Untitled - Power BI Desktop

Search

Sign in

File Home Insert Modeling View Optimize Help

Share

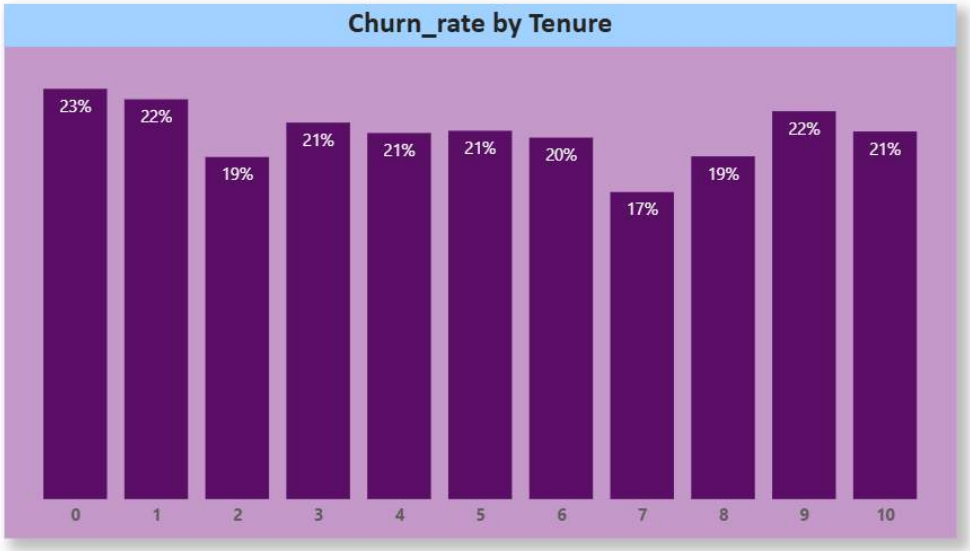
Paste Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Transform data Refresh New visual Text box More visuals New visual calculation New measure Quick measure Sensitivity Publish Prep data f AI

Clipboard Data Queries Insert Calculations Sensitivity Share Cop

1. What is the overall churn rate, and how does it trend over time (if applicable)?

Churn Rate

20%



Filters

Data

Search

Customer-Churn-Reco...

- Age
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- IsActiveMember
- NumOfProducts
- Point Earned
- RowNumber
- Satisfaction Sco...
- Surname
- Tenure

Page 1

Page 1 of 1

73%

Untitled - Power BI Desktop

Search

Sign in

File Home Insert Modeling View Optimize Help

Share

Paste Clipboard

Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Data

Transform data Refresh Queries

New visual Text box More visuals Insert

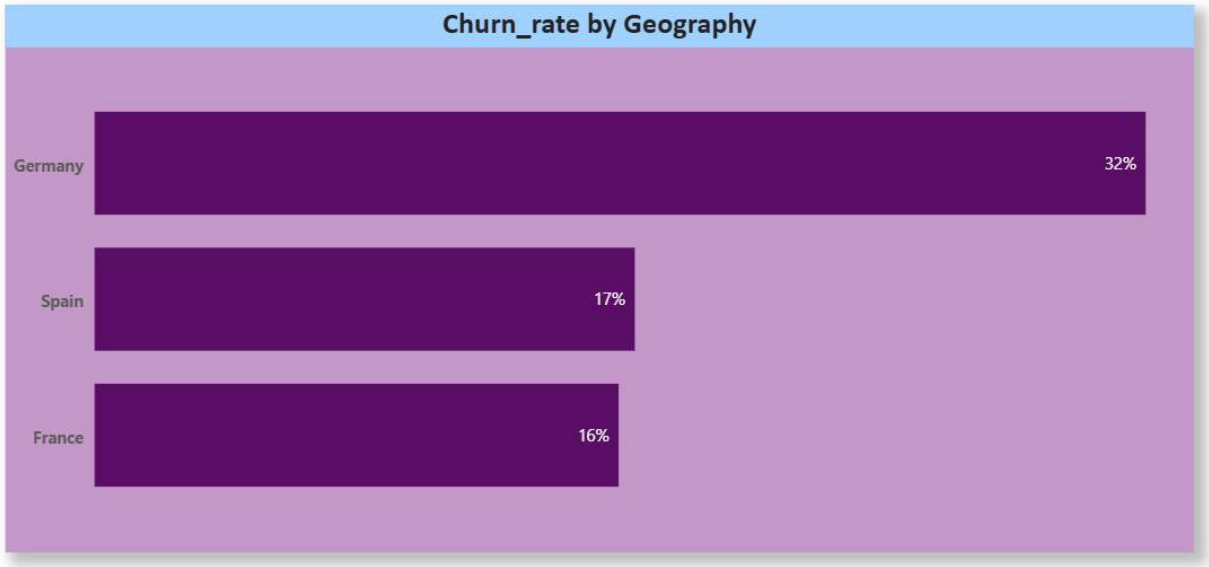
New visual calculation New measure Quick measure Calculations

Sensitivity Sensitivity

Publish Share

Prep data f AI Cop

2. How does churn vary by geography (France, Germany, Spain)?



Filters

Data

Search

Customer-Churn-Reco...

- Age
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- IsActiveMember
- NumOfProducts
- Point Earned
- RowNumber
- Satisfaction Sco...
- Surname
- Tenure

Page 1 Page 2

Page 2 of 2

73%

Untitled - Power BI Desktop

Search

Sign in

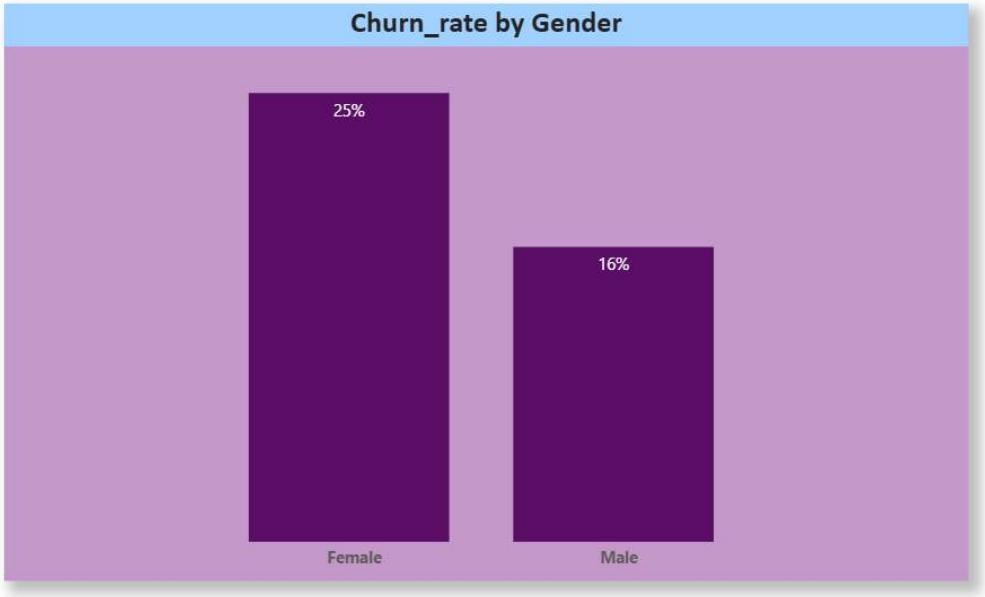
File Home Insert Modeling View Optimize Help

Share

Paste Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Transform data Refresh New visual Text box More visuals New visual calculation New measure Quick measure Sensitivity Publish Prep data f AI

Clipboard Data Queries Insert Calculations Sensitivity Share Cop

3. What are the differences in churn rates between male and female customers?



Filters

Data

Search

Customer-Churn-Reco...

- Age
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- IsActiveMember
- NumOfProducts
- Point Earned
- RowNumber
- Satisfaction Sco...
- Surname
- Tenure

Page 1 Page 2 Page 3

Page 3 of 3

73%

Untitled - Power BI Desktop

Search

Sign in

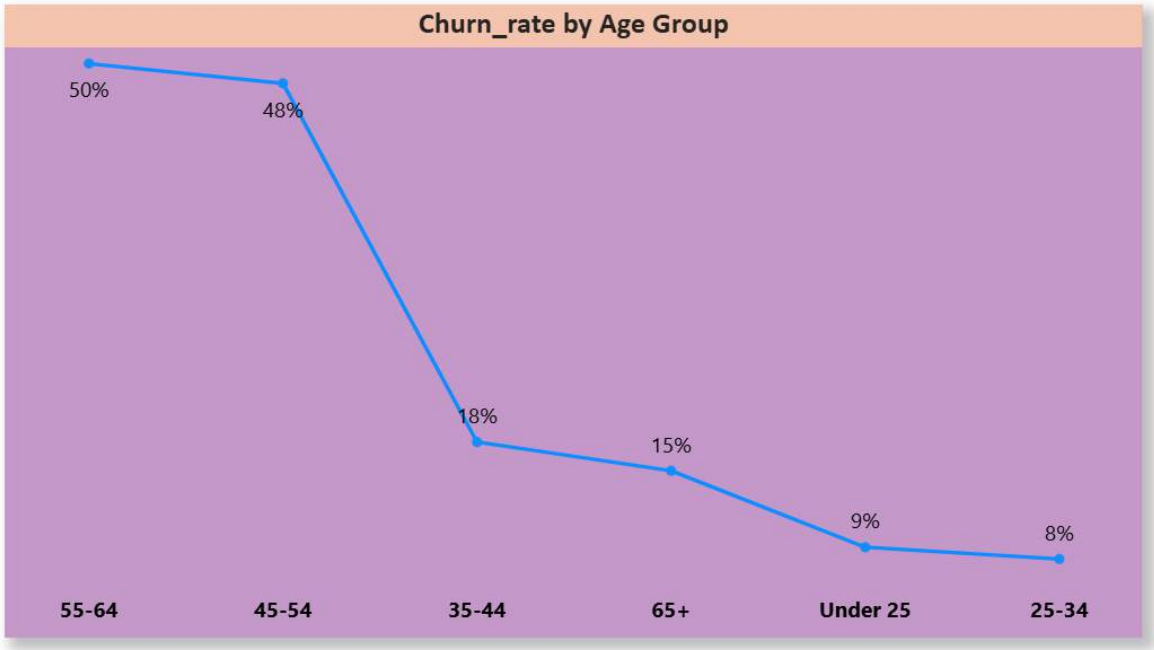
File Home Insert Modeling View Optimize Help

Share

Paste Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Transform data Refresh New visual Text box More visuals New visual calculation New measure Quick measure Sensitivity Publish Prep data f AI

Clipboard Data Queries Insert Calculations Sensitivity Share Cop

4. What is the churn rate distribution across different age groups?



Filters

Data

Search

Customer-Churn-Re...
- Age
- Age Group
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- IsActiveMemb...
- NumOfProdu...
- Point Earned
- RowNumber
- Satisfaction Sc...
- Surname

Page 1 Page 2 Page 3 Page 4

Page 4 of 4

73%

capstone BI • Last saved: Today at 10:54 AM

Search

Sign in

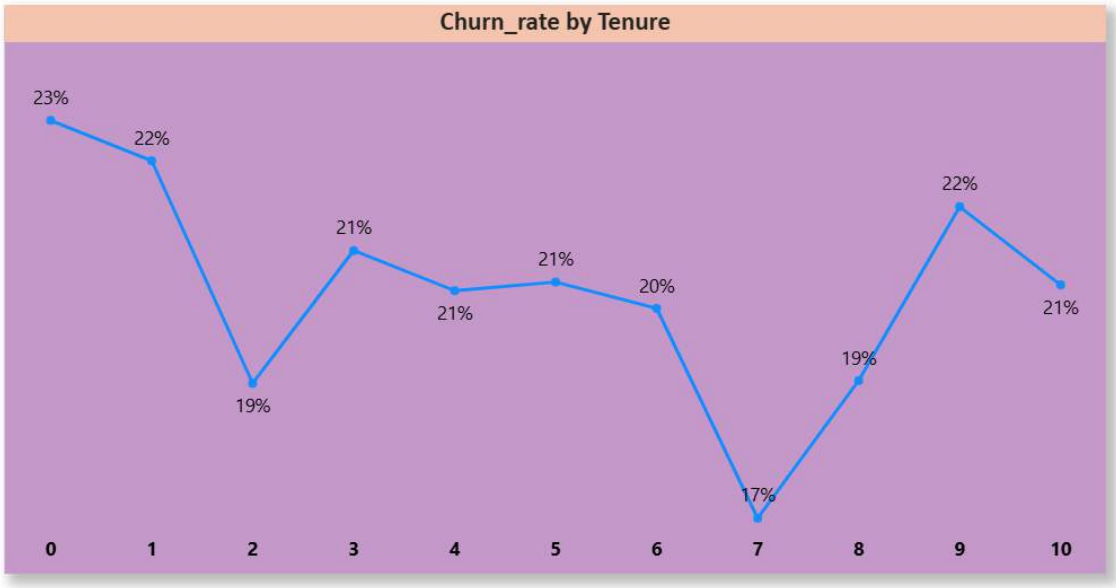
File Home Insert Modeling View Optimize Help

Share

Paste Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Transform data Refresh New visual Text box More visuals New visual calculation New measure Quick measure Sensitivity Publish Prep data f AI

Clipboard Data Queries Insert Calculations Sensitivity Share Cop

5. How does churn behavior change with tenure?



Filters

Format

Search

Page information

Canvas settings

Canvas background

Wallpaper

Filter pane

Filter cards

Data

Search

Customer-Churn-Re...

- Age
- Age Group
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- IsActiveMemb...
- NumOfProdu...
- Point Earned
- RowNumber
- Satisfaction Sc...
- Surname

Page 1 Page 2 Page 3 Page 4 Page 5

Page 5 of 5

63%

Untitled - Power BI Desktop

Search

Sign in

File Home Insert Modeling View Optimize Help

Share

Paste Clipboard

Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Data

Transform data Refresh Queries

New visual Text box More visuals Insert

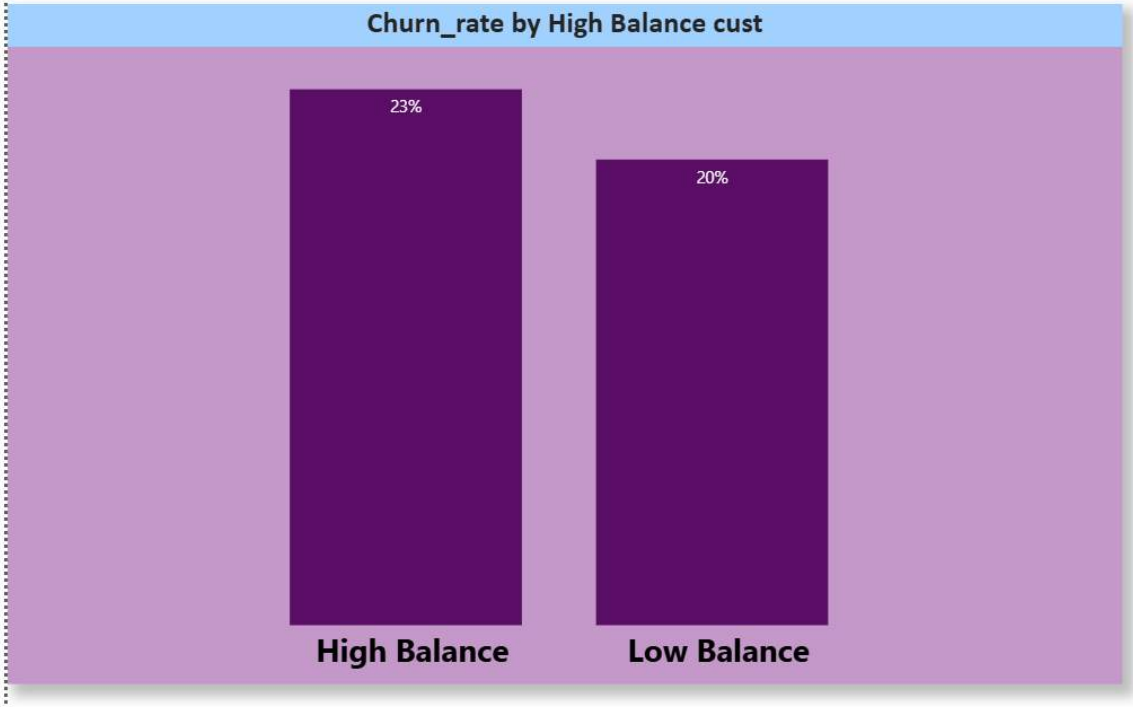
New visual calculation New measure Quick measure Calculations

Sensitivity Sensitivity

Publish Share

Prep data f AI Cop

6. What percentage of high-balance customers (e.g., >$100K) have churned?



High Balance churn(%)

23.12

Filters

Data

Search

Customer-Churn-Re...
- Age
- Age Group
- Avg_Balance
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- High balance ...
- High Balance ...
- IsActiveMemb...
- NumOfProdu...
- Point Earned

Page 1 Page 2 Page 3 Page 4 Page 5 Page 6

Page 6 of 6

73%

Untitled - Power BI Desktop

Search

Sign in

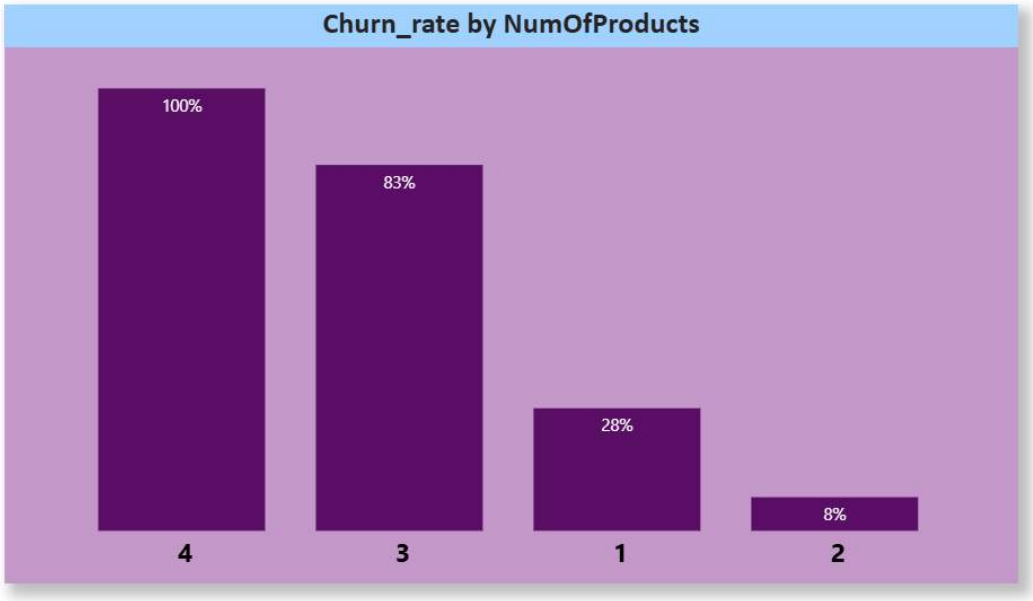
File Home Insert Modeling View Optimize Help

Share

Paste Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Transform data Refresh New visual Text box More visuals New visual calculation New measure Quick measure Sensitivity Publish Prep data f AI

Clipboard Data Queries Insert Calculations Sensitivity Share Cop

7. How do churn rates compare across different NumOfProducts?



Filters

Data

Search

Customer-Churn-Re...
- Age
- Age Group
- Avg_Balance
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- High balance ...
- High Balance ...
- IsActiveMemb...
- NumOfProdu...
- Point Earned

