

DV01 – Datasets. Basic Chart Types

1. Make some data

We will create our own dataset to practice with. Fill in the survey that you've got sent by email. You can download the responses [here](#) (File->Download as->csv)

2. Reproducible Research

Look at the example notebooks DV01-example-BasicCharts.ipynb and DV01-example-MoreCharts.ipynb

Use them as an inspiration to create a new similar basic analysis notebook for the group dataset. Clean the data as needed, make basic charts, document all steps.

Try to answer visually questions like:

- Are older students taller?
- Do Europeans have a smaller family?
- Make a histogram for every quantitative attribute. What are these charts saying?
- Is there any correlation between star sign and work energy?
- Which is the most creative star sign?
- Find a way to show the study profile, investment in DV and age, all-in one chart.

3. Attribute types

Examine each attribute in the group dataset. What is its datatype?

What about the attributes in the cities dataset?

What about your own chosen dataset (assignment 8)?

4. Scatter matrix and parallel coordinates*

Study, using example DV01-example-MoreCharts.ipynb and internet, what a scatter matrix is, what a parallel coordinates chart is, and how you can make them with pandas. Use these chart types in your group notebook.

5. More complex charts**

Copy/paste the group dataset on RAWGraphs (<http://rawgraphs.io/>) and experiment with even more chart types. Do you understand these more complex charts? Are there any interesting outcomes? Write down your findings.

6. Group selfie

Upload your most interesting charts, as a commented image, on the groupselfie padlet.

7. Get organised [= Step 0 of the main project]

8. Find your dataset! [= Step 1]

See the project description document and assignment A1 on canvas – deadline Sunday 17 september