

HW week 1

jhonny

6/16/2020

R basics

```
x <- c(2.23, 3.45, 1.87, 2.11, 7.33, 18.34, 19.23)
mean(x)
```

```
## [1] 7.794286
```

```
y <- 0

for (a in 1:25) {
  y[a] <- a^2
}
sum(y)
```

```
## [1] 5525
```

R data manipulation

Reading cvs file and saving to variable.

```
miceData <- read.csv("femaleMiceWeights.csv")
miceData
```

```
##      Diet Bodyweight
## 1  chow      21.51
## 2  chow      28.14
## 3  chow      24.04
## 4  chow      23.45
## 5  chow      23.68
## 6  chow      19.79
## 7  chow      28.40
## 8  chow      20.98
## 9  chow      22.51
## 10 chow      20.10
## 11 chow      26.91
## 12 chow      26.25
## 13  hf      25.71
## 14  hf      26.37
## 15  hf      22.80
## 16  hf      25.34
## 17  hf      24.97
## 18  hf      28.14
## 19  hf      29.58
## 20  hf      30.92
## 21  hf      34.02
## 22  hf      21.90
## 23  hf      31.53
## 24  hf      20.73
```

Get value of row 12 column 2.

```
miceData[12,2]
```

```
## [1] 26.25
```

```
names(miceData)
```

```
## [1] "Diet"      "Bodyweight"
```

Get mice weight and the amount of mice.

```
miceData$Bodyweight[11]
```

```
## [1] 26.91
```

```
length(miceData$Bodyweight)
```

```
## [1] 24
```

Search for "hf" and get the mean.

```
grep("hf",miceData$Diet)
```

```
## [1] 13 14 15 16 17 18 19 20 21 22 23 24
```

```
mean(miceData$Bodyweight[13:24])
```

```
## [1] 26.83417
```

Get a random number in the range 13–14 and print the weight of the mouse on that row.

```
intdump <- 13:24
set.seed(1)
sample(intdump, 1)
```

```
## [1] 21
```

```
miceData$Bodyweight[21]
```

```
## [1] 34.02
```

dplyr intro

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
mamSleep <- read.csv("msleep_ggplot2.csv")  
class(mamSleep)
```

```
## [1] "data.frame"
```

```
nrow(mamSleep)
```

```
## [1] 83
```

Select primates

```
primates <- filter(mamSleep, order=="Primates")  
nrow(primates)
```

```
## [1] 12
```

```
class(primates)
```

```
## [1] "data.frame"
```

Sleep total and mean

```
primates_sleep_total <- filter(mamSleep, order=="Primates") %>% select(sleep_  
total)  
class(primates_sleep_total)
```

```
## [1] "data.frame"
```

```
mean(unlist(primates_sleep_total))
```

```
## [1] 10.5
```

Use summarize to calculate mean.

```
primates
```

```
##           name           genus vore   order conservation sleep_total
## 1      Owl monkey          Aotus omni Primates          <NA>      17.0
## 2      Grivet Cercopithecus omni Primates           lc       10.0
## 3      Patas monkey Erythrocebus omni Primates           lc       10.9
## 4      Galago          Galago omni Primates          <NA>       9.8
## 5      Human          Homo omni Primates          <NA>       8.0
## 6  Mongoose lemur          Lemur herbi Primates          vu       9.5
## 7      Macaque          Macaca omni Primates          <NA>      10.1
## 8      Slow loris      Nyctibeus carni Primates          <NA>      11.0
## 9      Chimpanzee          Pan omni Primates          <NA>       9.7
## 10     Baboon          Papio omni Primates          <NA>       9.4
## 11     Potto  Perodicticus omni Primates           lc      11.0
## 12 Squirrel monkey      Saimiri omni Primates          <NA>       9.6
##   sleep_rem sleep_cycle awake brainwt bodywt
## 1      1.8           NA   7.0  0.0155  0.480
## 2      0.7           NA  14.0      NA  4.750
## 3      1.1           NA  13.1  0.1150 10.000
## 4      1.1  0.5500000  14.2  0.0050  0.200
## 5      1.9  1.5000000  16.0  1.3200 62.000
## 6      0.9           NA  14.5      NA  1.670
## 7      1.2  0.7500000  13.9  0.1790  6.800
## 8      NA           NA  13.0  0.0125  1.400
## 9      1.4  1.4166667  14.3  0.4400 52.200
## 10     1.0  0.6666667  14.6  0.1800 25.235
## 11     NA           NA  13.0      NA  1.100
## 12     1.4           NA  14.4  0.0200  0.743
```

