

US Domestic Air Travel Trends & Data Analysis



Presentation by Isaac Madera, Raymond Hung and Aashray Puri

4/20/2023

Intended Audience



Datasets & Guiding Questions

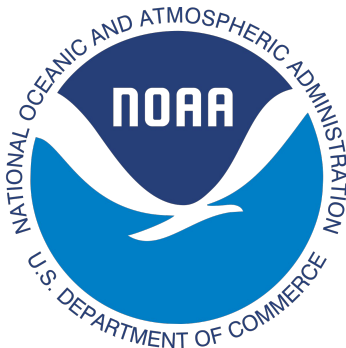
Primary Dataset:

- US Domestic Flight Data - Kaggle/USDOT

Supplementary Datasets:

- Airline Names - Kaggle
- Weather Data - NOAA
- Airport Geographical Information - Kaggle/NOAA

kaggle



1. Competitive landscape:

- How did COVID impact flight volume?
- Which airline has the most flights in the US?
- What are the busiest airports in the United States?

2. Delay landscape:

- Which airline has the fewest delayed flights in the US?
- Which airports are most likely to experience a delay?
- In which month/day/hour of day are delays most common?

3. Weather data:

- Which airports are more prone to experiencing weather events (rain, snow, thunderstorms, fog)?
- What is the impact of thunderstorms and snowstorms on flight operations?

Data Cleaning & Initial Exploration

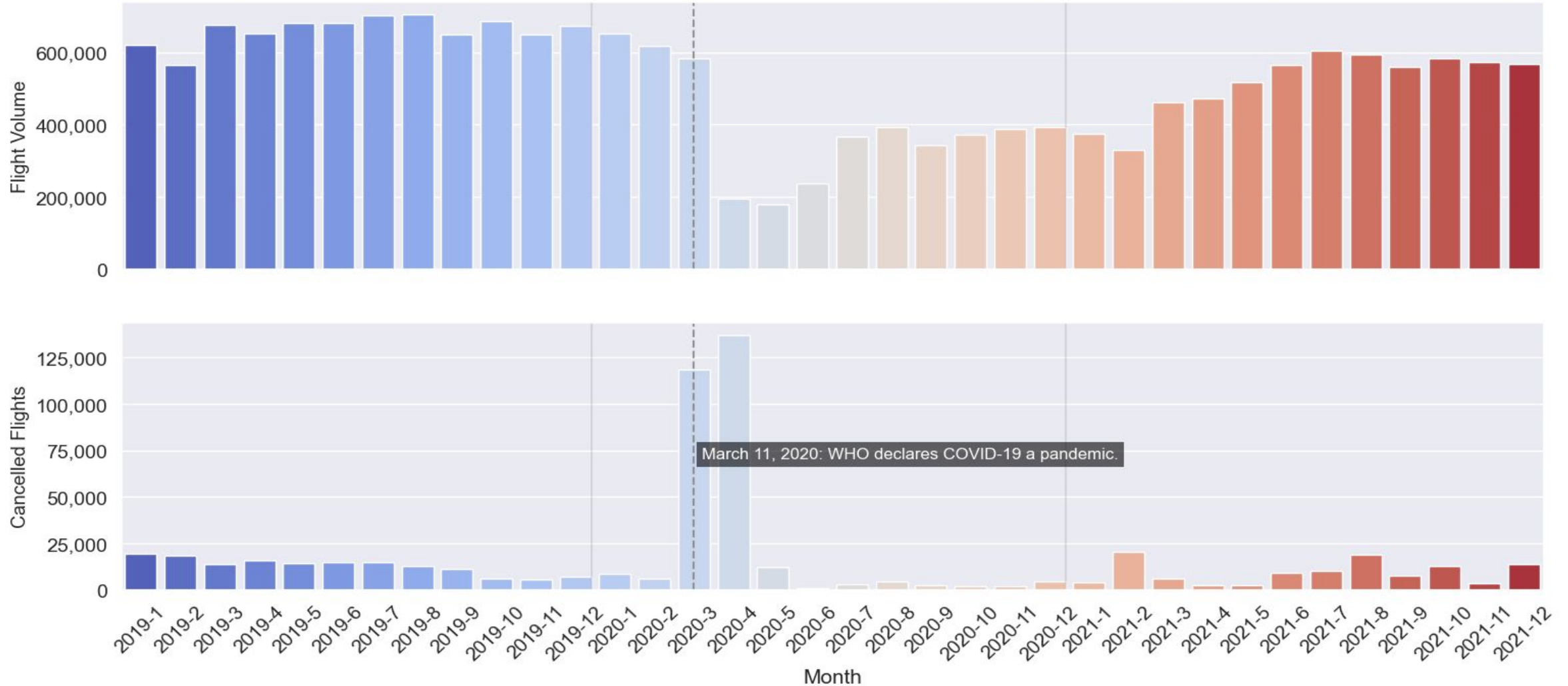
- Shape Check for consistency
- Column name and dtype check
- Grouping 2019-2021 data into 1 dataframe (combined_df).
- `combined_df.duplicated().sum()`
- `combined_df.isnull().sum()`
- `combined_df.describe().T`
- Dropping 41 columns and focusing on 20 columns (including new index)
- Merging combined_df with other datasets
- Making a subset of combined_df focusing on 2021



