```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import sklearn
from sklearn.preprocessing import StandardScaler
```

```
sheet_names = pd.ExcelFile('E:\DATA\quantinum forage\LLOYDS BANKING
GROUP ANALYSIS\Customer_Churn_Data_Large.xlsx').sheet_names
```

```
dfs = [ pd.read_excel('E:\DATA\quantinum forage\LLOYDS BANKING GROUP
ANALYSIS\Customer_Churn_Data_Large.xlsx', sheet_name=sheet) for sheet
in sheet_names ]
```

```
df_cd = dfs[0]
df_th = dfs[1]
df_cs = dfs[2]
df_oa = dfs[3]
df_churn = dfs[4]
```

```
df_cd
```

{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"Age","rawType":"int64","type":"integer"},
{"name":"Gender","rawType":"object","type":"string"},
{"name":"MaritalStatus","rawType":"object","type":"string"},
{"name":"IncomeLevel","rawType":"object","type":"string"}],"ref":"a6a5
8328-fcea-4950-9ac6-034784a2802d","rows":
[["0","1","62","M","Single","Low"],["1","2","65","M","Married","Low"],
["2","3","18","M","Single","Low"],["3","4","21","M","Widowed","Low"],
["4","5","21","M","Divorced","Medium"],
["5","6","57","F","Divorced","Medium"],
["6","7","27","F","Married","High"],["7","8","37","M","Single","Low"],
["8","9","39","M","Divorced","High"],
["9","10","68","M","Married","High"],
["10","11","54","M","Divorced","Medium"],
["11","12","41","F","Married","Low"],
["12","13","24","F","Divorced","Low"],
["13","14","42","M","Widowed","Medium"],
["14","15","42","M","Divorced","Medium"],
["15","16","30","F","Single","Medium"],
["16","17","19","M","Married","Low"],
["17","18","56","F","Married","High"],
["18","19","57","F","Divorced","High"],
["19","20","41","M","Single","Low"],
["20","21","64","M","Divorced","High"],
["21","22","42","F","Divorced","Low"],
["22","23","35","F","Divorced","Low"],
["23","24","55","M","Widowed","Medium"],

["24","25","43","F","Married","Low"],
["25","26","31","M","Single","Medium"],
["26","27","26","M","Widowed","Medium"],
["27","28","27","F","Widowed","High"],
["28","29","38","F","Divorced","Medium"],
["29","30","69","M","Divorced","High"],
["30","31","34","M","Divorced","Low"],
["31","32","69","F","Married","High"],
["32","33","23","M","Divorced","High"],
["33","34","33","M","Widowed","High"],
["34","35","65","M","Married","Low"],
["35","36","18","F","Widowed","Low"],
["36","37","36","F","Married","Medium"],
["37","38","53","F","Married","Low"],
["38","39","42","M","Married","Medium"],
["39","40","67","F","Widowed","Low"],
["40","41","69","F","Married","High"],
["41","42","47","M","Widowed","Medium"],
["42","43","37","M","Single","Low"],
["43","44","37","F","Widowed","Low"],
["44","45","32","F","Widowed","Low"],
["45","46","57","F","Married","High"],
["46","47","50","F","Married","Medium"],
["47","48","19","M","Divorced","Medium"],
["48","49","27","F","Single","Medium"],
["49","50","50","F","Single","Low"]],"shape":
{"columns":5,"rows":1000}}

```
print( df_churn.shape )
print( df_cd.shape )
print( df_cs.shape )
print( df_oa.shape )
print( df_th.shape )
```

```
(1000, 2)
(1000, 5)
(1002, 5)
(1000, 4)
(5054, 5)
```

```
df_th
```

{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"TransactionID","rawType":"int64","type":"integer"},
{"name":"TransactionDate","rawType":"datetime64[ns]","type":"datetime"
},{"name":"AmountSpent","rawType":"float64","type":"float"},
{"name":"ProductCategory","rawType":"object","type":"string"}],"ref":"
82ed36f3-a28b-4ad1-aff2-c90ee931781a","rows":[["0","1","7194","2022-
03-27 00:00:00","416.5","Electronics"],["1","2","7250","2022-08-08

```
00:00:00","54.96","Clothing"],["2","2","9660","2022-07-25
00:00:00","197.5","Electronics"],["3","2","2998","2022-01-25
00:00:00","101.31","Furniture"],["4","2","1228","2022-07-24
00:00:00","397.37","Clothing"],["5","2","8903","2022-01-09
00:00:00","285.21","Electronics"],["6","2","3527","2022-09-16
00:00:00","311.34","Electronics"],["7","2","9279","2022-11-19
00:00:00","199.73","Groceries"],["8","3","9839","2022-10-08
00:00:00","241.06","Books"],["9","3","7364","2022-08-15
00:00:00","51.07","Clothing"],["10","3","4967","2022-09-17
00:00:00","419.95","Furniture"],["11","3","1646","2022-07-09
00:00:00","287.72","Groceries"],["12","3","7806","2022-03-02
00:00:00","299.86","Groceries"],["13","3","7329","2022-02-11
00:00:00","403.32","Furniture"],["14","4","3394","2022-07-31
00:00:00","125.64","Electronics"],["15","4","1211","2022-12-01
00:00:00","123.59","Groceries"],["16","4","4235","2022-12-22
00:00:00","241.45","Electronics"],["17","4","3419","2022-12-27
00:00:00","44.22","Clothing"],["18","4","7670","2022-05-22
00:00:00","382.39","Furniture"],["19","5","1742","2022-02-21
00:00:00","475.69","Groceries"],["20","5","9349","2022-08-12
00:00:00","178.91","Furniture"],["21","5","7200","2022-12-02
00:00:00","373.51","Groceries"],["22","5","4938","2022-10-21
00:00:00","155.76","Electronics"],["23","5","5357","2022-07-04
00:00:00","308.8","Electronics"],["24","5","6791","2022-08-31
00:00:00","219.98","Furniture"],["25","5","7592","2022-12-21
00:00:00","218.98","Electronics"],["26","5","9249","2022-05-25
00:00:00","69.86","Groceries"],["27","6","6635","2022-08-11
00:00:00","365.57","Books"],["28","6","3772","2022-10-29
00:00:00","448.86","Clothing"],["29","6","6237","2022-08-09
00:00:00","41.7","Furniture"],["30","6","8661","2022-07-03
00:00:00","104.0","Furniture"],["31","6","9155","2022-10-18
00:00:00","204.16","Electronics"],["32","7","8762","2022-10-24
00:00:00","86.73","Books"],["33","8","4215","2022-10-25
00:00:00","263.24","Books"],["34","8","2500","2022-11-09
00:00:00","483.91","Clothing"],["35","8","1383","2022-07-06
00:00:00","240.92","Books"],["36","8","6171","2022-06-10
00:00:00","346.24","Clothing"],["37","8","1807","2022-07-13
00:00:00","253.68","Furniture"],["38","8","5759","2022-09-26
00:00:00","161.56","Electronics"],["39","8","9841","2022-03-10
00:00:00","297.33","Furniture"],["40","9","7332","2022-06-02
00:00:00","96.67","Electronics"],["41","9","4938","2022-06-11
00:00:00","472.21","Clothing"],["42","9","9647","2022-01-13
00:00:00","44.26","Groceries"],["43","9","9438","2022-10-19
00:00:00","378.14","Groceries"],["44","9","4713","2022-05-26
00:00:00","325.96","Clothing"],["45","10","3487","2022-09-07
00:00:00","5.86","Furniture"],["46","10","7614","2022-10-31
00:00:00","35.29","Clothing"],["47","10","5052","2022-08-02
00:00:00","303.4","Groceries"],["48","10","4074","2022-09-08
00:00:00","223.85","Furniture"],["49","10","7393","2022-01-19
00:00:00","471.18","Furniture"]],"shape":{"columns":5,"rows":5054}}
```

```
df_th = df_th.groupby( 'CustomerID').agg( total_amount =
                                  ('AmountSpent' , 'sum' ) ,
                                  countoftransaction = ('CustomerID' ,
'count')

                                  ).reset_index()

df_th
```

```
{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"total_amount","rawType":"float64","type":"float"},
{"name":"countoftransaction","rawType":"int64","type":"integer"}],"ref
":"90217265-d6b3-4173-9a7b-f718896a324a","rows":
[["0","1","416.5","1"],["1","2","1547.4199999999998","7"],
["2","3","1702.98","6"],["3","4","917.29","5"],
["4","5","2001.49","8"],["5","6","1164.29","5"],["6","7","86.73","1"],
["7","8","2046.88","7"],["8","9","1317.24","5"],
["9","10","1397.36","7"],["10","11","1733.83","5"],
["11","12","87.1","1"],["12","13","392.02","3"],
["13","14","2843.59","9"],["14","15","1914.02","8"],
["15","16","2019.79","9"],["16","17","1625.27","7"],
["17","18","951.3800000000001","5"],["18","19","1017.26","5"],
["19","20","2775.85","9"],["20","21","913.02","5"],
["21","22","1582.73","6"],["22","23","1490.74","4"],
["23","24","1143.22","5"],["24","25","764.77","3"],
["25","26","571.76","3"],["26","27","2210.6","8"],
["27","28","1702.26","6"],["28","29","853.0699999999999","3"],
["29","30","706.66","2"],["30","31","1108.56","5"],
["31","32","1225.52","5"],["32","33","689.0","2"],
["33","34","1886.4","9"],["34","35","779.8399999999999","3"],
["35","36","1321.5500000000002","6"],["36","37","3166.88","9"],
["37","38","874.86","3"],["38","39","1191.94","5"],
["39","40","1578.3500000000001","7"],["40","41","363.22","2"],
["41","42","530.61","3"],["42","43","2213.63","7"],
["43","44","687.8199999999999","3"],["44","45","300.42","1"],
["45","46","134.67","1"],["46","47","1466.37","7"],
["47","48","1597.29","7"],["48","49","641.48","3"],
["49","50","1514.62","4"]],"shape":{"columns":3,"rows":1000}}
```

```
print(df_churn['CustomerID'].nunique())
print(df_cd['CustomerID'].nunique())
print(df_cs['CustomerID'].nunique())
print(df_oa['CustomerID'].nunique())
```

```
1000
1000
668
1000
```

```
# the transaction history column  has 5000+ rows and the dat that it
hold is not useful for us
# we will drop it

df_final = df_cd.merge(df_churn , on = 'CustomerID' , how = 'left') \
                .merge(df_th  , on = 'CustomerID' , how = 'left') \
                .merge(df_oa , on = 'CustomerID' , how = 'left') \
                .merge(df_cs , on = 'CustomerID' , how = 'left')

df_final.shape
```

```
(1334, 15)
```

```
df_final.duplicated().sum()
```

```
np.int64(0)
```

```
df_final.columns
# we have combined the dat from all the sheets and now we will drop
the unwanted columsn which don't have any information
```

```
Index(['CustomerID', 'Age', 'Gender', 'MaritalStatus', 'IncomeLevel',
       'ChurnStatus', 'total_amount', 'countoftransaction',
'LastLoginDate',
       'LoginFrequency', 'ServiceUsage', 'InteractionID',
'InteractionDate',
       'InteractionType', 'ResolutionStatus'],
      dtype='object')
```

```
# the columsn like transaction date , transaction id
#  interaction , data , intereaction id
# gender  , are irrelevent and we will drop them

# checking for missing values
df_final.isnull().sum()
```

{"columns":[{"name":"index","rawType":"object","type":"string"},
{"name":"0","rawType":"int64","type":"integer"}],"ref":"ff3d17e3-a682-
4357-944d-c2c336310bcd","rows":[["CustomerID","0"],["Age","0"],
["Gender","0"],["MaritalStatus","0"],["IncomeLevel","0"],
["ChurnStatus","0"],["total_amount","0"],["countoftransaction","0"],
["LastLoginDate","0"],["LoginFrequency","0"],["ServiceUsage","0"],
["InteractionID","332"],["InteractionDate","332"],
["InteractionType","332"],["ResolutionStatus","332"]],"shape":
{"columns":1,"rows":15}}

```
df_final
```

{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"Age","rawType":"int64","type":"integer"},
{"name":"Gender","rawType":"object","type":"string"},

```
{"name":"MaritalStatus","rawType":"object","type":"string"},
{"name":"IncomeLevel","rawType":"object","type":"string"},
{"name":"ChurnStatus","rawType":"int64","type":"integer"},
{"name":"total_amount","rawType":"float64","type":"float"},
{"name":"countoftransaction","rawType":"int64","type":"integer"},
{"name":"LastLoginDate","rawType":"datetime64[ns]","type":"datetime"},
{"name":"LoginFrequency","rawType":"int64","type":"integer"},
{"name":"ServiceUsage","rawType":"object","type":"string"},
{"name":"InteractionID","rawType":"float64","type":"float"},
{"name":"InteractionDate","rawType":"datetime64[ns]","type":"datetime"
},{"name":"InteractionType","rawType":"object","type":"unknown"},
{"name":"ResolutionStatus","rawType":"object","type":"unknown"}],"ref"
:"778ffd05-0bbd-43af-b3e6-2f10dce0fa31","rows":
[["0","1","62","M","Single","Low","0","416.5","1","2023-10-21
00:00:00","34","Mobile App","6363.0","2022-03-31
00:00:00","Inquiry","Resolved"],
["1","2","65","M","Married","Low","1","1547.4199999999998","7","2023-
12-05 00:00:00","5","Website","3329.0","2022-03-17
00:00:00","Inquiry","Resolved"],
["2","3","18","M","Single","Low","0","1702.98","6","2023-11-15
00:00:00","3","Website","9976.0","2022-08-24
00:00:00","Inquiry","Resolved"],
["3","4","21","M","Widowed","Low","0","917.29","5","2023-08-25
00:00:00","2","Website","7354.0","2022-11-18
00:00:00","Inquiry","Resolved"],
["4","4","21","M","Widowed","Low","0","917.29","5","2023-08-25
00:00:00","2","Website","5393.0","2022-07-03
00:00:00","Inquiry","Unresolved"],
["5","5","21","M","Divorced","Medium","0","2001.49","8","2023-10-27
00:00:00","41","Website",null,null,null,null],
["6","6","57","F","Divorced","Medium","0","1164.29","5","2023-09-17
00:00:00","2","Website","2358.0","2022-05-05
00:00:00","Feedback","Resolved"],
["7","7","27","F","Married","High","0","86.73","1","2023-01-07
00:00:00","32","Mobile App",null,null,null,null],
["8","8","37","M","Single","Low","1","2046.88","7","2023-04-22
00:00:00","17","Online Banking","4191.0","2022-04-29
00:00:00","Inquiry","Unresolved"],
["9","8","37","M","Single","Low","1","2046.88","7","2023-04-22
00:00:00","17","Online Banking","8937.0","2022-09-13
00:00:00","Feedback","Unresolved"],
["10","9","39","M","Divorced","High","0","1317.24","5","2023-06-27
00:00:00","24","Website","7813.0","2022-08-15
00:00:00","Inquiry","Resolved"],
["11","10","68","M","Married","High","1","1397.36","7","2023-01-13
00:00:00","29","Online Banking",null,null,null,null],
["12","11","54","M","Divorced","Medium","0","1733.83","5","2023-01-22
00:00:00","30","Online Banking","6610.0","2022-10-28
00:00:00","Feedback","Resolved"],
["13","11","54","M","Divorced","Medium","0","1733.83","5","2023-01-22
00:00:00","30","Online Banking","7511.0","2022-01-26
00:00:00","Feedback","Unresolved"],
["14","12","41","F","Married","Low","0","87.1","1","2023-12-31
00:00:00","43","Website","8485.0","2022-11-29
00:00:00","Feedback","Resolved"],
["15","12","41","F","Married","Low","0","87.1","1","2023-12-31
00:00:00","43","Website","7872.0","2022-01-12
00:00:00","Inquiry","Unresolved"],
["16","13","24","F","Divorced","Low","0","392.02","3","2023-07-05
00:00:00","10","Mobile App","9681.0","2022-07-27
00:00:00","Complaint","Resolved"],
["17","14","42","M","Widowed","Medium","0","2843.59","9","2023-01-10
00:00:00","26","Online Banking","4459.0","2022-07-25
00:00:00","Complaint","Unresolved"],
["18","14","42","M","Widowed","Medium","0","2843.59","9","2023-01-10
00:00:00","26","Online Banking","8823.0","2022-07-29
00:00:00","Feedback","Resolved"],
["19","15","42","M","Divorced","Medium","0","1914.02","8","2023-06-15
00:00:00","44","Online Banking","6644.0","2022-10-18
00:00:00","Feedback","Unresolved"],
["20","15","42","M","Divorced","Medium","0","1914.02","8","2023-06-15
00:00:00","44","Online Banking","3914.0","2022-02-04
00:00:00","Complaint","Unresolved"],
["21","16","30","F","Single","Medium","1","2019.79","9","2023-03-15
00:00:00","41","Online Banking","5761.0","2022-06-30
00:00:00","Inquiry","Resolved"],
["22","16","30","F","Single","Medium","1","2019.79","9","2023-03-15
00:00:00","41","Online Banking","9278.0","2022-09-08
00:00:00","Inquiry","Resolved"],
["23","17","19","M","Married","Low","0","1625.27","7","2023-09-03
00:00:00","8","Online Banking","2246.0","2022-05-11
00:00:00","Inquiry","Unresolved"],
["24","17","19","M","Married","Low","0","1625.27","7","2023-09-03
00:00:00","8","Online Banking","7540.0","2022-05-06
00:00:00","Inquiry","Resolved"],
["25","18","56","F","Married","High","1","951.3800000000001","5","2023
-11-24 00:00:00","37","Website","9487.0","2022-11-08
00:00:00","Inquiry","Unresolved"],
["26","19","57","F","Divorced","High","0","1017.26","5","2023-03-08
00:00:00","39","Online Banking","9979.0","2022-11-28
00:00:00","Complaint","Unresolved"],
["27","19","57","F","Divorced","High","0","1017.26","5","2023-03-08
00:00:00","39","Online Banking","8180.0","2022-04-18
00:00:00","Complaint","Unresolved"],
["28","20","41","M","Single","Low","0","2775.85","9","2023-08-15
00:00:00","16","Website","9094.0","2022-07-19
00:00:00","Feedback","Resolved"],
["29","21","64","M","Divorced","High","0","913.02","5","2023-06-16
00:00:00","18","Website",null,null,null,null],
["30","22","42","F","Divorced","Low","0","1582.73","6","2023-05-06
00:00:00","42","Mobile App","2931.0","2022-01-07
00:00:00","Inquiry","Resolved"],
["31","22","42","F","Divorced","Low","0","1582.73","6","2023-05-06
00:00:00","42","Mobile App","6355.0","2022-08-06
00:00:00","Inquiry","Resolved"],
["32","23","35","F","Divorced","Low","0","1490.74","4","2023-07-07
00:00:00","19","Mobile App","4991.0","2022-03-18
00:00:00","Complaint","Unresolved"],
["33","23","35","F","Divorced","Low","0","1490.74","4","2023-07-07
00:00:00","19","Mobile App","6401.0","2022-02-23
00:00:00","Complaint","Resolved"],
["34","24","55","M","Widowed","Medium","0","1143.22","5","2023-10-04
00:00:00","12","Mobile App","2733.0","2022-11-02
00:00:00","Complaint","Resolved"],
["35","25","43","F","Married","Low","0","764.77","3","2023-02-15
00:00:00","27","Mobile App","8432.0","2022-12-06
00:00:00","Feedback","Resolved"],
["36","26","31","M","Single","Medium","1","571.76","3","2023-09-29
00:00:00","45","Mobile App","3912.0","2022-12-12
00:00:00","Complaint","Resolved"],
["37","27","26","M","Widowed","Medium","0","2210.6","8","2023-10-01
00:00:00","22","Mobile App","9414.0","2022-06-13
00:00:00","Feedback","Unresolved"],
["38","28","27","F","Widowed","High","0","1702.26","6","2023-10-26
00:00:00","42","Online Banking","4174.0","2022-04-14
00:00:00","Complaint","Unresolved"],
["39","28","27","F","Widowed","High","0","1702.26","6","2023-10-26
00:00:00","42","Online Banking","2029.0","2022-12-01
00:00:00","Inquiry","Unresolved"],
["40","29","38","F","Divorced","Medium","0","853.0699999999999","3","2
023-09-02 00:00:00","34","Website","8538.0","2022-09-13
00:00:00","Inquiry","Resolved"],
["41","29","38","F","Divorced","Medium","0","853.0699999999999","3","2
023-09-02 00:00:00","34","Website","3691.0","2022-04-14
00:00:00","Feedback","Unresolved"],
["42","30","69","M","Divorced","High","0","706.66","2","2023-07-12
00:00:00","38","Website","3289.0","2022-04-04
00:00:00","Complaint","Resolved"],
["43","30","69","M","Divorced","High","0","706.66","2","2023-07-12
00:00:00","38","Website","4606.0","2022-10-16
00:00:00","Complaint","Unresolved"],
["44","31","34","M","Divorced","Low","0","1108.56","5","2023-08-17
00:00:00","23","Online Banking","6635.0","2022-06-18
00:00:00","Inquiry","Resolved"],
["45","31","34","M","Divorced","Low","0","1108.56","5","2023-08-17
00:00:00","23","Online Banking","5130.0","2022-05-09
00:00:00","Complaint","Resolved"],
```