```
primates_sleep_total <- filter(mamSleep, order=="Primates") %>% summarize(mea
n = mean(sleep_total), n = n())
primates_sleep_total
```

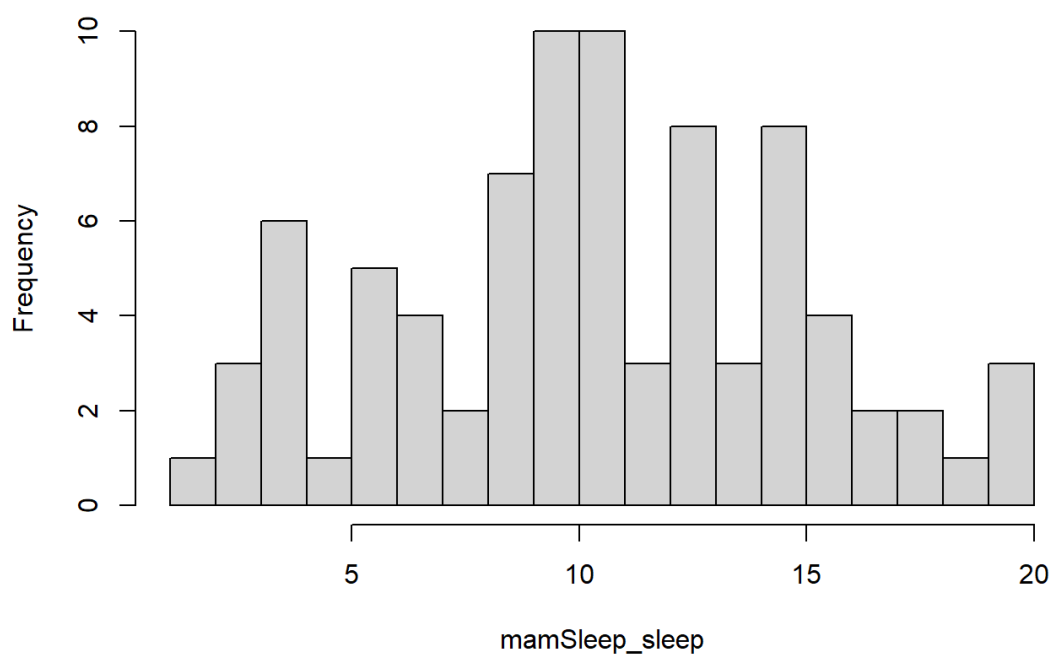
```
##   mean  n
## 1 10.5 12
```

Data exploration

Histograms

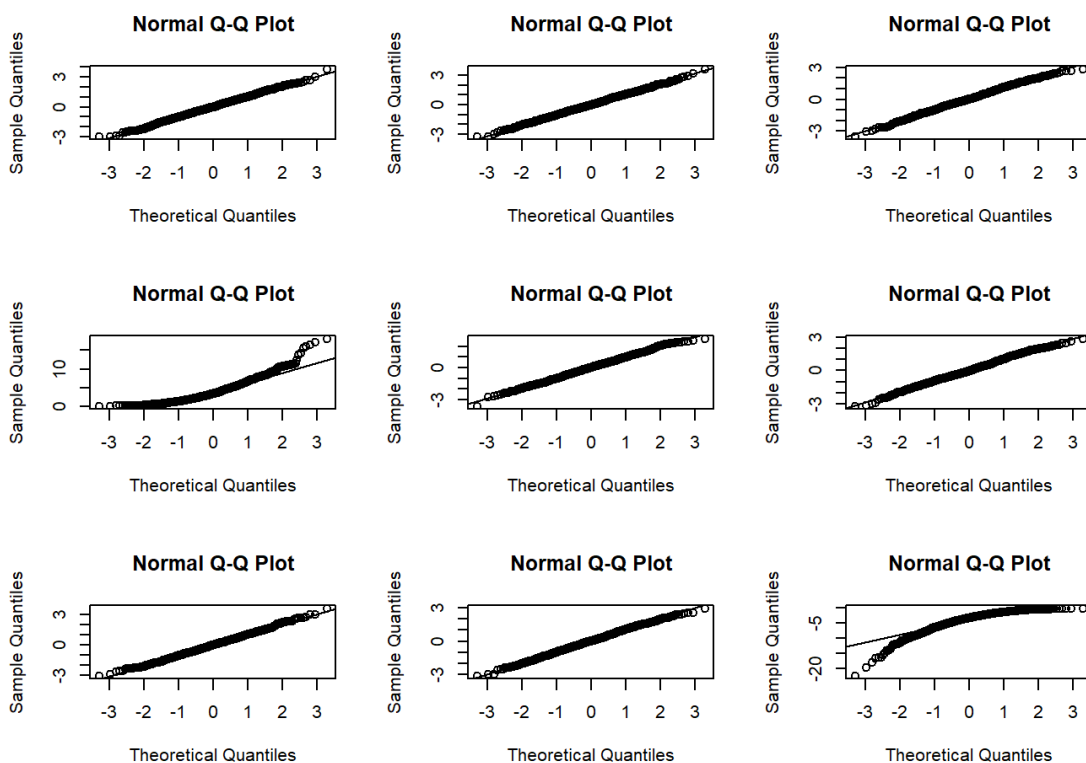
```
mamSleep_sleep <- mamSleep$sleep_total
hist(mamSleep_sleep,breaks=seq(floor(min(mamSleep_sleep)),ceiling(max(mamSlee
p_sleep))))
```

Histogram of mamSleep_sleep

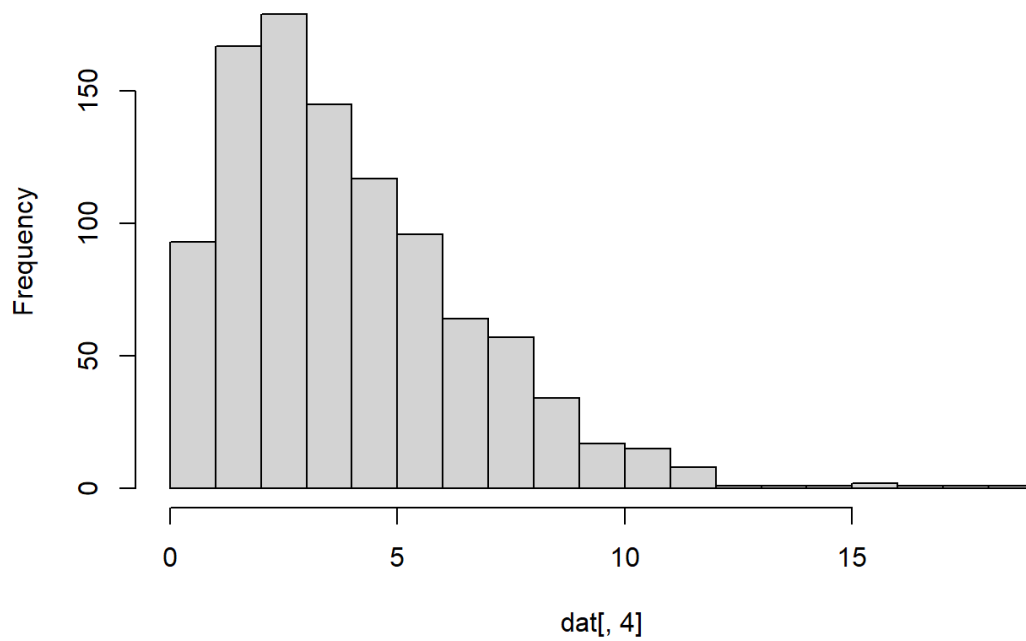


