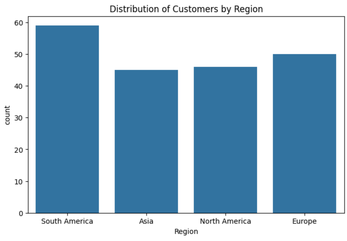
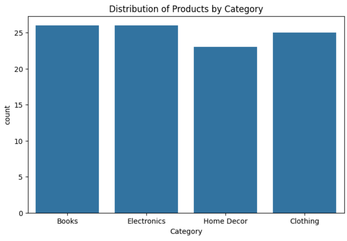
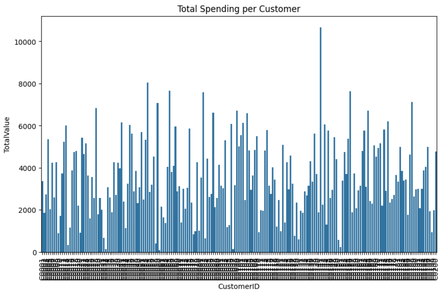
In [7]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns


customers_df = pd.read_csv(r"C:/Users/Administrator/Downloads/Customers.csv")  # Update with your file path
products_df = pd.read_csv(r"C:\Users\Administrator\Downloads\Products.csv")     # Update with your file path
transactions_df = pd.read_csv(r"C:\Users\Administrator\Downloads\Transactions.csv")  # Update with your file path


print("Customers Data:")
print(customers_df.head())

print("\nProducts Data:")
print(products_df.head())

print("\nTransactions Data:")


print("\nCustomers Data Shape:", customers_df.shape)
print("Products Data Shape:", products_df.shape)
print("Transactions Data Shape:", transactions_df.shape)


print("\nMissing values in Customers Data:")
print(customers_df.isnull().sum())

print("\nMissing values in Products Data:")
print(products_df.isnull().sum())

print("\nMissing values in Transactions Data:")
print(transactions_df.isnull().sum())


print(transactions_df.head())
print("\nCustomers Data Summary:")
print(customers_df.describe())

print("\nProducts Data Summary:")
print(products_df.describe())

print("\nTransactions Data Summary:")
print(transactions_df.describe())


plt.figure(figsize=(8, 5))
sns.countplot(x='Region', data=customers_df)
plt.title('Distribution of Customers by Region')
plt.show()


plt.figure(figsize=(8, 5))
sns.countplot(x='Category', data=products_df)
plt.title('Distribution of Products by Category')
plt.show()


transaction_summary = transactions_df.groupby('CustomerID')['TotalValue'].sum().reset_index()
plt.figure(figsize=(10, 6))
sns.barplot(x='CustomerID', y='TotalValue', data=transaction_summary)
plt.title('Total Spending per Customer')
plt.xticks(rotation=90)
plt.show()


print("\nUnique Regions in Customers:")
print(customers_df['Region'].unique())

print("\nUnique Categories in Products:")
print(products_df['Category'].unique())


correlation_matrix = transactions_df[['Price', 'Quantity', 'TotalValue']].corr()
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Correlation Matrix of Price, Quantity, and TotalValue')
plt.show()
```

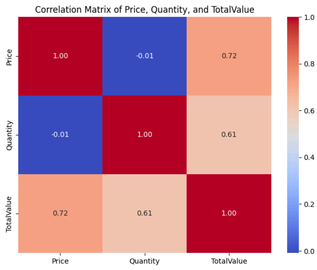
```
Customers Data:
  CustomerID        CustomerName          Region  SignupDate
0      C0001    Lawrence Carroll  South America  2022-07-10
1      C0002      Elizabeth Lutz           Asia  2022-02-13
2      C0003     Michael Rivera  South America  2024-03-07
3      C0004  Kathleen Rodriguez  South America  2022-10-09
4      C0005        Laura Weber           Asia  2022-08-15

Products Data:
  ProductID              ProductName     Category   Price
0      P001     ActiveWear Biography        Books  169.30
1      P002    ActiveWear Smartwatch  Electronics  346.30
2      P003  ComfortLiving Biography        Books   44.12
3      P004       BookWorld Rug   Home Decor   95.69
4      P005          TechPro T-Shirt     Clothing  429.31

Transactions Data:
  TransactionID CustomerID ProductID      TransactionDate  Quantity  \
0        T00001      C0199      P067  2024-08-25 12:38:23         1
1        T00112      C0146      P067  2024-05-27 22:23:54         1
2        T00166      C0127      P067  2024-04-25 07:38:55         1
3        T00272      C0087      P067  2024-03-26 22:55:37         2
4        T00363      C0070      P067  2024-03-21 15:10:10         3

   TotalValue   Price
0      300.68  300.68
1      300.68  300.68
2      300.68  300.68
3      601.36  300.68
4      902.04  300.68

Customers Data Shape: (200, 4)
Products Data Shape: (100, 4)
Transactions Data Shape: (1000, 7)

Missing values in Customers Data:
CustomerID      0
CustomerName    0
Region          0
SignupDate      0
dtype: int64

Missing values in Products Data:
ProductID      0
ProductName    0
Category       0
Price          0
dtype: int64

Missing values in Transactions Data:
TransactionID      0
CustomerID         0
ProductID          0
TransactionDate    0
Quantity           0
TotalValue         0
Price              0
dtype: int64

Customers Data Summary:
       CustomerID      CustomerName          Region  SignupDate
count         200               200             200         200
unique        200               200               4         179
top         C0001  Lawrence Carroll  South America  2022-04-16
freq            1                 1              59           3

Products Data Summary:
            Price
count  100.000000
mean   267.551700
std    143.219383
min     16.080000
25%    147.767500
50%    292.875000
75%    397.090000
max    497.760000

Transactions Data Summary:
          Quantity   TotalValue       Price
count  1000.000000  1000.000000  1000.00000
mean      2.537000   689.995560   272.55407
std       1.117981   493.144478   140.73639
min       1.000000    16.080000    16.08000
25%       2.000000   295.295000   147.95000
50%       3.000000   588.880000   299.93000
75%       4.000000  1011.660000   404.40000
max       4.000000  1991.040000   497.76000
```








```
Unique Regions in Customers:
['South America' 'Asia' 'North America' 'Europe']

Unique Categories in Products:
['Books' 'Electronics' 'Home Decor' 'Clothing']
```


Correlation Matrix of Price, Quantity, and TotalValue
Price
Quantity
TotalValue
1.00
-0.01
0.72
-0.01
1.00
0.61
0.72
0.61
1.00
Price
Quantity
TotalValue
1.0
0.8
0.6
0.4
0.2
0.0

In [ ]: