Project Goal

Our project is analyzing the impact of the Affordable Care Act on low income Americans in 5 New England states (Rhode Island, Massachusetts, Connecticut, New Hampshire, and Vermont). We are looking at the prevalence of preventative care before and after the ACA was implemented, using the 2011-2014 period as our before baseline and 2015-2018 as after. The metrics we are focusing on relate to checkup frequency, insurance, and cost barriers to seeking help. All of the states within our data set adopted the medicaid expansion and had no other major changes to healthcare within this time period.

Hypothesis

H_0: In the New England States, the Affordable Care Act did not change low-income residents' access to healthcare according to our metrics for success.

H_A: In New England states, the Affordable Care Act increased low-income residents' access to healthcare using the following metrics for success:

The first metric is **health coverage**.

H_0: There is no difference in the proportion of low income residents who have access to health care in 2011-2014 compared to 2015-2018.

H_A: There is an increase in the proportion of low income residents who have access to health care in 2015-2018 (post-ACA) compared to the residents surveyed in 2011-2014.

The ACA institutionalized subsidies and premium tax credits to help low income individuals purchase marketplace coverage. It also expanded the federal and state sponsored Medicaid program. We predict that the proportion of low income residents who have health insurance is significantly higher in the 2015-2018 period following the ACA than in 2011-2014. Similarly, the proportion of people who have at least one personal doctor or health care provider is higher in 2015-2018.

The second metric is the **cost barrier to seeking medical help.**

H_0: There is no difference in the proportion of low income residents who could not see a doctor because of cost in 2011-2014 compared to 2015-2018.

H_A: There is an decrease in the proportion of low income residents who could not see a doctor because of cost in 2015-2018 (post-ACA) compared to the residents surveyed in 2011-2014.

We will analyze the amount of times in which low-income residents did not see a doctor or take prescribed medication because of cost. We predict these instances should be significantly lower in the time period after ACA measures have been implemented.

The third metric is **checkup frequency.**
H_0: There is no difference in the length of time between the last checkup for low-income residents in 2011-2014 compared to 2015-2018.
H_A: There is an decrease in the length of time between the last checkup for low-income residents in 2015-2018 (post-ACA) compared to the residents surveyed in 2011-2014.

Part of the ruling under the ACA was that private health plans must provide coverage for a range of preventative services and may not impose cost-sharing (including deductibles or copayments) related to these services (Source 1). We predict, based on this ruling as well as the increase in accessible public insurance options, that the average length of time between the last checkup is significantly lower after the ACA.

The last metric we will check is **overall health.**
H_0: There is no difference in the general health or the number of days participants could not do usual activities because of poor health of low income residents in 2011-2014 compared to 2015-2018.
H_A: There is an increase in the general health of low income residents and a decrease in the number of days participants could not do usual activities because of poor health in 2015-2018 (post-ACA) compared to the residents surveyed in 2011-2014.

To analyze this we will focus on two questions in the survey. The first question, genhlth, asks the general health of the participants (with the options of ["Excellent", "Very Good", "Good", "Fair", or "Poor"]. The next question, poorhlth, asked the participants how many days did poor physical or mental keep you from doing your usual activities. We predict the number of days of poor health is significantly lower after the ACA, and the overall health improves.

Testing

Python Notebook with all Testing: https://github.com/theklam/just_scraping_by

Variable	Meaning
medcost	Was there a time in the past 12 months when you needed to see a doctor but could not because of cost?
medscost	Was there a time in the past 12 months when you did not take your medication as prescribed because of cost?
hlthpln1	Do you have any kind of health care coverage?
persdoc2	Do you have one person you think of as your personal doctor or health care provider?
checkup1	Length of time between your last checkup?
poorhlth	How many days did your poor physical or mental keep you from doing your usual activities?
genhlth	General health of the participants ("Excellent", "Very Good", "Good", "Fair", or "Poor")

The variables from the BRFSS that we tested for with meanings

For the majority of variables we tested (medcost, medscost, hlthpln1, persdoc2, checkup1, genhlth), we used the chi-squared test for independence because almost all of the variables we were testing were categorical and we were only looking for differences between the ranges 2011-2014 and 2015-2018. Thus by testing for independence between "before" and "after" the Affordable Care Act we can detect if there is a significant difference between the two populations. We can detect there was a significant improvement by looking at the distribution graphs of the data from the two time periods as well as the p-values from the metrics we chose. We considered a two sample t test, however, we cannot be sure that the two samples are independent and we cannot do a paired t test because we cannot identify the identity of the participants in the study.

We measured success of a chi squared test of having a p-value less than .05 for the metrics chosen: number of people who have access to health care, less people having trouble seeing a doctor because of cost, checkup frequency, and improved general health for low income Americans. We felt these metrics measured the results that correspond best to the intent of the ACA and what it promised to achieve. The Affordable Care Act was designed to lower the cost of healthcare and provide Medicaid to more people and families that qualify. This, in turn, should improve access to health care, lower cost barriers, and broadly increase general health.

There were a couple of challenges that we discovered from transferring the data from our R preprocessing file to python dataframes. The first issue dealt with re-organizing the data from our updates since the last data deliverable. After more thought, we decided to extend the baseline period to 2011-2014 instead of just using 2014. Our original hesitation against including these extra years was that the surveys were somewhat different, both with different questions as well as different variable naming conventions. However, after reading more about the ACA we realized a larger time period would better

represent health statistics prior to ACA implementation. This change meant there was more data work on renaming and organizing variables to ensure we had a direct comparison between the two time periods. Another issue we dealt with was determining what sample of our survey population qualified as low income. The survey responses included a variable on income and we wanted to ensure we were only running tests on low income Americans to answer our original hypothesis. To do this, we decided to categorize low income in a similar way that many national funding programs use: less than 200% of the federal poverty level. This also is within the amount covered by ACA eligibility. The ACA lowered for households earning between 100% and 400% of the federal poverty level and expanded Medicaid for adults earning less than 138% of the federal poverty level. The federal poverty level used was as defined by the Department of Health and Human Services (Source 2). We took the household income from the survey question (income2) and the family size (with total number of children/dependents) and compared it with the low income chart we made by taking 200% of the federal poverty level. We then filtered out responses that were not within this range.

Results

	Chi-Squared Values	Degrees of Freedom	P-value	Significant? (Is p < 0.05?)
medcost	298.3474	1	7.5477e-67	Yes
medscost	2.6580	1	0.10303	No
hlthpln1	426.9376	1	7.54078e-95	Yes
persdoc2	84.8191	1	3.269563e-20	Yes
checkup1	95.25679	3	1.62604e-20	Yes
poorhlth	4.44602	4	0.348997	No
genhlth	52.82029	4	9.292436e-11	Yes

Table 1: Results of Chi-squared tests. Code for all tests are in the attached Jupyter notebook.

1. **Health Coverage.** For both questions (proportion of low-income residents who are insured, proportion of low-income residents who have a personal doctor or health care provider), the p-value was essentially zero. It is safe to conclude that there is a relationship between the health insurance and access to personal doctors and the time period (pre or post ACA). Looking at the health insurance in Figure 1, it is clear that the number of people who have insurance is higher in 2015-2018, thus the chi-squared shows that this change is significant. This also goes for the access to a personal doctor, as Figure 2 displays that the number of people who have access to at least one personal doctor increased post ACA, and so the chi squared signifies this difference.

2. **Cost barriers.** For the number of people who could not see a doctor because of cost, the p-value was once again close to zero, signifying a strong relationship. Again looking at the graph (Figure 3), one can see the difference is the lowering of people who could not see a doctor because of cost in 2015-2018. However, for medication costs, we did not find a significant difference. For the question whether they did not take medication as prescribed because of cost, the p-value was 0.103. This is not underneath our alpha threshold so we cannot make assumptions about the relationship. This test might not be fully representative of the sample size of the population. The variable for medication cost barrier (medscost) was not included in the survey for multiple years (2011, 2012, and 2015), and even when it was included there were very few responses relative to the other questions. For 2015-2018, there were 1077 responses out of 28691 low income Americans surveyed. However, the cost barrier to seeing a doctor was included every year with a very reasonable number of responses. Therefore, although we didn't find a strong significant relationship for cost barriers to medication, there is strong evidence of a relationship between cost barriers to seeing a doctor and the time range.