Page 1 Page 2 Page 3 Page 4 Page 5 Page 6 Page 7

Page 7 of 7

73%

Untitled - Power BI Desktop

Search

Sign in

File Home Insert Modeling View Optimize Help

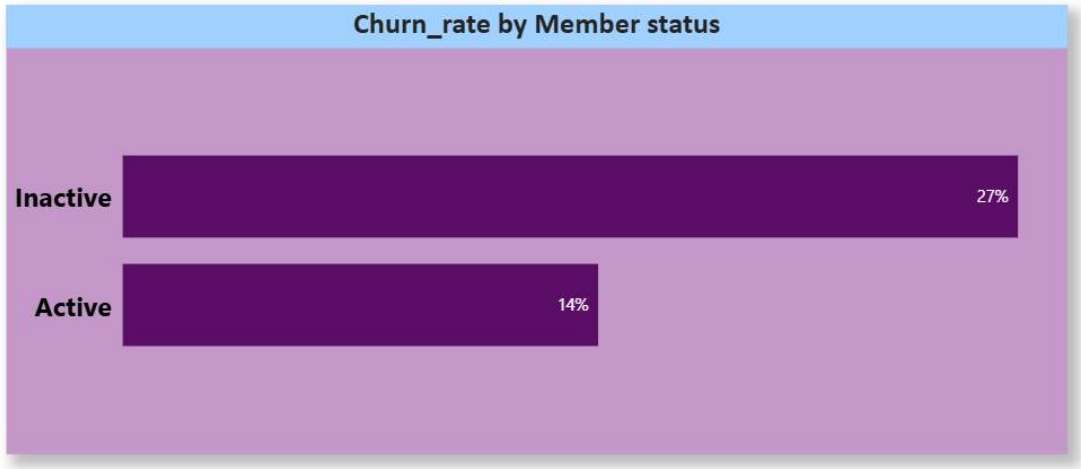
Share

Paste Get data Excel workbook OneLake catalog SQL Server Enter data Dataverse Recent sources Transform data Refresh New visual Text box More visuals New visual calculation New measure Quick measure Sensitivity Publish Prep data f AI

Clipboard Data Queries Insert Calculations Sensitivity Share Cop

8. Are inactive members (IsActiveMember = 0) more likely to churn?

Member status	Churn_rate
Active	14%
Inactive	27%
Total	**20%**



Filters

Data

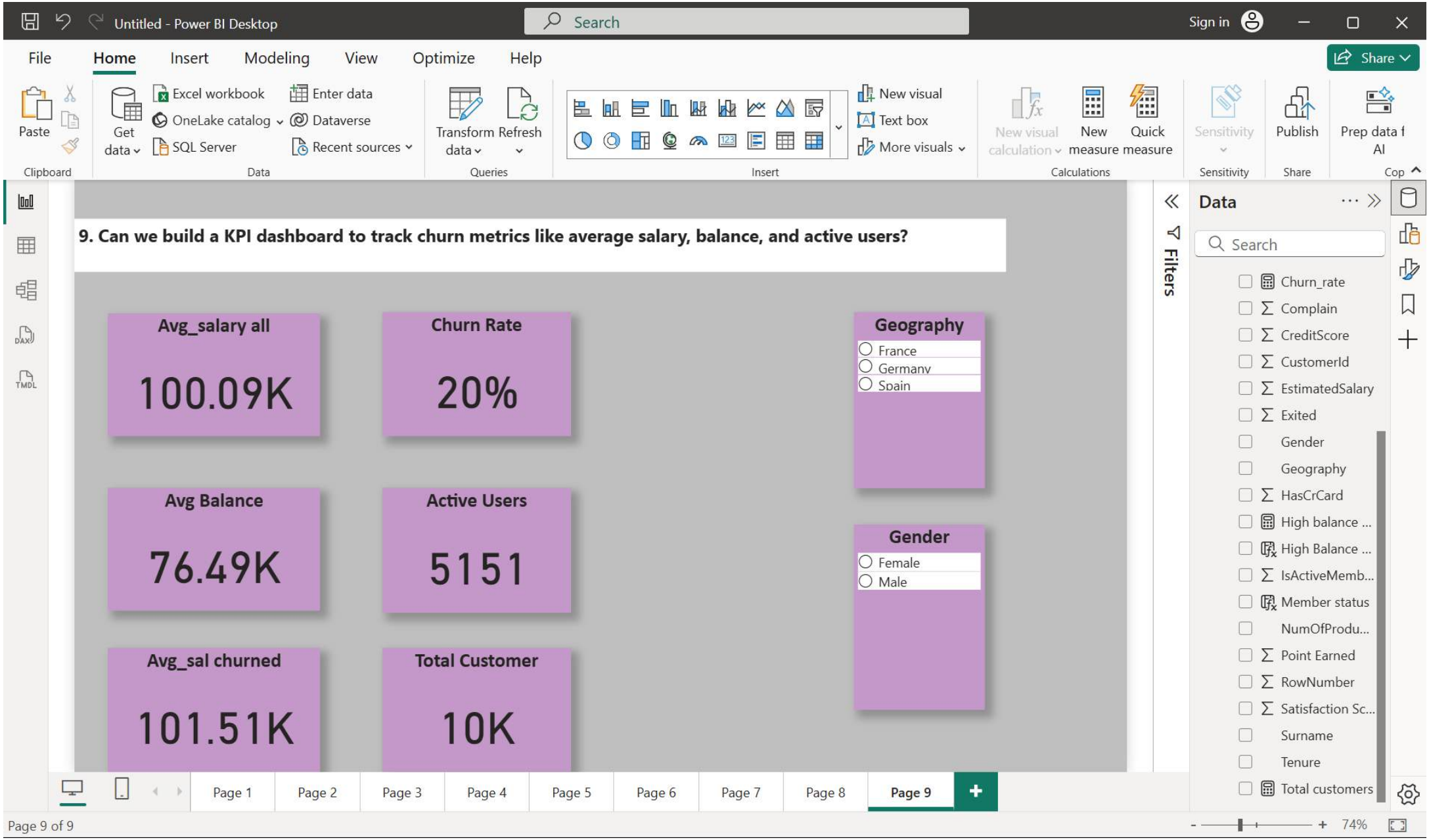
Search

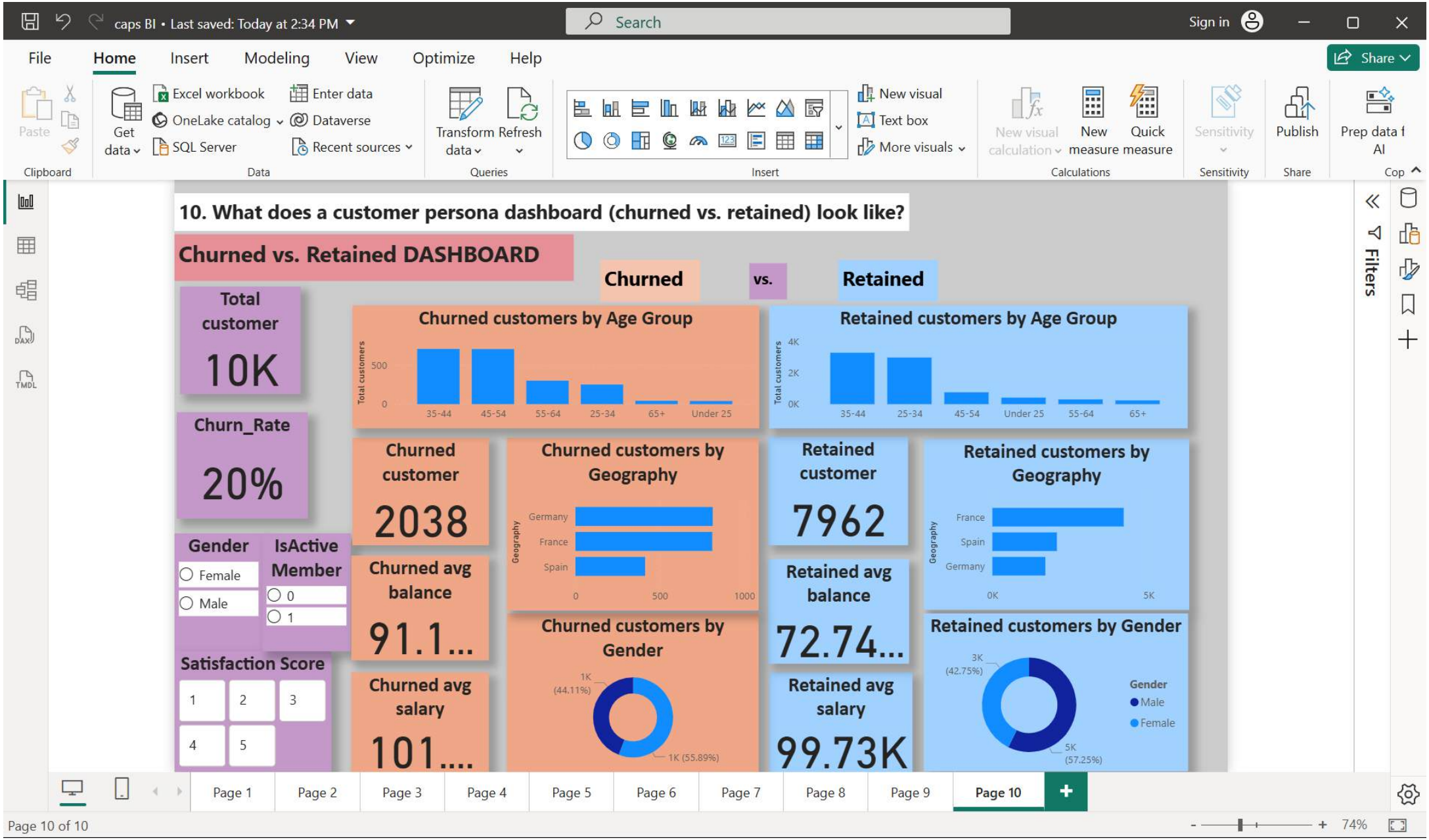
Customer-Churn-Re...
- Age
- Age Group
- Avg_Balance
- Balance
- Card Type
- Churn_rate
- Complain
- CreditScore
- CustomerId
- EstimatedSalary
- Exited
- Gender
- Geography
- HasCrCard
- High balance ...
- High Balance ...
- IsActiveMemb...
- Member status
- NumOfProdu...

