Assignment1_part2

Haoyi Wei

2022-12-19

Analyze the ToothGrowth data in the R datasets package.

Overview

This project analyze the ToothGrowth data in the R datasets package.

Load the ToothGrowth data and perform some basic exploratory data analyses

```
data(ToothGrowth)
```

Provide a basic summary of the data

```
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
```

```
dim(ToothGrowth)
```

```
## [1] 60  3
```

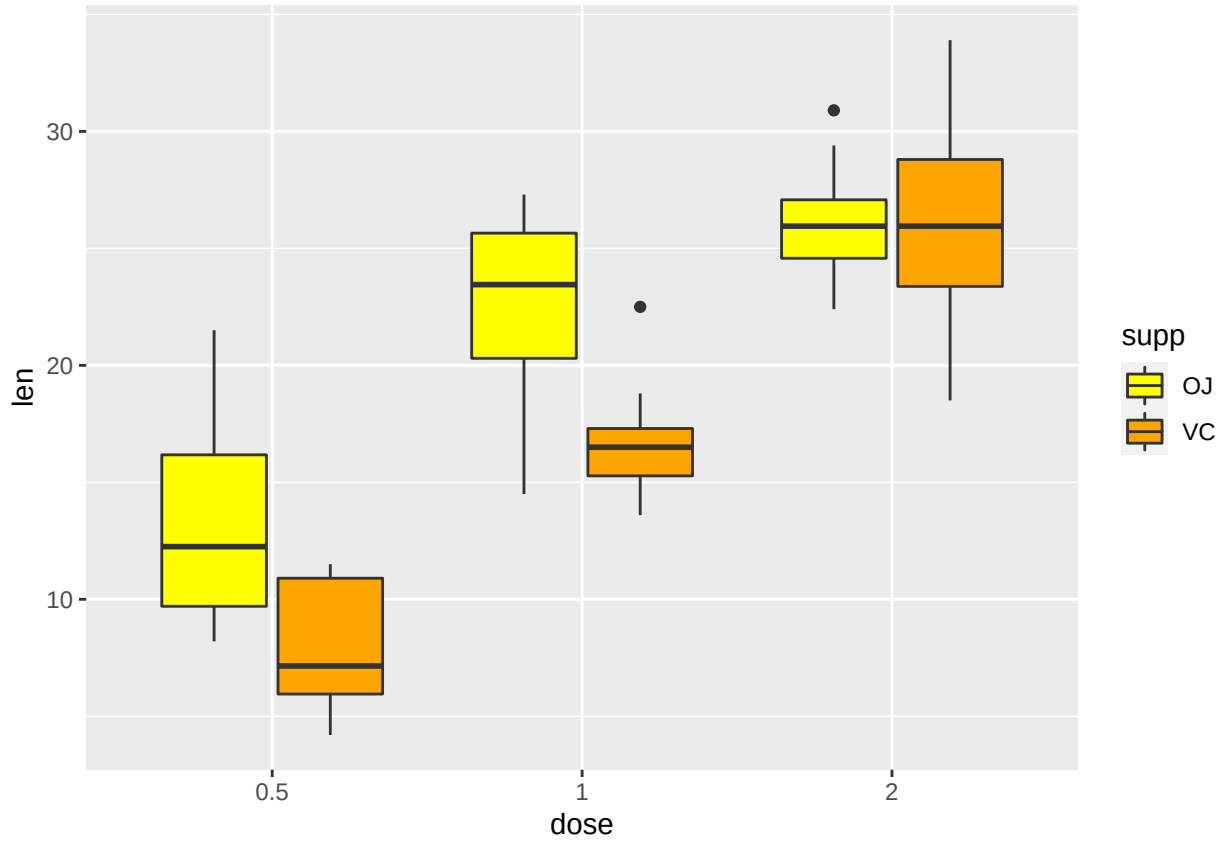
```
str(ToothGrowth)
```

```
## '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
```

```
library(ggplot2)
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
ggplot(ToothGrowth, aes(x=dose, y= len, fill=supp)) +
          geom_boxplot() +
          scale_fill_manual(values=c("yellow","orange"))
```



1. 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)

```
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
```

```
# Hypothesis Testing
dose.5 <- ToothGrowth %>%
          filter(dose ==0.5)
dose1 <- ToothGrowth %>%
          filter(dose ==1)
dose2 <- ToothGrowth %>%
          filter(dose ==2)
```

```
t.test(len ~ supp, dose.5)
```

t-test for 0.5 mg/day dose

```
##
##  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 between group OJ and group VC 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, dose1)
```

t-test for 1 mg/day dose

```
##
##  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 between group OJ and group VC 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, dose2)
```

t-test for 2 mg/day dose

```
##
##  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 between group OJ and group VC 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
```

Summarize the results Summarize the results in a table

```
dose <- c(0.5, 1.0, 2.0)
p_value <- c(0.0064, 0.0010, 0.9639)
cv <- c("1.72, 8.78", "2.80, 9.06", "-3.80, 3.64")
outcome <- c("Reject null", "Reject null", "Fail to reject null")
data.frame(dose, cv, p_value, outcome)
```

```
##   dose          cv p_value             outcome
## 1  0.5  1.72, 8.78  0.0064         Reject null
## 2  1.0  2.80, 9.06  0.0010         Reject null
## 3  2.0 -3.80, 3.64  0.9639 Fail to reject null
```

For dose 0.5 and dose 1, the p value are indistinguishable from zero, thus reject the null hypotheses that the supplement types don't have a difference on tooth growth.

For dose 2.0, fail to reject the null because the p-value is greater than 0.5.

Conclusion

For low dose such as 0.5 and 1, OJ is higher than VC, but for dose 2, OJ and VC's effect is inditingusihable.