# Project Report

Dhruv Relwani
University of Washington MSDS
Data 512, A7 Assignment

## Introduction

In this project report we evaluate the impact the Covid19 pandemic had on Clark County, Nevada. As we are already aware, this pandemic had an unequal impact on different sections of our population. As a software engineer, my work was disrupted to a lesser extent during the pandemic and hence I was privileged to continue my livelihood/ income because of the nature of my job; however, a majority of the world did not have this privilege and hence the pandemic had a direct impact on the livelihoods of a majority of the people. In this project, we will explore the effect Covid-19 pandemic had on the tourism industry in Clark County, Nevada. This analysis is interesting and important to explore because Clark county (specifically Las Vegas) has one of the biggest tourism based service sectors in the world and this industry is one of the primary employers in the region. Hence the progression of the pandemic on the tourism industry in Clark County, meant a direct impact on the livelihoods of the people living there. This research domain is human-centered since it helps us uncover the details of human-interactions during a world-wide crisis and helps us understand how human lives were practically impacted by the trends we see in our Covid-19 data. We will look at this pandemic's impact on the tourism industry through the lens of available Airbnb data in Clark county.

## Background/Related Work

This pandemic has been highly datafied, which has helped in making several research questions possible to explore. Research papers like, "*Exploring the impact of COVID-19 on tourism*" published in the ScienceDirect Journal have studied the consequences and settings of the COVID-19 pandemic and how innovation and change can contribute to the tourism industry's revival to the next normal. We explore the direct COVID-19 tourism impacts, attitudes, and practices in gaining the leisure industry's boom and recovery.

Reports published by the UN World Tourism Organisation show that international tourism was expected to decline over 70% in 2020 and return to the levels of tourism activity 30 years ago. The decline in the first ten months of 2020 represents 900 million fewer international tourist arrivals compared to the same period in 2019, and translates into a loss of US$ 935 billion in export revenues from international tourism, more than 10 times the loss in 2009 under the impact of the global economic crisis.

After reading such reports about the pandemic, I was inspired to focus on the city of Las Vegas within Clark County because of its significance as a tourist destination. Such research reports/ blogs raised questions in my mind which eventually led me to formulate hypotheses about the progression of the pandemic within Clark County. Here are some of the research questions/ hypotheses that I wanted to explore:

- How were the descriptions of Airbnb listings in Clark county affected by the pandemic?
- How was the income of Airbnb hosts in Clark county impacted by the pandemic?
- Hypothesis: For every 10% increase in the confirmed covid cases in Clark county, we notice a 2% decrease in Airbnb listings in that region.
- Hypothesis: The Airbnb listings in Las Vegas were quicker to bounce back after vaccinations were rolled out, than the listings in other parts of Clark county.

