Big Basket Sales and Product Analysis

Objective

The analysis emphasizes leveraging top-performing products through targeted marketing and bundling, addressing low-performing items via customer feedback and pricing strategies, and optimizing product portfolios across diverse categories. Additionally, it highlights the importance of customer reviews, inventory management, and pricing strategies to drive sales, improve profitability, and enhance brand reputation.

Submitted by: Rohit Sharma

Submitted to : Neeraj ma'am

Importing Libraries

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

1.LOAD DATASET.`

```
In [3]: df=pd.read_csv("C:/Users/123/Desktop/file han/BigBasket Products.csv")
```

2.Use head function to look for first 12 rows

```
In [5]: df.head(12)
```

	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...

3.Get Description of the data in the DataFrame.

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

Out[7]:

	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

4: Get Information about the DataFrame.

In [9]: `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
```

5. Find out Top & least sale products.

top sale produts

```
In [12]:  top_sale = df.sort_values(by='sale_price', ascending=False).head(10)
          top_sale
```

Out[12]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description
1249	1250	Beard Kit	Beauty & Hygiene	Men's Grooming	Uncle Tony	112475.00	3300.0	Combos & Gift Sets	NaN	The combination of a beard oil, a beard wash, ...
248	249	4mm Aluminium Induction Base Chapati Roti Tawa...	Kitchen, Garden & Pets	Cookware & Non Stick	HAZEL	111649.00	1289.0	Tawa & Sauce Pan	NaN	Hazel Aluminium Tawa has an ergonomic design f...
436	437	Balloon - Polka Dot, 12 Inch	Cleaning & Household	Party & Festive Needs	B Vishal	88899.00	129.0	Caps, Balloons & Candles	3.9	Whether it is a party in the office, a Christm...
288	289	Arrabbiata Tomato Pasta Sauce With Chilli	Gourmet & World Food	Sauces, Spreads & Dips	Montanini	22325.00	325.0	Mustard & Cheese Sauces	5.0	NaN
25301	25302	Bravura Clipper	Kitchen, Garden & Pets	Pet Food & Accessories	Wahl	12500.00	12500.0	Pet Cleaning & Grooming	NaN	The bravura clipper is a must-have clipper for...
21761	21762	Pet Food - N&D Team Breeder Puppy Top Farmina	Kitchen, Garden & Pets	Pet Food & Accessories	Farmina	10090.00	10090.0	Pet Meals & Treats	NaN	Dog Food Adult Health Nutritional Dog Food.
12669	12670	Epilator SE9-9961 Legs-Body-Face	Beauty & Hygiene	Feminine Hygiene	Braun	8184.44	10769.0	Hair Removal	NaN	This cordless epilator has a sonic exfoliation...
23082	23083	Gas Stove-4 Burner Royale Plus Schott Glass, B...	Kitchen, Garden & Pets	Cookware & Non Stick	Prestige	7999.00	12245.0	Gas Stove	NaN	Prestige Royale Plus Gas Stove. Add a Touch of...
2781	2782	Extra Virgin Olive Oil	Gourmet & World Food	Oils & Vinegar	Abbies	7299.00	7299.0	Extra Virgin Olive Oil	NaN	Suitable to cook Indian meals due to its neutr...
25797	25798	4 Burner Gas Stove - Marvel Plus Glass Tables,...	Kitchen, Garden & Pets	Cookware & Non Stick	Prestige	7270.00	9695.0	Gas Stove	NaN	Prestige Marvel Plus Gas Stove . A Smart and C...

least sale products

```
In [14]:  least_sale = df.sort_values(by='sale_price', ascending=True).head(10)
```

```
least_sale
```

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description
26976	26977	Curry Leaves	Fruits & Vegetables	Herbs & Seasonings	Fresho	2.45	15.00	Indian & Exotic Herbs	NaN	With dark green and glossy appearance, curry l...
21312	21313	Serum	Beauty & Hygiene	Hair Care	Livon	3.00	3.00	Hair Oil & Serum	2.5	Instantly Softens and Smoothens Dry, Rough, Ta...
11306	11307	Happy Happy Choco-Chip Cookies	Snacks & Branded Foods	Biscuits & Cookies	Parle	5.00	5.00	Glucose & Milk Biscuits	4.2	Baked to perfection and topped with delicious ...
2761	2762	Orbit Sugar-Free Chewing Gum - Lemon & Lime	Snacks & Branded Foods	Chocolates & Candies	Wrigleys	5.00	5.00	Mints & Chewing Gum	4.2	With Orbit Sugarfree Chewing Gums, there's no ...
3445	3446	Marie Light Biscuits - Active	Snacks & Branded Foods	Ready To Cook & Eat	Sunfeast	5.00	5.00	Breakfast & Snack Mixes	4.5	EAN Code: 8901725114848 Country of origin: In...
9971	9972	Tomato - Local, Organically Grown	Fruits & Vegetables	Fresh Vegetables	Fresho	5.00	6.25	Organic Vegetables	NaN	Fresho brings to you an exquisite range of loc...
27490	27491	50-50 Timepass Salted Biscuits	Snacks & Branded Foods	Biscuits & Cookies	Britannia	5.00	5.00	Salted Biscuits	4.2	Britannia 50-50 Timepass Salted Biscuits are a...
22072	22073	Tiger Chocolate Cream Biscuits	Snacks & Branded Foods	Biscuits & Cookies	Britannia	5.00	5.00	Cream Biscuits & Wafers	4.2	Britannia Tiger Chocolate Cream Biscuits will ...
18290	18291	Sugar Coated Chocolate	Snacks & Branded Foods	Chocolates & Candies	Cadbury Gems	5.00	5.00	Chocolates	4.2	Cadbury Gems is chocolate flavor, rich and cho...
21228	21229	Dish Shine Bar	Cleaning & Household	Detergents & Dishwash	Exo	5.00	5.00	Dishwash Bars & Powders	4.2	Exo Dish Shine Bar makes your vessels aromatic...

6. Measuring Discount

```python
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)
```

	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

7. Finding out the Missing Values from the Dataset.

```python
df.isna().sum()
```

```
Out[18]:  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
```

```python
In [19]:  # Checking NaN value in product column
          df[df['product'].isna()]
```

Out[19]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentag
14363	14364	NaN	Beverages	Coffee	Cothas Coffee	200.0	240.0	Ground Coffee	4.2	Cothas Specialty Blend Coffee and Chicory incl...	16.6

```python
In [20]:  # Checking NaN value in brand column
          df[df['brand'].isna()]
```

Out[20]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentag
9765	9766	Food Package - Medium	Cleaning & Household	Disposables, Garbage Bag	NaN	50.0	50.0	Aluminium Foil, Clingwrap	NaN	NaN	0

```python
In [22]:  # # Checking NaN value in sale price column
          df[df['sale_price'].isnull()]
```

Out[22]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_per
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

```python
In [27]:  # replacing NaN With "Unknown"
          df.loc[df['product'].isna(),'product']='Unknown'
```

```
df['product']
```

Out[27]:
```
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 [30]:
```
# Checking NaN value in product column
df[df['product'].isna()]
```

Out[30]:

index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentage

In [33]:
```
# reploicing NaN with "unknown"
df.loc[df['brand'].isna(),'brand'] = 'unknown'
df['brand']
```

Out[33]:
```
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 [37]:
```
# handling NaN value with median
df['sale_price'] = df['sale_price'].fillna(df['sale_price'].median())
```

In [40]:
```
df[df['sale_price'].isnull()]
```

Out[40]:

index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentage

In [42]:
```
# Handling NaN value with "Median" in rating column
df['rating'] = df['rating'].fillna(df['rating'].median())
```

In [46]:
```
# reploicing NaN with "unknown" in description column
df.loc[df['description'].isna(),'description'] = 'unknown'
```

In [49]:
```
# Checking NaN value in description column
df[df['description'].isna()]
```

Out[49]:

index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentage

In [51]:
```
# handling Nan Value with Median in discount percentage
df['discount_percentage'] = df['discount_percentage'].fillna(df['discount_percentage'].median())
```

In [53]:
```
# Checking NaN value in discount percentage column
df[df['discount_percentage'].isna()]
```

Out[53]:

index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_percentage

Additional

In [56]:
```
def handle_outliers(column_name):
    Q1 = df[column_name].quantile(0.25)
    Q3 = df[column_name].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5*IQR
    upper_bound = Q3 + 1.5*IQR
    outliers = df[(df[column_name]<lower_bound) | (df[column_name]>upper_bound)]
    df[column_name] = df[column_name].clip(lower_bound , upper_bound)


    print(f"Outliers in {column_name}:\n", outliers)
```

```
handle_outliers('rating')
```

```
Outliers in rating:
       index                                          product  \
1          2                             Water Bottle - Orange
2          3                         Brass Angle Deep - Plain, No.2
5          6                      Germ - Removal Multipurpose Wipes
6          7                                          Multani Mati
8          9  Biotin & Collagen Volumizing Hair Shampoo + Bi...
...      ...                                               ...
27518  27519            Gluten-Free Vegetable Millet Khichdi Mix
27533  27534               Herbal Lemon Facial Massage Cream
27534  27535                              Chia Quinoa Cookies
27535  27536                          Cookies - Butter Vanilla
27548  27549                        Apple Cider Vinegar Shampoo

                       category            sub_category             brand  \
1          Kitchen, Garden & Pets  Storage & Accessories        Mastercook
2             Cleaning & Household            Pooja Needs               Trm
5             Cleaning & Household    All Purpose Cleaners    Nature Protect
6                 Beauty & Hygiene              Skin Care         Satinance
8                 Beauty & Hygiene              Hair Care        StBotanica
...                           ...                    ...               ...
27518      Gourmet & World Food  Cooking & Baking Needs         Graminway
27533          Beauty & Hygiene              Skin Care            Olivia
27534      Gourmet & World Food  Sauces, Spreads & Dips           Prasukh
27535      Bakery, Cakes & Dairy  Cookies, Rusk & Khari         Momkhatai
27548          Beauty & Hygiene              Hair Care  Morpheme Remedies

       sale_price  market_price                        type  rating  \
1          180.00         180.0        Water & Fridge Bottles     2.3
2          119.00         250.0               Lamp & Lamp Oil     3.4
5          169.00         199.0  Disinfectant Spray & Cleaners     3.3
6           58.00          58.0                    Face Care     3.6
8         1098.00        1098.0        Shampoo & Conditioner     3.5
...           ...           ...                         ...     ...
27518      350.00         350.0          Flours & Pre-Mixes     1.0
27533      191.25         225.0                   Body Care     3.2
27534       76.50          85.0  Jams, Marmalade, Spreads     3.3
27535      120.00         120.0  Bakery Biscuits, Cookies     3.3
27548      499.00         499.0        Shampoo & Conditioner     5.0

                                      description  discount_percentage
1      Each product is microwave safe (without lid), ...                 0.00
2      A perfect gift for all occasions, be it your m...                52.40
5      Stay protected from contamination with Multipu...                15.08
6      Satinance multani matti is an excellent skin t...                 0.00
8      An exclusive blend with Vitamin B7 Biotin, Hyd...                 0.00
...                                           ...                  ...
27518  Graminway Little Millet Khichdi Mix makes your...                 0.00
27533  Get a toned complexion that's free from dirt a...                15.00
27534  The Energy Cookies get its name because they'r...                10.00
27535  Momkhatai Butter Vanilla are classic nankhatai...                 0.00
27548  Say no to dull, lifeless, dry and damaged hair...                 0.00

[6372 rows x 11 columns]
```

8.Finding the outliers from the dataset according to the columns

```python
# calculate the IQR
Q1=df['rating'].quantile(0.25)
Q3=df['rating'].quantile(0.75)
IQR=Q3-Q1

# define outlier range
lower_bound = Q1 - 1.5*IQR
upper_bound = Q3 + 1.5*IQR

# Identify outliers
outliers = df[(df['sale_price'] < lower_bound) | (df['sale_price'] > upper_bound)]

# Display the calculated values
print(f"Q1: {Q1}")
print(f"Q3: {Q3}")
print(f"IQR: {IQR}")
print(f"lower: {lower_bound}")
print(f"upper: {upper_bound}")
```

```
Q1: 4.0
Q3: 4.2
IQR: 0.20000000000000018
lower: 3.6999999999999997
upper: 4.5
```

9. Creating the Plots or visualizations.

1.Bar plot of top 10 sold products

```
In [66]: # Sort the data
         sorted_data = top_sale.sort_values(by='market_price',ascending=False)

         # Create the barplot
         plt.figure(figsize=(12,8))
         ax=sns.barplot(x='market_price', y='product', data=sorted_data, color="#FE9900")
         plt.xticks(rotation=90)
         plt.title("Top 10 Sale Products")
         for container in ax.containers:
             ax.bar_label(container)
         plt.show()
```
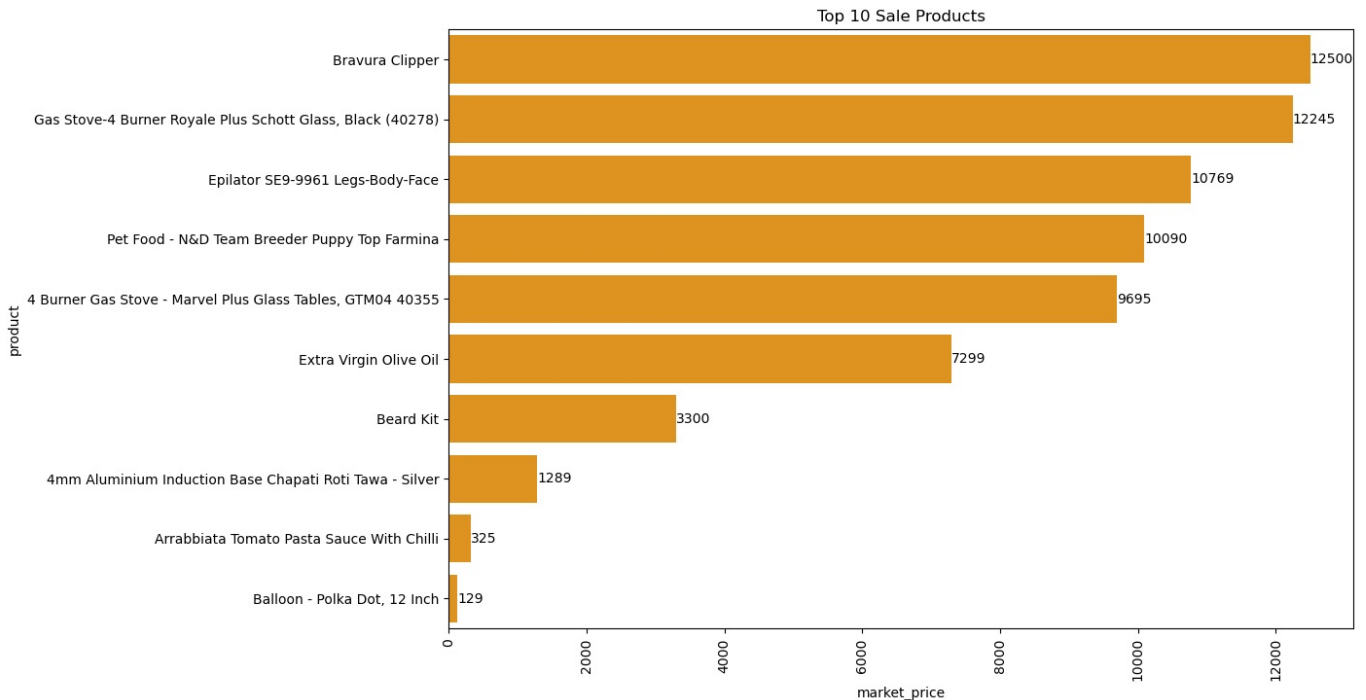

Key Takeaways:

Beard Kit: This is the clear top-selling product, with the highest sales volume among the 10 listed.

4mm Aluminium Induction Base Chapati Roti Tawa - Silver: This product ranks second in sales, indicating strong demand for kitchenware items.

Balloon - Polka Dot; 12 Inch: This product occupies the third position, suggesting popularity for party supplies or decorative items.

Arrabbiata Tomato Pasta Sauce With Chilli: This product has lower sales compared to the top three, suggesting a potential niche market or less demand for this specific flavor.

Bravura Clipper: This product has the lowest sales among the top 10, indicating it might be a less popular or niche product.

Recommendations:

Focus on Top Performers: Invest in marketing and promotions for the top-selling products like the Beard Kit and 4mm Aluminium Induction Base Chapati Roti Tawa to further boost their sales.

Analyze Underperformers: Investigate the reasons behind the low sales of products like Arrabbiata Tomato Pasta Sauce With Chilli and Bravura Clipper. Consider price adjustments, targeted marketing, or product improvements to increase their appeal.

Product Bundling: Create attractive product bundles combining top-selling items with underperforming ones to incentivize purchases and increase overall sales.

Customer Feedback: Gather customer feedback on the underperforming products to identify specific areas for improvement and tailor your strategies accordingly.

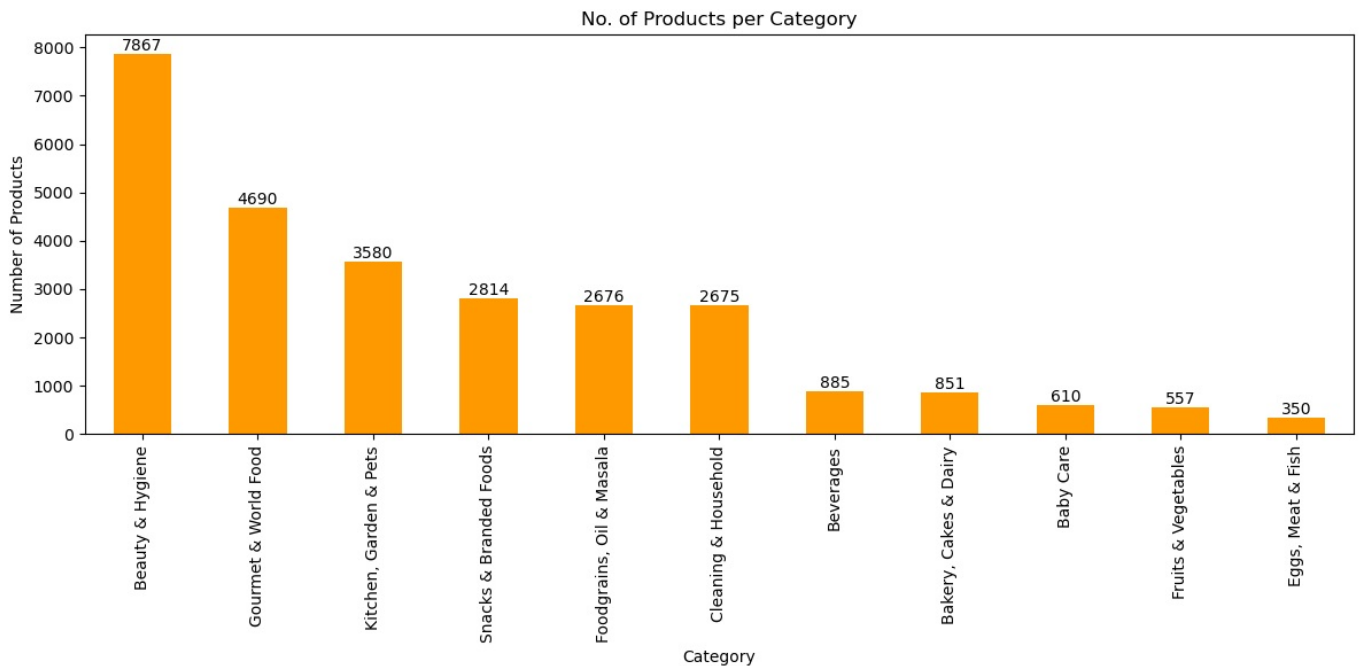### 2.Number of Products per Category

