R → Programming Language
Rstudio → IDE (Integrated Development environment)
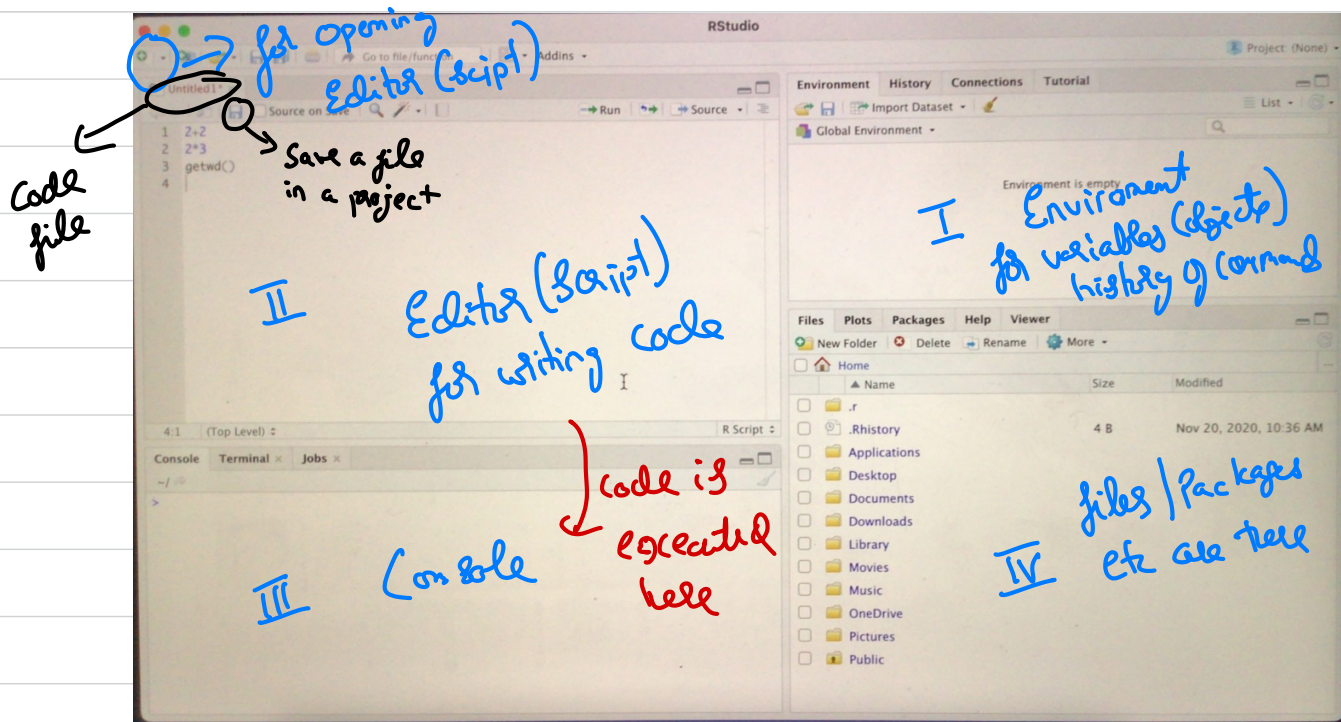
Create a project ⇒ Anything you do gets saved in one place
↳ File ⇒ New project ⇒ new directory ⇒ new project ⇒ directory name

\# helps in commenting

————————————→ Reading right to left.

ⓐ = 5
↓
variable/object

Command line Execution → Each & Every line came be executed Seperately

Declare Variable ⇒

Function writing :

Variable = function (x,y){
        result = x*y
        print (result)
}
Multiplier (4,6)

Result = 24

Data Types :→ (Nominal , ordinal , Internal & ratio) →Stats

for System data types ⇒ (Numeric, character, logical,
                                    Date, vector).

Numeric ← Integer (whole Number) (1,2,3 - -)
          Float (Decimal) (1.2, 3.2, - -)

for Specifying Integer numeric add L as suffix.

a = (5L) ⟶ integer.

