IVI

University of Applied Sciences Northwestern Switzerland

Boran Eker
BSc Data Science Student

December 23, 2024

Contents

Introduction

In an age where data grows at an unprecedented rate, interactive visualizations have emerged as a critical tool for uncovering insights and fostering informed decision-making. Echoing Ben Shneiderman's mantra, "Overview first, zoom and filter, details on demand," this document emphasizes the importance of creating visual tools that not only present data but actively engage users in exploring it.

This document aims to guide the general public through the essential aspects of designing and evaluating interactive visualizations, addressing both the technical and human-centered challenges inherent in the process. By focusing on performance optimization, dashboard design, human-computer interaction (HCI) principles, and usability evaluation, this document provides a comprehensive framework for crafting effective and meaningful visualizations.

Modern datasets often contain vast amounts of information spanning numerous variables, making static representations insufficient for exploration. Interactive visualizations bridge this gap, allowing users to filter, zoom, and manipulate data dynamically. However, the complexity of large datasets can also strain computational resources, highlighting the need for performance-conscious design. Additionally, effective visualizations must prioritize usability, ensuring that users can intuitively navigate and interact with the data.

To address these challenges, this document is structured around four key learning objectives:

- **Performance:** Learn to identify and mitigate performance issues in interactive visualizations, such as rendering delays or sluggish interactivity, by leveraging techniques like data tiling, GPU acceleration, and hierarchical data organization.
- **Dashboard Design Principles:** Explore the art of designing dashboards with multiple linked views, applying concepts like brushing and linking, focus+context displays, and Shneiderman's mantra to create intuitive user experiences.
- **HCI Basics:** Delve into the principles of interaction design, understanding how to optimize user input modalities (e.g., mouse, touch, voice) and ensure compatibility across devices while incorporating advanced techniques like gesture-based interactions.
- **Evaluation:** Develop a systematic approach to assessing usability and user experience through small-scale user studies, informed by quantitative and qualitative methods tailored to the scope of your projects.

By integrating these elements, this document equips the general public with the skills needed to create performant, user-friendly, and impactful interactive visualizations.

For further details on the technical implementation, the complete source code of the dashboard is available on my GitHub Repository. This includes the scripts as pdf and the whole code in Python.

LO 1 Performance

Slow data processing and delayed visualization results can significantly hinder user experience and application efficiency. When dealing with large datasets, optimizing the performance of data processing workflows and interactive visualizations becomes a critical challenge. Users expect data to load and process quickly, enabling real-time insights and seamless interaction. This demand is not limited to scientific or industrial applications but extends to sectors like finance and healthcare, where rapid decisions based on large datasets are crucial.

A specific example can be found in the analysis of natural disaster data, such as flood events, where extensive datasets need to be processed to provide accurate analyses and forecasts. In this context, performance affects not only user satisfaction but also the ability to make timely and informed decisions. This chapter explores the challenges of processing large datasets, with a particular focus on performance testing and optimization strategies.

1.1 Scaling Tests

Scaling tests are methods used to evaluate the performance of a system under varying levels of load. They aim to assess the efficiency and stability of software systems or applications, particularly as the number of users or the volume of data increases. In a scaling test, the system is first tested under normal conditions and then gradually subjected to increasing loads until its maximum capacity is reached or exceeded. The primary goal is to identify bottlenecks and ensure that the system operates reliably even under high loads.[6]

A distinction is often made between vertical scaling, where the performance is enhanced by adding more resources like CPU or RAM to a single server, and horizontal scaling, where additional servers or instances are added. A specific approach within horizontal scaling involves testing the system's processing capacity by incrementally increasing or decreasing the dataset size. This method is frequently employed to understand how well a system handles varying data volumes and to ensure scalability.[4]

1.2 Methodology and Tools

Scaling tests require a structured approach and specialized tools to simulate realistic scenarios and analyze the results accurately. A common method involves simulating data loads by varying the dataset size, as described in horizontal scaling. For instance, a dataset can be split into smaller portions, and the processing speed of the system is tested with each portion. During these tests, performance metrics such as CPU and memory usage, latency, error rates, and throughput are monitored to assess the system's efficiency.[16]

Conducting such tests not only helps to understand the technical limits of a system but also reveals optimization opportunities in its architecture. This is particularly important for applications that operate under demanding conditions, such as e-commerce platforms or streaming services, where the ability to handle large or fluctuating data volumes is critical. This ensures that developers can optimize the architecture to meet the expected load conditions.[1]

1.3 Self-Experiment: Performance Test with Reduced Data Sizes

The processing speed of data heavily depends on the size of the data being processed and the efficiency of the underlying code. To investigate this, I conducted a performance test measuring the runtime for varying data volumes. As an example, I used a dataset containing flood data and implemented a Python function to measure the processing time as a function of the data size. The data volume was gradually reduced (10%, 25%, 50%, 75%, and 100%), and the processing time was recorded using Python's time module.[14]

The results show a clear relationship between data size and processing time. As the data size increases, the processing time also increases. Interestingly, the differences for smaller data sizes were relatively minor, but a significant increase in processing time was observed for larger datasets. This trend is evident in the graphs below. The line chart illustrates the increase in processing time with growing data size, while the bar chart highlights the differences between the percentages more clearly.[10]

