

MVLU COLLEGE

Aim:

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

EXCEL

The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays the R command line output for version 4.1.2, showing the standard R welcome message and workspace loading.
- File/URL Tab:** Shows the path to the Excel file: C:/Users/mvluc/OneDrive/Desktop/ANSHRAH/DATA ANALYSIS WITH SASS SPSS R/CHARACTERS_DATASET.xlsx.
- Data Preview:** A grid view of the first 50 rows of the dataset, which contains information about characters such as Name, Initial_Age, Age, Gender, Gender_Number, Role, Alignment, Combat_Style, Status, Archetype, and Arc_Type.
- Import Options:** A panel for configuring the import:
 - Name: CHARACTERS_DATASET
 - Sheet: Default
 - Range: A1:D10
 - Max Rows: 0
 - Skip: 0
 - First Row as Names:
 - Open Data Viewer:
- Code Preview:** A code editor showing the R code used to import the data:

```
library(readxl)
CHARACTERS_DATASET <- read_excel("c:/users/mvluc/onedrive/desktop/anshrah/data analysis with sass spss r/CHARACTERS_DATASET.xlsx")
View(CHARACTERS_DATASET)
```
- Environment Tab:** Shows the global environment with an empty list.
- File Explorer:** A sidebar showing the local file system with various documents and projects.

ANSHRAH SHAIKH
S111
SYCS
DATA ANALYSIS IN SASS / SPSS / R

MVLU COLLEGE

CHARACTERS_DATASET

NAME	INITIAL_AGE	AGE	GENDER	GENDER_NUMBER	ROLE	ALIGNMENT	COMBAT_STYLE	STATUS	ARCHETYPE	ARC_TYPE
NEVEAH ELIXIR	9	16	Female	1	Protagonist	Good	Ranged	Alive	NA	NA
DAMIAN ESTEBETH	9	16	Male	0	Protagonist	Good	Close	Alive	NA	NA
ELIAS EVANDER	9	16	Male	0	Protagonist	Good	Stealth	Alive	NA	NA
CALE WERENUS	9	16	Male	0	Supporting	Neutral	Close	Alive	NA	NA
ADEN WERENUS	28	35	Male	0	Supporting	Neutral	Close	Alive	NA	NA
IDIRIA WERENUS	30	37	Female	1	Supporting	Neutral	None	Alive	NA	NA
VICTORIA CASABEL	10	17	Female	1	Supporting	Good	Stealth	Alive	NA	NA
JULIAN CASABEL	18	25	Male	0	Supporting	Good	Ranged	Missing	NA	NA
JULIA CASABEL	39	46	Female	1	Supporting	Good	Close	Dead	NA	NA
VECTOR CASABEL	41	48	Male	0	Antagonist	Evil	Stealth	Alive	NA	NA
AXEL	9	16	Male	0	Supporting	Good	Stealth	Alive	NA	NA
RAYDENCEYEL EDEVORY	9	16	Female	1	Supporting	Good	Stealth	Alive	NA	NA
RENOX EDEVORY	14	21	Male	1	Supporting	Good	Ranged	Alive	NA	NA
REVANA EDEVORY	19	26	Female	0	Supporting	Good	Close	Dead	NA	NA

Showing 1 to 15 of 33 entries, 11 total columns

Console Terminal Background Jobs

```
R - R 4.1.2 - ~/
```

R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support is 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.

```
8
9  read_excel("C:/users/mvluc/onedrive/Desktop/ANSHRAH/DATA ANALYSIS WITH SASS SPSS R/CHARACTERS_DATASET.xlsx")
10 View(CHARACTERS_DATASET)
11 lm(AGE~GENDER)
12 attach(CHARACTERS_DATASET)
```

11:1 (Top Level) ◊

Console Terminal Background Jobs

```
R - R 4.1.2 - ~/
```

```
> attach(CHARACTERS_DATASET)
> lm(AGE~GENDER)

Call:
lm(formula = AGE ~ GENDER)

Coefficients:
(Intercept)   GENDERMale
            33.786        -6.154
```

MVLU COLLEGE

CSV:

The screenshot shows the 'Import Dataset' dialog in RStudio. The 'Name' field is set to 'student_spending'. The 'Input File' pane displays the first few lines of the CSV data. The 'Encoding' is set to 'Automatic', 'Heading' is 'Yes', 'Row names' is 'Automatic', 'Separator' is 'Comma', 'Decimal' is 'Period', 'Quote' is 'Double ("')', and 'Comment' is 'None'. The 'na.strings' field is set to 'NA'. A checkbox for 'Strings as factors' is checked. Below the dialog, a preview of the 'Data Frame' is shown, displaying columns: X, age, gender, year_in_school, major, monthly_income, financial_aid, tuition, housing, food, transportation, books_supplies, entertainment. The data frame contains 15 rows of student information.

The screenshot shows the RStudio interface with the 'student_spending' dataset loaded. The environment pane shows variables like 'CHARACTERS_DA', 'student_spending', and 'text_data'. The data grid displays the same 15 rows of student data as the import dialog. The bottom console shows the command 'lm(formula = AGE ~ GENDER)' and its coefficients. The status bar at the bottom right indicates the system is at 31°C, sunny, with ENG IN and 11:58 on 24-Nov-2025.

