Peer Assessment on ToothGrowth data - Data Science Specialization
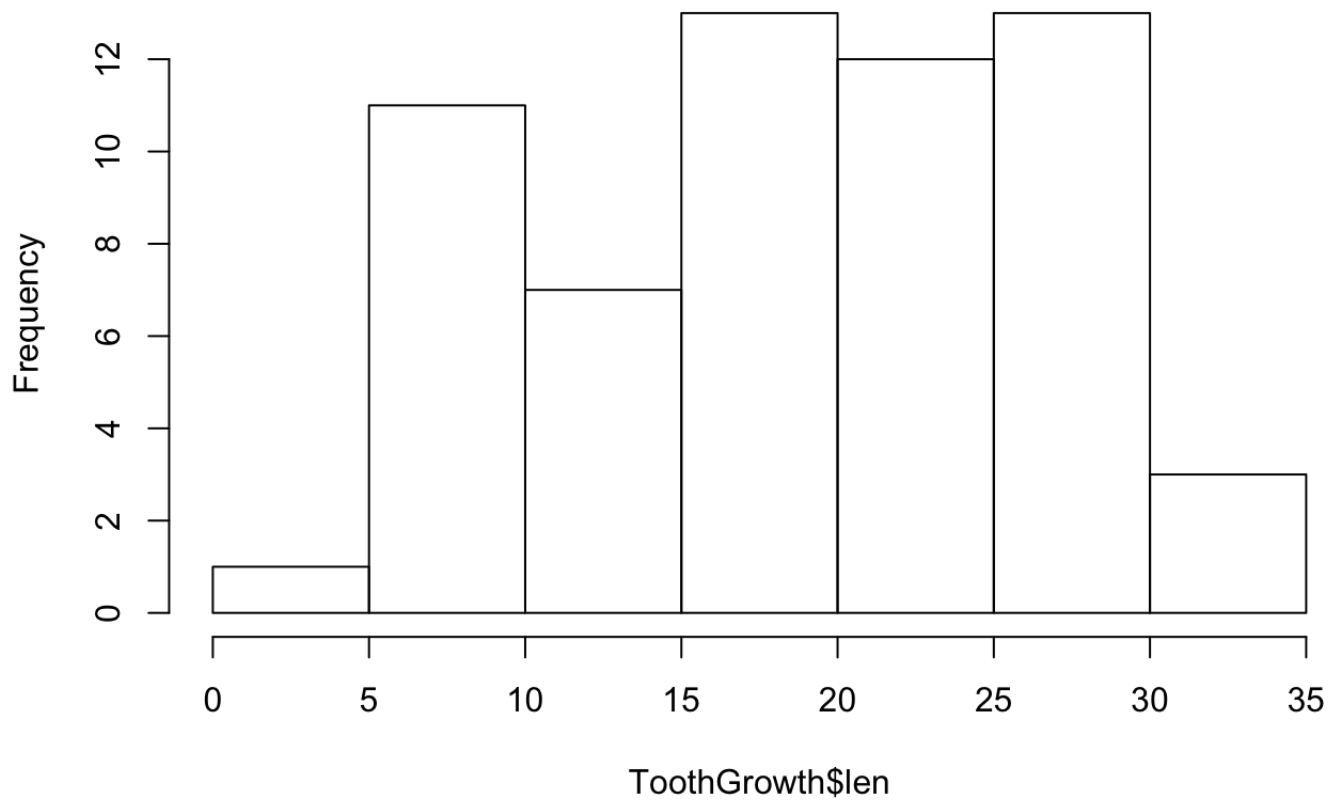
Michael Karp

1/22/2014
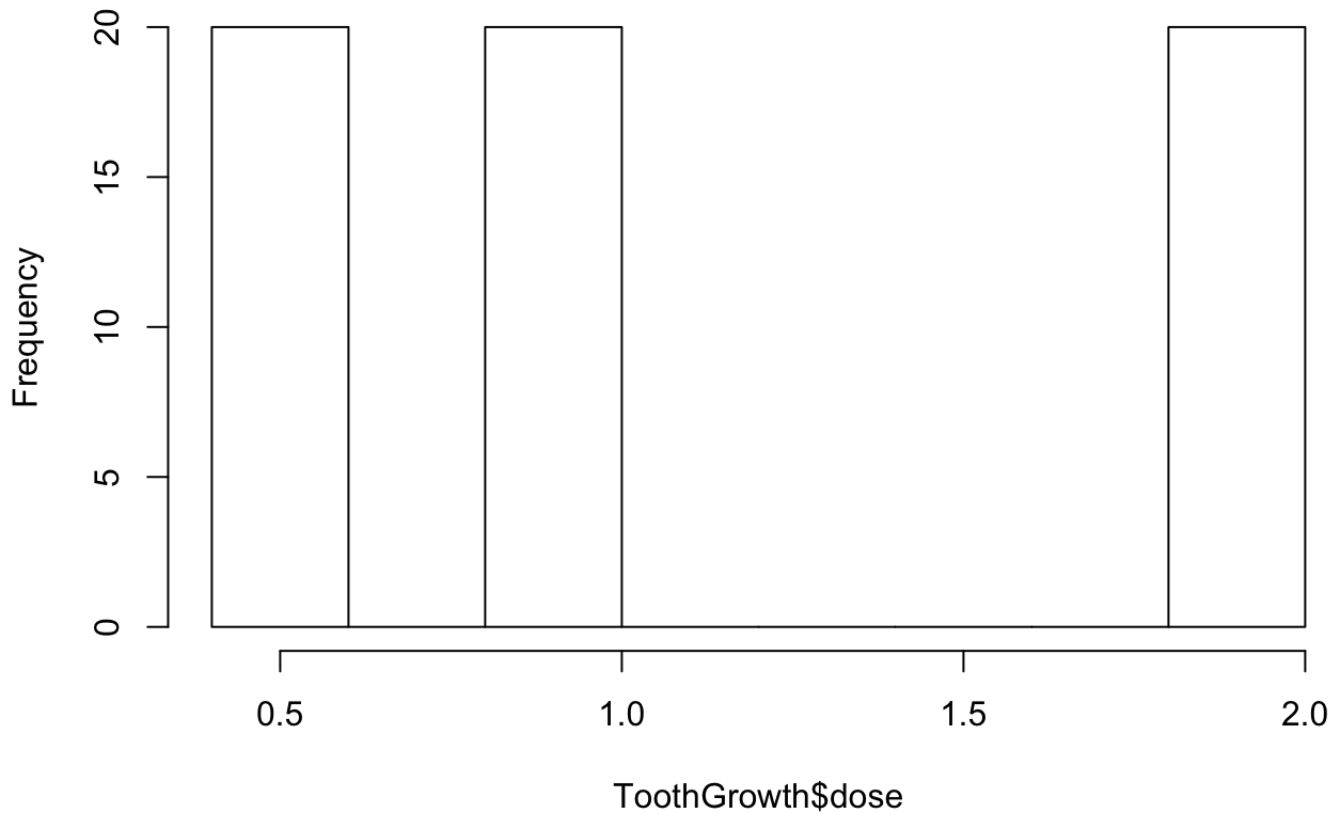
Investigating the tooth growth dataset

```
# set seed to make research reproducible
set.seed(1000)
# Load the ToothGrowth data and perform some basic exploratory data analyses
library(datasets)
hist(ToothGrowth$len)
```

Histogram of ToothGrowth$len


```
hist(ToothGrowth$dose)
```

Histogram of ToothGrowth$dose


```
table(ToothGrowth$supp == "VC", ToothGrowth$len > 18.81)
```
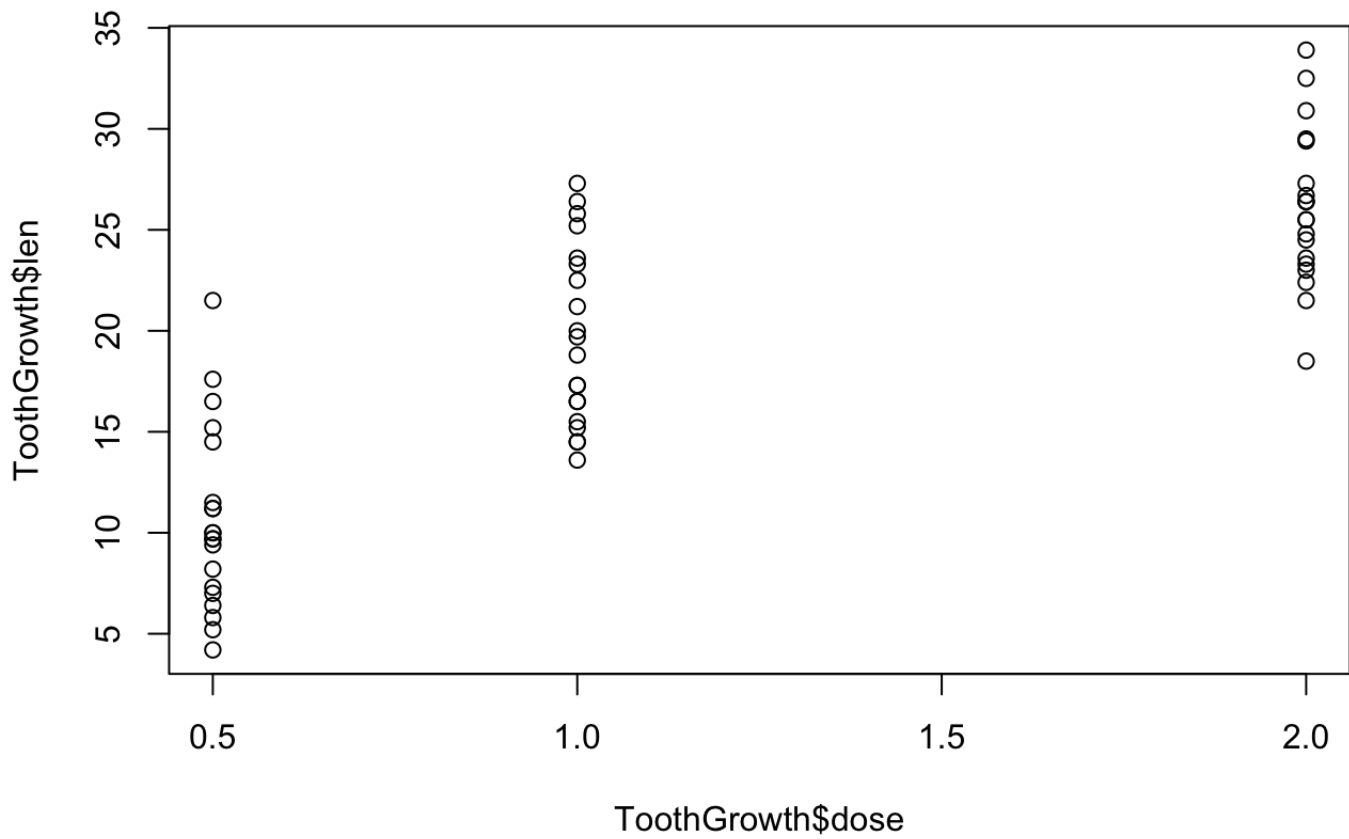
```
      FALSE TRUE
```

FALSE 10 20 TRUE 20 10

```
table(ToothGrowth$supp == "OJ", ToothGrowth$len > 18.81)
```
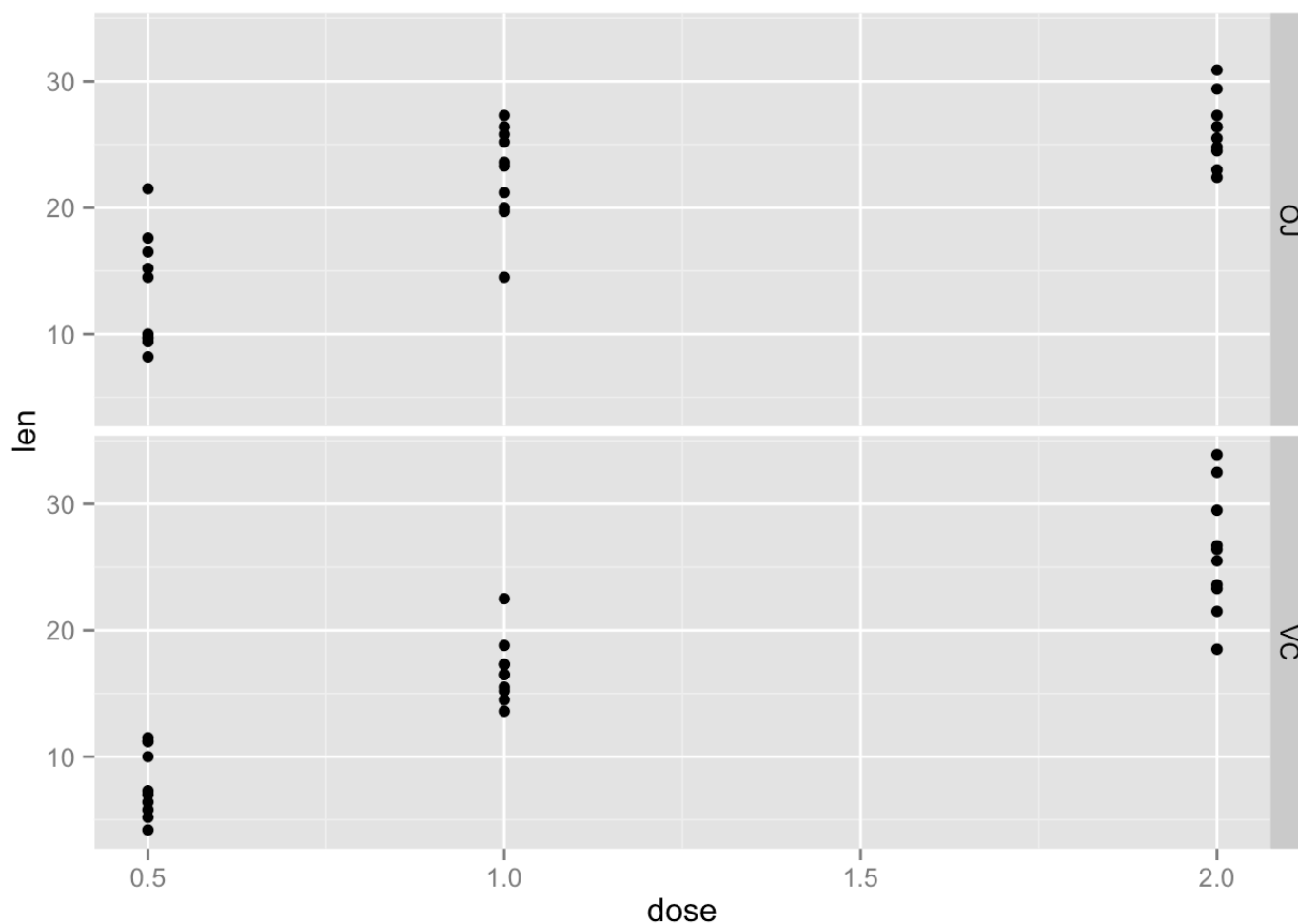
```
      FALSE TRUE
```

FALSE 20 10 TRUE 10 20

```
plot(ToothGrowth$len~ToothGrowth$dose)
library(ggplot2)
```

```
exp_plot <- ggplot(data = ToothGrowth, aes(x = dose, y = len)) + geom_point() + facet_grid(
supp~.)
exp_plot
```

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

```
# 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)
# H_0 <- mean(dose) == 1.167
# H_a <- mean(dose) > 30



# State your conclusions and the assumptions needed for your conclusions.
```