Phase III (deadline: May 31)

Make a Python notebook containing the following sections:

1. OLAP queries

    a. Check and revise the queries defined in Phase I. Ensure that the queries are meaningful for the processes at hand

    b. For each query type identify dimensions, attributes and measures for each one

    c. Write the SQL queries and execute them. Make sure the results are understandable

2. Bottlenecks

    a. Identify the most important bottlenecks and execute performance analysis

    b. Propose a plan for improving performance and implement it

        • NOTE, if for improving performance, the existing database schema is to be changed, just describe the proposed changes – do NOT implement it!

    c. Verify that the proposed modifications have better performance

    d. Discuss: With the full use of the data warehouse what are the expectation for its growth in 1, 5 and 10 years. Identify what approaches would be used to address such an increase in data

3. Do a critical assessment of this full project

    a. For each of the 3 phases of the project discuss what went right, wrong and what would you do to make it better?

    b. For each phase propose a grade (1-20) Justify your grade

The notebook should be produced with full output (that should not be longer than necessary!). **If images are present, ZIP the notebook with all the required images attached**

The Notebook can be written in Portuguese or English. It is to be submitted on the course's Moodle **until 23:59 of May 31**. On Section Zero it must contain the **group number, the Student Id's and names, and an estimation of how many hours each student contributed to this phase**