LOH Team - Healthcare and Elections

Lauren Kostich, Olan Trosky, Hubbert Smith

Contents

---

1. Political Context

Regardless of party, health insurance coverage remains a crucial campaign issue for candidates to resonate with voters on. CNN conducted a series of polls surrounding voting issues for 2018 voters. A 2018 pre-election poll asked likely voters how important an issue was to their vote and 52% said healthcare over other issues. This trend can be attributed to the failed repeal of the ACA in mid-2017. Following the 2018 general election, CNN conducted an exit poll asking voters the most important issue facing the country: healthcare, immigration, economy, and gun policy with 41% of respondents listing healthcare.

Most important issue facing the country

	Democrat	Republican	No Answer
Health care **41%**	75%	23%	2%
Immigration **23%**	23%	75%	2%
Economy **22%**	34%	63%	3%
Gun policy **10%**	70%	29%	1%

18778 Respondents

Looking at 2020, healthcare remains a crucial issue for voters. According to a poll conducted by Morning Consult and the Bipartisan Policy Centers, 56% of respondents selected health care as the most important issue in their vote choice for the presidential election. Within health policy: 66% of Democrats, 54% of Independents, and 46% of Republicans selected it as their most important issue. Healthcare remains incredibly important to voters in 2020, making it worth studying. The issue of healthcare has become incredibly contentious issue within party structures and between the two major parties. To better understand the importance of healthcare, it makes sense to look at the history and implementation of the Patient Protection and Affordable Care Act as the majority of current healthcare surrounds next steps for the legislation.

The Patient Protection and Affordable Care Act, commonly known as Obamacare. Signed in 2010, the ACA was signed into law to expand Medicaid with emphasis on making health insurance more affordable and supporting innovation to lower individual's overall cost. The major provisions of the ACA were implemented by 2014, covering an additional 20 to 24 million people. The 2010 legislation has been extremely divisive since it was passed. Liberals believe it did not go far enough in providing coverage, while conservatives have rallied against the individual mandate. This divisiveness provided voters in 2016, the first presidential election year since the bill's full implementation, a unique set of circumstances. We looked towards this election to get a better understanding of how healthcare coverage specifically impacts elections, attempting to bypass rhetoric about the issue and see how those directly impacted responded. We believe this data will give insight to the 2020 primaries where healthcare remains a prominent issue and over twenty candidates are currently vying for the Democratic nomination. By focusing on 2016 primaries, a year with a large number of Republican candidates, we hope to get a better understanding of the issue heading into 2020. As healthcare continues to be a key issue across demographics, this data will prove valuable to candidates on the federal and state level.

2. Literature Review

Currently, literature exists on Medicaid's impact on voter participation in general elections such as the 2008 Oregon election. This helped us better understand the potential for confounding variables in our own study and understand the viability of our own question.

In an attempt to better understand the interaction between government policy and political participation, researchers conducted a study on the impacts of Medicaid enrollment on voter turnout and registration. In 2008, Oregon used a lottery system to allocate 10,000 available slots in their Medicaid program to a waiting list of 90,000 low-income, previously uninsured adults (aged 19-64). The researchers linked administrative data on lottery participants and Medicaid enrollment with Oregon's statewide voter lists, allowing them to analyze voter turnout and registration for 2010. The researchers highlight the presence of confounding variables but conclude that the use of a randomized evaluate design avoids contamination by the confounding variables. These confounding variables include socio-economic status and health which could both impact voter participation (384). The study concluded that Medicaid enrollment impacted the 2008 election, but not subsequent elections. It should be noted that the 2008 election had much higher turnout the November 2010 election which can be attributed to 2008 being a Presidential election year. The study also notes that the Oregon expansion was not partisan or politicized while the national conversation was highly divisive.

3. Data and Methodology

Our data source was the American National Election Studies' (ANES) 2016 Time Series Study. The study is conducted through a collaboration between University of Michigan and Stanford University. The data includes various demographics and political variables taken from pre- and post-election surveys of the voting-age population in attempt to get a sample representative of the US adult population defined as individuals over the age of 18 with a valid mailing address. The data was collected through both face-to-face and internet responses conducted between September 7th and November 7th, 2016 for the pre-election survey and between November 9th and January 8th, 2017 for the post-election survey. It should be noted that respondents from Alaska and Hawaii were not included in the face-to-face model but were included in the Internet component. Individuals from Alaska and Hawaii make up less than one percent of the study population. It should also be noted that the survey was offered in both English and Spanish. We retrieved the data from the Inter-University Consortium for Political and Social Research.

