Untitled

Precious Ogunbekun

24/07/2021

ANOVA * Response variable (CI) * Two categorical variable; gender with two labels (male and female) & Bone with 3 labels (OO, OA AND ND)

```
Calcium <- read.table(file.choose(), header=TRUE)
names(Calcium)
```

```
## [1] "CI"     "Bone"   "Gender"
```

```
str(Calcium)
```

```
## 'data.frame':    24 obs. of  3 variables:
##  $ CI    : int  1200 1000 980 900 850 800 950 900 1000 900 ...
##  $ Bone  : chr  "ND" "ND" "ND" "ND" ...
##  $ Gender: chr  "M" "M" "M" "F" ...
```

```
attach(Calcium)
```

DESCRIPTIVE STATISTICS

- The mean it is obvious that the daily intake of calcium is most effective in adult male with normal bone density and the least effective is in adult female with osteoporosis
- I can assume adult female with normal bone density has a low variation from the mean, while adult female with osteoporosis has a large variation from the mean which could be due to select observations with high value as a result of other effect aside calcium intake

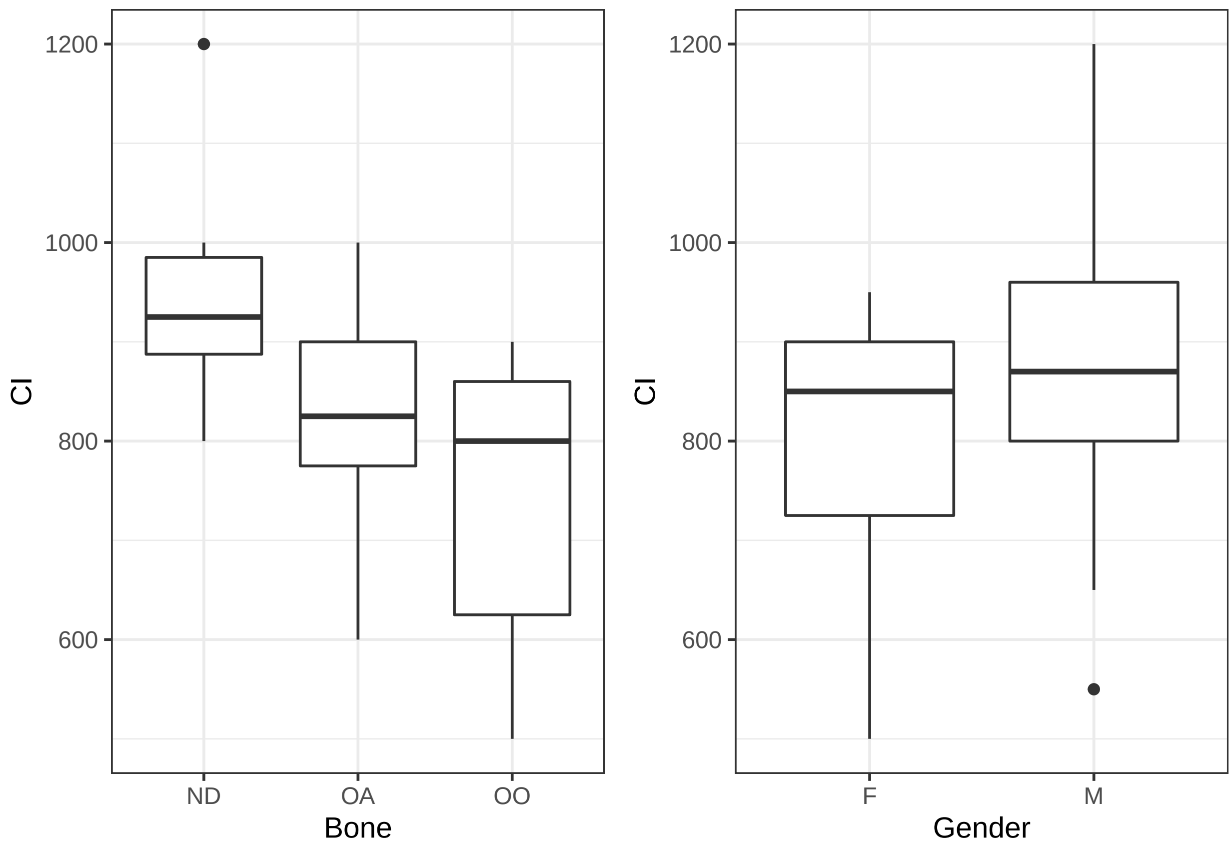
```
with(data = Calcium, expr = tapply(CI, list(Bone, Gender), mean))
```

```
##           F    M
## ND 880.0000 1060
## OA 733.3333  870
## OO 675.0000  765
```

