Week4 Project Part2

Melody Zhang

2022-06-27

Part2: Basic Inferential Data Analysis Instructions

1. Load the ToothGrowth data and perform some basic exploratory data analyses.

```
data("ToothGrowth")
library(ggplot2)
```

2.Provide a basic summary of 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
```

```
dim(ToothGrowth)
```

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

```
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
```

```
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 ...
```

```
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
```

3.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)
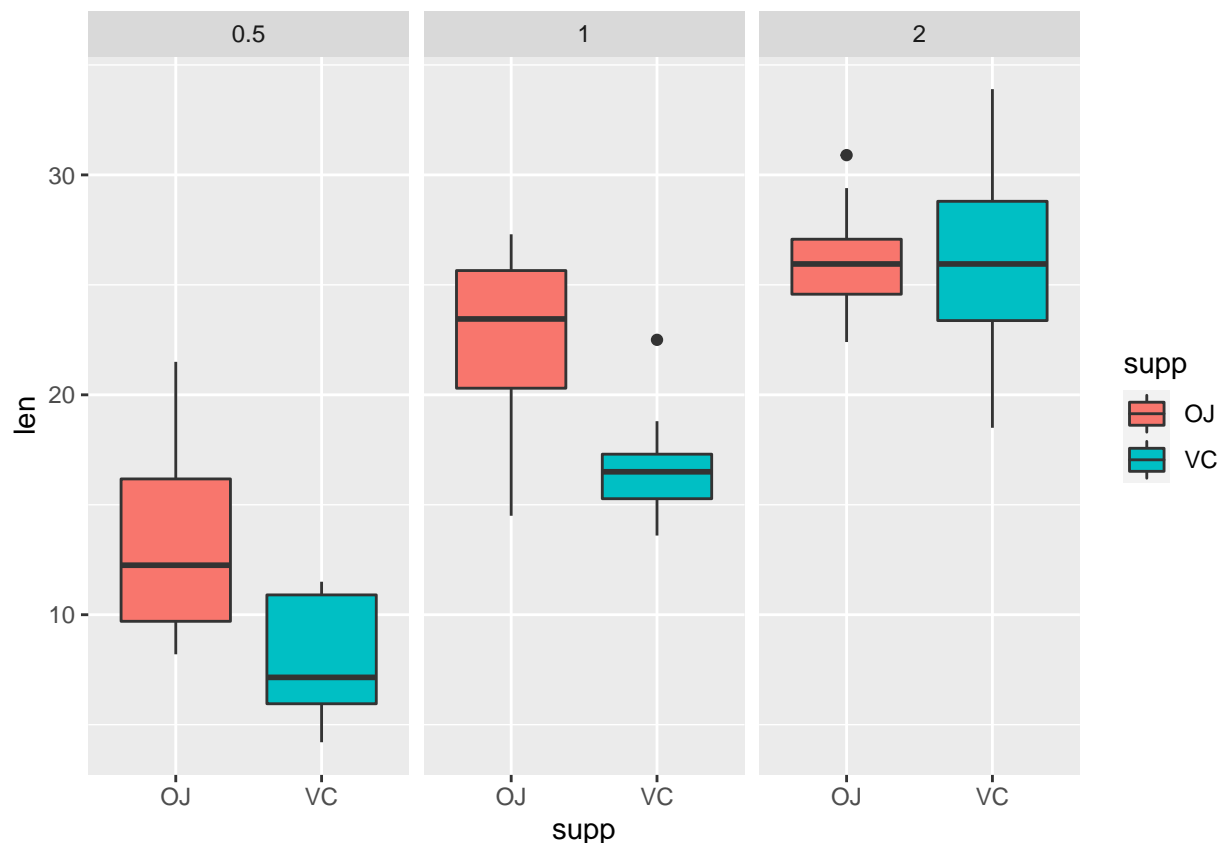
Firstly, analyse the relationship between tooth length and supplementary methods.

```
ggplot(ToothGrowth, aes(x=dose,y=len)) +
  geom_boxplot(aes(fill=dose)) +
  facet_grid(~ supp)
```
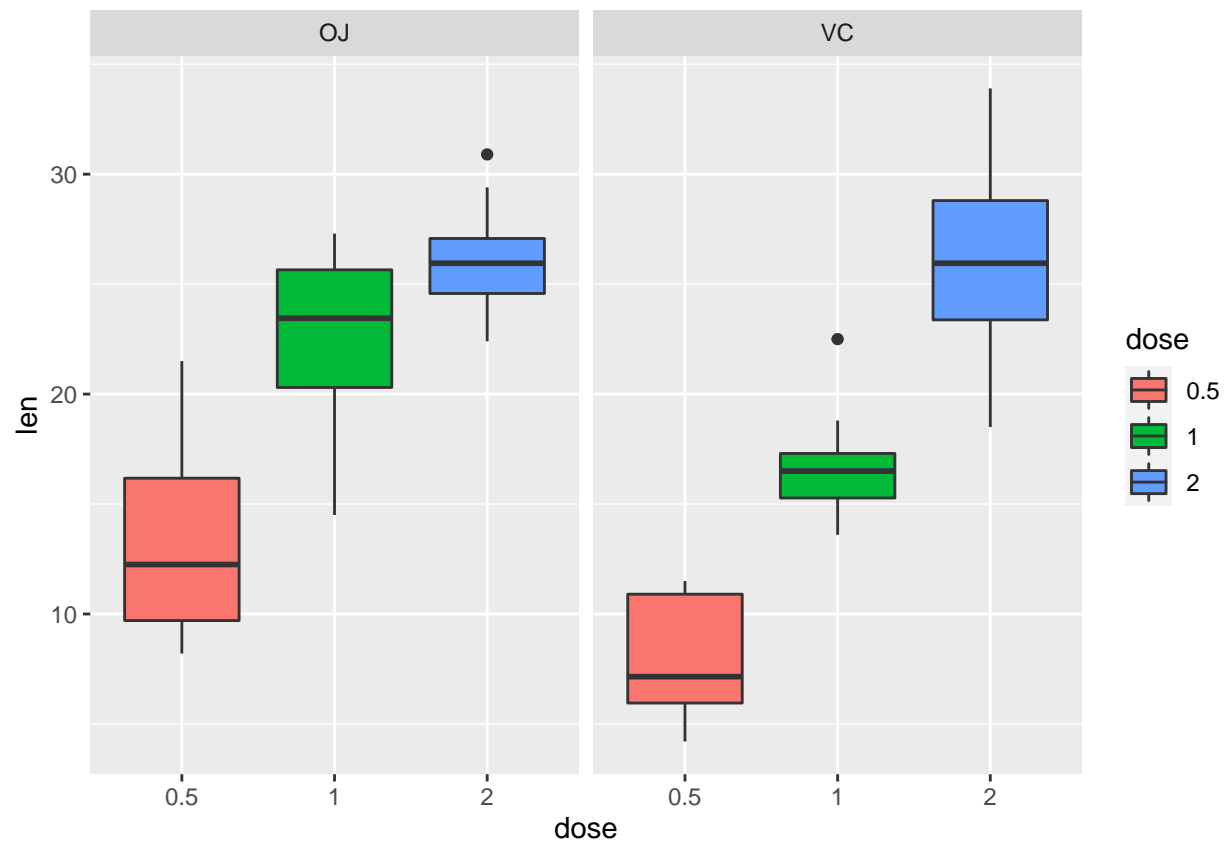

Secondly, analyze the relationship between tooth length and drug dose.

```
ggplot(ToothGrowth, aes(x=supp,y=len)) +
  geom_boxplot(aes(fill=supp)) +
  facet_grid(~ dose)
```

Thirdly, use t.test() to check the confidence interval.

```
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 between group OJ and group VC 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~dose,data = subset(ToothGrowth,dose == 0.5|dose == 1))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means between group 0.5 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5    mean in group 1
```

```
##                  10.605                  19.735
```

```
t.test(len~dose,data = subset(ToothGrowth,dose == 1|dose == 2))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```