ANOVA

Maksym Rud'

03 11 2020

Two-way ANOVA

Source: W.M. To, T.W. Yu, T.M. Lai, S.P. Li (2007). "Characterization of Commercial Clothes Dryers Based on Energy-Efficiency Analysis," International Journal of Clothing Science and Technology, Vol. 19, #5, pp. 270-290.

Desription: Replicates each of 3 clothing categories (1=Towels, 2=Jeans, 3=Thermal Clothing) for 4 dryer types (1=Electric Dryer, 2=Bi-directional Electric dryer, 3=Town Gas-Fired Dryer, 4=LPG-Fired dryer)

First of all, here are first 5 rows of the data

```
head(drying_data)
```

```
##   Cloth_cat dry_type energ_effect cycle_time
## 1    Towels   Electr        1.157      25.83
## 2    Towels   Electr        1.189      26.28
## 3    Towels   Electr        1.190      26.05
## 4    Towels   Bi-dir        1.236      26.61
## 5    Towels   Bi-dir        1.244      27.37
## 6    Towels   Bi-dir        1.264      27.08
```

Frequency table

We have 3X4 design cells with the factors being clothing categories and dryer types and 3 subjects in each cell. Here, we have a balanced design. In the next sections I'll analyse data from balanced designs

```
##
##                       Electr Bi-dir Gas-Fired LPG-Fired
##   Towels                   3      3         3         3
##   Jeans                    3      3         3         3
##   Thermal Clothing         3      3         3         3
```

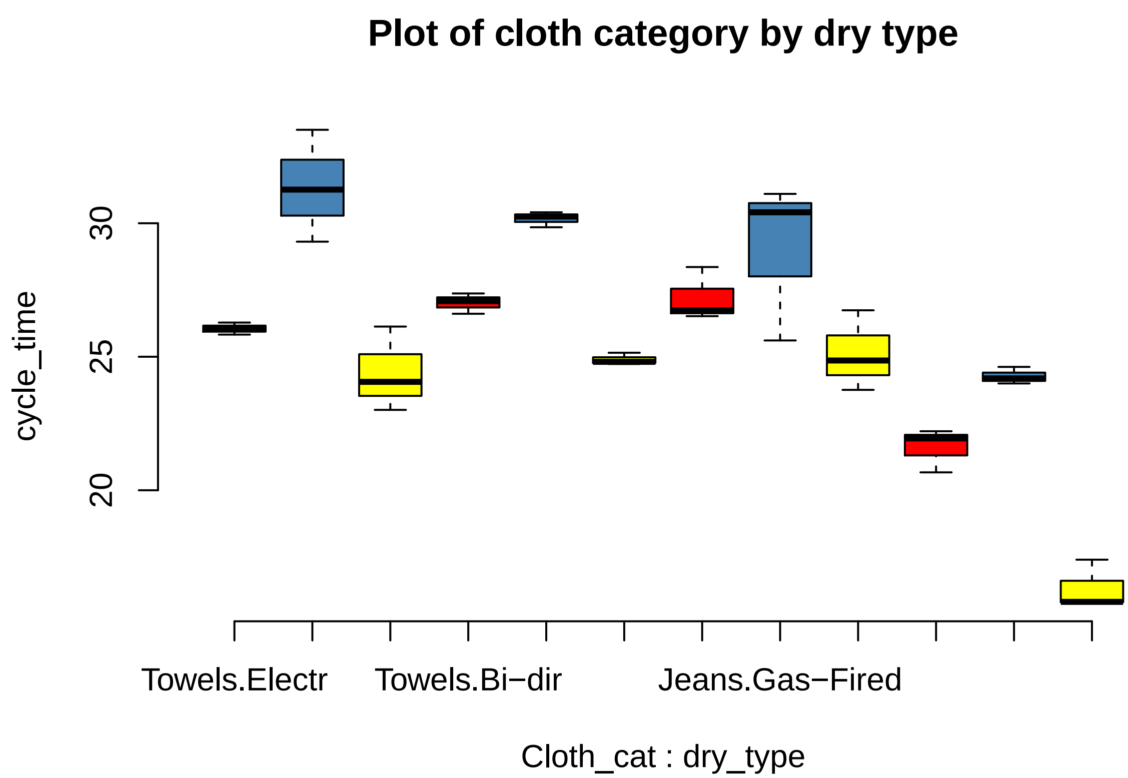
Visualize data

Two-way interaction plot, which plots the mean (or other summary) of the response for two-way combinations of factors, thereby illustrating possible interactions.

