Lab 3: Multiple linear regression

**Due date:** Thursday, Feb 27, 2025 submitted as Word document to Canvas ***Lab03***  link

This lab counts 9 % toward your total grade.

**Objectives:** In this lab, you will practice your skills in

1. Explore multiple regression
2. Multicollinearity
3. Model selection

**Format of answer:** Submit your answers as a **Word document** with graphs and verbal descriptions, properly labeled in the task sequence, with answers in red text and only relevant content included

Data Introduction - evals

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.

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

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 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.

Task 1: Exploring the data (4 pts)

1. 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 quick look at the relationship between one of these scores (bty\_f1lower) and the average beauty score(bty\_avg) using **plot()**.Please explain the relationship between these two variables. (0.5 pts)

library(carData);library(car);library(regclass);library(openintro);library(MASS)

evals = read.csv('evals.csv')

plot(evals$bty\_avg~evals$bty\_f1lower)

A graph with circles and numbers

AI-generated content may be incorrect.

The scatterplot shows relationship between beauty rating of professor from the lower level female( (1) lowest - (10) highest) ” bty\_f1lower” and “bty\_avg”, the average beauty rating of professor.There is a positive corelation between both variables. The discreteness in the dataset is observed as beauty ratings are given on a **1 to 10 scale**. Additionally, it means that similar ratings were given to the professor by multiple students. There are diffferent values of bty\_avg against bty\_f1lower.

1. Using **car::vif()** function to check the multicollinearity for independent variables (bty\_f1lower,bty\_f1upper,bty\_f2upper, bty\_m1lower, bty\_m1upper, bty\_m2upper, bty\_avg) and dependent variable(score). Please illustrate the multicollinearity relationship among them and justify the reason that use the average beauty score (bty\_avg) as a single representative of these variables. (1 pts)

lm01 = lm(score ~bty\_f1lower+bty\_f1upper+bty\_f2upper+ bty\_m1lower+ bty\_m1upper+ bty\_m2upper+bty\_avg, data = evals)

vif(lm01)

Variance inflation factor (VIF) function is used to check the colinearity . VIF value tells how much variance is increased because of multicollinearity. Multicollinearity occurs between independent variables, when variables are highly correlated in a regression model. Following is the VIF values

A blue rectangle with black text

AI-generated content may be incorrect.  
  
All values are above 10. When the VIF value is above 10, it is a strong signal that there is severe multicollinearity. This means that the independent variables are very closely related, which can confuse the results of the analysis. In this example, all variables are different characteristics of beauty perception. In order to remove confusion, we can either combine variables to a single variable. But as already average beauty score is available in the data, which is dependent on all variables we must use bty\_avg as a single variable.

1. The model you finalized in Task1.a is a bivariate model. 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 (gender) into the model. (0.5 pts)

m\_bty\_gen <- lm(score ~ bty\_avg + gender, data = evals)

summary(m\_bty\_gen)

A screenshot of a computer code

AI-generated content may be incorrect.

1. p-values and parameter estimates should only be trusted if the conditions for the regression are reasonable. Verify that the conditions for Task1.c model are reasonable using diagnostic plots (plot the histogram of residual to assess the normality of the residual; using residual plot (residual vs fitted value) to assess the variance in residual. (1 pts)

hist(m\_bty\_gen$residuals, main = 'Histogram of the residual')

plot(x = m\_bty\_gen$residuals, y = m\_bty\_gen$fitted.values,

xlab = 'residuals', ylab = 'fitted value')

To address that the p-values and parameter estimates are trusted if the conditions for the regression are reasonable, we must check the condition of the linear regression. First, the normality of the residuals is checked by histogram. Histogram reveals that residuals are slightly left skewed showing some outliers but giving a shape of a bell curve. Second, homoscedasticity is observed asresiduals have the constant variable.

A graph of a person with a number of bars

AI-generated content may be incorrect.

A graph of a number of dots

AI-generated content may be incorrect.

1. 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.). 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.

