# Statistical_Inference_Course_Project_Part_2

*Jason B.*

## Overview

**The purpose of this exercise is to perform basic inferential statistical analysis of the ToothGrowth data.**

**The documention states, "The Effect of Vitamin C on Tooth Growth in Guinea Pigs: The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice or ascorbic acid (a form of vitamin C and coded as VC)."**

## Load the data

```
library(ggplot2)
library(datasets)
data("ToothGrowth")
```

## Explore the data

```
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 ...
```

```
sum(is.na(ToothGrowth))
```

```
## [1] 0
```

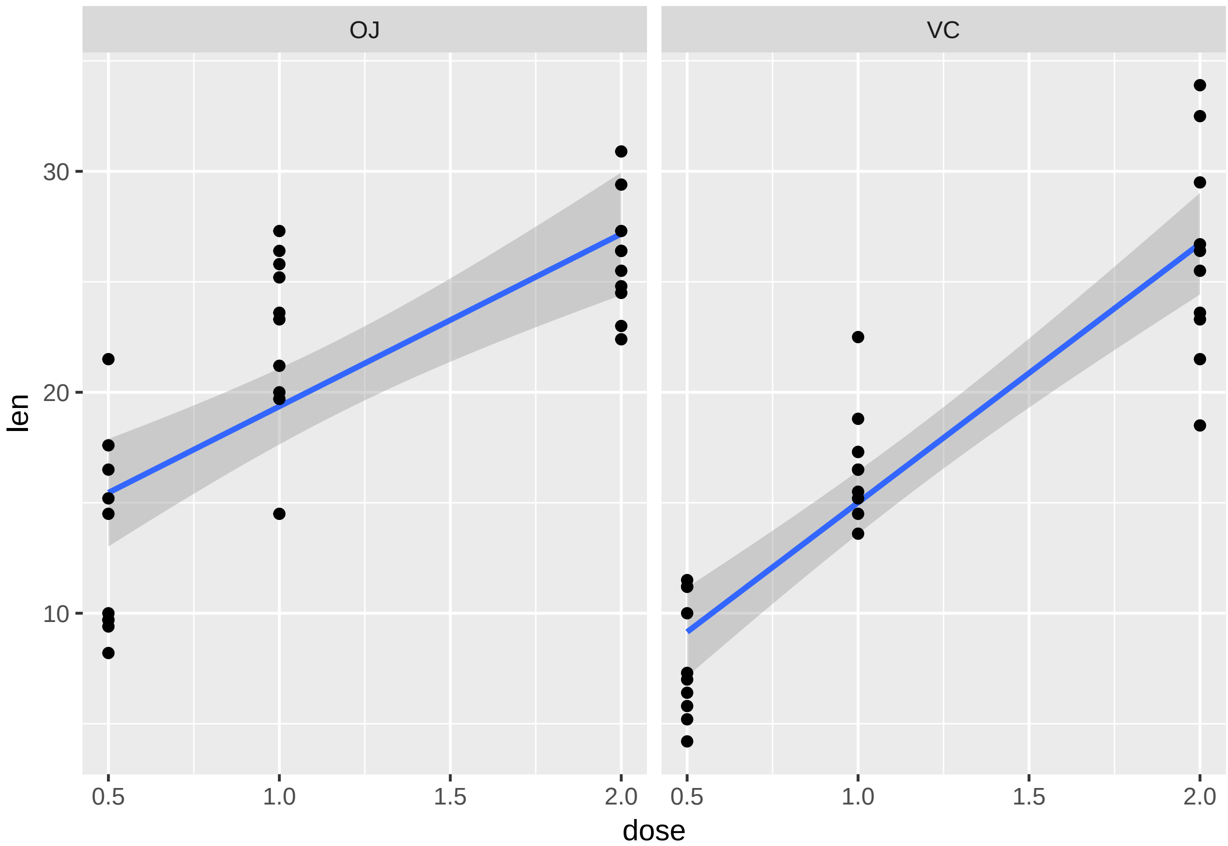
```
summary(ToothGrowth)
```

```
##       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
```

**There are 3 variables, length, supplement, and dose. There are no missing values and the data frame appears "tidy". The variable len and dose are numeric doubles and the variable supp is a factor with two levels.**

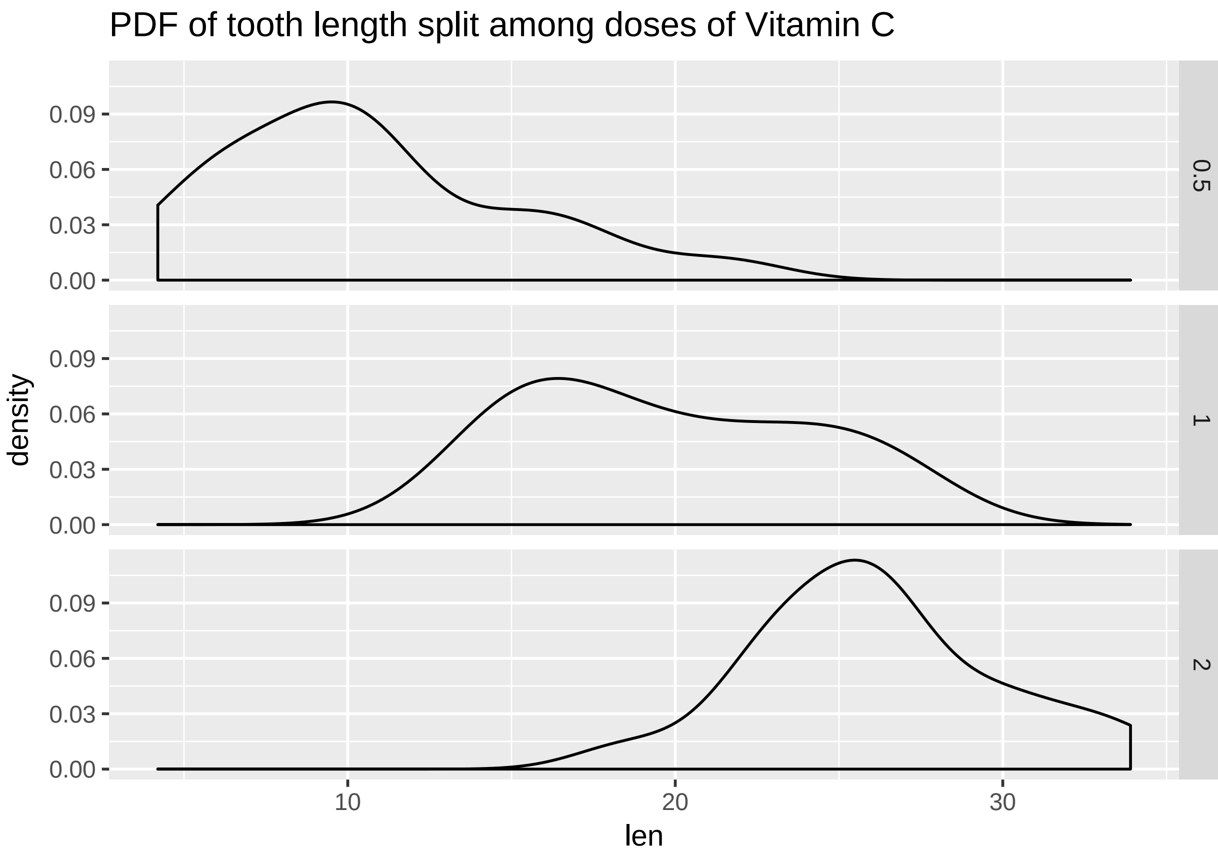
## Plot the data

```
ggplot(ToothGrowth, aes(dose, len))+
        geom_smooth(method = "lm")+
        geom_point()+
        facet_grid(.~supp)
```




**There is a general upward trend for tooth length as a function of dose. The slope of the regression line is steeper for delivery via vitamin C.**

```
ggplot(ToothGrowth, aes(len))+
       geom_density()+
       facet_grid(dose~.)+
       ggtitle("PDF of tooth length split among doses of Vitamin C")
```




**These density plots show the distribution of tooth length values as a function of dose irrespective of delivery method. The data are somewhat normally distributed. I will assume a normal distribution for inferential analysis.**

### Inferential analysis

```
t.test(len ~ supp, data = ToothGrowth, var.equal = F, alternative = "two.sided")
```

```
##
##  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
```

**This analysis compares the effect of delivery methods on tooth length. It is not clear if the authors had a directional hypothesis; so, I performed a two-tailed test. Also, to be conservative, I assumed unequal variances. The p-value is greater than alpha = 0.05; therefore, this analysis fails to reject the null hypothesis that there is no effect of delivery method.**

```r
#subset by dose
tg_low <- subset(ToothGrowth$len, ToothGrowth$dose == 0.5)
tg_mid <- subset(ToothGrowth$len, ToothGrowth$dose == 1.0)
tg_hi <- subset(ToothGrowth$len, ToothGrowth$dose == 2.0)
plm <- t.test(tg_low, tg_mid)$p.value
plh <- t.test(tg_low, tg_hi)$p.value
pmh <- t.test(tg_hi, tg_mid)$p.value
uncorr_p <- c(p_low_mid = plm, p_low_hi = plh, p_mid_hi = pmh)
corr_p <- 3*uncorr_p
corr_p
```

```
##    p_low_mid     p_low_hi     p_mid_hi
## 3.804902e-07 1.319257e-13 5.719289e-05
```

**This analysis comparing the effect of different doses of Vitamin C on tooth length reveals something interesting. The p-values from the comparisons of the low to mid, low to hi, and mid to hi doses are all less than alpha = 0.5. Again, the tests were performed assuming unequal variances and with a two-sided hypothesis. Note, that I used a Bonferroni correction to control for multiple hypothesis testing. Thus, this analysis rejects the null hypothesis. The average tooth length was 10.605 for the 0.5 dose, 19.735 for the 1.0 dose, and 26.1 for the 2.0 dose. Thus, higher doses were associated with increased length.**

## Conclusion

**In this study, my analysis suggests that increasing doses of Vitamin C increased the length of guinea pig teeth irrespective of delivery method.**