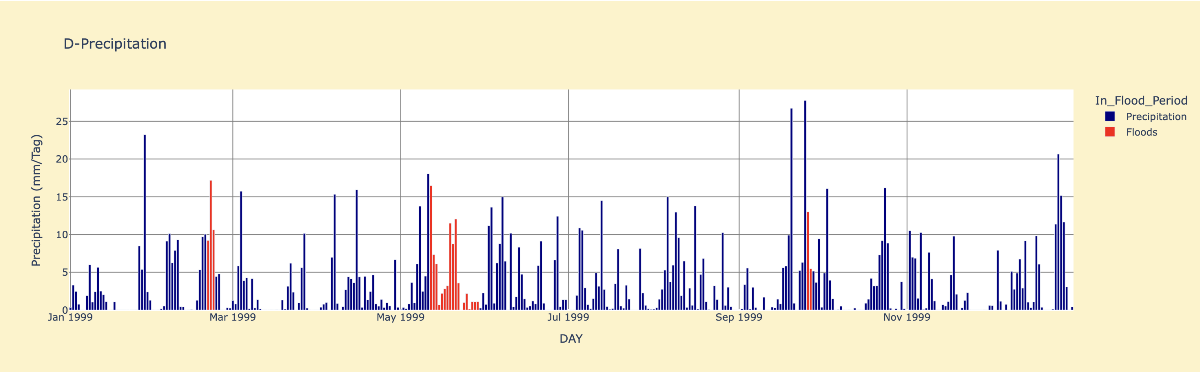


Figure 1.1: Interactive Flooding Histogram

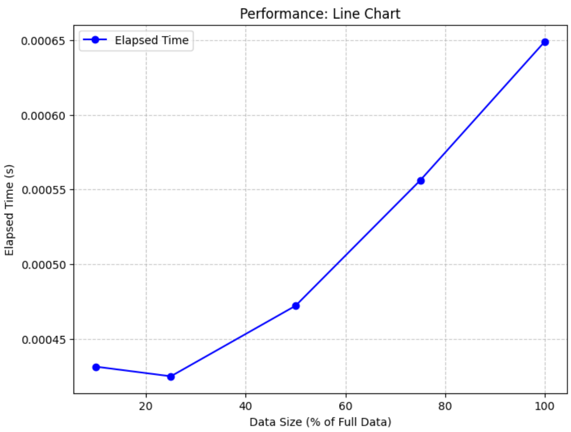


Figure 1.2: Line Chart Performance

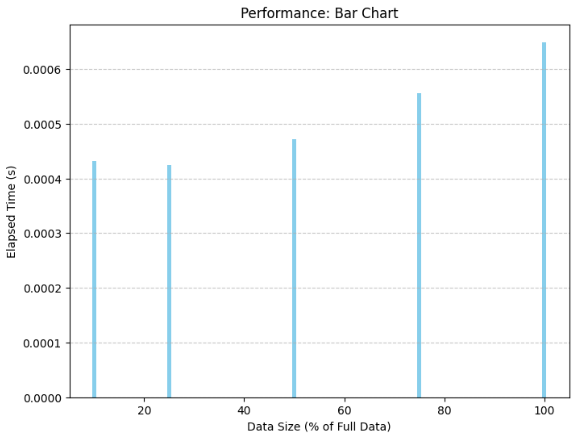


Figure 1.3: Bar Chart Performance

1.4 Optimization Strategies

The increase in processing time can be attributed to the growing data volume and the limitations of the hardware. To improve performance, several optimization strategies can be employed. One approach is to make the processing more efficient by optimizing calculations and eliminating unnecessary operations. For larger datasets, distributing tasks across multiple processor cores to utilize parallel processing can lead to substantial improvements.[13]

Alternatively, using more powerful hardware, such as a modern computer with a faster CPU and GPU, can significantly enhance performance. If local hardware is insufficient, cloud computing services like AWS or Google Colab can be used to access high-performance resources on demand. These services offer flexible scaling options, allowing large datasets to be processed quickly without requiring costly hardware upgrades.[5]

Finally, testing and refining the code can help identify bottlenecks and reduce runtime further. This is particularly relevant when large datasets need to be analyzed regularly, ensuring both efficiency and cost-effectiveness.

LO 2 Dashboard design principles

Ben Shneiderman, an American computer scientist, introduced the concept of the "Visual Information-Seeking Mantra" in 1996. This mantra outlines an effective sequence for interactive data exploration: *Overview first, zoom and filter, details on demand* [11]. This principle forms the foundation of modern interactive visualization tools and has been comprehensively implemented in our dashboard.

Shneiderman's Mantra – Overview

The overview stage provides a broad visualization of the data, offering users a general understanding of trends and patterns. In our dashboard, this is achieved through an interactive map that visualizes flooding events based on geographic coordinates. This allows users to quickly grasp the relationship between rainfall and flood occurrences across different regions. The use of tools like Plotly ensures clarity and interactivity in the initial representation of the data [8].

Shneiderman's Mantra – Filter

Filtering enables the reduction of displayed data to focus on specific relationships and patterns. In the dashboard, users can select a specific year using a slider to visualize flooding events from that period. Additionally, a dropdown menu allows users to restrict the view to a specific country. These filtering mechanisms dynamically update the visualizations, helping users immediately understand the impact of their selections. Furthermore, the precipitation histogram allows users to filter data by timeframe (daily, weekly, or monthly), offering a granular perspective on precipitation patterns and their relationship to flooding events. This capability enhances the user's ability to uncover specific trends and correlations. Such interactivity is central to effective data exploration, as highlighted by Shneiderman's emphasis on reducing complexity to enhance perception [11].

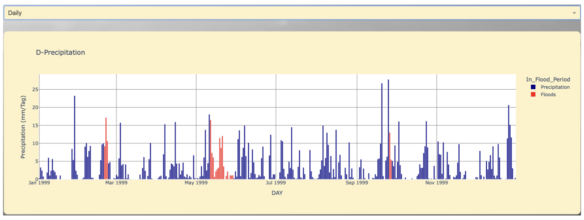


Figure 2.1: Daily Precipitation without filter

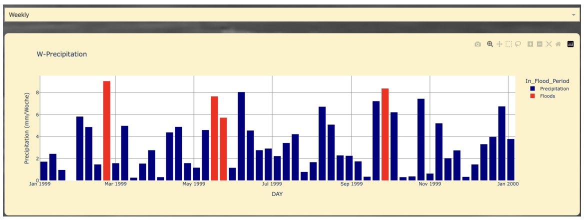


Figure 2.2: Weekly Precipitation with filter

Shneiderman's Mantra – Zoom

Zooming offers the ability to focus on specific areas of interest within the data. In our dashboard, users can zoom into the interactive map to examine flood events in detail, such as the extent of damages or the number of fatalities in highly affected regions. By focusing attention on relevant information while minimizing distractions, zooming enhances the analytical depth, aligning with Shneiderman's recommendation for intuitive focus [8].

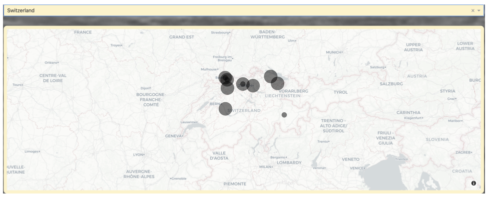


Figure 2.3: Zoomed-in Map of Switzerland

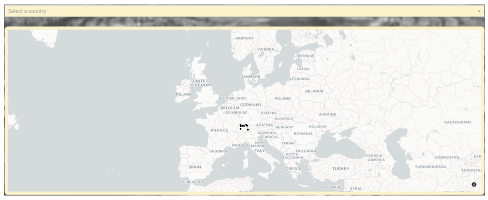


Figure 2.4: Overview Map of Europe

Shneiderman's Mantra – Details on Demand

Providing details on demand is a key component of interactive visualizations. In our dashboard, users can hover over specific points on the map to retrieve additional details about each flooding event, including location, duration, damage, and fatalities. This feature ensures that users receive relevant information only when required, thereby maintaining the simplicity and clarity of the overall visualization. Murray (2017) emphasizes the importance of providing contextual details dynamically to engage users and support deeper insights [8].

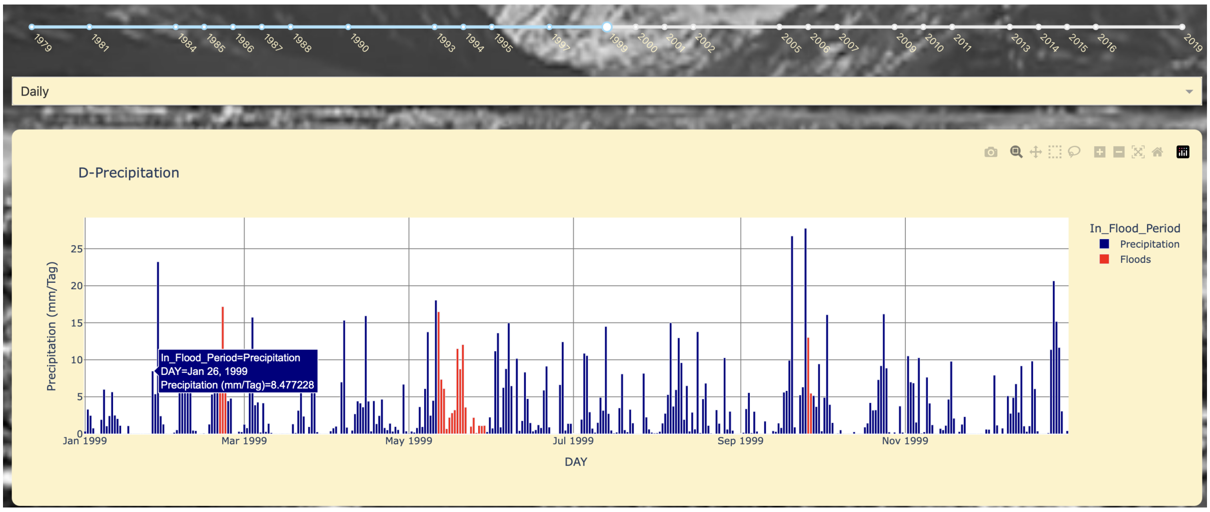


Figure 2.5: Histogram showing the number of flooding events filtered by year and country, dynamically updated based on user selections.

Brushing and Linking

Brushing and linking are essential techniques in interactive dashboards [8]. Brushing refers to the selective highlighting of data points within a visualization, while linking ensures that this selection is reflected across connected visualizations. In our dashboard, marking a specific time frame in the rainfall chart highlights corresponding flooding events on the map. This interconnected behavior simplifies the process of identifying relationships between variables and supports hypothesis-driven data analysis.

LO 3 HCI Basics

HCI stands for Human-Computer Interaction and focuses on how humans interact with computer technologies. The goal is to design intuitive and efficient user experiences, ranging from simple systems like traffic lights to complex dashboards. This section outlines the HCI principles applied in the development of an interactive dashboard for visualizing flood and precipitation data.

Interactive Features: Selection, Zoom, and Filtering

Basic interactive functions, such as selection, zoom, and filtering, play a critical role in dashboard usability. For instance, the dropdown menu in this dashboard allows users to select specific countries, while the year slider facilitates navigation through historical data. These features align with the principles of the *Cognitive Walkthrough* method, which evaluates potential usability issues step-by-step [9].

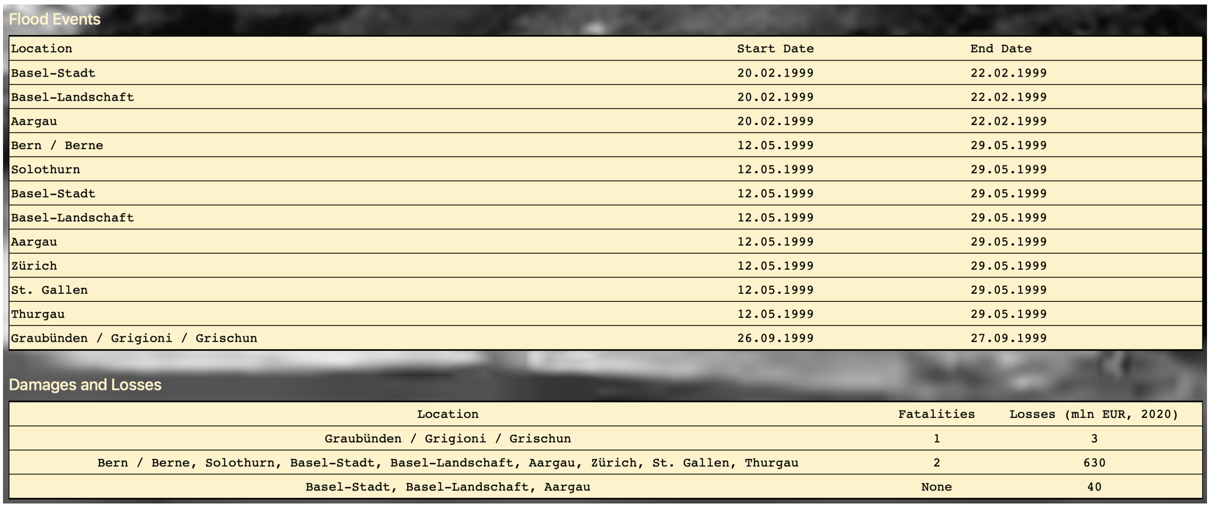
The map visualization tool supports zoom functionality, incorporating *Fitts' Law*, which states that larger and closer interaction targets are easier to access [3]. This is reflected in the use of large markers for flood events, simplifying user navigation.

Responsive Design and Input Modalities

This dashboard is optimized for desktop use but features a responsive design adaptable to various screen sizes. The primary input modalities are mouse and keyboard, as these are preferred for desktop environments. Future iterations could include touchscreen support, especially for tablets and smartphones. According to research in *Interaction Techniques*, multimodal inputs, such as voice and gesture controls, offer opportunities to enhance user experience [2].

Information Design: Miller's Law

The information presentation in the dashboard follows *Miller's Law*, which states that humans can store about 7 ± 2 items in short-term memory [7]. To reduce cognitive load, the dashboard avoids overcrowding the interface. Dropdown menus and tables are organized into manageable "chunks," making navigation intuitive and user-friendly.



Flood Events

Location	Start Date	End Date
Basel-Stadt	20.02.1999	22.02.1999
Basel-Landschaft	20.02.1999	22.02.1999
Aargau	20.02.1999	22.02.1999
Bern / Berne	12.05.1999	29.05.1999
Solothurn	12.05.1999	29.05.1999
Basel-Stadt	12.05.1999	29.05.1999
Basel-Landschaft	12.05.1999	29.05.1999
Aargau	12.05.1999	29.05.1999
Zürich	12.05.1999	29.05.1999
St. Gallen	12.05.1999	29.05.1999
Thurgau	12.05.1999	29.05.1999
Graubünden / Grigioni / Grischun	26.09.1999	27.09.1999

Damages and Losses

Location	Fatalities	Losses (mln EUR, 2020)
Graubünden / Grigioni / Grischun	1	3
Bern / Berne, Solothurn, Basel-Stadt, Basel-Landschaft, Aargau, Zürich, St. Gallen, Thurgau	2	630
Basel-Stadt, Basel-Landschaft, Aargau	None	40

Figure 3.1: Flood Events

Perception Psychology: Weber’s Law and Change Blindness

The visualizations are designed with *Weber’s Law* in mind, which describes the difficulty of perceiving small differences when stimuli intensity is high [15]. Significant data points, such as larger markers for greater flood damages, are highlighted to improve visibility.

Change Blindness, the phenomenon where users fail to notice large changes in a visual scene, is also addressed [12]. Clear color distinctions were implemented to differentiate flood periods (red) from normal precipitation (blue) in the bar charts, ensuring users can easily identify critical patterns.

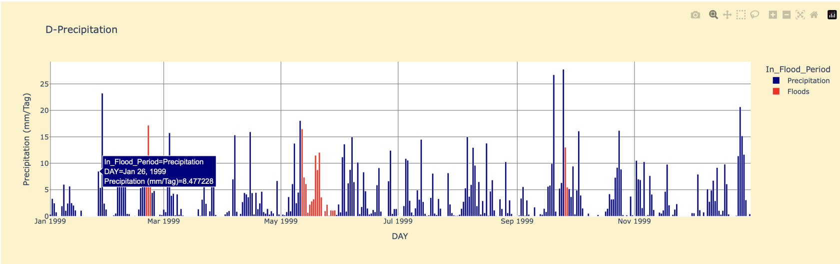


Figure 3.2: Precipitation Overview

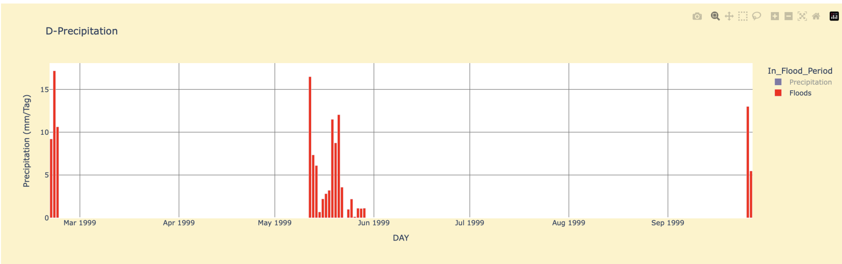


Figure 3.3: Floods Overview

Navigation with Animated Transitions and Breadcrumbs

Future improvements could include animated transitions between views, such as smooth changes when switching data for different years or countries. Breadcrumbs could also be introduced to display the user’s current context and enhance the traceability of their interactions.

Conclusion

The dashboard’s design adheres to established HCI principles, leveraging interactive features to provide an intuitive and efficient user experience. By incorporating *Fitts’ Law*, *Miller’s Law*, and *Weber’s Law*, and by applying the *Cognitive Walkthrough* method, the application ensures a user-friendly and accessible interface.

LO 4 Evaluation

Evaluating interactive visualizations is an essential step to ensure their usability and effectiveness. A dashboard must do more than just display data—it should empower users to interact with the information intuitively and gain meaningful insights. Usability testing provides valuable feedback on how well a dashboard meets these goals, highlighting areas that work well and those that require improvement.

There are various methods to evaluate visualizations, as outlined by Christian Rohrer (2008), which classify research methods based on the type of data (behavioral vs. attitudinal) and the approach (qualitative vs. quantitative). Methods such as usability lab studies, ethnographic field studies, and A/B testing allow researchers to gather behavioral data, while surveys and interviews often focus on attitudinal insights. Each method has its strengths and limitations, and the choice of approach depends on the specific goals of the evaluation, as well as the available resources.

In the context of this dashboard, which focuses on precipitation and flood data in Switzerland, the chosen evaluation approach was primarily qualitative, combining observational usability testing and user feedback. While quantitative methods, such as A/B testing or controlled experiments, could provide deeper insights into task efficiency and user performance metrics, these were not feasible due to resource constraints, such as limited access to a large and diverse participant pool and the technical setup required for advanced testing.

Instead, a focused usability test was conducted with a small group of participants, providing valuable insights into user interaction and identifying areas for improvement. This practical approach allowed us to align the evaluation process with the scope and constraints of the project while still adhering to key principles of user-centered design.

User Test

The user test was conducted to determine how intuitive and user-friendly the dashboard is for users with varying levels of expertise. The main goal was to evaluate the structure, design, functionality, and user behavior when interacting with the dashboard. Special attention was given to the accessibility of regional data for Switzerland and the clarity of visualizations related to precipitation and flood events.

A key focus was on ensuring that users could navigate the interface easily and understand the data without requiring extensive prior knowledge. Interactive elements such as the map, year slider, dropdown menus, and time frame selector were evaluated for their intuitiveness and efficiency. Additionally, we sought to identify potential barriers that users might encounter during their interaction with the dashboard, such as difficulty in interpreting charts or navigating through tables.

Test Procedure

The user test involved five participants aged between 18 and 25. Three participants were female, and two were male. Among them, one had in-depth knowledge of Swiss regions and climatic conditions, while the other four had only general awareness about rainfall and flooding. This diversity allowed us to observe the dashboard's usability for both experienced and inexperienced users.

The test was conducted on a laptop, and participants were given a brief introduction to the dashboard, including its purpose and an overview of its features. Afterward, participants were presented with six tasks designed to test various aspects of usability, including navigation, data interpretation, and interaction with the dashboard's features.

For example, one task required participants to use the year slider to display flood events for a specific year and analyze the patterns. Another task involved exploring the Swiss map to identify regions with the most flood events and accessing detailed information about those events. Participants were also tasked with correlating flood periods with precipitation levels using the interactive bar chart and analyzing the data tables for insights into damages and losses caused by floods.

Throughout the test, observations were made regarding the participants' interactions with the dashboard. Feedback was also collected through follow-up questions to understand their experience, any challenges they faced, and suggestions for improvement.

Test Results

In the first task, participants were asked to use the year slider to display flood events for a specific year. All participants successfully completed this task without difficulty. The functionality of the slider was described as intuitive, and the visual feedback it provided was considered helpful in understanding the selected year's data. This confirmed that the slider is an effective tool for temporal exploration.

In another task, participants used the map to explore flood data for specific Swiss regions. While four participants navigated the map seamlessly, one participant struggled with zooming in and out due to unfamiliarity with the trackpad. Despite this minor challenge, all participants successfully located the requested regions and understood the data associated with them. The map's interactive markers, which display detailed information on hover, were praised for their clarity.

The task involving the correlation of precipitation levels with flood periods proved to be one of the most engaging for participants. Using the bar chart, they could quickly identify spikes in precipitation during flood periods. The color-coded distinction between precipitation data and flood periods was especially appreciated for its clarity. However, one participant noted that the legend could be more prominently displayed to enhance accessibility.

Analyzing the data tables, participants were able to interpret and filter information about flood damages and losses. Four participants completed this task with ease, while one participant suggested that a search function would make filtering data faster and more convenient. This feedback highlighted an opportunity for further enhancing the dashboard's usability.

Concluding Remarks

The user test demonstrated that the dashboard provides a user-friendly and accessible platform for visualizing precipitation and flood data in Switzerland. Its interactive features, such as the map, year slider, and time frame selector, were well-received by participants. The design and functionality of the visualizations enabled users to gain meaningful insights, regardless of their prior knowledge of the subject matter. Importantly, the test confirmed that the dashboard effectively meets the needs of both novice and experienced users.

At the same time, the test identified areas for improvement. The addition of a search function, better prominence for the legend, and potential refinements to the map's zooming functionality could further enhance the dashboard's usability. These enhancements would address the minor challenges observed during the test and make the dashboard even more accessible.

Future Improvements

To improve the dashboard further, several enhancements are recommended. Introducing a search function would allow users to locate specific regions or events more efficiently. This feature would be particularly beneficial for tasks requiring detailed exploration of the dataset. Additionally, refining the legend's placement and design could make it more noticeable, helping users quickly interpret visualizations.

Providing a brief tutorial or an introductory video could assist first-time users in understanding the dashboard's features and functionality. This would reduce the learning curve and make the dashboard more welcoming for new users. Finally, future user tests could include a larger and more diverse group of participants to gain broader insights. Using a timer to measure task completion times could also provide valuable data on the efficiency of navigation and interaction.

Overall, the user test has confirmed the dashboard's strengths while highlighting actionable opportunities for improvement, ensuring its continued development as a powerful tool for visualizing precipitation and flood data in Switzerland.

