

## Chapter 3 – R and Rstudio: a general introduction with statistics.

### Quiz block 1:

```
setwd("C:/Users/x991625/OneDrive - NIAB/a). NIAB WORK/QMBP 2020") # will be different based on users Laptop
list.files()
## [1] "TG data for day 1.csv"
```

### Quiz block 2:

```
# 1.
my_second_result <- (9.75 * 80) - 34
my_second_result / 2

## [1] 373

# 2.

Combine <- "harvester"
Combine

## [1] "harvester"
```

### Quiz block 3:

```
#1.
#save excel sheet as txt file within excel and read in:
data_txt<-read.table("TG data for day 1.txt",header=T)
#2.
names(data_txt)

## [1] "variety"      "year"          "ORIGIN"        "Rht2"          "PpdD1"
## [6] "yield"        "CALLOW_2011"   "FRANCE_2010"   "FRANCE_2011"   "LGE_2010"
## [11] "LGE_2011"     "NIAB_2011"     "FT"            "HT"            "AWNS"

dim(data_txt)

## [1] 376 15
```

### Quiz block 4:

```
#1.
days<-c("Mon", "Tue", "Wed", "Thu", "Fri", "Sat", "Sun")
nums<- round(runif(7,1,10))
logical<-c(T,T,T,T,T,F,F)

# 2.
ex4<-data.frame(days,nums,logical)
names(ex4)<-c("day_of_week", "coffee_consumed", "attendance")

#3.

ex4$coffee_consumed<-as.factor(ex4$coffee_consumed)
class(ex4$coffee_consumed)
```

```
## [1] "factor"
```

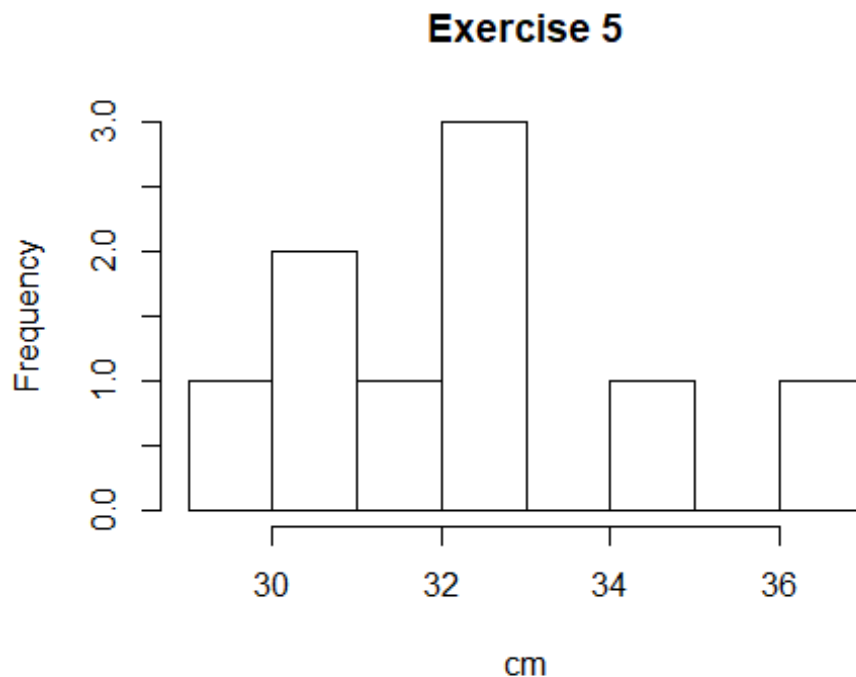
#### Quiz block 5:

#1.

```
leaf_sizes<-c(30.1, 32.3, 36.8, 34.1, 32.0, NA, 32.4, 30.1, 29.9, 32.6)
```

#2.

```
hist(leaf_sizes, main="Exercise 5", xlab = "cm")
```



#3.

```
mean(leaf_sizes, na.rm = T)
```

```
## [1] 32.25556
```

#4.

```
leaf_sizes[!is.na(leaf_sizes)]
```

```
## [1] 30.1 32.3 36.8 34.1 32.0 32.4 30.1 29.9 32.6
```

#### Quiz block 6

#1.

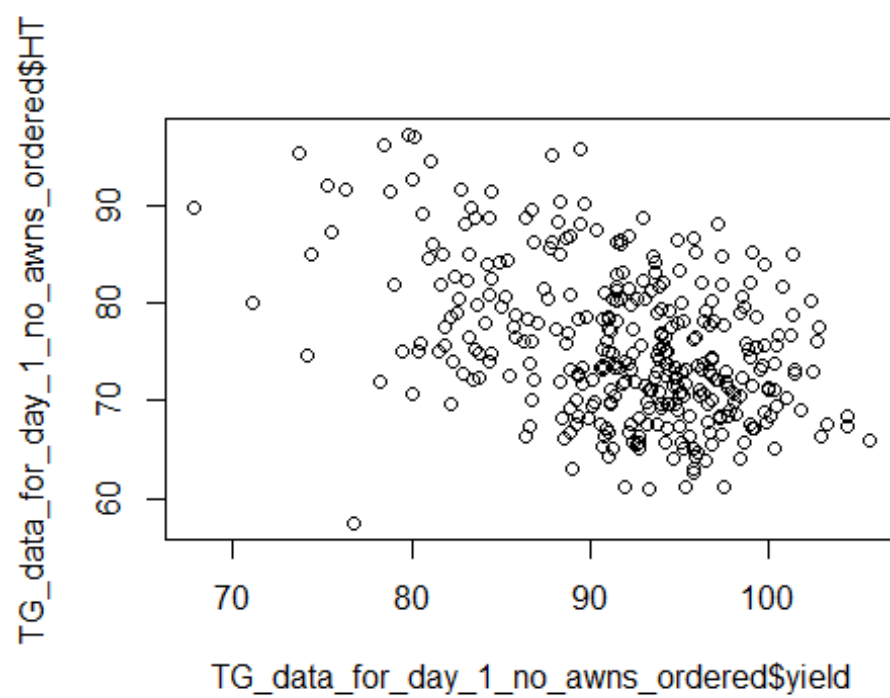
```
TG_data_for_day_1_no_awns<-subset(data_txt, AWNS==0)
```

#2.

```
TG_data_for_day_1_no_awns_ordered<-TG_data_for_day_1_no_awns[order(TG_data_for_day_1_no_awns$ORIGIN),]
```

#3.

```
plot(TG_data_for_day_1_no_awns_ordered$yield, TG_data_for_day_1_no_awns_ordered$HT)
```



```
#4.  
cor(TG_data_for_day_1_noawns_ordered$yield, TG_data_for_day_1_noawns_ordered$HT, use="pairwise.complete.obs")  
## [1] -0.4275771
```