## Methodology

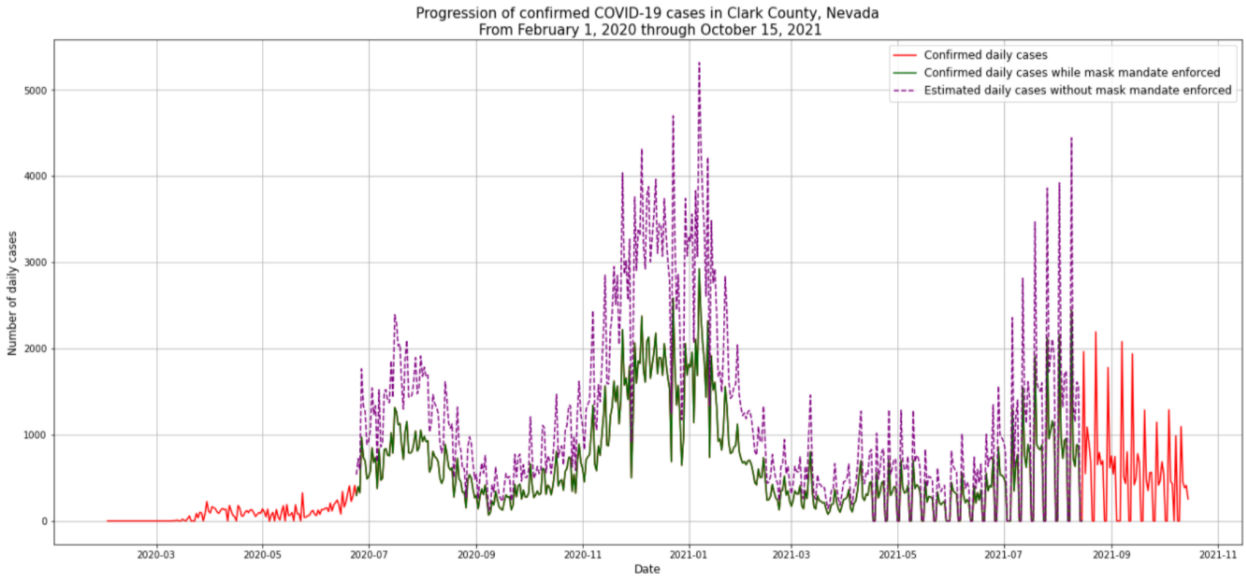
While plotting the progression of the confirmed Covid19 cases in Clark County as a timeseries, I combined the datasets provided by John Hopkins university (about confirmed cases count) and the masking mandate/ compliance datasets by the CDC and New York Times, into pandas dataframes and then selected just the relevant data for Clark county, NV for the timeframe of the assignment. I was only given cumulative case counts in the first dataset and hence had to calculate the daily counts myself by incrementally subtracting from the total count. Then using this info I was able to get a rate of change of cases every day for the second visualization. One additional derived field in this data was the estimated case counts if the mask mandates were not enforced. For this estimated case count I assumed a 55% efficacy of masks in preventing Covid-19 infection, a 10-day infection period from only the confirmed and alive cases (non-asymptomatic) and anyone in Clark county who was alive and not confirmed to be infected as “at-risk” for infection. The population data for this county was assumed by the Census data.

While evaluating the Inside Airbnb data, I was able to clean the available data in Tableau Prep Builder by removing columns that were not relevant, performing joins on datasets for different months of the pandemic and by aggregating data for every month. Then I utilized Tableau’s visualization software for exploratory data analysis of the clean dataset. I was able to encode information about prices and availability of listings in the colors of different plots and adjust them to an appropriate scale. Finally for my hypothesis testing, I used a large sample Z test and set a significance threshold of 5% to compare against.

During these transformations of the Airbnb data, the human centered considerations of data privacy were maintained as the datasets had been cleaned sufficiently to not include any personally identifiable information about the owners of the Airbnb listings. I also made sure to not include the Titles of the Airbnb listings in my analysis since they sometimes

contained very detailed information about the property (like lock mechanisms) which could potentially compromise their security. I wanted to use this data to specifically answer questions about human behaviours and livelihoods during the pandemic. I believe that my analysis could potentially help in advocating for financial aid for the individuals who relied very heavily on tourism within Clark County for their livelihoods and had to scramble to find new sources of income during this economic crisis.

## Findings

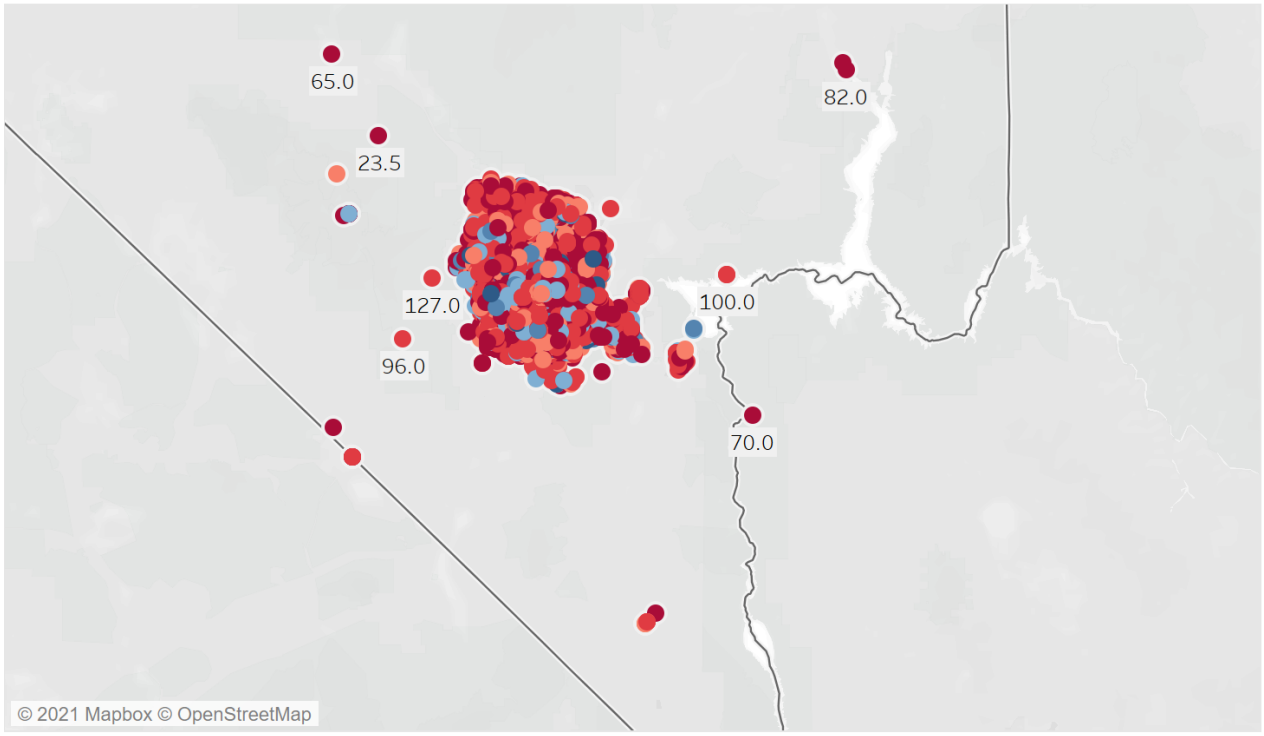



First I wanted to plot the progression of the confirmed Covid19 case count in Clark county so that we set the context for how this crisis could have impacted the people during those months. We can see the periods of the plot highlighted in green for the months when a mask mandate was enforced within Clark county. While answering the research question and developing the model I noticed that mask mandates were statistically significant in preventing worse outbreaks and curbing the rate of infection (from its potential) in Clark county. Even for a person who might consider masking as a non-deterministic atrocity with only a near 50% probability to help with containing Covid, this model suggests otherwise.

In this geospatial representation of different Airbnb listings during the months of December 2020 to October 2021, in Clark County Nevada, we can view their prices on a scale of 0-500 USD per day. There were very few outliers with luxury prices for their Airbnb listings (around $1-2k) that were filtered out for this representation because the color scale would then be completely distorted by these outliers. Unsurprisingly we see most of the Airbnb listings crowded around the City of Las Vegas and we can also see the average daily prices of each of these listings during the pandemic. We notice that listings with a higher

price point are mostly concentrated around the City of Las Vegas (when compared against those at the outskirts of Clark County).

Prices of Airbnb Listings in Clark County, NV (Dec 2020 - Oct 2021)

