

# Course Project for Statistical Inference Course

## Basic Inferential Data Analysis

By

PREETIKA SRIVASTAVA(04 Aug 2017)

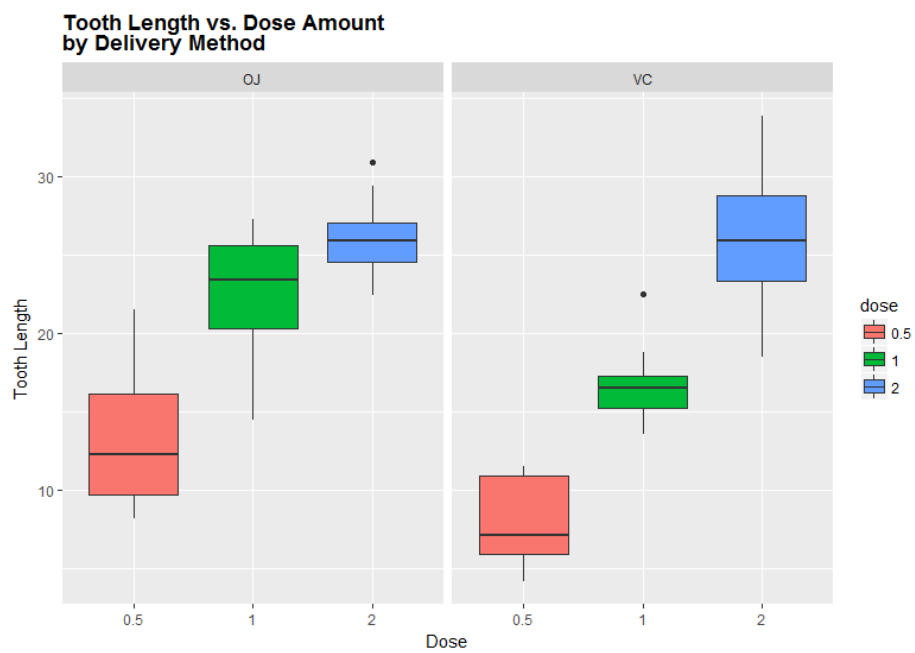
### 1. Exploratory Analysis

```
> data("ToothGrowth")
> 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
> View(ToothGrowth)
> 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
> unique(ToothGrowth$supp)
[1] VC OJ
Levels: OJ VC
> unique(ToothGrowth$dose)
[1] 0.5 1.0 2.0
> ToothGrowth$dose = as.factor(ToothGrowth$dose)
```

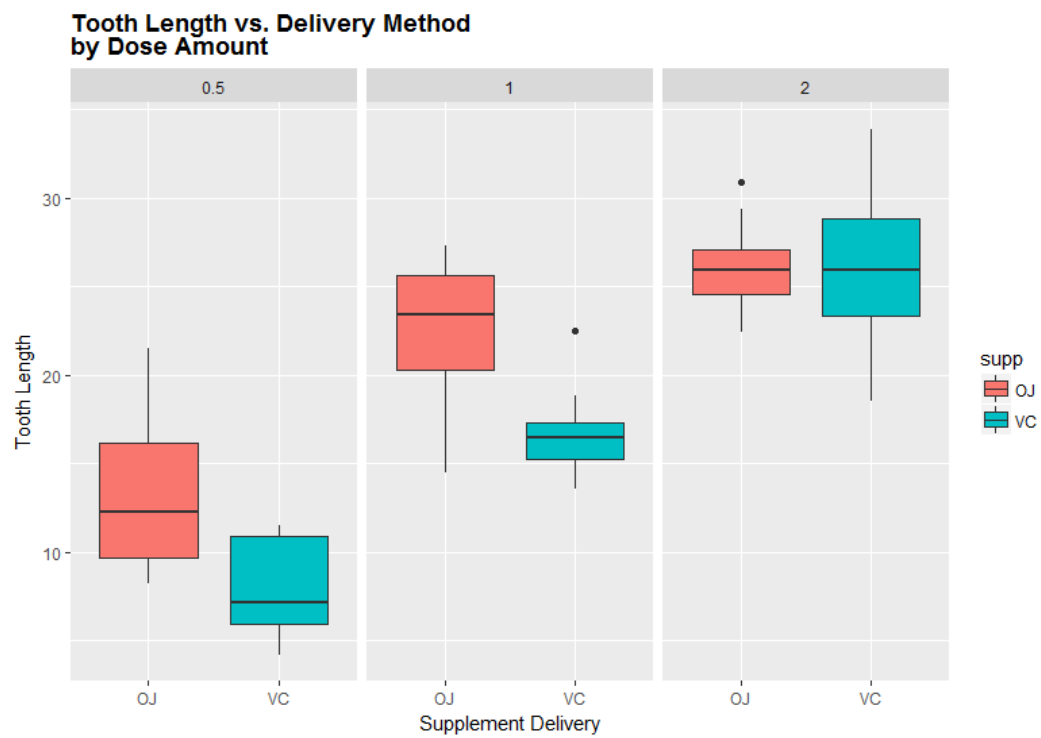
### Box Plot Analysis for better visualization of data

A. Len vs. Dose amount

```
> library(ggplot2)
warning message:
package 'ggplot2' was built under R version 3.3.2
> ggplot(aes(x=dose, y=len), data=ToothGrowth) + geom_boxplot(aes(fill=dose)) + xlab("Dose") + ylab("Tooth Length") + facet_grid(~ supp) + ggtitle("Tooth Length vs. Dose Amount \nby Delivery Method") +
+   theme(plot.title = element_text(lineheight=.8, face="bold"))
```



## 2. Len vs Supplement delivery method



## 3. Comparing Tooth Growth using t-Test

```
> t.test(len~supp,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
```

Here, it is evident that P-value is greater than 0.05, and confidence interval comprises value 0 in it, so we can say that Supplement type doesn't seem to show any growth on Tooth Growth length.

Now, comparing the length with dose values:

#### A. Comparing with 1.0 and 2.0 dose value

```
> t.test(len~dose,data=ToothGrowth_sub_1N2)
```

```
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
```

#### B. Comparing with 1.0 and 0.5 dose value

```
> ToothGrowth_sub_1N5 = subset(ToothGrowth, ToothGrowth$dose %in% c(1.0,0.5))
> t.test(len~dose,data=ToothGrowth_sub_1N5)
```

```
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

```

### C. Comparing with 2.0 and 0.5 dose value

```

> ToothGrowth_sub_1N2 = subset(ToothGrowth, ToothGrowth$dose %in% c(1.0,2.
0))
> t.test(len~dose,data=ToothGrowth_sub_1N2)

```

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

```

It is evident here that all three P-Values are essentially zero and none of the confidence interval constitutes zero in it, so Null Hypothesis can be rejected.

Hence, we can conclude that the average tooth length increases with an increasing dose, and therefore the null hypothesis can be rejected.