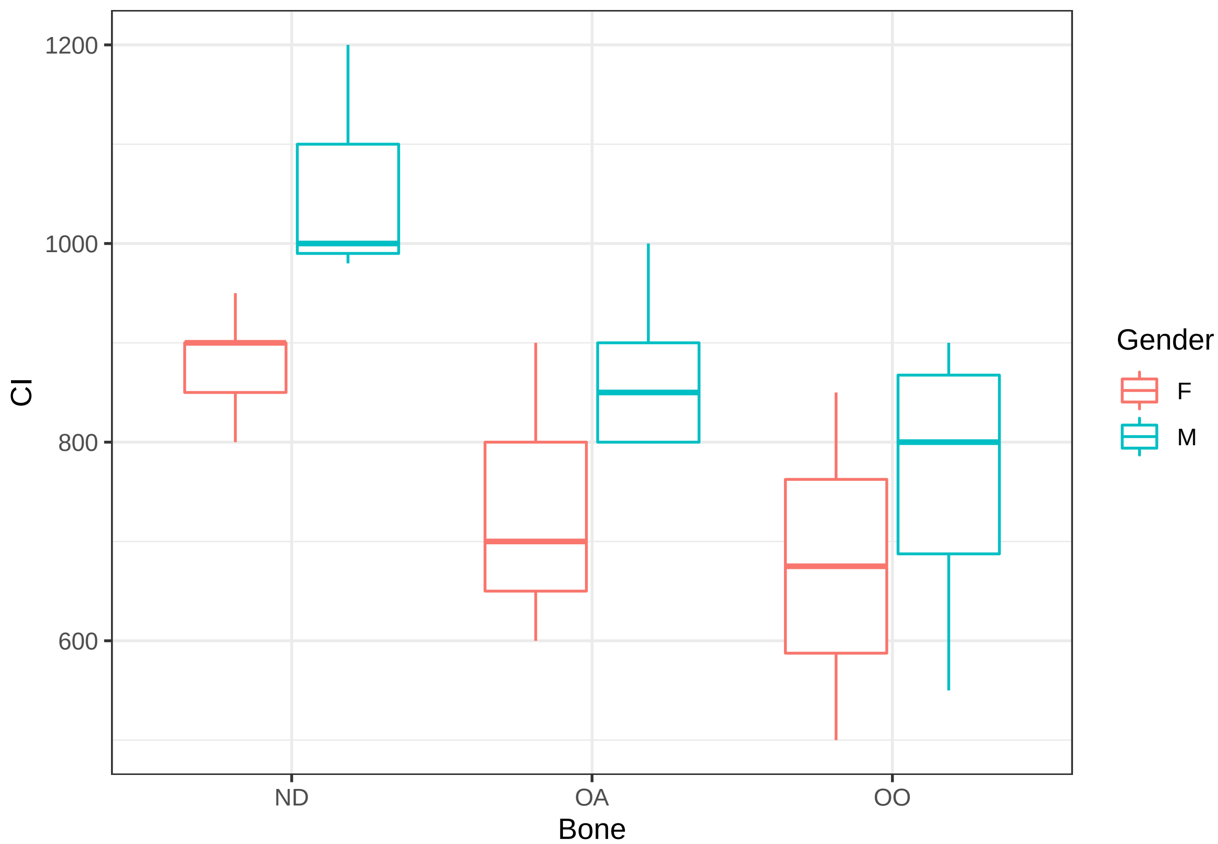
```
with(data = Calcium, expr = tapply(CI, list(Bone, Gender), sd))
```

```
##            F        M
## ND  57.00877 121.6553
## OA 152.75252  83.6660
## OO 247.48737 138.3835
```

```r
library(ggplot2)
library(gridExtra)
p1 <- ggplot(data = Calcium, mapping = aes(x = Bone, y = CI)) + geom_boxplot() +
  theme_bw()
p2 <- ggplot(data = Calcium, mapping = aes(x = Gender, y = CI)) + geom_boxplot() +
  theme_bw()
p3 <- ggplot(data = Calcium, mapping = aes(x = Bone, y = CI, colour = Gender)) +
  geom_boxplot() + theme_bw()

grid.arrange(p1, p2, ncol = 2)
```

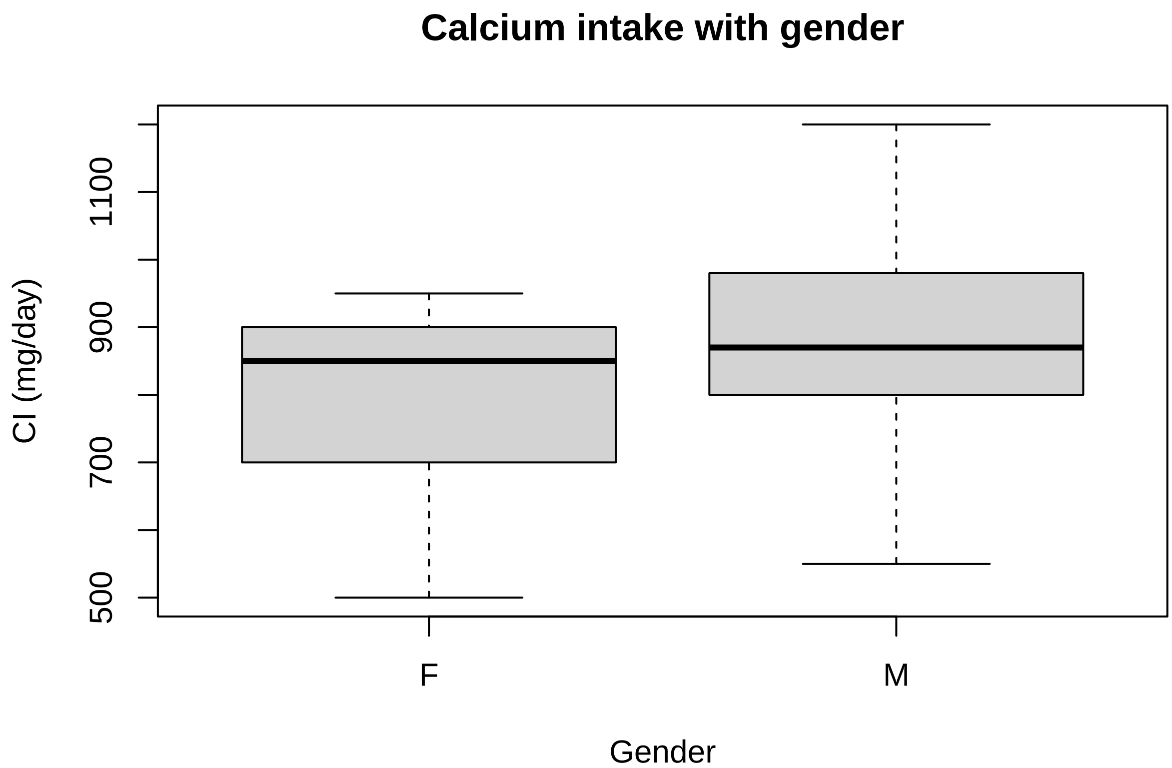



```r
p3
```

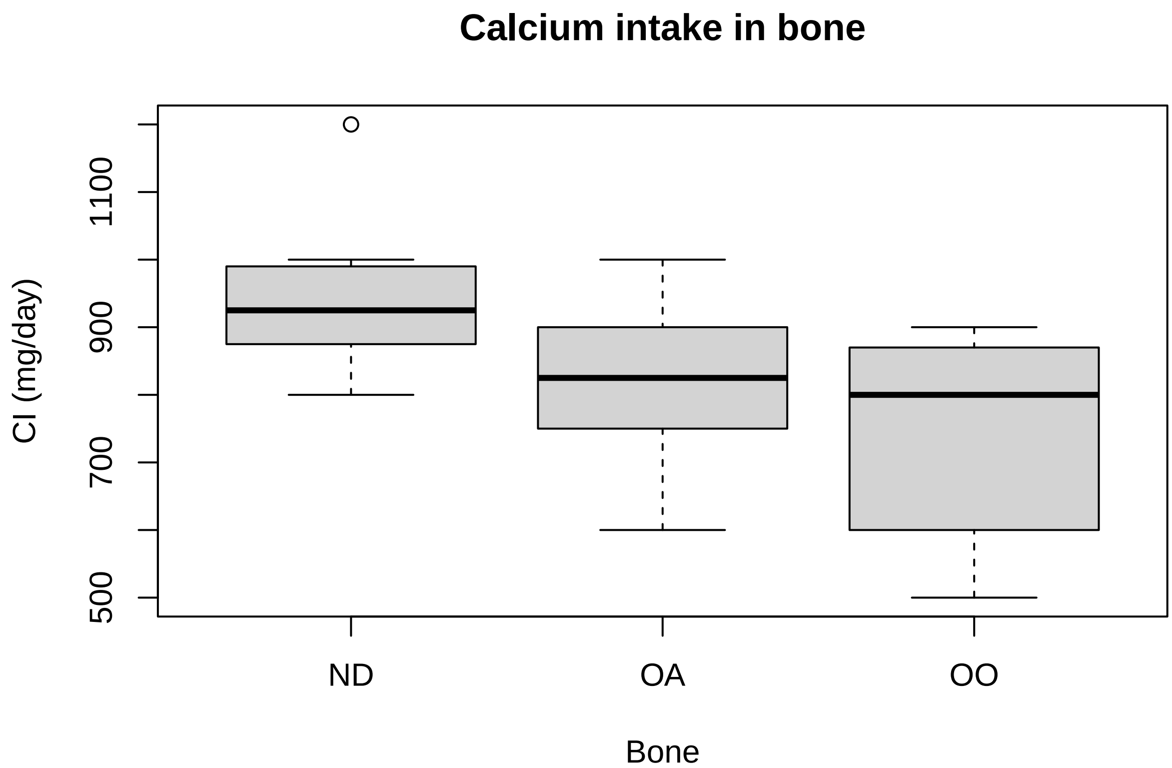


Calcium intake with gender - calcium intake in male seem higher than in female Calcium intake in bone - calcium intake in male seem higher than in female

```
boxplot(CI~Gender, ylab = "CI (mg/day)", xlab = "Gender", main = "Calcium intake with gender")
```




```
boxplot(CI~Bone, ylab = "CI (mg/day)", xlab = "Bone", main = "Calcium intake in bone")
```



Analysis of Variance There is significance different with calcium intake and bone also with gender because the P-value> 0.05. But no interaction between the two variable (gender and bone) so we simplify our model from multiplicative to a more pasimonic model which is the additive effect.

```
CI_Bone_Gender<-lm(CI~Bone*Gender)          ### (*) multiplicative model with interaction
anova(CI_Bone_Gender)
```

```
## Analysis of Variance Table
##
## Response: CI
##             Df Sum Sq Mean Sq F value  Pr(>F)
## Bone         2 171775   85888  5.6368 0.01257 *
## Gender       1 101157  101157  6.6389 0.01901 *
## Bone:Gender  2   6763    3382  0.2219 0.80312
## Residuals   18 274267   15237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
CI_Bone_Gender<-lm(CI~Bone+Gender)     ### (+) Additive model without interaction
anova(CI_Bone_Gender)
```

```
## Analysis of Variance Table
##
## Response: CI
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Bone       2 171775   85888  6.1123 0.00848 **
```

```
## Gender     1 101157  101157  7.1990 0.01430 *
## Residuals 20 281030   14052
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(CI_Bone_Gender)
```

