In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
bank_df=pd.read_csv("bank.csv",sep=";")
bank_df
```

Out[2]:

	age	job	marital	education	default	balance	housing	loan	contact	da
0	30	unemployed	married	primary	no	1787	no	no	cellular	
1	33	services	married	secondary	no	4789	yes	yes	cellular	
2	35	management	single	tertiary	no	1350	yes	no	cellular	
3	30	management	married	tertiary	no	1476	yes	yes	unknown	
4	59	blue-collar	married	secondary	no	0	yes	no	unknown	
...	...	...	...	...	...	...	...	...	...	
4516	33	services	married	secondary	no	-333	yes	no	cellular	
4517	57	self-employed	married	tertiary	yes	-3313	yes	yes	unknown	
4518	57	technician	married	secondary	no	295	no	no	cellular	
4519	28	blue-collar	married	secondary	no	1137	no	no	cellular	
4520	44	entrepreneur	single	tertiary	no	1136	yes	yes	cellular	

4521 rows × 17 columns

In [3]:
```
bank_df=pd.read_csv("bank.csv",sep=";")
unique_label=sorted(bank_df["job"].unique())
list1=[i for i in range(len(unique_label))]
dict1={i:j for i,j in zip(unique_label,list1)}
bank_df["job_update"]=bank_df["job"].map(dict1)
bank_df
```

Out[3]:

	age	job	marital	education	default	balance	housing	loan	contact	da
0	30	unemployed	married	primary	no	1787	no	no	cellular	
1	33	services	married	secondary	no	4789	yes	yes	cellular	
2	35	management	single	tertiary	no	1350	yes	no	cellular	
3	30	management	married	tertiary	no	1476	yes	yes	unknown	
4	59	blue-collar	married	secondary	no	0	yes	no	unknown	
...	...	...	...	...	...	...	...	...	...	
4516	33	services	married	secondary	no	-333	yes	no	cellular	
4517	57	self-employed	married	tertiary	yes	-3313	yes	yes	unknown	
4518	57	technician	married	secondary	no	295	no	no	cellular	
4519	28	blue-collar	married	secondary	no	1137	no	no	cellular	
4520	44	entrepreneur	single	tertiary	no	1136	yes	yes	cellular	

4521 rows × 18 columns

In [4]:
```
bank_df=pd.read_csv("bank.csv",sep=";")
cat_col=bank_df.select_dtypes(include="object").columns
for i in cat_col:
    unique_label=sorted(bank_df[i].unique())
    list1=[i for i in range(len(unique_label))]
    dict1={i:j for i,j in zip(unique_label,list1)}
    bank_df[i]=bank_df[i].map(dict1)
bank_df
```

Out[4]:

	age	job	marital	education	default	balance	housing	loan	contact	day	month
0	30	10	1	0	0	1787	0	0	0	19	10
1	33	7	1	1	0	4789	1	1	0	11	8
2	35	4	2	2	0	1350	1	0	0	16	0
3	30	4	1	2	0	1476	1	1	2	3	6
4	59	1	1	1	0	0	1	0	2	5	8
...	...	...	...	...	...	...	...	...	...	...	...
4516	33	7	1	1	0	-333	1	0	0	30	5
4517	57	6	1	2	1	-3313	1	1	2	9	8
4518	57	9	1	1	0	295	0	0	0	19	1
4519	28	1	1	1	0	1137	0	0	0	6	3
4520	44	2	2	2	0	1136	1	1	0	3	0

4521 rows × 17 columns

In [19]:

```
bank_df=pd.read_csv("bank.csv",sep=";")
cat_col=bank_df.select_dtypes(include="object").columns

from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
for i in cat_col:
    bank_df[f"{i}_new"]=le.fit_transform(bank_df[i])
bank_df
```

Out[19]:

	age	job	marital	education	default	balance	housing	loan	contact	da
0	30	unemployed	married	primary	no	1787	no	no	cellular	
1	33	services	married	secondary	no	4789	yes	yes	cellular	
2	35	management	single	tertiary	no	1350	yes	no	cellular	
3	30	management	married	tertiary	no	1476	yes	yes	unknown	
4	59	blue-collar	married	secondary	no	0	yes	no	unknown	
...	...	...	...	...	...	...	...	...	...	
4516	33	services	married	secondary	no	-333	yes	no	cellular	
4517	57	self-employed	married	tertiary	yes	-3313	yes	yes	unknown	
4518	57	technician	married	secondary	no	295	no	no	cellular	
4519	28	blue-collar	married	secondary	no	1137	no	no	cellular	
4520	44	entrepreneur	single	tertiary	no	1136	yes	yes	cellular	

4521 rows × 27 columns

In [21]:
```
bank_df=pd.read_csv("bank.csv",sep=";")
cat_col=bank_df.select_dtypes(include="object").columns

from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
for i in cat_col:
    bank_df[i]=le.fit_transform(bank_df[i])
bank_df
```

Out[21]:

	age	job	marital	education	default	balance	housing	loan	contact	day	month
0	30	10	1	0	0	1787	0	0	0	19	10
1	33	7	1	1	0	4789	1	1	0	11	8
2	35	4	2	2	0	1350	1	0	0	16	0
3	30	4	1	2	0	1476	1	1	2	3	6
4	59	1	1	1	0	0	1	0	2	5	8
...	...	...	...	...	...	...	...	...	...	...	...
4516	33	7	1	1	0	-333	1	0	0	30	5
4517	57	6	1	2	1	-3313	1	1	2	9	8
4518	57	9	1	1	0	295	0	0	0	19	1
4519	28	1	1	1	0	1137	0	0	0	6	3
4520	44	2	2	2	0	1136	1	1	0	3	0

4521 rows × 17 columns

In [31]:
```
bank_df=pd.read_csv("bank.csv",sep=";")
bank_df["housing"].unique()
```

Out[31]:
```
array(['no', 'yes'], dtype=object)
```

In [33]:
```
#np.where
con=bank_df["housing"]=="no"
np.where(con,0,1)
```

Out[33]:
```
array([0, 1, 1, ..., 0, 0, 1])
```

In [35]:
```
con=bank_df["housing"]=="no"
np.where(con,0,1)
```

Out[35]:
```
array([0, 1, 1, ..., 0, 0, 1])
```

In [ ]: