Course Inferential Statistics Madelon den Boeft March 15 2018
==========

Assignment: Analyzing the Tooth Growth data in R

load plotting library

```
library(ggplot2)
```

load tooth growth data

```
data(ToothGrowth)
```

Perform basic exploratory data analysis

```
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 ...
```

```
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
```

```
dim(ToothGrowth)
```

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

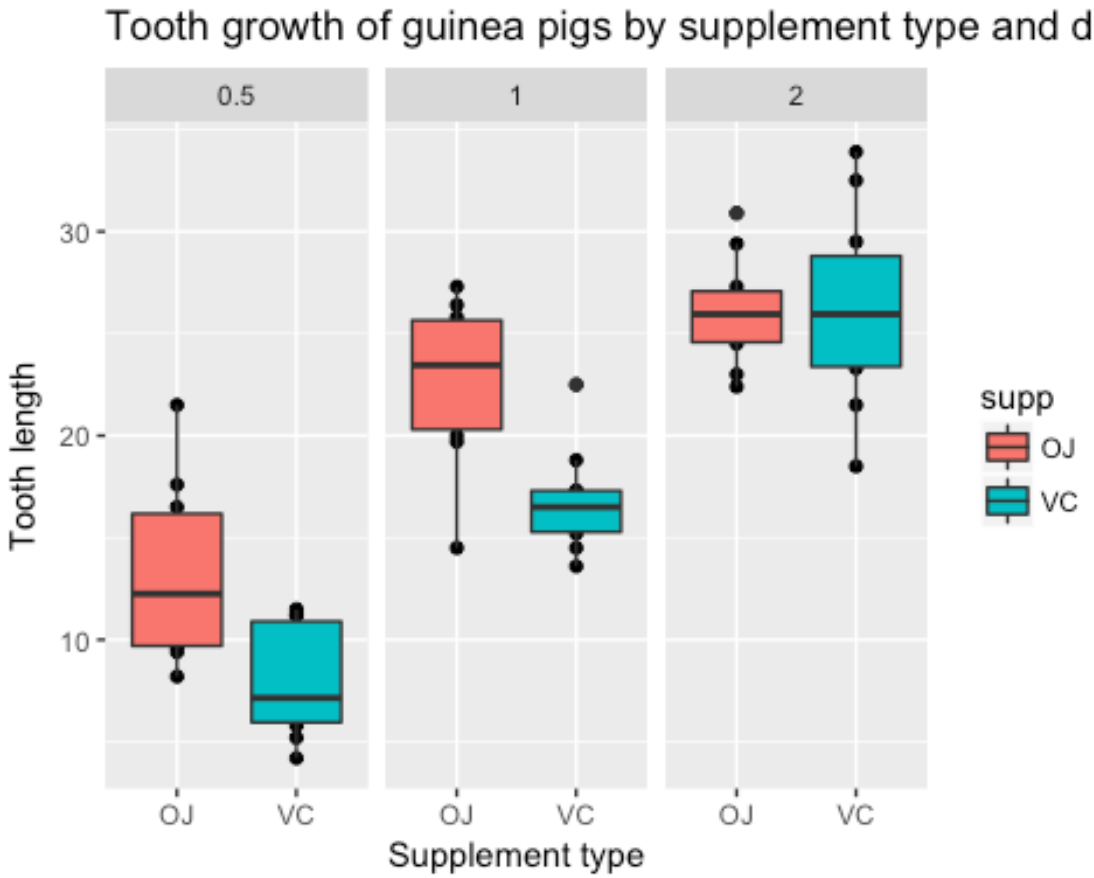
Provide a basic summary of the data

```
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
```

Use CI and/or hypothesis tests to compare tooth grow by supplement type and dosage

```
qplot(supp,len,data=ToothGrowth, facets=~dose, main="Tooth growth of guinea pigs by supplement type and dosage",xlab="Supplement type", ylab="Tooth length") + geom_boxplot(aes(fill = supp))
```



==> The dosage increases tooth growth (positive effect)

==> There is a lineair relationship between VC and tooth growth

==> In general, OJ induces more growth than VJ. However, for OJ dosage 2.0 there is less growth imrovement

Using confidence intervals (CI)

```
Lower <- subset(ToothGrowth, dose %in% c(0.5, 1.0))
Middle <- subset(ToothGrowth, dose %in% c(0.5, 2.0))
Upper <- subset(ToothGrowth, dose %in% c(1.0, 2.0))
```

Next we will do the t test on dosage and on supplement

```
t.test(len ~ dose, paired = F, var.equal = F, data = Lower)
```

```
##
##  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
```

```
t.test(len ~ dose, paired = F, var.equal = F, data = Middle)
```

```
##
##  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
```

```
t.test(len ~ dose, paired = F, var.equal = F, data = Upper)
```

```
##
##  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
```

```
t.test(len ~ supp, paired = F, var.equal = F, 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
```

To draw any conclusions, we must assume the following assumptions:

1. The poplulations are independent, the variances between populations are different and a random population was used

2. The population was comprised of similar guinea pigs.

3. Measurement error was accounted for with significant digits.

4. Double blind research methods were used.

Conclusions

1. Supplement type has no effect on tooth growth

2. Increasing the dosage level leads to increased tooth growth