Big Data and Analytical Lab

Lab Assignment – 0

(BCSE0183)

Packages & Library in R Programming Language

1) Install below packages using `install.package( )` function :

• Coin

```
package 'coin' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Glau\AppData\Local\Temp\Rtmpol5e0s\downloaded_packages
>
```

• XML

```
package 'XML' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Glau\AppData\Local\Temp\Rtmpol5e0s\downloaded_packages
```

• tidyr

```
package 'purrr' successfully unpacked and MD5 sums checked
package 'stringr' successfully unpacked and MD5 sums checked
package 'cpp11' successfully unpacked and MD5 sums checked
package 'tidyr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Glau\AppData\Local\Temp\Rtmpol5e0s\downloaded_packages
```

• Cluster

```
package 'cluster' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Glau\AppData\Local\Temp\Rtmpol5e0s\downloaded_packages
```

• Graphics

```
Warning in install.packages :
  package 'Graphics' is a base package, and should not be updated
```

• bioconductor packages using biocLite()

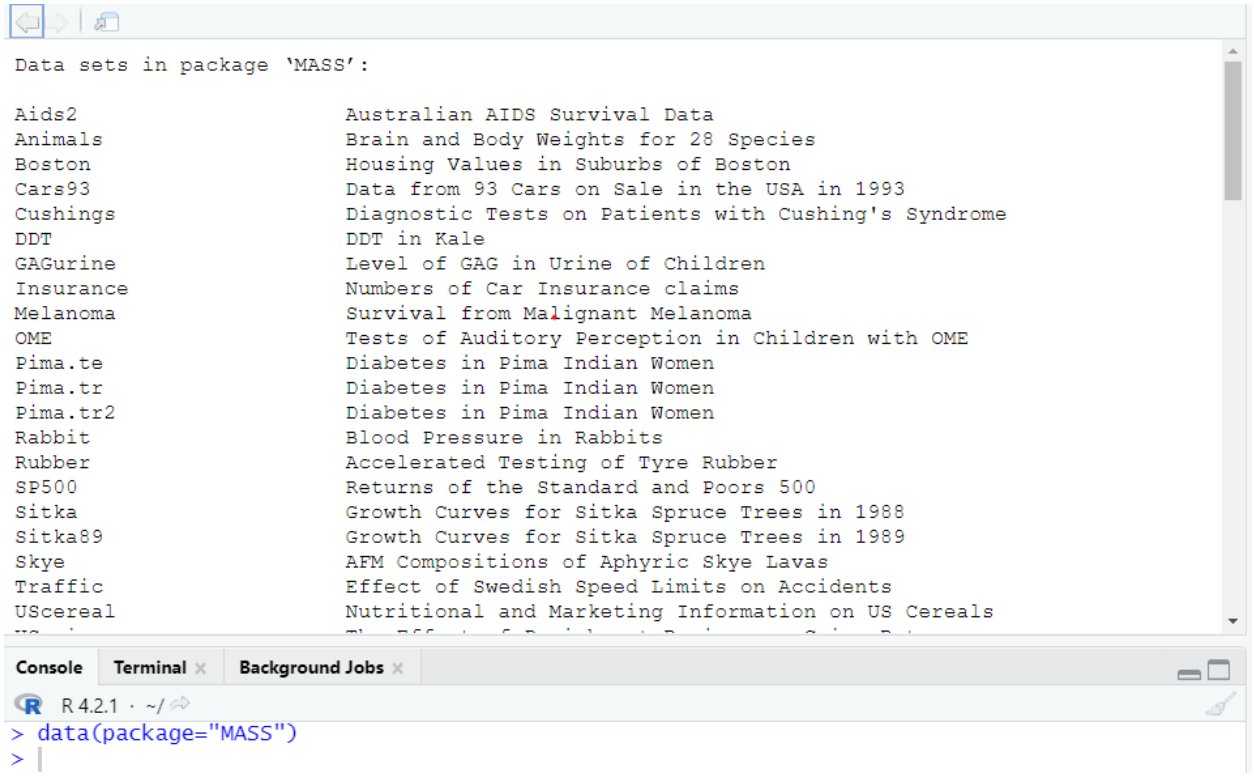
```
Warning in install.packages :
  package 'bioconductor packages using biocLite()' is not available for this version of R

A version of this package for your version of R might be available elsewhere,
see the ideas at
https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages
```

2) Display below datasets available under specific packages:

• Graphics

```
> data(package="graphics")
no data sets found
```

• MASS



- Datasets

```
Data sets in package 'datasets':

AirPassengers             Monthly Airline Passenger Numbers 1949-1960
BJsales                   Sales Data with Leading Indicator
BJsales.lead (BJsales)    Sales Data with Leading Indicator
BOD                       Biochemical Oxygen Demand
CO2                       Carbon Dioxide Uptake in Grass Plants
ChickWeight               Weight versus age of chicks on different diets
DNase                     Elisa assay of DNase
EuStockMarkets            Daily Closing Prices of Major European Stock
                          Indices, 1991-1998
Formaldehyde              Determination of Formaldehyde
HairEyeColor              Hair and Eye Color of Statistics Students
Harman23.cor              Harman Example 2.3
Harman74.cor              Harman Example 7.4
Indometh                  Pharmacokinetics of Indomethacin
InsectSprays              Effectiveness of Insect Sprays
JohnsonJohnson            Quarterly Earnings per Johnson & Johnson Share
LakeHuron                 Level of Lake Huron 1875-1972
LifeCycleSavings          Intercountry Life-Cycle Savings Data
Loblolly                  Growth of Loblolly pine trees
Nile                      Flow of the River Nile
Orange                    Growth of Orange Trees
OrchardSprays             Potency of Orchard Sprays
PlantGrowth               Results from an Experiment on Plant Growth
```

```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> data(package="datasets")
```

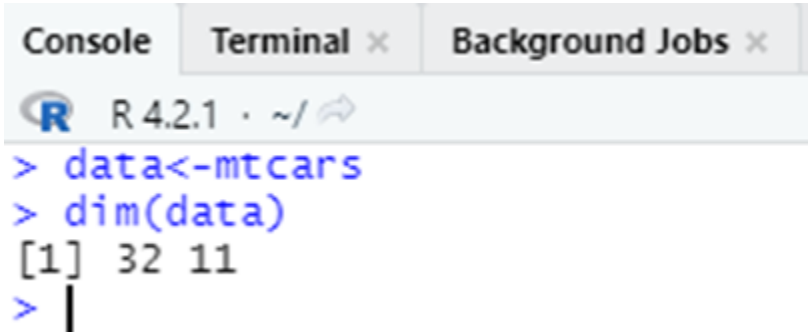
- Cluster

```
Data sets in package 'cluster':

agriculture               European Union Agricultural Workforces
animals                   Attributes of Animals
chorSub                   Subset of C-horizon of Kola Data
flower                    Flower Characteristics
plantTraits               Plant Species Traits Data
pluton                    Isotopic Composition Plutonium Batches
ruspini                   Ruspini Data
votes.repub               Votes for Republican Candidate in Presidential
                          Elections
xclara                    Bivariate Data Set with 3 Clusters
```

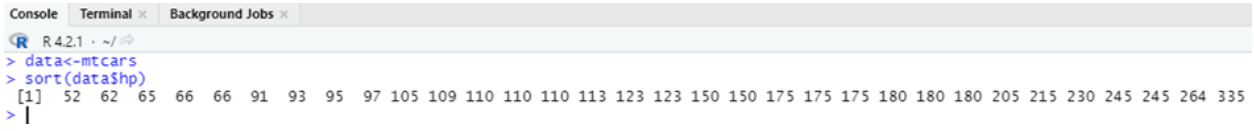
```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> data(package="cluster")
>
```

3) For given ‘mtcars’ dataset,use below functions:-

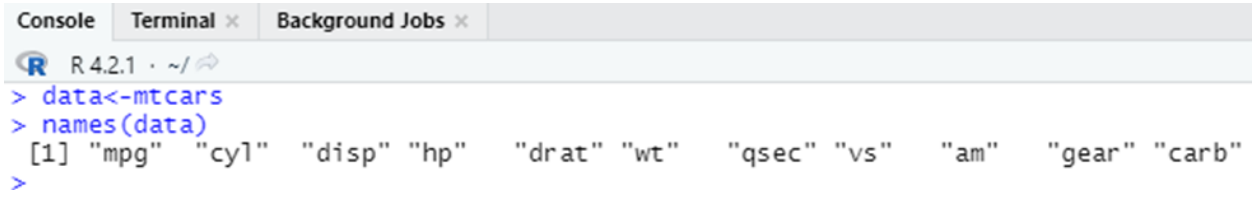
(a)dim()



(b) sort()

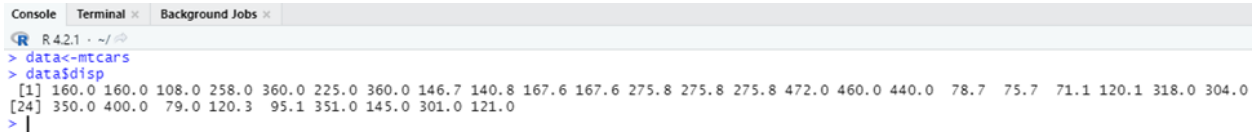


(c) names()



