Homework Week 11

Conor Craig

2023-11-14

Question 1

Estimate a constant-only model

```
women09 <- world$women09
women09_no_na <- na.omit (women09)
model_1 <- lm(women09 ~ 1)
summary (model_1)
```

```
##
## Call:
## lm(formula = women09 ~ 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.177  -7.477  -1.627   5.773  39.123
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.1772     0.8238   20.85   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.05 on 179 degrees of freedom
##   (11 observations deleted due to missingness)
```

Question 2

Per capita GDP (gdp_10_thou) as the main independent variable.

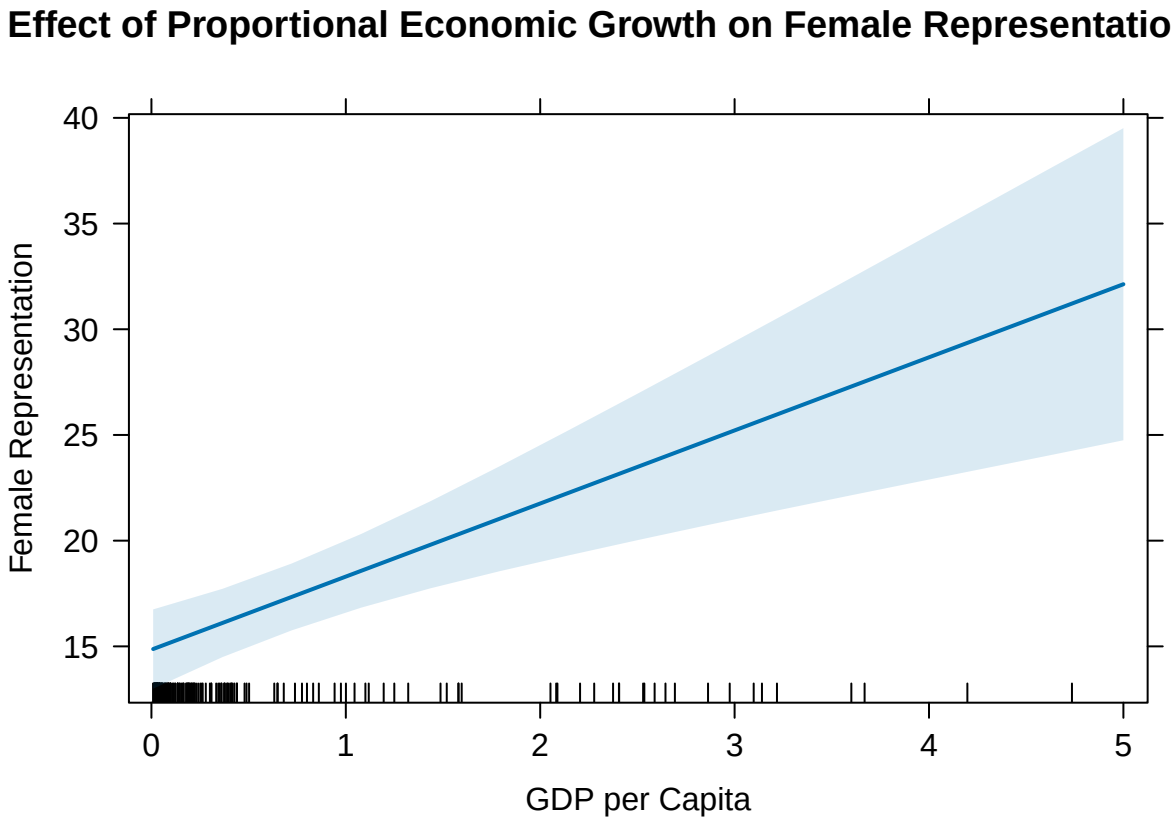
```r
gdp_10_thou <- world$gdp_10_thou
new_world <- data.frame (women09 = women09, gdp_10_thou = gdp_10_thou)
no_na_data <- na.omit(new_world)
regress_1 <- lm(women09 ~ gdp_10_thou, data = no_na_data)
summary(regress_1)
```

```
##
## Call:
## lm(formula = women09 ~ gdp_10_thou, data = no_na_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -24.74  -6.74  -1.62   5.78  41.38
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  14.8430     0.9542   15.56  < 2e-16 ***
## gdp_10_thou   3.4574     0.8351    4.14  5.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.38 on 167 degrees of freedom
## Multiple R-squared:  0.09308,    Adjusted R-squared:  0.08765
## F-statistic: 17.14 on 1 and 167 DF,  p-value: 5.501e-05
```

