

## \* Date & time

( Lubridate ) ( ~~tidyverse~~ )

Lubridate is a package  
Can convert into date & time format.

- install.packages("tidyverse")
- install.packages("lubridate")

ydm - hms ( Date )



## \* Data types in R \*

i) Vector  
ii) list  
iii) Matrices  
iv) Arrays  
v) Factors  
vi) Data frames

## \* tidyverse package \*

- ggplot2 - graphics
- dplyr - Data manipulations
- tidyr - data Cleaning
- readr - importing data
- purrr - Work with functions & vectors
- tibble - with for data frame
- stringr - "in your tools"
- forcats - provides ~~tools~~ that solve common problems with factors.

Factors  $\Rightarrow$  Store Categorical data in R where the data values are limited usually based on a finite group like Country or year.

( \* Pivot-Longer & pivot-Wide )

file file file file  
row col

# R

## \* Most Data

### \* Vectors & lists

- (most common data structures)
- \* Vectors
  - \* Data frames
  - \* Matrices
  - \* Arrays

### \* Primary Data type

1 Logical	true/false
2 Integers	+ - whole value
3 Double or float	32.4
4 Characters	"Delhi"

### \* Vector

Vector is by using c() (Called "Combine" function)

### \* List

#### list()

list(list(list(1, 2, 3)))

Explain  
\* Logical operators

- AND
- OR
- NOT

### Conditional Statements

- if()
  - else()
  - elseif()
- IF  
ELSE  
ELSE IF

yet another markup language (YAML)

# \*R\*

## • operators

- i) Arithmetic operators + - \* / %
- ii) Relational operators == != < > <= >=
- iii) Logical operators TRUE FALSE
- iv) Assignment operators = <-
- v) Miscellaneous operators

## Q12 • How to add a new column on data?

~~use (mutate) function to do it~~

\* mutate(mutate(diamonds, carat\_2 = condition))

exep:

mutate(diamonds, carat\_2 = carat \* 100)

## Q22 • How to create a table in R?

~~use tibble~~

\* tibble(Colname = c(1, 2, 3),  
Colname = c("a", "b", "c"),  
Colname = c('alpha', 'beta', 'gamma'))

(tip also can use variable for table.)

## \* Basic Concepts of R \*

- Function:  
A body of reusable code for performing specific task in R
- Argument:  
Transformation (Information) needed by a function in R inorder to run.
- Comment:  
Helpful text that describes or explains R Code, preceded by #
- Variable:  
Value that value which is can be stored for later use
- Data types:  
An attribute that describe a piece of data based on its value
- Vector:  
A group of data elements of the same type stored in a one-dimensional sequence in R
- Pipe:  
A tool in R for expressing a sequence of multiple operations, represented by %>%  
(Pipe shortcut key Ctrl + shift + m %>%)

## Convert Data frame to a tibble

i) Create a table in R  
tibble (colname = c(---),  
colname = c(---),  
colname = c(---))

ii) Convert data frame to a tibble  
as\_tibble (data frame)

iii) Convert tibble to a data frame  
as\_data\_frame (tibble)

## ★ Cleanup Columns ★

:- Some packages  
- here  
- skimr  
- Janitor

ir install.packages ("here")  
library ("here")

ii) install.packages ("skimr")  
library ("skimr")

iii) install.packages ("janitor")  
library ("janitor")

# ~~Summaries~~

## (Summaries about dataframes)

- skim - without charts (- )
  - glimpse. (- )
  - head (- )
  - select (- )

(Change column name)

- rename
    - $\therefore \text{rename } (\text{DF}, \text{new col name} = \text{col name})$
    - or
    - $\text{DF pipe } (\%>%)$
    - Select
    - $\text{Select } (\text{new name} = \text{name})$

- rename\_with (Upper & Lower)



## ( Data transform )

- separate()
  - unite()
  - rotate()

$\therefore \text{Separate}(\text{Datalogname}, \text{Colname}, \text{into} = C('1^{\text{st}} \text{ Colname}', '2^{\text{nd}} \text{ Colname}'), \text{Sep} = '-')$

# 'R' R Programming

- 1) Variable
- 2) Data types
- 3) Converting data types
- 4) Operators
- 5) If-else statement (if, else, else if)
- 6) Switch statement
- 7) Next & Break

## "User input in R"

```
name <- readline(prompt = "Your name?")
```

```
age <- readline(prompt = "Your age?")
```

```
print(paste("Your name is", name, "your age is", age))
```

Paste key combine the strings or value while you print

Sep = Sep key separate the values and user regular.

paste0 = This key remove space for words

(  
  tolower ( )  
  toupper ( )  
)

## (Handling data in 'R')

- arrange()
- group\_by()
- filter()

### Categorized data

#### Clean

- 1) ~~rename~~
- 2) `grename_with()`
- 3) `Select()`
- 3) `glimpse()`
- 4) `renamel()`
- 5) `clean_names()`
- 6) `skim_without_charts()`

#### Organize

- 1) `groupby()`
- 2) `arrange()`
- 3) `Max()`
- 4) `summarize()`
- 5) `mean()`
- 6) `filter()`
- 7) `drop_na()`

#### Transform

- separate()
- write()
- mutate()

# "R Visualization"

gg (grammar of graphics)

1: ggplot2

2: Plotly

3: Lattice

4: RGL

5: Dygraphs

6: Leaflet

7: Highcharter

8: Patchwork

9:

10: gganimate

11: ggiraph

## ★ Core concepts in ggplot2 ★

1: Aesthetics : A visual property of an object in your plot.  
(size, shape & color in your data points)

2: Geoms : The geometric object used to represent your data

3: Facets : Let you display smaller groups, or subsets, of your data.

4: Labels & annotations : Let you customize your plot

## (Count Value from vector or table)

Count count = 0

```
for(i in 1:nrow(table_name)){  
  if(table$colname[i] == 'Variable') {  
    Count = Count + 1  
  }  
}
```

Count



# (SQL)

SQL

1) Insert value in table

Insert into student (SId, Sname, age, subject) value (1, amit, 23, 'java');

DDL

1) Create

2) Alter

3) Drop

4) Truncate

5) Rename

DML

Insert

Update

Delete

Lock

TCL

Commit

Save point

Roll back

DEEL

Select

DCL

Grant

Revoke

2) Where

Where clause use for 3 conditions

1) Select

2) Delete

3) Update

ex: Select \* from table name  
where condition

update Table set column = ' '   
where condition

delete table where condition.

3) SQL Database operators

- Arithmetic operators (+ - x ÷)
- Logical operators ( AND, OR, NOT )
- Concatenate operator (||) update
- Relation operator ( $\geq$ ,  $\leq$ ,  $=\geq$ )

4) How to comment (/\* --)

~~Z~~ 17 What are the different data base management system?

- Distributed database : Amazon Redshift, Teradata
- NoSQL Database : MongoDB, cassandra
- Graph database : Amazon Neptune
- Cloud : Amazon webservices, Azure, oracle cloud

27 What is Entity Integrity?

It ensure Table must have column or set of columns which uniquely identify a row.

32 What is Domain integrity?

It ensure the column values should be defined.

41 What are constraints in SQL?

Rules concerning data.

- \* not null
- \* check
- \* default
- \* unique
- \* P key
- \* F. key.

51 P key?

This constraint uniquely identifies each row in table

Property: Unique  
              : not null

62 Types of anomalies;

- Insert anomalies
- update anomalies
- delete anomalies

71 Scalar & aggregate function?

agg = max(), min(), AVG(), count()

Scalar = upper(), length(), floor, ceil()