U.S. Aviation Trends in 2019-2020

CS3200 Final Project Report

William Hanvey, Ngoc Khanh Ve Le, Laura Nguyen, Triet Nguyen, Dustin Yang

June 24, 2021

1. Introduction

According to a News Article by the BBC, airlines around the world were set for the 'worst' year on record when the Coronavirus first hit. Our project seeks to explore what U.S. aviation trends existed both domestically and internationally during the years 2019-2020 to evaluate the true effects COVID-19 had on the country's air travel.

For our database, we gathered air travel data from the Bureau of Transportation Statistics (BTS). We combined it with COVID-19 data from both the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO). Finally, we used data from the UN's World Tourism Organization (UNWTO) to summarize the total arrival count into various countries within the years 2007-2019.

What actually happened to the U.S. aviation market when COVID-19 restrictions began? How did domestic aviation trends compare to those internationally? Through analysis of our created database, we hope to provide significant insights into what actually happened in U.S. aviation when the world maintained widespread lockdowns due to the pandemic.

2. Database Design
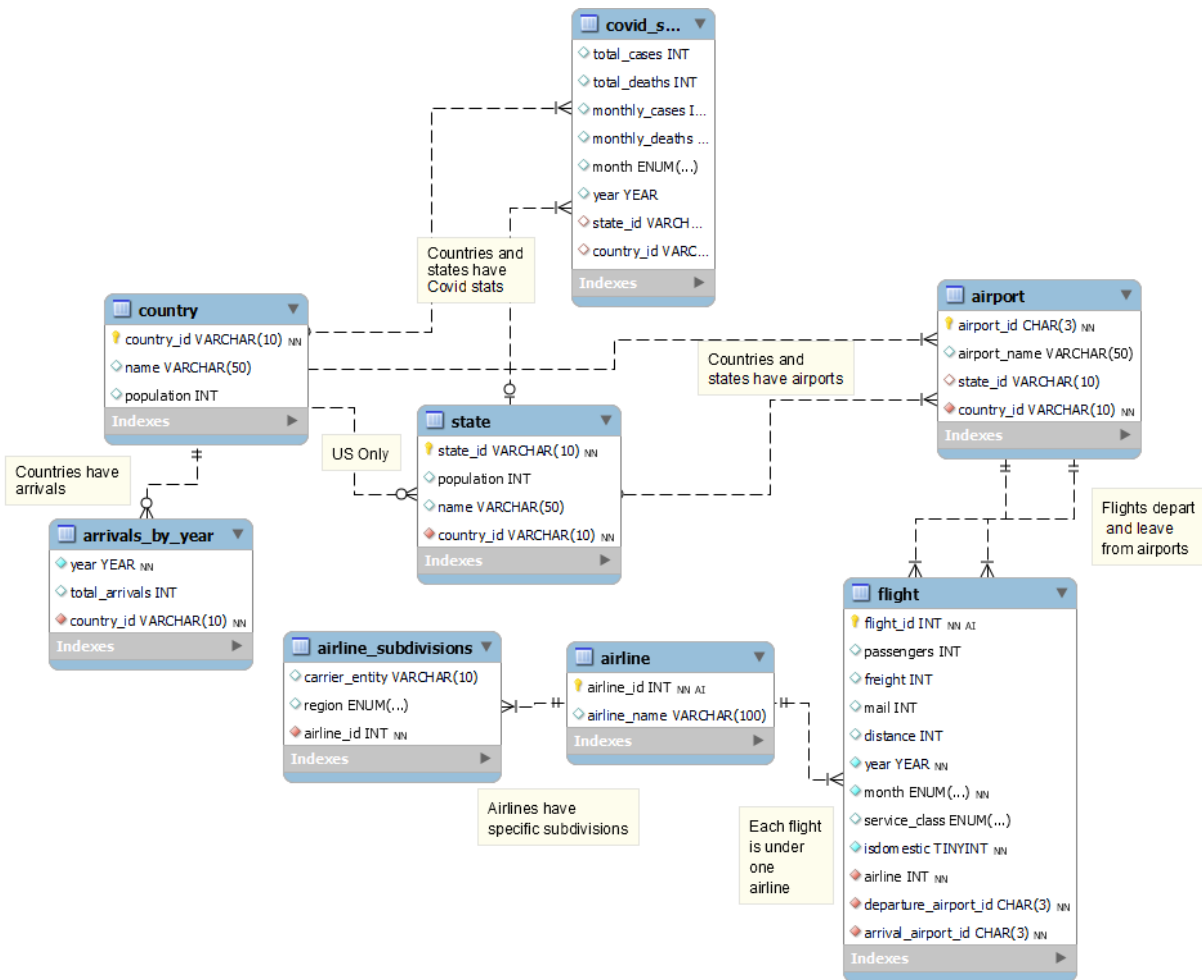
EER Model Diagram

3. Methods / Use Cases:

Our database model was constructed with four primary data sources detailed below. Because some sources had more information than was needed, we used Python pandas and Excel sheets to aggregate it and remove unnecessary elements. For importing into the database, we used both the Import Wizard and SQL 'LOAD DATA' command statements. All database details and files were locally standardized across the group to maintain data integrity and enable simultaneous remote collaboration.

