

VENTURES

Final Project Proposal

Guillaume Laliberté

CSCI E-171: Visualization

Prof. Alexander Lex

April 2015

Ventures - Proposition

1. Background and Motivation

As a business-inclined person, I have always been interested in how people create companies from scratch. I myself have founded a few companies, and as I am getting better at it, I try to understand how various industries work, and how funding and financing plays a role in companies' success.

This is why I decided to create a visualization that would capture various dimensions of the startup world. To me this project is not just school work: it is a tool I intend to use to comprehend how the startup scene is evolving, and how I can use that data to my advantage.

2. Project Objectives

The main objective of this project is to highlight patterns and trends about what kind of companies get funded, and uncover correlations between industry, state of the market, capital and successful exits/IPOs.

To do this I need to:

- Identify a few datasets that are related to each other (Series A, IPOs, VC, acquisitions, etc.)
- Implement interactions like brushing, industry filter, region filters, VCs
- Maybe match those dataset to another external dataset (i.e. S&P 500)

3. Data

Most of the data I will need is available from the *National Venture Capital Association* or the *Kauffman Foundation* and through *Quandl*. Additional industry and financial data is available from *Quandl* and sources like *Google Public Data Explorer*.

Among the dataset (from the four sources mentioned above) I think will be useful are:

- VC deals per sector
- Deals by stage
- VC-backed IPOs
- Kauffman indexes
- Stock market and other financial data

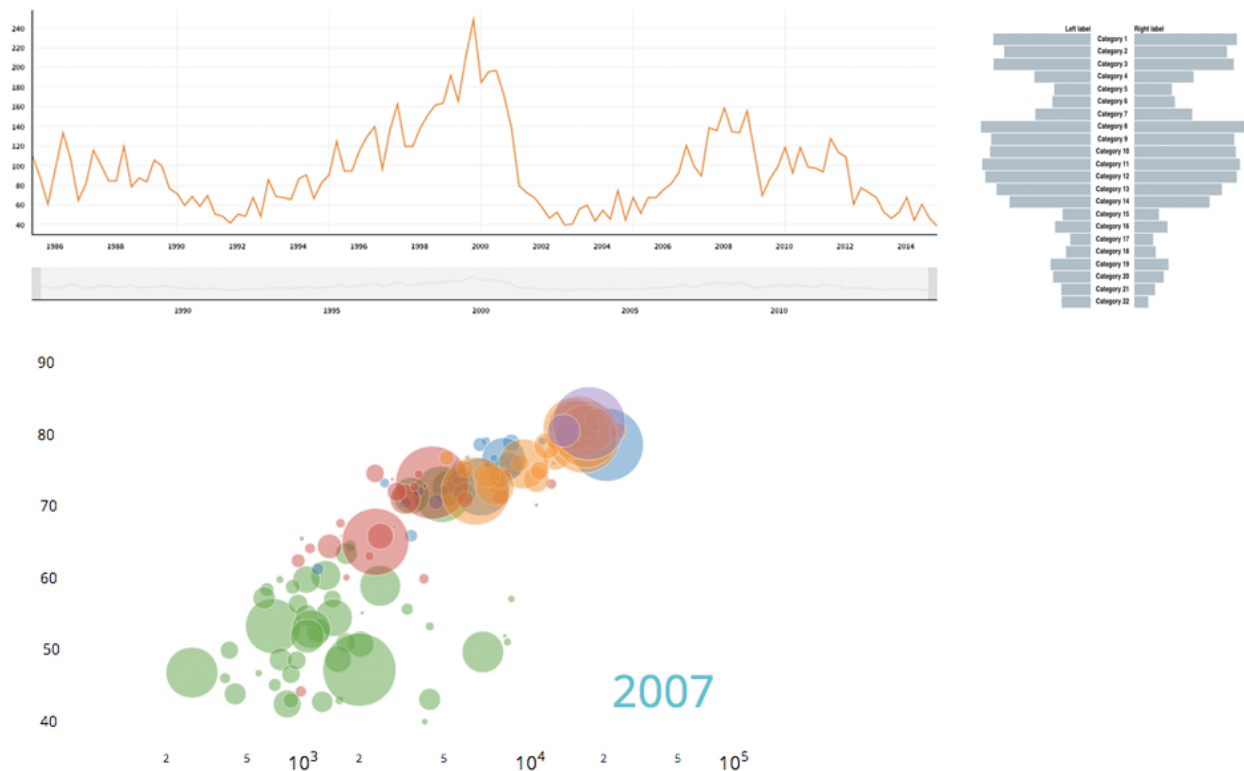
4. Visualization

There are a few element I think could be useful to identify patterns. It seems like one line chart or stacked area chart could be the centre piece where most of the interaction happens. It allows the use to compare a few datasets (with filters) and to select a time period (brush).

To compare the categories, a pyramid chart will be used. A bubble chart will also be a good solution to compare the institutional investors and startups in terms of deals count or size.

