Taggle: Scaling Table Visualization through Aggregation

Katarina Furmanova, Miroslava Jaresova, Bikram Kawan, Holger Stitz, Martin Ennemoser, Samuel Gratzl, Alexander Lex, Marc Streit

Figure 1: The Taggle interface consisting of (a) a table view and (b) a data selection panel showing a dataset on AIDS in several countries. The data selection panel allows users to filter out records by interacting with the attribute histograms. The selected rows indicate the relationship between new infections and death rate over time. The rows of countries (c) in Asia, Europe, and North America have been aggregated to histograms and box plots. Matrix columns (d) can be aggregated using statistical metrics. The visualizations of attributes can be changed via a toolbar (e) for both aggregated and non-aggregated data subsets.

Abstract

Tabular data visualizations are easy to understand and powerful at communicating patterns in datasets, especially when paired with interaction techniques such as sorting. In this work we present Taggle, a novel visualization technique for large and complex tables. We consider datasets that are composed of columns of categorical or numerical data, in addition to homogeneous matrices. The key contribution of Taggle is its ability to aggregate data subsets (rows and columns) on demand based on user-defined grouping rules.

Motivation

Visualization of tabular or multi-dimensional data is important in many application domains and a mainstay of visualization research. One significant drawback of tabular data visualization is that by drawing each item in a row, the available screen space is quickly exhausted. Our solution to this problem is the selective aggregation of rows encompassing diverse data types based on user driven selections.

Tabular Data Definition

Taggle is designed for exploration and presentation of tables that are a combination of both heterogeneous vectors and homogeneous matrices.

- **Vectors** are columns in the table where all associated records are of the same type and semantics.
- **Matrices** are composed of vectors of the same semantics and data type.

Implementation

Taggle was implemented as web application based on the Caleydo Phovea Platform using TypeScript.
Demo: **https://taggle.caleydoapp.org**

Taggle Concept

Taggles's main operations as illustrated in Figure 2 are:

- **Filtering:** Filters are defined by interacting with the histograms in the side panel.
- **Sorting and Ranking [1]:** Items can be hierarchically sorted, where ties can be broken using other vectors.
- **Grouping & Aggregation [2]:** Tables can be stratified into groups. These groups can be leveraged to aggregate the data.

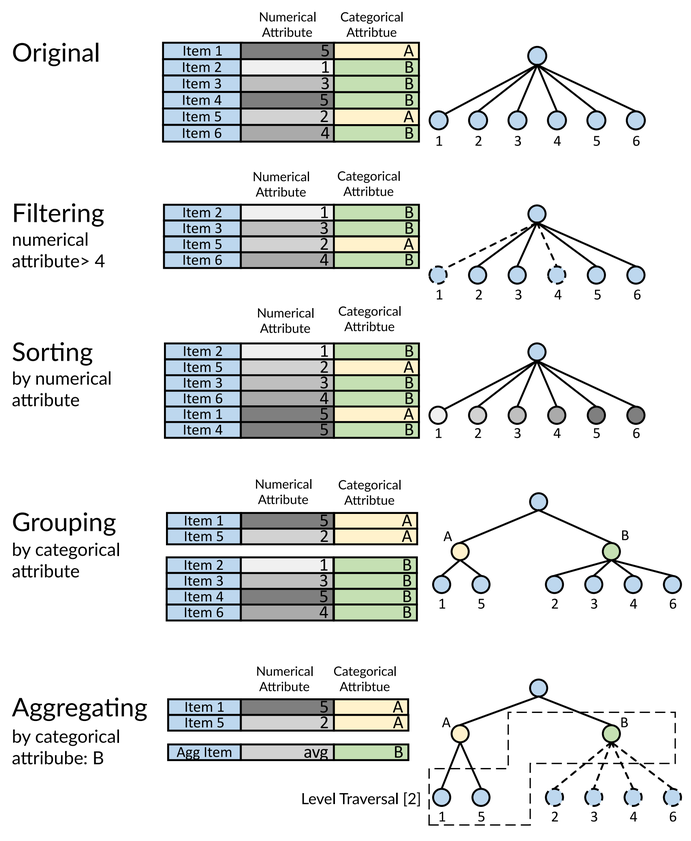


Figure 2: Topological operations in aggregation hierarchy.

Visualization and Interaction

The Taggle interface consists of two parts, as shown in Figure 1:

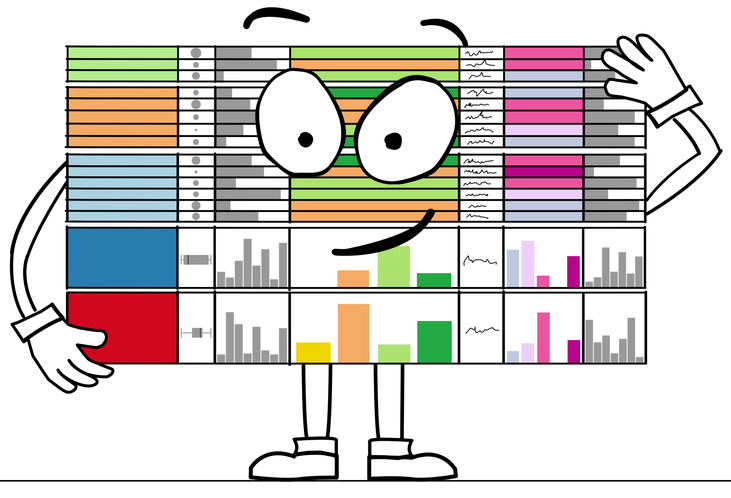
(a) Table View

The Table View is the main view and visualizes data vectors as columns. It encodes each selected vector or matrix using a mark appropriate for the data type, such as, bar plots for numerical data, colored cells for categorical data, and heat maps for matrices.

(b) Data Selection Panel

The Data Selection Panel is used for selecting the vectors and matrices to show in the Table View and for filtering. For each vector or matrix present in the Table View, the side panel shows a histogram of the data.

[1] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. LineUp: Visual Analysis of Multi-Attribute Rankings. IEEE Transactions on Visualization and Computer Graphics (InfoVis '13), 19(12):2277–2286, 2013.
[2] N. Elmqist, and J.D. Fekete. Hierarchical Aggregation for Information Visualization. IEEE Transactions on Visualization and Computer Graphics, pages 439–454. 2010.



Contact
furmanova@mail.muni.cz
caleydo.org
https://github.com/Caleydo/taggle

Acknowledgements
This work was supported by Boehringer Ingelheim Regional Center Vienna, the State of Upper Austria (FFG #851460), and the US National Institutes of Health (U01 CA198935).