

Visualizing personal spending data

Motivation

The core idea induced in the past, when I started to track my personal expenses. Without deeper knowledge about charts and visualizations, I created a simple project to record, preprocess and plot the gathered data in a primitive way. The main goal of this project is to visualize the data in a more sophisticated way, using more advanced visualization techniques and tools.

Description

The layout consist of several charts, each representing a different aspect of the data. Starting with a line chart showing the monthly sum over entire period when the data was recorded. The second chart shows comparison between last 12 months and previous 12 months. The next section is dedicated to filters. Available options are to select particular categories and/or to select a time range. The next two charts are mass chart and cumulative distribution over month. Another chart is violin plot presenting analysis for each category, followed by a barplot showing the number of records for each category during selected time range. Lastly a treemap is used to visualize how much money was spent in each category.

Dataset

The dataset is a collection of personal expenses, recorded in a CSV file. Considering the publicness of the project, the exact data is sampled and randomized from the original dataset.

Preprocessing

Preprocessing consists of calculating the stats for particular charts. The data is filtered based on the selected time range and categories. The data is then grouped by month and category, and the stats are calculated. The data is then used to plot the charts.

Fields are exactly as defined in the database.

Used Technology

The project is web-based, coded in TypeScript using React library. For data manipulation, data-forge library is used. For visualizations, Chart.js library is used. Additionally, the application is fully containerized using Docker.

Running the Project

Project is deployed at <https://cillik.eu>.

However, the project can also be run locally. The repository is publicly available at <https://github.com/xcillik/PV251-project>. The project is containerized, so the easiest way to run it is to use Docker. The Dockerfile is included in the repository. The project can be run using the following commands:

```
git clone https://github.com/xcillik/PV251-project.git
cd PV251-project
docker compose up --build
```

The application is then available at `http://localhost:8080`.

For more, detailed instructions, please refer to the README file in the repository.

Takeaways

Viewing personal expenses in a more sophisticated way can provide valuable insights. By examining the data, one can identify trends, anomalies, and patterns. This can help to optimize spending, set budgets, and plan for the future. The most obvious pattern is increasing spendings over time, which can be caused by inflation.

Screenshots

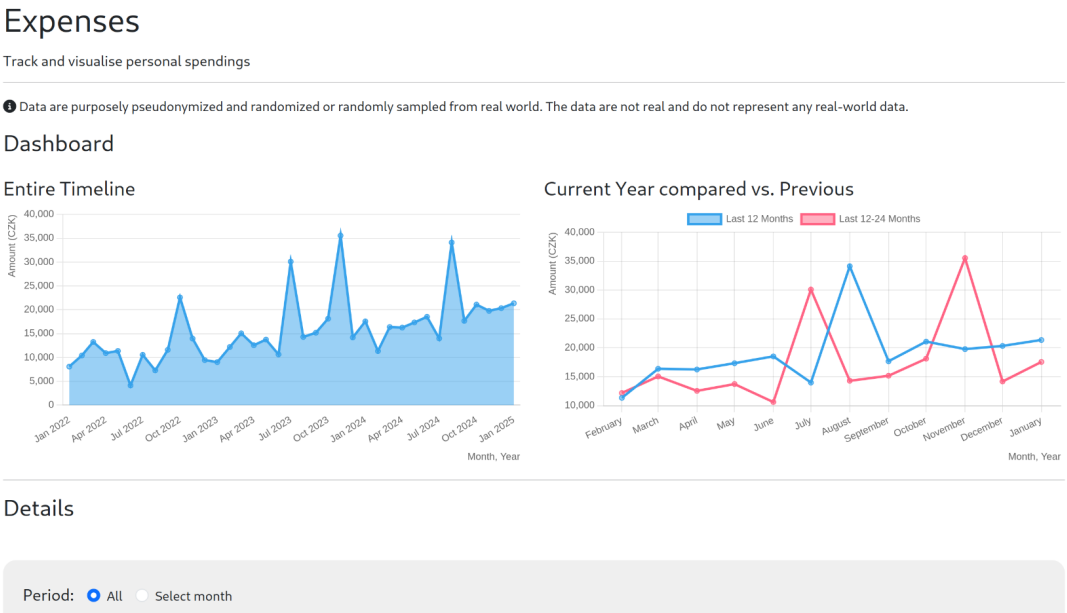


Figure 1: Entire Period

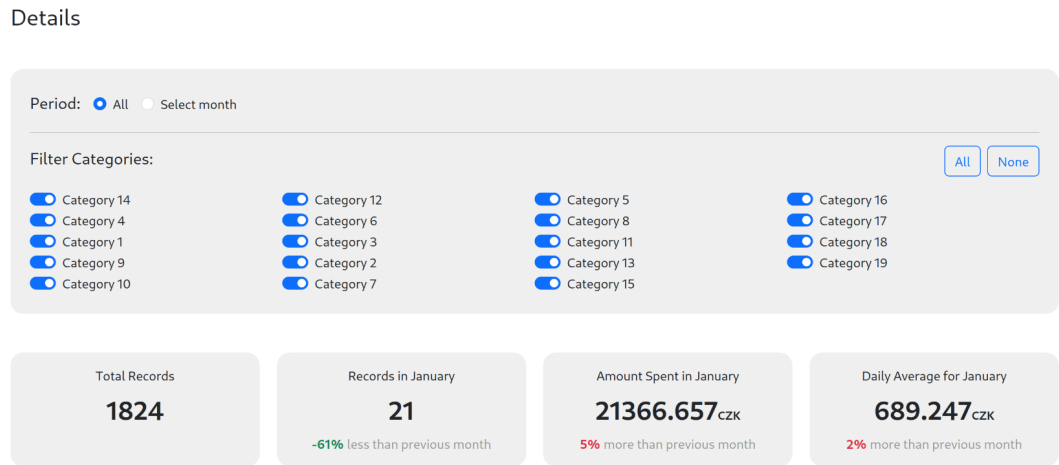
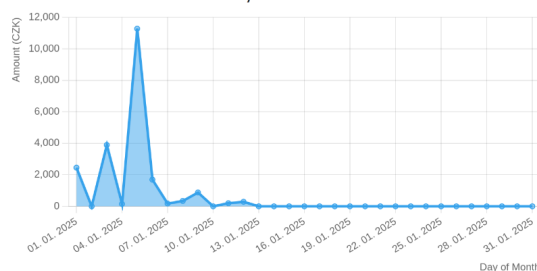
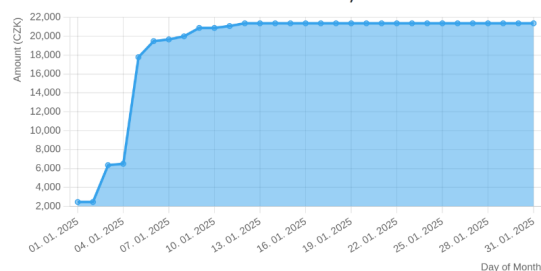


Figure 2: Filters

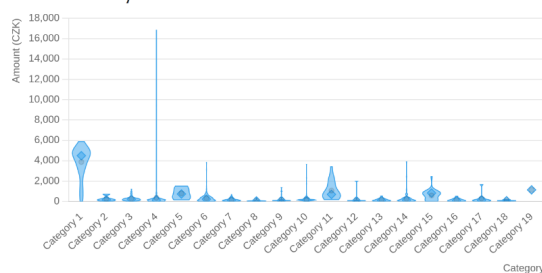
Mass Function over January 2025



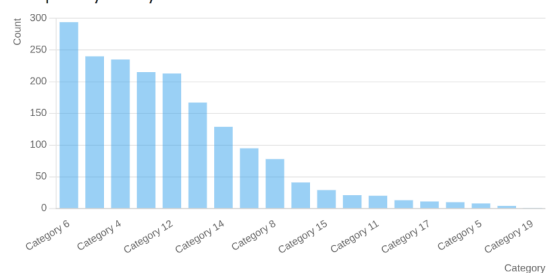
Cummulative Distribution over January 2025



Amount Analysis



Frequency Analysis



Part of the Whole

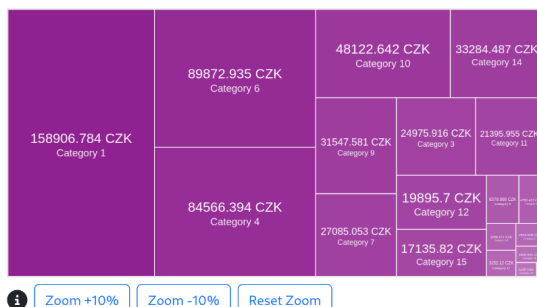


Figure 3: Additional Charts