Sarah & Joe Present: Does time spent doing homework in adolescence increasing reading and math achievement?

Packages

```
library(mlr)
library(here)
library(purrr)
library(dplyr)
library(tidyr)
library(haven)
library(forcats)
library(ggplot2)
library(stringr)
library(sjlabelled)
```

Read in data and filter out all those who were not in the wave

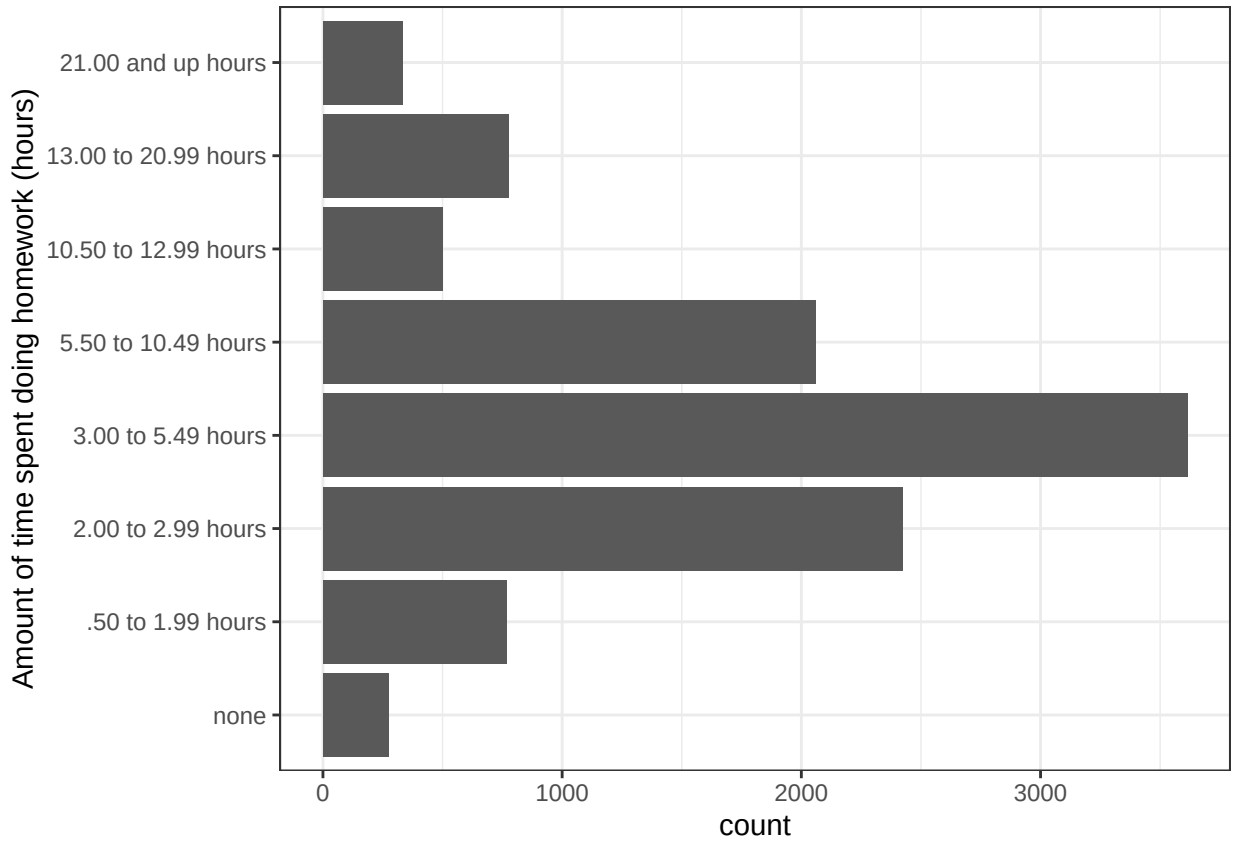
```
data <- read_dta(here("class data.dta"))

baseYearCohort <-
    data %>%
    select(matches("by|sex|race|^id$")) %>%
    filter(byhomewk != 99)
```

Number of students removed due to not being in the wave: 1397
Fraction of total students removed: 0.1150362

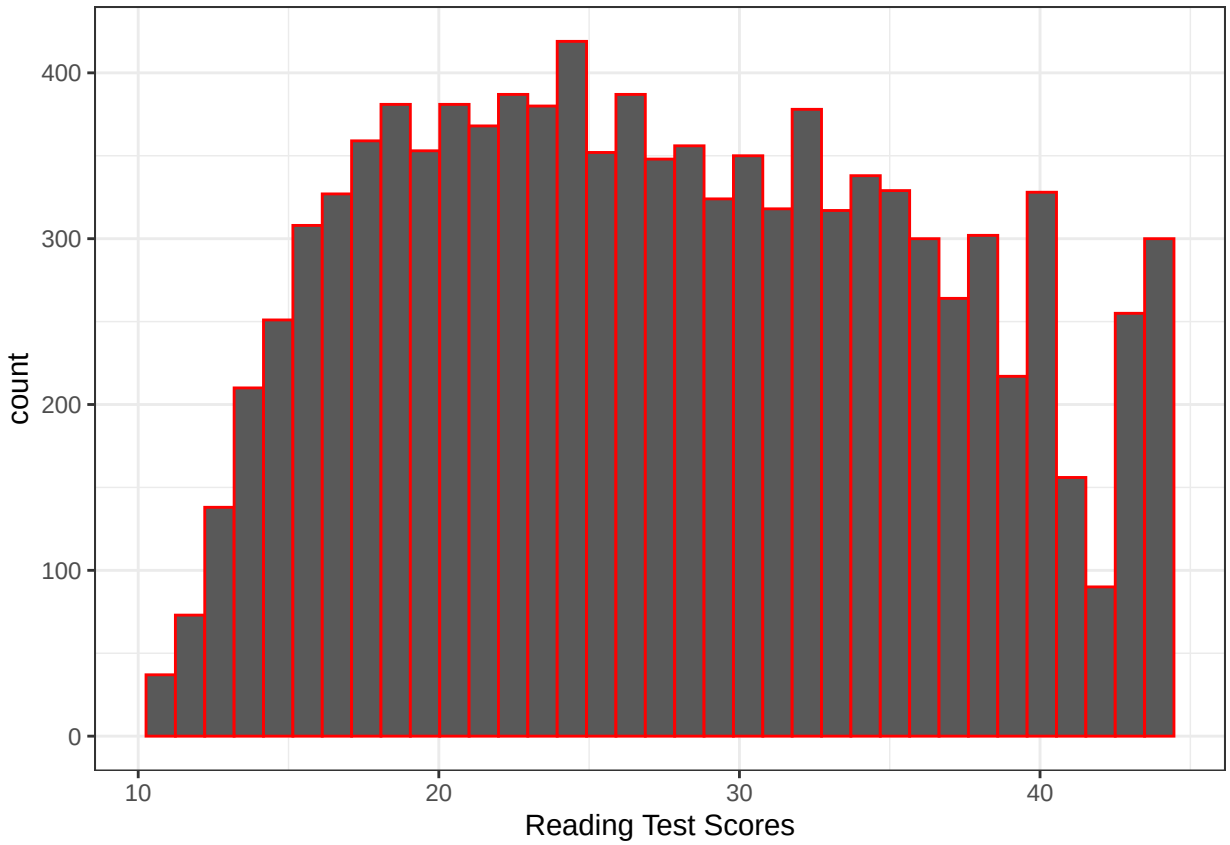
Get a sense for the data

```
ggplot(baseYearCohort, aes(x = as_label(byhomewk))) +
    geom_bar() +
    theme_bw() +
    coord_flip() +
    labs(x = "Amount of time spent doing homework (hours)")
```