Data Cleaning

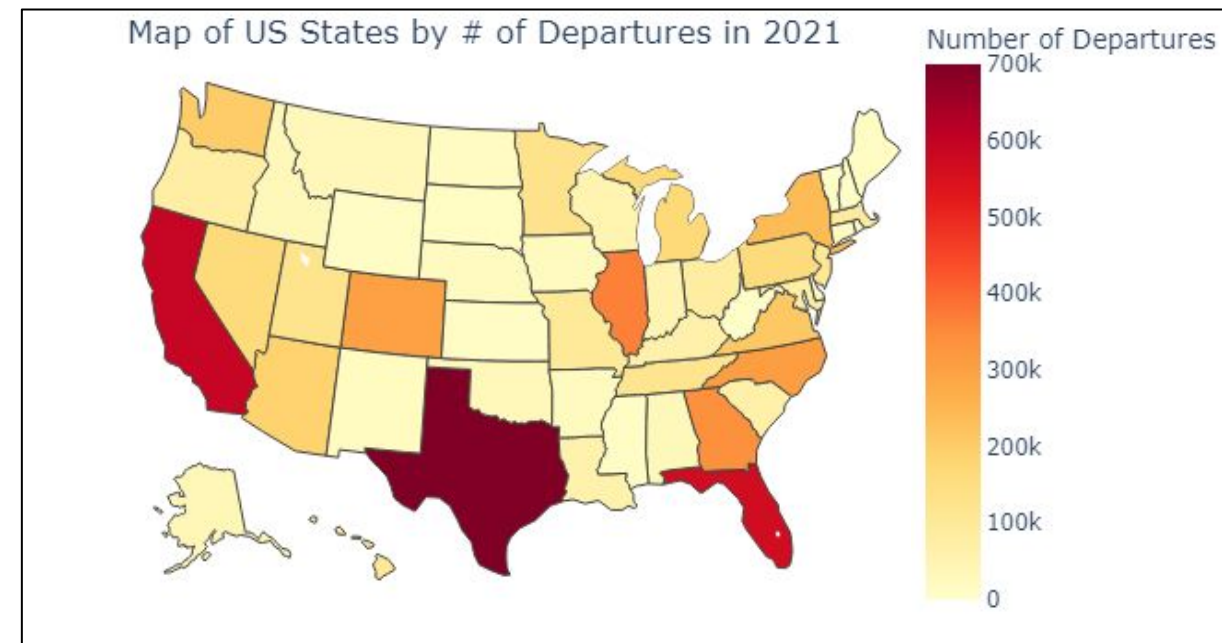
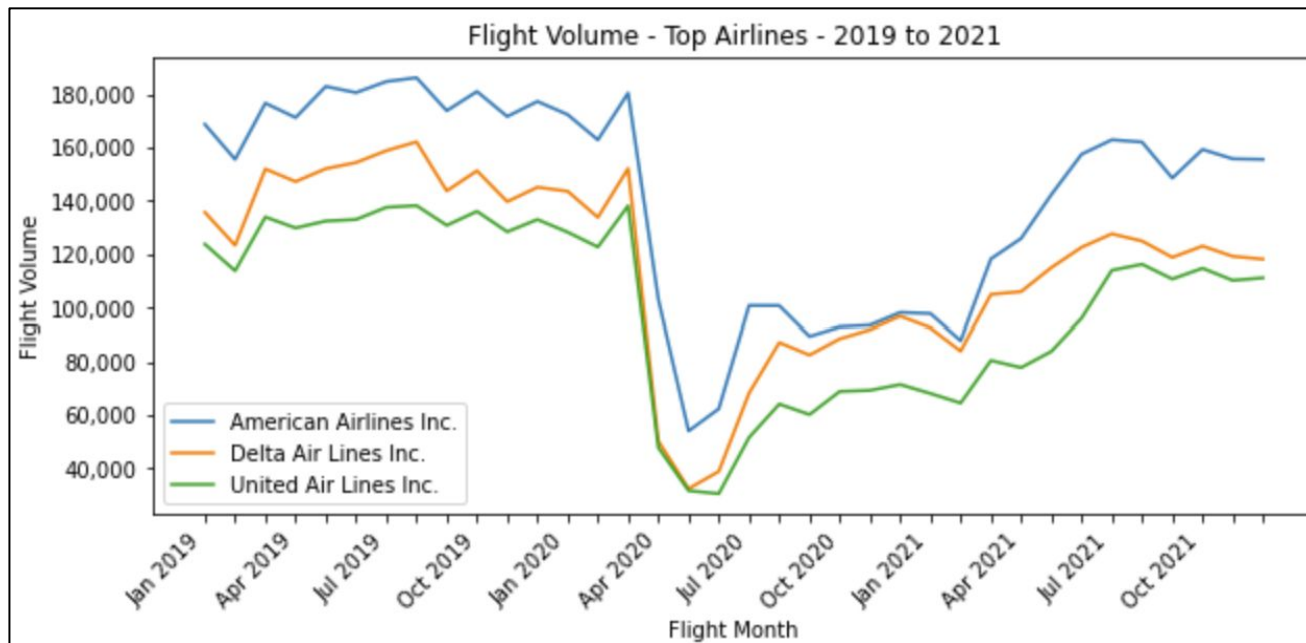
How did COVID impact Flight Volume?

