Repo name: Group_G_Covid_Housing_Prices

Title: Comparing Covid-induced changes in Housing & Hotel prices in two US states

Group Members: Kum Hyun Lee (khl2139), Mengying Xu (mx2238), Yi Hyun Kim (yk2906)

Abstract

This project aims to explore how the 2019 Covid-19 pandemic has impacted the housing and the hospitality market by focusing on comparing price changes in housing and hotel rooms across different US states. The project will track how waves of covid cases affected price fluctuations in housing and hotels in two regions selected via EDA explained below (see "Selection of Case Study Area").

The following questions will be examined:

- How has Covid affected the housing market via housing prices?
- How has Covid affected the hospitality industry via hotel prices?
- Was the housing market worse hit than the hospitality industry, vice versa?
- Were there any regional variations in housing price/ hotel room price fluctuations?
    - If so, how is the correlation between changes in housing prices and hotel room prices following the rise and fall of Covid cases different in different regions?

This project will use ggplot2, ggmap and tmap in R, and possibly geopandas and altair in Python while including interactive visualizations.

Covid & Housing/ Hotel prices

- Covid is a global pandemic that has impacted housing and hotel prices dramatically. It will be hypothesized that throughout time, higher instances of Covid-19 cases decreases the price in general, but may have small fluctuations before and after the recovery from a new variant.
- Also, it will be hypothesized that Covid-19 would influence housing and hotel prices in different ways with regards to demographic factors. For example, in areas with higher population densities, the price drop may be more severe compared to that of areas with lower population densities.

Selection of Case Study Area

Two case study areas will be selected after conducting exploratory data analysis (EDA). This will entail comparing the housing price, hotel price, and demographic data for different areas in the US. Two approaches will be used for the purpose of selecting the areas:

1) Comparing two cities within the same state that have different characteristics (e.g. city that is a popular tourist destination vs residential city)
2) Comparing two cities from different states that have similar characteristics (e.g. one city from the West Coast and another from the East Coast that are both popular tourist destinations)

Data

This project will require data on each of the following:

1) Covid-19

- Covid-19 cases and deaths at state level (CDC): https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36/data
- Covid-19 data at county level (New York Times): https://github.com/nytimes/covid-19-data

The Covid-19 data at the state level will be needed for EDA, and the data at the county level will be used after narrowing down the case study areas for analysis.

2) Housing prices

- FHFA House Price Index
  measures changes in single family home values (50 states, over 400 cities, 1970-) : weighted, repeat-sales index that measures average price changes in repeat sales or refinancings on the same properties
  https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index.aspx

3) Hotel prices

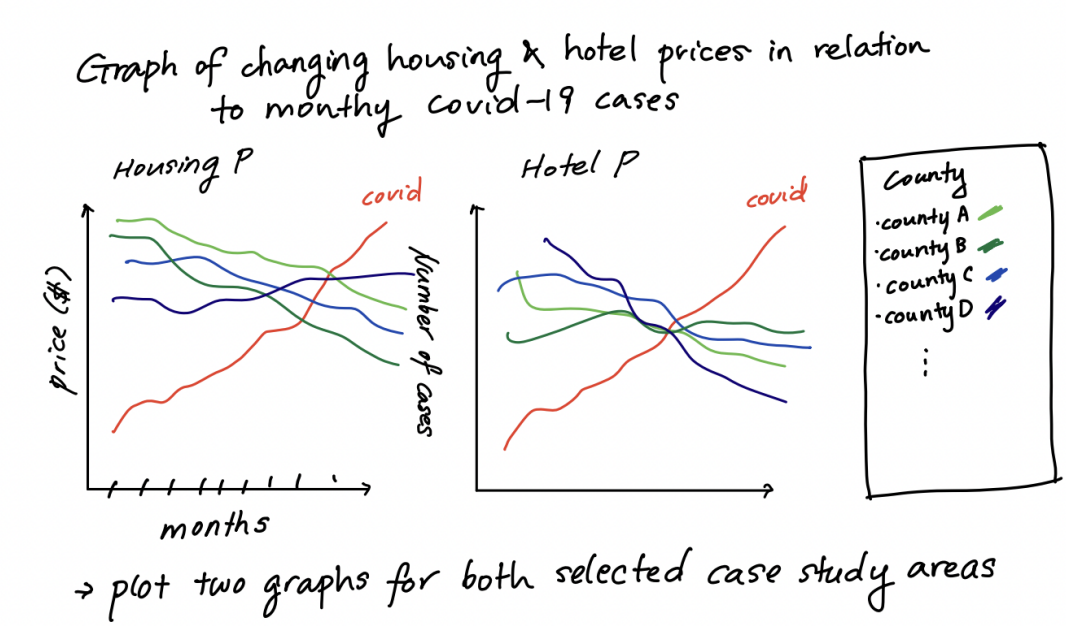
- *Request free sample:*
  https://datarade.ai/data-categories/hotel-rates-pricing-data

