In [15]:
```
import pandas as pd
```

In [16]:
```
df = pd.read_csv('new_test.csv')
```

In [17]:
```
df.head(10)
```

Out[17]:

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign
0	32	4	0	6	0	0	0	0	3	3	131	5
1	37	10	3	6	0	0	0	0	4	3	100	1
2	55	5	0	5	1	2	0	0	3	2	131	2
3	44	2	1	0	1	0	0	1	4	3	48	2
4	28	0	2	3	0	0	0	0	5	0	144	2
5	45	10	1	2	0	0	0	0	1	0	126	3
6	55	0	0	6	0	0	0	0	7	0	189	1
7	46	0	1	3	0	0	0	1	4	3	107	3
8	53	1	2	2	0	2	0	0	3	4	355	1
9	30	0	2	6	0	0	0	0	1	1	135	1

In [18]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8238 entries, 0 to 8237
Data columns (total 13 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   age          8238 non-null   int64
 1   job          8238 non-null   int64
 2   marital      8238 non-null   int64
 3   education    8238 non-null   int64
 4   default      8238 non-null   int64
 5   housing      8238 non-null   int64
 6   loan         8238 non-null   int64
 7   contact      8238 non-null   int64
 8   month        8238 non-null   int64
 9   day_of_week  8238 non-null   int64
 10  duration     8238 non-null   int64
 11  campaign     8238 non-null   int64
 12  poutcome     8238 non-null   int64
dtypes: int64(13)
memory usage: 836.8 KB
```

In [19]:
```
df.describe()
```

Out[19]:

	age	job	marital	education	default	housing	loan	c
count	8238.000000	8238.000000	8238.000000	8238.000000	8238.000000	8238.000000	8238.000000	8238.(
mean	39.613498	3.729425	1.170430	3.740592	0.201141	1.072833	0.320952	0.:
std	9.021371	3.597274	0.610371	2.134617	0.400878	0.985527	0.717918	0.4
min	26.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.(
25%	32.000000	0.000000	1.000000	2.000000	0.000000	0.000000	0.000000	0.(
50%	38.000000	2.000000	1.000000	3.000000	0.000000	2.000000	0.000000	0.(
75%	47.000000	7.000000	2.000000	6.000000	0.000000	2.000000	0.000000	1.(
max	55.000000	11.000000	3.000000	7.000000	1.000000	2.000000	2.000000	1.(

In [20]:
```
X = df.drop('education', axis=1)
```

In [21]:
```
X
```

Out[21]:

	age	job	marital	default	housing	loan	contact	month	day_of_week	duration	campaign	poutc
0	32	4	0	0	0	0	0	3	3	131	5	
1	37	10	3	0	0	0	0	4	3	100	1	
2	55	5	0	1	2	0	0	3	2	131	2	
3	44	2	1	1	0	0	1	4	3	48	2	
4	28	0	2	0	0	0	0	5	0	144	2	
...	...	...	...	...	...	...	...	...	...	...	...	...
8233	48	4	1	0	2	0	0	6	3	554	1	
8234	30	7	2	0	2	0	0	6	0	159	1	
8235	33	7	1	0	0	0	0	4	1	472	1	
8236	44	1	1	0	2	2	1	6	1	554	5	
8237	42	1	1	1	2	0	0	6	3	83	5	

8238 rows × 12 columns

In [22]:
```
y = df['education']
```

In [23]:
```
y
```

Out[23]:
```
0       6
1       6
2       5
3       0
4       3
       ..
8233    2
8234    3
8235    3
8236    1
8237    2
Name: education, Length: 8238, dtype: int64
```

In [24]:
```
from sklearn.model_selection import train_test_split
```

In [25]:
```
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2,random_state=42
```

In [26]:
```
from sklearn.tree import DecisionTreeClassifier
```

In [27]:
```
cd = DecisionTreeClassifier(random_state=42)
cd.fit(X_train, y_train)
```

Out[27]:
```
DecisionTreeClassifier
DecisionTreeClassifier(random_state=42)
```

In [28]:
```
y_predict = cd.predict(X_test)
```

In [29]:
```
from sklearn.metrics import accuracy_score
```

In [30]:
```
accu = accuracy_score(y_test, y_predict)
```

In [31]:
```
print("Accuracy:",accu)
```

```
Accuracy: 0.38106796116504854
```

In [ ]: