**Statistical procedures:**

**Example scripts**

**DESCRIPTIVE STATISTICS (ALL VARIABLES)**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(psych)

describe(df1,type=2)

**DESCRIPTIVE STATISTICS (SELECTED VARIABLES)**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(psych)

df11 <- subset(df1, ,select = c(EWB1, EWB2))

describe(df11,type=2)

**DESCRIPTIVE STATISTICS (SELECTED VARIABLES) BY GROUP**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(psych)

describeBy (df1$defp, df1$ncateg, type = 2)

**FREQUENCY DISTRIBUTION TABLE**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(psych)

transform(table(df1$EWB1), cumFreq = cumsum(Freq), pct = prop.table(Freq)\*100)

**HISTOGRAM**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(ggplot2)

ggplot(df1, aes(AGE)) + geom\_histogram(color = "black", fill = "gray", binwidth = 1) + ggtitle("Age of participants") + labs(x="Age (in years)", y="f") + theme\_bw()

**t-TEST FOR ONE MEAN**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(psych)

describe (df1$ageatdeath)

t.test(df1$ageatdeath, mu = 78)

**CONFIDENCE INTERVAL FOR THE MEAN**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(psych)

describe (df1$income)

t.test(df1$income, mu = 0)

**t-TEST FOR INDEPENDENT MEANS**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(psych)

describeBy (df1$defp, df1$gender)

t.test(df1$defp ~ df1$gender, var.equal=TRUE, paired=FALSE)

**t-TEST FOR DEPENDENT MEANS (REPEATED-MEASURES t-TEST)**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(riskgain,riskloss))

df11 <- na.omit(df11)

describe(df11)

t.test(df11$riskgain, df11$riskloss, var.equal=TRUE, paired=TRUE)

**ONE-WAY ANOVA WITH TUKEY POST-HOCS**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(psych)

df11 <- subset(df1, ,select = c(ncateg,moneyr))

df11 <- na.omit(df11)

df11$ncateg.f <- factor(df11$ncateg, labels = c("Control", "Life events", "No life events"))

describeBy (df11$moneyr,df11$ncateg.f)

oneway <- aov(moneyr ~ ncateg.f, data = df11)

summary(oneway)

TukeyHSD(oneway)

**ONE-WAY ANOVA WITH TREND ANALYSIS**

setwd("e:/R output")

df1 <- read.csv("TREND\_NA.csv")

library(psych)

df11 <- subset(df1, ,select = c(FACTORA,DEPVAR))

df11 <- na.omit(df11)

df11$FACTORA.f <- factor(df11$FACTORA, labels = c("Time1","Time2","Time3","Time4"))

describeBy (df11$DEPVAR, df11$FACTORA.f)

linear <- c(-3, -1, 1, 3)

quadratic <- c(1, -1, -1, 1)

cubic <- c(-1, 3, -3, 1)

matrix<- cbind(linear,quadratic,cubic)

contrasts(df11$FACTORA.f) <- matrix

oneway <- aov(DEPVAR ~ FACTORA.f, data = df11)

summary(oneway)

summary(oneway, split=list(FACTORA.f =list("Linear"=1,"Quadratic"=2, "Cubic"=3)))

**Single-factor within-subjects (repeated measures) ANOVA**

setwd("e:/R output")

df1 <- read.csv("SINGLEWS\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(TIME1,TIME2,TIME3,TIME4))

df11 <- na.omit(df11)

df11=stack(df11)

nsubjects= 8

nlevels= 4

df11=data.frame(recall=df11, subj=factor(rep(paste("subj", 1:nsubjects, sep=""), nlevels)))

colnames(df11) = c("DEPVAR","FACTORA","subject")

describeBy (df11$DEPVAR,df11$FACTORA)

singlewf = aov(DEPVAR ~ FACTORA + Error(subject/FACTORA), data= df11)

summary(singlewf)

**Two-factor between-subjects ANOVA**

setwd("e:/R output")

df1 <- read.csv("TWOFACTORBS\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(FACTORA,FACTORB,DEPVAR))

df11 <- na.omit(df11)

df11$FACTORA.f <- factor(df11$FACTORA,labels = c("A1","A2"))

df11$FACTORB.f <- factor(df11$FACTORB,labels = c("B1","B2","B3"))

describeBy (df11$DEPVAR*,* list(df11$FACTORA.f, df11$FACTORB.f))

describeBy (df11$DEPVAR*,* df11$FACTORA.f)

describeBy (df11$DEPVAR*,* df11$FACTORB.f)

twoway <- aov(DEPVAR ~ FACTORA.f \* FACTORB.f, data = df11)

summary(twoway)

**Two-factor between-subjects ANOVA (SIMPLE EFFECTS)**

setwd("e:/R output")

df1 <- read.csv("TWOFACTORBS\_NA.csv")

library(psych)

library(phia)

df11 <- subset(df1, ,select = c(FACTORA,FACTORB,DEPVAR))

df11 <- na.omit(df11)

df11$FACTORA.f <- factor(df11$FACTORA, labels = c("A1","A2"))

df11$FACTORB.f <- factor(df11$FACTORB, labels = c("B1","B2","B3"))

twoway <- aov(DEPVAR ~ FACTORA.f \* FACTORB.f, data = df11)

testInteractions(twoway, fixed="FACTORA.f", across="FACTORB.f", adjustment="none")

**Two-factor between-subjects ANOVA (SIMPLE COMPARISONS)**

setwd("e:/R output")

df1 <- read.csv("TWOFACTORBS\_NA.csv")

library(psych)

library(phia)

df11 <- subset(df1, ,select = c(FACTORA,FACTORB,DEPVAR))

df11 <- na.omit(df11)

df11$FACTORA.f <- factor(df11$FACTORA, labels = c("A1","A2"))

df11$FACTORB.f <- factor(df11$FACTORB, labels = c("B1","B2","B3"))

twoway <- aov(DEPVAR ~ FACTORA.f \* FACTORB.f, data = df11)

b1\_vs\_b2 <- list(FACTORB.f = c(1, -1, 0))

b2\_vs\_b3 <- list(FACTORB.f = c(0, 1, -1))

testInteractions(twoway, custom=c(b1\_vs\_b2), fixed="FACTORA.f", adjustment="none")

testInteractions(twoway, custom=c(b2\_vs\_b3), fixed="FACTORA.f", adjustment="none")

**Two-factor between-subjects ANOVA (MAIN COMPARISONS)**

setwd("e:/R output")

df1 <- read.csv("TWOFACTORBS\_NA.csv")

library(psych)

library(phia)

df11 <- subset(df1, ,select = c(FACTORA,FACTORB,DEPVAR))

df11 <- na.omit(df11)

df11$FACTORA.f <- factor(df11$FACTORA, labels = c("A1","A2"))

df11$FACTORB.f <- factor(df11$FACTORB, labels = c("B1","B2","B3"))

twoway <- aov(DEPVAR ~ FACTORA.f \* FACTORB.f, data = df11)

b1\_vs\_b2 <- list(FACTORB.f = c(1, -1, 0))

b2\_vs\_b3 <- list(FACTORB.f = c(0, 1, -1))

testInteractions(twoway, custom=c(b1\_vs\_b2), adjustment="none")

testInteractions(twoway, custom=c(b2\_vs\_b3), adjustment="none")

