

Analysing ToothGrowth Dataset

Hizban Sheik

2022-09-19

Exploratory Data Analysis

Lets see Average Length Growth for each supp

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.1.3
```

```
toothgrowth <- ToothGrowth
mean.len <- with(toothgrowth,tapply(len, supp, mean))
type <- c("OJ","VC")
avg.data <- data.frame(avg=mean.len,type=type)
g <- ggplot(avg.data,aes(x=type,y=avg)) + geom_bar(aes(fill=type),stat = 'identity') +
ylab("Average Length Grown")+ geom_text(aes(label=avg), position=
position_dodge(width=0.9), vjust=-0.25) + ggtitle("Average Growth of Tooth Length for each supp")
g
```

So,Average Length of Tooth Growth is higher for supp “OJ” than “VC”.

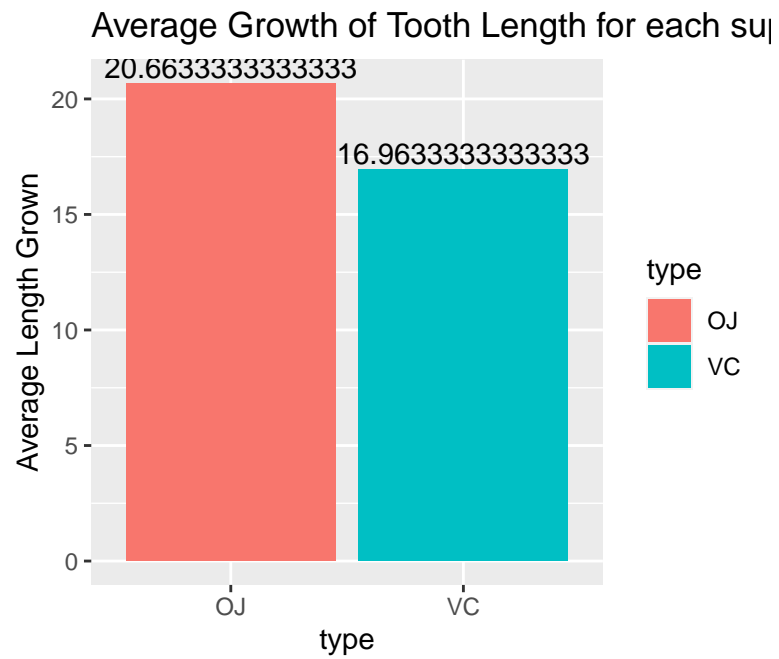
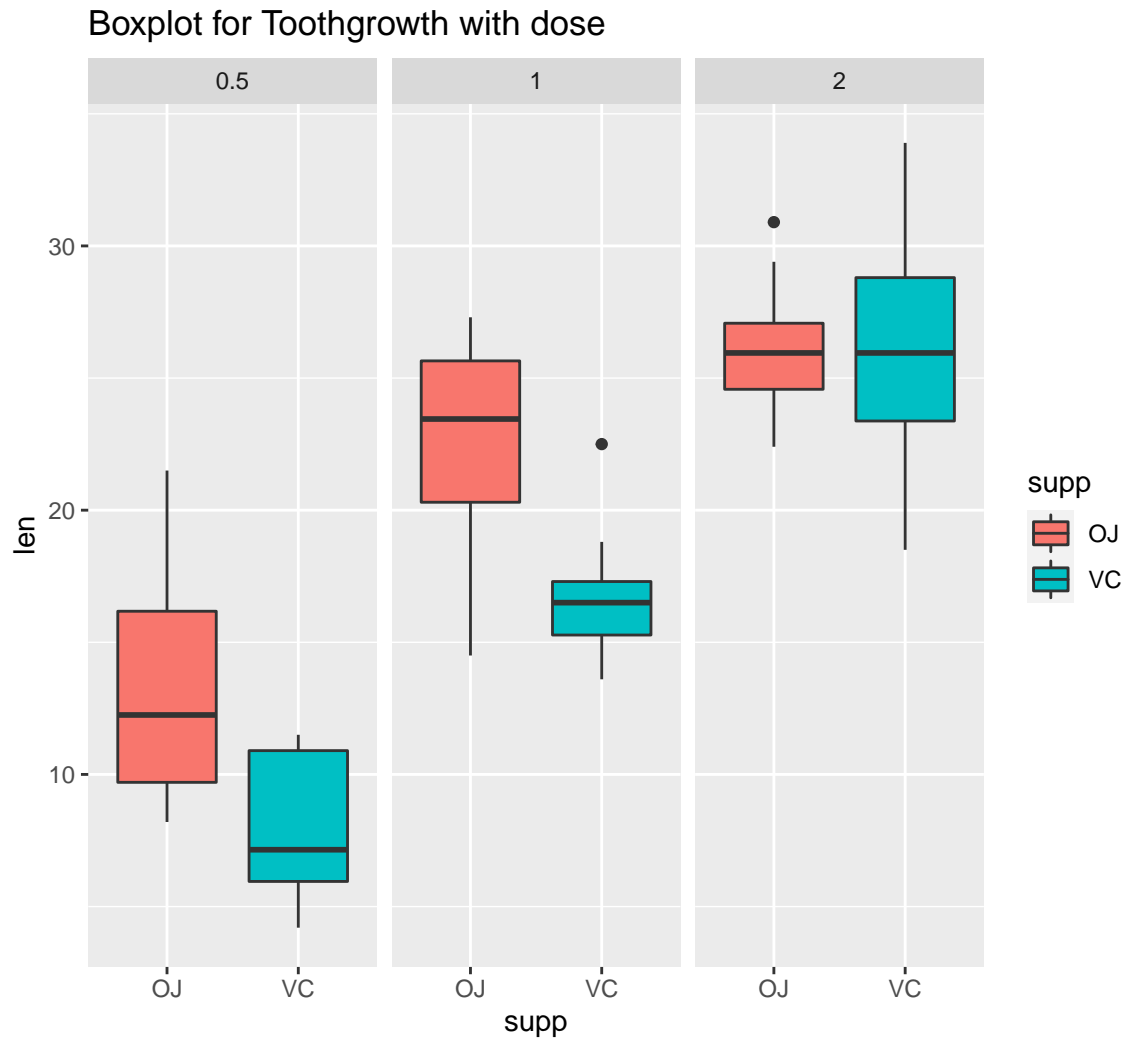


Figure 1: Mean of the Teeth Growth Length for each type

