Test of Association for Political Data between States

*Belle Joy, Christopher Davis, Jacob Mishler*

*12/11/19*

**Contents**

Introduction………………………………………………………………………..............2

About the data………………………………………………………………..……………2

Data Clean-Up…………………………………………………………………………….2

Data Testing……………………………………………………………………………….3

California Gender Voting Statistics……………………………………………………….3

California Racial Voting Statistics……………………………………………..………….4

Florida Gender Voting Statistics……………………………………………………..........5

Florida Racial Voting Statistics………………………………………………….………..5

Georgia Gender Voting Statistics…………………………………………………………6

Georgia Racial Voting Statistics………………………………………………………….7

Indiana Gender Voting Statistics…………………………………………………….........7

Indiana Racial Voting Statistics…………………………………………………………...8

Iowa Gender Voting Statistics…………………………………………………….............9

Iowa Racial Voting Statistics……………………………………………………………..9

Kentucky Gender Voting Statistics………………………………………………………10

Kentucky Racial Voting Statistics……………………………………………………….11

Michigan Gender Voting Statistics………………………………………………………11

Michigan Racial Voting Statistics……………………………………………………….12

Minnesota Gender Voting Statistics…………………………………………………......12

Minnesota Racial Voting Statistics………………………………………………………13

Missouri Gender Voting Statistics……………………………………………………….14

Missouri Racial Voting Statistics………………………………………………………...14

Nevada Gender Voting Statistics………………………………………………………...15

Nevada Racial Voting Statistics………………………………………………………….16

Limitations of Study………………………………………………………...…………...16

Conclusion and Final Thoughts………………………………………………………….16

References………………………………………………………………………………..17

**Introduction**

The 2016 United States Presidential race took place between Democrat Hillary Clinton and Republican Donald Trump. We collected exit-polling data from CNN to determine what factors separated voting preferences. While the data shows trends based on age, gender, race, education, and other socioeconomic factors, we collected only data for differences in gender and race. We then used the Chi-Square test to determine if there is a significant difference between the two factors.

**About the data**

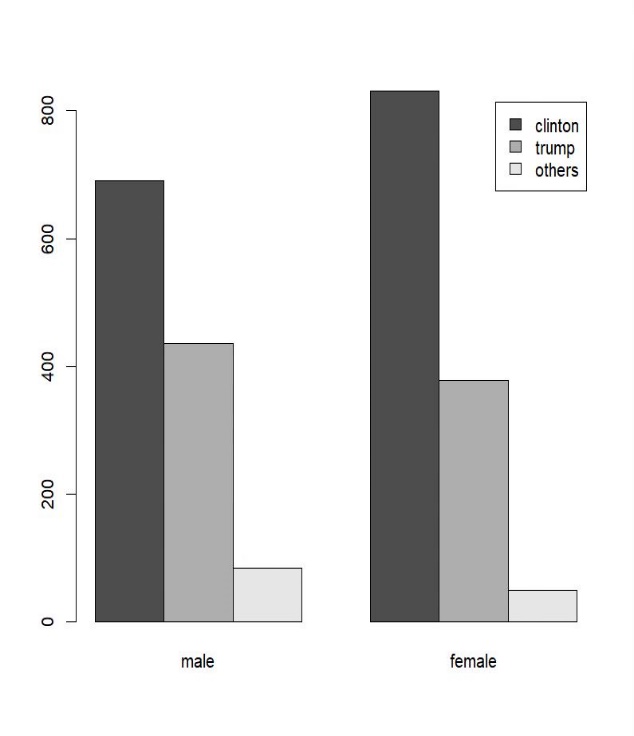
* The data being analyzed is the exit polls that were taken from the 2016 presidential election between Hillary Clinton and Donald Trump.
* All tests ran within the scope of this project will be assumed to have a 95 percent confidence interval (also considered to be an alpha of 0.05).
* The main hypothesis that will be tested will be:
  + H0: The voting proportions are equal and there is no difference between men and women, or between whites and nonwhites.
  + H1: The voting proportions are not equal and there is a difference between men and women, or between whites and nonwhites.
* Using inference from the graphs and geographic location, assumptions will be able to be made regarding how geography affects the voting process for the top candidates

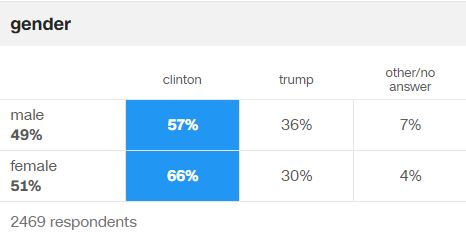
**Data Clean-Up**

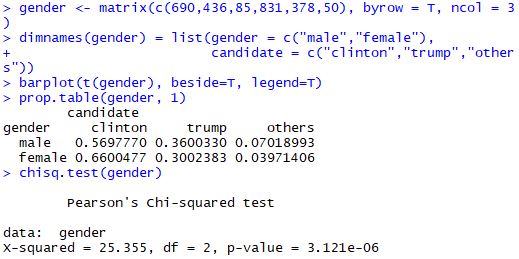
Within the dataset provided from CNN, there were imperfections that had to be considered with the focus of the specific hypothesis at hand. The main issue with the data was that there were some states that had a decent number of votes counting towards candidates, other than Trump and Clinton, that were not considered in the other category. Libertarian candidate Johnson and Green candidate Stein are candidates that have a significant number of votes, which could throw off the tests that will be conducted. For simplicity, we will include these two candidates with “other” votes.

**Data Testing**

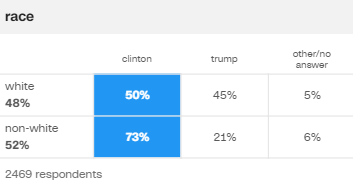
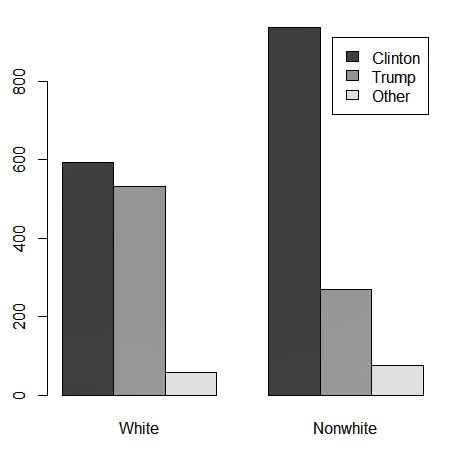
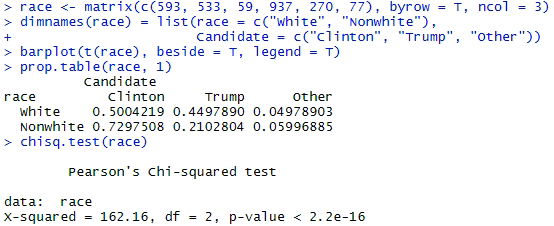
For each of the 10 states we collected data, we believed it was most appropriate to create a bar plot, proportion table, and run a Chi-Square test. Creating the bar plot will allow the reader to have an alternative to the table by seeing the amount of people that voted rather than a percentage. Creating a proportion table will show the precise proportions of the number of voters for each presidential candidate. The Chi-Square test will allow the reader to determine if the null hypothesis, which states that there is no difference in voting patterns, can be rejected or accepted.

**California Gender Voting Statistics**

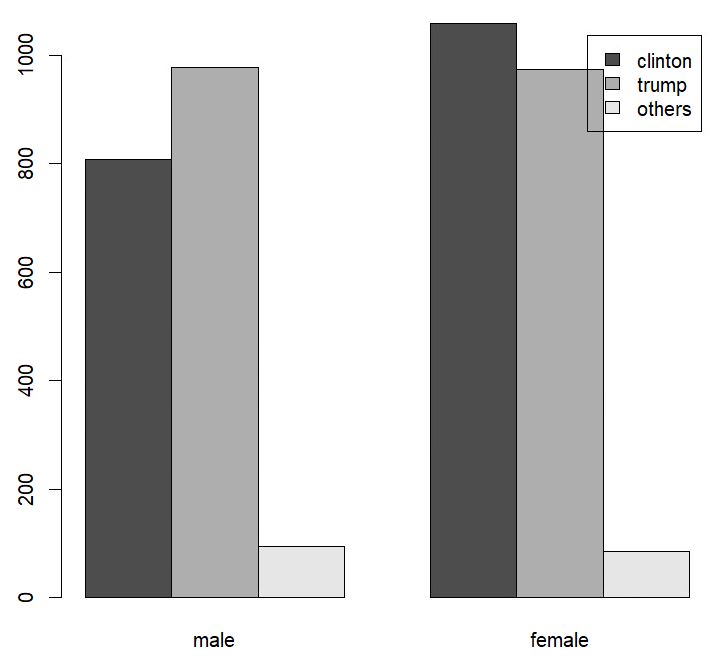


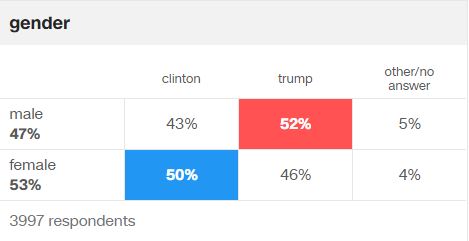
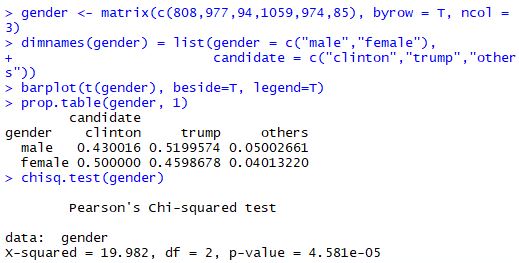
 The exit polling from the state of California suggests approximately 57% of men and 66% of women voted for Clinton, while 36% of men and 30% of women voted for Trump. Using the Chi-Square test, we received a p-value of 3.121e-6, which is much smaller than alpha. So, for California, we reject the null hypothesis.

**California Racial Voting Statistics**



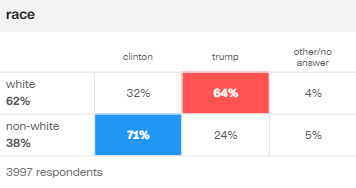
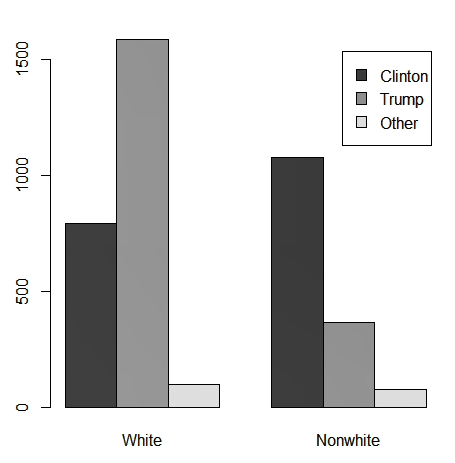
The exit polling from the state of California suggests approximately 50% of whites and 73% of nonwhites voted for Clinton, while 45% of whites and 21% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for California, we again reject the null hypothesis.