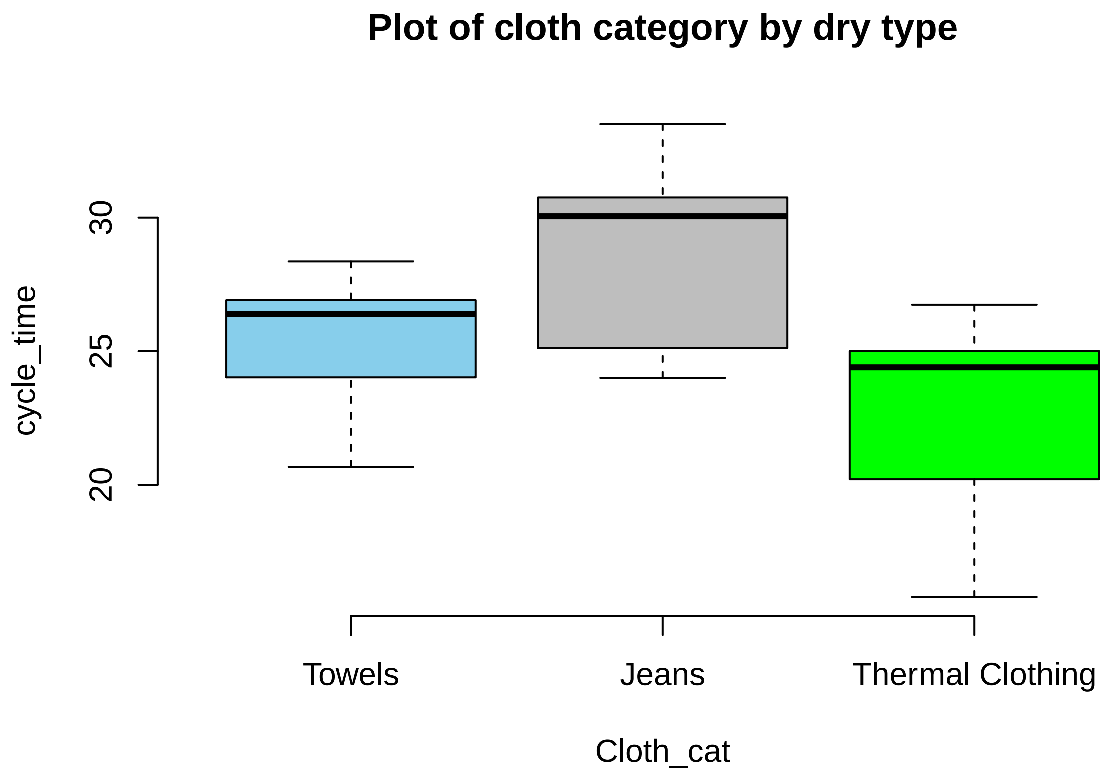
Where red is for Towels, Blue is for Jeans and Green - Thermal Clothing

Here we can see some descending tendency on time for spending for a drying cycle.

