From Static to Interactive: A Comparative Evaluation of ggplot2 and Plotly in Clinical Data Visualisation

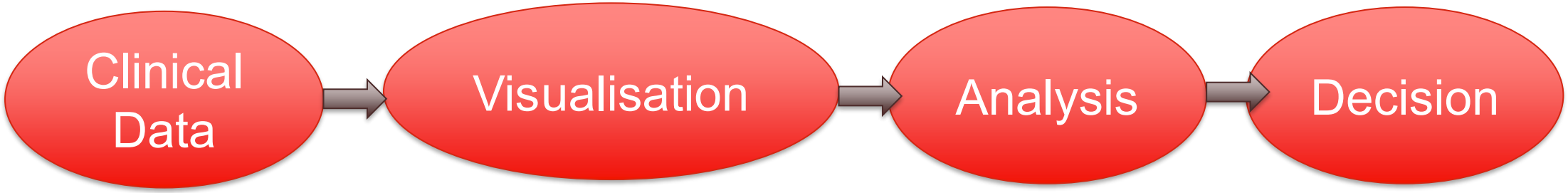
Mugdha Joshi
Senior Computational Statistician
Eli Lilly

"The information included in this presentation represents the opinion of the speaker, no way related to Eli Lilly and Company or its affiliates."

Why Data Visualisation Matters in Clinical Research

Clinical studies generate **complex, high-dimensional data**.

- Effective visualisation → better **interpretation, communication, and decision-making**.
- Regulatory, operational, and translational research contexts require **different levels of interactivity**.



R as a Visualisation Ecosystem

- R offers a **rich ecosystem**: ggplot2, Plotly, Shiny, lattice, highcharter, etc.
- ggplot2 and Plotly are **dominant for static vs interactive** paradigms.

ggplot	Plotly
Static Publication-ready	Interactive exploratory

ggplot2 Overview

- **ggplot2** is a powerful R visualization package built on the **Grammar of Graphics**, which treats plots as a combination of independent layers.
- Users map data to aesthetics (like x, y, color), add geometric objects (points, lines, bars), and customize scales, facets, and themes
- Strong integration with tidyverse
- layered approach makes ggplot2 especially strong for **complex, multi-variable plots**, **reproducible analysis**, and **high-quality static graphics** used in reports, papers, and presentations.

Plotly Overview

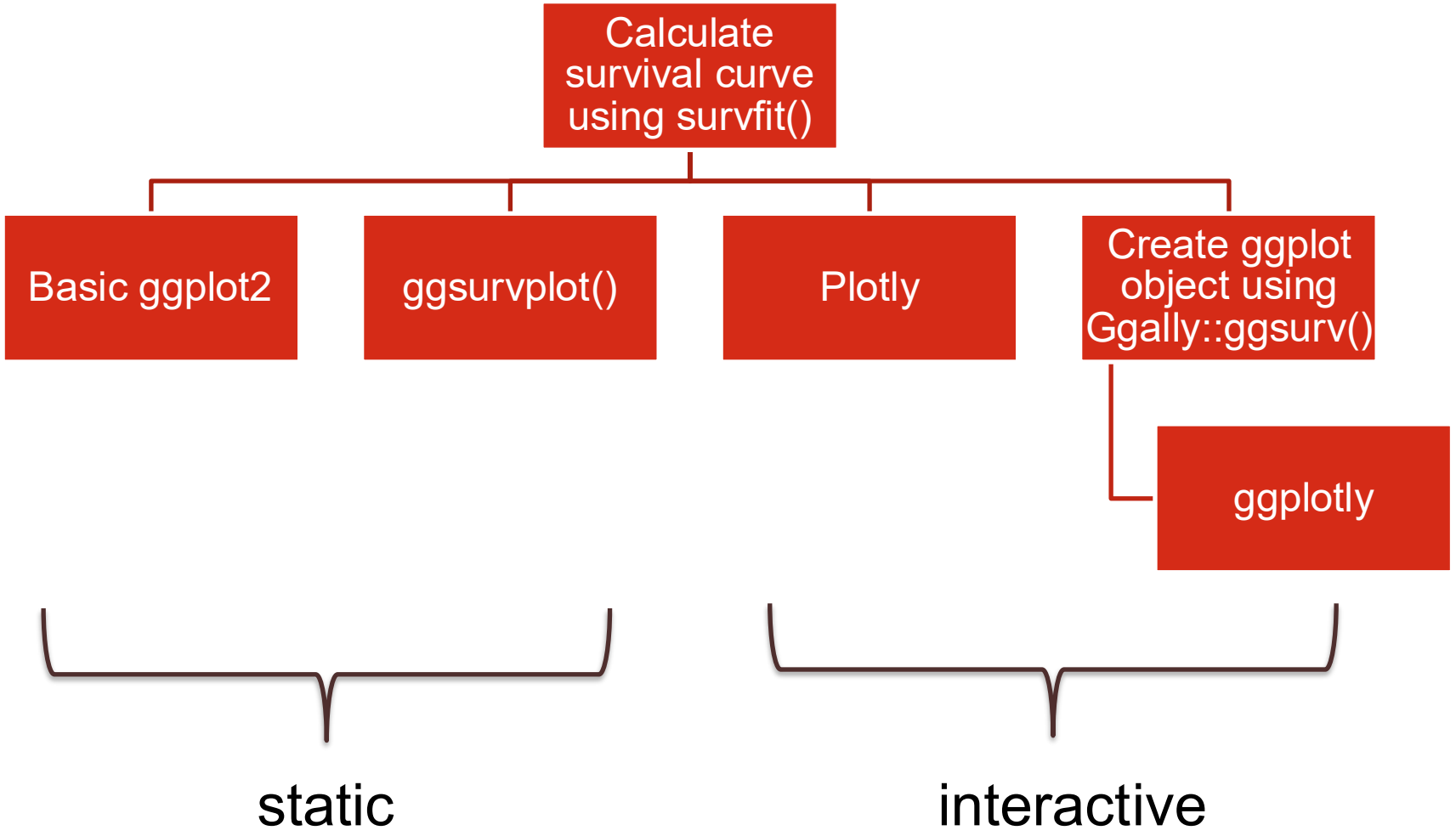
- **Plotly** is a data-visualization library that focuses on **interactive graphics**. It is an open-source package built from JavaScript, html and CSS
- It allows users to create plots with features like **hover tooltips, zooming, panning, and clickable legends**, making it especially useful for **exploratory analysis, dashboards, and web applications**