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

1. What is the equation corresponding to males? For two professors who received the same beauty rating, which gender tends to have the higher course evaluation score? Please explain the . (0.5 pts)

β₀ :Theintercept is 3.747

β₁: For each additional unit increase in beauty score, there is an increament in course evaluation score by 0.074, regardless of gender.

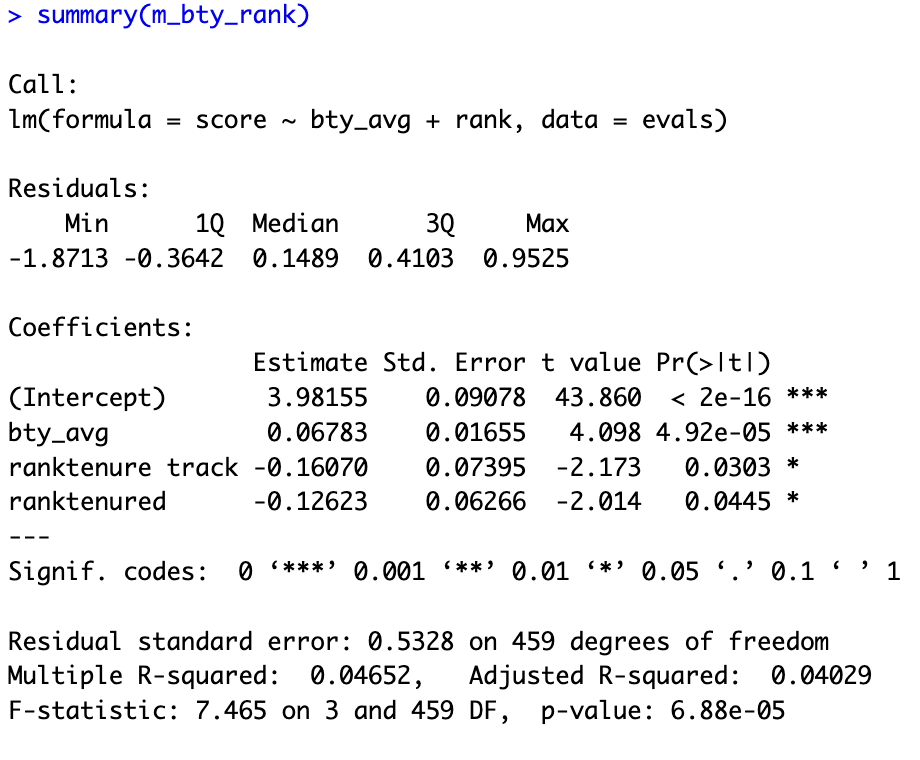
β₂: This is gendermale coefficient which reveals that the male professors receive an average 0.172 higher evaluation score than female professors, controlling for beauty.

Male professors receive an average 0.172 higher evaluation score than female professors.

1. 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. (0.5 pts)

m\_bty\_rank = lm(score ~ bty\_avg+rank, data=evals)

summary(m\_bty\_rank)



R automatically converts categorical variables into dummy (indicator) variables.The first category ‘Teaching’ is alphabetically chosen as the reference category, and the remaining levels get separate coefficients. The choosen reference category will be difference from the ranktenure track and ranktenured.

Task 2: The search for the best model (5 pts)

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.

*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)*

1. 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. (0.5 pts)

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)

The following variables related to students might not be statistically significant as score is evaluated based on professors

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.

1. Check your suspicions based on the summary from the full model. Include the model output in your response. (1 pts)

summary(m\_full)

From summary below we can say that number of professors teaching sections in course in sample- single “cls\_profssingle” is the highest p-value in this model. Since value of variable cls\_profssingle 0.77806 is greater than 0.05, we fail to reject the null hypothesis. Thus, cls\_profssingle variable is not statistically significant.

A screenshot of a computer

AI-generated content may be incorrect.

