```python
#Task 2: Lookalike Model

import pandas as pd
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.preprocessing import StandardScaler

# Load datasets
customers_url = 'https://drive.google.com/uc?id=1bu_--
mo79VdUG9oin4ybfFGRUSXAe-WE'
products_url = 'https://drive.google.com/uc?id=1IKuDizVapw-
hyktwfpoAoaGtHtTNHfd0'
transactions_url = 'https://drive.google.com/uc?id=1saEqdbBB-
vuk2hxoAf4TzDEsykdKlzbF'

customers = pd.read_csv(customers_url)
products = pd.read_csv(products_url)
transactions = pd.read_csv(transactions_url)

# Merge datasets
transactions['TransactionDate'] =
pd.to_datetime(transactions['TransactionDate'])
customers['SignupDate'] = pd.to_datetime(customers['SignupDate'])
merged_data = transactions.merge(customers,
on='CustomerID').merge(products, on='ProductID')

# Feature Engineering
customer_features = merged_data.groupby('CustomerID').agg(
    total_spending=('TotalValue', 'sum'),
    num_transactions=('TransactionID', 'count'),
    avg_transaction_value=('TotalValue', 'mean'),
    num_categories=('Category', 'nunique')
).reset_index()

# Normalize features for similarity calculation
scaler = StandardScaler()
normalized_features = scaler.fit_transform(customer_features.iloc[:,
1:])

# Calculate Cosine Similarity
similarity_matrix = cosine_similarity(normalized_features)
similarity_df = pd.DataFrame(similarity_matrix,
index=customer_features['CustomerID'],
columns=customer_features['CustomerID'])

# Generate Lookalike Recommendations
lookalike_data = {}
for customer_id in customer_features['CustomerID'][:20]:  # Customers
C0001 - C0020
    similar_customers =
similarity_df[customer_id].nlargest(4).iloc[1:]  # Exclude the
```

```python
# customer itself
    lookalike_data[customer_id] = list(zip(similar_customers.index,
similar_customers.values))

# Create Lookalike.csv
lookalike_output = []
for customer_id, similar_list in lookalike_data.items():
    for sim_customer, score in similar_list:
        lookalike_output.append([customer_id, sim_customer, score])

lookalike_df = pd.DataFrame(lookalike_output, columns=['CustomerID',
'SimilarCustomerID', 'SimilarityScore'])
lookalike_df.to_csv('Lookalike.csv', index=False)

# Display the result
print("Lookalike recommendations saved to 'Lookalike.csv'. Here are
the first few rows:")
print(lookalike_df.head())

Lookalike recommendations saved to 'Lookalike.csv'. Here are the first
few rows:
  CustomerID SimilarCustomerID  SimilarityScore
0      C0001             C0086         0.996560
1      C0001             C0189         0.994776
2      C0001             C0055         0.993965
3      C0002             C0199         0.998247
4      C0002             C0010         0.997953
```