Rita Kalo
Jaret King
Brandon Masiello
Isabella Plummer

Who Contributed to the Bernie Sanders Campaign?

Main Goal

The purpose of this research was to determine correlations between donations made to presidential candidate Bernie Sanders and socioeconomic measures predetermined. Within this research, a subset was aimed to learn how to properly analyze and display the results of research in a GIS format.

Background

Bernie Sanders, a senator from Vermont, is known for his progressive opinions on social equality, education and tax reform, and federal labor reform. During the presidential campaign, Bernie Sanders was adamant about raising the minimum wage, a belief that affects the occupations and incomes of citizens. Furthermore, a main component of the Sanders campaign was the desire to make all public colleges and universities tuition free, capturing the interest of many citizens who have attended or are currently attending college. This belief helped determine the second and third measure of research within this project: education and age.

Further into research into the belief system of the Sanders campaign, another belief of Bernie was aimed to expand social security, validating another factor of research within this project: age. Finally, Bernie fought to decrease the wealth gap throughout the entire United States, gaining the support of lower income individuals or "blue collar" occupations. This belief determined the final measure of correlation within this research: occupation. Among all the beliefs stated by the Sanders campaign, the final measures that were believed to show the strongest correlation to the number of donations were occupation, education, age, and income

Study Area

The two states that were chosen for comparative analysis were New York and Nebraska. These two states differ greatly in many significant ways. Politically, New York has been Democratic for almost three decades [1], whereas Nebraska has only voted for a Democrat seven times in the last 149 years [2]. New York has a population almost ten times greater than Nebraska: 19,745,289 and 1,907,116, respectively [3]. From further research with donation data from the Federal Elections Committee, Nebraska donated far less per person to the Bernie Sander's campaign when compared to New York. Also, New York was found to have a higher percentage of citizens in the Millennial generation and a higher percentage of citizens with bachelor's degrees, relative to that of Nebraska.

<u>Objectives & Criteria</u>

The objective of this project was to analyze—by zip code—the people who donated to Bernie Sanders' presidential campaign in 2016. Unlike other campaigns, Bernie was successfully funded almost entirely by "regular people". In this project, analysis was performed to determine what commonalities existed between those who funded his campaign. As previously stated, the Bernie Sanders campaign was run on issues that directly affected very specific sets of people in the United States. This project aimed to analyze the socioeconomic factors of the "regular people" in order to gain accurate representation. To do this, our project focused on four different criteria: occupation, education, income, and age.

For occupation, the analysis was split into two fields: blue collar and white collar. Data collected from *AmericanFactFinder.com* was divided into several fields, which were then combined for ease of analysis. Blue collar workers were defined as citizens with occupations in natural resources, construction, maintenance, production, transportation, and material industries. White collar workers were defined as citizens with occupations in management, business, science, arts, service, sales, and office industries.

For education, the analysis was also split into two fields: the percentage of people with high school diplomas and the percentage of people with at least a bachelor's degree. This data was also collected from *AmericanFactFinder.com* and required no manipulation for analysis.

For income, analysis was divided into lower, middle, and upper income classes. The analysis was made using the same income breaks for both Nebraska and New York. Lower class was defined as households earning $35,000 a year or less; middle class was defined as households earning between $35,000 - $100,000 a year. Upper class was defined as a household earning over $100,000 a year.

For age, analysis was divided into generational ages defined by Dr. Jill Novak at the University of Phoenix [4]. For simplification purposes, donors who fit into the generation "Baby Boomers" and older were simplified as "Baby Boomers+" in an attempt to visually represent differences between older and younger generations and to simplify the data set. The generations that were analyzed for this research are defined as: Millennial (age 18 - 34), Generation X (age 35 - 50), and Baby Boomers+ (age 50 and above).

Methodology

The general process by which analysis was run for research can be summarized in the flowchart below. After step one, each step was performed four times — once for each measure.
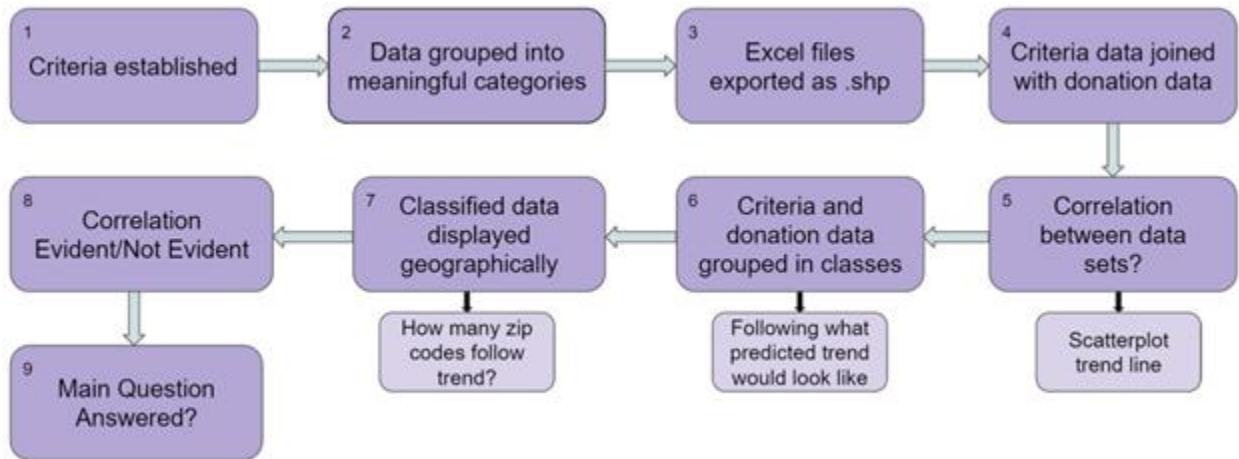

