

SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS /R

Practical No: 2

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

1) Importing CSV(.csv) file:

```

File Edit Code View Plots Session Build Debug Profile Tools Help
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Daily.Coffee.Intake.vs.Sleep.Duration
ID Age Gender Country Coffee_ Intake Caffeine_mg Sleep_Hours Sleep_Quality BMI Heart_Rate Str
1 1 40 Male Germany 3.5 328.1 7.5 Good 24.9 78 1
2 2 33 Male Germany 1.0 94.1 6.2 Good 20.0 67 1
3 3 42 Male Brazil 5.3 503.7 5.9 Fair 22.7 59 1
4 4 53 Male Germany 2.6 249.2 7.3 Good 24.7 71 1
5 5 32 Female Spain 3.1 298.0 5.3 Fair 24.1 76 1
6 6 32 Male Mexico 3.4 326.4 6.4 Good 27.0 82 1
7 7 53 Male France 2.7 252.1 7.8 Good 24.3 58 1
8 8 44 Female Canada 4.5 423.5 5.5 Fair 15.8 62 1
9 9 29 Male UK 1.7 162.0 7.1 Good 21.7 60 1
10 10 41 Female Switzerland 4.0 383.2 6.4 Good 30.4 69 1
11 11 29 Female Switzerland 4.5 427.5 8.1 Excellent 21.5 66 1
12 12 29 Male Netherlands 4.1 385.2 6.5 Good 23.1 80 1
13 13 37 Female Italy 4.7 441.9 6.9 Good 15.4 77 1
Showing 1 to 13 of 10,000 entries. 16 total columns
Console Terminal Background Jobs
R - R 4.5.2 - ~/ 
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, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
> daily.Coffee.Intake.vs.Sleep.Duration <- read.csv("~/datasets/Daily_Coffee_Intake_vs_Sleep Duration.csv")
> View(Daily.Coffee.Intake.vs.Sleep.Duration)
> View(Daily.Coffee.Intake.vs.Sleep.Duration)
> |
```

Select File to Import

File name: Daily Coffee Intake vs Sleep Duration

All Files (*.*) Open Cancel

11:52:16 AM 11/24/2025

Name	Status	Date modified	Type	Size
cars		7/28/2025 12:24 PM	Microsoft Excel Co...	2 KB
Daily Coffee Intake vs Sleep Duration		8/25/2025 11:46 AM	Microsoft Excel Co...	733 KB
heart		7/28/2025 11:48 AM	Microsoft Excel Co...	36 KB
House Price Prediction		9/21/2025 11:57 AM	Microsoft Excel Co...	308 KB
insurance_dataset		9/20/2025 8:58 PM	Microsoft Excel Co...	37 KB
international-airline-passengers		8/12/2025 9:16 AM	Microsoft Excel Co...	2 KB
iris		8/11/2025 10:51 AM	Microsoft Excel Co...	4 KB
penguins		8/11/2025 12:09 PM	Microsoft Excel Co...	19 KB
Student Mental health		11/13/2025 8:13 AM	Microsoft Excel Co...	8 KB
study hours vs student scores		8/25/2025 11:46 AM	Microsoft Excel Co...	733 KB
titanic		8/4/2025 11:05 AM	Microsoft Excel Co...	60 KB

The file you are attempting to preview could harm your computer. If you trust the file and the source you received it from, open it to view its contents.

11:52:32 AM 11/24/2025

SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS / R

2) Importing Excel(.xlsx) file:

SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS / R

Choose File

Documents > datasets

Name	Status	Date modified	Type	Size
cars		7/28/2025 12:24 PM	Microsoft Excel Co...	2 KB
Daily Coffee Intake vs Sleep Duration(excel version)		11/24/2025 12:07 PM	Microsoft Excel W...	805 KB
Daily Coffee Intake vs Sleep Duration(txt version)		11/24/2025 12:10 PM	Text Document	731 KB
Daily Coffee Intake vs Sleep Duration		8/25/2025 11:46 AM	Microsoft Excel Co...	733 KB
heart		7/28/2025 11:48 AM	Microsoft Excel Co...	36 KB
House Price Prediction		9/21/2025 11:57 AM	Microsoft Excel Co...	308 KB
insurance_dataset		9/20/2025 8:58 PM	Microsoft Excel Co...	37 KB
international-airline-passengers		8/12/2025 9:16 AM	Microsoft Excel Co...	2 KB
iris		8/11/2025 10:51 AM	Microsoft Excel Co...	4 KB
penguins		8/11/2025 12:09 PM	Microsoft Excel Co...	19 KB
Student Mental health		11/13/2025 8:13 AM	Microsoft Excel Co...	8 KB
study hours vs student scores		8/25/2025 11:46 AM	Microsoft Excel Co...	733 KB
titanic		8/4/2025 11:05 AM	Microsoft Excel Co...	60 KB