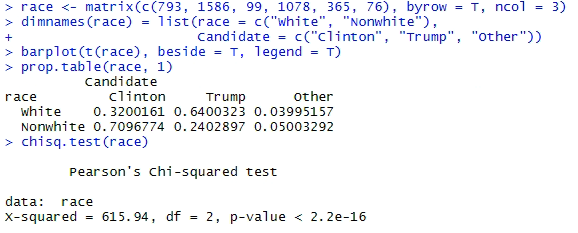
**Florida Gender Voting Statistics**



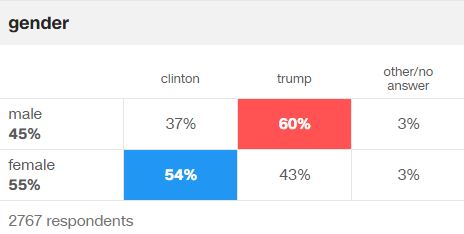
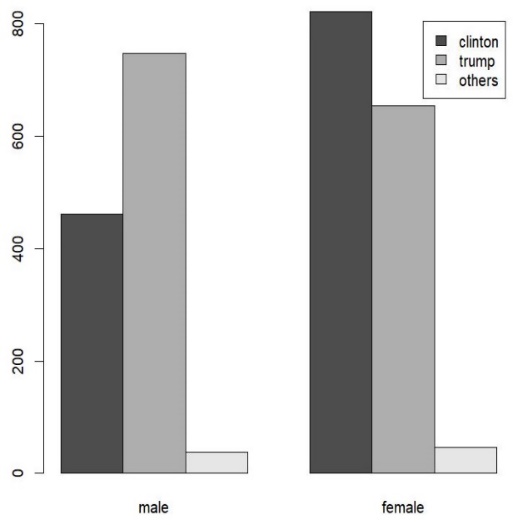
The exit polling from the state of Florida suggests approximately 43% of men and 50% of women voted for Clinton, while 52% of men and 46% of women voted for Trump. Using the Chi-Square test, we received a p-value of 4.581e-5, which is much smaller than alpha. So, for Florida, we reject the null hypothesis.

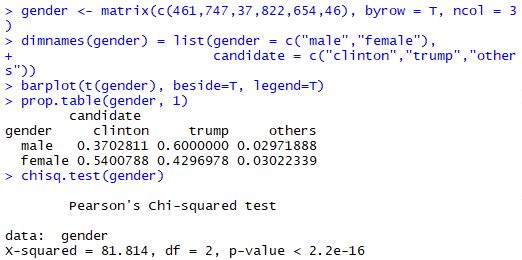
**Florida Racial Voting Statistics**



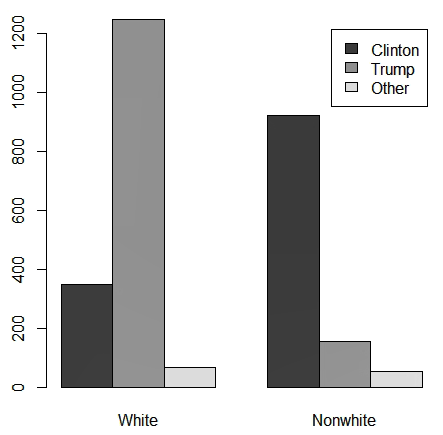
The exit polling from the state of Florida suggests approximately 32% of whites and 71% of nonwhites voted for Clinton, while 64% of whites and 24% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Florida, we again reject the null hypothesis.

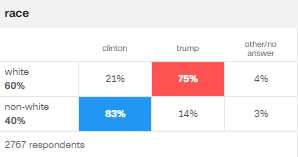
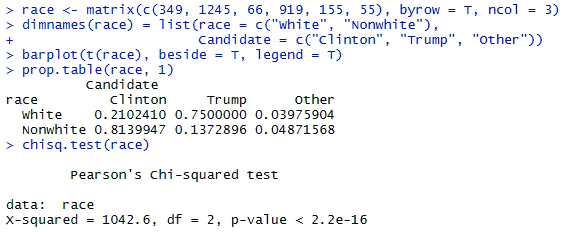
**Georgia Gender Voting Statistics**



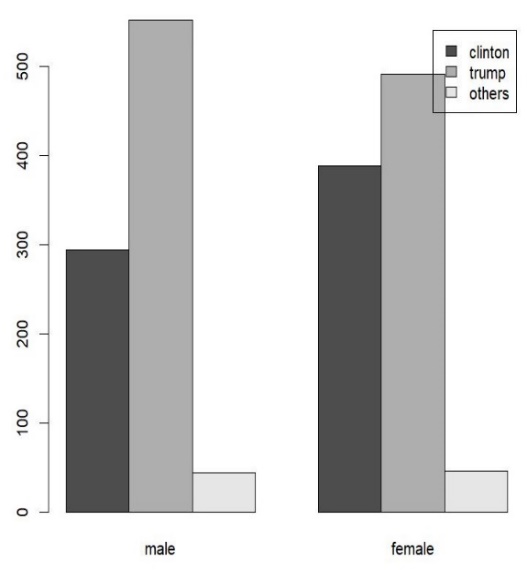


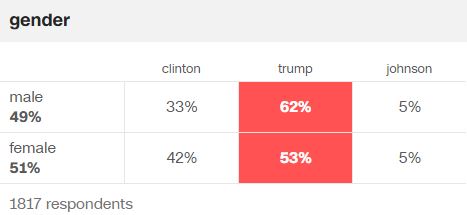
The exit polling from the state of Georgia suggests approximately 37% of men and 54% of women voted for Clinton, while 60% of men and 43% of women voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Georgia, we reject the null hypothesis.

