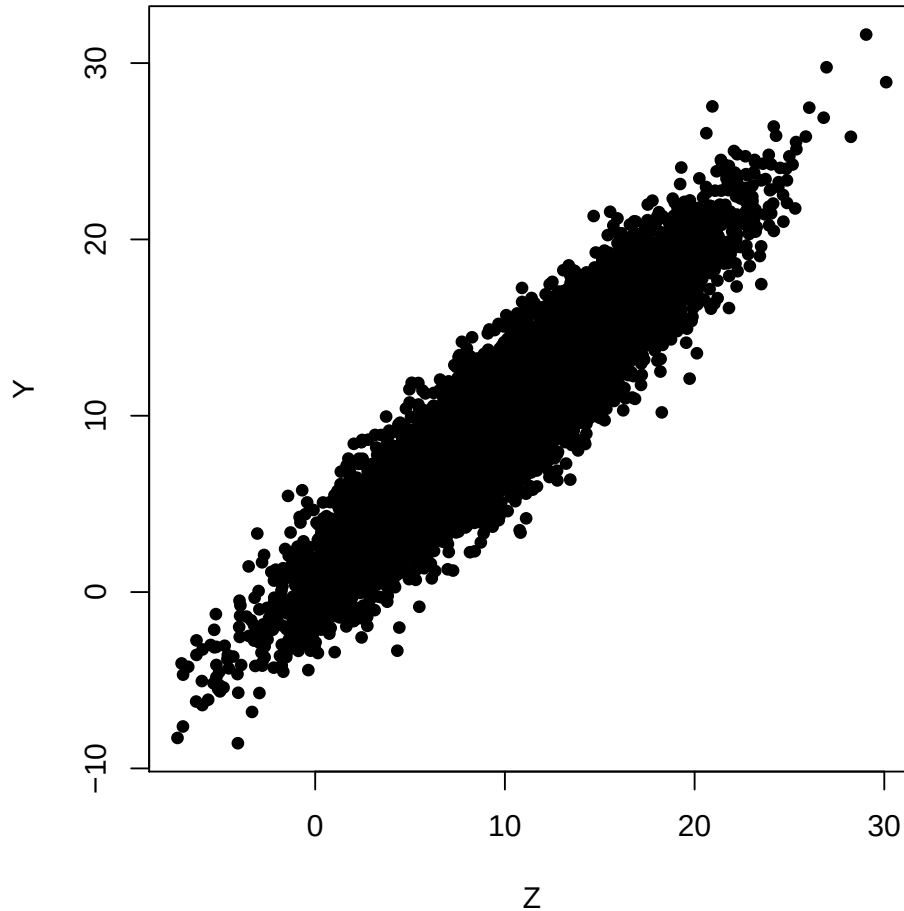
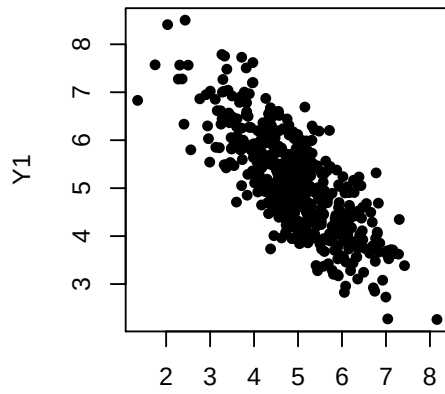


SIMPSON'S PARADOX WITH CONTINUOUS VARIABLES

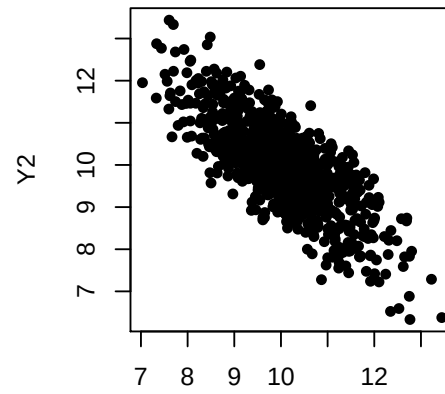
```
library(MASS)
set.seed(2101)                                     #Set the random number generator seed value
n<-10000                                           #Set the sample size
X<-rnorm(n,10,5)                                   #Generate the X random variables
Sig.YZ<-matrix(c(1,-0.9,-0.9,1),2,2)             #Conditional variance-covariance for Y,Z given X
YZ<-mvrnorm(n,mu=c(0,0),Sigma=Sig.YZ)           #Generate the Y,Z variables
Y<-X+YZ[,1];Z<-X+YZ[,2]                          #Change the mean according to X
par(mar=c(4,4,1,0),pty='s')                     #Set up the plotting margins
plot(Z,Y,pch=19,cex=0.8)
```



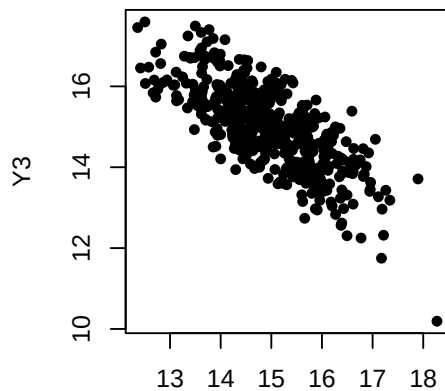
```
Y1<-Y[X>4.5 & X < 5.5];Z1<-Z[X>4.5 & X < 5.5]; #First subset analysis
Y2<-Y[X>9.5 & X < 10.5];Z2<-Z[X>9.5 & X < 10.5]; #Second subset analysis
Y3<-Y[X>14.5 & X < 15.5];Z3<-Z[X>14.5 & X < 15.5]; #Third subset analysis
par(mar=c(4,2,1,0),pty='s',mfrow=c(2,2))         #Set up the plotting margins
plot(Z1,Y1,pch=19,cex=0.8)
plot(Z2,Y2,pch=19,cex=0.8)
plot(Z3,Y3,pch=19,cex=0.8)
```



Z1



Z2



Z3

```
coef(summary(lm(Y1~Z1)))           #Group 1 regression

+               Estimate Std. Error  t value    Pr(>|t|)
+ (Intercept)   8.8879877  0.14946921  59.46367 1.830707e-222
+ Z1            -0.7680397  0.02959994 -25.94734 3.706014e-93

coef(summary(lm(Y2~Z2)))           #Group 2 regression

+               Estimate Std. Error  t value    Pr(>|t|)
+ (Intercept)  17.8555080  0.23106248  77.27567 0.000000e+00
+ Z2           -0.7859522  0.02285968 -34.38159 1.431588e-161

coef(summary(lm(Y3~Z3)))           #Group 3 regression

+               Estimate Std. Error  t value    Pr(>|t|)
+ (Intercept)  26.0170737  0.46111581  56.42199 1.830873e-213
+ Z3           -0.7387753  0.03073268 -24.03875 2.712403e-84

coef(summary(lm(c(Y1,Y2,Y3)~c(Z1,Z2,Z3)))) #Pooled regression

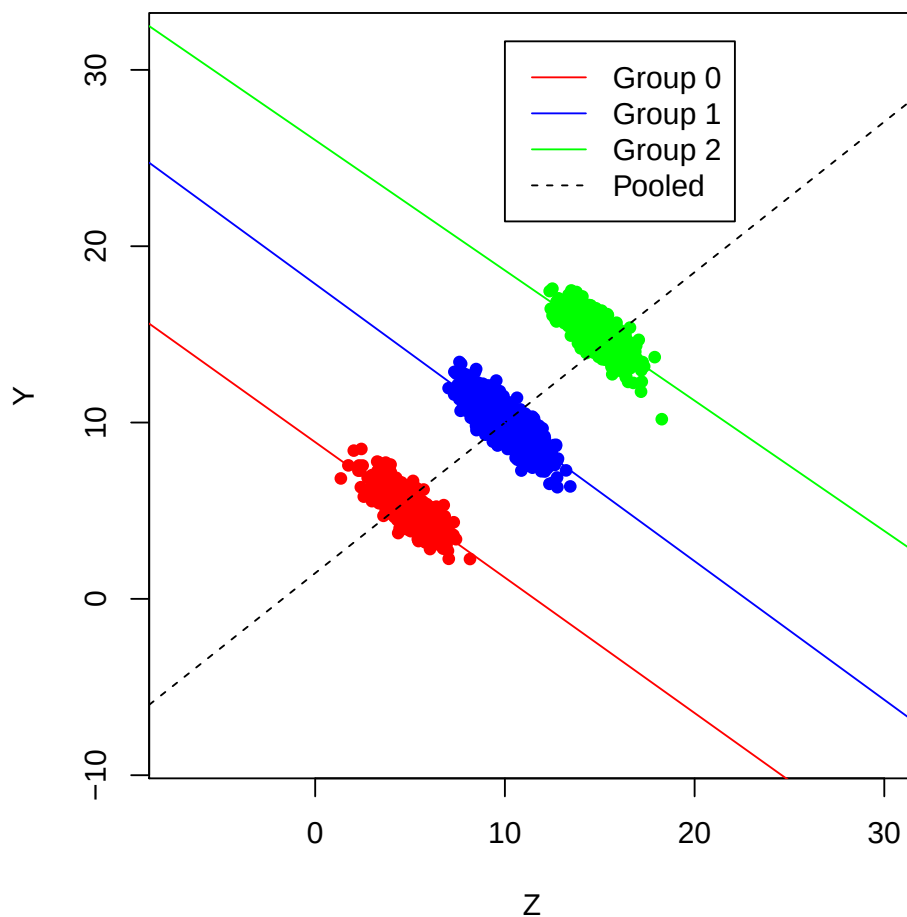
+               Estimate Std. Error  t value    Pr(>|t|)
+ (Intercept)   1.459732  0.12525325  11.65425 2.649385e-30
+ c(Z1, Z2, Z3)  0.853282  0.01169586  72.95588 0.000000e+00
```

```
par(mar=c(4,4,1,0),pty='s')
plot(Z,Y,type='n')
points(Z1,Y1,pch=19,cex=0.8,col='red')
```

```

points(Z2,Y2,pch=19,cex=0.8,col='blue')
points(Z3,Y3,pch=19,cex=0.8,col='green')
abline(lm(c(Y1,Y2,Y3)~c(Z1,Z2,Z3)),lty=2)
abline(lm(Y1~Z1),col='red')
abline(lm(Y2~Z2),col='blue')
abline(lm(Y3~Z3),col='green')
legend(10,max(Y),c('Group 0','Group 1','Group 2','Pooled'),
      col=c('red','blue','green','black'),lty=c(1,1,1,2))

```



```

cor(cbind(X,Y,Z))
+           X           Y           Z
+ X 1.0000000 0.9795457 0.9795717
+ Y 0.9795457 1.0000000 0.9230643
+ Z 0.9795717 0.9230643 1.0000000

```