Daily Coffee Intake

All Files (*.*) Open Cancel

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Import Excel Data

ID Age

1 1.40 Male
2 2.33 Male
3 3.42 Male
4 4.53 Male
5 5.32 Fem
6 6.32 Male
7 7.53 Male
8 8.44 Fem
9 9.29 Male
10 10.41 Fem
11 11.29 Fem
12 12.29 Ma
13 13.37 Fem

Data Preview:

ID	Age	Gender	Country	Coffee_	Caffeine_	Sleep_	Sleep_Quality	BMI	Heart_Rate	Stress_Level	Physical_Activity_Hour
1	40	Male	Germany	3.5	328.1	7.5	Good	24.9	78	Low	
2	33	Male	Germany	1.0	94.1	6.2	Good	20.0	67	Low	
3	42	Male	Brazil	5.3	503.7	5.9	Fair	22.7	59	Medium	
4	53	Male	Germany	2.6	249.2	7.3	Good	24.7	71	Low	
5	32	Female	Spain	3.1	298.0	5.3	Fair	24.1	76	Medium	
6	32	Male	Mexico	3.4	326.4	6.4	Good	27.0	82	Low	
7	53	Male	France	2.7	252.1	7.8	Good	24.3	58	Low	
8	44	Female	Canada	4.5	423.5	5.5	Fair	15.8	62	Medium	
9	29	Male	UK	1.7	162.0	7.1	Good	21.7	60	Low	
10	41	Female	Switzerland	4.0	383.2	6.4	Good	30.4	69	Low	
11	29	Female	Switzerland	4.5	427.5	8.1	Good	21.5	66	Low	

Showing 1 to 11

Console

```
R - R 4.5.2
Run r.rat
```

Import Options:

Name: Daily_Coffee_Intake_vs_Sleep_L Max Rows: First Row as Names:

Sheet: Default Skip: 0 Open Data Viewer:

Range: A1:D10 NA:

Code Preview:

```
library(readxl)
Daily_Coffee_Intake_vs_Sleep_Duration_excel_version_ <- read_excel("datasets/Daily_Coffee_Intake_vs_Sleep_Duration(excel version).xlsx")
View(Daily_Coffee_Intake_vs_Sleep_Duration_excel_version_)
```

Import Cancel

algorithm practical 5.py 1.1 kB Feb 7, 2025, 9:56 PM
algorithm practical no 4.pdf 224.9 kB Feb 21, 2025, 1:58 PM
algorithm practical no 3.py 322.9 kB Feb 21, 2025, 1:58 PM

SHETH LUJ AND SIR MV 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

Daily_Coffee_... vs Sleep.Duration

Daily.Coffee.Intake.vs.Sleep.Duration

Environment History Connections Tutorial

R Global Environment

Data

Daily_Coffee_... 10000 obs. of 16 variables

- \$ ID : num [1:10000] 1 2 3 4 5 6 7 8 9 10 ...
- \$ Age : num [1:10000] 40 33 42 53 32 32 53 44 29 41 ...
- \$ Gender : chr [1:10000] "Male" "Male" "Male" "Male" ...
- \$ Country : chr [1:10000] "Germany" "Germany" "Brazil" "Germany" ...
- \$ Coffee_Intake : num [1:10000] 3.5 1 5.3 2.6 3.1 3.4 2.7 4.5 1.7 4 ...
- \$ Caffeine_mg : num [1:10000] 328.1 94.1 503.7 249.2 298.4 ...
- \$ Sleep_Hours : num [1:10000] 7.5 6.2 5.9 7.3 5.3 6.4 7.8 5.5 7.1 ...
- \$ Sleep_Quality : chr [1:10000] "Good" "Good" "Fair" "Good" ...

Files Plots Packages Help Viewer Presentation

Console Terminal Background Jobs

```
R 4.5.2 --> Physical_Activity_Hours 12 10000 7.49 4.32 7.5 7.48 5.49 0 15.0 15.0
Health_Issues* 13 10000 2.24 0.95 3.0 2.30 0.00 1 4.0 3.0
Occupation* 14 10000 2.99 1.41 3.0 2.99 1.48 1 5.0 4.0
Smoking 15 10000 0.20 0.40 0.0 0.13 0.00 0 1.0 1.0
Alcohol_Consumption 16 10000 0.30 0.46 0.0 0.25 0.00 0 1.0 1.0
> skew kurtosis se
ID 0.00 -1.20 28.87
Age 0.36 -0.35 0.11
Gender* 0.33 -1.05 0.01
Country* 0.01 -1.21 0.06
Coffee_Intake 0.26 -0.28 0.01
Caffeine_mg 0.26 -0.28 1.38
Sleep_Hours -0.02 -0.14 0.01
Sleep_Quality* -0.59 -0.29 0.01
BMI 0.05 -0.17 0.04
Heart_Rate 0.10 -0.26 0.10
Stress_Level* 0.08 -0.28 0.01
Physical_Activity_Hours 0.00 -1.20 0.04
Health_Issues* -0.48 -1.70 0.01
Occupation* 0.02 -1.29 0.01
Smoking 1.50 0.24 0.00
Alcohol_Consumption 0.87 -1.24 0.00
> View(Daily_Coffee_..._version_)
```

