

```

rm(list = ls())
if(!require(devtools)) install.packages("devtools")
devtools::install_github("kassambara/survminer", build_vignettes = FALSE)
install.packages("remotes")
remotes::install_github("sachsmc/ggkm")
install.packages("jskm")
From github: latest version
install.packages("remotes")
remotes::install_github("jinseob2kim/jskm")
alldata2=read.csv("G:\\time to employment\\datasets\\whole_524.csv")
attach(alldata2)
names(alldata2)
table(College)
alldata2[alldata2==""] <- NA
require("survival")
#####
library(MASS)
library(survival)
library(survminer)
library(ggplot2)
library(devtools)
library(ggkm)
library(ggfortify)
library(GGally)
library("coxphw")
library("splines")
library(gplot)
#update.packages(ask = FALSE)
#=====
library(jskm)
fit0 <- coxph(Surv(Time, status) ~ 1, data = alldata2)
summary(fit0)

```

#R codes for graphical displays=====

#####

#null model=====

```
fit0 <- survfit(Surv(Time) ~ 1, data = alldata2)
```

#estimating unemployment by 12 months

```
summary(survfit(Surv(Time) ~ 1, data = alldata2), times = 12)
```

#time to crude=====

```
plot(fit, xlab="Time", ylab="Survival Probability", main="Kaplan-Meier plot")
```

```
ggsurvplot(fit0,data = alldata2, surv.median.line = "hv", # Add medians survival
```

```
  palette= 'blue', legend = "none", xlab = "Time since graduation in months",
```

```
  ylab = "Unemployment rate") legend= "none"
```

#####

```
ggsurvplot(fit0,data = alldata2, surv.median.line = "hv", # Add medians survival
```

```
  palette= 'blue', legend = "none", xlab = "Time since graduation in months",
```

```
  ylab = "Unemployment rate",cumcensor = F,conf.int=TRUE)
```

#####Data management===#####

#missing value removal=====

```
subset(alldata2, !is.na( field_preference))
```

```
alldata2[!is.na(alldata2$field_preference)]
```

#####

#####curves in the same panel=====

```
fitg <- survfit(Surv(Time) ~ Sex, data = subset(alldata2, !is.na(Sex)))
```

```
summary(fitg)
```

#####Model Full=====

```
str(alldata2)
```

```
fulfit0 <- coxph(Surv(Time, status) ~ College+cluster(ID), data =alldata2 , x = TRUE,method = "breslow")
```

```
gfit1.ph <- cox.zph(fit= fulfit0, transform = "km")
```

```
gfit1.ph
```

```
fulfitm <- coxph(Surv(Time, status) ~ Sex +age+as.factor(gpacat1)+College+field_preference+
```

```
residence2+cluster(ID), data =alldata2 , x = TRUE,method = "breslow") gfit1.ph2 <- cox.zph(fit= fulfitm,
```

```
transform = "km")
```

```

plot(gfit1.ph)
#####
#-----
fit <- list(PFS = fit0, OS = fitg)
ggsurvplot(fit, data = alldata2, surv.median.line = "hv", combine = TRUE, # Combine curves
risk.table = TRUE, tables.height = 0.20, # Adjust tables height
tables.theme = theme_cleantable(), # Clean risk table
censor = TRUE, palette = 'redblueblack', legend.title = "Key \n\n", legend.labs = c("Crude", "Female",
"Male"), linewidth = 4, xlab = "Time since graduation in months",
ylab = "Unemployment rate", linetype = 1:3, lwd = 4:7)
#####=====
#####College=====
table(College)
fitcol <- survfit(Surv(Time) ~ College, data = subset(alldata2, !is.na(College)))
#-----
table(stream)
fitst <- survfit(Surv(Time) ~ stream, data = subset(alldata2, !is.na(stream)))
#####
fitst2 <- survfit(Surv(Time) ~ stream2, data = subset(alldata2, !is.na(stream2)))
###GPA=====
fitgpa <- survfit(Surv(Time) ~ gpacat1, data = subset(alldata2, !is.na(gpacat)))
table(gpacat)
length(!is.na(gpacat))
####Region-----
fitreg <- survfit(Surv(Time) ~ regioncat, data = subset(alldata2, !is.na(regioncat)))
#-----
fitregbi <- survfit(Surv(Time) ~ Regiondico, data = subset(alldata2, !is.na(Regiondico)))
#=====
ggsurvplot(fitregbi, data = alldata2,
surv.median.line = "hv", # Add medians survival pval = T, # Change legends: title & labels
legend.title = "Region where graduates \n are originaly from \n\n",

```

```

legend.labs = c("Amhara","Out of Amhara"),risk.table = F, cumevents = FALSE, cumcensor = FALSE,
tables.height = 0.5, group.by = NULL, facet.by = NULL, add.all = FALSE, combine = FALSE,legend="top",
xlab = "Time since graduation in months", ylab = "Unemployment rate", ggtheme = theme_grey())
#Residence-----
fitres <- survfit(Surv(Time)~residence2, data = subset(alldata2, !is.na(residence2)))
#Father education
fitfat <- survfit(Surv(Time) ~fatheredu_new, data = subset(alldata2, !is.na(fatheredu_new)))
#Motrher edu-----
fitmat <- survfit(Surv(Time)~motheredu_new, data = subset(alldata2, !is.na( motheredu_new)))
#field preference==-----
fitpreffild <- survfit(Surv(Time) ~ field_preference, data = subset(alldata2, !is.na( field_preference)))
#####University preference-----
fituniv <- survfit(Surv(Time) ~ Preference_university, data = subset(alldata2,
!is.na(Preference_university)))

#got vocational training -----
fitcons <- survfit(Surv(Time) ~ Got_consultance, data = subset(alldata2, !is.na(Got_consultance)))
#####
#time to employment by gender
#####side by side plot#####
require("survival")
# List of ggsurvplots
require("survminer")
table(College)
splots <- list()
#=====
splots[[1]]=ggsurvplot(fitg,data = alldata2, surv.median.line = "hv", # Add medians survival
pval = 'p=0.0004',# Change legends: title & labels
legend.title = "Sex \n\n", legend.labs = c("Female", "Male"), linewidth=4, risk.table = F, cumevents =
FALSE, cumcensor = FALSE, tables.height = , group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE, xlab = ".", ylab = "Unemployment rate",ggtheme =
theme_grey(),title="A")

```

```
#####GPA=====,
splots[[2]]=ggsurvplot(fitgpa,data = alldata2,
  surv.median.line = "hv", # Add medians survival
  pval = T,# Change legends: title & labels
  legend.title = "CGPA", legend.labs = c("2-2.74","2.75-3.24","3.25-3.74","3.75-4.0"), linewidth=4,
  risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height = 0, group.by = NULL, facet.by =
  NULL, add.all = FALSE, combine = FALSE, legend="top", xlab = ".", ylab = "Unemployment rate",
  ggtheme = theme_grey(),title="B")
#####
splots[[3]] <- ggsurvplot(fitcol, data = alldata2,
  surv.median.line = "hv", # Add medians survival
  pval = T,# Change legends: title & labels
  legend.title = "College/Faculty", legend.labs = c("CANaRM","CBE","CHM","IEBS","Law","NCS","IOT"),
  linewidth=10, risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height = 0, group.by =
  NULL,
  facet.by = NULL, add.all = FALSE, combine = FALSE, legend= "top", xlab = "Time since graduation in
  months", ylab = "Unemployment rate",ggtheme = theme_grey(),title="c")
#####
#####
splots[[4]] <- ggsurvplot(fitreg,data = subset(alldata2, !is.na(regioncat)),
  surv.median.line = "hv", # Add medians survival
  pval = T,# Change legends: title & labels legend.title = "Region\n\n", legend.labs =
  c("AA","Ama","Oro","Others","SNNPR","Tigray"), risk.table = F, cumevents = FALSE, cumcensor =
  FALSE, tables.height = 0, group.by = NULL, facet.by = NULL, add.all = FALSE, combine = FALSE,
  legend="top", xlab = ".", ylab = ".", ggtheme = theme_grey(),title="D")
#=====
splots[[5]]=ggsurvplot(fitfat, data = alldata2, timeby=5000,
  surv.median.line = "hv", # Add medians survival
  pval=T, legend.title = "Graduates' \n father academic level \n\n", legend.labs = c("At most
  Secondary","At least High school"), risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height
  = 0, group.by = NULL, facet.by = NULL, add.all = FALSE, combine = FALSE, legend="top", xlab = ".",
  ylab = ".", ggtheme = theme_grey(),title="E")
```

```

#=====mothe edu=====
splots[[6]]=ggsurvplot(fitmat, data = alldata2, timeby=5000,
surv.median.line = "hv", # Add medians survival
pval=T, legend.title = "Graduates\n mother  eduaction attainment \n\n",
legend.labs = c("At most Secondary","At leat High school"),
risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height = 0, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE, legend="top", xlab = "Time since graduation in
months", ylab = ".", ggtheme = theme_grey(),title="F")
#####
#####Residence #####
splots[[7]] <- ggsurvplot(fitres,data = subset(alldata2, !is.na(residence2)),
surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels ,legend.title = "Residence", legend.labs = c("Rural","Urban"),
risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height = 0, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE, legend="top", xlab = ".", ylab = ".", ggtheme =
theme_grey(),title="G")
#=====preference and vocational =====
splots[[8]] <-ggsurvplot(fitpreffild, data = alldata2, timeby=5000, surv.median.line = "hv", # Add medians
survival pval=T,risk.table = F, cumevents = FALSE, legend.title = "Joining field of study \n in my own
choice", legend.labs = c("No","Yes"), cumcensor = FALSE, tables.height = 0, group.by = NULL, facet.by
= NULL, add.all = FALSE, combine = FALSE, legend="top", xlab = ".", ylab = ".",ggtheme =
theme_grey(),title="H")
#=====
splots[[9]] <-ggsurvplot(fituniv, data = subset(alldata2, !is.na(Preference_university))), timeby=5000,
surv.median.line = "hv", # Add medians survival
pval=T,risk.table = F, cumevents = FALSE, legend.title = "DMU was the preferred \n study location",
legend.labs = c("No","Yes"), cumcensor = FALSE, tables.height = 0, group.by = NULL, facet.by = NULL,
add.all = FALSE, combine = FALSE, legend="top", xlab = "Time since graduation", ylab = ".",ggtheme =
theme_grey(),title="I")
#=====9 plots in one=====
arrange_ggsurvplots(splots, print = TRUE, ncol = 3, nrow = 3, risk.table.height = 0.4, surv.plot.height=25)
#####

```

```
#####

muluye <- list()

#=====

muluye[[1]]=ggsurvplot(fitg,data = alldata2, surv.median.line = "hv", # Add medians survival
pval = 'p=0.0004',# Change legends: title & labels
legend.title = "Sex \n\n", legend.labs = c("Female", "Male"), linewidth=4,
risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height = , group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE, xlab = ".", ylab = "Unemployment rate",ggtheme =
theme_grey(),title="A")

#####GPA=====,
muluye[[2]]=ggsurvplot(fitgpa,data = alldata2,
surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels
legend.title = "CGPA", legend.labs = c("2-2.74","2.75-3.24","3.25-3.74","3.75-4.0"), linewidth=4,
risk.table = F, cumevents = FALSE, cumcensor = FALSE, tables.height = 0, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE, legend="top", xlab = "Time since graduation in
months", ylab = "Unemployment rate", ggtheme = theme_grey(),title="B")

#####college#####
muluye[[3]] <- ggsurvplot(fitcol, data = alldata2, surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels legend.title = "College/Faculty", legend.labs =
c("CANaRM","CBE","CHM","IEBS","Law","NCS","IOT"), linewidth=10, risk.table = F, cumevents = FALSE,
cumcensor = FALSE, tables.height = 0, group.by = NULL, facet.by = NULL, add.all = FALSE, combine =
FALSE, legend= "top", xlab = ".", ylab = "Unemployment rate",ggtheme = theme_grey(),title="C")

#####Region#####
muluye[[4]] <- ggsurvplot(fitreg,data = subset(alldata2, !is.na(regioncat)),
surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels
legend.title = "Region\n\n",
legend.labs = c("AA","Ama","Oro","Others","SNNPR","Tigray"),
risk.table = F, cumevents = FALSE,
cumcensor = FALSE, tables.height = 0, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE,
```

```

legend="top", xlab = "Time since graduation in months",
  ylab = ".", ggtheme = theme_grey(),title="D")
#####Residence #####
muluye[[5]] <- ggsurvplot(fitres,data = subset(alldata2, !is.na(residence2)),
  surv.median.line = "hv", # Add medians survival pval = T,# Change legends: title & labels legend.title =
"Residence", legend.labs = c("Rural","Urban"), risk.table = F, cumevents = FALSE, cumcensor = FALSE,
tables.height = 0, group.by = NULL, facet.by = NULL, add.all = FALSE, combine = FALSE, legend="top",
xlab = ".", ylab = ".", ggtheme = theme_grey(),title="E")
#=====preference and vocational
GTRAINING=====
muluye[[6]] <- ggsurvplot(fitpreffild, data = alldata2, timeby=5000,
  surv.median.line = "hv", # Add medians survival
  pval=T,risk.table = F, cumevents = FALSE,
  legend.title = "Joining field of study \n in my own choice",
  legend.labs = c("No","Yes"),
  cumcensor = FALSE, tables.height = 0, group.by = NULL,
  facet.by = NULL, add.all = FALSE, combine = FALSE,
  legend="top", xlab = "Time since graduation in months",
  ylab = ".",ggtheme = theme_grey(),title="F")
#=====
#####6 in one=====
arrange_ggsurvplots(muluye, print = TRUE, ncol = 2, nrow = 2,
  risk.table.height = 0.4, surv.plot.height=25,censor.plot.height=2)
#####
yosef <- list()
#=====
yosef[[1]]=ggsurvplot(fitg,data = alldata2, surv.median.line = "hv", # Add medians survival
  pval = 'p=0.0004',# Change legends: title & labels
  legend.title = "Sex \n\n",
  legend.labs = c("Female", "Male"), linewidth=4,
  risk.table = F, cumevents = FALSE,
  cumcensor = FALSE, tables.height = , group.by = NULL,

```



```

facet.by = NULL, add.all = FALSE, combine = FALSE, xlab = ".",
ylab = "Unemployment rate",ggtheme = theme_grey(),title="A")
#####GPA=====,
yosef[[2]]=ggsurvplot(fitgpa,data = alldata2,
surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels
legend.title = "CGPA",
legend.labs = c("2-2.74","2.75-3.24","3.25-3.74","3.75-4.0"), linewidth=4,
risk.table = F, cumevents = FALSE,
cumcensor = FALSE, tables.height = 0, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation in months",
ylab = "Unemployment rate", ggtheme = theme_grey(),title="B")
#####college#####
yosef[[3]] <- ggsurvplot(fitcol, data = alldata2,
surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels
legend.title = "College/Faculty",
legend.labs = c("CANAARM","CBE","CHM",
,"IEBS","Law","NCS","IOT"), linewidth=10,
risk.table = F, cumevents = FALSE,
cumcensor = FALSE, tables.height = 0, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE,
legend= "top", xlab = ".",
ylab = "Unemployment rate",ggtheme = theme_grey(),title="C")
#####Region#####
#####Residence #####
yosef[[4]] <- ggsurvplot(fitres,data = subset(alldata2, !is.na(residence2)),
surv.median.line = "hv", # Add medians survival
pval = T,# Change legends: title & labels
legend.title = "Residence",
legend.labs = c("Rural","Urban"),

