Cross Selling Credit Union

LISUM12: Data Glacier Final Project Presentation

Group Name: Team Coltenback

Name: Richard Coltenback

Email: rdcoltenback@gmail.com

Country: United States of America

College: Drew University

Specialization: Data Analyst

GitHub Repository Link: https://github.com/RColtenback/Cross-Selling-Data-Analysis.git

Problem Description:

XYZ credit union in Latin America is performing very well in selling the banking products (e.g.: credit card, deposit account, retirement account, safe deposit box etc.). However, their existing customers are not buying more than one product which means bank is not performing well in cross-selling (Bank is not able to sell their other offerings to existing customer). XYZ Credit Union decided to approach Team Coltenback to solve their problem.

Business Understanding:

We're trying to understand the advantages and disadvantages of cross-selling banking products as opposed to up-selling. In particular, we want to know the financial benefits of cross-selling to understand which products do well together and which should be sold alone. We have personal information about each customer for the bank. Thus, we can analyze their backgrounds and see if there is any correlation with certain personal traits (e.g.: race, age, sex, etc.).

Project Lifecycle and Deadline:

1. Business understanding

2. Data Understanding

3. Data Cleansing and Transformation

4. Exploratory data analysis

5. EDA Recommendation (PowerPoint)

6. Dashboard which should capture type of customer their count, segment wise (VIP, student, etc.) customer average age, and other KPIs which gives better business insight in taking decision.

7. Prepare a final presentation

The final deadline is scheduled for October 30th, 2022.

Data Understanding:

Train.csv

Total number of observations (row)	13647309
Total number of files	1
Total number of features (columns)	48
Base format of the file	.csv
Size of the data	2.29 GB

The dataset we will be using will be information given to us from the XYZ Credit Union. It has many different variables and lots of data that can help us solve our problem. The data has to deal with banking and different products that go along with it. Some examples of the variables are loans, pensions, mortgages, etc. Fortunately, due to the data collected being based on each customer, we also have access to personal information like their ages, sex, country of residence, etc. This dataset has a mix of different types of variables. Some variables like fecha_dato and fecha_alta are dates. Some variables like indext and indfall are Boolean or dichotomous as they are true/false or yes/no. Some variables like ind_empleado are string or nominal as they have letters for each customer. Finally, some variables like ncodpers and age are integer or scale as they have numbers for each customer.

There are some issues with the dataset. First, not all of the data is filled in. Some of the values of the variables are either missing or set as NA. This isn't too big of an issue for me as I'm planning to analyze the data in ways that don't limit me by losing variables due to their being missing data or NA. I do believe visuals might need some work around so they still depict what is wanting to be shown. Since the dataset is so big, I won't be able to use a friendly program like Tableau, which is easy to work in, so I'll have to be clever. A second issue that could arise is the large number of data being available to us causing slow loading times and trouble processing commands. I could use modin and dask instead of regular pandas, but there may be more that needs to be done to optimize the exploration of the dataset. Issues like outliers and skew won't be an issue for me as I will use means and medians in places that need it. I believe keeping the outliers is good practice so people who have the data explore all of it and know everything about it, painting the clearest picture possible. However, should I need to explore the dataset without outliers, I could also set minimum and maximum values on each variable before running analysis.

Data Cleansing and Transformation:

My first goal for cleaning up this data was changing the names of the column variables. The original names were in Spanish and didn't make sense to me. Using the column description, I changed the names of each variable to make them more understandable for an English speaking audience (I could change them back to Spanish after since this is a Latin American credit union). Next, I checked the types of variables and how many NA values there are in the dataset to help show me which variables need a work around. Then, I found the mean, median, mode, and other attributes of a couple variables to have experience with them. Using the median, I was able to fill the NA values for each variable, not changing anything about the data and the trends. Also, the amount of NA values for each variable I changed wouldn't have enough weight to statistically significantly impact the data. Finally, I changed some of the numeric variables from objects to integers or float to make exploring the data easier.

```python
df.rename(columns={'fecha_dato':'Partitioned Date',
                   'ncodpers':'Customer Code',
                   'ind_empleado':'Employee Index',
                   'pais_residencia':'Country Residence',
                   'sexo':'Sex',
                   'age':'Age',
                   'fecha_alta':'Date Joined',
                   'ind_nuevo':'New Customer Index',
                   'antiguedad':'Seniority',
                   'indrel':'Primary Customer',
                   'ult_fec_cli_1t':'Last Date as Primary Customer',
                   'indrel_1mes':'Customer Type at Beginning of Month',
                   'tiprel_1mes':'Customer Relation Type at Beginning of Month',
                   'indresi':'Resident Country is Bank Country',
                   'indext':'Birth Country Different Than Bank Country',
                   'conyuemp':'Spouse Index',
                   'canal_entrada':'Channel Used to Join',
                   'indfall':'Deceased',
                   'tipodom':'Address Type',
                   'cod_prov':'Province Code',
                   'nomprov':'Province Name',
                   'ind_actividad_cliente':'Activity Index',
                   'renta':'Gross Household Income',
                   'segmento':'Segmentation',
                   'ind_ahor_fin_ult1':'Savings Account',
                   'ind_aval_fin_ult1':'Guarantees',
                   'ind_cco_fin_ult1':'Current Accounts',
                   'ind_cder_fin_ult1':'Derivative Account',
                   'ind_cno_fin_ult1':'Payroll Account',
                   'ind_ctju_fin_ult1':'Junior Account',
                   'ind_ctma_fin_ult1':'More Partiuclar Account',
                   'ind_ctop_fin_ult1':'Particular Account',
                   'ind_ctpp_fin_ult1':'Particular Plus Account',
                   'ind_deco_fin_ult1':'Short-Term Deposits',
                   'ind_deme_fin_ult1':'Medium-Term Deposits',
                   'ind_dela_fin_ult1':'Long-Term Deposits',
                   'ind_ecue_fin_ult1':'E-Account',
                   'ind_fond_fin_ult1':'Funds',
                   'ind_hip_fin_ult1':'Mortgage',
                   'ind_plan_fin_ult1':'Pensions',
                   'ind_pres_fin_ult1':'Loans',
                   'ind_reca_fin_ult1':'Taxes',
                   'ind_tjcr_fin_ult1':'Credit Card',
                   'ind_valo_fin_ult1':'Securities',
                   'ind_viv_fin_ult1':'Home Account',
                   'ind_nomina_ult1':'Payroll',
                   'ind_nom_pens_ult1':'Pension',
                   'ind_recibo_ult1':'Direct Debit'}, inplace = True)
```

```
df['Sex'] = df['Sex'].replace(['H'],['1'])
df['Sex'] = df['Sex'].replace(['V'],['0'])
df['Sex'] = df['Sex'].astype(float)
df['Sex'] = df['Sex'].fillna(0)

df['Age'] = df['Age'].replace([' NA'],[' 39'])
df['Age'] = df['Age'].astype(int)

df['Seniority'] = df['Seniority'].replace(['     NA'],['50'])
df['Seniority'] = df['Seniority'].replace(['-999999'],['50'])
df['Seniority'] = df['Seniority'].astype(int)

df['Resident Country is Bank Country'] = df['Resident Country is Bank Country'].replace(['N'],['0'])
df['Resident Country is Bank Country'] = df['Resident Country is Bank Country'].replace(['S'],['1'])
df['Resident Country is Bank Country'] = df['Resident Country is Bank Country'].astype(float)
df['Resident Country is Bank Country'] = df['Resident Country is Bank Country'].fillna(1)

df['Birth Country Different Than Bank Country'] = df['Birth Country Different Than Bank Country'].replace(['N'],['0'])
df['Birth Country Different Than Bank Country'] = df['Birth Country Different Than Bank Country'].replace(['S'],['1'])
df['Birth Country Different Than Bank Country'] = df['Birth Country Different Than Bank Country'].astype(float)
df['Birth Country Different Than Bank Country'] = df['Birth Country Different Than Bank Country'].fillna(0)

df['Spouse Index'] = df['Spouse Index'].replace(['N'],['0'])
df['Spouse Index'] = df['Spouse Index'].replace(['S'],['1'])
df['Spouse Index'] = df['Spouse Index'].astype(float)

df['Deceased'] = df['Deceased'].replace(['N'],['0'])
df['Deceased'] = df['Deceased'].replace(['S'],['1'])
df['Deceased'] = df['Deceased'].astype(float)
df['Deceased'] = df['Deceased'].fillna(0)

df['Activity Index'] = df['Activity Index'].fillna(0)

df['Payroll'] = df['Payroll'].fillna(0)

df['Pension'] = df['Pension'].fillna(0)

df['New Customer Index'] = df['New Customer Index'].fillna(0)

df['Primary Customer'] = df['Primary Customer'].fillna(1)

df['Gross Household Income'] = df['Gross Household Income'].fillna(101850)
```

```
Dataframe Variable Types

Partitioned Date                                  object
Customer Code                                      int64
Employee Index                                    object
Country Residence                                 object
Sex                                               object
Age                                               object
Date Joined                                       object
New Customer Index                               float64
Seniority                                         object
Primary Customer                                 float64
Last Date as Primary Customer                     object
Customer Type at Beginning of Month               object
Customer Relation Type at Beginning of Month      object
Resident Country is Bank Country                  object
Birth Country Different Than Bank Country         object
Spouse Index                                      object
Channel Used to Join                              object
Deceased                                          object
Address Type                                     float64
Province Code                                    float64
Province Name                                     object
Activity Index                                   float64
Gross Household Income                           float64
Segmentation                                      object
Savings Account                                    int64
Guarantees                                         int64
Current Accounts                                   int64
Derivative Account                                 int64
Payroll Account                                    int64
Junior Account                                     int64
More Partiuclar Account                            int64
Particular Account                                 int64
Particular Plus Account                            int64
Short-Term Deposits                                int64
Medium-Term Deposits                               int64
Long-Term Deposits                                 int64
E-Account                                          int64
Funds                                              int64
Mortgage                                           int64
Pensions                                           int64
Loans                                              int64
Taxes                                              int64
Credit Card                                        int64
Securities                                         int64
Home Account                                       int64
Payroll                                          float64
Pension                                          float64
Direct Debit                                       int64
dtype: object
```

