

MVLU COLLEGE

Practical no :- 02

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

The screenshot shows the RStudio interface with the 'Student.Mental.health' dataset loaded. The console displays the following code and output:

```
R - R 4.5.2 - ~/R
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.

> student.Mental.health <- read.csv("C:/Users/itlab/downloads/S100/Student Mental health.csv")
> view(student.Mental.health)
```

The environment pane shows the dataset 'Student.Mental.health' with 101 observations and 11 variables. The data table is as follows:

Timestamp	Choose.your.gender	Age	What.is.your.course	Your.current.year.of.Study	What.is.your.CGPA.	Marital.status
8/7/2020 12:02	Female	18	Engineering	year 1	3.00 - 3.49	No
8/7/2020 12:04	Male	21	Islamic education	year 2	3.00 - 3.49	No
8/7/2020 12:05	Male	19	BIT	Year 1	3.00 - 3.49	No
8/7/2020 12:06	Female	22	Laws	year 3	3.00 - 3.49	Yes
8/7/2020 12:13	Male	23	Mathematics	year 4	3.00 - 3.49	No
8/7/2020 12:31	Male	19	Engineering	Year 2	3.50 - 4.00	No
8/7/2020 12:32	Female	23	Pendidikan Islam	year 2	3.50 - 4.00	Yes
8/7/2020 12:33	Female	18	BCS	year 1	3.50 - 4.00	No
8/7/2020 12:35	Female	19	Human Resources	Year 2	2.50 - 2.99	No
8/7/2020 12:39	Male	18	ikhs	year 1	3.50 - 4.00	No
8/7/2020 12:39	Female	20	Psychology	year 1	3.50 - 4.00	No
8/7/2020 12:39	Female	24	Engineering	Year 3	3.50 - 4.00	Yes
8/7/2020 12:40	Female	18	BCS	year 1	3.00 - 3.49	No
8/7/2020 12:41	Male	19	Engineering	year 1	3.00 - 3.49	No
8/7/2020 12:43	Female	18	KENIMS	Year 2	3.50 - 4.00	No
8/7/2020 12:43	Male	24	BCS	Year 3	3.50 - 4.00	No
8/7/2020 12:46	Female	24	Accounting	year 3	3.00 - 3.49	No

The screenshot shows the RStudio interface with the 'sales_data' dataset loaded. The console displays the following code and output:

```
R - R 4.5.2 - ~/R
> library(readr)
> sales_data <- read_csv("C:/Users/itlab/downloads/S100/sales_data.csv")
Rows: 1000 Columns: 14
Column specification
  delimeter: ","
chr (7): Sales_Rep, Region, Product_Category, Customer_Type, Payment_Method, Sales_Channel, R...
dbl (6): Product_ID, Sales_Amount, Quantity_Sold, Unit_Cost, Unit_Price, Discount
date (1): Sale_Date
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(sales_data)
```

The environment pane shows the dataset 'sales_data' with 1000 observations and 14 variables. The data table is as follows:

Product_ID	Sale_Date	Sales_Rep	Region	Sales_Amount	Quantity_Sold	Product_Category	Unit_Cost	Unit_Price	
1	1052	2023-02-03	Bob	North	5053.97	18	Furniture	152.75	261
2	1099	2023-04-21	Bob	West	4384.02	17	Furniture	3816.39	4208
3	1015	2023-09-21	David	South	4631.23	30	Food	261.56	371
4	1072	2023-08-24	Bob	South	2167.94	39	Clothing	4330.03	4461
5	1061	2023-03-24	Charlie	East	3750.20	13	Electronics	637.37	692
6	1021	2023-02-11	Charlie	West	3761.15	32	Food	900.79	1106
7	1083	2023-04-11	Bob	West	618.31	29	Furniture	2408.81	2624
8	1087	2023-01-06	Eve	South	7698.92	46	Furniture	3702.51	3964
9	1075	2023-06-29	David	South	4223.39	30	Furniture	738.06	1091
10	1075	2023-10-09	Charlie	West	8239.58	18	Clothing	2228.35	2682
11	1088	2023-11-16	Eve	North	8518.45	13	Furniture	2440.11	2511
12	1100	2023-08-14	Bob	West	2198.74	43	Food	1100.81	1131
13	1024	2023-11-11	Eve	West	6607.80	21	Food	622.01	641
14	1003	2023-12-31	Alice	South	4775.59	30	Furniture	4190.28	4272
15	1022	2023-08-17	Charlie	South	8813.55	21	Food	2537.20	2866
16	1053	2023-10-16	Bob	North	2235.83	48	Furniture	121.19	481
17	1002	2023-05-30	David	North	6810.35	17	Furniture	4024.76	4422

NANDINI PANDIT S100

Data Analysis with SAS / SPSS /R Practical

MVLU COLLEGE

R - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Environment History Connections Tutorial

R - Global Environment

Data 9994 obs. of 21 variables

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State
1	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gule	Consumer	United States	Henderson	Kentucky
2	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gule	Consumer	United States	Henderson	Kentucky
3	CA-2016-138688	2016-06-12	2016-06-16	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	California
4	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida
5	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida
6	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
7	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
8	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
9	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
10	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
11	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
12	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California
13	CA-2017-114412	2017-04-15	2017-04-20	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolin
14	CA-2016-161389	2016-12-05	2016-12-10	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle	Washington
15	US-2015-118983	2015-11-22	2015-11-26	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas
16	US-2015-118983	2015-11-22	2015-11-26	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas
17	CA-2014-105893	2014-11-11	2014-11-18	Standard Class	PK-19075	Pete Kilz	Consumer	United States	Madison	Wisconsin
18	CA-2014-167164	2014-05-13	2014-05-15	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan	Utah
19	CA-2014-143336	2014-08-27	2014-09-01	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California
20	CA-2014-143336	2014-08-27	2014-09-01	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California
21	CA-2014-143336	2014-08-27	2014-09-01	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California
22	CA-2016-137330	2016-12-09	2016-12-13	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska
23	CA-2016-137330	2016-12-09	2016-12-13	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska
24	US-2017-156909	2017-07-16	2017-07-18	Second Class	SF-20065	Sandra Flanagan	Consumer	United States	Philadelphia	Pennsylvania
25	CA-2015-106320	2015-09-25	2015-09-30	Standard Class	EB-13870	Emily Burns	Consumer	United States	Orem	Utah
26	CA-2016-121755	2016-01-16	2016-01-20	Second Class	BH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California

Showing 1 to 26 of 9,994 entries, 21 total columns

Console

Files Plots Packages Help Viewer Presentation

28°C Mostly clear 21:53 24-11-2025