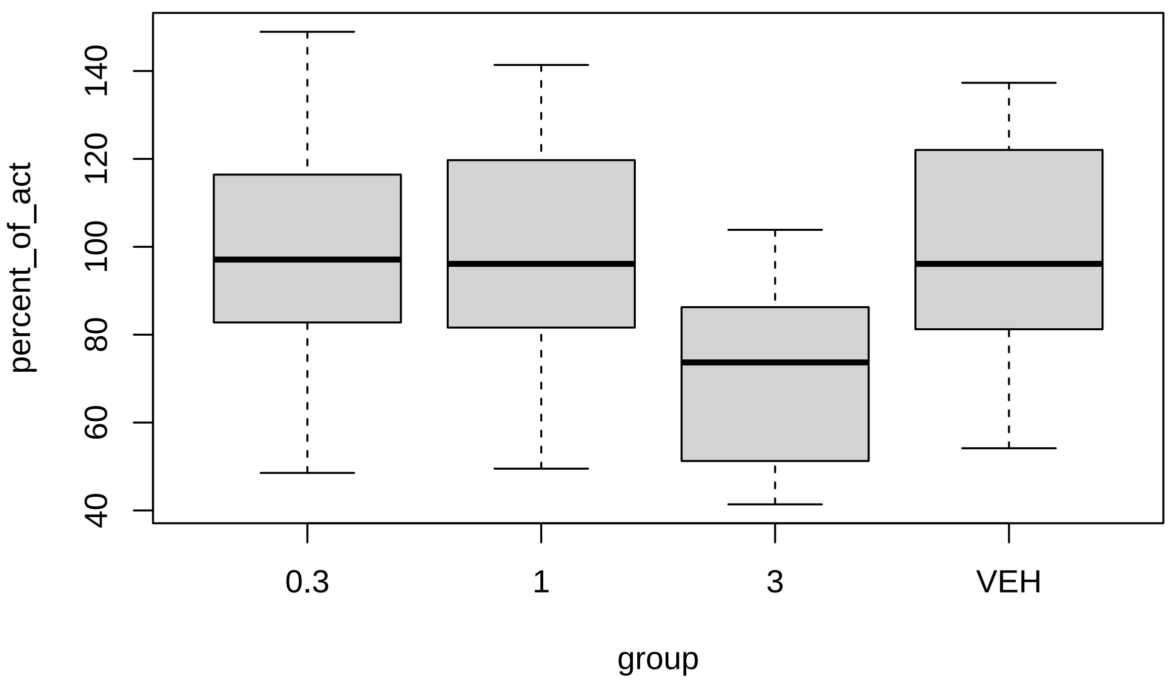
Assignment

Madalene_vanKoeverden_47098406

2022-10-17

#Q1a

```
dat=read.table("mice_pot.txt", header = TRUE)
boxplot(percent_of_act~group, data =dat)
```



The median for three groups (0.3 mg/kg, 1mg/kg and VEH) is approximately the same, although the range is not. Both the median and range of 3 mg/kg group is smaller than the other three groups. From these observations this suggests the higher dosage of 3 mg/kg THC decreases the total distance covered.

#Q1b

Null H0: Mean of percent_of_act is the same for all groups. (i.e. there is no difference in activities)

Alternative H1: Mean of percent_of_act is not the same for all groups. (i.e. there is a difference in activities)

```r
mice.aov=aov(percent_of_act~group,data=dat)
summary(mice.aov)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## group        3   6329  2109.7   3.126 0.0357 *
## Residuals   42  28344   674.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test Stat F obs = 3.126

Critical Value, if H0 is true Fobs behaves like F3,42 distribution. Using F table F3,40 for 0.05 sig = 2.84

P Value: P (F3,42 >_3.126)= 0.0357 <0.05

Conclusion: Reject H0 in favour of H1, because the P value is less than the significance level (5%), F obs > Critical Value. Therefore, there is evidence that there is a difference in activity levels across the groups.

#Q1c

General Contrast,dosages of 0.3, 1 and VEH as a group, compared to dosage of 3.

Null H0: C = 0 (where C denotes a contrast that is C = mu3- (mu0.3+mu1+muVEH)/3)

Alternative H1: C does not = 0

Test Statistic: tobs = c/s.e.(c)

```r
summary.lm(mice.aov)
```

```
##
## Call:
## aov(formula = percent_of_act ~ group, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -49.542 -17.907  -2.569  19.706  51.594
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   97.323      8.659  11.239 3.06e-14 ***
## group1         1.730     11.455   0.151   0.8807
## group3       -26.655     11.936  -2.233   0.0309 *
## groupVEH       2.677     10.953   0.244   0.8081
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 25.98 on 42 degrees of freedom
## Multiple R-squared:  0.1825, Adjusted R-squared:  0.1241
## F-statistic: 3.126 on 3 and 42 DF,  p-value: 0.0357
```

c = -26.655 + 33.91 = 7.255

s.e(c)=-2.789

T_obs = 7.255/-2.789 = -2.601

Critical Value Tn-g = T42 approx = 2.021

P-Value: P(|t42|>_ |-2.601290785|)= 2P(t42>_ 2.601290785)

```r
2*pt(2.601,42, lower.tail=FALSE)
```

```
## [1] 0.01278013
```

=0.0127

Conclusion: Reject H0 in favour of H1, because the P value is less than the significance level and |Tobs| is greater than T42(2.601 > 2.021). Therefore, using general contrast there is evidence that C does not =0. That is the mu3 is statistically significant (different from mu0.3, mu1 and muVEH).

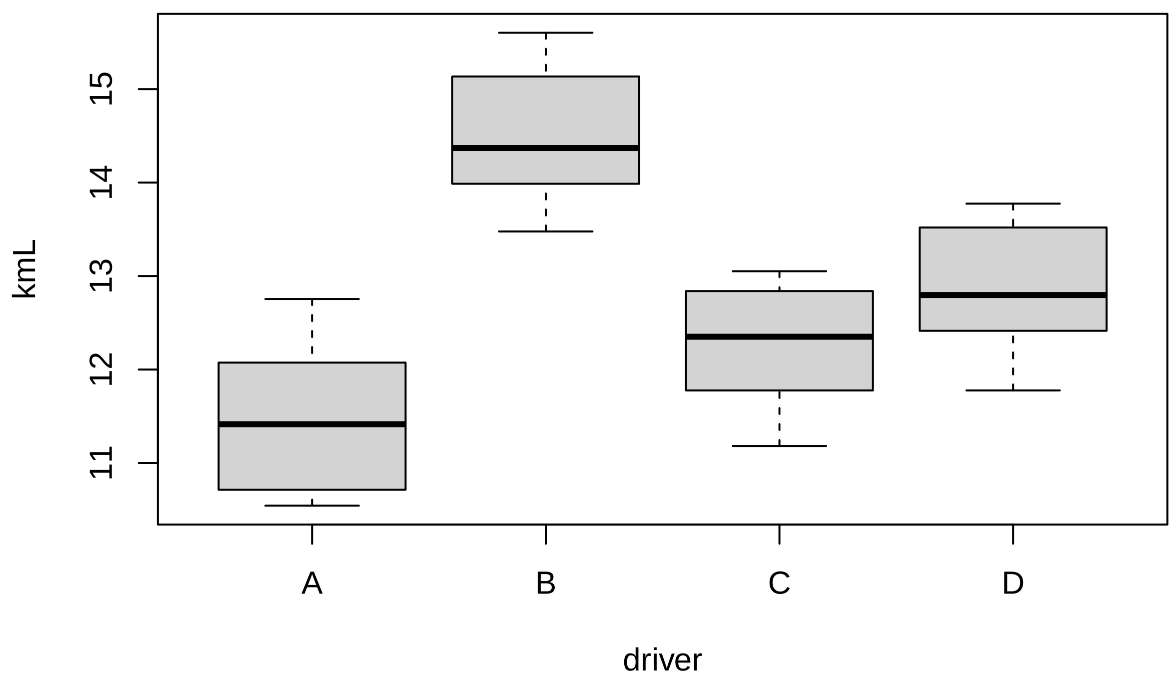
#Q2a

A study with the same number of replicates for each treatment is said to be balanced. Therefore, this study design is balanced because there are five cars and four drivers. Each drivers completes 10 drives, for a total of 40 drives. 2 drives per car, per driver (i.e., driver A drives car 1, twice and so on etc).

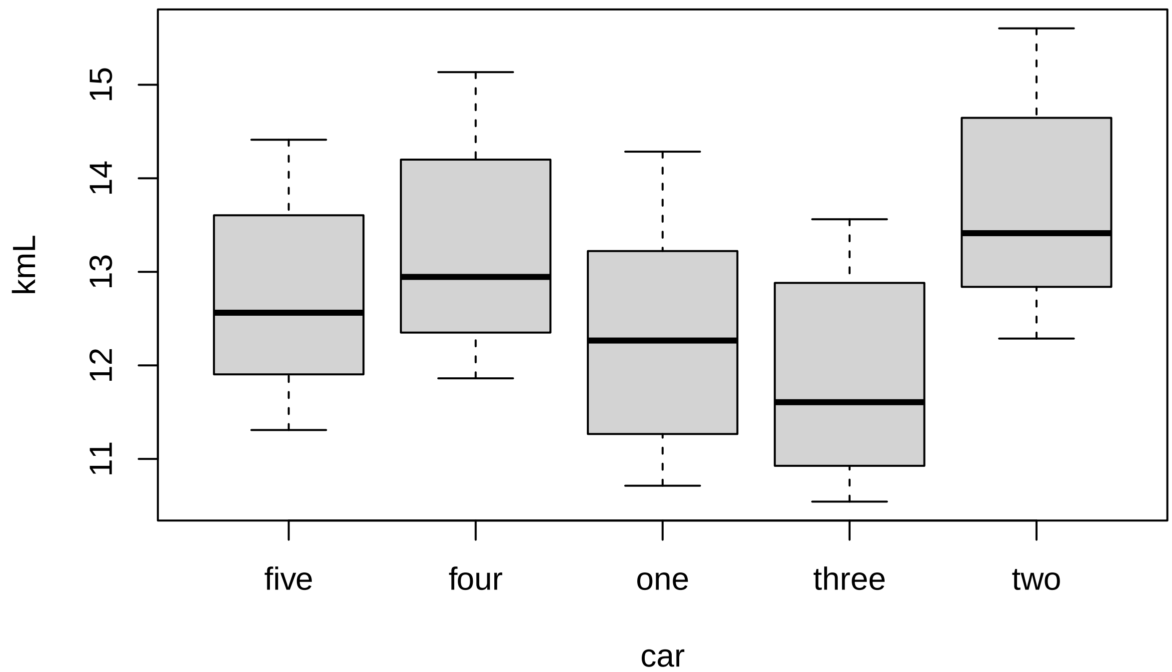
#Q2B

Testing the differences in fuel efficiency by car (i.e., do all of the 5 cars have the same fuel efficiency)and differences in fuel efficiency by driver (i.e., do all of the 4 drivers have the same fuel efficiency)

```r
dat.1=read.table("kml.dat", header =TRUE)
boxplot(kmL~driver, data = dat.1)
```




```r
boxplot(kmL~car, data=dat.1)
```



The boxplot comparing drivers, suggests there is significant difference from driver to driver relating to fuel efficiency. For example, driver A has a much lower range of fuel efficiency than driver B.

The boxplot comparing cars, suggests there is significant difference in fuel efficiency from car to car. However, in comparison to the previous boxplot the spread of data seems closer together, i.e. there is a smaller difference in fuel efficiency between cars and a larger difference between drivers.

#Q2c

Testing difference in fuel efficiency between drivers.

Null H0: Mean of kmL is the same for all drivers.

Alternative H1: Mean of kmL is not the same for all drivers.

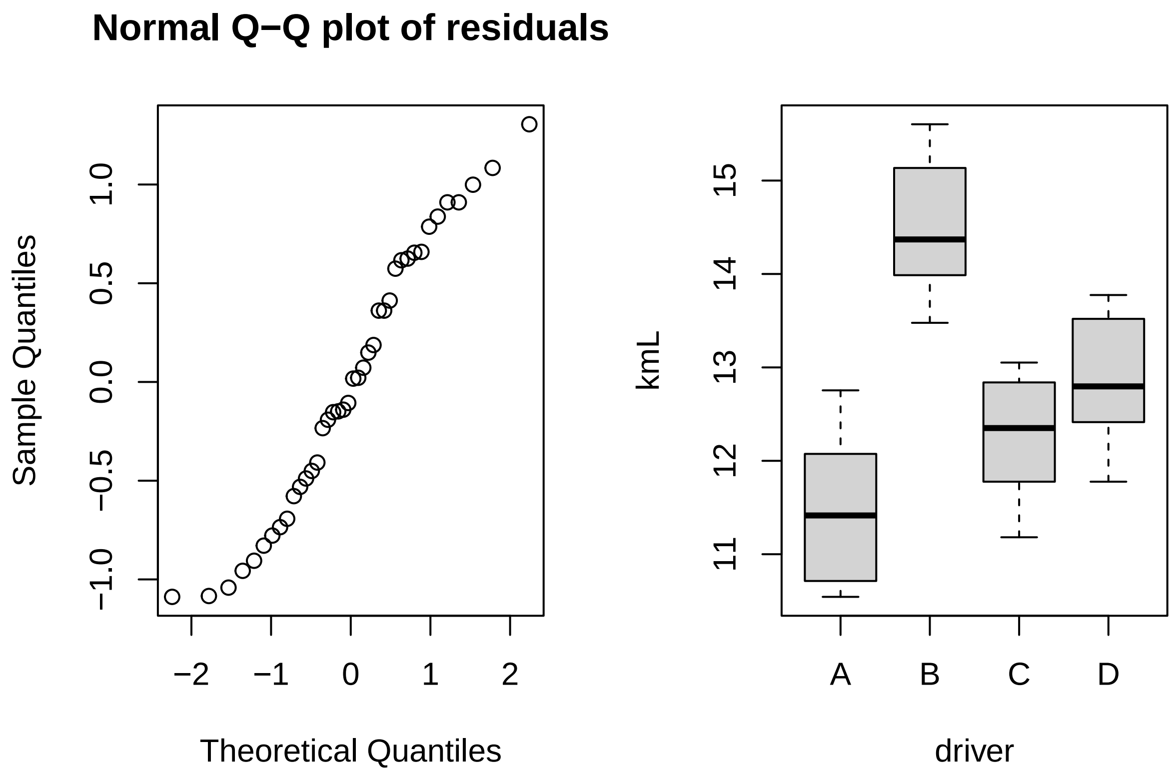
Assumptions: check difference in variance with boxplot & normality of residuals.

```
driver.aov=aov(kmL~driver, data=dat.1)

par(mfrow=c(1,2))

qqnorm(driver.aov$residuals, main ="Normal Q-Q plot of residuals")

boxplot(kmL~driver, data = dat.1)
```




```
summary(driver.aov)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## driver        3  50.66  16.887   33.41 1.67e-10 ***
## Residuals    36  18.20   0.505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test Statistic: Fobs = 33.41

Critical Value: if H0 is true Fobs behaves like F3,36 distribution.

P Value: P (F3,36 >_ 33.41)= 0.0000000000167 <0.05

Conclusion: Reject H0 in favour of H1, because the P value is less than the significance level (5%). Therefore, there is evidence that there is a difference in fuel efficiency among the drivers.

Testing difference in fuel efficiency between cars.

Null H0: Mean of kmL is the same for all cars.

