Project 3: Gentrification Within the U.S.

Meiling Traeger and Damian Franco

Introduction

According to a recent White House article, 2021 has seen the greatest increase in housing prices over the past year within the US (Bernstein et al, 2021). Although this is partially due to the pandemic, the rising cost of living is not a recent trend. In 2018, the Urban Institute found that millennials are becoming homeowners at lower rates than previous generations (Choi et al, 2018). Clearly, it has become harder for the average American to find a place to live. There are a variety of factors that can impact the price of homes: available inventory, interest rates, etc. However, in 2010, the Nation Bureau of Economic research documented a link between housing prices and gentrification (Guerrieri et al, 2010). Gentrification can be characterized as the process of high-income groups moving into low-income urban areas. This often results in existing residents, usually minority groups, becoming displaced and shifting the economic landscape of the area, i.e. more expensive businesses being established.

In our project, we aim to explore the connection between housing prices and socioeconomic shifts within the United States. By utilizing housing prices, median income value and racial population data, we will attempt to create models that can show where high-income groups might displace those in lower-income areas and the changes in housing prices which might follow. By predicting these changes, home buyers and renters can make more informed decisions as to where they should look to live. Additionally, gentrification is a very important topic that deserves attention which is a motivating reason for us to continue this research.

Goals

Last year, one of our group mates, Damian Franco, collected and correlated differences and similarities between every zip code area within the United States with the help of Nicholas Livingstone and Datenzing Tamang. Gentrified areas can be very difficult to categorize, but there was a creation of a set of criteria that the previous team settled on last year. This criterion is represented by Figure 1 in the Race and Gentrification Within the US paper (Franco et al, 2021) and will be expanded on in our planned paper for this project.

There were many correlations that were found through several clustering and other data analysis techniques. Since there was no implementation of data predication developed last year, our goal is to implement a model that could allow for the forecasting of which area may be considered gentrified within the next time series step. Models that we will plan on utilizing for data prediction include regression and neural networks models. We are most comfortable with those models, but plan on investigating other time series forecasting models as well.

The data that was analyzed last year was the U.S. Census data from 2010-2019, so we will try to predict data for the year 2020 and compare our results to the actual results from the U.S. Census data in 2020. Comparing the data will allow us to improve our predicting model and

analyze how realistic dynamics work of the human population within these areas. Comparison will most likely be done through plotting information and comparing integer values, but we will also try to calculate entropy and mutual information of the real vs predicted data. Zip codes were the areas that were investigated through the research last year. We will also attempt to change the data sets from the zip code area size to a smaller size with census tracts data sets. Census tracts will allow for more intricate results due to the smaller area size. The transition in data sets might be more difficult than expected, so we are going to attempt this but no promises for the transition over.

Gentrification would be very useful to investigate and predict, but we do expect our results to be different to the exact results because the human population is a very complex system and trying to predict the dynamics of it could result in either small or large differences. We would consider this a complex adaptive system problem because human population dynamics are always changing and could be somewhat chaotic. Many times, there are many factors that could influence where people choose to live and investigating this topic further will help prove how complex this realm could be.

What Success Looks Like

Success within our project will consist of the generation of a prediction of the dynamics of each area that was investigated. An analysis of these predictions and its comparisons between fabricated and real life dynamics will allow us to judge how successfully the models we implemented. Success to us will not be based on the actual numbers of the prediction, but rather how the predictions could be explained and the insight we gain from these comparisons.

Some insight that we hope to gain out of this project is to dive deeper into the dynamics of the human population, familiarize ourselves with time series forecasting models, and analyze how the human population is a complex system that can or cannot be predicted with the use of data forecasting.

We are keeping in mind that our prediction year could be influenced by the pandemic within the real data, which will not be accounted for within the predicted data. With the pandemic in mind, we can analyze exactly how much the pandemic has influenced gentrified areas within the United States.

Timeline

Below is a 5-week plan for the most optimal outcome for this research.

- Week 1:
    - Familiarize ourselves with others' research on gentrification within the U.S.
    - Attempt to translate zip code area size data with census tract area size data for more expansive and accurate results.
    - Decide on which time series forecasting model each of us will implement within our research.

    - Investigate and implement other factors that could relate to gentrification within the U.S.
- Week 2:
    - Begin implementing the time series forecasting model on our data.
    - Create examples (through the data) that represents gentrified and non-gentrified areas
    - Create examples (through the data) of the complexity of the human population within these areas.
- Week 3:
    - Continue and finalize implementing the time series forecasting model on our data.
    - Use the implemented model to predict dynamics of each area for 2021.
- Week 4:
    - Investigate and plot how our predicted data compares to the actual U.S. Census data.
    - Attempt to implement shared information and entropy calculating.
    - Start writing the report for this research.
- Week 5:
    - Prepare and finish the presentation on our topic.
    - Finalize the report.

References

Bernstein, Jared, et al. “Housing Prices and Inflation.” *The White House*, The United States Government, 9 Sept. 2021, https://www.whitehouse.gov/cea/blog/2021/09/09/housing-prices-and-inflation/.

Choi, Jung Hyun, et al. “The State of Millennial Homeownership.” *Urban Institute*, 18 July 2018, https://www.urban.org/urban-wire/state-millennial-homeownership.

Guerrieri, Veronica, et al. “Endogenous Gentrification and Housing Price Dynamics.” *NBER*, National Bureau of Economic Research, 27 July 2010, https://www.nber.org/papers/w16237.

Richardson, Jared, et al. “Gentrification and Disinvestment 2020: Do Opportunity Zones benefit or gentrify low-income neighborhoods?” *NCRC*, National Community Reinvestment Coalition

Franco, Damian, Livingstone, Nicholas, Tamang, Datenzing. “Race and Gentrification within the U.S.” University of New Mexico. 10, December 2021.