

Big_ Basket Sales Analysis



Submitted by: Manish Sharma

Submitted to: Neeraj Sharma

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.

1. Loading Data

```
In [15]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
```

```
In [16]: # Loading CSV file in Pandas
df = pd.read_csv("C:/Users/DELL/Downloads/BigBasket Products.csv")
df
```

Out[16]:

	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	220.00	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	180.00	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	119.00	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	149.00	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	162.00	162.0	Bathing Bars & Soaps	4.4	Nivea Creme Soft Soap gives your skin the best...
...
27550	27551	Wottagirl! Perfume Spray - Heaven, Classic	Beauty & Hygiene	Fragrances & Deos	Layerr	199.20	249.0	Perfume	3.9	Layer brings you Wottagirl Classic fragrant b...
27551	27552	Rosemary	Gourmet & World Food	Cooking & Baking Needs	Puramate	67.50	75.0	Herbs, Seasonings & Rubs	4.0	Puramate rosemary is enough to transform a dis...
27552	27553	Peri-Peri Sweet Potato Chips	Gourmet & World Food	Snacks, Dry Fruits, Nuts	FabBox	200.00	200.0	Nachos & Chips	3.8	We have taken the richness of Sweet Potatoes (...)
27553	27554	Green Tea - Pure Original	Beverages	Tea	Tetley	396.00	495.0	Tea Bags	4.2	Tetley Green Tea with its refreshing pure, ori...
27554	27555	United Dreams Go Far Deodorant	Beauty & Hygiene	Men's Grooming	United Colors Of Benetton	214.53	390.0	Men's Deodorants	4.5	The new mens fragrance from the United Dreams ...

27555 rows × 10 columns

2. Using Head function to looking for first 12 rows

In [40]:

```
df.head(12)
```

Out[40]:

	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 thorough...
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...

2. Using Tail function to looking for first 12 rows

In [50]: df.tail(12)

Out[50]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description
27543	27544	Popcorn - French Butter & Pink Salt	Gourmet & World Food	Snacks, Dry Fruits, Nuts	4700BC	31.50	35.0	Gourmet Popcorn	4.1	High-quality mushroom corn popped in olive oil...
27544	27545	Brass Dhoop Aarti With Wooden Handle,No.3	Cleaning & Household	Pooja Needs	Trm	239.00	495.0	Other Pooja Needs	4.0	Dhoobakal is a vessel used for burning the Dho...
27545	27546	Toilet Cleaning Brush - Round With Holder (Big)	Cleaning & Household	Mops, Brushes & Scrubs	Liao	189.00	349.0	Toilet & Other Brushes	3.8	This round toilet brush is made up of virgin q...
27546	27547	Organic Powder - Garam Masala	Foodgrains, Oil & Masala	Organic Staples	Organic Tattva	152.00	160.0	Organic Masalas & Spices	4.2	Organic Tattva Garam masala is a famous spice ...
27547	27548	Powder - Baking	Snacks & Branded Foods	Ready To Cook & Eat	Kwality	38.00	38.0	Home Baking	3.7	Kwality Baking Powder Directions for use: Use ...
27548	27549	Apple Cider Vinegar Shampoo	Beauty & Hygiene	Hair Care	Morpheme Remedies	499.00	499.0	Shampoo & Conditioner	5.0	Say no to dull, lifeless, dry and damaged hair...
27549	27550	Papad - Garlic Disco	Snacks & Branded Foods	Ready To Cook & Eat	Atish	61.00	61.0	Papads, Ready To Fry	4.0	Papads are prepared from urad dal flour and sp...
27550	27551	Wottagirl! Perfume Spray - Heaven, Classic	Beauty & Hygiene	Fragrances & Deos	Layerr	199.20	249.0	Perfume	3.9	Layerr brings you Wottagirl Classic fragrant b...
27551	27552	Rosemary	Gourmet & World Food	Cooking & Baking Needs	Puramate	67.50	75.0	Herbs, Seasonings & Rubs	4.0	Puramate rosemary is enough to transform a dis...
27552	27553	Peri-Peri Sweet Potato Chips	Gourmet & World Food	Snacks, Dry Fruits, Nuts	FabBox	200.00	200.0	Nachos & Chips	3.8	We have taken the richness of Sweet Potatoes (...)
27553	27554	Green Tea - Pure Original	Beverages	Tea	Tetley	396.00	495.0	Tea Bags	4.2	Tetley Green Tea with its refreshing pure, ori...
27554	27555	United Dreams Go Far Deodorant	Beauty & Hygiene	Men's Grooming	United Colors Of Benetton	214.53	390.0	Men's Deodorants	4.5	The new mens fragrance from the United Dreams ...

3. Getting Description of the data in the DataFrame

In [52]: df.describe()

Out[52]:	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. Getting info about the Data Frame

```
In [59]: 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. Finding out top and amp: least sold products

```
In [101]: # Top 10 sold products
top_sold = df.sort_values(by='sale_price', ascending=False).head(10)
top_sold
```

Out[101]

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discount_per
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...	-6
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...	

In [97]:

```
# lest 10 sold product
least_sold = df.sort_values(by='sale_price').head(10)
least_sold
```

Out[97]:

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	discoun
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 on a certain item.

In [106]:

```
df['discount'] = ((df['market_price'] - df['sale_price']) / df['market_price']) * 100
print(df[['product', 'market_price', 'sale_price', 'discount']])
```

	product	market_price \
0	Garlic Oil - Vegetarian Capsule 500 mg	220.0
1	Water Bottle - Orange	180.0
2	Brass Angle Deep - Plain, No.2	250.0
3	Cereal Flip Lid Container/Storage Jar - Assort...	176.0
4	Creme Soft Soap - For Hands & Body	162.0
...
27550	Wottagirl! Perfume Spray - Heaven, Classic	249.0
27551	Rosemary	75.0
27552	Peri-Peri Sweet Potato Chips	200.0
27553	Green Tea - Pure Original	495.0
27554	United Dreams Go Far Deodorant	390.0

	sale_price	discount
0	220.00	0.000000
1	180.00	0.000000
2	119.00	52.400000
3	149.00	15.340909
4	162.00	0.000000
...
27550	199.20	20.000000
27551	67.50	10.000000
27552	200.00	0.000000
27553	396.00	20.000000
27554	214.53	44.992308

[27555 rows x 4 columns]

7. Find out the Missing Values from the Dataset..

