

# SurvClusVis

Exploring survival data using tSNE

Lovedeep Gondara

[GitHub](#)  
[Live demo](#)

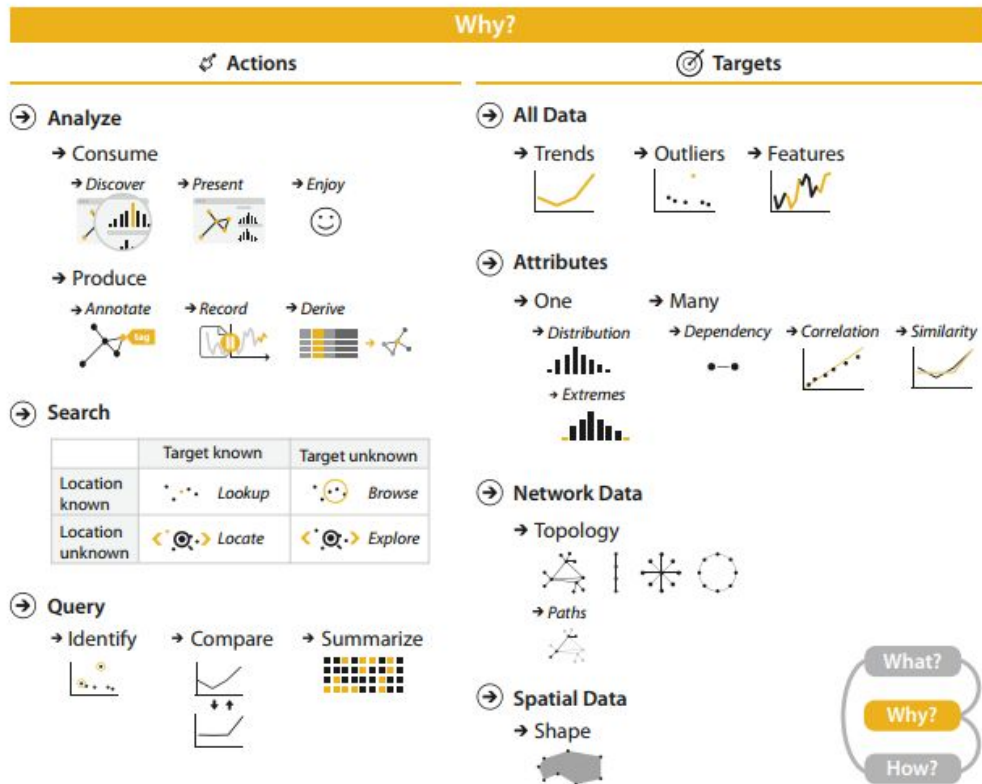
# Motivation

- 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



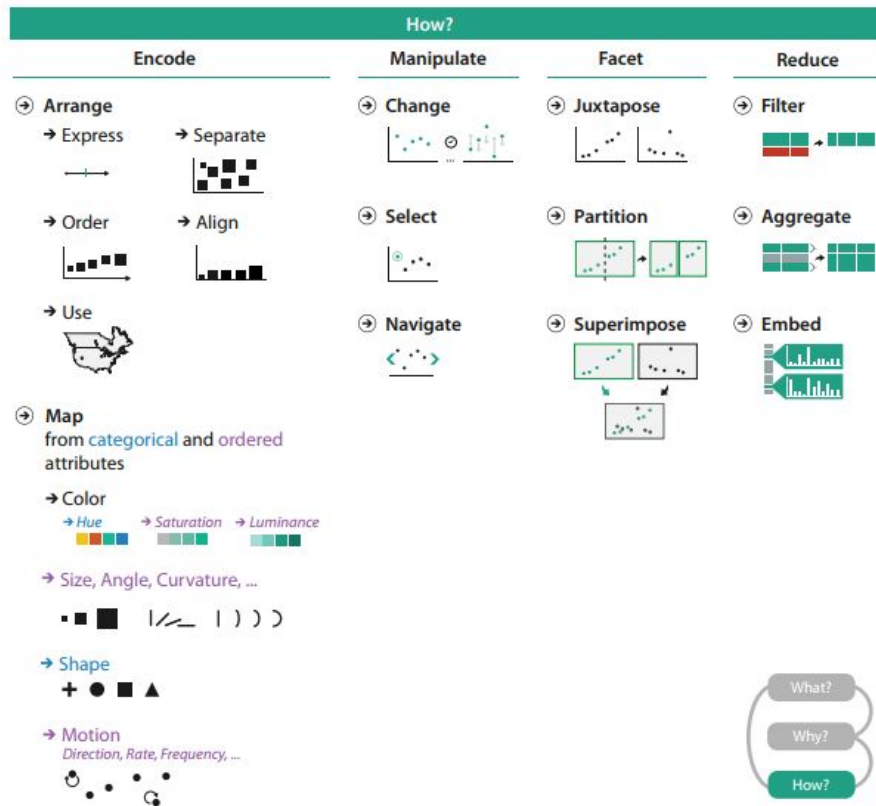
# Why

- 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

Exploratory data

Choose file to upload

Browse... No file selected

Header

Separator

Comma

Semicolon

Tab

Select cluster

1

Run

Dashboard

Explore

tSNE and cluster results

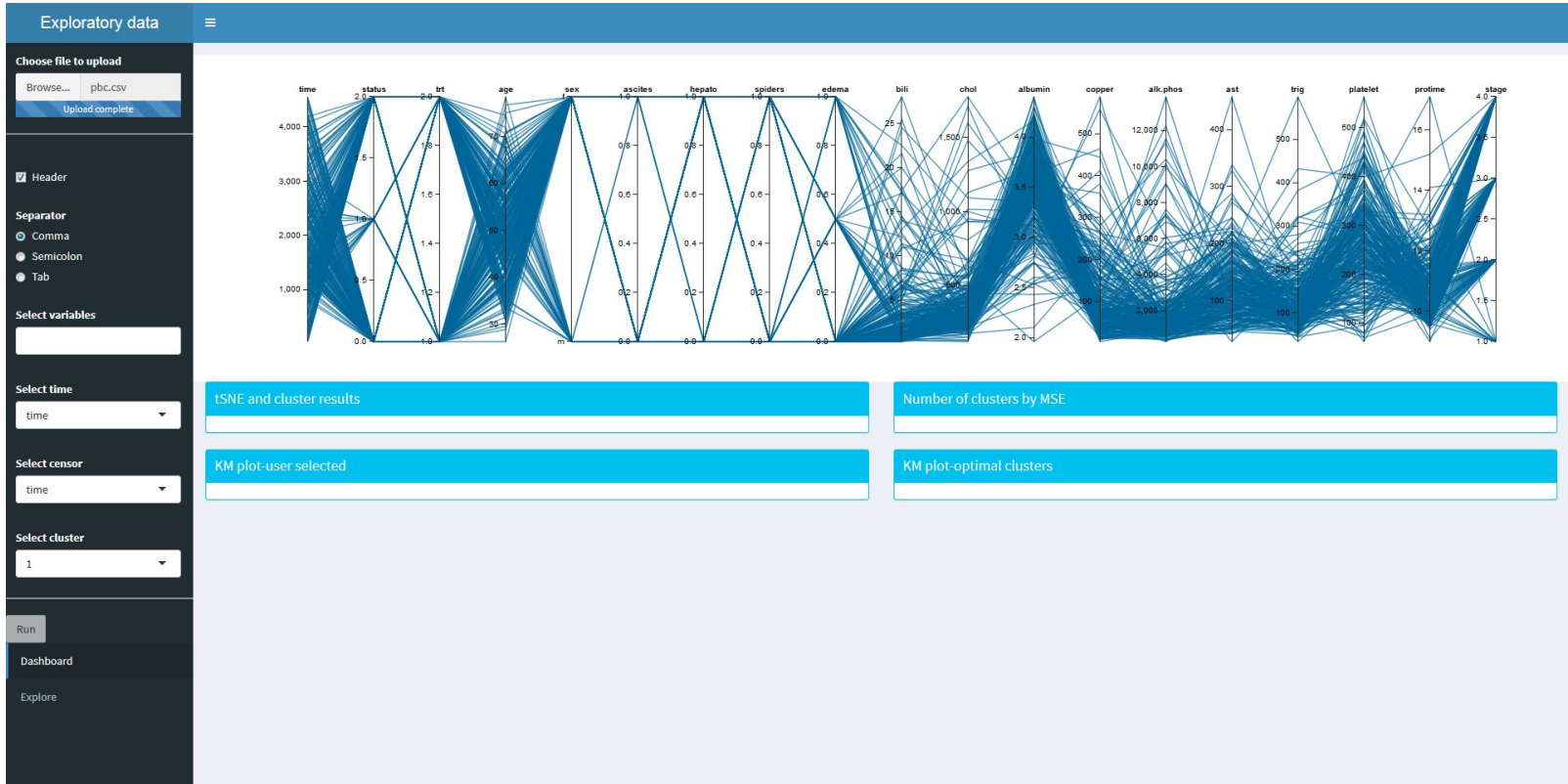
Number of clusters by MSE

KM plot-user selected

