НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Кафедра автоматизації проектування енергетичних процесів і систем

Лабораторна робота № 3

З дисципліни: «Big Data та аналіз даних»

**Виконав:**

студент 5 курсу

групи ТР-02мп, ТЕФ

Круглий Д.В.

**Перевірив:**

Матичин І.І.

Київ – 2021

**Постановка задачі**

1. Згенерувати 100 значень для двох предикторів *x*1 та *x*2. Можно seq.
2. Зафіксувати певні значення коефіцієнтів β0, β1, β2.
3. Згенерувати вектор ε зі 100 нормально розподілених похибок.
4. Обчислити вектор залежної змінної *y* за формулою *y* = β0 + β1*x*1 + β2*x*2
5. За допомогою лінійної регресії *y* на *x*1 та *x*2 оцінити значення коефіцієнтів β0, β1, β2. Для цього спочатку використати безпосередню формулу естиматора лінійної регресії, а потім вбудовану функцію lm. Порівняти результати.

**Представлення коду програми**

x1<-sample(0:1000, 100, replace=FALSE, prob=NULL)

x2<-sample(0:1000, 100, replace=FALSE, prob=NULL)

x\_matr<-matrix(c(rep(1, 100), x1, x2), ncol = 3)

b0<-3

b1<-5

b2<-7

eseq<-seq(0, 1, by = .01)

e<-dnorm(eseq, mean=2.5, sd=0.5)

y<-seq(1, 100, by = 1)

for(i in 1:100)

y[i]<-b0+b1\*x1[i]+b2\*x2[i]+e[i]

m1<-lm(y~x1+x2)

c1<-solve(t(x\_matr) %\*% x\_matr) %\*% t(x\_matr) %\*% y

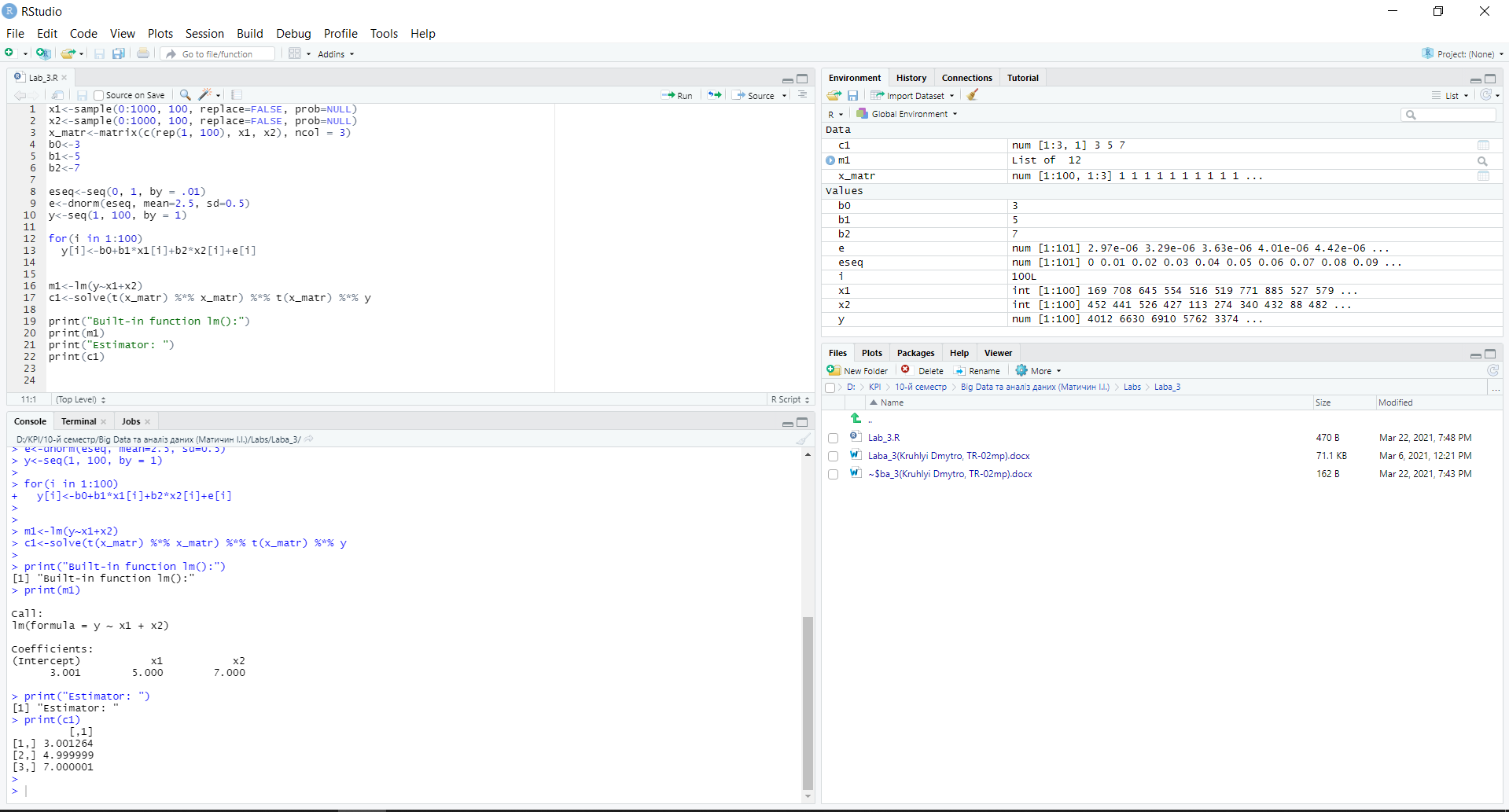
print("Built-in function lm():")

print(m1)

print("Estimator: ")

print(c1)

**Скріншот з результатами**



**Висновок:** виконавши дану лабораторну роботу, я згенерував 100 значень для двох предикторів, навчився фіксувати певні значення коефіцієнтів та згенерувати вектор зі 100 нормально розподілених похибок.