# Bank Churn Internal Report

**Background**
Pulled public data from https://www.kaggle.com/datasets/mathchi/churn-for-bank-customers licensed under CC0: Public Domain.

This is a dataset more often used in machine learning models to predict churn. It's also used for cleaning and analysis entirely within Tableau.

**Initial Preparation**
Even though the dataset is frequently cleaned and analyzed entirely with Tableau, I still loaded it into a dataframe in Python for basic cleaning and analysis.

**Loading, Validating and Prepping the Data**
Column names and data types look good. Nothing unusual in the basic descriptive statitistics for the numeric variables.

There are no N/A values and no duplicated rows.

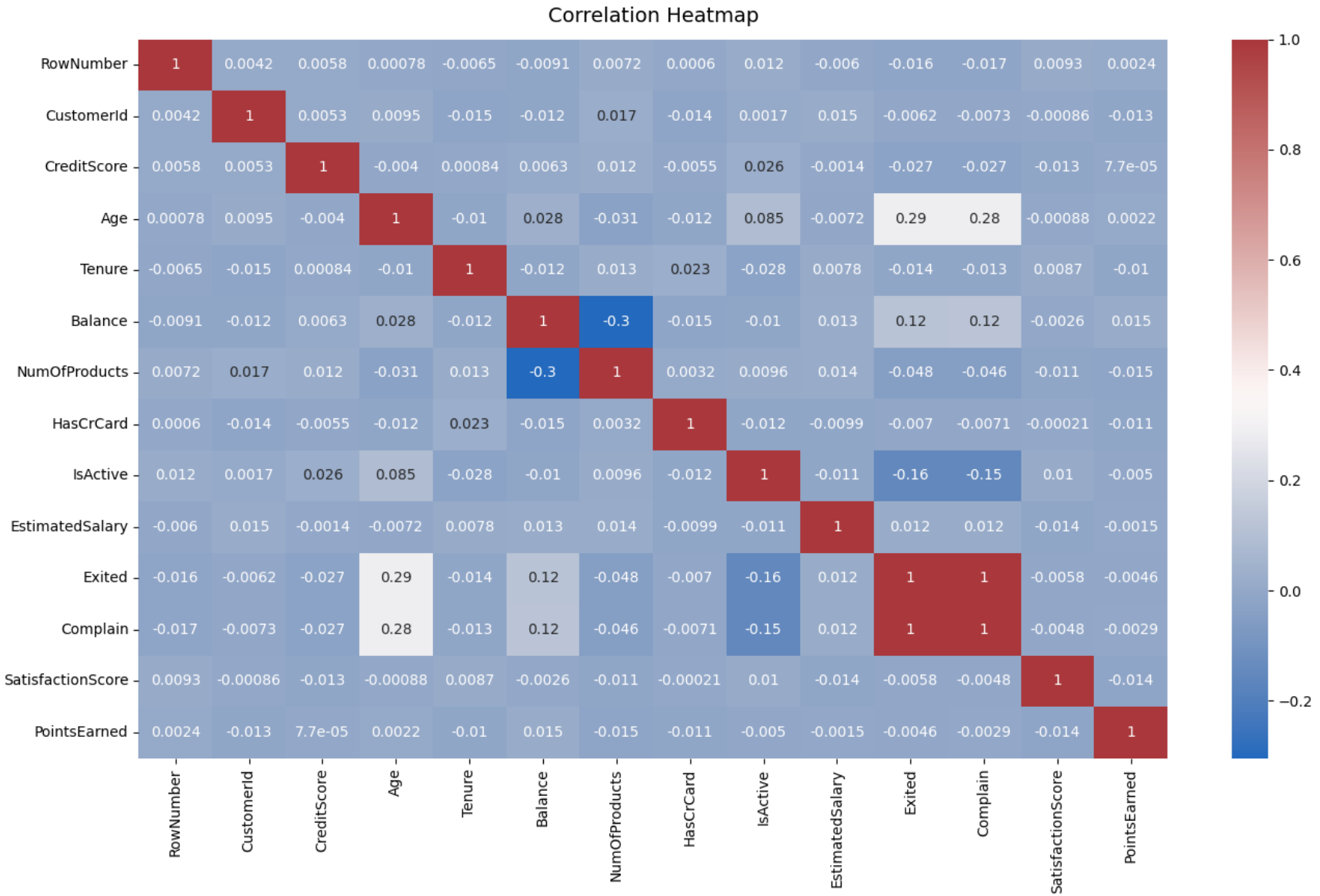
There are a few outliers for credit score and for age that we'll leave in as we need to see all the data for this analysis and they do not appear to skew the results.

**Verifying The Cleaned Data**
As the basic statistics look good, we'll continue with further analysis and visualization in Tableau.

One thing that stands out in the correlation heatmap is a 100% or 1:1 correlation between complained and exited. So, we know there's a serious customer service problem. If the customer complained, they left. We'll look at churn rates across several variables, but we know that we need emphasize the relationjship between complained and exited.

I saved the clean data to a CSV file for further analysis and visualization in Tableau.

Please see page 2 for the correlation heatmap.


Correlation Heatmap
RowNumber 1 0.0042 0.0058 0.00078 -0.0065 -0.0091 0.0072 0.0006 0.012 -0.006 -0.016 -0.017 0.0093 0.0024
CustomerId 0.0042 1 0.0053 0.0095 -0.015 -0.012 0.017 -0.014 0.0017 0.015 -0.0062 -0.0073 -0.00086 -0.013
CreditScore 0.0058 0.0053 1 -0.004 0.00084 0.0063 0.012 -0.0055 0.026 -0.0014 -0.027 -0.027 -0.013 7.7e-05
Age 0.00078 0.0095 -0.004 1 -0.01 0.028 -0.031 -0.012 0.085 -0.0072 0.29 0.28 -0.00088 0.0022
Tenure -0.0065 -0.015 0.00084 -0.01 1 -0.012 0.013 0.023 -0.028 0.0078 -0.014 -0.013 0.0087 -0.01
Balance -0.0091 -0.012 0.0063 0.028 -0.012 1 -0.3 -0.015 -0.01 0.013 0.12 0.12 -0.0026 0.015
NumOfProducts 0.0072 0.017 0.012 -0.031 0.013 -0.3 1 0.0032 0.0096 0.014 -0.048 -0.046 -0.011 -0.015
HasCrCard 0.0006 -0.014 -0.0055 -0.012 0.023 -0.015 0.0032 1 -0.012 -0.0099 -0.007 -0.0071 -0.00021 -0.011
IsActive 0.012 0.0017 0.026 0.085 -0.028 -0.01 0.0096 -0.012 1 -0.011 -0.16 -0.15 0.01 -0.005
EstimatedSalary -0.006 0.015 -0.0014 -0.0072 0.0078 0.013 0.014 -0.0099 -0.011 1 0.012 0.012 -0.014 -0.0015
Exited -0.016 -0.0062 -0.027 0.29 -0.014 0.12 -0.048 -0.007 -0.16 0.012 1 1 -0.0058 -0.0046
Complain -0.017 -0.0073 -0.027 0.28 -0.013 0.12 -0.046 -0.0071 -0.15 0.012 1 1 -0.0048 -0.0029
SatisfactionScore 0.0093 -0.00086 -0.013 -0.00088 0.0087 -0.0026 -0.011 -0.00021 0.01 -0.014 -0.0058 -0.0048 1 -0.014
PointsEarned 0.0024 -0.013 7.7e-05 0.0022 -0.01 0.015 -0.015 -0.011 -0.005 -0.0015 -0.0046 -0.0029 -0.014 1
RowNumber
CustomerId
CreditScore
Age
Tenure
Balance
NumOfProducts
HasCrCard
IsActive
EstimatedSalary
Exited
Complain
SatisfactionScore
PointsEarned
1.0
0.8
0.6
0.4
0.2
0.0
−0.2