["46","32","69","F","Married","High","0","1225.52","5","2023-01-06
00:00:00","48","Mobile App",null,null,null,null],
["47","33","23","M","Divorced","High","1","689.0","2","2023-05-19
00:00:00","24","Mobile App","8219.0","2022-05-30
00:00:00","Feedback","Resolved"],
["48","34","33","M","Widowed","High","0","1886.4","9","2023-04-03
00:00:00","44","Online Banking","7853.0","2022-07-02
00:00:00","Feedback","Resolved"],
["49","34","33","M","Widowed","High","0","1886.4","9","2023-04-03
00:00:00","44","Online Banking","5032.0","2022-05-19
00:00:00","Feedback","Unresolved"]],"shape":
{"columns":15,"rows":1334}}

```
df_final.drop(columns = [ 'InteractionDate'  ,
'InteractionID' ,'InteractionType' ,'LastLoginDate' ] , inplace = True
)
#we will drop use less columns
```

```
df_final.columns
# only relevant columns are left now
```

```
Index(['CustomerID', 'Age', 'Gender', 'MaritalStatus', 'IncomeLevel',
       'ChurnStatus', 'total_amount', 'countoftransaction',
'LoginFrequency',
       'ServiceUsage', 'ResolutionStatus'],
      dtype='object')
```

```
df_final
```

{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"Age","rawType":"int64","type":"integer"},
{"name":"Gender","rawType":"object","type":"string"},
{"name":"MaritalStatus","rawType":"object","type":"string"},
{"name":"IncomeLevel","rawType":"object","type":"string"},
{"name":"ChurnStatus","rawType":"int64","type":"integer"},
{"name":"total_amount","rawType":"float64","type":"float"},
{"name":"countoftransaction","rawType":"int64","type":"integer"},
{"name":"LoginFrequency","rawType":"int64","type":"integer"},
{"name":"ServiceUsage","rawType":"object","type":"string"},
{"name":"ResolutionStatus","rawType":"object","type":"unknown"}],"ref"
:"ddec546d-6a94-414d-823c-42df0ba24210","rows":
[["0","1","62","M","Single","Low","0","416.5","1","34","Mobile
App","Resolved"],
["1","2","65","M","Married","Low","1","1547.4199999999998","7","5","We
bsite","Resolved"],
["2","3","18","M","Single","Low","0","1702.98","6","3","Website","Reso
lved"],
["3","4","21","M","Widowed","Low","0","917.29","5","2","Website","Reso
lved"],

```
["4","4","21","M","Widowed","Low","0","917.29","5","2","Website","Unre
solved"],
["5","5","21","M","Divorced","Medium","0","2001.49","8","41","Website"
,null],
["6","6","57","F","Divorced","Medium","0","1164.29","5","2","Website",
"Resolved"],
["7","7","27","F","Married","High","0","86.73","1","32","Mobile
App",null],
["8","8","37","M","Single","Low","1","2046.88","7","17","Online
Banking","Unresolved"],
["9","8","37","M","Single","Low","1","2046.88","7","17","Online
Banking","Unresolved"],
["10","9","39","M","Divorced","High","0","1317.24","5","24","Website",
"Resolved"],
["11","10","68","M","Married","High","1","1397.36","7","29","Online
Banking",null],
["12","11","54","M","Divorced","Medium","0","1733.83","5","30","Online
Banking","Resolved"],
["13","11","54","M","Divorced","Medium","0","1733.83","5","30","Online
Banking","Unresolved"],
["14","12","41","F","Married","Low","0","87.1","1","43","Website","Res
olved"],
["15","12","41","F","Married","Low","0","87.1","1","43","Website","Unr
esolved"],
["16","13","24","F","Divorced","Low","0","392.02","3","10","Mobile
App","Resolved"],
["17","14","42","M","Widowed","Medium","0","2843.59","9","26","Online
Banking","Unresolved"],
["18","14","42","M","Widowed","Medium","0","2843.59","9","26","Online
Banking","Resolved"],
["19","15","42","M","Divorced","Medium","0","1914.02","8","44","Online
Banking","Unresolved"],
["20","15","42","M","Divorced","Medium","0","1914.02","8","44","Online
Banking","Unresolved"],
["21","16","30","F","Single","Medium","1","2019.79","9","41","Online
Banking","Resolved"],
["22","16","30","F","Single","Medium","1","2019.79","9","41","Online
Banking","Resolved"],
["23","17","19","M","Married","Low","0","1625.27","7","8","Online
Banking","Unresolved"],
["24","17","19","M","Married","Low","0","1625.27","7","8","Online
Banking","Resolved"],
["25","18","56","F","Married","High","1","951.3800000000001","5","37",
"Website","Unresolved"],
["26","19","57","F","Divorced","High","0","1017.26","5","39","Online
Banking","Unresolved"],
["27","19","57","F","Divorced","High","0","1017.26","5","39","Online
Banking","Unresolved"],
["28","20","41","M","Single","Low","0","2775.85","9","16","Website","R
esolved"],
["29","21","64","M","Divorced","High","0","913.02","5","18","Website",
null],
["30","22","42","F","Divorced","Low","0","1582.73","6","42","Mobile
App","Resolved"],
["31","22","42","F","Divorced","Low","0","1582.73","6","42","Mobile
App","Resolved"],
["32","23","35","F","Divorced","Low","0","1490.74","4","19","Mobile
App","Unresolved"],
["33","23","35","F","Divorced","Low","0","1490.74","4","19","Mobile
App","Resolved"],
["34","24","55","M","Widowed","Medium","0","1143.22","5","12","Mobile
App","Resolved"],
["35","25","43","F","Married","Low","0","764.77","3","27","Mobile
App","Resolved"],
["36","26","31","M","Single","Medium","1","571.76","3","45","Mobile
App","Resolved"],
["37","27","26","M","Widowed","Medium","0","2210.6","8","22","Mobile
App","Unresolved"],
["38","28","27","F","Widowed","High","0","1702.26","6","42","Online
Banking","Unresolved"],
["39","28","27","F","Widowed","High","0","1702.26","6","42","Online
Banking","Unresolved"],
["40","29","38","F","Divorced","Medium","0","853.0699999999999","3","3
4","Website","Resolved"],
["41","29","38","F","Divorced","Medium","0","853.0699999999999","3","3
4","Website","Unresolved"],
["42","30","69","M","Divorced","High","0","706.66","2","38","Website",
"Resolved"],
["43","30","69","M","Divorced","High","0","706.66","2","38","Website",
"Unresolved"],
["44","31","34","M","Divorced","Low","0","1108.56","5","23","Online
Banking","Resolved"],
["45","31","34","M","Divorced","Low","0","1108.56","5","23","Online
Banking","Resolved"],
["46","32","69","F","Married","High","0","1225.52","5","48","Mobile
App",null],
["47","33","23","M","Divorced","High","1","689.0","2","24","Mobile
App","Resolved"],
["48","34","33","M","Widowed","High","0","1886.4","9","44","Online
Banking","Resolved"],
["49","34","33","M","Widowed","High","0","1886.4","9","44","Online
Banking","Unresolved"]],"shape":{"columns":11,"rows":1334}}
```

```
# we will fill the missing values of resolution status  with no
problem
df_final['ResolutionStatus'].fillna('No Problem', inplace=True)
```

```
C:\Users\Ayush\AppData\Local\Temp\ipykernel_3312\983418897.py:2:
FutureWarning: A value is trying to be set on a copy of a DataFrame or
```

```
Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never
work because the intermediate object on which we are setting values
always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try
using 'df.method({col: value}, inplace=True)' or df[col] =
df[col].method(value) instead, to perform the operation inplace on the
original object.


  df_final['ResolutionStatus'].fillna('No Problem', inplace=True)
```

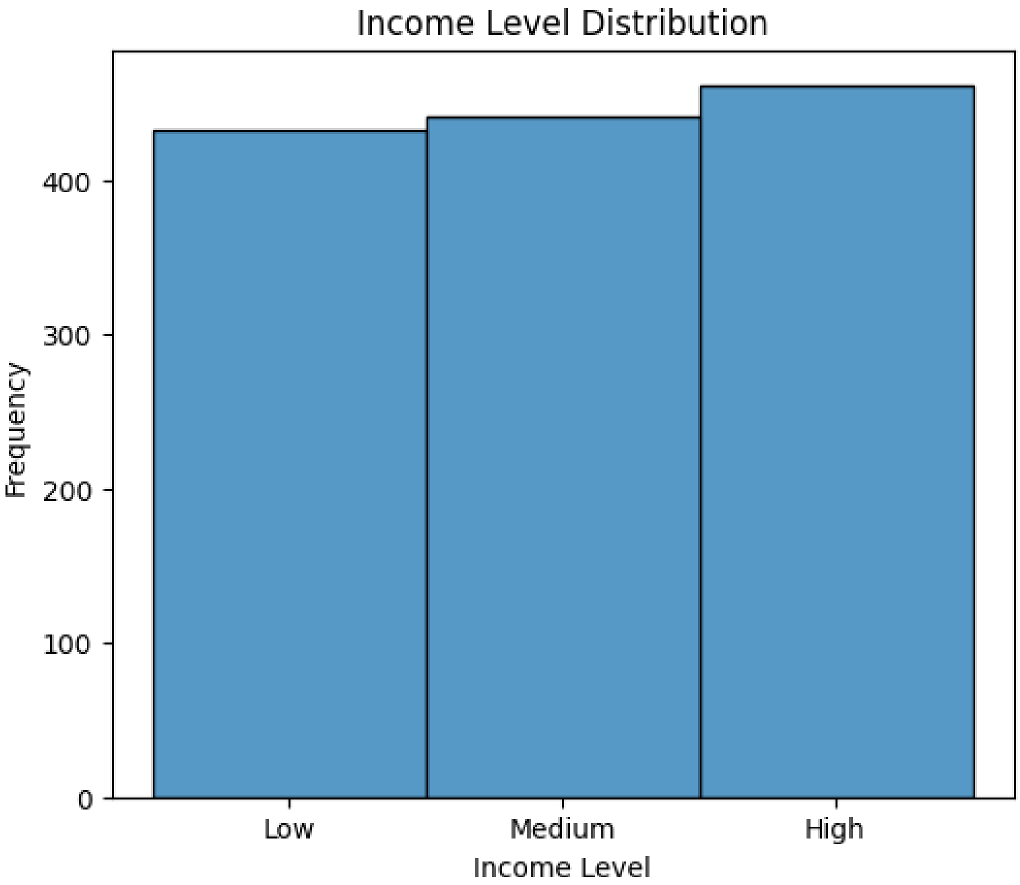
```
# now we will replace resolved and noproblem
```

EDA

```
df_final['CustomerID'].nunique()
```

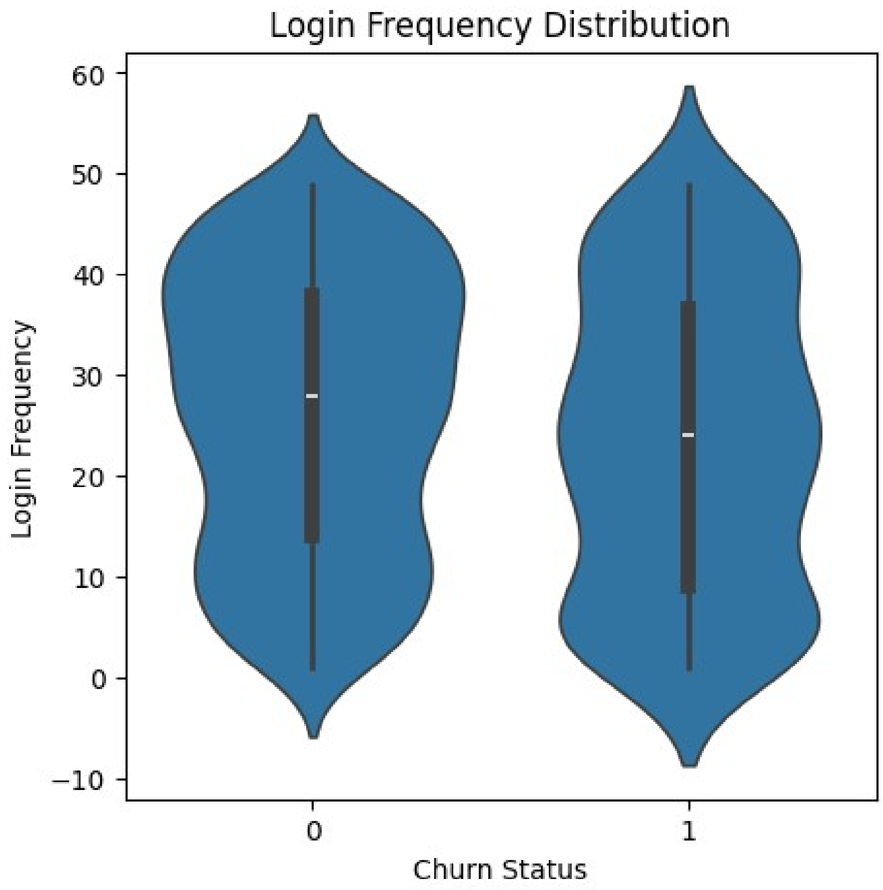
```
1000
```

```
# histogram for income column
plt.figure(figsize=(6 , 5 ))
sns.histplot(df_final['IncomeLevel'] )
plt.title('Income Level Distribution')
plt.xlabel('Income Level')
plt.ylabel('Frequency')
plt.show()
print( df_final['IncomeLevel'].value_counts())
```

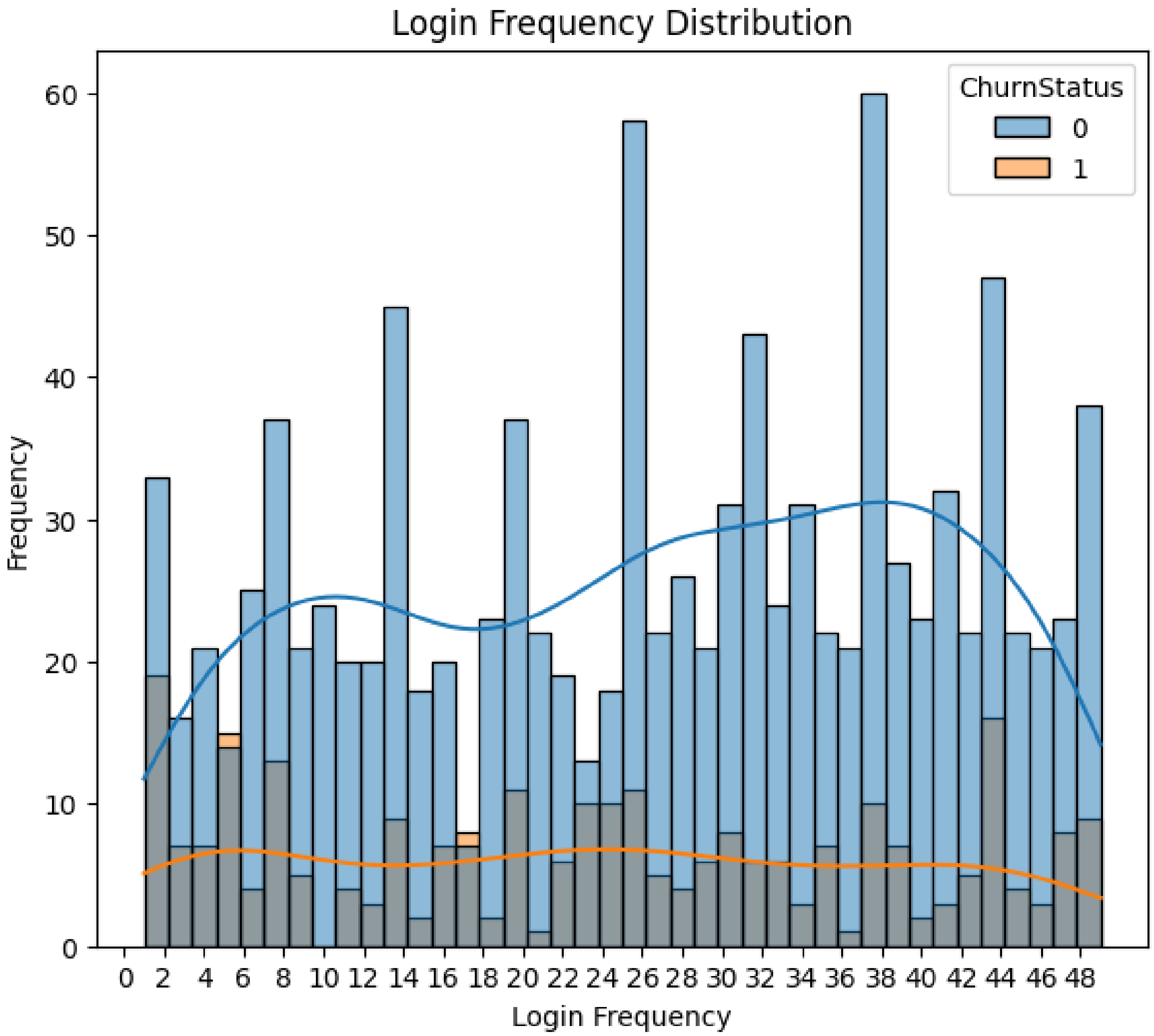


```
IncomeLevel
High      461
Medium    441
Low       432
Name: count, dtype: int64
```

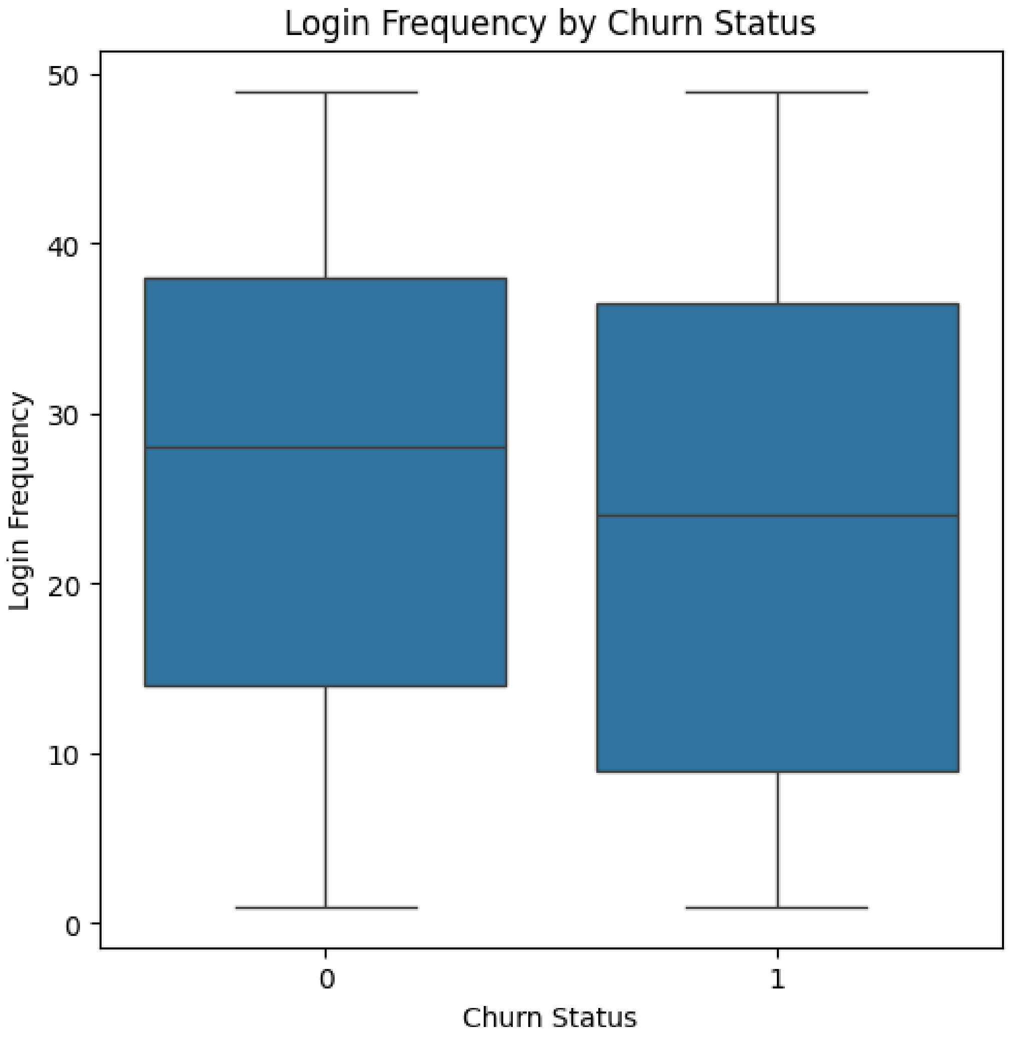
```
# violin plot for LoginFrequency
plt.figure(figsize=(5, 5))
sns.violinplot(x='ChurnStatus' , y = 'LoginFrequency'  , data=df_final
)
plt.title('Login Frequency Distribution')
plt.xlabel('Churn Status')
plt.ylabel('Login Frequency')
plt.show()
```