```

```

risk.table = F, cumevents = FALSE,
  cumcensor = FALSE, tables.height = 0, group.by = NULL,
  facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation in months",
  ylab = "Unemployment rate", ggtheme = theme_grey(),title="D")

```

```
#####4 in one=====
```

```

arrange_ggsurvplots(yosef, print = TRUE, ncol = 2, nrow = 2,
risk.table.height = 0.4, surv.plot.height=25,ncensor.plot.height=2)

```

```
#####
```

```
##Vocational counselling=====
```

```

pltspref[[3]] <-ggsurvplot(fitcons, data = subset(alldata2, !is.na(Got_consultance)), timeby=5000,
surv.median.line = "hv", # Add medians survival
pval=T,risk.table = F, cumevents = FALSE,

```

```

  legend.title = "Receiving vocational counselling\n how to search a job",
legend.labs = c("No","Yes"),
  cumcensor = FALSE, tables.height = 0.5, group.by = NULL,
  facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation",
  ylab = ".",ggtheme = theme_grey())

```

```
#####
```

```
#=====residence=====
```

```
##### combined plots #####
```

```
#=====three in one: grade, region, residence
```

```
#=====
```

```
splits <- list()
```

```
table(fatheredu_new)
```

```
#=====
```

```
#####
```

```
ggsurvplot(fitregbi,data = alldata2,
```

```
surv.median.line = "hv", # Add medians survival
```

```

pval = T,# Change legends: title & labels
legend.title = "Region where graduates \n are originaly from \n\n",
legend.labs = c("Amhara","Out of Amhra"),
risk.table = F, cumevents = FALSE,
  cumcensor = FALSE, tables.height = 0.5, group.by = NULL,
  facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation in months",
  ylab = "Unemployment rate", ggtheme = theme_grey())
#=====
table(residence2)
splts[[4]] <- ggsurvplot(fitres,data = subset(alldata2, !is.na(residence2)),
  surv.median.line = "hv", # Add medians survival
  pval = T,# Change legends: title & labels
  legend.title = "Graduates' original residence",
  legend.labs = c("Rural","Urban"),
  risk.table = F, cumevents = FALSE,
  cumcensor = FALSE, tables.height = 0., group.by = NULL,
  facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation in months",
  ylab = ".", ggtheme = theme_grey())
#=====
arrange_ggsurvplots(splts, print = TRUE, ncol = 3, nrow = 2, risk.table.height = 0.24)

#=====
arrange_ggsurvplots(parts, print = TRUE, ncol = 2, nrow = 1, risk.table.height = 0.4)
#####
#=====preference and vocational
GTRAINING=====

#####

pltspref <- list()
pltspref[[1]] <-ggsurvplot(fitpreffild, data = alldata2, timeby=5000,

```

```

surv.median.line = "hv", # Add medians survival
pval=T,risk.table = F, cumevents = FALSE,
legend.title = "Joining field of study \n in my own choice",
legend.labs = c("No","Yes"),
cumcensor = FALSE, tables.height = 0.5, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation",
ylab = ".",ggtheme = theme_grey())
#=====

pltspref[[2]] <-ggsurvplot(fituniv, data = subset(alldata2, !is.na(Preference_university)), timeby=5000,
surv.median.line = "hv", # Add medians survival
pval=T,risk.table = F, cumevents = FALSE,
legend.title = "DMU was the preferred study location",
legend.labs = c("No","Yes"),
cumcensor = FALSE, tables.height = 0.5, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation",
ylab = ".",ggtheme = theme_grey())
##Vocational counselling=====

pltspref[[3]] <-ggsurvplot(fitcons, data = subset(alldata2, !is.na(Got_consultance)), timeby=5000,
surv.median.line = "hv", # Add medians survival
pval=T,risk.table = F, cumevents = FALSE,
legend.title = "Receiving vocational counselling\n how to search a job",
legend.labs = c("No","Yes"),
cumcensor = FALSE, tables.height = 0.5, group.by = NULL,
facet.by = NULL, add.all = FALSE, combine = FALSE,
legend="top", xlab = "Time since graduation",
ylab = ".",ggtheme = theme_grey())
#####

arrange_ggsurvplots(pltspref, print = TRUE, ncol = 3, nrow = 1,
risk.table.height = 0.4, xlab = "Time since graduation",ylab ="Unemployment rate ")

```

```
#####
### Descriptive Analysis and Models=====
#####

fitc <- coxph(Surv(Time, status) ~ age+factor(Sex)+factor(College)+ factor(regioncat)+factor(gpacat)+
factor(residence2) +factor(field_preference)+factor(Got_consultance), data = alldata2)
summary(fitc)
res=cox.zph(fitc,terms=TRUE)
par(mfrow=c(3,3))
plot(res)
res2=cox.zph(fitc,terms=FALSE)
par(mfrow=c(1,5))
plot(res2)
#####

#AFT models using =====
# Fit weibull distribution
time <- alldata2$Time_actual[Time_actual>0]
table(is.na(time))
time>0
fitW <- fitdist(time, distr = "weibull")
summary(fitW)
fitg <- fitdist(time, "gamma")
summary(fitg)
fitln <- fitdist(time, "lnorm")
summary(fitln)
fitex <- fitdist(time, "exp")
summary(fitex)
require(actuar)
fitlogl <- fitdist(time, "llogis")
summary(fitlogl)
#####

cdfcomp(list(fitW, fitg, fitln,fitex), legendtext=c("Weibull", "gamma",
"lognormal","Exponential","Loglogistic"))
```

```

denscomp(list(fitW,fitlogl, fitg, fitln,fitex), legendtext=c("Weibull","Loglogistic", "gamma",
"lognormal","Exponential"))
qqcomp(list(fitW,fitlogl, fitg, fitln,fitex), legendtext=c("Weibull","Loglogistic", "gamma",
"lognormal","Exponential"))
ppcomp(list(fitW, fitlogl,fitg, fitln,fitex), legendtext=c("Weibull","Loglogistic", "gamma",
"lognormal","Exponential"))
gofstat(list(fitW,fitlogl, fitg, fitln,fitex), fitnames=c("Weibull","Loglogistic", "gamma",
"lognormal","Exponential"))

#####setting reference category category#####
alldata2$gpacat <- factor(alldata2$gpacat, levels = c("Pass","D","GD","VGD"))
alldata2$gpacat = relevel(alldata2$gpacat, ref = "VGD")
str(alldata2)

## Exponential model
survregExp <- survreg(Surv(timepara, status)~age+factor(Sex)+factor(regioncat)
+as.factor(gpacat1)+factor(regioncat)+factor(College), data = alldata2, dist = "exponential")
summary(survregExp)

#####
age <- survreg(Surv(timepara, status)~age
, data = alldata2, dist = "weibul")
summary( age)

=====

gender <- survreg(Surv(timepara, status)~factor(Sex)
, data = alldata2, dist = "weibul")
summary(gender)

#####
region <- survreg(Surv(timepara, status)~factor(regioncat)
, data = alldata2, dist = "weibul")
summary(region)

#####=====

gpa <- survreg(Surv(timepara, status)~+as.factor(gpacat1)
, data = alldata2, dist = "weibul")
summary(gpa)

```

```

#=====
survregWeibull <- survreg(Surv(timepara, status)~factor(stream)
, data = alldata2, dist = "weibul")
summary(survregWeibull)
#=====

residence <- survreg(Surv(timepara, status)~factor(residence2)
, data = alldata2, dist = "weibul")
summary(residence)
#=====

summary(survregWeibull)
summary(survregWeibull)
## Weibull model full model=====

survregWeibull <- survreg(Surv(timepara, status)~age+factor(Sex)+factor(regioncat)
+as.factor(gpacat1)+factor(regioncat)+factor(stream)+factor(residence2) , data = alldata2, dist =
"weibul")
summary(survregWeibull)

survregres<- survreg(Surv(timepara, status)~factor(residence2)
, data = alldata2, dist = "weibul")
summary(survregres)

#####

Weibulla <- survreg(Surv(timepara, status)~age , data = alldata2, dist = "weibul")
summary(Weibulla)

#####

Weibullg <- survreg(Surv(timepara, status)~factor(Sex) , data = alldata2, dist = "weibul")
summary(Weibullg)

#####

Weibullr <- survreg(Surv(timepara, status)~factor(regioncat) , data = alldata2, dist = "weibul")
Weibullr <- survreg(Surv(timepara, status)~factor(regioncat) , data = alldata2, dist = "weibul")
summary(Weibullr)

#####

Weibullgpa <- survreg(Surv(timepara, status)~factor(gpacat1)
, data = alldata2, dist = "weibul")

```

```

Weibullr <- survreg(Surv(timepara, status)~factor(regioncat)
, data = alldata2, dist = "weibul")
summary(Weibullgpa)
#####
Weibullgpa <- survreg(Surv(timepara, status)~factor(gpacat1) , data = alldata2, dist = "weibul")Weibulls
<- survreg(Surv(timepara, status)~factor(stream) , data = alldata2, dist = "weibul")
summary(Weibulls)
#####
Weibullgpa <- survreg(Surv(timepara, status)~factor(gpacat1)
, data = alldata2, dist = "weibul")
%%%%%%%%%%%%%%
Weibullr <- survreg(Surv(timepara, status)~factor(residence2)
, data = alldata2, dist = "weibul")
summary(Weibull)
#####all predictors=====
survregWeibull <- survreg(Surv(timepara, status)~age+factor(Sex)+factor(regioncat)
+as.factor(gpacat1)+factor(regioncat)+factor(stream)+factor(residence2) , data = alldata2, dist =
"weibul")
summary(survregWeibull)
survregWeibull <- survreg(Surv(timepara, status)~age+factor(Sex)+factor(regioncat)
+as.factor(gpacat1)+factor(regioncat)+factor(stream)+factor(residence2)+factor(Father_edu.1)
+factor(Mother_edu.1), data = alldata2, dist = "exponential")
summary(survregWeibull)
stepAIC(survregWeibull)
#####=====
fitg <- coxph(Surv(Time_actual, status) ~age+factor(Sex)+factor(gpacat), data = alldata2)
summary(fitg)
cox.zph(fitg)
#=====
fitc <- coxph(Surv(Time_actual, status) ~factor(College), data = alldata2)
summary(fitc)
cox.zph(fitc)

```



```

#=====
table(gpacat)
fitgp <- coxph(Surv(Time_actual, status) ~factor(gpacat1), data = alldata2)
summary(fitgp)
cox.zph(fitgp)
gpacat1
#####
fitreg <- coxph(Surv(Time_actual, status) ~factor(regioncat), data = alldata2)
summary(fitreg)
cox.zph(fitreg)
#####
fitres <- coxph(Surv(Time_actual, status) ~factor(residence2), data = alldata2)
summary(fitres)
#####
fitres <- coxph(Surv(Time_actual, status) ~factor(residence2), data = alldata2)
summary(fitres)
#####
fitfa <- coxph(Surv(Time_actual, status) ~factor(Father_edu.1), data = alldata2)
summary(fitfa)
#=====
fitmo <- coxph(Surv(Time_actual, status) ~factor(Mother_edu.1), data = alldata2)
summary(fitmo)
#####
fullm <- coxph(Surv(Time_actual,
status)~age+factor(Sex)+factor(stream)+as.factor(gpacat1)+factor(regioncat), data = alldata2)
cox.zph(fullm)
cox.zph(fullm)
###Fit a stratified model, clustered on college of =====
fullm2 <- coxph(Surv(Time_actual, status)~age+factor(Sex)+as.factor(gpacat)+factor(regioncat)+
strata(regioncat), data = alldata2)
summary(fullm2)
cox.zph(fullm2)

```

```
#####
```

```
(res.zph1 <- cox.zph(fullm))
```