Assignment-45488479

2022-10-28

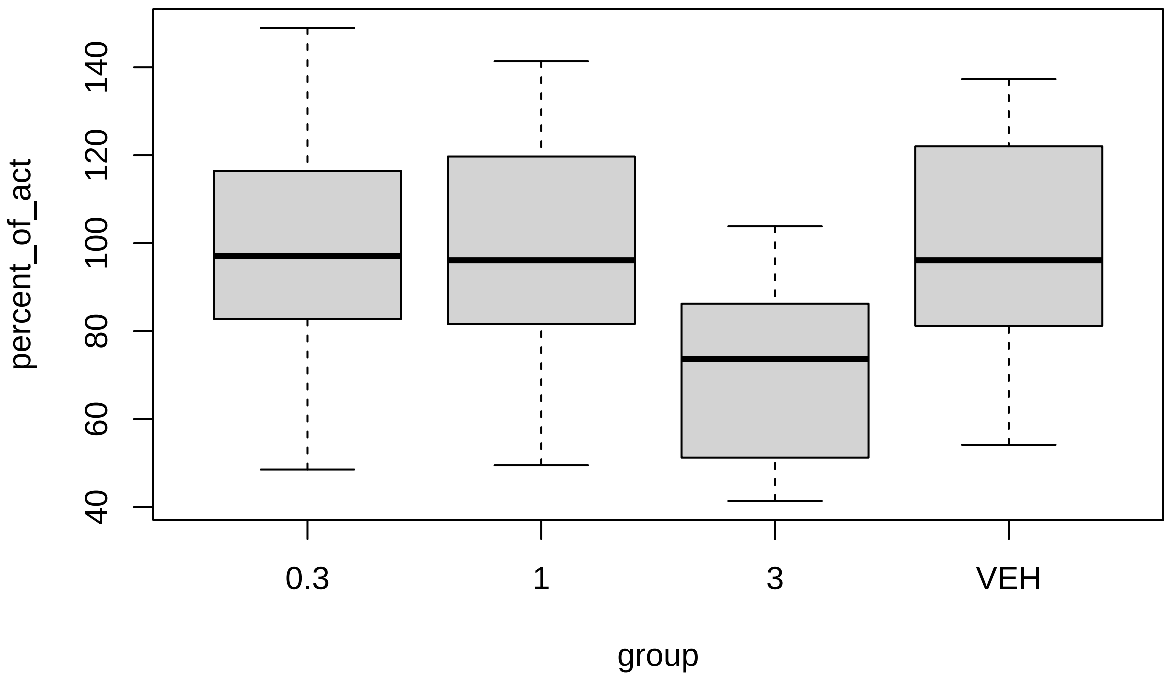
```r
library("tinytex")
library("tidyverse")
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.8     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
mice <- read.table("mice_pot.txt", header = TRUE)
head(mice)
```

```
##   group percent_of_act
## 1   0.3       98.82671
## 2   0.3      116.42599
## 3   0.3      132.09128
## 4   0.3       84.12842
## 5   0.3      148.91697
## 6   0.3       48.54306
```

```r
boxplot(percent_of_act ~ group, data = mice)
```



#a. ##The variances look larger in the 0.3mg/kg plot than 3mg/kg. ##The boxes appear to be equally sized implying that the different formations appear to have the same variability. ## Including Plots

You can also embed plots, for example:

```
mice.aov = aov(percent_of_act ~ group, data = mice)
anova(mice.aov)
```

```
## Analysis of Variance Table
##
## Response: percent_of_act
##           Df Sum Sq Mean Sq F value Pr(>F)
## group      3   6329 2109.65  3.1261 0.0357 *
## Residuals 42  28344  674.85
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##b. ##• Hypotheses: H0 :mu1 =mu2 =mu3 =mu4(no effect in the population); H1 :not all means are equal(an effect in the population). ##• Test statistic: Fobs = 3.1261 ##• Null distribution: If H0 is true, Fobs behaves like a F3,42 distribution ##• P-Value: P (F3,42 >= 3.1261) = 0.0357 < 0.05 ##• Conclusion: Since the P-Value is less than the significance level of 0.05 (5%) we have evidence to reject H0 in favour of H1. That is, we have evidence that the mice content is not all the same for the different formations.

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

```r
TukeyHSD(mice.aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = percent_of_act ~ group, data = mice)
##
## $group
##                  diff         lwr       upr     p adj
## 1-0.3       1.7298436 -28.9122207 32.371908 0.9987527
## 3-0.3     -26.6546330 -58.5829330  5.273667 0.1309483
## VEH-0.3     2.6774970 -26.6219205 31.976915 0.9947868
## 3-1       -28.3844765 -58.1381757  1.369223 0.0663434
## VEH-1       0.9476534 -25.9655801 27.860887 0.9996956
## VEH-3      29.3321300   0.9630909 57.701169 0.0403151
```

##c. ##t-test value is 3.1262. and assciated degrees of freedom is F3,42. ##Although P-vlaue = 0.0357 < 0.05, the H0 is rejected.

```r
kml <- read.table("kml.dat", header = TRUE)
head(kml)
```

```
##        kmL driver   car
## 1 10.75614      A   one
## 2 10.71363      A   one
## 3 12.28666      A   two
## 4 12.75432      A   two
## 5 10.54357      A three
## 6 10.67111      A three
```

```r
summary(aov(kmL ~ driver * car, data = kml))
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## driver       3  50.66  16.887  531.60  < 2e-16 ***
## car          4  17.12   4.280  134.73 3.66e-14 ***
## driver:car  12   0.44   0.037    1.16    0.371
## Residuals   20   0.64   0.032
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(aov(kmL ~ car * driver, data = kml))
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## car          4  17.12   4.280  134.73 3.66e-14 ***
## driver       3  50.66  16.887  531.60  < 2e-16 ***
## car:driver  12   0.44   0.037    1.16    0.371
## Residuals   20   0.64   0.032
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
checkbalance = kml; checkbalance[15, ] = c(90, "truck", "km")

summary(aov(kmL ~ driver * car, data=checkbalance))
```

```
##             Df Sum Sq Mean Sq   F value   Pr(>F)    
## driver       4   5869    1467 44128.610  < 2e-16 ***
## car          4     16       4   119.828 3.37e-13 ***
## driver:car  12      0       0     1.061     0.44    
## Residuals   19      1       0                       
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