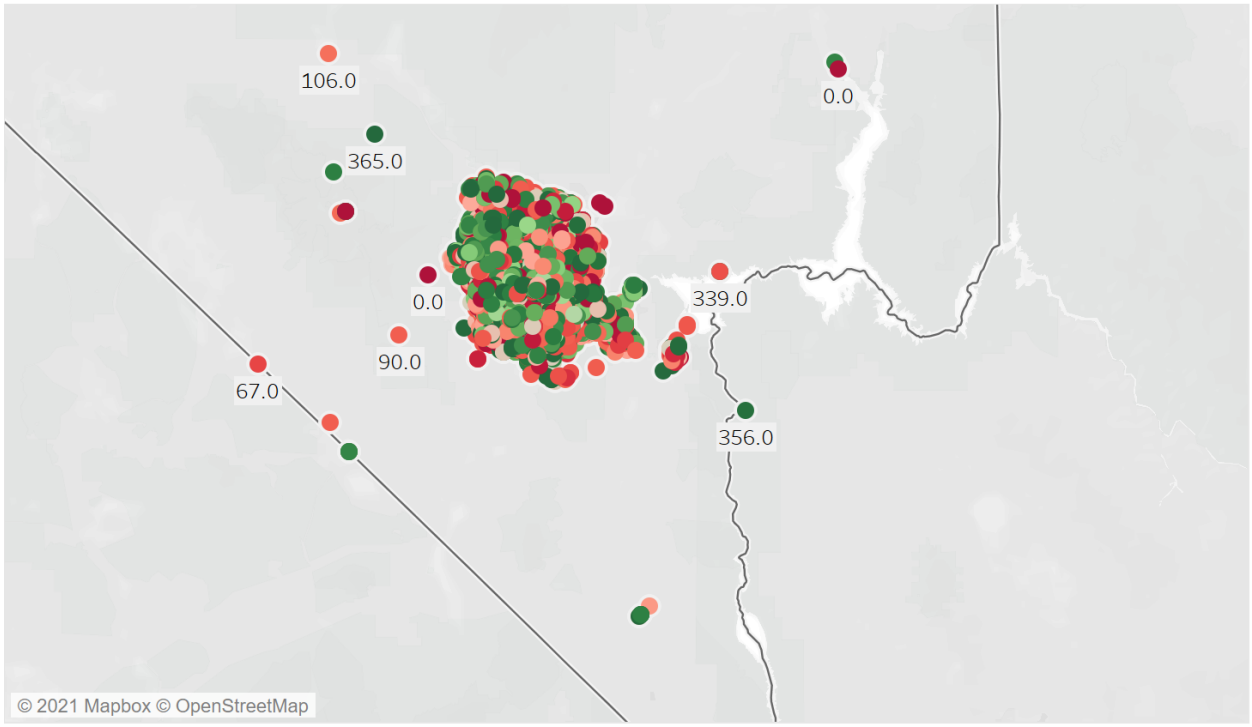



Map based on average of Longitude and average of Latitude. Color shows average of Price. The marks are labeled by average of Price. Details are shown for Id. The view is filtered on average of Price, which includes values less than or equal to 500.0.

Average Daily Price in USD


We were also able to get a similar representation of the Availability of these Airbnb listings (calculated as the number of days a particular listing is available out of 365 days of the year) and plot the Medians for each of these listings during the pandemic months. This gives us an idea about how the lockdowns, mask mandates and the general fear of travel impacted the inventory of unoccupied Airbnb listings and therefore affected a source of income of the people in Clark County.

Median Availability of Airbnb Listings in Clark County, NV (Dec 2020 - Oct 2021)

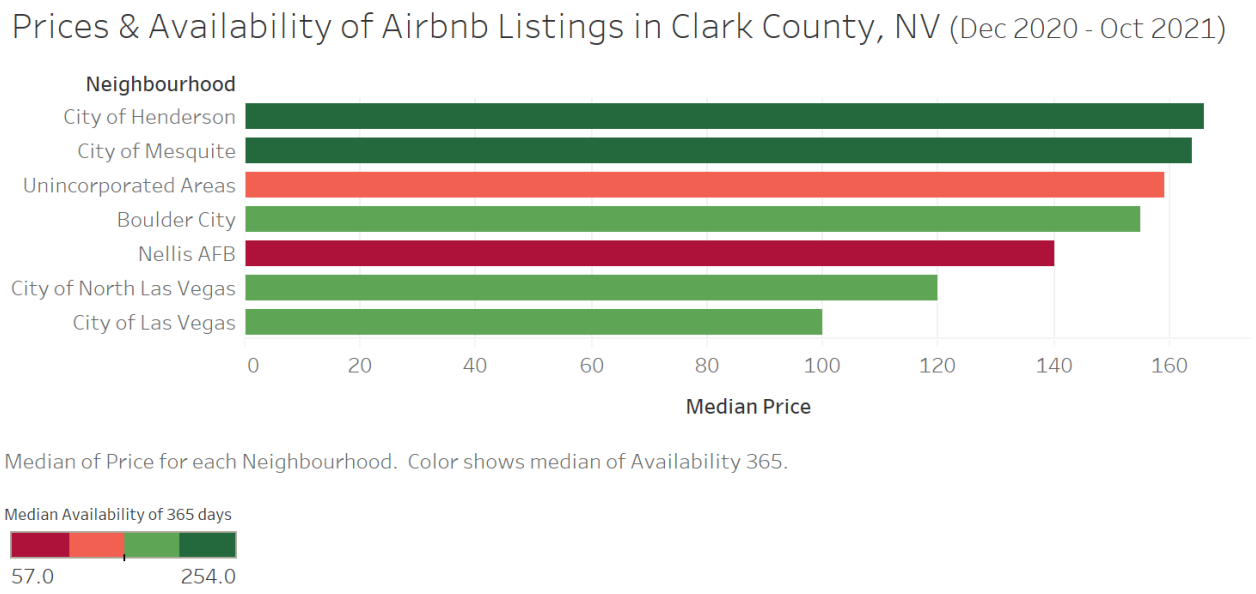



Map based on average of Longitude and average of Latitude. Color shows median of Availability 365. The marks are labeled by median of Availability 365. Details are shown for Id.

