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

In [11]:
```
#read the bank data
path=r"C:\Users\Mrityunjay\Desktop\Data science naresh it\Class notes by me\bank.cs
bank_df=pd.read_csv(path,sep=";")
bank_df
```

Out[11]:

	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 [15]:
```
colm=bank_df.columns
colm
```

Out[15]:
```
Index(['age', 'job', 'marital', 'education', 'default', 'balance', 'housing',
       'loan', 'contact', 'day', 'month', 'duration', 'campaign', 'pdays',
       'previous', 'poutcome', 'y'],
      dtype='object')
```

In [19]:
```
#pick the categorical colomns
cat_colm=[]
for key,value in dict(bank_df.dtypes).items():
    if value=="object":
        cat_colm.append(key)
print(cat_colm)
```

```
['job', 'marital', 'education', 'default', 'housing', 'loan', 'contact', 'month', 'p
outcome', 'y']
```

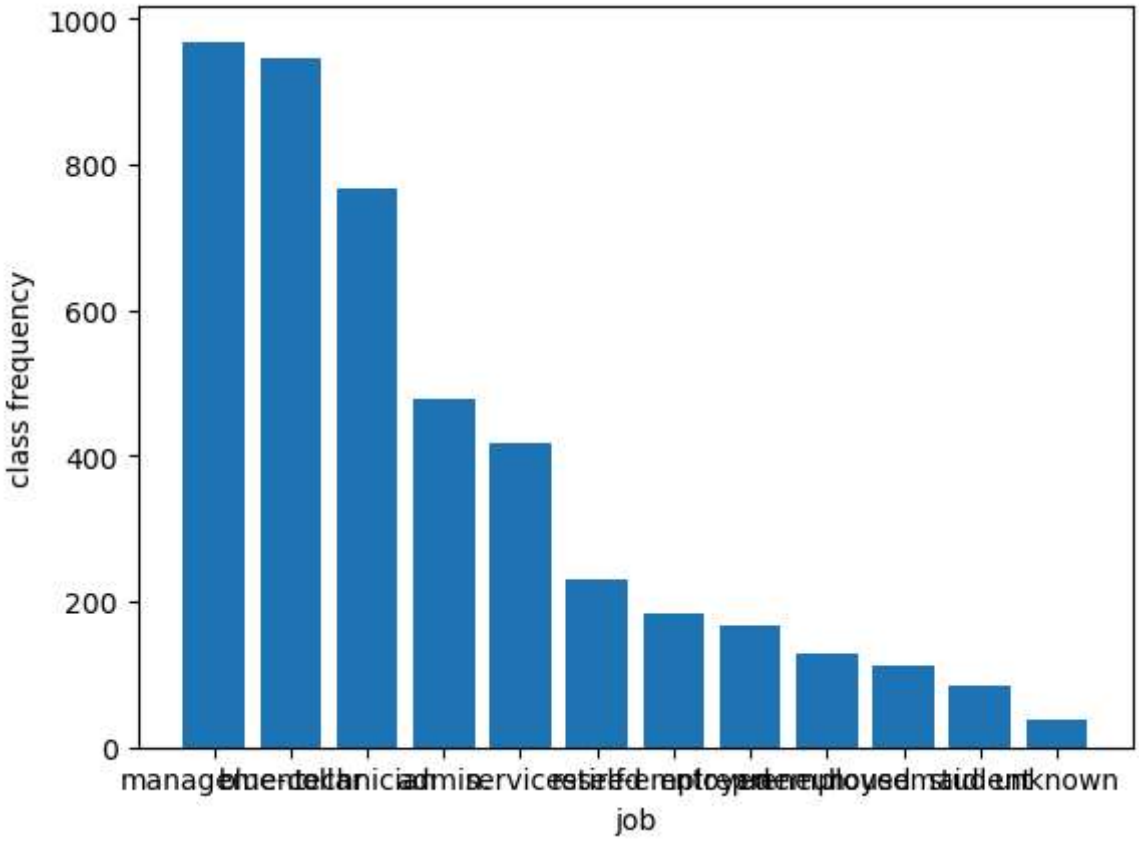
In [23]:
```
import os
os.getcwd()
```

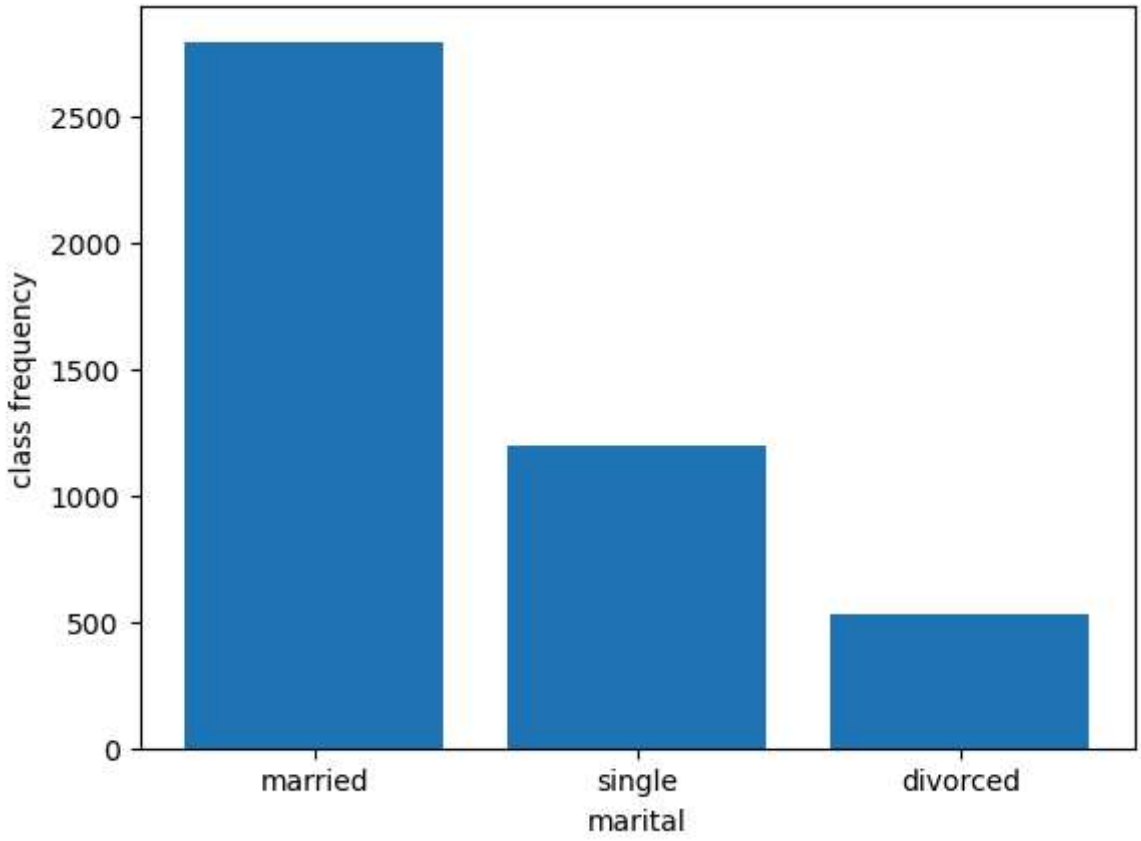
Out[23]:
```
'C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiments'
```

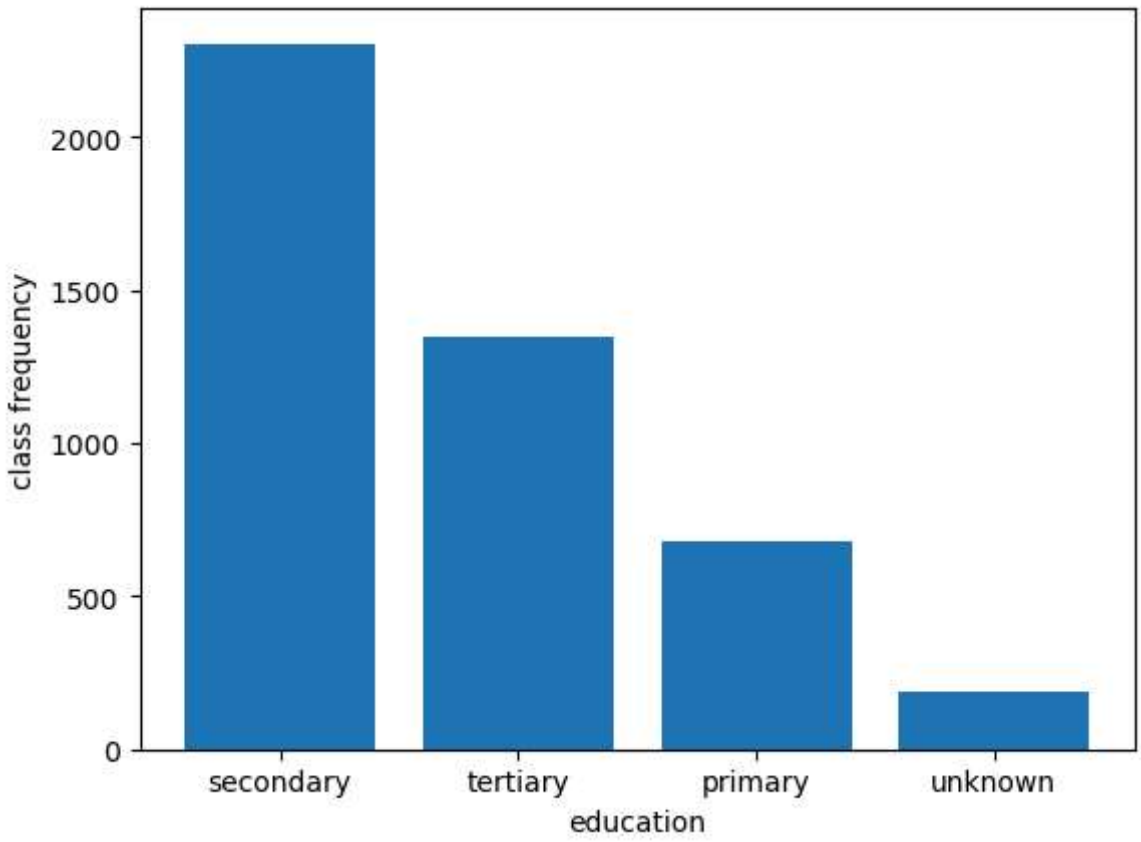
In [41]:
```
os.makedirs('C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiment
```

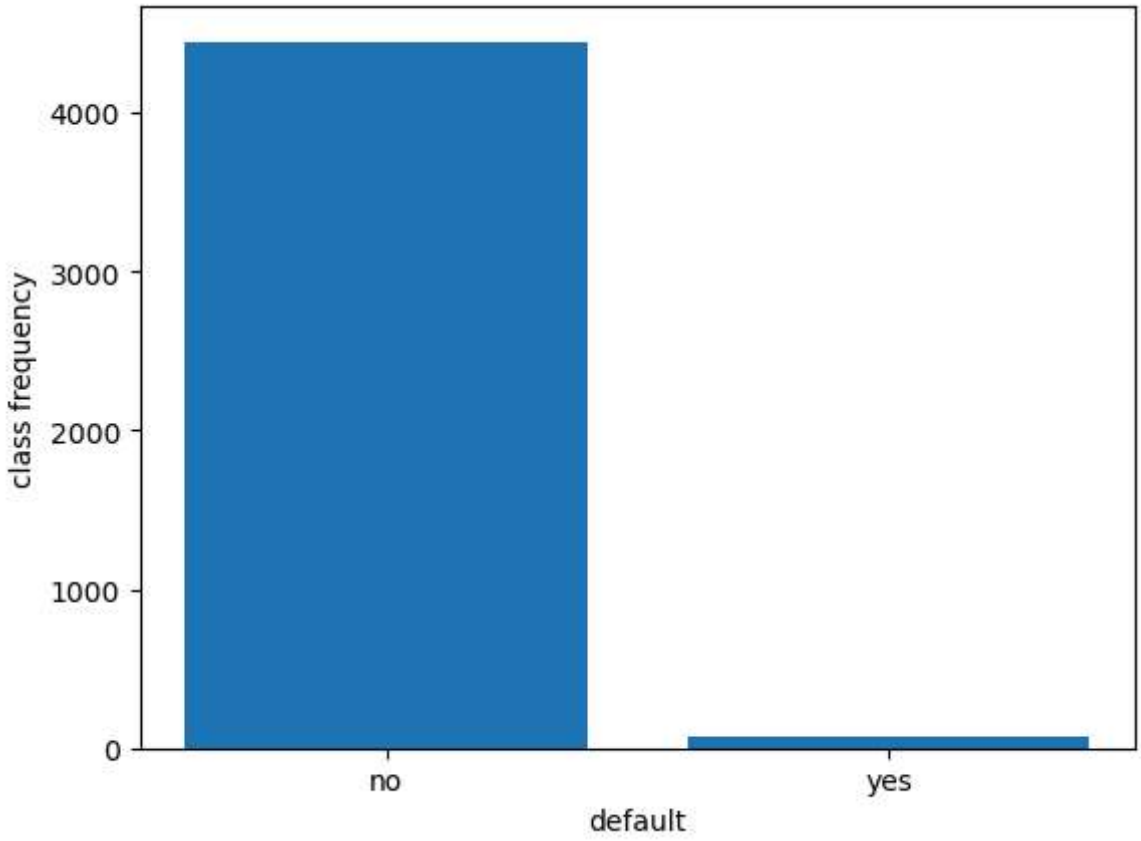
In [55]:
```
data_frame=r'C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiment
```

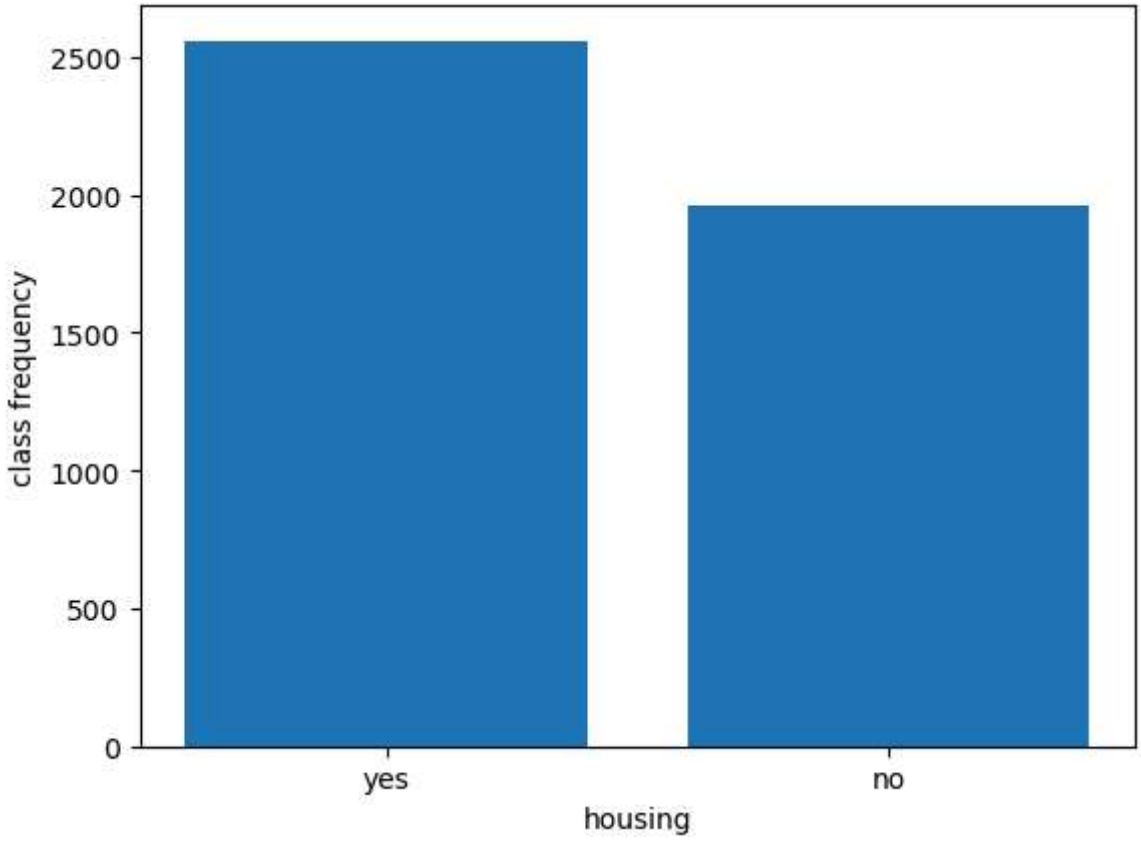
In [63]:
```
#change into dataframe and plot them
for i in cat_colm:
    dfn=bank_df[i].value_counts()
    keys=dfn.keys()
    value=dfn.values
    col=[i,"count"]
    df=pd.DataFrame(zip(keys,value),columns=col)
    df.to_csv(f'{data_frame}\\{i}.csv',index=False)

    plt.bar(i,"count",data=df)
    plt.xlabel(f"{i}")
    plt.ylabel("class frequency")
    plt.savefig(f"{data_frame}\\{i}_bank_df.jpg")
    plt.show()
```

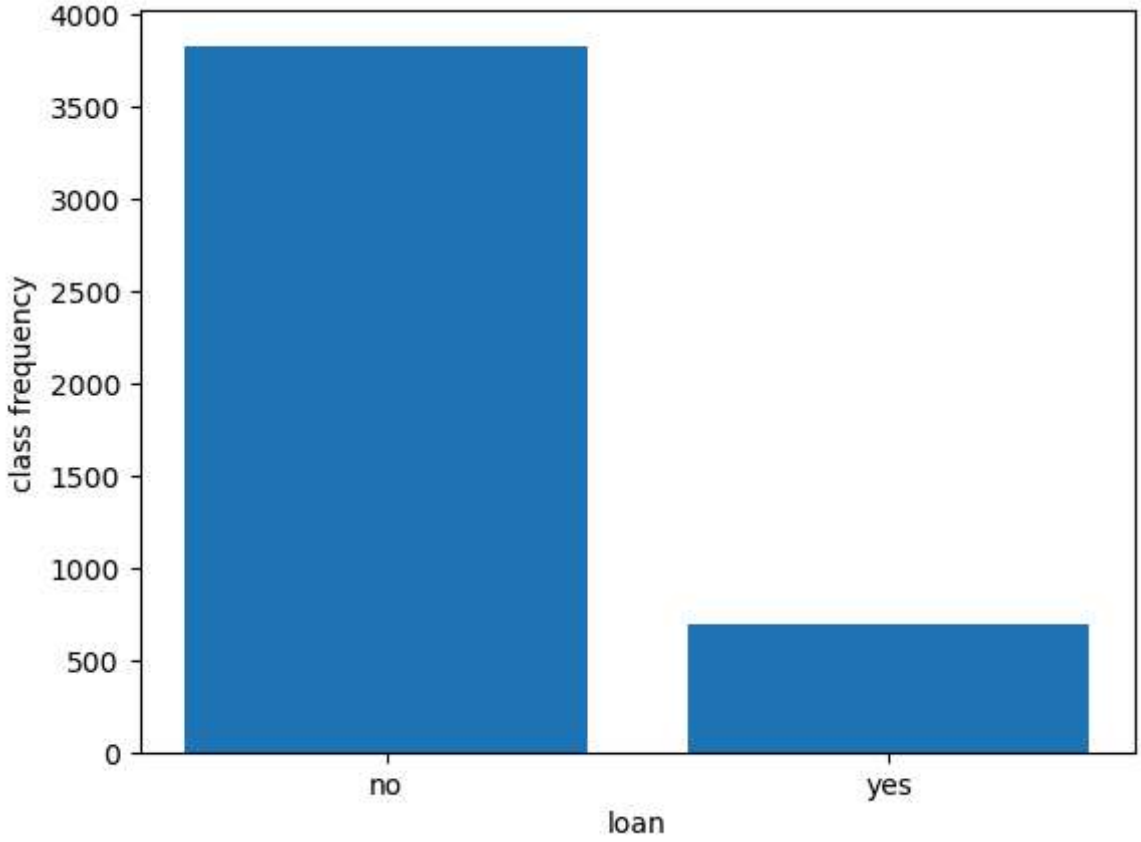


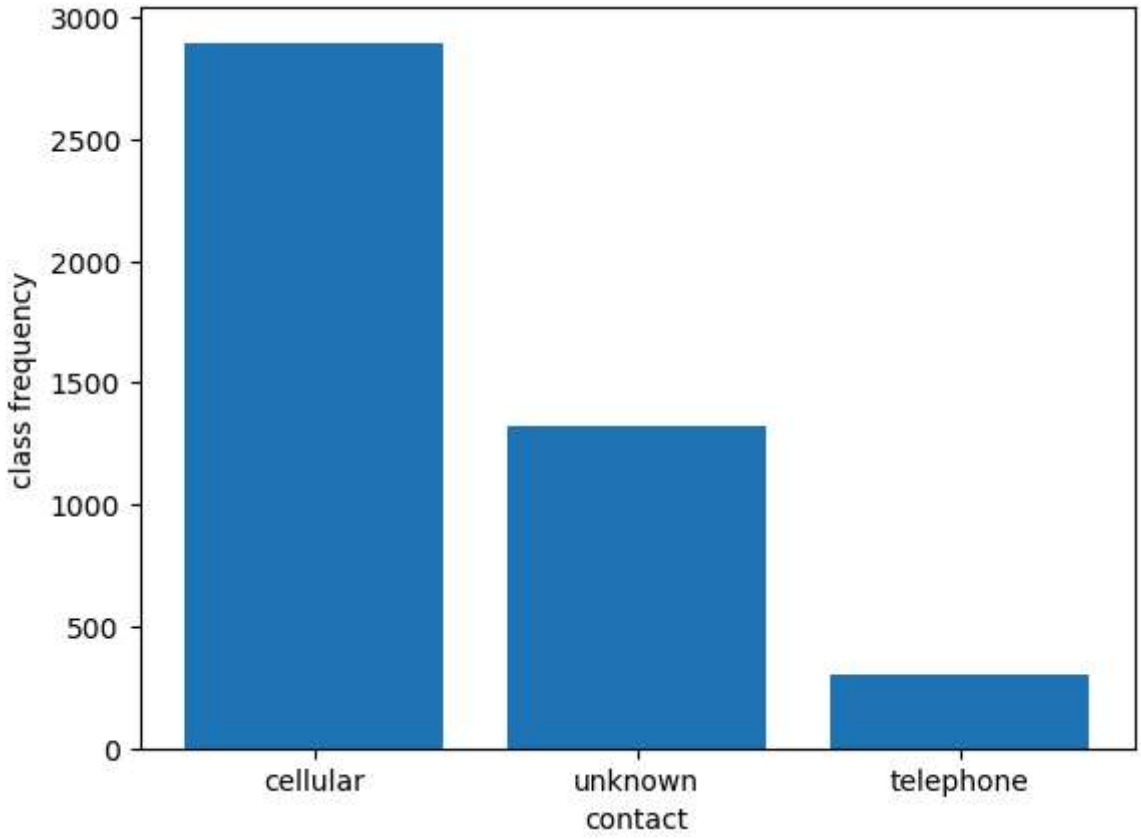

2500
2000
1500
1000
500
0
class frequency
married
single
divorced
marital

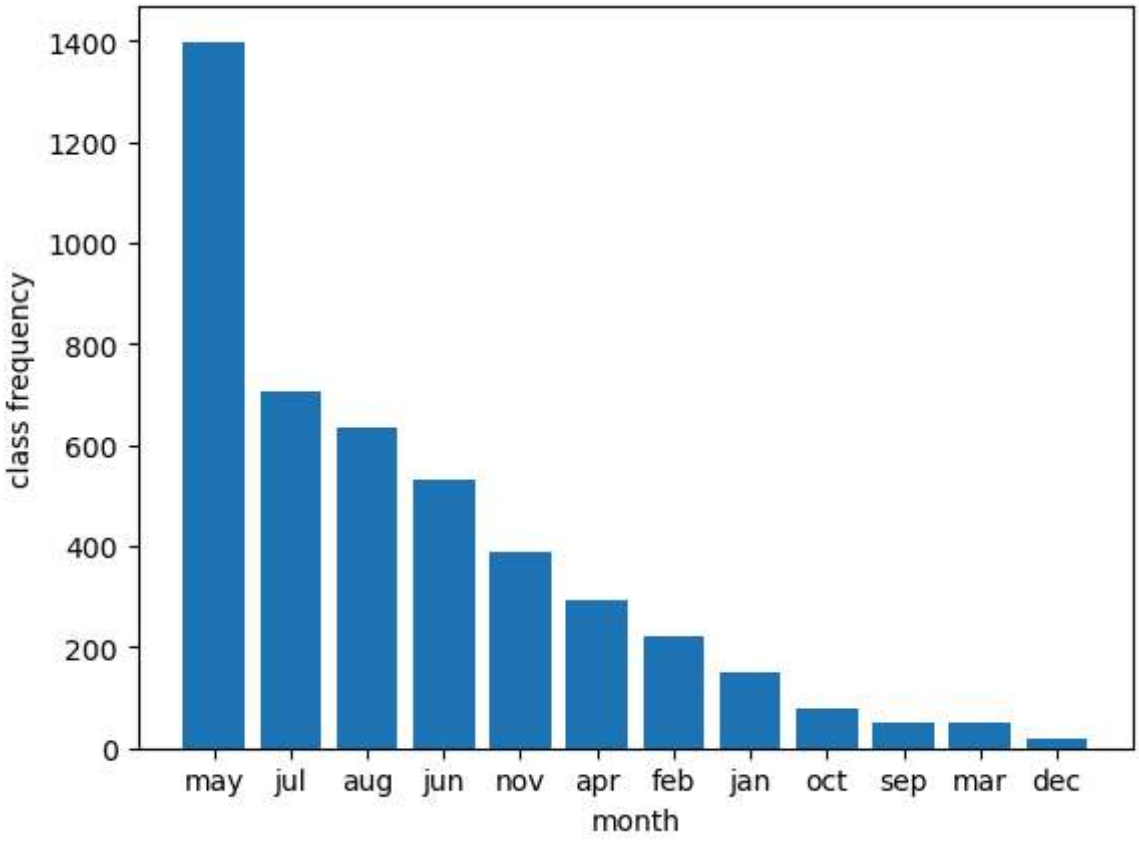

2000
1500
1000
500
0
class frequency
secondary
tertiary
primary
unknown
education

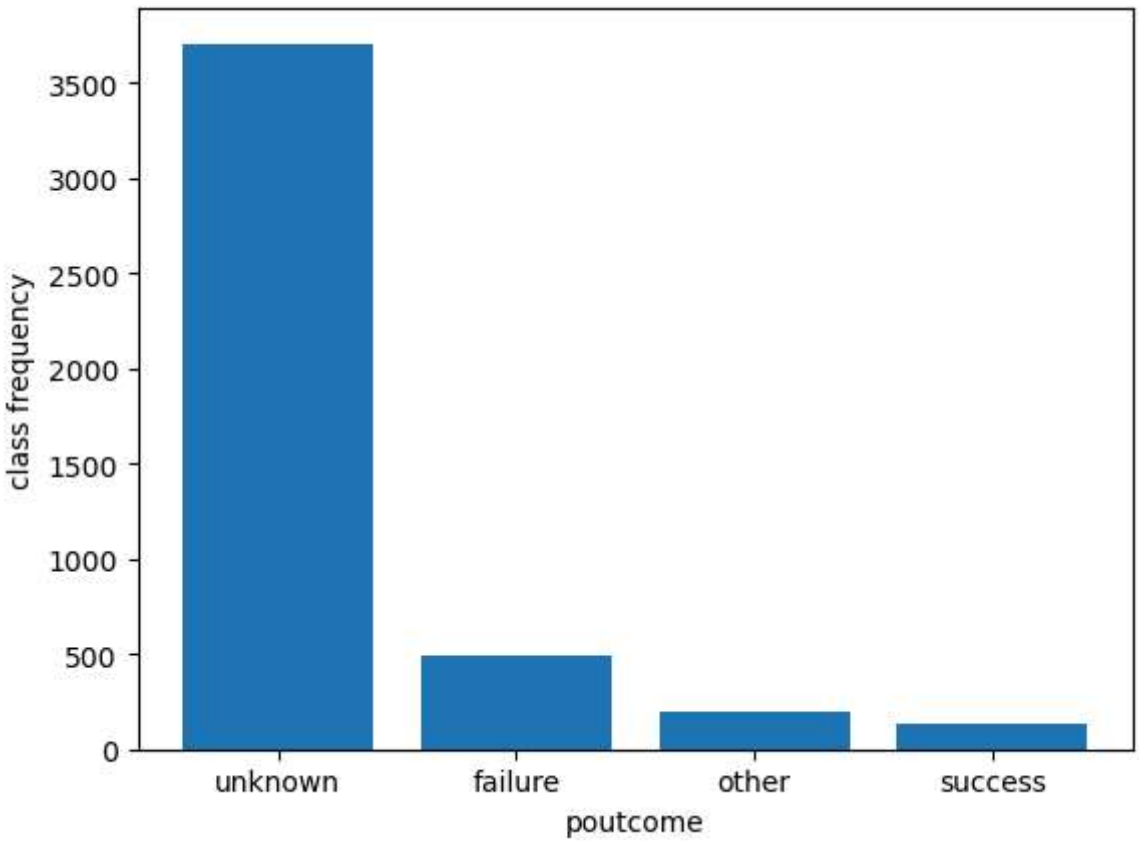

4000
3000
2000
1000
0
class frequency
no
yes
default

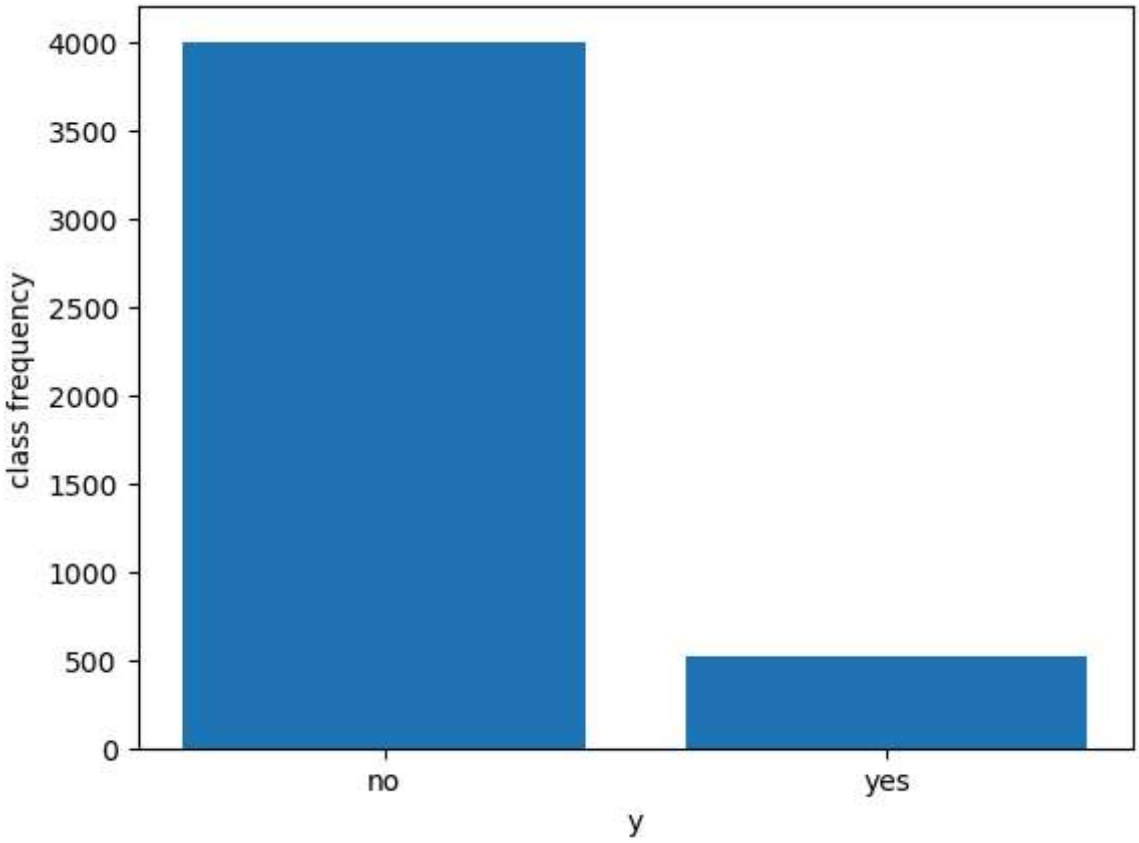

2500
2000
1500
1000
500
0
class frequency
yes
no
housing


4000
3500
3000
2500
2000
1500
1000
500
0
class frequency
no
yes
loan


3000
2500
2000
1500
1000
500
0
class frequency
cellular
unknown
telephone
contact


1400
1200
1000
800
600
400
200
0
class frequency
may
jul
aug
jun
nov
apr
feb
jan
oct
sep
mar
dec
month


3500
3000
2500
2000
1500
1000
500
0
class frequency
unknown
failure
other
success
poutcome



In [67]:
```
root_diractory=os.getcwd()
new_folder="matplot_graphs"
matplot_dir=os.path.join(root_diractory,new_folder)
os.makedirs(matplot_dir)
```

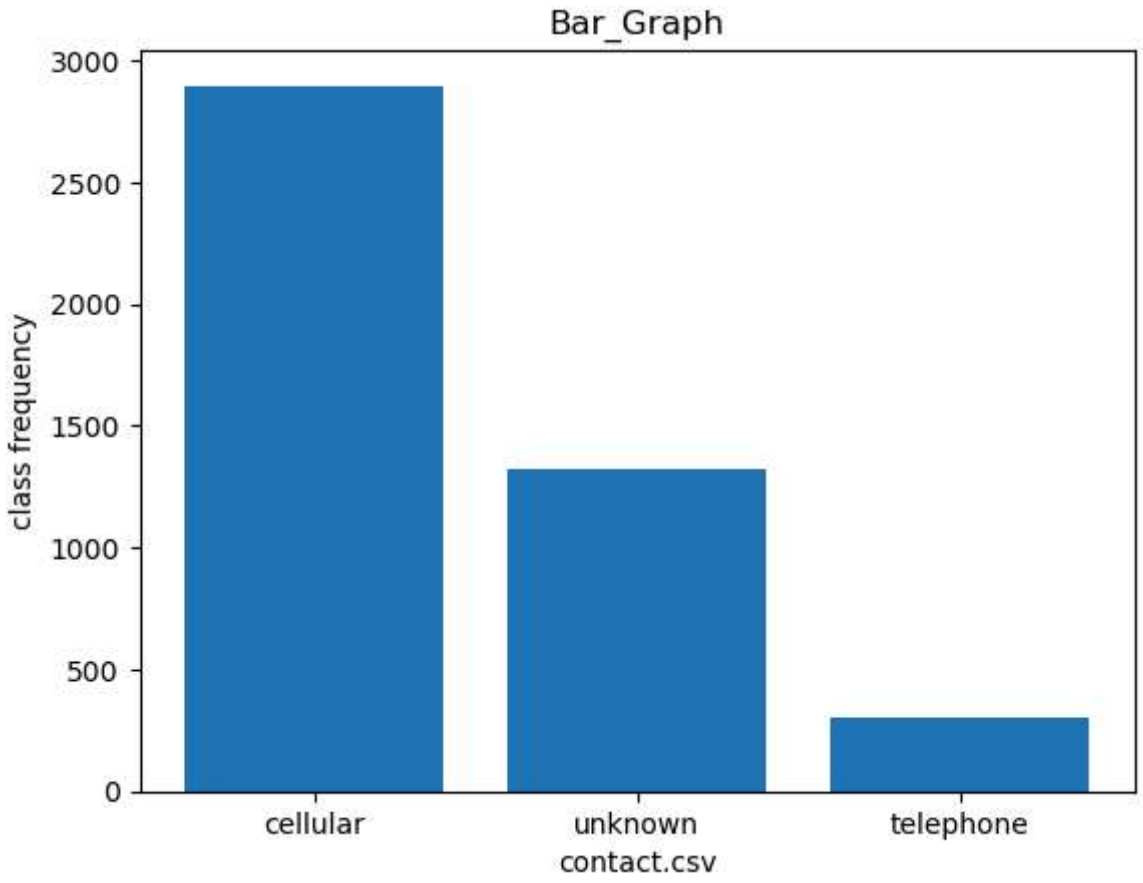
In [75]:
```
for i in cat_colm:
    dfn=bank_df[i].value_counts()
    keys=dfn.keys()
    value=dfn.values
    col=[i,"count"]
    df=pd.DataFrame(zip(keys,value),columns=col)
    df.to_csv(f'{matplot_dir}\\{i}.csv',index=False)
```

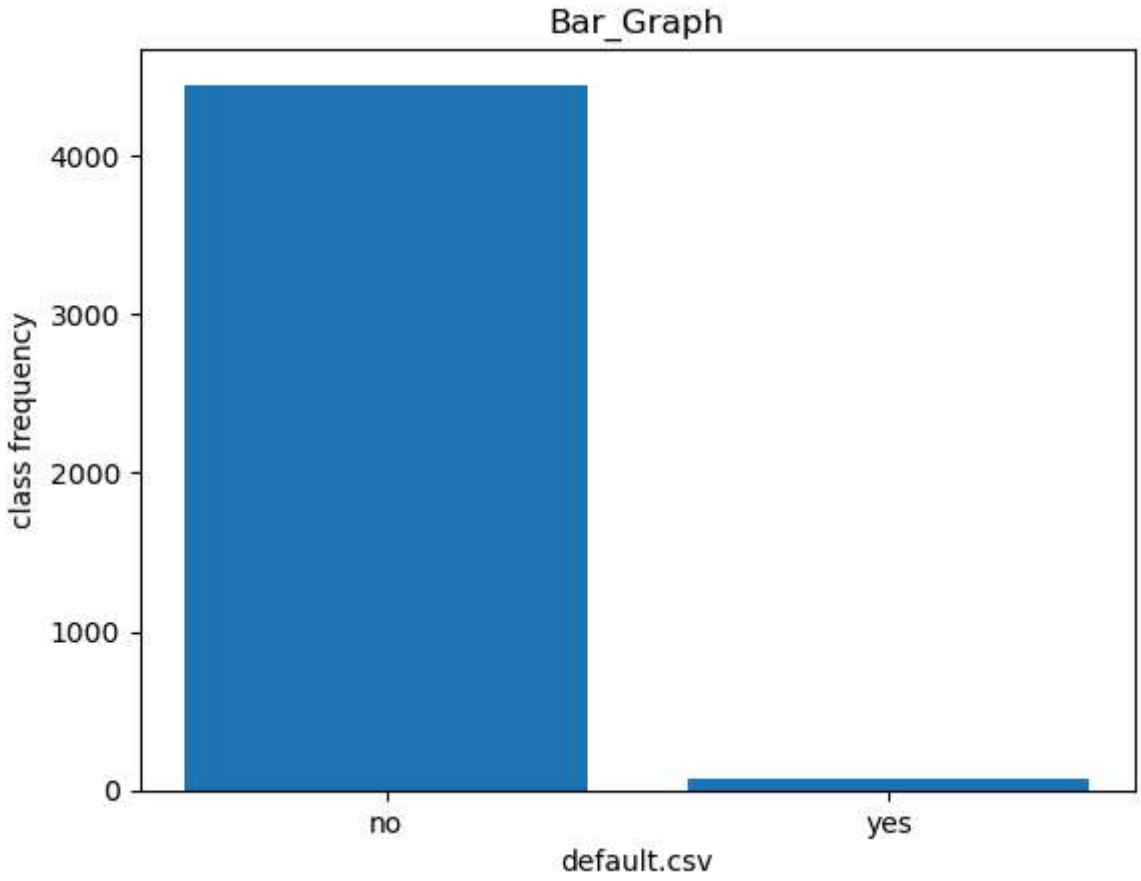
In [79]:
```
matplot_df=os.listdir(matplot_dir)
matplot_df
```

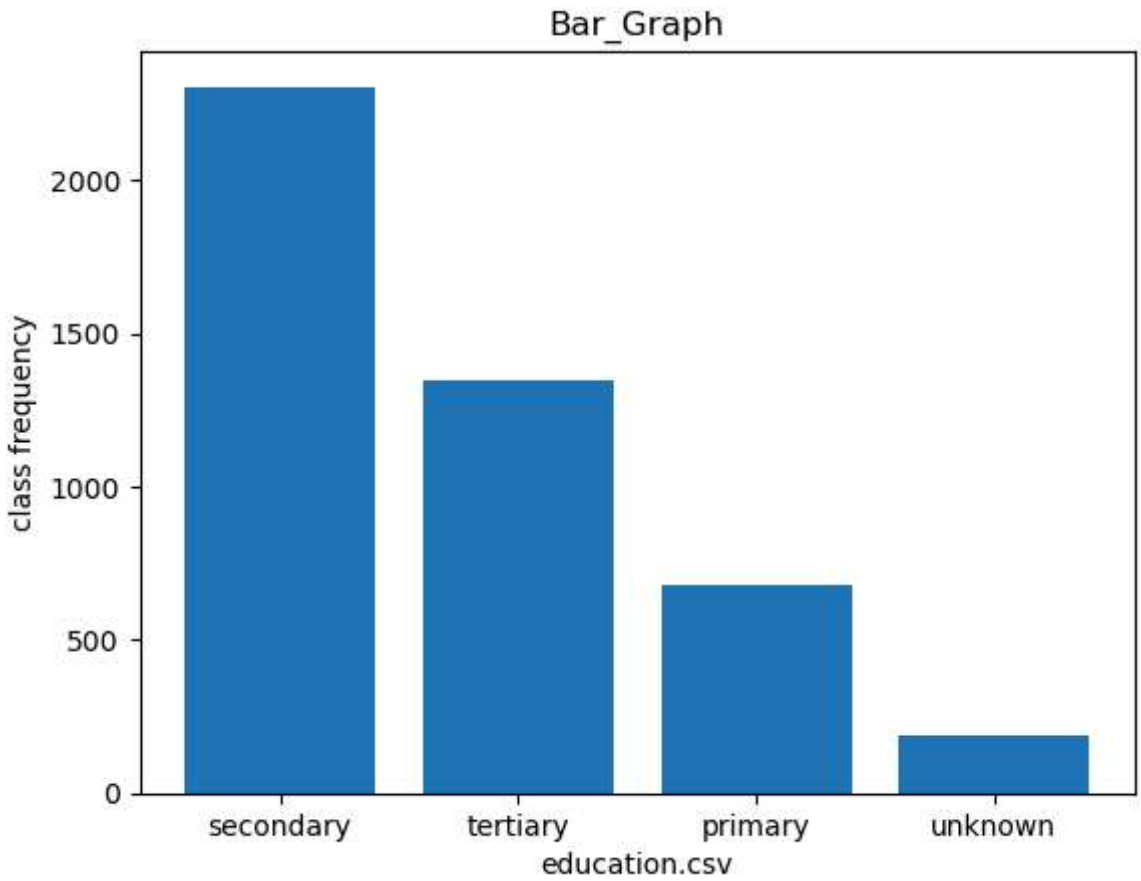
Out[79]:
```
['contact.csv',
 'default.csv',
 'education.csv',
 'housing.csv',
 'job.csv',
 'loan.csv',
 'marital.csv',
 'month.csv',
 'poutcome.csv',
 'y.csv']
```

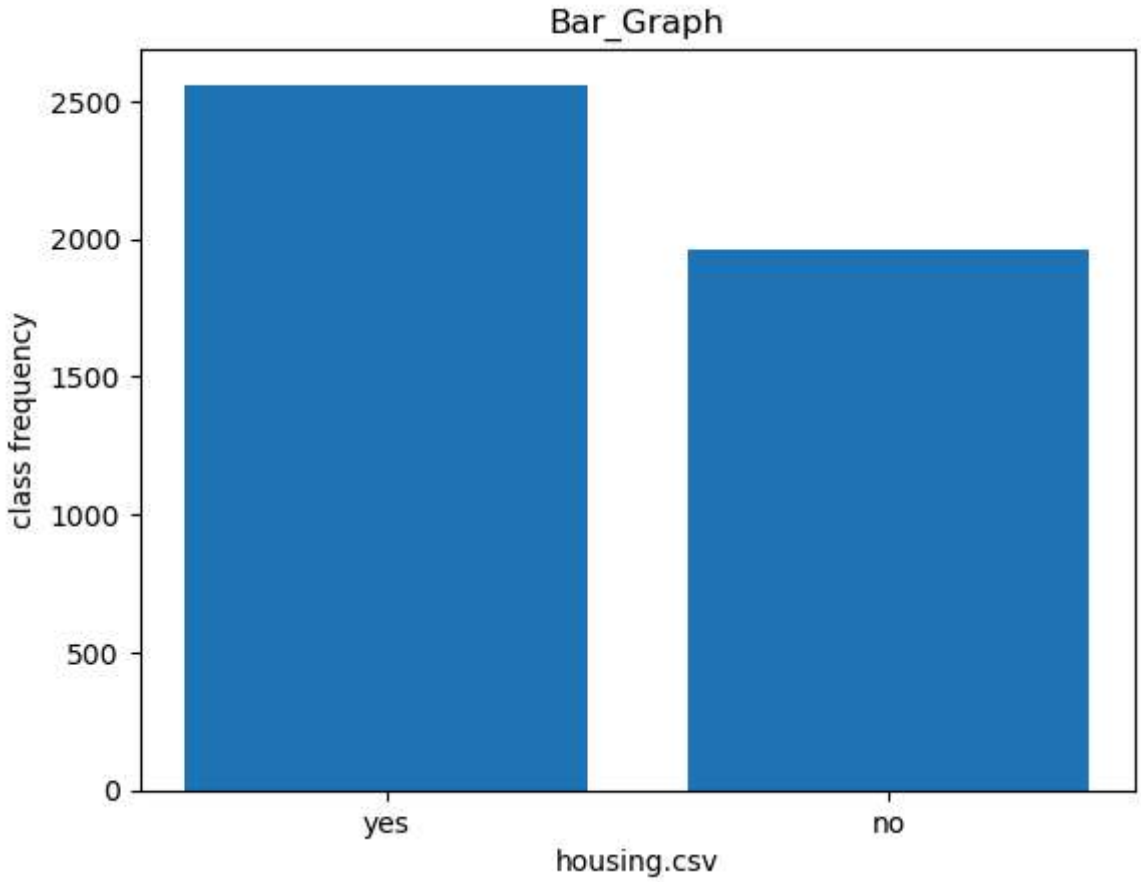
In [87]:
```
#pd.read_csv('contact.csv')
```

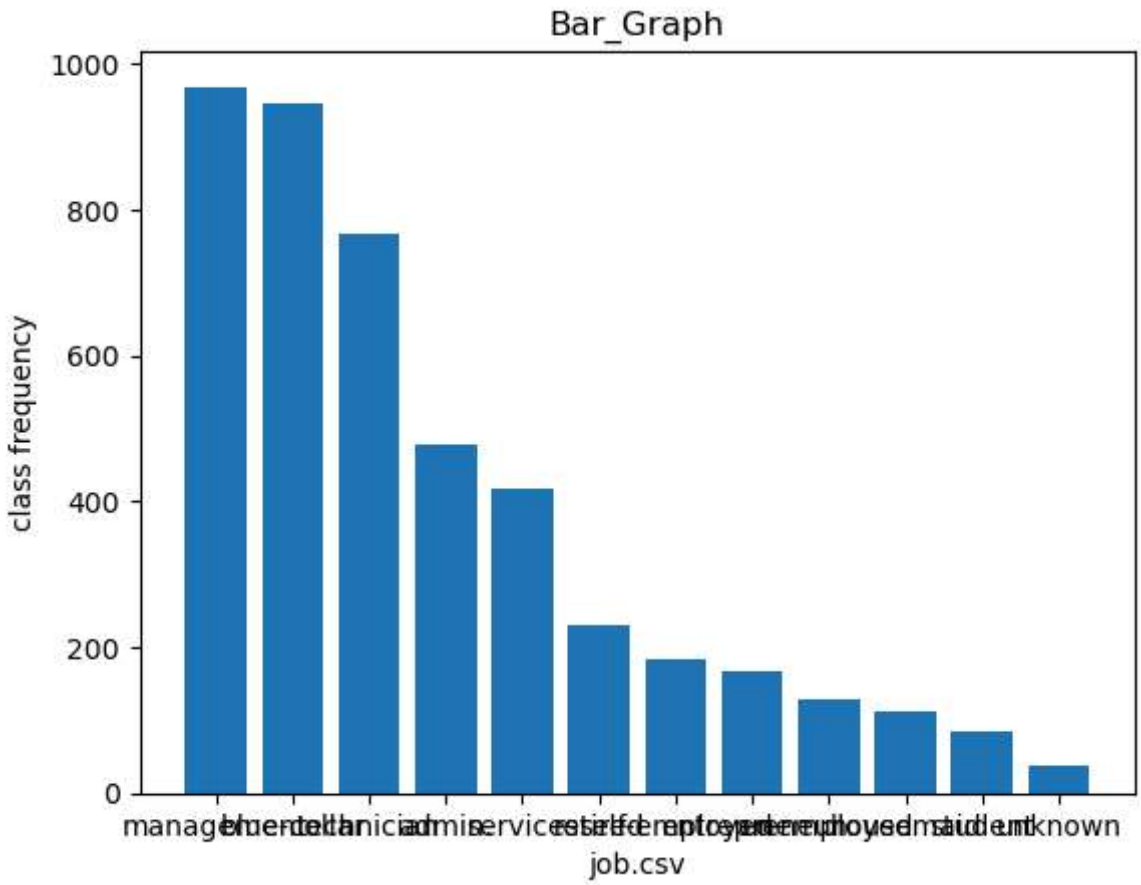
In [93]:
```
for i in matplot_df:
    dfs=pd.read_csv(f"{matplot_dir}\\{i}")
    plt.title("Bar_Graph")
    plt.bar(i[:-4],"count",data=dfs)
    plt.xlabel(f"{i}")
    plt.ylabel("class frequency")
    plt.savefig(f"{matplot_dir}\\{i}_matplot.jpg")
    plt.show()
```

