

BIGBASKET PROJECT

STEPS TO FOLLOW

Step 1: Load DataSet.

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

Step 3: Get Description of the data in the DataFrame.

Step 4: Find Information about the DataFrame.

Step 5: Find out Top & least sold products.

Step 6: Measuring discount on a certain item.

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

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

Step 9: Create Plots or visualizations.

Importing libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Step 1: Load DataSet.

```
df=pd.read_csv("C:/Users/hsr49/OneDrive/Desktop/BigBasket
Products.csv")
```

df

	index	product \
0	1	Garlic Oil - Vegetarian Capsule 500 mg
1	2	Water Bottle - Orange
2	3	Brass Angle Deep - Plain, No.2
3	4	Cereal Flip Lid Container/Storage Jar - Assort...
4	5	Creame Soft Soap - For Hands & Body
...
27550	27551	Wottagirl! Perfume Spray - Heaven, Classic

27551	27552	Rosemary
27552	27553	Peri-Peri Sweet Potato Chips
27553	27554	Green Tea - Pure Original
27554	27555	United Dreams Go Far Deodorant

	category	sub_category \
0	Beauty & Hygiene	Hair Care
1	Kitchen, Garden & Pets	Storage & Accessories
2	Cleaning & Household	Pooja Needs
3	Cleaning & Household	Bins & Bathroom Ware
4	Beauty & Hygiene	Bath & Hand Wash
...
27550	Beauty & Hygiene	Fragrances & Deos
27551	Gourmet & World Food	Cooking & Baking Needs
27552	Gourmet & World Food	Snacks, Dry Fruits, Nuts
27553	Beverages	Tea
27554	Beauty & Hygiene	Men's Grooming

	brand	sale_price	market_price \
0	Sri Sri Ayurveda	220.00	220.0
1	Mastercook	180.00	180.0
2	Trm	119.00	250.0
3	Nakoda	149.00	176.0
4	Nivea	162.00	162.0
...
27550	Layerr	199.20	249.0
27551	Puramate	67.50	75.0
27552	FabBox	200.00	200.0
27553	Tetley	396.00	495.0
27554	United Colors Of Benetton	214.53	390.0

	type	rating \
0	Hair Oil & Serum	4.1
1	Water & Fridge Bottles	2.3
2	Lamp & Lamp Oil	3.4
3	Laundry, Storage Baskets	3.7
4	Bathing Bars & Soaps	4.4
...
27550	Perfume	3.9
27551	Herbs, Seasonings & Rubs	4.0
27552	Nachos & Chips	3.8
27553	Tea Bags	4.2
27554	Men's Deodorants	4.5

	description
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 ...

[27555 rows x 10 columns]

```

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

```
df["product"].value_counts().head(12)
```

```

product
Turmeric Powder/Arisina Pudi      26
Extra Virgin Olive Oil            15
Cow Ghee/Tuppa                   14
Soft Drink                        12
Colorsilk Hair Colour With Keratin 12
Olive Oil - Extra Virgin          12
Coriander Powder                  11
Powder - Coriander                11
Ghee/Tuppa                       11
Peanut Butter - Creamy            10
Casting Creme Gloss Hair Color    10
Hand Sanitizer                    10
Name: count, dtype: int64

```

Step 3: Get Description of the data in the DataFrame.

```
df.describe()
```

	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.

```
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 & least sold products.

```
#Top Five selling product
```

```
df["product"].value_counts().head(5)
```

```
product
Turmeric Powder/Arisina Pudi      26
Extra Virgin Olive Oil            15
Cow Ghee/Tuppa                   14
Soft Drink                       12
Colorsilk Hair Colour With Keratin 12
Name: count, dtype: int64
```

```
#Least Five selling product
```

```
df["product"].value_counts().tail(5)
```

```
product
Pepper & Herb Salami Chicken      1
Nutmeg Powder                     1
Disney Mickey Mouse Plastic Kids Sipper Bottle - Coolest Dude  1
Chocolates-Roasted Peanut Chocolate 1
Green Tea - Pure Original          1
Name: count, dtype: int64
```

