Regression Project
Due Monday April 24th

Description of Dataset

The dataset in question contains data from 178 countries. It contains information about the countries, highest point, lowest point (Minimum.elevation), total area, land area, water area and population circa 2023. Throughout this project our aim will be to construct a model which estimates the highest (denoted as Maximum.elevation). The data comes from various wikipedia tables and some countries may be excluded simply for having multiple names. For purposes of this assignment we will assume that the list is complete

1. Single Variable regression: Use single variable regression to predict The height of the highest point in a country from. In each case discuss the model. How significant is the slope? Does the slope make intuitive sense? Does it correspond to your intuition from the scatter plot? Discuss how you'd deal with outliers?
   (a) The country's total area
   (b) The country's lowest point
2. Several Variable regression/ Use the techniques of multiple variable regression to fix a model using some of the variables given. For example, the lowest point, the total area, the land and water area and the population. Discuss why you left out certain variables and left others in.
3. Overall, if you had to accurately predict the maximum elevation, which model (from the above ones) would you choose and why?