Statistical Inference Course Project Part 2

Overview: In this project, we're going to analyze the ToothGrowth data in the R datasets package and study the ToothGrowth data set and compare the tooth growth by supplement type and dose.

Load the ToothGrowth data and perform some basic exploratory data analyses

```
head(ToothGrowth)
```

```
##    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
```

```
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 ...
```

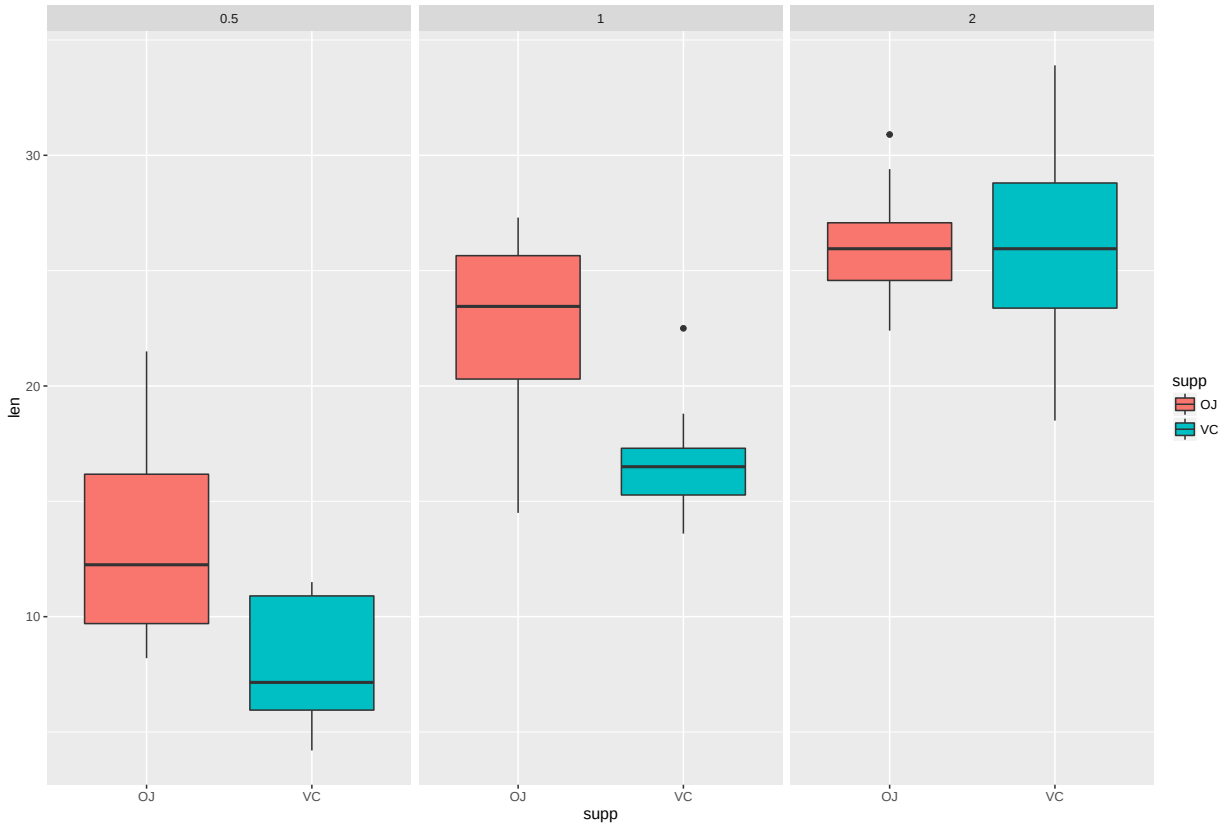
Provide a basic summary of the data

```
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
```

Boxplot comparison by delivery methods and dose level

```
ggplot(aes(x=supp, y=len), data=ToothGrowth)+geom_boxplot(aes(fill = supp))+facet_wrap(~dose)
```



The above graph indicates that the dosage has significant impact on teeth growth. At dose level 0.5 and 1, OJ seems more effective than VC. But it is hard to tell at the dose level 2.

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

Hypothesis: Different Supplements have no effect on tooth growth

```
t.test(len~supp, data = ToothGrowth)
```

```
##
##  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 confidence interval includes 0 and the p-value is above the 0.05 threshold. This indicates that we fail to reject the null hypothesis.

Hypothesis: Supplement doesage have no effect on tooth growth We will run differet groups of dose for this Hypothesis analysis: 0.5 & 1, 1 & 2, 0.5 &2

```r
dose1 <- subset(ToothGrowth, dose %in% c(0.5, 1))
t.test(len ~ dose, data = dose1)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

```r
dose2 <- subset(ToothGrowth, dose %in% c(1, 2))
t.test(len ~ dose, data = dose2)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

```r
dose3 <- subset(ToothGrowth, dose %in% c(0.5, 2))
t.test(len ~ dose, data = dose3)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean in group 0.5   mean in group 2
##            10.605            26.100
```

The confidence intervals for all the combination of dose pairs do not includes 0 and the p-value is below the 0.05 threshold. This indicates that we can reject this null hypothesis.

Hypothesis: OJ has same effect as VC for tooth growth at 1 dose level

```r
ToothGrowth.5 <- subset(ToothGrowth, dose == 0.5)
t.test(len ~ supp, data = ToothGrowth.5, paired = FALSE)
```

```
##
##  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
```

The confidence interval does not include 0 and the p-value is below the 0.05 threshold. This indicates that we can reject this null hypothesis.

Hypothesis: OJ has same effect as VC for tooth growth at 1 dose level

```
ToothGrowth.1 <- subset(ToothGrowth, dose == 1)
t.test(len ~ supp, data = ToothGrowth.1, paired = FALSE)
```

```
##
##  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
```

The confidence interval does not include 0 and the p-value is below the 0.05 threshold. This indicates that we can reject this null hypothesis.

Hypothesis: OJ is more effective than VC for tooth growth at 2 dose level

```
ToothGrowth.2 <- subset(ToothGrowth, dose == 2)
t.test(len ~ supp, data = ToothGrowth.2, paired = FALSE)
```

```
##
##  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
```

The confidence interval includes 0 and the p-value is above the 0.05 threshold. This indicates that we can not reject this null hypothesis.

Conclusions and Assumptions

Hypothesis of different supplements have no effect on tooth growth – fail to reject Hypothesis of supplement doesage have no effect on tooth growth – reject Hypothesis of OJ has same effect as VC for tooth growth

at 0.5 dose level – reject Hypothesis of OJ has same effect as VC for tooth growth at 1 dose level – reject Hypothesis of OJ has same effect as VC for tooth growth at 2 dose level – fail to reject

The above tests indicate that the doseage of suplements has significant impact on tooth growth. It is hard to tell if the type of supplements makes different on tooth growth in general. But for dose level 0.5 and 1, OJ is better than VC.

Assumptions:
1. assuming unequal variances between the two groups 2. conf.level = 0.95 3. alternative = two.sided