Lab 06 - Modeling Course Evaluations

Due: Thursday, Mar 05 at 11:59pm

Team gitData - Nathan Kim, Blossom Mojekwu, Avanti Shah

Packages

```
library(tidyverse)
library(ggpol)
library(broom)
```

Data

```
evals <- read_csv("data/evals-mod.csv")
```

Exercises

Exercise 1

```
evals <- evals %>%
  rowwise() %>%
  mutate(bty_avg = mean(c(bty_f1lower, bty_f1upper, bty_f2upper, bty_m1lower,
                         bty_m1upper, bty_m2upper))) %>%
  ungroup()
```

Exercise 2

```
professor_eval <- lm(score ~ bty_avg, data = evals)
  tidy(professor_eval) %>%
  select(term, estimate)
```

```
# A tibble: 2 x 2
  term        estimate
  <chr>          <dbl>
1 (Intercept)   3.88
2 bty_avg       0.0666
```

```
glance(professor_eval) %>%
  pull(r.squared)
```

```
[1] 0.03502322
```

```
glance(professor_eval) %>%
  pull(adj.r.squared)
```

```
[1] 0.03292999
```

The linear model that predicts the average professor evaluation score based on average beauty rating is:

$\widehat{score}$ = 3.88 + (0.066)*(bty_avg)

For this data set, the R^2 value is 0.350 while the adjusted R^2 value is 0.033.

Exercise 3

```
lm_score <- lm(score ~ bty_avg + gender, data = evals)
tidy(lm_score) %>%
  select(term, estimate)
```

```
# A tibble: 3 x 2
  term        estimate
  <chr>          <dbl>
1 (Intercept)   3.75
2 bty_avg       0.0742
3 gendermale    0.172
```

The expected score is as follows:

Expected Score = 3.747 + 0.0742*(bty_avg) + 0.172*(gendermale)

```
glance(lm_score) %>%
  pull(r.squared)
```

```
[1] 0.05912354
```

```
glance(lm_score) %>%
  pull(adj.r.squared)
```

```
[1] 0.05503277
```

The R^2 value is 0.059 and the adjusted R^2 value is 0.055.

Exercise 4

In this linear model, the intercept is approximately 3.747. This means that if the professor is female and the average beauty rating is 0, then the expected evaluation score will be 3.747. The slope is 0.0742, meaning that for every unit increase in average beauty rating, the expected evaluation score will increase by 0.0742. The positive slope means that there is a positive correlation between average beauty rating and evaluation score.

Exercise 5

The equation of the line corresponding to male professors would be as follows:

$\widehat{expectedScore}$ = 3.920 + 0.0742*(bty_avg)

This increase in the intercept is accounted for by the gender: male corresponding to 1, or a binary value of "yes".

Exercise 6

If two professors were to get the same beauty rating, male professors would on average have a higher course evaluation score. This is shown when comparing the two equations above. Men would automatically have a higher intercept, meaning equal beauty ratings would result in men having the higher overall score.

Exercise 7

How does the relationship between beauty and evaluation score vary between male and female professors?

```r
male_professors <- evals %>%
  filter(gender == "male")

female_professors <- evals %>%
  filter(gender == "female")
```

```r
female_bty_rank <- lm(score ~ bty_avg, data = female_professors)
female_bty_rank %>%
  tidy() %>%
  mutate(female_estimate = round(estimate, 4)) %>%
  select(term, female_estimate)
```

```
# A tibble: 2 x 2
  term        female_estimate
  <chr>                 <dbl>
1 (Intercept)          3.95
2 bty_avg              0.0306
```

```r
male_bty_rank <- lm(score ~ bty_avg, data = male_professors)
male_bty_rank %>%
  tidy() %>%
  mutate(male_estimate = round(estimate, 4)) %>%
  select(term, male_estimate)
```

```
# A tibble: 2 x 2
  term        male_estimate
  <chr>               <dbl>
1 (Intercept)         3.77
2 bty_avg             0.110
```

From this analysis, we can see that for female professors, the slope of the linear model relationship between score and beauty evaluation is 0.0306. This indicates a weak positive relationship, showing that for every increment in `bty_avg`, there is a 0.0306 increase in evaluation score. The linear model relationship for male professors has a slope of 0.1103, which indicates a stronger positive relationship. As a result, we can see that the relationship between beauty and evaluation score is stronger for male professors than female professors.

