Tooth Growth

Zoheb Nensey

Overview

The purpose of this analysis is to look at tooth growth. We test the differences between different dosages and supplement combinations.

Preliminary steps.

1. Load the data.

```
library(datasets)
data(ToothGrowth)
```

2. Load useful packages.

```
library(ggplot2)
library(lattice)
library(dplyr)
```

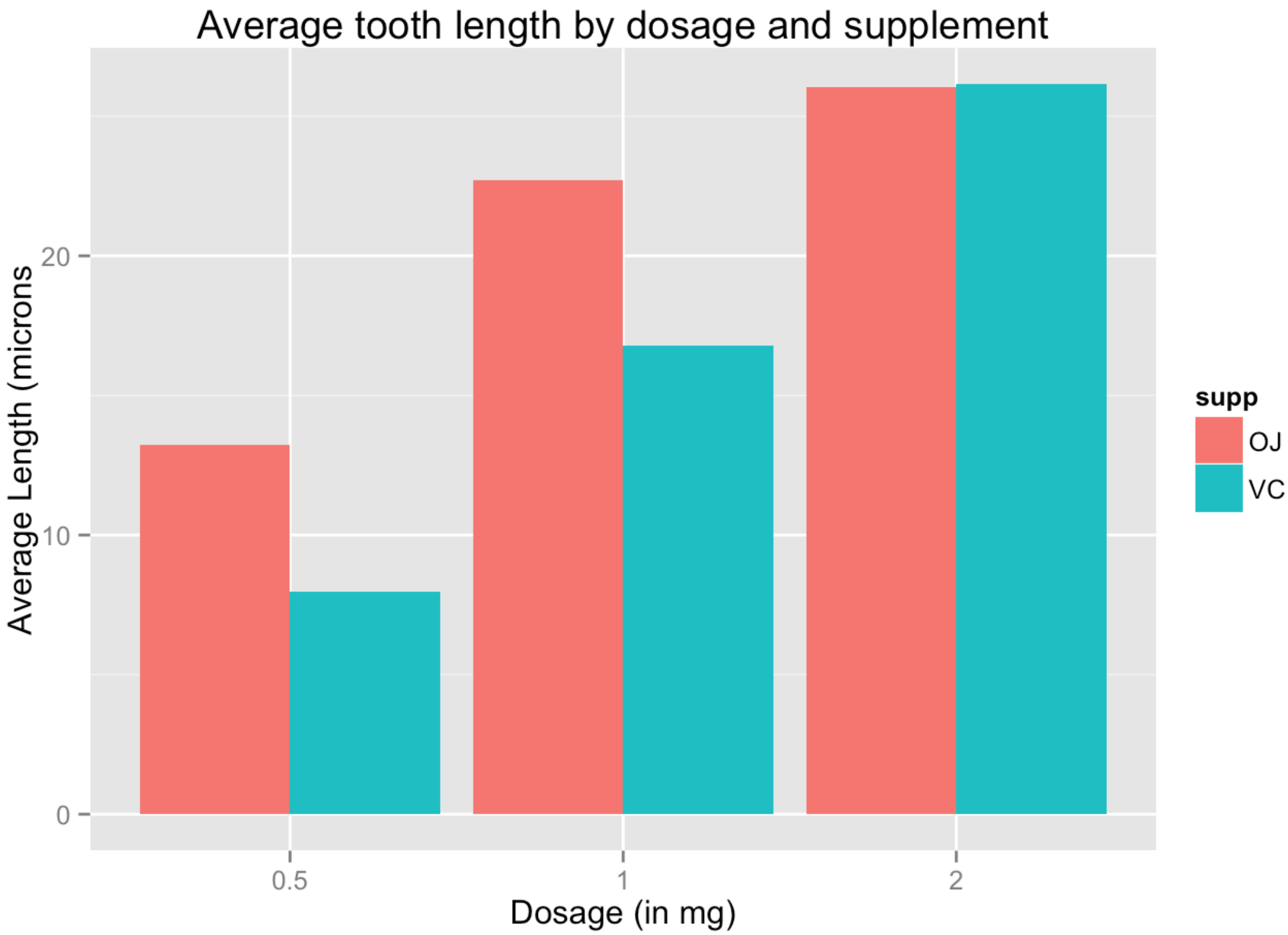
Exploratory analysis

First, it's worth noting that the data is already sorted in a useful fashion. The supp variable indicates the supplement being used, which can be orange juice (coded as OJ), or ascorbic acid (coded as VC). The length is measured in microns.

An exploratory analysis should look at the average tooth length by supplement and dose.

```
average_length <- ToothGrowth %>%
      group_by(supp,dose) %>%
      summarize(avg_len = mean(len), len_sd = sd(len), observations = n())
```

We can plot these on a chart.



It doesn’t seem surprising (logically) that tooth length grows up with the dosage of supplement. What is interesting, however, is how at some point the supplements are equal (at a dosage of 2 mg).

Basics on the data.

The best way to compare this data is going to be based on the group. The dplyr package already did the difficult work of averaging the data by group and giving us the number of observations in each group. You can see those results below.

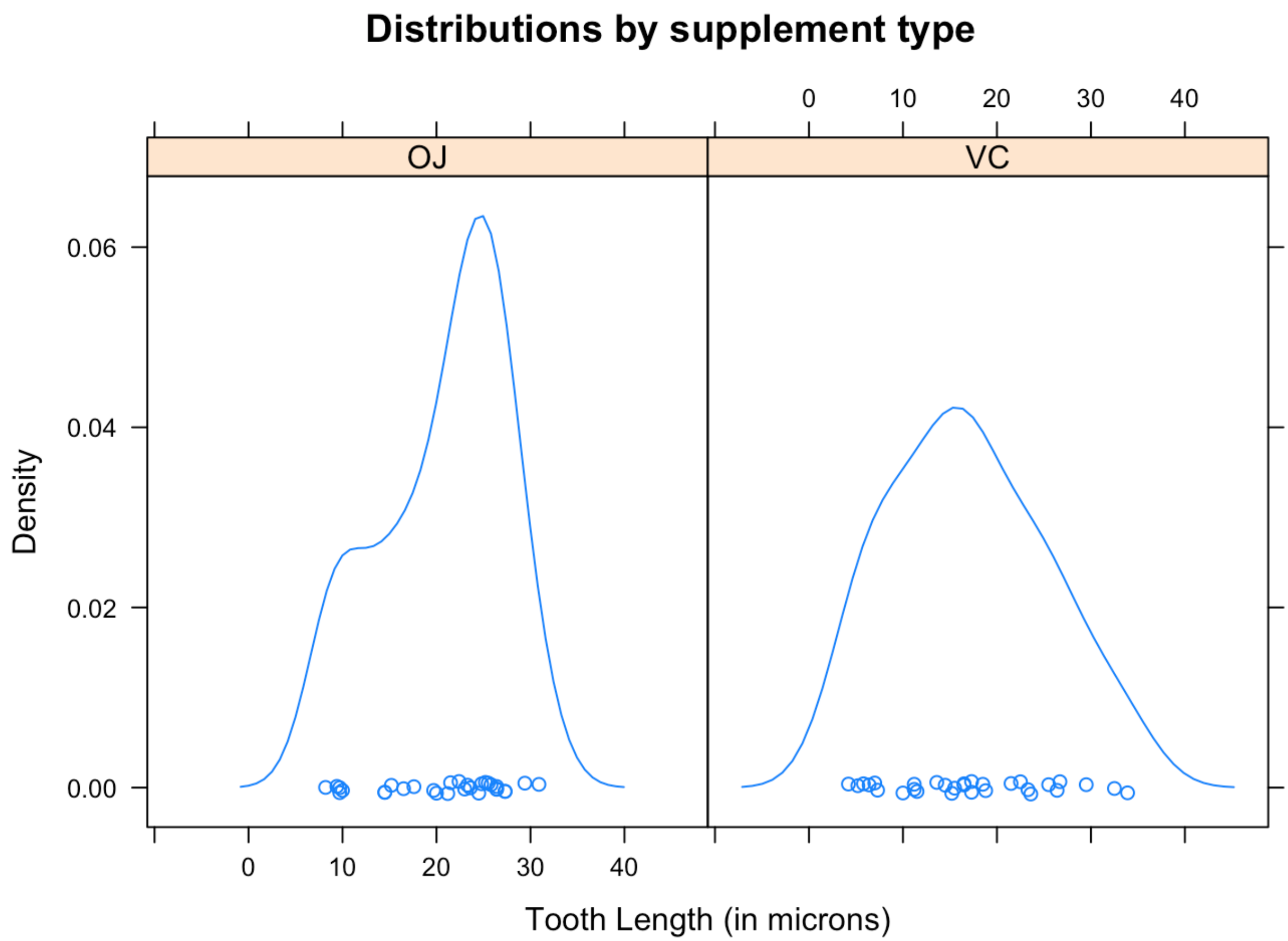
```
## Source: local data frame [6 x 5]
## Groups: supp
##
##   supp dose avg_len   len_sd observations
## 1   OJ  0.5   13.23 4.459709           10
## 2   OJ  1.0   22.70 3.910953           10
## 3   OJ  2.0   26.06 2.655058           10
## 4   VC  0.5    7.98 2.746634           10
## 5   VC  1.0   16.77 2.515309           10
## 6   VC  2.0   26.14 4.797731           10
```

This data is useful, as it can help us calculate inferential statistics and compare by group.

Analysis of the data

For this analysis, we'll most likely want to know if there's differences by supplement. The easiest way to do this is to see if the data is using a t-test, since the data is limited. To do this, however, we need to see if it's normally distributed. We can make this significantly easier by creating a histogram for each supplement.

```
densityplot(~ToothGrowth$len|ToothGrowth$supp,
            xlab = "Tooth Length (in microns)",
            main = "Distributions by supplement type")
```



This data is imperfect, but it does seem reasonably well distributed and not overly skewed, especially since the t-test is robust to this assumption.

Now, let's go ahead and compare the growth of teeth by supplement type. The null hypothesis for this will be that there is no difference between supplement types for tooth growth. The alternate hypothesis is 2-sided and says that there is a difference. For this test, we'll be setting the alpha at .05.

```
results <- t.test(len ~ supp,
       data=ToothGrowth,
       paired=FALSE,
       conf.level = 0.95)

pval <- results$p.value
```

Comparing these two groups, we can see that the 95% confidence interval for the 2-sided test crosses zero and is not significant. Further, the p-value is 0.0606345, which is lower than the specified alpha of .05.

We should also compare these supplements by dosage level, as this well tell us if there truly is a difference between the lower dosages as it appeared during exploratory data analysis.

We can start by subsetting the data by dosage level, then running t-tests on each subset.

For the 0.5 milligram group:

At this dosage level, we can see that there is a difference. The 95% confidence interval does not cross zero, and the p-value is is 0.0063586, which is less than the alpha of 0.05.

For the 1.0 milligram group:

At this dosage level, we can once again see a difference. The 95% confidence interval does not cross zero, and the p-value is is 0.0010384, which is less than the alpha of 0.05.

For the 2.0 milligram group:

At this level, however, everything changes. There is no difference when the dosage is 2.0 milligrams, as the 95% confidence intervaldoes cross zero and the p-value is 0.9638516, which is extremely high.

Conclusions

It appears that there isn’t a huge difference once dosages grow past 2.0 mg, but at smaller dosages there appears to be a large difference.