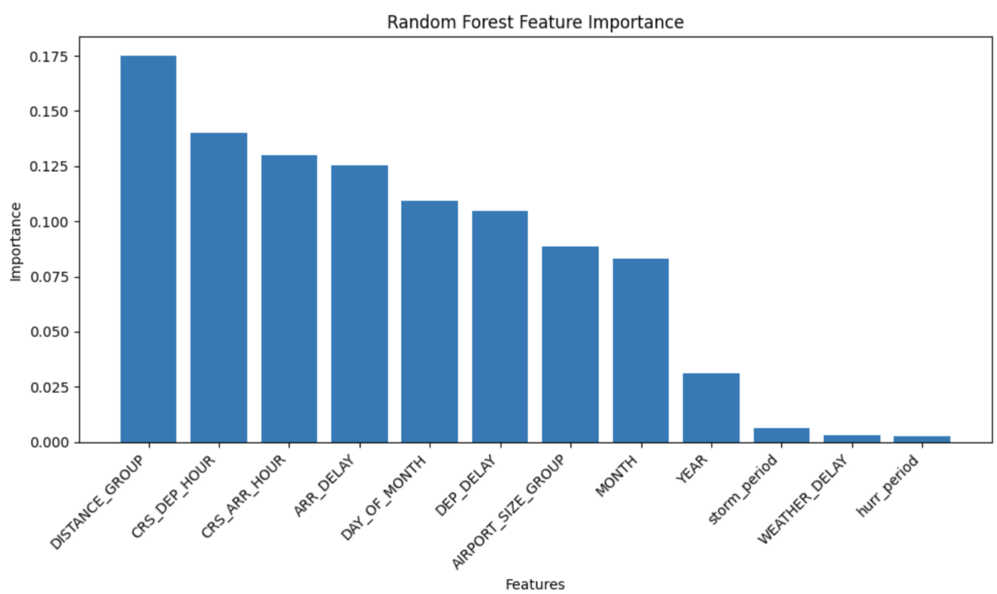
Using a random forest classifier, we see that the model accurately predicted whether the origin state was hurricane-prone in many cases. However, there are still a large number of false negatives.

Figure 15. Confusion Matrix for Hurricane-Prone Classification



Looking at the feature importance chart for this classification, we find that the distance group has the most importance in predicting whether a state is hurricane-prone or not. This makes sense, since many of the hurricane-prone states are close to Texas, so they will all be assigned to a smaller distance group. Interestingly, whether the flight was scheduled in a storm period, hurricane period, or experienced weather delay does not have much importance in predicting whether the origin state is hurricane-prone. It could be the case that while hurricane-prone states experience the worst weather delays, most flights do not experience delays because of weather.

Figure 16. Importance of Hurricane-Prone Features



Discussion

Model Performance

When exploring the relationship between airport sizes and delay times, it was found that larger airports, indicating to greater than 30,000 flights a year, had longer mean departure delays in comparison to smaller airports. **Figure 2** visualizes mean departure delay over time, and visually, we can see a separation between the larger and smaller airports. There are many ways to interpret this result, and thus can lead to future questions of what causes a shorter delay time for smaller airports, as well as what causes a longer delay time for larger airports. Overall, the graph allows a solid visualization of the differences between the two categories.

In addition, we explored what type of delay causes the longest departure time. After creating a density plot to compare the departure delay categories, as seen in **Figure 4**, it was found that weather delays had the longest departure delay time. This was a conclusion that was visually based, and after observing the graph, we found that weather extended the most with the highest density and longer departure delay time. This conclusion helps transition into the next question asked within the analysis. We further explored how weather delays pertain to certain times of the year. For this, the hurricane season was explored, and based on our findings in **Figure 6,** it was deemed that during hurricane months there were far higher percentages in delayed flights. This further analyzes how weather is a factor into delays and how it can impact specific times of the year.

Then, we looked at carrier delay by operating carrier. While we did not find a strong relationship between the size of the carrier and the average carrier delay, we did find that certain carriers experience longer delays than others. This is in contrast to our findings about how the size of an airport affects departure delays. We may expect larger airlines to experience more traffic and thus more operational difficulties than smaller carriers, but this is not necessarily the case.

To better understand the claims of our exploratory analysis, we performed a random forest classification followed by a confusion matrix seen in **Figure 10**. This matrix helped confirm how we can predict if a flight arrives late, and if so, what categories have the highest importance involved in prediction. These models help further motivate the ideas behind our data analysis, and confirm that there is a relationship present between variables of a flight and delay times. In particular, the departure delay of a flight seems to be an especially important feature in predicting whether a flight will arrive late or not. Additionally, we find that whether or not a flight is scheduled during a storm or hurricane period has little impact on whether there is an arrival delay.