Page 1 Page 2 Page 3 Page 4 Page 5 Page 6 Page 7 Page 8

Page 8 of 8

73%


Untitled - Power BI Desktop
Search
Sign in
File
Home
Insert
Modeling
View
Optimize
Help
Share
Paste
Get data
Excel workbook
OneLake catalog
SQL Server
Enter data
Dataverse
Recent sources
Transform data
Refresh
New visual
Text box
More visuals
New visual calculation
New measure
Quick measure
Sensitivity
Publish
Prep data f AI
Clipboard
Data
Queries
Insert
Calculations
Sensitivity
Share
Cop
9. Can we build a KPI dashboard to track churn metrics like average salary, balance, and active users?
Avg_salary all
100.09K
Churn Rate
20%
Avg Balance
76.49K
Active Users
5151
Avg_sal churned
101.51K
Total Customer
10K
Geography
France
Germany
Spain
Gender
Female
Male
Filters
Data
Search
Churn_rate
Complain
CreditScore
CustomerId
EstimatedSalary
Exited
Gender
Geography
HasCrCard
High balance ...
High Balance ...
IsActiveMemb...
Member status
NumOfProdu...
Point Earned
RowNumber
Satisfaction Sc...
Surname
Tenure
Total customers
Page 1
Page 2
Page 3
Page 4
Page 5
Page 6
Page 7
Page 8
Page 9
Page 9 of 9
74%


caps BI • Last saved: Today at 2:34 PM
Search
Sign in
File
Home
Insert
Modeling
View
Optimize
Help
Share
Paste
Get data
Excel workbook
OneLake catalog
SQL Server
Enter data
Dataverse
Recent sources
Transform data
Refresh
New visual
Text box
More visuals
New visual calculation
New measure
Quick measure
Sensitivity
Publish
Prep data f AI
Clipboard
Data
Queries
Insert
Calculations
Sensitivity
Share
Cop
10. What does a customer persona dashboard (churned vs. retained) look like?
Churned vs. Retained DASHBOARD
Churned
vs.
Retained
Total customer
10K
Churn_Rate
20%
Churned customer
2038
Churned avg balance
91.1...
Churned avg salary
101....
Churned customers by Age Group
Total customers
500
0
35-44
45-54
55-64
25-34
65+
Under 25
Churned customers by Geography
Germany
France
Spain
Geography
0
500
1000
Churned customers by Gender
1K (44.11%)
1K (55.89%)
Retained customer
7962
Retained avg balance
72.74...
Retained avg salary
99.73K
Retained customers by Age Group
4K
2K
0K
35-44
25-34
45-54
Under 25
55-64
65+
Retained customers by Geography
France
Spain
Germany
0K
5K
Retained customers by Gender
3K (42.75%)
5K (57.25%)
Gender
Male
Female
Gender
Female
Male
IsActive Member
0
1
Satisfaction Score
1
2
3
4
5
Filters
Page 1
Page 2
Page 3
Page 4
Page 5
Page 6
Page 7
Page 8
Page 9
Page 10
Page 10 of 10
74%

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

[3]:
```
df=pd.read_csv("Customer-churn-Records.csv")
```

[5]:
```
df.head()
```

[5]:

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited
0	1	15634602	Hargrave	619	France	Female	42	2	0.00	1	1	1	101348.88	1
1	2	15647311	Hill	608	Spain	Female	41	1	83807.86	1	0	1	112542.58	0
2	3	15619304	Onio	502	France	Female	42	8	159660.80	3	1	0	113931.57	1
3	4	15701354	Boni	699	France	Female	39	1	0.00	2	0	0	93826.63	0
4	5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1	1	1	79084.10	0

[6]:
```
df.shape
```

[6]:
```
(10000, 18)
```

[7]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 18 columns):
```

memory usage: 1.4+ MB

[ ]:

[ ]:
```
# 11. Check the percentage of missing data and handle accordingly.
```

[8]:
```
df.isnull().sum()
```

[8]:
```
RowNumber            0
CustomerId           0
Surname              0
CreditScore          0
Geography            0
Gender               0
Age                  0
Tenure               0
Balance              0
NumOfProducts        0
HasCrCard            0
IsActiveMember       0
EstimatedSalary      0
Exited               0
Complain             0
Satisfaction Score   0
Card Type            0
Point Earned         0
dtype: int64
```

[ ]:

Code

JupyterLab

Python [conda env:base] *

```
Balance              0
NumOfProducts        0
HasCrCard            0
IsActiveMember       0
EstimatedSalary      0
Exited               0
Complain             0
Satisfaction Score   0
Card Type            0
Point Earned         0
dtype: int64
```

[ ]:

[ ]:
```
# 12. How many rows and columns are there in the dataset?
```

[10]:
```
df.shape
```

[10]:
```
(10000, 18)
```

[13]:
```
rows, cols =df.shape
print(f"The dataset has {rows} rows and {cols} columns.")
```

```
The dataset has 10000 rows and 18 columns.
```

[ ]:

[ ]:

[ ]:

[ ]:
```
#13. What is the distribution of churned vs. non-churned customers?
```

[14]:
```
df['Exited'].value_counts()
```

[14]:
```
Exited
0    7962
1    2038
Name: count, dtype: int64
```

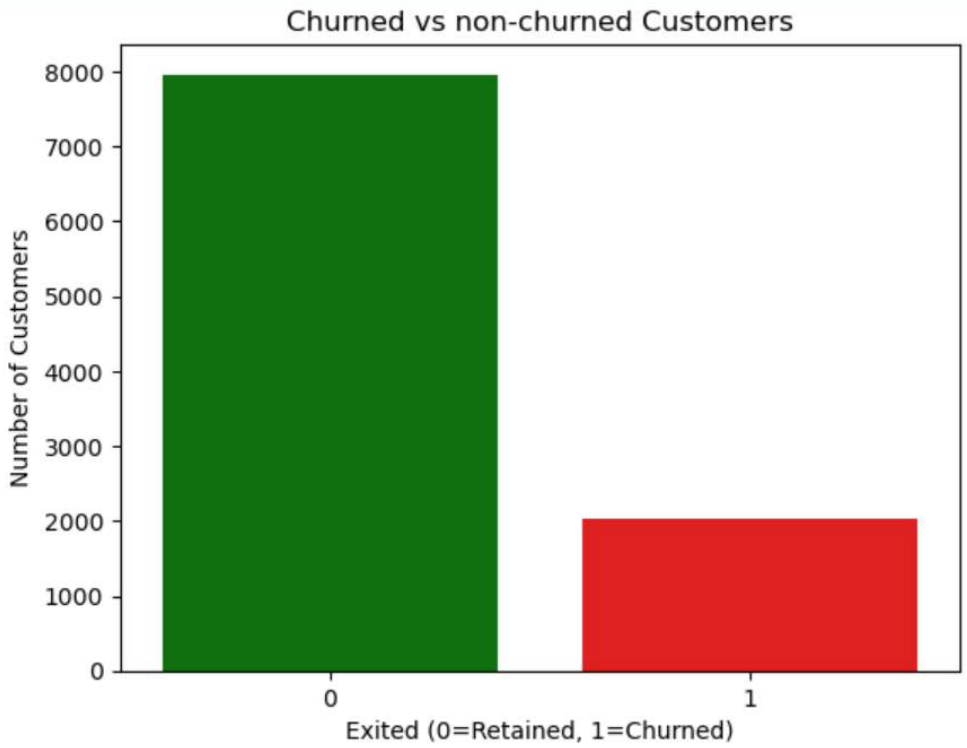
[15]:
```
df['Exited'].value_counts(normalize=True)*100
```

[15]:
```
Exited
0    79.62
1    20.38
Name: proportion, dtype: float64
```

[23]:
```
sns.countplot(x='Exited',data=df,palette=['green','red'])
plt.title('Churned vs non-churned Customers')
plt.xlabel('Exited (0=Retained, 1=Churned)')
plt.ylabel('Number of Customers')
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_19932\3929842554.py:1: FutureWarning:

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

  sns.countplot(x='Exited',data=df,palette=['green','red'])
```

```
sns.countplot(x='Exited',data=df,palette=['green','red'])
```




```
# About 20% customers have Churned , While 80% remain with Bank.
```

[ ]:
```
#14. What is the distribution of EstimatedSalary of churned and retained Customers?
```

[ ]:

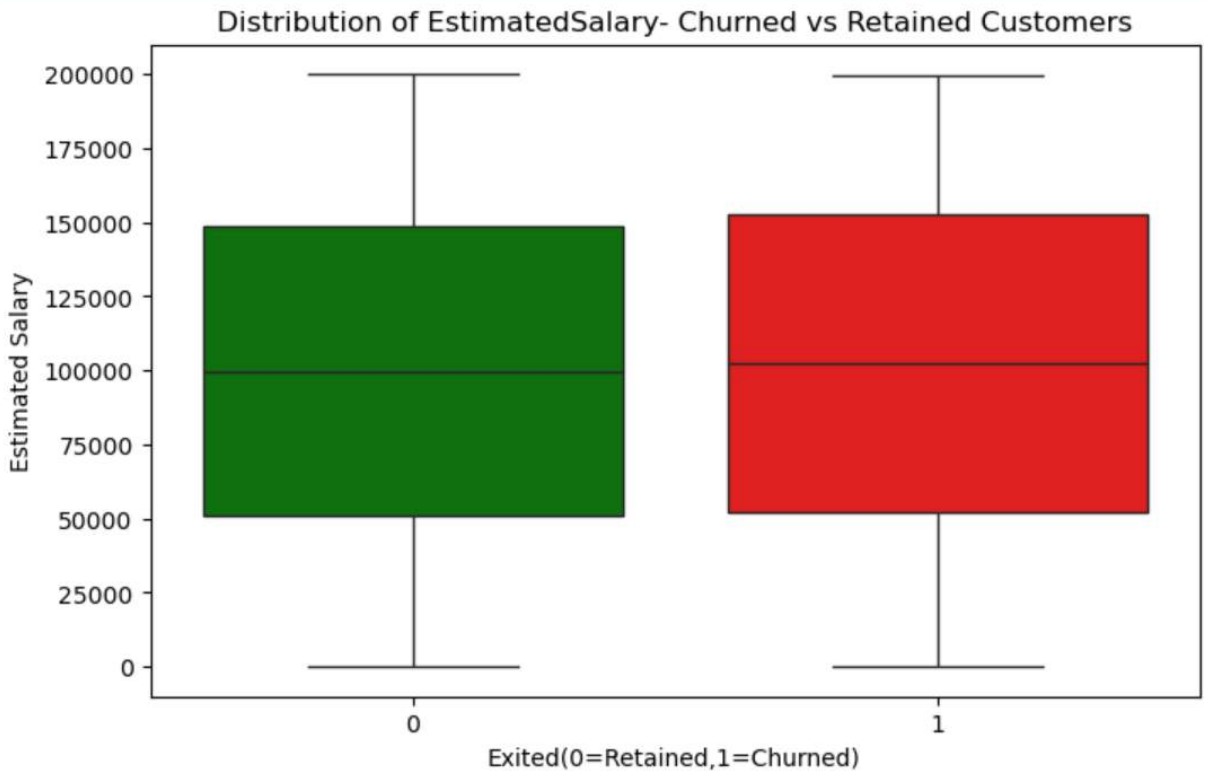
[25]:
```
df.groupby('Exited')['EstimatedSalary'].describe()
```

[25]:

Exited	count	mean	std	min	25%	50%	75%	max
0	7962.0	99726.853141	57399.956717	90.07	50783.295	99620.355	148602.4450	199992.48
1	2038.0	101509.908783	57932.623392	11.58	51924.020	102489.335	152443.8575	199808.10

[26]:
```
plt.figure(figsize=(8,5))
sns.boxplot(x='Exited',y='EstimatedSalary',data=df,palette=['green','red'])
plt.title('Distribution of EstimatedSalary- Churned vs Retained Customers')
plt.xlabel('Exited(0=Retained,1=Churned)')
plt.ylabel('Estimated Salary')
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_19932\3890800965.py:2: FutureWarning:

