Problem 1

Consider a situation where an anthropologist is studying a population of people who have lived in mountain isolation for several generations. She is interested in studying the relationship of the height at 18 for males, to the following variables:

Variables: length at birth, mother's age at 18, father's age at 18, and paternal and maternal grandparents age at 18 (7 predictors).

All heights and lengths are in inches. A random sample of 20 males of age 18 or more was obtain from the study population, and the you can find the dataset in "Problem 1" folder.

Are there any observations that have high leverage values? If so, what are the sample items numbers?

Are there any observations that are outliers? If so, what are the sample items numbers?

Are there any observations that are influential observations? If so, what are the sample items numbers?

Compute dffits and dfbetas and discuss your findings.

Write a short summary discussing unusual observations. Explain what an investigator should do about them.

Is there any indication of multicollinearity among the predictor variables?