

SurvClusVis

Exploring survival data using tSNE

Lovedeep Gondara

[GitHub](#)

[Live demo](#)

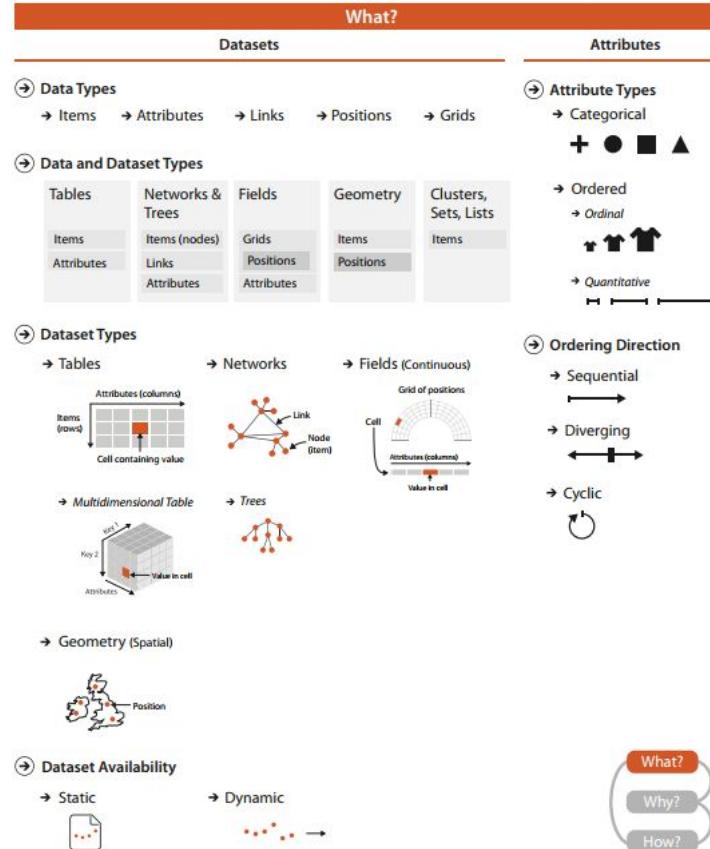
Motivation

- Physician researchers decide on a project based on datasets
- Current tools are limited for exploration
- Either expensive or too advanced for a naive audience
- An effective exploration can save money and time

