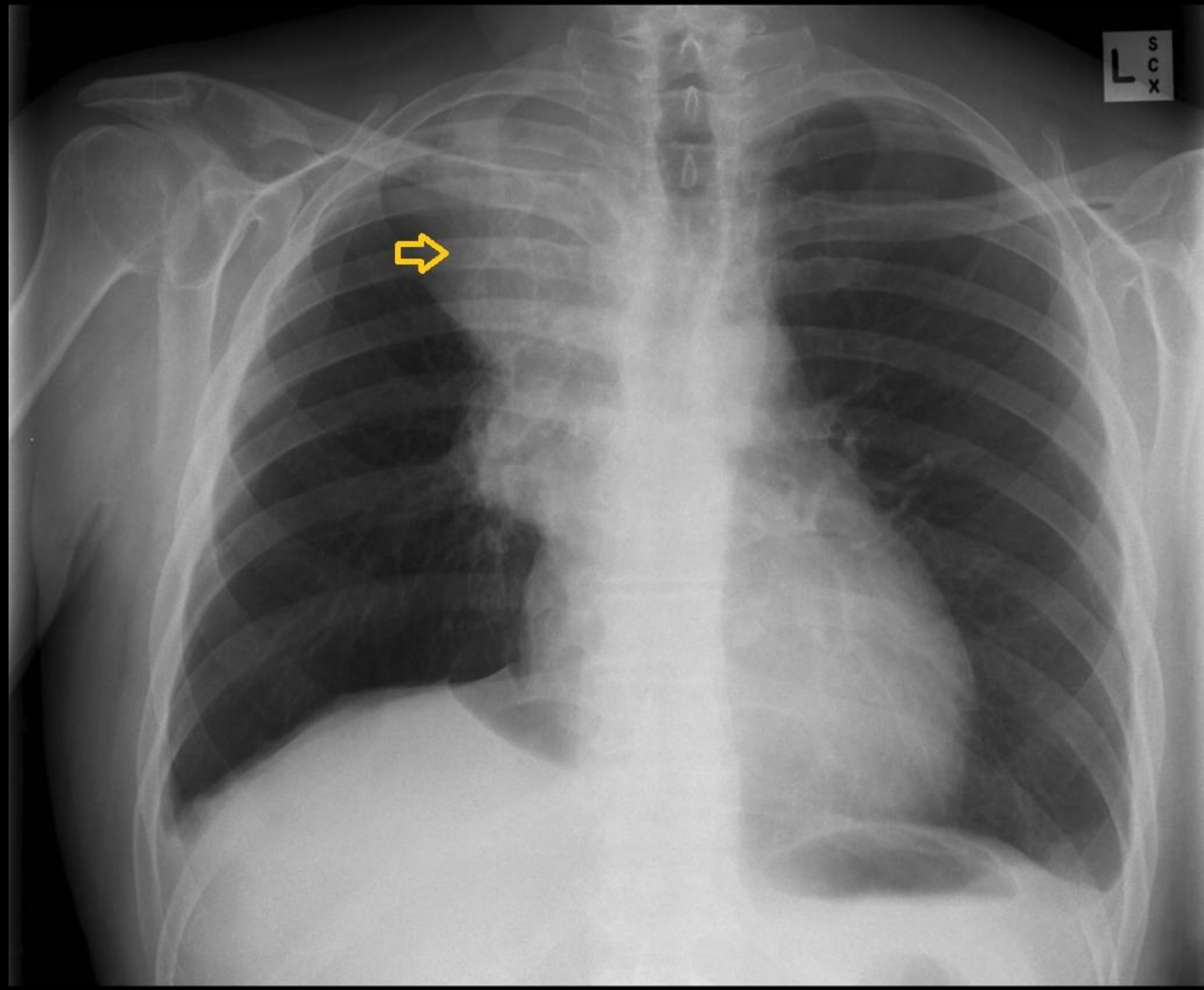
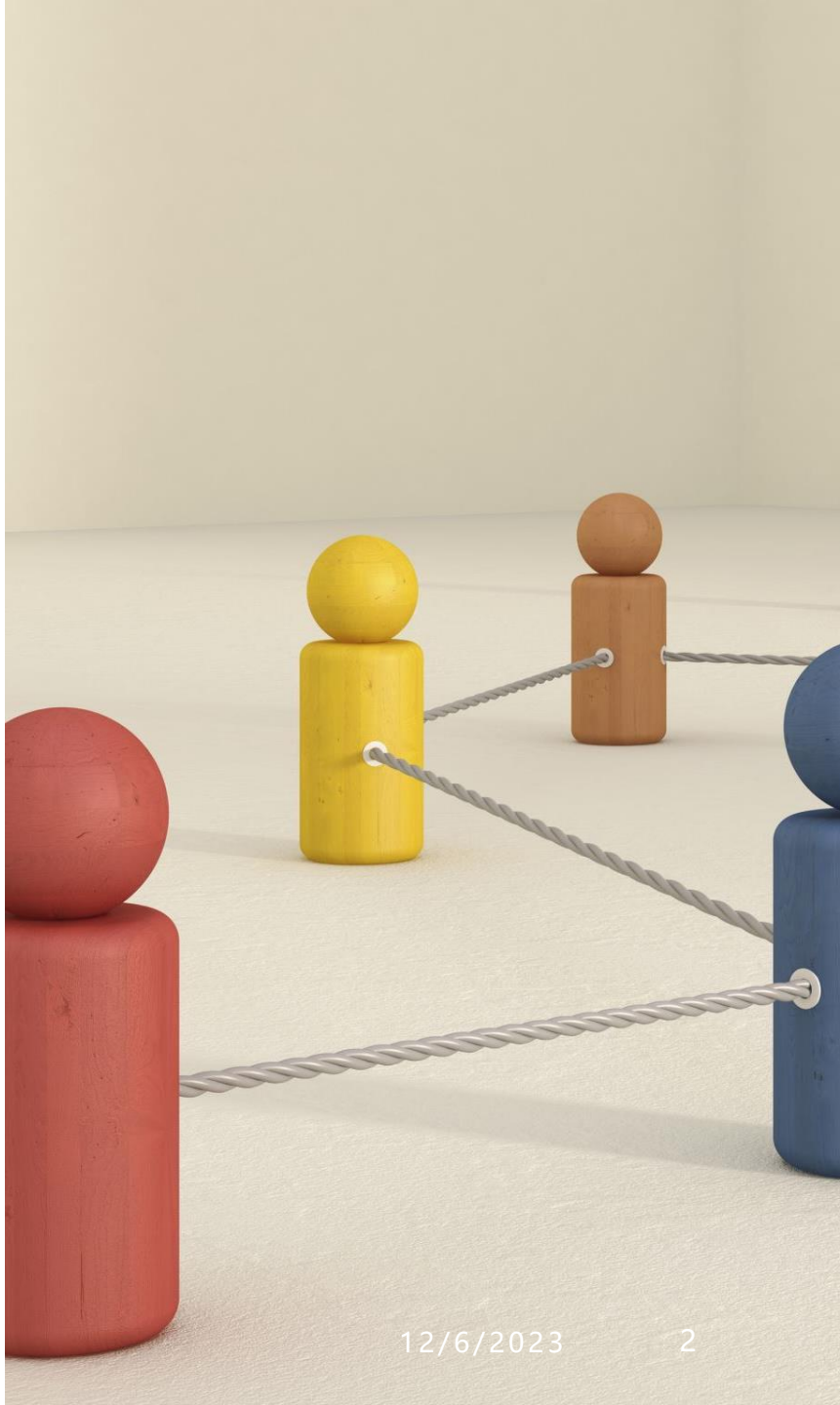
# Lung Cancer

Aryan Rezanezhad

Jiachen Pan

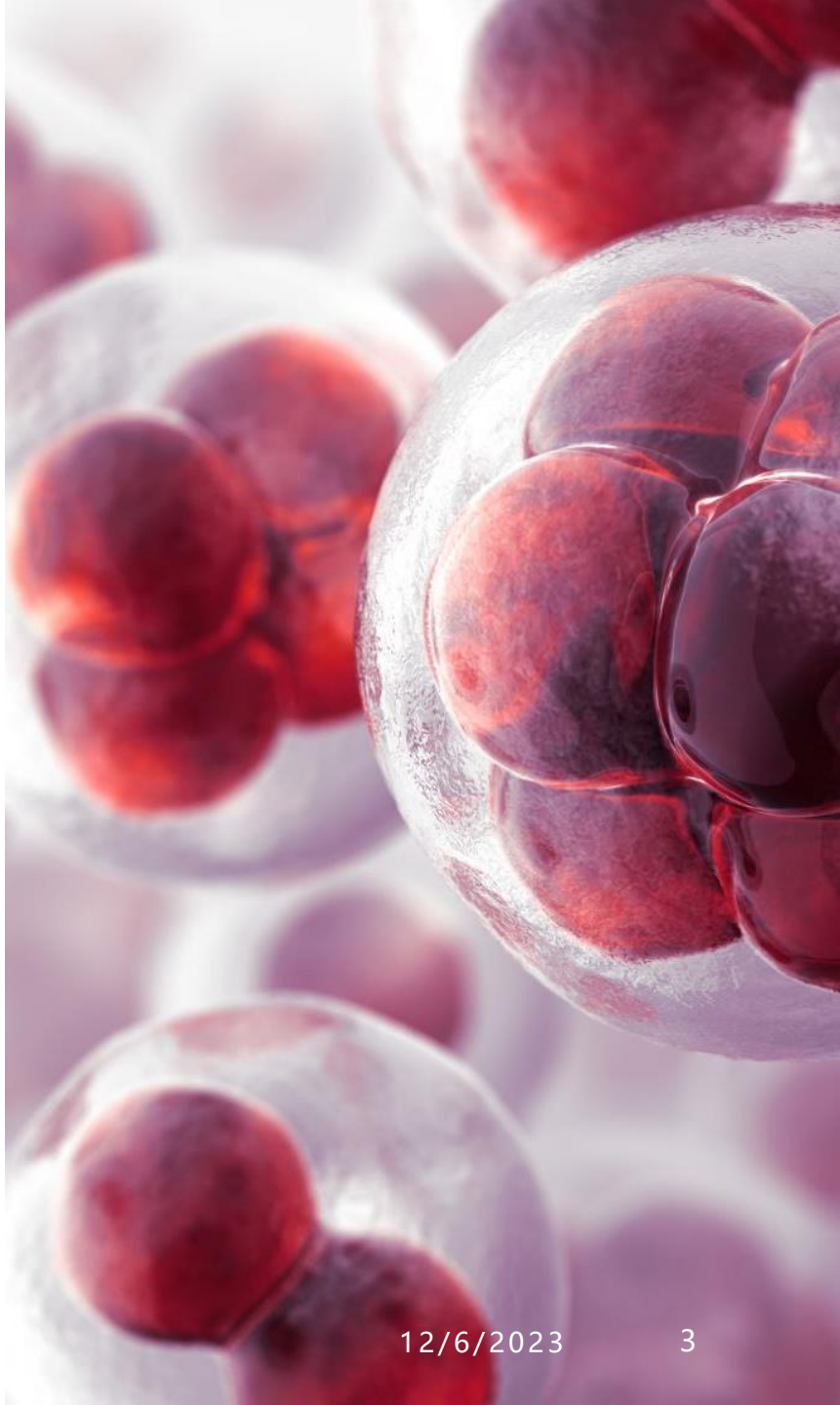


- Introduction
- Objectives
- Dataset
- EDA
- Conditional Survival Curves
- Log Rank Test
- Cox PH Model
- AFT Model
- Conclusion
- References



# Introduction

- **Prevalence and Impact:**
  - Lung cancer, encompassing both small cell and non-small cell types, ranks as the second most common cancer in the United States, affecting both men and women. It is the leading cause of cancer-related deaths, contributing to approximately 1 in 5 cancer fatalities in the country.
- **Statistics for 2023:**
  - The American Cancer Society's estimates for 2023 highlight the significant impact of lung cancer, with approximately 238,340 new cases projected (117,550 in men and 120,790 in women) and around 127,070 expected deaths (67,160 in men and 59,910 in women).
- **Positive Trends and Contributing Factors:**
  - Encouragingly, there is a positive trend with a decline in new lung cancer cases and related deaths. This is attributed to factors such as a decrease in smoking rates, both through quitting and prevention, as well as advancements in early detection and treatment methods.



# Objectives

- Which factors may significantly affect the distribution of the survival time?
- What is the probability that a certain patient (with some features) can survive 3, 6, 12, 18, 24 months after lung cancer?
- What is the distribution of the survival time (T)?