Character ← words/strings (Nominal)
            Classification (Gender - Male, female) (ordinal)
            level of classification

Level of classification - facts - - - - Involves level

Example —> Education Qualification => 10th, 12th, Graduation, Post Graduation (4 levels)

logical ──> True (1)
         ──> False (0)

Date
  Starting date (1970) - Numeric Value
  Date - mm/dd/yyyy
  POSIXct — Date plus time

date = as.Date("2012-06-28")

date = as.POSIXct("2012-06-28 17:42")

Vector
  → Convert everything into a matrix language.
  → Array & Matrix
       ↓   ↳ 2 dimension Array
   Collection y   ↓
   similar elemts   R X C
                              (Rows X Comnlums)
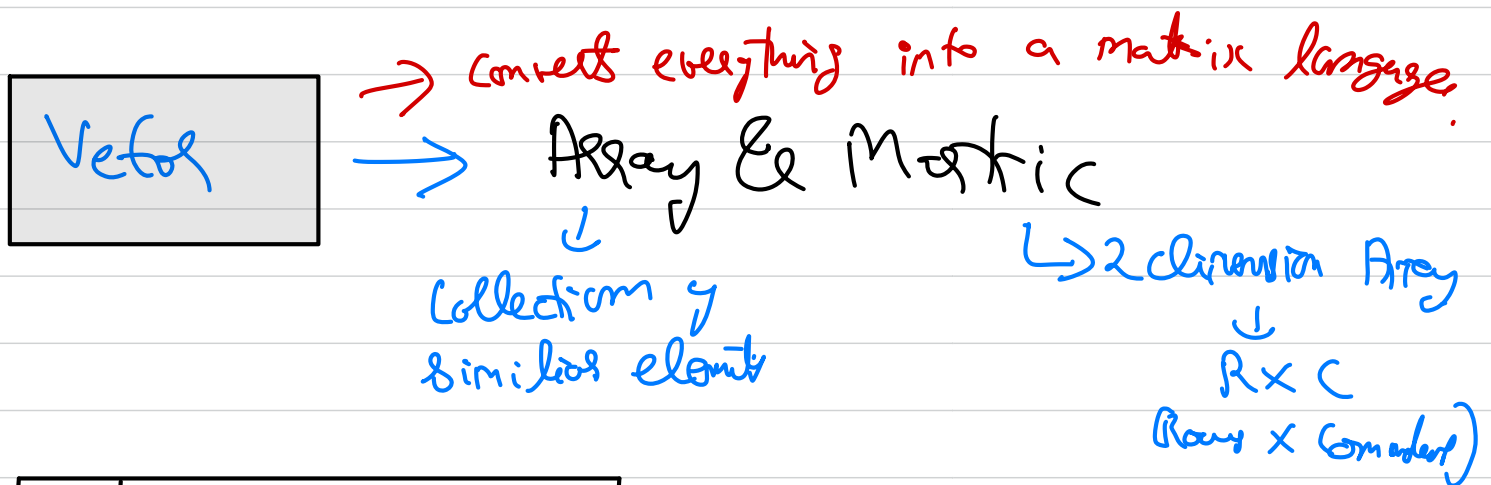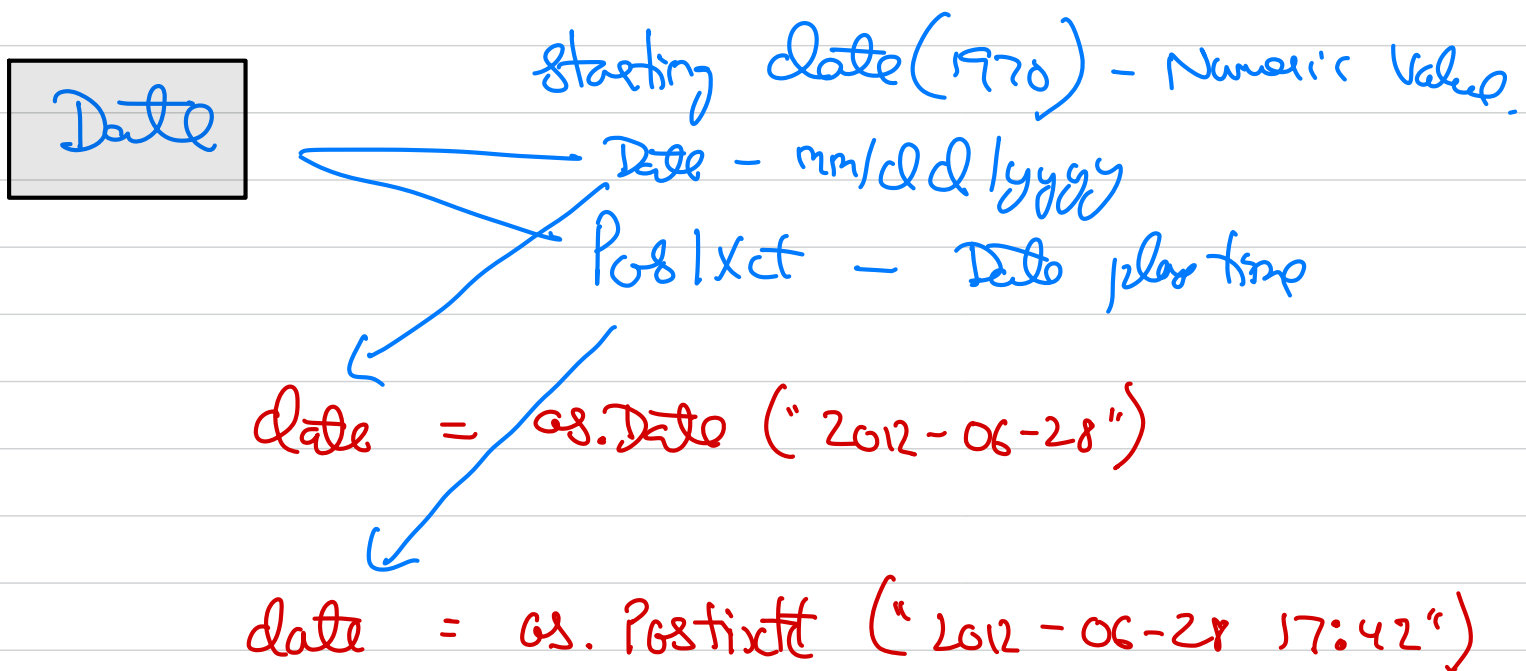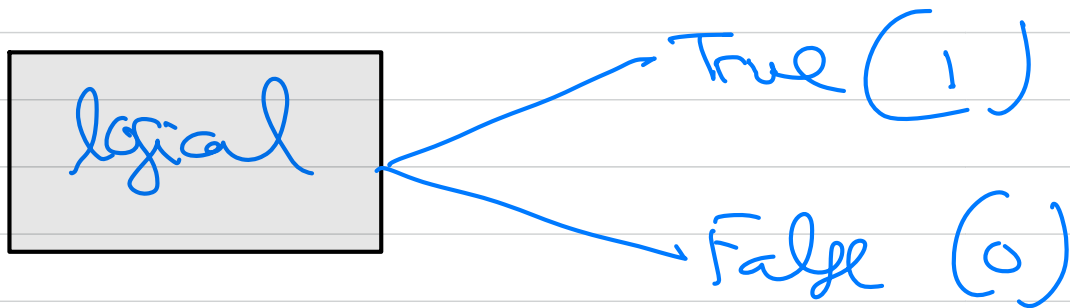
	v_1	v_2	v_3
1			
2			
3			
4			

$v = c(1,2,3,4,5)$

$S = v * 2$

S

$\hookrightarrow$ Result 2 4 6 8 10

A vector is collection of elements, all of same type.
A vector cannot be of mixed types.

Vector Operation. $(+, -, *, /)$

<div style="border:2px solid red; background:#fdd; padding:4px;">
Vector data should be of one type
</div>

R is also known as vectorised language

Vector - $a = c(1,2,3,4,5,6)$

$a = 1:10 \rightarrow$ | colon Command |

$b = 5:15$

$a+b$

Output = 6 8 10 12 14 - - - - -

Smaller vector will repeat

1:10 $\rightarrow$

$\leftarrow 3:4$

1	2	3	4	5	6	7	8	9	10
3	4	5	3	4	5	3	4	5	3

Output

4	6	8	7	9	11	10	14	14	13

$$a = 1:10$$
$$b = 10:20$$

Vector of Character

$$a = c("a", "aa", "aaa", "aaaa")$$

i) chars (a)

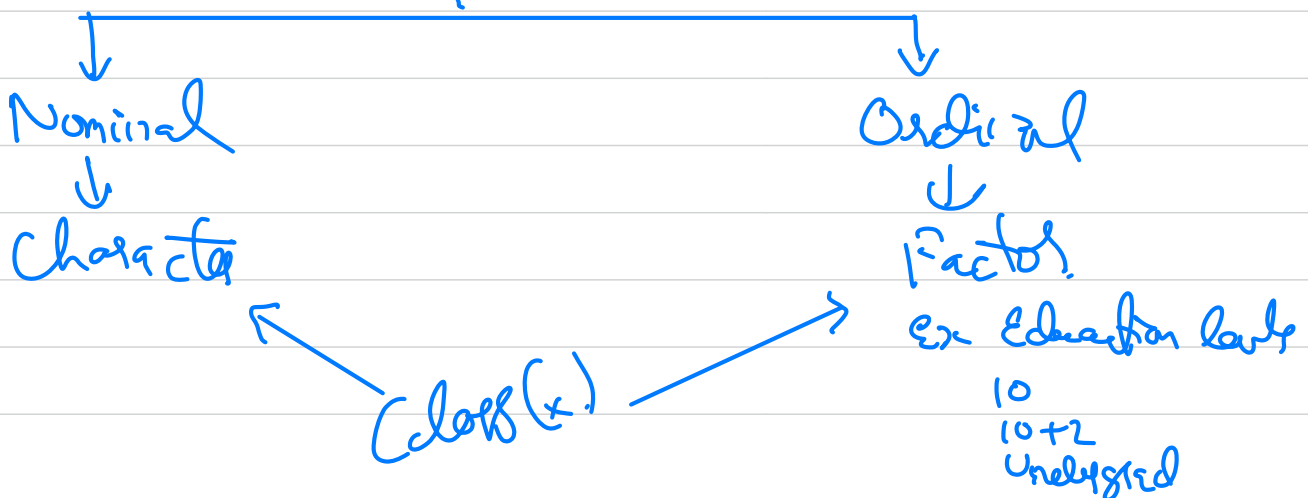## Subscripting

Accessing "individual elements"

$$a = 1:10$$
$$a[c(1:5, 9)]$$

Output = 1 2 3 4 5 9

Subscripting eg
big bracket.

Naming vects (Dictionary).

$$c(one = "a", Two = "y", Last = "r")$$

String.

Nominal → Character

Ordinal → Factor
Ex Education level
10
10+2
Unelysred

class (x)

Facts:

$q = c($"a", "aa", "aaa", "aaaa"$)$

$q_2 = c(q,$ "aa", "aaaa"$) \Rightarrow$ To add in a vector we vector and add element to vect

$q_2 = $ as.factor (q_2)

Missing Values:

NA $\Rightarrow$ Missing value

NULL $\Rightarrow$ Absence of anything. It is not exactly missing value, but nothingness.

Data structures — How a language store data

① Vector — Already discussed.

② Data frames $\rightarrow$ Excel Spreadsheet $\rightarrow$ Combination of different vectors

$R \times C$

	V_1	V_2	V_3
V_1			
V_5			
V_6			
V_7			

③ Matrix → Involves Vect as well as Data frames. → 2 Dimensions only.

④ Arrays → Array is n dimensional data set

In a way Matrix is a subset of Array.

⑤ List → It can store different type of data even a vector, other list ect.

Data frames:

Creating a data frame:

$X = 10:1$ (vector 1)
$Y = -4:5$ (vector 2)
$q = c('a', 'aa', 'aaa')$ (vector 3)

theDF = data.frame (X, Y, q)

Str (theDF) → Structure.
q = as.factor (q)

Variable names
→ column names
theDF = data.frames (First = X, Second = Y, third = q)

Accessing a specific Column / Row / data

Column => DF # Sport => Column. by name.

DF [3, 2] => 3rd row & 2nd Column

DF [3, 2:3] => " Through

see

in R files of sir

Matrix

A = matrix (1:10, nrow = 5)

Output

1	6
2	7
3	8
4	9
5	10

Gives Multiplication, addition ---- operating Calcula

$A \%*\% + (B.)$ → Transpose.

Multiplication like a Matrix.

R Markdown

File → New file → R mark down.

``` => grave Accent

Programs run
in markdown can
be stored in word
html & PDF.

# How to read data in R.

- Reading Comma Seperated Values (CSVs)

data set in UCi.edu → clean data
real world
data (Not
Simulate, Not
manupilated)

Jareddatadata.com

Creating a variable and give the entire path
in the

theURL = "htt://www - - - - - .CSV" ← first row contain header or not

tomatos = read.table (file =the url, header = TRUE,
                                    sep = ",")
                           ↳ separated by.

head (tomato)

Text file ⇒ Separator (Sep) becomes Space
                    i.e.  " "  .

R reserved backward slash (\) for
specific function but the directory
have backward slash so you use.
double back slash or single forward
slash.

read.table (file.choose()) ⇒ It lets you
                                          select the file
will always              No need for directory.
have header ="True."

# R & Packages:

R $\Rightarrow$ basic functions

R studio $\Rightarrow$ Histogram, Simple linear Regression
$\downarrow$
Special package

- Install Package
- Load the Package on Working Environment.

from R project.org.

Every package comes with its own dataset to explain the features of the package.

R also have built in data sets.
just type data() in R console.
$\downarrow$
Name the data set by
after seeing data() in the window
opened.

If the data set contain Null value

mean (y, na.rm = True) $\Rightarrow$ Remove the Null
value to find mean

Istall . package ("Name").

Heat map => (dealation → Search on google.

T-test

t.test (coloma nane  ⟶  check Syntax later