```
In [68]: # Count the number of products per category
         category_counts = df['category'].value_counts()

         # Plotting the number of products per category
         plt.figure(figsize=(12,6))
```

```
ax = category_counts.plot(kind='bar', color='#FE9900')
plt.title('No. of Products per Category')
plt.xlabel('Category')
plt.ylabel('Number of Products')
# Adding labels on top of the bars
for i in ax.patches: ax.text(i.get_x() + i.get_width() / 2, i.get_height() + 0.2,
                             f'{int(i.get_height())}', ha='center', va='bottom')

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


Observation

The distribution of products across categories is uneven, with some categories having a significantly larger number of products than others.

This uneven distribution might indicate different levels of demand or market size for different product categories.

It's important to consider the sales performance and profitability of each category to understand the overall business impact.

Recommendations:

Category Analysis: Analyze the sales and profit margins of each category to identify opportunities for improvement.

Product Portfolio Management: Review the product portfolio within each category to ensure a balanced mix of high-performing and low-performing items.

Targeted Marketing: Implement targeted marketing strategies for categories with high potential but lower product numbers.

3.Top 10 selling products by Volume

```
In [70]: counts = df['product'].value_counts()
         counts_p = pd.DataFrame({'product': counts.index, 'Counts': counts.values})[:10]
         bx = sns.barplot(x='product', y='Counts', data=counts_p, color='#FE9900')
         plt.xticks(rotation=75, ha='right')
         bx.set_xlabel('Products')
         bx.set_ylabel('No. of Products')
         bx.set_title('Top 10 selling products by Volume')
         for p in bx.patches:
             bx.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()
```

Top 10 selling products by Volume
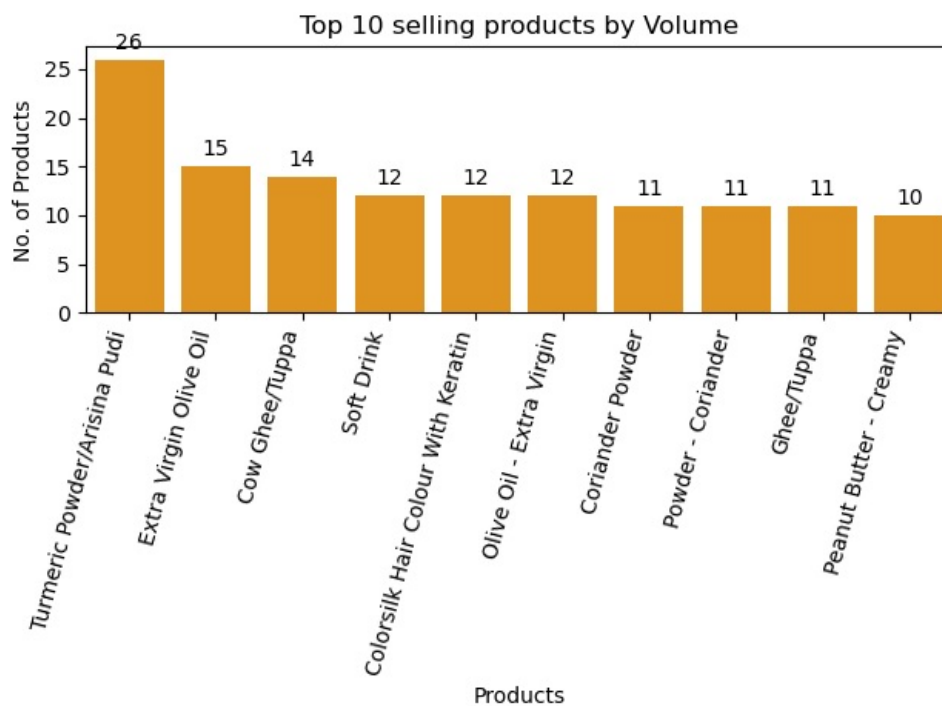
Observations:

The chart highlights a wide range of product categories, including food items, personal care products, and beverages.

The top-selling products are a mix of essential household items and popular consumer goods.

It's important to consider the sales volume in relation to the overall product portfolio and market demand.

Recommendations:

Inventory Management: Optimize inventory levels for the top-selling products to avoid stockouts and meet customer demand.

Marketing and Promotion: Continue to promote the top-selling products through effective marketing campaigns.

Customer Insights: Analyze customer purchasing behavior to identify trends and opportunities for cross-selling and upselling.

4.Which category sales more as a count and profit ?

```
In [80]: data = pd.DataFrame(df.groupby('category')['sale_price'].sum()).reset_index()
         data.columns = ['category', 'profit']
         sort_data = data.sort_values(by="profit",ascending=False)
         sort_data
```

Out[80]:

	category	profit
2	Beauty & Hygiene	3404749.24
9	Kitchen, Garden & Pets	1928014.76
8	Gourmet & World Food	1522115.31
4	Cleaning & Household	694141.73
6	Foodgrains, Oil & Masala	517060.55
10	Snacks & Branded Foods	364800.40
0	Baby Care	326317.17
3	Beverages	212186.67
1	Bakery, Cakes & Dairy	121525.14
5	Eggs, Meat & Fish	101114.12
7	Fruits & Vegetables	28345.36

```
In [90]: fig = plt.figure(figsize=(14,8))
         ax = fig.add_axes([0,0,1,1])
         ax = sns.barplot(data = sort_data, y = 'category', x = 'profit', color='#FE9900')
         plt.title('Categories Profit')
         plt.xlabel('Profit',fontsize=15, weight='semibold')
         plt.ylabel('Category',fontsize=15, weight='semibold')
         for container in ax.containers:
             ax.bar_label(container)
```

```
plt.show()
```


