Data Science Competition 2023

Introduction

Context

You're a group of young data scientists, freshly graduated from the Department of Informatics of the University of Bergen. A biology lab has contacted you to uncover the mysteries of the measurements they made in their last experiments. They had some vague objectives when doing the experiments, but don't have the expertise to analyse the dataset themselves and struggle to find the right words when trying to explain to you what to look for in the data. So here you are, ready to accomplish your first task as a data scientist!

You are allowed to work in groups (up to 3 people in a group) or alone and you can ask questions to your fellow data scientists on Discord, server: "Data Science Competition 2023" (invite link). You are also allowed to ask questions to the lab that contacted you, represented here by Natacha Galmiche (natacha.galmiche@uib.no) either on the "general-questions" and "technical-issues" channels or by direct messages on Discord if your question really needs to be private.

Tasks

[SPOILER] *This part will be available on Friday November 3 at 17:00* ***[SPOILER]***

Delivery

In research everything should be reproducible, therefore, the biology lab wants to get your source code. However, the biology researchers are not experts in programming, so they expect you to share your results and explain your methodology without having to run your code. Thus, rather than a simple python file (.py) they would prefer a **jupyter notebook** (.ipynb) file (using python or R, and potentially using Google Colab) **with the output of each cell visible and comments on your results and methodology throughout the notebook**. If you can only provide a python/R script without using a notebook, then they would need a report (as a .pdf file) explaining your general approach, and with your results and comments on your results.

Once all the deliveries are received, the lab will give you feedback! There will be winners in each category. The category of a group is defined as the highest category among its members (e.g. a group of 2 bachelor and 1 master students is in the master category).

It is planned to **announce the winners of the competition on Friday, November 17, from 14:00 in the Big Auditorium**. We will keep you updated on that.

What to submit:

- jupyter notebook (.ipynb) (+ optionally a .pdf)
- OR python/R script + pdf.

Where to submit: by email to natacha.galmiche@uib.no.

Deadline: Sunday, 5th November at 23:59.