What

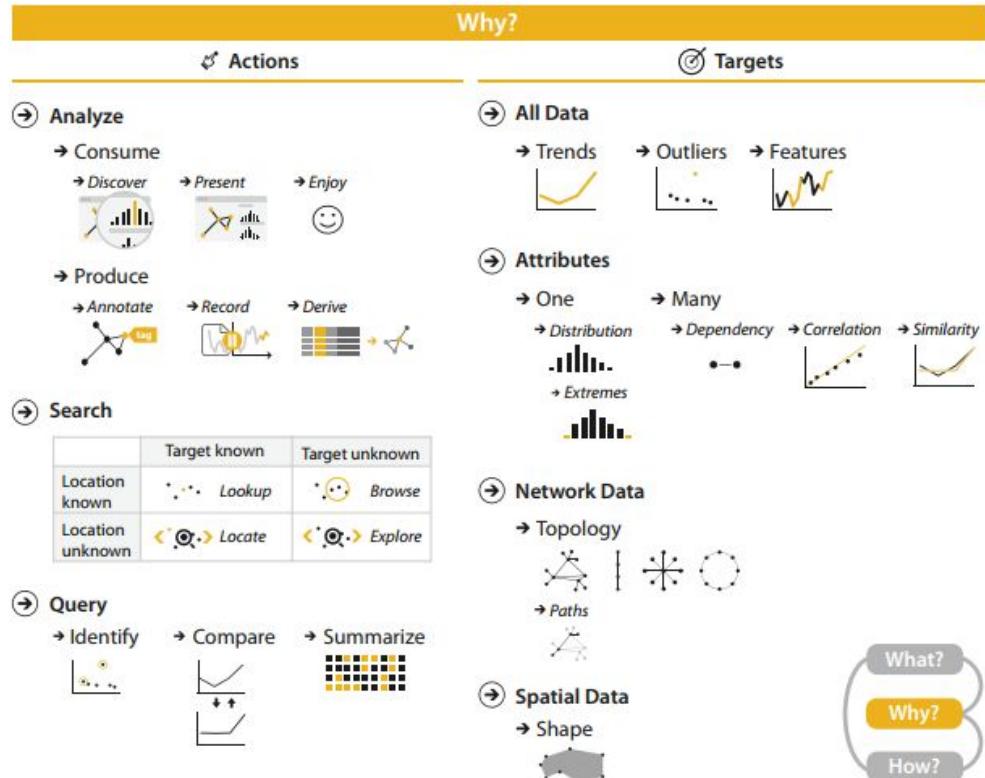
Survival datasets

- Time to event information
- Survival time
- Censoring information
- Other potential covariates



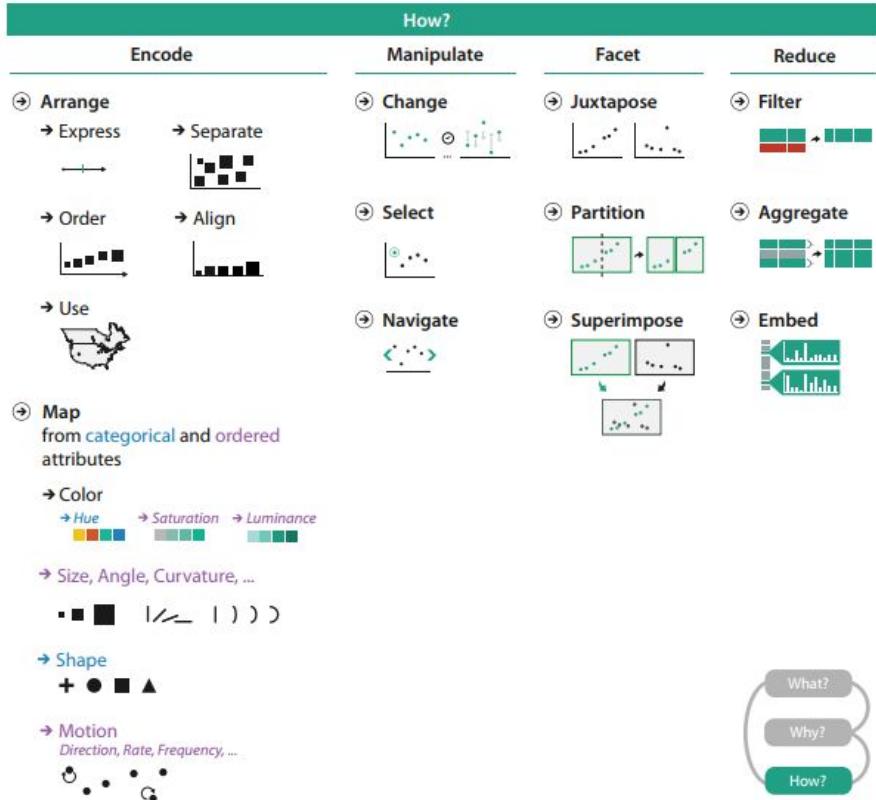
Why

- Current tools restricting
- Underlying structure is hidden
- Uncover low dim. manifold



How

- R, Shiny and D3
- Data → tSNE → Cluster → Explore
- [Online](#)
- Offline
- [Video](#)



Limitations

- Large datasets, problems with Barnes-Hut tSNE
 - Use other tSNE implementation
- Too much user input required
 - Automate

References

Visualization Analysis and Design, Tamara Munzner (A K Peters Visualization Series, CRC Press, 2014)