TO

```
Dataframe Variable Types

Partitioned Date                                  object
Customer Code                                      int64
Employee Index                                    object
Country Residence                                 object
Sex                                              float64
Age                                                int32
Date Joined                                       object
New Customer Index                               float64
Seniority                                          int32
Primary Customer                                 float64
Last Date as Primary Customer                     object
Customer Type at Beginning of Month               object
Customer Relation Type at Beginning of Month      object
Resident Country is Bank Country                 float64
Birth Country Different Than Bank Country        float64
Spouse Index                                     float64
Channel Used to Join                              object
Deceased                                         float64
Address Type                                     float64
Province Code                                    float64
Province Name                                     object
Activity Index                                   float64
Gross Household Income                           float64
Segmentation                                      object
Savings Account                                    int64
Guarantees                                         int64
Current Accounts                                   int64
Derivative Account                                 int64
Payroll Account                                    int64
Junior Account                                     int64
More Partiuclar Account                            int64
Particular Account                                 int64
Particular Plus Account                            int64
Short-Term Deposits                                int64
Medium-Term Deposits                               int64
Long-Term Deposits                                 int64
E-Account                                          int64
Funds                                              int64
Mortgage                                           int64
Pensions                                           int64
Loans                                              int64
Taxes                                              int64
Credit Card                                        int64
Securities                                         int64
Home Account                                       int64
Payroll                                          float64
Pension                                          float64
Direct Debit                                       int64
dtype: object
```

```
Dataframe NA Count

Partitioned Date                                      0
Customer Code                                         0
Employee Index                                    27734
Country Residence                                 27734
Sex                                               27804
Age                                                   0
Date Joined                                       27734
New Customer Index                                27734
Seniority                                             0
Primary Customer                                  27734
Last Date as Primary Customer                  13622516
Customer Type at Beginning of Month              149781
Customer Relation Type at Beginning of Month     149781
Resident Country is Bank Country                  27734
Birth Country Different Than Bank Country         27734
Spouse Index                                   13645501
Channel Used to Join                             186126
Deceased                                          27734
Address Type                                      27735
Province Code                                     93591
Province Name                                     93591
Activity Index                                    27734
Gross Household Income                          2794375
Segmentation                                     189368
Savings Account                                       0
Guarantees                                            0
Current Accounts                                      0
Derivative Account                                    0
Payroll Account                                       0
Junior Account                                        0
More Partiuclar Account                               0
Particular Account                                    0
Particular Plus Account                               0
Short-Term Deposits                                   0
Medium-Term Deposits                                  0
Long-Term Deposits                                    0
E-Account                                             0
Funds                                                 0
Mortgage                                              0
Pensions                                              0
Loans                                                 0
Taxes                                                 0
Credit Card                                           0
Securities                                            0
Home Account                                          0
Payroll                                           16063
Pension                                           16063
Direct Debit                                          0
dtype: int64
```

TO

```
Dataframe NA Count

Partitioned Date                                           0
Customer Code                                              0
Employee Index                                         27734
Country Residence                                      27734
Sex                                                        0
Age                                                        0
Date Joined                                            27734
New Customer Index                                         0
Seniority                                                  0
Primary Customer                                           0
Last Date as Primary Customer                       13622516
Customer Type at Beginning of Month                   149781
Customer Relation Type at Beginning of Month          149781
Resident Country is Bank Country                           0
Birth Country Different Than Bank Country                  0
Spouse Index                                        13645501
Channel Used to Join                                  186126
Deceased                                                   0
Address Type                                           27735
Province Code                                          93591
Province Name                                          93591
Activity Index                                             0
Gross Household Income                                     0
Segmentation                                          189368
Savings Account                                            0
Guarantees                                                 0
Current Accounts                                           0
Derivative Account                                         0
Payroll Account                                            0
Junior Account                                             0
More Partiuclar Account                                    0
Particular Account                                         0
Particular Plus Account                                    0
Short-Term Deposits                                        0
Medium-Term Deposits                                       0
Long-Term Deposits                                         0
E-Account                                                  0
Funds                                                      0
Mortgage                                                   0
Pensions                                                   0
Loans                                                      0
Taxes                                                      0
Credit Card                                                0
Securities                                                 0
Home Account                                               0
Payroll                                                    0
Pension                                                    0
Direct Debit                                               0
dtype: int64
```

Exploratory Data Analysis:

I began the exploratory data analysis (EDA) by taking the little experience I gained from cleaning and transforming the data and applying it by finding all the count values of each variable and the sum of each variable to get a vague sense of how many accounts there were in all the options that the bank offered.

Dataframe Count

```
Partitioned Date                                13647309
Customer Code                                   13647309
Employee Index                                  13619575
Country Residence                               13619575
Sex                                             13647309
Age                                             13647309
Date Joined                                     13619575
New Customer Index                              13647309
Seniority                                       13647309
Primary Customer                                13647309
Last Date as Primary Customer                      24793
Customer Type at Beginning of Month             13497528
Customer Relation Type at Beginning of Month    13497528
Resident Country is Bank Country                13647309
Birth Country Different Than Bank Country       13647309
Spouse Index                                        1808
Channel Used to Join                            13461183
Deceased                                        13647309
Address Type                                    13619574
Province Code                                   13553718
Province Name                                   13553718
Activity Index                                  13647309
Gross Household Income                          13647309
Segmentation                                    13457941
Savings Account                                 13647309
Guarantees                                      13647309
Current Accounts                                13647309
Derivative Account                              13647309
Payroll Account                                 13647309
Junior Account                                  13647309
More Partiuclar Account                         13647309
Particular Account                              13647309
Particular Plus Account                         13647309
Short-Term Deposits                             13647309
Medium-Term Deposits                            13647309
Long-Term Deposits                              13647309
E-Account                                       13647309
Funds                                           13647309
Mortgage                                        13647309
Pensions                                        13647309
Loans                                           13647309
Taxes                                           13647309
Credit Card                                     13647309
Securities                                      13647309
Home Account                                    13647309
Payroll                                         13647309
Pension                                         13647309
Direct Debit                                    13647309
dtype: int64
```

```
Dataframe Sum

Sex                                         6.195253e+06
Age                                         5.483361e+08
New Customer Index                          8.112070e+05
Seniority                                   1.082538e+09
Primary Customer                            1.607702e+07
Resident Country is Bank Country            1.358144e+07
Birth Country Different Than Bank Country   6.447360e+05
Deceased                                    3.476200e+04
Activity Index                              6.235185e+06
Gross Household Income                      1.741660e+12
Savings Account                             1.396000e+03
Guarantees                                  3.160000e+02
Current Accounts                            8.945588e+06
Derivative Account                          5.376000e+03
Payroll Account                             1.103620e+06
Junior Account                              1.292970e+05
More Partiuclar Account                     1.327420e+05
Particular Account                          1.760616e+06
Particular Plus Account                     5.910080e+05
Short-Term Deposits                         2.427500e+04
Medium-Term Deposits                        2.266800e+04
Long-Term Deposits                          5.863810e+05
E-Account                                   1.129227e+06
Funds                                       2.522840e+05
Mortgage                                    8.033600e+04
Pensions                                    1.251590e+05
Loans                                       3.585700e+04
Taxes                                       7.169800e+05
Credit Card                                 6.057860e+05
Securities                                  3.494750e+05
Home Account                                5.251100e+04
Payroll                                     7.459610e+05
Pension                                     8.100850e+05
Direct Debit                                1.745712e+06
dtype: float64
```

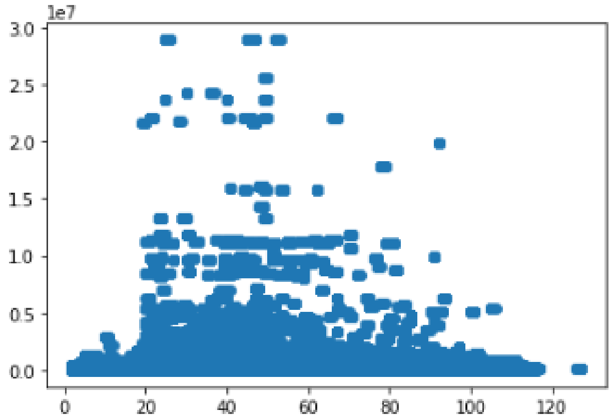
I then found the means and medians of each variable to see which accounts of the bank were doing better than the others.

```
Dataframe Mean

Sex                                                   0.453954
Age                                                  40.179064
New Customer Index                                    0.059441
Seniority                                            79.322469
Primary Customer                                      1.178036
Resident Country is Bank Country                      0.995174
Birth Country Different Than Bank Country             0.047243
Deceased                                              0.002547
Activity Index                                        0.456880
Gross Household Income                           127619.323986
Savings Account                                       0.000102
Guarantees                                            0.000023
Current Accounts                                      0.655484
Derivative Account                                    0.000394
Payroll Account                                       0.080867
Junior Account                                        0.009474
More Partiuclar Account                               0.009727
Particular Account                                    0.129008
Particular Plus Account                               0.043306
Short-Term Deposits                                   0.001779
Medium-Term Deposits                                  0.001661
Long-Term Deposits                                    0.042967
E-Account                                             0.082744
Funds                                                 0.018486
Mortgage                                              0.005887
Pensions                                              0.009171
Loans                                                 0.002627
Taxes                                                 0.052536
Credit Card                                           0.044389
Securities                                            0.025608
Home Account                                          0.003848
Payroll                                               0.054660
Pension                                               0.059359
Direct Debit                                          0.127916
dtype: float64
```

