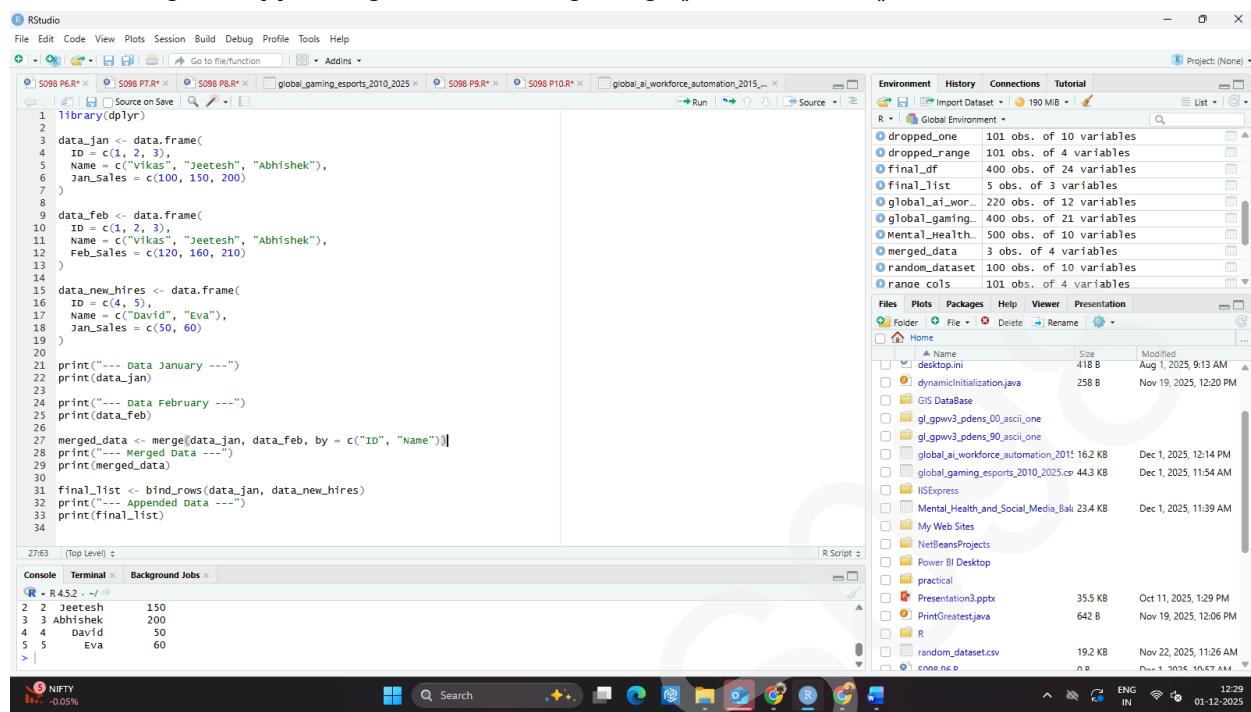


# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 6. Combining and appending datasets using merge() or bind\_rows() in R.



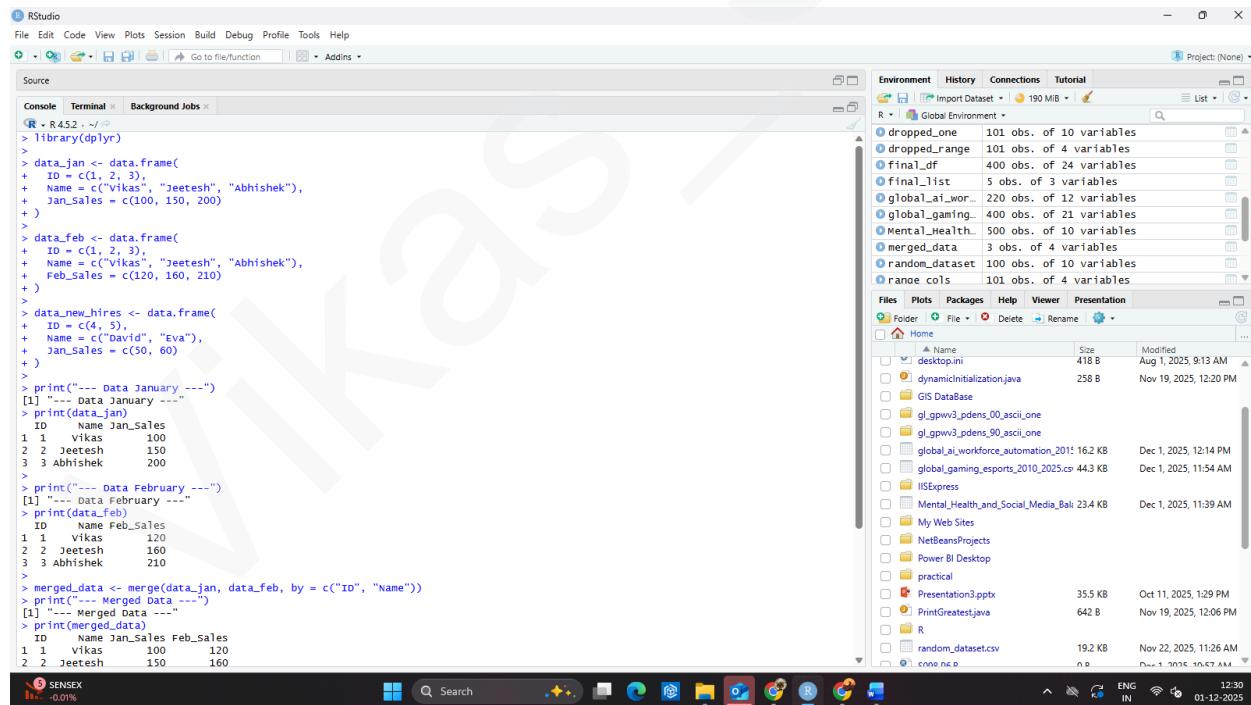
```

library(dplyr)
data_jan <- data.frame(
  ID = c(1, 2, 3),
  Name = c("Vikas", "Jeetesh", "Abhishek"),
  Jan_sales = c(100, 150, 200)
)
data_feb <- data.frame(
  ID = c(1, 2, 3),
  Name = c("Vikas", "Jeetesh", "Abhishek"),
  Feb_sales = c(120, 160, 210)
)
data_new_hires <- data.frame(
  ID = c(4, 5),
  Name = c("David", "Eva"),
  Jan_sales = c(50, 60)
)
print("--- Data January ---")
print(data_jan)
print("--- Data February ---")
print(data_feb)
merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
print("--- Merged Data ---")
print(merged_data)
final_list <- bind_rows(data_jan, data_new_hires)
print("--- Appended Data ---")
print(final_list)
  
```

Console output:

```

2 2 Jeetesh 150
3 3 Abhishek 200
4 4 David 50
5 5 Eva 60
  
```



```

library(dplyr)
data_jan <- data.frame(
  ID = c(1, 2, 3),
  Name = c("Vikas", "Jeetesh", "Abhishek"),
  Jan_sales = c(100, 150, 200)
)
data_feb <- data.frame(
  ID = c(1, 2, 3),
  Name = c("Vikas", "Jeetesh", "Abhishek"),
  Feb_sales = c(120, 160, 210)
)
data_new_hires <- data.frame(
  ID = c(4, 5),
  Name = c("David", "Eva"),
  Jan_sales = c(50, 60)
)
print("--- Data January ---")
[1] "--- Data January ---"
> print(data_jan)
  ID Name Jan_Sales
1 1 Vikas 100
2 2 Jeetesh 150
3 3 Abhishek 200
> print("--- Data February ---")
[1] "--- Data February ---"
> print(data_feb)
  ID Name Feb_Sales
1 1 Vikas 120
2 2 Jeetesh 160
3 3 Abhishek 210
> merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
> print("--- Merged Data ---")
[1] "--- Merged Data ---"
> print(merged_data)
  ID Name Jan_Sales Feb_Sales
1 1 Vikas 100 120
2 2 Jeetesh 150 160
  
```

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

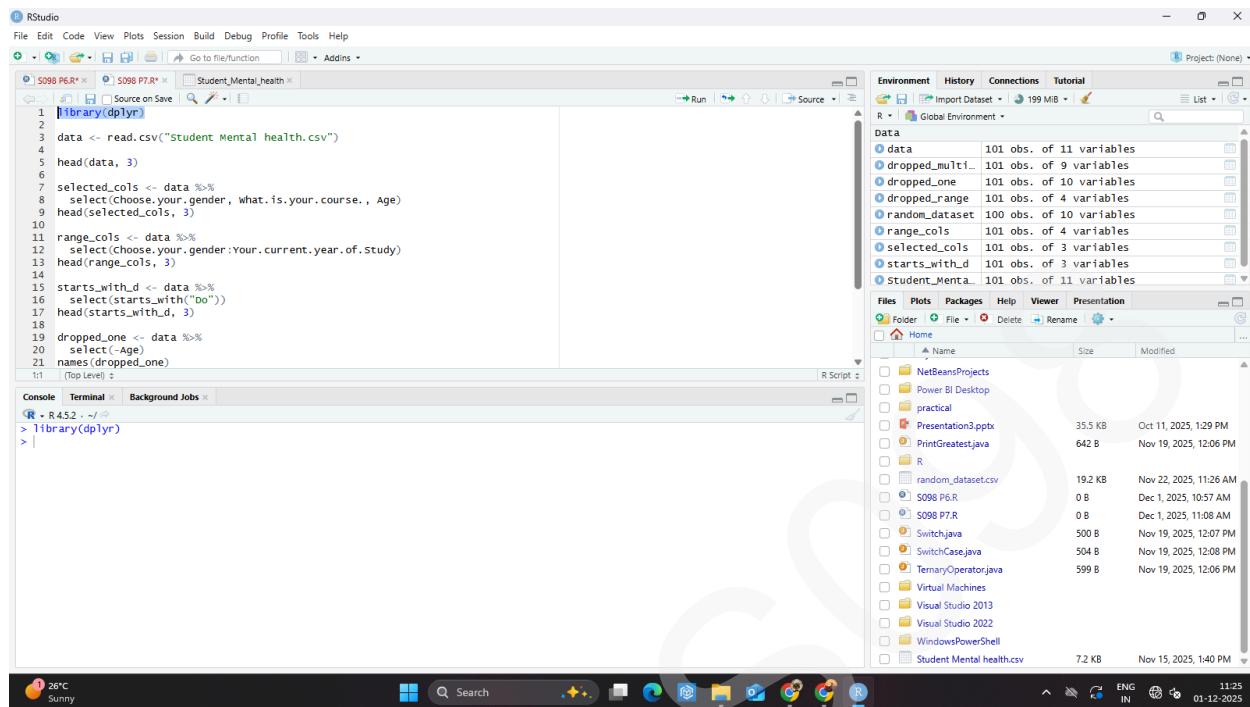
The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays R code for merging datasets. The code creates a new dataset `data\_new\_hires` from scratch, then prints `data\_january` and `data\_february` datasets, merges them into `merged\_data`, and finally appends `data\_new\_hires` to `merged\_data` to create `final\_list`.
- Environment Tab:** Shows the global environment with various objects: `dropped\_one` (101 obs. of 10 variables), `dropped\_range` (101 obs. of 4 variables), `final\_df` (400 obs. of 24 variables), `final\_list` (5 obs. of 3 variables), `global\_ali\_wor\_` (220 obs. of 12 variables), `global\_gaming\_` (400 obs. of 21 variables), `Mental\_Health` (500 obs. of 10 variables), `merged\_data` (3 obs. of 4 variables), `random\_dataset` (100 obs. of 10 variables), and `range\_cols` (101 obs. of 4 variables).
- File Explorer:** Shows a folder structure under "Home" containing files like `desktop.ini`, `dynamicInitialization.java`, `GIS DataBase`, `gl\_gpvw2\_pdens\_00\_ascii.one`, `gl\_gpvw2\_pdens\_90\_ascii.one`, `global\_ai\_workforce\_automation\_2011\_2012\_2013.csv`, `global\_gaming\_esports\_2010\_2012\_2013.csv`, `IISExpress`, `Mental\_Health\_and\_Social\_Media\_Bali`, `Presentation3.pptx`, `PrintGreatest.java`, `R`, `random\_dataset.csv`, and `cnn\_nic.o`.
- Taskbar:** Shows standard Windows taskbar icons for File Explorer, Search, Task View, Start, Taskbar settings, and system status (ENG IN, 01-12-2025, 12:30).

