Data Science - RStudio

Compulsory

```r
# Imagine the following scenario: You are a data analyst at an
organisation. You have been given a data set and asked to
# create a meaningful data visualisation using this data. Using the
ggplotin-built data sets in RStudio and the qplot function,
# get your creative juices flowing and create a meaningful and
impactful data visualisation using your preferred data set.

# 1. Write an R program to create three vectors a, b, c with 5
integers. Combine the three vectors to become a 3×5 matrix
# where each column represents a vector. Print the content of the
matrix. Plot a graph and label correctly.

a<-c(1,2,3)
b<-c(4,5,6)
c<-c(7,8,9)
m<-cbind(a,b,c)
print(m)

     a b c
[1,] 1 4 7
[2,] 2 5 8
[3,] 3 6 9

matplot(t(m),
        type = "l",
        lwd = 2,
        main="The Rows of a Matrix",
        ylab="Value")
```
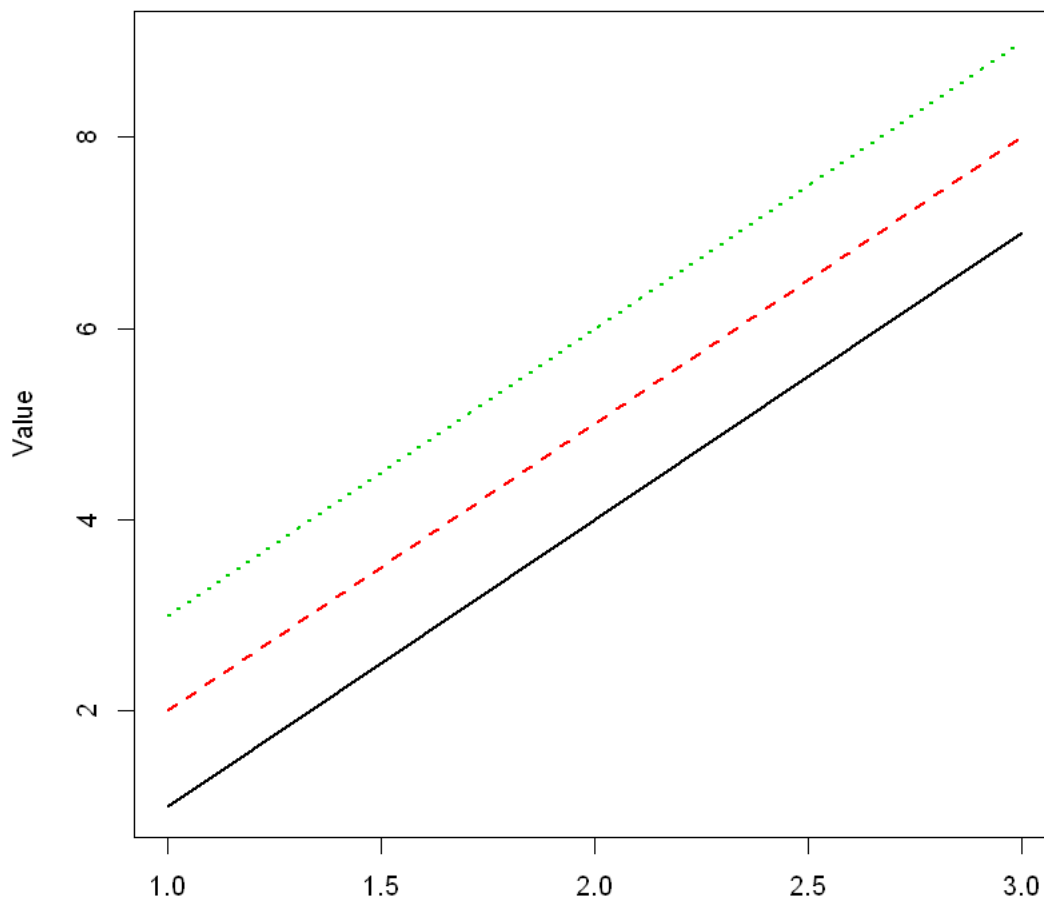
The Rows of a Matrix


```r
# 2. Write a R program to create a Data frames which contain details
of 5 employees and display the details.
# (Name, Age, Role and Length of service).

employee.data <- data.frame(
