# Tooth Growth Data Analysis

*Ng Ming Chun*

*Saturday, June 13, 2015*

## Tooth Growth Data Analysis

### 1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(datasets)
data (ToothGrowth) # load the data
head (ToothGrowth) # show the first 10 rows of data
```

```
##    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
```

```
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 ...
```

### 2. Provide a basic summary of the data.

```
summary (ToothGrowth) # print the summary of the data
```

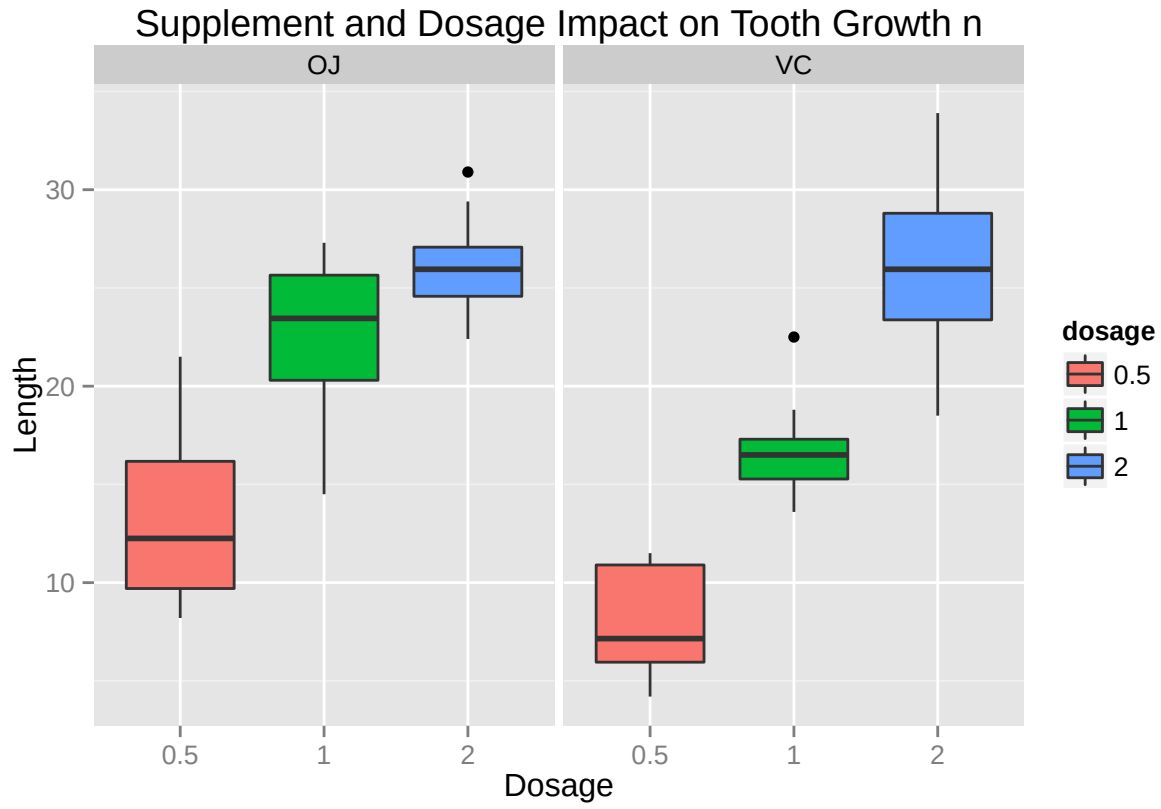
```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

It appears that the dataset contains the measurement of toothgrowth (i.e. length) with respect to 2 types of supplements of varing dosages. It is best to present the data in 2 panels for comparison of the 2 supplements. The y axis will represent the length, while the axis will respect the dosage of the respective supplement.

```
library (ggplot2)

dosage<-factor(ToothGrowth$dose)
plot <- ggplot(data=ToothGrowth, aes(x=dosage, y=len, fill=dosage))

plot + geom_boxplot(notch=F) + facet_grid(.~supp) + labs(title = "Supplement and Dosage Impact on Tooth
```




### 3. 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)

As there are different dosages, it is necessary to conduct 3 comparisons. I also assume non-pairwise comparisons.

```
for (d in levels(factor(ToothGrowth$dose))) {
  subset_d<- ToothGrowth[ToothGrowth$dose == d, ]
  print (t.test(data=subset_d, len ~ supp, paired=FALSE, var.equal=FALSE))
}
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
```

```
## 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 in group OJ mean in group VC
##            13.23             7.98
##
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##            22.70            16.77
##
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.0461, 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 in group OJ mean in group VC
##            26.06            26.14
```

Based on the t-test, the following observations were made: * At 0.5 dosage, supplement OJ is better than supplement VC. * At 1.0 dosage, supplement OJ is better than supplement VC. * At 2.0 dosage, both supplements OJ and VC are almost equally effective.

## 4. State your conclusions and the assumptions needed for your conclusions.

Both supplements are effective on promoting tooth growth, and increasing dosage always leads to faster growth. Below 2.0 dosage, supplement OJ is better than VC.

Assumptions *The sample data is not paired.* The sample variances for the supplements are not equal.