Motivation

Exploratory data analysis is an iterative process in which analysts explore a variety of hypotheses by constructing different visualisations.

The history of visualisations explored by an analyst in this process can be useful for helping analysts
• keep track of their progress
• identify promising directions for further exploration
• illustrate the data analysis process to other analysts.
• collaborate with other analysts to jointly and more efficiently explore the dataset.

However, current exploratory data analysis software do not support tracking and sharing exploration histories.

Solution

Chive is a system designed to unobtrusively capture the complete history of an analyst’s data analysis process, and to enable collaboration between analysts via the sharing of exploration histories.

Key Features

• Full tracking of exploration history
• Metadata (titles, description, creation times) for intermediate visualisations
• Searching of intermediate visualisations on title, description, variables or data visualised
• Exporting and importing of exploration history for collaboration
• Recommendations for similar views that may be of interest

Navigating History

Local History View

The local history view is displayed in main application frame. It shows the structure of the history tree near the current node.

In the local history view, the current node is highlighted in yellow and shown in the context of its siblings. Details of each node can be viewed on hovering, which also reveals the node's children. Left-clicking on any node reverts the visualisation to the corresponding state, and right-clicking allows the associated metadata to be edited.

Global History View

The global history view opens in a new window, and shows the history tree using the familiar filesystem metaphor. Each node shows information such as the number of descendants with titles, with more details on hover. The view is linked to view in the main application, and updates in sync to center on the appropriate node. For consistency, click interactions are exactly the same as the local history view. Analysts can also search through history using this view.

Future Work

• Real-time collaboration on the web without requiring manual exchange of histories
• History operations such as merging histories, comparing histories, replaying parts of histories from one branch on another branch
• Improvements to the data analysis layer – more data analysis features such as different encodings, aggregation, better filtering, faster data manipulation, etc.