Case 1: Wages of Workers

In the United States, age discrimination is illegal, but its occurrence is hard to prove (The New York Times, August 7, 2017). Even without discrimination, it is widely believed that wages of workers decline, as they get older. A young worker can expect wages to rise with age only up to a certain point, beyond which wages begin to fall. Ioannes Papadopoulos works in the human resources department of a large manufacturing firm and is examining the relationship between wages (in $), years of education, and age. Specifically, he wants to verify the quadratic effect of age on wages. He gathers data on 80 workers in his firm with information on their hourly wage, education, and age, which is included in the “Wages” sheet of InClassExercises.xlsx file.

1. Read the data into a new tibble called myData. (*Hint: We will use read\_excel() function, which is a part of “readxl” package. readxl package comes with tidyverse, but you will still need to load readxl explicitly, because it is not a core tidyverse package. If you do not have it listed under your packages, please install it and then load it. Once you load the package, you may want to look at the help menu for how to use read\_excel() function*)
2. Produce a scatterplot matrix using the function pairs(). What type of relationship(s) do you anticipate between the variables?
3. Create both a linear model and a quadratic model to predict wage. Analyze both models and identify which one is a better fit?
4. Use the best fitting model to predict hourly wages for someone with 16 years of education and age equal to 30, 50, or 70.
5. According to the best fitting model, at what age will someone with 16 years of education attain the highest wage? (*Hint: You may want to plot the estimated value line, and may need to use some basic calculus concepts*)