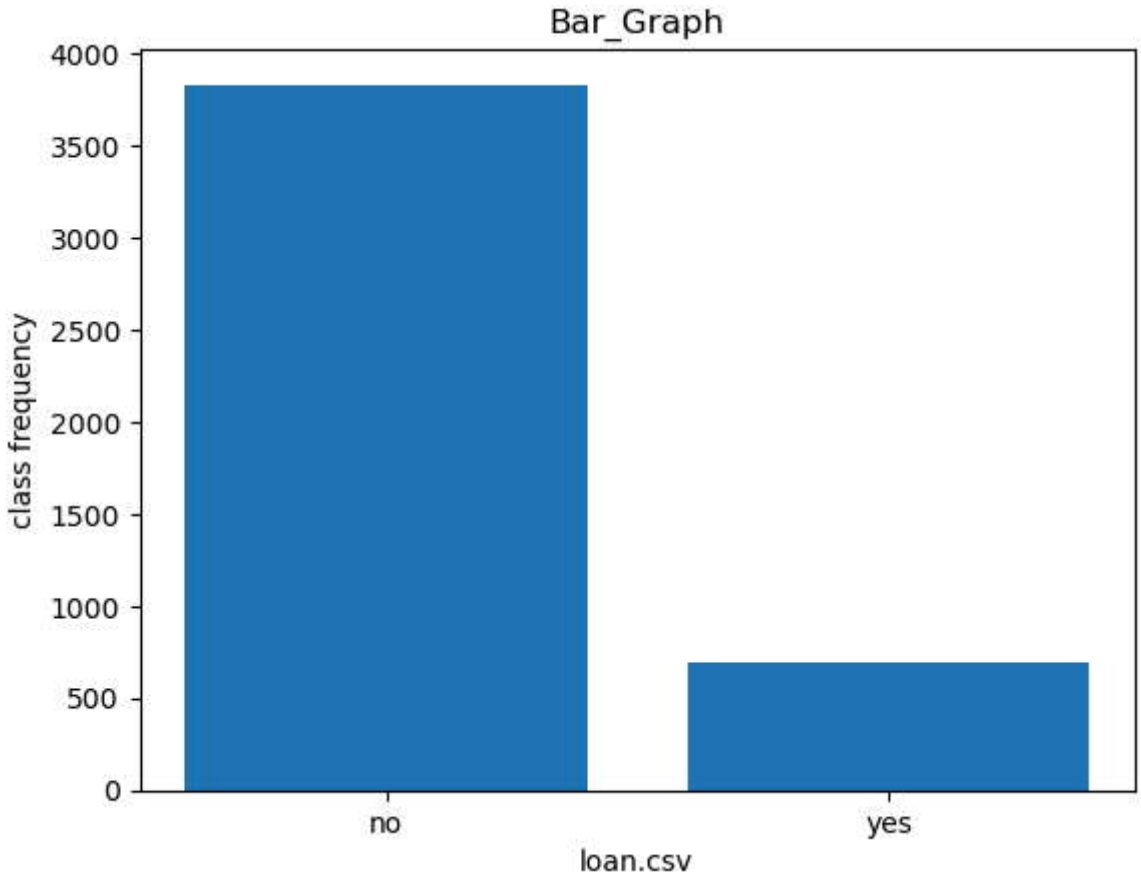


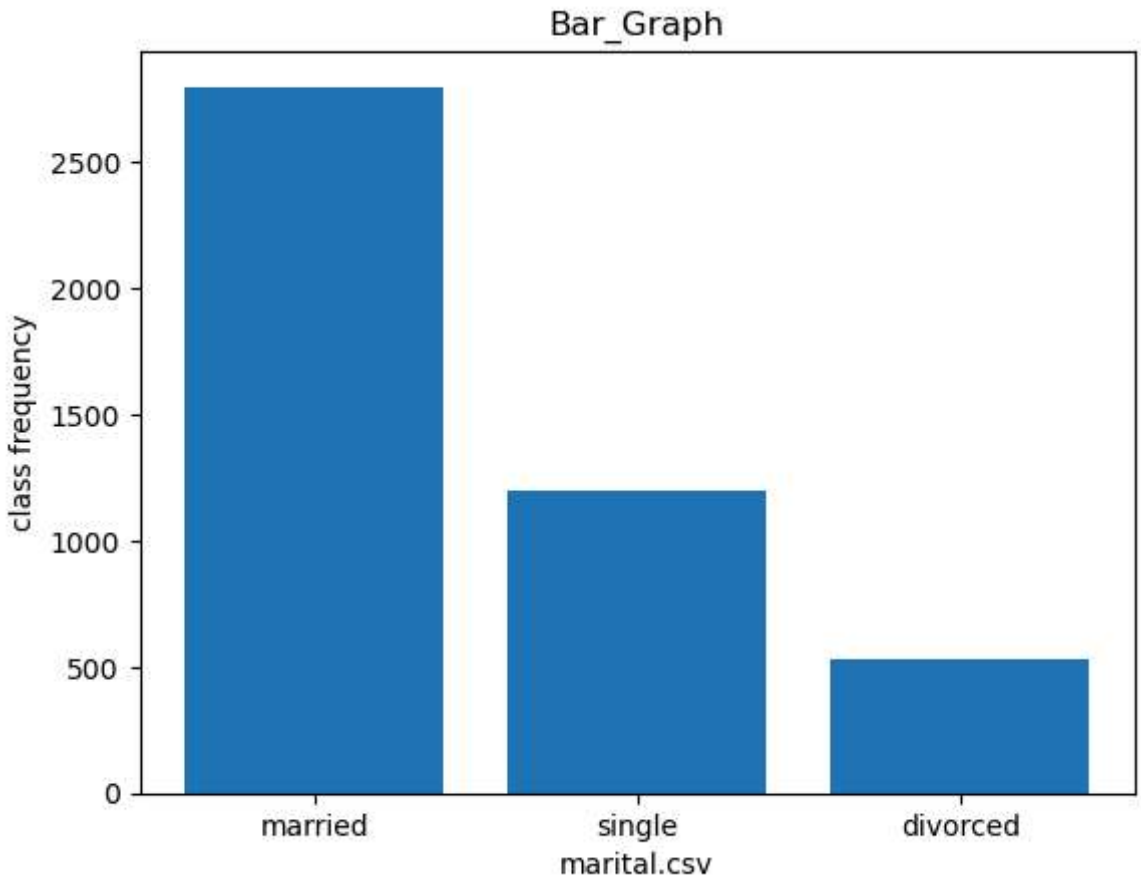

Bar_Graph
class frequency
4000
3000
2000
1000
0
no
yes
default.csv

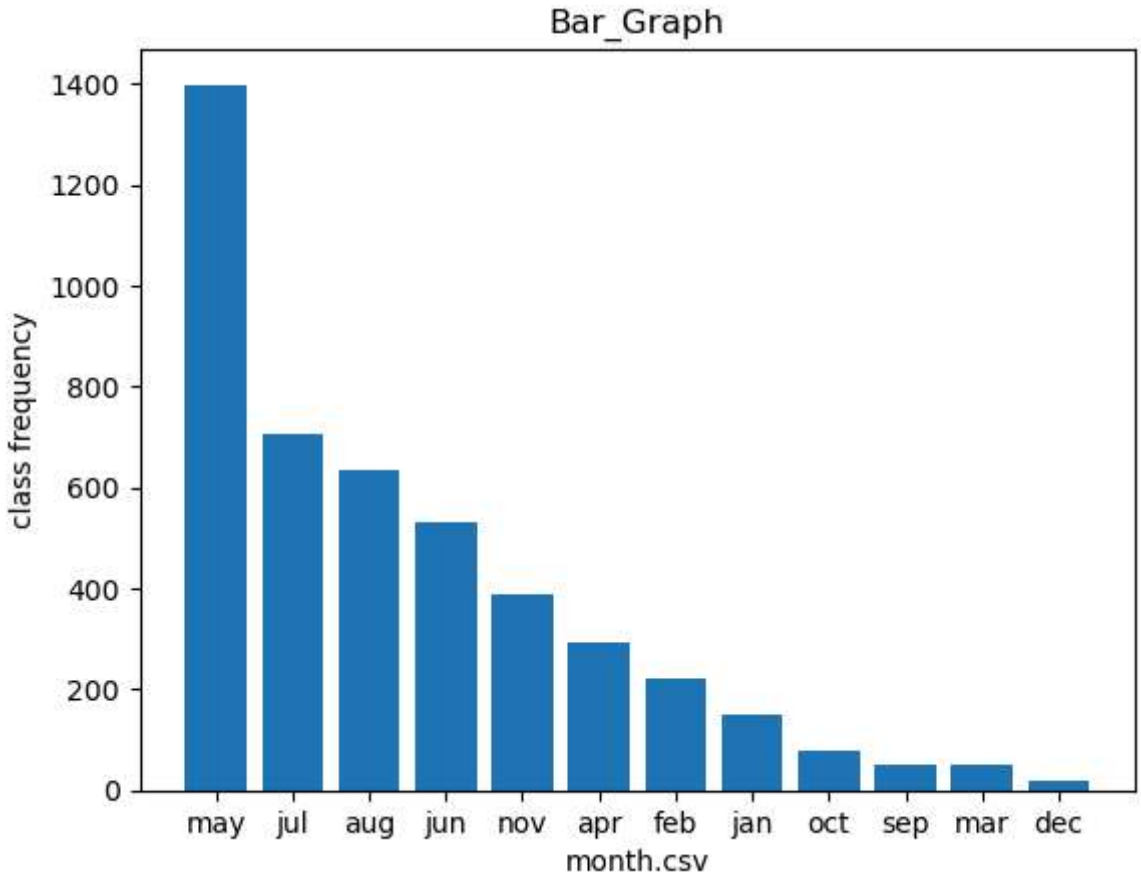

Bar_Graph
class frequency
2000
1500
1000
500
0
secondary
tertiary
primary
unknown
education.csv

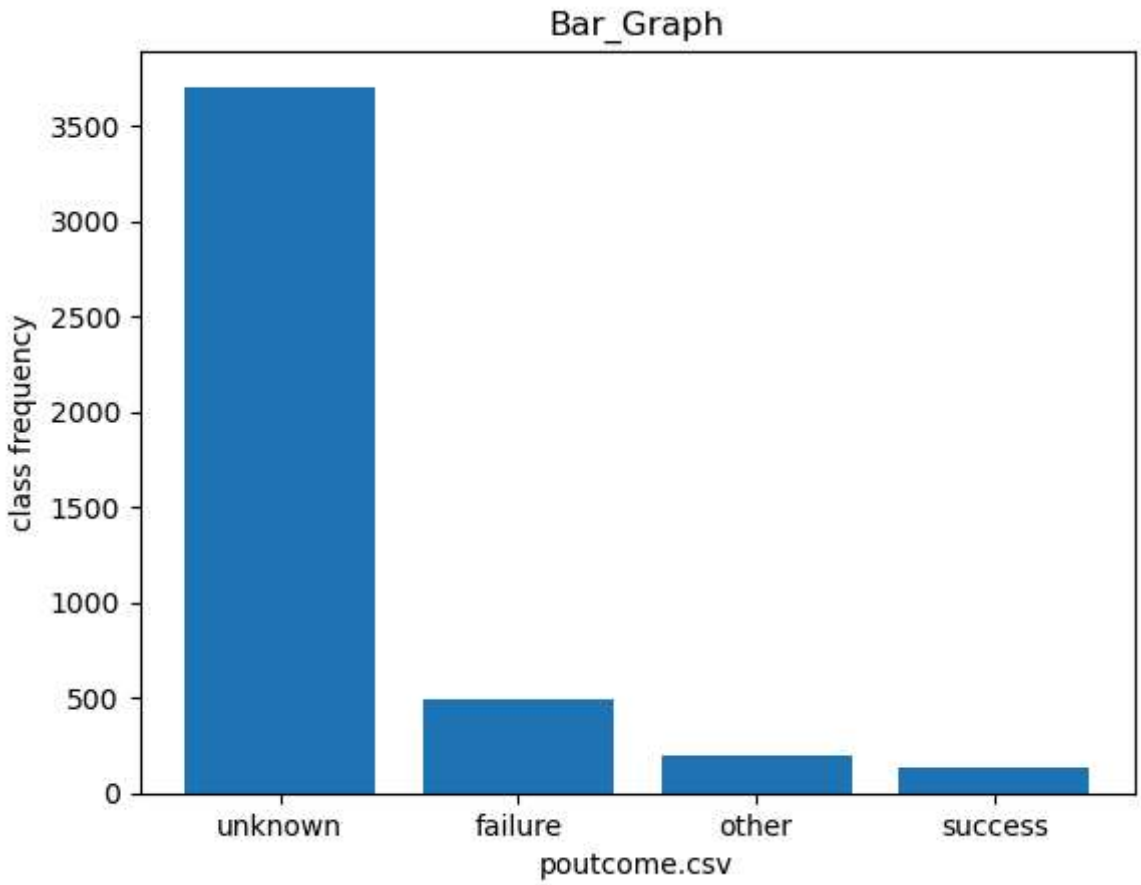

Bar_Graph
class frequency
2500
2000
1500
1000
500
0
yes
no
housing.csv

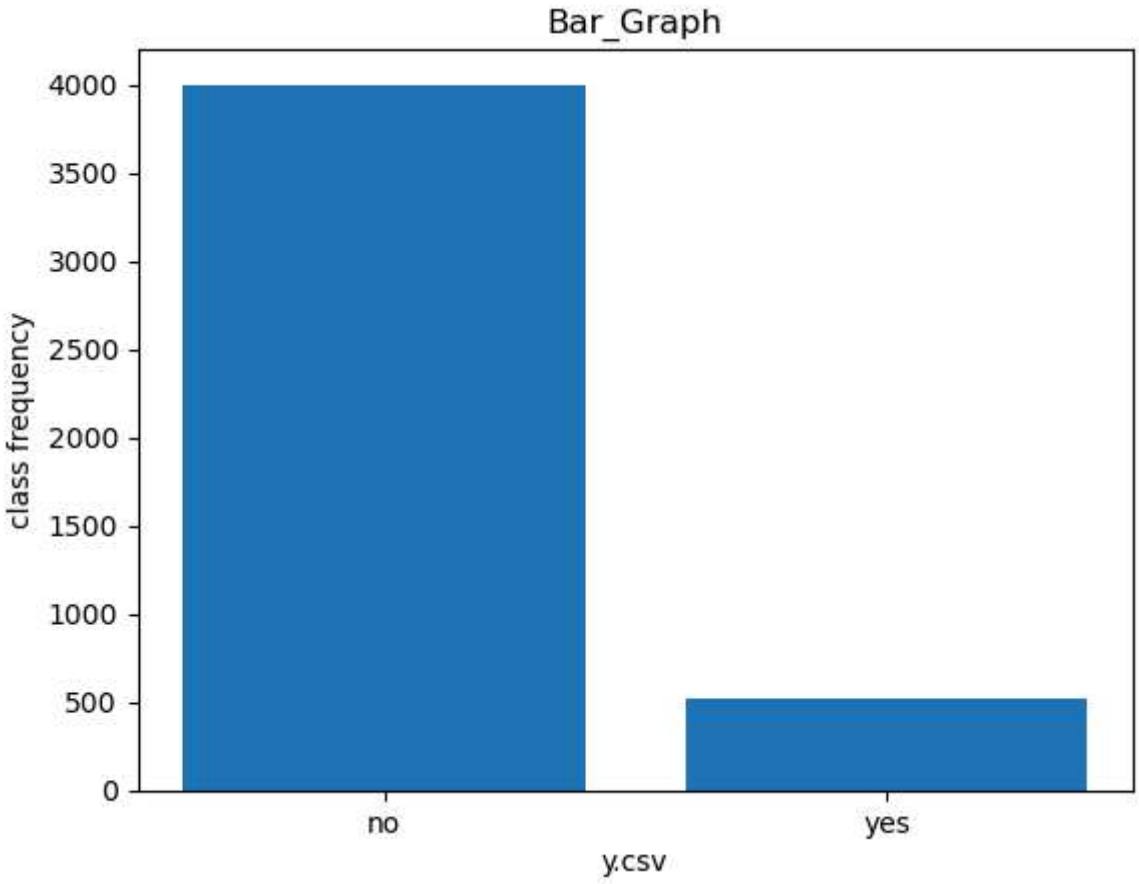

Bar_Graph
class frequency
1000
800
600
400
200
0
job.csv


Bar_Graph
4000
3500
3000
2500
2000
1500
1000
500
0
class frequency
no
yes
loan.csv


Bar_Graph
2500
2000
1500
1000
500
0
class frequency
married
single
divorced
marital.csv


Bar_Graph
class frequency
0
200
400
600
800
1000
1200
1400
may
jul
aug
jun
nov
apr
feb
jan
oct
sep
mar
dec
month.csv


Bar_Graph
class frequency
0
500
1000
1500
2000
2500
3000
3500
unknown
failure
other
success
poutcome.csv



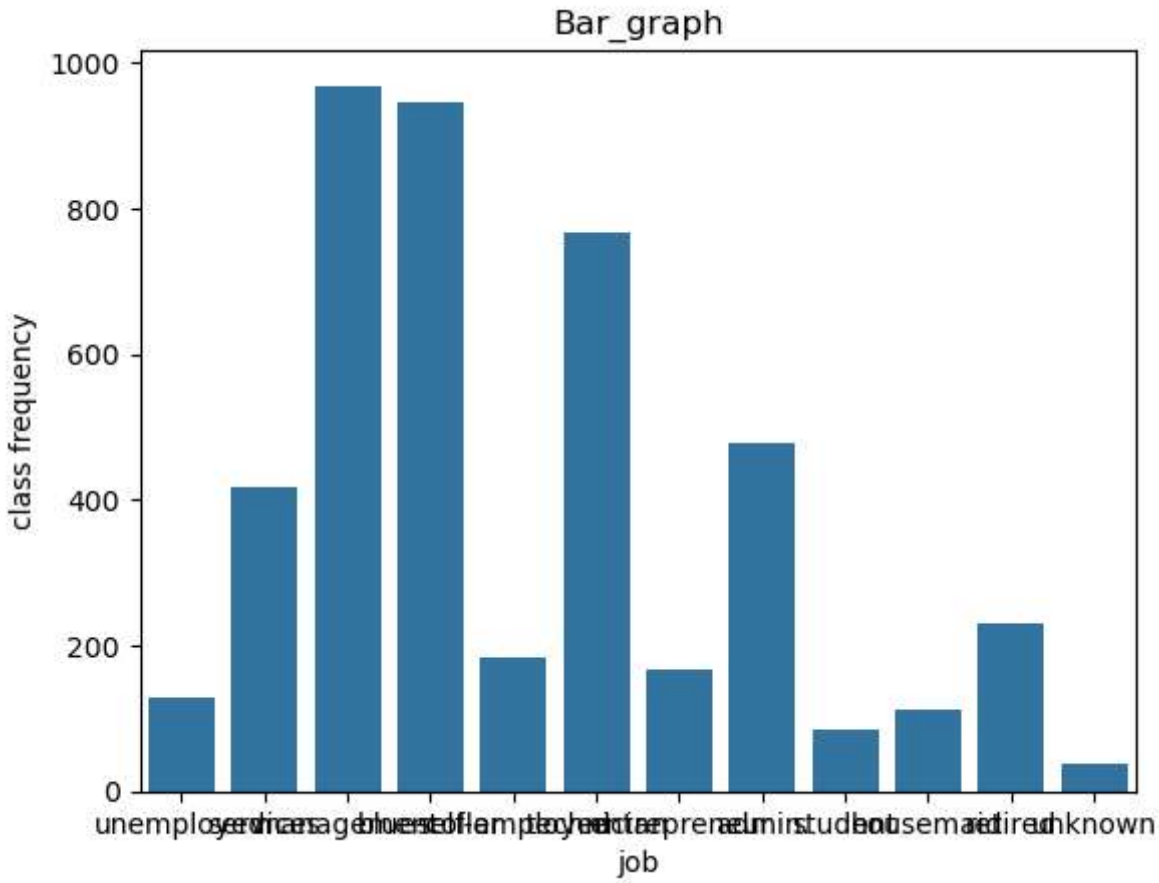
In [97]:
```
#create seaborn directory

root_sea=os.getcwd()
new_folder="seaborn_graphs"
seaborn_dir=os.path.join(root_sea,new_folder)
os.makedirs(seaborn_dir)
```

In [103…
```
#categorical columns
cat_colm
```

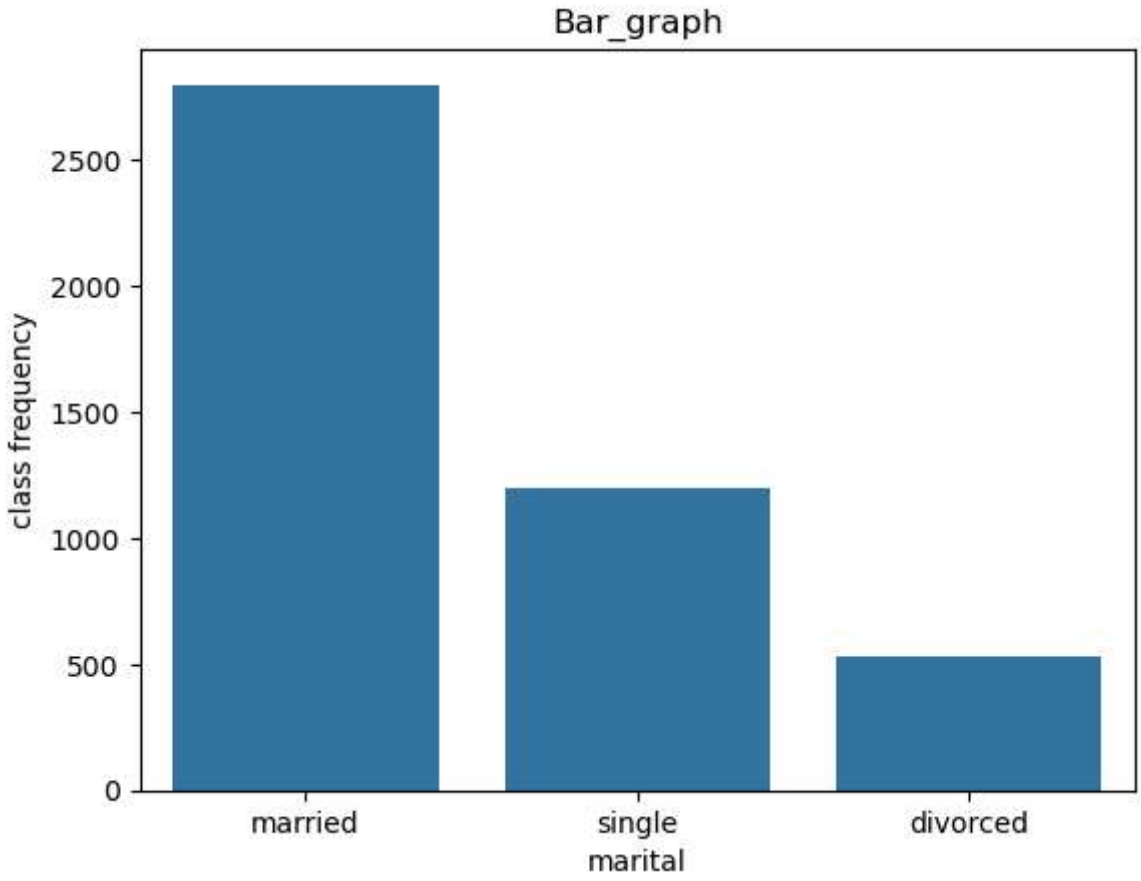
Out[103…
```
['job',
 'marital',
 'education',
 'default',
 'housing',
 'loan',
 'contact',
 'month',
 'poutcome',
 'y']
```

In [109…
```
for i in cat_colm:
    sns.countplot(data=bank_df,
                  x=i)
    plt.title("Bar_graph")
    plt.xlabel(f"{i}")
    plt.ylabel("class frequency")
    plt.figure(figsize=(8,5))
```

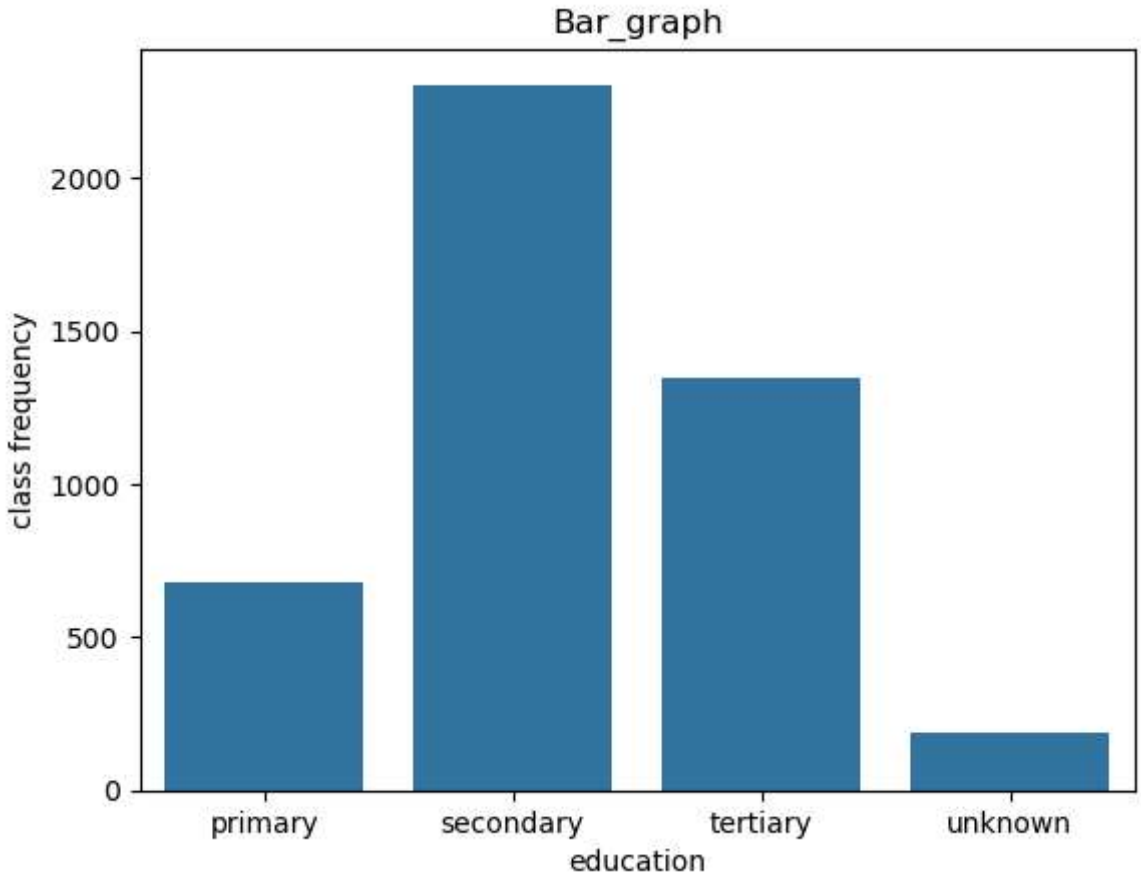
```
    plt.savefig(f"{seaborn_dir}\\{i}.seaborn.jpg")
    plt.show()
```



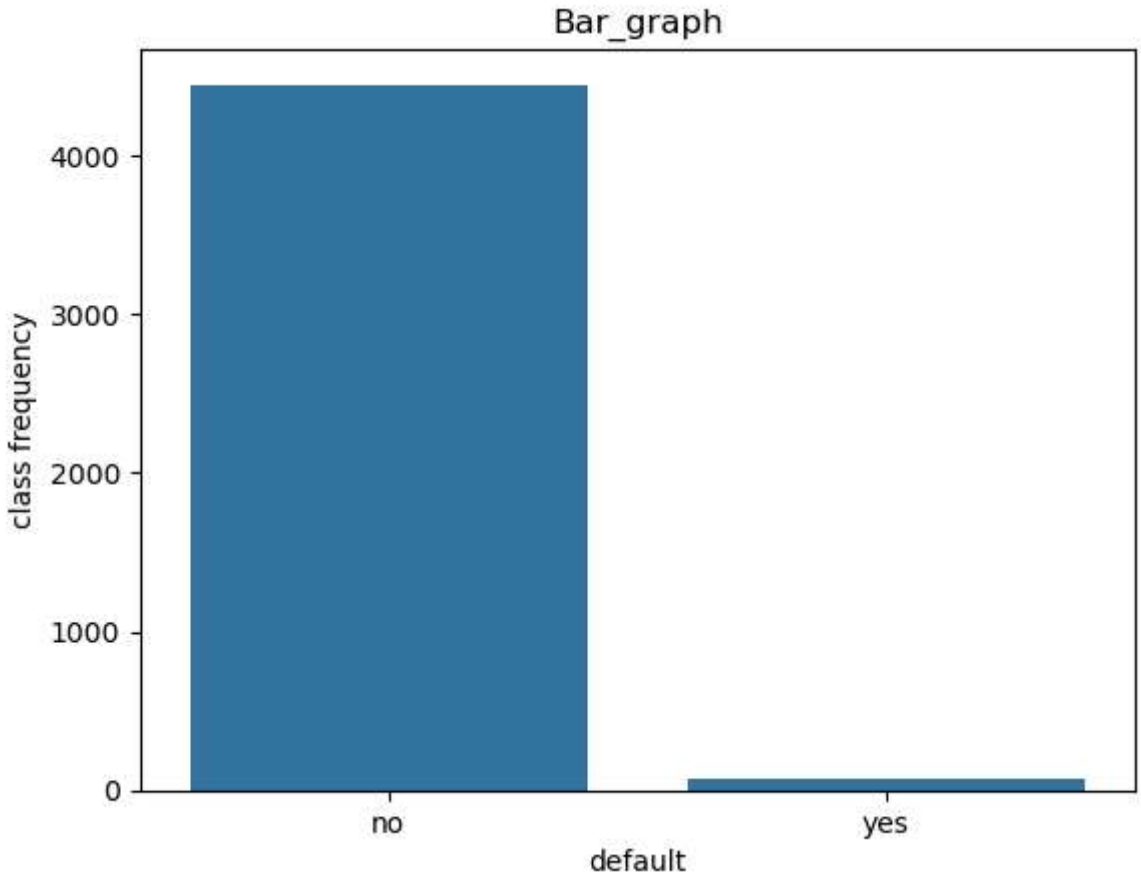

```
<Figure size 800x500 with 0 Axes>
```