Case Study: Survival Analysis

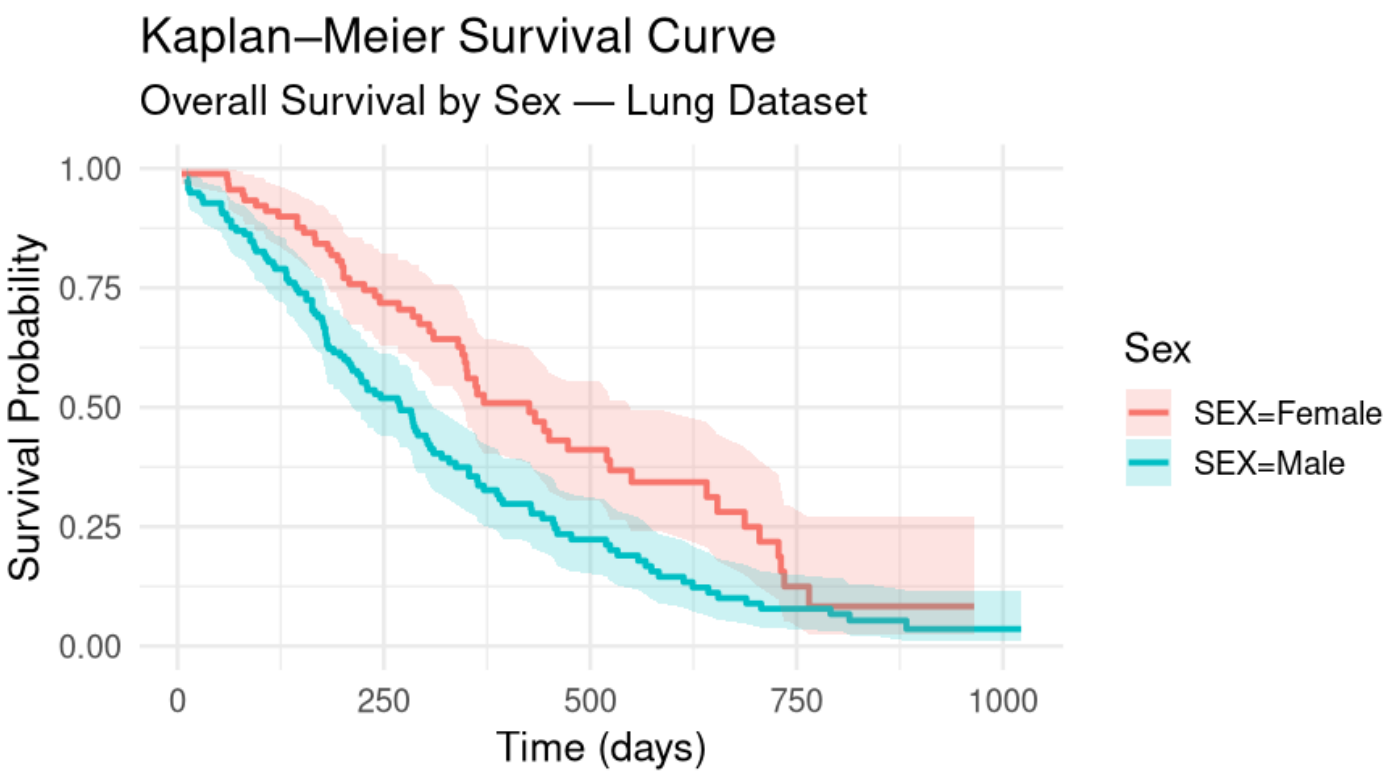
- **Survival analysis** consists of statistical methods that help us understand and predict how long it takes for an event to occur. These methods help researchers analyze time-to-event data, revealing not just if something happens, but when and how events unfold over time.

- We will use the lung dataset from the **survival** package as example data. The data contain subjects with advanced lung cancer from the North Central Cancer Treatment Group.

- We will visualise **Kaplan-Meier survival curves by gender** using ggplot2 and Plotly


Calculate survival curve using survfit()
Basic ggplot2
ggsurvplot()
Plotly
Create ggplot object using Ggally::ggsurv()
ggplotly
static
interactive

KM plot using ggplot2

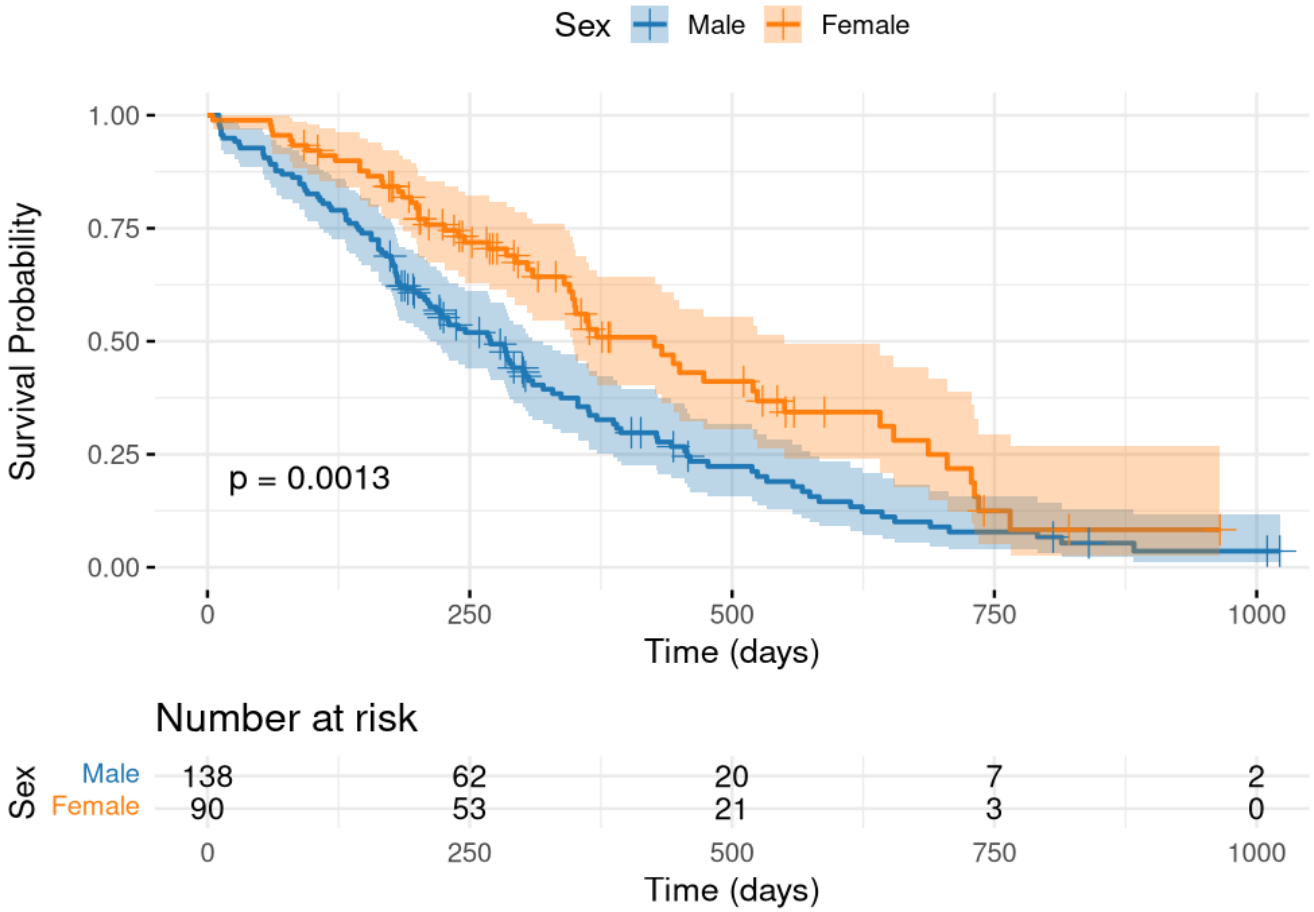


KM plot made using basic ggplot2

```r
ggplot(km_df, aes(time, surv, color = strata)) +
  geom_step(size = 1) +
  geom_ribbon(
    aes(ymin = lower, ymax = upper, fill = strata),
    alpha = 0.2,
    color = NA
  ) +
  labs(
    title = "Kaplan-Meier Survival Curve",
    subtitle = "Overall Survival by Sex – Lung Dataset",
    x = "Time (days)",
    y = "Survival Probability",
    color = "Sex",
    fill = "Sex"
  ) +
  scale_y_continuous(limits = c(0, 1)) +
  theme_minimal(base_size = 14)
```

ggsurvplot()

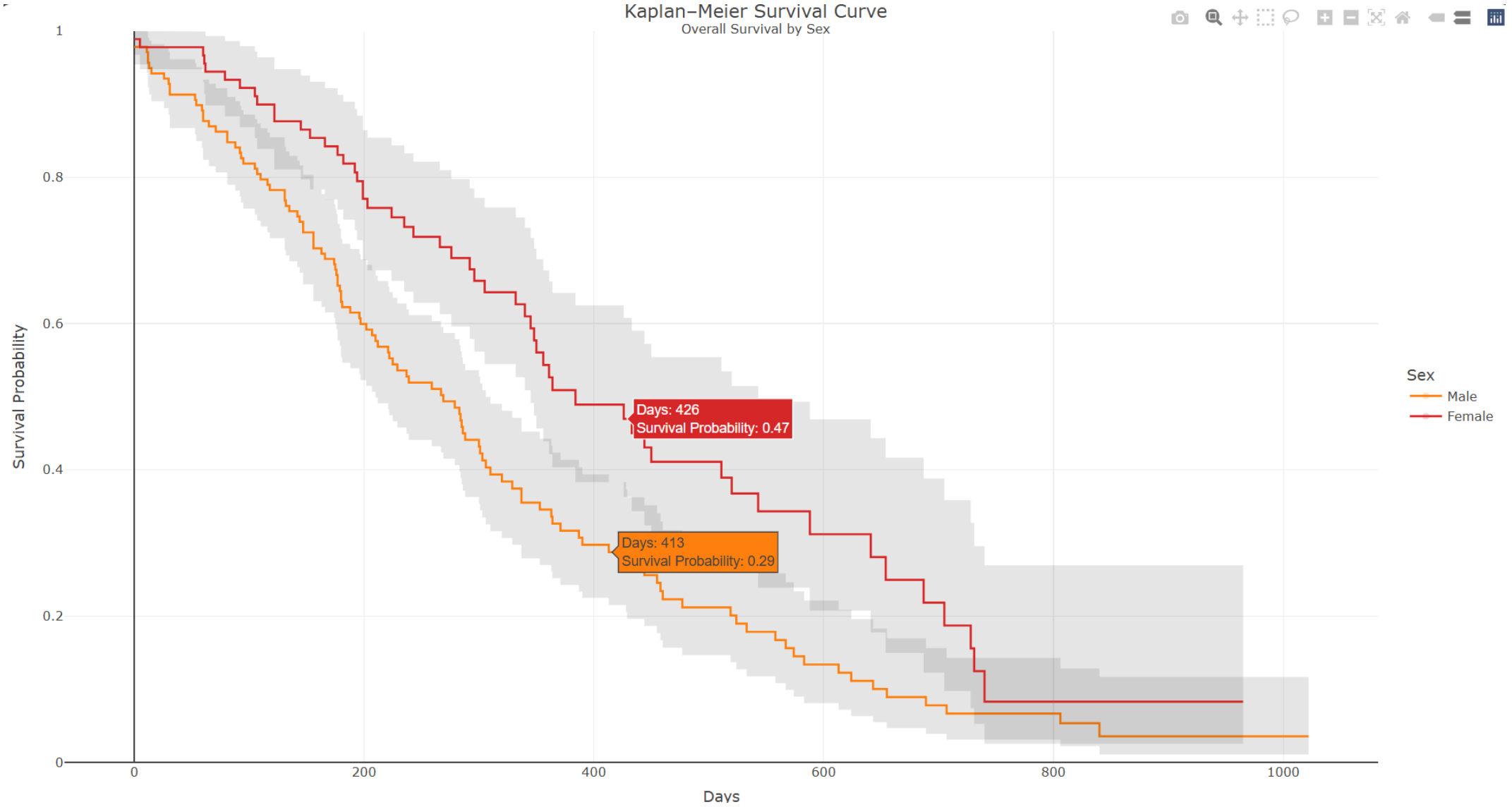
- Full-featured, high-level function of survminer package using ggplot capabilities
- Automatically adds: risk table, p-values, censoring ticks, titles, legends.
- Produces a **list of ggplot objects** (plot + risk table)
- Best for **exploratory work or quick clinical reports**

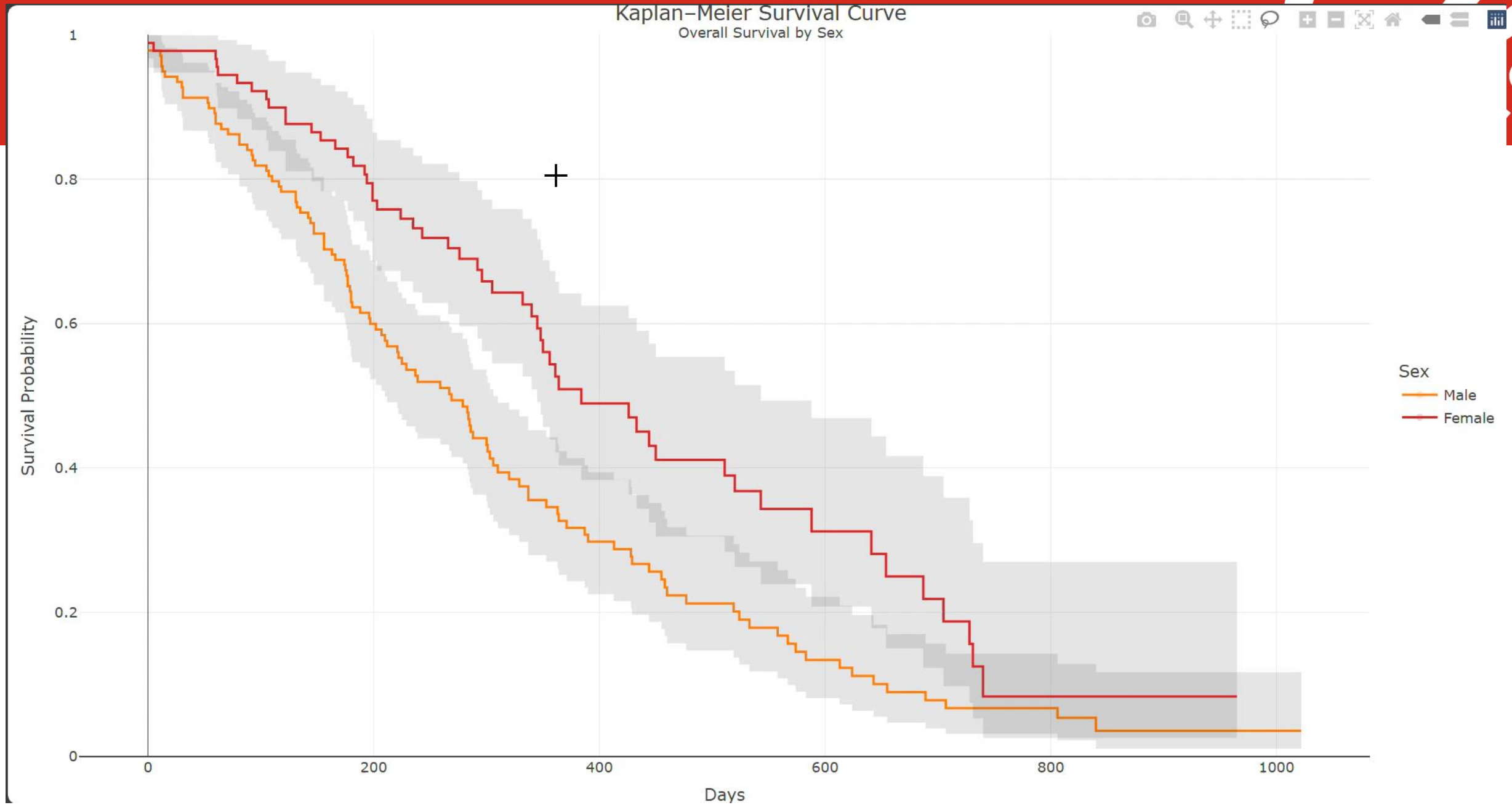



```r
ggsurv <- ggsurvplot(
  km_fit,
  data = adtte,
  risk.table = TRUE,          # Show risk table
  conf.int = TRUE,            # Show confidence interval
  pval = TRUE,                # Add log-rank p-value
  censor.shape = 3,           # Censoring tick shape
  censor.size = 4,            # Censor tick size
  legend.title = "Sex",
  legend.labs = levels(adtte$SEX),
  xlab = "Time (days)",
  ylab = "Survival Probability",
  palette = c("#1f77b4", "#ff7f0e"), # Color for Male/Female
  ggtheme = theme_minimal(base_size = 14)
)
```

