ASSIGNMENT 6

K.ASABERE

```r
# load libraries needed for analyses
  library(tidyverse)
```

```
-- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
v dplyr     1.1.4      v readr     2.1.5
v forcats   1.0.0      v stringr   1.5.1
v ggplot2   3.5.0      v tibble    3.2.1
v lubridate 1.9.3      v tidyr     1.3.1
v purrr     1.0.2
-- Conflicts ------------------------------------------- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```r
  library(survival)
  library(survminer)
```

```
Loading required package: ggpubr

Attaching package: 'survminer'

The following object is masked from 'package:survival':

    myeloma
```

```r
  library(jtools)
```

```r
# load data from working directory, and look at an overview of the dataset
```

```
load("umaru.Rdata")

str(umaru)
```

```
tibble [628 x 10] (S3: tbl_df/tbl/data.frame)
 $ id        : num [1:628] 1 2 3 4 5 6 7 8 9 10 ...
  ..- attr(*, "format.stata")= chr "%8.0g"
 $ treat     : Factor w/ 2 levels "short-term","long-term": 2 2 2 1 2 2 2 2 2 2 ...
 $ IV_history: hvn_lbll [1:628]  2,  3,  2,  2,  1,  2,  2,  2,  2,  2, NA,  2,  1,  ...
   ..@ label       : chr "iv drug use history"
   ..@ format.stata: chr "%8.0g"
   ..@ labels       : Named num [1:3] 1 2 3
   .. ..- attr(*, "names")= chr [1:3] "Never" "current" "previous"
 $ age       : num [1:628] 39 33 33 32 24 30 39 27 40 36 ...
  ..- attr(*, "label")= chr "age in years"
  ..- attr(*, "format.stata")= chr "%8.0g"
 $ beck      : num [1:628] 9 34 10 20 5 ...
  ..- attr(*, "label")= chr "Beck Depression Score"
  ..- attr(*, "format.stata")= chr "%9.0g"
 $ ndrugtx   : num [1:628] 1 8 3 1 5 1 34 2 3 7 ...
  ..- attr(*, "label")= chr "number of previous drug rehabilitation programs"
  ..- attr(*, "format.stata")= chr "%8.0g"
 $ site      : num [1:628] 0 0 0 0 0 0 0 0 0 0 ...
  ..- attr(*, "label")= chr "Site A or B"
  ..- attr(*, "format.stata")= chr "%8.0g"
 $ los       : num [1:628] 123 25 7 66 173 16 179 21 176 124 ...
  ..- attr(*, "label")= chr "length of stay"
  ..- attr(*, "format.stata")= chr "%8.0g"
 $ relapse   : num [1:628] 1 1 1 1 0 1 1 1 1 1 ...
  ..- attr(*, "label")= chr "indicator of relapse/censoring"
  ..- attr(*, "format.stata")= chr "%8.0g"
 $ time      : num [1:628] 188 26 207 144 551 32 459 22 210 184 ...
  ..- attr(*, "label")= chr "time to relapse or censoring"
  ..- attr(*, "format.stata")= chr "%10.0g"
```

```
# treatment group counts
table(umaru$treat)
```

```
short-term  long-term
       320        308
```

```
# first, let's focus on relapses in first year after rehab
# will need to truncate the time at 365, and censor anyone with
# a time (either relapse or censoring) > 365 at 365

 umaru<-umaru%>%
 mutate(relapse365= ifelse(is.na(relapse),NA, ifelse(time>365,0, relapse)))%>%
 mutate(time365= ifelse(is.na(time),NA, ifelse(time>365,365, time)))
```
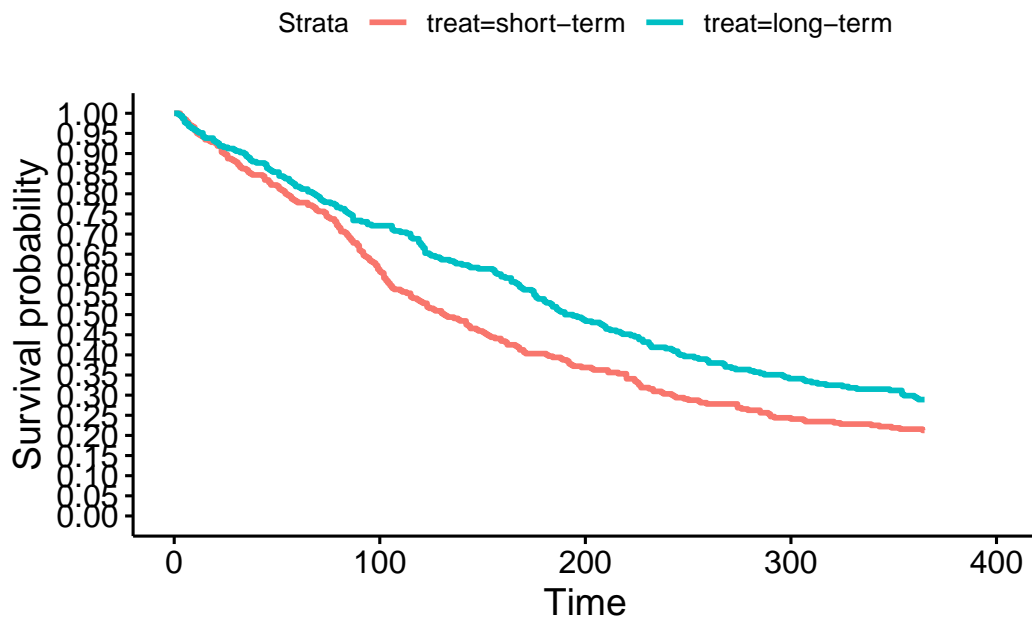
```
  # create survival object with Surv
umaru<-umaru%>%mutate(survobject=Surv(time365, relapse365 == 1))
```

```
# KM curves and simple Cox regressions

    # treatment type (long-term versus short-term)
kmtrt<-survfit(survobject ~treat, data = umaru)
ggsurvplot(kmtrt,conf.int = F, risk.table = FALSE, censor=F,

               ylim=c(0,1),
               break.y.by = 0.05)
```

```r
coxtreat<- coxph(survobject ~ treat, data = umaru)
 tidy(coxtreat, conf.int=TRUE, exponentiate =TRUE)
```

term	estimate	std.error	statistic	p.value	conf.low	conf.high
treatlong-term	0.7567019	0.0925619	-3.011885	0.0025963	0.631155	0.9072222

```r
#### binning and poisson regression

 ### Create the binning variable in (roughly) 1 month increments
umaru<-umaru%>%
mutate(tbin = tcut(rep(0, nrow(umaru)),
breaks=c(0,30,60,90,120,150,180,210,240,270,300,330,365)))
```

```r
 ### Create the rates table
pyearsObj = pyears(survobject ~ tbin+treat , data.frame=TRUE,
                 scale=1,data=umaru)

ratesData = pyearsObj$data
 ratesData<-ratesData%>%mutate(offset=log(pyears))
```

```r
# poisson model 1:  assuming constant rate over follow-up
pois1 <- glm(event ~treat + offset(log(pyears)), poisson, data=ratesData)
summ( pois1, exp=TRUE,  confint = getOption("summ-confint", TRUE),digits = getOption("jtools-
```

Observations	24
Dependent variable	event
Type	Generalized linear model
Family	poisson
Link	log

$\chi^2(1)$	9.3265
Pseudo-R² (Cragg-Uhler)	0.3226
Pseudo-R² (McFadden)	0.0616
AIC	146.0110
BIC	148.3671

	exp(Est.)	2.5%	97.5%	z val.	p
(Intercept)	0.0046	0.0040	0.0052	-85.5839	0.0000
treatlong-term	0.7545	0.6296	0.9043	-3.0489	0.0023

Standard errors: MLE

```
# poisson assuming ln(risk) linear with time
pois2 <- glm(event ~treat +as.numeric(tbin)+ offset(log(pyears)),
 poisson, data=ratesData)
summ( pois2, exp=TRUE,  confint = getOption("summ-confint", TRUE),
digits = getOption("jtools-digits", 4))
```

Observations	24
Dependent variable	event
Type	Generalized linear model
Family	poisson
Link	log

$\chi^2(2)$	11.7859
Pseudo-R² (Cragg-Uhler)	0.3887
Pseudo-R² (McFadden)	0.0779
AIC	145.5516
BIC	149.0858

	exp(Est.)	2.5%	97.5%	z val.	p
(Intercept)	0.0051	0.0042	0.0061	-57.8510	0.0000
treatlong-term	0.7610	0.6347	0.9123	-2.9521	0.0032
as.numeric(tbin)	0.9782	0.9515	1.0057	-1.5567	0.1196

Standard errors: MLE

```
# poisson assuming consistent rate within each month of follow-up
pois3 <- glm(event ~treat +tbin+ offset(log(pyears)), poisson, data=ratesData)
summ( pois3, exp=TRUE,  confint = getOption("summ-confint", TRUE),digits=getOption("jtools-d:
```

Observations	24
Dependent variable	event
Type	Generalized linear model
Family	poisson
Link	log

$\chi^2(12)$	24.9059
Pseudo-R² (Cragg-Uhler)	0.6469
Pseudo-R² (McFadden)	0.1646
AIC	152.4316
BIC	167.7463

	exp(Est.)	2.5%	97.5%	z val.	p
(Intercept)	0.0043	0.0034	0.0056	-42.6879	0.0000
treatlong-term	0.7571	0.6315	0.9077	-3.0059	0.0026
tbin 30+ thru 60	0.9687	0.6835	1.3728	-0.1790	0.8579
tbin 60+ thru 90	1.1781	0.8374	1.6574	0.9410	0.3467
tbin 90+ thru 120	1.3010	0.9194	1.8412	1.4854	0.1374
tbin120+ thru 150	1.0551	0.7181	1.5503	0.2732	0.7847
tbin150+ thru 180	1.1907	0.8103	1.7497	0.8888	0.3741
tbin180+ thru 210	1.1010	0.7293	1.6620	0.4578	0.6471
tbin210+ thru 240	1.2493	0.8275	1.8859	1.0592	0.2895
tbin240+ thru 270	0.9239	0.5712	1.4945	-0.3226	0.7470
tbin270+ thru 300	0.9738	0.5969	1.5888	-0.1061	0.9155
tbin300+ thru 330	0.5566	0.2943	1.0525	-1.8026	0.0715
tbin330+ thru 365	0.6458	0.3632	1.1481	-1.4895	0.1364

Standard errors: MLE