

# Data Pipeline Report

File: online\_retail\_customer\_churn.csv | Timestamp: 2025-08-23T14:45:27.712121

## 1. Data Validation Summary

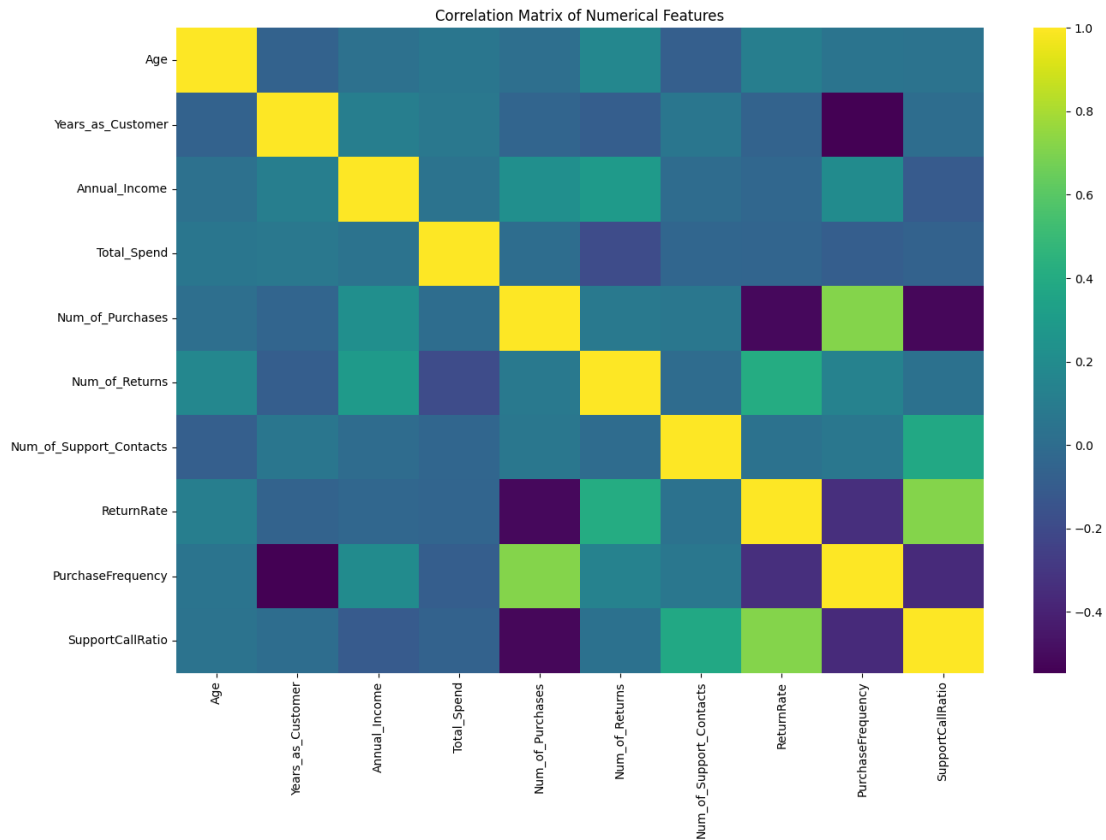
Total Rows	100
Total Columns	9
Missing Values Found	0
Duplicate Rows Found	0

## 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
-0.7	-1.29	0.37	0.42	-0.91	0.95	-0.1
-0.63	-1.7	1.12	1.34	0.12	0.95	1.39
-1.39	-0.47	1.12	-0.5	-0.4	-0.1	1.39
1.1	0.75	0.42	1.82	1.51	-0.8	0.64
-0.9	0.34	-0.66	-1.25	-0.69	-0.8	-1.59