

# Statistical Inference Course Project Part 2

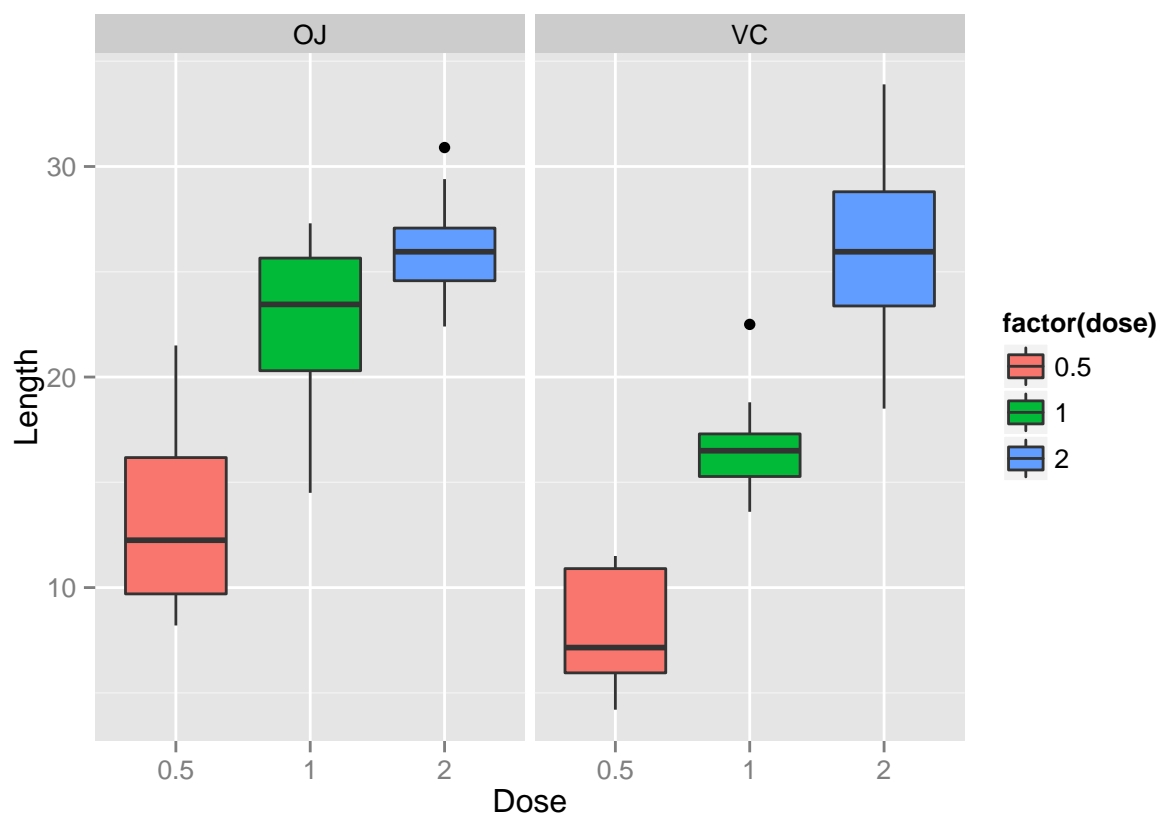
*Sven Koerbitz*

## Overview

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

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

```
library(datasets)
library(ggplot2)
data(ToothGrowth)
plot <- ggplot(ToothGrowth, aes(x=factor(dose), y=len, fill=factor(dose)))
plot + geom_boxplot() +
  facet_grid(.~supp) +
  scale_x_discrete("Dose") +
  scale_y_continuous("Length")
```



### 2. Provide a basic summary of the data.

```
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
```

```
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(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
```

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

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

3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.

```
# T-Test by Supp Type
t.test(len ~ supp, data = ToothGrowth)
```

```
##
## 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
```

```
# Subsetting Data by Dose
dose0.5 = subset(ToothGrowth, dose == 0.5)
dose1.0 = subset(ToothGrowth, dose == 1.0)
dose2.0 = subset(ToothGrowth, dose == 2.0)

# T-Test Supp by Dose
t.test(len~supp, data=dose0.5)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##      13.23      7.98
```

```
t.test(len~supp, data=dose1.0)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##      22.70      16.77
```

```
t.test(len~supp, data=dose2.0)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = -0.0461, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.79807 3.63807
## sample estimates:
## mean in group OJ mean in group VC
##      26.06      26.14
```

Given a dose of 0.5 the null hypothesis that the supplement type (OJ/VC) has no influence on the growth of teeth can be rejected because the p-value is 0.006359 which is way below a significance level of 0.05.

Given a dose of 1.0 the null hypothesis that the supplement type (OJ/VC) has no influence on the growth of teeth can be rejected because the p-value is 0.001038 which is way below a significance level of 0.05.

Given a dose of 2.0 the null hypothesis that the supplement type (OJ/VC) has no influence on the growth of teeth can not be rejected because the p-value is 0.9639 which is above a significance level of 0.05. The data is not conclusive.

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

Based on the basic exploratory data analyses the assumption was made that - the growth of teeth depends on the dose (Higher dose = longer theeth) and - the supplement type OJ has a higher effect on the growth of teeth than the supplement type VC

Based on the hypothesis tests it can be concluded that - the growth of teeth increases with the dose and - the supplement type OJ has a higher effect in the doses of 0.5 and 1.0 but not in doses of 2.0