Big Basket Mini Project

Step 1 : Load Dataset

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv("C:/Users/Administrator/Downloads/BigBasket Products.csv")
```

Step 2 : Use head function to look for first 12 rows

In [3]:
```
df.head(12)
```

Out[3]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description
0	1	Garlic Oil - Vegetarian Capsule 500 mg	Beauty & Hygiene	Hair Care	Sri Sri Ayurveda	220.0	220.0	Hair Oil & Serum	4.1	This Product contains Garlic Oil that is known...
1	2	Water Bottle - Orange	Kitchen, Garden & Pets	Storage & Accessories	Mastercook	180.0	180.0	Water & Fridge Bottles	2.3	Each product is microwave safe (without lid), ...
2	3	Brass Angle Deep - Plain, No.2	Cleaning & Household	Pooja Needs	Trm	119.0	250.0	Lamp & Lamp Oil	3.4	A perfect gift for all occasions, be it your m...

3	4	Cereal Flip Lid Container/Storage Jar - Assort...	Cleaning & Household	Bins & Bathroom Ware	Nakoda	149.0	176.0	Laundry, Storage Baskets	3.7	Multipurpose container with an attractive desi...
4	5	Creme Soft Soap - For Hands & Body	Beauty & Hygiene	Bath & Hand Wash	Nivea	162.0	162.0	Bathing Bars & Soaps	4.4	Nivea Creme Soft Soap gives your skin the best...
5	6	Germ - Removal Multipurpose Wipes	Cleaning & Household	All Purpose Cleaners	Nature Protect	169.0	199.0	Disinfectant Spray & Cleaners	3.3	Stay protected from contamination with Multipu...
6	7	Multani Mati	Beauty & Hygiene	Skin Care	Satinance	58.0	58.0	Face Care	3.6	Satinance multani matti is an excellent skin t...
7	8	Hand Sanitizer - 70% Alcohol Base	Beauty & Hygiene	Bath & Hand Wash	Bionova	250.0	250.0	Hand Wash & Sanitizers	4.0	70%Alcohol based is gentle of hand leaves skin...
8	9	Biotin & Collagen Volumizing Hair Shampoo + Bi...	Beauty & Hygiene	Hair Care	StBotanica	1098.0	1098.0	Shampoo & Conditioner	3.5	An exclusive blend with Vitamin B7 Biotin, Hyd...

9	10	Scrub Pad - Anti-Bacterial, Regular	Cleaning & Household	Mops, Brushes & Scrubs	Scotch brite	20.0	20.0	Utensil Scrub-Pad, Glove	4.3	Scotch Brite Anti-Bacterial Scrub Pad thoroug...
10	11	Wheat Grass Powder - Raw	Gourmet & World Food	Cooking & Baking Needs	NUTRASHIL	261.0	290.0	Flours & Pre-Mixes	4.0	Wheatgrass is a superfood potent health food w...
11	12	Butter Cookies Gold Collection	Gourmet & World Food	Chocolates & Biscuits	Sapphire	600.0	600.0	Luxury Chocolates, Gifts	2.2	Enjoy a tin full of delicious butter cookies m...

Step 3 : Get Description of the data in the DataFrame

In [5]:
```
df.describe()
```

Out[5]:

	index	sale_price	market_price	rating
count	27555.00000	27549.000000	27555.000000	18919.000000
mean	13778.00000	334.648391	382.056664	3.943295
std	7954.58767	1202.102113	581.730717	0.739217
min	1.00000	2.450000	3.000000	1.000000
25%	6889.50000	95.000000	100.000000	3.700000
50%	13778.00000	190.320000	220.000000	4.100000

75%	20666.50000	359.000000	425.000000	4.300000
max	27555.00000	112475.000000	12500.000000	5.000000

Step 4 : Find Information about the DataFrame

In [7]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 27555 entries, 0 to 27554
Data columns (total 10 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   index          27555 non-null  int64  
 1   product        27554 non-null  object 
 2   category       27555 non-null  object 
 3   sub_category   27555 non-null  object 
 4   brand          27554 non-null  object 
 5   sale_price     27549 non-null  float64
 6   market_price   27555 non-null  float64
 7   type           27555 non-null  object 
 8   rating         18919 non-null  float64
 9   description    27440 non-null  object 
dtypes: float64(3), int64(1), object(6)
memory usage: 2.1+ MB
```

Step 5 : Find out Top & amp; least sold products

(a). Top Sold Products

In [9]:
```
# Clean product names and group by product
df['product'] = df['product'].str.strip().str.lower()  # Clean spaces and convert to lowercase

# Group by 'product' and sum the 'sale_price' to get top 5 products
top_products = df.groupby('product')['sale_price'].sum().sort_values(ascending=False).head(5)
```

```
# Display the top products
print(top_products)
```

```
product
beard kit                                                 112475.00
4mm aluminium induction base chapati roti tawa - silver   112178.00
balloon - polka dot, 12 inch                               88899.00
extra virgin olive oil                                     24808.53
olive oil - extra virgin                                   22568.22
Name: sale_price, dtype: float64
```

(b). Least Sold Products

In [11]:
```
# Clean product names and group by product
df['product'] = df['product'].str.strip().str.lower()  # Clean spaces and convert to lowercase

# Group by 'product' and sum the 'sale_price' to get bottom 5 least sold products
least_sold_products = df.groupby('product')['sale_price'].sum().sort_values(ascending=True).head(5)

# Display the least sold products
print(least_sold_products)
```

```
product
battery aa 3ut hi top                              0.0
snack mix - dhokla                                 0.0
klassic plain cocktail napkins (22 x 22 cm)        0.0
puja flower wicks - puvvu vathulu batti            0.0
steel fork - medium, premium excel series, bbcl08  0.0
Name: sale_price, dtype: float64
```

Step 6 : Measuring discount on a certain item

In [13]:
```
df['discount_percentage'] = ((df['market_price'] - df['sale_price']) /
df['market_price']) * 100
df['discount_percentage']=df['discount_percentage'].round(2)
df[['product', 'market_price', 'sale_price',
'discount_percentage']].head(10)
```

Out[13]:

	product	market_price	sale_price	discount_percentage

0	garlic oil - vegetarian capsule 500 mg	220.0	220.0	0.00
1	water bottle - orange	180.0	180.0	0.00
2	brass angle deep - plain, no.2	250.0	119.0	52.40
3	cereal flip lid container/storage jar - assort...	176.0	149.0	15.34
4	creme soft soap - for hands & body	162.0	162.0	0.00
5	germ - removal multipurpose wipes	199.0	169.0	15.08
6	multani mati	58.0	58.0	0.00
7	hand sanitizer - 70% alcohol base	250.0	250.0	0.00
8	biotin & collagen volumizing hair shampoo + bi...	1098.0	1098.0	0.00
9	scrub pad - anti- bacterial, regular	20.0	20.0	0.00

Step 7 : Find out the Missing Values from the Dataset

In [15]:
```
df.isnull().sum()
```

Out[15]:
```
index                   0
product                 1
category                0
sub_category            0
brand                   1
sale_price              6
market_price            0
type                    0
rating               8636
description           115
discount_percentage     6
dtype: int64
```

In [17]:
```
df[df['product'].isna()]
```

Out[17]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentage
14363	14364	NaN	Beverages	Coffee	Cothas Coffee	200.0	240.0	Ground Coffee	4.2	Cothas Specialty Blend Coffee and Chicory incl...	16.67

In [19]:
```
df[df['brand'].isna()]
```

Out[19]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentage
9765	9766	food package - medium	Cleaning & Household	Disposables, Garbage Bag	NaN	50.0	50.0	Aluminium Foil, Clingwrap	NaN	NaN	0.0

In [21]:
```
df[df['sale_price'].isna()]
```

Out[21]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_pe
1719	1720	puja flower wicks - puvvu vathulu batti	Cleaning & Household	Pooja Needs	Jaji	NaN	20.0	Camphor & Wicks	NaN	Jaji Puja Flower Batti / Puvvu Vathulu is made...	
1720	1721	powder - sambar	Foodgrains, Oil & Masala	Masalas & Spices	Annapoorna	NaN	46.0	Blended Masalas	NaN	Annaporna Spices are the	

										preserve for Connoiss...
1721	1722	steel fork - medium, premium excel series, bbcl08	Kitchen, Garden & Pets	Crockery & Cutlery	BB Home	NaN	299.0	Cutlery, Spoon & Fork	NaN	BB Home provides fine and classy cutlery that ...
1722	1723	snack mix - dhokla	Snacks & Branded Foods	Ready To Cook & Eat	MTR	NaN	65.0	Breakfast & Snack Mixes	4.1	MTR Dhokla is extremely special in Gujarat. It...
2401	2402	battery aa 3ut hi top	Kitchen, Garden & Pets	Appliances & Electricals	Nippo	NaN	150.0	Battery & Electrical	4.0	This Nippo battery incorporates state-of-the-a...
2402	2403	klassic plain cocktail napkins (22 x 22 cm)	Cleaning & Household	Disposables, Garbage Bag	Origami	NaN	32.0	Toilet Paper	3.9	Klassic Plain Cocktail Napkins (22 X 22 cm) 10...

