



POLITECNICO
MILANO 1863

Brambilla Mattia
mattia.brambilla@polimi.it

MTAFFI 2023 - Assignment

Objectives of the assignment

1. Improve self abilities in Python programming
2. Practice with an unknown dataset by developing individual code from scratch (i.e., from an empty Colab notebook)
3. Acquire self competences in solving coding problems by reading the documentation
4. Move from didactical examples to hands-on real problem to acquire expertise in handling real data and derive conclusions from data analyses
5. Train presentation skills (also improving the quality of Python plotting)
6. Get up to +3 points in the final mark of the course

What you have

You have received for the first time the ClimateFish dataset, which collects data for 7 Mediterranean indigenous and 8 nonindigenous fishes, proposed as candidate indicators of climate change.

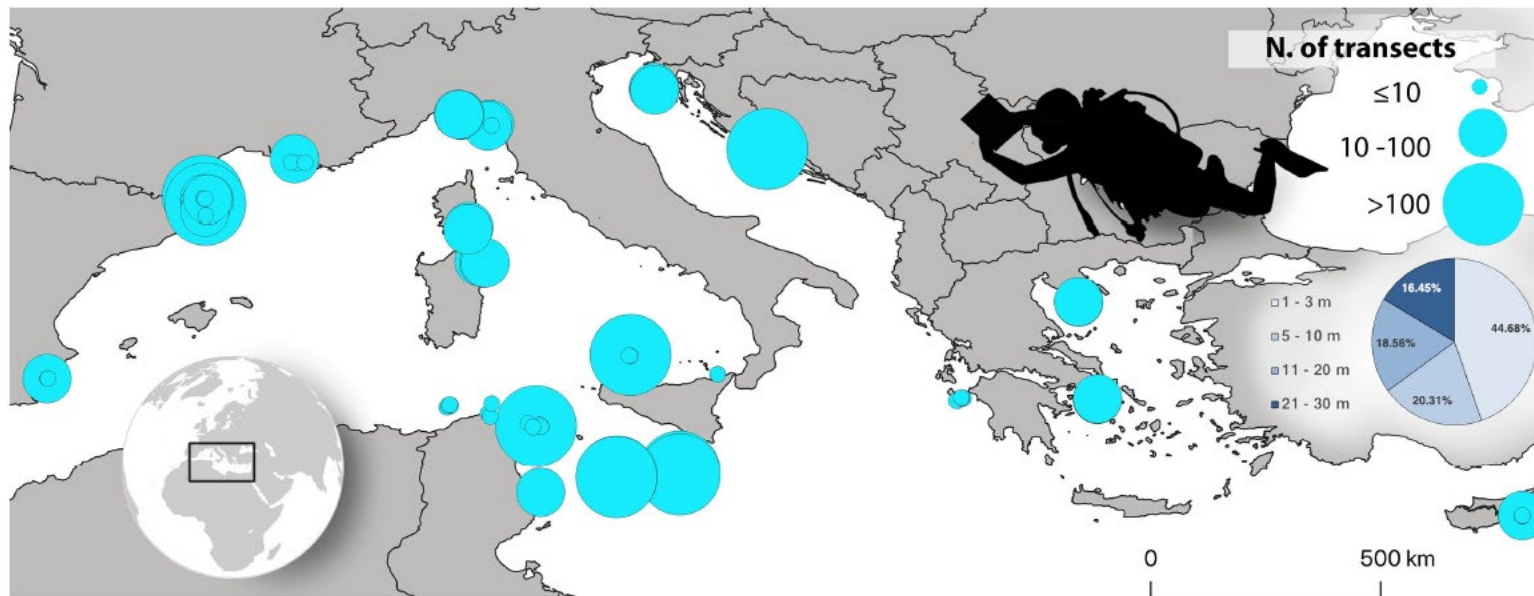
These species have been selected by a network of Mediterranean scientists based on their wide distribution, responsiveness to temperature conditions and easy identification.

Overall, there are 101771 observed individuals belonging to the 15 target species. Counts were realized along 3142 transects carried out in 7 Mediterranean countries between 2009 and 2021.

The dataset description is here: [Frontiers | ClimateFish: A Collaborative Database to Track the Abundance of Selected Coastal Fish Species as Candidate Indicators of Climate Change in the Mediterranean Sea \(frontiersin.org\)](https://www.frontiersin.org/articles/10.3389/fmars.2022.882111/full)

Fish censuses:

- 9 years in Malta (Cirkewa, Dwejra and Gnejna), Croatia (Ciovo Jug, Ciovo Punta, and Marjan), Sicily (Faraglioni, Mannarazza, and Pozzolana)
- 6 years in Spain (Les Cliques, Punta Tees Frares, and Sarnella)
- 5 years in Corse (Capu di Fora, La Vardiola, and Punta Arasu)
- and more...



Assignment_description: this set of slides

Dataset_ClimateFish: a .csv file containing all the 3142 transects to be used in Colab

Dataset_description: a summary description of each variable (i.e., each column) of the dataset

ClimateFish_Paper: the .pdf of the dataset description [Frontiers | ClimateFish: A Collaborative Database to Track the Abundance of Selected Coastal Fish Species as Candidate Indicators of Climate Change in the Mediterranean Sea \(frontiersin.org\)](https://www.frontiersin.org/articles/10.3389/fmars.2019.00041/full)

You have to present the results of your analyses to a board of scientists expert in climate change.
Your presentation is targeted to answer the following questions:

1. What is the dataset about? i.e., overview of the dataset (assume the scientists do not know anything)
2. Are the environmental conditions of the analyzed countries similar?
3. Are there significant changes from 2009 to 2021?

At the end of your analyses, discuss what are (if any) the most evident variations that deserve high attention for further researches. If yes, highlight them. If no, what do you conclude from the 10+ years data?

Consider both aggregated statistics (e.g., overall number of observations for the 15 species), state-dependent statistics (e.g., overall number of observations for the 15 species for each country), or even more granular analyses for each site.

Note that the above examples are just illustrative, you are requested to provide analyses according to your decisions.

Remember you have multiple variables on the days, months, years, country, study area, location, depth, projects. **FREEDOM TO CARRY OUT LOT OF DIFFERENT ANALYSES**

Notice that the censuses are carried out by different research groups and different locations, maybe there exist correlations.

Maybe there are some groups of similar sites where there are similarities in the observations.

Pay attention to the different number of observations for the sites (i.e., do not compare sites with few observations and sites with thousands of observations, unless...)

Remember to emphasize the visualization of the results by choosing the correct type of plot.

Decide by yourselves your team. No more than 4 students per team.

Create your team here: [MTAFFI - Assignment groups.xlsx](#)

Deadline for the team composition: 14 Nov.

The sooner you decide your team, the sooner you can start working in Colab.

Wednesday 15th Nov.: supervised class work. Meet here in class and carry on your work, with my supervision for assistance, suggestions, checks, doubts, etc.

Upload on WeBeep delivery folder your assignment by Tue 5th Dec.

Files to be uploaded: a working Colab notebook + presentation

Wednesday 6th Dec.: project presentation in class (max 10 min).

Evaluation modality: correctness of the analyses (50%) + quality of the presentation, both oral and written (50%)

Each group can get from +0 to +3 points (half points are valid as well)