

# Data Cleaning

David Sutton

2/19/2020

```
##
## Call:
## lm(formula = cbind(house12_vars$housevote10 == 1) ~ race_12 +
##     gender_12 + Age + educ_12 + faminc_12, data = house12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.6416 -0.4472 -0.4037  0.5443  0.6511
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.477072   0.175396   2.720  0.00677 **
## race_12      -0.010873   0.020799  -0.523  0.60136
## gender_12     0.047949   0.047174   1.016  0.30995
## Age          -0.001619   0.002136  -0.758  0.44889
## educ_12      -0.001164   0.016799  -0.069  0.94477
## faminc_12     0.001264   0.000891   1.419  0.15671
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4988 on 463 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.00852,    Adjusted R-squared:  -0.002187
## F-statistic: 0.7957 on 5 and 463 DF,  p-value: 0.5531

##
## Call:
## lm(formula = cbind(house12_vars$housevote10 == 1) ~ guncontrol12 +
##     climate12 + immg.status12 + immg.borderpatrol12 + immg.policeqs12 +
##     immg.finebiz12 + immg.pubgoods12 + immg.citzdenial12 + abortion12 +
##     jobs.envir12 + gaymarriage12 + affirmact12 + balbudget1.12 +
##     balbudget2.12, data = house12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7279 -0.4477 -0.3000  0.5136  0.7791
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.0014931  0.2796535  -0.005  0.9957
## guncontrol12 -0.0305756  0.0269121  -1.136  0.2565
## climate12    -0.0133712  0.0268146  -0.499  0.6183
```

```

## immg.status12      0.0573803  0.0562189   1.021   0.3080
## immg.borderpatrol12 0.0380340  0.0543630   0.700   0.4845
## immg.policeqs12    0.0717641  0.0574148   1.250   0.2120
## immg.finebiz12     0.0780664  0.0584913   1.335   0.1827
## immg.pubgoods12    -0.1115749  0.0571484  -1.952   0.0515 .
## immg.citzdenial12  0.0962707  0.0607429   1.585   0.1137
## abortion12        0.0131956  0.0266816   0.495   0.6212
## jobs.envir12       0.0105571  0.0214600   0.492   0.6230
## gaymarriage12      -0.0180192  0.0580658  -0.310   0.7565
## affirmact12        0.0254488  0.0317017   0.803   0.4225
## balbudget1.12      0.0570529  0.0331561   1.721   0.0860 .
## balbudget2.12      -0.0007952  0.0337554  -0.024   0.9812
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4969 on 438 degrees of freedom
## (17 observations deleted due to missingness)
## Multiple R-squared:  0.03779, Adjusted R-squared:  0.007032
## F-statistic: 1.229 on 14 and 438 DF, p-value: 0.2508

##
## Call:
## lm(formula = cbind(house12_vars$housevote10 == 1) ~ docvisit12 +
##     newchild12 + moveoutchild12 + newmarriage12 + newdivorce12,
##     data = house12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.6029 -0.4630 -0.3561  0.5370  0.6439
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.909525   0.624854   1.456   0.146
## docvisit12    -0.106840   0.077651  -1.376   0.170
## newchild12     0.070143   0.101007   0.694   0.488
## moveoutchild12 0.008985   0.093148   0.096   0.923
## newmarriage12 -0.109034   0.227882  -0.478   0.633
## newdivorce12  -0.139948   0.194752  -0.719   0.473
##
## Residual standard error: 0.4991 on 464 degrees of freedom
## Multiple R-squared:  0.006607, Adjusted R-squared: -0.004098
## F-statistic: 0.6172 on 5 and 464 DF, p-value: 0.6868

##
## Call:
## lm(formula = cbind(house12_vars$housevote10 == 1) ~ AAraceresentA12 +
##     AAraceresentB12 + Hispraceresent12, data = house12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5672 -0.4477 -0.4149  0.5471  0.6553
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)

```

```

## (Intercept)      0.589536    0.120297    4.901 0.00000132 ***
## AAraceresentA12   0.006056    0.034841    0.174    0.862
## AAraceresentB12  -0.032015    0.021791   -1.469    0.142
## Hispraceresent12 -0.020584    0.033175   -0.620    0.535
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4981 on 464 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.006341, Adjusted R-squared:  -8.333e-05
## F-statistic: 0.987 on 3 and 464 DF, p-value: 0.3986

##
## Call:
## lm(formula = cbind(pres12_vars$presvote12 == 2) ~ race_12 + gender_12 +
##     Age + educ_12 + faminc_12, data = pres12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7191 -0.5548  0.3244  0.3978  0.5193
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.67930728  0.20068839   3.385 0.000805 ***
## race_12      -0.00503514  0.02224360  -0.226 0.821069
## gender_12    -0.11690062  0.05648317  -2.070 0.039321 *
## Age           0.00207184  0.00235082   0.881 0.378831
## educ_12      -0.00302322  0.01995892  -0.151 0.879703
## faminc_12    -0.00001474  0.00101060  -0.015 0.988373
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4886 on 307 degrees of freedom
## Multiple R-squared:  0.01828, Adjusted R-squared:  0.00229
## F-statistic: 1.143 on 5 and 307 DF, p-value: 0.3375

##
## Call:
## lm(formula = cbind(pres12_vars$presvote12 == 2) ~ guncontrol12 +
##     climate12 + immg.status12 + immg.borderpatrol12 + immg.policeqs12 +
##     immg.finebiz12 + immg.pubgoods12 + immg.citzdenial12 + abortion12 +
##     jobs.envir12 + gaymarriage12 + affirmact12 + balbudget1.12 +
##     balbudget2.12, data = pres12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.97216 -0.39369  0.08641  0.38107  0.73181
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.417563    0.306012   1.365    0.1735
## guncontrol12  0.026292    0.028737   0.915    0.3610
## climate12     0.113367    0.027125   4.179 0.0000389 ***
## immg.status12  0.030626    0.060366   0.507    0.6123

```

```

## immg.borderpatrol12 -0.041743  0.064266 -0.650  0.5165
## immg.policeqs12     -0.084390  0.061455 -1.373  0.1708
## immg.finebiz12      -0.021028  0.075169 -0.280  0.7799
## immg.pubgoods12     -0.007692  0.058994 -0.130  0.8964
## immg.citzdenial12   -0.037574  0.058547 -0.642  0.5215
## abortion12         -0.047152  0.028553 -1.651  0.0998 .
## jobs.envir12        0.053345  0.023454  2.274  0.0237 *
## gaymarriage12       -0.014623  0.055212 -0.265  0.7913
## affirmact12         0.043730  0.034582  1.265  0.2071
## balbudget1.12       -0.046707  0.037513 -1.245  0.2141
## balbudget2.12       0.006703  0.036386  0.184  0.8540
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.44 on 286 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.2327, Adjusted R-squared:  0.1952
## F-statistic: 6.197 on 14 and 286 DF, p-value: 0.00000000008872

##
## Call:
## lm(formula = cbind(pres12_vars$presvote12 == 2) ~ docvisit12 +
##     newchild12 + moveoutchild12 + newmarriage12 + newdivorce12,
##     data = pres12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7837 -0.6112  0.3888  0.3888  0.4777
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.18045    0.72663   1.625   0.105
## docvisit12     -0.05677    0.08371  -0.678   0.498
## newchild12      0.08887    0.11003   0.808   0.420
## moveoutchild12 -0.09300    0.11688  -0.796   0.427
## newmarriage12  -0.22926    0.22353  -1.026   0.306
## newdivorce12   -0.02284    0.22518  -0.101   0.919
##
## Residual standard error: 0.4913 on 307 degrees of freedom
## Multiple R-squared:  0.007528, Adjusted R-squared: -0.008636
## F-statistic: 0.4657 on 5 and 307 DF, p-value: 0.8017

##
## Call:
## lm(formula = cbind(pres12_vars$presvote12 == 2) ~ AAraceresentA12 +
##     AAraceresentB12 + Hispraceresent12, data = pres12_vars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7428 -0.5214  0.2572  0.3760  0.7865
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.64755    0.13008   4.978 0.00000107 ***

```

```

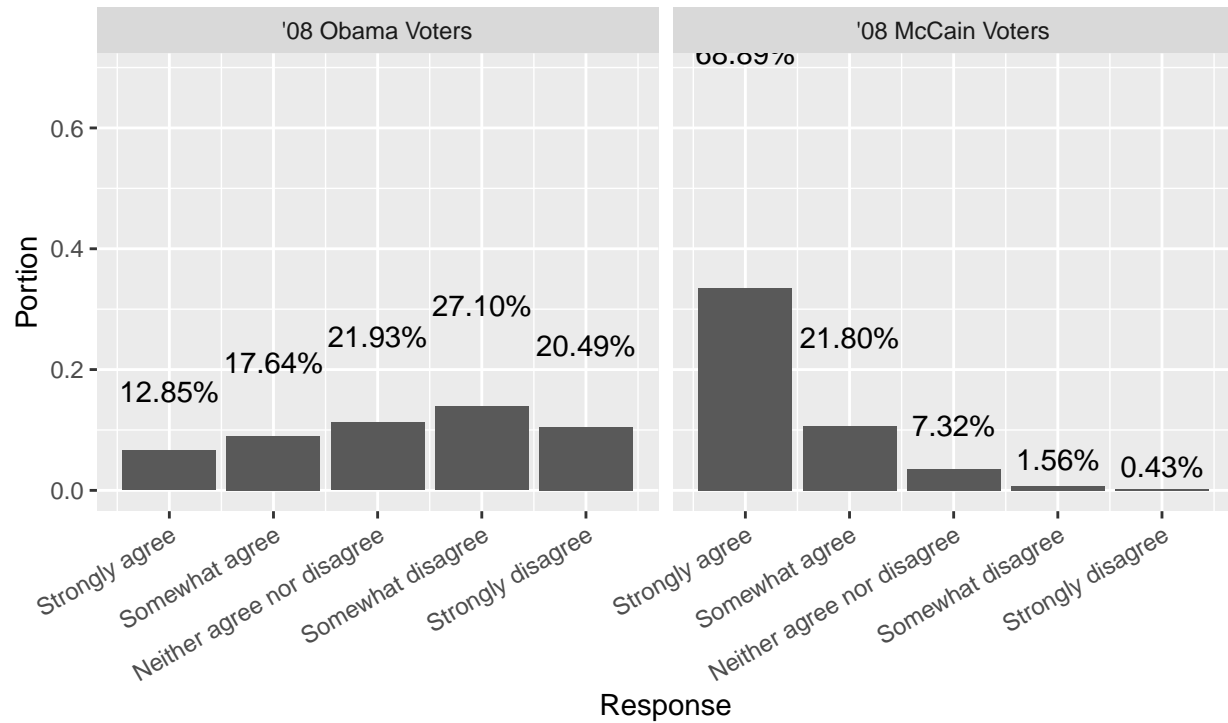
## AAraceresentA12 -0.02345    0.04427   -0.530    0.5966
## AAraceresentB12  0.03958    0.02424    1.633    0.1035
## Hispraceresent12 -0.07920    0.04570   -1.733    0.0841 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4742 on 308 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.07055,    Adjusted R-squared:  0.06149
## F-statistic: 7.792 on 3 and 308 DF,  p-value: 0.00004967

##
## Call:
## lm(formula = cbind(CCES_Panel_Full3waves_VV_V4$CC10_317 == 2) ~
##     CC10_422a + CC10_422b, data = CCES_Panel_Full3waves_VV_V4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8039 -0.3311  0.1416  0.1961  1.1416
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.440054   0.021665  20.31 <0.0000000000000002 ***
## CC10_422a   -0.136342   0.004295 -31.75 <0.0000000000000002 ***
## CC10_422b    0.100032   0.003990  25.07 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3944 on 8859 degrees of freedom
## (638 observations deleted due to missingness)
## Multiple R-squared:  0.376,    Adjusted R-squared:  0.3759
## F-statistic: 2669 on 2 and 8859 DF,  p-value: < 0.00000000000000022

```

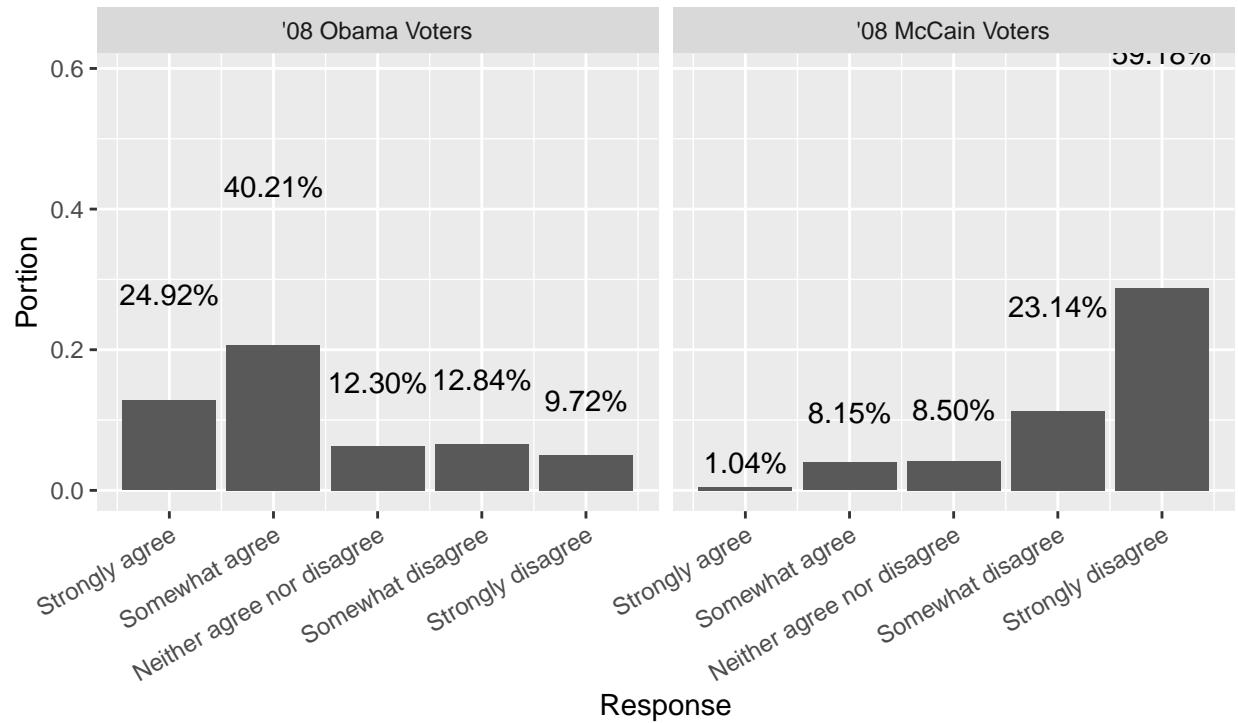
## Racial Resentment of Blacks – 1

The Irish, Italians, Jews and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors.



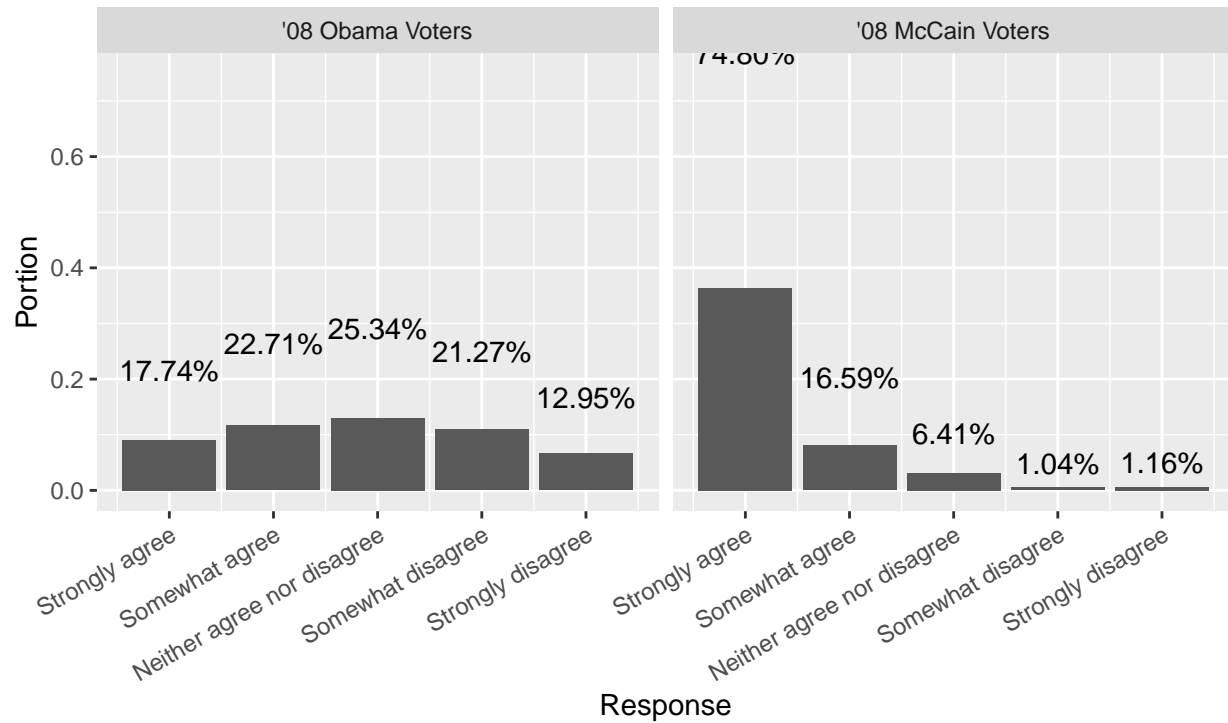
## Racial Resentment of Blacks – 2

Generations of slavery and discrimination have created conditions that make it difficult for to work their way out of the lower class.



## Racial Resentment of Hispanics – 1

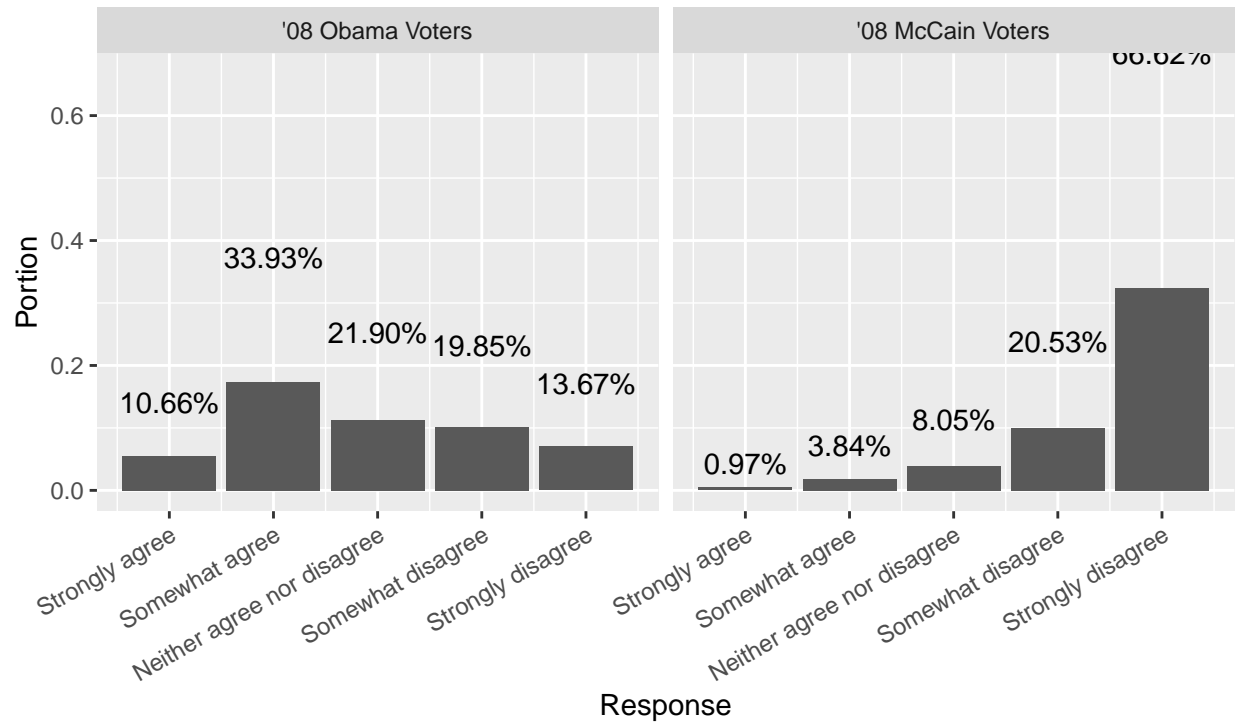
The Irish, Italians, Jews and many other minorities overcame prejudice and worked their way up. Hispanics should do the same without any special favors.



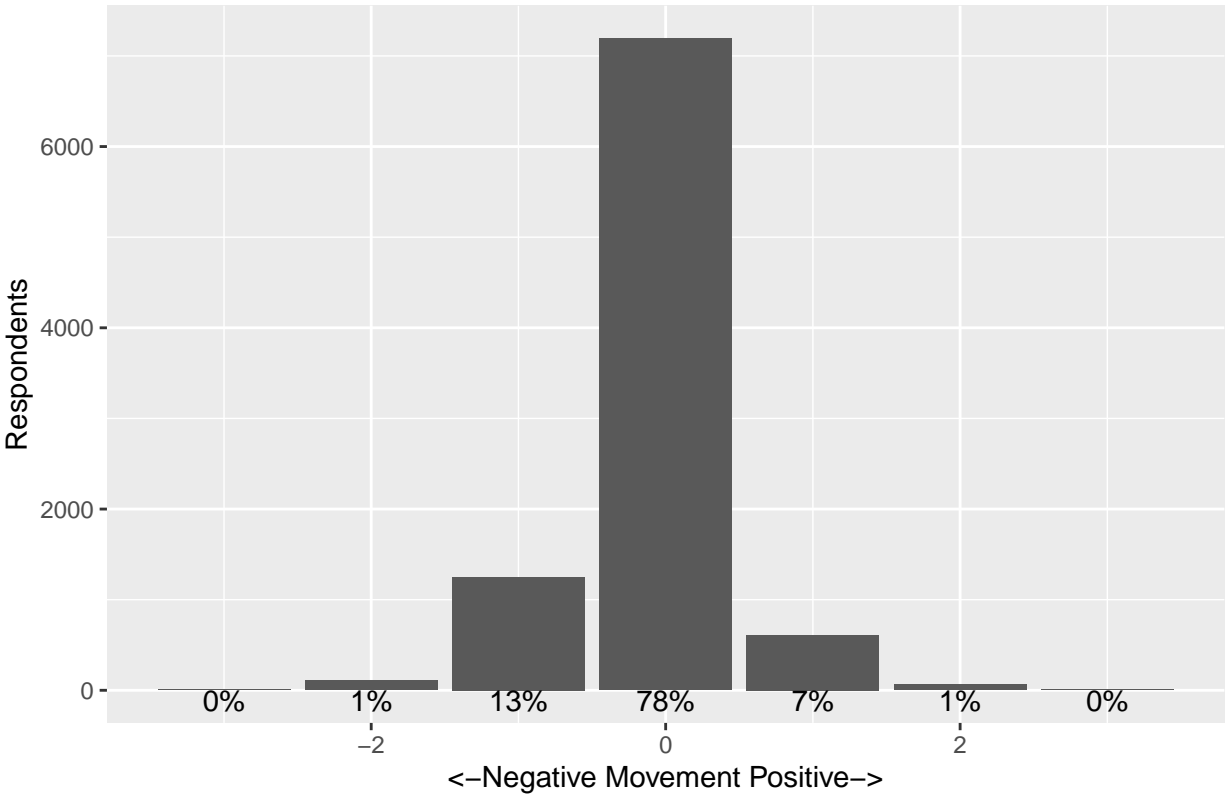


## Racial Resentment of Hispanics – 2

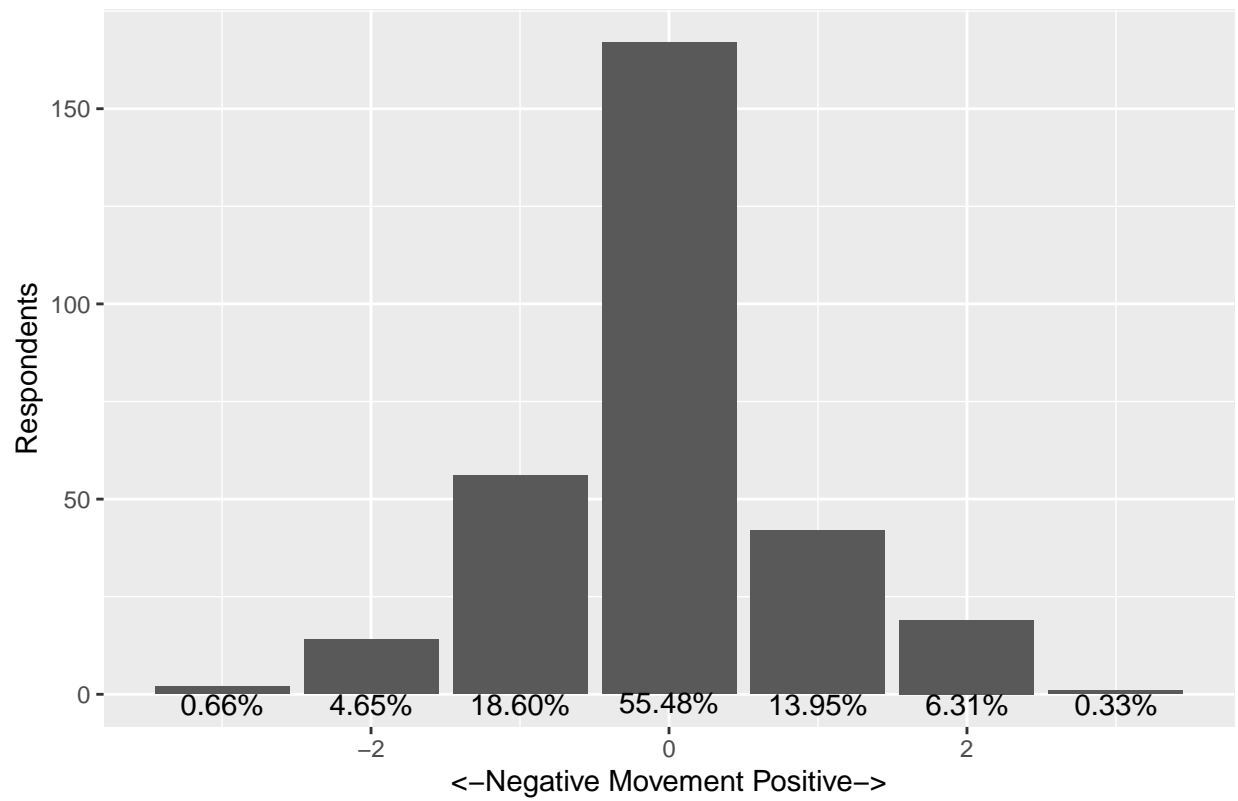
Generations of slavery and discrimination have created conditions that make it difficult for Hispanics to work their way out of the lower class.



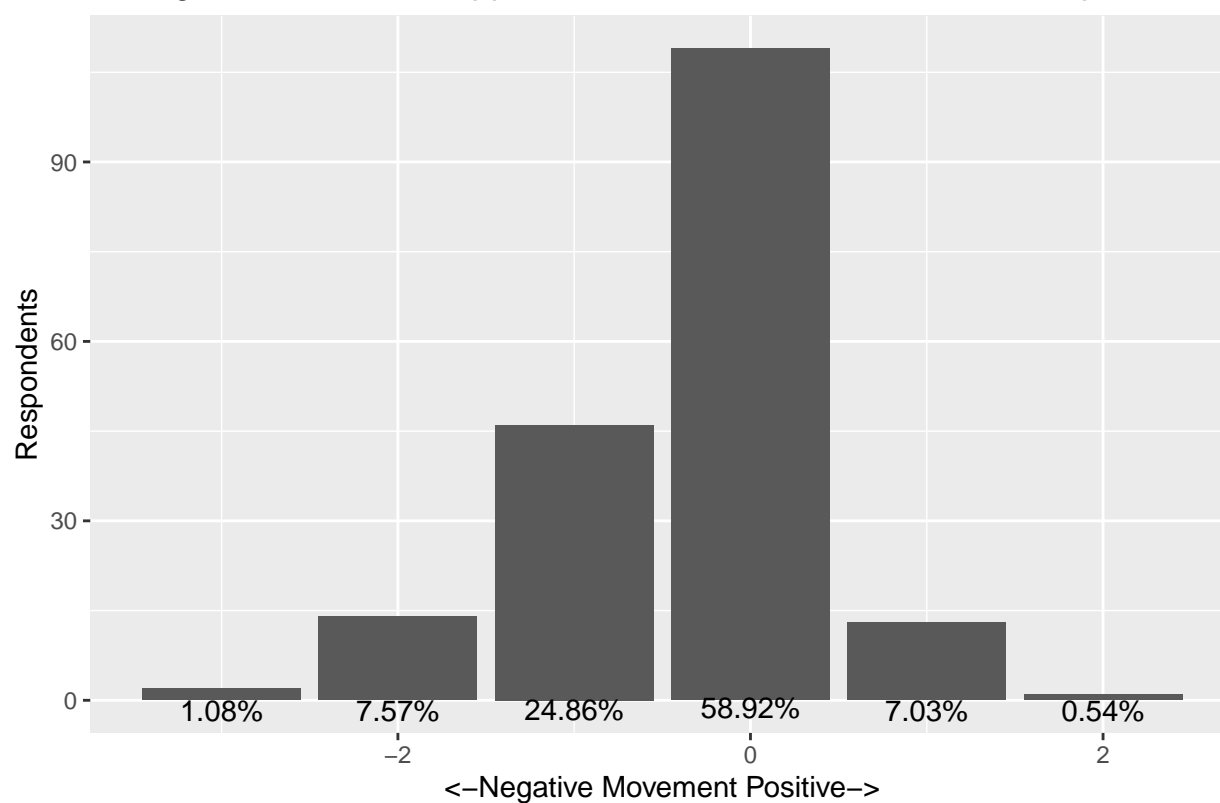
Change in Presidential Approval from 2010 to 2014: General Public



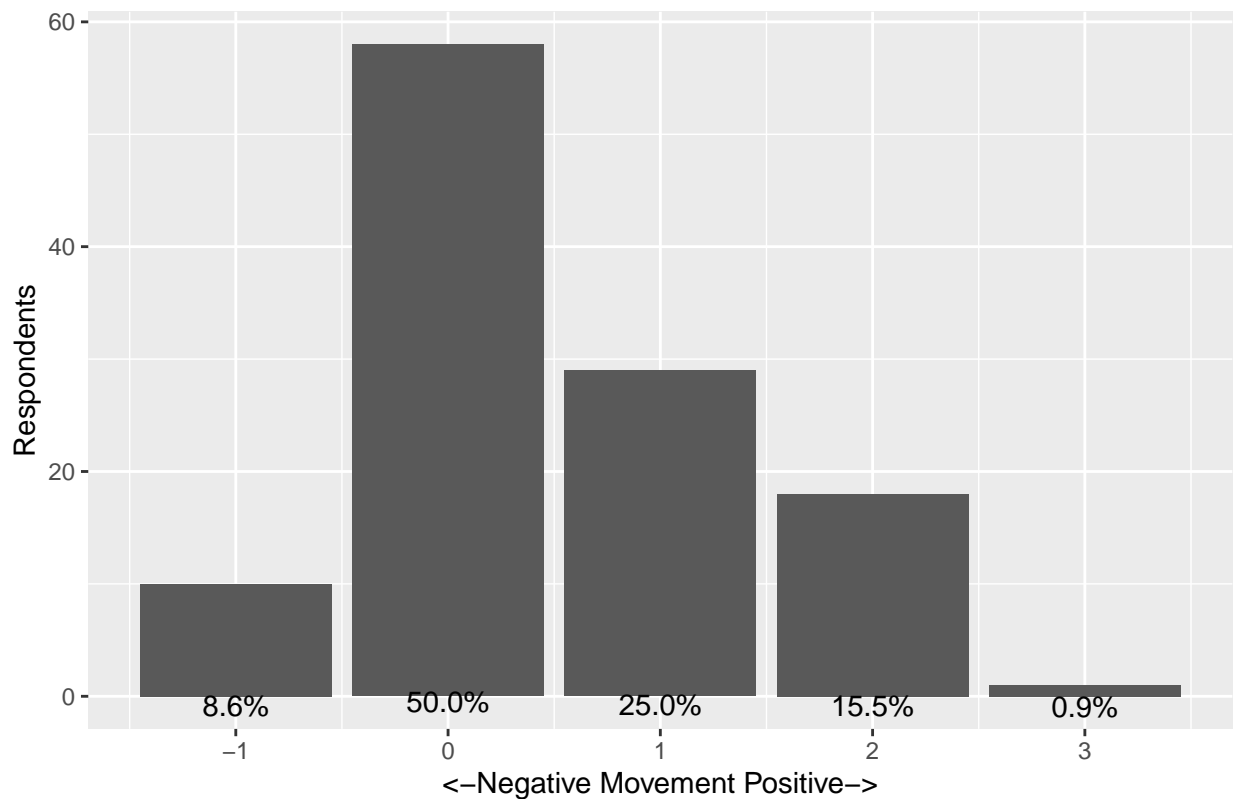
Change in Presidential Approval from 2010 to 2014: Switchers



