Process Book

Brian Park
Jared Nielson

Overview and Motivation

Income inequality has steadily grown since the mid 1970's. In 1975 the top 1% of earners captured 8.9% of new income created, their lowest share in the past 100 years. Since then the percentage of total income for top earners (both in the top 10% and top 1%) has increased, peaking at 50% and 22% respectively just prior to the 2008 financial crisis (Saez, 2016). Total income for top earners dropped post 2008, but quickly recovered. From 2008 to present the top 1% of earners have more than made up for losses in the crisis increasing their income from 2008 to 2012 by 34.7% while the bottom 99% saw incomes increase by a modest 0.8% during the same time period (Saez, 2016).

Why is this concerning? Income inequality isn't just about money, it is about quality of life. The gap in life expectancy between those in the top 1% and those in the bottom 1% is 14.6 years (Chetty, Stepner and Abraham, 2016). Millions of Americans cannot afford health insurance and lack access to health care, affordable housing, and even food. This is a life and death issue.

In December of 2017 congress passed and President Trump signed into law tax cuts for the wealthiest of Americans. Among many, especially those at the top there seems to be an unwillingness to accept the data at its face value, that as the richest prosper the middle and lower class struggle. One of these reasons may be that most people can't internally visualize just how skewed the income distribution is. Most papers on income inequality present graphics such as line and bar charts, but fail to create compelling and interactive visualizations that drive home the reality of income inequality. Our goal is to create a visualization that will drive home both the reality and the scale of income inequality.

Related Work

In the past, we have learned about disparities that are present in both class as well as race from schooling as well as politics. It is clear that a large percentage of our societal burdens are caused by factors that are a direct result of unfair income differences. In class we have also watched visualizations about the distribution of wealth in the USA and just how mind boggling it is.

Questions

When we started this project our main questions we wanted to answer:
1. What does the income distribution in the United States look like.
2. Do people have an understanding of what that distribution is before seeing it?
3. Are there areas of the country where economic inequality is greater than others?

As we dove further into the data we had available, as well as explored different visualizations our questions changed. We were able to see stark inequality between different racial groups as well as overwhelming aggregation of income at the top. Our questions we wanted to answer became:
1. How has income distribution changed over time in the USA? And which groups have made the most economic gains?
2. Just how large is the gap in income between various groups and can it be effectively visualized?
3. How does race factor into income in the USA?
4. What events have influenced wealth distribution in America?

In our design evolution section we will discuss how our designs evolved to answer these questions.

Data:

The all data was parsed from various excel sheets and data files from the Census Bureau website (www.census.gov). Not only did the census data include aggregate data, but it also included breakdowns by race. We decided that because of this availability we would include race in our visualization. We were able to show additional information in our project helping delineate some of the points that represent the thesis of the project about unjust disparities in society involving money. The raw data was manipulated by hand to generate well formed csv files and are then loaded into the project on initialization.

Exploratory Data Analysis

Looking at our data and exploring major historical events that happen in corresponding years we can see very interesting trends that have occured over time. There are several articles that discuss events and how this affected income, taxes, and the wealth gap. On further exploration, our data sources also had racial statistics available readily so we believed it would benefit our project to include them to further drive the thesis of unfair policies in the USA.