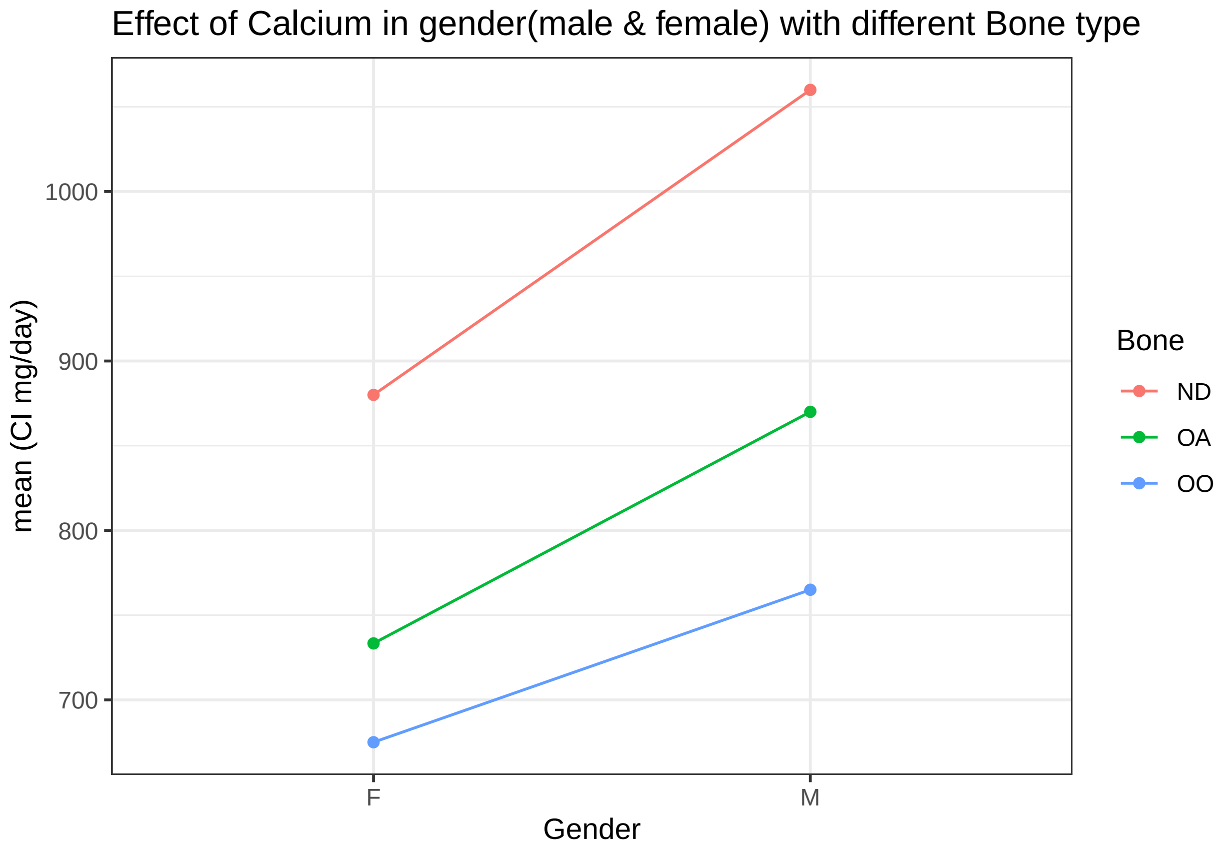
```
##
## Call:
## lm(formula = CI ~ Bone + Gender)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -227.202  -70.804   -8.125   69.115  211.607
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   895.45      46.18  19.389 1.95e-14 ***
## BoneOA       -163.45      60.66  -2.694 0.013949 *
## BoneOO       -257.05      62.36  -4.122 0.000529 ***
## GenderM       138.81      51.73   2.683 0.014296 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 118.5 on 20 degrees of freedom
## Multiple R-squared:  0.4927, Adjusted R-squared:  0.4166
## F-statistic: 6.475 on 3 and 20 DF,  p-value: 0.003064
```

INTERACTION PLOT The calcium intake effect is highest in adult with normal bone density and in general the effect of calcium intake is highest in male than in female for the three different bone. The pattern are the same

The pattern of the effect of calcium intake is the same for male and female which is that as the calcium intake increases the effect deceases from normal bone to low bone density(OA) and finally to ostereoporosis(OO).