Step 6: Measuring discount on a certain item.

```
df["discount"]=(df["market_price"]-df["sale_price"]/
df["market_price"]*100).round(2)
```

df

	index	product \
0	1	Garlic Oil - Vegetarian Capsule 500 mg
1	2	Water Bottle - Orange
2	3	Brass Angle Deep - Plain, No.2
3	4	Cereal Flip Lid Container/Storage Jar - Assort...
4	5	Creme Soft Soap - For Hands & Body
...
27550	27551	Wottagirl! Perfume Spray - Heaven, Classic
27551	27552	Rosemary
27552	27553	Peri-Peri Sweet Potato Chips
27553	27554	Green Tea - Pure Original
27554	27555	United Dreams Go Far Deodorant

	category	sub_category \
0	Beauty & Hygiene	Hair Care
1	Kitchen, Garden & Pets	Storage & Accessories
2	Cleaning & Household	Pooja Needs
3	Cleaning & Household	Bins & Bathroom Ware
4	Beauty & Hygiene	Bath & Hand Wash
...
27550	Beauty & Hygiene	Fragrances & Deos
27551	Gourmet & World Food	Cooking & Baking Needs
27552	Gourmet & World Food	Snacks, Dry Fruits, Nuts
27553	Beverages	Tea
27554	Beauty & Hygiene	Men's Grooming

	brand	sale_price	market_price \
0	Sri Sri Ayurveda	220.00	220.0
1	Mastercook	180.00	180.0
2	Trm	119.00	250.0
3	Nakoda	149.00	176.0
4	Nivea	162.00	162.0
...
27550	Layerr	199.20	249.0
27551	Puramate	67.50	75.0
27552	FabBox	200.00	200.0
27553	Tetley	396.00	495.0
27554	United Colors Of Benetton	214.53	390.0

	type	rating \
0	Hair Oil & Serum	4.1
1	Water & Fridge Bottles	2.3
2	Lamp & Lamp Oil	3.4

3	Laundry, Storage Baskets	3.7
4	Bathing Bars & Soaps	4.4
...
27550	Perfume	3.9
27551	Herbs, Seasonings & Rubs	4.0
27552	Nachos & Chips	3.8
27553	Tea Bags	4.2
27554	Men's Deodorants	4.5

	description	discount
0	This Product contains Garlic Oil that is known...	120.00
1	Each product is microwave safe (without lid), ...	80.00
2	A perfect gift for all occasions, be it your m...	202.40
3	Multipurpose container with an attractive desi...	91.34
4	Nivea Creme Soft Soap gives your skin the best...	62.00
...
27550	Layerr brings you Wottagirl Classic fragrant b...	169.00
27551	Puramate rosemary is enough to transform a dis...	-15.00
27552	We have taken the richness of Sweet Potatoes (...)	100.00
27553	Tetley Green Tea with its refreshing pure, ori...	415.00
27554	The new mens fragrance from the United Dreams ...	334.99

[27555 rows x 11 columns]

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

```
df.isnull().sum()
```

```
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
```

```
df[df["product"].isnull()]
```

	index	product	category	sub_category	brand
sale_price \					
14363	14364	NaN	Beverages	Coffee	Cothas Coffee

200.0

	market_price		type	rating	\
14363	240.0	Ground Coffee		4.2	

			description	discount
14363	Cothas	Specialty Blend Coffee and Chicory incl...		156.67

df[df["brand"].isnull()]

	index		product		category	\
9765	9766	Food Package - Medium	Cleaning & Household			

		sub_category	brand	sale_price	market_price	\
9765	Disposables, Garbage Bag		NaN	50.0	50.0	

			type	rating	description	discount
9765	Aluminium Foil, Clingwrap			NaN	NaN	-50.0

df[df["sale_price"].isnull()]

	index		product	\
1719	1720	Puja Flower Wicks - Puvvu Vathulu Batti		
1720	1721	Powder - Sambar		
1721	1722	Steel Fork - Medium, Premium Excel Series, BBCL08		
1722	1723	Snack Mix - Dhokla		
2401	2402	Battery AA 3UT Hi Top		
2402	2403	Klassic Plain Cocktail Napkins (22 X 22 cm)		

		category	sub_category	
brand	\			
1719	Cleaning & Household		Pooja Needs	Jaji
1720	Foodgrains, Oil & Masala		Masalas & Spices	Annapoorna
1721	Kitchen, Garden & Pets		Crockery & Cutlery	BB Home
1722	Snacks & Branded Foods		Ready To Cook & Eat	MTR
2401	Kitchen, Garden & Pets	Appliances & Electricals		Nippo
2402	Cleaning & Household	Disposables, Garbage Bag		Origami

	sale_price	market_price		type	rating	\
1719	NaN	20.0	Camphor & Wicks		NaN	
1720	NaN	46.0	Blended Masalas		NaN	
1721	NaN	299.0	Cutlery, Spoon & Fork		NaN	
1722	NaN	65.0	Breakfast & Snack Mixes		4.1	
2401	NaN	150.0	Battery & Electrical		4.0	
2402	NaN	32.0	Toilet Paper		3.9	

		description	discount
1719	Jaji Puja Flower Batti / Puvvu Vathulu is made...		NaN
1720	Annaporna Spices are the preserve for Connoiss...		NaN
1721	BB Home provides fine and classy cutlery that ...		NaN
1722	MTR Dhokla is extremely special in Gujarat. It...		NaN
2401	This Nippo battery incorporates state-of-the-a...		NaN
2402	Klassic Plain Cocktail Napkins (22 X 22 cm) 10...		NaN

```
df[df["rating"].isnull()]
```

	index	product \
55	56	Soothing Cucumber Facial Scrub With Apricot Seeds
59	60	Corporate Planner Diary With Premium PU Leathe...
65	66	Ayurvedic Anti-Tan Face Pack
68	69	Organic Carom Seeds/Ajwain/0m Kalu
69	70	Padded Harness - 3/4 inch, Grey Colour
...
27509	27510	Deluxe Crackers - Veg
27511	27512	Specialist Stain Remover Pen & Marker
27514	27515	Verge & Sheer Perfume For Pair
27530	27531	Tick'et to Fleadom Dry Shampoo For Dogs
27531	27532	Vanilla Ice Cream

	brand \	category	sub_category
55		Beauty & Hygiene	Skin Care
TJORI			
59		Cleaning & Household	Stationery Prozo
Plus			
65		Beauty & Hygiene	Skin Care
TJORI			
68		Foodgrains, Oil & Masala	Masalas & Spices
Earthon			
69		Kitchen, Garden & Pets	Pet Food & Accessories
Glenand			
...	
.			
27509		Gourmet & World Food	Chocolates & Biscuits
Kerk			
27511		Cleaning & Household	All Purpose Cleaners
365			
27514		Beauty & Hygiene	Fragrances & Deos Skinn by
Titan			
27530		Kitchen, Garden & Pets	Pet Food & Accessories Captain
Zack			
27531		Gourmet & World Food	Dairy & Cheese Haagen-
Dazs			

sale_price	market_price	type	rating \
------------	--------------	------	----------

55	299.4	499.0	Face Care	NaN
59	399.0	399.0	Notebooks, Files, Folders	NaN
65	269.4	449.0	Face Care	NaN
68	72.0	72.0	Whole Spices	NaN
69	840.0	840.0	Pet Collars & Leashes	NaN
...
27509	150.0	150.0	Cookies, Biscotti, Wafer	NaN
27511	449.0	449.0	Imported Cleaners	NaN
27514	1615.5	1795.0	Perfume	NaN
27530	99.0	99.0	Pet Cleaning & Grooming	NaN
27531	225.0	225.0	Gourmet Ice Cream	NaN

		description	discount
55		The soothing feel of cucumber meets the gentle...	439.0
59		A5 Size (210x150mm) \n192 Pages Premium Natura...	299.0
65		A nourishing face pack that removes tan and br...	389.0
68		Earthon's Ajwain is Best quality, organically ...	-28.0
69		These are soft padded harness for your active ...	740.0
...	
27509		Kerk Biscuits-has been a household name synony...	50.0
27511		Mightier than the pen. The pen may be mightier...	349.0
27514		VERGE for men paints a picture of a classy out...	1705.0
27530		1) No Rinse Defence Against Ticks and Fleas: C...	-1.0
27531		Some things are so good in their simplest form...	125.0