  employee_id = c (1:5),
  employee_name = c("Charlie","Harry","Laura","Eloise","Simon"),
  age = c(30,28,40,48,27),
  role = c("business analyst", "sales executive", "CEO", "COO",
"CFO"),
  length_of_service = (c(3, 5, 8, 4, 6)),
  stringsAsFactors = FALSE
)

print(employee.data)
```

```
  employee_id employee_name age                role length_of_service
1           1       Charlie  30 business analyst                    3
2           2         Harry  28  sales executive                    5
3           3         Laura  40              CEO                     8
4           4        Eloise  48              COO                     4
5           5         Simon  27              CFO                     6
```

```r
# 3. Import the GGPLOT 2 library and plot a graph using the qplot
function. X axis is the sequence of 1:20 and the y axis
# is the x ^ 2. Label the graph appropriately.
install.packages("ggplot2", dependencies = TRUE)

# Import Libraries
library(ggplot2)

x <- c(1:20)
print(x)
```
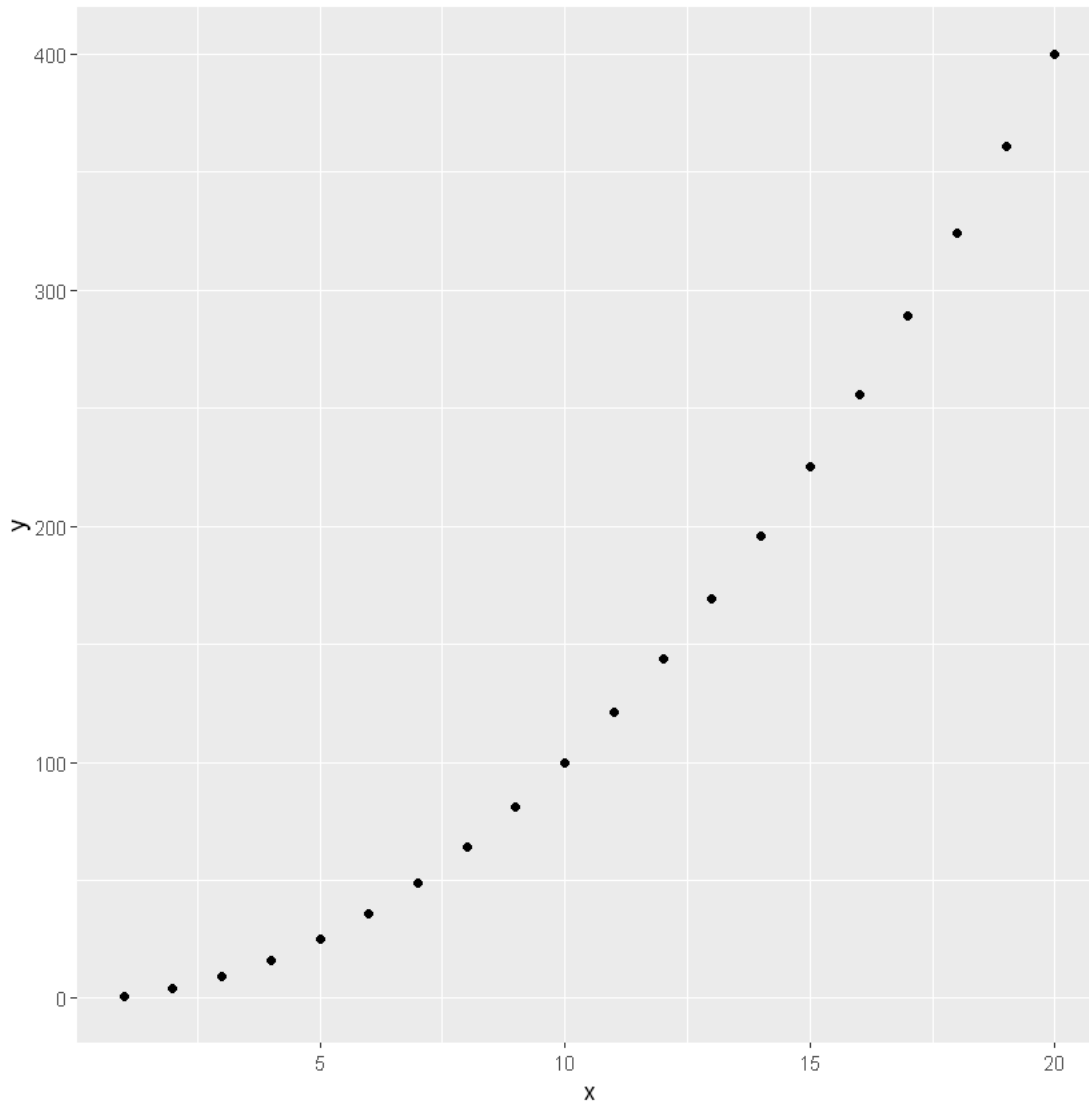
```
 [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```