Question 3

Graph estimated effect of per capita on female representation

```r
effect_plot_1 <- effect(term = "gdp_10_thou", mod = regress_1)
plot (effect_plot_1,
main = "Effect of Proportional Economic Growth on Female Representation",
xlab = "GDP per Capita",
ylab = "Female Representation")
```



Question 4

Estimate electoral system as the main independent variable.

```
new_world $ pr_sys <- ifelse(world$pr_sys == "Yes", 1, 0)

no_na_data_2 <- na.omit(new_world)
no_na_data_2$pr_sys <- as.factor(no_na_data_2$pr_sys)
regress_pr_sys <- lm(women09 ~ pr_sys, data = no_na_data_2)
summary (regress_pr_sys)
```

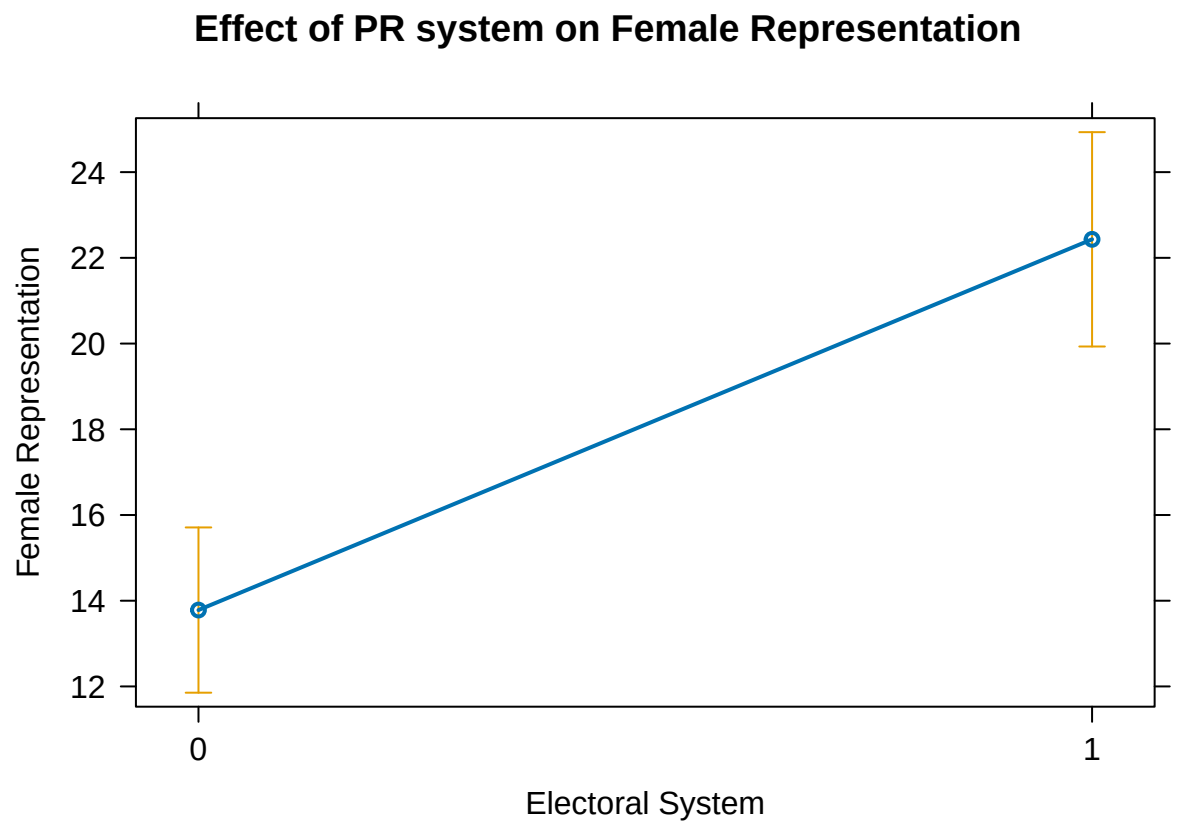
```
##
## Call:
## lm(formula = women09 ~ pr_sys, data = no_na_data_2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -16.632  -7.732  -2.082   7.018  33.868
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.7821     0.9765  14.114  < 2e-16 ***
## pr_sys1       8.6497     1.5993   5.408 2.17e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.05 on 167 degrees of freedom
## Multiple R-squared:  0.149,  Adjusted R-squared:  0.144
## F-statistic: 29.25 on 1 and 167 DF,  p-value: 2.172e-07
```

