ECON 3161 Homework 2 (Fall 2023)

Due date: September 24 (Sunday, by the end of the day)

Instruction: (1) There are 50 points in total; (2) This HW involves data analysis. You're expected to use some statistical software to conduct your analysis. Stata is highly recommended. However, you can also use other software such as R, SAS or Python. Excel is not recommended here, and HW using Excel will only receive 70% of the grade; (3) If you are off campus, you can use Gatech VLAB to access Stata: https://mycloud.gatech.edu/; (4) Here is an introduction video to Stata: https://www.youtube.com/watch?v=gdnDkjoPJTM&t=3633s. I also posted some tutorial slides and a sample code in the HW folder. I'll go over it in class.

In this HW you'll use the house price data (I have uploaded HPRICE.csv and HPRICE.dta in HW2 folder) to answer the following questions:

1. (10 points) How many variables in this dataset? What is the sample size?
2. (5 points) How many houses are priced between \$300,000 and \$700,000?
3. (15 points) Present the summary statistics of the five variables price, assess, bdrms, lotsize, sqrft, colonial. Here the summary statistics should include mean, standard deviation, min, max, and sample size. You should present your results in a table. I give an example how your table could look like here (note that my example is using a different dataset, and your table could be organized slightly different):

Table 1: Summary Statistics

variables	mean	sd	min	max	obs
wage	957.9455	404.3608	115	3078	935
educ	13.46845	2.196654	9	18	935
exper	11.56364	4.374586	1	23	935
age	33.08021	3.107803	28	38	935
black	.1283422	.3346495	0	1	935
married	.8930481	.3092174	0	1	935
south	.3411765	.4743582	0	1	935
urban	.7176471	.4503851	0	1	935
N	935				

4. (10 points) Plot the histogram of variable price.
5. (10 points) Scatter plot price against bdrms.