Statistical Inference Course Project Part 2

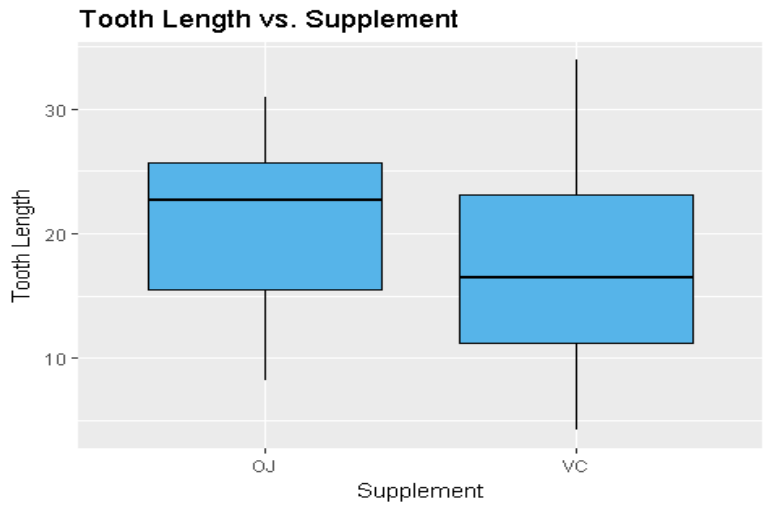
Charles Eads

Part 2: Basic Inferential Data Analysis

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. (Only use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

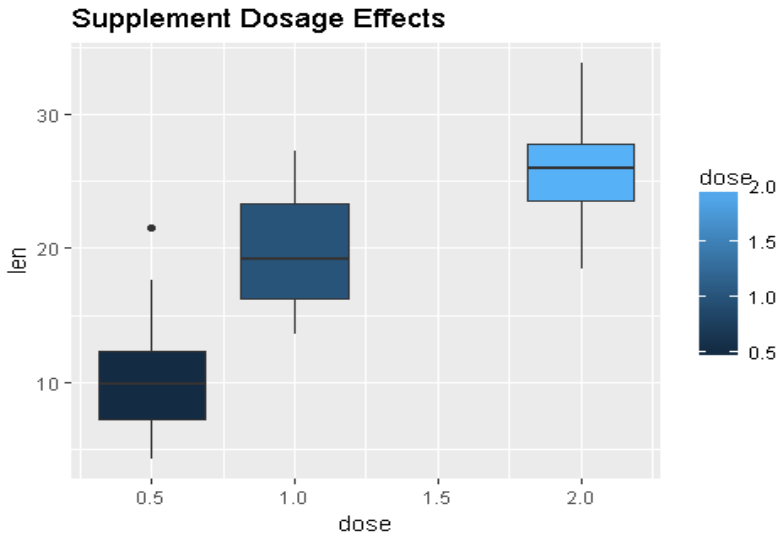
Questions 1 and 2

```
library(datasets)
data("ToothGrowth")
#getting a description of the data
?ToothGrowth

head(ToothGrowth)

#Summary of the data
summary(ToothGrowth)

library(ggplot2)

ggplot(aes(x = supp, y = len), data = ToothGrowth) + xlab("Supplement") +
ylab("Tooth Length") + geom_boxplot(fill='#56B4E9', color="black") +
ggtitle("Tooth Length vs. Supplement")
```



It appears the length of the Guinea Pigs teeth are longer for those given the OJ supplement.
#Dosage of Supplement Effects

```r
#Creating another plot for dosage
ggplot(ToothGrowth, aes(x=dose, y=len, fill=dose, group = dose)) +
geom_boxplot() + ggtitle("Supplement Dosage Effects")
```



Question 3

```r
#t test for Supplements
t.test(data = ToothGrowth, len ~ supp, paired = FALSE, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

The p-value is very low at around 6% and the confidence interval has one that is -.17(close to 0). We cannot reject the NULL hypothesis for supplements not effecting tooth length at this time.

```r
t.test(len ~ supp, paired = FALSE, var.equal = TRUE, data =
subset(ToothGrowth, dose ==.5))
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 18, p-value = 0.005304
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  1.770262 8.729738
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

```
t.test(len ~ supp, paired = FALSE, var.equal = TRUE, data =
subset(ToothGrowth, dose ==1))
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 18, p-value = 0.0007807
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.840692 9.019308
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

The confidence intervals on the above 2 are 1.77 and 8.73 at .5mg/day and 2.84 and 9 at 1 mg/day so we can reject the NULL hypothesis.

```
t.test(len ~ supp, paired = FALSE, var.equal = TRUE, data =
subset(ToothGrowth, dose ==2))
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 18, p-value = 0.9637
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.722999  3.562999
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

The confidence interval here is -3.72 and 3.46 at 2.0mg/day so we can't reject the NULL hypothesis.

Question 4

It appears that the more dosage you have of either supplement leads to more tooth growth based on the means of the groups above on the t-tests. It also appears that Orange Juice gives more tooth growth than Vitamin C at the .5mg/day and 1mg/day dosages. However, both Vitamin C and Orange Juice have the same amount of tooth growth for 2mg/day dosages. This leads me to believe I can't conclude that Orange Juice is not a better way to increase tooth growth at higher dosages than Vitamin C, but is at lower dosages(below 2.0mg/day).