Information Visualization
4DT909 – 2025

Final Project

Deadline for this assignment is October 20, 2025 at 23:55.

Activity 1 *Find datasets*
Many research and learning platforms provide access to real-world datasets that can be used for analysis and visualization tasks.

EU Referendum (2016) `http://www.bbc.com/news/uk-politics-36616028`

US Presidential Election (2016) `http://elections.huffingtonpost.com/2016/results/president`

Free and Occupied Beds for Covid-19 Patients in Germany (2020) `https://coronavis.dbvis.de/en/overview/map/bedcapacities`

The datasets listed above and those shown in class are not allowed for this assignment. You must find your own dataset from a reliable, real-world source that is relevant and suitable for academic analysis. For example, you could look for datasets that allow you to explore topics such as social, economic, environmental, or entertainment trends. However, do not use or modify any of the datasets mentioned earlier.

Important requirements:

- The datasets must be **real** (it is forbidden to use generated or partially synthetic data!).
- The datasets must come from a **reliable and reputable source** (e.g accademic peer reviewed research paper). Provide the proper **reference** for the datasets used.
- They should have a **relevant context** for academic use.
- After merging and eventual cleaning, the final dataset must contain at least **100 relevant rows** and **10 relevant columns**.
- The dataset must include different variable types: **continuous, discrete, nominal, and/or ordinal**.

Note: In your presentation, you must clearly demonstrate how the merged dataset satisfies all of these requirements. If any of the requirements are not met, you will fail on the assignment.

Activity 2 *Task Definition*
In this activity, you must clearly define the **analytical questions (tasks)** that your visualization aims to answer. These tasks guide the purpose of your design, determine which variables to highlight, and establish how interactivity will support insight generation.
You should define at least **five tasks** that your visualization will address. Below are example tasks that can be adapted to your chosen dataset.

- *Does average income increase with education level, or do CO_2 emissions vary by region and year?*
- *Is there a clear correlation between GDP per capita and life expectancy?*
- *Which countries or states perform best in renewable energy adoption?*

- *Are house prices normally distributed, or do we observe distinct market segments?*
- *How does the relationship between variables differ when filtering for urban vs. rural regions?*

Note: The final result should make it clear how the visualization enables users to answer these questions effectively.

Activity 3 *Visualization that Answers the Tasks*
In this activity, you will implement at least **five visualizations** that adequately answers the analytical questions defined in the previous activity. Your goal is to ensure that each visualization provides clear, interpretable, and evidence-based answers to the proposed questions. It's recommended implement visualizations and interactions additional to those taught in the classroom.

Objectives:

- Create one or more visualizations that directly respond to the five analytical tasks defined earlier.
- Choose appropriate visualization types that best fit the nature of each question and data type.
- Clearly explain in your presentation how each visualization supports the reasoning behind the answers.

Activity 4 *Core interaction techniques*
In this task you should implement **core interactions techniques** within D3.js. You must choose **at least three** of the following approaches:

- Lasso Selection.
- Marquee Selection.
- Axis-drag based Selection.
- Off-centered Pointer.

Start with a simple scatterplot (for example, Median Income vs Median House Value). Implement one of the techniques above to select subsets of data. The selected subset should be highlighted and used to update another linked visualization (for instance, a histogram of house values).

Activity 5 *Multiple Coordinated Views (MCV)*
In this task you will build a **dashboard with interconnected visualizations** within D3.js. You must choose at least **four** of the following coordination strategies:

- Juxtaposition.
- Integrated.
- Superimpose.
- Brushing & Linking.
- Dynamic Queries.

Start simple: create two static views (for example, a scatterplot and a histogram). Then add interaction only when it improves interpretability. For instance, brushing points in the scatterplot could filter the histogram, or a dropdown menu could update all views simultaneously.

Activity 6 *Enhance Visualization Quality*
In this task, you will focus on refining your visualizations to make them clear, objective, and effective for communicating insights.

Objectives:

- Apply the design guidelines discussed in class to enhance readability, clarity, and usability of your visualizations.
- Ensure that each visualization conveys its intended message quickly and accurately.
- Optimize visual encoding (colors, shapes, sizes, labels, and axes) to avoid ambiguity or misleading impressions.
- Maintain consistency across visualizations to allow easy comparison and interpretation.
- Make the dashboard visually appealing while keeping it professional and focused on the data.

Note: Pay attention to all aspects of good visualization practice, including layout, interactivity, labeling, and color schemes. The goal is to produce a dashboard that is both informative and visually coherent.

Activity 7 *Case Study*
In this activity, you will define your visualization from a **user-centered perspective**. This involves identifying who the user is, what goals they pursue, and how the visualization supports their decision-making or understanding through interaction and design.

Expected Content:

- **User Profile:** Describe the intended user (e.g., policy analyst, environmental researcher, data journalist, educator, or business manager). Include relevant characteristics such as domain expertise, analytical needs, and data literacy level.
- **User Story:** Write a concise narrative that expresses the user's goal and motivation, for example: *As a [type of user], I want to [perform an action] so that I can [achieve a goal].* This statement should clearly express why the user values the visualization.
- **User Scenario:** Describe a realistic scenario in which the user interacts with the visualization to achieve their goals. Illustrate how the user explores the data, interacts with the interface, and gains insights or makes decisions based on the visualization. Focus on the narrative of the user's experience rather than a strict step-by-step process.

Example:

- **User Profile:** *Public health officer with moderate data literacy.*
- **User Story:** *As a public health officer, I want to explore infection rates by region and time so that I can identify areas needing additional resources.*
- **User Scenario:** *A public health officer accesses an interactive dashboard displaying infection rates and hospital capacity across regions. They notice a spike in infections in certain areas and use brushing to focus on a specific time period. By filtering the data by age group, the officer identifies the most affected populations and regions, enabling them to prioritize the allocation of medical resources efficiently. Throughout this process, the dashboard allows the officer to explore patterns and derive actionable insights quickly.*

Activity 8 *Video Presentation*
In this task, you will create a **recorded video presentation** of your project, lasting between 10 and 12 minutes. The video should clearly demonstrate both the final visualization system and the logic behind its implementation in D3.js.

Structure of the video:

- The first **2 or 3 minutes** must present the application itself. Show how the visualization works, highlight the main interactions, and briefly demonstrate how it answers the analytical questions defined earlier.

- The remaining **8 to 9 minutes** must focus on the **code explanation**. This part should emphasize the logic of the implementation in `D3.js`: how the data is processed, how elements are bound to the DOM, how interactivity is implemented, and how multiple views are coordinated. Avoid reading or scrolling through the code line by line; instead, explain the reasoning and structure behind key parts of the code.

Note: If working in pairs, each person must talk for approximately the same amount of time.

Activity 9 *In-Person Presentation*
In this task, you will prepare and deliver an **in-person presentation** of your final project, lasting approximately **7 minutes**. The presentation should clearly demonstrate your visualization, highlight tasks, and explain your design and interaction choices. The presentation must be prepared in PDF or PowerPoint format.

Note: You will have to present your work on October 21 or 22, 13:15-15:00, in person. Both the presenters and the order of presentations will be determined randomly in class on the scheduled days.

You must do it with the same partner from the seminar, if it was done in pairs. Check if this PDF really belongs to the course you are enrolled in.

Note: If you do not submit your assignment on time and/or do not present in person, you will fail the assignment.
Students must submit a single ZIP file containing all D3.js code, the slides presentation, and the recorded video. The evaluation will be based on the quality of the analyses, the clarity and effectiveness of the presentations, the correctness and logic of the code, and adherence to the stipulated time limits. Any error or omission in any of the required activities will be penalized.

If you have questions you can contact Benjamin Powley via email (benjamin.powley@lnu.se).

Upload it to Moodle by the given deadline!

Any kind of plagiarism is not acceptable!