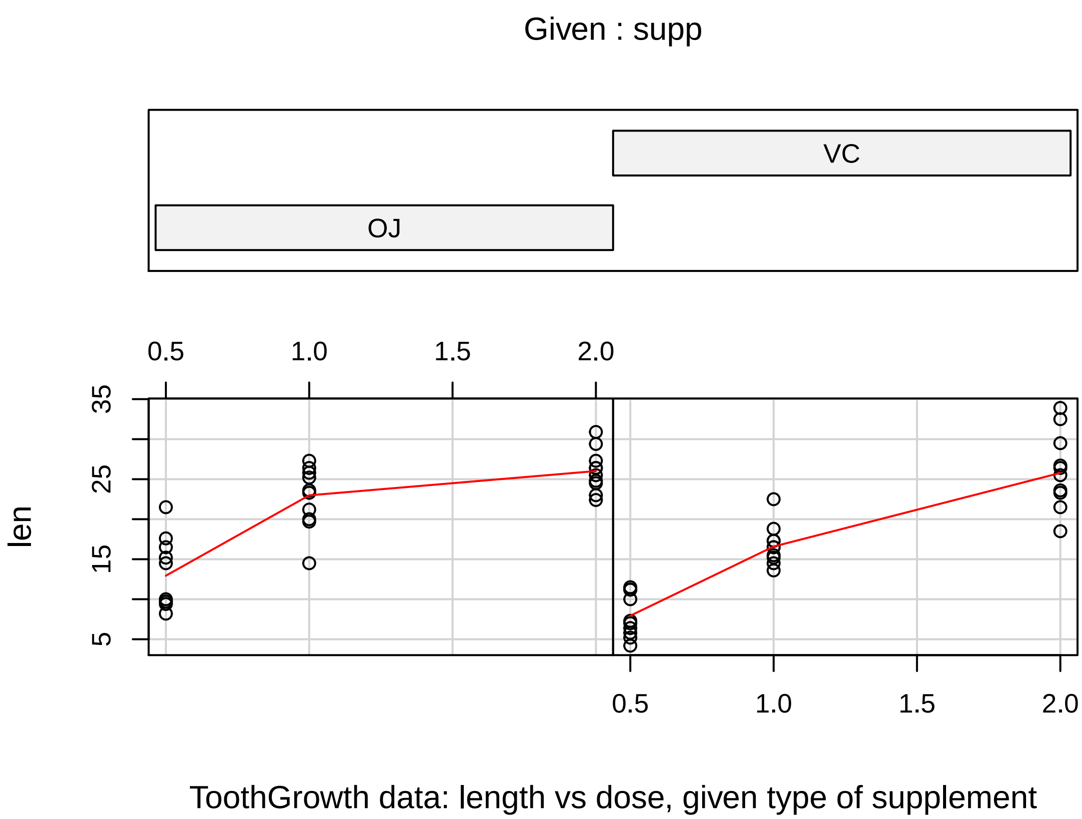
# Tooth Growth Analysis

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

```
library(datasets)
attach(ToothGrowth)
require(graphics)
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 ...
```

```
 coplot(len ~ dose | supp, data = ToothGrowth, panel = panel.smooth,
xlab = "ToothGrowth data: length vs dose, given type of supplement")
```




### Basic summary of the data

```
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
```

From the graph above it appears there is a clear effect on tooth growth of dosage. Less clear is the effect of supplement method. Now we will look at data aggregated by factors and the mean and standard deviation of each:

```r
dataGroup <- as.data.frame(split(len,list(supp,dose)))
dataGroup
```

```
##    OJ.0.5 VC.0.5 OJ.1 VC.1 OJ.2 VC.2
## 1    15.2    4.2 19.7 16.5 25.5 23.6
## 2    21.5   11.5 23.3 16.5 26.4 18.5
## 3    17.6    7.3 23.6 15.2 22.4 33.9
## 4     9.7    5.8 26.4 17.3 24.5 25.5
## 5    14.5    6.4 20.0 22.5 24.8 26.4
## 6    10.0   10.0 25.2 17.3 30.9 32.5
## 7     8.2   11.2 25.8 13.6 26.4 26.7
## 8     9.4   11.2 21.2 14.5 27.3 21.5
## 9    16.5    5.2 14.5 18.8 29.4 23.3
## 10    9.7    7.0 27.3 15.5 23.0 29.5
```

```r
sapply(dataGroup,mean)
```

```
## OJ.0.5 VC.0.5   OJ.1   VC.1   OJ.2   VC.2
##  13.23   7.98  22.70  16.77  26.06  26.14
```

```r
sapply(dataGroup,sd)
```

```
##   OJ.0.5   VC.0.5     OJ.1     VC.1     OJ.2     VC.2
## 4.459709 2.746634 3.910953 2.515309 2.655058 4.797731
```

It appears at lower doses, OJ has a greater effect than VC, but that effect diminshes as dosage is raised.

### Confidence Intervals and Tests

We will first test for effect of supplement method, disregarding the dosage. First we consider unequal variances:

```r
t.test(len ~ supp, paired = F, var.equal = F)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

and next equal variances:

```r
t.test(len ~ supp, paired = F, var.equal = T)
```

```
##
```

```
##  Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1670064  7.5670064
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

There is almost no difference between the two tests. Both show results that are not significant at the 5% level.

Next, we can consider whether the dose has an effect on tooth length. First, we test dosages of 0.5 and 1:

```
d1 <- ToothGrowth[dose==0.5 | dose==1,]
t.test(len ~ dose, paired = F, var.equal = F,data = d1)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

and now between 1 and 2:

```
d2 <- ToothGrowth[dose==1 | dose==2,]
t.test(len ~ dose, paired = F, var.equal = F,data = d2)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

and now between 0.5 and 2:

```
d1 <- ToothGrowth[dose==0.5 | dose==2,]
t.test(len ~ dose, paired = F, var.equal = F,data = d1)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean in group 0.5   mean in group 2
##            10.605            26.100
```

Clearly all three tests are significant and have very low p-values, indicating we should reject that hypothesis that there is no difference in tooth length between the dosage levels.

### Conclusions and Assumptions

From the analysis above, we conclude that supplement method does not significantly effect tooth length, so OJ and VC provide similar effectiveness. However dosage level does appear to significantly effect tooth length, and we conclude that higher dosages lead to greater tooth length.

We assumed that the population variances of the quantities being tested were not equal throughout. This is a mild assumption that seemed justified by the EDA. We also assume all of the subjects are independent from each other, represent the complete population of guinea pigs and were randomly selected from that population.