**STAT 40001/MA59800 Statistical Computing Fall 2017**

**Lab-6**

Q.N. 1) The competing risk analysis data are provided in the link below

<http://www.umass.edu/statdata/statdata/data/comprisk.dat>

1. Import the data in R
2. Determine the number of variables included in the data.

Q.N. 2) The Weight of Euro Coins: Its Distribution Might Not Be As Normal As You Would Expect written by Shkedy et al. is an article published in Journal of Statistics Education Volume 14, Number 2 (2006).

Go to <http://www.amstat.org/publications/jse/datasets/euroweight.dat.txt>

a) Import the ***euroweight.dat.txt*** data in R.  
> data = read.table("http://www.amstat.org//publications//jse//datasets//euroweight.dat.txt")

b) Select the third column batch of the coins.   
> attach(data)  
third\_col = V3

c) Create the frequency table of the batch of the coins.

> table(V3)

V3

1 2 3 4 5 6 7 8

250 250 250 250 250 250 250 250

Q.N. 3) Create a pie chart displaying the information given below and save it

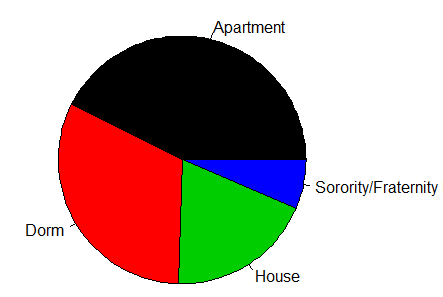
|  |  |
| --- | --- |
| Types of Housing | Frequency |
| Apartment | 20 |
| Dorm | 15 |
| House | 9 |
| Sorority/Fraternity House | 5 |

> house = c(20,15,9,3)

> names(house) = c("Apartment","Dorm","House","Sorority/Fraternity House")

> pie(house)

> pie(house,col = c(1,2,3,4))  
  
Q.N. 4) Go to <http://www.amstat.org/publications/jse/jse_data_archive.htm>



a) Import the babyboom.dat.txt data

> data = read.table("http://ww2.amstat.org//publications//jse//datasets//babyboom.dat.txt")

> head(data)

V1 V2 V3 V4

1 5 1 3837 5

2 104 1 3334 64

3 118 2 3554 78

4 155 2 3838 115

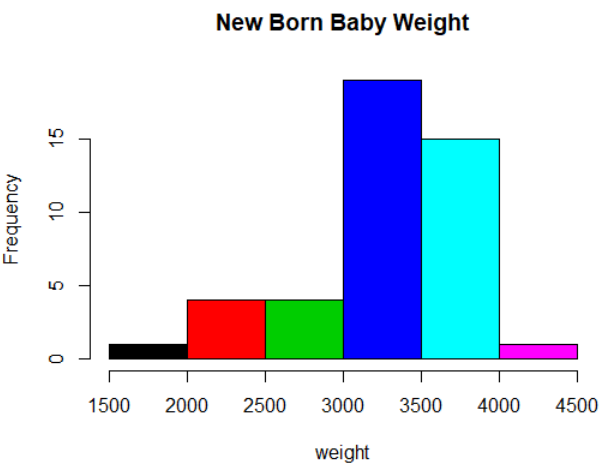
5 257 2 3625 177

6 405 1 2208 245

b) Select the column with the birth weight of new born babies.  
> weight = data$V3

c) Create a histogram of the subject data.

> hist(weight,main="New Born Baby Weight",col = c(1,2,3,4,5,6))



Q.N. 5) The link below provides a data file *homes* which includes monthly data regarding the number of new single-family houses sold in the U.S. in thousands(homes) and 30 year conventional mortgage rate (irate) from January, 1992 to March, 2010.

<http://www.principlesofeconometrics.com/poe4/poe4stata.htm>

1. Import the data in R

> data = read.dta("C:\\Users\\Administrator\\Desktop\\homes.dta")

1. Calculate the five number summary of homes and irate

> fivenum(data$homes)

[1] 324 654 840 964 1389

> fivenum(data$irate)

[1] 4.810 6.090 6.950 7.715 9.200

1. Draw a scatterplot to display the data.

> plot(data$homes, data$irate,xlab="homes",ylab="irate",main="dataset distribution")

