

Multiple linear regression

Jimmy Ng

November 2, 2018

Grading the professor

Many college courses conclude by giving students the opportunity to evaluate the course and the instructor anonymously. However, the use of these student evaluations as an indicator of course quality and teaching effectiveness is often criticized because these measures may reflect the influence of non-teaching related characteristics, such as the physical appearance of the instructor. The article titled, "Beauty in the classroom: instructors' pulchritude and putative pedagogical productivity" (Hamermesh and Parker, 2005) found that instructors who are viewed to be better looking receive higher instructional ratings. (Daniel S. Hamermesh, Amy Parker, Beauty in the classroom: instructors pulchritude and putative pedagogical productivity, *Economics of Education Review*, Volume 24, Issue 4, August 2005, Pages 369-376, ISSN 0272-7757, 10.1016/j.econedurev.2004.07.013. <http://www.sciencedirect.com/science/article/pii/S0272775704001165>.)

In this lab we will analyze the data from this study in order to learn what goes into a positive professor evaluation.

The data

The data were gathered from end of semester student evaluations for a large sample of professors from the University of Texas at Austin. In addition, six students rated the professors' physical appearance. (This is a slightly modified version of the original data set that was released as part of the replication data for *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Gelman and Hill, 2007).) The result is a data frame where each row contains a different course and columns represent variables about the courses and professors.

```
load("more/evals.RData")
```

variable	description
score	average professor evaluation score: (1) very unsatisfactory - (5) excellent.
rank	rank of professor: teaching, tenure track, tenured.
ethnicity	ethnicity of professor: not minority, minority.
gender	gender of professor: female, male.
language	language of school where professor received education: english or non-english.
age	age of professor.
cls_perc_eval	percent of students in class who completed evaluation.

cls_did_eval	number of students in class who completed evaluation.
cls_students	total number of students in class.
cls_level	class level: lower, upper.
cls_profs	number of professors teaching sections in course in sample: single, multiple.
cls_credits	number of credits of class: one credit (lab, PE, etc.), multi credit.
bty_f1lower	beauty rating of professor from lower level female: (1) lowest - (10) highest.
bty_f1upper	beauty rating of professor from upper level female: (1) lowest - (10) highest.
bty_f2upper	beauty rating of professor from second upper level female: (1) lowest - (10) highest.
bty_m1lower	beauty rating of professor from lower level male: (1) lowest - (10) highest.
bty_m1upper	beauty rating of professor from upper level male: (1) lowest - (10) highest.
bty_m2upper	beauty rating of professor from second upper level male: (1) lowest - (10) highest.
bty_avg	average beauty rating of professor.
pic_outfit	outfit of professor in picture: not formal, formal.
pic_color	color of professor's picture: color, black & white.

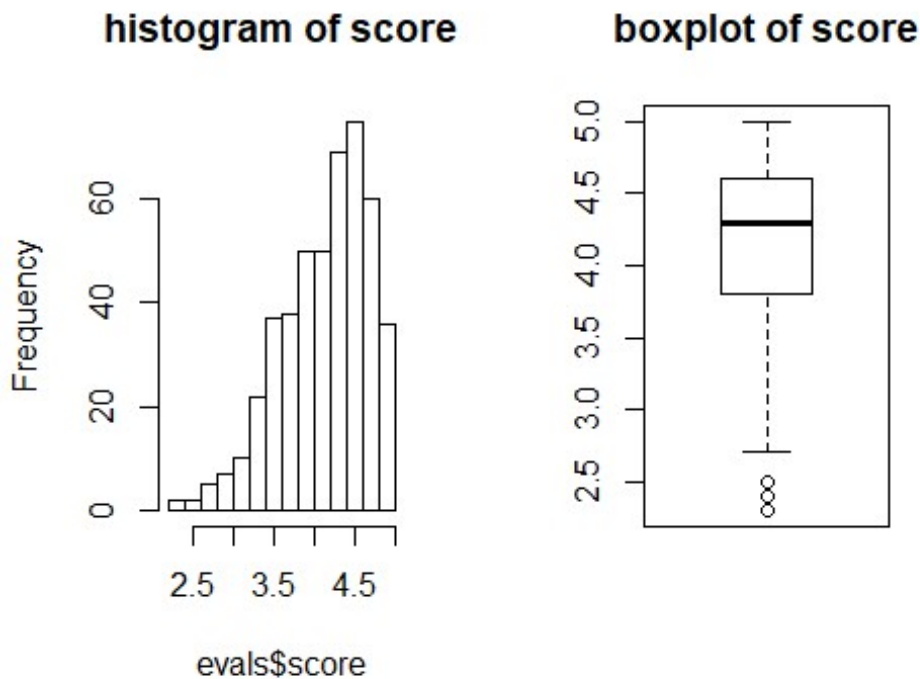
Exploring the data

1. Is this an observational study or an experiment? The original research question posed in the paper is whether beauty leads directly to the differences in course evaluations. Given the study design, is it possible to answer this question as it is phrased? If not, rephrase the question.

JN: strictly speaking, this is an observational study because we do not control about assigning students to different professors/classes. What if all English professors are considered to be significantly more attractive than professors teaching mathematics? Neither do we control the “attractiveness” of professors in this student. We cannot do randomized control trial. However, we can still answer the question by carrying out a correlation or quasi-experimental study.

2. Describe the distribution of score. Is the distribution skewed? What does that tell you about how students rate courses? Is this what you expected to see? Why, or why not?

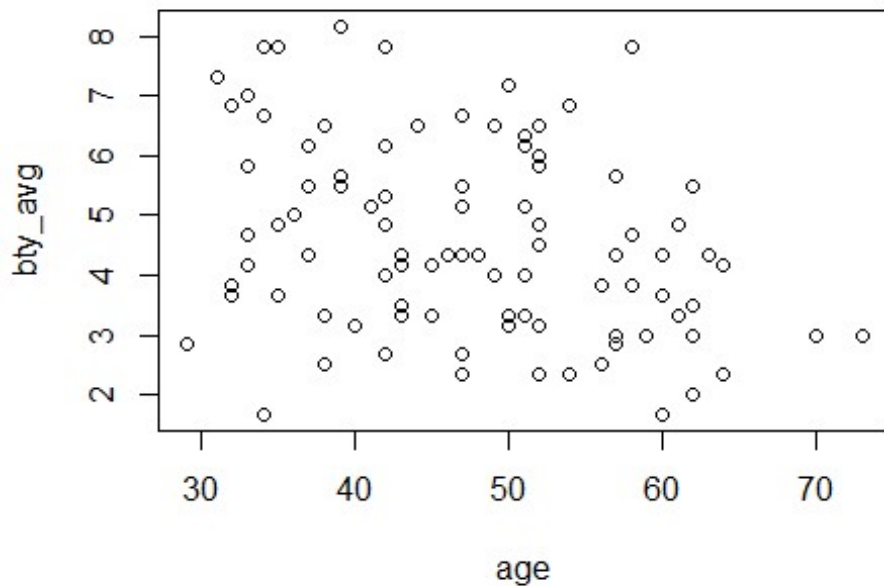
```
par(mfrow = c(1, 2))  
hist(evals$score, main = "histogram of score")  
boxplot(evals$score, main = "boxplot of score")
```



JN: Yes that’s heavily left skewed. That means majority of students have positive rating of their professors, with only few exceptions that rate their professors negatively (or below average). That’s what I expected - majority of students have positive learning experience.

3. Excluding score, select two other variables and describe their relationship using an appropriate visualization (scatterplot, side-by-side boxplots, or mosaic plot).

```
with(evals,  
  plot(bty_avg ~ age))
```



```
with(evals,
      cor.test(bty_avg, age))
##
## Pearson's product-moment correlation
##
## data:  bty_avg and age
## t = -6.8664, df = 461, p-value = 2.137e-11
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.3850454 -0.2195681
## sample estimates:
##          cor
## -0.3046034
```

JN: there's a very weak negative correlation between a professor's age and his/her average beauty rating.

Simple linear regression

The fundamental phenomenon suggested by the study is that better looking teachers are evaluated more favorably. Let's create a scatterplot to see if this appears to be the case:

```
plot(evals$score ~ evals$bty_avg)
```

Before we draw conclusions about the trend, compare the number of observations in the data frame with the approximate number of points on the scatterplot. Is anything awry?

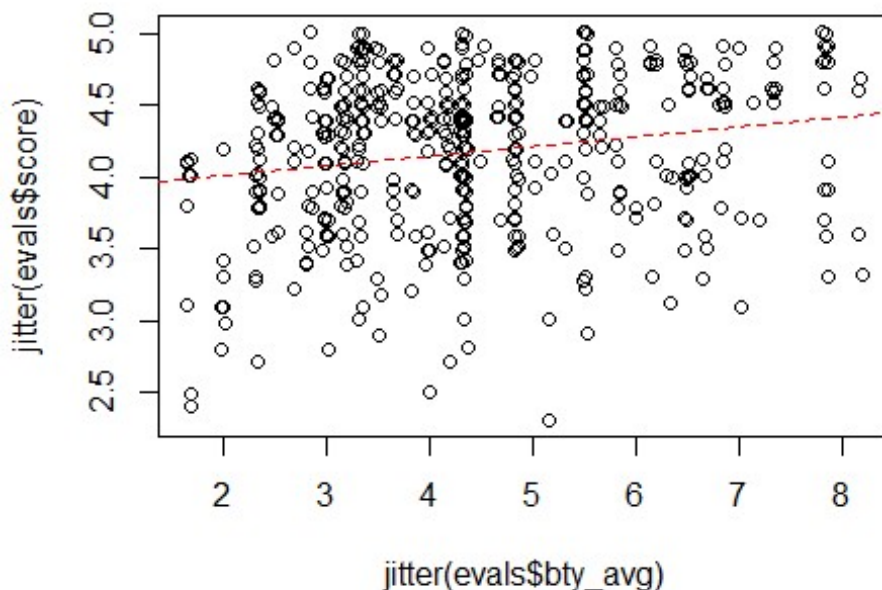
4. Replot the scatterplot, but this time use the function `jitter()` on the y - or the x -coordinate. (Use `?jitter` to learn more.) What was misleading about the initial scatterplot?

```
plot(evals$score ~ jitter(evals$bty_avg))
```

JN: what's misleading about the scatterplot is that, although the two variables look like “continuous”, they are actually not. If we run `library(tidyverse)`, and then “`unique(evals$score) %>% length`”, we will see that we only have 27 uniques for score and 35 for `btj_avg`. In other words, there are “spaces” or “gaps” between dots and as a result, the scatterplot depicts the relationship between the two variables visually less convincing.

5. Let's see if the apparent trend in the plot is something more than natural variation. Fit a linear model called `m_bty` to predict average professor score by average beauty rating and add the line to your plot using `abline(m_bty)`. Write out the equation for the linear model and interpret the slope. Is average beauty score a statistically significant predictor? Does it appear to be a practically significant predictor?

```
m_bty <- lm(score ~ bty_avg, data = evals)
plot(jitter(evals$score) ~ jitter(evals$bty_avg))
abline(m_bty, lty = 2, col = "red")
```



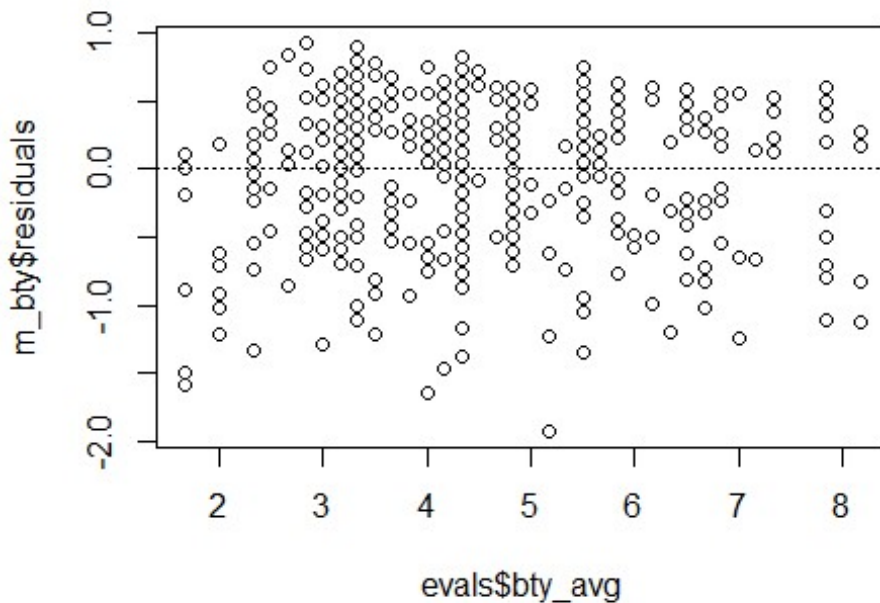
```
summary(m_bty)
##
## Call:
## lm(formula = score ~ bty_avg, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9246 -0.3690  0.1420  0.3977  0.9309
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.88034    0.07614   50.96 < 2e-16 ***
## bty_avg        0.06664    0.01629    4.09 5.08e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5348 on 461 degrees of freedom
```

```
## Multiple R-squared:  0.03502,    Adjusted R-squared:  0.03293
## F-statistic: 16.73 on 1 and 461 DF,  p-value: 5.083e-05
```

JN: the equation is $y = 0.06664 * x + 3.88034$ // the slope (0.06664) is positive, meaning every single increase in `bty_avg` would bring an increase of score by 0.06664. // Yes, the `bty_avg` is a significant predictor with p-value far below 0.05, i.e. 5.08e-05. However, it may not be practically significant or useful because the estimate is so small (0.06664).

6. Use residual plots to evaluate whether the conditions of least squares regression are reasonable. Provide plots and comments for each one (see the Simple Regression Lab for a reminder of how to make these).

```
plot(m_bty$residuals ~ evals$bty_avg)
abline(h = 0, lty = 3) # adds a horizontal dashed line at y = 0
```



JN: yes the condition is met. The above residual plot shows a random pattern - the randomness indicating the satisfaction of homoscedasticity.

Multiple linear regression

The data set contains several variables on the beauty score of the professor: individual ratings from each of the six students who were asked to score the physical appearance of the professors and the average of these six scores. Let's take a look at the relationship between one of these scores and the average beauty score.

```
plot(evals$bty_avg ~ evals$bty_f1lower)
cor(evals$bty_avg, evals$bty_f1lower)
```

As expected the relationship is quite strong - after all, the average score is calculated using the individual scores. We can actually take a look at the relationships between all beauty variables (columns 13 through 19) using the following command:

```
plot(evals[,13:19])
```

These variables are collinear (correlated), and adding more than one of these variables to the model would not add much value to the model. In this application and with these highly-correlated predictors, it is reasonable to use the average beauty score as the single representative of these variables.

