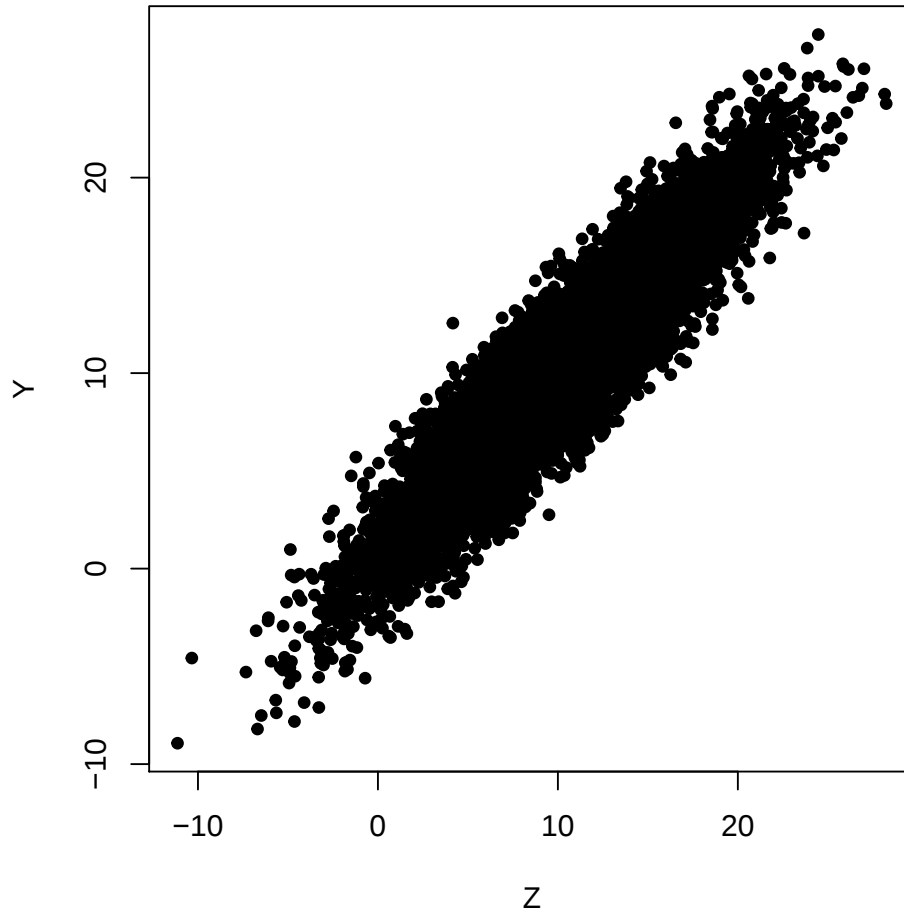
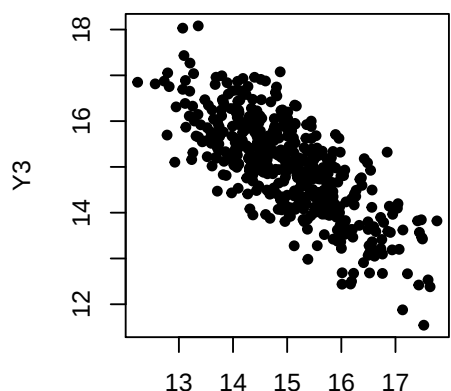
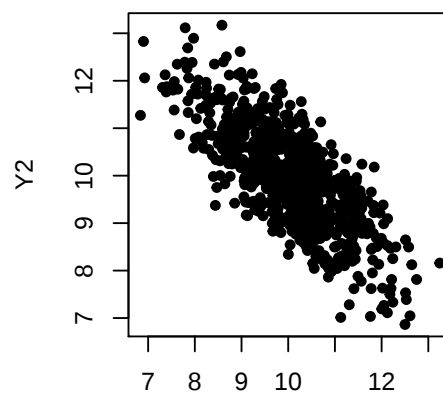
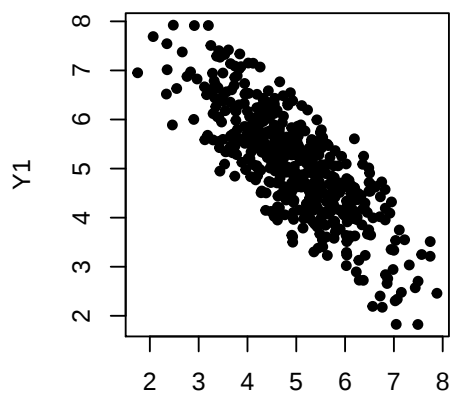