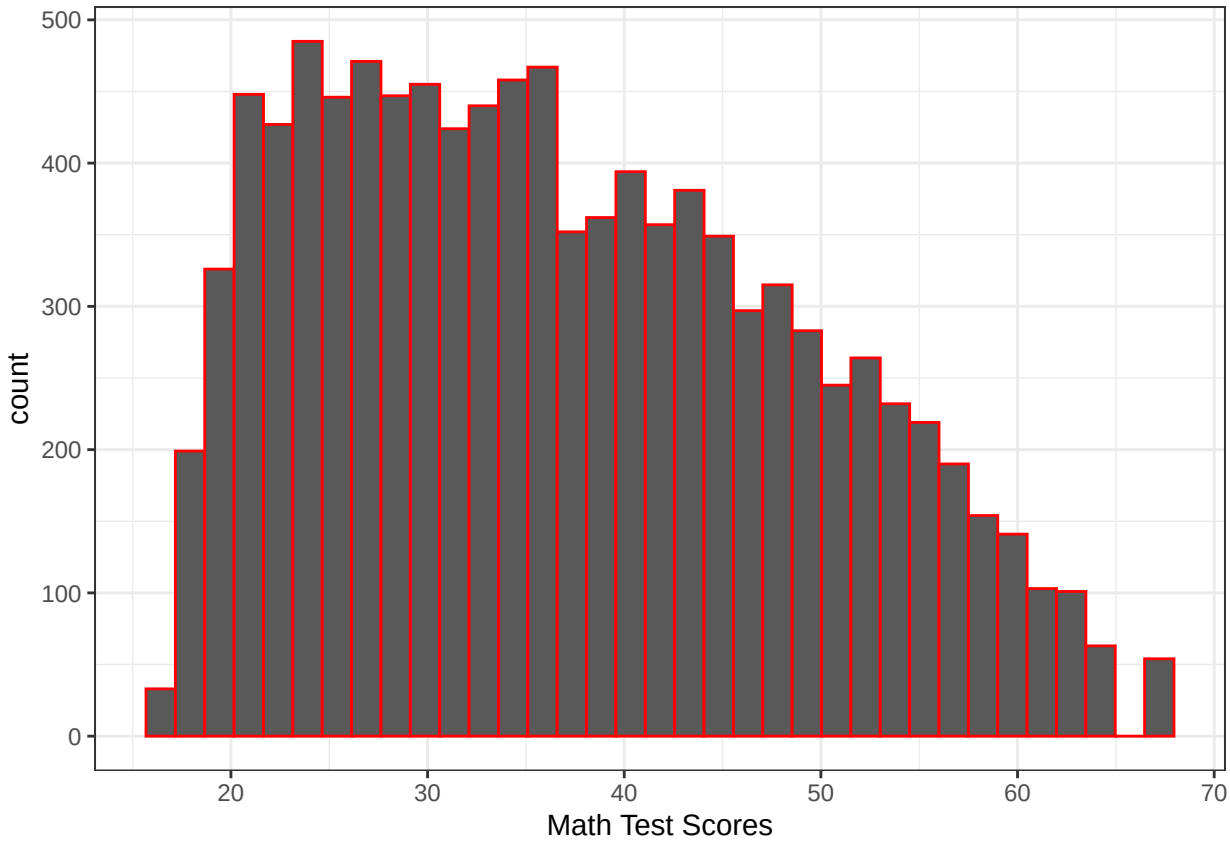

```
ggplot(baseYearCohort, aes(x = as.numeric(by2xrirr))) +
    geom_histogram(bins = 35, color = "red") +
    labs(x = "Reading Test Scores") +
    theme_bw()
```

```
## Warning: Removed 366 rows containing non-finite values (stat_bin).
```



