Hackathon: Data Acquisition, Management, and Visualization

Emanuele Guidotti[*]

Description

In this 2-day, full-time lab, you will learn how to create, deploy, and maintain a real-world data-driven platform. Through the case study of the COVID-19 Data Hub, we will see how to design and develop an international platform that now serves thousands of users worldwide with more than 5 million downloads. You will work on similar challenging projects in the hackathon that we start at the end of day 1, work overnight, and finish with your pitch at the end of day 2.

From a technical side, you will gain hands-on experience with code management in Github, cloud computing with Google Cloud Platform, shell scripting in Linux, database management, advanced R/Rshiny and Python programming.

This is an advanced and intensive lab. Students who apply should be confident with at least 75% of the topics covered and autonomously gain a basic understanding of the remaining 25% before the beginning of the experience. This 2-day, full-time lab, will be held on the weekend so that 1) you will have no other commitments and 2) only highly motivated students will be willing to sacrifice their weekend and join us.

Duration

2-day full time + overnight between day 1 and 2.

Calendar

- Saturday 28 May 2022
    - 9:00 – 12:30 on site at the University (room 22)
    - 15:00 – 00:00 online on Teams
- Sunday 29 May 2022
    - 00:00 – 19:00 online on Teams

Contents

Saturday

I will illustrate in detail the workflow used to build the COVID-19 Data Hub from scratch. In particular:

- workflow design
- data acquisition
- managing the code on Github
- managing the data on the cloud
- automating tasks via cron jobs

From this setup, we will see how to create an SQL database and an interactive visualization tool on top of it.

At the end of the morning, we will discuss and select the projects you will start to work on in the hackathon in groups of 2-3 people. I will propose some projects, but you are encouraged to bring your own ideas.

---

[*]Institute of Financial Analysis, University of Neuchâtel: Rue Abram-Louis-Breguet 2, 2000 Neuchâtel, Switzerland. E-mail: emanuele.guidotti@unine.ch

You will start to work on the projects in the afternoon. We can connect on Teams any time to answer questions and help out with the projects. Around 18:30, we meet all together on Teams to take stock of the situation.

Overnight

Have fun with your projects! I'll be available remotely.

Sunday

Complete your project and pitch it at the end of the day.

Eligible students

Max 10 students.

Requirements

This is an advanced and intensive lab. Students who apply should be confident with at least 75% of the following topics and autonomously gain a basic understanding of the remaining 25% before the beginning of the experience. The topics are:

- Github
- Shell scripting (in Linux)
- SQL
- R or Python

Moreover, by the beginning of the lab you should have created:

- your account on GitHub (free)
- your account on Google Cloud Platform (free tier up to $300, definitely enough for our purposes. The free credits last 90 days, do not create the account too early.)

How to apply

Send an email at dse@unimi.it, with your CV and cover letter explaining why you want to join this lab and how you meet the requirements. Please include in your CV all relevant links to you Github account, previous projects, your website or portfolio.

- *email*: dse@unimi.it
- *subject*: Hackathon: Data Acquisition, Management, and Visualization
- *attachments*: CV and cover letter

Deadline to apply: Sunday 24 April 2022 at 23:59