**STAT 40001/ MA 59800 Statistical Computing Fall 2017**

**Lab-4**

1) Go to <http://www.amstat.org/publications/jse/jse_data_archive.htm>

a) Import the ***babyboom.dat.txt*** data which has Time of Birth, Sex, and Birth Weight of Babies and number of minutes after midnight.

b) How many observations are recorded?  
> babyboom = read.table("http://ww2.amstat.org//publications//jse//datasets//babyboom.dat.txt");

> colnames(babyboom) <- c("Time of Birth", "Sex","Birth Weight","number of minutes after midnight");  
> dim(babyboom)

[1] 44 4

c) Print first 5 observations  
> head(babyboom,5)  
 Time of Birth Sex Birth Weight number of minutes after midnight

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

d) Print last 5 observations  
> tail(babyboom,5)  
 Time of Birth Sex Birth Weight number of minutes after midnight

40 2104 2 2121 1264

41 2123 2 3150 1283

42 2217 1 3866 1337

43 2327 1 3542 1407

44 2355 1 3278 1435

2) Install *UsingR* pacakage in your computer and access the data *BushApproval* . How many variables are listed in the data set?  
> dim(BushApproval)

[1] 323 3

3) Murder rates for some Southern US cities is provided in the data set named  *south* in the *UsingR* package.

a) How many states are included in the data set?

> length(south)

[1] 30

b) Print the variable list.

> south

[1] 12 10 10 13 12 12 14 7 16 18 8 29 12 14 33 10 6 18 11 25 8 16 14 11 10 20 14 11 12 13

4) For 1997-2002, from police-reported car crashes in which there is a harmful event (people or property), and from which at least one vehicle was towed (Data are restricted to front-seat occupants, include only a subset of the variables recorded, and are restricted in other ways also), are provided as Item “nassCDS” under Reich source of data in Useful link.

a) Import the data in R using appropriate R code

data = read.table("https://raw.githubusercontent.com/vincentarelbundock/Rdatasets/master/csv/DAAG/nassCDS.csv",header=T,sep = ',')

b) Print the first 5 observations.

> head(data,5)

X dvcat weight dead airbag seatbelt frontal sex ageOFocc yearacc yearVeh abcat occRole deploy injSeverity caseid

1 1 25-39 25.069 alive none belted 1 f 26 1997 1990 unavail driver 0 3 2:3:1

2 2 10-24 25.069 alive airbag belted 1 f 72 1997 1995 deploy driver 1 1 2:3:2

3 3 10-24 32.379 alive none none 1 f 69 1997 1988 unavail driver 0 4 2:5:1

4 4 25-39 495.444 alive airbag belted 1 f 53 1997 1995 deploy driver 1 1 2:10:1

5 5 25-39 25.069 alive none belted 1 f 32 1997 1988 unavail driver 0 3 2:11:1

5) Consider a data set where the columns are separated by $

Col1$Col2$Col3

1$2$3

4$5$6

7$8$9

a$b$c

Save the data in a local drive and import it in R by removing $ sign.

> data = read.delim("C:\\Users\\Administrator\\Desktop\\task5.txt",sep = "$");

> data

Col1 Col2 Col3

1 1 2 3

2 4 5 6

3 7 8 9

4 a b c