Exam 1

Henrique

March 11, 2019

Question 1.a)

```
getwd()

## [1] "C:/Users/henri/Documents/R158"

Stat_class_data = read.csv("Stat_class_data.csv", header = TRUE)
```

The columms are sex with m and f as paramaters,color with a string with color name, Height with a integer number and number with a a numerical number assigned.

Question 1.b)

```
attach(Stat_class_data)

sum(Sex=="M")

## [1] 60

sum(Sex=="F")

## [1] 73
```

There are 60 males and 73 females.

Questin 1.c)

```
attach(Stat_class_data)

## The following objects are masked from Stat_class_data (pos = 3):
##
##     Color, Height, Number, Sex

table(Color)

## Color
##  Black   Blue  Brown  Green Orange Purple    Red   Teal  White Yellow
##      2     52      3     27      8     19      8      7      1      6

detach(Stat_class_data)
```

Top colours: 1-Blue 2-Green 3-Purple

Question 1.d)

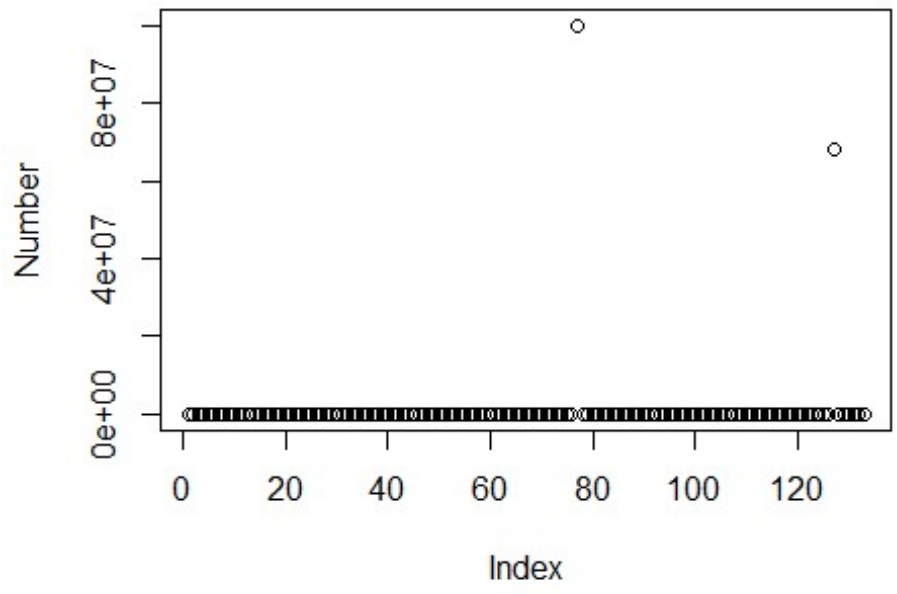
```
attach(Stat_class_data)
```

```
## The following objects are masked from Stat_class_data (pos = 3):
##
##     Color, Height, Number, Sex

summary(Height)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   60.00   66.00   68.00   68.35   71.00   79.00

detach(Stat_class_data)
```

Median height = 68.00

Question 1.e)

```
attach(Stat_class_data)

## The following objects are masked from Stat_class_data (pos = 3):
##
##     Color, Height, Number, Sex

plot(Number)
```




```
detach(Stat_class_data)
```

numbers is quite randomized.

QUestion 1.f)

```
attach(Stat_class_data)

## The following objects are masked from Stat_class_data (pos = 3):
##
##     Color, Height, Number, Sex

sum(Number>5000)

## [1] 7

detach(Stat_class_data)
```

there are 7 numbers above 5000, which is about 5% of the students.

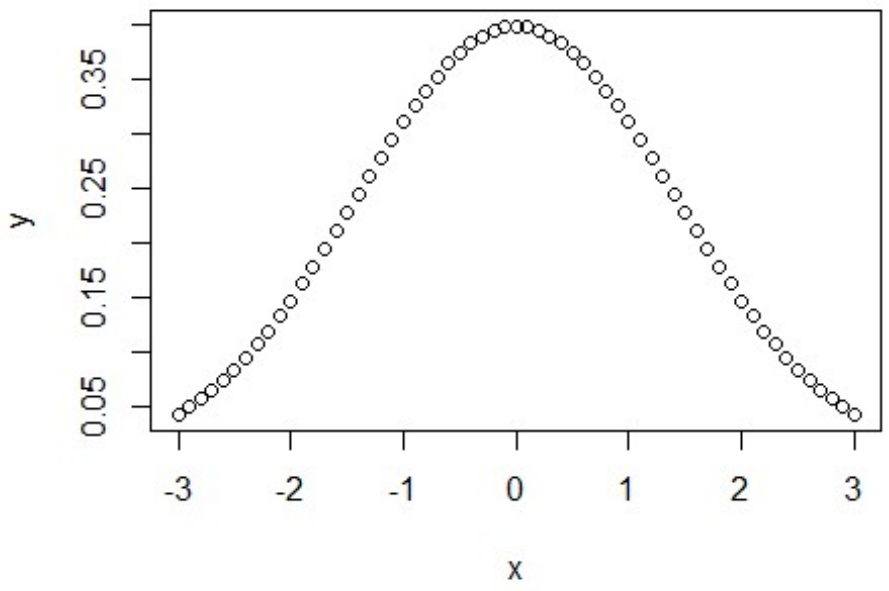
question 1.g)

```
NEX <- Number[c(-77,-127)]

mean(NEX)

## [1] 841.4564

mean(Number)

## [1] 1263612
```

Mean number without 77 and 127: 841.4564 Mean of number: 1,263,612

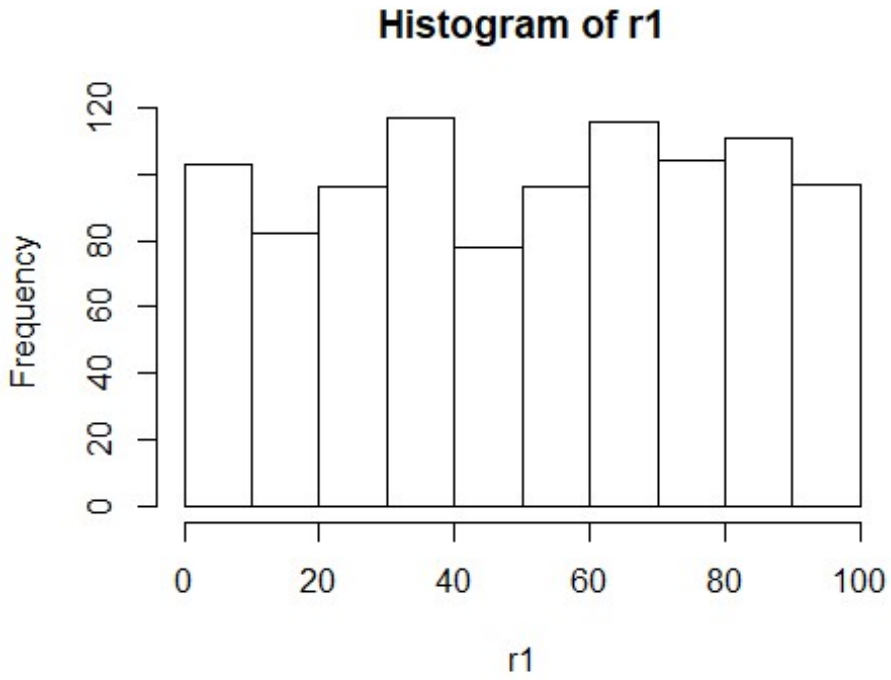
Question 2

```
x <- seq(-3,3,0.1)

b1=1/sqrt(2*pi)

b2=-((x/2)^2)

y = b1*exp(b2)


plot(x,y)
```



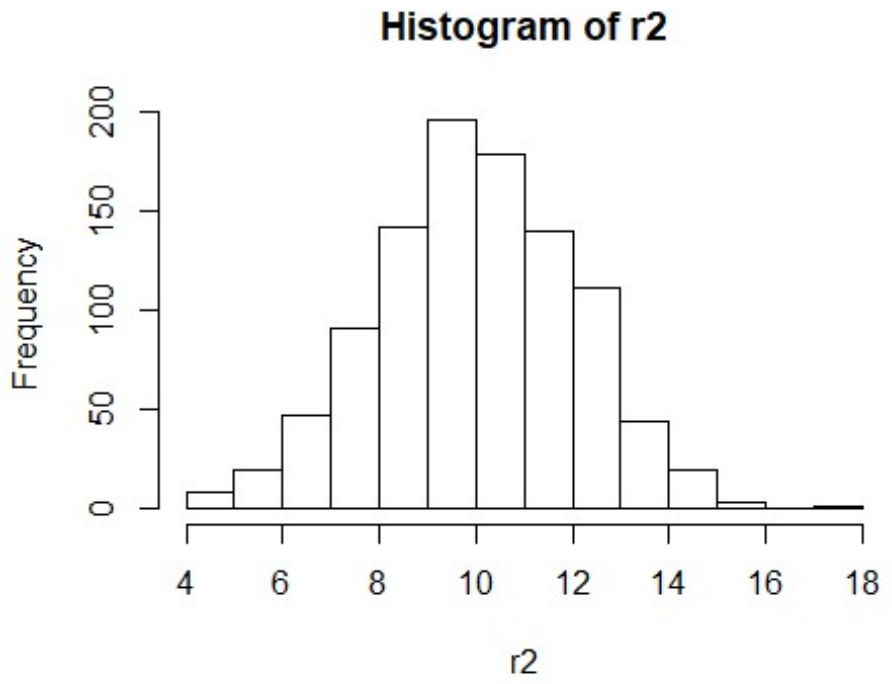
Question 3.a)

```
r1 <- sample(1:100,1000, replace = TRUE, pro = NULL)

hist(r1)
```



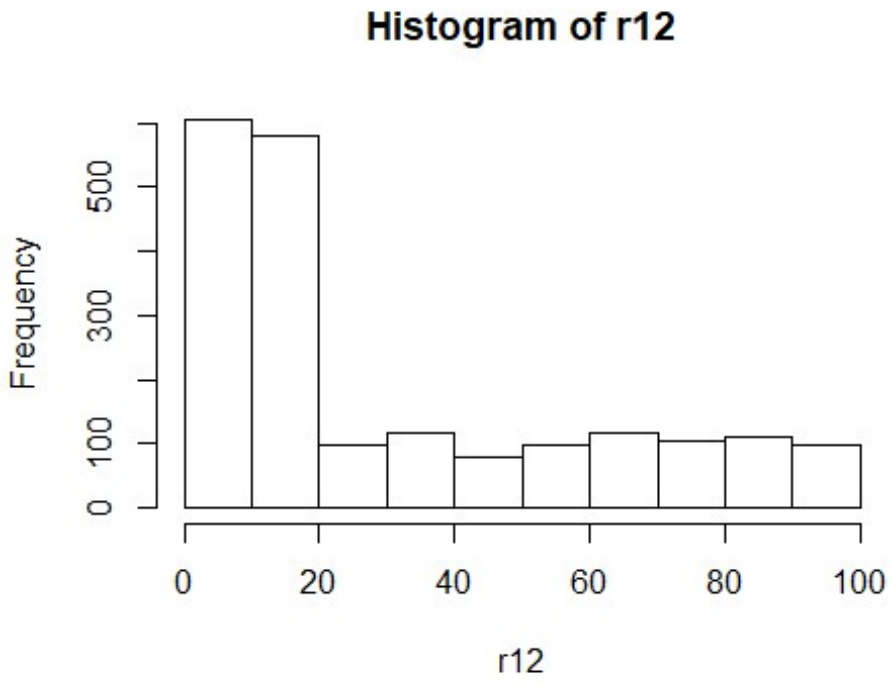
Question 3.b)

```
r2 <- rnorm(1000,10,2)

hist(r2)
```



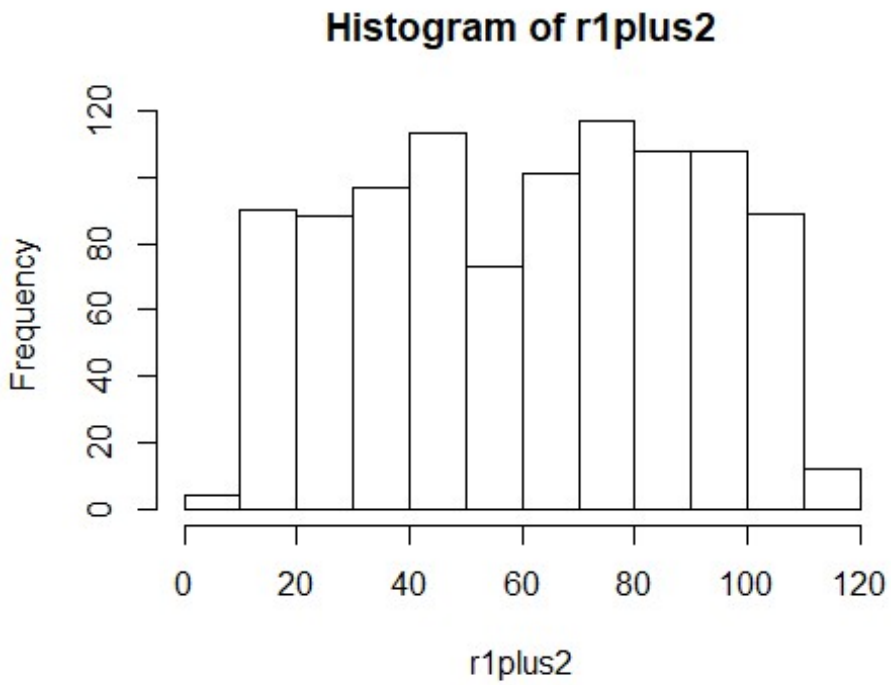
Question 3.c)

```
r12 <- c(r1,r2)

hist(r12)
```



Question 3.d)

```
r1plus2 <- c(r2+r1)

hist(r1plus2)
```



question 3.e)

they are different because when when the values are combined they skew the normal distribution to the left, and when they are added they just randomize, so there is no patterm.