Conclusions

Contents

Overview

This script specializes in analyzing the models and drawing inferences for the research question. The research question is whether voting blocs, echo nest music features and migration patterns can explain the points and voting patterns of the 2016 ESC. This can be converted into a statistical problem using multiple linear regression whereby determining whether voting blocs, echo nest music features and migration patterns significantly explain the points and voting patterns.

This question will be answer by: 1. Constricting t-tests 2. Observing the signs of the estimated coefficients 3. measuring the increase in variance explained (R-sq) with the addition of a predictor variable

```
# load relevant libraries
library(rmarkdown)
library(knitr)
library(car)
```

```
## Loading required package: carData
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:car':
##
##      recode


## The following objects are masked from 'package:stats':
##
##      filter, lag


## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
```

```r
# load in the historic voting data for deriving the voting blocs
processed_data <- read.csv(file = "./data/arch/processed_data.csv", header = T)
# split the televote data
televote_data <- processed_data %>% filter(Voting_Method_J == 0)
# split out the jury vote data
jury_data <- processed_data %>% filter(Voting_Method_J == 1)
```
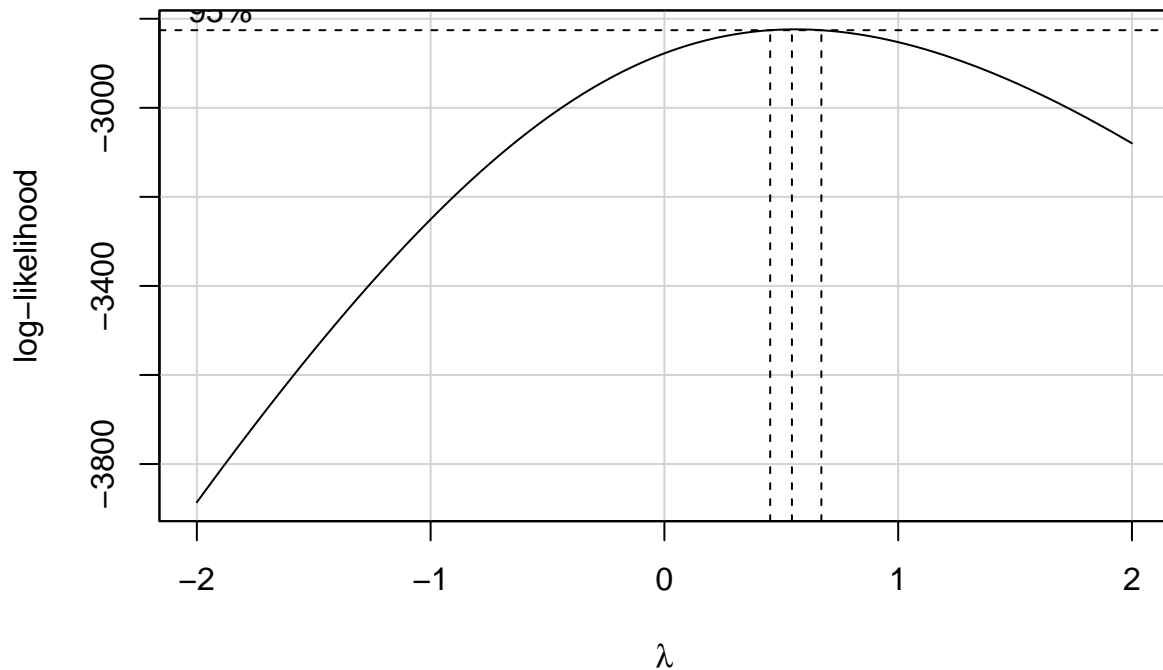
Overall Model

```r
#-- Fit the Final Model --#

# define the model formula for the final model
my_model_overall_form <- Points ~ Average_Points + VBlocs1_TC_3 + CAP_DIST_km +
                              FC_NonCitzens + ComLANGFAM_y + liveness + key_3 + METRIC_Citizens +
                              TC_PerfType_Solo + key_2 + VBlocs1_TC_13 + key_6 + time_signature_4 +
                              ComVBlocs1_y + VBlocs1_TC_1 + key_5 + OOA + speechiness
# fit the final overall model
my_model_overall <- lm(formula = my_model_overall_form, data = processed_data)
# perform box-cox transformation
bct <- boxCox(object = my_model_overall)
```

Profile Log−likelihood


```r
# return the optimal power transformation
p <- bct$x[which.max(x = bct$y)]
# transform the response variable
bctPoints <- (((processed_data$Points)^p) - 1)/(p)
# redefine the model formula for the final model
my_model_overall_bct_from <- bctPoints ~ Average_Points + VBlocs1_TC_3 + CAP_DIST_km +
                                 FC_NonCitzens + ComLANGFAM_y + liveness + key_3 + METRIC_Citi
                                 TC_PerfType_Solo + key_2 + VBlocs1_TC_13 + key_6 + time_signa
                                 ComVBlocs1_y + VBlocs1_TC_1 + key_5 + OOA + speechiness
# refit the model with the power transformation
my_model_overall <- lm(my_model_overall_bct_from, data = processed_data)
# generate final model summary
summary(my_model_overall)
```

```
##
## Call:
## lm(formula = my_model_overall_bct_from, data = processed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6714 -1.0056  0.0282  0.9735  3.1707
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.20107    0.45207   4.869 1.42e-06 ***
## Average_Points 0.20808    0.06479   3.212 0.001386 **
```

```
## VBlocs1_TC_3      0.58153    0.19211    3.027 0.002568 **
## CAP_DIST_km       0.12954    0.06434    2.013 0.044490 *
## FC_NonCitzens     0.17855    0.07048    2.533 0.011538 *
## ComLANGFAM_y      0.37800    0.13986    2.703 0.007061 **
## liveness         -0.29223    0.06762   -4.322 1.79e-05 ***
## key_3             0.51047    0.23228    2.198 0.028329 *
## METRIC_Citizens   0.09698    0.06789    1.428 0.153639
## TC_PerfType_Solo  0.94943    0.26757    3.548 0.000416 ***
## key_2            -0.92510    0.45034   -2.054 0.040363 *
## VBlocs1_TC_13    -1.99317    1.10071   -1.811 0.070641 .
## key_6            -1.38873    0.44355   -3.131 0.001822 **
## time_signature_4 -0.88157    0.36291   -2.429 0.015408 *
## ComVBlocs1_y     -0.40286    0.17563   -2.294 0.022128 *
## VBlocs1_TC_1      0.50874    0.17131    2.970 0.003092 **
## key_5            -0.48380    0.19256   -2.512 0.012236 *
## OOA               0.47418    0.23786    1.994 0.046630 *
## speechiness       0.10033    0.06662    1.506 0.132541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.43 on 639 degrees of freedom
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.1867
## F-statistic: 9.379 on 18 and 639 DF,  p-value: < 2.2e-16
```

```
# write the model to disk
saveRDS(object = my_model_overall, file = './models/arch/overall_final_model.RDS')
```

T-tests

Make sure to report 1. the adjusted R-sq value 2. the degrees of freedom 3. the f statistic for f test 4. the p-value for f test

NOTE: include notation for levels of significance * for 0.05 ** for 0.01 *** for 0.001

```
summary(my_model_overall)
```

```
##
## Call:
## lm(formula = my_model_overall_bct_from, data = processed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6714 -1.0056  0.0282  0.9735  3.1707
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.20107    0.45207    4.869 1.42e-06 ***
## Average_Points   0.20808    0.06479    3.212 0.001386 **
## VBlocs1_TC_3     0.58153    0.19211    3.027 0.002568 **
## CAP_DIST_km      0.12954    0.06434    2.013 0.044490 *
## FC_NonCitzens    0.17855    0.07048    2.533 0.011538 *
## ComLANGFAM_y     0.37800    0.13986    2.703 0.007061 **
## liveness        -0.29223    0.06762   -4.322 1.79e-05 ***
```

4

```
## key_3              0.51047    0.23228   2.198 0.028329 *
## METRIC_Citizens    0.09698    0.06789   1.428 0.153639
## TC_PerfType_Solo   0.94943    0.26757   3.548 0.000416 ***
## key_2             -0.92510    0.45034  -2.054 0.040363 *
## VBlocs1_TC_13     -1.99317    1.10071  -1.811 0.070641 .
## key_6             -1.38873    0.44355  -3.131 0.001822 **
## time_signature_4  -0.88157    0.36291  -2.429 0.015408 *
## ComVBlocs1_y      -0.40286    0.17563  -2.294 0.022128 *
## VBlocs1_TC_1       0.50874    0.17131   2.970 0.003092 **
## key_5             -0.48380    0.19256  -2.512 0.012236 *
## OOA                0.47418    0.23786   1.994 0.046630 *
## speechiness        0.10033    0.06662   1.506 0.132541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.43 on 639 degrees of freedom
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.1867
## F-statistic: 9.379 on 18 and 639 DF,  p-value: < 2.2e-16
```