The COVID-stats table and attributes were populated with data from the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO). A similar process was used for the arrivals_by_year table based on data from the UN's World Tourism Organization (UNWTO), which summarized the total arrival count into various countries within the years 2007-2019. These tables served as supplementary

staging tables for our analysis. Additional info on name abbreviations and population for countries and U.S. states were consolidated into referential entities, with data provided by external CSV sources.

For aviation-related tables, such as airports, airlines, and flights, we first created a primary staging table using air travel data from the Bureau of Transportation Statistics (BTS), specifically the T-100 Domestic Market and T-100 International Market files in 2019 and 2020 (altogether, four files). With this staging table, we could then easily populate the related attributes for these aviation entities (i.e., origin airports, destination, passenger & freight count, etc.).

With all the data imported in, we then analyzed the database using SQL queries to look at aviation trends during COVID. We used Excel and Python to generate visualizations based on the results. For a use case example, we queried our database to find out airports with the highest passenger volume, retrieved their volume change during the pandemic and their states, then used that state info to aggregate COVID-19 cases data in order to find a correlation between case count and passenger count.

4. Changes to the Data

While selecting from the staging table to insert data into our aviation-related tables, we found that there were some inconsistencies in the data that needed updating. For example, there was an airport_id listed as 'XWA' that belonged to two airports- one in North Dakota and one in Alaska. However, we realized that this was an error and subsequently changed the airport info for the Alaskan airport to 'GSZ,' as the data was actually referring to Granite Mountain Air Station.

5. Analysis and Visualizations

COVID Cases

The amount of COVID-19 cases for the states in the U.S. that had the most passenger arrivals/departures in 2019 or 2020 are graphed in the figure below. January 2020 marked the start of the spread of COVID-19 in the U.S., and several state lockdowns and mask mandates being introduced around March - April 2020 helped slow the spread of the disease at the start of the pandemic. However, starting May - July 2020, there was a large spike in the number of cases and again in October - December 2020. This might be attributed to a lot more people traveling during the holiday season. In combination with fewer holidays and traveling at the start of 2021, the introduction of vaccines in multiple waves may have caused the number of cases to drop significantly starting February 2021 and again in April 2021 when the 3rd wave of the vaccine was released to the public.
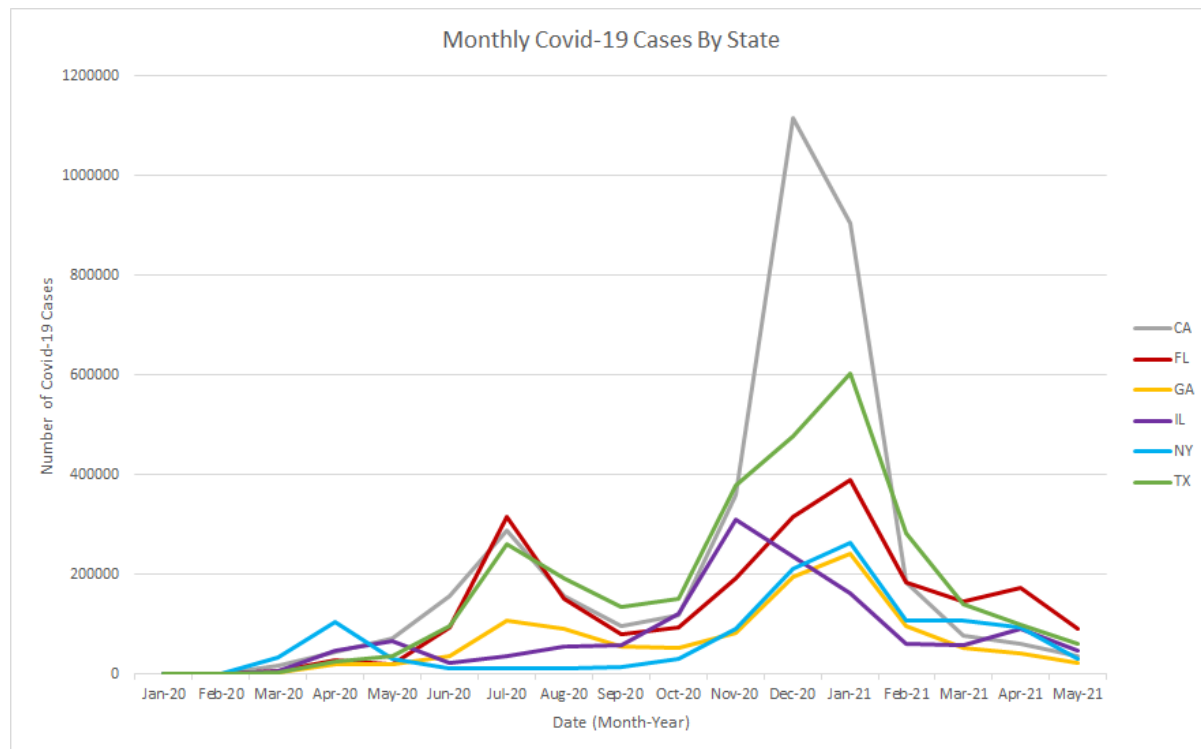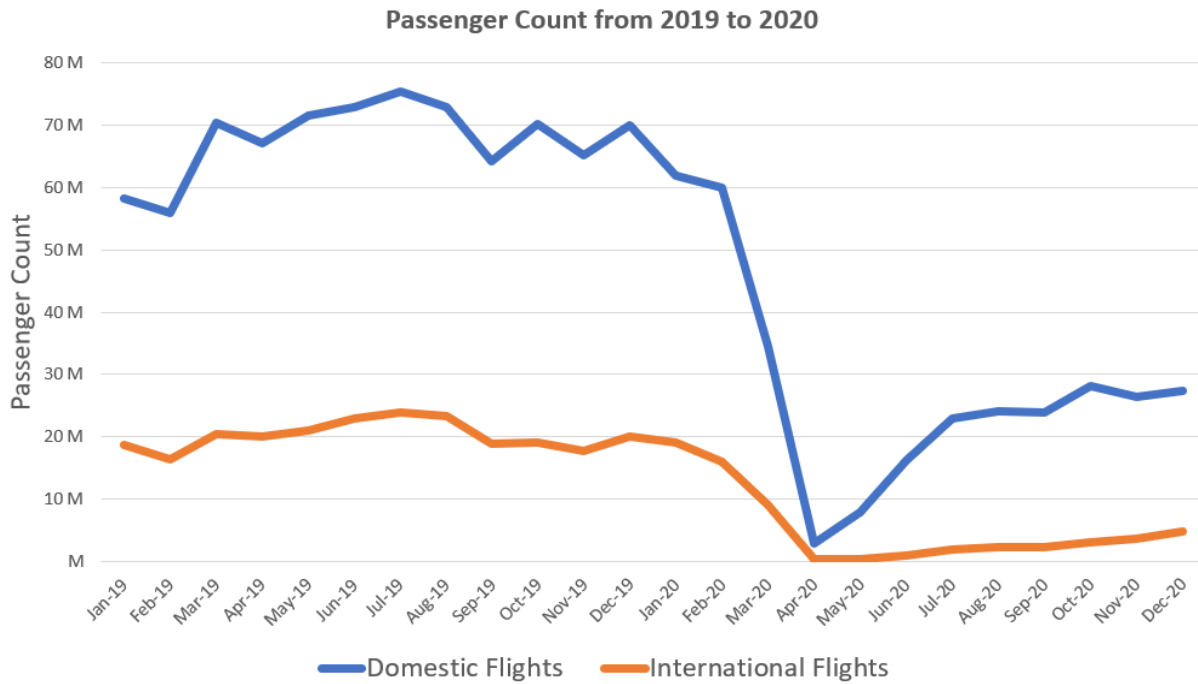
Passenger Count Trends

There was a significant drop in the number of monthly passengers domestically and internationally pre- and post-lockdown. Starting in February until April 2020, there was a drop in the number of monthly passengers to an all-time low of <5 million monthly passengers in the 21st century. The reason for this large decrease was because many states in the U.S. and other countries began locking down and stopping flights starting March 2020 due to the rapid spread of the virus. The monthly domestic passenger count in the U.S. rose at a steady but quick pace from May - July 2020, which reflected a similar rise in the number of COVID-19 cases during that time.

Something to note is that at the height of the second spike in COVID-19 cases around July 2020, the rate of the number of monthly passengers fell and plateaued in the following months of July - September 2020. However, in October - December 2020, the number of passengers began to rise, and the number of cases of COVID-19 rose shortly thereafter, beginning the third spike of COVID-19 cases. This suggests some correlation between the number of passengers flying and the number of COVID-19 cases.

In contrast to the rapid rise of domestic flights from June to December 2020, there was a slight increase in international passengers. The number of international passengers decreased because some countries closed their borders, and the U.S. banned flights from some countries.

4

Passenger Count from 2019 to 2020

Passenger Count vs Departure Count

Airline Name	% Decrease of Passengers	Airport ID	% Decrease of Departures
Southwest	-58.34	EWR (NJ)	-39.05
Delta	-66.05	JFK (NY)	-33.54
American	-57.80	MSP (MN)	-31.80
United	-67.41	IAH (TX)	-28.53
JetBlue	-66.60	LAX (CA)	-27.16
SkyWest	-52.00	ATL (GA)	-26.97
Alaska	-65.44	DEN (CO)	-25.55
Spirit	-45.94	ORD (IL)	-25.33
Frontier	-50.62	MIA (FL)	-14.36
Republic	-53.44	DFW (TX)	-11.01

We wanted to analyze whether there was a correlation between passenger numbers and the number of departures. The table on the left shows how much the top 10 airlines in 2019 were affected during 2020. United had the highest reduction in passengers, at 68%. On the other hand, the table on the right shows how much the top 10 U.S. Airports in 2019 were affected during 2020. 2020 clearly had a negative impact on how many departures airports had during the year. The values on the right table are essentially half of the values on the left side. In other words, the ratio of departures to passengers became significantly uneven during COVID. Flights were still likely to take off even with a great reduction in passengers.

This can largely be attributed to cost. Spaces for airlines to take off are pre-bought, so even with fewer passengers, there is a necessity to take off even if they won't be making nearly as much money during 2019 or if it doesn't seem cost-efficient. Airlines spend most of their money upfront, which means flying a mostly empty plane is still worthwhile to try and recover some of their costs.

Freight Volume Trends

While airlines struggle to maintain their passenger volume at profitable levels after the lockdown began, air freight volumes show virtually no sign of change - in fact, air shipping volumes were higher year-over-year when looking at the months during which people had to stay home. This phenomenon is likely caused by increased consumer demand - people were purchasing office supplies for working from home, workout machines for home gyms, or just generally buying things out of boredom.