US Operating Flight Volume vs Cancellations
2019-2021



Which airline has the most flights in the US?

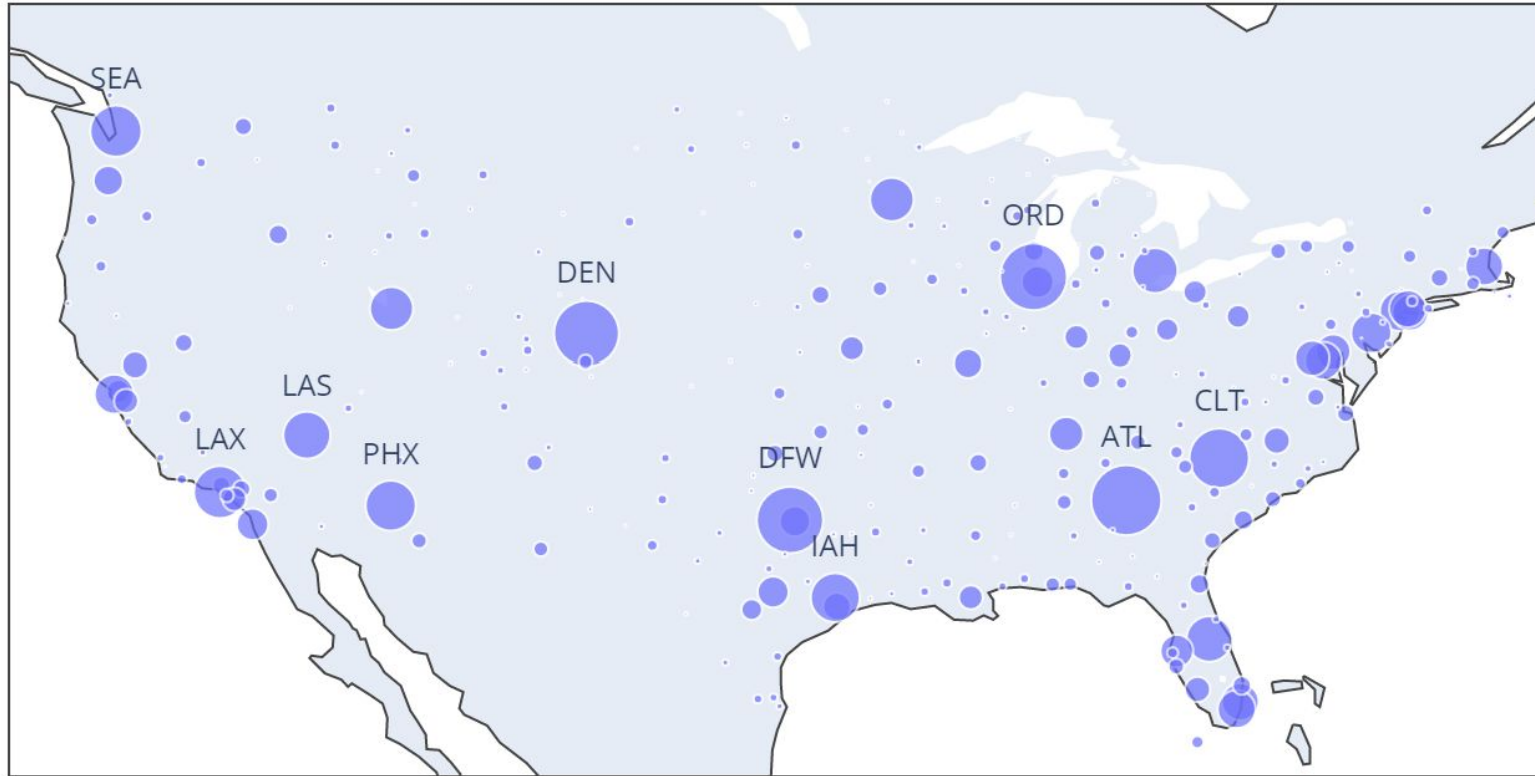
	American Airlines Inc.	Delta Air Lines Inc.	United Air Lines Inc.	Southwest Airlines Co.	Alaska Airlines Inc.	JetBlue Airways	Spirit Air Lines	Frontier Airlines Inc.	Allegiant Air	Hawaiian Airlines Inc.
Flight Volume	5,096,872	4,190,263	3,605,095	3,389,862	1,085,859	644,276	531,308	363,860	319,675	198,882



