

Financing Tech Startups

CS 171 FINAL PROJECT PROCESS BOOK

By Andrew Raftery, David Taitz & Jason Hirschhorn

Table of Contents

Project Proposal
Overview and Motivation
Related Work
Research Questions
Data
Design and Implementation Evolution
 Meeting Notes
 Visualizations

Project Proposal

Background and Motivation

With continuous evolution in the mobile and web technology marketplace, the prevalence (and rise and fall) of startup businesses in the United States is undeniable and inescapable. Yet with so many companies raising money to fund their new ventures and with so few of these businesses actually succeeding, we are curious to better understand 1) whether or not investor capital is being allocated wisely, and 2) the financial trajectories of successful startup businesses, sectors within the tech industries, and investment firms. As three people who are interested in pursuing careers in startups and venture capital, it seems prudent to better understand trends in the market.

Project Objectives

Our series of visualizations will answer the following questions for the n largest tech firms:

- How much funding have they received and in how many rounds? From whom? When?
- How active has a particular venture capital firm been?
- What sectors within the tech industry receive the most attention?
- How large are different rounds of funding (both in absolute and relative terms)?

With our visualization, we hope to learn significantly more about funding patterns and trends in the industry.

Data and Data Processing

We will collect data on all venture deals completed by 20 top firms in the last 5 years from [CrunchBase's API](#). The data includes the company invested in, the firms involved in the deal, the size of the round raised, the company category, the company's founding year, and more. For the data collection and processing, we'll use Python. We

don't anticipate to have to clean the data substantially (since the API is made for easy data access), though some will certainly be required (converting strings like '\$50M' into 50000000, for example).

Visualization and Features

The basic visualization will be a large rectangle, broken (and colored) into sections to represent each company. The size of the section will correspond to the total amount of money raised by that company to date. Each company subsection will be further broken down into sections that represent the different rounds (series A, B, C, etc.) that the company has raised, again, size indicating the dollar amount of the round. Though the underlying data is different, the general design of the visualization will be similar to a [treemap](#). Below the large rectangle view will be a graph showing all 20 firms and some data on them (how much money they've invested over the last 5 years, how many companies they've invested in, etc.). Hovering over the bar of a firm in the bar graph will highlight all of the rounds they've participated in in the rectangle visualization. We may also include a feature allowing users to search for company, which, if found, would be highlighted in the rectangle visualization. Our must-have feature will be the rectangle chart; implementing that will be our top priority, though we also hope to show macro-level data represented in bar or pie charts below. The company search is also an "optional" feature.

Project Schedule

- Monday, March 24th
 - Data collected and cleaned
- Monday, March 31st
 - Prototype of the rectangle visualization done
- Friday, April 4th
 - Full prototype complete

Overview and Motivation

The process of starting and financing a startup can be very difficult as it lacks much guidelines and structures. It is the duty of the founder to use their knowledge and ability to use their resources efficiently. In order to stay afloat, they must only explore routes that will accelerate their businesses and avoid wasting time and money pursuing those without sufficient outcomes.

One of the most important routes that a startup founder must consider is financing. This is a critical decision as the quality of investor and quantity of capital given by them can make or break a startup. But finding the right fit of investor is a very complex process. Some things that need to be considered are the industries that the venture capital fund invests in, the typical check size that they invest, typical rounds they participate in, and whether or not they have conflicts of interests. While some firms might be interested in the field your startup is investing in, they might not invest at an early stage as your company or have a conflict interest.

Most of the information about fit of venture capital to company can be obtained and inferred by previous companies that the firm has invested in. This data exists on crunchbase, but its user interface is designed for the user to look up venture capital funds and companies individually. Instead, we aim to use this information to create a visualization that aggregates and filters this information by user input, so that a user can compare venture capital funds side by side when making this decision.

As students who are passionate about the tech industry, and are interested in or have already pursued funding for their ventures, we thought this project was very pertinent. We believe we can make a visualization that would help any entrepreneur at any stage of development find out which venture capital firms they should pursue for funding based on which firms are a match.

Related Work

Our approach is similar to the visualization made by the <http://newsmap.jp> visualization that was presented to us in class. We really enjoyed the idea of having a dynamic treemap where you can apply filters to get different make-ups of the “picture.” In this visualization, you can filter the most read articles by newspaper sections (World, National, Business, Entertainment, Sports etc) and also by country. This visualization is also naturally color coded by the sections.

We believed that this would be a good visualization to model after because our individual nodes in our dataset represent a part of the picture, or the pie. In our case, for example, the objective size of Dropbox’s Series A round is important, but more importantly is the relative size of it's Series A round to the rest of its funding and the rest of the industry. As we are trying to help the audience understand which venture capital fund is a fit based on their investments, relative area makes more sense for us to visualize significant investments of the fund than absolute area. Thus, we can know the relative size of the investments the venture capital funds give compared to other similar financing rounds and compared to the entire industry.

Research Questions

The intended audience of this visualization is an entrepreneur or startup founder who is looking to find venture capital firms that fit their current funding needs based on the funding round and industry.

Thus, the main question we are trying to answer is: Which venture capital firms would be good fits for my next round of funding given my industry and round?

This question evolved since our project proposal as there were many interesting applications of the Crunchbase data. For example, the Crunchbase data can also give you an idea of the most wide-reaching companies in terms of funding. Through internal meetings, meetings with our TF, and our design studio, we decided to position the visualization so that included information mentioned above but included this information as to give context of the main question.

As a result our visualization answers other questions of the end user, some intentionally and others happenstance. They include but are not limited to the following:

- Which companies have raised the most amount of capital? Which companies have raised the most amount of capital in their industries?
- What is the breakdown of investments within each venture capital fund in each major industry (social, cloud storage, education, etc)?
- What is the breakdown of investments within each venture capital fund in each round (seed, series a, series b, series c, etc)?
- What is the timeline between funding rounds for major tech startups that have secured venture funding?
- What are the relative sizes of rounds of financing for startups?
- What are the relative sizes of rounds of financing for startups by industry?

Design Evolution and Implementation

Meeting Notes, Thursday, April 3rd, 2014

The three of us discussed the feedback we received on our project proposal and how we could effectively focus our implementation to answer these questions.

Project Objectives

We narrowed our project objective to better define terms. For example, our project looks at how “active” venture capital firms are in certain industries. We define this to be how much they have invested (in millions of dollars) in the last 5 years. We are interested in the percentage of the venture capital total investment expenditure in each market sector.

In terms of measuring performance, our visualization will not attempt to judge the value of a venture capital portfolio. This is because many startups do not have clear metrics on evaluations and these evaluations are often based on future value (instead of actual ROI). There is also much gray area in how to evaluate poor performance in venture capital as public evaluations are often only revealed in financing rounds. As a result, companies that are not performing might be evaluated at the last round of financing, but truly may be on their way to bankruptcy (and just not have public data to support). Instead our visualization, will look at performance as whether or not a venture capital firm chose to participate in future rounds of a company’s financing. This is a good indication if a firm believes the investment’s performance has met or exceeded expectations.

The intended audience of the data visualization would be someone who is looking to finance their startup or someone interested in the field. For the founder, the visualization should provide information on whether or not certain venture capital firms seem to be a fit as an investor. They can judge whether the firm’s preferred sectors/industries match theirs, whether their average investment size is comparable to

Data Spec

Financial Orgs

- Array of financial org objects, each in the following form:

Key	Description	Sample Value
name	Proper name of the financial organization.	Accel Partners
total_invested	Total capital invested. Note: this is just a sum of the amounts of <i>all</i> rounds the firm was involved in. So, this is a very inflated estimate (particularly for firms involved in later-stage investing).	9428541785
num_companies	Integer value, total number of companies invested in.	301

Example JSON

```
{  
  "name": "Accel Partners",  
  "total_invested": 9428541785,  
  "num_companies": 301  
}
```

Industries

- Array of industry objects, which hold arrays of company objects, each in the following form:

Industries, industry objects

Industries, industry objects

Key	Description	Sample Value
name	Industry name	network_hosting
children	Array of company objects (see below)	(see below)

Industries, company objects

Key	Description	Sample Value
name	Company proper name	Dropbox
	Company permalink, for	

permalink	lookup in other data, or from the Crunchbase API (sometimes this is different altogether from the plaintext name)	dropbox
founded	Integer value, year founded, may be null.	2007
industry	Industry category	network_hosting
tags	Comma-delimited list of tags describing the company. May be null.	techcrunch50, tc50, file-storage
overview	Description of the company. May be null.	Dropbox was founded in 2007 by Drew Houston and Arash Ferdowsi. Frustrated by working from multiple computers, Drew was inspired to create a service [...]
city	City the company is located in. May be null.	San Francisco
state	State the company is located in (US postal code). May be null.	CA
country	Country the company is located in (abbreviation). May be null.	USA
twitter_handle	Company's Twitter handle. May be null.	Dropbox
crunchbase_url	URL to the company's Crunchbase profile.	http://www.crunchbase.com/company/dropbox
url	URL to the company's website. May be null.	http://www.dropbox.com/
	URL to the	http://www.crunchbase.com/assets/images/resized/0001/1969/11969v19-

"Goldman Sachs",
"Sequoia Capital",
"Valiant Capital Partners",
"Glynn Capital Management",
"Index Ventures",
"Accel Partners",
"RIT Capital Partners",
"BlackRock",
"SV Angel",
"Benchmark"

],
"tags":"techcrunch50, tc50, file-storage",
"url":"http://www.dropbox.com",
"overview":"Dropbox was founded in 2007 by Drew Houston and Arash Ferdowsi. Frustrated by working from multiple computers, Drew was inspired to create a service that would let people bring all their files anywhere, with no need to email around attachments. Drew created a demo of Dropbox and showed it to fellow MIT student Arash Ferdowsi, who dropped out with only one semester left to help make Dropbox a reality. Guiding their decisions was a relentless focus on crafting a simple and reliable experience across every computer and phone. Drew and Arash moved to San Francisco in fall 2007, secured seed funding from Y Combinator, and set about building a world-class engineering team. In fall 2008, Sequoia Capital led a \$7.2M Series A with Accel Partners to help bring Dropbox to people everywhere.",
"industry":"network_hosting",
"people": [
 {
 "name": "Drew Houston",
 "title": "Founder & CEO"
 },
 {
 "name": "Arash Ferdowsi",
 "title": "Founder & CTO"
 }
,
 "state": "CA",
 "children": [
 {
 "date": "09/04/08",
 "currency": "USD",
 "amount": 1200000,
 "name": "seed",
 "investors": [
 "Sequoia Capital"
]
 },
 {
 "date": "11/24/09",
 "currency": "USD",
 "amount": 6000000,
 "name": "a",
 "investors": [
 "Sequoia Capital",
 "Accel Partners"
]
 },
 {
 "date": "10/18/11",
 "currency": "USD",
 "amount": 10000000,
 "name": "b",
 "investors": [
 "Accel Partners",
 "Benchmark",
 "Index Ventures",
 "RIT Capital Partners",
 "Sequoia Capital",
 "SV Angel",
 "Valiant Capital Partners",
 "BlackRock",
 "Bloomberg",
 "Goldman Sachs",
 "Glynn Capital Management",
 "Index Ventures",
 "Accel Partners",
 "RIT Capital Partners",
 "BlackRock",
 "SV Angel",
 "Benchmark"
]
 }
]
 }
]

```
"amount":250000000,
"name":"b",
"investors":[
    "Benchmark",
    "SV Angel",
    "RIT Capital Partners",
    "Accel Partners",
    "Index Ventures",
    "Glynn Capital Management",
    "Valiant Capital Partners",
    "Sequoia Capital",
    "Institutional Venture Partners",
    "Greylock Partners",
    "Goldman Sachs"
]
},
{
    "date":"01/24/14",
    "currency":"USD",
    "amount":350000000,
    "name":"c",
    "investors":[
        "BlackRock",
        "T. Rowe Price"
    ]
},
],
"founded":2007,
"total_raised":607200000,
"company_permalink":"dropbox",
"image_url":"http://www.crunchbase.com/assets/images/resized/0001/1969/11969v19-max-450x450.png",
"country":"USA",
"crunchbase_url":"http://www.crunchbase.com/company/dropbox",
"employees":642
}
]
```

{}

the amount of money they are raising, and whether the firm is a long term fit based on whether the firm has historically reinvested in many of their early round investments.

Data

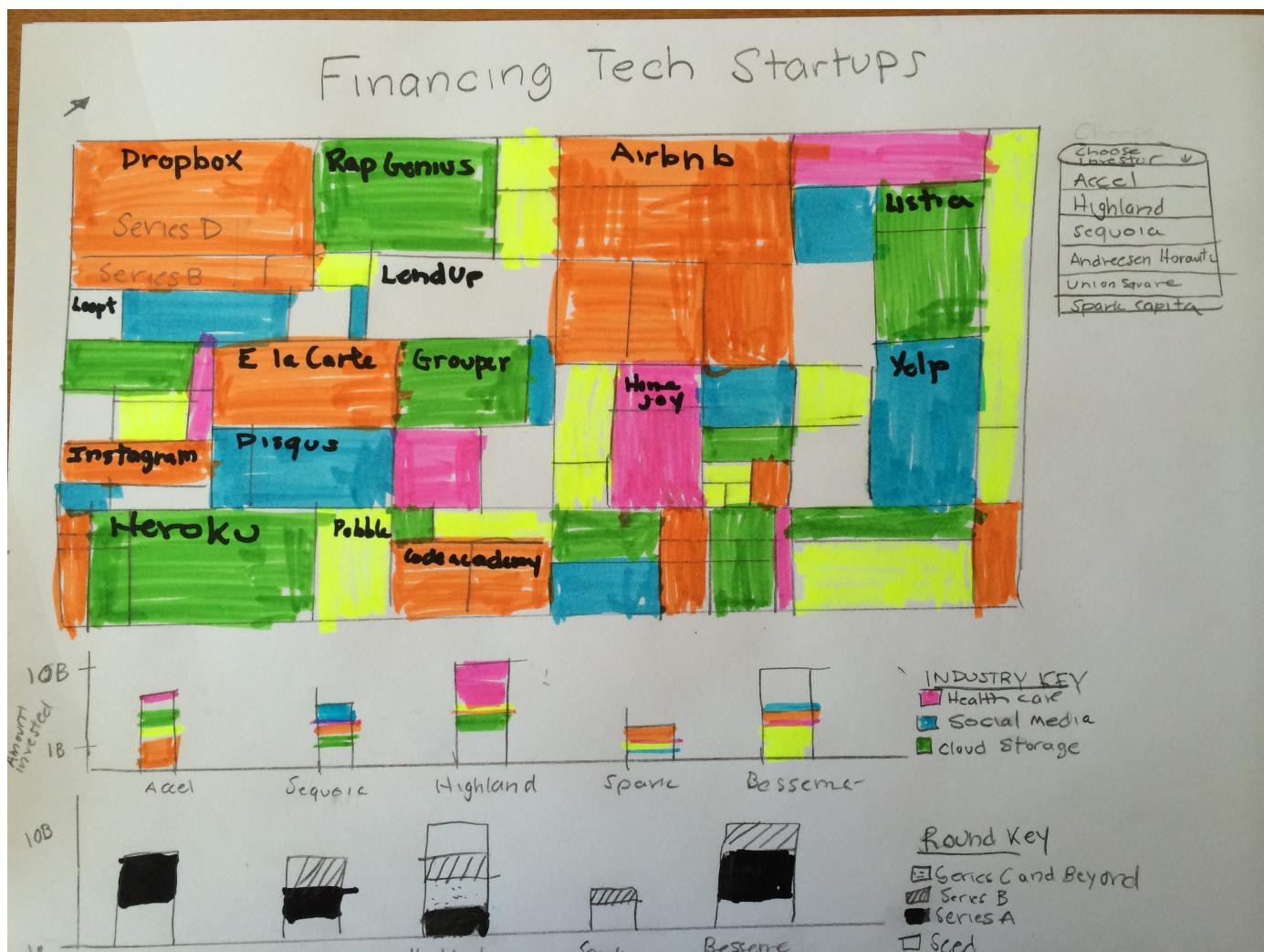
At this point, we have successfully scraped the data from the Crunchbase API and parsed it into an actionable format for our visualization. As the data is quite expansive, we decided to download it locally and break it down into separate files to store and match information.

We parsed the API with python and converted the data into three JSON files. These JSON files focus the information around company (startup), financial organization (venture capital firm), and rounds (e.g. seed, series A, etc).

Visualizations

We continued to flesh out the structure of our visualization and talked about how to increase our scope through the interaction with our visualization and supporting graphs. We include the first visualization (Visualization 1) created, below:

Financing Tech Startups



In this visualization, we show our main tree layout visualization, and the two subgraphs. In addition, we include an dropdown menu that allows the user to focus its attention on different areas of the graph.

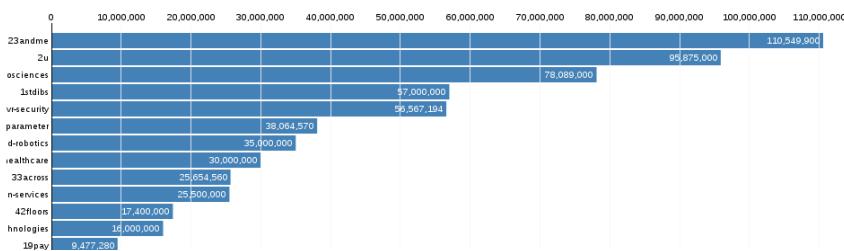
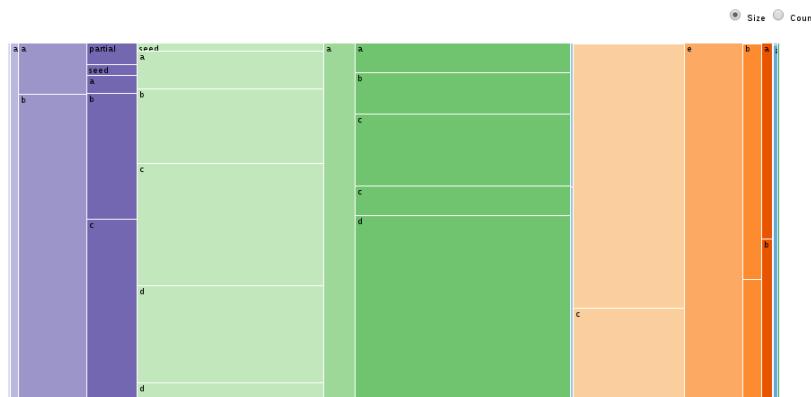
Meeting Notes, Sunday, April 6th, 2014

At this meeting, we met as a group discuss the progress of the data scraping, process book, and visualization as a whole. We also had a skype call with Alain Ibrahim.

Visualization

We had implemented our visualization based on the data structures described above, but the format was not exactly as we expected. Based on the way we obtained our data, a company in our visualization was represented by a column rectangle, instead of just a box. This is because we formatted our funding data in such a way that it only had one parent: the company. We realized that this visualization needed to be improved as the large lengths of the visualization often hid important information or made that information unreadable. This also obfuscated the relative size of rounds, as a company with only one or two rounds of funding filled an entire column

The visualization looked as followed:



We decided to restructure the data, so that it had two parents, industry, and company, and thus be 2-dimensional boxes, instead of columns. The new data structures are included in a separate document.

User Interaction

One of the major discussions of this meeting was how to handle user interaction with the visualization. At this point, we had scraped the data, and created a basic treemap, but we still need to interact with the data in such a way that it meets our end goal of helping the user decipher information about whether a certain venture capital fund is a fit.

While the subdetail visualization above is informative, it made our visualization look like it is oriented such that we are trying to tell our audience about companies and not about venture capital firms.

As a result, we came up with a few detail graphs and interactions to help user flow and provide more detail to our user:

1. When you hover over a venture firm in the subdetail graphs in Visualization 1, it will highlight the specific rounds within the companies that this venture firm has participated in on the treemap
2. When you hover over a round on the treemap, a tooltip will appear that will display the size of each round and possibly the participants
3. A user has the ability to choose the industry to explore in a drop down menu
4. A user has the ability to choose multiple rounds to explore by checking them off. Thus giving the user the ability to see standalone round comparisons, or round comparisons relatives to other rounds

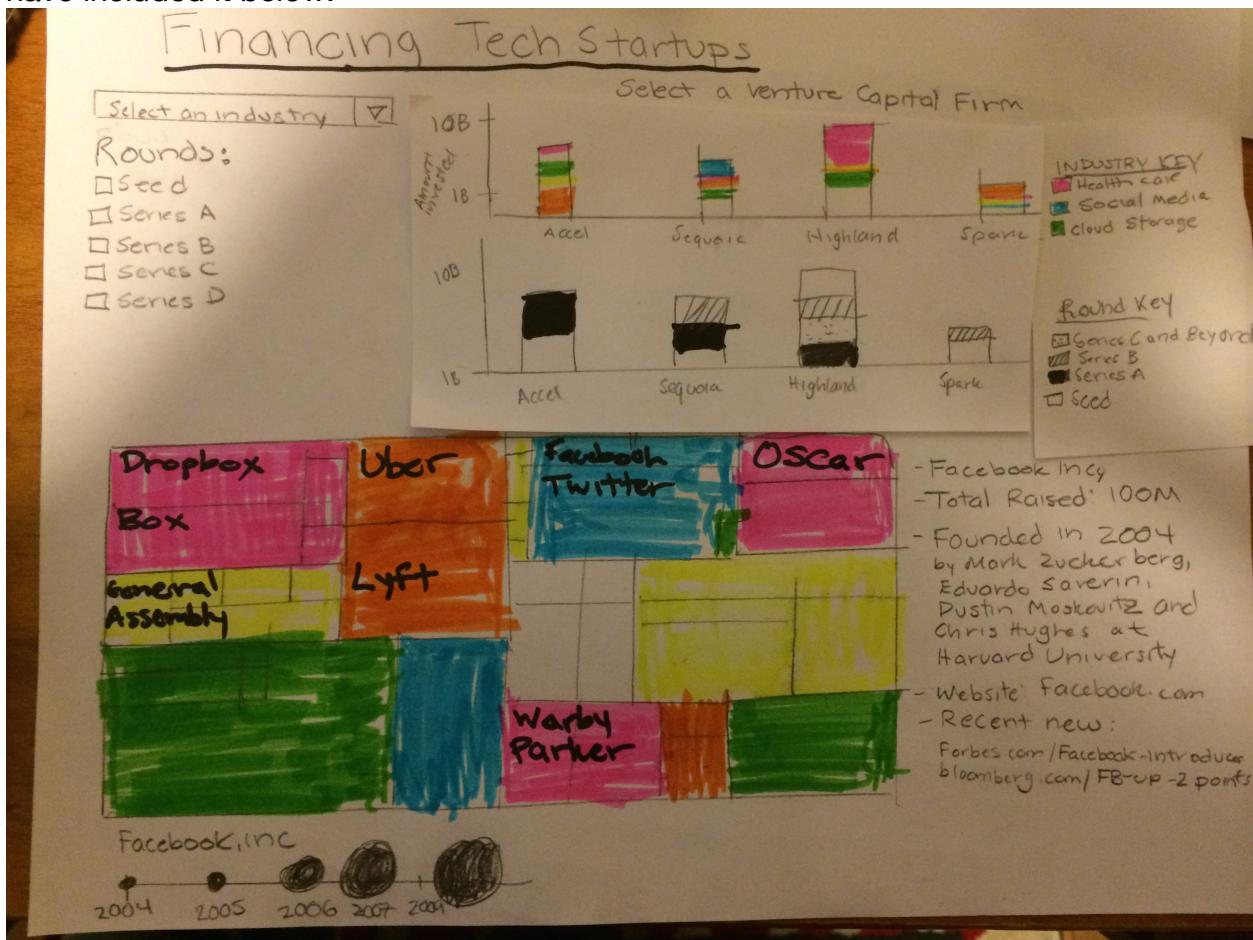
Design Studio, Tuesday, April 8th, 2014

See the design studio document to understand the reasoning behind the changes we decided to implement based on the feedback from our classmates.

As discussed, the most important visualization changes we have chose to make based on the feedback are as follows:

- Add Titles to Guide the User to Choose A Venture Capital Firm
- Include Venture Capital stacked bar charts before the treemap so that the user clicks on a venture capital firm and isolate portfolio companies on the tree map
- Include a timeline graph where size of the round (in millions of \$) is displayed in the radius of the circle.
- Include tooltips with additional details about the companies, such as recent financing events, founders, websites, founding date, other recent news, etcs.

We sketched our new vision of our visualization based on these actionable items and have included it below:



As you can see, the treemap is now broken down by an extra dimension (industry) and is below the subdetail graphs of the venture capital firms so that the end user can select the venture capital firm. This visualization also includes our selectors (industry, and financing rounds). Finally, this redesign has a timeline graph of the highlighted company as well as a story of that company on the right that includes its total capital raised to date, founders, founding date, and recent news.