R Course: Lesson 5

Cédric HASSEN-KHODJA

June 28, 2016

Introduction

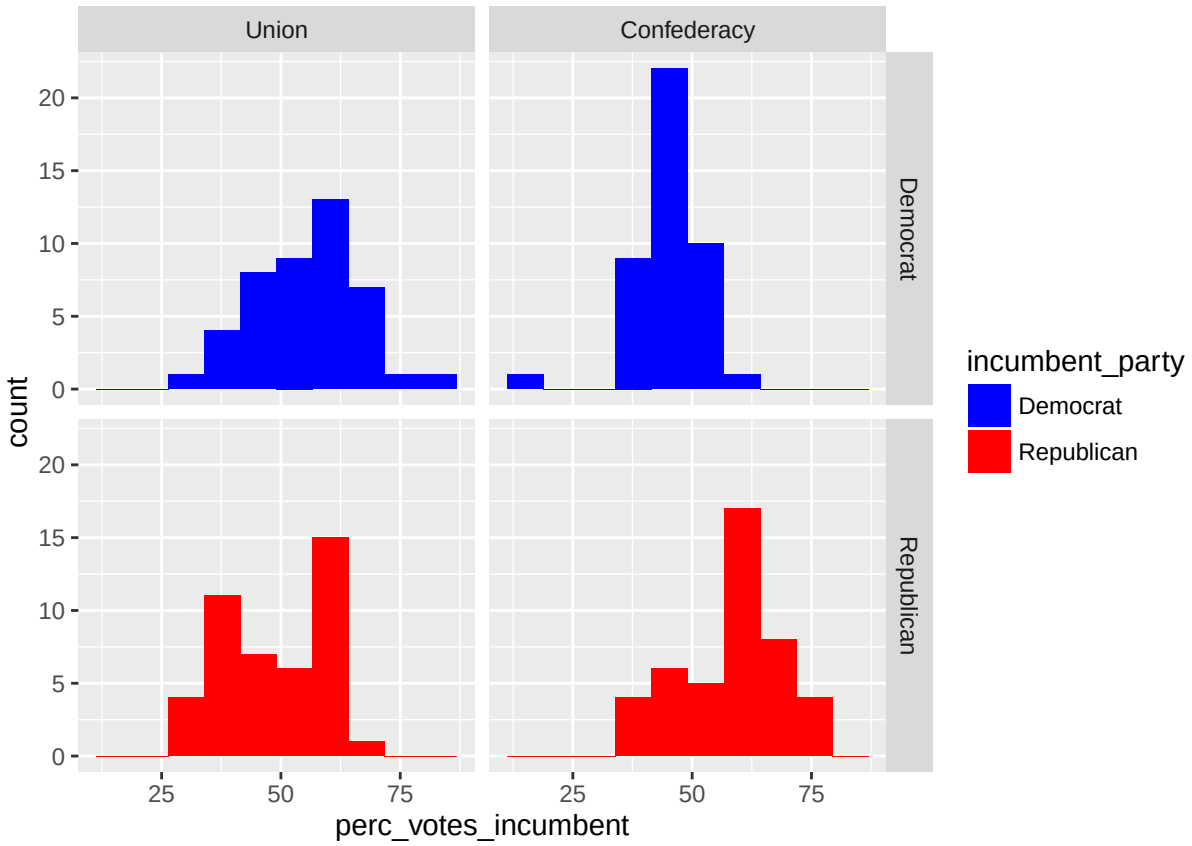
Today I looked at election data from eight United States presidential elections (1964, 1972, 1980, 1984, 1992, 1996, 2004, 2012). Specifically, I looked at elections where an incumbent was running for president. I wanted to see if the percentage of the population that voted for the incumbent varied by the political party of the incumbent (Democrat, Republic) and whether the state was part of the Union or the Confederacy during the civil war.

