Analyze ToothGrowth Dataset

York Huang

Sunday, March 13, 2016

Overview

In the project, we're going to analyze the ToothGrowth data in the R datasets package

Prepare Data

Load data from R ToothGrowth package and convert to data frame.

```
library(ggplot2)

data(ToothGrowth)
tg <- data.frame(ToothGrowth)
```

Exploration Data

There are 60 observations.

```
dim(tg)
```

```
## [1] 60  3
```

```
head(tg)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

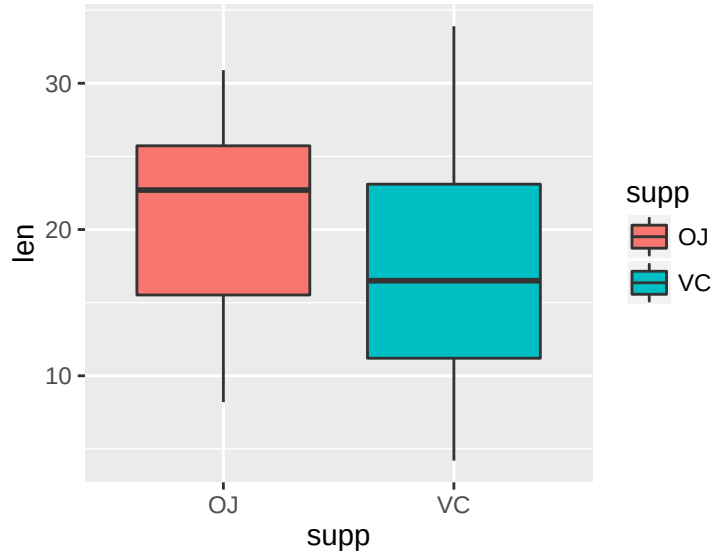
Summary of data (tg) -

```
summary(tg)
```

```
##       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
```

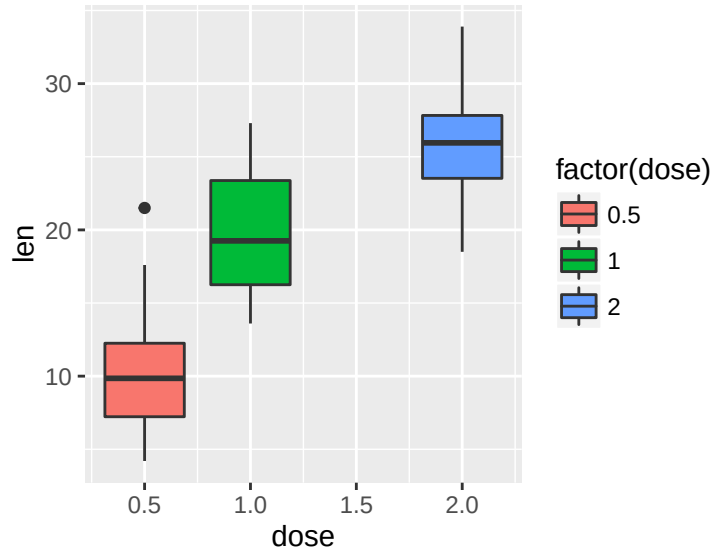
Plot tooth length vs supplement.

```
ggplot(data=tg, aes(x=supp, y=len)) + geom_boxplot(aes(fill=supp))
```



Plot tooth length vs dose.

```
ggplot(data=tg, aes(x=dose, y=len)) + geom_boxplot(aes(fill=factor(dose)))
```



Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose

Compare between supplements.

```
t.test(tg$len[tg$supp=="OJ"], tg$len[tg$supp=="VC"], paired = FALSE, var.equal = FALSE)$conf
```

```
## [1] -0.1710156  7.5710156
## attr(,"conf.level")
## [1] 0.95
```

The 95 percent confidence interval (-0.1710156, 7.5710156) include 0. We are not conficent to compare the impacts between OJ and VC supplements.

Compare between dose 0.5 and 1.

```
t.test(tg$len[tg$dose==0.5], tg$len[tg$dose==1], paired = FALSE, var.equal = FALSE)$conf
```

```
## [1] -11.983781  -6.276219
## attr(,"conf.level")
## [1] 0.95
```

The 95 percent confidence interval (-11.983781, -6.276219) does NOT include 0. We are conficent to say it is helpful to tooth length when increasing does from 0.5 to 1.

Compare between dose 1 and 2.

```
t.test(tg$len[tg$dose==1], tg$len[tg$dose==2], paired = FALSE, var.equal = FALSE)$conf
```

```
## [1] -8.996481 -3.733519
## attr(,"conf.level")
## [1] 0.95
```

The 95 percent confidence interval (-8.996481, -3.733519) does NOT include 0. We are conficent to say it is helpful to tooth length when increasing does from 1 to 2.

State your conclusions and the assumptions needed for your conclusions

We can summarize the following. * We can't find the impact of supplements on tooth length. * Increasing dose is helpful to tooth length. We also note that the observation size is still small. More data may help us better understand the impacts.