(d) rownames

(e) Print Variable values using ‘$’ sign



(f) summary()

```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> data<-mtcars
> summary(data)
      mpg             cyl             disp             hp             drat             wt             qsec             vs
 Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0   Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000
 1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5   1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000
 Median :19.20   Median :6.000   Median :196.3   Median :123.0   Median :3.695   Median :3.325   Median :17.71   Median :0.0000
 Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7   Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375
 3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0   3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000
 Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0   Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000
       am              gear            carb
 Min.   :0.0000   Min.   :3.000   Min.   :1.000
 1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000
 Median :0.0000   Median :4.000   Median :2.000
 Mean   :0.4062   Mean   :3.688   Mean   :2.812
 3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000
 Max.   :1.0000   Max.   :5.000   Max.   :8.000
>
```

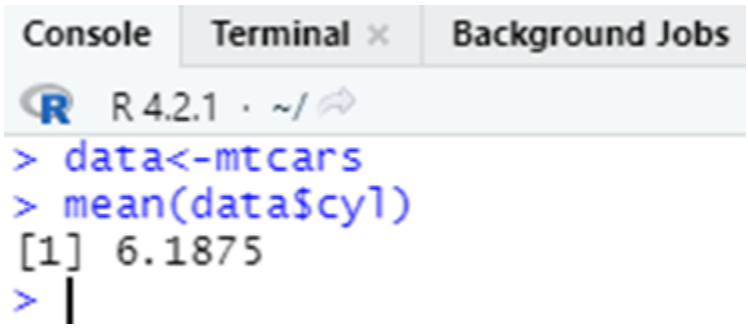
(g) max() and which.max()

```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> data<-mtcars
> max(data)
[1] 472
> which.max(data$cyl)
[1] 5
>
```

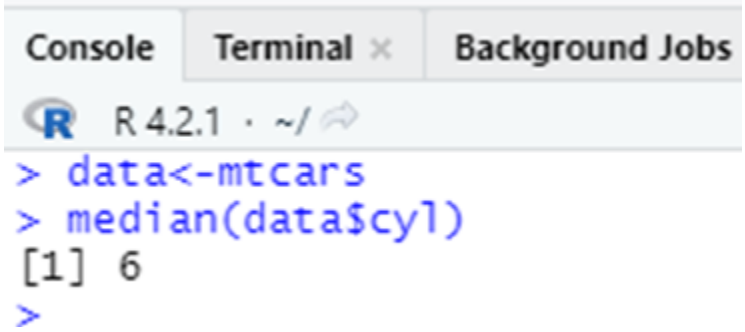
(h) min() and which.min()

```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> data<-mtcars
> min(data)
[1] 0
> which.min(data$cyl)
[1] 3
>
```

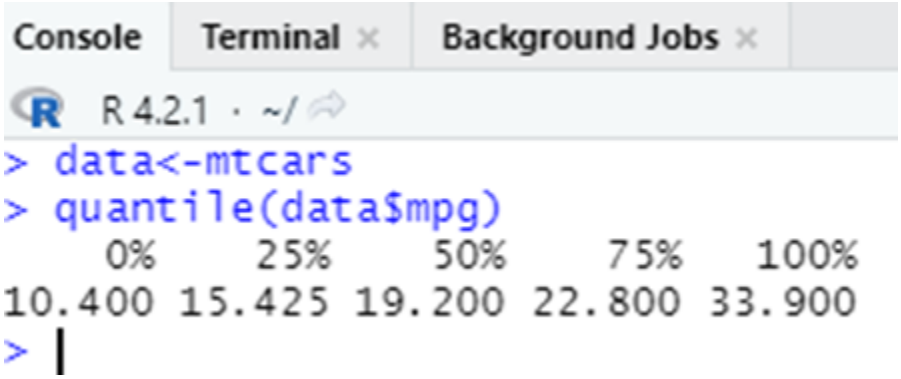
(i) mean()



(j) median()



