02441 Applied Statistics and Statistical Software

Exercise 2B - Brain

The dataset brainweight contains measurements of the weight of both brain and body for different mammals

Variable name	Description
art	type of mammal
body	weigth of body (kg)
brain	weight of brain (gram)

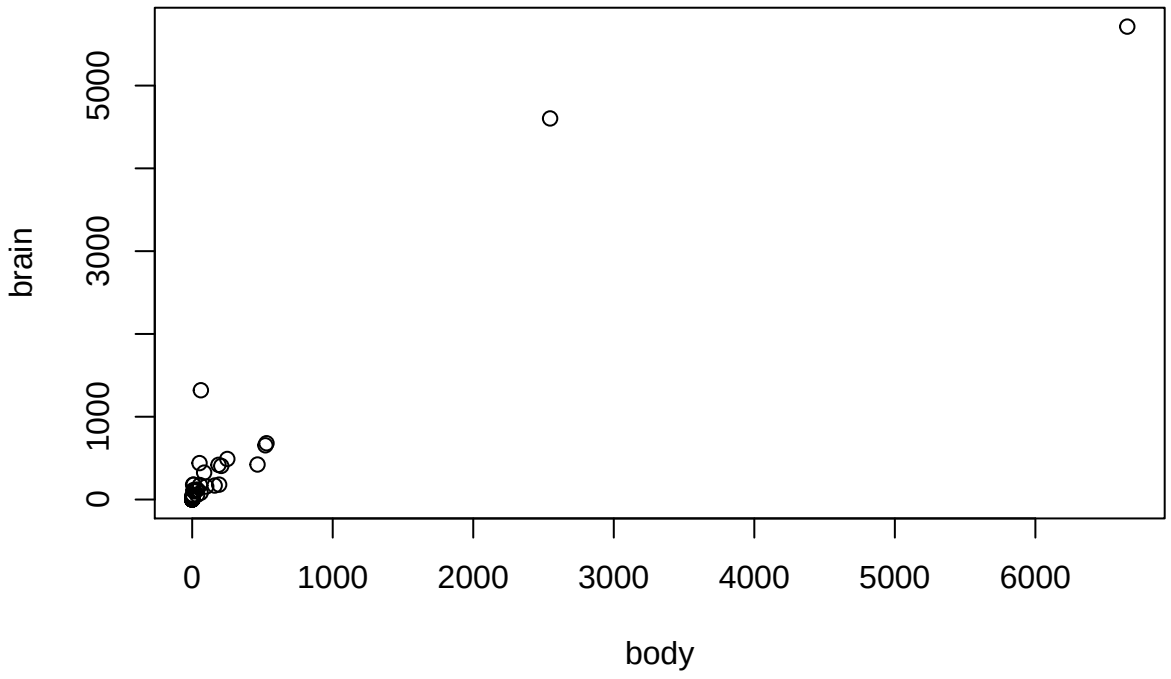
1. Make a scatterplot of body against brain. Do you see any correlation?

Start by loading the data

```
br <- read.table("brainweight.txt", header = TRUE)
```

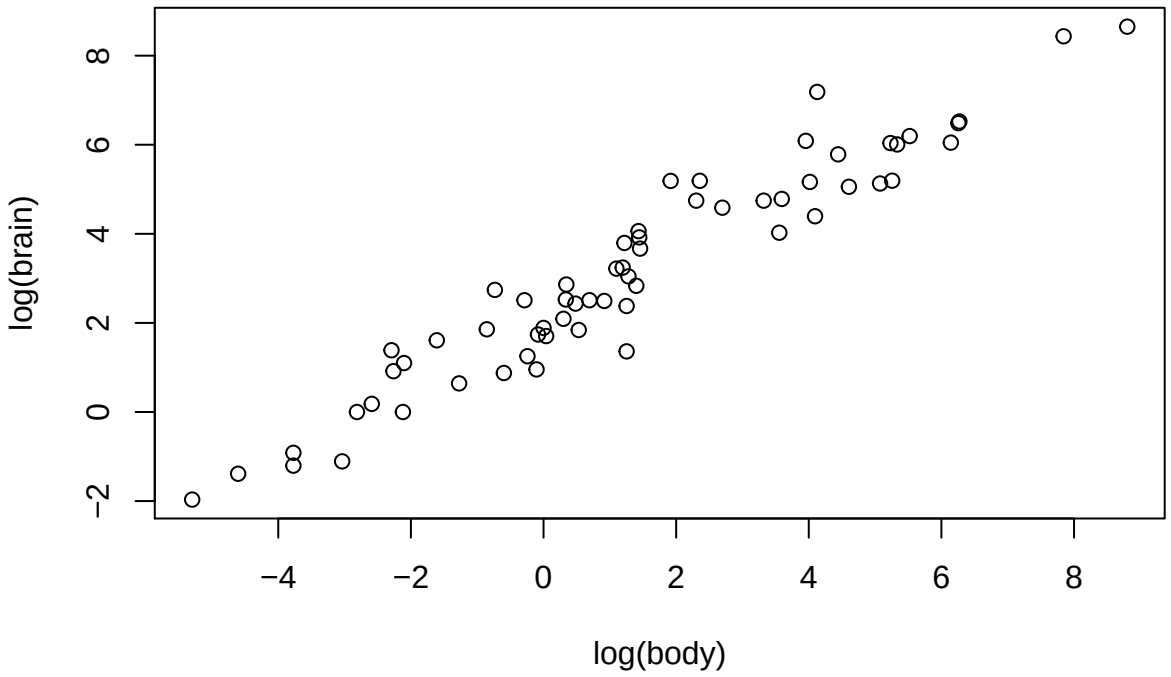
Plot body against brain

```
plot(brain~body, data = br)
```



2. Make a log transform of both body and brain. Make a scatterplot of the transformed variables. Compute the correlation.

```
plot(log(brain)~log(body), data = br)
```




```
cor(log(br$body), log(br$brain))
```

```
## [1] 0.9595748
```

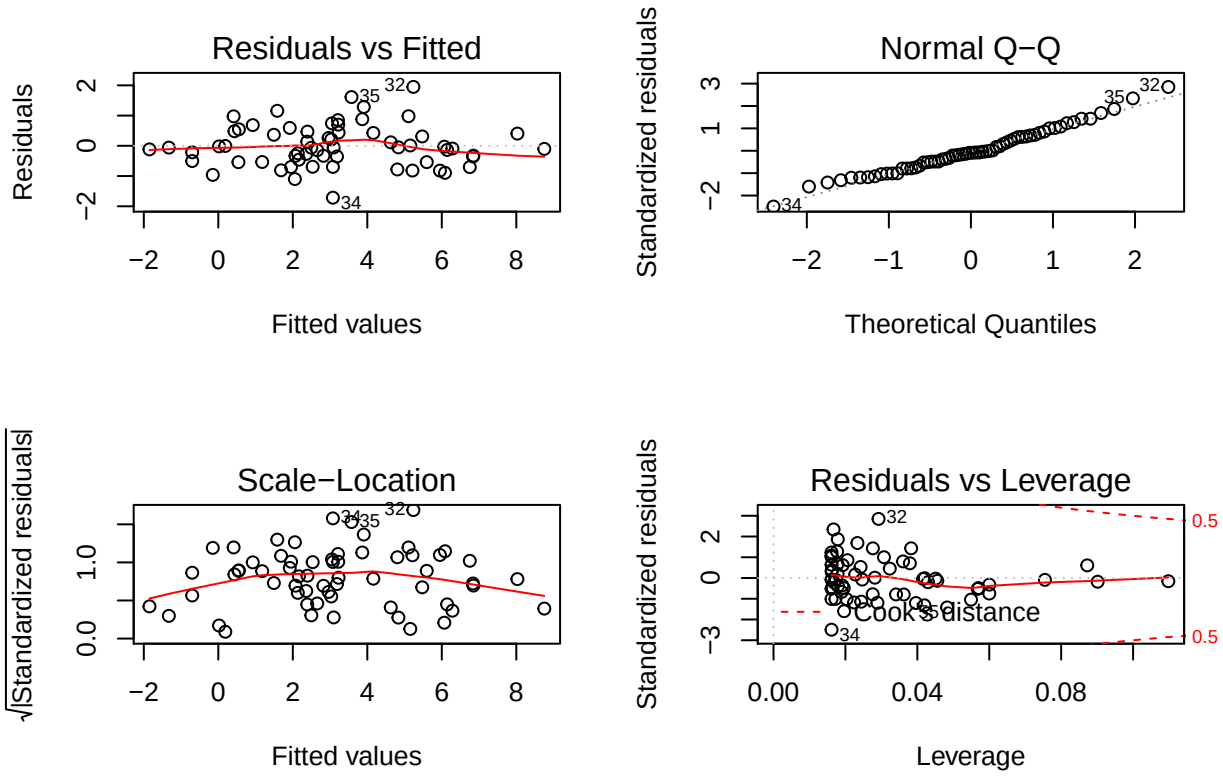
3. Fit a regression model between log(body) and log(brain)

```
lm1 <- lm(log(brain)~log(body), data = br)
summary(lm1)
```

```
##
## Call:
## lm(formula = log(brain) ~ log(body), data = br)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.71550 -0.49228 -0.06162  0.43597  1.94829
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.13479    0.09604   22.23   <2e-16 ***
## log(body)    0.75169    0.02846   26.41   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6943 on 60 degrees of freedom
## Multiple R-squared:  0.9208, Adjusted R-squared:  0.9195
## F-statistic: 697.4 on 1 and 60 DF,  p-value: < 2.2e-16
```

4. Is there any outlier in the data?

```
par(mfrow=c(2,2))
plot(lm1)
```




```
library(kableExtra)
kable(br[c(32,34,35),])
```

	art	body	brain
32	Human	62.0	1320.0
34	Water.opossum	3.5	3.9
35	Rhesus.monkey	6.8	179.0

The outliers do not surprise ;)

5. How would you evaluate the fit of the model?

Most of the variance in brain weight can be explained by body weight. R^2 is quite high and the model parameters are highly significant as it can be seen in the lm1 model summary.