KM plot made using ggsurvplot

KM plot using Plotly




Kaplan–Meier Survival Curve
Overall Survival by Sex
Survival Probability
1
0.8
0.6
0.4
0.2
0
0
200
400
600
800
1000
Days
Sex
Male
Female

ggsurv() and ggplotly()

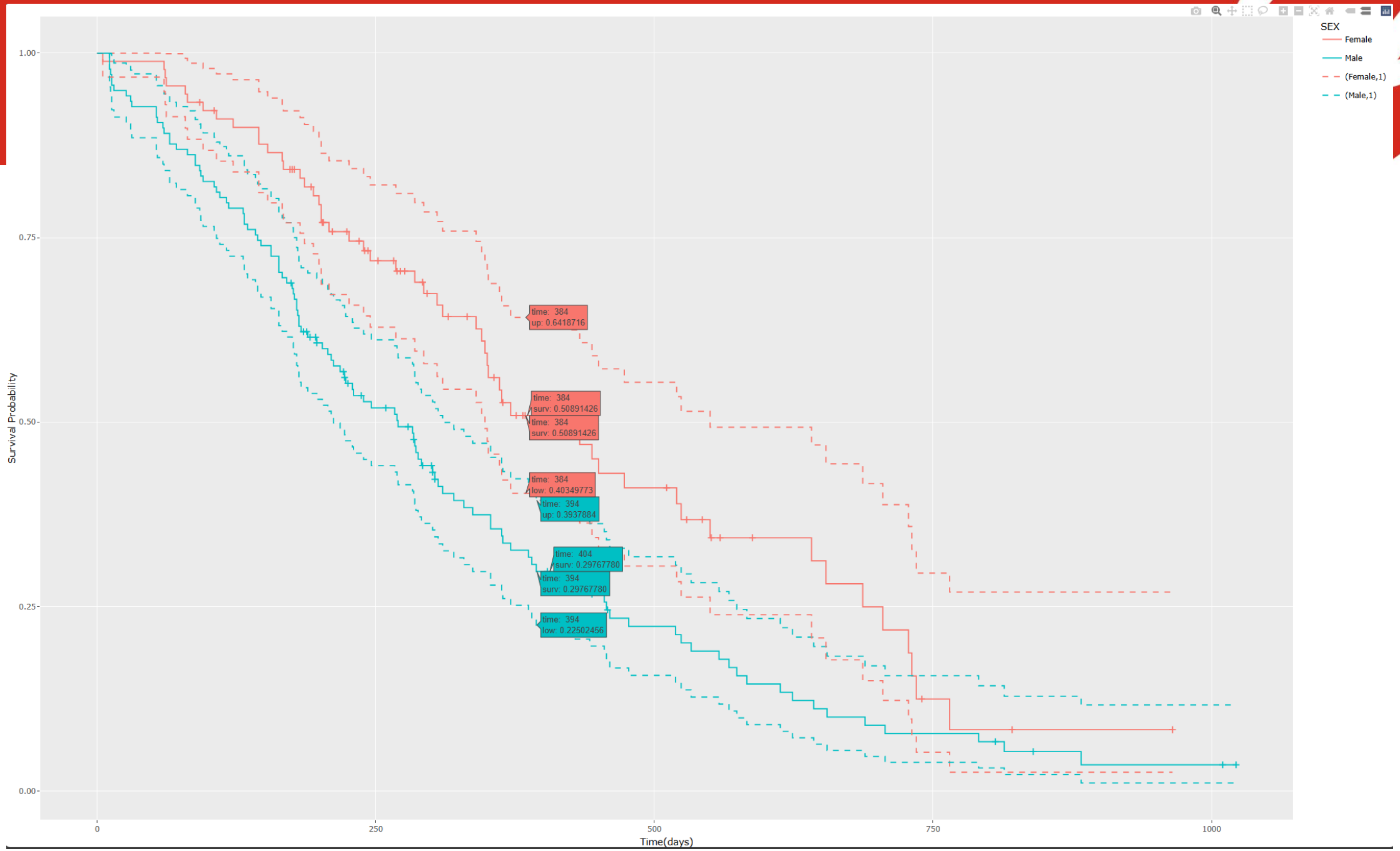
- ggsurv() is part of the Ggally package and produces Kaplan-Meier plots using **ggplot2**. As a first argument it needs a survfit object, created by the survival package
- Essentially, ggsurv() is a **wrapper that returns a raw ggplot object** of the Kaplan–Meier curve.
- You can think of it as **“ggplot-ready survival data”**: it produces a **clean ggplot object of survival steps**, and you can **modify it like any ggplot2 plot**.
- ggplotly() is part of the plotly package. ggplotly() is a function from the **plotly** package in R that converts a static **ggplot2** visualization into an interactive Plotly graph. It enables features like zooming, panning, and hover tooltips without changing the original ggplot code.

