BHao_Assign11

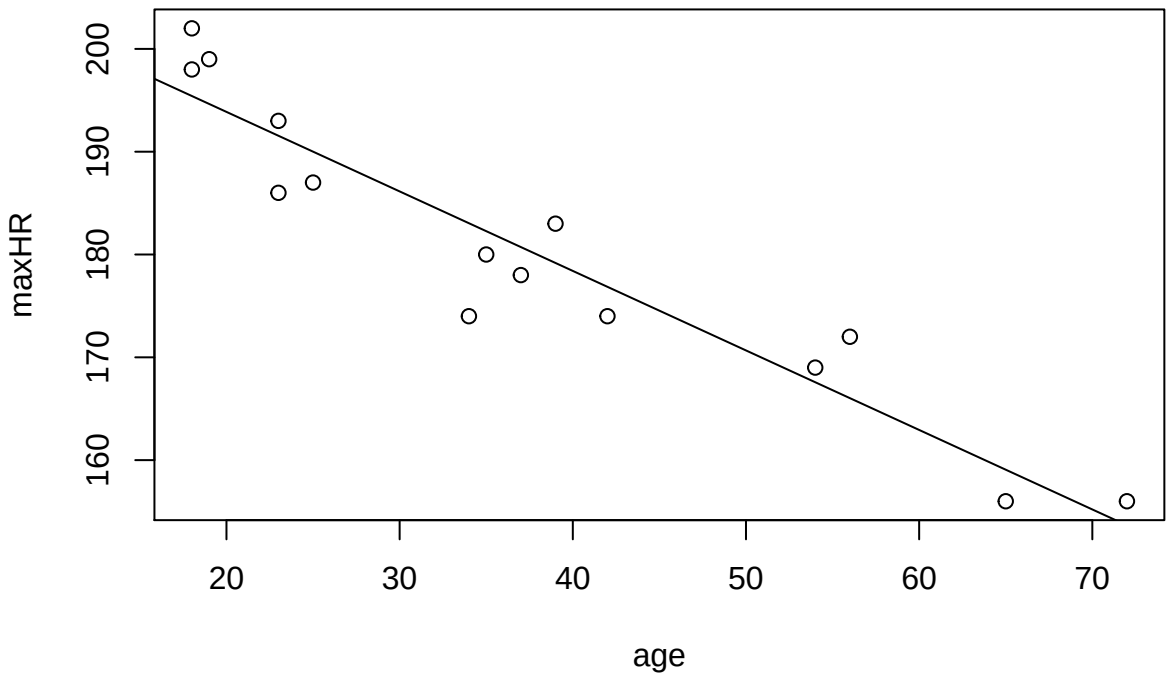
- Fitting max heart rate to age using lm results in the following equation: max heart rate = 209.3416 - 0.7744 * age
- The effect of age on max heart rate is significant
- The significant level is at the 0.001 level

```
age = c(18,23,25,35,65,54,34,56,72,19,23,42,18,39,37)
maxHR = c(202,186,187,180,156,169,174,172,156,199,193,174,198,183,178)

age_hr = lm(maxHR ~ age)
summary(age_hr)
```

```
##
## Call:
## lm(formula = maxHR ~ age)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.045 -2.932  1.424  3.201  6.580
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 209.34158    2.90977   71.94  < 2e-16 ***
## age          -0.77344    0.07101  -10.89 6.62e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.646 on 13 degrees of freedom
## Multiple R-squared:  0.9012, Adjusted R-squared:  0.8936
## F-statistic: 118.6 on 1 and 13 DF,  p-value: 6.617e-08
```

```
plot(maxHR ~ age)
abline(age_hr$coefficients[1], age_hr$coefficients[2])
```



- Fitting a multiple regression the auto data set results in the equation below: mpg = 45.2511 - 0.0060 * disp - 0.0436 * hp - 0.0053 * weight - 0.0231 * accel
- hp and weight have significant impacts
- hp and weight have 0.001 and 0.01 significance levels, respectively
- the standard errors for disp, hp, weight and accel are shown below or 0.0067, 0.0166, 0.0008 and 0.1256, respectively
- unsurprisingly the significance levels and standard errors are much higher for the smaller sample vs. the full data set

```
mpg = read.table('auto-mpg.data')
names(mpg) = c('disp', 'hp', 'weight', 'accel', 'mpg')
str(mpg)
```

```
## 'data.frame':    392 obs. of  5 variables:
##  $ disp  : num  307 350 318 304 302 429 454 440 455 390 ...
##  $ hp    : num  130 165 150 150 140 198 220 215 225 190 ...
##  $ weight: num  3504 3693 3436 3433 3449 ...
##  $ accel : num  12 11.5 11 12 10.5 10 9 8.5 10 8.5 ...
##  $ mpg   : num  18 15 18 16 17 15 14 14 14 15 ...
```

```
lm_model = lm(mpg ~ ., data = mpg)
summary(lm_model)
```

```
##
```

```
## Call:
## lm(formula = mpg ~ ., data = mpg)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -11.378  -2.793  -0.333   2.193  16.256
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 45.2511397  2.4560447  18.424  < 2e-16 ***
## disp        -0.0060009  0.0067093  -0.894  0.37166
## hp          -0.0436077  0.0165735  -2.631  0.00885 **
## weight      -0.0052805  0.0008109  -6.512  2.3e-10 ***
## accel       -0.0231480  0.1256012  -0.184  0.85388
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.247 on 387 degrees of freedom
## Multiple R-squared:  0.707,  Adjusted R-squared:  0.704
## F-statistic: 233.4 on 4 and 387 DF,  p-value: < 2.2e-16
```

```
mpg_sample = mpg[sample(nrow(mpg), 40), ]
lm_model_sample = lm(mpg ~ ., data = mpg_sample)
summary(lm_model_sample)
```

```
##
## Call:
## lm(formula = mpg ~ ., data = mpg_sample)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.4757 -3.5730  0.0131  2.4070 10.8089
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 31.464195   7.391581   4.257 0.000148 ***
## disp         0.004713   0.019930   0.237 0.814421
## hp           0.007880   0.054953   0.143 0.886805
## weight      -0.007119   0.002792  -2.549 0.015326 *
## accel        0.705022   0.399834   1.763 0.086582 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.487 on 35 degrees of freedom
## Multiple R-squared:  0.7054, Adjusted R-squared:  0.6717
## F-statistic: 20.95 on 4 and 35 DF,  p-value: 6.868e-09
```