R_Analyze

AUTHOR
ZhongKai Han, WenTing Huang

load spss data (.saw)

加载spss进行分析

```
library(memisc)
```

```
Loading required package: lattice

Loading required package: MASS


Attaching package: 'memisc'

The following objects are masked from 'package:stats':

    contr.sum, contr.treatment, contrasts

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

    as.array
```

```
data0 = as.data.set(spss.system.file("254359000_36_36.sav"))
```

```
File character set is 'UTF-8'.

Converting character set to UTF-8.
```

```
data = as.data.frame(data0)
#str(data)
#knitr::kable(data)
```

描述性分析性别,年龄,输入时间

```
#table(data["totalseconds"],data["gender"],data["age"])
#table(data$totalseconds,data$gender,data$age)
options(digits=10)
options(scipen = 999)
class(data$age)
```

```
[1] "ordered" "factor"
```

```
levels(data$age)
```

```
 [1] "26" "29" "30" "31" "35" "36" "37" "38" "39" "40" "41" "42" "43" "45" "46"
[16] "48" "50"
```

```
data$index <- as.numeric(as.character(data$index))
#data$gender <- as.character(data$gender)
data$gender <- as.numeric(data$gender)-1
data$age <- as.numeric(as.character(data$age))
data$totalseconds <- as.numeric(as.character(data$totalseconds))
#sd(data$totalseconds)
#summary(data$gender)
summary(data$totalseconds, digits = 4)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
   25.0    59.5   104.0   113.8   140.0   271.0 
```

```
data$totalseconds
```

```
 [1]  84  55 100 252 113 230  89 137  54  25 271 130  53 257  70 157  61  93 108
[20] 128  53 137 111 151  55  96 138  34 109  46  63  51  80 211 146 147
```

```
#sd(data$totalseconds)
#sd(data$totalseconds)/sqrt(length(data$totalseconds))
#summary(data$age, digits = 4)
#sd(data$age)
#sd(data$age)/sqrt(length(data$age))
#mean(as.numeric(as.character(data$age)))
#median(as.numeric(as.character(data$age)))
#summary(data$age, digits = 4)
#table(data$gender,data$age)
#table(data$gender)
#table(data$age)


#data.frame( as.numeric(data$totalseconds),as.numeric(data$gender),as.numeric(data$age
#  summary()
#summary( as.numeric(data$totalseconds),as.numeric(data$gender),as.numeric(data$age) )
```

离差平方和 (Sum of Squares of Deviations, SS)

衡量数据离散程度的指标，它是每个数据点与其平均值之间的距离的平方和。

$$SS = \sum_{i=1}^{n}(x_i - \bar{x})^2$$

方差(Var, Variance)

方差是数据集中每个数据点与样本均值的偏差的平方的平均值。

$$\text{总体}Var = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{SS}{n}\text{样本}Var = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{SS}{n-1}$$

```
var(data$age)
```

```
[1] 34.02857143
```

```
var(data$totalseconds)
```

```
[1] 4182.821429
```

-

标准差(Population Standard Deviation,PSD | Standard Deviation,SD)

标准差表示样本数据的离散程度。它是每个数据点与样本均值的偏差的平方的平均值的平方根。 标准差中要注意: 总体标准差,样本标准差 总体标准差是整个总体数据计算出来,这需要所有样本 样本标准差是从总体中抽取出来的,通常实际应用都使用样本标准差

$$总体:PSD=\sqrt{Var}=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}, 样本:SD=\sqrt{Var}=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$$

```
sd(data$age)
```

```
[1] 5.83340136
```

```
sd(data$totalseconds)
```

```
[1] 64.67473563
```

标准误(Standard Error of the Mean，SEM)

表示样本均值估计的精度。它是样本标准差除以样本大小的平方根。

$$总体SEM=\frac{SD}{\sqrt{n}}=\frac{\sqrt{Var}}{\sqrt{n}}=\frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}}{\sqrt{n}}=(\sum_{i=1}^{n}(x_i-\bar{x})^2*n^{-1})^{\frac{1}{2}}*n^{-\frac{1}{2}}$$

$$样本SEM=\frac{SD}{\sqrt{n-1}}=\frac{\sqrt{Var}}{\sqrt{n-1}}=\frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n-1}}=(\sum_{i=1}^{n}(x_i-\bar{x})^2*(n-1)^{-1})^{\frac{1}{2}}*(n-1)^{-\frac{1}{2}}$$

```
sd(data$age)/sqrt(length(data$totalseconds))
```

```
[1] 0.97223356
```

```
sd(data$totalseconds)/sqrt(length(data$totalseconds))
```

```
[1] 10.7791226
```

协方差

协方差衡量的是两个变量的平均离差程度。

$$Cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

标准误差(Standard Error, SE)

用来衡量样本统计量与总体参数之间的差异的一种估计量。在回归分析中，标准误差用于衡量回归系数估计值与真实参数之间的差异。

按性别统计

```
table(data$gender)
```

```

 0  1 
17 19 
```

```
summary(data$gender, digits = 4)
```

```
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
0.0000  0.0000  1.0000  0.5278  1.0000  1.0000 
```

按性别分组查看年龄分布

```
library(dplyr)
```

```

Attaching package: 'dplyr'

The following objects are masked from 'package:memisc':

    collect, recode, rename, syms

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

    select

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
summary(data$age, digits = 4)
```

```
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 26.00   35.00   38.00   37.50   40.25   50.00 
```

```
#data[data$gender=="男",]
#data[data$gender == "男",]$age
summary(data[data$gender=="男",]$age, digits = 4)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

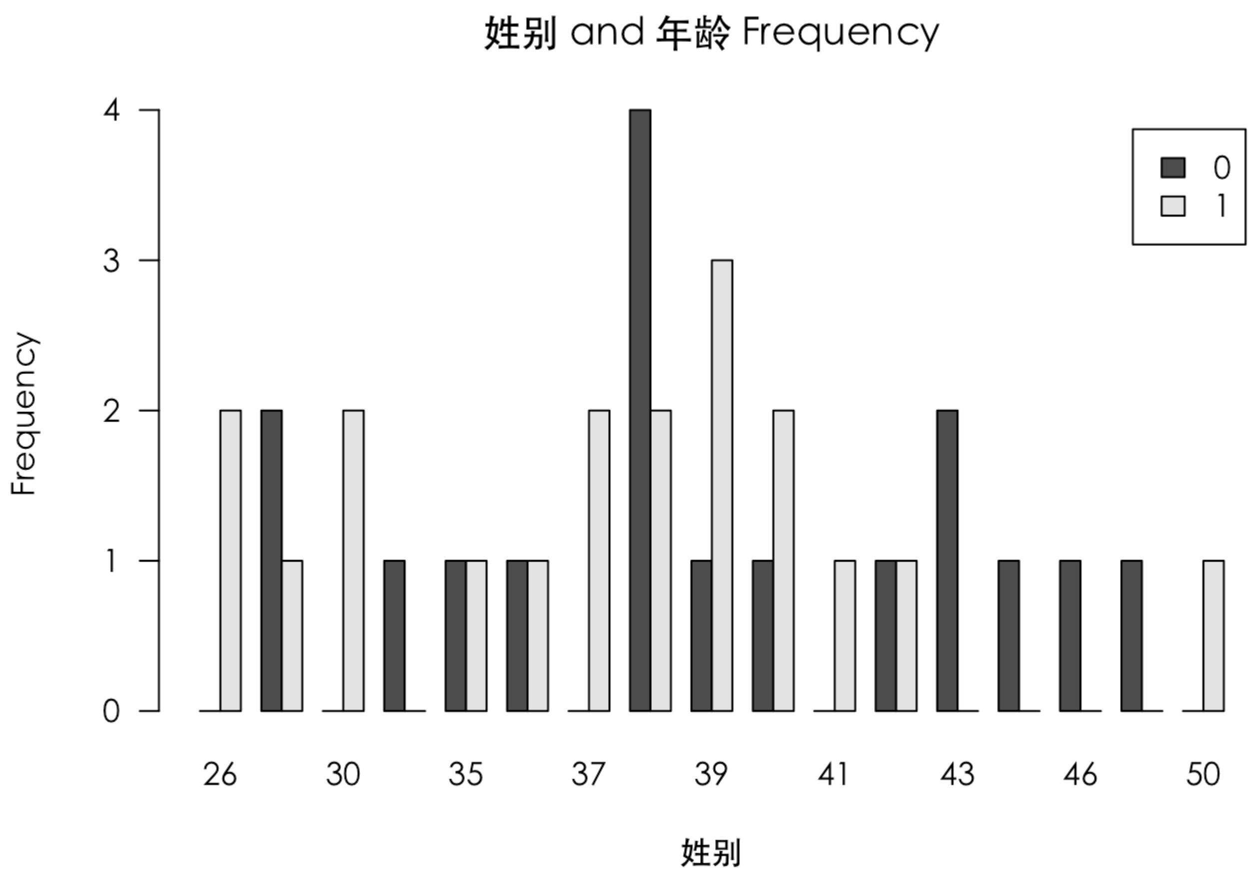
```
summary(data[data$gender=="女",]$age, digits = 4)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
#summary(data[data$age=="1",], digits = 4)
table(data$gender,data$age)
```

```

    26 29 30 31 35 36 37 38 39 40 41 42 43 45 46 48 50
  0  0  2  0  1  1  1  0  4  1  1  0  1  2  1  1  1  0
  1  2  1  2  0  1  1  2  2  3  2  1  1  0  0  0  0  1
```

```
#addmargins(prop.table(table_),2)*100

par(family = "STHeiti")
barplot(table(data$gender,data$age), beside = TRUE, legend = TRUE,
        main = "姓别 and 年龄 Frequency",
        xlab = "姓别",
        ylab = "Frequency",
        las=1)
```




```
summarise(group_by(data,gender),count=n(),percentage = count / sum(count),min_var=min(
```

```
# A tibble: 2 × 6
  gender count percentage min_var median_var max_var
   <dbl> <int>      <dbl>   <dbl>      <dbl>   <dbl>
1      0    17          1      29         38      48
2      1    19          1      26         38      50
```

```
#summarise(group_by(data,gender),count=n(),min_var=min(age),median_var=median(age),max

# 打印输出结果
#print(cross_table_percentage_with_margins)
#knitr::kable( table(data$gender,data$age) )

#group_by(data,gender, age) %>%
#   summarise(count = n()) %>%
#   mutate(percentage = count / sum(count) * 100)
```

T检验 T-test

独立样本T检验 Independent Samples t-test

提出假设：首先，提出零假设（H0）和备择假设（H1）。零假设通常是两组样本的均值相等，备择假设则是两组样本的均值不相等。 计算差异：计算两组样本的均值差异，这是独立样本 t 检验的核心。如果差异较大，那么可能存在显著的差异；如果差异较小，那么可能不存在显著差异。 计算 t 统计量：根据

样本数据计算 t 统计量。t 统计量表示了两组样本均值之间的差异相对于它们的方差的大小。如果 t 统计量越大，表示两组样本的均值差异越显著。 确定显著性水平：选择显著性水平（通常为 0.05），这是决定是否拒绝零假设的标准。 做出决策：将计算得到的 t 统计量与 t 分布的临界值进行比较，根据显著性水平判断是否拒绝零假设。如果计算得到的 t 统计量大于临界值，就拒绝零假设，认为两组样本的均值存在显著差异；否则，接受零假设，认为两组样本的均值没有显著差异。 ***对正态性假设较为敏感。当样本来自非正态分布时，Levene's Test 的结果可能不准确。

计算过程:

先做 Levene's Test通过计算两组样本的绝对离差平方和的平均值之比来检验方差是否相等。该比值称为Levene统计量。 Levene's Test的结果是一个p值。如果p值小于显著性水平（通常为0.05），则可以拒绝原假设，即两组样本的方差不相等。 如果两组样本的方差不相等 则使用Welch's t检验,否则使用传统t检验 #### !! Levene's Test 1.计算每个样本的绝对离差平方和 2.计算总体绝对离差平方和,计算所有样本的绝对离差平方和之总和 3.计算组间离差平方和 4.计算 Levene 统计量 5.计算p值,在自由度为 k - 1 和 N - k 的 F 分布下，计算 Levene 统计量的 p 值。

$$
\begin{aligned}
SSW_i(\text{第}i\text{组的绝对离差平方和}) &= \sum (x_i - \bar{x}_i)^2 \\
SSW &= \sum (SSW_i) \\
SSB = \sum n_i(\text{每组样本数})(\bar{x_i} - \bar{x}(\text{所有组均值总和}))^2 \\
F = \frac{(SSB/(k(\text{总组数}) - 1))}{(SSW/(N(\text{总样本数, 每组样本之和}) - k))} \\
or \\
W = \frac{N-k}{k-1} * \frac{\sum_{i=1}^{k} N_i(\bar{Z_i} - \bar{Z_{all}})^2}{\sum_{i=1}^{k}\sum_{j=1}^{N_i}(Z_{ij} - \bar{Z}_i)^2}
\end{aligned}
$$

```
library(car)
```

```
Loading required package: carData


Attaching package: 'car'

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

    recode

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

    recode
```

```
library(onewaytests)
# 创建示例数据
groupA <- c(23, 25, 27, 29, 31)
groupB <- c(22, 26, 28, 30, 32)
# 执行Levene's Test

# 输出测试结果
#print(levene_test)
getLeveneTest <-function(data0,data1){
```

```
  #levene_test <- leveneTest(data0, data1)
  #print(levene_test)
  n0 <- length(data0)
  n1 <- length(data1)
  N <- n0+n1
  k = 2
  mean0 <- mean(data0)
  mean1 <- mean(data1)
  mean_A <- (mean0+mean1)/2
  SSW0 <-sum((data0-mean0)^2)
  SSW1 <-sum((data1-mean1)^2)
  SSW <- SSW0+SSW1
  SSB <- n0*(mean0-mean_A)^2 + n1*(mean1-mean_A)^2
  W = (N-k)/(k-1)*SSB/SSW
  cat(W,pf(W, k - 1, N - k,lower.tail=T,log.p=T),"\n")
  #计算每组总离差平方和

}
for(col_name in colnames(data)){
  if(col_name!="index"&&col_name!="gender"){
    data0 <- data[data$gender=="0",][[col_name]]
    data1 <- data[data$gender=="1",][[col_name]]
    cat(col_name,":")
    getLeveneTest(data0,data1)
  }
}
```

```
totalseconds :0.0838266055 -1.486934706
age :1.396181788 -0.2817833004
Q3_Row1 :1.464073971 -0.2673940725
Q3_Row2 :0.00006244189435 -5.07379101
Q3_Row3 :0.2867271334 -0.9058728088
Q3_Row4 :0.1911253851 -1.092866649
Q3_Row5 :0.1516857341 -1.201836674
Q3_Row6 :0.5978919017 -0.5882942547
Q3_Row7 :2.760899055 -0.1118131793
Q4_Row1 :0.164239275 -1.164181668
Q4_Row2 :1.085054731 -0.3637285917
Q4_Row3 :13.15140427 -0.0009312929049
Q4_Row4 :0.0372944941 -1.883978141
Q4_Row5 :0.3155156534 -0.8627484673
Q4_Row6 :1.111511239 -0.3555162193
Q4_Row7 :1.889265717 -0.1963521985
Q5_Row1 :0.4939191345 -0.6674184285
Q5_Row2 :0.4097299222 -0.7473858671
Q5_Row3 :0.7923272036 -0.4774759305
Q5_Row4 :0.02113089415 -2.165272272
Q5_Row5 :0.7928771633 -0.4772124864
Q5_Row6 :3.735615963 -0.06360928168
Q5_Row7 :0.1319357364 -1.268269732
depression :0.7236006097 -0.5123543116
anxiety :1.991791024 -0.1830093456
pressure :0.3794953916 -0.7808336662
Zdepression :0.7236006097 -0.5123543116
```

```
Zanxiety :1.991791024 -0.1830093456
Zage :1.396181788 -0.2817833004
ZageXZdepression :0.3595587377 -0.8045856984
ZageXZanxiety :0.2630909099 -0.945002196
```

传统t检验

1.计算样本均值 2.计算样本标准差 3.计算样本方差 4.计算标准误差 5.计算t统计量:独立样本t统计量(t-statistic) 6.确定自由度 (Degrees of Freedom, df)

$$4.\ SE_A = \frac{S_A}{\sqrt{n_A}}, SE_B = \frac{S_B}{\sqrt{n_B}} 5.\ t = \frac{(\bar{X}_A - \bar{X}_B)}{\sqrt{SE_A^2 + SE_B^2}} 6.\ df = n_A + n_B - 2$$

```
#str(data)
#xx <- data[data$gender=="0",]$age
#yy <- data[data$gender=="1",]$age
#t.test(xx,yy)
#t.test(data[data$gender=="0",], data[data$gender=="1",] )
#t.test(data[data$gender=="0",]$age, data[data$gender=="1",]$age, var.equal=T)
#t.test(data[data$gender=="0",]$age, data[data$gender=="1",]$age, )
#result =t.test(data[data$gender=="0",]$age, data[data$gender=="1",]$age, )
for(col_name in colnames(data)){
  if(col_name!="index"&&col_name!="gender"){
    mean0 <- mean(data[data$gender=="0",][[col_name]])
    mean1 <- mean(data[data$gender=="1",][[col_name]])
    for(i in 1:0){
      result = t.test(data[data$gender=="0",][[col_name]], data[data$gender=="1",][[co
    cat(col_name,":","method",result$method,"t",result$statistic,"df",result$parameter
    #print(result)
    #result = t.test(data[data$gender=="0",][[col_name]], data[data$gender=="1",][[col
    #print(result)
    }

  }
}
```

```
totalseconds : method  Two Sample t-test t -0.2890810983 df 34 Sig.(2-tailed)
0.7742746954 stderr 21.87994108 Mean Difference -6.325077399 estimate 110.4117647
116.7368421 null.value 0 conf.int -50.79046753 38.14031273

totalseconds : method Welch Two Sample t-test t -0.2929262235 df 33.52341782 Sig.(2-
tailed) 0.7713841411 stderr 21.59273186 Mean Difference -6.325077399 estimate
110.4117647 116.7368421 null.value 0 conf.int -50.22977826 37.57962346

age : method  Two Sample t-test t 1.179776498 df 34 Sig.(2-tailed) 0.2462783068 stderr
1.936663195 Mean Difference 2.284829721 estimate 38.70588235 36.42105263 null.value 0
conf.int -1.650943422 6.220602865

age : method Welch Two Sample t-test t 1.184407313 df 33.92851333 Sig.(2-tailed)
0.2444828842 stderr 1.929091197 Mean Difference 2.284829721 estimate 38.70588235
36.42105263 null.value 0 conf.int -1.635859551 6.205518994
```

```
Q3_Row1 : method  Two Sample t-test t 1.208120534 df 34 Sig.(2-tailed) 0.2353371224
stderr 0.3869582935 Mean Difference 0.4674922601 estimate 2.941176471 2.473684211
null.value 0 conf.int -0.3189016072 1.253886127

Q3_Row1 : method Welch Two Sample t-test t 1.222188032 df 33.72719989 Sig.(2-tailed)
0.2301096077 stderr 0.3825043675 Mean Difference 0.4674922601 estimate 2.941176471
2.473684211 null.value 0 conf.int -0.3100818001 1.24506632

Q3_Row2 : method  Two Sample t-test t 0.007889814475 df 34 Sig.(2-tailed) 0.9937510172
stderr 0.392401525 Mean Difference 0.003095975232 estimate 2.529411765 2.526315789
null.value 0 conf.int -0.7943598694 0.8005518199

Q3_Row2 : method Welch Two Sample t-test t 0.008046044263 df 32.22608723 Sig.(2-
tailed) 0.9936298587 stderr 0.384782277 Mean Difference 0.003095975232 estimate
2.529411765 2.526315789 null.value 0 conf.int -0.7804642783 0.7866562287

Q3_Row3 : method  Two Sample t-test t 0.5346420981 df 34 Sig.(2-tailed) 0.5963766567
stderr 0.3879798117 Mean Difference 0.2074303406 estimate 2.470588235 2.263157895
null.value 0 conf.int -0.5810395015 0.9959001826

Q3_Row3 : method Welch Two Sample t-test t 0.5392086114 df 33.95790883 Sig.(2-tailed)
0.5932609446 stderr 0.3846940427 Mean Difference 0.2074303406 estimate 2.470588235
2.263157895 null.value 0 conf.int -0.5743977107 0.9892583918

Q3_Row4 : method  Two Sample t-test t 0.4365037135 df 34 Sig.(2-tailed) 0.6652308149
stderr 0.3830039868 Mean Difference 0.1671826625 estimate 2.588235294 2.421052632
null.value 0 conf.int -0.6111750867 0.9455404118

Q3_Row4 : method Welch Two Sample t-test t 0.4421427597 df 33.57532517 Sig.(2-tailed)
0.6612209455 stderr 0.3781191908 Mean Difference 0.1671826625 estimate 2.588235294
2.421052632 null.value 0 conf.int -0.6016061531 0.9359714782

Q3_Row5 : method  Two Sample t-test t 0.3888670316 df 34 Sig.(2-tailed) 0.6997992343
stderr 0.3980763321 Mean Difference 0.1547987616 estimate 2.470588235 2.315789474
null.value 0 conf.int -0.6541896785 0.9637872018

Q3_Row5 : method Welch Two Sample t-test t 0.3931022578 df 33.80146068 Sig.(2-tailed)
0.6967123924 stderr 0.393787516 Mean Difference 0.1547987616 estimate 2.470588235
2.315789474 null.value 0 conf.int -0.645646932 0.9552444552

Q3_Row6 : method  Two Sample t-test t 0.7720405147 df 34 Sig.(2-tailed) 0.4454238511
stderr 0.3248197329 Mean Difference 0.2507739938 estimate 1.882352941 1.631578947
null.value 0 conf.int -0.409339125 0.9108871126

Q3_Row6 : method Welch Two Sample t-test t 0.7815222048 df 33.65534851 Sig.(2-tailed)
0.4399589106 stderr 0.3208789108 Mean Difference 0.2507739938 estimate 1.882352941
1.631578947 null.value 0 conf.int -0.4015764758 0.9031244634

