

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

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

RStudio

```
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R4.5.2 ~/ ...
> library(dplyr)
> library(tidyverse)
> df <- read.csv("Student Stress Factors (2).csv",
+   na.strings = c("", "NA"),
+   check.names = FALSE)
> names(df) <- make.names(names(df))
> print(names(df))
[1] "Kindly.Rate.your.Sleep.Quality..."
[2] "How.many.times.a.week.do.you.suffer.headaches...."
[3] "How.would.you.rate.your.academic.performance....."
[4] "How.would.you.rate.your.study.load."
[5] "How.many.times.a.week.you.practice.extracurricular.activities...."
[6] "How.would.you.rate.your.stress.levels."
> df_clean <- df %>%
+   mutate(
+     Kindly.Rate.your.Sleep.Quality... =
+       replace_na(Kindly.Rate.your.Sleep.Quality..., 0),
+     How.many.times.a.week.do.you.suffer.headaches.... =
+       replace_na(How.many.times.a.week.do.you.suffer.headaches..., 0),
+     How.would.you.rate.your.academic.performance..... =
+       replace_na(How.would.you.rate.your.academic.performance..., 0),
+     how.would.you.rate.your.study.load. =
+       replace_na(How.would.you.rate.your.study.load., 0),
+     How.many.times.a.week.you.practice.extracurricular.activities.... =
+       replace_na(How.many.times.a.week.you.practice.extracurricular.activities..., 0),
+     How.would.you.rate.your.stress.levels. =
+       replace_na(How.would.you.rate.your.stress.levels., 0)
+   )
> print("--- Cleaned Dataset (First Rows) ---")
[1] "--- Cleaned Dataset (First Rows) ---"
> print(head(df_clean))
Kindly.Rate.your.Sleep.Quality... How.many.times.a.week.do.you.suffer.headaches...
1 3 1
2 4 1
3 2 1
4 3 2
5 2 3
6 3 1
How.would.you.rate.your.academic.performance..... how.would.you.rate.your.study.load.
1 3 4
2 2 3
3 2 1
4 3 2
5 1 5
6 3 2
How.many.times.a.week.you.practice.extracurricular.activities...
1 2
2 3
3 4
4 3
5 5
6 1
How.would.you.rate.your.stress.levels.
1 3
2 2
3 4
4 3
5 3
6 1
> df_calc <- df_clean %>%
+   mutate(
+     Sleep_Health_Score =
+       Kindly.Rate.your.Sleep.Quality... -
+       How.would.you.rate.your.stress.levels.
+   )
> print("--- Method A: Sleep Health Score ---")
[1] "--- Method A: Sleep Health Score ---"
```

Environment History Connections Tutorial

df_text 520 obs. of 7 variables

dropped_multiple 20000 obs. of 13 variables

dropped_one 20000 obs. of 14 variables

dropped_range 20000 obs. of 5 variables

energy 5000 obs. of 6 variables

final_data 200002 obs. of 14 variables

final_dataset 520 obs. of 9 variables

Food_Delivery_Route_ 200 obs. of 10 variables

high_cost 1101 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Home

30°C Sunny

ENG IN 12:48 01-12-2025

RStudio

```
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R4.5.2 ~/ ...
+ )
> print("--- Cleaned Dataset (First Rows) ---")
[1] "--- Cleaned Dataset (First Rows) ---"
> print(head(df_clean))
Kindly.Rate.your.Sleep.Quality... How.many.times.a.week.do.you.suffer.headaches...
1 3 1
2 4 1
3 2 1
4 3 2
5 2 3
6 3 1
How.would.you.rate.your.academic.performance..... how.would.you.rate.your.study.load.
1 3 4
2 2 3
3 2 1
4 3 2
5 1 5
6 3 2
How.many.times.a.week.you.practice.extracurricular.activities...
1 2
2 3
3 4
4 3
5 5
6 1
How.would.you.rate.your.stress.levels.
1 3
2 2
3 4
4 3
5 3
6 1
> df_calc <- df_clean %>%
+   mutate(
+     Sleep_Health_Score =
+       Kindly.Rate.your.Sleep.Quality... -
+       How.would.you.rate.your.stress.levels.
+   )
> print("--- Method A: Sleep Health Score ---")
[1] "--- Method A: Sleep Health Score ---"
```

Environment History Connections Tutorial

df_text 520 obs. of 7 variables

dropped_multiple 20000 obs. of 13 variables

dropped_one 20000 obs. of 14 variables

dropped_range 20000 obs. of 5 variables

energy 5000 obs. of 6 variables

final_data 200002 obs. of 14 variables

final_dataset 520 obs. of 9 variables

Food_Delivery_Route_ 200 obs. of 10 variables

high_cost 1101 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Home

30°C Sunny

ENG IN 12:48 01-12-2025

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

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

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
> print("")
[1] "--> Method A: Sleep Health Score --"
> df_calc %>%
+   select(
+     Kindly.Rate.your.Sleep.Quality.,
+     How.would.you.rate.your.stress.levels.,
+     Sleep_Health_Score
+   ) %>% head()
+ )
Kindly.Rate.your.Sleep.Quality... How.would.you.rate.your.stress.levels. Sleep_Health_Score
1 3 3 0
2 4 2 2
3 2 4 -2
4 3 3 0
5 2 3 -1
6 3 1 2
> df_logic <- df_clean %>%
+   mutate(
+     Stress_Category = ifelse(
+       How.would.you.rate.your.stress.levels. > 7,
+       "High Stress", "Normal"
+     ),
+     StudyLoad_Category = ifelse(
+       how.would.you.rate.your.study.load. > 6,
+       "Heavy Study Load", "Light Study Load"
+     )
+ )
> print("")
[1] "--> Method B: Conditional Labels --"
> print(
+   df_logic %>%
+     select(
+       How.would.you.rate.your.stress.levels.,
+       Stress_Category,
+       how.would.you.rate.your.study.load.,
+       StudyLoad_Category
+     ) %>% head()
+ )
How.would.you.rate.your.stress.levels. Stress_Category how.would.you.rate.your.study.load.
```

Environment History Connections Tutorial

df_text 520 obs. of 7 variables
dropped_multiple 20000 obs. of 13 variables
dropped_one 20000 obs. of 14 variables
dropped_range 20000 obs. of 5 variables
energy 5000 obs. of 6 variables
final_data 200002 obs. of 14 variables
final_dataset 520 obs. of 9 variables
Food_Delivery_Route... 200 obs. of 10 variables
high_cost 1101 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Home

Name	Size	Modified
S093_CHEATAN_MANDAVKAR FEATURE ENGINEERING SUBMISSION - Code	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 AM
WhatsApp Image 2025-11-25 at 19.51.57_2b9447da.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	696.8 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
Practical_No_10_R.R	2.8 KB	Dec 1, 2025, 12:45 PM
Student Stress Factors (2).csv	6.9 KB	Dec 1, 2025, 12:34 PM

ENG IN WiFi 01-12-2025 12:49

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
> print("")
[1] "--> Method B: Conditional Labels --"
> print(
+   df_logic %>%
+     select(
+       How.would.you.rate.your.stress.levels.,
+       Stress_Category,
+       how.would.you.rate.your.study.load.,
+       StudyLoad_Category
+     ) %>% head()
+ )
How.would.you.rate.your.stress.levels. Stress_Category how.would.you.rate.your.study.load.
1 3 Normal 4
2 2 Normal 3
3 4 Normal 1
4 3 Normal 2
5 3 Normal 5
6 1 Normal 2
StudyLoad_Category
1 Light Study Load
2 Light Study Load
3 Light Study Load
4 Light Study Load
5 Light Study Load
6 Light Study Load
> df_text <- df_clean %>%
+   mutate(
+     Student_Summary = paste(
+       "Sleep Quality:", Kindly.Rate.your.Sleep.Quality.,
+       "I Stress:", How.would.you.rate.your.stress.levels.,
+       "I Headaches/week:", How.many.times.a.week.do.you.suffer.headaches....
+     )
+ )
> print("")
[1] "--> Method C: Text Transformation --"
> print(head(df_text$Student_Summary))
[1] "Sleep Quality: 3 | Stress: 3 | Headaches/week: 1"
[2] "Sleep Quality: 4 | Stress: 2 | Headaches/week: 1"
[3] "Sleep Quality: 2 | Stress: 4 | Headaches/week: 1"
```

