

# 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 :-

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Project: (None)
Environment History Connections Tutorial
R - Global Environment
describe_vars 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 Plots Packages Help Viewer Presentation
Install Update
Package Descri... Source Version
User Library
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 1.0.11
bslib Cust... 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 Transl... 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

> print("=== PRACTICAL: DESCRIPTIVE STATISTICS ON AVOCADO DATASET ===")
[1] "=== PRACTICAL: DESCRIPTIVE STATISTICS ON AVOCADO DATASET ==="
> print("1. Structure of Dataset")
[1] "1. Structure of Dataset"
> str(df)
'data.frame': 18249 obs. of 14 variables:
 $ X: int 0 1 2 3 4 5 6 7 8 9 ...
 $ Date: chr "2015-12-27" "2015-12-20" "2015-12-13" "2015-12-06" ...
 $ AveragePrice: num 1.33 1.35 0.93 1.08 1.28 1.26 0.99 0.98 1.02 1.07 ...
 $ Total.Volume: num 64237 54877 118220 78992 51040 ...
 $ X4046: num 1037 674 795 1132 941 ...
 $ X4225: num 54455 44639 109150 71976 43838 ...
 $ X4770: num 48.2 58.3 130.5 72.6 75.8 ...
 $ Total.Bags: num 8697 9506 8145 5811 6184 ...
 $ Small.Bags: num 8604 9408 8042 5677 5986 ...
 $ Large.Bags: num 93.2 97.5 103.1 133.8 197.7 ...
 $ XLarge.Bags: num 0 0 0 0 0 0 0 0 0 ...
 $ type: chr "conventional" "conventional" "conventional" "conventional"
 ...
 $ year: int 2015 2015 2015 2015 2015 2015 2015 2015 2015 2015 ...
 $ region: chr "Albany" "Albany" "Albany" "Albany" ...
> print("2. First 10 Rows of the Dataset")
[1] "2. First 10 Rows of the Dataset"
```

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Project: (None)
Environment History Connections Tutorial
R - Global Environment
describe_vars 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 Plots Packages Help Viewer Presentation
Install Update
Package Descri... Source Version
User Library
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 1.0.11
bslib Cust... 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 Transl... 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

> 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
  Small.Bags Large.Bags XLarge.Bags type year region
1      8603.62      93.25           0 conventional 2015 Albany
2      9408.07      97.49           0 conventional 2015 Albany
3      8042.21     103.14           0 conventional 2015 Albany
4      5677.40     133.76           0 conventional 2015 Albany
5      5986.26     197.69           0 conventional 2015 Albany
6      6556.47     127.44           0 conventional 2015 Albany
7      8196.81     122.05           0 conventional 2015 Albany
8      6266.85     562.37           0 conventional 2015 Albany
9     11104.53     283.83           0 conventional 2015 Albany
10     8061.47     564.45           0 conventional 2015 Albany
```

Name :- Mithil Kadam

Roll No :- S083

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

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

RStudio interface showing data analysis results. The console displays summary statistics for variables X, Date, AveragePrice, Total.Volume, X4046, X4225, X4770, and Total.Bags.

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
X	0.00	10.00	24.00	24.23	38.00	52.00
Date	Length:18249	Class:character	Mode:character			
AveragePrice	0.440	1.100	1.370	1.406	1.660	3.250
Total.Volume	85	10839	107377	850644	432962	62505647
X4046	0.000e+00	8.541e+02	8.645e+03	2.930e+05	1.110e+05	2.274e+07
X4225	0	3009	29061	295155	150207	20470573
X4770	0	185	22840	6243	2546439	19373134
Total.Bags	0	5089	39744	239639	110783	19373134

RStudio interface showing data analysis results. The console displays summary statistics for variables X, Date, AveragePrice, Total.Volume, X4046, X4225, X4770, and Total.Bags.

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
X	0.00	10.00	24.00	24.23	38.00	52.00
Date	Length:18249	Class:character	Mode:character			
AveragePrice	0.440	1.100	1.370	1.406	1.660	3.250
Total.Volume	85	10839	107377	850644	432962	62505647
X4046	0.000e+00	8.541e+02	8.645e+03	2.930e+05	1.110e+05	2.274e+07
X4225	0	3009	29061	295155	150207	20470573
X4770	0	185	22840	6243	2546439	19373134
Total.Bags	0	5089	39744	239639	110783	19373134

Name :- Mithil Kadam  
Roll No :- S083