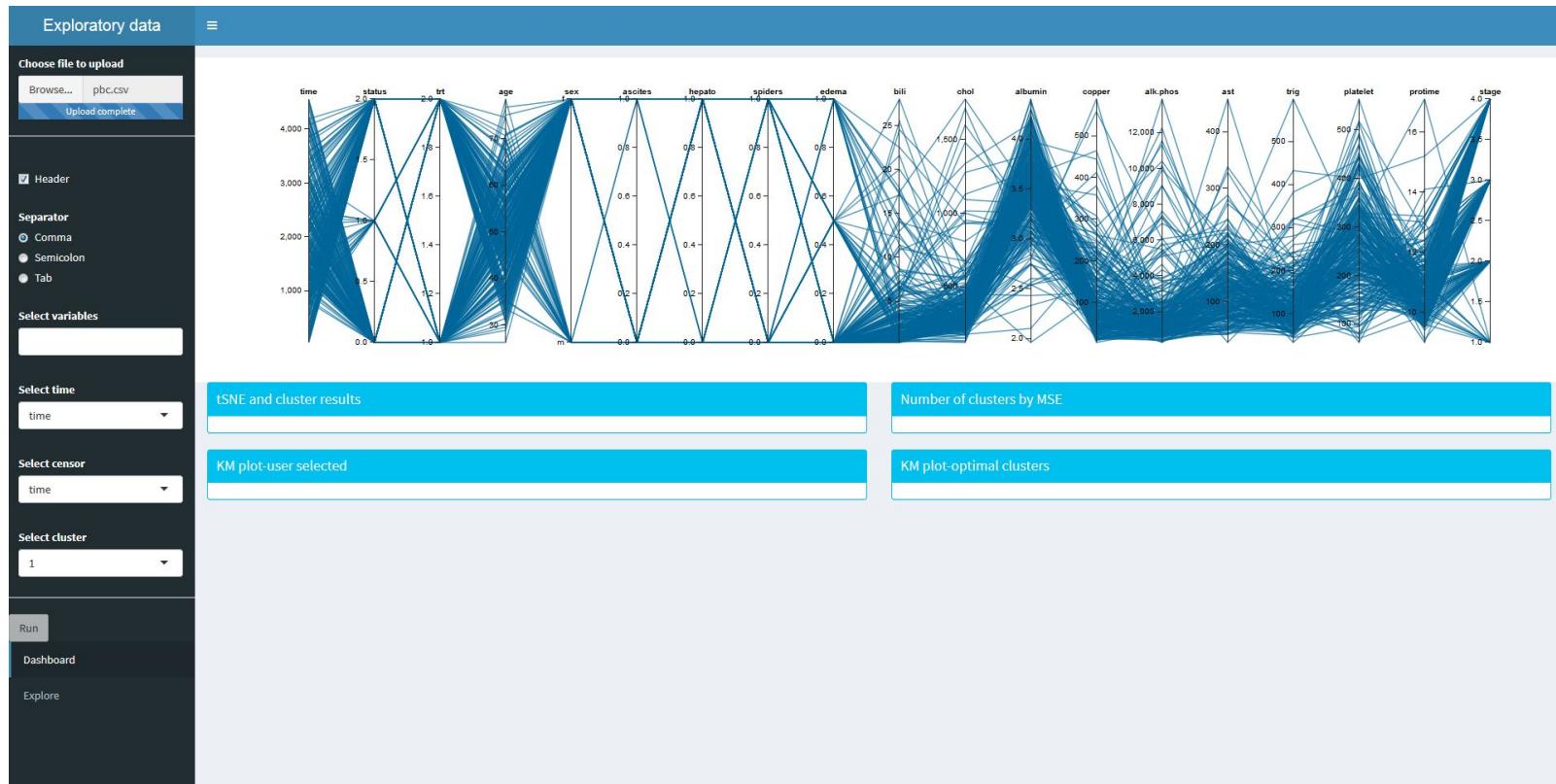
Maaten, Laurens Van Der, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of Machine Learning Research 9.Nov (2008): 2579-2605.

Screenshots- Start screen

The screenshot shows the start screen of an exploratory data analysis application. The interface is divided into several sections:

- Left Sidebar (Dark Theme):** Labeled "Exploratory data".
 - File Upload:** "Choose file to upload" with "Browse..." and "No file selected" buttons.
 - Data Settings:** "Header" checkbox (checked), "Separator" radio buttons for "Comma", "Semicolon", and "Tab", and a "Select cluster" dropdown set to "1".
 - Run Options:** "Run" button.
 - Navigation:** "Dashboard" and "Explore" buttons.
- Main Content Area:** A large, mostly empty white space.
- Plots and Results:** Four blue-tinted rectangular boxes arranged in a 2x2 grid.
 - tSNE and cluster results:** An empty plot area.
 - Number of clusters by MSE:** An empty plot area.
 - KM plot-user selected:** An empty plot area.
 - KM plot-optimal clusters:** An empty plot area.

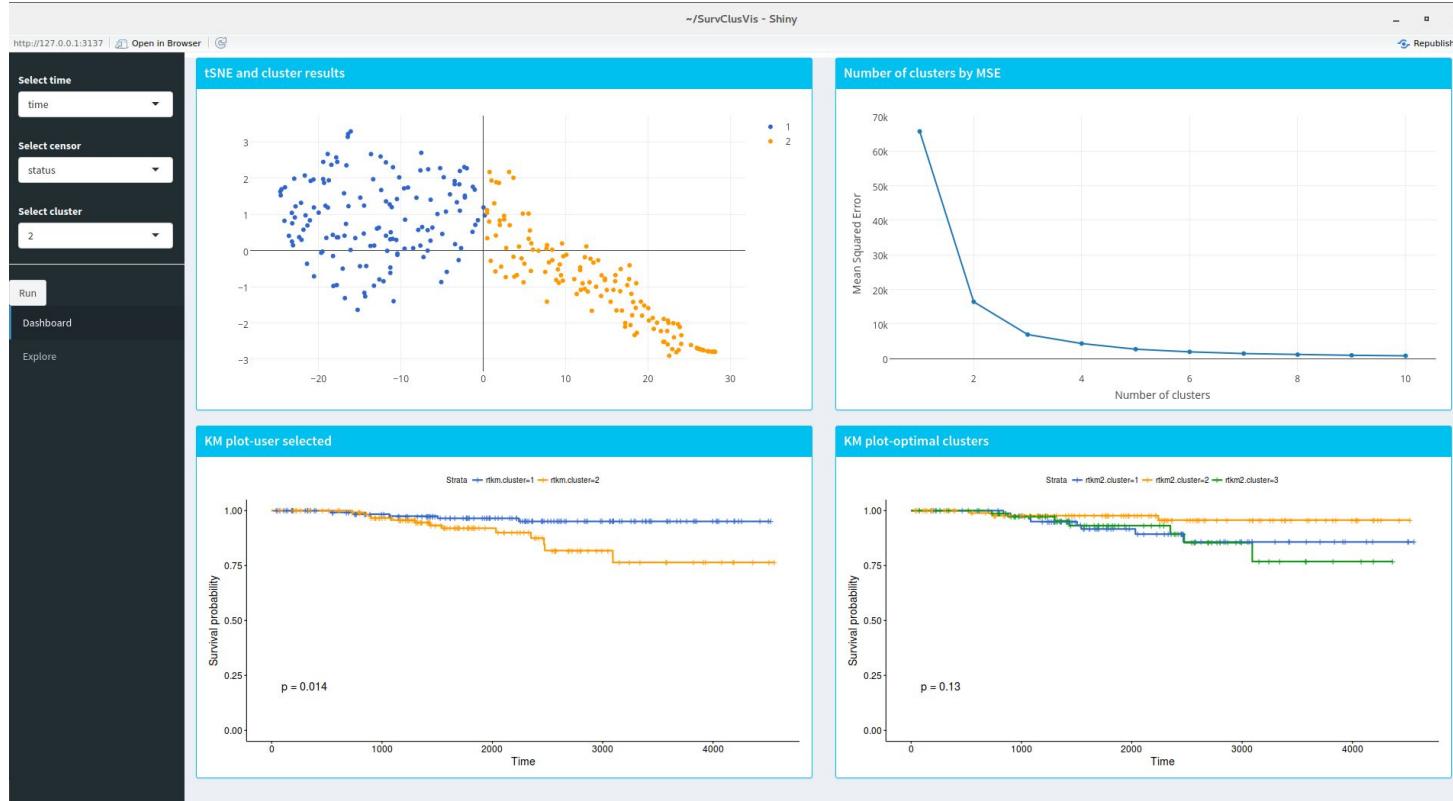
Screenshots - Initial variable distribution



