

SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT :- DATA ANALYSIS WITH SAS / SPSS / R

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R4.1.2 - ~/
The downloaded source packages are in
'c:\Users\info\AppData\Local\Temp\Rtmpcb4sqh\downloaded_packages'
> my_data <- read.csv("c:\\Users\\info\\Downloads\\house_prices_texas_counties.csv")
> head(my_data)
  1..Location Bedroom Total.Sqft Bathroom Price X X.1 X.2
1 Anderson 1 650 1 110000 NA NA NA
2 Andrews 1 700 1 153200 NA NA NA
3 Angelina 1 800 1 110500 NA NA NA
4 Aransas 1 650 1 183200 NA NA NA
5 Archer 1 700 1 156100 NA NA NA
6 Armstrong 1 800 1 106000 NA NA NA
> tail(my_data)
  967 winkler 4 2560 4 390000 NA NA NA
968 wood 4 2570 4 450000 NA NA NA
969 yoakum 4 2580 4 420000 NA NA NA
970 young 4 2590 4 390000 NA NA NA
971 zapata 4 2600 4 510000 NA NA NA
972 Zavala 4 2350 4 570000 NA NA NA
> dim(my_data)
[1] 972 8
> cat("Dimensions (rows, columns): ", dim(my_data), "\n")
Dimensions (rows, columns): 972 8
> str(my_data)
'data.frame': 972 obs. of 8 variables:
 $ 1..Location: chr "Anderson" "Andrews" "Angelina" "Aransas" ...
 $ Bedroom: int 1 1 1 1 1 1 1 1 1 1 ...
 $ Total.Sqft: int 650 700 800 650 700 800 650 700 800 800 ...
 $ Bathroom: int 1 1 1 1 1 1 1 1 1 1 ...
 $ Price: num 110000 153200 110500 183200 156100 ...
 $ X: logi NA NA NA NA NA NA ...
 $ X.1: logi NA NA NA NA NA NA ...
 $ X.2: logi NA NA NA NA NA NA ...
> summary(my_data)
 1..Location Bedroom Total.Sqft Bathroom Price X X.1 X.2
Length:972 Min. :1.00 Min. : 650 Min. :1.00 Min. : 62560 Mode:logical
Class :character 1st Qu.:1.75 1st Qu.: 950 1st Qu.:1.75 1st Qu.: 180000 NA's:972
Mode :character Median :2.50 Median :1560 Median :2.50 Median : 270000
Mean :2.50 Mean :1576 Mean :2.50 Mean : 302922
3rd Qu.:3.25 3rd Qu.:2010 3rd Qu.:3.25 3rd Qu.: 393565
Max. :4.00 Max. :2640 Max. :4.00 Max. :1350000
> names(my_data)
[1] "1..Location" "Bedroom" "Total.Sqft" "Bathroom" "Price" "X" "X.1" "X.2"
```

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R4.1.2 - ~/
The downloaded source packages are in
'c:\Users\info\AppData\Local\Temp\Rtmpcb4sqh\downloaded_packages'
> my_data <- read.csv("c:\\Users\\info\\Downloads\\house_prices_texas_counties.csv")
> head(my_data)
  1..Location Bedroom Total.Sqft Bathroom Price X X.1 X.2
1 Anderson 1 650 1 110000 NA NA NA
2 Andrews 1 700 1 153200 NA NA NA
3 Angelina 1 800 1 110500 NA NA NA
4 Aransas 1 650 1 183200 NA NA NA
5 Archer 1 700 1 156100 NA NA NA
6 Armstrong 1 800 1 106000 NA NA NA
> tail(my_data)
  967 winkler 4 2560 4 390000 NA NA NA
968 wood 4 2570 4 450000 NA NA NA
969 yoakum 4 2580 4 420000 NA NA NA
970 young 4 2590 4 390000 NA NA NA
971 zapata 4 2600 4 510000 NA NA NA
972 Zavala 4 2350 4 570000 NA NA NA
> dim(my_data)
[1] 972 8
> cat("Dimensions (rows, columns): ", dim(my_data), "\n")
Dimensions (rows, columns): 972 8
> str(my_data)
'data.frame': 972 obs. of 8 variables:
 $ 1..Location: chr "Anderson" "Andrews" "Angelina" "Aransas" ...
 $ Bedroom: int 1 1 1 1 1 1 1 1 1 1 ...
 $ Total.Sqft: int 650 700 800 650 700 800 650 700 800 800 ...
 $ Bathroom: int 1 1 1 1 1 1 1 1 1 1 ...
 $ Price: num 110000 153200 110500 183200 156100 ...
 $ X: logi NA NA NA NA NA NA ...
 $ X.1: logi NA NA NA NA NA NA ...
 $ X.2: logi NA NA NA NA NA NA ...
> summary(my_data)
 1..Location Bedroom Total.Sqft Bathroom Price X X.1 X.2
Length:972 Min. :1.00 Min. : 650 Min. :1.00 Min. : 62560 Mode:logical
Class :character 1st Qu.:1.75 1st Qu.: 950 1st Qu.:1.75 1st Qu.: 180000 NA's:972
Mode :character Median :2.50 Median :1560 Median :2.50 Median : 270000
Mean :2.50 Mean :1576 Mean :2.50 Mean : 302922
3rd Qu.:3.25 3rd Qu.:2010 3rd Qu.:3.25 3rd Qu.: 393565
Max. :4.00 Max. :2640 Max. :4.00 Max. :1350000
> names(my_data)
[1] "1..Location" "Bedroom" "Total.Sqft" "Bathroom" "Price" "X" "X.1" "X.2"
> cat("Column Names: ", names(my_data), "\n")
Column Names: 1..Location Bedroom Total.Sqft Bathroom Price X X.1 X.2
> summary(my_data)
 1..Location Bedroom Total.Sqft Bathroom Price X X.1 X.2
Length:972 Min. :1.00 Min. : 650 Min. :1.00 Min. : 62560 Mode:logical
Class :character 1st Qu.:1.75 1st Qu.: 950 1st Qu.:1.75 1st Qu.: 180000 NA's:972
Mode :character Median :2.50 Median :1560 Median :2.50 Median : 270000
Mean :2.50 Mean :1576 Mean :2.50 Mean : 302922
3rd Qu.:3.25 3rd Qu.:2010 3rd Qu.:3.25 3rd Qu.: 393565
Max. :4.00 Max. :2640 Max. :4.00 Max. :1350000
> |
```