**Two-factor MIXED FACTORIAL ANOVA**

setwd("e:/R output")

df1 <- read.csv("TWOMIXED\_NA.csv")

library(psych)

df11 <- subset(df1, ,select = c(TIME1,TIME2,TIME3,TIME4))

df11 <- na.omit(df11)

df11=stack(df11)

nlevelsw= 4

ngroup= 10

ntotal= 30

df11 =data.frame(recall= df11, BETWEEN=factor(rep(c("Group1", "Group2","Group3"),c(ngroup,ngroup,ngroup))),subj=factor(rep(paste("subj", 1:ntotal, sep=""), nlevelsw)))

colnames(df11) = c("DEPVAR","WITHIN","BETWEEN","subject")

describeBy (df11$DEPVAR, list(df11$BETWEEN, df11$WITHIN))

describeBy (df11$DEPVAR,df11$BETWEEN)

describeBy (df11$DEPVAR,df11$WITHIN)

mixed = aov(DEPVAR ~ BETWEEN\*WITHIN + Error(subject/WITHIN), data= df11)

summary(mixed)

**Two-factor WITHIN-SUBJECTS ANOVA**

setwd("e:/R output")

df1 <- read.csv("TWOFACTORWS\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(a1b1,a1b2,a1b3,a2b1,a2b2,a2b3))

df11 <- na.omit(df11)

df11=stack(df11)

ncases= 3

nobs= 6

df11=data.frame(recall=df11,subj=factor(rep(paste("subj", 1: ncases, sep=""), nobs)), WSFACTORA=factor(rep(rep(c("A1","A2"), c(ncases\*3, ncases \*3)))),WSFACTORB=factor(rep(c("B1","B2","B3"), c(ncases, ncases, ncases))))

colnames(df11) = c("DEPVAR","ABCOMB","subject","WSFACTORA","WSFACTORB")

describeBy (df11$DEPVAR, list(df11$WSFACTORA, df11$WSFACTORB))

describeBy (df11$DEPVAR*,* df11$WSFACTORA)

describeBy (df11$DEPVAR*,* df11$WSFACTORB)

twowithin= aov(DEPVAR ~ WSFACTORA \* WSFACTORB + Error(subject/(WSFACTORA \* WSFACTORB)), data=df11)

summary(twowithin)

**THREE-factor BETWEEN-SUBJECTS ANOVA**

setwd("e:/R output")

df1 <- read.csv("THREEFACTORBS\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(FACTORA,FACTORB,FACTORC,DEPVAR))

df11 <- na.omit(df11)

df11$FACTORA.f <- factor(df11$FACTORA, labels = c("A1","A2"))

df11$FACTORB.f <- factor(df11$FACTORB, labels = c("B1","B2","B3"))

df11$FACTORC.f <- factor(df11$FACTORC, labels = c("C1","C2"))

describeBy (df11$DEPVAR*,* list(df11$FACTORA.f, df11$FACTORB.f*,* df11$FACTORC.f))

describeBy (df11$DEPVAR, list(df11$FACTORA.f, df11$FACTORB.f))

describeBy (df11$DEPVAR, list(df11$FACTORA.f, df11$FACTORC.f))

describeBy (df11$DEPVAR, list(df11$FACTORB.f, df11$FACTORC.f))

describeBy (df11$DEPVAR,df11$FACTORA.f)

describeBy (df11$DEPVAR,df11$FACTORB.f)

describeBy (df11$DEPVAR,df11$FACTORC.f)

threeway <- aov(DEPVAR ~ FACTORA.f\* FACTORB.f \* FACTORC.f, data = df11)

summary(threeway)

**PEARSON CORRELATION (ONE CORRELATION)**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

cor.test(~ EWB1 +EWB2,na.action="na.exclude", data = df1)

**PEARSON CORRELATION (CORRELATION MATRIX) – PAIRWISE DELETION**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(EWB1, EWB2, EWB3, EWB4))

corr.test(df11)

**PEARSON CORRELATION (CORRELATION MATRIX) – LISTWISE DELETION**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(psych)

df11 <- subset(df1,,select = c(EWB1, EWB2, EWB3, EWB4))

df11 <- na.omit(df11)

corr.test(df11)

