

Statistical Inference Project Part 2

Contents

0.1	Synopsis	1
0.2	Data Description	1
0.3	Data Processing	1
0.4	Anaylsis	2
0.5	Conslusions and Assumptions	5

0.1 Synopsis

This is the project for the statistical inference class. This is the second part of the project which is an analysis of the ToothGrowth data.

In the second portion of the class, I am going to analyze the ToothGrowth data in the R datasets package.

0.2 Data Description

The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

[,1] len numeric Tooth length
[,2] supp factor Supplement type (VC or OJ).
[,3] dose numeric Dose in milligrams.
(source: R Doumentation)

0.3 Data Processing

0.3.1 Loading packages and libraries

```
library(ggplot2)
library(datasets)
library(reshape2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

0.3.2 Loading dataset

```
toothGrowth = tbl_df(ToothGrowth)
```

0.4 Anaylsis

0.4.1 Summarizing Data

```
glimpse(toothGrowth)
```

```
## 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....
```

```
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
```

```
head(toothGrowth,3); tail(toothGrowth,3)
```

```
## Source: local data frame [3 x 3]
##
##      len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
```

```
## Source: local data frame [3 x 3]
##
##      len supp dose
## 58 27.3   OJ    2
## 59 29.4   OJ    2
## 60 23.0   OJ    2
```

0.4.2 Exploratory Data Analysis

0.4.2.1 Tooth length vs supplement type First, I will explore the relaion between the supplement type and the lenth of tooth

```
ggplot(data=toothGrowth, aes(x=len)) +
  geom_histogram(binwidth=2) +
  facet_grid( . ~ supp) +
  xlab("Tooth Length") +
  ylab("Count")
```

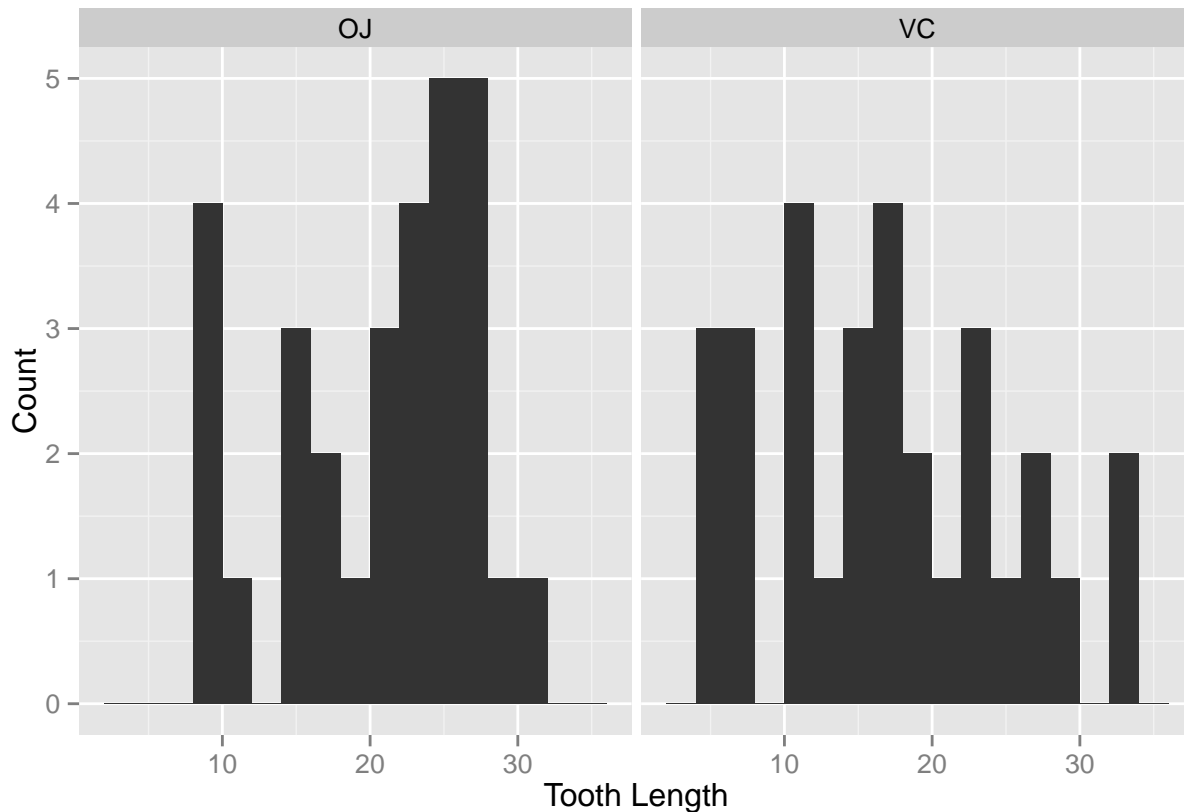


Figure 1:

It seems that both supplements have similar effects on the length of tooth, however, the orange juice, OJ, has more profound effect on most of the teeth

0.4.2.2 Tooth length vs dosage I will then explore the effect of the dosage quantity on the length of the tooth while examining this effect for both of the supplements.