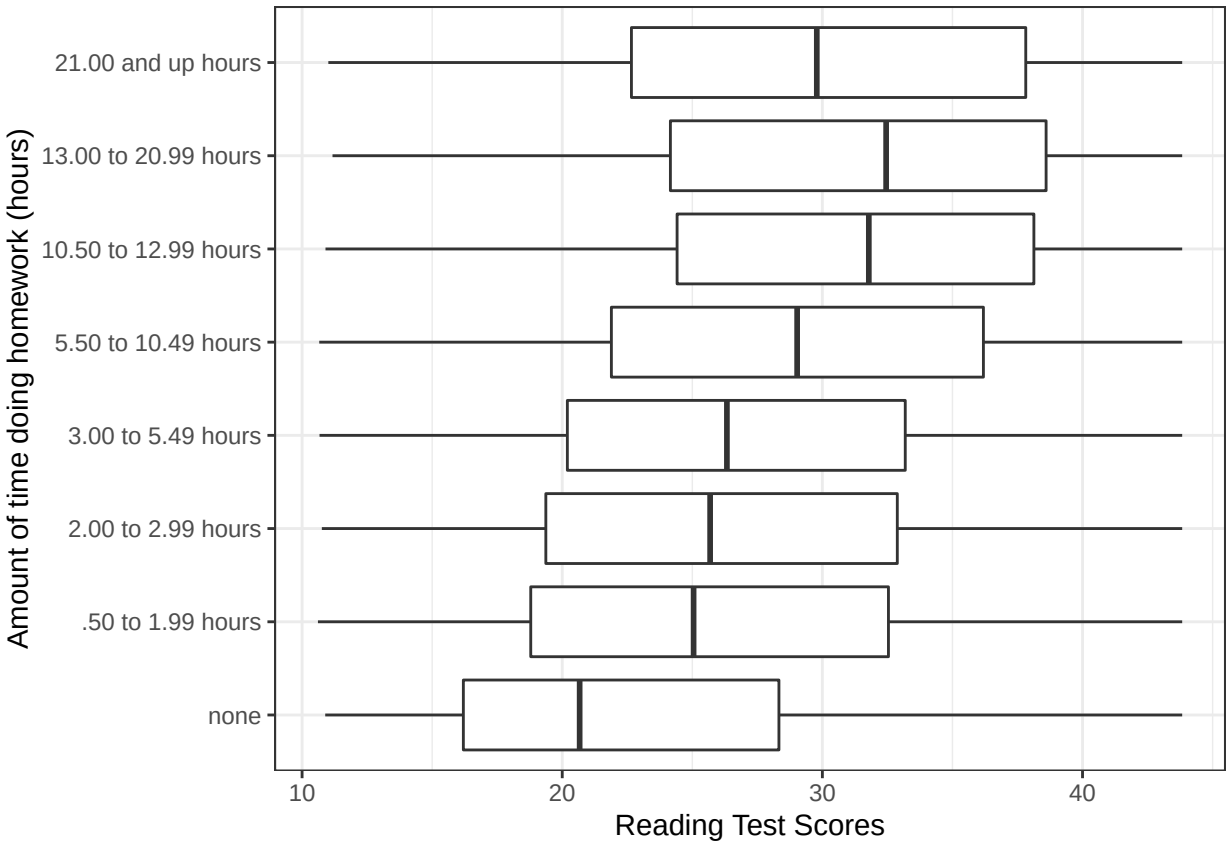

```
ggplot(baseYearCohort, aes(x = as.numeric(by2xmirr))) +
    geom_histogram(bins = 35, color = "red") +
    labs(x = "Math Test Scores") +
    theme_bw()
```

```
## Warning: Removed 365 rows containing non-finite values (stat_bin).
```



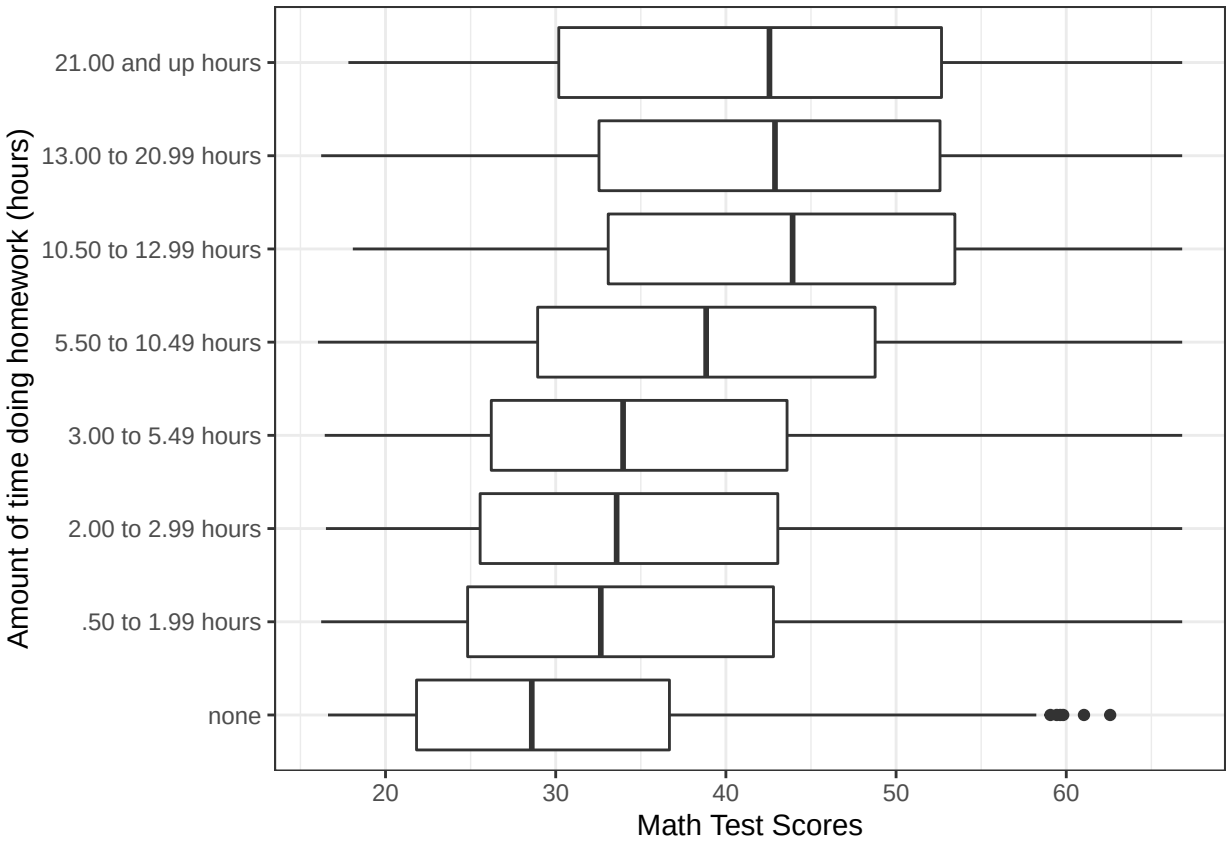

```
ggplot(baseYearCohort, aes(x = as_label(byhomewk), y = as.numeric(by2xrirr))) +
    geom_boxplot() +
    theme_bw() +
    coord_flip() +
    labs(x = "Amount of time doing homework (hours)", y = "Reading Test Scores")
```

```
## Warning: Removed 366 rows containing non-finite values (stat_boxplot).
```