```
##qq plot
```

```
load("./data.RData")
par(mfrow = c(3,3))
for (i in 1:9) {
  qqnorm(dat[,i])
  qqline(dat[,i])
}
```

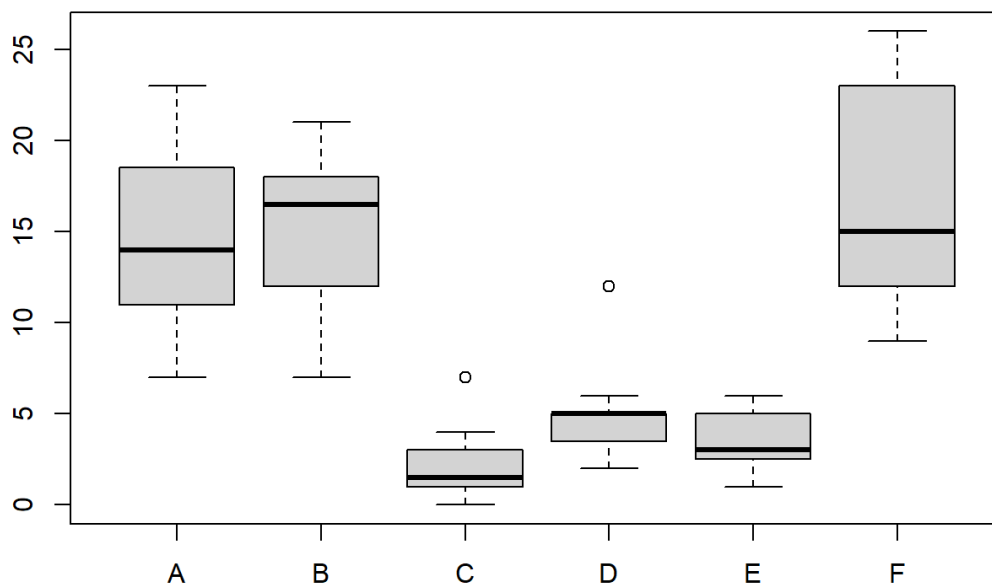