To expand the flowchart, the first step in our methodology was deciding what criteria would yield meaningful results. As previously explained, the factors that were chosen were age, occupation, education and income as measures of correlation. The next step was to group the measure data into fewer categories. The data that was retrieved was broken up into too many subcategories for analysis, so each measure was merged into fewer categories. This merging of fields was performed in Microsoft Excel. For the next stage, the Excel file was converted into a shapefile using ArcMap. The data in Excel was then joined area shapefiles for visual projection. The next phase of analysis involved determining if a correlation existed between any of the measures and the donation per capita data. Before analysis was performed, an assessment was performed into the natural breaks in the data. Each measure was broken into even quartiles of percentages. For dividing the donation data, the distribution of the average donations per capita for each state was analyzed and divided according to a high, semi-high, semi-low, and low average donation amount. The divisions that were chose are displayed in the legends below.
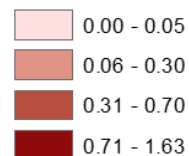
Legend

NY Donation Data
Average Donation Per Capita

	0.00 - 0.30
	0.31 - 2.00
	2.01 - 8.00
	8.01 - 20.56
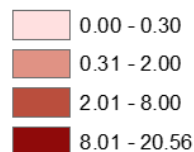
Legend

NE Donation Data
Average Donation Per Capita

|	
	0.00 - 0.05
	0.06 - 0.30
	0.31 - 0.70
	0.71 - 1.63

Scatterplots were created from the attribute table that displayed the donation data against each measure individually. A trendline was added to each scatter plot that statistically revealed any correlation (negative, positive or none) among the data. From the trends that were developed, zip codes that followed the trends were selected and displayed. For example, the
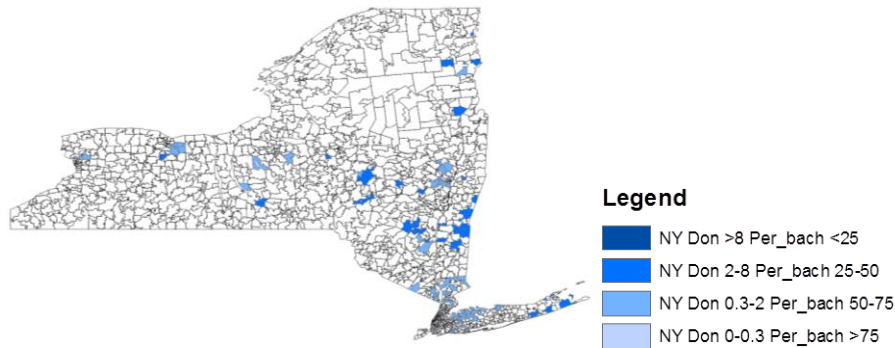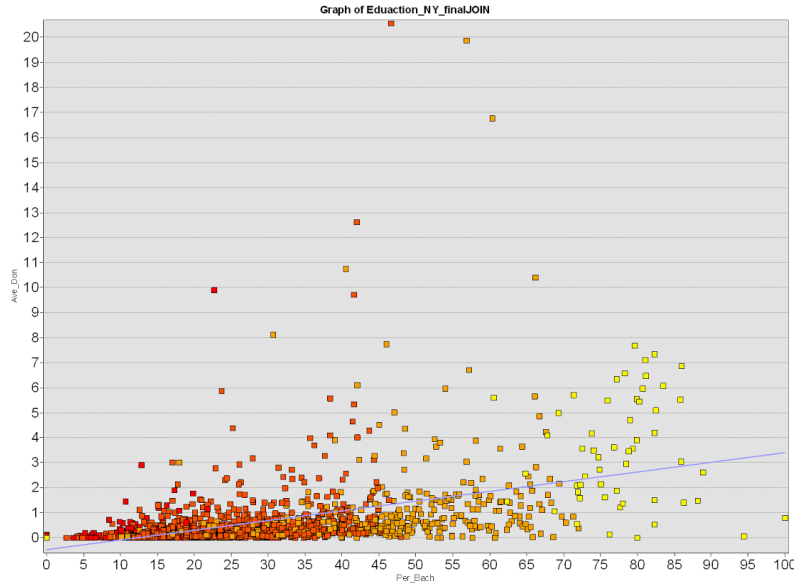
percent of people with high school diplomas was shown to have a negative correlation with the average donation amount per capita. So, if the trend was very strong, the zip codes in the lowest quartile of percentage of people with high school diplomas should also be the zip codes with the highest average donation amount per capita, and so on. By creating these "trending" zip codes, visual representation of the strength of a trend could be seen. After both the correlation of data and the strength of that correlation was determined, the initial hypothesis of each measure could be assessed and a conclusion could be drawn.

Education Results

One hypothesis for the education measure was that zip codes with higher percentages of people with high school diplomas would donate more per capita. Both states yielded interesting results for education as a measure of how likely a person was to donate. Below are the results for the percentage of people in New York that had high school diplomas. The results for Nebraska were similar—in both cases, as the percentage of high school diplomas decreased, the donations per capita increased. These results refute the original hypothesis, but can be considered fairly significant according to the scatterplot and the geographical representation of the zip codes that followed the trend (Green shaded map; 31% of zip codes considered to strongly follow trend).

Legend
▉	NY Don 2-8 Per_high 25-50
▉	NY Don 0.3-2 Per_high 50-75
▢	NY Don 0-0.3 Per_high >75
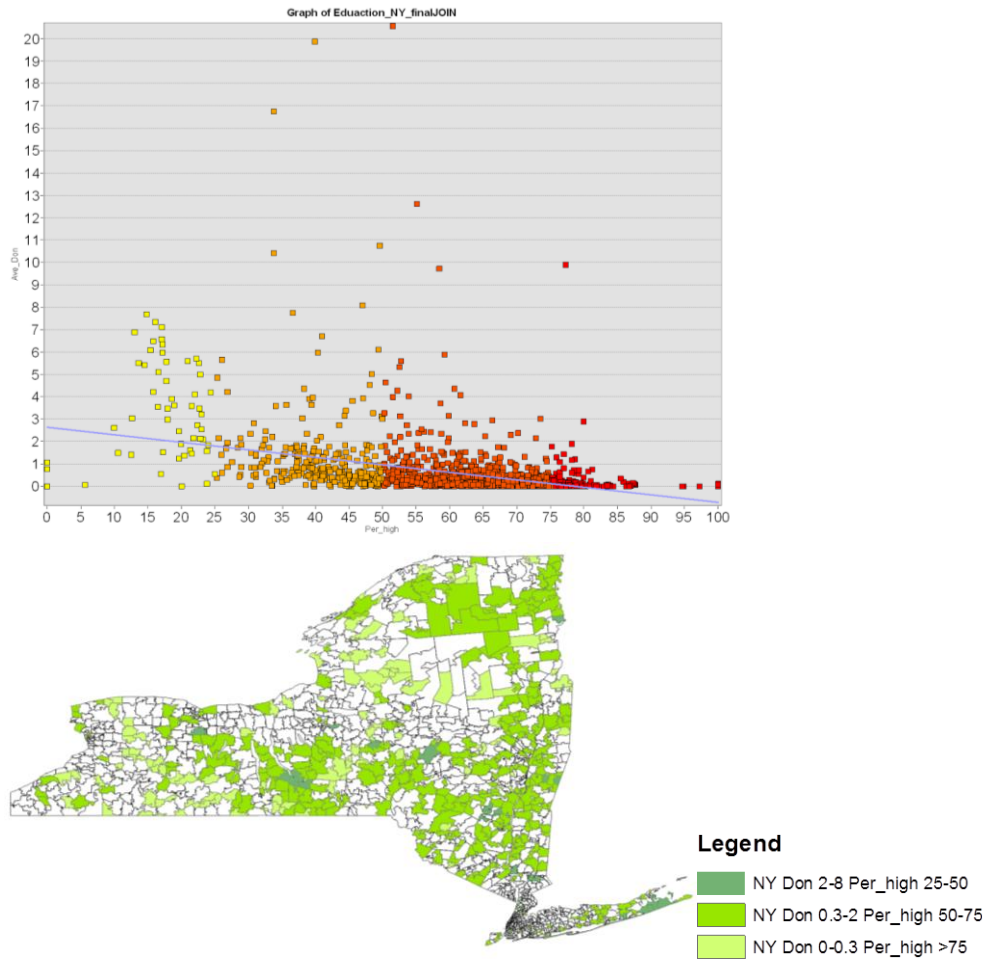
Another hypothesis for the education measure was that zip codes with higher percentages of people with Bachelor's degrees would donate more per capita. Below are the results of this analysis. Again, the results for Nebraska were similar. Both states showed that zip codes with a higher percentage of people with Bachelor's degrees donated more per capita. The results for this analysis confirmed the original hypothesis, however while the relationship was apparent in the scatterplot, the geographical representation of the trending zip codes (Blue shaded map) showed that only 9% of zip codes strongly followed the trend.