Exercise 8

The adjusted R^2 in Exercise 2 is 0.03292999 and it means that the variability of the average professor evaluation is roughly 3.29% based on average beauty rating (bty_avg) only. The adjusted R^2 in Exercise 3 is 0.05503277 and it reflects the variability of the average professor evaluation as roughly 5.50% based on average beauty rating and gender together, representing an increase of 2.21% in adjusted R^2 when gender is added to the model. Since adjusted R^2 increases when gender is added, gender relates to score and provides useful new information to the model.

Exercise 9

`professor_eval: score = 3.88033262 + 0.06663701*(bty_avg)   lm_score: score = 3.74733459 + 0.07415493*(bty_avg) + 0.17238871*(gender)`

The slope for `bty_avg` under the linear model professor_eval is 0.06664 and the slope under `lm_score` is 0.07415. The addition of `gender` to the model increased the parameter estimate(`slope`) for `bty_avg` because the consideration of `gender` weighed less than `bty_avg`.

Exercise 10

```
m_bty_rank <- lm(score ~ rank + bty_avg, data = evals)

m_bty_rank %>%
  tidy() %>%
  mutate(estimate = round(estimate, 4)) %>%
  select(term, estimate)
```

```
# A tibble: 4 x 2
  term             estimate
  <chr>               <dbl>
1 (Intercept)        3.98
2 ranktenure track  -0.161
3 ranktenured       -0.126
4 bty_avg            0.0678
```

where `b0 = 3.9815`; `b1 = -0.1607`; `b2 = -0.1262`; `b3 = 0.0678`

Linear Model:

$\widehat{score}$ = `3.9815 - 0.1607*(ranktenure track) - 0.1262*(ranktenured) + 0.0678*(bty_avg)`

`3.9815` is the intercept which represents the score when the rank is teaching and the average beauty rating is set to 0.

`-0.1607` means that the score has an average decrease of 0.1607 point if a professor's rank is on the tenure track.

`-0.1262` means that the score has an average decrease of 0.1262 point if a professor's rank is tenure.

`0.0678` represents an average increase of 0.0678 point in score for every additional average beauty rating.

Exercise 11

We would expect **`cls_students`** alone to be the worst predictor of evaluation score because it is unlikely that students considerably factor in the total number of students in class when evaluating a professor, and other factors such as **`ethnicity`** and **`age`** are expected to bias students more.

Exercise 12

```
lm_cls_students <- lm(score ~ cls_students, data = evals)

lm_cls_students %>%
  tidy() %>%
  select(term, estimate)
```

```
# A tibble: 2 x 2
  term         estimate
  <chr>           <dbl>
1 (Intercept)  4.16
2 cls_students 0.000188
```

```
evals %>%
  summarise(mean = mean(score))
```

```
# A tibble: 1 x 1
   mean
  <dbl>
```

```
1  4.17
```

Our suspicion is warranted because the average professor evaluation score is approximately equal to the intercept of the model `lm_cls_did_eval`. This shows that that when the total number of students in a class (`cls_student`) is 0, the average score with for all possible class sizes is almost equal. The slope for `cls_students` is also very low at 0.019% showing a small predication estimate for professor evaluation score.

Exercise 13

We would not include `cls_did_eval`, because since we are already included the percentage who completed the evaluation and the total number of students in the class, it would be redundant to add a variable that contributes the same information as the combination of the other two. Since `cls_perc_eval` and `cls_students` can be multiplied to get `cls_did_eval`, it would not add any useful additional information to the model.

