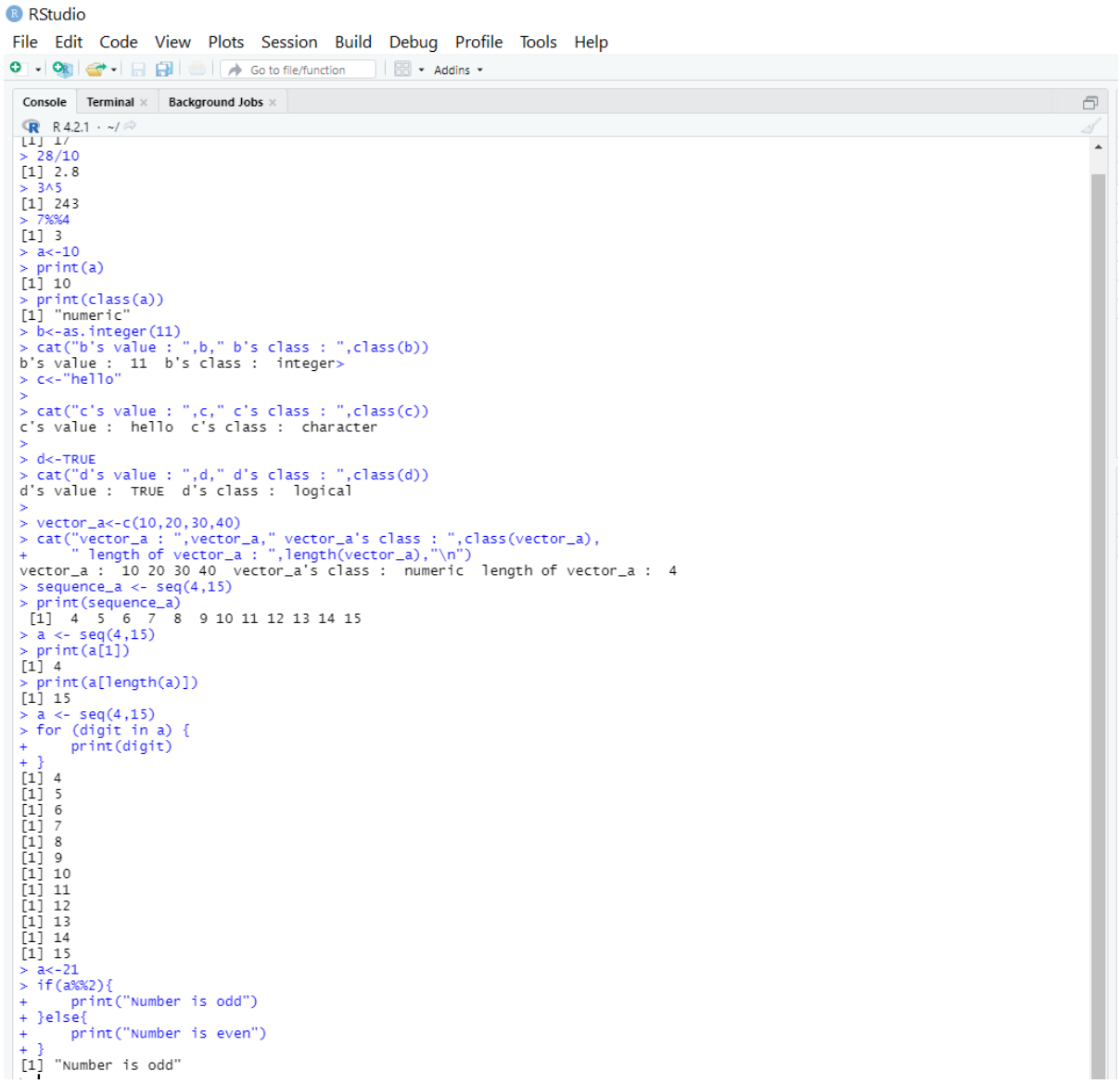
Data Analytics Assignment - 01

Name : Adarsh Subhas Nayak
SRN : PES1UG20CS620
Section : K
Date : 17-08-2022



```
package 'gtable' successfully unpacked and MD5 sums checked
package 'isoband' successfully unpacked and MD5 sums checked
package 'rlang' successfully unpacked and MD5 sums checked
package 'scales' successfully unpacked and MD5 sums checked
package 'tibble' successfully unpacked and MD5 sums checked
package 'withr' successfully unpacked and MD5 sums checked
package 'ggplot2' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Hp\AppData\Local\Temp\RtmpKuM86c\downloaded_packages
> library(ggplot2)
> search()
 [1] ".GlobalEnv"        "package:ggplot2"
 [3] "tools:rstudio"     "package:stats"
 [5] "package:graphics"  "package:grDevices"
 [7] "package:utils"     "package:datasets"
 [9] "package:methods"   "Autoloads"
[11] "org:r-lib"         "package:base"
>
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins

Console | Terminal | Background Jobs

R 4.2.1 · ~/

```
The downloaded binary packages are in
        C:\Users\Hp\AppData\Local\Temp\RtmpKuM86c\downloaded_packages
> library(ggplot2)
> search()
 [1] ".GlobalEnv"        "package:ggplot2"
 [3] "tools:rstudio"     "package:stats"
 [5] "package:graphics"  "package:grDevices"
 [7] "package:utils"     "package:datasets"
 [9] "package:methods"   "Autoloads"
[11] "org:r-lib"         "package:base"
>
> df <- txhousing
> head(df)
# A tibble: 6 × 9
  city     year month sales volume median listi…¹ inven…²
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Abilene  2000     1    72 5.38e6  71400     701     6.3
2 Abilene  2000     2    98 6.50e6  58700     746     6.6
3 Abilene  2000     3   130 9.28e6  58100     784     6.8
4 Abilene  2000     4    98 9.73e6  68600     785     6.9
5 Abilene  2000     5   141 1.06e7  67300     794     6.8
6 Abilene  2000     6   156 1.39e7  66900     780     6.6
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names ¹listings, ²inventory
# ℹ Use `colnames()` to see all variable names
> tail(df)
# A tibble: 6 × 9
  city     year month sales volume median listi…¹ inven…²
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Wichit…  2015     2   100 1.16e7  94000     795     6.8
2 Wichit…  2015     3   152 1.67e7  89200     818     6.8
3 Wichit…  2015     4   129 1.55e7 105300     760     6.4
4 Wichit…  2015     5   174 1.92e7 100000     776     6.4
5 Wichit…  2015     6   143 1.88e7 118800     770     6.2
6 Wichit…  2015     7   172 2.39e7 116700     811     6.5
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names ¹listings, ²inventory
# ℹ Use `colnames()` to see all variable names
> colnames(df)
[1] "city"      "year"      "month"     "sales"
[5] "volume"    "median"    "listings"  "inventory"
[9] "date"
>
> dim(df)
[1] 8602    9
>
> top5 <- df[1:5,]
> top5
# A tibble: 5 × 9
  city     year month sales volume median listi…¹ inven…²
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Abilene  2000     1    72 5.38e6  71400     701     6.3
2 Abilene  2000     2    98 6.50e6  58700     746     6.6
3 Abilene  2000     3   130 9.28e6  58100     784     6.8
4 Abilene  2000     4    98 9.73e6  68600     785     6.9
5 Abilene  2000     5   141 1.06e7  67300     794     6.8
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names ¹listings, ²inventory
# ℹ Use `colnames()` to see all variable names
>
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

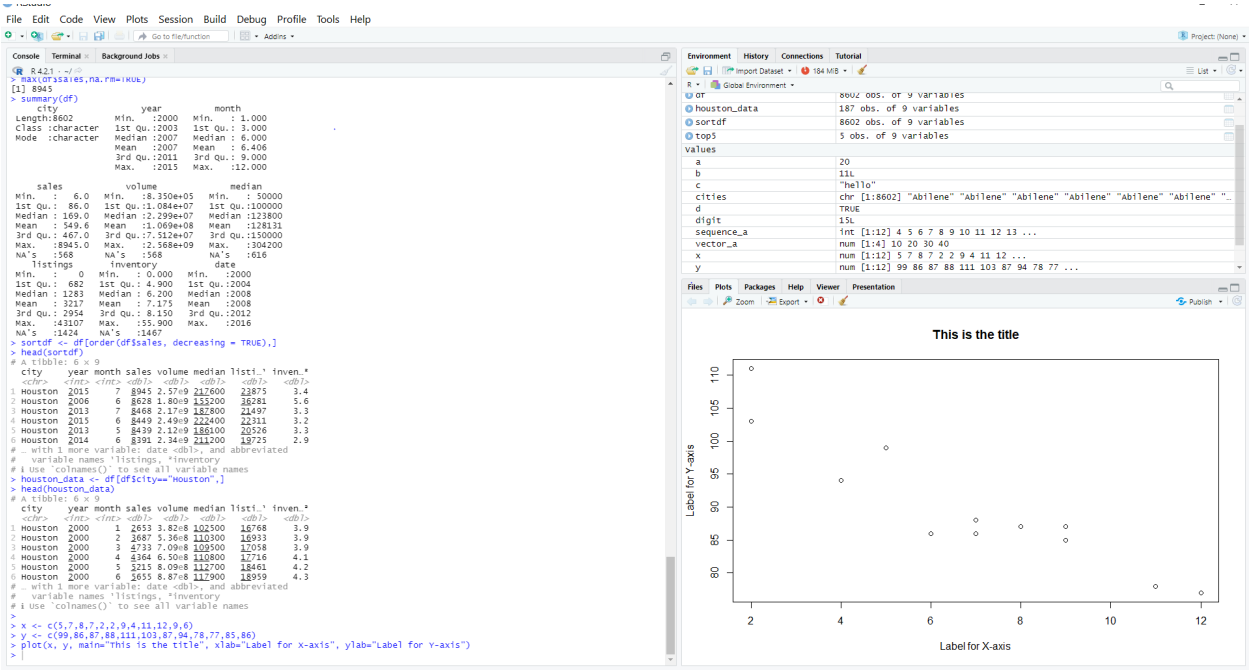
Go to file/function | Addins

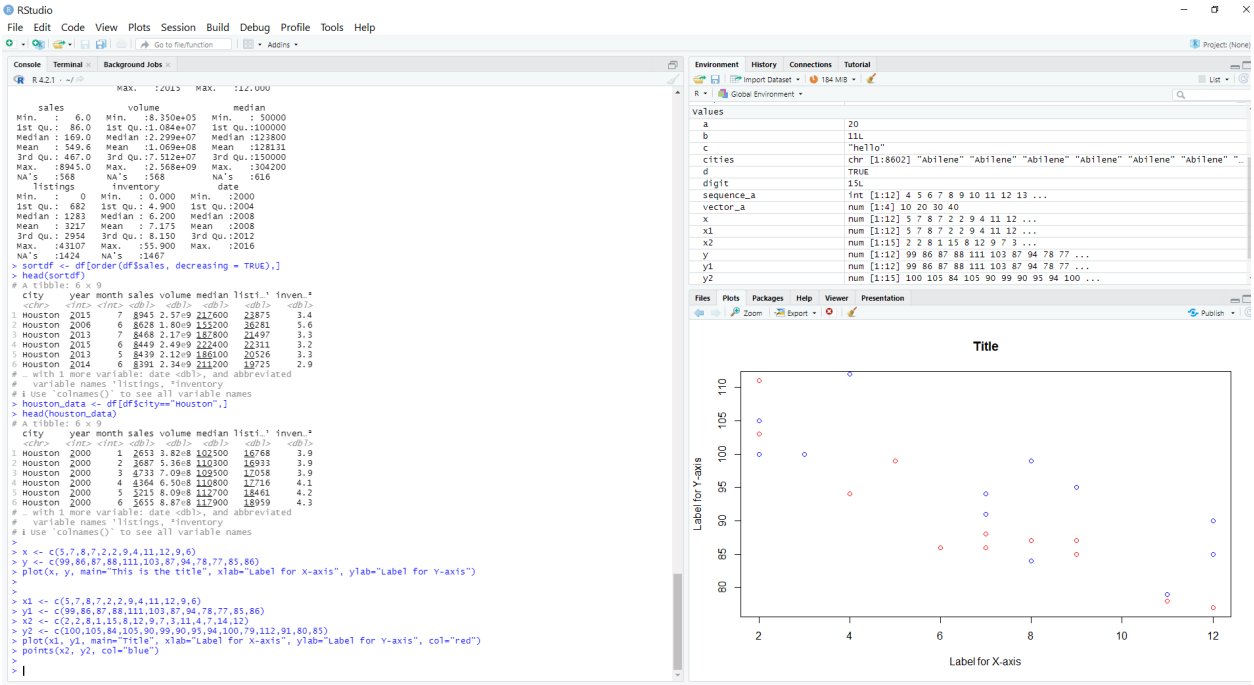
Console | Terminal | Background Jobs

R 4.2.1 · ~/

```
> median(df$sales,na.rm=TRUE)
[1] 169
> min(df$sales,na.rm=TRUE)
[1] 6
> max(df$sales,na.rm=TRUE)
[1] 8945
> summary(df)
     city                year          month
 Length:8602        Min.   :2000   Min.   : 1.000
 Class :character   1st Qu.:2003   1st Qu.: 3.000
 Mode  :character   Median :2007   Median : 6.000
                    Mean   :2007   Mean   : 6.406
                    3rd Qu.:2011   3rd Qu.: 9.000
                    Max.   :2015   Max.   :12.000

     sales            volume              median
 Min.   :   6.0   Min.   :8.350e+05   Min.   : 50000
 1st Qu.:  86.0   1st Qu.:1.084e+07   1st Qu.:100000
 Median : 169.0   Median :2.299e+07   Median :123800
 Mean   : 549.6   Mean   :1.069e+08   Mean   :128131
 3rd Qu.: 467.0   3rd Qu.:7.512e+07   3rd Qu.:150000
 Max.   :8945.0   Max.   :2.568e+09   Max.   :304200
 NA's   :568      NA's   :568         NA's   :616
    listings       inventory            date
 Min.   :    0   Min.   : 0.000   Min.   :2000
 1st Qu.:  682   1st Qu.: 4.900   1st Qu.:2004
 Median : 1283   Median : 6.200   Median :2008
 Mean   : 3217   Mean   : 7.175   Mean   :2008
 3rd Qu.: 2954   3rd Qu.: 8.150   3rd Qu.:2012
 Max.   :43107   Max.   :55.900   Max.   :2016
 NA's   :1424    NA's   :1467
> sortdf <- df[order(df$sales, decreasing = TRUE),]
> head(sortdf)
# A tibble: 6 × 9
  city     year month sales volume median listi…¹ inven…²
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Houston  2015     7  8945 2.57e9 217600   23875     3.4
2 Houston  2006     6  8628 1.80e9 155200   36281     5.6
3 Houston  2013     7  8468 2.17e9 187800   21497     3.3
4 Houston  2015     6  8449 2.49e9 222400   22311     3.2
5 Houston  2013     5  8439 2.12e9 186100   20526     3.3
6 Houston  2014     6  8391 2.34e9 211200   19725     2.9
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names ¹listings, ²inventory
# ℹ Use `colnames()` to see all variable names
> houston_data <- df[df$city=="Houston",]
> head(houston_data)
# A tibble: 6 × 9
  city     year month sales volume median listi…¹ inven…²
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Houston  2000     1  2653 3.82e8 102500   16768     3.9
2 Houston  2000     2  3687 5.36e8 110300   16933     3.9
3 Houston  2000     3  4733 7.09e8 109500   17058     3.9
4 Houston  2000     4  4364 6.50e8 110800   17716     4.1
5 Houston  2000     5  5215 8.09e8 112700   18461     4.2
6 Houston  2000     6  5655 8.87e8 117900   18959     4.3
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names ¹listings, ²inventory
# ℹ Use `colnames()` to see all variable names
>
```


File Edit Code View Plots Session Build Debug Profile Tools Help
> max(df$sales,na.rm=TRUE)
[1] 8945
> summary(df)
city year month
Length:8602 Min. :2000 Min. : 1.000
Class :character 1st Qu.:2003 1st Qu.: 3.000
Mode :character Median :2007 Median : 6.000
Mean :2007 Mean : 6.406
3rd Qu.:2011 3rd Qu.: 9.000
Max. :2015 Max. :12.000
sales volume median
Min. : 6.0 Min. :8.350e+05 Min. : 50000
1st Qu.: 86.0 1st Qu.:1.084e+07 1st Qu.:100000
Median : 169.0 Median :2.299e+07 Median :123800
Mean : 549.6 Mean :1.069e+08 Mean :128131
3rd Qu.: 467.0 3rd Qu.:7.512e+07 3rd Qu.:150000
Max. :8945.0 Max. :2.568e+09 Max. :304200
NA's :568 NA's :568 NA's :616
listings inventory date
Min. : 0 Min. : 0.000 Min. :2000
1st Qu.: 682 1st Qu.: 4.900 1st Qu.:2004
Median : 1283 Median : 6.200 Median :2008
Mean : 3217 Mean : 7.175 Mean :2008
3rd Qu.: 2954 3rd Qu.: 8.150 3rd Qu.:2012
Max. :43107 Max. :55.900 Max. :2016
NA's :1424 NA's :1467
> sortdf <- df[order(df$sales, decreasing = TRUE),]
> head(sortdf)
A tibble: 6 × 9
city year month sales volume median listi…¹ inven.²
<chr> <int> <int> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Houston 2015 7 8945 2.57e9 217600 23875 3.4
2 Houston 2006 6 8628 1.80e9 155200 16281 5.6
3 Houston 2013 7 8468 2.17e9 187800 21497 3.3
4 Houston 2015 6 8449 2.49e9 222400 22311 3.2
5 Houston 2013 5 8439 2.12e9 186100 20526 3.3
6 Houston 2014 6 8391 2.34e9 211200 19725 2.9
… with 1 more variable: date <dbl>, and abbreviated
variable names ¹listings, ²inventory
ℹ Use `colnames()` to see all variable names
> houston_data <- df[df$city=="Houston",]
> head(houston_data)
A tibble: 6 × 9
city year month sales volume median listi…¹ inven.²
<chr> <int> <int> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Houston 2000 1 2653 3.82e8 102500 16768 3.9
2 Houston 2000 2 3687 5.36e8 110300 16933 3.9
3 Houston 2000 3 4733 7.09e8 109500 17058 3.9
4 Houston 2000 4 4364 6.50e8 110800 17716 4.1
5 Houston 2000 5 5215 8.09e8 112700 18461 4.2
6 Houston 2000 6 5655 8.87e8 117900 18959 4.3
… with 1 more variable: date <dbl>, and abbreviated
variable names ¹listings, ²inventory
ℹ Use `colnames()` to see all variable names
>
> x <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> plot(x, y, main="This is the title", xlab="Label for X-axis", ylab="Label for Y-axis")
Environment History Connections Tutorial
Import Dataset 184 MiB
Global Environment
houston_data 187 obs. of 9 variables
sortdf 8602 obs. of 9 variables
top5 5 obs. of 9 variables
values
a 20
b 11L
c "hello"
cities chr [1:8602] "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "...
d TRUE
digit 15L
sequence_a int [1:12] 4 5 6 7 8 9 10 11 12 13 ...
vector_a num [1:4] 10 20 30 40
x num [1:12] 5 7 8 7 2 2 9 4 11 12 ...
y num [1:12] 99 86 87 88 111 103 87 94 78 77 ...