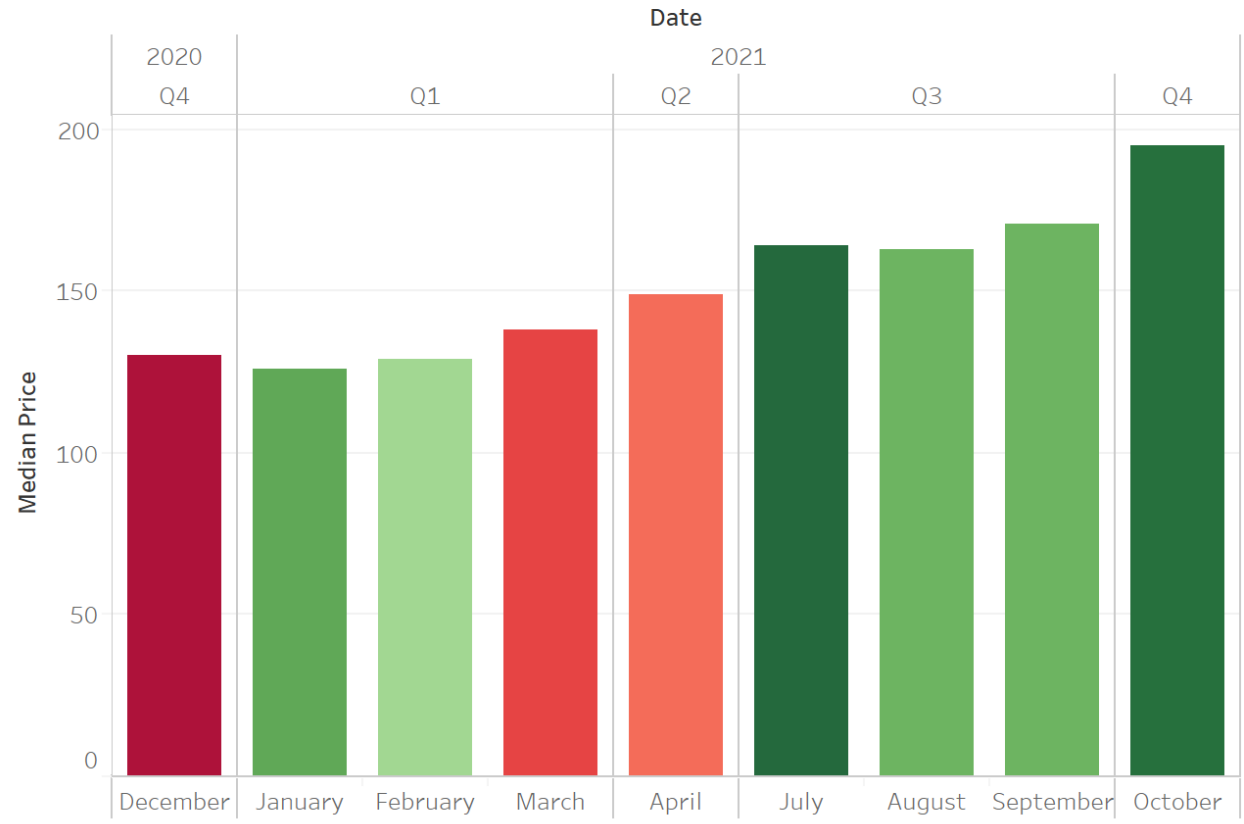
Availability out of 365 days

0.0 365.0

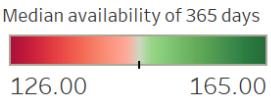
This information led me to investigate how different neighborhoods within Clark County were affected by the pandemic and highlight any inequalities. We notice that on an average the listings in the Cities of Henderson and Mesquite had a higher price point than the listings in the City of Las Vegas, however these cities also had most of their listings available for a greater number of days of the year (indicating that tourists were unlikely to spend a lot of money on travel accommodations during the pandemic).




I then wanted to explore the progression of the median daily prices and median availability of these listings over the months of the pandemic. We notice that at the heart of the pandemic in December 2020, we had very little availability of Airbnb listings and the prices of the ones that were available were low (around $130). Once the vaccines become available and approved by the FDA in the months of January and February, we notice a trend when many of these listings are available on the market again (indicating an anticipation of tourism and increased comfort of the Airbnb owners to allow guests). This trend in increased availability of Airbnb listings continues till we learn about the Delta variant in March and April 2021. However, the prices of these listings continue to climb up. Starting in Q3 of 2021 (July 2021) we notice a significant increase in the median price and availability of these listings (indicating that many people were comfortable to travel again after their vaccination shots and that led to an increase in demand for these listings). Increased availability in the second half of 2021 could also indicate a new wave of homeowners who bought a place during the pandemic because of low mortgage interest rates and listed those properties as Airbnb listings.




Median of Price for each Date Month broken down by Date Year and Date Quarter. Color shows median of Availability 365.




Finally, this led me to conduct a hypothesis test to confirm whether the average price of the Airbnb listings were higher since the beginning of Q3 of 2021 when compared against those from December 2020 to April 2021.
**Null Hypothesis**: The average prices of Airbnb listings have remained the same in Las Vegas since the beginning of Q3 of 2021.

**Alternate Hypothesis**: The average prices of Airbnb listings have increased in Las Vegas since the beginning of Q3 of 2021.

Our P value < alpha (0.05) so we reject the null hypothesis with 95% confidence. This means that the average prices of Airbnb listings in Las Vegas have truly increased from July - October 2021 when compared to the average prices from December 2020 - April 2021.

## Discussion/Implications

These findings help us understand some dimensions of how the pandemic progressed in Clark County, Nevada and how the tourism industry has been impacted there through the lens of the Airbnb listings in the region. This helps lay the foundation for potential research to study the pandemic's impact on the livelihoods of the people in Clark County. We could potentially get more information about flight schedules to and from Clark County during the pandemic and combine that with datasets about casinos, hotel and entertainment industry in the region. This would help in making this research project more comprehensive in evaluating the impact on tourism from a holistic perspective.

Human centered data science principle of fairness helped me in narrowing my research focus to the impact of the pandemic on the livelihoods of the people living in Clark County. It helped me focus on a topic that could inspire discussions about a fair financial aid for the people who lost their sources of income during this economic crisis.

Secondly, the concept of reproducibility and intuitiveness helped guide my decisions while recording every step I made in cleaning the raw data in Tableau Prep Builder and storing that as an Output.hyper file. While selecting the visualizations in Tableau I made sure to make them as intuitive and understandable as possible on the first glance to any user (for example: by including a detailed legend and selecting appropriate graphing choices for different data measurement types: ordinal/ nominal/ geospatial). These choices would help another researcher easily trace my steps and have access to raw data with detailed descriptions to reproduce my research.

Lastly, the principle of data privacy helped guide my decision to not include the Airbnb listing titles in my research as they sometimes contained information that could impact the privacy of the Airbnb owners.

## Limitations

The most significant limitations for this project include:

- For the Inside Airbnb data, not all data was able publicly for the duration of the pandemic. To access data archives for Clark County, Nevada for the duration of Feb 2020 - November 2020, I would have had to submit a data retrieval request to the activist organization maintaining this dataset and pay for it. Also the months of May and June 2021 were omitted in their publicly available records for Clark County.
- Lack of raw datasets that included information which could help me answer some of my original hypothesis questions from the A5 assignment. For example: Their weren't datasets which provided information whether an Airbnb listing within Clark county was a new listing or an existing one that became available after a particular date. This information was indirectly represented by a calculated field provided in the data set called "the availability of an Airbnb listing out of 365 days in a year".

Hence I modified my hypothesis question to test for daily prices of these listings before and after Q3 of 2021.

- The dataset that provided information about daily prices of Airbnb listings was normally distributed for the sample that was available to us. Our assumption for the validity of the Large Sample Z Test was that the rest of the data set was also normally distributed.

## Conclusion

This research project helped us understand how the pandemic's confirmed case counts progressed in Clark County, Nevada and how the mask mandates that were enforced, potentially helped in keeping the case counts lower. We also learned that most of the Airbnb listing in this county are unsurprisingly concentrated in the areas near Las Vegas. Visually we can see that the listings with a higher price point (closer to $500 per day) are also more concentrated around Las Vegas. We see that the median availability of these listings was at a low in December 2020 and then they increased in Q1 of 2021 before dropping again towards the end of the same quarter when concerns about the Delta variant and case counts increased. The prices and the availability of the listings then significantly increased starting in July 2021, which indicated a comfort of the Airbnb owners to host new guests and the increased willingness to travel of the tourists after vaccinations were rolled out throughout the country. This study helps us understand human centered data science as it fundamentally sheds light on questions about human behaviours and livelihoods during the pandemic. It shows us how human emotions of fear and decision making impact economies at a macro level and therefore the income of the people in that region.

## References


- Jaffar Abbas, Riaqa Mubeen, Paul Terhemba Iorember, Saqlain Raza, Gulnara Mamirkulova, Exploring the impact of COVID-19 on tourism: transformational potential and implications for a sustainable recovery of the travel and leisure industry, Current Research in Behavioral Sciences, Volume 2, 2021, 100033, ISSN 2666-5182, https://doi.org/10.1016/j.crbeha.2021.100033. (https://www.sciencedirect.com/science/article/pii/S2666518221000206).
- Unwto.org. 2021. Impact assessment of the COVID-19 outbreak on international tourism | UNWTO. [online] Available at: https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism [Accessed 15 December 2021].

## Data Sources

- Inside Airbnb data: The data is a collection of Airbnb data that is publicly available on its website, across multiple cities. The data is aimed at providing a 360-degree insight into Airbnb's presence in a city. Dataset URL: Get the Data - Inside Airbnb. Adding data to the debate.
- The RAW_us_confirmed_cases.csv file from the Kaggle repository of John Hopkins University COVID-19 data. This is where I got information about confirmed Covid-19 case counts.
- The CDC dataset of masking mandates by county. Here I got information about the dates during which a mask mandate was enforced in Clark county.
- The New York Times mask compliance survey data. This dataset provided information about mask mandate compliance probabilities.