

# ToothGrowth Analysis

*Gene Kaufman*

*February 9, 2016*

## Overview

The ToothGrowth dataset was collected from a study researching the Effect of Vitamin C on Tooth Growth in Guinea Pigs, and published in 1947. The study looked at the effect of 3 dosage strengths (**dose**) administered one of two ways (**supp**). Each of the 6 combinations of dose x supp were administered to 10 Guinea Pigs. This paper will document some basic data analysis and make some conclusions on that data.

## Loading data, basic exploratory data analysis

Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(datasets)
data(ToothGrowth)

# set dose as a factor
ToothGrowth$dose<-as.factor(ToothGrowth$dose)

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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

## Basic data summary

Provide a basic summary of the data.

```
summary(ToothGrowth)

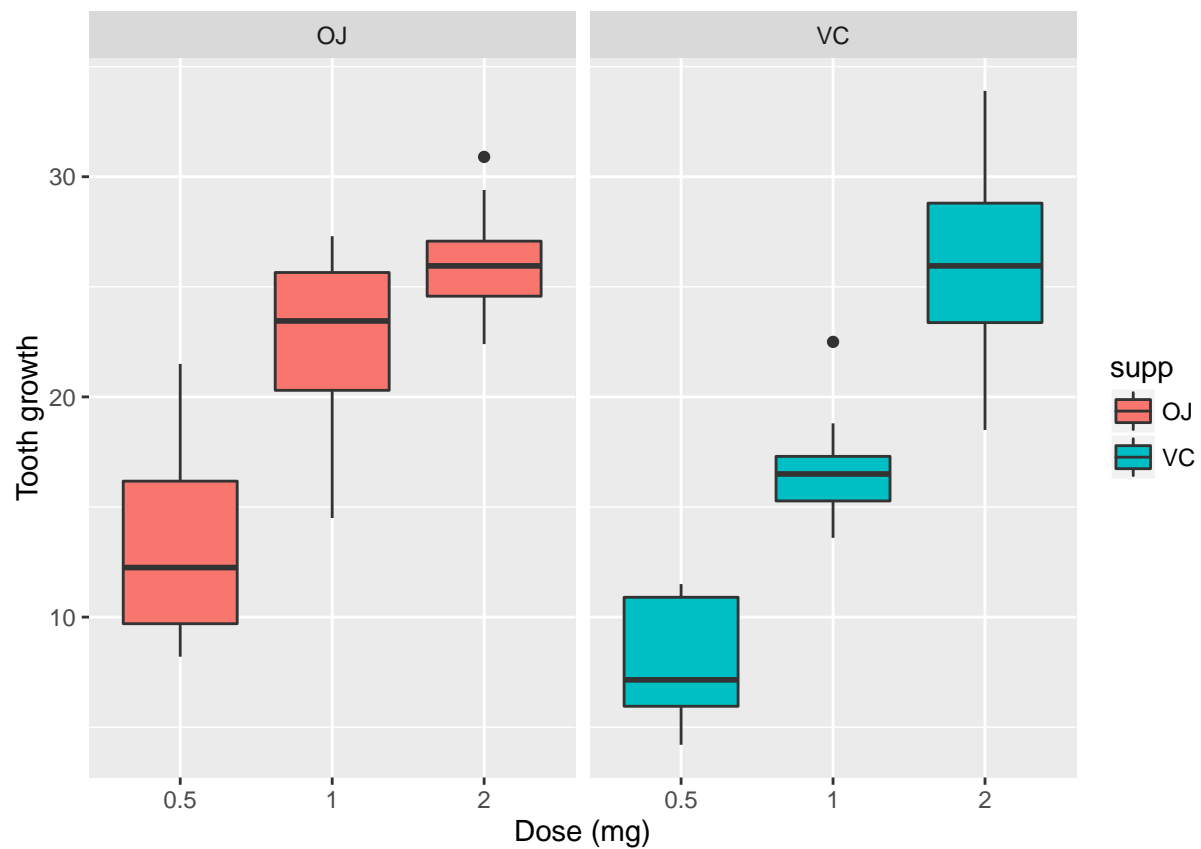
##      len      supp      dose
## Min.   : 4.20    OJ:30    0.5:20
## 1st Qu.:13.07    VC:30     1 :20
## Median :19.25           2 :20
## Mean   :18.81
## 3rd Qu.:25.27
## Max.   :33.90

table(ToothGrowth$supp, ToothGrowth$dose)
```

```
##
##      0.5  1  2
##    OJ  10 10 10
##    VC  10 10 10
```

Boxplot of dose x length per supp

```
library(ggplot2)
g<-ggplot(data=ToothGrowth,
          aes(x=dose,y=len,fill=supp)) +
  ylab("Tooth growth") +
  xlab("Dose (mg)") +
  geom_boxplot() +
  facet_grid(.~supp)
print(g)
```



## Compare Tooth growth by Supplement Type and Dosage

### Comparing Supplement Type

Null hypothesis: The supplement type will have no effect on tooth growth

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

```
##
## 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
```

The confidence Interval of -0.1710156 , 7.5710156 contains 0 and the p-value is greater than 0.5, so we cannot reject the null hypothesis.

## Comparing Dosage

For this comparison, we need to compare each of the three dosages against each other in pairs.

### Compare 0.5mg, 1.0mg

Null hypothesis: A dosage of 0.5mg will show the same growth as a dosage of 1.0mg

```
t.test(len ~ dose, paired = F, var.equal = F,
      data = subset(ToothGrowth, dose %in% c(0.5, 1.0)))
```

```
##
## 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
```

The confidence Interval of -11.983781 , -6.276219 does NOT contain 0 and the p-value is less than 0.5, so we can reject the null hypothesis.

### Compare 0.5mg, 2.0mg

Null hypothesis: A dosage of 0.5mg will show the same growth as a dosage of 2.0mg

```
t.test(len ~ dose, paired = F, var.equal = F,
      data = subset(ToothGrowth, dose %in% c(0.5, 2.0)))
```

```
##
## 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
```

The confidence Interval of -18.15617 , -12.83383 does NOT contain 0 and the p-value is less than 0.5, so we can reject the null hypothesis.

### Compare 1.0mg, 2.0mg

Null hypothesis: A dosage of 1.0mg will show the same growth as a dosage of 2.0mg

```
t.test(len ~ dose, paired = F, var.equal = F,
       data = subset(ToothGrowth, dose %in% c(1.0, 2.0)))
```

```
##
## 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
```

The confidence Interval of -8.996481 , -3.733519 does NOT contain 0 and the p-value is less than 0.5, so we can reject the null hypothesis.

## Conclusions & Assumptions

### Conclusions

The data analysis shows that the evidence wasn't strong enough to suggest a difference in Supplement delivery types. However, the t-tests did show a positive correlation between dosage levels and tooth growth: the higher the dosage, the bigger the growth.

### Assumptions

1. The sample guinea pigs were representative of the entire population.
2. The sample animals were used only once for any combination of dosage and supplement.
3. The samples were independently distributed
4. The variances were different between the sample populations.