

# Statistical testing

Weeks 2 and 3 LOVE'R course

Tania L. Maxwell

email: [tania.maxwell@inrae.fr](mailto:tania.maxwell@inrae.fr)

twitter: [@tania\\_maxwell7](https://twitter.com/tania_maxwell7)

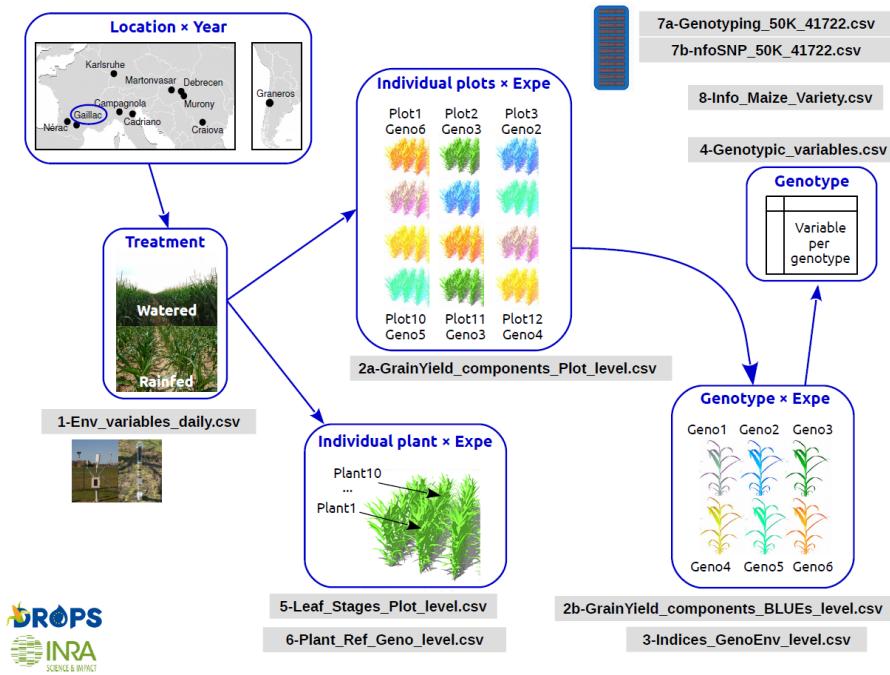
website: [tania-maxwell.github.io](https://tania-maxwell.github.io)

12-11-2021

# Topics for today

- Explain the dataset we will be working with
- Data format
- Visual data
- Treatment effects
- What to extract from your analysis

# Dataset: A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios



Citation: Millet, Emilie J.; Pommier, Cyril; Buy, Mélanie; Nagel, Axel; Kruijer, Willem; Welz-Bolduan, Therese; Lopez, Jeremy; Richard, Cécile; Racz, Ferenc; Tanzi, Franco; Spirkot, Tamas; Canè, Maria-Angela; Negro, Sandra S.; Coupel-Ledru, Aude; Nicolas, Stéphane D.; Palaffre, Carine; Bauland, Cyril; Praud, Sébastien; Ranc, Nicolas; Presterl, Thomas; Bedo, Zoltan; Tuberosa, Roberto; Usadel, Björn; Charcosset, Alain; van Eeuwijk, Fred A.; Draye, Xavier; Tardieu, François; Welcker, Claude, 2019, "A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios", <https://doi.org/10.15454/IASSTN>, Portail Data INRAE, V2, UNF:6:zF9w0A2f+MHeW7maeeXJWA== [fileUNF]

# Experiment detail

A panel of 256 maize hybrids was grown:

- with two water regimes (irrigated or rainfed),
- in seven sites spread along a climatic transect from western to eastern Europe
- in 2012 and 2013

which equals 29 experiments, as the combination of one year, one site and one water regime, with two and three repetitions for watered and rainfed treatments, respectively.

Measurements:

- Hourly records of micrometeorological data and soil water status, and associated with precise measurement of phenology
- **Grain yield and its components at the end of the experiment**

# Statistical Analysis

What is the **first** step to do when you start your statistical analysis?

## State the specific research question(s)

Your statistical test should be based on the biological question.

1. Is there a relation between the grain weight and the tassel height in the Gaillac site?
2. What is the effect of the water treatment and Variety ID on plant height in 2012?

# Today's objectives

- Go through the questions together, then individually
- Look into detail our analyses
- Meanwhile, go through important concepts to keep in mind

# Data Import

```
library(tidyr)
library(dplyr)
library(corrplot)
library(corrgram)
library(multcomp)
library(nlme)
library(lme4)

setwd("C:/Users/tmaxwell/Documents/Teaching/M2_B2AS_R-stats")
maize_data <- read.table("2a-GrainYield_components_Plot_level-1.csv",
                         header = T, sep = ",")
```

Note: Because we are importing a `.csv` file, the `sep =` will be a comma.



What kind of data do we have here?

Hint: Click on the data in your Global Environmentm, and use the `str()` function,

```
str(maize_data)
```

```
## 'data.frame': 19358 obs. of 21 variables:  
## $ Site : chr "Gaillac" "Gaillac" "Gaillac" "Gaillac" ...  
## $ year : int 2012 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...  
## $ Experiment : chr "Gai12R" "Gai12R" "Gai12R" "Gai12W" ...  
## $ plotID : chr "G72-01-1-1" "G72-01-2-1" "G72-01-3-1" "G72-02-1-1" ...  
## $ treatment : chr "rainfed" "rainfed" "rainfed" "watered" ...  
## $ Replicate : int 1 2 3 1 2 1 2 3 1 2 ...  
## $ block : int 9 7 3 26 18 15 8 22 11 12 ...  
## $ Row : int 3 11 19 7 14 4 11 24 3 12 ...  
## $ Column : int 5 16 21 11 13 20 27 13 20 27 ...  
## $ Accession : chr "B73_H" "B73_H" "B73_H" "B73_H" ...  
## $ Code_ID : int 3001 3001 3001 3001 3001 3002 3002 3002 3002 3002 ...  
## $ Variety_ID : chr "B73" "B73" "B73" "B73" ...  
## $ plant.height : num 170 170 165 255 280 165 165 175 230 230 ...  
## $ tassel.height : num 230 220 235 315 340 205 220 215 285 285 ...  
## $ ear.height : num 105 105 90 115 160 105 105 80 110 105 ...  
## $ anthesis : num 64.8 69.2 67 66 69.2 ...  
## $ silking : num 78.9 78.9 81.7 64.8 68.1 ...  
## $ anthesis.silking.interval: num -14.12 -9.71 -14.63 1.19 1.07 ...  
## $ grain.number : num 1492 1521 1960 4462 3987 ...  
## $ grain.yield : num 3.41 3.24 3.4 12.55 10.56 ...  
## $ grain.weight : num 229 213 174 281 265 ...
```

# Converting our data into the proper format

You could write out each individual column in the following way:

```
maize_data$year <- as.factor(maize_data$year)
```

```
maize_data$Replicate <- as.factor(maize_data$Replicate)
```

etc...

OR you could use `mutate_if()` from the `dplyr` package. Here, the function states "If a column is a character, convert it as a factor" for the entire data frame.

```
maize_data <- maize_data %>%  
  mutate_if(is.character, as.factor)
```

- It is important to do this at the beginning to reduce future problems with analyses.
- Tip: try to name your columns in a consistent manner (i.e. all variables with capital first letter), because R is case-sensitive

# Exploring the data



How many observations do we have per experiment?

Hint: use the `table()` function

```
table(maize_data$Experiment)
```

```
##  
## Bol12R Bol12W Cam11R Cam11W Cam12R Cam12W Cam13R Cam13W Cra12R Cra12W Deb11R Deb11W Deb12  
##    756     449     108      72     756     504     756     504     753     500     105      70      75  
## Gai12R Gai12W Gai13R Gai13W Gra13R Gra13W Kar11R Kar11W Kar12R Kar12W Kar13R Kar13W Mar13  
##    756     504     756     504     759     506     148     100     747     504     723     487     75  
## Mur13W Ner11R Ner11W Ner12R Ner12W Ner13R Ner13W  
##    504     285     190     756     504     756     504
```

What do we see here? Does this make sense? Think of the experiment set up.

**Question 1: Is there a relation between the grain weight and the tassel height in the Gaillac site?**

# Question 1: Is there a relation between the grain weight and the tassel height in the Gaillac site?

Key parts to this question:

- Relation between two **numerical** variables
- What are the units?

`grain.weight`: individual grain weight (mg)

`tassel.height`: plant height including tassel, from ground level to the highest point of the tassel (cm)

Thus, they are both **continuous** variables

# Reminder

## CONTINUOUS

measured data, can have  $\infty$  values within possible range.



I AM 3.1" TALL  
I WEIGH 34.16 grams

## DISCRETE

OBSERVATIONS CAN ONLY EXIST  
AT LIMITED VALUES, OFTEN  
COUNTS.



I HAVE 8 LEGS  
and  
4 SPOTS!

@allison\_horst

# Data Frame



First, separate the data frame to only include values from the Gaillac site. Hint: use piping `%>%` and the `filter()` function.

```
maize_Gaillac <- maize_data %>%
  filter(Site == "Gaillac")
```

# Visualize the relation between grain weight and tassel height

```
plot(maize_Gaillac$grain.weight ~ maize_Gaillac$tassel.height)
```

# Do the statistical test

Which test do you think we should use? Why?

## Pearson's correlation

Why?