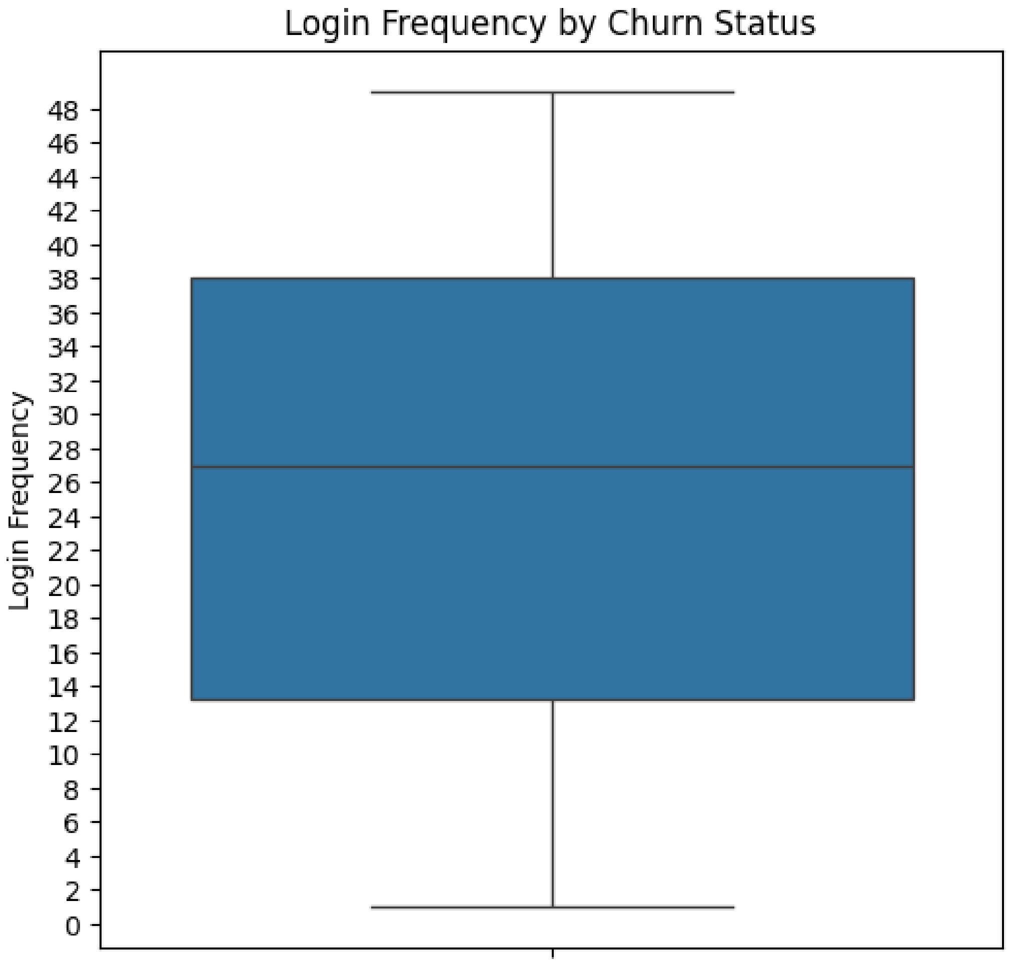
```
# histogram for login frequqncy
plt.figure(figsize=(7, 6))
sns.histplot( data = df_final , x = 'LoginFrequency' ,  bins=40 ,
kde=True , hue = 'ChurnStatus' )
plt.title('Login Frequency Distribution')
plt.xlabel('Login Frequency')
plt.ylabel('Frequency')
plt.xticks( np.arange( 0 ,50 , 2 )   )
plt.show()
```



```
plt.figure(figsize=(6, 6))
sns.boxplot(data=df_final, x='ChurnStatus', y='LoginFrequency')
plt.title('Login Frequency by Churn Status')
plt.xlabel('Churn Status')
plt.ylabel('Login Frequency')
plt.show()

# the people who are not leaving ar have higher login frequnce on an
average
```

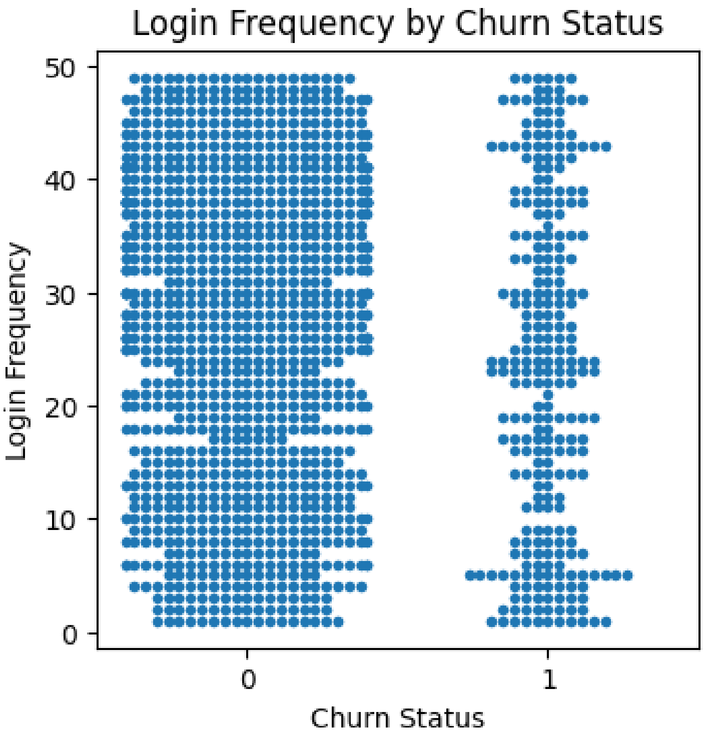


```
plt.figure(figsize=(6, 6))
sns.boxplot(data=df_final, y = 'LoginFrequency')
plt.title('Login Frequency by Churn Status')
plt.ylabel('Login Frequency')
plt.yticks( np.arange( 0 ,50 , 2 )   )
plt.show()
```



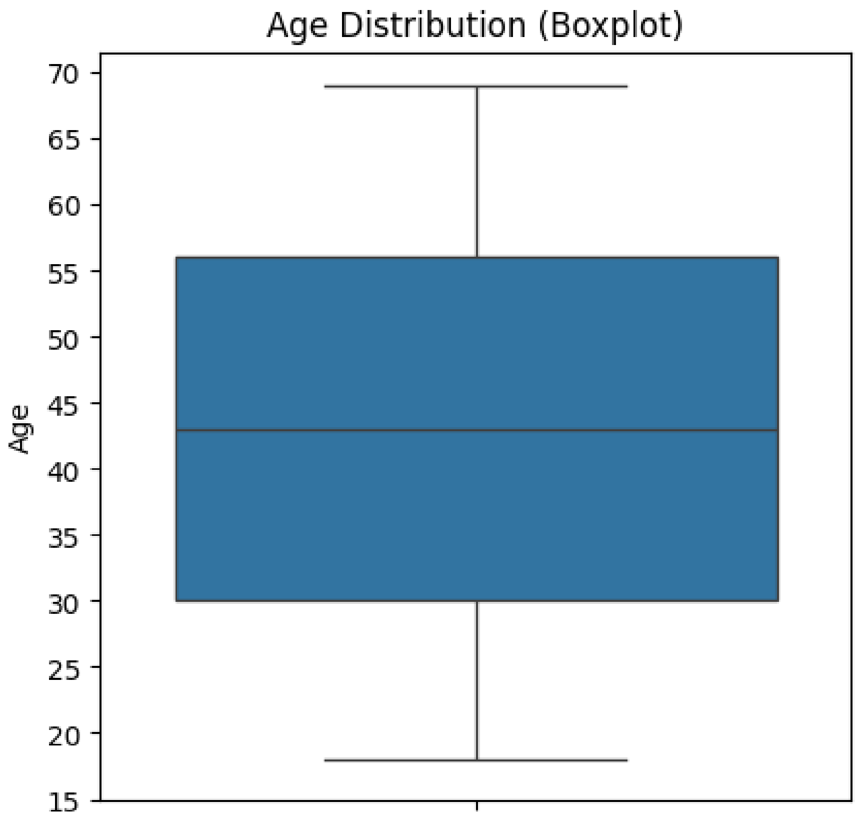
```
plt.figure(figsize=(4, 4))
sns.swarmplot(data=df_final, x='ChurnStatus', y='LoginFrequency' ,
size =4  )
plt.title('Login Frequency by Churn Status')
plt.xlabel('Churn Status')
plt.ylabel('Login Frequency')
plt.show()
# the people who are  leaving have lower  login frequency
```

```
e:\DATA\quantinum forage\.venv\Lib\site-packages\seaborn\
categorical.py:3399: UserWarning: 11.0% of the points cannot be
placed; you may want to decrease the size of the markers or use
stripplot.
  warnings.warn(msg, UserWarning)
```



```
# box plot and histogram for age
q1 = df_final['Age'].quantile(0.25)
q2 = df_final['Age'].quantile(0.50)
q3 = df_final['Age'].quantile(0.75)
print( "25 % quartile :" , q1 , "       50 % quartile :"  , q2 ,  "
75 % quartile:"  , q3 )

plt.figure(figsize=(5, 5))
sns.boxplot(data=df_final, y='Age')
plt.title('Age Distribution (Boxplot)')
plt.ylabel('Age')
plt.yticks( np.arange( 15, 75 , 5  ) )
plt.show()
```

```
25 % quartile : 30.0       50 % quartile : 43.0         75 % quartile:
56.0
```

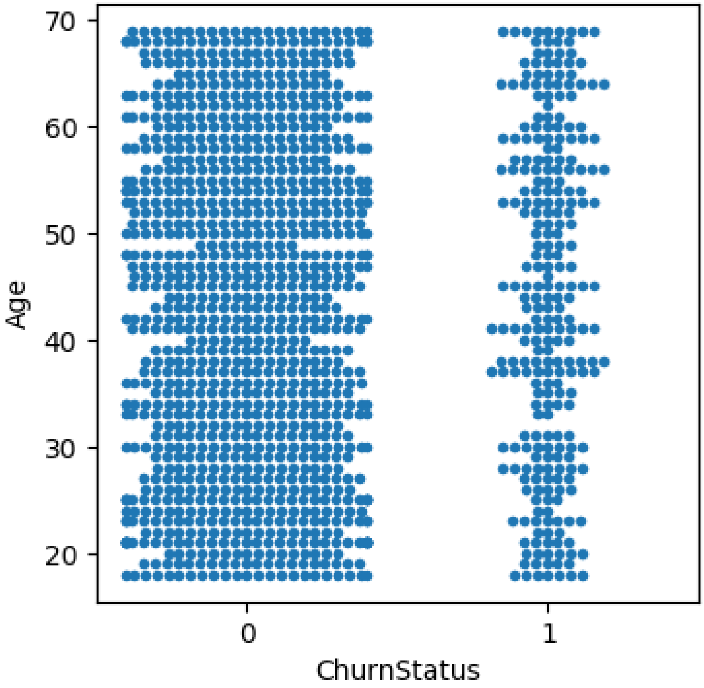


bi variate analysis using the target variable

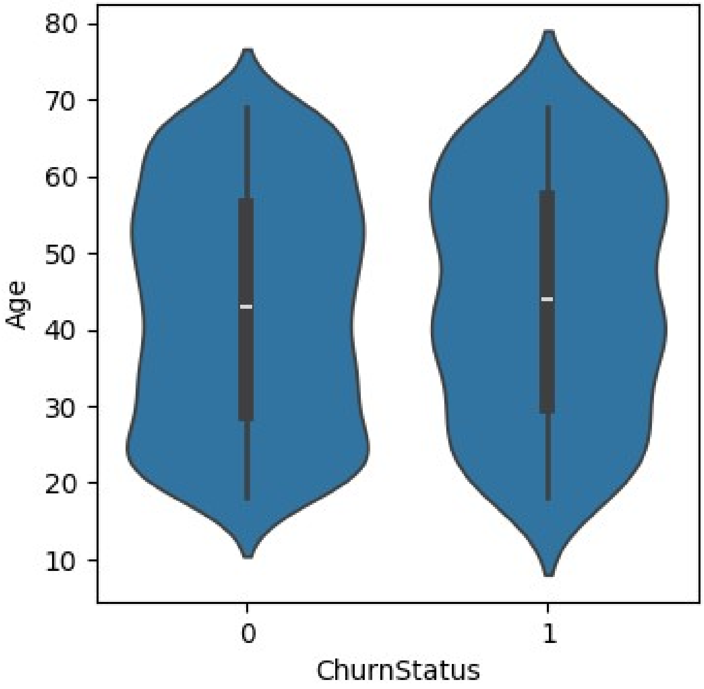
```
# swarm plot for churn vs age
plt.figure( figsize= ( 4 ,4 ))
sns.swarmplot( data = df_final , x = 'ChurnStatus'  , y = 'Age'  ,
size = 4 )
```

```
<Axes: xlabel='ChurnStatus', ylabel='Age'>
```

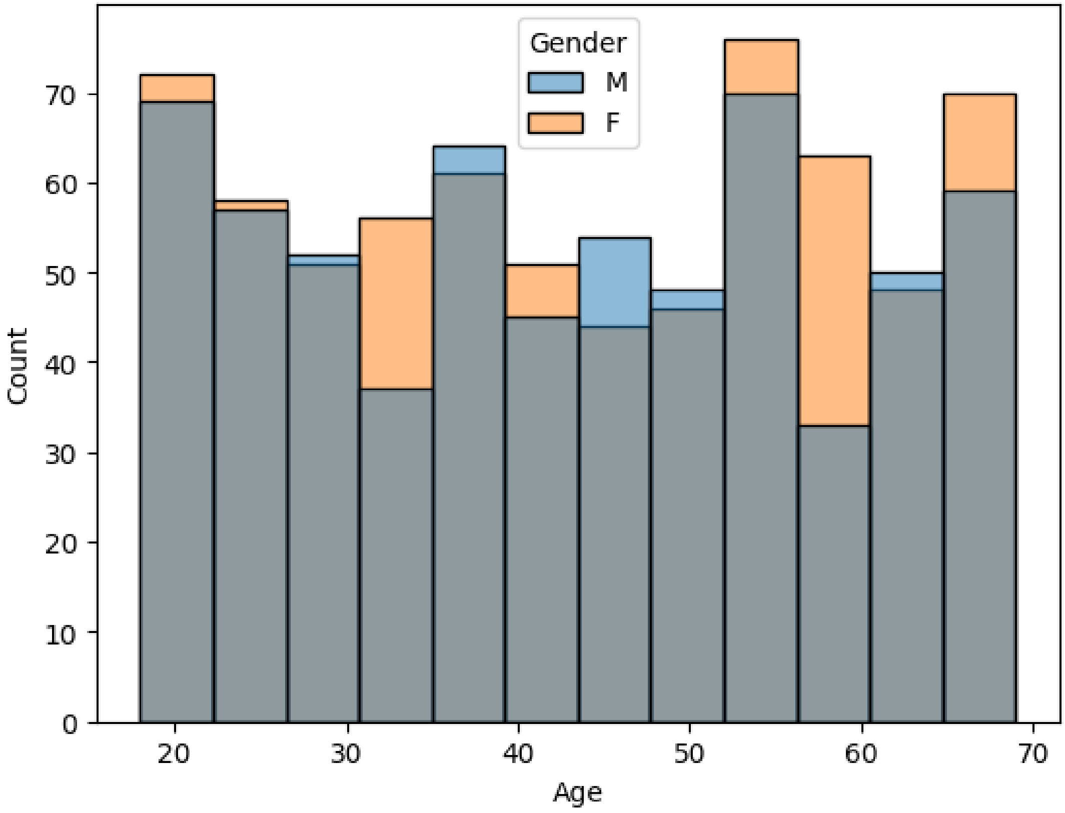
```
e:\DATA\quantinum forage\.venv\Lib\site-packages\seaborn\
categorical.py:3399: UserWarning: 7.3% of the points cannot be placed;
you may want to decrease the size of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
```



```
plt.figure(  figsize = ( 4 ,4 ) )
sns.violinplot(  x = 'ChurnStatus' , y = 'Age' , data = df_final )

<Axes: xlabel='ChurnStatus', ylabel='Age'>
```

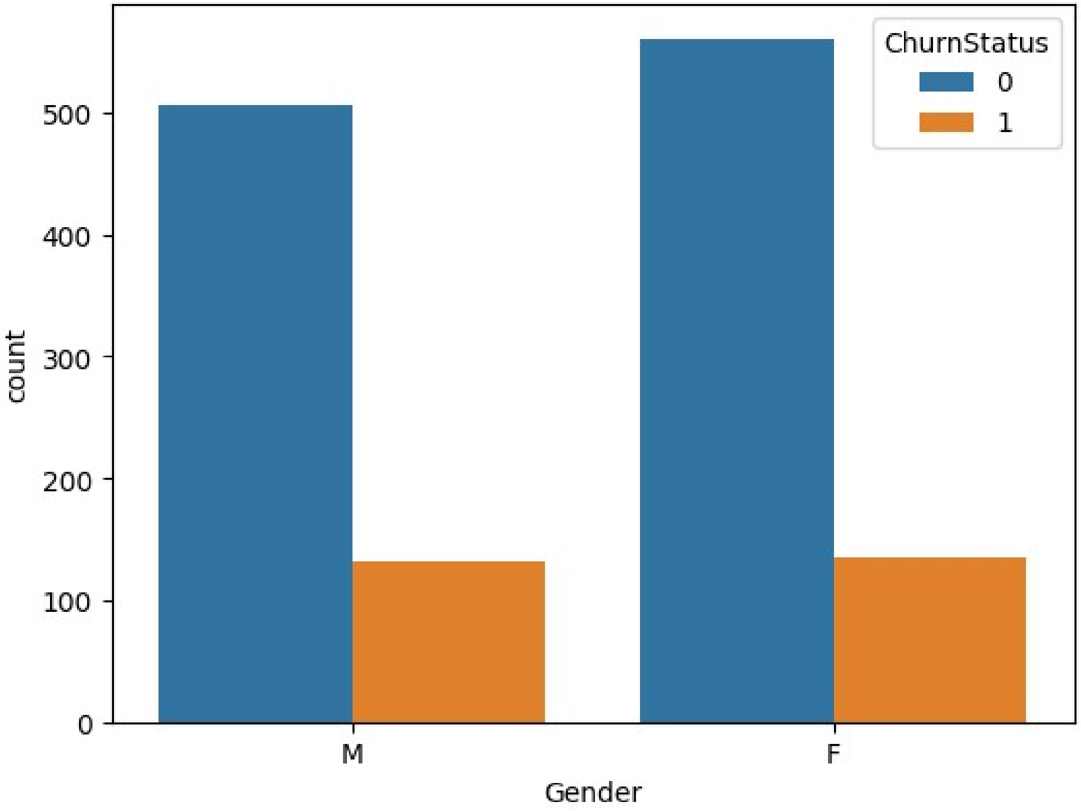


```
plt.Figure( figsize= ( 4 , 4 ))
sns.histplot(data=df_final , x = 'Age' , hue = 'Gender' )  # split by
category (optional)

<Axes: xlabel='Age', ylabel='Count'>
```