- All 3 major legacy airlines saw Flight Volume fall because of the COVID-19 pandemic.
- American Airlines saw the Highest Flight Volume 2019-2021, despite COVID.
- In 2021, Texas, California and Florida saw the highest number of flight departures

What are the busiest airports in the US?

US Airports by 2021 Flight Volume (bubble size) - Top 10 Labeled



- In 2021, **ATL (Atlanta)** was the busiest airport in the US
- A lot of traffic is concentrated on the **East** side of the country
- The size of some airports is largely because of airlines' strategic decision on where to set up their base
- Table below provides insights into airlines claiming their territory (**'hubs'** in the airline industry)

Strategic airline hub placement- A look at Flight Volume by airline & airport:

Origin	ATL	ORD	DFW	DEN	CLT	LAX	SEA	PHX	IAH	LAS
MajorAirline										
American Airlines Inc.	12,351	112,188	242,784	9,962	205,427	37,862	6,487	75,465	8,806	12,333
Delta Air Lines Inc.	246,399	12,188	9,889	11,487	8,576	43,770	42,620	11,467	7,537	13,749
United Air Lines Inc.	6,057	136,074	7,674	141,561	4,739	28,274	6,679	8,273	120,265	9,910

Which airline has the fewest delayed flights in the US?

	MajorAirline	2021 Avg Delay %
1	Hawaiian Airlines Inc.	8.80
2	Delta Air Lines Inc.	10.99
3	Alaska Airlines Inc.	13.24
4	American Airlines Inc.	15.93
5	United Air Lines Inc.	16.26
6	Spirit Air Lines	20.25
7	Frontier Airlines Inc.	21.81
8	JetBlue Airways	26.23
9	Allegiant Air	26.47
10	Southwest Airlines Co.	26.86

