# Project 2

Data Visualization – LECD, LIACD 2025/26

# Introduction

The main goal of the information visualization project is the development and application of a set of skills that allow the construction of visualizations capable of **solving analytical problems** and those related to the **communication of information**. The work is organised in a manner simulating the conditions that occur in a **real-world project**. This includes activities such as communication with the client (teacher) and obtaining feedback, working with data, identification of the problem(s) and possible solutions, sketching and prototyping, implementation of visualization, and validation. It is also intended that the students demonstrate skills in analysis and synthesis, organisation and planning, critical reasoning, application of knowledge in practice, investigation, autonomous learning, and group work.

The project must integrate the following characteristics and requirements:
1. All the code should be produced by the students;
2. The students are free to use any visualization tools and frameworks for initial data exploration;
3. The students should use Plotly frameworks (or similar, e.g., Seaborn or Matplotlib) and Python to implement the core visualization;
4. The students are free to use any additional libraries/modules to support the main code;
5. The students are free to choose the datasets, but they should be approved by the teachers;
6. Interactivity is optional;
7. The visualization should follow best practices of design and visualization;
8. The students should document all the steps of the project in a single progress report.

# Groups

The project should be developed in **groups of 2 or 3 elements**.
It is expected that the groups of 3 elements to deliver projects of higher volume and quality.

# Milestones

The project is divided into several milestones with respective deadlines:
- **Group formation**: We expect you to form a team during the PL class.
- **Project proposal**: For this milestone, you should have completed the following topics of your progress report (see Project section): Authors, Introduction, Related Work, and Data. This will be the first part of your progress report.
- **Midterm review and feedback**: For this milestone, you should have completed the Design Requirements and Design steps of your project. It is not mandatory to entirely complete the design phase, but it is highly recommended. It is your opportunity to show your progress so far and get feedback from the teachers.
- **Final submission** (deadline: **consult Inforestudante**): For the final submission you should have completed all the remaining steps of the progress report and must hand in the following items (see details in Project section): i) progress report and ii) the code.
- **Project presentation and defences** (during the last T and PL class): There will be a final demonstration of your projects. Begin with a formal presentation of your project, followed by a Q&A session. The defences are mandatory. The projects will be evaluated according to predefined grading criteria (see respective Section). The defences require registration for a time slot, which will be opened in the *Inforestudante* closer to the defence date.

# Evaluation

The grading for Project 2 is equal to 35% of the final grade, which is equivalent to 7 out of 20.

**Important notice!** It is not possible to improve the grade of the project.
**Another important notice!** The submissions made after the deadline will not be accepted.

## Grading criteria

The following elements will be evaluated:
- Data:
    - Data preparation (Does your final dataset include only the data items needed for visualization? Was the data stored in the appropriate format and structure? Was the data cleaned and filtered according to the good practices of data preparation?).
- Solution:
    - Application of visualization principles and practices of good design (Was it designed following the principles studied in the course?);
    - Effectiveness and appropriateness of the solution (Is your visualization effective in answering the defined analytical questions? Is your visualization effective in

communicating data/insights? Does your solution employ appropriate visualization methods and encodings?).
- Implementation: Quality of the implementation (Is your visualization robust? Is your application optimised to run in real-time? Does your code follow conventions and overall good programming practices?)
- Presentation and defence: Clear, engaging, and effective presentation. Did you answer all the questions during the defence? The score for this component is individual.
- Progress report: Are your project and the progress well documented in the report?

## Minimum grade

To get approval in the Data Visualization course, it is necessary to obtain a classification equal to at least 40% of the final grade of the project.

# Project

The project consists of creating a visualization for data analysis and/or communication. The students are faced with real-world datasets with the goal of identifying analytical and/or communication problems and developing a visualization that solves these problems. The visualization should be effective, should contain multiple visualizations, and should incorporate appropriate visualization techniques for data analysis and/or communication, depending on the path each student chooses. The visualization should follow good practices of visual design, as well as the visualization principles and theories studied in this course.

## Progress report (the article)

Keeping track of your progress is an important part of your project. Remember, it is equally important how you get your final results! Therefore, each project should be documented in the progress report, providing details about each part of your project, including any additional information that you find relevant (e.g., alternative designs you tried, the problems that you faced, or the insights that you learned). It is also important to underline that your report book is not a logbook or diary. You should present your work in an informative and concise way, without unneeded details that do not contribute to the understanding of your project. Instead, include many figures in your progress report (e.g., different designs you tried, photos of your sketches, or sketches you made using different visualization tools, etc.). Ultimately, the goal is to convey the progress and rationale of your project in a clear and informative way.

The progress report should take the form of an **article** with a typical structure, which should include the following topics:
- **Authors**: list the members of the team, including their email addresses.

**Project**

- **Introduction**: Provide an overview of your project, its goals, and motivation. Try to define the problem being addressed, the target users, and what they would want to know about the data.
- **Related work**: Include and briefly describe any visualization that you used as a reference or inspiration for your project (e.g., an article, a visualization you found on the internet, examples you saw in class, etc).
- **Data**: Describe the chosen data, its characteristics, the data type, the structure, and data sources, among others.
- **Analytical questions**: Provide a description of analítical questions or insights to communicate for your project.
- **Design requirements**: Provide a list of design requirements and the tasks for your visualization.
- **Design**: Describe the rationale behind your designs, supported by the mockups/sketches of tested solutions. It is important to use the theory that you have learned in the course to justify your choices, i.e., what design principles, perceptual theories, or colour theories informed your decisions.
- **Implementation**: Provide the key functionalities of your application. Briefly describe the algorithms for visualization construction (if any). You can resort to diagrams to explain the architectural aspects of your applications.
- **Validation and Reflection**: This section should include a critical reflection about your visualization. Provide an overall assessment of your final visualization. Was your visualization effective in getting insights about your data? Was it effective in answering the defined questions? What could you learn about your data that wouldn't be possible without it? Try to identify what works well in your visualization and what does not, and how would you improve it.

The progress report should not exceed 8 pages in length (plus references). We suggest including 3-4 pages of text and 5-4 pages of figures.

## Data

The students are free to use any data and datasource. As an example, you can use the dataset from the first project: google drive repository. The students are allowed to use any additional datasources to enrich the datasets at hand, for instance, geographic information about places http://centraldedados.pt/codigos_postais/. Try to be creative and employ all the skills you have earned in the other courses.

**Important note: the students are not allowed to use the same dataset from Project 1**

Regarding the data preparation, the students are free to choose any API, programming language, any visualization tool, or framework. The teachers may restrain to support only the tools they use. Consider discussing your choice before starting the project.

To submit big datasets **do not use** *Inforestudante*, instead use external file-sharing services (e.g. Dropbox, Wetransfer). Include the link in your report. Recall that Inforestudante has a 500 Mb upload limit.

## Code

You should use Plotly (or similar, e.g., Seaborn, Matplotlib) and Python to implement the core functionalities of the project. The students are free to use any additional APIs and frameworks as supportive tools (e.g., to compute graph layouts), but the teachers will only answer the questions regarding Plotly and Python.

All the code should be submitted to the *Inforestudante* platform.

## Suggestion for Project activities

As a suggestion, this section lists several core activities that the students may perform to accomplish the project. This includes the following:
- Definition of target users and a problem that the visualization should solve;
- Developing analytical questions for data analysis and/or communication;
- Definition of data and tasks, as well as the specification of the requirements;
- Acquisition of the necessary data and its preparation;
- Identification of an appropriate visualization method for each of the views of the visualization application (e.g., a time-series representation, a map-based visualization);
- Implementation of the chosen views;
- Refinement and definition of visual encodings;
- Implementation of the data transformation functions;
- Refinement of the visualizations;
- Addition of the supportive graphical elements, such as title, captions, legends, etc.
- Implementation of the basic interaction, such as element selection/hovering, range selection, etc.

# Plagiarism and fraud policy

According to the Article 28.º of the disciplinar reglament for the UC students (https://www.uc.pt/regulamentos/ga/vigentes/Regulamento_Disciplinar_Estudantes_UC)

 *"2 — A fraude cometida em sede de avaliação de uma unidade curricular implica a anulação imediata dessa avaliação e leva à reprovação liminar do estudante na inscrição na unidade curricular em causa, devendo ser registada na plataforma informática da UC e averbada no processo individual do estudante."*