Appendix: Visualizations

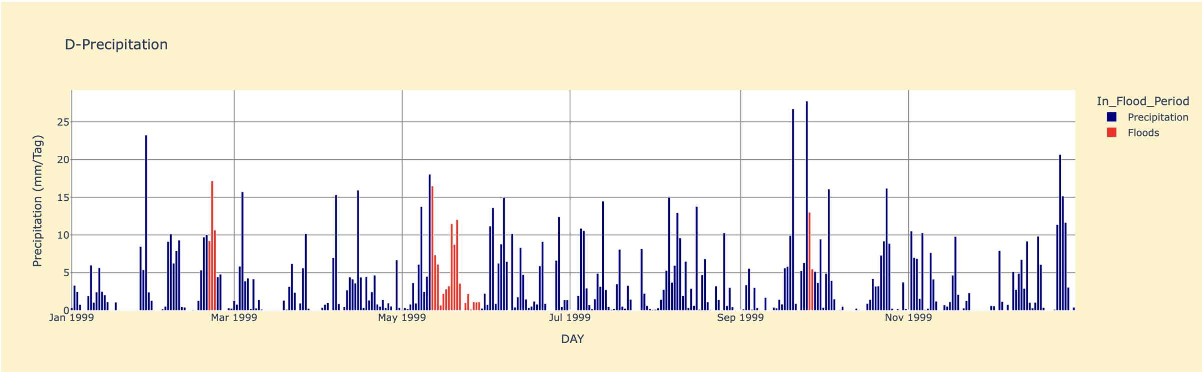


Figure 4.1: Interactive Flooding Histogram

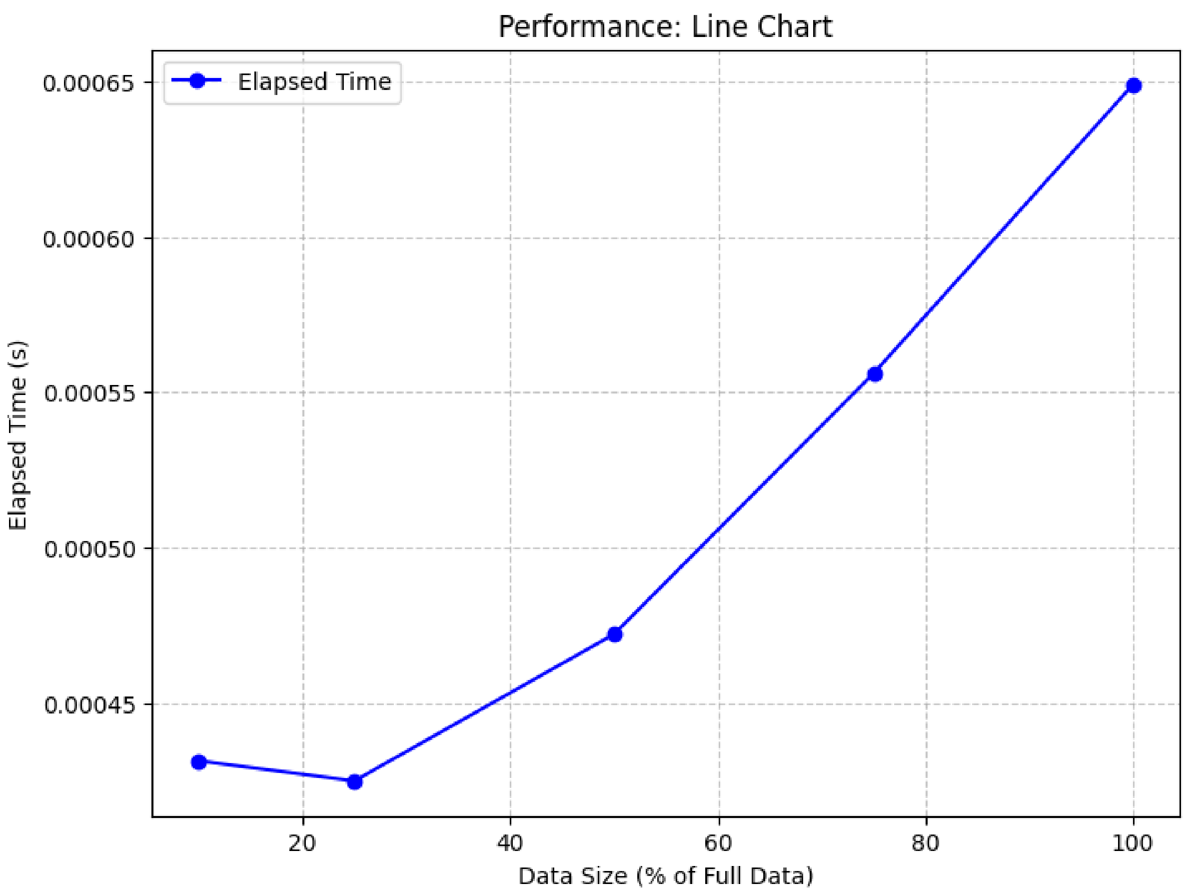


Figure 4.2: Line Chart Performance

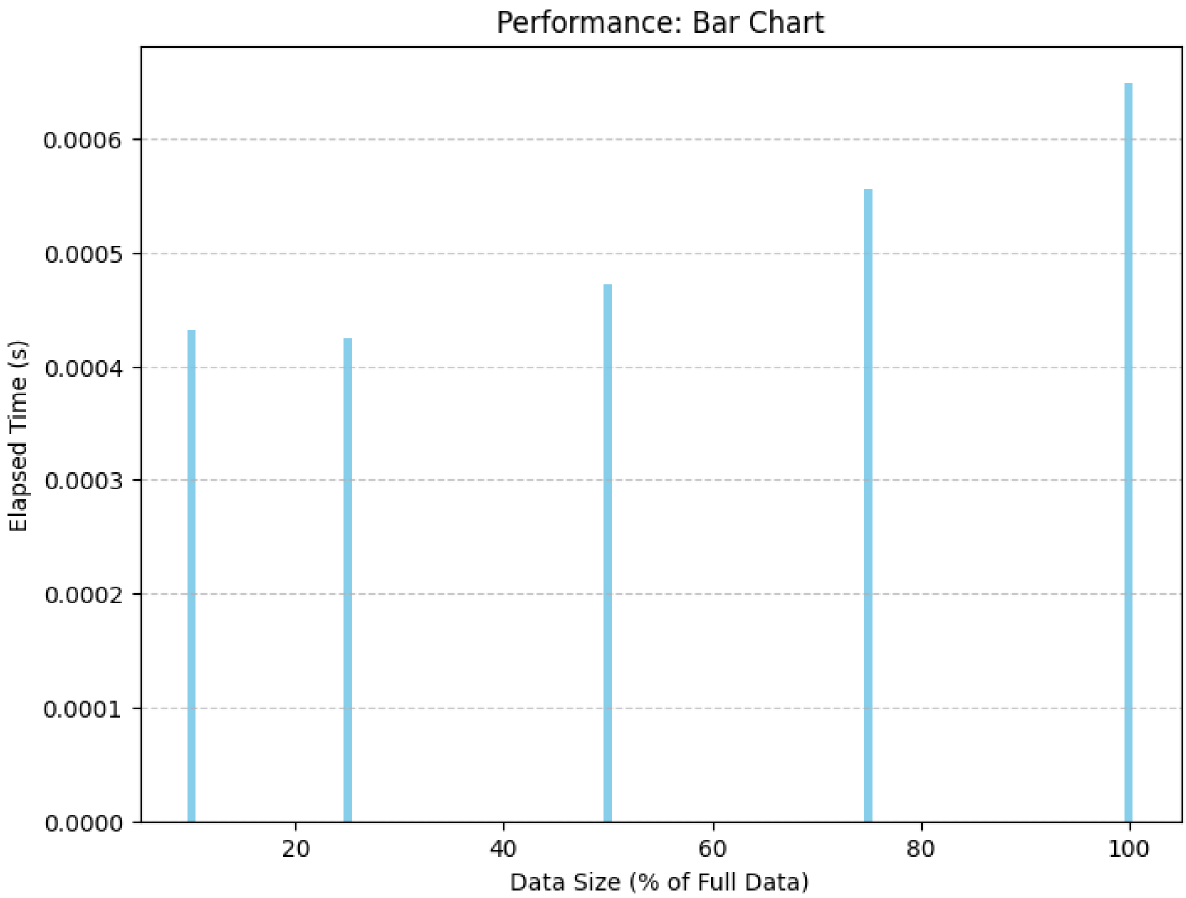


Figure 4.3: Bar Chart Performance

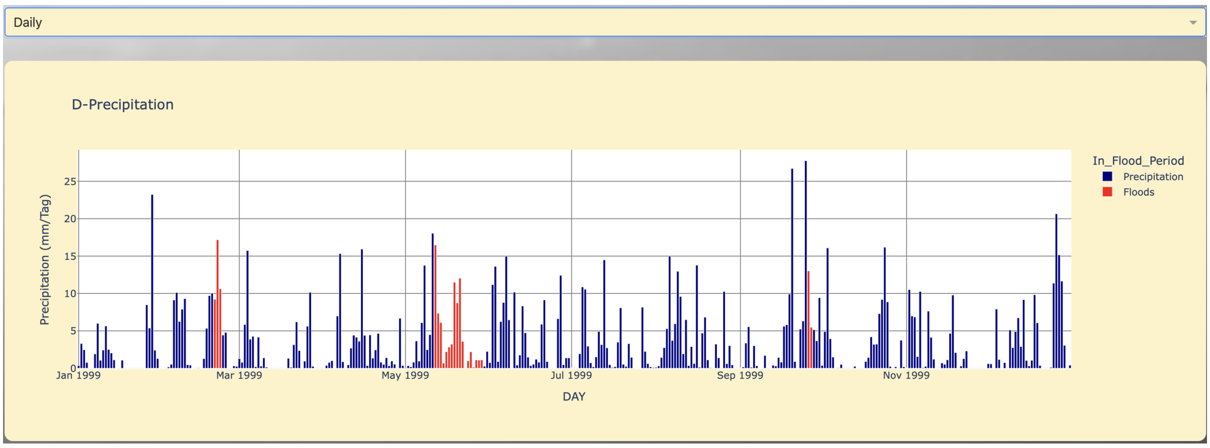


Figure 4.4: Daily Precipitation

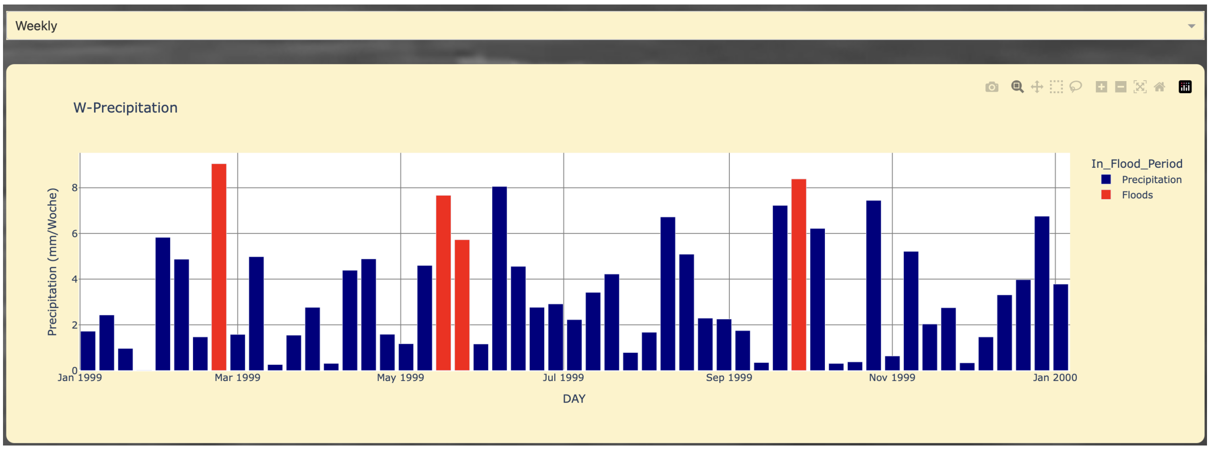


Figure 4.5: Weekly Precipitation

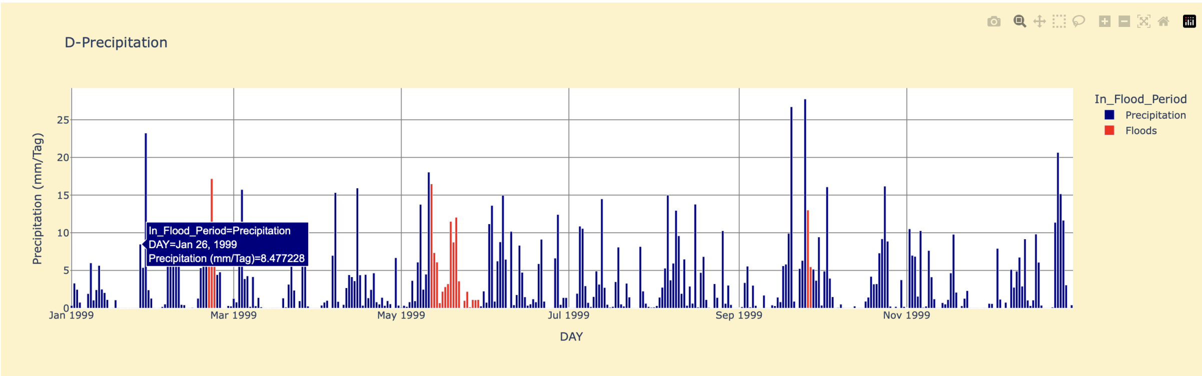


Figure 4.6: Precipitation Overview

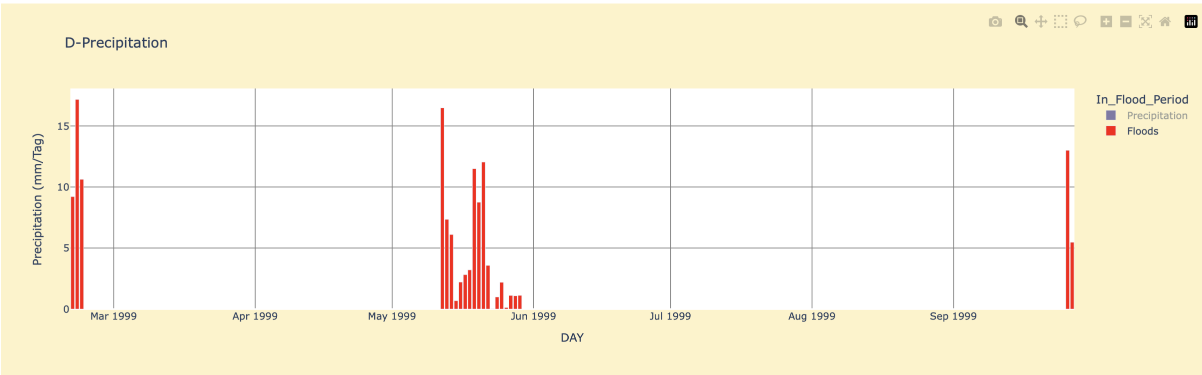


Figure 4.7: Floods Overview

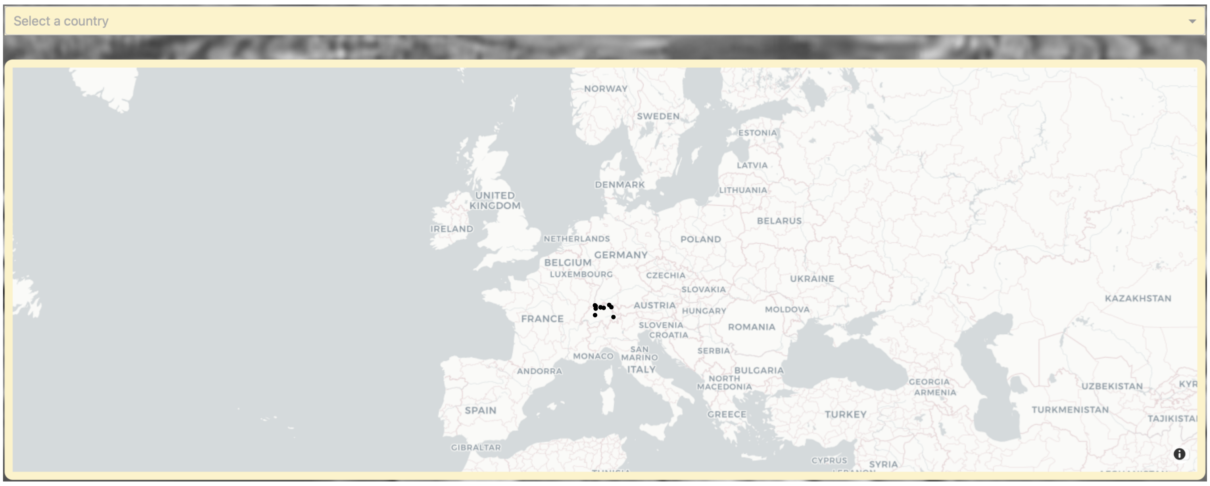


Figure 4.8: Overview Map of Europe

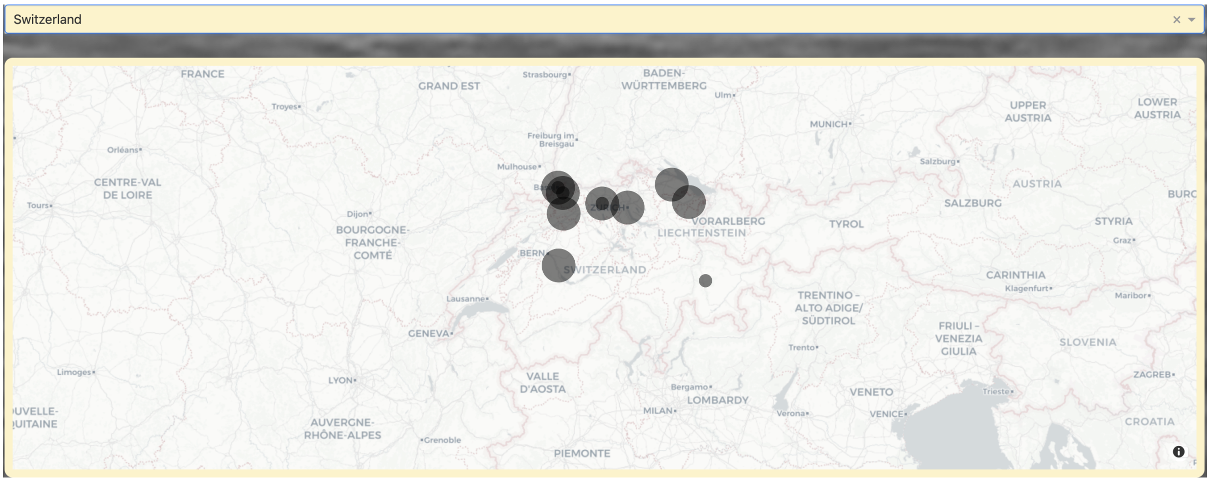


Figure 4.9: Zoomed-in Map of Switzerland

Flood Events

Location	Start Date	End Date
Basel-Stadt	20.02.1999	22.02.1999
Basel-Landschaft	20.02.1999	22.02.1999
Aargau	20.02.1999	22.02.1999
Bern / Berne	12.05.1999	29.05.1999
Solothurn	12.05.1999	29.05.1999
Basel-Stadt	12.05.1999	29.05.1999
Basel-Landschaft	12.05.1999	29.05.1999
Aargau	12.05.1999	29.05.1999
Zürich	12.05.1999	29.05.1999
St. Gallen	12.05.1999	29.05.1999
Thurgau	12.05.1999	29.05.1999
Graubünden / Grigioni / Grischun	26.09.1999	27.09.1999

Damages and Losses

Location	Fatalities	Losses (mln EUR, 2020)
Graubünden / Grigioni / Grischun	1	3
Bern / Berne, Solothurn, Basel-Stadt, Basel-Landschaft, Aargau, Zürich, St. Gallen, Thurgau	2	630
Basel-Stadt, Basel-Landschaft, Aargau	None	40

Figure 4.10: Flood Events

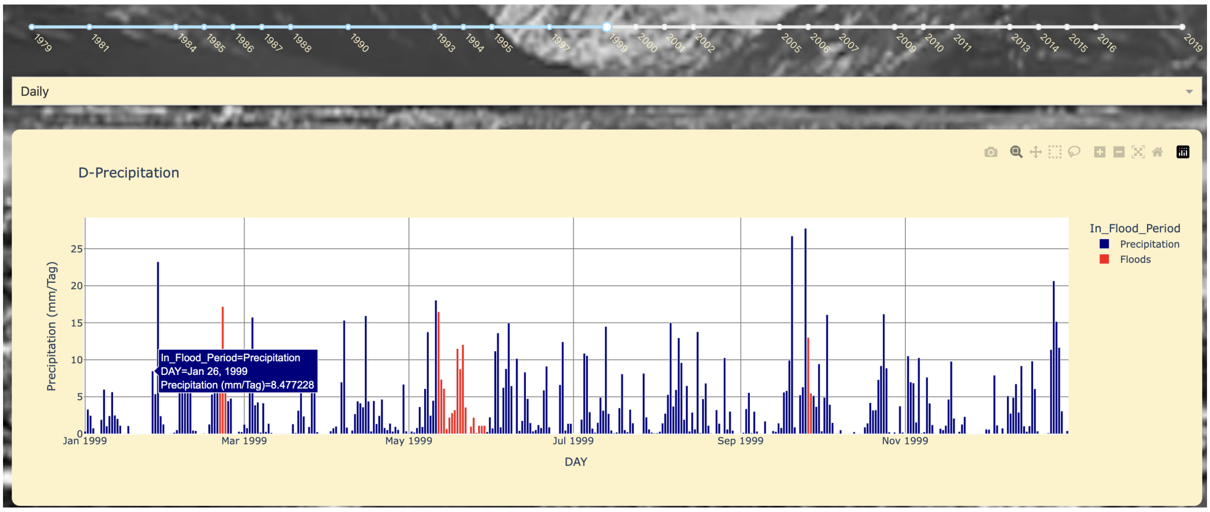


Figure 4.11: Histogram showing the number of flooding events filtered by year and country, dynamically updated based on user selections.

Bibliography

[1] Acquisa. Skalierbarkeit: Ein umfassender Überblick, 2024. Accessed: 2024-12-10.

[2] Wikipedia Contributors. Interaction technique, 2022. Accessed: 2024-12-04.

[3] Paul M. Fitts. Fitts' law, 2022. Accessed: 2024-12-04.

[4] GeeksforGeeks. Overview of scaling: Vertical and horizontal scaling, 2024. Accessed: 2024-12-10.

[5] KIT Technologies. Cloud computing einfach erklärt: Ein umfassender leitfaden, 2024. Accessed: 2024-12-10.

[6] LambdaTest. What is scalability testing: Examples and best practices, 2024. Accessed: 2024-12-10.

[7] George A. Miller. Miller's law, 2018. Accessed: 2024-12-04.

[8] Scott Murray. *Interactive Data Visualization for the Web*. O'Reilly Media, 2nd edition, 2017.

[9] Usability Body of Knowledge. Cognitive walkthrough, 2022. Accessed: 2024-12-04.

[10] Real Python. Python timer: Introduction to time and timeit libraries, 2024. Accessed: 2024-12-10.

[11] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. *Proceedings of the IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[12] Daniel J. Simons and Daniel T. Levin. Change blindness. *Trends in Cognitive Sciences*, 1(7):261–267, 1997.

[13] Statistics Easily. Was ist parallelverarbeitung in der datenwissenschaft?, 2024. Accessed: 2024-12-10.

[14] Statistics Easily. What is computation time? explained for data science, 2024. Accessed: 2024-12-10.

[15] Ernst Heinrich Weber. Weber's law, 2022. Accessed: 2024-12-04.

[16] Wikipedia. Scalability testing, 2024. Accessed: 2024-12-10.