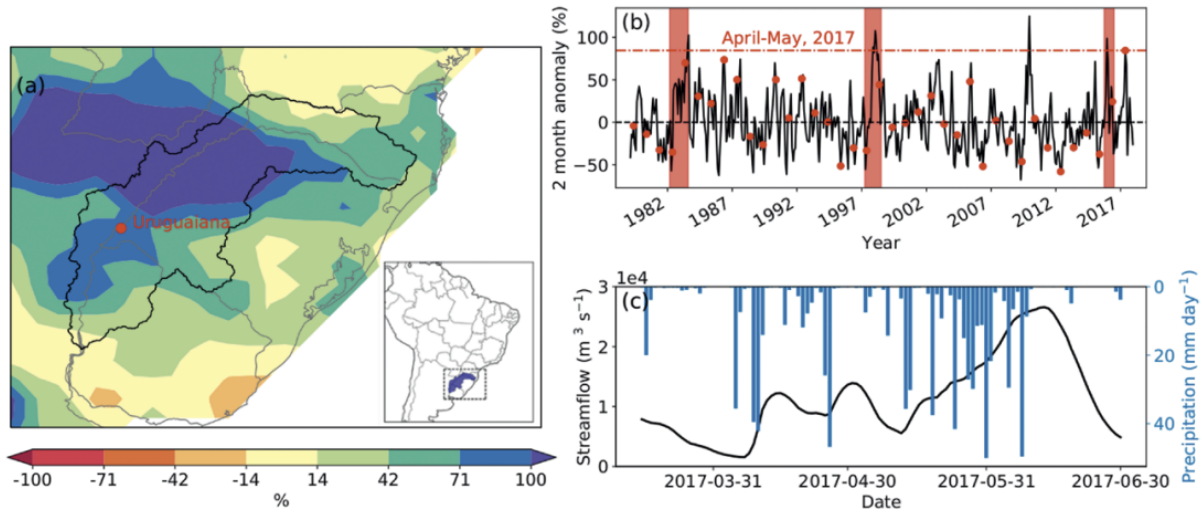
University of Trento --- M.Sc. Data Science

Data Visualization Lab Exam
09 September 2022

> Create a Jupyter and/or R Notebook, named
> *name_surname.{ipynb,Rmd}*

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 data about anomalous precipitations in the Uruguay basin.



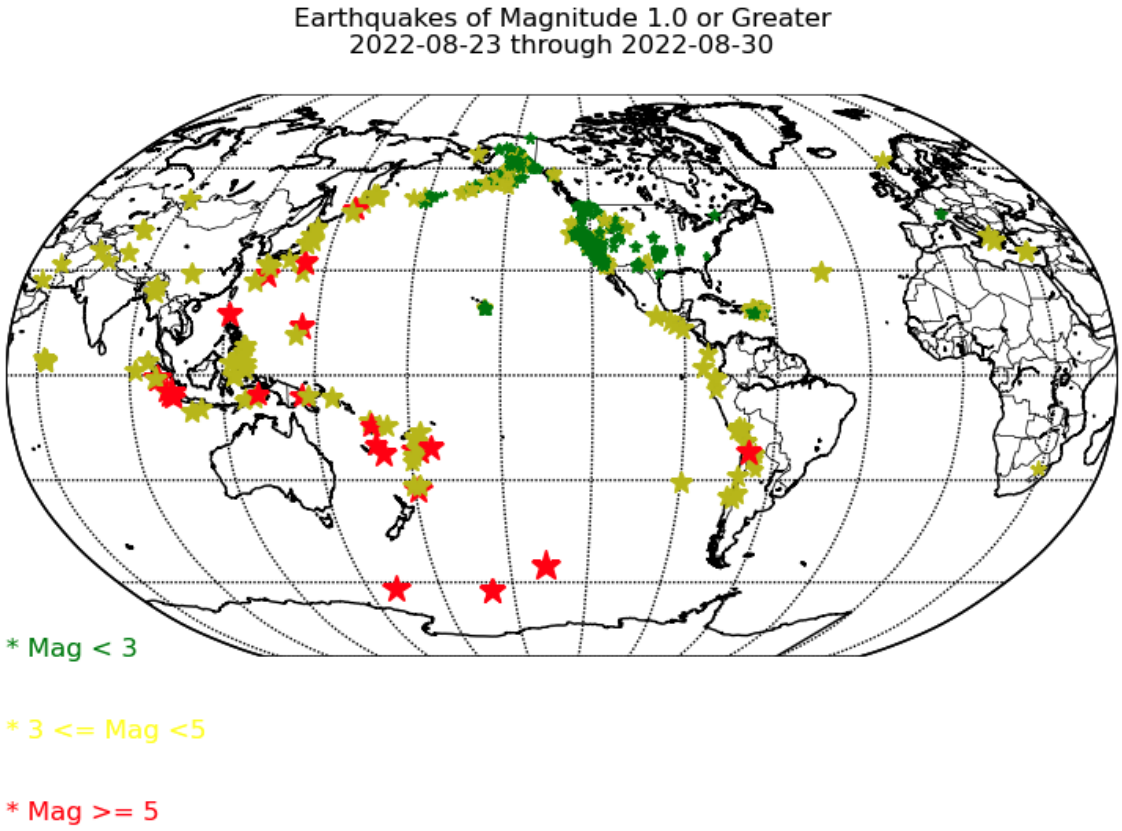
2. [0-5 points] Discuss in detail the use of text in data visualization. Then present two dataviz examples, different from those used in the course slides, showing good and bad use of text, respectively.

3. [0-5 points] Discuss in detail the PCA algorithm and its pros and cons when used as a dimensionality reduction method. Provide also two examples, different from those used in the course slides, highlighting these pros and cons.

4. [0-7 points] The datafile AU_Car_Crash collects information about car accidents in Australia in the period 1989-2021. Prepare an infographic with two panels describing some patterns inferred by (a subset of) the dataset, including a map.

5. [0-7 points] Consider the World_Happiness_Index dataset, where 146 countries are ranked by decreasing happiness. Prepare 4 dimension reduction data visualization using PCA, MDS, tSNE and UMAP with optimized parameters, comment on differences and similarities among the four plots, and discuss the obtained clusters with respect to the original ranking.

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



Email the notebook(s) to *giuseppe.jurman@unitn.it* and please **wait for confirmation of correct receipt of the files before leaving the room**.

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.