LSC Project 2021/22
Catania, Delzanno, Guerrini
B.Y.O.L: Build Your Own Lab

Goal

The goal of the LSC practical project is to prepare the material for a possible lab activity for other LM students on a topic related to the LSC course, e.g.,

- APIs of the Spark ecosystems: advanced aspects of RDD/Dataframe/ SparkSQL/streaming Spark, libraries such as GraphX and MLlib
- Optimisations issues such as custom partitioners, parameter tuning, performance monitoring, etc
- Integration with data ingestion/analytics systems/services
- Analysis of a particular problem (either batch or real-time)

The proposed exercises must be complementary to those presented in the course.

Assignment

Each group (max 3 students) is required to submit a short document (max 5 pages, notebook or PDF) divided in 3 parts:

- P1: a presentation of the chosen topic and rationale behind the proposal.

- P2: a report describing at least 5 programming questions related to the topic with links to the source code (e.g. on the user_lsc_XX home folder).

- P3: solutions and timings of the proposed exercises both on local and cluster mode (when applicable).

Submission

For each exam session (January, February, June, July, September), the project material must be submieed via Aulaweb at least 5 days before the date of the wrieen exam.