```
ggplot(data=toothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
  geom_bar(stat="identity") +
  xlab("Dosage (mg)") +
  ylab("Tooth Length") +
  guides(fill=guide_legend(title="Supplement Type"))
```

As the dosage increases so does the length of the tooth, however, again the orange juice, the OJ supplement has a stronger effect on the tooth length.

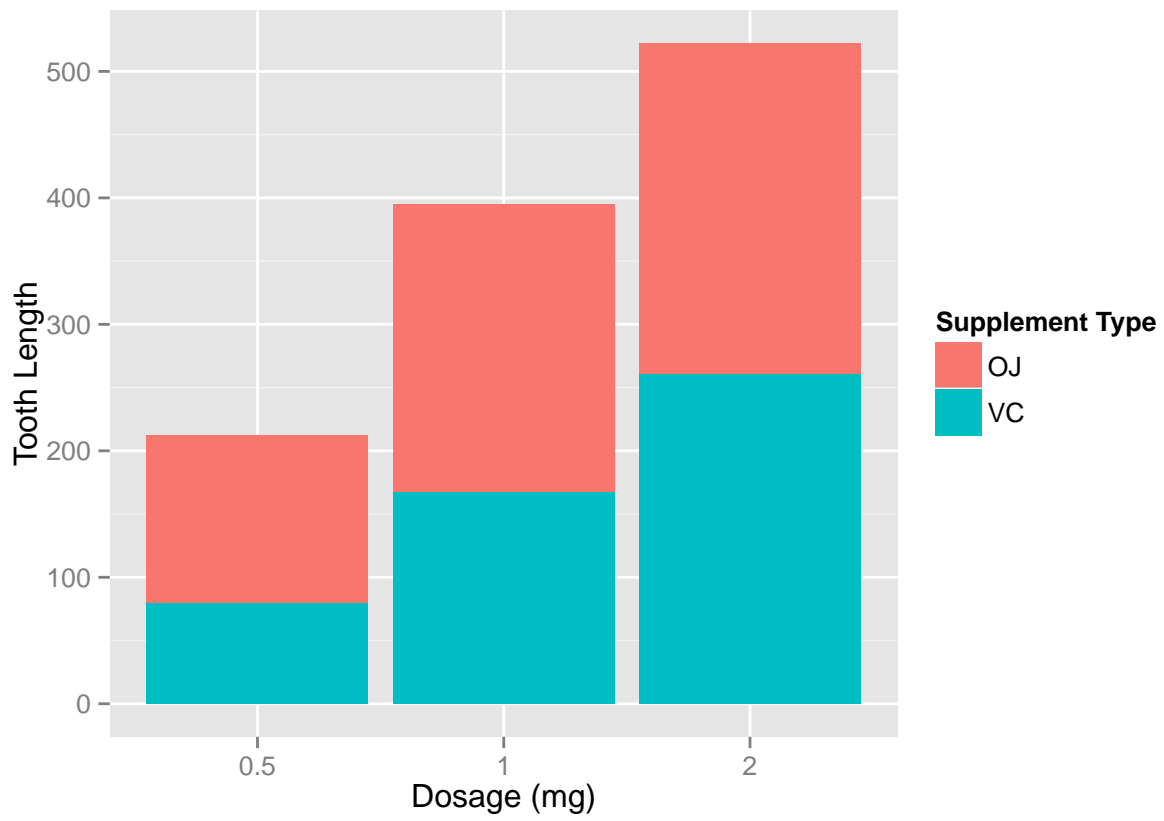


Figure 2:

0.4.3 Comparison of tooth growth by supp and dose

Did the student perform some relevant confidence intervals and/or tests? Were the results of the tests a

0.5 Conslusions and Assumptions

Did the student describe the assumptions needed for their conclusions?