

# MVLU COLLEGE

**AIM:**Creating datasets from raw data (text files, CSV files, Excel sheets) and importing data into SAS/SPSS/R.

Importing data set:

The screenshot shows two instances of the RStudio interface. The top instance is used for installing the 'Rcpp' package, while the bottom instance is used for importing an Excel dataset.

**Top Instance (Console):**

```
Installing 'Rcpp' for Excel Import...
Installing package into 'c:/users/mvluc/onedrive/documents/R/win-library/4.1'
(as 'lib' is unspecified)

There is a binary version available but the source version is later:
  binary source needs_compilation
Rcpp 1.0.10 1.1.0 TRUE

Binaries will be installed
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.1/Rcpp_1.0.10.zip'
Content type 'application/zip' length 3291415 bytes (3.1 MB)
=====
downloaded 3.1 MB

package 'Rcpp' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\mvluc\AppData\Local\Temp\Rtmpg1v5dK\downloaded_packages

<U+2714> Package 'Rcpp' successfully installed.
```

**Bottom Instance (Console):**

```
Installing 'ImportExcelData'
Installing package into 'C:/Users/mvluc/OneDrive/Desktop/ankita tiwari/Data Analysis/student data.xls.xlsx'
(as 'lib' is unspecified)

There is a binary version available but the source version is later:
  binary source needs_compilation
Rcpp 1.0.10 1.1.0 TRUE

Binaries will be installed
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.1/ImportExcelData_0.1.0.zip'
Content type 'application/zip' length 1000000 bytes (1 MB)
=====
downloaded 1.0 MB

package 'ImportExcelData' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\mvluc\OneDrive\Desktop\ankita tiwari\Data Analysis
```

**Environment View:**

The environment view shows the Global Environment tab, which is currently empty.

**Data Preview View:**

ID (double)	Age (double)	Study Hours (double)	Attendance (%) (double)	Score (%) (double)
1	17	2.5	85	76
2	16	3.0	90	82
3	17	1.0	70	54
4	18	4.0	95	88
5	16	2.0	80	72
6	17	3.5	92	85
7	18	1.5	75	60
8	17	2.8	88	78
9	16	0.5	65	40
10	18	4.5	98	93
11	17	3.2	87	81
12	16	2.7	82	74
13	18	1.2	68	55

**Code Preview:**

```
library(readxl)
student_data_xls <- read_excel("c:/users/mvluc/onedrive/desktop/ankita tiwari/Data Analysis/student data.xls.xlsx")
View(student_data_xls)
```

**Buttons:**

Import Cancel

**ANKITA TIWARI**

**SYCS**

**S122**

**Data Analysis with SAS / SPSS /R PRACTICAL NO.2**

# MVLU COLLEGE

The screenshot shows two instances of the RStudio interface running side-by-side on a Windows desktop. Both instances have identical configurations and content.

**Top Panel (Environment View):**

- Import Dataset: 234 MB
- Global Environment
- Data: student\_data.xls (20 obs. of 5 variables)

**Bottom Panel (File Explorer):**

- Files: .Rhistory, Custom Office Templates, desktop.ini, ExamOffice tools.accdb, GIS DataBase, IISExpress, My Web Sites, NetBeansProjects, NetBeansProjectstr, R, Virtual Machines, Visual Studio 2022
- Size: 31 B, 418 B, 716 KB
- Modified: Sep 30, 2025, 3:20 PM, Aug 2, 2025, 10:01 AM, Oct 11, 2025, 1:29 PM

**Console Output:**

```
R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale
```

**Background Jobs:**

```
R - R.4.1.2 - ~
```

**System Tray:**

- 29°C Sunny
- ENG IN
- 11:17
- 24-11-2025

# MVLU COLLEGE

CSV:

Showing 1 to 20 of 20 entries, 5 total columns

Console Terminal Background Jobs

R - R4.1.2 - ~

You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale.