Exercise 14

```
lm_full_model <- lm(score ~ rank + ethnicity + language + age + cls_perc_eval
                    + cls_students + cls_level + cls_profs + cls_credits +
                      bty_avg, data = evals)

lm_full_model %>%
  tidy() %>%
  mutate(myestimate = round(estimate, 3)) %>%
  select(term, myestimate)
```

```
# A tibble: 12 x 2
   term                  myestimate
   <chr>                      <dbl>
 1 (Intercept)                3.51
 2 ranktenure track          -0.106
 3 ranktenured               -0.009
 4 ethnicitynot minority      0.221
 5 languagenon-english       -0.101
 6 age                       -0.005
 7 cls_perc_eval              0.006
 8 cls_students               0.001
 9 cls_levelupper             0.005
10 cls_profssingle            0.004
11 cls_creditsone credit      0.557
12 bty_avg                    0.059
```

$\widehat{score}$ = 3.512 - 0.106*(ranktenure track) - 0.009*(ranktenured) + 0.221*(ethnicitynot minority) - 0.101*(languagenon-english) - 0.005*(age) + 0.006(cls_perc_eval) + 0.001(cls_students) + 0.005(cls_levelupper) + 0.004(cls_profssingle) + 0.557*(cls_creditsone credit) + 0.059(bty_avg)

Exercise 15

```
backward_full <- lm_full_model %>%
  step(direction = "backward")
```

```
backward_full %>%
  tidy()
```

```
# A tibble: 6 x 5
```

```
  term                  estimate std.error statistic  p.value
  <chr>                    <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)           3.17      0.151        21.0  4.74e-69
2 ethnicitynot minority 0.255     0.0714        3.57 3.90e- 4
3 cls_perc_eval         0.00591   0.00156       3.79 1.68e- 4
4 cls_students          0.000607  0.000343      1.77 7.75e- 2
5 cls_creditsone credit 0.573     0.105         5.44 8.64e- 8
6 bty_avg               0.0632    0.0160        3.95 9.04e- 5
```

```
glance(lm_full_model) %>%
  pull(r.squared)
```

```
[1] 0.1411963
```

```
glance(lm_full_model) %>%
  pull(adj.r.squared)
```

```
[1] 0.1202498
```

Answer:

$\widehat{bestScoreModel} =$ `3.169 - 0.2553*(ethnicitynot minority)` + `0.0059(cls_perc_eval) + 0.0006(cls_students)` + `0.573*(cls_creditsone credit)` + `0.063(bty_avg)`

For the full model, R^2 is 0.1411963 and adjusted R^2 is 0.1202498.

Exercise 16

`cls_perc_eval` is a continuous variable with a slope of 0.0059, which means that for every additional unit increase in `cls_perc_eval`, there is an *increment* of 0.0059 to `score` (holding everything else constant). This displays a weak positive relationship. `ethnicity` is a categorical variable with two categories: minority or not minority. The slope of this variable is 0.2553, which means which means that for every additional unit increase in `ethnicity`, there is an *increment* of 0.2553 to `score` (holding everything else constant), indicating a stronger relationship. Since this is a categorical variable, this increment depends on the factor values of `ethnicity` and the baseline, so it can mean that the difference in `score` for minority and non-minority professors differ by 0.2553 if every other variable is held constant.

Exercise 17

We would assess the linearity assumption using a diagnostic plot of residuals against fitted values and observe whether the spread of the values on either side of `residual = 0` is somewhat uniform, with points are that are random and uncorrelated. If this is observed, we can say that there is homoscedasticity and the linearity assumption is satisfied. Since this model is specifically a linear model assessing the **linear** relationship between the variables, if the linearity assumption is severely violated, it would not make sense to use a linear model to predict values, and we would instead have to use a more appropriate type of regression model to extrapolate values accurately.

Exercise 18

Since this survey only included students at UT Austin, we would not be comfortable making predictions for professors at any school. This is because since this was a convenience sample (chosen at only one school since that is an inexpensive and convenient way to recruit participants). However, we do not know if this sample is representative of the population (whether this is Texan universities, public universities, or American universities; let alone any univeristy anywhere). Moreover, we are told that only "six students rated the professors' physical appearance", which itself is a subjective and biased measure and can change simply if these six students change. As a result, we are unaware of other confounding factors that potentially exist at this university that contributed to the results, so unless we know that the sample is representative, we cannot use this model for professors at any university.