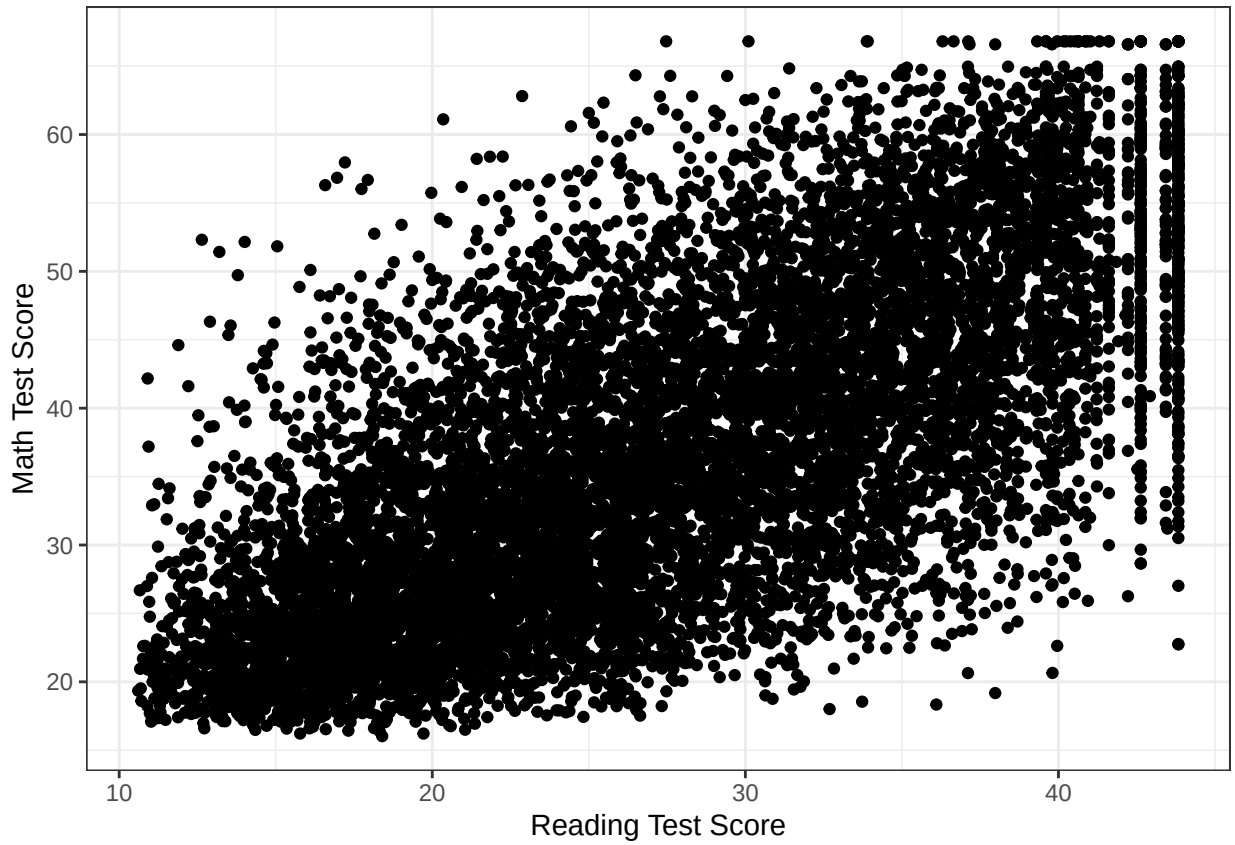

```
ggplot(baseYearCohort, aes(x = as_label(byhomewk), y = as.numeric(by2xmirr))) +
    geom_boxplot() +
    theme_bw() +
    coord_flip() +
    labs(x = "Amount of time doing homework (hours)", y = "Math Test Scores")
```

```
## Warning: Removed 365 rows containing non-finite values (stat_boxplot).
```




```
ggplot(baseYearCohort, aes(x = as.numeric(by2xrirr), y = as.numeric(by2xmirr))) +
    labs(x = "Reading Test Score", y = "Math Test Score") +
    geom_point() +
    theme_bw()
```

```
## Warning: Removed 380 rows containing missing values (geom_point).
```




```
cor(baseYearCohort$by2xrirr, baseYearCohort$by2xmirr, use = "complete.obs")
```

```
## [1] 0.7032982
```

Let's run some basic regressions

```
baseYearCohortWide <-
    baseYearCohort %>%
    mutate(byhomewk_label = as_label(byhomewk),
           byhomewk = as.numeric(as.logical(as.numeric(byhomewk)))) %>%
    pivot_wider(names_from = byhomewk_label,
                values_from = byhomewk,
                values_fill = F)

readingTestRegression <-
    lm(by2xrirr ~ `.50 to 1.99 hours` + `2.00 to 2.99 hours` +
           `3.00 to 5.49 hours` + `5.50 to 10.49 hours` +
           `10.50 to 12.99 hours` + `13.00 to 20.99 hours` +
           `21.00 and up hours`,
       data = baseYearCohortWide)
summary(readingTestRegression)
```

```
##
## Call:
## lm(formula = by2xrirr ~ `.50 to 1.99 hours` + `2.00 to 2.99 hours` +
##     `3.00 to 5.49 hours` + `5.50 to 10.49 hours` + `10.50 to 12.99 hours` +
```

```
##     `13.00 to 20.99 hours` + `21.00 and up hours`, data = baseYearCohortWide)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.0096  -6.8776  -0.4125   6.7024  20.9657
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              22.8643     0.5196  44.001  < 2e-16 ***
## `.50 to 1.99 hours`       3.0414     0.6050   5.027 5.07e-07 ***
## `2.00 to 2.99 hours`      3.5682     0.5482   6.508 7.95e-11 ***
## `3.00 to 5.49 hours`      4.0433     0.5388   7.504 6.70e-14 ***
## `5.50 to 10.49 hours`     6.0781     0.5534  10.984  < 2e-16 ***
## `10.50 to 12.99 hours`    8.0453     0.6474  12.428  < 2e-16 ***
## `13.00 to 20.99 hours`    8.2398     0.6058  13.602  < 2e-16 ***
## `21.00 and up hours`      6.8078     0.7077   9.620  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.459 on 10373 degrees of freedom
##   (366 observations deleted due to missingness)
## Multiple R-squared:  0.04154,    Adjusted R-squared:  0.0409
## F-statistic: 64.23 on 7 and 10373 DF,  p-value: < 2.2e-16
```

```
mathTestRegression <-
    lm(by2xmirr ~ `.50 to 1.99 hours` + `2.00 to 2.99 hours` +
           `3.00 to 5.49 hours` + `5.50 to 10.49 hours` +
           `10.50 to 12.99 hours` + `13.00 to 20.99 hours` +
           `21.00 and up hours`,
       data = baseYearCohortWide)
summary(mathTestRegression)
```

```
##
## Call:
## lm(formula = by2xmirr ~ `.50 to 1.99 hours` + `2.00 to 2.99 hours` +
##     `3.00 to 5.49 hours` + `5.50 to 10.49 hours` + `10.50 to 12.99 hours` +
##     `13.00 to 20.99 hours` + `21.00 and up hours`, data = baseYearCohortWide)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -26.349  -9.662  -1.080   8.684  32.325
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              30.8433     0.7170  43.019  < 2e-16 ***
## `.50 to 1.99 hours`       3.6421     0.8355   4.359 1.32e-05 ***
## `2.00 to 2.99 hours`      4.1789     0.7566   5.523 3.41e-08 ***
## `3.00 to 5.49 hours`      4.7171     0.7435   6.344 2.33e-10 ***
## `5.50 to 10.49 hours`     8.3756     0.7636  10.968  < 2e-16 ***
## `10.50 to 12.99 hours`   12.0024     0.8942  13.423  < 2e-16 ***
## `13.00 to 20.99 hours`   11.7257     0.8366  14.016  < 2e-16 ***
## `21.00 and up hours`     11.0460     0.9766  11.311  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 11.69 on 10374 degrees of freedom
##   (365 observations deleted due to missingness)
## Multiple R-squared:  0.05817,    Adjusted R-squared:  0.05753
## F-statistic: 91.52 on 7 and 10374 DF,  p-value: < 2.2e-16
```