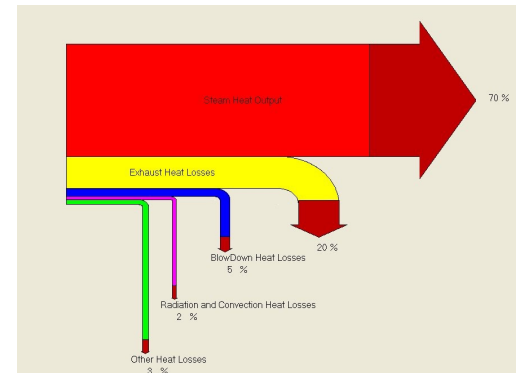
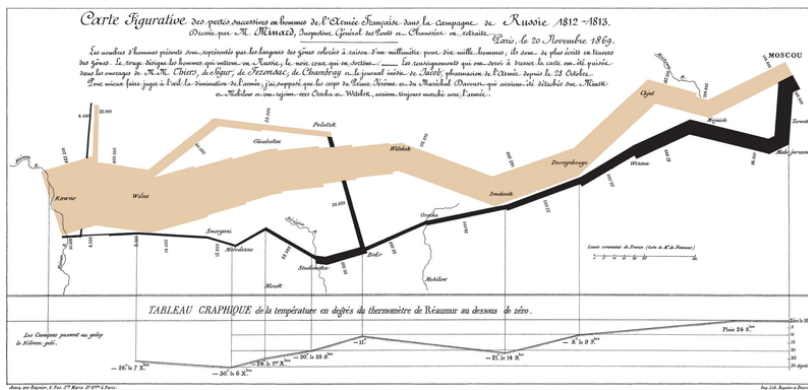
One thing I would really like to achieve is a Sankey diagram: it could be useful to visualize how many companies go bankrupt at each round of funding (from early stage to IPO). Drawing that data will make it easy to see how startups fail or succeed. This will be an optional feature because of time constraints, but it will be the first I complete if I have enough time.

This is a first, very rough sketch using dummy charts¹ to explore ways to layout the information. I have more work to do on the data in order to refine the visualization. One thing I would like to improve already is to have only a small line chart with a big Sankey chart and Bubble chart.

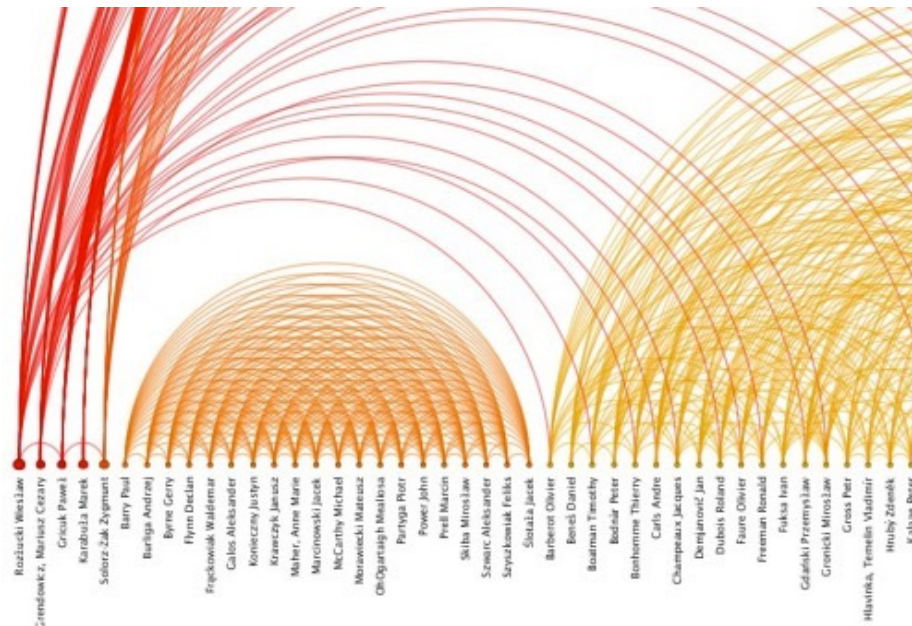


¹ Quandl, plot.ly and <http://www.jasondavies.com/d3-pyramid/>,

There are a few other graphs that I think are really inspiring and which could be used in this project. The first one is a classic, one of my favourite graphs featured in-class. I think it could be really useful in showing how startups "die" over time². Similar graphs are often used to illustrate energy consumption³:



If I can find the data needed, another interesting alternative would be to illustrate the relationships between startups and investment firms. This would allow the user to quickly identify how firms rely on each other. It could be done with an arc connection graph:



² <http://tomtunguz.com/images/sankey-minard.png>

³ http://3.bp.blogspot.com/_XswVxxXMFWE/TJdHL-IK-ol/AAAAAAAAABq/Wlbe3ZvHzgE/s1600/SankeyDiagram.bmp

5. Must-Have Features

Because of time constraints, I will focus first on the interaction and basic graphs that display the most relevant data. Hopefully I will be able to add a Sankey graph or arc connection graph.

- Line / area chart with brush (main control for the visualization)
- At least 3 relevant datasets
- Bubble chart to compare startups or investors (performance indicator and comparison)
- Pyramid to compare sectors (market trends and comparison)

6. Optional Features

- Sankey diagram or arc connection graph
- "Zoom brush" (<http://bl.ocks.org/mbostock/1667367>)
- Stock market data
- More datasets
- A way to include any dataset by calling an API on-demand (Quandl api for example)

7. Project Schedule

Proposed schedule for the project:

Date	Goal	Turn in
Apr 5, 2015	Project Proposal Due	Proposal
Apr 9, 2015	Project proposal update. More sketches.	
Apr 10, 2015	Refined sketches, Initial datasets selected, at least one API call working and chart view working	
Apr 17, 2015	Completed data acquisition, Data structure ready, working prototype with views working, no interaction	Process Book Code
Apr 24, 2015	Basic interaction possible, website working + style	
May 1, 2015	Cleaned up code, Screen-cast, integrate final feedback	
May 5, 2015	Due Date, hand in everything. Includes optional features if possible.	Process Book Code Screen-cast Website Data + other stuff