- Looking for the **relation** between two variables
- There is not a clear cause and effect relationship between the variables (linear regression)
- There may be a third non-measured variable which acts on both the variables in the correlation

# Correlation vs linear regression

## Correlation

- used to associate two variables (no cause and effect)
- mostly used as a "pre-analysis" phase to determine relations between explanatory variables (note: in different )

## Linear regression

- often used for observational studies
- a cause and effect relationship is difficult to assure without having controled for other pertinente variables
- be careful with interpretation

# What are the test assumptions for Pearson's correlation?

1. each variable should be continuous
2. related pairs: each value should have both a grain yield and a tassel height value
3. absence of outliers: can skew
4. linearity: the shape of the dots is a line and not curved

Tip: you can find test assumptions online

# Do we satisfy the test assumptions?

- [x] 1. Both variables are continuous

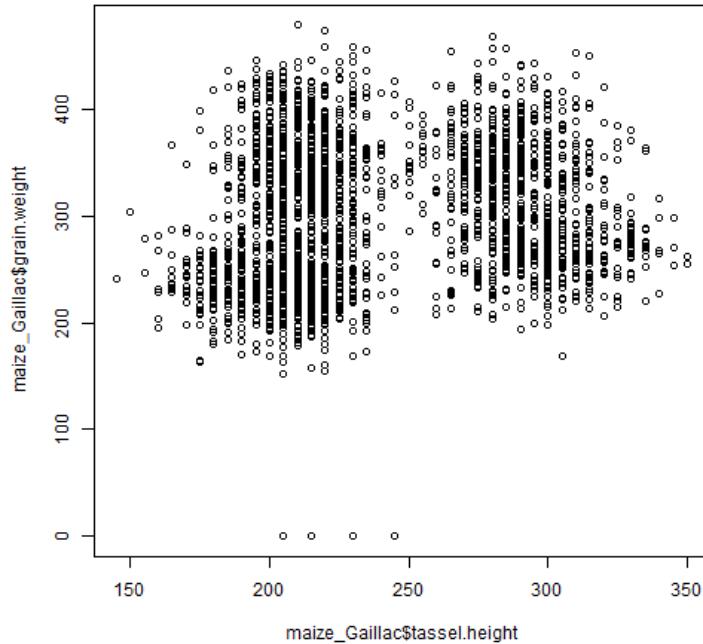
Related pairs? Let's remove the na values from the data frame using the `drop_na()` function

```
maize_Gaillac <- maize_Gaillac %>%
  drop_na("grain.weight", "tassel.height")
```

- [x] 2. NA values removed

Are there outliers which could skew the correlation?

```
plot(maize_Gaillac$grain.weight ~ maize_Gaillac$tassel.height)
```



Visually, we can determine to remove the values at the bottom of the graph, with grain weights lower than 100mg.



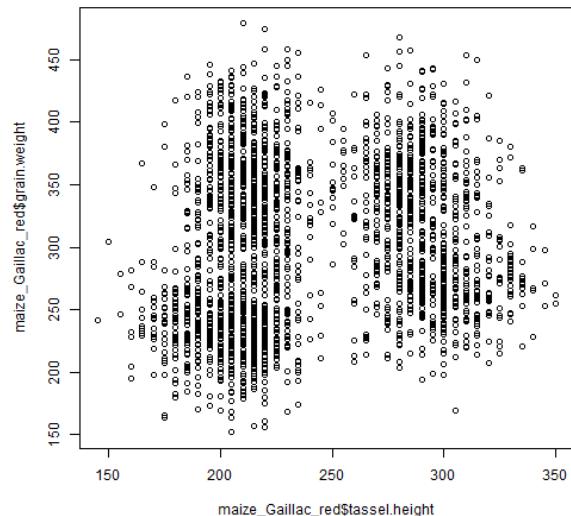
Remove the values with a grain weigh lower than 100 mg to create a new data frame for the analyses.

Hint: use the `filter()` function.

```
maize_Gaillac_red <- maize_Gaillac %>%
  filter(grain.weight>100)
```

Check the number of observations in each data frame in your Global Environment. You can also check with a graph:

```
plot(maize_Gaillac_red$grain.weight ~ maize_Gaillac_red$tassel.height)
```



[x] 3. Outliers removed

[x] 4. Linearity (no curved shape within the dots)

# Pearson's correlation

Here the p-value corresponds to the statistical test with

$$H_0 : r = 0 \text{ (there is no correlation)} \quad H_1 : r \neq 0$$

```
cor.test(maize_Gaillac_red$grain.weight, maize_Gaillac_red$tassel.height)
```

```
##  
##      Pearson's product-moment correlation  
##  
## data: maize_Gaillac_red$grain.weight and maize_Gaillac_red$tassel.height  
## t = 6.8529, df = 2471, p-value = 9.103e-12  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
##  0.0976787 0.1750430  
## sample estimates:  
##      cor  
## 0.1365691
```

The Pearson's coreltation shows a significant positive (`cor` value is positive) relationship.

However, a significant result is not very informative (with a large number of samples, anything can become significant)

It is better to look at the strength of the correlation ( $r^2$ )

## Finding the $r^2$

The  $r^2$  is the r correlation value, squared

```
cor(maize_Gaillac_red$grain.weight, maize_Gaillac_red$tassel.height)^2  
## [1] 0.01865111
```

We can conclude that **1.87%** of the variation of the grain weight is associated to the tassel height.