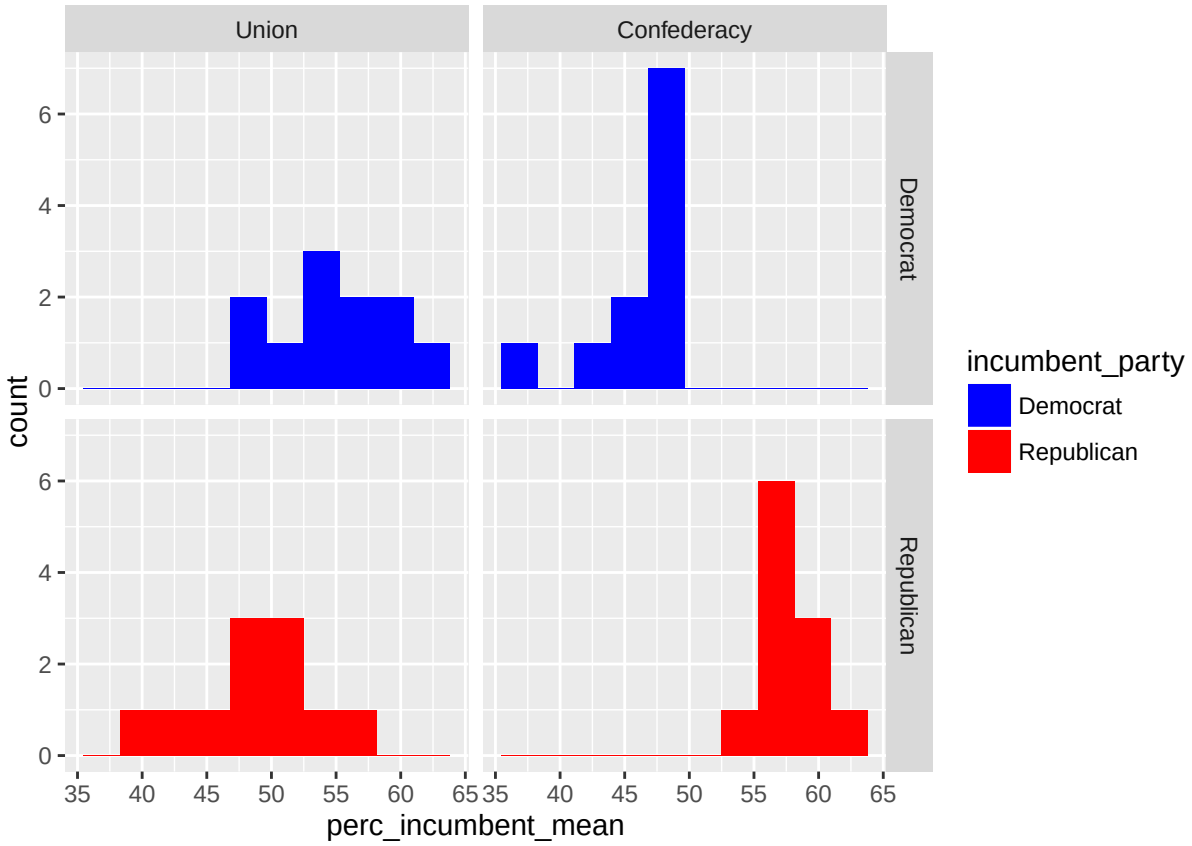
Data

Since I'm using an ANOVA today, I needed to make sure my data set was balanced. So, instead of taking all states that were officially states during the civil war, I made sure it was the same number in each group (Union, Confederacy). There were only 11 Confederacy states, so to get a matched sample of Union states I used data from the first 11 Union states that were admitted to the United States. For example, California was not included because it joined the United States later than other states.

