

Cat Anatomy

Samuel Dummer

8/24/2021

Abstract

This document summarizes a report about the anatomy of cats. This includes the heart weight (grams), body weight (kilograms), and the sex

Model Setup

Here setup the model by cleaning up the data, setting up the directory, and loading extra functions

```
rm(list=ls())
setwd("C:/Users/isabe/Desktop/RFLoder")
source("myfunctions.R")
```

Loading Libraries

In this model, we used the MASS package (Modern Applied Statistics with S, from Venables and Ripley). The MASS package includes many data sets including the one we used, cats.

```
# load data sets from MASS package
library(MASS)
```

Data Demographics

In this section, we observed the structure of the data. This includes, the “head” (first few variables) of the data, the names of the columns, the structure of the data, the dimensions of the data (# of rows and columns), and the summary of the data.

```
head(cats)
```

```
##   Sex Bwt Hwt
## 1  F 2.0 7.0
## 2  F 2.0 7.4
## 3  F 2.0 9.5
## 4  F 2.1 7.2
## 5  F 2.1 7.3
## 6  F 2.1 7.6
```

```
names(cats)
```

```
## [1] "Sex" "Bwt" "Hwt"
```

```
str(cats)
```

```
## 'data.frame': 144 obs. of 3 variables:
## $ Sex: Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...
## $ Bwt: num 2 2 2 2.1 2.1 2.1 2.1 2.1 2.1 2.1 ...
## $ Hwt: num 7 7.4 9.5 7.2 7.3 7.6 8.1 8.2 8.3 8.5 ...
```

```
dim(cats)
```

```
## [1] 144 3
```

```
summary(cats)
```

```
## Sex      Bwt      Hwt
## F:47  Min.   :2.000  Min.   : 6.30
## M:97  1st Qu.:2.300  1st Qu.: 8.95
##      Median :2.700  Median :10.10
##      Mean   :2.724  Mean   :10.63
##      3rd Qu.:3.025  3rd Qu.:12.12
##      Max.   :3.900  Max.   :20.50
```

Good looking summary table

The knitr/kable package is very useful to display tables in a format that is easier to comprehend. The summary of the dataset is displayed in this table.

```
knitr::kable(
  summary(cats),
  caption = "Table 1: Summary table of cat anatomy data"
)
```

Table 1: Table 1: Summary table of cat anatomy data

Sex	Bwt	Hwt
F:47	Min. :2.000	Min. : 6.30
M:97	1st Qu.:2.300	1st Qu.: 8.95
NA	Median :2.700	Median :10.10
NA	Mean :2.724	Mean :10.63
NA	3rd Qu.:3.025	3rd Qu.:12.12
NA	Max. :3.900	Max. :20.50

Data Munging (Renaming Tables and Data)

Data munging is a process where one goes through the data and changes the names of the columns and data. This makes it easier for anyone to understand what the values mean. For example, this dataset originally used Sex, Bwt, and Hwt and we changed them to Gender, Body Weight, and Heart Weight to make them easier to read. Lastly, we changed F/M to Female/Male to, again, make it easier to comprehend the data.

```
colnames(cats) <- c("Gender", "BodyWeight", "HeartWeight")
names(cats)
```

```
## [1] "Gender"      "BodyWeight"  "HeartWeight"
```

```
cats$Gender <- factor(cats$Gender, levels = c("F", "M"), labels = c("Female", "Male"))
head(cats)
```

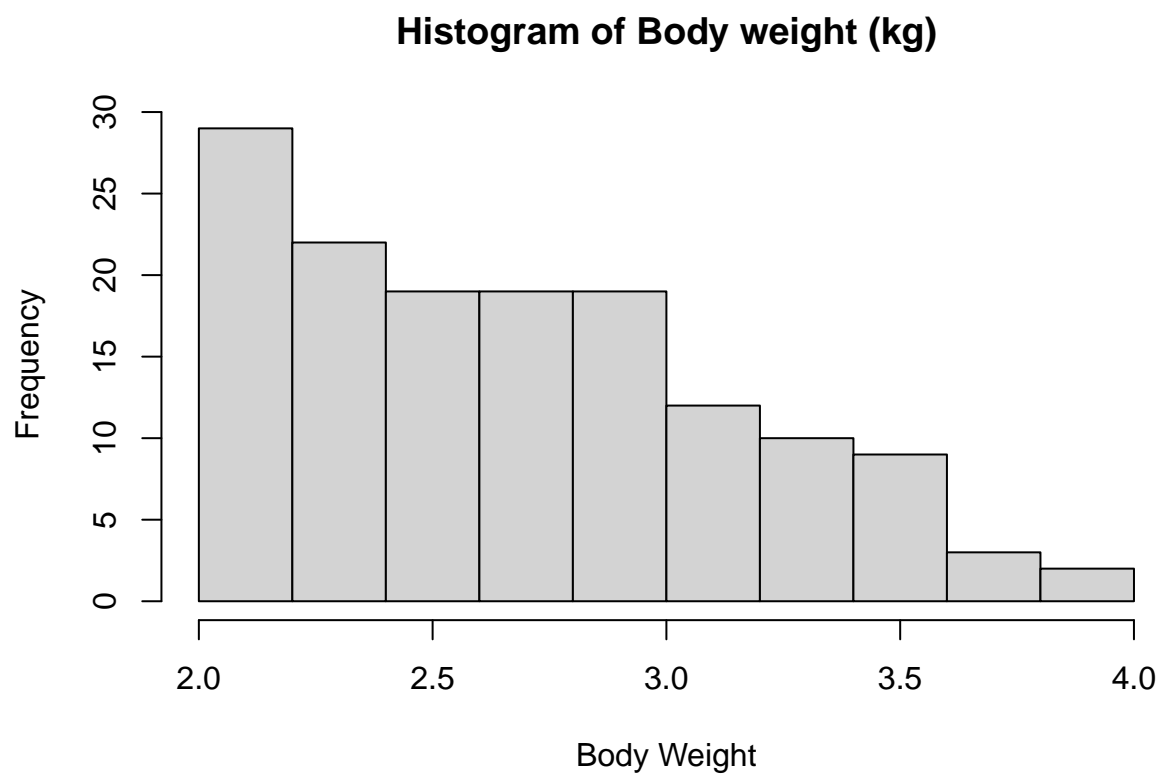
```
##   Gender BodyWeight HeartWeight
## 1 Female         2.0          7.0
## 2 Female         2.0          7.4
## 3 Female         2.0          9.5
## 4 Female         2.1          7.2
## 5 Female         2.1          7.3
## 6 Female         2.1          7.6
```

Graphical Representations

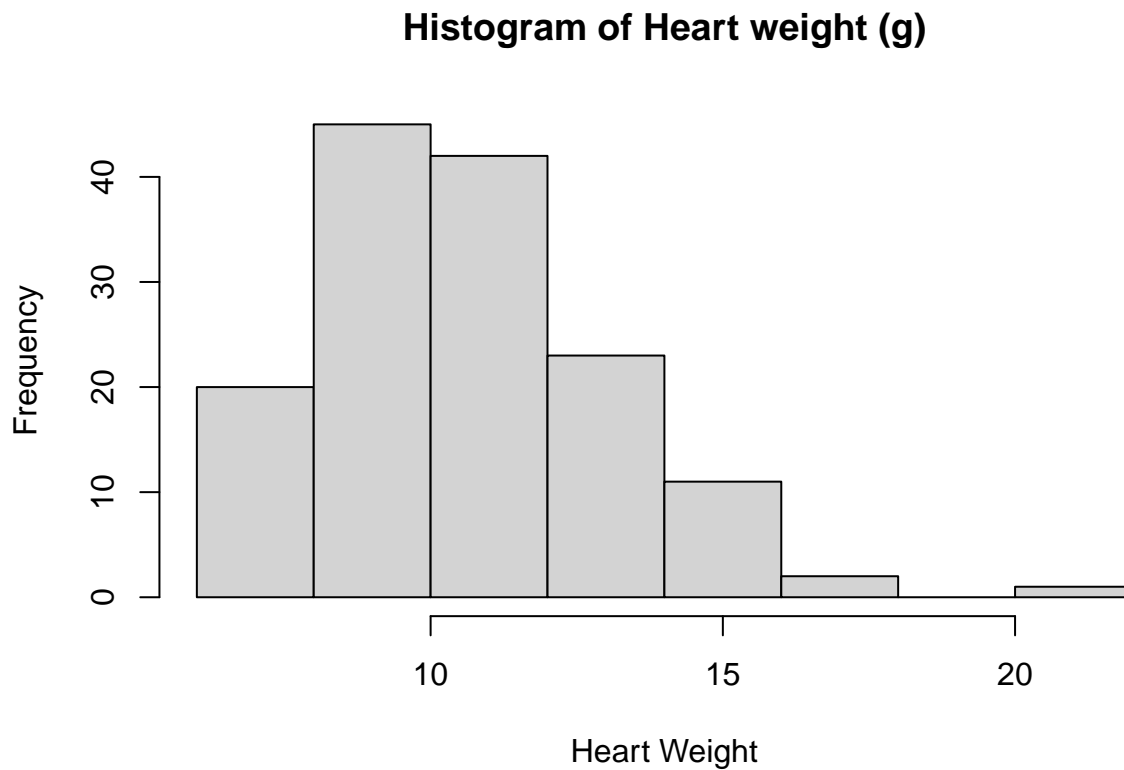
When we use Exploratory Data Analysis(EDA), making plots is incredibly helpful to define what the general idea of the data is and trends/behaviors we can observe. We made histograms, pairwise correlation plots, and a scatter plot with a regression line.

Histogram Histograms are very helpful when you are looking at data. They help to tell you if the data set follows normal distribution or not.

```
hist(cats$BodyWeight, main="Histogram of Body weight (kg)", xlab = "Body Weight")
```

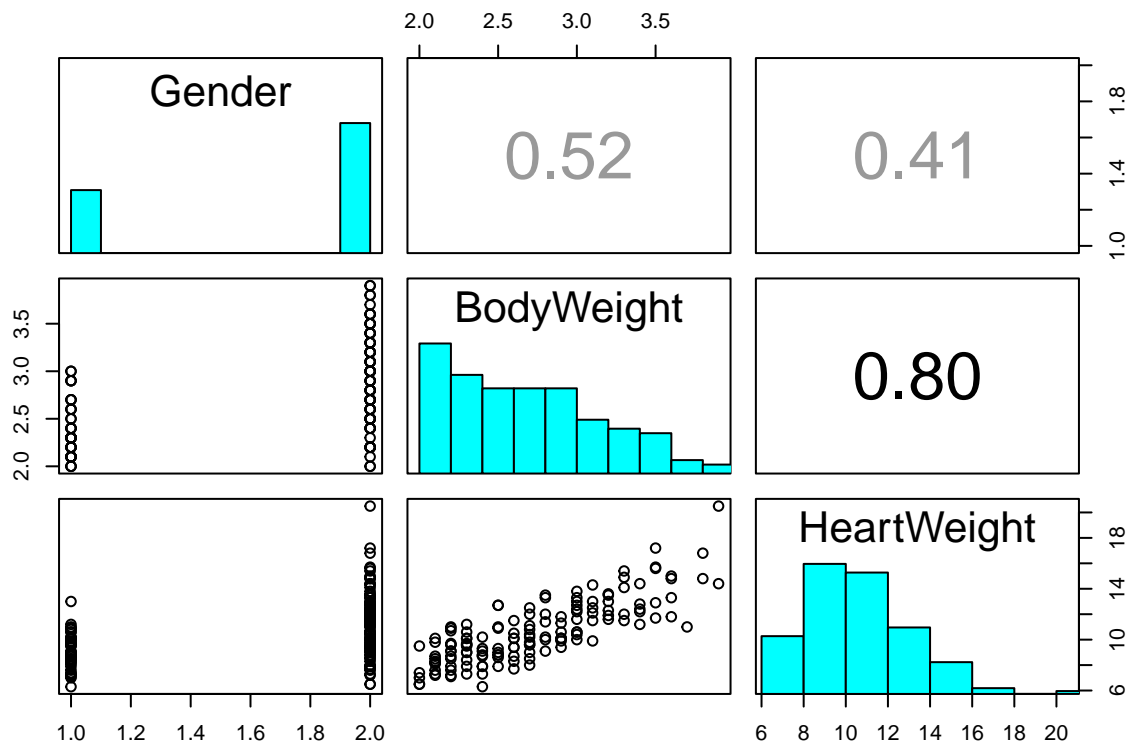


```
hist(cats$HeartWeight, main="Histogram of Heart weight (g)", xlab = "Heart Weight")
```



Pairwise plot We used one of the built-in functions that we had “myfuntions.R” to help us create a pairwise plot. This plot helps to show histograms of the different data, scatter plots, and the correlation coefficient between each pair of data.

```
pairs(cats,upper.panel=panel.cor,diag.panel=panel.hist)
```



Scatter Plot: Body Weight vs. Heart Weight (Includes linear regression) The last plot that we made was a simple scatter plot, but we also added a linear regression. In the plot the data is actually fairly linear which you can observe because many of the data points match up with the linear regression

```
plot(cats$HeartWeight~cats$BodyWeight, main="Simple plot of heart weight \n as a function of body weight")
fit<- lm(cats$HeartWeight~cats$BodyWeight)
abline(fit)
```

**Simple plot of heart weight
as a function of body weight**

