COMP4447 – Data Tools 1 Final Project

Authors: Sammantha Firestone and Jacquelyn Noyes

**Datasets and Motivation**

Currently in the United States, one of the largest economic barriers that exists is the unaffordability of housing. Historically, the average cost of a single-family home was around five times more than that of the average yearly household income; however, in recent years, that ratio has increased to the average house costing over seven times more than the average household income.

The motivation of this analysis was to further understand the disproportionate increase of housing prices compared to household incomes. To accomplish this, we pulled from multiple data sources to construct a set that fit our needs.

The data for median household income was collected from the [US Census](https://www2.census.gov/programs-surveys/cps/tables/time-series/historical-income-households/h08.xlsx). The data is organized based on median income per year, from 1984 to 2021, for each state in the United States. There is also a row that shows the median income throughout the years for the entire country.

The data for the average price of a house was collected from Zillow. One dataset listed the [median price](https://files.zillowstatic.com/research/public_csvs/zhvi/Metro_zhvi_uc_sfrcondo_tier_0.33_0.67_month.csv?t=1664126527) of a house based on year and region of the United States. The other dataset contained an algorithm for [predicting price](https://files.zillowstatic.com/research/public_csvs/zhvf_growth/Metro_zhvf_growth_uc_sfrcondo_tier_0.33_0.67_month.csv?t=1664126527) increases and/or decreases based on previous trends.

For this analysis, we will clean and combine the data taken from the above datasets to create a singular data frame that includes median household income and average house price based on year and state. We will then work to visualize the growing gap between household wages and housing market prices.

**Literature review**

Unsurprisingly, there have been numerous other investigations of this topic.

<https://www.jchs.harvard.edu/blog/price-to-income-ratios-are-nearing-historic-highs>

<https://anytimeestimate.com/research/housing-prices-vs-inflation/>

The above sources have looked at other factors affecting inability to afford a house (tuition, consumer price index increase) rather than just the relationship between household income and housing prices.