

R

R is a programming language and open-source software used for statistical computing and data analysis. It was specifically designed for data analysis and statistical modeling, making it a popular choice among statisticians, data analysts, and data scientists.

In this sheets the Basic Concepts for Beginners ❤

<https://github.com/qusaybtoush>

<https://www.linkedin.com/in/qusayal-btoush>

<https://www.kaggle.com/qusaybtoush1990>

In [1]:

```
# Define the Variables
#using '<- ' OR '='

a = 9      #numeric
x <- 5.5    #numeric
y = "one"   #character
z= TRUE     #boolean  TRUE or FALSE - you can writing T or F
q = a + 9+5.5    # you can add multi variables

a  # you can writing  print (a or x ...)

x
y
z
q
```

9

5.5

'one'

TRUE

23.5

In [2]:

```
# show the type
class(x) # you can writing class(a or y or z)
```

'numeric'

In [3]:

```
#create a sequence of elements
x <- seq (1,10)
x
y <- seq (1,10,3)    #you can add step any number
y

#select element index 3          * the index start from 1
x[3]  # index 3
x[-1] # all element without the index 1 , you can writing -2 to -10 if you don't need
```

1·2·3·4·5·6·7·8·9·10

1·4·7·10

3

2·3·4·5·6·7·8·9·10

In [4]:

```
# you can Create any mathematical equation
x + y
x * y
x - y
```

Warning message in x + y:

“longer object length is not a multiple of shorter object length”

2·6·10·14·6·10·14·18·10·14

```
Warning message in x * y:  
“longer object length is not a multiple of shorter object length”  
1·8·21·40·5·24·49·80·9·40  
  
Warning message in x - y:  
“longer object length is not a multiple of shorter object length”  
0··2··4··6·4·2·0··2·8·6
```

In [5]:

```
# Define the List  
x <- list (1 : 10)  
#or  
y<- list ( 5,9, 15 ,1)  
x  
y
```

```
1. 1·2·3·4·5·6·7·8·9·10
```

```
1.5  
2.9  
3.15  
4.1
```

In [6]:

```
# Define the Factors           * using the factor with category data  
F <- c("M","M","F","M","F")  
F<-factor(F)  
summary(F)
```

F:2 M:3

In [7]:

```
# How to create a Vectors  
# A vector can be created using an in-built function in R called c(). Elements must be  
  
x <- c(10, 5, 7, 15,7 )      #number  
y <- c("A","XX","one")        #string  
z = c(10, 1.5, "Cat", TRUE , seq(5,9))  #you can add any type  
  
x  
y  
z
```

```
10·5·7·15·7
```

```
'A' · 'XX' · 'one'  
'10' · '1.5' · 'Cat' · 'TRUE' · '5' · '6' · '7' · '8' · '9'
```

In [8]:

```
# sum, sort, mean, ...  
  
R <- c(10, 5, 7, 15,7,1,5 )  
sort( R) # sort the emelent  
mean( R) # the average
```

```
max (R)    # the maximum
min (R)    # the minimum
```

1 · 5 · 7 · 7 · 10 · 15

7.14285714285714

15

1

In [9]:

```
#To create a matrix in R you need to use the function called matrix()
#The arguments to this      = matrix(data, number of rows, number of columns ,filling
row <- c('A','B','C','P') # you can writing any name
col <- c ('col 1', 'col 2', 'col 3','col 4','col 5') # you can writing any name

Mat <- matrix(c(1:20), nrow= 4, ncol =5 , byrow = T , dimnames =list(row,col))

Mat
```

A matrix: 4 × 5 of type int

	col 1	col 2	col 3	col 4	col 5
A	1	2	3	4	5
B	6	7	8	9	10
C	11	12	13	14	15
P	16	17	18	19	20

In [10]:

```
# select the element in the matrix / just choice the index [ row index , column index ]
Mat[ 3,2]

Mat[ 2,5] = 100 # you can change the number inside the matrix just choice the index
Mat
```

12

A matrix: 4 × 5 of type dbl

	col 1	col 2	col 3	col 4	col 5
A	1	2	3	4	5
B	6	7	8	9	100
C	11	12	13	14	15
P	16	17	18	19	20

In [11]:

```
# IF statement
M <- 75

if (M >= 50){                      # you can writing if Like = if(x>=90 && x<100 ) ...
  print ("Good")
}else {
  print ("Bad")
}
```

```
[1] "Good"
```