3. **Checkup frequency.** The length of time between the subject's last checkup is significantly lower after the ACA.The p-value from the chi-squared analysis was essentially zero, signifying that there is a relationship between time period and the time between check ups for patients. Looking at the distribution of individuals in the bar chart of this data in Figure 5, one can see that the number of respondents who saw a doctor less than 12 months ago is higher for the 2015-18 population. In turn, the number of people who saw a doctor over 1 year ago, 2 years ago, and over 5 years ago is lower in the 2015-18 respondents. Thus, the relationship that was proven significant was that those post enactment of the ACA, there were less people who had to wait over a year for a checkup.

4. **General health.** We found a statistically significant relationship for the genhlth variable. The p-value of the chi-squared test was 9.9292e-11, which is less than 0.05. This relationship can also be seen by the bar graph in Figure 7, in which the counts for Excellent and Very Good categories are higher and the Poor Health category is lower in 2015-2018 than 2011-2014. The chi-squared test in poorhlth was conducted by splitting the variable into 5 bins. This resulted in a p-value of 0.348997, so we fail to reject can reject the null hypothesis that there is no relationship of poorhlth between the two time frames. However, the test of genhlth showed a promising p-value so there might be a relationship of health with the policy changes from the ACA.

Discussion

Considering we have 5/7 chi squared tests with a p value lower than .05, we can reject the null hypothesis that the Affordable Care Act did not change low-income residents' access to healthcare according to our metrics for success. Thus, there is strong evidence to support our alternate hypothesis that the Affordable Care Act had a statistically significant beneficial effect to low income Americans in

the New England area. Intuitively this seems accurate as opening up opportunities for low income participants would lead to increased health care and in turn better general health for those populations. We are confident in the results, however, there are always confounding variables in tests over time, so it is hard to tell if everything is due to the enactment of the ACA. For example, incomes may have gone up within the time period we are testing which would affect how many people we filter within the income ranges. The income value is also dependent on the location. Not all of New England is homogenous, as some states or cities have higher purchasing power than others. Also we are not accounting for any rural/urban distinction which could affect access to healthcare and incomes. There was also about a 2% inflation rate every year in the past decade (Source 3) but we did not include that within our model.

There are also confounding variables likely within the questions posed. For example, we did not take into account any increases in medication costs over the past decade which could influence responses to the medscost question. In addition, there are differences in how people perceive their own health over time (as opposed to actual changes in their health). The difference between what they would consider to be "excellent" vs "very good" vs "good" is very subjective. Lastly, the survey was conducted every year over the phone but we suspect that phone line availability for low income Americans is likely greater in 2015-2018 than 2011-2014.

