In [26]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
```

In [27]:
```
dt = pd.read_csv('Titanic.csv')
```

In [28]:
```
dt.head(10)
```

Out[28]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708

In [29]:
```
dt.drop(['PassengerId','Name','SibSp','Parch','Ticket','Cabin','Embarked'],
```

In [30]:
```
dt.head(10)
```

Out[30]:

	Survived	Pclass	Sex	Age	Fare
0	0	3	male	22.0	7.2500
1	1	1	female	38.0	71.2833
2	1	3	female	26.0	7.9250
3	1	1	female	35.0	53.1000
4	0	3	male	35.0	8.0500
5	0	3	male	NaN	8.4583
6	0	1	male	54.0	51.8625
7	0	3	male	2.0	21.0750
8	1	3	female	27.0	11.1333
9	1	2	female	14.0	30.0708

In [31]:
```
Survive = dt.drop('Survived' , axis = 'columns')
sur_vive = dt.Survived
```

In [32]:
```
Survive.Sex = Survive.Sex.map({'male' : 1 , 'female' :0 })
```

In [33]:
```
Survive.Age[:10]
```

Out[33]:
```
0    22.0
1    38.0
2    26.0
3    35.0
4    35.0
5     NaN
6    54.0
7     2.0
8    27.0
9    14.0
Name: Age, dtype: float64
```

In [34]:
```
Survive.Age = Survive.Age.fillna(Survive.Age.mean())
```

In [35]:
```
Survive.head(10)
```

Out[35]:

	Pclass	Sex	Age	Fare
0	3	1	22.000000	7.2500
1	1	0	38.000000	71.2833
2	3	0	26.000000	7.9250
3	1	0	35.000000	53.1000
4	3	1	35.000000	8.0500
5	3	1	29.699118	8.4583
6	1	1	54.000000	51.8625
7	3	1	2.000000	21.0750
8	3	0	27.000000	11.1333
9	2	0	14.000000	30.0708

In [38]:
```
X_train, X_test, y_train, y_test = train_test_split(Survive,sur_vive,test_s
```

In [39]:
```
len(X_train)
```

Out[39]: 712

In [40]:
```
len(X_test)
```

Out[40]: 179

In [41]:
```
from sklearn import tree
model = tree.DecisionTreeClassifier()
```

In [42]:
```
model.fit(X_train,y_train)
```

Out[42]: DecisionTreeClassifier()

In [43]:
```
model.score(X_test,y_test)
```

Out[43]: 0.7486033519553073