In [6]:
```
import pandas as pd
import matplotlib.pyplot as plt
import plotly.express as px

url_customer = "https://drive.google.com/uc?export=download&id=1bu_--mo79VdUG9oin4ybf
url_product = "https://drive.google.com/uc?export=download&id=1IKuDizVapw-hyktwfpoAoa
url_transaction = "https://drive.google.com/uc?export=download&id=1saEqdbBB-vuk2hxoAf

customers = pd.read_csv(url_customer)
products = pd.read_csv(url_product)
transactions = pd.read_csv(url_transaction)

print("Customers Dataset:")
display(customers.head())

print("Products Dataset:")
display(products.head())

print("Transactions Dataset:")
display(transactions.head())

data = transactions.merge(customers, on="CustomerID").merge(products, on="ProductID")

print("Merged Dataset:")
display(data.head())

print(f"Dataset Shape: {data.shape}")
print("Columns:", data.columns)

print("Missing Values:")
print(data.isnull().sum())
print("Basic Statistics:")
print(data.describe())
```

Customers Dataset:

	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 Dataset:

	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 Dataset:

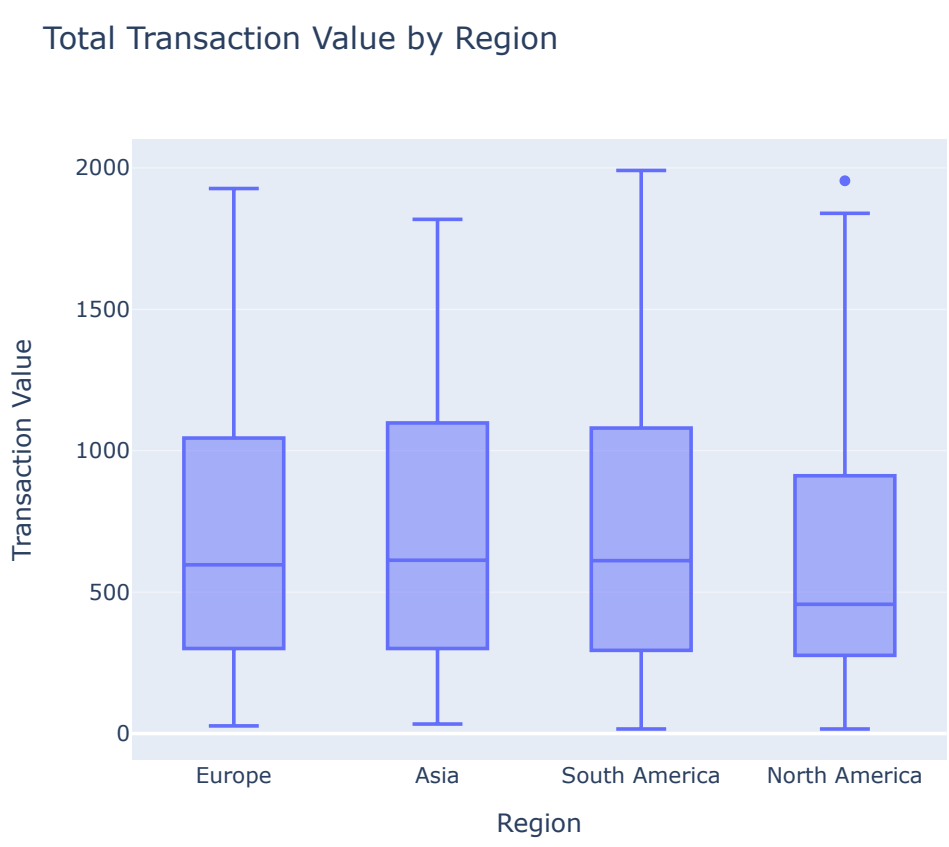
	TransactionID	CustomerID	ProductID	TransactionDate	Quantity	TotalValue	Price
0	T00001	C0199	P067	2024-08-25 12:38:23	1	300.68	300.68
1	T00112	C0146	P067	2024-05-27 22:23:54	1	300.68	300.68
2	T00166	C0127	P067	2024-04-25 07:38:55	1	300.68	300.68
3	T00272	C0087	P067	2024-03-26 22:55:37	2	601.36	300.68
4	T00363	C0070	P067	2024-03-21 15:10:10	3	902.04	300.68

Merged Dataset:

