

Body Fat - Rule of Thumb

The data set contains 17 measurements from 252 observations.

In this report, our goal is to get a clear and robust equation to estimate body fat.

Overall Finding

$$Y = 2 * X + 1111$$

where x is ... and y is ...

Data Cleaning

```
BodyFat <- read.csv("./data/BodyFat.csv"); attach(BodyFat)
```

We will build a linear model to get the numerical result for the body fat estimation.

Thus the outcome is BODYFAT (If not valid, can be approximately translated from DENSITY)[@Siri]

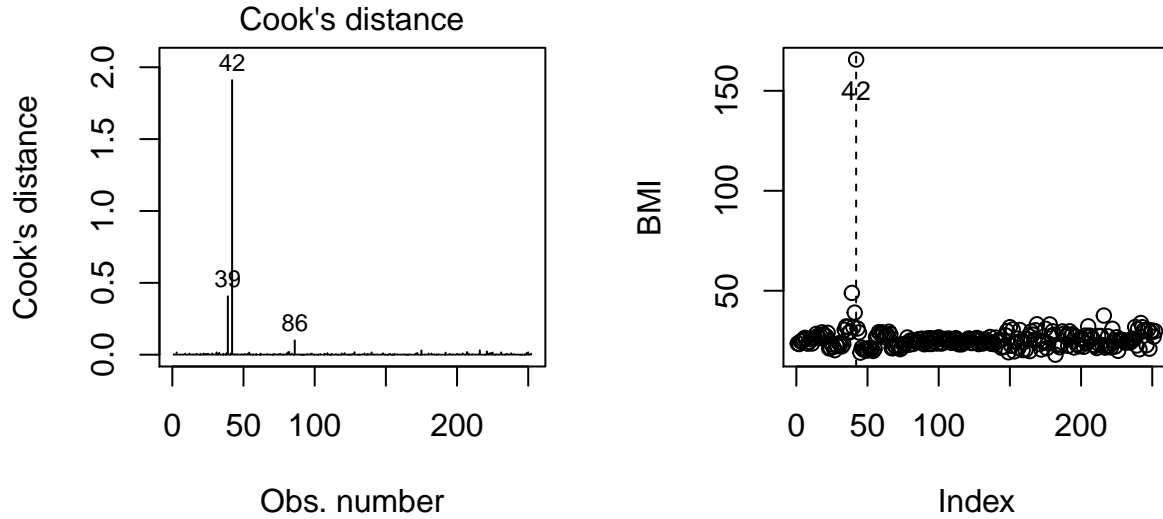
Invalid observation, Outlier and Influential Points

```
dev.new(width = 5, height = 3)
plot(sort(BODYFAT), xlab = "", ylab = "Body Fat")
```

By analysing the BODYFAT variable, we can see that the ID=182 and ID=216 observations have extreme values of body fat, separately 0 and 45.1.

We consider that these two values are not valid and try to transform the second-class variable DENSITY to BODYFAT with equation [Siri] $B = 495/D - 450$ and get -3.61 and 47.49 separately, which are still not in a normal range of body fat. Thus we **delete** these two observations from the data set.

```
fit <- lm(BODYFAT ~ .-IDNO-DENSITY, data = BodyFat)
par(mfrow=c(1,2))
plot(fit, which=4) ## Cook's distance
BMI <- WEIGHT * 0.453592 / (HEIGHT*2.54*0.01)^2
plot(BMI); abline(v = 42, lty = 2); text(42,150, labels = "42", cex = 0.8)
```



By analysing the Cook's distance of the linear model build from `lm(BODYFAT ~ .)` (all possible X despite of ID and **DENSITY**), we look into details of the outlier candidates with ID 42, 39 and 86.

Observation with ID=42 is weighted 205 lb and 29.5 in tall. When we calculate their BMI, we can see that this observation is very abnormal, so we consider this observation as an invalid record and delete it from the model.

Obs. 39 is a very fat guy with most of his record as the maximum one in the data and Obs. 86 looks normal.

For other variables with extreme values, we also do some explanatory analysis. The IDs of these observations are listed below. They will be checked in the following sections.

ID	Abnormal	Action
182	BodyFat = 0 (failed to be recalculated from Density)	Delete
42	BMI = 165 (Very Extreme BMI)	Delete
39	Weight = 363.15 (Extreme Fat)	Check
79	Age = 81 (Extreme Old)	Check

Statistical Modeling