Directional / Sign effects

Report the sign effects of each significant coefficient + indicates the predictor variable has a positive effect on the dependent variable - indicates the predictor variables has a negative effect on the dependent variable

```
summary(my_model_overall)
```

```
##
## Call:
## lm(formula = my_model_overall_bct_from, data = processed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6714 -1.0056  0.0282  0.9735  3.1707
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        2.20107    0.45207   4.869 1.42e-06 ***
## Average_Points     0.20808    0.06479   3.212 0.001386 **
## VBlocs1_TC_3       0.58153    0.19211   3.027 0.002568 **
## CAP_DIST_km        0.12954    0.06434   2.013 0.044490 *
## FC_NonCitzens      0.17855    0.07048   2.533 0.011538 *
## ComLANGFAM_y       0.37800    0.13986   2.703 0.007061 **
## liveness          -0.29223    0.06762  -4.322 1.79e-05 ***
## key_3              0.51047    0.23228   2.198 0.028329 *
## METRIC_Citizens    0.09698    0.06789   1.428 0.153639
## TC_PerfType_Solo   0.94943    0.26757   3.548 0.000416 ***
## key_2             -0.92510    0.45034  -2.054 0.040363 *
## VBlocs1_TC_13     -1.99317    1.10071  -1.811 0.070641 .
## key_6             -1.38873    0.44355  -3.131 0.001822 **
## time_signature_4  -0.88157    0.36291  -2.429 0.015408 *
## ComVBlocs1_y      -0.40286    0.17563  -2.294 0.022128 *
## VBlocs1_TC_1       0.50874    0.17131   2.970 0.003092 **
## key_5             -0.48380    0.19256  -2.512 0.012236 *
```

```
## OOA                 0.47418     0.23786    1.994 0.046630 *
## speechiness          0.10033     0.06662    1.506 0.132541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.43 on 639 degrees of freedom
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.1867
## F-statistic: 9.379 on 18 and 639 DF,  p-value: < 2.2e-16
```

Explained Variance

Investigate the increase of variance explained by incorporating specific predictor variables. Observe the increase in R-sq when a predictor variable is included / excluded from the model. Do this for voting blocs, Echo Nest music factors and Migration patterns

```
# define the model formula
omodel_ex.vblocs_form <- bctPoints ~ Average_Points + CAP_DIST_km +
                                FC_NonCitzens + ComLANGFAM_y + liveness + key_3 + METRIC_Citizens +
                                TC_PerfType_Solo + key_2 + key_6 + time_signature_4 +
                                key_5 + OOA + speechiness
# fit the linear model
omodel_ex.vblocs <- lm(formula = omodel_ex.vblocs_form, data = processed_data)
# generate model summary
summary(omodel_ex.vblocs)
```

```
##
## Call:
## lm(formula = omodel_ex.vblocs_form, data = processed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6544 -1.0223  0.0817  1.0404  3.1696
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.22835    0.45398   4.909 1.16e-06 ***
## Average_Points    0.21609    0.06485   3.332 0.000912 ***
## CAP_DIST_km       0.15612    0.06145   2.540 0.011305 *
## FC_NonCitzens     0.13242    0.06668   1.986 0.047477 *
## ComLANGFAM_y      0.33952    0.13695   2.479 0.013422 *
## liveness         -0.30582    0.06714  -4.555 6.28e-06 ***
## key_3             0.79740    0.21742   3.668 0.000265 ***
## METRIC_Citizens   0.15658    0.06686   2.342 0.019483 *
## TC_PerfType_Solo  0.99223    0.26671   3.720 0.000216 ***
## key_2            -1.11985    0.44859  -2.496 0.012797 *
## key_6            -1.27711    0.44667  -2.859 0.004385 **
## time_signature_4 -0.71024    0.36227  -1.960 0.050369 .
## key_5            -0.32412    0.18072  -1.794 0.073358 .
## OOA               0.35308    0.23108   1.528 0.127012
## speechiness       0.09159    0.06593   1.389 0.165250
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

6

```
## Residual standard error: 1.449 on 643 degrees of freedom
## Multiple R-squared:  0.1837, Adjusted R-squared:  0.1659
## F-statistic: 10.34 on 14 and 643 DF,  p-value: < 2.2e-16
```

```r
# define the model formula
omodel_ex.music_form <- bctPoints ~ Average_Points + VBlocs1_TC_3 + CAP_DIST_km +
                                    FC_NonCitzens + ComLANGFAM_y + METRIC_Citizens +
                                    TC_PerfType_Solo + VBlocs1_TC_13 +
                                     ComVBlocs1_y + VBlocs1_TC_1 + OOA
# fit the linear model
omodel_ex.music <- lm(omodel_ex.music_form, data = processed_data)
# investigate the R-sq value
summary(omodel_ex.music)
```