Environment History Connections Tutorial

df_text 520 obs. of 7 variables
dropped_multiple 20000 obs. of 13 variables
dropped_one 20000 obs. of 14 variables
dropped_range 20000 obs. of 5 variables
energy 5000 obs. of 6 variables
final_data 200002 obs. of 14 variables
final_dataset 520 obs. of 9 variables
Food_Delivery_Route... 200 obs. of 10 variables
high_cost 1101 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Home

Name	Size	Modified
S093_CHEATAN_MANDAVKAR FEATURE ENGINEERING SUBMISSION - Code	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 AM
WhatsApp Image 2025-11-25 at 19.51.57_2b9447da.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	696.8 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
Practical_No_10_R.R	2.8 KB	Dec 1, 2025, 12:45 PM
Student Stress Factors (2).csv	6.9 KB	Dec 1, 2025, 12:34 PM

ENG IN WiFi 01-12-2025 12:49

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

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

Screenshot of RStudio showing the console output of a script named 'Final_Dataset_Output.R'. The script prints a summary of the final dataset and its head. It also prints a summary of the 'Sleep_Health_Score' dataset.

```
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 ~/ ...
> print("--- Final Dataset Output ---")
[1] "--- Final Dataset Output ---"
> print(head(final_dataset))
Kindly.Rate.your.Sleep.Quality... How.many.times.a.week.do.you.suffer.headaches...
1 3 1
2 4 1
3 2 1
4 3 2
5 2 3
6 3 1
How.would.you.rate.your.academic.performance..... how.would.you.rate.your.study.load.
1 3 4
2 2 3
3 2 1
4 3 2
5 1 5
6 3 2
How.many.times.a.week.you.practice.extracurricular.activities...
1 2
2 3
3 4
4 3
5 5
6 1
How.would.you.rate.your.stress.levels. Sleep_Health_Score Good_Sleeper
1 3 0 FALSE
2 2 2 TRUE
3 4 -2 FALSE
4 3 0 FALSE
5 3 -1 FALSE
6 1 2 TRUE
>
Summary
1 Performance=3 | Study Load=4
2 Performance=2 | Study Load=3
3 Performance=2 | Study Load=1
4 Performance=3 | Study Load=2
5 Performance=1 | Study Load=5
6 Performance=3 | Study Load=2
```

Environment History Connections Tutorial

df_text 520 obs. of 7 variables
dropped_multiple 200000 obs. of 13 variables
dropped_one 200000 obs. of 14 variables
dropped_range 200000 obs. of 5 variables
energy 5000 obs. of 6 variables
final_data 200002 obs. of 14 variables
final_dataset 520 obs. of 9 variables
Food_Delivery_Route_ 200 obs. of 10 variables
high_cost 1101 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Home

Name	Size	Modified
S093_CHEIAN MANDAVKAR FEATURE ENGINEERING SUBMISSION - Colab	2.3 MB	Jul 27, 2025, 4:18 PM
S093_Chetan_Mandavkar_F_P_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 AM
Virtual Machines		
WhatsApp Image 2025-11-25 at 19.51.57_2b9447db.jpg	17.9 KB	Nov 25, 2025, 7:52 PM
Attendance_Prediction	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
Practical_No_10.R.R	2.8 KB	Dec 1, 2025, 12:45 PM
Student Stress Factors (2).csv	6.9 KB	Dec 1, 2025, 12:34 PM

ENG IN WiFi 01-12-2025 12:49

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