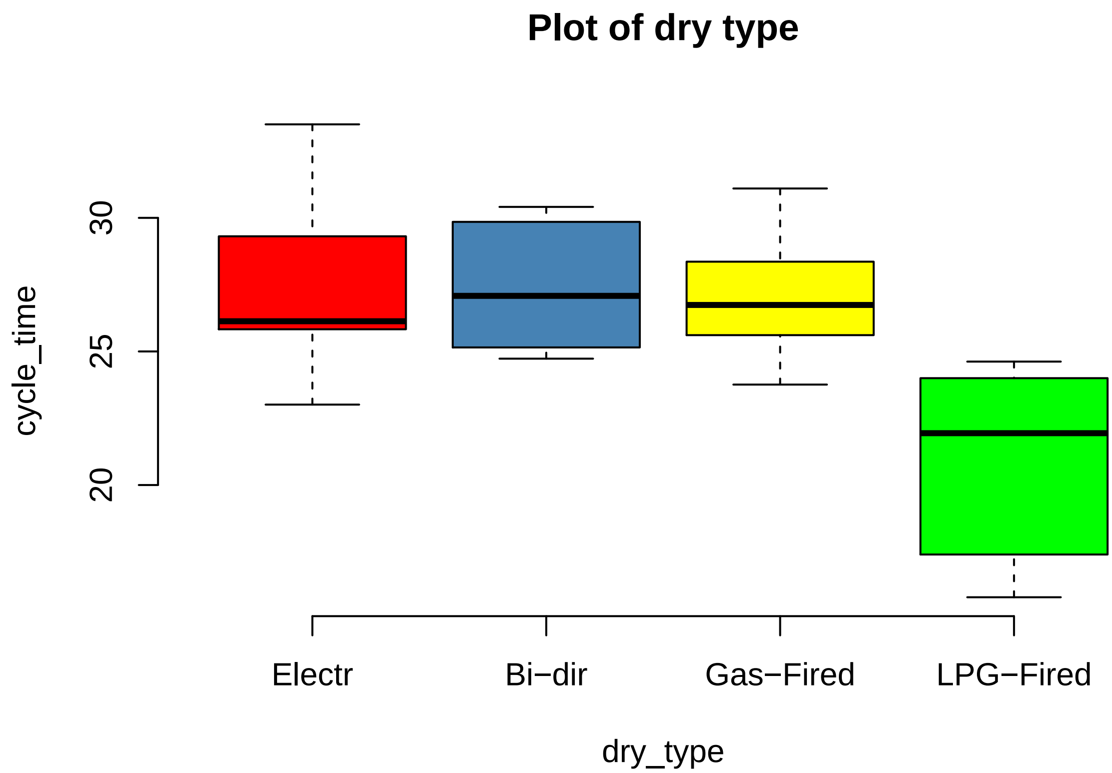
```r
boxplot(cycle_time ~ Cloth_cat*dry_type , data = drying_data,
        main = "Plot of cloth category by dry type",
        col = c("red", "steelblue", "yellow"), frame = FALSE)
```




```r
boxplot(cycle_time ~ Cloth_cat , data = drying_data,
        main = "Plot of cloth category by dry type",
        col = c("skyblue", "grey", "green"), frame = FALSE)
```

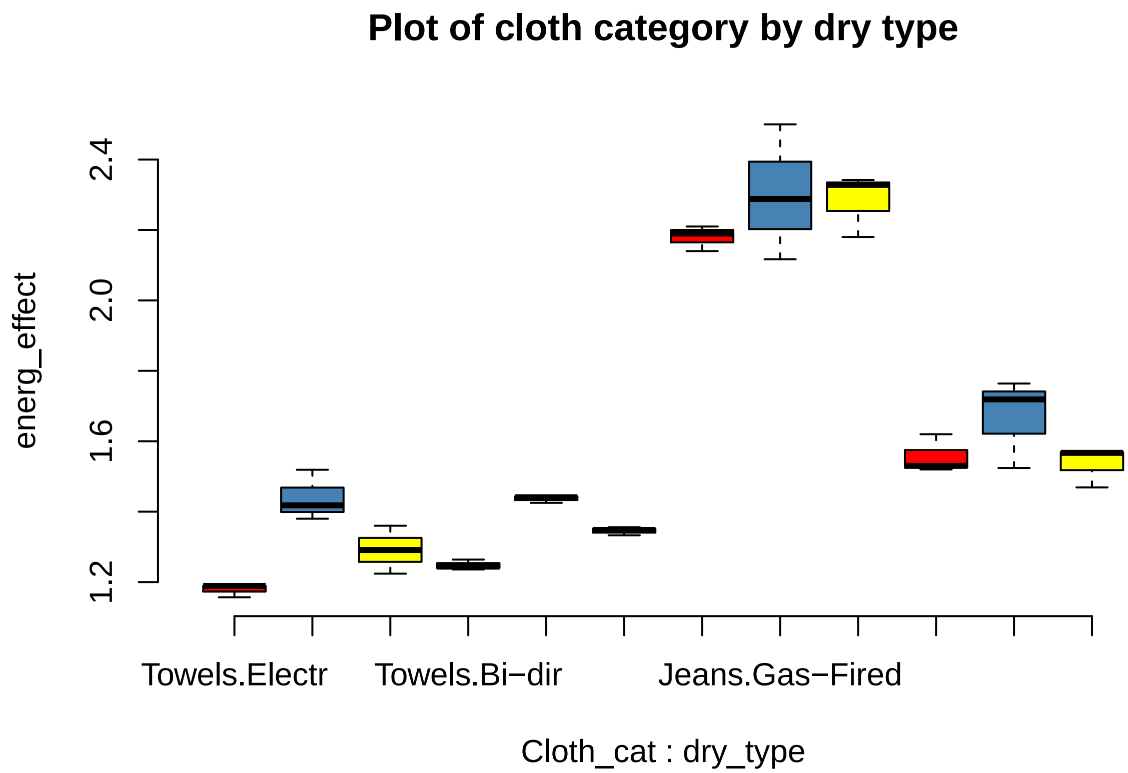



```
boxplot(cycle_time ~ dry_type , data = drying_data,
        main = "Plot of dry type",
        col = c("red", "steelblue", "yellow", "green"), frame = FALSE)
```



Also we can conclude that obviously Town Gas-Fired dryers need more energy per kilogram of a cloth.

```
boxplot(energ_effect ~ Cloth_cat*dry_type , data = drying_data,
        main = "Plot of cloth category by dry type",
        col = c("red", "steelblue", "yellow"), frame = FALSE)
```

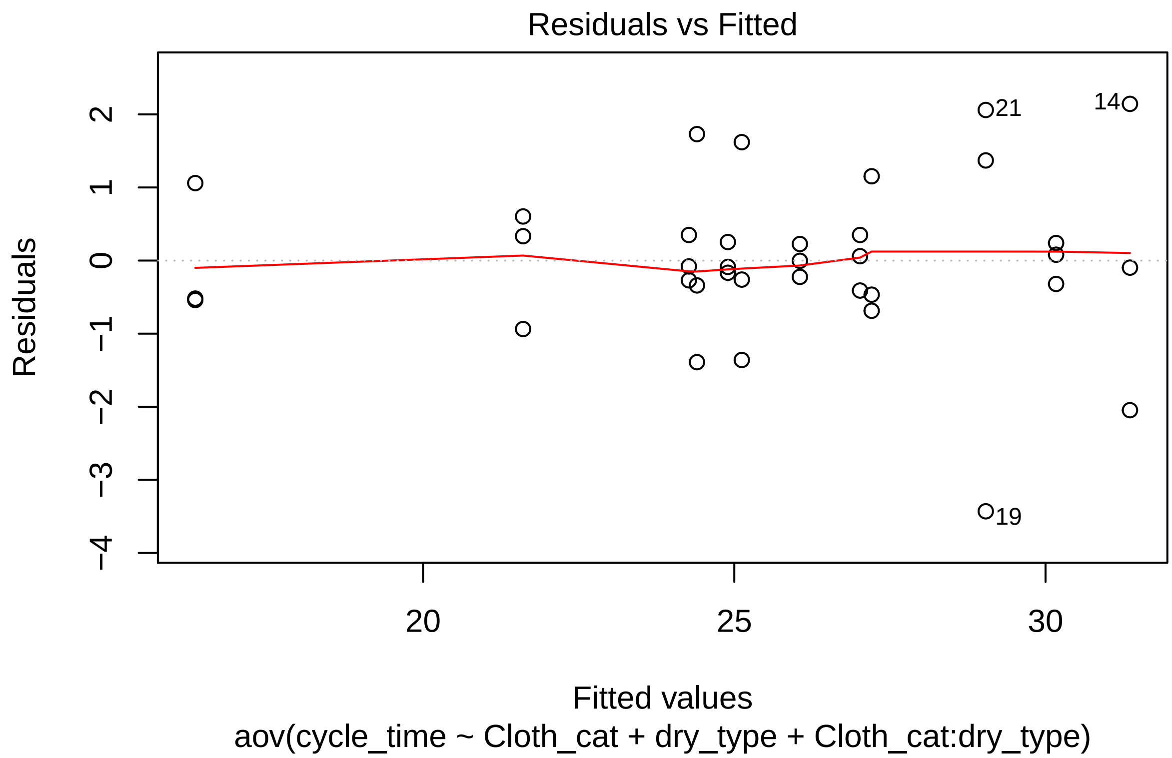


Check the homogeneity of variance assumption

ANOVA assumes that the data are normally distributed and the variance across groups are homogeneous. We can check that with some diagnostic plots.

