

Data Pipeline Report

File: online_retail_customer_churn.csv | Timestamp: 2025-08-23T14:54:33.766341

1. Data Validation Summary

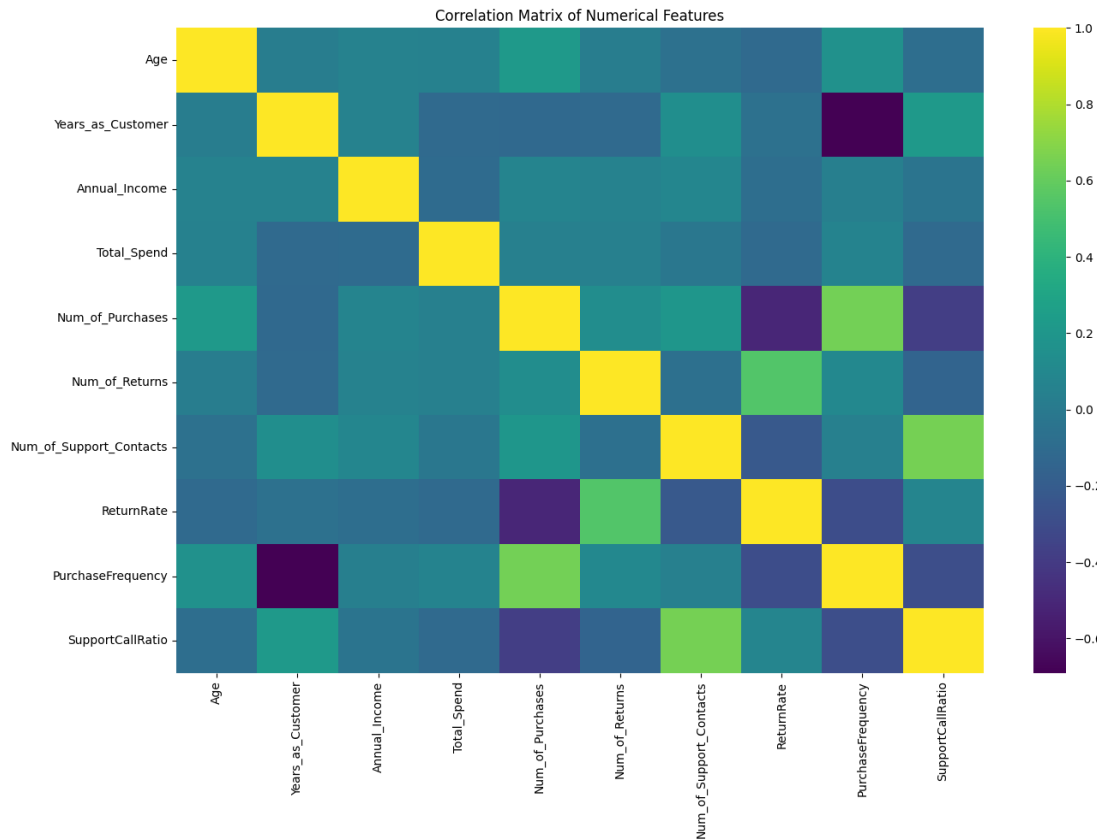
Total Rows	100
Total Columns	9
Missing Values Found	0
Duplicate Rows Found	0

2. Feature Engineering Summary

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

3. Exploratory Data Analysis

Correlation Heatmap



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
1.25	0.96	1.66	-0.42	0.09	-0.83	-1.37
-0.8	0.6	-1.28	-1.64	-0.22	1.59	0.07
-0.55	1.32	-0.84	0.22	-0.52	-0.14	0.79
0.38	0.96	-0.9	-1.15	0.84	-0.14	-0.65
-0.18	0.96	-1.06	0.82	-1.05	-0.49	0.07