Monthly US Inbound Air Freight from Top 5 Countries - 2019 to 2020
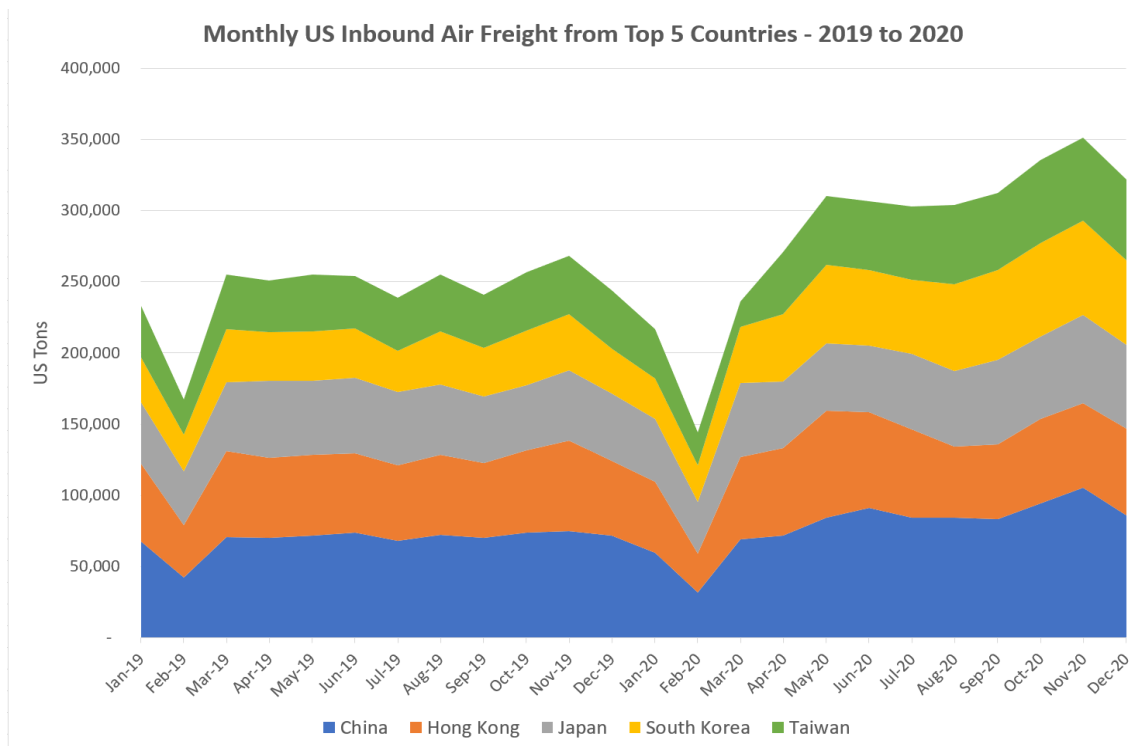
Correlational Data

Unfortunately, there was not as huge a correlation as we were hoping to find. There was a weak positive correlation between COVID cases and the percent change in flights for domestic flights and a very weak negative correlation for international flights. For U.S. states, states with a lower decrease in flights were more likely to have a higher number of COVID cases, which can make sense if you assume people less worried about COVID would be more likely to travel.

Internationally, however, this is the opposite- countries with a smaller decrease in the number of flights were more likely to have a higher number of COVID cases,

which, again, can make sense if you assume countries with lower amounts of flights were less likely to have COVID as an accidental import. However, the reason for these two correlations being the opposite is a mystery. One theory we had is that the U.S. had a high proportion of COVID deniers who wanted to keep traveling but were only able to do so domestically.
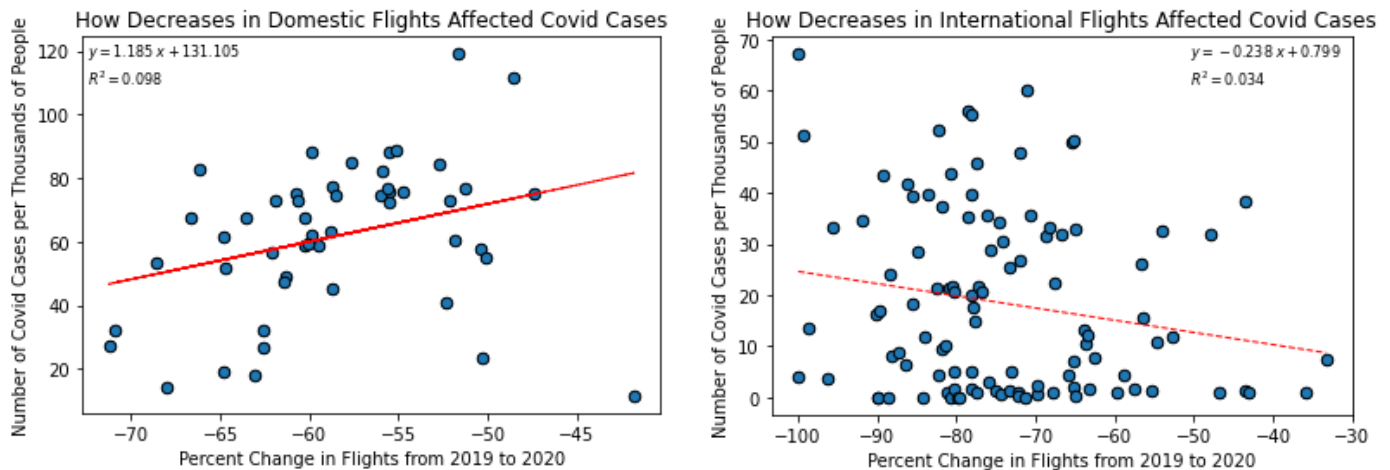
6. Conclusions

COVID-19 was primarily responsible for a huge decrease in commercial flights and movement of people between 2019 and 2020, but it was also responsible for a slight increase in the movement of physical goods. With borders closed and people stuck at home, excess income was often spent on physical goods, and empty planes that were still flying due to sunk airline costs would have more space available for freight. International statistics each saw heavier drops than their domestic counterparts due largely to many other countries closing their borders completely to the outside world. For example, international passenger numbers started a heavy drop in February when countries like China began to close their borders, while domestic passenger numbers didn't experience a harsh drop until lockdown restrictions and the severity of the virus became more widespread in April.

While the percent change in the number of flights and the number of COVID cases in a nation don't have an incredibly strong correlation, a correlation still exists, especially among domestic flights where all the data is available. Part of the reason for the lack of strong correlation in international data could be because countries outside the U.S. only have flight data to and from the U.S., and we suspect that if we had flight data between international destinations, there would be a significantly higher correlation. Another factor could be inadequate COVID data- many countries

underreport or misrepresent their true number of coronavirus cases and deaths, which certainly has a negative effect on the integrity of our data and our conclusions.

Overall, COVID-19 clearly had a negative impact on the airline industry. However, with vaccinations being globally distributed and cases declining, 2021 and 2022 will likely be the beginning of a slow turnaround for aviation.

7. Citations (Data Sources)

COVID-19 Case Surveillance Public Use Data. Centers for Disease Control and Prevention, www.data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf.

Data Library: Aviation. Bureau of Transportation Statistics, www.transtats.bts.gov/databases.asp?Z1qr_VQ=E&Z1qr_Qr5p=N8vn6v1 0&f7owrp6_VQF=D.

INTERNATIONAL TOURISM AND COVID-19. United Nations World Tourism Organization, www.unwto.org/international-tourism-and-COVID-19.

WHO Coronavirus (COVID-19) Dashboard. World Health Organization, www.COVID19.who.int/info.

Countries in the World by Population (2021). Worldometer, www.worldometers.info/world-population/population-by-country.