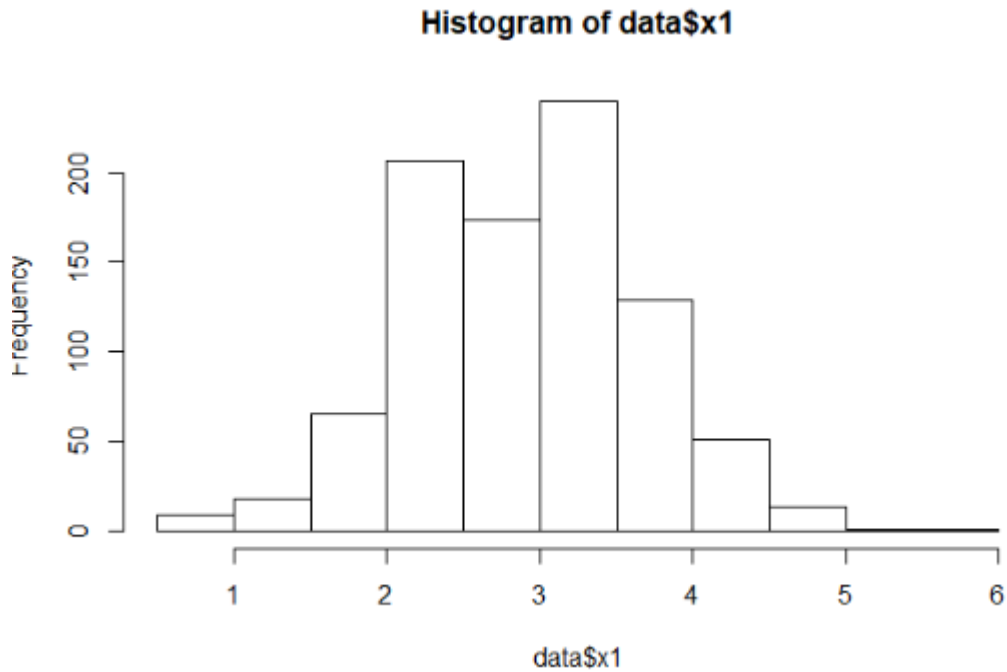


ML-HW2

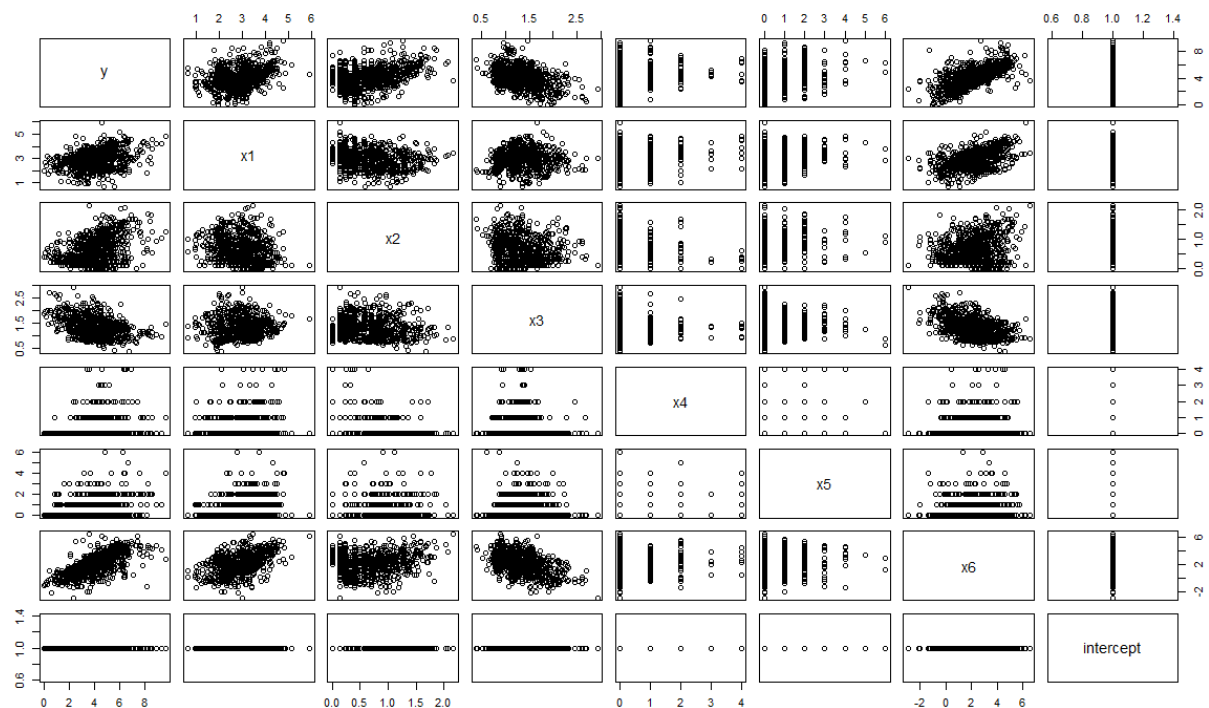
Problem 7: Programming: prediction of acute aquatic toxicity

(a) explore the data.

```
> hist(data$x1)
```



```
> pairs(data)
```



(b) implement the analytical solution to multiple linear regression and compare the obtained estimates with those of the built-in function `lm` in R.

```
>fit <- lm(y ~ x1 + x2 + x3 + x4 + x5 + x6 , toxicity_data)

##Coefficients:
(Intercept)      x1      x2      x3      x4      x5      x6
  2.17456    0.38563    1.25562   -0.74641    0.41355   -0.06433    0.39005
```

`lm` : a function for Fitting Linear Models in R

`AS_MLR`: My analytical function for Fitting multiple Linear Models in R

```
> AS_MLR(data)
      [,1]
x1      0.38562621
x2      1.25562193
x3     -0.74641356
x4      0.41355009
x5      0.06433409
x6      0.39005251
intercept 2.17456016
```

(c) compare cross validation errors of KNN and multiple linear regression on the above dataset.

```
>CV_error_KNN(data,5,1)
[1] 1.342034

> CV_error_MLR(data,5)
[1] 0.6894322
```

`CV_error_MLR > CV_error_KNN`