

# Data Pipeline Report

**File:** online\_retail\_customer\_churn.csv | **Timestamp:** 2025-08-23T14:39:31.684888

## 1. Data Validation Summary

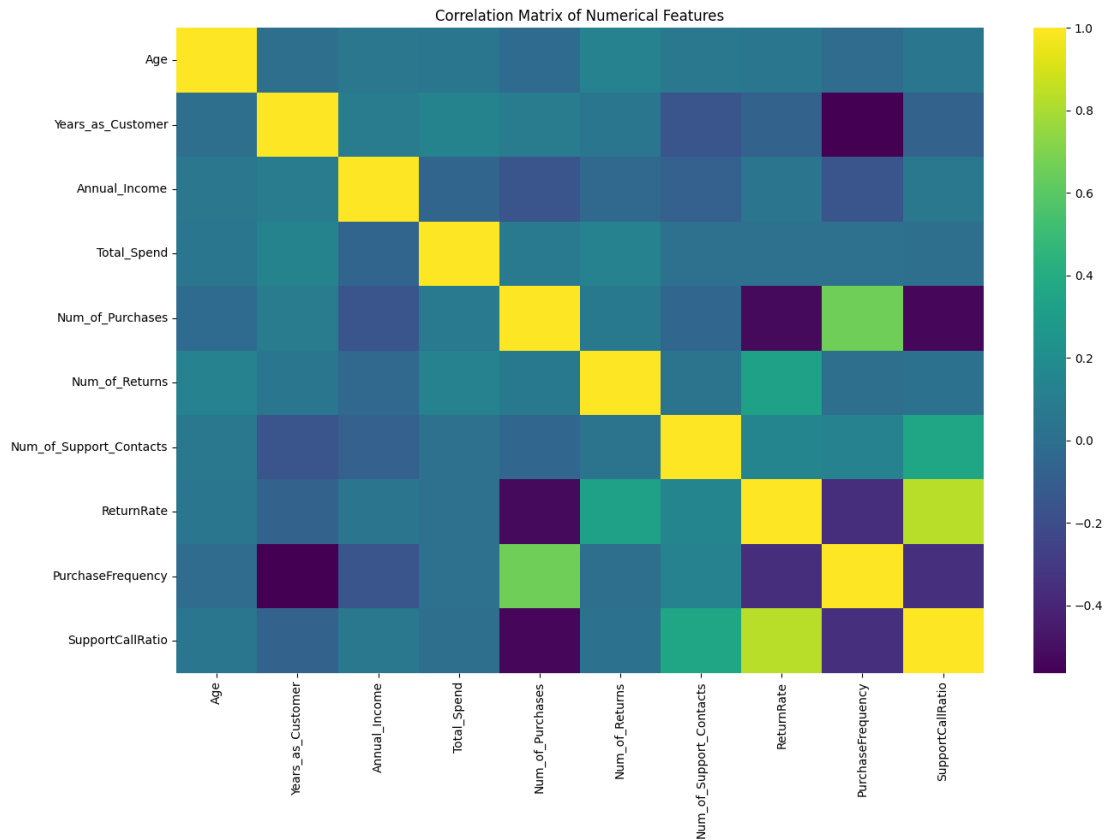
Total Rows	100
Total Columns	9
Missing Values Found	0
Duplicate Rows Found	0

## 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
-0.84	1.12	-0.08	1.27	-1.01	-0.69	1.31
-1.68	1.53	0.92	1.74	1.26	1.53	-0.11
-0.42	1.12	-0.75	1.84	-1.23	0.79	-0.11
1.4	1.12	0.83	0.41	1.33	1.16	-1.53
0.07	1.12	-0.25	0.59	1.48	1.16	0.6