



# Real Estate Market Analysis

Project Report - Spring 2023  
BAN 612 - Data Analytics  
**Group-3**

Feb 26, 2023





**Team Introduction**



**Project Overview**



**Data Collection**



**Exploratory Data  
Analysis**



**Predictive Analysis**

## Project Team Members



CAL STATE  
EAST BAY



Parthiv Patel  
NetId - hk9113



Swati Sharma  
NetId - xn2486



Megh Dave  
NetId- cc3118



Sam Thomas  
NetId - my4704



Akhil  
NetId - to6616

## Technologies



## Problem Statement

The housing price or rental price in San Francisco Bay Area accelerating at shocking rates. From 2010 to 2017, the median price of a single family home in San Francisco has gone from approximately & 775,000 to \$1.5 million. This project presents detail analysis on the real estate in San Francisco Bay area region.

- To collect information on more than 600 distinct properties from Redfin Real Estate website
- To clean, prepare, organize and handle the missing data, outliers, incorrect record and so on
- Prepare a statistical report to analyse the current trend in the market for real estate
- Propose several research questions and use visualization techniques to identify the researched questions

## Data Resources

- Redfin Real Estate - <https://www.redfin.com>

## Attributes



State



Price



Status



Area



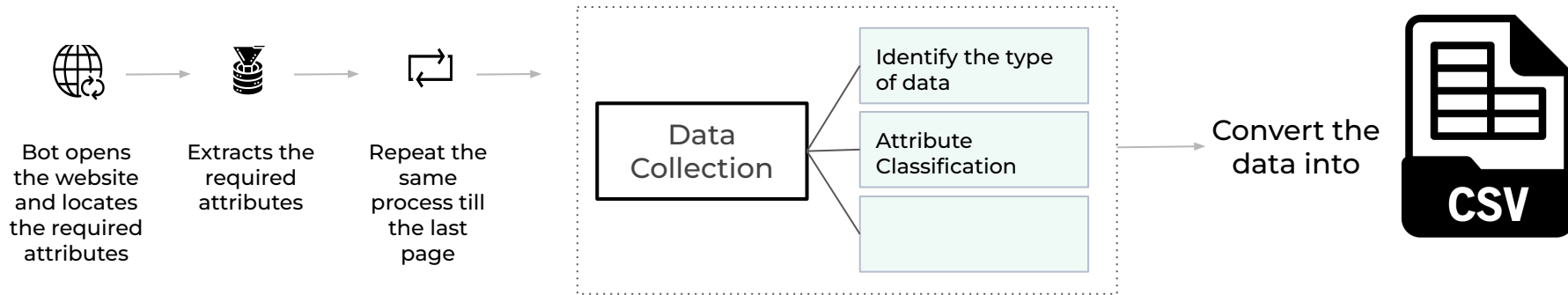
# Part 1: Data Collection



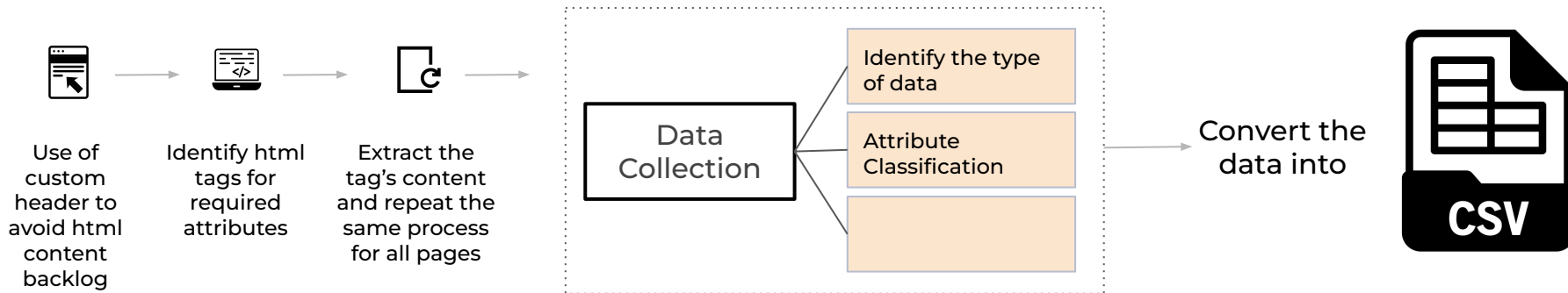


# Flowchart - Web Scraping Steps

## Selenium Automation



## HTML Parser





## Data Cleaning Process:

- We extracted over **1200** data from six different county having six different pages. After cleaning, the total data boiled down to around **430**.
- Data is cleaned by removing null values, by making function to split all alphanumeric data into categorical and quantitative columns for exploratory analysis and prediction.

