Statistical Inference Project

Part: 2

Albion Dervishi

March 21, 2015

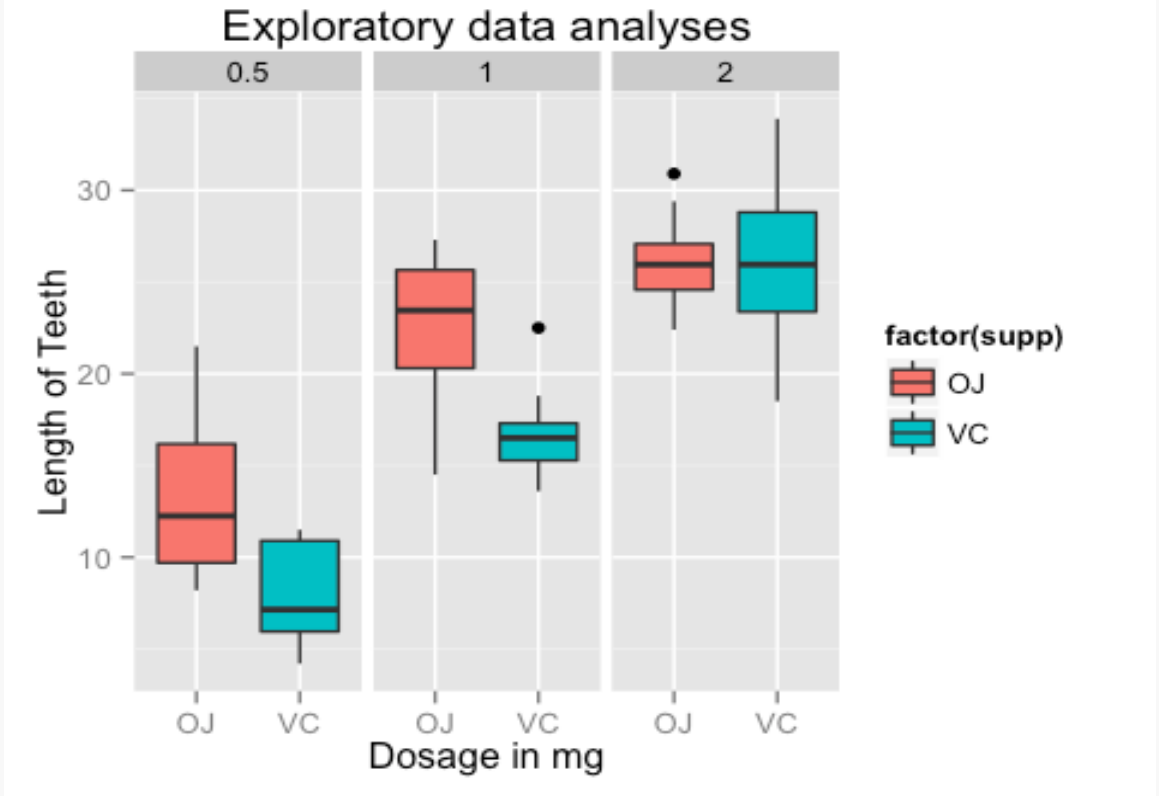
Question 1-Load the ToothGrowth data and perform some basic exploratory data analyses

```
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
```

```
library(ggplot2)
```

```
g1 <-ggplot(ToothGrowth, aes(x=factor(supp),y=len, fill=factor(supp)))
g1 + geom_boxplot(width=1) + facet_grid(.~dose) +
  scale_x_discrete("Dosage in mg") +
  scale_y_continuous("Length of Teeth") +
  ggtitle("Explorat ory data analyses")
```



Question 2-Provide a basic summary of the data.

```
ToothGrowth$dose<-as.factor(ToothGrowth$dose)
summary(ToothGrowth)
```

```
##       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
```

Question 3- 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)

```
d1<- subset(ToothGrowth, subset = ToothGrowth$dose=="1")
d2<- subset(ToothGrowth, subset = ToothGrowth$dose=="2")
d0.5<- subset(ToothGrowth, subset = ToothGrowth$dose=="0.5")
OJ<- subset(ToothGrowth, subset = ToothGrowth$supp== "OJ")
VC<- subset(ToothGrowth, subset= ToothGrowth$supp=="VC")
```

I will run a t-test comparing the difference in tooth length between group d0.5 and d1

```
t.test(d1-d2)
##  One Sample t-test
##
## data:  d1 - d2
## t = -4.6046, df = 19, p-value = 0.0001934
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -9.258186 -3.471814
## sample estimates:
## mean of x
##    -6.365
```

As per results confidence interval does not contain zero. we have enough evidence to reject H0

```
t.test(d0.5-d2)
##  One Sample t-test
##
## data:  d0.5 - d2
## t = -11.2915, df = 19, p-value = 7.19e-10
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -18.3672 -12.6228
## sample estimates:
## mean of x
##   -15.495
```

Also confidence interval does not contain zero, although the p-value less than first and hereby it is caution to reject the null hypothesis.

```
t.test(d0.5-d1)
##  One Sample t-test
##
## data:  d0.5 - d1
## t = -6.9669, df = 19, p-value = 1.225e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -11.872879  -6.387121
## sample estimates:
## mean of x
##     -9.13
```

As per confidence interval in this analysis difference between the two treatment groups is not zero, we have enough evidence to reject H0

```
t.test(OJ-VC)
##  One Sample t-test
##
## data:  OJ - VC
## t = 3.3026, df = 29, p-value = 0.00255
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.408659 5.991341
## sample estimates:
## mean of x
##       3.7
```

As per confidence interval in this analysis difference between the two treatment groups is not zero. I reject null hypothesis.

Question 4-Conclusion:

There is a positive correlation between the teeth length and the dose levels of Vitamin C. Rejection of null hypothesis in group d0.5-d2(dose 0.5 and does 2) have positive influence on this correlation. Rejection of null hypothesis in the group OJ-VC(Orange Juice and Vitamin C) have clear output of the results over this correlation on teeth growth.