# part2

*Saurabh Yadav*

# PART-2 Basic inferential data analysis

## Loading the tooth data

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

```
## Warning: package 'ggplot2' was built under R version 3.5.1
```

## Subsetting the tooth growth data on the basis of the supp

```
vc<-subset(ToothGrowth,supp=="VC")
oj<-subset(ToothGrowth,supp=="OJ")
```

## Arranging the datasets on the basis of doses

```
newvc<-c()
newvc$dose0.5<-vc$len[vc$dose==0.5]
newvc$dose1<-vc$len[vc$dose==1]
newvc$dose2<-vc$len[vc$dose==2]
newvc<-data.frame(newvc)
newoj<-c()
newoj$dose0.5<-oj$len[oj$dose==0.5]
newoj$dose1<-oj$len[oj$dose==1]
newoj$dose2<-oj$len[oj$dose==2]
newoj<-data.frame(newoj)
```

## Taking the mean of the length of the tooth growth for different doses

```
lengthmeanvc<-apply(newvc, 2, mean)
lengthmeanoj<-apply(newoj, 2, mean)
```

## converting it into a marix

```
lengthmeanvc<-data.frame(matrix(lengthmeanvc))
lengthmeanoj<-data.frame(matrix(lengthmeanoj))
```

## giving new column names

```
colnames(lengthmeanvc)<-"len"
rownames(lengthmeanvc)<-c("dose0.5","dose1","dose2")
colnames(lengthmeanoj)<-"len"
rownames(lengthmeanoj)<-c("dose0.5","dose1","dose2")
```

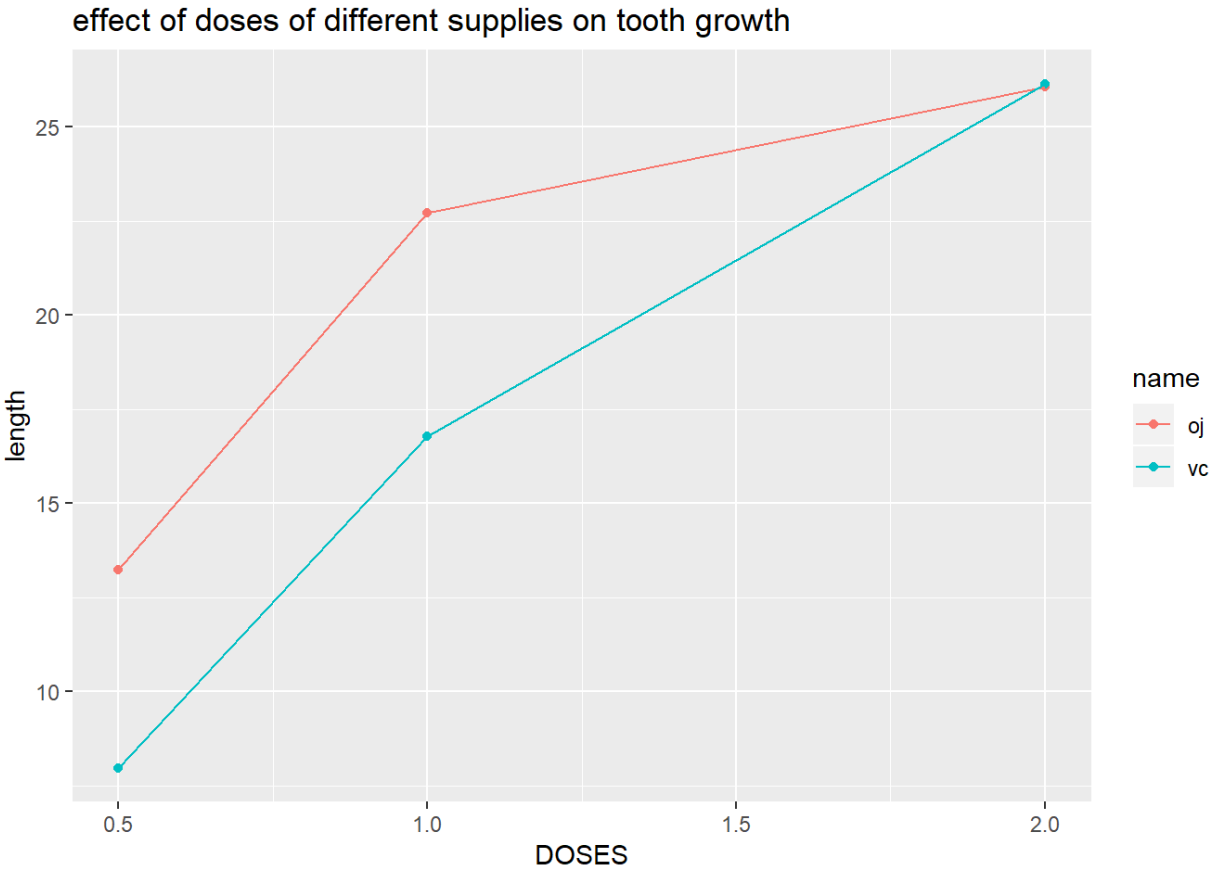
## adding new coloumn

```
lengthmeanvc$dose<-c(0.5,1,2)
lengthmeanvc$name<-c("vc","vc","vc")
lengthmeanoj$dose<-c(0.5,1,2)
```

```
lengthmeanoj$name<-c("oj","oj","oj")
finaldata<-rbind(lengthmeanoj,lengthmeanvc)
```

plotting

```
ggplot(finaldata,aes(x=dose,y=len,color=name))+geom_point()+geom_line()+labs(title="effect of doses of different supplies on tooth growth")+labs(x="DOSES",y="length")
```




```
len<-lengthmeanvc
colnames(len)<-c("lenVC","dose")
len$lenOJ<-lengthmeanoj$len
```

Using confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose

# 1. Confidence interval for dose 0.5

```
t.test(oj$len[oj$dose==0.5],vc$len[vc$dose==0.5],paired = F)$conf.int
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

RESULT: The t test between the supp vc and oj for dose=0.5 gives the interval 1.719057 to 8.780943. Which means the dose of 0.5 for oj is way more effective for tooth growth.

# 2. Confidence interval for dose 1

```
t.test(oj$len[oj$dose==1],vc$len[vc$dose==1],paired = F)$conf.int
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

RESULT: The t test between the supp vc and oj for dose=1 gives the interval 2.802148 to 9.057852. Which means the dose of 1 for oj is way more effective for tooth growth.

## 3. Confidence interval for dose 2

```
t.test(oj$len[oj$dose==2],vc$len[vc$dose==2],paired = F)$conf.int
```

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95
```

RESULT: The t test between the supp vc and oj for dose=2 gives the interval -3.79807 to 3.63807. Which means the dose of 2 for both vc and oj are approximately same. This can also be shown by the plot which clearly shows that dose 2 for both affect the tooth growth in a same way.

## Conclusion

1. The effect of two medications are similar when the dose is 2.
2. The oj is more effective in tooth growth when the doses are 0.5 and 1