Homework assignment 14

Use black text (if possible) for everything you include in this document. Keep both your answers and the original questions. Save this document in PDF format and submit it on Canvas. Include your last name, the course number and the module number in the name of your file.

1. Show a documentation header. The documentation header is a description of who wrote the program, when it was written, what the purpose of the program is (briefly), and what restrictions (if any) that you may place on the program. For SPSS, you can just type the documentation as free format text. For other programs, you might use the comment feature (such as /\* and \*/ in SAS).

2. Download the file [data-05-housing.txt](https://github.com/pmean/classes/blob/master/biostats-1/data/data-05-housing.txt) and review [data-05-housing-dictionary.yaml](https://github.com/pmean/classes/blob/master/biostats-1/data/data-05-housing-dictionary.yaml). You may already have this data from a previous assignment. Import the data into SPSS. Be careful of the missing value codes and convert them from the R language standard (NA) to the SPSS standard for missing (a dot). Display the first ten rows and five columns of data.

3. Calculate the means and standard deviations for price, sqft, and age. Count the number of missing values for these variables. Display your results below.

4. Calculate the correlations (Pearson correlations) among price, sqft, and age. Use the pairwise deletion option. Interpret the three correlations.

5. Draw a scatterplot with price on the y-axis and sqft on the x-axis. Include a linear regression trendline on your plot. Display the plot below. Is this plot consistent with the correlation that you computed earlier?

6. Repeat the previous question with age on the x-axis. Is this plot consistent with the correlation?

7. Calculate a linear regression with price as the dependent variable and sqft as the independent variable. Interpret the slope, intercept, and R-squared.

8. Find the F-ratio in the ANOVA table and the associated p-value. What hypothesis does the F-ratio test?

9. Calculate a 95% confidence interval for the slope and interpret it. Does the confidence interval indicate a positive relationship between price and sqft?

10. Compute and save the residuals (unstandardized residuals) from the previous regression model. Draw a Q-Q plot of the residuals. Display it below and interpret it.

11. Calculate a linear regression with price as the dependent variable and with sqft and age as the two independent variables. Interpret the intercept and the two slopes.