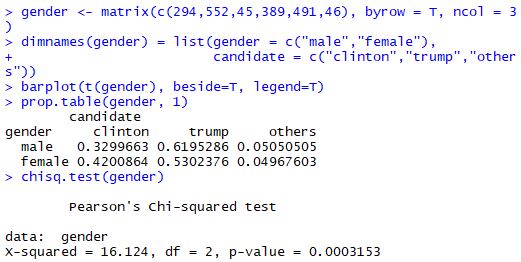
**Georgia Racial Voting Statistics**

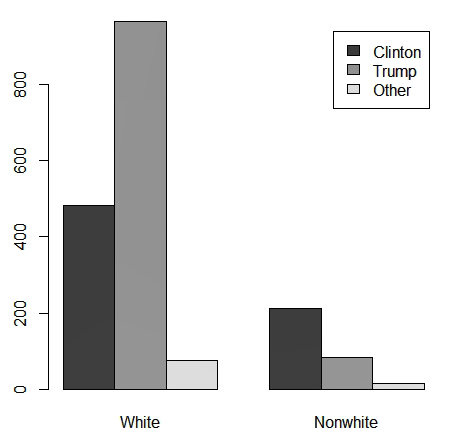


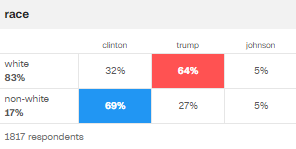
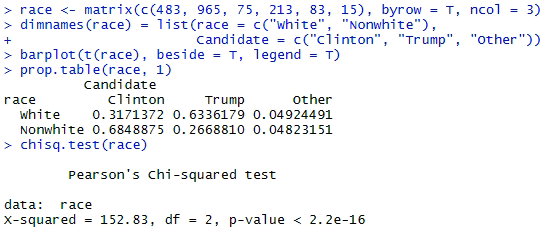
The exit polling from the state of Georgia suggests approximately 21% of whites and 83% of nonwhites voted for Clinton, while 75% of whites and 14% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Georgia, we again reject the null hypothesis.

**Indiana Gender Voting Statistics**

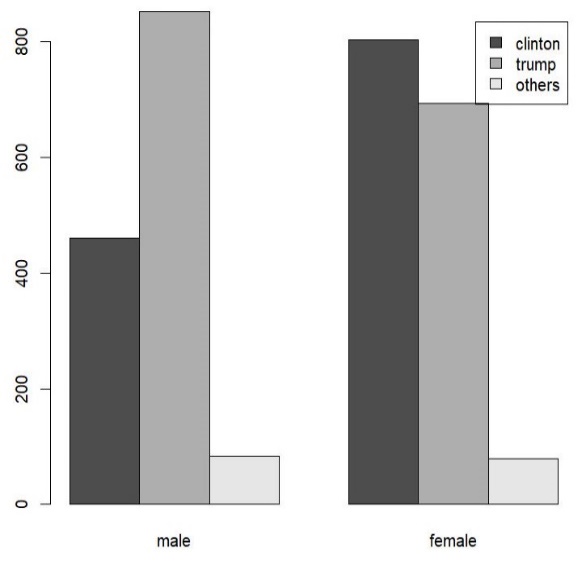


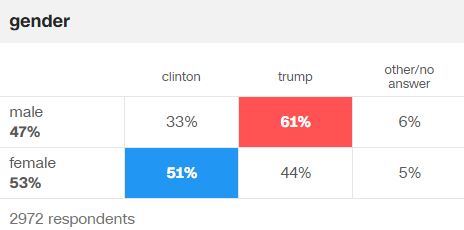
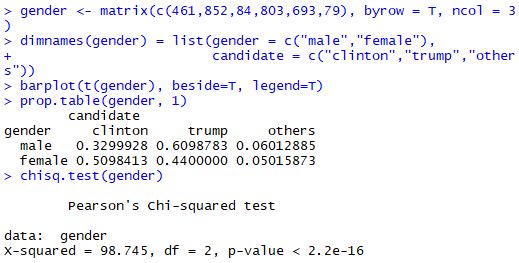
The exit polling from the state of Indiana suggests approximately 33% of men and 42% of women voted for Clinton, while 62% of men and 53% of women voted for Trump. Using the Chi-Square test, we received a p-value of 3.153e-4, which is much smaller than alpha. So, for Indiana, we reject the null hypothesis.

**Indiana Racial Voting Statistics**

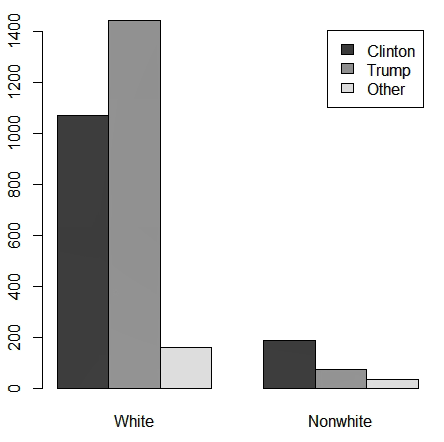


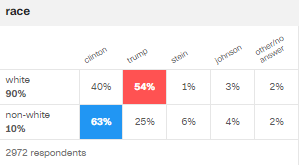
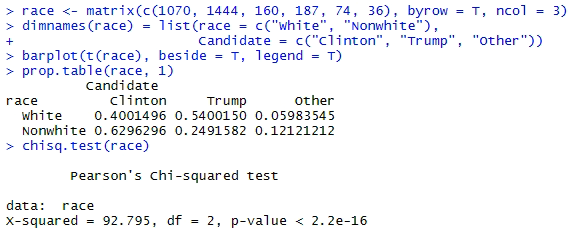
The exit polling from the state of Indiana suggests approximately 32% of whites and 69% of nonwhites voted for Clinton, while 64% of whites and 27% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Indiana, we reject the null hypothesis.

**Iowa Gender Voting Statistics**

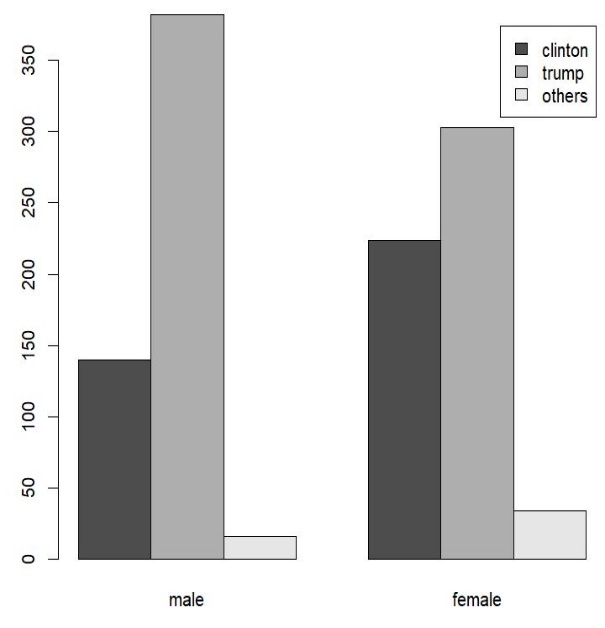


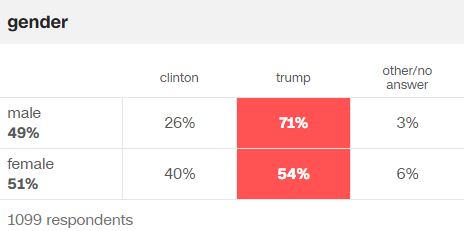
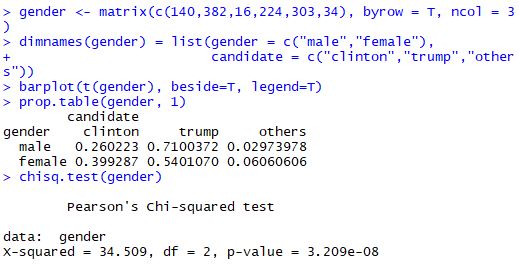
The exit polling from the state of Iowa suggests approximately 33% of men and 51% of women voted for Clinton, while 61% of men and 44% of women voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Iowa, we reject the null hypothesis.

**Iowa Racial Voting Statistics**

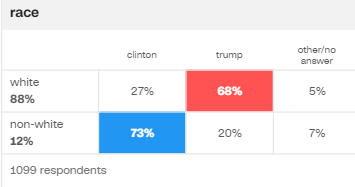
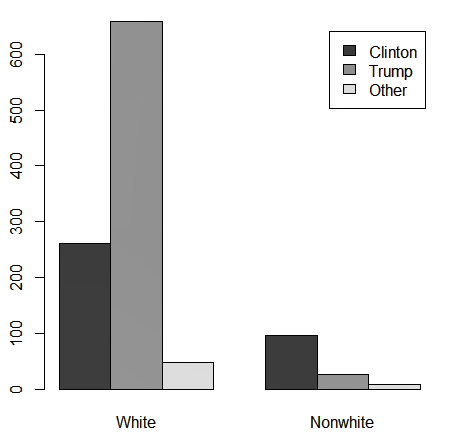


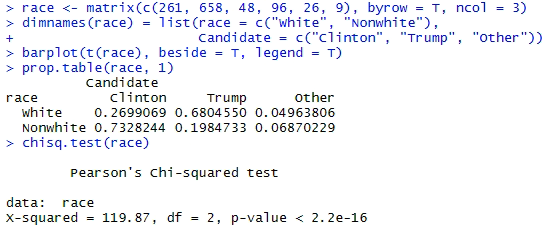
The exit polling from the state of Iowa suggests approximately 40% of whites and 63% of nonwhites voted for Clinton, while 54% of whites and 25% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Iowa, we again reject the null hypothesis.

**Kentucky Gender Voting Statistics**

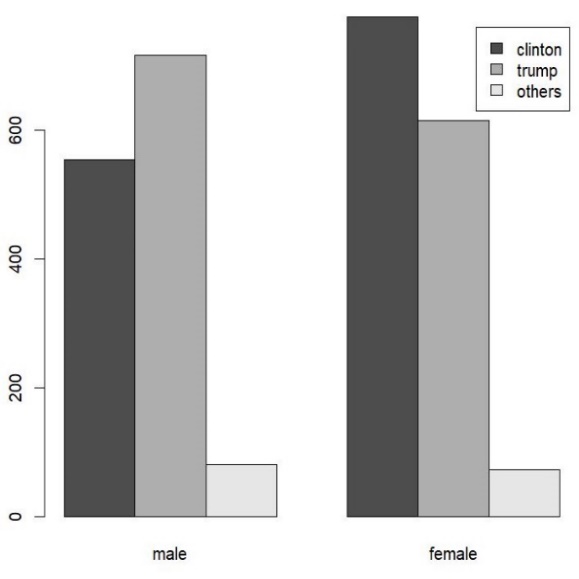


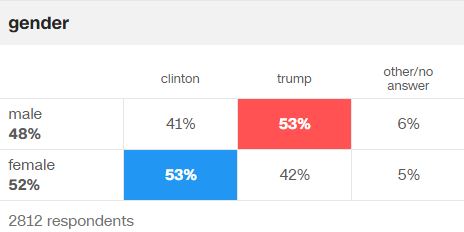
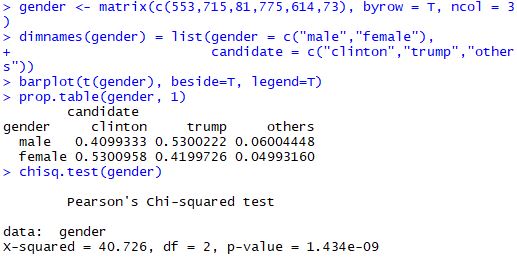
The exit polling from the state of Kentucky suggests approximately 26% of men and 40% of women voted for Clinton, while 71% of men and 54% of women voted for Trump. Using the Chi-Square test, we received a p-value of 3.209e-8, which is much smaller than alpha. So, for Kentucky, we reject the null hypothesis.

**Kentucky Racial Voting Statistics**

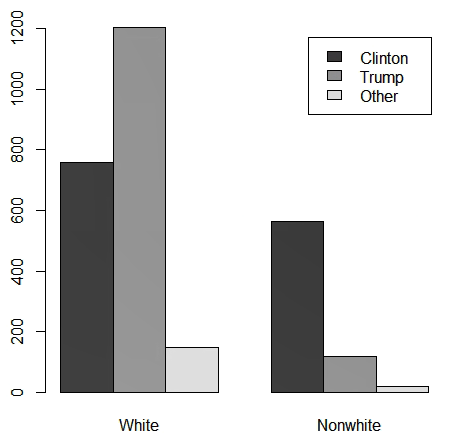


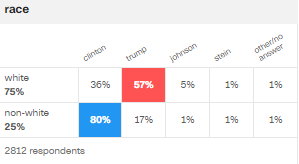
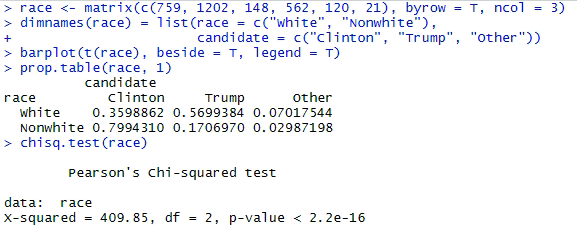
The exit polling from the state of Kentucky suggests approximately 27% of whites and 73% of nonwhites voted for Clinton, while 68% of whites and 20% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Kentucky, we again reject the null hypothesis.

**Michigan Gender Voting Statistics**



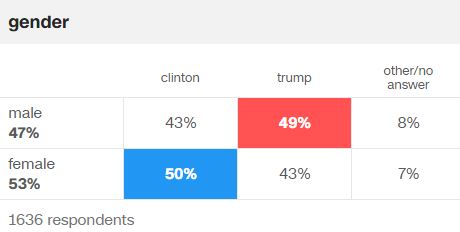
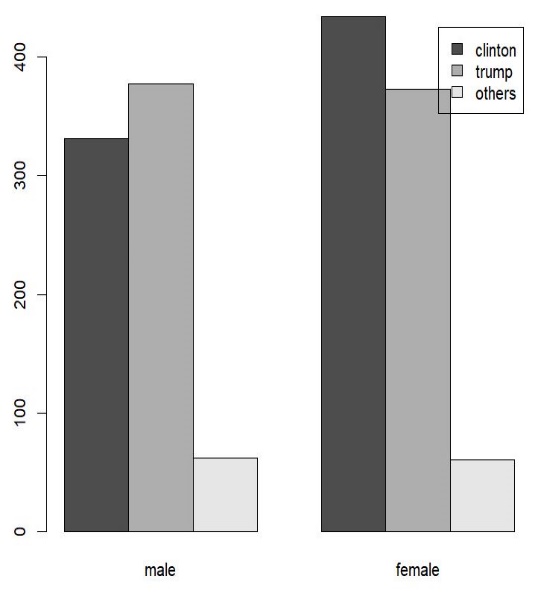
The exit polling from the state of Michigan suggests approximately 41% of men and 53% of women voted for Clinton, while 53% of men and 42% of women voted for Trump. Using the Chi-Square test, we received a p-value of 1.434e-9, which is much smaller than alpha. So, for Michigan, we reject the null hypothesis.

