Hey ChatGPT, can you visualize my data?

A Multi-Dimensional Study on Using an LLM for Constructing Data Visualizations

We explored the effectiveness of an LLM in creating data visualizations across a spectrum of scenarios, characterized by three key dimensions: the complexity of the underlying data, the user's data visualization competencies, and the requirements of the resulting visualization. Based on an empirical study, we offer insights into the potential role of LLMs as tools for empowering users with varied expertise to effectively visualize and interpret data.

hochschule mannheim

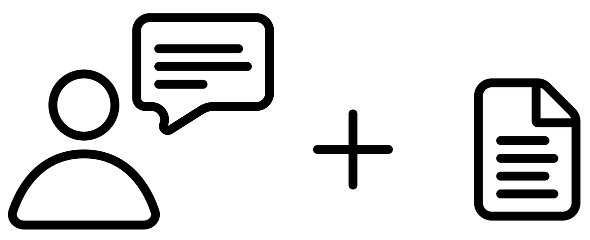
Mara Ströbel, Kai Eckert, Till Nagel
Hochschule Mannheim

human data interaction lab

Acknowledgement
Poster design by Sophie Humbert.

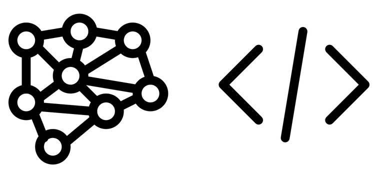
methodology

01 Provide prompt and data set



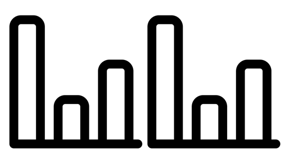
Each scenario consisted of a prompt and a pre-selected dataset. All prompts were simulated.

02 Generate code



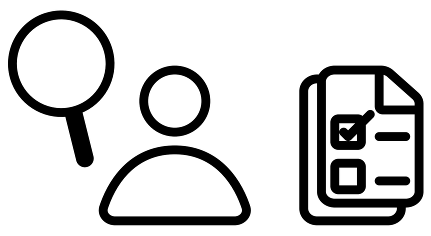
Next, ChatGPT 4 was asked to generate Python code based on the prompt and the dataset.

03 Render visualization



The code was used to render the visualization in a Python environment and exported as static image.

04 Evaluate



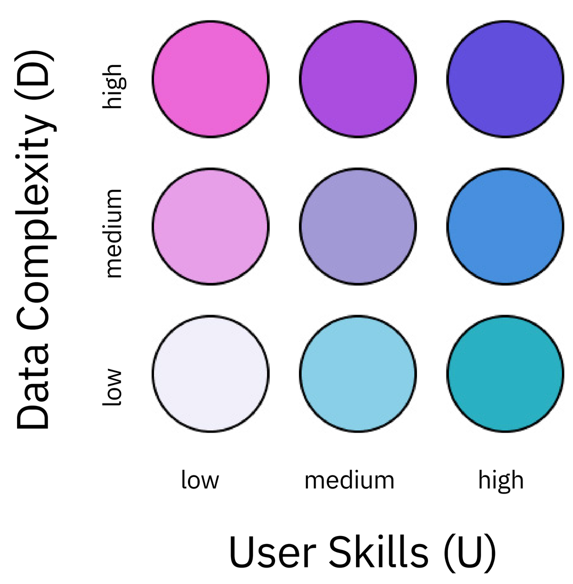
An expert rated each visualization based on the ICE-T heuristics and the Data Visualization Checklist.

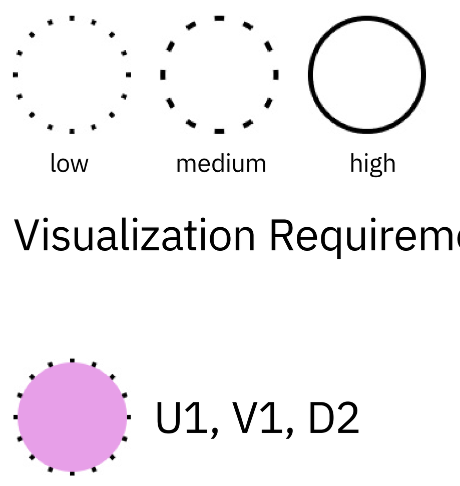
visualization construction factors

We structured the experiments around three relevant factors: dataset complexity (D), user's experience with datavis (U), and the visualization requirements (V). Each of the factors is categorized into three levels (high to low), leading to 3×3×3 = 27 combinations. In our study, we systematically explore how these factors influence the effectiveness of LLM-generated visualizations.

In an experiment, a novice user (U1) might ask for a basic visualization (V1) of a dataset with medium complexity (D2).

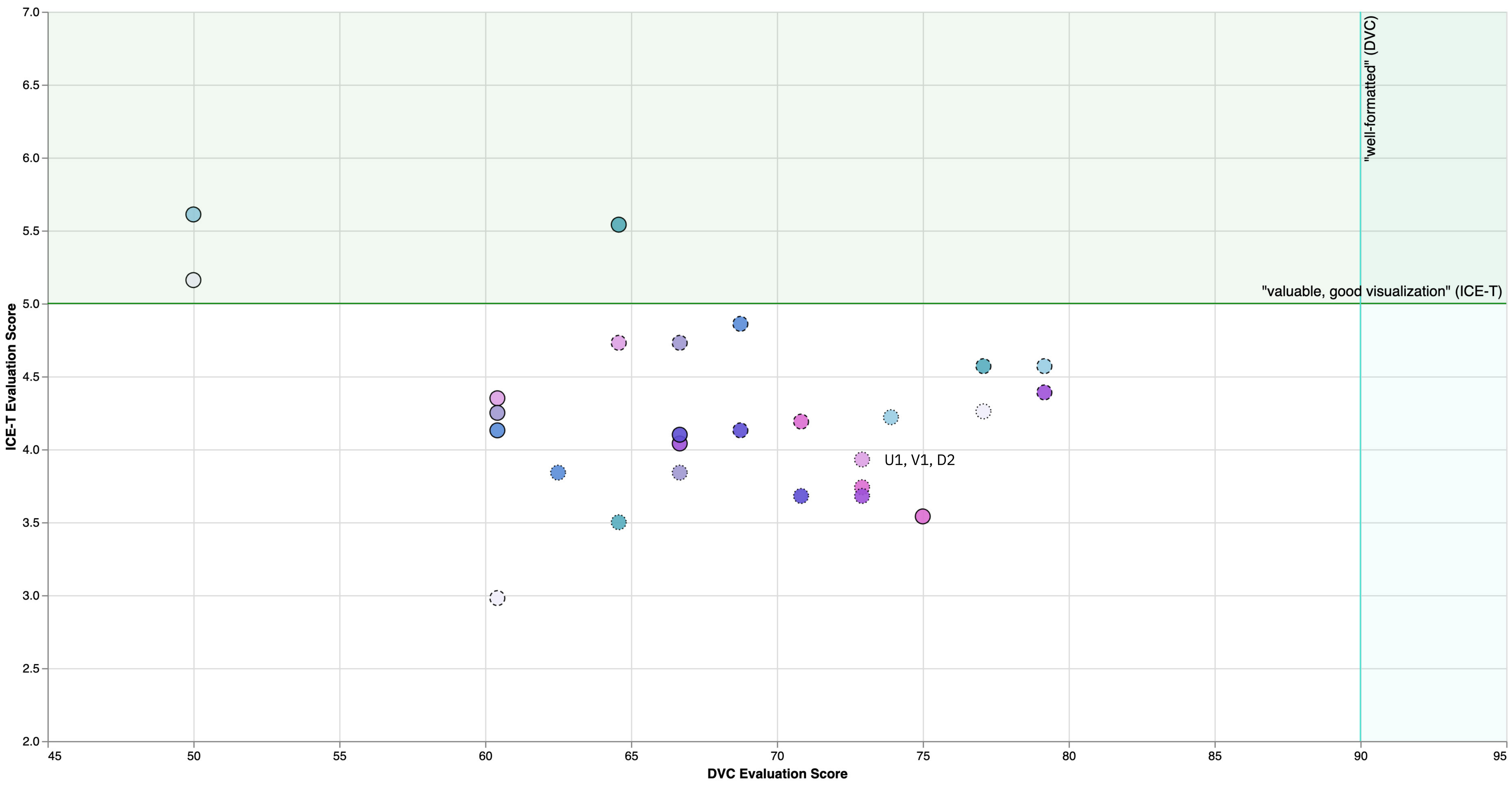
Prompts for user group U1 were kept broad (e.g. *"Create a simple, meaningful data visualization"*) to explore visualization generation for users with no prior skills, without specifying data tasks or visualization properties.





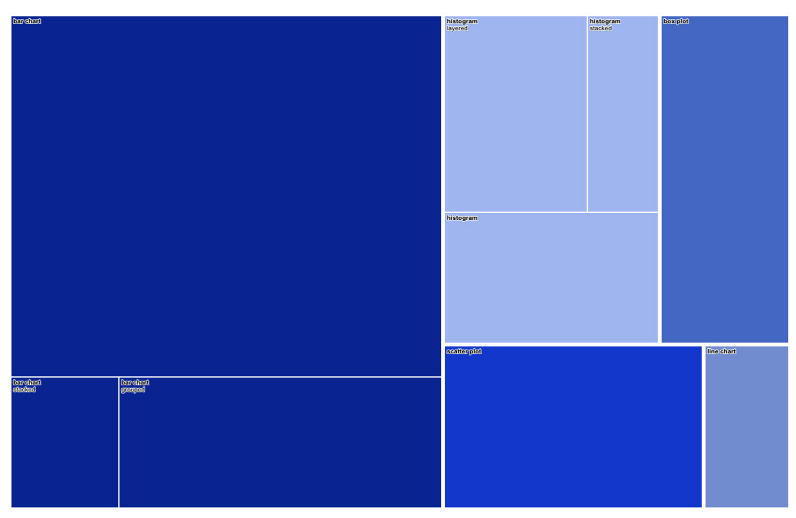
Experiment 112: A novice user (U1) asks for a basic visualization (V1) of a dataset with medium complexity (D2).

results



Scatterplot of evaluation results for the 27 LLM-generated visualizations according to ICE-T and Data Visualization Checklist scores. Each circle denotes an individual visualization, with hue representing dataset complexity levels and user competencies (see bivariate color schema above). The dashed strokes encode visualization requirement levels. On the top and right edges, the lines represent the thresholds for valuable and well-formed visualizations.

Boxplots showing distribution of ICE-T scores for each of the 23 generated visualizations, grouped by dataset complexity.



Treemap showing the frequency of generated visualization techniques, aggregating all variations (e.g. grouped and stacked bar charts).

conclusion

Our study demonstrates that LLMs for generating data visualizations present a promising tool suitable for users across all levels of competencies, including those without prior experience in data visualization.
Throughout our experiments,

ChatGPT-generated visualizations reliably achieved a basic level of quality both in supporting data analysis as well a design. However, they often lack the analytics excellence and refined design formatting necessary for final, publication-ready visualizations.

Given these findings, we advocate for the use of LLM-generated visualizations primarily in the exploratory data analysis (EDA) phase, where the ability to rapidly generate and iterate on visualizations can enhance productivity and insight discovery.

outlook

Our study highlights the role of LLMs as a valuable starting point in the data visualization process, pointing towards a future where further advancements could broaden their applicability and effectiveness.