library(xlsx)

data=read.csv('data1.csv')

colnames(data)[4:9] <-c("age","gender","mar","edu","rank","jobt")

GQ=1:522

SQ=1:522

PR=1:522

fan=c(10,13,18,20,22,25,28,29,32,35,37,41,42,47,55,56)#需反向的列数

data[,fan]=6-data[,fan]#进行反向计分

for (i in 1:522) { ##测试i=1

GQ[i]=sum(data[i,10:37]) ##共情

SQ[i]=sum(data[i,38:57]) ##述情

PR[i]=sum(data[i,58:92]) ##职业压力

}

score=data.frame(GQ,SQ,PR)##整合得分矩阵

#异常值检查

par(mfrow=c(1,3))

boxplot(GQ,main = "共情能力",cex.axis=2,cex.main=2)

boxplot(SQ,main = "述情障碍",cex.axis=2,cex.main=2,cex=1.5)##1个异常值

boxplot(PR,main = "职业压力",cex.axis=2,cex.main=2,cex=1.5)##14个异常值

par(mfrow=c(1,1))

outvalue1=boxplot(GQ)$out

OUT1=which(GQ %in% outvalue1)##GQ异常值行

outvalue2=boxplot(SQ)$out

OUT2=which(SQ %in% outvalue2)##SQ异常值行

outvalue3=boxplot(PR)$out

OUT3=which(PR %in% outvalue3)##PR异常值行

OUT=c(OUT1,OUT2,OUT3)#整合异常值行

OUTo=OUT[order(OUT)]#排序

OUTf=OUTo[!duplicated(OUTo)]#去重

score=score[-OUTf,]##得分矩阵排除异常值

data=data[-OUTf,]##原始数据排除异常值

GQ=score[,1]

SQ=score[,2]

PR=score[,3]

save(data,file="F:/data.Rdata")##储存一下关键的结果

save(score,file="F:/score.Rdata")

##小条目

F1=1:508

F2=F1

F3=F1

R1=F1

R2=F1

R3=F1

R4=F1

R5=F1

T1=F1

T2=F1

T3=F1

for (i in 1:508) { ##测试i=1

T1[i]=sum(data[i,c(1,3,5,6,7,9,10,11,16,17,20,21,23,25,26)+9])

T2[i]=sum(data[i,c(12,22,24,27,28)+9])

T3[i]=sum(data[i,c(2,4,8,13,14,15,18,19)+9])

}

for (i in 1:508) { ##测试i=1

F1[i]=sum(data[i,c(1,3,6,7,9,13,14)+37]) ##共情

F2[i]=sum(data[i,c(2,4,11,12,17)+37])

F3[i]=sum(data[i,c(5,8,10,15,16,18,19,20)+37])

R1[i]=sum(data[i,1:7+57])

R2[i]=sum(data[i,8:12+57])

R3[i]=sum(data[i,13:15+57])

R4[i]=sum(data[i,16:26+57])

R5[i]=sum(data[i,27:35+57])

}

nscore=data.frame(T1,T2,T3,F1,F2,F3,R1,R2,R3,R4,R5)

nmeanen=1

nstden=1

npercent=1

inter=1

full=c(75,25,40,35,25,40,35,25,15,55,45)

for (i in 1:11) {

nmeanen[i]=mean(nscore[,i])

nstden[i]=sd(nscore[,i])

npercent[i]=mean(nscore[,i]/full[i])

}

tiaomu=data.frame(nmeanen,nstden,npercent)

write.xlsx(tiaomu,file = "F:/小分.xlsx")

cor(nscore)

write.xlsx(cor(nscore),file = "F:/小相关系数.xlsx")

##描述性统计

library(scatterplot3d)

scatterplot3d(score,highlight.3d=TRUE,type="h",xlab = "共情能力",ylab = "述情障碍"，zlab="职业压力")

meanen=1

varen=1

sden=1

for (i in 1:3) {

meanen[i]=mean(score[,i])

sden[i]=sd(score[,i])

}

meanen

sden##求得均值和方差

Miaoshu=data.frame(meanen,varen,stden)

write.xlsx(Miaoshu,file = "F:/Miaoshu.xlsx")

table(data$age)

table(data$gender)

table(data$mar)

table(data$edu)

table(data$rank)

table(data$jobt)

#求相关系数

cor(score)

cor.test(score[,1],score[,2])$p.value

cor.test(score[,1],score[,3])$p.value

cor.test(score[,2],score[,3])$p.value

##次要变量回归及相关系数

attach(data)

sqf=data.frame(SQ,age,edu,rank,jobt)

cor(sqf)

cor.test(sqf$SQ,sqf$age)

cor.test(sqf$SQ,sqf$edu)

cor.test(sqf$SQ,sqf$rank)

cor.test(sqf$SQ,sqf$jobt)

gqf=data.frame(GQ,age,edu,rank,jobt)

cor(gqf)

cor.test(gqf$GQ,gqf$age)

cor.test(gqf$GQ,gqf$edu)

cor.test(gqf$GQ,gqf$rank)

cor.test(gqf$GQ,gqf$jobt)

prf=data.frame(PR,age,edu,rank,jobt)

cor(prf)

cor.test(prf$PR,prf$age)

cor.test(prf$PR,prf$edu)

cor.test(prf$PR,prf$rank)

cor.test(prf$PR,prf$jobt)

allg=lm(PR~SQ+GQ+age+edu+rank+jobt)

summary(allg)

detach(data)

##输出表格

Xiao=matrix(1:24,nrow=3)

for (i in 1:3) {

Xiao[i,1]=cor(score[,i],data$age)

Xiao[i,2]=cor.test(score[,i],data$age)$p.value

Xiao[i,3]=cor(score[,i],data$edu)

Xiao[i,4]=cor.test(score[,i],data$edu)$p.value

Xiao[i,5]=cor(score[,i],data$rank)

Xiao[i,6]=cor.test(score[,i],data$rank)$p.value

Xiao[i,7]=cor(score[,i],data$jobt)

Xiao[i,8]=cor.test(score[,i],data$jobt)$p.value

}

write.xlsx(Xiao,file = "F:/小变量相关性.xlsx")

##假设检验年龄对述情

a=sqf$SQ[sqf$age==1|sqf$age==2] ##分组三组

b=sqf$SQ[sqf$age==3]

c=sqf$SQ[sqf$age==4|sqf$age==5|sqf$age==6]

d=c(a,b,c)

e=factor(c(rep(1,193),rep(2,171)))

f=factor(c(rep(1,193),rep(2,158)))

g=factor(c(rep(1,171),rep(2,158)))

bartlett.test(d[c(1:364)]~e) ##方差齐性检验

t.test(a,b,paired = FALSE) ##0.1901(<30 30-40)

bartlett.test(d[c(1:193,365:522)]~f)

t.test(a,c,paried = FALSE) ##0.7648(<30 40<)

bartlett.test(d[c(194:522)]~g)

t.test(b,c,paired = FALSE) ##0.1208(30-40 40<)

##假设检验年龄对共情

aa=gqf$GQ[gqf$age==1|gqf$age==2] ##分组三组

bb=gqf$GQ[gqf$age==3]

cc=gqf$GQ[gqf$age==4|gqf$age==5|gqf$age==6]

dd=c(aa,bb,cc)

ee=factor(c(rep(1,193),rep(2,171)))

ff=factor(c(rep(1,193),rep(2,158)))

gg=factor(c(rep(1,171),rep(2,158)))

bartlett.test(dd[c(1:364)]~ee) ##方差齐性检验

t.test(aa,bb,paired = FALSE) ##0.09628(<30 30-40)

bartlett.test(dd[c(1:193,365:522)]~ff)

t.test(aa,cc,paried = FALSE) ##2.397e-07(<30 40<)

bartlett.test(dd[c(194:522)]~gg)

t.test(bb,cc,paired = FALSE) ##0.001107(30-40 40<)

boxplot(aa,bb,cc,names=c("30岁以下","30-40岁","40岁以上"),main="各年龄组共情能力",cex.axis=1.25)##箱线图

lines(1:3,c(mean(aa),mean(bb),mean(cc)),lty=5,col="blue")######################

##假设检验年龄对职业

aaa=prf$PR[prf$age==1|prf$age==2] ##分组三组

bbb=prf$PR[prf$age==3]

ccc=prf$PR[prf$age==4|prf$age==5|prf$age==6]

ddd=c(aaa,bbb,ccc)

eee=factor(c(rep(1,193),rep(2,171)))

fff=factor(c(rep(1,193),rep(2,158)))

ggg=factor(c(rep(1,171),rep(2,158)))

bartlett.test(ddd[c(1:364)]~eee) ##方差齐性检验

t.test(aaa,bbb,paired = FALSE) ##0.1462(<30 30-40)

bartlett.test(ddd[c(1:193,365:522)]~fff)

t.test(aaa,ccc,paried = FALSE) ##0.9376(<30 40<)

bartlett.test(ddd[c(194:522)]~ggg)

t.test(bbb,ccc,paired = FALSE) ##0.1328(30-40 40<)

##假设检验婚姻对述情

allm=data.frame(SQ,GQ,PR,data$mar)

attach(allm)

h=SQ[data.mar==1]

i=SQ[data.mar==2]

j=c(h,i)

k=factor(c(rep(1,124),rep(2,388)))

bartlett.test(j~k)

t.test(h,i,paired = FALSE) ##0.6326

boxplot(h,i,names=c("未婚","已婚"),main="述情能力",cex.axis=1.25)##箱线图婚姻

mean(h)

mean(i)

sd(h)

sd(i)

##假设检验婚姻对共情

hh=GQ[data.mar==1]

ii=GQ[data.mar==2]

jj=c(hh,ii)

kk=factor(c(rep(1,124),rep(2,388)))

bartlett.test(jj~kk)

t.test(hh,ii,paired = FALSE) ##0.002944

boxplot(hh,ii,names=c("未婚","已婚"),main="共情能力",cex.axis=1.25)##箱线图婚姻

lines(1:2,c(mean(hh),mean(ii)),lty=5,col="blue")##############################

mean(hh)

mean(ii)

sd(hh)

sd(ii)

##假设检验婚姻对职业压力

hhh=PR[data.mar==1]

iii=PR[data.mar==2]

mean(hhh)

mean(iii)

sd(hhh)

sd(iii)

jjj=c(hhh,iii)

kkk=factor(c(rep(1,124),rep(2,388)))

bartlett.test(jjj~kkk)

t.test(hhh,iii,paired = FALSE) ##0.6561

boxplot(hhh,iii,names=c("未婚","已婚"),main="职业压力")##箱线图婚姻

detach(allm)

##假设检验学历对述情

alle=data.frame(SQ,GQ,PR,data$edu)

attach(alle)

l=SQ[data.edu==1|data.edu==2]

m=SQ[data.edu==3]

n=SQ[data.edu==4]

o=c(l,m,n)

p=factor(c(rep(1,25),rep(2,404)))

q=factor(c(rep(1,25),rep(2,93)))

r=factor(c(rep(1,404),rep(2,93)))

bartlett.test(o[c(1:429)]~p)

t.test(l,m,paired = FALSE) ##0.0008615(<本科 本科)

bartlett.test(o[c(1:25,430:522)]~q)

t.test(l,n,paired = FALSE) ##1.651e-05(<本科 硕士)

bartlett.test(o[c(26:522)]~r)

t.test(m,n,paired = FALSE) ##0.6669 (本科 硕士)

boxplot(l,m,n,names=c("本科以下","本科","硕士"),main="不同受教育程度组的述情障碍",cex.axis=1.25)##箱线图

lines(1:3,c(mean(l),mean(m),mean(n)),lty=5,col="blue")

##假设检验学历对共情

ll=GQ[data.edu==1|data.edu==2]

mm=GQ[data.edu==3]

nn=GQ[data.edu==4]

oo=c(ll,mm,nn)

pp=factor(c(rep(1,25),rep(2,404)))

qq=factor(c(rep(1,25),rep(2,93)))

rr=factor(c(rep(1,404),rep(2,93)))

var.test(oo[c(1:429)]~pp) ##方差非齐性 (<本科 本科)

bartlett.test(oo[c(1:25,430:522)]~qq)

t.test(ll,nn,paired = FALSE) ##0.3491(<本科 硕士)

bartlett.test(oo[c(26:522)]~rr)

t.test(mm,nn,paired = FALSE) ##0.4709(本科 硕士)

boxplot(ll,mm,nn,names=c("本科以下","本科","硕士"),main="不同受教育程度组的共情能力")##箱线图

##假设检验学历对职业压力

lll=PR[data.edu==1|data.edu==2]

mmm=PR[data.edu==3]

nnn=PR[data.edu==4]

ooo=c(lll,mmm,nnn)

ppp=factor(c(rep(1,25),rep(2,404)))

qqq=factor(c(rep(1,25),rep(2,93)))

rrr=factor(c(rep(1,404),rep(2,93)))

bartlett.test(ooo[c(1:429)]~ppp)

t.test(lll,mmm,paired = FALSE) ##0.3686(<本科 本科)

bartlett.test(ooo[c(1:25,430:522)]~qqq)

t.test(lll,nnn,paired = FALSE) ##0.3463(<本科 硕士)

bartlett.test(ooo[c(26:522)]~rrr)

t.test(mmm,nnn,paired = FALSE) ##0.8607(本科 硕士)

##样本量太小，造成误差

######最终回归

ma=data$mar-1#婚姻虚拟变量

ge=data$gender-1#性别虚拟变量

totalline=lm(PR~GQ+SQ+data$age+data$edu+ma+ge+0)

summary(totalline)

plot()

###最终回归修正

retotalline=lm(SQ~PR+GQ+data$edu+(data$gender-1)+0)

summary(retotalline)