Group Project

Development of a Scientific Poster on Data-Based Decision Making

The project is carried out in groups of 2. The enrolment is done on the moodle platform, where it is possible to access the identification number of each group.

Objective: The aim of this project is to apply the concepts and techniques explored throughout the course to develop a scientific poster that presents the results of a data science project focussed on decision-making. The group will have to identify one or more concrete problems that can be solved by analysing data, using data science tools and methodologies and applying the concepts of decision making to substantiate the recommendations and conclusions.

Main Tasks:

Listed below are the main tasks of the project. Other tasks can and should be carried out to complement those listed below.

T1. Select one or more problems, using the data provided, whose focus is decision support. Justify the choice of problem and its relevance in the context of decision-making.

T2. Prepare the data to answer the problem(s).

T3. Apply Data Mining techniques to extract insights from the data, using Python to perform the analyses.

T4. Create data visualisations that allow clear communication of the results, such as graphs, tables and diagrams. Visualisations should be developed based on data visualisation principles discussed in class.

T5. Interpret the results obtained and relate them to the problem identified. Based on the data analysed, present a decision proposal supported by the evidence extracted from the data.

T6. Consider the ethical challenges related to the use of data in the project. Describe how issues such as privacy, transparency and accountability should be incorporated throughout the process.

T7. Presentation and communication of the results supported by a scientific poster during a poster session. Evaluation, based on the subject taught, of a poster from another group.

Deliverables:

E1. **Scientific poster** illustrating the work carried out with an emphasis on results and conclusions:
  - Size A0, vertical.
  - In PDF format, file name: **GrupoX_poster.pdf** here **X** is the Group's identification number.

E2. **Project report** describing the work carried out on all the tasks:
  - Maximum 10 pages, including cover, table of contents and references, annexes accounted for separately.
  - The structure of the report is up to each group, but it must include:
    - Introduction and objectives
    - Description of data and methods
    - Results and visualisations
    - Discussion and conclusions
  - In PDF format, file name: **GrupoX_report.pdf**

E3. **Code** developed to realise the proposed project:
  - Jupyter notebook file, or Python, with the name: **GrupoX_code**

E4. If you make changes using Excel (or another spreadsheet) to the data provided, you must send the new version of the spreadsheet and mention the changes made in the report. The name of the new data file must include the identification of the group (**GroupX**). The format must be csv or xls.

E5. **Peer review report** of a poster:
  - Maximum 2 pages.
  - PDF format, file name: **GrupoX_peer_review.pdf.**
  - Alternatively, if it is handwritten, the format can be an image (photograph), make sure it is legible. If this is not the case, this component will not be assessed.
  - The structure of the report should follow the template found on the moodle platform under the name **peer_review_doc.doc**.

If the file names and formats do not correspond to the request, the projects submitted may not be evaluated.

Dates:

- **13/11/2024** until 23h59: Deliverables E1 to E4, via moodle, in a zip file with the name **GrupoX.zip**
- **16/11/2024** at 9h30: Presentation of scientific posters, it is mandatory for all students to attend the final presentation of practical projects, students who do not attend will not be assessed.
- **16/11/2024** until 13h00: Deliverable E5, via moodle.

Main assessment criteria:

Report:

- Organisation, coherence and cohesion: Logical and clear structure, with well-articulated ideas.
- Definition of problems and objectives: Clarity and relevance of the objectives and how they fit into the decision-making process.
- Data preparation: description of data processing, justified.
- Methodology and Analysis Techniques: description of the exploratory analysis of the data, as well as the modelling techniques applied, justifying their use.
- Interpretation of results: depth of analysis, consistency with the proposed objectives, critical discussion and consideration of ethical issues.

Scientific Poster:

- Visual clarity and design: good visualisation practices
- Communication: clarity in oral presentation and ability to summarise
- Consistency with the report

Peer review: Internal and external critical capacity.

Datasets:

The datasets are available on the moodle platform. The following table shows their distribution among the groups.

Group	Members	Dataset
1	Mariana Baptista Vaag Mkrtchyan	Obesity Risk Dataset
2	Nuno Sousa Erwann Miloux	Employee Satisfaction Survey Data
4	António Teixeira João Pandaio	Customer Satisfaction in Airline
5	Pedro Conceição Ricardo Mororó	Road Accident Severity in India
6	Carolina Paixão Afonso Peças	TikTok User Engagement Data
7	David Paulino Bruno Martins	Obesity Risk Dataset
8	Pedro Victorino Diogo Silva	Employee Satisfaction Survey Data
9	Madalena Sérgio Diogo Cardoso	Customer Satisfaction in Airline
10	Francisco Rodrigues André Santos	Road Accident Severity in India
11	Rafael Etelvino Ana Rita Ferreira	TikTok User Engagement Data