PS2 Final Copy

Nathaniel Williams

2025-10-22

```
set.seed(123)
Rv1<- rnorm(20)
rv2<- rnorm(20)
#correlation
cor(Rv1, rv2)
```

```
## [1] -0.09172278
```

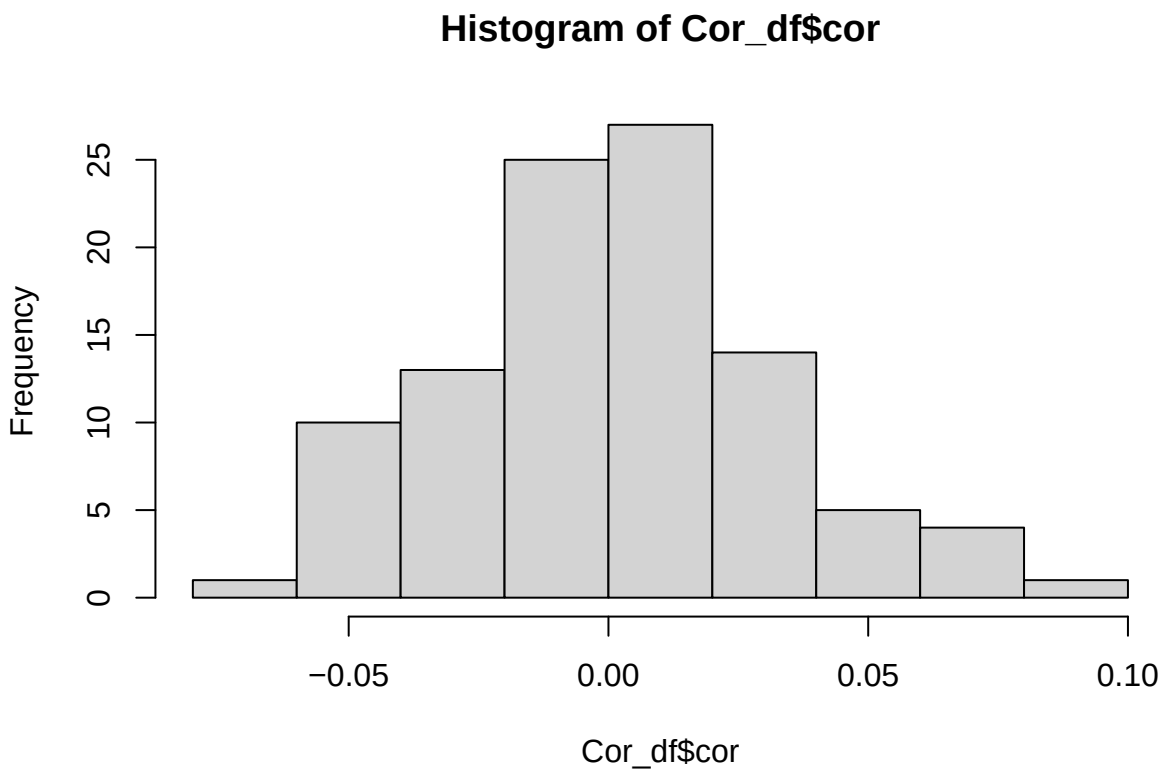
```
#Weak Negative correlation -0.16
#storage
Cor_df<- data.frame(cor=numeric(100))
#Create a loop
for(i in 1:100){
  V1<-rnorm(1000)
  V2<-rnorm(1000)
  Cor_df$cor[i] <- cor(V1,V2)
}
Cor_df
```

```
##                cor
## 1    8.369403e-02
## 2    4.891108e-02
## 3    8.274482e-03
## 4    4.031930e-02
## 5   -3.181623e-03
## 6   -3.915630e-02
## 7   -5.425100e-03
## 8    8.089084e-03
## 9    3.594701e-03
## 10  -7.333699e-03
## 11   9.808989e-03
## 12   2.511210e-02
## 13   1.000853e-02
## 14  -4.882554e-02
## 15  -4.447849e-02
## 16  -4.347861e-02
## 17   6.733134e-03
## 18  -1.499176e-02
## 19  -4.236107e-03
## 20  -9.153423e-03
## 21   1.106125e-05
## 22  -7.542216e-02
## 23   9.136657e-03
## 24  -1.455386e-02
```

```
## 25  -2.492491e-02
## 26   1.130248e-02
## 27   1.716731e-02
## 28   5.806208e-03
## 29   2.917039e-02
## 30   4.278073e-02
## 31   1.252033e-02
## 32   2.559613e-02
## 33  -3.371786e-02
## 34  -4.134163e-04
## 35  -1.522054e-02
## 36   2.004992e-03
## 37   7.670034e-02
## 38  -5.976531e-02
## 39  -2.399439e-02
## 40   1.282669e-02
## 41   2.561024e-02
## 42  -2.631632e-03
## 43  -2.878010e-02
## 44   1.141358e-02
## 45  -4.347872e-02
## 46   6.775752e-02
## 47  -2.556180e-02
## 48   6.964997e-02
## 49  -2.885624e-02
## 50  -1.346079e-02
## 51   1.901782e-02
## 52   2.713323e-02
## 53   4.678055e-02
## 54  -1.165793e-02
## 55  -4.698992e-02
## 56   7.133638e-02
## 57   1.374859e-02
## 58  -3.396349e-02
## 59  -4.534589e-02
## 60   9.412510e-03
## 61   3.310979e-02
## 62   3.348000e-02
## 63   1.914375e-02
## 64   3.323119e-02
## 65  -2.486496e-03
## 66  -9.576811e-03
## 67  -5.723777e-02
## 68  -4.481005e-02
## 69  -1.101335e-02
## 70  -1.002203e-02
## 71   1.664538e-02
## 72   1.126176e-03
## 73   2.362170e-02
## 74   3.029059e-02
## 75  -2.543768e-03
## 76   1.595370e-02
## 77   3.417173e-02
## 78   1.702087e-02
```

