```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns


# Load Data
customers = pd.read_csv("Customers.csv")
products = pd.read_csv("Products.csv")
transactions = pd.read_csv("Transactions.csv")


# Convert date columns to datetime
customers['SignupDate'] = pd.to_datetime(customers['SignupDate'])
transactions['TransactionDate'] = pd.to_datetime(transactions['TransactionDate'])


# Merge datasets for analysis
df = transactions.merge(customers, on="CustomerID").merge(products, on="ProductID")


df.head()
```

	TransactionID	CustomerID	ProductID	TransactionDate	Quantity	TotalValue	Price_x	CustomerName	Region	SignupDate	ProductNa
0	T00001	C0199	P067	2024-08-25 12:38:23	1	300.68	300.68	Andrea Jenkins	Europe	2022-12-03	ComfortLiv Blueto Spea
1	T00112	C0146	P067	2024-05-27 22:23:54	1	300.68	300.68	Brittany Harvey	Asia	2024-09-04	ComfortLiv Blueto Spea
2	T00166	C0127	P067	2024-04-25 07:38:55	1	300.68	300.68	Kathryn Stevens	Europe	2024-04-04	ComfortLiv Blueto Spea
3	T00272	C0087	P067	2024-03-26	2	601.36	300.68	Travis	South	2024-04-11	ComfortLiv Blueto

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.preprocessing import StandardScaler


# Add Customer Profile Info
customer_features = customer_features.merge(customers, on="CustomerID")


# Scaling Features
scaler = StandardScaler()
scaled_features = scaler.fit_transform(customer_features[['TotalValue', 'Quantity', 'Price_x']])


# Similarity Calculation
similarity_matrix = cosine_similarity(scaled_features)
similarity_df = pd.DataFrame(similarity_matrix, index=customer_features['CustomerID'], columns=customer_features['CustomerID'])


# Find Top 3 Lookalikes
lookalikes = {}
for customer in similarity_df.index[:20]:
    similar_customers = similarity_df[customer].sort_values(ascending=False)[1:4]
    lookalikes[customer] = list(similar_customers.items())


# Save Lookalikes as CSV
lookalike_csv = []
for cust_id, sim_list in lookalikes.items():
    for sim_cust_id, score in sim_list:
        lookalike_csv.append({"cust_id": cust_id, "similar_cust_id": sim_cust_id, "score": score})

lookalike_df = pd.DataFrame(lookalike_csv)
lookalike_df.to_csv("Lookalike.csv", index=False)


Start coding or generate with AI.
```