Assignment 08

Analysis of Covariance in Multiple Regression: Part II

The file *colleges-bordering-mn.csv* contains institutional data for 104 colleges and universities in the five state area (MN, IA, WI, ND, and SD). These data were collected by the Department of Education for the 2013 College Score Card. The variables are:

- `name`: Name of college/university
- `tuition_in_state`: In-state tuition and fees
- `tuition_out_state`: Out-of-state tuition and fees
- `state`: State postal abbreviation (IA = Iowa, MN = Minnesota, ND = North Dakota, SD = South Dakota, WI = Wisconsin)
- `public`: Is this a public school? (1 = yes, 0 = no)
- `main`: Is this the main campus? (1 = yes, 0 = no)
- `admission`: Admission rate
- `act75`: 75th percentile of the ACT cumulative scores
- `avg_fac_salary`: Average faculty salary (per month)
- `completion`: Four-year completion rate for first-time, full-time students
- `pct_pell`: Percentage of undergraduates who receive a Pell grant

In this assignment, you are going to focus on the question of whether there are differences in the average in-state tuition across the five states. Use these data to answer each of the following questions. Each question is worth one point unless otherwise noted. The entire assignment is worth 12 points.

Create Dummy Variables

Create five dummy variables, one for each state, for the analysis.

Unadjusted Group Differences Model: ANOVA

Fit the regression model that uses the dummy predictors for state to predict variation in average in-state tuition. In this model, use Minnesota as the reference group.

1. Write the fitted regression equation based on the `summary()` output. Write the regression equation using Equation Editor (or some other program that correctly types mathematical expressions). Be sure the equation is labeled and numbered according to the APA format.
2. Are there any statistically reliable differences in the average in-state tuition between Minnesota and its bordering states? Explain.
3. Report and interpret the R^2 value for this model.
4. Which group difference(s) is/are not represented in this fitted model?

Adjusted Group Differences Model: ANCOVA

Again, fit the regression model that uses the dummy predictors for state to predict variation in average in-state tuition, but this time control for differences in (1) sector (public vs. private), (2) ACT scores, and (3) percentage of students on Pell grants. Again, use Minnesota as the reference group.

5. Write the fitted regression equation based on the `summary()` output. Write the regression equation using Equation Editor (or some other program that correctly types mathematical expressions). Be sure the equation is labeled and numbered according to the APA format.
6. Are there any statistically reliable differences in the average in-state tuition between Minnesota and its bordering states after controlling for these other predictors? Explain
7. Report and interpret the R^2 value for this model.
8. Create a table (suitable for publication) that presents each of the possible pairwise contrasts (hypothesis) of interest, the unadjusted p-values, and the Benjamani–Hochberg adjusted p-values for the controlled differences. (Note: To obtain all of these, you may need to fit additional models.) **(2pts.)**
9. Based on the Benjamani–Hochberg adjusted p-values, which states have a statistically reliable difference in their average in-state tuition after controlling for the other predictors in the model?
10. Create a visualization of the information that you reported in the table in Question 8. From the visualization, a reader should be able to cull information about each state's average in-state tutiion value and also how the five states differ (or don't) from each other. **(2pts.)**