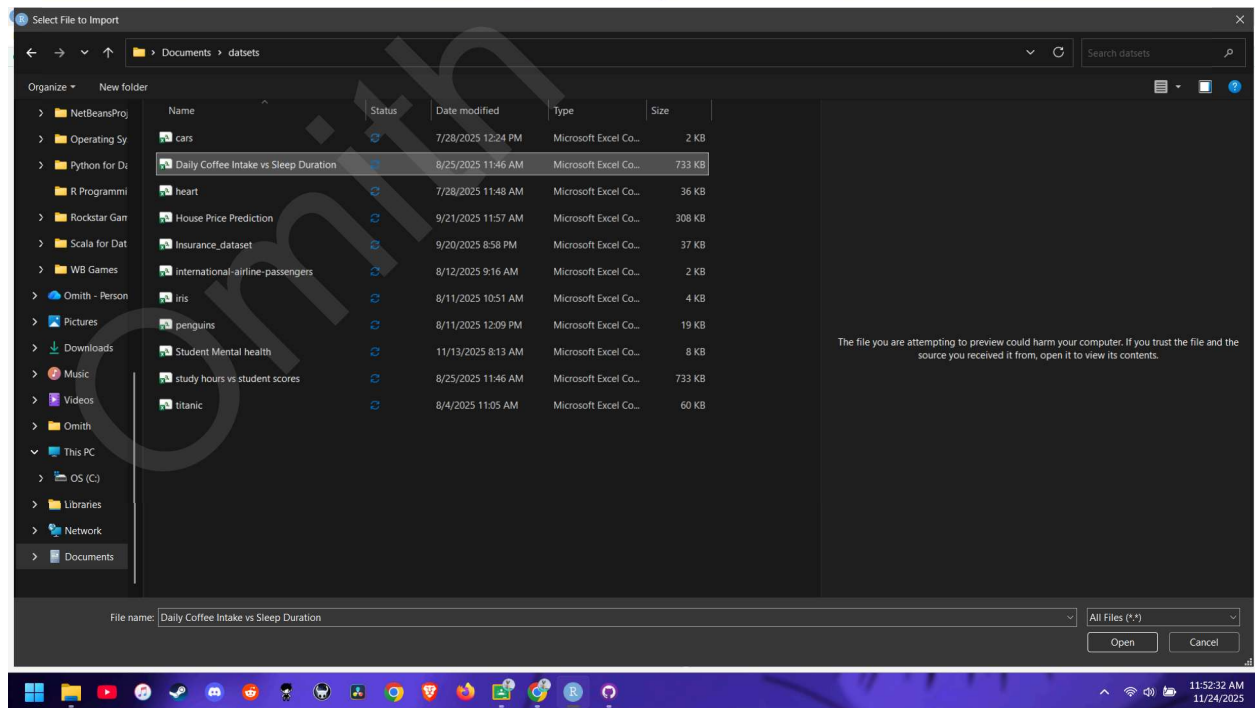


Practical No: 2

1) Importing CSV(.csv) file:



SHETH LUJ AND SIR MV COLLEGE

Subject: Data Analysis with SAS / SPSS / R

RStudio interface showing a dataset named "Daily.Coffee.Intake.vs.Sleep.Duration" with 13 columns: ID, Age, Gender, Country, Coffee_Intake, Caffeine_mg, Sleep_Hours, Sleep_Quality, BMI, and Heart_Rate. The console shows R code for reading the CSV file and viewing its structure.

```
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)  
>
```

2) Importing Excel(.xlsx) file:

RStudio interface showing the process of importing an Excel file. The console shows R code for installing the "readr" and "psych" packages, reading the Excel file, and describing its structure. The Environment pane shows the imported data as a tibble.

```
R - R 4.5.2 - ~/ -  
install.packages(c("readr", "psych"))  
library(readr)  
library(psych)  
Caffeine_mg <- read.csv("Daily.Coffee.Intake.vs.Sleep.Duration.csv")  
head(Daily.Coffee.Intake.vs.Sleep.Duration)  
tail(Daily.Coffee.Intake.vs.Sleep.Duration)  
dim(Daily.Coffee.Intake.vs.Sleep.Duration)  
> library(readr)  
> library(psych)  
> describe(Daily.Coffee.Intake.vs.Sleep.Duration)  
vars      n mean      sd median trimmed      mad min      max range  
ID      1 10000 5000.50 2886.90 5000.5 5000.50 3706.50 1 10000.0 9999.0  
Age      2 10000 34.95 11.16 34.0 34.48 11.86 18 80.0 62.0  
Gender*  3 10000 1.52 0.54 1.0 1.50 0.00 1 3.0 2.0  
Country* 4 10000 10.47 5.74 10.0 10.47 7.41 1 20.0 19.0  
Coffee_Intake 5 10000 2.51 1.45 2.5 2.47 1.48 0 8.2 8.2  
Caffeine_mg 6 10000 238.41 137.75 235.4 235.07 143.29 0 780.3 780.3  
Sleep_Hours 7 10000 6.64 1.22 6.6 6.64 1.19 3 10.0 7.0  
Sleep_Quality* 8 10000 2.62 0.84 3.0 2.66 0.00 1 4.0 3.0  
BMI      9 10000 23.99 3.91 24.0 23.99 4.00 15 38.2 23.2  
Heart_Rate 10 10000 70.62 9.82 71.0 70.55 10.38 50 109.0 59.0  
Stress_Level* 11 10000 2.11 0.54 2.0 2.13 0.00 1 3.0 2.0  
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  
ID      0.00 skew kurtosis se  
Age      0.36 -0.35 0.11
```

