

Statistical Inference Course Project – Part II

Scott Shepard

October 26, 2014

This is the second of two reports for the Statistical Inference Course Project.

Load the ToothGrowth Dataset and view the first few rows.

```
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
```

```
summary(ToothGrowth)
```

```
##      len      supp      dose
## Min.   : 4.2    OJ:30  Min.   :0.50
## 1st Qu.:13.1   VC:30  1st Qu.:0.50
## Median :19.2           Median :1.00
## Mean   :18.8           Mean   :1.17
## 3rd Qu.:25.3           3rd Qu.:2.00
## Max.   :33.9           Max.   :2.00
```

Do some exploratory analysis on the set.

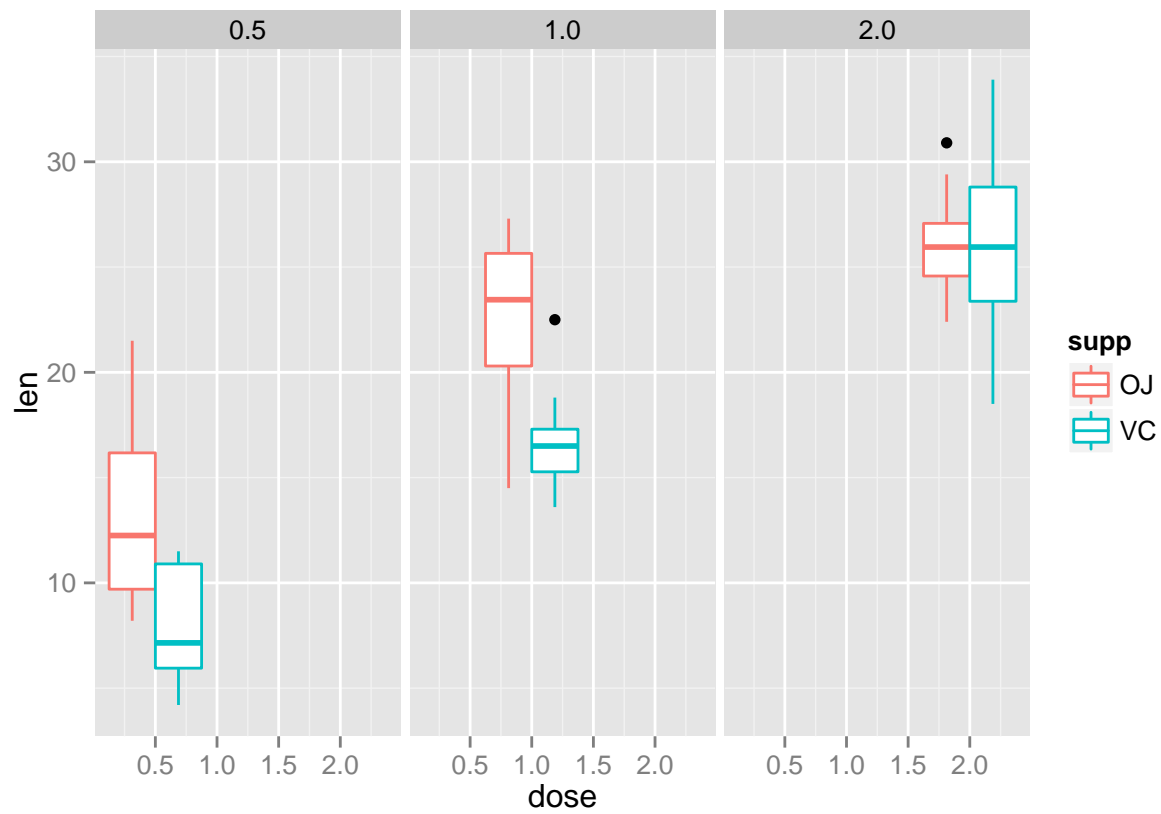
1. What does the distribution of dose by supplement type look like?
2. Does dose of supplement have any correlation with tooth length?

```
with(ToothGrowth, table(supp, dose))
```

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

So there were 10 Guinea Pigs at each dose and supplement.

```
ggplot(ToothGrowth, aes(x=dose, y=len, color=supp)) + geom_boxplot() + facet_wrap(~ dose)
```



So the dose level does have an effect on tooth length. And it seems like OJ supplement has greater tooth lengths at the 0.5 and 1.0 doses but not at the 2.0 level.