Q3_Row7 : method  Two Sample t-test t 1.659029162 df 34 Sig.(2-tailed) 0.1063081444
stderr 0.2929834643 Mean Difference 0.4860681115 estimate 2.117647059 1.631578947
null.value 0 conf.int -0.1093459253 1.081482148

Q3_Row7 : method Welch Two Sample t-test t 1.634552264 df 29.89991325 Sig.(2-tailed)
0.1126295104 stderr 0.2973707982 Mean Difference 0.4860681115 estimate 2.117647059
```

1.631578947 null.value 0 conf.int -0.1213293379 1.093465561

Q4_Row1 : method Two Sample t-test t 0.4046385593 df 34 Sig.(2-tailed) 0.688277676 stderr 0.3596069445 Mean Difference 0.1455108359 estimate 2.882352941 2.736842105 null.value 0 conf.int -0.5852984025 0.8763200743

Q4_Row1 : method Welch Two Sample t-test t 0.4079751979 df 33.96941279 Sig.(2-tailed) 0.6858519988 stderr 0.3566658872 Mean Difference 0.1455108359 estimate 2.882352941 2.736842105 null.value 0 conf.int -0.579345496 0.8703671678

Q4_Row2 : method Two Sample t-test t 1.040050863 df 34 Sig.(2-tailed) 0.3056561158 stderr 0.3125591364 Mean Difference 0.3250773994 estimate 2.588235294 2.263157895 null.value 0 conf.int -0.3101191893 0.9602739881

Q4_Row2 : method Welch Two Sample t-test t 1.047773663 df 33.99176353 Sig.(2-tailed) 0.3021362299 stderr 0.3102553641 Mean Difference 0.3250773994 estimate 2.588235294 2.263157895 null.value 0 conf.int -0.3054429882 0.955597787

Q4_Row3 : method Two Sample t-test t 3.620885736 df 34 Sig.(2-tailed) 0.0009454810217 stderr 0.2984063665 Mean Difference 1.080495356 estimate 2.764705882 1.684210526 null.value 0 conf.int 0.4740606562 1.686930056

Q4_Row3 : method Welch Two Sample t-test t 3.586563008 df 31.53936543 Sig.(2-tailed) 0.001116440771 stderr 0.3012620589 Mean Difference 1.080495356 estimate 2.764705882 1.684210526 null.value 0 conf.int 0.4664929343 1.694497778

Q4_Row4 : method Two Sample t-test t 0.1928195727 df 34 Sig.(2-tailed) 0.8482474587 stderr 0.417464653 Mean Difference 0.08049535604 estimate 2.764705882 2.684210526 null.value 0 conf.int -0.767894893 0.928885605

Q4_Row4 : method Welch Two Sample t-test t 0.1962209044 df 32.74009678 Sig.(2-tailed) 0.845650162 stderr 0.4102282388 Mean Difference 0.08049535604 estimate 2.764705882 2.684210526 null.value 0 conf.int -0.7543720407 0.9153627528

Q4_Row5 : method Two Sample t-test t 0.5608402978 df 34 Sig.(2-tailed) 0.5785840541 stderr 0.3588158534 Mean Difference 0.2012383901 estimate 2.411764706 2.210526316 null.value 0 conf.int -0.5279631579 0.9304399381

Q4_Row5 : method Welch Two Sample t-test t 0.5681341647 df 33.56383007 Sig.(2-tailed) 0.5737254111 stderr 0.3542092742 Mean Difference 0.2012383901 estimate 2.411764706 2.210526316 null.value 0 conf.int -0.5189461846 0.9214229648

Q4_Row6 : method Two Sample t-test t 1.052654097 df 34 Sig.(2-tailed) 0.2999235294 stderr 0.3588158534 Mean Difference 0.3777089783 estimate 2.588235294 2.210526316 null.value 0 conf.int -0.3514925696 1.106910526

Q4_Row6 : method Welch Two Sample t-test t 1.060005504 df 33.99793296 Sig.(2-tailed) 0.2966149795 stderr 0.3563273746 Mean Difference 0.3777089783 estimate 2.588235294 2.210526316 null.value 0 conf.int -0.3464369939 1.101854951

Q4_Row7 : method Two Sample t-test t 1.372382837 df 34 Sig.(2-tailed) 0.178930767 stderr 0.365457782 Mean Difference 0.5015479876 estimate 2.764705882 2.263157895 null.value 0 conf.int -0.2411515833 1.244247559

Q4_Row7 : method Welch Two Sample t-test t 1.372825213 df 33.60116644 Sig.(2-tailed) 0.1789002926 stderr 0.3653400177 Mean Difference 0.5015479876 estimate 2.764705882 2.263157895 null.value 0 conf.int -0.2412370021 1.244332977

Q5_Row1 : method Two Sample t-test t 0.70170841 df 34 Sig.(2-tailed) 0.4876368782 stderr 0.3794366238 Mean Difference 0.26625387 estimate 2.529411765 2.263157895 null.value 0 conf.int -0.5048541253 1.037361865

Q5_Row1 : method Welch Two Sample t-test t 0.7025032572 df 33.69724743 Sig.(2-tailed) 0.4871902303 stderr 0.3790073103 Mean Difference 0.26625387 estimate 2.529411765 2.263157895 null.value 0 conf.int -0.5042366343 1.036744374

Q5_Row2 : method Two Sample t-test t -0.6391129194 df 34 Sig.(2-tailed) 0.5270321219 stderr 0.3633131726 Mean Difference -0.2321981424 estimate 2.294117647 2.526315789 null.value 0 conf.int -0.9705393427 0.5061430579

Q5_Row2 : method Welch Two Sample t-test t -0.6530736923 df 31.6733954 Sig.(2-tailed) 0.5184216014 stderr 0.3555466177 Mean Difference -0.2321981424 estimate 2.294117647 2.526315789 null.value 0 conf.int -0.9567159014 0.4923196165

Q5_Row3 : method Two Sample t-test t 0.888752918 df 34 Sig.(2-tailed) 0.3803808098 stderr 0.4145370948 Mean Difference 0.3684210526 estimate 3 2.631578947 null.value 0 conf.int -0.4740196822 1.210861787

Q5_Row3 : method Welch Two Sample t-test t 0.8979796625 df 33.84511541 Sig.(2-tailed) 0.3755403497 stderr 0.4102777246 Mean Difference 0.3684210526 estimate 3 2.631578947 null.value 0 conf.int -0.4655041682 1.202346273

Q5_Row4 : method Two Sample t-test t 0.145140192 df 34 Sig.(2-tailed) 0.8854571987 stderr 0.3839567345 Mean Difference 0.05572755418 estimate 2.529411765 2.473684211 null.value 0 conf.int -0.7245664114 0.8360215198

Q5_Row4 : method Welch Two Sample t-test t 0.1470008298 df 33.58829702 Sig.(2-tailed) 0.8840101888 stderr 0.3790968681 Mean Difference 0.05572755418 estimate 2.529411765 2.473684211 null.value 0 conf.int -0.715037959 0.8264930674

Q5_Row5 : method Two Sample t-test t 0.8890613092 df 34 Sig.(2-tailed) 0.380217413 stderr 0.3795703365 Mean Difference 0.3374613003 estimate 2.705882353 2.368421053 null.value 0 conf.int -0.4339184319 1.108841033

Q5_Row5 : method Welch Two Sample t-test t 0.8917508424 df 33.87117036 Sig.(2-tailed) 0.3788179858 stderr 0.378425547 Mean Difference 0.3374613003 estimate 2.705882353 2.368421053 null.value 0 conf.int -0.4316996978 1.106622298

Q5_Row6 : method Two Sample t-test t 1.929789181 df 34 Sig.(2-tailed) 0.0620108892 stderr 0.372199337 Mean Difference 0.7182662539 estimate 2.823529412 2.105263158 null.value 0 conf.int -0.03813380518 1.474666313

Q5_Row6 : method Welch Two Sample t-test t 1.926738183 df 33.32646095 Sig.(2-tailed) 0.06257472908 stderr 0.3727887162 Mean Difference 0.7182662539 estimate 2.823529412 2.105263158 null.value 0 conf.int -0.03989594828 1.476428456

Q5_Row7 : method Two Sample t-test t -0.3626686192 df 34 Sig.(2-tailed) 0.7190972837 stderr 0.3926859208 Mean Difference -0.1424148607 estimate 2.647058824 2.789473684 null.value 0 conf.int -0.9404486672 0.6556189458

Q5_Row7 : method Welch Two Sample t-test t -0.3664928229 df 33.83173653 Sig.(2-tailed) 0.7162795222 stderr 0.3885884028 Mean Difference -0.1424148607 estimate 2.647058824 2.789473684 null.value 0 conf.int -0.9322662049 0.6474364835

depression : method Two Sample t-test t 0.8493334295 df 34 Sig.(2-tailed) 0.401636706 stderr 0.2921353289 Mean Difference 0.2481203008 estimate 2.428571429 2.180451128 null.value 0 conf.int -0.3455701173 0.8418107188

depression : method Welch Two Sample t-test t 0.8577340232 df 33.88209733 Sig.(2-tailed) 0.3970662452 stderr 0.2892741736 Mean Difference 0.2481203008 estimate 2.428571429 2.180451128 null.value 0 conf.int -0.3398309105 0.836071512

anxiety : method Two Sample t-test t 1.409128639 df 34 Sig.(2-tailed) 0.1678780367 stderr 0.2749494798 Mean Difference 0.3874391862 estimate 2.680672269 2.293233083 null.value 0 conf.int -0.1713253844 0.9462037568

anxiety : method Welch Two Sample t-test t 1.408371934 df 33.4855244 Sig.(2-tailed) 0.1682385926 stderr 0.2750972075 Mean Difference 0.3874391862 estimate 2.680672269 2.293233083 null.value 0 conf.int -0.1719421625 0.9468205349

pressure : method Two Sample t-test t 0.6150805715 df 34 Sig.(2-tailed) 0.5425966849 stderr 0.3185452655 Mean Difference 0.195931004 estimate 2.647058824 2.45112782 null.value 0 conf.int -0.4514308628 0.8432928708

pressure : method Welch Two Sample t-test t 0.621008104 df 33.89925701 Sig.(2-tailed) 0.5387478213 stderr 0.3155047458 Mean Difference 0.195931004 estimate 2.647058824 2.45112782 null.value 0 conf.int -0.4453219775 0.8371839855

Zdepression : method Two Sample t-test t 0.8493334295 df 34 Sig.(2-tailed) 0.401636706 stderr 0.3351857926 Mean Difference 0.2846844987 estimate 0.1502501521 -0.1344343466 null.value 0 conf.int -0.3964949878 0.9658639853

Zdepression : method Welch Two Sample t-test t 0.8577340232 df 33.88209733 Sig.(2-tailed) 0.3970662452 stderr 0.3319030038 Mean Difference 0.2846844987 estimate 0.1502501521 -0.1344343466 null.value 0 conf.int -0.389910024 0.9592790215

Zanxiety : method Two Sample t-test t 1.409128639 df 34 Sig.(2-tailed) 0.1678780367 stderr 0.3292451412 Mean Difference 0.4639487577 estimate 0.2448618444 -0.2190869134 null.value 0 conf.int -0.2051578728 1.133055388

Zanxiety : method Welch Two Sample t-test t 1.408371934 df 33.4855244 Sig.(2-tailed) 0.1682385926 stderr 0.3294220414 Mean Difference 0.4639487577 estimate 0.2448618444 -0.2190869134 null.value 0 conf.int -0.2058964492 1.133793965

Zage : method Two Sample t-test t 1.179776498 df 34 Sig.(2-tailed) 0.2462783068 stderr 0.3319955331 Mean Difference 0.3916805274 estimate 0.2067202784 -0.1849602491 null.value 0 conf.int -0.2830155719 1.066376627

Zage : method Welch Two Sample t-test t 1.184407313 df 33.92851333 Sig.(2-tailed) 0.2444828842 stderr 0.3306974916 Mean Difference 0.3916805274 estimate 0.2067202784 -0.1849602491 null.value 0 conf.int -0.2804297956 1.06379085