SIMPSON'S PARADOX WITH CONTINUOUS VARIABLES

```
library(MASS)
set.seed(2111)                                     #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)
```



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

+               Estimate Std. Error   t value      Pr(>|t|)
+ (Intercept)   9.1336287 0.14557132  62.74333 4.733761e-236
+ Z1            -0.8184506 0.02868607 -28.53129 3.872360e-106

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

+               Estimate Std. Error   t value      Pr(>|t|)
+ (Intercept)  17.6160738 0.23954157  73.54078 0.000000e+00
+ Z2           -0.7629346 0.02374499 -32.13033 3.096694e-146

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

+               Estimate Std. Error   t value      Pr(>|t|)
+ (Intercept)  26.5516171 0.47620626  55.75655 4.768122e-208
+ Z3          -0.7695225 0.03163317 -24.32644 6.308926e-85

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

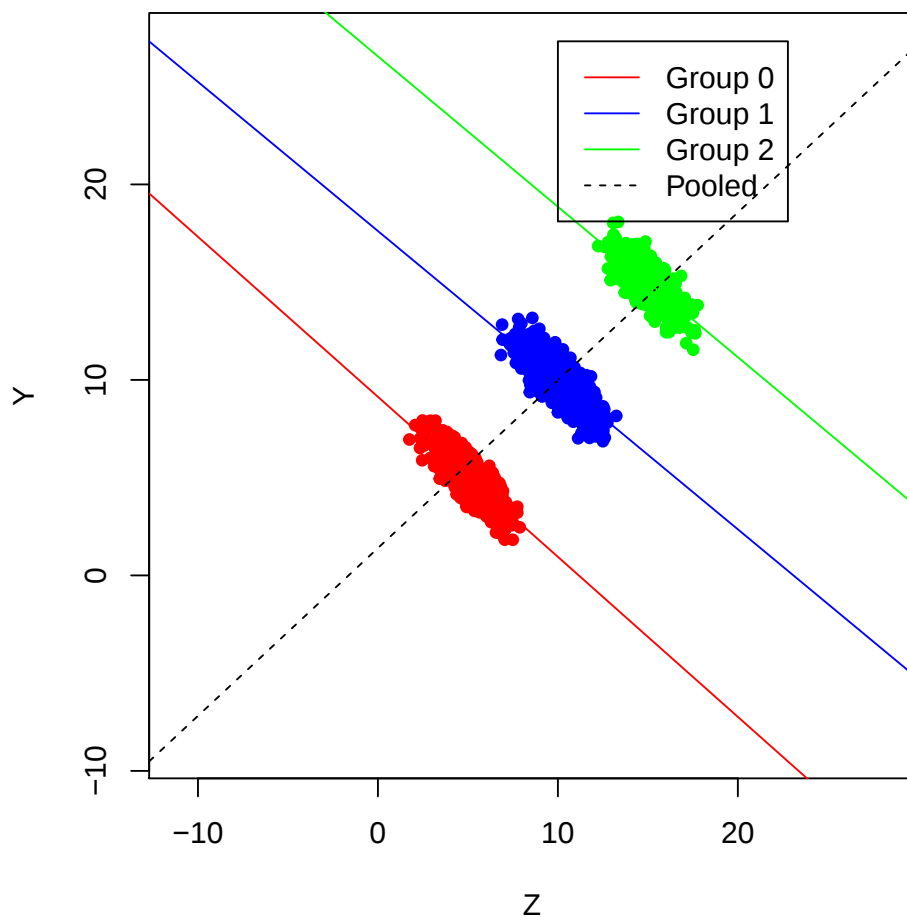
+               Estimate Std. Error   t value      Pr(>|t|)
+ (Intercept)   1.3976736 0.12496988  11.18408 4.216643e-28
+ c(Z1, Z2, Z3) 0.8583356 0.01172109  73.23004 0.000000e+00
```

```
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 1.000000 0.9805354 0.9807260
+ Y 0.9805354 1.0000000 0.9270599
+ Z 0.9807260 0.9270599 1.0000000

```