Let's throw in some controls and make these regressions a bit more complicated

```
CreateWideControls <- function(df, test) {
    df %>%
        mutate(across(c(byhomewk, sex, race, bypared, byfaminc),
                      list(label = compose(as.character, as_label))),
               by2xrirr = as.numeric(by2xrirr),
               by2xmirr = as.numeric(by2xmirr)) %>%
        select(matches(paste0("label|^id$|", test))) %>%
        filter(across(everything(), ~!is.na(.x))) %>%
        mutate(byfaminc_label = if_else(byfaminc_label == "none",
                                        "$0",
                                        byfaminc_label)) %>%
        mutate(across(c(byhomewk_label, sex_label, race_label, bypared_label,
                        byfaminc_label),
                      as.factor))
}

RelocatePretty <- function(df) {
    df %>%
        relocate(".50.to.1.99.hours", "2.00.to.2.99.hours",
                 "3.00.to.5.49.hours", "5.50.to.10.49.hours",
                 "10.50.to.12.99.hours", "13.00.to.20.99.hours",
                 "21.00.and.up.hours", "less.than..1.000",
                 ".1.000....2.999", ".3.000....4.999", ".5.000....7.499",
                 ".7.500....9.999", ".10.000..14.999", ".15.000..19.999",
                 ".20.000..24.999", ".25.000..34.999", ".35.000..49.999",
                 ".50.000..74.999", ".75.000..99.999", ".100.000.199.999",
                 ".200.000.or.more", "..hs.....4yr.deg", "h.s..grad.or.ged",
                 "college.graduate", "m.a..equivalent", "ph.d...m.d...other") %>%
        select(-id, -black.not.hispanic, -male, -none, -didn.t.finish.hs, -`.0`)
}

# Reading Scores
readScoresWide <-
    CreateWideControls(baseYearCohort, "by2xrirr") %>%
    createDummyFeatures()

colnames(readScoresWide) <- str_replace(colnames(readScoresWide),
                                        "by.*_label.|race_label.|sex_label.",
                                        "")

readScoresWide <- RelocatePretty(readScoresWide)

# Math Scores
mathScoresWide <-
    CreateWideControls(baseYearCohort, "by2xmirr") %>%
    createDummyFeatures()
```

```r
colnames(mathScoresWide) <- str_replace(colnames(mathScoresWide),
                                        "by.*_label.|race_label.|sex_label.",
                                        "")

mathScoresWide <- RelocatePretty(mathScoresWide)
```

Number of rows removed due to missing values in the control variables: 1330
Fraction removed: 0.1237555

The math and reading scores differ by 1 in terms of missing values.

```r
readingTestRegressionControls <- lm(by2xrirr ~ ., data = readScoresWide)
summary(readingTestRegressionControls)
```

```
##
## Call:
## lm(formula = by2xrirr ~ ., data = readScoresWide)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.9812  -5.7091  -0.1828   5.7170  23.9387
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              13.3665     1.4060   9.507  < 2e-16 ***
## `.50.to.1.99.hours`       1.9674     0.5731   3.433 0.000600 ***
## `2.00.to.2.99.hours`      2.2867     0.5211   4.389 1.15e-05 ***
## `3.00.to.5.49.hours`      2.6679     0.5133   5.197 2.06e-07 ***
## `5.50.to.10.49.hours`     4.0680     0.5273   7.715 1.34e-14 ***
## `10.50.to.12.99.hours`    5.5650     0.6174   9.014  < 2e-16 ***
## `13.00.to.20.99.hours`    5.5530     0.5783   9.603  < 2e-16 ***
## `21.00.and.up.hours`      3.7900     0.6824   5.554 2.87e-08 ***
## less.than..1.000         -1.1449     1.5644  -0.732 0.464272
## `.1.000....2.999`         0.7805     1.4435   0.541 0.588737
## `.3.000....4.999`         2.0186     1.4133   1.428 0.153248
## `.5.000....7.499`         1.1078     1.3600   0.815 0.415374
## `.7.500....9.999`         1.6471     1.3520   1.218 0.223165
## `.10.000..14.999`         2.6244     1.3084   2.006 0.044898 *
## `.15.000..19.999`         2.8018     1.3094   2.140 0.032400 *
## `.20.000..24.999`         3.7696     1.3027   2.894 0.003817 **
## `.25.000..34.999`         4.1595     1.2932   3.216 0.001303 **
## `.35.000..49.999`         4.7112     1.2934   3.642 0.000272 ***
## `.50.000..74.999`         4.6418     1.3021   3.565 0.000366 ***
## `.75.000..99.999`         5.1583     1.3500   3.821 0.000134 ***
## `.100.000.199.999`        5.0039     1.3593   3.681 0.000233 ***
## `.200.000.or.more`        4.9441     1.4472   3.416 0.000638 ***
## ..hs.....4yr.deg          3.3669     0.3088  10.904  < 2e-16 ***
## h.s..grad.or.ged          1.7150     0.3311   5.180 2.26e-07 ***
## college.graduate          6.0249     0.3688  16.337  < 2e-16 ***
## m.a..equivalent           7.8615     0.4124  19.061  < 2e-16 ***
## ph.d...m.d...other        8.9555     0.5058  17.704  < 2e-16 ***
## female                    1.6613     0.1592  10.434  < 2e-16 ***
## amer.ind.ak.native       -0.3873     0.8831  -0.439 0.660982
## asian.pacific.islndr      3.0678     0.4185   7.331 2.48e-13 ***
```

```
## hispanic                 1.0098     0.3617   2.792 0.005247 **
## white.not.hispanic       3.1870     0.2976  10.708  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.636 on 9385 degrees of freedom
## Multiple R-squared:  0.2175, Adjusted R-squared:  0.2149
## F-statistic: 84.13 on 31 and 9385 DF,  p-value: < 2.2e-16
```

