Problem Set 5
Chang Gao

Q3

Please see the Jupyter Notebook for details.

I use SelectorGadget and found ".trs_word_table p" for the housing data I wanted from the sample html from website of bureau of statistics of China, however, it was hard for me to turn then into a well-organized csv. I always got one single column or a row.

In my current research, I use javascript to process the monthly htmls that contains the housing data. I downloaded all the htmls by hand, which can be imporved in the future. Then run the following code under node.json environment to extract the tables. There are some errors in display (all red) if I translate the code into English, so I kept the original one.

Each html contains around 4-12 tables. One difficulty in extracting tables is that for some months and some specific table such as price for small, medium or large houses, the table look too long so it was broken into two parts. I address this issue by letting claude use javascript to check whether one table contains name of two specific cities, if not, it should merge the following table.

Then I did data cleaning for these table, I didn't put the code here since they are very long and contains a lot of steps.

I'm interested in China's housing price market. Since housing accounts for over 50% of asset composition in Chinese household finance, rising prices have improved many people's lives while making life more difficult for others. I want to analyze their impact on welfare measures such as income distribution and wealth distribution, as well as the effects of upward and downward price shocks. Is this truly a bubble? I hope my research can remind the government or speculative-investors to avoid past practices.

Q4

I scraped data from Yahoo finance by API and put the data with the city housing prices together.

I've compared China's housing market prices, stock indices, US housing price indices, and stock indices together. Initially it was just for fun, but interestingly, China's housing market returns are similar to the US, while China's stock indices perform poorly, weaker than housing returns. In contrast, US stock indices perform well. From another perspective, do investors believe that only the housing market is worth investing in China, while US investors tend to prefer stock markets? Then why are housing market returns similar between the two countries? Especially given that real estate assets constitute a large portion of household wealth in China.