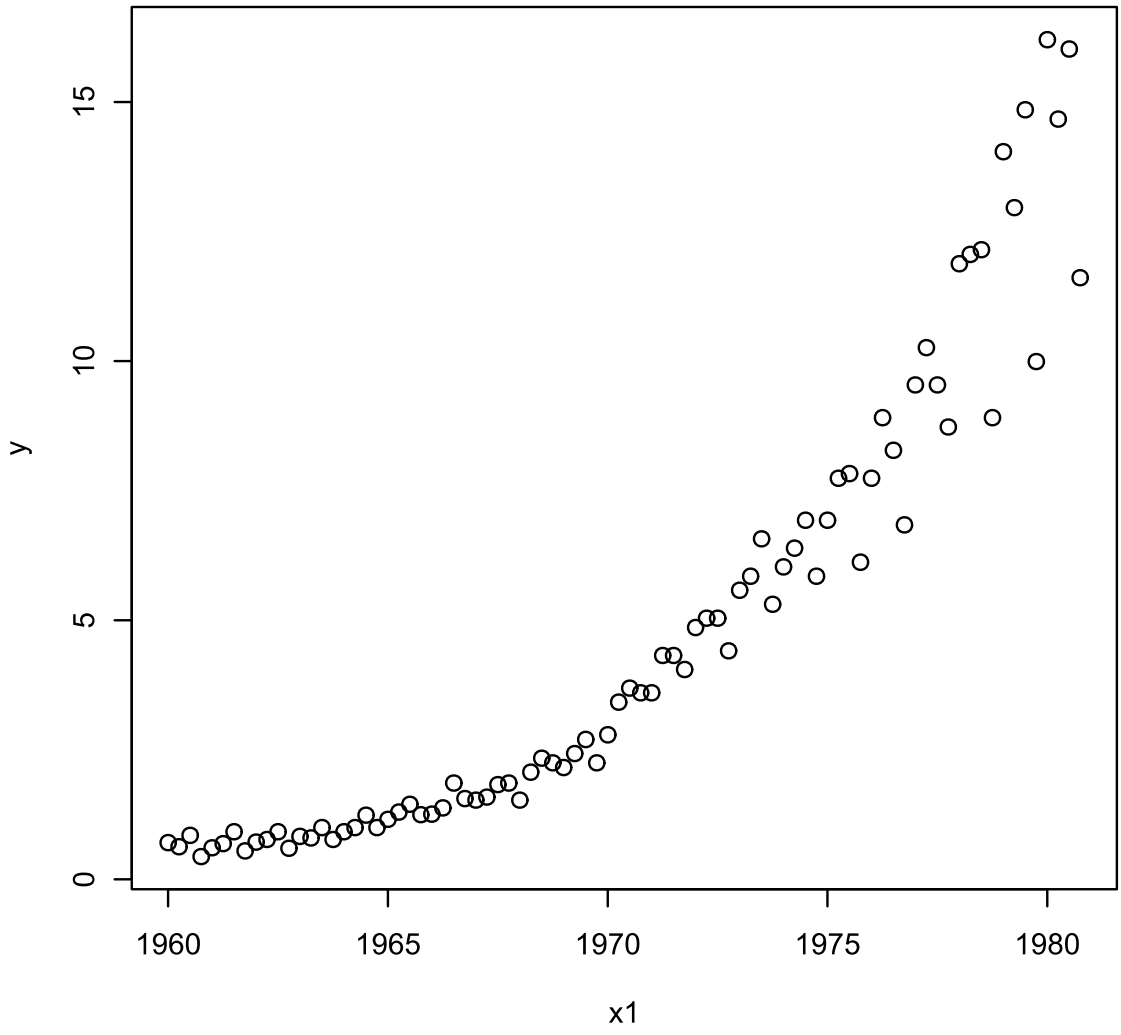
Quick look at assignment datasets

In [85]:
```
data <- read.csv('dataset/Task1A.csv')
length(data)
dim(data)
head(data)
plot(data)
```

2

84 2

	x1	y
1	1960.00	0.71
2	1960.25	0.63
3	1960.50	0.85
4	1960.75	0.44
5	1961.00	0.61
6	1961.25	0.69

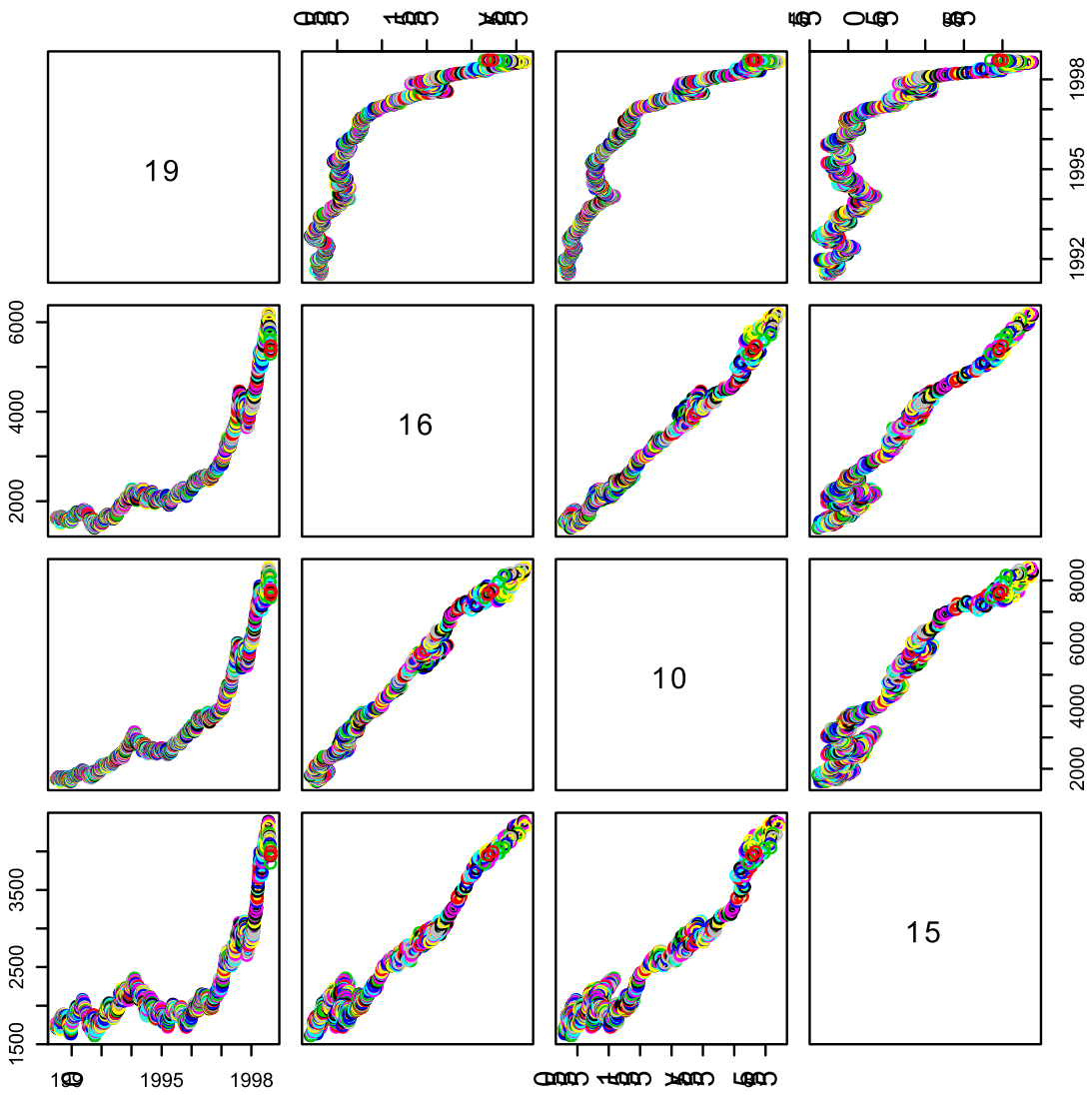


In [102]:
```
data <- read.csv('dataset/Task1B.csv')
length(data)
dim(data)
head(data)
plot(data[,-5], col=data$y+3)
```

5

1860 5

	x1	x2	x3	x4	y
1	1991.496	1628.750	1678.100	1772.800	2443.600
2	1991.50	1613.63	1688.50	1750.50	2460.20
3	1991.504	1606.510	1678.600	1718.000	2448.200
4	1991.508	1621.040	1684.100	1708.100	2470.400
5	1991.512	1618.160	1686.600	1723.100	2484.700
6	1991.515	1610.610	1671.600	1714.300	2466.800

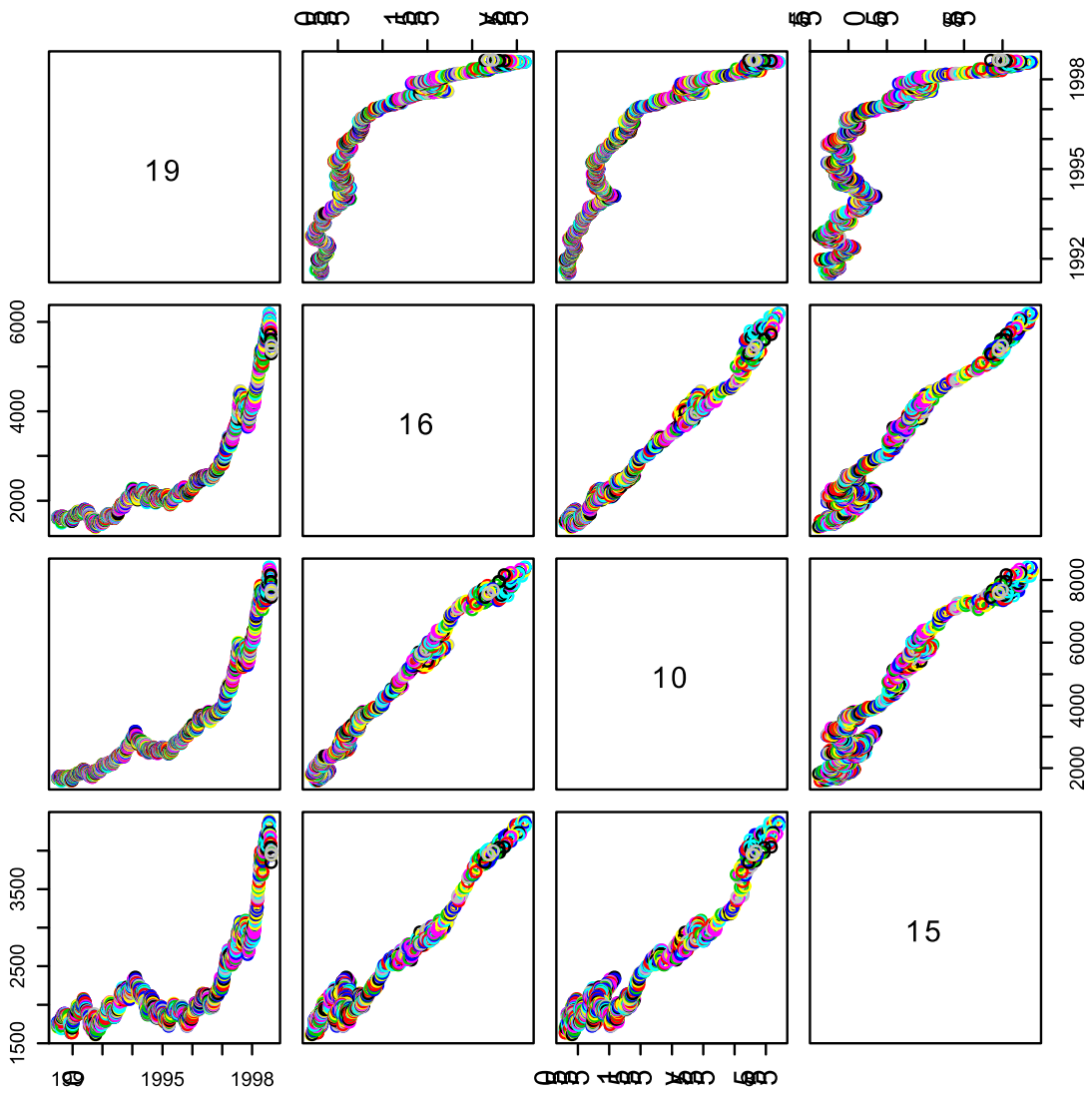


In [101]:
```
data <- read.csv('dataset/Task2A.csv')
length(data)
dim(data)
head(data)
plot(data[,-5], col=data$y+1)
```

5

1860 5

	x1	x2	x3	x4	y
1	1991.496	1628.750	1678.100	1772.800	2443.600
2	1991.50	1613.63	1688.50	1750.50	2460.20
3	1991.504	1606.510	1678.600	1718.000	2448.200
4	1991.508	1621.040	1684.100	1708.100	2470.400
5	1991.512	1618.160	1686.600	1723.100	2484.700
6	1991.515	1610.610	1671.600	1714.300	2466.800

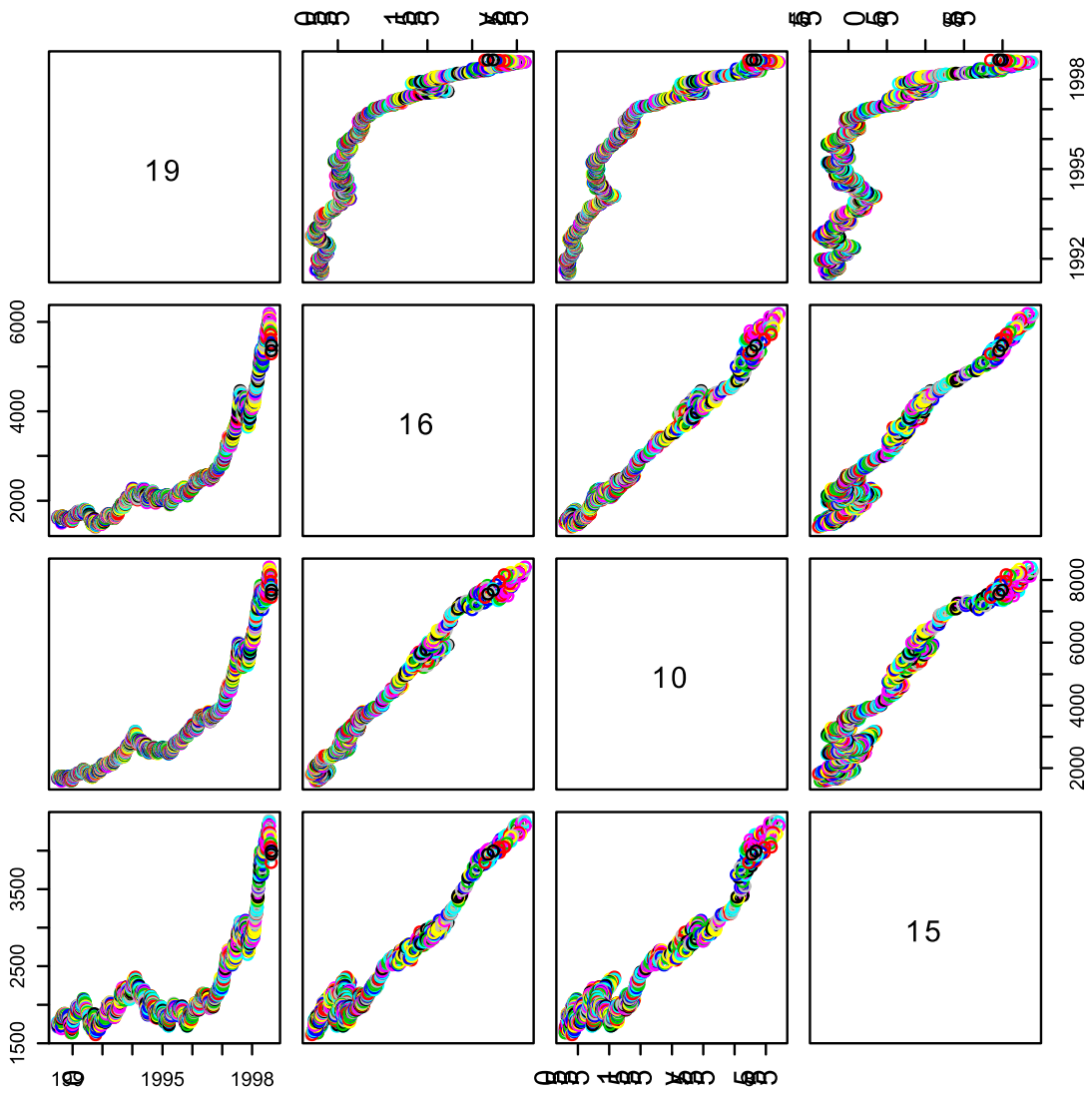


In [100]:
```
data <- read.csv('dataset/Task2B.csv')
length(data)
dim(data)
head(data)
plot(data[,-5], col=data$y+2)
```

5

1860 5

	x1	x2	x3	x4	y
1	1991.496	1628.750	1678.100	1772.800	2443.600
2	1991.50	1613.63	1688.50	1750.50	2460.20
3	1991.504	1606.510	1678.600	1718.000	2448.200
4	1991.508	1621.040	1684.100	1708.100	2470.400
5	1991.512	1618.160	1686.600	1723.100	2484.700
6	1991.515	1610.610	1671.600	1714.300	2466.800