In [12]:

```
# while Loop
L =1
while (L<10){
  print (L)
  L<- L+4
} # using this Loop I can writing if statement and for Loop inside this
```

```
[1] 1
[1] 5
[1] 9
```

In [13]:

```
# for Loop
for (i in 1:10) # this Loop print from range 1 to 10
  print (i)
```

```
[1] 1
[1] 2
[1] 3
[1] 4
[1] 5
[1] 6
[1] 7
[1] 8
[1] 9
[1] 10
```

In [14]:

```
# Define the Function
# name on function <- writing function (here the argument) { here the body function w
Fun <- function (y ){
  return (y * 50) # you can writing a lot of things here, for Loop or if ...
}
Fun(10) # 10 this = y
```

500

Import Data Set

In [15]:

```
# Import Data
# You can Import data from multiple sources / excel, csv, sql, web page, spss ,...
# you can join more the 2 table using function inner_join or left_join, right_join

df = read.csv('/kaggle/input/task-sales/Task-1.csv') # now, import from CSV , first upl
```

In [16]:

```
#show the DataSet
df
```

A data.frame: 9648 × 12

Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold
<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>

Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold
<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>
Foot Locker	1185732	01-01-20	Northeast	New York	New York	Men's Street Footwear	50	1200
Foot Locker	1185732	02-01-20	Northeast	New York	New York	Men's Athletic Footwear	50	1000
Foot Locker	1185732	03-01-20	Northeast	New York	New York	Women's Street Footwear	40	1000
Foot Locker	1185732	04-01-20	Northeast	New York	New York	Women's Athletic Footwear	45	850
Foot Locker	1185732	05-01-20	Northeast	New York	New York	Men's Apparel	60	900
Foot Locker	1185732	06-01-20	Northeast	New York	New York	Women's Apparel	50	1000
Foot Locker	1185732	07-01-20	Northeast	New York	New York	Men's Street Footwear	50	1250
Foot Locker	1185732	08-01-20	Northeast	New York	New York	Men's Athletic Footwear	50	900
Foot Locker	1185732	21-01-20	Northeast	New York	New York	Women's Street Footwear	40	950
Foot Locker	1185732	22-01-20	Northeast	New York	New York	Women's Athletic Footwear	45	825
Foot Locker	1185732	23-01-20	Northeast	New York	New York	Men's Apparel	60	900
Foot Locker	1185732	24-01-20	Northeast	New York	New York	Women's Apparel	50	1000
Foot Locker	1185732	25-01-20	Northeast	New York	New York	Men's Street Footwear	50	1220
Foot Locker	1185732	26-01-20	Northeast	New York	New York	Men's Athletic Footwear	50	925
Foot Locker	1185732	27-01-20	Northeast	New York	New York	Women's Street Footwear	40	950
Foot Locker	1185732	28-01-20	Northeast	New York	New York	Women's Athletic Footwear	45	800

Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold
<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>
Foot Locker	1185732	29-01-20	Northeast	New York	New York	Men's Apparel	NA	850
Foot Locker	1185732	30-01-20	Northeast	New York	New York	Women's Apparel	50	950
Foot Locker	1185732	31-01-20	Northeast	New York	New York	Men's Street Footwear	50	1200
Foot Locker	1185732	01-02-20	Northeast	New York	New York	Men's Athletic Footwear	50	900
Foot Locker	1185732	02-02-20	Northeast	New York	New York	Women's Street Footwear	40	900
Foot Locker	1185732	03-02-20	Northeast	New York	New York	Women's Athletic Footwear	45	825
Foot Locker	1185732	04-02-20	Northeast	New York	New York	Men's Apparel	60	825
Foot Locker	1185732	05-02-20	Northeast	New York	New York	Women's Apparel	50	950
Foot Locker	1185732	06-02-20	Northeast	New York	New York	Men's Street Footwear	60	1220
Foot Locker	1185732	07-02-20	Northeast	New York	New York	Men's Athletic Footwear	55	925
Foot Locker	1185732	08-02-20	Northeast	New York	New York	Women's Street Footwear	50	900
Foot Locker	1185732	09-02-20	Northeast	New York	New York	Women's Athletic Footwear	50	850
Foot Locker	1185732	10-02-20	Northeast	New York	New York	Men's Apparel	60	875
Foot Locker	1185732	03-03-20	Northeast	New York	New York	Women's Apparel	65	1000
:	:	:	:	:	:	:	:	:
Amazon	1185732	17-09-21	Northeast	New Hampshire	Manchester	Women's Athletic Footwear	41	74
Amazon	1185732	17-09-21	Northeast	New Hampshire	Manchester	Men's Apparel	48	70
Amazon	1185732	17-09-21	Northeast	New Hampshire	Manchester	Women's Apparel	53	104

Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold
<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>
Foot Locker	1185732	19-10-21	Northeast	New Hampshire	Manchester	Men's Street Footwear	52	162
Foot Locker	1185732	19-10-21	Northeast	New Hampshire	Manchester	Men's Athletic Footwear	44	99
Foot Locker	1185732	19-10-21	Northeast	New Hampshire	Manchester	Women's Street Footwear	52	60
Foot Locker	1185732	19-10-21	Northeast	New Hampshire	Manchester	Women's Athletic Footwear	44	61
Foot Locker	1185732	19-10-21	Northeast	New Hampshire	Manchester	Men's Apparel	50	58
Foot Locker	1185732	19-10-21	Northeast	New Hampshire	Manchester	Women's Apparel	64	90
Foot Locker	1185732	18-11-21	Northeast	New Hampshire	Manchester	Men's Street Footwear	59	149
Foot Locker	1185732	18-11-21	Northeast	New Hampshire	Manchester	Men's Athletic Footwear	42	114
Foot Locker	1185732	18-11-21	Northeast	New Hampshire	Manchester	Women's Street Footwear	51	114
Foot Locker	1185732	18-11-21	Northeast	New Hampshire	Manchester	Women's Athletic Footwear	43	114
Foot Locker	1185732	18-11-21	Northeast	New Hampshire	Manchester	Men's Apparel	59	96
Foot Locker	1185732	18-11-21	Northeast	New Hampshire	Manchester	Women's Apparel	61	120
Foot Locker	1185732	17-12-21	Northeast	New Hampshire	Manchester	Men's Street Footwear	49	200
Foot Locker	1185732	17-12-21	Northeast	New Hampshire	Manchester	Men's Athletic Footwear	49	128
Foot Locker	1185732	17-12-21	Northeast	New Hampshire	Manchester	Women's Street Footwear	47	128
Foot Locker	1185732	17-12-21	Northeast	New Hampshire	Manchester	Women's Athletic Footwear	49	116

Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold
<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>
Foot Locker	1185732	17-12-21	Northeast	New Hampshire	Manchester	Men's Apparel	50	123
Foot Locker	1185732	17-12-21	Northeast	New Hampshire	Manchester	Women's Apparel	61	144
Foot Locker	1185732	24-01-21	Northeast	New Hampshire	Manchester	Men's Street Footwear	36	165
Foot Locker	1185732	24-01-21	Northeast	New Hampshire	Manchester	Men's Athletic Footwear	36	93
Foot Locker	1185732	24-01-21	Northeast	New Hampshire	Manchester	Women's Street Footwear	35	99
Foot Locker	1185732	24-01-21	Northeast	New Hampshire	Manchester	Women's Athletic Footwear	33	51
Foot Locker	1185732	24-01-21	Northeast	New Hampshire	Manchester	Men's Apparel	50	64
Foot Locker	1185732	24-01-21	Northeast	New Hampshire	Manchester	Women's Apparel	41	105
Foot Locker	1185732	22-02-21	Northeast	New Hampshire	Manchester	Men's Street Footwear	41	184
Foot Locker	1185732	22-02-21	Northeast	New Hampshire	Manchester	Men's Athletic Footwear	42	70
Foot Locker	1185732	22-02-21	Northeast	New Hampshire	Manchester	Women's Street Footwear	29	83

In [17]:

```
#show the columns name with data type using function str(the name dataframe )
str(df)
```

```
'data.frame': 9648 obs. of 12 variables:
 $ Retailer      : chr "Foot Locker" "Foot Locker" "Foot Locker" ...
 $ Retailer.ID   : int 1185732 1185732 1185732 1185732 1185732 1185732 1185732 ...
 $ Invoice.Date  : chr "01-01-20" "02-01-20" "03-01-20" "04-01-20" ...
 $ Region        : chr "Northeast" "Northeast" "Northeast" "Northeast" ...
 $ State         : chr "New York" "New York" "New York" "New York" ...
 $ City          : chr "New York" "New York" "New York" "New York" ...
 $ Product       : chr "Men's Street Footwear" "Men's Athletic Footwear" "Women's Street Footwear" "Women's Athletic Footwear" ...
 $ Price.per.Unit: int 50 50 40 45 60 50 50 40 45 ...
 $ Units.Sold    : int 1200 1000 1000 850 900 1000 1250 900 950 825 ...
 $ Total.Sales   : int 60000 50000 40000 38250 54000 50000 62500 45000 38000 37125 ...
 ...
```

```
$ Operating.Profit: int 30000 15000 14000 13388 16200 12500 31250 13500 13300 12994
...
$ Sales.Method : chr "In-store" "In-store" "In-store" "In-store" ...
```

In [18]:

```
# show the shape the dataset, how many row and column using function dim(the name data)
dim(df)
```

9648 · 12

In [19]:

```
# show the summary / description about the dataset using function summary (the name data
# you can see the min, max, mean, data type ... everything about the data
summary(df)
```

Retailer	Retailer.ID	Invoice.Date	Region
Length:9648	Min. :1128299	Length:9648	Length:9648
Class :character	1st Qu.:1185732	Class :character	Class :character
Mode :character	Median :1185732	Mode :character	Mode :character
	Mean :1173850		
	3rd Qu.:1185732		
	Max. :1197831		
State	City	Product	Price.per.Unit
Length:9648	Length:9648	Length:9648	Min. : 7.00
Class :character	Class :character	Class :character	1st Qu.: 35.00
Mode :character	Mode :character	Mode :character	Median : 45.00
			Mean : 45.22
			3rd Qu.: 55.00
			Max. :110.00
			NA's :2
Units.Sold	Total.Sales	Operating.Profit	Sales.Method
Min. : 0.0	Min. : 0.0	Min. : 0	Length:9648
1st Qu.: 106.0	1st Qu.: 425.8	1st Qu.: 192	Class :character
Median : 176.0	Median : 958.0	Median : 437	Mode :character
Mean : 256.9	Mean : 9327.4	Mean : 3443	
3rd Qu.: 350.0	3rd Qu.:15000.0	3rd Qu.: 5206	
Max. :1275.0	Max. :82500.0	Max. :39000	

In [20]:

```
#show the number of row first or last, head(df, the number of rows, the Default = 6 rows)
head(df)
```

A data.frame: 6 × 12

	Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold	Total
	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	
1	Foot Locker	1185732	01-01-20	Northeast	New York	New York	Men's Street Footwear	50	1200	
2	Foot Locker	1185732	02-01-20	Northeast	New York	New York	Men's Athletic Footwear	50	1000	
3	Foot Locker	1185732	03-01-20	Northeast	New York	New York	Women's Street Footwear	40	1000	

	Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold	Total
	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	
4	Foot Locker	1185732	04-01-20	Northeast	New York	New York	Women's Athletic Footwear	45	850	
5	Foot Locker	1185732	05-01-20	Northeast	New York	New York	Men's Apparel	60	900	
6	Foot Locker	1185732	06-01-20	Northeast	New York	New York	Women's Apparel	50	1000	

In [21]:

```
# Rename the Column - just write rename(data frame , new name = old name)
#rename(df, "new co"= "City")
```

In [22]:

```
# check if there duplication in the data set / sum - mean the total
sum(duplicated(df))
```

0

In [23]:

```
# check the unique value or in all data set
#sum(unique(df))
```

In [24]:

```
# check if there missing value in the data set
sum(is.na(df))
```

2

In [25]:

```
# drop the nan / missing value and save the data after drop the missing
df <- na.omit(df)
```

In [26]:

```
# function names to show the names of columns in the data set
names(df)
```

'Retailer' · 'Retailer.ID' · 'Invoice.Date' · 'Region' · 'State' · 'City' · 'Product' · 'Price.per.Unit' · 'Units.Sold' · 'Total.Sales' · 'Operating.Profit' · 'Sales.Method'

In [27]:

```
sum(is.na(df))
```

0

In [28]:

```
# add new column - price * units sold column // you can choich any operation + / - *
# the name new column r
df$r <- df$Price.per.Unit * df$Units.Sold
```

In [29]:

```
# show the first 6 rows and show the new column r
head(df)
```

A data.frame: 6 × 13

	Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold	Total
	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	<int>
1	Foot Locker	1185732	01-01-20	Northeast	New York	New York	Men's Street Footwear	50	1200	
2	Foot Locker	1185732	02-01-20	Northeast	New York	New York	Men's Athletic Footwear	50	1000	
3	Foot Locker	1185732	03-01-20	Northeast	New York	New York	Women's Street Footwear	40	1000	
4	Foot Locker	1185732	04-01-20	Northeast	New York	New York	Women's Athletic Footwear	45	850	
5	Foot Locker	1185732	05-01-20	Northeast	New York	New York	Men's Apparel	60	900	
6	Foot Locker	1185732	06-01-20	Northeast	New York	New York	Women's Apparel	50	1000	

dplyr

In [30]:

```
# dplyr
#is a grammar of data manipulation, providing a consistent set of verbs that help you s

# mutate() adds new variables that are functions of existing variables
# select() picks variables based on their names.
# filter() picks cases based on their values.
# summarise() reduces multiple values down to a single summary.
# arrange() changes the ordering of the rows.
# group by

#how you can install and load any library in r

#install.packages("dplyr") #install Library
#library(dplyr)           # Load Library
```

In [31]:

```
install.packages("dplyr")
```

Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

Warning message:

“unable to access index for repository http://cran.rstudio.com/src/contrib:
cannot open URL ‘http://cran.rstudio.com/src/contrib/PACKAGES’”

Warning message:

“package ‘dplyr’ 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>

In [32]:

```
library(dplyr)
```

Attaching package: ‘dplyr’

The following objects are masked from ‘package:stats’:

filter, lag

The following objects are masked from ‘package:base’:

intersect, setdiff, setequal, union

In [33]:

```
# make gorup by column Retailer and the sum of profit / you can write in the end on line
df%>%
  group_by(Retailer)%>%
  summarize(sum(Operating.Profit))
```

A tibble: 6 × 2

Retailer sum(Operating.Profit)

<chr>	<int>
Amazon	2881885
Foot Locker	8057050
Kohl's	3681153
Sports Direct	7427383
Walmart	2578226
West Gear	8566869

In [34]:

```
# here show the min, max and average the price and create new name
df%>%
  summarize(MAX = max(Price.per.Unit), MIN = min(Price.per.Unit), AVG = mean(Price.per.U
```

A data.frame: 1 × 3

MAX MIN AVG

<int>	<int>	<dbl>
110	7	45.21563

In [35]:

```
#show the unique function and how many unique value in this column/ using this $ after
unique(df$Sales.Method)
```

'In-store' · 'Outlet' · 'Online'

In [36]:

```
# here we using filter function / filter on column Sales method value = online or insto
# and using select function to choich 3 column and show the 3 rows using head

df %>%
filter(Sales.Method=="Online" || Sales.Method=="In-store" ) %>%
select(Region,Product,Price.per.Unit )%>%
head(3)
```

A data.frame: 3 × 3

	Region	Product	Price.per.Unit
			<chr>
1	Northeast	Men's Street Footwear	50
2	Northeast	Men's Athletic Footwear	50
3	Northeast	Women's Street Footwear	40

In [37]:

```
# how you can sort the data using arrange function
df%>%
arrange(Price.per.Unit)%>%
head(3)
```

A data.frame: 3 × 13

	Retailer	Retailer.ID	Invoice.Date	Region	State	City	Product	Price.per.Unit	Units.Sold		
	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>		
1	Sports Direct	1197831	25-08-21	South	Alabama	Birmingham	Men's Street Footwear	7	224		
2	West Gear	1185732	12-09-21	West	Utah	Salt Lake City	Women's Street Footwear	7	105		
3	West Gear	1185732	12-09-21	West	Utah	Salt Lake City	Women's Athletic Footwear	7	111		

◀ ▶

```
# function Mutate to creating new column using any operation math

df%>%
mutate (new_c = Price.per.Unit * 1.5 ) %>%
select (Retailer,Product,Price.per.Unit,Sales.Method,new_c)%>%
head(3)
```

A data.frame: 3 × 5

Retailer	Product	Price.per.Unit	Sales.Method	new_c
<chr>	<chr>	<int>	<chr>	<dbl>

Retailer	Product	Price.per.Unit	Sales.Method	new_c
<chr>	<chr>	<int>	<chr>	<dbl>
1 Foot Locker	Men's Street Footwear	50	In-store	75
2 Foot Locker	Men's Athletic Footwear	50	In-store	75
3 Foot Locker	Women's Street Footwear	40	In-store	60

In [39]:

```
# Now show the total sales with group on Sales Method based on Retailer == "Foot Locker"
df%>%
filter(Retailer == "Foot Locker")%>%
group_by(Sales.Method)%>%
summarize(sum(Total.Sales))
```

A tibble: 3 × 2

Sales.Method sum(Total.Sales)

<chr>	<int>
In-store	7652500
Online	7294363
Outlet	7011665

ggplot

In [40]:

```
# ggplot is a popular data visualization package in R programming.

#install.packages("ggplot2")      install
#library("ggplot2")              Loud

# the formal the draw visualization / ggplot(dataframe, aes(x="column name",y="column n
```

In [41]:

```
install.packages("ggplot2")
```

Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

Warning message:
“unable to access index for repository http://cran.rstudio.com/src/contrib:
 cannot open URL ‘http://cran.rstudio.com/src/contrib/PACKAGES’”

Warning message:
“package ‘ggplot2’ is not available for this version of R

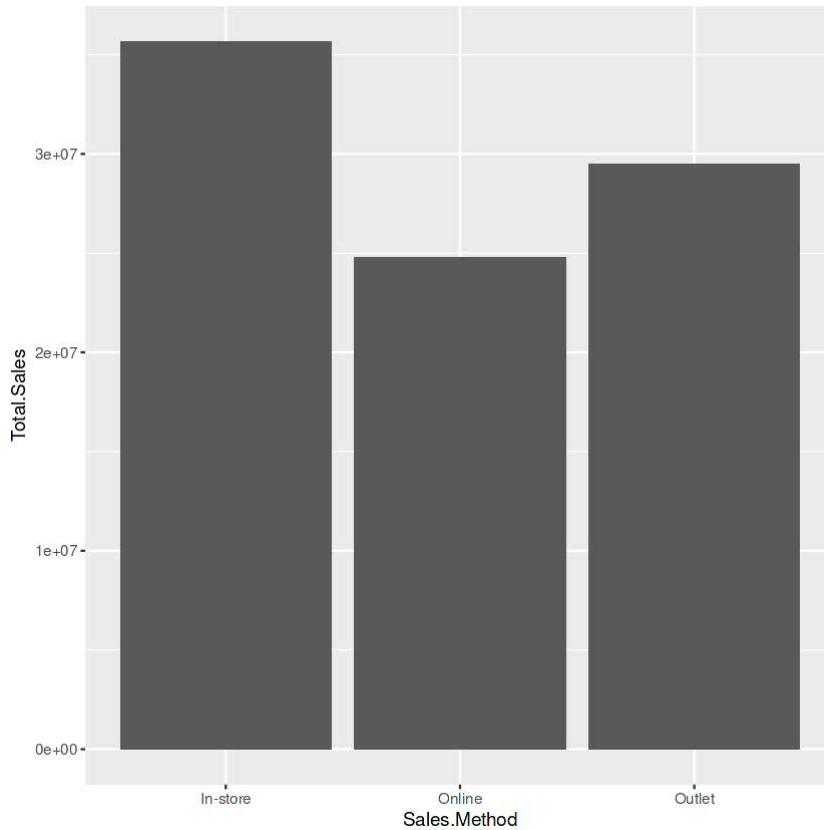
A version of this package for your version of R might be available elsewhere,
see the ideas at
<https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages>”

In [42]:

```
library("ggplot2")
```

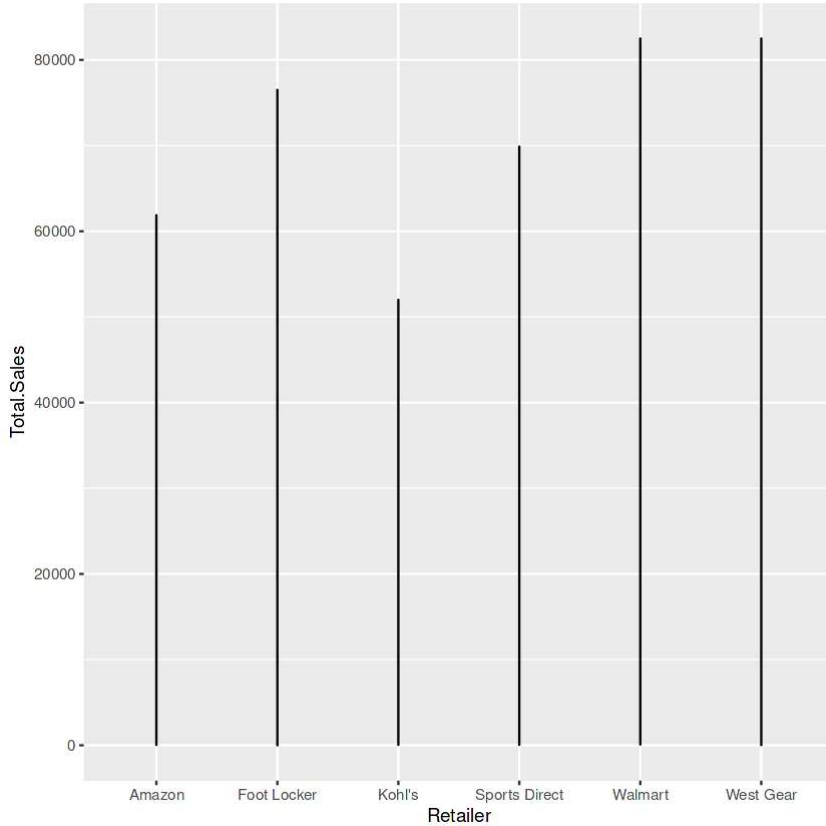
In [43]:

```
#draw column Sales Method with Total sales / basic visualization  
df%>%  
ggplot(aes(x=Sales.Method,y= Total.Sales))+geom_bar(stat="identity")
```

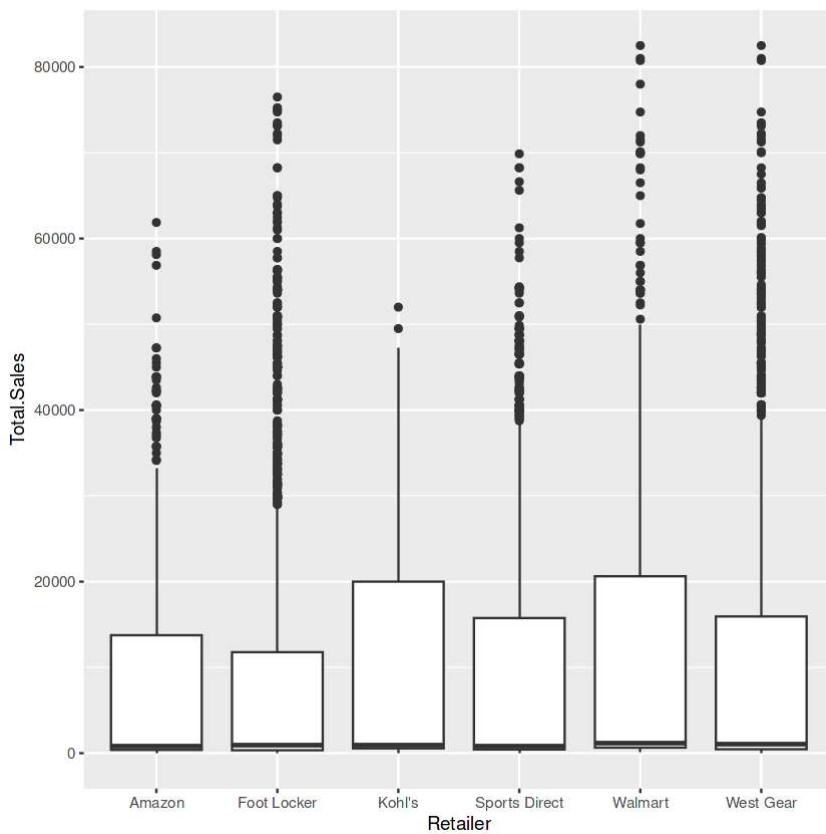


In [44]:

```
#draw column Retailer with Total sales / basic visualization  
df%>%  
ggplot(aes(x=Retailer,y= Total.Sales))+geom_line(stat="identity")
```

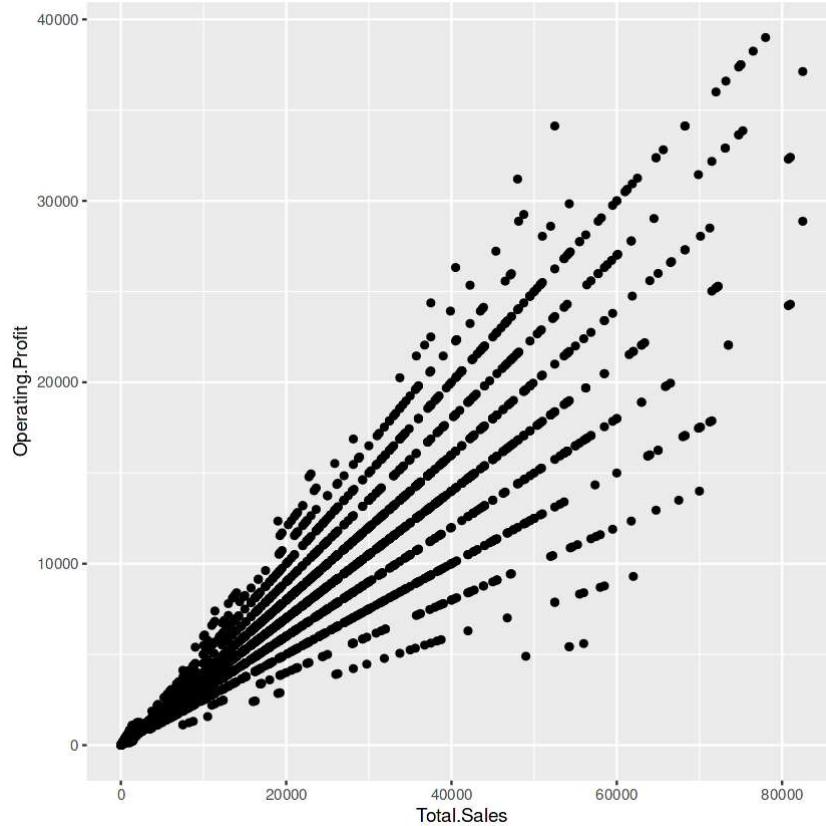


```
In [45]: #draw column Retailer with Total sales / basic visualization  
df%>%  
ggplot(aes(x=Retailer,y= Total.Sales))+geom_boxplot()
```



In [46]:

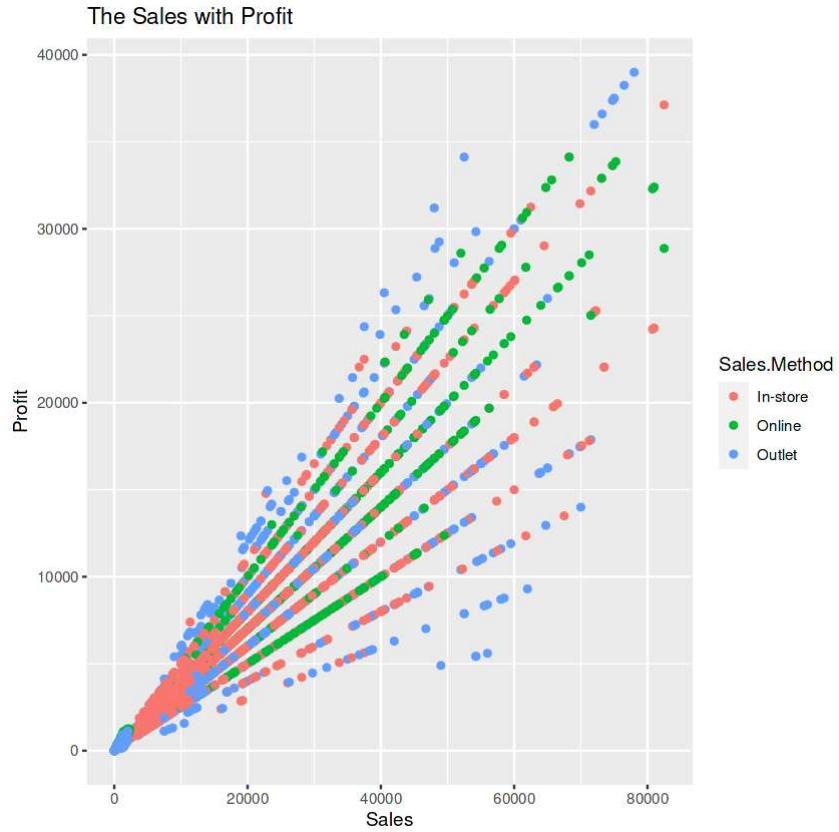
```
df%>%
ggplot(aes(x=Total.Sales,y= Operating.Profit))+geom_point()
```



In [47]:

```
# now using color in this graph / you can replace Sales method to any color you want
#Labs if you want add title and x =the name x Label and the same in y

df%>%
ggplot(aes(x=Total.Sales,y= Operating.Profit, color=Sales.Method))+geom_point()+labs(ti
```



End the first version

If you have any questions, feel free to contact me.

Qusay AL-Btoush

<https://www.kaggle.com/quasybtoush1990>

<https://github.com/quasybtoush>

<https://www.linkedin.com/in/quusal-al-btoush/>