

**SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE  
SUBJECT - Data Analysis with SAS / SPSS / R**

Aim :- Generating descriptive statistics using summary() or describe() ®

## Output :-

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help, and a Project dropdown set to '(None)'. The left sidebar has tabs for Untitled1\*, Untitled2\*, R data sets, Source on Save, Run, Source, and Environment. The main area contains a script editor with R code for printing dataset descriptions and summaries, followed by output from the console showing the structure and first 10 rows of the avocado dataset. The right sidebar features an Environment browser with tabs for Environment, History, Connections, and Tutorial. The Environment tab lists various datasets like describe\_vars, df, and iris, each with details such as number of observations and variables. Below the Environment browser is a global environment viewer. At the bottom, there's a file browser, a help viewer, and a presentation viewer.

RStudio interface showing R code, console output, and environment browser.

**R Script:**

```
print(describe(describe_vars))  
print("8. Grouped Summary by Type")  
grouped_type <- df %>%  
  group_by(type) %>%  
  summarise(  
    .by_group = TRUE  
  )
```

**Console:**

```
R - R 4.5.2 - ./ -> 20 rows of the dataset  
> print(head(df, 10))
```

	X	Date	AveragePrice	Total.VOLUME	X4046	X4225	X4770	Total.Bags
1	0	2015-12-27	1.33	64236.62	1036.74	54454.85	48.16	8696.87
2	1	2015-12-20	1.35	54876.98	674.28	44638.81	58.33	9505.56
3	2	2015-12-13	0.93	118220.22	794.70	109149.67	130.50	8145.35
4	3	2015-12-06	1.08	78992.15	1132.00	71976.41	72.58	5811.16
5	4	2015-11-29	1.28	51039.60	941.48	43838.39	75.78	6183.95
6	5	2015-11-22	1.26	55979.78	1184.27	48067.99	43.61	6683.91
7	6	2015-11-15	0.99	83453.76	1368.92	73672.72	93.26	8318.86
8	7	2015-11-08	0.98	109428.33	703.75	101815.36	80.00	6829.22
9	8	2015-11-01	1.02	99811.42	1022.15	87315.57	85.34	11388.36
10	9	2015-10-25	1.07	74338.76	842.40	64757.44	113.00	8625.92

**Environment:**

Global Environment	18249 obs. of 6 variables
df	18249 obs. of 14 variables
duplicates_report	13 obs. of 4 variables
female_non_smoke...	54 obs. of 7 variables
fri_sat_tips	106 obs. of 7 variables
high_bill_subset	32 obs. of 7 variables
iris	150 obs. of 5 variables

**Files:**

- Install
- Update

Package	Description	Source	Version
askpass	Passw...	https://cran.rstudio.com	1.2.1
backports	Reimp...	https://cran.rstudio.com	1.5.0
base64enc	Tools f...	https://cran.rstudio.com	0.1-3
bit	Classe...	https://cran.rstudio.com	4.6.0
bit64	A S3 C...	https://cran.rstudio.com	4.6.0-1
blob	A Sim...	https://cran.rstudio.com	1.2.4
broom	Conve...	https://cran.rstudio.com	0.10.11
bslib	Custo...	https://cran.rstudio.com	0.9.0
cachem	Cache...	https://cran.rstudio.com	1.1.0
callr	Call R...	https://cran.rstudio.com	3.7.6
cellranger	Transf...	https://cran.rstudio.com	1.1.0
cli	Helper...	https://cran.rstudio.com	3.6.5
clipr	Read ...	https://cran.rstudio.com	0.8.0

Name :- Mithil Kadam  
Roll No :- S083

# SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE

## SUBJECT - Data Analysis with SAS / SPSS / R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1\* Untitled2\* R data sets

```

29 print(describe(describe_vars))
30
31 print("8. Grouped Summary by Type")
32 grouped_type <- df %>%
33   group_by(type) %>%
34   summarise(
  
```

(Top Level) R Script

Console Terminal Background Jobs

```

> print(summary(df))
      X          Date        AveragePrice    Total.Volume
Min. : 0.00  Length:18249  Min. :0.440  Min. : 85
1st Qu.:10.00  Class :character  1st Qu.:1.100  1st Qu.: 10839
Median :24.00  Mode  :character  Median :1.370  Median : 107377
Mean  :24.23  Mean   :1.406  Mean   : 850644
3rd Qu.:38.00  3rd Qu.:1.660  3rd Qu.: 432962
Max. :52.00   Max. :3.250  Max. :162505647
X4046      X4225      X4770      Total.Bags
Min. :0.000e+00  Min. : 0  Min. : 0  Min. : 0
1st Qu.:8.541e+02 1st Qu.: 3009 1st Qu.: 0  1st Qu.: 5089
Median :8.645e+03  Median : 29061  Median : 185  Median : 39744
Mean  :12.930e+05  Mean   : 295155  Mean   : 22840  Mean   : 239639
3rd Qu.:1.110e+05 3rd Qu.: 150207 3rd Qu.: 6243  3rd Qu.: 110783
Max. :2.274e+07  Max. :20470573  Max. :2546439  Max. :19373134
Small.Bags Large.Bags XLarge.Bags type
Min. : 0  Min. : 0  Min. : 0  Length:18249
1st Qu.: 2849 1st Qu.: 127.5 1st Qu.: 0.0  Class :character
Median : 26363  Median : 2647.7  Median : 0.0  Mode  :character
Mean  : 182195  Mean   : 54338.1  Mean   : 3106.4
3rd Qu.: 83338 3rd Qu.: 22029.2 3rd Qu.: 132.5
Max. :13384587  Max. :5719096.6  Max. :551693.7
year      region

```

ENG IN 12:22 PM 12/15/2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1\* Untitled2\* R data sets

```

29 print(describe(describe_vars))
30
31 print("8. Grouped Summary by Type")
32 grouped_type <- df %>%
33   group_by(type) %>%
34   summarise(
  
```

(Top Level) R Script

Console Terminal Background Jobs

```

> print("4. Summary of AveragePrice")
[1] "4. Summary of AveragePrice"
> print(summary(df$AveragePrice))
  Min. 1st Qu. Median  Mean 3rd Qu.  Max.
0.440  1.100  1.370  1.406  1.660  3.250
> print("5. Summary of Total Volume")
[1] "5. Summary of Total Volume"
> print(summary(df$Total.Volume))
  Min. 1st Qu. Median  Mean 3rd Qu.  Max.
85  10839  107377  850644  432962  162505647
> print("6. Summary of Small Bags")
[1] "6. Summary of Small Bags"
> print(summary(df$Small.Bags))
  Min. 1st Qu. Median  Mean 3rd Qu.  Max.
0  2849  26363  182195  83338  13384587
> print("7. Detailed Description for Price and Volume Variables")
[1] "7. Detailed Description for Price and Volume Variables"
> describe_vars <- df %>%
+   select(AveragePrice, Total.Volume, X4046, X4225, X4770, Total.Bags)

```

ENG IN 12:22 PM 12/15/2025