

# Rbasics

PhD toolbox - 40th PhD cycle



Part II - How to manage data in spreadsheets

# Data Organization in Spreadsheets

**Now you should know the basic R syntax and you're ready to start to import real datasets in R!**

**but**

**Your data requires a clear structure**

Spreadsheets (mostly Excel) are useful tools for data entry but not suitable for reproducible research

Example: statistical procedures in Excel are manual. If you need to change one parameter of your analysis you'll have to redo all your job.

# Data Organization in Spreadsheets

Do not treat your data spreadsheet as your lab book!

- Your data needs to be correctly read and interpreted by your Computer (not by your supervisor!)
- Additional notes and graphic layout of your data are useless most of the time
- keep your spreadsheet as tidy as possible

Some operative TIPS according to <https://datacarpentry.org/spreadsheet-ecology-lesson/>

# Data Organization in Spreadsheets

Some cardinal rules to correctly compile your data spreadsheet


- 1) variables in column, observations in rows

| Observations  | Factor_A | Factor_B | Measure_1 | Measure_2 |
|---------------|----------|----------|-----------|-----------|
| Observation_1 | X        | 1        |           |           |
| Observation_2 | Y        | 1        |           |           |
| Observation_3 | X        | 2        |           |           |
| Observation_4 | X        | 2        |           |           |

# Data Organization in Spreadsheets

Some cardinal rules to correctly compile your data spreadsheet

- 2) Don't mix multiple information in one cell



| Plot | Species-Sex | Weight |
|------|-------------|--------|
| 1    | DM-M        | 40     |
| 1    | DM-F        | 36     |
| 1    | DS-F        | 135    |
| 1    | DM-F        | 39     |
| 2    | DM-M        | 43     |

| Plot | Species | Sex | Weight |
|------|---------|-----|--------|
| 1    | DM      | M   | 40     |
| 1    | DM      | F   | 36     |
| 1    | DS      | F   | 135    |
| 1    | DM      | F   | 39     |
| 2    | DM      | M   | 43     |

# Data Organization in Spreadsheets

Some cardinal rules to correctly compile your data spreadsheet

- 3) **NEVER** touch the raw data! If needed make a copy and modify it.
- 4) Export and store your data as a text-based file (csv, tsv...)