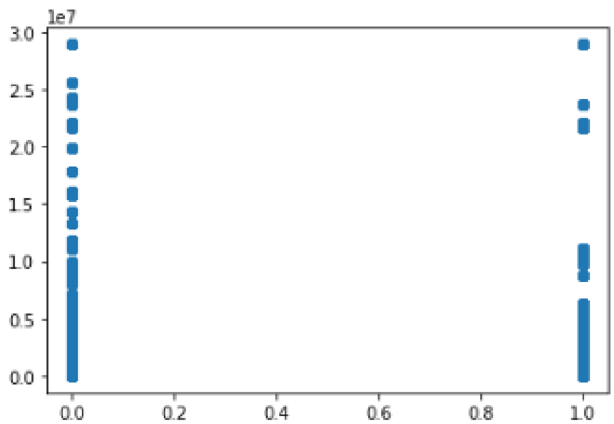
```
Dataframe Median

Sex                                             0.0
Age                                            39.0
New Customer Index                              0.0
Seniority                                      50.0
Primary Customer                                1.0
Resident Country is Bank Country                1.0
Birth Country Different Than Bank Country       0.0
Deceased                                        0.0
Activity Index                                  0.0
Gross Household Income                     101850.0
Savings Account                                 0.0
Guarantees                                      0.0
Current Accounts                                1.0
Derivative Account                              0.0
Payroll Account                                 0.0
Junior Account                                  0.0
More Partiuclar Account                         0.0
Particular Account                              0.0
Particular Plus Account                         0.0
Short-Term Deposits                             0.0
Medium-Term Deposits                            0.0
Long-Term Deposits                              0.0
E-Account                                       0.0
Funds                                           0.0
Mortgage                                        0.0
Pensions                                        0.0
Loans                                           0.0
Taxes                                           0.0
Credit Card                                     0.0
Securities                                      0.0
Home Account                                    0.0
Payroll                                         0.0
Pension                                         0.0
Direct Debit                                    0.0
dtype: float64
```

Then, I created some visualizations with the variables and compared them to gross household income. I did this because I assumed the greater the wealth of a family, the greater the possible amount of money that could be trusted with the bank. Though the visualizations helped a little, they couldn't paint a full picture.

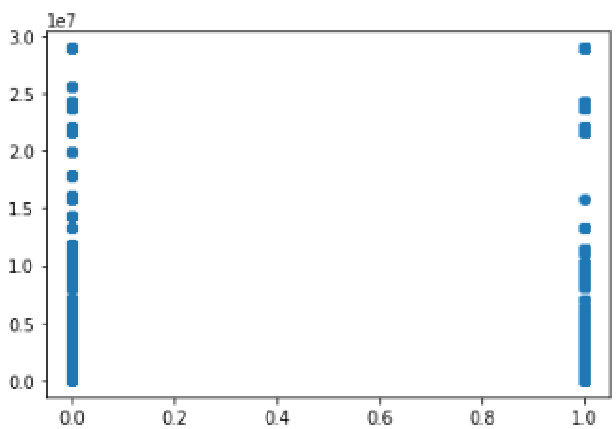
Gross Household Income by Age



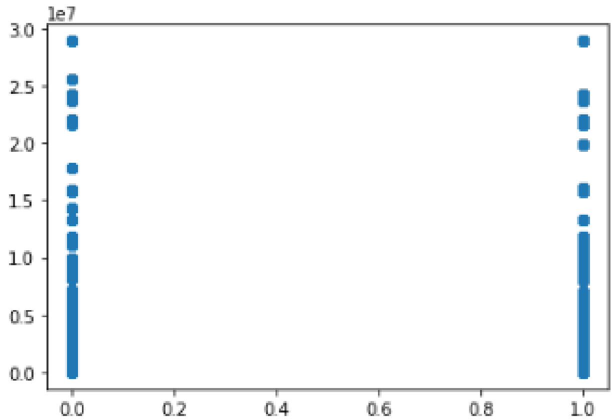
Gross Household Income by E-Account



Gross Household Income by Direct Debit



Gross Household Income by Activity Index



I decided to find the correlation (Pearson's r) values between each of the variables, with special interest in gross household income, to see if I could find any correlations between them. Although it wasn't giving me all the information I wanted, specifically for the variables of sex and age, I changed a couple of the variables to make them able to be correlated. I then analyzed all the statistically significant correlations and they helped me come to a final conclusion. Therefore, with the information I received before from the means and medians, and the recent information with the correlation coefficients, I was able to come to a final recommendation.

Full Correlation Table

	Sex	Age	New Customer Index	Seniority	Primary Customer	Resident Country is Bank Country	Birth Country Different Than Bank Country	Spouse Index	Deceased	Activity Index	Gross Household Income	Savings Account	Guarantees
Sex	1	-0.12359	0.024902	-0.00396	-0.00059	0.012855	0.010159	0.113889	-0.01483	-0.05212	-0.007568	-0.00427	-0.001921
Age	-0.12359	1	-0.15279	0.105567	-0.00087	-0.02946	-0.01652	-0.08281	0.109344	0.135249	0.067631	0.004951	0.000912
New Customer Index	0.024902	-0.15279	1	-0.01114	0.063021	0.01694	0.016996	-0.02621	-0.01263	-0.00424	-0.030742	-0.00255	-0.001019
Seniority	-0.00396	0.105567	-0.01114	1	-0.00103	-0.00219	-0.0009	0.030794	0.002182	0.005027	0.002672	0.000477	0.000226
Primary Customer	-0.00059	-0.00087	0.063021	-0.00103	1	0.002058	0.004602	-0.00562	0.017776	-0.03303	-0.000249	-0.00043	-0.000206
Resident Country is Bank Country	0.012855	-0.02946	0.01694	-0.00219	0.002058	1	-0.246658	NaN	-0.00038	-0.01376	-0.00148	0.000706	-0.0034
Birth Country Different Than Bank Country	0.010159	-0.01652	0.016996	-0.0009	0.004602	-0.24666	1	-0.01244	-0.00712	-0.02109	-0.005723	0.000578	-0.001074
Spouse Index	0.113889	-0.08281	-0.02621	0.030794	-0.00562	NaN	-0.012439	1	-0.00397	0.041176	-0.054439	NaN	-0.009492
Deceased	-0.01483	0.109344	-0.01263	0.002182	0.017776	-0.00038	-0.007117	-0.00397	1	0.001147	0.008964	-0.00051	-0.000244
Activity Index	-0.05212	0.135249	-0.00424	0.005027	-0.03303	-0.01376	-0.021085	0.041176	0.001147	1	0.060746	0.005574	0.005242
Gross Household Income	-0.00757	0.067631	-0.03074	0.002672	-0.00025	-0.00148	-0.005723	-0.05444	0.008964	0.060746	1	0.006067	0.004039
Savings Account	-0.00427	0.004951	-0.00255	0.000477	-0.00043	0.000706	0.000578	NaN	-0.00051	0.005574	0.006067	1	-0.000049
Guarantees	-0.00192	0.000912	-0.00102	0.000226	-0.00021	-0.0034	-0.001074	-0.00949	-0.00024	0.005242	0.004039	-4.9E-05	1
Current Accounts	0.036655	-0.16767	0.002846	0.036575	-0.04695	-0.01251	-0.023655	-0.11853	-0.01394	0.170152	-0.053887	-0.00108	-0.001767
Derivative Account	-0.01263	0.013363	-0.00459	0.000597	-0.00085	0.000479	-0.003244	NaN	0.000241	0.017899	0.004521	-0.0002	-0.000096
Payroll Account	-0.02719	0.068959	-0.04699	0.009866	-0.01154	0.016264	-0.004743	0.172047	-0.01057	0.31544	0.034652	0.001146	0.010302
Junior Account	0.005916	-0.15239	-0.01122	0.000278	-0.00256	0.006632	-0.019236	NaN	-0.00495	0.071464	0.02402	-0.00099	-0.000471
More Partiuclar Account	-0.01049	0.035484	0.041261	-0.00363	-0.00096	0.003698	-0.000673	-0.0103	-0.00247	0.066172	0.007335	-0.00071	0.001695
Particular Account	-0.06242	0.317002	-0.09691	0.012087	-0.01425	-0.02021	-0.034163	-0.06007	0.031835	0.169366	0.085797	0.009616	0.001828
Particular Plus Account	-0.02937	0.117488	-0.05356	0.007013	-0.00802	0.00404	-0.012588	-0.05426	0.001911	0.156211	0.051297	0.004184	0.008702
Short-Term Deposits	-0.00207	0.017725	0.087791	-0.01518	0.00666	0.000918	-0.005495	-0.00459	0.00025	0.036943	0.000305	-0.00043	-0.000203