Categories Profit

Category	Profit
Beauty & Hygiene	3.40475e+06
Kitchen, Garden & Pets	1.92801e+06
Gourmet & World Food	1.52212e+06
Cleaning & Household	694142
Foodgrains, Oil & Masala	517061
Snacks & Branded Foods	364800
Baby Care	326317
Beverages	212187
Bakery, Cakes & Dairy	121525
Eggs, Meat & Fish	101114
Fruits & Vegetables	28345.4
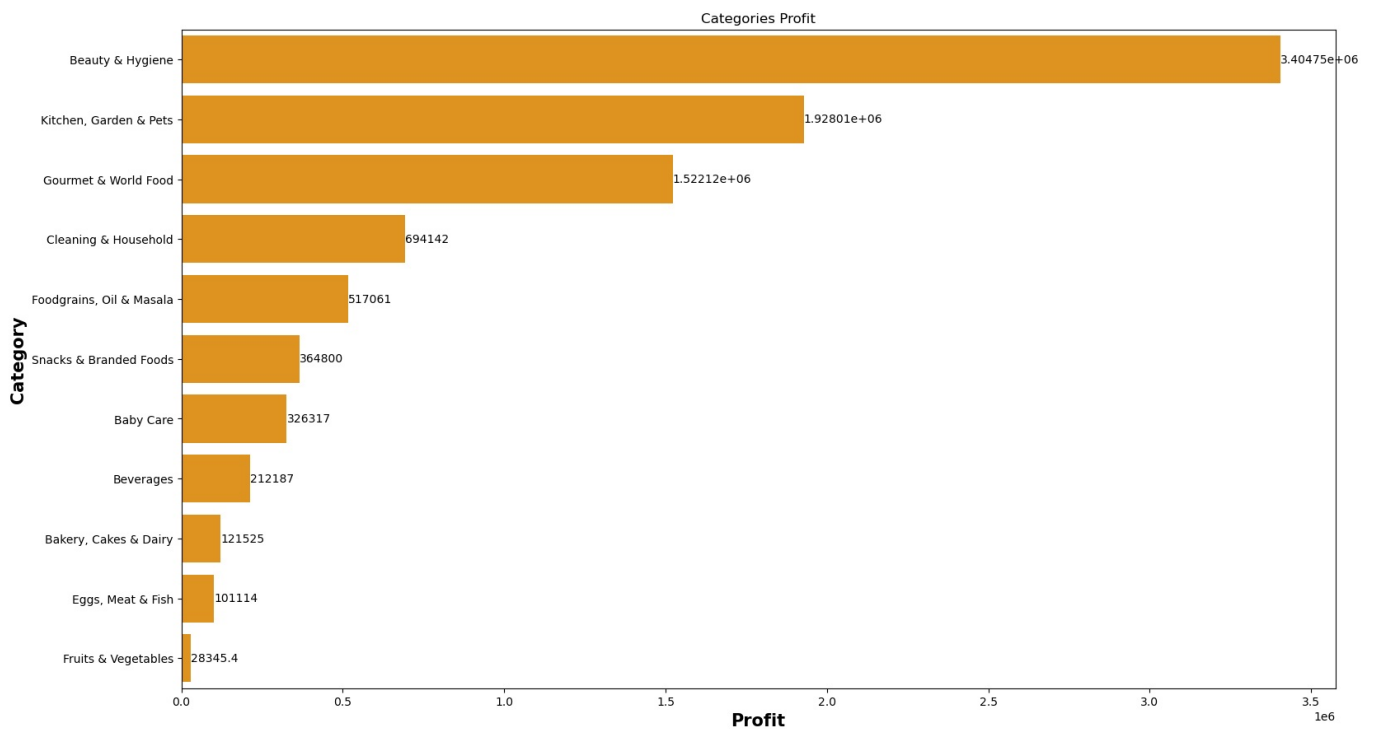
Observation

Highest Profit: The "Beauty & Hygiene" category has the highest profit, indicating strong sales or high margins in this c ory.

Moderate Profit: Categories such as "Gourmet & World Food" and "Snacks & Branded Foods" also show significant profits, though not as high as "Beauty Hygine".

Lowest Profit: The "Baby Care" and "Beverages" categories have the lowest profits, suggesting either lower sales or lower profit margins in are

**Category Spread**: The profits are spread across a range of categories, reflecting diversity in revnue# sources

- Your business might benefit from analyzing why "Beauty & Hygiene" outperforms other categories. Is it due to higher demand, better marketing, or h## i

pnts

- Conversely, you might want to investigate if there are opportunities to boost profits in lower-performing categories like "Baby Care" and ecific category!ecific category!

5. Top 5 Products By Rating

```
In [82]: top_5_products = df.nlargest(7, 'rating')[['product', 'rating']]
         # Plot histogram of ratings
         plt.figure(figsize=(8,6))
         plt.bar(top_5_products['product'], top_5_products['rating'],color="#FE9900")
         plt.title('Top 5 Products By Rating')
         plt.xlabel('Product')
         plt.ylabel('Rating')
         plt.xticks(rotation=45,ha='right')
         plt.tight_layout()
         plt.show()
```
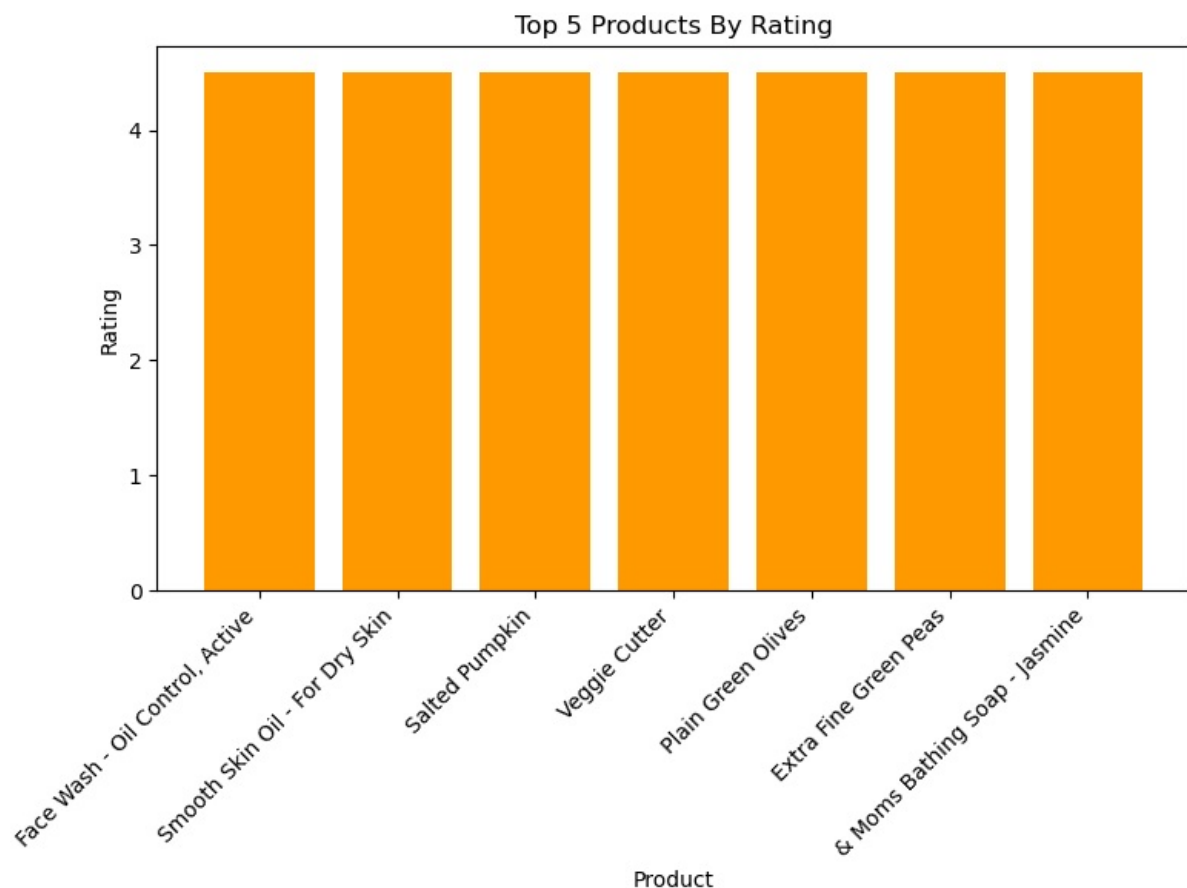
Top 5 Products By Rating

Observations:

The chart highlights the importance of customer reviews and ratings in assessing product performance.

Positive ratings can significantly impact sales and brand reputation.

It's important to maintain consistent quality and customer service to sustain high ratings.

Recommendations:

Customer Feedback: Continuously monitor customer feedback and reviews to identify areas for improvement.

Quality Control: Implement rigorous quality control measures to ensure product consistency.

Marketing and Promotion: Leverage positive ratings in marketing and promotional materials to attract new customers.

6.Top 25 Products By Sales

```python
top_25_products = df[['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="#FE9900")
ax.set_xlabel('Sale Price')
ax.set_ylabel('Product')
ax.set_title('Top 25 Products by Sale Price')
plt.tight_layout()
plt.show()
```

Top 25 Products by Sale Price
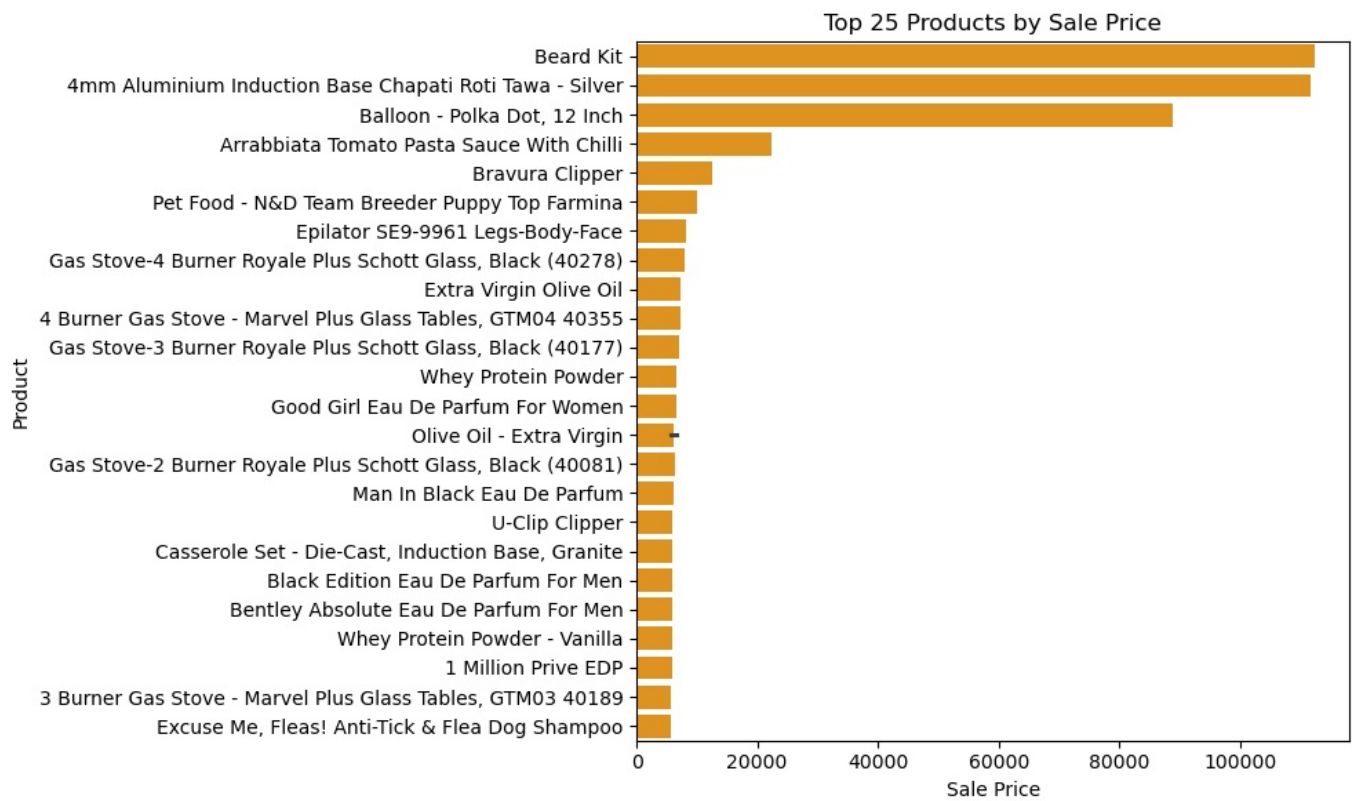
Observations:

The chart highlights the diversity of products offered, ranging from household appliances to personal care items and food products.