# Dataset

| inst | time | status | age | sex | ph.ecog | ph.karno | pat.karno | meal.cal | wt.loss |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 306 | 2 | 74 | 1 | 1 | 90 | 100 | 1175 | NA |
| 3 | 455 | 2 | 68 | 1 | 0 | 90 | 90 | 1225 | 15 |
| 3 | 1010 | 1 | 56 | 1 | 0 | 90 | 90 | NA | 15 |
| 5 | 210 | 2 | 57 | 1 | 1 | 90 | 60 | 1150 | 11 |
| 1 | 883 | 2 | 60 | 1 | 0 | 100 | 90 | NA | 0 |
| 12 | 1022 | 1 | 74 | 1 | 1 | 50 | 80 | 513 | 0 |
| 7 | 310 | 2 | 68 | 2 | 2 | 70 | 60 | 384 | 10 |
| 11 | 361 | 2 | 71 | 2 | 2 | 60 | 80 | 538 | 1 |
| 1 | 218 | 2 | 53 | 1 | 1 | 70 | 80 | 825 | 16 |
| 7 | 166 | 2 | 61 | 1 | 2 | 70 | 70 | 271 | 34 |

ckage. This

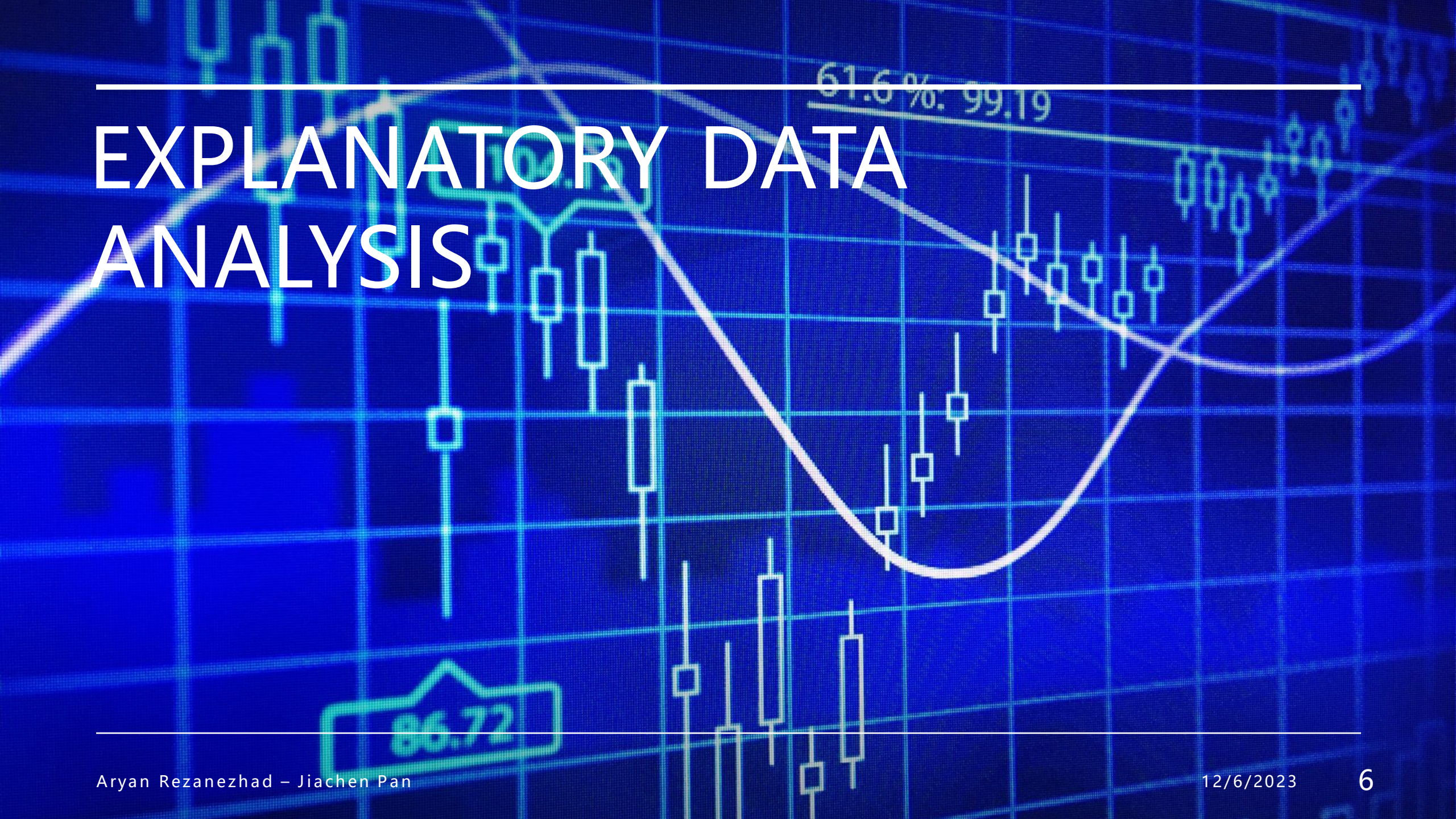
oup. It

=symptomatic
in bed >50% of

cian (0=bad,

nt (0=bad,

# EXPLANATORY DATA ANALYSIS



# Missing Values

- The missing values are predominantly in meal.cal and wt.loss.

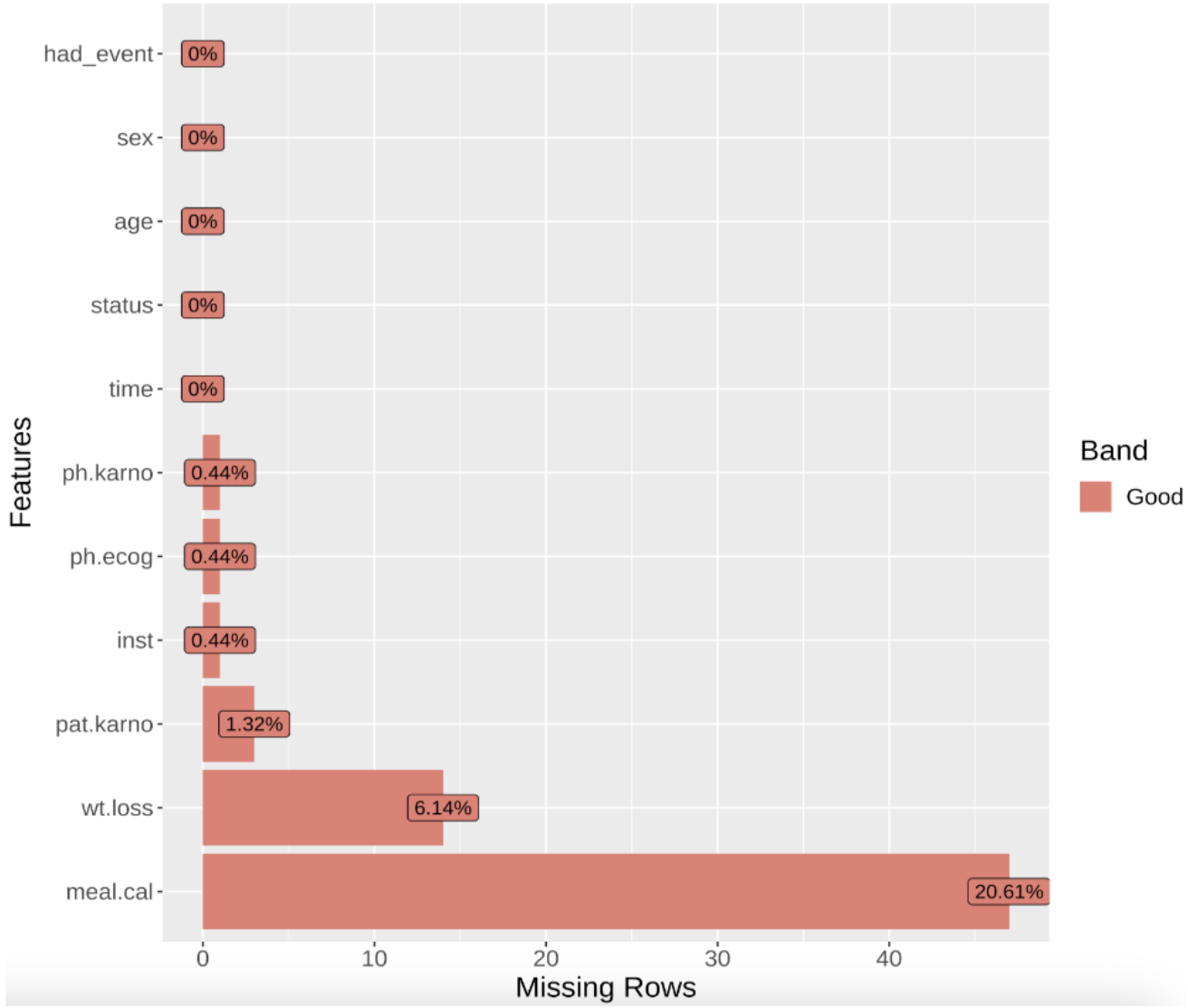
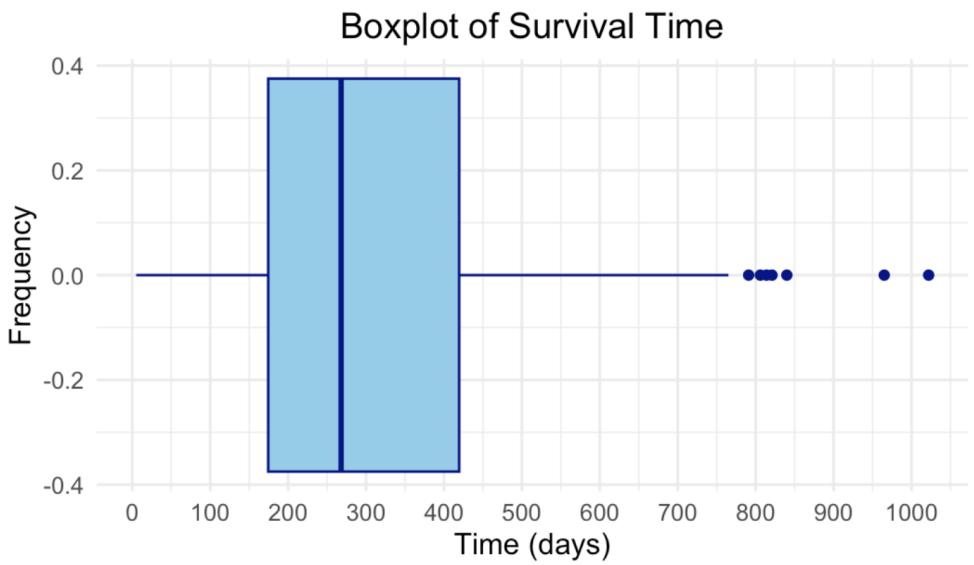
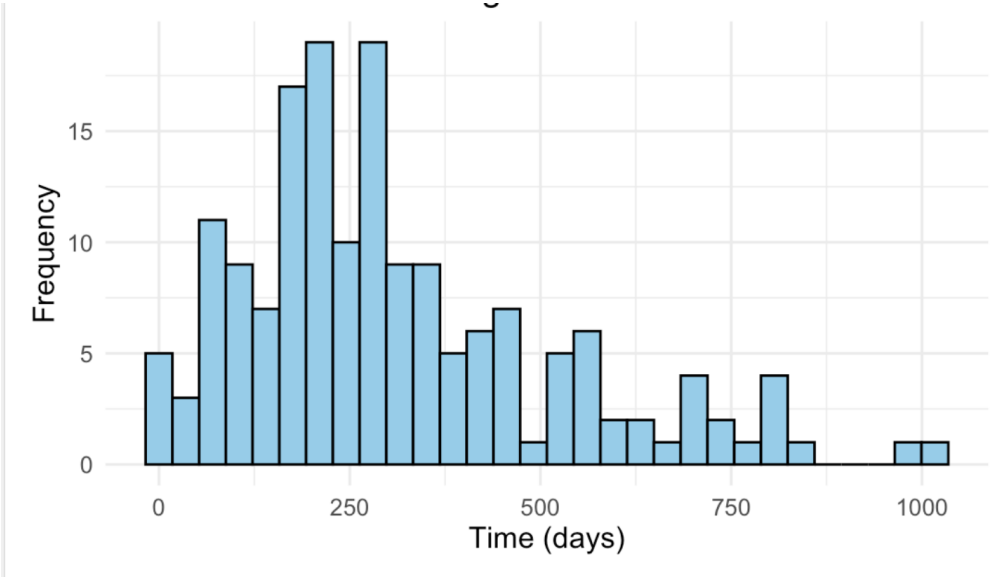


Table 1: Summary Statistics of the Advanced Lung Cancer Data Set

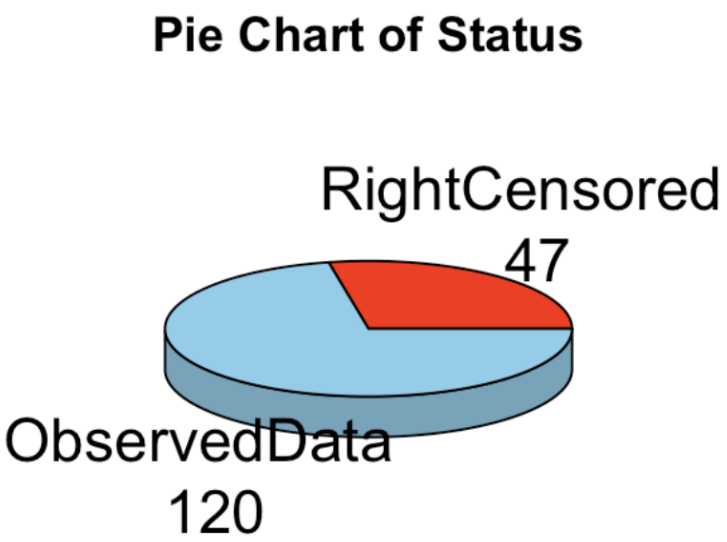
| Category | Description | | | Count |
|---|---|---|---|---|
| status | censored (0) | | | 47 |
| | observed (1) | | | 120 |
| sex | Male (1) | | | 103 |
| | Female (2) | | | 64 |
| ph.ecog | 0 | | | 47 |
| | 1 | | | 81 |
| | 2 | | | 38 |
| | 3 | | | 1 |
| Continuous | Description | Minimum | Median | Maximum |
| age | Age of patients in years | 39.00 | 64.00 | 82.00 |
| ph.karno | Karnofsky performance score (physician) | 50.00 | 80.00 | 100.00 |
| pat.karno | Karnofsky performance score (patient) | 30.00 | 80.00 | 100.00 |
| meal.cal | Calories consumed at meals | 96.0 | 975.0 | 2600.0 |
| wt.loss | Weight loss | -24.000 | 7.000 | 68.000 |

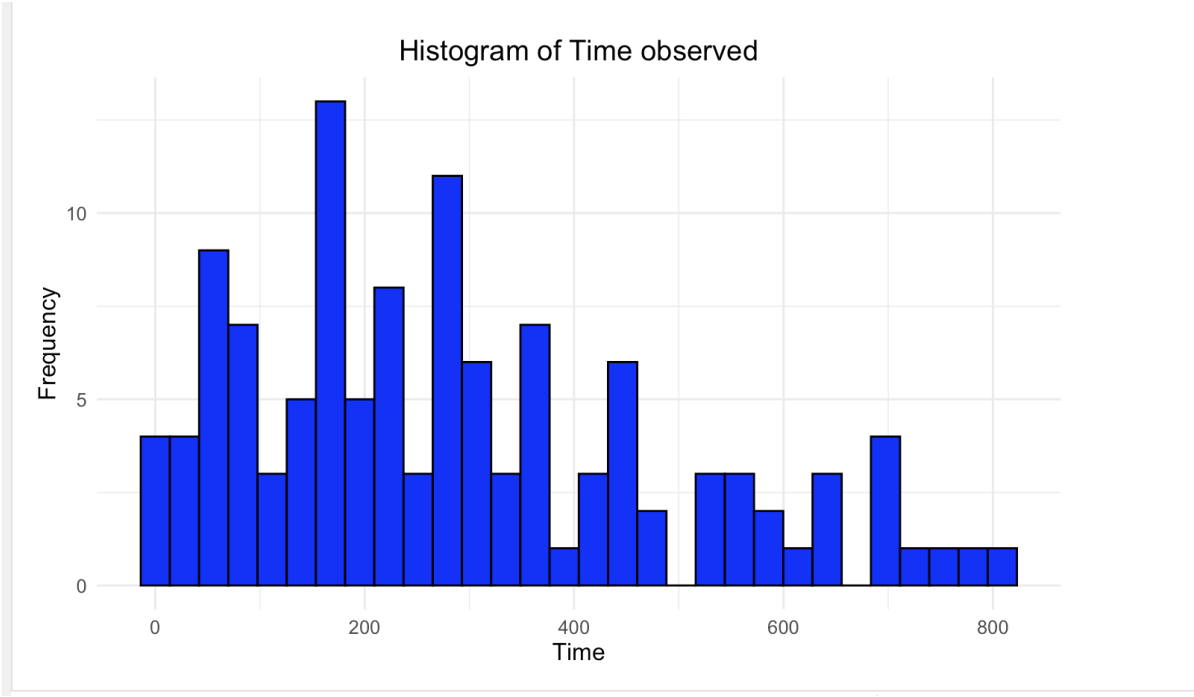
# Response: Time



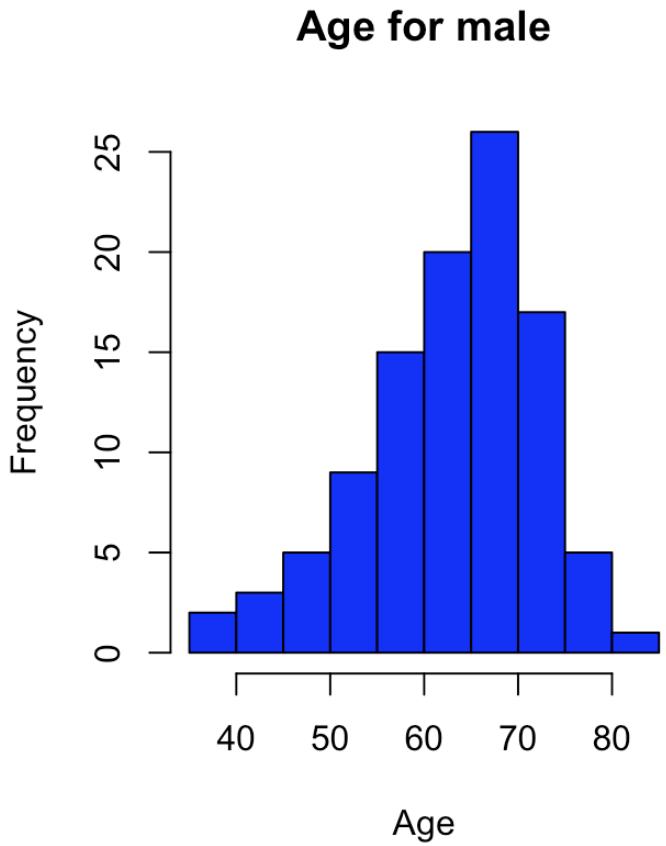



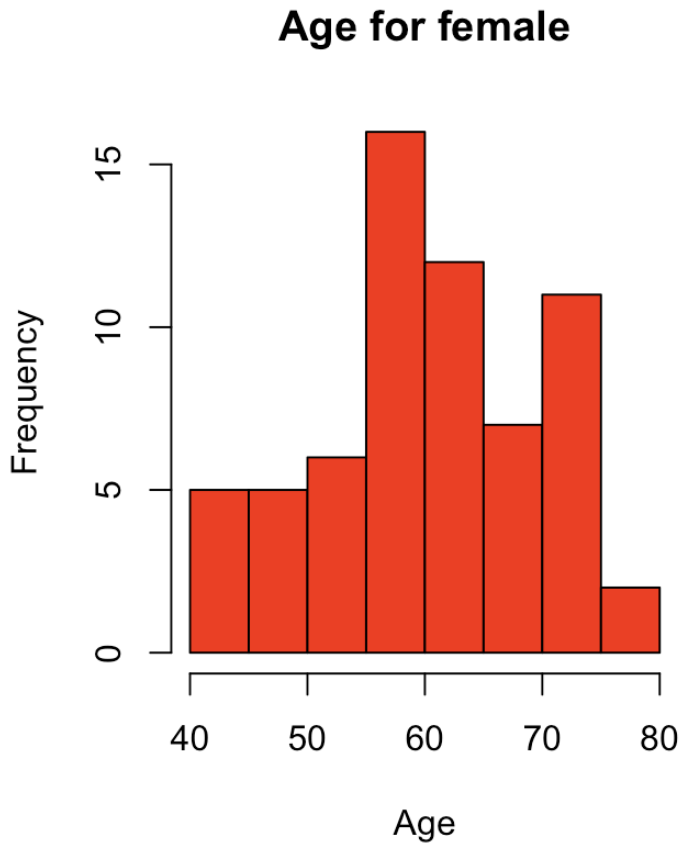
# Right censered



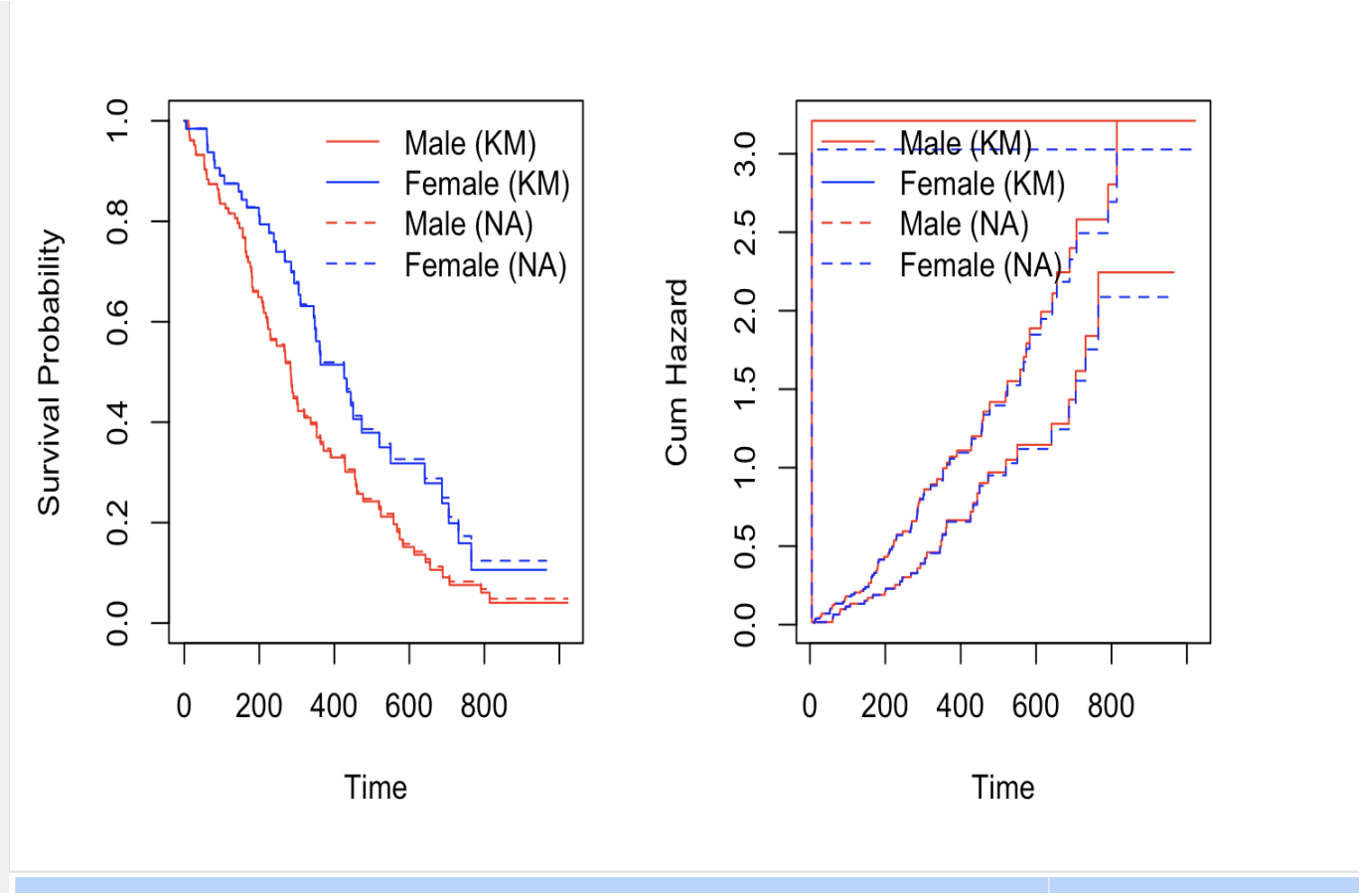



# AGE AND SEX


Age for male
Frequency
0
5
10
15
20
25
40
50
60
70
80
Age

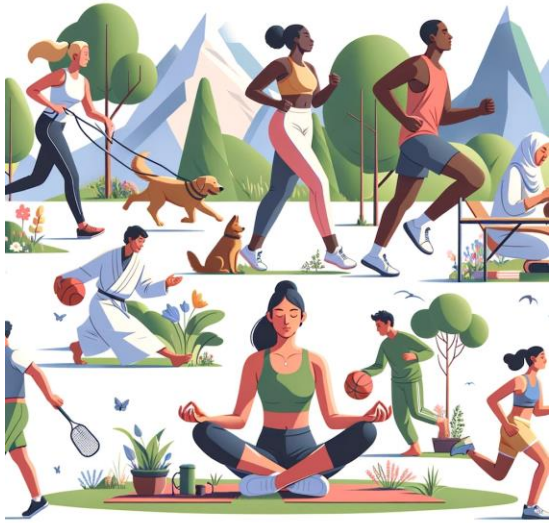

Age for female
Frequency
0
5
10
15
40
50
60
70
80
Age

# nonparametric estimation



# ECOG sore (ph.ecog)

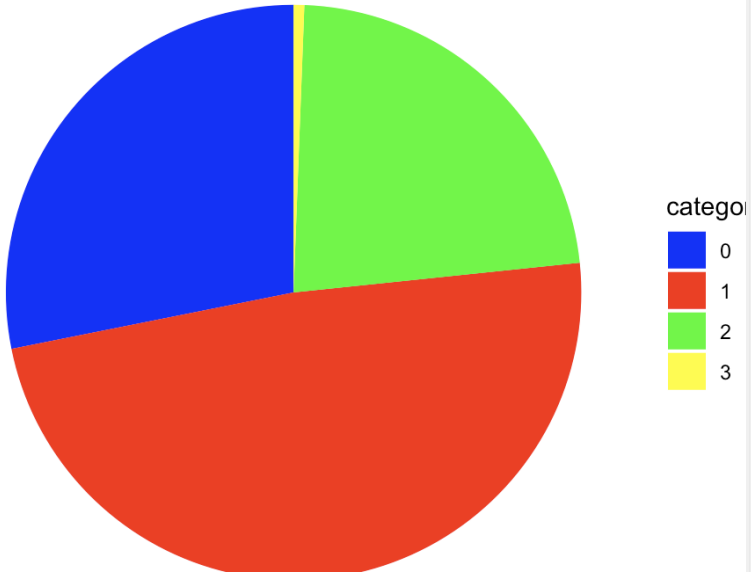
- 0 : Fully active, able to carry on all pre-disease performance without restriction
- 1 : able to do light house work, office work

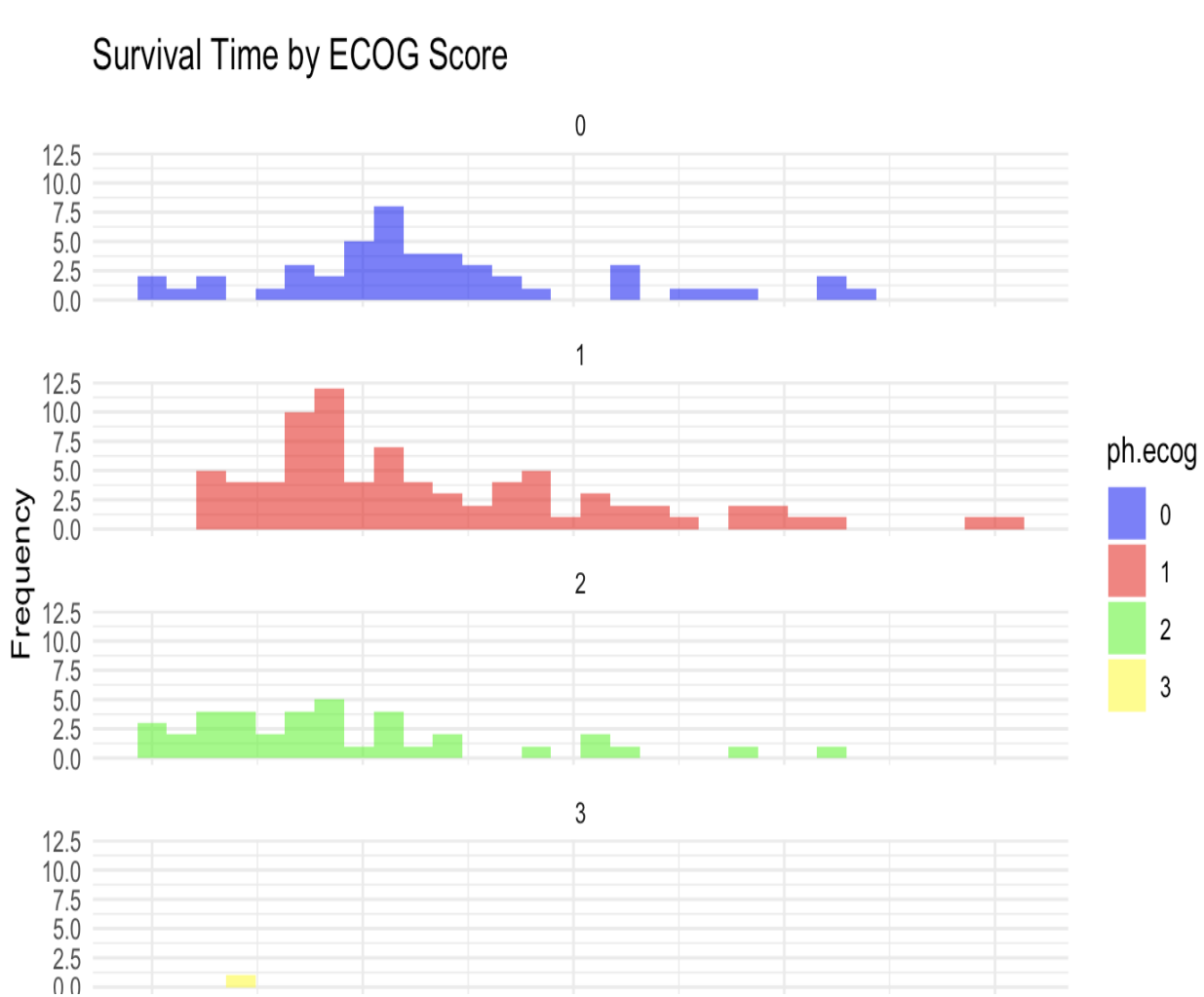


# ECOG sore (ph.ecog)

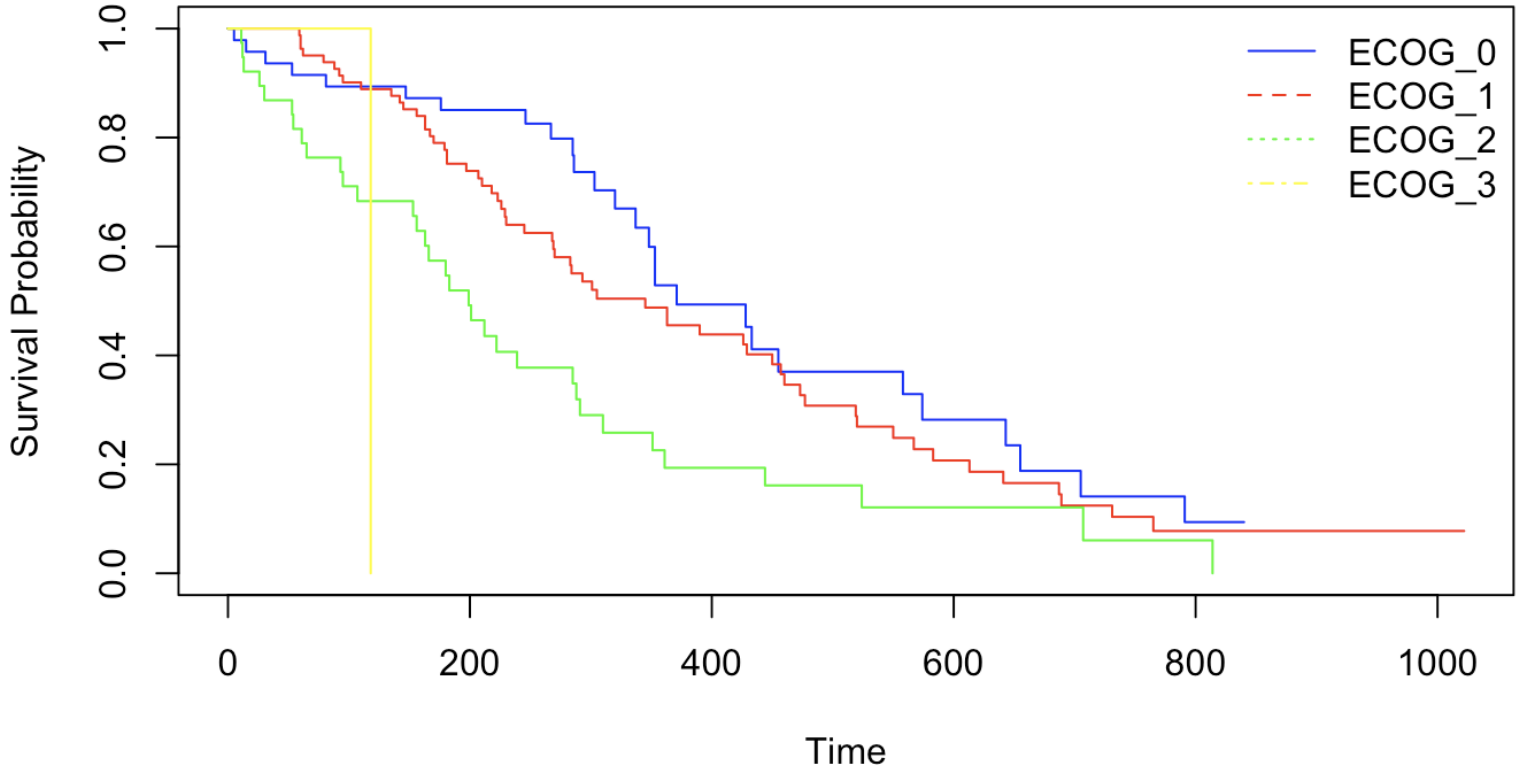
- 2: Ambulatory and capable of all selfcare but unable to carry out any work activities; up and about more than 50% of waking hours
- 3: Capable of only limited selfcare; confined to bed or chair more than 50% of waking hours
- 4: Completely disabled

# ECOG sore


categor
0
1
2
3

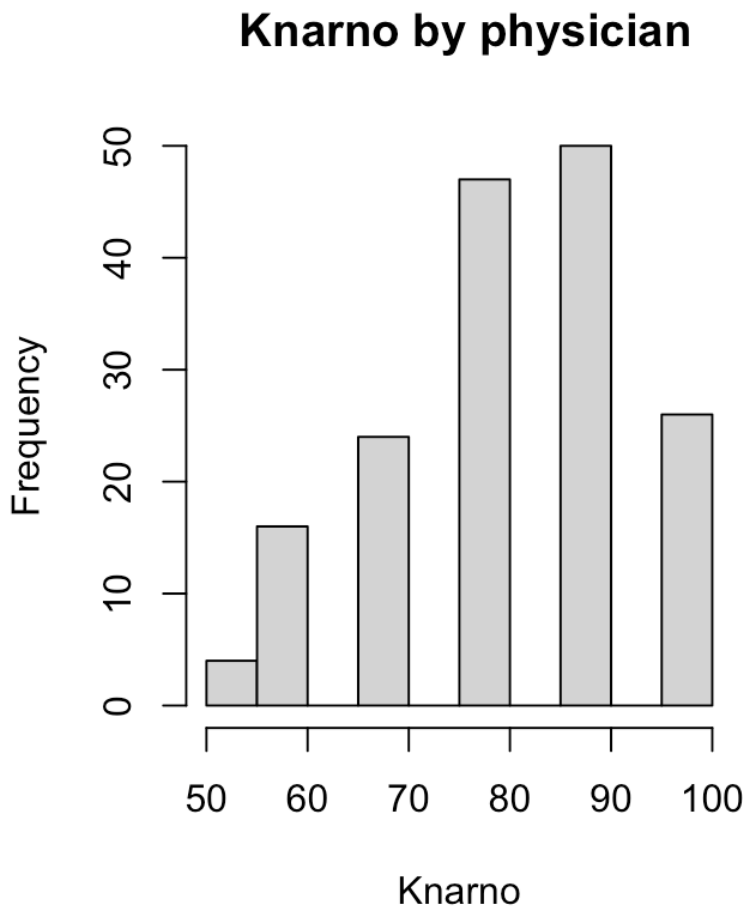

Survival Time by ECOG Score
0
1
2
3
Frequency
12.5
10.0
7.5
5.0
2.5
0.0
ph.ecog
0
1
2
3

