Lookalike Model

It includes:

1. Dataset Processing
2. Similarity calculation.
3. Recommendations System
4. Validation of results.

1. Dataset Processing

1. **Load the datasets**:
    - Read the Customers.csv, Transactions.csv, and Products.csv files into dataframes.
2. **Merge data**:
    - Combine customer profile data (Customers.csv) with transaction data (Transactions.csv) and product data (Products.csv) to create a single dataset.
3. **Feature engineering**:
    - Generate new features based on transaction and product data:

2. Similarity Calculation

4. **Normalize features**:
    - Scale the feature values to a common range (e.g., 0 to 1). This ensures no single feature dominates the similarity calculation.
5. **Cosine similarity**:
    - Uses cosine similarity to measure how closely customers resemble each other based on normalized features.

3. Recommendation System

6. **Generate recommendations for the first 20 customers:**
    - Outputs the results for the first 20 customers (C0001–C0020) into Lookalike.csv.
7. **Reusable Functionality**

- Includes a function recommend lookalikes(customer_id) to dynamically retrieve lookalike recommendations for any customer.

4. Validation of Results

8. **Output results**:

   - Focus on Customer IDs C0001 to C0020 and generate their top 3 lookalikes along with similarity scores.

   - Save the top 3 lookalike recommendations for each of the first 20 customers into a CSV file (Lookalike.csv) for final evaluation and use.

```
c:\Users\Aditya\Zeotap>
c:\Users\Aditya\Zeotap> c: && cd c:\Users\Aditya\Zeotap && cmd /C "c:\Users\Aditya\AppData\Local\Programs\Python\Python313\python.exe c:\Users\Aditya\.vscod
n.debugpy-2024.14.0-win32-x64\bundled\libs\debugpy\adapter/../..\debugpy\launcher 50256 -- C:\Users\Aditya\Zeotap\Lookalike_Model.py "
Example Recommendation: {'customer_id': 'C0001', 'recommendations': [('C0152', np.float64(0.9999891450678953)), ('C0056', np.float64(0.9996620544408745)), (
.9992334447527148))]}
Mean Similarity Score (Overall): 0.8583
Customer C0001 - Avg Recommendation Score: 0.9996
Customer C0002 - Avg Recommendation Score: 0.9978
Customer C0003 - Avg Recommendation Score: 0.9997
Customer C0004 - Avg Recommendation Score: 0.9971
Customer C0005 - Avg Recommendation Score: 0.9999
Customer C0006 - Avg Recommendation Score: 0.9977
Customer C0007 - Avg Recommendation Score: 0.9981
Customer C0008 - Avg Recommendation Score: 0.9970
Customer C0009 - Avg Recommendation Score: 0.9980
Customer C0010 - Avg Recommendation Score: 0.9966
Customer C0011 - Avg Recommendation Score: 0.9999
Customer C0012 - Avg Recommendation Score: 0.9970
Customer C0013 - Avg Recommendation Score: 0.9981
Customer C0014 - Avg Recommendation Score: 1.0000
Customer C0015 - Avg Recommendation Score: 0.9976
Customer C0016 - Avg Recommendation Score: 0.9999
Customer C0017 - Avg Recommendation Score: 0.9989
Customer C0018 - Avg Recommendation Score: 0.9988
Customer C0019 - Avg Recommendation Score: 0.9871
Customer C0020 - Avg Recommendation Score: 1.0000
Recommendations for C0001: [('C0152', np.float64(0.9999891450678953)), ('C0056', np.float64(0.9996620544408745)), ('C0107', np.float64(0.9992334447527148))]
```