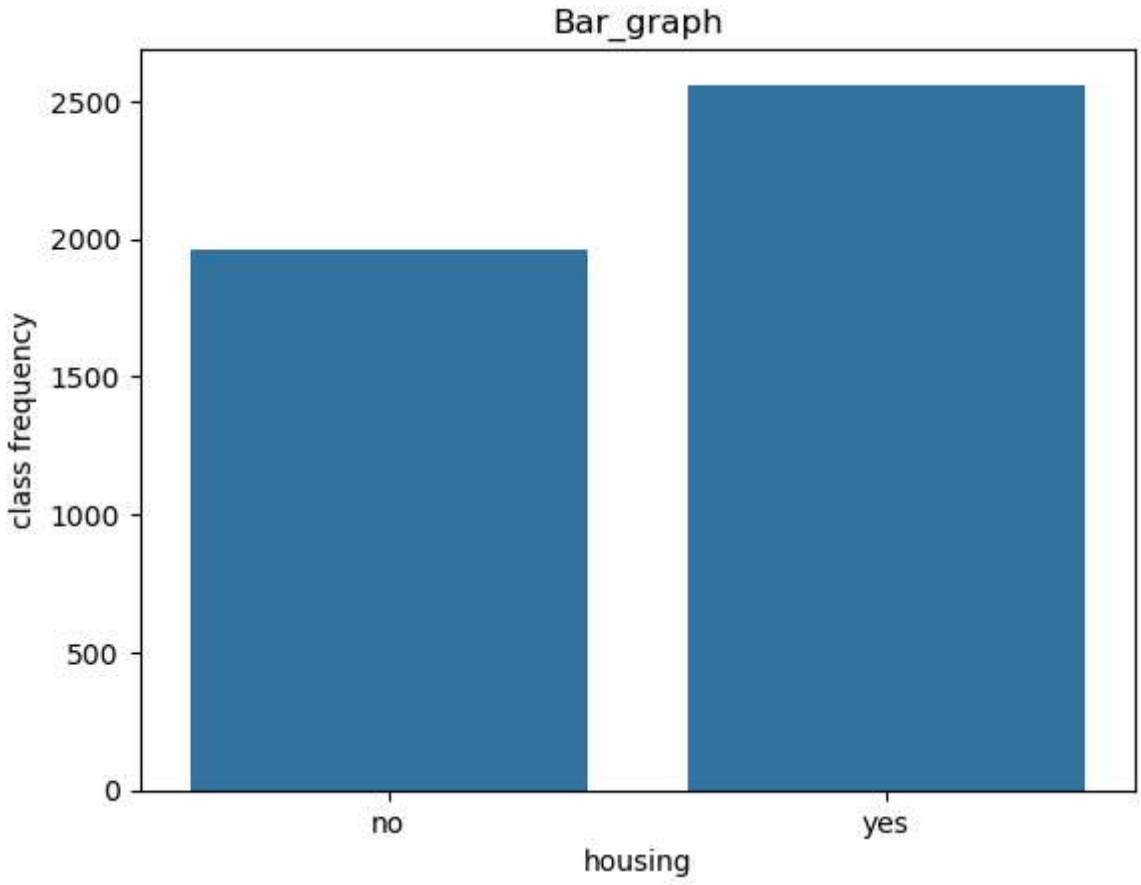

```
<Figure size 800x500 with 0 Axes>
```



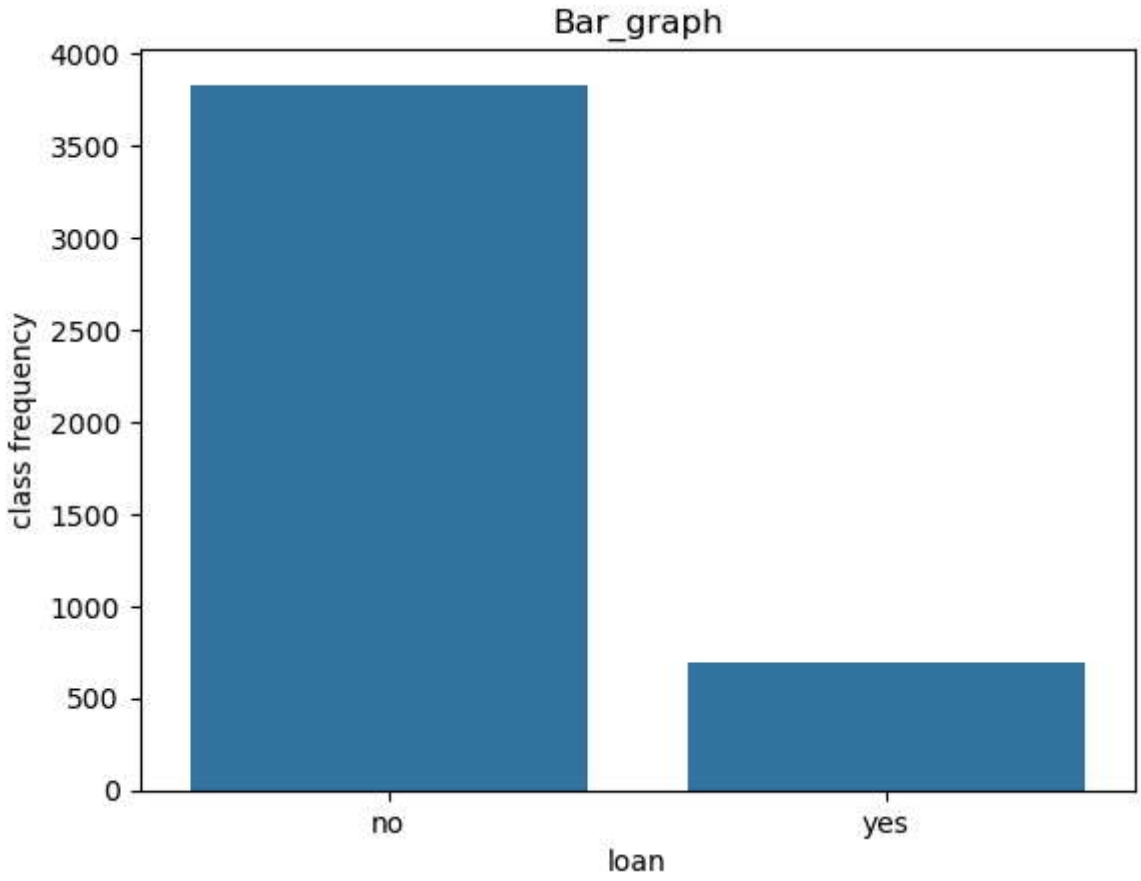

```
<Figure size 800x500 with 0 Axes>
```



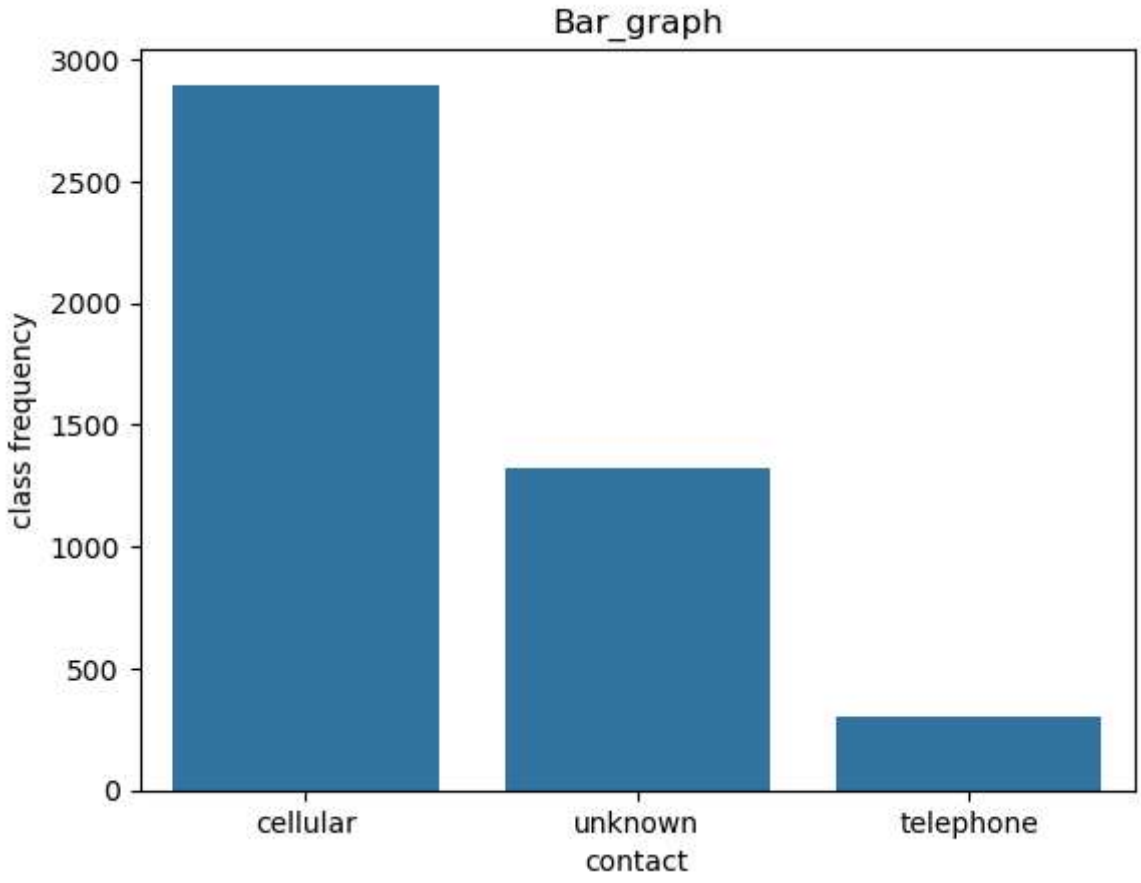

```
<Figure size 800x500 with 0 Axes>
```



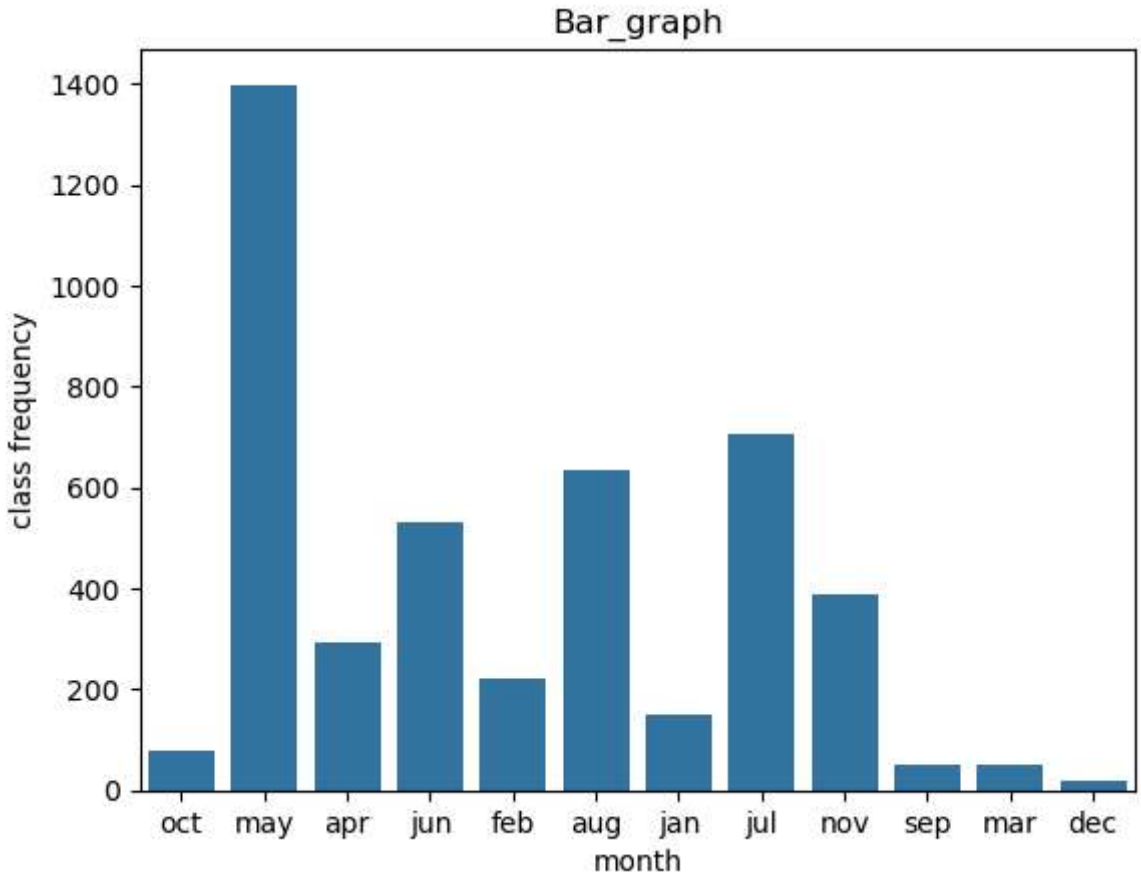

```
<Figure size 800x500 with 0 Axes>
```



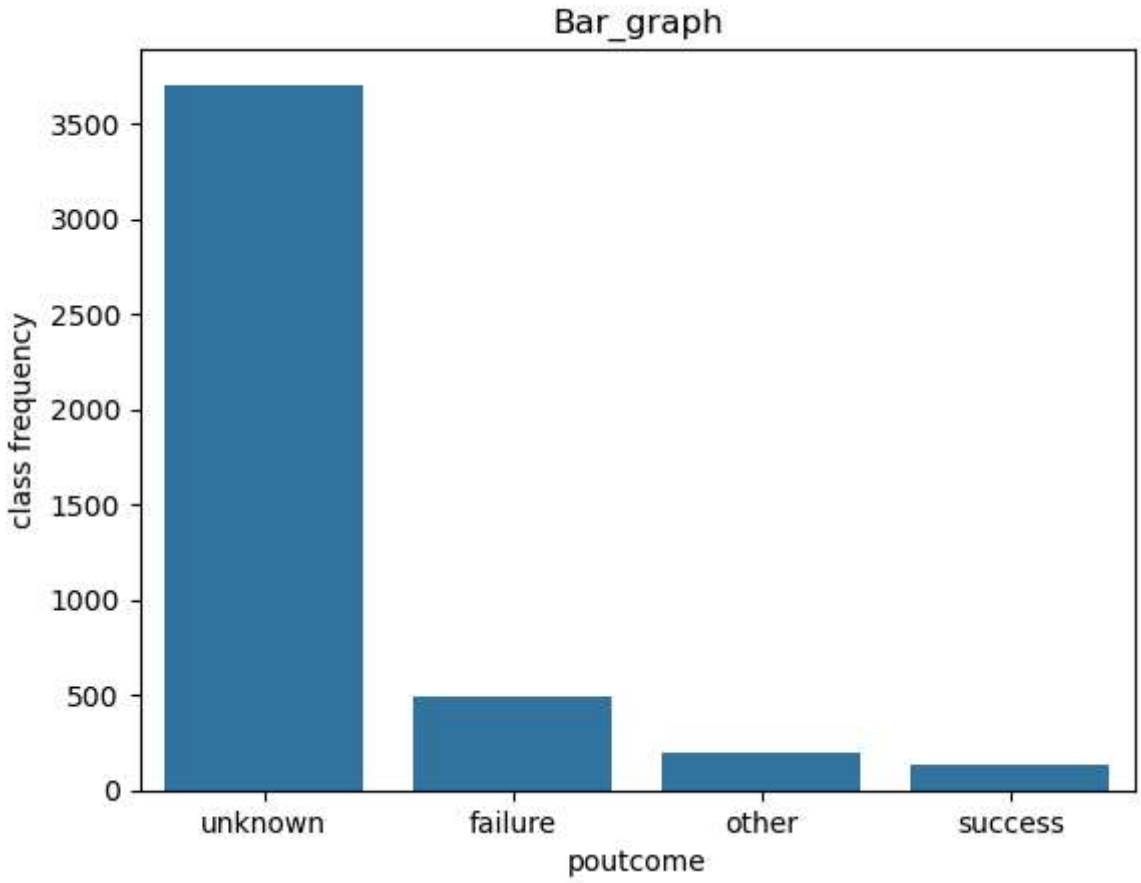

```
<Figure size 800x500 with 0 Axes>
```



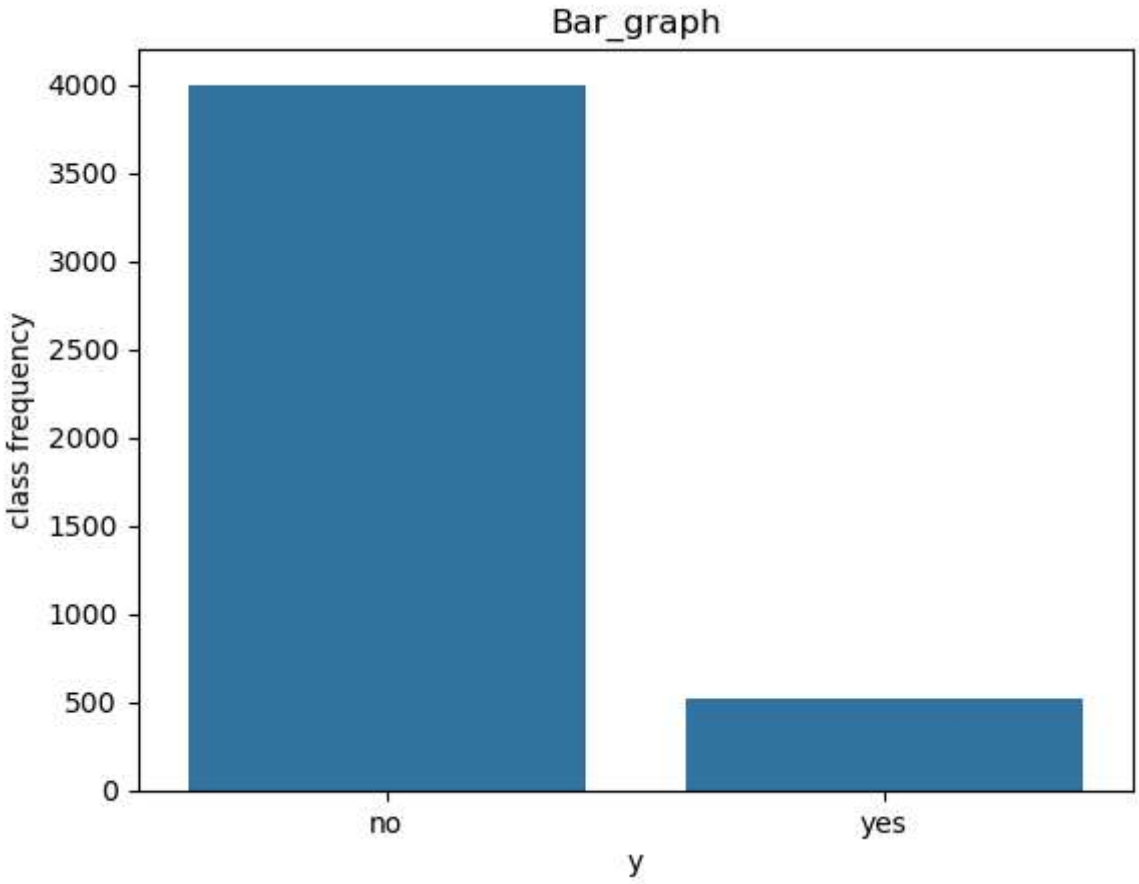

```
<Figure size 800x500 with 0 Axes>
```



<Figure size 800x500 with 0 Axes>



<Figure size 800x500 with 0 Axes>




```
<Figure size 800x500 with 0 Axes>
```

In [111…
```
#all cloumns of bankdatasets
colm
```

Out[111…
```
Index(['age', 'job', 'marital', 'education', 'default', 'balance', 'housing',
       'loan', 'contact', 'day', 'month', 'duration', 'campaign', 'pdays',
       'previous', 'poutcome', 'y'],
      dtype='object')
```

In [113…
```
#again try
root_dir=os.getcwd()
new_folder="data_frames"
new_dir=os.path.join(root_dir,new_folder)
os.makedirs(new_dir)
```

In [119…
```
for i in colm:
    cnf=bank_df[i].value_counts()
    keys=cnf.keys()
    value=cnf.values
    df=pd.DataFrame(zip(keys,value),columns=[i,"count"])
    df.to_csv(f"{new_dir}\\{i}.csv")
```

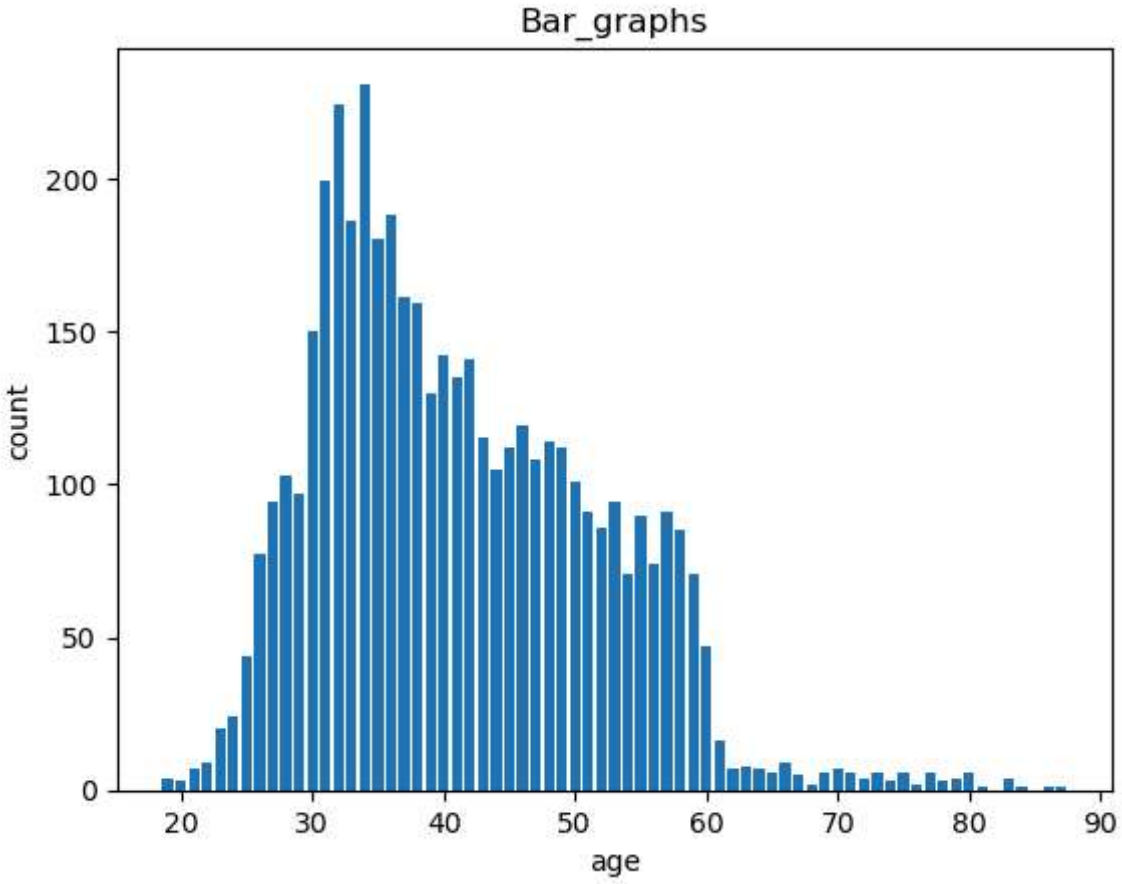
In [125…
```
os.getcwd()
```

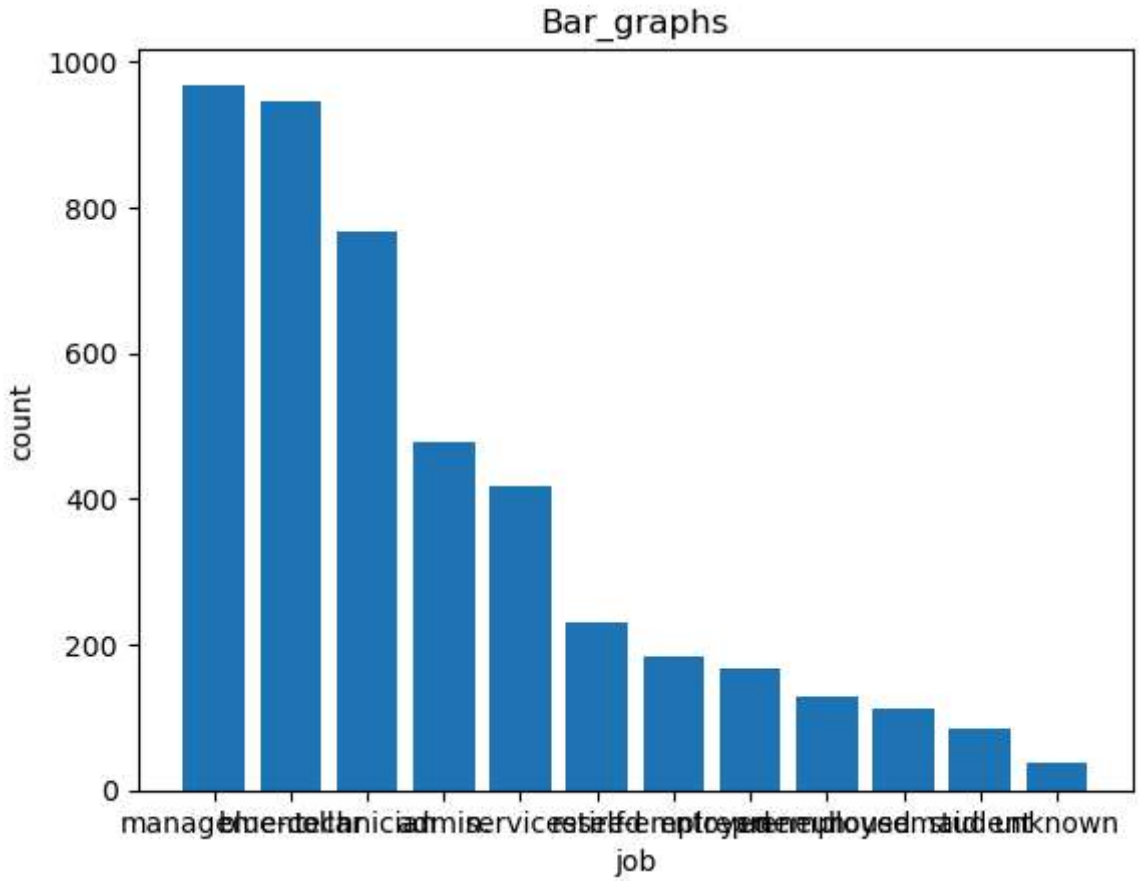
Out[125…
```
'C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiments'
```

