Part 2 Appendix

Volha Leusha

```
#Load the ToothGrowth data:
library(datasets)
data("ToothGrowth")

#first glance on 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
```

```
#look into data description:
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
```

```
#check for unique dose values
unique(ToothGrowth$dose)
```

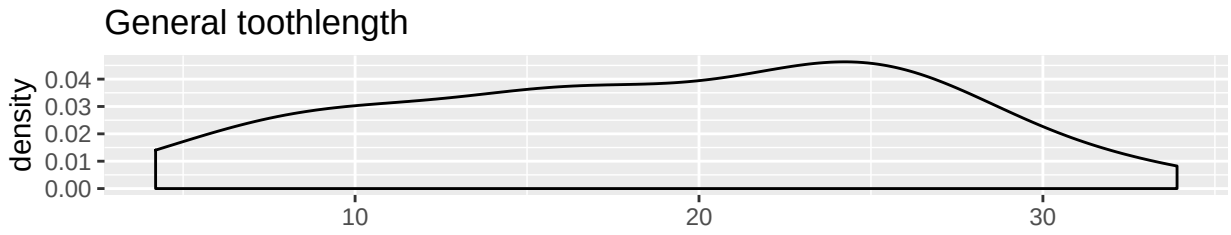
```
## [1] 0.5 1.0 2.0
```

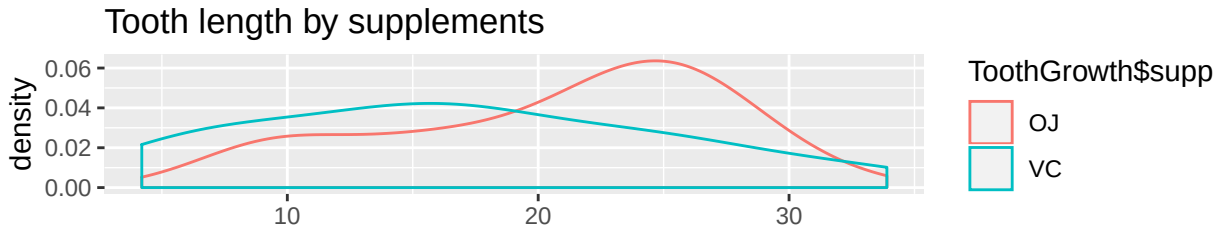
```
#find standard deviation for length
sd(ToothGrowth$len)
```

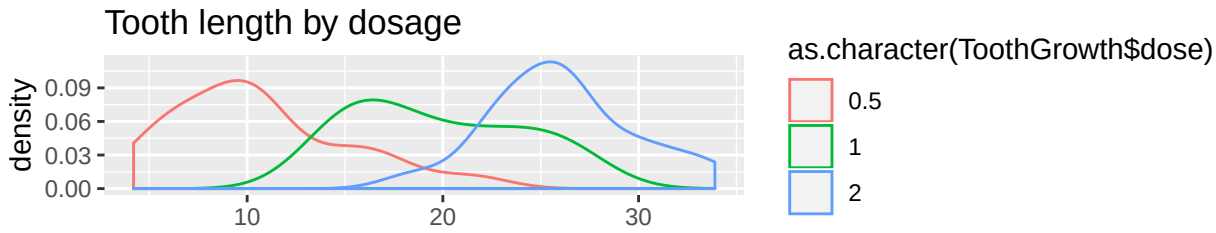
```
## [1] 7.649315
```

```
#open ggplot2 and gripExtra libraries:
library(ggplot2)
library(gridExtra)

#Plot general tooth length:
g1<- ggplot(data=ToothGrowth, aes(ToothGrowth$len))+geom_density()+
    labs(title="General toothlength", x="")

#Plot tooth length by supplements
g2<- ggplot(data=ToothGrowth, aes(ToothGrowth$len, color=ToothGrowth$supp))+
```

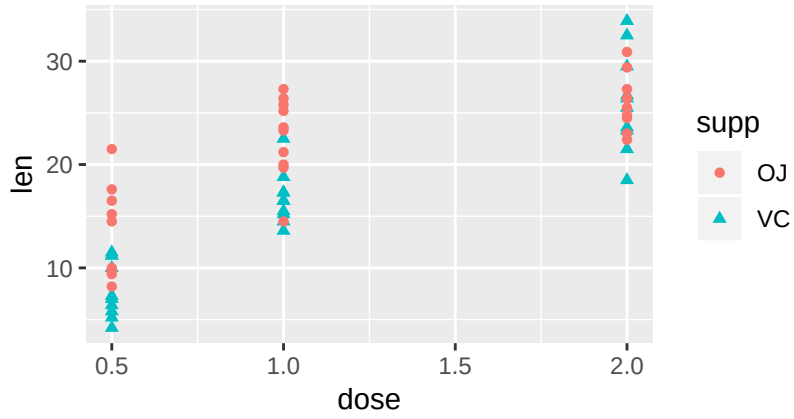
```
    geom_density()+labs(title="Tooth length by supplements ", x="")

#Plot tooth length by dosage
g3<- ggplot(data=ToothGrowth, aes(ToothGrowth$len, color =as.character(ToothGrowth$dose)))+
    geom_density()+labs(title="Tooth length by dosage", x="")

#arrange all three graphs on one page
grid.arrange(g1, g2, g3, nrow = 3)
```








```
#Plot dataset grouped by supplement type and dosage
ggplot(ToothGrowth, aes(x=dose, y=len, shape=supp, color=supp)) +geom_point()+
    theme(plot.margin = margin(3,3,3,3, "cm"))
```




```
#1. First test:
#Ho: there is no difference between tooth length with different supplements
```

```r
#Ha: tooth length for supplement OJ in higher
##------------------------------------------------------------------------------
#create new variable diff1 equal to difference between values of length with OJ
#and VC supplements (each supplement has equal amount of values in dataset= 30):
diff1<-ToothGrowth$len[ToothGrowth$supp=="OJ"]-ToothGrowth$len[ToothGrowth$supp=="VC"]
#perform t test
t.test(diff1, data= ToothGrowth)$conf
```

```
## [1] 1.408659 5.991341
## attr(,"conf.level")
## [1] 0.95
```

```r
#Second test:
#Ho: there is no difference between dose0.5 and 1.0 for tooth length
#Ha: tooth length for dose 1.0 is higher than for dose 0.5
#------------------------------------------------------------------------------
#create new variable diff2 equal to difference between values of dose 1.0
# and dose 0.5 (each dose has equal amount of values in dataset = 20):
diff2<- ToothGrowth$len[ToothGrowth$dose==1]-ToothGrowth$len[ToothGrowth$dose==.5]
#perform t test
t.test(diff2, data= ToothGrowth)
```

```
##
##  One Sample t-test
##
## data:  diff2
## t = 6.9669, df = 19, p-value = 1.225e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   6.387121 11.872879
## sample estimates:
## mean of x
##      9.13
```

```r
# Third test:
#Ho: there is no difference between dose 1.0 and 2.0 for tooth length
#Ha: tooth length for dose 2.0 is higher than for dose 1.0
#------------------------------------------------------------------------------
#create new variable diff2 equal to difference between values of dose 2.0
# and dose 1.0 (each dose has equal amount of values in dataset = 20):
diff3<- ToothGrowth$len[ToothGrowth$dose==2]-ToothGrowth$len[ToothGrowth$dose==1]
#perform t test
t.test(diff3, data= ToothGrowth)
```

```
##
##  One Sample t-test
##
## data:  diff3
## t = 4.6046, df = 19, p-value = 0.0001934
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.471814 9.258186
## sample estimates:
## mean of x
##     6.365
```