

# Tooth Growth and Supplemen

*Cliff Lau*

## Exploratory Analysis and Summary

In the analysis we'll use the ToothGrowth dataset:

```
library(ggplot2)
library(dplyr)

data("ToothGrowth")
tooth <- tbl_df(ToothGrowth)
```

Some quick summary statistics of the dataset:

```
summary(tooth)
```

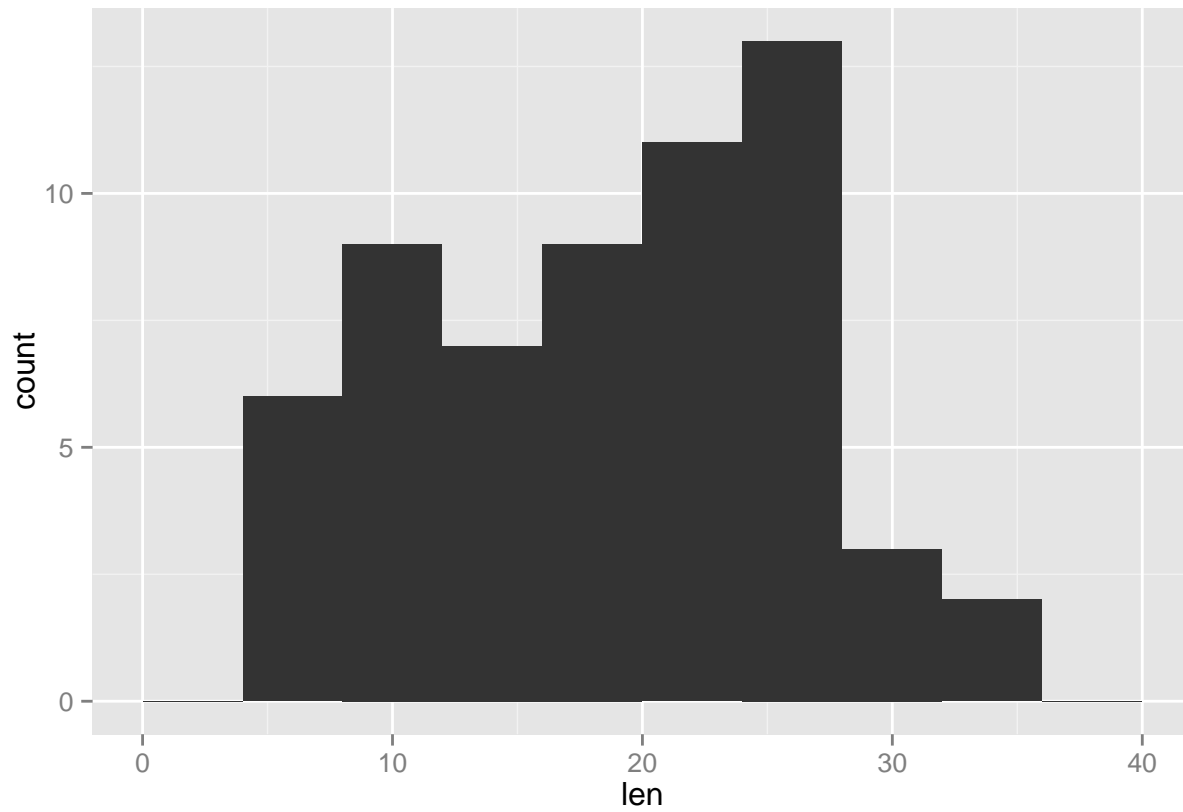
##	len	supp	dose
## Min.	: 4.20	OJ:30	Min. :0.500
## 1st Qu.	:13.07	VC:30	1st Qu.:0.500
## Median	:19.25		Median :1.000
## Mean	:18.81		Mean :1.167
## 3rd Qu.	:25.27		3rd Qu.:2.000
## Max.	:33.90		Max. :2.000

```
glimpse(tooth)
```