```
In [91]: data.isna().sum()
```

```
Out[91]: index          0
product          1
category         0
sub_category     0
brand            1
sale_price       6
market_price     0
type            0
rating          8636
description      115
discount         6
dtype: int64
```

```
In [129]: # Finding Blank values
blank_values = (df==" ").sum()
print('Blank Values count:\n',blank_values)
```

```
Blank Values count:
index          0
product        0
category       0
sub_category   0
brand         0
sale_price    0
market_price  0
type         0
rating       0
description   0
discount_percent 0
discount     0
dtype: int64
```

```
In [131]: df[df.isna().any(axis=1)]
```


Out[131..

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description
55	56	Soothing Cucumber Facial Scrub With Apricot Seeds	Beauty & Hygiene	Skin Care	TJORI	299.4	499.0	Face Care	NaN	The soothing feel of cucumber meets the gentle...
59	60	Corporate Planner Diary With Premium PU Leathe...	Cleaning & Household	Stationery	Prozo Plus	399.0	399.0	Notebooks, Files, Folders	NaN	A5 Size (210x150mm) \n192 Pages Premium Natura...
65	66	Ayurvedic Anti-Tan Face Pack	Beauty & Hygiene	Skin Care	TJORI	269.4	449.0	Face Care	NaN	A nourishing face pack that removes tan and br...
68	69	Organic Carom Seeds/Ajwain/Om Kalu	Foodgrains, Oil & Masala	Masalas & Spices	Earthon	72.0	72.0	Whole Spices	NaN	Earthon's Ajwain is Best quality, organically ...
69	70	Padded Harness - 3/4 inch, Grey Colour	Kitchen, Garden & Pets	Pet Food & Accessories	Glenand	840.0	840.0	Pet Collars & Leashes	NaN	These are soft padded harness for your active ...
...
27509	27510	Deluxe Crackers - Veg	Gourmet & World Food	Chocolates & Biscuits	Kerk	150.0	150.0	Cookies, Biscotti, Wafer	NaN	Kerk Biscuits-has been a household name synony...
27511	27512	Specialist Stain Remover Pen & Marker	Cleaning & Household	All Purpose Cleaners	365	449.0	449.0	Imported Cleaners	NaN	Mightier than the pen. The pen may be mightier...
27514	27515	Verge & Sheer Perfume For Pair	Beauty & Hygiene	Fragrances & Deos	Skin by Titan	1615.5	1795.0	Perfume	NaN	VERGE for men paints a picture of a classy out...
27530	27531	Tick'et to Fleadom Dry Shampoo For Dogs	Kitchen, Garden & Pets	Pet Food & Accessories	Captain Zack	99.0	99.0	Pet Cleaning & Grooming	NaN	1) No Rinse Defence Against Ticks and Fleas: C...
27531	27532	Vanilla Ice Cream	Gourmet & World Food	Dairy & Cheese	Haagen-Dazs	225.0	225.0	Gourmet Ice Cream	NaN	Some things are so good in their simplest form...

8728 rows × 12 columns

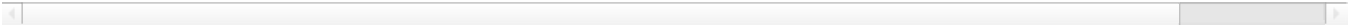


In [141..

```
# Checking NaN value in Product column
df[df['product'].isna()]
```

Out[141..

	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



In [143..

```
# Handling NaN value with "Unknown"
df.loc[df['product'].isna(),'product']='Unknown'
df['product']
```

Out[143..

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 [147..

```
# Checking NaN value in Brand column
df[df['brand'].isna()]
```

```

Out[147...
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

In [149...
# Handling NaN value with "Unknown"
df.loc[df['brand'].isna(), 'brand'] = 'Unknown'
df['brand']

Out[149...
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 [157...
# Checking NaN value in Sale_Price column
df[df['sale_price'].isna()]

Out[157...
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...

In [175...
# Handling NaN value with "Median"
df['sale_price'] = df['sale_price'].fillna(df['sale_price'].median())

In [177...
df.isna().sum()

```

Out[177... index 0
product 0
category 0
sub_category 0
brand 0
sale_price 0
market_price 0
type 0
rating 8636
description 115
discount_percentage 6
discount 6
dtype: int64

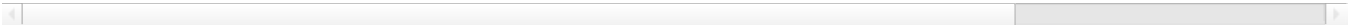
```
In [181... # Checking NaN value in rating column
df[df['rating'].isna()]
```

Out[181...

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	description	
	55	56	Soothing Cucumber Facial Scrub With Apricot Seeds	Beauty & Hygiene	Skin Care	TJORI	299.4	499.0	Face Care	NaN	The soothing feel of cucumber meets the gentle...
	59	60	Corporate Planner Diary With Premium PU Leather...	Cleaning & Household	Stationery	Prozo Plus	399.0	399.0	Notebooks, Files, Folders	NaN	A5 Size (210x150mm) \n192 Pages Premium Natural...
	65	66	Ayurvedic Anti-Tan Face Pack	Beauty & Hygiene	Skin Care	TJORI	269.4	449.0	Face Care	NaN	A nourishing face pack that removes tan and br...
	68	69	Organic Carom Seeds/Ajwain/Om Kalu	Foodgrains, Oil & Masala	Masalas & Spices	Earthon	72.0	72.0	Whole Spices	NaN	Earthon's Ajwain is Best quality, organically ...
	69	70	Padded Harness - 3/4 inch, Grey Colour	Kitchen, Garden & Pets	Pet Food & Accessories	Glenand	840.0	840.0	Pet Collars & Leashes	NaN	These are soft padded harness for your active ...

	27509	27510	Deluxe Crackers - Veg	Gourmet & World Food	Chocolates & Biscuits	Kerk	150.0	150.0	Cookies, Biscotti, Wafer	NaN	Kerk Biscuits-has been a household name synony...
	27511	27512	Specialist Stain Remover Pen & Marker	Cleaning & Household	All Purpose Cleaners	365	449.0	449.0	Imported Cleaners	NaN	Mightier than the pen. The pen may be mightier...
	27514	27515	Verge & Sheer Perfume For Pair	Beauty & Hygiene	Fragrances & Deos	Skinn by Titan	1615.5	1795.0	Perfume	NaN	VERGE for men paints a picture of a classy out...
	27530	27531	Tick'et to Fleadom Dry Shampoo For Dogs	Kitchen, Garden & Pets	Pet Food & Accessories	Captain Zack	99.0	99.0	Pet Cleaning & Grooming	NaN	1) No Rinse Defence Against Ticks and Fleas: C...
	27531	27532	Vanilla Ice Cream	Gourmet & World Food	Dairy & Cheese	Haagen-Dazs	225.0	225.0	Gourmet Ice Cream	NaN	Some things are so good in their simplest form...

8636 rows × 12 columns