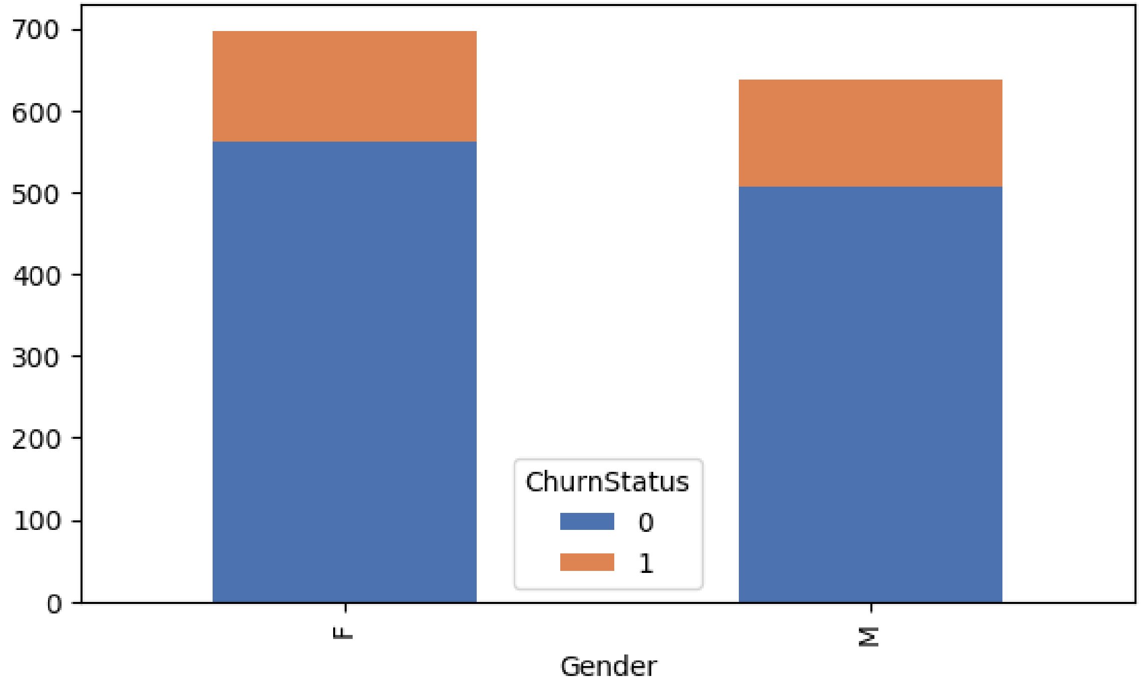
categorical vs target variable

```
plt.Figure( figsize= ( 4 , 4 ))
sns.countplot( data = df_final  , x = 'Gender'  , hue = 'ChurnStatus'
)

<Axes: xlabel='Gender', ylabel='count'>
```



```
counts = pd.crosstab(df_final['Gender'], df_final['ChurnStatus'])

# Plot stacked bar chart (normal, not 100%)
counts.plot(kind='bar', stacked=True, figsize=(7,4), color=['#4C72B0',
'#DD8452'])
```

```
<Axes: xlabel='Gender'>
```



```
## 100 percent stacked bar chart for marital status  vs churnstatus
crosstab = pd.crosstab( df_final['MaritalStatus'] ,
df_final['ChurnStatus'] )
crosstab2 = pd.crosstab( df_final['ResolutionStatus'] ,
df_final['ChurnStatus'] )
crosstab3 = pd.crosstab( df_final['IncomeLevel'] ,
df_final['ChurnStatus'] )


crosstab2
```

{"columns":
[{"name":"ResolutionStatus","rawType":"object","type":"string"},
{"name":"0","rawType":"int64","type":"integer"},
{"name":"1","rawType":"int64","type":"integer"}],"ref":"ee1981b9-da0f-4b4e-b7b2-d2fe7f57f5b5","rows":[["No Problem","271","61"],
["Resolved","416","107"],["Unresolved","380","99"]],"shape":
{"columns":2,"rows":3}}

```
proportion = crosstab.div( crosstab.sum(axis =1 )   , axis = 0 )
proportion2 = crosstab2.div( crosstab2.sum(axis =1 )   , axis = 0 )
proportion3 = crosstab3.div( crosstab3.sum(axis =1 )   , axis = 0 )

proportion3
```

{"columns":[{"name":"IncomeLevel","rawType":"object","type":"string"},
{"name":"0","rawType":"float64","type":"float"},
{"name":"1","rawType":"float64","type":"float"}],"ref":"6c4ca22d-8f45-4de6-9162-ee057d5331d7","rows":

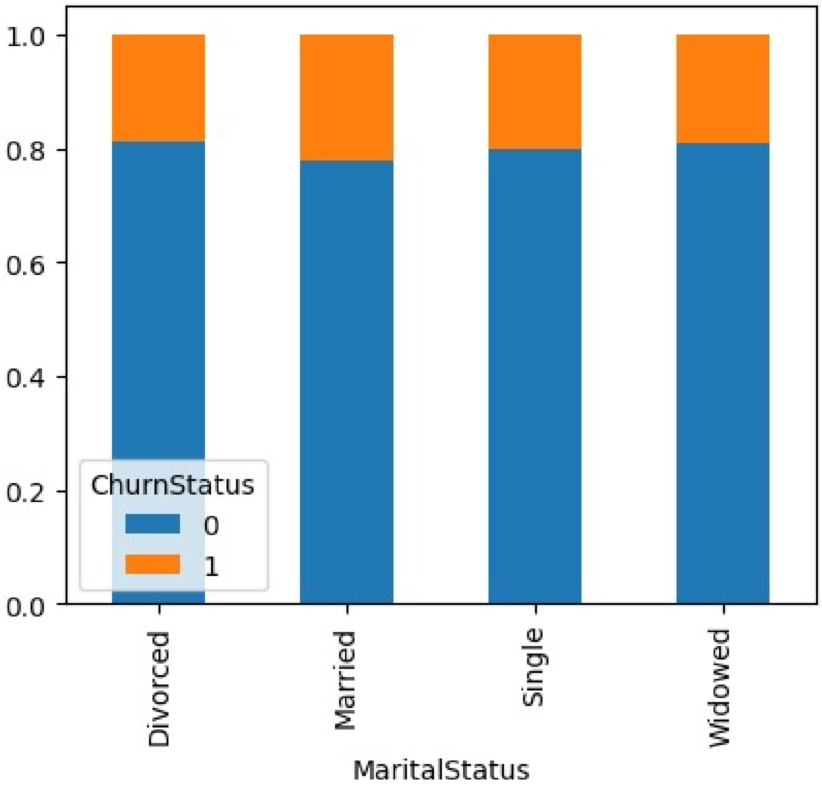
```
[["High","0.8156182212581344","0.1843817787418655"],
["Low","0.7870370370370371","0.21296296296296297"],
["Medium","0.7959183673469388","0.20408163265306123"]],"shape":
{"columns":2,"rows":3}}
```

```
proportion2
```

```
{"columns":
[{"name":"ResolutionStatus","rawType":"object","type":"string"},
{"name":"0","rawType":"float64","type":"float"},
{"name":"1","rawType":"float64","type":"float"}],"ref":"f57b98dd-53a5-
49a8-8112-cf99c3f10d6f","rows":[["No
Problem","0.8162650602409639","0.18373493975903615"],
["Resolved","0.7954110898661568","0.2045889101338432"],
["Unresolved","0.7933194154488518","0.20668058455114824"]],"shape":
{"columns":2,"rows":3}}
```

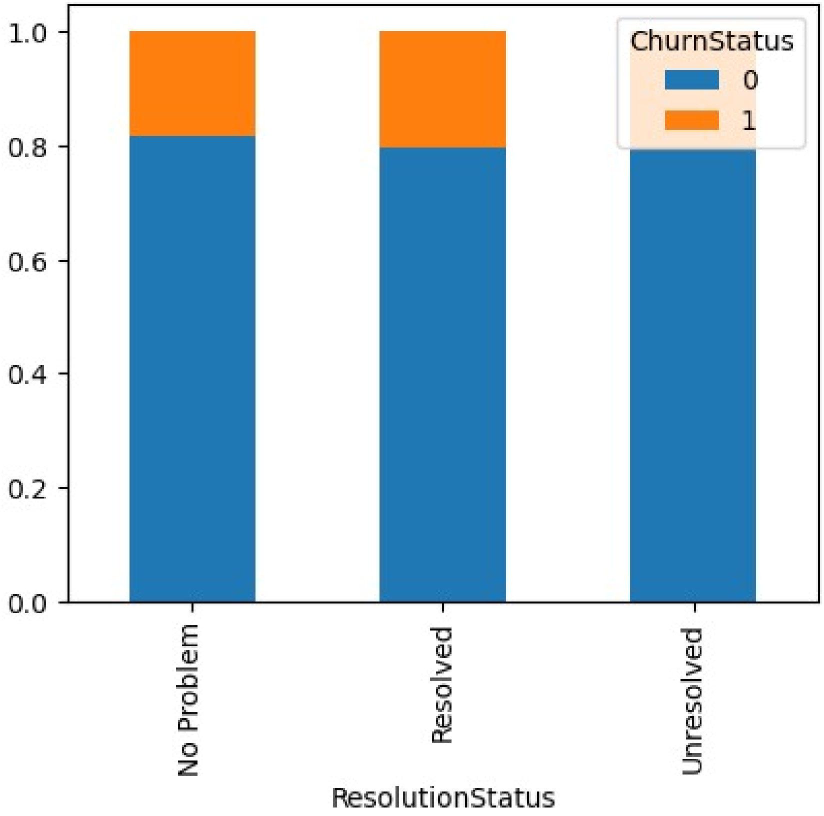
```
proportion.plot( kind = 'bar'  , stacked= True  ,  figsize=( 5, 4 ) )

# the churn rate across all of them is almost equal nothing to find
out from  here
```

```
<Axes: xlabel='MaritalStatus'>
```

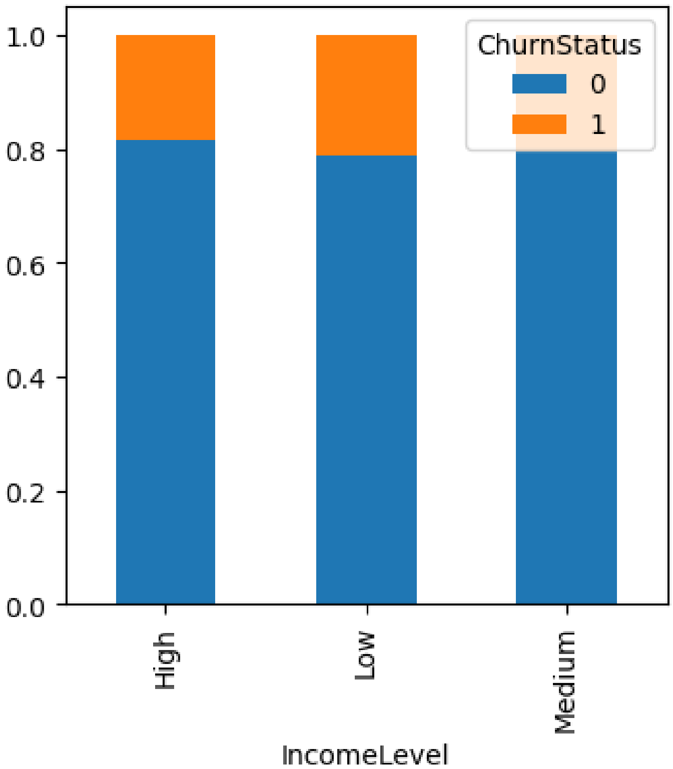


```
proportion2.plot( kind = 'bar'  , stacked= True  ,  figsize=( 5, 4 ) )
```

```
<Axes: xlabel='ResolutionStatus'>
```



```
proportion3.plot( kind= 'bar'  , stacked = True  , figsize = ( 4 , 4 )
)

<Axes: xlabel='IncomeLevel'>
```



numerical vs target

```
df_final
```

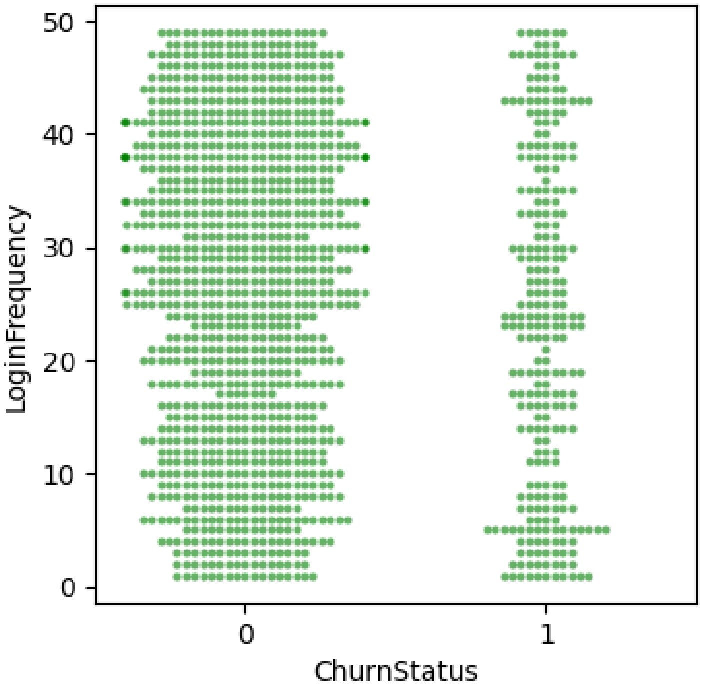
```
{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"Age","rawType":"int64","type":"integer"},
{"name":"Gender","rawType":"object","type":"string"},
{"name":"MaritalStatus","rawType":"object","type":"string"},
{"name":"IncomeLevel","rawType":"object","type":"string"},
{"name":"ChurnStatus","rawType":"int64","type":"integer"},
{"name":"total_amount","rawType":"float64","type":"float"},
{"name":"countoftransaction","rawType":"int64","type":"integer"},
{"name":"LoginFrequency","rawType":"int64","type":"integer"},
{"name":"ServiceUsage","rawType":"object","type":"string"},
{"name":"ResolutionStatus","rawType":"object","type":"string"}],"ref":
"97d68717-7adc-454c-a147-1466423583b3","rows":
[["0","1","62","M","Single","Low","0","416.5","1","34","Mobile
App","Resolved"],
["1","2","65","M","Married","Low","1","1547.4199999999998","7","5","We
bsite","Resolved"],
["2","3","18","M","Single","Low","0","1702.98","6","3","Website","Reso
lved"],
["3","4","21","M","Widowed","Low","0","917.29","5","2","Website","Reso
lved"],
["4","4","21","M","Widowed","Low","0","917.29","5","2","Website","Unre
solved"],
["5","5","21","M","Divorced","Medium","0","2001.49","8","41","Website"
,"No Problem"],
["6","6","57","F","Divorced","Medium","0","1164.29","5","2","Website",
"Resolved"],
["7","7","27","F","Married","High","0","86.73","1","32","Mobile
App","No Problem"],
["8","8","37","M","Single","Low","1","2046.88","7","17","Online
Banking","Unresolved"],
["9","8","37","M","Single","Low","1","2046.88","7","17","Online
Banking","Unresolved"],
["10","9","39","M","Divorced","High","0","1317.24","5","24","Website",
"Resolved"],
["11","10","68","M","Married","High","1","1397.36","7","29","Online
Banking","No Problem"],
["12","11","54","M","Divorced","Medium","0","1733.83","5","30","Online
Banking","Resolved"],
["13","11","54","M","Divorced","Medium","0","1733.83","5","30","Online
Banking","Unresolved"],
["14","12","41","F","Married","Low","0","87.1","1","43","Website","Res
olved"],
["15","12","41","F","Married","Low","0","87.1","1","43","Website","Unr
esolved"],
["16","13","24","F","Divorced","Low","0","392.02","3","10","Mobile
App","Resolved"],
["17","14","42","M","Widowed","Medium","0","2843.59","9","26","Online
Banking","Unresolved"],
["18","14","42","M","Widowed","Medium","0","2843.59","9","26","Online
Banking","Resolved"],
["19","15","42","M","Divorced","Medium","0","1914.02","8","44","Online
Banking","Unresolved"],
["20","15","42","M","Divorced","Medium","0","1914.02","8","44","Online
Banking","Unresolved"],
["21","16","30","F","Single","Medium","1","2019.79","9","41","Online
Banking","Resolved"],
["22","16","30","F","Single","Medium","1","2019.79","9","41","Online
Banking","Resolved"],
["23","17","19","M","Married","Low","0","1625.27","7","8","Online
Banking","Unresolved"],
["24","17","19","M","Married","Low","0","1625.27","7","8","Online
Banking","Resolved"],
["25","18","56","F","Married","High","1","951.3800000000001","5","37",
"Website","Unresolved"],
["26","19","57","F","Divorced","High","0","1017.26","5","39","Online
Banking","Unresolved"],
["27","19","57","F","Divorced","High","0","1017.26","5","39","Online
Banking","Unresolved"],
["28","20","41","M","Single","Low","0","2775.85","9","16","Website","R
esolved"],
["29","21","64","M","Divorced","High","0","913.02","5","18","Website",
"No Problem"],
["30","22","42","F","Divorced","Low","0","1582.73","6","42","Mobile
App","Resolved"],
["31","22","42","F","Divorced","Low","0","1582.73","6","42","Mobile
App","Resolved"],
["32","23","35","F","Divorced","Low","0","1490.74","4","19","Mobile
App","Unresolved"],
["33","23","35","F","Divorced","Low","0","1490.74","4","19","Mobile
App","Resolved"],
["34","24","55","M","Widowed","Medium","0","1143.22","5","12","Mobile
App","Resolved"],
["35","25","43","F","Married","Low","0","764.77","3","27","Mobile
App","Resolved"],
["36","26","31","M","Single","Medium","1","571.76","3","45","Mobile
App","Resolved"],
["37","27","26","M","Widowed","Medium","0","2210.6","8","22","Mobile
App","Unresolved"],
["38","28","27","F","Widowed","High","0","1702.26","6","42","Online
Banking","Unresolved"],
["39","28","27","F","Widowed","High","0","1702.26","6","42","Online
Banking","Unresolved"],
["40","29","38","F","Divorced","Medium","0","853.0699999999999","3","3
4","Website","Resolved"],
["41","29","38","F","Divorced","Medium","0","853.0699999999999","3","3
4","Website","Unresolved"],
["42","30","69","M","Divorced","High","0","706.66","2","38","Website",
"Resolved"],
["43","30","69","M","Divorced","High","0","706.66","2","38","Website",
"Unresolved"],
["44","31","34","M","Divorced","Low","0","1108.56","5","23","Online
Banking","Resolved"],
["45","31","34","M","Divorced","Low","0","1108.56","5","23","Online
Banking","Resolved"],
["46","32","69","F","Married","High","0","1225.52","5","48","Mobile
App","No Problem"],
["47","33","23","M","Divorced","High","1","689.0","2","24","Mobile
App","Resolved"],
["48","34","33","M","Widowed","High","0","1886.4","9","44","Online
Banking","Resolved"],
["49","34","33","M","Widowed","High","0","1886.4","9","44","Online
Banking","Unresolved"]],"shape":{"columns":11,"rows":1334}}
```

```
# churn status  vs incoem level
df_final.columns
```

```
Index(['CustomerID', 'Age', 'Gender', 'MaritalStatus', 'IncomeLevel',
       'ChurnStatus', 'total_amount', 'countoftransaction',
'LoginFrequency',
       'ServiceUsage', 'ResolutionStatus'],
      dtype='object')
```

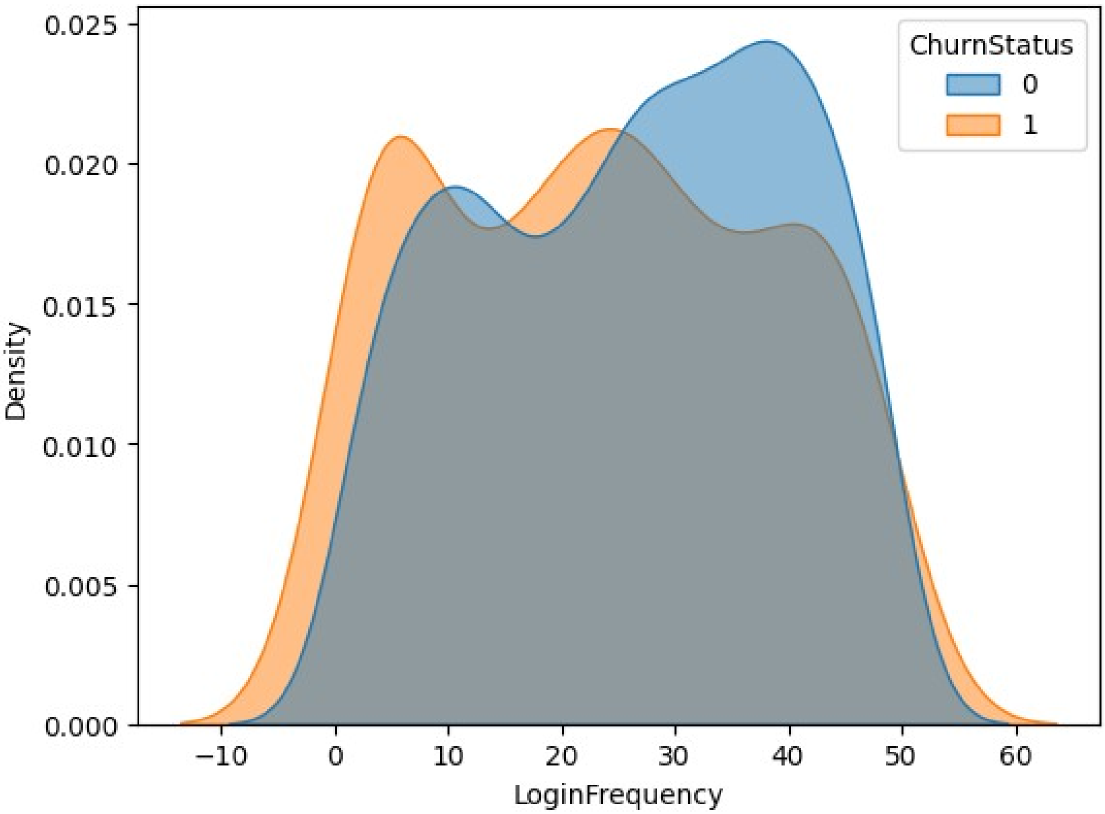
```
plt.figure( figsize= ( 4 , 4 ))
sns.swarmplot(x='ChurnStatus', y='LoginFrequency', data=df_final,
color='g', size=3, alpha=0.6)
# people with t=less login frequency are leaving
```

```
<Axes: xlabel='ChurnStatus', ylabel='LoginFrequency'>
```



