

Showcase of d3.js and its possibilities in
infographics using refugee data of the current
Ukraine conflict

Luis Rothenhäusler (20202459)

Brandenburg, May 30, 2022

Abstract - German

Englische Arbeiten brauchen eine Zusammenfassung auf Deutsch. Mal abgesehen davon, dass wenn die Zusammenfassung interessant ist man ohne English eh nicht weiter kommt...

Contents

| | | |
|----------|---------------------------------------|-----------|
| 1 | Introduction | 3 |
| 2 | Basics | 4 |
| 2.1 | Data | 4 |
| 2.1.1 | Types of Data | 4 |
| 2.1.2 | Datasets | 5 |
| 2.2 | Data Visualization | 5 |
| 2.3 | Diagrams | 5 |
| 2.3.1 | Refugees per country | 5 |
| 2.3.2 | Refugees over time | 6 |
| 2.4 | D3.js | 6 |
| 2.4.1 | What is it? | 6 |
| 2.4.2 | How does it work? | 6 |
| 3 | Implementation | 8 |
| 3.1 | Showcase | 8 |
| 3.1.1 | Integration of each diagram | 8 |
| 3.1.2 | Data Updates | 8 |
| 3.2 | Diagrams | 8 |
| 3.2.1 | Bar Chart | 8 |
| 4 | Conclusion | 10 |
| 5 | Sources | 11 |
| 6 | Appendix | 12 |

1. Introduction

Describe the background of the thesis, why it is important, what do we want to achieve. Why is this thesis? What do we want to do? What is the status quo? What benefits will result from this thesis? Something about the importance of infographics and comprehensible data.

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.

2. Basics

Everything necessary to understand the implementation as well as anything which is done beforehand, will come up here

In the following, all concepts, technologies and required backgrounds for understanding this thesis are explained.

2.1 Data

We'll talk about data a bit. Where does it come from? How is it structured? What kind of attributes? What even are attributes?

Since ancient times, humans have recorded data. Recording the ins and outs of available resources was one of the driving factors behind the conceptualization of writing. (TODO: Check sources of the beer brewing video series) With the introduction of computers the amounts of gathered data have grown drastically. Vast amounts of data are gathered across all aspects of life.

2.1.1 Types of Data

Even though data comes from a huge variety of sources and can express a plethora of things, there are only four different types of data. They are split into equal pairs. Two types of categorical data and two types of numerical data. In the following each of the types of data will be explained.

Categorical

What is categorical data? Nominal and ordinal data

Categorical data is information collected in groups. It is often of descriptive nature. Whilst the values can be represented in numbers, they do not allow for arithmetic operations. There are two types of categorical data. Nominal and ordinal data.

Nominal data is mostly descriptive in nature. They are independent and have no inherit order. Examples are 'Country', 'Color', 'Brand'.

Ordinal data is mostly similar to nominal data. Yet the data does have some sort of internal order. For example different dates each describe a day, but one day also comes after another.

Numeric

What is numeric data? Continuous and ???

2.1.2 Datasets

What are our datasets about? Where do they come from?

Preprocessing

What is done in preprocessing? Python script which removes all excess / maybe do that in JS as well..?

Data Types

Which data types can be found in our data-sets? Where?

2.2 Data Visualization

What is it? Where is it? Why is it important?

2.3 Diagrams

What diagrams exists? Which are the most common? What possibilities do they offer for encoding data? Which considerations for readability? Why do some diagrams not make as much sense? Which considerations where made for fulfilling the showcase requirements?

2.3.1 Refugees per country

Which ones did I choose? Why? Which data attributes do they encode?

Bar chart

What is it? Why is it here?

Pie chart

What is it? Why is it here?

Tree map

What is it? Why is it here?

Sankey

What is it? Why is it here?

2.3.2 Refugees over time

Which ones did I choose? Why? Which data attributes do they encode?

Area graph

What is it? Why is it here?

Circle graph

What is it? Why is it here?

2.4 D3.js

This is all about d3. What is it? Where does it come from? What is it used for? Who uses it? Why should it be used? How does it work? Enter, update and exit pattern. Something about the modular structure of D3 as well. Might be worth mentioning "observables" as well.

2.4.1 What is it?

General definition. What is it, where does it come from, what is it for?

2.4.2 How does it work?

General functioning of D3.

General Update Pattern

What is it? What can it do? Describe data joins and dom element links.

Modules

The way D3 is split up into modules, the core package and what kind of extensions are there.

3. Implementation

How are the chosen diagrams implemented? Which D3 modules have been used? How was the implementation done?

3.1 Showcase

How is the showcase structured? How can you get there? Why does it exist? Who might benefit? How can you reuse a part the interesting parts?

3.1.1 Integration of each diagram

How is each diagram integrated? How can you access them? Where can you grab them standalone?

3.1.2 Data Updates

How can you simulate data changes? Why is this useful?

3.2 Diagrams

describe all the diagrams and why they are special and what makes them tick

3.2.1 Bar Chart

How does it work? Which d3 features does it use? how do they work?

Pie chart

How does it work? Which d3 features does it use? how do they work?

Tree map

How does it work? Which d3 features does it use? how do they work?

Sankey

How does it work? Which d3 features does it use? how do they work?

Area graph

How does it work? Which d3 features does it use? how do they work?

Circle graph

How does it work? Which d3 features does it use? how do they work?

4. Conclusion

How well did it work? Was it worth the effort? What could be improved?

5. Sources

This should ideally be done automatically by bibtex.

6. Appendix

I guess this should contain all the source code. Maybe there are ways to import it automatically too?