Assignment 03

More Simple Regression

The file *teengamb.csv* contains data collected on a study of teenage gambling in Britain. The data consists of five variables and 47 observations. The variables are:

- `sex`: 0 = male; and 1 = female
- `status`: Socioeconomic status score based on parents' occupation
- `income`: Income in British pounds per week
- `verbal`: verbal score in words out of 12 correctly defined
- `gamble`: expenditure on gambling in British pounds (per year)

These source of these data is: Ide-Smith, S. G., & Lea, S. E. G. (1988). Gambling in young adolescents. *Journal of Gambling Behavior, 4* (2), 110–118.

Use the data in *teengamb.csv* to answer each of the following questions. In this homework assignment, you will be focusing on the relationship between income (predictor) and expenditures on gambling (outcome). This assignment is worth 15 points. Each question is worth 1 point unless otherwise noted.

To begin, fit a regression model using income to predict expenditures on gambling (i.e., regress expenditure on gambling on income).

1. Compute and report the value for R^2.
2. Interpret the value of R^2 using the context of the data.
3. Using symbols, write the omnibus null hypothesis related to the intercept and slope paramters used in the regression model that is tested by the F-statistic in this analysis.
4. Using symbols, write the null hypothesis related to the variance accounted for that is tested by the F-statistic in this analysis.
5. 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 income and expenditures on gambling. **(2pts.)**
6. Using symbols, write the null hypothesis that is tested by the t-statistic for the intercept.
7. Based on the results of the t-test, what do the data suggest about the tenability of the null hypothesis for the intercept? Explain. **(2pts.)**
8. Using symbols, write the null hypothesis that is tested by the t-statistic for the slope.
9. Based on the results of the t-test, what do the data suggest about the tenability of the null hypothesis for the slope? Explain. **(2pts.)**
10. Would the inferences in this analysis apply to all teenagers in Britain? Explain.
11. Create a publication quality plot that displays the regression line based on the analysis. This plot should also include the data plotted as a scatterplot. The data should be semi-transparent, and the regression line should be completely opaque (non-transparent). Be sure to appropriately number, label, and re-size the plot for publication. **(2pts.)**