Department of Computer Science and Media Technology

Linnæus University

Autumn 2021 – SP2

Course Assignments for Advanced Information Visualization and Applications 4DV806 – Autumn 2021

Assignment 1

Deadline for this assignment is Nov 22, 2021 at 23:59.

Task 1 *Visualize a Data Set*

The first task is to develop a program for visualizing a data set. *Note that you are not allowed to modify the format or the contents of the provided data set file with third-party tools* (you are allowed to process the data within your implementation, though)! You can download the data set from Moodle or via the following URL:

`https://cs.lnu.se/isovis/courses/fall21/4dv806/assignments/publications-stats.json`

(1) Take a look at the data, find out what the data is about, think of a good way to visualize the complete data set to present the overview of the aggregated data as well as details about particular data items (e.g., a treemap), and implement it. (2) In addition you should discuss what interaction possibilities would make sense and why, based on the available attributes in the data set.
You can use any programming language and library of your choice to visualize the data sets (JavaScript and D3.js for instance, https://d3js.org). Please note that this task requires programming rather than usage of existing visualization tools or environments such as Tableau.

Task 2 *Review a Visualization Tool*

The second task is to try out a visualization tool titled Embedding Projector and write a short review. You should start by reading an article at `https://distill.pub/2016/misread-tsne/` that discusses a well-known dimensionality reduction technique titled t-SNE, and then look at the documentation at `https://tensorflow.org/tensorboard/tensorboard_projector_plugin#analysis`.

The Embedding Projector tool is available at `https://projector.tensorflow.org/`

You should try to use the Word2Vec 10K data set with the several dimensionality reduction techniques supported by this tool and investigate the available visual representations and interactions. Include your review in your report. Try to think about the pros and cons of this tool as well as particular design choices that you would have potentially changed, for instance (please motivate your statements!).

Prepare a short report (2–4 pages, A4, 12pt) and a presentation (5–10 minutes). You will have to present your solution after the deadline.
Please prepare a ZIP archive with your report (PDF) and implementation and upload it to Moodle by the given deadline! If you have questions, you can contact Angelos Chatzimparmpas via email (angelos.chatzimparmpas@lnu.se) or use the Moodle discussion forum. You will have to present your work on Nov 23, 15:15–17:00, online via Zoom.
Please note: any kind of plagiarism is not acceptable!