**The Changing Landscape of the American Economy**

Katherine Dai and Patrick Xu

**I. Initial Proposal**

**Background and Motivation**

In 1908, Henry Ford began producing and selling the Ford Model T, decimating the formerly robust buggy and whip industry. More recently, Amazon has displaced booksellers and companies in the technology space have dominated the economy in both in terms of overall market cap and growth. We are interested in seeing how the American economy has evolved over the years, both in terms of what the most successful companies do and how big they are.

**Project Objectives**

By analyzing trends in the companies in the S&P 500 market index, which is composed of the 500 largest companies by market cap on the NYSE, we hope to not only visualize the overall growth in the American economy but also to break down the components of this growth.

* How big is the American economy each year?
* What are the ten largest companies in the economy for each year?
* What do these companies do?
* Where are these companies located?

There are several benefits to this project. First, we will learn how to geomap data and create a visualization using a new structure. Additionally, we’ll be able to strengthen our API manipulation skills. Finally, we’ll be able to discern trends in industry movements and get a better understanding of where the American economy has historically been and what it seems to be trending towards.

**Data**

We are planning to collect our data primarily from Baker Library, where we can gain access to historical S&P 500 indices and to the Wharton Research Data Services (WRDS) to collect company financial information. Additionally, we plan on using the Yahoo! Finance API to help better categorize each company by industry and sector.

**Data Processing**

Although we are getting data from multiple sources, the actual cleanup process should be relatively painless. We plan on grabbing the top 500 companies by year according to the S&P 500, from 1955 (or potentially earlier) to current day, and putting it in CSV format. From Yahoo! Finance, we will retrieve the industry for each of these companies, as well as the market cap during the fiscal year; this can be directly downloaded in CSV format. Finally, WRDS will provide us the headquarters of the companies, from which we will likely only store the state name.

**Visualization** How will you display your data? Provide some general ideas that you have for the visualization design. Include sketches of your design.

There are a number of features that we are planning on implement to visualize the S&P 500 and its historical growth. First, we want to represent current and historical market cap with a line chart of the S&P 500. This line chart will also include interactivity - users will be able to focus in on a year to better understand growth and market cap, and hovering over any year will reveal a tooltip with numerical data on the S&P 500. Next, we want to focus in on the top 10 companies in the S&P 500 (by market cap), and create a bar chart that will show each individual market cap. Additionally, we want to color the bars based off of their industries. Finally, we want to geomap the companies by state onto the US. We want to see how the headquarters of the S&P 500 companies have changed over time.

**Must-Have Features**

* The line chart with zoom in and tool tip capabilities
* Bar chart that has industry/market cap information
* Map of the US that depicts number of companies by states

**Optional Features**

* Features that would be nice to implement:
* Play button that shows changes over time
* Companies mapped by cities rather than states
* Pie chart that shows industry breakdown
* Markers for recessions

**Project Schedule**

Week of March 23 - Aggregate data, map geographic location to companies and their headquarters (using multiple APIs), create a line chart showing overall growth of the US economy

Week of March 30 - Associate industries with company, create initial US map, create bar chart representing market cap of top 10 companies in the US (color bars by industry)

Week of April 6 (first version due on Thursday, April 10) - Create a play button that will allow viewers to see how headquarters/industries have changed over time.

**II. Initial Sketch**



**III. Project Milestone: Revised Proposal**

Changes in data

* using U.S. GDP data in place of S&P 500
  + the S&P 500 is only representative of the stock market, whereas GDP gives a much bigger picture of the American economy and is more relevant to the trends we’re trying to portray
* data processed via different process than expected
  + not all 500 companies on the Forbes list every year were used - only the top 200 were used, since it’s a good representation of that particular year and more data points per year would crowd the visualization
  + we were also constrained by time

Changes in visuals

* front-end inspiration found at <http://remittances.herokuapp.com/?en>
  + the layout and sleek design appealed to us, and we hope to construct a visualization with a similar color palette

What we have finished

* line that draws as time moves
* functionality for play button that cycles through years
* initial list of top 10 companies (will be dynamic in future)
* complete dataset of top 200 companies from 1955-2005 (including industry)

What we are looking to complete soon

* pie chart that is attached to the play button to show changes in industries
* implement a front end design that will increase readability and aesthetics of visualization
* interactivity with the map, so users can hone in on what they want to learn (possibly the largest part we need to implement)
* choropleth map in addition to dots to represent company density
  + a toggle would allow the user to switch between the two - choropleth allows for easier comparison while dots show the more detailed data (and more movement)

Concerns

* data readability
  + there’s a high density of graphs on the screen, which makes the overall visualization seems a bit fragmented (though interactivity will help increase the sense of cohesion)
* emphasizing movement
  + one thought was to implement a visual element that shows the path of a company moving from New York to California

**IV. Updated Sketch**