CLEANING THE MISSING DATA

In [23]:
```
# replacing NaN with "Unknown"
df.loc[df['product'].isna(),'product']='Unknown'
df['product']
```

Out[23]:
```
0                    garlic oil - vegetarian capsule 500 mg
1                                         water bottle - orange
2                               brass angle deep - plain, no.2
3          cereal flip lid container/storage jar - assort...
4                            creme soft soap - for hands & body
                                    ...
27550           wottagirl! perfume spray - heaven, classic
27551                                                rosemary
27552                             peri-peri sweet potato chips
27553                                green tea - pure original
27554                             united dreams go far deodorant
Name: product, Length: 27555, dtype: object
```

In [25]:
```
# replacing NaN with "Unknown"
df.loc[df['brand'].isna(),'brand']='Unknown'
df['brand']
```

Out[25]:
```
0                  Sri Sri Ayurveda
1                         Mastercook
2                                Trm
3                             Nakoda
4                              Nivea
                      ...
27550                         Layerr
27551                       Puramate
27552                         FabBox
27553                         Tetley
27554    United Colors Of Benetton
Name: brand, Length: 27555, dtype: object
```

In [43]:
```
if 'discount_percentage' not in df.columns:
    df['discount_percentage'] = np.nan  # Create the column if missing
```

In [27]:
```
# Replace NaN values in 'sale_price' with the median of 'sale_price'
df['sale_price'] = np.where(df['sale_price'].isna(), df['sale_price'].median(), df['sale_price'])

# Replace NaN values in 'discount_percentage' with the median of 'discount_percentage'
df['discount_percentage'] = np.where(df['discount_percentage'].isna(), df['discount_percentage'].median(),

# Optional: Display the updated columns
```

```
print(df['sale_price'])
print(df['discount_percentage'])
```

```
0         220.00
1         180.00
2         119.00
3         149.00
4         162.00
           ...
27550     199.20
27551      67.50
27552     200.00
27553     396.00
27554     214.53
Name: sale_price, Length: 27555, dtype: float64
0          0.00
1          0.00
2         52.40
3         15.34
4          0.00
          ...
27550     20.00
27551     10.00
27552      0.00
27553     20.00
27554     44.99
Name: discount_percentage, Length: 27555, dtype: float64
```

In [29]:
```
df_c=df.dropna(how='any')
```

In [31]:
```
df_c.isna().sum()
```

Out[31]:
```
index                0
product              0
category             0
sub_category         0
brand                0
sale_price           0
market_price         0
type                 0
rating               0
description          0
```

```
discount_percentage    0
dtype: int64
```

In [33]:
```
df_c.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 18831 entries, 0 to 27554
Data columns (total 11 columns):
 #   Column               Non-Null Count  Dtype  
---  ------               --------------  -----  
 0   index                18831 non-null  int64  
 1   product              18831 non-null  object 
 2   category             18831 non-null  object 
 3   sub_category         18831 non-null  object 
 4   brand                18831 non-null  object 
 5   sale_price           18831 non-null  float64
 6   market_price         18831 non-null  float64
 7   type                 18831 non-null  object 
 8   rating               18831 non-null  float64
 9   description          18831 non-null  object 
 10  discount_percentage  18831 non-null  float64
dtypes: float64(4), int64(1), object(6)
memory usage: 1.7+ MB
```

Step 8 : Find out the outliers from the dataset according to the columns and fill them with the mean

In [17]:
```
mean_price = df['sale_price'].mean()
df['sale_price'] = df['sale_price'].fillna(mean_price)
```

Step 9 : Create Plots or visualizations

9.1) Number of Products per Category :->

A bar plot showing the count of products within each category, providing insight into the distribution of products across various categories.

In [71]:

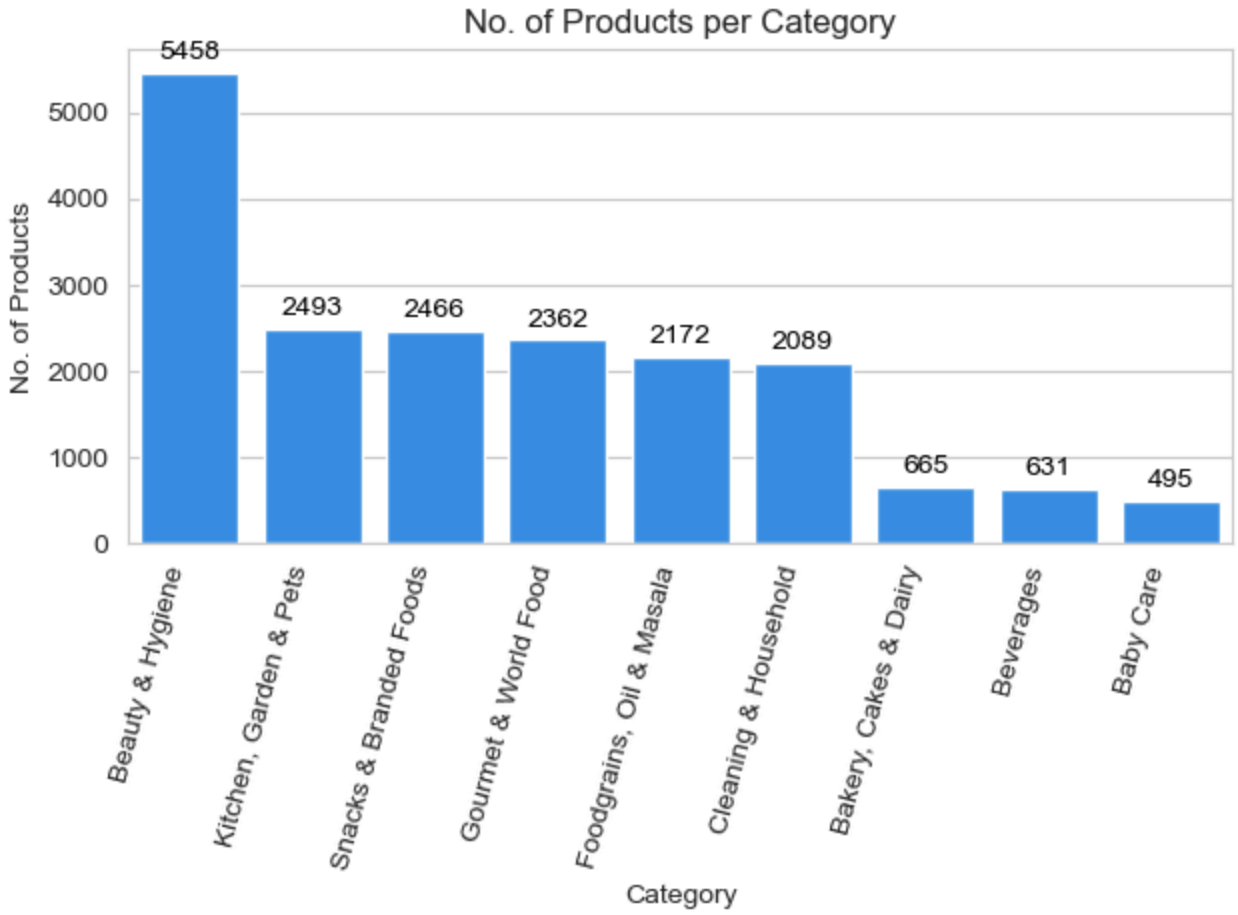
```
counts = df_c['category'].value_counts()
counts_df = pd.DataFrame({'Category': counts.index, 'Counts': counts.values})

ax = sns.barplot(x='Category', y='Counts', data=counts_df, color='dodgerblue')

plt.xticks(rotation=75, ha='right')
ax.set_xlabel('Category')
ax.set_ylabel('No. of Products')
ax.set_title('No. of Products per Category')

for p in ax.patches:
    ax.annotate(f'{int(p.get_height())}',
                (p.get_x() + p.get_width() / 2., p.get_height()),
                ha='center', va='baseline', fontsize=10,
                color='black', xytext=(0, 5),
                textcoords='offset points')

plt.tight_layout()
plt.show()
```



Key Takeaways:

- **Beauty & Hygiene** dominates with the highest number of products (5458).
- **Kitchen, Garden & Pets** follows closely with 2493 products.
- **Snacks & Branded Foods** and **Gourmet & World Food** have a similar product count, around 2400.
- **Foodgrains, Oil & Masala** and **Cleaning & Household** have a moderate number of products, just over 2000.
- **Bakery, Cakes & Dairy** has significantly fewer products (665).
- **Beverages** and **Baby Care** have the lowest product counts, around 600 and 500, respectively.

Recommendations:

- **Focus on High-Performing Categories**:
  Invest in marketing and promotions for Beauty & Hygiene, Kitchen, Garden & Pets, and Snacks & Branded Foods to drive sales.
- **Optimize Low-Performing Categories**:
  Analyze the product mix and customer demand for Bakery, Cakes & Dairy, Beverages, and Baby Care to identify opportunities for improvement.
- **Consider Product Expansion**:
  Explore opportunities to expand the product range in categories with lower product counts to attract a wider customer base.

9.2) Top 5 selling Brands by Volume :->

This bar plot illustrates the five brands with the highest sales volume, highlighting the most popular brands.