Residuals plot

```
plot(res.aov3, 1)
```



Points 19, 21 and 14 are detected as outliers, which can severely affect normality and homogeneity of variance

Levene's test

```
leveneTest(cycle_time ~ Cloth_cat*dry_type, data = drying_data)
```

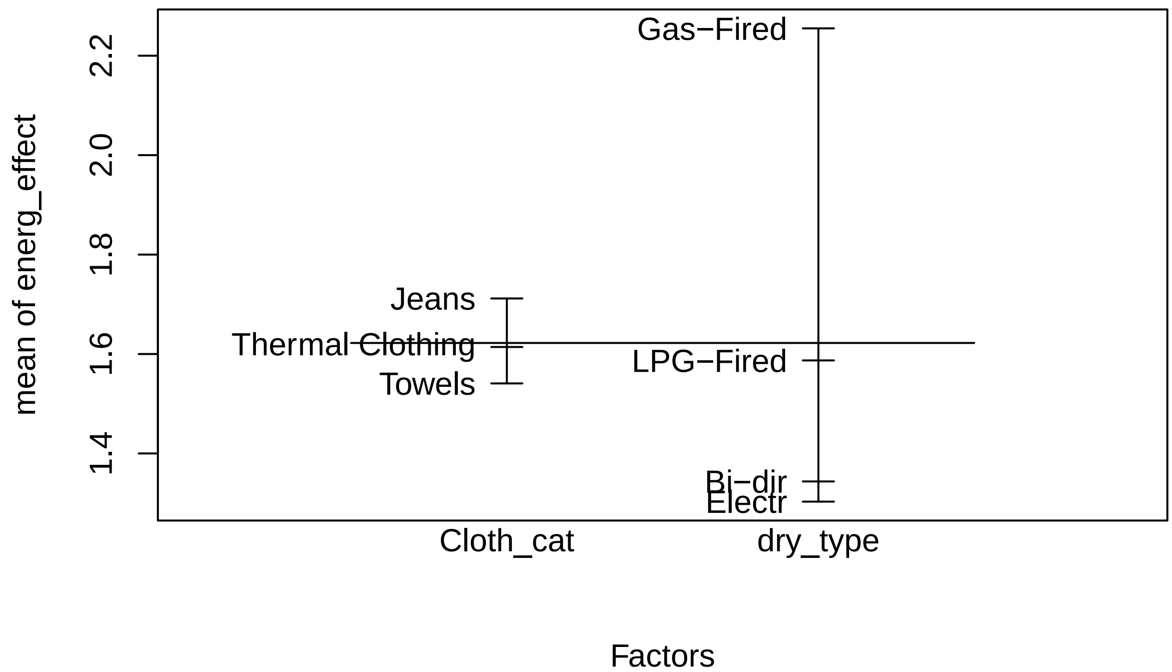
```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group 11  0.8783 0.5721
##       24
```

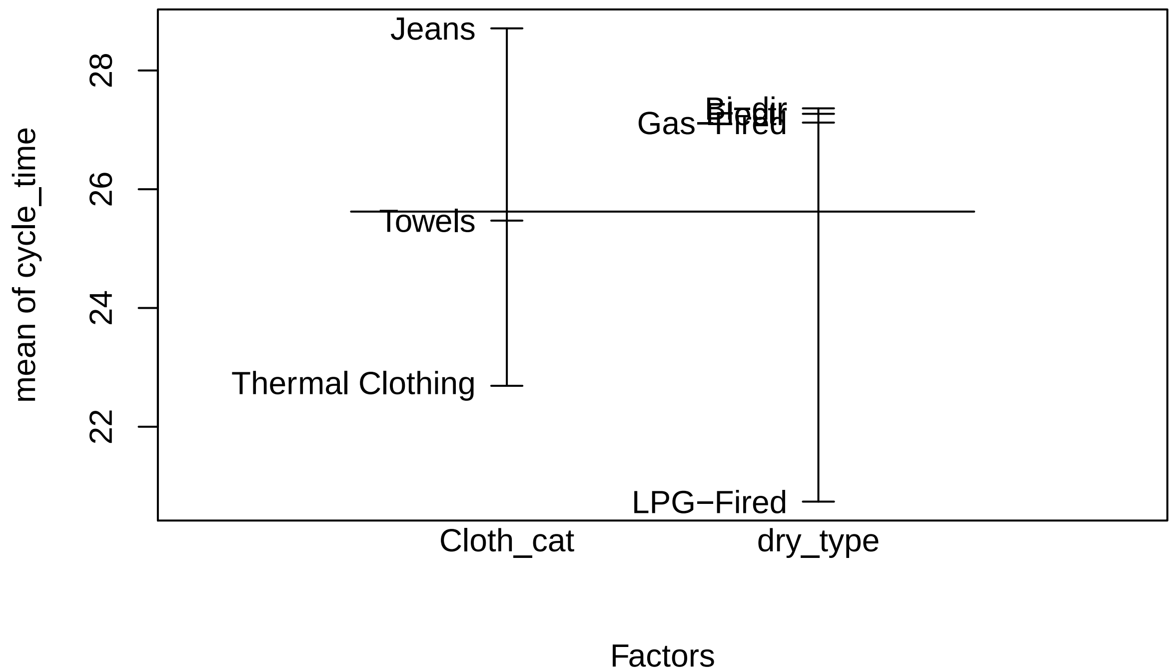
From the output above we can see that the p-value is not less than the significance level of 0.05. This means that there is no evidence to suggest that the variance across groups is statistically significantly different. Therefore, we can assume the homogeneity of variances in the different treatment groups.

Compute two-way ANOVA test

We want to know if tooth length depends on clothing category and drying type.

```
plot.design(drying_data)
```


Gas−Fired
Jeans
Thermal Clothing
Towels
LPG−Fired
Bi−dir
Electr
2.2
2.0
1.8
1.6
1.4
mean of energ_effect
Cloth_cat
dry_type
Factors




```
res.aov3 <- aov(energ_effect ~ Cloth_cat + dry_type + Cloth_cat:dry_type, data = drying_data)
summary(res.aov3)
```

```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## Cloth_cat           2  0.176  0.0882  13.332 0.000128 ***
## dry_type            3  5.230  1.7432 263.397  < 2e-16 ***
## Cloth_cat:dry_type  6  0.036  0.0060   0.912 0.502992
## Residuals          24  0.159  0.0066
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
res.aov3 <- aov(cycle_time ~ Cloth_cat + dry_type + Cloth_cat:dry_type, data = drying_data)
summary(res.aov3)
```

```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## Cloth_cat           2 217.86  108.93  61.979 3.32e-10 ***
## dry_type            3 286.56   95.52  54.349 7.53e-11 ***
## Cloth_cat:dry_type  6  24.43    4.07   2.317    0.066 .
## Residuals          24  42.18    1.76
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the ANOVA table we can conclude that both clothing category and drying type are statistically significant for both predicted values. However in Energy spending drying type is the most significant factor variable. These results would lead us to believe that changing the way of drying clothes or the clothe, will impact significantly the mean drying time.

Based on the p-values and a significance level of 0.05

```
TukeyHSD(res.aov3, which = "dry_type")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = cycle_time ~ Cloth_cat + dry_type + Cloth_cat:dry_type, data = drying_data)
##
## $dry_type
##                            diff       lwr       upr     p adj
## Bi-dir-Electr        0.09222222 -1.631762  1.816206 0.9988210
## Gas-Fired-Electr    -0.14777778 -1.871762  1.576206 0.9952181
## LPG-Fired-Electr    -6.53111111 -8.255095 -4.807127 0.0000000
## Gas-Fired-Bi-dir    -0.24000000 -1.963984  1.483984 0.9802826
## LPG-Fired-Bi-dir    -6.62333333 -8.347317 -4.899349 0.0000000
## LPG-Fired-Gas-Fired -6.38333333 -8.107317 -4.659349 0.0000000
```

```
plot(TukeyHSD(res.aov3, which = "dry_type"), las = 1)
```