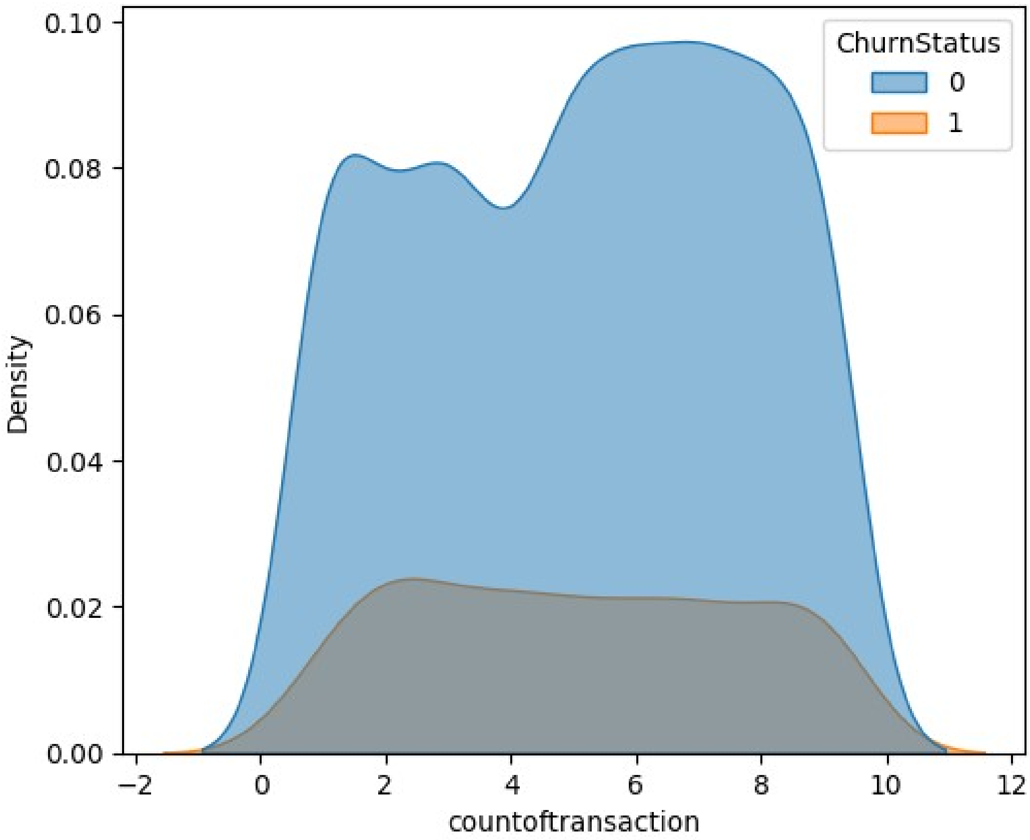
```
sns.kdeplot(data=df_final, x='LoginFrequency', hue='ChurnStatus',
fill=True, common_norm=False, alpha=0.5)
```

```
<Axes: xlabel='LoginFrequency', ylabel='Density'>
```

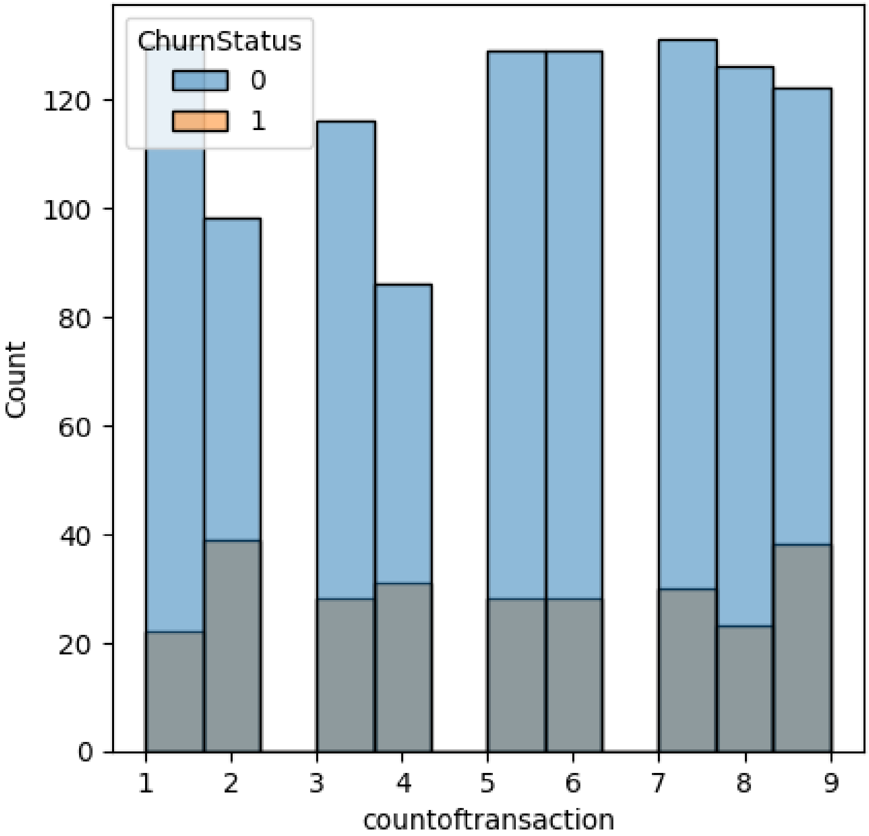


```
plt.figure( figsize = ( 6 , 5 ))
sns.kdeplot( data = df_final ,  x = 'countoftransaction'  , hue =
'ChurnStatus'  ,  fill = True , alpha = 0.5 )
```

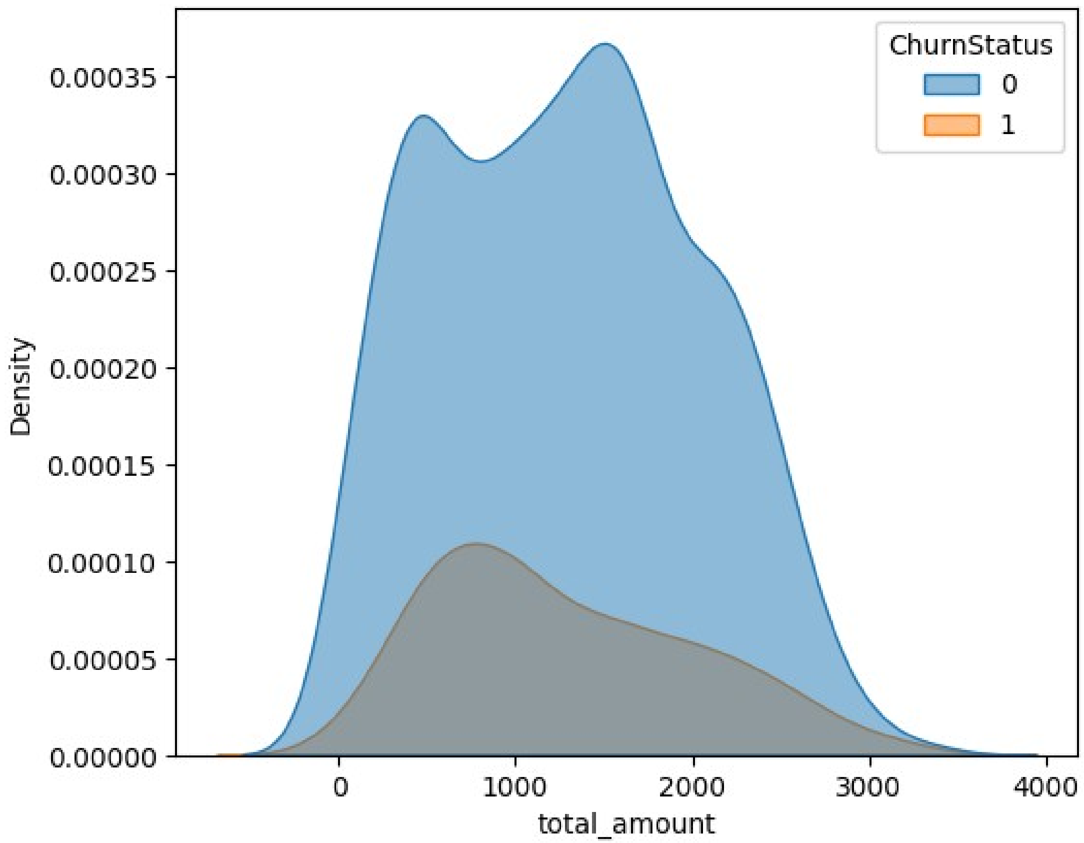
```
<Axes: xlabel='countoftransaction', ylabel='Density'>
```



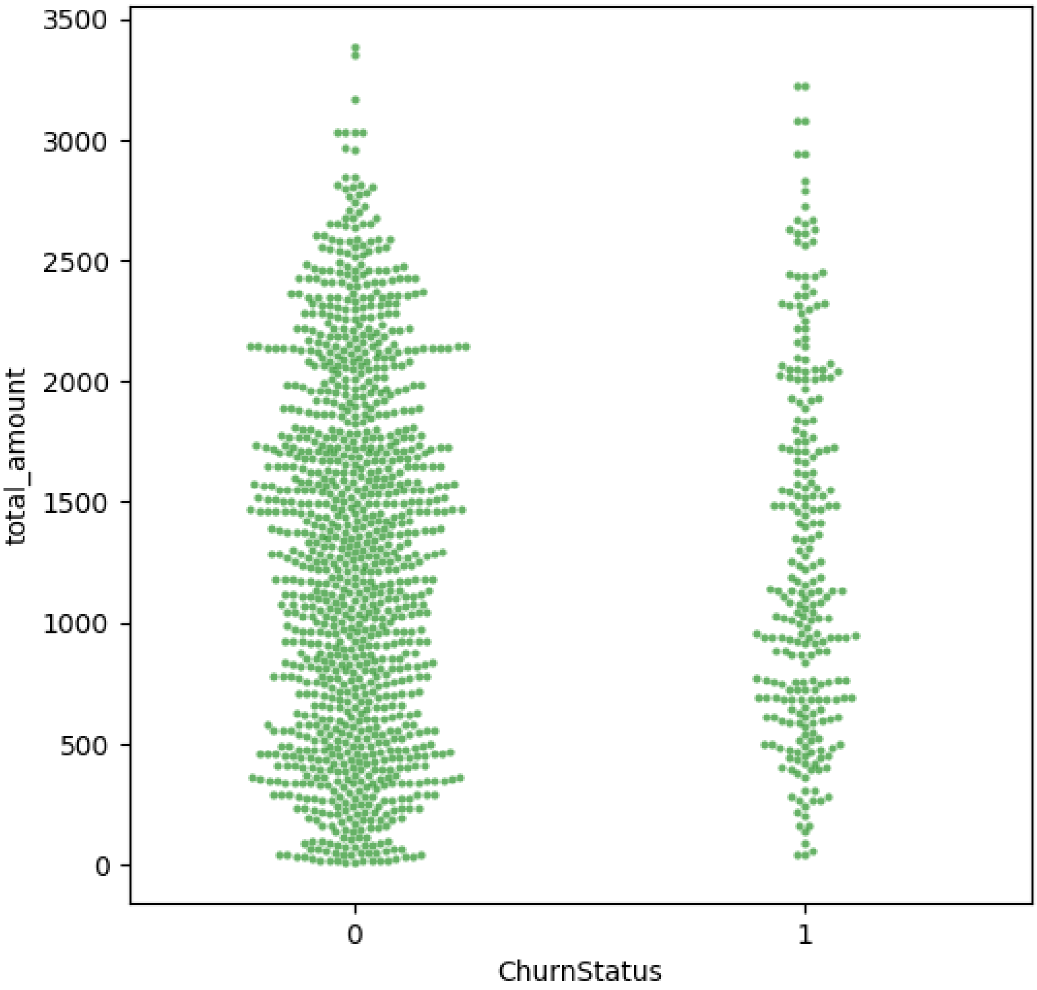
```
plt.figure( figsize = ( 5 , 5 ))
sns.histplot( data = df_final ,  x = 'countoftransaction'  , hue =
'ChurnStatus'  ,  fill = True , alpha = 0.5 )

<Axes: xlabel='countoftransaction', ylabel='Count'>
```



```
plt.figure( figsize = ( 6 , 5 ))
sns.kdeplot( data = df_final ,  x = 'total_amount'  , hue =
'ChurnStatus'  ,  fill = True , alpha = 0.5 )

<Axes: xlabel='total_amount', ylabel='Density'>
```

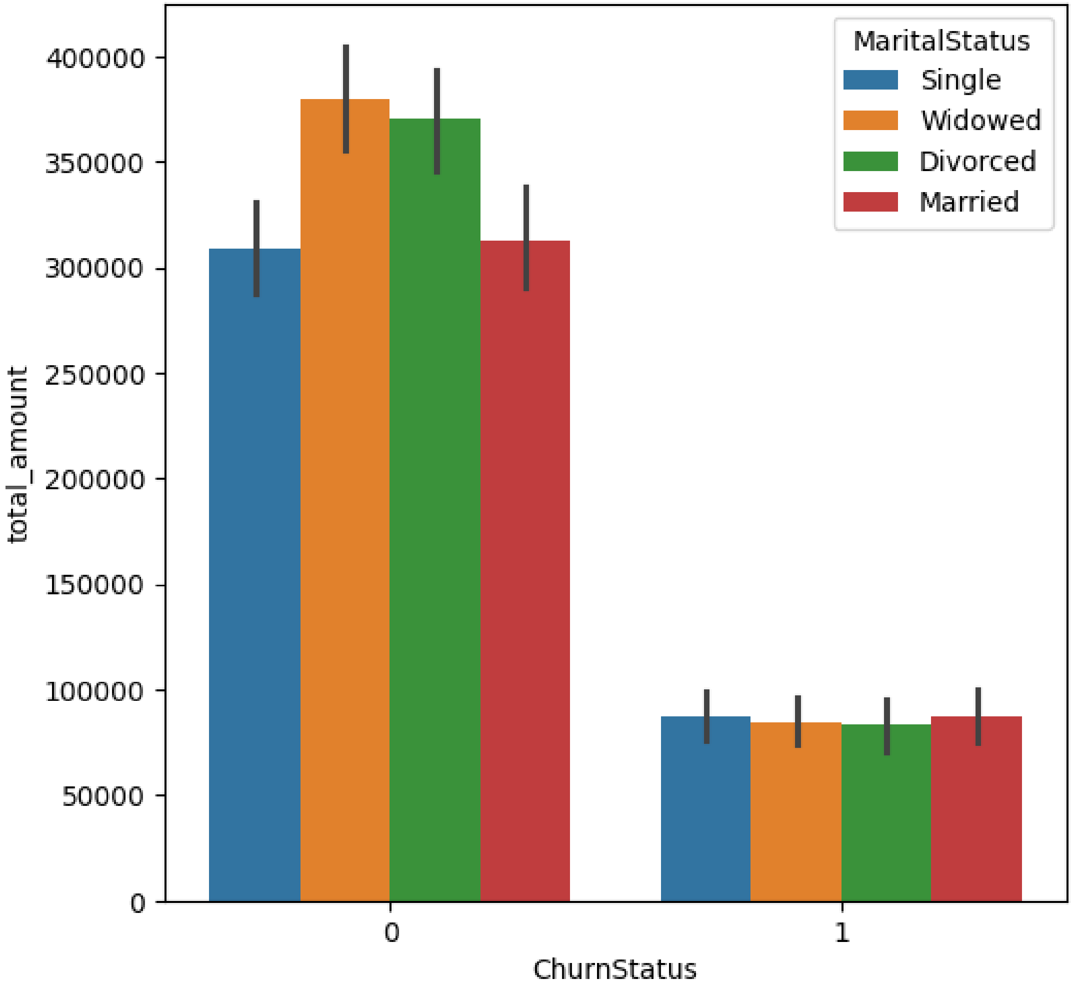


```
plt.figure( figsize= ( 6 , 6 ))
sns.swarmplot(x='ChurnStatus', y='total_amount', data=df_final,
color='g', size=3, alpha=0.6)

<Axes: xlabel='ChurnStatus', ylabel='total_amount'>
```



```
plt.figure( figsize= ( 6 , 6 ))
sns.barplot( data = df_final ,  x =  'ChurnStatus'  , y =
'total_amount'  , hue = 'MaritalStatus' , estimator = sum )
```

```
<Axes: xlabel='ChurnStatus', ylabel='total_amount'>
```

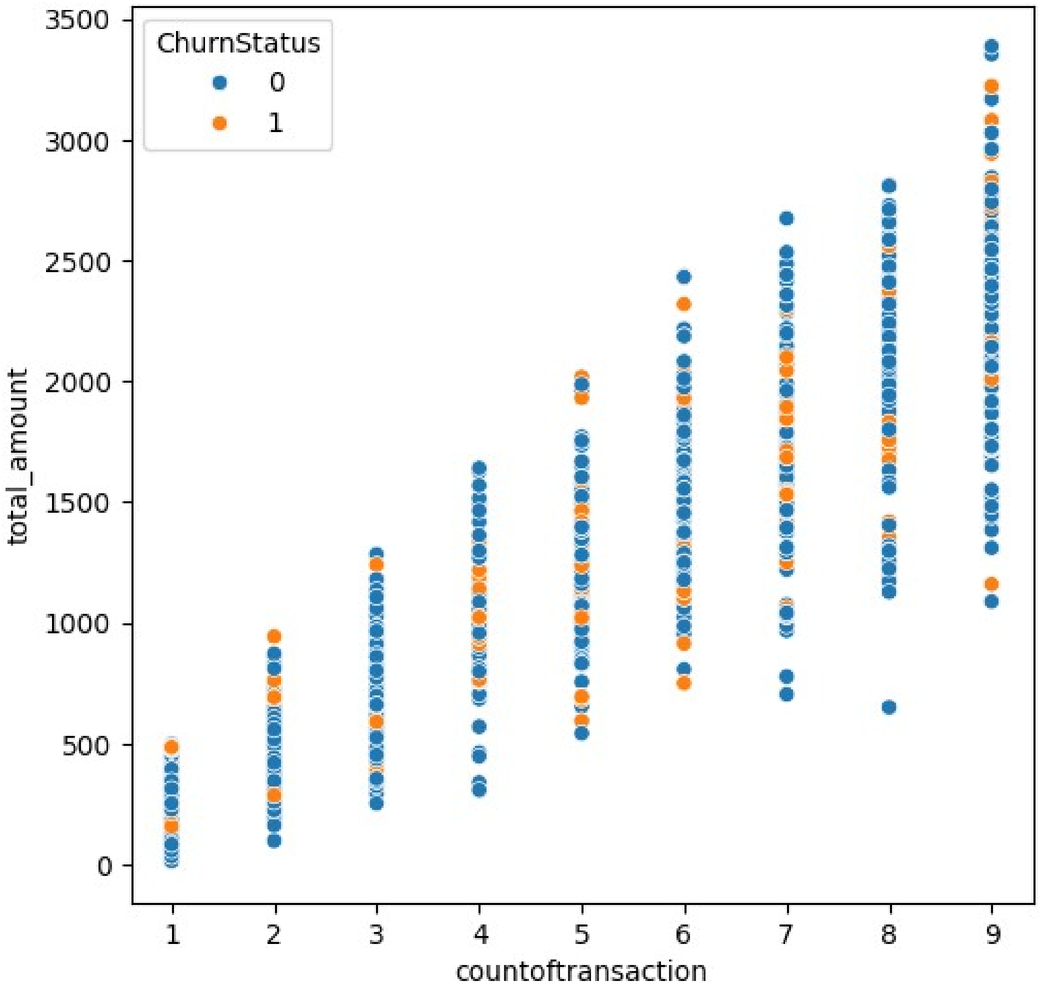


```
'''['CustomerID', 'Age', 'Gender', 'MaritalStatus', 'IncomeLevel',
       'ChurnStatus', 'total_amount', 'countoftransaction',
'LoginFrequency',
       'ServiceUsage', 'ResolutionStatus'] '''
```

```
"['CustomerID', 'Age', 'Gender', 'MaritalStatus', 'IncomeLevel',\n
'ChurnStatus', 'total_amount', 'countoftransaction',
'LoginFrequency',\n       'ServiceUsage', 'ResolutionStatus'] "
```

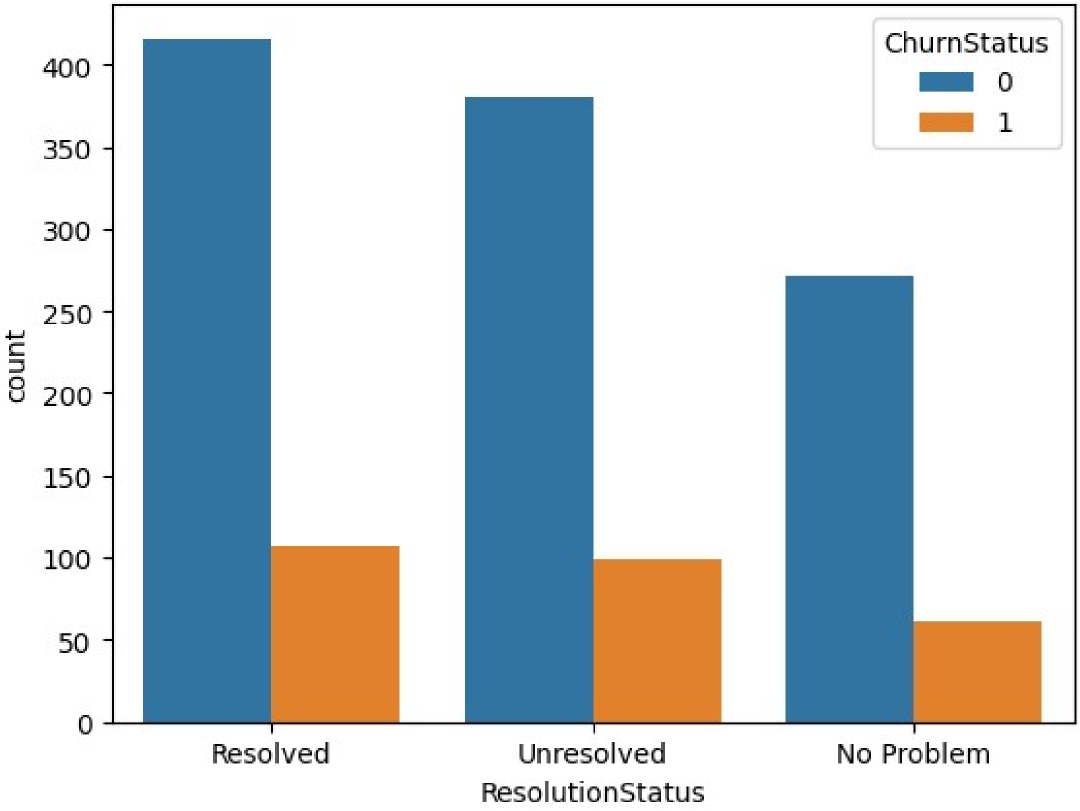
```
plt.figure( figsize= ( 6 , 6 ))
sns.scatterplot( data = df_final ,  x =  'countoftransaction' , y =
'total_amount'  , hue = 'ChurnStatus' )
```

```
<Axes: xlabel='countoftransaction', ylabel='total_amount'>
```



```
sns.countplot(data=df_final , x = 'ResolutionStatus',  hue =
'ChurnStatus' )
```

```
<Axes: xlabel='ResolutionStatus', ylabel='count'>
```



```
django = pd.crosstab( df_final['ResolutionStatus'] ,
df_final['ChurnStatus'] )

django
```

{"columns":
[{"name":"ResolutionStatus","rawType":"object","type":"string"},
{"name":"0","rawType":"int64","type":"integer"},
{"name":"1","rawType":"int64","type":"integer"}],"ref":"5f20da98-7336-4109-a3d8-587773110858","rows":[["No Problem","271","61"],
["Resolved","416","107"],["Unresolved","380","99"]],"shape":
{"columns":2,"rows":3}}

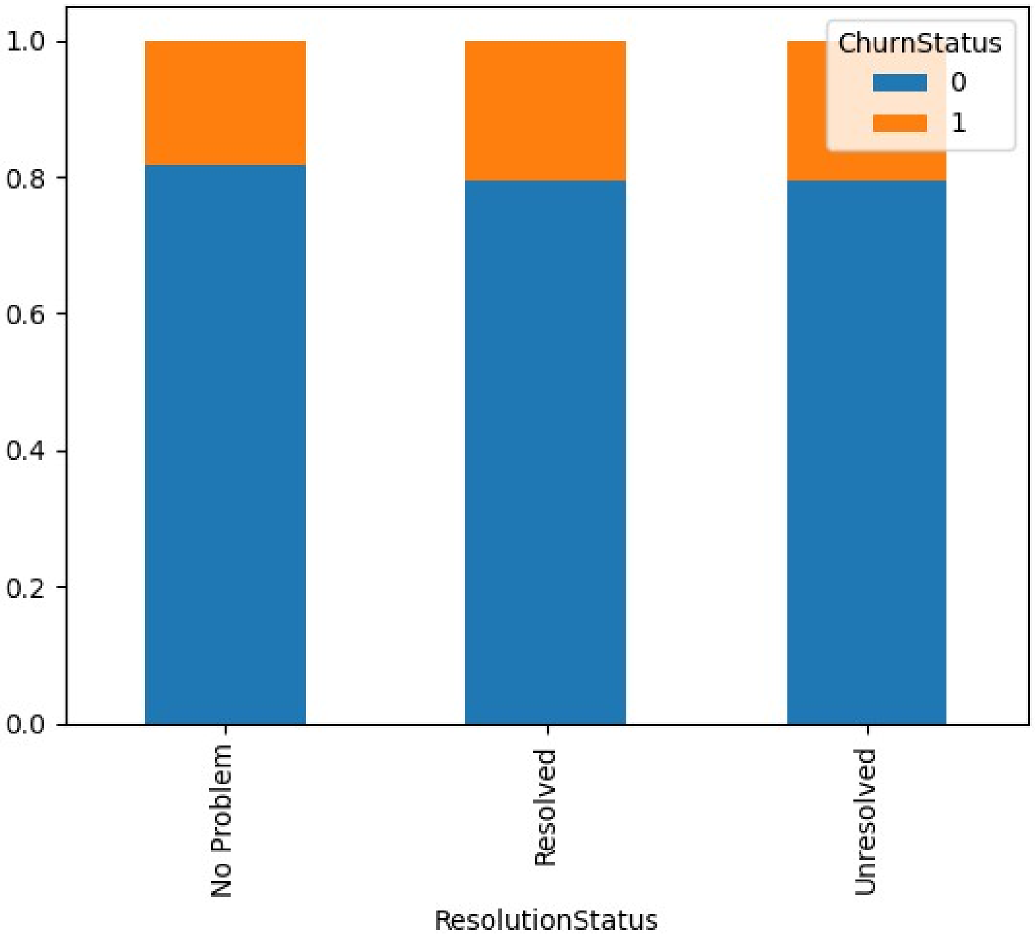
```
django_normal = django.div( django.sum(axis= 1 )  , axis = 0 )

django_normal
```

{"columns":
[{"name":"ResolutionStatus","rawType":"object","type":"string"},
{"name":"0","rawType":"float64","type":"float"},
{"name":"1","rawType":"float64","type":"float"}],"ref":"571962d8-4590-4020-88e9-554ef9ca1878","rows":[["No
Problem","0.8162650602409639","0.18373493975903615"],
["Resolved","0.7954110898661568","0.2045889101338432"],

```
["Unresolved","0.7933194154488518","0.20668058455114824"]],"shape":
{"columns":2,"rows":3}}
```

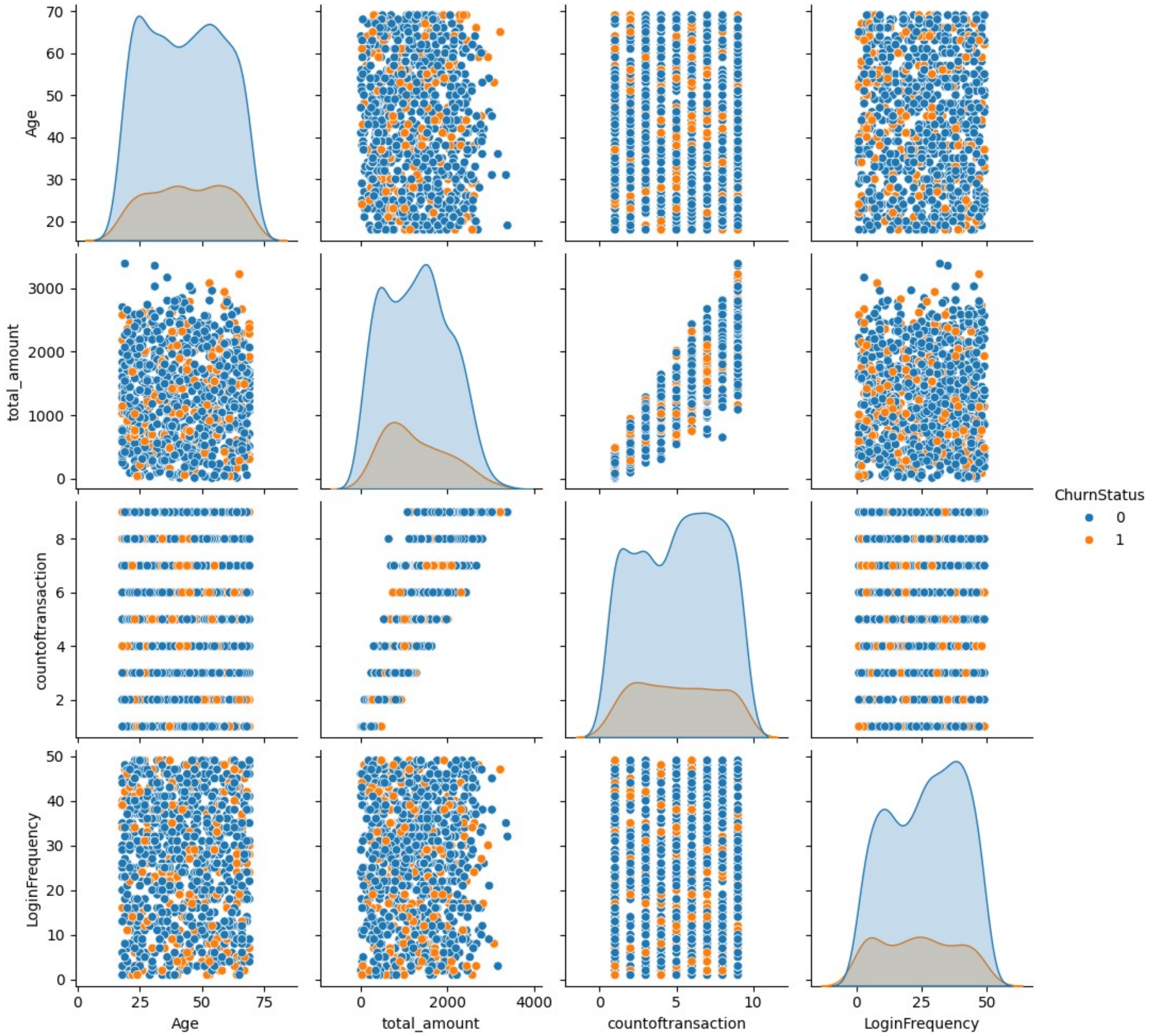
```
django_normal.plot( kind = 'bar'  , stacked = True )
```

```
<Axes: xlabel='ResolutionStatus'>
```



```
numeric_cols = [
    'Age', 'total_amount', 'countoftransaction',
    'LoginFrequency'
]

sns.pairplot(df_final[numeric_cols + ['ChurnStatus']],
hue='ChurnStatus')
```

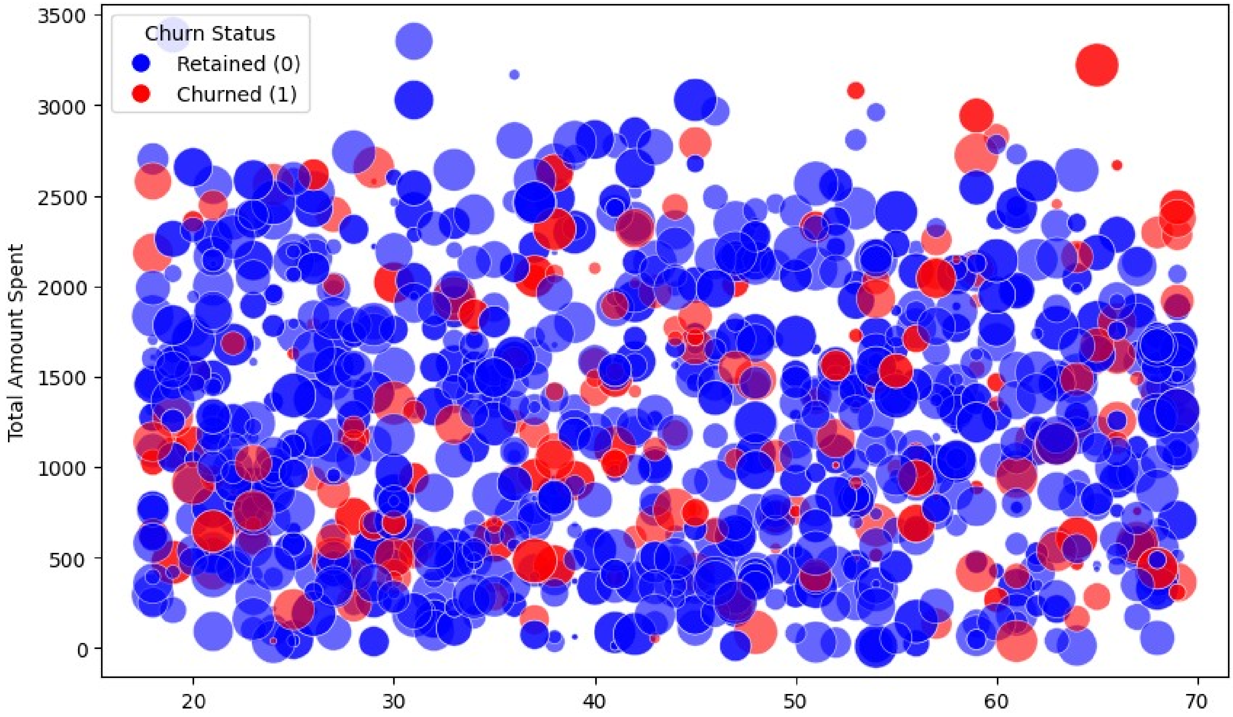
```
<seaborn.axisgrid.PairGrid at 0x24475408710>
```



```
palette = {0: 'blue', 1: 'red'}

plt.figure(figsize=(10, 6))
scatter = plt.scatter(
    df_final['Age'] ,
    df_final['total_amount'],
    s=df_final['LoginFrequency'] * 10,  # Adjust multiplier for
visible bubble sizes
    c=df_final['ChurnStatus'].map(palette),
    alpha=0.6,
    edgecolors='w',
    linewidth=0.5
)

# Custom legend for ChurnStatus
```

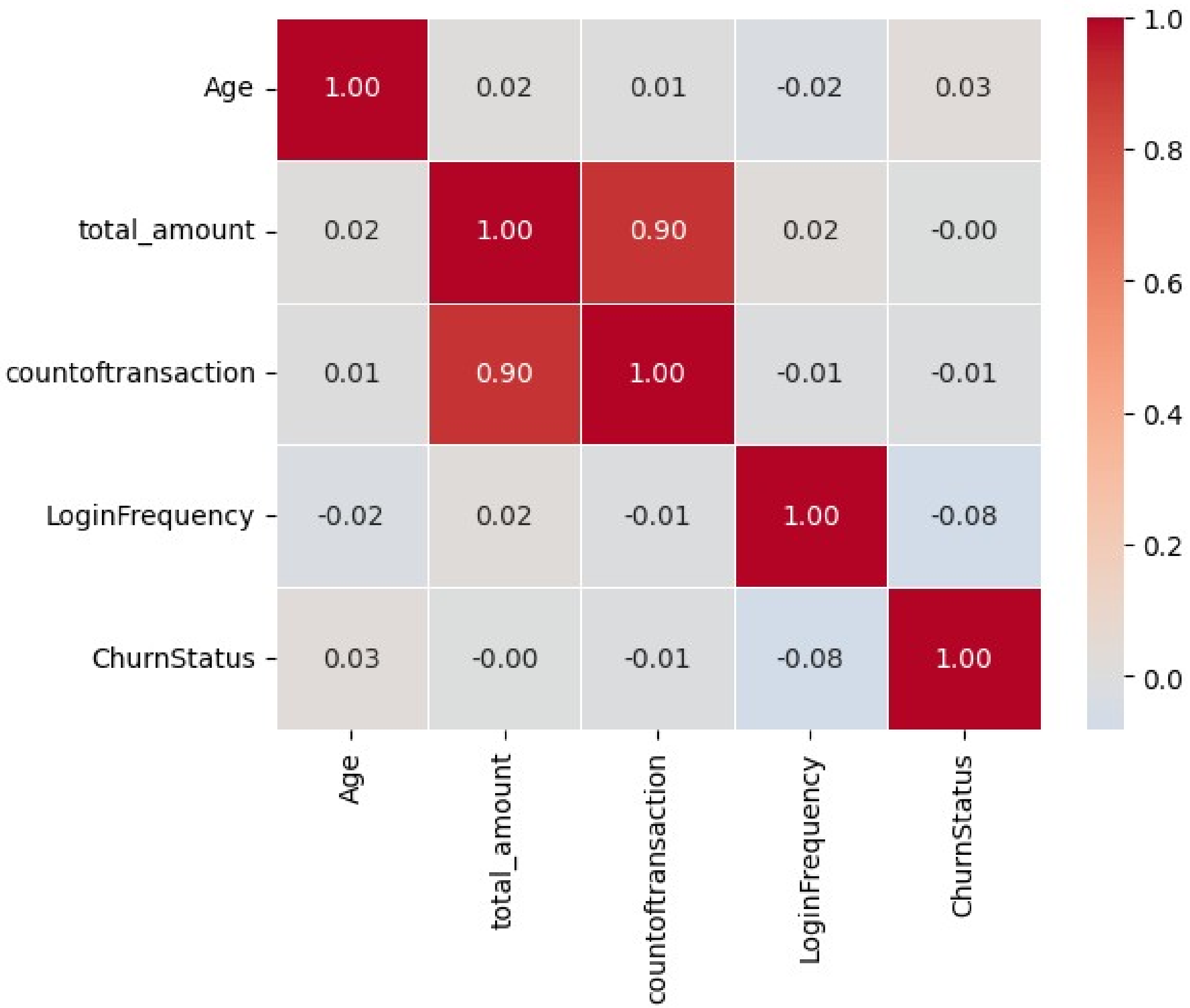
```
from matplotlib.lines import Line2D
legend_elements = [
    Line2D([0], [0], marker='o', color='w', label='Retained (0)',
markerfacecolor='blue', markersize=10),
    Line2D([0], [0], marker='o', color='w', label='Churned (1)',
markerfacecolor='red', markersize=10)
]

plt.legend(handles=legend_elements, title="Churn Status")
plt.ylabel('Total Amount Spent')

# we are losing low spending young customers
```

```
Text(0, 0.5, 'Total Amount Spent')
```



```
corr_matrix = df_final[numeric_cols +
['ChurnStatus']].corr(method='pearson')
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', center=0,
linewidths=.5, fmt=".2f")
```

```
<Axes: >
```



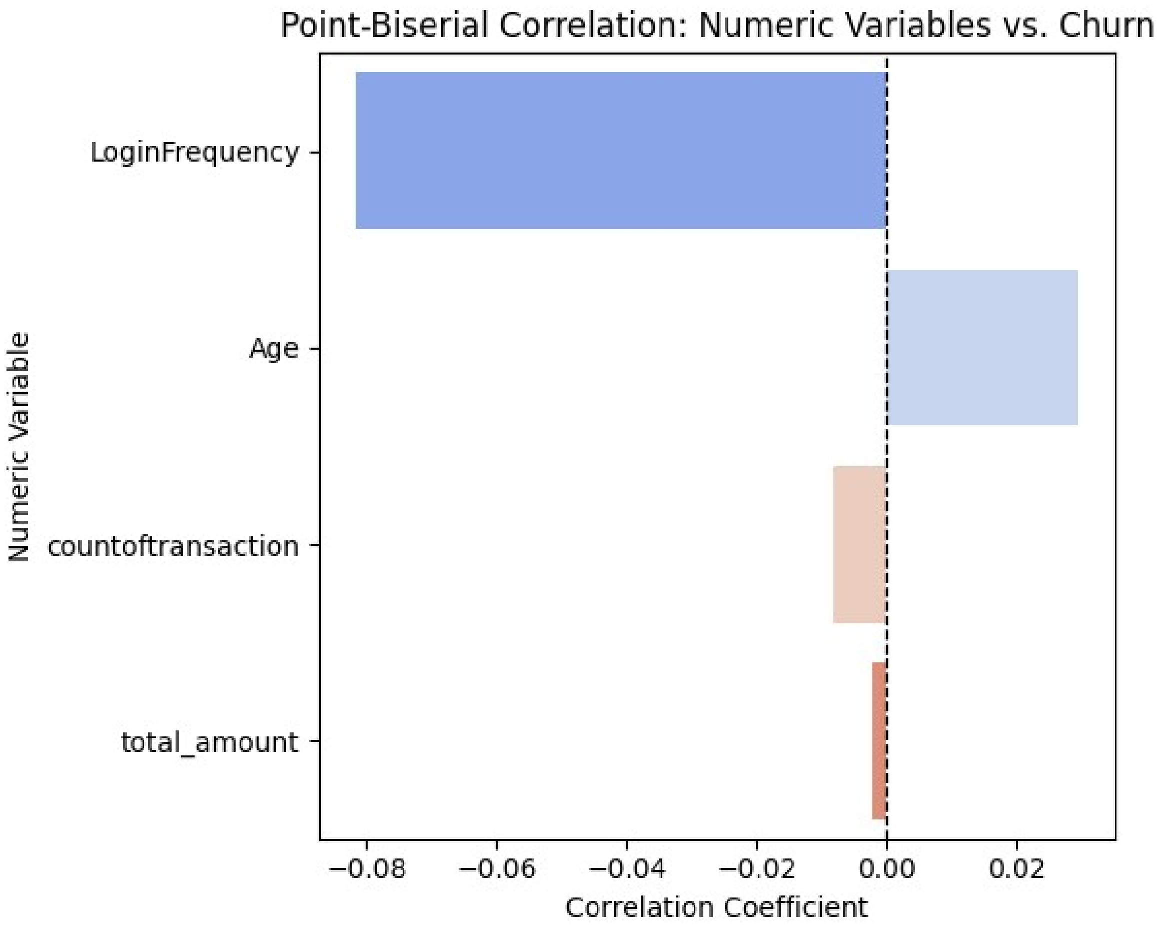
```
from scipy.stats import pointbiserialr

# List your numeric columns

correlations = {}
for col in numeric_cols:
    corr, _ = pointbiserialr(df_final['ChurnStatus'], df_final[col])
    correlations[col] = corr

# Convert to DataFrame for plotting
corr_df = pd.DataFrame(list(correlations.items()),
columns=['Variable', 'PointBiserialCorrelation'])
corr_df = corr_df.sort_values('PointBiserialCorrelation', key=abs,
ascending=False)

corr_df
```

{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"Variable","rawType":"object","type":"string"},

```
{"name":"PointBiserialCorrelation","rawType":"float64","type":"float"}
],"ref":"9b7acbc7-d71f-4274-9d9a-55ddb8f06d6d","rows":
[["3","LoginFrequency","-0.08159047725998914"],
["0","Age","0.029585983696895548"],["2","countoftransaction","-
0.00821776456349772"],["1","total_amount","-
0.0022439614369999823"]],"shape":{"columns":2,"rows":4}}
```

```python
plt.figure(figsize=(6, 5))
sns.barplot(
    data=corr_df,
    x='PointBiserialCorrelation',
    y='Variable',
    palette='coolwarm',
    orient='h'
)
plt.title('Point-Biserial Correlation: Numeric Variables vs. Churn')
plt.xlabel('Correlation Coefficient')
plt.ylabel('Numeric Variable')
plt.axvline(0, color='k', linestyle='--', linewidth=1)
plt.tight_layout()
plt.show()
```

```
C:\Users\Ayush\AppData\Local\Temp\ipykernel_3312\1030485422.py:2:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `y` variable to `hue` and set
`legend=False` for the same effect.

  sns.barplot(
```



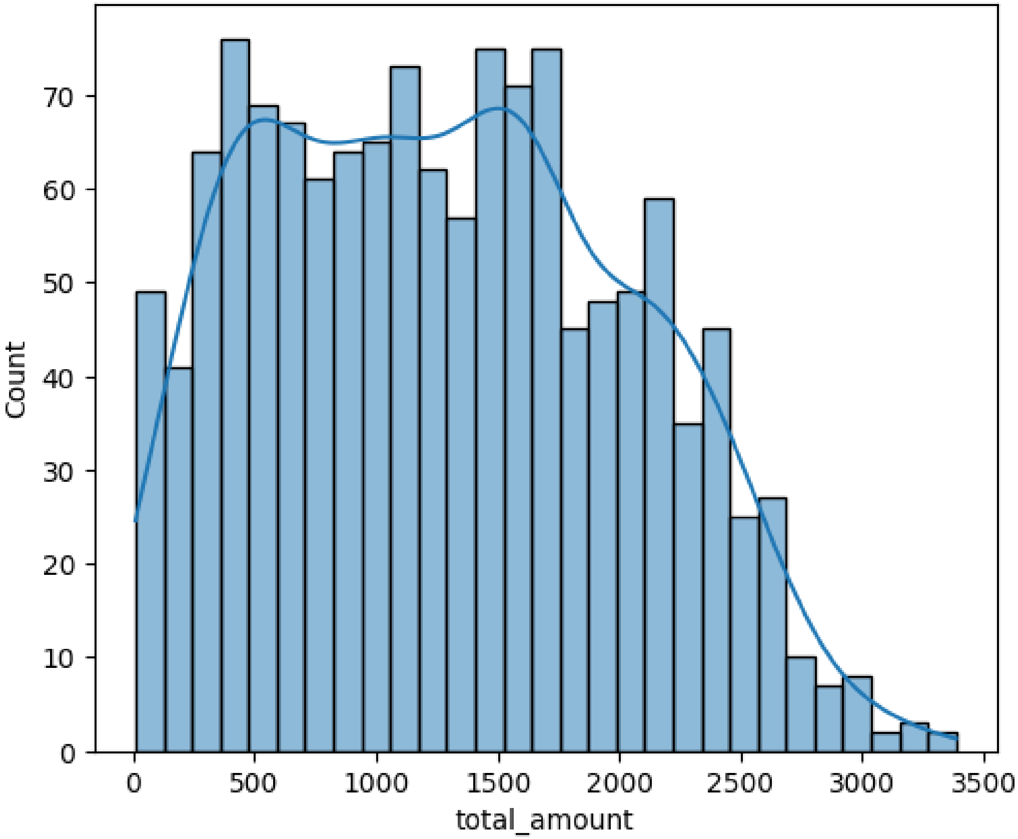
```
df_final.columns
```

```
Index(['CustomerID', 'Age', 'Gender', 'MaritalStatus', 'IncomeLevel',
       'ChurnStatus', 'total_amount', 'countoftransaction',
'LoginFrequency',
       'ServiceUsage', 'ResolutionStatus'],
      dtype='object')
```

outliers , imputing etc

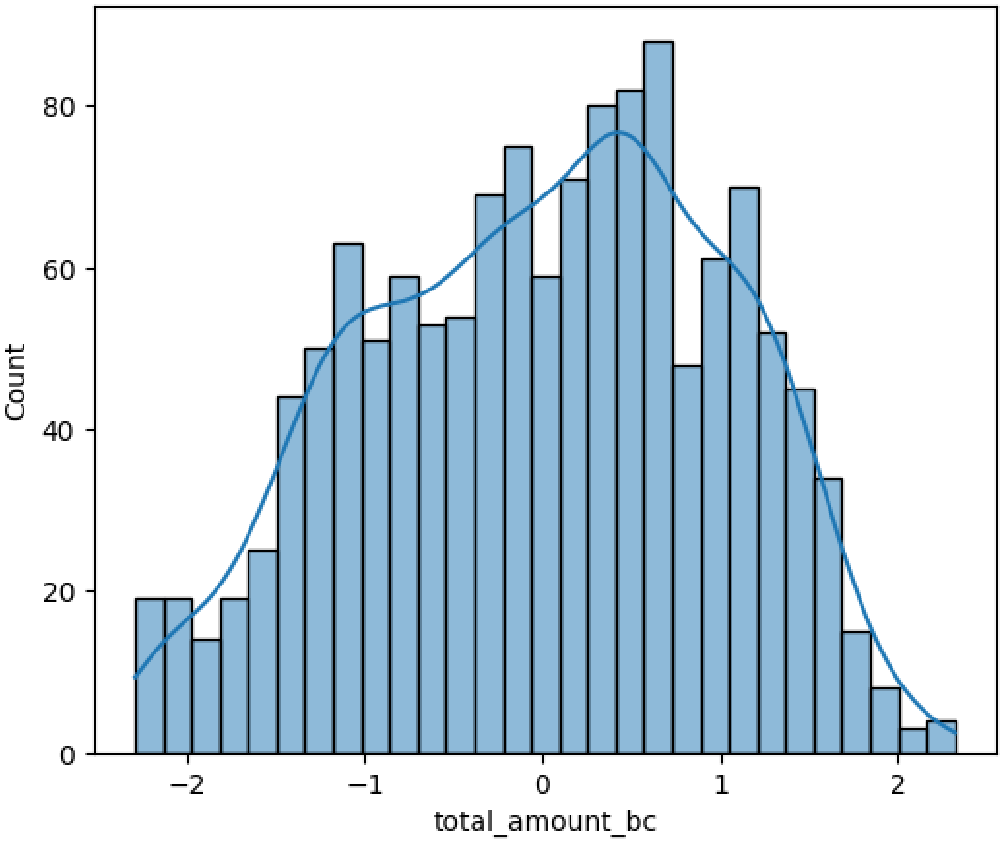
```
plt.figure(figsize=(6, 5))
sns.histplot( data = df_final  , x ='total_amount' , bins = 29 , kde=
True )
```

```
<Axes: xlabel='total_amount', ylabel='Count'>
```



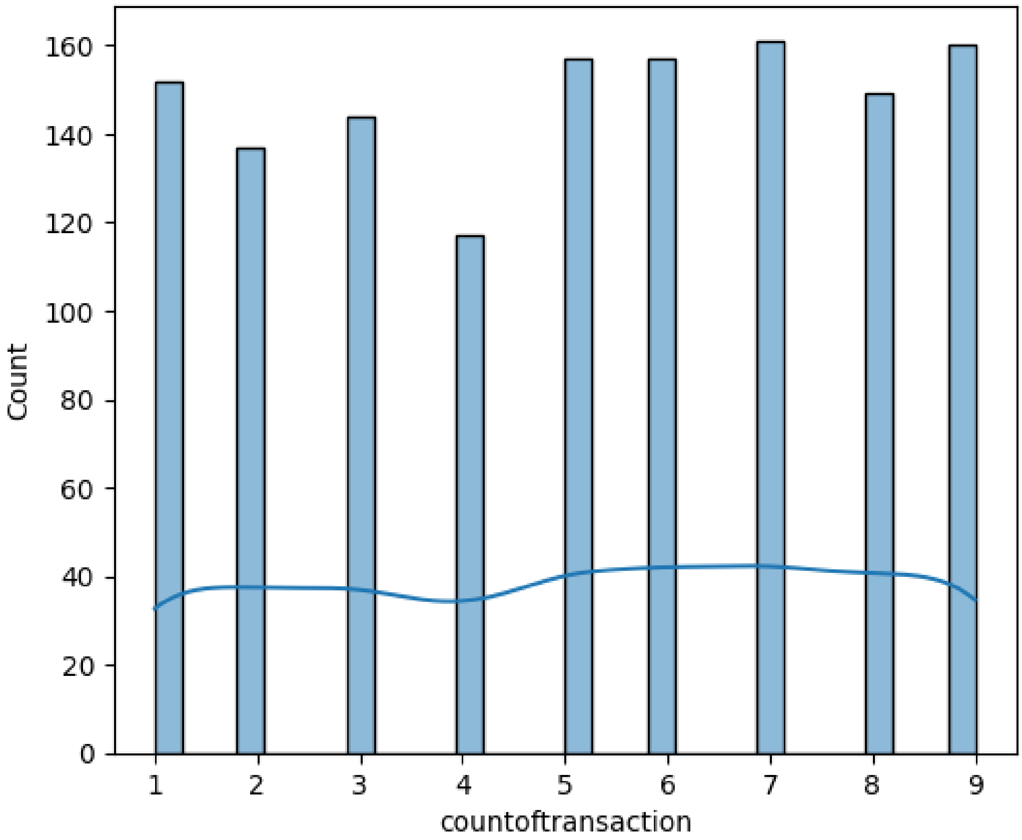
```python
from sklearn.preprocessing import PowerTransformer

# For strictly positive data (Box-Cox)
pt = PowerTransformer(method='box-cox')
df_final['total_amount_bc'] =
pt.fit_transform(df_final[['total_amount']])

plt.figure(figsize=(6, 5))
sns.histplot( data = df_final  , x ='total_amount_bc' , bins = 29 ,
kde= True )
```

```
<Axes: xlabel='total_amount_bc', ylabel='Count'>
```

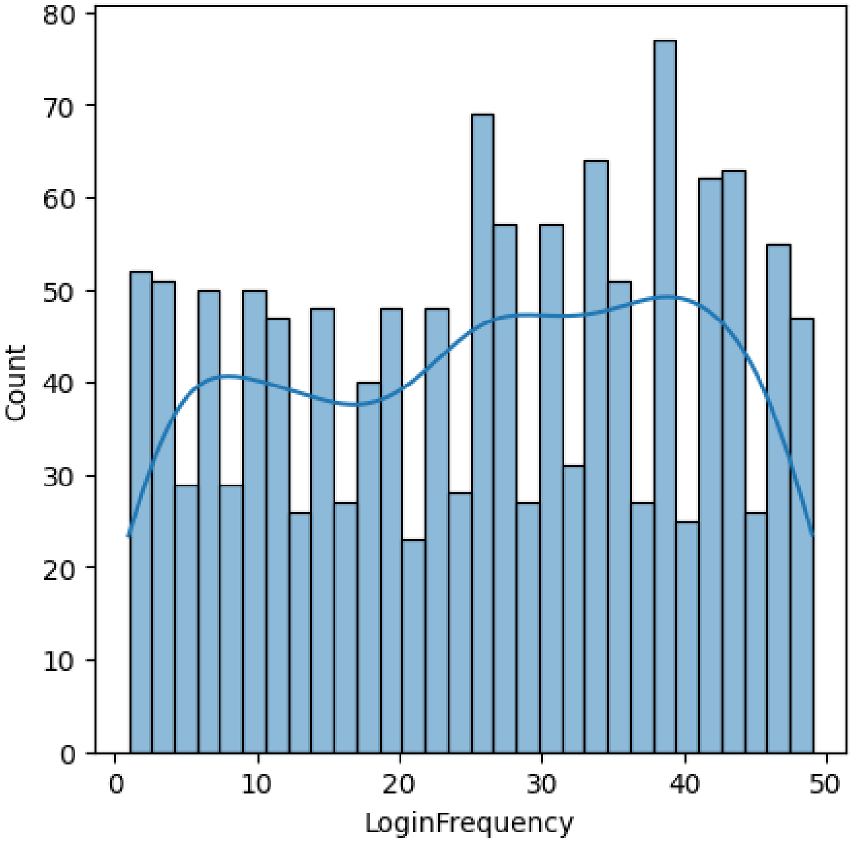


```
plt.figure(figsize=(6, 5))
sns.histplot( data = df_final  , x ='countoftransaction' , bins = 30,
kde= True )
```

