

# Study of Tooth Growth

Raghav Pradhan

2024-12-09

## Synopsis

This project is concerned with the ToothGrowth data set in R where I will be performing a bit of **Exploratory Data Analysis** to get the gist of Data & then **statistics** to compare the growth of the teeth by the level of Supplement & Dose given.

## Exploring the Dataset

```
dat <- datasets::ToothGrowth #loading the data set into a variable  
str(dat) #checking the structure of the dataset
```

```
## '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 ...
```

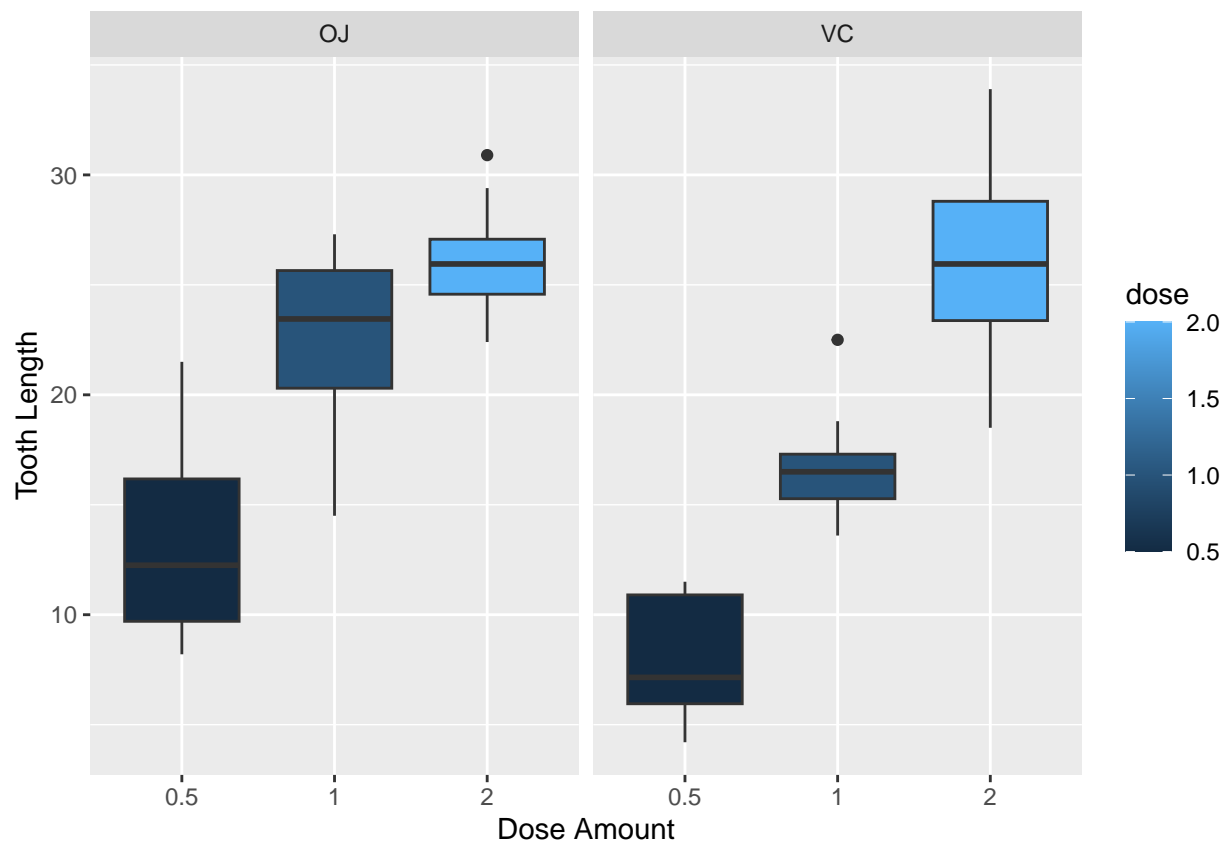
We see that the length & Dose are supposedly **numeric** whereas the supplement is **Factor**,so let's check the factors present in supplement.

```
table(dat$supp)
```

```
##  
## OJ VC  
## 30 30
```

Now after exploring the data set,let's make a plot to visualize the data set to get the gist of the data clearly.

```
library(ggplot2) #using the ggplot package  
ggplot(aes(x = as.factor(dose), y=len), data=dat) + geom_boxplot(aes(fill = dose)) +  
  xlab("Dose Amount") + ylab("Tooth Length") + facet_grid(~supp)
```



With this plot we have visualized the data in the data set, looking clearly we can deduce that the amount of Dose given plays a huge role in its growth, but it's still very much unclear about the difference in supplement, so we will be using stats for further analysis.

