Sama Issa Basic inferential data analysis

Overview

Now in the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.
4. State your conclusions and the assumptions needed for your conclusions.

inferential data analysis

Load the ToothGrowth data and perform some basic exploratory data analyses

First we load the ToothGrowth dataset

```
data(ToothGrowth)
```

Provide a basic summary of the data

basic summary of data

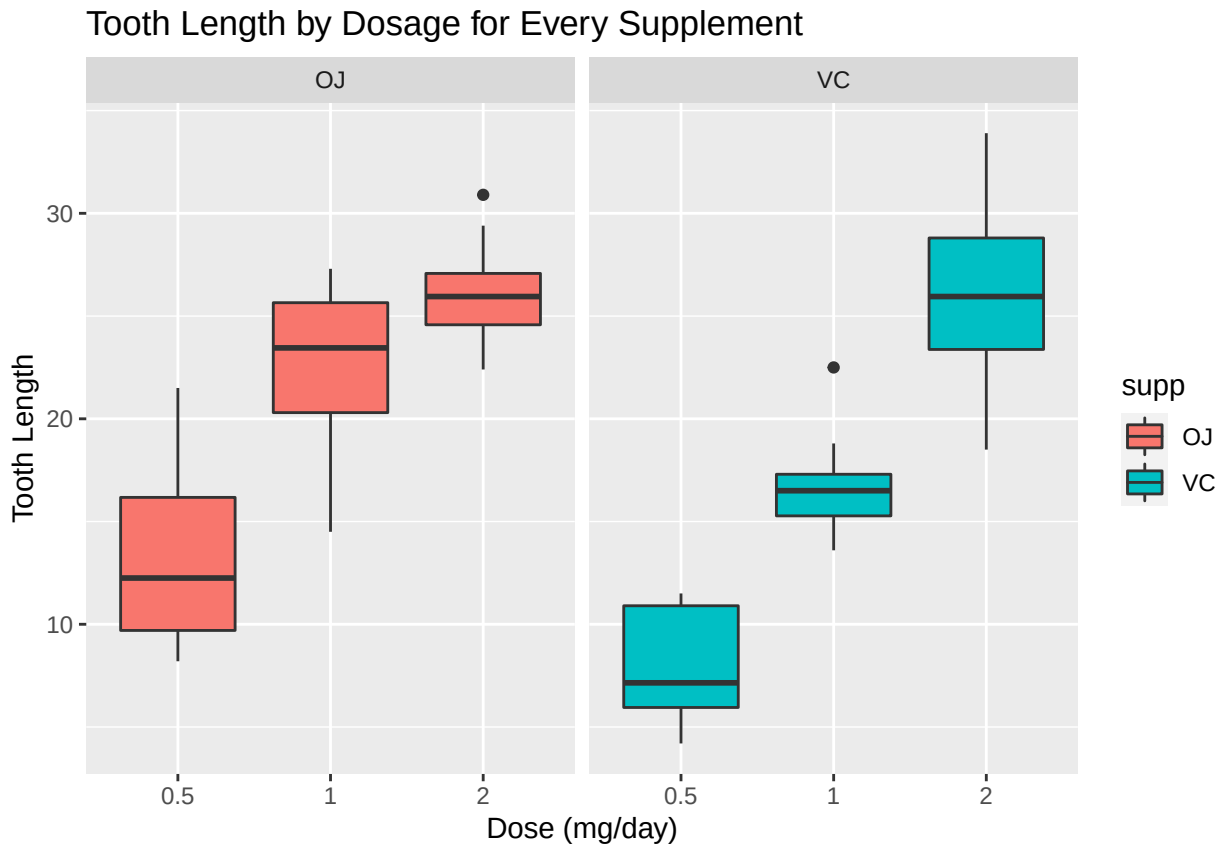
```
str(ToothGrowth)
```

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

```
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
```

```r
library(ggplot2)
g <- ggplot(ToothGrowth, aes(x = factor(dose), y = len))
g <- g + facet_grid(.~supp)
g <- g + geom_boxplot(aes(fill = supp))
g <- g + labs(title = "Tooth Length by Dosage for Every Supplement")
g <- g + labs(x = "Dose (mg/day)", y = "Tooth Length")
print(g)
```



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

```r
# Comparison by delivery method for the same dosage
t05 <- subset(ToothGrowth, dose == "0.5")
t1 <- subset(ToothGrowth, dose == "1")
t2 <- subset(ToothGrowth, dose == "2")

## Testing Dose wise
t05 <- t.test(len ~ supp, data = t05)
t1 <- t.test(len ~ supp, data = t1)
t2 <- t.test(len ~ supp, data = t2)

summaryBYsupp <- data.frame(
      "p-value" = c(t05$p.value, t1$p.value, t2$p.value),
      "Conf.Low" = c(t05$conf.int[1],t1$conf.int[1], t2$conf.int[1]),
```

```
      "Conf.High" = c(t05$conf.int[2],t1$conf.int[2], t2$conf.int[2]),
      row.names = c("Dosage .05","Dosage 1","Dosage 2"))

# Show the data table
print("Dosage .05 :")
```

```
## [1] "Dosage .05 :"
```

```
print(paste("    p-value :", t05$p.value))
```

```
## [1] "    p-value : 0.0063586067640968"
```

```
print(paste("    Conf.Low :", t05$conf.int[1]))
```

```
## [1] "    Conf.Low : 1.71905727146767"
```

```
print(paste("    Conf.High :", t05$conf.int[2]))
```

```
## [1] "    Conf.High : 8.78094272853233"
```

```
print("Dosage 1 :")
```

```
## [1] "Dosage 1 :"
```

```
print(paste("    p-value :", t1$p.value))
```

```
## [1] "    p-value : 0.00103837587229988"
```

```
print(paste("    Conf.Low :", t1$conf.int[1]))
```

```
## [1] "    Conf.Low : 2.80214824916537"
```

```
print(paste("    Conf.High :", t1$conf.int[2]))
```

```
## [1] "    Conf.High : 9.05785175083463"
```

```
print("Dosage 2 :")
```

```
## [1] "Dosage 2 :"
```

```
print(paste("    p-value :", t2$p.value))
```

```
## [1] "    p-value : 0.963851588723373"
```

```r
print(paste("    Conf.Low :", t2$conf.int[1]))
```

```
## [1] "    Conf.Low : -3.79807046333516"
```

```r
print(paste("    Conf.High :", t2$conf.int[2]))
```

```
## [1] "    Conf.High : 3.63807046333515"
```

State your conclusions and the assumptions needed for your conclusions

OJ delivers more tooth growth than VC for dosages 0.5 & 1.0. OJ and VC deliver the same amount of tooth growth for dose amount 2.0 mg/day. For the entire data set we cannot conclude OJ is more effective that VC.