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help,  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

```
> library(readxl)
Warning message:
package 'readxl' was built under R version 4.1.3
> student_data.xls <- read_excel("C:/users/mvluc/OneDrive/Desktop/ankita tiwari/Data Analysis/student.xls")
> View(student_data.xls)
> View(student_data.xls)
> View(student_data.xls)
```

Showing 1 to 9 of 101 entries, 11 total columns

Console Terminal Background Jobs

R - R4.1.2 - ~

Natural language support but running in an English locale.

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help,  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

```
> library(readxl)
Warning message:
package 'readxl' was built under R version 4.1.3
> student_data.xls <- read_excel("C:/users/mvluc/OneDrive/Desktop/ankita tiwari/Data Analysis/student.xls")
> View(student_data.xls)
> View(student_data.xls)
> View(student_data.xls)
> Student.Mental.health <- read.csv("C:/users/mvluc/OneDrive/Desktop/ankita tiwari/python/Student Mental Health.csv", stringsAsFactors=TRUE)
> View(Student.Mental.health)
```

**ANKITA TIWARI**

**SYCS**

**S122**

**Data Analysis with SAS / SPSS /R PRACTICAL NO.2**

# MVLU COLLEGE

The screenshot shows the RStudio interface. In the top-left pane, there is an R script with the following code:

```

1: read.csv("C:/Users/mvluc/OneDrive/Desktop/ankita tiwari/python/student Mental.health.csv", stringsAsFactors = TRUE)
2: View(Student.Mental.health)
3: lm(choose.your.gender ~ Age)
4: attach(Student.Mental.health)
5:
6

```

An error message is displayed in the console:

```

> lm(choose.your.gender ~ Age)
Error in formula.default(object, env = baseenv()) : invalid formula

```

The environment pane shows two datasets:

- Data**
  - student\_data.xls | 20 obs. of 5 variables
  - Student.Mental.health | 101 obs. of 11 variables

The file browser pane shows the following files:

- Rhistory (31 B)
- Custom Office Templates
- desktop.ini (418 B)
- ExamOffice tools.accdb
- GIS DataBase
- IISExpress
- My Web Sites
- NetBeansProjects
- NetBeansProjectsrser
- R (716 KB)
- Virtual Machines
- Visual Studio 2022

## Text:

The screenshot shows the RStudio interface. In the top-left pane, there is an R script with the following code:

```

1: read.csv("C:/Users/mvluc/OneDrive/Desktop/ankita tiwari/python/student Mental.health.csv", stringsAsFactors = TRUE)
2: View(Student.Mental.health)
3: lm(choose.your.gender ~ Age)
4: attach(Student.Mental.health)
5:
6

```

An error message is displayed in the console:

```

> lm(choose.your.gender ~ Age)
Error in formula.default(object, env = baseenv()) : invalid formula

```

A modal dialog titled "Import Dataset" is open in the center of the screen. It shows the following settings:

- Name: Text\_data
- Input File: student\_data.xls
- Encoding: Automatic
- Heading: Yes
- Row names: Automatic
- Separator: Whitespace
- Decimal: Period
- Quote: Double ("")
- Comment: None
- na.strings: NA
- Strings as factors: checked

The "Data Frame" section shows the first few rows of the dataset:

ID	Name	Age	Gender	Country	Occupation
1	Alice Johnson	24	F	USA	Data Analyst
2	Mike Turner	32	M	Canada	Software Engineer
3	Priya Sharma	28	F	India	Designer
4	Tom Müller	35	M	Germany	Accountant
5	Sarah Lee	32	F	South Korea	Teacher
6	David Brown	40	M	UK	Project Manager
7	Lina Garcia	27	F	Spain	Nurse
8	John Parker	38	M	USA	Artist
9	Mei Chen	29	F	China	Graphic Designer
10	Olivia Taylor	25	F	UK	Journalist
11	Anna Yusuf	26	M	Kenya	Sales Representative
12	Lucas Silva	30	M	Brazil	Sales Representative
13	James Wilson	33	F	Australia	Lawyer
14	Noah Thompson	45	M	USA	Researcher
15	Hana Takashii	34	F	Japan	Architect
16	Omar Hassan	32	M	Egypt	Architect