| Column Attributes | Before Cleaning | After Cleaning |
|-------------------|-----------------|----------------|
| Price             | 598,000         | 598000.0       |
| Lot_Area          | 5,200 Acres     | 5200.0         |
| HOA Due           | \$449           | 449.0          |
| Redfin_estimate   | 498,286         | 498286.0       |



## Final Result

|    | State | County     | City         | Street        | Zipcode | Property_Typ | Status | Price      | Year_Built | Beds | Bath | Sqft | Acre       | Lot Area      | HOA Due     | School_ratin | WalkScore | Bike_Score | Garage_avai | Redfin_estimate |
|----|-------|------------|--------------|---------------|---------|--------------|--------|------------|------------|------|------|------|------------|---------------|-------------|--------------|-----------|------------|-------------|-----------------|
| 0  | CA    | Sonoma-Cou | Santa Rosa   | 1830 Leafgre  | 95405   | Single Famil | Active | 750,000    | 1961       | 3    | 2    | 1425 | 1,425Sq Ft | 7,022 Sq. Ft. |             | 6            | 58        | 57         | Yes         | 752,455         |
| 1  | CA    | Sonoma-Cou | Sonoma       | 20745 Denm    | 95476   | Multi-Family | Active | 11,500,000 | 2022       | 4    | 5.5  | 6005 | 6,005Sq Ft | 4.58 Acres    |             |              | 27        | 38         | Yes         |                 |
| 2  | CA    | Sonoma-Cou | Glen Ellen   | 6409 - 6411   | 95442   | Multi-Family | Active | 3,295,000  | 1983       | 6    | 5    | 4078 | 4,078Sq Ft | 22.68 Acres   |             | 5            | 0         | 7          | No          |                 |
| 3  | CA    | Sonoma-Cou | Santa Rosa   | 121 Oakwoo    | 95409   | Mobile/Man   | Active | 130,000    | 1971       | 2    | 2    | 960  | 960Sq Ft   |               |             | 7            | 56        | 64         | No          | 128,174         |
| 4  | CA    | Sonoma-Cou | The Sea Ran  | 487 Drovers   | 95497   | Single Famil | Active | 1,175,000  | 1993       | 2    | 2    | 1849 | 1,849Sq Ft | 0.73 Acres    | \$292/month | 4            | 0         | 3          | Yes         | 1,186,436       |
| 5  | CA    | Sonoma-Cou | Petaluma     | 214 Jacquely  | 94952   | Single Famil | Active | 1,425,000  | 2010       | 4    | 2.5  | 3141 | 3,141Sq Ft | 7,148 Sq. Ft. |             | 8            | 32        | 13         | Yes         | 1,431,730       |
| 6  | CA    | Sonoma-Cou | Rohnert Park | 4579 Harmoi   | 94928   | Single Famil | Active | 700,000    | 1986       | 3    | 2    | 1470 | 1,470Sq Ft | 4,500 Sq. Ft. |             | 6            | 30        | 64         | Yes         | 701,179         |
| 7  | CA    | Sonoma-Cou | Santa Rosa   | 55 Coronado   | 95409   | Mobile/Man   | Active | 99,000     | 1966       | 2    | 2    | 1040 | 1,040Sq Ft |               |             | 7            | 21        | 39         | No          | 97,785          |
| 8  | CA    | Sonoma-Cou | Santa Rosa   | 510 Todd Rd   | 95407   | Multi-Family | Active | 1,250,000  | 1985       | 5    | 2    | 1080 | 1,080Sq Ft | 19 Acres      |             | 4            | 19        | 32         | No          | 1,241,725       |
| 9  | CA    | Sonoma-Cou | The Sea Ran  | 37238 Schoo   | 95497   | Single Famil | Active | 859,000    | 1988       | 1    | 1.5  |      | 1,425Sq Ft | 0.61 Acres    | \$292/month | 4            | 0         | 4          | Yes         | 857,863         |