4. Statistical Methods

Our independent variable x tracked whether or not an individual had healthcare coverage. Our independent variable, y, tracked whether or not an individual had voted in the 2016 presidential primary. Our control variables are gender, race, and education level. Based on the data provided and scope of our question, we presented the following hypothesis:

Linear Regressions:

$$y = a_1x_1 + a_2x_2 + \cdots a_nx_n$$

$$y = a_1(insurance)$$

$$(dep) = a_1(insurance) + a_2(ACA) + a_n(age) + a_n(edu) + a_n(job) + a_n(party) + a_n(sex) + a_n(socio)$$

Based on these equations, we ran the following output functions:

```
output1 <- lm(primary2016 ~ V161112, data=anes_subset03)
summary(output1)

output4 <- lm(primary2016 ~ insurance + job + age + ACA +
edu + socio + sex, data = anes_subset03)
```

We used the following variables to focus our study on:

V161277 - PRE: Initial R employment status, start of occupation module Recoded to "employment status"

V161267 – PRE: Respondent age Recoded as "age"

V161113 – PRE: Favor or oppose 2010 health care law recoded as "favor or oppose ACA"

V161112 – PRE: Does R have health insurance recoded as "insurance"

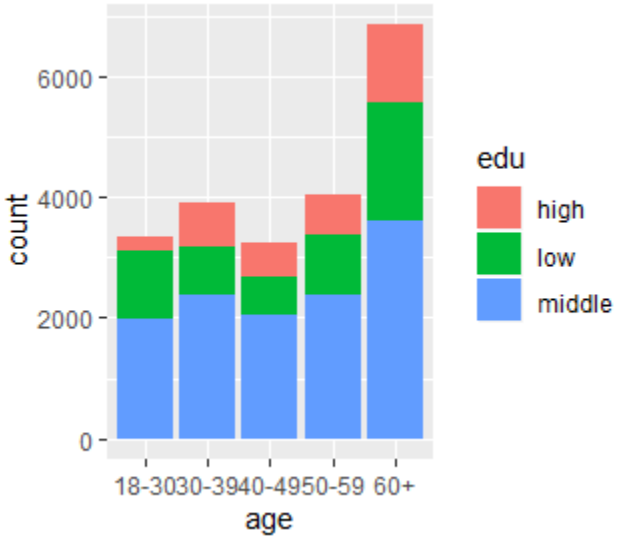
V161270 – PRE: Highest level of Education recoded as "education"

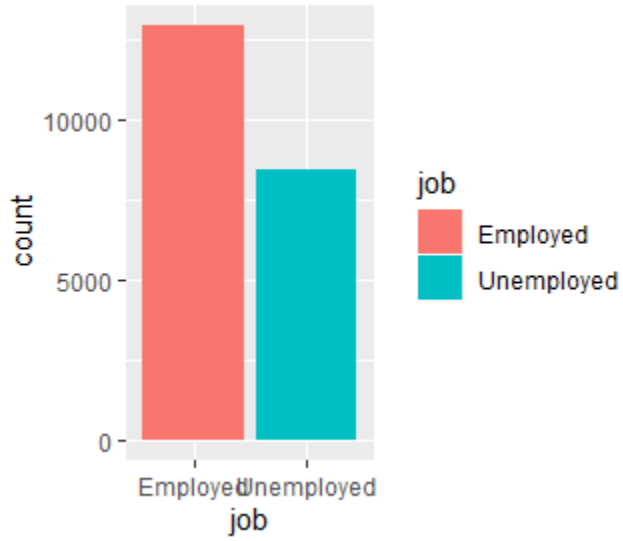
V161307 – PRE: Social class (2-question version) recoded as "socio"

V161342 – PRE FTF CASI/WEB: R self-identified gender recoded as "gender"

V16021 – PRE: Did R vote in a Presidential primary or caucus Recoded as "voted in presidential primary"

V161019 – PRE: Party of Registration Recoded as "party" ### ———————————— ### 5. Results





