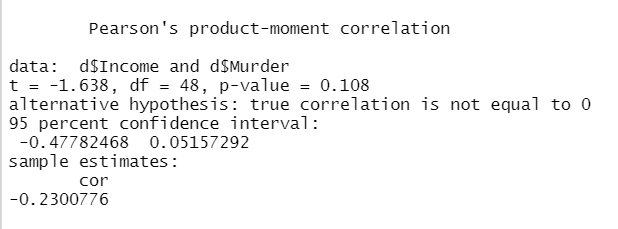
|  |  |
| --- | --- |
| **Name** | **VAISHALI M** |
| **Sub name &Slot** | **MAT LAB**  **L31, L32** |
| **Assignment name and Date** | **Correlation and regression** |

1. find correlation between any two variables (using 3 different methods)

**Method 1: for Income and Murder**

cor.test(d$Income,d$Murder,method ="pearson" )



**Method 2: for population and murder**

r=cor(d$Population,d$Murder)

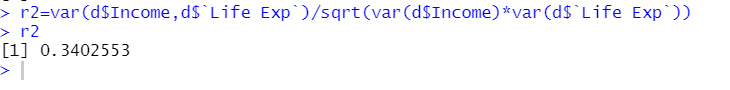
r



**Method 3: for income and life exp**

r2=var(d$Income,d$`Life Exp`)/sqrt(var(d$Income)\*var(d$`Life Exp`))

r2



Observation :

* correlation between Income and murder is WEAK NEGATIVE CORRELATION
* correlation between Population and murder is WEAK POSITIVE CORRELATION
* correlation between income and life exp is WEAK POSITIVE CORRELATION

2. Find regression equations and draw regression line

Here correlation between income and illiteracy is -0.4370752 which is STRONG NEGATIVE CORRELATION.

x=d$Income

y=d$Illiteracy

plot(x,y)

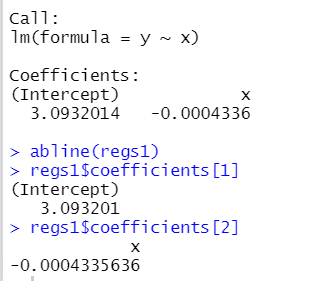
regs1<-lm(y~x) # lm means linear model

regs1

abline(regs1)

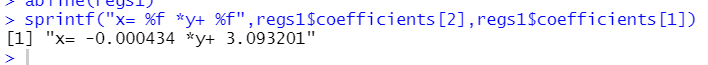
regs1$coefficients[1]

regs1$coefficients[2]

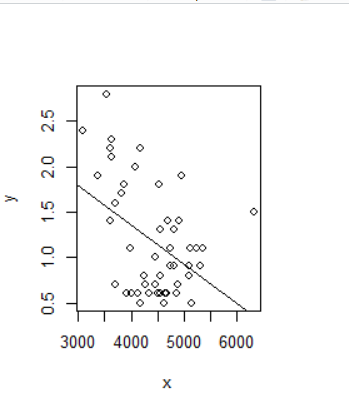


**Regression equation**

**sprintf("x= %f \*y+ %f",regs2$coefficients[2],regs2$coefficients[1])**



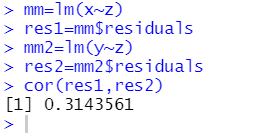
The regression line for income and il-literacy is shown below



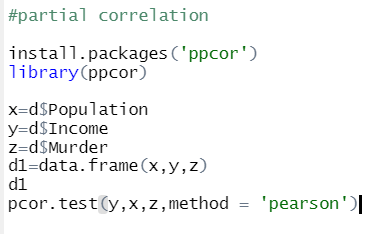
3. partial correlation for any three variables in the data set

**Method 1**—Finding regression and then correlation

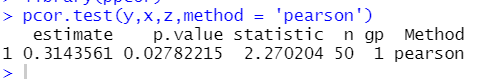
This is correlation of x and y keeping Z constant



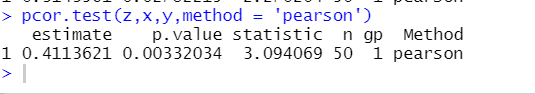
**Method 2**



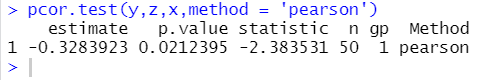
Result : 1. Keeping z as constant



2. Keeping y as constant



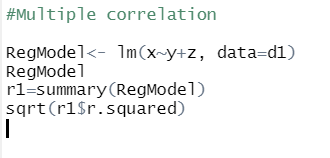
3. Keeping x as constant

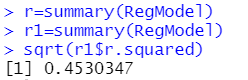


4. multiple correlation

Square root of R square.

**#Multiple correlation for x as dependent**





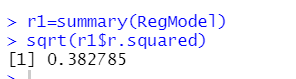
**#Multiple correlation for y as dependent**

RegModel<- lm(y~x+z, data=d1)

RegModel

r1=summary(RegModel)

sqrt(r1$r.squared)



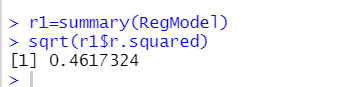
**#Multiple correlation for z as dependent**

RegModel<- lm(z~y+x, data=d1)

RegModel

r1=summary(RegModel)

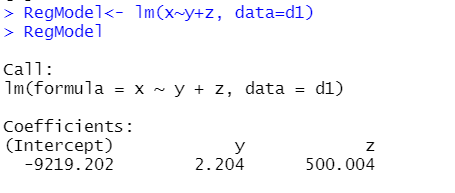
sqrt(r1$r.squared)



5. multiple regression

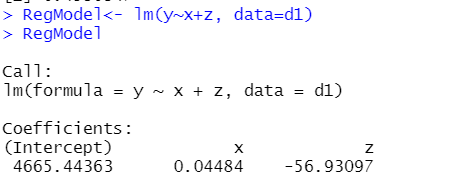
1. RegModel<- lm(x~y+z, data=d1)

RegModel



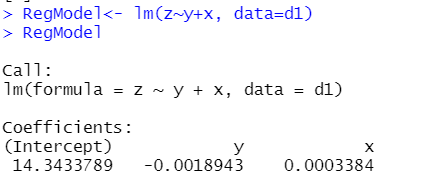
2. RegModel<- lm(y~x+z, data=d1)

RegModel



3. RegModel<- lm(z~x+y, data=d1)

RegModel



6. diagnostic measures