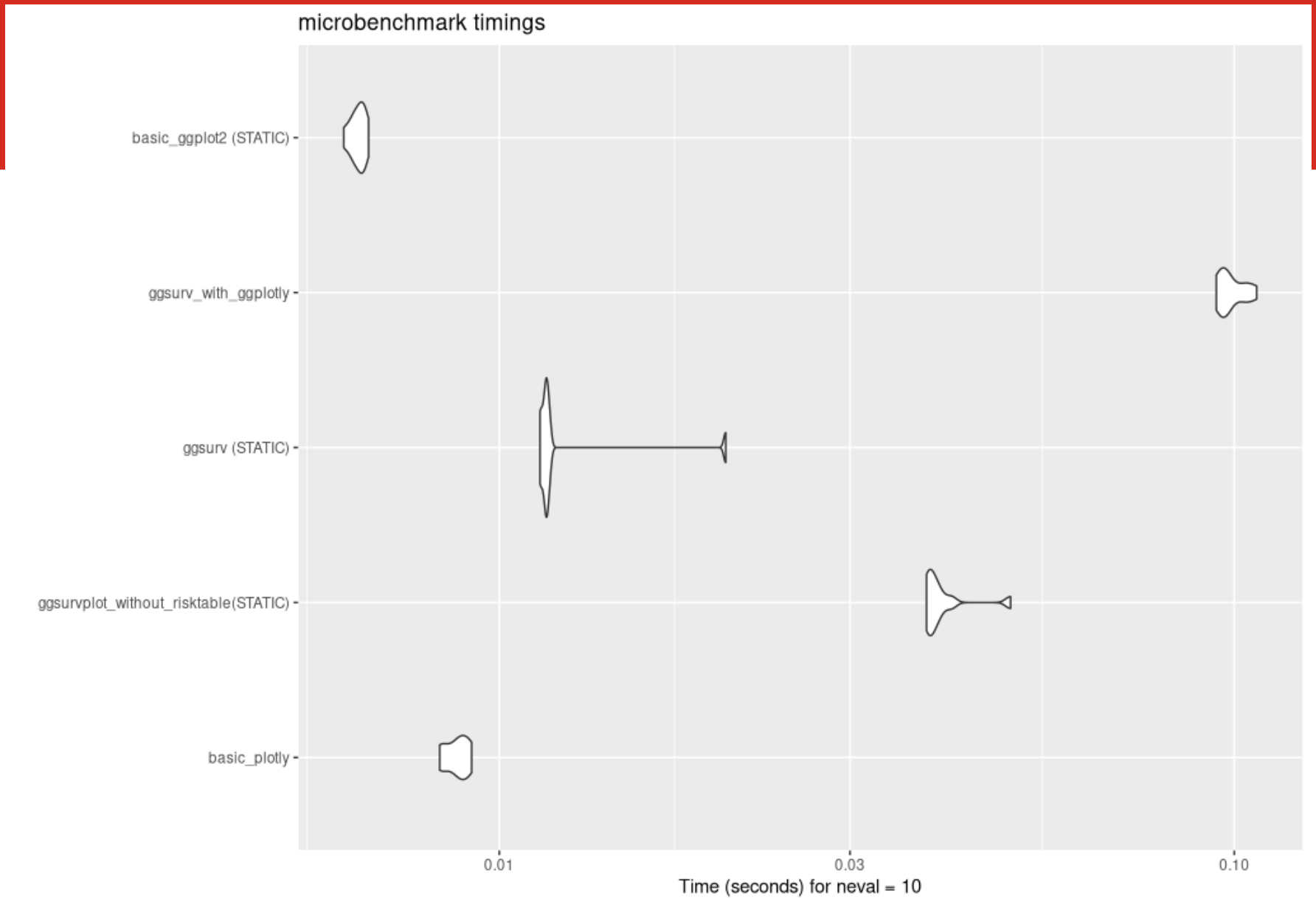
```r
km_fit <- survfit(Surv(AVAL, CNSR == 0) ~ SEX, data = adtte)

# ---- 3. Generate ggplot2 object using ggsurv ----

gg_km <- GGally::ggsurv(s=km_fit,
                CI = TRUE,          # Show confidence intervals
                lty.est = 1,        # Line type
                cens.size = 2,
               xlab="Time(days)",
               ylab="Survival Probability")# Size of censoring ticks
               # censor.shape = 3)    # Shape of censoring ticks

# ---- 4. Convert to interactive Plotly plot ----
plotly_km <- ggplotly(gg_km, tooltip = c("x", "y"))
```


SEX
Female
Male
(Female,1)
(Male,1)
Survival Probability
1.00
0.75
0.50
0.25
0.00
0
250
500
750
1000
Time(days)
time: 384
up: 0.6418716
time: 384
surv: 0.50891426
time: 384
surv: 0.50891426
time: 384
low: 0.40349773
time: 394
up: 0.3937884
time: 404
surv: 0.29767780
time: 394
surv: 0.29767780
time: 394
low: 0.22502456



Runtime calculated for different methods using microbenchmark

Static vs Interactive: What's the Difference?

Criteria	ggplot2	Plotly
Usability	Widely known in R community	Intuitive, but JS-oriented
Flexibility	Very flexible via extensions	Strong for dashboards
Customization	Themes, layers, aesthetics	Layout, hover, and callbacks
Interactivity	None (static)	Rich (zoom, hover, filters)
Performance	Lightweight	Slightly heavier on large data

Practical Considerations

Scenario	Recommended Tool
Regulatory report	ggplot2
Peer-reviewed publication	ggplot2
Exploratory analysis	Plotly
Interactive dashboards	Plotly
Mixed static/interactive	ggplotly() hybrid

Alternatives methods

- **Esquisse** is an R package that offers a **GUI (drag-and-drop) interface** for building **ggplot2** plots. It's mainly used for **exploratory analysis and learning**, because it lets users interactively design plots and then exports clean, reproducible ggplot code.
- **Lattice** is an older R visualization system designed for **trellis (multi-panel) graphics**. It excels at displaying relationships across many groups using consistent scales and layouts, but it is **less flexible and less customizable** than ggplot2 and has limited interactivity.

Recommendations

Summary:

- **ggplot2** → best for reproducibility and compliance.
- **Plotly** → best for communication and dynamic exploration.
- **Combination** of both can serve full project lifecycle: from analysis → exploration → presentation.

References

- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*
- Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny*
- Emily C. Zabor, "Survival Analysis in R". (https://www.emilyzabor.com/tutorials/survival_analysis_in_r_tutorial.html)
- Survminer cheatsheet (https://rpkgs.datanovia.com/survminer/survminer_cheatsheet.pdf)
- Carson Sievert, 2019, “Interactive web-based data visualization with R, plotly, and shiny”. Available at https://plotly-r.com/index.html
- FDA guidance on data visualisation for clinical review reports

THANK YOU!

Mugdha Joshi
Eli Lilly
1st Floor, Building Primrose(7B) Wing B. Embassy Tech Village, Outer Ring Road, Devarabisanahalli, Bengaluru-560103
Email: joshi_mugdha@lilly.com