Bottom 3 Airlines

Southwest[®]

allegiant[®]

jetBlue

Top 3 Airlines

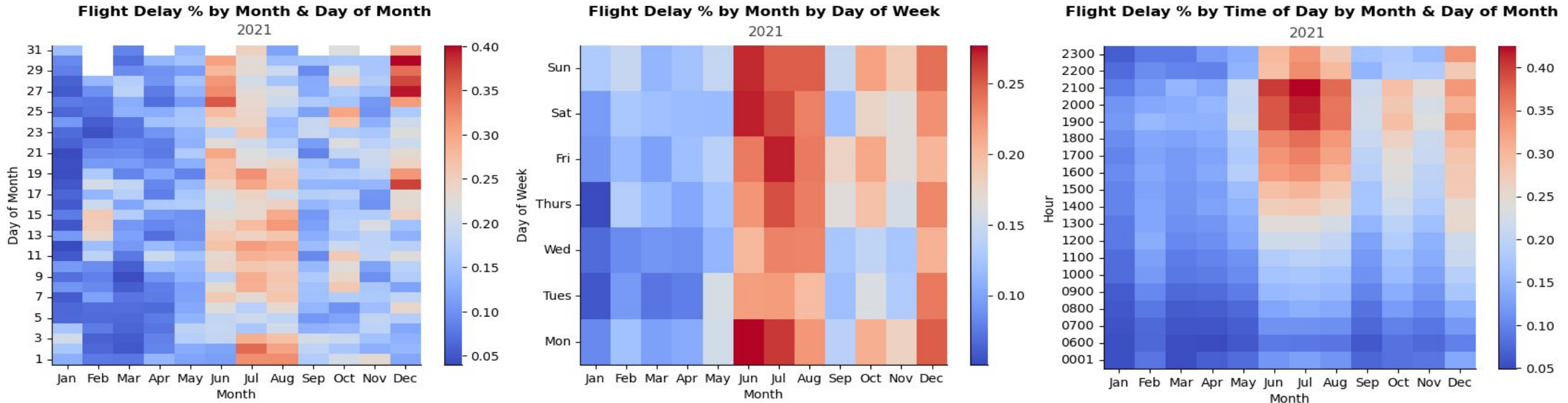
HAWAIIAN[®]
AIRLINES[®]

 DELTA

Alaska[®]
AIRLINES[®]

- Flight Delay % = (# of Flight Departures that were > 15 mins Delayed) / (Total # of Flight Departures)
- Low cost airlines had the worst flight delay % in 2021 among all airlines, while legacy carriers performed above-average

In which months/day/hour of the week are delays most common?



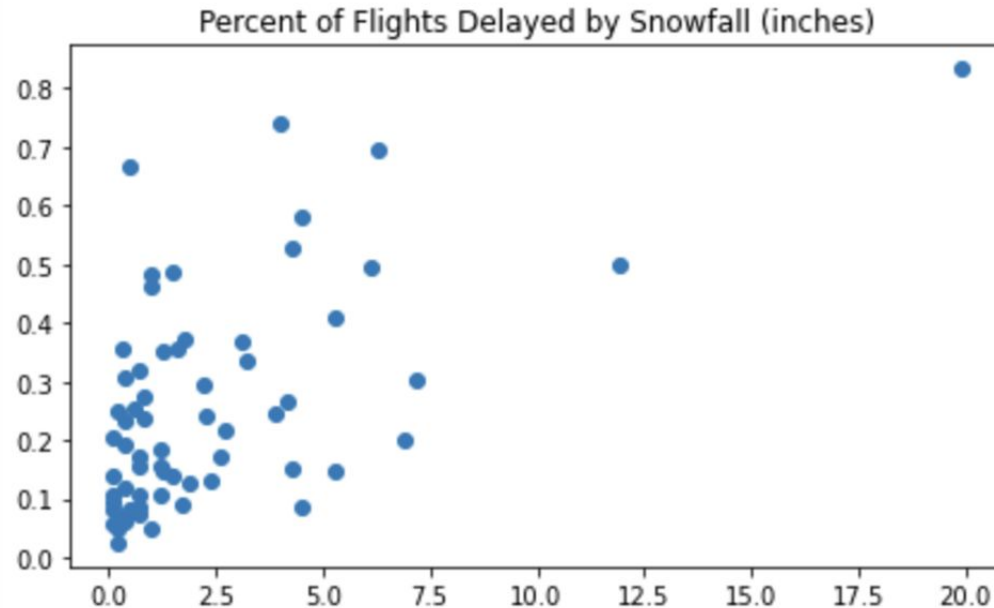
- **Summer months** had the highest Flight Delay % overall
- The worst days to fly in 2021 were **Fridays, Saturdays, Sundays, and Mondays**
- The worst week to fly in 2021 was **between Christmas and New Years**
- The worst time to fly during Summer 2021 was during the **late afternoon - evening hours**

Which airports are more prone to experiencing weather events?