# Data Organization in Spreadsheets

## Some common errors

|    | A                     | B       | C    | D    | E    | F       | G   | H      | I   | J                     | K       | L    | M    | N    | O       | P      | Q        | R                     | S    | T    | U    | V   | W | X       | Y       | Z                     | AA      | AB    | AC   | AD   | AE  | AF     | AG      |       |        |     |
|----|-----------------------|---------|------|------|------|---------|-----|--------|-----|-----------------------|---------|------|------|------|---------|--------|----------|-----------------------|------|------|------|-----|---|---------|---------|-----------------------|---------|-------|------|------|-----|--------|---------|-------|--------|-----|
| 1  |                       |         |      |      |      |         |     |        |     |                       |         |      |      |      |         |        |          |                       |      |      |      |     |   |         |         |                       |         |       |      |      |     |        |         |       |        |     |
| 2  | lake site May 29 2012 |         |      |      |      |         |     | 29-May |     | lake site Jun 12 2012 |         |      |      |      |         | 12-Jun |          | lake site Jun 19 2012 |      |      |      |     |   | 19-Jun  |         | lake site Jun 26 2012 |         |       |      |      |     | 26-Jun |         |       |        |     |
| 3  |                       |         | bug1 | bug2 |      |         |     | avr    | SEM |                       | plot    | bug1 | bug2 | gene |         | avr    | SEM      | plot                  | bug1 | bug2 | gene |     |   |         | plot    | bug1                  | bug2    | gene  |      | avr  | SEM |        |         |       |        |     |
| 4  | 1                     | T1      | 1    | 1    | 2    | T1      | 2.6 | 0.51   |     | 1                     | T1      | 6    | 85   | 91   | T1      | 30.4   | 15       | 47                    | 126  |      |      |     |   | 1       | T1      | 17                    | 80      | 97    |      | avr  | SEM |        |         |       |        |     |
| 5  | 2                     | T1      | 1    | 2    | 3    | T2      | 0.2 | 0.2    |     | 2                     | T1      | 8    | 13   | 21   | T2      | 0.2    | 0.2      |                       |      |      |      |     |   | T1      | 77.8    | 30                    | 384865  | 2     | T1   | 50   | 270 | 320    | T1      | 141.6 | 60     | 313 |
| 6  | 3                     | T1      | 1    | 3    | 4    | control | 0.2 | 0.2    |     | 3                     | T1      | 11   | 0    | 11   | control | 0.6    | 0.6      |                       |      |      |      |     |   | T2      | 1.8     | 1                     | 5620499 | 3     | T1   | 6    | 0   | 6      | T2      | 0.2   | 0.2    |     |
| 7  | 4                     | T1      | 1    | 0    | 1    |         |     |        |     | 4                     | T1      | 0    | 6    | 6    |         |        |          |                       |      |      |      |     |   | control | 0.4     | 0                     | 244949  | 4     | T1   | 0    | 39  | 39     | control | 0     | 0      |     |
| 8  | 5                     | T1      | 0    | 3    | 3    |         |     |        |     | 5                     | T1      | 3    | 20   | 23   |         |        |          |                       |      |      |      |     |   | 5       | T1      | 4                     | 96      | 100   |      |      |     |        |         |       |        |     |
| 9  | 6                     | T2      | 1    | 0    | 1    |         |     |        |     | 6                     | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 6       | T2      | 0                     | 1       | 1     |      |      |     |        |         |       |        |     |
| 10 | 7                     | T2      | 0    | 0    | 0    |         |     |        |     | 7                     | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 7       | T2      | 0                     | 0       | 1     |      |      |     |        |         |       |        |     |
| 11 | 8                     | T2      | 0    | 0    | 0    |         |     |        |     | 8                     | T2      | 1    | 0    | 1    |         |        |          |                       |      |      |      |     |   | 8       | T2      | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 12 | 9                     | T2      | 0    | 0    | 0    |         |     |        |     | 9                     | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 9       | T2      | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 13 | 10                    | T2      | 0    | 0    | 0    |         |     |        |     | 10                    | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 10      | T2      | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 14 | 11                    | control | 0    | 0    | 0    |         |     |        |     | 11                    | control | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 11      | control | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 15 | 12                    | control | 0    | 0    | 0    |         |     |        |     | 12                    | control | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 12      | control | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 16 | 13                    | control | 0    | 0    | 0    |         |     |        |     | 13                    | control | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 13      | control | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 17 | 14                    | control | 0    | 0    | 0    |         |     |        |     | 14                    | control | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 14      | control | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 18 | 15                    | control | 1    | 0    | 1    |         |     |        |     | 15                    | control | 3    | 0    | 3    |         |        |          |                       |      |      |      |     |   | 15      | control | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 19 |                       |         |      |      |      |         |     |        |     |                       |         |      |      |      |         |        |          |                       |      |      |      |     |   |         |         |                       |         |       |      |      |     |        |         |       |        |     |
