ANOVA

Amira Ibrahim

October 16, 2019

Attach Data

```
Diet_Data <- read.table(file.choose() , header = T , sep = "\t")
attach(Diet_Data)
```

check normality

```
library(moments)
skewness(WeightLoss)

## [1] -0.4944242
```

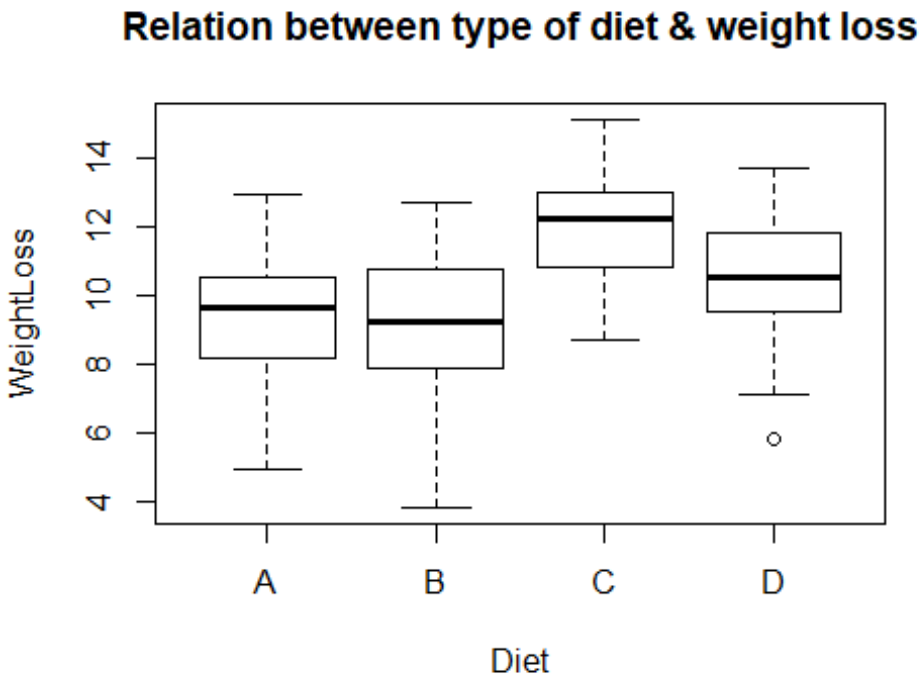
accepted level from -1 to +1

```
kurtosis(WeightLoss)

## [1] 2.884693
```

accepted level from -2 to +2 may to +3

```
boxplot(WeightLoss~Diet , main = "Relation between type of diet & weight
loss")
```



visually it is normally distributed

```
shapiro.test(WeightLoss)

##
##  Shapiro-Wilk normality test
##
## data:  WeightLoss
## W = 0.96993, p-value = 0.1447
```

p-value > 0.05 , Fail to reject H0 , Data is normally distributed

Analysis of relation between type of diet and weight lose :

comparing one categorical variable(Diet) contain more than 2 levels and one numerical variable :

using one way ANOVA

H0 : Mean weight lose is the same for all types of diet

```
ANOVA1 <- aov(formula = WeightLoss~Diet , data=Diet_Data)
summary(ANOVA1)
```

```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## Diet         3  97.33   32.44   6.118 0.00113 **
## Residuals   56 296.99    5.30
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p-value < 0.05 , reject H0

I need to Know which type of Diet is significant :

```
TukeyHSD(ANOVA1)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = WeightLoss ~ Diet, data = Diet_Data)
##
## $Diet
##           diff        lwr       upr     p adj
## B-A -0.2733333 -2.4999391 1.9532725 0.9880087
## C-A  2.9333333  0.7067275 5.1599391 0.0051336
## D-A  1.3600000 -0.8666058 3.5866058 0.3773706
## C-B  3.2066667  0.9800609 5.4332725 0.0019015
## D-B  1.6333333 -0.5932725 3.8599391 0.2224287
## D-C -1.5733333 -3.7999391 0.6532725 0.2521236
```

```
plot(TukeyHSD(ANOVA1), las=1)
```

Diffrence in mean between C-A , C-B is significant Diffrence in mean C-B is more significant than C-A