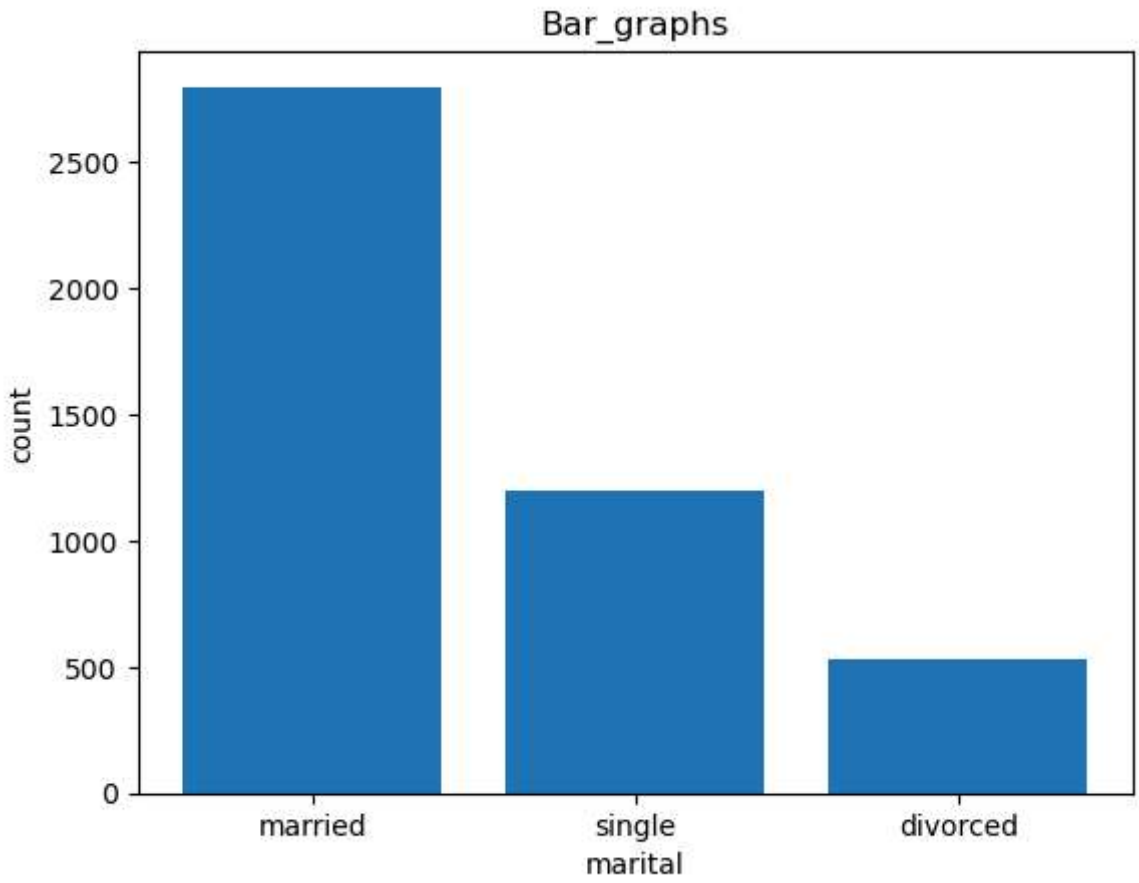
In [129…
```
os.makedirs('C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiment
```

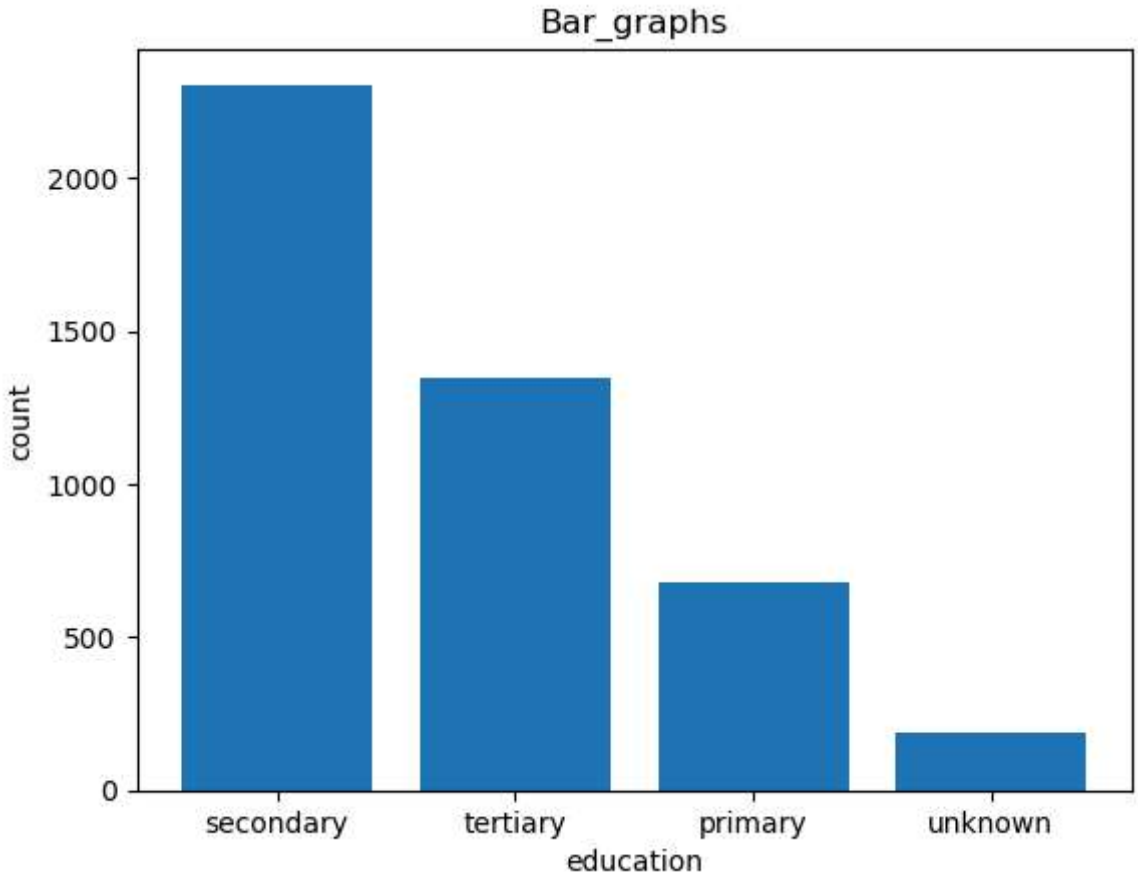
In [137…
```
mat_lib=r'C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiments\\
```

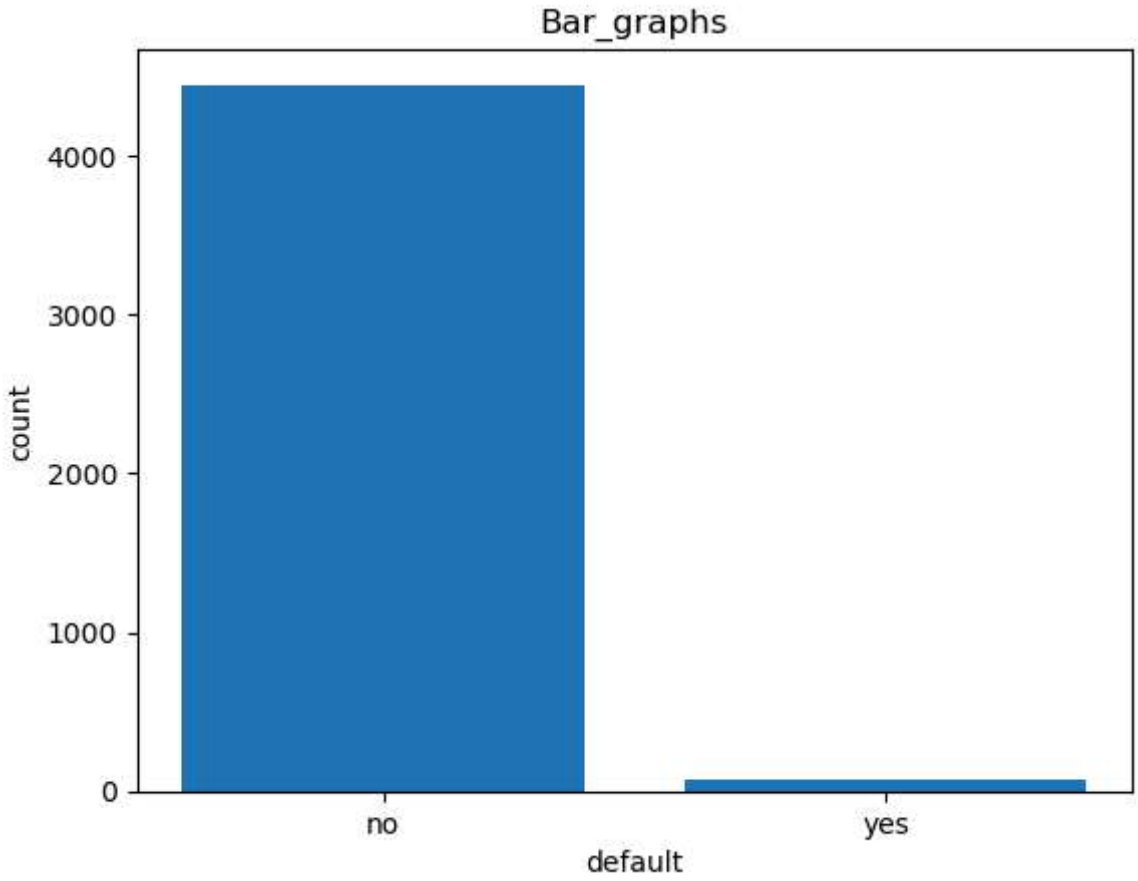
In [139…
```
for i in colm:
    dfs=pd.read_csv(f"{new_dir}\\{i}.csv")
    plt.bar(i,"count",data=dfs)
    plt.title("Bar_graphs")
    plt.xlabel(f"{i}")
    plt.ylabel("count")
    plt.savefig(f"{mat_lib}\\{i}.jpg")
    plt.show()
```

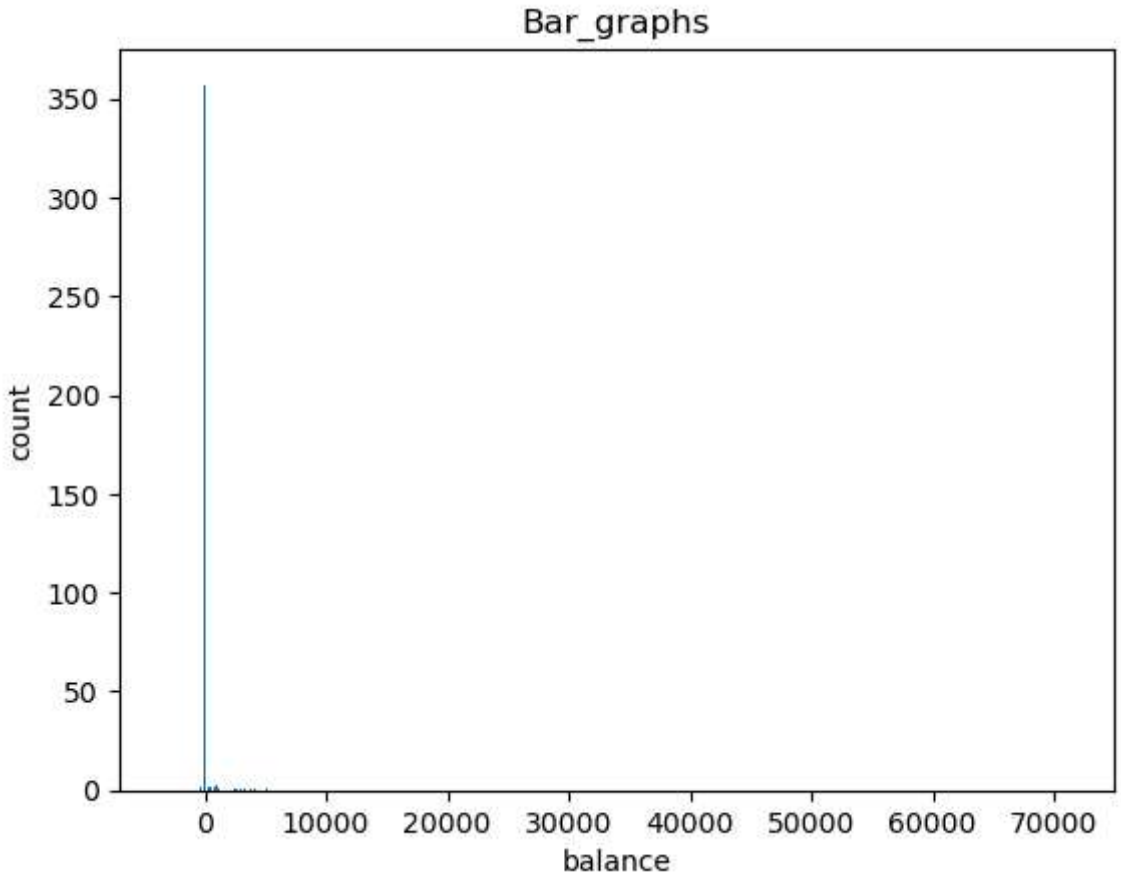


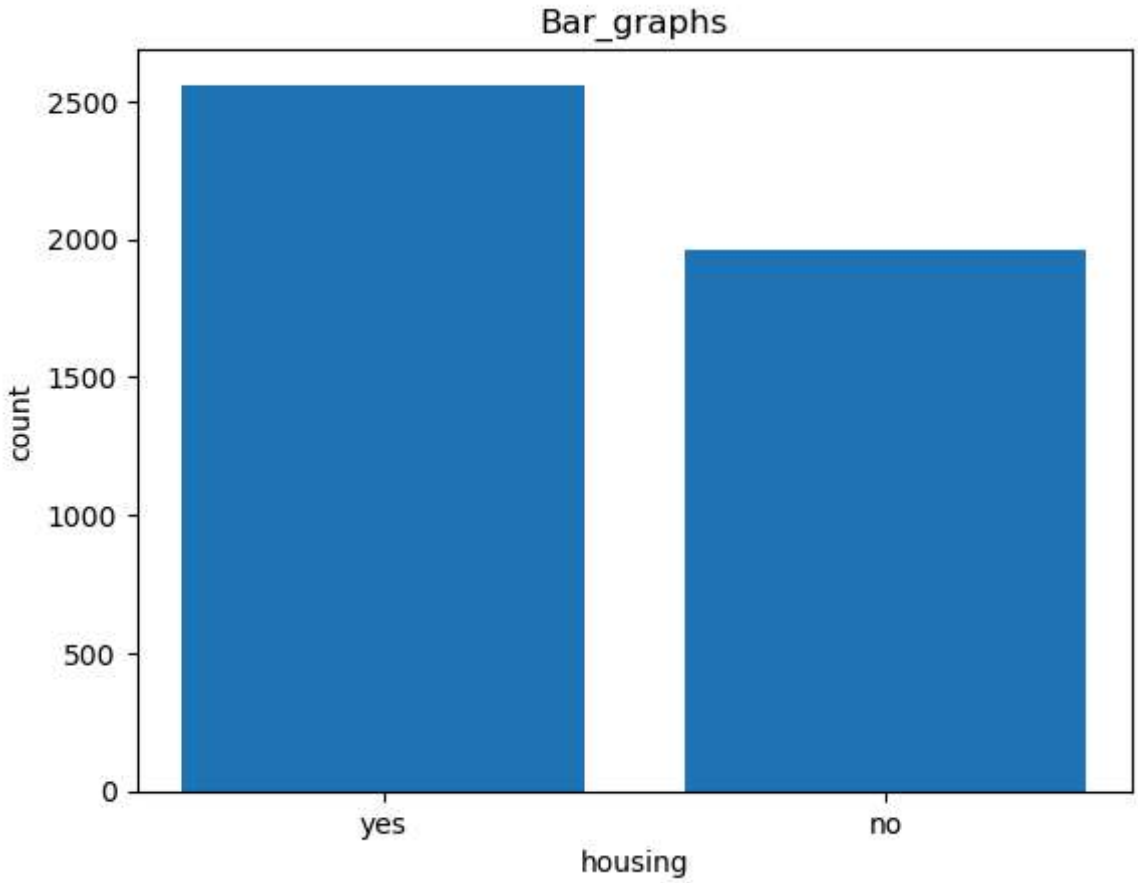

Bar_graphs
1000
800
600
400
200
0
count
job
unknown

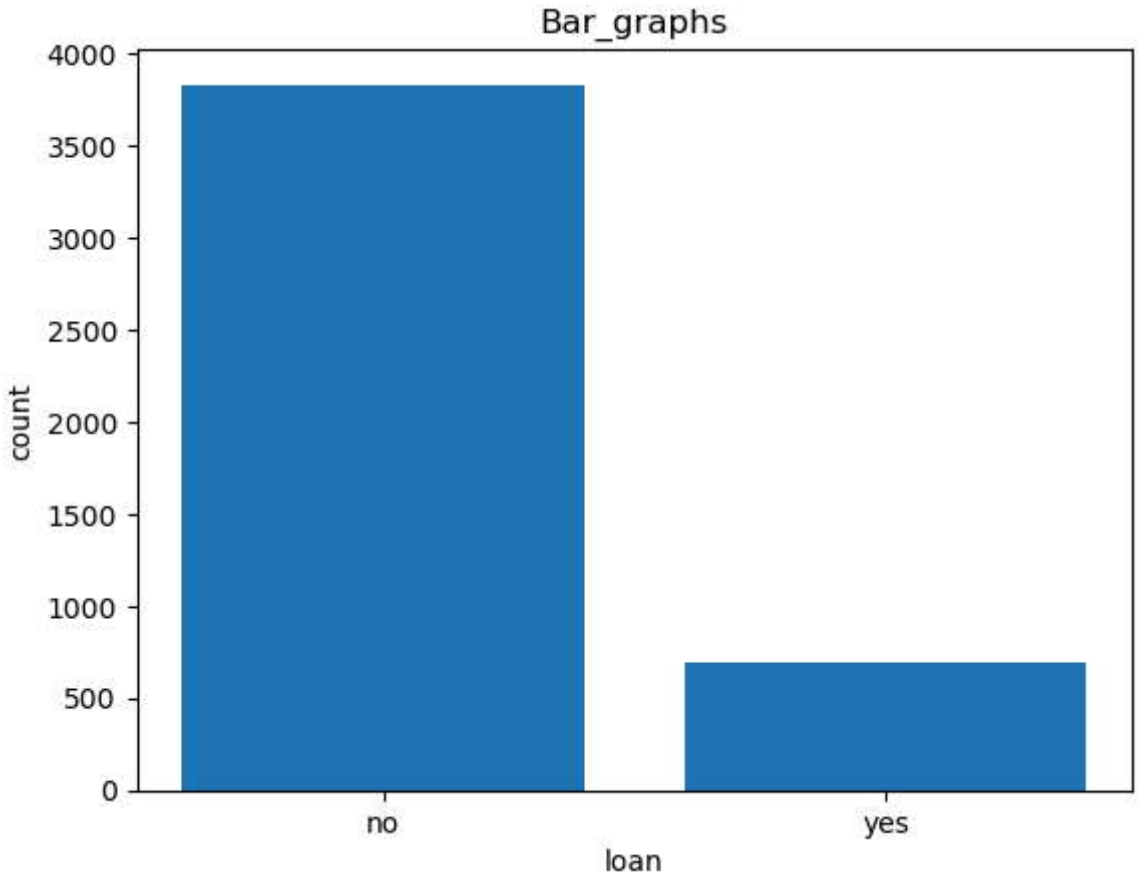

Bar_graphs
2500
2000
1500
1000
500
0
count
married
single
divorced
marital

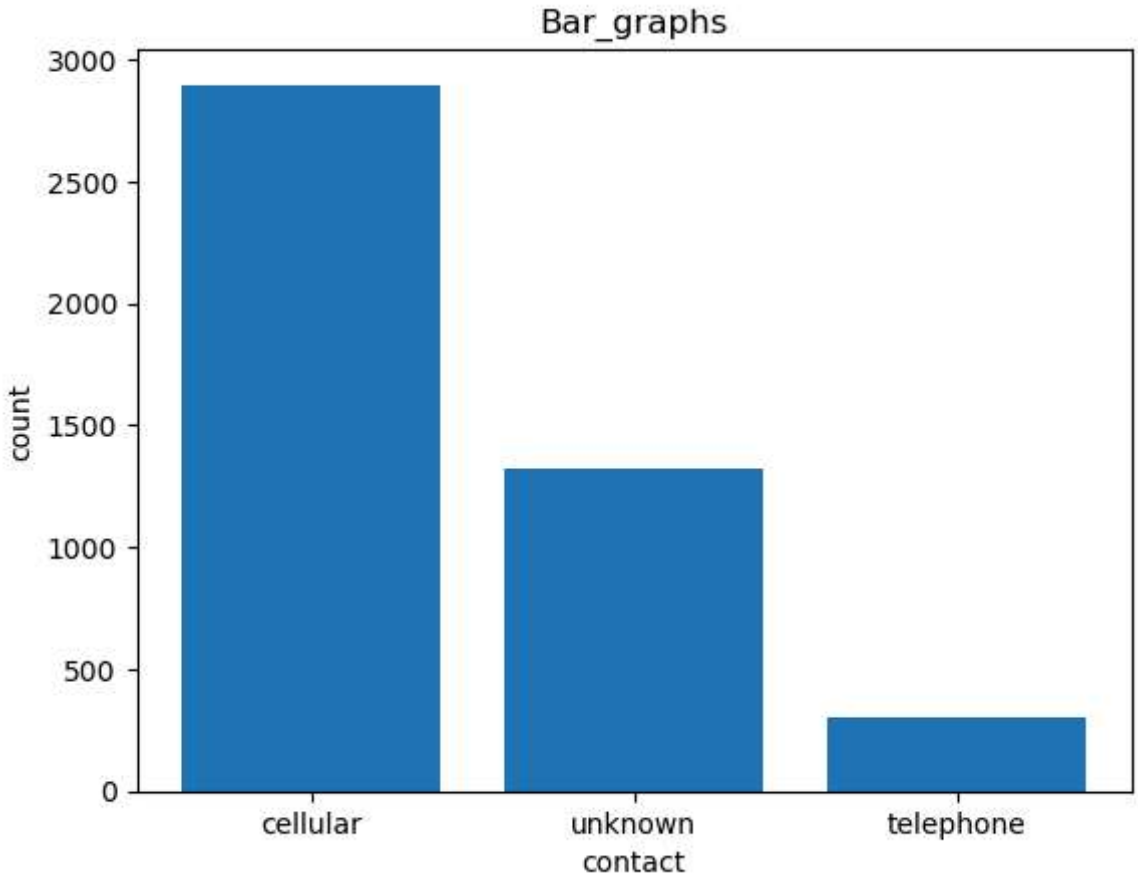

Bar_graphs
2000
1500
1000
500
0
count
secondary
tertiary
primary
unknown
education