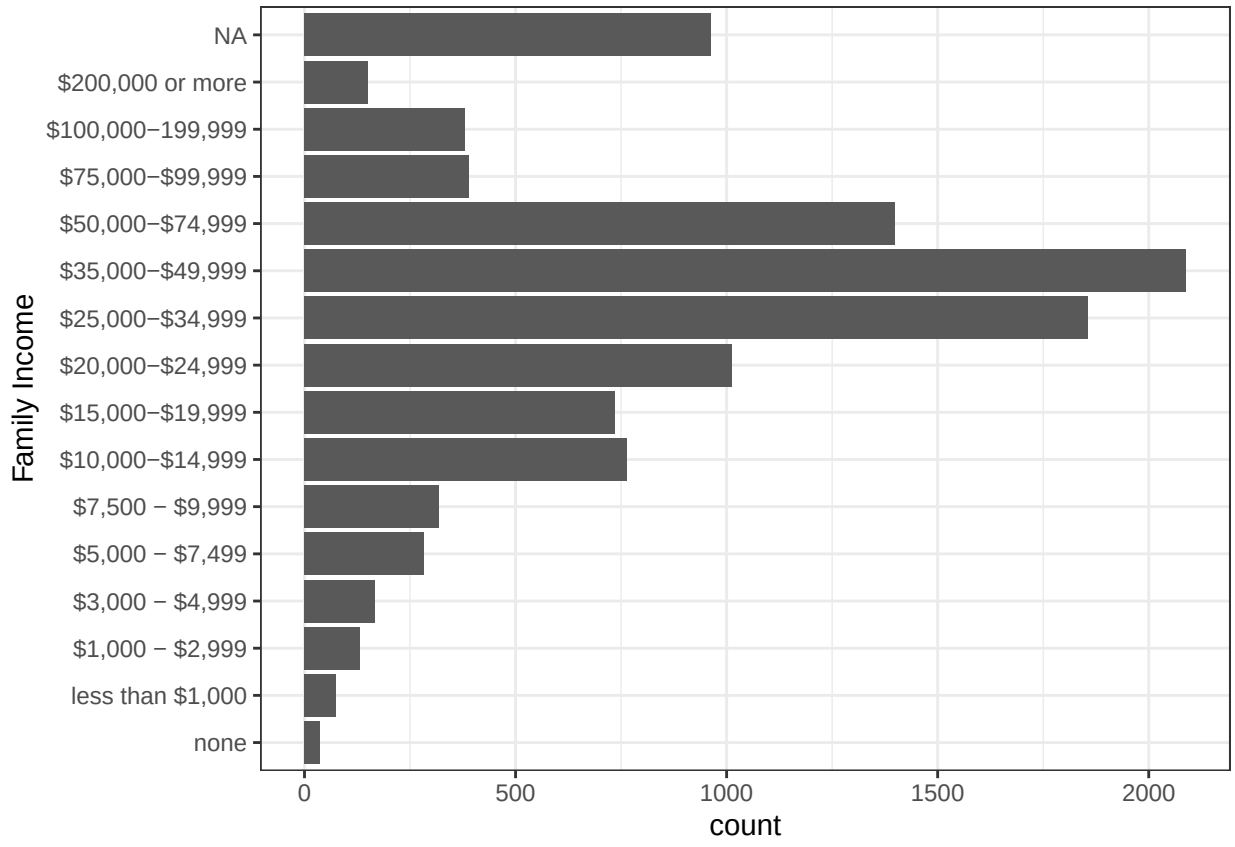
```
mathTestRegressionControls <- lm(by2xmirr ~ ., data = mathScoresWide)
summary(mathTestRegressionControls)
```

```
##
## Call:
## lm(formula = by2xmirr ~ ., data = mathScoresWide)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -33.140  -7.591  -0.734   7.182  41.526
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              18.43707    1.91536   9.626  < 2e-16 ***
## `.50.to.1.99.hours`       2.55852    0.77152   3.316 0.000916 ***
## `2.00.to.2.99.hours`      2.68930    0.70109   3.836 0.000126 ***
## `3.00.to.5.49.hours`      3.15862    0.69059   4.574 4.85e-06 ***
## `5.50.to.10.49.hours`     5.90061    0.70946   8.317  < 2e-16 ***
## `10.50.to.12.99.hours`    8.78759    0.83106  10.574  < 2e-16 ***
## `13.00.to.20.99.hours`    7.64216    0.77866   9.814  < 2e-16 ***
## `21.00.and.up.hours`      6.57541    0.91792   7.163 8.46e-13 ***
## less.than..1.000         -1.43566    2.12860  -0.674 0.500034
## `.1.000....2.999`         0.02738    1.96734   0.014 0.988896
## `.3.000....4.999`         2.17074    1.92708   1.126 0.260007
## `.5.000....7.499`         1.68679    1.85580   0.909 0.363411
## `.7.500....9.999`         1.70624    1.84520   0.925 0.355149
## `.10.000..14.999`         3.26858    1.78749   1.829 0.067493 .
## `.15.000..19.999`         3.75011    1.78889   2.096 0.036079 *
## `.20.000..24.999`         4.11680    1.78012   2.313 0.020763 *
## `.25.000..34.999`         4.97223    1.76759   2.813 0.004918 **
## `.35.000..49.999`         5.78632    1.76784   3.273 0.001068 **
## `.50.000..74.999`         6.32239    1.77937   3.553 0.000382 ***
## `.75.000..99.999`         7.81042    1.84312   4.238 2.28e-05 ***
## `.100.000.199.999`        8.22769    1.85542   4.434 9.34e-06 ***
## `.200.000.or.more`        8.65577    1.97254   4.388 1.16e-05 ***
## ..hs.....4yr.deg          4.39204    0.41631  10.550  < 2e-16 ***
## h.s..grad.or.ged          2.21530    0.44625   4.964 7.02e-07 ***
## college.graduate          8.14218    0.49728  16.374  < 2e-16 ***
## m.a..equivalent          11.32637    0.55603  20.370  < 2e-16 ***
## ph.d...m.d...other       12.38697    0.68202  18.162  < 2e-16 ***
## female                   -0.73636    0.21460  -3.431 0.000603 ***
## amer.ind.ak.native        0.18456    1.19027   0.155 0.876781
## asian.pacific.islndr      8.52191    0.56398  15.110  < 2e-16 ***
## hispanic                  1.98162    0.48694   4.070 4.75e-05 ***
## white.not.hispanic        5.51501    0.40050  13.770  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.29 on 9386 degrees of freedom
## Multiple R-squared:  0.2728, Adjusted R-squared:  0.2704
## F-statistic: 113.6 on 31 and 9386 DF,  p-value: < 2.2e-16
```