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 7. Selecting and dropping variables using select() in R. import dataset.



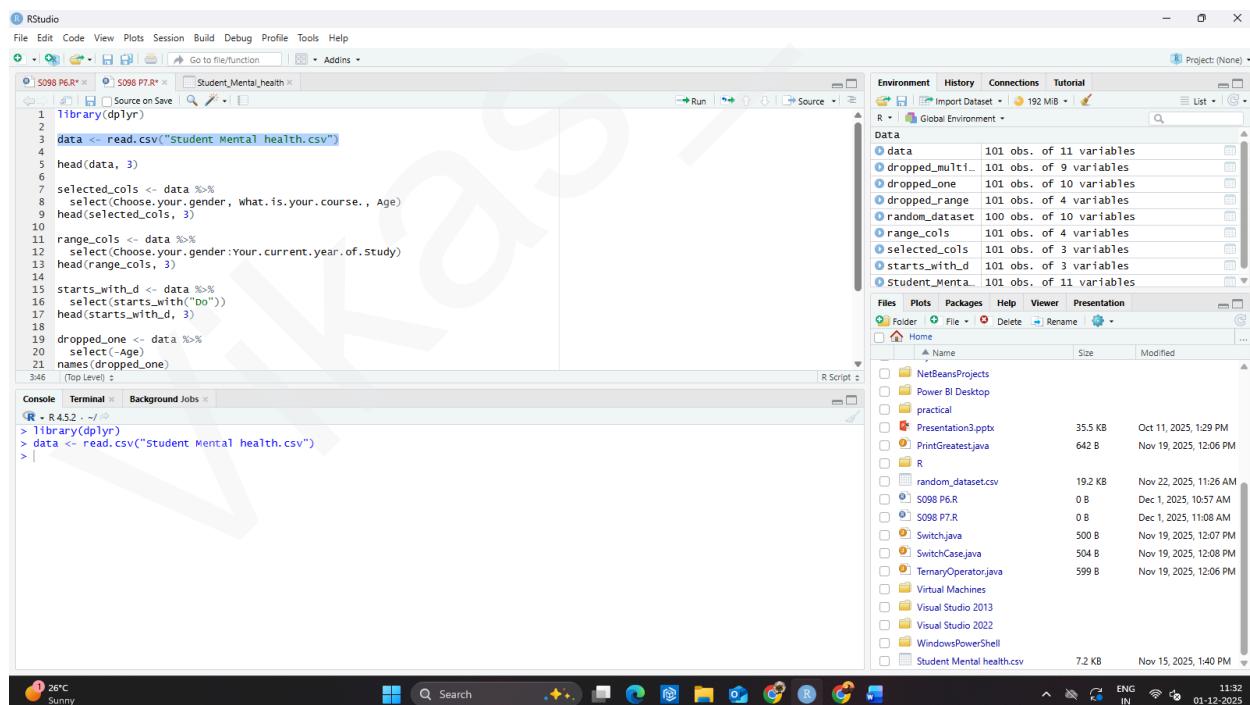
The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project: (None)
- Code Editor:** An R script titled "S098 P7.R" containing the following code:

```

1 library(dplyr)
2
3 data <- read.csv("Student_Mental_health.csv")
4
5 head(data, 3)
6
7 selected_cols <- data %>%
8   select(choose.your.gender, what.is.your.course., Age)
9 head(selected_cols, 3)
10
11 range_cols <- data %>%
12   select(choose.your.gender:Your.current.year.of.Study)
13 head(range_cols, 3)
14
15 starts_with_d <- data %>%
16   select(starts_with("do"))
17 head(starts_with_d, 3)
18
19 dropped_one <- data %>%
20   select(-Age)
21 names(dropped_one)

```
- Console:** Shows the command `> library(dplyr)` and the output of the R code.
- Environment:** Shows the global environment with objects like `data`, `dropped\_multi\_`, etc.
- Files:** Shows a file tree with various files including "Presentation3.pptx", "PrintGreatest.java", "random\_dataset.csv", "S098 P6.R", "S098 P7.R", "Switch.java", "SwitchCase.java", "TernaryOperator.java", "Virtual Machines", "Visual Studio 2013", "Visual Studio 2022", "WindowsPowerShell", and "Student Mental health.csv".
- System:** Shows the desktop environment with icons for various applications like NetBeans, Power BI Desktop, practical, etc.



The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project: (None)
- Code Editor:** An R script titled "S098 P7.R" containing the same R code as the first screenshot.
- Console:** Shows the command `> library(dplyr)` and the output of the R code.
- Environment:** Shows the global environment with objects like `data`, `dropped\_multi\_`, etc.
- Files:** Shows a file tree with various files including "Presentation3.pptx", "PrintGreatest.java", "random\_dataset.csv", "S098 P6.R", "S098 P7.R", "Switch.java", "SwitchCase.java", "TernaryOperator.java", "Virtual Machines", "Visual Studio 2013", "Visual Studio 2022", "WindowsPowerShell", and "Student Mental health.csv".
- System:** Shows the desktop environment with icons for various applications like NetBeans, Power BI Desktop, practical, etc.

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio interface showing R script and environment.

```

library(dplyr)
data <- read.csv("Student_Mental_health.csv")
head(data, 3)
selected_cols <- data %>%
  select(choose.your.gender, what.is.your.course., Age)
head(selected_cols, 3)
range_cols <- data %>%
  select(choose.your.gender:your.current.year.of.Study)
head(range_cols, 3)
starts_with_d <- data %>%
  select(starts_with("do"))
head(starts_with_d, 3)
dropped_one <- data %>%
  select(-Age)
names(dropped_one)

```

Console output:

```

> library(dplyr)
> data <- read.csv("student Mental health.csv")
> head(data, 3)
  Timestamp Choose.your.gender Age What.is.your.course. Your.current.year.of.Study what.is.your.CGPA. Marital.status
1 08-07-2020 12:02   Female 18      Engineering          year 1    3.00 - 3.49       No
2 08-07-2020 12:04     Male 21  Islamic education          year 2    3.00 - 3.49       No
3 08-07-2020 12:05     Male 19          BIT             Year 1    3.00 - 3.49       No
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack. Did.you.seek.any.specialist.for.a.treatment.
1 Yes           No        Yes           No
2 No            Yes       No           No
3 Yes           Yes      Yes           No

```

Environment pane shows variables and files:

- Data: data (101 obs. of 11 variables), dropped\_multi\_ (101 obs. of 9 variables), dropped\_one (101 obs. of 10 variables), dropped\_range (101 obs. of 4 variables), random\_dataset (100 obs. of 10 variables), range\_cols (101 obs. of 4 variables), selected\_cols (101 obs. of 3 variables), starts\_with\_d (101 obs. of 3 variables), Student\_Mental (101 obs. of 11 variables)
- Files: S098 P6.R, S098 P7.R, Switch.java, SwitchCase.java, TernaryOperator.java, random\_dataset.csv, S098 P6.R, S098 P7.R, PrintGreatest.java, Home

RStudio interface showing R script and environment.

```

library(dplyr)
data <- read.csv("Student_Mental_health.csv")
head(data, 3)
selected_cols <- data %>%
  select(choose.your.gender, what.is.your.course., Age)
head(selected_cols, 3)
range_cols <- data %>%
  select(choose.your.gender:your.current.year.of.Study)
head(range_cols, 3)
starts_with_d <- data %>%
  select(starts_with("do"))
head(starts_with_d, 3)
dropped_one <- data %>%
  select(-Age)
names(dropped_one)

```

Console output:

```

> library(dplyr)
> data <- read.csv("student Mental health.csv")
> head(data, 3)
  Timestamp Choose.your.gender Age What.is.your.course. Your.current.year.of.Study what.is.your.CGPA. Marital.status
1 08-07-2020 12:02   Female 18      Engineering          year 1    3.00 - 3.49       No
2 08-07-2020 12:04     Male 21  Islamic education          year 2    3.00 - 3.49       No
3 08-07-2020 12:05     Male 19          BIT             Year 1    3.00 - 3.49       No
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack. Did.you.seek.any.specialist.for.a.treatment.
1 Yes           No        Yes           No
2 No            Yes       No           No
3 Yes           Yes      Yes           No
> selected_cols <- data %>%
+   select(choose.your.gender, what.is.your.course., Age)
> |

```

Environment pane shows variables and files:

- Data: data (101 obs. of 11 variables), dropped\_multi\_ (101 obs. of 9 variables), dropped\_one (101 obs. of 10 variables), dropped\_range (101 obs. of 4 variables), random\_dataset (100 obs. of 10 variables), range\_cols (101 obs. of 4 variables), selected\_cols (101 obs. of 3 variables), starts\_with\_d (101 obs. of 3 variables), Student\_Mental (101 obs. of 11 variables)
- Files: S098 P6.R, S098 P7.R, Switch.java, SwitchCase.java, TernaryOperator.java, random\_dataset.csv, S098 P6.R, S098 P7.R, PrintGreatest.java, Home

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio interface showing R code in the script pane and its output in the console pane. The environment pane shows various objects and their details.

```

library(dplyr)
data <- read.csv("Student Mental health.csv")
head(data, 3)

selected_cols <- data %>%
  select(choose.your.gender, what.is.your.course., Age)
head(selected_cols, 3)

range_cols <- data %>%
  select(choose.your.gender:your.current.year.of.Study)
head(range_cols, 3)

starts_with_d <- data %>%
  select(starts_with("do"))
head(starts_with_d, 3)

dropped_one <- data %>%
  select(-Age)
names(dropped_one)

> |
```

```

[R - R4.5.2 - ~]
> library(dplyr)
> data <- read.csv("student Mental health.csv")
> head(data, 3)
  Timestamp Choose.your.gender Age What.is.your.course. Your.current.year.of.Study what.is.your.CGPA. Marital.status
1 08-07-2020 12:02   Female 18     Engineering      year 1    3.00 - 3.49       NO
2 08-07-2020 12:04   Male  21   Islamic education      year 2    3.00 - 3.49       NO
3 08-07-2020 12:05   Male  19        BIT          Year 1    3.00 - 3.49       NO
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack. Did.you.seek.any.specialist.for.a.treatment.
1 Yes           NO       Yes           NO           NO
2 No            Yes      No            No           NO
3 Yes           Yes      Yes           Yes          NO

> selected_cols <- data %>%
+   select(choose.your.gender, what.is.your.course., Age)
> head(selected_cols, 3)
  Choose.your.gender what.is.your.course. Age
1   Female           Engineering 18
2   Male             Islamic education 21
3   Male             BIT 19
> |
```

