

CSC 805 Data Visualization
Phase 3 Project Implementation
Fall 2023

Visualizing the Housing Landscape
in San Francisco

Group 3
Khalid Mehtab Khan
Joanne Pak
Durga Siva Lokesh Telaprolu



San Francisco Housing Landscape



Team Members:
Khalid Mehtab Khan
Joanne Pak
Durga Siva Lokesh Telaprolu

I. Purpose

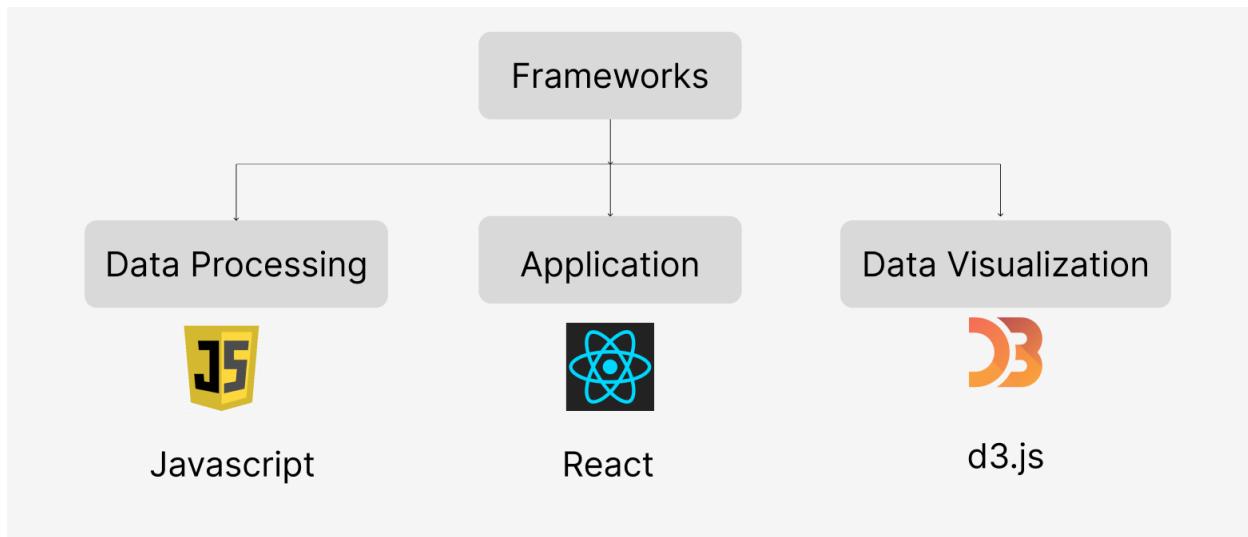
Choosing housing arrangements requires taking into account various factors, especially in a vibrant city like San Francisco. The city has evolved over the years, making it a popular tourist destination and attracting a diverse population. However, this has created a major problem for the small city, overpopulation and other factors have caused housing prices to hike up immensely over the past few years.

For our project, the goal was to find the average home values and rent of each zip-code area within San Francisco. We created an interactive dashboard that provides users with key factors that shape the competitive San Francisco housing market.

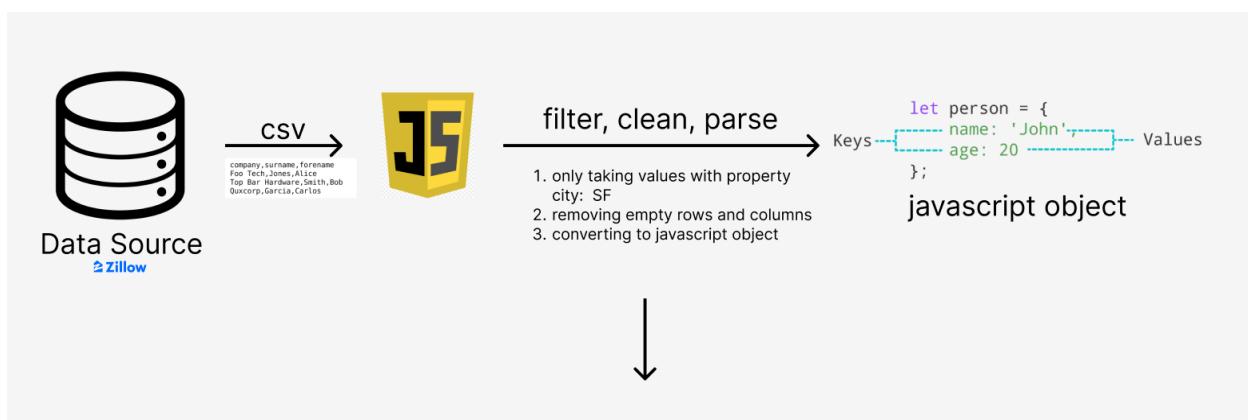
II. System Architecture

(figma link: <https://www.figma.com/file/rmoP0KTN5FPGTRdLj5HPpK/Untitled?type=design&node-id=0-1&mode=design&t=f8V2nyShxdrZWhfq-0>)

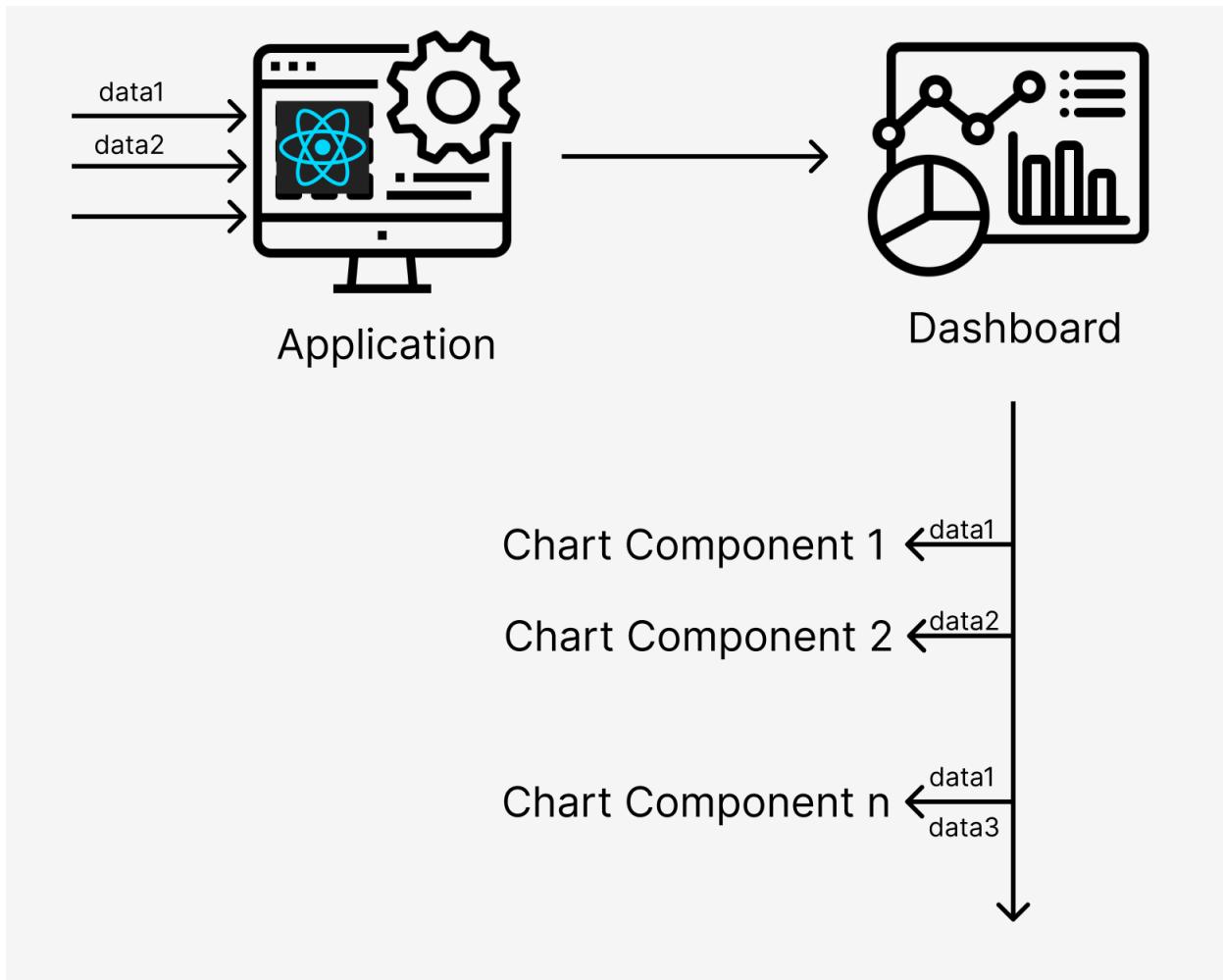
Framework:



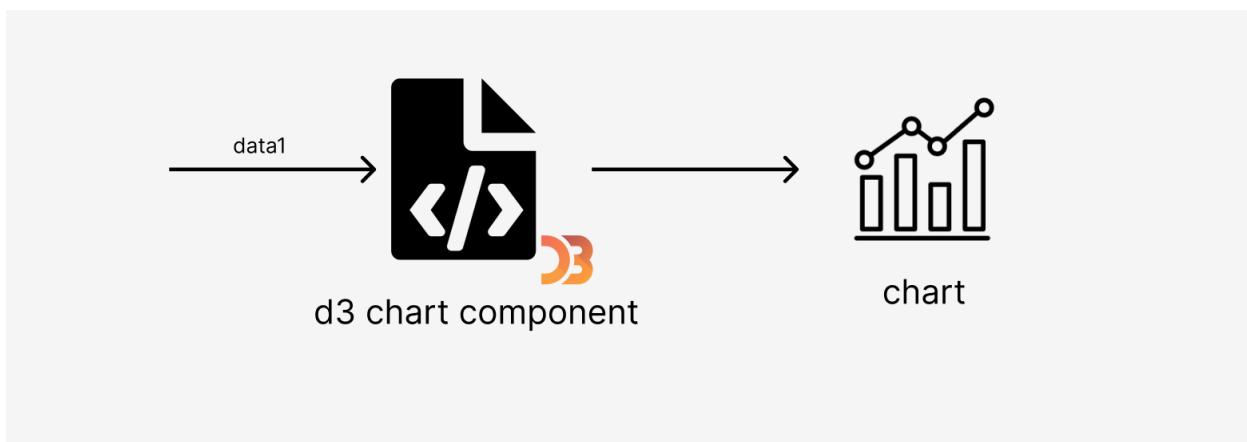
Data Processing:



Application:



Data Visualization:



III. Data Set Description

We used a few different types of datasets for this project. There were many factors that could affect housing prices such as crime, population, age, income, education, and energy.

The Zillow dataset provided us with many time series of home value index and rentals. We've extracted the time series for condos, family homes, and homes with 1 to 5 and above bedrooms of only the zipcodes within the San Francisco area.

The SF crime data allows us to see a time series of different crime types in different areas within SF.

Dataset:

- Housing Data on Zillow
 - <https://www.zillow.com/research/data/>
- SF Crime Data from sfgov.org
 - <https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-2018-to-Present/wg3w-h783>
- Average income, age, education from Simple Maps
 - <https://simplemaps.com/city/san-francisco/zips/age-median>
- Education data from GreatSchools
 - <https://www.greatschools.org/california/san-francisco/schools/?distance=5&gradeLevels%5B%5D=h&lat=37.7605&locationType=city&lon=-122.443>

IV. System Description

Introductory page

The screenshot shows a dark-themed web page with white text. At the top, it says "Visualizing the Housing Landscape in San Francisco". Below that is a brief description: "Discover the intricate landscape of San Francisco's housing market. Dive into key factors that shape the city's living experience." A section titled "Key Factors:" lists six items with icons: Crime (person walking), Population (two people), Age (person icon), Income (dollar sign), Education (graduation cap), and Energy (lightning bolt). Another section titled "Data Sources:" lists three sources: SFData, Zillow, and Kaggle. A "Data Analysis:" section contains a paragraph about exploring data from various perspectives and drawing conclusions. Finally, a "Project Goals:" section states the aim is to visualize complex housing data in an accessible way, providing valuable insights.

Visualizing the Housing Landscape in San Francisco

Discover the intricate landscape of San Francisco's housing market. Dive into key factors that shape the city's living experience.

Key Factors:

- Crime
- Population
- Age
- Income
- Education
- Energy

Data Sources:

- SFData
- Zillow
- Kaggle

Data Analysis:

Through in-depth analysis, we explore the data from different perspectives and pivots. This is an exploratory data visualization dashboard that revolves around neighbourhoods and allows users to interact with the data and draw their own conclusions. Our aim is to enable uncover patterns, trends, and correlations that contribute to a comprehensive understanding of the housing in San Francisco. The model can be extended to other cities and regions. The modular behaviour of our code allows for easy integration of new data sources and data types.

Project Goals:

Our primary goal is to visualize and communicate complex housing data in an accessible and meaningful way. By leveraging interactive charts and visualizations, we aim to provide valuable insights to both researchers and the general public.

For our project, the user would be able to view house pricing trends along with the different key factors for each neighborhood and zip code area within San Francisco. We created an interactive dashboard that is straight-forward and user-friendly.

The first visualization is a time series of housing prices for different housing units using an abstract circular calendar chart. The legend displays different colors for each neighborhood and the user will be able to compare prices with other neighborhoods in a single year. There are two drop-down lists for the user to select a certain year and the type of housing unit which ranges from condos to a 5 bedroom unit. The user can also hover over the spider chart to see the tooltip of the region name and the average housing price value.

Next, we have a map of comparisons using a choropleth map of the zip code areas in San Francisco. Here, the user can explore which areas are more active for a given attribute. The attributes list includes crime rate, average median income, average age, population, population density, percentage of college

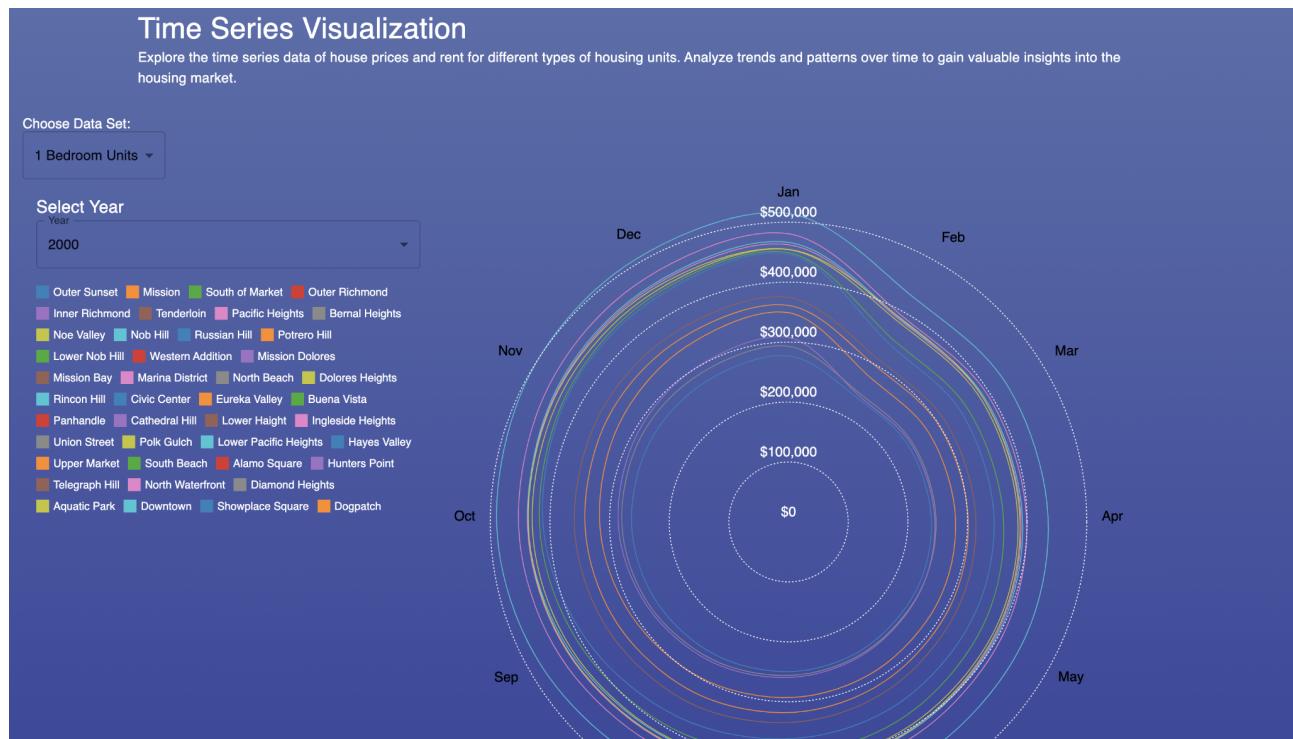
graduates, and active enrollment in CleanPowerSF. Multiple attributes can be dragged and dropped onto the two maps and are easily removable.

Then, we have a stacked line chart for the time series of crime in zip code areas. There is a drop-down list to select a crime type which will show the trends of each area from 2018 to 2023. The stacked line chart also has a tooltip for the year/month which compares the number of crimes committed in each zip code area.

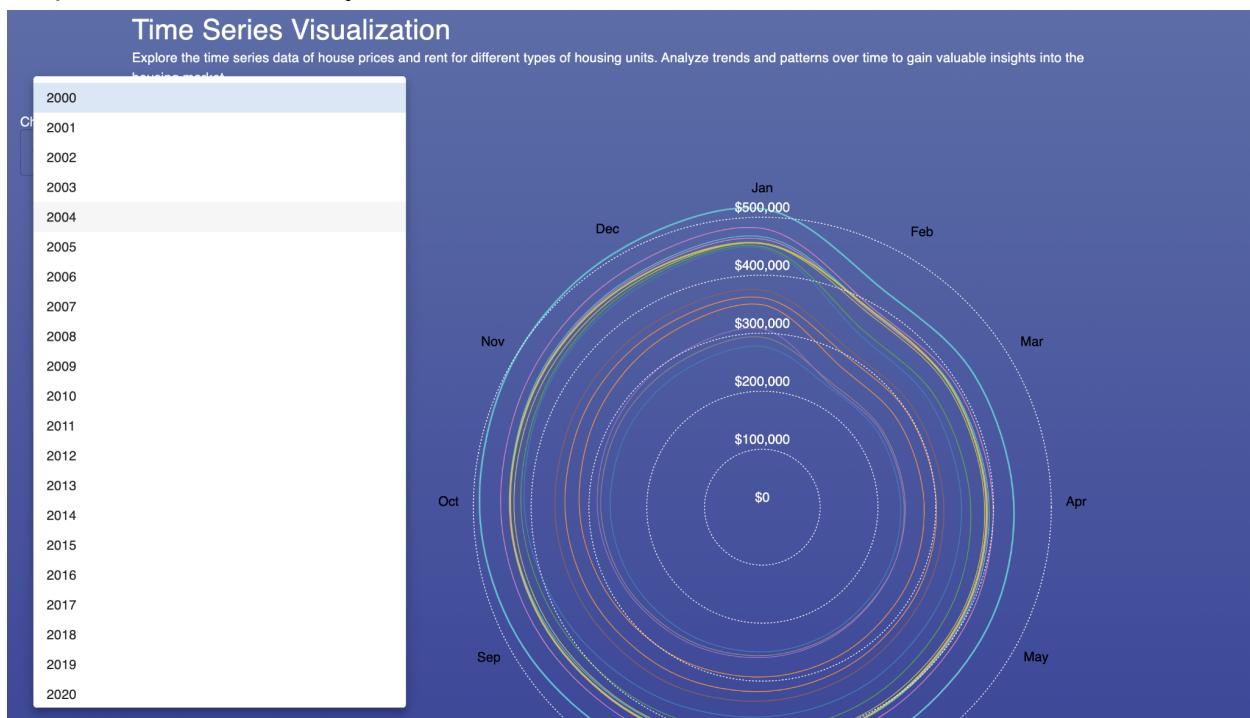
Finally, we have a stacked bar chart which displays the population and rent for each zip code area along with a drop-down list to select the year. The blue bars represent the population number while the orange bars represent the average rent. When hovered over a colored bar, we will see a tooltip with the average rent value.

Screenshots:

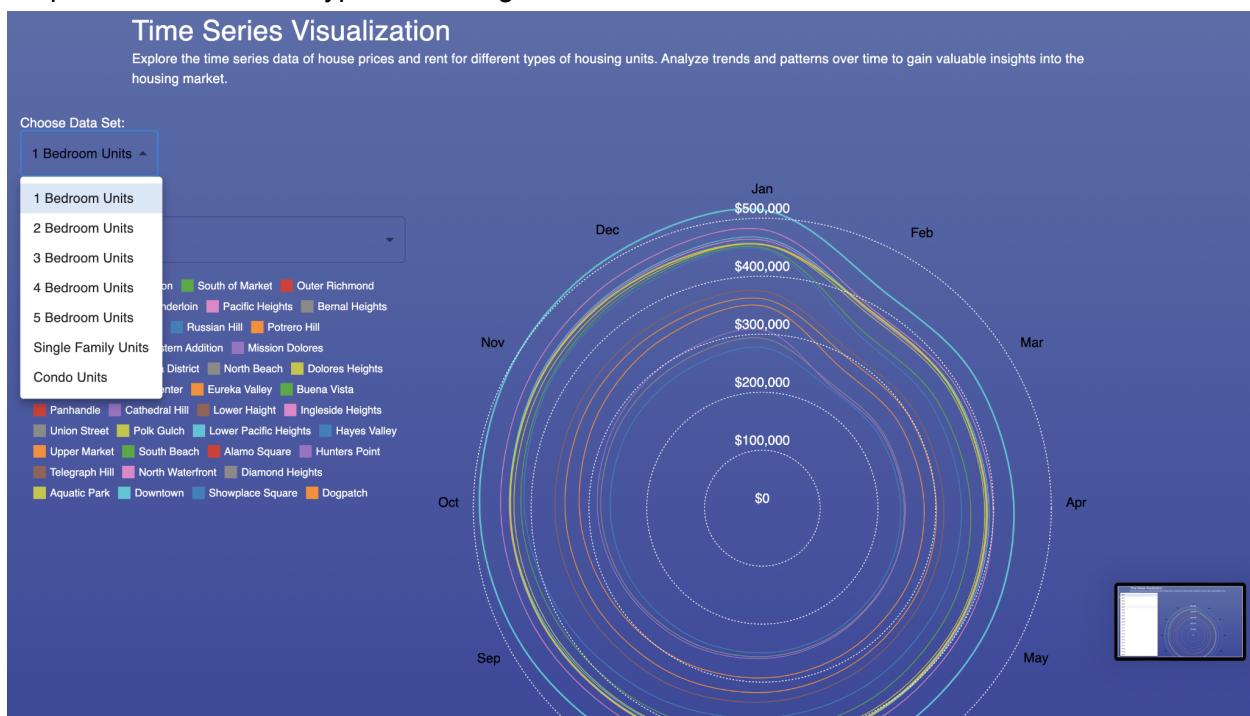
Circular Calendar chart for the time series of housing prices in the neighborhoods of SF



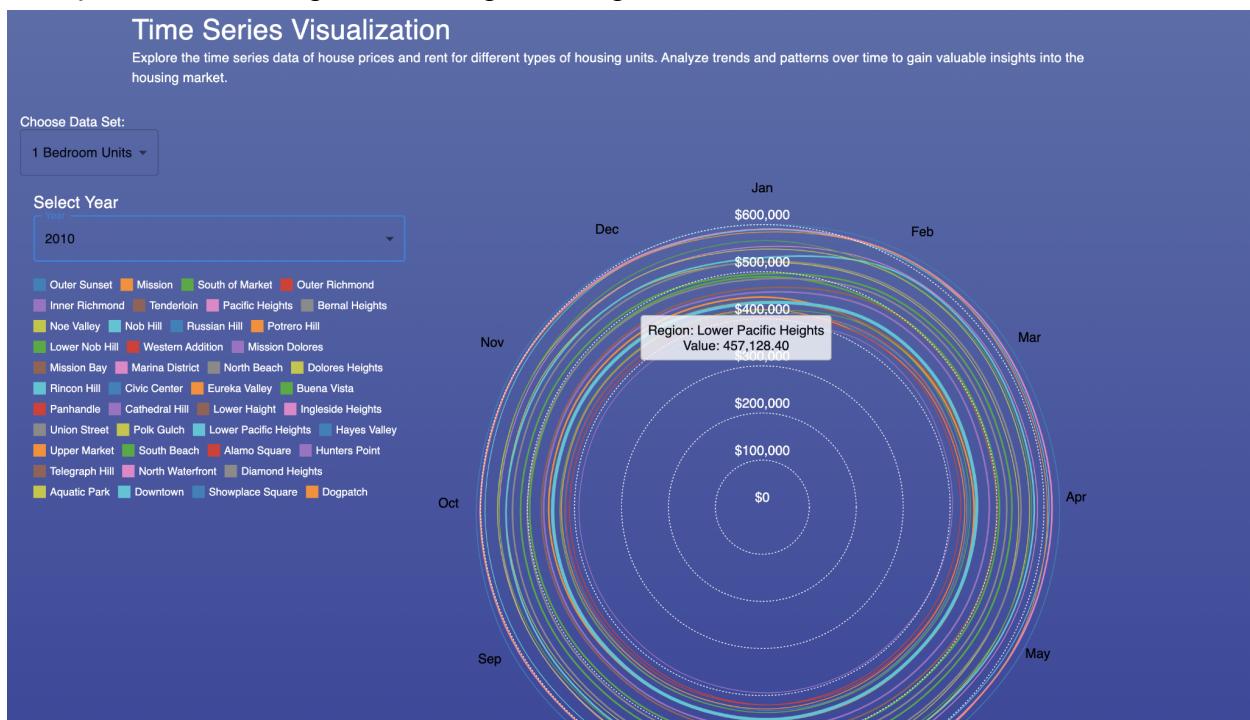
Drop-down list to select year.



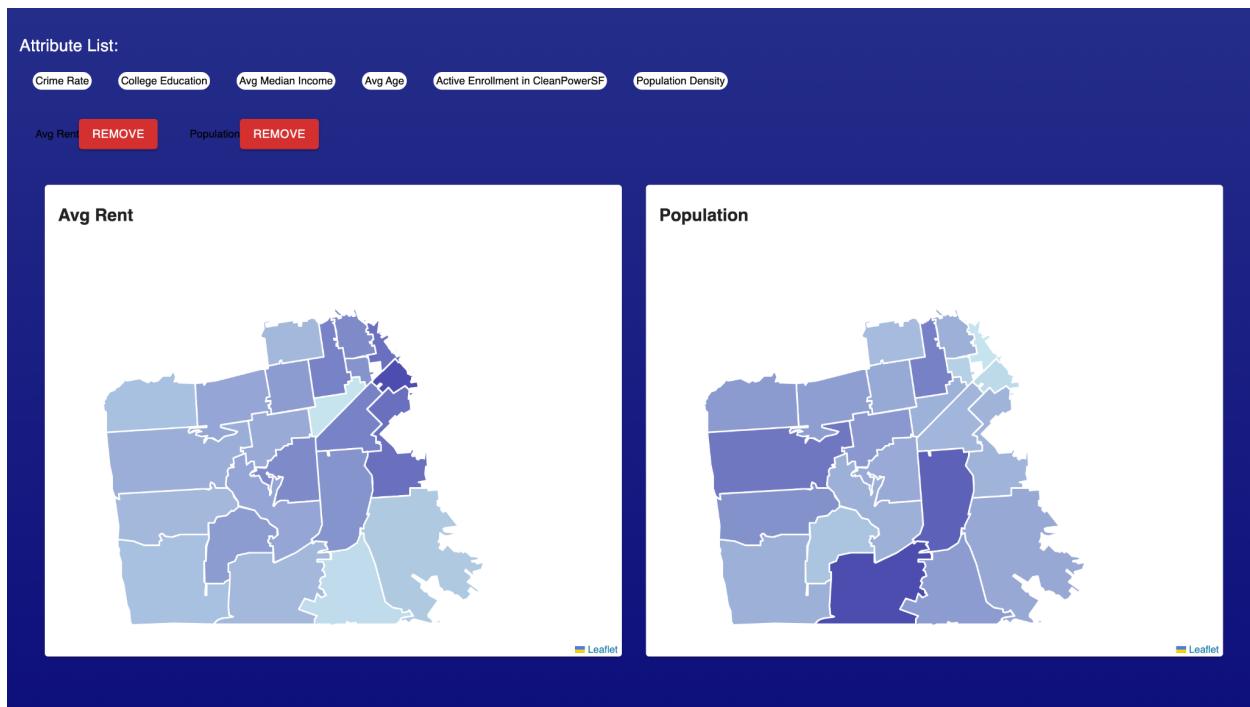
Drop-down list to select type of housing.



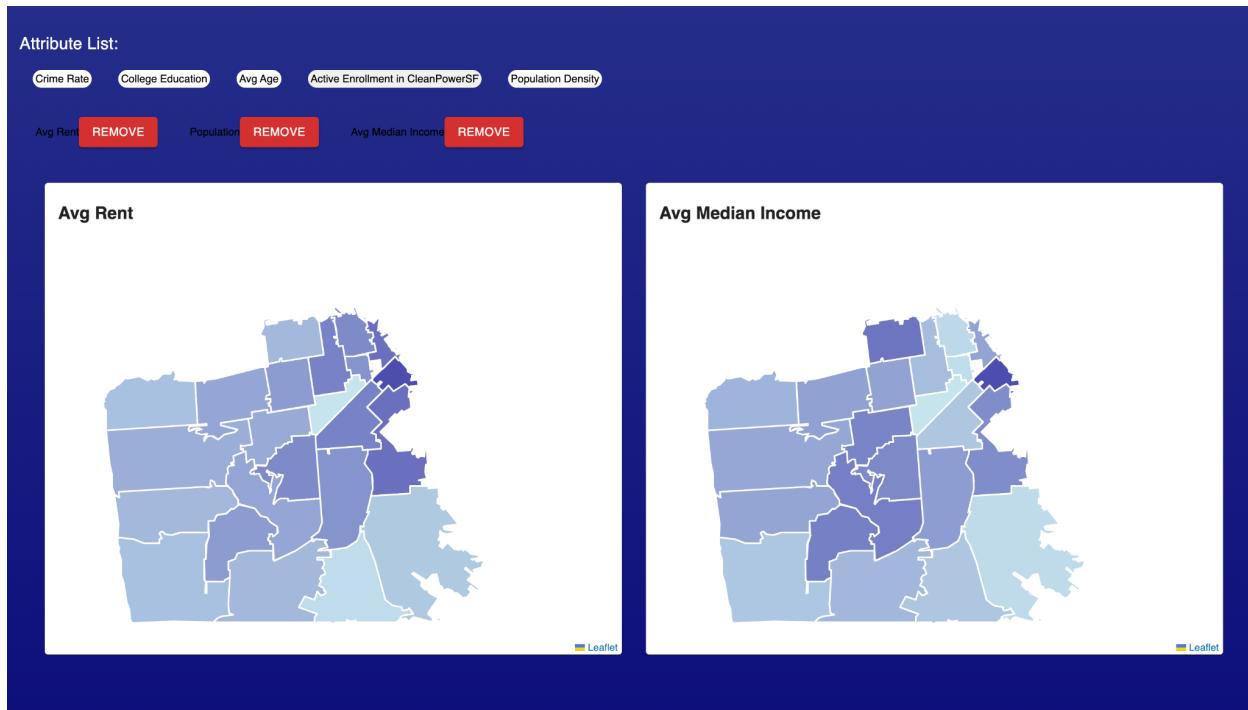
Tooltip to view housing value of a given neighborhood.



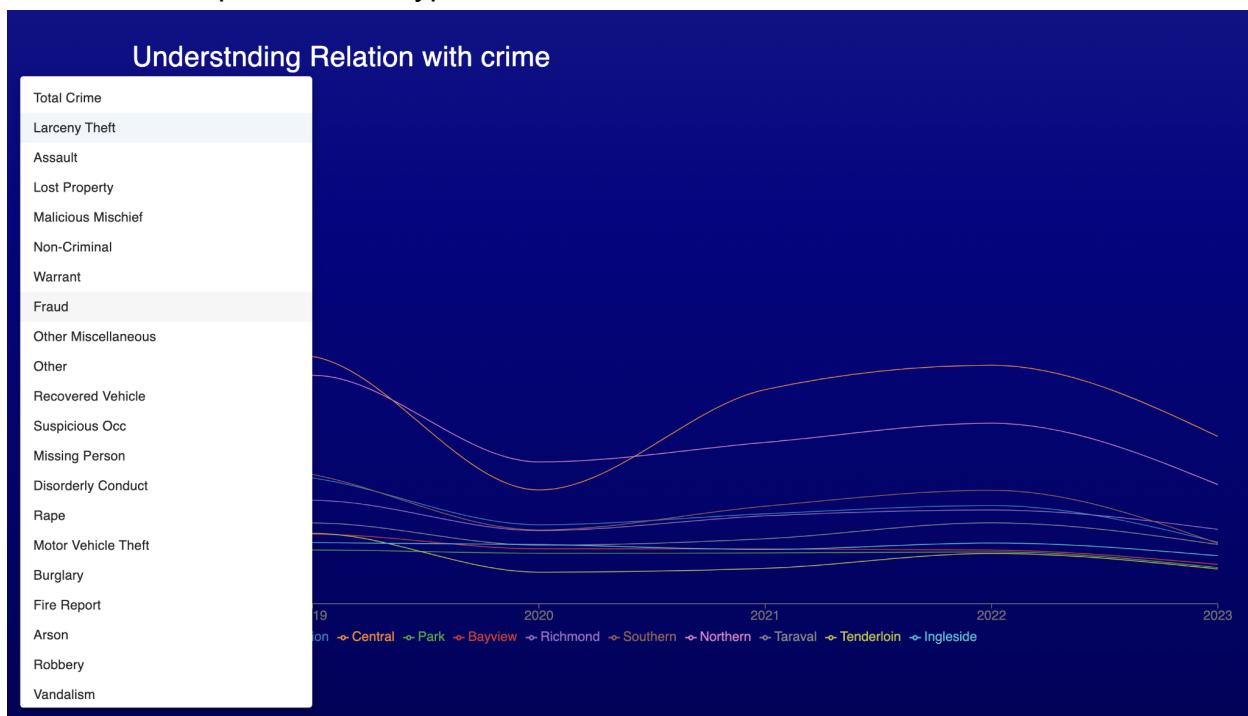
Choropleth map to compare key factors.



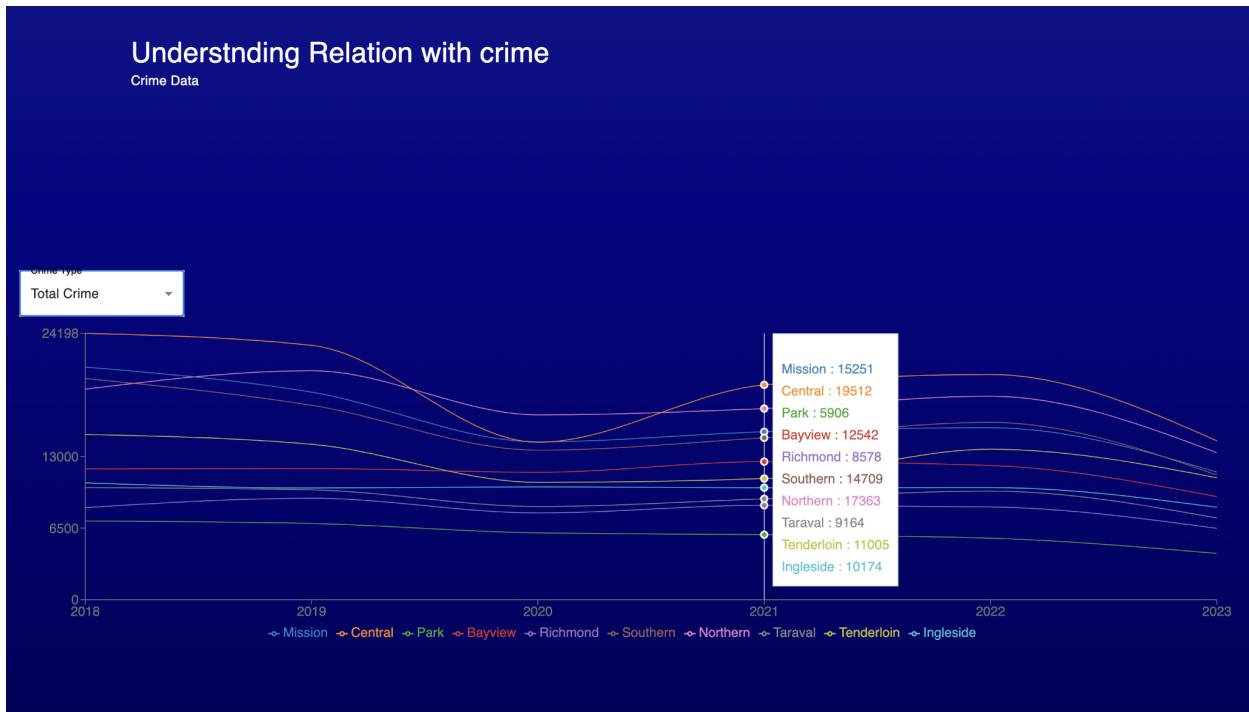
User can drag and drop the attributes onto the choropleth maps, along with the option to remove them.



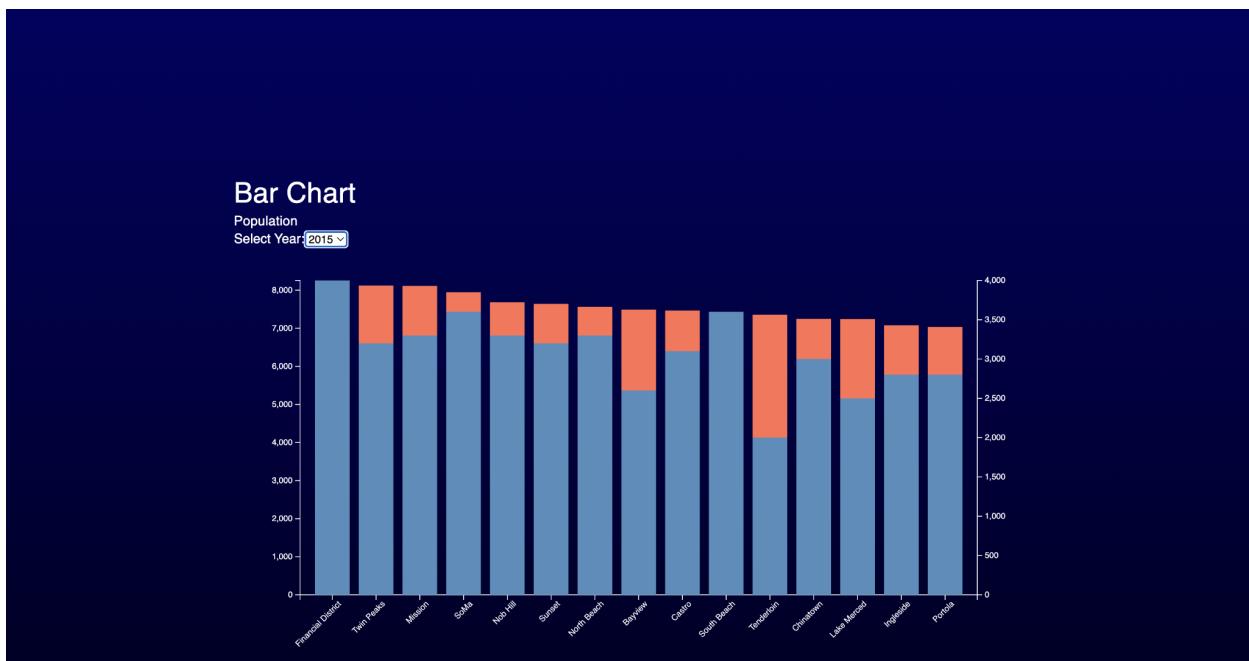
Stacked line chart for the crime time series in different areas in SF. Drop-down list to select total or specific crime type.



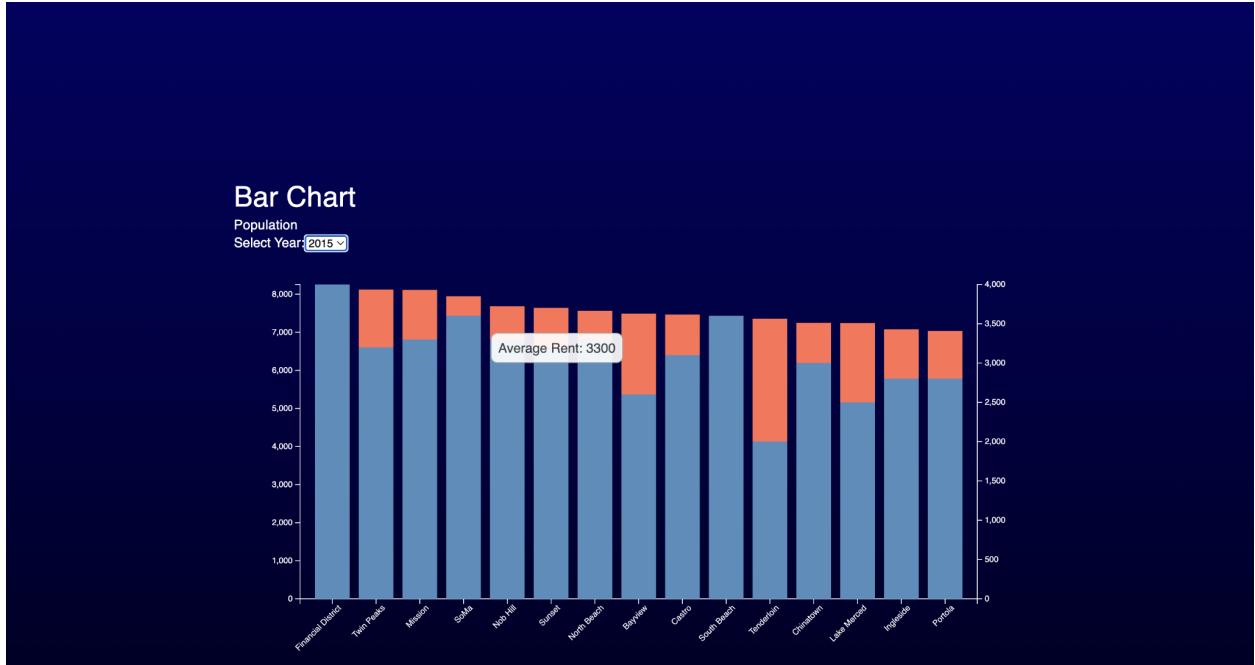
Hover for tooltip to show comparison of the areas with the number of crimes committed for the selected crime.



Stacked bar chart of the population and average rent in given area.



Hover for tooltip to view average rent value.



V. Demo Video and Github Repo

Github link

<https://github.com/Kahl-d/DataWizard>

Demo Video:

https://drive.google.com/file/d/1GALbvpyFRbwwm_pOHgpVs9bqmTpdyG1/view?usp=sharing