**SCATTERPLOT**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(ggplot2)

ggplot(df1, aes(x=satmath, y=cgpa)) + geom\_point(shape=1) +ggtitle("SAT Math and college GPA") + xlim(200, 800) + ylim(0, 4) + labs(x="SAT Math",y="College GPA") + theme\_bw()

**INTERNAL CONSISTENCY (CRONBACH’S ALPHA)**

setwd("e:/R output")

df1 <- read.csv("EWB\_NA.csv")

library(psych)

df11 <- subset(df1, ,select = c(EWB1,EWB2,EWB3,EWB4,EWB5))

df11 <- na.omit(df11)

alpha(df11)

**PRINCIPAL COMPONENTS ANALYSIS (VARIMAX ROTATION)**

setwd("e:/R output")

df1 <- read.csv("ABCD\_NA.csv")

library(psych)

df11 <- subset(df1, ,select = c(ABCD1, ABCD2,ABCD3,ABCD4,ABCD5, ABCD6, ABCD7,ABCD8,ABCD9,ABCD10,ABCD11))

df11 <- na.omit(df11)

KMO (df11)

cortest.bartlett(df11)

principal(df11, nfactors = 11, rotate = "none")

principal(df11, nfactors = 3, rotate = "varimax")

**PRINCIPAL COMPONENTS ANALYSIS (OBLIQUE ROTATION)**

setwd("e:/R output")

df1 <- read.csv("ABCD\_NA.csv")

library(psych)

library(GPArotation)

df11 <- subset(df1,,select = c(ABCD1, ABCD2,ABCD3,ABCD4, ABCD5, ABCD6,ABCD7,ABCD8,ABCD9,ABCD10,ABCD11))

df11 <- na.omit(df11)

KMO (df11)

cortest.bartlett(df11)

principal(df11, nfactors = 11, rotate = "none")

principal(df11, nfactors = 3, rotate = "oblimin")

**FACTOR ANALYSIS (PRINCIPAL AXIS FACTORING)**

setwd("e:/R output")

df1 <- read.csv("ABCD\_NA.csv")

library(psych)

library(GPArotation)

df11 <- subset(df1,,select = c(ABCD1, ABCD2,ABCD3,ABCD4, ABCD5, ABCD6,ABCD7,ABCD8,ABCD9,ABCD10,ABCD11))

df11 <- na.omit(df11)

KMO (df11)

cortest.bartlett(df11)

fa(df11, nfactors = 11, fm ="pa",SMC = FALSE,rotate = "none")

fa(df11, nfactors = 3, fm = "pa",SMC = FALSE,rotate = "varimax")

**LINEAR REGRESSION**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

linreg *<-* lm(cgpa ~ satmath, data = df1)

summary(linreg)

**MULTIPLE REGRESSION (STANDARD)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

stanmrc <- lm(cgpa ~ satmath + satverb, data = df1)

stanmrcbetas <- lm.beta(stanmrc)

summary(stanmrcbetas)

**MULTIPLE REGRESSION (HIERARCHICAL WITH TWO STEPS)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

library(lmSupport)

df11 <- subset(df1, ,select = c(cgpa*,*satmath*,*satverb))

df11 <- na.omit(df11)

step1 <- lm(cgpa ~ satmath, data = df11)

step1b <- lm.beta(step1)

summary(step1b)

step2 <- lm(cgpa ~ satmath + satverb, data = df11)

step2b <- lm.beta(step2)

summary(step2b)

lm.deltaR2(step1b, step2b)

**MULTIPLE REGRESSION (HIERARCHICAL WITH THREE STEPS)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

library(lmSupport)

df11 <- subset(df1,,select = c(cgpa,hsgpa,satmath,satverb,sshada,sshawm))

df11 <- na.omit(df11)

step1 *<-* lm(cgpa ~ hsgpa, data = df11)

step1b *<-* lm.beta(step1)

summary(step1b)

step2 *<-* lm(cgpa ~ hsgpa + satmath + satverb, data = df11)

step2b *<-* lm.beta(step2)

summary(step2b)

lm.deltaR2(step1b, step2b)

step3 *<-* lm(cgpa ~ hsgpa +satmath +satverb +sshada +sshawm, data = df11)

step3b *<-* lm.beta(step3)

summary(step3b)

lm.deltaR2(step2b, step3b)

**MULTIPLE REGRESSION (MODERATOR VARIABLE)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

library(lmSupport)

df11 <- subset(df1,,select = c(cgpa, satmath,satverb))

df11 <- na.omit(df11)

maineffs <- lm(cgpa ~ satmath + satverb, data = df11)

maineffsb <- lm.beta(maineffs)

summary(maineffsb)

moderator <- lm(cgpa ~ satmath + satverb + satmath \* satverb, data = df11)

moderatorb <- lm.beta(moderator)

summary(moderatorb)

lm.deltaR2(maineffsb, moderatorb)

**MULTIPLE REGRESSION (PORTRAYING MODERATING EFFECT)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

library(lmSupport)

library(ggplot2)

df11 <- subset(df1,,select = c(cgpa, satmath,satverb))

df11 <- na.omit(df11)

transform(table(df11$satverb), cumFreq = cumsum(Freq), pct = prop.table(Freq)\*100)

lowsatverb <- df11 [which(df11$satverb <= 550) ,]

highsatverb <- df11 [which(df11$satverb > 550) ,]

lowsatverbreg <- lm(cgpa ~ satmath, data = lowsatverb)

summary(lowsatverbreg)

highsatverbreg <- lm(cgpa ~ satmath, data = highsatverb)

summary(highsatverbreg)

df11$satverbg <- cut(df11$satverb, breaks=c(200, 550, 800), labels=c("Low SAT Verbal"," High SAT Verbal"))

ggplot(df11, aes(x = satmath, y = cgpa, linetype = satverbg)) + geom\_smooth(method=lm,se=FALSE,color="black", fullrange=TRUE) + xlim(200,800)+ ylim(0,4) + labs(x = "SAT Math", y = "College GPA") + theme\_bw() + theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank()) + theme(legend.title = element\_blank())+ theme(legend.key = [element\_rect](http://docs.ggplot2.org/0.9.2.1/element_rect.html)(color = "white"))

**MULTIPLE REGRESSION (STEPWISE)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

df11 <- subset(df1, ,select = c(cgpa,satmath,satverb,sshada, sshawm,sshata,sshaea))

allvars <- lm(cgpa ~ satmath + satverb + sshada + sshawm + sshata + sshaea, data = df11)

stepmrc<-step(allvars,test="F",direction="both")

stepmrcbetas<- lm.beta(stepmrc)

summary(stepmrcbetas)

**MULTIPLE REGRESSION (BACKWARD)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

df11<- subset(df1,,select = c(cgpa,satmath,satverb,sshada,sshawm, sshata))

allvars <- lm(cgpa ~ satmath + satverb + sshada + sshawm + sshata, data = df11)

backmrc<-step(allvars,test="F",direction="backward")

backmrcbetas<- lm.beta(backmrc)

summary(backmrcbetas)

**MULTIPLE REGRESSION (FORWARD)**

setwd("e:/R output")

df1 <- read.csv("COLLEGE\_NA.csv")

library(lm.beta)

df11 <- subset(df1, ,select = c(cgpa,satmath,satverb,sshada,sshata, sshaea))

df11 <- na.omit(df11)

nullmodel <- lm(cgpa ~ 1, data = df11)

fullmodel <- lm(cgpa~satmath+satverb+sshada+sshata+sshaea,data=df11)

forwmrc <- step(nullmodel, scope = list(lower = nullmodel,upper = fullmodel), test="F", direction = "forward")

forwmrcbetas<- lm.beta(forwmrc)

summary(forwmrcbetas)

**CANONICAL CORRELATION ANALYSIS**

setwd("e:/R output")

df1 <- read.csv("PSYCH\_NA.csv")

library(CCA)

library(candisc)

df11 <- subset(df1,,select = c(fplan1,fplan2,fplan3, fplan4,fplan5,fplan6,instruct, faculty,relfac,tguide, tpsk))

df11 <- na.omit(df11)

major <- df11[, 7:11]

outcome <- df11[,1:6]

cc1 <- cc(major, outcome)

ev <- (1 - cc1$cor^2) # conduct dimension reduction analysis

n <- dim(major)[1]

p <- length(major)

q <- length(outcome)

k <- min(p, q)

m <- n - 3/2 - (p + q)/2

w <- rev(cumprod(rev(ev)))

d1 <- d2 <- f <- vector("numeric", k)

for (i in 1:k) {

 s <- sqrt((p^2 \* q^2 - 4)/(p^2 + q^2 - 5))

 si <- 1/s

 d1[i] <- p \* q

 d2[i] <- m \* s - p \* q/2 + 1

 r <- (1 - w[i]^si)/w[i]^si

 f[i] <- r \* d2[i]/d1[i]

 p <- p - 1

 q <- q - 1}

pv <- pf(f, d1, d2, lower.tail = FALSE)

(dmat <- cbind(WilksL = w, F = f, df1 = d1, df2 = d2, p = pv))

cc1$cor # calculate canonical correlation for each dimension (root)

cc2<- [cancor](http://inside-r.org/r-doc/stats/cancor)(major, outcome) # calculate redundancy index

redundancy(cc2)

s1 <- diag(sqrt(diag(cov(major)))) # standardized coefficients (set1)

s1 %\*% cc1$xcoef

cc3<- comput(major, outcome, cc1) # structure coefficients (set1)

cc3[3:3]

s2 <- diag(sqrt(diag(cov(outcome)))) # standardized coefficients (set2)

s2 %\*% cc1$ycoef

cc3[6:6] # structure coefficients (set2)

**DISCRIMINANT ANALYSIS (TWO GROUPS)**

setwd("e:/R output")

df1 <- read.csv("PSYCH\_NA.csv")

library(MASS)

df11 <- subset(df1,,select = c(doagain,instruct,faculty,relfac,tguide))

df11 <- na.omit(df11)

discrim <- lda(doagain ~ instruct +faculty +relfac +tguide, data = df11, prior = c(1,1)/2)

discrim

predict <- predict(discrim) # predict group membership

table(df11$doagain, predict$class) # classification summary table

classtable <- table(df11$doagain, predict$class)

diag(prop.table(classtable, 1)) # % correct predictions for each group

sum(diag(prop.table(classtable))) # % correct predictions for total sample

**DISCRIMINANT ANALYSIS (THREE GROUPS)**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(MASS)

df11 <- subset(df1,,select = c(ncateg,riskgain,riskloss,moneyr,age))

df11 <- na.omit(df11)

discrim <- lda(ncateg ~ riskgain+riskloss+moneyr+age, data = df11, prior = c(1,1,1)/3)

discrim

predict <- predict(discrim) # predict group membership

table(df11$ncateg, predict$class) # classification summary table

classtable <- table(df11$ncateg, predict$class)

diag(prop.table(classtable, 1)) # % correct predictions for each group

sum(diag(prop.table(classtable))) # % correct predictions for total sample

**CHI-SQUARE TEST OF INDEPENDENCE**

setwd("e:/R output")

df1 <- read.csv("MISC\_NA.csv")

library(gmodels)

df1$nage.f <- factor(df1$nage, labels = c("< 30", "30+"))

df1$ncateg.f <- factor(df1$ncateg, labels = c("Control", "Life events", "No life events"))

CrossTable(df1$nage.f, df1$ncateg.f, digits = 3, prop.r=TRUE, prop.c=TRUE, prop.t=FALSE, prop.chisq=FALSE, chisq=TRUE)