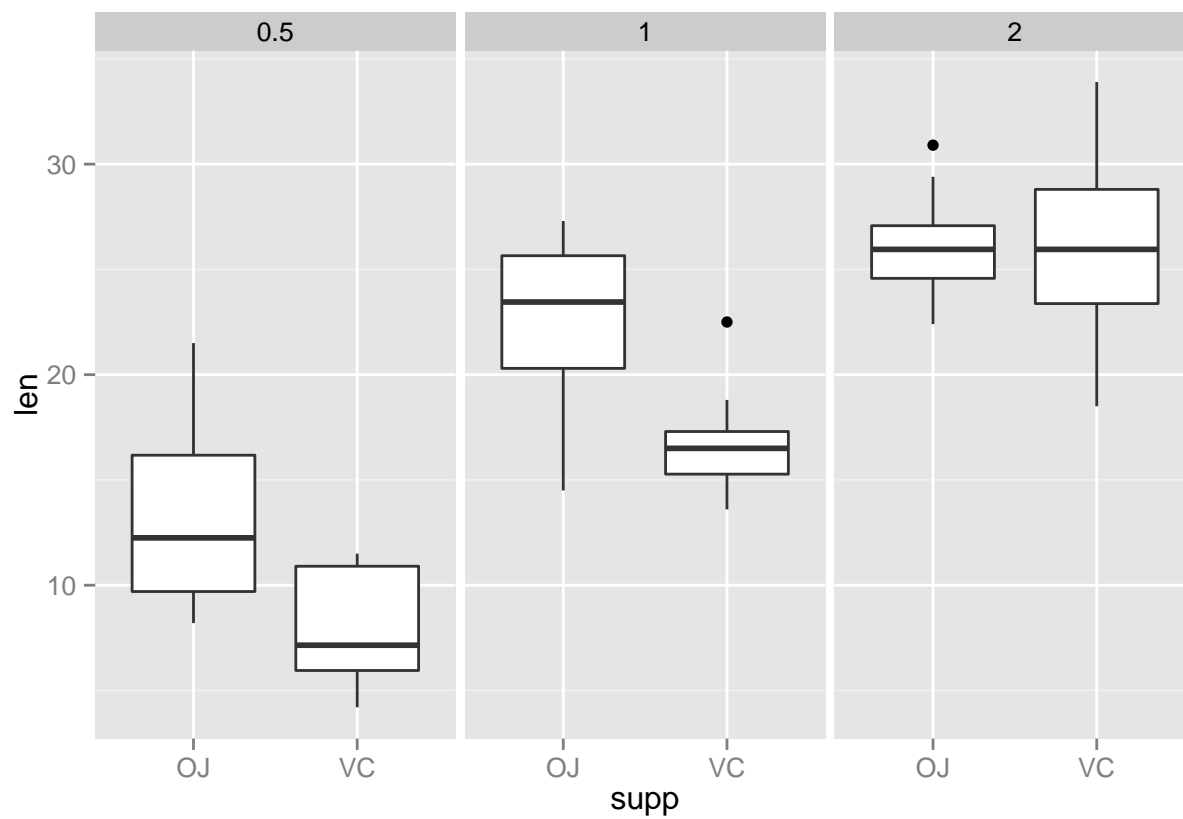
```
## Observations: 60
## Variables:
## $ len (dbl) 4.2, 11.5, 7.3, 5.8, 6.4, 10.0, 11.2, 11.2, 5.2, 7.0, 16....
## $ supp (fctr) VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, ...
## $ dose (dbl) 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1.0, 1....
```

Some exploratory graphs of the dataset:

```
ggplot(tooth, aes(len)) + geom_histogram(binwidth = 4)
```



```
ggplot(tooth, aes(supp, len, factor(dose))) + geom_boxplot() + facet_grid(. ~ dose)
```



The histogram shows the distribution of tooth length. We can see that it's mostly in the range of 4.2 to 33.9, with median of 19.25.

## Hypothesis Testing and Confidence Interval

To compare length of tooth based on supplement type:

```
t.test(len ~ supp, paired = FALSE, data = tooth)

##
##  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
```

Based on the results above, with p-value > 5%, we do not reject the null hypothesis that the true difference is equal to 0. Also, based on the CI provided, since it includes 0, which means that it's possible that the difference is 0.

```
tooth2 <-
  tooth %>%
  filter(dose != 2.0)
t.test(len ~ supp, paired = FALSE, data = tooth2)

##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.0503, df = 36.553, p-value = 0.004239
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.875234 9.304766
## sample estimates:
## mean in group OJ mean in group VC
##      17.965      12.375
```

```
tooth3 <-
  tooth %>%
  filter(dose != 0.5)
t.test(len ~ supp, paired = FALSE, data = tooth3)

##
##  Welch Two Sample t-test
##
## data:  len by supp
```

```
## t = 1.8397, df = 31.273, p-value = 0.07533
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.3166175  6.1666175
## sample estimates:
## mean in group OJ mean in group VC
##           24.380           21.455
```

Based on the above, it seems there is evidence that there is a greater than 0 effect with the supplement comparing half a dose to 1 dose. However, when comparing 1 to 2 dose, we found not statistical evidence that there is a effect on tooth length.

## Conclusion

Based on our analysis above, we can conclude the following:

1. There is no difference between the 2 types of suppliment, VC or OJ in tooth length
2. There is an increase in tooth length when we compare 0.5 to 1 dose of suppliment, however, there are no evidence of difference between 1 and 2 dose.