ToothGrowth Analysis

- Load the ToothGrowth data and perform some basic exploratory data analyses

```
suppressMessages(library(datasets))
suppressMessages(library(ggplot2))
```

```
## Warning: package 'ggplot2' was built under R version 3.2.2
```

```
suppressMessages(library(dplyr))
```

```
## Warning: package 'dplyr' was built under R version 3.2.2
```

```
data(ToothGrowth)
str(ToothGrowth)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
ToothGrowth
```

```
##     len supp dose
## 1   4.2   VC  0.5
## 2  11.5   VC  0.5
## 3   7.3   VC  0.5
## 4   5.8   VC  0.5
## 5   6.4   VC  0.5
## 6  10.0   VC  0.5
## 7  11.2   VC  0.5
## 8  11.2   VC  0.5
## 9   5.2   VC  0.5
## 10  7.0   VC  0.5
## 11 16.5   VC  1.0
## 12 16.5   VC  1.0
## 13 15.2   VC  1.0
## 14 17.3   VC  1.0
## 15 22.5   VC  1.0
## 16 17.3   VC  1.0
## 17 13.6   VC  1.0
## 18 14.5   VC  1.0
## 19 18.8   VC  1.0
## 20 15.5   VC  1.0
## 21 23.6   VC  2.0
## 22 18.5   VC  2.0
## 23 33.9   VC  2.0
## 24 25.5   VC  2.0
## 25 26.4   VC  2.0
```

```
## 26 32.5   VC  2.0
## 27 26.7   VC  2.0
## 28 21.5   VC  2.0
## 29 23.3   VC  2.0
## 30 29.5   VC  2.0
## 31 15.2   OJ  0.5
## 32 21.5   OJ  0.5
## 33 17.6   OJ  0.5
## 34  9.7   OJ  0.5
## 35 14.5   OJ  0.5
## 36 10.0   OJ  0.5
## 37  8.2   OJ  0.5
## 38  9.4   OJ  0.5
## 39 16.5   OJ  0.5
## 40  9.7   OJ  0.5
## 41 19.7   OJ  1.0
## 42 23.3   OJ  1.0
## 43 23.6   OJ  1.0
## 44 26.4   OJ  1.0
## 45 20.0   OJ  1.0
## 46 25.2   OJ  1.0
## 47 25.8   OJ  1.0
## 48 21.2   OJ  1.0
## 49 14.5   OJ  1.0
## 50 27.3   OJ  1.0
## 51 25.5   OJ  2.0
## 52 26.4   OJ  2.0
## 53 22.4   OJ  2.0
## 54 24.5   OJ  2.0
## 55 24.8   OJ  2.0
## 56 30.9   OJ  2.0
## 57 26.4   OJ  2.0
## 58 27.3   OJ  2.0
## 59 29.4   OJ  2.0
## 60 23.0   OJ  2.0
```

Length of ToothGrowth$len

```r
length(ToothGrowth$len)
```

```
## [1] 60
```

Mean of ToothGrowth Dataset

```r
aggregate(ToothGrowth$len,list(ToothGrowth$supp,ToothGrowth$dose),mean)
```

```
##   Group.1 Group.2     x
## 1      OJ     0.5 13.23
## 2      VC     0.5  7.98
## 3      OJ     1.0 22.70
## 4      VC     1.0 16.77
## 5      OJ     2.0 26.06
## 6      VC     2.0 26.14
```

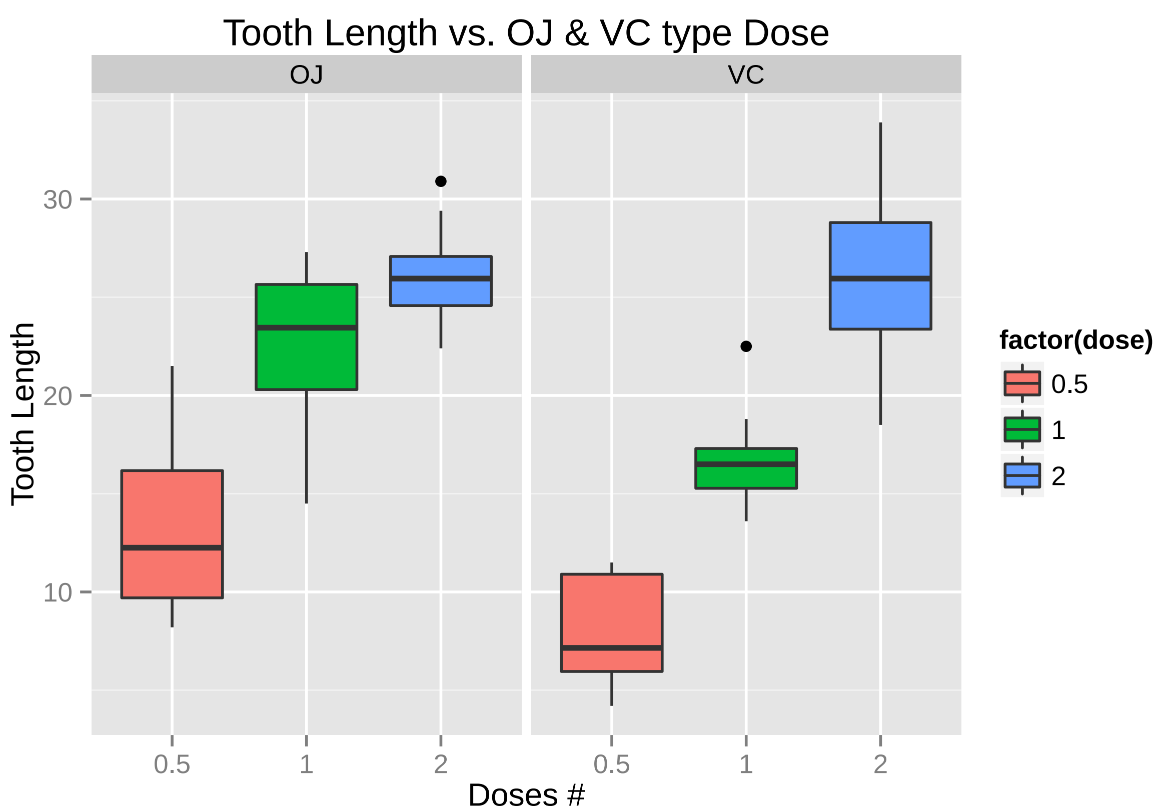
Standard Deviation of ToothGrowth Dataset

```
aggregate(ToothGrowth$len,list(ToothGrowth$supp,ToothGrowth$dose),sd)
```

```
##   Group.1 Group.2        x
## 1      OJ     0.5 4.459709
## 2      VC     0.5 2.746634
## 3      OJ     1.0 3.910953
## 4      VC     1.0 2.515309
## 5      OJ     2.0 2.655058
## 6      VC     2.0 4.797731
```

BoxPlot of the quantile

```
    ggplot(ToothGrowth, aes(x = factor(dose), y = len, fill = factor(dose)))+
                  geom_boxplot()+
                  facet_grid(.~supp)+
                  labs(title = "Tooth Length vs. OJ & VC type Dose",
                  x = "Doses #", y = "Tooth Length")
```



- Provide a basic summary of the data.

```r
summary(ToothGrowth)
```

```
##       len        supp         dose      
##  Min.   : 4.20   OJ:30   Min.   :0.500  
##  1st Qu.:13.07   VC:30   1st Qu.:0.500  
##  Median :19.25           Median :1.000  
##  Mean   :18.81           Mean   :1.167  
##  3rd Qu.:25.27           3rd Qu.:2.000  
##  Max.   :33.90           Max.   :2.000
```

- Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)

If we refer to the box plot , OJ seems better with 0.5 dose and 1 which effects on the teeth growth than VC. We can assume by making hypothesis that mean of OJ and VC will not more than 0

a) 0.5 Dose

With 0.95% confident rate , the boundary of from 1.719057 and 8.780943 contains the difference between the two population. Because of the boundary does not have 0 value , the possibility that the two population means are not equal.

```r
oj_dose95 <- ToothGrowth %>% filter(dose=="0.5" & supp=="OJ")
vc_dose95 <- ToothGrowth %>% filter(dose=="0.5" & supp=="VC")
t.test(oj_dose95$len,vc_dose95$len)
```

```
## 
##  Welch Two Sample t-test
## 
## data:  oj_dose95$len and vc_dose95$len
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean of x mean of y 
##     13.23      7.98
```

b) 1 Dose

With 0.95% confident rate , the boundary of from 2.802148 and 9.057852 contains the difference between the two population. Because of the boundary does not have 0 value , the possibility that the two population means are not equal.

```r
oj_dose95 <- ToothGrowth %>% filter(dose=="1" & supp=="OJ")
vc_dose95 <- ToothGrowth %>% filter(dose=="1" & supp=="VC")
t.test(oj_dose95$len,vc_dose95$len)
```

```
## 
##  Welch Two Sample t-test
## 
## data:  oj_dose95$len and vc_dose95$len
```

```
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean of x mean of y
##     22.70     16.77
```

c) 2 Dose

With 0.95% confident rate , the boundary of from -3.79807 and 3.63807 contains the difference between the two population. Because of the boundary does not have 0 value , the possibility that the two population means are equal.

```r
oj_dose95 <- ToothGrowth %>% filter(dose=="2" & supp=="OJ")
vc_dose95 <- ToothGrowth %>% filter(dose=="2" & supp=="VC")
t.test(oj_dose95$len,vc_dose95$len)
```

```
##
##  Welch Two Sample t-test
##
## data:  oj_dose95$len and vc_dose95$len
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean of x mean of y
##     26.06     26.14
```

- State your conclusions and the assumptions needed for your conclusions.

We can conclude that With 95% rate confident , OJ dose with 0.5 and 1 will have longer tooth length than VC with the same 0.5 and 1 dose. We also can assume with the dose of 2 , there is no significant different effects between for both OV and VC.