Body Fat Project Presentation

Group-20

October 20, 2022

Summary of Data Cleaning

▶ We drop the sample whose estimated body fat is negative. The estimated fat is infered from Siri's equation:
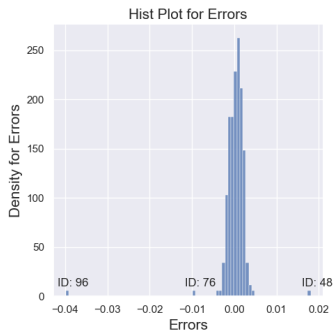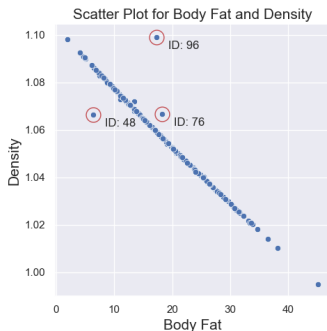
$$\text{Fat} = \frac{495}{\text{Density}} - 450$$

The sample's statistics are listed below:

IDNO	BODYFAT	DENSITY	Estimated Body Fat
182	0.0	1.1089	-3.6117

Summary of Data Cleaning

▶ From the scatter plot of body fat and density, we find there are 3 obvious outliers. We use the Siri's equation to compute the estimated body fat and substract it by corresponding the original BODYFAT to get errors.

Summary of Data Cleaning

▶ We use $3 - \sigma$ criterion on the errors to find out these outliers. Finally, we use estimated body fat from Siri's equation to replace the original body fat for these samples.

IDNO	BODYFAT	Estimated Body Fat	Error
48	6.4	14.1350	7.7350
76	18.3	14.0915	-4.2085
96	17.3	0.3685	-16.9315