One of the first visualizations we did (which we include in a more polished form in the final project) was a simple line chart shown below. This visualization allowed us to explore the data in various combinations and really highlighted income disparity in the United States.
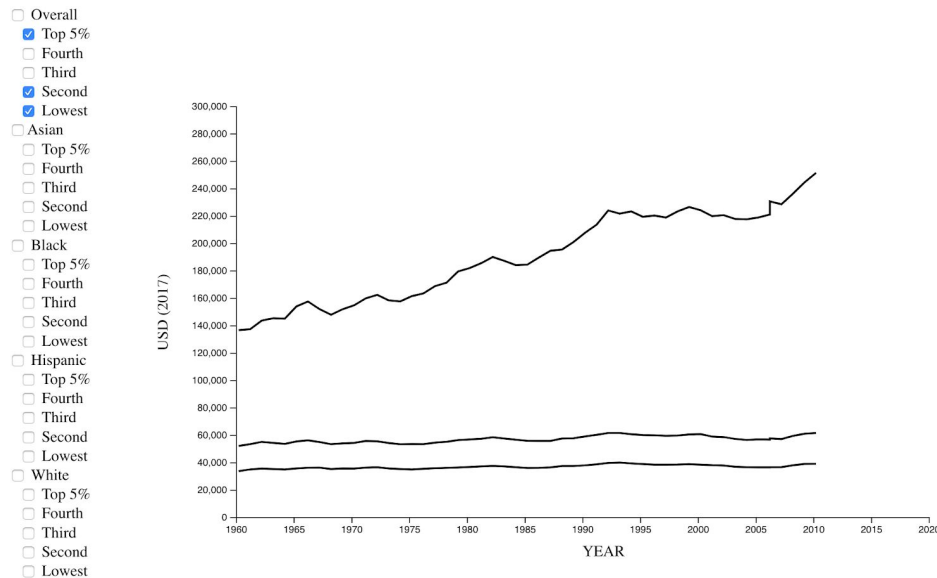

Design Evolution

Although wealth disparity is more of a problem in the USA compared to income disparity, wealth visualizations have been done frequently, for example the video about wealth that we watched in class is a great example. We decided to focus more on income and tax rates, with some highlights about wealth. Our main data focus however, changed to income, and with the ease of access racial disparities in income.
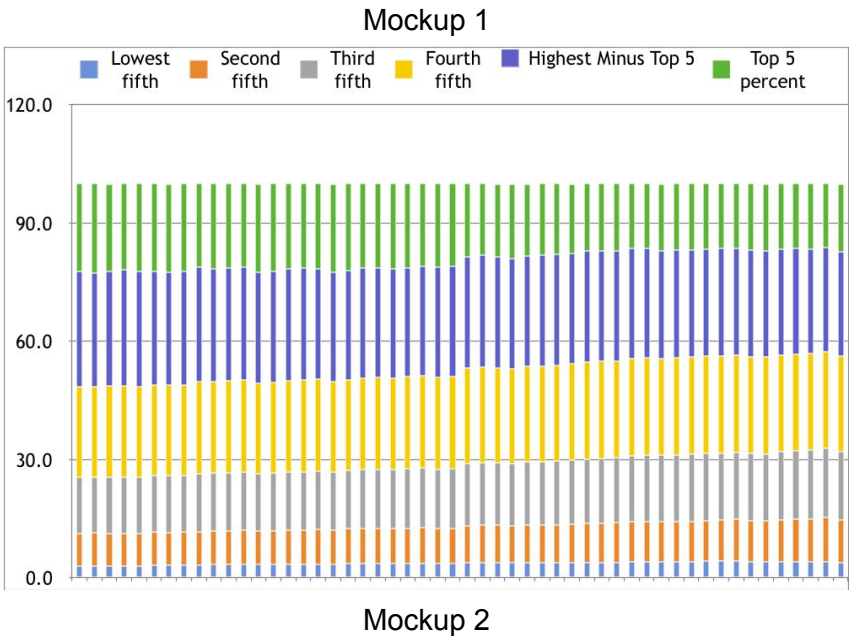
The data we were able to get from the census bureau's website we felt was best visualized with a line chart. However, there were many different racial categories for income (Overall, White, Hispanic, Black, Asian) and within each category 4 quartiles and the top 5%. This led to 25 different sets of historical income data. An interested user might want to explore or compare any number of these data sets, for example the top 5% overall vs the top 5% of black earners. Rather than clutter our line chart with 25 different lines we decided to implement selectors so users can select which sets of data they want to see. Color encodes the category while y position encodes income and x position encodes time. Unwanted data is hidden away.
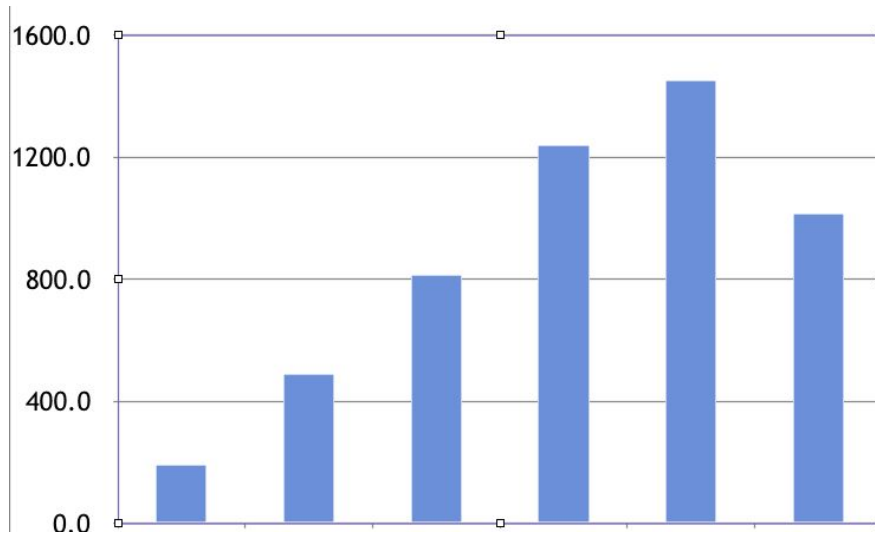
The first prototype of the line chart is shown above. The final version is much more polished. We decided to use a step line generator since that made more sense. We also encoded category with hugh and pentile with saturation added a legend, brushing to show wealth gaps and the

poverty line for 2017. The controls from the original design are not shown but are below and are used to link this chart with the aggregate income chart.
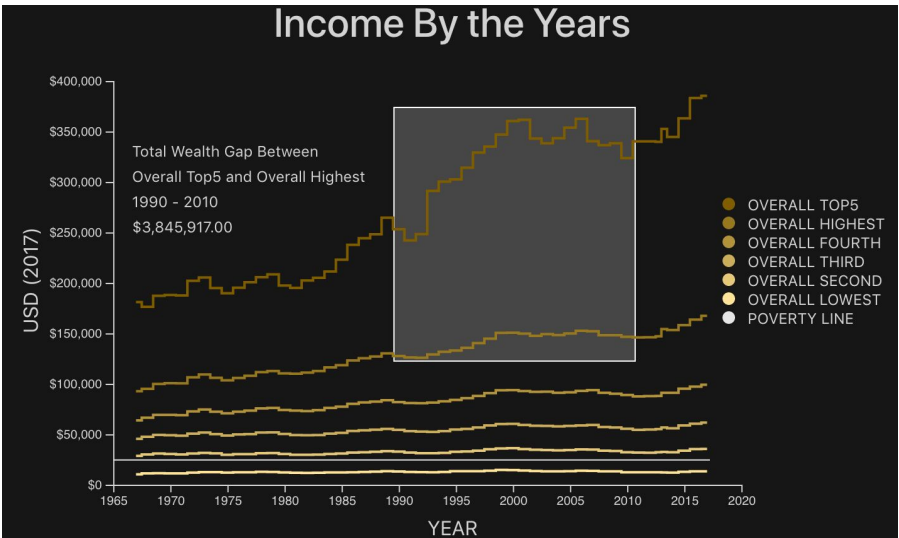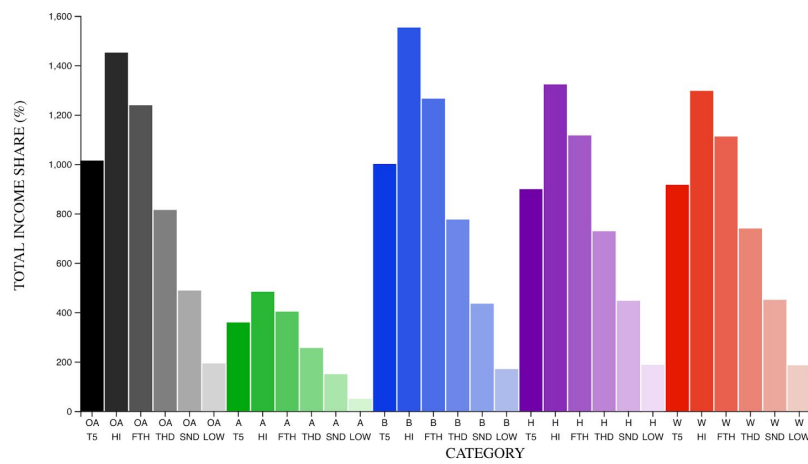

The Census Bureau had similarly formatted data for total income shares that we wanted to use as well. It made sense to link another view to our checkbox controller since the demographic categories were identical. Originally we wanted to do a stacked bar chart (where each demographic category makes up a percentage of the whole), but that ended up mirroring the line chart data and being uninteresting. We thought it might be more interesting to show total income shares by demographic instead. Below are our mockups for this process and the final result.
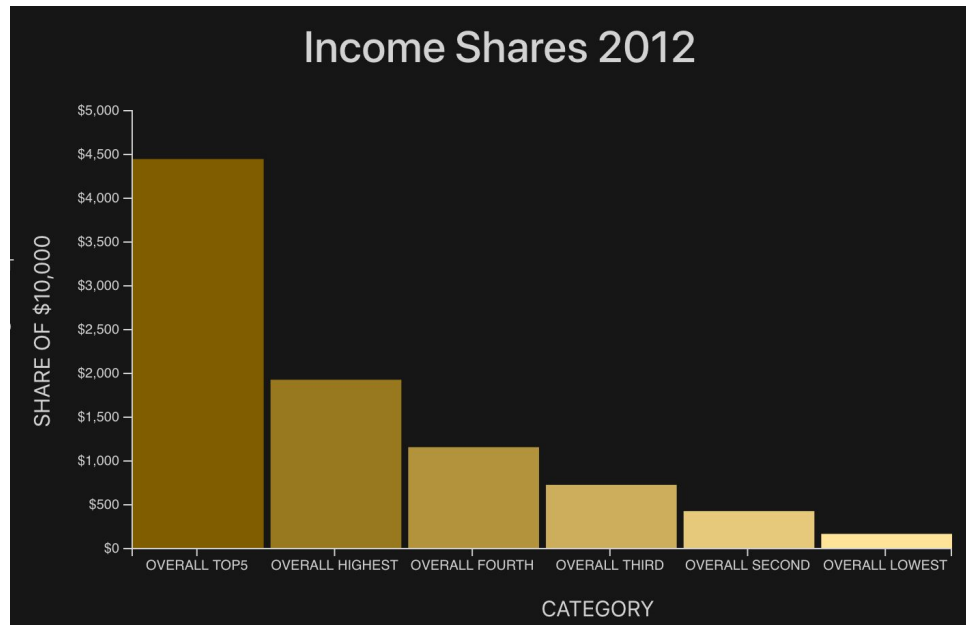
Mockup 1

Mockup 2

First Pass

Aggregated Income Shares

TOTAL INCOME SHARE (%)

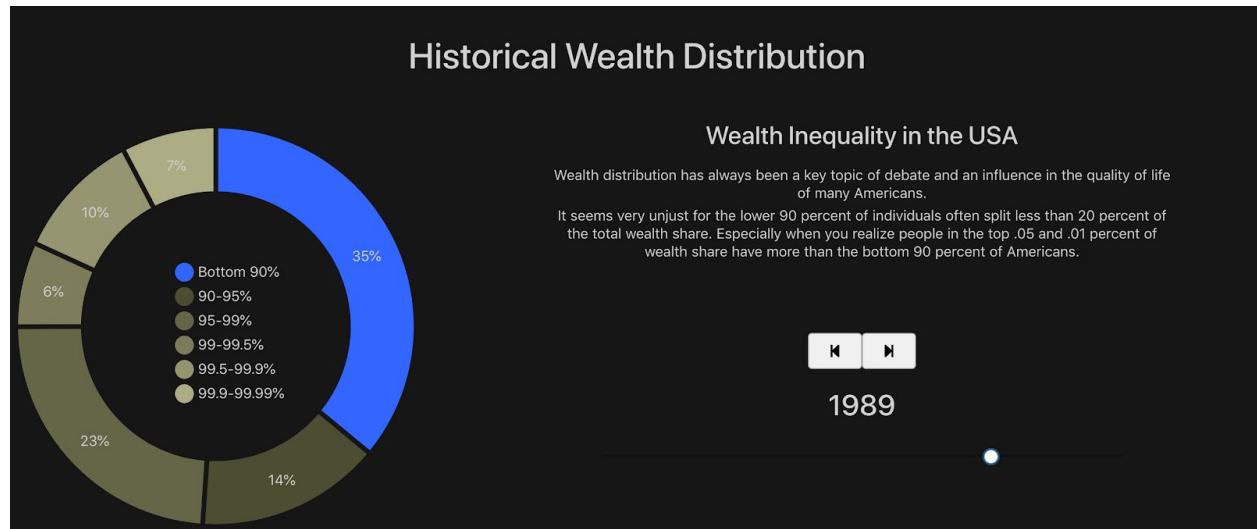| OA T5 | OA HI | OA FTH | OA THD | OA SND | OA LOW | A T5 | A HI | A FTH | A THD | A SND | A LOW | B T5 | B HI | B FTH | B THD | B SND | B LOW | H T5 | H HI | H FTH | H THD | H SND | H LOW | W T5 | W HI | W FTH | W THD | W SND | W LOW |

CATEGORY

Hightest contains the 80th-94th percentiles

Final

The final version is a different take on the problem. We manipulated the data to show how much an individual in each pentile and the top 5% would earn for every $10,000 of income generated. This fixed a few problems. First it normalized the data so Asian earners didn't look under represented because we were missing historical data. Second it really drives home the stark difference between top earners and the rest of the country. We were able to include this in our income story.

To visualize wealth shares we wanted to easily show the proportion. While angle isn't the greatest visual encoding, it seemed appropriate for this case as the biggest comparison we wanted to show was between the groups in the top 10 and the bottom 90. Below is our first version of this chart.
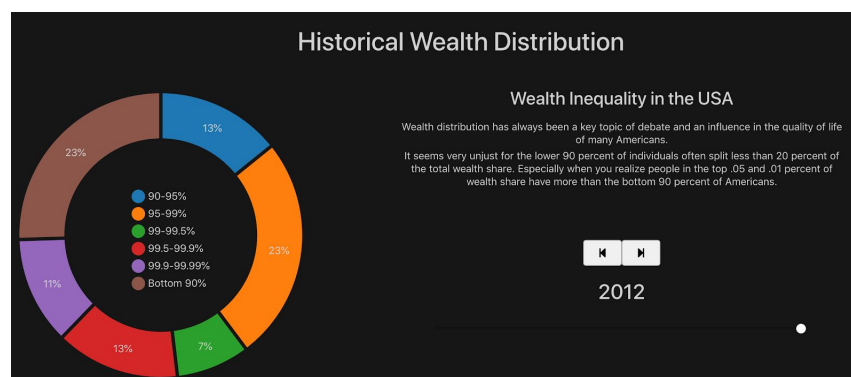
Notice how difficult it is to tell the bottom 90 from the top 10. We thought it would be better if we changed how we used our color encoding for category. We decided to use one hue for all groups in the 10% and another for the bottom 90. The groups in the top 10 are then distinguished by value. Our final version is below.

Historical Wealth Distribution

Wealth Inequality in the USA

Wealth distribution has always been a key topic of debate and an influence in the quality of life of many Americans.

It seems very unjust for the lower 90 percent of individuals often split less than 20 percent of the total wealth share. Especially when you realize people in the top .05 and .01 percent of wealth share have more than the bottom 90 percent of Americans.

1989

- Bottom 90%
- 90-95%
- 95-99%
- 99-99.5%
- 99.5-99.9%
- 99.9-99.99%

With this new color encoding it becomes much easier to see what we want to point out (that the bottom 90% has much less than the top 10%).

For our story we wanted to take the martini glass approach. We have two stories to tell, one about income, the other about wealth. In each case users can step through the story and then explore the data themselves.

Our final design definitely deviated from our proposal. Originally we had planned to use a cargogram to show income inequality by state. However our final project remains true to the ideals that we had set when we first started. We wanted to highlight income inequality in America.

Implementation

For our main view we wanted to make selecting categories easy. There are 25 different possible data sets to select. We implemented a way to quickly select demographics. These checkboxes drive both the Income Year Chart and the Income Shares Chart. Because both charts had handlers for the change it actually made implementing our story logic fairly easy. More on that later.

Each view was implemented in it's own class (We have an IncomeYearChart class, AggregateIncome class and a Wealth class) to drive its own display. We then have an init.js file that loads all of the data, does any mangling and then instantiates each view and sets any necessary callbacks. This way each view is basically encapsulated by itself and we just have one global driver function.

For the Income chart, in addition to filtering with the check boxes there are several interactive feeatures

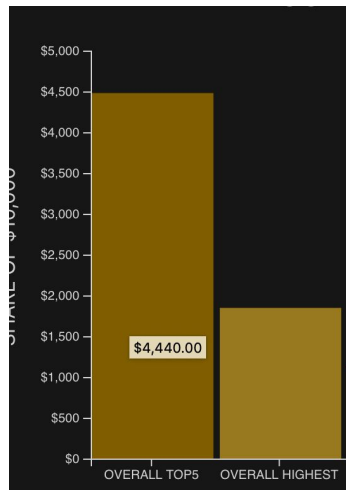- Hover highlighting for both lines and legend
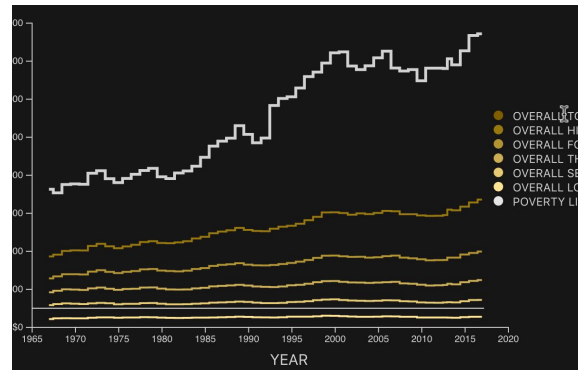
- Hover tooltip on lines
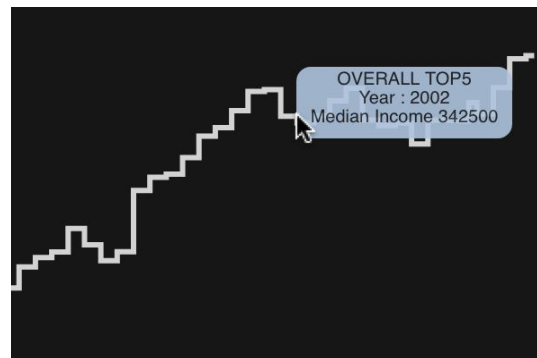
- Brush for income gap.

The Income Shares Chart has both a time slider and tool tip.

The wealth share chart has a timeline to allow scrubbing and seeing how wealth distribution has changed over time.

The stories were implemented using a Story class that is reusable (we use different instances of the class for each story). The class takes selectors for elements to modify as well as a side effects function that makes necessary changes on the page. Because of this design we were able to add a second story with only a few additional lines of code. Below is an example of what a story card looks like.

One challenge for the Income Shares Chart was getting the binned axis labels to behave well. We were able to set a condition to reformat the labels themselves if the number of bins were too high, and to abbreviate them if they were even higher. This way we were able to display every category without getting overlap in our labels.

The brush implementation on the income chart was pretty tricky. We had to find a way to calculate the area between two brushed curves. Ultimately we decided to only use the highest and lowest curves caught by the left end of the brush. From there we could loop through the data and calculate the wealth gap.

Evaluation

Both of us are well aware of income and wealth inequality in the United States. One of the reasons we chose this topic was because we both see it as an issue. Neither of us knew the extent of the problem though. Both of us were extremely surprised by the racial disparity in income. In fact we made it one of the key parts of our story. One thing that stood out as well was that a person in the top 5% will earn nearly 30x as much, that is staggering.

Questions 1-3 were answered with our first two visualizations (Income Chart and Income Share Chart). From the Income Chart it's very easy to see which groups are making the most economic gains and the disparity between economic and racial groups. The brushing on the chart as well as the Income Share chart show just how big the gap is. Question 4 was answered with our wealth chart.

Our visualization works very well. I feel it presents the data in a fair and accurate way. It also paints a picture of the problems we face in America. If we were to improve it we might add transitions to the Income charts. It would also be nice to allow for re-ordering of the bars in the Income Shares chart.