

# ToothGrowth

Sabank

April 26, 2015

## Exploratory Analysis

```
# load R libraries
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.1.3

library(plyr)
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:plyr':
##
##      arrange, count, desc, failwith, id, mutate, rename, summarise,
##      summarize
##
## The following object is masked from 'package:stats':
##
##      filter
##
## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union

library(reshape2)
library(knitr)
```

### 1. Load data

```
# load data
data(ToothGrowth)
```

## 2. Statistics

```
# summary statistics
```

```
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
```

```
### analysis of tooth growth by supp and dose
```

```
x <- melt(ToothGrowth, id = c("supp", "dose"), measure.vars = "len")
```

```
stats_x <- dcast(x, supp + dose ~ variable, mean)
```

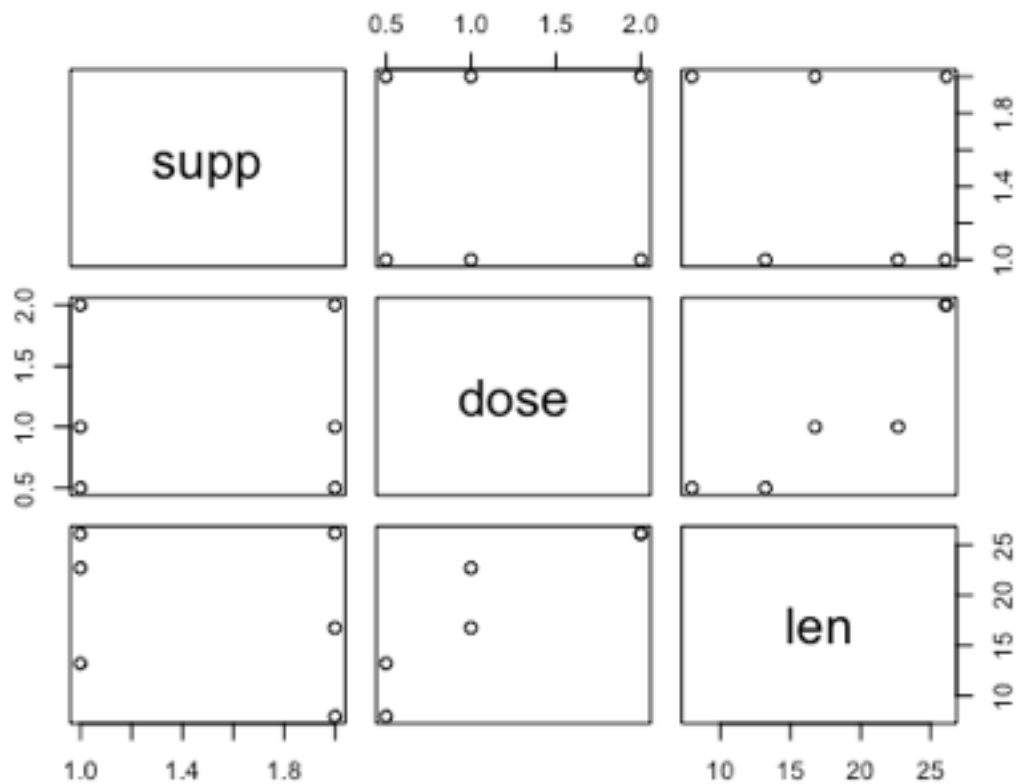
```
kable(head(stats_x), format = "html", align = "c",
       caption = "Average teeth growth")
```

Average teeth growth

supp	dose	len
OJ	0.5	13.23
OJ	1.0	22.70
OJ	2.0	26.06
VC	0.5	7.98
VC	1.0	16.77
VC	2.0	26.14

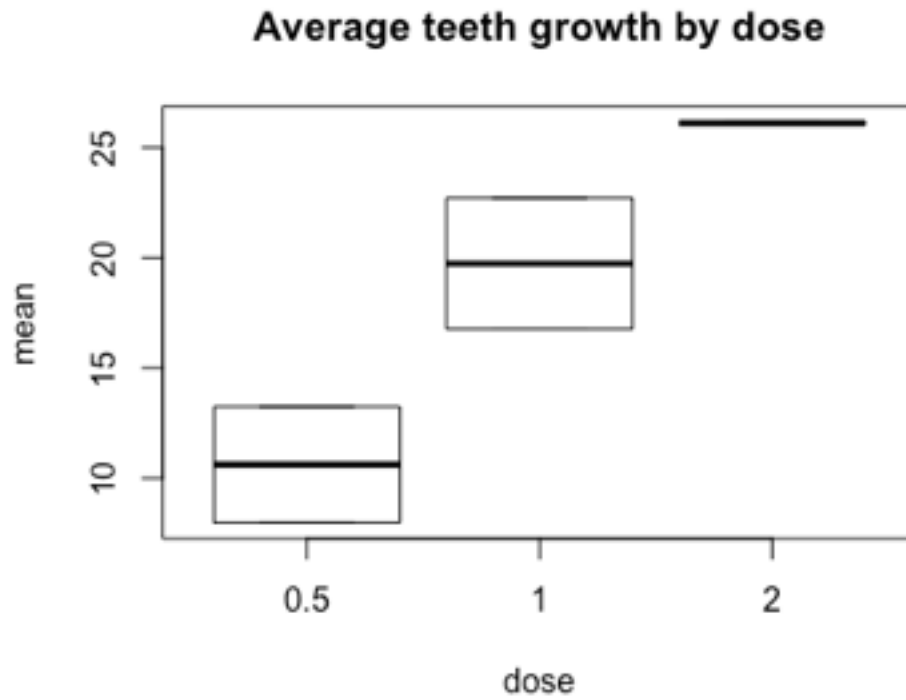
```
## plot stats overview
```

```
plot(stats_x)
```



## 2.1 Analysis of data "dose"

```
boxplot(len ~ dose, data = stats_x, main = "Average teeth growth by
dose",
        xlab = "dose", ylab = "mean")
```



```
d_05 <- filter(stats_x, dose == 0.5)
d_1 <- filter(stats_x, dose == 1)
d_2 <- filter(stats_x, dose == 1)

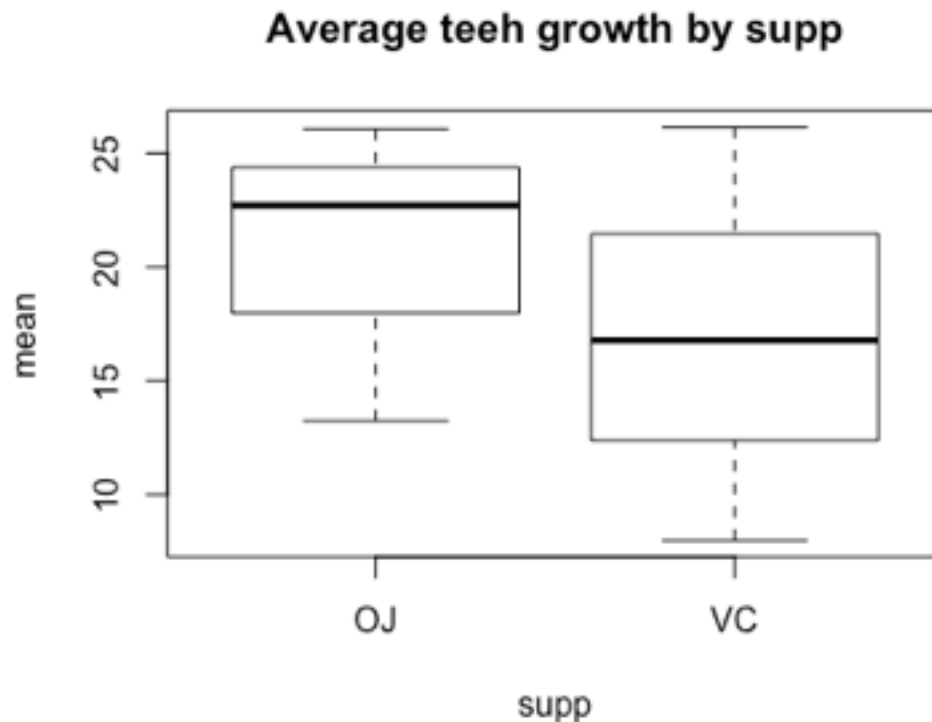
d_05_conf <- (mean(d_05$len) + c(-1,1) * qnorm(.975) * sd(d_05$len)/
sqrt(length(d_05$len)))
d_1_conf <- (mean(d_1$len) + c(-1,1) * qnorm(.975) * sd(d_1$len)/
sqrt(length(d_1$len)))
d_2_conf <- (mean(d_2$len) + c(-1,1) * qnorm(.975) * sd(d_2$len)/
sqrt(length(d_2$len)))
```

Confidence interval 95%:

- dose 0.5 = 5.4600945, 15.7499055
- dose 1 = 13.9237068, 25.5462932
- dose 2 = 13.9237068, 25.5462932

## 2.2 Analysis of data "supp"

```
boxplot(len ~ supp, data = stats_x, main = "Average teeh growth by  
supp",  
        xlab = "supp", ylab = "mean")
```



```
s_oj <- filter(stats_x, supp == "OJ")  
s_vc <- filter(stats_x, supp == "VC")  
  
s_oj_conf <- (mean(s_oj$len) + c(-1,1) * qnorm(.975) * sd(s_oj$len)/  
sqrt(length(s_oj$len)))  
s_vc_conf <- (mean(s_vc$len) + c(-1,1) * qnorm(.975) * sd(s_vc$len)/  
sqrt(length(s_vc$len)))
```

Confidence interval 95%:

- supp OJ = 13.1348233, 28.1918433
- supp VC = 6.6867882, 27.2398785