This data is from Kaggle, used to predict churn. The origin is https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers/data

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
# Load bank churners data
churners_data = pd.read_csv('BankChurners.csv')
# Display the first few rows of the dataframe to understand its
structure
print("First few rows of the churners data:")
print(churners data.head())
First few rows of the churners data:
   CLIENTNUM
                 Attrition Flag Customer Age Gender Dependent count
0
  768805383
                                            45
                                                                     3
              Existing Customer
                                                    М
                                                                     5
1 818770008
              Existing Customer
                                            49
  713982108
             Existing Customer
                                            51
                                                    М
                                                                     3
  769911858
              Existing Customer
                                            40
                                                                     4
                                                                     3
                                            40
4 709106358
              Existing Customer
                                                    М
  Education_Level Marital_Status Income_Category Card_Category \
      High School
0
                         Married
                                      $60K - $80K
                                                           Blue
1
         Graduate
                          Single
                                 Less than $40K
                                                           Blue
2
                                    $80K - $120K
         Graduate
                         Married
                                                           Blue
3
      High School
                         Unknown
                                  Less than $40K
                                                           Blue
       Uneducated
                         Married
                                      $60K - $80K
                                                           Blue
                        Credit Limit Total Revolving Bal
   Months on book
Avg_Open_To_Buy \
               39
                             12691.0
                                                       777
11914.0
               44
                              8256.0
                                                       864
7392.0
                                                         0
               36
                              3418.0
3418.0
               34
                              3313.0
                                                      2517
796.0
               21
                              4716.0
                                                         0
4716.0
```

Total_Amt Total Ct Chn		otal_Trans_Amt	Total_Trans_	Ct
0	1.335	1144		42
1.625 1	1.541	1291		33
3.714 2	2.594	1887		20
2.333	1.405	1171		20
2.333				
4 2.500	2.175	816		28
Avg_Utili 0 1 2 3	zation_Ratio 0.061 0.105 0.000 0.760 0.000	\		
		rition_Flag_Car ion_Level_Month		ntacts_Count_12_ _mon_1 \
1			0.000057	
2			0.000021	
3			0.000134	
4			0.000022	
		rition_Flag_Car ion_Level_Month		ntacts_Count_12_ 2_mon_2
1			0.99994	
2			0.99998	
3			0.99987	
4			0.99998	
[5 rows x 23	columns]			

#Statistics Summary

```
# Summary statistics of numerical columns
print("Summary statistics of numerical columns:")
print(churners data.describe())
Summary statistics of numerical columns:
                      Customer Age
                                     Dependent count
                                                       Months on book
          CLIENTNUM
       1.012700e+04
                      10127.000000
                                        10127.000000
                                                         10127.000000
count
       7.391776e+08
                         46.325960
                                                            35.928409
                                            2.346203
mean
                                             1.298908
std
       3.690378e+07
                          8.016814
                                                              7.986416
       7.080821e+08
                         26.000000
                                            0.00000
                                                            13.000000
min
                         41.000000
25%
       7.130368e+08
                                             1.000000
                                                            31.000000
50%
       7.179264e+08
                         46.000000
                                             2.000000
                                                            36.000000
75%
       7.731435e+08
                         52.000000
                                            3.000000
                                                            40.000000
                         73.000000
max
       8.283431e+08
                                            5.000000
                                                            56.000000
       Total Relationship Count
                                   Months Inactive 12 mon
                    10127.000000
                                              10127.000000
count
                        3.812580
                                                  2.341167
mean
std
                        1.554408
                                                  1.010622
                        1.000000
                                                  0.00000
min
25%
                        3.000000
                                                  2.000000
50%
                        4.000000
                                                  2.000000
                        5.000000
75%
                                                  3.000000
                        6.000000
                                                  6.000000
max
                                Credit Limit
                                               Total Revolving Bal
       Contacts Count 12 mon
                 10127.000000
                                10127.000000
                                                      10127.000000
count
                     2.455317
                                 8631.953698
                                                       1162.814061
mean
std
                     1.106225
                                 9088.776650
                                                        814.987335
                                                          0.000000
min
                     0.000000
                                 1438.300000
25%
                     2.000000
                                 2555.000000
                                                        359.000000
                                 4549,000000
                                                       1276,000000
50%
                     2.000000
75%
                     3.000000
                                11067.500000
                                                       1784.000000
                     6.000000
                               34516.000000
                                                       2517.000000
max
       Avg_Open_To Buy
                         Total Amt Chng Q4 Q1
                                                 Total Trans Amt
Total_Trans_Ct \
          10127.000000
                                  10127.000000
count
                                                    10127.000000
10127.000000
           7469.139637
                                      0.759941
                                                     4404.086304
mean
64.858695
std
                                      0.219207
                                                     3397,129254
           9090.685324
23.472570
min
               3.000000
                                      0.000000
                                                      510.000000
10.000000
25%
           1324.500000
                                      0.631000
                                                     2155.500000
45.000000
50%
           3474.000000
                                      0.736000
                                                     3899,000000
67.000000
75%
           9859.000000
                                      0.859000
                                                     4741.000000
```

```
81.000000
          34516.000000
                                      3.397000
                                                    18484.000000
max
139.000000
       Total Ct Chng 04 01
                              Avg Utilization Ratio \
count
               10127.000000
                                       10127.000000
mean
                   0.712222
                                            0.274894
                   0.238086
                                            0.275691
std
                   0.000000
                                            0.000000
min
25%
                   0.582000
                                            0.023000
50%
                   0.702000
                                            0.176000
75%
                   0.818000
                                            0.503000
                   3.714000
max
                                            0.999000
Naive Bayes Classifier Attrition Flag Card Category Contacts Count 12
mon_Dependent_count_Education_Level_Months_Inactive_12_mon_1 \
count
                                               10127.0\overline{0}00\overline{0}0
mean
                                                   0.159997
std
                                                   0.365301
                                                   0.000008
min
25%
                                                   0.000099
50%
                                                   0.000181
75%
                                                   0.000337
                                                   0.999580
max
Naive Bayes Classifier Attrition Flag Card Category Contacts Count 12
mon Dependent count Education Level Months Inactive 12 mon 2
count
                                               10127.000000
                                                   0.840003
mean
                                                   0.365301
std
min
                                                   0.000420
25%
                                                   0.999660
50%
                                                   0.999820
75%
                                                   0.999900
```

max 0.999990

Summary based on the provided numerical columns:

Client Statistics:

Mean client number: 7.391776e+08Mean customer age: 46.33 years

• Mean dependent count: 2.35

• Mean months on book: 35.93

Mean total relationship count: 3.81

Mean months inactive in the last 12 months: 2.34

• Mean contacts count in the last 12 months: 2.46

Credit Statistics:

Mean credit limit: 8631.95

Mean total revolving balance: 1162.81

Mean average open to buy: 7469.14

Mean total amount change Q4-Q1: 0.76

Mean total transaction amount: 4404.09

Mean total transaction count: 64.86

Mean total count change Q4-Q1: 0.71

Mean average utilization ratio: 0.27

Naive Bayes Classifier Statistics:

- Mean Naive Bayes Classifier (Attrition Flag Card Category Contacts Count 12 mon -Dependent count - Education Level - Months Inactive 12 mon - 1): 0.159997
- Mean Naive Bayes Classifier (Attrition Flag Card Category Contacts Count 12 mon -Dependent count - Education Level - Months Inactive 12 mon - 2): 0.840003

These statistics provide an overview of the numerical features in the dataset, including client demographics, credit-related information, and classifier statistics.

```
# Check for missing values
print("Missing values:")
print(churners_data.isnull().sum())

Missing values:
CLIENTNUM
0
Attrition_Flag
0
Customer_Age
0
Gender
0
```

```
Dependent count
Education Level
Marital Status
Income_Category
Card Category
Months_on_book
Total Relationship Count
Months_Inactive_12_mon
Contacts_Count_12_mon
Credit Limit
Total Revolving Bal
Avg Open To Buy
Total Amt Chng Q4 Q1
Total_Trans_Amt
Total Trans Ct
Total Ct Chng Q4 Q1
Avg Utilization Ratio
Naive Bayes Classifier Attrition Flag Card Category Contacts Count 12
mon Dependent count Education Level Months Inactive 12 mon 1
Naive Bayes Classifier Attrition Flag Card Category Contacts Count 12
mon Dependent count Education Level Months Inactive 12 mon 2
dtype: int64
```

#Correlations Matrix

```
# Correlation matrix
correlation_matrix = churners_data.corr()
print("Correlation matrix:")
print(correlation_matrix)