# Subsetting for a correlation matrix

We can get a quick overview of how all of our numeric columns value relate to one another

First, we need to subset only the numeric columns. One way we can do this is by viewing our data table and noting the numeric columns.

```
maize_Gaillac_num <- maize_Gaillac[,c(13:21)]
```

Or, we could select columns which are numeric using `select(where(is.numeric))`

```
maize_Gaillac_num <- maize_Gaillac %>%
  dplyr::select(where((is.numeric)))
```

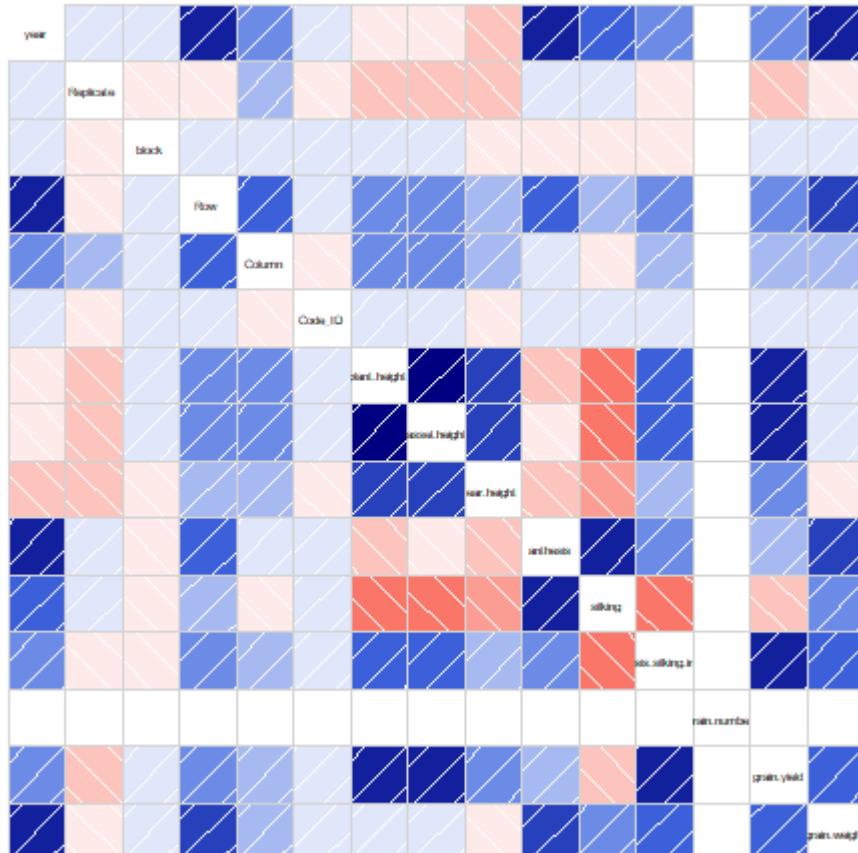
Note: the `select()` function is present in several packages, so by using `dplyr::` we are telling R in which package to take the function.

We need to remove NA values from the data frame. Do so, we can use the `na.omit()` function.

```
maize_Gaillac_num <- na.omit(maize_Gaillac_num)
```

We will use a function which needs a correlation matrix from our values. To do so, we will use the `corrgram()` function.

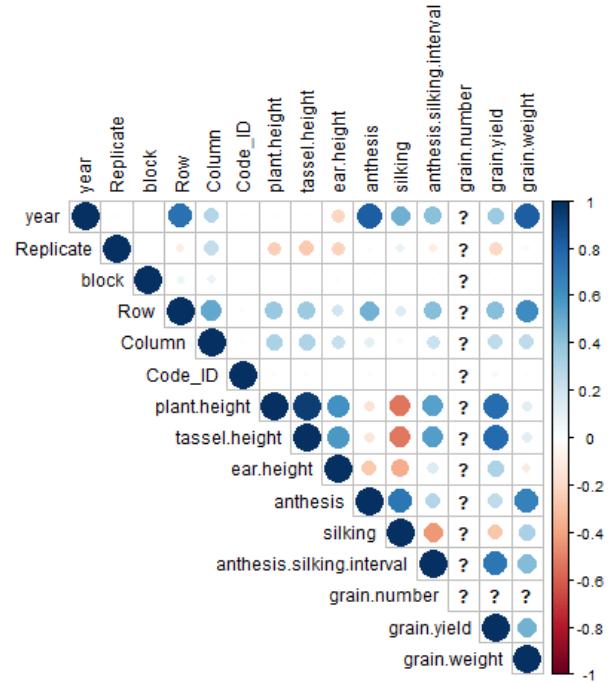
```
corr_maize <- corrgram(maize_Gaillac_num)
```



# Correlation matrix

Positive correlations are displayed in blue and negative correlations in red color. Color intensity and the size of the circle are proportional to the correlation coefficients.

```
corrplot(corr_maize, method = "circle", type = "upper", tl.col= "black")
```



# Your turn



Using the previous data table, what is the relation between grain weight and tassel height within the waterred site? And within the rainfed site?

# Solution

```
maize_Gaillac_wat <- maize_Gaillac_red %>%
  filter(treatment == "watered")

cor(maize_Gaillac_wat$grain.weight, maize_Gaillac_wat$tassel.height)^2

## [1] 0.0748907
```

We can conclude that **7.49%** of the variation of the grain weight is associated to the tassel height if we only include the watered sites.

```
maize_Gaillac_rain <- maize_Gaillac_red %>%
  filter(treatment == "rainfed")

cor(maize_Gaillac_rain$grain.weight, maize_Gaillac_rain$tassel.height)^2

## [1] 0.04363096
```

We can conclude that **4.36%** of the variation of the grain weight is associated to the tassel height if we only include the rainfed sites.

