

How Silicon Valley Companies Reshaped its Surroundings

Jeehee Choi
012565696
jeehee.choi@sjsu.edu

Tim Johnsen
014278914
tim.johnsen@sjsu.edu

Hyein Nam
014567475
hyein.nam@sjsu.edu

1. Introduction

FAANG (Facebook, Amazon, Apple, Netflix, Google), a movie ‘The Social Network’, TV series ‘Silicon Valley’ are the representatives of Silicon Valley. We, San Jose State University students, live in this area surrounded by the global companies, yet do not know well about the history of companies in Silicon Valley and its surroundings.

We are curious why and which industries found Silicon Valley as a favorable place to have headquarters; what are the distinctive points of people who reside in Silicon Valley; what are the results of increased population in Silicon Valley and how it has changed over time. This is why we have chosen this topic and want to learn about companies, residents, and houses in Silicon Valley.

2. Related Work

The San Francisco bay area has the largest concentration of high-tech companies, the highest average high-tech salary in the [United States](#) at \$144,800, and the San Jose metropolitan area has the third-highest GDP per capita in the world [1]. With the large number of high-salary jobs in Silicon Valley, the large influx of high-tech-related people have moved to Silicon Valley. With higher demand for housing for increasing numbers of people, housing shortage and high rent price has become a big issue.

Figure 9 shows the severity of the housing issues in the Bay Area. The median housing price in San Francisco is \$1.5 million, while it is \$0.58 million in Seattle, \$0.63 million in Los Angeles [2].

3. Narrative

Fortune 500

Nowadays, Silicon Valley is an important region in the US economy, because many IT companies, such as

Apple, Google and Intel, started and grew in this area. They became mega-corporations in the US. We investigated when the Silicon Valley economy started to grow, how many big corporations are located in Silicon Valley, and which industry leads the economic growth of Silicon Valley. We selected the companies in Silicon Valley from Fortune 500, because it shows the top 500 companies in the US as measured by revenue. They are significant influencers in the US economy. In 2019, there are 38 Fortune companies located in Silicon Valley. We categorized them into 6 industrial sectors: Technology, Financials & Services, Retailing & Wholesalers, Technology Related, Energy, and Others. In order to explain their history, we created a geographic visualization to present the locations, Fortune rank, and industrial category from 1960 to 2019 with a 10-years interval. From the map, many companies located in the Peninsula area entered the Fortune 500 in 2010. Then they are continuously growing and become mega-corporations. The majority of them are technology companies (Figure 1).

The map presents the qualitative trends of big companies in Silicon Valley. However, it does not show how much the growth of them affects the Silicon Valley economy. We draw the graphs to show the changes in revenues and profits of the companies over time (Figure 2). From the graphs, the number of big companies and their revenues and profits radically increased from 2010. As shown in the map, technology companies lead the economic growth of Silicon Valley. Most big tech companies are located in the Peninsula and south bay. In 2019, 50% of big corporations are in Santa Clara county. They participate in 66% of revenues and 75% of profits for 38 fortune companies in Silicon Valley. Thus, Santa Clara County is the economic center of Silicon Valley.

Housing

To no surprise, the number of housing units has linearly increased every decade since 1960. There doesn't seem to

be much drastic, exponential, change since the advent of large companies. Figure 3 plots time versus housing units.

As the number of houses increase, however, the percent of owners who live in those houses has decreased, Figure 4. This can reflect an increasing trend to capitalize the influx of employees in need of housing. This is especially true for temporary positions such as internships, contractors, and consultants that get drawn into the bay.

With the number of people flooding in, both the value and rent of housing has skyrocketed. This is visible in figures 3 and 4. Interestingly, the rent seems to scale up for all five counties studied but the value of the two east bay counties (Alameda and Contra Costa) seems to level out as reflected in the recent 2019 surveys. This shows that these two counties are the most profitable for owners who rent but not owners who live in their housing units.

The number of people who stay in one housing unit decreases until 2010, reflecting the number of high-paying employees being sent into the area. However, recently, from 2010 to 2019, the number of people per housing unit has started to increase. This could show that the density of people in the area is reaching a capacity, or that there is an influx of lower income employees. See figure 5.

The maps presented show how these housing trends all vary by county, showing shifts in housing situations. This is done using colors, sizes, and shapes centered around the location of the counties.

Demography

Using Census data, I collected a date range from 1960 to 2019 with a 10-year interval (1960, 1970, 1980, 1990, 2000, 2010, 2019). The locations selected are the five counties in Silicon Valley: San Francisco, San Mateo, Santa Clara, Alameda, Contra Costa.

Figure 10 and 11 show the population growth in Silicon Valley. The population of Silicon Valley has grown from around 3 million in 1960 to 6.4 million in 2019. The number of people in Santa Clara County and Alameda County has increased more than two times.

Figure 12 shows changes in race ratio in Silicon Valley by year. The Asian, Pacific Islander, and Hispanic populations were less than 5% in 1960, while its number grew up to 62% in 2019. The White population shrank from 97% in 1960 to 31% in 2019. The Black, Native Indian, and other ethnic groups' population grew up from 1960 to 1990, then has declined afterward.

Figure 13 shows changes in educational attainment in Silicon Valley by year. The high school graduation rate had been decreasing from around 30% in 1960 to 15% in 2019. On the other hand, the number of people with a bachelor's or a master's degree has increased; the highly educated population consists of above 50% of Silicon Valley.

Figure 14 shows changes in occupations in Silicon Valley by year. Diverse companies established headquarters in Silicon Valley, and the industry has become distinct. With this influence, the occupations also have changed. While other-occupations (Operatives, Craftsmen, foremen, Occupation not reported) are the major careers in 1960, the management and service-related occupations have become the core in 2019. The popularity of sales and construction occupations has been decreasing since 2000. The production occupation rate has been steady since 2000.

The Census data were not available for educational attainment and occupations data in 1970, 1980, and 1990.

Figure 15 shows changes in median family income in Silicon Valley. The median family income in Silicon Valley has grown exponentially. The median family income was less than \$100k in 1960; the median family income increased more than \$900k in Alameda and Contra Costa County in 2019; the median family income increased more than \$1100k in San Mateo and Santa Clara County in 2019.

4. Implementation of Your Story

We collect 3 datasets to present to explain the growth of Silicon Valley: Fortune 500 companies, demographics, and housing. All datasets are obtained from public websites [Table 1]. The data acquired from the website is the most challenging process in the project. Fortune 500 is not open-sourced data, but Fortune website shows the detailed list for the last 5 years. We manually collected Fortune 500 lists for between 2015 and 2019 by web scraping. Old Fortune 500 lists were obtained from GitHub. We search the demographics data in the US and California state census, and TownCharts website then manually collected the data from each website because each website has different attributes and different geological levels of data. The US census provides population, race, income, poverty data by county, and Educational data for each county is obtained from TownChart and US census educational attainment.

Further, the census data has a 10-year interval. We decided to present all data with a 10-year interval.

After collecting the raw data, we performed data wrangling, then inserted the georgical data to create an interactive map using ArcGIS online. The map is generated by adding the layer containing the data. To increase the expressiveness and effectiveness, we apply 2 visual variables: size and color into the maps. Fortune 500 company map uses the longitude and latitude to present the location. The symbol size and the colors represent the Fortune 500 rank of the companies and their industry sector. The demographic map uses the color to show the county area and symbol size to represent their pollution. To present the county area, we created a hierarchical layer of county boundaries for each county then linked to the layer having the democratic data. Unlike other maps, the Housing map applies the color scale to present the persons per household, and symbol size shows the year. In this map, we changed the background map to improve the visibility of the symbols. Finally, a story map is generated to present the visualizations with our story using an ArcGIS web application, Classic Story Maps. This application also provides templates and a builder user interface. Most templates consist of a main stage and side panel. In the main stage, we can add any visualizations, such as the map in ArcGIS online, images, video on YouTube, and a Website link. The ArcGIS map on the main stage provides the interactive mode like zooming and selecting the location, but not for selecting the layer on the maps. The side panel is used for adding the description like reports or articles.

5. Results

Fortune 500

The map of Fortune 500 companies in Silicon Valley from 1960 to 2019 presents that the growth of technology companies in Santa Clara county makes Silicon Valley an essential economic region of the United States. As explained in the previous section, we apply color to inform the industrial sector of each company. According to Fortune 500 rank, there are 11 industrial sectors in the data. However, 11 distinct colors on geographic visualization do not effectively convey the information. We re-categorized the companies into 6 industry sectors, followed by Miller's magic number. In this map, the distinct color is selected because it represents the qualitative values. The size of symbols is used to represent the rank of the companies. Unfortunately, many companies are located in Santa Clara County and the city

3 Choi, Johnsen, Nam

of San Francisco. We faced the problem of overlapped symbols. We tried to change the shape and size of the symbol, but the problem is not solved. Zooming is the best option to show the cloud area.

Another problem is that ArcGIS online version does not provide a time slider function to visualize temporal data or to create an animated map. We generated the map composed of multiple layers, which are presenting the data for each year. Users can select the layers to see the map for a year of interest. Unfortunately, the layer selection is not working when the map is embedded into a story map. We used QGIS to create an animated map and inserted it into the story map as a video format. QGIS is open-sourced software and has many plugins to add more functions and attributes like time slider and 3D visualization. It is an excellent alternative to ArcGIS. We use a time manager plugin to create the animated map for Fortune 500 companies in Silicon Valley. When we created the map using QGIS, we followed the Gestalt principles of similarity. The map generated using QGIS is applied to 2 visual variables, symbol size, and color to present Fortune 500 ranks and industrial sectors. The color of symbols on the maps generated by the different tools are consistent.

Provide your visual story major scenes and explanation of rules and concepts that you used and innovation in visualizations. Results of your system and how evaluated your story(based on the concept that we are going to discuss in upcoming sessions.

Housing

The ArcGIS figure 8 maps housing statistics against time and across the 5 bay counties of interest.

The ArcGIS maps show the metrics: number of housing units, percent of owners who live in their housing units, median rent, median value, and persons per household. The given metrics are measured for each of 5 counties: San Francisco, Contra Costa, Alameda, and Santa Clara. The metrics are also measured from the surveys, for data which was available, during 1960, 2000, 2010, and 2019. The metric is indicated by the shape of a marker placed near the county on the map. The year of the survey is mapped to the size of the marker. The value of the metric is heat mapped so that larger values are darker red.

The ArcGISmaps were evaluated by asking three observers several questions, via zoom, and the answers are displayed below with the question. The time was also recorded for the user to be able to find the answer.

1. Which county had the highest number of persons per household in 1960?
Contra Costa (5s), Contra Costa (2s), Santa Clara or Contra Costa (5s).

These responses are the correct answer, and within a small period of time. This shows that the number of persons arcGIS map in figure 8 is robust.

2. How many housing units were in Alameda after 2010?
More than 600k (4s), between 650-700k (10s), more than 678k (25s).

The answers are nearly correct, though the participants expressed and appeared confused since the color was near the top of the heat map. They were not sure if it was higher than this value or at it.

3. Did the value of housing increase or decrease over time in San Mateo?
Increase (10s), increase (4s), increase (10s)

This was easy info to extract from the map.

4. What year had the highest owner occupied percent in Santa Clara?
1960 (4s), 1960 (2s), 1960 (7s)

This too seemed to be easy info to extract from the map, even with this map being the most complex of all in figure 8.

The feedback was rather positive. The only negative seems to be looking at regions near the edge of the heat map. This can be fixed by a larger gradient in color.

Demography

Considering Miller's magic number 7 (+- 2), 5 counties (San Francisco, San Mateo, Santa Clara, Alameda, Contra Costa), 7 years with a 10 interval (1960, 1970, 1980, 1990, 2000, 2010, 2019), 5 types of attributes (population growth, race, educational attainment, occupations, and median family income) data were collected and preprocessed. Each county's color was color-coded with that of the same color used in the Future 500's county. Santa Clara with red, San Mateo with gray, San Francisco with yellow, Contra Costa with green, and Alameda with blue.

Instead of showing all county's attributes, I chose the average of the 5 counties for each attribute to increase the

effectiveness and efficiency of overall understanding for readers. I used bar charts to compare values with each county and stacked bar charts to show the overall quantity (%) of each variable.

I wanted to show animated charts on the ArcGIS map, but the tools were limited to do so. Instead, I created a time-series of population growth using ArcGIS Pro, did a video-screenshot of the animation, and created a Youtube video.

Clicking each county's circle in the ArcGIS online map shows detailed information of all attributes by year. Figure 16 illustrates a pop-up with details. The detail shows selected data's attributes and an arrow and switches to different years.

6. Discussion

We have shown where and what rate large company headquarters have moved into the bay area. There was a large increase from 2000 to 2019. People who view the maps will understand the story of how these companies affected housing situations and demographics.

The number of housing units linearly increased, with an initial drop in people per house followed by a recent increase. The drop is assumed to be from highly paid employees moving in with the headquarters. The recent increase is assumed to be from many temporary positions and lower paid employees flocking to the now booming bay area.

Housing unit values have skyrocketed from 2000, including median rent. Interestingly, the value of housing units in Contra Costa and Alameda seem to reach an asymptote while the other three counties continue to rise. However, the median rent for all 5 counties continues to rise at the same rate. This shows that owners who rent out units in Contra Costa and Alameda are set to get profits versus the value of their homes, however the owners who live in their units are not gaining much capital.

We expect the reader to learn that demographics have changed with the advent of large companies. The percent of people with higher education has drastically increased, even though surprisingly the number of people who have graduated highschool is decreasing. The diversity has increased due to more jobs being pulled in from around the world. Finally, family incomes have risen exponentially across the bay area.

7. Future Work

In this project, we created geographic valuations and story maps to present the history of Silicon Valley. Initially, we planned to create an animated map or interactive map with a time slider to convey the information and present the story. Unfortunately, we faced the technical limitation of ArcGIS online. To solve this problem, we use an alternative product, QGIS, to create an animated map, but the QGIS map is not 100% compatible with ArcGIS. We can add the animated map as a video format. Thus, the animated map is not interactive. For archiving the initial project goal, we plan to use different versions like ArcGIS pro to provide the time slider and other interactive options to users for improving the expressiveness and effectiveness.

We investigated the history of Silicon Valley using 3 different data. Here we focused on presenting the qualitative relationship between the economic growth in Silicon Valley and changes in demographics and housing for telling the story. To describe the economic trend in Silicon Valley, we selected 38 big companies from Fortune 500. Most mega-corporations in Silicon Valley, such as Apple, Alphabet (Google), and HP started as startups, grew, and became giant companies. The history and statistics of big companies in Silicon Valley are not enough to explain why many technology-based companies are established and how many potential mega-corporations are located in Silicon Valley. We will include startup data in the story map to provide the comprehensive information of the Silicon Valley economy. Furthermore, statistical methods and machine learning algorithms will be applied to discover the correlation between them and predict future trends. Recently, the COVID-19 pandemic affects not only individuals but also mega-corporations. Many companies lay off or plan to lay off. It might impact the economic growth in the entire county. We have a plan to collect the data related to present how COVID-19 pandemic affects Silicon Valley.

8. What did you learn

We learned the story of the bay area when related to large business headquarters, with resulting demographics and housing situations. Other than this we learned how to use ArcGIS, and gained better insights to working in a group. We especially gained experience with working in a group online.

We found out that ArcGIS is powerful, however the Pro version is really needed to do anything detailed. The version we had access to is very limited, and had a couple bugs.

We also learned more about markers, and how to display multiple levels of information within one figure.

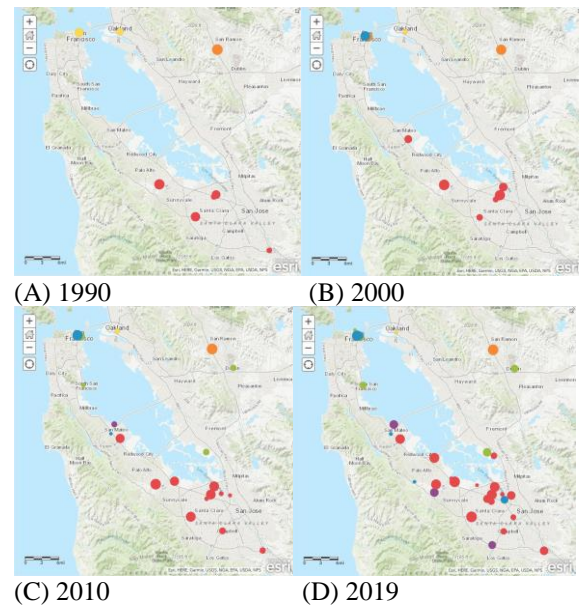


Figure 1. Fortune companies in Silicon Valley.

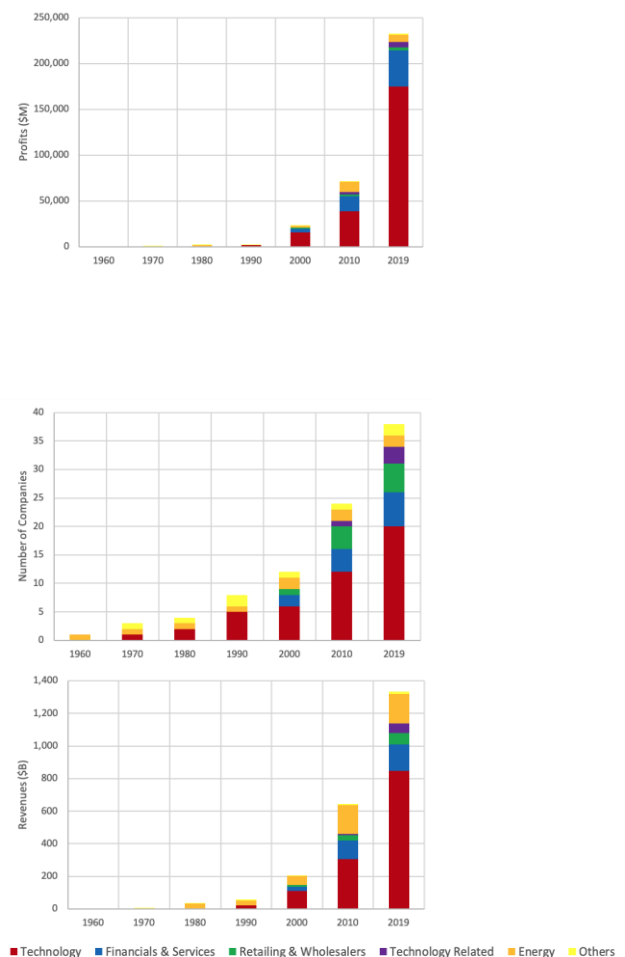


Figure 2- (A) Industrial Sector

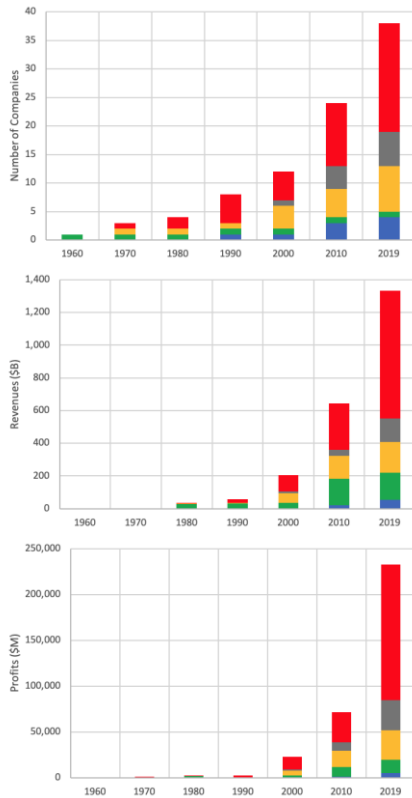


Figure 2- (B) County in Silicon Valley
 Figure 2. Statistics of Fortune companies in Silicon Valley

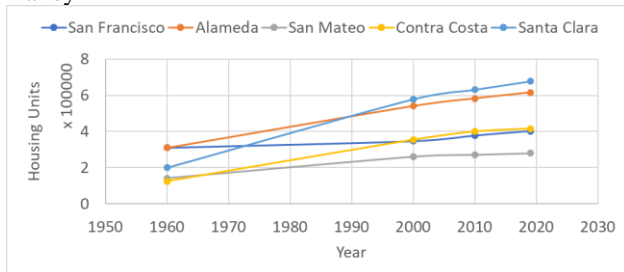


Figure 3: Housing units versus time.

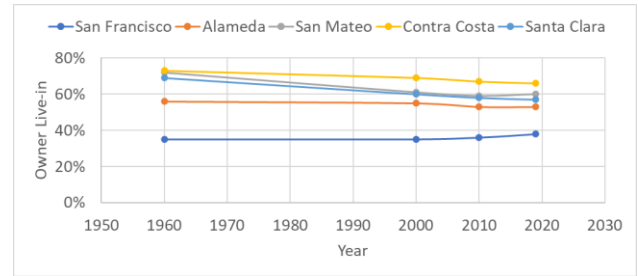


Figure 4: Owner live-in percents versus time.

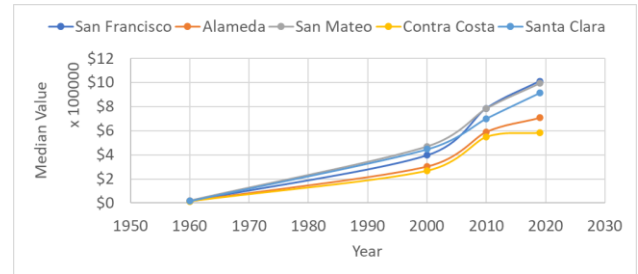


Figure 5: Housing value vs time.

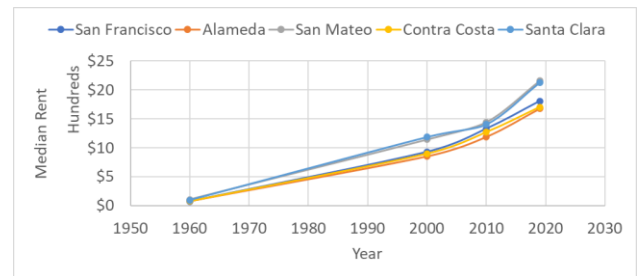


Figure 6: Housing rent vs time.

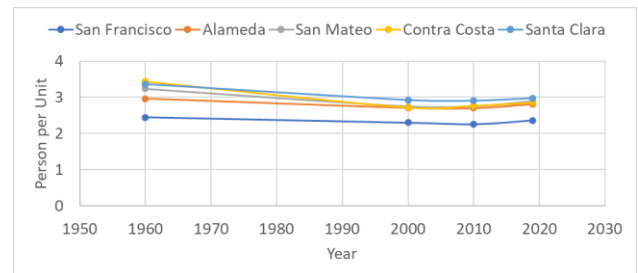


Figure 7: Person per housing unit vs time.

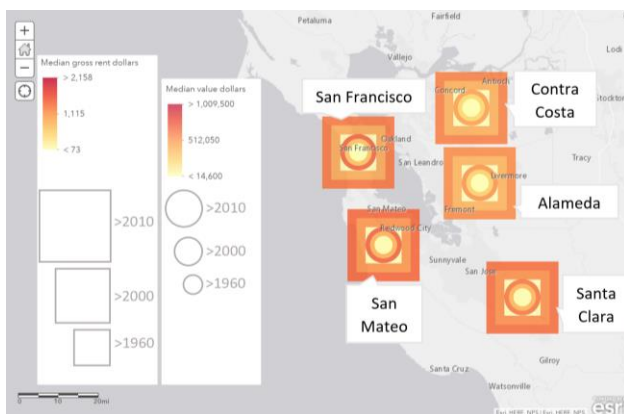
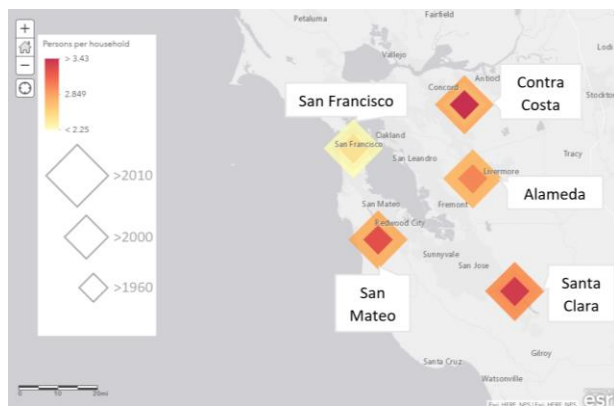
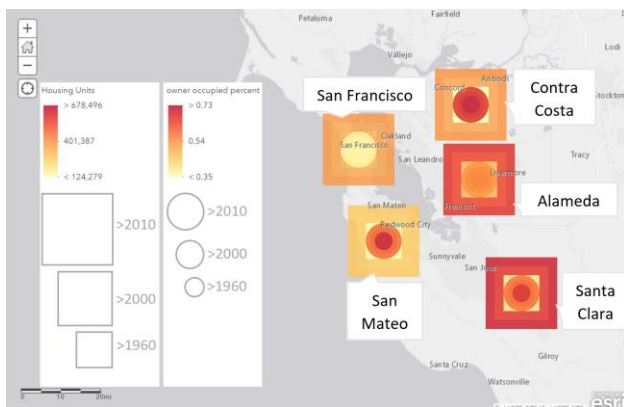