Change in Presidential Approval from 2010 to 2014: Dem-to-Rep Switchers



## Change in Presidential Approval from 2010 to 2014: Rep-to-Dem Switchers



## Milestone 5

```
##
## Call:
## lm(formula = cbind(CCES_Panel.vars$pres.switch == 1) ~ race_12 +
##   gender_12 + Age_12 + educ_12 + faminc_12 + guncontrol12 +
##   climate12 + immg.status12 + immg.borderpatrol12 + immg.policeqs12 +
##   immg.finebiz12 + immg.pubgoods12 + immg.citzdenial12 + abortion12 +
##   jobs.envir12 + gaymarriage12 + affirmact12 + balbudget1.12 +
##   balbudget2.12 + docvisit12 + newchild12 + moveoutchild12 +
##   newmarriage12 + newdivorce12 + AAraceresentA12 + AAraceresentB12 +
##   Hispraceresent12, data = CCES_Panel.vars)
##
## Coefficients:
##      (Intercept)          race_12          gender_12
##      0.10285988      0.00331345      0.00600541
##           Age_12           educ_12           faminc_12
##      -0.00047997     -0.00250672     -0.00002552
##   guncontrol12      climate12      immg.status12
##      0.00401248     -0.01778270      0.00399880
## immg.borderpatrol12 immg.policeqs12 immg.finebiz12
##      -0.00063371     -0.01619733     -0.01115589
```

```

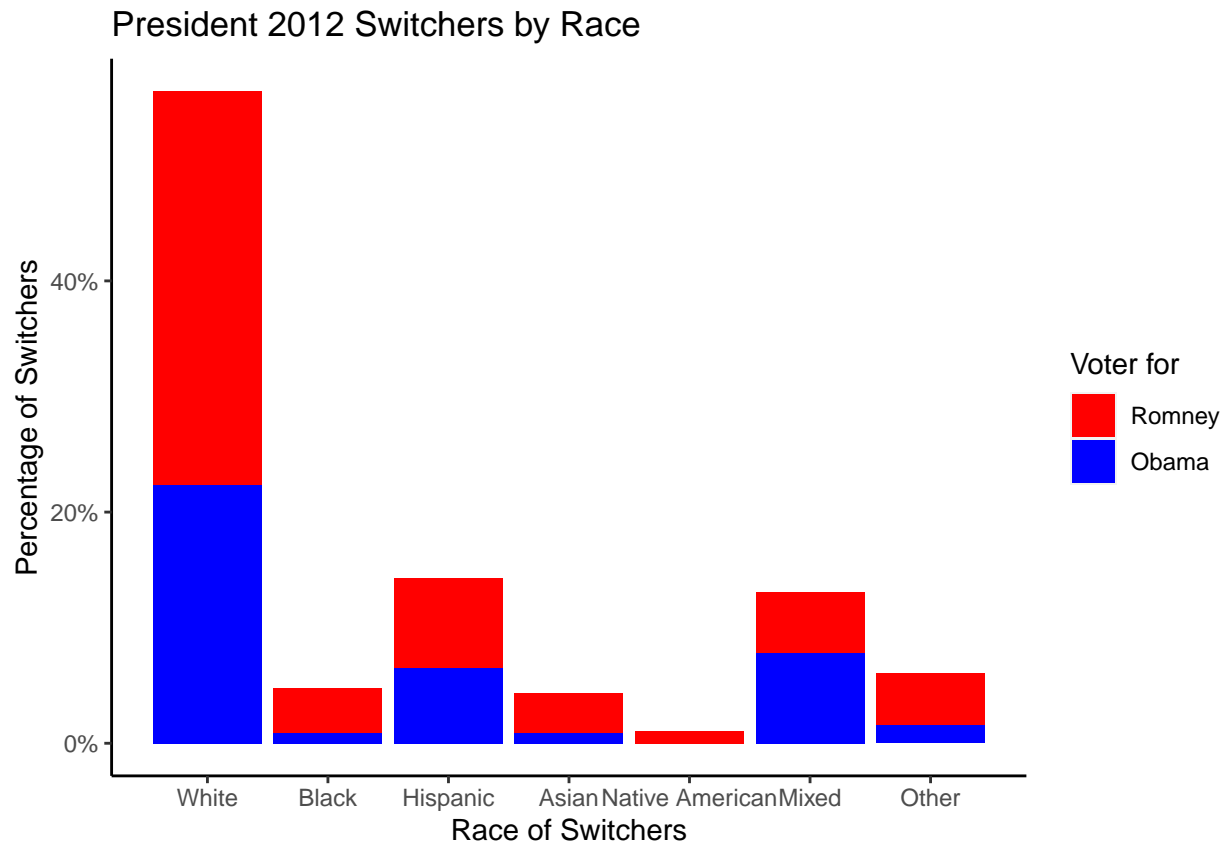
##      immg.pubgoods12      immg.citzdenial12      abortion12
##      -0.00012914          0.01216344          0.00259913
##      jobs.envir12        gaymarriage12        affirmact12
##      0.00554541          -0.00309504          0.00241928
##      balbudget1.12       balbudget2.12        docvisit12
##      -0.00197432         0.00092926          0.00803172
##      newchild12          moveoutchild12        newmarriage12
##      -0.00824828         0.00365670          -0.00160469
##      newdivorce12        AAraceresentA12      AAraceresentB12
##      0.00364126          -0.00792498          0.00079688
##      Hispraceresent12
##      -0.00168937

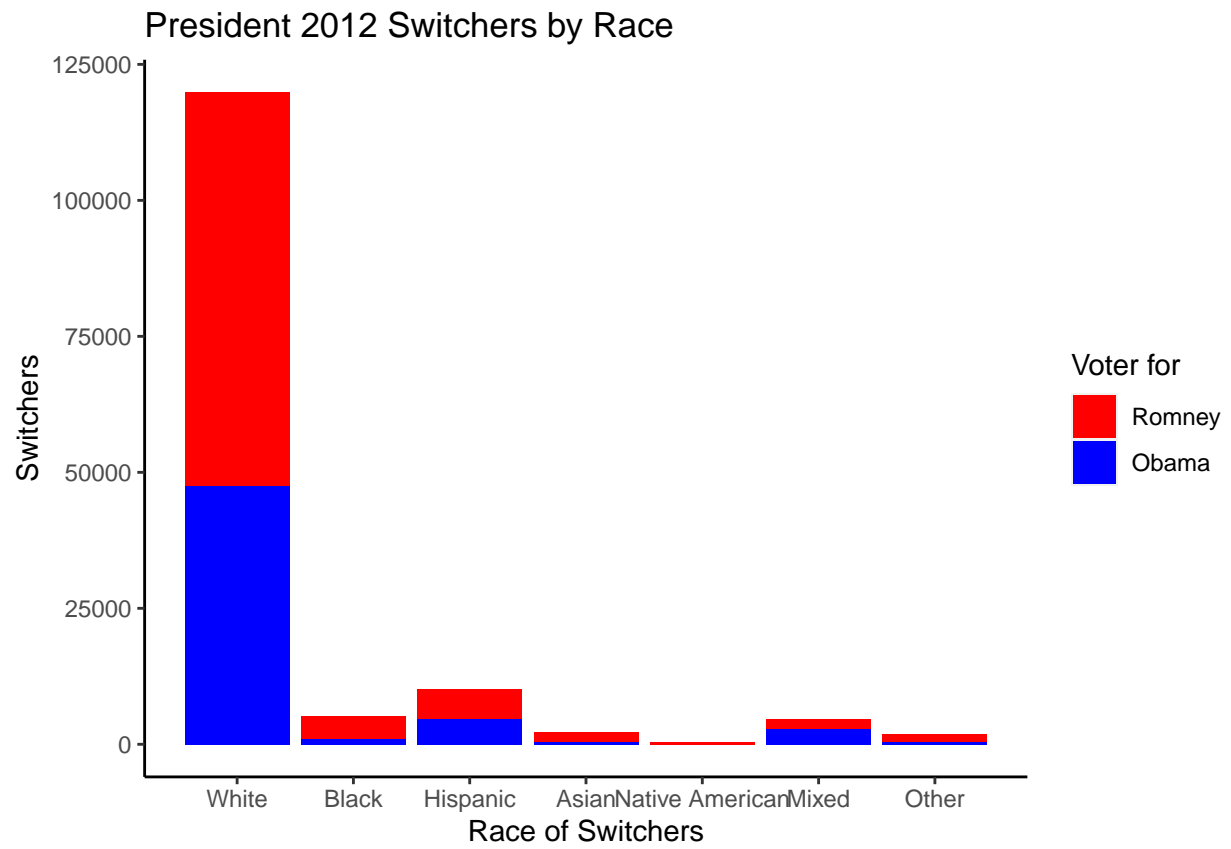
##
## Call:
## lm(formula = cbind(CCES_Panel.vars$presvote12 == 2 & CCES_Panel.vars$presvote08 ==
## 1) ~ race_12 + gender_12 + Age.12 + educ_12 + faminc_12 +
## guncontrol12 + climate12 + immg.status12 + immg.borderpatrol12 +
## immg.policeqs12 + immg.finebiz12 + immg.pubgoods12 + immg.citzdenial12 +
## abortion12 + jobs.envir12 + gaymarriage12 + affirmact12 +
## balbudget1.12 + balbudget2.12 + docvisit12 + newchild12 +
## moveoutchild12 + newmarriage12 + newdivorce12 + AAraceresentA12 +
## AAraceresentB12 + Hispraceresent12, data = CCES_Panel.vars)
##
## Coefficients:
##      (Intercept)          race_12          gender_12
##      0.05727681          0.00159138          0.00048551
##      Age.12            educ_12            faminc_12
##      -0.00017224        -0.00066496        -0.00002317
##      guncontrol12       climate12          immg.status12
##      0.00286797         -0.00601014          0.00192058
## immg.borderpatrol12    immg.policeqs12    immg.finebiz12
##      -0.00106909        -0.01489068        -0.00781855
## immg.pubgoods12       immg.citzdenial12    abortion12
##      -0.00027128         0.01084663          0.00180336
##      jobs.envir12       gaymarriage12      affirmact12
##      0.00468359          0.00038447          0.00013991
##      balbudget1.12      balbudget2.12      docvisit12
##      -0.00395711         0.00040308          0.00296617
##      newchild12         moveoutchild12     newmarriage12
##      -0.00041311        -0.00048618        -0.00957048
##      newdivorce12       AAraceresentA12    AAraceresentB12
##      0.00594179         -0.00316024          0.00038022
##      Hispraceresent12
##      -0.00220480

##
## Call:
## lm(formula = cbind(CCES_Panel.vars$presvote12 == 1 & CCES_Panel.vars$presvote08 ==
## 2) ~ race_12 + gender_12 + Age.12 + educ_12 + faminc_12 +
## guncontrol12 + climate12 + immg.status12 + immg.borderpatrol12 +
## immg.policeqs12 + immg.finebiz12 + immg.pubgoods12 + immg.citzdenial12 +
## abortion12 + jobs.envir12 + gaymarriage12 + affirmact12 +
## balbudget1.12 + balbudget2.12 + docvisit12 + newchild12 +

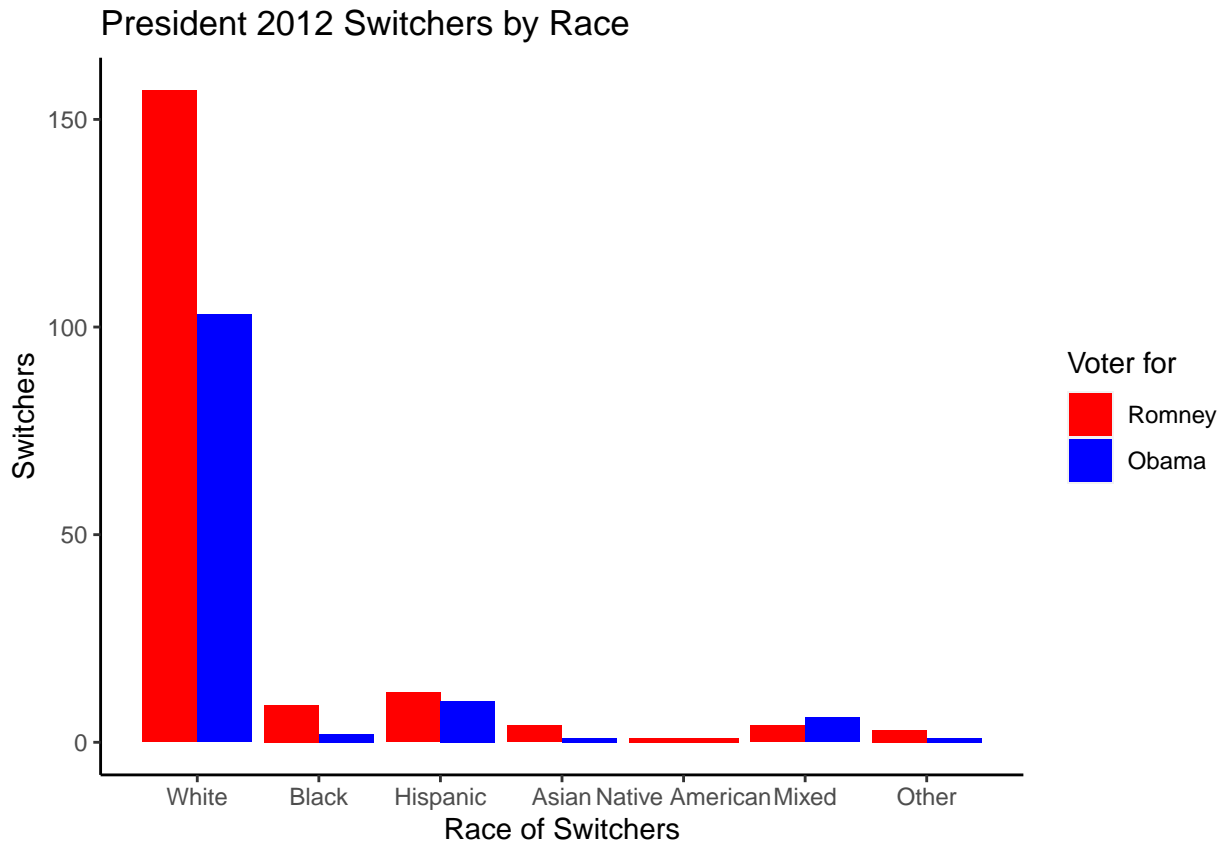
```

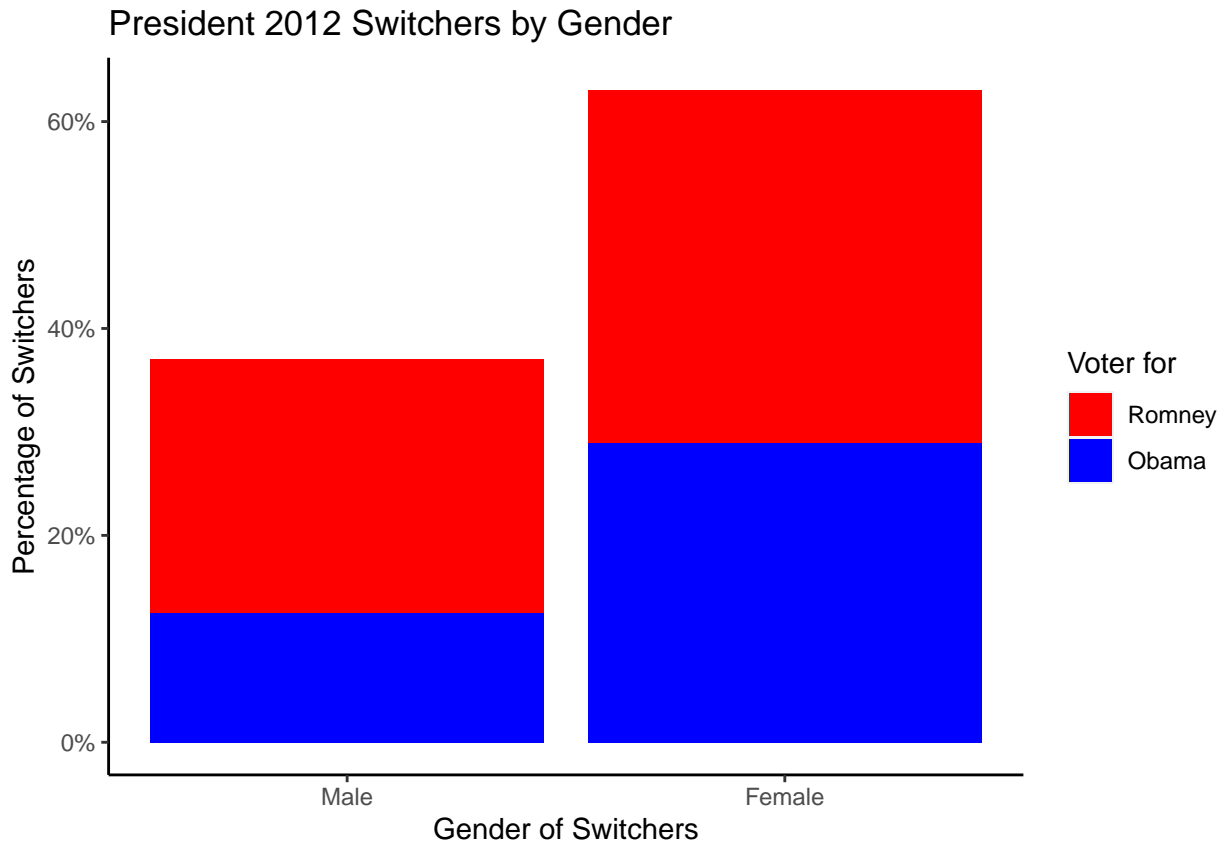
```
##      moveoutchild12 + newmarriage12 + newdivorce12 + AAraceresentA12 +
##      AAraceresentB12 + Hispraceresent12, data = CCES_Panel.vars)
##
## Coefficients:
##      (Intercept)           race_12           gender_12
##      0.039921807         0.001450671         0.004333229
##      Age_12             educ_12             faminc_12
##      -0.000244067        -0.001412600         0.000001886
##      guncontrol12        climate12          immg.status12
##      0.001107988         -0.010496125         0.001876854
## immg.borderpatrol12      immg.policeqs12      immg.finebiz12
##      0.000643220         -0.000825109        -0.003505932
##      immg.pubgoods12     immg.citzdenial12     abortion12
##      0.000074885         0.001277439         0.000822181
##      jobs.envir12        gaymarriage12        affirmact12
##      0.000465327         -0.003558943         0.002276884
##      balbudget1.12       balbudget2.12       docvisit12
##      0.002254485         0.000326017         0.004726600
##      newchild12          moveoutchild12        newmarriage12
##      -0.007436042         0.003603333         0.007742221
##      newdivorce12        AAraceresentA12      AAraceresentB12
##      -0.002438655        -0.004425435         0.000133923
##      Hispraceresent12
##      0.000660610
```



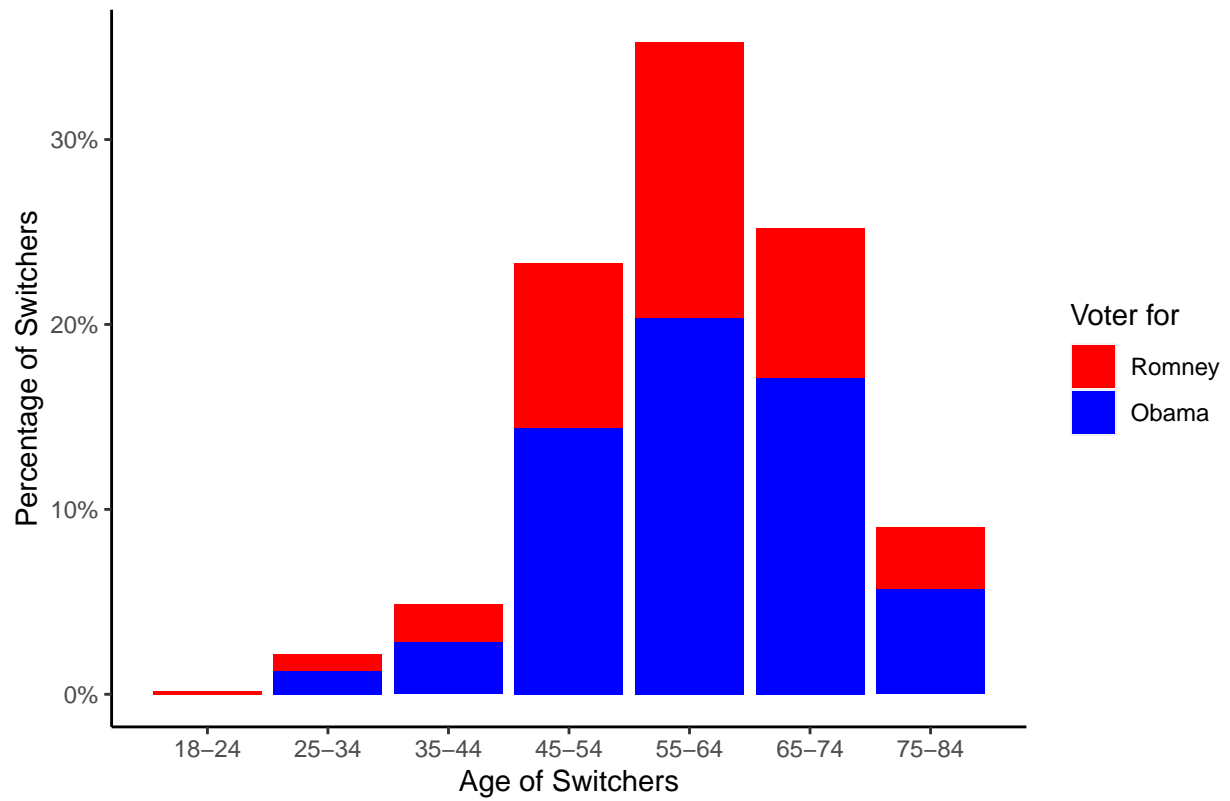




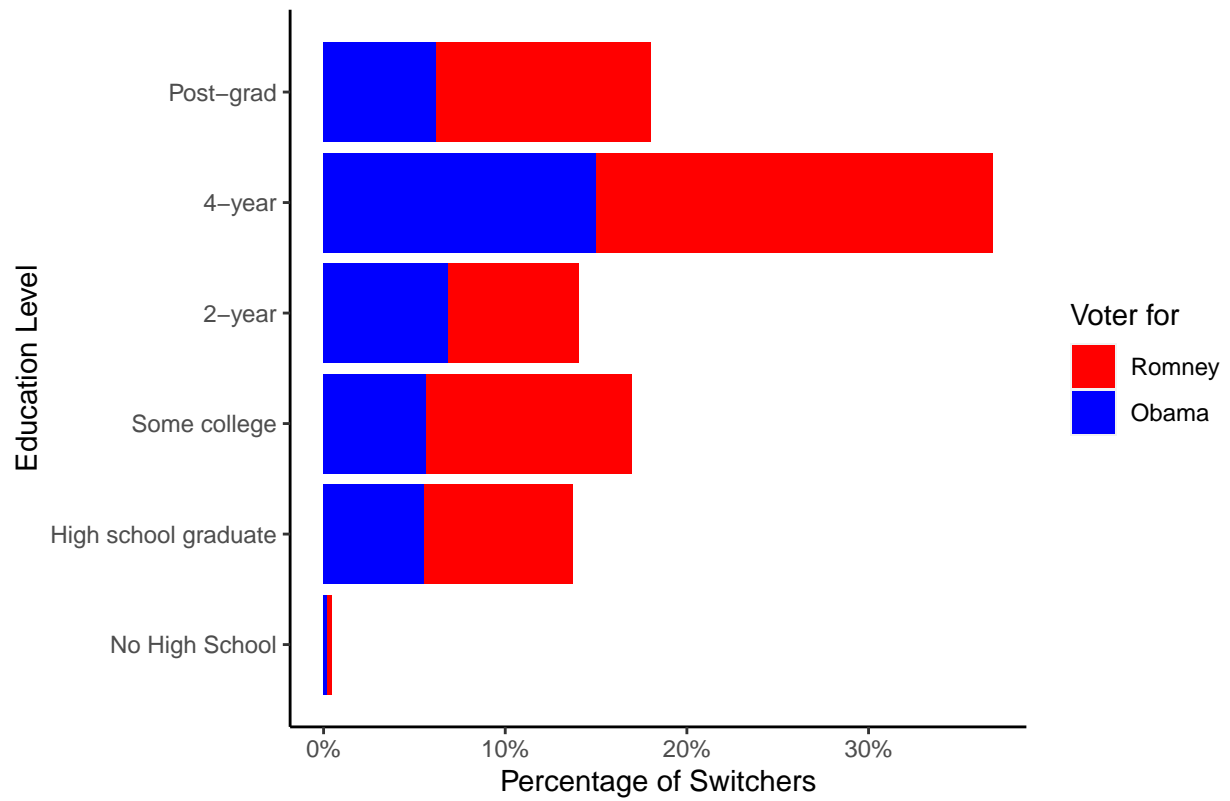




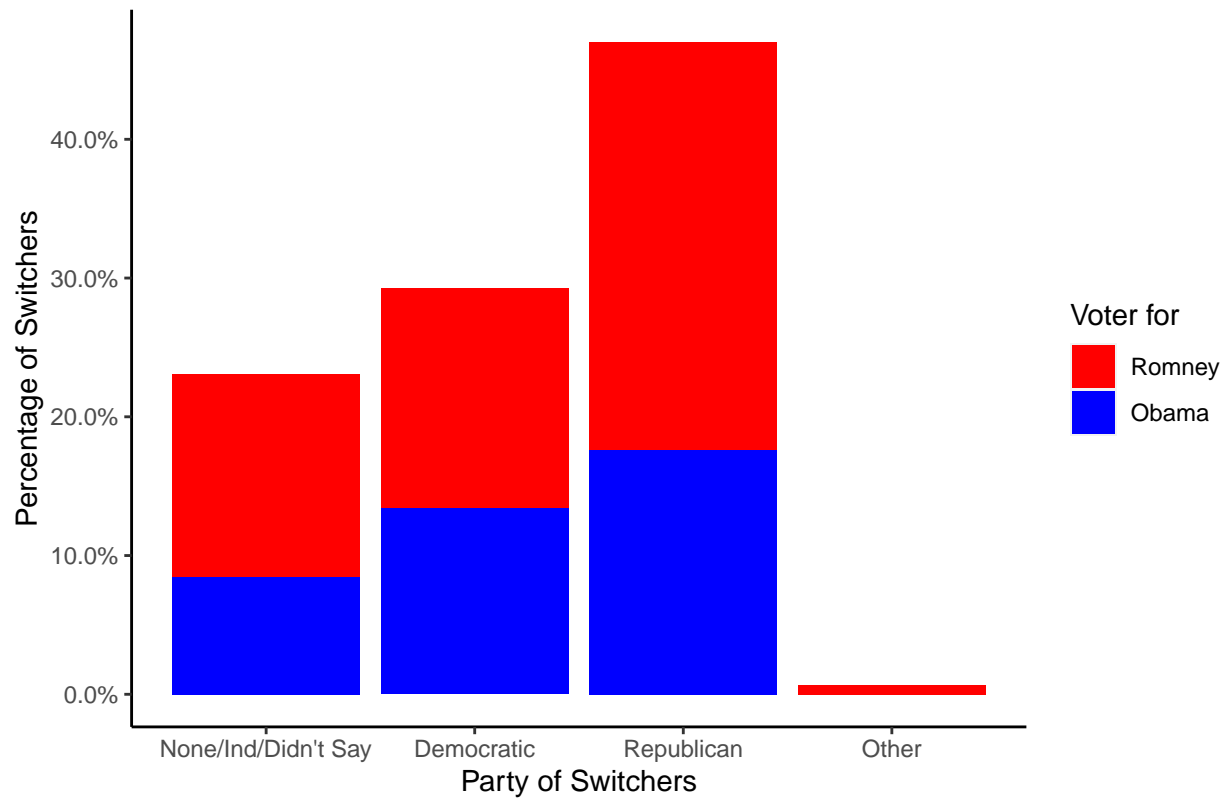
President 2012 Switchers by Age

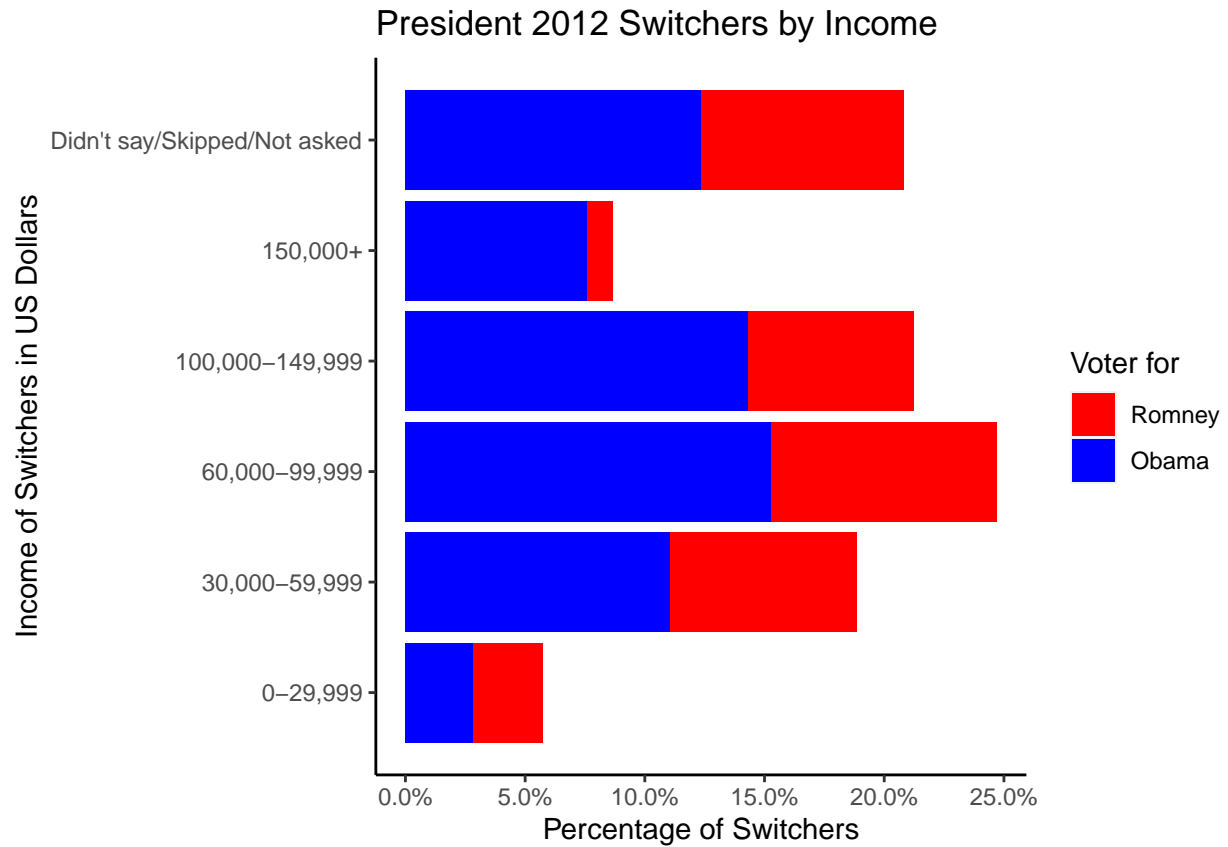


President 2012 Switchers by Education



President 2012 Switchers by Party





Descriptive stuff in one pane, and maybe a drop down panel to pick one, and maybe something that will allow the user to compare two or three demographic/policy groups groups