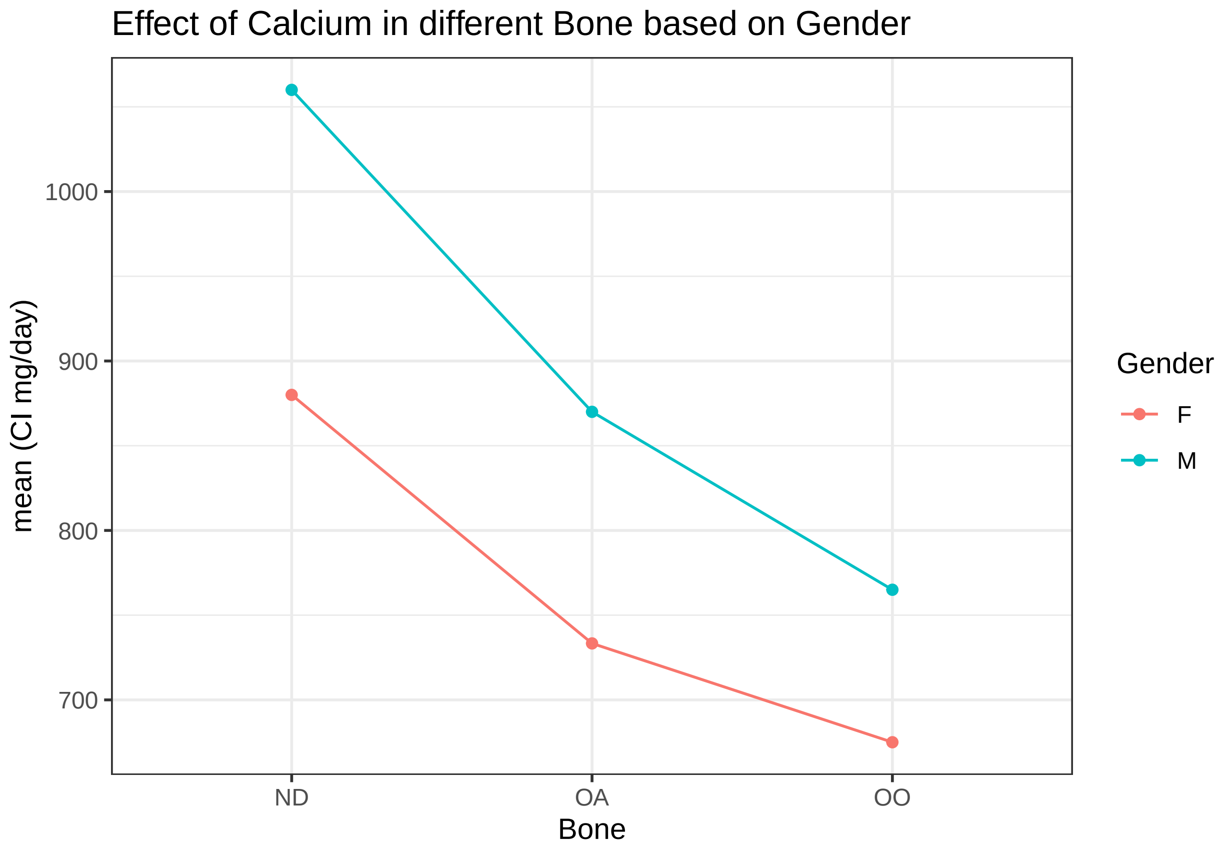
```
ggplot(data = Calcium, aes(x = Gender, y = CI, colour =Bone, group = Bone)) + ggtitle("Effect of Calciu
  stat_summary(fun.y = mean, geom = "point") + stat_summary(fun.y = mean,
                                                          geom = "line") + labs(y = "mean (CI mg/day)
```

```
## Warning: `fun.y` is deprecated. Use `fun` instead.

## Warning: `fun.y` is deprecated. Use `fun` instead.
```



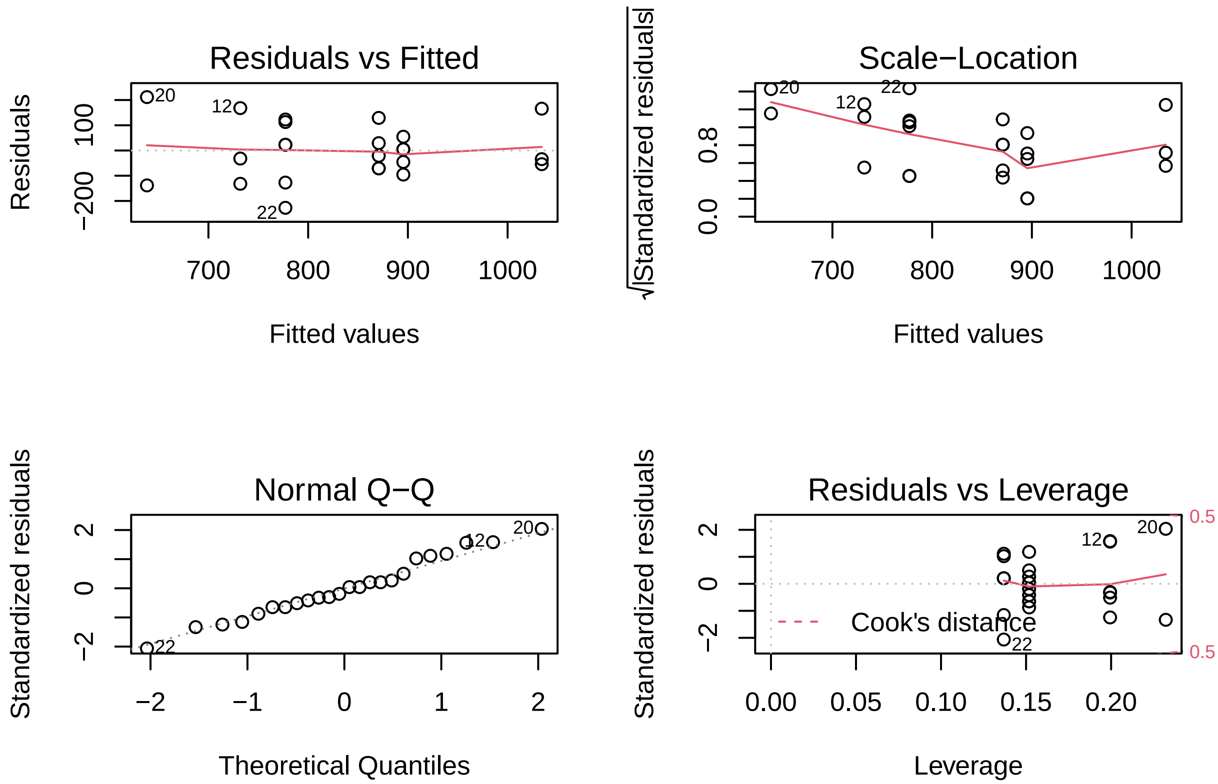

```
ggplot(data = Calcium, aes(x = Bone, y = CI, colour = Gender, group = Gender)) + ggtitle("Effect of Cal
  stat_summary(fun.y = mean, geom = "point") + stat_summary(fun.y = mean,
                                                            geom = "line") + labs(y = "mean (CI mg/day)
```

```
## Warning: `fun.y` is deprecated. Use `fun` instead.

## Warning: `fun.y` is deprecated. Use `fun` instead.
```



ASSUMPTIONS OF MODEL The normality and heteroscedasticity looks acceptable from the plot

```
par(mfcol=c(2,2))
plot(CI_Bone_Gender)
```



Shapiro test (test for normality in residuals "the variation among sampling units within each sample): the normality is greater than 0.05

```
#leveneTest(CI_Bone_Gender) ###test heteroscedasticity
shapiro.test(residuals(CI_Bone_Gender)) ###test normality in residuals "the variation amongst sampling
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(CI_Bone_Gender)
## W = 0.97698, p-value = 0.8343
```