| 10 | CA    | Sonoma-Cou | Sonoma       | 22033 Bonne   | 95476   | Single Famil | Active | 4,500,000  | 1989       | 4    | 4    | 3309 | 3,309Sq Ft | 10 Acres      |             | 5            | 13        | 29         | Yes         | 4,493,606       |
| 11 | CA    | Sonoma-Cou | Petaluma     | 1619 Northst  | 94954   | Single Famil | Active | 700,000    | 1986       | 3    | 2.5  | 1421 | 1,421Sq Ft | 4,400 Sq. Ft. | \$69/month  | 7            | 52        | 47         | Yes         | 700,119         |
| 12 | CA    | Sonoma-Cou | Petaluma     | 30 Acorn Cir  | 94952   | Townhouse    | Active | 545,000    | 1983       | 2    | 1.5  | 1242 | 1,242Sq Ft | 2,431 Sq. Ft. | \$620/month | 7            | 41        | 35         | Yes         | 545,717         |
| 13 | CA    | Sonoma-Cou | Santa Rosa   | 3664 Fir Ridg | 95403   | Vacant Land  | Active | 430,000    | 1983       |      |      |      | 0.27       | 0.27 Acres    | \$75/month  | 6            | 3         | 6          | No          | 427,766         |
| 14 | CA    | Sonoma-Cou | Santa Rosa   | 3767 Saint A  | 95403   | Single Famil | Active | 2,500,000  | 2020       | 4    | 4.5  | 3560 | 3,560Sq Ft | 0.32 Acres    | \$75/month  | 5            | 3         | 6          | Yes         |                 |
| 15 | CA    | Sonoma-Cou | Santa Rosa   | 336 Pythian I | 95409   | Single Famil | Active | 995,000    | 1978       | 3    | 2    | 2029 | 2,029Sq Ft | 8,281 Sq. Ft. | \$113/month | 7            | 11        | 33         | Yes         |                 |
| 16 | CA    | Sonoma-Cou | The Sea Ran  | 100 Lupine C  | 95497   | Vacant Land  | Active | 179,000    | 1978       |      |      |      | 0.66       | 0.66 Acres    | \$292/month | 4            | 0         | 14         | No          |                 |
| 17 | CA    | Sonoma-Cou | Petaluma     | 433 Garfield  | 94954   | Single Famil | Active | 795,000    | 1969       | 4    | 2    | 1882 | 1,882Sq Ft | 6,242 Sq. Ft. |             | 7            | 27        | 52         | Yes         | 798,524         |
| 18 | CA    | Sonoma-Cou | Sonoma       | 180 Serres D  | 95476   | Single Famil | Active | 3,895,000  | 2005       | 4    | 3.5  | 2369 | 2,369Sq Ft | 1.04 Acres    |             | 5            | 42        | 40         | Yes         | 3,870,615       |
| 19 | CA    | Sonoma-Cou | Kenwood      | 8099 Hwy 12   | 95409   | Single Famil | Active | 2,549,000  | 1947       | 4    | 4    | 2662 | 2,662Sq Ft | 2.94 Acres    |             | 7            | 11        | 34         | No          | 2,523,686       |
| 20 | CA    | Sonoma-Cou | Santa Rosa   | 580 Los Alar  | 95409   | Single Famil | Active | 695,000    | 1986       | 3    | 2    | 1761 | 1,761Sq Ft | 0.56 Acres    |             | 9            | 12        | 21         | Yes         | 698,068         |
| 21 | CA    | Sonoma-Cou | Healdsburg   | 2707 N Fitch  | 95448   | Single Famil | Active | 1,349,000  | 1930       | 3    | 3    | 1753 | 1,753Sq Ft | 0.37 Acres    |             | 7            | 6         | 12         | Yes         | 1,352,552       |
| 22 | CA    | Sonoma-Cou | Santa Rosa   | 2013 Elson V  | 95403   | Single Famil | Active | 650,000    | 2022       | 3    | 2.5  | 1808 | 1,808Sq Ft |               |             | 3            | 66        | 57         | Yes         | 654,073         |
| 23 | CA    | Sonoma-Cou | Petaluma     | 340 Wilson S  | 94952   | Single Famil | Active | 699,000    | 1900       | 2    | 1    | 838  | 838Sq Ft   | 4,051 Sq. Ft. |             | 7            | 91        | 54         | Yes         | 704,512         |