PandAIRmic

An analysis of the impacts of air travel on the spread of COVID-19

Jian Cong Loh, Joshua Neronha, Stephen Sun, Tzuhwan Seet

Hypothesis

We hypothesize that the number of air passenger arrivals in a US state is positively correlated with the number of COVID-19 cases.

Data

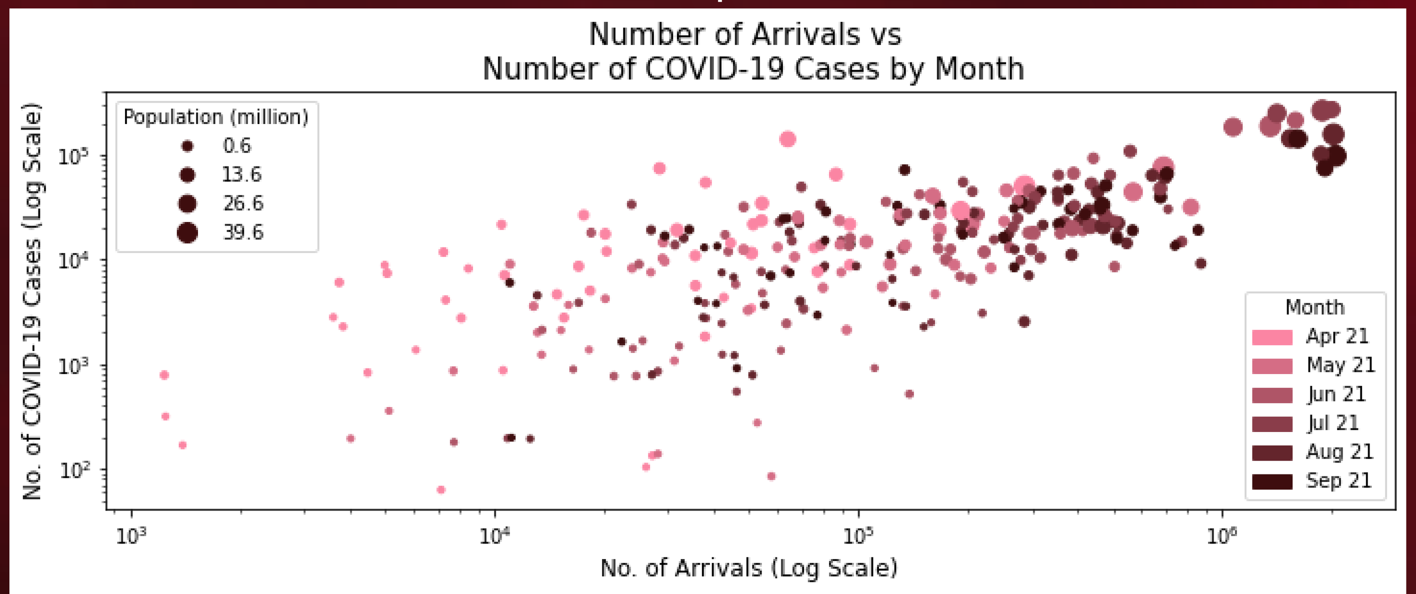
- Our **COVID-19 data** is obtained from *The Atlantic's* **The COVID Tracking Project**. It gives the daily number of positive cases for each of the 50 states, five territories and D.C., and our data spans Jan 20 and Feb 21.
- Our **air passenger arrival data** is compiled from two datasets from the USDOT. The **DB1B** gives a 10% sample of all itineraries by quarter, and covers Q1 to Q3 2020. The **T100** breaks down the number of passengers flying between two airports (including layovers) by month, and covers Jan to Nov 2020. We combined the two datasets to estimate the number of passengers travelling between every pair of states per month, excluding layovers.
- Our **policy data** comes from the **COVID-19 US State Policy Database** developed by the Boston University School of Public Health. It tracks the dates when each US state implemented COVID-19-related social safety net, economic, and social distancing policies.

Methodology

To match the monthly arrivals data, we aggregated the cases from the 15th of each month to the 14th of the next, given the 2-week incubation period. This gave us **294 data points** from **49 states** (excl. DE) and **6 months** (Apr – Sep 20).

Our exploratory data analysis (see below) gave us the following insights:

- States with large pop. had more arrivals and cases – pop. is a possible causal factor for both
- Linear r/s in log-log plot – possible polynomial r/s
- Cases increased over time – possible fixed time effects



We ran simple linear regression with different versions of our variables:

- **Model 1A** – No. of arrivals vs no. of cases
- **Model 1B** – Change in arrivals vs change in cases
- **Model 1C** – Per capita arrivals vs per capita cases
- **Model 1C(W)** – Weighted p.c. arrivals vs p.c. cases

Using Model 1C(W) as the base, we ran several more complex regression models:

- **Model 2** – Polynomial regression
- **Model 3** – Multiple linear regression (w/ policy as controls)
- **Model 4** – Fixed effects model (entity & time)

Simple Linear Regression

Recently, avocados have been specifically tied to the millennial generation, due to its increasing presence in social media and public platforms. While this connection is widely acknowledged, we wanted to see whether it could be statistically supported that millenials consume more avocados, even though increased consumption may have increased prices of this unique fruit.

Polynomial Regression

Multiple Linear Regression

Fixed Effects Model

Map Visualization

Recently, avocados have been specifically tied to the millennial generation, due to its increasing presence in social media and public platforms. While this connection is widely acknowledged, we wanted to see whether it could be statistically supported that millenials consume more avocados, even though increased consumption may have increased prices of this unique fruit.

Conclusions

Recently, avocados have been specifically tied to the millennial generation, due to its increasing presence in social media and public platforms. While this connection is widely acknowledged, we wanted to see whether it could be statistically supported that millenials consume more avocados, even though increased consumption may have increased prices of this unique fruit.

Limitations / Future Directions

Recently, avocados have been specifically tied to the millennial generation, due to its increasing presence in social media and public platforms. While this connection is widely acknowledged, we wanted to see whether it could be statistically supported that millenials consume more avocados, even though increased consumption may have increased prices of this unique fruit.