```
ggplot(ToothGrowth,aes(x=supp,y=len)) + geom_boxplot(aes(fill=supp)) + facet_grid(.~dose) +
  ggtitle("Boxplot for Toothgrowth with dose ")
```



So the Box Plot shows that as concentration of dose increases Tooth Growth also increases. This is visible for both “OJ” and “VC”. For the case of “VC” there is significant improvement from dosage of 1 to 2, while that is not the case for “OJ”.

Hypothesis Testing

There are two features in this dataset on which Hypothesis Testing can be performed: 1) Hypothesis Test on “supp” i.e supplement type 2) Hypothesis Test on the “dose”

Hypothesis Test on “supp”

Let our Null Hypothesis be that there is no difference in toothgrowth due to supplement type.

Null Hypothesis: Toothgrowth due to “OJ” = Toothgrowth due to “VC”

Alternative Hypothesis: Toothgrowth due to “OJ” > Toothgrowth due to “VC”

```
oj.data <- subset(ToothGrowth,supp=="OJ")
vc.data <- subset(ToothGrowth,supp=="VC")
t.test(x=oj.data$len,y=vc.data$len,alternative='greater',var.equal=FALSE,paired=FALSE)
```

```
##
## Welch Two Sample t-test
##
## data:  oj.data$len and vc.data$len
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687      Inf
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

So, the p-value is 0.03 which is less than $\alpha=0.05$ (Type I error rate) so we reject the null hypothesis. This means that supplement type “OJ” gives better Toothgrowth than the supplement type “VC”. ## Hypothesis Test on Dosage **Null Hypothesis:** Toothgrowth does not change with the concentration of dose

Alternative Hypothesis: Toothgrowth due to “OJ” > Toothgrowth increases with concentration of dose

```
dose1 <- subset(ToothGrowth,dose=="0.5")
dose2 <- subset(ToothGrowth,dose=="1")
dose3 <- subset(ToothGrowth,dose=="2")
```

First for dose of 0.5mg and 1mg

```
t.test(x=dose1$len,y=dose2$len,alternative="less",paired=FALSE,var.equal=FALSE)
```

```
##
## Welch Two Sample t-test
##
## data:  dose1$len and dose2$len
## t = -6.4766, df = 37.986, p-value = 6.342e-08
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -6.753323
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

So p-value is 6.342e-08 which is very less than $\alpha=0.05$ so we reject the null hypothesis, this means that Toothgrowth for dose of 0.5mg is less than Toothgrowth for dose of 1mg.

First for dose of 1mg and 2mg

```
t.test(x=dose2$len,y=dose3$len,alternative="less",paired=FALSE,var.equal=FALSE)
```

```
##
## Welch Two Sample t-test
##
```

```
## data: dose2$len and dose3$len
## t = -4.9005, df = 37.101, p-value = 9.532e-06
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -4.17387
## sample estimates:
## mean of x mean of y
##      19.735    26.100
```

The p-value is $9.532e-06$ again it is very less than $\alpha=0.05$ so we reject the null hypothesis, this means Toothgrowth for dose of 1mg is less than Toothgrowth for 2mg.

Assumptions

- 1) Variance between supplement types and dose is not equal.
- 2) The Data is non-paired i.e experiments are done on different subjects for each supp and dose.
- 3) The variables are independent and identically distributed.

Conclusions

- 1) Supplement type “OJ” gives more Toothgrowth than the supplement type “VC”.
- 2) Toothgrowth increases as the concentration or power of dosage increases.