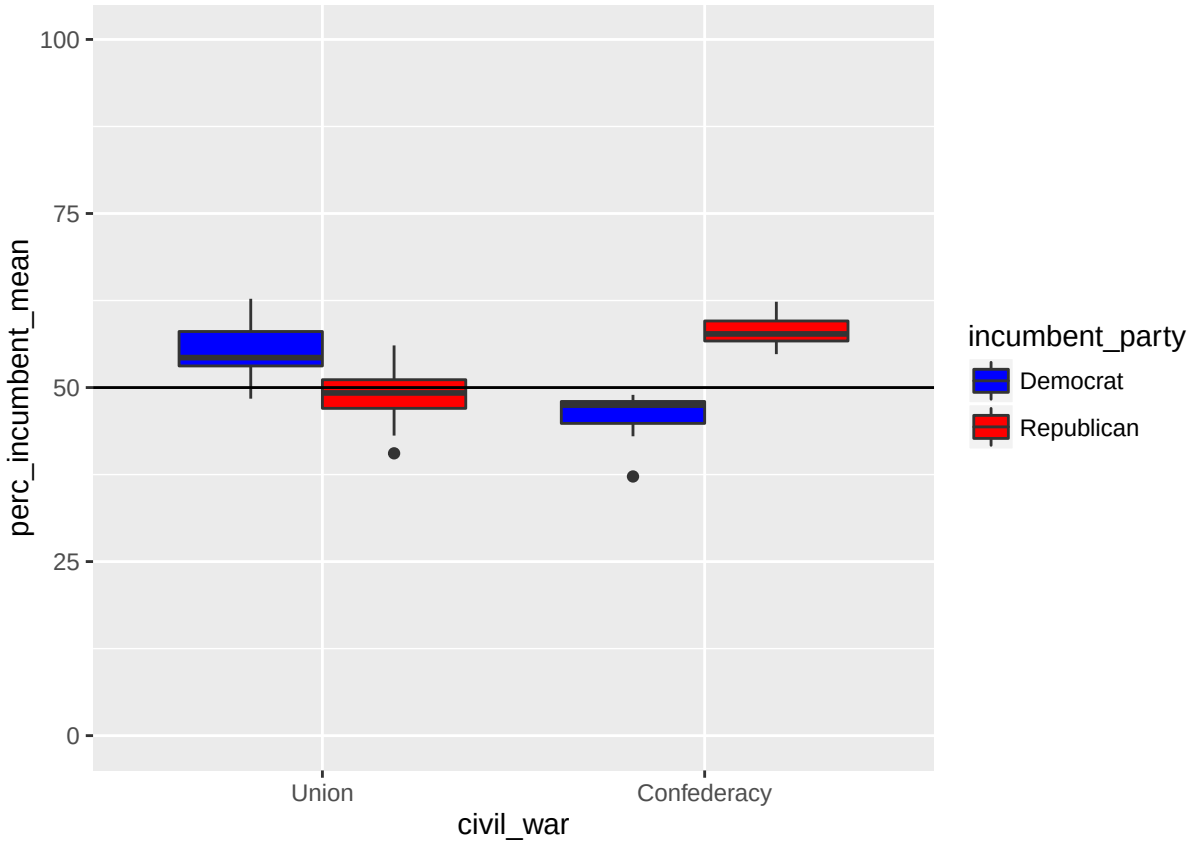
Results

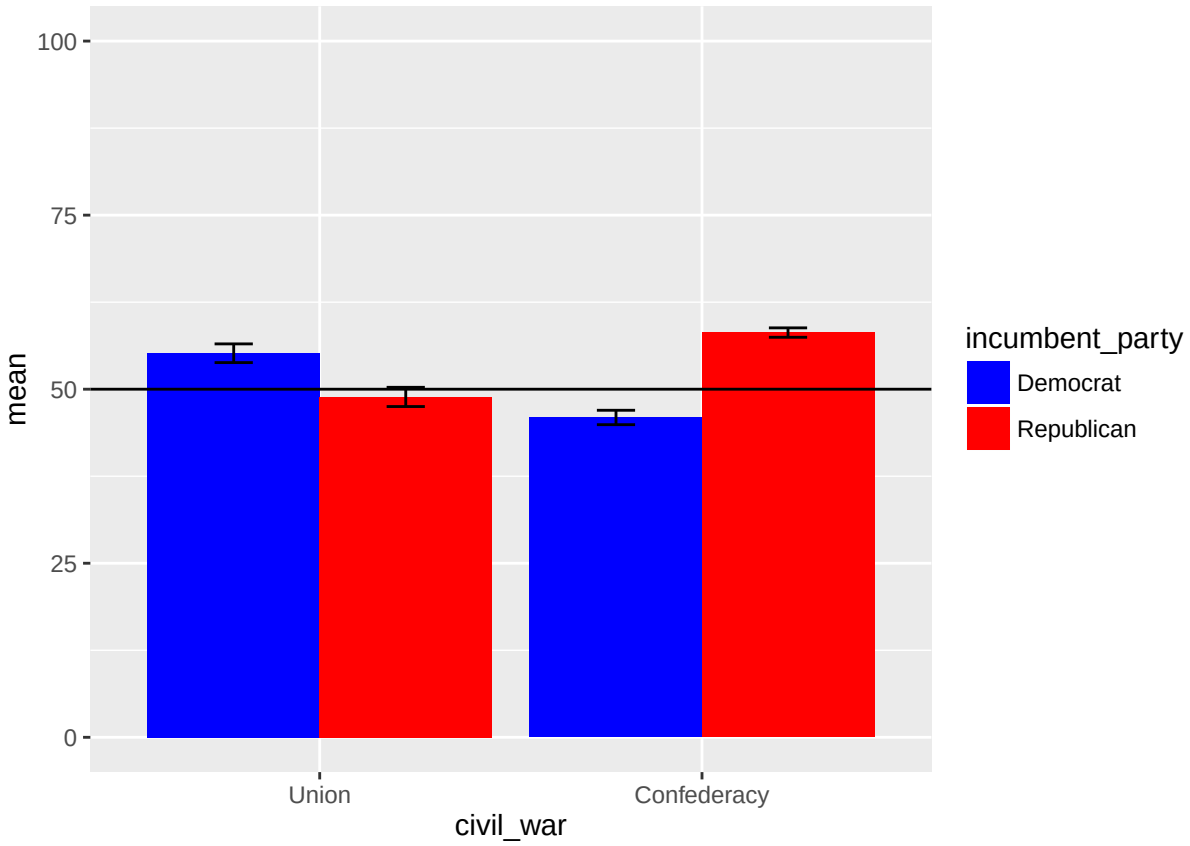
First I plotted histograms of our data separated by our two variables to be sure they had a normal distribution. I did this over the entire data set and separating the data by my categorical variables. The distributions do not look completely normal (particularly for the summarized data), but enough such that I am going to more forward with the analysis without transforming my data.


Union
Confederacy
Democrat
Republican
count
20
15
10
5
0
25
50
75
perc_votes_incumbent
incumbent_party
Democrat
Republican


Union
Confederacy
Democrat
Republican
count
6
4
2
0
35
40
45
50
55
60
65
perc_incumbent_mean
incumbent_party
Democrat
Republican

I've plotted my data both as a boxplot and a barplot, but I think the boxplot is better. Both suggest that states that stayed in the Union vote more for incumbent Democrats than incumbent Republicans, conversely, states that joined the Confederacy show the opposite effect. This suggests there may be an interaction in the data.





To test this effect I ran an analysis of variance (ANOVA). The dependent variable was percentage of the population voting for the incumbent, the independent variables were incumbent party (Democrat, Republican) and the state's country during the civil war (Union, Confederacy). Incumbent party was included as a within-state variable. The model did not find either main effect to be significant, although the effect of incumbent party was trending [$F(1, 20) = 3.59$, $p = 0.07$]. There was a significant interaction of incumbent party and civil war [$F(1, 20) = 34.94$, $p < 0.001$] .

```
incumbent.aov_sum
```

```
##
## Error: state
##           Df Sum Sq Mean Sq F value Pr(>F)
## civil_war  1   0.00  0.0007       0  0.985
## Residuals 20  39.92  1.9961
##
## Error: state:incumbent_party
##                            Df Sum Sq Mean Sq F value   Pr(>F)
## incumbent_party             1   96.4    96.4   3.592   0.0726 .
## incumbent_party:civil_war   1  937.7   937.7  34.941 8.81e-06 ***
## Residuals                  20  536.7    26.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

To better understand the interaction of incumbent party and civil war, I ran t-tests looking at my two main effects within subsets of the data. To account for my multiple tests, I did Bonferroni correction, making my new p-value for significance 0.0125. Looking first within civil war country, I ran paired t-tests to see if

either group showed a difference of incumbent party. I found that for Union states there was not a significant effect given my p-value correction ($p = 0.0408$). However, for states from the Confederacy the effect was very strong ($p < 0.001$), showing that Confederacy states have a strong preference for Republican incumbents. By looking at the mean of the differences for each test we can further say that Confederacy states' preference is indeed much larger than Union states' preference.

```
incumbent_union.ttest
```

```
##
##  Paired t-test
##
## data:  data_union_stats$democrat and data_union_stats$republican
## t = 2.3477, df = 10, p-value = 0.0408
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.3195664 12.2258881
## sample estimates:
## mean of the differences
##                6.272727
```

```
incumbent_confederacy.ttest
```

```
##
##  Paired t-test
##
## data:  data_confederacy_stats$democrat and data_confederacy_stats$republican
## t = -7.5325, df = 10, p-value = 1.987e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -15.799975  -8.586389
## sample estimates:
## mean of the differences
##               -12.19318
```

I can also look at the results within party, to see if there is a difference depending on the status of a state during the civil war. To do this I ran unpaired t-tests within each party. For Democrats there was a significant effect ($p < 0.001$), where Democratic incumbents get a much higher percentage of votes from Union states than Confederacy state. Republicans show a significant effect in the reverse direction ($p < 0.001$), getting a higher percentage of votes from Confederacy states.

```
incumbent_democrat.ttest
```

```
##
##  Welch Two Sample t-test
##
## data:  perc_incumbent_mean by civil_war
## t = 5.4386, df = 18.774, p-value = 3.14e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   5.671881 12.778119
## sample estimates:
##       mean in group union mean in group confederacy
##                  55.16364                  45.93864
```

```
incumbent_republican.ttest
```

```
##
##  Welch Two Sample t-test
##
## data:  perc_incumbent_mean by civil_war
## t = -6.0086, df = 14.529, p-value = 2.741e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -12.528222  -5.953596
## sample estimates:
##       mean in group union mean in group confederacy
##                  48.89091                  58.13182
```

Conclusion

A state's status during the civil war has a large effect on how it votes. However, the effect is stronger for states from the Confederacy, who have a much stronger Republican bias than Union states have a Democrat bias.