DAY 2 LAB MANNUAL PART 2- FEB 9
ITA0443-STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS
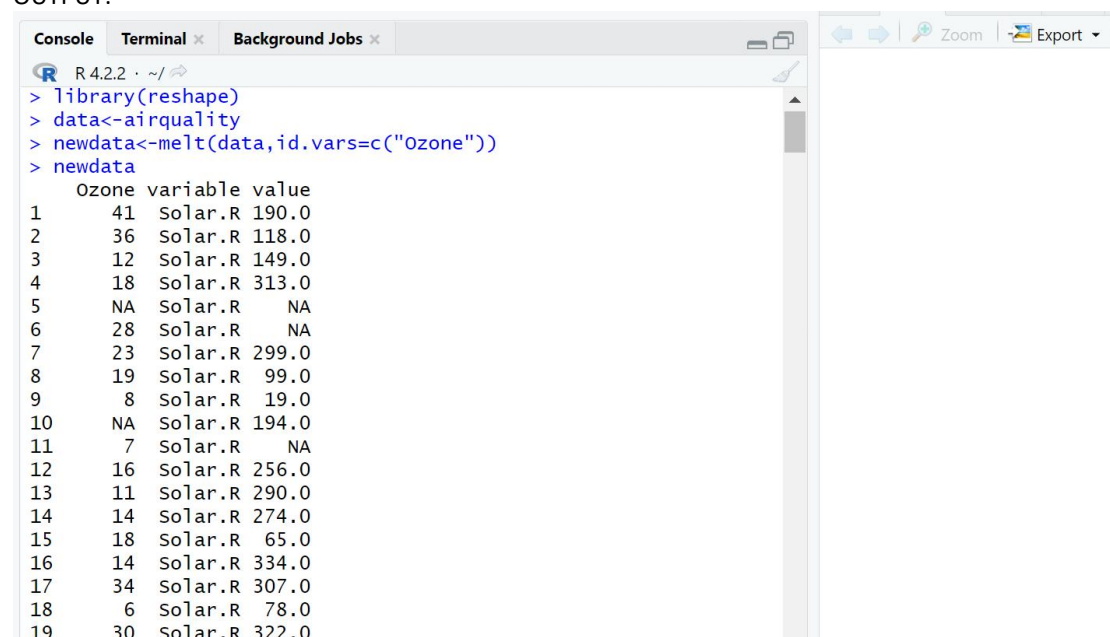GITHUBLINK:-https://github.com/Vamsim29/ITA0443-STATISTICS-WITH-R-PROGRAMMING

1. Melt airquality data set and display as a long – format data ?

PROGRAM:-
```
library(reshape)
data<-airquality
newdata<-melt(data,id.vars=c("Ozone"))
newdata
```

OUTPUT:-

2. Melt airquality data and specify month and day to be "ID variables"?

PROGRAM:-

```
library(reshape)
data<-airquality
newdata<-melt(data,id.vars=c("Month","Day"))
Newdata
```

OUTPUT:-

```
Console   Terminal ×   Background Jobs ×                        Zoom
  R  R 4.2.2 · ~/
> library(reshape)
> data<-airquality
> newdata<-melt(data,id.vars=c("Month","Day"))
> newdata
   Month Day variable value
1      5   1    Ozone    41
2      5   2    Ozone    36
3      5   3    Ozone    12
4      5   4    Ozone    18
5      5   5    Ozone    NA
6      5   6    Ozone    28
7      5   7    Ozone    23
8      5   8    Ozone    19
9      5   9    Ozone     8
10     5  10    Ozone    NA
11     5  11    Ozone     7
12     5  12    Ozone    16
13     5  13    Ozone    11
14     5  14    Ozone    14
15     5  15    Ozone    18
16     5  16    Ozone    14
17     5  17    Ozone    34
18     5  18    Ozone     6
19     5  19    Ozone    30
```

3.Cast the molten airquality data set .

PROGRAM:-
library(reshape)
data<-airquality
longdata<-melt(data,id.vars=c("Ozone"))
shortdata<-cast(longdata,Ozone~variable,sum)
shortdata

OUTPUT:-

```
Source                                                    Files   Plots   Pa

Console   Terminal ×   Background Jobs ×                          Zoom
  R  R 4.2.2 · ~/
> library(reshape)
> data<-airquality
> longdata<-melt(data,id.vars=c("Ozone"))
> shortdata<-cast(longdata,Ozone~variable,sum)
> shortdata
   Ozone Solar.R  Wind Temp Month Day
1      1       8   9.7   59     5  21
2      4      25   9.7   61     5  23
3      6      78  18.4   57     5  18
4      7      NA  31.5  223    21  50
5      8      19  20.1   61     5   9
6      9      84  39.0  224    25  38
7     10     264  14.3   73     7  12
8     11     654  35.5  201    15  55
9     12     269  24.1  147    11  22
10    13     514  44.7  287    33  74
11    14     819  53.3  270    28  83
12    16     541  32.0  307    29  56
13    18     733  46.5  263    28  65
14    19      99  13.8   59     5   8
15    20     593  40.2  295    31  81
16    21     939  56.8  305    32  58
```
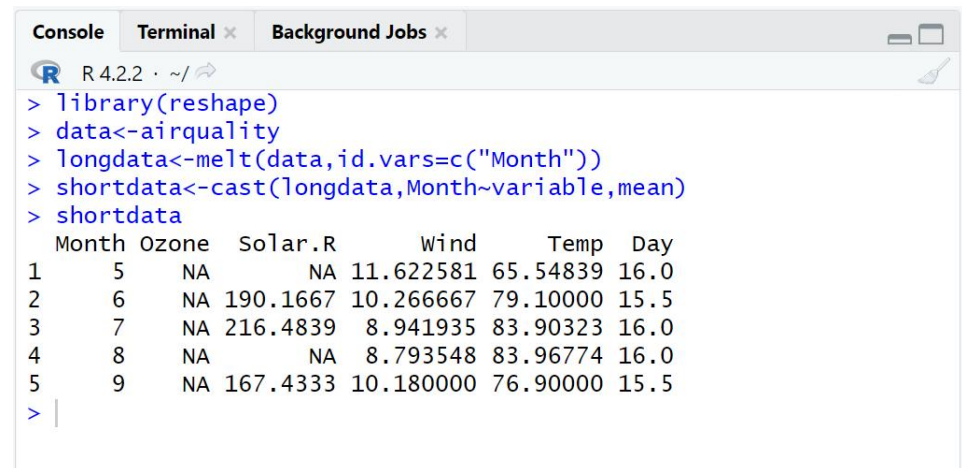
4.Use cast function appropriately and compute the average of Ozone, Solar.R , Wind and temperature per month ?

PROGRAM:-
```
library(reshape)
data<-airquality
longdata<-melt(data,id.vars=c("Month"))
shortdata<-cast(longdata,Month~variable,mean)
shortdata
```

OUTPUT:-

```
Console   Terminal ×   Background Jobs ×                              ─ □

R   R 4.2.2 · ~/ ⇗
> library(reshape)
> data<-airquality
> longdata<-melt(data,id.vars=c("Month"))
> shortdata<-cast(longdata,Month~variable,mean)
> shortdata
  Month Ozone  Solar.R      Wind     Temp Day
1     5    NA       NA 11.622581 65.54839 16.0
2     6    NA 190.1667 10.266667 79.10000 15.5
3     7    NA 216.4839  8.941935 83.90323 16.0
4     8    NA       NA  8.793548 83.96774 16.0
5     9    NA 167.4333 10.180000 76.90000 15.5
> |
```
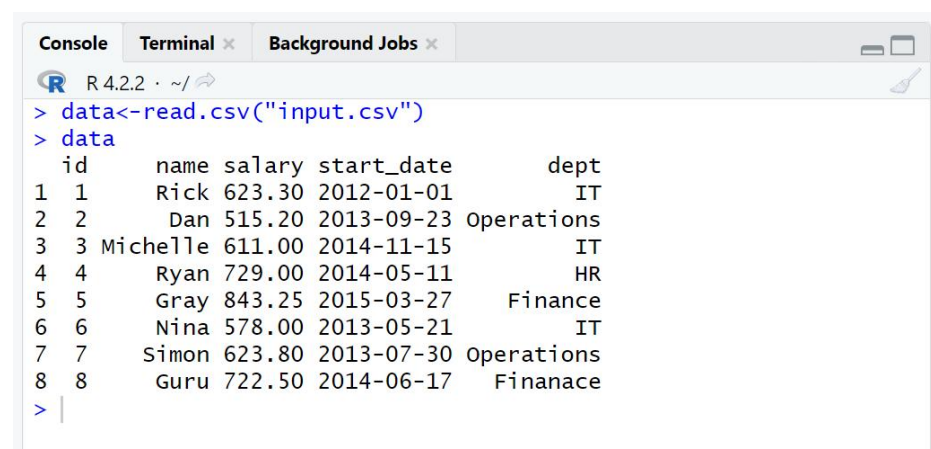
8 FILE MANUPULATION IN R