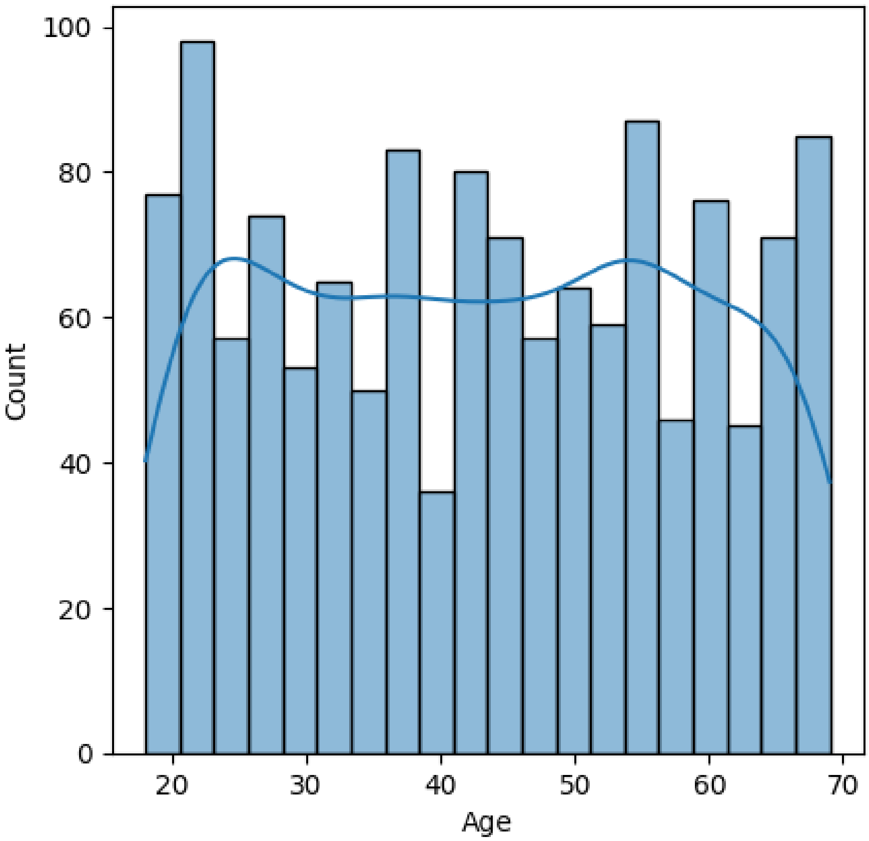
```
<Axes: xlabel='countoftransaction', ylabel='Count'>
```



```
plt.figure(figsize=(5, 5))
sns.histplot( data = df_final  , x ='LoginFrequency' , bins = 30, kde=
True )

<Axes: xlabel='LoginFrequency', ylabel='Count'>
```



```
plt.figure(figsize=(5, 5))
sns.histplot( data = df_final  , x ='Age' , bins = 20, kde= True )

<Axes: xlabel='Age', ylabel='Count'>
```



```
from scipy.stats import zscore

# List your numeric columns
numeric_cols = [
     'total_amount', 'countoftransaction',
    'LoginFrequency'
]

outlier_results = {}

for col in numeric_cols:
    data = pd.to_numeric(df_final[col], errors='coerce').dropna()
    # IQR Method
    Q1 = data.quantile(0.25)
    Q3 = data.quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR
    iqr_outliers = (data < lower_bound) | (data > upper_bound)

    # Z-score Method
    z_scores = zscore(data)
    z_outliers = np.abs(z_scores) > 3
    # Combine outlier indices
    outlier_indices_iqr = data.index[iqr_outliers].tolist()
    outlier_indices_z = data.index[z_outliers].tolist()

    outlier_results[col] = {
        'iqr_outlier_indices': outlier_indices_iqr,
        'zscore_outlier_indices': outlier_indices_z,
        'num_iqr_outliers': len(outlier_indices_iqr),
        'num_zscore_outliers': len(outlier_indices_z),
    }
    print(f"{col} - IQR: {len(outlier_indices_iqr)} outliers, Z-score:
{len(outlier_indices_z)} outliers")

# Example: To view outlier rows for a specific column using IQR
# df_final.loc[outlier_results['total_amount']['iqr_outlier_indices']]
```

```
total_amount - IQR: 0 outliers, Z-score: 0 outliers
countoftransaction - IQR: 0 outliers, Z-score: 0 outliers
LoginFrequency - IQR: 0 outliers, Z-score: 0 outliers
```

```
df_final.drop( columns= ['total_amount_log'] , inplace= True )
```

```
df_final
```

```
{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"Age","rawType":"int64","type":"integer"},
{"name":"Gender","rawType":"object","type":"string"},
{"name":"MaritalStatus","rawType":"object","type":"string"},
{"name":"IncomeLevel","rawType":"object","type":"string"},
{"name":"ChurnStatus","rawType":"int64","type":"integer"},
{"name":"total_amount","rawType":"float64","type":"float"},
{"name":"countoftransaction","rawType":"int64","type":"integer"},
{"name":"LoginFrequency","rawType":"int64","type":"integer"},
{"name":"ServiceUsage","rawType":"object","type":"string"},
{"name":"ResolutionStatus","rawType":"object","type":"string"},
{"name":"total_amount_bc","rawType":"float64","type":"float"}],"ref":"
33547094-0303-4369-bc76-1c6b00344b6e","rows":
[["0","1","62","M","Single","Low","0","416.5","1","34","Mobile
App","Resolved","-1.1901642609199812"],
["1","2","65","M","Married","Low","1","1547.4199999999998","7","5","We
bsite","Resolved","0.43822023200276483"],
["2","3","18","M","Single","Low","0","1702.98","6","3","Website","Reso
lved","0.6206325962206389"],
["3","4","21","M","Widowed","Low","0","917.29","5","2","Website","Reso
lved","-0.3796921171977037"],
["4","4","21","M","Widowed","Low","0","917.29","5","2","Website","Unre
solved","-0.3796921171977037"],
["5","5","21","M","Divorced","Medium","0","2001.49","8","41","Website"
,"No Problem","0.9551734421229886"],
["6","6","57","F","Divorced","Medium","0","1164.29","5","2","Website",
"Resolved","-0.041026979273344984"],
["7","7","27","F","Married","High","0","86.73","1","32","Mobile
App","No Problem","-1.961883066701129"],
["8","8","37","M","Single","Low","1","2046.88","7","17","Online
Banking","Unresolved","1.0044663399625502"],
["9","8","37","M","Single","Low","1","2046.88","7","17","Online
Banking","Unresolved","1.0044663399625502"],
["10","9","39","M","Divorced","High","0","1317.24","5","24","Website",
"Resolved","0.15605453836881017"],
["11","10","68","M","Married","High","1","1397.36","7","29","Online
Banking","No Problem","0.25608631618162225"],
["12","11","54","M","Divorced","Medium","0","1733.83","5","30","Online
Banking","Resolved","0.656104788648745"],
["13","11","54","M","Divorced","Medium","0","1733.83","5","30","Online
Banking","Unresolved","0.656104788648745"],
["14","12","41","F","Married","Low","0","87.1","1","43","Website","Res
olved","-1.9606782443744375"],
["15","12","41","F","Married","Low","0","87.1","1","43","Website","Unr
esolved","-1.9606782443744375"],
["16","13","24","F","Divorced","Low","0","392.02","3","10","Mobile
App","Resolved","-1.2367978100623787"],
["17","14","42","M","Widowed","Medium","0","2843.59","9","26","Online
Banking","Unresolved","1.8161825611228544"],
["18","14","42","M","Widowed","Medium","0","2843.59","9","26","Online
Banking","Resolved","1.8161825611228544"],
["19","15","42","M","Divorced","Medium","0","1914.02","8","44","Online
Banking","Unresolved","0.8590672945046037"],
["20","15","42","M","Divorced","Medium","0","1914.02","8","44","Online
Banking","Unresolved","0.8590672945046037"],
["21","16","30","F","Single","Medium","1","2019.79","9","41","Online
Banking","Resolved","0.9750934186066263"],
["22","16","30","F","Single","Medium","1","2019.79","9","41","Online
Banking","Resolved","0.9750934186066263"],
["23","17","19","M","Married","Low","0","1625.27","7","8","Online
Banking","Unresolved","0.5302701236774611"],
["24","17","19","M","Married","Low","0","1625.27","7","8","Online
Banking","Resolved","0.5302701236774611"],
["25","18","56","F","Married","High","1","951.3800000000001","5","37",
"Website","Unresolved","-0.3312106037377801"],
["26","19","57","F","Divorced","High","0","1017.26","5","39","Online
Banking","Unresolved","-0.23920075300576915"],
["27","19","57","F","Divorced","High","0","1017.26","5","39","Online
Banking","Unresolved","-0.23920075300576915"],
["28","20","41","M","Single","Low","0","2775.85","9","16","Website","R
esolved","1.7505691628647226"],
["29","21","64","M","Divorced","High","0","913.02","5","18","Website",
"No Problem","-0.38580868952859837"],
["30","22","42","F","Divorced","Low","0","1582.73","6","42","Mobile
App","Resolved","0.4801662949072248"],
["31","22","42","F","Divorced","Low","0","1582.73","6","42","Mobile
App","Resolved","0.4801662949072248"],
["32","23","35","F","Divorced","Low","0","1490.74","4","19","Mobile
App","Unresolved","0.3701825145092396"],
["33","23","35","F","Divorced","Low","0","1490.74","4","19","Mobile
App","Resolved","0.3701825145092396"],
["34","24","55","M","Widowed","Medium","0","1143.22","5","12","Mobile
App","Resolved","-0.06886566901905382"],
["35","25","43","F","Married","Low","0","764.77","3","27","Mobile
App","Resolved","-0.6048482455780341"],
["36","26","31","M","Single","Medium","1","571.76","3","45","Mobile
App","Resolved","-0.9138823188174997"],
["37","27","26","M","Widowed","Medium","0","2210.6","8","22","Mobile
App","Unresolved","1.179186215571774"],
["38","28","27","F","Widowed","High","0","1702.26","6","42","Online
Banking","Unresolved","0.6198020560429668"],
["39","28","27","F","Widowed","High","0","1702.26","6","42","Online
Banking","Unresolved","0.6198020560429668"],
["40","29","38","F","Divorced","Medium","0","853.0699999999999","3","3
4","Website","Resolved","-0.47276867303680686"],
["41","29","38","F","Divorced","Medium","0","853.0699999999999","3","3
4","Website","Unresolved","-0.47276867303680686"],
["42","30","69","M","Divorced","High","0","706.66","2","38","Website",
"Resolved","-0.6946836844866168"],
["43","30","69","M","Divorced","High","0","706.66","2","38","Website",
"Unresolved","-0.6946836844866168"],
["44","31","34","M","Divorced","Low","0","1108.56","5","23","Online
Banking","Resolved","-0.11505288285365234"],
["45","31","34","M","Divorced","Low","0","1108.56","5","23","Online
Banking","Resolved","-0.11505288285365234"],
["46","32","69","F","Married","High","0","1225.52","5","48","Mobile
App","No Problem","0.03889366204341104"],
["47","33","23","M","Divorced","High","1","689.0","2","24","Mobile
App","Resolved","-0.7224903487606392"],
["48","34","33","M","Widowed","High","0","1886.4","9","44","Online
Banking","Resolved","0.8284039033416816"],
["49","34","33","M","Widowed","High","0","1886.4","9","44","Online
Banking","Unresolved","0.8284039033416816"]],"shape":
{"columns":12,"rows":1334}}
```

```
cols_to_scale = ['Age', 'total_amount_bc', 'LoginFrequency',
'countoftransaction']

scaler = StandardScaler()
x_scaled = scaler.fit_transform( df_final[cols_to_scale])
```

```
xsdf = pd.DataFrame( x_scaled  , columns= [ col +'_scaled'  for col in
cols_to_scale ] , index = df_final.index )

xsdf
```

```
{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"Age_scaled","rawType":"float64","type":"float"},
{"name":"total_amount_bc_scaled","rawType":"float64","type":"float"},
{"name":"LoginFrequency_scaled","rawType":"float64","type":"float"},
{"name":"countoftransaction_scaled","rawType":"float64","type":"float"
}],"ref":"38a4e535-7756-4d91-ac0c-86413fa8744a","rows":
[["0","1.230462561333705","-
1.1901642609199814","0.5863376053144339","-1.5770396009897798"],
["1","1.426908532948078","0.4382202320027649","-
1.4787622307374777","0.7272000788791865"],["2","-
1.6507450223437656","0.620632596220639","-
1.6211829090858854","0.3431601322343588"],["3","-
1.4542990507293925","-0.37969211719770374","-1.6923932482600892","-
0.040879814410468895"],["4","-1.4542990507293925","-
0.37969211719770374","-1.6923932482600892","-0.040879814410468895"],
["5","-
1.4542990507293925","0.9551734421229887","1.0848099795338608","1.11124
00255240142"],["6","0.9030526086430833","-0.041026979273345","-
1.6923932482600892","-0.040879814410468895"],["7","-
1.0614071075006466","-1.9618830667011293","0.4439169269660262 5","-
1.5770396009897798"],["8","-
0.4065872021194033","1.0044663399625504","-
0.6242381606470314","0.7272000788791865"],["9","-
0.4065872021194033","1.0044663399625504","-
0.6242381606470314","0.7272000788791865"],["10","-
0.2756232210431546","0.156054538368 8102","-0.12576578642760453","-
0.040879814410468895"],
["11","1.623354504562451","0.2560863161816223","0.23028590944341473","
0.7272000788791865"],
["12","0.7066066370287103","0.6561047886487451","0.3014962486176186","
-0.040879814410468895"],
["13","0.7066066370287103","0.6561047886487451","0.3014962486176186","
-0.040879814410468895"],["14","-0.14465923996690597","-
1.9606782443744377","1.2272306578822685","-1.5770396009897798"],
["15","-0.14465923996690597","-
1.9606782443744377","1.2272306578822685","-1.5770396009897798"],
["16","-1.2578530791150195","-1.236797810062378 9","-
1.1227105348664583","-0.8089597077001243"],["17","-
0.07917724942878164","1.8161825611228546","0.016654891920803182","1.49
5279972168842"],["18","-
0.07917724942878164","1.8161825611228546","0.016654891920803182","1.49
5279972168842"],["19","-
0.07917724942878164","0.8590672945046038","1.2984409970564725","1.1112
400255240142"],["20","-
0.07917724942878164","0.8590672945046038","1.2984409970564725","1.1112
400255240142"],["21","-
0.8649611358862735","0.9750934186066265","1.0848099795338608","1.49527
9972168842"],["22","-
0.8649611358862735","0.9750934186066265","1.0848099795338608","1.49527
9972168842"],["23","-1.5852630318056413","0.5302701236774612","-
1.265131213214866","0.7272000788791865"],["24","-
1.5852630318056413","0.5302701236774612","-
1.265131213214866","0.7272000788791865"],["25","0.837570618104959","-
0.33121060373778016","0.7999686228370455","-0.040879814410468895"],
["26","0.9030526086430833","-
0.23920075300576918","0.9423893011854532","-0.040879814410468895"],
["27","0.9030526086430833","-
0.23920075300576918","0.9423893011854532","-0.040879814410468895"],
["28","-0.14465923996690597","1.7505691628647229","-
0.6954484998212352","1.495279972168842"],["29","1.3614265424099536","-
0.3858086895285984","-0.5530278214728276","-0.040879814410468895"],
["30","-
0.07917724942878164","0.48016629490722484","1.1560203187080647","0.343
160132234358 8"],["31","-
0.07917724942878164","0.48016629490722484","1.1560203187080647","0.343
160132234358 8"],["32","-0.537551183195652","0.37018251450923967","-
0.48181748229862376","-0.4249197610552966"],["33","-
0.537551183195652","0.37018251450923967","-0.48181748229862376","-
0.4249197610552966"],["34","0.7720886275668346","-
0.06886566901905383","-0.9802898565180507","-0.040879814410468895"],
["35","-0.013695258890657307","-
0.6048482455780342","0.08786523109500703","-0.8089597077001243"],
["36","-0.7994791453481492","-
0.9138823188174998","1.3696513362306764","-0.8089597077001243"],
["37","-1.126889098038771","1.1791862155717743","-
0.2681864647760122","1.1112400255240142"],["38","-
1.0614071075006466","0.6198020560429669","1.1560203187080647","0.34316
01322343588"],["39","-
1.0614071075006466","0.6198020560429669","1.1560203187080647","0.34316
01322343588"],["40","-0.34110521158127893","-
0.4727686730368069","0.5863376053144339","-0.8089597077001243"],
["41","-0.34110521158127893","-
0.4727686730368069","0.5863376053144339","-0.8089597077001243"],
["42","1.6888364951005752","-
0.6946836844866169","0.8711789620112493","-1.192999654344952"],
["43","1.6888364951005752","-
0.6946836844866169","0.8711789620112493","-1.192999654344952"],
["44","-0.6030331737337763","-0.11505288285365235","-
0.19697612560180836","-0.040879814410468895"],["45","-
0.6030331737337763","-0.11505288285365235","-0.19697612560180836","-
0.040879814410468895"],
["46","1.6888364951005752","0.03889366204341104","1.5832823537532879",
"-0.040879814410468895"],["47","-1.3233350696531438","-
0.7224903487606393","-0.12576578642760453","-1.192999654344952"],
["48","-
0.6685151642719006","0.8284039033416817","1.2984409970564725","1.49527
9972168842"],["49","-
0.6685151642719006","0.8284039033416817","1.2984409970564725","1.49527
9972168842"]],"shape":{"columns":4,"rows":1334}}
```

```
df_final = pd.concat([df_final, xsdf], axis=1 )

df_final.drop( columns = cols_to_scale +['total_amount']  , inplace=
True )

df_final
```

```
{"columns":[{"name":"index","rawType":"int64","type":"integer"},
{"name":"CustomerID","rawType":"int64","type":"integer"},
{"name":"Gender","rawType":"object","type":"string"},
{"name":"MaritalStatus","rawType":"object","type":"string"},
{"name":"IncomeLevel","rawType":"object","type":"string"},
{"name":"ChurnStatus","rawType":"int64","type":"integer"},
{"name":"ServiceUsage","rawType":"object","type":"string"},
{"name":"ResolutionStatus","rawType":"object","type":"string"},
{"name":"Age_scaled","rawType":"float64","type":"float"},
{"name":"total_amount_bc_scaled","rawType":"float64","type":"float"},
{"name":"LoginFrequency_scaled","rawType":"float64","type":"float"},
{"name":"countoftransaction_scaled","rawType":"float64","type":"float"
}],"ref":"0c83a093-dc6d-4025-a653-7fe061fa39cc","rows":
[["0","1","M","Single","Low","0","Mobile
App","Resolved","1.230462561333705","-
1.1901642609199814","0.5863376053144339","-1.5770396009897798"],
["1","2","M","Married","Low","1","Website","Resolved","1.4269085329480
78","0.4382202320027649","-1.4787622307374777","0.7272000788791865"],
["2","3","M","Single","Low","0","Website","Resolved","-
1.6507450223437656","0.620632596220639","-
1.6211829090858854","0.3431601322343588"],
["3","4","M","Widowed","Low","0","Website","Resolved","-
1.4542990507293925","-0.37969211719770374","-1.6923932482600892","-
0.040879814410468895"],
["4","4","M","Widowed","Low","0","Website","Unresolved","-
1.4542990507293925","-0.37969211719770374","-1.6923932482600892","-
0.040879814410468895"],
["5","5","M","Divorced","Medium","0","Website","No Problem","-
1.4542990507293925","0.9551734421229887","1.0848099795338608","1.11124
00255240142"],
["6","6","F","Divorced","Medium","0","Website","Resolved","0.903052608
6430833","-0.041026979273345","-1.6923932482600892","-
0.040879814410468895"],["7","7","F","Married","High","0","Mobile
App","No Problem","-1.0614071075006466","-
1.9618830667011293","0.4439169269660262 5","-1.5770396009897798"],
["8","8","M","Single","Low","1","Online Banking","Unresolved","-
0.4065872021194033","1.0044663399625504","-
```

```
0.6242381606470314","0.7272000788791865"],
["9","8","M","Single","Low","1","Online Banking","Unresolved","-
0.4065872021194033","1.0044663399625504","-
0.6242381606470314","0.7272000788791865"],
["10","9","M","Divorced","High","0","Website","Resolved","-
0.2756232210431546","0.1560545383688102","-0.12576578642760453","-
0.040879814410468895"],["11","10","M","Married","High","1","Online
Banking","No
Problem","1.623354504562451","0.2560863161816223","0.23028590944341473
","0.7272000788791865"],["12","11","M","Divorced","Medium","0","Online
Banking","Resolved","0.7066066370287103","0.6561047886487451","0.30149
62486176186","-0.040879814410468895"],
["13","11","M","Divorced","Medium","0","Online
Banking","Unresolved","0.7066066370287103","0.6561047886487451","0.301
4962486176186","-0.040879814410468895"],
["14","12","F","Married","Low","0","Website","Resolved","-
0.14465923996690597","-1.9606782443744377","1.2272306578822685","-
1.5770396009897798"],
["15","12","F","Married","Low","0","Website","Unresolved","-
0.14465923996690597","-1.9606782443744377","1.2272306578822685","-
1.5770396009897798"],["16","13","F","Divorced","Low","0","Mobile
App","Resolved","-1.2578530791150195","-1.2367978100623789","-
1.1227105348664583","-0.8089597077001243"],
["17","14","M","Widowed","Medium","0","Online Banking","Unresolved","-
0.07917724942878164","1.8161825611228546","0.016654891920803182","1.49
5279972168842"],["18","14","M","Widowed","Medium","0","Online
Banking","Resolved","-
0.07917724942878164","1.8161825611228546","0.016654891920803182","1.49
5279972168842"],["19","15","M","Divorced","Medium","0","Online
Banking","Unresolved","-
0.07917724942878164","0.8590672945046038","1.2984409970564725","1.1112
400255240142"],["20","15","M","Divorced","Medium","0","Online
Banking","Unresolved","-
0.07917724942878164","0.8590672945046038","1.2984409970564725","1.1112
400255240142"],["21","16","F","Single","Medium","1","Online
Banking","Resolved","-
0.8649611358862735","0.9750934186066265","1.0848099795338608","1.49527
9972168842"],["22","16","F","Single","Medium","1","Online
Banking","Resolved","-
0.8649611358862735","0.9750934186066265","1.0848099795338608","1.49527
9972168842"],["23","17","M","Married","Low","0","Online
Banking","Unresolved","-1.5852630318056413","0.5302701236774612","-
1.265131213214866","0.7272000788791865"],
["24","17","M","Married","Low","0","Online Banking","Resolved","-
1.5852630318056413","0.5302701236774612","-
1.265131213214866","0.7272000788791865"],
["25","18","F","Married","High","1","Website","Unresolved","0.83757061
8104959","-0.33121060373778016","0.7999686228370455","-
0.040879814410468895"],["26","19","F","Divorced","High","0","Online
```

```
Banking","Unresolved","0.9030526086430833","-
0.23920075300576918","0.9423893011854532","-0.040879814410468895"],
["27","19","F","Divorced","High","0","Online
Banking","Unresolved","0.9030526086430833","-
0.23920075300576918","0.9423893011854532","-0.040879814410468895"],
["28","20","M","Single","Low","0","Website","Resolved","-
0.14465923996690597","1.7505691628647229","-
0.6954484998212352","1.495279972168842"],
["29","21","M","Divorced","High","0","Website","No
Problem","1.3614265424099536","-0.3858086895285984","-
0.5530278214728276","-0.040879814410468895"],
["30","22","F","Divorced","Low","0","Mobile App","Resolved","-
0.07917724942878164","0.48016629490722484","1.1560203187080647","0.343
1601322343588"],["31","22","F","Divorced","Low","0","Mobile
App","Resolved","-
0.07917724942878164","0.48016629490722484","1.1560203187080647","0.343
1601322343588"],["32","23","F","Divorced","Low","0","Mobile
App","Unresolved","-0.537551183195652","0.37018251450923967","-
0.48181748229862376","-0.4249197610552966"],
["33","23","F","Divorced","Low","0","Mobile App","Resolved","-
0.537551183195652","0.37018251450923967","-0.48181748229862376","-
0.4249197610552966"],["34","24","M","Widowed","Medium","0","Mobile
App","Resolved","0.7720886275668346","-0.06886566901905383","-
0.9802898565180507","-0.040879814410468895"],
["35","25","F","Married","Low","0","Mobile App","Resolved","-
0.013695258890657307","-0.6048482455780342","0.08786523109500703","-
0.8089597077001243"],["36","26","M","Single","Medium","1","Mobile
App","Resolved","-0.7994791453481492","-
0.9138823188174998","1.3696513362306764","-0.8089597077001243"],
["37","27","M","Widowed","Medium","0","Mobile App","Unresolved","-
1.126889098038771","1.1791862155717743","-
0.2681864647760122","1.1112400255240142"],
["38","28","F","Widowed","High","0","Online Banking","Unresolved","-
1.0614071075006466","0.6198020560429669","1.1560203187080647","0.34316
01322343588"],["39","28","F","Widowed","High","0","Online
Banking","Unresolved","-
1.0614071075006466","0.6198020560429669","1.1560203187080647","0.34316
01322343588"],
["40","29","F","Divorced","Medium","0","Website","Resolved","-
0.34110521158127893","-0.4727686730368069","0.5863376053144339","-
0.8089597077001243"],
["41","29","F","Divorced","Medium","0","Website","Unresolved","-
0.34110521158127893","-0.4727686730368069","0.5863376053144339","-
0.8089597077001243"],
["42","30","M","Divorced","High","0","Website","Resolved","1.688836495
1005752","-0.6946836844866169","0.8711789620112493","-
1.192999654344952"],
["43","30","M","Divorced","High","0","Website","Unresolved","1.6888364
951005752","-0.6946836844866169","0.8711789620112493","-
```

```
1.192999654344952"],["44","31","M","Divorced","Low","0","Online
Banking","Resolved","-0.6030331737337763","-0.11505288285365235","-
0.19697612560180836","-0.040879814410468895"],
["45","31","M","Divorced","Low","0","Online Banking","Resolved","-
0.6030331737337763","-0.11505288285365235","-0.19697612560180836","-
0.040879814410468895"],["46","32","F","Married","High","0","Mobile
App","No
Problem","1.6888364951005752","0.03889366204341104","1.583282353753287
9","-0.040879814410468895"],
["47","33","M","Divorced","High","1","Mobile App","Resolved","-
1.3233350696531438","-0.7224903487606393","-0.12576578642760453","-
1.192999654344952"],["48","34","M","Widowed","High","0","Online
Banking","Resolved","-
0.6685151642719006","0.8284039033416817","1.2984409970564725","1.49527
9972168842"],["49","34","M","Widowed","High","0","Online
Banking","Unresolved","-
0.6685151642719006","0.8284039033416817","1.2984409970564725","1.49527
9972168842"]],"shape":{"columns":11,"rows":1334}}
```

```python
cat_cols = ['Gender' , 'MaritalStatus'  ,'IncomeLevel' ,
'ServiceUsage' ,'ResolutionStatus'   ]

dummies = pd.get_dummies( df_final[cat_cols]  , prefix= 'cat_')


df_final = pd.concat( [df_final ,dummies]  , axis =1 )

df_final.drop( columns=cat_cols , inplace=True )
```