---

code run

```
# data with Amelia code
library(ggplot2)
library(tidyverse)
library(Amelia)
library(tibble)
library(ggplot2)
library(dplyr)
library(repmis)

source_data("https://github.com/hubbertsmith/up/blob/master/36824-0001-Data.rda?raw=true")

ANES2016 <- da36824.0001

anes_subset02 <- subset(ANES2016, select = c(V161021, V161112, V161342, V161307, V161019,
                                             V161270, V161113, V161267, V161277, V161004))

anes_subset02$primary2016 <- recode(anes_subset02$V161021,
                                    "(1) 1. Yes, voted in primary or caucus" = 1,
                                    "(2) 2. No, didn't vote in primary or caucus" = 0)

anes_subset02$V161342 <- recode(anes_subset02$V161342,
                                "(2) 2. Female" = 1, "(1) 1. Male" = 0)

anes_subset02$V161307 <- recode(anes_subset02$V161307,
                                "(1) 1. Lower class" = 0,
                                "(2) 2. Working class" = 0,
                                "(3) 3. Middle class" = 1,
                                "(4) 4. Upper class" = 1)
```

```
anes_subset02$V161019 <- recode(anes_subset02$V161019,
                                "(1) 1. Democratic party" = 0,
                                "(2) 2. Republican party" = 1,
                                "(4) 4. None or 'independent" = 2,
                                "(5) 5. Other SPECIFY" = 2)

anes_subset02$V161270 <- recode(anes_subset02$V161270,
                                "(02) 2. 1st, 2nd, 3rd or 4th grade" = 2,
                                "(03) 3. 5th or 6th grade" = 3,
                                "(04) 4. 7th or 8th grade" = 4,
                                "(05) 5. 9th grade" = 5,
                                "(06) 6. 10th grade" = 6,
                                "(07) 7. 11th grade" = 7,
                                "(08) 8. 12th grade no diploma" = 8,
                                "(09) 9. High school graduate- high school diploma or equivalent (for e
                                "(10) 10. Some college but no degree" = 10,
                                "(11) 11. Associate degree in college - occupational/vocational program
                                "(12) 12. Associate degree in college -- academic program" = 12,
                                "(13) 13. Bachelor's degree (for example: BA, AB, BS)" = 13,
                                "(14) 14. Master's degree (for example: MA, MS, MENG, MED, MSW, MBA)" =
                                "(15) 15. Professional school degree (for example: MD, DDS, DVM, LLB, J
                                "(16) 16. Doctorate degree (for example: PHD, EDD)" = 16)
anes_subset02$V161112 <- recode(anes_subset02$V161112, "(1) 1. Yes" = 1,
                                "(2) 2. No" = 0)
anes_subset02$V161113 <- recode(anes_subset02$V161113, "(1) 1. Favor" = 1,
                                "(2) 2. Oppose" = 2,
                                "(3) 3. Neither favor nor oppose" = 3)
anes_subset02$V161277 <- recode(anes_subset02$V161277, "(1) 1. Initial employment status: working now" =
                                "2) 2. Initial employment status: temporarily laid off" = 0,
                                "(4) 4. Initial employment status: unemployed" = 0,
                                "(5) 5. Initial employment status: retired" = 0,
                                "(6) 6. Initial employment status: permanently disabled" = 0,
                                "(7) 7. Initial employment status: homemaker" = 0,
                                "(8) 8. Initial employment status: student" = 0)
anes_subset02$V161004 <- recode(anes_subset02$V161004, "(1) 1. Very much interested" = 3,
                                "(2) 2. Somewhat interested" = 2,
                                "(3) 3. Not much interested" = 1)

am_output <- amelia(anes_subset02, idvars = "primary2016", noms = c("V161307","V161113", "V161277", "V1
                    ords =c ("V161112", "V161270", "V161019", "V161267"))

am_output2 <- bind_rows(unclass(am_output$imputations), .id = "m") %>%
  group_by(m) %>%
  nest()

am_clean <- rbind(am_output2[[2]][[1]],am_output2[[2]][[2]],am_output2[[2]][[3]],am_output2[[2]][[4]],a

#rename

#dependent variable
#pre 2016: voted in presidential primary
am_clean$dep <- -1
am_clean$dep[am_clean$primary2016== 1] <- "voted"
```

```r
am_clean$dep[am_clean$primary2016== 0] <- "abstained"
am_clean$dep <- factor(am_clean$dep)


#new variable: gender
am_clean$sex <- -1
am_clean$sex[am_clean$V161342== 1] <- "female"
am_clean$sex[am_clean$V161342== 0] <- "male"
am_clean$sex <- factor(am_clean$sex)

#new variable: socioeconomic status
am_clean$socio <- -1
am_clean$socio[am_clean$V161307== 0] <- "low"
am_clean$socio[am_clean$V161307== 1] <- "high"
am_clean$socio <- factor(am_clean$socio)

#new variable: Education
#need to get rid of "others" (90 & 95)
am_clean$edu <- -1
am_clean$edu[am_clean$V161270 <= 9] <- "low"
am_clean$edu[am_clean$V161270 >= 10]<- "middle" #10-13
am_clean$edu[am_clean$V161270 >= 14] <- "high" #14-16
am_clean$edu <- factor(am_clean$edu)

#independent variable: health insurance
#does the subject have health insurance?
am_clean$insurance <- -1
am_clean$insurance[am_clean$V161112== 1] <- "coverage"
am_clean$insurance[am_clean$V161112== 0] <- "no coverage"
am_clean$insurance <- factor(am_clean$insurance)

#new variable: favor or oppose ACA
am_clean$ACA <- -1
am_clean$ACA[am_clean$V161113== 1] <- "favor"
am_clean$ACA[am_clean$V161113== 2] <- "oppose"
am_clean$ACA[am_clean$V161113== 3] <- "neutral"
am_clean$ACA <- factor(am_clean$ACA)

#new variable: age
am_clean$age <- -1
am_clean$age[am_clean$V161267 >= 18] <- "18-30"
am_clean$age[am_clean$V161267 >= 30] <- "30-39"
am_clean$age[am_clean$V161267 >= 40] <- "40-49"
am_clean$age[am_clean$V161267 >= 50] <- "50-59"
am_clean$age[am_clean$V161267 >= 60] <- "60+"
am_clean$age <- factor(am_clean$age)

#new variable: employment status
am_clean$job <- -1
am_clean$job[am_clean$V161277== 1] <- "Employed"
am_clean$job[am_clean$V161277== 0] <- "Unemployed"
am_clean$job <- factor(am_clean$job)
```

```r
am_clean$party <- -1
am_clean$party[am_clean$V161019== 0] <- "dem"
am_clean$party[am_clean$V161019== 1] <- "rep"
am_clean$party[am_clean$V161019== 2] <- "ind_other"
am_clean$party[am_clean$V161019== 3] <- "ind_other"
am_clean$party <- factor(am_clean$party)
```

Regressions

```r
#regressions

am_clean$insurance <- relevel(am_clean$insurance, ref = "no coverage")

output3 <- lm(primary2016 ~ insurance, data = am_clean)
summary(output3)
```

```
##
## Call:
## lm(formula = primary2016 ~ insurance, data = am_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4593 -0.4593 -0.2540  0.5406  0.7460
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.25399    0.01138   22.32   <2e-16 ***
## insurancecoverage  0.20536    0.01192   17.23   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4931 on 21323 degrees of freedom
##   (30 observations deleted due to missingness)
## Multiple R-squared:  0.01374,    Adjusted R-squared:  0.01369
## F-statistic:   297 on 1 and 21323 DF,  p-value: < 2.2e-16
```

```r
output4 <- lm(primary2016 ~ insurance + job + age + ACA + edu + socio + sex, data = am_clean)
summary(output4)
```

```
##
## Call:
## lm(formula = primary2016 ~ insurance + job + age + ACA + edu +
##     socio + sex, data = am_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8478 -0.4283 -0.2017  0.4833  0.9855
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)        0.472601   0.033577  14.075  < 2e-16 ***
## insurancecoverage  0.078188   0.011841   6.603 4.12e-11 ***
## jobUnemployed      0.016699   0.007644   2.185   0.0289 *
## age30-39           0.046024   0.011254   4.089 4.34e-05 ***
## age40-49           0.109408   0.011783   9.285  < 2e-16 ***
## age50-59           0.169118   0.011193  15.110  < 2e-16 ***
## age60+             0.280290   0.010794  25.968  < 2e-16 ***
## ACAneutral        -0.108975   0.009007 -12.099  < 2e-16 ***
## ACAoppose         -0.045635   0.007377  -6.186 6.29e-10 ***
## edulow            -0.169259   0.011005 -15.381  < 2e-16 ***
## edumiddle         -0.041243   0.009378  -4.398 1.10e-05 ***
## sociolow          -0.060910   0.006921  -8.801  < 2e-16 ***
## sexfemale         -0.118918   0.029650  -4.011 6.07e-05 ***
## sexmale           -0.122922   0.029682  -4.141 3.47e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4721 on 21311 degrees of freedom
##   (30 observations deleted due to missingness)
## Multiple R-squared:  0.09679,    Adjusted R-squared:  0.09624
## F-statistic: 175.7 on 13 and 21311 DF,  p-value: < 2.2e-16
```

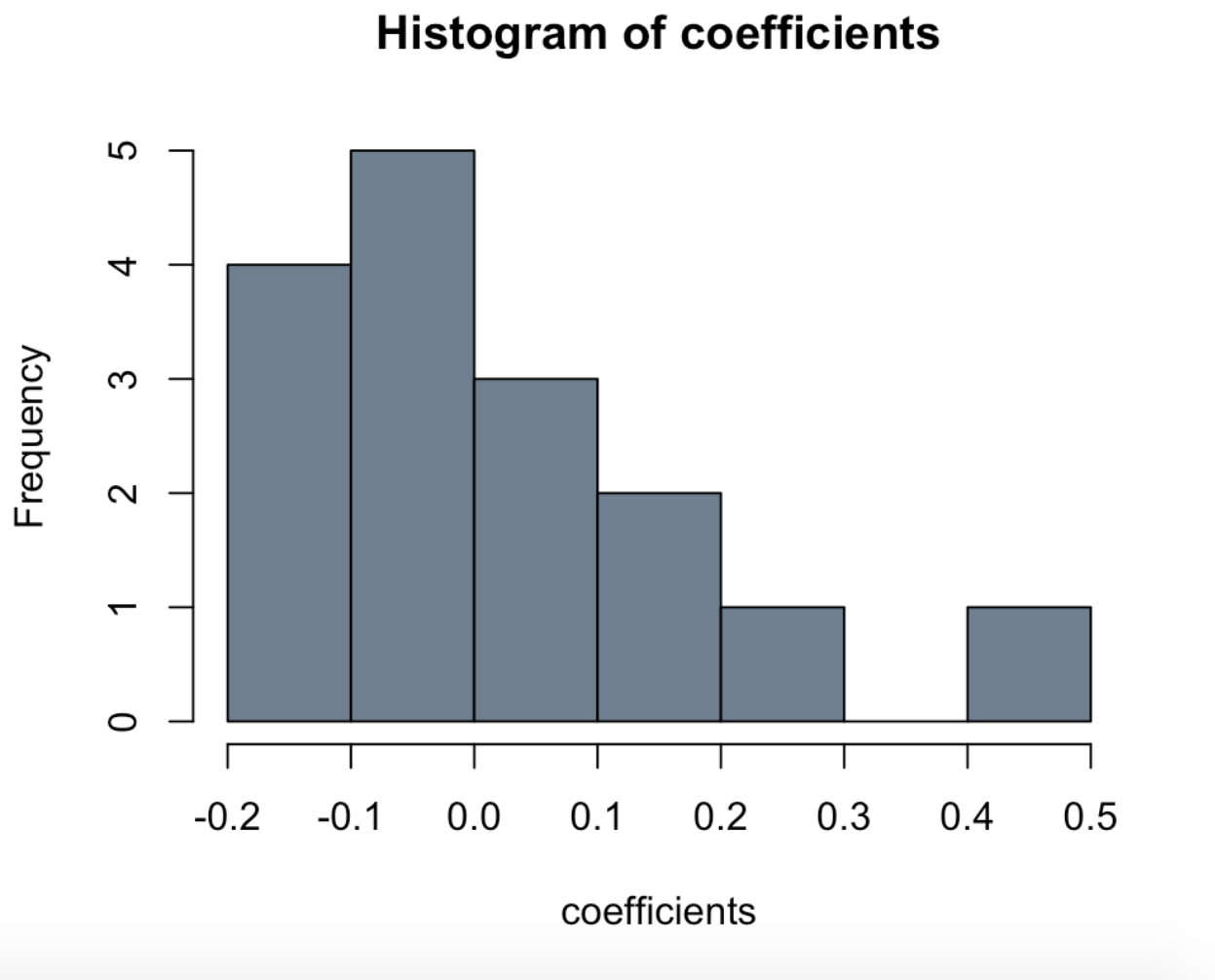
Output 1:

The bivariate regression shows a fairly symmetrical distribution of residuals which demonstrates that the model predicts points that occur pretty closely to the actual observed points. The expected value of the coefficient estimate for the intercept demonstrates that 25% of all those surveyed in the dataset to show up to vote. Essentially, 25% of those surveyed regardless of possessing insurance are going to show up to vote in the presidential primary. The coefficient estimate demonstrates a positive relationship for respondents with coverage and they have 20% greater odds than those without insurance to vote in a presidential primary. This finding accepts the null hypothesis which is people that don't have healthcare are less likely to vote in the presidential primary.