RStudio interface showing R code in the script pane and its output in the console pane. The environment pane shows various objects and their details.

```

library(dplyr)
data <- read.csv("Student Mental health.csv")
head(data, 3)

selected_cols <- data %>%
  select(choose.your.gender, what.is.your.course., Age)
head(selected_cols, 3)

range_cols <- data %>%
  select(choose.your.gender:your.current.year.of.Study)
head(range_cols, 3)

starts_with_d <- data %>%
  select(starts_with("do"))
head(starts_with_d, 3)

dropped_one <- data %>%
  select(-Age)
names(dropped_one)

> |
```

```

[R - R4.5.2 - ~]
> head(data, 3)
  Timestamp Choose.your.gender Age What.is.your.course. Your.current.year.of.Study what.is.your.CGPA. Marital.status
1 08-07-2020 12:02   Female 18     Engineering      year 1    3.00 - 3.49       NO
2 08-07-2020 12:04   Male  21   Islamic education      year 2    3.00 - 3.49       NO
3 08-07-2020 12:05   Male  19        BIT          Year 1    3.00 - 3.49       NO
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack. Did.you.seek.any.specialist.for.a.treatment.
1 Yes           NO       Yes           NO           NO
2 No            Yes      No            No           NO
3 Yes           Yes      Yes           Yes          NO

> selected_cols <- data %>%
+   select(choose.your.gender, what.is.your.course., Age)
> head(selected_cols, 3)
  Choose.your.gender what.is.your.course. Age
1   Female           Engineering 18
2   Male             Islamic education 21
3   Male             BIT 19
> range_cols <- data %>%
+   select(choose.your.gender:your.current.year.of.study)
> |
```

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows an RStudio interface with the following details:

- File Explorer:** Shows a project named "S098 PER".
- Code Editor:** An R script titled "Student\_Mental\_health.R" is open, containing code to read a CSV file, select specific columns, and filter data by gender and year.
- Console:** Displays the output of the R script, including the resulting data frames and their contents.
- Environment:** Shows the global environment with various objects and their characteristics.
- Plots:** No plots are currently displayed.
- Packages:** No packages are currently displayed.
- Help:** No help pages are currently displayed.
- Viewer:** No files are currently displayed.
- Presentation:** No presentations are currently displayed.

The screenshot shows an RStudio interface with several panes:

- Code Editor:** Displays R code for reading a CSV file, selecting specific columns, and filtering data based on gender and course.
- Environment:** Shows the global environment with various objects and their details.
- Console:** Displays the output of the R code, including the resulting data frame.
- File Bar:** Includes "File", "Edit", "Code", "View", "Plot", "Session", "Build", "Debug", "Profile", "Tools", "Help", and "Addins".
- System Tray:** Shows the date (Nov 15, 2025), time (1:40 PM), battery level (26°C), and network status (ENG IN).

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio Environment Screenshot:

```

library(dplyr)
data <- read.csv("Student Mental health.csv")
head(data, 3)
selected_cols <- data %>%
  select(choose.your.gender, what.is.your.course., Age)
head(selected_cols, 3)
range_cols <- data %>%
  select(choose.your.gender:Your.current.year.of.study)
head(range_cols, 3)
starts_with_d <- data %>%
  select(starts_with("Do"))
head(starts_with_d, 3)
dropped_one <- data %>%
  select(-Age)
names(dropped_one)

```

Console Output:

```

Choose.your.gender what.is.your.course. Age
1 Female Engineering 18
2 Male Islamic education 21
3 Male BIT 19
> range_cols <- data %>%
+ select(choose.your.gender:Your.current.year.of.study)
> head(range_cols, 3)
Choose.your.gender Age what.is.your.course. Your.current.year.of.study
1 Female 18 Engineering year 1
2 Male 21 Islamic education year 2
3 Male 19 BIT Year 1
> starts_with_d <- data %>%
+ select(starts_with("Do"))
> head(starts_with_d, 3)
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack.
1 Yes No Yes
2 No Yes No
3 Yes Yes Yes

```

File Explorer:

- Project: (None)
- data (101 obs. of 11 variables)
- dropped\_multi\_ (101 obs. of 9 variables)
- dropped\_one (101 obs. of 10 variables)
- dropped\_range (101 obs. of 4 variables)
- random\_dataset (100 obs. of 10 variables)
- range\_cols (101 obs. of 4 variables)
- selected\_cols (101 obs. of 3 variables)
- starts\_with\_d (101 obs. of 3 variables)
- Student\_Mental\_ (101 obs. of 11 variables)

File Explorer Details:

Name	Size	Modified
NetBeansProjects		
Power BI Desktop		
practical		
Presentation3.pptx	35.5 KB	Oct 11, 2025, 1:29 PM
PrintGreatest.java	642 B	Nov 19, 2025, 12:06 PM
R		
random_dataset.csv	19.2 KB	Nov 22, 2025, 11:26 AM
S098 P6.R	0 B	Dec 1, 2025, 10:57 AM
S098 P7.R	0 B	Dec 1, 2025, 11:08 AM
Switch.java	500 B	Nov 19, 2025, 12:07 PM
SwitchCase.java	504 B	Nov 19, 2025, 12:08 PM
TernaryOperator.java	599 B	Nov 19, 2025, 12:06 PM
Virtual Machines		
Visual Studio 2013		
Visual Studio 2022		
WindowsPowerShell		
Student Mental health.csv	7.2 KB	Nov 15, 2025, 1:40 PM

RStudio Environment Screenshot:

```

head(selected_cols, 3)
range_cols <- data %>%
  select(choose.your.gender:Your.current.year.of.study)
head(range_cols, 3)
starts_with_d <- data %>%
  select(starts_with("Do"))
head(starts_with_d, 3)
dropped_one <- data %>%
  select(-Age)
names(dropped_one)
dropped_multiple <- data %>%
  select(-(Do.you.have.Depression., -Do.you.have.Anxiety.))
names(dropped_multiple)
dropped_range <- data %>%
  select(-(what.is.your.course.:Do.you.have.Panic.attack.))
names(dropped_range)

```

Console Output:

```

Choose.your.gender what.is.your.course. Age
1 Female Engineering 18
2 Male Islamic education 21
3 Male BIT 19
> range_cols <- data %>%
+ select(choose.your.gender:Your.current.year.of.study)
> head(range_cols, 3)
Choose.your.gender Age what.is.your.course. Your.current.year.of.study
1 Female 18 Engineering year 1
2 Male 21 Islamic education year 2
3 Male 19 BIT Year 1
> starts_with_d <- data %>%
+ select(starts_with("Do"))
> head(starts_with_d, 3)
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack.
1 Yes No Yes
2 No Yes No
3 Yes Yes Yes

```

File Explorer:

- Project: (None)
- data (101 obs. of 11 variables)
- dropped\_multi\_ (101 obs. of 9 variables)
- dropped\_one (101 obs. of 10 variables)
- dropped\_range (101 obs. of 4 variables)
- random\_dataset (100 obs. of 10 variables)
- range\_cols (101 obs. of 4 variables)
- selected\_cols (101 obs. of 3 variables)
- starts\_with\_d (101 obs. of 3 variables)
- Student\_Mental\_ (101 obs. of 11 variables)

File Explorer Details:

Name	Size	Modified
NetBeansProjects		
Power BI Desktop		
practical		
Presentation3.pptx	35.5 KB	Oct 11, 2025, 1:29 PM
PrintGreatest.java	642 B	Nov 19, 2025, 12:06 PM
R		
random_dataset.csv	19.2 KB	Nov 22, 2025, 11:26 AM
S098 P6.R	0 B	Dec 1, 2025, 10:57 AM
S098 P7.R	0 B	Dec 1, 2025, 11:08 AM
Switch.java	500 B	Nov 19, 2025, 12:07 PM
SwitchCase.java	504 B	Nov 19, 2025, 12:08 PM
TernaryOperator.java	599 B	Nov 19, 2025, 12:06 PM
Virtual Machines		
Visual Studio 2013		
Visual Studio 2022		
WindowsPowerShell		
Student Mental health.csv	7.2 KB	Nov 15, 2025, 1:40 PM

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface with the following details:

- File:\*\* S098 P6.R\*\***: Contains R code for data analysis.
- Console Output:**

```
R > head(range_cols, 3)
#> #> choose.your.gender Age what.is.your.course Your.current.year.of.Study
#> #> Female 18 Engineering year 1
#> #> Male 21 Islamic education year 2
#> #> Male 19 BIT Year 1
#> starts_with_d <- data %>%
#>   select(starts_with("Do"))
#> head(starts_with_d, 3)
#>
#> dropped_one <- data %>%
#>   select(-Age)
#> names(dropped_one)
#>
#> dropped_multiple <- data %>%
#>   select(~do.you.have.Depression., -do.you.have.Anxiety.)
#> names(dropped_multiple)
#>
#> dropped_range <- data %>%
#>   select(~(what.is.your.course, do.you.have.Panic.attack.))
#> names(dropped_range)
```
- Output:**

```
Female 18 Engineering year 1
Male 21 Islamic education year 2
Male 19 BIT Year 1
[1] "choose.your.gender"
[2] "your.current.year.of.study"
[3] "Marital.status"
[4] "Do.you.have.Depression."
[5] "Do.you.have.Anxiety."
[6] "did.you.seek.any.specialist.for.a.treatment."
```
- Environment View:** Shows variables like data, dropped\_multil\_, dropped\_one, etc.
- File Explorer:** Lists files including S098 P6.R, S098 P7.R, and various Java files.
- System Tray:** Shows weather (26°C, Sunny), system icons, and a taskbar with various application icons.

The screenshot shows the RStudio interface with the following details:

- File:\*\* S098 P6.R\*\***: Contains R code for data analysis.
- Console Output:**

```
R > head(range_cols, 3)
#> #> choose.your.gender Age what.is.your.course Your.current.year.of.Study
#> #> Female 18 Engineering year 1
#> #> Male 21 Islamic education year 2
#> #> Male 19 BIT Year 1
#> starts_with_d <- data %>%
#>   select(starts_with("Do"))
#> head(starts_with_d, 3)
#>
#> dropped_one <- data %>%
#>   select(-Age)
#> names(dropped_one)
#>
#> dropped_multiple <- data %>%
#>   select(~do.you.have.Depression., -do.you.have.Anxiety.)
#> names(dropped_multiple)
#>
#> dropped_range <- data %>%
#>   select(~(what.is.your.course, do.you.have.Panic.attack.))
#> names(dropped_range)
```
- Output:**

```
Female 18 Engineering year 1
Male 21 Islamic education year 2
Male 19 BIT Year 1
[1] "choose.your.gender"
[2] "your.current.year.of.study"
[3] "Marital.status"
[4] "Do.you.have.Depression."
[5] "Do.you.have.Anxiety."
[6] "did.you.seek.any.specialist.for.a.treatment."
```
- Environment View:** Shows variables like data, dropped\_multil\_, dropped\_one, etc.
- File Explorer:** Lists files including S098 P6.R, S098 P7.R, and various Java files.
- System Tray:** Shows weather (26°C, Sunny), system icons, and a taskbar with various application icons.

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows the RStudio interface with two panes. The left pane displays R code and its execution results. The right pane shows the R environment, history, and connections.

**Code and Console Output:**

```
9 head(selected_cols, 3)
10 range_cols <- data %>
11   select(choose.your.gender:Your.current.year.of.Study)
12 head(range_cols, 3)
13 
14 starts_with_d <- data %>
15   select(starts_with("Do"))
16 head(starts_with_d, 3)
17 
18 dropped_one <- data %>
19   select(Age)
20 names(dropped_one)
21 
22 dropped_multiple <- data %>
23   select(-Do.you.have.Depression., -Do.you.have.Anxiety.)
24 names(dropped_multiple)
25 
26 dropped_range <- data %>
27   select(-(what.is.your.course.:Do.you.have.Panic.attack.))
28 names(dropped_range)
29 
30 #> [Top Level]:
```

**R Script:**

```
R 4.5.2 - ~/
```

```
Do.you.have.Depression. Do.you.have.Anxiety. Do.you.have.Panic.attack.
1 Yes No Yes
2 No Yes No
3 Yes Yes Yes

> names(dropped_one)
[1] "Timestamp"                               "Choose.your.gender"
[3] "what.is.your.course."                   "Your.current.year.of.study"
[5] "what.is.your.CGPA."                     "Marital.status"
[7] "Do.you.have.Depression."                "Do.you.have.Anxiety."
[9] "Do.you.have.Panic.attack."              "Did.you.seek.any.specialist.for.a.treatment."
> dropped_multiple <- data %>
+   select(-Do.you.have.Depression., -Do.you.have.Anxiety.)
> names(dropped_multiple)
[1] "Timestamp"                               "Choose.your.gender"
[3] "Age"                                     "what.is.your.course."
[5] "Your.current.year.of.Study"              "what.is.your.CGPA."
[7] "Marital.status"                          "Do.you.have.Panic.attack."
[9] "Did.you.seek.any.specialist.for.a.treatment."
> |
```

**Environment:**

Object	Type	Size	Modified
data	101 obs. of 11 variables		
dropped_multiple	101 obs. of 9 variables		
dropped_one	101 obs. of 10 variables		
dropped_range	101 obs. of 4 variables		
random_dataset	100 obs. of 10 variables		
range_cols	101 obs. of 4 variables		
selected_cols	101 obs. of 3 variables		
starts_with_d	101 obs. of 3 variables		
Student_Mental	101 obs. of 11 variables		

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface. In the top-left, there are two tabs: "S098 P6.R" and "S098 P7.R". The "S098 P7.R" tab is active, displaying R script code. The code performs several operations on a dataset named "Student\_Mental\_health": it creates a subset of columns, drops specific rows based on gender, selects rows where "Do you have Depression.", and so on. It also creates variables like "dropped\_one", "dropped\_multiple", and "dropped\_range". The "Console" tab at the bottom shows the output of these commands. In the top-right, the "Project" pane shows a list of files and datasets. The "Files" tab is selected, showing items like "data", "dropped\_multi\_1", etc. The "Data" tab lists datasets such as "data", "dropped\_multi\_1", "dropped\_one", etc. The "Global Environment" tab shows objects like "data", "dropped\_multi\_1", etc. At the bottom right, the system tray shows the date and time as Nov 15, 2025, 1:40 PM.

```
head(selected_cols, 3)
range_cols <- data %>%
  select(choose.your.gender:Your.current.year.of.Study)
head(range_cols, 3)

starts_with_d <- data %>%
  select(starts_with("Do"))
head(starts_with_d, 3)

dropped_one <- data %>%
  select(-Age)
names(dropped_one)

dropped_multiple <- data %>%
  select(~Do.you.have.Depression., -Do.you.have.Anxiety.)
names(dropped_multiple)

dropped_range <- data %>%
  select(-(what.is.your.course.:Do.you.have.Panic.attack.))
names(dropped_range)

[1] "Timestamp"                      "choose.your.gender"
[2] "What.is.your.course."           "Your.current.year.of.study"
[3] "What.is.your.CGPA."             "Marital.status"
[4] "Age"                            "Do.you.have.Anxiety."
[5] "Your.current.year.of.study"     "Did.you.seek.any.specialist.for.a.treatment."
[6] "Marital.status"                 "Do.you.have.Panic.attack."
[7] "Did.you.seek.any.specialist.for.a.treatment."
[8] "Timestamp"                      "choose.your.gender"
[9] "Age"                            "Did.you.seek.any.specialist.for.a.treatment."
> |
```

26°C  
Sunny

11:35  
01-12-2025

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 8. Applying basic data cleaning functions: handling missing values using na.omit()/replace\_na() in R. import dataset.

The screenshot shows the RStudio interface. The left pane contains an R script with the following code:

```

1 library(dplyr)
2 library(tidyverse)
3
4 data <- read.csv("Mental_Health_and_Social_Media_Balance_dataset.csv",
5   na.strings = c("", "NA"))
6
7 head(data)
8
9 colsum(is.na(data))
10 clean.omit <- na.omit(data)
11
12 print(nrow(data))
13 print(nrow(clean.omit))
14 head(clean.omit)
15
16 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
17
18 clean.replace <- data %>%
19   replace_na(list(
20     Age = mean(data$Age, na.rm = TRUE),
21
22   (Top Level) : 
  
```

The right pane shows the Global Environment and a file browser. The file browser lists various files including 'Mental\_Health\_and\_Social\_Media\_Bali' and 'Presentation3.pptx'. The status bar at the bottom indicates it's 11:45 AM on 01-12-2025.

This screenshot is identical to the one above, showing the RStudio interface with the same R script, file browser, and system status information.

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows the RStudio interface with several windows open:

- Code Editor:** Displays R code for reading a CSV file and performing initial data cleaning.
- Console:** Shows the execution of the R code, resulting in a data frame with columns: User\_ID, Age, Gender, Daily\_Screen\_Time.hrs, Sleep\_Quality.1..10, Stress\_Level.1..10, Days\_without\_Social\_Media, Exercise\_Frequency.week, Social\_Media, Reform\_Happiness\_Index.1..10, and Platform.
- Data View:** A tree view of the workspace showing various objects like data, dropped variables, and specific datasets.
- File Explorer:** Shows the project structure with files like .Rhistory, desktop.ini, and various CSV and Java files.
- Help:** A sidebar with links to Environment, History, Connections, and Tutorial.

The status bar at the bottom indicates the system temperature (26°C), battery level (Sunny), and system time (01-12-2025).

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project Bar:** S098 P6.R\*, S098 P7.R\*, S098 P8.R\*, Mental\_Health\_and\_Social\_Media\_Balance\_Dataset.csv
- Code Editor:** Displays R code for reading a CSV file, handling missing values, calculating average CGPA, and creating a new variable Age. The code is as follows:

```
1 library(dplyr)
2 library(tidyverse)
3 
4 data <- read.csv("Mental_Health_and_Social_Media_Balance_Dataset.csv",
5   na.strings = c("", "NA"))
6 
7 head(data)
8 
9 colSums(is.na(data))
10 
11 clean.omit <- na.omit(data)
12 
13 print(nrow(data))
14 print(nrow(clean.omit))
15 head(clean.omit)
16 
17 avg_cgpa <- mean(data$What_is_your.CGPA., na.rm = TRUE)
18 
19 clean.replace <- data %>%
20   replace_na(list(
21     Age = mean(data$Age, na.rm = TRUE),
22     ))
23 
```

- Console:** Shows the output of the R code, including the original data rows and the modified dataset 'clean.replace'.
- Environment:** Shows the global environment with objects like data, dropped\_multi, dropped\_one, etc.
- Files:** Shows the project structure with files like J.Rhistory, AI\_Impact\_on\_jobs\_2030.csv, and Mental\_Health\_and\_Social\_Media\_Balance\_Dataset.csv.

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio interface showing R code in the script pane and data analysis results in the console pane.

```

library(dplyr)
library(tidyr)
data <- read.csv("Mental_Health_and_Social_Media_Balance_Dataset.csv",
na.strings = c("", "NA"))
head(data)
colSums(is.na(data))
clean.omit <- na.omit(data)
print(nrow(data))
print(nrow(clean.omit))
head(clean.omit)
avg_cgpa <- mean(data$what.your.CGPA., na.rm = TRUE)
clean.replace <- data %>%
  replace_na(list(
    Age = mean(data$Age, na.rm = TRUE),
    Sleep_Quality = mean(data$Sleep_Quality, na.rm = TRUE),
    Stress_Level = mean(data$Stress_Level, na.rm = TRUE),
    Days_without_Social_Media = mean(data$Days_without_Social_Media, na.rm = TRUE),
    Exercise_Frequency = mean(data$Exercise_Frequency, na.rm = TRUE),
    Facebook = mean(data$Facebook, na.rm = TRUE),
    LinkedIn = mean(data$LinkedIn, na.rm = TRUE),
    YouTube = mean(data$YouTube, na.rm = TRUE),
    TikTok = mean(data$TikTok, na.rm = TRUE),
    Twitter = mean(data$Twitter, na.rm = TRUE),
    Happiness_Index = mean(data$Happiness_Index, na.rm = TRUE)))
colSums(is.na(data))
User_ID      Age   Gender Daily_Screen_Time.hrs.
Sleep_Quality 1.10. Stress_Level.1.10. Days_without_Social_Media Exercise_Frequency.week.
Social_Media_Platform Happiness_Index.1.10.
> clean.omit <- na.omit(data)

```

Environment pane shows various objects and files:

- Data: clean.omit, data, dropped\_multi\_, dropped\_one, dropped\_range, Mental\_Health\_, random\_dataset, range\_cols, selected\_cols
- Files: Home, Rhistory, AI\_Impact\_on\_labs\_2030.csv, Custom Office Templates, Database1.accd, desktop.ini, dynamicInitialization.java, GIS DataBase, gl\_gpwv3\_pdens\_00\_ascii\_one, gl\_gpwv3\_pdens\_90\_ascii\_one, IISExpress, Mental\_Health\_and\_Social\_Media\_Bali, My Web Site, NetBeansProjects, Power BI Desktop, practical, Presentation3.pptx, PrintGreatest.java

