Statistical Inference Project Part 2 - Basic Inferential Data Analysis On ToothGrowth

loading library

```
library(ggplot2)
library(datasets)
data(ToothGrowth)
str(ToothGrowth)
```

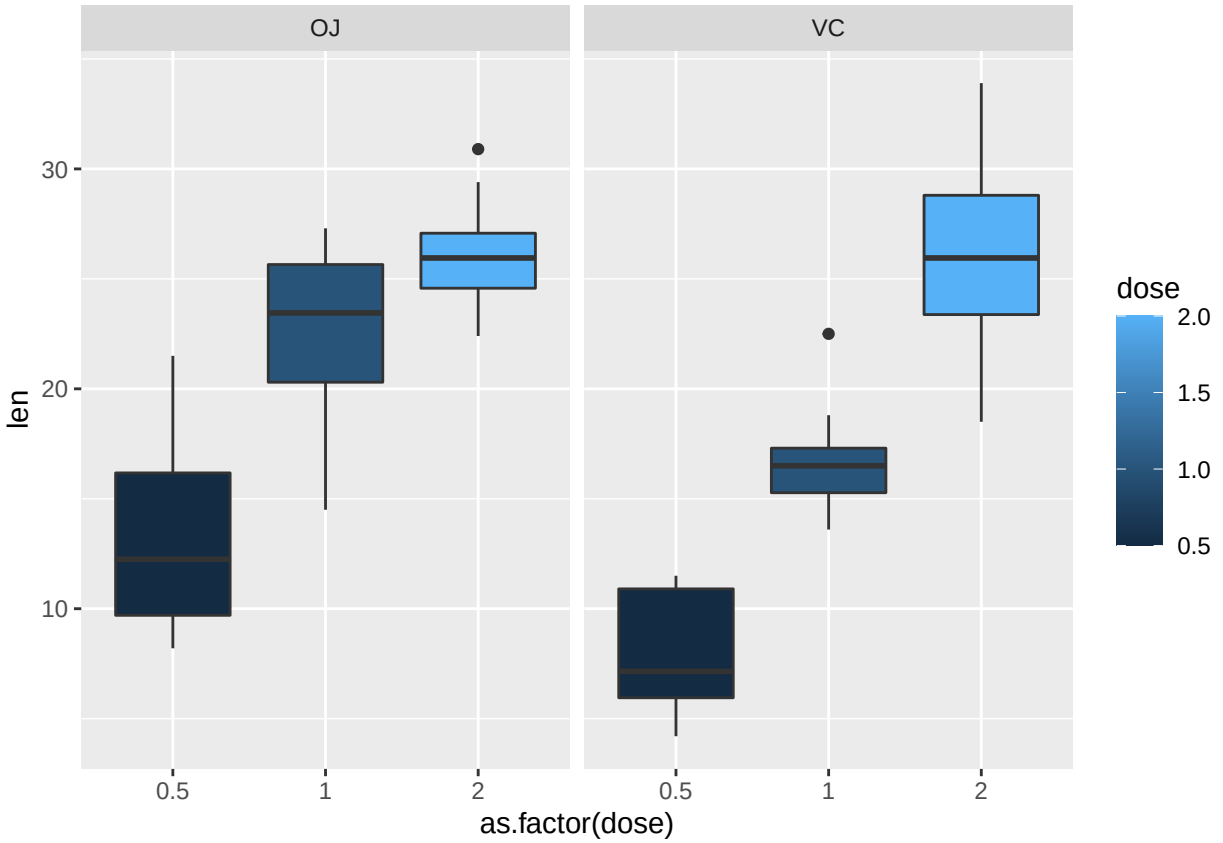
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

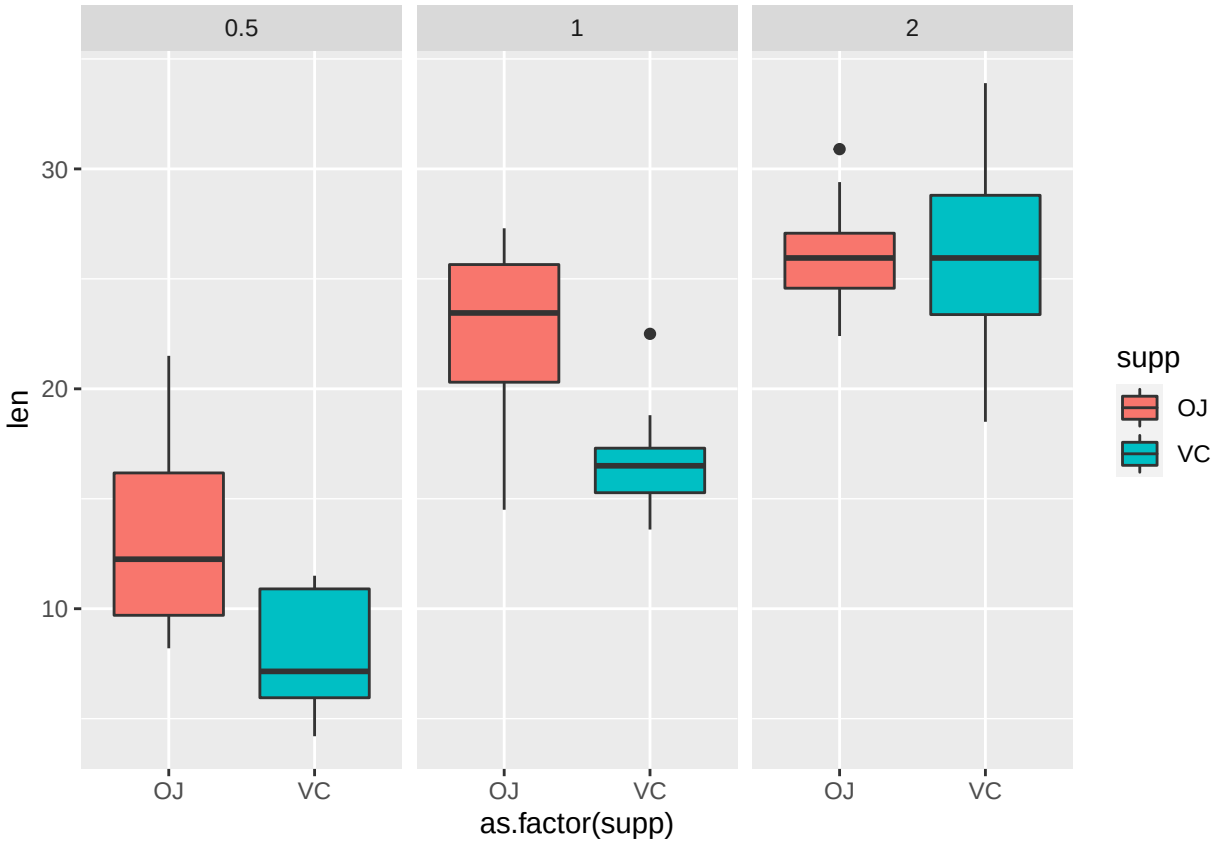
Plotting length vs dose boxplots for different supp

```
ggplot(data = ToothGrowth , aes(x = as.factor(dose),y = len)) + geom_boxplot(aes(fill= dose)) + facet_g
```



length vs supplementary

```
ggplot(data = ToothGrowth , aes(x = as.factor(supp),y = len)) + geom_boxplot(aes(fill= supp)) + facet_g
```



###T test to compare length for different supplements

```
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
```

Since the p value is greater then 0.05 , we can say that the null hypothesis is true

###t test to compare length for different doses

comparing 1.0 and 2.0

```
subsett <- ToothGrowth[ToothGrowth$dose %in% c(1.0,2.0),]
t.test(len~as.factor(dose),data = subsett)
```

```
##
```

```
##  Welch Two Sample t-test
##
## data:  len by as.factor(dose)
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means 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
```

comparing 1.0 and 0.5

```
subsett <- ToothGrowth[ToothGrowth$dose %in% c(1.0,0.5),]
t.test(len~as.factor(dose),data = subsett)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by as.factor(dose)
## 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 in group 0.5   mean in group 1
##            10.605            19.735
```

comparing 2.0 and 0.5

```
subsett <- ToothGrowth[ToothGrowth$dose %in% c(2.0,0.5),]
t.test(len~as.factor(dose),data = subsett)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by as.factor(dose)
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean in group 0.5   mean in group 2
##            10.605            26.100
```

for all the cases the p values is nearly zero . Therefore the null hypothesis can be rejected . Hence tooth length increases with an increase in the dose