| 20 |                       |         |      |      |      |         |     |        |     |                       |         |      |      |      |         |        |          |                       |      |      |      |     |   |         |         |                       |         |       |      |      |     |        |         |       |        |     |
| 21 | Barn site May 29 2012 |         |      |      |      |         |     | 29-May |     | Barn site Jun 12 2012 |         |      |      |      |         | 12-Jun |          | Barn site Jun 19 2012 |      |      |      |     |   | 19-Jun  |         | Barn Site Jun 26 2012 |         |       |      |      |     | 26-Jun |         |       |        |     |
| 22 |                       | plot    | bug1 | bug2 | gene |         |     |        |     | plot                  | bug1    | bug2 | gene |      |         |        | plot     | bug1                  | bug2 | gene |      |     |   | plot    | bug1    | bug2                  | gene    |       | avr  | SEM  |     |        |         |       |        |     |
| 23 | 1                     | T1      | 3    | 3    | 6    |         |     |        |     | 1                     | T1      | 21   | 0    | 21   |         |        | 1        | T1                    | 5    | 0    | 5    |     |   |         | 1       | T1                    | 0       | 0     | 0    |      | avr | SEM    |         |       |        |     |
| 24 | 2                     | T1      | 1    | 4    | 5    |         | avr | SEM    |     | 2                     | T1      | 36   | 74   | 110  |         | avr    | SEM      | 2                     | T1   | 65   | 502  | 567 |   | avr     | SEM     | 2                     | T1      | 44    | 2057 | 2101 | T1  | 431.8  | 417.33  |       |        |     |
| 25 | 3                     | T1      | 0    | 0    | 0    | T1      | 2.4 | 1.288  |     | 3                     | T1      | 13   | 0    | 13   | T1      | 30.6   | 20       | 10124                 |      |      |      |     |   | T1      | 119.4   | 5                     | 111     | 92882 | 3    | T1   | 12  | 20     | 32      | T2    | 0.4    | 0.4 |
| 26 | 4                     | T1      | 0    | 0    | 0    | T2      | 0.4 | 0.245  |     | 4                     | T1      | 7    | 0    | 7    | T2      | 1      | 0.774597 |                       |      |      |      |     |   | T2      | 5       | 2                     | 1908902 | 4     | T1   | 0    | 16  | 16     | control | 1.2   | 0.5831 |     |
| 27 | 5                     | T1      | 0    | 1    | 1    | control | 1   | 0.316  |     | 5                     | T1      | 2    | 0    | 2    | control | 2.2    | 1.714643 |                       |      |      |      |     |   | control | 2.8     | 0                     | 969536  | 5     | T1   | 0    | 10  | 10     |         |       |        |     |
| 28 | 6                     | T2      | 0    | 0    | 0    |         |     |        |     | 6                     | T2      | 1    | 0    | 1    |         |        |          |                       |      |      |      |     |   | 6       | T2      | 0                     | 8       | 8     |      |      |     |        |         |       |        |     |
| 29 | 7                     | T2      | 0    | 0    | 0    |         |     |        |     | 7                     | T2      | 0    | 4    | 4    |         |        |          |                       |      |      |      |     |   | 7       | T2      | 0                     | 12      | 12    |      |      |     |        |         |       |        |     |
| 30 | 8                     | T2      | 0    | 1    | 1    |         |     |        |     | 8                     | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 8       | T2      | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 31 | 9                     | T2      | 0    | 1    | 1    |         |     |        |     | 9                     | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 9       | T2      | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 32 | 10                    | T2      | 0    | 0    | 0    |         |     |        |     | 10                    | T2      | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 10      | T2      | 2                     | 0       | 2     |      |      |     |        |         |       |        |     |
| 33 | 11                    | control | 0    | 0    | 0    |         |     |        |     | 11                    | control | 1    | 0    | 1    |         |        |          |                       |      |      |      |     |   | 11      | control | 0                     | 5       | 5     |      |      |     |        |         |       |        |     |
| 34 | 12                    | control | 0    | 1    | 1    |         |     |        |     | 12                    | control | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 12      | control | 1                     | 1       | 2     |      |      |     |        |         |       |        |     |
| 35 | 13                    | control | 0    | 1    | 1    |         |     |        |     | 13                    | control | 0    | 0    | 0    |         |        |          |                       |      |      |      |     |   | 13      | control | 0                     | 0       | 0     |      |      |     |        |         |       |        |     |
| 36 | 14                    | control | 0    | 1    | 1    |         |     |        |     | 14                    | control | 8    | 1    | 9    |         |        |          |                       |      |      |      |     |   | 14      | control | 0                     | 5       | 5     |      |      |     |        |         |       |        |     |
| 37 | 15                    | control | 0    | 2    | 2    |         |     |        |     | 15                    | control | 0    | 1    | 1    |         |        |          |                       |      |      |      |     |   | 15      | control | 0                     | 2       | 2     |      |      |     |        |         |       |        |     |
| 38 |                       |         |      |      |      |         |     |        |     |                       |         |      |      |      |         |        |          |                       |      |      |      |     |   |         |         |                       |         |       |      |      |     |        |         |       |        |     |
| 39 |                       |         |      |      |      |         |     |        |     |                       |         |      |      |      |         |        |          |                       |      |      |      |     |   |         |         |                       |         |       |      |      |     |        |         |       |        |     |

### 1) Using multiple tables

The computer reads your table "by row".

Here, a computer will assign to the same sample values from 4 different samples!

# Data Organization in Spreadsheets

## 2) Using multiple tabs

This can look tidy but does not allow you to make data communicating in different tabs. Sooner or later you'll need to collapse all your data in a single table.

## 3) Do not properly indicate real zeros and missing data

- write always all the real zeros
- leave blank (or fill with **NA** values) if data is missing



# Data Organization in Spreadsheets

## 4) Do not use formatting to convey information!

- it will be lost when exporting your table in a text file

### Solution:

Add a new variable encoding which observation will need to be excluded from the analysis.

### **More in general:**

**Don't be afraid to add as much as variables are needed to properly annotate your sample**

| Date collect | Species | Sex | Weight | Calibrated |
|--------------|---------|-----|--------|------------|
| 1/8/14       | NA      |     |        |            |
| 1/8/14       | DM      | M   | 44     | Y          |
| 1/8/14       | DM      | M   | 38     | Y          |
| 1/8/14       | OL      |     |        |            |
| 1/8/14       | PE      | M   | 22     | Y          |
| 1/8/14       | DM      | M   | 38     | Y          |
| 1/8/14       | DM      | M   | 48     | Y          |
| 1/8/14       | DM      | M   | 43     | Y          |
| 1/8/14       | DM      | F   | 35     | Y          |
| 1/8/14       | DM      | M   | 43     | Y          |
| 1/8/14       | DM      | F   | 37     | Y          |
| 1/8/14       | PF      | F   | 7      | Y          |
| 1/8/14       | DM      | M   | 45     | Y          |
| 1/8/14       | OT      |     |        |            |
| 1/8/14       | DS      | M   | 157    | N          |
| 1/8/14       | OX      |     |        |            |
| 2/18/14      | NA      | M   | 218    | N          |
| 2/18/14      | PF      | F   | 7      | Y          |
| 2/18/14      | DM      | M   | 52     | Y          |

# Data Organization in Spreadsheets

## **5) Do not merge cells!**

It will create artifacts or issues when exporting into a text file.

Solution: re-structure your data such as merging cells is not required

- In my experience this is commonly used in table headers!

# Data Organization in Spreadsheets

## 6) Headers should be one line

- see the previous point
- column names should avoid problematic characters
  - symbols (°, ?, %, !, +, [], () )
  - spaces
- use underscore ( \_ ) or **camel case** notations

Example:

Root diameter (mm) ->    Root\_diameter    or    RootDiameter

- keep it as simple as possible: e.g. RD.

You'll need an annotation file to track the meaning of your codes!

# Data Organization in Spreadsheets

## 6) **do not includes measure units in your data spreadsheet**

Measure units are essential, but:

- do not include in your data (your observations can have all the same measure unit).

If not so: can you convert them to the same unit? Otherwise add a variable indicating the measure unit for each of your observation.

- do not include in your column header.

Compile e README file writing annotation of your column names.

# Data Organization in Spreadsheets

## 6) **do not includes measure units in your data spreadsheet**

Measure units are essential, but:

- do not include in your data (your observations can have all the same measure unit).

If not so: can you convert them to the same unit? Otherwise add a variable indicating the measure unit for each of your observation.

- do not include in your column header.

Compile e README file writing annotation of your column names.

# Data Organization in Spreadsheets

| R   | S   | T   |
|---|---|---|
| L media lato corto dente                        | L media lato lungo dente                      | Ampiezza media rachide                        |
| m 0,106 cm + 0,121 cm + 0,094 cm / 3 = 0,107 cm | 0,294 cm + 0,342 cm + 0,237 cm / 3 = 0,291 cm | 0,093 cm + 0,296 cm + 0,253 cm / 3 = 0,214 cm |
| m 0,085 cm + 0,071 cm + 0,111 cm / 3 = 0,089 cm | 0,234 cm + 0,251 cm + 0,329 cm / 3 = 0,271 cm | 0,155 cm + 0,189 cm + 0,218 cm / 3 = 0,187 cm |
| m 0,086 cm + 0,115 cm + 0,100 cm / 3 = 0,100 cm | 0,323 cm + 0,242 cm + 0,367 cm / 3 = 0,311 cm | 0,150 cm + 0,320 cm + 0,244 cm / 3 = 0,238 cm |
| m 0,072 cm + 0,069 cm + 0,074 cm / 3 = 0,072 cm | 0,299 cm + 0,383 cm + 0,257 cm / 3 = 0,313 cm | 0,089 cm + 0,242 cm + 0,313 cm / 3 = 0,214 cm |
| m 0,109 cm + 0,144 cm + 0,201 cm / 3 = 0,151 cm | 0,391 cm + 0,437 cm + 0,415 cm / 3 = 0,414 cm | 0,153 cm + 0,298 cm + 0,365 cm / 3 = 0,272 cm |
| m 0,040 cm + 0,058 cm + 0,029 cm / 3 = 0,042 cm | 0,215 cm + 0,345 cm + 0,216 cm / 3 = 0,258 cm | 0,187 cm + 0,247 cm + 0,230 cm / 3 = 0,221 cm |
| m 0,026 cm + 0,037 cm + 0,025 cm / 3 = 0,029 cm | 0,210 cm + 0,238 cm + 0,148 cm / 3 = 0,199 cm | 0,193 cm + 0,310 cm + 0,295 cm / 3 = 0,266 cm |
| m 0,033 cm + 0,027 cm + 0,048 cm / 3 = 0,036 cm | 0,169 cm + 0,165 cm + 0,238 cm / 3 = 0,190 cm | 0,107 cm + 0,283 cm + 0,328 cm / 3 = 0,239 cm |
| m 0,031 cm + 0,044 cm + 0,036 cm / 3 = 0,037 cm | 0,188 cm + 0,251 cm + 0,208 cm / 3 = 0,216 cm | 0,165 cm + 0,247 cm + 0,367 cm / 3 = 0,260 cm |
|   |   |   |
| m 0,078 cm + 0,042 cm + 0,048 cm / 3 = 0,056 cm | 1,795 cm + 1,748 cm + 0,062 cm / 3 = 1,201 cm | 0,075 cm + 0,229 cm + 0,338 cm / 3 = 0,214 cm |
| m 0,040 cm + 0,031 cm + 0,051 cm / 3 = 0,041 cm | 2,038 cm + 1,062 cm + 1,083 cm / 3 = 1,394 cm | 0,111 cm + 0,218 cm + 0,353 cm / 3 = 0,277 cm |
| m 0,100 cm + 0,033 cm + 0,018 cm / 3 = 0,050 cm | 0,158 cm + 0,202 cm + 2,273 cm / 3 = 0,878 cm | 0,120 cm + 0,359 cm + 0,539 cm / 3 = 0,339 cm |
| m 0,079 cm + 0,051 cm + 0,035 cm / 3 = 0,055 cm | 0,169 cm + 0,129 cm + 1,645 cm / 3 = 0,648 cm | 0,144 cm + 0,220 cm + 0,422 cm / 3 = 0,262 cm |
| m 0,029 cm + 0,032 cm + 0,042 cm / 3 = 0,034 cm | 0,107 cm + 0,140 cm + 1,501 cm / 3 = 0,583 cm | 0,146 cm + 0,324 cm + 0,325 cm / 3 = 0,265 cm |
| m 0,030 cm + 0,039 cm + 0,037 cm / 3 = 0,035 cm | 0,188 cm + 0,201 cm + 0,065 cm / 3 = 0,151 cm | 0,102 cm + 0,199 cm + 0,291 cm / 3 = 0,197 cm |
| m 0,033 cm + 0,023 cm + 0,034 cm / 3 = 0,030 cm | 0,061 cm + 0,099 cm + 0,195 cm / 3 = 0,118 cm | 0,053 cm + 0,149 cm + 0,238 cm / 3 = 0,147 cm |
| m 0,036 cm + 0,023 cm + 0,014 cm / 3 = 0,024 cm | 0,205 cm + 0,169 cm + 0,820 cm / 3 = 0,398 cm | 0,118 cm + 0,182 cm + 0,262 cm / 3 = 0,187 cm |
| m 0,064 cm + 0,040 cm + 0,038 cm / 3 = 0,047 cm | 0,131 cm + 0,149 cm + 2,123 cm / 3 = 0,801 cm | 0,130 cm + 0,322 cm + 0,436 cm / 3 = 0,296 cm |
