
d3.js and its potential in data visualization

Creating a diagram showcase using ukrainian refugee data

Luis Rothenhäusler

20202459



Bachelorarbeit - Exposé

Fachbereich Informatik
und Medien
Technische Hochschule Brandenburg

Betreuer: Prof. Julia Schnitzer
2. Betreuer: Hier könnte Ihr Name stehen

Brandenburg, den 24.05.2022
Bearbeitungszeit: dd.mm.yyyy - dd.mm.yyyy

Table of contents

1	Introduction	2
2	Status Quo	2
3	Goals	2
4	Methodology	2
5	Data	3
6	Diagrams	3
7	Preliminary Structure	3

1 Introduction

The postmodern world produces huge amounts of data every second. Analyzing this data leads to better-informed decision-making in every sector. Yet the vast amounts of created data is often hard to comprehend with the human mind. Data visualization is about finding ways to represent this data in visually appealing, yet easily understandable visual representations. Doing this quickly and always up to date can be crucial. There are many tools available to help with the creation of infographics. Some of these tools have a graphical-user-interface, are programming based. This thesis will be a deep dive into the possibilities of one of these tools, the 'd3.js'(D3) library for JavaScript. To show its capabilities and potential D3 will be used to create a showcase containing several different graphics.

The following sections describe what the goals of this thesis are, as well as why they should be pursued and how they will be achieved.

2 Status Quo

Since the first notation of the development of D3 in 2011, the library has seen continuous development. During this time many different resources and usage examples have been created for D3. Yet a lot of these openly available examples are created in online environments with D3 preloaded, are lacking documentation, use different styles of JavaScript, do not react to changing data or are not easily comparable or adaptable for one's own purposes.

3 Goals

The main result of this thesis will be the creation of a showcase of eight diagrams created from two datasets in D3. The showcase will help new users get an overview of D3's possibilities and complexity. We will also answer the question of: 'When and why is D3 a good choice for data visualization?'

Furthermore there are certain requirements the showcase will fulfill. Each of the approximately eight diagrams should be easily adaptable to one's personal use and use some aspect of D3 which is not used by the other diagrams. They will be comparable to each other, to allow fast comprehension and understanding of the differences in the creation process. Furthermore all diagrams should be able to react to data changes, so they can make full use of D3's possibilities.

4 Methodology

To achieve these goals, there are certain limitations. Besides using a consistent code style and proper documentation, each diagram should be fully independent of the other graphics. This will allow easier adaptation and an easier comparison in effort between the

graphics. To further simplify the comparison, only two datasets are used. Furthermore the diagrams should be implemented using only JavaScript, HTML, CSS and the D3 library, and not rely on a specific framework. This makes them lightweight and easier to adapt them to all kinds of web based projects. To mimic changing data, the showcase has an option to manipulate the used data-sets.

To make sure the full extend of the D3 library is leveraged, as many different sections of the D3's API[1] as possible is used.

5 Data

The two datasets used in this thesis, contain data of all of the four different kinds. Our first dataset contains the refugees in absolute people(discreet) per country(nominal). We can also express this in refugees in percent(continuos) per country. The second dataset is about the total refugee count(discreet) over time(ordinal).

All data comes from the UNHCR, the UN Refugee Agency[2].

6 Diagrams

There are many considerations to make about which way might be the best to represent data. The chosen shapes, colors and diagram types[3] are perceived differently. It can also differ depending on the types of data which are presented[4]. As the showcase is about the possibilities of D3, the best didactic choices are not considered. Instead the diagrams are chosen to cover a wide area of D3's functions. Yet the chosen combinations of data and diagrams should still be reasonable and realistically usable.

7 Preliminary Structure

This is the preliminary structure of the actual bachelors thesis:

1. German Abstract
2. Introduction
3. Basics
 - Infographics
 - D3.js
 - What is it?
 - How does it work?
 - * Modules
 - * General Update Pattern

- Data
 - Types of Data
 - * Categorical
 - * Numeric
 - Datasets
 - * Preprocessing
 - * Data-Types
 - Diagrams
4. Implementation
- Diagrams
 - For each diagram title
 - * How does it work?
 - * What parts of D3 does it use?
 - * Pro and Con considerations
 - Showcase
 - Data Updates
 - Integration of diagrams
5. Conclusion
6. Sources
7. Appendix

References

- [1] M. Bostock, “D3/api.md at main · d3/d3,” Apr 2022. [Online]. Available: <https://github.com/d3/d3/blob/main/API.md>
- [2] UNHCR, “Operational data portal,” accessed:12.05.2022. [Online]. Available: <https://data2.unhcr.org/en/situations/ukraine>
- [3] J. Heer and M. Bostock, “Crowdsourcing graphical perception: using mechanical turk to assess visualization design,” in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2010, pp. 203–212.
- [4] J. Mackinlay, “Automating the design of graphical presentations of relational information,” *Acm Transactions On Graphics (Tog)*, vol. 5, no. 2, pp. 110–141, 1986.