Visualizations

For our visualizations, we chose to use both bar graphs and colored tables. We picked bar graphs to display how our variables changed from the before (2011-2014) to after (2015-2018) the ACA was implemented, since bar charts make it easier to compare two different groups. Considering our bar charts are relatively easy to follow, they can stand alone. With the title and the axis labels, one can tell the realtionship that we are trying to display.
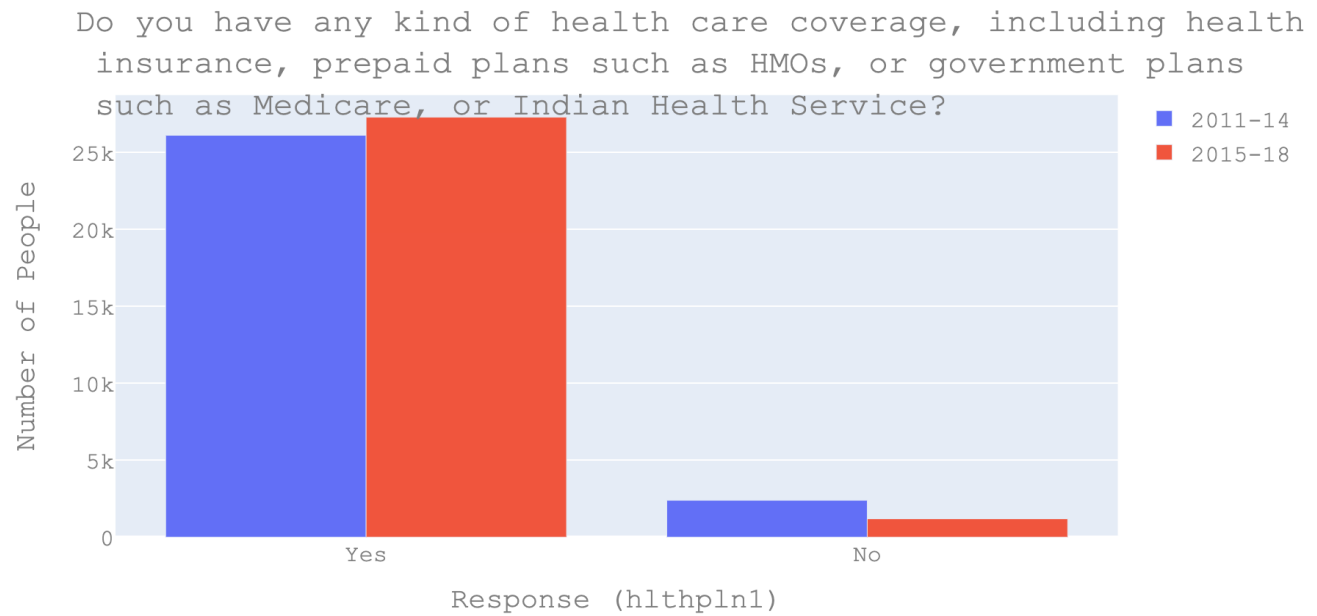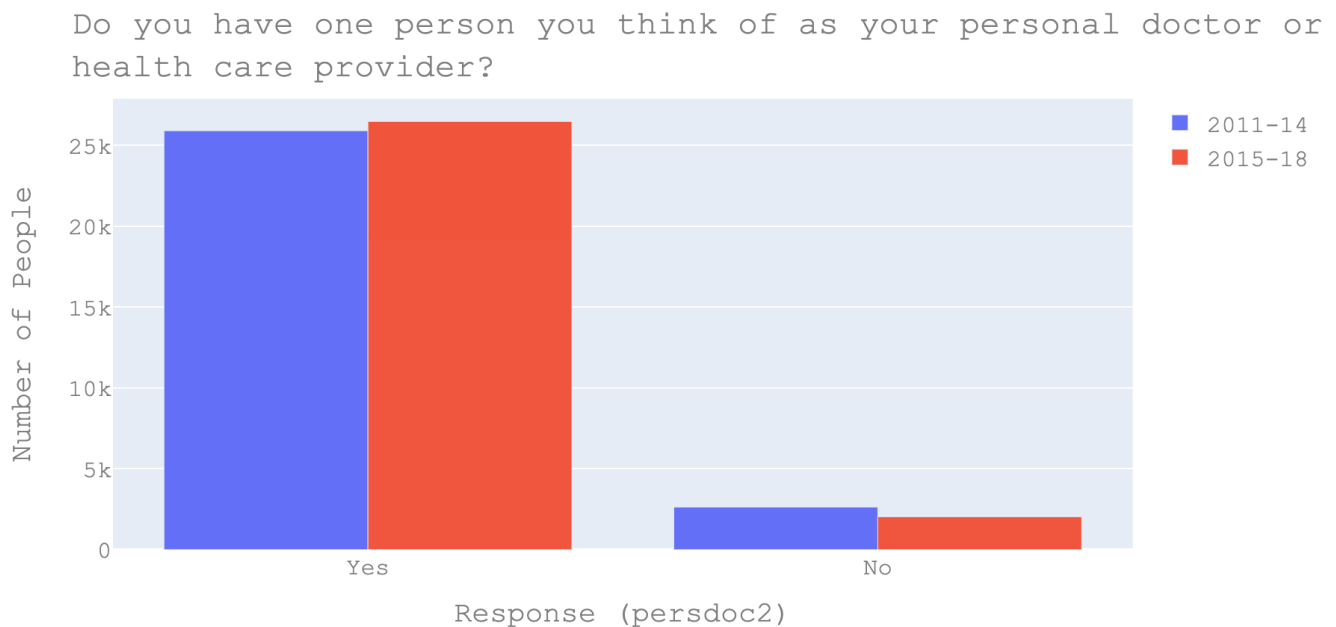
Figure 1: Hlthpln1