```
In [183... # Handling NaN value with "Median"
df['rating'] = df['rating'].fillna(df['rating'].median())
```

```
In [191... # Checking NaN value in description column
df[df['description'].isna()]
```

Out[191...

	index	product	category	sub_category	brand	sale_price	market_price	type	rating	de
288	289	Arrabbiata Tomato Pasta Sauce With Chilli	Gourmet & World Food	Sauces, Spreads & Dips	Montanini	22325.00	325.0	Mustard & Cheese Sauces	5.0	
480	481	Pumice Stone & Emery Foot Filer - FFL3407, Col...	Beauty & Hygiene	Bath & Hand Wash	Kaiv	175.00	175.0	Bathing Accessories	3.8	
780	781	Peanut Butter - Chunky	Gourmet & World Food	Sauces, Spreads & Dips	American Garden	299.00	299.0	Chocolate, Peanut Spread	4.1	
845	846	Sarasaparilla/Nannari/Sugandhi Root	Beverages	Fruit Juices & Drinks	NaturoBell	140.00	140.0	Syrups & Concentrates	2.5	
1047	1048	Natura Rice & Walnut Drink, Vegan	Gourmet & World Food	Drinks & Beverages	BORGES	275.01	349.0	Health Drinks	4.5	
...
25510	25511	BBPopular Almond/Badam Californian 500gm + BBR...	Foodgrains, Oil & Masala	Dry Fruits	bb Combo	443.00	725.0	Almonds	4.1	
25583	25584	Chai/Chai/Tea/ Coffee Bone China Microwave Ser...	Kitchen, Garden & Pets	Crockery & Cutlery	BP Bharat	269.00	373.0	Cups, Mugs & Tumblers	2.5	
26116	26117	Tea Strainer - Stainless Steel, Medium	Kitchen, Garden & Pets	Kitchen Accessories	Ritu	99.00	135.0	Strainer, Ladle, Spatula	4.3	
26835	26836	Cold Pressed Virgin Groundnut Oil	Foodgrains, Oil & Masala	Edible Oils & Ghee	Chekko	325.00	325.0	Groundnut Oil	4.2	
27303	27304	Zipper Storage - Gallon	Cleaning & Household	Disposables, Garbage Bag	Glad	350.00	350.0	Aluminium Foil, Clingwrap	4.7	

115 rows × 12 columns

In [193...

```
# Handling NaN value with "Unknown"
df.loc[df['description'].isna(),'description']='Unknown'
df['description']
```

Out[193...

```
0      This Product contains Garlic Oil that is known...
1      Each product is microwave safe (without lid), ...
2      A perfect gift for all occasions, be it your m...
3      Multipurpose container with an attractive desi...
4      Nivea Creme Soft Soap gives your skin the best...
...
27550   Layerr brings you Wottagirl Classic fragrant b...
27551   Puramate rosemary is enough to transform a dis...
27552   We have taken the richness of Sweet Potatoes (...
27553   Tetley Green Tea with its refreshing pure, ori...
27554   The new mens fragrance from the United Dreams ...
Name: description, Length: 27555, dtype: object
```

In [199...

```
# Checking NaN value in discount_percentage column
df[df['discount_percentage'].isna()]
```

Out[199..	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	190.32	20.0	Camphor & Wicks	4.1	Jaji Puja Flower Batti / Puvvu Vathulu is made...
	1720	1721	Powder - Sambar	Foodgrains, Oil & Masala	Masalas & Spices	Annapoorna	190.32	46.0	Blended Masalas	4.1	Annaporna Spices are the preserve for Connoiss...
	1721	1722	Steel Fork - Medium, Premium Excel Series, BBCL08	Kitchen, Garden & Pets	Crockery & Cutlery	BB Home	190.32	299.0	Cutlery, Spoon & Fork	4.1	BB Home provides fine and classy cutlery that ...
	1722	1723	Snack Mix - Dhokla	Snacks & Branded Foods	Ready To Cook & Eat	MTR	190.32	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	190.32	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	190.32	32.0	Toilet Paper	3.9	Klassic Plain Cocktail Napkins (22 X 22 cm) 10...

```
In [201.. # Handling NaN value with "Median"
df['discount_percentage'] = df['discount_percentage'].fillna(df['discount_percentage'].median())
```

```
In [203.. df.isna().sum()
```

Out[203..	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
	discount	6
	dtype: int64	

```
In [205.. # Checking NaN value in discount column
df[df['discount'].isna()]
```

Out [205...	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	190.32	20.0	Camphor & Wicks	4.1	Jaji Puja Flower Batti / Puvvu Vathulu is made...
	1720	1721	Powder - Sambar	Foodgrains, Oil & Masala	Masalas & Spices	Annapoorna	190.32	46.0	Blended Masalas	4.1	Annaporna Spices are the preserve for Connoiss...
	1721	1722	Steel Fork - Medium, Premium Excel Series, BBCL08	Kitchen, Garden & Pets	Crockery & Cutlery	BB Home	190.32	299.0	Cutlery, Spoon & Fork	4.1	BB Home provides fine and classy cutlery that ...
	1722	1723	Snack Mix - Dhokla	Snacks & Branded Foods	Ready To Cook & Eat	MTR	190.32	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	190.32	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	190.32	32.0	Toilet Paper	3.9	Klassic Plain Cocktail Napkins (22 X 22 cm) 10...

```
In [207... # Handling NaN value with "Median"
df['discount'] = df['discount'].fillna(df['discount'].median())
```

```
In [209... df.isna().sum()
```

```
Out[209... 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
discount        0
dtype: int64
```

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

```
In [237... # calculate the IQR
Q1=df['sale_price'].quantile(0.25)
Q3=df['sale_price'].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: 95.0
Q3: 359.0
IQR: 264.0
lower: -301.0
upper: 755.0

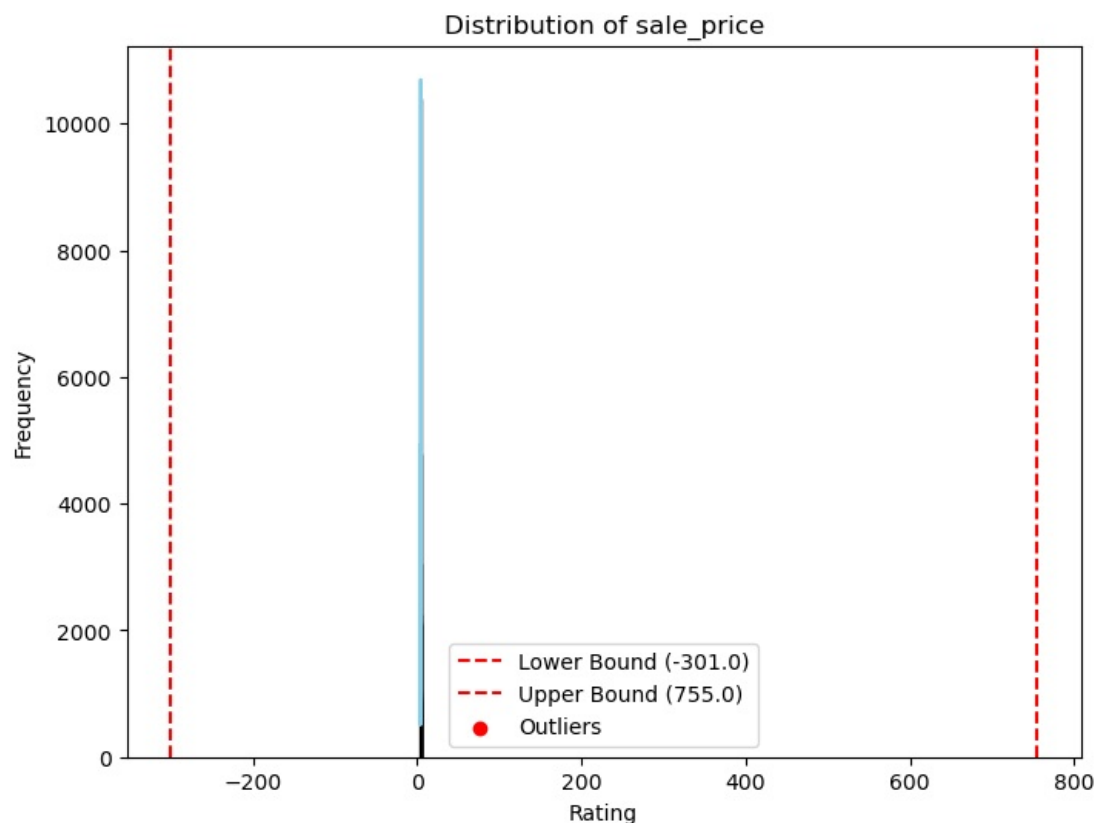
```
In [239].. # Create a histogram to visualize the distribution of the data
plt.figure(figsize=(8, 6))
sns.histplot(df['rating'], bins=10, kde=True, color='skyblue', edgecolor='black')

# Add lines for the lower and upper bounds
plt.axvline(x=lower_bound, color='red', linestyle='--', label=f'Lower Bound ({lower_bound})')
plt.axvline(x=upper_bound, color='red', linestyle='--', label=f'Upper Bound ({upper_bound})')

# Highlight the outliers
outlier_values = df[(df['rating'] < lower_bound) | (df['rating'] > upper_bound)]['rating']
plt.scatter(outlier_values, np.zeros_like(outlier_values), color='red', label='Outliers', zorder=5)

# Add title and labels
plt.title('Distribution of sale_price')
plt.xlabel('Rating')
plt.ylabel('Frequency')
plt.legend()

# Show the plot
plt.show()
```