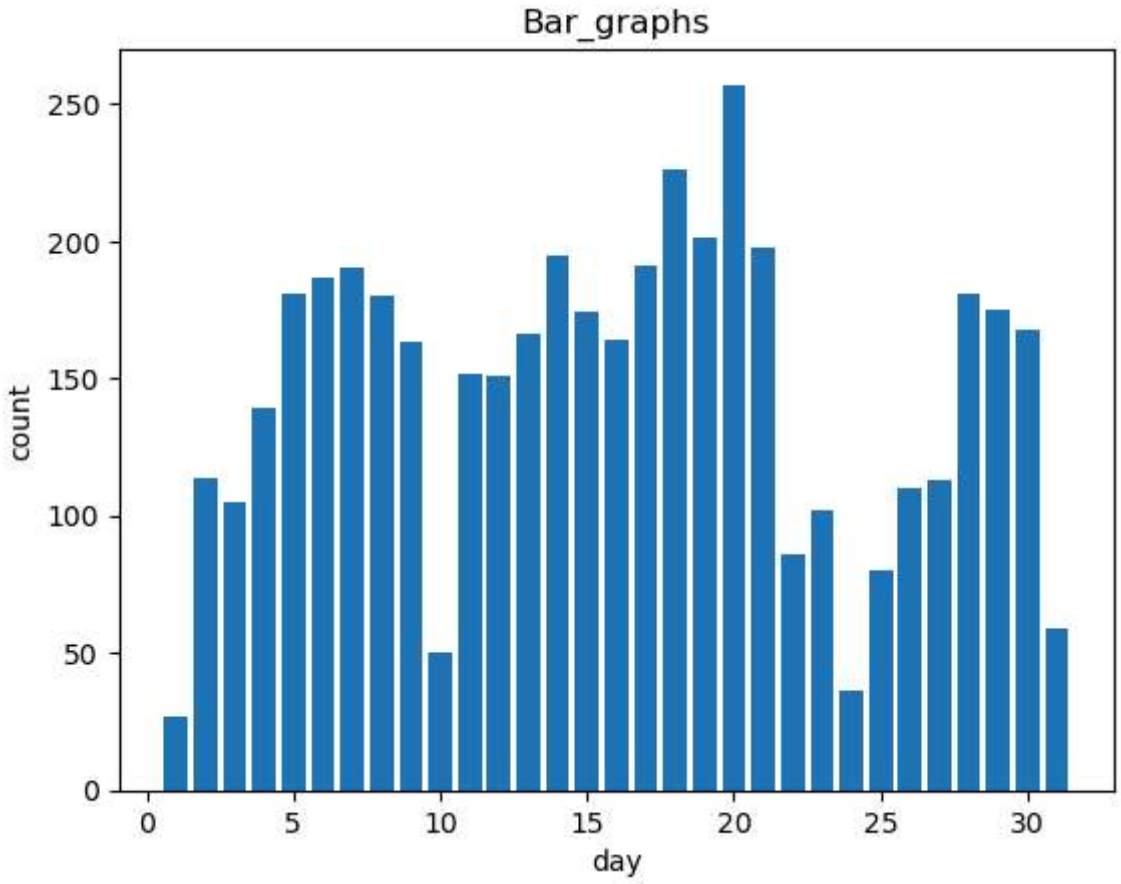

Bar_graphs
4000
3000
2000
1000
0
count
no
yes
default

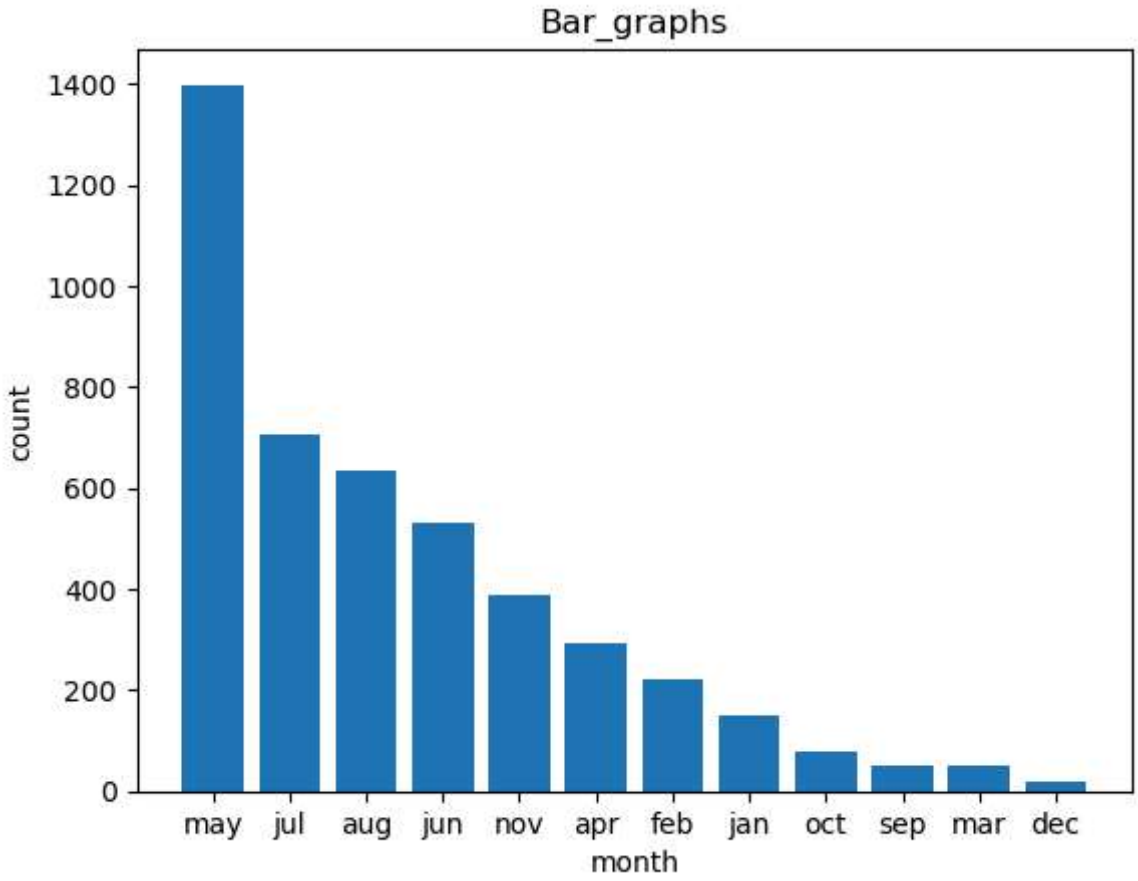

Bar_graphs
count
balance
0
50
100
150
200
250
300
350
0
10000
20000
30000
40000
50000
60000
70000

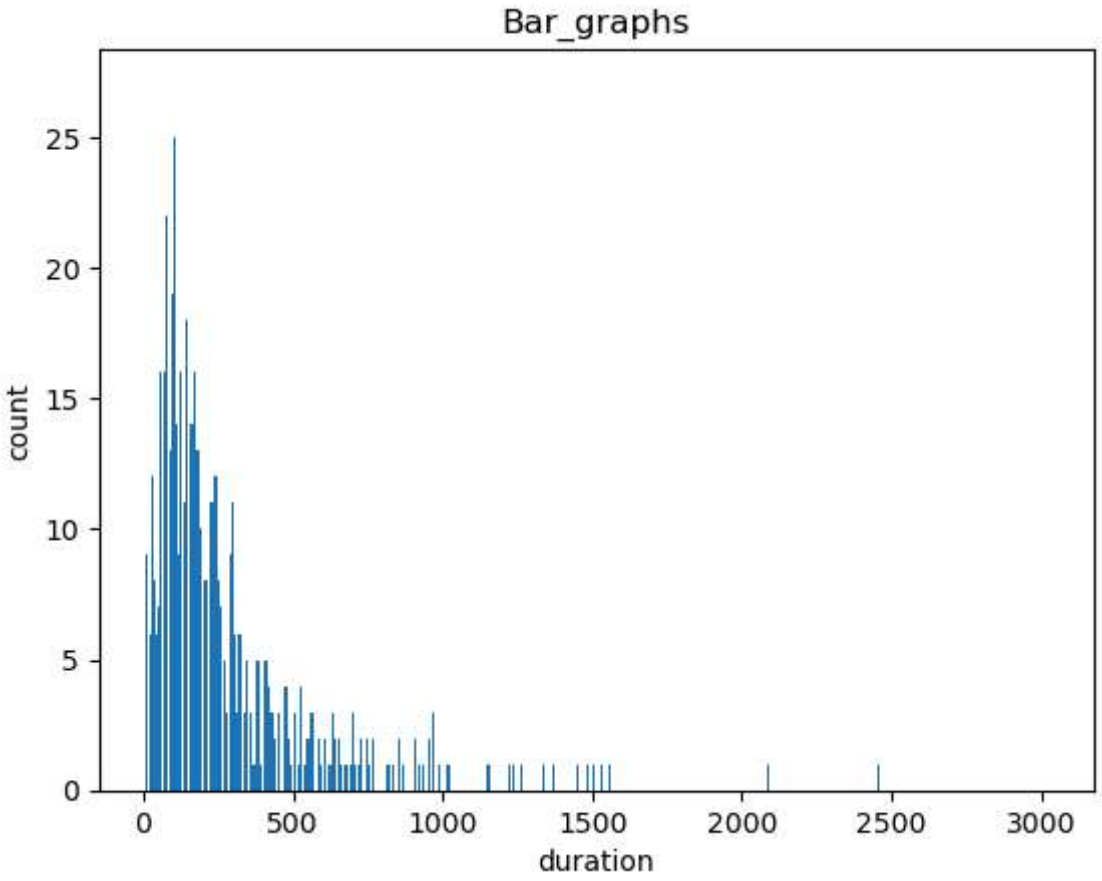

Bar_graphs
count
housing
0
500
1000
1500
2000
2500
yes
no

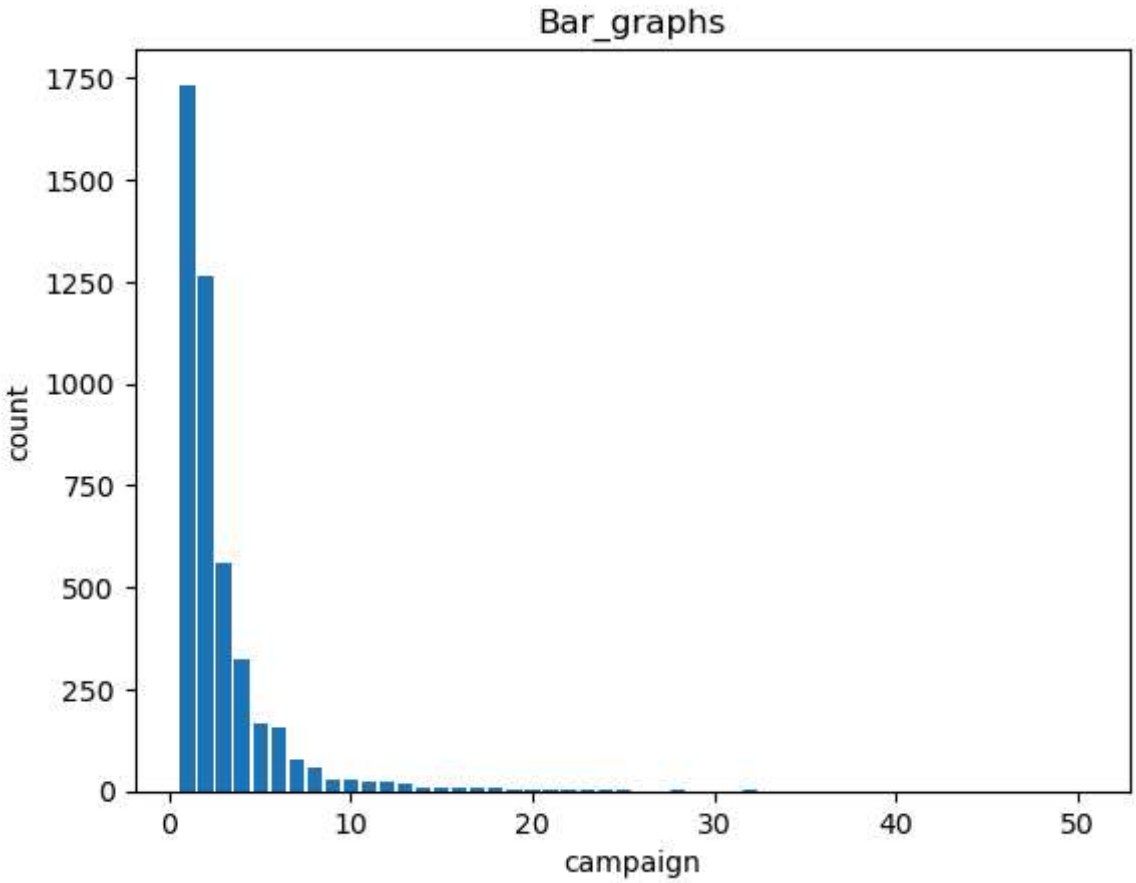

Bar_graphs
4000
3500
3000
2500
2000
1500
1000
500
0
count
no
yes
loan

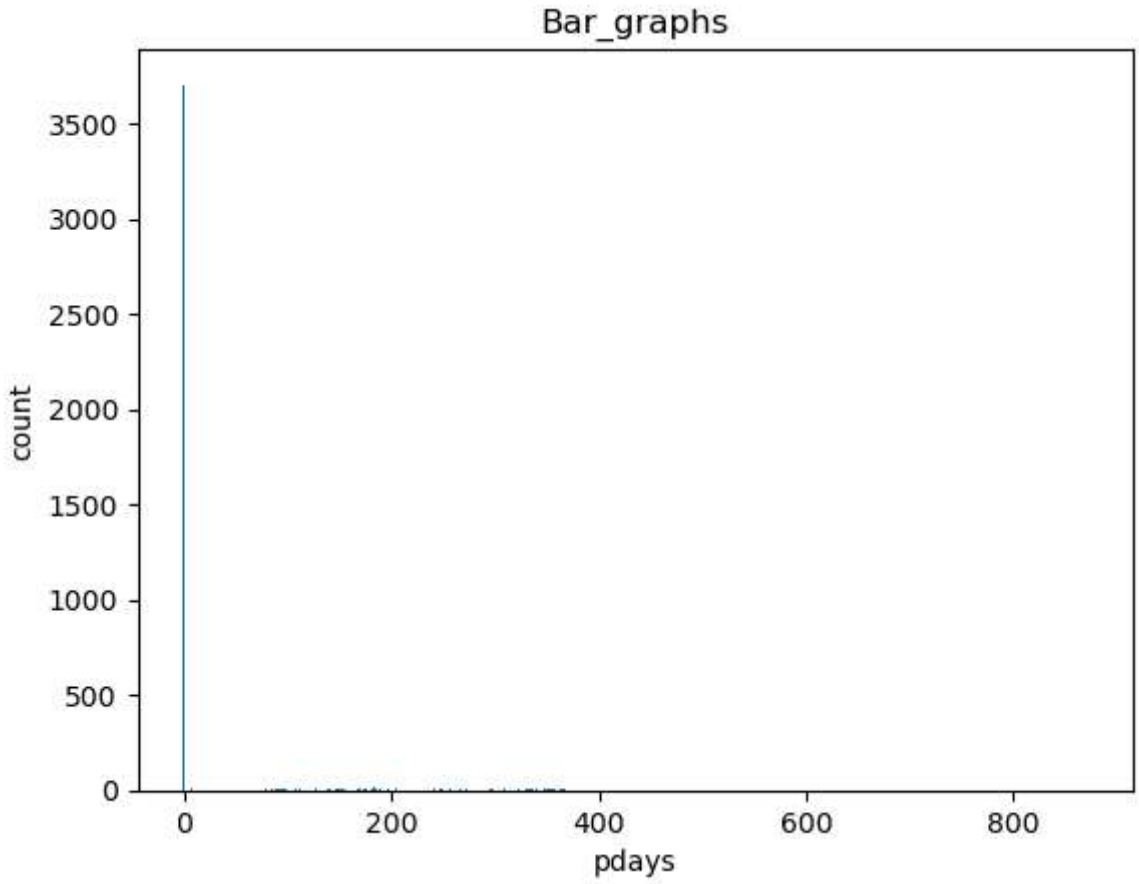

Bar_graphs
3000
2500
2000
1500
1000
500
0
count
cellular
unknown
telephone
contact

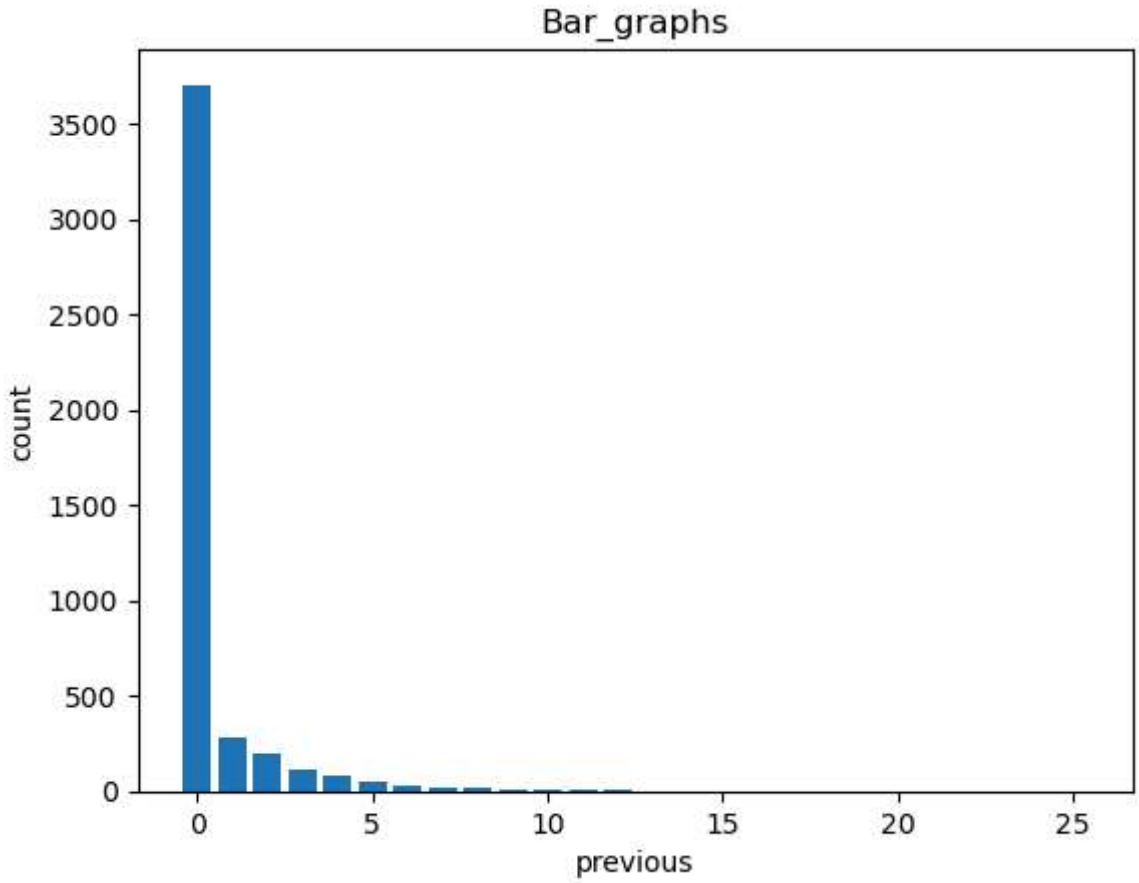

Bar_graphs
count
day
0
50
100
150
200
250
0
5
10
15
20
25
30

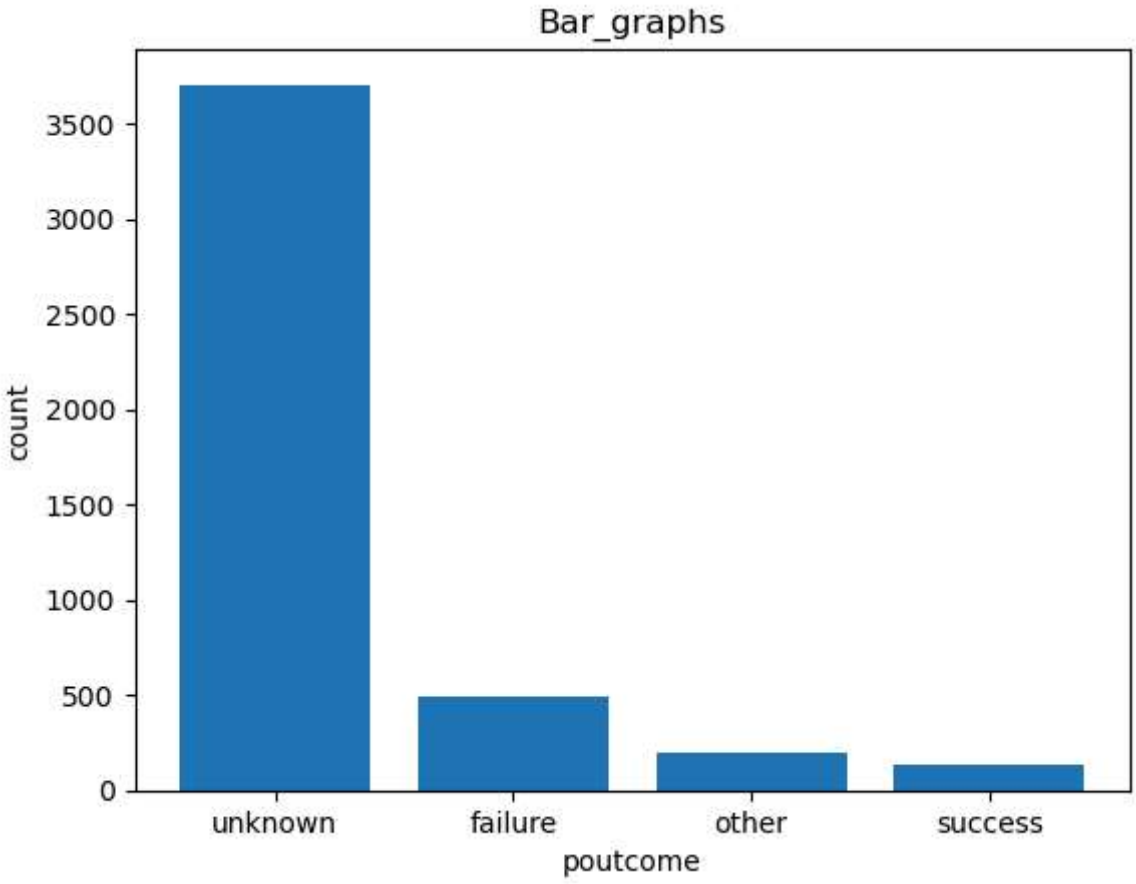

Bar_graphs
count
month
0
200
400
600
800
1000
1200
1400
may
jul
aug
jun
nov
apr
feb
jan
oct
sep
mar
dec

