Statistical Methods 2 Homework 1

2023-01-19

Loadings possible useful libraries

```
library(ggplot2)
library(Sleuth3)
library(MASS)
```

Problem 2)

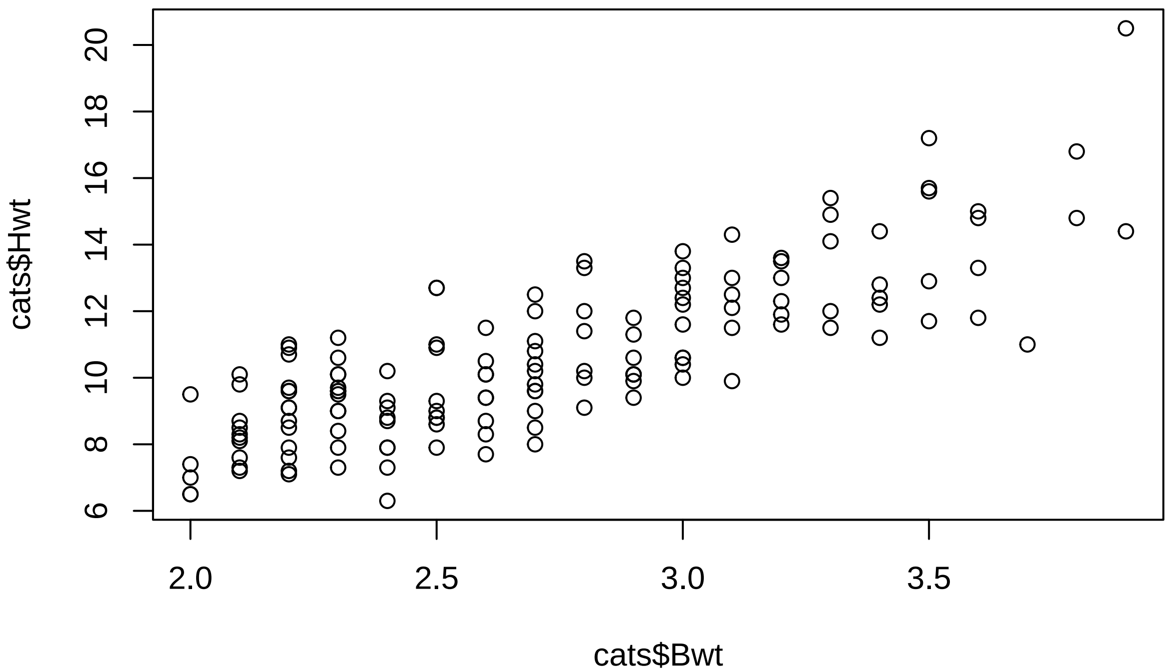
```
cats <- cats
```

a)

The explanatory variable here would be the body weight (Bwt). The response variable would be hearth weight (Hwt). This is because we are trying to investigate if body weight is a good variable to predict heart weight in cats.

b)

```
plot(cats$Bwt,cats$Hwt)
```



There appears to be a somewhat positive linear relationship between cats body weight and heart weight. There appears to be less variance the lower the weight of the cat, and higher variance as the body weight increases. However, this effect doesn't seem too large.

c)

Yes, I think using a linear regression model would be appropriate as it has a fairly positive linear relationship and the variance is fairly consistent (and possibly normal) for the distribution around heart weight. There is also a reasonably large sample size. It also likely reasonably fair to assume independence between observations in this study.

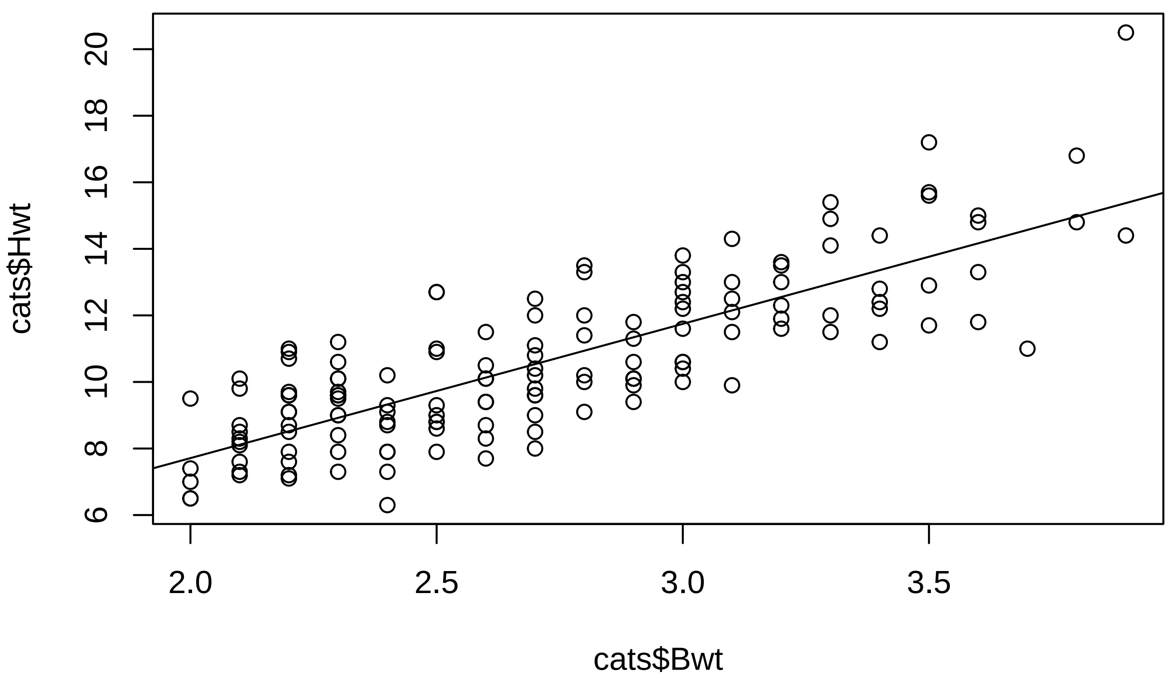
d)

```
fit <- lm(Hwt ~ Bwt, data = cats)
summary(fit)
```

```
##
## Call:
## lm(formula = Hwt ~ Bwt, data = cats)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5694 -0.9634 -0.0921  1.0426  5.1238
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.3567     0.6923  -0.515    0.607
## Bwt           4.0341     0.2503  16.119   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.452 on 142 degrees of freedom
## Multiple R-squared:  0.6466, Adjusted R-squared:  0.6441
## F-statistic: 259.8 on 1 and 142 DF,  p-value: < 2.2e-16
```

```
plot(cats$Bwt,cats$Hwt)
abline(fit)
```



The intercept parameter was estimated to be $\hat{\beta}_1 = -0.3567$ and $\hat{\beta}_0 = 4.0341$.

e)

In cats, for every 1 unit increase in body weight (Bwt) in kg it is associated with mean heart weight increases (Hwt) by 4.0341 grams. With a base case of 0 kg body weight for a cat, the heart weight would be -.3567 grams, which doesn't make sense to have a negative weight however this is okay as it also doesn't make sense to have a 0 weight.

f)

```
B_1_hypothesis <- 3.75
B_1_estimated <- fit$coefficients[2]
difference <- B_1_estimated - B_1_hypothesis
se <- summary(fit)$coefficients[2, 2]

t_stat <- difference/se
t_stat
```

```
##      Bwt
## 1.135064
```

```
p_value <- 2*(1-pt(t_stat, fit$df.residual))
p_value
```

```
##      Bwt
## 0.258261
```

$$H_0 : \beta_1 = 4$$

Here, the test statistic = 1.135 and the p-value = .258. We would fail to reject the null hypothesis with an alpha level .05 that B_1 is equal to 4. To determine a linear relationship between the explanatory variable and the response variable, it is not very useful because any non zero number can be considered a non linear relationship.

g)

```
intercept <- fit$coefficients[1]
predicted <- intercept + B_1_estimated*2.25
predicted
```

```
## (Intercept)
##    8.719979
```

Based on the fitted regression model in part (d), we would predict a mean heart weight of 8.72 grams for a cat with a body weight of 2.25 kg.