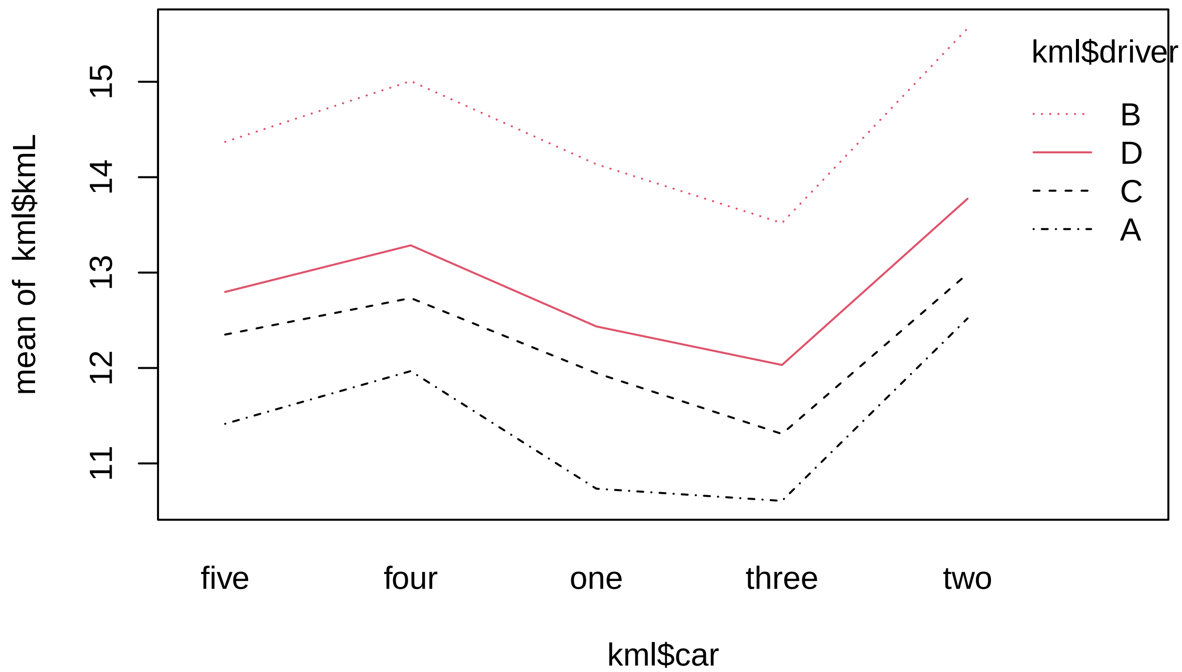
```r
summary(aov(kmL ~ car * driver, data=checkbalance))
```

```
##             Df Sum Sq Mean Sq   F value Pr(>F)    
## car          5   5837  1167.4 35111.696 <2e-16 ***
## driver       3     48    15.9   478.424 <2e-16 ***
## car:driver  12      0     0.0     1.061   0.44    
## Residuals   19      1     0.0                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

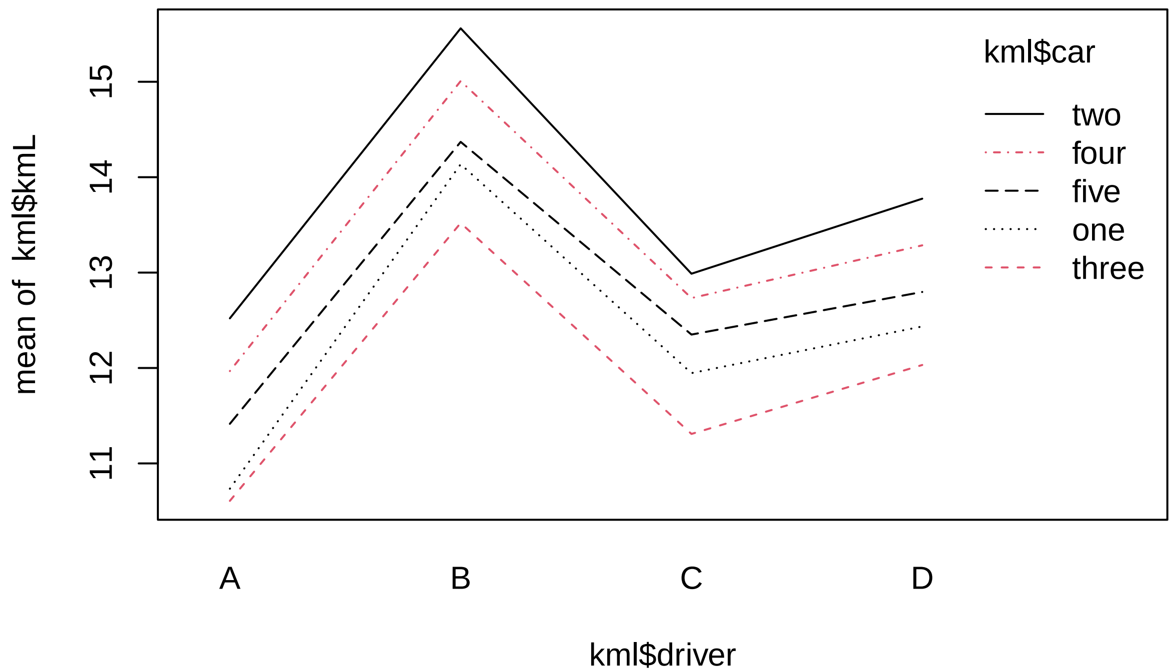
##Q2 ##a.Yes the design is balanced since there are two entries for a combination of each of the Car and driver

##b.

```r
interaction.plot(kml$car, kml$driver, kml$kmL, col = 1:2)
```




```
interaction.plot(kml$driver, kml$car, kml$kmL, col = 1:2)
```




```
kml1 = lm(kmL ~ driver, data = kml)
anova(kml1)
```

```
## Analysis of Variance Table
##
## Response: kmL
##           Df Sum Sq Mean Sq F value    Pr(>F)
## driver     3 50.661 16.8869  33.409 1.672e-10 ***
## Residuals 36 18.197  0.5055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
kml2 = lm(kmL ~ car, data = kml)
anova(kml2)
```

```
## Analysis of Variance Table
##
## Response: kmL
##           Df Sum Sq Mean Sq F value  Pr(>F)
## car        4 17.119  4.2798  2.8952 0.03597 *
## Residuals 35 51.738  1.4782
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##c. ##Null hypothesis : There is no Dependence of drivers. ##Alternative hypothesis : There is some dependence on drivers and fuel efficiencies. ##From kml1, t-stat is 33.409, P-vlaue is 1.672e-10. ##Since P-value is less than 0.05, null hypothesis is rejected.

##Null hypothesis: There is no dependence on cars and fuel efficiencies. ##Alternate hypothesis:There is some dependence on cars and fuel efficiencies. ##From kml2, t-stat is 2.8952, P-value is 0.03597. ##Since P-value is less than 0.05, null hypothesis is rejected.

##d ##In both the Cases null hypothesis is rejected (See Q2-c). ##This implies that Cars and drivers on fuel efficiencies choice of in Kml.