7:53:23 PM 11/24/2025

3) Importing text(txt)file:

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Daily_Coffee_... vs Sleep.Duration

Daily.Coffee.Intake.vs.Sleep.Duration

Environment History Connections Tutorial

R Global Environment

Data

Daily_Coffee_... 10000 obs. of 16 variables

- \$ ID : num [1:10000] 1 2 3 4 5 6 7 8 9 10 ...
- \$ Age : num [1:10000] 40 33 42 53 32 32 53 44 29 41 ...
- \$ Gender : chr [1:10000] "Male" "Male" "Male" "Male" ...
- \$ Country : chr [1:10000] "Germany" "Germany" "Brazil" "Germany" ...
- \$ Coffee_Intake : num [1:10000] 3.5 1 5.3 2.6 3.1 3.4 2.7 4.5 1.7 4 ...
- \$ Caffeine_mg : num [1:10000] 328.1 94.1 503.7 249.2 298.4 ...
- \$ Sleep_Hours : num [1:10000] 7.5 6.2 5.9 7.3 5.3 6.4 7.8 5.5 7.1 ...
- \$ Sleep_Quality : chr [1:10000] "Good" "Good" "Fair" "Good" ...

Files Plots Packages Help Viewer Presentation

Console Terminal Background Jobs

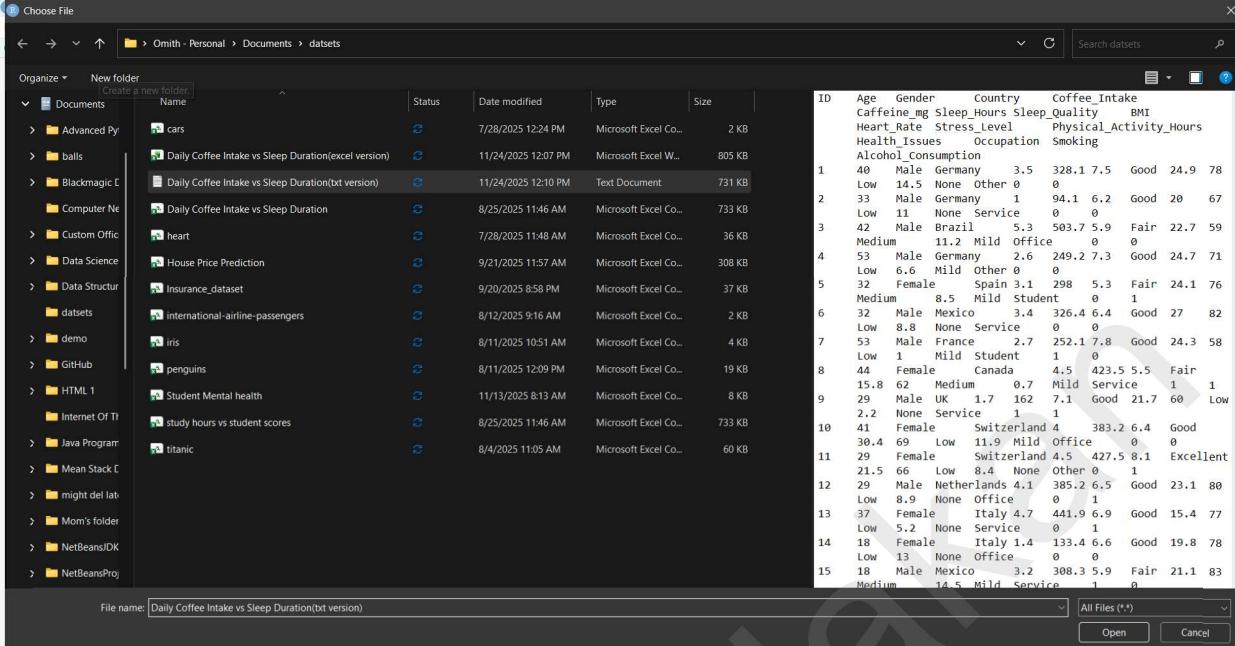
From Text (base)... From Text (read...)

