

Introduction to R

Chatchawan Panraksa

October 2, 2018

What is R?

- ① Statistical Programming Language
- ② Free and open source

Download R

- 1 Download R <https://www.r-project.org/>
- 2 Download RStudio
<https://www.rstudio.com/products/rstudio/download/>

Note that using R does not require RStudio.

You need R before using RStudio.

However, RStudio has better interface.

Simple manipulations; numbers and vectors

- Vectors and assignment

```
1+1
```

```
## [1] 2
```

Simple manipulations; numbers and vectors

- Vectors and assignment

```
1+1
```

```
## [1] 2
```

```
x<-c(10.4, 5.6, 3.1, 6.4, 21.7)  
x[3]
```

```
## [1] 3.1
```

```
assign("x", c(10.4, 5.6, 3.1, 6.4, 21.7))  
x
```

```
## [1] 10.4 5.6 3.1 6.4 21.7
```

```
1/x
```

```
## [1] 0.09615385 0.17857143 0.32258065 0.15625000 0.04608295
```

```
x^2
```

```
## [1] 108.16 31.36 9.61 40.96 470.89
```

```
y <- c(x, 0, x)
```

```
y
```

```
## [1] 10.4 5.6 3.1 6.4 21.7 0.0 10.4 5.6 3.1 6.4 21.7
```

```
1/x
```

```
## [1] 0.09615385 0.17857143 0.32258065 0.15625000 0.04608295
```

```
x^2
```

```
## [1] 108.16 31.36 9.61 40.96 470.89
```

```
y <- c(x, 0, x)
```

```
y
```

```
## [1] 10.4 5.6 3.1 6.4 21.7 0.0 10.4 5.6 3.1 6.4 21.7
```

-Vector arithmetic

```
v <- y + y + 1
```

```
v
```

```
## [1] 21.8 12.2 7.2 13.8 44.4 1.0 21.8 12.2 7.2 13.8 44.4
```

```
sum(x)
```

```
## [1] 47.2
```

```
sum((x-mean(x))^2)/(length(x)-1)
```

```
## [1] 53.853
```

```
1:18
```

```
## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
```

```
a<-1:18
```

```
a[5]
```

```
## [1] 5
```


How to download data?

- 1 put the file in the same folder as the R file (.csv or other types).
- 2 Read the file:

#Example

```
stress=read.csv("stress.csv") #read csv file
```

```
seq(-5, 5, by=1.6) -> s3
```

```
s3
```

```
## [1] -5.0 -3.4 -1.8 -0.2  1.4  3.0  4.6
```

```
a<-1:4
```

```
rep(a,times=3)
```

```
## [1] 1 2 3 4 1 2 3 4 1 2 3 4
```

```
rep(a,each=3)
```

```
## [1] 1 1 1 2 2 2 3 3 3 4 4 4
```

```
x
```

```
## [1] 10.4  5.6  3.1  6.4 21.7
```

```
temp<-x<6
```

```
temp
```

```
## [1] FALSE  TRUE  TRUE FALSE FALSE
```

```
z <- c(1:3,NA); ind <- is.na(z)
```

```
z
```

```
## [1] 1 2 3 NA
```

```
ind
```

```
## [1] FALSE FALSE FALSE TRUE
```

```
attach(faithful)
summary(eruptions)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.600   2.163   4.000   3.488   4.454   5.100
```

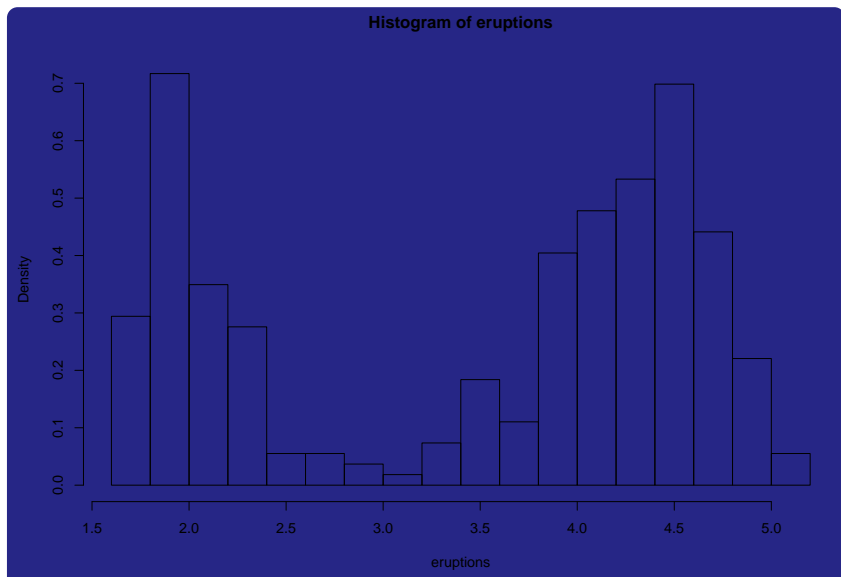
```
fivenum(eruptions)
```

```
## [1] 1.6000 2.1585 4.0000 4.4585 5.1000
```

stem(eruptions)

```
##  
## The decimal point is 1 digit(s) to the left of the |  
##  
## 16 | 070355555588  
## 18 | 000022233333335577777777888822335777888  
## 20 | 00002223378800035778  
## 22 | 0002335578023578  
## 24 | 00228  
## 26 | 23  
## 28 | 080  
## 30 | 7  
## 32 | 2337  
## 34 | 250077  
## 36 | 0000823577  
## 38 | 2333335582225577  
## 40 | 0000003357788888002233555577778  
## 42 | 03335555778800233333555577778  
## 44 | 02222335557780000000023333357778888  
## 46 | 0000233357700000023578
```

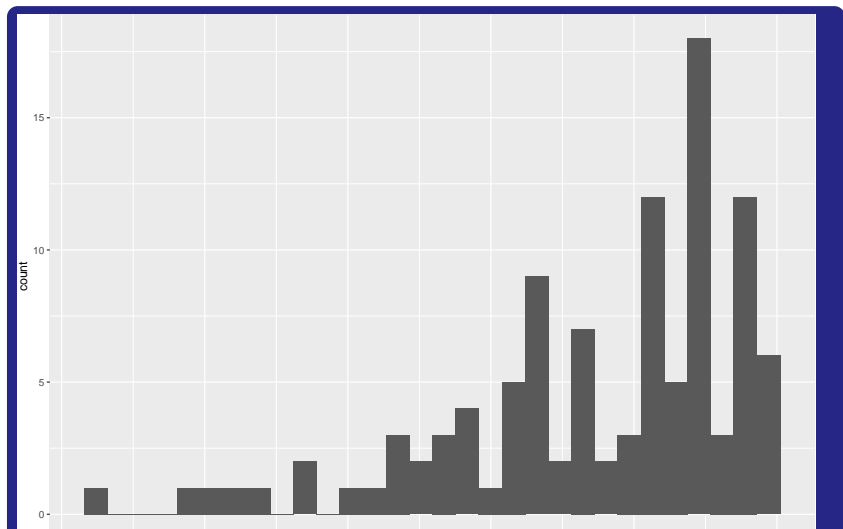
```
hist(eruptions, seq(1.6, 5.2, 0.2), prob=TRUE)
```



A simple standard plot:

```
my_plot2 <- ggplot(admit, aes(gre.quant)) + geom_histogram()  
print(my_plot2)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwi
```



Exercises

```
#Load Data mtcars  
mc=mtcars  
summary(mc)
```

```
##           mpg           cyl           disp           hp  
## Min.      :10.40    Min.      :4.000    Min.      : 71.1    Min.      : 52.  
## 1st Qu.:15.43    1st Qu.:4.000    1st Qu.:120.8    1st Qu.: 96.  
## Median :19.20    Median :6.000    Median :196.3    Median :123.  
## Mean     :20.09    Mean     :6.188    Mean     :230.7    Mean     :146.  
## 3rd Qu.:22.80    3rd Qu.:8.000    3rd Qu.:326.0    3rd Qu.:180.  
## Max.     :33.90    Max.     :8.000    Max.     :472.0    Max.     :335.  
##           drat           wt           qsec           vs  
## Min.      :2.760    Min.      :1.513    Min.      :14.50    Min.      :0.00  
## 1st Qu.:3.080    1st Qu.:2.581    1st Qu.:16.89    1st Qu.:0.00  
## Median :3.695    Median :3.325    Median :17.71    Median :0.00  
## Mean     :3.597    Mean     :3.217    Mean     :17.85    Mean     :0.43  
## 3rd Qu.:3.920    3rd Qu.:3.610    3rd Qu.:18.90    3rd Qu.:1.00  
## Max.     :4.930    Max.     :5.424    Max.     :22.90    Max.     :1.00  
##           am           gear           carb  
## Min.      :0.0000    Min.      :2.000    Min.      :1.000
```


Exercises

- 1 Find the mean, variance of wt (weight).
- 2 Plot scatter plot of wt (weight) and qsec (quarter mile per second).
- 3 From 2, what is your conclusion?
- 4 Plot scatter plot of wt (weight) and mpg (mile per gallon).
- 5 What is your conclusion?
- 6 Compute correlation of wt vs qsec.
- 7 Compute correlation of wt vs mpg.