Output 2:

Again, the residuals for second regression model, with all selected variables, shows a symmetrical distribution which affirms the predicted model fit to the actual data. When including all selected variables, the intercept coefficient increases from .25 to .48. The main variable of interest, whether or not the respondent has healthcare coverage, remains stronger for those who possess insurance. However, the factor has decreased by 65%. In terms of employment, this model shows that employment status and gender have little to no affect on whether or not the respondent are likely to vote in the presidential primary. As expected, the model reaffirms that the higher the level of education and socioeconomic status the more likely that respondent is to vote in a presidential primary. In general, older populations are a larger percentage of the voting bloc in the United States. The multivariate regression reaffirms that as respondents increase in age, measured here by decade, the more likely they are to participate in a presidential primary. In regard to the variable that measures the respondent's opinion regarding implementation of the Affordable Care Act (ACA), this model shows that those who either oppose or are in favor in the ACA are more likely to vote in the presidential primary than those with a neutral position.

These findings are consistent with literature review. ### Histogram



For the second output, a multivariable regression, we ran a histogram to get a visual representation of the coefficients. The majority of our coefficients had a weak correlation, but we did find that age had a significant impact on the regression.
———————————

Sources

Enten, H. (2019, March 26). Unlike Mueller, health care will likely be a top issue in 2020. Retrieved from https://www.cnn.com/2019/03/26/politics/health-care-2020-campaign-issue/index.html.

Baicker, K., & Finkelstein, A. (2019). The Impact of Medicaid Expansion on Voter Participation: Evidence from the Oregon Health Insurance Experiment. Quarterly Journal of Political Science, 14(4), 383–400. Retrieved from https://economics.mit.edu/files/18481

Summary

THE END <<<

Workflow

assign issues

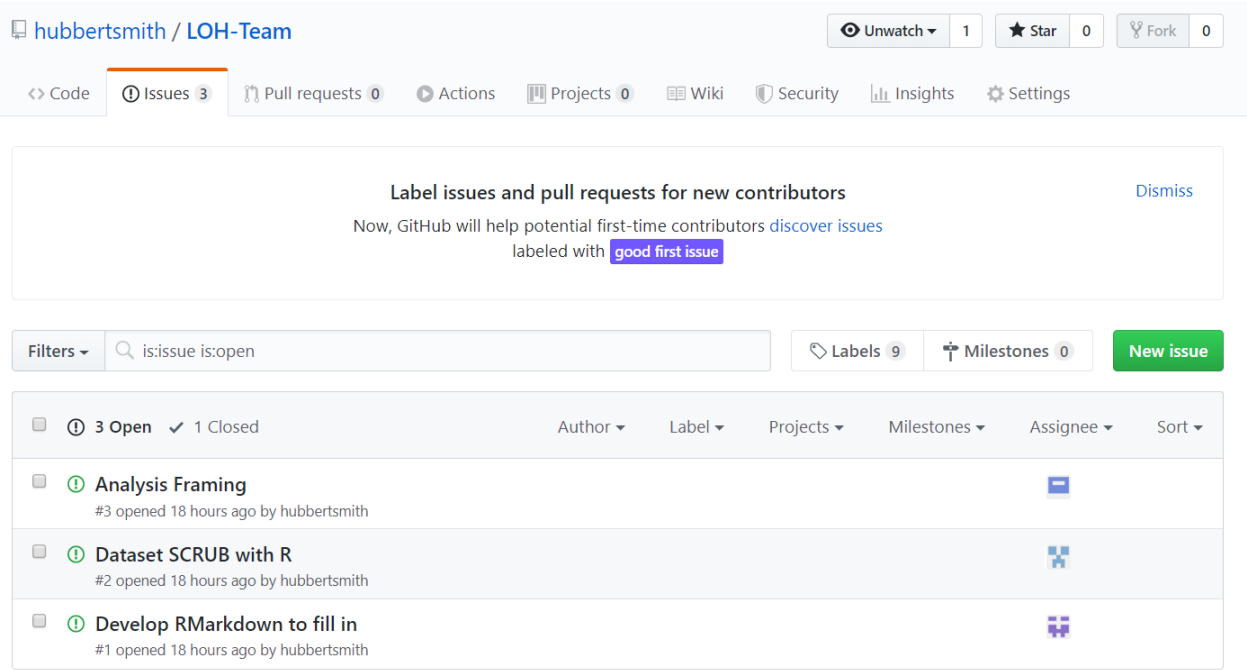
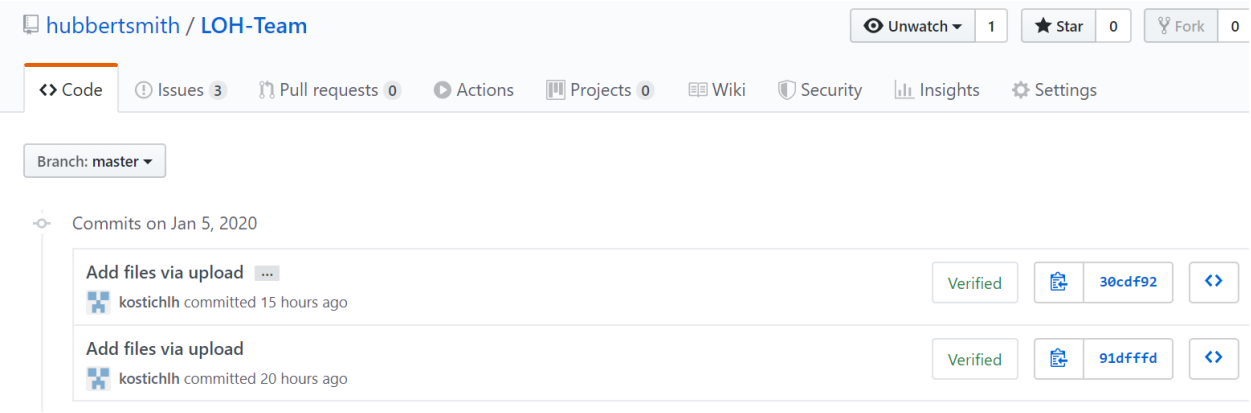
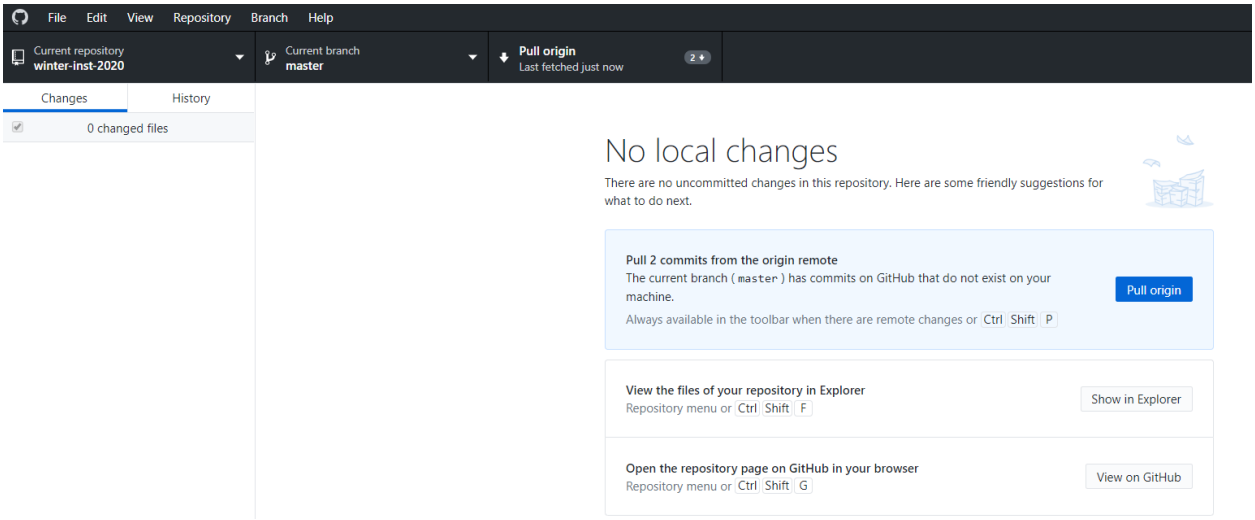


Figure 1: create GITHUB issues, who does what .

Lauen – create and upload files



Olan and Hubbert — Pull files



Lauren — Submit files ss