Perform One-Way Anova

Aniket Roy and Krutik Shah

01/10/2020

data1 <- read.csv('One-Way(Rp).csv')  
data1

## CBA CS BDA  
## 1 44 55 70  
## 2 54 63 75  
## 3 78 58 65  
## 4 46 45 50  
## 5 43 76 73  
## 6 50 78 85  
## 7 45 48 75  
## 8 76 76 65  
## 9 58 53 73  
## 10 70 76 80

## Performing One-Way Anova

One-way ANOVA (Analysis of Variance) is used to examine the differences between means of more than two independent samples. It is used when we have a categorical independent variable (treatment) (with more than two categories) and a normally distributed interval or ratio dependent variable.

H0:There is no difference between means of all three groups.

Ha:Difference among maens is significant.

## To Check Normality of data

result <- shapiro.test(data1$CBA)  
result

##   
## Shapiro-Wilk normality test  
##   
## data: data1$CBA  
## W = 0.8535, p-value = 0.06392

result <- shapiro.test(data1$CS)  
result

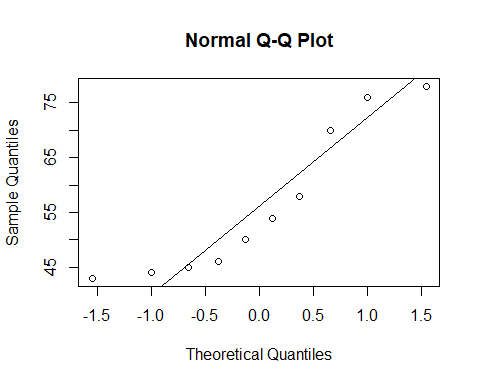
##   
## Shapiro-Wilk normality test  
##   
## data: data1$CS  
## W = 0.87468, p-value = 0.1133

result <- shapiro.test(data1$BDA)  
result

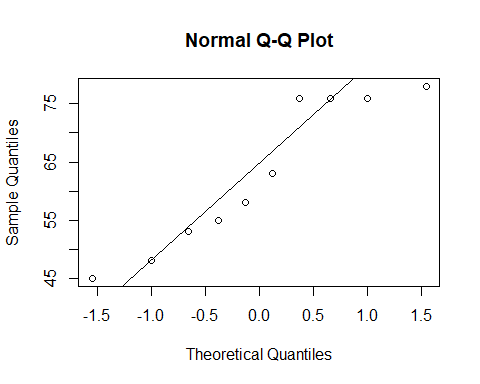
##   
## Shapiro-Wilk normality test  
##   
## data: data1$BDA  
## W = 0.92825, p-value = 0.4309

**As we can see that p-value of CBA,CS and BDA is as follows 0.06392,0.1133 and 0.4309.From this value we can say that all values are more than 0.05 so here We accept NULL Hypothesis.**

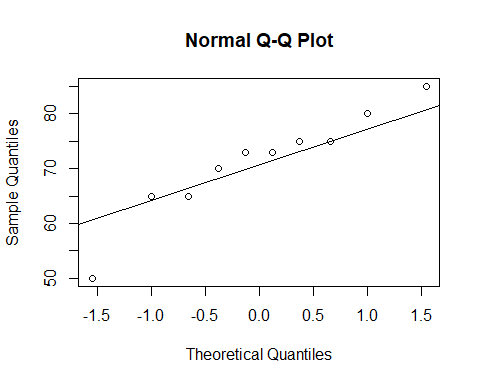
qqnorm(data1$CBA)  
qqline(data1$CBA, distribution=qnorm)



qqnorm(data1$CS)  
qqline(data1$CS, distribution=qnorm)



qqnorm(data1$BDA)  
qqline(data1$BDA, distribution=qnorm)



## Testing Homogeneity of Variances

**Bartlett’s test is used to test if k samples have equal variances. Equal variances across samples is called homogeneity of variances.Before applying ANOVA we have to check the assumption about the homogeneity of variances.**

result <- data.frame(data1$CBA,data1$CS,data1$BDA)  
res.var <- bartlett.test(list(data1$CBA,data1$CS,data1$BDA))  
res.var

##   
## Bartlett test of homogeneity of variances  
##   
## data: list(data1$CBA, data1$CS, data1$BDA)  
## Bartlett's K-squared = 1.082, df = 2, p-value = 0.5822

As p value is greater than 0.05 .We fail to reject the null hypothesis H0 at 5% level of significance. We conclude that the variances are equal across samples.

Homogeneity of variance is an assumption underlying both t tests and F tests (analyses of variance, ANOVAs) in which the population variances (i.e., the distribution, or “spread,” of scores around the mean) of two or more samples are considered equal.

## One Way Anova

st\_result <- stack(result)  
res <- oneway.test(values~ind, data = st\_result)  
res

##   
## One-way analysis of means (not assuming equal variances)  
##   
## data: values and ind  
## F = 4.0126, num df = 2.000, denom df = 17.547, p-value = 0.03673

Since p-value which 0.03673 is very less than 0.05 we Accept the Null Hypothesis.

## Pair-wise Comparison

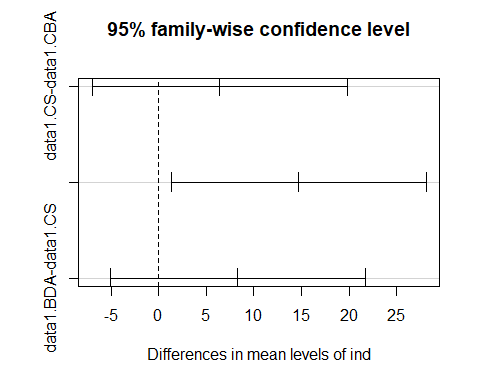
res.anova <- aov(values~ind, data = st\_result)  
summary(res.anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## ind 2 1086 543.2 3.707 0.0378 \*  
## Residuals 27 3957 146.6   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

TK <- TukeyHSD(res.anova)  
TK

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = values ~ ind, data = st\_result)  
##   
## $ind  
## diff lwr upr p adj  
## data1.CS-data1.CBA 6.4 -7.023333 19.82333 0.4737245  
## data1.BDA-data1.CBA 14.7 1.276667 28.12333 0.0297128  
## data1.BDA-data1.CS 8.3 -5.123333 21.72333 0.2917773

plot(TK)



## Linear Model

model1 <- lm(values ~ ind,data=st\_result)  
anova(model1)

## Analysis of Variance Table  
##   
## Response: values  
## Df Sum Sq Mean Sq F value Pr(>F)   
## ind 2 1086.5 543.23 3.7068 0.03781 \*  
## Residuals 27 3956.9 146.55   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

tk1 <- TukeyHSD(aov(model1))  
tk1

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = model1)  
##   
## $ind  
## diff lwr upr p adj  
## data1.CS-data1.CBA 6.4 -7.023333 19.82333 0.4737245  
## data1.BDA-data1.CBA 14.7 1.276667 28.12333 0.0297128  
## data1.BDA-data1.CS 8.3 -5.123333 21.72333 0.2917773

plot(tk1)

