Project 1 Report

TA Mentor: Haewon Hwang (hh474)
Team Members: Josh Zheng (zz459), Val Mack (dgm97), Jiangjie Man (jm2559)

Introduction

A **unicorn** is a start-up company valued at over $1 billion. The billion-dollar technology startup was once the stuff of myth. Today they're seemingly everywhere, backed by a bull market and a new generation of disruptive technology. There are a few early clues that a unicorn will be successful, such as "Is the business in a hot sector, one where many other new startups are also focusing?" "Has it raised funding at a quick pace?" Based on those criteria and others, we studied unicorns (worldwide) valuation and venture capital fundings and built visualizations using financial data from CB Insights and Quandl.com.

Data Description

- **Top 10 Unicorns Funding History and Valuation Data**

This data has all funding and valuation data from 2006 to 2016 of top 10 unicorn companies scraped from CBinsights. We used an ipython notebook and selenium to collect this data from the CBInsights web pages (legally obtained!). We then used python to export the data as json and csv files.

- **National Venture Capital Association Data (NVCA)**

This data is a comprehensive annual data on the venture capital industry in the U.S. covering investments, deal-making, startup valuations, as well as IPOs and Acquisitions. It is available on https://www.quandl.com/data/NVCA-National-Venture-Capital-Association-Data.

For the second graph, we used Venture Capital Investments By Stage dataset from NVCA yearbook, which includes 5 variables: *"Date","Seed","Early Stage","Expansion","Later Stage"*. The tricky part is that the format of the data is an array of arrays, e.g., [["2015-01-01", 1, 2, 3, 4], ["2016-01-01", 5, 6, 7, 8], ...], which means we have to transform it into an array of objects, e.g. [{date: 2015-01-01, Seed: 1, Early Stage: 2, Expansion: 3, Later Stage: 4}, {date: 2016-01-01, Seed: 5, Early Stage: 6, Expansion: 7, Later Stage: 8}, ...] by using *map( )* in the callback function of *d3.json( )*.

For the third graph, same techniques were used. We transformed an array of arrays into an array of objects for plotting multi-line graph. For this one, we used Venture Capital

Investments By Industry dataset but only selected a limited number of industries related to IT and Finance areas instead of 17 different industries, which are *"Business Products and Services", "Consumer Products and Services", "Financial Services", "Industrial/Energy", "Retailing/Distribution", "Software", "Telecommunications"*.

- **World Map Data (world-50m.json)**

This map shows the Natural Earth "Admin 0 - Countries" shapefile at 50m resolution. Converted to *TopoJSON*, only 140K gzipped.

Mapping from Data to Visual Elements

- **Position (X, Y axis)**

We used D3's date scale function for X-axis since we want to measure companies' growth over time. For Y-axis, the values are in USD million. It shows the valuation for each unicorn or Venture Capital fundings for each stage or industry.

- **Color**

For aesthetic quality, we choose bright colors. Also use D3's *schemeCategory10* and *schemeCategory20* library for multi-line graph.

- **Shape (circles)**

Circles are calculated using square root function so that they can be scaled properly.

- **Transformations**

The transform attribute was used to position different elements within each visualization to convey information in a clear manner. This was particularly important in our map to organize the small multiples, and in the legend, to display each component in an easily digestible way.

The Story

- **Top 10 Unicorns Worldwide**

This visualization displays the dramatic rise in funding for tech companies that occurred within a very short amount of time. The aim is to show the viewer how quickly some of these companies are raising money, and also give context around who these companies are and where they are located. This information informs the viewer about the content we are introducing to them. It is meant to be surprising to see how the funding amounts differ based on regions of the world as well. For example, the United States has older unicorns than Asia.

- **Venture Capital Investments By Stage**

Our second visualization tells the viewer when companies are raising the most money. Venture Capital (VC) is the most heard of, and in many cases important for of capital for startups, so we chose to focus on VC investments. This graph is surprising because it shows how dramatically funding has declined since the dot com crash of 2001. It also makes the viewer curious, because it shows that VC investments have decreased over time, while the prominence of unicorns has increased over time. We also break down the funding stages (seed, early, expansion, later) so that the viewer can see what stage of a company gets the most funding and how funding per stage has changed over time.

- **Venture Capital Fundings By Industry**

Our final visualization breaks down VC funding in terms of industry. It is similar in shape to the one above but detailed in a different way. The surprising part is that it shows the strength of software startups in terms of the amount of investment they get and have gotten compared to all other industries who are flatlining. The viewer who knows a little more about startups and their valuation may be worried that another crash, similar to the one seen in 2001, is coming.