

Information Visualisation

2024/2025 – 1st Semester

Practical Assignment

Application for the visual exploration of a data set

Introduction

This work consists in the conceptualization and implementation of a prototype of a simple application for visual exploration. The visualisation must include at least two distinct views that clearly illustrate characteristics of the data and allow the user to visually explore the data, helping him understand the phenomenon and respond to a set of relevant questions.

Each group must choose/propose a work theme that allows the visualisation of data/information using the d3.js tool. The use of data sets resulting from work developed by students in other curricular units is encouraged.

The use of uncommon representations and the availability of interaction features will be valued in the evaluation.

A website must be developed to present the generated visualisations in each context — for example, making an introduction, describing the data, presenting the visualisations and conclusions obtained, etc. A human-centred approach should be used, with several iterations of design, implementation, and evaluation (at least two).

The first iteration must include:

- Analysis of the data and the phenomenon they represent, of users, context of use and main questions that the application should help to answer.
- Justified choice of the visualisation techniques selected (for representation, presentation, and interaction)
- Development of a low-fidelity prototype illustrative of the application's conceptual model (on paper or using rapid prototyping S/W, such as Balsamiq)
- Usability testing of the prototype with users (can be done in class with classmates - Do not forget to define the most important tasks that should be tested with users.).

The second iteration must correspond to the integration of the results obtained through the evaluation with development of the prototype in d3.js, The prototype must be tested through a Heuristic Evaluation and again with users as representative as possible of the target users.

Groups may propose the use of other libraries or Visualization S/W (such as Power BI), however this possibility should be discussed and justified with the lab teachers.

The work must be carried out in a group of 2 elements.

Evaluation of the first iteration

In classes on 28 and 30/10. Each group should bring a simple presentation (< 10 min) and the conceptual model they propose to discuss with Prof. in class, low-fidelity prototypes must be tested with the help of colleagues.

Final presentation

Each presentation, whose duration should not exceed 15 minutes (10+5 min questions), will be performed by the group members in the class on November 25th and 27th. Presenters must use careful language, speak clearly and not exceed time.

Each group should produce a sober presentation and a short video of their work as well as a report describing it. The report must be succinct and written in the format provided. Do not forget in the report to clearly explain what data was used and its source, indicate the target users of the visualisations, in what context they will use them, what are their objectives in terms of visualisation/data analysis, including a set of questions that must be answered using visualisations.

Delivery

Each group must deliver through e-learning the report, presentation slides, video demo and code developed by the delivery date (27/11/2023).

Use the following rule for file names: TPnnmec1+nmec2, where n is the class number and the nmec1 and nmec2 are the student numbers of the two elements of the group. For example: a group from the TP1 class that has developed the AppXYZ application and is made up of students with nmecs 44444 and 55555 must submit a file with the name: TP1-AppXYZ-44444+55555

The final presentation must also include the indication of the **percentage of work of each student** to the assignment.

Warning

The developed visualisation must be original. In addition to the files distributed in the practical classes, some code or existing libraries may be used (occasionally), but this fact must be clearly referenced in the code/presentation and report.

Works that use a large volume of non-original code, without explicitly referencing it, will be void.

The tools created must be available for use, either through a code repository (github, gitlab) or zip file, indicating any non-original code or library used. Instructions for downloading or using dependencies, APIs or commands necessary for its operation must be explicit in the README file. If the tool has more than one page, also indicate which page(s) generate the visualisations planned in the design phase.

Themes

Students are encouraged to choose and propose topics of their own interest (datasets that result from work developed in other curricular units, etc...), however, here are some proposals/links where data can be found to use in the assignment:

	<p>Visualisation of data related to the origin of students from DETI courses and similar.</p> <p>See the origin choices of students who entered DETI courses: Compare them with similar courses in the country. In addition to information regarding grades, see also which options the students choose (where they entered and which options followed). An assignment can be made for each of the courses.</p> <p>https://www.dges.gov.pt/coloc/2024</p> <p>https://www.dges.gov.pt/pt/pagina/informacao-geral-publico</p>	
	<p>Históricos de notas de milhares de alunos da UA disponibilizados de forma anónima.</p> <p>Colaboração: Diogo Gomes</p>	
	<p>Visualização de dados de movimentação e consumos de uma casa particular (Diogo Gomes)</p> <p>Colaboração: Diogo Gomes</p>	
	<p>Home comfort dataset: dados ambientes recolhidos em diversas casas de Portugal com alguma informação de preferência dos habitantes (se estava temperatura correta ou não, etc...)</p> <p>Colaboração: Mário Antunes</p>	
	<p>Apollo pothole dataset: dataset com diversos trajetos de carros, com velocidade, posição e acelerómetro.</p> <p>Colaboração: Mário Antunes</p>	
	<p>Visualisation of data available on the kaggle site about machine learning: MDB, Football, road accidents, detection of credit card fraud, historical information about bitcoin, related to fires, various diseases, etc...</p> <p>http://archive.ics.uci.edu/ml/</p>	
	<p>Visualisation of data from questionnaires prepared by the Visualization Society:</p> <p>https://github.com/data-visualization-society/data_visualization_survey/commit/70415aef02ca913d2f781a152ddc91ef9837252b</p>	
	<p>Data available from the National Institute of Statistics, Portugal (INE):</p> <p>Various themes are possible regarding agriculture, environmental, justice, population, prices, etc...</p> <p>http://www.ine.pt</p> <p>A theme, for example, can use the samples from censuses from 2001 to 2011 and make visualisations about different aspects of the evolution of the population in Portugal.</p> <p>https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_pufs&menuBOUI=13707095&contexto=up&selTab=tab5</p>	

	Visualisation of data available in PORDATA: Data on population, education, employment, Science and Technology, etc... http://www.pordata.pt/	
	*** Students can propose other assignments in line with their interests	

Autor: [Paulo Dias](#)

Data: 07 October 2024