SHETH LUJ AND SIR MV COLLEGE

Subject: Data Analysis with SAS / SPSS / R

File name: Daily_Coffee_Intake_vs_Sleep_Duration(excel version)

File/URL: ~/datasets/Daily_Coffee_Intake_vs_Sleep_Duration(excel version).xlsx

ID (double)	Age (double)	Gender (character)	Country (character)	Coffee_Intake (double)	Caffeine_mg (double)	Sleep_Hours (double)	Sleep_Quality (character)	BMI (double)	Heart_Rate (double)	Stress_Level (character)	Physical_Activity_Hour (double)
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	Excellent	21.5	66	Low	

Import Options:

Name: Daily_Coffee_Intake_vs_Sleep_I Max Rows: ☒ First Row as Names

Sheet: Default Skips: 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_)
```


SHETH LUJ AND SIR MV COLLEGE

Subject: Data Analysis with SAS / SPSS / R

The screenshot shows the RStudio interface with a dataset named 'Daily_Coffee_Intake_vs_Sleep_Duration' loaded. The dataset has 10,000 observations and 16 variables. The console displays summary statistics for the dataset.

Variable	Mean	SD	Min	Q1	Median	Q3	Max
Physical_Activity_Hours	12.0000	7.49	4.32	7.5	7.48	5.49	15.0
Health_Issues*	13.0000	2.24	0.95	3.0	2.30	0.00	1.0
Occupation*	14.0000	2.99	1.41	3.0	2.99	1.48	1.0
Smoking	15.0000	0.20	0.40	0.0	0.13	0.00	1.0
Alcohol_Consumption	16.0000	0.30	0.46	0.0	0.25	0.00	1.0

The console also displays summary statistics for the variables:

Variable	Mean	SD	Min	Q1	Median	Q3	Max
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				

3) Importing text(txt)file:

The screenshot shows the RStudio interface with the 'Import Dataset' dialog box open. The 'From Text (readr)...' option is selected. The console displays summary statistics for the dataset.

Variable	Mean	SD	Min	Q1	Median	Q3	Max
Physical_Activity_Hours	12.0000	7.49	4.32	7.5	7.48	5.49	15.0
Health_Issues*	13.0000	2.24	0.95	3.0	2.30	0.00	1.0
Occupation*	14.0000	2.99	1.41	3.0	2.99	1.48	1.0
Smoking	15.0000	0.20	0.40	0.0	0.13	0.00	1.0
Alcohol_Consumption	16.0000	0.30	0.46	0.0	0.25	0.00	1.0

The console also displays summary statistics for the variables:

Variable	Mean	SD	Min	Q1	Median	Q3	Max
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				

SHETH LUJ AND SIR MV COLLEGE

Subject: Data Analysis with SAS / SPSS / R

The image shows a Windows file explorer window titled 'Choose File' with the path 'Omith - Personal > Documents > datasets'. It lists various files, including 'Daily Coffee Intake vs Sleep Duration(excel version)' and 'Daily Coffee Intake vs Sleep Duration(txt version)'. The 'Daily Coffee Intake vs Sleep Duration(txt version)' file is selected.

Below the file explorer, the RStudio interface is visible. The 'Environment' pane shows the loaded dataset 'Daily_Coffee_Intake_vs_Sleep_Duration' with 10000 observations and 16 variables. The 'Files' pane shows the file structure of the dataset, including 'Daily_Coffee_Intake_vs_Sleep_Duration_excel_version' and 'Daily_Coffee_Intake_vs_Sleep_Duration_txt_version'.

The console shows the following R code and output:

```
> 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: 16
Column specification
Delimiter: ","
chr (1): ID Age Gender Country Coffee_Intake Caffeine_mg Sleep_Hours Sleep_Qual...
ty BMI Heart_Rate Stress_Level Physical_Activity_Hours Health_Issues Occupa...

I 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_)
>
```