Preparing a Markdown about a dataset

Rdouglas

23/1/2022

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Dataset : iris

The Iris flower data set or Fisher's Iris data set is a multivariate data set introduced by the British statistician and biologist Ronald Fisher in his 1936 paper The use of multiple measurements in taxonomic problems as an example of linear discriminant analysis.[1] It is sometimes called Anderson's Iris data set because Edgar Anderson collected the data to quantify the morphologic variation of Iris flowers of three related species.[2] Two of the three species were collected in the Gaspé Peninsula "all from the same pasture, and picked on the same day and measured at the same time by the same person with the same apparatus"

Attribute Information:

1. sepal length in cm
2. sepal width in cm
3. petal length in cm
4. petal width in cm
5. class: – Iris Setosa – Iris Versicolour – Iris Virginica ## import libraries and packages

```
summary(iris)
```

```
##   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width   
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100  
##  1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300  
##  Median :5.800   Median :3.000   Median :4.350   Median :1.300  
##  Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199  
##  3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800  
##  Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500  
##        Species  
##  setosa    :50  
##  versicolor:50  
##  virginica :50  
##                 
##                 
## 
```

```
colnames(iris)
```

```
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

```
str(iris)
```

```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

Checking if there are any NA's values

```
is.na(iris)
```

```
##        Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##   [1,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [2,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [3,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [4,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [5,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [6,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [7,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [8,]        FALSE       FALSE        FALSE       FALSE   FALSE
##   [9,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [10,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [11,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [12,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [13,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [14,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [15,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [16,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [17,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [18,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [19,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [20,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [21,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [22,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [23,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [24,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [25,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [26,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [27,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [28,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [29,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [30,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [31,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [32,]        FALSE       FALSE        FALSE       FALSE   FALSE
```

```
## [33,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [34,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [35,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [36,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [37,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [38,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [39,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [40,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [41,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [42,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [43,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [44,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [45,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [46,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [47,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [48,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [49,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [50,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [51,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [52,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [53,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [54,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [55,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [56,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [57,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [58,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [59,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [60,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [61,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [62,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [63,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [64,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [65,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [66,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [67,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [68,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [69,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [70,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [71,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [72,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [73,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [74,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [75,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [76,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [77,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [78,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [79,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [80,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [81,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [82,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [83,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [84,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [85,]      FALSE      FALSE      FALSE      FALSE  FALSE
## [86,]      FALSE      FALSE      FALSE      FALSE  FALSE
```

```
##  [87,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [88,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [89,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [90,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [91,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [92,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [93,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [94,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [95,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [96,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [97,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [98,]        FALSE       FALSE        FALSE       FALSE   FALSE
##  [99,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [100,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [101,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [102,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [103,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [104,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [105,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [106,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [107,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [108,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [109,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [110,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [111,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [112,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [113,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [114,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [115,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [116,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [117,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [118,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [119,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [120,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [121,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [122,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [123,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [124,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [125,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [126,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [127,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [128,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [129,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [130,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [131,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [132,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [133,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [134,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [135,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [136,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [137,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [138,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [139,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [140,]        FALSE       FALSE        FALSE       FALSE   FALSE
```

```
## [141,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [142,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [143,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [144,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [145,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [146,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [147,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [148,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [149,]        FALSE       FALSE        FALSE       FALSE   FALSE
## [150,]        FALSE       FALSE        FALSE       FALSE   FALSE
```

what about Species?

```
unique(iris$Species)
```

```
## [1] setosa     versicolor virginica
## Levels: setosa versicolor virginica
```

Measures of central tendency

Means values for Sepal.Length and Sepal.Width

Check two columns Sepal.Length and Sepal.Width

```
head(iris[c(1,2)])
```

```
##   Sepal.Length Sepal.Width
## 1          5.1         3.5
## 2          4.9         3.0
## 3          4.7         3.2
## 4          4.6         3.1
## 5          5.0         3.6
## 6          5.4         3.9
```

mean values = 5.84 Sepal.Length and 3.05 for Sepal.Width

```
mean(iris$Sepal.Length, na.rm = TRUE)
```

```
## [1] 5.843333
```

```
mean(iris$Sepal.Width, na.rm = TRUE)
```

```
## [1] 3.057333
```

mean values for each characteristic

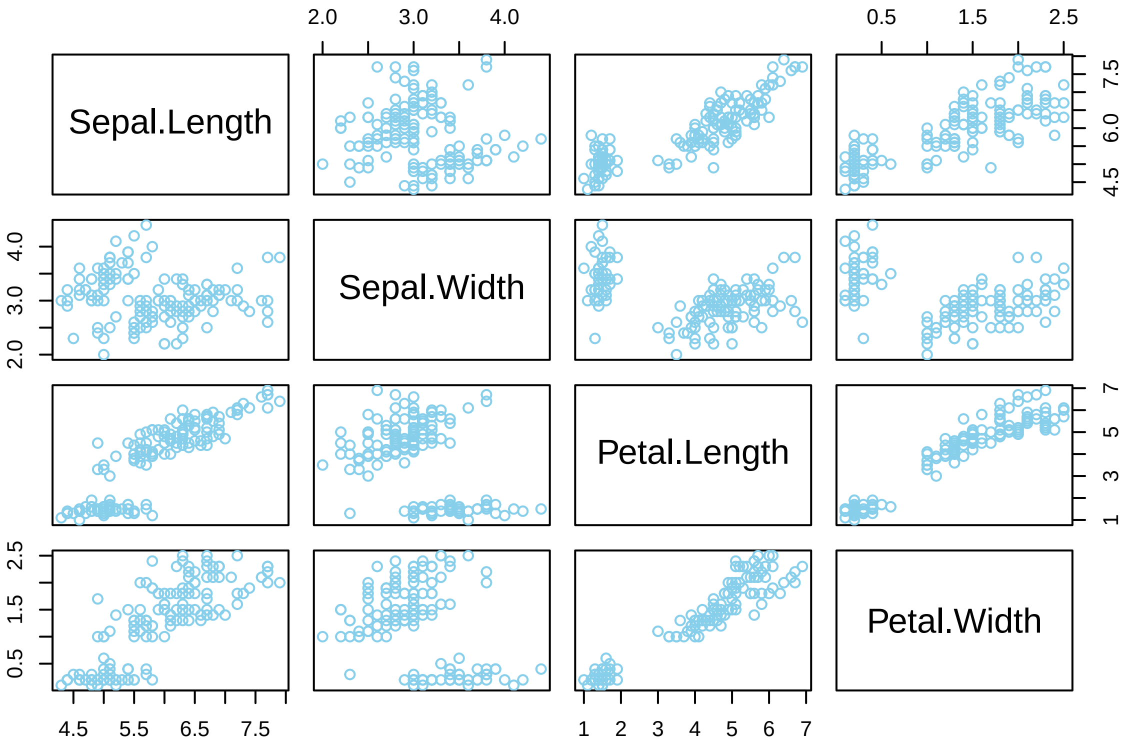
```
lapply(iris, mean)
```

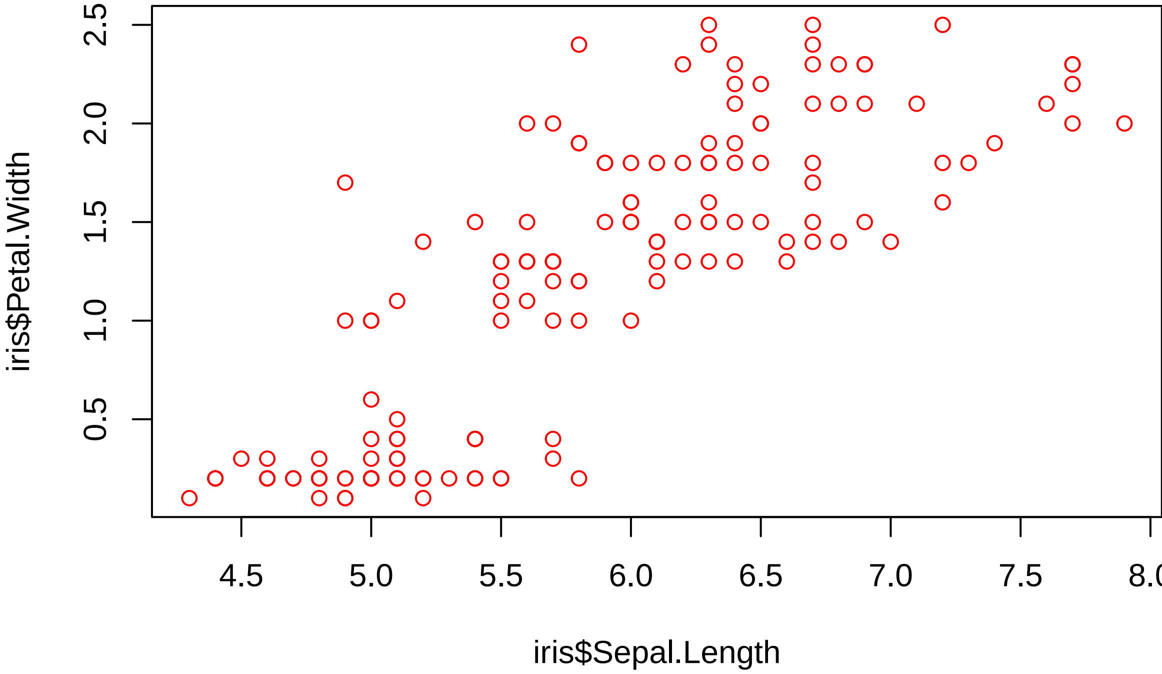
```
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA
```

```
## $Sepal.Length
## [1] 5.843333
##
## $Sepal.Width
## [1] 3.057333
##
## $Petal.Length
## [1] 3.758
##
## $Petal.Width
## [1] 1.199333
##
## $Species
## [1] NA
```

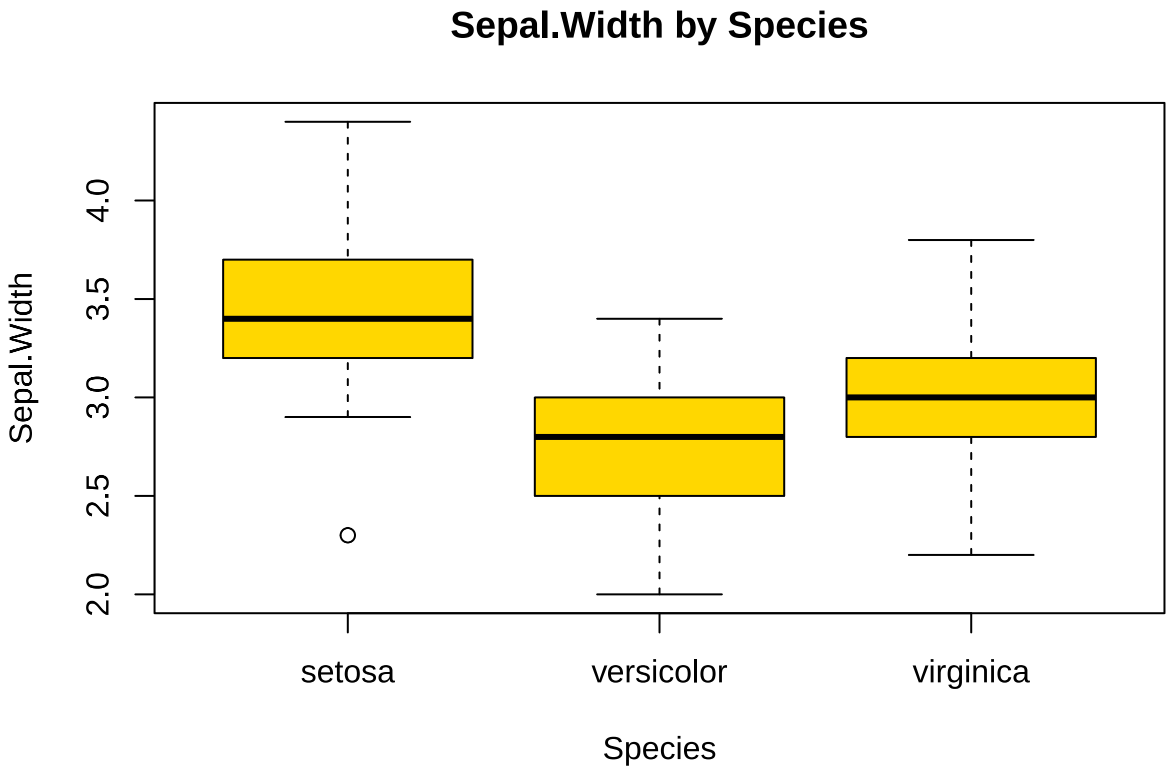
Plots and Hist

You can also embed plots, for example:

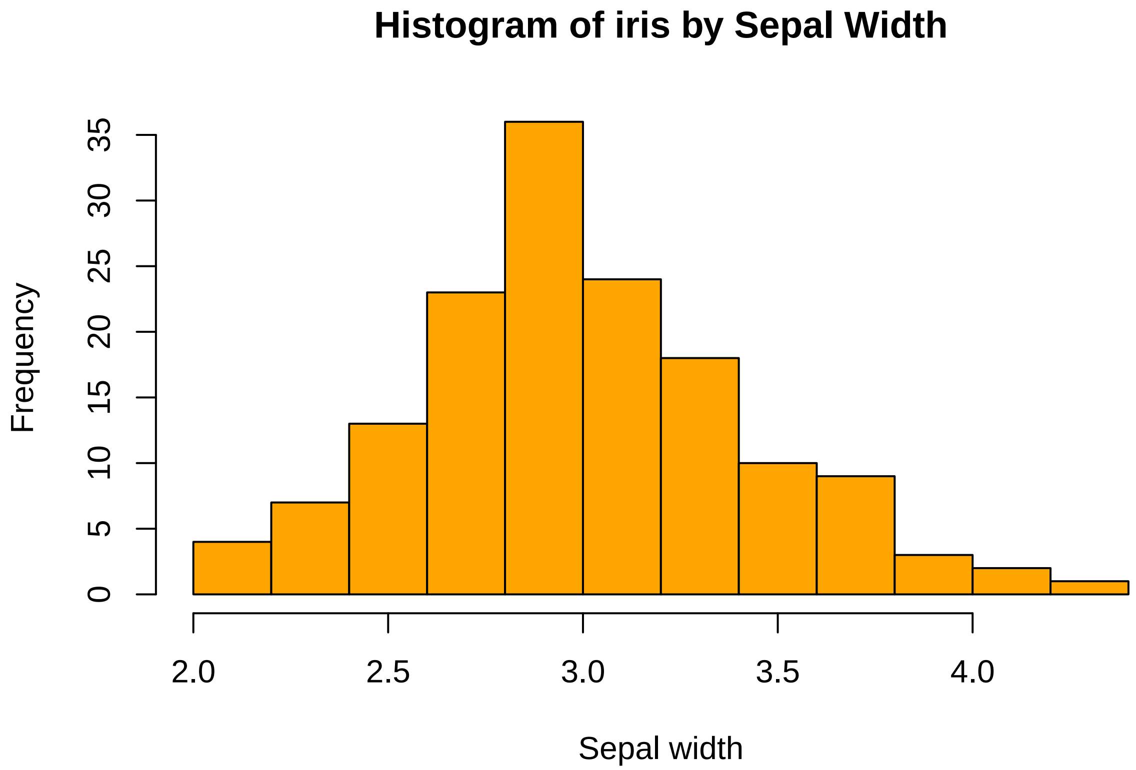




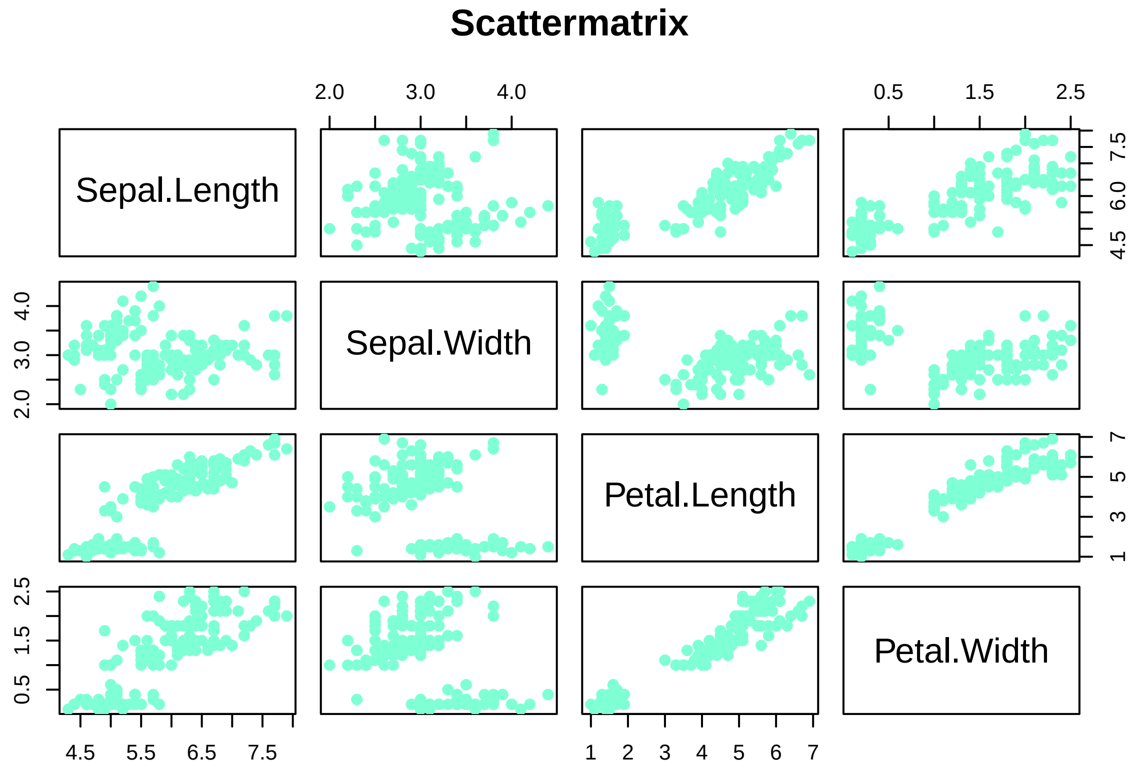
Try an specific plot



View a Boxplot
Histogram



Scatter Plot Matrices - R Base Graphs check : http://www.sthda.com/english/wiki/scatter-plot-matrices-r-base-graph



chek for ggtplot information https://www.r-graph-gallery.com/ggplot2-color.html

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.