Data Pipeline Report

File: online_retail_customer_churn.csv | **Timestamp:** 2025-08-23T14:59:53.385710

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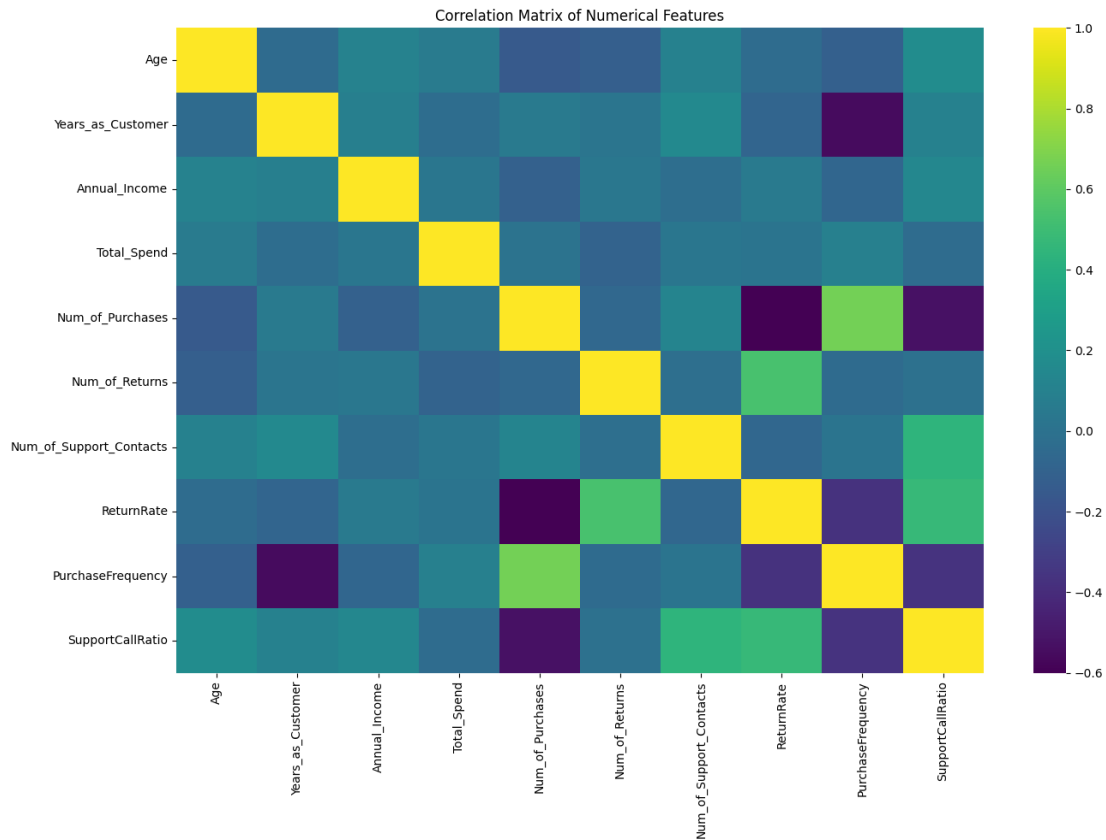
Total Rows	100
Total Columns	9
Missing Values Found	0
Duplicate Rows Found	0

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
-1.08	0.16	1.6	-0.8	-0.49	-1.51	-1.52
0.84	0.16	-0.64	1.78	0.64	-0.85	-0.08
-1.01	1.38	1.34	-1.14	0.79	-0.52	-0.8
1.57	0.57	-0.84	-0.08	0.94	-1.51	-0.08
-0.42	-1.47	-0.46	0.32	-0.04	1.45	-1.52