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

  sns.boxplot(x='Exited',y='EstimatedSalary',data=df,palette=['green','red'])
```



[ ]: `# Both Churned and Retained customers have similar salary distributions,So estimated salary is not a major factor in customer churn for data.`

[ ]:

[ ]:

```
[ ]:
```

```
[20]: # 15. How do churn rates vary by Gender, Geography, and IsActiveMember?
```

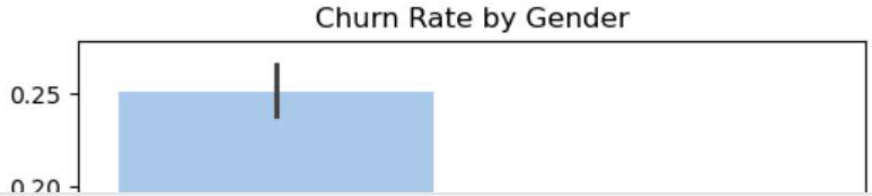
```
[21]: #churn rate vary by Genger
```

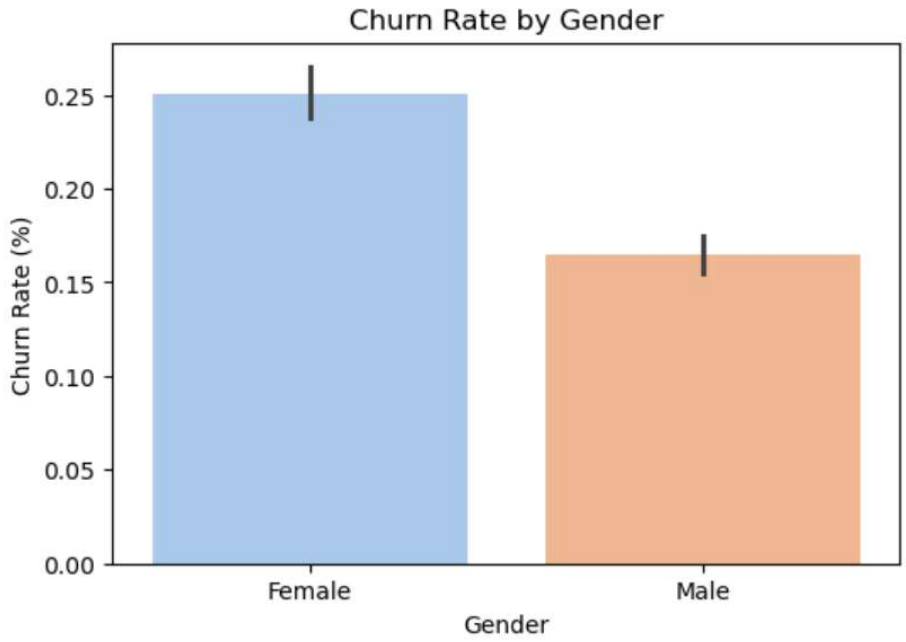
```
[22]: churn_gender=df.groupby('Gender')['Exited'].mean()*100
      print(churn_gender)
```

```
Gender
Female    25.071539
Male      16.474253
Name: Exited, dtype: float64
```

```
[23]: plt.figure(figsize=(6,4))
      sns.barplot(x='Gender',y='Exited',data=df,estimator='mean',palette='pastel')
      plt.title('Churn Rate by Gender')
      plt.ylabel('Churn Rate (%)')
      plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2934609559.py:2: FutureWarning:

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

  sns.barplot(x='Gender',y='Exited',data=df,estimator='mean',palette='pastel')
```





[24]:
```
#Female Customers have a higher churn Rate.
```

[25]:
```
 #churn rate vary by Geography
```

[30]:
```
churn_geography=df.groupby('Geography')['Exited'].mean()*100
print(churn_geography)
```

```
Geography
France     16.174711
Germany    32.443204
```

Code JupyterLab Python [conda env:base] *

[25]:
```
#churn rate vary by Geography
```

[30]:
```
churn_geography=df.groupby('Geography')['Exited'].mean()*100
print(churn_geography)
```

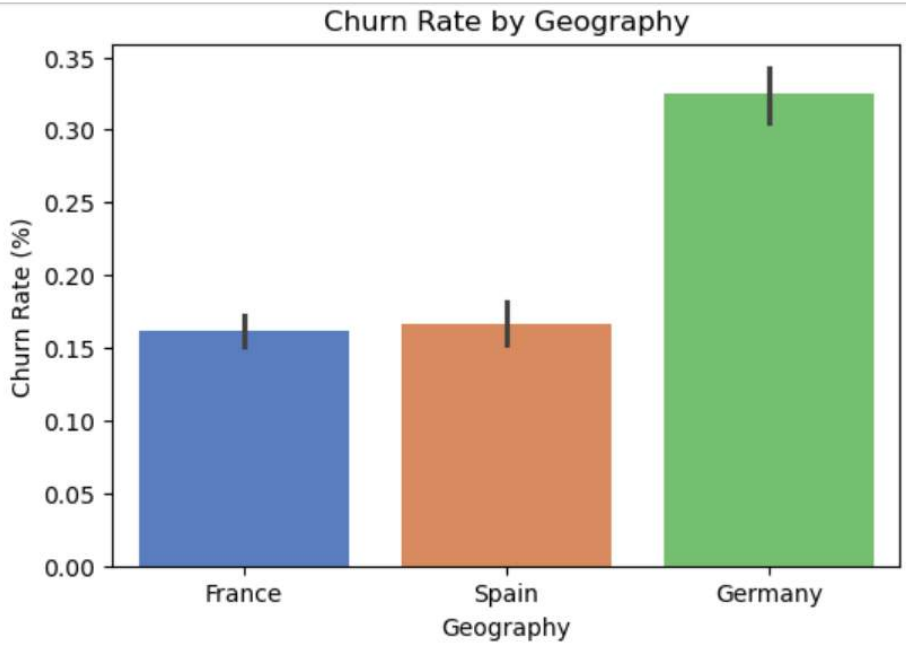
```
Geography
France     16.174711
Germany    32.443204
Spain      16.673395
Name: Exited, dtype: float64
```

[26]:
```
plt.figure(figsize=(6,4))
sns.barplot(x='Geography',y='Exited',data=df,estimator='mean',palette='muted')
plt.title('Churn Rate by Geography')
plt.ylabel('Churn Rate (%)')
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\707093323.py:2: FutureWarning:

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

  sns.barplot(x='Geography',y='Exited',data=df,estimator='mean',palette='muted')
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *



[27]:
```
#Churn rate is highest among German Customers.
```

[28]:
```
 #churn Rate by Active Membership
```

[31]:
```
churn_Activemembership=df.groupby('IsActiveMember')['Exited'].mean()*100
print(churn_Activemembership)
```

```
IsActiveMember
0    26.871520
1    14.269074
```

Capstone Last Checkpoint: yesterday

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

[28]:
```
#churn Rate by Active Membership
```

[31]:
```
churn_Activemembership=df.groupby('IsActiveMember')['Exited'].mean()*100
print(churn_Activemembership)
```

```
IsActiveMember
0    26.871520
1    14.269074
Name: Exited, dtype: float64
```

[29]:
```
plt.figure(figsize=(6,4))
sns.barplot(x='IsActiveMember',y='Exited',data=df,estimator='mean',palette='coolwarm')
plt.title('Churn Rate by Active Membership')
plt.ylabel('Churn Rate (%)')
plt.xticks([0,1],['Inactive','Active'])
plt.show()
```

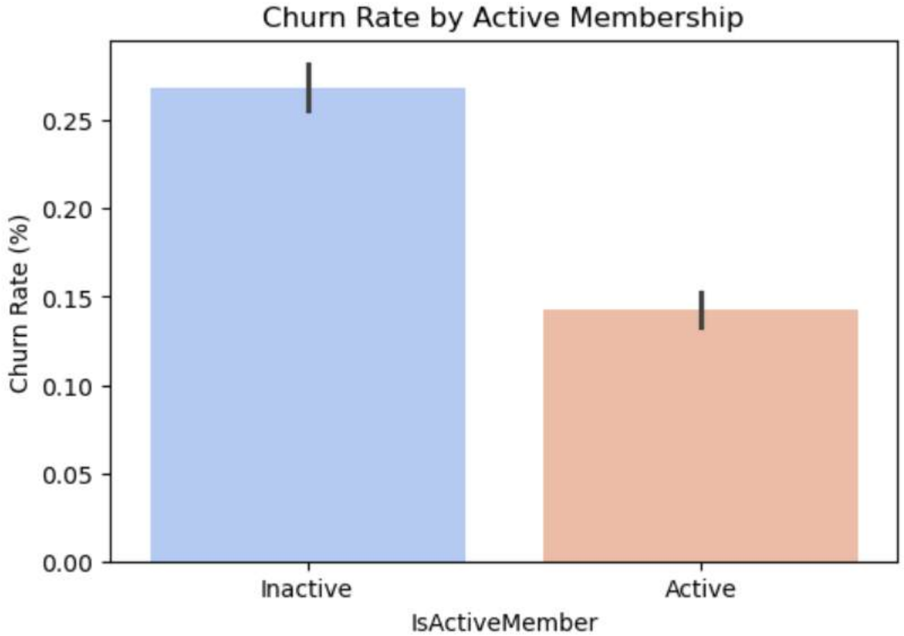
```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2505570173.py:2: FutureWarning:

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

  sns.barplot(x='IsActiveMember',y='Exited',data=df,estimator='mean',palette='coolwarm')
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

