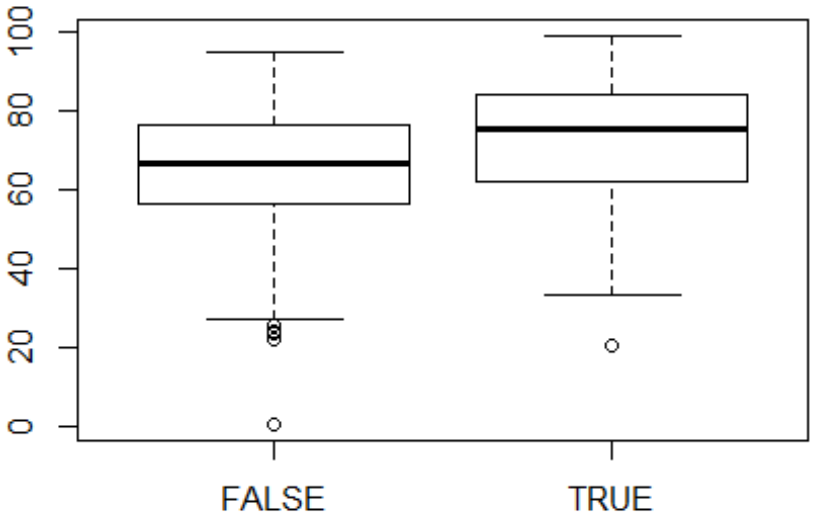
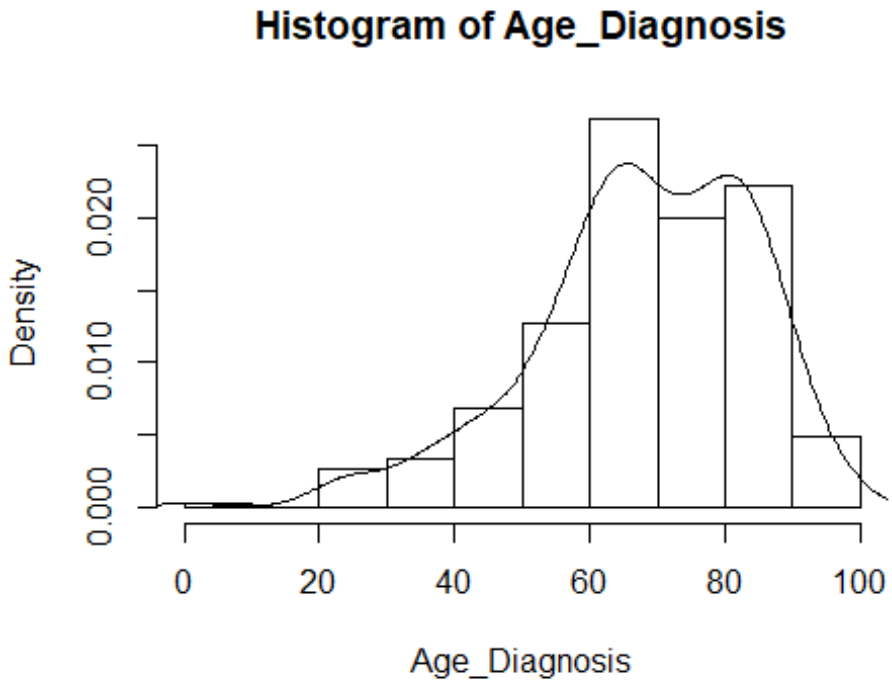
SurvivalAnalysis

W. Grossmann

```
Surv3 <- read.csv("Surv3.csv", sep=";", dec=",")
#---------------------------------------------------------
# Define subset for region and known stadium
#---------------------------------------------------------
vie1<-subset(Surv3, (Region == "VIE")&(Stadium != "unknown"))

surv.dat<-attach(vie1)

#---------------------------------------------------------
#Data description
#---------------------------------------------------------
table(Survived)

## Survived
##  no yes
## 137 168

table(MultipleTumor)

## MultipleTumor
##   0   1
## 262  43

table(Event,Survived)

##        Survived
## Event    no yes
##   FALSE   0 168
##   TRUE  137   0

table(Stadium)

## Stadium
## disseminized    localized     regional      unknown
##           88          179           38            0

summary(Age_Diagnosis)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.4167 58.8333 68.5000 67.9112 81.0000 98.9167

boxplot(Age_Diagnosis~Event)
```




```
hist(Age_Diagnosis, freq=FALSE)
lines(density(Age_Diagnosis))
```



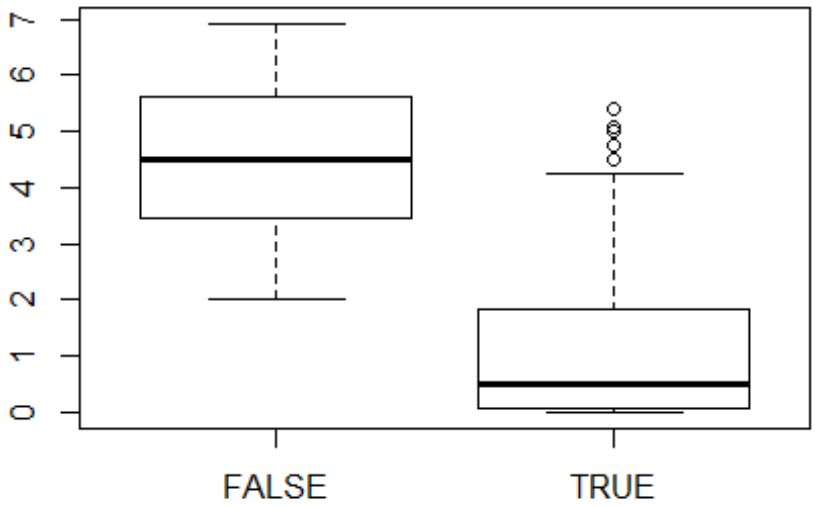

```
summary(Year_diagnosis)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2006    2007    2008    2008    2009    2010
```

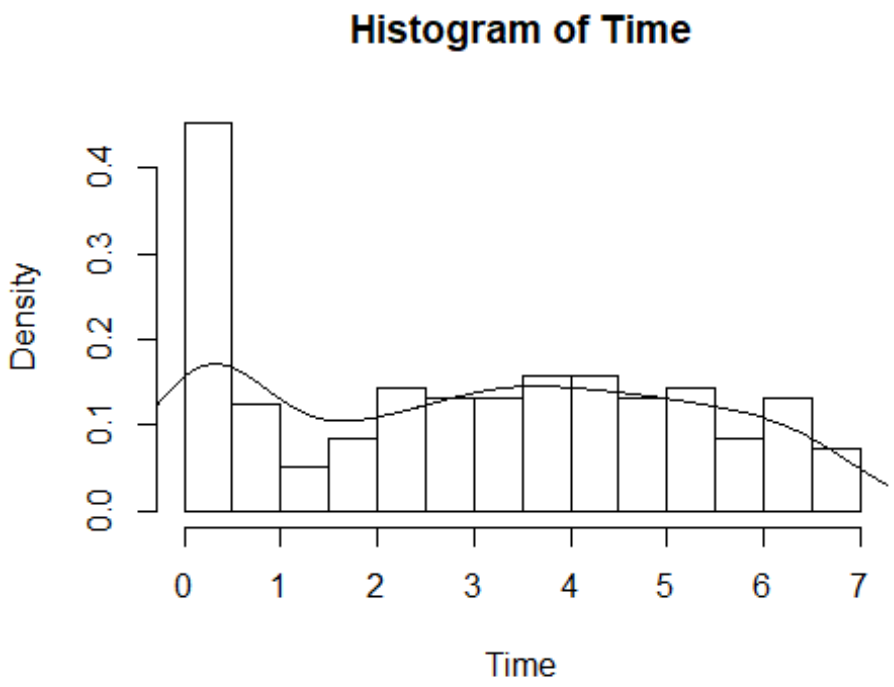
```
summary(Time)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.750   3.083   2.995   4.750   6.917
```

```
boxplot(Time~Event)
```




```
hist(Time, freq=FALSE)
lines(density(Time))
```

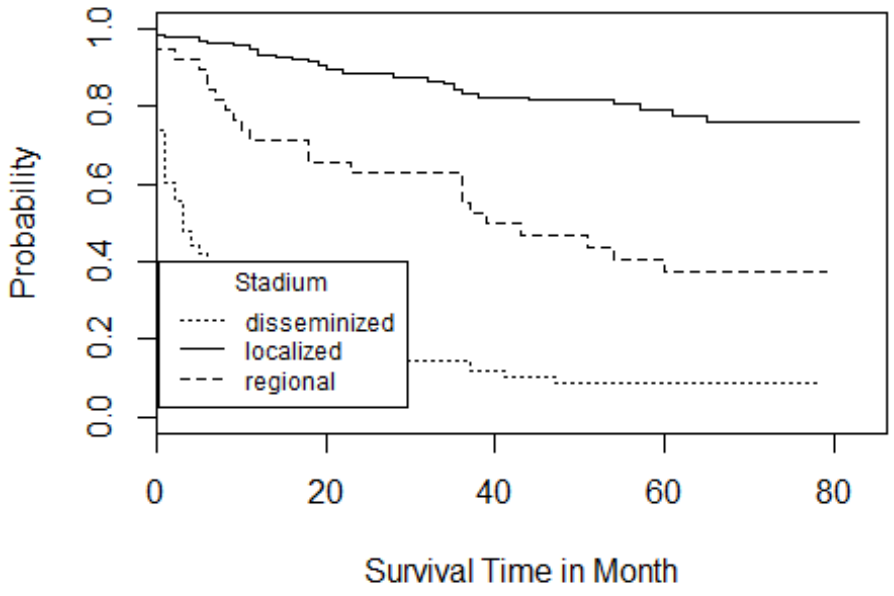



```
#-------------------------------------------------------
#Survival Kaplan Meier without group considerations
#-------------------------------------------------------
KM.years<-survfit(Surv(round(Time),Event)~1, data = vie1)
summary(KM.years)

## Call: survfit(formula = Surv(round(Time), Event) ~ 1, data = vie1)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     0    305      69    0.774  0.0240        0.728        0.822
##     1    236      23    0.698  0.0263        0.649        0.752
##     2    213      19    0.636  0.0275        0.584        0.692
##     3    174      16    0.578  0.0286        0.524        0.637
##     4    136       6    0.552  0.0292        0.498        0.612
##     5     86       4    0.526  0.0305        0.470        0.590

#-------------------------------------------------------
# Survival Kaplan-Meier with grouping by stadium
#-------------------------------------------------------
KM.Stadium<-survfit(Surv(round(Time*12),Event)~Stadium, data = vie1)
summary(KM.Stadium)

## Call: survfit(formula = Surv(round(Time * 12), Event) ~ Stadium, data = vi
e1)
##
##                 Stadium=disseminized
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     0     88      23   0.7386  0.0468       0.6523        0.836
```

```
##     1     65      12   0.6023  0.0522       0.5082         0.714
##     2     53       4   0.5568  0.0530       0.4621         0.671
##     3     49       7   0.4773  0.0532       0.3835         0.594
##     4     42       3   0.4432  0.0530       0.3506         0.560
##     5     39       2   0.4205  0.0526       0.3290         0.537
##     6     37       5   0.3636  0.0513       0.2758         0.479
##     7     32       2   0.3409  0.0505       0.2550         0.456
##     8     30       1   0.3295  0.0501       0.2446         0.444
##     9     29       4   0.2841  0.0481       0.2039         0.396
##    10     25       1   0.2727  0.0475       0.1939         0.384
##    11     24       1   0.2614  0.0468       0.1840         0.371
##    14     23       2   0.2386  0.0454       0.1643         0.347
##    18     21       1   0.2273  0.0447       0.1546         0.334
##    20     20       1   0.2159  0.0439       0.1450         0.322
##    22     19       1   0.2045  0.0430       0.1355         0.309
##    25     17       2   0.1805  0.0412       0.1154         0.282
##    26     15       1   0.1684  0.0401       0.1056         0.269
##    28     14       1   0.1564  0.0390       0.0959         0.255
##    29     13       1   0.1444  0.0378       0.0864         0.241
##    37     11       2   0.1181  0.0352       0.0659         0.212
##    41      8       1   0.1034  0.0338       0.0545         0.196
##    47      6       1   0.0861  0.0322       0.0414         0.179
##
##                 Stadium=localized
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     0    179       3    0.983 0.00959        0.965        1.000
##     1    176       1    0.978 0.01105        0.956        1.000
##     5    175       2    0.966 0.01345        0.940        0.993
##     6    173       1    0.961 0.01449        0.933        0.990
##     9    172       1    0.955 0.01544        0.926        0.986
##    11    171       1    0.950 0.01633        0.918        0.982
##    12    170       3    0.933 0.01869        0.897        0.970
##    14    167       1    0.927 0.01940        0.890        0.966
##    16    166       1    0.922 0.02007        0.883        0.962
##    18    165       1    0.916 0.02071        0.876        0.958
##    19    164       2    0.905 0.02191        0.863        0.949
##    20    162       2    0.894 0.02302        0.850        0.940
##    22    160       2    0.883 0.02405        0.837        0.931
##    28    142       1    0.876 0.02467        0.829        0.926
##    32    137       2    0.864 0.02592        0.814        0.916
##    34    133       1    0.857 0.02652        0.807        0.911
##    35    131       2    0.844 0.02769        0.792        0.900
##    36    128       2    0.831 0.02878        0.776        0.889
##    38    121       1    0.824 0.02935        0.768        0.884
##    44    104       1    0.816 0.03012        0.759        0.877
##    54     74       1    0.805 0.03167        0.745        0.870
##    57     64       1    0.793 0.03358        0.729        0.861
##    61     55       1    0.778 0.03593        0.711        0.852
##    65     47       1    0.762 0.03879        0.689        0.841
##
##                 Stadium=regional
```

```
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     0     38       2    0.947  0.0362        0.879        1.000
##     2     36       1    0.921  0.0437        0.839        1.000
##     5     35       1    0.895  0.0498        0.802        0.998
##     6     34       2    0.842  0.0592        0.734        0.966
##     7     32       1    0.816  0.0629        0.701        0.949
##     8     31       1    0.789  0.0661        0.670        0.930
##     9     30       1    0.763  0.0690        0.639        0.911
##    10     29       1    0.737  0.0714        0.609        0.891
##    11     28       1    0.711  0.0736        0.580        0.870
##    18     27       2    0.658  0.0770        0.523        0.827
##    23     25       1    0.632  0.0783        0.495        0.805
##    36     24       3    0.553  0.0807        0.415        0.736
##    37     21       1    0.526  0.0810        0.389        0.712
##    39     20       1    0.500  0.0811        0.364        0.687
##    43     17       1    0.471  0.0815        0.335        0.661
##    51     15       1    0.439  0.0819        0.305        0.633
##    54     14       1    0.408  0.0818        0.275        0.604
##    60     12       1    0.374  0.0818        0.244        0.574
```

```r
plot(survfit(Surv(round(Time*12),Event)~Stadium, data = vie1),
     lty=c(3,1,2), mark.time=FALSE,
     ylab = "Probability", xlab = "Survival Time in Month")
legend(0.2,0.4,legend =c("disseminized", "localized","regional"),
            lty=c(3,1,2), title = "Stadium", bty ="y", cex = 0.75)
```



```r
# Test for difference in groups

survdiff(Surv(round(Time),Event)~Stadium)
```

```
## Call:
## survdiff(formula = Surv(round(Time), Event) ~ Stadium)
##
##                         N Observed Expected (O-E)^2/E (O-E)^2/V
## Stadium=disseminized   88       79     26.5    104.39    169.04
## Stadium=localized     179       35     92.8     35.98    138.97
## Stadium=regional       38       23     17.8      1.54      2.08
##
##  Chisq= 184  on 2 degrees of freedom, p= 0
```

```r
#-----------------------------------------------------
#Coxregression
#-----------------------------------------------------
melan.ph<-coxph(Surv(Time,Event)~Age_Diagnosis+MultipleTumor + Stadium, data=
vie1)
```

```
## Warning in coxph(Surv(Time, Event) ~ Age_Diagnosis + MultipleTumor +
## Stadium, : X matrix deemed to be singular; variable 5
```

```r
melan.ph
```

```
## Call:
## coxph(formula = Surv(Time, Event) ~ Age_Diagnosis + MultipleTumor +
##     Stadium, data = vie1)
##
##                       coef exp(coef) se(coef)      z       p
## Age_Diagnosis      0.02989   1.03034  0.00656   4.55 5.3e-06
## MultipleTumor      0.00902   1.00906  0.21763   0.04    0.97
## Stadiumlocalized -2.64414   0.07107  0.21412 -12.35 < 2e-16
## Stadiumregional  -1.41093   0.24392  0.24571  -5.74 9.3e-09
## Stadiumunknown          NA        NA  0.00000     NA      NA
##
## Likelihood ratio test=186  on 4 df, p=0
## n= 305, number of events= 137
```