Final project - Part 2

Raphael Carvalho

In this paper, we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.
4. State the conclusions and the assumptions needed for the conclusions.

Loading the packages

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

## Getting a glimpse of the data:

```
## Observations: 60
## Variables: 3
## $ len  <dbl> 4.2, 11.5, 7.3, 5.8, 6.4, 10.0, 11.2, 11.2, 5.2, 7.0, 16....
## $ supp <fct> VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, V...
## $ 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....
```

```
print(paste("Rows with NA values:", sum(!complete.cases(ToothGrowth))))
print(summary(ToothGrowth))
```

```
## [1] "Rows with NA values: 0"
##       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
```
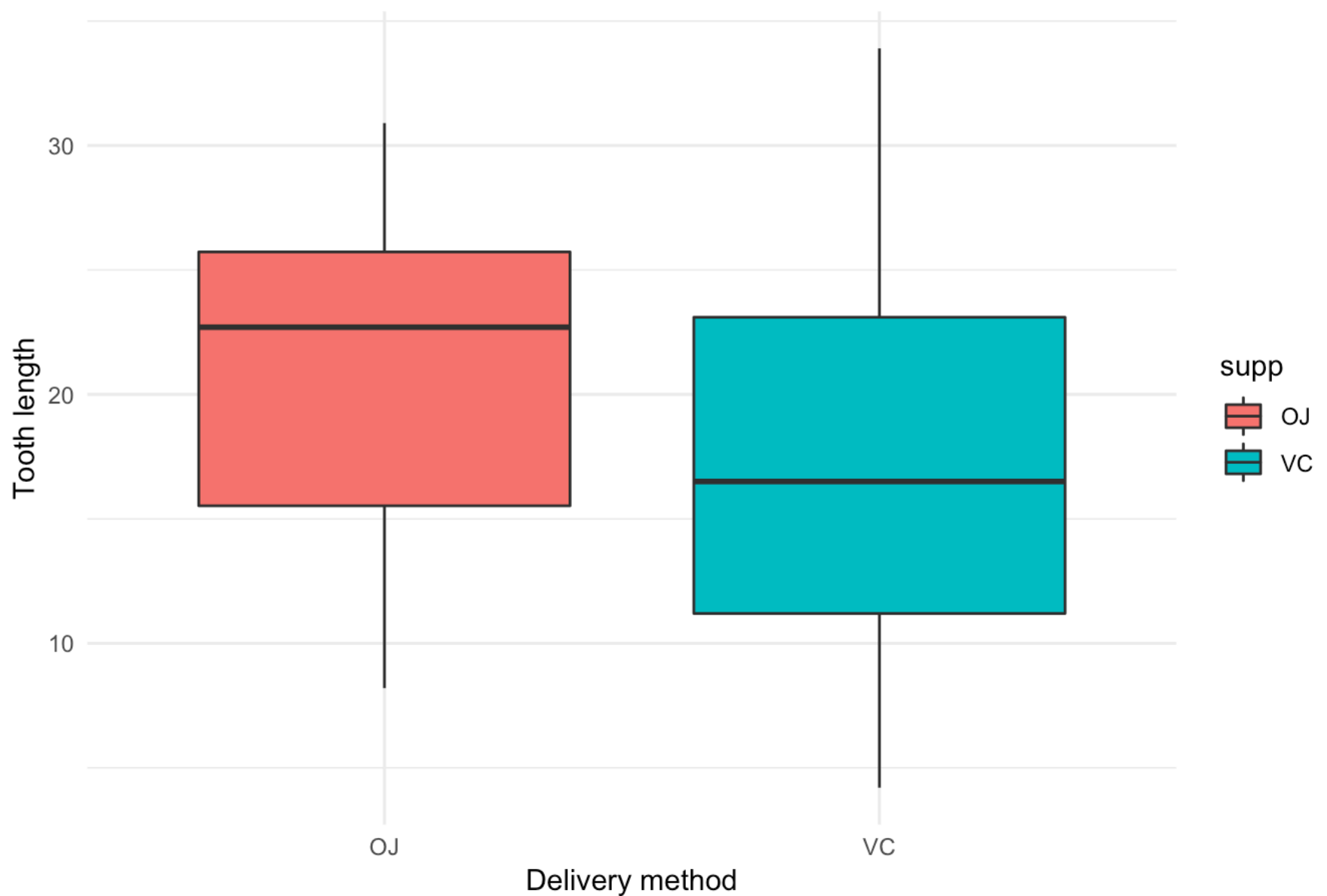
Exploring the dataset