[8636 rows x 11 columns]

```
df[df["description"].isnull()]
```

	index	product \
288	289	Arrabbiata Tomato Pasta Sauce With Chilli
480	481	Pumice Stone & Emery Foot Filer - FFL3407, Col...
780	781	Peanut Butter - Chunky
845	846	Squash - Sarasaparilla/Nannari/Sugandhi Root
1047	1048	Natura Rice & Walnut Drink, Vegan
...
25510	25511	BBPopular Almond/Badam Californian 500gm + BBR...
25583	25584	Chai/Chai/Tea/ Coffee Bone China Microwave Ser...
26116	26117	Tea Strainer - Stainless Steel, Medium
26835	26836	Cold Pressed Virgin Groundnut Oil
27303	27304	Zipper Storage - Gallon

	category	sub_category
brand \		
288	Gourmet & World Food	Sauces, Spreads & Dips
Montanini		
480	Beauty & Hygiene	Bath & Hand Wash
Kaiv		
780	Gourmet & World Food	Sauces, Spreads & Dips American
Garden		

845	Beverages	Fruit Juices & Drinks	
NaturoBell			
1047	Gourmet & World Food	Drinks & Beverages	
BORGES			
...	
...			
25510	Foodgrains, Oil & Masala	Dry Fruits	bb
Combo			
25583	Kitchen, Garden & Pets	Crockery & Cutlery	BP
Bharat			
26116	Kitchen, Garden & Pets	Kitchen Accessories	
Ritu			
26835	Foodgrains, Oil & Masala	Edible Oils & Ghee	
Chekko			
27303	Cleaning & Household	Disposables, Garbage Bag	
Glad			

	sale_price	market_price	type	rating	\
288	22325.00	325.0	Mustard & Cheese Sauces	5.0	
480	175.00	175.0	Bathing Accessories	3.8	
780	299.00	299.0	Chocolate, Peanut Spread	4.1	
845	140.00	140.0	Syrups & Concentrates	2.5	
1047	275.01	349.0	Health Drinks	4.5	
...	
25510	443.00	725.0	Almonds	NaN	
25583	269.00	373.0	Cups, Mugs & Tumblers	2.5	
26116	99.00	135.0	Strainer, Ladle, Spatula	4.3	
26835	325.00	325.0	Groundnut Oil	4.2	
27303	350.00	350.0	Aluminium Foil, Clingwrap	4.7	

	description	discount
288	NaN	-6544.23
480	NaN	75.00
780	NaN	199.00
845	NaN	40.00
1047	NaN	270.20
...
25510	NaN	663.90
25583	NaN	300.88
26116	NaN	61.67
26835	NaN	225.00
27303	NaN	250.00

[115 rows x 11 columns]

```
df[df["discount"].isnull()]
```

	index	product	\
1719	1720	Puja Flower Wicks - Puvvu Vathulu Batti	
1720	1721	Powder - Sambar	

1721	1722	Steel Fork - Medium, Premium Excel Series, BBCL08
1722	1723	Snack Mix - Dhokla
2401	2402	Battery AA 3UT Hi Top
2402	2403	Klassic Plain Cocktail Napkins (22 X 22 cm)

	category	sub_category	brand \
1719	Cleaning & Household	Pooja Needs	Jaji
1720	Foodgrains, Oil & Masala	Masalas & Spices	Annapoorna
1721	Kitchen, Garden & Pets	Crockery & Cutlery	BB Home
1722	Snacks & Branded Foods	Ready To Cook & Eat	MTR
2401	Kitchen, Garden & Pets	Appliances & Electricals	Nippo
2402	Cleaning & Household	Disposables, Garbage Bag	Origami

	sale_price	market_price	type	rating \
1719	NaN	20.0	Camphor & Wicks	NaN
1720	NaN	46.0	Blended Masalas	NaN
1721	NaN	299.0	Cutlery, Spoon & Fork	NaN
1722	NaN	65.0	Breakfast & Snack Mixes	4.1
2401	NaN	150.0	Battery & Electrical	4.0
2402	NaN	32.0	Toilet Paper	3.9

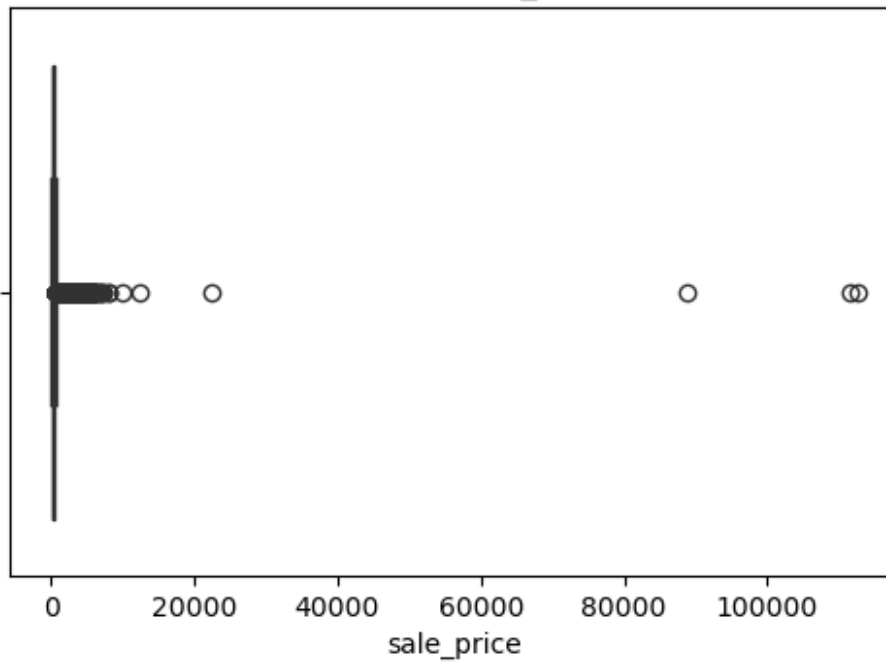
	description	discount
1719	Jaji Puja Flower Batti / Puvvu Vathulu is made...	NaN
1720	Annaporna Spices are the preserve for Connoiss...	NaN
1721	BB Home provides fine and classy cutlery that ...	NaN
1722	MTR Dhokla is extremely special in Gujarat. It...	NaN
2401	This Nippo battery incorporates state-of-the-a...	NaN
2402	Klassic Plain Cocktail Napkins (22 X 22 cm) 10...	NaN

Step 8: Find out the outliers from the dataset according to the columns

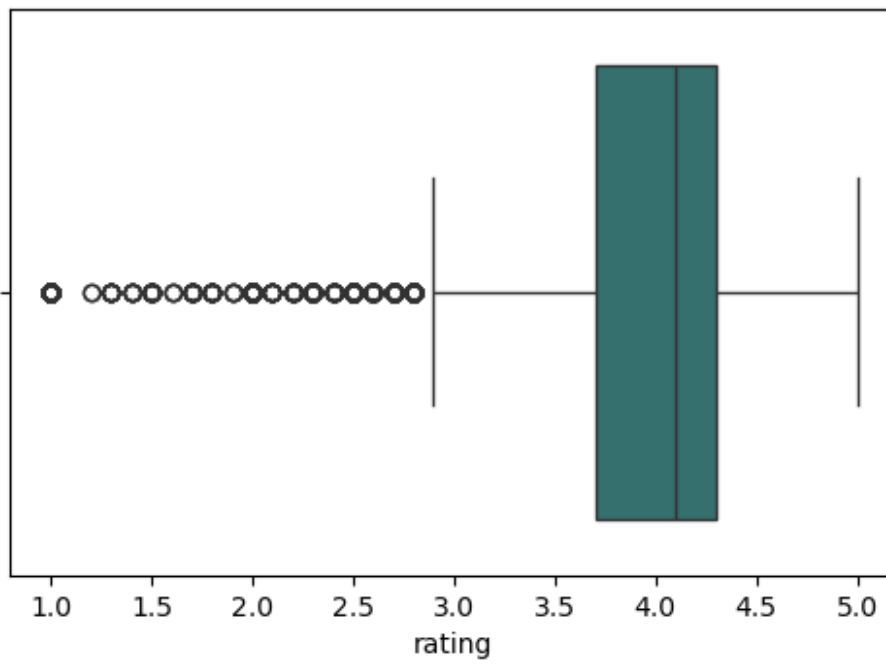
```
numerical_columns = ['sale_price', 'rating', 'market_price',
                     'discount']
plt.figure(figsize=(5, 15))
for i, column in enumerate(numerical_columns, 1):
    plt.subplot(len(numerical_columns), 1, i)
    sns.boxplot(df[column], orient='h', color='#2B7A78')
    plt.title(f'Box Plot - {column}')
```

```
plt.xlabel(column)
plt.tight_layout()
```