Graph of Eduaction_NY_finalJOIN

Legend

▮	NY Don >8 Per_bach <25
▮	NY Don 2-8 Per_bach 25-50
▮	NY Don 0.3-2 Per_bach 50-75
▮	NY Don 0-0.3 Per_bach >75

Income Results

To aid in simplifying the comparison of New York and Nebraska, the income definitions for lower, middle and upper income classes were kept the same. Because New York's cost of living is much higher than Nebraska's, someone who is considered middle income by the standards of this project in Nebraska might actually be in the same position as a low-income person in New York.

The hypothesis for the income measure was that zip codes with a high percentage of lower and middle income households would tend to donate the most per capita. As seen in the scatter plots for Nebraska below, there is a trend line indicating that average donations per capita appears to decrease as the percentage of lower and middle income households increases.

However, the map of Nebraska below shows that this trend is not strongly followed by most zip codes. For lower incomes (shown below), and middle incomes (not shown) very few zip codes proved to fit the criteria for a strong correlation. The actual data shows that the most donations per capita tended to occur in zip codes that had no majority of low or middle income households. In fact, for lower and middle income households the percentage of zip codes strongly following the trend of decreasing donations per capita was only 4% and 12.5%, respectively. For upper income households, the trend shows that there is an increase in donations per capita as the percentage of upper income households increases. Since this trend was followed in Nebraska by 72% of the zip codes, it initially appears as if this is accurate. However, considering that the majority of zip codes in Nebraska have less than 25% of its households that are upper income, there were very few zip codes involved in the analysis. In fact, there are 87 zip codes that have over 25% of upper income households, and only 19 of those zip codes followed the trend. This may have caused the results to be skewed. Because of the factors that affected the accuracy of the upper income data, the results do not have a strong connection to the trend, and are thus inconclusive.

Nebraska Results
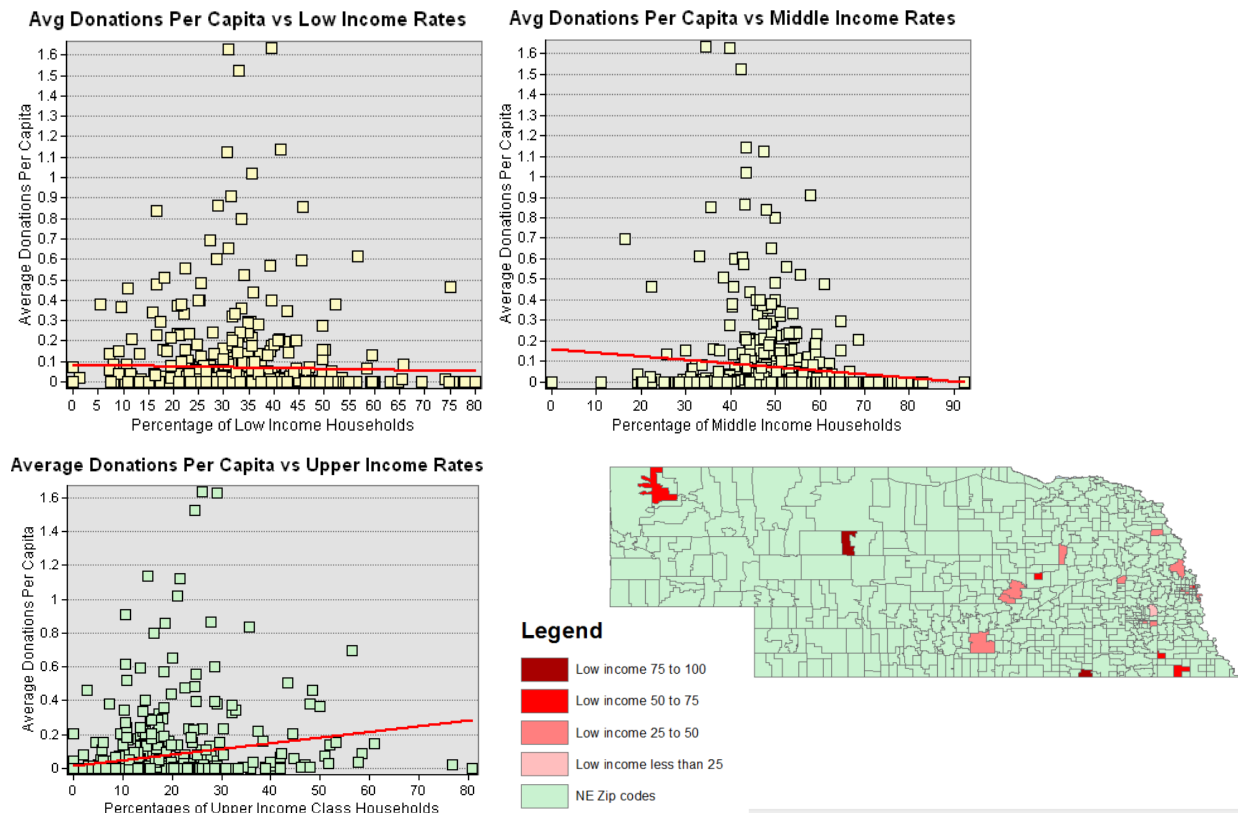
As seen in the graphs below, the same trends in Nebraska appear in New York, as well. When the income and donation criteria are combined, the trend of donations per capita increasing as rates of lower and middle income households increased is not followed much. Only

