Date: 29/Feb/2024	ANALYSIS OF VARIANCE(ANNOVA)
EXPERIMENT – 06	

**AIM:** To perform analysis of variance(annova)

SOFTWARE REQUIRED: RStudio

R CODE:

```
rm(list=ls())

data<-data.frame(sale.count=c
(40,60,70,30,50,30,30,10,70,60,50,60,30,20,20),type=c("Can-
A","Can-A","Can-A","Can-A","Can-A","Can-B","Can-B","Can-B","Can-
B","Can-B","Can-C","Can-C","Can-C","Can-C","Can-C"))
library(dplyr)
group_by(data,type)                                        %>%
summarise(count=n(),mean=mean(sale.count,na.rm=TRUE))
result<-aov(sale.count~type,data=data)
summary(result)
data<-PlantGrowth
group_by(data,group)                                       %>%
summarise(count=n(),mean=mean(weight,na.rm=TRUE))
result<-aov(weight~group,data=data)
summary(result)
TukeyHSD(result)
plot(result,1)
plot(result,2)
kruskal.test(weight~group,data=data)
data<- read.csv("iris.csv",row.names=1)
```

OUTPUT:

Q-Q Residuals

Theoretical Quantiles
aov(weight ~ group)
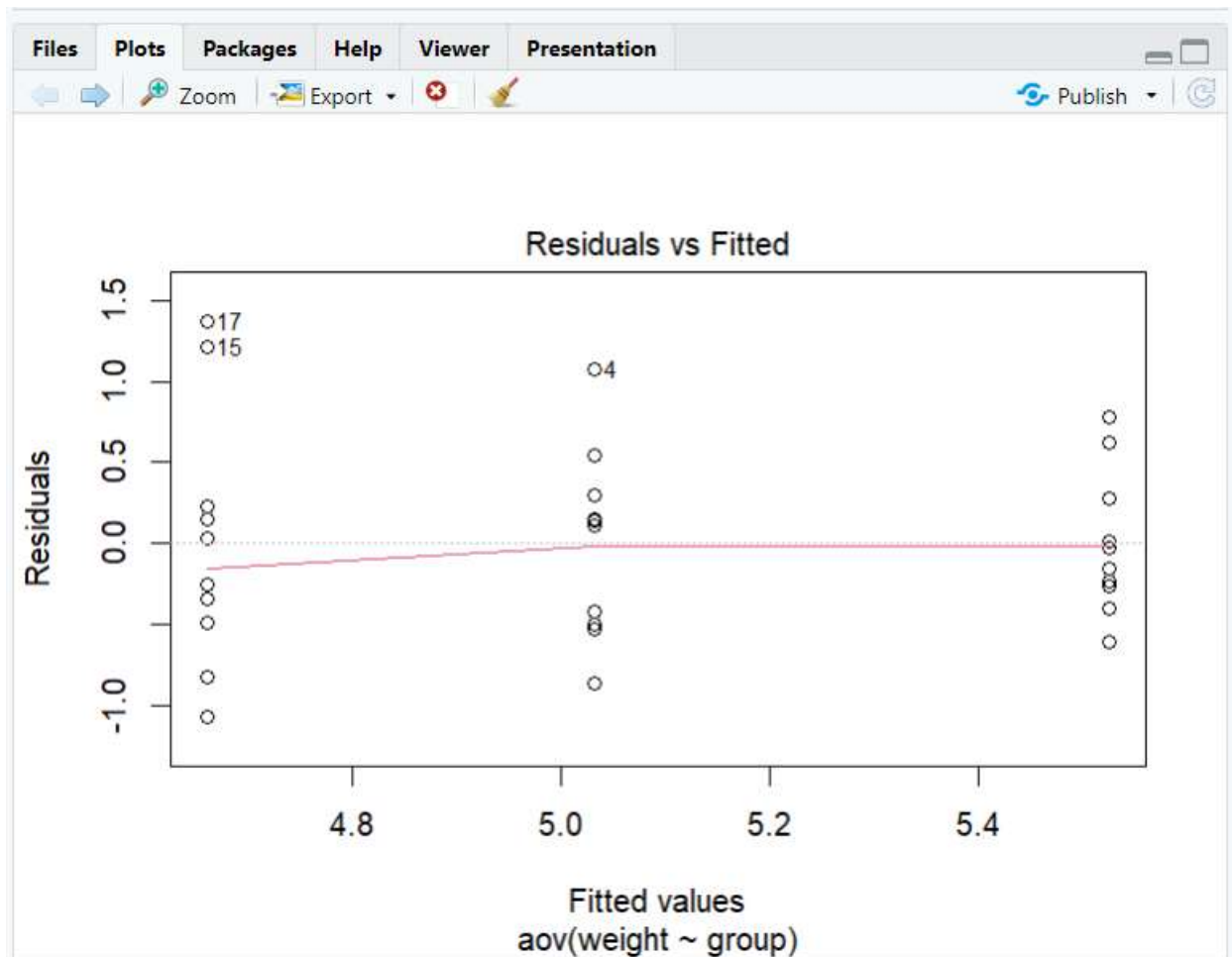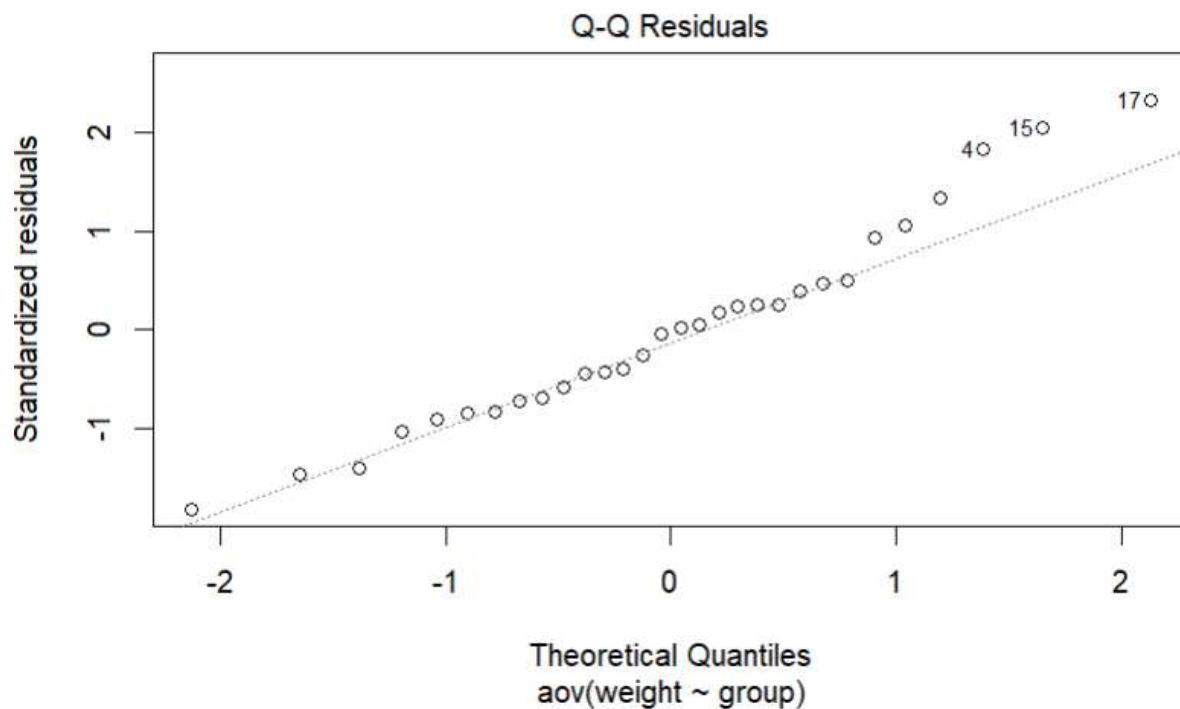
```
> fit<- kmeans(df,3)
>
> fit$size
[1] 47 53 50
> fit$withinss
[1] 47.45019 44.08754 47.35062
> fit$tot.withinss  # Within Cluster Sum of Squares (WCSS)
[1] 138.8884
```

