MULTI-VIEW LAYOUT IN R

CHARLES WIREDU

2/22/2022

INSTALLATION OF REQUIRED PACKAGES

The following packages have to be installed first before the lab coding start.

```r
setwd("~/Desktop/Data science Assignments")


#install.packages("ggplot2")

#install.packages("tidyverse")

#install.packages("colorspace")
#install.packages("gridExtra")
#install.packages("egg")
#install.packages("gtable")
#install.packages("grid")
#install.packages("lattice")
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(colorspace)
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```r
library(egg)
library(gtable)
library(grid)
library(lattice)
```

LOADING IRIS DATASET

```
data(iris)

names(iris)

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

head(iris)

##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

DESCRIPTION OF DATASET

The data used for lab work is known as iris dataset, which is built in R. The data set contains measurements on four different attributes for flowers from three different species. The species are namely; setosa, versicolor and virginica. The attributes measurement are in centimeters. The data set has five variables coupled with 150 observations.This dataset was chosen because of its quality which has no missing values and can be deem as normally distributed. The function below gives a summary of this data set

```
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
##
##
##
```

MULTI -VIEW PLOT WITH IRIS DATASET

The function below illustrate a multi-view plot on the iris data set.The plot is made of boxplot, histogram, scatterplot, dotplot and facetted plot.The multi-view plot helps to facilitate control over layout.The variables used in this plots include sepal.width, sepal.length, petal.width,petal.length and species. The functions below illustrate the multi-view plot;

```
p1 <- qplot(Sepal.Length, Sepal.Width, data = iris, colour = Species)+
  ggtitle("Scatter plot on sepal.length")

p2 <- qplot(Sepal.Length, data = iris) +
  ggtitle("Histogram on Sepal.length")

p3 <- qplot(Sepal.Length, data = iris, geom = "dotplot")+
```

```
  ggtitle("Dotplot on sepal.length")

p4<- ggplot(iris, aes(x=Petal.Length, y=Petal.Width, fill=Species)) +
  geom_boxplot()+ggtitle("Boxplot on Petal-Iris data ")

p5 <- qplot(x = Sepal.Length,data = iris,binwidth = 0.2,fill = Species, xlab = "Sepal.Length (cm)")+
  ggtitle("Sepal.length Histogram")

p6 <-p1 + facet_wrap( ~ Species, nrow = 1) + theme(legend.position = "none") +
  ggtitle("facetted plot")

grid.arrange(p1,p2,p3,p4,p5,p6, nrow=3)
```
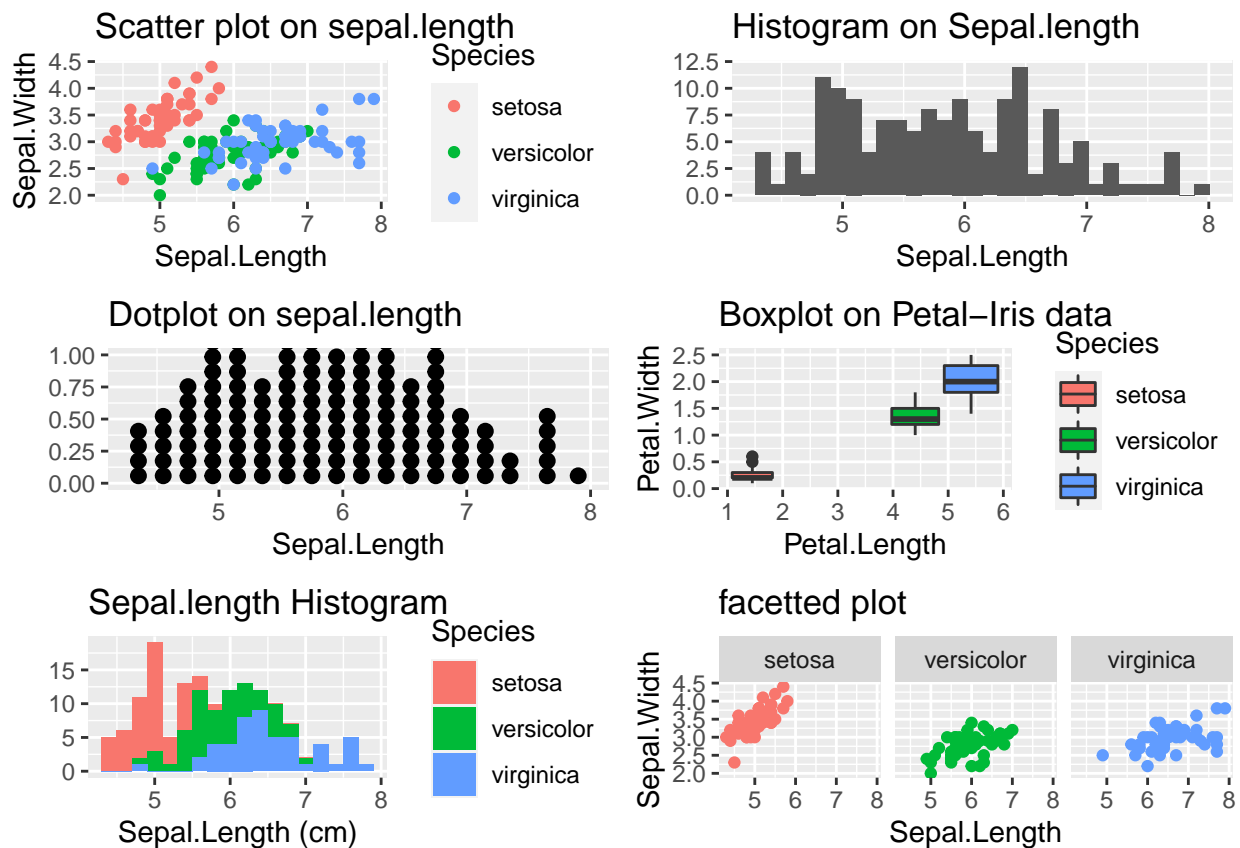
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

Bin width defaults to 1/30 of the range of the data. Pick better value with `binwidth`.

The plot above shows visualization of the iris data using different functions in R. The scatter plot shows the relationship between the two variables sepal.length and sepal.width. It can be observed that there is a positive relationship between sepal.length and sepal.width since most of the points on the scatter plot spread towards the right area of the plot.