Part 2: Basic inferential data analysis.

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

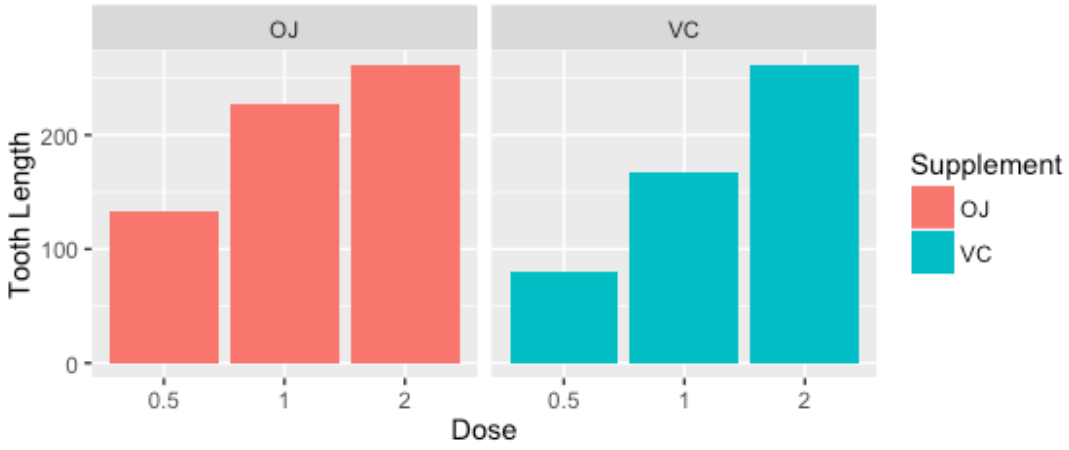
1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
data(ToothGrowth)
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 ...
```

2. 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
```

The plot shows that the length of the tooth appears to be related to the amount of dosage.



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).

```
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 p-value is equal to 0.06063. The 95 percent confidence interval is -.1710156 to 7.5710156. This contains 0 for the t-test. This means we fail to reject that the other supplement has no affect on tooth growth.

Split the doses into the 3 subgroups: 0.5 - 1 (A), 1 - 2 (B), and 0.5 - 2 (C).

```
ToothGrowth.dose_A <- subset (ToothGrowth, dose %in% c(0.5, 1.0))
ToothGrowth.dose_B <- subset (ToothGrowth, dose %in% c(1.0, 2.0))
ToothGrowth.dose_C <- subset (ToothGrowth, dose %in% c(0.5, 2.0))
```

Run the t-test for all 3 subgroups: 0.5 - 1 (A), 1 - 2 (B), and 0.5 - 2 (C)

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

##
##  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
```

The p-value is less than 0.05, confidence interval does not contain 0. The mean in group 0.5 = 10.605 and mean in group 1 = 19.735

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

##
##  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
```

The p-value is less than 0.05, confidence interval does not contain 0. The mean in group 1 = 19.735 and mean in group 2 = 26.100

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

##
##  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 p-value is less than 0.05, confidence interval does not contain 0. The mean in group 0.5 = 10.605 and mean in group 2 = 26.100

The null hypothesis can be rejected. As the dose increases from 0.5 to 2, the mean length of the tooth also increases from 10.605 to 26.100.

4. State your conclusions and the assumptions needed for your conclusions.

We can conclude that increasing the dose of a supplement (regardless of type) will increase tooth length for guinea pigs.

We assumed that the populations of guinea pigs are similar and randomly selected.