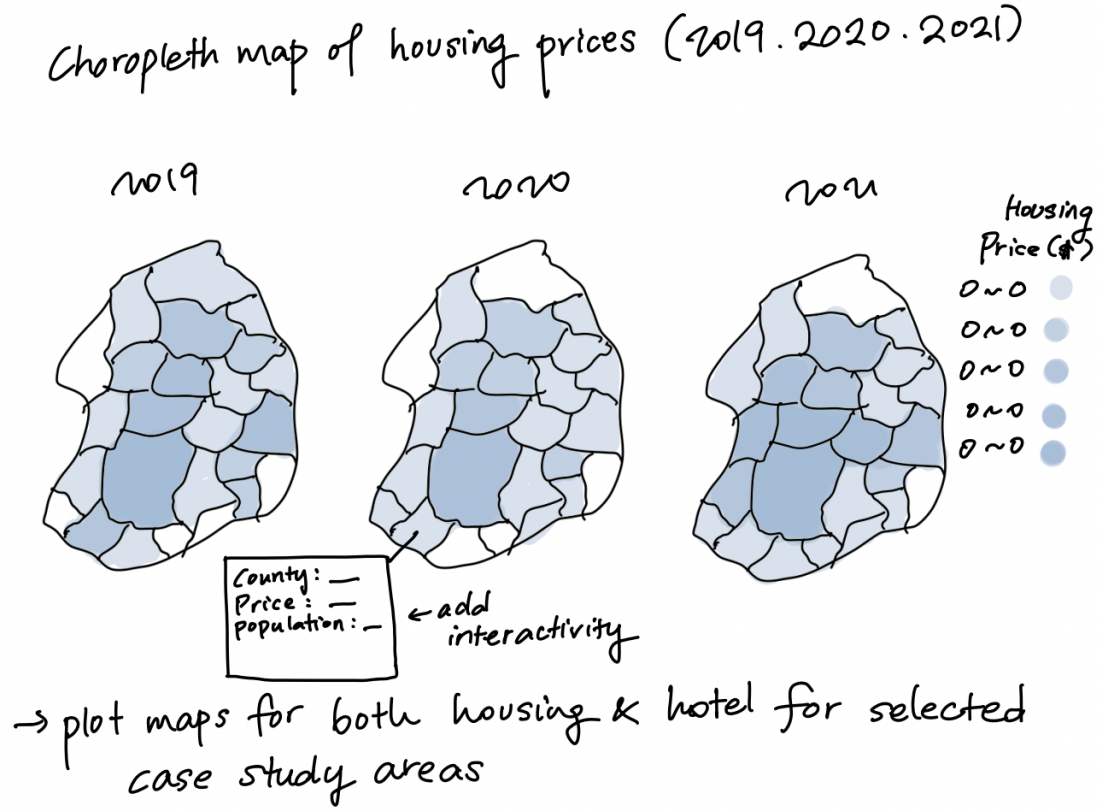
4) Demographic data

- American Community Survey 5-Year Data (2009-2019):
  https://www.census.gov/data/developers/data-sets/acs-5year.html

This dataset can be retrieved in Python using an API key.

Ideas for Visualizations





Hurdles

- Data availability for hotel prices
    - Limited Access
    - Lacking geographical indicator