4% and 12% of zip codes, respectively, strongly follow the trend. However, if the overall spread of the data in the scatter plots is considered instead, it is obvious that there is a notable decrease in donations per capita as lower and middle income households increase.

The trend of upper income household percentages increasing as average donations per capita increased was followed much more so than lower and middle income trends. The trend for upper income was strongly followed by 58.5% of zip codes in New York, and the spread of data overall shows that the trend was indeed followed. As seen by the graph below for New York upper income rates, this trend is more accurate than Nebraska's due to a more even distribution of upper income rates between 0 and 75%.

New York Results
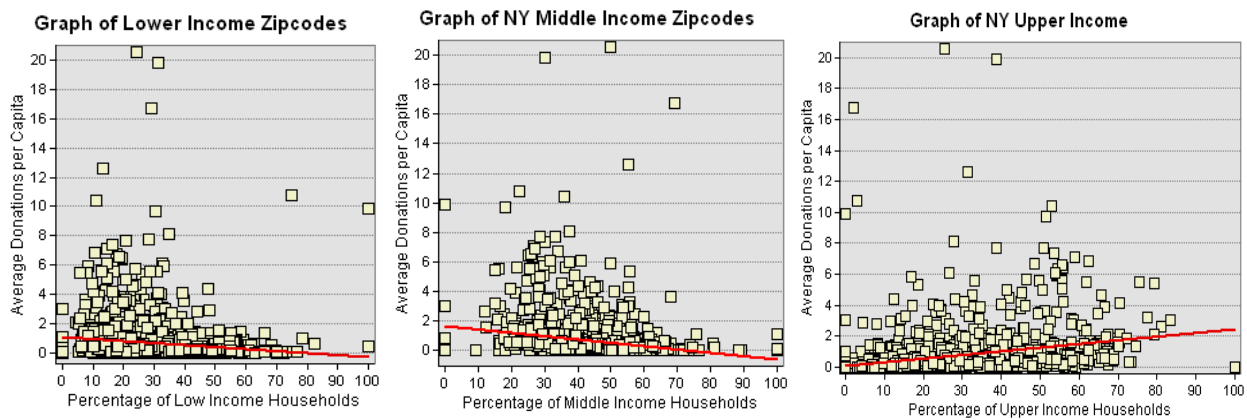
Graph of Lower Income Zipcodes | Graph of NY Middle Income Zipcodes | Graph of NY Upper Income
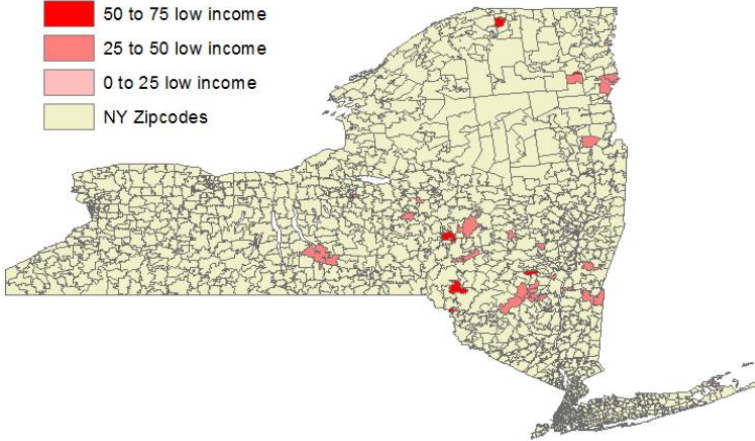
Occupation Results