Correlation matrix:
CLIENTNUM
```

Customer_Age \ CLIENTNUM	1 000000	
0.007613	1.000000	
Customer Age	0.007613	
1.000000	0.007013	
Dependent count	0.006772 -	
0.122254	0.000772	
Months on book	0.134588	
0.788912	01131300	
Total Relationship Count	0.006907 -	
0.010931	0100007	
Months_Inactive_12_mon	0.005729	
0.054361	01005725	
Contacts Count 12 mon	0.005694 -	
0.018452		
Credit Limit	0.005708	
0.002476		
Total Revolving Bal	0.000825	
0.014780		
Avg Open To Buy	0.005633	
0.001151		
Total Amt Chng Q4 Q1	0.017369 -	
0.062042		
Total Trans Amt	-0.019692 -	
0.046446		
Total Trans Ct	-0.002961 -	
$0.067\overline{0}97$		
Total_Ct_Chng_Q4_Q1	0.007696 -	
0.012143		
Avg_Utilization_Ratio	0.000266	
0.007114		
Naive_Bayes_Classifier_Attrition_Flag_Card_Cate	-0.046411	
0.018189		
Naive Bayes Classifier Attrition Flag Card Cate	0.046410 -	
0.018189		
	Dependent_count	\
CLIENTNUM	0.006772	
Customer_Age	-0.122254	
Dependent_count	1.000000	
Months_on_book	-0.103062	
Total_Relationship_Count	-0.039076	
Months_Inactive_12_mon	-0.010768	
Contacts_Count_12_mon	-0.040505	
Credit_Limit	0.068065	
Total_Revolving_Bal	-0.002688	
Avg_Open_To_Buy	0.068291	
Total_Amt_Chng_Q4_Q1	-0.035439	
Total_Trans_Amt	0.025046	

Total_Trans_Ct Total_Ct_Chng_Q4_Q1 Avg_Utilization_Ratio Naive_Bayes_Classifier_Attrition_Flag_Card_Cate Naive_Bayes_Classifier_Attrition_Flag_Card_Cate	0.049912 0.011087 -0.037135 0.019189 -0.019189	
CLIENTNUM Customer_Age Dependent_count Months_on_book Total_Relationship_Count Months_Inactive_12_mon Contacts_Count_12_mon Credit_Limit Total_Revolving_Bal Avg_Open_To_Buy Total_Amt_Chng_Q4_Q1 Total_Trans_Amt Total_Trans_Ct Total_Trans_Ct Total_Ct_Chng_Q4_Q1 Avg_Utilization_Ratio Naive_Bayes_Classifier_Attrition_Flag_Card_Cate Naive_Bayes_Classifier_Attrition_Flag_Card_Cate	Months_on_book 0.134588 0.788912 -0.103062 1.000000 -0.009203 0.074164 -0.010774 0.007507 0.008623 0.006732 -0.048959 -0.038591 -0.049819 -0.014072 -0.007541 0.013693 -0.013694	
Total_Relationship_Count \ CLIENTNUM 0.006907		
Customer_Age 0.010931	-	
Dependent_count	-	
0.039076 Months on book	<u>-</u>	
0.009203 Total_Relationship_Count		
1.000000 Months Inactive 12 mon	-	
0.003675		
Contacts_Count_12_mon 0.055203		
Credit_Limit	-	
0.071386 Total_Revolving_Bal 0.013726		
Avg_Open_To_Buy	-	
0.072601 Total_Amt_Chng_Q4_Q1 0.050119		
Total_Trans_Amt	-	

```
0.347229
Total Trans Ct
0.241891
Total Ct Chng Q4 Q1
0.040831
Avg Utilization Ratio
0.067663
Naive Bayes Classifier Attrition Flag Card Cate...
0.149981
Naive Bayes Classifier Attrition Flag Card Cate...
0.149981
Months Inactive 12 mon \
CLIENTNUM
0.005729
Customer_Age
0.054361
Dependent count
0.010768
Months on book
0.074164
Total Relationship Count
0.003675
Months Inactive 12 mon
1.000000
Contacts Count 12 mon
0.029493
Credit Limit
0.020394
Total Revolving Bal
0.042210
Avg_Open_To_Buy
0.016605
Total Amt Chng Q4 Q1
0.032247
Total Trans Amt
0.036982
Total_Trans_Ct
0.042787
Total_Ct_Chng_Q4_Q1
0.038989
Avg Utilization Ratio
0.007503
Naive Bayes Classifier Attrition Flag Card Cate...
0.153452
Naive_Bayes_Classifier_Attrition_Flag_Card_Cate...
0.153452
```