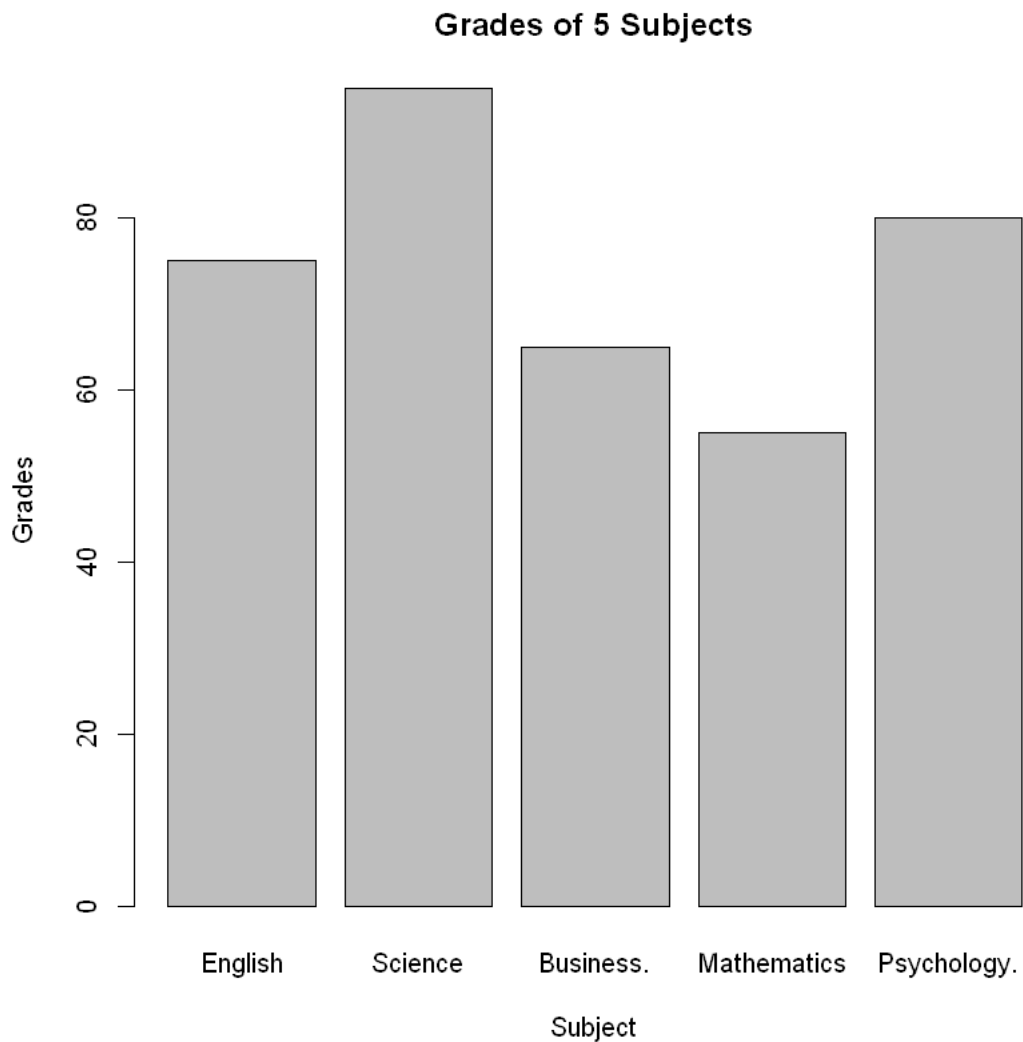
```r
y <- x^2
print(y)
```

```
 [1]   1   4   9  16  25  36  49  64  81 100 121 144 169 196 225 256
289 324 361
[20] 400
```

```r
qplot(x, y)
```

```r
# 4. Create a simple bar plot of five subjects

Grades = c(75, 95, 65, 55, 80)
barplot(Grades,
        main = "Grades of 5 Subjects",
        xlab = "Subject",
        ylab = "Grades",
        names.arg = c("English", "Science", "Business.",
"Mathematics", "Psychology."),
        col = "gray",
        horiz = FALSE)
```

Grades of 5 Subjects

Additional

```
# 1. Write a R program to take input from the user (name and age) and
display the values.

name = readline(prompt="Input your name: ")
age =  readline(prompt="Input your age: ")
print(paste("My name is",name, "and I am",age ,"years old."))

# 2. Write a R program to create a sequence of numbers from 20 to 50
and find the mean of numbers from 20 to 50 and sum of
# numbers.

print(seq(20,50))
print(mean(20:50))
print(sum(20:50))
```

```r
# 3. Write a R program to create a vector which contains 10 random
integer values between -50 and +50

x = sample(-50:50, 10)
print(x)
```