## Task 2: Lookalike Model

```
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import NearestNeighbors
```

```
customers = pd.read_csv('/content/Customers.csv')
```

```
products = pd.read_csv('/content/Products.csv')
```

```
transactions = pd.read_csv('/content/Transactions.csv')
```

```
# Data Preparation
customer_transactions = transactions.merge(customers, on='CustomerID')
customer_product_matrix = customer_transactions.pivot_table(
    index='CustomerID', columns='ProductID', values='Quantity', fill_value=0
)
```

```
# Normalize data
scaler = StandardScaler()
customer_product_matrix_scaled = scaler.fit_transform(customer_product_matrix)
```

```
# KNN Model
model = NearestNeighbors(n_neighbors=4, metric='cosine')
model.fit(customer_product_matrix_scaled)
```

```
NearestNeighbors
NearestNeighbors(metric='cosine', n_neighbors=4)
```

```
# Finding lookalikes
lookalikes = {}
for i, cust_id in enumerate(customer_product_matrix.index[:20]):
    distances, indices = model.kneighbors([customer_product_matrix_scaled[i]], n_neighbors=4)
    lookalikes[cust_id] = [(customer_product_matrix.index[indices[0][j]], distances[0][j]) for j in range(1, 4)]
```

```
# Saving lookalikes to CSV
import csv
with open('lookalike_customers.csv', 'w', newline='') as csvfile:
    fieldnames = ['CustomerID', 'Lookalikes']
    writer = csv.writer(csvfile)
    writer.writerow(fieldnames)
    for cust_id, lookalikes_list in lookalikes.items():
        lookalike_str = ','.join([f'{l[0]} ({l[1]:.2f})' for l in lookalikes_list])
        writer.writerow([cust_id, lookalike_str])
```