```
#Copyright (c) 2018 Jouke Profijt.
#Licensed under GPLv3. See gpl.md
```

```
BirdBones <- read.csv("../data/bird.csv",header = T, sep = ",")
```

Introduction

Research Question

What bone or group of bones that most birds have in common, is the most significant for the function in the diffrent ecological groups?

Data

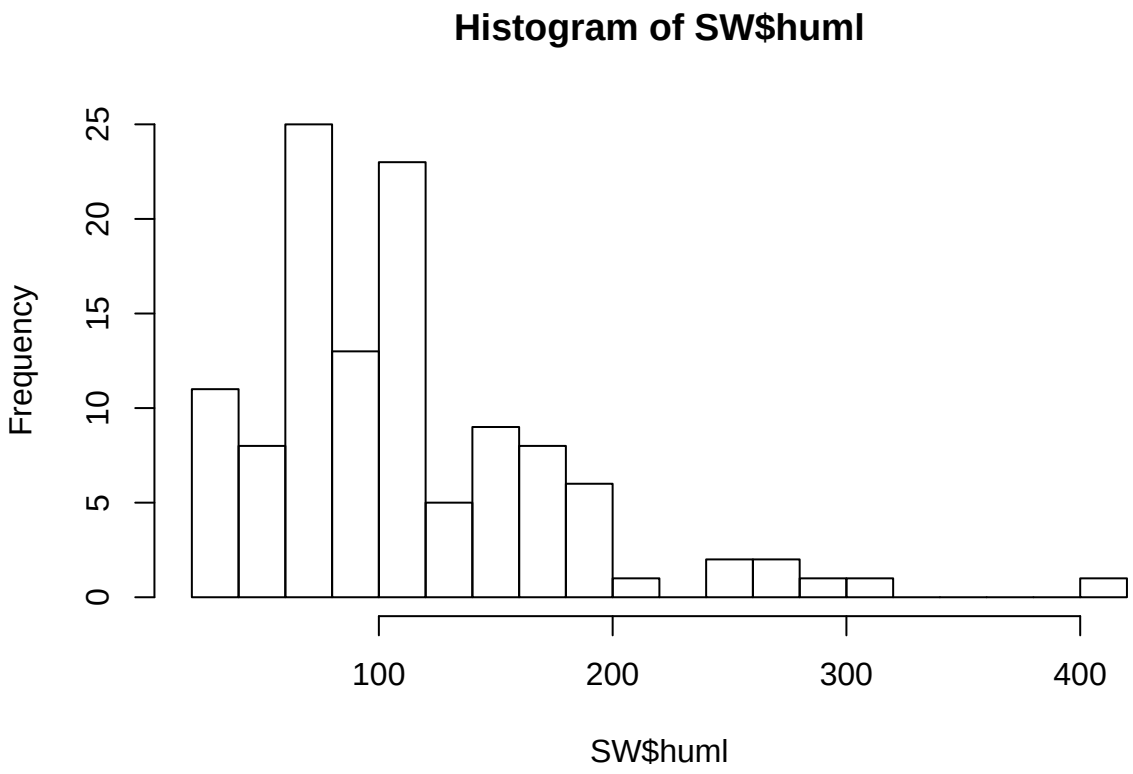
```
Data recieved from  [Birds' Bones and Living Habits](https://www.kaggle.com/zhangjuefei/birds-bones-and-

Bone measurements were measured from a skeleton collection of Natural History Museum of Los Angeles Cou
provided by Dr. D. Liu of beijing Museaum of Natural History
```

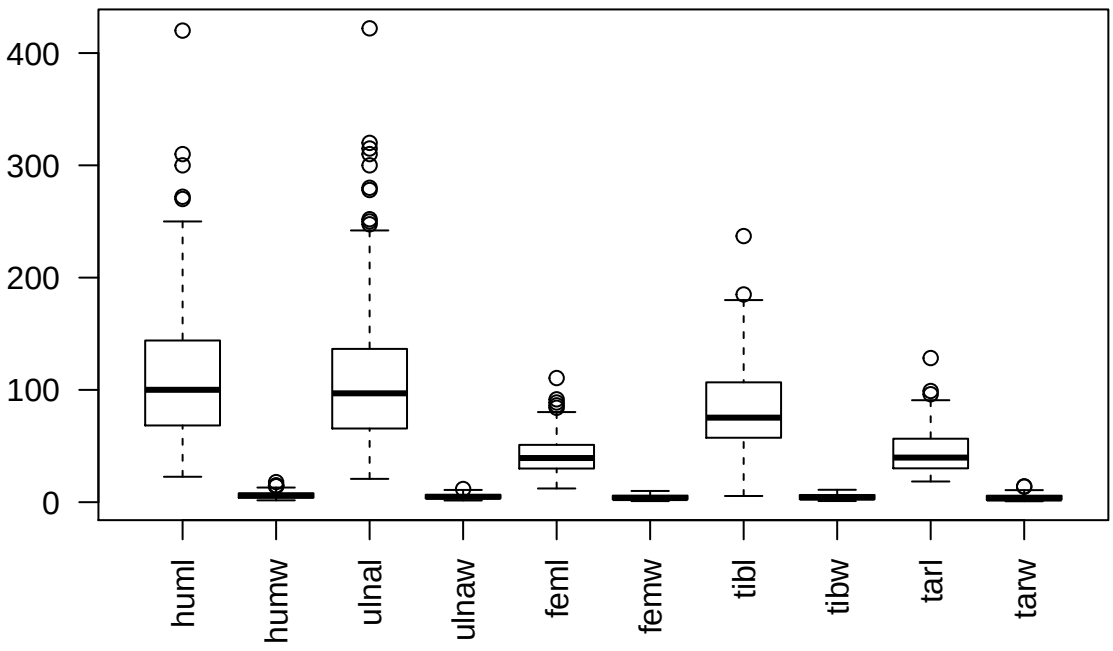
Exploratory Data Analyses

```
# Swimming
SW <- BirdBones[BirdBones$type == "SW", 1:11]
# Wading
W <- BirdBones[BirdBones$type == "W", 1:11]
# Terrestial
TER <- BirdBones[BirdBones$type == "T", 1:11]
# Raptors
R <- BirdBones[BirdBones$type == "R", 1:11]
# Scansorial
P <- BirdBones[BirdBones$type == "P",1:11]
# Singing
SO <- BirdBones[BirdBones$type == "SO",1:11]
```

```
source("../scripts/BoneMeans.R")
dataMeans <- BoneMeans(BirdBones)
hist(SW$huml, breaks = 15)
```



