```r
install.packages("BiocManager")
install.packages("survival")
install.packages("survminer")
BiocManager::install("SummarizedExperiment")
library(survival)
library(survminer)
library(SummarizedExperiment)

# Fit survival curves
fit <- survfit(Surv(time, status) ~ sex, data = lung)

# Basic survival curves
ggsurvplot(fit, data = lung)

# Customized survival curves
ggsurvplot(fit, data = lung,
           surv.median.line = "hv", # Add medians survival

           # Change legends: title & labels
           legend.title = "Sex",
           legend.labs = c("Male", "Female"),
           # Add p-value and tervals
           pval = TRUE,

           conf.int = TRUE,
           # Add risk table
           risk.table = TRUE,
           tables.height = 0.2,
           tables.theme = theme_cleantable(),

           # Color palettes. Use custom color: c("#E7B800", "#2E9FDF"),
           # or brewer color (e.g.: "Dark2"), or ggsci color (e.g.: "jco")
           palette = c("#E7B800", "#2E9FDF"),
           ggtheme = theme_bw() # Change ggplot2 theme
)
```
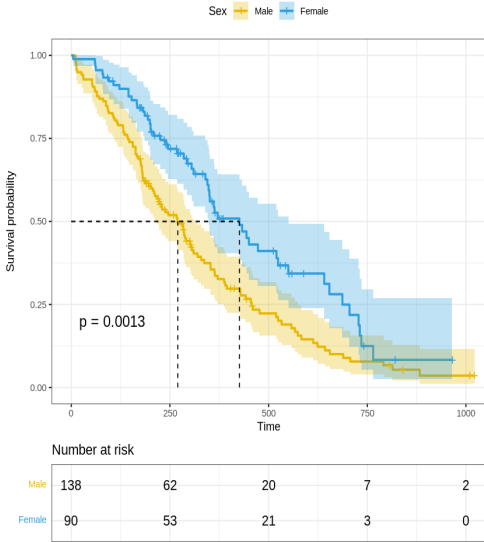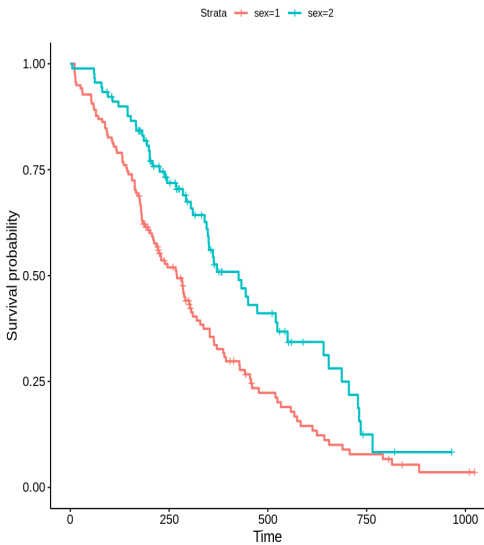
Arquivos ✕   •••

📤 🗁 △ 🚫

⬆ ..
▾ 📁 data
       📄 tcgaLIHCdata_preprocessed.RData
▾ 📁 sample_data
       📄 README.md
       📄 anscombe.json
       📄 california_housing_test.csv
       📄 california_housing_train.csv
       📄 mnist_test.csv
       📄 mnist_train_small.csv
       📄 tcgaLIHCdata_preprocessed.RData
   📄 tcgaLIHCdata_preprocessed.RData

rowMedians

The following objects are masked from 'package:m

anyMissing, rowMedians

```
load(file = "./tcgaLIHCdata_preprocessed.RData")
class(tcgaLIHCdata)
```

> 'RangedSummarizedExperiment'

```
dim(tcgaLIHCdata)
```

> 29885 · 371

```
gexp <- assay(tcgaLIHCdata)
rowAnnotation <- rowData(tcgaLIHCdata)
colAnnotation <- colData(tcgaLIHCdata)
```

```
class(gexp)
```

> 'matrix' · 'array'

```
class(rowAnnotation)
```

> 'DFrame'

```
names(colAnnotation)
```

> 'bcr_patient_barcode' · 'gender' · 'Age' ·
> 'ajcc_pathologic_tumor_stage' · 'OS' · 'OS.time' · 'PFI' ·
> 'PFI.time' · 'OS.time.months' · 'PFI.time.months' ·

```
colAnnotation$ajcc_stage <- colAnnotation$ajcc_pathologic_tumor_stage
idx <- colAnnotation$ajcc_stage%in%c("Stage IIIA","Stage IIIB","Stage IIIC")
colAnnotation$ajcc_stage[idx] <- "Stage III"
idx <- colAnnotation$ajcc_stage%in%c("Stage IVA","Stage IVB")
colAnnotation$ajcc_stage[idx] <- "Stage IV"
idx <- colAnnotation$ajcc_stage%in%c("[Discrepancy]")
colAnnotation$ajcc_stage[idx] <- NA
table(colAnnotation$ajcc_stage, useNA="ifany")
```

```
     Stage I  Stage II Stage III  Stage IV      <NA>
         171        86        85         5        24
```

```
fit <- survfit(Surv(PFI.time.months, OS) ~ ajcc_stage, data = colAnnotation)
ggsurvplot(fit, data=colAnnotation, legend="right")
```
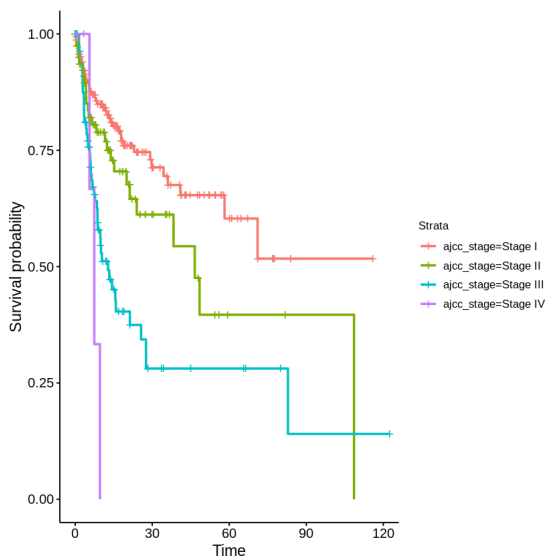


```
fit <- coxph(Surv(OS.time, OS) ~ ajcc_stage, data = colAnnotation)
fit
```