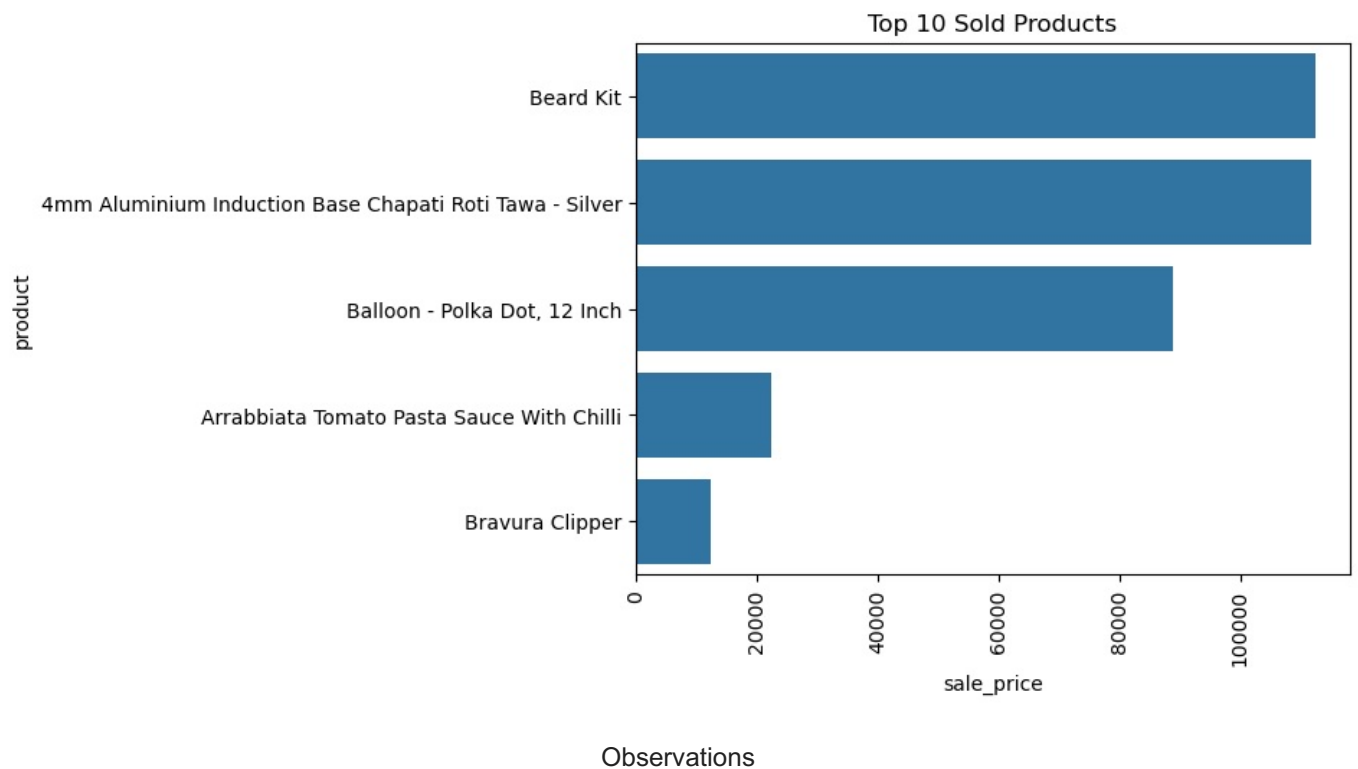


Create Plots or visualizations.

9. Create Plots or visualizations.

```
In [99]: import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [101].. # Bar plot of top 10 sold products
sns.barplot(x='sale_price', y='product', data=top_sold)
plt.xticks(rotation=90)
plt.title("Top 10 Sold Products")
plt.show()
```



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:

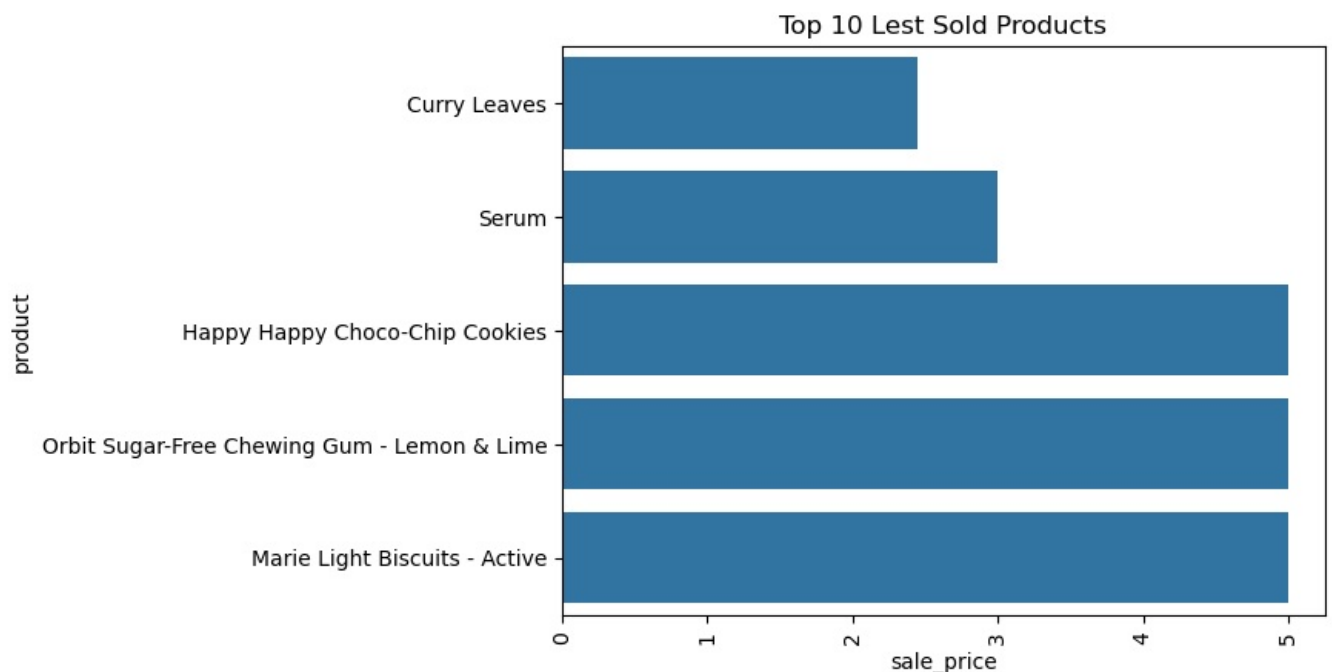
Focus on Top Performers: Invest in marketing and promotions for the top-selling products like the Beard Kit and 4mm Aluminium Induction Base Chapati Roti Tawa to further boost their sales.

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.

```
In [107...] sns.barplot(x='sale_price', y='product', data=least_sold)
plt.xticks(rotation=90)
plt.title("Top 10 Lest Sold Products")
plt.show()
```

Observations:

Product Performance:

Marie Light Biscuits - Active and Orbit Sugar-Free Chewing Gum - Lemon & Lime are the least sold products, with sales significantly lower than the others.

Happy Happy Choco-Chip Cookies and Serum have slightly higher sales compared to the top two least sold products.

Curry Leaves is the most sold product among the top 10 least sold items.