As for predicting whether a state is hurricane-prone, we found that our model was somewhat successful, but there were still a large number of misclassifications. The most important feature in predicting whether the origin state of a flight is hurricane prone was the distance group, which makes sense because most hurricane-prone states are close to Texas, causing flights from hurricane-prone states to be placed in a smaller distance group. However, it seems that weather delay and the presence of a hurricane or storm near the time of a flight has little importance in predicting whether a state is hurricane-prone or not. This could be because most flights are not during periods with hurricanes, so these features do not make much distinction between hurricane-prone and non-hurricane-prone states in most cases.

Limitations

Our models held a variety of limitations regarding the possible accuracy and claims. When setting up the predictions to our exploratory analysis, it is likely that some features may inadvertently reflect whether the delay has already occurred. As an example, when forecasting delays for a flight that already landed, we end up with features that would only be present post-departure. This becomes unsuitable for our predictive model, causing inability to form real-time predictions. Initially, we considered causes of delay like weather and security as potential feature variables. However, since we cannot separate delay causes into components that affect the departure and arrival, we could not use these variables as features in our predictive model. If our goal is to predict arrival times for the efficiency of airports, then these features are not appropriate to include in our model because they could be telling us something about the arrival delay that we did not know at the time of departure. When predicting the duration of a delay, we ran into a similar problem.

Additionally, we see that our model for predicting arrival delays has a strong dependence on departure delay times. Looking at the predicted against actual arrival times in Figure 12, we see that our model only makes very accurate predictions when the arrival delay is long, and these long delays are almost assuredly affected by long departure delays. There may be other features that can increase the accuracy of our model. Otherwise, for classifying hurricane-prone states, we find that the most important feature is the distance group between the origin and destination airports. This means that our model is limited to Texas flights since these distances are only relevant for Texas airports.

There are also limitations present in **Figure 6** and **Figure 7** regarding the relationship between specific times of the year and weather delays. It may have been an ineffective approach to focus on only California and Florida to form a conclusion. We know that Florida experienced a harsh hurricane season in 2022, but we do not know if we will find the same effect every year or if other hurricane-prone states follow a similar pattern.

In the case of classifying late arrivals and hurricane-prone states, we found that the categorization of flights into groups characterized by whether or not a hurricane recently landed was not an important feature. Here, we used a time period of 1 day before to 5 days after the hurricane or storm landed as the period of time for flights to be affected by a hurricane or storm. Perhaps for smaller and larger storms, we could have made a better distinction. Additionally, we only considered days within the same month, so if a hurricane or storm landed at the end of a month, the first days in the next month would not be considered. Additionally, there is also the possibility of hurricanes causing flights to be canceled rather than delayed, so we cannot explore the impact of hurricanes on flight departures as deeply as we would like. Finally, we did not consider the localized area of impact that a hurricane or state would have in a particular state. Especially in the case of hurricanes or storms with weaker winds, it makes sense that only a few airports rather than the entire state are affected.

Ethics Framework

Airlines are networks of people, working to provide a mass transportation service to more people. They are incapable of functioning without the input of airline employees and partnerships, whom they are in part responsible for. We must therefore approach our research responsibly, starting with anticipating and reflecting upon the possible implications of our observations. By assessing and comparing the delay times of large and small airports in Texas, and observing the carrier delay of airlines, we are inadvertently creating a perception of the management and efficiency of these airlines. It should be noted that this was not the goal of our project; our goals were to investigate trends and relationships causing flight delays with our dataset. While carrier delays are in part representative of an airline's organization, flight delays should not be the only indicator of whether an airline is well-organized or efficient. With this, our project may have the **unintended consequence** of creating a false perception of what airlines are the "best" and "worst" to the public eye. This misinterpretation of the data by the public could drive customer choice, therebying posing **project risks** and possibly imposing on **stakeholder involvement**. Our project may put airline employees of seemingly lower-performing airlines (those that experience less flight delays, including carrier delays, on average) at risk of losing their job, especially for those in the line personnel department. Should there be an influx of customers to airlines with seemingly higher-performing airlines (those that experience more flight delays, including carrier delays, on average), it is possible that airline alliances and airline loyalty programs may become dismantled and stretched thin as both are dependent on customer engagement. To avoid the unintended consequence of creating a false perception of airline efficiency, there should be a greater degree of **openness** in future analyses of flight delays. Our project's observations and attempts to predict flight delay should be coupled with data pertaining to weather forecast, airline crew availability and quality of work, and aircraft maintenance. Doing so would be a more transparent approach to addressing the growing frequency of flight delays not only in Texas but also across the U.S..

