**Project Summary**

Alex C, Erik C, Max D, Bronson J, Lekun W - Team PRISM

**Introduction**

COVID-19 has greatly impacted international travels, especially between the U.S. and China. Greater Los Angeles is the most popular destination among both Chinese tourists and scholars. Therefore, in this project, we seek to explore what the pandemic might mean to these international communities, as well as local ones, through analyzing air passenger data between LAX and Chinese airports.

**Data** **Sources** (from Jan 2009 to May 2020):

<https://www.transtats.bts.gov/Tables.asp?DB_ID=111&DB_Name=Air%20Carrier%20Statistics%20%28Form%2041%20Traffic%29-%20All%20Carriers&DB_Short_Name=Air%20Carriers> (*T-100 International Segment-All Carriers*, Jan 2009 – Oct 2019)

<https://www.flightera.net/en/search> (info by flight number, Jan - May 2020)

<https://planefinder.net/> (flight capacity info)

<https://zhuanlan.zhihu.com/p/39121498> (list of flights from and to LAX from China)

<https://www.npr.org/2019/05/31/728590535/chinese-tourism-to-u-s-is-down-after-years-of-booming-growth> (# of Chinese tourists visiting LA)

**Methodology**

Passenger volume estimation had COVID-19 not happen:

We explored the model strengths of two popular forecasting methods on seasonal time series data: seasonal ARIMA and STL decomposed forecasting. We judge model strength mainly by their RMSE, MAPE, MASE in their validations, as well as other statistics such as AICc and Q-value. STL methods have shown much better fit and accuracy in training and validation, however, choosing the after-decomposition forecasting method was a bit trickier, as ETS is extremely accurate while random walk does better in preserving the overall growing trend. We eventually decided on STL+ARIMA(0,1,0) but also left STL+ETS as a reference.

Passenger volume estimation with COVID-19 ongoing:

First, we estimated passenger volumes since January 2020. Nov and Dec 2019 are intentionally left out because the pandemic had not yet started then. No accurate passenger data was found, we looked up a list of flights between LAX and Chinese airports, then multiplied each of them by their seat capacity. This would provide a quite reasonable estimate for two reasons. 1) There is usually a high traffic in January each year, so we can assume that flights are nearly full; 2) From Feb onwards, travel restrictions clearly took place as the number of flights dropped noticeably, thus each flight must be almost full, given that demand should not have changed so drastically.

Then we estimated future volume. China has been limiting the number of flights per week for each country since Mar 29, so we have only a few weeks of data to work with. It is hard to give any accurate prediction for when/where the volume will stabilize, so we took a naïve approach in averaging the available data since April, and multiply it into a monthly average.

**Conclusion**

We have selected Sept 2020, Jan 2021, and Spring 2021 as 3 key timeframes within the next 12 months and analyzed what could happen to different communities. With the strong seasonality, we said that Mar-Jun traffic is mostly due to tourism, and the spikes around Jan and Sept are due to international students. With estimated cumulative traffic impact/delays, we have come up with some possible outlooks should the pandemic end in each of these timeframes chosen. In short, Sept 2020 would be the scenario with the least overall impact. However, we should also be well-aware and prepared should the pandemic last, or should there be a resurgence.