ROSANA LIN HO - STAT 521

```
### LAB 2 – Data input and Output ###

# Question 1 #
> quakes[1:5,]
     lat   long depth mag stations
1 -20.42 181.62   562 4.8       41
2 -20.62 181.03   650 4.2       15
3 -26.00 184.10    42 5.4       43
4 -17.97 181.66   626 4.1       19
5 -20.42 181.96   649 4.0       11

# a #
> summary(quakes)
      lat              long            depth            mag          stations
 Min.   :-38.59   Min.   :165.7   Min.   : 40.0   Min.   :4.00   Min.   : 10.
00
 1st Qu.:-23.47   1st Qu.:179.6   1st Qu.: 99.0   1st Qu.:4.30   1st Qu.: 18.
00
 Median :-20.30   Median :181.4   Median :247.0   Median :4.60   Median : 27.
00
 Mean   :-20.64   Mean   :179.5   Mean   :311.4   Mean   :4.62   Mean   : 33.
42
 3rd Qu.:-17.64   3rd Qu.:183.2   3rd Qu.:543.0   3rd Qu.:4.90   3rd Qu.: 42.
00
 Max.   :-10.72   Max.   :188.1   Max.   :680.0   Max.   :6.40   Max.   :132.
00
> dim(quakes)
[1] 1000    5          # 5 Variables and 1000 Observations
> names(quakes)
[1] "lat"      "long"     "depth"    "mag"      "stations"  # Variable's names
> str(quakes)          # Types and descriptions/units
'data.frame':  1000 obs. of  5 variables:
 $ lat     : num  -20.4 -20.6 -26 -18 -20.4 ...
 $ long    : num  182 181 184 182 182 ...
 $ depth   : int  562 650 42 626 649 195 82 194 211 622 ...
 $ mag     : num  4.8 4.2 5.4 4.1 4 4 4.8 4.4 4.7 4.3 ...
 $ stations: int  41 15 43 19 11 12 43 15 35 19 ...


# b - 100st obs#
> quakes[100,]
       lat   long depth mag stations
100 -24.57 179.92   484 4.7       33


# c - Subset of quakes #
> bigquakes =subset(quakes, quakes$mag > 5)
> View(bigquakes)
> # c - Subset of quakes #
> bigquakes =subset(quakes, quakes$mag > 5)
> View(bigquakes)
> dim(bigquakes)
[1] 151   5            # 151 observations in new dataset
> sum(quakes$mag)      #Sum Magnitude of original data
[1] 4620.4
> sum(bigquakes$mag)   #Sum Magnitude of Big quakes
[1] 804.7
> 100*(sum(bigquakes$mag)/sum(quakes$mag))  # % of magnitude from big quakes
[1] 17.41624
```

```
# d #
# Menu:  Session -> Set Working Directory
setwd("Learning/MS_Applied_Stat/STAT521/")

# saving a dataset in text format
write.csv(bigquakes, file = 'bigquakes_csv.csv')

# saving a dataset in SAS format
library(foreign)
> write.foreign(bigquakes, datafile ="bigquakes_sas.sasdat", codefile ="bigqu
akes_sas_codefile", package = "SAS"


# saving a dataset in Excell format
library("openxlsx")
> write.xlsx(bigquakes,file = "bigquakes_xlsx.xlsx", sheetName = "Sheet1", co
l.names = TRUE, row.names = TRUE, append = FALSE)

# Question 2 #
# a #
> df2 = list(volcano=volcano,rivers=rivers,rock=rock)
> summary(df2)
        Length Class      Mode
volcano 5307   -none-     numeric
rivers   141   -none-     numeric
rock       4   data.frame list

# b #
# Menu:  Session -> Set Working Directory
setwd("Learning/MS_Applied_Stat/STAT521/")

dput(df2, file = "df2")

# c #
It does not look usable because the structure is not based on rows and column

# d #
> df2_d = dget("df2")
> summary(df2_d)
        Length Class      Mode
volcano 5307   -none-     numeric
rivers   141   -none-     numeric
rock       4   data.frame list
```