Data Analysis of Roll Call Data

The summary statistics and plots of the data can be found in the plots.qmd file in the same directory.

Simple Linear Probability Model

```
df_no_change <- df %>% filter(Vote_change_dummy == 0)

# get all reps who voted only pos. or only neg. and regress with all contributions
df_no_change <- df %>% filter(Vote_change_dummy == 0)
df_no_change$"all_votes" <- 0
# create a col that indicates if the rep voted all pos or neg. votes
for (i in 1:nrow(df_no_change)) {
    if (!is.na(any(df_no_change[i, c("Vote3", "Vote4", "Vote51", "Vote52", "Vote6", "Vote7")]
        df_no_change[i, "all_votes"] <- "+"
    } else if (!is.na(any(df_no_change[i, c("Vote3", "Vote4", "Vote51", "Vote52", "Vote6", "
        df_no_change[i, "all_votes"] <- "-"
    }
}

df_no_change <- dummy_cols(df_no_change, select_columns = "all_votes")
df_no_change <- df_no_change %>%
    rename("all_votes_plus" = "all_votes_+") %>%
    rename("all_votes_minus" = "all_votes_-")

df_no_change <- df_no_change %>% select(-c(
    last_name, first_name, name, District, Vote_count, Vote_change,
    Vote_change, Vote_change_dummy, member_id, Vote3, Vote4, Vote51, Vote52, Vote6, Vote7
))
```

```r
view(df_no_change)

ols_1 <- lm(all_votes_minus ~ . - all_votes - all_votes_plus - state, data = df_no_change)
summary(ols_1)
```

```
Call:
lm(formula = all_votes_minus ~ . - all_votes - all_votes_plus -
    state, data = df_no_change)

Residuals:
     Min       1Q   Median       3Q      Max
-0.48825 -0.20836 -0.02519  0.04117  0.95039

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)    -3.284e-02  1.577e-01  -0.208 0.835383
partyR          2.715e-01  6.803e-02   3.991 0.000107 ***
GeographicalMW -5.491e-04  1.666e-01  -0.003 0.997375
GeographicalNE  4.937e-02  1.762e-01   0.280 0.779787
GeographicalSO  7.210e-02  1.646e-01   0.438 0.662150
GeographicalWE  6.524e-02  1.643e-01   0.397 0.692019
pro_env.113    -1.978e-06  3.882e-06  -0.509 0.611251
anti_env.113   -8.252e-07  9.276e-07  -0.890 0.375228
pro_env.114    -1.968e-07  3.867e-06  -0.051 0.959476
anti_env.114   -8.718e-07  1.034e-06  -0.843 0.400755
pro_env.115     4.792e-06  6.013e-06   0.797 0.426846
anti_env.115    1.245e-06  1.213e-06   1.027 0.306322
pro_env.116    -4.340e-06  4.842e-06  -0.896 0.371593
anti_env.116    2.173e-06  8.181e-07   2.656 0.008848 **
pro_env.117    -1.834e-07  2.681e-06  -0.068 0.945564
anti_env.117   -1.957e-06  5.530e-07  -3.540 0.000550 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3102 on 135 degrees of freedom
  (386 observations deleted due to missingness)
Multiple R-squared:  0.309, Adjusted R-squared:  0.2322
F-statistic: 4.024 on 15 and 135 DF,  p-value: 5.234e-06
```

LPM per session

```r
# Vote 3
df_113 <- filter_session_data(df, "113")
ols_113 <- lm(Vote3_minus ~ . - Vote3 - Vote3_plus, data = df_113)
summary(ols_113)
```

```
Call:
lm(formula = Vote3_minus ~ . - Vote3 - Vote3_plus, data = df_113)

Residuals:
     Min       1Q   Median       3Q      Max
-0.10197 -0.04349 -0.01876  0.00988  0.94468

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     2.065e-02  6.473e-02   0.319   0.7500
partyR          9.066e-01  2.386e-02  37.993   <2e-16 ***
anti_env.113    2.910e-07  1.392e-07   2.091   0.0377 *
pro_env.113    -1.044e-06  9.369e-07  -1.115   0.2661
GeographicalMW  4.863e-02  6.765e-02   0.719   0.4729
GeographicalNE  9.553e-03  6.893e-02   0.139   0.8899
GeographicalSO  6.692e-02  6.673e-02   1.003   0.3170
GeographicalWE  2.649e-02  6.640e-02   0.399   0.6903
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1427 on 223 degrees of freedom
  (133 observations deleted due to missingness)
Multiple R-squared:   0.92, Adjusted R-squared:  0.9175
F-statistic: 366.2 on 7 and 223 DF,  p-value: < 2.2e-16
```

```r
# Vote 4
df_114 <- filter_session_data(df, "114")
ols_114 <- lm(Vote4_minus ~ . - Vote4 - Vote4_plus, data = df_114)
summary(ols_114)
```

```
Call:
```

```
lm(formula = Vote4_minus ~ . - Vote4 - Vote4_plus, data = df_114)

Residuals:
     Min       1Q   Median       3Q      Max
-0.98817 -0.02622 -0.00144  0.01460  0.93854

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     1.359e-02  6.612e-02   0.206   0.8373
partyR          9.091e-01  2.288e-02  39.733   <2e-16 ***
anti_env.114    2.561e-07  1.206e-07   2.123   0.0347 *
pro_env.114    -3.578e-07  7.709e-07  -0.464   0.6430
GeographicalMW  4.902e-02  6.924e-02   0.708   0.4796
GeographicalNE -1.523e-02  7.026e-02  -0.217   0.8286
GeographicalSO  6.114e-02  6.838e-02   0.894   0.3721
GeographicalWE  1.172e-02  6.845e-02   0.171   0.8641
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.147 on 257 degrees of freedom
  (155 observations deleted due to missingness)
Multiple R-squared:  0.9141,    Adjusted R-squared:  0.9117
F-statistic: 390.6 on 7 and 257 DF,  p-value: < 2.2e-16
```

```r
# Vote 51
df_1151 <- filter_session_data(df, "1151")
ols_1151 <- lm(Vote51_minus ~ . - Vote51 - Vote51_plus, data = df_1151)
summary(ols_1151)
```

```
Call:
lm(formula = Vote51_minus ~ . - Vote51 - Vote51_plus, data = df_1151)

Residuals:
     Min       1Q   Median       3Q      Max
-0.97783 -0.02038  0.00638  0.02071  0.94366

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     1.134e-02  6.099e-02   0.186   0.8527
partyR          9.191e-01  2.272e-02  40.454   <2e-16 ***
anti_env.115    2.516e-07  1.491e-07   1.687   0.0928 .
```

```
pro_env.115    -2.324e-07  8.706e-07  -0.267   0.7898
GeographicalMW  5.141e-02  6.414e-02   0.802   0.4236
GeographicalNE -3.332e-02  6.470e-02  -0.515   0.6070
GeographicalSO  4.533e-02  6.347e-02   0.714   0.4757
GeographicalWE  9.000e-03  6.306e-02   0.143   0.8866
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1477 on 256 degrees of freedom
  (151 observations deleted due to missingness)
Multiple R-squared:  0.9141,    Adjusted R-squared:  0.9117
F-statistic: 389.1 on 7 and 256 DF,  p-value: < 2.2e-16
```

```r
# Vote 52
df_1152 <- filter_session_data(df, "1152")
ols_1152 <- lm(Vote52_minus ~ . - Vote52 - Vote52_plus, data = df_1152)
summary(ols_1152)
```

```

Call:
lm(formula = Vote52_minus ~ . - Vote52 - Vote52_plus, data = df_1152)

Residuals:
     Min       1Q   Median       3Q      Max
-0.96624 -0.01687  0.03388  0.05833  0.94984

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     1.960e-02  8.748e-02   0.224    0.823
partyR          9.133e-01  2.957e-02  30.889   <2e-16 ***
anti_env.115   -1.509e-08  1.953e-07  -0.077    0.938
pro_env.115    -2.346e-07  1.156e-06  -0.203    0.839
GeographicalMW  9.420e-03  9.105e-02   0.103    0.918
GeographicalNE -3.456e-02  9.177e-02  -0.377    0.707
GeographicalSO  3.345e-02  9.009e-02   0.371    0.711
GeographicalWE -7.358e-04  9.013e-02  -0.008    0.993
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1934 on 251 degrees of freedom
  (152 observations deleted due to missingness)
Multiple R-squared:  0.8542,    Adjusted R-squared:  0.8501
```

```
F-statistic:   210 on 7 and 251 DF,  p-value: < 2.2e-16
```

```r
# Vote 6
df_116 <- filter_session_data(df, "116")
ols_116 <- lm(Vote6_minus ~ . - Vote6 - Vote6_plus, data = df_116)
summary(ols_116)
```

```

Call:
lm(formula = Vote6_minus ~ . - Vote6 - Vote6_plus, data = df_116)

Residuals:
     Min       1Q   Median       3Q      Max
-0.97615 -0.00761  0.01674  0.03295  0.96990

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     7.591e-03  6.108e-02   0.124    0.901
partyR          9.580e-01  2.129e-02  45.008   <2e-16 ***
anti_env.116    5.239e-08  1.128e-07   0.465    0.643
pro_env.116     1.686e-09  6.585e-07   0.003    0.998
GeographicalMW -5.010e-04  6.404e-02  -0.008    0.994
GeographicalNE -2.477e-02  6.459e-02  -0.384    0.702
GeographicalSO -3.775e-03  6.328e-02  -0.060    0.952
GeographicalWE  1.574e-02  6.348e-02   0.248    0.804
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1361 on 260 degrees of freedom
  (140 observations deleted due to missingness)
Multiple R-squared:  0.9277,    Adjusted R-squared:  0.9258
F-statistic: 476.9 on 7 and 260 DF,  p-value: < 2.2e-16
```

```r
# Vote 7
df_117 <- filter_session_data(df, "117")
ols_117 <- lm(Vote7_minus ~ . - Vote7 - Vote7_plus, data = df_117)
summary(ols_117)
```

```

Call:
lm(formula = Vote7_minus ~ . - Vote7 - Vote7_plus, data = df_117)
```

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.97484 -0.01100  0.02600  0.03423  0.08917

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     1.105e-02  6.358e-02   0.174    0.862
partyR          9.585e-01  2.002e-02  47.882   <2e-16 ***
anti_env.117   -4.818e-08  1.003e-07  -0.480    0.631
pro_env.117     5.306e-08  4.739e-07   0.112    0.911
GeographicalMW -5.898e-04  6.585e-02  -0.009    0.993
GeographicalNE -5.115e-02  6.652e-02  -0.769    0.443
GeographicalSO  5.827e-03  6.542e-02   0.089    0.929
GeographicalWE -1.258e-04  6.534e-02  -0.002    0.998
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1265 on 235 degrees of freedom
  (112 observations deleted due to missingness)
Multiple R-squared:  0.937, Adjusted R-squared:  0.9352
F-statistic: 499.7 on 7 and 235 DF,  p-value: < 2.2e-16
```

LPM of only mind changers

```r
df_subsample <- df_subsample %>% select(-c(
    last_name, first_name, name, District, Vote_count, Vote_change,
    Vote_change, Vote_change_dummy, votes, member_id, Vote3, Vote4, Vote51, Vote52, Vote6, Vo
))
# view(df_subsample)

ols_subsample <- lm(vote_change_to_anti ~ ., data = df_subsample)
summary(ols_subsample)
```

```
Call:
lm(formula = vote_change_to_anti ~ ., data = df_subsample)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-0.5000 -0.3333  0.0000  0.3750  0.6667

Coefficients: (7 not defined because of singularities)
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     8.761e-02  2.024e+00   0.043    0.967
partyR         -2.293e+00  3.812e+00  -0.601    0.570
GeographicalNE  1.160e+01  1.580e+01   0.734    0.491
GeographicalSO  4.598e-01  1.920e+00   0.239    0.819
GeographicalWE  4.830e+00  6.500e+00   0.743    0.486
pro_env.113     8.931e-04  1.083e-03   0.825    0.441
anti_env.113   -4.994e-06  6.710e-06  -0.744    0.485
pro_env.114    -9.263e-04  1.202e-03  -0.771    0.470
anti_env.114           NA         NA      NA       NA
pro_env.115            NA         NA      NA       NA
anti_env.115           NA         NA      NA       NA
pro_env.116            NA         NA      NA       NA
anti_env.116           NA         NA      NA       NA
pro_env.117            NA         NA      NA       NA
anti_env.117           NA         NA      NA       NA

Residual standard error: 0.6009 on 6 degrees of freedom
  (17 observations deleted due to missingness)
Multiple R-squared:  0.2417,    Adjusted R-squared:  -0.6431
F-statistic: 0.2732 on 7 and 6 DF,  p-value: 0.9432
```

Why warning message?