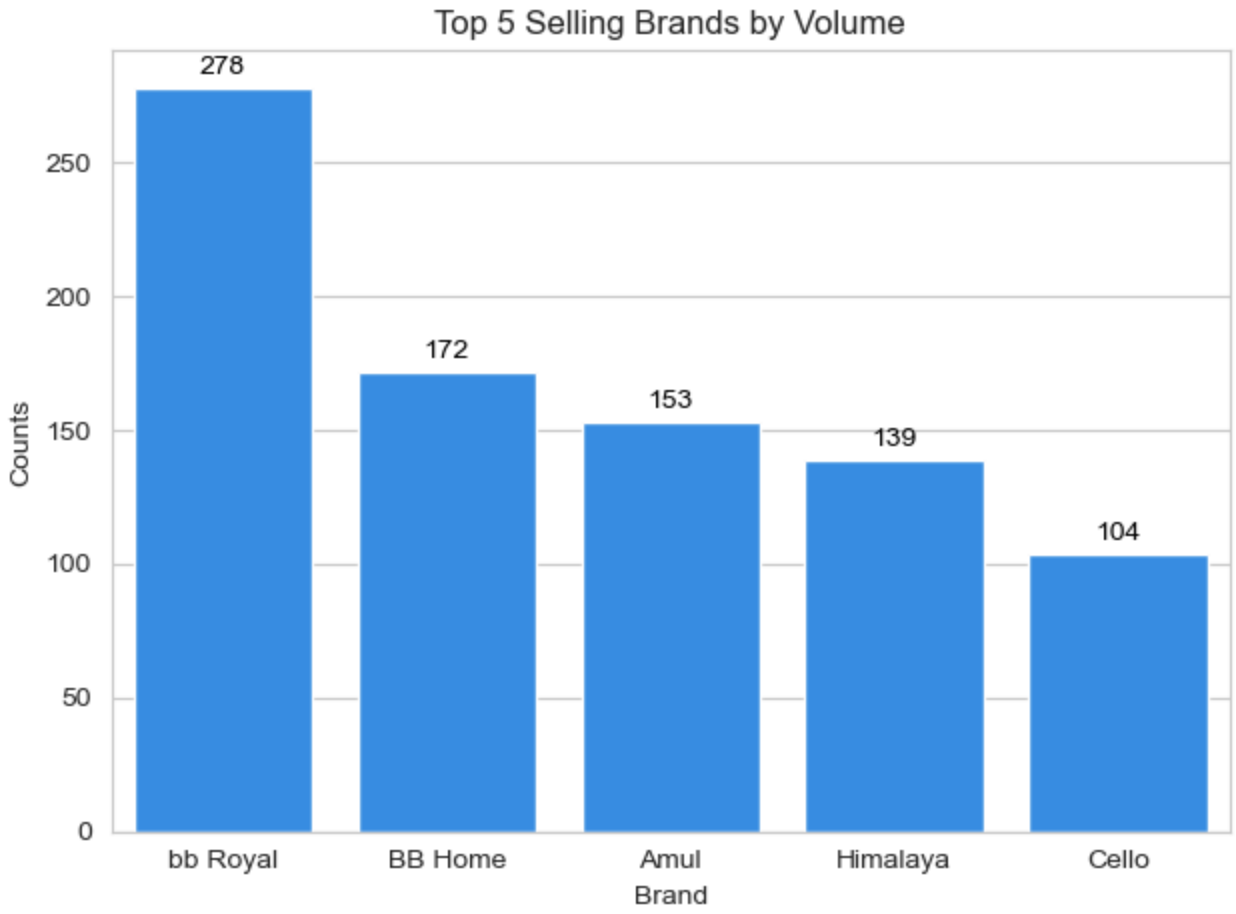
In [101…

```
counts = df_c['brand'].value_counts()
counts_b = pd.DataFrame({'brand': counts.index, 'Counts': counts.values})[:5]

cx = sns.barplot(x='brand', y='Counts', data=counts_b, color='dodgerblue')
cx.set_xlabel('Brand')
cx.set_title('Top 5 Selling Brands by Volume')

for p in cx.patches:
    cx.annotate(f'{int(p.get_height())}',
                (p.get_x() + p.get_width() / 2., p.get_height()),
                ha='center', va='baseline', fontsize=10,
                color='black', xytext=(0, 5), textcoords='offset points')

plt.tight_layout()
plt.show()
```



Key Takeaways:

- **bb Royal** is the top-selling brand, with a volume of 278 units.
- **BB Home** follows closely with 172 units.
- **Amul** and **Himalaya** have a moderate volume, around 150 and 140 units, respectively.
- **Cello** has the lowest volume among the top 5 brands, with 104 units.

Recommendations:

- **Focus on Top-Performing Brands**: Continue to promote and invest in **bb Royal** and **BB Home** to maintain their sales momentum.
- **Optimize Brand Mix**: Analyze the performance of brands with lower volumes (**Cello**, **Himalaya**, **Amul**) to identify opportunities for improvement, such as pricing adjustments or targeted promotions.
- **Consider Brand Expansion**: Explore opportunities to expand the product range of top-performing brands to attract a wider customer base.

9.3) Top 10 products with highest discount percentage :->

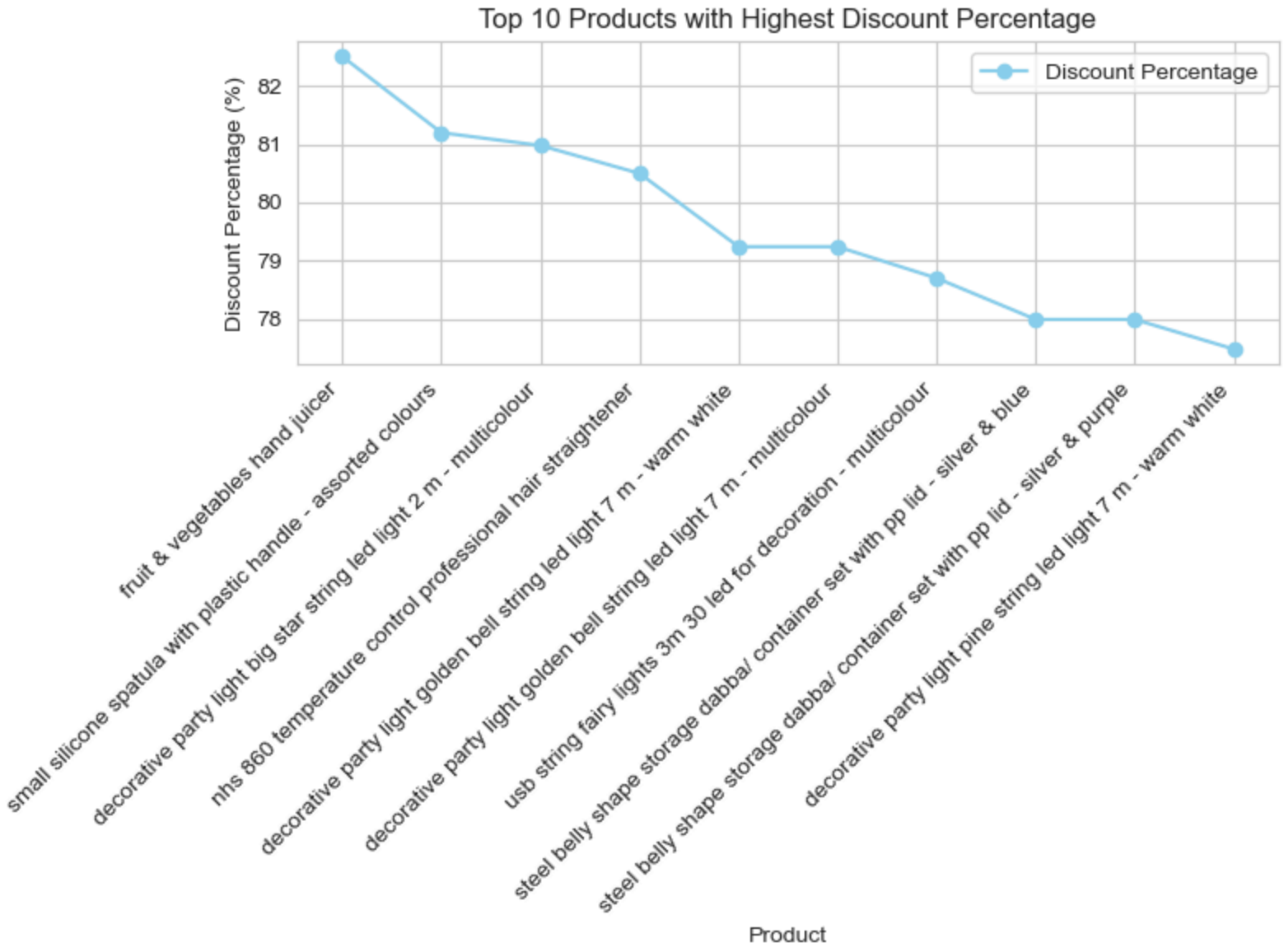
A line plot showing the products with the largest discounts, helping identify the items with the highest markdown.s

In [107…

```
top_10_discounts = df_c.sort_values(by='discount_percentage', ascending=False).head(10)

plt.figure(figsize=(8, 6))
plt.plot(top_10_discounts['product'],
         top_10_discounts['discount_percentage'],
         marker='o', color='skyblue', linestyle='-', label='Discount Percentage')

plt.xticks(rotation=45, ha='right')
plt.xlabel('Product')
plt.ylabel('Discount Percentage (%)')
plt.title('Top 10 Products with Highest Discount Percentage')
plt.grid(True)
plt.legend()
plt.tight_layout()
plt.show()
```



Key Takeaways:

- **Fruit & Vegetables Hand Juicer** has the highest discount percentage at 82%.
- **Small Silicone Spatula** and **Decorative Party Light Big Star String LED Light 2 M - Multicolour** follow closely, with discount percentages around 81%.

- **NHS 860 Temperature Control Professional Hair Straightener** and **Decorative Party Light Golden Bell String LED Light 7 M - Warm White** have a similar discount percentage, around 80%.
- **Decorative Party Light Golden Bell String LED Light 7 M - Multicolour** and **USB String Fairy Lights 3M 30 LED For Decoration - Multicolour** have a discount percentage of 79%.
- **Steel Belly Shape Storage Dabba/ Container Set With PP Lid - Silver & Blue** and **Steel Belly Shape Storage Dabba/ Container Set With PP Lid - Silver & Purple** have a discount percentage of 78%.
- **Decorative Party Light Pine String LED Light 7 M - Warm White** has the lowest discount percentage among the top 10 products, at 77%.

Recommendations:

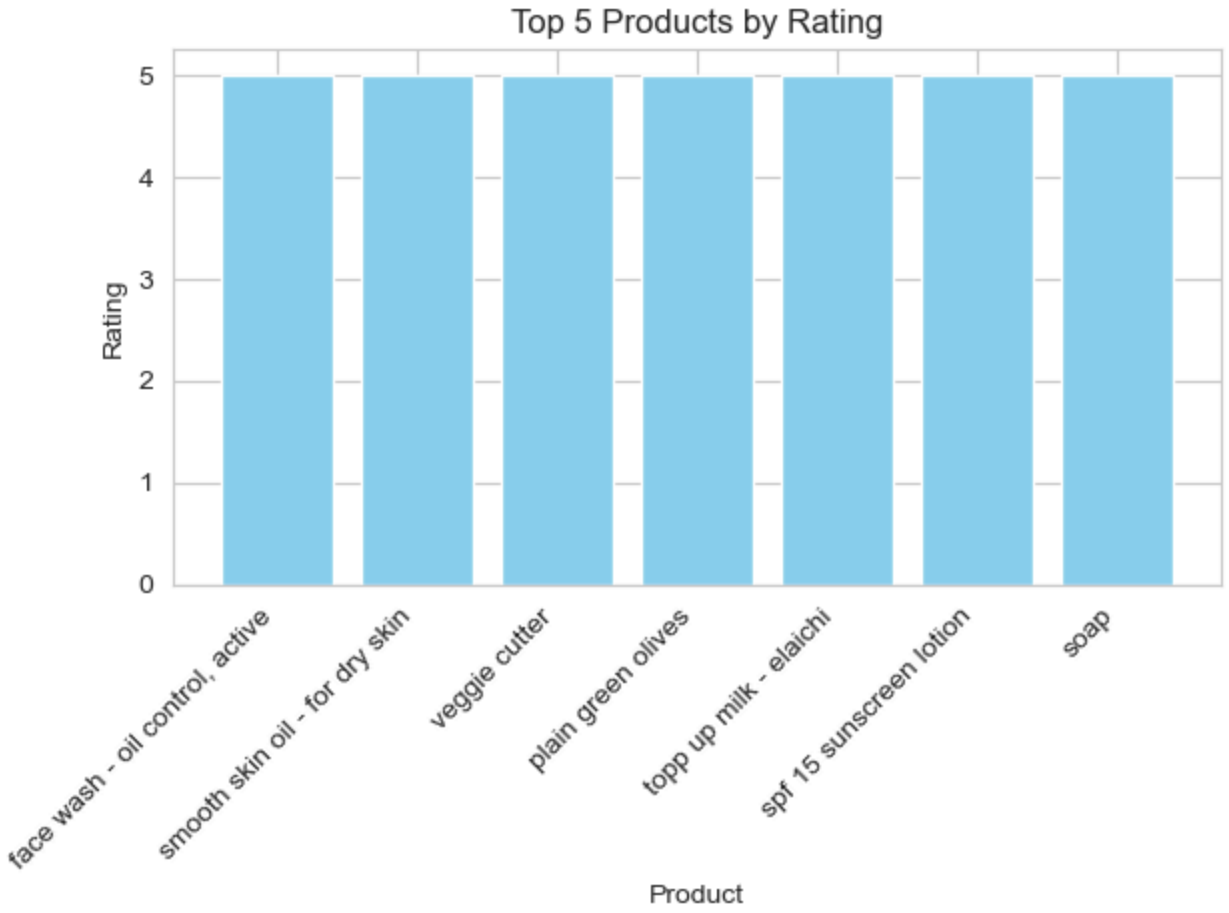
- **Leverage High-Discount Products**: Promote products with the highest discount percentages (**Fruit & Vegetables Hand Juicer**, **Small Silicone Spatula**, **Decorative Party Light Big Star String LED Light 2 M - Multicolour**) to attract customers and boost sales.
- **Optimize Discounts**: Analyze the impact of discounts on different product categories and adjust discount strategies accordingly to maximize profitability.
- **Monitor Competitor Pricing**: Stay informed about competitor pricing and adjust your discount strategy to remain competitive.

9.4) Top 7 Products by Rating :->

A bar plot of the highest-rated products, showing the top 5 products with the best ratings.

In [132…

```
top_7_products = df_c.nlargest(7, 'rating')[['product', 'rating']]
plt.bar(top_5_products['product'], top_5_products['rating'],
color='skyblue')
plt.xlabel('Product')
plt.ylabel('Rating')
plt.title('Top 5 Products by Rating')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



Key Takeaways:

- **Face Wash Oil Control, Active**, **Smooth Skin Oil - For Dry Skin**, **Veggie Cutter**, **Plain Green Olives**, **Topp Up Milk - Elaichi**, **SPF 15 Sunscreen Lotion**, and **Soap** are the top 5 products by rating, all with a perfect score of 5.

Recommendations:

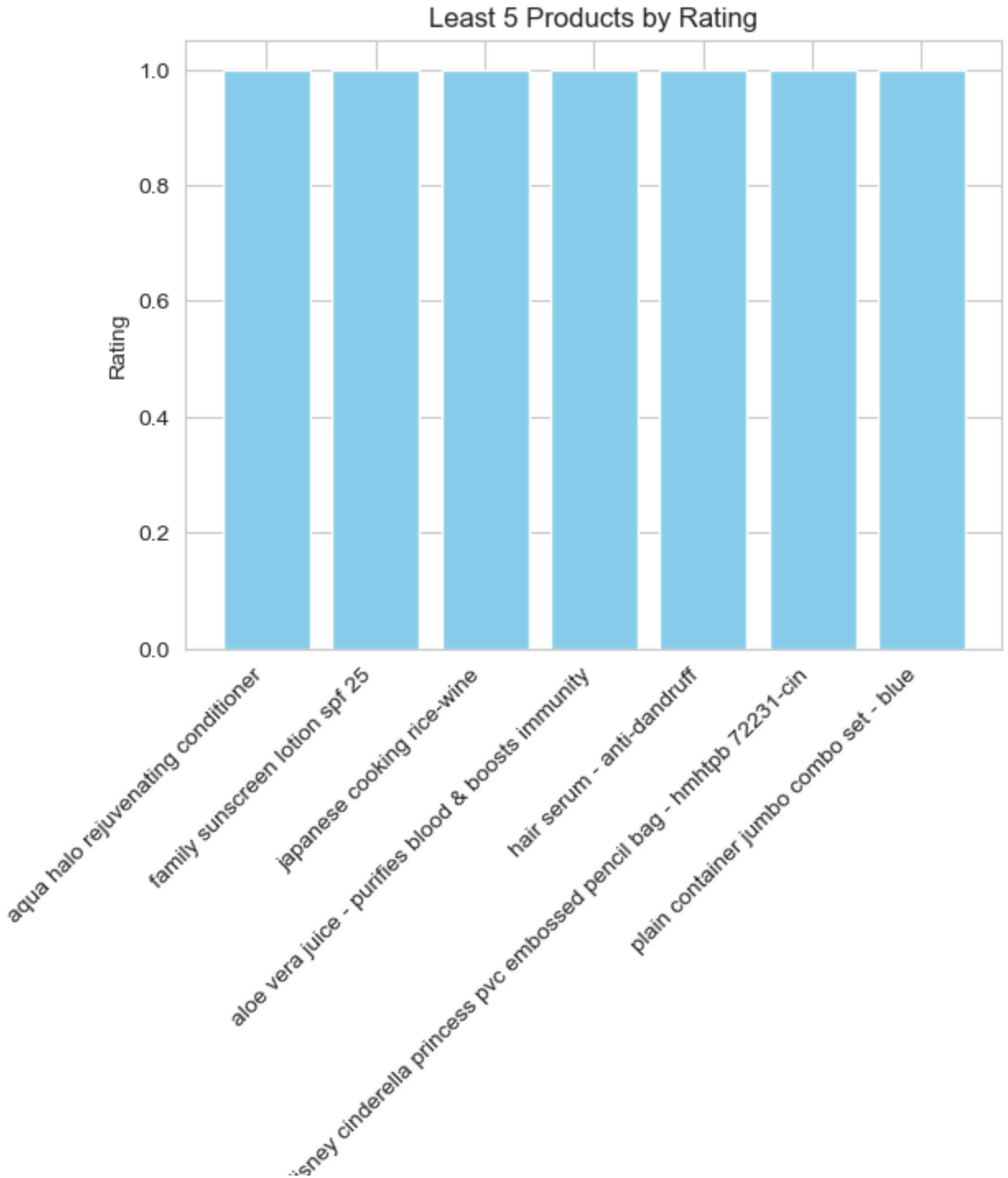
- **Leverage High-Rated Products**: Promote these top-rated products to attract new customers and increase sales.
- **Maintain Product Quality**: Ensure consistent quality and performance to maintain the high ratings and customer satisfaction.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high ratings.

9.5) Last 7 Products by Rating :->

A bar plot of the least-rated products, showing the top 5 products with the least ratings.

In [124…