| m 0,056 cm + 0,025 cm + 0,027 cm / 3 = 0,036 cm | 0,126 cm + 0,198 cm + 0,179 cm / 3 = 0,168 cm | 0,182 cm + 0,297 cm + 0,302 cm / 3 = 0,260 cm |
| m 0,033 cm + 0,045 cm + 0,041 cm / 3 = 0,040 cm | 0,134 cm + 0,201 cm + 0,183 cm / 3 = 0,173 cm | 0,111 cm + 0,143 cm + 0,189 cm / 3 = 0,148 cm |
| m 0,088 cm + 0,086 cm + 0,068 cm / 3 = 0,081 cm | 0,495 cm + 0,241 cm + 0,214 cm / 3 = 0,317 cm | 0,121 cm + 0,199 cm + 0,245 cm / 3 = 0,183 cm |
| m 0,024 cm + 0,020 cm + 0,027 cm / 3 = 0,024 cm | 0,328 cm + 0,243 cm + 0,316 cm / 3 = 0,296 cm | 0,078 cm + 0,134 cm + 0,201 cm / 3 = 0,138 cm |
| m 0,126 cm + 0,057 cm + 0,034 cm / 3 = 0,072 cm | 0,138 cm + 1,354 cm + 1,203 cm / 3 = 0,895 cm | 0,059 cm + 0,124 cm + 0,235 cm / 3 = 0,139 cm |
| m 0,055 cm + 0,050 cm + 0,075 cm / 3 = 0,060 cm | 0,163 cm + 0,265 cm + 0,249 cm / 3 = 0,226 cm | 0,165 cm + 0,219 cm + 0,263 cm / 3 = 0,216 cm |
| m 0,088 cm + 0,100 cm + 0,029 cm / 3 = 0,072 cm | 0,221 cm + 0,212 cm + 1,975 cm / 3 = 0,803 cm | 0,132 cm + 0,324 cm + 0,617 cm / 3 = 0,358 cm |
| m 0,101 cm + 0,049 cm + 0,054 cm / 3 = 0,068 cm | 0,205 cm + 0,130 cm + 2,370 cm / 3 = 0,902 cm | 0,117 cm + 0,433 cm + 0,663 cm / 3 = 0,404 cm |
| m 0,185 cm + 0,037 cm + 0,086 cm / 3 = 0,103 cm | 0,102 cm + 1,315 cm + 0,165 cm / 3 = 0,527 cm | 0,168 cm + 0,382 cm + 0,452 cm / 3 = 0,334 cm |
| m 0,072 cm + 0,046 cm + 0,065 cm / 3 = 0,061 cm | 0,196 cm + 0,165 cm + 1,636 cm / 3 = 0,666 cm | 0,161 cm + 0,489 cm + 0,675 cm / 3 = 0,441 cm |
| m 0,056 cm + 0,066 cm + 0,037 cm / 3 = 0,053 cm | 2,316 cm + 0,172 cm + 1,907 cm / 3 = 1,465 cm | 0,112 cm + 0,197 cm + 0,359 cm / 3 = 0,223 cm |
| m 0,053 cm + 0,040 cm + 0,054 cm / 3 = 0,049 cm | 0,160 cm + 0,881 cm + 0,149 cm / 3 = 0,397 cm | 0,069 cm + 0,171 cm + 0,238 cm / 3 = 0,159 cm |
| m 0,034 cm + 0,040 cm + 0,029 cm / 3 = 0,024 cm | 0,083 cm + 0,084 cm + 1,206 cm / 3 = 0,458 cm | 0,091 cm + 0,143 cm + 0,247 cm / 3 = 0,160 cm |
| m 0,092 cm + 0,045 cm + 0,041 cm / 3 = 0,059 cm | 0,152 cm + 0,302 cm + 1,298 cm / 3 = 1,752 cm | 0,126 cm + 0,241 cm + 0,345 cm / 3 = 0,237 cm |
| m 0,042 cm + 0,047 cm + 0,045 cm / 3 = 0,045 cm | 0,137 cm + 0,061 cm + 0,055 cm / 3 = 0,084 cm | 0,166 cm + 0,267 cm + 0,242 cm / 3 = 0,218 cm |
| m 0,030 cm + 0,043 cm + 0,030 cm / 3 = 0,034 cm | 0,070 cm + 0,077 cm + 0,082 cm / 3 = 0,076 cm | 0,085 cm + 0,175 cm + 0,210 cm / 3 = 0,157 cm |
| m 0,036 cm + 0,035 cm + 0,025 cm / 3 = 0,032 cm | 0,087 cm + 0,129 cm + 1,442 cm / 3 = 0,553 cm | 0,080 cm + 0,165 cm + 0,260 cm / 3 = 0,168 cm |
| m 0,059 cm + 0,042 cm + 0,029 cm / 3 = 0,043 cm | 0,112 cm + 0,182 cm + 1,615 cm / 3 = 0,636 cm | 0,069 cm + 0,136 cm + 0,241 cm / 3 = 0,149 cm |
| m 0,086 cm + 0,067 cm + 0,036 cm / 3 = 0,063 cm | 0,134 cm + 0,149 cm + 0,166 cm / 3 = 0,150 cm | 0,109 cm + 0,212 cm + 0,295 cm / 3 = 0,205 cm |

# Data Organization in Spreadsheets

## 7) Write your annotations for every sample

- Computers are very literal. If you do not write in each row sample information, your computer won't understand where is the sample from

| SampleID | Site   | plot | root_weight |
|----------|--------|------|-------------|
| Plant_1  | Site 1 | 1    | 0.56        |
| Plant_2  | ?      | 2    | 0.8         |
| Plant_3  | ?      | 3    | 0.59        |
| Plant_1  | Site 2 | 1    | 0.7         |
| Plant_2  | ?      | 2    | 0.69        |
| Plant_3  | ?      | 3    | 0.92        |

Each row must  
be unique!

## 8) Include your replicate number, but only for tracking purposes

Most of the analyses do not require a replicate number!

Often they are stored along with the sample name -> split in a new variable!

# Some notes about date/hour formatting

- Storing dates/times in one field in the format (“15/01/2024”) can cause compatibility issues between softwares
- Storing dates as YEAR, MONTH, DAY in separate columns eliminates any ambiguities!
- as a single string YYYYMMDDhhmmss format (or YYYYMMDD for date only)
- as YEAR, DAY-OF-YEAR (**DOI**):

“=A1-DATE(YEAR(A1);1;0)” where A2 is the date”

see

> **library**(anytime) # in R for format conversion!



# Data Organization in Spreadsheets

**Do your exercise!**

```
>download.file("https://ndownloader.figshare.com/files/2252083",  
               "survey_data_spreadsheet_messy.xls")
```