****

**Prepared By:** Sandeep Singh

**Student ID:** C0727422

**Subject:** 2018S-T2 BDM 1053 - Big Data Algorithms and Statistics 01

**Title:** Assignment 7

**Week:** 09 July - 15 July

**Phone:** +1 705-977-1868

**Email (Internal):** c0727422@mylambton.ca

**Email (External):** official.sandeep93@gmail.com

**Submitted To:** Mr. [Pedram Habibi](https://moodle.queenscollege.ca/moodle/user/view.php?id=370&course=1590)

**Submission Date:** 19 July, 2018

**Contents**

[Explanation 3](#_Toc519770141)

[Pre-requisites 3](#_Toc519770142)

[Question 1 3](#_Toc519770143)

[Question 2 3](#_Toc519770144)

[Question 3 5](#_Toc519770145)

# Explanation

# Pre-requisites

Before performing this exercise, the following libraries must be installed in R:

* RcmdrPlugin.KMggplot2
* Ggplot2
* RColorBrewer
* Plotrix
* Cowplot

# Question 1

***Visualize the diamonds in diamonds dataset.***

library("RcmdrPlugin.KMggplot2", lib.loc="~/R/win-library/3.5")

df\_diamonds = diamonds

In this question, I have tried to visualize the prices of diamonds using scatter plot of price and carat using R commander. After importing the library *RcmdrPlugin.KMggplot2*, R commander opens up and we select necessary fields to plot the graphs. Screenshots have been attached for greater view.

# Question 2

***Create the following graphs for mtcars dataset: » Line graphs  » Scatterplots and scatterplot matrix  » Pie chart  » Barplot  Change colour of background to yellow, and of graph to green, and label axes and change title of the default graph.***

#Line Plots

ggplot(mtcars) + geom\_line(aes(x = value, y = mpg, colour = "mpg")) +

geom\_line(aes(x = value, y = cyl, colour = "cyl")) +

geom\_line(aes(x = value, y = gear, colour = "gear")) +

labs(title = "mtcars Visualization - Line Plot", x ="Value", y = "cyl, gear, mpg") +

theme(panel.background = element\_rect(fill = "springgreen"),

plot.background = element\_rect(fill = "yellow2"))

#Scatter Plots

ggplot(mtcars) + geom\_point(aes(mpg,cyl, colour = "mpg vs cyl")) +

geom\_point(aes(mpg,am, colour = "mpg vs am")) +

geom\_point(aes(mpg, gear, colour = "mpg vs gear")) +

labs(title = "mtcars Visualization - Scatter Plot", x = "mpg", y = "cyl, gear, mpg")+

theme(panel.background = element\_rect(fill = "springgreen"),

plot.background = element\_rect(fill = "yellow2"))

#Scatter plot matrices

plotmatrix(mtcars[, c(1,2,9,10)], colour = "species")

pairs(mtcars[, c(1,2,9,10)], col = "sandybrown", main = "mtcars Visualization - Scatter Plot Matrix") +

theme(panel.background = element\_rect(fill = "springgreen"),

plot.background = element\_rect(fill = "yellow2"))+

labs(title = "mtcars Visualization - Scatter Plot Matrix")

#Pie Chart

library(plotrix)

pie3D(x = table(mtcars$cyl), labels = c(8,4,6), explode = 0.2, main = "mtcars Visualization - 3D Plot", sub = "Number of Cylinders vs Vehicle Count")+

theme(panel.background = element\_rect(fill = "springgreen"),

plot.background = element\_rect(fill = "yellow2"))

#Bar Plot

ggplot(mtcars) + geom\_histogram(aes(x = mpg), binwidth = 2, boundary = 20, colour = "black", fill = "cyan") +

facet\_wrap(~cyl) +

theme(panel.background = element\_rect(fill = "springgreen"),

plot.background = element\_rect(fill = "yellow2")) +

labs(title = "mtcars Visualization - Bar Plot [No. of Cylinders with mpg]", x ="mpg", y = "Count", sub = "abc")

The requested plots have been plotted fulfilling the given conditions such as Title, legends, background colour, and graph colour.

# Question 3

***Create the following histogram graphs for VADeaths with suitable color palettes.***

df = data.frame(VADeaths)

histA = ggplot(df) + geom\_histogram(aes(x = Rural.Male), colour = "black", fill = "red")

histB = ggplot(df) + geom\_histogram(aes(x = Rural.Female), colour = "black", fill = "cyan")

histC = ggplot(df) + geom\_histogram(aes(x = Urban.Male), colour = "black", fill = "blue")

histD = ggplot(df) + geom\_histogram(aes(x = Urban.Female), colour = "black", fill = "pink")

library(cowplot)

plot\_grid(histA,histB, histC, histD)

VADeaths dataset has been converted to data frame. Histograms for each column has been plotted and finally combined all using plot\_grid() to compensate all the plots in single plot for better visualization.

\*\*\*\*\*