In order to see if beauty is still a significant predictor of professor score after we've accounted for the gender of the professor, we can add the gender term into the model.

```
m_bty_gen <- lm(score ~ bty_avg + gender, data = evals)
summary(m_bty_gen)

##
## Call:
## lm(formula = score ~ bty_avg + gender, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8305 -0.3625  0.1055  0.4213  0.9314
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.74734    0.08466  44.266  < 2e-16 ***
## bty_avg        0.07416    0.01625   4.563 6.48e-06 ***
## gendermale     0.17239    0.05022   3.433 0.000652 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5287 on 460 degrees of freedom
## Multiple R-squared:  0.05912,    Adjusted R-squared:  0.05503
## F-statistic: 14.45 on 2 and 460 DF,  p-value: 8.177e-07
```

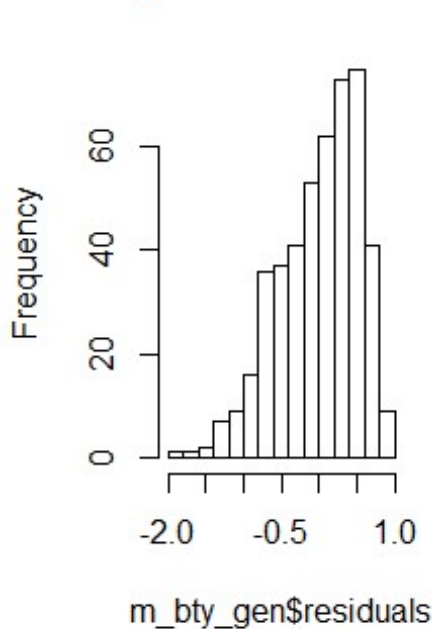
7. P-values and parameter estimates should only be trusted if the conditions for the regression are reasonable. Verify that the conditions for this model are reasonable using diagnostic plots.

```
library(car)

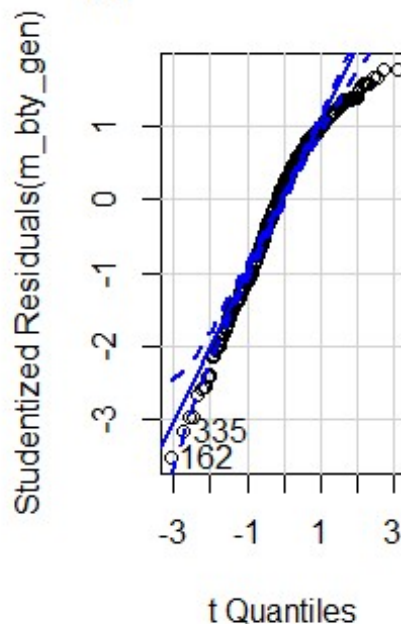
## Loading required package: carData

par(mfrow = c(1, 2))
hist(m_bty_gen$residuals, main = "histogram of the residuals")
car::qqPlot(m_bty_gen, main = "qqplot of the residuals")
```

histogram of the residuals



qqplot of the residuals



```
## [1] 162 335
```

JN: the residuals are nearly normally distributed (with left skewed).

8. Is `bty_avg` still a significant predictor of score? Has the addition of gender to the model changed the parameter estimate for `bty_avg`?

JN: Yes, it still is a significant predictor with p-value 6.48e-06. Adding “gender” to the model also changed the parameter estimate of “bty_avg” from 0.06664 to 0.07416.

Note that the estimate for gender is now called `gendermale`. You’ll see this name change whenever you introduce a categorical variable. The reason is that R recodes gender from having the values of `female` and `male` to being an indicator variable called `gendermale` that takes a value of 0 for females and a value of 1 for males. (Such variables are often referred to as “dummy” variables.)

As a result, for females, the parameter estimate is multiplied by zero, leaving the intercept and slope form familiar from simple regression.

$$\begin{aligned} \text{score} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{bty_avg} + \hat{\beta}_2 \times (0) \\ &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{bty_avg} \end{aligned}$$

We can plot this line and the line corresponding to males with the following custom function.

```
multiLines(m_bty_gen)
```

9. What is the equation of the line corresponding to males? (*Hint:* For males, the parameter estimate is multiplied by 1.) For two professors who received the same beauty rating, which gender tends to have the higher course evaluation score?

JN: $y = 3.74734 + x_1 * (0.07416) + 1 * (0.17239)$. “Male” would tend to have higher score. If that’s “female”, the estimate (0.17239) would be multiplied by 0, and that means it will not be added or contributed to the increase in score.

The decision to call the indicator variable `gendermale` instead of `genderfemale` has no deeper meaning. R simply codes the category that comes first alphabetically as a 0. (You can change the reference level of a categorical variable, which is the level that is coded as a 0, using the `relevel` function. Use `?relevel` to learn more.)

10. Create a new model called `m_bty_rank` with `gender` removed and `rank` added in. How does R appear to handle categorical variables that have more than two levels? Note that the `rank` variable has three levels: `teaching`, `tenure track`, `tenured`.

```
m_bty_rank <- lm(score ~ bty_avg + rank, data = evals)
summary(m_bty_rank)
##
## Call:
## lm(formula = score ~ bty_avg + rank, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8713 -0.3642  0.1489  0.4103  0.9525
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.98155    0.09078   43.860 < 2e-16 ***
## bty_avg         0.06783    0.01655    4.098 4.92e-05 ***
## ranktenure track -0.16070    0.07395   -2.173  0.0303 *
## ranktenured     -0.12623    0.06266   -2.014  0.0445 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5328 on 459 degrees of freedom
## Multiple R-squared:  0.04652,    Adjusted R-squared:  0.04029
## F-statistic: 7.465 on 3 and 459 DF,  p-value: 6.88e-05
```

JN: the categorical variable would be “dummified”. If it has 3 levels, then you will see it appear twice (k - 1) as shown above.

The interpretation of the coefficients in multiple regression is slightly different from that of simple regression. The estimate for `bty_avg` reflects how much higher a group of professors is expected to score if they have a beauty rating that is one point higher *while holding all other variables constant*. In this case, that translates into considering only professors of the same rank with `bty_avg` scores that are one point apart.

The search for the best model

We will start with a full model that predicts professor score based on rank, ethnicity, gender, language of the university where they got their degree, age, proportion of students that filled out evaluations, class size, course level, number of professors, number of credits, average beauty rating, outfit, and picture color.

11. Which variable would you expect to have the highest p-value in this model? Why? *Hint:* Think about which variable would you expect to not have any association with the professor score.

JN: I expect that `cls_profs` would have the highest p-value, simply because the number of professors teaching the course should have nothing to do with the score.

Let's run the model...

```
m_full <- lm(score ~ rank + ethnicity + gender + language + age + cls_perc_eval
             + cls_students + cls_level + cls_profs + cls_credits + bty_avg
             + pic_outfit + pic_color, data = evals)
summary(m_full)
```

12. Check your suspicions from the previous exercise. Include the model output in your response.

JN: Yes, my suspicion is confirmed - `cls_profssingle` has the highest p-value of "0.77806". That means, the variable is highly irrelevant to the regression model.

13. Interpret the coefficient associated with the ethnicity variable.

JN: once again, the p-value is not significant, i.e. 0.11698, meaning ethnicity of professor does not significantly impact the score. To interpret the coefficient, we can state that as when a professor is "not minority", we would add 0.1234929 to the score, holding everything else constant.

14. Drop the variable with the highest p-value and re-fit the model. Did the coefficients and significance of the other explanatory variables change? (One of the things that makes multiple regression interesting is that coefficient estimates depend on the other variables that are included in the model.) If not, what does this say about whether or not the dropped variable was collinear with the other explanatory variables?

```
m_parital <- lm(score ~ rank + ethnicity + gender + language + age + cls_perc_eval
                + cls_students + cls_level + cls_credits + bty_avg
                + pic_outfit + pic_color, data = evals)
summary(m_parital)
```

JN: Yes, the coefficients and significance of the other explanatory variables change after dropping "`cls_profs`".

15. Using backward-selection and p-value as the selection criterion, determine the best model. You do not need to show all steps in your answer, just the output for the final model. Also, write out the linear model for predicting score based on the final model you settle on.

```
m_all <- lm(score ~., data = evals)

library(MASS)
MASS::stepAIC(m_all, direction = "backward")
## Start:  AIC=-631.44
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##       cls_did_eval + cls_students + cls_level + cls_profs + cls_credits +
##       bty_f1lower + bty_f1upper + bty_f2upper + bty_m1lower + bty_m1upper +
##       bty_m2upper + bty_avg + pic_outfit + pic_color
##
##              Df Sum of Sq  RSS    AIC
## - cls_profs    1    0.0076 107.66 -633.41
## - cls_level    1    0.0756 107.73 -633.12
## - cls_students  1    0.1325 107.78 -632.87
## - bty_m2upper   1    0.1646 107.81 -632.73
## - bty_m1lower   1    0.1649 107.82 -632.73
```

```

## - bty_m1upper      1      0.1653 107.82 -632.73
## - bty_avg          1      0.1655 107.82 -632.73
## - bty_f1lower      1      0.1657 107.82 -632.73
## - bty_f2upper      1      0.1662 107.82 -632.73
## - bty_f1upper      1      0.1669 107.82 -632.72
## - cls_did_eval     1      0.1974 107.85 -632.59
## - pic_outfit       1      0.3270 107.98 -632.04
## <none>              107.65 -631.44
## - rank             2      1.0319 108.68 -631.02
## - cls_perc_eval    1      0.5699 108.22 -631.00
## - ethnicity        1      0.9156 108.57 -629.52
## - language         1      1.2320 108.88 -628.17
## - age              1      1.5956 109.25 -626.63
## - pic_color        1      2.2190 109.87 -623.99
## - cls_credits      1      4.4803 112.13 -614.56
## - gender           1      5.6448 113.30 -609.78
##
## Step: AIC=-633.41
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##         cls_did_eval + cls_students + cls_level + cls_credits + bty_f1lower +
##         bty_f1upper + bty_f2upper + bty_m1lower + bty_m1upper + bty_m2upper +
##         bty_avg + pic_outfit + pic_color
##
##           Df Sum of Sq    RSS    AIC
## - cls_level      1      0.0751 107.73 -635.09
## - cls_students   1      0.1299 107.79 -634.85
## - bty_m2upper     1      0.1578 107.82 -634.73
## - bty_m1lower     1      0.1581 107.82 -634.73
## - bty_m1upper     1      0.1585 107.82 -634.73
## - bty_avg         1      0.1587 107.82 -634.73
## - bty_f1lower     1      0.1590 107.82 -634.72
## - bty_f2upper     1      0.1594 107.82 -634.72
## - bty_f1upper     1      0.1601 107.82 -634.72
## - cls_did_eval    1      0.1929 107.85 -634.58
## - pic_outfit      1      0.3673 108.03 -633.83
## <none>              107.66 -633.41
## - rank           2      1.0311 108.69 -632.99
## - cls_perc_eval   1      0.5958 108.25 -632.85
## - ethnicity        1      0.9207 108.58 -631.47
## - language         1      1.2392 108.90 -630.11
## - age              1      1.6029 109.26 -628.57
## - pic_color        1      2.2156 109.87 -625.98
## - cls_credits      1      4.5061 112.16 -616.42
## - gender           1      5.6685 113.33 -611.65
##
## Step: AIC=-635.09
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##         cls_did_eval + cls_students + cls_credits + bty_f1lower +
##         bty_f1upper + bty_f2upper + bty_m1lower + bty_m1upper + bty_m2upper +
##         bty_avg + pic_outfit + pic_color
##
##           Df Sum of Sq    RSS    AIC
## - bty_m2upper     1      0.1353 107.87 -636.50
## - bty_m1lower     1      0.1356 107.87 -636.50
## - bty_m1upper     1      0.1359 107.87 -636.50
## - bty_avg         1      0.1361 107.87 -636.50
## - bty_f1lower     1      0.1364 107.87 -636.50
## - bty_f2upper     1      0.1368 107.87 -636.50

```

```

## - bty_f1upper      1      0.1374 107.87 -636.49
## - cls_students     1      0.1910 107.92 -636.27
## - cls_did_eval     1      0.2465 107.98 -636.03
## - pic_outfit       1      0.4078 108.14 -635.34
## <none>              107.73 -635.09
## - rank             2      1.0007 108.73 -634.80
## - cls_perc_eval    1      0.5664 108.30 -634.66
## - ethnicity        1      0.9869 108.72 -632.86
## - language         1      1.1731 108.91 -632.07
## - age              1      1.5633 109.30 -630.41
## - pic_color        1      2.1435 109.88 -627.96
## - cls_credits      1      4.4849 112.22 -618.20
## - gender           1      5.6057 113.34 -613.60
##
## Step: AIC=-636.5
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##         cls_did_eval + cls_students + cls_credits + bty_f1lower +
##         bty_f1upper + bty_f2upper + bty_m1lower + bty_m1upper + bty_avg +
##         pic_outfit + pic_color
##
##           Df Sum of Sq    RSS    AIC
## - bty_m1lower    1      0.0319 107.90 -638.37
## - bty_m1upper    1      0.1637 108.03 -637.80
## - cls_students   1      0.2207 108.09 -637.56
## - cls_did_eval   1      0.2710 108.14 -637.34
## - pic_outfit     1      0.4084 108.28 -636.75
## <none>           107.87 -636.50
## - bty_f1lower    1      0.4873 108.36 -636.42
## - cls_perc_eval  1      0.5417 108.41 -636.18
## - bty_avg        1      0.6009 108.47 -635.93
## - rank           2      1.1260 109.00 -635.70
## - ethnicity       1      0.8591 108.73 -634.83
## - language       1      1.1373 109.01 -633.65
## - bty_f2upper    1      1.1690 109.04 -633.51
## - age            1      1.5120 109.38 -632.06
## - bty_f1upper    1      1.8540 109.72 -630.61
## - pic_color      1      2.3311 110.20 -628.61
## - cls_credits    1      4.3939 112.26 -620.02
## - gender         1      5.8105 113.68 -614.21
##
## Step: AIC=-638.37
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##         cls_did_eval + cls_students + cls_credits + bty_f1lower +
##         bty_f1upper + bty_f2upper + bty_m1upper + bty_avg + pic_outfit +
##         pic_color
##
##           Df Sum of Sq    RSS    AIC
## - bty_m1upper    1      0.1347 108.03 -639.79
## - cls_students   1      0.2258 108.13 -639.40
## - cls_did_eval   1      0.2842 108.19 -639.15
## - pic_outfit     1      0.4271 108.33 -638.54
## <none>           107.90 -638.37
## - bty_f1lower    1      0.5241 108.42 -638.12
## - cls_perc_eval  1      0.5243 108.42 -638.12
## - rank           2      1.0984 109.00 -637.68
## - ethnicity       1      0.9341 108.83 -636.38
## - language       1      1.1373 109.04 -635.51
## - bty_avg        1      1.1563 109.06 -635.43

```

```

## - bty_f2upper      1      1.4875 109.39 -634.03
## - age              1      1.6158 109.52 -633.49
## - bty_f1upper      1      2.3451 110.25 -630.41
## - pic_color        1      2.4870 110.39 -629.82
## - cls_credits      1      4.3675 112.27 -622.00
## - gender           1      5.8001 113.70 -616.13
##
## Step:  AIC=-639.79
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##       cls_did_eval + cls_students + cls_credits + bty_f1lower +
##       bty_f1upper + bty_f2upper + bty_avg + pic_outfit + pic_color
##
##              Df Sum of Sq    RSS    AIC
## - cls_students  1      0.2111 108.25 -640.89
## - cls_did_eval  1      0.2718 108.31 -640.63
## - pic_outfit    1      0.4059 108.44 -640.05
## - bty_f1lower   1      0.4326 108.47 -639.94
## <none>                                108.03 -639.79
## - cls_perc_eval 1      0.6118 108.65 -639.18
## - rank           2      1.1107 109.15 -639.05
## - ethnicity      1      0.8677 108.90 -638.09
## - language       1      1.0897 109.12 -637.14
## - bty_avg        1      1.2696 109.31 -636.38
## - bty_f2upper    1      1.3535 109.39 -636.03
## - age            1      1.8207 109.86 -634.05
## - bty_f1upper    1      2.2334 110.27 -632.32
## - pic_color      1      2.5154 110.55 -631.13
## - cls_credits    1      4.8529 112.89 -621.45
## - gender         1      5.7018 113.74 -617.98
##
## Step:  AIC=-640.89
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##       cls_did_eval + cls_credits + bty_f1lower + bty_f1upper +
##       bty_f2upper + bty_avg + pic_outfit + pic_color
##
##              Df Sum of Sq    RSS    AIC
## - cls_did_eval  1      0.1415 108.39 -642.28
## - pic_outfit    1      0.3573 108.60 -641.36
## - bty_f1lower   1      0.4667 108.71 -640.89
## <none>                                108.25 -640.89
## - rank           2      1.1821 109.43 -639.86
## - ethnicity      1      0.9919 109.24 -638.66
## - language       1      0.9963 109.24 -638.64
## - bty_avg        1      1.3404 109.59 -637.19
## - bty_f2upper    1      1.5770 109.82 -636.19
## - age            1      1.8395 110.09 -635.08
## - bty_f1upper    1      2.1827 110.43 -633.64
## - pic_color      1      2.3708 110.62 -632.85
## - cls_perc_eval  1      2.4497 110.70 -632.52
## - cls_credits    1      4.8839 113.13 -622.45
## - gender         1      5.5358 113.78 -619.79
##
## Step:  AIC=-642.28
## score ~ rank + ethnicity + gender + language + age + cls_perc_eval +
##       cls_credits + bty_f1lower + bty_f1upper + bty_f2upper + bty_avg +
##       pic_outfit + pic_color
##
##              Df Sum of Sq    RSS    AIC

```

```
## <none> 108.39 -642.28
## - pic_outfit 1 0.5065 108.89 -642.12
## - bty_f1lower 1 0.5212 108.91 -642.06
## - rank 2 1.2018 109.59 -641.18
## - ethnicity 1 1.0217 109.41 -639.94
## - language 1 1.0886 109.48 -639.65
## - bty_avg 1 1.4304 109.82 -638.21
## - bty_f2upper 1 1.7306 110.12 -636.95
## - age 1 1.9987 110.39 -635.82
## - pic_color 1 2.2613 110.65 -634.72
## - bty_f1upper 1 2.2891 110.68 -634.60
## - cls_perc_eval 1 2.3203 110.71 -634.47
## - cls_credits 1 4.9069 113.30 -623.78
## - gender 1 5.7484 114.14 -620.35
##
## Call:
## lm(formula = score ~ rank + ethnicity + gender + language + age +
##     cls_perc_eval + cls_credits + bty_f1lower + bty_f1upper +
##     bty_f2upper + bty_avg + pic_outfit + pic_color, data = evals)
##
## Coefficients:
## (Intercept) ranktenure track ranktenured
## 4.08717 -0.18412 -0.07980
## ethnicitynot minority gendermale languagenon-english
## 0.15616 0.25493 -0.22848
## age cls_perc_eval cls_creditsone credit
## -0.00913 0.00447 0.50273
## bty_f1lower bty_f1upper bty_f2upper
## 0.03884 0.07990 0.05849
## bty_avg pic_outfitnot formal pic_colorcolor
## -0.14036 -0.10030 -0.20547
```

JN: as shown above, we can use the *AIC* (Akaike’s Information Criteria) to determine the best model. Using direction = “backward” and starting with *m_all* (by including every explanatory variable), we found the best model (or stopped removing variable) with the smallest AIC returned.

