Data Pipeline Report

File: online_retail_customer_churn.csv | **Timestamp:** 2025-08-23T19:42:24.000362

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
1.38	-0.47	-1.51	-0.09	-0.99	0.67	0.73
-1.57	-0.06	-1.03	-1.72	0.07	-1.55	-0.68
-1.3	-0.06	0.56	-0.62	1.59	1.41	0.73
0.71	0.35	1.53	0.78	-1.74	-1.92	1.43
0.57	-0.47	-0.9	0.31	-0.84	0.3	0.02