Wealth Disparity in the United States

Group members:

Gaurav Keswani (gk368)
Rui Liang (rl522)
Taihang Ye (ty322)

1. Data Source

The data set of wealth distribution of U.S. from 2000~2011 was found at http://www.census.gov/people/wealth/

We retrieved data sets for multiple years between 2000 and 2011, but went with 3 i.e. 2000, 2005 and 2011. The reason for doing so was to compare the status of the society at the beginning of the millennium to the time before and after the great recession of 2008.

2. Data description & data binding

2.2 Mean Net Worth Per Quintile

The mean net worth per quintile is a comparison between the mean net worths of the different quintiles across 3 years - 2000, 2005 and 2011. Since the difference between the maximum and minimum mean net worth was huge, we decided to map the values to a scale of 0 to 1 where 1 represents the total mean net worth of all quintiles put together in 2011 (can be seen by the 2011 bar ending at 1.0). Also, the negative mean worth values have been mapped to 0 since we are not considering debt. The 5 quintiles are represented by 5 colors - dark green (5th quintiles i.e. richest) to red (1st quintile i.e. poorest)

As can be seen from the graph, the 5th quintile owned about 74% of the wealth in the United States at the beginning of the millennium. This number shrunk to 72% by 2005 and probably would have shrunk further if not for the recession in 2008. The recession in 2008 led to losses for everyone, but it is clear that the poor were affected in the worst possible way while the 5th quintile's share of the mean net worth grew to 79.5%

The data for this visualization was taken from the 2000, 2005 and 2011 data files from the "Monthly Household Income" section of the "Mean Net Worth" sheet. In the data set, each of the net worth quintile was itself divided into 5 quintiles based on the monthly incomes. But we manually calculated the overall mean worth for each quintile of net worth by using the individual means values and number of households data given in the sheet. These values were then manually

mapped to "quintile_data.csv"

2.2 Age-Quintile-Wealth distribution in different years

Since we saw such a huge gap between the 5th quintiles and the other quintiles, we decided to add the age variable to the equation to see if there were particular age groups that were affected.

The heat maps for "Age-Quintile Wealth distribution" are in 3 parts - one for each year i.e. 2000, 2005 and 2011. In each of these heat maps we have age ("<35", "35-44", "45-54", "55-64", ">65") and quintiles (1, 2, 3, 4, 5) where quintile 5 is the richest top 20% group. The purpose of these heat maps is to analyze the distribution of mean net worth.

The data for this is coming from the "Distribution 2000.xlsx", "Distribution 2005.xlsx" and "Distribution 2011.xlsx" files. We selected the age groups sections in "Distribution of Net Worth By Net Worth Quintiles" table and manually converted into "age_quintile_wealth.csv".

For "Age-Quintile-Wealth distribution" chart, we divide mean value into three ranges, which are "wealth_min to zero", "zero to two third of wealth range", and "two third of wealth range to wealth_max". Then we use colorScale to linearly map mean value to different colors, red for the poorest and green for the richest. We use rectangles for the heat map, and label the age groups and years at the bottom and the quintiles on the left side of chart.

The purpose of this chart is to show that the more and more wealth is being accumulated by the 5th quintile over the years while the 1st quintile is in a worse state.

2.3 Mean and Median Net Worth By Region

In the "Median and Mean net worth comparison by regions" chart, we concentrate on the data from 2011. We use the region section from the data set and include both mean and median values for each quintile, (manually converted into into "region.csv"), which has four variables: region ("North east", "Mid west", "South", "West"), quintile (2, 3, 4, 5), median and mean.

In this region bar chart, we use the linear mapping to map the both mean and median wealth values to the length of the bars and differentiate between the different quintiles using different colors, like red for quintile 2(60~80%), orange for quintile 3(40~60%), light green for quintile 2(20~40%), dark green for quintile 1(0~20%).

2.4 Mean Net Worth By Race

In the "Wealth Distribution within Racial Groups" chart, we divided each race group equally into 5 quintiles, from the richest to the poorest. Each quintile contains 20% of the population of that race.

The data has been edited in following way:

i.   Add up the numbers from 5 quintiles to calculate the total mean net worth of the four racial groups: White, Asian, Black, and Hispanic;

ii.   Calculate the percentage of mean net worth of each quintile against the total net worth of that race;

iii.   Because the positive net worth and negative net worth would be reflected on two different dollar signs, we then take the percentage of mean net worth of positive quintile and divide that by the sum of percentage of positive net worth. This would give us a new percentage to display on graph. (Detailed calculation can be found on "racial-data.csv")

The size of dollar sign represents the average net worth owned by each racial group, and the span of color represents the percentage of mean net worth owned by each quintiles within that race. The dollar sign under x-axis indicates a negative number in net worth. As shown, the majority of the wealth is owned by the richest 20%, while the poorest 20% has a negative net worth.

3. Story

Wealth disparity is a reality in our society and there has always been a gap between the rich and the poor. But when it comes to the United States, we always heard that the disparity is huge. Our aim for this project was to dive deep into the data and figure out whether these claims were actually true. This topic is even more relevant today keeping in mind that the general elections are coming up and this could decide how the next 5 years go.

The first thing we looked at was a comparison of the mean net worth across 3 years - 2000, 2005 and 2011. The reason for choosing these 3 years was to get a view on the state of the people at the beginning of the millennium and a few years before and after the effect of the recession of 2008 had settled.

What we observed was that there was a marked difference in the status of the two groups with the top quintile holding about 74% of the total wealth. That being said, what we saw is that the lower classes had started to make up some ground on the higher ones by 2005 but the recession stopped this upward trend. While it is true that everyone lost money because of the recession and this includes the higher classes as well, it is very clear that the poor were affected the most. This can very well be seen from the fact that in 2005, the 5th quintile was holding 73% of the wealth while in 2011 they were holding almost 80%.

Looking at this, we decided to analyze the data further and chose 3 different angles - age, region and race. With the age analysis, we could clearly see that the wealth was being accumulated by the 5th quintile while the situation of the people in the 1st quintile was worsening by the day. This visualization backed up the claim that the rich were becoming richer and the poor were becoming poorer.

By now it was clear to us that the majority of the wealth was with the top 20% but we wanted to see how evenly it was distributed in the top 20%. For this reason, we did the mean vs median analysis at a region level and saw that in the west, the mean net worth was 3.5 times the median income. What this means is that even in the highest quintile, it is the only the top few who have accumulated most of the wealth. This fact could not have been uncovered from an individual mean analysis or median analysis but the combined analysis made it very apparent.

The last analysis was the mean net worth analysis at the race level where we were aiming to see if this trend was prevalent across all races. Our analysis showed that even though the wealth was concentrated in the highest quintile for most races, the lower classes of the Black and Hispanic races were the worst affected while those of the White and Asian races were still doing a little better.

What our study has shown is that the gap between the rich and the poor is indeed huge and that steps need to be taken to bridge it.