

# Statistical Inference - Course Project 2

Author: Danijel Bara

```
## Run time: 2015-08-22 21:37:57
## R version: R version 3.2.0 (2015-04-16)
```

## Overview

In the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

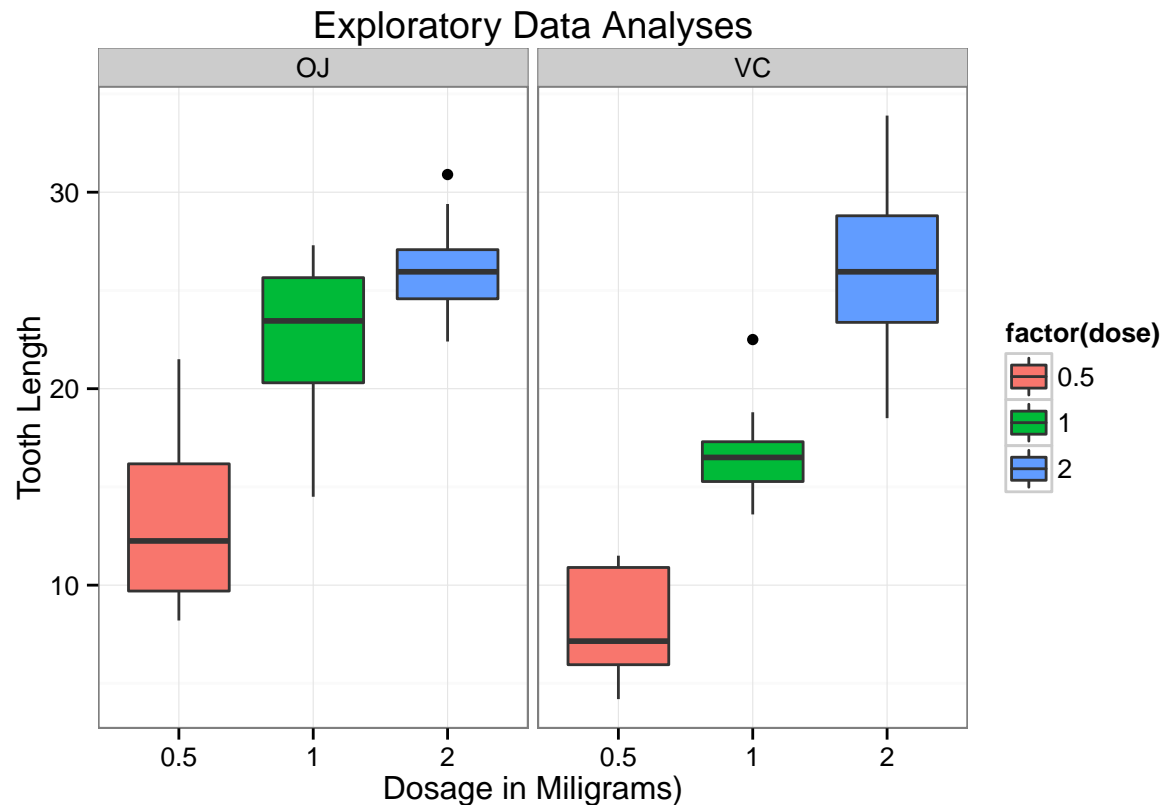
### Question 1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(ggplot2)
library(datasets)
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 ...
```

Description of data, found on web site: <https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/ToothGrowth.html> The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC)). We can analyse the data on plot:

```
plot <- ggplot(ToothGrowth, aes(x=factor(dose),y=len,fill=factor(dose)))
plot + geom_boxplot(notch=F) + facet_grid(.~supp) +
  scale_x_discrete("Dosage in Miligrams") +
  scale_y_continuous("Tooth Length") +
  ggtitle("Exploratory Data Analyses") +
  theme_bw()
```



Question 2. Provide a basic summary of the data.

```
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
```

```
dose<-as.factor(ToothGrowth$dose)
table(ToothGrowth$supp, ToothGrowth$dose)
```

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

The ToothGrowth dataset explains the relation between the growth of teeth of guinea pigs at each of three dose levels of Vitamin C (0.5, 1 and 2 mg) with each of two delivery methods (orange juice and ascorbic acid).

**Question 3.** 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).

```
fit <- lm(len ~ dose + supp, data=ToothGrowth)
confint(fit)
summary(fit)
# T Test by supplement type
t.test(len ~ supp, data = ToothGrowth)
# T test by dose level
data_05.1 <- subset(ToothGrowth, dose %in% c(0.5,1))
data_05.2 <- subset(ToothGrowth, dose %in% c(0.5,2))
data_1.2 <- subset(ToothGrowth, dose %in% c(1,2))
t.test(len ~ dose, paired=F, var.equal=F, data=data_05.1)
t.test(len ~ dose, paired=F, var.equal=F, data=data_05.2)
t.test(len ~ dose, paired=F, var.equal=F, data=data_1.2)
# Testing delivery method
t.test(len ~ supp, paired=F, var.equal=F, data=ToothGrowth[ToothGrowth$dose==0.5,])
t.test(len ~ supp, paired=F, var.equal=F, data=ToothGrowth[ToothGrowth$dose==1,])
t.test(len ~ supp, paired=F, var.equal=F, data=ToothGrowth[ToothGrowth$dose==2,])
```

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

We can conclude, after the analysis that:

1. The higher the dosage, the longer the teeth. When doses are larger, there is different growth of the tooth and 2mg dose has greater impact on tooth growth than those of 1mg and 0.5mg.
2. There is importance in delivery method when dosage is 0.5mg and 1mg, and orange juice and vitamin C have different impact on tooth growth. When dosage is 2mg there is no obvious importance on delivery method.