The distribution of prices suggests that the platform caters to a wide range of customer budgets and preferences.

It's important to consider the sales volume and profit margins of each product to understand the overall revenue impact.

Recommendations:

Pricing Strategy: Analyze the pricing strategy for high-priced items to ensure they are competitively priced and aligned with customer expectations.

Product Mix: Consider expanding the product mix to include more mid-priced items to attract a wider range of customers.

Marketing and Promotion: Implement targeted marketing campaigns to promote high-priced items and highlight their unique value proposition.

7.Average Rating of Top Brands

```python
brand_ratings = df.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='#FE9900')
plt.xticks(rotation=90)
plt.xlabel('Brand')
plt.ylabel('Average Rating')
plt.title('Average Rating of top Brands')
plt.tight_layout()
plt.show()
```
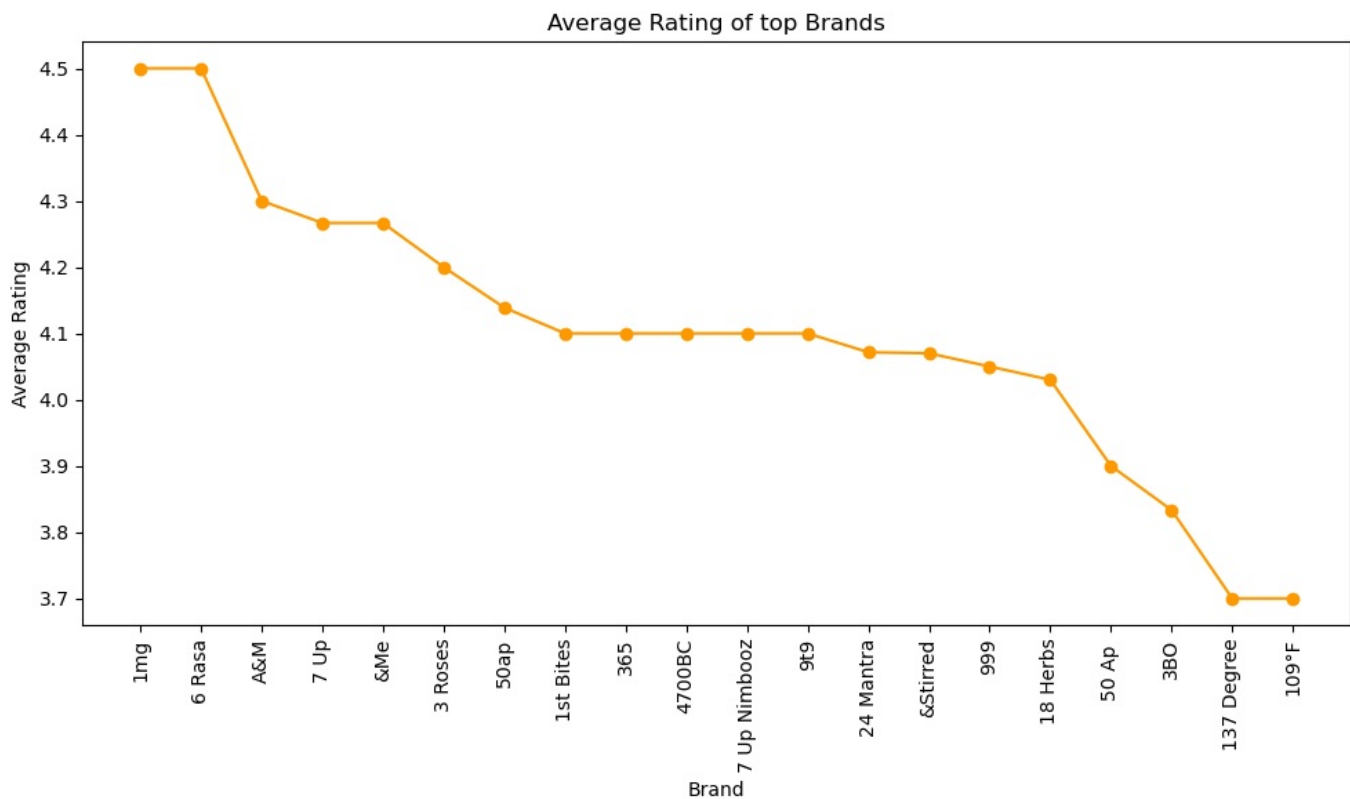
Average Rating of top Brands

Observations:

The chart highlights the importance of customer reviews and ratings in assessing brand performance.

Positive ratings can significantly impact brand reputation and customer trust.

It's important for brands to actively monitor and respond to customer feedback to maintain high ratings.

Recommendations:

Customer Feedback: Continuously monitor customer feedback and reviews to identify areas for improvement.

Quality Control: Implement rigorous quality control measures to ensure product consistency.

Marketing and Promotion: Leverage positive ratings in marketing and promotional materials to attract new customers.

Overall Conclusion

The analysis highlights a diverse range of products and categories, revealing disparities in sales performance and demand across product types. Top-selling products like Beard Kit and Curry Leaves suggest strong demand in specific niches, while underperforming items like Bravura Clipper and Arrabbiata Sauce point to potential issues in pricing, marketing, or market alignment. The skewed sales distribution toward lower-priced items emphasizes the need for a balanced pricing strategy and product mix. To drive growth, the focus should be on optimizing inventory, targeted marketing for high-potential categories, leveraging customer feedback, and promoting positive product ratings to enhance brand trust and visibility.

In [ ]:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js