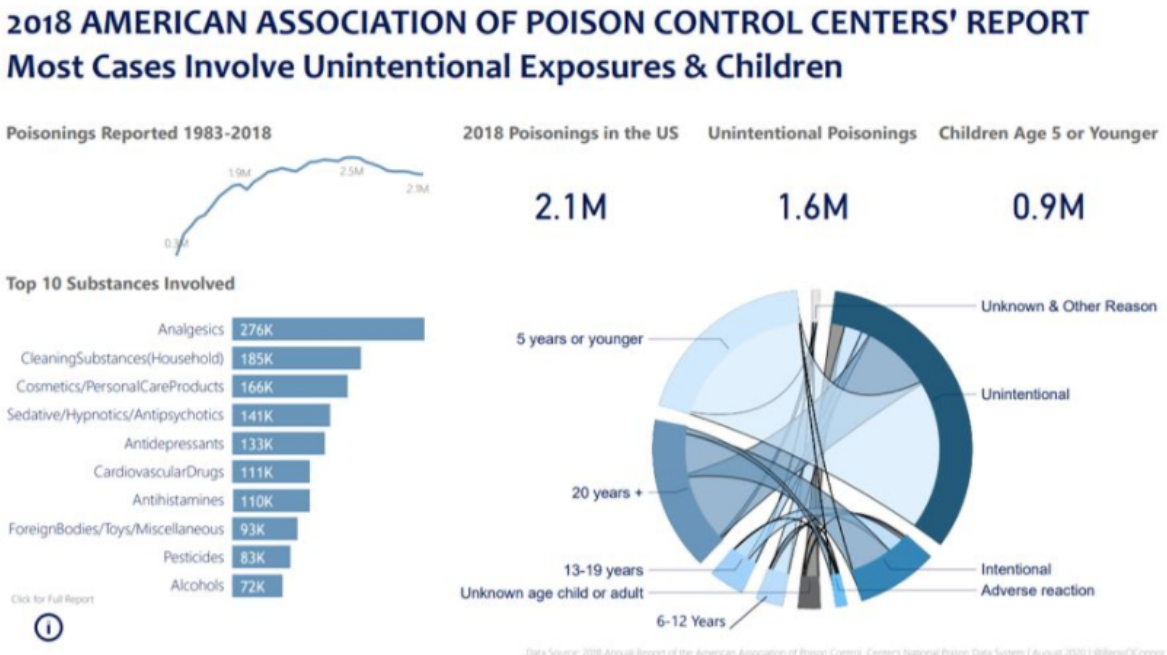
University of Trento --- M.Sc. Data Science

Data Visualization Lab Exam
14 July 2023

> *Using either a Jupyter and/or a R Notebook generate the HTML files to submit.*

Answer the questions (in a Markdown cell / as plain text) and solve the exercises listed hereafter:

1. [0-5 points] Describe in detail the meaning of the visual encoding elements in the following infographic reporting statistics on children poisoning in 2018



2. [0-5 points] Discuss in detail the concept of geometric projection in map representation. Explain pros and cons of the different types of

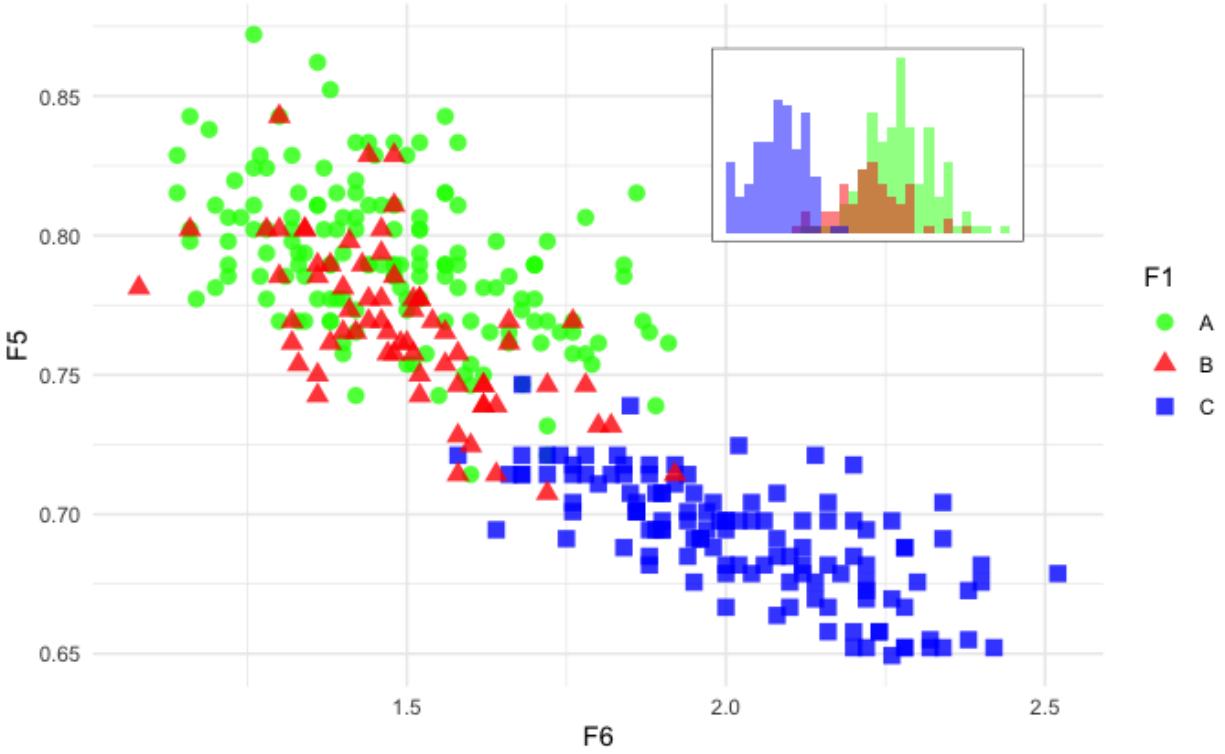
projections with graphical examples different from those shown in the course slides.

3. [0-5 points] Discuss in detail the parameters of the UMAP function (number of neighbors, minimum distance, number of components, metrics…) and show their impact on the UMAP projection with graphical examples different from those shown in the course slides.

4. [0-7 points] The datafile Food_Production collects information about crop production in EU countries across several years. Prepare an infographic with two panels describing some patterns inferred by (a subset of) the dataset, including at least one geographic map.

5. [0-7 points] Consider the Breast_cancer dataset, where a cohort of samples (on the rows) are described by the values of their gene expression levels (columns 3 to the end). Column 1 is the sample name, column 2 is the sample class (DCIS or IBC). Use a 2D dimensionality reduction dataset with optimized parameters to separate the two sets of samples DCIS and IBC, motivating the choice of the algorithm and the optimization of the parameter.

6. [0-7 points] Using the dataset Ex6, try to replicate the following plot.



Send by email to *giuseppe.jurman@unitn.it*

1. **The HTML files produced by the notebooks (compulsory): please name it as** *name_surname.html* or *name_surname_1.html* (if you have more than one HTML file).
2. The notebook(s) (Python,R)
3. All the files needed to correctly compile the notebooks (images, .json, etc)

Notes:

- Exam is passed when at least 18 points are earned.
- If more than 30 points are achieved, the corresponding mark will be "30 cum laude"
- Use of the internet is allowed, but the candidate is expected to work individually.
- Check your mailbox in the next few hours, I may contact you for missing files or replacing corrupted files.