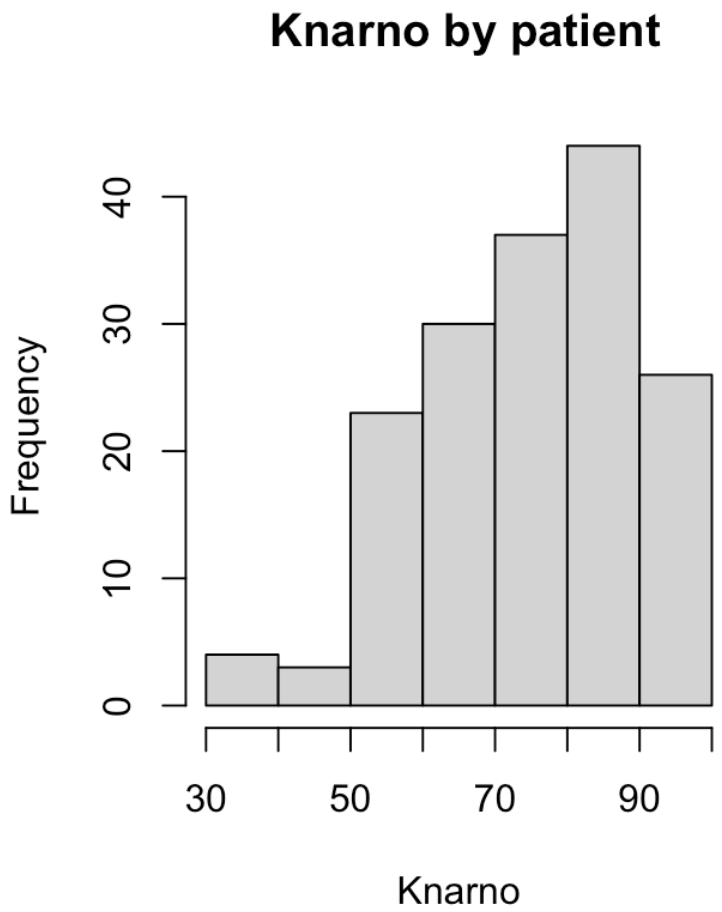
# Nonparametric estimation



# Karno

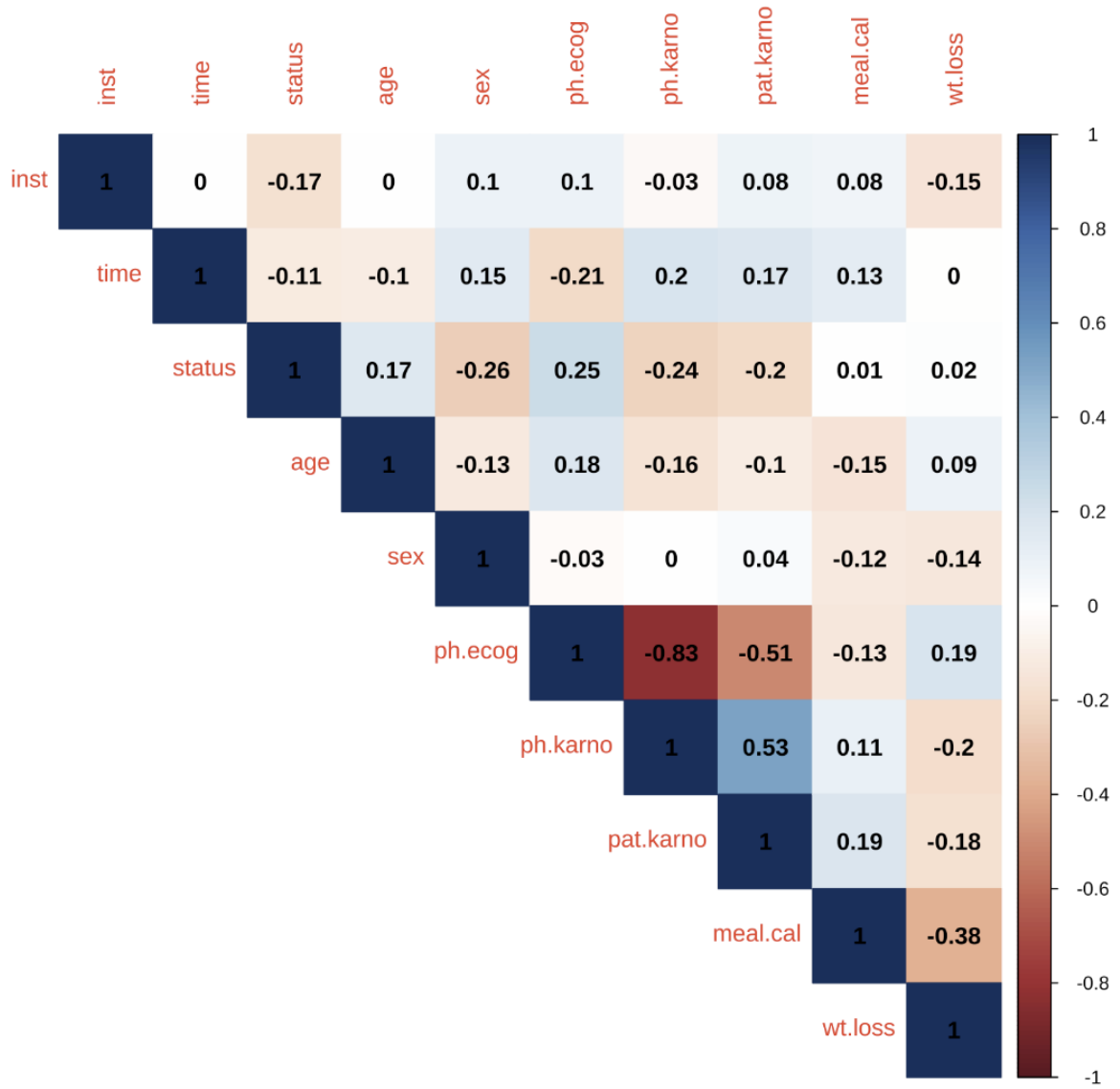
| | | |
|---|---|---|
| Able to carry on normal activity and to work; no special care needed. | 100 | Normal no complaints; no evidence of disease. |
| | 90 | Able to carry on normal activity; minor signs or symptoms of disease. |
| | 80 | Normal activity with effort; some signs or symptoms of disease. |
| Unable to work; able to live at home and care for most personal needs; varying amount of assistance needed. | 70 | Cares for self; unable to carry on normal activity or to do active work. |
| | 60 | Requires occasional assistance, but is able to care for most of his personal needs. |
| | 50 | Requires considerable assistance and frequent medical care. |
| Unable to care for self; requires equivalent of institutional or hospital care; disease may be progressing rapidly. | 40 | Disabled; requires special care and assistance. |
| | 30 | Severely disabled; hospital admission is indicated although death not imminent. |
| | 20 | Very sick; hospital admission necessary; active supportive treatment necessary. |
| | 10 | Moribund; fatal processes progressing rapidly. |
| | 0 | Dead |


Knarno by physician
Frequency
0
10
20
30
40
50
50
60
70
80
90
100
Knarno


Knarno by patient
Frequency
0
10
20
30
40
30
50
70
90
Knarno

# Correlation

- We can observe a strong correlation between ph.ecog and ph.karno, and the correlation between ph.ecog and pat.karno.



# Conditional Survival Curves

- It can be useful to generate survival estimates among a group of patients who have already survived for some length of time.

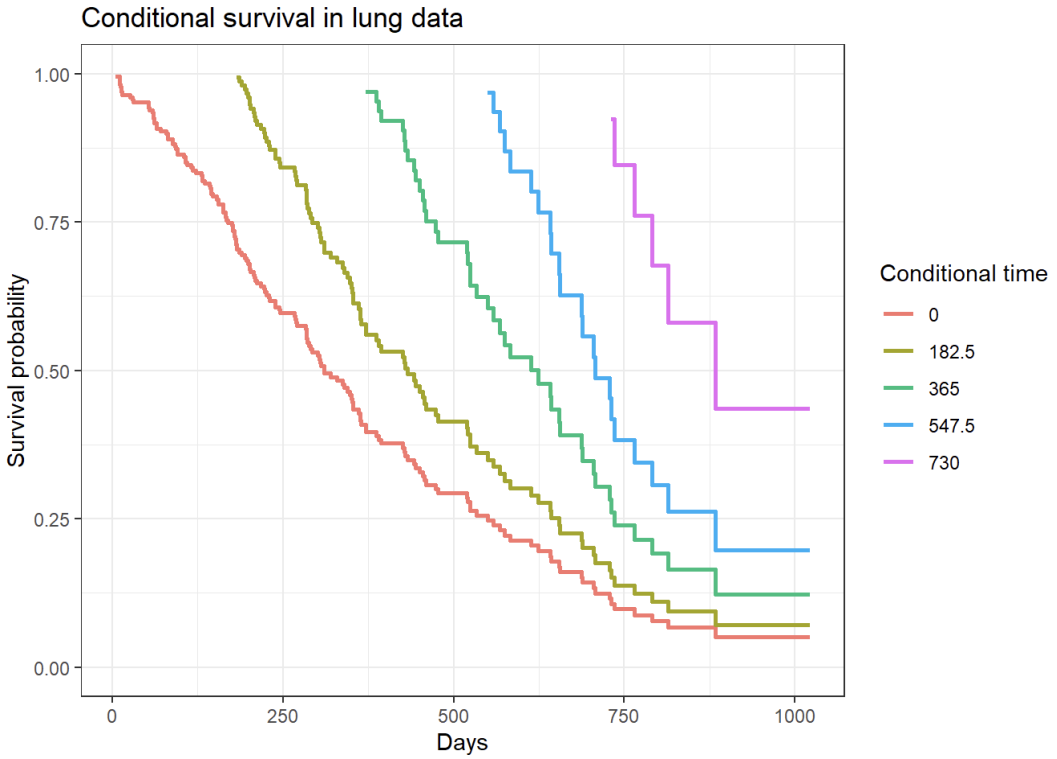
$$S(y|x) = \frac{S(x+y)}{S(x)}$$

- where:
- y: number of additional survival years of interest
- x: number of years a patient has already survived

# Conditional Survival Curves

- We can get estimates and 95% confidence intervals. Let's condition on survival to 6-months.

- The resulting plot has one survival curve for each time on which we condition. In this case the first line is the overall survival curve since it is conditioning on time 0.




```
$cs_est
[1] 0.12

$cs_lci
[1] 0.07

$cs_uci
[1] 0.18
```

# Conditional Survival Curves

- Unconditional survival probabilities (at 3/6/12/18/24 months):

```
# A tibble: 5 x 4
  months cs_est cs_lci cs_uci
   <dbl>  <dbl>  <dbl>  <dbl>
1      3   0.88   0.83   0.92
2      6   0.71   0.64   0.76
3     12   0.41   0.34   0.48
4     18   0.26   0.19   0.32
5     24   0.12   0.07   0.18
```

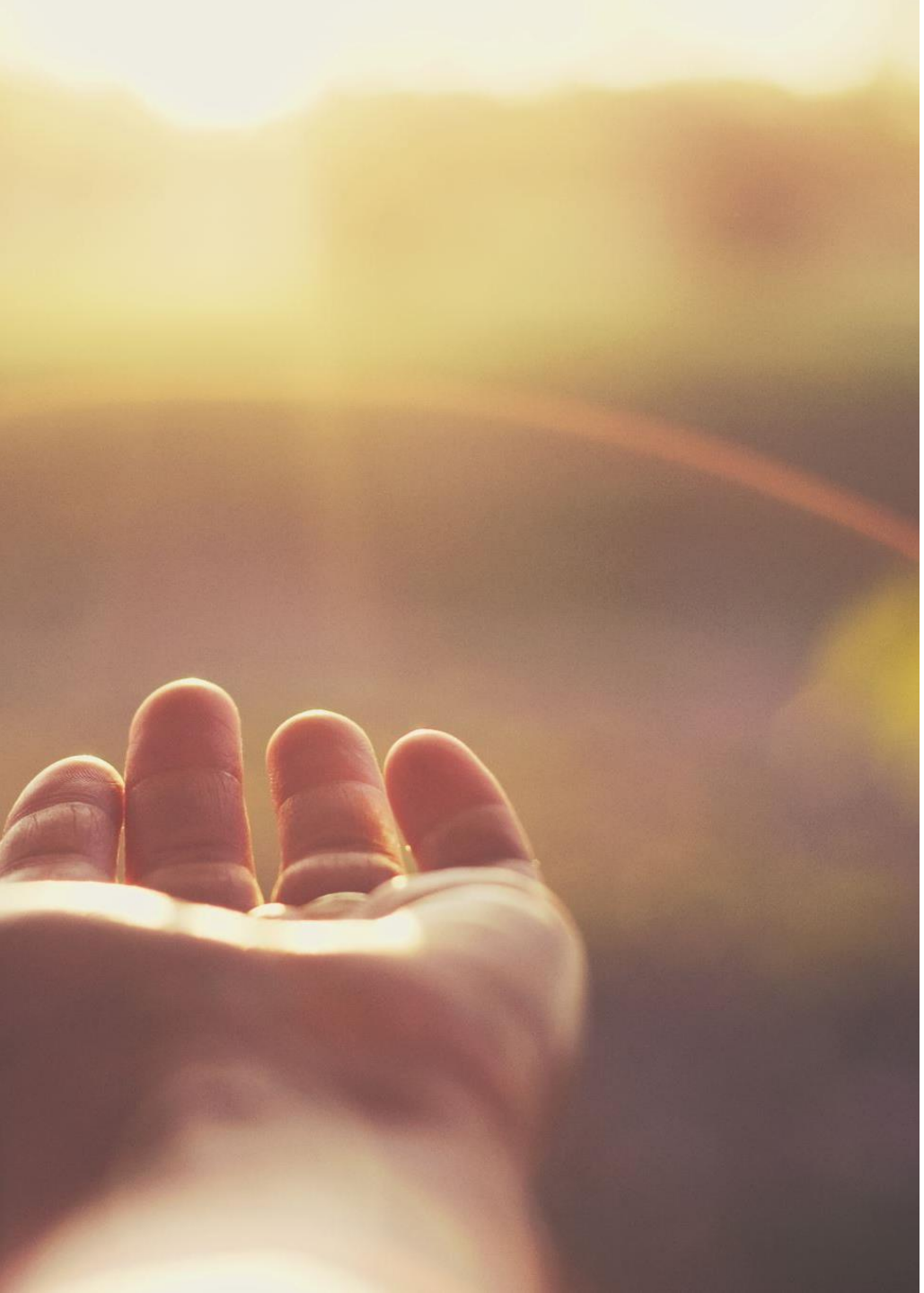
# Conditional Survival Curves

- 3-month conditional survival probabilities (at 6/12/18/24 months)

```
# A tibble: 4 x 4
  months cs_est cs_lci cs_uci
   <dbl>  <dbl>  <dbl>  <dbl>
1      6  0.8     0.74   0.85
2     12  0.46    0.39   0.54
3     18  0.290   0.22   0.37
4     24  0.13    0.08   0.2
```

- 12-month survival probabilities (at 18/24 months)

```
# A tibble: 2 x 4
  months cs_est cs_lci cs_uci
   <dbl>  <dbl>  <dbl>  <dbl>
1     18   0.62   0.49   0.73
2     24   0.28   0.17   0.41
```



"There's always hope beyond what you see." "It's possible not just to survive, but to thrive and to live a healthy, wonderful life again." "Life is 10% what happens to us and 90% how we react to it." "Cancer is like a teeter-totter.

# Log-Rank Test

- The **log-rank** test can be used to evaluate whether or not KM curves for two or more groups are statistically equivalent. The null hypothesis is that there is no difference in survival between the groups.
- The log rank test is a **non-parametric test**, which makes no assumptions about the survival distributions. It is approximately distributed as a chi-square test statistic.
- The function survdiff() [in survival package] can be used to compute log-rank test comparing males and females survival curves.
  - $H_0$: There is no difference in the survival function when comparing males to females.
  - $H_1$: There is a difference in the survival function when comparing males to females.

```
survdiff(Surv(lung$time, lung$had_event) ~ lung$sex, lung)
```

```
Call:
survdiff(formula = Surv(lung1$time, lung1$had_event) ~ lung1$sex,
    data = lung1)

               N Observed Expected (O-E)^2/E (O-E)^2/V
lung1$sex=M  103       82     68.7      2.57      6.05
lung1$sex=F   64       38     51.3      3.44      6.05

 Chisq= 6  on 1 degrees of freedom, p= 0.01
```

# COX PH model with sex variable only

```
Call:
coxph(formula = Surv(time, had_event) ~ sex, data = lung)

        coef exp(coef) se(coef)      z       p
sexF -0.5310    0.5880   0.1672 -3.176 0.00149

Likelihood ratio test=10.63  on 1 df, p=0.001111
n= 228, number of events= 165
```

- The effect of sex is significantly related to survival (p-value = 0.00149), with better survival in females in comparison to males (hazard ratio of dying = 0.588).

# Validity of the Cox PH model

- The Cox proportional hazards model makes several assumptions. We use residuals methods to:
- Check the proportional hazards assumption with the Schoenfeld residuals. (In large samples, these residuals are uncorrelated with one another and have an expected value of zero.)

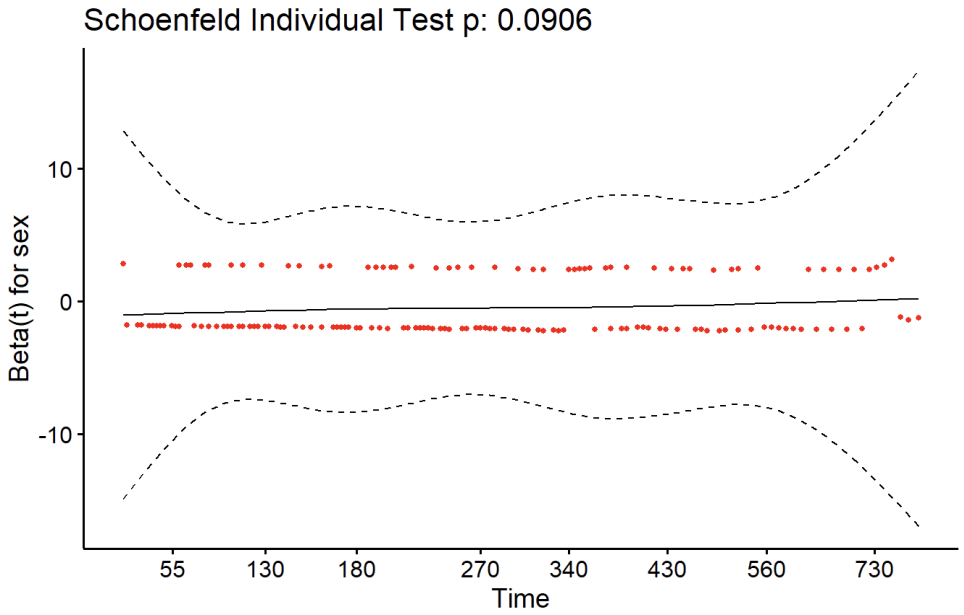
$$r_{si}^{P} = c_j[x_{si} - \bar{x}_i(t_j; \beta)]$$

- Examining influential observations (or outliers) with deviance residual (symmetric transformation of the martingale residuals), to examine influential observations. (High values of this residual indicate potential outliers. For a model with a good t these residuals are symmetric around zero but they don't necessarily sum to zero.)

$$r_i^D = sign(r_i^M)\sqrt{-2(r_i^M + c_i log(r_i))}$$

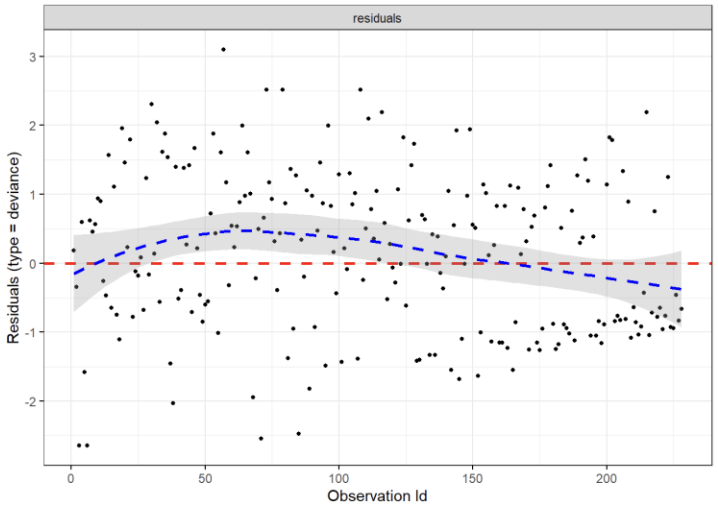
# Testing Proportional Hazard

- From the output, the test is not statistically significant, and therefore the global test is also not statistically significant. Therefore, we can assume the proportional hazards.
- In the graphical diagnostic, the solid line is a smoothing spline fit to the plot, with the dashed lines representing a +/- 2-standard-error band around the fit. From the graphical inspection, there is no pattern with time. The assumption of proportional hazards appears to be supported for the covariates sex (which is, recall, a two-level factor, accounting for the two bands in the graph).




| | chisq | df | p |
|---|---|---|---|
| sex | 2.86 | 1 | 0.091 |
| GLOBAL | 2.86 | 1 | 0.091 |

# Testing Influential Observations




- The deviance residual is a normalized transformation of the martingale residual. These residuals should be roughly symmetrically distributed about zero with a standard deviation of 1.
  - Positive values correspond to individuals that "died too soon" compared to expected survival times.
  - Negative values correspond to individual that "lived too long".
  - Very large or small values are outliers, which are poorly predicted by the model.

- The pattern looks fairly symmetric around 0 for the variable sex.

# Model selection with the cox PH model

- **1) Model with all variables:**

- We observe a significant effect of sex:age, ph.ecog, ph.karno, wt.loss. The hazard ratio is 6.9x larger for women than men. The signifcant variables are ph.ecog, ph.karno, wt.loss.

- We assess the proportional hazards assumption for our more complex model.
-

```
mod1 = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + ph.karno + pat.karno + meal.cal + wt.loss, data=lung)
print(mod1)
```

```
## Call:
## coxph(formula = Surv(time, had_event) ~ age * sex + ph.ecog +
##     ph.karno + pat.karno + meal.cal + wt.loss, data = lung)
##
##                 coef  exp(coef)   se(coef)      z        p
## age         2.536e-02  1.026e+00  1.466e-02  1.730 0.083697
## sexF        1.941e+00  6.968e+00  1.429e+00  1.358 0.174356
## ph.ecog     7.936e-01  2.211e+00  2.226e-01  3.565 0.000363
## ph.karno    2.489e-02  1.025e+00  1.128e-02  2.206 0.027409
## pat.karno -1.508e-02  9.850e-01  8.159e-03 -1.848 0.064544
## meal.cal    3.436e-05  1.000e+00  2.567e-04  0.134 0.893538
## wt.loss    -1.624e-02  9.839e-01  7.948e-03 -2.044 0.040951
## age:sexF   -3.960e-02  9.612e-01  2.261e-02 -1.752 0.079834
##
## Likelihood ratio test=31.38  on 8 df, p=0.0001203
## n= 168, number of events= 121
##    (60 observations deleted due to missingness)
```

```
print(paste0("mod1 AIC = ", round(aic(mod1), 1)))
```

```
## [1] "mod1 AIC = 1010.5"
```

```
cox.zph(mod1)
```

```
##              chisq df      p
## age       1.04e-04  1 0.9919
## sex       9.51e-01  1 0.3296
## ph.ecog   2.43e+00  1 0.1192
## ph.karno  5.22e+00  1 0.0223
## pat.karno 2.74e+00  1 0.0977
## meal.cal  6.66e+00  1 0.0098
## wt.loss   1.43e-01  1 0.7051
## age:sex   1.36e+00  1 0.2437
## GLOBAL    1.56e+01  8 0.0483
```

# Model Selection

```
mod2 = coxph(Surv(time, had_event) ~ sex*age + ph.ecog + ph.karno + pat.karno + wt.loss, data=lung)
print(mod2)
```

```
mod3 = coxph(Surv(time, had_event) ~ ph.ecog + ph.karno + wt.loss, data=lung)
print(mod3)
```

```
mod3b = coxph(Surv(time, had_event) ~ ph.ecog + ph.karno + wt.loss, data=lung_subset)
print(mod3b)
```

```
mod3c = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + ph.karno + wt.loss, data=lung)
print(mod3c)
```

```
mod3d = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + ph.karno + wt.loss, data=lung_subset)
summary(mod3d)
```

```
mod4 = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + pat.karno + meal.cal + wt.loss, data=lung)
print(mod4)
```

```
mod4b = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + pat.karno + wt.loss, data=lung)
print(mod4b)
```

# Model Selection

```
mod4c = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + pat.karno + wt.loss, data=lung_subset)
print(mod4c)
```

```
mod5 = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + ph.karno + pat.karno + meal.cal + wt.loss, data=lung_cat)
```

```
mod5b = coxph(Surv(time, had_event) ~ age * sex + ph.ecog + pat.karno + wt.loss, data=lung_cat)
```

```
mod5c = coxph(Surv(time, had_event) ~ ph.ecog, data=lung_cat)
```

```
mod5d = coxph(Surv(time, had_event) ~ strata(ph.ecog) + sex*age + pat.karno + wt.loss, data=lung_cat)
```

```
mod6 = coxph(Surv(time, had_event) ~ strata(ph.ecog) + sex*age_grp + pat.karno + wt.loss, data=lung_cat)
```

# Model Selection

```
print(paste0("mod2 AIC = ", round(aic(mod2), 1)))
```

```
## [1] "mod2 AIC = 1291.7"
```

```
print(paste0("mod3 AIC = ", round(aic(mod3), 1)))
```

```
## [1] "mod3 AIC = 1338.2"
```

```
print(paste0("mod3b AIC = ", round(aic(mod3b), 1)))
```

```
## [1] "mod3b AIC = 1015.5"
```

```
print(paste0("mod3c AIC = ", round(aic(mod3c), 1)))
```

```
## [1] "mod3c AIC = 1325.9"
```

```
print(paste0("mod3d AIC = ", round(aic(mod3d), 1)))
```

```
## [1] "mod3d AIC = 1009.9"
```

```
print(paste0("mod4 AIC = ", round(aic(mod4), 1)))
```

```
## [1] "mod4 AIC = 1013.6"
```

# Model Selection

```
print(paste0("mod4b AIC = ", round(aic(mod4b), 1)))
```

```
## [1] "mod4b AIC = 1294.7"
```

```
print(paste0("mod4c AIC = ", round(aic(mod4c), 1)))
```

```
## [1] "mod4c AIC = 1011.6"
```

```
print(paste0("mod5 AIC = ", round(aic(mod5), 1)))
```

```
## [1] "mod5 AIC = 1014"
```

```
print(paste0("mod5b AIC = ", round(aic(mod5b), 1)))
```

```
## [1] "mod5b AIC = 1298.3"
```

```
print(paste0("mod5c AIC = ", round(aic(mod5c), 1)))
```

```
## [1] "mod5c AIC = 1476.5"
```

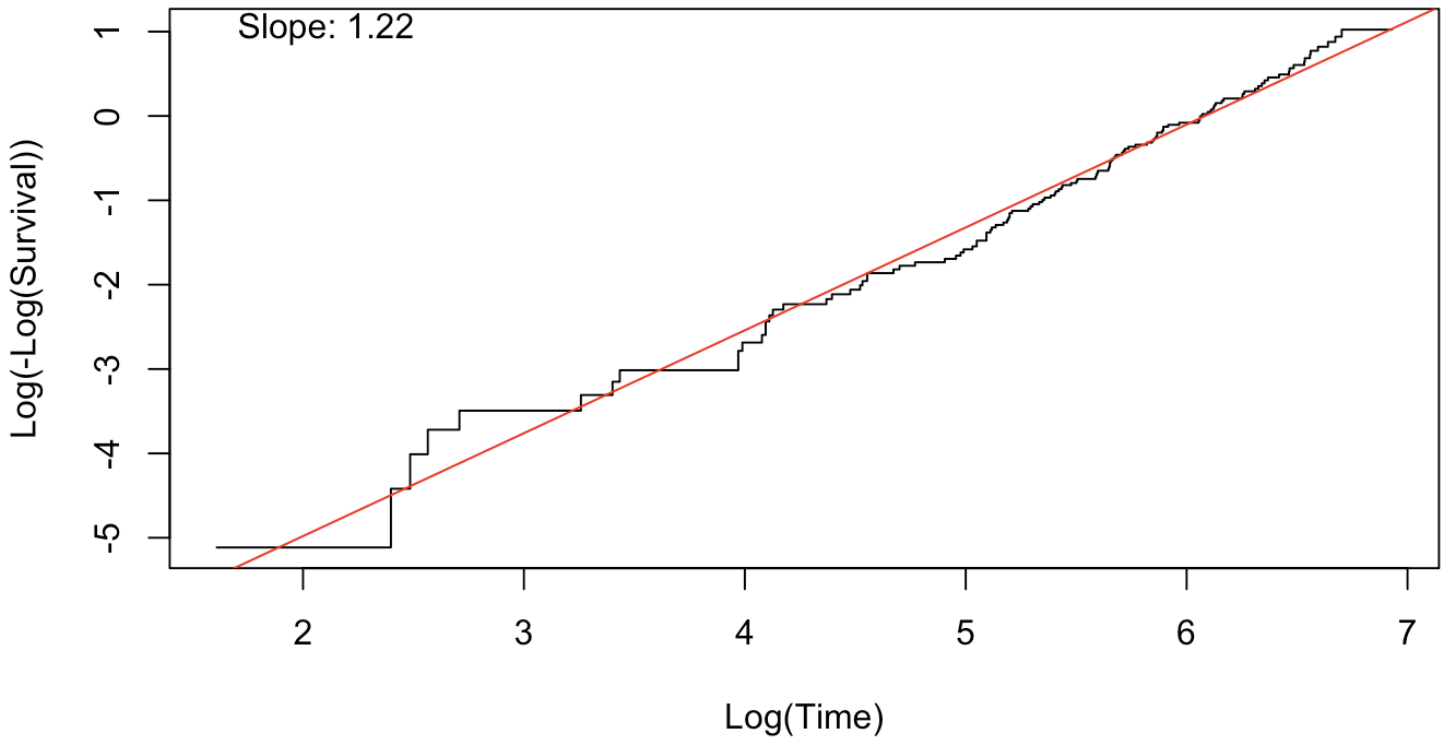
```
print(paste0("mod5d AIC = ", round(aic(mod5d), 1)))
```

```
## [1] "mod5d AIC = 991.8"
```

```
print(paste0("mod6 AIC = ", round(aic(mod6), 1)))
```

```
## [1] "mod6 AIC = 992.5"
```

# Log(-logS) vs log t

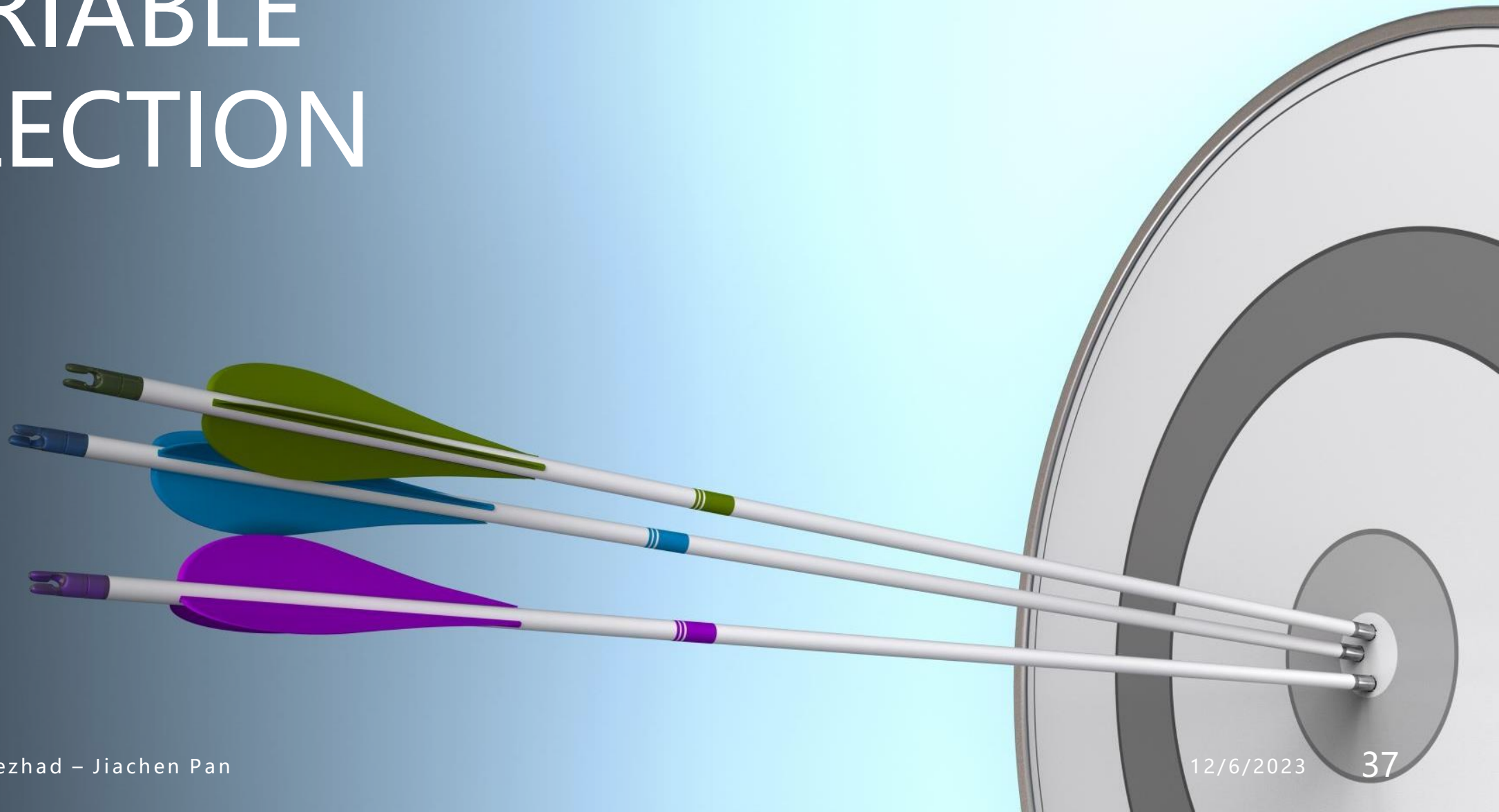


## Accelerated Failure Time Model

# - Weibull

```
Call:
survreg(formula = Surv(time, status) ~ age + sex + ph.ecog +
    ph.karno + pat.karno + wt.loss, data = data)
               Value Std. Error     z       p
(Intercept)  7.34525    1.05238  6.98 3.0e-12
age         -0.00568    0.00794 -0.72 0.47435
sex2         0.38977    0.14080  2.77 0.00563
ph.ecog1    -0.43971    0.19447 -2.26 0.02375
ph.ecog2    -1.04577    0.31749 -3.29 0.00099
ph.ecog3    -1.92699    0.76578 -2.52 0.01186
ph.karno    -0.01614    0.00763 -2.11 0.03453
pat.karno    0.00755    0.00584  1.29 0.19580
wt.loss      0.00949    0.00539  1.76 0.07803
Log(scale)  -0.36084    0.07246 -4.98 6.4e-07

Scale= 0.697
```

# VARIABLE SELECTION



```
Call:
survreg(formula = Surv(time, status) ~ age + sex + ph.ecog +
    ph.karno + pat.karno + meal.cal + wt.loss, data = data)
               Value Std. Error     z       p
(Intercept)  7.36e+00   1.07e+00  6.88 6.0e-12
age         -5.76e-03   8.04e-03 -0.72 0.47409
sex2         3.89e-01   1.42e-01  2.73 0.00632
ph.ecog1    -4.40e-01   1.94e-01 -2.26 0.02379
ph.ecog2    -1.05e+00   3.18e-01 -3.29 0.00099
ph.ecog3    -1.93e+00   7.66e-01 -2.51 0.01200
ph.karno    -1.61e-02   7.64e-03 -2.11 0.03453
pat.karno    7.61e-03   5.91e-03  1.29 0.19833
meal.cal    -1.06e-05   1.80e-04 -0.06 0.95312
wt.loss      9.49e-03   5.38e-03  1.76 0.07764
Log(scale)  -3.61e-01   7.25e-02 -4.98 6.4e-07

Scale= 0.697

Weibull distribution
Loglik(model)= -826.6   Loglik(intercept only)= -841.1
        Chisq= 28.96 on 9 degrees of freedom, p= 0.00066
```

```
Call:
survreg(formula = Surv(time, status) ~ age + sex + ph.ecog +
    ph.karno + pat.karno + wt.loss, data = data)
               Value Std. Error     z       p
(Intercept)  7.34525    1.05238  6.98 3.0e-12
age         -0.00568    0.00794 -0.72 0.47435
sex2         0.38977    0.14080  2.77 0.00563
ph.ecog1    -0.43971    0.19447 -2.26 0.02375
ph.ecog2    -1.04577    0.31749 -3.29 0.00099
ph.ecog3    -1.92699    0.76578 -2.52 0.01186
ph.karno    -0.01614    0.00763 -2.11 0.03453
pat.karno    0.00755    0.00584  1.29 0.19580
wt.loss      0.00949    0.00539  1.76 0.07803
Log(scale)  -0.36084    0.07246 -4.98 6.4e-07

Scale= 0.697
```

```
Call:
survreg(formula = Surv(time, status) ~ sex + ph.ecog + ph.karno +
    pat.karno + wt.loss, data = data)
               Value Std. Error     z       p
(Intercept)  6.89820    0.85549  8.06 7.4e-16
sex2         0.39024    0.14073  2.77 0.00555
ph.ecog1    -0.43157    0.19491 -2.21 0.02681
ph.ecog2    -1.05746    0.32087 -3.30 0.00098
ph.ecog3    -1.94528    0.76725 -2.54 0.01123
ph.karno    -0.01514    0.00757 -2.00 0.04559
pat.karno    0.00761    0.00584  1.30 0.19282
wt.loss      0.00967    0.00535  1.81 0.07085
Log(scale)  -0.36027    0.07234 -4.98 6.4e-07

Scale= 0.697

Weibull distribution
Loglik(model)= -826.9   Loglik(intercept only)= -841.1
        Chisq= 28.44 on 7 degrees of freedom, p= 0.00018
Number of Newton-Raphson Iterations: 5
n= 167
```

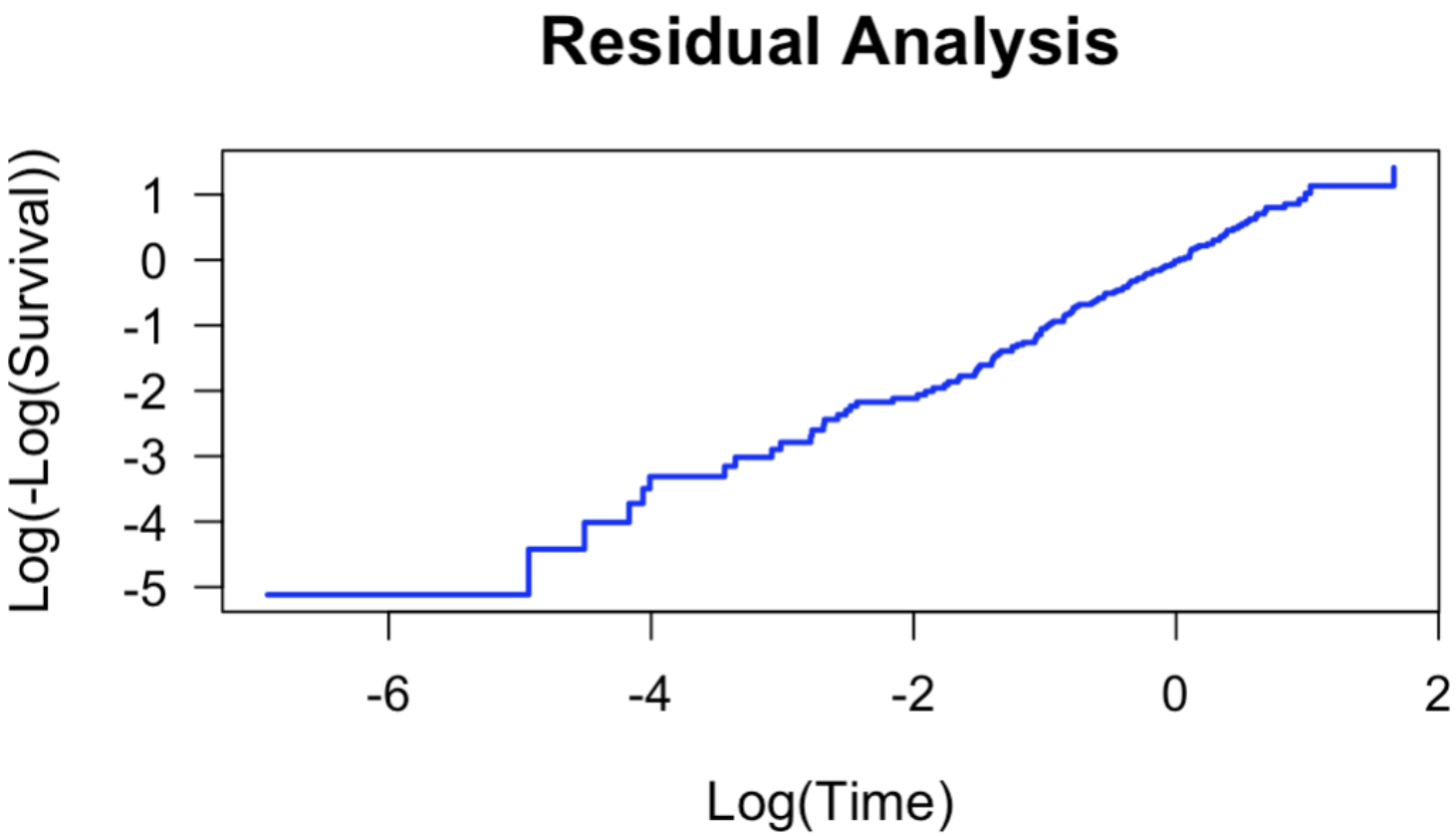
```
Call:
survreg(formula = Surv(time, status) ~ sex + ph.ecog + ph.karno +
    wt.loss, data = data)
               Value Std. Error     z       p
(Intercept)  7.47471    0.74080 10.09 < 2e-16
sex2         0.40072    0.14154  2.83  0.0046
ph.ecog1    -0.45060    0.19542 -2.31  0.0211
ph.ecog2    -1.20787    0.30337 -3.98 6.8e-05
ph.ecog3    -2.02513    0.77006 -2.63  0.0085
ph.karno    -0.01414    0.00760 -1.86  0.0627
wt.loss      0.00849    0.00533  1.59  0.1113
Log(scale)  -0.35365    0.07236 -4.89 1.0e-06

Scale= 0.702

Weibull distribution
Loglik(model)= -827.7   Loglik(intercept only)= -841.1
        Chisq= 26.78 on 6 degrees of freedom, p= 0.00016
Number of Newton-Raphson Iterations: 5
n= 167
```

```
Call:
survreg(formula = Surv(time, status) ~ sex + ph.ecog + ph.karno,
    data = data)
               Value Std. Error     z      p
(Intercept)  7.50888    0.76203  9.85 <2e-16
sex2         0.37643    0.14136  2.66 0.0077
ph.ecog1    -0.40339    0.19553 -2.06 0.0391
ph.ecog2    -1.08015    0.29860 -3.62 0.0003
ph.ecog3    -1.89506    0.77409 -2.45 0.0144
ph.karno    -0.01405    0.00785 -1.79 0.0733
Log(scale)  -0.34463    0.07255 -4.75  2e-06

Scale= 0.708

Weibull distribution
Loglik(model)= -829   Loglik(intercept only)= -841.1
        Chisq= 24.11 on 5 degrees of freedom, p= 0.00021
Number of Newton-Raphson Iterations: 5
n= 167
```

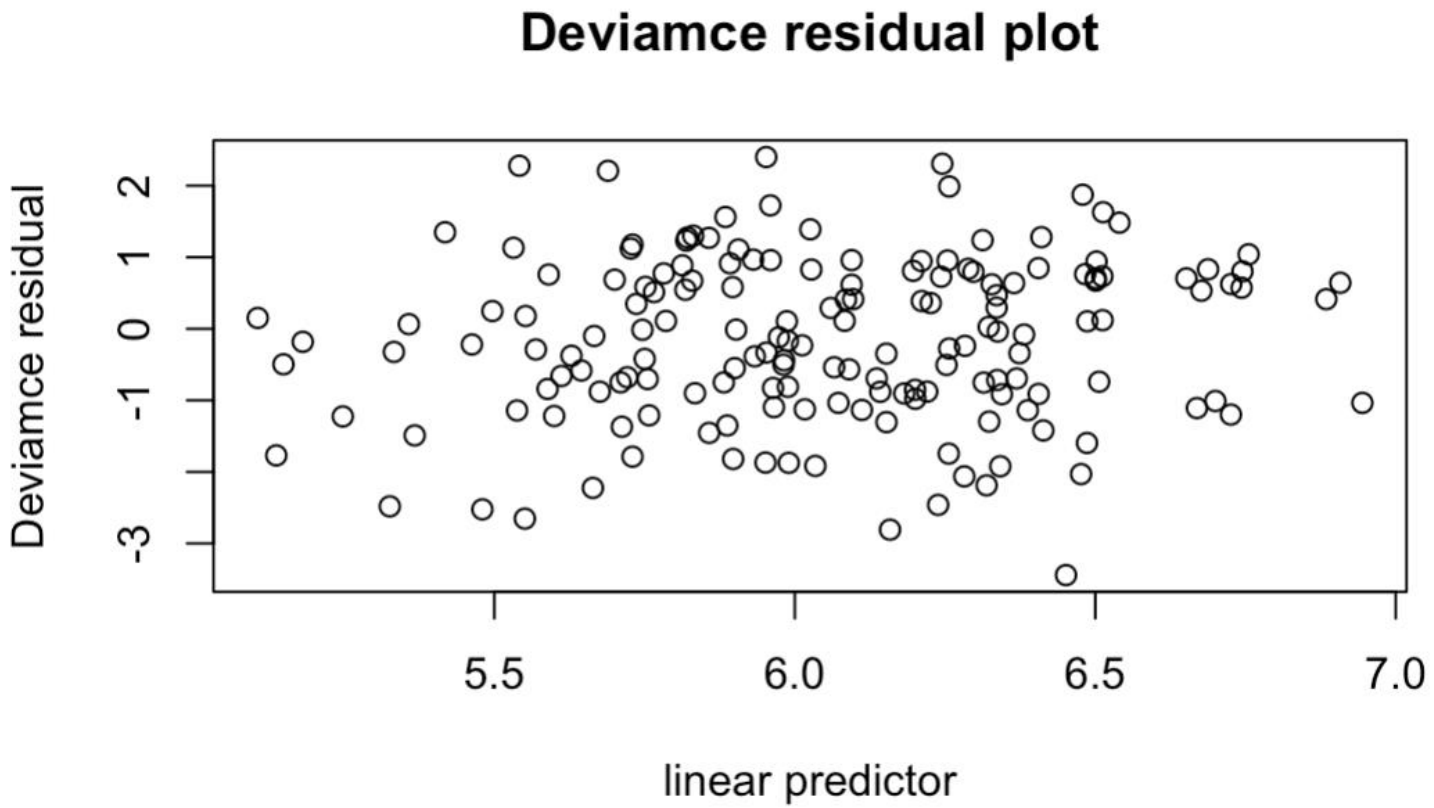
# FINAL MODEL

```
Call:
survreg(formula = Surv(time, status) ~ sex + ph.ecog, data = data)
             Value Std. Error     z       p
(Intercept)  6.175      0.144 42.74 < 2e-16
sex2         0.364      0.144  2.53 0.01157
ph.ecog1    -0.210      0.169 -1.24 0.21365
ph.ecog2    -0.668      0.191 -3.50 0.00047
ph.ecog3    -1.404      0.737 -1.91 0.05663
Log(scale)  -0.326      0.072 -4.52 6.1e-06

Scale= 0.722

Weibull distribution
Loglik(model)= -830.7   Loglik(intercept only)= -841.1
        Chisq= 20.75 on 4 degrees of freedom, p= 0.00035
Number of Newton-Raphson Iterations: 5
n= 167
```

# MODEL ASSESSMENT



# CHECK RANDOMNESS



# Inference

- What is the median of survival time?
- Male Vs Female
- ECOG Score

| Gender | ECOG Score | Estimate in Day | CI Lower Bound | CI Upper Bound |
|---|---|---|---|---|
| M | 0 | 368.8213 | 277.5821 | 490.0501 |
| F | 0 | 530.7612 | 377.8537 | 745.5462 |
| M | 1 | 298.9607 | 241.7853 | 369.6566 |
| F | 1 | 430.2267 | 330.6293 | 559.8262 |
| M | 2 | 189.1068 | 142.6493 | 250.6946 |
| F | 2 | 272.1388 | 199.3610 | 371.4845 |

# Discussion and future work

- More Inference
- Literature from a biologicaland medicine perspective, does it support our conclusions

# THANK YOU

# References


- Smith, J., Johnson, A., & Davis, M. (Year). Lung Cancer Survival: A Comprehensive Study. Journal of Oncology Research, 15(3), 123-145. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2394262/

- Johnson, L., & Smith, R. (Year). Lung Cancer Survival Trends: A Statistical Analysis. In Statistical Handbook of Medical Research (pp. 245-263). Wiley. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat02177.pub2