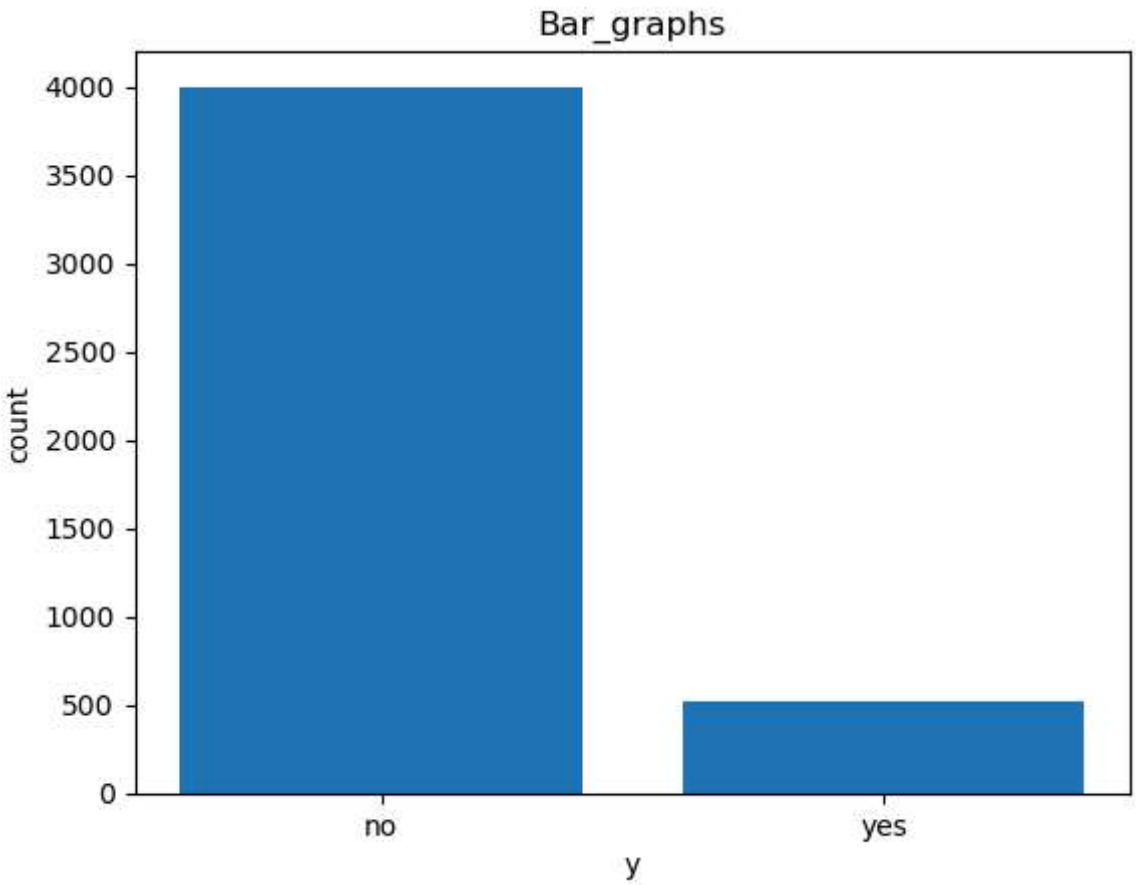

Bar_graphs
count
duration
0
5
10
15
20
25
0
500
1000
1500
2000
2500
3000


Bar_graphs
count
campaign
0
250
500
750
1000
1250
1500
1750
0
10
20
30
40
50


Bar_graphs
3500
3000
2500
2000
1500
1000
500
0
count
0
200
400
600
800
pdays


Bar_graphs
3500
3000
2500
2000
1500
1000
500
0
count
0
5
10
15
20
25
previous

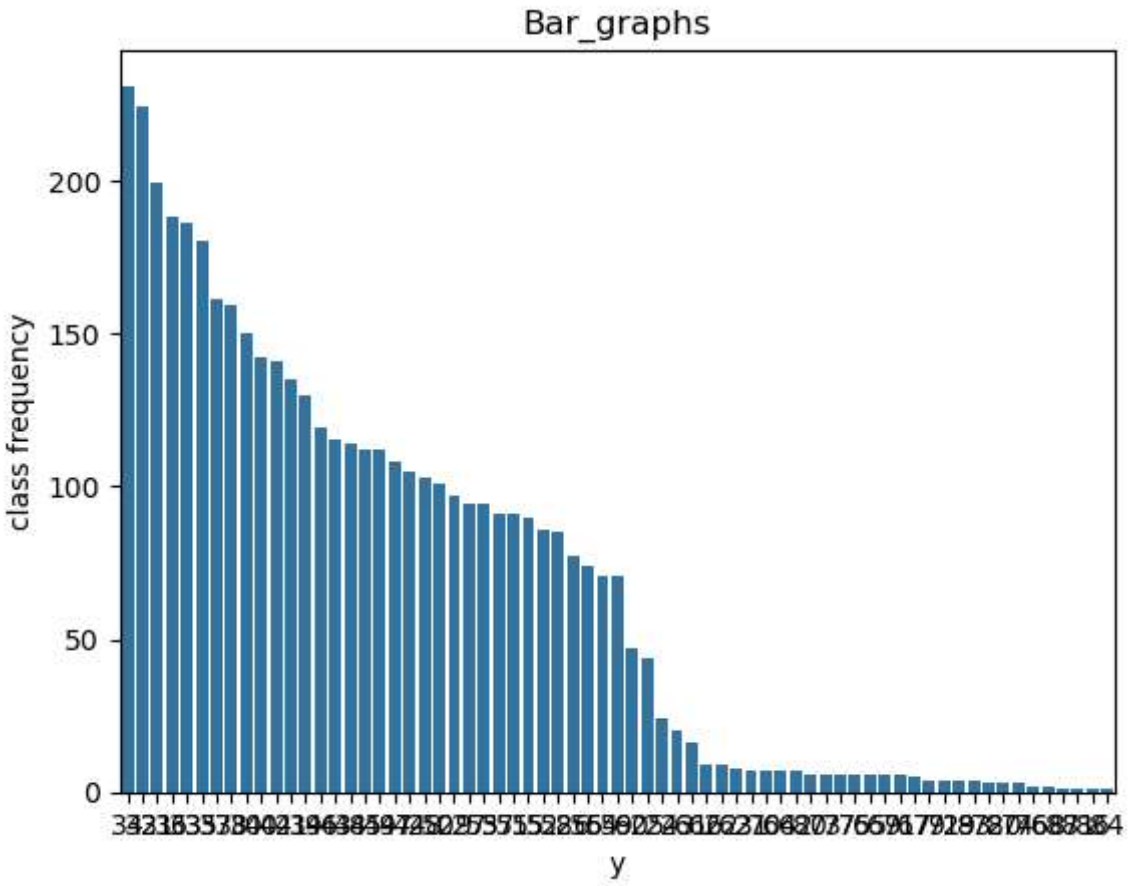


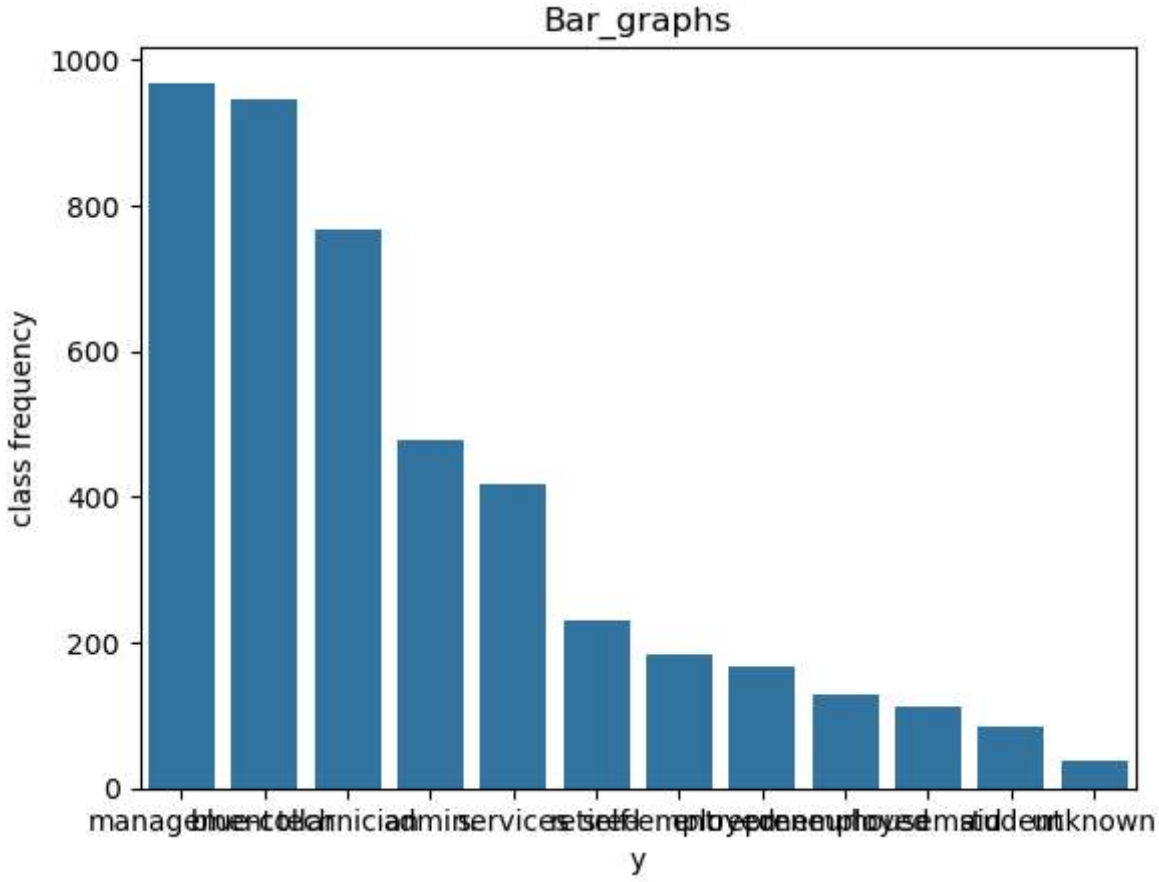


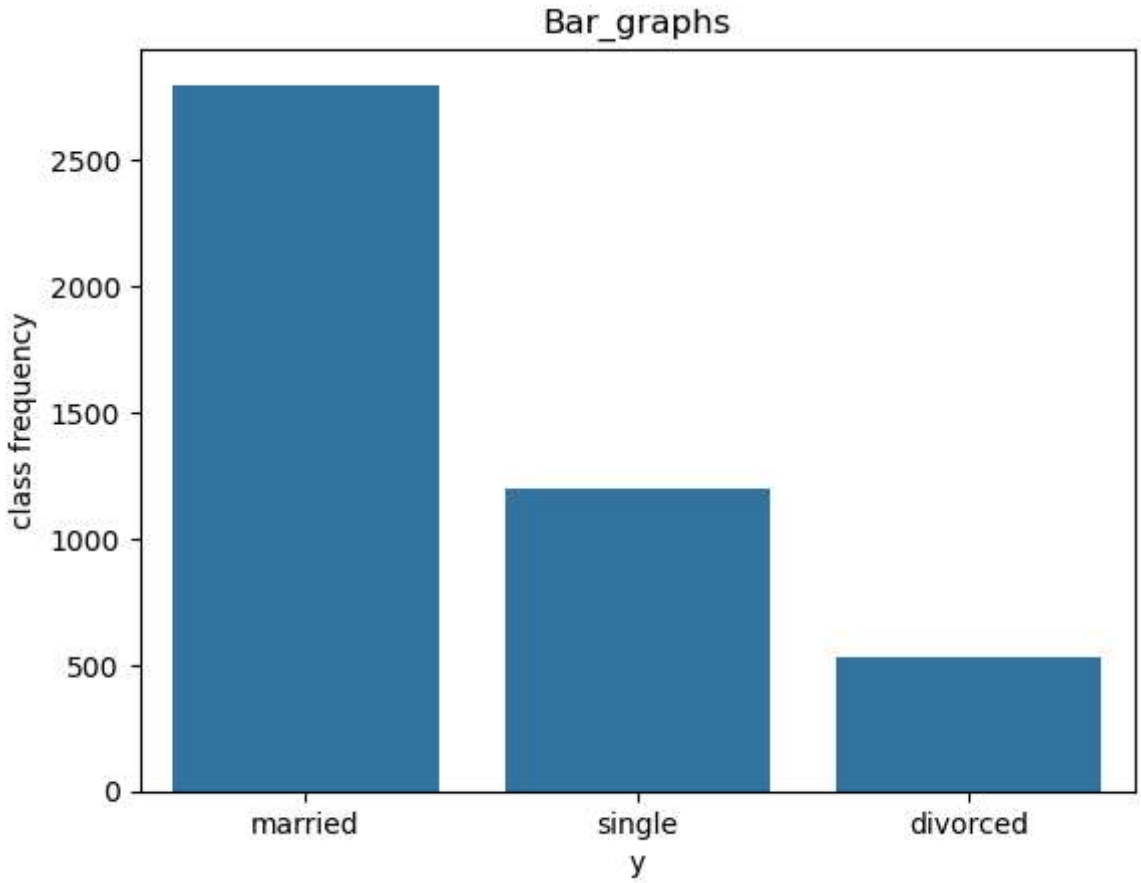
In [143…
```
os.makedirs('C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiment
```

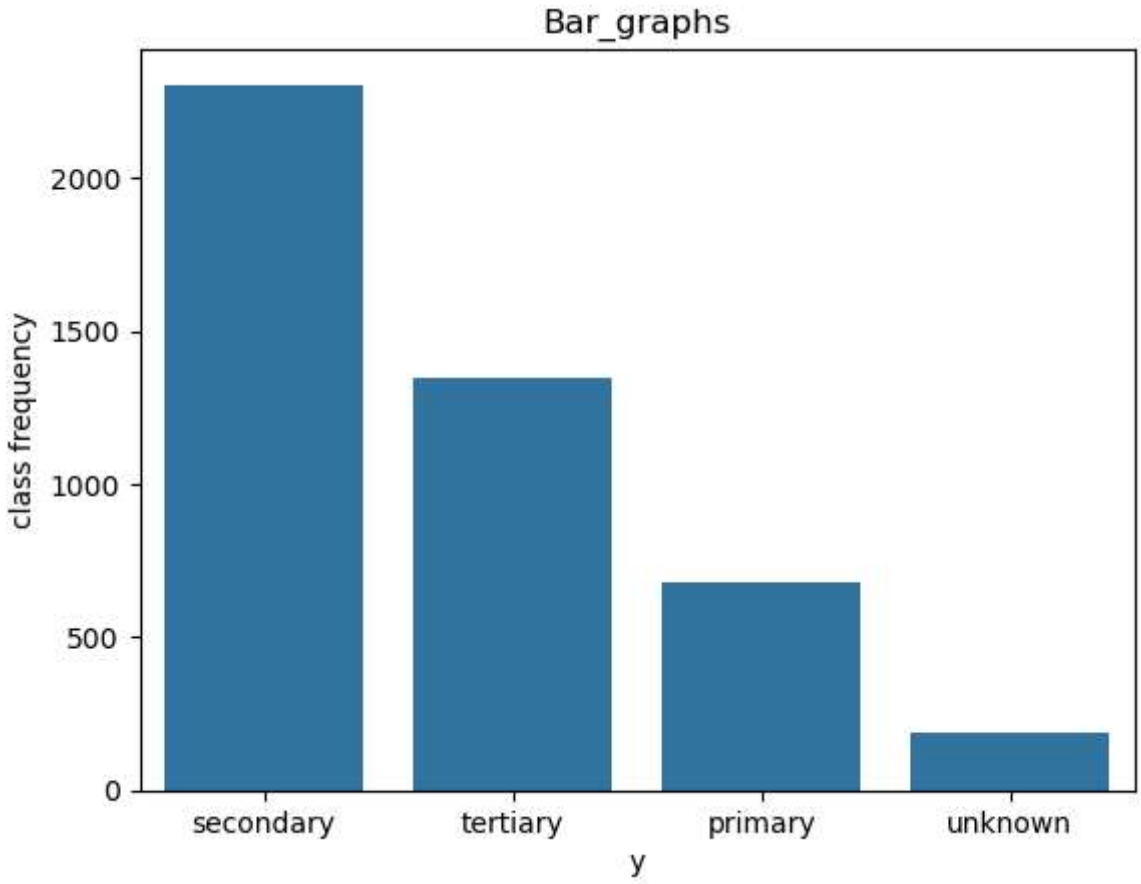
In [145…
```
sea_born=r'C:\\Users\\Mrityunjay\\Desktop\\Data science naresh it\\Class Assiments\
```

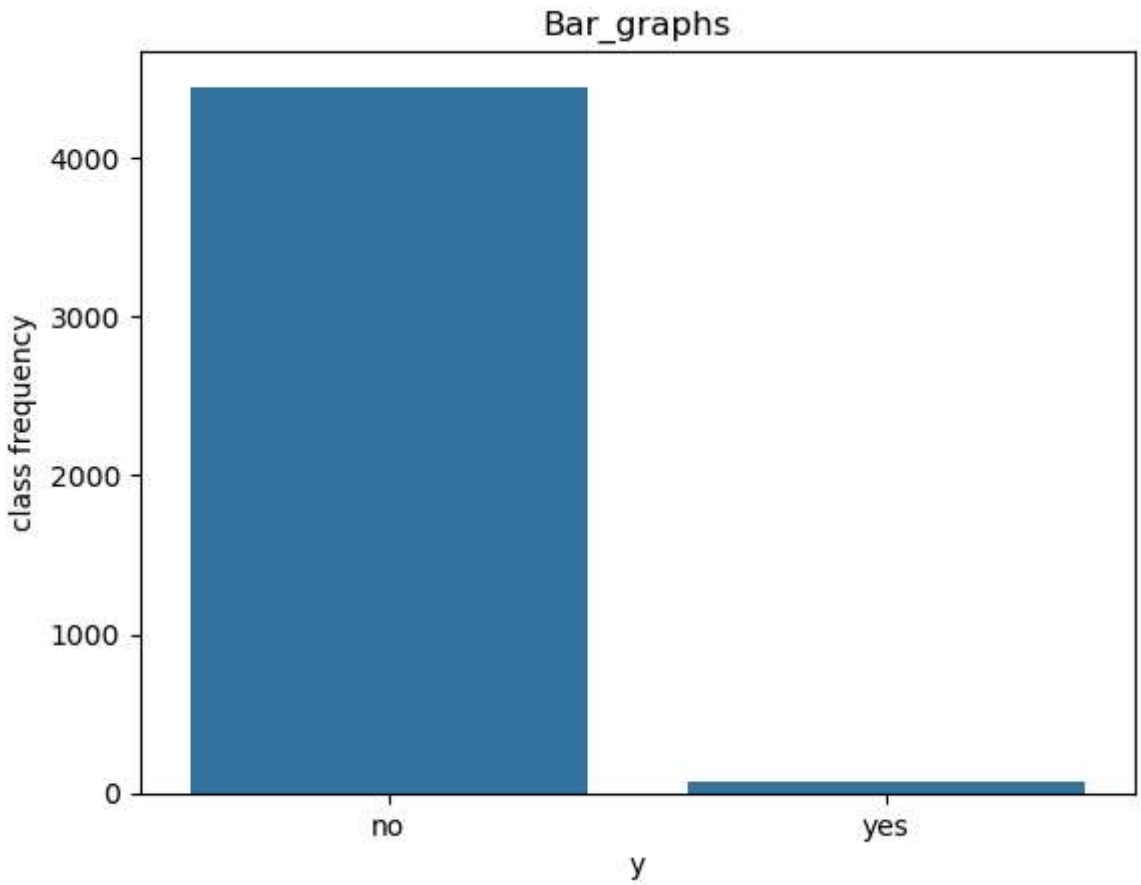
In [175…
```
orders={}
for i in colm:
    cnf=bank_df[i].value_counts()

    orders[i]=cnf.keys()


for j in colm:

    sns.countplot(data=bank_df,
                  x=j,
                  order=orders[j])
    plt.title("Bar_graphs")
    plt.xlabel(f"{i}")
    plt.ylabel("class frequency")
    plt.show()
```

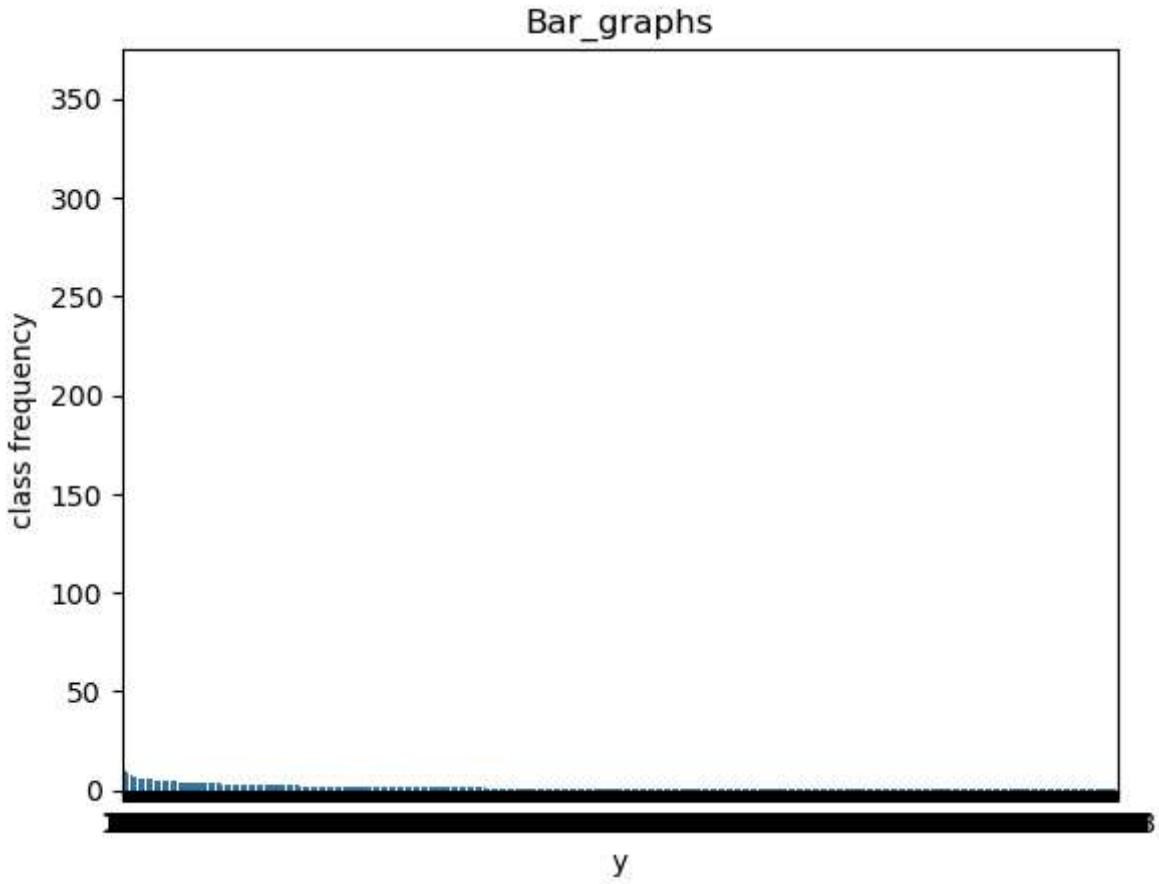


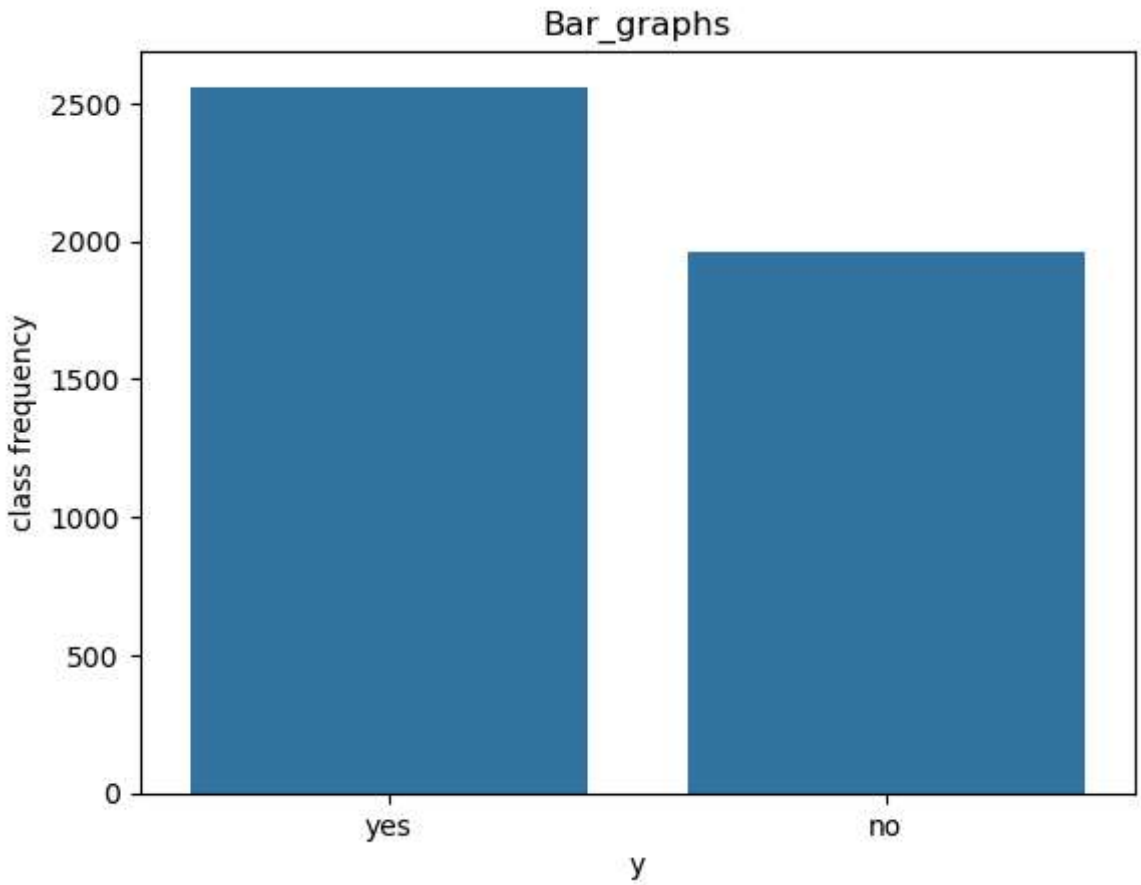

Bar_graphs
1000
800
600
400
200
0
class frequency
y
unknown

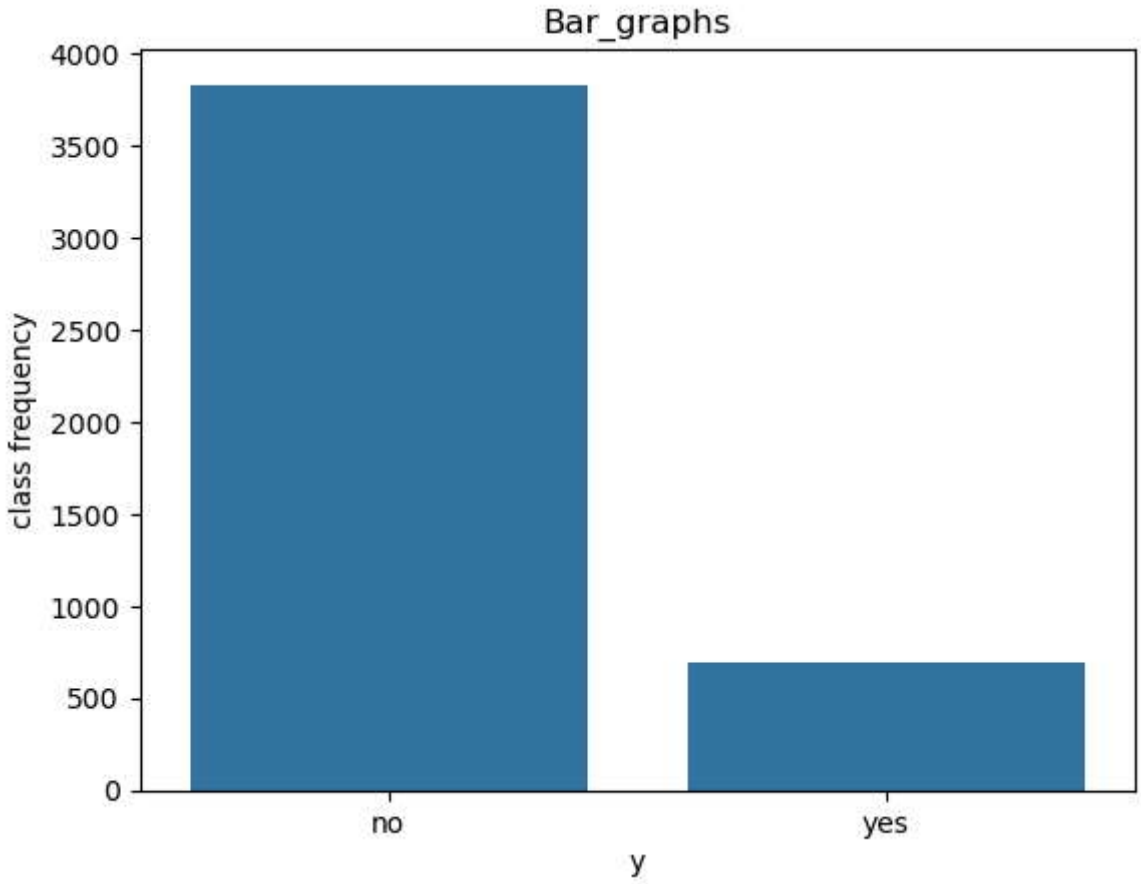

Bar_graphs
2500
2000
1500
1000
500
0
class frequency
married
single
divorced
y

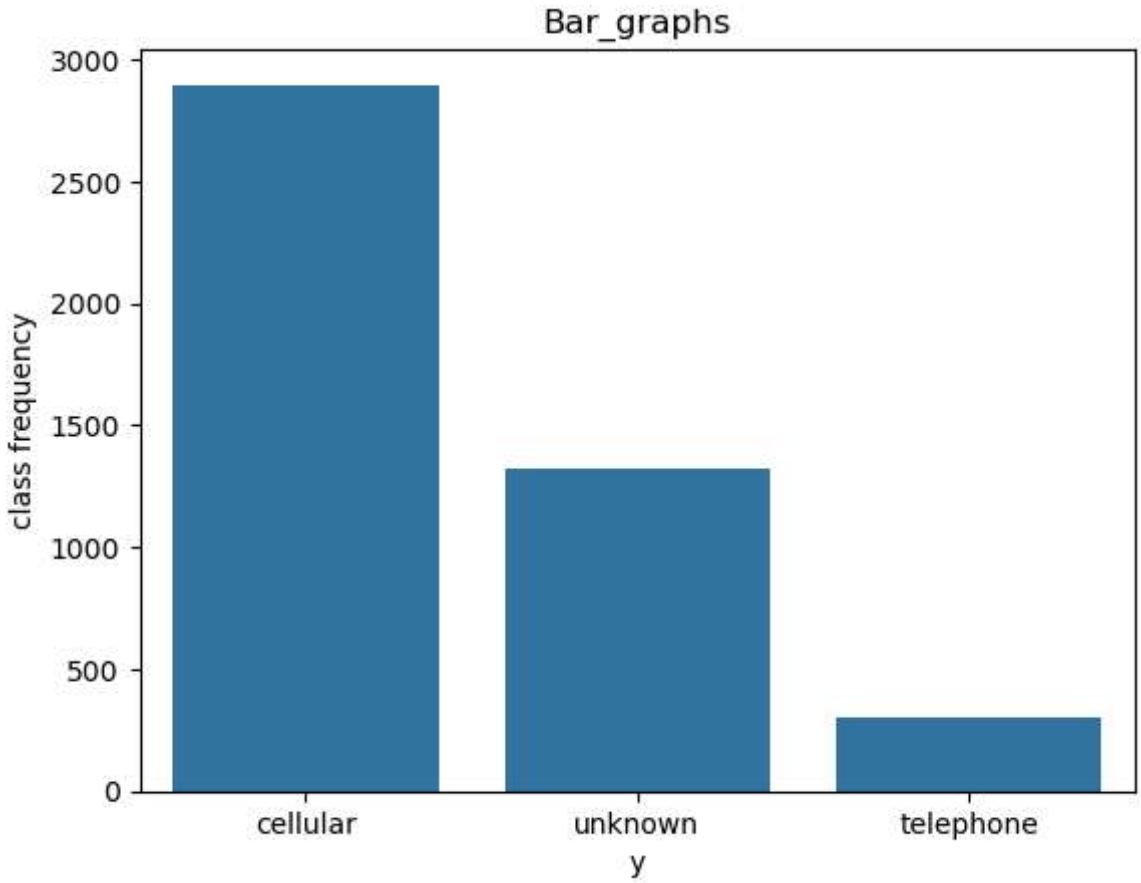

Bar_graphs
class frequency
2000
1500
1000
500
0
secondary
tertiary
primary
unknown
y

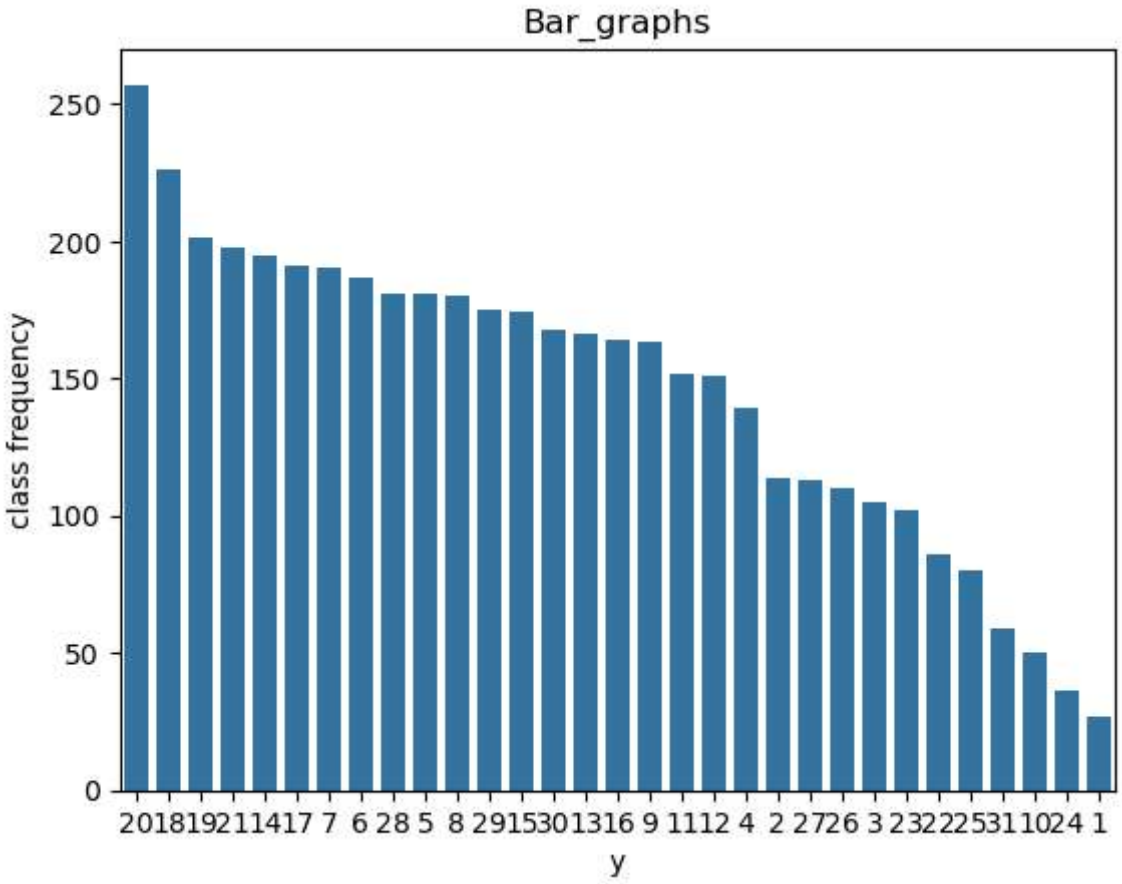

Bar_graphs
class frequency
4000
3000
2000
1000
0
no
yes
y

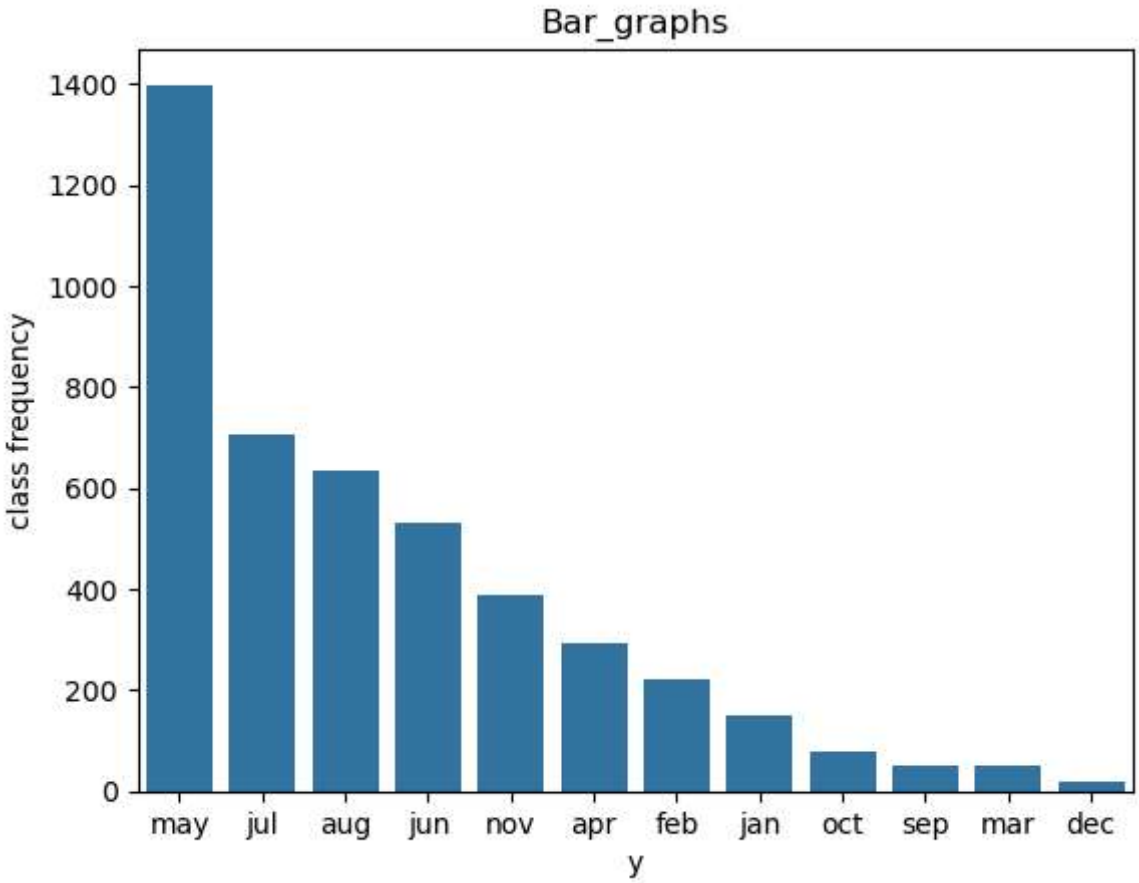

Bar_graphs
350
300
250
200
150
100
50
0
class frequency
y

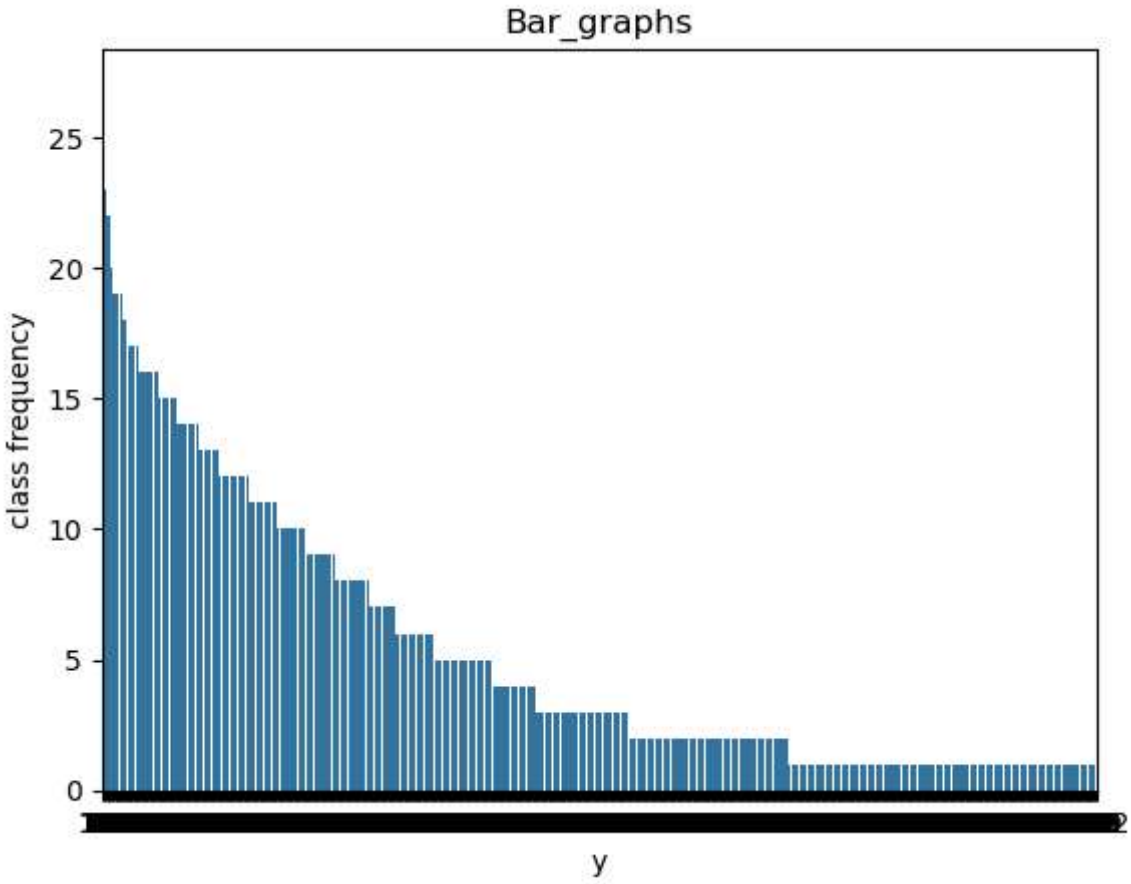

Bar_graphs
2500
2000
1500
1000
500
0
class frequency
yes
no
y

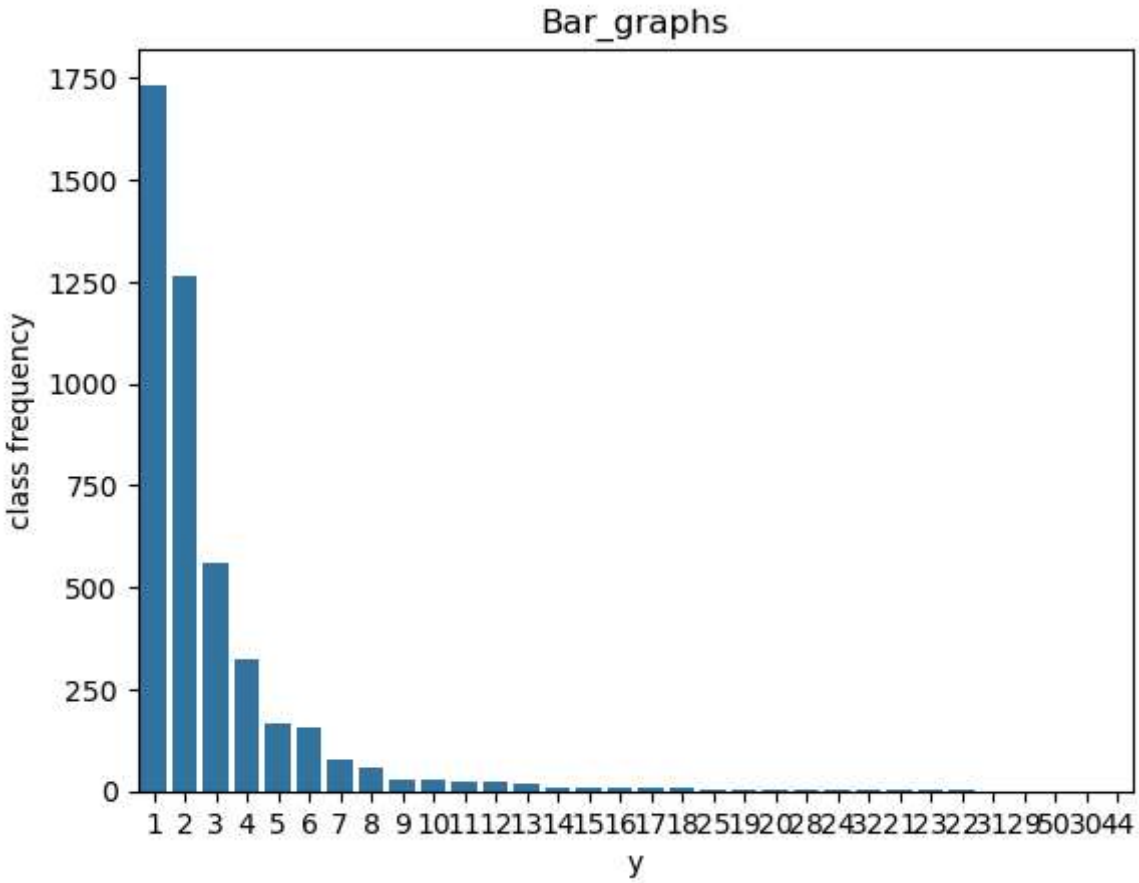

Bar_graphs
4000
3500
3000
2500
2000
1500
1000
500
0
class frequency
no
yes
y

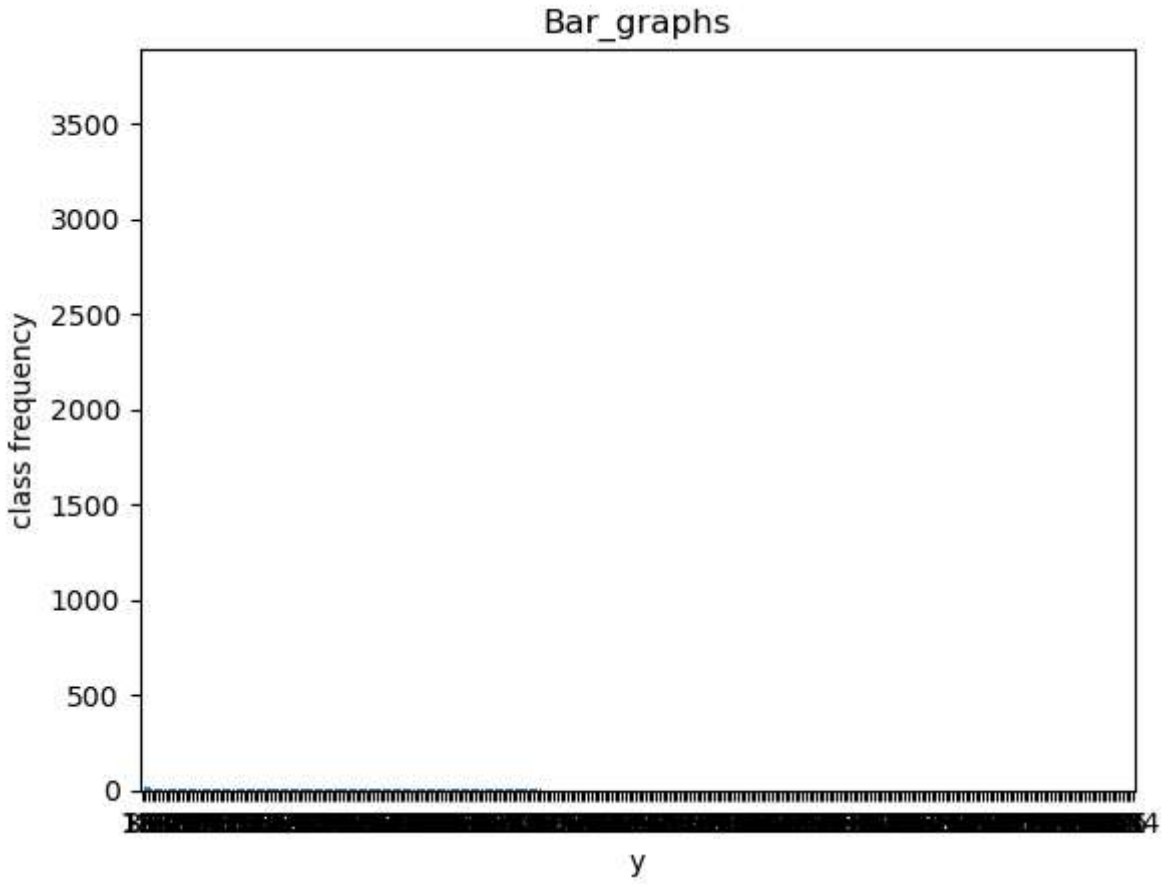

Bar_graphs
3000
2500
2000
1500
1000
500
0
class frequency
cellular
unknown
telephone
y

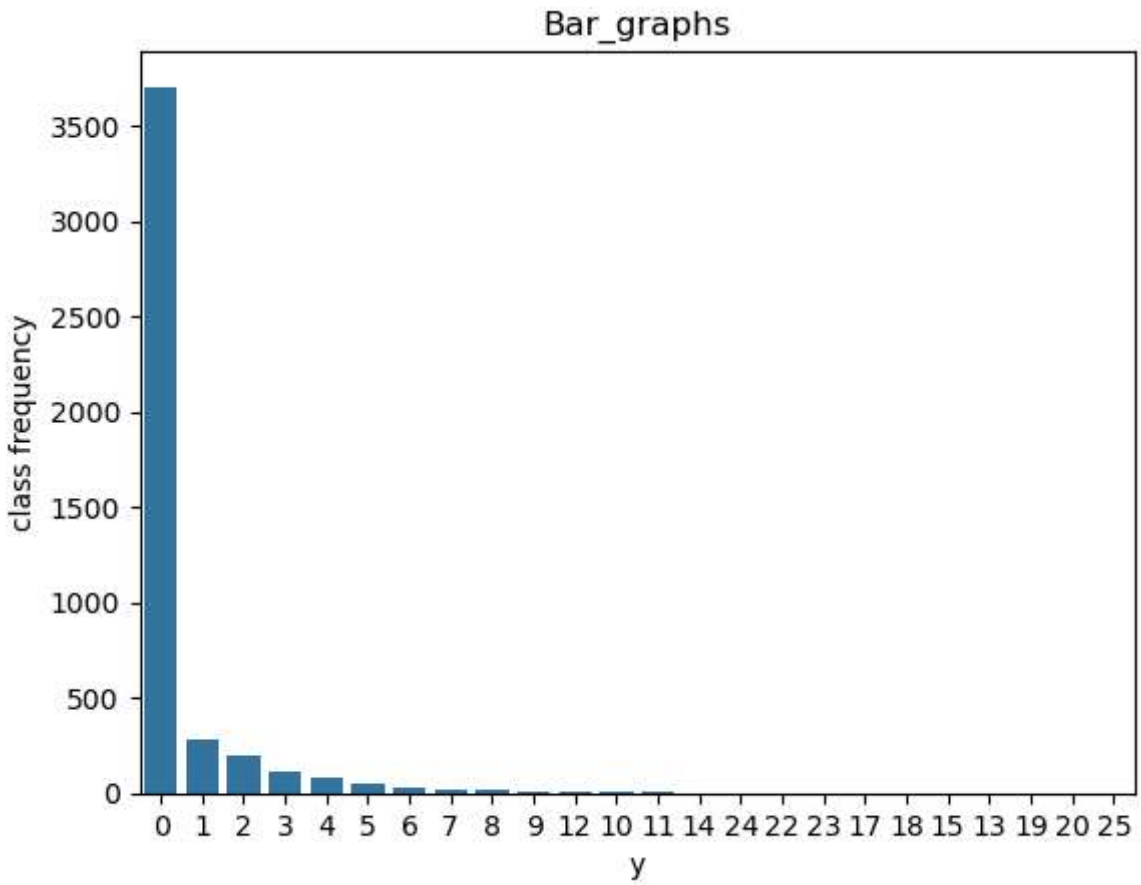

Bar_graphs
class frequency
0
50
100
150
200
250
20 18 19 2 11 4 17 7 6 28 5 8 29 15 30 13 16 9 11 12 4 2 27 26 3 23 22 25 31 10 24 1
y

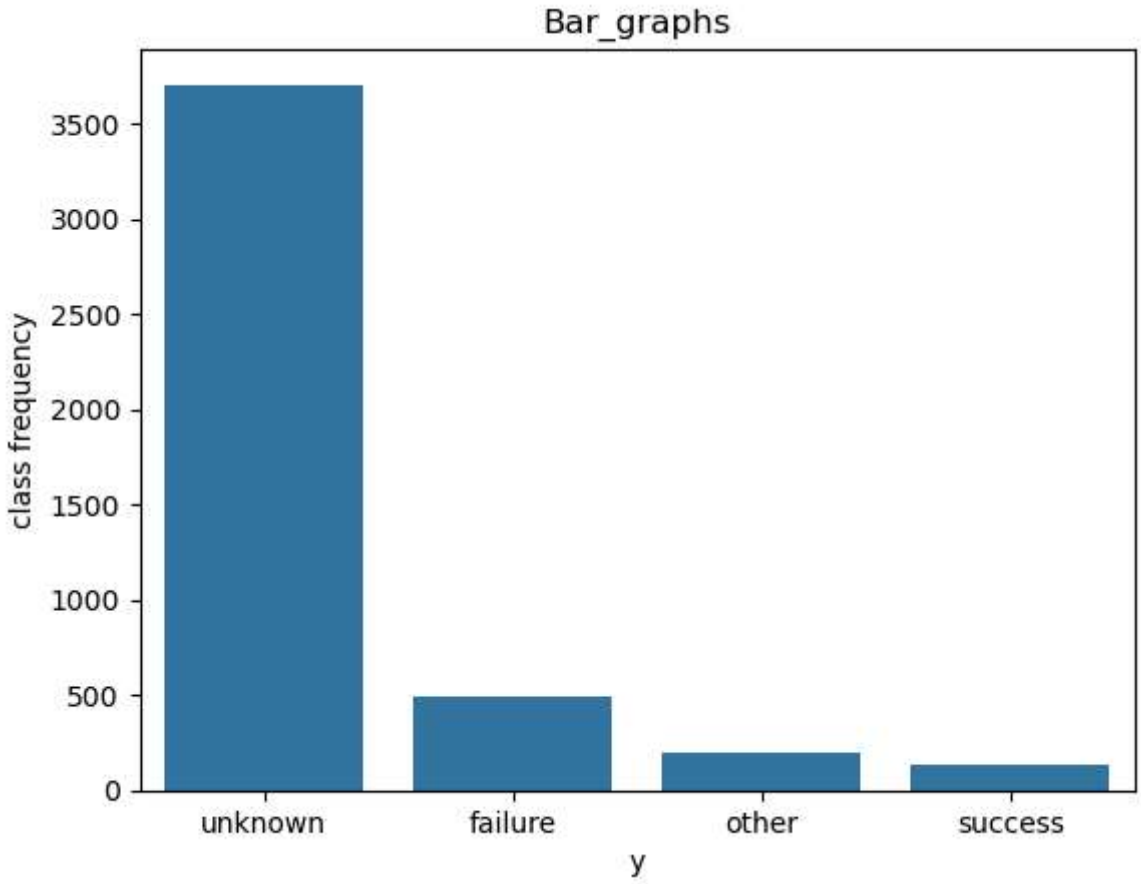

Bar_graphs
class frequency
0
200
400
600
800
1000
1200
1400
may jul aug jun nov apr feb jan oct sep mar dec
y

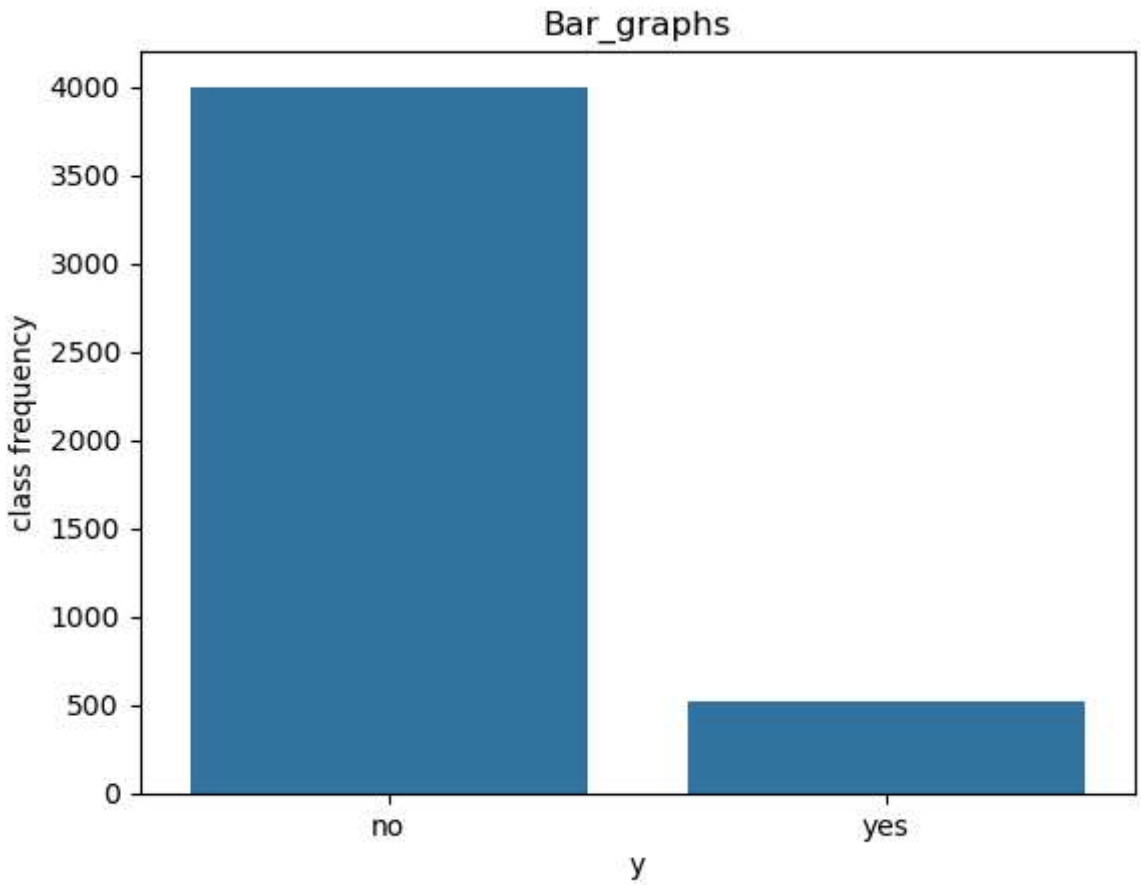

Bar_graphs
class frequency
y
0
5
10
15
20
25


Bar_graphs
class frequency
y
0
250
500
750
1000
1250
1500
1750
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 25 19 20 28 24 32 22 12 32 23 12 95 03 44


Bar_graphs
class frequency
3500
3000
2500
2000
1500
1000
500
0
y


Bar_graphs
class frequency
3500
3000
2500
2000
1500
1000
500
0
0 1 2 3 4 5 6 7 8 9 12 10 11 14 24 22 23 17 18 15 13 19 20 25
y


Bar_graphs
class frequency
3500
3000
2500
2000
1500
1000
500
0
unknown
failure
other
success
y


Bar_graphs
class frequency
4000
3500
3000
2500
2000
1500
1000
500
0
no
yes
y

In [ ]: