***Comparing Mean Income with Median House Sales Price:***

***Adjusted for Inflation***

**Brittany Roberts, Sebastian Rolett & Kevin Arnold**

# DSC450-T301 Fall 2023

# <https://brittlaur.github.io/>

# <https://emberanimus.github.io/DSC450/> <https://kevinqrnold.github.io/>

# Which Domain?

*What domain is this data going to come from? Please list 10 references (with a brief annotation) to use to make sense of what you’re doing with these data.*

[FRED Economic Data Database](https://fred.stlouisfed.org/) - Economic Research from the Federal Reserve Bank of St. Louis

[Inflation & Cost of Living](https://www.forbes.com/advisor/investing/inflation-cost-of-living/#:~:text=%E2%80%9CWhen%20inflation%20rises%2C%20so%20do,housing%2C%20medications%20and%20transportation.%E2%80%9D) – Information on how inflation effects cost of living

[Inflation & Minimum Wage](https://www.forbes.com/advisor/personal-finance/minimum-wage-vs-inflation/) - Information on inflation and minimum wage

[Home Prices & Household Incomes](https://listwithclever.com/research/home-price-v-income-historical-study/) - Information on home sale prices and household incomes

# Which Data?

*What is the dataset you’ll be examining? Please provide a codebook if there is one or a link to the dataset as well as a detailed description.*

<https://fred.stlouisfed.org/series/MSPUS> - Median Sales Price of Houses Sold for the United States

<https://fred.stlouisfed.org/series/MAFAINUSA646N> - Mean Family Income in the United States

<https://fred.stlouisfed.org/series/FPCPITOTLZGUSA> - Inflation, consumer prices for the United States

We will be looking at these three datasets from the FRED Economic Data database provided by the Federal Reserve Bank of St. Louis. The data provided goes from 1953 to the present time. The Median Sales Price of Houses was collected quarterly.

# Research Questions? Benefits? Why analyze these data?

*How are you proposing to analyze this dataset? This is about your approach. Here, you’ll be proposing your research questions as well as justifications for why you’d offer these data in this way.*

Our analysis will be looking at what the effect of inflation has on a family’s income and what house sale prices would be with that adjustment in inflation. We would like to see how inflation changes what a family would be able to afford with their income.

# What Method?

*What methods will you be using? What will those methods provide in terms of analysis? How is this useful?*

We will first merge the three datasets together to make it easier to work with the data all together. This will also allow us to create more categories to work with, without deleting replacing existing categories, and make data cleansing/wrangling easier to accomplish.

We will be working with a supervised learning linear regression to predict house sale prices through inflation adjusted incomes.

# Potential Issues?

*What challenges do you anticipate having? What could cause this project to go off schedule?*

All three datasets are under one hundred entries, making them small datasets. This could cause the data to be skewed by any outlier in the data. It would be difficult to eliminate them from the analysis because it would give us an even smaller dataset to work with.

The datasets do not consider any cost of living or number of incomes for each family. It also does not include the number of people in each household.

# Concluding Remarks

*Tie it all together. Think of this section as your final report’s abstract.*

As previously stated, we are looking to see what effect inflation has on a family’s income and house sale prices. With the increases in inflation and thus cost of living, we want to see how affordable it would be to buy a house based on people’s household income.

Variables being used will be mean household income, median sales price of houses, inflation rate and inflation adjusted incomes. Factors that will not be included in the analysis will be the number of people in the household, current prices of everyday necessities, or number of incomes due to no data collected.