And the model is described below,

```
lm(formula = score ~ rank + ethnicity + gender + language + age + cls_perc_eval + cls_credits + bty_f1lower +
bty_f1upper + bty_f2upper + bty_avg + pic_outfit + pic_color, data = evals)
```

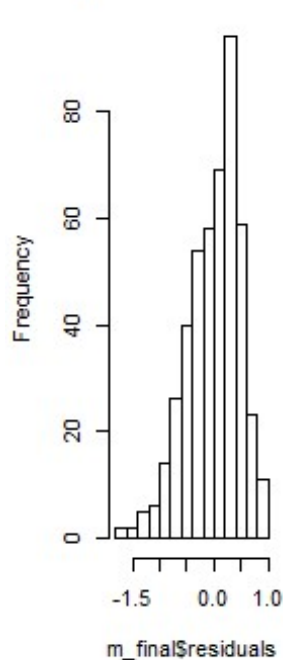
16. Verify that the conditions for this model are reasonable using diagnostic plots.

```
m_final <- lm(formula = score ~ rank + ethnicity + gender + language + age +
  cls_perc_eval + cls_credits + bty_f1lower + bty_f1upper +
  bty_f2upper + bty_avg + pic_outfit + pic_color, data = evals)
summary(m_final)
##
## Call:
## lm(formula = score ~ rank + ethnicity + gender + language + age +
##     cls_perc_eval + cls_credits + bty_f1lower + bty_f1upper +
##     bty_f2upper + bty_avg + pic_outfit + pic_color, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7317 -0.3070  0.1020  0.3535  0.9283
```

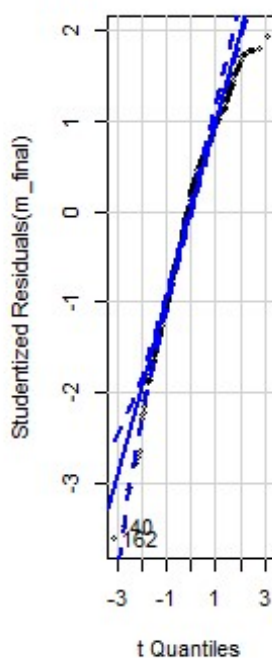
```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.087168   0.284915  14.345 < 2e-16 ***
## ranktenure track -0.184119   0.082619  -2.229  0.02634 *
## ranktenured    -0.079796   0.065691  -1.215  0.22511
## ethnicitynot minority 0.156156   0.075989   2.055  0.04046 *
## gendermale      0.254929   0.052300   4.874 1.52e-06 ***
## languagenon-english -0.228479   0.107710  -2.121  0.03445 *
## age            -0.009130   0.003177  -2.874  0.00424 **
## cls_perc_eval    0.004470   0.001444   3.097  0.00208 **
## cls_creditsone credit 0.502732   0.111631   4.504 8.54e-06 ***
## bty_f1lower      0.038838   0.026460   1.468  0.14287
## bty_f1upper      0.079899   0.025975   3.076  0.00223 **
## bty_f2upper      0.058486   0.021868   2.674  0.00776 **
## bty_avg         -0.140359   0.057725  -2.432  0.01543 *
## pic_outfitnot formal -0.100300   0.069317  -1.447  0.14860
## pic_colorcolor   -0.205474   0.067209  -3.057  0.00237 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4919 on 448 degrees of freedom
## Multiple R-squared:  0.2068, Adjusted R-squared:  0.1821
## F-statistic: 8.345 on 14 and 448 DF,  p-value: 4.79e-16
```

```
par(mfrow = c(1, 3))
hist(m_final$residuals, main = "histogram of the residuals")
car::qqPlot(m_final, main = "qqplot of the residuals")
## [1] 40 162
plot(m_final$residuals, main = "residuals' plot")
```

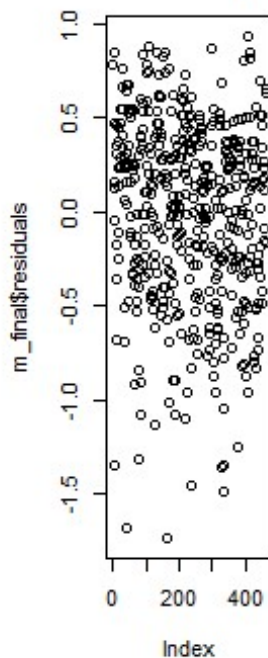
histogram of the residuals



qqplot of the residuals



residuals' plot



JN: Yes, every condition is checked out!!!

17. The original paper describes how these data were gathered by taking a sample of professors from the University of Texas at Austin and including all courses that they have taught. Considering that each row represents a course, could this new information have an impact on any of the conditions of linear regression?

JN: it won't, consider the sample (in this case, professors) were randomly selected. There's no reason to believe that the courses that were taught by these random samples contain bias or misrepresent the general population.

18. Based on your final model, describe the characteristics of a professor and course at University of Texas at Austin that would be associated with a high evaluation score.

JN: Most likely, a professor is male, teaching, not minority, received education from school that was taught in English, relatively young and less likely to be rated high in `bty_avg`.

19. Would you be comfortable generalizing your conclusions to apply to professors generally (at any university)? Why or why not?

JN: I would feel uncomfortable because the adjusted R-squared is only 0.1821. That's too weak a relationship between all independent variables and the dependent variable. There is still over 80% of the variance not explained by the model. We need to collect different data or redesign our study by coming up with better hypothesis for testing.

This is a product of OpenIntro that is released under a [Creative Commons Attribution-ShareAlike 3.0 Unported](#). This lab was written by Mine Çetinkaya-Rundel and Andrew Bray.