1. Consider the following data present. Create this file using windows notepad . Save the file as input.csv using the save As All files(*.*) option in notepad.
2.Use appropriate R commands to read input.csv file.

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

OUTPUT:-

```
Console   Terminal ×   Background Jobs ×                              ─ □

R   R 4.2.2 · ~/ ⇗
> data<-read.csv("input.csv")
> data
  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  5     Gray 843.25 2015-03-27    Finance
6  6     Nina 578.00 2013-05-21         IT
7  7    Simon 623.80 2013-07-30 Operations
8  8     Guru 722.50 2014-06-17   Finanace
> |
```
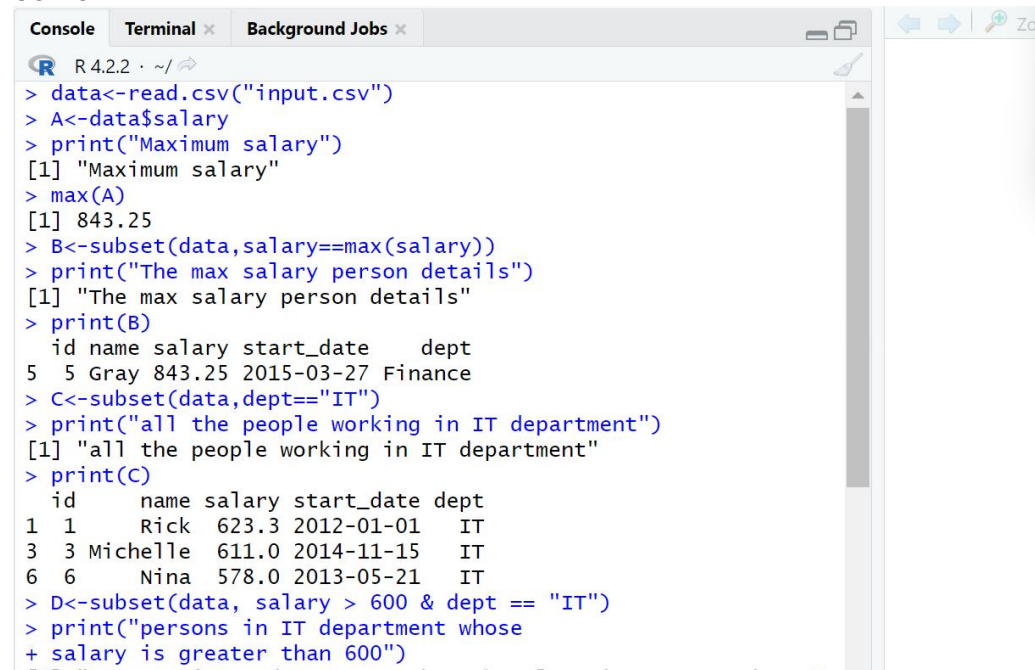
3.Analyze the CSV File and compute the following.
a. Get the maximum salary
b. Get the details of the person with max salary

c. Get all the people working in IT department
d. Get the persons in IT department whose salary is greater than 600
e. Get the people who joined on or after 2014