Dispersion of tooth growth by dose
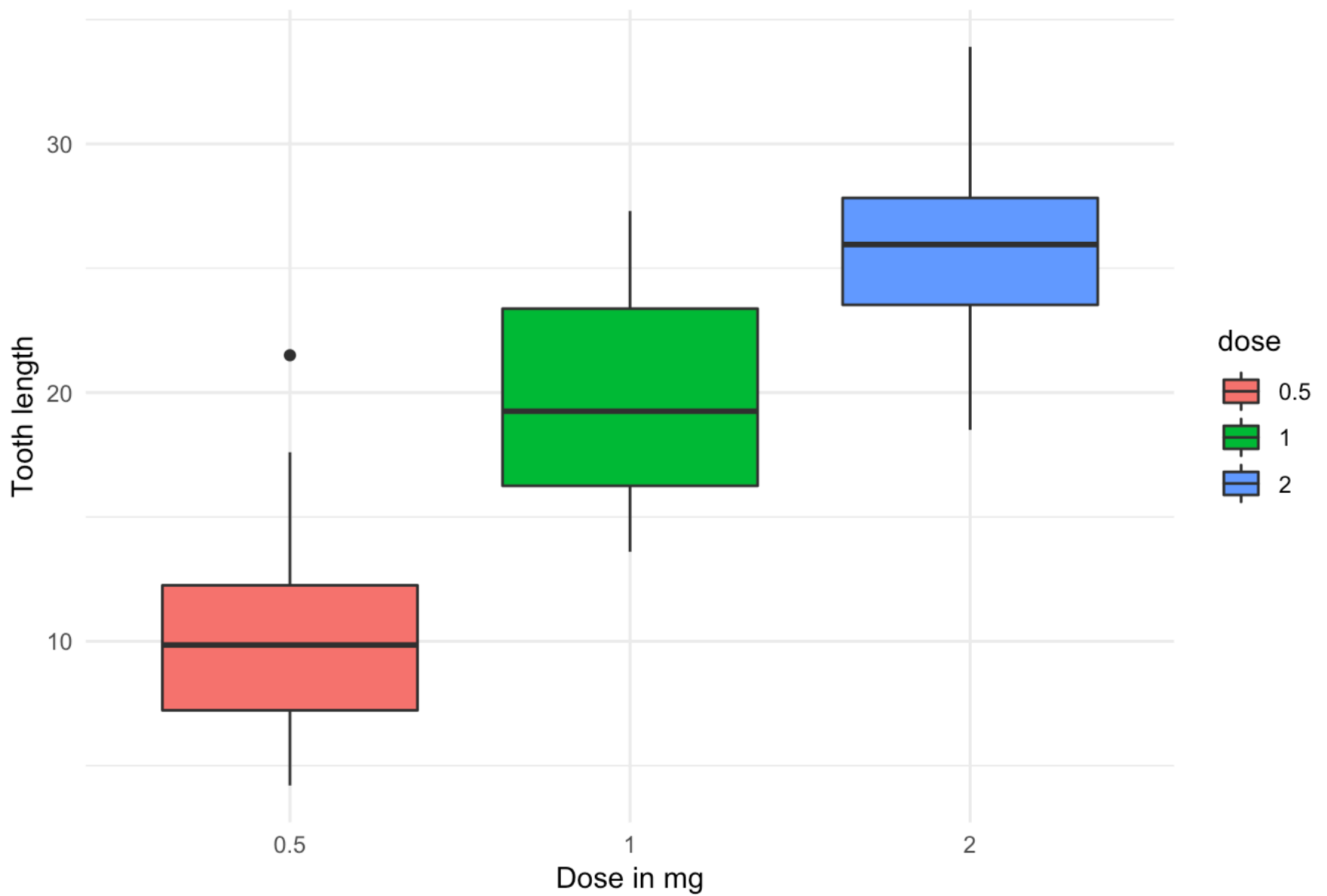
We can see that the longer theeth received higher doses. We can also notice that the dispersion of teeth length that were given a 1 mg dose is very high compared to the teeth subimitted to other doses. Lastly, we can see an outlier tooth at the group of 0.5 mg.

Dispersion of tooth growth by delivery method

The tooth length is similar for both delivery methods. However, we can see that the dispersion for Orange Juice is much lower than the dispersion observed for Vitamin C.

Testing some hyphotesis

Setting the variables:

```
dose <- ToothGrowth$dose
supp <- ToothGrowth$supp
len <-  ToothGrowth$len
```

Does the delivery mode has an influence on the tooth growth?

```
t.test(len[supp == "VC"],len[supp == "OJ"], paired=FALSE)
```

```
## 
##  Welch Two Sample t-test
## 
## data:  len[supp == "VC"] and len[supp == "OJ"]
## 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:
##  -7.5710156  0.1710156
## sample estimates:
## mean of x mean of y
##  16.96333  20.66333
```

This test shows us that:

1. We shouldn't reject the null hypothesis as the p-value is bigger than .05.
2. The confidence interval contains 0, so the test is not really significant.

Does the dose mode has an influence on the tooth growth?

```
t.test(len[dose == 0.5],len[dose == 1], paired=FALSE)
```

```
## 
##  Welch Two Sample t-test
## 
## data:  len[dose == 0.5] and len[dose == 1]
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

This test shows us that:

1. We should reject the null hypothesis as the p-value is almost 0.
2. The confidence interval does not contain 0, so the test is significant.

Conclusion

1. The dose of Vitamin C is proved to be a factor of teeth's growth for Guinea pigs;
2. The delivery mode does not appear to have an effect on teeth's growth.