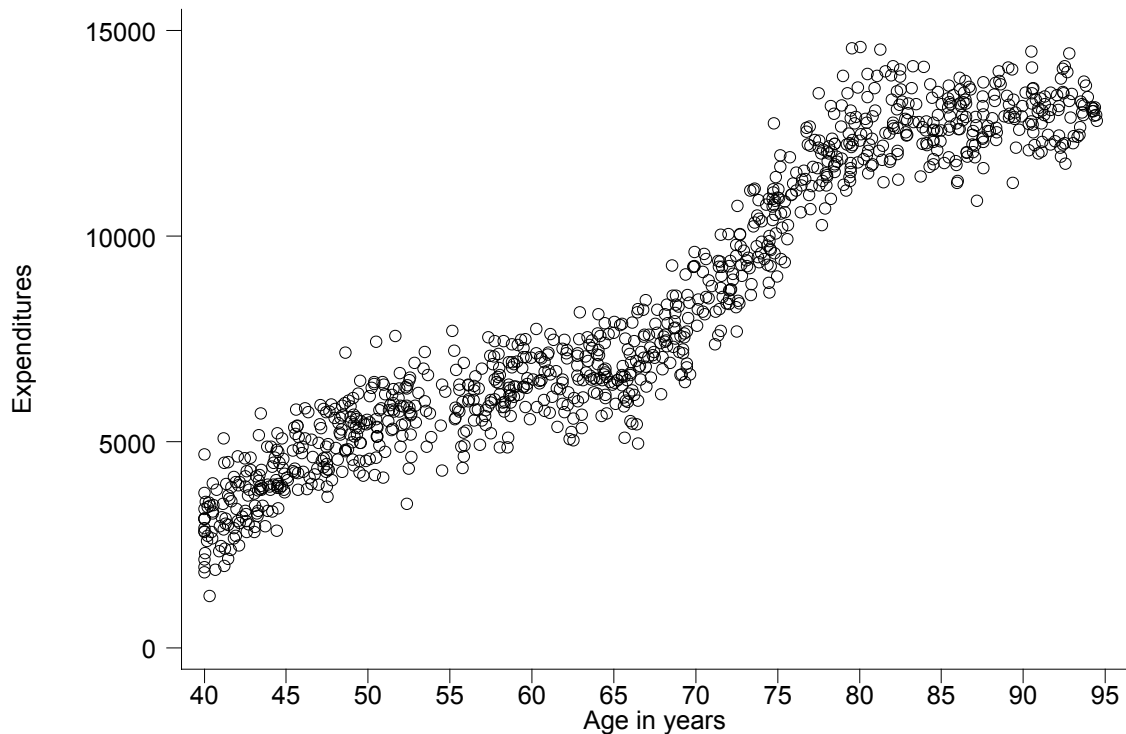


## Some more work with Linear Splines

Suppose you are interested in modeling the total health care expenditures for persons aged 40 to 94. You are able to take a random sample of 1000 US non-institutionalized citizens and obtain all their medical records in the past year. For each subject, you record the total health care expenditures and the subject's age. The graph below displays the data.



You know that after a person reaches the age of 65, Medicare begins to supplement medical care. You also have information to suggest that doctors tend not to perform major surgeries on patients who are over 80 years of age.

You have two specific scientific questions in mind:

Question 1: Is the rate at which total medical expenditures increase with age different for persons less than 65 years of age compared to persons who are 65 to 80 years of age?

Question 2: Is the rate at which total medical expenditures change with age different for persons 65 to 80 years of age compared to persons over 80?

Based on this information, propose a multiple linear regression model to describe the association between total medical expenditures and age. Write out the form of your proposed model, defining all explanatory variables, and interpret the regression coefficients in your model. How would you answer the two scientific questions?

Model:

Interpretation of Regression Coefficients