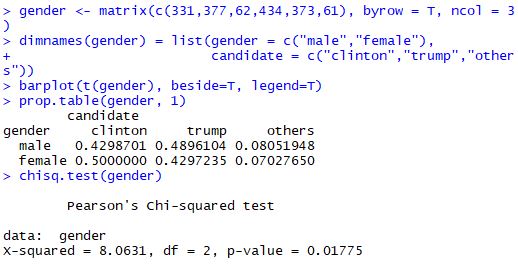
**Michigan Racial Voting Statistics**

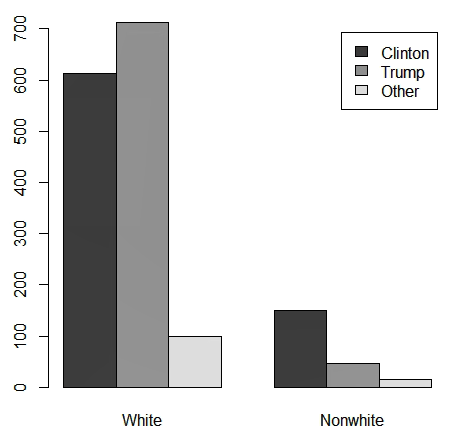


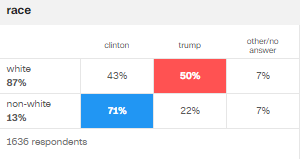
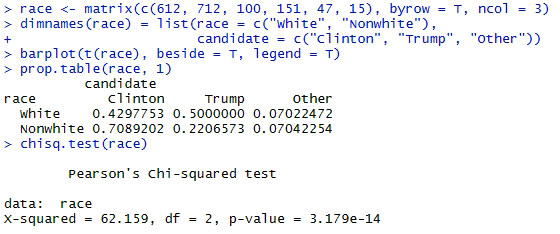
The exit polling from the state of Michigan suggests approximately 36% of whites and 80% of nonwhites voted for Clinton, while 57% of whites and 17% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Michigan, we again reject the null hypothesis.

**Minnesota Gender Voting Statistics**



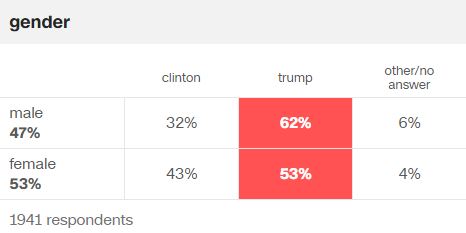
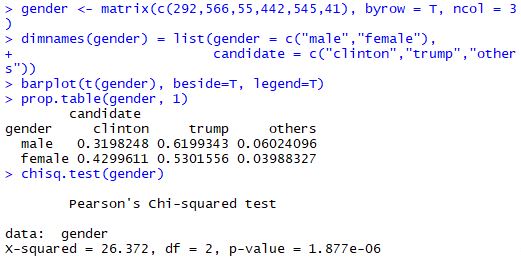
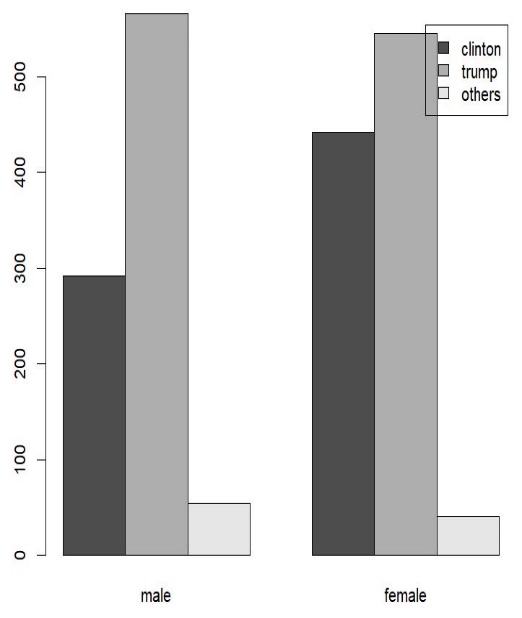
The exit polling from the state of Minnesota suggests approximately 43% of men and 50% of women voted for Clinton, while 49% of men and 43% of women voted for Trump. Using the Chi-Square test, we received a p-value of .01775, which is much smaller than alpha. So, for Minnesota, we reject the null hypothesis.

**Minnesota Racial Voting Statistics**

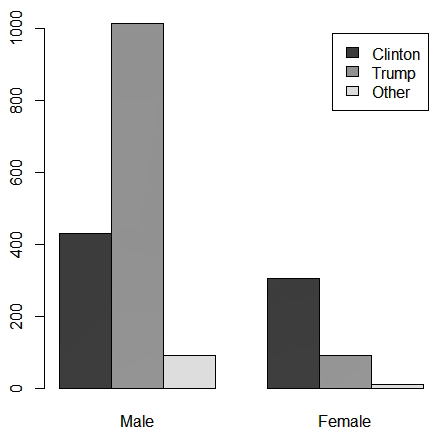


The exit polling from the state of Minnesota suggests approximately 43% of whites and 71% of nonwhites voted for Clinton, while 50% of whites and 22% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value of 3.179e-14, which is much smaller than alpha. So, for Minnesota, we again reject the null hypothesis.