Potential Insights

Product Category: The chart includes a mix of food items (biscuits, cookies, gum, curry leaves) and a non-food item (serum). This suggests that the underperformance might not be solely due to product category.

Price Point: It's important to consider the pricing of these products compared to similar items. Lower-priced items might have lower sales due to lower profit margins.

Marketing and Promotion: The lack of effective marketing and promotion for these products could be contributing to their low sales.

Customer Preferences: It's possible that these products don't align with current customer preferences or trends.

Recommendations:

Targeted Marketing: Implement targeted marketing campaigns to promote these products to specific customer segments.

Price Optimization: Analyze pricing strategies and consider adjustments to make these products more competitive.

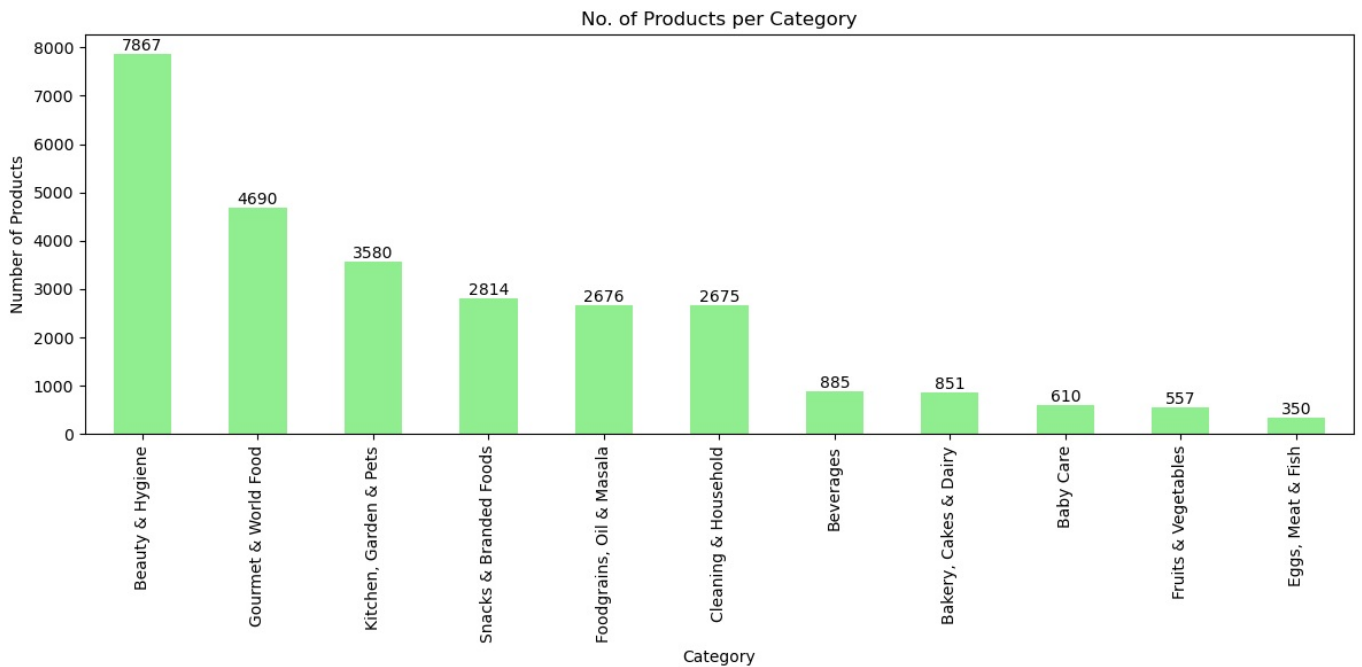
Product Bundling: Create attractive product bundles that include these items to encourage sales.

Customer Feedback: Gather customer feedback to understand why these products are not performing well and make necessary improvements.

```
In [257... # 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='lightgreen')
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()
```



Observations:

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.

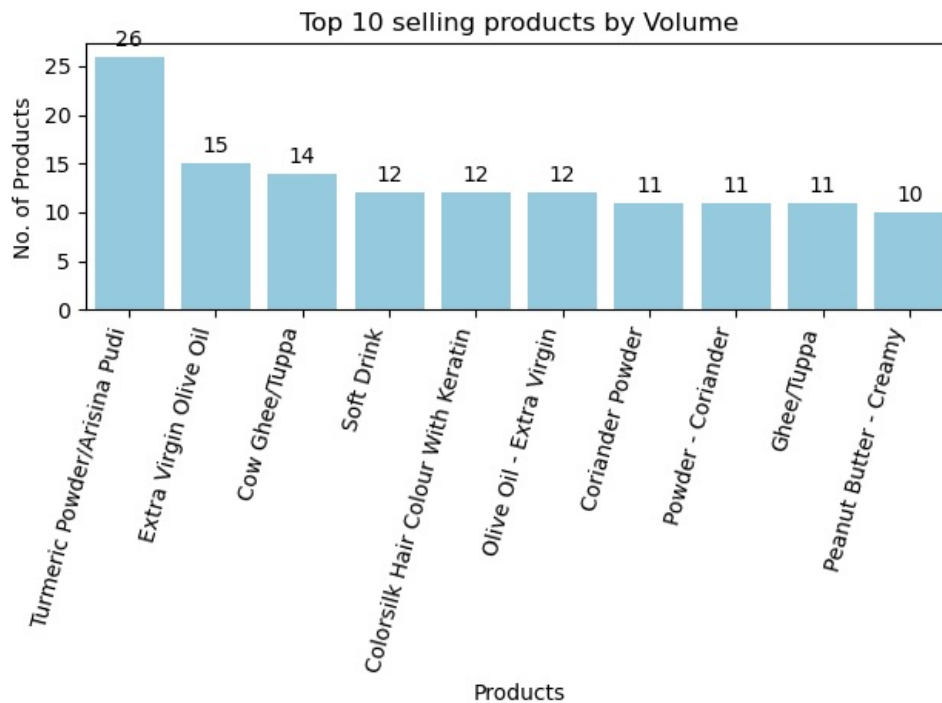
Recommendations:

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.

```
In [266]: 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='skyblue')
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 point')
plt.tight_layout()
plt.show()
```



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.

```
In [272]: # Plot histogram of sale prices
plt.figure(figsize=(8,6))
plt.hist(df['sale_price'], bins=20, color='green', edgecolor='black')
plt.title('Distribution of Sale Prices')
plt.xlabel('Sale Price')
plt.ylabel('Frequency')
plt.grid(True)
plt.show()
```



Observations:

The distribution is right-skewed, suggesting that most sales are for lower-priced items.

This pattern is common in retail settings where a large portion of sales comes from lower-priced products.

The distribution might be influenced by factors such as product mix, target market, and pricing strategy.

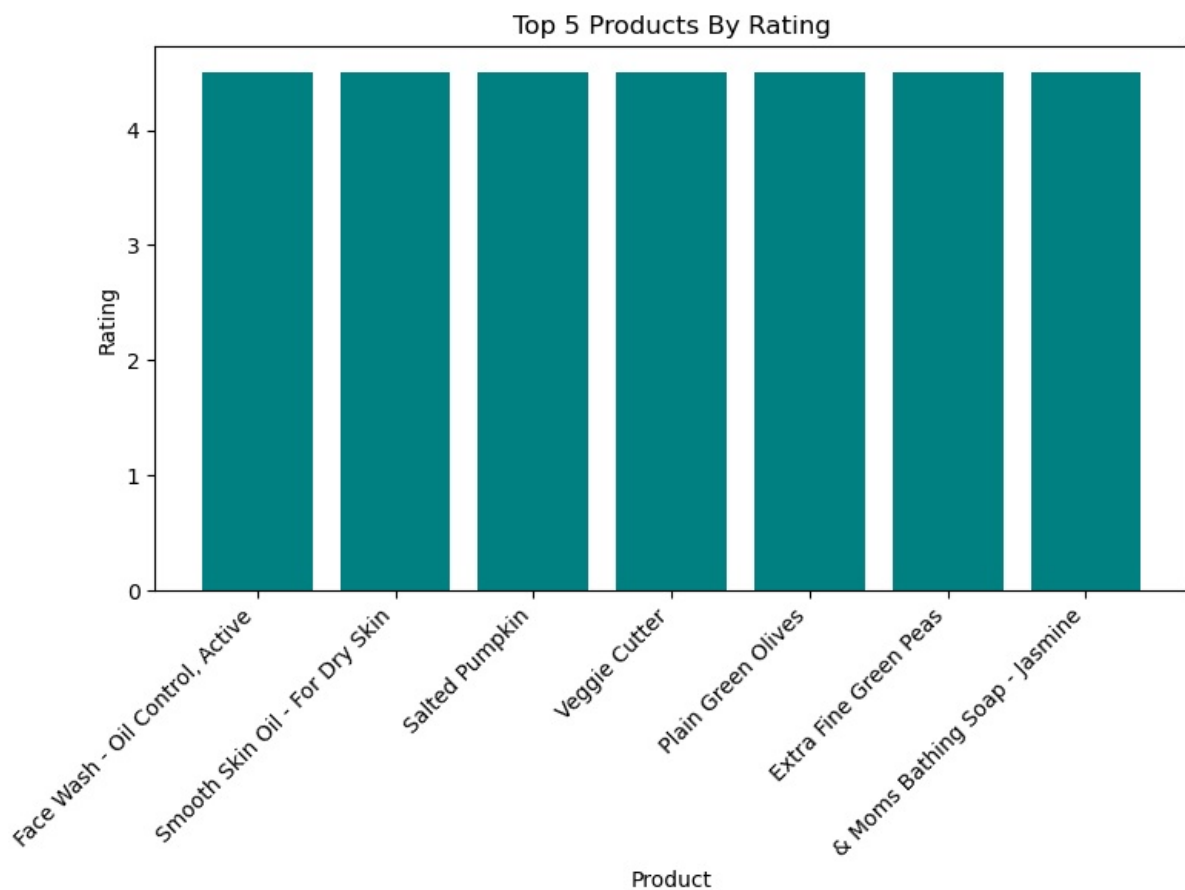
Recommendations:

Pricing Strategy: Analyze the pricing strategy for high-priced items to identify potential adjustments that could increase sales.

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.

```
In [275...] 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="#008080")
plt.title('Top 5 Products By Rating')
plt.xlabel('Product')
plt.ylabel('Rating')
plt.xticks(rotation=45,ha='right')
plt.tight_layout()
plt.show()
```



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.

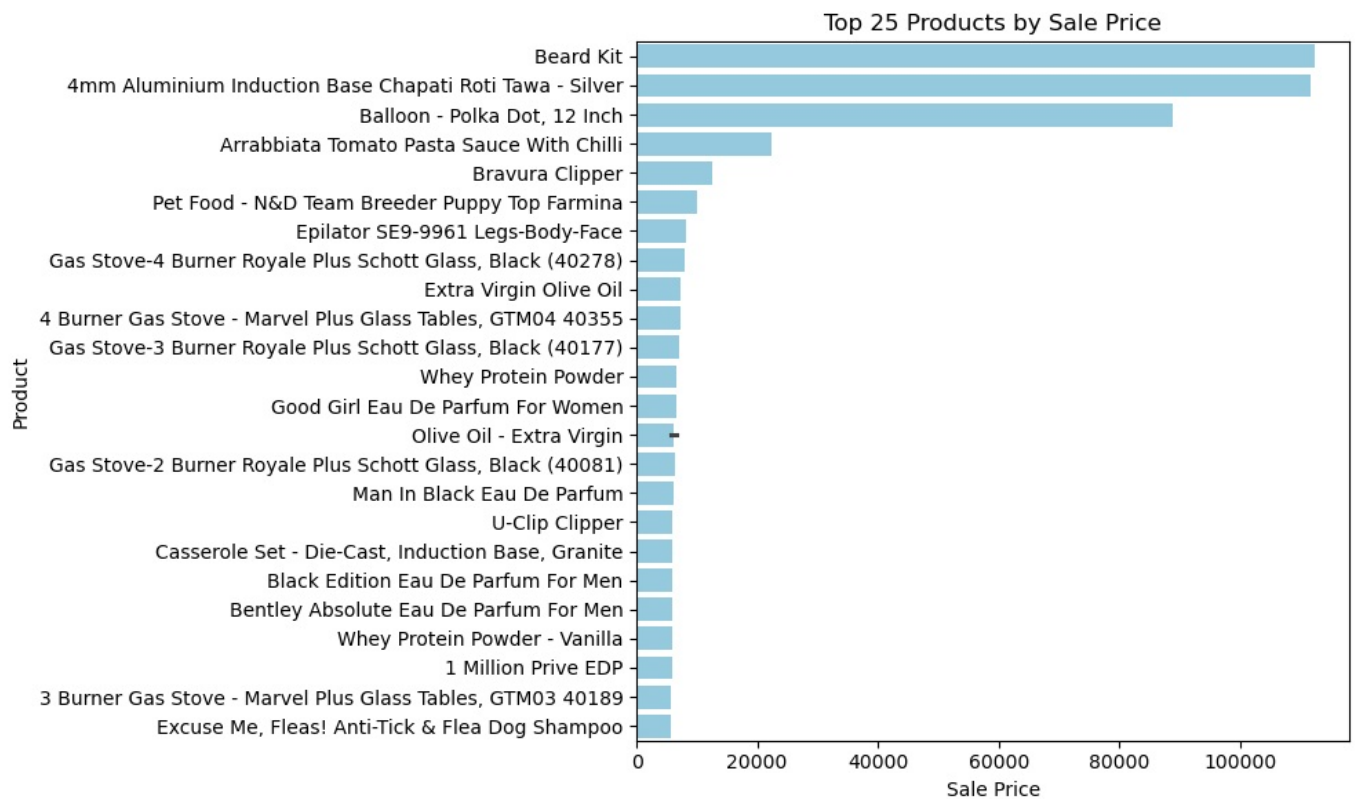
Recommendations:

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.

```
In [280]: 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='skyblue')  
ax.set_xlabel('Sale Price')  
ax.set_ylabel('Product')  
ax.set_title('Top 25 Products by Sale Price')  
plt.tight_layout()  
plt.show()
```



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.

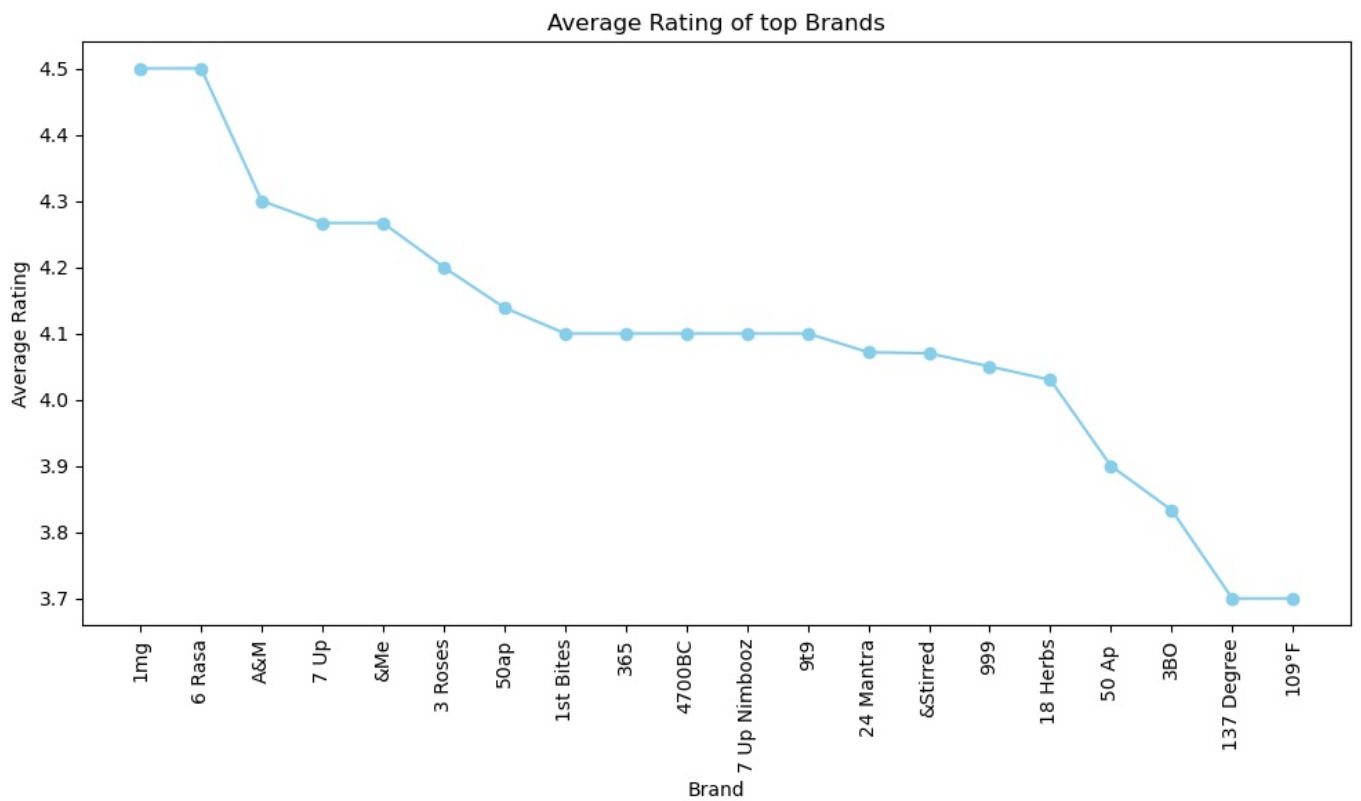
Recommendations:

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.

```
In [283.. 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='skyblue')
plt.xticks(rotation=90)
plt.xlabel('Brand')
plt.ylabel('Average Rating')
plt.title('Average Rating of top Brands')
plt.tight_layout()
plt.show()
```



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 []: