Homework 1
Due Thursday, January 28 10 AM (pdf format please)

For all questions, please show your work or include a copy of the output, whichever is relevant. Please write your answers in report form including relevant statistics, as if you were describing results in a published study. Be sure to describe each finding in terms of the research problem. Most responses to a particular question should be approximately one paragraph in length. Text should be typed, but please do not waste your time typing hand computations. All answers should be in your own words.

1. The data for the questions below come from a Stanford University structured play intervention study that investigated the effects of conflict resolution at recess, class time, and after school in low income schools (download from http://web.pdx.edu/~newsomj/data.htm). The data set contains five variables from the teacher survey portion of the study. The dependent variable for our analysis will be the emotional regulation subscale of a social competence scale (emoreg), which has values that range between 0 and 3. The remaining variables in the data set include treat, an indicator of the conflict resolution intervention (treat = 1) or control (treat = 0), the teacher's perceived bullying atmosphere at school (bullying), with higher scores indicating more bullying, timetran, the estimated minutes between recess and classroom learning, and playsup, rating of support for play time at school (“It is important for students to have the opportunity to play during the school day”) on a 5-point scale of agreement.

a. Use R and SPSS to obtain a correlation matrix of all of the correlations in the data set. (Skip any scatterplots for now—we will use some later.) Describe and explain your results. Include both printouts with your response (but only one write-up of the results is needed).

b. Use R and SPSS to conduct a simple regression analysis in which support (playsup) predicts emotional regulation (emoreg). Be sure to obtain standardized and unstandardized coefficients, R-squared, and confidence intervals. Report your results, including standardized and unstandardized coefficients, significance tests, confidence intervals, $R^2$, $F$-test, and provide an interpretation of the coefficients and $R^2$.

c. Using R and SPSS, conduct a simultaneous multiple regression analysis using play support and transition time (timetran) as predictors of emotional regulation. Report your results, including standardized and unstandardized coefficients, significance tests, confidence intervals, $R^2$, $F$-test, and provide an interpretation of the coefficients and $R^2$.

d. Using R and SPSS, conduct a hierarchical multiple regression, entering play support in the first block and then transition time in the second block. Report your results, including standardized and unstandardized coefficients, significance tests, confidence intervals, $R^2$, $F$-test, $R^2$ change, and $F$ change, and provide an interpretation of the coefficients and $R^2$. Include a comment that compares your results to those obtained in the simultaneous regression above. How are they similar or different and why? (Be specific, give values)

e. Based on your results from the hierarchical multiple regression in the previous problem, sketch a Venn diagram and label it with the approximate percentage of variance in the dependent variable accounted for by each of the predictors and the percentage variance overlap in the two predictors (be sure to clearly indicate which quantity from the output goes with which percentage).

f. Using R or SPSS, test a simple regression model with treatment (treat) predicting emotional regulation. Report your results, including standardized and unstandardized coefficients, significance tests, confidence intervals, $R^2$, $F$-test, and provide an interpretation of the coefficients and $R^2$. What are the special interpretations of the intercept and slope in this model (be specific, give values)?
g. Use R or SPSS to test a multiple regression with the treatment (treat) predicting transition time (timetran) and perceived bullying atmosphere (bullying) using a hierarchical regression with treatment entered first. Before testing the model, center the bullying variable. Report your results of just the final step, including an interpretation of the intercept, unstandardized slopes, slope significance tests, \( R \)-squared, and \( F \)-test. Discuss the effect of adding bullying to the model in terms of its impact treatment effect and describe what centering the bullying variable did to the interpretation of the results.

h. In R or SPSS, conduct an ANCOVA analysis comparing treatment groups (treat), controlling for bullying (use the original bullying variable, not the centered variable), to examine mean differences in transition time. Report and interpret your findings, including the adjusted means. Describe how the output from the ANCOVA and the regression model from the last problem compare, giving specific values to support your statements about the similarities and differences. ANCOVA

2. Use your own data for the following problems (if you don’t have a data set you can use, I can find one for you). Make sure your dependent variable is a continuous variable (ordinal with at least 5 ordinal values, interval, or ratio scale). You will need to have at least two continuous predictors.

a. Use SPSS or R to test a simple regression with one predictor of your outcome. Provide a very brief statement about the purpose of the analysis (no longer than one paragraph). In your analyses, be sure to obtain and interpret standardized and unstandardized coefficients, R-square, R-square significance, confidence intervals, and a scatterplot in terms of the research problem.

b. Using R or SPSS, run the same simple regression analysis but obtain a histogram of residuals and residual plot using \( X \) values (x-axis) with standardized residual values (y-axis), and make an assessment about nonlinearity, homoscedasticity, and outliers. Request the casewise regression diagnostics that were discussed in class, obtaining at least one \( X \) outlier diagnostic index, one \( Y \) outlier diagnostic index, and one influence statistic. Use syntax/code to print the diagnostics or save the diagnostics to a data file to obtain the diagnostics for each case. [Please include only relevant excerpts of the individual case diagnostics to save space]. Interpret the regression output results, diagnostics, and discuss any outliers or other potential problems with the data.

c. Use SPSS or R to test a simultaneous multiple regression with the predictor used in the previous problem plus one other predictor. Report your results, including standardized and unstandardized coefficients, significance tests, confidence intervals, \( R \)-squared, \( F \)-test, and provide an interpretation of the coefficients and \( R \)-squared.

3. Read one of the following articles (copies available from the class website (http://web.pdx.edu/~newsomj/mvclass/) and write two paragraphs summarizing the article. First, describe the study design (e.g., randomized experiment, non-equivalent control group design, cross-sectional survey; for a quick refresher, see http://sphweb.bumc.bu.edu/otlt/mpi/programevaluation/ProgramEvaluation7.html) and purpose of the study in your own words. Be sure to include who/what was studied (e.g., who were the participants?) and the number of cases. Then, choose one regression analysis used in the article that you have learned about in the course so far (i.e., simple or multiple regression), and, in your own words, describe the hypothesis that is being tested, the results obtained, and what the findings mean. Be sure to include the relevant statistical values and whether the results were significant. Write your paragraphs as if you were describing results in a published article or reporting someone else’s results as in a review article.