```
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2505570173.py:2: FutureWarning:

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

  sns.barplot(x='IsActiveMember',y='Exited',data=df,estimator='mean',palette='coolwarm')
```



[ ]:

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

[ ]:

[ ]:
```
# 16. What is the average CreditScore, Balance, and EstimatedSalary of churned vs. retained customers?
```

[33]:
```
df.head()
```

[33]:

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited
0	1	15634602	Hargrave	619	France	Female	42	2	0.00	1	1	1	101348.88	1
1	2	15647311	Hill	608	Spain	Female	41	1	83807.86	1	0	1	112542.58	0
2	3	15619304	Onio	502	France	Female	42	8	159660.80	3	1	0	113931.57	1
3	4	15701354	Boni	699	France	Female	39	1	0.00	2	0	0	93826.63	0
4	5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1	1	1	79084.10	0

[34]:
```
avg_values=df.groupby('Exited')[['CreditScore','Balance','EstimatedSalary']].mean()
print(avg_values)
```

```
        CreditScore       Balance  EstimatedSalary
Exited
0        651.837855  72742.750663     99726.853141
1        645.414622  91109.476006    101509.908783
```

[ ]:
```
# Note: 0 is for retained and 1 is for Churned
```

[ ]:

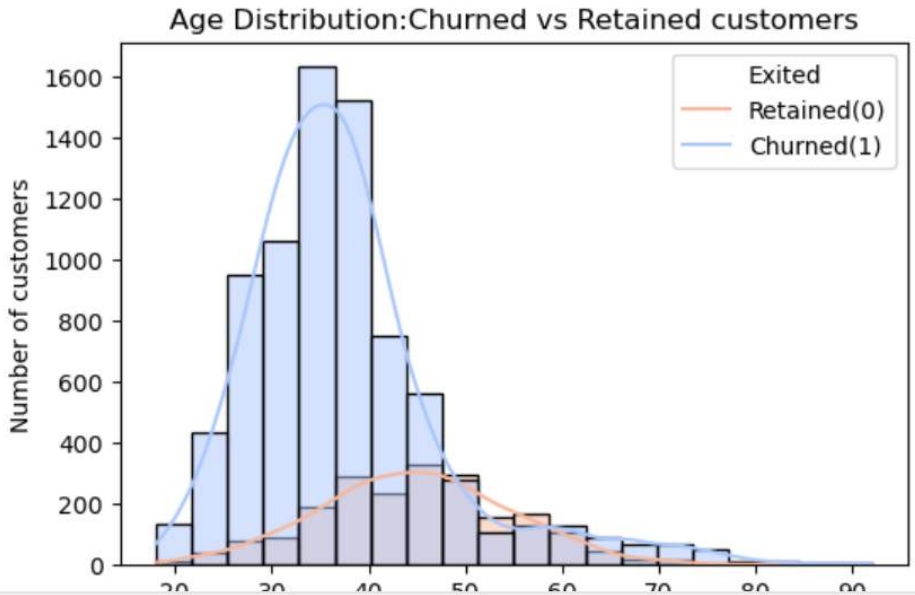
[ ]:
```
17. #How does Age impact churn? Plot histograms and boxplots for churned and non-churned groups.
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

[ ]: 17. #How does Age impact churn? Plot histograms and boxplots for churned and non-churned groups.

•[35]:
```
#Histogram
plt.figure(figsize=(6,4))
sns.histplot(data=df,x='Age',hue='Exited',kde=True,bins=20,palette='coolwarm')
plt.title('Age Distribution:Churned vs Retained customers')
plt.xlabel('Age')
plt.ylabel('Number of customers')
plt.legend(title='Exited',labels=['Retained(0)','Churned(1)'])
plt.show()
```

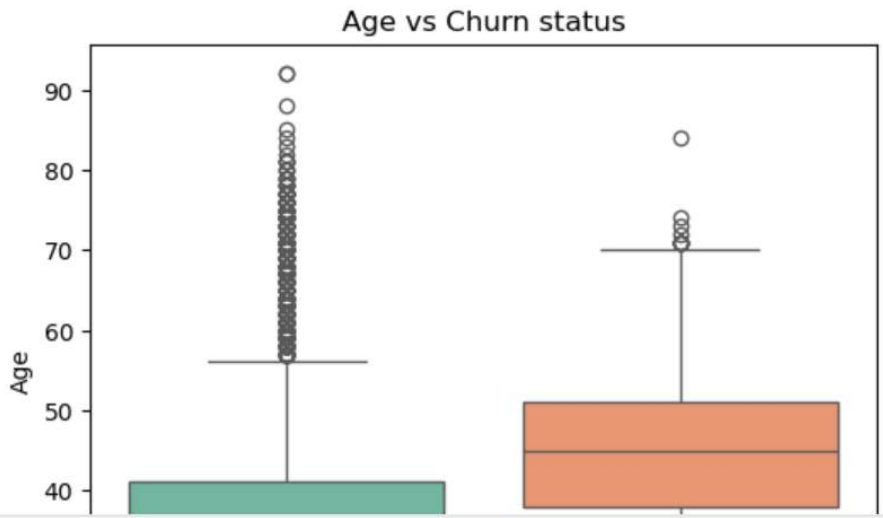


File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

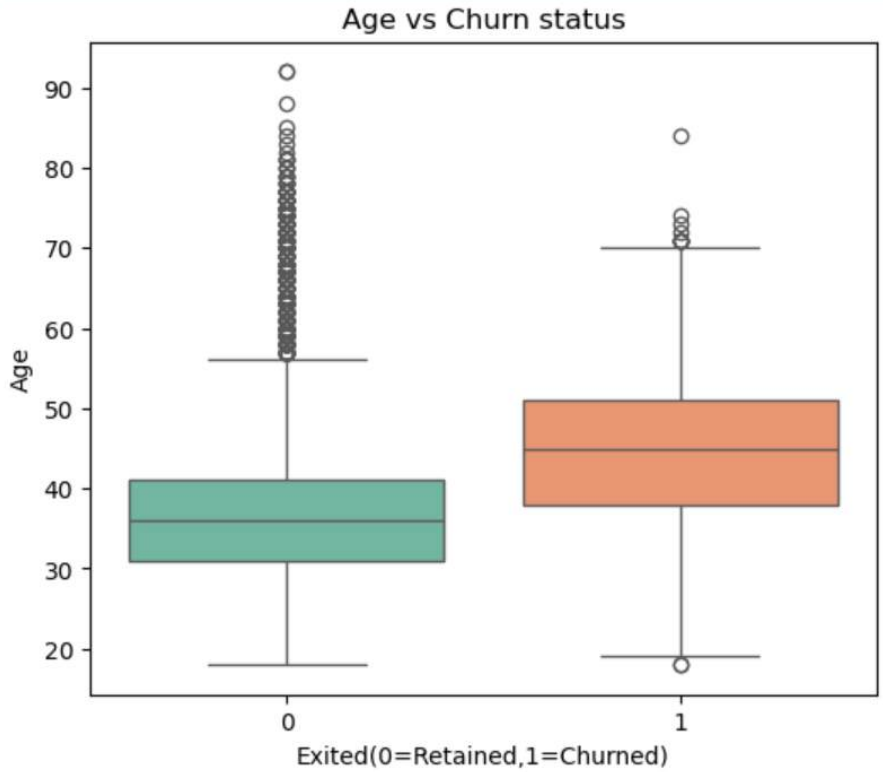
[37]:
```
#Boxplot
plt.figure(figsize=(6,5))
sns.boxplot(x='Exited',y='Age',data=df,palette='Set2')
plt.title('Age vs Churn status')
plt.xlabel('Exited(0=Retained,1=Churned)')
plt.ylabel('Age')
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2785628990.py:3: FutureWarning:

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

  sns.boxplot(x='Exited',y='Age',data=df,palette='Set2')
```



File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

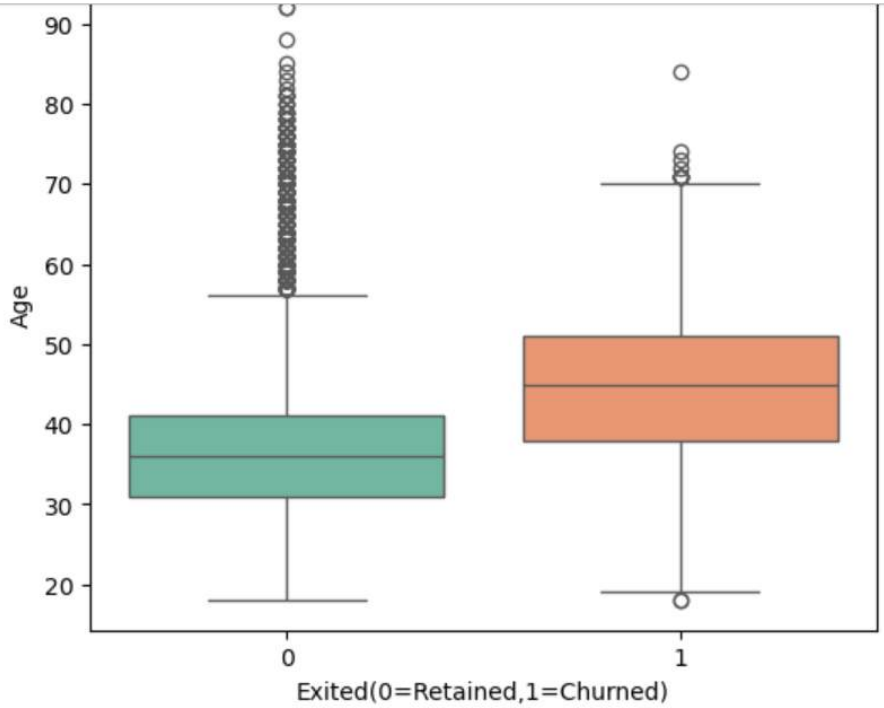


[39]:
```
df.groupby('Exited')['Age'].mean()
```

[39]:
```
Exited
0    37.408063
1    44.835623
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *



[39]:
```
df.groupby('Exited')['Age'].mean()
```

[39]:
```
Exited
0    37.408063
1    44.835623
Name: Age, dtype: float64
```

[ ]:
```
#Churned customers have a average age og 44.8 compared to 37.4 for Retained.Older Customers are more likely to churn.
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

```
1    44.835623
Name: Age, dtype: float64
```

