On réalise une analyse de variance (ANOVA) sur la base de données mtcars disponible dans R.

```
# Chargement des données
data(mtcars)

# affichage des premières lignes
head(mtcars)
                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1

# Transformation des facteurs
mtcars$am <- factor(mtcars$am, labels = c("Automatique", "Manuelle"))
mtcars$cyl <- factor(mtcars$cyl)

# Vérification
table(mtcars$am, mtcars$cyl)

               4  6  8
  Automatique  3  4 12
  Manuelle     8  3  2
```

Interprétation de la vérification des conditions de l'ANOVA :

Avant d’interpréter les résultats de l’ANOVA, il faut vérifier les conditions nécessaires à sa validité :

```
# Test de normalité des résidus
shapiro.test(residuals(anova_model))

        Shapiro-Wilk normality test

data:  residuals(anova_model)
W = 0.96277, p-value = 0.3263

# test d'independance
durbinWatsonTest(anova_model)
 lag Autocorrelation D-W Statistic p-value
   1      0.07110978      1.670616   0.188
 Alternative hypothesis: rho != 0
# test d'Homogenité
bartlett.test(mpg ~ interaction(am, cyl), data = mtcars)

        Bartlett test of homogeneity of variances

data:  mpg by interaction(am, cyl)
Bartlett's K-squared = 10.341, df = 5, p-value = 0.06614
```

Normalité des résidus :

Le test de Shapiro-Wilk appliqué aux résidus donne une p-value de 0.3263 (> 0.05). L'hypothèse de normalité est donc acceptée.

Indépendance des résidus :

Le test de Durbin-Watson donne une p-value de 0.188 (> 0.05). Cela indique que les résidus suivent une loi normale.

Homogénéité des variances :

Le test de Bartlett donne une p-value de 0.06614 (> 0.05). L'hypothèse d' homogénéité est donc acceptée.

Conclusion :

Les conditions principales de validité de l'ANOVA sont globalement respectées. Nous pouvons donc interpréter les résultats de l'ANOVA de manière fiable.

```
summary(anova_model)
            Df Sum Sq Mean Sq F value    Pr(>F)    
am           1  405.2   405.2  44.064 4.85e-07 ***
cyl          2  456.4   228.2  24.819 9.35e-07 ***
am:cyl       2   25.4    12.7   1.383    0.269    
Residuals   26  239.1     9.2                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

L'ANOVA à deux facteurs (avec interaction) a été réalisée pour étudier l'effet du type de boîte de vitesses (am) et du nombre de cylindres (cyl) sur la consommation (mpg).

Effet principal de am (boîte de vitesses) :

Le type de boîte a un effet significatif sur la consommation (p-value = 4.85e-07). Cela signifie que la consommation moyenne diffère significativement entre les véhicules automatiques et manuels.

Effet principal de cyl (nombre de cylindres) :

Le nombre de cylindres influence également significativement la consommation (p-value = 9.35e-07). Les véhicules ayant un nombre différent de cylindres ont des consommations moyennes différentes.

Interaction am:cyl :

L'interaction entre le type de boîte et le nombre de cylindres n'est pas significative (p-value = 0.269). Cela signifie que l'effet de la boîte sur la consommation ne dépend pas significativement du nombre de cylindres.

Classer les traitements par groupes qui sont siginficativement different (le test Tukey)

```
tukey <- TukeyHSD(anova_model)
print(tukey)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = mpg ~ am * cyl, data = mtcars)

$am
                         diff     lwr      upr p adj
Manuelle-Automatique 7.244939 5.00149 9.488388 5e-07

$cyl
         diff        lwr        upr     p adj
6-4 -4.756706  -8.399753 -1.1136596 0.0087506
8-4 -7.329581 -10.365453 -4.2937086 0.0000072
8-6 -2.572874  -6.060826  0.9150773 0.1788313

$`am:cyl`
                                  diff        lwr        upr     p adj
Manuelle:4-Automatique:4      5.175000  -1.132282 11.4822821 0.1546661
Automatique:6-Automatique:4  -3.775000 -10.890574  3.3405739 0.5871784
Manuelle:6-Automatique:4     -2.333333  -9.940202  5.2735351 0.9315095
Automatique:8-Automatique:4  -7.850000 -13.863758 -1.8362425 0.0054390
Manuelle:8-Automatique:4     -7.500000 -16.004737  1.0047375 0.1072775
Automatique:6-Manuelle:4     -8.950000 -14.655151 -3.2448487 0.0006955
Manuelle:6-Manuelle:4        -7.508333 -13.815615 -1.2010512 0.0129262
Automatique:8-Manuelle:4    -13.025000 -17.277369 -8.7726313 0.0000000
Manuelle:8-Manuelle:4       -12.675000 -20.040319 -5.3096813 0.0002083
Manuelle:6-Automatique:6      1.441667  -5.673907  8.5572406 0.9883098
Automatique:8-Automatique:6  -4.075000  -9.453868  1.3038683 0.2192160
Manuelle:8-Automatique:6     -3.725000 -11.793302  4.3433024 0.7158963
Automatique:8-Manuelle:6     -5.516667 -11.530424  0.4970909 0.0859484
Manuelle:8-Manuelle:6        -5.166667 -13.671404  3.3380708 0.4436999
Manuelle:8-Automatique:8      0.350000  -6.765574  7.4655739 0.9999875
```

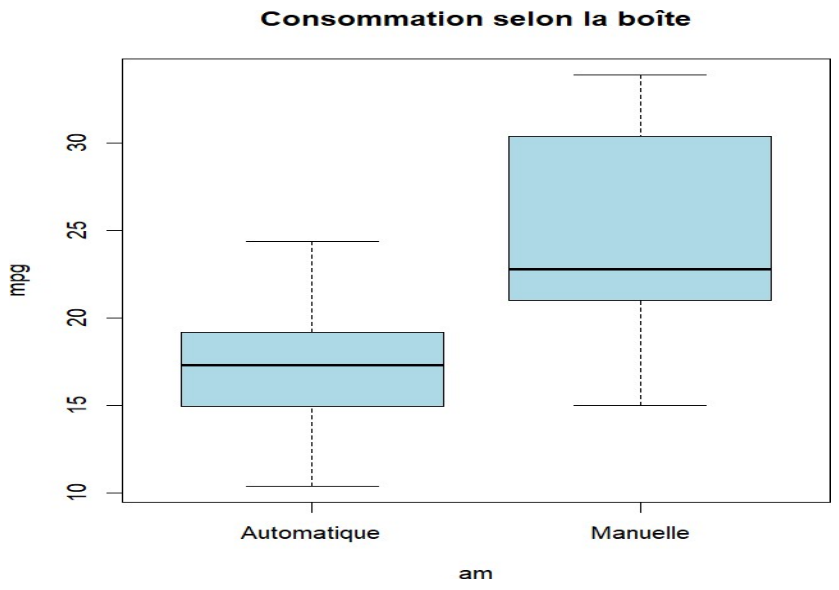
Le type de transmission a un effet significatif : les voitures manuelles consomment moins (plus de mpg) que les automatiques.

Le nombre de cylindres a aussi un effet significatif : plus de cylindres = plus de consommation.

L'interaction entre transmission et cylindres n'est pas significative globalement, mais quelques comparaisons montrent que les voitures automatiques 8 cylindres sont particulièrement inefficaces en consommation.

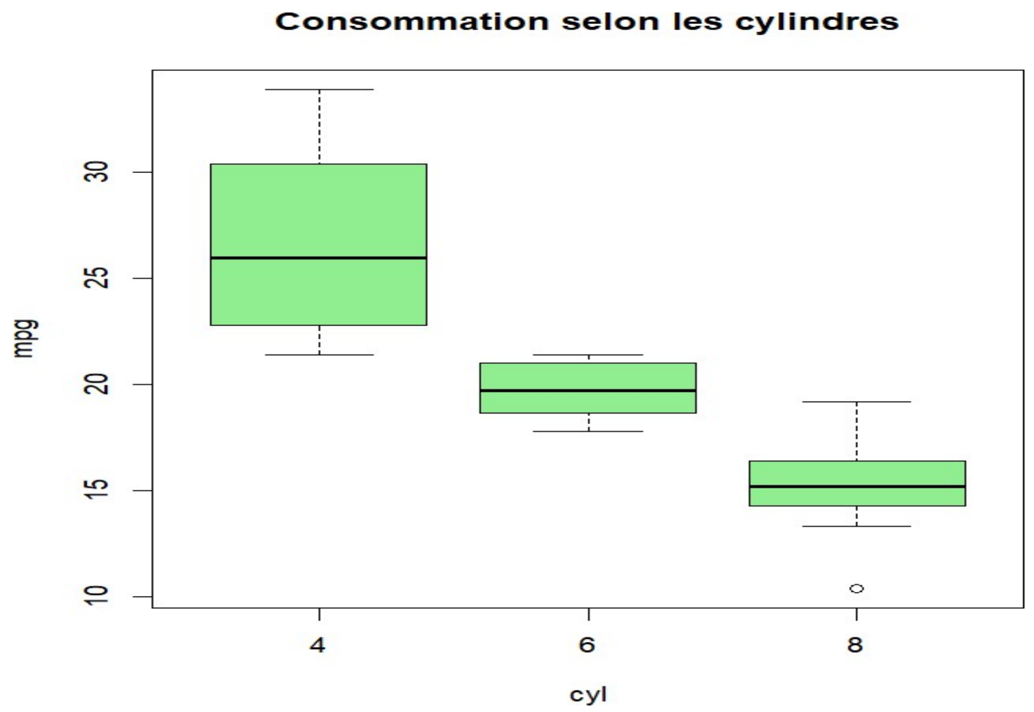
Représentation Graphique

```
# Visualisation 1 : Boxplot de mpg selon le type de boîte
boxplot(mpg ~ am, data = mtcars, main = "Consommation selon la boîte", col
="lightblue")
```



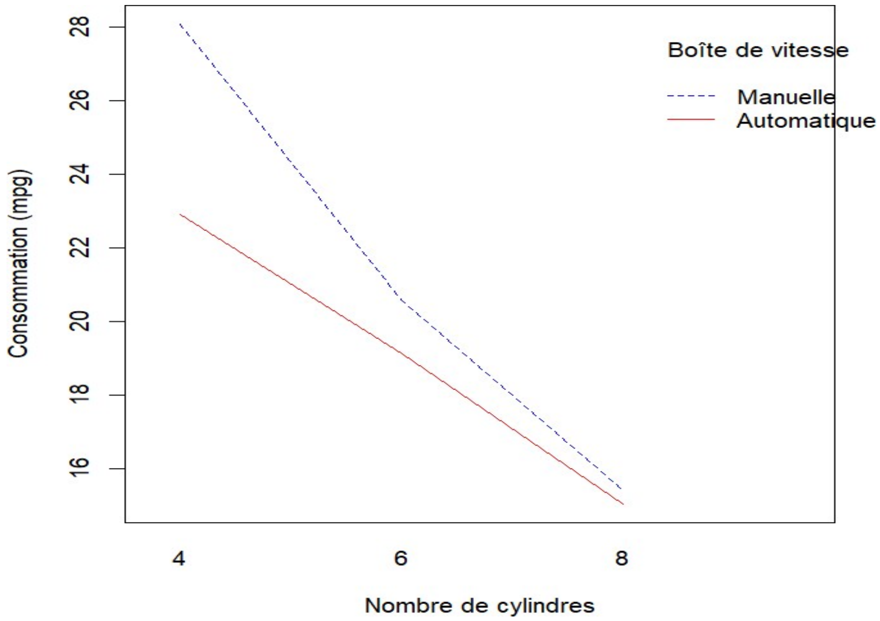
Le boxplot montre que les voitures manuelles ont en moyenne une meilleure consommation (plus haut mpg) que les voitures automatiques.

```
# Visualisation 2 : Boxplot de mpg selon le nombre de cylindres
boxplot(mpg ~ cyl, data = mtcars, main = "Consommation selon les cylindres
", col="lightgreen")
```



Le nombre de cylindres est inversement lié à la consommation d’essence : plus il y a de cylindres, plus le véhicule consomme.

```
# Visualisation 3 : Interaction Plot
interaction.plot(mtcars$cyl, mtcars$am, mtcars$mpg, col=c("red", "blue"),
lty=1:2,xlab="Nombre de cylindres", ylab="Consommation (mpg)", trace.label
="Boîte de vitesse")
```



La consommation augmente avec le nombre de cylindres, indépendamment du type de transmission. La boîte manuelle reste plus économe que l'automatique pour chaque nombre de cylindres.