Lookalike Model

In [7]:
```python
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import davies_bouldin_score
import pandas as pd
url_customer = "https://drive.google.com/uc?export=download&id=1bu_--mo79VdUG9oin4ybfFGRUSXAe-
url_product = "https://drive.google.com/uc?export=download&id=1IKuDizVapw-hyktwfpoAoaGtHtTNHfc
url_transaction = "https://drive.google.com/uc?export=download&id=1saEqdbBB-vuk2hxoAf4TzDEsykc

customers = pd.read_csv(url_customer)
products = pd.read_csv(url_product)
transactions = pd.read_csv(url_transaction)
data = transactions.merge(customers, on="CustomerID").merge(products, on="ProductID")
user_product_matrix = data.pivot_table(index="CustomerID", columns="ProductID", values="TotalV
similarity_matrix = cosine_similarity(user_product_matrix)
similarity_df = pd.DataFrame(similarity_matrix, index=user_product_matrix.index, columns=user_
print(similarity_df.head())

lookalike_results = {}
for cust_id in user_product_matrix.index[:20]:
    similar_customers = similarity_df.loc[cust_id].sort_values(ascending=False).iloc[1:4]
    lookalike_results[cust_id] = list(zip(similar_customers.index, similar_customers.values))
lookalike_df = pd.DataFrame({
    "CustomerID": lookalike_results.keys(),
    "Recommendations": [str(v) for v in lookalike_results.values()]
})
lookalike_df.to_csv("Lookalike_Recommendations.csv", index=False)
print("Lookalike Recommendations:")
display(lookalike_df)


data["AvgPrice"] = data[["Price_x", "Price_y"]].mean(axis=1)
features = data.groupby("CustomerID").agg({
    "TotalValue": "sum",
    "Quantity": "sum",
    "AvgPrice": "mean"
}).reset_index()

print(features.head())

lookalike_df = pd.DataFrame({
    "CustomerID": lookalike_results.keys(),
    "Recommendations": [str(v) for v in lookalike_results.values()]
})
lookalike_df.to_csv("Ganesha_D_Lookalike.csv", index=False)
```

```
CustomerID C0001 C0002     C0003     C0004     C0005 C0006     C0007  \
CustomerID
C0001      1.0   0.0   0.000000  0.000000  0.000000    0.0  0.203038
C0002      0.0   1.0   0.000000  0.000000  0.000000    0.0  0.000000
C0003      0.0   0.0   1.000000  0.139782  0.347737    0.0  0.000000
C0004      0.0   0.0   0.139782  1.000000  0.186362    0.0  0.000000
C0005      0.0   0.0   0.347737  0.186362  1.000000    0.0  0.000000

CustomerID     C0008 C0009     C0010  ...     C0191    C0192 C0193  \
CustomerID                            ...
C0001       0.000000   0.0  0.000000  ...  0.13837  0.000000   0.0
C0002       0.095163   0.0  0.000000  ...  0.00000  0.000000   0.0
C0003       0.004856   0.0  0.000000  ...  0.00000  0.000000   0.0
C0004       0.016953   0.0  0.071485  ...  0.00000  0.000000   0.0
C0005       0.000000   0.0  0.000000  ...  0.00000  0.641821   0.0

CustomerID     C0194 C0195 C0196 C0197 C0198     C0199 C0200
CustomerID
C0001       0.410078   0.0   0.0   0.0   0.0  0.119181   0.0
C0002       0.000000   0.0   0.0   0.0   0.0  0.000000   0.0
C0003       0.000000   0.0   0.0   0.0   0.0  0.000000   0.0
C0004       0.073930   0.0   0.0   0.0   0.0  0.038244   0.0
C0005       0.020658   0.0   0.0   0.0   0.0  0.000000   0.0

[5 rows x 199 columns]
Lookalike Recommendations:
```

	CustomerID	Recommendations
0	C0001	[('C0050', 0.5326650751975152), ('C0100', 0.52...
1	C0002	[('C0109', 0.5734695282308275), ('C0079', 0.52...
2	C0003	[('C0181', 0.6200214220870754), ('C0186', 0.55...
3	C0004	[('C0063', 0.45920773392365705), ('C0070', 0.4...
4	C0005	[('C0096', 0.6683801091411397), ('C0192', 0.64...
5	C0006	[('C0058', 0.6833174088001752), ('C0040', 0.61...
6	C0007	[('C0020', 0.5798921734396695), ('C0031', 0.44...
7	C0008	[('C0165', 0.4818084247771058), ('C0169', 0.42...
8	C0009	[('C0140', 0.7865578016626215), ('C0112', 0.71...
9	C0010	[('C0029', 0.8903528709045399), ('C0083', 0.84...
10	C0011	[('C0135', 0.6044043336963852), ('C0173', 0.42...
11	C0012	[('C0091', 0.7650329732156202), ('C0066', 0.62...
12	C0013	[('C0058', 0.5569830949313298), ('C0040', 0.49...
13	C0014	[('C0128', 0.5740534126137472), ('C0159', 0.44...
14	C0015	[('C0073', 0.5473952661786068), ('C0036', 0.47...
15	C0016	[('C0052', 0.774391222794736), ('C0064', 0.666...
16	C0017	[('C0146', 0.49419227564570306), ('C0147', 0.3...
17	C0018	[('C0059', 0.37252427170007346), ('C0115', 0.3...
18	C0019	[('C0064', 0.639336891964859), ('C0049', 0.566...
19	C0020	[('C0026', 0.6544854567469435), ('C0007', 0.57...

```
  CustomerID  TotalValue  Quantity    AvgPrice
0      C0001     3354.52        12  278.334000
1      C0002     1862.74        10  208.920000
2      C0003     2725.38        14  195.707500
3      C0004     5354.88        23  240.636250
4      C0005     2034.24         7  291.603333
```