	Current Accounts	Derivative Account	Payroll Account	Junior Account	More Partiuclar Account	Particular Account	Particular Plus Account	Short-Term Deposits	Medium-Term Deposits	Long-Term Deposits	E-Account	Funds	Mortgage	Pensions
Sex	0.036655	-0.01263	-0.02719	0.005916	-0.01049	-0.06242	-0.02937	-0.002071	-0.00329	-0.023505	-0.042111	-0.04157	-0.02664	-0.01625
Age	-0.16767	0.013363	0.068959	-0.15239	0.035484	0.317002	0.117488	0.017725	0.028532	0.175474	0.142814	0.115399	0.037736	0.065969
New Customer Index	0.002846	-0.00459	-0.04699	-0.01122	0.041261	-0.09691	-0.05356	0.087791	-0.01005	-0.03955	-0.059726	-0.03274	-0.01931	-0.02328
Seniority	0.036575	0.000597	0.009866	0.000278	-0.00363	0.012087	0.007013	-0.015184	0.001694	0.008428	0.006571	0.004875	-0.00126	0.003134
Primary Customer	-0.04695	-0.00085	-0.01154	-0.00256	-0.00096	-0.01425	-0.00802	0.00666	-0.00115	-0.008047	-0.009846	-0.0053	-0.00315	-0.00402
Resident Country is Bank Country	-0.01251	0.000479	0.016264	0.006632	0.003698	-0.02021	0.00404	0.000918	-0.00596	0.001642	0.004881	-0.00793	0.0033	-0.00175
Birth Country Different Than Bank Country	-0.02366	-0.00324	-0.00474	-0.01924	-0.00067	-0.03416	-0.01259	-0.005495	-0.00303	-0.03154	-0.028999	-0.0168	-0.01143	-0.00495
Spouse Index	-0.11853	NaN	0.172047	NaN	-0.0103	-0.06007	-0.05426	-0.004588	-0.00562	-0.04302	0.053851	-0.04354	-0.03199	0.143582
Deceased	-0.01394	0.000241	-0.01057	-0.00495	-0.00247	0.031835	0.001911	0.00025	-0.0001	-0.001407	-0.00603	0.009688	0.003282	0.002177
Activity Index	0.170152	0.017899	0.31544	0.071464	0.066172	0.169366	0.156211	0.036943	0.044162	0.228881	0.284509	0.147586	0.083339	0.103684
Gross Household Income	-0.05389	0.004521	0.034652	0.02402	0.007335	0.085797	0.051297	0.000305	0.006347	0.045854	0.054914	0.055592	0.008116	0.044143
Savings Account	-0.00108	-0.0002	0.001146	-0.00099	-0.00071	0.009616	0.004184	-0.000427	-0.00041	0.002037	0.001565	0.001355	0.002442	0.005639
Guarantees	-0.00177	-9.6E-05	0.010302	-0.00047	0.001695	0.001828	0.008702	-0.000203	0.000552	0.001609	0.011931	0.003975	-0.00037	0.002253
Current Accounts	1	0.00321	-0.34547	-0.13482	-0.02465	0.003798	-0.05078	-0.003055	0.012498	0.010251	-0.076882	0.010201	-0.04697	-0.03073
Derivative Account	0.00321	1	0.008509	-0.00194	-0.00152	0.012585	0.007151	0.000652	0.001095	0.004279	0.010484	0.017096	0.001513	0.013004
Payroll Account	-0.34547	0.008509	1	-0.02901	0.024416	0.021291	0.186533	-0.002743	0.014493	0.095624	0.306321	0.085948	0.155558	0.102922
Junior Account	-0.13482	-0.00194	-0.02901	1	-0.00969	-0.03764	-0.02081	-0.004128	-0.00399	-0.019685	-0.029327	-0.01342	-0.00753	-0.00941
More Partiuclar Account	-0.02465	-0.00152	0.024416	-0.00969	1	-0.03603	-0.0183	0.016547	-0.00074	0.007782	-0.011362	-0.00215	-0.00507	-0.00264
Particular Account	0.003798	0.012585	0.021291	-0.03764	-0.03603	1	0.037871	0.000567	0.04065	0.077666	-0.0102	0.087223	0.018857	0.049246
Particular Plus Account	-0.05078	0.007151	0.186533	-0.02081	-0.0183	0.037871	1	-0.000412	0.022891	0.074684	0.067205	0.067732	0.07065	0.062644
Short-Term Deposits	-0.00306	0.000652	-0.00274	-0.00413	0.016547	0.000567	-0.00041	1	0.050163	0.014158	-0.0034	0.003708	-0.00132	-0.00198

Loans	Taxes	Credit Card	Securities	Home Account	Payroll	Pension	Direct Debit
-0.01452	-0.05108	-0.047341	-0.05463	-0.01021	-0.01986	-0.02106	-0.04729
0.019566	0.087455	0.10082	0.110106	0.017366	0.043252	0.07142	0.08864
-0.01258	-0.05501	-0.051815	-0.038225	-0.01561	-0.03688	-0.03937	-0.053103
-0.04	-0.06337	0.008151	0.005861	0.002215	0.004929	0.005156	0.012985
-0.00188	-0.00602	-0.008518	-0.006373	-0.00266	-0.00926	-0.00965	-0.011989
0.002083	0.007568	0.003258	-0.018717	0.003174	0.014824	0.014919	0.011704
0.031036	0.007773	-0.001367	-0.017942	-0.00279	-0.00689	-0.00856	-0.003418
-0.02504	-0.05459	-0.074667	0.132255	-0.01349	0.227092	0.21249	0.049243
0.003296	0.00101	-0.010262	0.007082	-0.00183	-0.01147	-0.01102	-0.012925
0.036457	0.222989	0.23232	0.17332	0.039814	0.260518	0.272069	0.412218
-0.01063	0.072117	0.05606	0.067869	0.023001	0.026596	0.032057	0.038665
-0.00052	0.00762	0.001444	0.00625	0.014351	0.000242	0.000093	0.00322
-0.00025	0.011016	0.010423	0.013198	-0.0003	0.011704	0.011163	0.011151
-0.02876	-0.09158	-0.095949	0.014066	-0.00551	-0.25773	-0.27089	-0.069308
0.001434	0.011231	0.009599	0.033319	0.00479	0.006868	0.0074	0.009136
0.002517	0.319811	0.386592	0.122338	0.030378	0.757957	0.790128	0.535626
-0.00502	-0.02241	-0.021078	-0.015855	-0.00608	-0.02327	-0.02322	-0.037449
-0.00417	0.013283	0.02031	0.001634	-0.00564	0.021845	0.021794	0.029004
-0.00911	0.061071	0.063697	0.096055	0.062589	0.027068	0.028114	0.04304
-0.00185	0.168267	0.167066	0.092291	0.040633	0.172543	0.179819	0.201764
-0.00217	-0.00152	-0.001279	0.00105	-0.00116	-0.00128	-0.00157	-0.004947

Medium-Term Deposits	-0.00329	0.028532	-0.01005	0.001694	-0.00115	-0.00596	-0.003027	-0.00562	-0.0001	0.044162	0.006347	-0.00041	0.000552	0.012498	0.001095	0.014493	-0.00399	-0.00074	0.04065	0.022891	0.050163	1	0.079801	0.016717	0.03027	0.004506	0.011215
Long-Term Deposits	-0.02351	0.175474	-0.03955	0.008428	-0.00805	0.001642	-0.03154	-0.04302	-0.00141	0.228881	0.045854	0.002037	0.001609	0.010251	0.004279	0.095624	-0.01969	0.007782	0.077666	0.074684	0.014158	0.079801	1	0.22212	0.177314	0.016609	0.060747
E-Account	-0.04211	0.142814	-0.05973	0.006571	-0.00985	0.004881	-0.028999	0.053851	-0.00603	0.284509	0.054914	0.001565	0.011931	-0.07688	0.010484	0.306321	-0.02933	-0.01136	-0.0102	0.067205	-0.0034	0.016717	0.22212	1	0.142475	0.066474	0.090767
Funds	-0.04157	0.115399	-0.03274	0.004875	-0.0053	-0.00793	-0.016804	-0.04354	0.009688	0.147586	0.055592	0.001355	0.003975	0.010201	0.017096	0.085948	-0.01342	-0.00215	0.087223	0.067732	0.003708	0.03027	0.177314	0.142475	1	0.019263	0.126414
Mortgage	-0.02664	0.037736	-0.01931	-0.00126	-0.00315	0.0033	-0.011434	-0.03199	0.003282	0.083339	0.008116	0.002442	-0.00037	-0.04697	0.001513	0.155558	-0.00753	-0.00507	0.018857	0.07065	-0.001316	0.004506	0.016609	0.066474	0.019263	1	0.041512
Pensions	-0.01625	0.065969	-0.02328	0.003134	-0.00402	-0.00175	-0.004954	0.143582	0.002177	0.103684	0.044143	0.005639	0.002253	-0.03073	0.013004	0.102922	-0.00941	-0.00264	0.049246	0.062644	-0.001982	0.011215	0.060747	0.090767	0.126414	0.041512	1
Loans	-0.01452	0.019566	-0.01258	-0.04	-0.00188	0.002083	0.031036	-0.02504	0.003296	0.036457	-0.010625	-0.00052	-0.000247	-0.02876	0.001434	0.002517	-0.00502	-0.00417	-0.00911	-0.00185	-0.002167	-0.0015	-0.009541	-0.008074	0.001723	0.010664	0.005768
Taxes	-0.05108	0.087455	-0.05501	-0.06337	-0.00602	0.007568	0.007773	-0.05459	0.00101	0.222989	0.072117	0.00762	0.011016	-0.09158	0.011231	0.319811	-0.02241	0.013283	0.061071	0.168267	-0.001522	0.01155	0.063987	0.200133	0.069055	0.089504	0.082939
Credit Card	-0.04734	0.10082	-0.05182	0.008151	-0.00852	0.003258	-0.001367	-0.07467	-0.01026	0.23232	0.05606	0.001444	0.010423	-0.09595	0.009599	0.386592	-0.02108	0.02031	0.063697	0.167066	-0.001279	0.015925	0.106513	0.255686	0.100685	0.1192	0.110744
Securities	-0.05463	0.110106	-0.03823	0.005861	-0.00637	-0.01872	-0.017942	0.132255	0.007082	0.17332	0.067869	0.00625	0.013198	0.014066	0.033319	0.122338	-0.01586	0.001634	0.096055	0.092291	0.00105	0.017594	0.081671	0.150803	0.179663	0.044399	0.116294
Home Account	-0.01021	0.017366	-0.01561	0.002215	-0.00266	0.003174	-0.002788	-0.01349	-0.00183	0.039814	0.023001	0.014351	-0.000299	-0.00551	0.00479	0.030378	-0.00608	-0.00564	0.062589	0.040633	-0.001163	0.007144	0.007073	0.01038	0.018472	0.006017	0.022068
Payroll	-0.01986	0.043252	-0.03688	0.004929	-0.00926	0.014824	-0.006887	0.227092	-0.01147	0.260518	0.026596	0.000242	0.011704	-0.25773	0.006868	0.757957	-0.02327	0.021845	0.027068	0.172543	-0.001281	0.015789	0.078139	0.263043	0.072952	0.16493	0.09907
Pension	-0.02106	0.07142	-0.03937	0.005156	-0.00965	0.014919	-0.008559	0.21249	-0.01102	0.272069	0.032057	0.000093	0.011163	-0.27089	0.0074	0.790128	-0.02322	0.021794	0.028114	0.179819	-0.001565	0.016167	0.090496	0.281967	0.082387	0.162163	0.099784
Direct Debit	-0.04729	0.08864	-0.0531	0.012985	-0.01199	0.011704	-0.003418	0.049243	-0.01293	0.412218	0.038665	0.00322	0.011151	-0.06931	0.009136	0.535626	-0.03745	0.029004	0.04304	0.201764	-0.004947	0.014645	0.09392	0.296159	0.093503	0.169577	0.104648
# of Greens		1				1	1	2		8				3		8			1	1			2	8			

