In [42]:
```python
import pandas as pd
from sklearn.model_selection import train_test_split

data=pd.read_csv("Data-Week2.csv")
labels=data["Low Income"]
#split ratio defines the ratio of data that is considered for testing(Lets consider 70
split_ratio = 0.7

# Splitting data and labels into train and test sets
X_train, X_test, y_train, y_test = train_test_split(data, labels, test_size=1 - split_
X_train
```

Out[42]:

	State-County FIPS	Total Active Borrowers	Very Low Income	Low Income	Married	American Indian/Alaskan Native	Asian	Black/African American	Hawaiin Pacific Islander
1933	38041.0	2	1.0	1.0	1	0.0	0.0	0.0	0.0
2372	47059.0	202	72.0	131.0	59	0.0	0.0	10.0	0.0
655	17187.0	54	23.0	31.0	13	1.0	1.0	1.0	0.0
1659	31131.0	28	13.0	15.0	14	0.0	0.0	0.0	0.0
999	21093.0	49	24.0	25.0	10	0.0	0.0	6.0	0.0
...	...	...	...	...	...	...	...	...	...
3092	72115.0	94	50.0	43.0	42	0.0	0.0	0.0	0.0
1095	22045.0	101	45.0	56.0	20	0.0	2.0	62.0	0.0
1130	22115.0	101	44.0	57.0	23	2.0	0.0	46.0	0.0
1294	27051.0	10	5.0	5.0	4	0.0	0.0	0.0	0.0
860	20011.0	54	24.0	30.0	18	2.0	0.0	5.0	0.0

2180 rows × 13 columns

In [43]:
```python
X_test
```

Out[43]:

	State-County FIPS	Total Active Borrowers	Very Low Income	Low Income	Married	American Indian/Alaskan Native	Asian	Black/African American	Hawaiin Pacific Islander
1510	29147.0	42	13.0	29.0	20	0.0	0.0	0.0	0.0
1068	21231.0	112	40.0	73.0	41	1.0	0.0	1.0	0.0
567	17003.0	47	19.0	27.0	9	0.0	0.0	18.0	0.0
2901	54039.0	160	72.0	89.0	49	1.0	0.0	8.0	0.0
1467	29061.0	15	2.0	13.0	8	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...
2337	46123.0	15	4.0	10.0	5	0.0	0.0	0.0	0.0
2826	51580.0	9	4.0	5.0	1	0.0	0.0	1.0	0.0
626	17129.0	9	6.0	3.0	3	0.0	0.0	0.0	0.0
347	12109.0	161	50.0	105.0	28	1.0	3.0	44.0	0.0
278	8113.0	6	6.0	0.0	3	0.0	0.0	0.0	0.0

935 rows × 13 columns

In [ ]: