Assignment 3: Human Factors (10 Points)

General instructions.

1. Restate your chosen SDG in your submission and complete all tasks by visualizing phenomena that speak to your chosen SDG.
2. You can reuse the dataset from the previous assignment or use one or more other datasets, provided that they are all connected to your SDG and contain the information needed to perform the assignment tasks.
3. Always specify which datasets you used for which tasks and provide links to the sources of all datasets.
4. Always specify which visualization tools (e.g., programming languages and libraries) you used for which tasks.
5. You are allowed to use generative AI *to support your learning*. If you opt to use generative AI, you *must* state which model(s) you used for which tasks, describe how and why you used them, and provide a critical reflection on how they supported you in understanding the course material and carrying out your assignment tasks. Undisclosed usage of generative AI, if discovered, will be considered cheating.

If you have questions or feedback on the assignments, please share these in the course forum.
See also our MyCourses page for further general instructions.

Specific instructions for Assignment 3 To complete each of the following tasks, please

1. create **two** visualizations as described in the task,
2. describe and interpret the visualizations with a view to understanding the progress toward your SDG,
3. motivate your visualization-design choices and describe the thinking that went into your design, and
4. report any challenges you may have encountered and how you overcame them.

Your answers to 2.–4. should be at least one paragraph of text each. The target length of your text (excluding graphics) is between two-thirds of a page and one page, but you can remain below that if your answers are concise or go over that if you have more to share (no need to optimize the layout to make your answers look longer or shorter).

Task 1: Maps and Colormaps (5 Points)

Select a dataset pertaining to your SDG that contains observations for distinct geographical units (e.g., countries around the world, countries in the EU, or Finnish regions). Create two choropleth maps with the same underlying data, one using a sequential colormap and one using a diverging color map. If necessary, ensure that the data is suitable for mapping to a sequential resp. diverging color palette by employing suitable data transformations (recall the material on re-expression). Use the colormap principles you learned in the course to inform your colormap choice. Elaborate on the perceptual differences between the two choropleth maps in your discussion (specific instructions, subtask 3).

Task 2: Heatmaps and Clustermaps (5 Points)

Select a dataset pertaining to your SDG that contains observations for ≥ 10 continuous variables, measured at ≥ 10 distinct points in time (e.g., yearly over 10 years, monthly over 12 months, or hourly over 24 hours, depending on the time scale of interest).

First, create a heatmap of your data, with alphabetically ordered variables as rows and sequentially ordered time stamps as columns. Use the colormap principles you learned in the course to inform your colormap choice.

Second, familiarize yourself with the idea of a hierarchically clustered heatmap (see seaborn documentation) and the choices that need to be made in creating such a so-called *clustermap*, and then create a clustermap of your data that highlights the similarity between variables but maintains the original temporal structure. Use the colormap principles you learned in the course to inform your colormap choice. Elaborate on your parameter choices for the clustermap as well as on the perceptual differences between the heatmap and the clustermap in your discussion (specific instructions, subtask 3).