```
##
## Call:
## lm(formula = omodel_ex.music_form, data = processed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.5771 -1.0536  0.1137  1.0166  3.3280
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.65789    0.30711   5.398 9.45e-08 ***
## Average_Points    0.19497    0.06128   3.182  0.00153 **
## VBlocs1_TC_3      0.72053    0.17954   4.013 6.69e-05 ***
## CAP_DIST_km       0.20068    0.06109   3.285  0.00107 **
## FC_NonCitzens     0.21016    0.07179   2.928  0.00354 **
## ComLANGFAM_y      0.45916    0.13999   3.280  0.00109 **
## METRIC_Citizens   0.10930    0.06850   1.596  0.11106
## TC_PerfType_Solo  0.54352    0.26489   2.052  0.04058 *
## VBlocs1_TC_13    -1.91138    1.12688  -1.696  0.09034 .
## ComVBlocs1_y     -0.43479    0.18095  -2.403  0.01655 *
## VBlocs1_TC_1      0.45805    0.16015   2.860  0.00437 **
## OOA               0.49416    0.21639   2.284  0.02272 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.476 on 646 degrees of freedom
## Multiple R-squared:  0.1481, Adjusted R-squared:  0.1336
## F-statistic: 10.21 on 11 and 646 DF,  p-value: < 2.2e-16
```

```r
# define the model formula
omodel_ex.mig_form <- Points ~ Average_Points + VBlocs1_TC_3 + CAP_DIST_km +
                               ComLANGFAM_y + liveness + key_3 +
                               TC_PerfType_Solo + key_2 + VBlocs1_TC_13 + key_6 + time_signature_4 +
                               ComVBlocs1_y + VBlocs1_TC_1 + key_5 + OOA + speechiness
# fit the linear model
omodel_ex.mig <- lm(formula = omodel_ex.mig_form, data = processed_data)
# generate the model summary
summary(omodel_ex.mig)
```

```
##
```

```
## Call:
## lm(formula = omodel_ex.mig_form, data = processed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.2863 -2.2401 -0.3431  2.1879  7.4326
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         4.7632     0.9656   4.933 1.03e-06 ***
## Average_Points      0.6159     0.1354   4.548 6.49e-06 ***
## VBlocs1_TC_3        1.5416     0.3965   3.888 0.000112 ***
## CAP_DIST_km         0.1943     0.1342   1.448 0.148080
## ComLANGFAM_y        0.6732     0.2970   2.267 0.023721 *
## liveness           -0.5601     0.1444  -3.880 0.000115 ***
## key_3               0.9461     0.4900   1.931 0.053968 .
## TC_PerfType_Solo    2.0257     0.5712   3.546 0.000419 ***
## key_2              -1.6246     0.9621  -1.689 0.091762 .
## VBlocs1_TC_13      -1.2822     2.1927  -0.585 0.558927
## key_6              -3.1631     0.9456  -3.345 0.000870 ***
## time_signature_4   -2.0330     0.7748  -2.624 0.008902 **
## ComVBlocs1_y       -0.8538     0.3740  -2.283 0.022772 *
## VBlocs1_TC_1        1.3193     0.3620   3.644 0.000290 ***
## key_5              -1.3732     0.3967  -3.461 0.000573 ***
## OOA                 1.0335     0.5066   2.040 0.041746 *
## speechiness         0.2187     0.1423   1.537 0.124815
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.056 on 641 degrees of freedom
## Multiple R-squared:  0.1934, Adjusted R-squared:  0.1732
## F-statistic: 9.604 on 16 and 641 DF,  p-value: < 2.2e-16
```
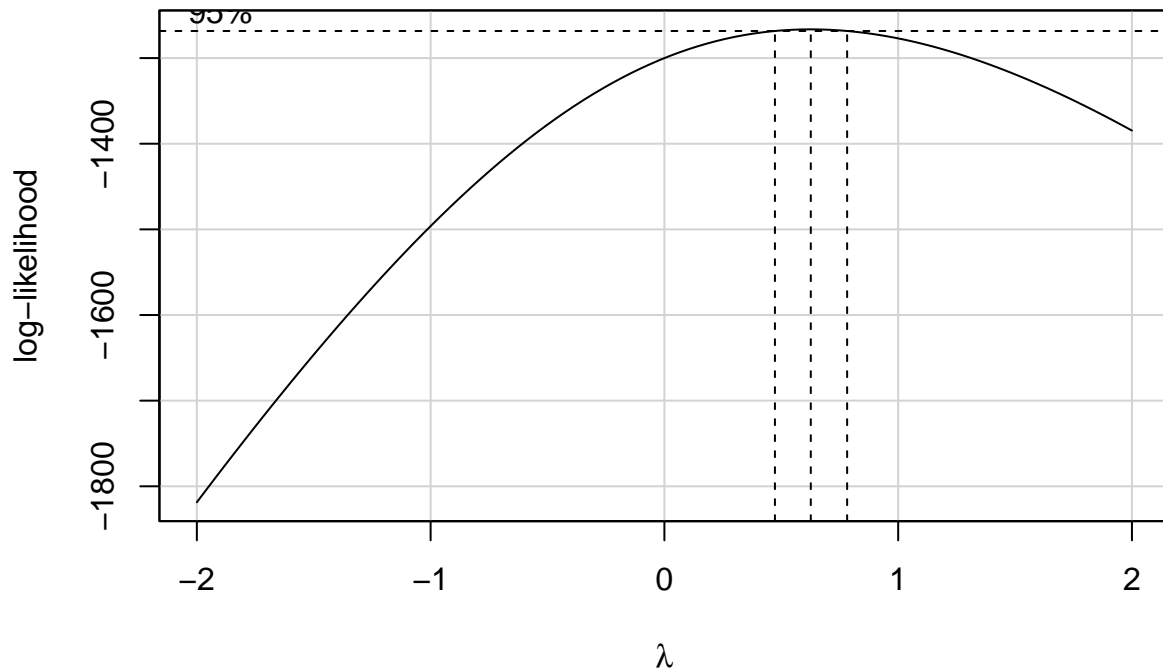
TELEVOTE MODEL

```r
# define the model formula
my_model_tele_form <- Points ~ METRIC_Citizens + Average_Points + VBlocs1_TC_3 +
                         VBlocs2_TC_1 + mode_1 + key_11 + OOA + acousticness + danceability +
                         key_7 + VBlocs1_TC_13 + ComLANGFAM_y
# fit the linear model
my_model_tele <- lm(formula = my_model_tele_form, data = televote_data)
# perform box-cox transformation
bct <- boxCox(object = my_model_tele)
```

Profile Log–likelihood


```r
# return the optimal power transformation
p <- bct$x[which.max(x = bct$y)]
# transform the response variable
bctPoints <- (((televote_data$Points)^p) - 1)/(p)
# define the model formula
my_model_tele_bct <- bctPoints ~ METRIC_Citizens + Average_Points + VBlocs1_TC_3 +
                            VBlocs2_TC_1 + mode_1 + key_11 + OOA + acousticness + danceability +
                            key_7 + VBlocs1_TC_13 + ComLANGFAM_y
# refit the model with the power transformation
my_model_tele <- lm(my_model_tele_bct, data = televote_data)
# generate model summary
summary(my_model_tele)
```

```
##
## Call:
## lm(formula = my_model_tele_bct, data = televote_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6953 -1.0485  0.1163  1.0374  3.7539
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.26304    0.24167   9.364  < 2e-16 ***
## METRIC_Citizens  0.34461    0.07753   4.445 1.22e-05 ***
## Average_Points   0.39603    0.08918   4.441 1.24e-05 ***
```

```
## VBlocs1_TC_3      1.07370     0.36185    2.967 0.003236 **
## VBlocs2_TC_1     -0.77013     0.24756   -3.111 0.002037 **
## mode_1            0.70492     0.18910    3.728 0.000229 ***
## key_11            0.99686     0.31422    3.173 0.001661 **
## OOA               0.65497     0.35667    1.836 0.067249 .
## acousticness      0.32378     0.13542    2.391 0.017397 *
## danceability     -0.30145     0.12654   -2.382 0.017800 *
## key_7            -0.69771     0.38737   -1.801 0.072639 .
## VBlocs1_TC_13    -1.87918     1.14005   -1.648 0.100287
## ComLANGFAM_y      0.30175     0.20574    1.467 0.143465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.519 on 314 degrees of freedom
## Multiple R-squared:  0.3206, Adjusted R-squared:  0.2946
## F-statistic: 12.35 on 12 and 314 DF,  p-value: < 2.2e-16
```

```
# write the model to disk
saveRDS(object = my_model_tele, file = './models/arch/televote_final_model.RDS')
```

T-tests

Make sure to report 1. the adjusted R-sq value 2. the degrees of freedom 3. the f statistic for f test 4. the p-value for f test

NOTE: include notation for levels of significance * for 0.05 ** for 0.01 *** for 0.001

```
summary(my_model_tele)
```

```
##
## Call:
## lm(formula = my_model_tele_bct, data = televote_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6953 -1.0485  0.1163  1.0374  3.7539
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.26304    0.24167    9.364  < 2e-16 ***
## METRIC_Citizens  0.34461    0.07753    4.445 1.22e-05 ***
## Average_Points   0.39603    0.08918    4.441 1.24e-05 ***
## VBlocs1_TC_3     1.07370    0.36185    2.967 0.003236 **
## VBlocs2_TC_1    -0.77013    0.24756   -3.111 0.002037 **
## mode_1           0.70492    0.18910    3.728 0.000229 ***
## key_11           0.99686    0.31422    3.173 0.001661 **
## OOA              0.65497    0.35667    1.836 0.067249 .
## acousticness     0.32378    0.13542    2.391 0.017397 *
## danceability    -0.30145    0.12654   -2.382 0.017800 *
## key_7           -0.69771    0.38737   -1.801 0.072639 .
## VBlocs1_TC_13   -1.87918    1.14005   -1.648 0.100287
## ComLANGFAM_y     0.30175    0.20574    1.467 0.143465
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.519 on 314 degrees of freedom
## Multiple R-squared:  0.3206, Adjusted R-squared:  0.2946
## F-statistic: 12.35 on 12 and 314 DF,  p-value: < 2.2e-16
```

Directional / Sign effects

Report the sign effects of each significant coefficient + indicates the predictor variable has a positive effect on the dependent variable - indicates the predictor variables has a negative effect on the defendant variable

```
summary(my_model_tele)
```

```
##
## Call:
## lm(formula = my_model_tele_bct, data = televote_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6953 -1.0485  0.1163  1.0374  3.7539
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     2.26304    0.24167   9.364  < 2e-16 ***
## METRIC_Citizens 0.34461    0.07753   4.445 1.22e-05 ***
## Average_Points  0.39603    0.08918   4.441 1.24e-05 ***
## VBlocs1_TC_3    1.07370    0.36185   2.967 0.003236 **
## VBlocs2_TC_1   -0.77013    0.24756  -3.111 0.002037 **
## mode_1          0.70492    0.18910   3.728 0.000229 ***
## key_11          0.99686    0.31422   3.173 0.001661 **
## OOA             0.65497    0.35667   1.836 0.067249 .
## acousticness    0.32378    0.13542   2.391 0.017397 *
## danceability   -0.30145    0.12654  -2.382 0.017800 *
## key_7          -0.69771    0.38737  -1.801 0.072639 .
## VBlocs1_TC_13  -1.87918    1.14005  -1.648 0.100287
## ComLANGFAM_y    0.30175    0.20574   1.467 0.143465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.519 on 314 degrees of freedom
## Multiple R-squared:  0.3206, Adjusted R-squared:  0.2946
## F-statistic: 12.35 on 12 and 314 DF,  p-value: < 2.2e-16
```

Explained Variance

Investigate the increase of variance explained by incorporating specific predictor variables. Observe the increase in R-sq when a predictor variable is included / excluded from the model. Do this for voting blocs, Echo Nest music factors and Migration patterns

```
# define model formula
tmodel_ex.vblocs_form <- bctPoints ~ METRIC_Citizens + Average_Points +
                                     mode_1 + key_11 + OOA + acousticness + danceability +
```

```
                              key_7 + ComLANGFAM_y
# fit linear model
tmodel_ex.vblocs <- lm(tmodel_ex.vblocs_form, data = televote_data)
# investigate the R-sq value
summary(tmodel_ex.vblocs)
```

```
##
## Call:
## lm(formula = tmodel_ex.vblocs_form, data = televote_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7686 -1.0841  0.1345  1.0881  3.2252
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.29790    0.24016   9.568  < 2e-16 ***
## METRIC_Citizens  0.37740    0.07695   4.905 1.50e-06 ***
## Average_Points   0.40399    0.09069   4.454 1.17e-05 ***
## mode_1           0.72012    0.19239   3.743 0.000216 ***
## key_11           0.96355    0.28618   3.367 0.000854 ***
## OOA              0.43418    0.34958   1.242 0.215156
## acousticness     0.50816    0.11960   4.249 2.83e-05 ***
## danceability    -0.21988    0.12175  -1.806 0.071870 .
## key_7           -1.02868    0.37187  -2.766 0.006004 **
## ComLANGFAM_y     0.23183    0.19775   1.172 0.241926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.547 on 317 degrees of freedom
## Multiple R-squared:  0.2881, Adjusted R-squared:  0.2679
## F-statistic: 14.26 on 9 and 317 DF,  p-value: < 2.2e-16
```

```
# define model formula
tmodel_ex.music_form <- bctPoints ~ METRIC_Citizens + Average_Points + VBlocs1_TC_3 +
                            VBlocs2_TC_1 + OOA + VBlocs1_TC_13 + ComLANGFAM_y
# fit linear model
tmodel_ex.music <- lm(tmodel_ex.music_form, data = televote_data)
# investigate the R-sq value
summary(tmodel_ex.music)
```

```
##
## Call:
## lm(formula = tmodel_ex.music_form, data = televote_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.793 -1.085  0.144  1.020  3.462
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.47465    0.19794  12.502  < 2e-16 ***
## METRIC_Citizens  0.42701    0.07645   5.585 5.00e-08 ***
```

```
## Average_Points    0.37322     0.09078    4.111 5.01e-05 ***
## VBlocs1_TC_3      1.42719     0.28383    5.028 8.27e-07 ***
## VBlocs2_TC_1     -0.77383     0.22735   -3.404  0.00075 ***
## OOA               0.79361     0.33509    2.368  0.01846 *
## VBlocs1_TC_13    -1.63414     1.12931   -1.447  0.14887
## ComLANGFAM_y      0.41105     0.20291    2.026  0.04362 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.558 on 319 degrees of freedom
## Multiple R-squared:  0.273,  Adjusted R-squared:  0.257
## F-statistic: 17.11 on 7 and 319 DF,  p-value: < 2.2e-16
```

```r
# define model formula
tmodel_ex.mig_form <- bctPoints ~  Average_Points + VBlocs1_TC_3 +
                                   VBlocs2_TC_1 + mode_1 + key_11 + OOA + acousticness + danceability +
                                   key_7 + VBlocs1_TC_13 + ComLANGFAM_y
# fit linear model
tmodel_ex.mig <- lm(tmodel_ex.mig_form, data = televote_data)
# investigate the R-sq value
summary(tmodel_ex.mig)
```

```
##
## Call:
## lm(formula = tmodel_ex.mig_form, data = televote_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0791 -1.1216  0.0716  1.0725  3.6974
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.23537    0.24868   8.989  < 2e-16 ***
## Average_Points 0.45656    0.09072   5.033 8.15e-07 ***
## VBlocs1_TC_3   1.37863    0.36571   3.770 0.000195 ***
## VBlocs2_TC_1  -0.91166    0.25271  -3.608 0.000359 ***
## mode_1         0.84530    0.19191   4.405 1.45e-05 ***
## key_11         1.27310    0.31705   4.015 7.42e-05 ***
## OOA            0.60631    0.36696   1.652 0.099483 .
## acousticness   0.34146    0.13934   2.451 0.014803 *
## danceability  -0.41039    0.12778  -3.212 0.001456 **
## key_7         -0.77320    0.39835  -1.941 0.053149 .
## VBlocs1_TC_13 -1.28113    1.16530  -1.099 0.272436
## ComLANGFAM_y   0.27895    0.21171   1.318 0.188600
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.563 on 315 degrees of freedom
## Multiple R-squared:  0.2778, Adjusted R-squared:  0.2526
## F-statistic: 11.02 on 11 and 315 DF,  p-value: < 2.2e-16
```

JURY MODEL

```r
# power transformation of 3/4
# define model formula
# Note: thesis shows output for non-power transformed case (i.e. Points ~ VBlocs2_TC_4 + key_3 + ... )
my_model_jury_form <- (Points)^(3/4) ~ VBlocs2_TC_4 + key_3 + TC_PerfType_Solo +
                                        liveness + ComVBlocs1_y + ComLANGFAM_y
# fit linear model
my_model_jury <- lm(formula = my_model_jury_form, data = jury_data)
# generate model summary
summary(my_model_jury)
```

```
##
## Call:
## lm(formula = my_model_jury_form, data = jury_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0963 -1.1107 -0.0648  1.0644  3.6173
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.83982    0.41895   4.391 1.53e-05 ***
## VBlocs2_TC_4      1.69551    0.27077   6.262 1.21e-09 ***
## key_3             1.00919    0.26191   3.853 0.000141 ***
## TC_PerfType_Solo  1.53092    0.42200   3.628 0.000332 ***
## liveness         -0.26215    0.07656  -3.424 0.000696 ***
## ComVBlocs1_y     -0.43580    0.21889  -1.991 0.047320 *
## ComLANGFAM_y      0.39571    0.19859   1.993 0.047142 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.476 on 324 degrees of freedom
## Multiple R-squared:  0.2127, Adjusted R-squared:  0.1981
## F-statistic: 14.59 on 6 and 324 DF,  p-value: 9.53e-15
```

```r
# write the model to disk
saveRDS(object = my_model_jury, file = './models/arch/jury_final_model.RDS')
```

T-tests

Make sure to report 1. the adjusted R-sq value 2. the degrees of freedom 3. the f statistic for f test 4. the p-value for f test

NOTE: include notation for levels of significance * for 0.05 ** for 0.01 *** for 0.001

```r
summary(my_model_jury)
```

```
##
## Call:
## lm(formula = my_model_jury_form, data = jury_data)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0963 -1.1107 -0.0648  1.0644  3.6173
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.83982    0.41895   4.391 1.53e-05 ***
## VBlocs2_TC_4      1.69551    0.27077   6.262 1.21e-09 ***
## key_3            1.00919    0.26191   3.853 0.000141 ***
## TC_PerfType_Solo 1.53092    0.42200   3.628 0.000332 ***
## liveness         -0.26215    0.07656  -3.424 0.000696 ***
## ComVBlocs1_y     -0.43580    0.21889  -1.991 0.047320 *
## ComLANGFAM_y      0.39571    0.19859   1.993 0.047142 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.476 on 324 degrees of freedom
## Multiple R-squared:  0.2127, Adjusted R-squared:  0.1981
## F-statistic: 14.59 on 6 and 324 DF,  p-value: 9.53e-15
```

Directional / Sign effects

Report the sign effects of each significant coefficient + indicates the predictor variable has a positive effect on the dependent variable - indicates the predictor variables has a negative effect on the dependent variable

```
summary(my_model_jury)
```

```
##
## Call:
## lm(formula = my_model_jury_form, data = jury_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0963 -1.1107 -0.0648  1.0644  3.6173
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.83982    0.41895   4.391 1.53e-05 ***
## VBlocs2_TC_4      1.69551    0.27077   6.262 1.21e-09 ***
## key_3            1.00919    0.26191   3.853 0.000141 ***
## TC_PerfType_Solo 1.53092    0.42200   3.628 0.000332 ***
## liveness         -0.26215    0.07656  -3.424 0.000696 ***
## ComVBlocs1_y     -0.43580    0.21889  -1.991 0.047320 *
## ComLANGFAM_y      0.39571    0.19859   1.993 0.047142 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.476 on 324 degrees of freedom
## Multiple R-squared:  0.2127, Adjusted R-squared:  0.1981
## F-statistic: 14.59 on 6 and 324 DF,  p-value: 9.53e-15
```

Explained Variance

Investigate the increase of variance explained by incorporating specific predictor variables. Observe the increase in R-sq when a predictor variable is included / excluded from the model. Do this for voting blocs, Echo Nest music factors and Migration patterns

```r
# define model formula
jmodel_ex.vblocs_form <- (Points)^(3/4) ~ key_3 + TC_PerfType_Solo +
                                          liveness + ComVBlocs1_y + ComLANGFAM_y
# fit linear model
jmodel_ex.vblocs <- lm(formula = jmodel_ex.vblocs_form, data = jury_data)
# investigate the R-sq value
summary(jmodel_ex.vblocs)
```

```
##
## Call:
## lm(formula = jmodel_ex.vblocs_form, data = jury_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9769 -1.1906 -0.1433  1.1388  3.1711
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        2.2309     0.4380   5.094 5.95e-07 ***
## key_3              0.9310     0.2766   3.366 0.000854 ***
## TC_PerfType_Solo   1.2335     0.4433   2.783 0.005707 **
## liveness          -0.3490     0.0796  -4.384 1.58e-05 ***
## ComVBlocs1_y      -0.0646     0.2228  -0.290 0.771987
## ComLANGFAM_y       0.6183     0.2065   2.994 0.002967 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.561 on 325 degrees of freedom
## Multiple R-squared:  0.1174, Adjusted R-squared:  0.1038
## F-statistic: 8.645 on 5 and 325 DF,  p-value: 1.046e-07
```

```r
# define model formula
jmodel_ex.music_form <- (Points)^(3/4) ~ VBlocs2_TC_4 + TC_PerfType_Solo +
                                         ComVBlocs1_y + ComLANGFAM_y
# fit linear model
jmodel_ex.music <- lm(formula = jmodel_ex.music_form, data = jury_data)
# investigate the R-sq value
summary(jmodel_ex.music)
```

```
##
## Call:
## lm(formula = jmodel_ex.music_form, data = jury_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0311 -1.1908 -0.0857  0.9681  3.6289
##
```

```
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.9945     0.4313   4.625 5.41e-06 ***
## VBlocs2_TC_4       1.7987     0.2744   6.554 2.19e-10 ***
## TC_PerfType_Solo   1.4758     0.4352   3.391 0.000781 ***
## ComVBlocs1_y      -0.5562     0.2239  -2.484 0.013495 *
## ComLANGFAM_y       0.3727     0.2045   1.823 0.069291 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.525 on 326 degrees of freedom
## Multiple R-squared:  0.1545,	Adjusted R-squared:  0.1442
## F-statistic:  14.9 on 4 and 326 DF,  p-value: 3.423e-11
```