```
hist(dat[,4],breaks=seq(floor(min(dat[,4])),ceiling(max(dat[,4]))))
```

Histogram of dat[, 4]

```
##boxplot
```

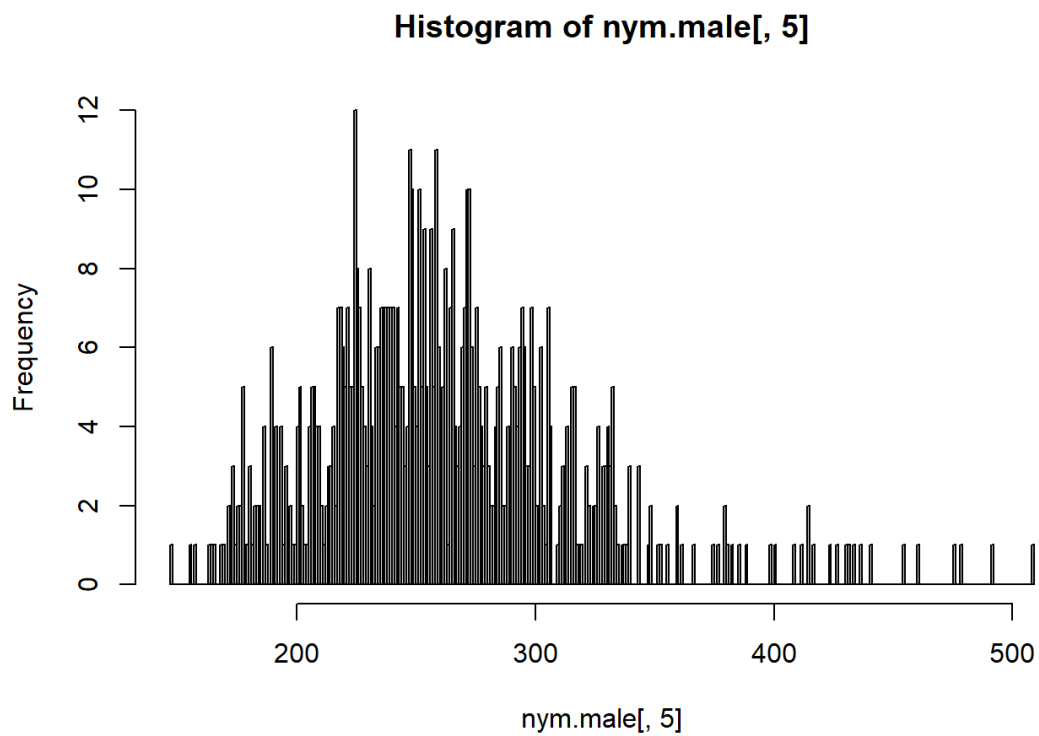
```
boxplot(split(InsectSprays[,1],InsectSprays[,2]))
```



```
data(nym.2002, package="UsingR")
head(nym.2002)
```

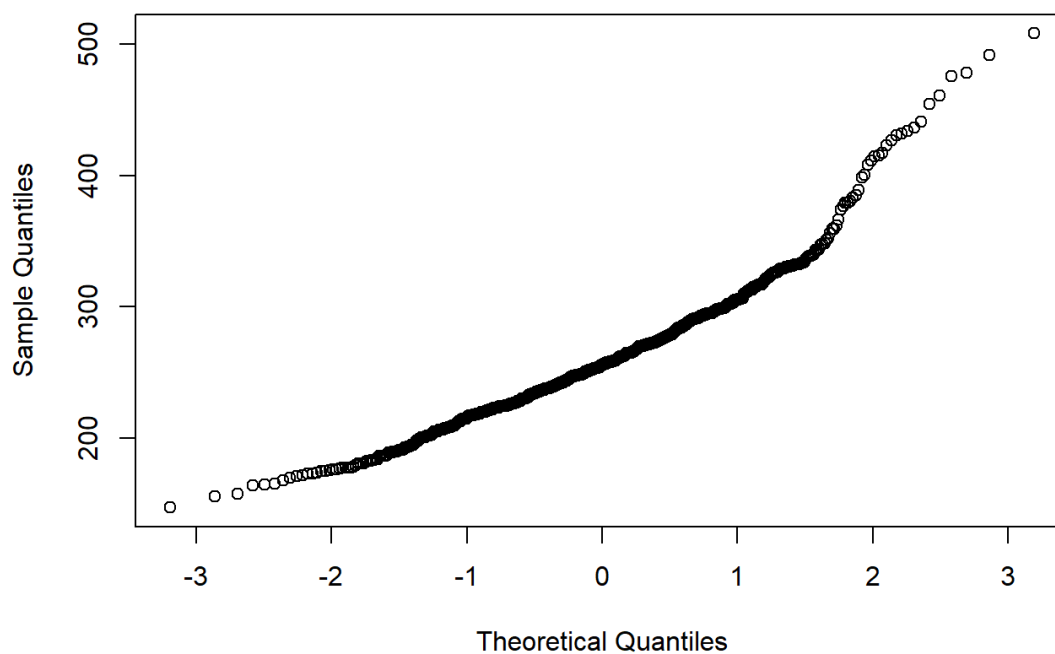
```
##      place gender age home      time
## 3475   3592   Male  52  GBR 217.4833
## 13594 13853 Female  40   NY 272.5500
## 12012 12256   Male  31  FRA 265.2833
## 10236 10457 Female  33   MI 256.1500
##  9476  9686   Male  33   NY 252.2500
##  1720  1784   Male  40   NJ 201.9667
```

```
nym.male <- filter(nym.2002, gender=="Male")
hist(nym.male[,5],breaks=seq(floor(min(nym.male[,5])),ceiling(max(nym.male[,5]
))))
```

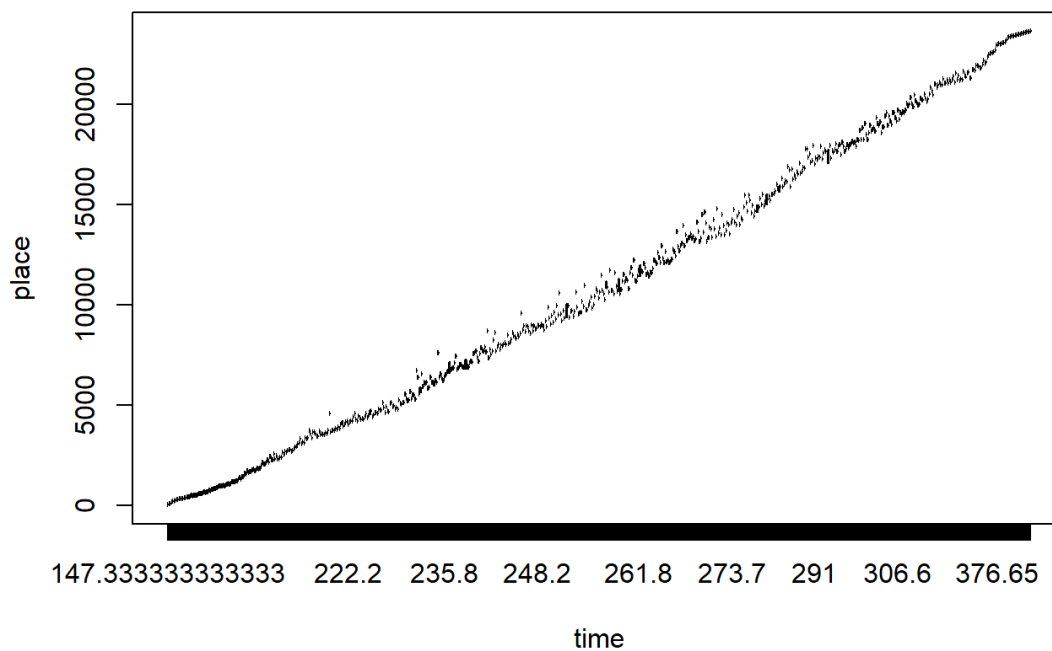


```
qqnorm(nym.male[,5])
```

Normal Q-Q Plot



```
boxplot(place~time,data=nym.male)
```



```
nym.female <- filter(nym.2002, gender=="Female")
hist(nym.female[,5],breaks=seq(floor(min(nym.female[,5])),ceiling(max(nym.female[,5]))))
```