

Figure 9: This figure shows the same features and interactions for the Constellations system as Figure 3 in the main paper. We include this figure to allow for easier reference to implemented system components and functionalities.

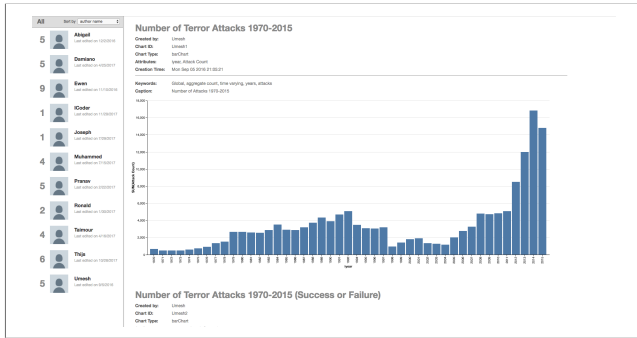


Figure 10: A screenshot of the Kaggle Kernels interface implemented for the user study.

10. User Study Text and Task Descriptions

This section contains the presented text and questions used in the different stages of the user study.

10.1. Training Stage

The Training Stage did not have an explicit script as participants were encouraged to ask questions during the system tutorial. Instead, after introducing the idea of “collaborative visual analytics” for large and deep datasets, the following lists itemize the necessary features and interactions that had to be covered to describe each system.

Kernels features and interactions:

- Ask if subject is familiar with data notebook IDEs; describe a data notebook if not
- Introduce the top pane (function board):
 - “Log In” link, “Task” buttons (Start Task 1, Stop Task 1, etc.)
- Introduce the left pane (user list board):
 - Elements: author picture, chart count, user name, last edit timestmep
 - Sorting options: name, chart count, and last edit time.
 - Clicking on a user selects them
 - “Show All” button shows all charts
- Introduce the right pane (chart board):
 - Charts ordering by creation timestamp
 - Chart elements: title, keywords, caption

Outline of Constellations features and interactions:

- Collab View Projection:
 - One circle per chart, colored by suer.
 - Circles positioned so that “similar” charts are close to each other (via chart encodings, keywords, dimensions)
 - Hovering to show tooltip, click to see Chart View
- Chart View:
 - Chart elements: title, keywords, caption, Vega-Lite specs, SQL panel, “Go Back” link
- Collab View, more functions:

- Top bar actions: Toggle image previews, toggle chart names, toggle author polylines, “Log In” link, “Task” buttons (Start Task 1, Stop Task 1, etc.)
- Filter board, time slider, lassos: How to add and delete filters and update the display
- Data coverage panel: treemap, Pearson correlation dropdown
- Cluster board: what clusters indicate, available stats in cluster board

10.2. Task Stage

The Task Stage instructed users to perform a set of three tasks (in order). The introductions for the stage and each task were printed on separates sheets of paper. These were individually handed to the participant, who read the instructions and subsequently began the appropriate task. The text for each document is as follows:

Task Stage

In this stage, you’re going to investigate a dataset called the **Terrorism Dataset**. The terrorism dataset shows information about terrorist attacks around the world since 1970. It’s a big dataset, with lots of data dimensions. Several users have already investigated it and have created a bunch of charts.

In the 1st part of this study, you will complete 3 tasks. For each task, I’ll give you a sheet of paper that describes the task. After you read it, I’ll say “Go,” and you click the “Start Task 1” (or 2 or 3) label. When you’re done, tell me “I’m done,” so I know to stop my stopwatch, and you also click the “Stop Task 1” (or 2 or 3) label. Then, verbally give me your answer. After that, we do the same process for task 2 and task 3.

You can use this sheet of paper as a scratchpad if you want. If you are stuck, confused, or need help, you can ask me!

Task 1

What are the three most commonly-used chart types? This goes for all users in the system. So find out and tell me you’re done. Then click the “Stop Task 1” label and give me the answer in descending order: first, second, and third.

Task 2

Several charts use the *year* attribute, either as an axis, for something like color or size, etc. For the charts that encode *year*, which other attribute(s) is/are used most frequently with this one, besides Attack Count? That is, *year* + other dimension(s) both appear in the chart, ignoring the Attack Count attribute.

Task 3

Between the charts made by Thijs and Damiano, which of the two charts are the most similar based on chart encoding choices and the data attributes used?

10.3. Freeform Stage

After completing the Task Stage, each participant was handed a single sheet of paper with the following instructions printed at top:

Freeform Analysis Stage

This next stage is called the **freeform analysis stage**.

Let's suppose you are a data analyst who is the leader of a team of analysts investigating the terrorism dataset. (Info about terrorist attacks since 1970.) Luckily, your collaborators have already investigated the dataset and have created a large number of charts!

Your job as the team lead is to **review** and **investigate** the work done by your teammates. You are trying to identify the insights of your collaborators, and understand their themes of investigation.

How you do this (i.e., your strategy) is up to you!

As an example insight, in Task 1 you found the *most common chart type*. But there's also potentially other insights that can be made: *How many charts did a user create? What are the analysis patterns of users?* And so forth... The point is you want to investigate about what collaborators' work.

While you do this, you're going to use a technique called **think-aloud protocol**. Verbally speak whatever comes into your mind as you're using the system. This might include what you're looking at, what you're thinking, the action you're doing, what you're feeling, any new facts you're learning, etc. etc.

You have 10 minutes for this. You can also use your paper and pencil if you want. We'll tell you when time's up. Afterwards, I might ask you some follow up questions about what you found during your investigation. (No pressure! There's no wrong answers and you can't do anything wrong here!)

- Demographics Information.
 - Interface Style (options = [Collab, Kaggle])
 - Gender
 - Age
 - Major
 - I am very familiar with using charts to display data.
 - I was familiar with the Terrorism dataset (prior to this study).
 - I have created Kernels on Kaggle before. (options = [No, >5, 10-20, 20+])
- General System Impressions.

Overall, the system was:

 - easy to learn
 - easy to use
- Reflect on the types of insights the system provided.

I felt the system was really successful at:

 - Letting me review all of the charts
 - Letting me compare charts to each other
 - Showing the high-level trends and patterns of users
 - Organizing the charts into meaningful groupings
 - Showing the data attribute used by the charts
 - Showing the temporal ordering of charts
 - Encouraging me to think about user strategies for chart creation and analysis
- Please reflect on your investigation strategies.

My investigation was really helped by:

 - Looking at charts individually (ie, one by one)
 - Filtering the set of shown charts
 - Looking at chart creation times
 - Looking at keywords
 - Looking at chart encodings
 - Looking at the data attributes used by charts
 - (Constellations only) Looking at data attribute coverages directly
 - (Constellations only) Looking at chart clusterings
- (Constellations only) Interface Actions / Events.

I found it really useful during my investigation to:

 - Adjust the distance weight metrics for the MDS plot
 - Filter charts based on keywords, users, data attributes, and encodings
 - Filter charts based on creation times
 - Draw lassos to select a subset of the charts
 - View the progress lines of individual users
 - View charts in context of the cluster bubbles
 - Interact with the data coverage panel

10.4. Review Stage

A Google survey was used to ask the following set of questions to each participant following the Freeform Stage. Each “agreement” question (i.e., *I am very familiar with...*) was answered using a Likert Scale, with a range of 1 (strongly agree) to 7 (strongly disagree).