```
Least_7_products = df_c.nsmallest(7, 'rating')[['product', 'rating']]
plt.bar(top_5_products['product'], top_5_products['rating'],
color='skyblue')
plt.xlabel('Product')
plt.ylabel('Rating')
plt.title('Least 5 Products by Rating')
plt.xticks(rotation=45, ha='right')
plt.show()
```


Least 5 Products by Rating
Rating
1.0
0.8
0.6
0.4
0.2
0.0
aqua halo rejuvenating conditioner
family sunscreen lotion spf 25
japanese cooking rice-wine
aloe vera juice - purifies blood & boosts immunity
hair serum - anti-dandruff
sney cinderella princess pvc embossed pencil bag - hmhtpb 72231-cin
plain container jumbo combo set - blue

Product

Key Takeaways:

- **Aqua Halo Rejuvenating Conditioner**, **Family Sunscreen Lotion SPF 25**, **Japanese Cooking Rice-Wine**, **Aloe Vera Juice - Purifies Blood & Boosts Immunity**, **Hair Serum - Anti-Dandruff**, **Disney Cinderella Princess PVC Embossed Pencil Bag - HMHTPB 72231-CIN**, and **Plain Container Jumbo Combo Set - Blue** are the lowest-rated products, all with a rating of 1.

Recommendations:

- **Review and Improve Low-Rated Products**: Analyze the feedback for these products to identify areas for improvement, such as quality, packaging, or product description.
- **Consider Product Removal**: If the issues with low-rated products cannot be resolved, consider removing them from the inventory.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high raings.

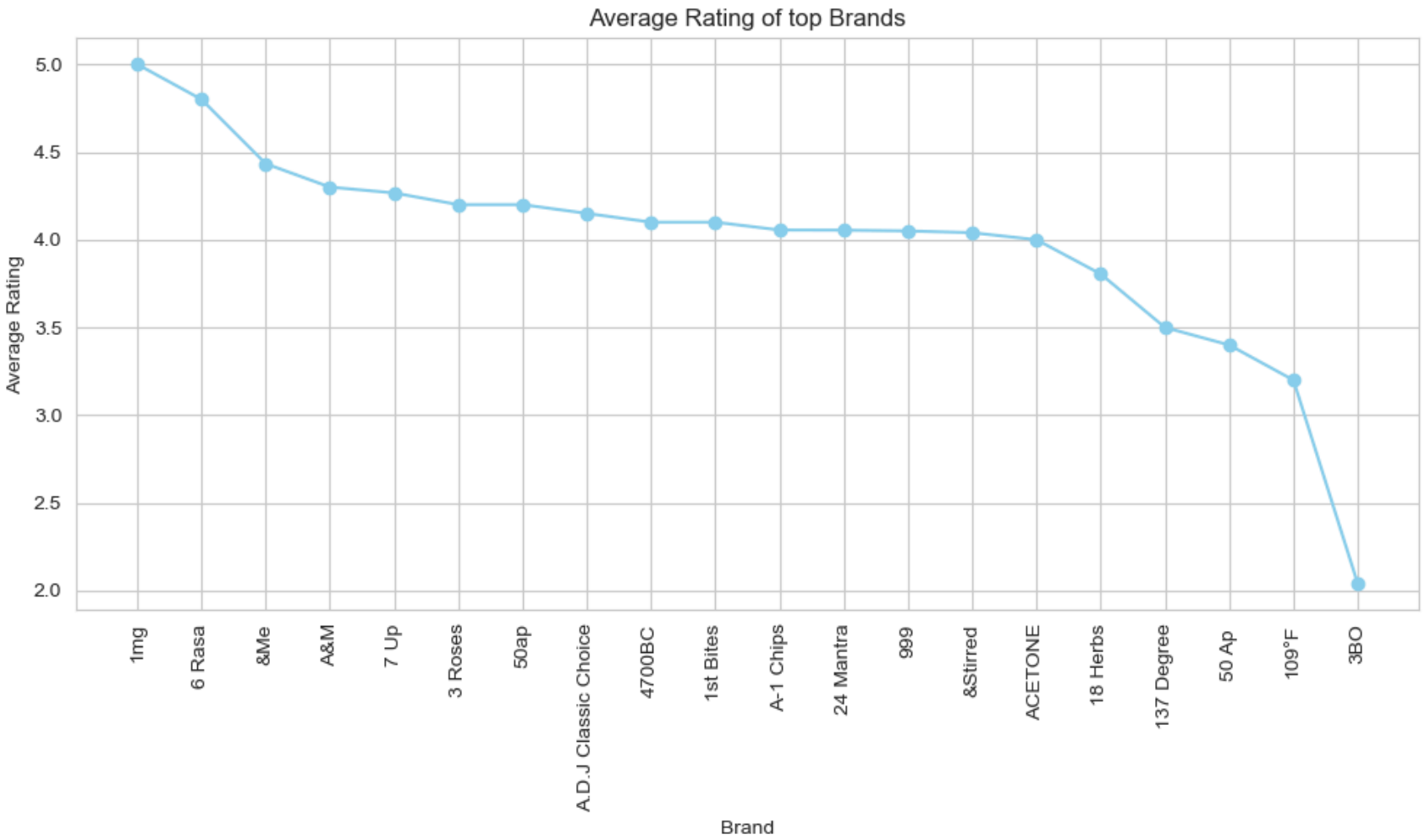
9.6) Average rating of Top Brands :->

A line plot that presents the average ratings of the top brands, offering insights into brand reputation based on customer ratings.

In [147…

```
brand_ratings = df_c.groupby('brand')['rating'].mean().reset_index().head(20)
brand_ratings_sorted = brand_ratings.sort_values(by='rating', ascending=False)

plt.figure(figsize=(10,6))
plt.plot(brand_ratings_sorted['brand'], brand_ratings_sorted['rating'], marker='o', color='skyblue')
plt.xticks(rotation=90)
plt.xlabel('Brand')
plt.ylabel('Average Rating')
plt.title('Average Rating of top Brands')
plt.tight_layout()
plt.show()
```



Key Takeaways:

- **1mg** and **6 Rasa** have the highest average ratings, scoring above 4.5.
- **&Me** and **A&M** follow closely with average ratings above 4.0.
- **7 Up**, **3 Roses**, and **50ap** have average ratings around 4.0.
- The average rating gradually declines for the remaining brands, with **380** having the lowest average rating.

Recommendations:

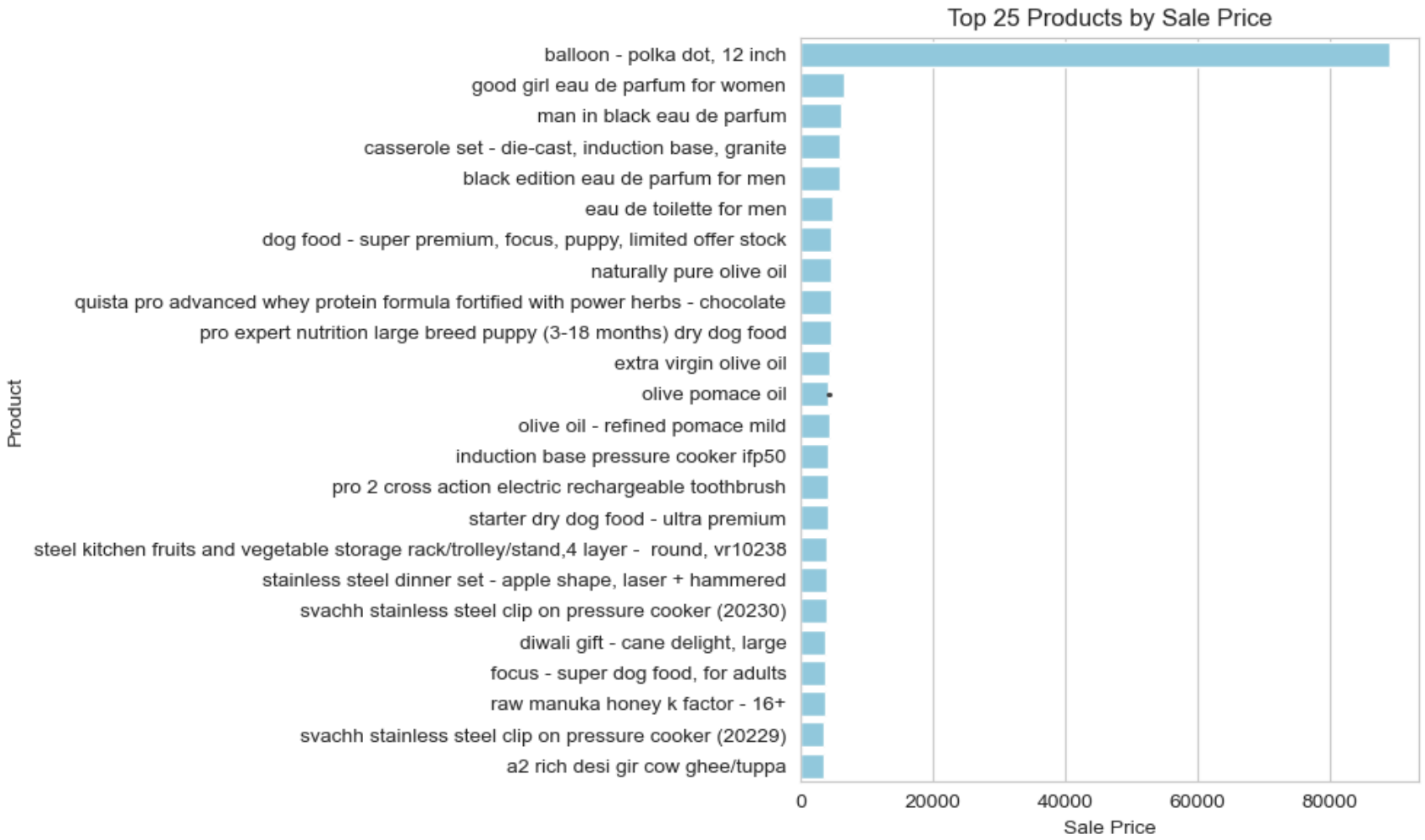
- **Leverage High-Rated Brands**: Promote brands with the highest average ratings (**1mg**, **6 Rasa**, **&Me**, **A&M**) to attract new customers and increase sales.
- **Analyze Low-Rated Brands**: Investigate the reasons for low ratings for brands like **380** and take steps to improve their performance, such as improving product quality, customer service, or marketing.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high ratings for all brands.

9.7) Top 25 Products by Sales :->

A bar plot displaying the top 25 products by sale price, which identifies the highest revenuegenerating items in the dataset.

In [153…

```
top_25_products = df_c[['product','sale_price']].sort_values(by='sale_price', ascending=False).head(25)
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='sale_price', y='product', data=top_25_products,color='skyblue')
ax.set_xlabel('Sale Price')
ax.set_ylabel('Product')
ax.set_title('Top 25 Products by Sale Price')
plt.tight_layout()
plt.show()
```



Key Takeaways:

- **Balloon - Polka Dot, 12 Inch** is the most expensive product with a sale price of 120,000.
- **Good Girl Eau De Parfum For Women** and **Man In Black Eau De Parfum** are the second and third most expensive products, priced around 90,000 and 80,000, respectively.
- **Casserole Set - Die-Cast, Induction Base, Granite** and **Black Edition Eau De Parfum For Men** have a similar sale price, around 70,000.
- The remaining products have significantly lower sale prices, ranging from 10,000 to 30,000.

Recommendations:

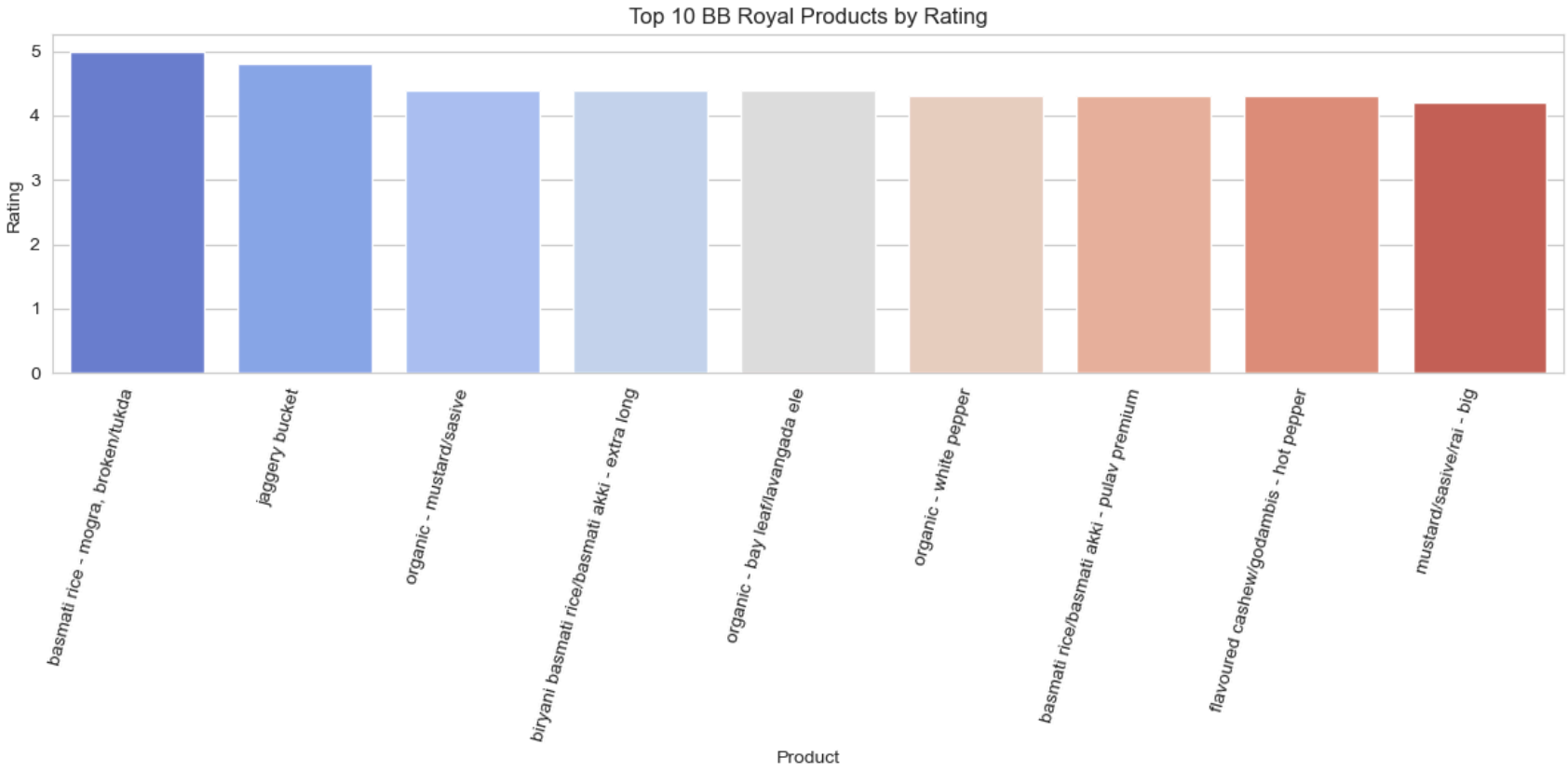
- **Analyze High-Price Products**: Investigate the factors contributing to the high prices of the top products (**Balloon - Polka Dot, 12 Inch**, **Good Girl Eau De Parfum For Women**, **Man In Black Eau De Parfum**). Consider if these prices are justified by the product's value or if adjustments are needed.
- **Optimize Product Mix**: Evaluate the product mix and consider whether the high-priced products are aligned with the overall business strategy and target customer needs.
- **Monitor Competitor Pricing**: Stay informed about competitor pricing for similar products and adjust your pricing strategy accrdingly.

9.8) Top 10 BigBasket Royal Products by Rating :->

A bar plot of the highest-rated products from the in house brand for groceries 'BB Royal', showing the top 10 products with the best rating.

In [180…

```
bb_royal_top10 = df_c[df_c['brand'] == 'bb Royal'].nlargest(10,
'rating')
# Check if there are enough entries
if not bb_royal_top10.empty:
 plt.figure(figsize=(12, 6))
 sns.barplot(x='product', y='rating', data=bb_royal_top10,hue='product', legend=False, palette='coolwarm')
 plt.xticks(rotation=75, ha='right')
 plt.xlabel('Product')
 plt.ylabel('Rating')
 plt.title('Top 10 BB Royal Products by Rating')
 plt.tight_layout()
 plt.show()
else:
 print("No products found for brand 'bb Royal'.")
```



Key Takeaways:

- **Basmati Rice - Mogra, Broken/Tukda** is the top-rated BB Royal product with a rating of 5.
- **Jaggery Bucket** follows closely with a rating of 4.75.
- **Organic - Mustard/Sasive** and **Biryani Basmati Rice/Basmati Akki - Extra Long** have a similar rating, around 4.5.
- **Organic - Bay Leaf/Lavangada Ele** and **Organic - White Pepper** have a rating of 4.25.
- **Basmati Rice/Basmati Akki - Pulav Premium** and **Flavoured Cashew/Godambis - Hot Pepper** have a rating of 4.0.
- **Mustard/Sasive/Rai - Big** has the lowest rating among the top 10 BB Royal products, with a rating of 3.75.

Recommendations:

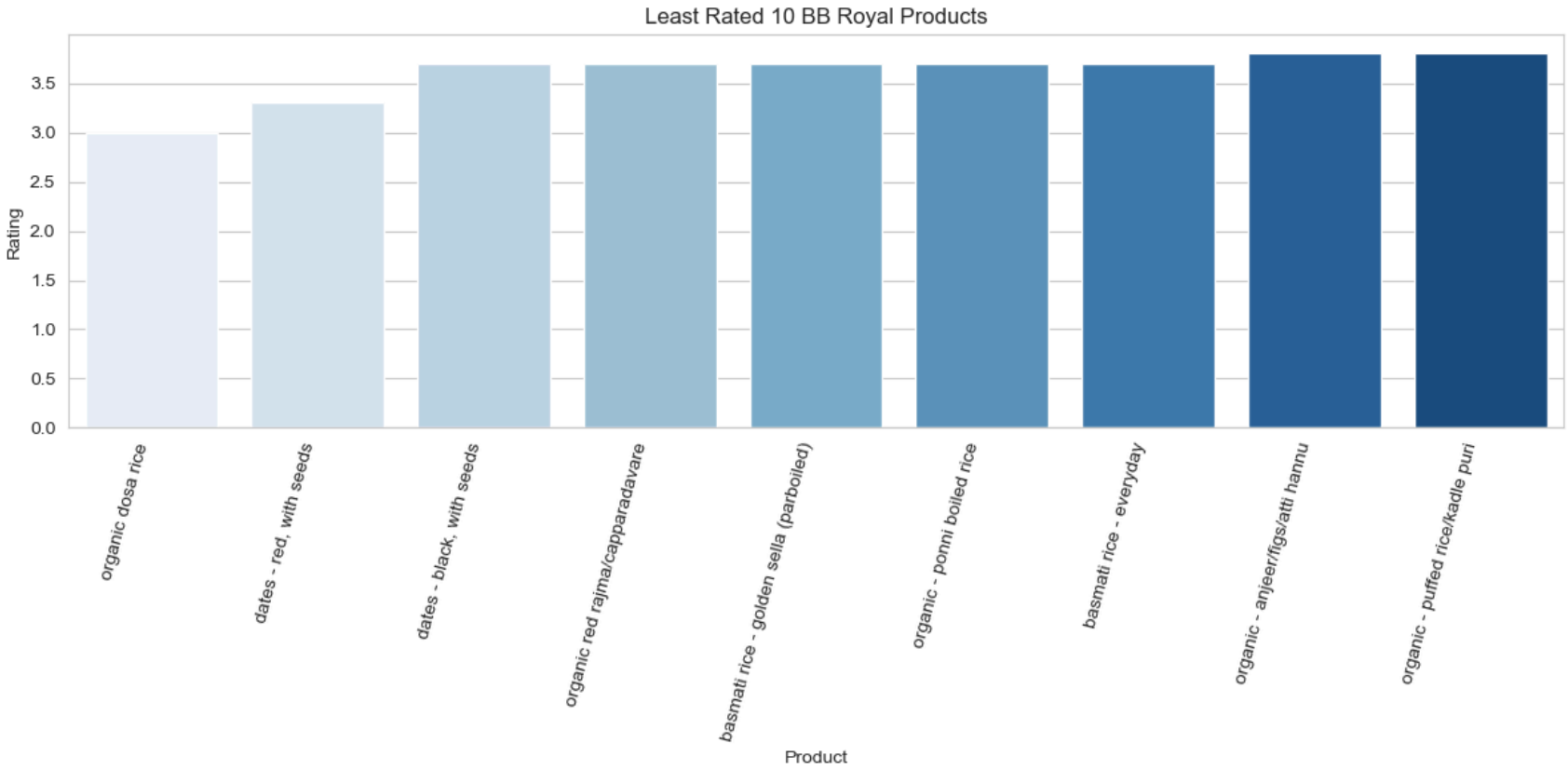
- **Leverage High-Rated Products**: Promote the top-rated products (**Basmati Rice - Mogra, Broken/Tukda**, **Jaggery Bucket**, **Organic - Mustard/Sasive**, **Biryani Basmati Rice/Basmati Akki - Extra Long**) to attract new customers and increase sales.
- **Analyze Low-Rated Products**: Investigate the reasons for the lower rating of **Mustard/Sasive/Rai - Big** and take steps to improve its performance, such as improving product quality or customer service.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high ratings for all BB Roya products.

9.9) Least Rated 10 BigBasket Royal Products :->

A bar plot of the least-rated products from the in house brand for groceries 'BB Royal'.

In [189…

```
bb_royal_top10 = df_c[df_c['brand'] == 'bb Royal'].nsmallest(10,
'rating')
# Check if there are enough entries
if not bb_royal_top10.empty:
 plt.figure(figsize=(12, 6))
 sns.barplot(x='product', y='rating', data=bb_royal_top10,hue='product', legend=False, palette='Blues')
 plt.xticks(rotation=75, ha='right')
 plt.xlabel('Product')
 plt.ylabel('Rating')
 plt.title('Least Rated 10 BB Royal Products')
 plt.tight_layout()
 plt.show()
else:
 print("No products found for brand 'bb Royal'.")
```



Key Takeaways:

- **Organic Dosa Rice** has the lowest rating among the top 10 BB Royal products, with a rating of 3.0.
- **Dates Red, With Seeds** and **Dates Black, With Seeds** follow closely with ratings of 3.25.
- **Organic Red Rajma/Capparadavare** has a rating of 3.5.
- **Basmati Rice - Golden Sella (Parboiled)** and **Organic - Ponni Boiled Rice** have a rating of 3.75.
- **Basmati Rice - Everyday**, **Organic - Anjeer/Figs/Atti Hannu**, and **Organic - Puffed Rice/Kadle Puri** have the highest ratings among the top 10 BB Royal products, with a rating of 4.0.

Recommendations:

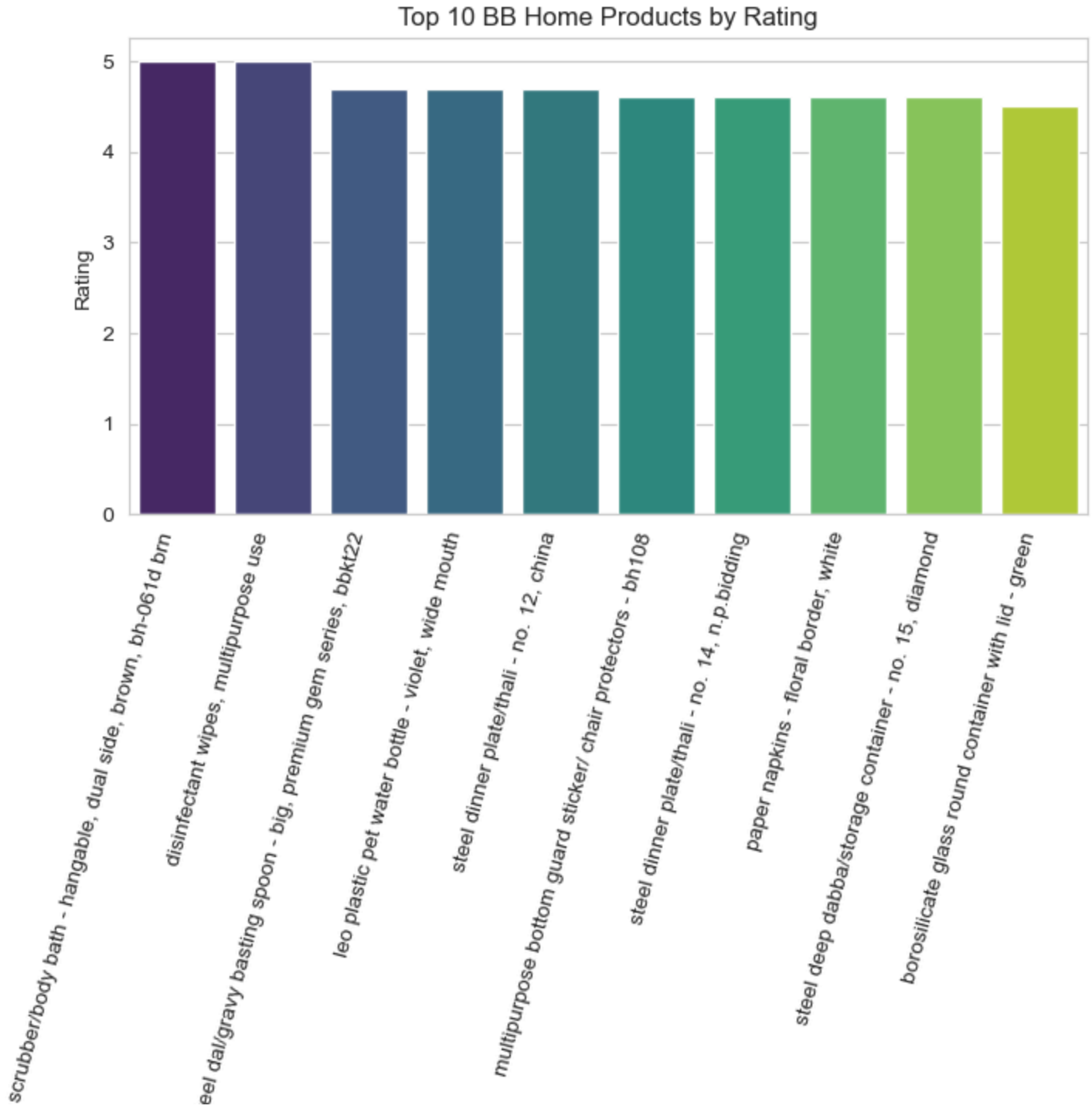
- **Analyze Low-Rated Products**: Investigate the reasons for the low ratings of **Organic Dosa Rice**, **Dates Red, With Seeds**, and **Dates Black, With Seeds** and take steps to improve their performance, such as improving product quality or customer service.
- **Leverage High-Rated Products**: Promote the top-rated products (**Basmati Rice - Everyday**, **Organic - Anjeer/Figs/Atti Hannu**, **Organic - Puffed Rice/Kadle Puri**) to attract new customers and increase sales.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high ratings for all BB Royal products.

9.10) Top 10 BigBasket Home Products by Rating :->

A bar plot of the highest-rated products from the in house brand for household items 'BB Home', showing the top 10 products with the best ratings.

In [195…

```
bb_royal_top10 = df_c[df_c['brand'] == 'BB Home'].nlargest(10,'rating')
# Check if there are enough entries
if not bb_royal_top10.empty:
 plt.figure(figsize=(8, 4))
 sns.barplot(x='product', y='rating', data=bb_royal_top10,hue='product', legend=False, palette='viridis')
 plt.xticks(rotation=75, ha='right')
 plt.xlabel('Product')
 plt.ylabel('Rating')
 plt.title('Top 10 BB Home Products by Rating')
 plt.show()
else:
 print("No products found for brand 'bb Royal'.")
```


Top 10 BB Home Products by Rating
Rating
0
1
2
3
4
5
scrubber/body bath - hangable, dual side, brown, bh-061d brn
disinfectant wipes, multipurpose use
eel dal/gravy basting spoon - big, premium gem series, bbkt22
leo plastic pet water bottle - violet, wide mouth
steel dinner plate/thali - no. 12, china
multipurpose bottom guard sticker/ chair protectors - bh108
steel dinner plate/thali - no. 14, n.p.bidding
paper napkins - floral border, white
steel deep dabba/storage container - no. 15, diamond
borosilicate glass round container with lid - green

Key Takeaways:

- **Leaf Scrubber/Body Bath - Hangable, Dual Side, Brown, BH-061D BRN** is the top-rated BB Home product with a rating of 5.
- **Disinfectant Wipes, Multipurpose Use**, **Steel Dal/Gravy Basting Spoon, Big, Premium Gem Series, BBKT22**, and **Leo Plastic Pet Water Bottle - Violet, Wide Mouth** follow closely with ratings of 4.75.
- **Steel Dinner Plate/Thali - No. 12, China** and **Multipurpose Bottom Guard Sticker/Chair Protectors - BH108** have a rating of 4.5.
- **Steel Dinner Plate/Thali - No. 14, N.P.Bidding** and **Paper Napkins - Floral Border, White** have a rating of 4.25.
- **Steel Deep Dabba/Storage Container - No. 15, Diamond** and **Borosilicate Glass Round Container With Lid - Green** have the lowest ratings among the top 10 BB Home products, with a rating of 4.0.

Recommendations:

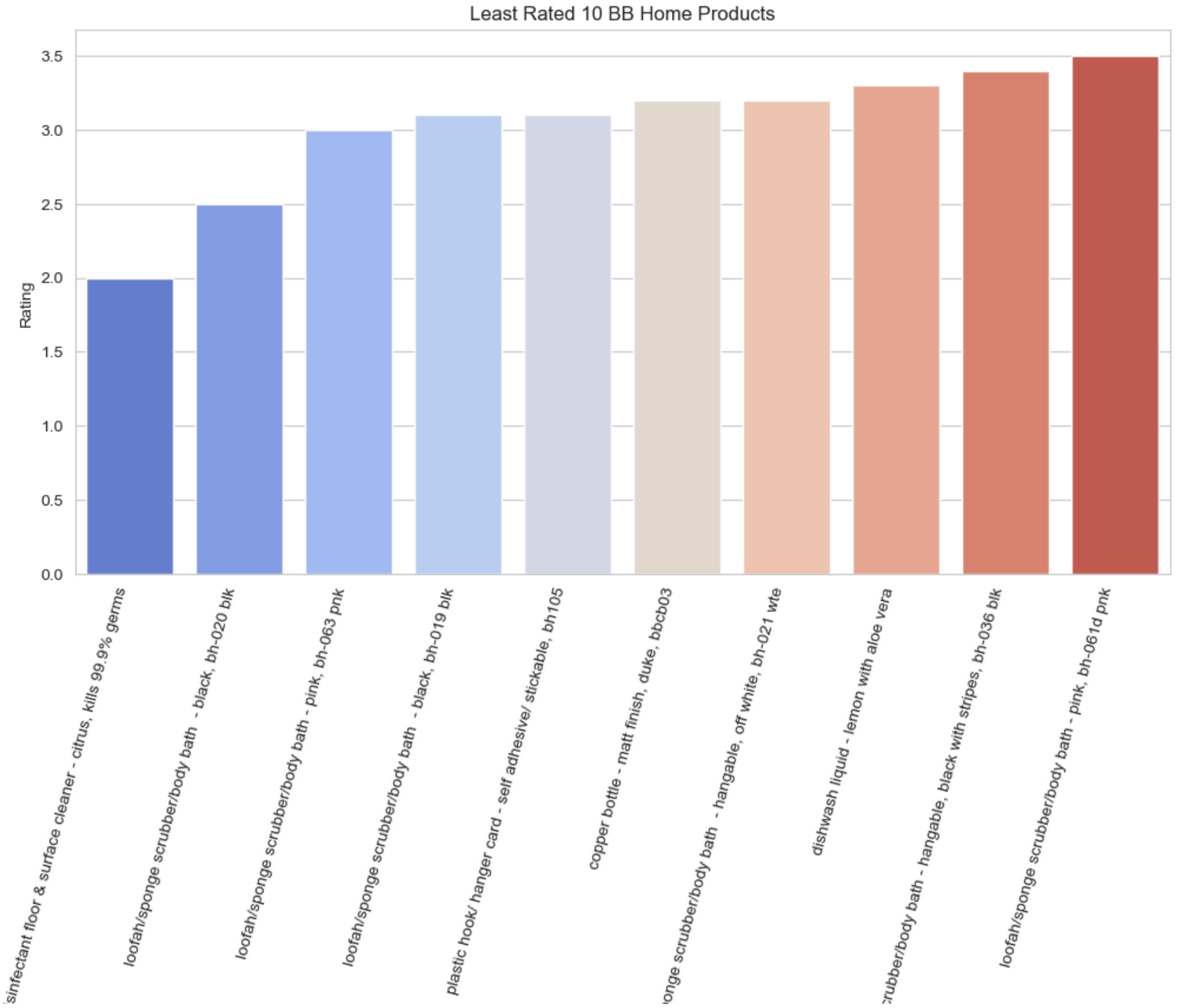
- **Leverage High-Rated Products**: Promote the top-rated products (**Leaf Scrubber/Body Bath - Hangable, Dual Side, Brown, BH-061D BRN**, **Disinfectant Wipes, Multipurpose Use**, **Steel Dal/Gravy Basting Spoon, Big, Premium Gem Series, BBKT22**, **Leo Plastic Pet Water Bottle - Violet, Wide Mouth**) to attract new customers and increase sales.
- **Analyze Low-Rated Products**: Investigate the reasons for the lower ratings of **Steel Deep Dabba/Storage Container - No. 15, Diamond** and **Borosilicate Glass Round Container With Lid - Green** and take steps to improve their performance, such as improving product quality or customer service.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high ratings for all BB Home products.

9.11) Least Rated 10 BigBasket Home Products :->

A bar plot of the Least-rated products from the in house brand for household items 'BigBasket Home'.

In [204…

```
bb_royal_top10 = df_c[df_c['brand'] == 'BB Home'].nsmallest(10,'rating')
# Check if there are enough entries
if not bb_royal_top10.empty:
 plt.figure(figsize=(12, 6))
 sns.barplot(x='product', y='rating', data=bb_royal_top10,hue='product', legend=False, palette='coolwarm')
 plt.xticks(rotation=75, ha='right')
 plt.xlabel('Product')
 plt.ylabel('Rating')
 plt.title('Least Rated 10 BB Home Products')
 plt.show()
else:
 print("No products found for brand 'BB Home'.")
```


Least Rated 10 BB Home Products
Rating
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
sinfectant floor & surface cleaner - citrus, kills 99.9% germs
loofah/sponge scrubber/body bath - black, bh-020 blk
loofah/sponge scrubber/body bath - pink, bh-063 pnk
loofah/sponge scrubber/body bath - black, bh-019 blk
plastic hook/ hanger card - self adhesive/ stickable, bh105
copper bottle - matt finish, duke, bbcb03
onge scrubber/body bath - hangable, off white, bh-021 wte
dishwash liquid - lemon with aloe vera
rubber/body bath - hangable, black with stripes, bh-036 blk
loofah/sponge scrubber/body bath - pink, bh-061d pnk

Key Takeaways:

- **Disinfectant Floor & Surface Cleaner - Citrus, Kills 99.9% Germs** has the lowest rating among the top 10 BB Home products, with a rating of 2.0.
- **Loofah/Sponge Scrubber/Body Bath - Black, BH-020 BLK** and **Loofah/Sponge Scrubber/Body Bath - Pink, BH-063 PNK** follow closely with ratings of 2.5.
- **Loofah/Sponge Scrubber/Body Bath - Black, BH-019 BLK** and **Plastic Hook/Hanger Card - Self Adhesive/Stickable, BH105** have a rating of 3.0.
- The remaining products have ratings between 3.0 and 3.5.

Recommendations:

- **Analyze Low-Rated Products**: Investigate the reasons for the low ratings of **Disinfectant Floor & Surface Cleaner - Citrus, Kills 99.9% Germs**, **Loofah/Sponge Scrubber/Body Bath - Black, BH-020 BLK**, and **Loofah/Sponge Scrubber/Body Bath - Pink, BH-063 PNK** and take steps to improve their performance, such as improving product quality or customer service.
- **Gather Customer Feedback**: Collect feedback from customers to identify areas for improvement and maintain high ratings for all BB Home poducts.

Conclusion : Big Basket Sales and Product Analysis ->

Key Insights:

1. Product Categories

- **Beauty & Hygiene**: Dominates with the highest product count, followed by **Kitchen**, **Garden & Pets**.

- **Beverages** and **Baby Care**: Have fewer products, suggesting room for potential expansion.

2. Top-Selling Products

- **Turmeric Powder** and **Cow Ghee** are among the top-selling items, showing strong customer demand for essential food products.
- **Recommendation**: Focus promotions on high-demand products like these to maximize sales.

3. Popular Brands

- **bb Royal** and **BB Home**: Leading in sales volume, reflecting customer loyalty toward Big Basket's in-house brands.
- **Amul** and **Himalaya**: Perform well, indicating strong market trust.

4. Discounted Products

- **Highest Discounts**: Products like the **Fruit & Vegetables Hand Juicer** have discounts up to **82%**.
- **Recommendation**: Leverage these high-discount items to attract price-sensitive customers.

5. Product Ratings

- **Highly Rated Products**: **Face Wash** and **Veggie Cutter** receive top ratings, showing high customer satisfaction.
  - **Recommendation**: Promote these products to boost Big Basket's brand image.
- **Low-Rated Products**: Items such as **Aqua Halo Rejuvenating Conditioner** received low ratings, suggesting potential quality issues.
  - **Recommendation**: Address quality concerns in low-rated products to meet customer expectations.

6. Pricing Strategy

- **High-Priced Items**: Luxury products like **perfumes** stand out.
- **Recommendation**: Monitor competitor pricing to ensure these items align with Big Basket’s target market.

7. Brand Ratings

- **Top Brands by Rating**: Brands like **1mg** and **6 Rasa** have high average ratings.
- **Low-Rated Brands**: Should consider quality improvements to enhance customer satisfaction.

Overall Conclusion

Big Basket's strategy effectively balances high-frequency essentials, strategic discounting, and the development of in-house brands to foster customer loyalty. To further strengthen its market position, there are opportunities for improvement, such as:

- **Addressing quality concerns** for lower-rated products to ensure consistent customer satisfaction.
- **Expanding the product range** in underperforming categories to capture untapped market potential.

By continuously adapting to customer needs and enhancing product quality, Big Basket can further solidify its leadership in the competitive e-commerce space.