

Term Project Deliverable 1
Joshua Greenert
7/30/2022

Introduction

With housing prices climbing quicker than they ever have, the trepidation within the market has never been more tangible. As a homeowner myself, I wonder whether there is going to be a massive crash in the market and the housing market soon myself. Since there doesn't seem to be much of a conclusion amongst professionals, this paper intends to answer the question of whether the housing market will crash, and what major factors describe this scenario. Since the data regarding homeownership is massive with multiple facets, this is a massive data science problem that needs insight and answers.

Research Questions

1. Are there comparable data that express what possible variables might be at play in the housing market?
2. What was the value of a dollar in the world economy in previous declines?
3. Are houses on the market currently overpriced?
4. What can be done to ensure that the housing market remains afloat?
5. How significant does the overall market play into the housing market specifically?
6. Would inflation play a part in the housing market prices and potential crash?
7. Were there particular situations during previous housing market crashes that are still occurring?
8. What is encompassed by the grouping of the housing market?

Approach and How It Addresses the Problem

To address if the housing market is going to crash, we will need to rely heavily on previous data from other instances of housing market crashes. Likewise, we'll need to review the datasets that trend parallel to our housing market data to ensure that we can confirm the correlation. After learning about when and why previous markets crashed, we can discover what other variables were applicable during these times to see if those variables share a similar

occurrence in our current time period. Once all data has been collected, an assessment can be made to answer the question: “Will the housing market crash?”

Data

These data sources are what will be used to gather the initial findings; however, the amount of data sources — due to the complexity of this problem — will most likely increase beyond this initial set of sources.

- [Home Price Index Dataset](#)
 - Kaggle (Link included)
 - The data source selected covers instances of the Home Price Index (HPI) over several years. This data will assist in isolating instances of crashes to predict future occurrences.
- [US Minimum Wage by State](#)
 - Kaggle (Link included)
 - This data source will provide the minimum wage for each state from the years 1968 to 2020. This data assists in outlining the effects of inflation and poverty amongst the housing market, and how that impact correlates.
- [US Dollar Index Data](#)
 - Kaggle (Link included)
 - The source data provided with the index of the US dollar should show correlations between inflation and the housing market. Similarly, there should be additional interactions taking place between the US dollar's price value and the impact that took place on several markets.

Required Packages

The packages listed here may expand due to the future altered requirements but should encompass the entire project overall.

1. Readxl
2. Plyr
3. Ggplot2
4. QuantPsyc
5. Car
6. Rmarkdown
7. Knitr
8. Dplyr

Plots and Table Needs

While there are several plots that will assist with the needs of this objective, the main ones used will involve Histograms, Scatter Plots, Bar plots, and potentially box plots. The use of these visuals will be explained when introduced to ensure that the audience is able to grasp the data and its connection to the problem statement.

Questions for Future Steps

- Is there a correlation between housing market values and the US dollar?
- Is there a correlation between housing market values and inflation rates?
- Are there outliers in any of the yearly datasets that will impact the overall results?
- Do the variables discovered describe the variability to a large degree, or is there more information that is needed to obtain the explanation of the data's behavior?