Data Pipeline Report

File: online_retail_customer_churn.csv | **Timestamp:** 2025-08-23T15:06:33.773951

1. Data Validation Summary
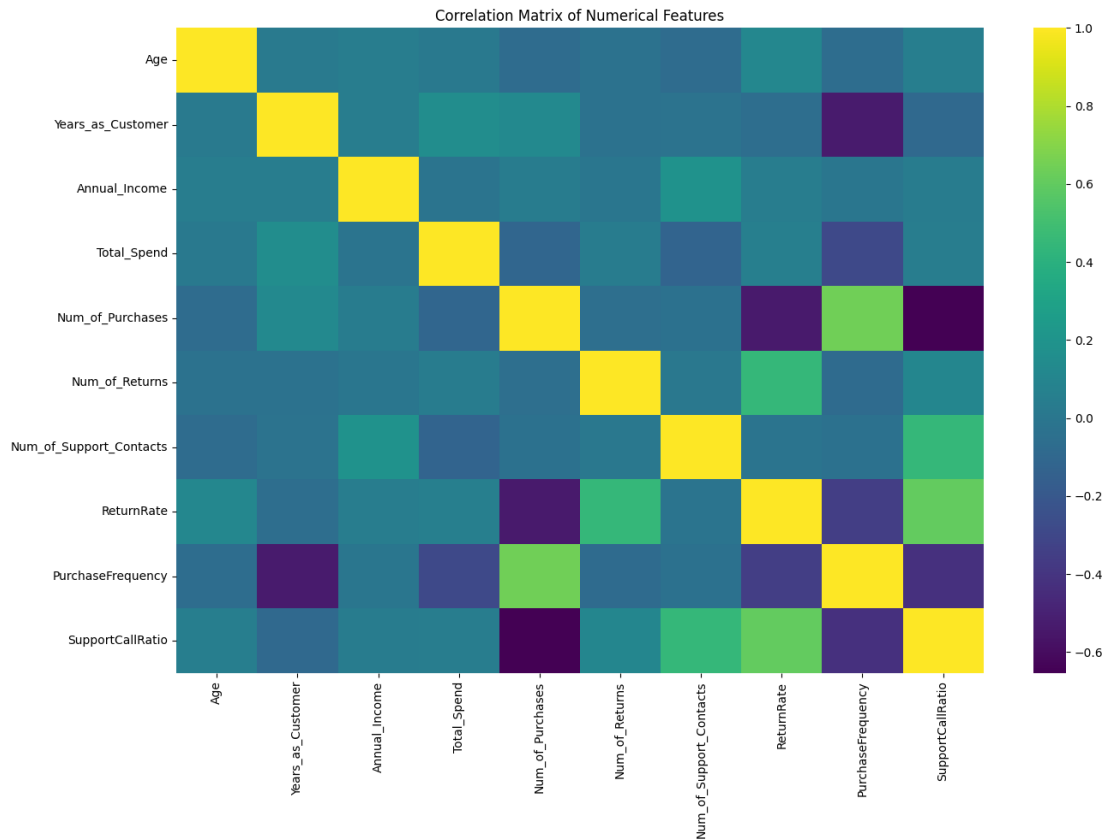
Total Rows	100
Total Columns	9
Missing Values Found	0
Duplicate Rows Found	0

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: (100, 11)

Processed DataFrame Head:

Age	Years_as_Customer	Annual_Income	Total_Spend	Num_of_Purchases	Num_of_Returns	Num_of_Support_Contacts
-1.18	-1.56	1.08	-0.9	0.4	0.81	1.32
0.11	0.76	-0.04	1.48	-1.58	-1.44	-0.19
-1.46	-0.01	-0.29	0.65	-0.33	-1.44	-1.69
-0.25	-1.56	-0.97	-1.47	1.32	1.13	0.56
0.83	-0.4	-0.96	0.87	-1.31	1.45	0.56