airport	ATL	ORD	DFW	DEN	CLT	LAX	SEA	PHX	IAH	LAS
rain_flag	0.37	0.34	0.27	0.18	0.35	0.10	0.44	0.13	0.39	0.11
t_storm_flag	0.15	0.11	0.16	0.10	0.13	0.01	0.02	0.07	0.21	0.05
snow_flag	0.00	0.13	0.01	0.13	0.01	0.00	0.03	0.00	0.01	0.00
freeze_flag	0.00	0.02	0.01	0.05	0.00	0.00	0.01	0.00	0.01	0.00
fog_flag	0.06	0.04	0.04	0.08	0.07	0.11	0.03	0.01	0.16	0.00

- We leveraged airport weather data to understand likelihood of various conditions in the data
- **ORD (Chicago) and Denver (DEN)** have the most volatile weather, prone to both snowstorms and thunderstorms.
- Rain is not necessarily indicative of delays as much as thunderstorms (see Seattle)

What is the impact of snowstorms and thunderstorms on flight operations?



- The scatterplot shows a **positive correlation** between snow (inches) and flight delays.
- Most airports are severely impacted by thunderstorms. Often **2-3x more delays when there are thunderstorms**. Only West coast airports (LAX, SEA) are not significantly affected

Flight Delay % on days with Thunderstorms vs Without

Origin	ATL	ORD	DFW	DEN	CLT	LAX	SEA	PHX	IAH	LAS
t-storms	0.25	0.34	0.35	0.42	0.27	0.23	0.17	0.31	0.31	0.37
no t-storms	0.12	0.15	0.19	0.22	0.11	0.17	0.14	0.19	0.14	0.23

In Conclusion....

If you are a consumer who wants to maximize odds of not getting delayed:



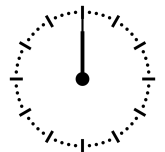
Book a flight for Q1 or early Q2!

Avoid holiday travel!



Avoid low-cost airlines!

Stay away from Denver & Chicago!



Fly before Noon!

Summary of Findings

Top 3 Airlines by Flight Volumes 2021: American Airlines, Delta, United

Top 3 States by Flight Volumes 2021: Texas, California & Florida

Top 3 Airports by Flight Volumes 2021: ATL, ORD, & DFW

Top 3 Airlines with least Flight Delay %: Hawaiian Airlines, Delta Airlines, & Alaska Airlines

Bottom 3 Airlines with most Flight Delay %: Southwest Airlines, Allegiant Air, JetBlue Airways

Top 3 Airports with lowest Flight Delay %: ATL, CLT, SEA

Bottom 3 Airports with the highest Flight Delay %: DEN, LAS, DFW

Best Month to Fly 2021: January

Best Days of the Month to Fly 2021: 4th/5th

Best Days of the Week to Fly 2021: Tuesday/Wednesday

Best Hour of the Day to Fly 2021: 6 am

Airports in 2021 most impacted by poor weather: Denver/Chicago

Impact of Snow & Thunderstorms on Flight Operations: Both delay Flights, but Thunderstorms are worse than Snow

Works Cited

1. CAPA - Centre for Aviation. "US Big 3 Airlines Take Varying Approaches to Recovery." CAPA - Centre for Aviation, 10 Nov. 2021, <https://centreforaviation.com/analysis/reports/us-big-3-airlines-take-varying-approaches-to-recovery-581256>.
2. Bonnie, Jin. "Airport Code and Geographical Information." *Kaggle*, 5 Nov. 2020, <https://www.kaggle.com/datasets/jinbonnie/airport-information>.
3. "Flight Status Prediction." *Kaggle*, <https://www.kaggle.com/datasets/robikscube/flight-delay-dataset-20182022?resource=download>.
4. Girl, Tensor. "Kaggle Global Trends." *Kaggle*, 18 Apr. 2021, <https://www.kaggle.com/datasets/usharengaraju/kaggle-global-trends>.

Image Links

<https://www.rd.com/article/how-high-do-planes-fly/>

<https://www.houstoniamag.com/travel-and-outdoors/2018/07/flight-delay-tips>

<https://www.istockphoto.com/photo/man-waiting-for-flight-in-airport-lounge-gm862026202-143149629>

<https://www.kaggle.com/>

<https://www.noaa.gov/>

<https://www.transportation.gov/>

<https://www.vecteezy.com/vector-art/6094871-data-cleaning-in-flat-outline-icon-editable-vector>