

Bodyfat Prediction

Yilun Chen
Jiyun Chen
Huiyu Jiang

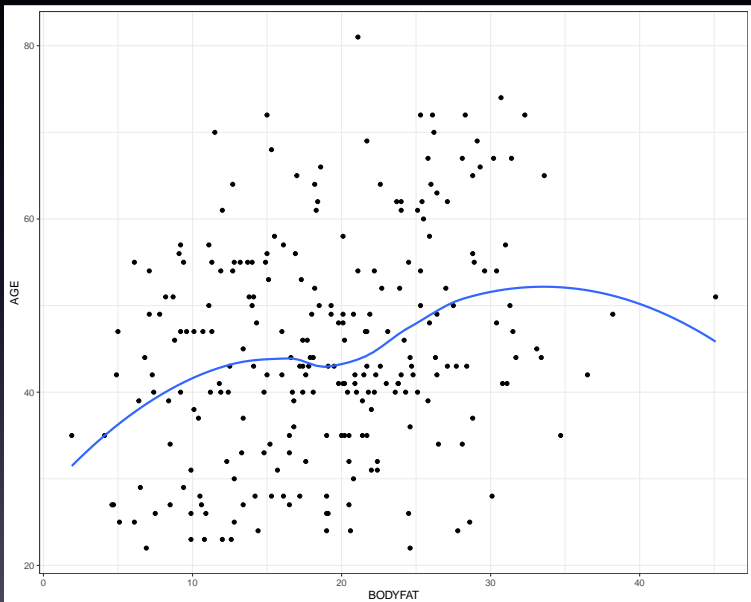
University of Wisconsin-Madison

February 7, 2018

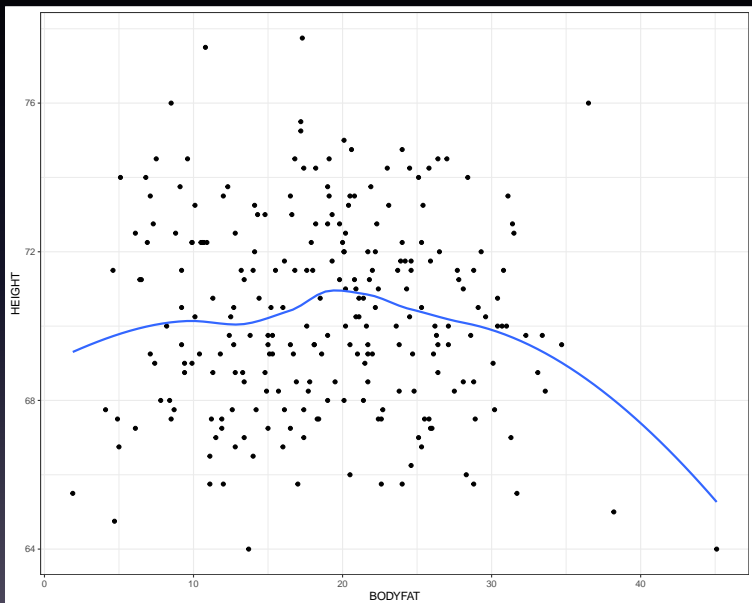
Outline

- Background and Data Description
- Model Selection and Data Processing
- Variable Selection
- Model Diagnostic and Prediction