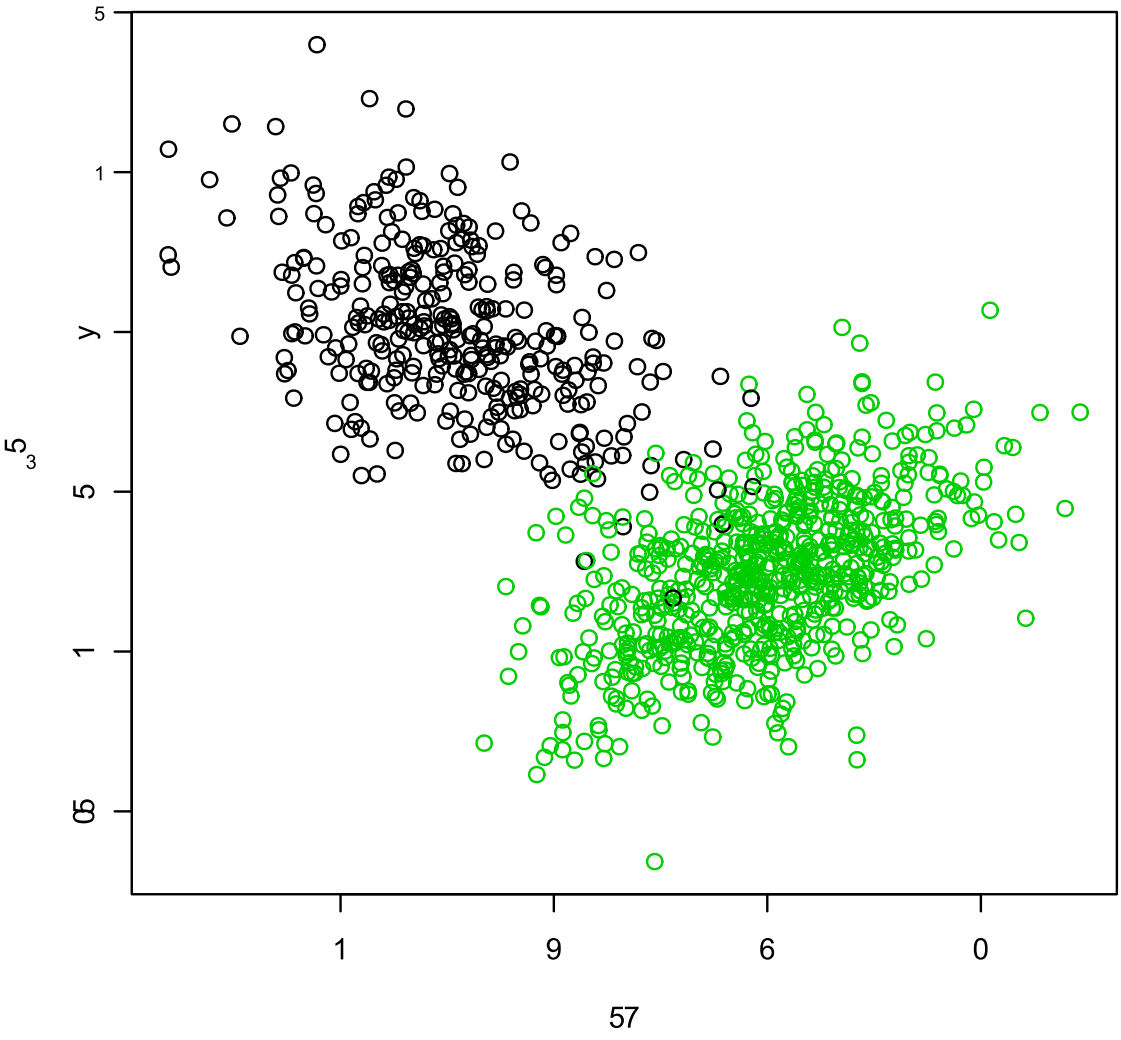


In [99]:
```
data <- read.csv('dataset/Task2D.csv')
length(data)
dim(data)
head(data)
plot(data[,-3], col=data$y+2)
```

3

1000 3

	x1	x2	y
1	-0.4367052	3.1714511	-1.0000000
2	3.4781571	-0.5154699	1.0000000
3	5.104897	1.227815	1.000000
4	4.8959781	0.9532055	1.0000000
5	3.997452	1.049502	1.000000
6	-0.5783557	5.4455353	-1.0000000



In [93]:
```
data <- read.csv('dataset/Task3A.txt')
length(data)
dim(data)
head(data)
```

1

2372 1

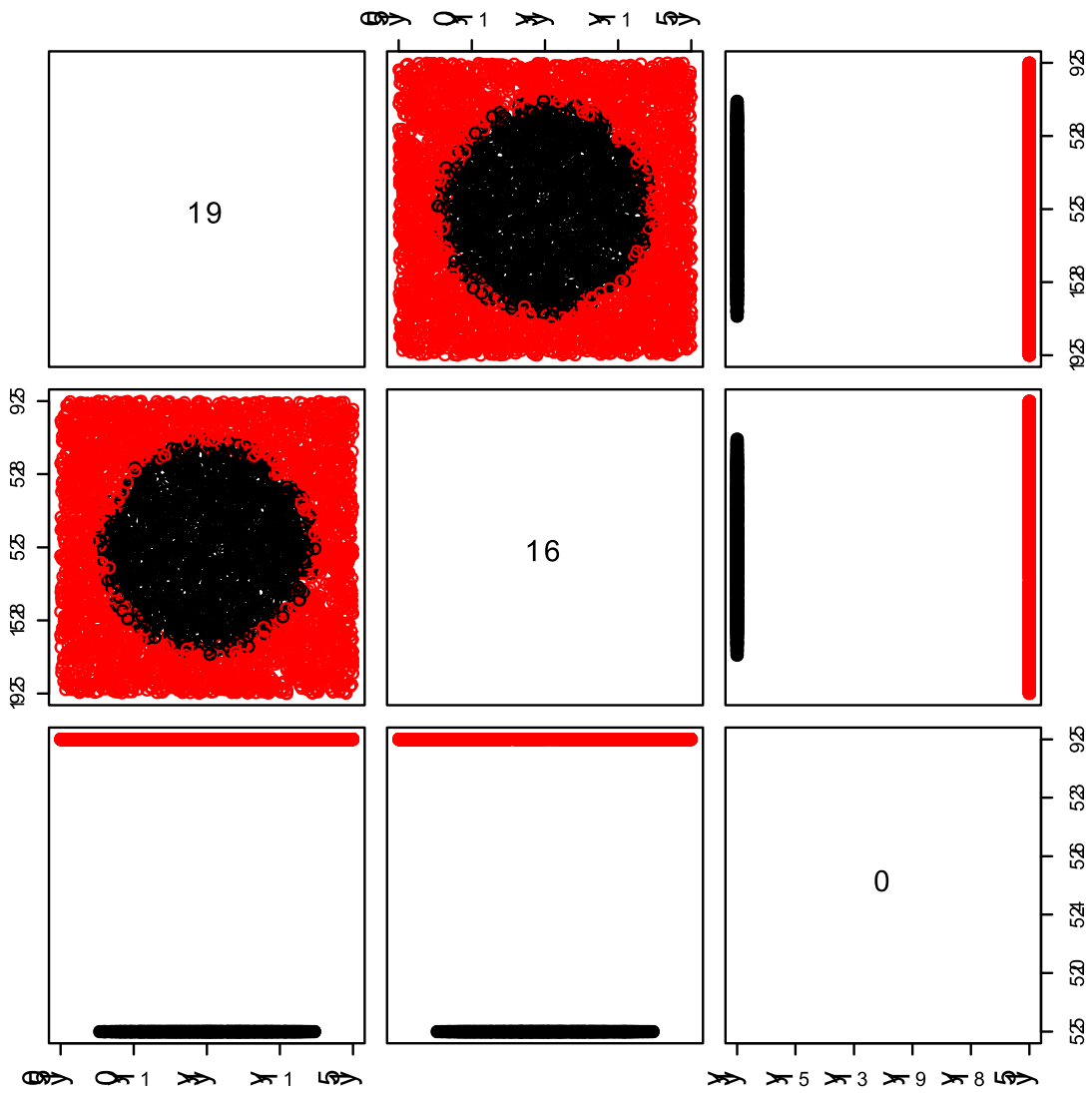
	sci.crypt.ripem.frequently.asked.questions.archive.name.ripem.faq.last.update.sun.m
1	sci.crypt ripem frequently asked questions archive name ripem faq last update mar about thi length etc what is pgp pgp is another cryptographic mail program called pretty good privacy
2	sci.crypt ripem frequently noted vulnerabilities archive name ripem attacks last update mar s
3	sci.crypt certifying authority question answered if you have access to ftp try ftping to rsa com
4	sci.crypt rubber hose cryptanalysis some sick part of me really liked that phrase actually mer
5	sci.crypt declassifying media there are many urban legends maybe this ought to be in the cr
6	sci.crypt re hard drive security for fbi targets from res colnet cmhnet org rob stampfli separat

In [97]:
```
data <- read.csv('dataset/Task3B.csv')
length(data)
dim(data)
head(data)
plot(data, col=data$y+1)
```

3

5000 3

	x1	x2	y
1	-0.4365731	0.3199448	0.0000000
2	0.4973951	-0.5492695	1.0000000
3	-0.3351625	-0.6077436	1.0000000
4	0.2139925	0.6141635	0.0000000
5	0.8914051	-0.1117573	1.0000000
6	-0.4656182	0.5093878	1.0000000



In [103]:
```
data <- read.csv('dataset/Task4A.csv')
length(data)
dim(data)
head(data)
plot(data)
```

4

150 4

	x1	x2	x3	x4
1	5.1	3.5	1.4	0.2
2	4.9	3.0	1.4	0.2
3	4.7	3.2	1.3	0.2
4	4.6	3.1	1.5	0.2
5	5.0	3.6	1.4	0.2
6	5.4	3.9	1.7	0.4