

```
In [24]: import numpy as np
import pandas as pd
df = pd.read_csv("suv_data.csv")
df.head()
```

```
Out[24]:
```

	User ID	Gender	Age	EstimatedSalary	Purchased
0	15624510	Male	19	19000	0
1	15810944	Male	35	20000	0
2	15668575	Female	26	43000	0
3	15603246	Female	27	57000	0
4	15804002	Male	19	76000	0

```
In [25]: gender = pd.get_dummies(df["Gender"],drop_first=True)
```

```
In [26]: gender.head()
```

```
Out[26]:
```

	Male
0	1
1	1
2	0
3	0
4	1

```
In [27]: # df.insert(2,"Male",gender)
```

```
In [28]: # df
```

```
Out[28]:
```

	User ID	Gender	Male	Age	EstimatedSalary	Purchased
0	15624510	Male	1	19	19000	0

	User ID	Gender	Male	Age	EstimatedSalary	Purchased
1	15810944	Male	1	35	20000	0
2	15668575	Female	0	26	43000	0
3	15603246	Female	0	27	57000	0
4	15804002	Male	1	19	76000	0
...
395	15691863	Female	0	46	41000	1
396	15706071	Male	1	51	23000	1
397	15654296	Female	0	50	20000	1
398	15755018	Male	1	36	33000	0
399	15594041	Female	0	49	36000	1

400 rows × 6 columns

```
In [64]: df.drop("Male",axis=1,inplace=True)
```

```
In [65]: df.shape
```

```
Out[65]: (400, 3)
```

```
In [74]: df.head()
```

```
Out[74]:
```

	Age	EstimatedSalary	Purchased
0	19	19000	0
1	35	20000	0
2	26	43000	0
3	27	57000	0
4	19	76000	0

```
In [78]: x = df.iloc[:,[0,1]].values  
y = df.iloc[:,2].values  
y
```

```
Out[78]: array([0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1,  
1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,  
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1,  
0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0,  
1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0,  
1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1,  
0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1,  
1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1,  
0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0,  
1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1,  
0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1,  
1, 1, 0, 1], dtype=int64)
```

```
In [79]: from sklearn.model_selection import train_test_split  
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size = 0.2,random_state=0)
```

```
In [80]: from sklearn.preprocessing import StandardScaler  
sc = StandardScaler()
```

```
In [81]: x_train = sc.fit_transform(x_train)  
x_test = sc.transform(x_test)
```

```
In [82]: from sklearn.linear_model import LogisticRegression  
rissh = LogisticRegression()
```

```
In [83]: rissh.fit(x_train,y_train)
```

```
Out[83]: LogisticRegression()
```

```
In [84]: Purchase_predict = rissh.predict(x_test)
```

```
In [102... Purchase_predict
```

```
Out[102... array([0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1,
        0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
        1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1,
        0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1], dtype=int64)
```

```
In [89]: from sklearn.metrics import accuracy_score
accuracy_score(y_test,Purchase_predict)*100
```

```
Out[89]: 92.5
```

```
In [99]: # Predicting a person of 24 years and earning 100000
rissh.predict([[24,100000]])
```

```
Out[99]: array([1], dtype=int64)
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```