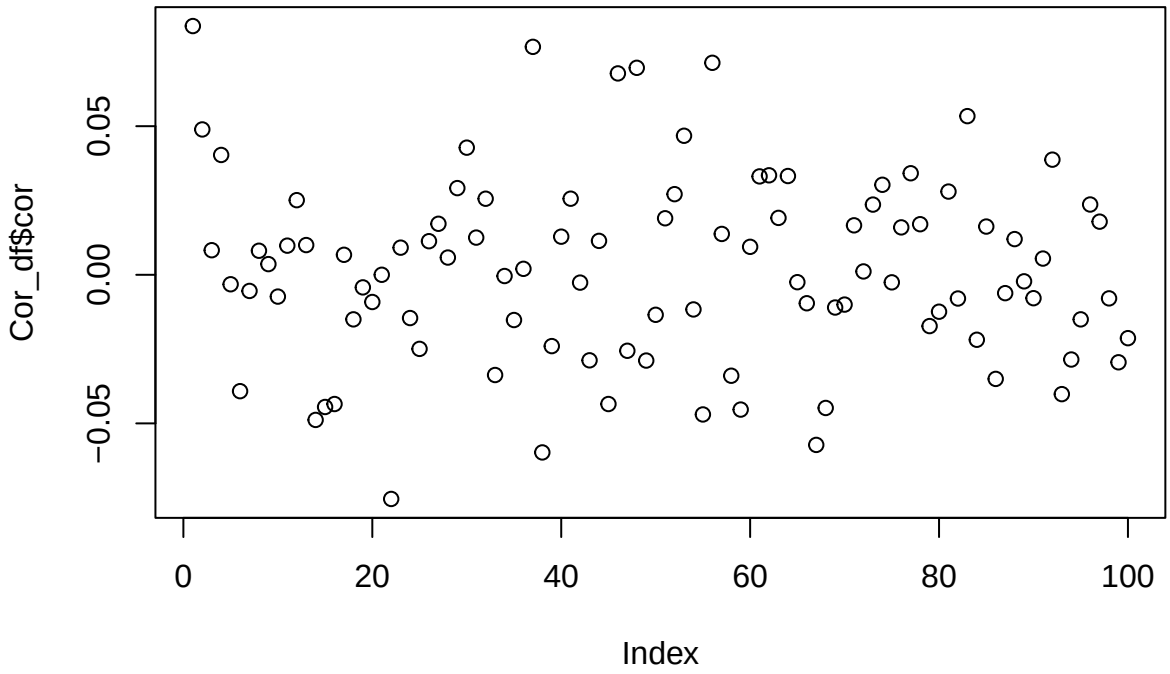
```
## 79  -1.725826e-02
## 80  -1.241931e-02
## 81   2.800401e-02
## 82  -7.976868e-03
## 83   5.338716e-02
## 84  -2.186086e-02
## 85   1.623297e-02
## 86  -3.502888e-02
## 87  -6.176843e-03
## 88   1.201720e-02
## 89  -2.199683e-03
## 90  -7.869507e-03
## 91   5.461537e-03
## 92   3.874570e-02
## 93  -4.015532e-02
## 94  -2.850479e-02
## 95  -1.499518e-02
## 96   2.365430e-02
## 97   1.788602e-02
## 98  -7.929560e-03
## 99  -2.943969e-02
## 100 -2.131575e-02
```

```
hist(Cor_df$cor)
```



We cluster around 0. Which is to be expected as both are just functions of some random noise. No relationship

```r
# We cluster around zero, no correlation, just random noise
plot(Cor_df$cor)
```




```r
mean(Cor_df$cor)
```

```
## [1] 0.001647925
```

```r
#approach zero, true random process, essentially no bias
median(Cor_df$cor)
```

```
## [1] 0.0005686185
```

```r
#average and median are essentially zero
```

```r
#Create random Variable Z
Z<- rnorm(1000)
#X and Y are functions of Z plus noise (more randomness)
X<- Z + rnorm(1000)
Y<- Z + rnorm(1000)
cor(Z, X)
```

```
## [1] 0.7167199
```

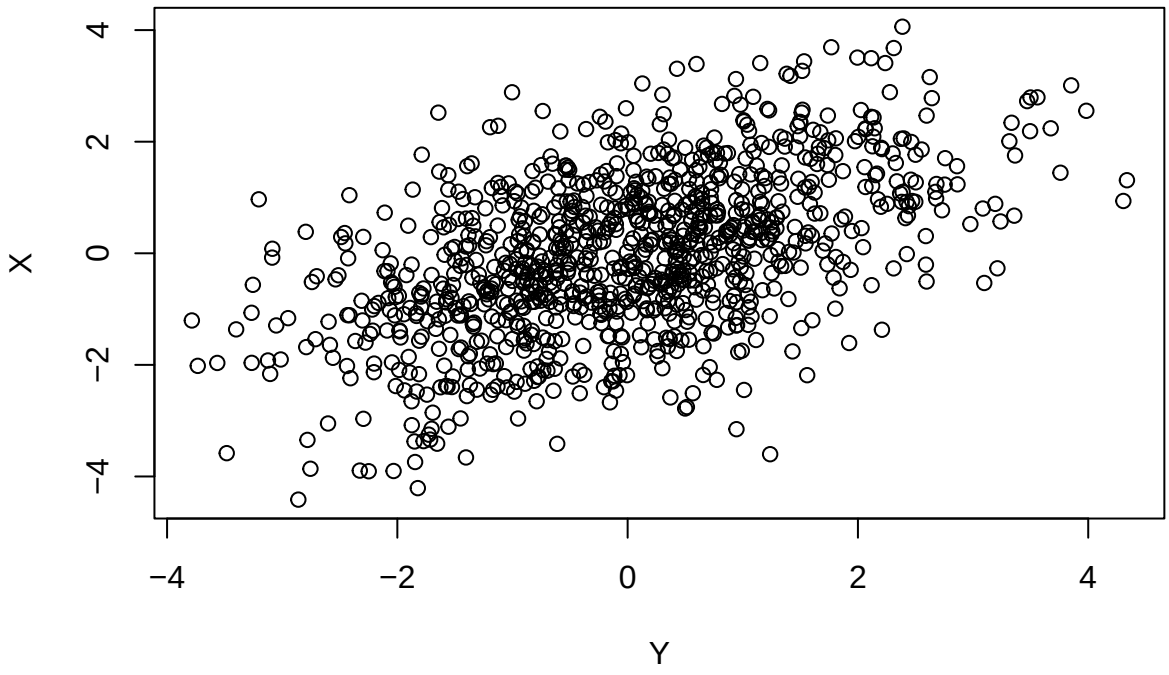
```r
cor(Y, Z)
```

```
## [1] 0.7095271
```