Alternative H1: Mean of kmL is not the same for all cars.

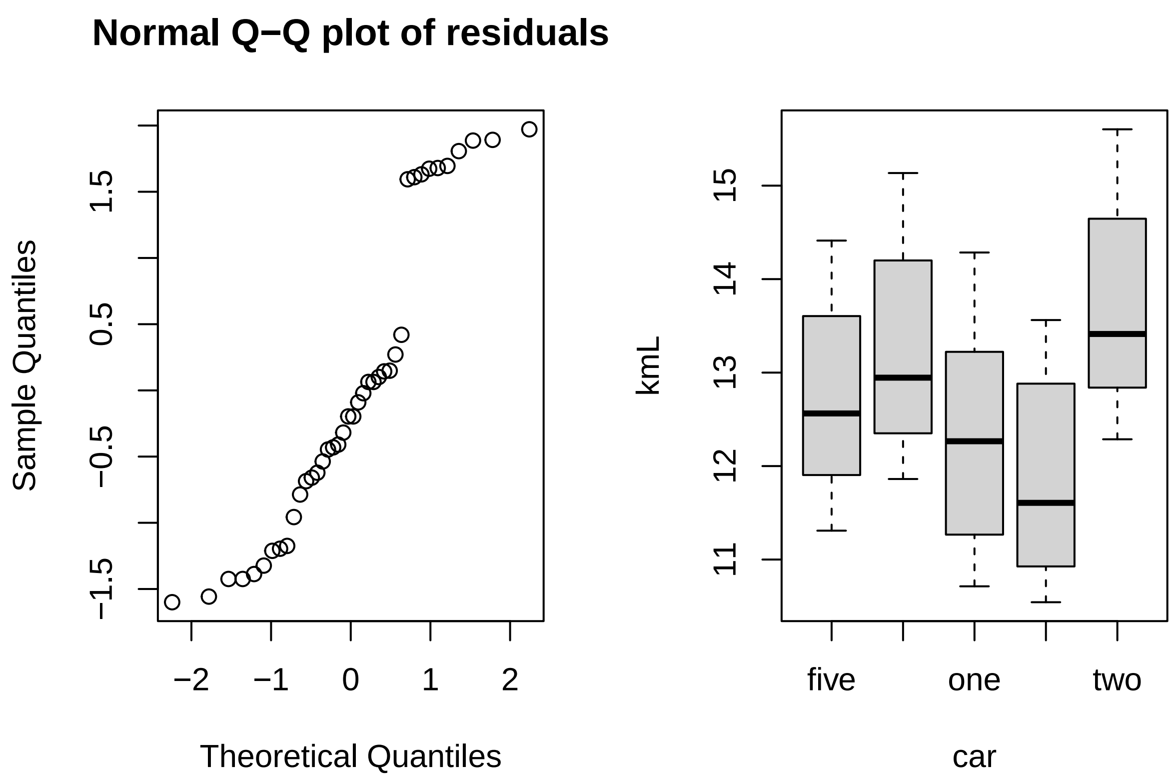
Assumptions: check difference in variance with boxplot & normality of residuals.

```
car.aov=aov(kmL~car, data=dat.1)

par(mfrow=c(1,2))
```

```r
qqnorm(car.aov$residuals, main ="Normal Q-Q plot of residuals")

boxplot(kmL~car, data = dat.1)
```




```r
summary(car.aov)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## car          4  17.12   4.280   2.895  0.036 *
## Residuals   35  51.74   1.478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test Statistic: Fobs = 2.895

Critical Value: if H0 is true Fobs behaves like F4,35 distribution.

P Value: P (F4,35 >_ 2.895)= 0.036 <0.05

Conclusion: Reject H0 in favour of H1, because the P value is less than the significance level (5%). Therefore, there is evidence that there is a difference in fuel efficiency among the cars.

#Q2d

Both the driver and the car impact the level of fuel efficiency. That is the observed fuel efficiency is different for each driver and each car (i.e., fuel efficiency is not the same among the five cars or four drivers).