

Novel Approaches for Visualising Time-to-Event Data

Victor Shvamsundar¹

Background and Problem

Visualising time-to-event data is crucial in fields like medicine and public health. However, traditional methods, such as Kaplan-Meier curves, have limitations in conveying complex patterns in survival data. This project explores novel techniques to enhance the clarity and interpretability of time-to-event visualisations.

Objectives of Project

- Identify challenges in visualising survival data.
- Develop alternative graphical methods to address these challenges.
- Test and evaluate these methods using real-world datasets.

Data Sources and Datasets

This study uses the **lung dataset** from the **survival** package:

- Lung Cancer Data¹**: Includes time-to-event data for patients with advanced lung cancer.

Variables:

- time**: Survival time in days.
- status**: Censoring indicator (1: death, 0: censored).
- sex**: Gender (1: male, 2: female).
- age**: Age of the patient.
- ph.ecog**: ECOG performance score (0 = good, 5 = poor).

#survival functions formulas for the poster

#Relationship between distribution F(t) and survival function S(t)

$$F(t) = P(T < t) = \int_0^t f(u)du$$

$$S(t) = P(T \geq t) = 1 - F(t)$$

#survival analysis include non parametric (Kaplan Maier estimate) and parametric functions #Life table estimates

$$S^*(t) = \prod_{i=1}^{j-1} \left(\frac{n'_i - d_i}{n'_i} \right)$$

Overall survival formula for non parametric method (KM)

$$\hat{S}(t) = \prod_{j=1}^k \left(\frac{n_j - d_j}{n_j} \right)$$

Early Results / Descriptive Statistics of Datasets

Kaplan-Meier Curves

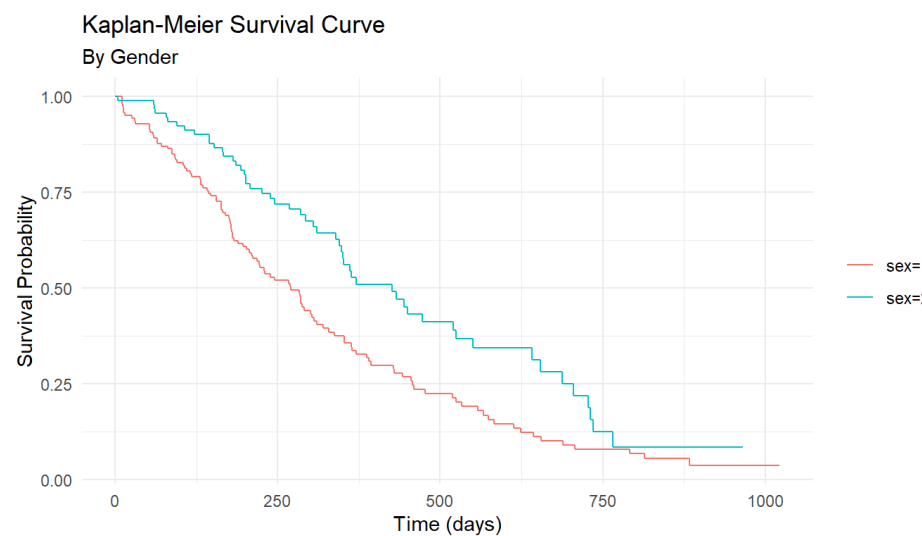


Figure 1: Kaplan-Meier curve for Lung Cancer dataset

Novel Visualisation: Interactive Plot

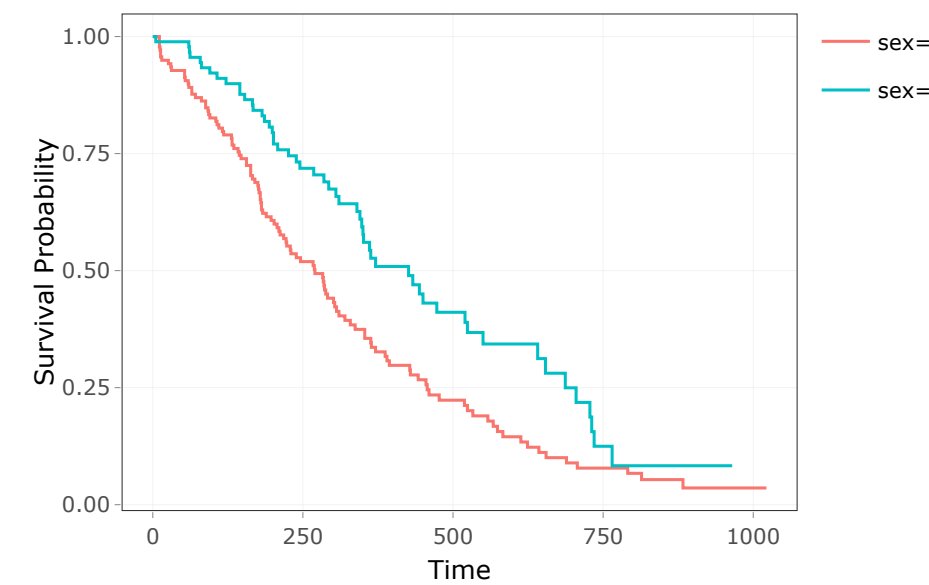


Figure 2: Interactive survival plot

Table Summary

Table 1: Summary of Survival by Gender								
	records	n.max	n.start	events	rmean	se(rmean)	median	0.95LCL
sex=1	138	138	138	112	326.0841	22.91156	270	212
sex=2	90	90	90	53	460.6473	34.68985	426	348

Next Project Steps

We will further investigate:

- Alternative visualisation methods like heatmaps and animations.
- Application of machine learning models to predict time-to-event outcomes.
- Evaluation of new methods with user studies to determine interpretability and utility.

GitHub

The code and datasets for this project can be viewed at our GitHub repository here: <https://github.com/>

References

- Therneau, T. (2015). A Package for Survival Analysis in R. Mayo Foundation.
- Kalbfleisch, J.D., & Prentice, R.L. (1980). The Statistical Analysis of Failure Time Data. Wiley.
- Murphy, S.A. et al. (2005). Randomized trial on intervention methods. doi: 10.1186/1468-6708-6-11

- Terry Therneau (2015). A Package for Survival Analysis in R↵