-0.0015	0.01155	0.015925	0.017594	0.007144	0.015789	0.016167	0.014645
0.00954	0.063987	0.106513	0.081671	0.007073	0.078139	0.090496	0.09392
0.00807	0.200133	0.255686	0.150803	0.01038	0.263043	0.281967	0.296159
.001723	0.069055	0.100685	0.179663	0.018472	0.072952	0.082387	0.093503
.010664	0.089504	0.1192	0.044399	0.006017	0.16493	0.162163	0.169577
.005768	0.082939	0.110744	0.116294	0.022068	0.09907	0.099784	0.104648
1	0.029758	0.011665	0.008507	-0.00037	0.005232	0.005542	0.004047
.029758	1	0.269161	0.113165	0.052966	0.277961	0.294729	0.343852
.011665	0.269161	1	0.140177	0.033704	0.356682	0.370464	0.37466
.008507	0.113165	0.140177	1	0.033142	0.113341	0.122808	0.144288
0.00037	0.052966	0.033704	0.033142	1	0.030462	0.028839	0.035847
.005232	0.277961	0.356682	0.113341	0.030462	1	0.957215	0.471138
.005542	0.294729	0.370464	0.122808	0.028839	0.957215	1	0.490888
.004047	0.343852	0.37466	0.144288	0.035847	0.471138	0.490888	1
	7	7			9	9	8

Statistically Significant Findings

	Age	Resident Country is Bank Country	Birth Country Different Than Bank Country	Spouse Index	Activity Index	Current Accounts	Payroll Account	Particular Account	Particular Plus Account	Long-Term Deposits	E-Account	Taxes	Credit Card	Payroll	Pension	Direct Debit
Age	1	-0.02946	-0.01652	-0.08281	0.135249	-0.16767	0.068959	0.317002	0.117488	0.175474	0.142814	0.087455	0.10082	0.043252	0.07142	0.08864
Resident Country is Bank Country	-0.02946	1	-0.246658	NaN	-0.01376	-0.01251	0.016264	-0.02021	0.00404	0.001642	0.004881	0.007568	0.003258	0.014824	0.014919	0.011704
Birth Country Different Than Bank Country	-0.01652	-0.24666	1	-0.01244	-0.02109	-0.02366	-0.00474	-0.03416	-0.01259	-0.03154	-0.028999	0.007773	-0.001367	-0.00689	-0.00856	-0.003418
Spouse Index	-0.08281	NaN	-0.012439	1	0.041176	-0.11853	0.172047	-0.06007	-0.05426	-0.04302	0.053851	-0.05459	-0.074667	0.227092	0.21249	0.049243
Activity Index	0.135249	-0.01376	-0.021085	0.041176	1	0.170152	0.31544	0.169366	0.156211	0.228881	0.284509	0.222989	0.23232	0.260518	0.272069	0.412218
Current Accounts	-0.16767	-0.01251	-0.023655	-0.11853	0.170152	1	-0.34547	0.003798	-0.05078	0.010251	-0.076882	-0.09158	-0.095949	-0.25773	-0.27089	-0.069308
Payroll Account	0.068959	0.016264	-0.004743	0.172047	0.31544	-0.34547	1	0.021291	0.186533	0.095624	0.306321	0.319811	0.386592	0.757957	0.790128	0.535626
Particular Account	0.317002	-0.02021	-0.034163	-0.06007	0.169366	0.003798	0.021291	1	0.037871	0.077666	-0.0102	0.061071	0.063697	0.027068	0.028114	0.04304
Particular Plus Account	0.117488	0.00404	-0.012588	-0.05426	0.156211	-0.05078	0.186533	0.037871	1	0.074684	0.067205	0.168267	0.167066	0.172543	0.179819	0.201764
Long-Term Deposits	0.175474	0.001642	-0.03154	-0.04302	0.228881	0.010251	0.095624	0.077666	0.074684	1	0.22212	0.063987	0.106513	0.078139	0.090496	0.09392
E-Account	0.142814	0.004881	-0.028999	0.053851	0.284509	-0.07688	0.306321	-0.0102	0.067205	0.22212	1	0.200133	0.255686	0.263043	0.281967	0.296159
Taxes	0.087455	0.007568	0.007773	-0.05459	0.222989	-0.09158	0.319811	0.061071	0.168267	0.063987	0.200133	1	0.269161	0.277961	0.294729	0.343852
Credit Card	0.10082	0.003258	-0.001367	-0.07467	0.23232	-0.09595	0.386592	0.063697	0.167066	0.106513	0.255686	0.269161	1	0.356682	0.370464	0.37466
Payroll	0.043252	0.014824	-0.006887	0.227092	0.260518	-0.25773	0.757957	0.027068	0.172543	0.078139	0.263043	0.277961	0.356682	1	0.957215	0.471138
Pension	0.07142	0.014919	-0.008559	0.21249	0.272069	-0.27089	0.790128	0.028114	0.179819	0.090496	0.281967	0.294729	0.370464	0.957215	1	0.490888
Direct Debit	0.08864	0.011704	-0.003418	0.049243	0.412218	-0.06931	0.535626	0.04304	0.201764	0.09392	0.296159	0.343852	0.37466	0.471138	0.490888	1
# of Greens	1	1	1	2	8	3	8	1	1	2	8	7	7	9	9	8

Final Recommendation:

I would like to begin this final recommendation with the saying that, "Correlation does not imply causation". Although trends say one thing, you are not always guaranteed the outcome you seek. With that being said, my final recommendation is this. In the prompt, we are told the XYZ Credit Union is performing very well in the selling of their banking products. Although they aren't where they'd like to be with cross selling to their existing customers, they are still doing well selling individual products. They might not be maximizing their revenue, but their business model is drawing in new customers and keeping them long term. Pushing new accounts on your existing customers could turn them away. I would say you could try cross selling to new customers that enter your credit union, but not the ones who have been loyal for many years. Despite this, there are ways of making more profit by cross selling certain accounts together. So, if you did want to try making more revenue through cross selling, this is how you could do it.

Shockingly, from the correlation coefficient findings, gross household income is not a variable which should be tested with bank success. It wasn't statistically significant with any other of the variables we used. The variables best correlated with other banking variables were activity index, payroll account, e-account, taxes, credit card, payroll, pensions, and direct debit. Looking deeper into a few of those variables, starting with activity index, we can see that it's statistically significantly positively correlated with payroll accounts, long-term deposits, e-accounts, taxes, credit card, payroll, pensions, and direct debit. In short, this means that customers who are active with their account tend to also have multiple types of banking accounts, specifically the accounts listed above. This means that if you were trying to cross sell, do it with active customers. Try getting them signed up for one of these account types, with a priority in direct debit, e-accounts, payroll accounts, pensions, payroll, taxes, and credit card, as these are the most used accounts out of all their customers. Moreover, these accounts are all statistically significantly positively correlated with each other.

If you have an older customer, the statistics show that they are more likely to have a particular account with the credit union. Also, particular accounts are one of the most popular accounts to have in this credit union. Thus, I would try pushing those types of accounts to older customers who don't already have one.

Customers that already have an account at the credit union tend to not have a payroll account or pension. This probably means that they use other banking competitors to handle those accounts. Maybe with some more open minded customers, you could be able to convince them to switch over to your credit union. This would make those variables more highly correlated, especially if they are an active customer, and could create more usage for your credit union.

Particular plus accounts have a statistically significantly positive relationship with direct debit. This makes sense as particular plus accounts deal with checking accounts and so do debit cards. Therefore, if someone is trying to sign up for a checking (particular plus) account, try offering direct debit, as it would line up with the wants of the customer.

To keep up with an increasing technologically driven world, the statistics and I highly recommend pushing e-accounts. It is positively correlated with many of the highest used banking accounts, and it provides simplicity and convenience for many, if not all, of your customers.

There are no statistically significant differences between gender and banking. Although 55% of the customers are male, there is no difference in the amount of accounts that both males and females use. Thus, you shouldn't treat them differently, and you shouldn't offer different accounts based on sex.

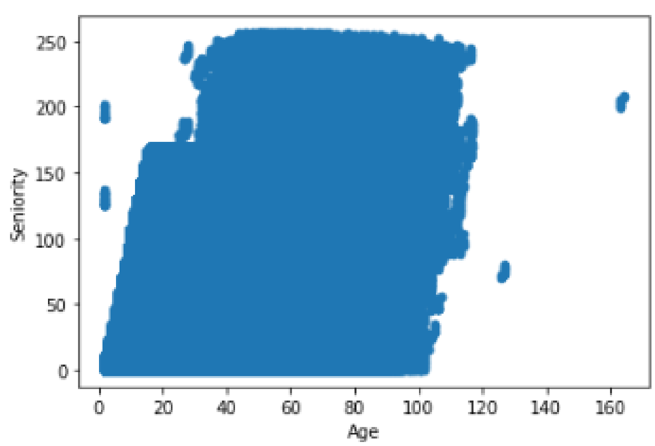
Finally, using the means of each of the different banking account types, I have found that the top five most used accounts are: particular account, direct debit, e-account, payroll account, and pension; and the five least used accounts are: guarantees, savings account, derivative account, medium-term deposits, and short-term deposits. With this information, you can promote what is working best for the XYZ Credit Union, or try promoting the lesser used accounts in hopes of attracting different types of customers. However, I wouldn't try and deviate too far from what is already being done, since the XYZ Credit Union is doing well already.

Modeling Technique:

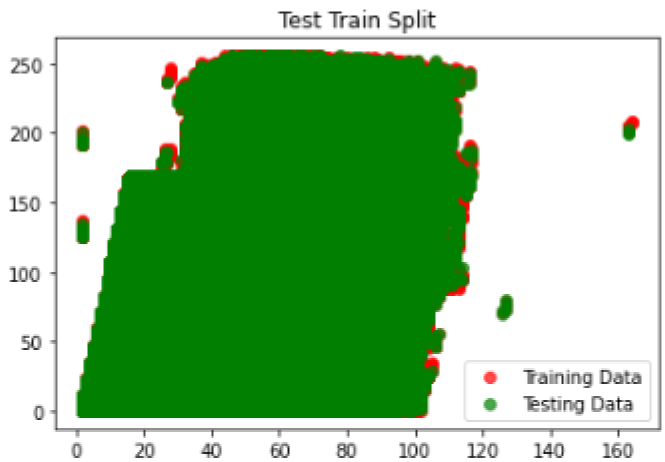
The best models for this problem/dataset should be predictive models. Specifically, models that use linear and multiple regression. Using the statistically significant values from the correlation coefficient table, along with means, medians, counts, sums, and regression equations, we can attempt to predict the best combination of cross selling techniques for the XYZ Credit Union.

Linear Regression Model:

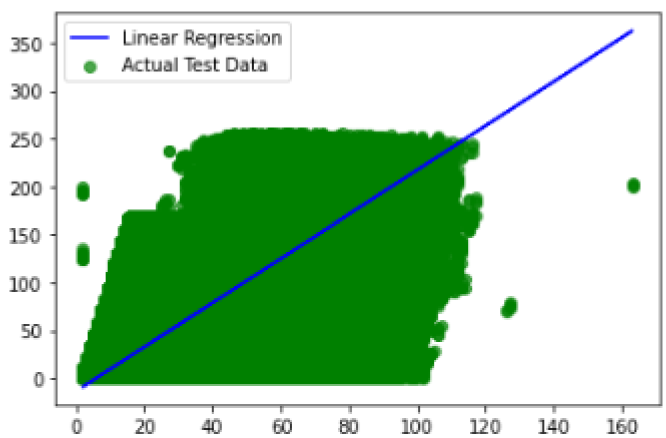
The first predictive model I used to attempt to predict the best combination of cross selling techniques was linear regression. This is an example of what an output would look like. I started with a regular scatterplot to get an overall understanding of each of the variables being used.



I then used the train, test, split as my method of prediction, and made a graph.



Next, I plotted the actual test data with a linear regression line to make an accurate graph.



Finally, I created a regression equation, with a linear regression score, and was able to make a prediction based on inputting a number for x.

```
Prediction 0.18993595689775208
Intercept -0.12016394257993027
Slope [0.006202]
Regression Equation: y = [0.006202] x + -0.12016394257993027
Linear Regression Score 0.10090753170178268
```

Multiple Regression:

The second predictive model I used was multiple regression. I focused less on visualizations for multiple regression as it is similar to linear regression, except it has more regression lines. I took the variables with two or more other variables that were statistically significantly correlated to it and tested the multiple regression. I was able to make a prediction for each dependent variable, along with finding the regression equation with the slope values and an intercept. Here is an example of what the output looks like:

```
Prediction [0.60601549]
Coefficients [ 0.07142447  0.10559897  0.20986961  0.06764832  0.13505848 -0.08433393
  0.13916534  0.07420132]
Intercept 0.009862632917611497
Regression Equation: y = [ 0.07142447  0.10559897  0.20986961  0.06764832  0.13505848 -0.08433393
  0.13916534  0.07420132] x + 0.009862632917611497
```

Ensemble Model:

I noticed while using linear regression that the prediction score was very low for most of the variables. Multiple regression would make that number a little better but not enough to feel confident in an accurate prediction. Thus, I decided to use an ensemble model to see if I could increase prediction accuracy. Similar to linear regression, I used a test, train, split model to find a baseline prediction. Then, using a bagging classifier and increasing the number, and therefore accuracy, of parameters like n estimators, I was able to output a better score of prediction. Here is an example of an output. The top number is the model using the train values of the variables, and the bottom number is the model using the test values of the variables. We can see that the outcomes are very good, all producing an accuracy score of 95% or better.

```
bag_model = BaggingClassifier(
    base_estimator = DecisionTreeClassifier(),
    n_estimators = 10,
    max_samples = 0.8,
    oob_score = True,
    random_state = 0)

bag_model.fit(A_train, b_train)
print(bag_model.oob_score_)

print(bag_model.score(A_test, b_test))
```

```
0.9979651176139157
0.9982795146765898
```

Boosted Model (Gradient Boosting Model):

Although the numbers of the ensemble model were great, I think it's important to have multiple accurate models to use in case one isn't available to us, or to see if we can make the accuracy score better still. For this model, I used a boosted model (gradient boosting model) to test the accuracy of the model. This one too used a train, test, split approach, then, using a gradient boosting classifier, I was able to make two predictions. Here is an example of an output. The top prediction based on the variables' tested accuracy score, and the bottom prediction testing it against the gradient boosting model's predicted variables. The model generated very accurate scores, much more accurate than the linear regression scores. However, not as accurate as the scores from the ensemble model. I would still have confidence using the boosted model.

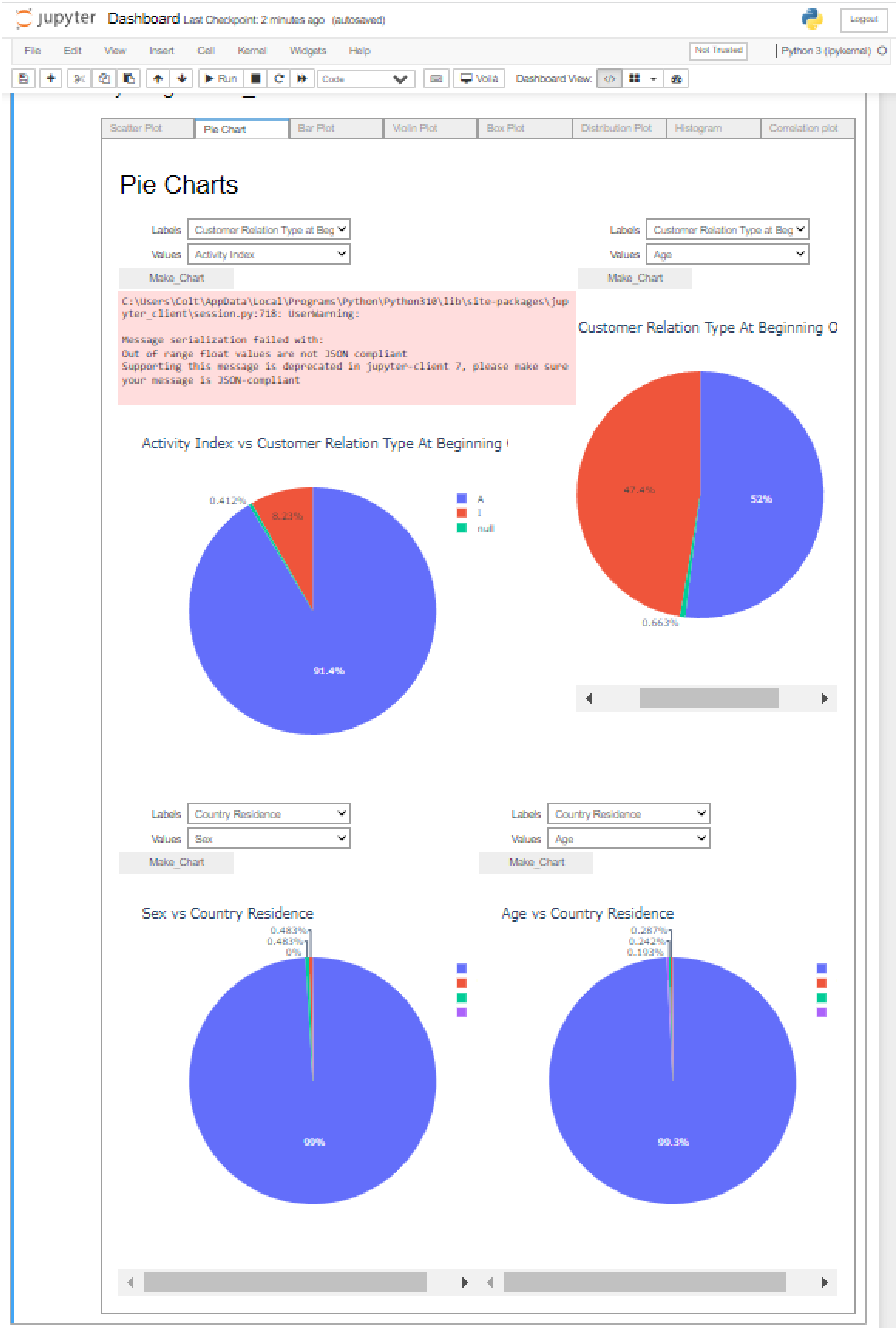
```
gbModel = GradientBoostingClassifier(random_state = 42)

gbModel.fit(S_train, t_train)

t_pred_gb = gbModel.predict(S_test)

print(gbModel.score(S_test, t_test))

print(accuracy_score(t_test, t_pred_gb))
```

```
0.9256824238622849
0.9256824238622849
```

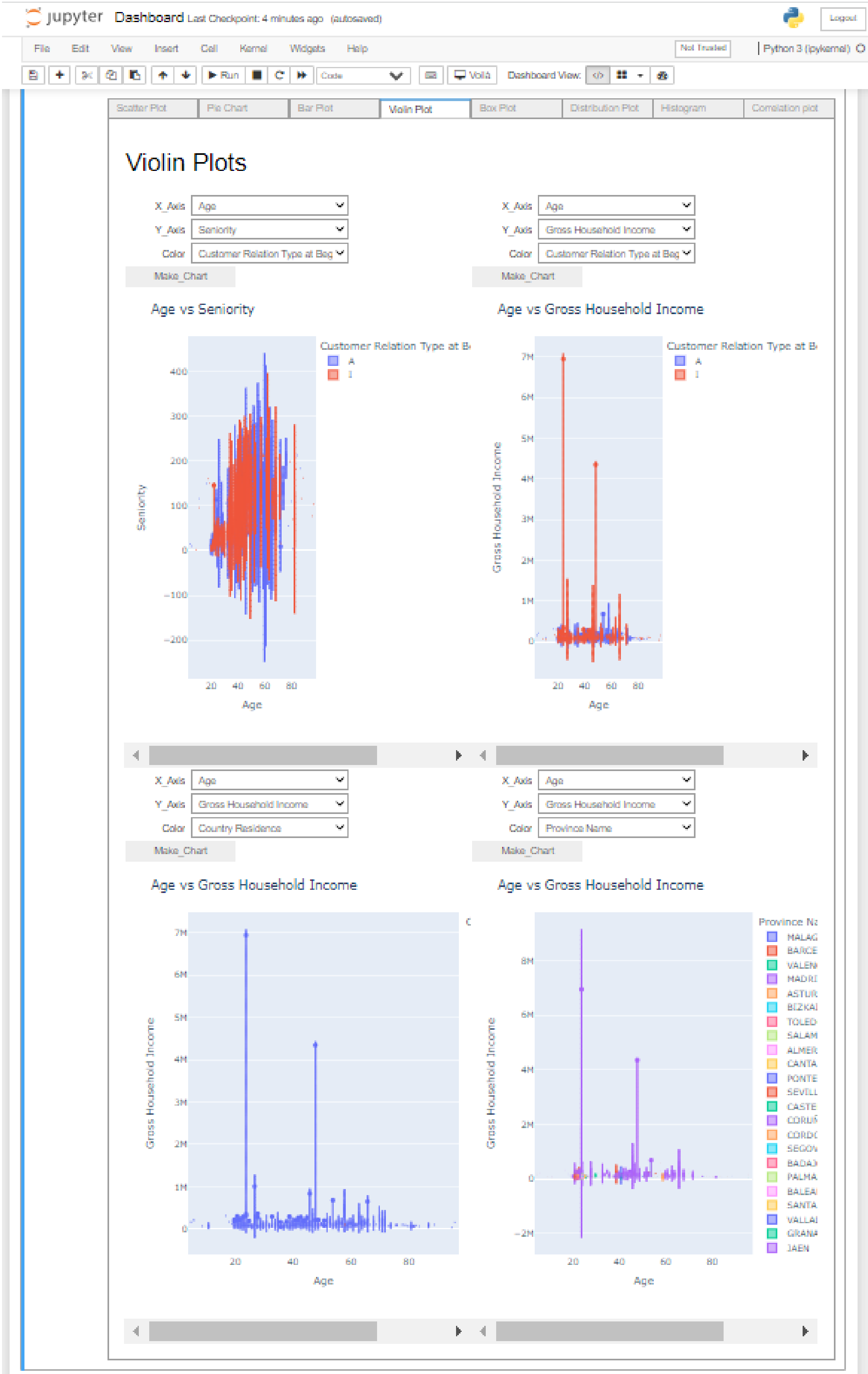
Dashboard Visualizations:

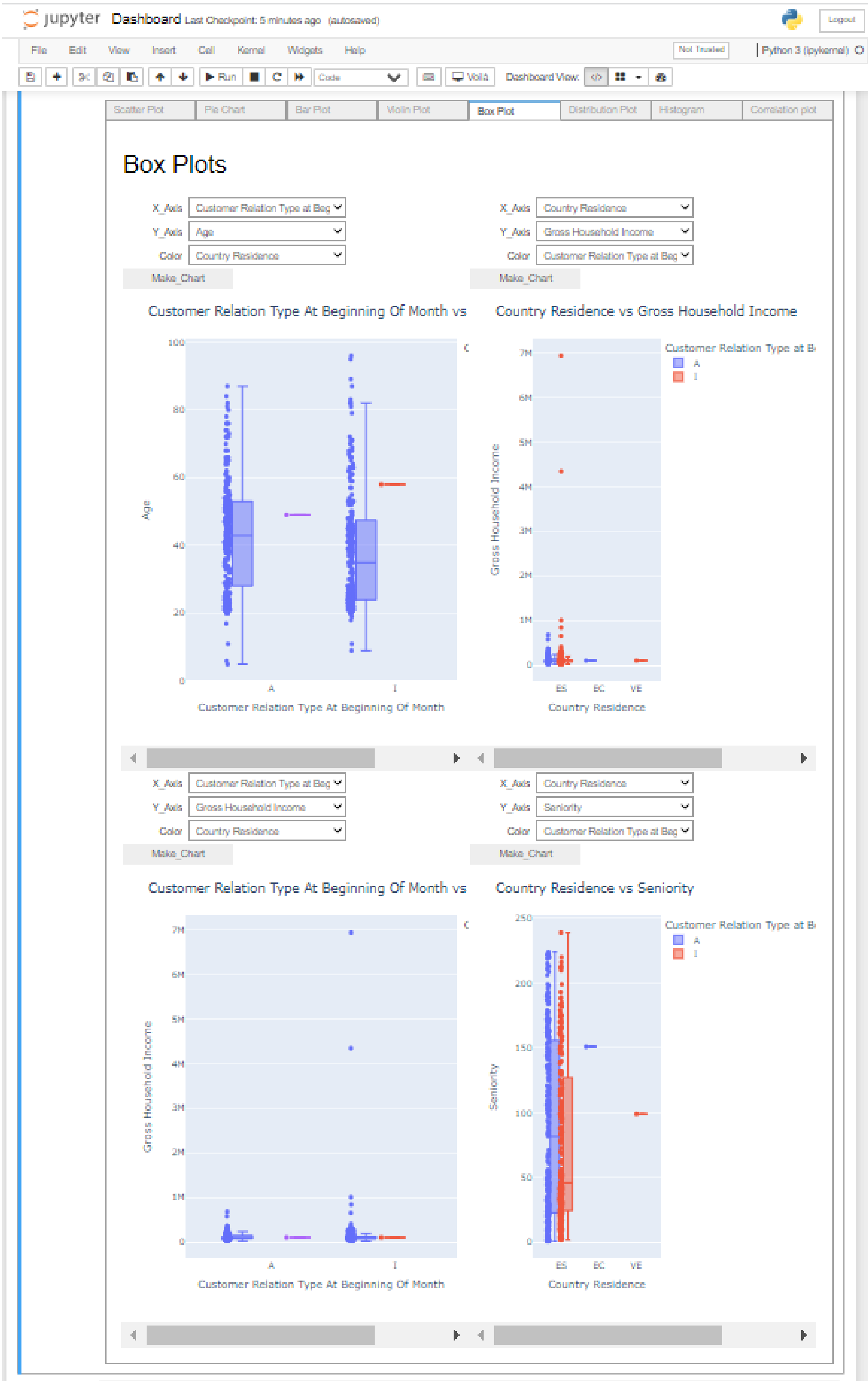
The next couple pictures will be visualizations I created for my dashboard. I won't be repeat the visualizations seen earlier in the presentation, although they should be considered part of the dashboard. More graphs can be created through Python; however, these were the graphs I chose to use as part of the presentation.


jupyter Dashboard Last Checkpoint: a minute ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Scatter Plots
X_Axis Age
Y_Axis Seniority
Color Activity Index
Make_Chart
X_Axis Age
Y_Axis Gross Household Income
Color E-Account
Make_Chart
Seniority vs Age
Gross Household Income vs Age
X_Axis Age
Y_Axis Gross Household Income
Color Sex
Make_Chart
X_Axis Seniority
Y_Axis Gross Household Income
Color Particular Account
Make_Chart
Gross Household Income vs Age
Gross Household Income vs Seniority

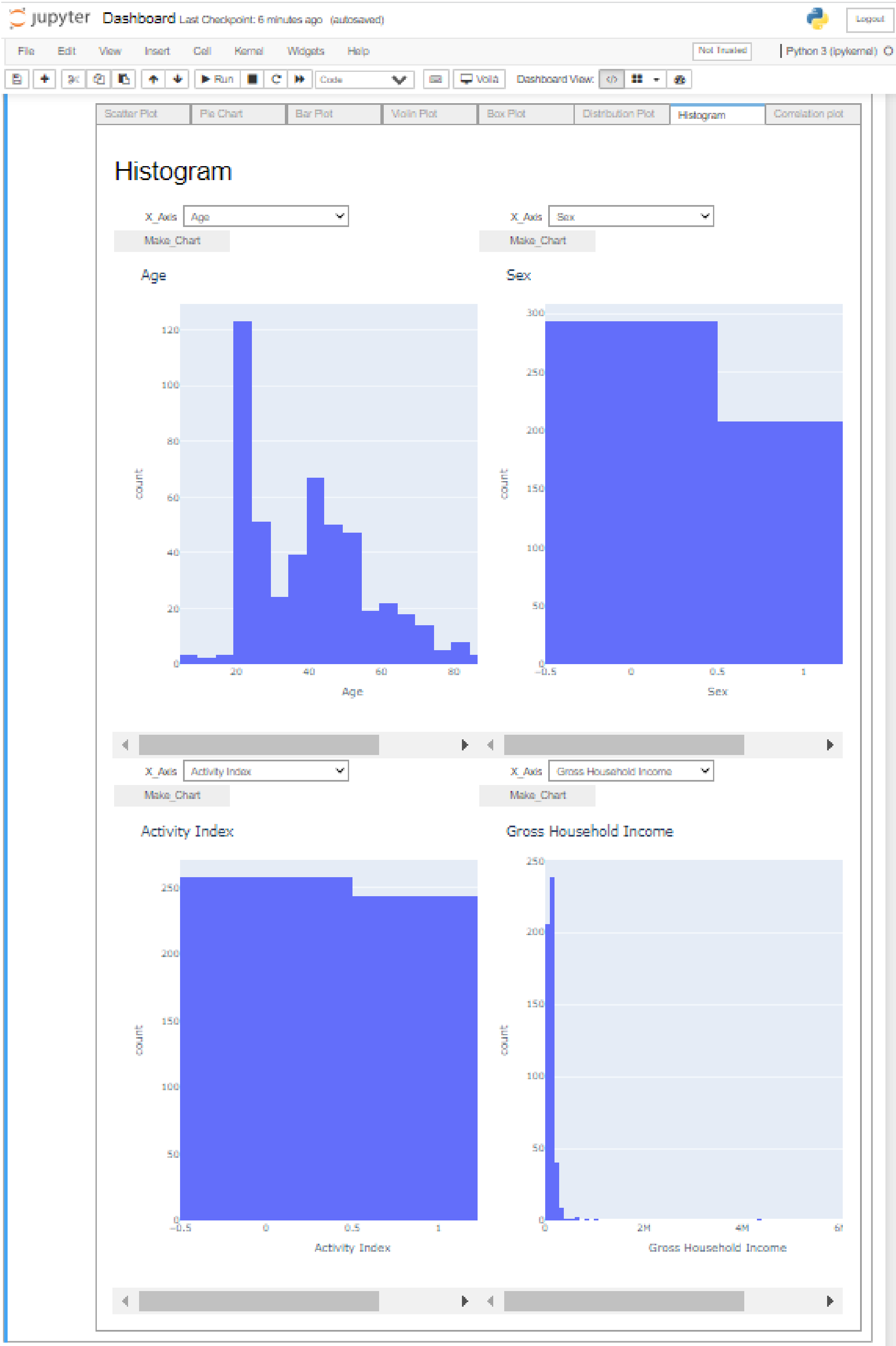

jupyter Dashboard Last Checkpoint: 2 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Pie Charts
Labels Customer Relation Type at Beg
Values Activity Index
Make_Chart
Labels Customer Relation Type at Beg
Values Age
Make_Chart
C:\Users\Colt\AppData\Local\Programs\Python\Python310\lib\site-packages\jupyter_client\session.py:718: UserWarning:
Message serialization failed with:
Out of range float values are not JSON compliant
Supporting this message is deprecated in jupyter-client 7, please make sure your message is JSON-compliant
Activity Index vs Customer Relation Type At Beginning
0.412%
8.23%
91.4%
A
I
null
Customer Relation Type At Beginning O
47.4%
52%
0.663%
Labels Country Residence
Values Sex
Make_Chart
Labels Country Residence
Values Age
Make_Chart
Sex vs Country Residence
0.483%
0.483%
0%
99%
Age vs Country Residence
0.287%
0.242%
0.193%
99.3%


jupyter Dashboard Last Checkpoint: 3 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Bar Plots
X_Axis Country Residence
Y_Axis Gross Household Income
Color Customer Relation Type at Beg
Make_Chart
X_Axis Province Name
Y_Axis Age
Color Country Residence
Make_Chart
Country Residence vs Gross Household Income
Gross Household Income
Country Residence
Province Name vs Age
Age
Province Name
X_Axis Customer Relation Type at Beg
Y_Axis Seniority
Color Province Name
Make_Chart
X_Axis Country Residence
Y_Axis Age
Color Customer Relation Type at Beg
Make_Chart
Customer Relation Type At Beginning Of Month vs
Seniority
Customer Relation Type At Beginning Of Month
Country Residence vs Age
Age
Country Residence

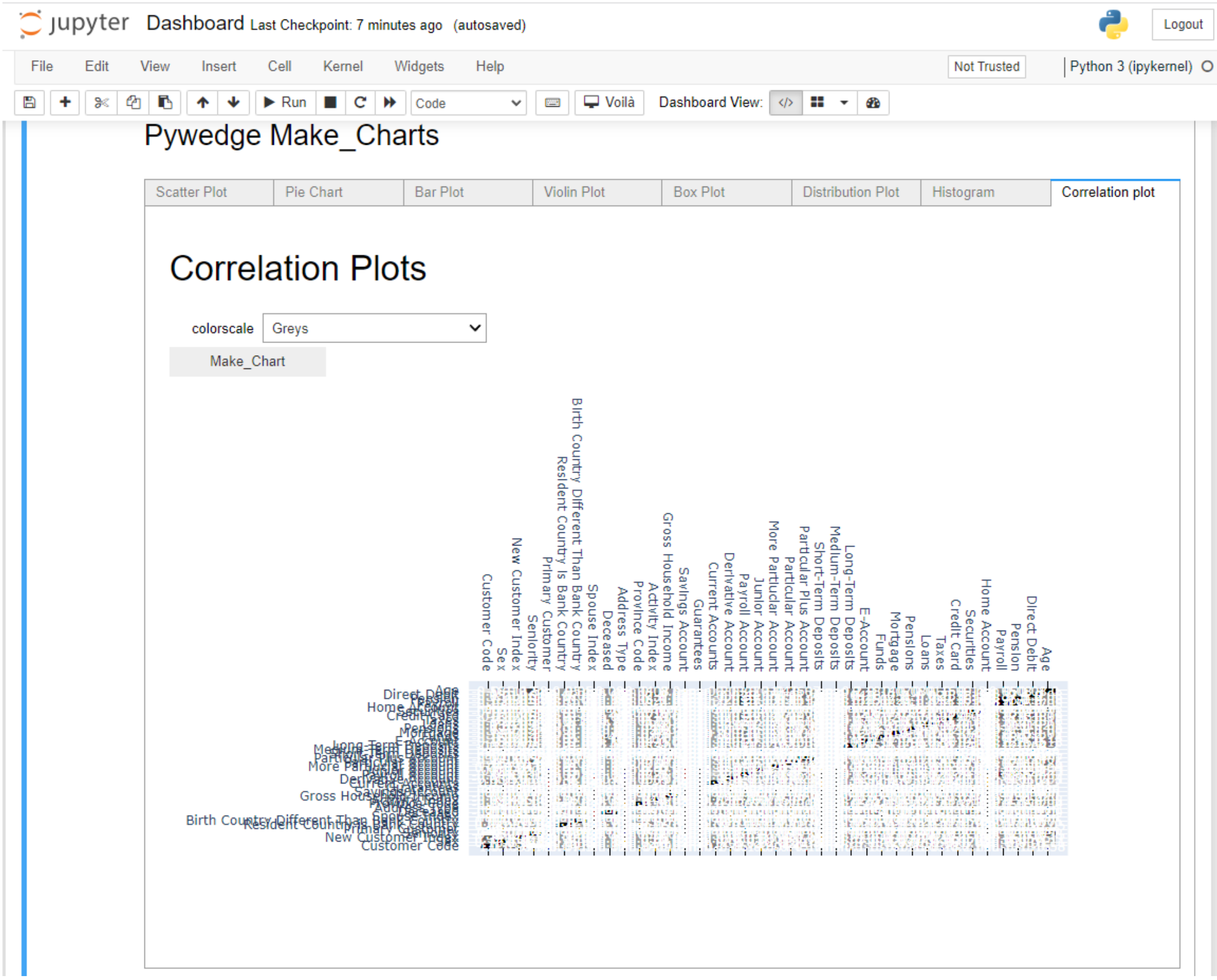

jupyter Dashboard Last Checkpoint: 4 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Violin Plots
X_Axis Age
Y_Axis Seniority
Color Customer Relation Type at Beg
Make_Chart
X_Axis Age
Y_Axis Gross Household Income
Color Customer Relation Type at Beg
Make_Chart
Age vs Seniority
Customer Relation Type at B
A
I
Seniority
Age
Age vs Gross Household Income
Customer Relation Type at B
A
I
Gross Household Income
Age
X_Axis Age
Y_Axis Gross Household Income
Color Country Residence
Make_Chart
X_Axis Age
Y_Axis Gross Household Income
Color Province Name
Make_Chart
Age vs Gross Household Income
Gross Household Income
Age
Age vs Gross Household Income
Province Na
MALAG
BARCE
VALEN
MADRI
ASTUR
BIZKAI
TOLED
SALAM
ALMER
CANTA
PONTE
SEVILL
CASTE
CORUÑ
CORD
SEGOV
BADAJ
PALMA
BALEA
SANTA
VALLA
GRANA
JAEN
Gross Household Income
Age


jupyter Dashboard Last Checkpoint: 5 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Box Plots
X_Axis Customer Relation Type at Beg
Y_Axis Age
Color Country Residence
Make_Chart
X_Axis Country Residence
Y_Axis Gross Household Income
Color Customer Relation Type at Beg
Make_Chart
Customer Relation Type At Beginning Of Month vs
Age
Customer Relation Type At Beginning Of Month
Country Residence vs Gross Household Income
Customer Relation Type at B
Gross Household Income
Country Residence
X_Axis Customer Relation Type at Beg
Y_Axis Gross Household Income
Color Country Residence
Make_Chart
X_Axis Country Residence
Y_Axis Seniority
Color Customer Relation Type at Beg
Make_Chart
Customer Relation Type At Beginning Of Month vs
Gross Household Income
Customer Relation Type At Beginning Of Month
Country Residence vs Seniority
Customer Relation Type at B
Seniority
Country Residence


jupyter
Dashboard Last Checkpoint: 6 minutes ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Distribution Plots
X_Axis
Age
Y_Axis
Gross Household Income
Color
Customer Relation Type at Beg
Make_Chart
X_Axis
Seniority
Y_Axis
Gross Household Income
Color
Country Residence
Make_Chart
Age vs Gross Household Income
Customer Relation Type at B
A
I
sum of Gross Household Income
Age
Seniority vs Gross Household Income
sum of Gross Household Income
Seniority
X_Axis
Age
Y_Axis
Seniority
Color
Customer Relation Type at Beg
Make_Chart
X_Axis
Sex
Y_Axis
Age
Color
Country Residence
Make_Chart
Age vs Seniority
Customer Relation Type at B
A
I
sum of Seniority
Age
Sex vs Age
sum of Age
Sex



In [38]:
```
# In case the graph doesn't save, run the Dashboard setup code, run the Dashboard of Graphs code, and input these settings:
#
# Scatterplot: X_Axis: Age                    X_Axis: Age
#              Y_Axis: Seniority              Y_Axis: Gross Household Income
#              Color: Activity Index          Color: E-Account
#
#              X_Axis: Age                    X_Axis: Age
```


jupyter
Dashboard Last Checkpoint: 7 minutes ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Voilà
Dashboard View:
Pywedge Make_Charts
Scatter Plot
Pie Chart
Bar Plot
Violin Plot
Box Plot
Distribution Plot
Histogram
Correlation plot
Correlation Plots
colorscale
Greys
Make_Chart
Customer Code
Sex
New Customer Index
Seniority
Primary Customer
Resident Country is Bank Country
Birth Country Different Than Bank Country
Spouse Index
Deceased
Address Type
Province Code
Activity Index
Gross Household Income
Savings Account
Guarantees
Current Accounts
Derivative Account
Payroll Account
Junior Account
More Particular Account
Particular Account
Particular Plus Account
Short-Term Deposits
Medium-Term Deposits
Long-Term Deposits
E-Account
Funds
Mortgage
Pensions
Loans
Taxes
Credit Card
Securities
Home Account
Payroll
Pension
Direct Debit
Age