

The lab today will use several datasets:

- `data1-bees.csv`: Yearly observations of bee species in Belgium (gbif.org)
- `data2-pesticides.csv`: Pesticides used per hectare in selected countries (fao.org)
- `data3-climate.csv`: Daily weather recordings from Ireland (data.europa.eu)

Note: This data is cyclic, which means there are implied patterns. Statistics on this dataset may be described per year (aggregating days) or per day (aggregating years).

Your submission today should be written text and visuals, as a brief report about your dataset.

1. Give *descriptive* statistics about the data set:
 - (a) The number of parameters, the number of samples in each parameter
 - (b) The smallest, largest, average value
 - (c) Visualize your data as a histogram and as a boxplot
 - (d) Identify any peculiar characteristics (outliers, several dense “clusters”, long tail, ...)
2. Give *inferential* statistics about pairs of parameters in your data set:
 - (a) Perform a linear regression on pairs of parameters.
 - (b) Read about the *t*-test here: statology.org/welchs-t-test, and determine if the *t*-test applies to pairs of parameters in your dataset.
 - (c) Visualize the linear regression of two parameters that have the highest *r*-squared value.
 - (d) **Bonus:** Perform a linear regression with a polynomial (instead of a line)
3. Submit your work as a PDF on ORTUS. There should be (at least) two written paragraphs, each with an accompanying diagram.