Assignment 03

Introduction to Multiple Regression

Research. Teaching. Service. The trifecta upon which that almost every university instructor is evaluated, and, ultimately compensated. There has been substantial research to indicate that higher quality teaching is associated with increased faculty salararies. One way which academic administrators judge teaching quality is through teachers' course evaluations. While we know evaluation scores are not perfectly measures of teaching quality, nonetheless, they do play a role in the tenure and promotion process. Hamermesh & Parker (2005) examined whether, adjusting for other determinants, whether instructor attractiveness explain differences in course evaluation scores—and thus on earnings differences. In this assignment, you will use a sample of their data to examine this question.

Please submit your responses to each of the questions below in a printed document. Also, please adhere to the following guidelines for further formatting your assignment:

- All graphics should be resized so that they do not take up more room than necessary and should have an appropriate caption and labels.
- Any typed mathematics (equations, matrices, vectors, etc.) should be appropriately typeset within the document using Equation Editor, Markdown, or LaTeX.

This assignment is worth 20 points. Each question is worth one point unless otherwise noted.

For this assignment, you will use the file *beauty.csv*. This file contains data collected from student evaluations of instructors' beauty and teaching quality for several courses at the University of Texas. The teaching evaluations were conducted at the end of the semester, and the beauty judgments were made later, by six students who had not attended the classes and were not aware of the course evaluations. The variables are:

- `prof`: Professor ID number
- `avgeval`: Average course rating
- `btystdave`: Measure of the professor's beauty composed of the average score on six standardized beauty ratings
- `tenured`: 0 = non-tenured; 1 = tenured
- `nonenglish`: 0 = native English speaker; 1 = non-native English speaker
- `age`: Professor's age (in years)
- `female`: 0 = male; 1 = female
- `students`: Number of students enrolled in the course
- `percentevaluating`: Percentage of enrolled students who completed an evaluation

These source of these data is: Hamermesh, D. S. & Parker, A. M. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review, 24*, 369–376. The data were made available by: Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Preparation

Fit the following regression models using R. You will use the output from the fitted models to answer the questions in the assignment.

- Model 1: `avgeval ~ btystdave`
- Model 2: `avgeval ~ btystdave + percentevaluating`
- Model 3: `avgeval ~ btystdave + percentevaluating + students`

---

1. Create a table that reports the results for the three regression models fitted in this assignment. This table should be similar in form to the tables of multiple regression results we created in class—with models in columns and predictors in rows. You might also want to consult the book *Presenting Your Findings*. Be sure to include the coefficients, SE, level of significance, and model summaries (RMSE and R^2) for each fitted model. Also, be sure that the table is appropriately captioned. If the table is too long, change the page orientation in your word processing program to "Landscape", rather than changing the size of the font. **(3pts.)**
2. Based on results presented in the regression table, explain why it would be appropriate to adopt Model 3.

Use the results from Model 3 to answer the remainder of the questions on this assignment.

Description

3. Report the regression equation from fitting Model 3. Use Equation Editor (or some other program that correctly types mathematical expressions) to typeset the equation correctly.
4. Using output from the ANOVA table, compute and report the value for R^2. Show your work for full credit.
5. Interpret the value of R^2 using the context of the data.
6. Interpret each of the coefficient values for the three variables in Model 3. **(3pts.)**

Model-Level Inference

7. Using symbols, write the omnibus null hypothesis that is tested by the F-statistic in this analysis in two different manners: (1) using the coefficient parameters used in the regression model, and (2) using the variance accounted for parameter. **(2pts.)**
8. Write no more than 2–3 sentences (to be included in a publication) that summarizes the results of the omnibus analysis. A summarization of the results includes a written description of what is being tested by the F-test and the statistical results. At a minimum report the F-statistic, *df*, and p-value. A summary should also indicate what the statistical results suggest about the tenability of the null hypothesis and what this means about the potential relationship between the predictors and the outcome. **(2pts.)**

Coefficient-Level Inference

9. Using symbols, write the null hypothesis that is tested by the t-statistic for the `btystdave` coefficient.

10. Based on the results of the t-test, what do the data suggest about the effect of beauty on course evaluations? Explain. **(2pts.)**

Displaying Results

11. Create a publication quality plot that displays the results from Model 3. For this plot, put the beauty predictor on the x-axis. Control out the effects of the percentage of students evaluating the course by setting its value to the mean. Show two separate lines to show the effect of class size (a small and large class). These lines should have different linetypes or use different colors so that they can easily be differentiated. **(3pts.)**