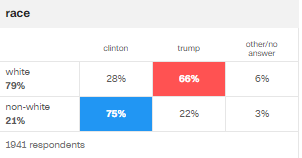
**Missouri Gender Voting Statistics**

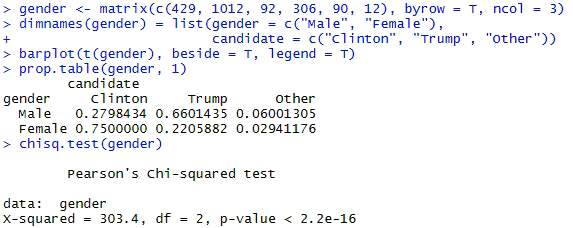


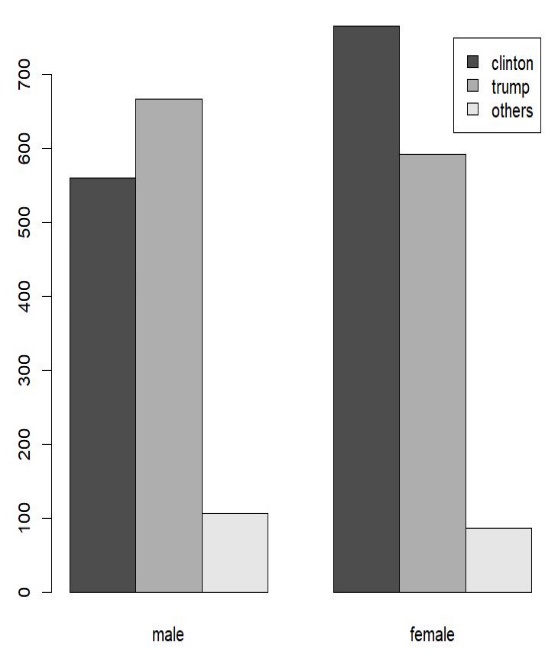
The exit polling from the state of Missouri suggests approximately 32% of men and 43% of women voted for Clinton, while 62% of men and 53% of women voted for Trump. Using the Chi-Square test, we received a p-value of 1.877e-6, which is much smaller than alpha. So, for Missouri, we reject the null hypothesis.



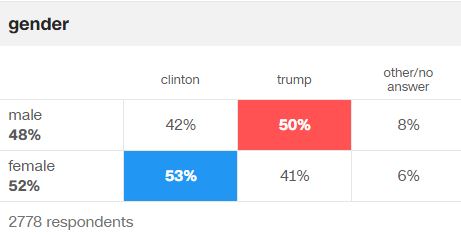
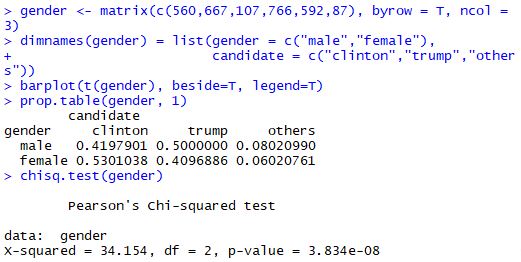
**Missouri Racial Voting Statistics**



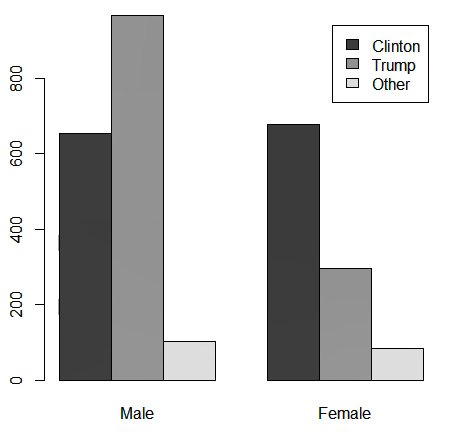
The exit polling from the state of Missouri suggests approximately 28% of whites and 75% of nonwhites voted for Clinton, while 66% of whites and 22% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16, which is much smaller than alpha. So, for Missouri, we again reject the null hypothesis.

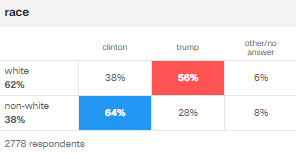
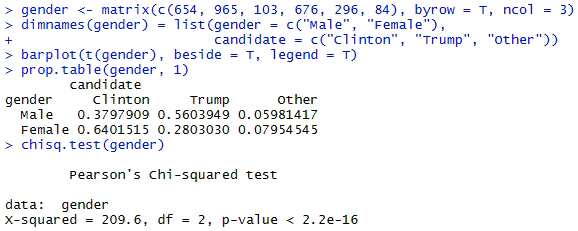


**Nevada Gender Voting Statistics**



The exit polling from the state of Nevada suggests approximately 42% of men and 53% of women voted for Clinton, while 50% of men and 41% of women voted for Trump. Using the Chi-Square test, we received a p-value of 3.834e-8, which is much smaller than alpha. So, for Nevada, we reject the null hypothesis.

**Nevada Racial Voting Statistics**



The exit polling from the state of Nevada suggests approximately 38% of whites and 64% of nonwhites voted for Clinton, while 56% of whites and 28% of nonwhites voted for Trump. Using the Chi-Square test, we received a p-value less than 2.2e-16 , which is much smaller than alpha. So, for Nevada, we again reject the null hypothesis.

**Limitations of the study**

The exit poll data only gives limited number of observations. It is a very convenient sampling and not a probabilistic sampling method. So, the information we extract from this is limited to only this study. The findings of a poll data cannot be generalized beyond the 2016 elections but with proper, well represented data this is possible.

**Conclusion and Final Thoughts**

For this project, we collected CNN exit polling data from 10 states from the 2016 presidential race, and conducted the Chi-Squared test on the data to see if there was a difference in the voting patterns of different genders. The tests we conducted from the CNN exit polls of the 2016 election show that there is in fact a difference between the voting patterns of men and women. The p-value for every state we collected data on is less than alpha(0.05), leading us to reject the null hypothesis that states there is no difference between gender. On the whole, men voted more Republican leaning, while women tended to vote more Democratic leaning. Women tended to vote for Clinton by a margin of approximately 10% more than men did.

**References**

<https://www.cnn.com/election/2016/results/exit-polls/>