## Statistical Tests

For the statistical testing part, I will be performing a series of *Hypothesis-Testing* for both Doses & Supplements.

### Hypothesis testing of Doses

- **NULL hypothesis** - Amount(MG) of Doses given does not influence length of tooth
- **ALTERNATE hypothesis** - Amount(MG) of Doses given does influence length of tooth

```
dose0.5 <- subset(dat, dat$dose == 0.5) #creating a new data set that only has value(0.5) in Dose
dose1.0 <- subset(dat, dat$dose == 1)   #creating a new data set that only has value(1) in Dose
dose2.0 <- subset(dat, dat$dose == 2)   #creating a new data set that only has value(2) in Dose
# by subsetting Dosage into different data set, we can conduct tests between them easily
```

Now after creating different data sets that host each value of dosages, let's do some Testing between Dosages.

```
t.test(dose1.0$len, dose0.5$len, alternative = "greater") #test b/w dose 1 mg and dose 0.5 mg
```

```
##
## Welch Two Sample t-test
##
## data: dose1.0$len and dose0.5$len
## t = 6.4766, df = 37.986, p-value = 6.342e-08
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.753323      Inf
## sample estimates:
## mean of x mean of y
##    19.735    10.605

t.test(dose2.0$len,dose1.0$len,alternative = "greater") #test b/w dose 2 mg and dose 0.5 mg
```

```
##
## Welch Two Sample t-test
##
## data: dose2.0$len and dose1.0$len
## t = 4.9005, df = 37.101, p-value = 9.532e-06
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  4.17387      Inf
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```

we can see that while performing both the test with *95% confidence interval*, the **P-Value\*** are less than **0.05**, infact it is less than **0.01** so with this we can clearly reject the NULL hypothesis\*\* and go with the **ALTERNATIVE** hypothesis that higher Dose(MG) does lead to more growth of teeth.

### Hypothesis testing of Supplement

- **NULL hypothesis** - Both OJ & VC have the same influence on the growth of tooth.
- **ALTERNATE hypothesis** - OJ has more influence on the growth of the tooth.

For this we will be performing tests between supplements for each Dose(0.5,1,2) MG.

```
doseOJ0.5 <- subset(dose0.5,dose0.5$supp == "OJ") #OJ & VC data set for 0.5 MG tests
doseVC0.5 <- subset(dose0.5,dose0.5$supp == "VC")

doseOJ1.0 <- subset(dose1.0,dose1.0$supp == "OJ") #OJ & VC data set for 1 MG tests
doseVC1.0 <- subset(dose1.0,dose1.0$supp == "VC")

doseOJ2.0 <- subset(dose2.0,dose2.0$supp == "OJ") #OJ & VC data set for 2 MG tests
doseVC2.0 <- subset(dose2.0,dose2.0$supp == "VC")
```

Now let's do the testing of Supplements for each respective Dose on the basis of their length.

```
t.test(doseOJ0.5$len,doseVC0.5$len,alternative = "greater")
```

```
##
## Welch Two Sample t-test
##
## data: doseOJ0.5$len and doseVC0.5$len
## t = 3.1697, df = 14.969, p-value = 0.003179
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.34604      Inf
## sample estimates:
## mean of x mean of y
##    13.23      7.98
```

```
t.test(doseOJ1.0$len,doseVC1.0$len,alternative = "greater")
```

```
##
## Welch Two Sample t-test
##
## data: doseOJ1.0$len and doseVC1.0$len
## t = 4.0328, df = 15.358, p-value = 0.0005192
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  3.356158      Inf
## sample estimates:
## mean of x mean of y
##    22.70     16.77
```

```
t.test(doseOJ2.0$len,doseVC2.0$len,alternative = "greater")
```

```
##
## Welch Two Sample t-test
##
## data: doseOJ2.0$len and doseVC2.0$len
## t = -0.046136, df = 14.04, p-value = 0.5181
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -3.1335      Inf
## sample estimates:
## mean of x mean of y
##    26.06     26.14
```

while performing all the tests with *95% confidence interval*, the **P-Value\*** are less than **0.05**, so with **this we can clearly reject the NULL hypothesis\*\*** and go with the **ALTERNATIVE hypothesis** that OJ clearly has more influence on teeth growth compared to VC.

## Conclusion

After exploring the data set and performing appropriate statistical tests, we can conclude that the higher the Dose(MG) more does it influence the growth of the tooth. For the supplement we can deduce that OJ promotes the growth of the tooth far better as compared to the VC.

*Thank you*