In [2]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

data = pd.read_csv("veriler.csv")
Y = data["cinsiyet"].values
X1 = data["kilo"].values
X2 = data["yas"].values
X3 = data["boy"].values
data
```

Out[2]:

	boy	kilo	yas	cinsiyet
0	129	38	12	1
1	180	90	30	1
2	190	80	25	1
3	175	90	35	1
4	177	60	22	0
5	185	105	33	1
6	165	55	27	0
7	155	50	44	0
8	160	58	39	0
9	162	59	41	0
10	167	62	55	0
11	174	70	47	1
12	193	90	23	1
13	187	80	27	1
14	183	88	28	1
15	159	40	29	0

In [3]:

```
X_123 = data.drop("cinsiyet",axis=1)
X_123
```

Out[3]:

	boy	kilo	yas
0	129	38	12
1	180	90	30
2	190	80	25
3	175	90	35
4	177	60	22
5	185	105	33
6	165	55	27
7	155	50	44
8	160	58	39

	boy	kilo	yas
9	162	59	41
10	167	62	55
11	174	70	47
12	193	90	23
13	187	80	27
14	183	88	28
15	159	40	29

In [4]:
```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(X_123,Y,test_size=0.3)
```

In [7]:
```
from sklearn.linear_model import LogisticRegression
log_reg = LogisticRegression()
log_reg.fit(x_train,y_train)
tahmin = log_reg.predict(x_test)
tahmin
```

Out[7]:
```
array([0, 0, 1, 1, 1], dtype=int64)
```

In [8]:
```
y_test
```

Out[8]:
```
array([0, 1, 1, 1, 1], dtype=int64)
```

In [9]:
```
tahminler = pd.Series(tahmin)
gercek_degerler = pd.Series(y_test)
data = dict(cinsiyet=gercek_degerler, tahmin=tahminler)
comparison_df = pd.DataFrame(data)
comparison_df
```

Out[9]:

	cinsiyet	tahmin
0	0	0
1	1	0
2	1	1
3	1	1
4	1	1

In [10]:
```
# sınıflandırma değerlendirmesi CONFUSION MATRIX
from sklearn.metrics import confusion_matrix
result = confusion_matrix(y_test,tahmin)
result
```

Out[10]:
```
array([[1, 0],
       [1, 3]], dtype=int64)
```

In [12]:
```
# toplam veriyi hesapla: sütunları topla;
#  (1+1) + (0+3) = 5
# toplam doğruyu hesapla: diyagonel sayıları topla;
```

```
#  3+1 = 4
# doğruluk hesapla;
#  4/5 = 0.8 --> %80
```

In [ ]: