

# SHETH L.U.J. AND SIR M.V. COLLEGE

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

The screenshot shows the RStudio interface with the following details:

- Console:** Displays R code and its output. It shows the installation of 'stringr' and 'tidyverse' packages, followed by the loading of a CSV file named "Synthetic\_Transportation\_Dataset\_Expanded\_v2.csv".
- Environment:** Shows the global environment with objects like region\_filter, robot\_related, selected\_cols, sleep\_df, special\_buildings, split\_matrix, starts\_with\_s, tidy\_transport, and transport\_df.
- Files:** Shows a file tree under "Home" containing various R scripts (e.g., Practical\_No\_9.R, Practical\_No\_9.R.R), a Microsoft Word document, and a PDF report.
- System:** Shows the taskbar with various application icons and system status.

The screenshot shows the RStudio interface with the following details:

- Console:** Displays continued R code for text manipulation, including the use of str\_sub() and str\_split() functions to extract specific columns from the dataset.
- Environment:** Shows the global environment with objects like region\_filter, robot\_related, selected\_cols, sleep\_df, special\_buildings, split\_matrix, starts\_with\_s, tidy\_transport, and transport\_df.
- Files:** Shows a file tree under "Home" containing various R scripts (e.g., Practical\_No\_9.R, Practical\_No\_9.R.R), a Microsoft Word document, and a PDF report.
- System:** Shows the taskbar with various application icons and system status.

NAME:- CHETAN MANDAVKAR  
ROLL NO. S093  
SUBJECT:- Data Analysis with SAS / SPSS /R

# SHETH L.U.J. AND SIR M.V. COLLEGE

The screenshot shows the RStudio interface. The left pane displays an R script with several lines of code related to data manipulation and splitting datasets. The right pane shows a file browser with a list of files in the 'Home' directory, including CSV files, a Word document, and a PPTX file. The bottom taskbar shows system icons and the date/time.

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Source
Console Terminal Background Jobs
R 4.5.2 - ~/d
2 09-06-2019 08:01 09 06 2019 08
3 08-01-2020 01:38 08 01 2020 01
4 18-10-2022 05:43 18 10 2022 05
5 29-06-2023 23:11 29 06 2023 23
6 02-03-2020 15:55 02 03 2020 15
> split_matrix <- str_split(transport_df$weather, " ", simplify = TRUE)
> transport_df$weather_Main <- split_matrix[, 1]
> if (ncol(split_matrix) == 2) {
+   transport_df$weather_Detail <- split_matrix[, 2]
+ } else {
+   transport_df$weather_Detail <- NA
+ }
> print(" --- After Safe weather Split ---")
[1] " --- After Safe weather Split ---"
> print(transport_df %>% select(weather, weather_Main, weather_Detail) %>% head())
# A tibble: 6 × 3
  Weather Weather_Main Weather_Detail
  <chr>    <chr>        <chr>
1 Fog      Fog          NA
2 Rain     Rain          NA
3 Storm    Storm         NA
4 Clear   Clear         NA
5 Storm    Storm         NA
6 Snow    Snow          NA
> tidy_transport <- transport_df %>%
+   separate(Timestamp, into = c("Date", "Time"), sep = " ")
> print(" --- Using separate() to split Timestamp into Date & Time ---")
[1] " --- Using separate() to split Timestamp into Date & Time ---"
> print(tidy_transport %>% select(Date, Time) %>% head())
# A tibble: 6 × 2
  Date      Time
  <chr>    <chr>
1 03-10-2023 09:50
2 09-06-2019 08:01
3 08-01-2020 01:38
4 18-10-2022 05:43
5 29-06-2023 23:11
6 02-03-2020 15:55
>
>
```

Environment History Connections Tutorial  
Import Dataset 6 MB Global Environment  
region\_filter 3278 obs. of 6 variables  
robot\_related 1 obs. of 3 variables  
selected\_cols 20000 obs. of 4 variables  
sleep\_df 62 obs. of 8 variables  
special\_buildings 3007 obs. of 6 variables  
split\_matrix chr [1:5500, 1] "Fog" "Rain" "Storm" "Clear" "Storm" "S...  
starts\_with\_s 20000 obs. of 3 variables  
tidy\_transport 5500 obs. of 24 variables  
transport\_df 5500 obs. of 23 variables

Files Plots Packages Help Viewer Presentation  
New Folder New File Delete Rename More  
Home  
Name Size Modified  
Practical\_No\_5\_R.R 672 B Nov 24, 2023, 6:55 PM  
Processed\_Student\_Mental\_Health.csv 6.1 KB Nov 18, 2025, 10:58 AM  
S093 CHETAN MANDAVKAR FEATURE ENGINEERING SUBMISSION - Colab 2.3 MB Jul 27, 2025, 4:18 PM  
S093\_Chetan\_Mandavkar\_FP\_Soft\_Copy\_Documentation\_1[1].docx 45.1 KB Oct 15, 2025, 8:30 PM  
SHRUTI MANDAVKAR EVS REPORT.pdf 357.3 KB Nov 21, 2025, 7:34 PM  
Structure and functions of pancreas Shruti Mandavkar.pptx 639.7 KB Aug 24, 2025, 11:35 PM  
Virtual Machines  
WhatsApp Image 2025-11-25 at 19.51.57\_b9447da.jpg 17.9 KB Nov 25, 2025, 7:52 PM  
Attendance\_Prediction.csv 1.6 MB Dec 1, 2025, 10:48 AM  
product\_sales\_dataset\_final.csv 25.3 MB Dec 1, 2025, 10:52 AM  
Practical\_No\_6\_R.R 696 B Dec 1, 2025, 11:07 AM  
synthetic\_fraud\_dataset.csv 847.7 KB Dec 1, 2025, 11:09 AM  
Practical\_No\_7\_R.R 1.1 KB Dec 1, 2025, 11:49 AM  
dataset\_2191\_sleep.csv 1.7 KB Dec 1, 2025, 11:39 AM  
Practical\_No\_8\_R.R 1.2 KB Dec 1, 2025, 11:51 AM  
Synthetic\_Transportation\_Dataset\_Expanded\_v2.csv 596.6 KB Dec 1, 2025, 12:10 PM  
Practical\_No\_9\_R.R 1.3 KB Dec 1, 2025, 12:23 PM

ENG IN 12:29 01-12-2025

NAME:- CHETAN MANDAVKAR  
ROLL NO. S093  
SUBJECT:- Data Analysis with SAS / SPSS /R