Data Pipeline Report

File: churn.csv | **Timestamp:** 2025-08-23T15:00:19.177090

1. Data Validation Summary

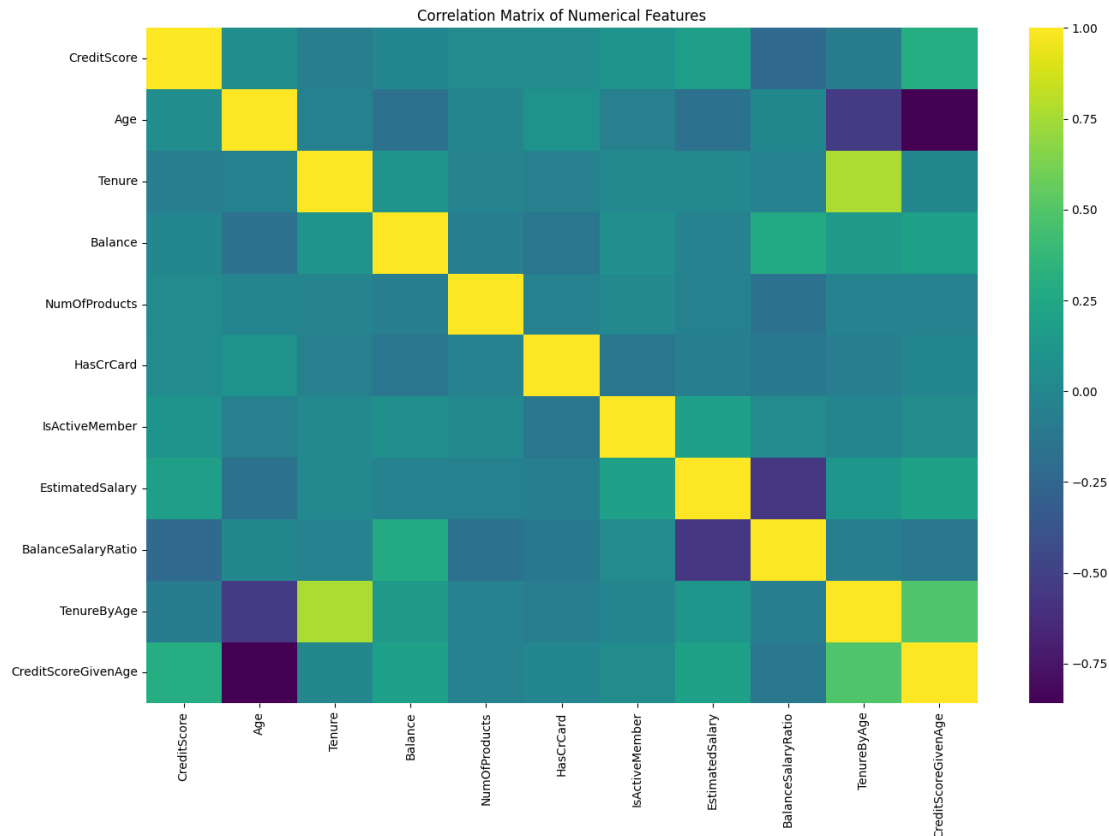|	
Total Rows	100
Total Columns	14
Missing Values Found	0
Duplicate Rows Found	0

2. Feature Engineering Summary

Columns Dropped: RowNumber, CustomerId, Surname
Columns Created: BalanceSalaryRatio, TenureByAge, CreditScoreGivenAge

3. Exploratory Data Analysis

Correlation Heatmap

Correlation Matrix of Numerical Features

4. Final Processed Data

Shape of Final DataFrame: (100, 15)

Processed DataFrame Head:

CreditScore	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember
0.57	-0.09	1.31	-0.14	0.04	0.9	0.9
-1.41	-1.16	0.63	0.43	0.04	0.9	-1.11
-0.22	-1.03	-0.39	-1.56	-1.2	-1.11	0.9
-0.03	-1.25	-1.08	1.66	-1.2	-1.11	0.9
1.12	0.54	-0.39	-0.04	0.04	0.9	0.9