From Excel... From SPSS... From SAS... From Stata... From CSV...

skew kurtosis se
ID 0.00 -1.20 28.87
Age 0.36 -0.35 0.11
Gender* 0.33 -1.05 0.01
Country* 0.01 -1.21 0.06
Coffee_Intake 0.26 -0.28 0.01
Caffeine_mg 0.26 -0.28 1.38
Sleep_Hours -0.02 -0.14 0.01
Sleep_Quality* -0.59 -0.29 0.01
BMI 0.05 -0.17 0.04
Heart_Rate 0.10 -0.26 0.10
Stress_Level* 0.08 -0.28 0.01
Physical_Activity_Hours 0.00 -1.20 0.04
Health_Issues* -0.48 -1.70 0.01
Occupation* 0.02 -1.29 0.01
Smoking 1.50 0.24 0.00
Alcohol_Consumption 0.87 -1.24 0.00
> View(Daily_Coffee_..._version_)

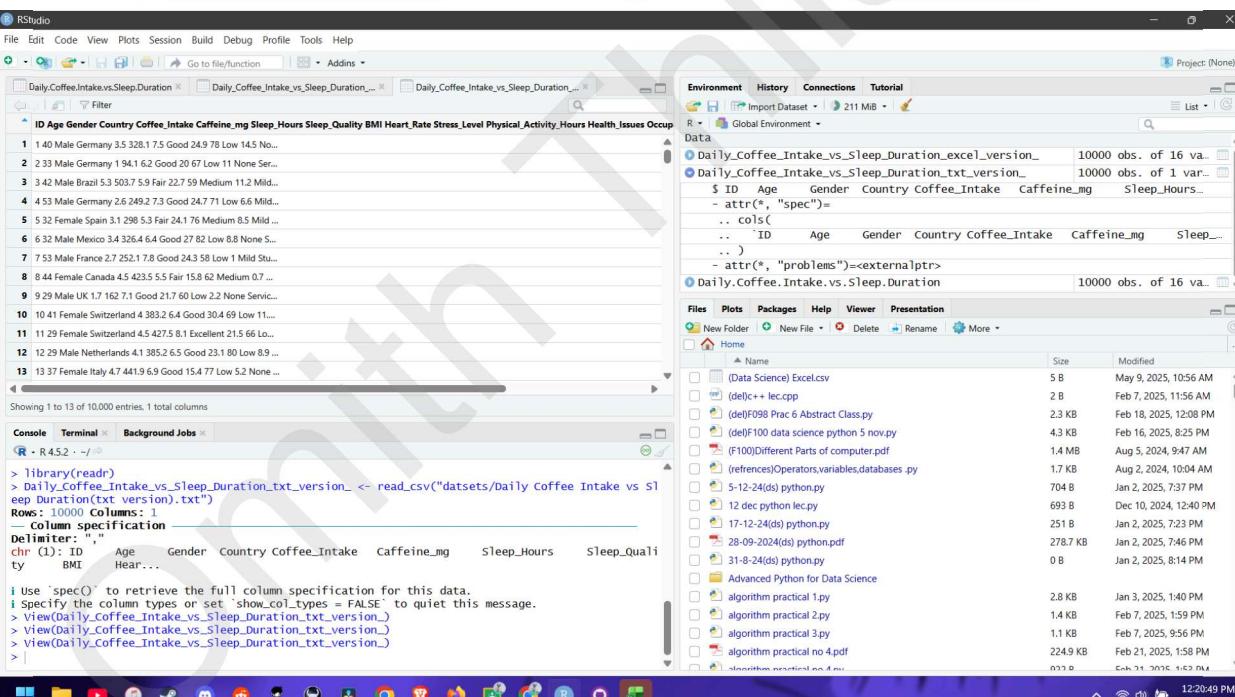
7:54:12 PM 11/24/2025

SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS / R



File name: Daily Coffee Intake vs Sleep Duration(txt version)

Open Cancel



```

library(readr)
Daily_Coffee_Intake_vs_Sleep_Duration_txt_version_ <- read_csv("datasets/Daily_Coffee_Intake_vs_Sleep_Duration(txt version).txt")
Rows: 10000 Columns: 1
--- Column specification ---
Delimiter: ","
chr (1): ID    Age   Gender  Country Coffee_ Intake  Caffeine_mg  Sleep_Hours  Sleep_Quali
ty  BMI   Hear...
Use `spec()` to retrieve the full column specification for this data.
I Specify the column types or set `show_col_types = FALSE` to quiet this message.
> View(Daily_Coffee_Intake_vs_Sleep_Duration_txt_version_)
> View(Daily_Coffee_Intake_vs_Sleep_Duration_txt_version_)
> View(Daily_Coffee_Intake_vs_Sleep_Duration_txt_version_)
> View(Daily_Coffee_Intake_vs_Sleep_Duration_txt_version_)

```