

# Data Pipeline Report

File: churn.csv | Timestamp: 2025-08-24T04:45:16.222443

## 1. Data Validation Summary

Total Rows	10000
Total Columns	14
Missing Values Found	0
Duplicate Rows Found	0

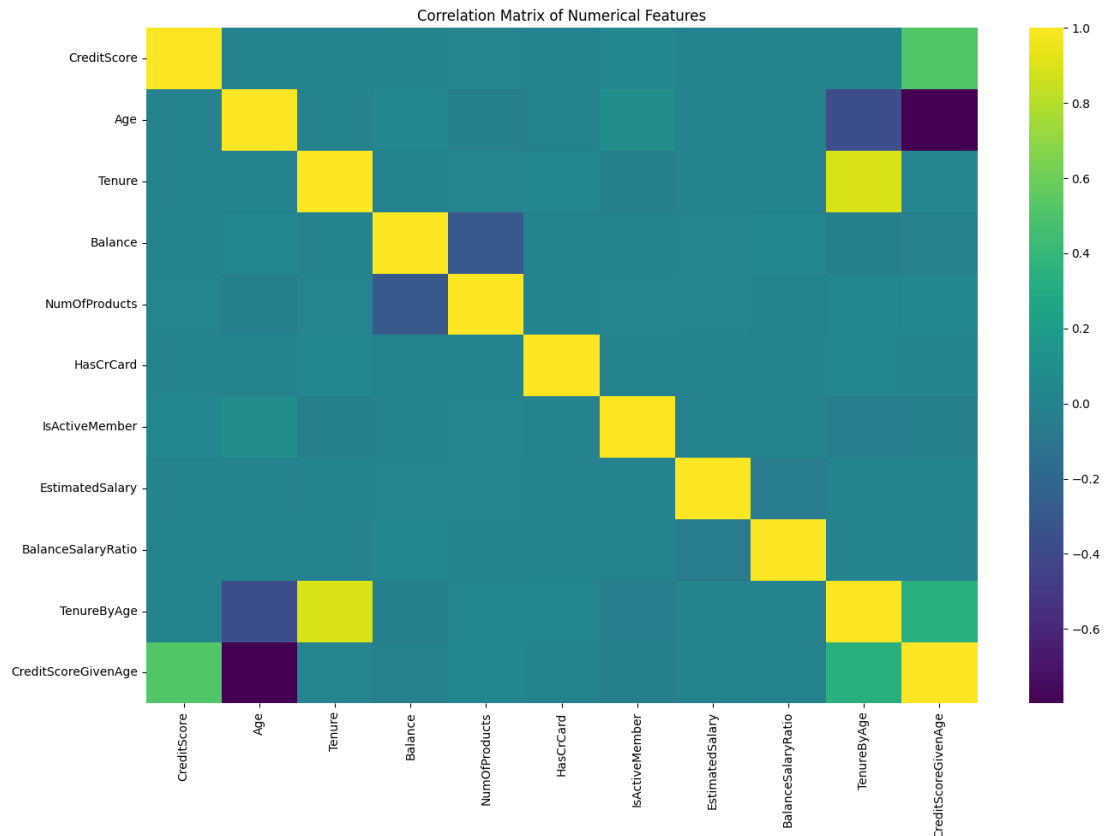
## 2. Feature Engineering Summary

**Columns Dropped:** RowNumber, CustomerId, Surname

**Columns Created:** BalanceSalaryRatio, TenureByAge, CreditScoreGivenAge

## 3. Exploratory Data Analysis

*Correlation Heatmap*



## 4. Final Processed Data

Shape of Final DataFrame: (10000, 15)

**Processed DataFrame Head:**

CreditScore	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember
-0.33	0.29	-1.04	-1.23	-0.91	0.65	0.97
-0.44	0.2	-1.39	0.12	-0.91	-1.55	0.97
-1.54	0.29	1.03	1.33	2.53	0.65	-1.03
0.5	0.01	-1.39	-1.23	0.81	-1.55	-1.03
2.06	0.39	-1.04	0.79	-0.91	0.65	0.97