Goals:

This project is about scraping [Zillow](https://www.zillow.com/homes/New-York,-NY_rb/) webpage and get data for available apartments in New York city. I did the scraping through using Selenium. I used [PyCharm](https://www.jetbrains.com/pycharm/) IDE for python coding.

I extracted 797 apartments from the website. For each apartment, the following items were extracted:

1. Price
2. Number of bedrooms
3. Number of bathrooms
4. Area
5. Address
6. Link to underlaying page
7. Type of home (multifamily, house, condo, …)
8. Year built
9. Heating system
10. Parking
11. HOA
12. Lot
13. Principal and interest
14. Property tax
15. Insurance
16. Nearby schools
17. Ranking for nearby schools
18. GPS coordinate



1. Develop a database for home available in the market
   1. Home owners: how much my property worths?
   2. Agents: What items are available in the market?/What is the price range for any specific location?/
   3. Architect: How many bedrooms/bathrooms should I design for the upcoming project? Does it worth to assign a parking? What is the balance between number of bedrooms/bathrooms and area?
2. Qualitative study on the effect of different factors on the price