

# Statistical Analysis of the ToothGrowth Data

*Arun K Viswanathan*

*January 17, 2015*

## Overview

The report analyzes the ToothGrowth dataset in R.

## Summary of the Tooth Growth dataset

The *Tooth Growth* dataset in R shows the effect of Vitamin C on tooth growth in guinea pigs. The data provides is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

The data contains three variables:

Variable	Data Type	Description
len	numeric	Tooth length
supp	factor	Supplement type (VC or OJ)
dose	numeric	Dose in milligrams

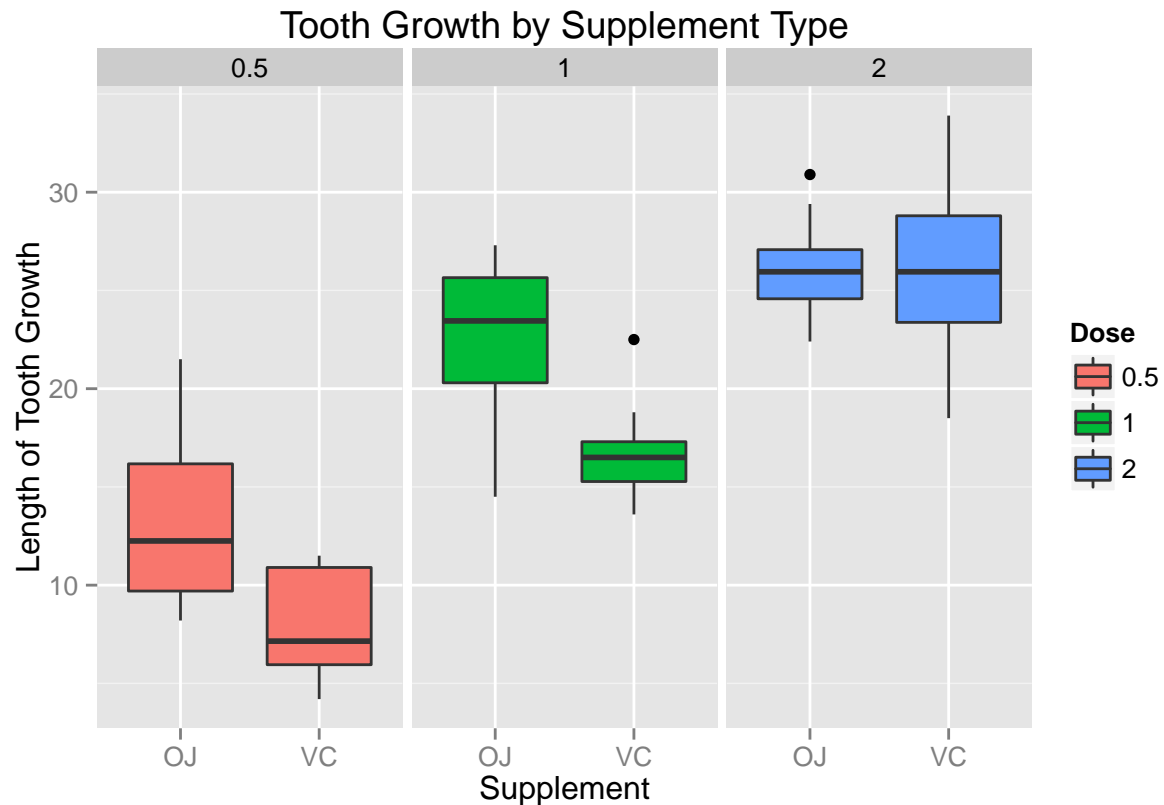
The dataset contains a total of 60 observations. Here's a summary of the raw data.

```
data <- transform(ToothGrowth, dose = as.factor(dose))
summary(data)
```

```
##      len      supp      dose
## Min.   : 4.20   OJ:30   0.5:20
## 1st Qu.:13.07   VC:30    1 :20
## Median :19.25           2 :20
## Mean   :18.81
## 3rd Qu.:25.27
## Max.   :33.90
```

The plot below shows the tooth growth ranges for each supplement at each dosage level.

```
library(ggplot2)
g <- ggplot(data, aes(x = supp, y = len, fill = dose)) +
  geom_boxplot() +
  facet_grid(. ~ dose) +
  ggtitle("Tooth Growth by Supplement Type") +
  xlab("Supplement") +
  ylab("Length of Tooth Growth") +
  guides(fill = guide_legend(title = "Dose"))
print(g)
```



## Tooth growth analysis by supplement

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

## Tooth growth analysis by dose

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

Did the student perform some relevant confidence intervals and/or tests? Were the results of the tests and/or intervals interpreted in the context of the problem correctly? Did the student describe the assumptions needed for their conclusions?