



ACADGILD

SESSION 7: Basic Statistics

Assignment 3

Submitted by: Munmun Ghosal

Login Id: munmun55@gmail.com

(M):+91-8007178659

Table of Contents

1. Problem Statement 3

2. Solution 3

1. Problem Statement

1. Create a **box and whisker plot** by class using **mtcars** dataset.

2. Solution

1. Create a **box and whisker plot** by class using **mtcars** dataset.

The R-script for the given problem is as follows:

```
library(readr)
library(ggplot2)
library(dplyr)
mtcars <- read_csv("E:/munmun_acadgild/acadgild data analytics/supporting
files/mtcars.csv")
View(mtcars)
str(mtcars)
mtcars1 <- mutate(mtcars,
  cyl = as.factor(cyl),
  disp = as.factor(disp),
  vs = as.factor(vs),
  am = as.factor(am),
  gear = as.factor(gear),
  carb = as.factor(carb),
  mpg = mpg, hp = hp, drat = drat, qsec=qsec)
str(mtcars1)
boxplot(mpg~carb, data = mtcars1, col =
c("Red","Green","Blue","Pink","yellow","orange"),main="Boxplot showing
distribution of mpg for each carb")
```

The output of the R-Script (from Console window) is given as follows:

The mtcars dataset is shown as follows:

	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
1	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
2	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
3	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
4	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
5	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
6	Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
7	Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
8	Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
9	Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
10	Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
11	Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
12	Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
13	Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
14	Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3

Showing 1 to 14 of 32 entries

```

> library(readr)
> library(ggplot2)
> library(dplyr)
> mtcars <- read_csv("E:/munmun_acadgild/acadgild data analytics/supporting
files/mtcars.csv")
Parsed with column specification:
cols(
  model = col_character(),
  mpg = col_double(),
  cyl = col_double(),
  disp = col_double(),
  hp = col_double(),
  drat = col_double(),
  wt = col_double(),
  qsec = col_double(),
  vs = col_double(),
  am = col_double(),
  gear = col_double(),
  carb = col_double()
)
> View(mtcars)
> str(mtcars)
Classes 'spec_tbl_df', 'tbl_df', 'tbl' and 'data.frame':    32 obs. of  12
variables:
 $ model: chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
 $ mpg  : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl  : num  6 6 4 6 8 6 8 4 4 6 ...
 $ disp : num  160 160 108 258 360 ...
 $ hp   : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat : num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt   : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec : num  16.5 17 18.6 19.4 17 ...
 $ vs   : num  0 0 1 1 0 1 0 1 1 1 ...
 $ am   : num  1 1 1 0 0 0 0 0 0 0 ...
 $ gear : num  4 4 4 3 3 3 3 4 4 4 ...
 $ carb : num  4 4 1 1 2 1 4 2 2 4 ...
- attr(*, "spec")=
 .. cols(
 ..   model = col_character(),
 ..   mpg = col_double(),
 ..   cyl = col_double(),
 ..   disp = col_double(),
 ..   hp = col_double(),
 ..   drat = col_double(),
 ..   wt = col_double(),
 ..   qsec = col_double(),
 ..   vs = col_double(),
 ..   am = col_double(),
 ..   gear = col_double(),
 ..   carb = col_double()
 .. )
> mtcars1 <- mutate(mtcars,
+                   cyl = as.factor(cyl),
+                   disp = as.factor(disp),
+                   vs = as.factor(vs),
+                   am = as.factor(am),

```

```

+           gear = as.factor(gear),
+           carb = as.factor(carb),
+           mpg = mpg, hp = hp, drat = drat, qsec=qsec)
> str(mtcars1)
Classes 'tbl_df', 'tbl' and 'data.frame':   32 obs. of  12 variables:
 $ model: chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
 $ mpg  : num   21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl  : Factor w/ 3 levels "4","6","8": 2 2 1 2 3 2 3 1 1 2 ...
 $ disp : Factor w/ 27 levels "71.1","75.7",...: 13 13 6 16 23 15 23 12 10 14
...
 $ hp   : num   110 110 93 110 175 105 245 62 95 123 ...
 $ drat : num    3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt   : num    2.62 2.88 2.32 3.21 3.44 ...
 $ qsec : num   16.5 17 18.6 19.4 17 ...
 $ vs   : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
 $ am   : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
 $ gear : Factor w/ 3 levels "3","4","5": 2 2 2 1 1 1 1 2 2 2 ...
 $ carb : Factor w/ 6 levels "1","2","3","4",...: 4 4 1 1 2 1 4 2 2 4 ...
>
> boxplot(mpg~carb, data = mtcars1, col =
c("Red","Green","Blue","Pink","yellow","orange"),main="Boxplot showing
distribution of mpg for each carb")

```