1. Interpret the coefficient associated with the ethnicity variable. (1 pts)

The coefficient associated with ethnicity is ethnicitynot minority=0.1234929

It means, non-minority professors receive, on average, 0.123 higher course evaluation scores as compared to minority professors.

1. Using different model selection methods (best subset, backward stepwise, forward stepwise), determine the best model based on AIC as the selection criterion (1 pts)

**Best Subset Model**

library(leaps)  
# Best subset selection

best\_subset <- regsubsets(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, nbest = 1, nvmax = 13)

# Summary of results

summary\_best <- summary(best\_subset)

# Print model details

print(summary\_best)

# Find the best model based on Cp (Mallows' Cp)

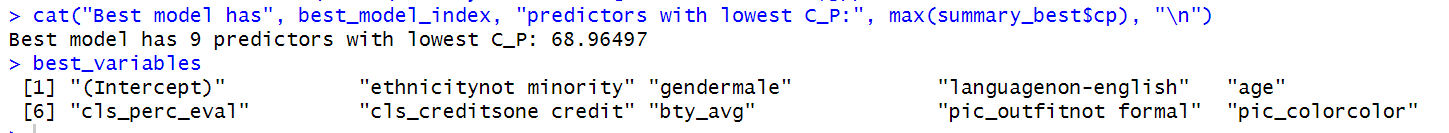
best\_model\_index <- which.min(summary\_best$cp)

# Extract variable names for the best model (excluding intercept)

best\_variables <- names(which(summary\_best$which[best\_model\_index,]))

print(best\_variables)

cat("Best model has", best\_model\_index, "predictors with lowest C\_P:", max(summary\_best$cp), "\n")



**Backward Model:**

full\_model = 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)

backward\_model <- step(full\_model,direction = "backward")

summary\_backward <- summary(backward\_model)

A screenshot of a computer

AI-generated content may be incorrect.

**Forward model:**

forward\_model = regsubsets(score ~ rank + ethnicity + gender + language + age + cls\_perc\_eval+ cls\_students + cls\_level + cls\_profs + cls\_credits + bty\_avg

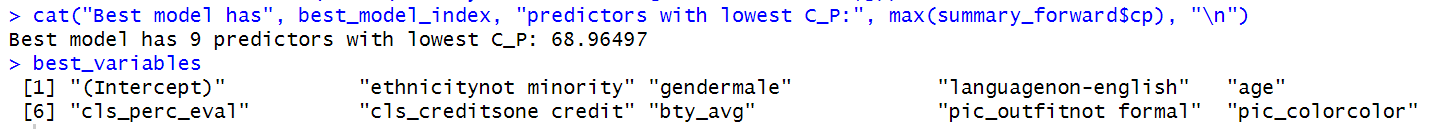
+ pic\_outfit + pic\_color, data = evals, nvmax = 13, method = 'forward')

summary\_forward <- summary(forward\_model)

best\_model\_index <- which.min(summary\_forward$cp)

best\_variables <- names(which(summary\_forward$which[best\_model\_index,]))

cat("Best model has", best\_model\_index, "predictors with lowest C\_P:", max(summary\_forward$cp), "\n")



**Results**

Based on three model selection model method, all of them detect the same variables. So the final model is :

final\_model = lm(score ~ ethnicity + gender + language + age + cls\_perc\_eval+ cls\_credits + bty\_avg+ pic\_outfit + pic\_color, data = evals)

summary(final\_model)

1. 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. (0.5 pts)

summary(backward\_model)

A professor who is male, not a minority, a native English speaker, younger, has a high beauty score, and teaches a one-credit course with a high percentage of student participation in evaluations is more likely to receive a high evaluation score.

A screenshot of a computer program

AI-generated content may be incorrect.

1. Would you be comfortable generalizing your conclusions to apply to professors generally (at any university)? Why or why not? (1 p

No, I will not be comfortable for generalizing conclusions to apply to professors generally to any other university because of following reasons

1. Data sample can vary
2. Survey responses might vary based on cultural differences of the university and structure of the individual professor’s course teaching and grading.
3. University policies might be different