ANSHRAH SHAIKH
S111
SYCS
DATA ANALYSIS IN SASS / SPSS / R

MVLU COLLEGE

RStudio

```

CHARACTERS_DATASET student_spending Untitled1*
1 read.csv("C:/users/mvluc/onedrive/desktop/ANSHRAH/ADVANCE PYTHON FOR DATA SCIENCE/student_spending.csv", stringsAsFactors=TRUE)
2 View(student_spending)
3 lm(age~gender)
4 attach(student_spending)

> attach(student_spending)
> lm(age~gender)

Call:
lm(formula = age ~ gender)

Coefficients:
(Intercept) gendermale genderNon-binary
21.59752      0.07102      0.16260

```

CHARACTERS_DA... 33 obs. of 11 variables
student_spended 1000 obs. of 18 variables

values

	difference	group1	group2	instal.packag...	product	quotient	?
group1	10	num [1:10]	25 25 10 34 30 23 28 29 25 27	"bada"	200	2	?
group2	22	22 24 19 31 27 21 26 28 24 26					

Files Plots Packages Help Viewer Presentation

Home

Name	Size	Modified
RData	237 B	Nov 24, 2025, 10:17 AM
Rhistory	8 B	Nov 24, 2025, 11:17 AM
class printString.java	125 B	Nov 11, 2025, 10:03 AM
Computer Network prat1.docx	1.1 MB	Nov 15, 2025, 9:44 AM
Custom Office Templates		
Deep GIS Prac.docx	488.4 KB	Nov 18, 2025, 9:33 AM
desktop.ini	418 B	Aug 4, 2025, 11:19 AM
GIS DataBase		
GIS Prac.docx	225.7 KB	Nov 18, 2025, 8:41 AM
IISExpress		
My Web Sites		
NetBeansProjects		
Office Tools Practical_1.accdb	604 KB	Oct 11, 2025, 1:32 PM
R		
Virtual Machines		
Visual Studio 2022		

Text:

RStudio

```

CHARACTERS_DATASET student_spending prac2.R text_data

> lm(formula = Score ~ Group)
Call:
lm(formula = Score ~ Group)

Coefficients:
(Intercept) GroupA   GroupB
202.315     10.000    -10.000

```

ID	Name	Age	Gender	Height_cm	Weight_kg	Score	Group	Joined	Active
1	Alice	29	F	165	58	88	A	2022-03-12	TRUE
2	Bob	34	M	178	82	73	B	2021-11-01	FALSE
3	Charlie	22	M	172	70	91	A	2023-01-22	TRUE
4	Diana	27	F	160	55	85	B	2022-06-15	TRUE
5	Edward	31	M	181	90	67	A	2020-08-09	FALSE
6	Fiona	25	F	166	62	95	A	2023-02-17	TRUE
7	George	40	M	175	85	72	B	2019-04-30	FALSE
8	Helen	33	F	170	68	89	A	2021-09-13	TRUE
9	Ian	28	M	182	88	76	B	2022-12-05	TRUE
10	Jenny	24	F	158	52	94	A	2023-05-10	TRUE
11	Karl	30	M	179	79	81	B	2021-03-25	FALSE
12	Lisa	26	F	167	61	90	A	2022-08-04	TRUE
13	Mari	35	M	183	92	66	B	2020-11-19	FALSE
14	Nina	23	F	162	56	93	A	2023-03-28	TRUE
15	Oscar	29	M	176	77	78	B	2021-01-10	TRUE
16	Paula	32	F	171	63	87	A	2022-10-01	TRUE
17	Quinn	27	M	174	73	82	B	2021-06-14	FALSE
18	Rachel	25	F	159	54	96	A	2023-04-19	TRUE
19	Sam	38	M	180	89	69	B	2019-09-07	FALSE
20	Tina	28	F	166	60	92	A	2022-01-23	TRUE

Show 1 to 20 of 20 entries, 10 total columns

```

> lm(formula = Score ~ Group)
Call:
lm(formula = Score ~ Group)

Coefficients:
(Intercept) GroupA   GroupB
202.315     10.000    -10.000

```

CHARACTERS_DA... 33 obs. of 11 variables
student_spended 1000 obs. of 18 variables
text_data 20 obs. of 10 variables

values

	difference	group1	group2	instal.packag...	product	quotient	?
group1	10	num [1:10]	25 25 10 34 30 23 28 29 25 27	"bada"	200	2	?
group2	22	22 24 19 31 27 21 26 28 24 26					

Files Plots Packages Help Viewer Presentation

Home

Name	Size	Modified
RData	237 B	Nov 24, 2025, 10:17 AM
Rhistory	8 B	Nov 24, 2025, 11:17 AM
class printString.java	125 B	Nov 11, 2025, 10:03 AM
Computer Network prat1.docx	1.1 MB	Nov 15, 2025, 9:44 AM
Custom Office Templates		
Deep GIS Prac.docx	488.4 KB	Nov 18, 2025, 9:33 AM
desktop.ini	418 B	Aug 4, 2025, 11:19 AM
GIS DataBase		
GIS Prac.docx	225.7 KB	Nov 18, 2025, 8:41 AM
IISExpress		
My Web Sites		
NetBeansProjects		
Office Tools Practical_1.accdb	604 KB	Oct 11, 2025, 1:32 PM
R		
Virtual Machines		
Visual Studio 2022		

MVLU COLLEGE

The screenshot shows the RStudio IDE interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help, and Addins. The main workspace contains three tabs: 'CHARACTERS_DATASET', 'student_spending', and 'text_data'. The 'prac2.R' tab is currently active, displaying R code for reading CSV and Excel files, attaching datasets, and fitting a linear model. The 'Console' tab shows the execution of this code, including the output of the model coefficients. The right sidebar features the 'Environment' tab, which lists various R objects like 'CHARACTERS_DA', 'student_spending', and 'text_data' along with their details. Below the environment is a file browser showing local files like 'RData', 'Rhistory', and various Microsoft Office documents. The bottom status bar shows system information including the date and time.

ANSHRAH SHAIKH
S111
SYCS
DATA ANALYSIS IN SASS / SPSS / R