Statistical Inference Part 2

Gill Collier

07/11/2020

Gill Collier - 01 November 2020

Coursera: Statistical Inference - Course Project

Part 2: Basic Inferential Data Analysis

Now in the second portion of the project, we're going to analyze the

ToothGrowth data in the R datasets package.

Libraries

```
library(ggplot2)
```

```
## Warning: replacing previous import 'vctrs::data_frame' by 'tibble::data_frame'
## when loading 'dplyr'
```

Load the ToothGrowth data and perform some basic exploratory data analyses

```
data = ToothGrowth
?"ToothGrowth"
```

```
## starting httpd help server ... done
```

The ToothGrowth data shows the Effect of Vitamin C on Tooth Growth in Guinea Pigs

Description

The response is the length of odontoblasts (cells responsible for tooth

growth) in 60 guinea pigs. Each animal received one of three dose levels of

vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange

juice or ascorbic acid (a form of vitamin C and coded as VC).

```
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 ...
```

```
head(ToothGrowth)
```

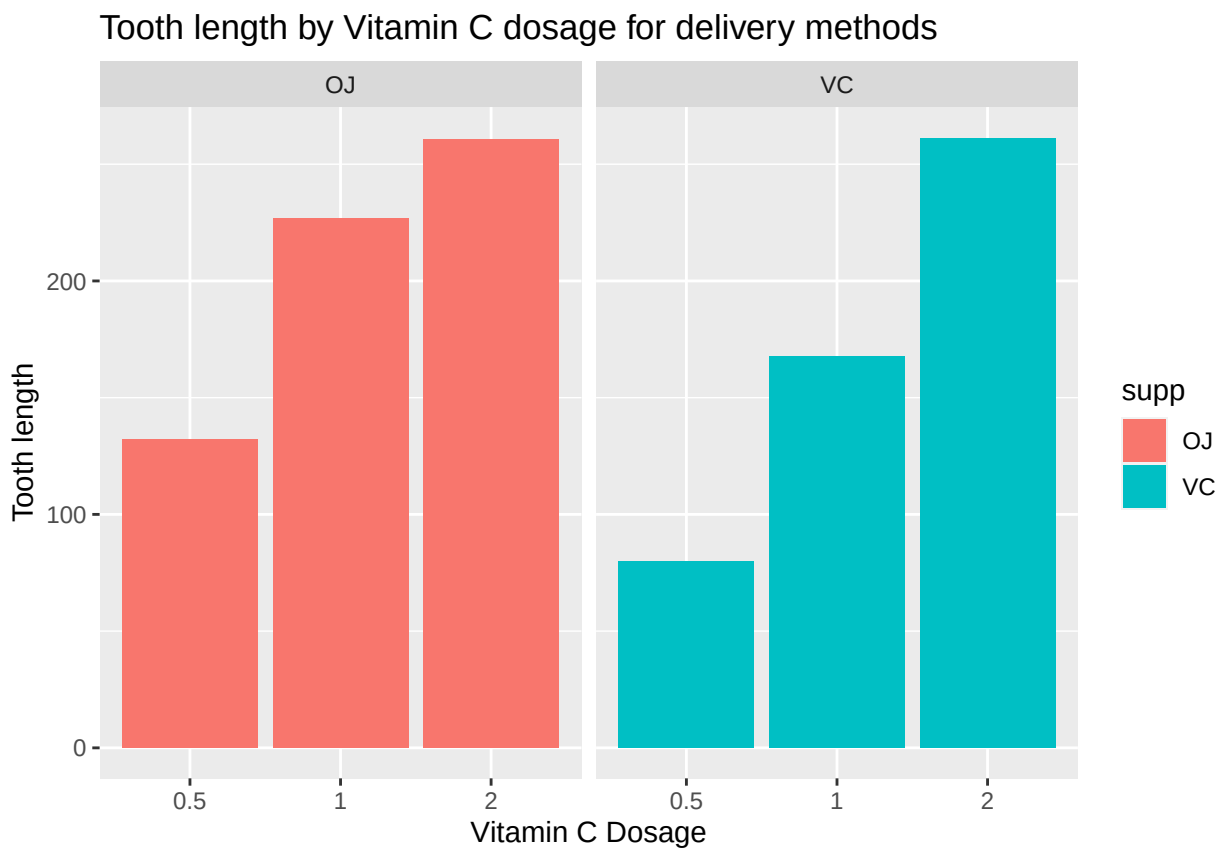
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

Summarise the data

```
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
```

```r
ggplot(data=ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
        geom_bar(stat="identity") +
        facet_grid(. ~ supp) +
        ggtitle("Tooth length by Vitamin C dosage for delivery methods") +
        xlab("Vitamin C Dosage") +
        ylab("Tooth length")
```



Use confidence intervals and/or hypothesis tests to compare tooth growth

by supp and dose.

```r
t.test(len ~ supp, data = ToothGrowth)
```

```
##
##  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
```

```
t.test(len ~ supp, data = subset(ToothGrowth, dose == 0.5))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

```
t.test(len ~ supp, data = subset(ToothGrowth, dose == 1))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

```
t.test(len ~ supp, data = subset(ToothGrowth, dose == 2))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

State your conclusions and the assumptions needed for your conclusions.

Tooth Length does increase as the dosage is increase for both delivery methods

When the lower dosages are administered (0.5 and 1) the most effective

method seems to the OJ but this is not the case with the higher dosage

The t-test analysis shows that the delivery method has no real effect

on tooth growth.