Conclusion

Main Findings

Overall, we found that larger Texas airports have increased mean departure delay, which may be due to the increased volume of people and moving parts that go into the operations of a large airport. We also found that delays caused by weather are on average longer than other delays and weather delay frequency can be specific to certain months. Specifically, hurricane season sees a much higher rate of weather delays in Florida. These findings make sense considering weather is an uncontrollable phenomenon that cannot be avoided through improvement or optimization. While airport management and staff may be able to improve in other areas of delay such as carrier delay, improving on weather delays would be much more logistically difficult. With that said, mean weather delay is not indicative of being hurricane prone, most likely because there are many other weather events that can cause delays which may be influencing the prediction.

Next Steps

Possible continuations of this study could incorporate additional weather and climate data to see the impact of hurricane season on flight delays through multiple years. Future study could also benefit from controlling for confounding weather events through methods such as GLMs or multilinear regression. It would also be advantageous to expand the analysis to include flights arriving in and departing from states besides Texas to compare findings, compare data between specific airlines, or compare airline data with data concerning other public modes of mass transit.

Acknowledgements

Contribution Breakdown	Team Member	Contribution	Total Contribution	Contribution Scaled
• Data Preparation → 15% • Graphs → 15% • Machine Learning → 15% • Presentation Prep→ 10% • Presentation Delivery → 20% • Report Writing → 15%	**Devin Xiao**	**Data Preparation: 3.75% Machine Learning: 6% Graphs: 6% Report Writing: 1.2%**	**16.95**	**100**
	Feifei Cao	**Presentation Prep: 3% Presentation Delivery: 5% Report Writing: 4.95% Machine Learning: 1.8% Graphs: 1.8%**	**16.55**	**97.64**
	Joy Richardson	**Data Preparation: 3.75%**	**16.85**	**99.41**

● Website Development → 10%		**Presentation Prep: 3% Presentation Delivery: 5% Report Writing: 1.5% Machine Learning: 1.8% Graphs: 1.8%**		
	Sofia Bautista	**Presentation Delivery: 5% Data Preparation: 3.75% Machine Learning: 1.8% Graphs: 1.8% Report Writing: 2% Presentation Prep: 2.2%**	**16.55**	**97.64**
	Vedika Shirtekar	**Presentation Delivery: 5% Data Preparation: 3.75% Presentation Prep: 1.8% Machine Learning: 1.8% Graphs: 1.8% Report Writing: 2.4**	**16.55**	**97.64**
	Yasmeen Luna	**Website Development: 10% Report Writing: 2.95 Machine Learning: 1.8% Graphs: 1.8%**	**16.55**	**97.64**

References

1. 2023 Atlantic Hurricane Season Wraps Up. (2023, November 28). National Environmental Satellite, Data, and Information Service. https://www.nesdis.noaa.gov/news/2023-atlantic-hurricane-season-wraps
2. Continental United States Hurricane Impacts/Landfalls 1851-2022. (2023, April). *National Oceanic and Atmospheric Administration.* https://www.aoml.noaa.gov/hrd/hurdat/All_U.S._Hurricanes.html
3. Continental United States Tropical Storms Impacts/Landfalls 1851-1970, 1983-2022. (2023, April). *National Oceanic and Atmospheric Administration.* https://www.aoml.noaa.gov/hrd/hurdat/uststorms.html
4. Full Year 2022 U.S. Airline Traffic Data. (2023, March 16). *Bureau of Transportation Statistics*. https://www.bts.gov/newsroom/full-year-2022-us-airline-traffic-data
5. Heil, J. (n.d.). 10 U.S. States Where Hurricanes Hit Most Often. *Universal Property & Casualty Insurance Company.* https://universalproperty.com/united-states-where-hurricanes-hit-most/
6. Potter, K. (2023, March 17). Report: Flight Disruptions in 2022 Were Some of Worst in Years. *Thrifty Traveler*. https://thriftytraveler.com/news/airlines/2022-flight-delays-cancellations/
7. Record-breaking Atlantic hurricane season draws to an end. (2021, June 10). *National Oceanic and Atmospheric Administration.* https://www.noaa.gov/media-release/record-breaking-atlantic-hurricane-season-draws-to-end
8. Saffir-Simpson Hurricane Scale. (n.d.). *National Weather Service*. https://www.weather.gov/mfl/saffirsimpson