These past couple of years, the United States housing market has gotten more and more expensive. Goldman’s Sach’s reports that housing prices will rise by 4.5% this year and 4.4% the following year which is an increase from the previous estimation of 4.2% and 3.2%. They also reported that affordability for these houses are at an all time low since they started collecting that data in the 1980s (Goldman Sachs, 2024). And while there is some hope with lowering prices with Forbes stating there was a decrease in August this year in monthly prices by 0.13% (the first decrease in August since 2022), the majority of new home buyers are struggling to find houses for an affordable price (Forbes, 2024). CNN reported that only 2.5% of houses in the United States actually changed hands this year in the first 8 months; that is the lowest turnover rate in the past 30 years (Delouya, 2024). What would be beneficial to these buyers would be figuring out what factors of real estate have the biggest impact on the price of a house. So, I would like to answer the following question: what parts of a house have the most influence on the price. Is it the number of bedrooms, the number of acres of land it’s on, or maybe the location? The difference between thousands of dollars in savings could be decided by just removing one bathroom from a buyer’s wish list. My hypothesis is the biggest impact is the location (specifically state) and the number of bedrooms and bathrooms. Location would play a role as a higher-in-demand location will raise the price. And more bedrooms and bathrooms means the house is bigger, which usually raises prices as well. But I’ll look into all the variables to see what has the biggest impact.

I will be using the following dataset for my final project:  
<https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset/data>

The dataset has 12 variables (one of them being the dependent variable, price). It also has quantitative (number of bedrooms, number of bathrooms, lot size in acres, etc.) and categorical (housing status, city, state, etc.) predictor variables. It also meets the minimum number of observations as well. I believe some potential independent variables are state, bed, bath, and acre\_lot. There’s also potential to simplify some of the variables into indicator variables (for example, split states into the 25 most expensive and 25 least expensive. If the state is more expensive, assign a 1. If it’s less expensive, assign a 0). Data was collected from realtor.com. It’s a real estate listing website operated by the News Corp subsidiary Move, Inc. and based in Santa Clara, California. It is the second most visited real estate listing website in the United States as of 2024, with over 100 million monthly active users. The data was collected via web scraping using python libraries. The data was originally collected November, 2023.

As for the methods of analysis, I’ll experiment with a mix of:

1. simple linear regression model (Y = B0 + B1X)
2. more complicated linear regression model (Y = B0 + B1X1 + B2X2)
3. quadratic models
4. second order models (with one, two, and three predictor variables)
5. interaction terms, if necessary

I will analyze the health of those different models including analyzing the residuals, R^2 terms, and utilize tests with F and t scores along with anova tables. There will also need to be some cleaning of the data (removing of NA values, removing outliers, etc) before starting the analysis.

References

*US house prices are forecast to rise more than 4% next year*. Goldman Sachs. (2024, September 11). <https://www.goldmansachs.com/insights/articles/us-house-prices-are-forecast-to-rise-more-than-4-percent-next-year>

Forbes Magazine. (2024, November 8). *Housing market predictions for 2025: When will home prices drop?*. Forbes. <https://www.forbes.com/advisor/mortgages/real-estate/housing-market-predictions/>

Delouya, S. (2024, September 30). *A new report illustrates just how stuck the housing market is | CNN business*. CNN. <https://www.cnn.com/2024/09/30/economy/housing-market-home-sales-redfin-report/index.html>