The hypothesis that was formulated for the occupation measure was that zip codes with a greater percentage of blue collar workers would donate less. Although Bernie Sanders may have appealed more to blue collar workers, white collar workers are generally more politically-involved, and so it was assumed that they would donate more. As the scatter plots below shows, there was a slight trend in the data for both states. It appears that as the percentage of blue collar

workers increases, the average amount donated per capita decreased. This observation agrees with the hypothesis that was formed, however when the maps of the zip codes that were shown to strongly follow the trend are considered, the significance of the results decreases. In fact, the percent of zip codes that strongly followed the trend in Nebraska was 4% and in New York it was only 2%. Ultimately, there was a slight correlation between the type of occupation and the amount donated per capita, but it was not a significant correlation.
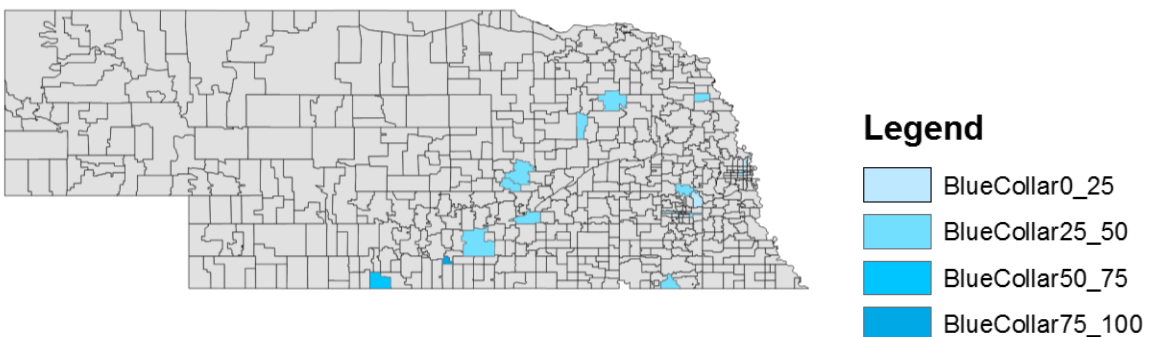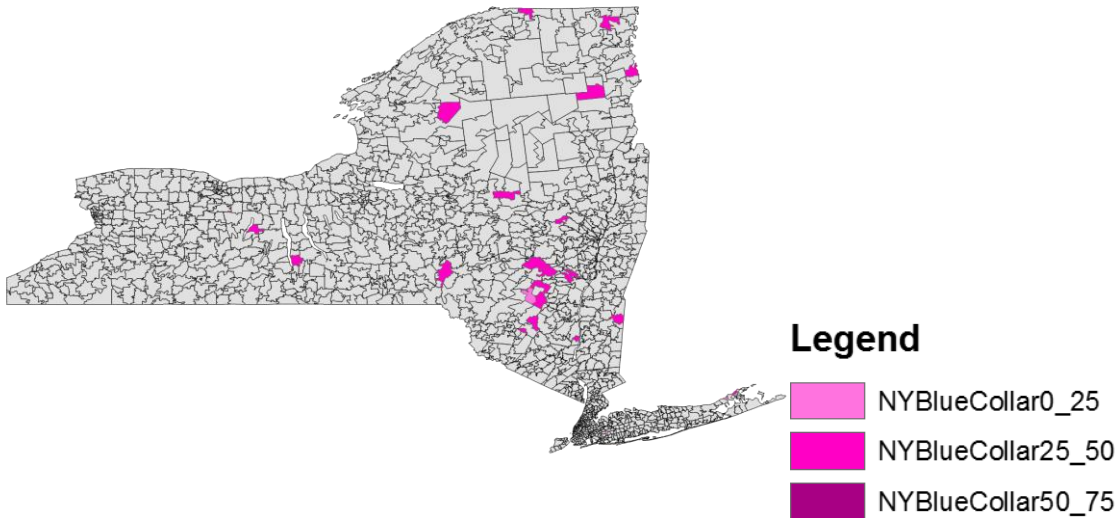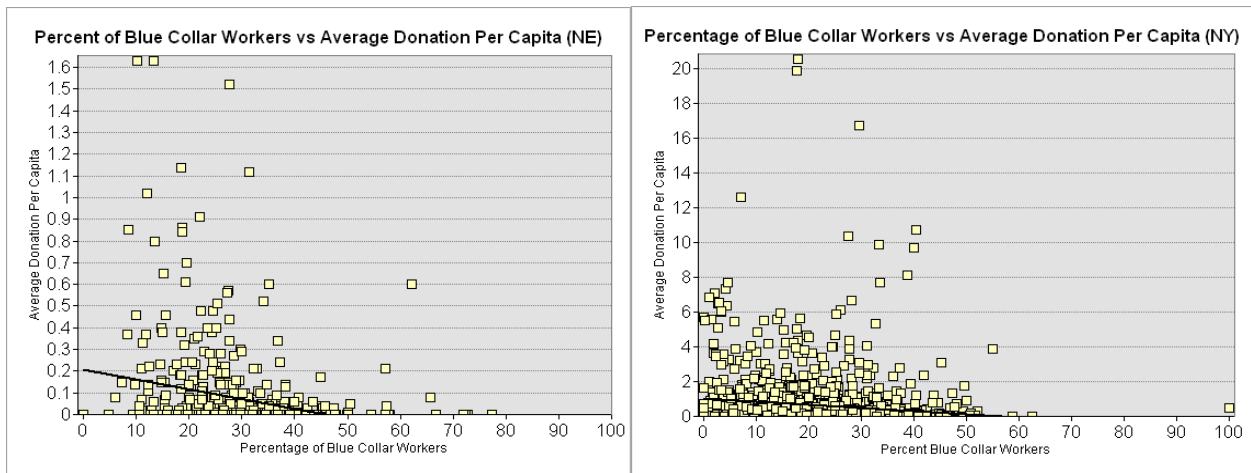
Because of the political stances that Senator Bernie Sanders has regarding free college funding and human rights, a hypothesis was developed stating that a zip code with a higher percentage of the population in the Millennial generation would be more likely to donate to the campaign. The opposing hypothesis for older generations or "Baby Boomers+" was developed as well. For the state of Nebraska, an inverse trend among all three generational ages developed in the data. In other words, as the percentage of Millennials increased in a zip code, the average donation per capita decreased—with the same correlation occurring among the other two generations. The figure below shows a sample trend for the "Generation X" in Nebraska.
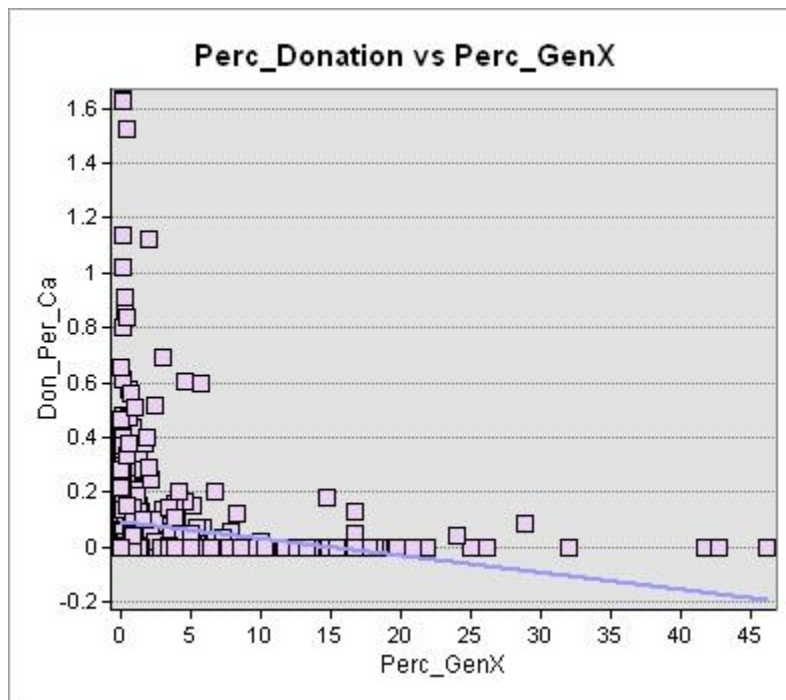
An analysis within the state of New York was also performed, but no trend was developed between the generation data and the average donation per person for all three generations. However, when the data was numerically analyzed, an increase in donations per capita occurred when the zip code had an equal distribution of generations in the population.

Although the correlations between each state's generational percentage and average donation per capita have been stated, the strength of those correlations proved to be rather weak. When selecting the counties that had a strong correlation to the stated trend, the number of zip codes that were selected were extremely insignificant. The map below of trending counties in Nebraska for Millennials gives an example of how few zip codes were selected for all indicators. (The rest of the maps for all measures can be found in the presentation.)

Legend

NEPercMillenials_75+

NEPercMillenials_0_25
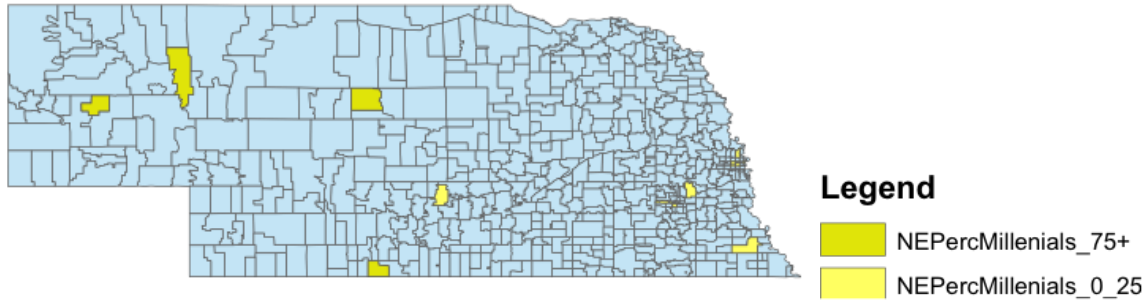
General Discussion

