

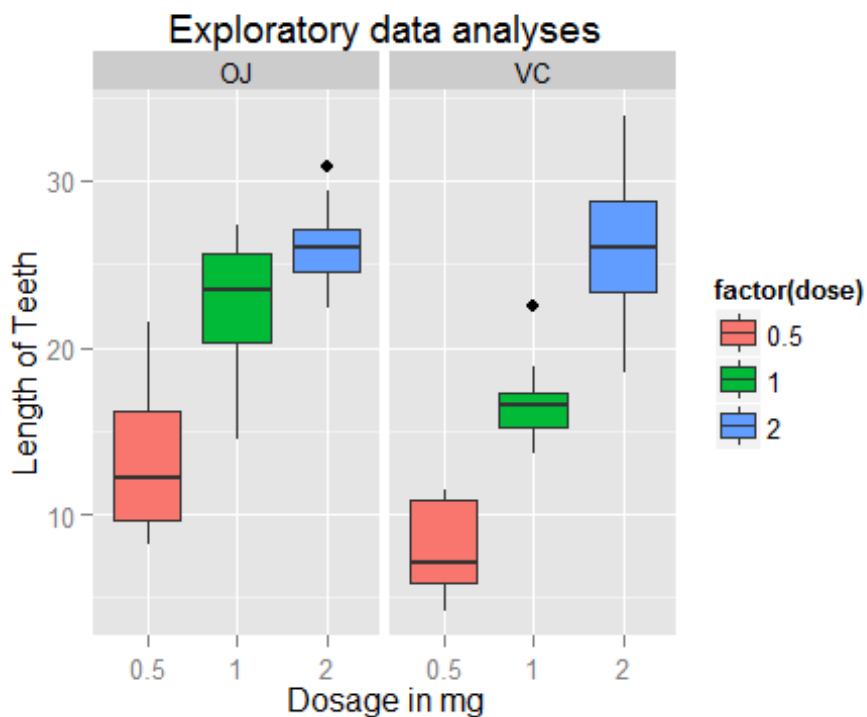
project, Question 2

1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(ggplot2)
library(datasets)
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

attach(ToothGrowth)
g1 <- ggplot(ToothGrowth, aes(x=factor(dose),y=len,fill=factor(dose)))
g1 + geom_boxplot(notch=F) + facet_grid(.~supp) +
  scale_x_discrete("Dosage in mg") +
  scale_y_continuous("Length of Teeth") +
  ggtitle("Exploratory data analyses")
```



2. Provide a basic summary of the data.

```
dose <- as.factor(dose)
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

table(supp, dose)

##      dose
## supp 0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

3. Use confidence intervals and hypothesis tests to compare tooth growth by supp and dose. (Use the techniques from class even if there's other approaches worth considering)

```
test.supp.equ <- t.test(len~supp, paired=F, var.equal=T, data=ToothGrowth)
test.supp.nequ <- t.test(len~supp, paired=F, var.equal=F, data=ToothGrowth)

data.frame("p-value"=c(test.supp.equ$p.value, test.supp.nequ$p.value), "Conf-
Low"=c(test.supp.equ$conf[1], test.supp.nequ$conf[1]),
           "Conf-
High"=c(test.supp.equ$conf[2], test.supp.nequ$conf[2]), row.names=c("Equal
Var", "Unequal Var"))

##           p.value  Conf.Low Conf.High
## Equal Var  0.06039337 -0.1670064  7.567006
## Unequal Var 0.06063451 -0.1710156  7.571016
```

The result shows that there is no difference between two supplement groups because the p value is more than .05 and point zero is included in the interval confidence range.

```
dose.05 <- ToothGrowth[which(dose==.5),1]
dose.10 <- ToothGrowth[which(dose==1),1]
dose.20 <- ToothGrowth[which(dose==2),1]
```

Test between dose=0.5 and dose=1

```
dose0510.t1 <- t.test(dose.05, dose.10, paired=F, var.equal=T)
dose0510.t2 <- t.test(dose.05, dose.10, paired=F, var.equal=F)
dose0510.result <- data.frame("p-value"=c(dose0510.t1$p.value,
dose0510.t2$p.value), "Conf-
Low"=c(dose0510.t1$conf[1], dose0510.t2$conf[1]), "Conf-
High"=c(dose0510.t1$conf[2], dose0510.t2$conf[2]), row.names=c("Equal
Var", "Unequal Var"), "Dose"="0.5 to 1")
```

test between dose=0.5 and dose=2

```
dose0520.t1 <- t.test(dose.05, dose.20, paired=F, var.equal=T)
dose0520.t2 <- t.test(dose.05, dose.20, paired=F, var.equal=F)
dose0520.result <- data.frame("p-value"=c(dose0520.t1$p.value,
dose0520.t2$p.value), "Conf-
Low"=c(dose0520.t1$conf[1],dose0520.t2$conf[1]), "Conf-
High"=c(dose0520.t1$conf[2],dose0520.t2$conf[2]), row.names=c("Equal
Var", "Unequal Var"), "Dose"="0.5 to 2")
```

test between dose=2 and dose=1

```
dose1020.t1 <- t.test(dose.10, dose.20, paired=F, var.equal=T)
dose1020.t2 <- t.test(dose.10, dose.20, paired=F, var.equal=F)
dose1020.result <- data.frame("p-value"=c(dose1020.t1$p.value,
dose1020.t2$p.value), "Conf-
Low"=c(dose1020.t1$conf[1],dose1020.t2$conf[1]), "Conf-
High"=c(dose1020.t1$conf[2],dose1020.t2$conf[2]), row.names=c("Equal
Var", "Unequal Var"), "Dose"="1 to 2")
rbind(dose0510.result,dose0520.result,dose1020.result)

##                p.value   Conf.Low  Conf.High      Dose
## Equal Var      1.266297e-07 -11.983748  -6.276252 0.5 to 1
## Unequal Var    1.268301e-07 -11.983781  -6.276219 0.5 to 1
## Equal Var1     2.837553e-14 -18.153519 -12.836481 0.5 to 2
## Unequal Var1   4.397525e-14 -18.156167 -12.833833 0.5 to 2
## Equal Var2     1.810829e-05  -8.994387  -3.735613  1 to 2
## Unequal Var2   1.906430e-05  -8.996481  -3.733519  1 to 2
```

the result show that the alternative hypothesis is accepted because P values are less than .05 and confidence intervals do not contain point zero. Therefore, it is concluded that groups with higher dosage have higher tooth length.

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

It is concluded that: 1) The group with dose equal 2 has larger impact on tooth length than other doses, 2) The group with dose equal 1 has higher tooth length than dose 0.5 . 3) There is no difference between two supplements of OJ and VC with respect to the impact they have on tooth length.