

COMPARING GROUPS: PART II

EPsy 8251

Assignment #4

In this assignment, you will again examine the salary data from 28 school districts in Minnesota, however in this assignment, you will focus on the regression model. These data are in the *MNsalary.csv* file. Please submit your responses to each of the questions below in a printed document. All graphics should be resized so that they do not take up more room than necessary and all should have an appropriate caption. Any equations should be appropriately typeset within the document. There are 11 points possible for the assignment (each question is worth one point).

DUMMY CODING

1. Create a dummy variable for the district type (`DistType`) variable in which public school districts take the value 1. Include the R syntax you used to create this variable.
2. Compute the mean for your dummy variable. Explain what that value tells us.
3. Fit a regression model using your dummy variable to predict teacher salary. 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.
4. Interpret the value of the intercept from the regression equation using the context of the data.
5. Interpret the value of the slope from the regression equation using the context of the data.
6. What do the results of the regression analysis tell us about whether or not there are statistically reliable differences in the average teacher salary between public and non-public school districts? Explain.
7. Report and interpret the R^2 value from the fitted regression.

EFFECTS CODING

8. Create a variable for the district type (`DistType`) variable in which public school districts take the value 1 and non-public school districts have the value -1 . Fit a regression model using this variable to predict teacher salary. 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.
9. Interpret the value of the intercept from the regression equation using the context of the data.
10. Interpret the value of the slope from the regression equation using the context of the data.
11. Explain why the effect of the predictor (i.e., slope) in this model is twice as large as the effect from your dummy coded model.