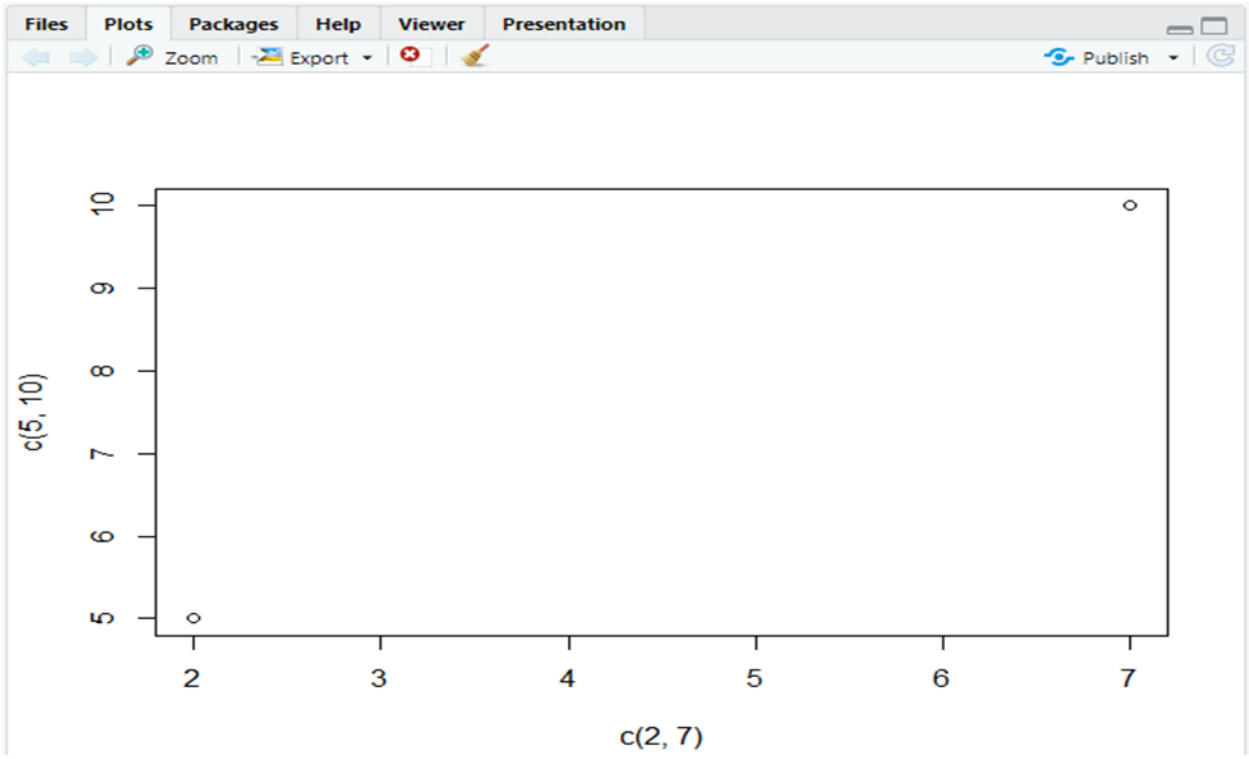
(k) quantile()



4) For given function plot ( ), use below operations:

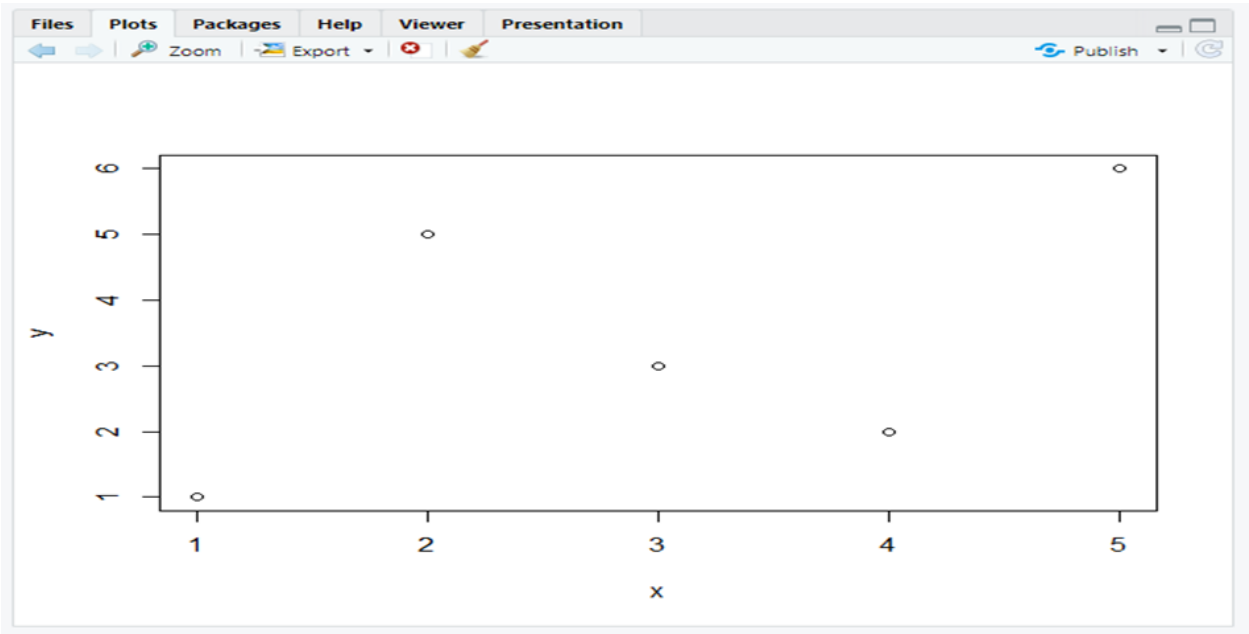
(a) Draw two points with position (2, 7) and (5, 10).

```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> plot(c(2,7),c(5,10))
>
```

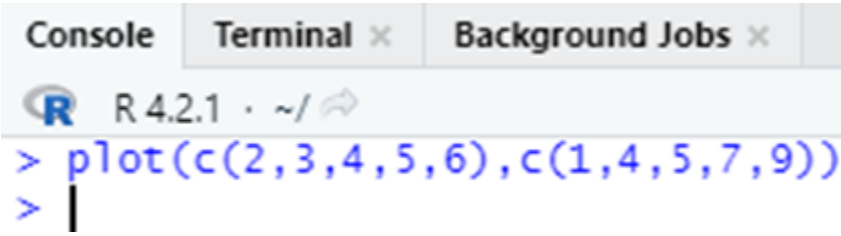


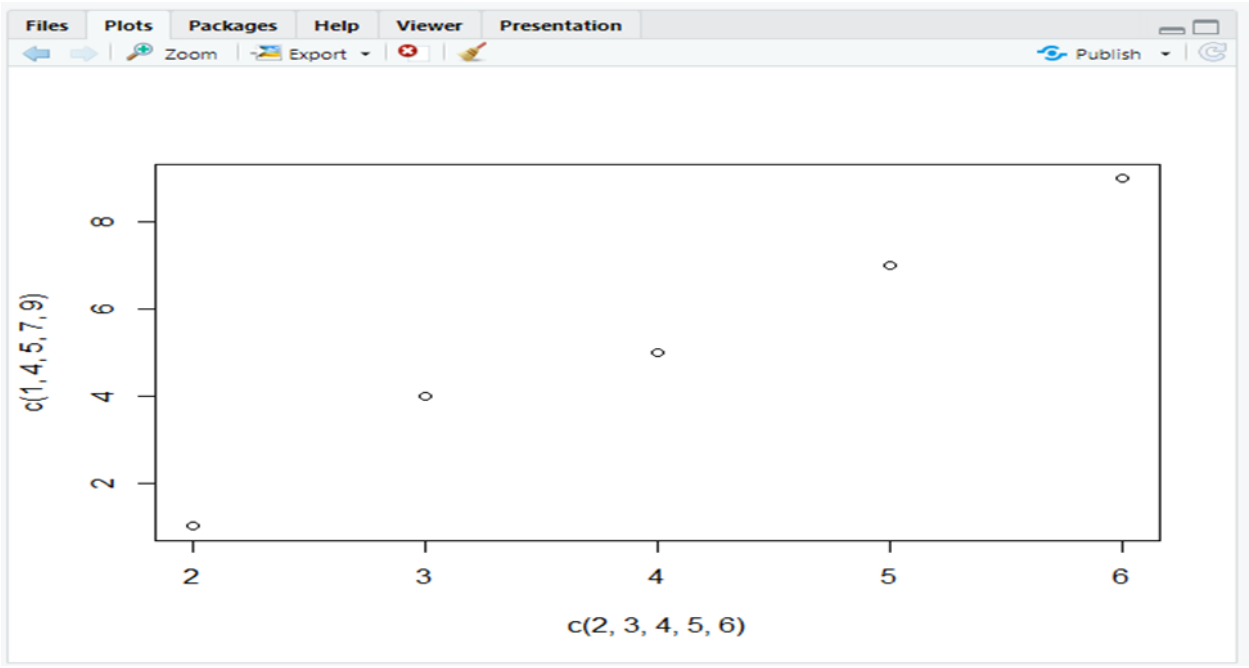
(b) Draw multiple points with variable positions

```
Console  Terminal  Background Jobs
R 4.2.1 · ~/
> x<-c(1,2,3,4,5)
> y<-c(1,5,3,2,6)
> plot(x,y)
>
```

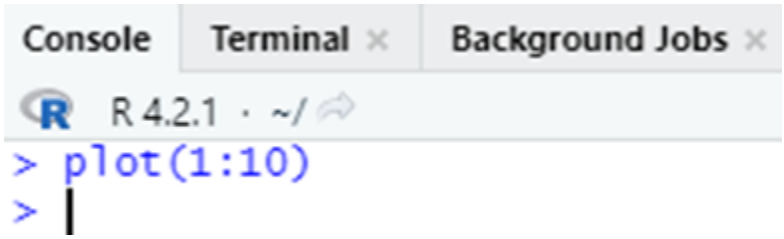


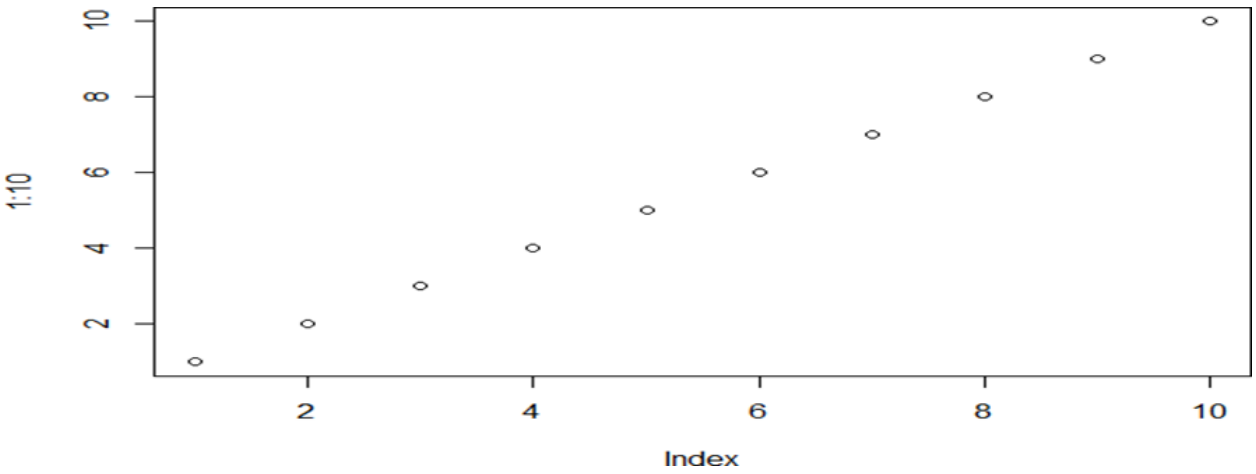
© Draw multiple points with vector





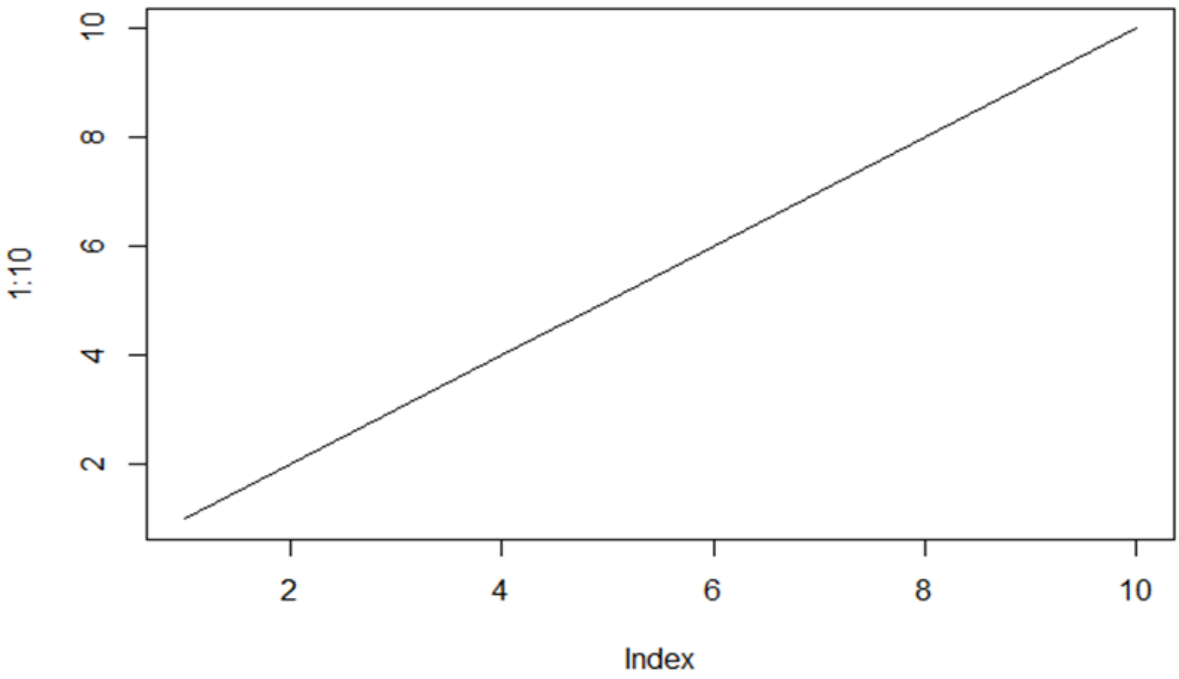
(d) Draw sequences of point





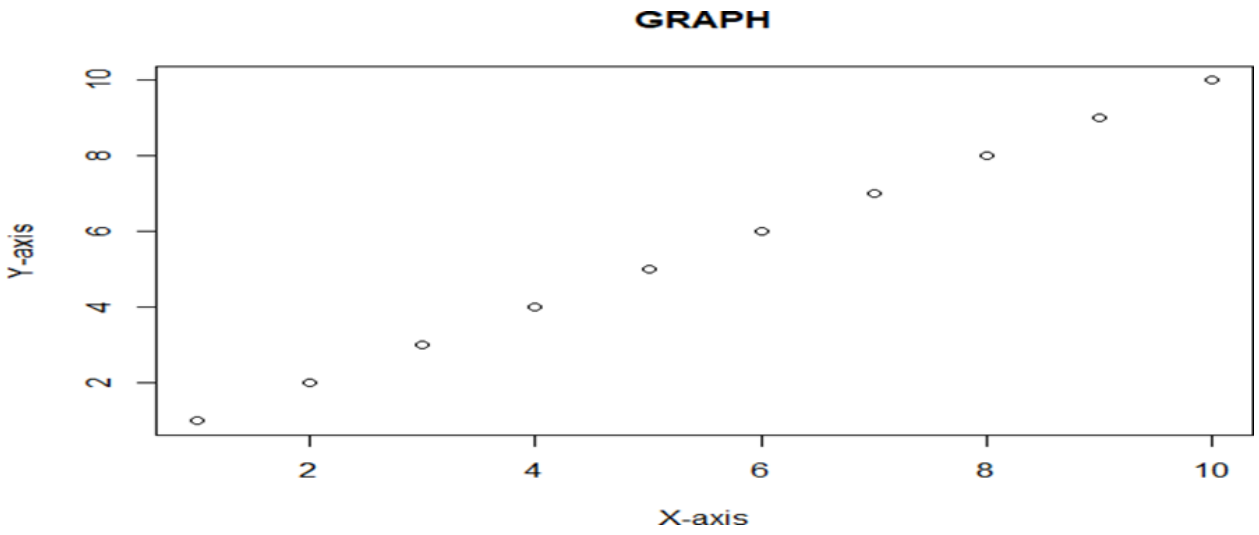
(e) Draw a line for sequences of points

```
Console   Terminal ×   Background Jobs
R 4.2.1 · ~/
> plot(1:10, type="l")
>
```

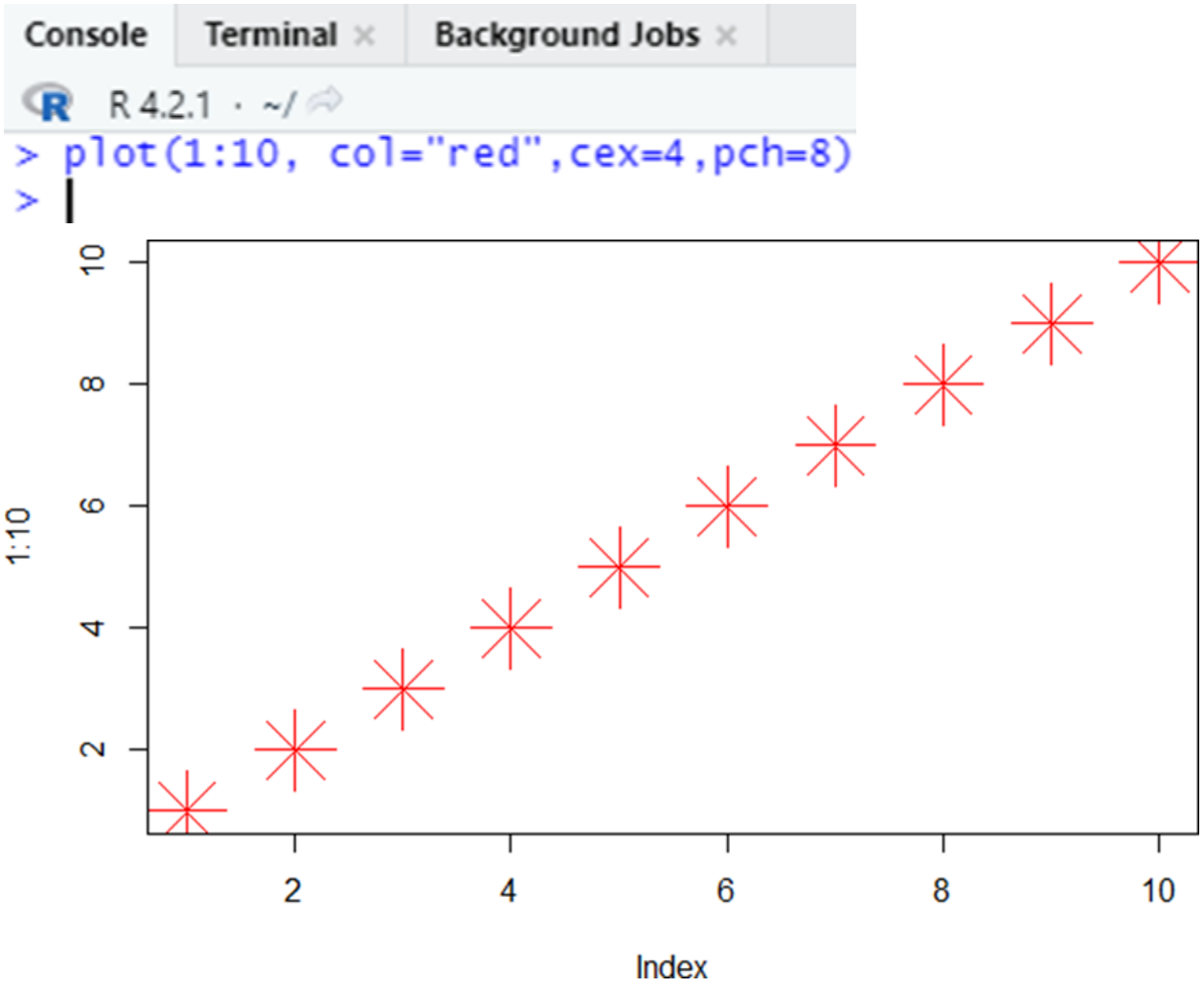


(f) Draw multiple points with parameters like main, xlab and ylab

```
Console   Terminal ×   Background Jobs ×
R 4.2.1 · ~/
> plot(1:10, main="GRAPH",xlab="X-axis",ylab="Y-axis")
>
```



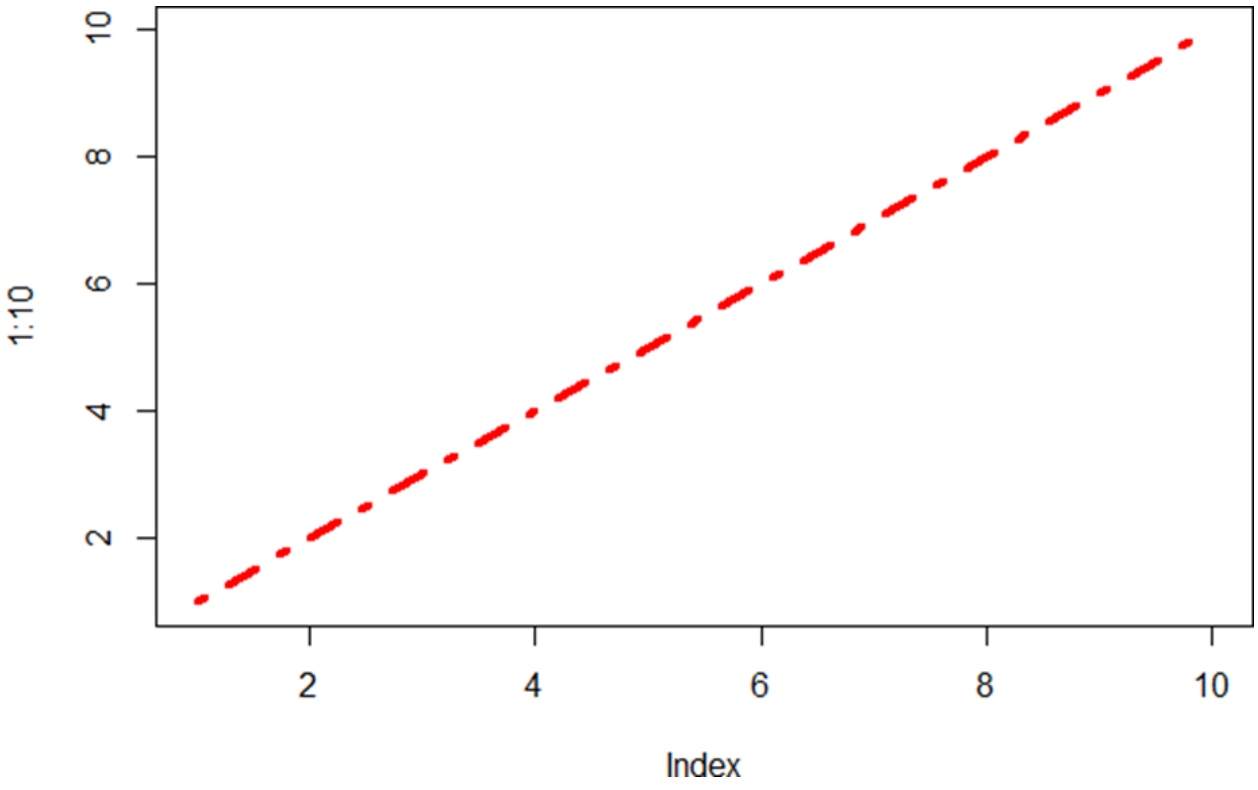
(g) Draw multiple points with point shape, colors and size



5) Compare two plots on following criteria:

(a) Draw Multiple Lines with style, width and color.

```
Console   Terminal   Background Jobs
R 4.2.1 · ~/
> plot(1:10, type="l",col="red",lwd=4,lty=4)
```



(b) Draw the observations for two different dataset using points( ) function.

Console | Terminal | Background Jobs

R 4.2.1 · ~/

```
> x<-c(1,2,3,4,5)
> y<-c(2,4,6,3,8)
>
> a<-c(3,4,6,5,7)
> b<-c(9,7,6,8,5)
>
> plot(x,y,main="My Observation",xlab="X-axis",ylab="Y-axis",col="red",cex=3)
> points(a,b,col="blue",cex=3)
>
```