

## ITDS 6<sup>th</sup> Practical

The first screenshot shows the RStudio interface with the following code in the script editor:

```
#Step 1: Install and Load Tidyverse
library(tidyverse)

#Step 2: Create a sample dataset
students <- tibble(
  Name = c("Anjali", "Rahul", "Priya", "Aman", "Sneha", "Ravi"),
  Age = c(20, 22, 19, 21, 20, 23),
  Gender = c("F", "M", "F", "M", "F", "M"),
  Marks = c(85, 78, 92, 88, 95, 60)
)

#install.packages("tidyverse")

#view the original dataset
print("Original Data: ")
print(students)
```

The console output shows the original dataset:

```
> #view the original dataset
> print("Original Data: ")
[1] "Original Data: "
> print(students)
# A tibble: 6 x 4
  Name      Age Gender Marks
<chr> <dbl> <chr> <dbl>
1 Anjali    20 F      85
2 Rahul     22 M      78
3 Priya     19 F      92
4 Aman      21 M      88
5 Sneha     20 F      95
6 Ravi      23 M      60
```

The Environment pane shows the installed packages:

Name	Description	Version
dplyr	A Grammar of Data Manipulation	1.1.4
dbplyr	A 'dplyr' Back End for Databases	2.5.0
dtplyr	Data Table Back-End for 'dplyr'	1.3.1

The second screenshot shows the RStudio interface with the following code in the script editor:

```
#install.packages("tidyverse")

#view the original dataset
print("Original data:")
print(students)

#a. Using select(): Select name and marks columns
selected_data <- students %>%
  select(Name, Marks)
print("Selected Name and Marks: ")
print(selected_data)

#a. Using select(): Select name and marks columns
selected_data <- students %>%
  select(Name, Marks)
print("Selected Name and Marks: ")
print(selected_data)

#a. Using select(): Select name and marks columns
selected_data <- students %>%
  select(Name, Marks)
print("Selected Name and Marks: ")
print(selected_data)
```

The console output shows the selected data:

```
> select(Name, Marks)
# A tibble: 6 x 2
  Name      Marks
<chr> <dbl>
1 Anjali    85
2 Rahul     78
3 Priya     92
4 Aman      88
5 Sneha     95
6 Ravi      60
```

The Environment pane shows the installed packages:

Name	Size	Modified
RData	4.4 KB	Jul 31, 2025, 12:21 PM
.Rhistory	8.5 KB	Jul 31, 2025, 12:21 PM
18.cpp	1.1 KB	Jun 30, 2025, 11:54 AM
18.exe	1.8 MB	Jun 30, 2025, 12:32 PM
9133 div Q.xlsx	12.5 KB	Apr 3, 2025, 8:39 AM
254038		
Adobe		
Desktop - Shortcut Link	740 B	Jul 17, 2025, 8:22 AM
desktop.ini	402 B	Feb 1, 2023, 3:56 PM
EDA-B2		
edaPSAB.txt	842 B	Jul 17, 2025, 8:49 AM

## ITDS 6<sup>th</sup> Practical

The image displays two screenshots of the RStudio environment, showing the progression of an R script for data manipulation using the dplyr package.

**Top Screenshot:**

- Code Editor:** The script defines a data frame 'students' and performs two operations:
  - a. Using select():** Selects 'Name' and 'Marks' columns, storing the result in 'selected\_data'.
  - b. Using filter():** Filters students with marks greater than 85, storing the result in 'high\_scorers'.
- Console:** Shows the execution of the filter operation, resulting in a tibble with 3 rows and 4 columns:
 

	Name	Age	Gender	Marks
1	Priya	19	F	92
2	Aman	21	M	88
3	Sneha	20	F	95
- Environment:** Lists files in the current directory, including .RData, .Rhistory, 18.cpp, 18.exe, 9133 div Q.xlsx, 254038, Adobe, Desktop - ShortcutLink, desktop.ini, EDA-B2, and edaPSAB.txt.

**Bottom Screenshot:**

- Code Editor:** The script continues with:
  - c. Using mutate():** Adds a new column 'Result' to the 'students\_with\_result' data frame, where 'Pass' is assigned for marks >= 75 and 'Fail' for marks < 75.
- Console:** Shows the execution of the mutate operation, resulting in a tibble with 6 rows and 5 columns:
 

	Name	Age	Gender	Marks	Result
1	Anjali	20	F	85	Pass
2	Rahul	22	M	78	Pass
3	Priya	19	F	92	Pass
4	Aman	21	M	88	Pass
5	Sneha	20	F	95	Pass
6	Ravi	23	M	60	Fail
- Environment:** Similar to the top screenshot, showing the same file list.

## ITDS 6<sup>th</sup> Practical

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for data manipulation using `mutate()` and summarization using `group_by()` and `summarise()`.
- Environment:** Shows the current environment with variables like `students_with_result` and `avg_marks_by_gender`.
- Console:** Displays the output of the R code, including the average marks by gender.
- Files:** Shows a list of files in the current directory, including `RData`, `.Rhistory`, `18.cpp`, `18.exe`, `9133 div Q.xlsx`, `254038`, `Adobe`, `Desktop - ShortcutLink`, `desktop.ini`, `EDA-B2`, and `edaP5AB.txt`.

```
#c. using mutate(): Add a new column 'Result' as Pass/Fail
students_with_result<- students %>%
  mutate(Result = ifelse(Marks >= 75, "Pass", "Fail"))
print("Data with Result column: ")
print(students_with_result)

#d. using group_by() and summarise(): Average Marks by Gender
avg_marks_by_gender <- students %>%
  group_by(Gender) %>%
  summarise(Average_Marks = mean(Marks))
print("Average Marks by Gender: ")
print(avg_marks_by_gender)
```

Console Output:

```
R 4.4.1 ~\j
6 Ravi 23 M 60 Fail
> #d. using group_by() and summarise(): Average Marks by Gender
> avg_marks_by_gender <- students %>%
+   group_by(Gender) %>%
+   summarise(Average_Marks = mean(Marks))
>
> print("Average Marks by Gender: ")
[1] "Average Marks by Gender: "
> print(avg_marks_by_gender)
# A tibble: 2 x 2
  Gender Average_Marks
  <chr>      <dbl>
1 F         90.7
2 M         75.3
```