Files Plots Packages Help Viewer Presentation
Zoom Export Publish
This is the title
Label for Y-axis
Label for X-axis
80 85 90 95 100 105 110
2 4 6 8 10 12

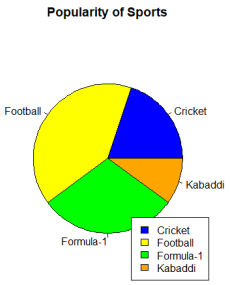

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Max. :2015 Max. :12.000
sales volume median
Min. : 6.0 Min. :8.350e+05 Min. : 50000
1st Qu.: 86.0 1st Qu.:1.084e+07 1st Qu.:100000
Median : 169.0 Median :2.299e+07 Median :123800
Mean : 549.6 Mean :1.069e+08 Mean :128131
3rd Qu.: 467.0 3rd Qu.:7.512e+07 3rd Qu.:150000
Max. :8945.0 Max. :2.568e+09 Max. :304200
NA's :568 NA's :568 NA's :616
listings inventory date
Min. : 0 Min. : 0.000 Min. :2000
1st Qu.: 682 1st Qu.: 4.900 1st Qu.:2004
Median : 1283 Median : 6.200 Median :2008
Mean : 3217 Mean : 7.175 Mean :2008
3rd Qu.: 2954 3rd Qu.: 8.150 3rd Qu.:2012
Max. :43107 Max. :55.900 Max. :2016
NA's :1424 NA's :1467
> sortdf <- df[order(df$sales, decreasing = TRUE),]
> head(sortdf)
A tibble: 6 × 9
city year month sales volume median listi…¹ inven.²
<chr> <int> <int> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Houston 2015 7 8945 2.57e9 217600 23875 3.4
2 Houston 2006 6 8628 1.80e9 155200 16281 5.6
3 Houston 2013 7 8468 2.17e9 187800 21497 3.3
4 Houston 2015 6 8449 2.49e9 222400 22311 3.2
5 Houston 2013 5 8439 2.12e9 186100 20526 3.3
6 Houston 2014 6 8391 2.34e9 211200 19725 2.9
… with 1 more variable: date <dbl>, and abbreviated
variable names ¹listings, ²inventory
ℹ Use `colnames()` to see all variable names
> houston_data <- df[df$city=="Houston",]
> head(houston_data)
A tibble: 6 × 9
city year month sales volume median listi…¹ inven.²
<chr> <int> <int> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Houston 2000 1 2653 3.82e8 102500 16768 3.9
2 Houston 2000 2 3687 5.36e8 110300 16933 3.9
3 Houston 2000 3 4733 7.09e8 109500 17058 3.9
4 Houston 2000 4 4364 6.50e8 110800 17716 4.1
5 Houston 2000 5 5215 8.09e8 112700 18461 4.2
6 Houston 2000 6 5655 8.87e8 117900 18959 4.3
… with 1 more variable: date <dbl>, and abbreviated
variable names ¹listings, ²inventory
ℹ Use `colnames()` to see all variable names
>
> x <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> plot(x, y, main="This is the title", xlab="Label for X-axis", ylab="Label for Y-axis")
>
>
> x1 <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y1 <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> x2 <- c(2,2,8,1,15,8,12,9,7,3,11,4,7,14,12)
> y2 <- c(100,105,84,105,90,99,90,95,94,100,79,112,91,80,85)
> plot(x1, y1, main="Title", xlab="Label for X-axis", ylab="Label for Y-axis", col="red")
> points(x2, y2, col="blue")
>
Environment History Connections Tutorial
Import Dataset 184 MiB
Global Environment
values
a 20
b 11L
c "hello"
cities chr [1:8602] "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "...
d TRUE
digit 15L
sequence_a int [1:12] 4 5 6 7 8 9 10 11 12 13 ...
vector_a num [1:4] 10 20 30 40
x num [1:12] 5 7 8 7 2 2 9 4 11 12 ...
x1 num [1:12] 5 7 8 7 2 2 9 4 11 12 ...
x2 num [1:15] 2 2 8 1 15 8 12 9 7 3 ...
y num [1:12] 99 86 87 88 111 103 87 94 78 77 ...
y1 num [1:12] 99 86 87 88 111 103 87 94 78 77 ...
y2 num [1:15] 100 105 84 105 90 99 90 95 94 100 ...
Files Plots Packages Help Viewer Presentation
Zoom Export Publish
Title
Label for Y-axis
Label for X-axis
80 85 90 95 100 105 110
2 4 6 8 10 12

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

```
 3rd Qu.: 467.0   3rd Qu.:7.512e+07   3rd Qu.:150000
 Max.   :8945.0   Max.   :2.568e+09   Max.   :304200
 NA's   :568      NA's   :568        NA's   :616
   listings         inventory            date
 Min.   :    0   Min.   : 0.000   Min.   :2000
 1st Qu.:  682   1st Qu.: 4.900   1st Qu.:2004
 Median : 1283   Median : 6.200   Median :2008
 Mean   : 3217   Mean   : 7.175   Mean   :2008
 3rd Qu.: 2954   3rd Qu.: 8.150   3rd Qu.:2012
 Max.   :43107   Max.   :55.900   Max.   :2016
 NA's   :1424    NA's   :1467
> sortdf <- df[order(df$sales, decreasing = TRUE),]
> head(sortdf)
# A tibble: 6 x 9
  city     year month sales volume median listi…' inven…*
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Houston  2015     7  8945 2.57e9 217600   23875     3.4
2 Houston  2006     6  8628 1.80e9 155200   36281     5.6
3 Houston  2013     7  8468 2.17e9 187800   21497     3.3
4 Houston  2015     6  8449 2.49e9 222400   22311     3.2
5 Houston  2013     5  8439 2.12e9 186100   20526     3.3
6 Houston  2014     6  8391 2.34e9 211200   19725     2.9
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names 'listings, *inventory
# ℹ Use `colnames()` to see all variable names
> houston_data <- df[df$city=="Houston",]
> head(houston_data)
# A tibble: 6 x 9
  city     year month sales volume median listi…' inven…*
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Houston  2000     1  2653 3.82e8 102500   16768     3.9
2 Houston  2000     2  3687 5.36e8 110300   16933     3.9
3 Houston  2000     3  4733 7.09e8 109500   17058     3.9
4 Houston  2000     4  4364 6.50e8 110800   17716     4.1
5 Houston  2000     5  5215 8.09e8 112700   18461     4.2
6 Houston  2000     6  5655 8.87e8 117900   18959     4.3
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names 'listings, *inventory
# ℹ Use `colnames()` to see all variable names
>
> x <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> plot(x, y, main="This is the title", xlab="Label for X-axis", ylab="Label for Y-axis")
>
>
> x1 <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y1 <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> x2 <- c(2,2,8,1,15,8,12,9,7,3,11,4,7,14,12)
> y2 <- c(100,105,84,105,90,99,90,95,94,100,79,112,91,80,85)
> plot(x1, y1, main="Title", xlab="Label for X-axis", ylab="Label for Y-axis", col="red")
> points(x2, y2, col="blue")
>
>
> #Pie chart
> x <- c(20,40,30,10)
> mylabel <- c("Cricket", "Football", "Formula-1", "Kabaddi")
> colors <- c("blue", "yellow", "green", "orange")
> pie(x, label = mylabel, main = "Popularity of Sports", col = colors)
> legend("bottomright", mylabel, fill = colors)
>
```

b	11L
c	"hello"
cities	chr [1:8602] "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "…
colors	chr [1:4] "blue" "yellow" "green" "orange"
d	TRUE
digit	15L
mylabel	chr [1:4] "Cricket" "Football" "Formula-1" "Kabaddi"
sequence_a	int [1:12] 4 5 6 7 8 9 10 11 12 13 ...
vector_a	num [1:4] 10 20 30 40
x	num [1:4] 20 40 30 10
x1	num [1:12] 5 7 8 7 2 2 9 4 11 12 ...
x2	num [1:15] 2 2 8 1 15 8 12 9 7 3 ...
y	num [1:12] 99 86 87 88 111 103 87 94 78 77 ...
y1	num [1:12] 99 86 87 88 111 103 87 94 78 77 ...
y2	num [1:15] 100 105 84 105 90 99 90 95 94 100 ...



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

```
 1st Qu.:  682   1st Qu.: 4.900   1st Qu.:2004
 Median : 1283   Median : 6.200   Median :2008
 Mean   : 3217   Mean   : 7.175   Mean   :2008
 3rd Qu.: 2954   3rd Qu.: 8.150   3rd Qu.:2012
 Max.   :43107   Max.   :55.900   Max.   :2016
 NA's   :1424    NA's   :1467
> sortdf <- df[order(df$sales, decreasing = TRUE),]
> head(sortdf)
# A tibble: 6 x 9
  city     year month sales volume median listi…' inven…*
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Houston  2015     7  8945 2.57e9 217600   23875     3.4
2 Houston  2006     6  8628 1.80e9 155200   36281     5.6
3 Houston  2013     7  8468 2.17e9 187800   21497     3.3
4 Houston  2015     6  8449 2.49e9 222400   22311     3.2
5 Houston  2013     5  8439 2.12e9 186100   20526     3.3
6 Houston  2014     6  8391 2.34e9 211200   19725     2.9
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names 'listings, *inventory
# ℹ Use `colnames()` to see all variable names
> houston_data <- df[df$city=="Houston",]
> head(houston_data)
# A tibble: 6 x 9
  city     year month sales volume median listi…' inven…*
  <chr>   <int> <int> <dbl>  <dbl>  <dbl>   <dbl>   <dbl>
1 Houston  2000     1  2653 3.82e8 102500   16768     3.9
2 Houston  2000     2  3687 5.36e8 110300   16933     3.9
3 Houston  2000     3  4733 7.09e8 109500   17058     3.9
4 Houston  2000     4  4364 6.50e8 110800   17716     4.1
5 Houston  2000     5  5215 8.09e8 112700   18461     4.2
6 Houston  2000     6  5655 8.87e8 117900   18959     4.3
# … with 1 more variable: date <dbl>, and abbreviated
#   variable names 'listings, *inventory
# ℹ Use `colnames()` to see all variable names
>
> x <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> plot(x, y, main="This is the title", xlab="Label for X-axis", ylab="Label for Y-axis")
>
>
> x1 <- c(5,7,8,7,2,2,9,4,11,12,9,6)
> y1 <- c(99,86,87,88,111,103,87,94,78,77,85,86)
> x2 <- c(2,2,8,1,15,8,12,9,7,3,11,4,7,14,12)
> y2 <- c(100,105,84,105,90,99,90,95,94,100,79,112,91,80,85)
> plot(x1, y1, main="Title", xlab="Label for X-axis", ylab="Label for Y-axis", col="red")
> points(x2, y2, col="blue")
>
>
> #Pie chart
> x <- c(20,40,30,10)
> mylabel <- c("Cricket", "Football", "Formula-1", "Kabaddi")
> colors <- c("blue", "yellow", "green", "orange")
> pie(x, label = mylabel, main = "Popularity of Sports", col = colors)
> legend("bottomright", mylabel, fill = colors)
>
> #Bar Plot
> x <- c("RCB", "MI", "CSK", "KKR")
> y <- c(5, 4, 3, 2)
> barplot(y, names.arg = x, col = c("red","blue","yellow","purple"),main="Successful IPL teams in India")
>
```

b	11L
c	"hello"
cities	chr [1:8602] "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "Abilene" "…
colors	chr [1:4] "blue" "yellow" "green" "orange"
d	TRUE
digit	15L
mylabel	chr [1:4] "Cricket" "Football" "Formula-1" "Kabaddi"
sequence_a	int [1:12] 4 5 6 7 8 9 10 11 12 13 ...
vector_a	num [1:4] 10 20 30 40
x	chr [1:4] "RCB" "MI" "CSK" "KKR"
x1	num [1:12] 5 7 8 7 2 2 9 4 11 12 ...
x2	num [1:15] 2 2 8 1 15 8 12 9 7 3 ...
y	num [1:4] 5 4 3 2
y1	num [1:12] 99 86 87 88 111 103 87 94 78 77 ...
y2	num [1:15] 100 105 84 105 90 99 90 95 94 100 ...