In [ ]:
```
#Vidhi Rane C-11
```

In [15]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [5]:
```
df = pd.read_csv("C:\TE sem 7\data modeling and visualization\customer_shoppir
df.head()
```

```
<>:1: SyntaxWarning: invalid escape sequence '\T'
<>:1: SyntaxWarning: invalid escape sequence '\T'
C:\Users\vidhi\AppData\Local\Temp\ipykernel_29112\3181797168.py:1: SyntaxWarnin
g: invalid escape sequence '\T'
  df = pd.read_csv("C:\TE sem 7\data modeling and visualization\customer_shoppi
ng_data.csv")
```

Out[5]:

	invoice_no	customer_id	gender	age	category	quantity	price	payment_
0	I138884	C241288	Female	28	Clothing	5	1500.40	Cre
1	I317333	C111565	Male	21	Shoes	3	1800.51	De
2	I127801	C266599	Male	20	Clothing	1	300.08	
3	I173702	C988172	Female	66	Shoes	5	3000.85	Cre
4	I337046	C189076	Female	53	Books	4	60.60	

In [7]:
```
df['sales_amount'] = df['quantity'] * df['price']
```

In [11]:
```
# Step 4: Group by Shopping Mall and calculate total sales
mall_sales = df.groupby('shopping_mall')['sales_amount'].sum().reset_index().s
print(mall_sales)
```

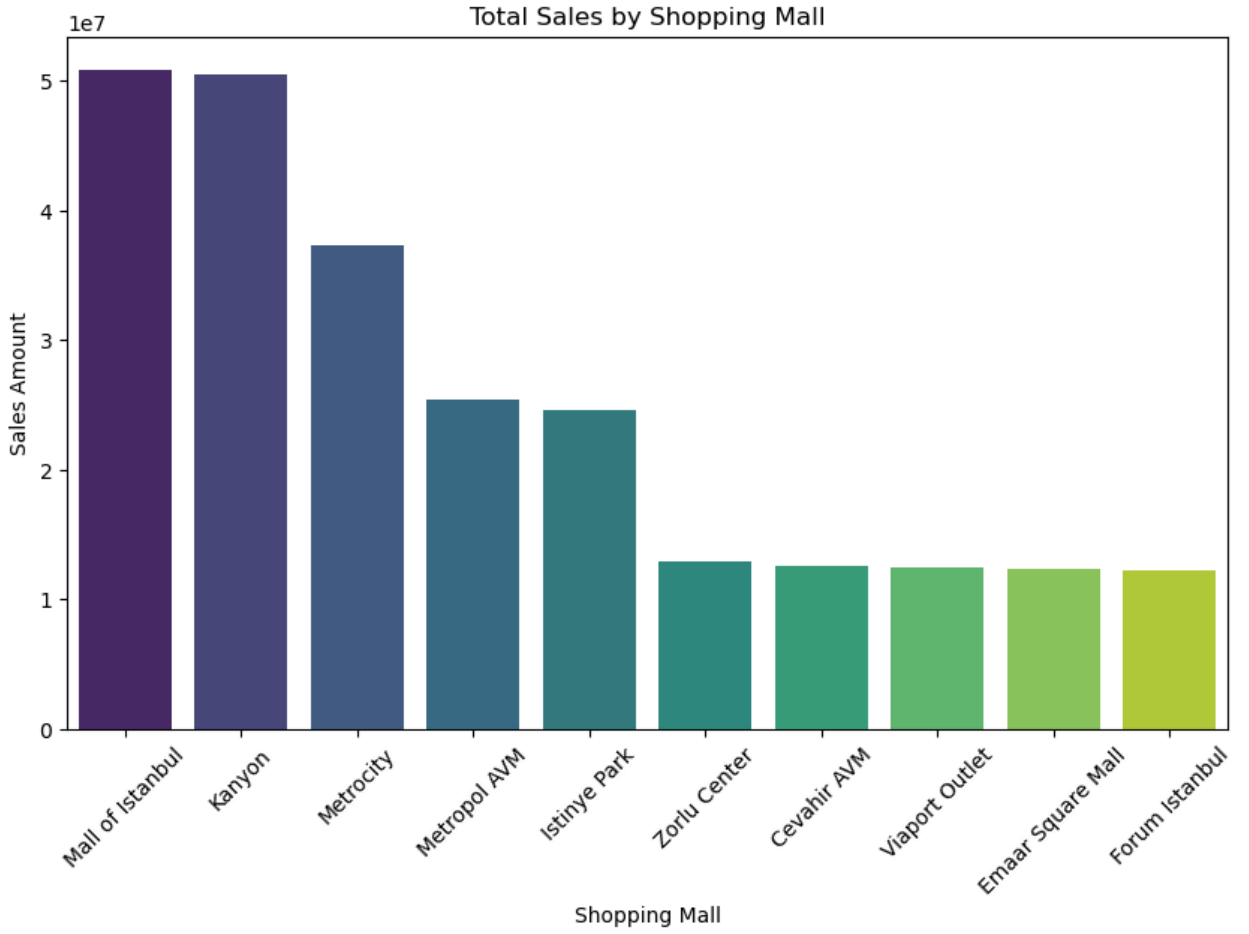
```
      shopping_mall  sales_amount
5   Mall of Istanbul   50872481.68
4            Kanyon   50554231.10
6          Metrocity   37302787.33
7       Metropol AVM   25379913.19
3       Istinye Park   24618827.68
9       Zorlu Center   12901053.82
0        Cevahir AVM   12645138.20
8     Viaport Outlet   12521339.72
1  Emaar Square Mall   12406100.29
2     Forum Istanbul   12303921.24
```

In [17]:
```
# Step 5a: Bar Plot of sales by Shopping Mall
plt.figure(figsize=(10,6))
sns.barplot(x='shopping_mall', y='sales_amount', data=mall_sales, palette='vir
plt.title("Total Sales by Shopping Mall")
plt.xlabel("Shopping Mall")
plt.ylabel("Sales Amount")
plt.xticks(rotation=45)
plt.show()
```

```
C:\Users\vidhi\AppData\Local\Temp\ipykernel_29112\93956800.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in
v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same e
ffect.

  sns.barplot(x='shopping_mall', y='sales_amount', data=mall_sales, palette='vi
ridis')
```



In [19]:
```
# Step 6: Identify top-performing shopping malls
top_malls = mall_sales.head(3)
print("Top Performing Shopping Malls:")
print(top_malls)
```

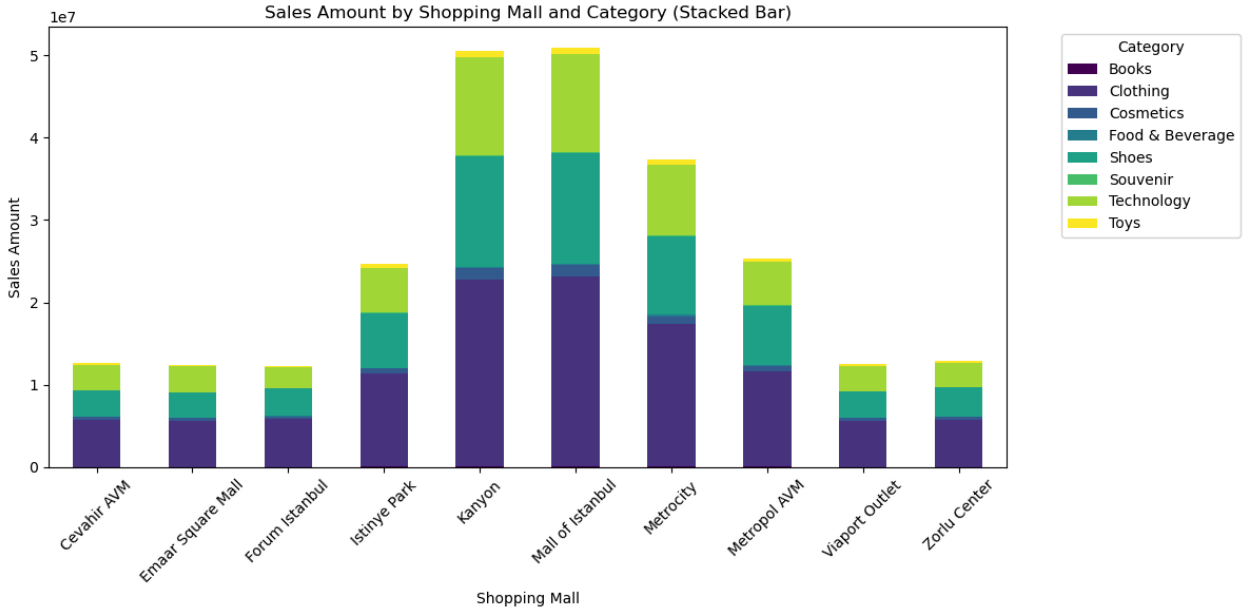
```
Top Performing Shopping Malls:
      shopping_mall  sales_amount
5  Mall of Istanbul   50872481.68
4           Kanyon    50554231.10
6         Metrocity   37302787.33
```

In [21]:
```
# Step 7: Group by Shopping Mall and Product Category
mall_category_sales = df.groupby(['shopping_mall', 'category'])['sales_amount'
print(mall_category_sales)
```

```
   shopping_mall         category  sales_amount
0    Cevahir AVM            Books      44541.00
1    Cevahir AVM         Clothing    5706321.28
2    Cevahir AVM        Cosmetics     321214.00
3    Cevahir AVM  Food & Beverage      44010.45
4    Cevahir AVM            Shoes    3243918.85
..           ...              ...          ...
75  Zorlu Center  Food & Beverage      41955.06
76  Zorlu Center            Shoes    3535601.47
77  Zorlu Center         Souvenir      28996.56
78  Zorlu Center       Technology    2987250.00
79  Zorlu Center             Toys     197550.08

[80 rows x 3 columns]
```

In [27]:
```
# Step 8b: Grouped Bar Plot
mall_category_pivot = mall_category_sales.pivot(index='shopping_mall', columns
mall_category_pivot.plot(kind='bar', stacked=True, figsize=(12,6), colormap='v
plt.title("Sales Amount by Shopping Mall and Category (Stacked Bar)")
plt.ylabel("Sales Amount")
plt.xlabel("Shopping Mall")
plt.xticks(rotation=45)
plt.legend(title='Category', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()
```



In [ ]: