MST 4053 Statistics Laboratory

Assignment No: 01   22/01/2010

1) Following are the number of items export by several countries during the last year.

Country	No of items
US	10
UK	12
Australia	4
Germany	16
France	8

   a) Draw a pie chart illustrating number of items export by each country.
   b) Put a suitable title and label the chart.
   c) Change the default colors of slice as you wish.
   d) Modify the above chart to annotate percentages.
   e) Draw a 3D pie chart to represent above data.
       (Hint : Download **plotrix** package from CRAN and use available help with it).

2) Consider the following three data sets and do the analysis separately for each one.

   8, 13, 16, 25, 26, 29, 30, 32, 37, 38, 40, 41, 44, 47, 49, 51, 54, 55, 58, 61, 63, 67, 75, 78, 82, 86, 95

   12, 0.4, 5, 2, 50, 8, 3, 1, 4, 0.25

   2 3 16 23 14 12 4 13 2 0 0 0 6 28 31 14 4 8 2 5

   a) Use **scan()** function to read data.
   b) Construct stem and leaf plot.
   c) Draw histogram.
   d) What are the advantages of stem and leaf plot comparing with histogram.
   e) What can you say about distribution of data?
   f) Draw box plot.
   g) Are there any outliers?
   h) Obtain descriptive statistics.

3) Type attach (*faithful*) command in the prompt and then the database is attached to the R search
   path. This means that the database is searched by R when evaluating a variable, so objects in
   the database can be accessed by simply giving their names. Where *faithful* is a built in data set in
   R and *eruptions* and *waiting* are two variables of it.

   a) Draw a frequencies histogram of variable *eruptions*.
   b) Put a suitable title, axis names and change the color of it as you wish.
   c) Add a normal curve to the above histogram.

d) Draw density histogram.

e) What can you say about distribution of data?

4) Use the available data set *mtcars*. Where variable *gear* represent the number of gear of a particular vehicle.

a) Count the number of vehicle belongs to each gear category.

b) Draw a bar plot to represent cars belong to each of three categories.

c) Draw horizontal bar plot for above case.

d) Variable *vs* represents the fuel categories and in which 1 indicates petrol vehicles. Draw stacked bar plot to represent above classification with respect to number of gears.

e) Draw a grouped bar plot for (d).

5) *mtcars* is a built in data set in R.

a) Draw scatter plot between *wt* and *mpg* variables in *mtcars*.

b) Draw regression line with red color.

c) Draw Scatterplot Matrices of variables *mpg, disp, drat and wt*.

d) You can create a 3D scatter plot by using **scatterplot3d** package. Use the function scatterplot3d(*x,y,z*) in the above package to draw 3D scatter plot.

e) Construct vertical line and regression plane.

6) Followings are the number of minutes spent by several people to complete some task.

81, 85, 93, 93, 99, 76, 75, 84, 78, 84, 81, 82, 89, 81, 96, 82, 74, 70, 84, 86, 80, 70, 131, 75, 88, 102, 115, 89, 82, 79, 106

a) Enter above data in an Excel worksheet and import into R working environment.

b) Draw histogram for above data.

c) Are there any skewness ?

d) Construct stem and leaf plot for above data?

e) Find mean, median, standard deviation, quantile, interquartile range and trim mean.

f) Draw a box plot and see weather outliers exist or not. If outliers exist remove those outliers and see what happens to the distribution.

7) Draw the frequency polygon for below data.

.314, .289, .282, .279, .275, .267, .266, .265, .256, .250, .249, .211, .161

Note

x=1:10

how many elements ?	length(x)
ith element	x[2] (i = 2)
all but ith element	x[-2] (i = 2)
first k elements	x[1:5] (k = 5)
last k elements	x[(length(x)-5):length(x)] (k = 5)
specific elements.	x[c(1,3,5)] (First, 3rd and 5th)
all greater than some value	x[x>3] (the value is 3)
bigger than or less than some values	x[x< -2	x > 2]
which indices are largest	which(x == max(x))