One limitation that was discovered after performing all the analyses was the criteria for the "trending" zip codes. Because the definitions used for the donation data were so narrow, it is highly possible that the percentage of zip codes that followed all the trends shown in the scatter plots should have been higher. For example, the highest level of donation data for New York was 8-20.56 donations per capita. All three scatter plots for the income measure show that the majority of the zip codes had 8 or less average donations per capita. This means that many zip codes that followed the general trends were likely not included due to our selection criteria. Instead of creating standardized criteria for donation data, it may have been more effective to create the donation classes on a case-by-case basis. That way, more zip codes that were following the trend could have been included in all the maps. In creating a standard donation criteria, it was effective at isolating zip codes that strongly followed a standard trend, but not all trends showed the same degree of strength. It may have been better for each measure to have its own trend standard in order to capture more of the zip codes that followed the trend less strongly.

Conclusions

The hypothesis for high school education was that for zip codes that had higher percentages of individuals with high school diplomas, there would be more donations per capita. The results of our research for both states disproved this relationship, and showed that zip codes with lower percentages of high school degrees were more likely to donate more per person. For the analysis on Bachelor's degrees, the hypothesis was that zip codes with a higher percentage of Bachelor's degrees would donate more per capita. The findings we gathered for both states supported this hypothesis, although the geospatial representation of the zip codes that strongly followed the trend did not indicate this was a strong trend. Overall, the results indicate that both zip codes with the lowest percentage of high school diplomas and zip codes with the highest percentage of Bachelor's degrees tended to donate the most per capita.

For income, the initial hypothesis was that more donations per capita would be found in zip codes that had a majority of lower and middle income class households. However, our data did not reflect this. In Nebraska, zip codes that had between 25-50% of low income households tended to donate the most. Zip codes that had between 40-60% of middle income households also donated the most. Zip codes that had between 10-25% of upper income households had the highest donations, with the amount donated dropping off significantly after any more of an

increase in upper income households per zip code. In New York, a trend that was opposite of the hypothesis was much more noticeable. As the percentage of low income households increased in a zip code, the data and the trend line showed a decrease in average donations per capita. Similar to the results in Nebraska, the average donations per capita tended to be highest when the percentage of middle income households per zip code ranged from 25-50%. There was a surprising trend showing that as upper income households increased, the average donations per capita also increased. The data and trend line showed this to be true until an area had over 75% of upper income households. Once the 75% threshold was reached, the data became less conclusive.

Overall, in Nebraska it appears that areas that had no real majority of lower and middle income households tended to donate the most per capita. Upper income households are purposely neglected in this conclusion. This is because although areas that had 10-25% of upper income households donated the most per capita as well, the vast majority of zip codes in Nebraska only have 0-25% of households that are upper income, rendering this aspect of the data inconclusive. In New York, donations per capita were the most where lower income households were at a minimum, upper income households were at a maximum, and middle income households were between 25-50%.

For occupation, both states indicated that as the percent of blue collar workers in a zip code increased, the average donation per capita decreased, which agrees with the occupation hypothesis. And since the measure of occupation is effectively binary, it can also be stated that as the percentage of white collar workers in a zip code increased, the average donation per capita also increased. Although the trend was apparent in the scatter plot, only a small percentage of zip codes were found to strongly agree with the correlation in both states. As discussed previously, the limitations on the donation criteria could have been the reason why so few zip codes were shown to follow the trend, however, it is better to conclude that these results were not significant than to falsely claim that they are. So, it can be concluded that there was a correlation, but not to a significant degree.

Finally, for age, the trends for both Nebraska and New York had weak strengths for all trends that were stated. In Nebraska, all three indicators showed a negative trend, but few zip codes proved to strongly follow the trend. In New York, there was basically no trend at all for any indicators and very few zip codes were selected for evidence of a strong trend. Therefore, no definitive correlation between the two measures in this section could be determined. It can be concluded that for the age measure, the most donations per capita came from zip codes that had no majority of Millennials, Generation X, and Baby Boomers+, which did not agree with the age hypotheses.

Sources

[1] "New York Presidential Election Voting History." *270towin.com*, n.d. Web. 18 Apr. 2017.
<http://www.270towin.com/states/New_York>.

[2] "Nebraska Presidential Election Voting History." *270towin.com,* n.d. Web. 18 Apr. 2017
<http://www.270towin.com/states/Nebraska>.

[3] United States Census Bureau. "QuickFacts."

[4] Novak, Jill, Dr. "The Six Living Generations In America." *Marketing Teacher*. N.p., 05 Mar.
2017.Web.18.Apr.2017.<http://www.marketingteacher.com/the-six-living-generations-in-
america/>.