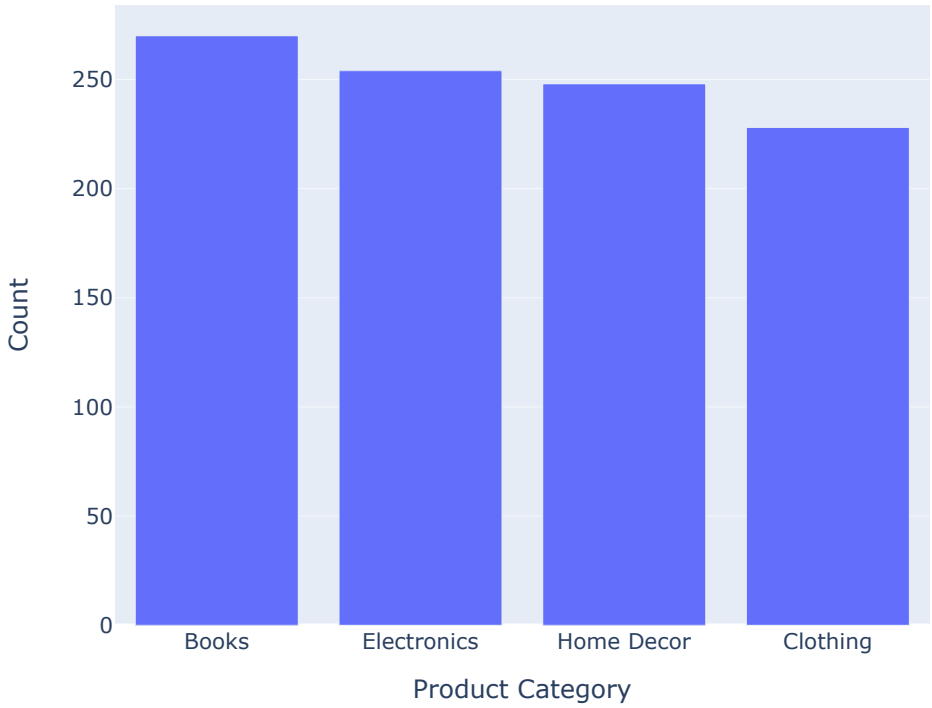
	TransactionID	CustomerID	ProductID	TransactionDate	Quantity	TotalValue	Price_x	CustomerName
0	T00001	C0199	P067	2024-08-25 12:38:23	1	300.68	300.68	Andrea Jenkins
1	T00112	C0146	P067	2024-05-27 22:23:54	1	300.68	300.68	Brittany Harvey
2	T00166	C0127	P067	2024-04-25 07:38:55	1	300.68	300.68	Kathryn Stevens
3	T00272	C0087	P067	2024-03-26 22:55:37	2	601.36	300.68	Travis Campbell
4	T00363	C0070	P067	2024-03-21 15:10:10	3	902.04	300.68	Timothy Perez

```
Dataset Shape: (1000, 13)
Columns: Index(['TransactionID', 'CustomerID', 'ProductID', 'TransactionDate',
       'Quantity', 'TotalValue', 'Price_x', 'CustomerName', 'Region',
       'SignupDate', 'ProductName', 'Category', 'Price_y'],
      dtype='object')
Missing Values:
TransactionID      0
CustomerID         0
ProductID          0
TransactionDate    0
Quantity           0
TotalValue         0
Price_x            0
CustomerName       0
Region             0
SignupDate         0
ProductName        0
Category           0
Price_y            0
dtype: int64
Basic Statistics:
          Quantity   TotalValue     Price_x     Price_y
count  1000.000000  1000.000000  1000.00000  1000.00000
mean      2.537000   689.995560   272.55407   272.55407
std       1.117981   493.144478   140.73639   140.73639
min       1.000000    16.080000    16.08000    16.08000
25%       2.000000   295.295000   147.95000   147.95000
50%       3.000000   588.880000   299.93000   299.93000
75%       4.000000  1011.660000   404.40000   404.40000
max       4.000000  1991.040000   497.76000   497.76000
```

In [5]:
```
data["SignupDate"] = pd.to_datetime(data["SignupDate"])

fig1 = px.box(data, x="Region", y="TotalValue",
              title="Total Transaction Value by Region",
              labels={"Region": "Region", "TotalValue": "Transaction Value"})
fig1.show()
category_count = data["Category"].value_counts().reset_index()
category_count.columns = ["Category", "Count"]

fig2 = px.bar(category_count, x="Category", y="Count",
              title="Product Categories Count",
              labels={"Category": "Product Category", "Count": "Count"})
fig2.show()
signup_trends = data.groupby(data["SignupDate"].dt.to_period("M")).size().reset_index
signup_trends.columns = ["SignupMonth", "Count"]

signup_trends["SignupMonth"] = signup_trends["SignupMonth"].astype(str)

fig3 = px.line(signup_trends, x="SignupMonth", y="Count",
               title="Customer Signup Trends Over Time",
               labels={"SignupMonth": "Signup Month", "Count": "Number of Signups"})
fig3.show()
```



Product Categories Count



Customer Signup Trends Over Time