[ ]:
```
#Churned customers have a average age og 44.8 compared to 37.4 for Retained.Older Customers are more likely to churn.
```

[ ]:

[ ]:
```
#18. Is there any correlation among numeric features like CreditScore, Balance, and EstimatedSalary?
```

[43]:
```
#select numeric columns
num_cols=['CreditScore','Balance','EstimatedSalary']
#calculate correlation matrix
corr_matrix=df[num_cols].corr()
print(corr_matrix)
```

```
                CreditScore   Balance  EstimatedSalary
CreditScore        1.000000  0.006268        -0.001384
Balance            0.006268  1.000000         0.012797
EstimatedSalary   -0.001384  0.012797         1.000000
```

[ ]:

[44]:
```
plt.figure(figsize=(6,5))
sns.heatmap(corr_matrix,annot=True,cmap='coolwarm',fmt='.2f')
plt.title('correlation among CreditScore, Balance, and EstimatedSalary')
plt.show()
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

[44]:
```
plt.figure(figsize=(6,5))
sns.heatmap(corr_matrix,annot=True,cmap='coolwarm',fmt='.2f')
plt.title('correlation among CreditScore, Balance, and EstimatedSalary')
plt.show()
```

Code JupyterLab Python [conda env:base] *

[ ]:
```
#19. What does a heatmap reveal about feature interactions with churn?
```

[49]:
```
Num_cols=['CreditScore','Age','Tenure','Balance','NumOfProducts','HasCrCard','IsActiveMember','EstimatedSalary','Exited']

corr=df[Num_cols].corr()
print(corr)
```

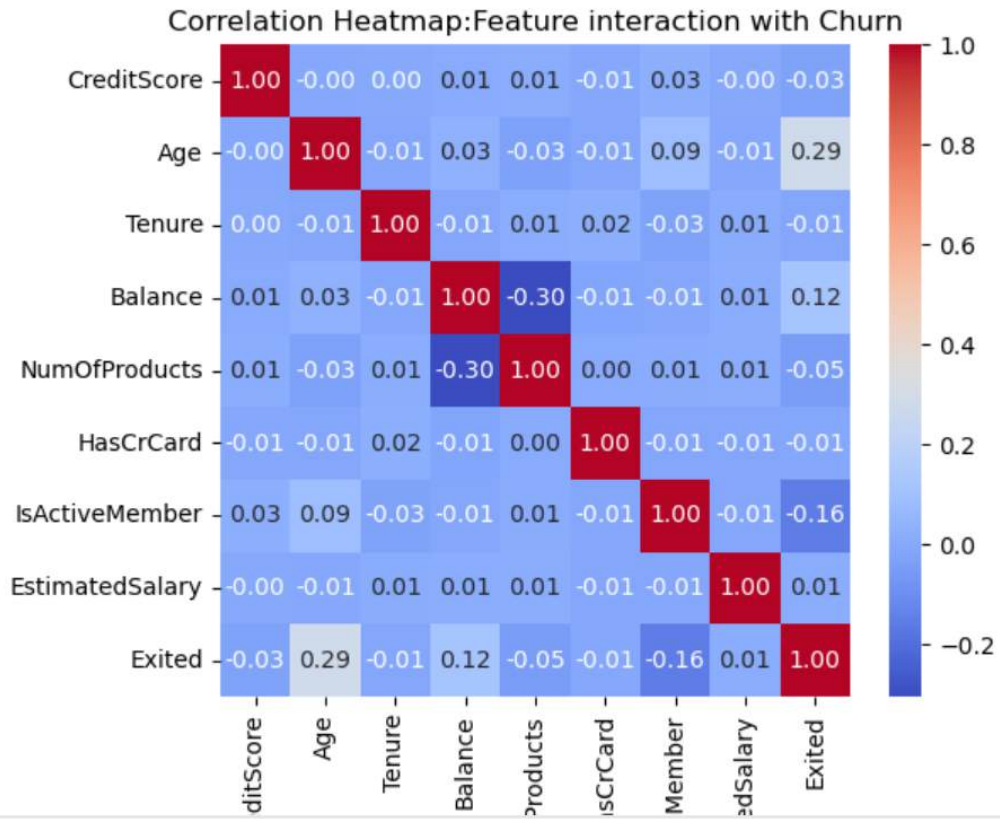
```
                 CreditScore       Age    Tenure   Balance  NumOfProducts  \
CreditScore         1.000000 -0.003965  0.000842  0.006268       0.012238
Age                -0.003965  1.000000 -0.009997  0.028308      -0.030680
Tenure              0.000842 -0.009997  1.000000 -0.012254       0.013444
Balance             0.006268  0.028308 -0.012254  1.000000      -0.304180
NumOfProducts       0.012238 -0.030680  0.013444 -0.304180       1.000000
HasCrCard          -0.005458 -0.011721  0.022583 -0.014858       0.003183
IsActiveMember      0.025651  0.085472 -0.028362 -0.010084       0.009612
EstimatedSalary    -0.001384 -0.007201  0.007784  0.012797       0.014204
Exited             -0.026771  0.285296 -0.013656  0.118577      -0.047611

                 HasCrCard  IsActiveMember  EstimatedSalary    Exited
CreditScore      -0.005458        0.025651        -0.001384 -0.026771
Age              -0.011721        0.085472        -0.007201  0.285296
Tenure            0.022583       -0.028362         0.007784 -0.013656
Balance          -0.014858       -0.010084         0.012797  0.118577
NumOfProducts     0.003183        0.009612         0.014204 -0.047611
HasCrCard         1.000000       -0.011866        -0.009933 -0.006976
IsActiveMember   -0.011866        1.000000        -0.011421 -0.156356
EstimatedSalary  -0.009933       -0.011421         1.000000  0.012490
Exited           -0.006976       -0.156356         0.012490  1.000000
```

[52]:
```
plt.figure(figsize=(6,5))
sns.heatmap(corr,annot=True,cmap='coolwarm',fmt='.2f')
plt.title('Correlation Heatmap:Feature interaction with Churn')
```

File Edit View Run Kernel Settings Help Trusted

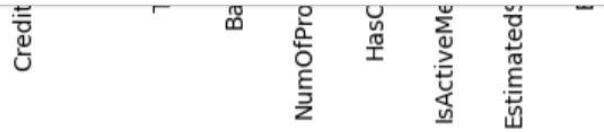
Code JupyterLab Python [conda env:base] *

[52]:
```
plt.figure(figsize=(6,5))
sns.heatmap(corr,annot=True,cmap='coolwarm',fmt='.2f')
plt.title('Correlation Heatmap:Feature interaction with Churn')
plt.show()
```



File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *



[ ]:
```
'''Age and Balance show a positive correlation with Churn, indicating that older customers and with higher account balance are more likely to leave the B
IsActivemember and Numofproducts have a negative correlation with churn, means customers who are active are less likely to Churn.
EstimatedSalary and Tenure show almost no correlation with churn,meaning they have minimal influence on wheather a customer stay or leave.'''
```

[ ]:

[ ]:
```
#20. Are there outliers in Balance, CreditScore, or Age that are mostly associated with churn?
```

[54]:
```
numeric_cols=['CreditScore','Age','Balance']
plt.figure(figsize=(15,5))
for i, col in enumerate(numeric_cols,1):
    plt.subplot(1,3,i)
    sns.boxplot(x='Exited',y=col,data=df,palette='Set2')
    plt.title(f'cols vs Churn')
plt.tight_layout()
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2120548627.py:5: FutureWarning:

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

  sns.boxplot(x='Exited',y=col,data=df,palette='Set2')
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2120548627.py:5: FutureWarning:

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

File Edit View Run Kernel Settings Help

Trusted

Code

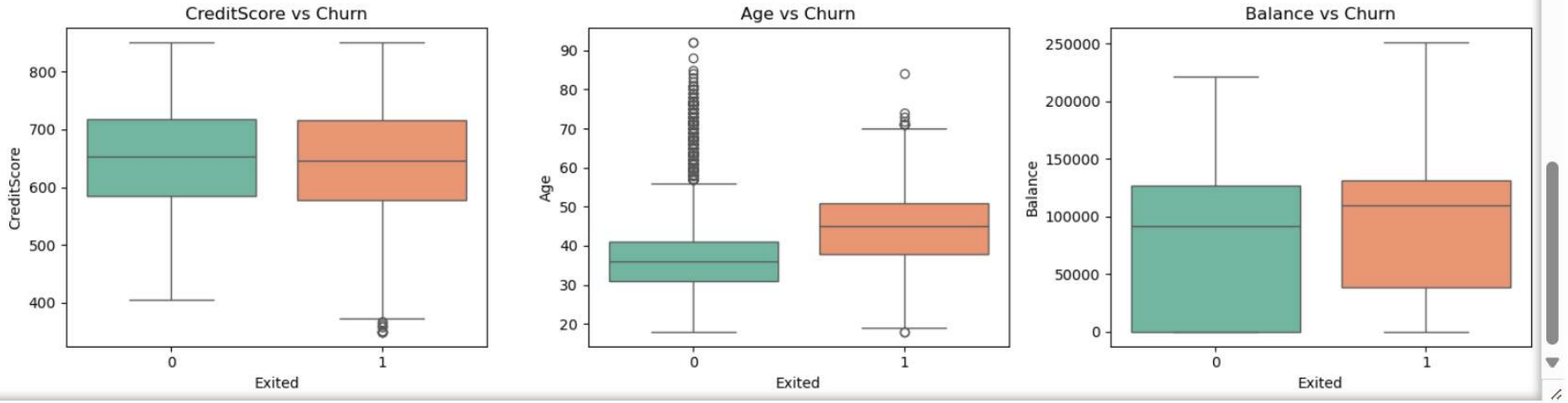
JupyterLab

Python [conda env:base] *

[114]:
```
#20. Are there outliers in Balance, CreditScore, or Age that are mostly associated with churn?
```