Figure 8: ArcGIS with Housing Statistics

Why are San Franciscans leaving? Highest housing prices in the West

MEDIAN HOME PRICES, JUNE 2019



SOURCE Redfin
George Petras/USA TODAY

Figure 9: Median Housing Price in the West Coast

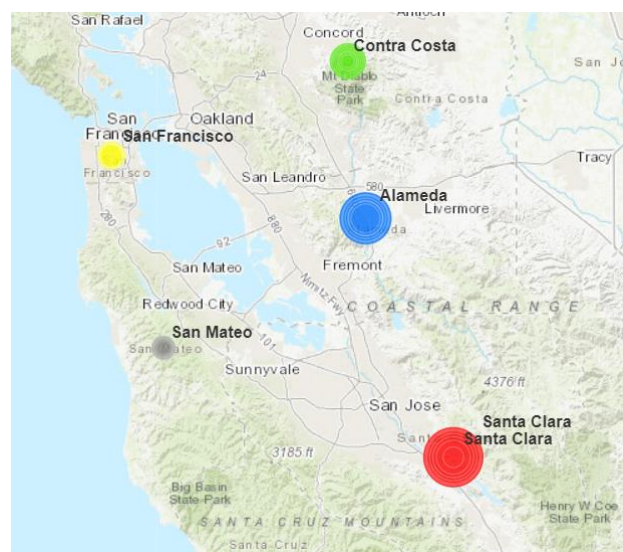


Figure 10: Population Growth in each County by Year

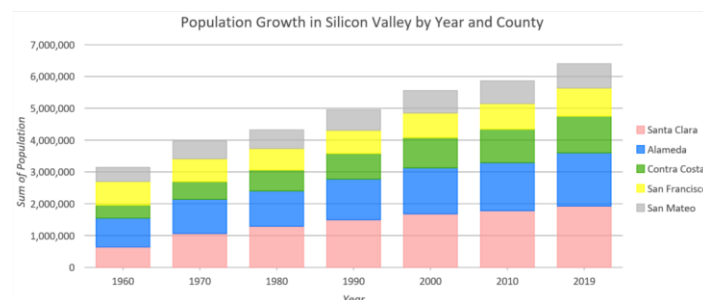


Figure 11: Population growth in Silicon Valley by Year and County

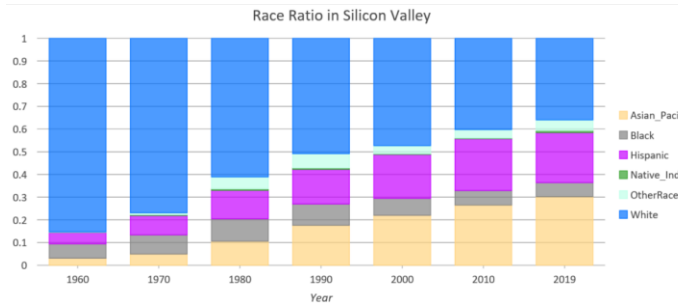


Figure 12: Race Ratio in Silicon Valley by Year

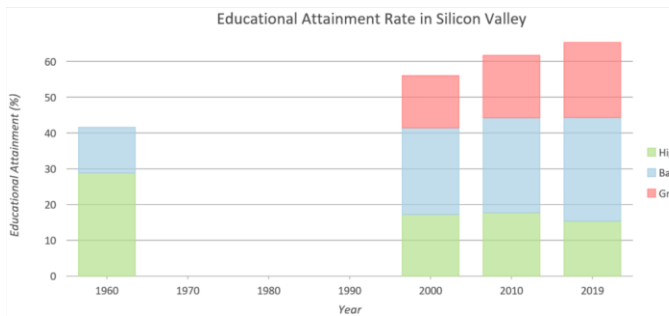


Figure 13: Educational Attainment Rate in Silicon Valley by Year

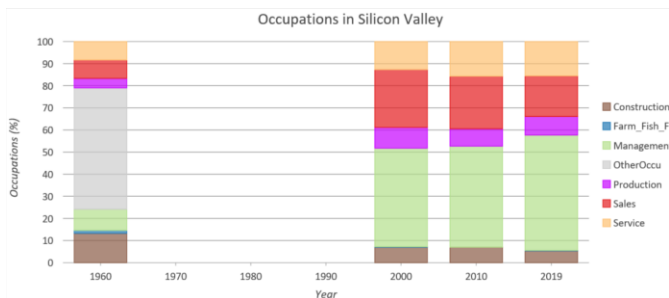


Figure 14: Occupations in Silicon Valley by Year

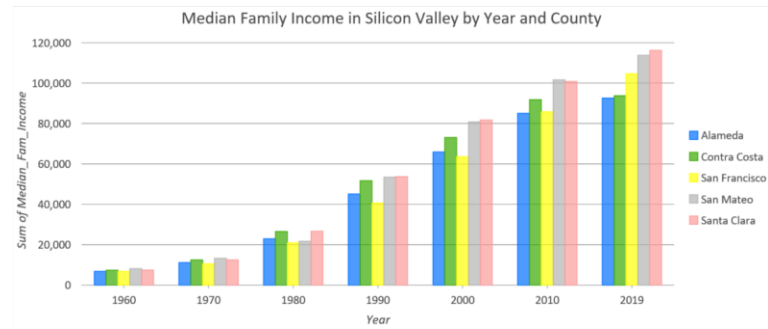


Figure 15: Median Family Income in Silicon Valley by Year and County

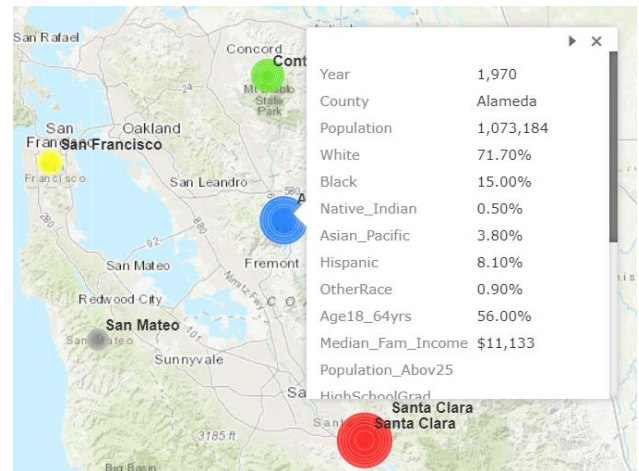


Figure 16: An Pop-up Shows Details about Selected Data with Each Year by Clicking an Arrow

Table1. Dataset and Data Source

Dataset	Data Source	URL
Fortune 500 from 2015 to 2019	Fortune 500	https://fortune.com/fortune500/ https://github.com/cmusam/fortune500
Fortune 500 from 1955 to 2014	GitHub	https://github.com/cmusam/fortune500).
Headquarter location of Fortune 500 companies in 2018	Homeland Security website	https://hifld-geoplatform.opendata.arcgis.com/datasets/fortune-500-corporate-headquarters
Demographics data for Bay Area	Bay Area Census data website	http://www.bayareacensus.ca.gov/bayarea.htm
Demographics and Housing data for the county in Silicon Valley	US Census Bureau website	https://www.census.gov/quickfacts/fact/dashboard/sanfranciscocountycalifornia/PST045219

Education Data	Towncharts	https://www.towncharts.com/California/Education/San-Francisco-County-CA-Education-data.html
Education attainment data in 2019	US Census Bureau website	https://data.census.gov/cedsci/table?t=Educational%20Attainment&tid=ACSST5Y2018.S1501&hidePreview=true&cid=S1501_C01_001E&vintage=2018&moe=false&layer=VT_2018_050_00_PY_D1&g=0500000US06013

[1] En.wikipedia.org. 2020. *Silicon Valley*. [online]
Available at:
<https://en.wikipedia.org/wiki/Silicon_Valley>
[Accessed 8 May 2020].

[2] Usatoday.com. 2020. [online] Available at:
<<https://www.usatoday.com/story/news/nation/2019/10/19/california-housing-crisis-residents-flee-san-francisco-because-costs/3985196002/>>
[Accessed 8 May 2020].