RStudio interface showing R code in the script pane and data analysis results in the console pane.

```

library(dplyr)
library(tidyr)
data <- read.csv("Mental_Health_and_Social_Media_Balance_Dataset.csv",
na.strings = c("", "NA"))
head(data)
colSums(is.na(data))
clean.omit <- na.omit(data)
print(nrow(data))
print(nrow(clean.omit))
head(clean.omit)
avg_cgpa <- mean(data$what.your.CGPA., na.rm = TRUE)
clean.replace <- data %>%
  replace_na(list(
    Age = mean(data$Age, na.rm = TRUE),
    Sleep_Quality = mean(data$Sleep_Quality, na.rm = TRUE),
    Stress_Level = mean(data$Stress_Level, na.rm = TRUE),
    Days_without_Social_Media = mean(data$Days without_Social_Media, na.rm = TRUE),
    Exercise_Frequency = mean(data$Exercise_Frequency, na.rm = TRUE),
    Facebook = mean(data$Facebook, na.rm = TRUE),
    LinkedIn = mean(data$LinkedIn, na.rm = TRUE),
    YouTube = mean(data$YouTube, na.rm = TRUE),
    TikTok = mean(data$TikTok, na.rm = TRUE),
    Twitter = mean(data$Twitter, na.rm = TRUE),
    Happiness_Index = mean(data$Happiness_Index, na.rm = TRUE)))
colSums(is.na(data))
User_ID      Age   Gender Daily_Screen_Time.hrs.
Sleep_Quality 1.10. Stress_Level.1.10. Days without_Social_Media Exercise_Frequency.week.
Social_Media_Platform Happiness_Index.1.10.
> clean.omit <- na.omit(data)
> print(nrow(data))
[1] 500

```

Environment pane shows various objects and files:

- Data: clean.omit, data, dropped\_multi\_, dropped\_one, dropped\_range, Mental\_Health\_, random\_dataset, range\_cols, selected\_cols
- Files: Home, Rhistory, AI\_Impact\_on\_labs\_2030.csv, Custom Office Templates, Database1.accd, desktop.ini, dynamicInitialization.java, GIS DataBase, gl\_gpwv3\_pdens\_00\_ascii\_one, gl\_gpwv3\_pdens\_90\_ascii\_one, IISExpress, Mental\_Health\_and\_Social\_Media\_Bali, My Web Site, NetBeansProjects, Power BI Desktop, practical, Presentation3.pptx, PrintGreatest.java

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio interface showing R script and environment.

```

1 library(dplyr)
2 library(tidyverse)
3
4 data <- read.csv("Mental_Health_and_Social_Media_Balance_Dataset.csv",
5   na.strings = c("", "NA"))
6
7 head(data)
8
9 colSums(is.na(data))
10 clean.omit <- na.omit(data)
11
12 print(nrow(data))
13 print(nrow(clean.omit))
14 head(clean.omit)
15
16 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
17
18 clean.replace <- data %>%
19   replace_na(list(
20     Age = mean(data$Age, na.rm = TRUE),
21     ...
22   ))
23
24 > clean.replace
      5           Facebook:    10
      3           LinkedIn:    10
      3           YouTube:     6
      1           TikTok:      8
      1           X (Twitter): 8
      3           LinkedIn:    8
> columns(is.na(data))
  User_ID       Age       Gender Daily_Screen_Time.hrs.
Sleep_Quality.1.10 Stress_Level.1.10 Days_Without_Social_Media Exercise_Frequency.week.
  0               0             0                  0
  0               0             0                  0
Social_Media_Platform Happiness_Index.1.10.
  0                   0
> clean.omit <- na.omit(data)
> print(nrow(data))
[1] 500
> print(nrow(clean.omit))
[1] 500
>

```

Environment pane shows:

- Data: clean.omit (500 obs. of 10 variables), data (500 obs. of 10 variables), dropped\_multi\_ (101 obs. of 9 variables), dropped\_one (101 obs. of 10 variables), dropped\_range (101 obs. of 4 variables), Mental\_Health\_ (500 obs. of 10 variables), random\_dataset (100 obs. of 10 variables), range\_cols (101 obs. of 4 variables), selected\_cols (101 obs. of 3 variables)
- Files: AI\_Impact\_on\_Jobs\_2030.csv, Database1.accdb, desktop.ini, dynamicInitialization.java, gl\_gpwv3\_pdens\_00\_ascii\_one, gl\_gpwv3\_pdens\_90\_ascii\_one, IISExpress, Mental\_Health\_and\_Social\_Media\_Bali, My Web Site, NetBeansProjects, Power BI Desktop, practical, Presentation3.pptx, PrintGreatest.java

RStudio interface showing R script and environment.

```

4 data <- read.csv("Mental_Health_and_Social_Media_Balance_Dataset.csv",
5   na.strings = c("", "NA"))
6
7 head(data)
8
9 colSums(is.na(data))
10 clean.omit <- na.omit(data)
11
12 print(nrow(data))
13 print(nrow(clean.omit))
14 head(clean.omit)
15
16 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
17
18 clean.replace <- data %>%
19   replace_na(list(
20     Age = mean(data$Age, na.rm = TRUE),
21     Your.current.year.of.study = "Unknown",
22     marital.status = "Unknown",
23     Choose.your.order = "Unknown",
24     ...
25   ))
26
27 > clean.replace
      5           Facebook:    10
      3           LinkedIn:    10
      3           YouTube:     6
      1           TikTok:      8
      1           X (Twitter): 8
      3           LinkedIn:    8
> columns(is.na(data))
  User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality.1.10 Stress_Level.1.10 Days_Without_Social_Media
1  U001 44 Male   3.1          7          6          2
2  U002 30 Father  5.1          7          8          5
3  U003 23 Other   4.4          6          7          1
4  U004 36 Female  5.7          7          8          1
5  U005 34 Female  7.0          4          7          5
6  U006 38 Male   6.6          5          7          4
> head(clean.omit)
  User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality.1.10 Stress_Level.1.10 Days_Without_Social_Media
1  U001 44 Male   3.1          7          6          2
2  U002 30 Father  5.1          7          8          5
3  U003 23 Other   4.4          6          7          1
4  U004 36 Female  5.7          7          8          1
5  U005 34 Female  7.0          4          7          5
6  U006 38 Male   6.6          5          7          4
> Exercise_Frequency.week. Social_Media_Platform Happiness_Index.1.10.
      5           Facebook:    10
      3           LinkedIn:    10
      3           YouTube:     6
      1           TikTok:      8
      1           X (Twitter): 8
      3           LinkedIn:    8
>

```

Environment pane shows:

- Data: clean.omit (500 obs. of 10 variables), data (500 obs. of 10 variables), dropped\_multi\_ (101 obs. of 9 variables), dropped\_one (101 obs. of 10 variables), dropped\_range (101 obs. of 4 variables), Mental\_Health\_ (500 obs. of 10 variables), random\_dataset (100 obs. of 10 variables), range\_cols (101 obs. of 4 variables), selected\_cols (101 obs. of 3 variables)
- Files: AI\_Impact\_on\_Jobs\_2030.csv, Database1.accdb, desktop.ini, dynamicInitialization.java, gl\_gpwv3\_pdens\_00\_ascii\_one, gl\_gpwv3\_pdens\_90\_ascii\_one, IISExpress, Mental\_Health\_and\_Social\_Media\_Bali, My Web Site, NetBeansProjects, Power BI Desktop, practical, Presentation3.pptx, PrintGreatest.java

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio interface showing R script and environment browser.

```

4 data <- read.csv("Mental_Health_and_Social_Media_Balance_dataset.csv",
5 na.strings = c("", "NA"))
6
7 head(data)
8
9 colSums(is.na(data))
10
11 clean.omit <- na.omit(data)
12
13 print(nrow(data))
14 print(nrow(clean.omit))
15 head(clean.omit)
16
17 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
18
19 clean.replace <- data %>%
20 replace_na(list(
21   Age = mean(data$Age, na.rm = TRUE),
22   Your.current.year.of.study = "Unknown",
23   marital.status = "Unknown",
24   Choose.your.gender = "Unknown",
25   what.is.your.course. = "Unknown",
26   what.is.your.CGPA. = avg_cgpa,
27   Do.you.have.depression. = "No",
28   Do.you.have.Anxiety. = "No",
29   Do.you.have.Panic.attack. = "No",
30   Did.you.seek.any.specialist.for.a.treatment. = "No"
31 ))
32
33 head(clean.replace)
34
35 colSums(is.na(clean.replace))

```

Environment browser shows various datasets and files:

- Global Environment: dropped\_multi\_, dropped\_one, dropped\_range, Mental\_Health, random\_dataset, range\_cols, selected\_cols, starts\_with\_d, Student\_Mental.
- Files: Rhistory, AI\_Impact\_on\_Jobs\_2030.csv, Custom Office Templates, desktop.ini, dynamicInitialization.java, GIS DataBase, gl\_gpwv3\_pdens\_00\_ascii\_one, gl\_gpwv3\_pdens\_90\_ascii\_one, IISExpress, Mental\_Health\_and\_Social\_Media\_Bali, My Web Site, NetBeansProjects, Power BI Desktop, practical, Presentation3.pptx, PrintGreatest.java.

RStudio interface showing R script and environment browser.

```

15 head(clean.omit)
16
17 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
18
19 clean.replace <- data %>%
20 replace_na(list(
21   Age = mean(data$Age, na.rm = TRUE),
22   Your.current.year.of.study = "Unknown",
23   marital.status = "Unknown",
24   Choose.your.gender = "Unknown",
25   what.is.your.course. = "Unknown",
26   what.is.your.CGPA. = avg_cgpa,
27   Do.you.have.depression. = "No",
28   Do.you.have.Anxiety. = "No",
29   Do.you.have.Panic.attack. = "No",
30   Did.you.seek.any.specialist.for.a.treatment. = "No"
31 ))
32
33 head(clean.replace)
34
35 colSums(is.na(clean.replace))

```

Environment browser shows various datasets and files:

- Global Environment: clean\_replace, data, dropped\_multi\_, dropped\_one, dropped\_range, Mental\_Health, random\_dataset, range\_cols, selected\_cols, starts with d.
- Files: Rhistory, AI\_Impact\_on\_Jobs\_2030.csv, Custom Office Templates, desktop.ini, dynamicInitialization.java, GIS DataBase, gl\_gpwv3\_pdens\_00\_ascii\_one, gl\_gpwv3\_pdens\_90\_ascii\_one, IISExpress, Mental\_Health\_and\_Social\_Media\_Bali, My Web Site, NetBeansProjects, Power BI Desktop, practical, Presentation3.pptx, PrintGreatest.java.

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Source on Save Go to file/function Addins
S098 PGR* S098 P7R* S098 P8R* Mental_Health_and_Social_Media_Balan...
15 head(clean.omit)
16
17 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
18
19 clean_replace <- data %>%
20   replace_na(list(
21     Age = mean(data$Age, na.rm = TRUE),
22     Your.current.year.of.study = "Unknown",
23     marital.status = "Unknown",
24     Choose.your.gender = "Unknown",
25     what.is.your.course. = "Unknown",
26     what.is.your.CGPA. = avg_cgpa,
27     Do.you.have.Depression. = "No",
28     Do.you.have.Anxiety. = "No",
29     Do.you.have.Panic.attack. = "No",
30     Did.you.seek.any.specialist.for.a.treatment. = "No"
31   ))
32
33 head(clean.replace)
34
35 colSums(is.na(clean.replace))
33:20 (Top Level) c

```

Console Terminal Background Jobs

```

R > 4.52: ~/ ~
+ Do.you.have.Panic.attack. = "No",
+ Did.you.seek.any.specialist.for.a.treatment. = "No"
+ )
> head(clean.replace)
User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality.1.10. Stress_Level.1.10. Days_without_Social_Media
1 U001 44 Male 3.1 7 6 2
2 U002 30 Other 5.1 7 8 5
3 U003 23 Other 7.4 6 7 1
4 U004 38 Female 5.7 7 8 1
5 U005 34 Female 7.0 4 7 5
6 U006 38 Male 6.6 5 7 4
Exercise_Frequency.week. Social_Media_Platform Happiness_Index.1.10.
1 5 Facebook 10
2 3 LinkedIn 10
3 3 YouTube 6
4 1 TikTok 8
5 1 X (Twitter) 8
6 3 LinkedIn 8
> colSums(is.na(clean.replace))
User_ID Age Gender Daily_Screen_Time.hrs.
0 0 0 0
Sleep_Quality.1.10. Stress_Level.1.10. Days_without_Social_Media Exercise_Frequency.week.
0 0 0 0
Social_Media_Platform Happiness_Index.1.10.
0 0

```

Environment History Connections Tutorial

Project: (None)

Files Plots Packages Help Viewer Presentation

Home

26°C Sunny 11:51 01-12-2025

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Source on Save Go to file/function Addins
S098 P6.R* S098 P7.R* S098 P8.R* Mental_Health_and_Social_Media_Balan...
15 head(clean.omit)
16
17 avg_cgpa <- mean(data$what.is.your.CGPA., na.rm = TRUE)
18
19 clean_replace <- data %>%
20   replace_na(list(
21     Age = mean(data$Age, na.rm = TRUE),
22     Your.current.year.of.study = "Unknown",
23     marital.status = "Unknown",
24     Choose.your.gender = "Unknown",
25     what.is.your.course. = "Unknown",
26     what.is.your.CGPA. = avg_cgpa,
27     Do.you.have.Depression. = "No",
28     Do.you.have.Anxiety. = "No",
29     Do.you.have.Panic.attack. = "No",
30     Did.you.seek.any.specialist.for.a.treatment. = "No"
31   ))
32
33 head(clean.replace)
34
35 colSums(is.na(clean.replace))
35:30 (Top Level) c

```

Console Terminal Background Jobs

```

R > 4.52: ~/ ~
3 U003 23 Other 7.4 6 7 1
4 U004 36 Female 5.7 7 8 1
5 U005 34 Female 7.0 4 7 5
6 U006 38 Male 6.6 5 7 4
Exercise_Frequency.week. Social_Media_Platform Happiness_Index.1.10.
1 5 Facebook 10
2 3 LinkedIn 10
3 3 YouTube 6
4 1 TikTok 8
5 1 X (Twitter) 8
6 3 LinkedIn 8
> colSums(is.na(clean.replace))
User_ID Age Gender Daily_Screen_Time.hrs.
0 0 0 0
Sleep_Quality.1.10. Stress_Level.1.10. Days_without_Social_Media Exercise_Frequency.week.
0 0 0 0
Social_Media_Platform Happiness_Index.1.10.
0 0

```

Environment History Connections Tutorial

Project: (None)

Files Plots Packages Help Viewer Presentation

Home

26°C Sunny 11:51 01-12-2025

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 9. Performing text manipulation using str\_sub(), str\_split() (R). import dataset.

The screenshot shows the RStudio interface. The code editor contains the following R script:

```

library(stringr)
library(dplyr)
library(tidyverse)

data <- read.csv("global_gaming_esports_2010_2025.csv")
head(data, 3)

data$user_prefix <- str_sub(data$user_ID, 1, 1)
data$user_number <- str_sub(data$user_ID, 2, 4)
data$Platform_short <- str_sub(data$Social_Media_Platform, 1, 3)

split_matrix <- str_split(data$Social_Media_Platform, "o", simplify = TRUE)
data$part1 <- split_matrix[, 1]
data$part2 <- split_matrix[, 2]
data$combined <- paste(data$Gender, data$Social_Media_Platform, sep = "-")

```

The console window shows the execution of the script. The environment sidebar lists various datasets like data, dropped\_multi, etc. The file browser sidebar shows files such as AI\_Impact\_on\_jobs\_2030.csv, Database1.accdb, and global\_gaming\_esports\_2010\_2025.csv. The system tray at the bottom indicates it's 26°C and sunny.

This screenshot is identical to the first one, showing the RStudio interface with the same R script, environment, and file browser. The system tray at the bottom indicates it's 26°C and sunny.

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None).
- Global Environment:** Shows variables like `data`, `dropped_multi`, `dropped_om`, etc.
- Console:** Displays R code and its output. The code reads a CSV file and prints its contents. The output includes columns such as `Gaming_Revenue_BillionsUSD`, `Esports_Revenue_MillionUSD`, `Active_Players_Million`, `Esports_Viewers_Million`, `Avg_Spending_USD`, `Top_Genre`, `Top_Platform`, `Esports_Tournaments_Count`, `Pro_Players_Count`, `INTERNET_Penetration_Percent`, `Avg_Latency_ms`, `AR_VR_Adoption_Index`, `Streaming_Influence_Index`, `covid_Impact_Index`, `Female_Gamer_Percent`, `Mobile_Gaming_Share`, `Esports_PrizePool_millionUSD`, and `Gaming_Companies_Count`.
- Terminal:** Shows the command `R > data <- read.csv("global_gaming_esports_2010_2025.csv")`.
- Plots:** No plots are visible.
- Help:** No help is visible.
- Viewer:** No viewer is visible.
- Presentation:** No presentation is visible.
- System:** Shows the date as 01-26-2025 and the time as 11:56 AM.

The screenshot shows the RStudio interface with the following details:

- Global Environment** pane on the right containing objects like `data`, `dropped_multi`, `dropped_0m`, etc.
- Code Editor** pane showing R code for reading a CSV file, extracting platform and genre information, and splitting the data into platforms and genres.
- Console** pane at the bottom displaying the output of the R code, including a table of gaming statistics.
- System tray** at the bottom left showing NIFTY +0.02%.
- Taskbar** at the bottom right with icons for File Explorer, Task View, and Start.

**MVLU COLLEGE**  
**Subject:-Data Analysis with SAS / SPSS /R**

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function    Addins

S098 P6.R\* S098 P7.R\* S098 P8.R\* global\_gaming\_esports\_2010\_2025 S098 P9.R\*

Source on Save Run Source R Script

4

5 data <- read.csv("global\_gaming\_esports\_2010\_2025.csv")

6

8 & (Top Level) 3

Console Terminal Background Jobs

R 4.5.2 - ~

```
+ separate(Game_Info, into = c("Plat", "Gen"), sep = "-")
> head(data)
#> #> #> #> #> #>
```

	Year	Country	Region	Gaming_Revenue_BillionUSD	Esports_Revenue_MillionUSD	Active_Players_Million
1	2010	United States	North America	6.24	409.6	11.0
2	2011	United States	North America	7.74	627.7	32.4
3	2012	United States	North America	9.00	928.3	51.9
4	2013	United States	North America	11.91	1321.8	50.3
5	2014	United States	North America	14.77	869.1	13.4
6	2015	United States	North America	18.31	944.4	26.6

Esports\_Viewers\_Million Avg\_Spending\_USD Top\_Genre Top\_Platform Esports\_Tournaments\_Count Pro\_Players\_Count

	1	2	3	4	5	6
1	17.9	133.9	Strategy	Mobile	104	15912
2	76.7	64.9	Sports	console	63	13797
3	14.4	84.1	MOBA	Mobile	22	3982
4	92.9	76.9	MOBA	Mobile	54	8824
5	21.3	69.0	Strategy	Mobile	84	24024
6	60.8	65.2	RPG	Console	67	12596

Internet\_Penetration\_Percent Avg\_Latency\_ms AR\_VR\_Adoption\_Index Streaming\_Influence\_Index Covid\_Impact\_Index

	1	2	3	4	5	6
1	82.5	95.4	0.13	0.10	0	0
2	70.5	83.9	0.08	0.16	0	0
3	63.7	32.3	0.11	0.22	0	0
4	51.4	88.2	0.14	0.28	0	0
5	93.4	70.5	0.14	0.34	0	0
6	84.7	30.6	0.06	0.40	0	0

Female\_Gamer\_Percent Mobile\_Gaming\_Share Esports\_PrizePool\_MillionUSD Gaming\_Companies\_Count Plat Gen Platform\_Code

	1	2	3	4	5	6
1	13.0	20	34.0	395	Mobile	Strategy
2	10.7	23	10.2	60	Console	Sports
3	13.3	26	8.5	349	Mobile	MOBA
4	12.8	29	18.3	53	Mobile	MOBA
5	17.0	32	18.3	50	Mobile	Strategy
6	17.3	35	13.4	240	Console	RPG

Genre\_End Platform\_Start Genre\_Split

	1	2	3	4	5	6
1	Mobile	Strategy				
2	rts	Console	Sports			
3	MOBA	Mobile	MOBA			
4	MOBA	Mobile	MOBA			
5	egy	Mobile	Strategy			
6	RPG	Console	RPG			

NIFTY +0.02%

Project: (None)

Environment History Connections Tutorial

Import Dataset 206 MB

Global Environment

data dropped\_multi\_101 obs. of 9 variables

dropped\_101 obs. of 10 variables

dropped\_range 101 obs. of 4 variables

global\_gaming\_400 obs. of 21 variables

mental\_health\_500 obs. of 10 variables

random\_dataset 100 obs. of 10 variables

range\_cols 101 obs. of 4 variables

selected\_cols 101 obs. of 3 variables

split\_matrix chr[1:400, 1:2] "Mobile" "Console" ...

File Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home Rhistory AL\_Impact\_on\_labs\_2030.csv Custom Office Templates Database1.accdb desktop.ini dynamicInitialization.java GIS DataBase gl\_gpwv3\_pdens\_00\_asci\_one gl\_gpwv3\_pdens\_90\_asci\_one global\_gaming\_esports\_2010\_2025.csv IISExpress Mental\_Health\_and\_Social\_Media\_Balk My Web Sites NetBeansProjects Power BI Desktop practical Presentation3.pptx

7.8 KB Nov 29, 2025, 1:56 PM 2964 KB Nov 24, 2025, 12:23 PM 336 KB Oct 10, 2025, 9:45 AM 418 B Aug 1, 2025, 9:13 AM 258 B Nov 19, 2025, 12:20 PM 7.8 KB Nov 29, 2025, 1:56 PM 343 KB Dec 1, 2025, 11:54 AM 23.4 KB Dec 1, 2025, 11:39 AM 35.5 KB Oct 11, 2025, 1:29 PM

ENGLISH IN 12/20 01-12-2025

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 10. Creating new variables using transformations and calculations in R. import dataset.

```

library(dplyr)
df <- read.csv("global_gaming_esports_2010_2025.csv")
df_clean <- df %>%
  mutate(
    Gaming_Revenue_BillionUSD = ifelse(is.na(Gaming_Revenue_BillionUSD), 0, Gaming_Revenue_BillionUSD),
    Esports_Revenue_MillionUSD = ifelse(is.na(Esports_Revenue_MillionUSD), 0, Esports_Revenue_MillionUSD),
    Avg_Spending_USD = ifelse(is.na(Avg_Spending_USD), 0, Avg_Spending_USD),
    Active_Players_Million = ifelse(is.na(Active_Players_Million), 0, Active_Players_Million)
  )
head(df_clean)
df_calc <- df_clean %>%
  mutate(
    Revenue_Per_Player = Gaming_Revenue_BillionUSD * 1000 / Active_Players_Million,
    Total_Revenue_Million = Gaming_Revenue_BillionUSD * 1000 + Esports_Revenue_MillionUSD
  )
df_calc %>% select(Year, Gaming_Revenue_BillionUSD, Active_Players_Million, Revenue_Per_Player)
df_logic <- df_clean %>%
  mutate(
    Marker_Level = ifelse(Gaming_Revenue_BillionUSD > 50, "High", "Medium"),
    Player_Scale = ifelse(Active_Players_Million > 200, "Large", "Normal")
  )
  
```

R 4.5.2 - ~

```

> library(dplyr)
> df <- read.csv("global_gaming_esports_2010_2025.csv")
> df_clean <- df %>%
+   mutate(
+     Gaming_Revenue_BillionUSD = ifelse(is.na(Gaming_Revenue_BillionUSD), 0, Gaming_Revenue_BillionUSD),
+     Esports_Revenue_MillionUSD = ifelse(is.na(Esports_Revenue_MillionUSD), 0, Esports_Revenue_MillionUSD),
+     Avg_Spending_USD = ifelse(is.na(Avg_Spending_USD), 0, Avg_Spending_USD),
+     Active_Players_Million = ifelse(is.na(Active_Players_Million), 0, Active_Players_Million)
+   )
> |
```

```

library(dplyr)
df <- read.csv("global_gaming_esports_2010_2025.csv")
df_clean <- df %>%
  mutate(
    Gaming_Revenue_BillionUSD = ifelse(is.na(Gaming_Revenue_BillionUSD), 0, Gaming_Revenue_BillionUSD),
    Esports_Revenue_MillionUSD = ifelse(is.na(Esports_Revenue_MillionUSD), 0, Esports_Revenue_MillionUSD),
    Avg_Spending_USD = ifelse(is.na(Avg_Spending_USD), 0, Avg_Spending_USD),
    Active_Players_Million = ifelse(is.na(Active_Players_Million), 0, Active_Players_Million)
  )
head(df_clean)
  
```

Year	Country	Region	Gaming_Revenue_BillionUSD	Esports_Revenue_MillionUSD	Active_Players_Million			
1	2010	United States	North America	6.24	409.6	11.0		
2	2011	United States	North America	7.74	622.7	32.4		
3	2012	United States	North America	9.60	928.3	51.9		
4	2013	United States	North America	11.91	1321.8	50.3		
5	2014	United States	North America	14.77	869.1	13.4		
6	2015	United States	North America	18.31	944.4	26.6		
			Esports_Viewers_Million	Avg_Spending_USD	Top_Genre	Platform	Esports_Tournaments_Count	Pro_Players_Count
1			17.9	133.9	Strategy	Mobile	104	15912
2			76.7	64.9	Sports	Console	63	13707
3			148.4	84.1	MOBA	Mobile	22	3982
4			92.9	76.9	MOBA	Mobile	54	8262
5			21.3	69.0	Strategy	Mobile	84	24024
6			60.8	65.2	RPG	Console	67	12596
			Internet_Penetration_Percent	Avg_Latency_ms	AR_VR_Adoption_Index	Streaming_Influence_Index	Covid_Impact_Index	
1			82.5	95.4	0.13	0.10	0	
2			70.5	83.9	0.08	0.05	0	
3			63.7	32.3	0.11	0.22	0	
4			51.4	88.2	0.14	0.28	0	
5			93.4	70.5	0.14	0.34	0	
6			84.7	30.6	0.06	0.40	0	
			Female_Gamer_Percent	Mobile_Gaming_Share	Esports_PrizePool_MillionUSD	Gaming_Companies_Count		
1			13.0	20	34.0	395		
2			10.7	23	10.2	60		
3			13.3	26	8.5	349		
4			12.8	29	18.3	53		
5			17.0	32	18.3	50		
6			17.3	35	13.4	240		

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows an RStudio interface with the following components:

- Top Bar:** RStudio, File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project Bar:** Go to file/function, Addins.
- Code Editor:** A large R script containing code for data cleaning, calculating revenue per player, and creating logic based on market level and player scale. The code includes imports from 'dplyr' and 'tidyverse' packages.
- Environment Tab:** Shows a list of objects in the environment, including 'df\_clean', 'df\_calc', and various datasets like 'global\_gaming\_automation' and 'global\_ai\_workforce'.
- Files Tab:** Shows a file tree with files like 'desktop.ini', 'dynamicalInitialization.java', 'GIS DataBase', and various CSV and PPTX files.
- Console Tab:** Displays the R session, showing the execution of code and an error message about an unexpected symbol.
- Terminal Tab:** Shows the command line history.
- Background Jobs:** Shows no active jobs.
- Bottom Status Bar:** Shows the date and time (12/21/2025), battery status (30°C, Sunny), and system icons.

The screenshot shows the RStudio interface. The left pane displays a large dataset in the 'Source' tab, specifically a subset of the 'Gaming\_Revenue\_BillionUSD' dataset. The right pane shows a file browser with numerous files listed under 'Environment'. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help, and Addins. The bottom status bar shows the number of files (194 MB) and the current date and time (12:22 01/12/2025).

Year	Gaming_Revenue_BillionUSD	Active_Players_Million	Revenue_Per_Player
1	2010	6.24	11.0
2	2011	7.74	32.4
3	2012	9.60	51.9
4	2013	11.91	50.3
5	2014	14.77	13.4
6	2015	18.31	26.6
7	2016	22.71	28.1
8	2017	28.17	4.3
9	2018	34.93	56.7
10	2019	43.32	9.1
11	2020	53.72	33.0
12	2021	66.62	36.8
13	2022	82.62	25.1
14	2023	102.46	7.6
15	2024	127.07	59.0
16	2025	157.58	53.5
17	2010	14.63	2945.42056
18	2011	16.19	59.7
19	2012	17.92	58.4
20	2013	19.84	4.3
21	2014	21.96	51.4
22	2015	24.30	38.4
23	2016	26.89	33.5
24	2017	29.77	32.8
25	2018	32.94	12.7
26	2019	36.46	57.4
27	2020	40.36	29.0
28	2021	44.57	52.3
29	2022	49.44	11.4
30	2023	54.71	42.2
31	2024	60.56	47.0
32	2025	67.02	19.3
33	2010	9.03	41.0
34	2011	10.62	58.2
35	2012	12.49	59.2
36	2013	14.70	45.0
37	2014	17.79	20.7
38	2015	20.34	28.4
39	2016	23.93	39.5
40	2017	28.14	9.7
			2901.03093

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Console Terminal Background Jobs

```
R > R 4.5.2 - ~/
```

```
211 2012      18.23    17.4     1047.70115
212 2013      21.99    41.2     533.73786
213 2014      26.53    45.4     584.36123
214 2015      32.01    10.2     3138.23529
215 2016      38.62    29.3     1318.08874
216 2017      46.59    57.0     817.68462
217 2018      56.20    41.1     1167.39539
218 2019      67.80    27.3     2483.51648
219 2020      81.80    11.4     7175.43860
220 2021      98.68    38.8     2543.29897
221 2022     119.05    38.9     3060.41131
222 2023     143.62    15.9     9032.70440
223 2024     173.27    35.7     4853.50140
224 2025     209.03    49.7     4205.83301
225 2026      1.97    50.5     39.00000
226 2011      2.43    32.7     72.10862
227 2012      3.01    44.3     67.94582
228 2013      3.73    23.5     158.72340
229 2014      4.62    17.6     262.50000
230 2015      5.72    58.4     97.94521
231 2016      7.08    10.2     694.11765
232 2017      8.77    44.6     196.63677
233 2018     10.86    58.1     186.91910
234 2019      13.45    38.3     331.7493
235 2020      16.65    24.7     674.00877
236 2021      20.61    5.3     3888.67925
237 2022      25.52    5.9     4325.42373
238 2023      31.60    13.9     2273.38129
239 2024      39.13    28.6     1368.18182
240 2025      48.44    52.2     927.96935
241 2010      6.23    33.4     186.52695
242 2011      6.60    31.7     208.20189
243 2012      7.00    2.8     2500.00000
244 2013      7.42    50.6     146.64032
245 2014      7.87    55.4     142.05776
246 2015      8.34    26.8     311.19403
247 2016      8.84    29.2     302.73973
248 2017      9.37    41.7     224.70024
249 2018      9.94    55.0     180.72727
250 2019     10.53    41.3     254.96368
[ reached 'max' / getoption("max.print") -- omitted 150 rows ]
```

Environment History Connections Tutorial

Import Dataset 194 MB

R df\_clean 400 obs. of 23 variables  
df\_logic 400 obs. of 23 variables  
dropped\_multi\_ 101 obs. of 9 variables  
dropped\_one 101 obs. of 10 variables  
dropped\_range 101 obs. of 4 variables  
global\_ai\_wor... 220 obs. of 12 variables  
global\_gaming\_ 400 obs. of 21 variables  
Mental\_Health\_ 500 obs. of 10 variables  
random\_dataset 100 obs. of 10 variables  
range\_cols 101 obs. of 4 variables

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home Name Size Modified

desktop.ini 418 B Aug 1, 2025, 9:13 AM  
dynamicInitialization.java 258 B Nov 19, 2025, 12:20 PM  
GIS DataBase  
gl\_gpww2\_pdens\_00\_ascii\_one  
gl\_gpww2\_pdens\_90\_ascii\_one  
global\_ai\_workforce\_automation\_2015 16.2 KB Dec 1, 2025, 12:14 PM  
global\_gaming\_esports\_2010\_2025.cp 44.3 KB Dec 1, 2025, 11:54 AM  
IISExpress  
My Web Sites  
NetBeansProjects  
Power BI Desktop  
practical  
Presentation3.pptx 35.5 KB Oct 11, 2025, 1:29 PM  
PrintGreatest.java 642 B Nov 19, 2025, 12:06 PM  
R  
random\_dataset.csv 19.2 KB Nov 22, 2025, 11:26 AM  
cnn.nco

Top Stories GRAP-IV curbs t...  
Search ENG IN 12:22 01-12-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source on Save

```
S098 P6.R* S098 P7.R* S098 P8.R* global_gaming_esports_2010_2025.x S098 P9.R* S098 P10.R* global_ai_workforce_automation_2015... R Script
```

```
20
21 df_calc %>% select(Year, Gaming_Revenue_BillionUSD, Active_Players_Million, Revenue_Per_Player)
22
23 df_logic <- df_clean %>%
24   mutate(
25     Market_Level = ifelse(Gaming_Revenue_BillionUSD > 50, "High", "Medium"),
26     Player_Scale = ifelse(Active_Players_Million > 200, "Large", "Normal")
27   )
28
29 df_logic %>% select(Year, Gaming_Revenue_BillionUSD, Market_Level, Active_Players_Million, Player_Scale)
30
31 df_text <- df_clean %>%
32   mutate(
33     Summary = paste(Country, "has", Gaming_Revenue_BillionUSD, "B USD gaming revenue with", Active_Players_Million, "M active p
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
[ reached 'max' / getoption("max.print") -- omitted 150 rows ]
```

```
> df_logic <- df_clean %>%
+   mutate(
+     Market_Level = ifelse(Gaming_Revenue_BillionUSD > 50, "High", "Medium"),
+     Player_Scale = ifelse(Active_Players_Million > 200, "Large", "Normal")
+   )
+ |
```

Environment History Connections Tutorial

Import Dataset 190 MB

R df\_clean 400 obs. of 21 variables  
df\_logic 400 obs. of 23 variables  
dropped\_multi\_ 101 obs. of 9 variables  
dropped\_one 101 obs. of 10 variables  
dropped\_range 101 obs. of 4 variables  
global\_ai\_wor... 220 obs. of 12 variables  
global\_gaming\_ 400 obs. of 21 variables  
Mental\_Health\_ 500 obs. of 10 variables  
random\_dataset 100 obs. of 10 variables  
range\_cols 101 obs. of 4 variables

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home Name Size Modified

desktop.ini 418 B Aug 1, 2025, 9:13 AM  
dynamicInitialization.java 258 B Nov 19, 2025, 12:20 PM  
GIS DataBase  
gl\_gpww2\_pdens\_00\_ascii\_one  
gl\_gpww2\_pdens\_90\_ascii\_one  
global\_ai\_workforce\_automation\_2015 16.2 KB Dec 1, 2025, 12:14 PM  
global\_gaming\_esports\_2010\_2025.cp 44.3 KB Dec 1, 2025, 11:54 AM  
IISExpress  
My Web Sites  
NetBeansProjects  
Power BI Desktop  
practical  
Presentation3.pptx 35.5 KB Oct 11, 2025, 1:29 PM  
PrintGreatest.java 642 B Nov 19, 2025, 12:06 PM  
R  
random\_dataset.csv 19.2 KB Nov 22, 2025, 11:26 AM  
cnn.nco

USD/INR +0.37%  
Search ENG IN 12:23 01-12-2025

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows an RStudio environment with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None) is selected.
- Script Editor:** The main pane displays R code for data manipulation and analysis. The code includes:
  - Reading datasets: `global\_gaming\_esports\_2010\_2025`, `global\_ai\_workforce\_automation\_2015\_2024`.
  - Creating a summary variable: `df\_text`.
  - Calculating final values: `final\_df`.
  - Final output: A table showing Year, Gaming\_Revenue\_BillionUSD, Market\_Level, Active\_Players\_Million, and Player\_Scale.
- Console:** Shows the execution of the R code, including the creation of the `final\_df` data frame and its contents.
- Environment:** A sidebar showing the global environment with various objects and their sizes.
- Plots:** No plots are currently visible.
- Help:** Help is available for the current session.
- Viewer:** No files are currently open.
- Presentation:** No presentations are currently open.
- System:** The bottom bar shows system icons for battery, signal, and time (12:23 PM).

The screenshot shows an RStudio interface with multiple tabs open. The main code editor tab contains R code for summarizing gaming revenue and player data across different countries. The preview pane below shows a sample of the dataset with columns for Year, Revenue\_Per\_Player, Market\_Level, Active\_Players\_Million, and Player\_Scale.

```
df %>% select(-Year, -Gaming_Revenue_BillionUSD, -Market_Level, -Active_Players_Million, -Player_Scale)
df_text <- df_clean %>
  mutate(
    summary = paste(Country, "has", Gaming_Revenue_BillionUSD, "B USD gaming revenue with", Active_Players_Million, "M active p
  )
)
head(df_text$summary)
final_df <- df_clean %>
  mutate(
    Revenue_Per_Player = Gaming_Revenue_BillionUSD * 1000 / Active_Players_Million,
    Influence_Score = round((Internet_Penetration_Percent + Streaming_Influence_Index + AR_VR_Adoption_Index) / 3, 2),
    Region_Report = paste(Region, "-", Top_Genre)
)

```

Year	Revenue_Per_Player	Market_Level	Active_Players_Million	Player_Scale
179 2012	6.29	Medium	18.1	Normal
180 2013	7.20	Medium	26.5	Normal
181 2014	8.24	Medium	8.6	Normal
182 2015	9.43	Medium	39.5	Normal
183 2016	10.80	Medium	2.7	Normal
184 2017	12.36	Medium	24.5	Normal
185 2018	14.14	Medium	7.7	Normal
186 2019	16.50	High	39.5	Normal
187 2020	18.53	Medium	11.7	Normal
188 2021	21.21	Medium	30.3	Normal
189 2022	24.28	Medium	21.7	Normal
190 2023	27.79	Medium	21.9	Normal
191 2024	31.80	Medium	44.6	Normal
192 2025	36.40	Medium	22.5	Normal
193 2010	14.29	Medium	19.5	Normal
194 2011	16.47	Medium	35.3	Normal
195 2012	18.98	Medium	50.3	Normal
196 2013	21.88	Medium	44.5	Normal
197 2014	25.21	Medium	10.3	Normal
198 2015	29.06	Medium	5.8	Normal
199 2016	33.50	Medium	16.0	Normal
200 2017	38.61	Medium	35.8	Normal

# **MVLU COLLEGE**

## **Subject:-Data Analysis with SAS / SPSS /R**

The screenshot shows the RStudio interface with several windows open:

- Code Editor:** Displays R code for data manipulation and analysis.
- File Browser:** Shows the project structure with files like `desktop.ini`, `dynamicalInitialization.java`, `GIS DataBase`, `gl_gpwv3_pdens_00_asci_one`, `gl_gpwv3_pdens_90_asci_one`, `global_ai_workforce_automation_2016`, `global_gaming_esports_2010_2025.cs`, `IISExpress`, `Mental_Health_and_Social_Media_Bal`, `My Web Sites`, `NetBeansProjects`, `Power BI Desktop`, `practical`, `Presentation3.pptx`, `PrintGreatestJava`, `R`, `random_dataset.csv`, and `random_nc.nc`.
- Console:** Displays the output of the R code, including a table of data and a message indicating 200 rows were omitted.
- Terminal:** Shows the command `R -e "source('~/Desktop/Project/Project.R')"`.
- Background Jobs:** Shows no active jobs.
- System Tray:** Includes icons for USD/INR exchange rate (4.37%), battery level (80%), signal strength, and system status.

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

RStudio interface showing R code in the script pane and a file browser in the right pane.

```

32   mutate(
33     Summary = paste(Country, "has", Gaming_Revenue_BillionUSD, "B USD gaming revenue with", Active_Players_Million, "M active p
34   )
35 
36 head(df_text$Summary)
37 
38 final_df <- df_clean %>%
39   mutate(
40     Revenue_Per_Player = Gaming_Revenue_BillionUSD * 1000 / Active_Players_Million,
41     Influence_Score = round((Internet_Penetration_Percent + Streaming_Influence_Index + AR_VR_Adoption_Index) / 3, 2),
42     Region_Report = paste(Region, "_", Top_Genre)
43   )
44 
45 head(final_df)
46 
```

Console output:

```

197 2014      25.21    Medium      10.3    Normal
198 2015      29.06    Medium       5.8    Normal
199 2016      33.50    Medium      16.0    Normal
200 2017      38.61    Medium      35.8    Normal
[ reached 'max' / getoption("max.print") -- omitted 200 rows ]
> df_text <- df_clean %>%
+   mutate(
+     Summary = paste(Country, "has", Gaming_Revenue_BillionUSD, "B USD gaming revenue with", Active_Players_Million, "M active play
ers,")
+   )
> head(df_text$Summary)
[1] "United States has 6.24 B USD gaming revenue with 11 M active players."
[2] "United States has 7.74 B USD gaming revenue with 32.4 M active players."
[3] "United States has 9.6 B USD gaming revenue with 51.9 M active players."
[4] "United States has 14.91 B USD gaming revenue with 50.3 M active players."
[5] "United States has 13.4 B USD gaming revenue with 13.4 M active players."
[6] "United States has 18.31 B USD gaming revenue with 26.6 M active players."
> final_df <- df_clean %>%
+   mutate(
+     Revenue_Per_Player = Gaming_Revenue_BillionUSD * 1000 / Active_Players_Million,
+     Influence_Score = round((Internet_Penetration_Percent + Streaming_Influence_Index + AR_VR_Adoption_Index) / 3, 2),
+     Region_Report = paste(Region, "_", Top_Genre)
+   )
> | 
```

File browser:

- df\_text
- dropped\_multi\_
- dropped\_one\_
- dropped\_range\_
- final\_df
- global\_ai\_wor...
- global\_gaming...
- Mental\_Health...
- random\_dataset
- range\_cols

RStudio interface showing R code in the script pane and a file browser in the right pane.

```

31 df_text <- df_clean %>%
32 
33 head(final_df)
Year Country Region Gaming_Revenue_BillionUSD Esports_Revenue_MillionUSD Active_Players_Million
1 2010 United States North America      6.24          409.6           11.0
2 2011 United States North America      7.74          622.7            32.4
3 2012 United States North America      9.60          928.3            51.9
4 2013 United States North America     11.91         1321.8            50.3
5 2014 United States North America     14.77          869.1            13.4
6 2015 United States North America     18.31          944.4            26.6
Esports_Viewers_Million Avg_Spending_USD Top_Genre Top_Platform Esports_Tournaments_Count Pro_Players_Count
1 17.9          133.9   Strategy   Mobile          104          15912
2 76.7          64.9    Sports    Console          63          13797
3 148.4          84.1   MOBA     Mobile          22          3982
4 92.9          76.9    MOBA     Mobile          54          8262
5 21.3          69.0    Strategy  Mobile          84          24024
6 60.8          65.2    RPG      Console          67          12596
Internet_Penetration_Percent Avg_Latency_ms AR_VR_Adoption_Index Streaming_Influence_Index Covid_Impact_Index
1 25.3          4.0        0.12          0.12          0
2 70.5          83.9       0.08          0.16          0
3 63.7          32.3       0.11          0.22          0
4 51.4          88.2       0.14          0.28          0
5 93.4          70.5       0.14          0.34          0
6 84.7          30.6       0.06          0.40          0
Female_Gamer_Percent Mobile_Gaming_Share Esports_PrizePool_MillionUSD Gaming_Companies_Count Revenue_Per_Player
1 13.0           20          34.0          395          567.2727
2 10.7           23          10.2           60          238.8889
3 13.3           26          8.5            349          184.9211
4 12.8           29          18.3            53          236.2793
5 17.0           32          18.3            50          1102.2388
6 17.3           35          13.4            240          688.3459
Influence_Score Region_Report
1 27.58 North America - Strategy
2 23.58 North America - Sports
3 21.34 North America - MOBA
4 17.27 North America - MOBA
5 31.29 North America - Strategy
6 28.39 North America - RPG
| 
```

File browser:

- df\_text
- dropped\_multi\_
- dropped\_one\_
- dropped\_range\_
- final\_df
- global\_ai\_wor...
- global\_gaming...
- Mental\_Health...
- random\_dataset
- range\_cols