

# Statistical Inference Project Part02

*Sandeep Anand*

*March 27, 2017*

## Part-2 Of the Statistical Inference Project:

- Analyze the ToothGrowth data in the R datasets package
- 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 the conclusions

```
data(ToothGrowth)
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 ...

# Summary
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
```

## Description of analysis with Graphs

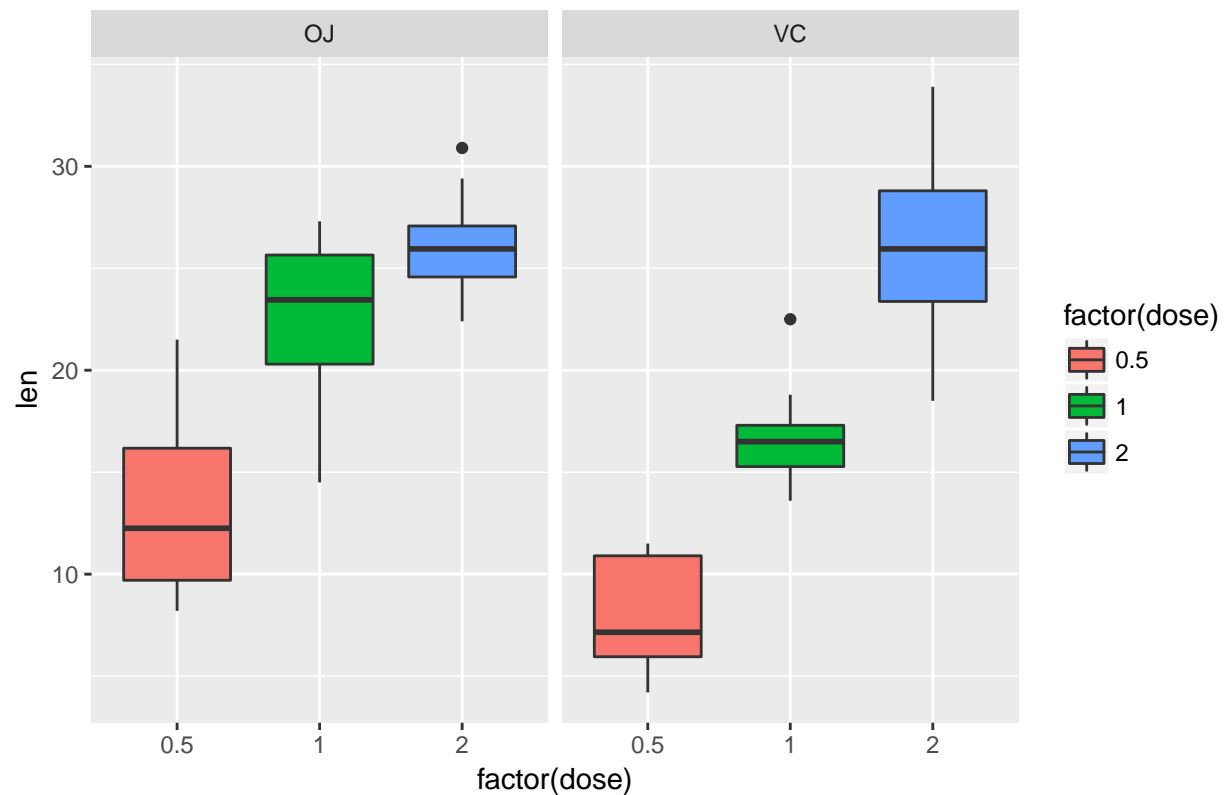
Toothgrowth comparision with respect to two groups OJ and VC

```
library(ggplot2)

g4<-ggplot(ToothGrowth, aes(x=factor(dose), y=len, fill=factor(dose))) +
  geom_boxplot()+facet_grid(.~supp)+
  ggtitle("Analyzing Toothgrowth data based on Type of Dose")

print(g4)
```

## Analyzing Toothgrowth data based on Type of Dose



## Toothgrowth comparison with averages

```
avg<-aggregate(ToothGrowth$len,by=list(dose=ToothGrowth$dose,supp=ToothGrowth$supp),FUN=mean)
```

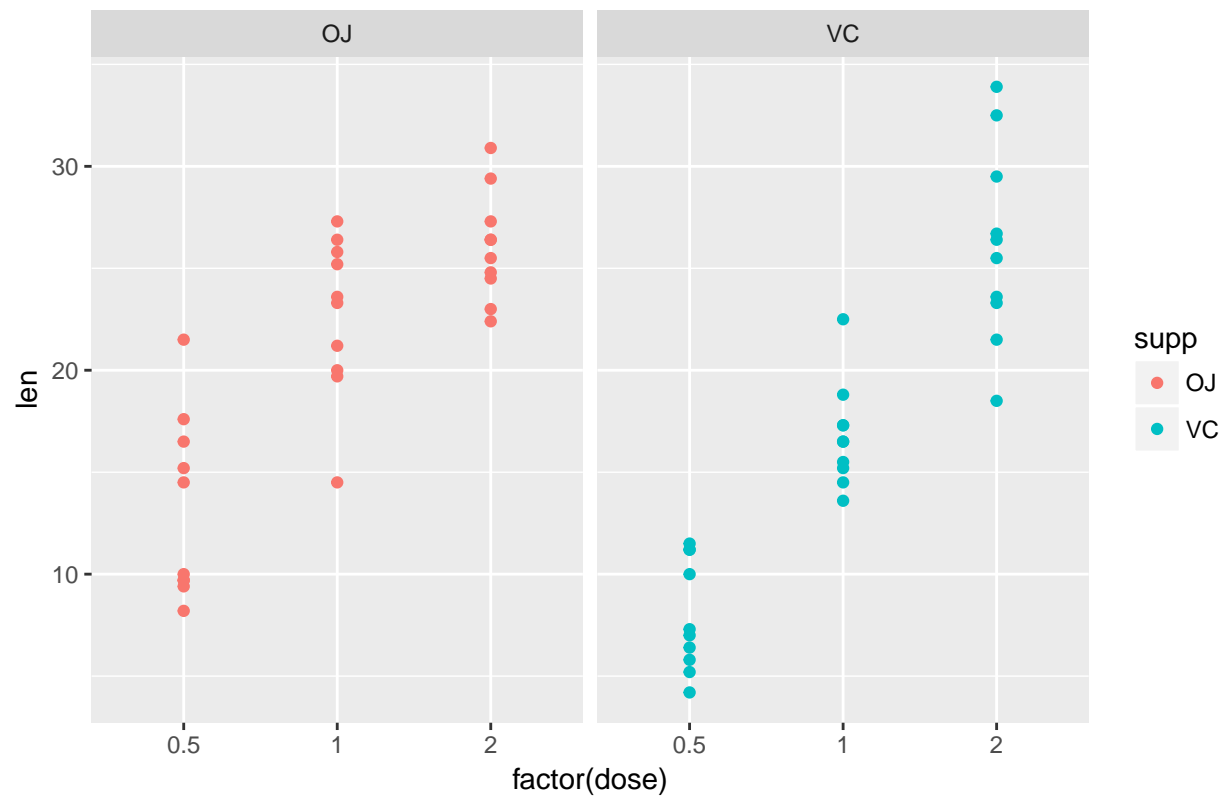
```
avg
```

```
##   dose supp      x
## 1  0.5   OJ 13.23
## 2  1.0   OJ 22.70
## 3  2.0   OJ 26.06
## 4  0.5   VC  7.98
## 5  1.0   VC 16.77
## 6  2.0   VC 26.14
```

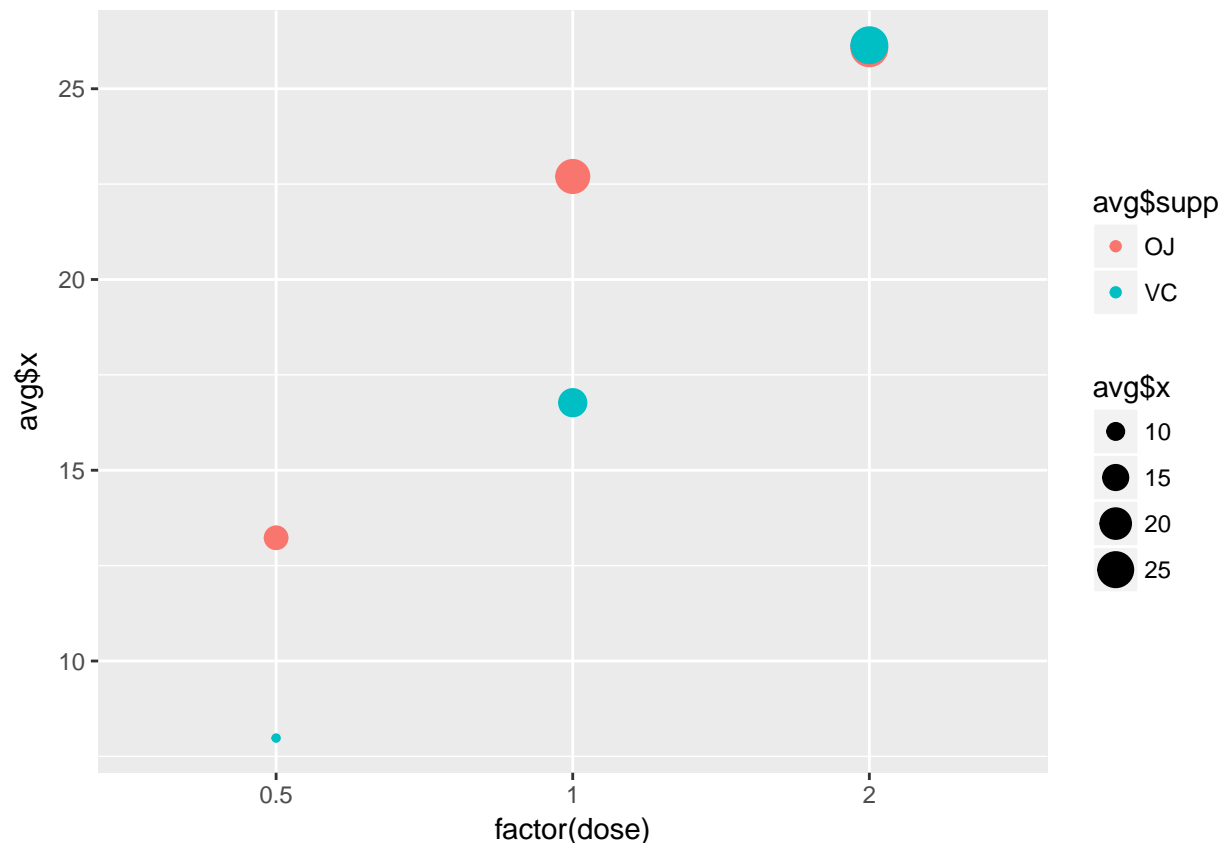
```
g5<-ggplot(ToothGrowth, aes(x=factor(dose), y=len)) +
  geom_point(aes(color=supp))+facet_grid(.~supp)+
  ggtitle("Analyzing Toothgrowth data trends based on Dosage")
```

```
print(g5)
```

## Analyzing Toothgrowth data trends based on Dosage



```
g<- ggplot(avg, aes(x=factor(dose), y=avg$x, color=avg$supp))+  
  geom_point(aes(size=avg$x))  
print(g)
```



Using Confidence Intervals and hypothesis tests to compare tooth growth by Supplement and Dosage.

```
print(t.test(len~supp,data = ToothGrowth, paired = FALSE))

##
## 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
```

## Conclusions

The boxplots show that the growth is faster in case of dose = VC than dose = OJ . The 95% Confidence interval is got as (-0.171,7.571)

As tooth size increases , dosage is also higher