Load the ToothGrowth data and perform some basic exploratory data analyses

```r
data("ToothGrowth")
```

Provide a basic summary of the data.

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

```r
library(ggplot2)
ggplot(ToothGrowth, aes(x=len, fill=supp)) + facet_wrap(~dose) + xlab('tooth length') + ggtitle('distril
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```

distribution of guinea pig tooth lengths
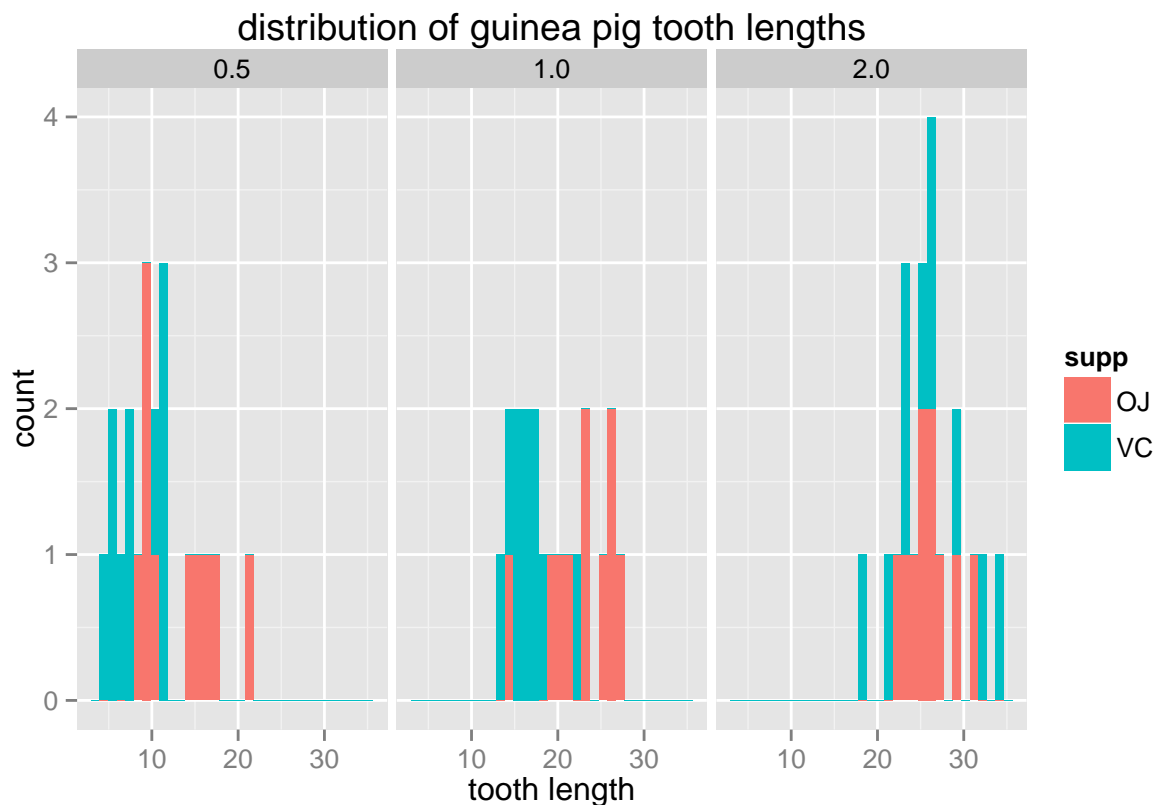
1. test the len with the dos

```r
t.test(ToothGrowth$len, ToothGrowth$dose, paired=FALSE, var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len and ToothGrowth$dose
## t = 17.81, df = 59.798, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  15.66453 19.62881
## sample estimates:
## mean of x mean of y
## 18.813333  1.166667
```

The 95% percent confidence interval doesn't contain 0, that means we can reject the null hypothesis. Means tooth len is related to the dose.

2. test the len with the supp in low does

```
d1<-ToothGrowth[ToothGrowth$dose==0.5,]
t.test(d1[d1$supp=="VC",]$len, d1[d1$supp=="OJ",]$len, paired=FALSE, var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  d1[d1$supp == "VC", ]$len and d1[d1$supp == "OJ", ]$len
## 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:
##  -8.780943 -1.719057
## sample estimates:
## mean of x mean of y
##      7.98     13.23
```

The 95% percent confidence interval doesn't contain 0, that means we can reject the null hypothesis. Means in low dose, tooth len is related to suppliment.

3. test the len with the supp in median dose

```
d2<-ToothGrowth[ToothGrowth$dose==1.0,]
t.test(d2[d2$supp=="VC",]$len, d2[d2$supp=="OJ",]$len, paired=FALSE, var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  d2[d2$supp == "VC", ]$len and d2[d2$supp == "OJ", ]$len
## 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:
##  -9.057852 -2.802148
## sample estimates:
## mean of x mean of y
##     16.77     22.70
```

The 95% percent confidence interval doesn't contain 0, that means we can reject the null hypothesis. Means in median dose, tooth len is related to the suppliment.

4. test the len with the supp in high dose

```
d3<-ToothGrowth[ToothGrowth$dose==2.0,]
t.test(d3[d3$supp=="VC",]$len, d3[d3$supp=="OJ",]$len, paired=FALSE, var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  d3[d3$supp == "VC", ]$len and d3[d3$supp == "OJ", ]$len
```

```
## 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:
##  -3.63807  3.79807
## sample estimates:
## mean of x mean of y
##     26.14    26.06
```

The 95% percent confidence interval contains 0, that means we can't reject the null hypothesis. Means in high dose, we can't say tooth len relate to supp.

Conclusion

1. ToothGrowth is related to dose.
2. In low and medium dose, ToothGrowth len is related to supp.
3. In high dose == 2.0, there is no obvious relationship with supp.