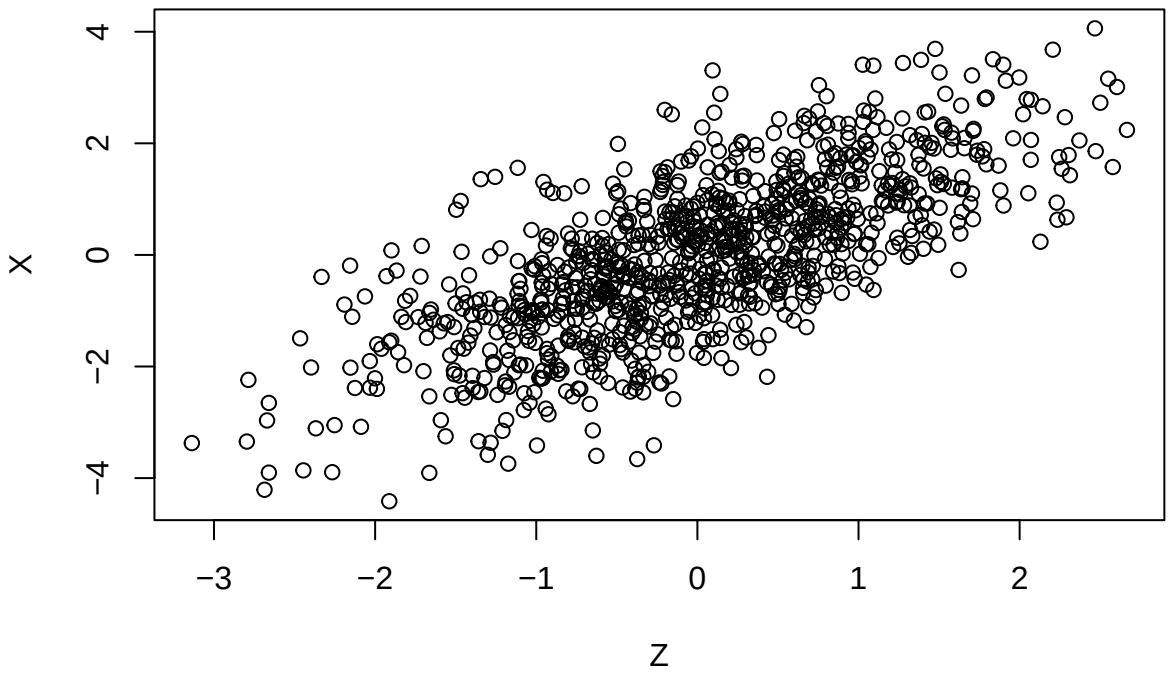
```r
cor(Y,X)
```

```
## [1] 0.5227975
```

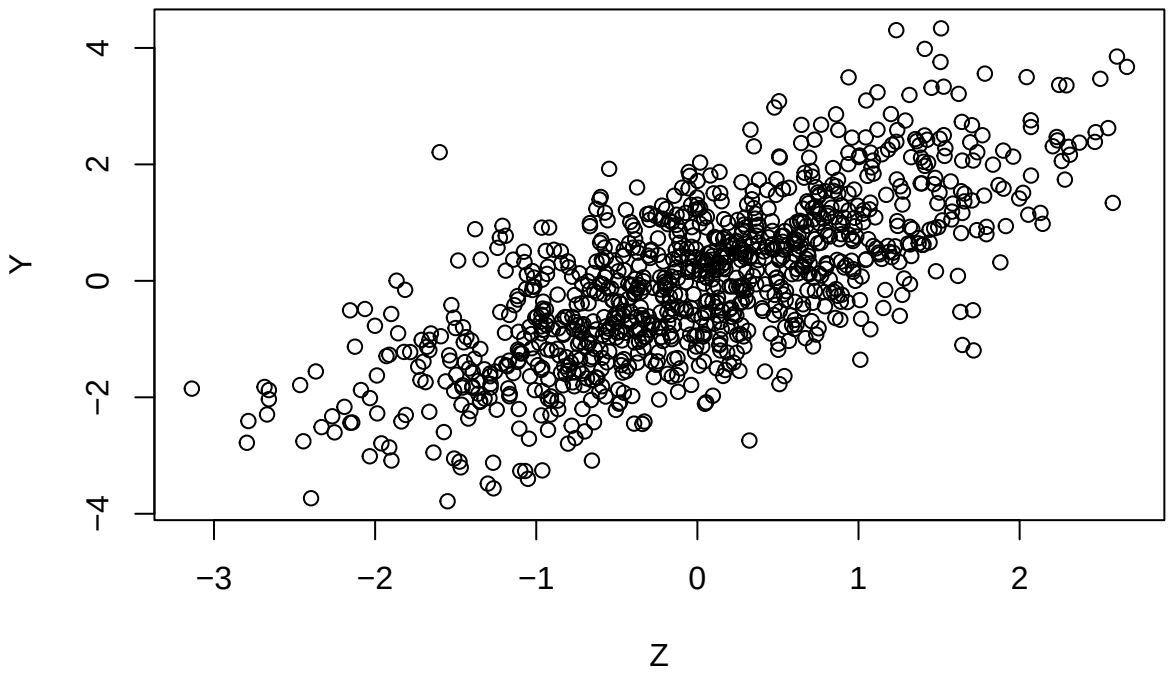
```r
plot(Y,X)
```




```r
plot(Z,X)
```




```
plot(Z,Y)
```



The relationship between z and x is a spurios one as we know that they are not factors of each other and the real unit of interest is Z