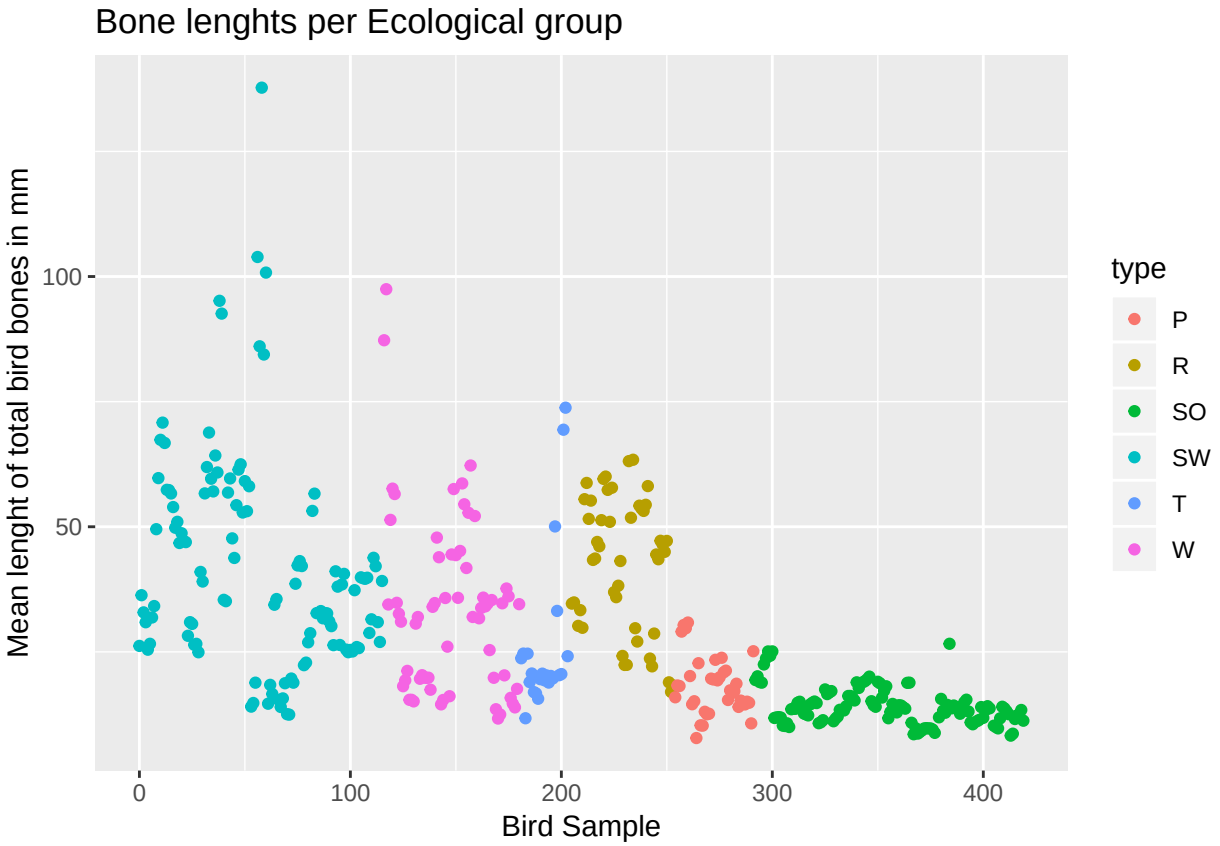

```
boxplot(SW[2:11], las = 2)
```



I'm creating a dotplot which displays the bonelengths on x axis an the Id on y colorcoded by their ecological group. by evaluating this we can see if some groups have overall larger or smaller bones.

```
library(ggplot2)
library(reshape)
source("../scripts/BoneMeans.R")
BirdBones <- BoneMeans(BirdBones)
ggplot(data = BirdBones, aes(id, means, colour = type)) +
  ggtitle("Bone lenghts per Ecological group")+
  ylab("Mean lenght of total bird bones in mm") +
  xlab("Bird Sample")+
  geom_point()
```

```
## Warning: Removed 7 rows containing missing values (geom_point).
```



As seen above swimming birds have the biggest bones, but also shown is that there are a lot more samples in that group where there is a lot of variation. I can look into cleaning up the data and removing the biggest outliers in this group. Singing birds also have a lot of samples but there is much less variation and so more certanty.

For the rest of the birds there are not a lot of sample so maby we could try and normalizing the data so there is an even amount of samples per group.