**The Rise of Silicon Valley**  
An Interactive Tour

*W209-4 Final Project, Fall 2017  
Phat Doan, Ted Pham, and Yisang Yoon*

INTRODUCTION

Silicon Valley is an area in northern California, southwest of San Francisco in the Santa Clara valley region, called from the silicon wafers employed in the many high-technology design and manufacturing companies in the semiconductor industry.

Since its humble beginnings decades ago, Silicon Valley has become a household name and profoundly impacted the world economy. Technologies emerged from the area changed the way many people live. While the term was coined to describe a cluster of semiconductor related companies, it is now often used synonymously to describe tech-based startup entrepreneurship in general.

The GDP generated by the region dwarfs that of many countries, and it remains a top choice of location for entrepreneurs to seek funding. Through our project, we want to take you through a visual history of this impressive Rise of the Silicon Valley as a significant hub of startups and a major economic powerhouse.

USER-TASK-DATA

The three of us in the took a lead in each of the 3 steps to bring together this project. Yisang was in charge of exploring, gathering and cleaning the necessary **data**. Phat led the efforts in converting the data into visualizations – deciding which visualization conveyed the data best, and implementing the infographics. This was the “**task**”. And finally, Ted brought together the visualizations into a website format to tell the final story, connecting the data and task to the end **users**.

*Data*

Silicon Valley of course is not defined by its economy alone. But as our data suggests, our project is focused on delivering information to anyone interested in the startup companies in Silicon Valley, particularly their finances. So by and large, the data is of various financial information of the companies headquartered in the Silicon Valley area.

The sources of these data are attributed to several sources, including Bloomberg, Crunchbase, Bureau of Economic Analysis, Federal Reserve Bank, and Kaggle. The data contains historical information of companies headquartered in the Silicon Valley area. Information includes valuation, revenue, acquisitions, investments, GDP, sector (industry), product, and location.

Some data required cleaning – such as realigning and removing seemingly erroneous and/or missing values. There were also datasets that we conducted exploratory data analysis, but decided leaving out as we found some of the information, while interesting, was not necessary for the sake of our intended users.

It should be noted that all information contained within our project is publicly available information, some of which are available from multiple sources. In addition, while we believe the validity of the data, we cannot guarantee their completeness and accuracy.

Below table provides descriptions of the data that we used, and their respective sources:

|  |  |
| --- | --- |
| **Data** | **Source** |
| Historical quarterly valuation information of publicly traded companies within the Silicon Valley, including stock ticker, price, volume, and total outstanding valuation, through 9/30/17 | Bloomberg |
| Latest profiles of publicly traded companies headquartered in Silicon Valley: ticker, name, city of domicile, market cap, trailing 12 month revenue, IPO date, bankruptcy date (if available), number of employees, GICS sector, CEO compensation in USD, CEO tenure and Board tenure (both in number of years): 1,844 company profiles collected, 384 valuations available, through 9/30/17 | Bloomberg |
| Historical company information regarding acquisitions and investments | Crunchbase |
| Historical global GDP information for each countries | World Bank |
| Historical GDP information for Silicon Valley area, including sector breakdown | Bureau of Economic Analysis |
| Historical GDP information for State of California | Federal Reserve Bank |
| Silicon Valley startup information, including name, website, main product, location, and target market | Kaggle |

Our team define the borders of Silicon Valley before we set out to collect data. The name Silicon Valley describes the South Bay of the San Francisco Area. This term derived as a result of the many silicon chip innovators and manufacturers that were based in the Santa Clara Valley. However, since its emergence in the 70’s, the breadth of the companies founded in the region has expanded, as well as the border that defined the region. Hence, we have included counties and cities beyond the traditional Silicon Valley Area for the purpose of our project.

While Santa Clara County, San Francisco, and the Berkeley-Emeryville area still constitute the strongest centers, the following counties have been included: Santa Clara County, San Mateo County, Alameda County, Santa Cruz County, San Francisco County, and Contra Costa County.

The following are the counties and cities we included (6 counties, 34 cities):

**Santa Clara County**

* Campbell
* Cupertino
* Los Altos
* Los Altos Hills
* Los Gatos
* Milpitas
* Monte Sereno
* Morgan Hill
* Mountain View
* Palo Alto
* San Jose
* Santa Clara
* Saratoga
* Sunnyvale

**San Mateo County**

* Belmont
* Burlingame
* East Palo Alto
* Foster City
* Menlo Park
* Millbrae
* Redwood City
* San Bruno
* San Carlos
* San Mateo
* South San Francisco

**Alameda County**

* Emeryville
* Fremont
* Newark
* Oakland
* Union City

**Santa Cruz County**

* Santa Cruz
* Scotts Valley

**San Francisco County**

* San Francisco

**Contra Costa County**

* San Ramon

*User*

We focus our data collection on financial information of the many companies that inhabit the Silicon Valley area. Naturally, our project aim to deliver insightful information to anyone interested in the startup companies in the area. Since Silicon Valley area can be of interest to a wide range of audience, and it would be difficult to satisfy everyone that had the slightest “interest”. Our team narrow our users more specifically, mostly to guide the team build a project that delivers information in the most useful and efficient way.

We envisioned several scenarios that would fit a user’s interest to our story. Some of these scenarios involved very specific prototype users:

(1) an entrepreneur who is seeking opportunities to establish a company in the Silicon Valley area;

(2) a foreign investment professional who is surveying trends in the Silicon Valley area to be informed of the general environment; and,

(3) an angel that is researching ways to garner additional interest of foreign investors.

Our team keep these model users in mind when selecting participants for our user tests. While we could not find exact matches, we are able to find proxy users based on background as well as interests. For example, one of our participants was a software engineer currently employed in the Silicon Valley, instead of an entrepreneur. Another participant was an investment analyst, instead of a foreign investor.

Using the user tests’ results, we enhance the project in various areas, something as small as catching a minor typo, to something larger such as adding a theme. Below are some changes we integrated based on the user tests:

* Reorganizing themes to better fit the storyline:
  + Growth: the humble rise of Silicon Valley
  + Funding: attracting investments and talents
  + Acquisition: growing up and spreading it influence
  + Adding a fourth theme (legacy): At the time we were conducting user tests, we did not have the page titled “Legacy”. However, one of the users pointed out that the site would be more interesting if there was a comparison of the size of Silicon Valley’s economy to other countries or the world. This would help put things in perspective as the relative size of Silicon Valley’s economy would help point out its rank in the world economy.
* Adding short descriptions to the infographics: At the time we were conducting user tests, we did not have much commentary for the infographics. However, a user complained that the missing narrative would have helped him get a sense of what was going on.

*Task*

With the above model users in mind, our team identifies specific information that our model users would seek after. For example, an entrepreneur may be most curious about his prospects of securing enough funding. What industries attracted the most funding? Is the trend such that these industries are receiving more funding in the recent years, or is the funding shrinking from its peak interest? The entrepreneur may also have interest in the acquisition environment, as she may be intending to sell her company later on. Which companies are most aggressive in acquiring other companies, and what industries are they expanding to? Where are most other companies located?

In another scenario, a Chinese analyst working for a ultra-high-net-worth family office may be interested to know about a few companies he is researching to recommend investments in, recent startups headquartered in the Silicon Valley area. What amount of funding did these companies receive, and when? How well are the industries these companies belong to represented in the area? Do they belong to one that has been attracting a lot of capital recently? Or one that is often the target of acquisition by recent trends?

Lastly, an angel fund agency is recruiting investors to build up its cash base, to be deployed when an attractive investment opportunity is found. Considering the latest trend, what information would be most compelling? The very impressive rise in the valuations of the companies? Perhaps how the aggregate valuation compares to other countries or even the world?

VISUALIZATIONS & INSIGHTS

As mentioned above, our story has 4 major themes that together paints the rise of Silicon Valley. Significant things we observe:

* The number of companies founded in the Silicon Valley has increased significantly since 2000’s. There are good reasons to launch your startup in Silicon Valley. Securing funding and mentorship, building a team of qualified individuals, and finding clients are all reasons why startups cluster in the Silicon Valley area despite soaring living costs.
* Silicon Valley as we defined it is quite a large area, covering from its traditional roots in San Jose all the way up to the Berkeley-Emeryville tech hub, and the area between them. Santa Clarita County, San Francisco, and Berkeley remain the top three centers, with San Francisco comprising the largest basket.

*Growth (the rise of Silicon Valley);*

Silicon Valley is the world’s leading center for venture capital investment attracting nearly $11 billion, more than a quarter of all global venture investment of $42 billion. There are good reasons to launch your startup in Silicon Valley. Securing funding and mentorship, building a team of qualified individuals, and finding clients are all reasons why startups cluster in the Silicon Valley area despite soaring living costs.

*Funding (raining money);*

The breadth of industries that glitter Silicon Valley has expanded so widely, “Silicon Valley” may now be an outdated misnomer. Silicon Valley is no longer just a hub for tech companies. Software industry has been a constant top performer when it comes to generating investments. But growing number of various industries have secured funding for their ventures over the years.

*Acquisitions (mature companies invest in other startups)*

An acquisition is the purchase of one company by another in which no new company is formed. Acquisitions are a part of the daily Silicon Valley environment. Acquisitions can be a way to grow business, add more products, or add more customers. They can also allow founders to cash out or become a part of a larger company. Sometimes, the combination of two companies forms a new company, and this is termed a “merger”.

Cisco leads this game with 194 acquisitions to date. Google is not far behind. The rise of Silicon Valley has affected the many parts of the country, coast to coast.

*Forging a Legacy*

The Silicon Valley economy surpasses GDP of many nations. Technology giants Apple and Google collectively employ almost 40,000 people in the South Bay Area and can be credited with much of the income generated. It is important to remember, however, that other tech firms including Intel, Cisco, Adobe, Hewlett Packard, Advanced Micro Devices, and Lockheed Martin continue to contribute to the South Bay’s economy. Not to mention the engineering firms like, Salas O’Brien Engineering and URS Corporation, who benefit from the area’s continued growth.

TOOLS AND TECHNIQUES

Our team uses Web technologies (HTML and Javascript) for the website, SQLite for data storage, Python for data cleaning and processing, and Tableau for all visualizations.

AUTHORS

*Phat Doan*

Phat is a Senior Statistician with the Population Health Analytics Group at Intermountain Healthcare, the largest, integrated healthcare system in Utah.

*Yisang Yoon*

Yisang is a Senior Investment Analyst with the Institutional Consulting Group at Bank of America Merrill Lynch.

*Ted Pham*

Ted is a Data Scientist at Hearts & Science specializing in marketing consulting with the high-profile client group.

ACKNOWLEDGEMENTS

We would like to thank the School of Information at UC Berkeley and Professor John Alexis Guerra Gómez in supervising our group through this project. The outstanding classmates in the W209 class of Fall 2017, Andrea Pope in particular, who provided exceptional feedback, as well as participants of the user tests need to be recognized as well.