Question 5

Graph of estimated effect of electoral system on female representation

```
effect_plot_1 <- effect(term = "pr_sys", mod = regress_pr_sys)
plot (effect_plot_1,
main = "Effect of PR system on Female Representation",
xlab = "Electoral System",
ylab = "Female Representation")
```



Question 6

regression of per capita GDP and electoral system

```
regress_2 <- lm(women09 ~ pr_sys + gdp_10_thou, data = no_na_data_2)
summary (regress_2)
```

```
##
## Call:
## lm(formula = women09 ~ pr_sys + gdp_10_thou, data = no_na_data_2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.371  -7.872  -1.266   6.399  36.148
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.3922     1.0248  12.092  < 2e-16 ***
## pr_sys1       7.7007     1.5713   4.901 2.25e-06 ***
## gdp_10_thou   2.7864     0.7948   3.506 0.000585 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.73 on 166 degrees of freedom
## Multiple R-squared:  0.2077, Adjusted R-squared:  0.1982
## F-statistic: 21.76 on 2 and 166 DF,  p-value: 4.049e-09
```

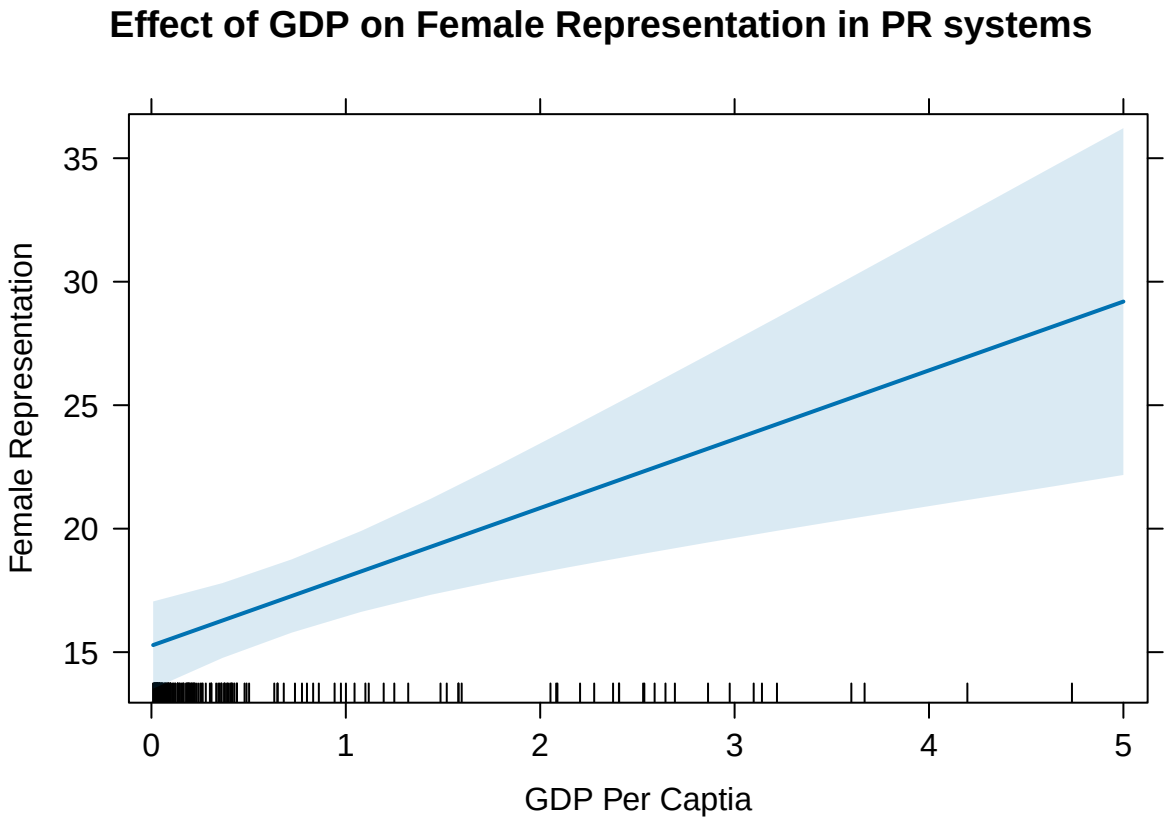
Quetion 7

Which is the best model of the four above?

Question 8

Graph of per capita GDP on female representation when considering PR systems

```
effect_plot_2 <- effect(term = "gdp_10_thou", mod = regress_2,
                        given.value = list(pr_sys = 1))
plot (effect_plot_2,
main = "Effect of GDP on Female Representation in PR systems",
xlab = "GDP Per Captia",
ylab = "Female Representation")
```



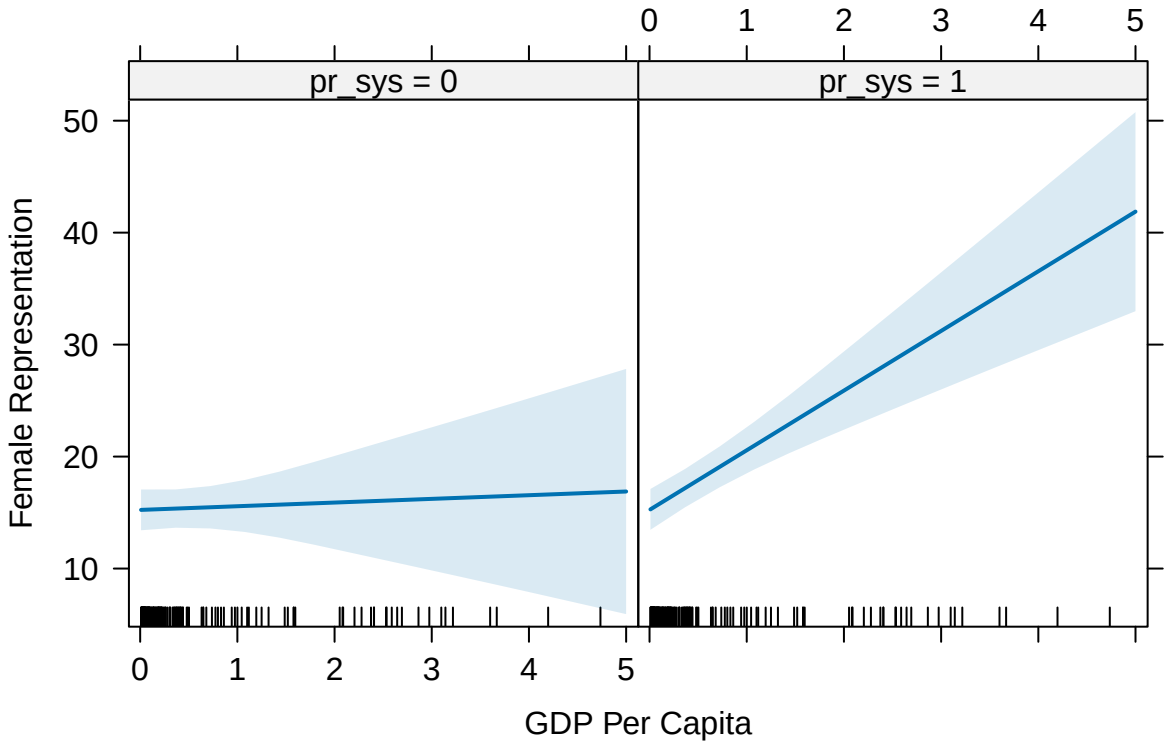
Question 9

Try creating the same graph by providing "gdp_10_thou:pr_sys" as the term.

```
regress_3 <- lm(women09 ~ gdp_10_thou:pr_sys, data = no_na_data_2)
effect_plot_3 <- effect(term = "gdp_10_thou:pr_sys", mod = regress_3)

plot (effect_plot_3,
main = "GDP Per Captia's Effect on Female Representation Given if PR system",
xlab = "GDP Per Capita",
ylab = "Female Representation")
```

GDP Per Captia's Effect on Female Representation Given if PR system



Question 10

Regression of region on female representation

```
new_world $ region <- world$region
no_na_data_3 <- na.omit (new_world)

region_regress <- lm(women09 ~ region, data = no_na_data_3)
summary (region_regress)
```

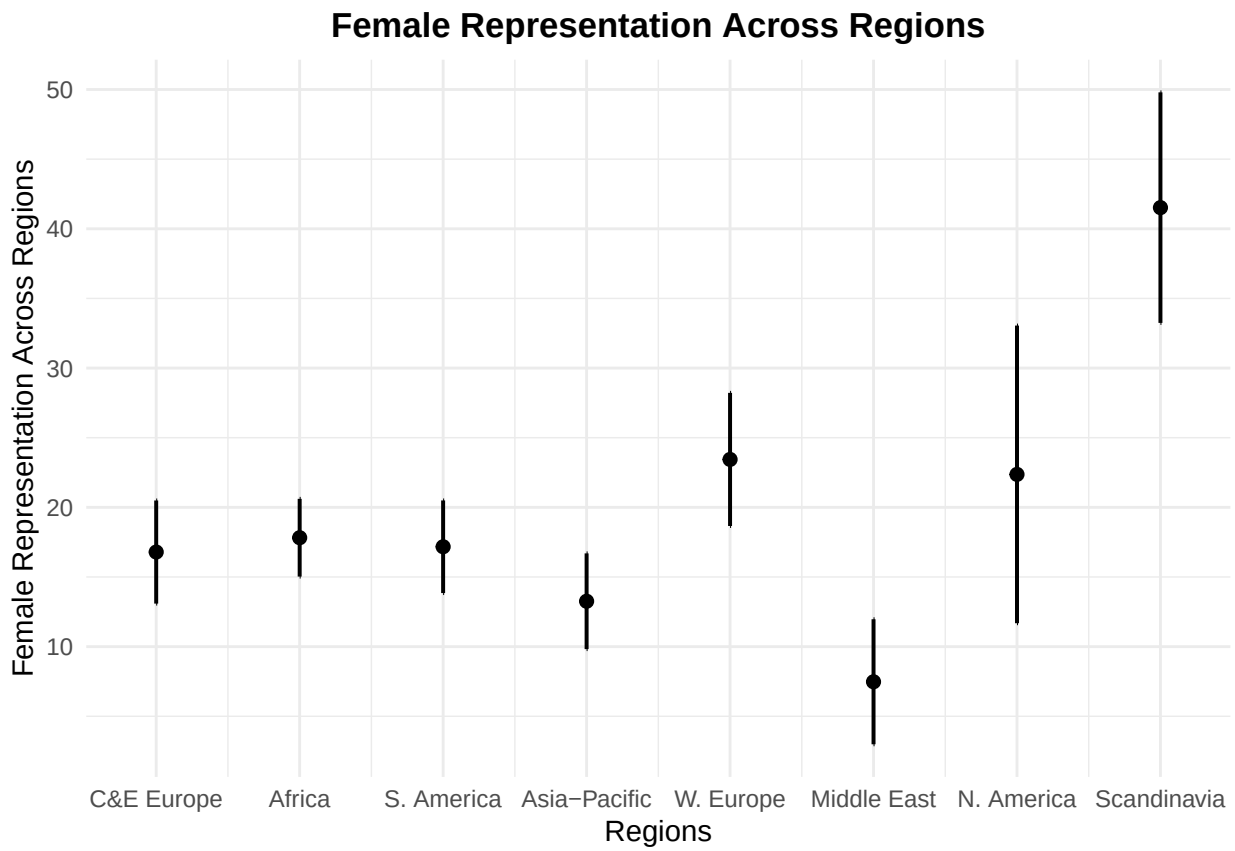
```
##
## Call:
## lm(formula = women09 ~ region, data = no_na_data_3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.171  -6.759  -1.959   4.924  38.480
##
## Coefficients:
```

```
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         17.8205     1.4129  12.613  < 2e-16 ***
## regionAsia-Pacific  -4.5618     2.2416  -2.035 0.043487 *
## regionC&E Europe    -1.0285     2.3472  -0.438 0.661861
## regionMiddle East  -10.3440     2.6764  -3.865 0.000161 ***
## regionN. America     4.5462     5.5923   0.813 0.417453
## regionS. America    -0.6495     2.1976  -0.296 0.767962
## regionScandinavia   23.6995     4.4230   5.358 2.86e-07 ***
## regionW. Europe      5.6195     2.8021   2.005 0.046587 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.372 on 161 degrees of freedom
## Multiple R-squared:  0.2871, Adjusted R-squared:  0.2561
## F-statistic: 9.261 on 7 and 161 DF,  p-value: 1.303e-09
```

Question 11

Graph of region on female representation.

```
library(ggeffects)
test_plot <- ggpredict(region_regress, terms = "region")
regional_plot <- plot(test_plot)+
  labs (title = "Female Representation Across Regions",
        x = "Regions",
        y = "Female Representation Across Regions") +
  theme_minimal() +
  theme (plot.title = element_text(hjust = 0.5, face = "bold"))
regional_plot
```



Question 12

Regression model of female representation on per cpaita GDP that controls for region

```
region_regress_2 <- lm(women09 ~ gdp_10_thou + region, data = no_na_data_3)
summary (region_regress_2)
```

```
##
## Call:
## lm(formula = women09 ~ gdp_10_thou + region, data = no_na_data_3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.123  -6.959  -1.845   4.348  38.563
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         17.7141     1.4194  12.480  < 2e-16 ***
## gdp_10_thou          1.0914     1.2707   0.859 0.391693
```

```
## regionAsia-Pacific  -4.9372     2.2856  -2.160 0.032254 *
## regionC&E Europe    -1.2315     2.3610  -0.522 0.602678
## regionMiddle East  -11.0675     2.8079  -3.942 0.000121 ***
## regionN. America     2.2842     6.1856   0.369 0.712410
## regionS. America    -0.9546     2.2279  -0.428 0.668890
## regionScandinavia   20.3981     5.8626   3.479 0.000648 ***
## regionW. Europe      3.1906     3.9828   0.801 0.424264
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.38 on 160 degrees of freedom
## Multiple R-squared:  0.2903, Adjusted R-squared:  0.2549
## F-statistic: 8.183 on 8 and 160 DF,  p-value: 2.991e-09
```

```
# GDP per capita does not appear to be an important determinat of female
# representation. When controlling for region, the variable is not
# statistically significant; indicating, we can't determine if the effect
# is different from zero
```

Question 13.

Regression model of female representation on frac_eth3

```
new_world $ frac_eth3 <- world$frac_eth3
new_world $ frac_high <- ifelse(world$frac_eth3 == "High", 1, 0)
new_world $ frac_mid <- ifelse(world$frac_eth3 == "Medium", 1, 0)
new_world $ frac_low <- ifelse(world$frac_eth3 == "Low", 1, 0)

no_na_data_4 <- na.omit (new_world)

eth_frac_regress <- lm(women09 ~ frac_mid + frac_high, data = no_na_data_4)
summary (eth_frac_regress)
```

```
##
## Call:
## lm(formula = women09 ~ frac_mid + frac_high, data = no_na_data_4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.785  -7.260  -1.435   5.090  39.863
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.7847     1.4160  12.560   <2e-16 ***
## frac_mid     -1.3482     2.0688  -0.652    0.516
## frac_high    -0.7993     2.0386  -0.392    0.696
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.88 on 163 degrees of freedom
## Multiple R-squared:  0.002652,   Adjusted R-squared:  -0.009585
## F-statistic: 0.2167 on 2 and 163 DF,  p-value: 0.8054
```

Question 14.

Do you think ethnic fractionalization has an impact on female representation?

```
# The results indicate that no, ethnic fractionalization has no effect on
# female representation. A substantive effect is reported, but it is impossible
# to tell if this actually differs from zero or not (that is, the findings
# lack statistical significance)

# This determination was made from the summary of the regression model above.
# Summary code not shown in this code chunk.
```