Week 3 - READING AND WRITING DIFFERENT TYPES OF DATASETS

a. Reading different types of data sets (.txt, .csv) from web and disk and writing in file in specific disk location.

```
library(utils)
data<- read.csv("input.csv")
data
```

Output :-

```
id, name, salary, start_date, dept 1
1 Rick 623.30 2012-01-01 IT
2 2 Dan 515.20 2013-09-23 Operations
3 3 Michelle 611.00 2014-11-15 IT 4 4 Ryan 729.00
2014-05-11 HR
5 NA Gary 843.25 2015-03-27 Finance
6 6 Nina 578.00 2013-05-21 IT
7 7 Simon 632.80 2013-07-30 Operations
8 8 Guru 722.50 2014-06-17 Finance
```

```
data<- read.csv("input.csv")

print(is.data.frame(data))
print(ncol(data))
print(nrow(data))
```

Output:-

```
[1] TRUE
[1] 5
[1] 8
```

```
# Create a data frame.
data<- read.csv("input.csv")
# Get the max salary from data frame.
sal<- max(data$salary)
sal
```

Output:-

```
[1] 843.25
```

```
# Create a data frame.
data<- read.csv("input.csv")
```

```
# Get the max salary from data frame.
sal<- max(data$salary)

# Get the person detail having max salary.
retval<- subset(data, salary == max(salary))
retval
```

Output:-

```
id    name  salary  start_datedept 5     NA
Gary  843.25  2015-03-27    Finance
```

Get all the people working in IT department

```
# Create a data frame.
data<- read.csv("input.csv")

retval<- subset( data, dept == "IT")
retval
```

Output:-

```
id   name      salary   start_datedept 1
1    Rick      623.3    2012-01-01   IT
3     3    Michelle  611.0    2014-11-15   IT
6     6    Nina      578.0    2013-05-21   IT
```

```
#Create a data frame.

data<- read.csv("input.csv")
retval<- subset(data, as.Date(start_date) >as.Date("2014-01-01"))

# Write filtered data into a new file.

write.csv(retval,"output.csv")
newdata<- read.csv("output.csv")
newdata
```

Output:-

```
  X      id   name      salary   start_datedept
1 3      3    Michelle  611.00   2014-11-15    IT
2 4      4    Ryan      729.00   2014-05-11    HR
3 5     NA    Gary      843.25   2015-03-27    Finance
```

4 8 8 Guru 722.50 2014-06-17 Finance

b. Reading Excel data sheet in R.

```
install.packages("xlsx")
library("xlsx")
data<- read.xlsx("input.xlsx", sheetIndex = 1)
data
```

Output:-

```
id, name, salary, start_date, dept 1
1 Rick 623.30 2012-01-01 IT
2 2 Dan 515.20 2013-09-23 Operations
3 3 Michelle 611.00 2014-11-15 IT 4 4 Ryan 729.00
2014-05-11 HR
5 NA Gary 843.25 2015-03-27 Finance
6 6 Nina 578.00 2013-05-21 IT
7 7 Simon 632.80 2013-07-30 Operations
8 8 Guru 722.50 2014-06-17 Finance
```

c. Reading XML dataset in R.

```
install.packages("XML")
library("XML") library("methods")
result<- xmlParse(file =
"input.xml")
result
```

Output:-

```
1
  Rick
  623.3
  1/1/2012
  IT


  2
  Dan
  515.2
```

9/23/2013
Operations

3
Michelle
611
11/15/2014
IT

4
Ryan
729
5/11/2014
HR

5
Gary
843.25
3/27/2015
Finance

6
Nina
578
5/21/2013
IT

7
Simon
632.8
7/30/2013
Operations

8
Guru
722.5
6/17/2014
Finance