CAIM: Cerca i Anàlisi d'Informació Massiva

Grau en Enginyeria Informàtica, UPC

September 08, 2025

Instructors

- Marta Arias (most theory + lab group 13)
  - marta.arias.v@upc.edu
  - Omega 322
- David García (some theory + lab groups 11,12)
  - david.garcia.soriano@upc.edu
  - Omega 217

for questions, ask in class or email

Material

- Announcements, course slides and lab coursework via the Racó; please check regularly!

Development

Theory:

- ▶ instructor explains theory; slides are not fully self-contained
- ▶ for each topic, one set of problems to be solved mostly *on your own*
- ▶ *some* problems will be solved in class

Lab:

- ▶ **python** notebooks and scripts
- ▶ to be solved in pairs; part during lab, *part outside class*
- ▶ hand report of each lab project, detailed rules for what/when/how to submit solutions in scripts
- ▶ hopefully feedback will be given early
- ▶ reports to be delivered through the Racó

Evaluation

Check "guia docent"

- ▶ P1: primer parcial
- ▶ P2: segon parcial
- ▶ F: grade for final exam
- ▶ L: average grade of lab reports

Final grade computed as:

$$\max(30\% P1 + 30\% P2, 60\% F) + 40\% L$$

*The grade assigned to the competencia transversal **autonomous learning** will be computed from exam responses and/or information reflected on project reports from a topic proposed by the instructor that students will have to learn on their own.*

Contents (tentative)

First half:

- ▶ Search/Analysis of Textual Information
- ▶ Search/Analysis of Linked Structures
- ▶ Aspects of the architecture of a Web Search System

Second half:

- ▶ Aspects of *Big Data* systems
- ▶ Social Network Analysis
- ▶ Recommender Systems