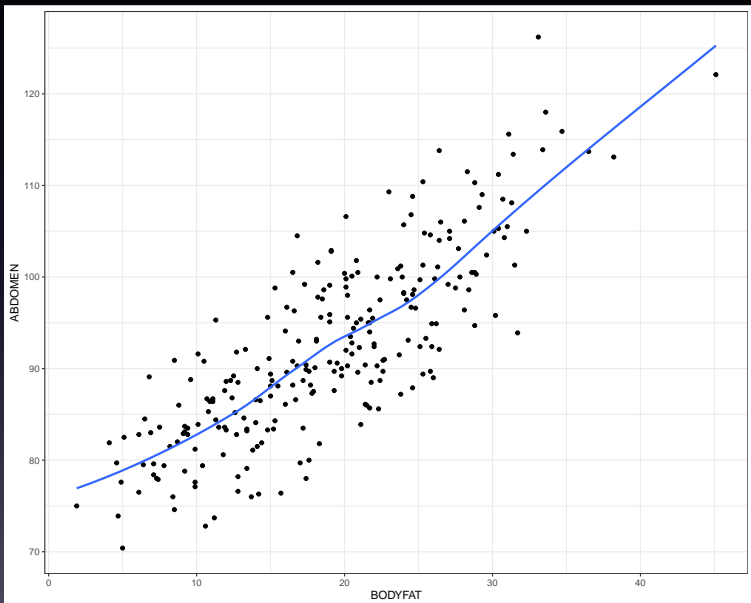
Model Selection



Model Selection



Model Selection



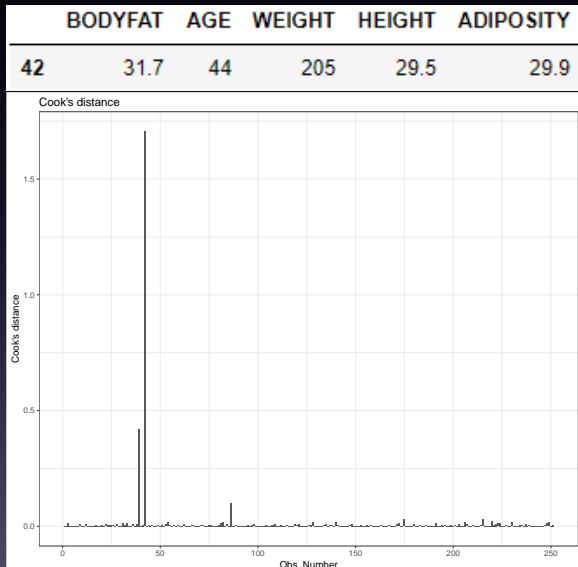
Data Processing

- false record:

	BODYFAT	AGE	WEIGHT	HEIGHT	ADIPOSITY
182	0	40	118.5	68	18.1

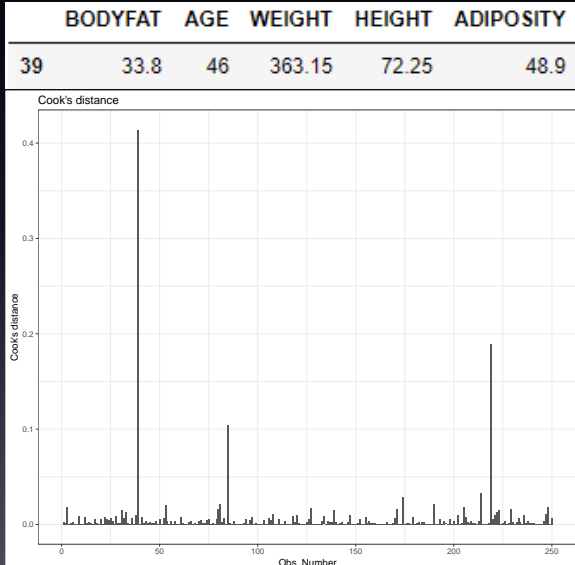
Data Processing

- check outliers:



Data Processing

- check outliers:



Variable Selection

- AIC Stepwise:

```
Call:
lm(formula = BODYFAT ~ ABDOMEN + WEIGHT + WRIST + BICEPS + AGE,
    data = data_body)

Residuals:
    Min       1Q   Median       3Q      Max
-9.4993 -2.6962 -0.2706  3.1972  8.7266

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -23.71669     7.06737  -3.356 0.000918 ***
ABDOMEN       0.83796     0.06314  13.272 < 2e-16 ***
WEIGHT      -0.08539     0.03055  -2.795 0.005607 **
WRIST       -1.65891     0.45052  -3.682 0.000285 ***
BICEPS       0.27709     0.13981   1.982 0.048619 *
AGE          0.03881     0.02615   1.484 0.139047

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.95 on 243 degrees of freedom
Multiple R-squared:  0.7356,    Adjusted R-squared:  0.7301
F-statistic: 135.2 on 5 and 243 DF,  p-value: < 2.2e-16
```

Variable Selection

