

Ashley Deibler

Dr. Fadi Alsaleem

DSC 530 Data Exploration & Analysis

Final Project - Week 12

For this project I focused on Texas real estate statistics to try and determine whether square footage of a property was the primary factor in determining listing prices. The variables present in this analysis are listing price, square footage, number of bedrooms, number of bathrooms, and the year the home was built. After completing the EDA, I discovered that square footage does not correlate with listing price, as we can tell using the R-squared value of 0.03.

I feel as though I missed some opportunities to more deeply explore some of the different variables and how they relate to listing price. Some questions that can be explored in a follow-up analysis include:

- Do the number of bedrooms impact listing price?
- Do the number of bathrooms increase with the number of bedrooms?
- By what percentage does property size increase with each additional bedroom?  
Bathroom?
- How does the year in which the property was built affect listing price?

I believe that I could have performed a better analysis of what truly impacts property listing value if I had variables such as location to nearby cities, lot size, view, safety of the neighborhood, and community amenities.

One challenge I faced was determining correlation values, such as Pearson's and Spearman's correlations. I used many resources in order to attempt this, but whether it be due to

lack of understanding of the concepts, or due to technical difficulties, I was unsuccessful.

However I was able to obtain the R-squared value through a regression analysis, which was useful in determining the lack of correlation between listing price and square footage.