The "Presentation" pane shows the following files:

- Rhistory (31 B)
- Custom Office Templates
- desktop.ini (418 B)
- ExamOffice tools.accdb
- GIS DataBase
- IISExpress
- My Web Sites
- NetBeansProjects
- NetBeansProjectsrser
- R (716 KB)
- Virtual Machines
- Visual Studio 2022

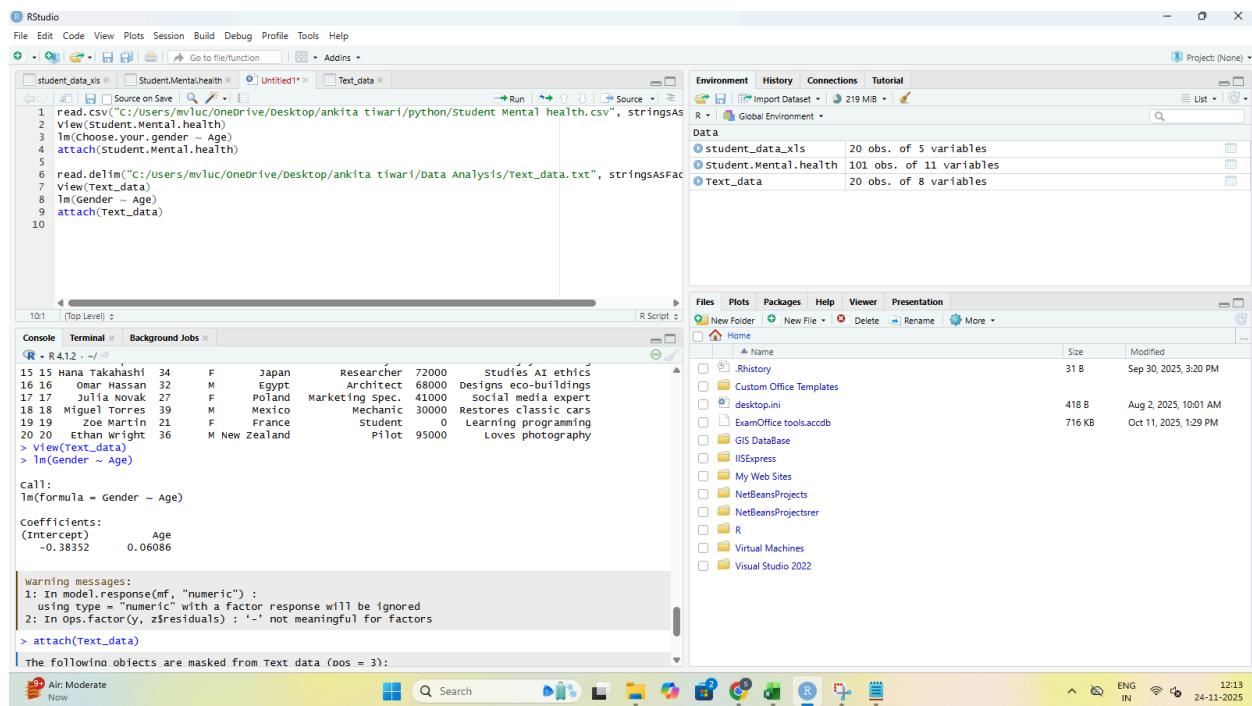
ANKITA TIWARI

SYCS

S122

Data Analysis with SAS / SPSS /R PRACTICAL NO.2

**MVLU COLLEGE**



**ANKITA TIWARI**  
**SYCS**  
**S122**  
**Data Analysis wi**