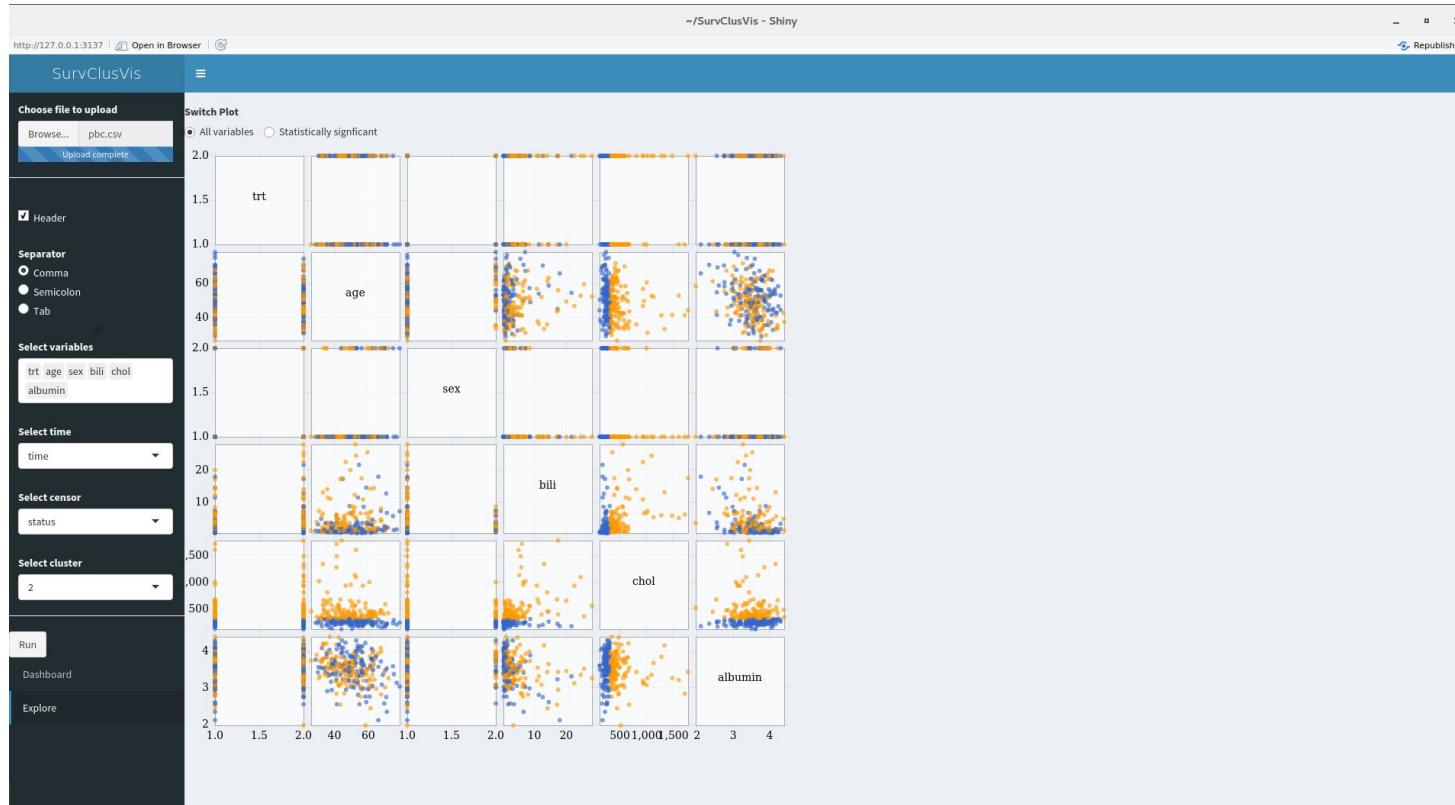
Screenshots - After selecting variables



Screenshots - Kaplan-Meier survival plots



Screenshots - Scatterplot matrix



Screenshots - Scatterplot matrix, significant vars

