Part2project

Simone Van Taylor

8/11/2019

Lets Load In and Explore the Data!!

```
## -- Attaching packages ------------------------------------------------------------------- tidyverse
## v tibble  2.1.1       v purrr   0.3.2
## v tidyr   0.8.3       v dplyr   0.8.0.1
## v readr   1.1.1       v stringr 1.3.1
## v tibble  2.1.1       v forcats 0.3.0
## Warning: package 'tibble' was built under R version 3.5.2
## Warning: package 'tidyr' was built under R version 3.5.2
## Warning: package 'purrr' was built under R version 3.5.2
## Warning: package 'dplyr' was built under R version 3.5.2
## -- Conflicts ---------------------------------------------------------------------- tidyverse_confli
## x tidyr::extract()   masks magrittr::extract()
## x dplyr::filter()    masks stats::filter()
## x dplyr::lag()       masks stats::lag()
## x purrr::set_names() masks magrittr::set_names()
```

```
#load in data
toothgrowth <- ToothGrowth
#explore 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
```

```
unique(toothgrowth$supp)
```

```
## [1] VC OJ
## Levels: OJ VC
```

```
unique(toothgrowth$dose)
```

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

```
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
```

```
#whats the mean of the toothgrowth by supplement?
toothgrowth %>%
  group_by(supp) %>%
  summarise(mean(len))
```

```
## # A tibble: 2 x 2
##   supp  `mean(len)`
##   <fct>       <dbl>
## 1 OJ           20.7
## 2 VC           17.0
```

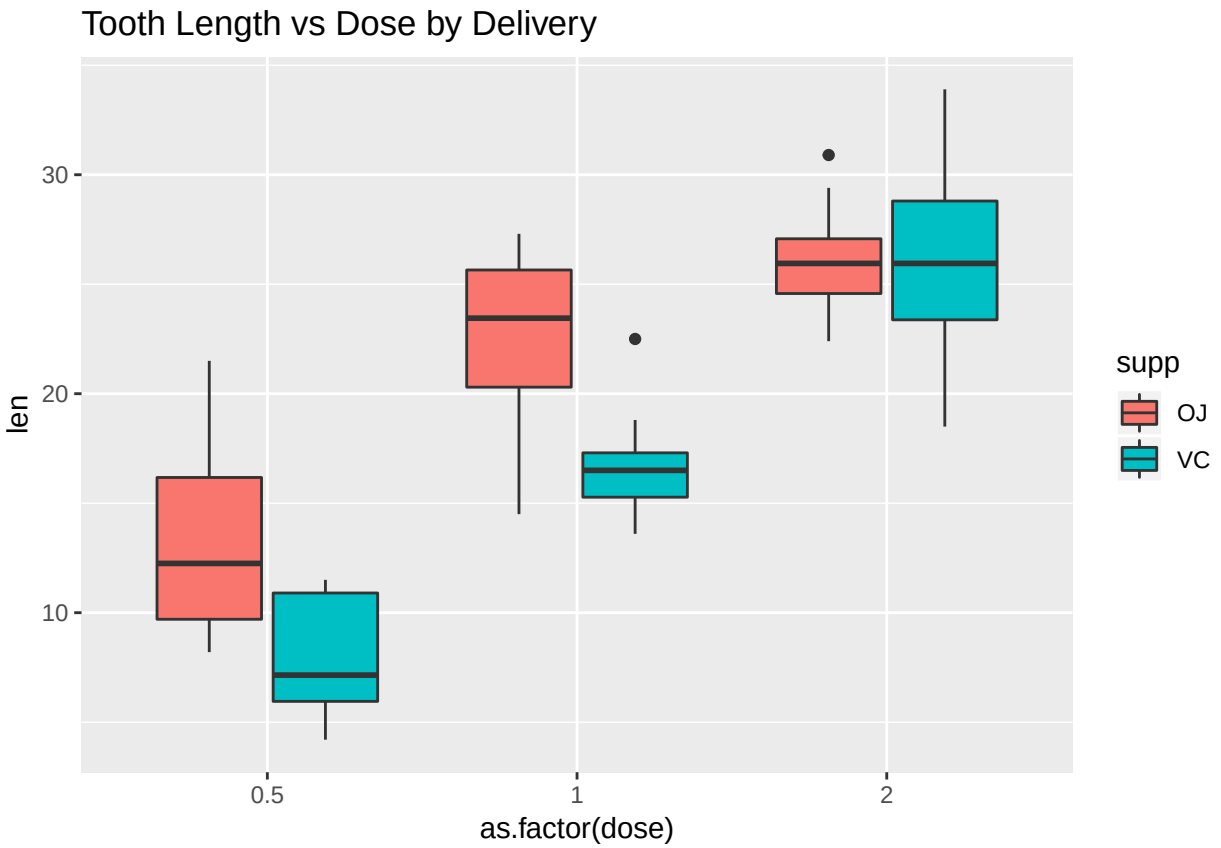
```
##seems like its a little longer for OJ

##whats the mean by dose?
toothgrowth %>%
  group_by(dose) %>%
  summarise(mean(len))
```

```
## # A tibble: 3 x 2
##    dose `mean(len)`
##   <dbl>       <dbl>
## 1   0.5        10.6
## 2   1          19.7
## 3   2          26.1
```

seems like the higher the dose the better

Lets Visualize this Exploratory Analysis

```
#lets visualize this:
toothgrowth %>%
  ggplot(aes(x = as.factor(dose), y = len, fill = supp))+
  geom_boxplot()+
  labs(xlab = "dose", ylab = "length", title = "Tooth Length vs Dose by Delivery")
```



looks like theres a basic trend that as dose goes up length goes up. But is it significant? and is there an effect of supplement?

T Tests

```
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
```

confidence interval crosses 0 and p is not less than .05 therefore we fail to reject the null hypothesis. Looks to be no significant difference between supplement type

```
toothgrowth_sub <- filter(ToothGrowth, ToothGrowth$dose == 1.0|ToothGrowth$dose == 0.5)
t.test(len~dose,data=toothgrowth_sub)
```

```
##
##  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 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
```

```
toothgrowth_sub <- filter(ToothGrowth, ToothGrowth$dose == 0.5|ToothGrowth$dose == 2.0)
t.test(len~dose,data=toothgrowth_sub)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by 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
```

```
toothgrowth_sub <- filter(ToothGrowth, ToothGrowth$dose == 1.0|ToothGrowth$dose == 2.0)
t.test(len~dose,data=toothgrowth_sub)
```

```
##
##  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 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
```

for all t tests the confidence int did not pass through zero and p was less than .05 therefore we reject the null hypothesis and can conclude dose significantly effected tooth length