[122]:
```
numeric_cols=['CreditScore','Age','Balance']
plt.figure(figsize=(15,4))
for i, col in enumerate(numeric_cols,1):
    plt.subplot(1,3,i)
    sns.boxplot(x='Exited',y=col,data=df,palette='Set2')
    plt.title(f'{col} vs Churn')
plt.tight_layout()
plt.show()
```

```
sns.boxplot(x='Exited',y=col,data=df,palette='Set2')
```



[ ]:

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab

Python [conda env:base] *

[173]:
```
#21. Group customers into age brackets (e.g.,18-30 as Adults, 30-50 as middle age and 50-100 as seniors.). How does churn rate change across them?
```

[174]:
```
bins=[18,30,50,100]
labels=['Adult(18-30)','Middle Age(30-50)','Senior(50-100)']
df['AgeGroups']=pd.cut(df['Age'],bins=bins,labels=labels,right=False)
```

[175]:
```
df.head()
```

[175]:

ıgraphy	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited	Complain	Satisfaction Score	Card Type	Point Earned	AgeGroups
France	Female	42	2	0.00	1	1	1	101348.88	1	1	2	DIAMOND	464	Middle Age(30-50)
Spain	Female	41	1	83807.86	1	0	1	112542.58	0	1	3	DIAMOND	456	Middle Age(30-50)
France	Female	42	8	159660.80	3	1	0	113931.57	1	1	3	DIAMOND	377	Middle Age(30-50)
France	Female	39	1	0.00	2	0	0	93826.63	0	0	5	GOLD	350	Middle Age(30-50)
Spain	Female	43	2	125510.82	1	1	1	79084.10	0	0	5	GOLD	425	Middle Age(30-50)

[176]:
```
Churn_by_agegroup=df.groupby('AgeGroups')['Exited'].mean()*100
print(Churn_by_agegroup)
```

```
AgeGroups
Adult(18-30)          7.556368
Middle Age(30-50)    18.380241
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

Age(30-50)

[176]:
```
Churn_by_agegroup=df.groupby('AgeGroups')['Exited'].mean()*100
print(Churn_by_agegroup)
```

```
AgeGroups
Adult(18-30)          7.556368
Middle Age(30-50)    18.380241
Senior(50-100)       45.448029
Name: Exited, dtype: float64
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2116857766.py:1: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  Churn_by_agegroup=df.groupby('AgeGroups')['Exited'].mean()*100
```

[177]:
```
#Visulize churn rate by Age groups
```

[178]:
```
plt.figure(figsize=(6,4))
sns.barplot(x=Churn_by_agegroup.index,y=Churn_by_agegroup.values,palette='Set2')
plt.title('Churn Rate by AgeGroup')
plt.ylabel('Churn Rate(%)')
plt.xlabel('Age Group')
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\1509556778.py:2: FutureWarning:

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

  sns.barplot(x=Churn_by_agegroup.index,y=Churn_by_agegroup.values,palette='Set2')
```

jupyter Capstone Last Checkpoint: 3 minutes ago

File Edit View Run Kernel Settings Help

Trusted

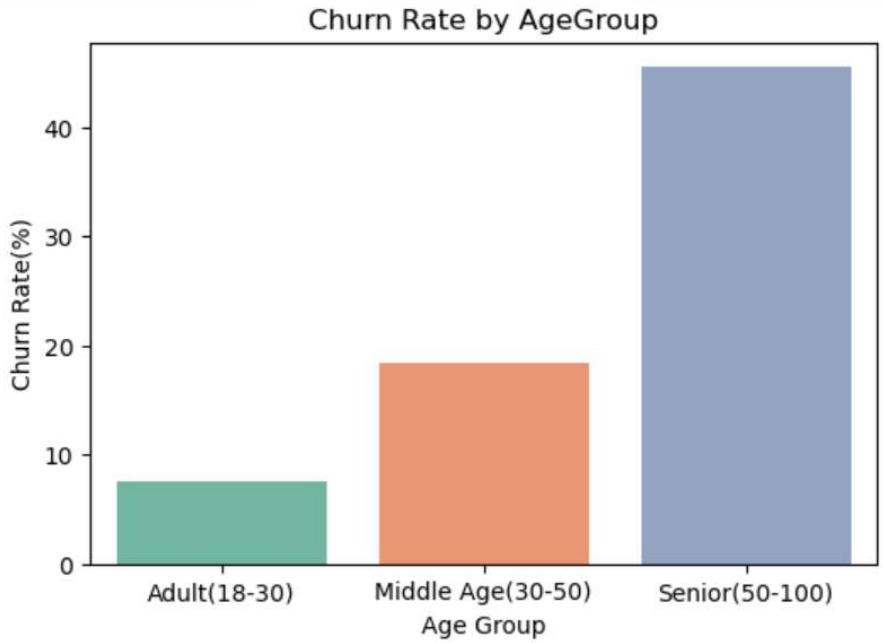
Code

JupyterLab

Python [conda env:base] *

```
plt.show()
```

```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\1509556778.py:2: FutureWarning:

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

  sns.barplot(x=Churn_by_agegroup.index,y=Churn_by_agegroup.values,palette='Set2')
```



[ ]:

Code JupyterLab Python [conda env:base] *

[ ]:

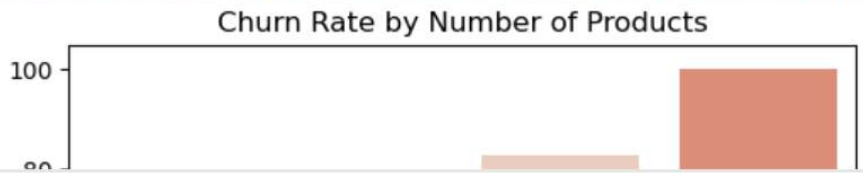
[179]:
```
#22. Are customers with only one product (NumOfProducts = 1) more likely to churn than those with multiple?
```

[180]:
```
Churn_by_products=df.groupby('NumOfProducts')['Exited'].mean()*100
print(Churn_by_products)
```

```
NumOfProducts
1     27.714398
2      7.603486
3     82.706767
4    100.000000
Name: Exited, dtype: float64
```

[181]:
```
plt.figure(figsize=(6,4))
sns.barplot(x=Churn_by_products.index,y=Churn_by_products.values,palette='coolwarm')
plt.title('Churn Rate by Number of Products')
plt.xlabel('Number of Products')
plt.xlabel('Churn Rate(%)')
plt.show()
```

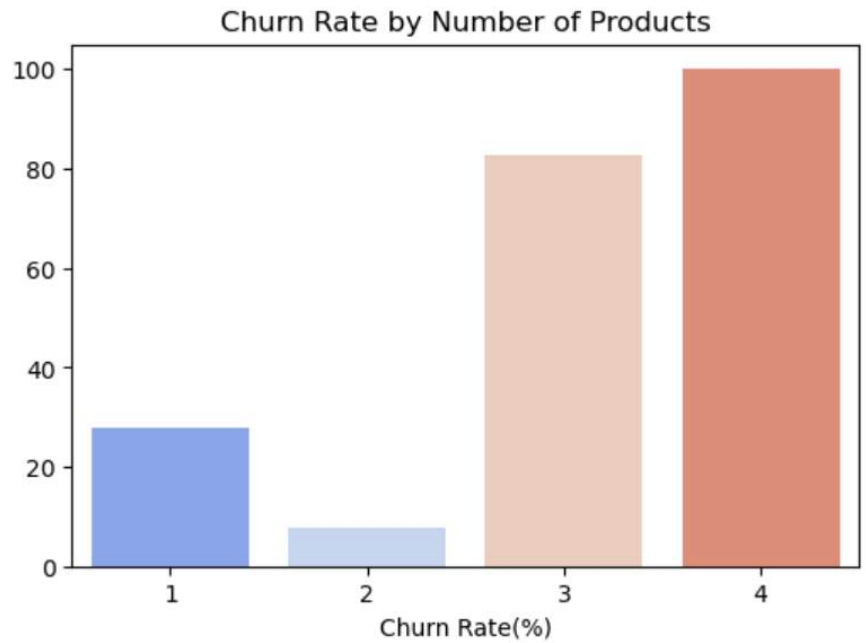
```
C:\Users\nayak\AppData\Local\Temp\ipykernel_6884\2976775891.py:2: FutureWarning:

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

  sns.barplot(x=Churn_by_products.index,y=Churn_by_products.values,palette='coolwarm')
```



File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *

```
sns.barplot(x=Churn_by_products.index,y=Churn_by_products.values,palette='coolwarm')
```

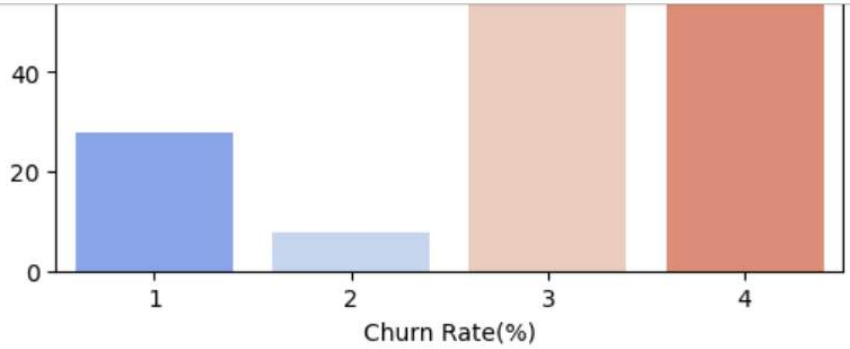


[182]:
```
df['NumOfProducts'].value_counts()
```

[182]:
```
NumOfProducts
1    5084
2    4590
3     266
4      60
Name: count, dtype: int64
```

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python [conda env:base] *



[182]:
```
df['NumOfProducts'].value_counts()
```

[182]:
```
NumOfProducts
1    5084
2    4590
3     266
4      60
Name: count, dtype: int64
```

[184]:
```
'''ChurnRate decreases from 1 to 2 products but spikes sharply for 3 and 4 products.That means moderate engagement(2 products) promotes loyalty.
and under-engagement(1 product)and over-engagement(3 and 4 products)may lead to customer dissatisfaction.'''
```

[184]:
```
'ChurnRate decreases from 1 to 2 products but spikes sharply for 3 and 4 products.That means moderate engagement(2 products) promotes loyalty.\nand under
-engagement(1 product)and over-engagement(3 and 4 products)may lead to customer dissatisfaction.'
```

[ ]:

[ ]:

[ ]: