# Data for “Rent”

U.S. rental market on Craigslist over a given month

# Table of Content

# Background

The data is downloaded from Kaggle:
https://www.kaggle.com/datasets/austinreese/usa-housing-listings

I remember during college and early career I always struggle to find a perfect place to rent. Now with my data analytic skills, I can look deep into the rental of houses, apartments, or even single rooms across the country. Therefore, I find this dataset of Craigslist posts of all rentals for a given month, It is definitely not a perfect dataset and I wish there could be more valuable data added to it. However, for the purpose of this project, it is good enough to work with.

The conclusions drew from this dataset might be different from nowadays common senses as there is no time data and it might be outdated and out of updates. All in all, everything serves for my skill demonstration.

# Questions & Assumptions

1. Which state has the highest average rental?

   *California or New York from my best guess*

2. Are California rentals pet-friendly?

   *Yes, I see people walking pets around all the time*

3. What type of house contributes the most of rental market?

   *Either house or apartment*

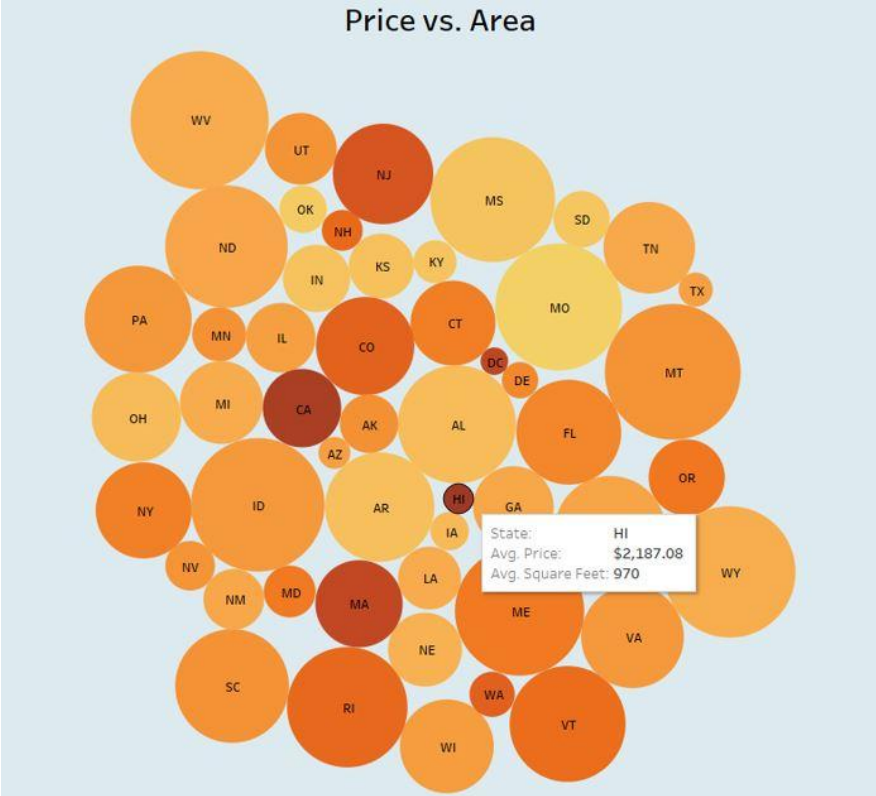
4. Is my current rental above or below average California market?

   *My current rental is a 4 beds 3 baths, I guess my price is below average*

# Methods

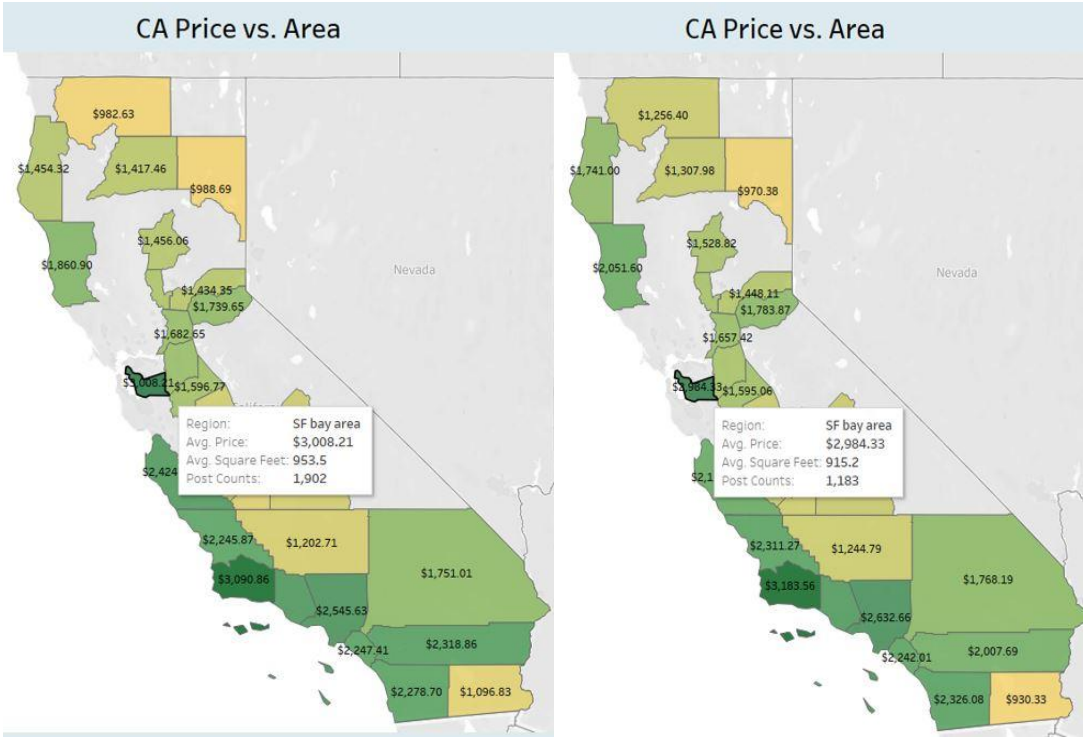
1. Download data from Kaggle.com
2. Import data to MySQL server
3. Perform data cleaning in MySQL workbench
4. Export data to local file for Tableau visualization

# Viz of price vs. area by state



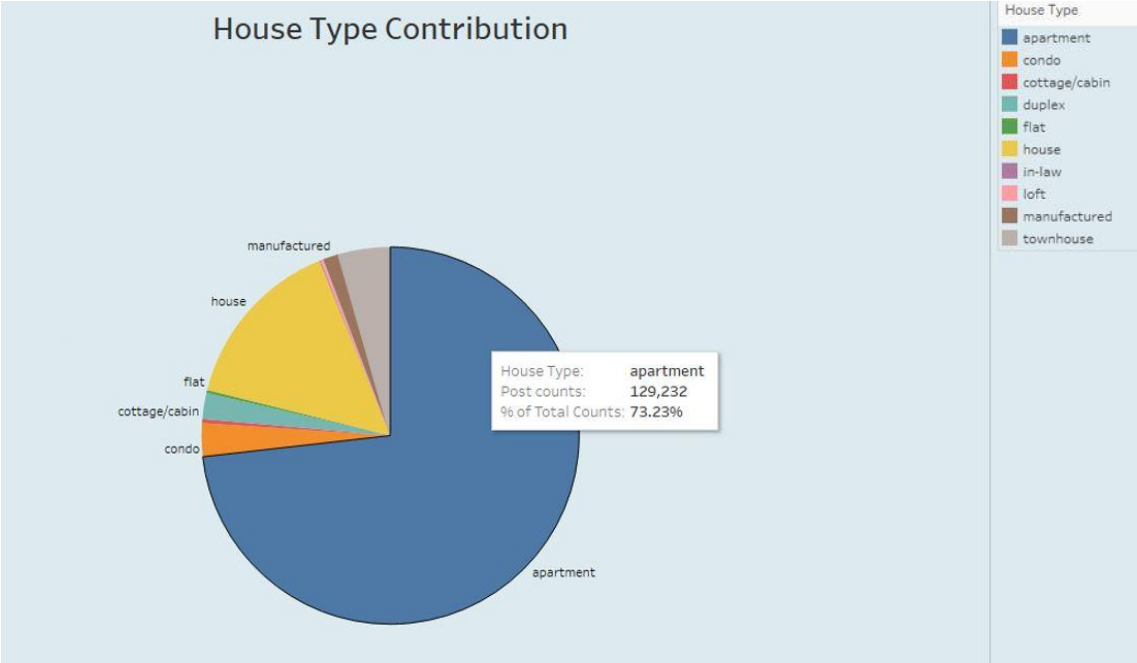

This Viz shows the average price and average square feet in each state. The bigger size of the circle, the larger area it represents. The darker color indicates the more expensive states. California is quite dark but to my surprise, Hawaii has the most expensive price among across the country.

# Viz of California rental information




After applying filters of cats allowed and dogs allowed, we have 1183 out of 1902 properties which allow pets. It is roughly 60% which is pretty good!

# Viz of house type contribution




This Viz clearly show that almost 73% of the rental type is apartment.

# Viz of price by bedroom and bathroom


From this Viz after applying the bathroom filter, I am very sure that this data was from 5 to 10 years back or it is just not accurate enough. However, considering nothing else but the data, my current rental is way far above the average!

# Reviews

1. Which state has the highest average rental?

   **To my surprise, Hawaii is more expensive than California back in the day.**

2. Are California rentals pet-friendly?

   **Even though the data is outdated, 60% pet-friendly is still pretty good!**

3. What type of house contributes the most of rental market?

   **I think even nowadays apartments are still the most popular in rental market.**

4. Is my current rental above or below average California market?

   **Unfortunately, I am paying way more than I should, if I was in the timeline of the data.**

# What To Improve

First of all, I need to choose my data more wisely in the future. I think rather than getting modified data from Kaggle, I might want to get data from websites like data.gov or census.gov. The data might look a lot more raw at the beginning, but I believe the process of cleaning data is very crucial to data analysis.

Furthermore, I did not use JOIN in data cleaning as there was only one table. In my next project, I will find datasets with multiple tables so that I could showcase my understanding of JOINs.

I also found a mistake during importing data to mySQL server. The column of “baths” in the raw data contains floating numbers instead of integers, because there are half bathrooms in a lot of houses. I need to observe the data more carefully.

All together, it is a pretty lousy project. As a self-taught beginner, I tried hard to put everything together. Hopefully it will be a good start for my future analysis career.