Does this make sense?

**Question 2: What is the effect of the water treatment  
and Variety ID on plant height in 2012?**

# Question 2: What is the effect of the water treatment on plant height in 2012?

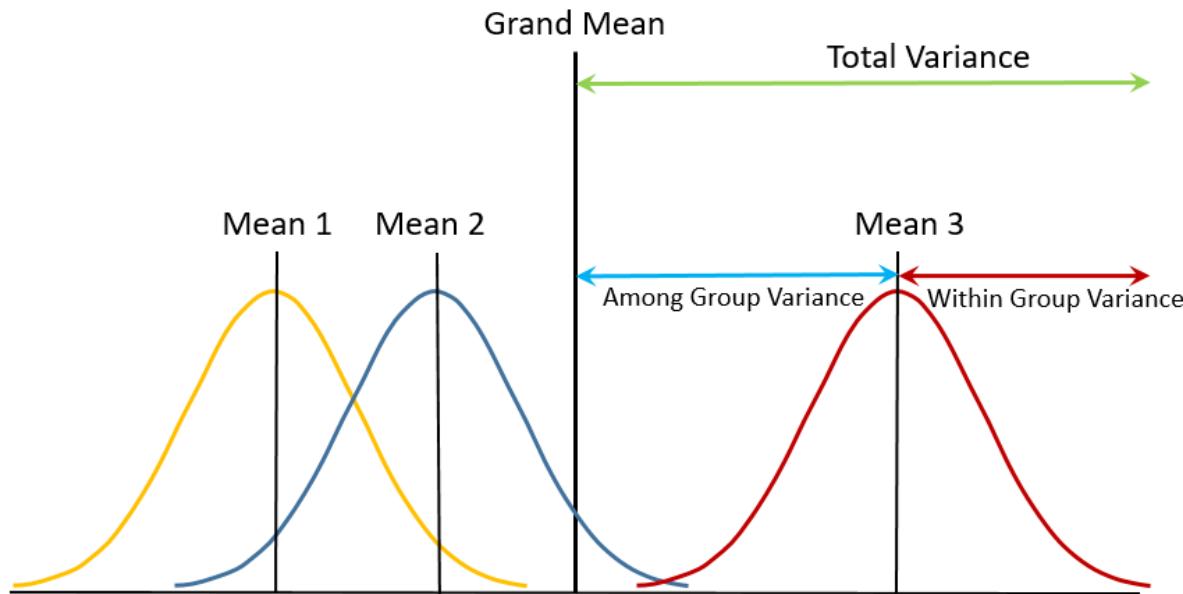
Key parts to this question:

- Effect of **two factors** (treatment and Variety) on a **numerical** variable (plant height)
- **plant.height** : from ground level to the base of the flag leaf (highest) leaf (cm), continuous variable

Which test should we use? Why?

# ANOVA (Analysis of Variance)

- Variance can be used to test whether means of treatments differed: as the means become more different, the variances increase



# ANOVA reminder: $x \sim a+b+a:b$

<code>aov()</code>	<code>anova(lm())</code>
Analysis of Variance	Analysis of variance of a linear model
Type 1 Sum of Squares:	Type 3 Sum of Squares:
$SS(A)$ for factor A	$SS(A / B, AB)$ for factor A
$SS(B / A)$ for factor B	$SS(B / A, AB)$ for factor B
$SS(AB / B, A)$ for interaction AB	
Will give different results if unbalanced data depending which factor is written first	Allows for unbalanced data.

```
mod_aov <- aov(x ~ a + b + a:b)  
mod_aov output is directly the Analysis of Variance Table
```

```
mod_lm <- lm(x ~ a + b + a:b)  
anova(mod_lm) gives the Analysis of Variance Table
```

# Example

We will decipher the Analysis of Variance table in detail with the example.

Note: an anova is basically an extended t-test (which compares the variance of two groups)

# Reducing the maize\_data dataframe



Separate the data frame.

- Only the year 2012
- To simplify our question, we will focus on three Varieties: "HMV5422", "11430" and "F712"

Hint: use the | symbol to include several values for Variety\_ID

- Remove the NA values from the "plant.height" column

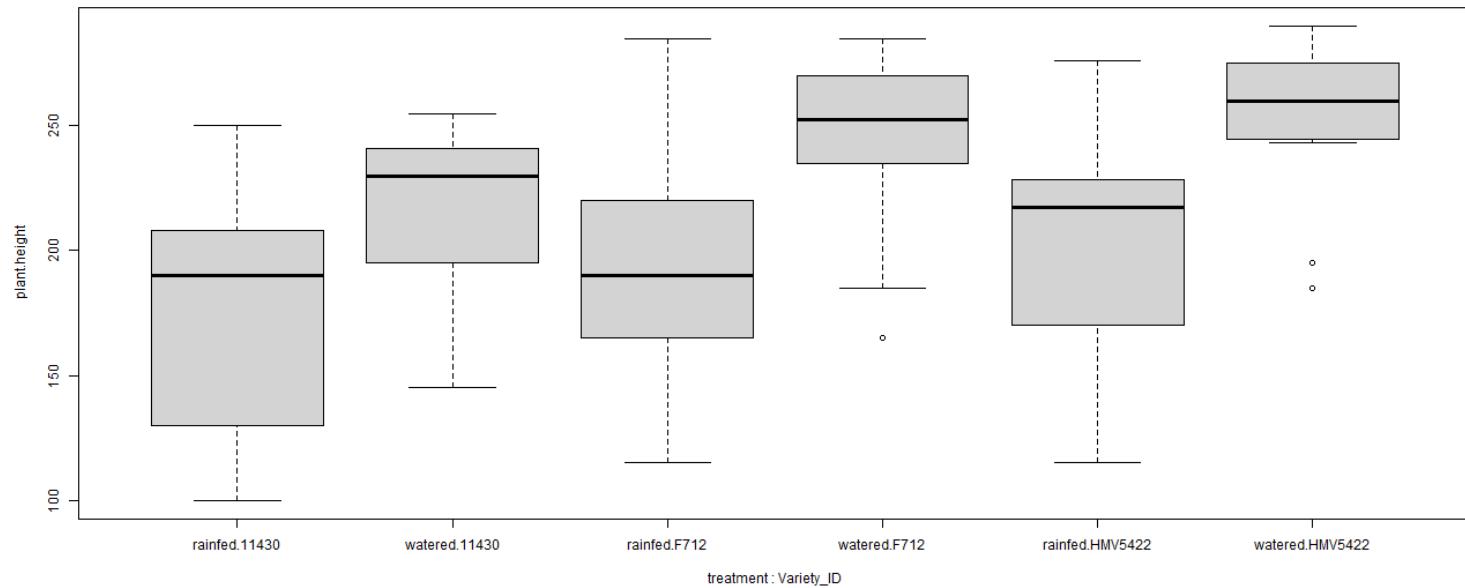
# Answer

```
maize_2012 <- maize_data %>%
  filter(year == "2012") %>%
  filter(Variety_ID == "HMV5422" | Variety_ID == "11430" | Variety_ID == "F712")%>%
  drop_na("plant.height") %>%
  droplevels() #we need to add this to remove the other levels of Variety
str(maize_2012)
```

```
## 'data.frame': 103 obs. of 21 variables:
## $ Site : Factor w/ 7 levels "Bologna","Campagnola",...: 5 5 5 5 5 5 ...
## $ year : int 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...
## $ Experiment : Factor w/ 14 levels "Bol12R","Bol12W",...: 9 9 9 10 10 9 9 ...
## $ plotID : Factor w/ 103 levels "113037","113057",...: 59 62 65 68 71 6 ...
## $ treatment : Factor w/ 2 levels "rainfed","watered": 1 1 1 2 2 1 1 1 2 ...
## $ Replicate : int 1 2 3 1 2 1 2 3 1 2 ...
## $ block : int 9 2 22 26 1 16 27 11 12 35 ...
## $ Row : int 3 10 24 7 10 4 16 21 3 18 ...
## $ Column : int 4 10 8 10 1 26 21 21 24 19 ...
## $ Accession : Factor w/ 3 levels "11430_H","F712_H",...: 3 3 3 3 3 2 2 2 ...
## $ Code_ID : int 3037 3037 3037 3037 3037 3057 3057 3057 3057 3057 ...
## $ Variety_ID : Factor w/ 3 levels "11430","F712",...: 3 3 3 3 3 2 2 2 ...
## $ plant.height : num 170 170 170 270 280 180 190 180 255 275 ...
## $ tassel.height : num 225 220 230 335 330 210 235 210 325 305 ...
## $ ear.height : num 105 105 100 120 140 105 105 110 150 120 ...
## $ anthesis : num 70.3 62.1 67 66 69.2 ...
## $ silking : num 80.2 64.8 74.9 66 68.1 ...
## $ anthesis.silking.interval: num -9.86 -2.72 -7.89 0 1.07 ...
## $ grain.number : num 2174 1336 1352 NA 5026 ...
## $ grain.yield : num 4.9 3.3 3.35 NA 13.17 ...
```

# Looking at data visually

```
boxplot(plant.height ~ treatment*Variety_ID, data = maize_2012)
```



What can we see?

-There seems to be a positive effect of watering plants on plant height (makes sense). We don't see a clear treatment effect, but there might be.

# Testing models

Can we use `aov()`? Hint: use the `table()` function for our factors to see if we have a balanced dataset for our different factors in our dataset.

```
table(maize_2012$Variety_ID)
```

```
##  
##    11430     F712  HMV5422  
##      34       35      34
```

```
table(maize_2012$treatment)
```

```
##  
## rainfed  watered  
##      62       41
```

Because we have unbalanced data, it is better to use the `lm()` function.

# Using the lm() function

```
mod_lm <- lm(plant.height ~ treatment + Variety_ID, data = maize_2012)
anova(mod_lm)
```

```
## Analysis of Variance Table
##
## Response: plant.height
##           Df Sum Sq Mean Sq F value    Pr(>F)
## treatment   1 56338  56338 31.2972 1.971e-07 ***
## Variety_ID  2 16905   8452  4.6955  0.01127 *
## Residuals  99 178211    1800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

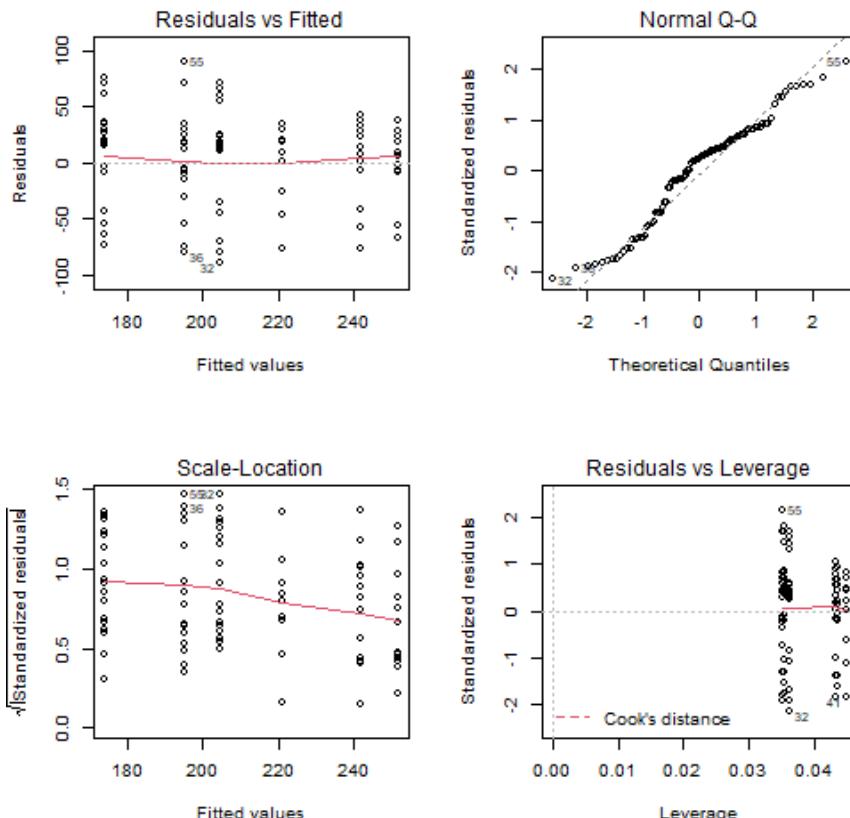
# What are the test assumptions

Before analyzing the data, check the following assumptions:

1. Homogeneity (or homoscedasticity) of variances (look at residuals vs fitted)
2. Normality of residues (but normality isn't as important as homoscedasticity for an ANOVA)
3. No outliers
4. The plant.height data are randomly selected and are independent (will go over this later)

# Check the assumptions

```
par(mfrow = c(2, 2))
plot(mod_lm)
```



# Looking at the analysis of variance table

```
anova(mod_lm)
```

```
## Analysis of Variance Table
##
## Response: plant.height
##           Df Sum Sq Mean Sq F value    Pr(>F)
## treatment   1 56338  56338 31.2972 1.971e-07 ***
## Variety_ID  2 16905    8452  4.6955  0.01127 *
## Residuals  99 178211     1800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Source	SS	df	MS	F
Effect A	given	a-1	SS/df	MS(A) / MS(W)
Effect B	given	b-1	SS/df	MS(B) / MS(W)
Within (residuals)	given		SS/df	

# Let's break down the table output

```
anova(mod_lm)
```

```
## Analysis of Variance Table
##
## Response: plant.height
##           Df Sum Sq Mean Sq F value    Pr(>F)
## treatment   1 56338  56338 31.2972 1.971e-07 ***
## Variety_ID  2 16905    8452  4.6955  0.01127 *
## Residuals  99 178211     1800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Df: degrees of freedom = the number of values in the final calculation of a statistic that are free to vary

- Each main effect has  $k-1$  Df

ex: `Variety_ID` has 3 levels ( $k=3$ ) so  $df = 3-1 = 2$

and `treatment` has two levels ( $m=2$ ) so  $df = 2-1 = 1$

```
anova(mod_lm)
```

```
## Analysis of Variance Table
##
## Response: plant.height
##           Df Sum Sq Mean Sq F value    Pr(>F)
## treatment     1  56338   56338 31.2972 1.971e-07 ***
## Variety_ID    2  16905    8452  4.6955  0.01127 *
## Residuals   99 178211    1800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- The residuals df is the total number of observations (N) - 1 - the df of each effect (k-1) and (m-1).
- You can find N by calculating the length of plant.height in the maize\_2012 data table, or by looking in the Global Environment data values

```
length(maize_2012$plant.height)
```

```
## [1] 103
```

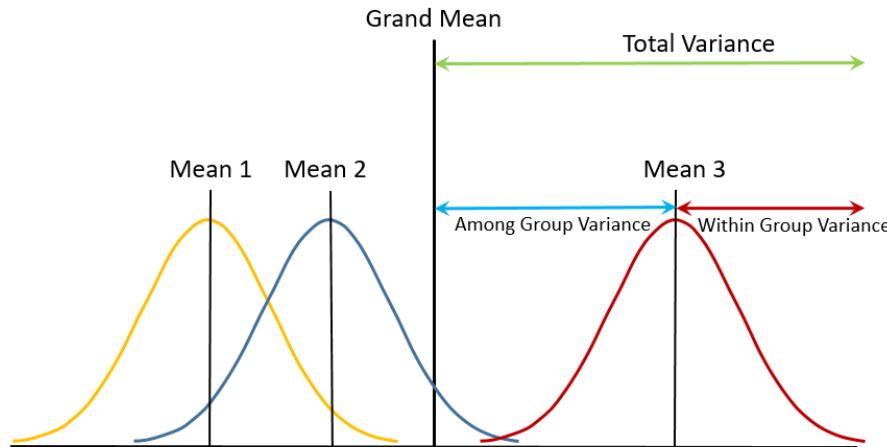
$$= N - 1 - (k-1) - (m-1) = 103 - 1 - 2 - 1 = 99$$

Why is this important? You can use this to better understand tables published in papers. You can also easily see if authors have removed data points from their analysis (the N.)

# Fisher's test

Sum of Squares of group / Sum of squares residuals (within)

Ex: If the `Variety_ID` means of plant height are more different (vary more between the different levels) than the plant height vary within the different Variety levels, the F-test ratio would be greater than an F-value expected by chance.



# The p-value

```
anova(mod_lm)
```

```
## Analysis of Variance Table
##
## Response: plant.height
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## treatment   1  56338  56338 31.2972 1.971e-07 ***
## Variety_ID  2   16905    8452  4.6955  0.01127 *
## Residuals  99 178211     1800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R then compares the F ratio calculated to the expected value given the two degrees of freedom (between the groups and within the groups), with alpha set to 0.05%.

If the p-value is less than the chosen significance level, the data provide sufficient evidence to conclude that your model fits the data better than a model with no independent variables.

Be cautious with interpreting p-values (will go over next week)

# Summary()

```
summary(mod_lm)
```

```
##  
## Call:  
## lm(formula = plant.height ~ treatment + Variety_ID, data = maize_2012)  
##  
## Residuals:  
##      Min      1Q Median      3Q      Max  
## -89.708 -32.332   9.154  27.020  90.044  
##  
## Coefficients:  
##                               Estimate Std. Error t value Pr(>|t|)  
## (Intercept)            173.865    7.976  21.799 < 2e-16 ***  
## treatmentwatered     47.138    8.543   5.518 2.75e-07 ***  
## Variety_IDF712        21.091   10.218   2.064  0.04161 *  
## Variety_IDHMV5422     30.843   10.293   2.996  0.00345 **  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 42.43 on 99 degrees of freedom  
## Multiple R-squared:  0.2913,    Adjusted R-squared:  0.2698  
## F-statistic: 13.56 on 3 and 99 DF,  p-value: 1.753e-07
```

# Post-hoc tests

The F-test doesn't reveal which Variety\_ID means are different, only that the Variety\_ID means are different.

Therefore, we can do a multiple comparison test, also called a posthoc test

```
modposthoc = lm(plant.height ~ Variety_ID+treatment,  
                 data = maize_2012, na.action=na.omit)  
tuk <- glht(modposthoc , linfct = mcp(Variety_ID= "Tukey"))
```

```
summary(tuk)
```

```
##  
##      Simultaneous Tests for General Linear Hypotheses  
##  
## Multiple Comparisons of Means: Tukey Contrasts  
##  
##  
## Fit: lm(formula = plant.height ~ Variety_ID + treatment, data = maize_2012,  
##        na.action = na.omit)  
##  
## Linear Hypotheses:  
##                         Estimate Std. Error t value Pr(>|t|)  
## F712 - 11430 == 0     21.091    10.218   2.064  0.10263  
## HMV5422 - 11430 == 0   30.843    10.293   2.996  0.00961 **  
## HMV5422 - F712 == 0    9.752    10.217   0.954  0.60728  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
## (Adjusted p values reported -- single-step method)
```

We can also use the `cld()` function to get the classic letter display.

```
tuk.cld <- cld(tuk) # letter-based display  
tuk.cld
```

```
## 11430 F712 HMV5422  
## "a" "ab" "b"
```

# Your turn

In groups of 2:

- State a research question for this data set (a response variable, and then one or more )
- Divide the original data set by site, year, treatment and/or variety
- Choose the (or several) fixed effect factors and the response variable according to your research question
- Write your model, test the assumptions and state your results
- Calculate the n, mean, and standard error for your groups

# Next week

- Pseudoreplication
- Heritability
- Multivariate analyses
- Rmarkdown basics

# References

Schumaker & Tomek 2013. Understanding Statistics Using R. *Springer*

<http://www.sthda.com/english/wiki/visualize-correlation-matrix-using-correlogram>

<https://howecoresearch.blogspot.com/2019/01/using-analysis-of-variance-anova-in.html>

Ressources:

[https://www.youtube.com/watch?v=-yQb\\_ZJnFXw](https://www.youtube.com/watch?v=-yQb_ZJnFXw)