KM plot-optimal clusters



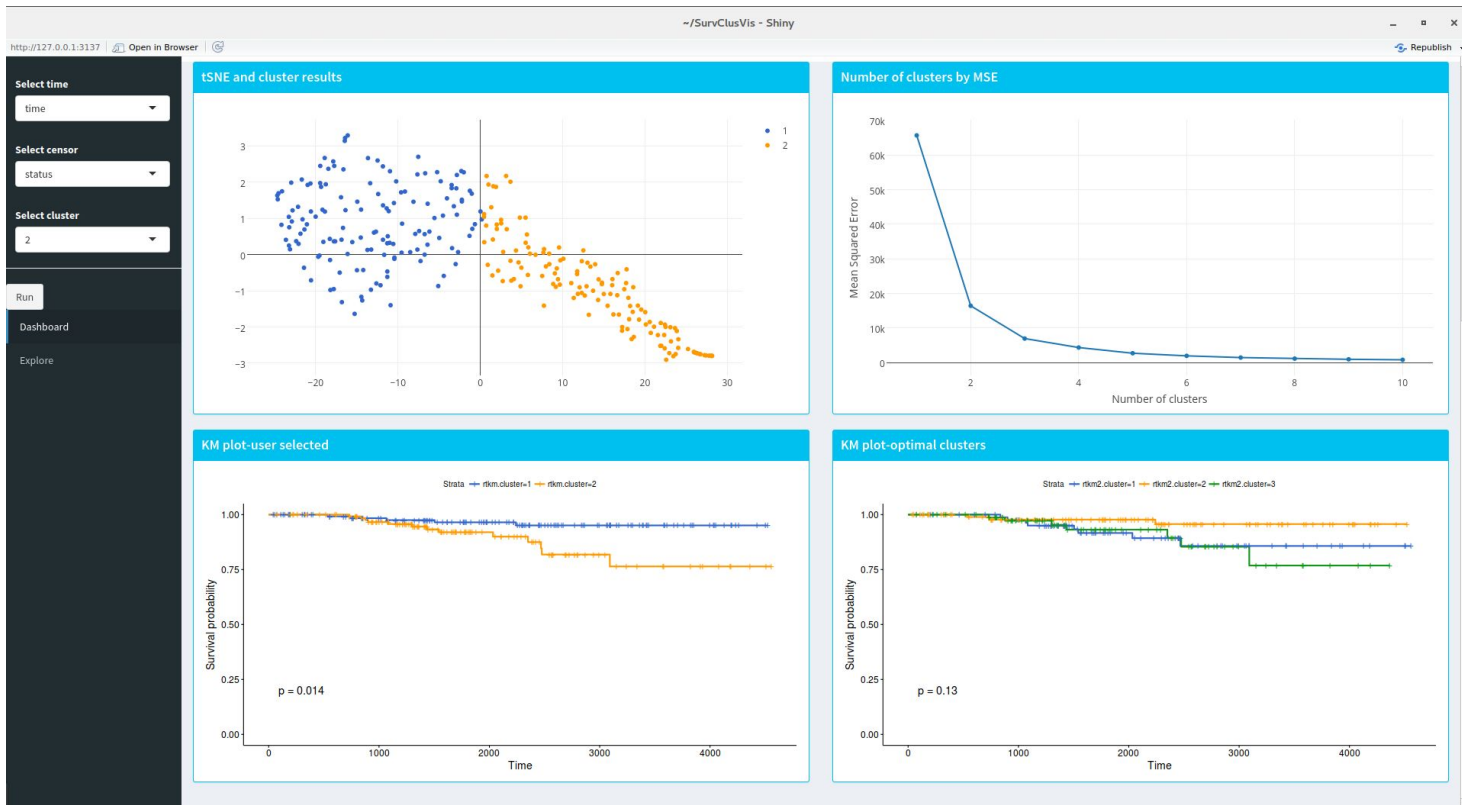
# Screenshots - Initial variable distribution



# Screenshots - After selecting variables



# Screenshots - Kaplan-Meier survival plots



# Screenshots - Scatterplot matrix



# Screenshots - Scatterplot matrix, significant vars

