1 2 9 0 FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE D COIMBRA

dei25 departamento de engenharia informática 1995 – 2020

Project

Advanced Data Visualization MECD 2020/21

Introduction

The main goal of the Information visualization project is the development and application of a set of skills that allow the construction of interactive visualizations capable of solving analytical problems and those related to communication of information. The work is organized 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 and possible solutions, sketching and prototyping, implementation of interactive visualization, and validation. It is also intended that the students demonstrate skills in analysis and synthesis, organization and planning, critical reasoning, application of knowledge in practice, investigation, autonomous learning and group work.

The project must integrate the following characteristics and the requirements:

1. All the code should be produced by the students;
2. The only frameworks allowed are **Plotly, Dash, Pandas, Numpy** (the use of additional frameworks should be approved by the teachers);
3. The project should be based on the provided datasets, which are equal for all the students (the use of additional datasets should be approved by the teachers);
4. The visualization should be an analytical application with **multiple coordinated views**;
5. The visualization should implement, at least, **three distinct views**: a time-series representation, a map, and one identified by the students (e.g., node-link representation, scatter plot, among many others covered during the semester);
6. Make use of the **linking and brushing** interaction technique;
7. Include other **interactive elements**, such as hover, range selection, among many others covered during the semester;
8. Contain the necessary additional elements, such as **legends, title**, etc.

Groups

The project should be developed individually.

Milestones

Conditions for the project**:**

- The project is divided in 3 milestones:

    - Milestone 1: planning, working with data - phase 1 (analysing, filtering, parsing), and prototyping (basic representation)
    - Milestone 2: working with data - phase 2 (datamining) and multiple views;
    - Milestone 3: Interactivity and refinement.
- The submissions are made through the Inforestudante platform;
- The submissions made after the deadline will not be accepted;
- The deadlines for submissions should be consulted on the Inforestudante;
- The defenses are mandatory.

Submission deadline

- Milestone 1: **26 of March**
- Milestone 2: **30 of April**
- Milestone Final: **21 of May**

The project defenses take place the next week after the submissions.

Evaluation

The grading for the project is equal to **60%** of the final grade, which is equivalent to **12** out of **20**.

Important notice! It is not possible to improve the grade of the project at any phase of the exam epoch.

Milestones

Grading breakdown:

- Milestone 1: **15%**
- Milestone 2: **50%**
- Milestone 3: **35%**

Grading criteria

The following elements will be evaluated:

- Identification of the problem and the appropriateness of the solution (e.g., analytical questions that answer the problem, appropriate visualization methods, appropriate visual encodings, among others);
- Work with data (e.g., cleaning, filtering, parsing);
- Interactivity (e.g., zoom/pan,direct selection of elements, advanced interaction methods, GUI, etc.);
- Application of principles and good practices of visual design;
- Implementation;
- Written report (the report is submitted for each milestone).

Minimum grade

In order to get approval in the Advanced 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 should solve these problems. The visualization should be effective, should contain multiple coordinated views, and should provide appropriate visualization tools for data analysis and/or communication, depending on the path each student chooses. The visualization should follow good practices of visual design

Data

You are provided with the data about Airbnb listings around the world. The data comes from an independent project that collects publicly available **Airbnb** data (http://insideairbnb.com/index.html). The city of attention will be the city of Porto, Portugal. The respective datasets are available at http://insideairbnb.com/get-the-data.html, through a manual search of the term “Porto, Norte, Portugal”. Additionally, the students are allowed to use the zip codes provided in this link http://centraldedados.pt/codigos_postais/

Milestone 1 - Planning, data works (phase 1), and prototyping

The first part of the project consists of analysing the provided data and working on the proposal for the project. This includes the following activities:

- Definition of target users and a problem that the visualization should solve – conceptualization;
- Developing of analytical questions for a data analysis and/or communication,
- Definition of data and task abstractions, as well as specification of the requirements;
- Acquisition of the necessary data, its parsing and filtering;
- Identification of an appropriate solution for each of the views of the visualization application – a time-series representation, a map based visualization, and one chosen by students;
- Implementation of one of the chosen views in Plotly (no Dash required).

The students should deliver a written report (1-2 pages) that addresses the above topics, including the relevant details of the implementation that each student finds important. The report and the code should be compressed into a single zip file and submitted to the Inforestudante platform.

Milestone 2 - Data works (phase 2), design and implementation

In the second part of the project the students should develop a working prototype of the visualization, which integrates all the non-interactive views based on basic visualization techniques. This includes the following steps:

- Refinement and definition of visual encodings;
- Implementation of the data transformation functions;
- Implementation of all of the views using Plotly + Dash;
- No interaction needed.

The students should deliver a written report (1 page), which should be a continuation of the report from the previous stage. The report should include a brief documentation of the relevant backend functionalities and the description of the frontend visualizations and their visual encodings. The report should also include print screens of the visualization (+1 page). The report and the code should be compressed into a single zip file and submitted to the Inforestudante platform.

Milestone 3 - Interaction, refinement, and validation

In the final stage of the project, the students should refine their visualizations and add interactivity. This includes the following tasks:

- Refinement of the basic visualizations employed at the previous stage;
- Inclusion of the additional embellishments, such as title, captions, color legends, etc.
- Implementation of the basic interaction, such as element selection/hovering, range selection, etc.;
- Implementation of the linking and brushing interaction method;
- Validation of the employed models through a written critical discussion.

The student should deliver a written report (1-2 pages), which should be a continuation of the previous report. The report should include a description of all the parts that were improved in regard to the milestone 2, and a brief description of any additional elements that were included. In regards to the interaction, the report should include enough details, technical and functional aspects, to provide a clear picture of the visualization capabilities. Additionally, the visualization methods employed in the project should be validated through a written critical discussion (1 paragraph), using the tools, conceptual and theoretical, acquired during the semester. The report and the code should be compressed into a single zip file and submitted to the Inforestudante platform.

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.”