

# HOMEWORK 1: GRAPHICAL METHODS FOR HIGH DIMENSIONAL DATA IN STATISTICS

ANIL A. AKSU

2-D and 3-D random data are generated by using built-in functions of *R* programming language and Car data is obtained by using *car* library in *R*. The particular script is given in Appendix A. In this homework, the graphical methods given by Grinstein et al. [2001] are used. These are Heat Map, 3-D scatter plot and Matrix of scatter plots.

## Heat Map

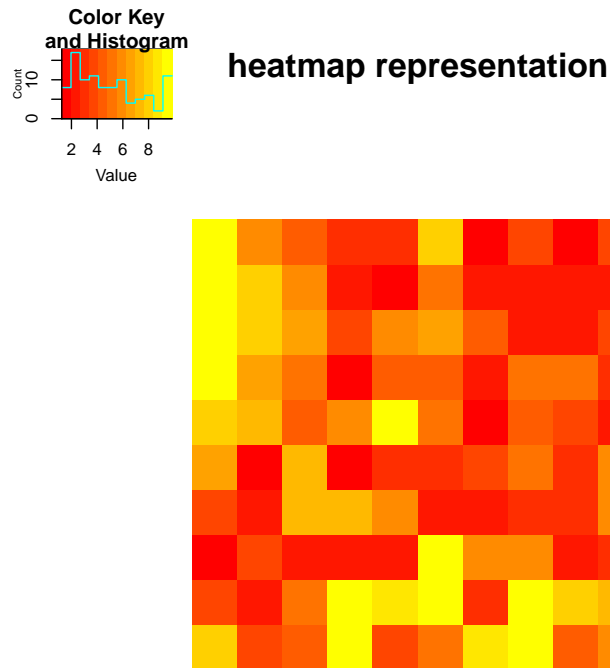


FIGURE 1. Heat map of random data set.

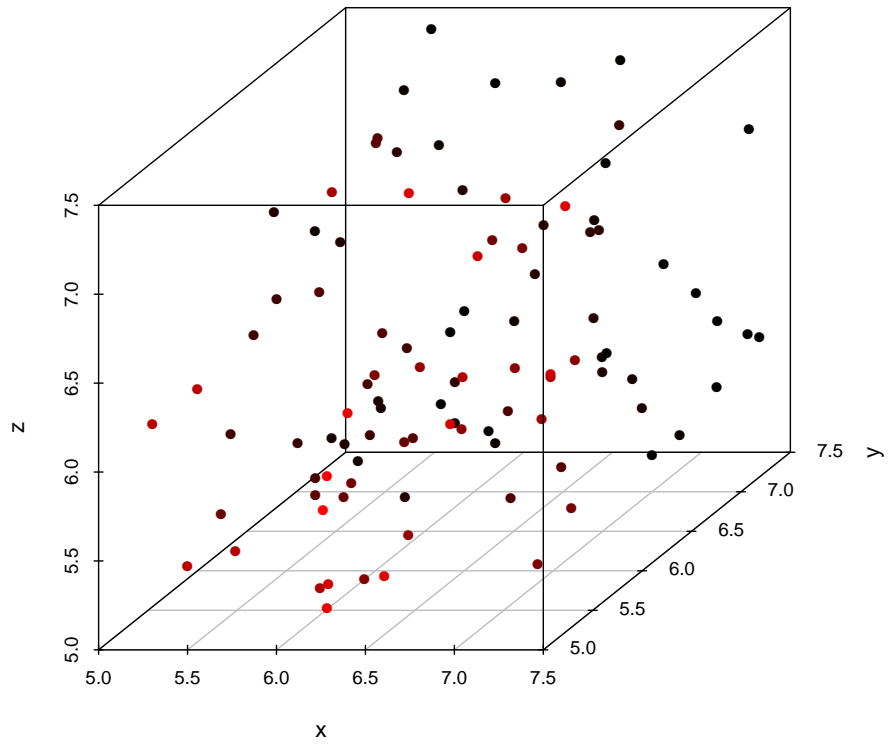
**3-D Scatter Plot****3D Scatterplot**

FIGURE 2. 3-D of Scatter Plot.

## Matrix of Scatter plots

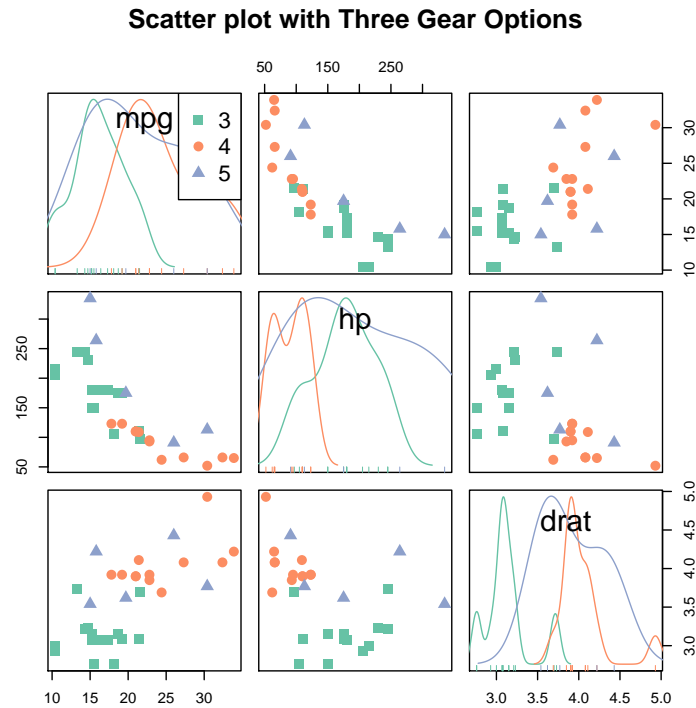


FIGURE 3. Matrix of Scatter plots.

## APPENDIX A. R CODES

```

1
2 #####
3 #
4 #      2-D Random Data by Anil Aksu
5 #      It is developed to show some graphical methods in R
6 #
7 #
8 #####
9
10 ## libraries
11 library(gplots)
12 library(rgl)
13 library(car)
14 library(RColorBrewer)
15 library(plotly)
16 library(scatterplot3d)

```

```
17
18 ## random data generation
19 size <- 100          #length of random number vectors
20 set.seed(1)
21 x <- runif(size, 5.0, 7.5)      # generate samples from ...
    uniform distribution (0.0, 1.0)
22 y <- runif(size, 5.0, 7.5)
23 z <- runif(size, 5.0, 7.5)
24 df <- matrix( runif(size,1., 10), 10, 10)
25
26 my_palette <- colorRampPalette(c("red", "yellow", ...
    "green"))(n = 299)
27
28 x11()
29 #scatter3d(x,y,z, main="PDF Scatterplot Example")
30 plot_ly(mtcars, x = ~wt, y = ~hp, z = ~qsec, color = ~am, ...
    colors = c('#BF382A', '#0C4B8E'))
31 x11()
32 heatmap.2(
33   df
34   , Colv = T
35   , Rowv = T
36   , trace="none"
37   , col = colorRampPalette(c('red', 'yellow'))(12)
38   , labRow=F
39   , labCol=F
40   , dendrogram="none"
41   , margins = c(5,5)
42   , main = "heatmap representation"
43 )
44
45 # Let's use the car dataset proposed by R
46 data=mtcars
47
48 # Make the plot
49 my_colors <- brewer.pal(nlevels(as.factor(data$cyl)), "Set2")
50
51 x11()
52 scatterplotMatrix(~mpg+hp+drat|gear, data=data , ...
    reg.line="", smoother="", col=my_colors , ...
    smoother.args=list(col="grey") , cex=1.5 , ...
    pch=c(15,16,17) , main="Scatter plot with Three Gear ...
    Options")
53 x11()
54 pdf("C:/Users/anil.aksu/Dropbox/Metu Phd Work/Courses/Fall ...
    2017/Bayesian Statistics/HW1/scatter3D.pdf")
55 scatterplot3d(x,y,z, pch=16, highlight.3d=TRUE, main="3D ...
    Scatterplot")
56 dev.off()
```

REFERENCES

Georges Grinstein, Marjan Trutschl, and Urška Cvek. High-dimensional visualizations. *Data mining conference KDD workshop*, pages 1–14, 2001. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.528{%&}rep=rep1{%&}type=pdf>.