Figure 2: Persdoc2

Was there a time in the past 12 months when you
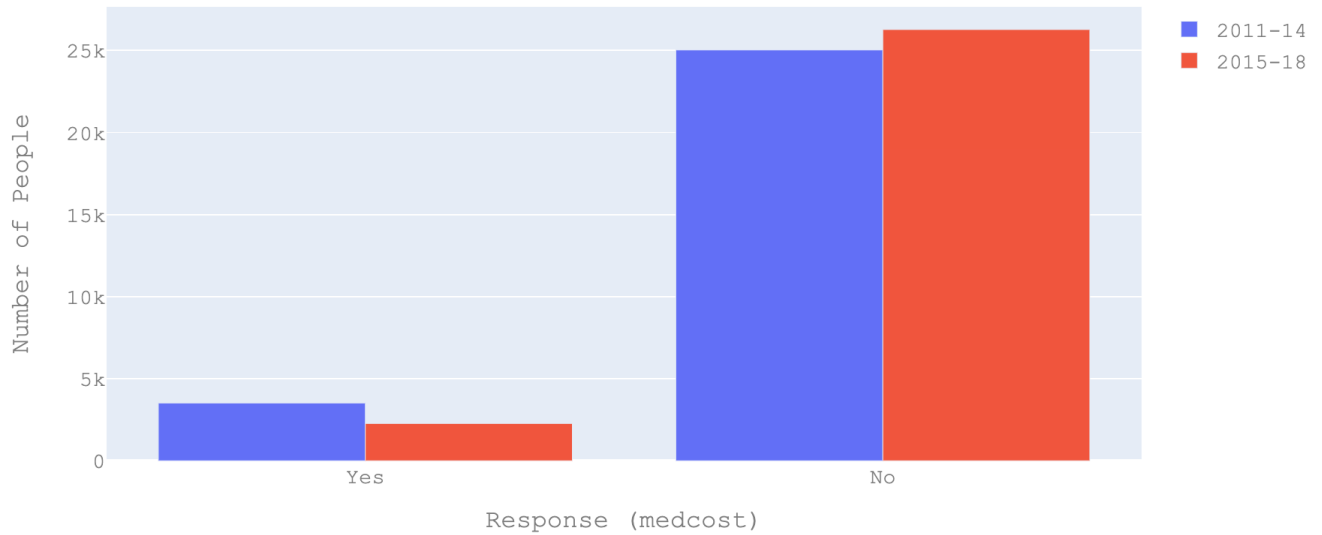 needed to see a doctor but could not because of cost?

Figure 3: Medcost

Not including over-the-counter (OTC) medications, was there a
 time in the past 12 months when you did not take your
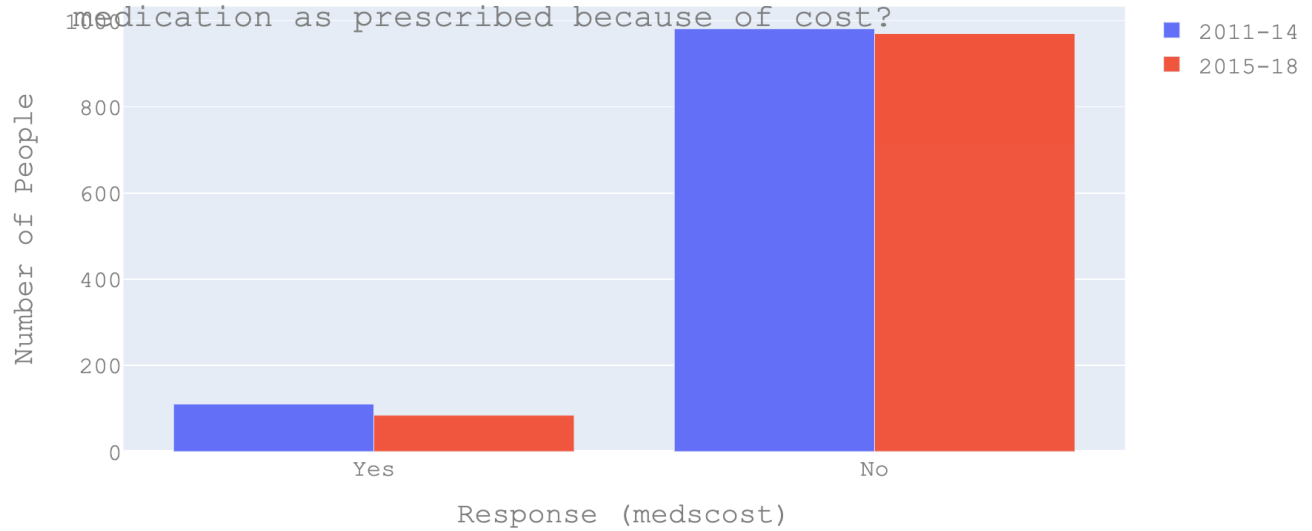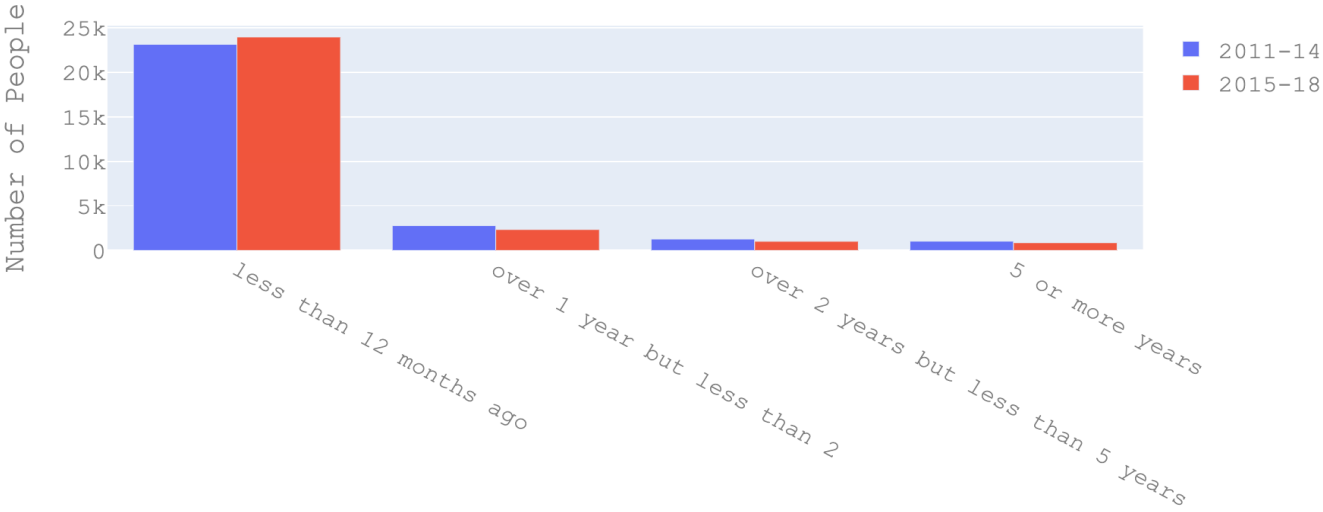medication as prescribed because of cost?

Figure 4: Medscost

Figure 5: Checkup1

Figure 6: Poorhlth

general health of the participants
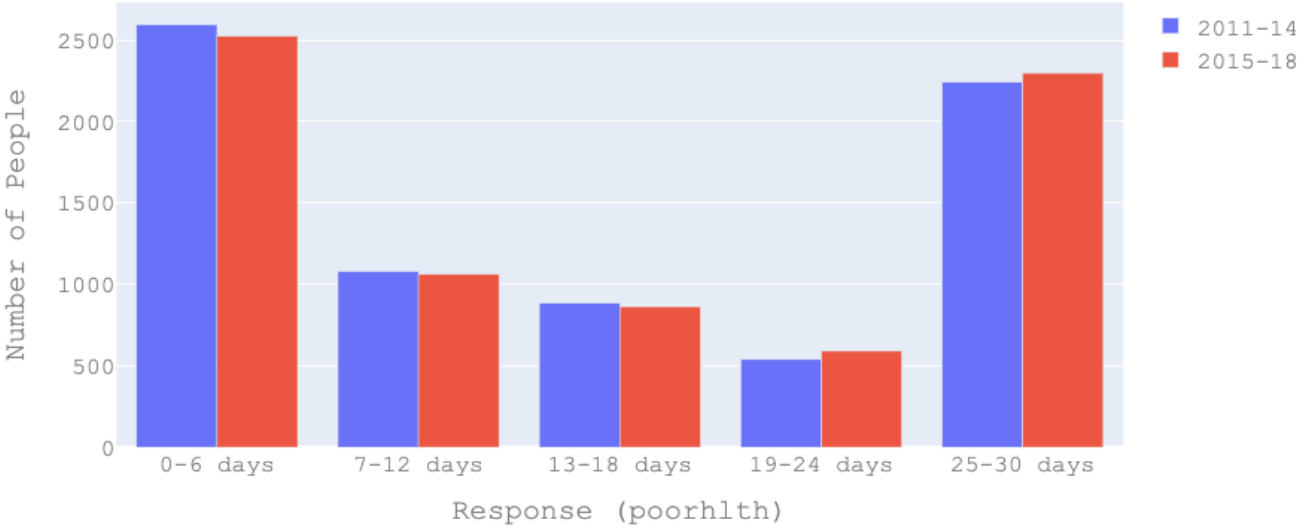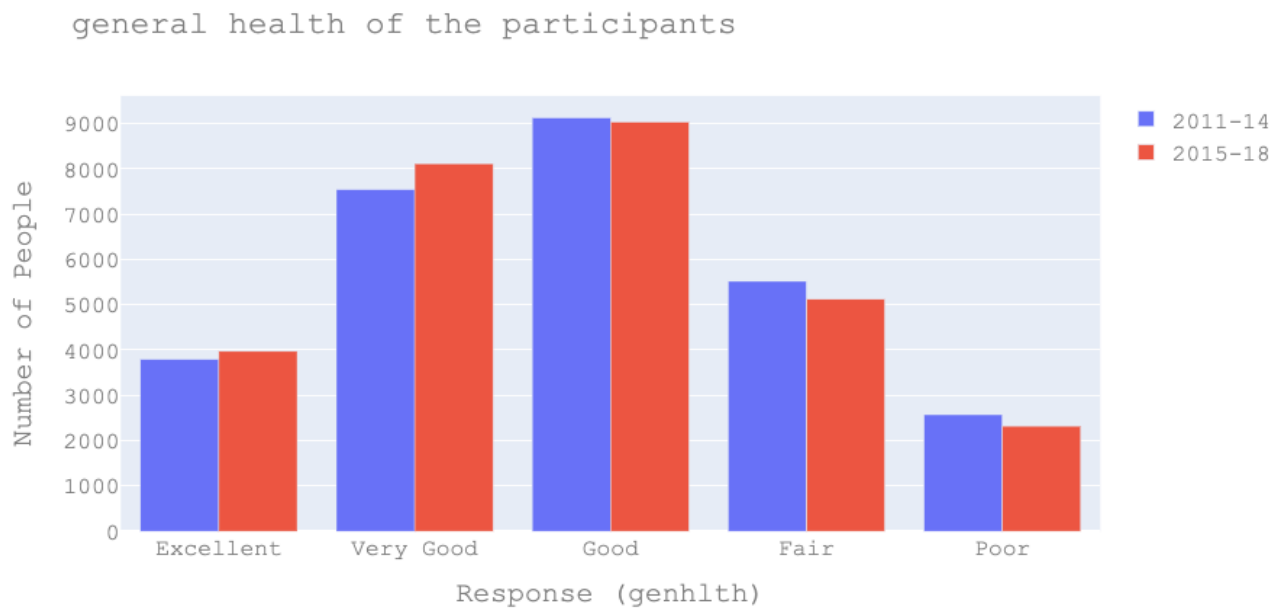
Figure 7: Genhlth

Future Directions

For our poster, we are planning to include a summary of what we were studying (the goal, methodology, etc.) and tables/graphs displaying our results. We are considering running more tests on screenings (such as mammogram, colonoscopy, or pap test) which might show some interesting results about access to preventative screens before and after the ACA.

As of now, all of our tests are done using chi-squared, since all of our variables are categorical except poorhlth which is continuous. We thought about using a t-test for poorhlth but one of the main assumptions for running a t-test is that the samples are independent. We cannot guarantee that the sample from 2011-2014 and 2015-2018 are independent, so we instead chose to divide the variable in bins and run a chi-squared. For future directions, we might try to look into what other analysis could be available for that variable instead of chi-squared.

Another goal we have is improving our data visualizations. We are hoping after completing the Data Viz lab and homework that we can better represent our data in graphs and tables.

Sources

(1)https://www.kff.org/health-reform/fact-sheet/preventive-services-covered-by-private-health-plans/
(2)https://aspe.hhs.gov/2019-poverty-guidelines
(3)https://www.usinflationcalculator.com/inflation/current-inflation-rates/