```
ZageXZdepression : method  Two Sample t-test t 0.5987060952 df 34 Sig.(2-tailed)
0.5533371345 stderr 0.3266342545 Mean Difference 0.1955579191 estimate 0.1004793349
-0.0950785842 null.value 0 conf.int -0.4682427511 0.8593585892

ZageXZdepression : method Welch Two Sample t-test t 0.6000915474 df 33.81422049 Sig.
(2-tailed) 0.552445926 stderr 0.3258801426 Mean Difference 0.1955579191 estimate
0.1004793349 -0.0950785842 null.value 0 conf.int -0.4668442613 0.8579600994

ZageXZanxiety : method  Two Sample t-test t 0.5121317223 df 34 Sig.(2-tailed)
0.6118696446 stderr 0.4031371621 Mean Difference 0.2064593291 estimate 0.1382902235
-0.06816910567 null.value 0 conf.int -0.612813955 1.025732613

ZageXZanxiety : method Welch Two Sample t-test t 0.5112397199 df 33.30019833 Sig.(2-
tailed) 0.6125564868 stderr 0.403840549 Mean Difference 0.2064593291 estimate
0.1382902235 -0.06816910567 null.value 0 conf.int -0.6148791587 1.027797817
```

相关性分析 Pearson Correlation

1.计算协方差 2.计算标准差 3.计算Pearson相关系数

$$
\begin{aligned}
1.\ Cov(x,y) &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_j-\bar{y})}{n-1} \\
2.\ SD(x) &= \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \\
3.\ r &= \frac{Cov(x,y)}{SD(x)\cdot SD(y)} \\
&or \\
r &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}
\end{aligned}
$$

```r
getPearsonCorrelationCoeffcient <-function(data0,data1){
  #levene_test <- leveneTest(data0, data1)
  #print(levene_test)
  n0 <- length(data0)
  n1 <- length(data1)
  if(n0!=n1){
    print("error")
  }
  n <- n0
  mean0 <- mean(data0)
  mean1 <- mean(data1)
  cov_ <- sum((data0-mean0)*(data1-mean1))/(n-1)
  sd_x <- sqrt(sum((data0-mean0)^2)/(n-1))
  sd_y <- sqrt(sum((data1-mean1)^2)/(n-1))
  r <- cov_/(sd_x*sd_y)
  df <- n-2
  t <- r / sqrt( (1-r^2) / df)
  #2.0322
  #p_value = dt(t,df=df)
```

```
  p_value = 2*pt(-abs(t),df=df)#dt(t,df=df)

  #p_value = 2* pf( r^2 * df / (1 - r^2) , df1=df, df2=Inf)
  cat("r:",r,"t:",t,"df",df,"p-value",p_value ,"\n")
}

getPearsonCorrelationCoeffcient(data$depression,data$anxiety)
```

```
r: 0.7786999274 t: 7.237148279 df 34 p-value 0.00000002241535713
```

```
getPearsonCorrelationCoeffcient(data$depression,data$pressure)
```

```
r: 0.8161253647 t: 8.234979265 df 34 p-value 0.000000001311410587
```

```
getPearsonCorrelationCoeffcient(data$anxiety,data$pressure)
```

```
r: 0.8670901293 t: 10.14944844 df 34 p-value 0.0000000000007967586114
```

```
###
cor.test(data$depression,data$anxiety)
```

```

    Pearson's product-moment correlation

data:  data$depression and data$anxiety
t = 7.2371483, df = 34, p-value = 0.00000002241536
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6049217871 0.8816758633
sample estimates:
         cor
0.7786999274
```

```
cor.test(data$depression,data$pressure)
```

```

    Pearson's product-moment correlation

data:  data$depression and data$pressure
t = 8.2349793, df = 34, p-value = 0.000000001311411
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6662210424 0.9026396333
sample estimates:
         cor
0.8161253647
```

```
cor.test(data$anxiety,data$pressure)
```

```

    Pearson's product-moment correlation
```

```
data:  data$anxiety and data$pressure
t = 10.149448, df = 34, p-value = 0.000000000007967586
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7530861008 0.9305422170
sample estimates:
         cor
0.8670901293
```

回归分析

一元回归分析

最小二乘法推导beta0 beta1

$$y_i = \beta_0 + \beta_1 x_i + \mu_i$$
$$\mu\text{的期望值为零，}x\text{和}\mu\text{之间的协方差也为零}$$
$$E(\mu) = 0 \to E(y - \beta_0 - \beta_1 x) = 0$$
$$Cov(x, \mu) = E(x\mu) = 0 \to E(x(y - \beta_0 - \beta_1 x)) = 0$$
$$\text{对一组样本, 估计}\hat{\beta}_0, \hat{\beta}_1\text{可得方程：}$$
$$n^{-1}\sum_{i=0}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \to n^{-1}\sum_{i=0}^{n}(y_i) - n^{-1}\sum_{i=0}^{n}(\hat{\beta}_0) - n^{-1}\sum_{i=0}^{n}(\hat{\beta}_1 x_i) = 0$$
$$\to \bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} = 0$$
$$\text{由此可得：}\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$
$$n^{-1}\sum_{i=0}^{n} x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \to n^{-1}\sum_{i=0}^{n} x_i(y_i - (\bar{y} - \hat{\beta}\bar{x}) - \hat{\beta}_1 x_i) = 0$$
$$\to n^{-1}\sum_{i=0}^{n} x_i(y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i) = 0$$
$$\to n^{-1}\sum_{i=0}^{n} x_i(y_i - \bar{y}) = n^{-1}\sum_{i=0}^{n} x_i(\hat{\beta}_1 x_i - \hat{\beta}_1 \bar{x})(*\text{从左到右变号}, x_i, \bar{x}\text{换位}) =$$
$$n^{-1}\hat{\beta}_1\sum_{i=0}^{n} x_i(x_i - \bar{x}) -> \hat{\beta}_1 = \frac{n^{-1}\sum_{i=0}^{n} x_i(y_i - \bar{y})}{n^{-1}\sum_{i=0}^{n} x_i(x_i - \bar{x})} \to \hat{\beta}_1 = \frac{\sum_{i=0}^{n} x_i(y_i - \bar{y})}{\sum_{i=0}^{n} x_i(x_i - \bar{x})}$$

求合算子性质化简：

$$\sum_{i=0}^{n} x_i(y_i - \bar{y}) \to \sum_{i=0}^{n}(x_i y_i - x_i\bar{y}) \to \sum_{i=0}^{n} x_i y_i - \sum_{i=0}^{n} x_i\bar{y} \to \sum_{i=0}^{n} x_i y_i - n\bar{y}\sum_{i=0}^{n} x_i \to$$

$$\sum_{i=0}^{n} x_i y_i - n\bar{y}\bar{x} \to \sum_{i=0}^{n} x_i y_i - \sum_{i=0}^{n} x_i\bar{y}\ or \sum_{i=0}^{n} x_i y_i - \sum_{i=0}^{n} y_i\bar{x} \to$$

$$\sum_{i=0}^{n} x_i(y_i - \bar{y})\ or\ \sum_{i=0}^{n} y_i(x_i - \bar{x})$$

$$\sum_{i=0}^{n} x_i y_i - n\bar{y}\bar{x} \to \sum_{i=0}^{n} x_i y_i - n\bar{y}\bar{x} - n\bar{y}\bar{x} + n\bar{y}\bar{x} \to$$

$$\sum_{i=0}^{n} x_i y_i - \sum_{i=0}^{n} x_i\bar{y} - \sum_{i=0}^{n} \bar{x}y_i + \sum_{i=0}^{n} \bar{x}\bar{y}$$

$$\sum_{i=0}^{n}(x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y}) \to \sum_{i=0}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

$$\sum_{i=0}^{n} x_i(x_i - \bar{x}) \to \sum_{i=0}^{n}(x_i^2 - x_i\bar{x}) \to \sum_{i=0}^{n} x_i^2 - \sum_{i=0}^{n} x_i\bar{x} \to \sum_{i=0}^{n} x_i^2 - \bar{x}\sum_{i=0}^{n} x_i \to$$

$$\sum_{i=0}^{n} x_i^2 - n\bar{x}^2 \to \sum_{i=0}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \to \sum_{i=0}^{n}(x_i^2 - 2\bar{x}^2 + \bar{x}^2) \to \sum_{i=0}^{n}(x_i - \bar{x})^2$$

代入化简公式：

$$\hat{\beta_1} = \frac{\sum_{i=0}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=0}^{n}(x_i - \bar{x})^2}$$

观察发现相近公式：

$$\text{协方差}Cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}\ and\ \text{方差}Var(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

分子同除$n-1$可抵消得：

$$\hat{\beta_1} = \frac{Cov(x,y)}{var(x)}, \hat{\beta_0} = \bar{y} - \hat{\beta_1}\bar{x}$$

求回归标准误差

Std. Error of the Estimate / standard error of the regression, SER Residual Sum of Squares ,RSS 是回归分析中用于衡量观测值与回归模型预测值之间差异的一个重要指标。它是残差的平方和，表示了模型未能解释的观测值的方差。

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y})^2$$

$$MSE = \frac{SSE}{n-k}$$

$$SER = \sqrt{MSE}$$

$$SER = (\sum_{i=1}^{n}(y_i - \hat{y})^2 * (n-k)^{-1})^{\frac{1}{2}}$$

beta0, beta1标准差及t值,显著性

$$SE(\hat{\beta}_0) = \sqrt{\frac{MSE * \sum_{i=1}^{N} x_i^2}{N * \sum_{i=1}^{n}(x_i - \bar{x})^2}} \rightarrow SE(\hat{\beta}_0) = \sqrt{\frac{\frac{\sum_{i=1}^{n} *(y_i - \hat{y})^2}{n-k} * \sum_{i=1}^{N} x_i^2}{N * \sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$SE(\hat{\beta}_1) = \sqrt{\frac{MSE}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \rightarrow SE(\hat{\beta}_1) = \sqrt{\frac{\frac{\sum_{i=1}^{n} *(y_i - \hat{y})^2}{n-k}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$T(\beta_0) = \frac{\beta 0}{SE(\beta_0)}$$

$$T(\beta_1) = \frac{\beta 1}{SE(\beta_1)}$$

$$PValue(\beta 0) = 2 * pt(-|T(\beta 0)|, df)$$

$$PValue(\beta 1) = 2 * pt(-|T(\beta 1)|, df)$$

R,R square

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

$$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$SSE = \sum_{i=1}^{n}(\hat{y} - \bar{y})^2,$$

$$SSR = \sum_{i=1}^{n}(y_i - \hat{y})^2$$

$$R^2 = 1 - \frac{SSR(\text{残差平方和})}{SST(\text{总平方和})} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y} - \bar{\hat{y}})}{(\sum_{i=1}^{n}(y_i - \bar{y})^2)(\sum_{i=1}^{n}(\hat{y} - \bar{\hat{y}})^2)}$$

$$Adjusted\ R^2 = 1 - \frac{SSR}{df_e} / \frac{SST}{df_t}$$

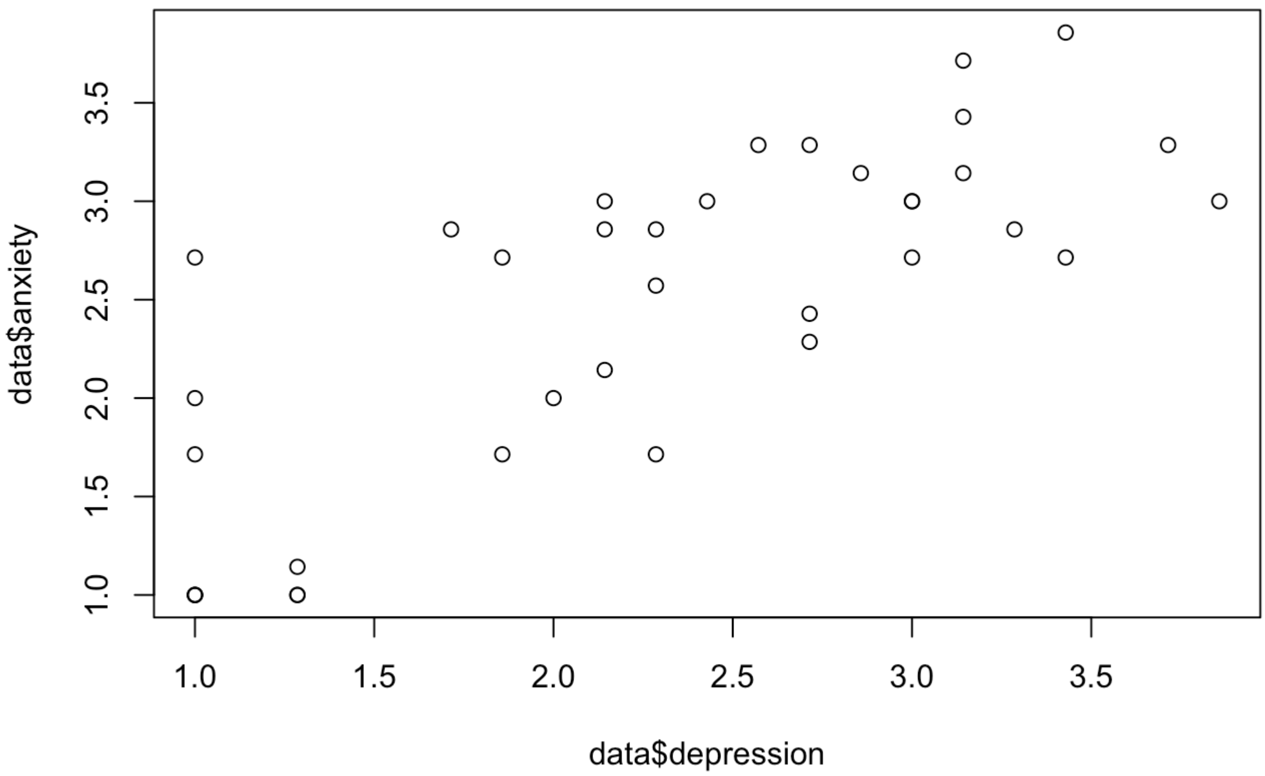
$$df_e = n(\text{样本数}) - k(\text{参数个数, 一元中有两个}beta),\ df_t = n - 1$$

```
SimpleLinearRegression <- function(datax,datay){
  x_m = mean(datax)
  y_m = mean(datay)
  cov = sum((datax-x_m)*(datay-y_m))
  var = sum((datax-x_m)^2)
  beta1 = cov/var
  beta0 = y_m - beta1*x_m
  N = length(datay)
  df =  N-2
  SS_res = sum( (datay-(beta0+beta1*datax))^2)
  SSX = sum((datax-x_m)^2)
  MSE = SS_res/df
  mu = sqrt( MSE )
  beta0_SE = sqrt(MSE/N)
  beta0_SE = sqrt( MSE*sum(datax^2) / (N*sum((datax-x_m)^2) ) )
  beta1_SE = sqrt(MSE/SSX)
  beta0_t = beta0/beta0_SE
  beta1_t = beta1/beta1_SE
```

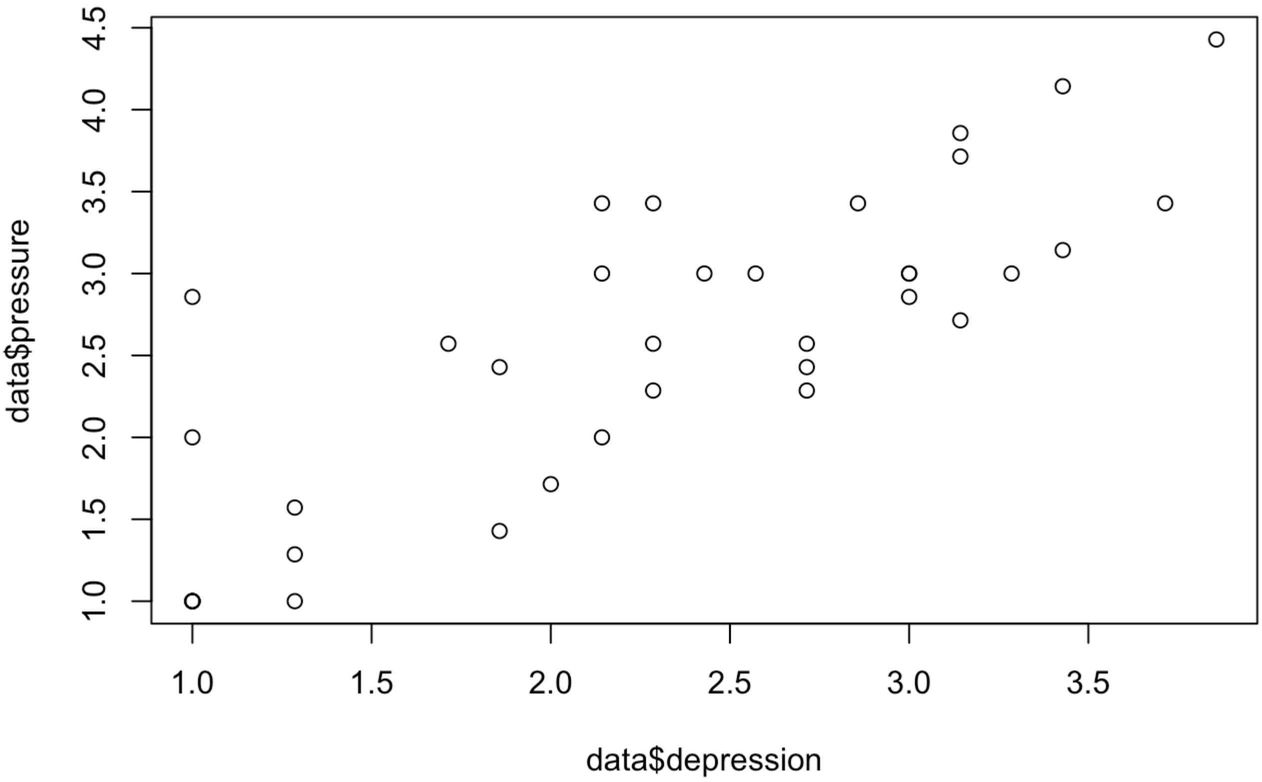
```
  beta0_pValue = 2*pt(-abs(beta0_t),df=df)
  beta1_pValue = 2*pt(-abs(beta1_t),df=df)
  r = cov / (sqrt(var)* sqrt(sum((datay-y_m)^2)) )
  r_square = 1- SS_res / (sum((datay-y_m)^2))
  adjusted_r_square = 1- (SS_res/df) /  (sum((datay-y_m)^2)/(N-1) )
  cat("beta0:",beta0,"beta1",beta1,"mu:",mu,"MSE:",MSE,"beta0_SE:",beta0_SE,"beta1_SE:

}
SimpleLinearRegression(data$depression,data$anxiety)
```

```
beta0: 0.7619047619 beta1 0.7461139896 mu: 0.5315819344 MSE: 0.282579353 beta0_SE:
0.2528997623 beta1_SE: 0.103095026 beta0_t: 3.012674883 beta1_t: 7.237148279
beta0_pValue: 0.004863096682 beta1_pValue: 0.00000002241535713 r: 0.7786999274 r
square 0.6063735769 adjusted_r_square 0.5947963291
```

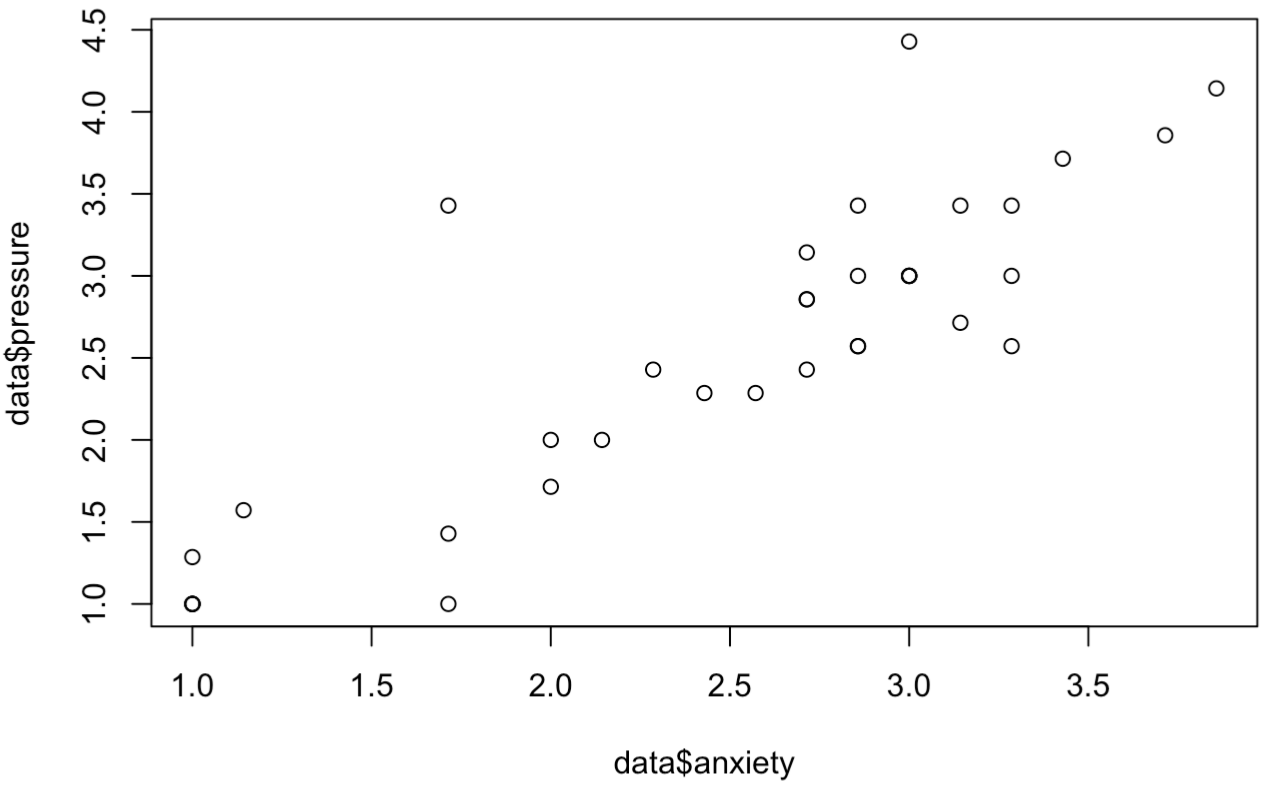
```
#r内置计算方法
#getPearsonCorrelationCoeffcient(data$depression,data$anxiety)
#getPearsonCorrelationCoeffcient(data$depression,data$pressure)
#getPearsonCorrelationCoeffcient(data$anxiety,data$pressure)
plot(data$depression,data$anxiety)
```




```
plot(data$depression,data$pressure)
```




```
plot(data$anxiety,data$pressure)
```



```
model<-lm(depression~anxiety,data=data)
summary(model)
```

```
Call:
lm(formula = depression ~ anxiety, data = data)

Residuals:
        Min          1Q      Median          3Q         Max 
-1.49112120 -0.30144131  0.04447364  0.28336519  1.13381908 

Coefficients:
             Estimate Std. Error t value       Pr(>|t|)    
(Intercept) 0.2851967  0.2930394 0.97324        0.33731    
anxiety     0.8127090  0.1122969 7.23715 0.000000022415 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5547984 on 34 degrees of freedom
Multiple R-squared:  0.6063736, Adjusted R-squared:  0.5947963 
F-statistic: 52.37632 on 1 and 34 DF,  p-value: 0.00000002241536
```

```
anova(model)
```

```
Analysis of Variance Table

Response: depression
          Df    Sum Sq    Mean Sq  F value         Pr(>F)    
anxiety    1 16.121493 16.1214934 52.37632 0.000000022415 ***
Residuals 34 10.465241  0.3078012                            
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(model)$r.squared
```

```
[1] 0.6063735769
```

```
#data$anxiety-predict(model)
```

多元回归分析

最小二乘法推导beta

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{xi} + \ldots + \beta_n x_{ni} + \mu_i$$

$$\mu = \sum_{i=1}^{n} \mu_i^2 = 0 = \sum_{i=1}^{n} (y_i - \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{xi} + \ldots + \beta_n x_{ni})^2$$

```
model2<-lm(pressure ~ depression + anxiety, data=data )
summary(model2)
```

```
Call:
lm(formula = pressure ~ depression + anxiety, data = data)

Residuals:
        Min          1Q      Median          3Q         Max 
-0.67337265 -0.20885061 -0.07938469  0.17758980  1.39712088 

Coefficients:
               Estimate  Std. Error t value    Pr(>|t|)    
(Intercept) 0.001536355 0.231883966 0.00663   0.9947535    
depression  0.388442401 0.133856235 2.90194   0.0065594 ** 
anxiety     0.666193411 0.139702300 4.76866 0.000036414 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4330253 on 33 degrees of freedom
Multiple R-squared:  0.802297,  Adjusted R-squared:  0.790315 
F-statistic:  66.9585 on 2 and 33 DF,  p-value: 0.000000000002422258
```

```
vif(model2)
```

```
 depression     anxiety 
2.540479859 2.540479859 
```

```
#anova(model)
#summary(model)$r.squared
```