What Influences ToothGrowth

Thomas Jo

2021 6 12

Synopt

First We have to load the dataset **"ToothGrowth"**.

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

```
## Warning: package 'ggplot2' was built under R version 4.0.4
```

```
data("ToothGrowth")
```

Let's see the inside of ToothGrowth Dataset.

```
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 ...
```

We notice there are 60 observations and 3 variables in the dataset. Let's use other functions to search more information.

- **Summary function**

```
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  
```

There are 2 levels of factor in supp, it's **OJ** AND **VC**, each has 30 observations.

- **Unique function**

```
unique(ToothGrowth$dose)
```

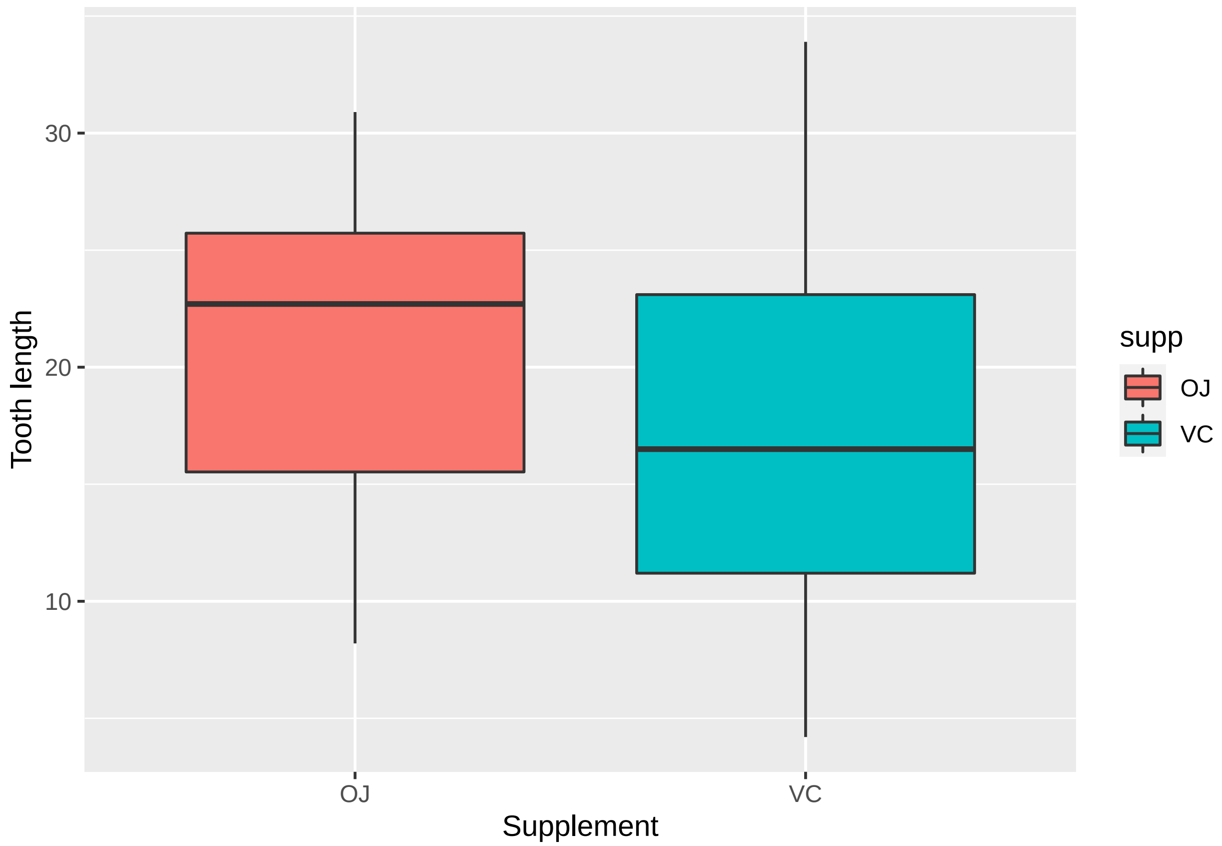
```
## [1] 0.5 1.0 2.0
```

We notice that they only took 0.5, 1.0, 2.0 doses in this measurement.

So far we found enough information to analyze the ToothGrowth Dataset.

1. Differences between two Supplements

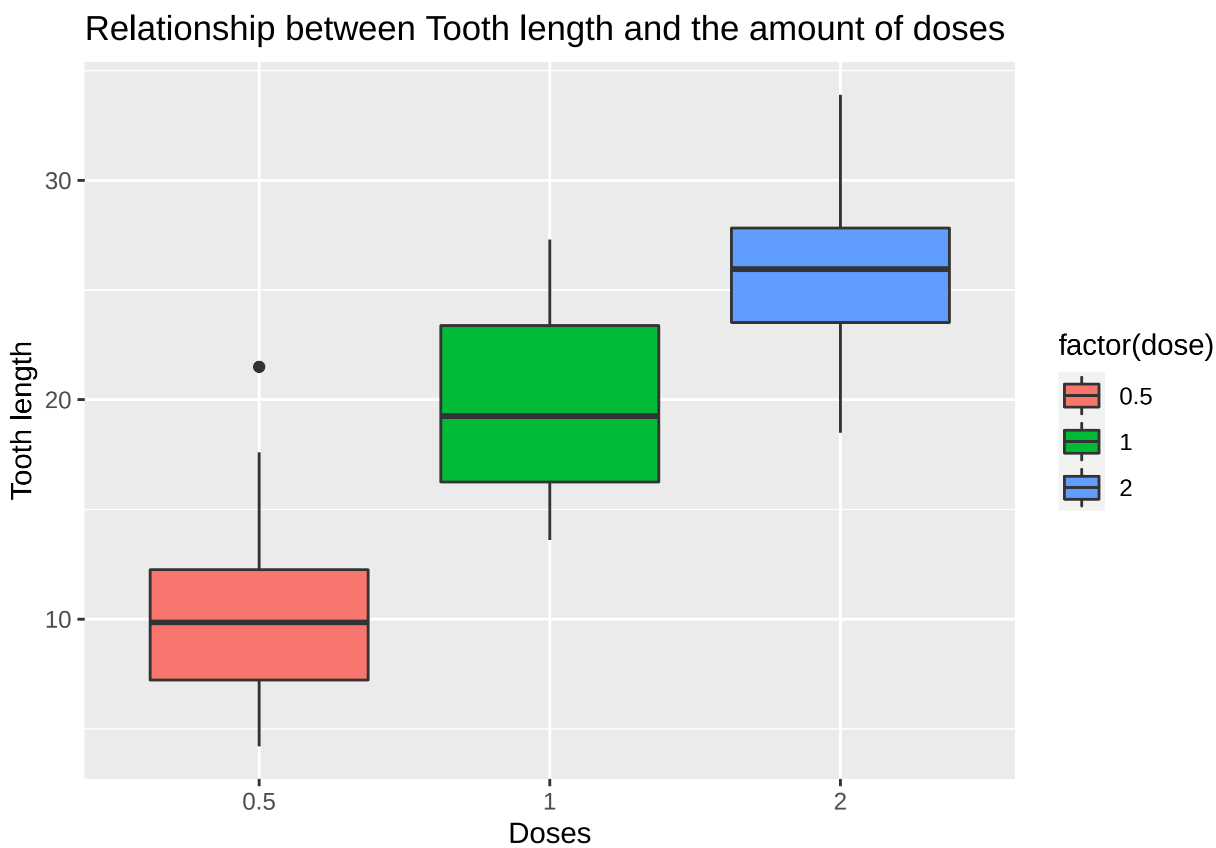
```
ggplot(ToothGrowth,aes(supp,len)) + geom_boxplot(aes(fill=supp)) + xlab("Supplement")+ ylab("Tooth len
```



2. Differences from the amount of doses

```
g <- ggplot(ToothGrowth,aes(x=factor(dose),y=len)) + geom_boxplot(aes(fill=factor(dose))) +
  xlab("Doses")+ylab("Tooth length")

g + labs(title="Relationship between Tooth length and the amount of doses")
```



We can see that the increase of Dosage influence the Tooth Length.

3. Confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose

- **.5 Dose**

```
t.test(len~supp,ToothGrowth[ToothGrowth$dose == .5,])
```

```
##
##  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
```

- **1.0 Dose**

```
t.test(len~supp,ToothGrowth[ToothGrowth$dose == 1,])
```

```
##
##  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
```

- **2.0 Dose**

```
t.test(len~supp,ToothGrowth[ToothGrowth$dose == 2,])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##            26.06            26.14
```

4. Conclusion

Only the usage of 2.0 dose fails to reject the null hypothesis. Other dosage of T-test P-value are less than 0.5, .5 dose P-value is **0.006359** and 1.0 dose P-value is **0.001038**. But the 2.0 dose's p-value is **0.9639**.

It means that using .5 dose and 1.0 dose of the supplement makes a significant difference between the group **OJ** AND **VC**. But when the dosage increases up to 2.0, there was no significant difference between group **OJ** AND **VC**. That is why they failed to reject the null hypothesis.