BCSE352E–Essentials of Data Analytics – Lab [Winter Semester 2023–24]
Name of the Student: **Atharv Patel**
Register Number: **21BLC1560**

```
> group_by(data,type) %>% summarise(count=n(),mean=mean(sale.c
ount,na.rm=TRUE))
# A tibble: 3 × 3
  type  count  mean
  <chr> <int> <dbl>
1 Can-A     5    50
2 Can-B     5    40
3 Can-C     5    36
>
```

```
>
> group_by(data,group) %>% summarise(count=n(),mean=mean(weigl
t,na.rm=TRUE))
# A tibble: 3 × 3
  group count  mean
  <fct> <int> <dbl>
1 ctrl     10  5.03
2 trt1     10  4.66
3 trt2     10  5.53
>
```

```
> result<-aov(weight~group,data=data)
> summary(result)
            Df Sum Sq Mean Sq F value Pr(>F)
group        2  3.766  1.8832   4.846 0.0159 *
Residuals   27 10.492  0.3886
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

BCSE352E–Essentials of Data Analytics – Lab [Winter Semester 2023–24]
Name of the Student: **Atharv Patel**
Register Number: **21BLC1560**

```
> TukeyHSD(result)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = weight ~ group, data = data)

$group
            diff        lwr       upr      p adj
trt1-ctrl -0.371 -1.0622161 0.3202161 0.3908711
trt2-ctrl  0.494 -0.1972161 1.1852161 0.1979960
trt2-trt1  0.865  0.1737839 1.5562161 0.0120064
```

```
> kruskal.test(weight~group,data=data)

        Kruskal-Wallis rank sum test

data:  weight by group
Kruskal-Wallis chi-squared = 7.9882, df = 2, p-value
= 0.01842

>
```

```
> result<-aov(sale.count~type,data=data)
> summary(result)
            Df Sum Sq Mean Sq F value Pr(>F)
type         2    520   260.0   0.661  0.534
Residuals   12   4720   393.3
>
```

BCSE352E–Essentials of Data Analytics – Lab [Winter Semester 2023–24]
Name of the Student: **Atharv Patel**
Register Number: **21BLC1560**

	weight	group			sale.count	type
1	4.17	ctrl		1	40	Can-A
2	5.58	ctrl		2	60	Can-A
3	5.18	ctrl		3	70	Can-A
4	6.11	ctrl		4	30	Can-A
5	4.50	ctrl		5	50	Can-A
6	4.61	ctrl		6	30	Can-B
7	5.17	ctrl		7	30	Can-B
8	4.53	ctrl		8	10	Can-B
9	5.33	ctrl		9	70	Can-B
10	5.14	ctrl		10	60	Can-B
11	4.81	trt1		11	50	Can-C
12	4.17	trt1		12	60	Can-C
13	4.41	trt1		13	30	Can-C
14	3.59	trt1		14	20	Can-C
15	5.87	trt1		15	20	Can-C