Data Pipeline Report

File: online_retail_customer_churn.csv | **Timestamp:** 2025-08-24T04:30:31.058905

1. Data Validation Summary
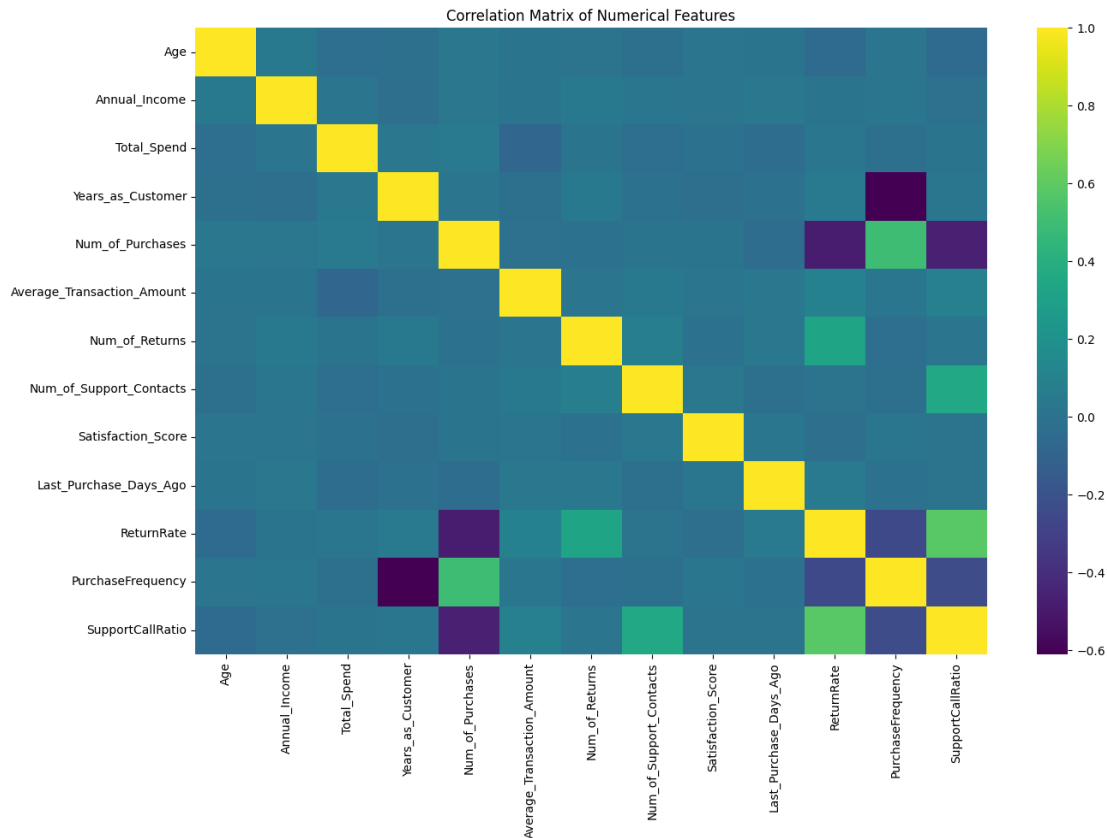
Total Rows	1000
Total Columns	15
Missing Values Found	0
Duplicate Rows Found	0

2. Feature Engineering Summary

Columns Dropped: Customer_ID
Columns Created: ReturnRate, PurchaseFrequency, SupportCallRatio

3. Exploratory Data Analysis

Correlation Heatmap

Correlation Matrix of Numerical Features

4. Final Processed Data

Shape of Final DataFrame: (1000, 19)

Processed DataFrame Head:

Age	Annual_Income	Total_Spend	Years_as_Customer	Num_of_Purchases	Average_Transaction_Amount	Num_of_Returns
1.23	-1.26	0.28	-0.85	-0.96	1.28	-0.9
1.43	-0.61	1.38	0.59	0.97	-1.67	-0.9
-1.66	-1.57	-1.56	0.59	0.76	-1.48	0.13
-1.46	-0.61	1.41	-1.22	-0.58	0.99	0.13
-1.46	-0.65	0.11	0.95	-0.23	-1.14	-0.56