Let's recode by family income

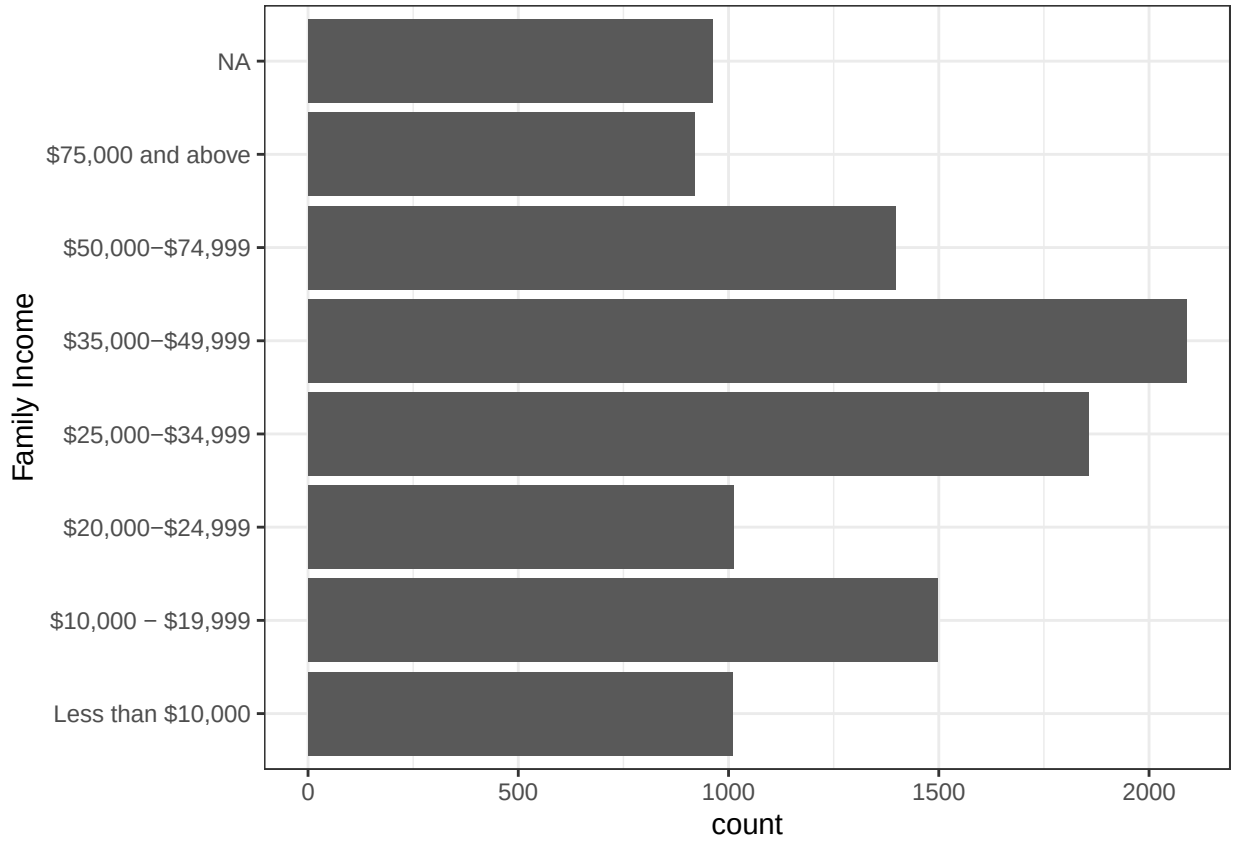
```
ggplot(baseYearCohort, aes(x = as_label(byfaminc))) +
    geom_bar() +
    theme_bw() +
    coord_flip() +
    labs(x = "Family Income")
```




```
baseYearCohortRecode <-
    baseYearCohort %>%
    mutate(byfaminc = as.character(as_label(byfaminc)),
           byfaminc = fct_recode(byfaminc,
                                 "Less than $10,000" = "none",
                                 "Less than $10,000" = "less than $1,000",
                                 "Less than $10,000" = "$1,000 - $2,999",
                                 "Less than $10,000" = "$3,000 - $4,999",
                                 "Less than $10,000"= "$5,000 - $7,499",
                                 "Less than $10,000" = "$7,500 - $9,999",
                                 "$10,000 - $19,999" = "$10,000-$14,999",
                                 "$10,000 - $19,999" = "$15,000-$19,999",
```

```
                                  "$75,000 and above" = "$75,000-$99,999",
                                  "$75,000 and above" = "$100,000-199,999",
                                  "$75,000 and above" = "$200,000 or more"),
           byfaminc = fct_relevel(byfaminc,
                                  "Less than $10,000",
                                  "$10,000 - $19,999",
                                  "$20,000-$24,999",
                                  "$25,000-$34,999",
                                  "$35,000-$49,999",
                                  "$50,000-$74,999",
                                  "$75,000 and above"))

ggplot(baseYearCohortRecode, aes(x = byfaminc)) +
    geom_bar() +
    theme_bw() +
    coord_flip() +
    labs(x = "Family Income")
```




```
CreateWideControls <- function(df, test) {
    df %>%
        mutate(across(c(byhomewk, sex, race, bypared, byfaminc),
                      list(label = compose(as.character, as_label))),
               by2xrirr = as.numeric(by2xrirr),
               by2xmirr = as.numeric(by2xmirr)) %>%
        select(matches(paste0("label|^id$|", test))) %>%
        filter(across(everything(), ~!is.na(.x))) %>%
```

```
        mutate(byfaminc_label = if_else(byfaminc_label == "none",
                                        "$0",
                                        byfaminc_label)) %>%
        mutate(across(c(byhomewk_label, sex_label, race_label, bypared_label,
                        byfaminc_label),
                      as.factor))
}

RelocatePrettyRecode <- function(df) {
    df %>%
        relocate(".50.to.1.99.hours", "2.00.to.2.99.hours",
                 "3.00.to.5.49.hours", "5.50.to.10.49.hours",
                 "10.50.to.12.99.hours", "13.00.to.20.99.hours",
                 "21.00.and.up.hours", ".10.000....19.999", ".20.000..24.999",
                 ".25.000..34.999", ".35.000..49.999", ".50.000..74.999",
                 ".75.000.and.above", "..hs.....4yr.deg", "h.s..grad.or.ged",
                 "college.graduate", "m.a..equivalent", "ph.d...m.d...other") %>%
        select(-id, -black.not.hispanic, -male, -none, -didn.t.finish.hs,
               -`Less.than..10.000`)
}

# Reading Scores
readScoresWideRecode <-
    CreateWideControls(baseYearCohortRecode, "by2xrirr") %>%
    createDummyFeatures()

colnames(readScoresWideRecode) <- str_replace(colnames(readScoresWideRecode),
                                        "by.*_label.|race_label.|sex_label.",
                                        "")

readScoresWideRecode <- RelocatePrettyRecode(readScoresWideRecode)

readingTestRegressionControlsRecode <- lm(by2xrirr ~ .,
                                          data = readScoresWideRecode)
summary(readingTestRegressionControlsRecode)
```

```
##
## Call:
## lm(formula = by2xrirr ~ ., data = readScoresWideRecode)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.9019  -5.6995  -0.1516   5.7411  24.4293
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            14.5129     0.6264  23.167  < 2e-16 ***
