Note: "OIS" refers to OpenIntro Statistics, the textbook for this class.

Problem 1 (20 points)

Before beginning the problems, please load the `prestige` dataset into R using `prestige_df <- read.csv("https://vincentarelbundock.github.io/Rdatasets/csv/carData/Prestige.csv")`– alternatively, you can download and install the `carData` package and load the dataset using `data(Duncan)`. You can read more about this dataset at https://vincentarelbundock.github.io/Rdatasets/doc/carData/Prestige.html.

(a) How many variables are there in this dataset? How many are numerical, and how many are categorical? (1 pt)

(b) Fit a model predicting prestige from education and type. (1 pt)

(c) State the model's adjusted R^2. (1 pt)

(d) Predict the prestige of a white collar position an average of 12 years education. (1 pt)

(e) Run a hypothesis test for the variable education at $\alpha = 0.05$. Be sure to include the appropriate hypotheses, report the test statistic and p-value, and clearly interpret your decision in the context of the problem. (4 pts)

(f) Interpret the point estimate for the coefficient on education. Is increased education expected to lead to an increase or decrease in perceived prestige? How do you know? (2 pts)

(g) Plot the residuals for this model against the fitted values. Is this model appropriate for the data? Explain. (3 pts)

(h) Consider the variable type. What is the reference level? Interpret the point estimate for `typeprof`. (2 pts)

(i) Fit models using `education` alone and `type` alone (2 pts). Out of the three models you have picked, which one describes the data best? (1 pt) Why? (2 pts)