```
Contacts Count 12 mon \
CLIENTNUM
0.005694
Customer Age
0.018452
Dependent count
0.040505
Months on book
0.010774
Total Relationship Count
0.055203
Months_Inactive_12_mon
0.029493
Contacts Count 12 mon
1.000000
Credit Limit
0.020817
Total Revolving Bal
0.053913
Avg_Open_To_Buy
0.025646
Total Amt Chng Q4 Q1
0.024445
Total Trans Amt
0.112774
Total Trans_Ct
0.152213
Total Ct Chng Q4 Q1
0.094997
Avg Utilization Ratio
0.055471
Naive Bayes Classifier Attrition Flag Card Cate...
0.205772
Naive Bayes Classifier Attrition Flag Card Cate...
0.205772
                                                      Credit_Limit \
CLIENTNUM
                                                          0.005708
Customer Age
                                                          0.002476
Dependent count
                                                          0.068065
Months on book
                                                          0.007507
Total Relationship Count
                                                         -0.071386
Months Inactive 12 mon
                                                         -0.020394
Contacts Count 12 mon
                                                          0.020817
Credit Limit
                                                          1.000000
Total Revolving Bal
                                                          0.042493
Avg Open To Buy
                                                          0.995981
Total Amt Chng Q4 Q1
                                                          0.012813
Total_Trans_Amt
                                                          0.171730
```

```
Total Trans Ct
                                                         0.075927
Total Ct Chng Q4 Q1
                                                        -0.002020
Avg Utilization Ratio
                                                        -0.482965
Naive Bayes Classifier Attrition Flag Card Cate...
                                                        -0.023853
Naive Bayes Classifier Attrition Flag Card Cate...
                                                         0.023853
Total Revolving Bal \
CLIENTNUM
0.000825
Customer Age
0.014780
Dependent count
0.002688
Months on book
0.008623
Total Relationship_Count
0.013726
Months Inactive 12 mon
0.042210
Contacts Count 12 mon
0.053913
Credit Limit
0.042493
Total Revolving Bal
1.000000
Avg_Open_To_Buy
0.047167
Total Amt Chng Q4 Q1
0.058174
Total Trans Amt
0.064370
Total_Trans_Ct
0.056060
Total Ct Chng Q4 Q1
0.089861
Avg Utilization Ratio
0.624022
Naive Bayes Classifier Attrition Flag Card Cate...
0.263033
Naive_Bayes_Classifier_Attrition_Flag_Card_Cate...
0.263032
                                                     Avg Open To Buy \
CLIENTNUM
                                                            0.005633
Customer Age
                                                            0.001151
Dependent count
                                                            0.068291
Months on book
                                                            0.006732
Total Relationship Count
                                                           -0.072601
Months Inactive 12 mon
                                                           -0.016605
```

```
Contacts Count 12 mon
                                                             0.025646
Credit Limit
                                                             0.995981
Total Revolving Bal
                                                            -0.047167
Avg Open To Buy
                                                             1.000000
Total Amt Chnq Q4 Q1
                                                             0.007595
Total_Trans_Amt
                                                             0.165923
Total Trans Ct
                                                             0.070885
Total Ct Chng Q4 Q1
                                                            -0.010076
Avg Utilization Ratio
                                                            -0.538808
Naive Bayes Classifier Attrition Flag Card Cate...
                                                            -0.000267
Naive Bayes Classifier Attrition Flag Card Cate...
                                                            0.000267
Total_Amt_Chng_Q4_Q1 \
CLIENTNUM
0.017369
Customer Age
0.062042
Dependent count
0.035439
Months on book
0.048959
Total Relationship Count
0.050119
Months Inactive 12 mon
0.032247
Contacts Count 12 mon
0.024445
Credit Limit
0.012813
Total Revolving Bal
0.058174
Avg_Open_To_Buy
0.007595
Total Amt Chng Q4 Q1
1.000000
Total Trans Amt
0.039678
Total Trans Ct
0.005469
Total_Ct_Chng_Q4_Q1
0.384189
Avg Utilization Ratio
0.035235
Naive Bayes Classifier Attrition Flag Card Cate...
0.131207
Naive_Bayes_Classifier_Attrition_Flag_Card_Cate...
0.131207
                                                     Total Trans Amt \
```

```
CLIENTNUM
                                                            -0.019692
Customer Age
                                                            -0.046446
Dependent count
                                                             0.025046
Months on book
                                                            -0.038591
Total Relationship Count
                                                            -0.347229
Months Inactive 12 mon
                                                            -0.036982
Contacts Count 12 mon
                                                            -0.112774
Credit Limit
                                                             0.171730
Total Revolving Bal
                                                             0.064370
Avg Open To Buy
                                                             0.165923
Total Amt Chng Q4 Q1
                                                             0.039678
Total Trans Amt
                                                             1.000000
Total_Trans_Ct
                                                             0.807192
Total Ct Chng 04 01
                                                             0.085581
Avg Utilization Ratio
                                                            -0.083034
Naive Bayes Classifier Attrition Flag Card Cate...
                                                            -0.168642
Naive Bayes Classifier Attrition Flag Card Cate...
                                                            0.168642
                                                     Total Trans Ct \
CLIENTNUM
                                                           -0.002961
Customer Age
                                                           -0.067097
Dependent count
                                                           0.049912
Months on book
                                                           -0.049819
Total Relationship Count
                                                           -0.241891
Months Inactive 12 mon
                                                           -0.042787
Contacts Count 12 mon
                                                           -0.152213
Credit Limit
                                                           0.075927
Total Revolving Bal
                                                           0.056060
Avg Open To Buy
                                                           0.070885
Total Amt Chng Q4 Q1
                                                            0.005469
Total Trans Amt
                                                           0.807192
Total Trans Ct
                                                            1.000000
Total Ct Chng Q4 Q1
                                                            0.112324
Avg Utilization Ratio
                                                           0.002838
Naive Bayes Classifier Attrition Flag Card Cate...
                                                          -0.371403
Naive Bayes Classifier Attrition Flag Card Cate...
                                                           0.371403
Total Ct Chng Q4 Q1 \
CLIENTNUM
0.007696
Customer Age
0.012143
Dependent count
0.011087
Months on book
0.014072
Total Relationship Count
0.040831
Months Inactive 12 mon
```

```
0.038989
Contacts_Count_12_mon
0.094997
Credit Limit
0.002020
Total Revolving Bal
0.089861
Avg Open To Buy
0.010076
Total Amt Chng Q4 Q1
0.384189
Total_Trans_Amt
0.085581
Total Trans Ct
0.112324
Total Ct Chng Q4 Q1
1.000000
Avg Utilization Ratio
0.074143
Naive Bayes Classifier Attrition Flag Card Cate...
0.290115
Naive Bayes Classifier Attrition Flag Card Cate...
0.290115
Avg Utilization Ratio \
CLIENTNUM
0.000266
Customer Age
0.007114
Dependent count
0.037135
Months on book
0.007541
Total Relationship Count
0.067663
Months Inactive 12 mon
0.007503
Contacts Count 12 mon
0.055471
Credit Limit
0.482965
Total Revolving_Bal
0.624022
Avg_Open_To_Buy
0.538808
Total_Amt_Chng_Q4_Q1
0.035235
Total Trans Amt
```

```
0.083034
Total Trans Ct
0.002838
Total Ct Chng Q4 Q1
0.074143
Avg Utilization Ratio
1.000000
Naive Bayes Classifier_Attrition_Flag_Card_Cate...
0.178405
Naive Bayes Classifier Attrition Flag Card Cate...
0.178405
Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_
mon Dependent count Education Level Months Inactive 12 mon 1 \
CLIENTNUM
-0.046411
Customer Age
0.018189
Dependent count
0.019189
Months on book
0.013693
Total_Relationship_Count
-0.149981
Months Inactive 12 mon
0.1534\overline{5}2
Contacts_Count_12_mon
0.205772
Credit Limit
-0.023853
Total Revolving Bal
-0.263033
Avg_Open_To_Buy
-0.\overline{000267}
Total_Amt_Chng_Q4_Q1
-0.131207
Total Trans Amt
-0.168642
```

```
Total_Trans_Ct
-0.371403
Total_Ct_Chng_Q4_Q1
-0.29\overline{0}11\overline{5}
Avg_Utilization_Ratio
-0.\overline{1}78405
Naive_Bayes_Classifier_Attrition_Flag_Card_Cate...
1.000000
Naive_Bayes_Classifier_Attrition_Flag_Card_Cate...
-1.000000
Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_
mon Dependent count Education Level Months Inactive 12 mon 2
CLIENTNUM
0.046410
Customer Age
-0.01818\overline{9}
Dependent_count
-0.019189
Months_on_book
-0.013694
Total Relationship Count
0.149\overline{9}81
Months_Inactive_12_mon
-0.153452
Contacts Count 12 mon
-0.20577\overline{2}
Credit Limit
0.0238\overline{5}3
Total_Revolving_Bal
0.263\overline{0}32
Avg_Open_To_Buy
0.000267
Total_Amt_Chng_Q4_Q1
0.131\overline{2}07
```

```
Total Trans Amt
0.168642
Total Trans Ct
0.371403
Total_Ct_Chng_Q4_Q1
0.290115
Avg Utilization Ratio
0.178405
Naive Bayes Classifier Attrition Flag Card Cate...
-1.000000
Naive Bayes Classifier Attrition Flag Card Cate...
1.000000
<ipython-input-10-ca4bdf0fda51>:2: FutureWarning: The default value of
numeric only in DataFrame.corr is deprecated. In a future version, it
will default to False. Select only valid columns or specify the value
of numeric only to silence this warning.
  correlation matrix = churners data.corr()
```

The provided correlation matrix represents the pairwise correlations between numerical features in the dataset. Here's a summary:

Positive Correlations:

- Customer Age is positively correlated with Months on Book (0.788912).
- Total Transaction Amount is positively correlated with Total
- Transaction Count (0.807192).
- Total Revolving Balance is positively correlated with Avg Utilization Ratio (0.624022).

Negative Correlations:

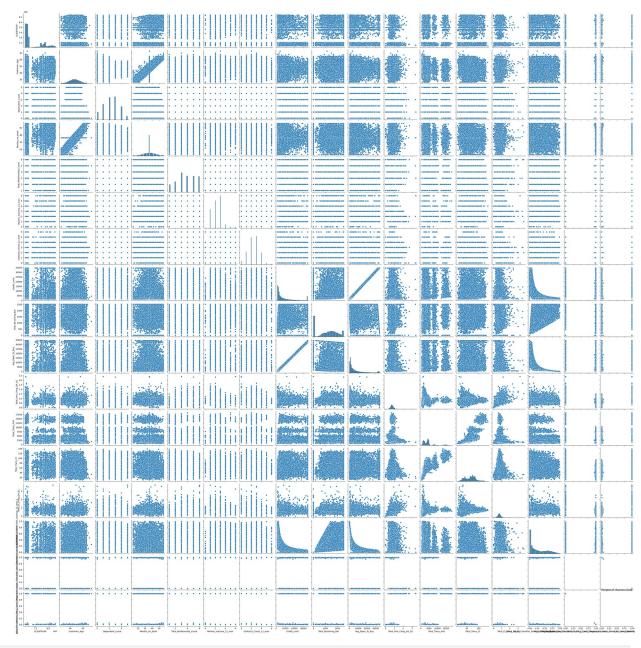
- Avg Utilization Ratio is negatively correlated with Credit Limit (-0.482965).
- Total Relationship Count is negatively correlated with Total
- Transaction Amount (-0.347229).
- Total Transaction Count is negatively correlated with Naive Bayes Classifier (Attrition Flag - Card Category - Contacts Count 12 mon - Dependent count - Education Level -Months Inactive 12 mon - 1) (-0.371403).

Weak Correlations:

There are weak correlations (close to 0) between many pairs of features, indicating little linear relationship between them. This correlation matrix provides valuable insights into how features are related to each other within the dataset. However, it's important to note that correlation

doesn't imply causation, and further analysis may be needed to understand the relationships between variables more deeply.

```
# Visualization: Pairplot
sns.pairplot(churners_data)
plt.title('Pairplot of churners Data')
plt.show()
```



```
# Visualization: Correlation Heatmap
plt.figure(figsize=(10, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm',
fmt=".2f")
```

plt.title('Correlation Heatmap of churners Data') plt.show()

