

# Analysis of the San Francisco Housing Development Pipeline

Thomas Su

## Table of Contents

Table of Contents .....	1
Executive Summary .....	1
Data Report .....	2
Findings .....	3
Conclusion .....	7
Appendix .....	8

## Executive Summary

The Housing Crisis in San Francisco has been a very contentious issue for the last few decades in the city. The trajectory of house prices in San Francisco have more than doubled in the last 10 years with no end in sight.<sup>1</sup> Many existing residents are being driven out of their homes and many are unable to move to the city because of its prohibitive entry price. Simply put, housing prices in SF are so high because demand far outstrips supply. The analysis of the San Francisco Development Pipeline is important to be able identify deficiencies in city planning to address this crisis.

---

<sup>1</sup> <https://www.bayareamarketreports.com/trend/3-recessions-2-bubbles-and-a-baby>

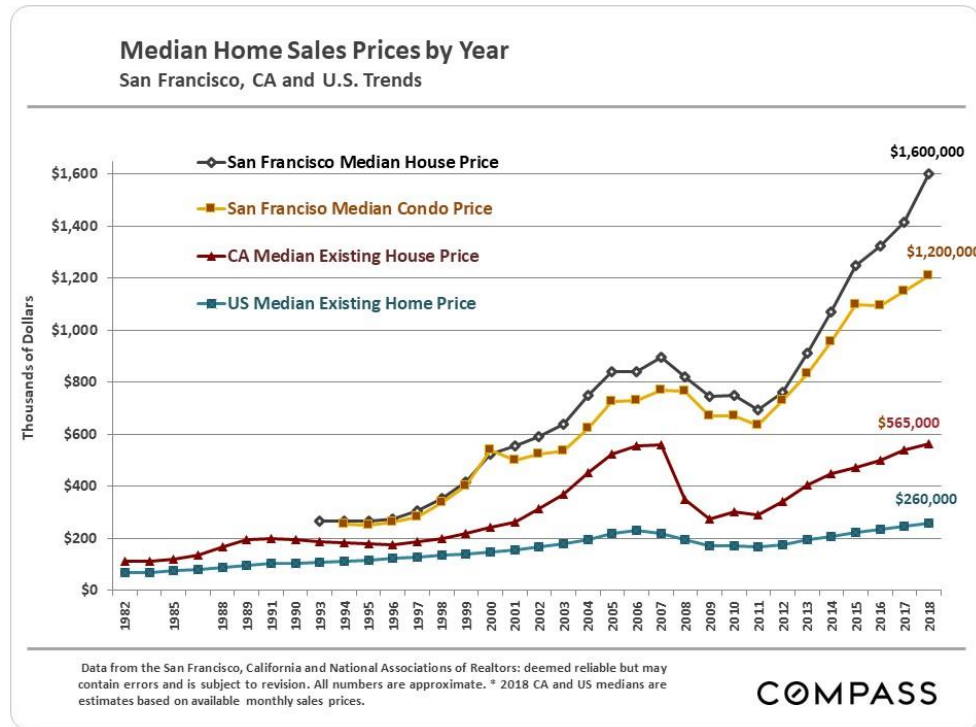


Figure 1: Housing Prices in San Francisco Relative to the Rest of the USA

From analysis of the dataset, we discover that there is a strong correlation between the neighborhoods with development permits approved and income. We find that low-income neighborhoods have more housing permits approved and high-income neighborhoods have less housing permit approved. Additionally, our study identifies the most influential individuals to contact that have the largest influence of permit approval and permit submission.

## Data Report

All data used in this project is from <https://sfplanning.org/project/pipeline-report#about>. The website and data are from the San Francisco Planning department. The data I used for my study is obtained specifically from approved building permits. As a result, we can expect much of the data to be without errors. Upon visual inspection, I could not find any input errors in the data.

There are many unfilled portions of our data. Many fields/columns in our table had to be set to null because not all the field in the housing applications were filled. For example, all the dimension tables do not have one to one mapping to the fact table.

All the data I used for this project came from one single Excel sheet. To create a relational database, I manually segmented the data. In Figure 2 we can see the relationship diagram for our data. It is a simple, traditional start schema. The center of the database is the data fact table. This table contains all the specifics of the housing project. This includes square footage of the building, the number of units, the address of the building and much more. The contact dimension table contains the contact information of the person who is requested the building permit. The case dimension table contains the any pending cases attached to a permit

that need to be resolved. The supervisor dimension table contains information about the supervisor for the district in where the building will be built/modified.

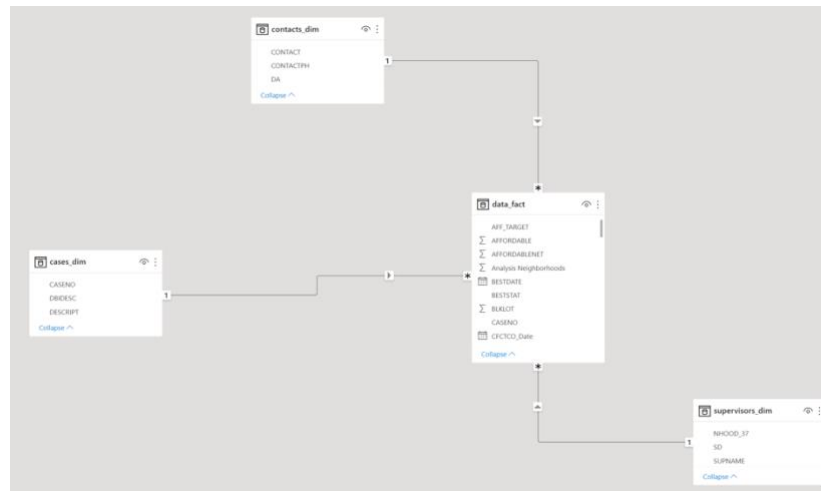


Figure 2: Relationship Diagram

Table 1 explains some of the key features used in this study.

Feature/Column Name	Explanation
UNITSNET	Net added Dwelling Units
SUPNAME	Project Sponsor (Used as primary key in supervisors_dim table)
CASENO	Planning Department Case Number (Used as primary key in cases_dim table)
CONTACT	Project Contact Name (Used as primary key in contacts_dim table)

Table 1: Feature Explanation

## Findings

Using the SQL database that I created, I used Python to create a series of plots using the seaborn and matplotlib Python packages. I choose to use Python because of the speed that I could process the data, the ability customize queries using the Python language, and because of my familiarity with the programming language.

In Figure 3 we can see the number of neighborhoods with the most approved number of permits. Many of the permits are being issued in the less affluent neighborhoods of SF such as Bayview and the Western Addition. We can also see very few permits being issued to affluent areas and neighborhoods that have strict zoning laws such as the Presidio and Chinatown respectively.

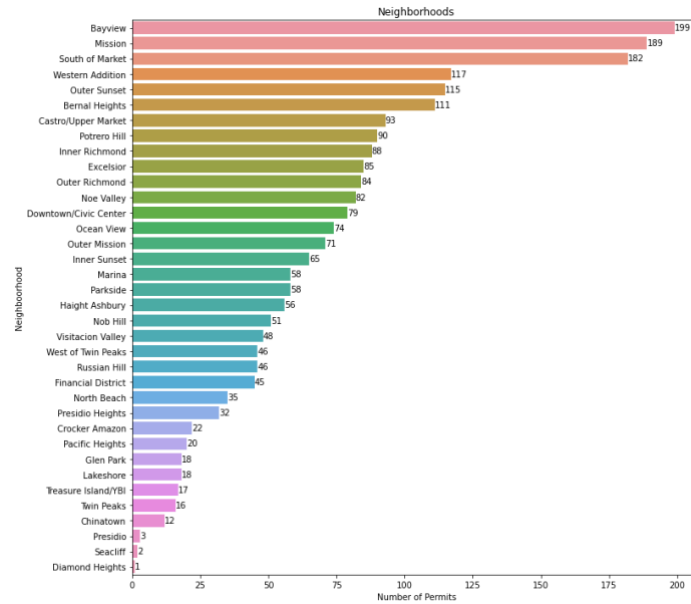


Figure 3: Neighborhoods with Most Amount of Permits

In Figure 4 we can see the number of affording housing units approved by neighborhood. Not that unlike the previous figure, this graph reflects number of units approved, not number of permits. A disproportionate number of units are approved in the Bayview, which is not necessarily reflected in Figure 3.

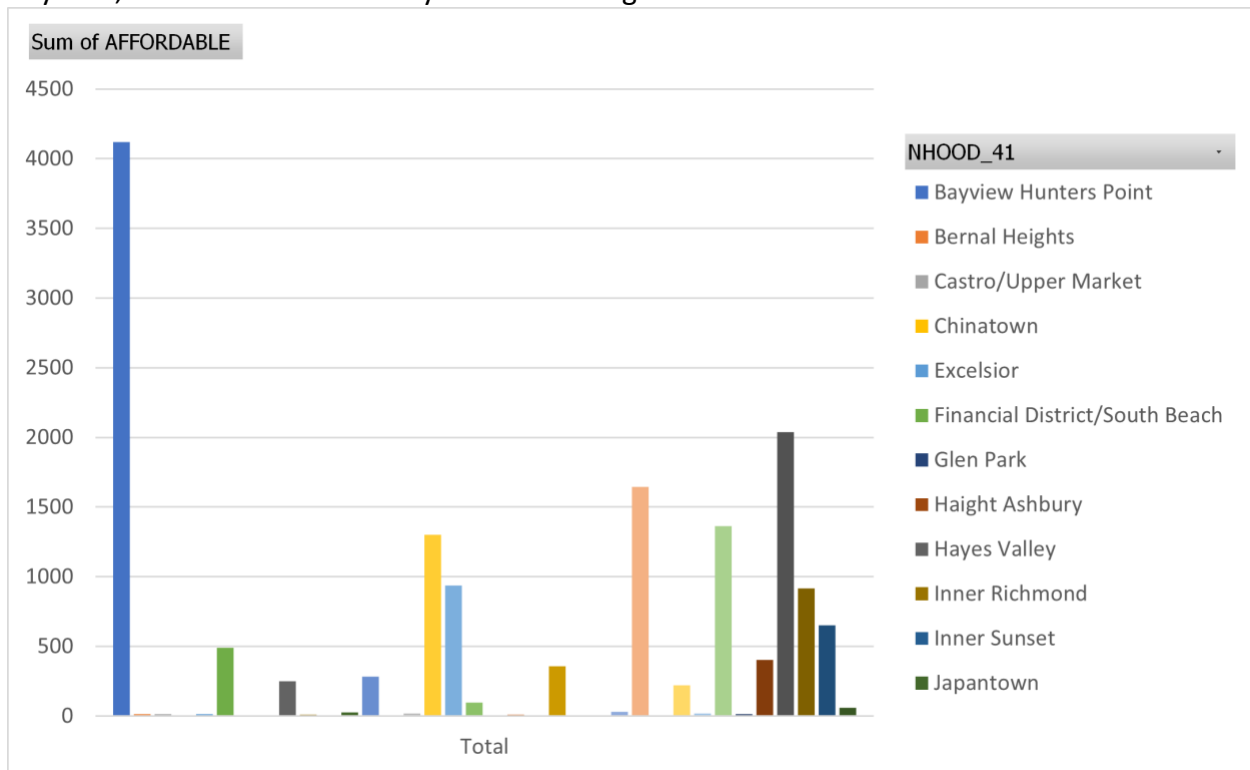


Figure 4: Affordable Housing Units

Figure 5 is a heatmap generated by the Folium Python library. The heatmap is generated by plotting the longitude and latitude from the fact table on a map. Locations are low in density and of new permits are shaded and blue. Locations that are high density are shaded in red.

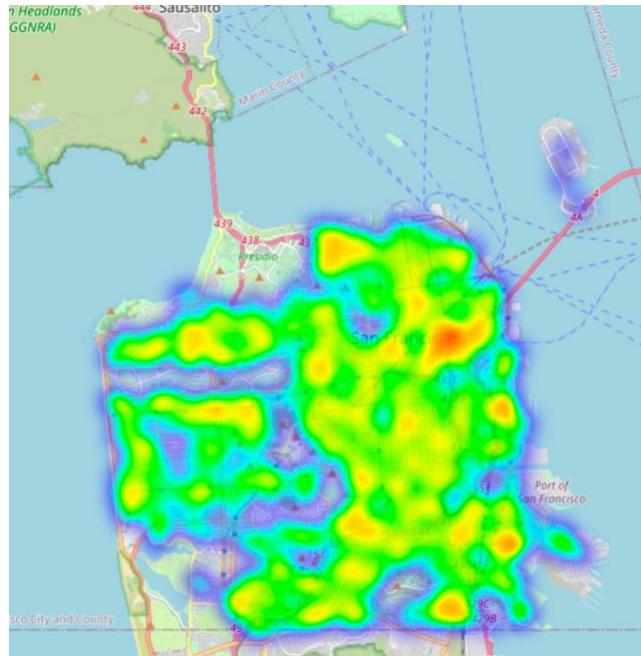


Figure 5: Permit Heatmap

Figure 6 shows a map of income disparity in San Francisco from Nick Conway, an independent blogger. Notice the strong overlap of red in Figure 5 and Figure 6. Neighborhood such as the Tenderloin (Known as South of Market in city data) and Bayview have large dark shades of red. Looking at these two graphs, we can observe a strong correlation between the number of housing permits and lower-income families.

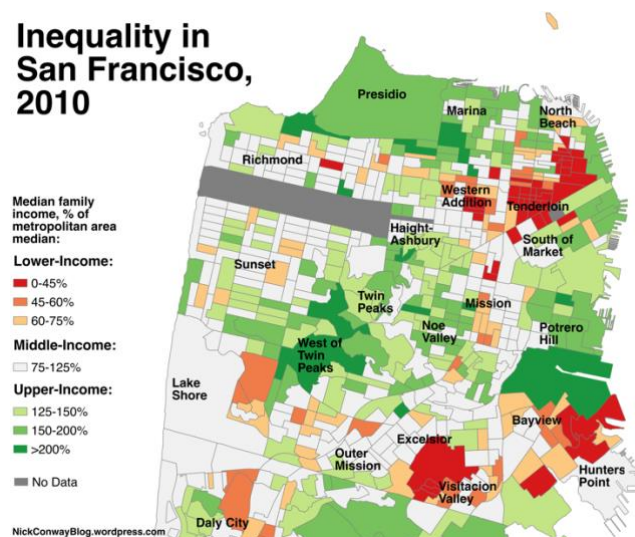


Figure 6: Inequality in San Francisco<sup>1</sup>

In Figure 7 we can see the number permits where units were added. The axis was limited 0 to 10 to make the graph viewable. The mean number of units added is 31.11 units. The median number of units is 1 unit. The max number of units added is 9119 units. The minimum number of units is -3 units. It is important to note not permits results in the creation of units. In fact, some permits request the demolition of some units, resulting in a negative number of units being created.

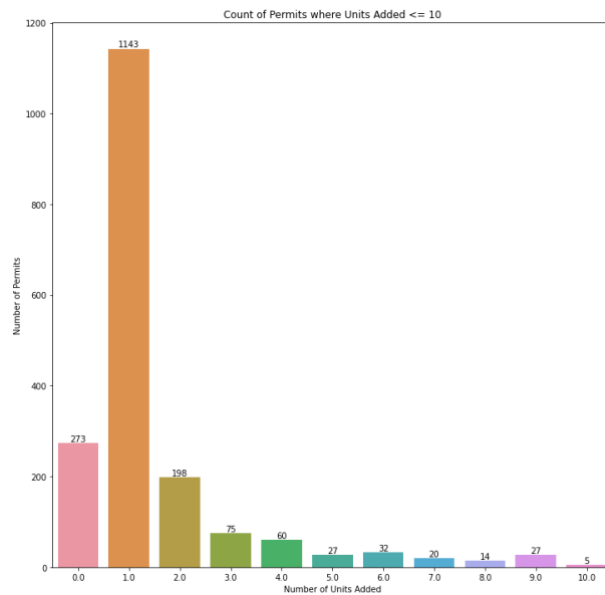


Figure 7: Counts of Where Permits Were Added ( $n \leq 10$ )

In Figure 8 we can view the individuals who are frequently request permits. Jim Abrams is a Lawyer who specializes in Real Estate development in the Bay Area.<sup>2</sup> Jeremy Schaub is founder of a small architecture firm who specializes in buildings in San Francisco.<sup>3</sup> John Kevlin is a Lawyer who specializes exclusively in permitting.

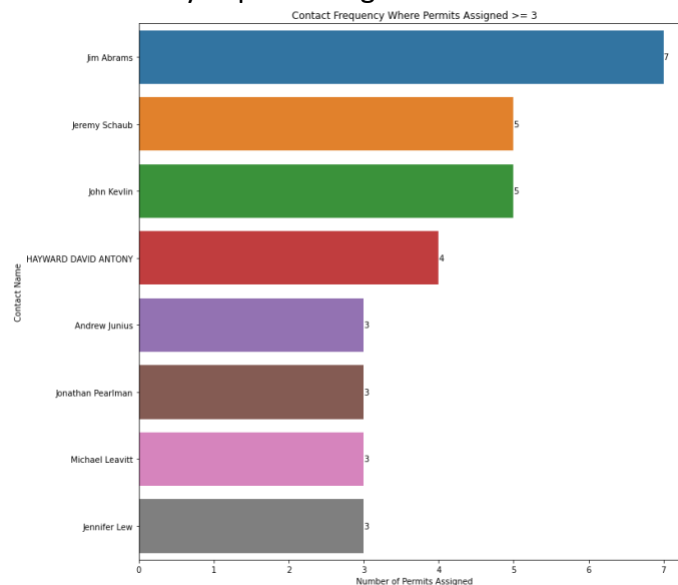


Figure 8: Number of Permits Requested ( $n \geq 3$ )

In Figure 9 we see the District Supervisors who have the greatest number of permits assigned to them. Looking at the data, we can see that some supervisors have permits assigned to them out of their elected district. In future research, it would be interesting to see the cause of this.

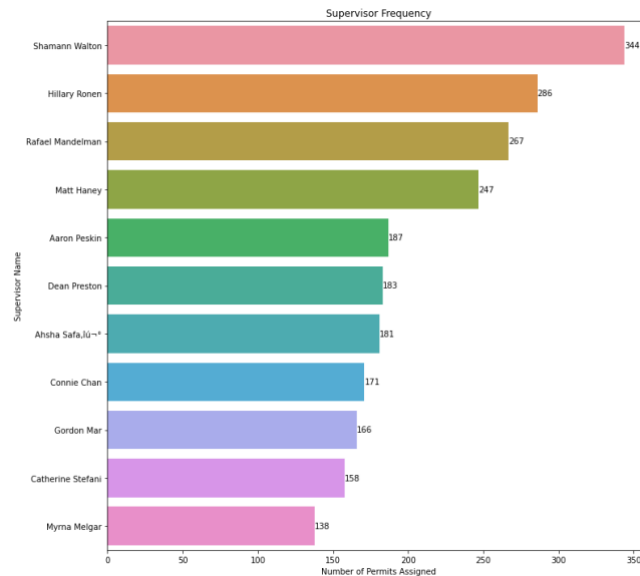


Figure 9: Number of Permits Assigned to District Supervisor

## Conclusion

From our findings, we can see a clear correlation between issued permits and lower-income neighborhoods. Intuitively, this makes sense. Housing pricing the San Francisco are too high because there is not enough housing in San Francisco, so more housing is developed. Lower cost units are created in lower income neighborhoods hence the disproportionate amount of units approved in those neighborhoods.

Is it fair that housing units are being developed at such a high concentration in less wealthy neighborhoods? Can affordable housing units be opened in a more diverse range of locations? What should be the goal of the city? Is it to decrease overall housing prices or to increase the population of the city through cheaper housing? These are some of the questions that should be asked and further researched from the data.

What is clear is that the current plan is to increase the amount of lower cost housing units to increase the supply of housing in the city. This strains low-income neighborhoods and exacerbates the wealth gap in the city. Units in high-income neighborhoods will continue to diminish in supply at a faster rate than low-income neighborhoods. This approach may be the best practical way to increase the number of affordable units, but there are externalities that must be considered like the ones above.

Lastly, our study identifies the most influential people in San Francisco housing development pipeline. Figure 8 and Figure 9 identifies the most influential individuals who submit housing permits and oversee them. Those who believe the current housing development pipeline is unjust should contact the individuals in the figures.

## Appendix

Figure 1: Housing Prices in San Francisco Relative to the Rest of the USA .....	2
Figure 2: Relationship Diagram.....	3
Figure 3: Neighborhoods with Most Amount of Permits .....	4
Figure 4: Permit Heatmap.....	5
Figure 5: Inequality in San Francisco .....	5
Figure 6: Counts of Where Permits Were Added ( $n \leq 10$ ) .....	6
Figure 7: Number of Permits Requested ( $n \geq 3$ ) .....	6
Table 1: Feature Explanation .....	3

---

<sup>1</sup> <https://sf.curbed.com/2016/12/9/13899936/income-inequality-map-san-francisco>

<sup>2</sup> <https://www.jabramslaw.com/about-forte/>

<sup>3</sup> <http://www.slasf.com/home>