PROGRAM:-

```
data<-read.csv("input.csv")
A<-data$salary
print("Maximum salary")
max(A)
B<-subset(data,salary==max(salary))
print("The max salary person details")
print(B)
C<-subset(data,dept=="IT")
print("all the people working in IT department")
print(C)
D<-subset(data, salary > 600 & dept == "IT")
print("persons in IT department whose
salary is greater than 600")
print(D)
print("Get the people who joined on or after 2014")
retval <- subset(data, as.Date(start_date) > as.Date("2014-01-01"))
print(retval)
print(retval)
```

OUTPUT:-

```
Console   Terminal ×   Background Jobs ×

R  R 4.2.2 · ~/
> data<-read.csv("input.csv")
> A<-data$salary
> print("Maximum salary")
[1] "Maximum salary"
> max(A)
[1] 843.25
> B<-subset(data,salary==max(salary))
> print("The max salary person details")
[1] "The max salary person details"
> print(B)
  id name salary start_date    dept
5  5 Gray 843.25 2015-03-27 Finance
> C<-subset(data,dept=="IT")
> print("all the people working in IT department")
[1] "all the people working in IT department"
> print(C)
  id     name salary start_date dept
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
> D<-subset(data, salary > 600 & dept == "IT")
> print("persons in IT department whose
+ salary is greater than 600")
```

```
+ salary is greater than 600 )
[1] "persons in IT department whose \nsalary is greater than 60
0"
> print(D)
  id      name salary start_date dept
1  1      Rick  623.3 2012-01-01   IT
3  3  Michelle  611.0 2014-11-15   IT
> print("Get the people who joined on or after 2014")
[1] "Get the people who joined on or after 2014"
> retval <- subset(data, as.Date(start_date) > as.Date("2014-01-
01"))
> print(retval)
  id      name salary start_date      dept
3  3  Michelle 611.00 2014-11-15        IT
4  4      Ryan 729.00 2014-05-11        HR
5  5      Gray 843.25 2015-03-27   Finance
8  8      Guru 722.50 2014-06-17  Finanace
```

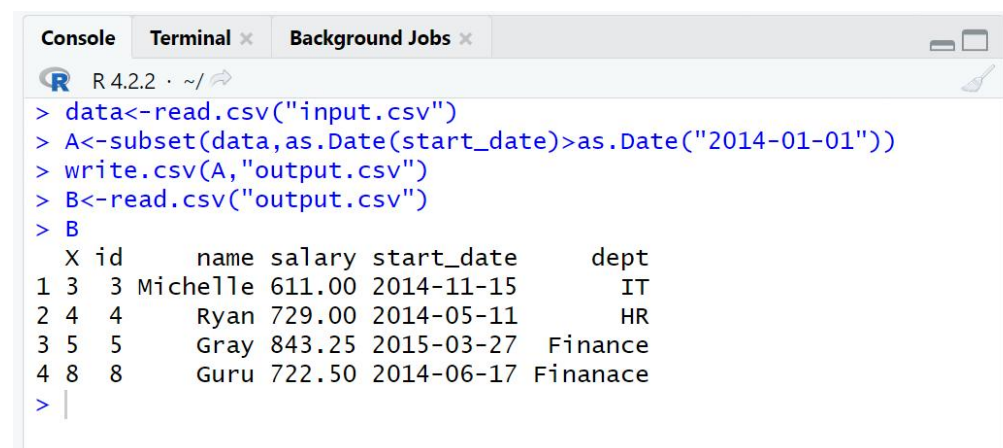4.Get the people who joined on or after 2014 and write the output onto a file called
output.csv

PROGRAM:-
data<-read.csv("input.csv")
A<-subset(data,as.Date(start_date)>as.Date("2014-01-01"))
write.csv(A,"output.csv")
B<-read.csv("output.csv")
B

OUTPUT:-

```
Console   Terminal ×   Background Jobs ×

R  R 4.2.2 · ~/
> data<-read.csv("input.csv")
> A<-subset(data,as.Date(start_date)>as.Date("2014-01-01"))
> write.csv(A,"output.csv")
> B<-read.csv("output.csv")
> B
  X id      name salary start_date      dept
1 3  3  Michelle 611.00 2014-11-15        IT
2 4  4      Ryan 729.00 2014-05-11        HR
3 5  5      Gray 843.25 2015-03-27   Finance
4 8  8      Guru 722.50 2014-06-17  Finanace
> |
```