

# Part 2 - An Extension Plan

## Background

During the last three years, we all have gone through a global pandemic that has not only been tragic but has also disrupted almost everything globally in addition to taking a deep toll on many individuals and their families.

However, from the perspective of data and analysis, one aspect that has been hard to miss in the last two years is the datafication of the pandemic. That is, many aspects of the pandemic have been collected, aggregated and re-represented as data. This datafication gives us the privilege to examine the pandemic from potentially many different perspectives to understand how it has changed lives and how it has changed society. To be honest, we are actually at the very beginning of understanding and comprehending these impacts.

During our Course Project we are going to begin taking a look at some of the social aspects of the pandemic by conducting a human centered data science analysis of some available COVID-19 data. My specific assignment for the Common Analysis was in Shelby county, Tennessee.

## Motivation/Problem Statement

For this project, I am planning to analyze the impact of covid on the housing market. I strongly think that one specific area that specifically experienced the brunt of this unprecedented crisis is the housing market. During the mid of the pandemic towards the end of the year 2020, I witnessed that suddenly almost all of my friends bought houses. Even the ones who had no plans to buy one ended up buying one. This was mainly due to the excellent housing market conditions that suddenly came up during that time- sudden large supply in the market and the reduced prices along with the very low interest rates. But as we moved on, the housing market from the start of the year 2021 created an immense anxiety and worry overall especially among the people who were

planning to buy a home for the first time. This was due to the fact that the demand of the houses had piled up, however, that led to an automatic decline in inventory and sharp escalations in the house prices.

Even now, the housing market is on the verge of a breaking point with a lot of uncertainty and increased interest rates. From my initial analysis, it is neither a buyer's market nor the seller's market right now as nobody knows about the exact housing market situation. A lot of the home-buyers across the country are frustrated with this increasing trend in housing prices. This topic is particularly human-centered as it affects a vast majority of Americans who are looking to purchase homes. If the prices keep increasing at this rate, a lot of the people from lower-income brackets will suffer as they will not be able to afford to house, or will be pushed to outskirts or suburbs thus causing a widening of the wealth gap between the rich and the poor. Because homeownership is an important tool for building long-term wealth and children of homeowners are likely to become homeowners, this trend can further exacerbate wealth inequality for future generations.

By performing this analysis I hope to learn if there was any correlation between the number of covid cases and housing prices. This analysis is of particular scientific interest as it will help others see trends between housing prices and a fluctuating economy during a pandemic.

## Research questions and/or hypotheses

With the progression in the pandemic, we clearly saw several waves in terms of covid cases. There was an indefinite impact on the businesses specifically as they were shut down resulting in people losing their jobs for a longer duration of time. This definitely created a lot of anxiety as we saw a sudden increase in the inventory with sharp reductions in the list prices for the houses. Meanwhile, this was followed by a series of interest rate reductions. Though this cycle reversed as anxiety increased among the buyers as well as the sellers, what remains still uncertain is the fact if the changes that we saw in the housing market were due to the impact of the pandemic or were they

triggered due to other reasons. Even now, the housing market remains shaky and on the edge.

As part of this project, I plan to explore the following questions to help understand the trend in the housing market since the start of the pandemic.

1. During the COVID-19 pandemic did the housing prices go up or down?
2. Did the number of COVID-19 cases and deaths have an impact on the housing prices?
3. Did changes in 30-year fixed Interest Rates cause any fluctuations in Median Sale price and number of homes sold in the market?
4. Were there any other trends in the data related to covid cases and the housing market or did the trends that we saw in the housing market due to other economic reasons?

## Data to be used

To perform this analysis, I will be using the data from Redfin's Data center for housing market data along with the datasets that we earlier used in the class for performing the common analysis:

- [John Hopkins University COVID-19 data](#)
- [Masking Mandates by County](#)
- [Mask Compliance Survey](#)
- [The weekly housing market data from Redfin's data center](#)

Redfin provides a rich dataset on the housing market that I think would provide a great context and base for performing my analysis.

## Unknowns and Dependencies

- As of now, I do not have any dependencies except the fact that since for my county, we did not have the data regarding the masks mandates, I had relied on the national CDC

guidelines for the masking dates in order to perform a time series analysis with respect to the mask mandate dates.

- Also, I am still deciding on the time period to use just this year's data or to go with the time period that we had used for performing the common analysis.

## Methodology

Since, I think this is purely a data science problem, I will use several basic statistical methods to perform my analysis:

- I will be checking if there is any correlation between covid cases and housing prices by performing exploratory data analysis. I would do some plots as well to clearly see any trends, etc.
- I will be performing linear regression to predict housing prices for my time period and compare it to the housing prices that we actually saw during that time period.
- I am also planning to perform some other basic analysis and visualizations to look at the trends that we saw in the housing market and to see if those were correlated to other factors that were impacted indirectly by the covid19 pandemic

## Timeline to completion

For me, the overall timeline to complete this task would be at least 2 weeks with the following breakdown of my tasks:

### **Data Analysis (6 days):**

- Performing initial analysis with respect to housing - 3 days
- Correlating the analysis with Shelby county and performing deep analysis - 3 days
- Drawing conclusions and compiling the results - 1 day

### **Presentations/Part 3 (1.5 days):**

- Creating presentations and preparing for the presentation using the PechaKucha format - 1.5 days

### **Final Report submission/Part 4 (5.5 days):**

- Incorporating feedback from presentations - 2 days
- Preparing the final report - 3 days
- Submission on github repo - 0.5 days