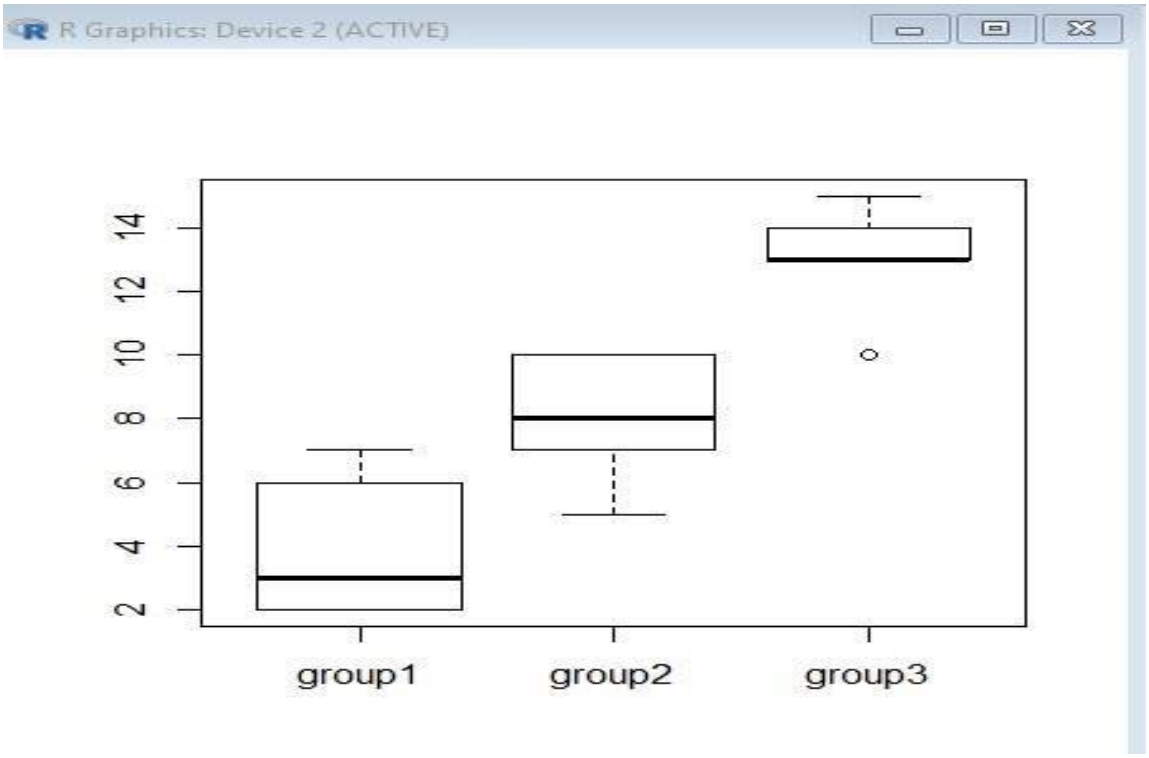
Name: Sumit Singh
Roll No: 380
Class: TYBSC CS A
Subject: Data Science
Practical No : 5

Aim: Demonstration of Analysis of Variance. Code:

#CREATE THE DATA IN TO THREE GROUPS #TO PRINT THE BOXPLOT

```
> group1=c(2,3,7,2,6)
> group2=c(10,8,7,5,10)
> group3=c(10,13,14,13,15)
> cg=data.frame(cbind(group1,group2,group3))
> cg
  group1 group2 group3
1      2     10     10
2      3      8     13
3      7      7     14
4      2      5     13
5      6     10     15
> boxplot(cg)
>
```

BOXPLOT



#TO PRINT THE DATA INTO STACK FORMATE

```
> stacked_g=stack(cg)
> stacked_g
   values     ind
1       2  group1
2       3  group1
3       7  group1
4       2  group1
5       6  group1
6      10  group2
7       8  group2
8       7  group2
9       5  group2
10     10  group2
11     10  group3
12     13  group3
13     14  group3
14     13  group3
15     15  group3
> |
```

TAKE ANOTHER DATASET AND WORK ON THAT.

 # CREATE THE DATA IN TO THREE GROUPS

```
> av=aov(values~ind,data=stacked_g)
> summary(av)
            Df Sum Sq Mean Sq F value   Pr(>F)
ind          2  203.3   101.7   22.59 8.54e-05 ***
Residuals   12   54.0     4.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> g1=c(29,30,31,31,29)
> g2=c(28,29,27,30,29)
> g3=c(25,28,29,27,29)
> cg1=data.frame(cbind(g1,g2,g3))
> cg1
  g1 g2 g3
1 29 28 25
2 30 29 28
3 31 27 29
4 31 30 27
5 29 29 29
```

```
> stacked_g= stack(cgl)
> stacked_g
   values ind
1      29  g1
2      30  g1
3      31  g1
4      31  g1
5      29  g1
6      28  g2
7      29  g2
8      27  g2
9      30  g2
10     29  g2
11     25  g3
12     28  g3
13     29  g3
14     27  g3
15     29  g3
```

```
> av=aov(values~ind,data=stacked_g)
> avl=aov(values~ind,data=stacked_g)
> summary(avl)
            Df Sum Sq Mean Sq F value Pr(>F)
ind          2  14.53   7.267   4.275 0.0397 *
Residuals   12  20.40   1.700
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

Conclusion: Hence, we successfully performed Analysis of variance.