## `.50.to.1.99.hours`     1.9679     0.5729   3.435 0.000595 ***
## `2.00.to.2.99.hours`    2.2902     0.5209   4.396 1.11e-05 ***
## `3.00.to.5.49.hours`    2.6715     0.5133   5.205 1.98e-07 ***
## `5.50.to.10.49.hours`   4.0741     0.5273   7.727 1.21e-14 ***
## `10.50.to.12.99.hours`  5.5744     0.6172   9.032  < 2e-16 ***
## `13.00.to.20.99.hours`  5.5500     0.5782   9.600  < 2e-16 ***
## `21.00.and.up.hours`    3.7890     0.6823   5.553 2.88e-08 ***
```

```
## `.10.000....19.999`      1.5340     0.3254   4.715 2.46e-06 ***
## `.20.000..24.999`        2.5862     0.3602   7.179 7.53e-13 ***
## `.25.000..34.999`        2.9740     0.3242   9.173  < 2e-16 ***
## `.35.000..49.999`        3.5242     0.3263  10.800  < 2e-16 ***
## `.50.000..74.999`        3.4540     0.3595   9.608  < 2e-16 ***
## `.75.000.and.above`      3.8759     0.4214   9.198  < 2e-16 ***
## ..hs.....4yr.deg         3.3974     0.3081  11.027  < 2e-16 ***
## h.s..grad.or.ged         1.7487     0.3302   5.295 1.21e-07 ***
## college.graduate         6.0639     0.3680  16.477  < 2e-16 ***
## m.a..equivalent          7.8967     0.4119  19.171  < 2e-16 ***
## ph.d...m.d...other       8.9763     0.5025  17.864  < 2e-16 ***
## female                   1.6508     0.1591  10.373  < 2e-16 ***
## amer.ind.ak.native      -0.4818     0.8826  -0.546 0.585138
## asian.pacific.islndr     3.0435     0.4177   7.286 3.46e-13 ***
## hispanic                 1.0162     0.3610   2.815 0.004886 **
## white.not.hispanic       3.2023     0.2970  10.781  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.637 on 9393 degrees of freedom
## Multiple R-squared:  0.2165, Adjusted R-squared:  0.2146
## F-statistic: 112.9 on 23 and 9393 DF,  p-value: < 2.2e-16
```

```r
# Math Scores
mathScoresWideRecode <-
    CreateWideControls(baseYearCohortRecode, "by2xmirr") %>%
    createDummyFeatures()

colnames(mathScoresWideRecode) <- str_replace(colnames(mathScoresWideRecode),
                                      "by.*_label.|race_label.|sex_label.",
                                      "")

mathScoresWideRecode <- RelocatePrettyRecode(mathScoresWideRecode)

mathTestRegressionControlsRecode <- lm(by2xmirr ~ ., data = mathScoresWideRecode)
summary(mathTestRegressionControlsRecode)
```

```
##
## Call:
## lm(formula = by2xmirr ~ ., data = mathScoresWideRecode)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -33.517  -7.637  -0.715   7.199  42.005
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             19.64781    0.84309  23.304  < 2e-16 ***
## `.50.to.1.99.hours`      2.53524    0.77126   3.287 0.001016 **
## `2.00.to.2.99.hours`     2.68635    0.70093   3.833 0.000128 ***
## `3.00.to.5.49.hours`     3.15617    0.69051   4.571 4.92e-06 ***
## `5.50.to.10.49.hours`    5.90086    0.70940   8.318  < 2e-16 ***
## `10.50.to.12.99.hours`   8.79682    0.83086  10.588  < 2e-16 ***
## `13.00.to.20.99.hours`   7.63895    0.77850   9.812  < 2e-16 ***
## `21.00.and.up.hours`     6.58368    0.91784   7.173 7.89e-13 ***
```

```
## `.10.000....19.999`     2.23956    0.43850   5.107 3.33e-07 ***
## `.20.000..24.999`       2.84071    0.48578   5.848 5.15e-09 ***
## `.25.000..34.999`       3.69246    0.43723   8.445  < 2e-16 ***
## `.35.000..49.999`       4.50212    0.43997  10.233  < 2e-16 ***
## `.50.000..74.999`       5.03291    0.48456  10.387  < 2e-16 ***
## `.75.000.and.above`     6.81164    0.56811  11.990  < 2e-16 ***
## ..hs.....4yr.deg        4.44126    0.41542  10.691  < 2e-16 ***
## h.s..grad.or.ged        2.26285    0.44513   5.084 3.78e-07 ***
## college.graduate        8.19643    0.49629  16.516  < 2e-16 ***
## m.a..equivalent        11.37179    0.55535  20.477  < 2e-16 ***
## ph.d...m.d...other     12.51155    0.67749  18.467  < 2e-16 ***
## female                 -0.74660    0.21449  -3.481 0.000502 ***
## amer.ind.ak.native      0.07383    1.18962   0.062 0.950515
## asian.pacific.islndr    8.49866    0.56298  15.096  < 2e-16 ***
## hispanic                2.00073    0.48609   4.116 3.89e-05 ***
## white.not.hispanic      5.55500    0.39973  13.897  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.3 on 9394 degrees of freedom
## Multiple R-squared:  0.272,  Adjusted R-squared:  0.2702
## F-statistic: 152.6 on 23 and 9394 DF,  p-value: < 2.2e-16
```