# The Effect of Vitamin C on Tooth Growth in Guinea Pigs

*Ajay Ghanti*

*11/6/2016*

## Overview

In this project, we do an analysis of the ToothGrowth data in the R datasets package. The goal of the analysis is to accomplish the following.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Establish the key assumptions we make about the data
3. Perform hypothesis tests to compare tooth growth by supplement and dose
4. State the conclusions based on our findings

## Exploratory Data Analysis

Lets take a look at the data summary.

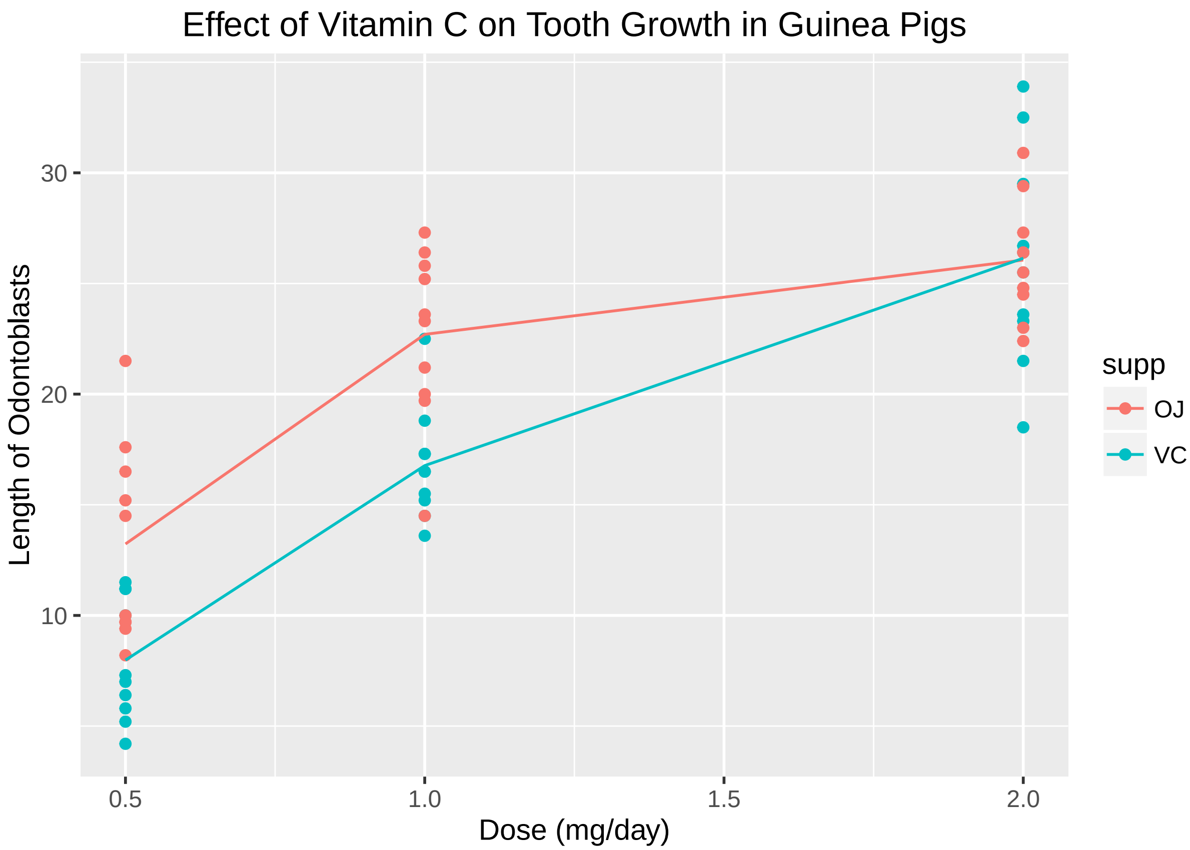
```
str(ToothGrowth)
```

```
## '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 ...
```

As can be seen, there are 60 observations, i.e., 60 guinea pigs which received Vitamin C supplements. 30 pigs received orange juice (code OJ) as the supplement (variable name: 'supp'), whereas the other 30 received ascorbic acid (code VC). Within each group of 30 pigs, 10 each received different doses of Vitamin C - 0.5, 1.0, and 2.0 mg/day (variable name: 'dose'). The response is the length of odontoblasts (cells responsible for tooth growth) in these 60 guinea pigs (variable name: 'len').

### Plot 1

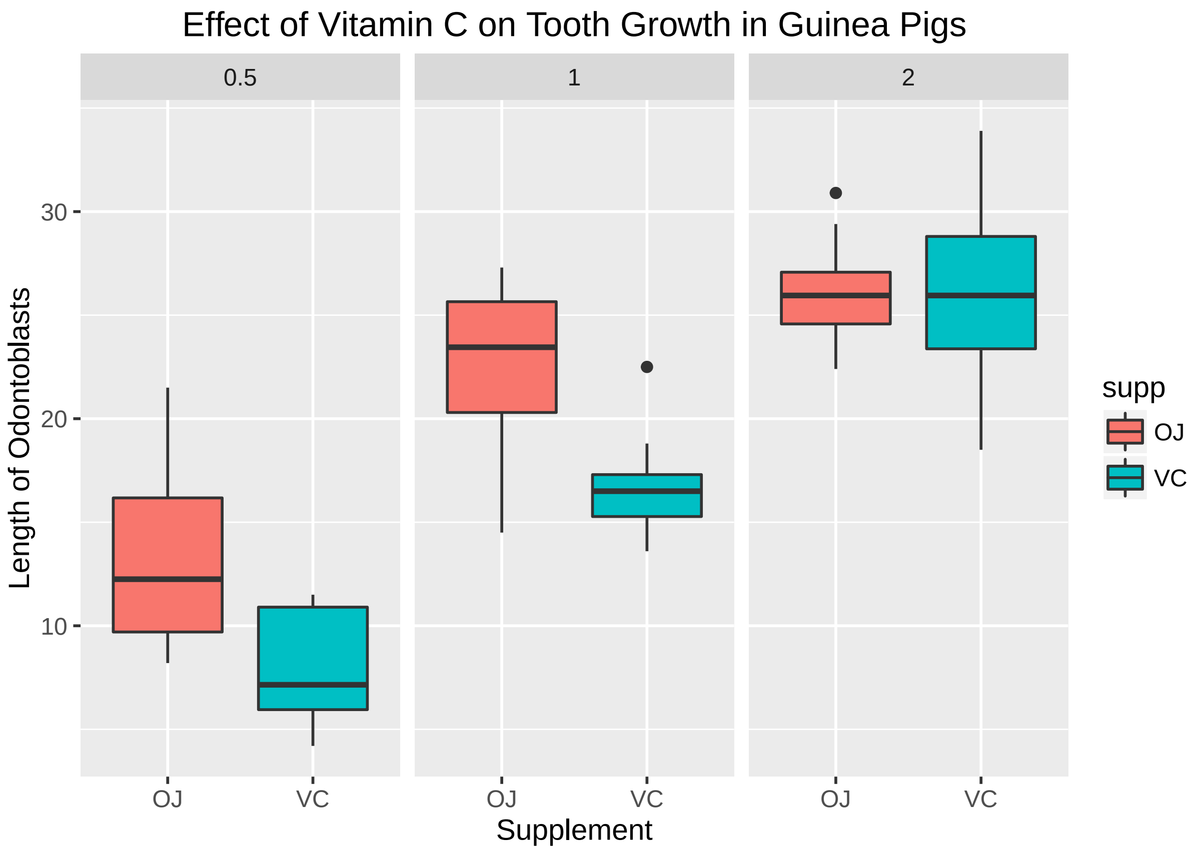
```
ggplot(ToothGrowth, aes(dose, len, group = supp)) +
    geom_point(aes(color=supp)) +
    stat_summary(fun.y=mean, geom="line", aes(group=supp, color=supp)) +
    labs(x = "Dose (mg/day)") + labs(y = "Length of Odontoblasts") +
    labs(title = "Effect of Vitamin C on Tooth Growth in Guinea Pigs")
```




The above scatterplot shows the the effect of dose on the length, for each supplement. The lines connect the averages for each dose. There appears to be an increase in length of tooth growth when the dosage increases, irrespective of the supplement.

#### Plot 2

```
ggplot(ToothGrowth, aes(supp, len, group = supp)) +
    facet_grid(~dose) + geom_boxplot(aes(fill = supp)) +
    labs(x = "Supplement") + labs(y = "Length of Odontoblasts") +
    labs(title = "Effect of Vitamin C on Tooth Growth in Guinea Pigs")
```




The boxplot shown above compares the effect of each supplement on tooth growth, for each dose. At the outset, orange juice seems to be more effective than ascorbic acid at lower doses, and at a higher dose, there doesnt seem to be any difference between the supplements.

## Key Assumptions

We will split the dataset equally into 6 groups of 10 pigs each, based on the dose (20 obs each) and the supplement type (10 obs each for each dose). So, for each dose, 10 pigs received orange juice and another 10 received ascorbic acid.

```
ojDose1 <- with(ToothGrowth, ToothGrowth[which(supp=='OJ' & dose==0.5), ]$len)
ojDose2 <- with(ToothGrowth, ToothGrowth[which(supp=='OJ' & dose==1.0), ]$len)
ojDose3 <- with(ToothGrowth, ToothGrowth[which(supp=='OJ' & dose==2.0), ]$len)
vcDose1 <- with(ToothGrowth, ToothGrowth[which(supp=='VC' & dose==0.5), ]$len)
vcDose2 <- with(ToothGrowth, ToothGrowth[which(supp=='VC' & dose==1.0), ]$len)
vcDose3 <- with(ToothGrowth, ToothGrowth[which(supp=='VC' & dose==2.0), ]$len)
```

1. Since it is known that the 60 guinea pigs used in the study were all distinct, the samples for each of the tests are assumed independent and random
2. So for each dose (0.5, 1.0, & 2.0), we will have two groups of 10 guinea pigs each
3. Since we plan to conduct a two-sided t-test, it is OK to have a small sample size (n=10)
4. We also assume that the variances of the two groups for each dose are unequal
5. We will determine the 95% Confidence Interval to test our hypotheses
6. If the p-value $< 0.05$, we reject the null hypothesis, else we will fail to reject the null hypothesis

## Hypothesis Tests

Let us state our null hypothesis ($H_0$) and alternate hypotheses ($H_a$)

***$H_0$: $X_{OJ}$ - $X_{VC}$ = 0***
There is NO difference in means for each supplement type (orange juice vs. ascorbic acid), i.e., orange juice and ascorbic acid both lead to similar effects in tooth growth

***$H_a$: $X_{OJ}$ - $X_{VC}$ != 0***
There is a difference in means for each supplement type (orange juice vs. ascorbic acid), i.e., orange juice and ascorbic acid DO NOT lead to similar effects in tooth growth

```
# Conduct two-sided t-tests on all three sets of data
D1 <- t.test(ojDose1-vcDose1, paired=FALSE, var.equal=FALSE)
D2 <- t.test(ojDose2-vcDose2, paired=FALSE, var.equal=FALSE)
D3 <- t.test(ojDose3-vcDose3, paired=FALSE, var.equal=FALSE)
# Round off the 95% CIs
ciD1 <- round(D1$conf.int, 3)
ciD2 <- round(D2$conf.int, 3)
ciD3 <- round(D3$conf.int, 3)
# Round of the p-values
pD1 <- round(D1$p.value, 3)
pD2 <- round(D2$p.value, 3)
pD3 <- round(D3$p.value, 3)
```

| Dose (mg/day) | 95% Confidence Interval | P-Value | Outcome |
|---|---|---|---|
| 0.5 | [1.263, 9.237] | 0.015 | Reject $H_0$ |
| 1.0 | [1.952, 9.908] | 0.008 | Reject $H_0$ |
| 2.0 | [-4.329, 4.169] | 0.967 | Fail to Reject $H_0$ |

As seen from the table above, for the doses of 0.5mg/day and 1.0mg/day, since the p-value is $< 0.05$, we reject the null hypothesis $H_0$, which means there was a difference in tooth growth for each supplement type. For the dose of 2.0 mg/day, the p-value is 0.967, hence we fail to reject the null hypothesis $H_0$, which means there was no noticeable difference on which supplement had more effect.

We subtracted the means as **Orange Juice - Ascorbic Acid**. By looking at the 95% confidence intervals, we can figure out which of the supplements had a positive effect on tooth growth. For the doses of 0.5mg/day and 1.0mg/day, the lower and upper values of the CI are positive, confirming that orange juice is more effective at these doses. For the dose of 2.0 mg/day, the lower and upper values of the CI are similar (-/+) and 0 falls in the interval, indicating that neither supplements have a higher efficacy.

## Conclusions

1. Increase in dosage has higher effect on tooth growth
2. At the dosages of 0.5mg/day and 1.0mg/day, orange juice has higher efficacy than ascorbic acid
3. At the dosage of 2.0mg/day, both supplements seem to have similar efficacy