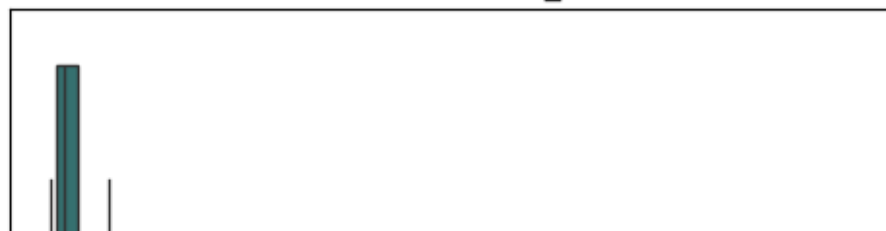
Box Plot - sale_price



Box Plot - rating



Box Plot - market_price



```
df1=df.select_dtypes([float])
df1
```

	sale_price	market_price	rating	discount
0	220.00	220.0	4.1	120.00
1	180.00	180.0	2.3	80.00
2	119.00	250.0	3.4	202.40
3	149.00	176.0	3.7	91.34
4	162.00	162.0	4.4	62.00
...
27550	199.20	249.0	3.9	169.00
27551	67.50	75.0	4.0	-15.00
27552	200.00	200.0	3.8	100.00
27553	396.00	495.0	4.2	415.00
27554	214.53	390.0	4.5	334.99

```
[27555 rows x 4 columns]
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 27555 entries, 0 to 27554
```

```
Data columns (total 11 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
10	discount	27549 non-null	float64

```
dtypes: float64(4), int64(1), object(6)
```

```
memory usage: 2.3+ MB
```

```
def detect_outliers (df,colm):
    q1=df[colm].quantile(0.25)
    q3=df[colm].quantile(0.75)
    iqr= q3-q1
    lower_limit=max(q1-1.5*iqr,0)
    upper_limit=(q3+1.5*iqr)
    return lower_limit,upper_limit
```

```
a=["sale_price","market_price","rating","discount"]
for i in a:
    print (detect_outliers (df,i))
```

```
(0, 755.0)
(0, 912.5)
(2.8000000000000007, 5.199999999999999)
(0, 827.5)
```

higher value outliers

```
df[df["sale_price"]>694.1209780754292]
```

	index	product \
8	9	Biotin & Collagen Volumizing Hair Shampoo + Bi...
47	48	Colour Catcher Sheets
51	52	Peach Syrup
69	70	Padded Harness - 3/4 inch, Grey Colour
86	87	Cheese - Feta Greek, Block
...
27498	27499	Juicer - Fruit & Vegetable, Light Green
27505	27506	Virgin Coconut Oil
27514	27515	Verge & Sheer Perfume For Pair
27538	27539	Quista Pro Advanced Whey Protein Formula forti...
27542	27543	Embellish Skin Lightening Serum

	category	sub_category
brand \		
8	Beauty & Hygiene	Hair Care
StBotanica		
47	Cleaning & Household	All Purpose Cleaners
Dylon		
51	Gourmet & World Food	Drinks & Beverages
Pekers		
69	Kitchen, Garden & Pets	Pet Food & Accessories
Glenand		
86	Gourmet & World Food	Dairy & Cheese
Signature		Fresho
...
...		
27498	Kitchen, Garden & Pets	Kitchen Accessories
Ganesh		
27505	Foodgrains, Oil & Masala	Edible Oils & Ghee
Merkera		
27514	Beauty & Hygiene	Fragrances & Deos
Skin by		
Titan		
27538	Beauty & Hygiene	Health & Medicine
Himalaya		
27542	Beauty & Hygiene	Skin Care
Organic		
Harvest		

	sale_price	market_price	type
rating \			
8	1098.00	1098.0	Shampoo & Conditioner
3.5			
47	799.00	799.0	Imported Cleaners
4.0			
51	850.00	850.0	Gourmet Juices & Drinks
4.2			
69	840.00	840.0	Pet Collars & Leashes
NaN			
86	695.70	773.0	International Cheese
NaN			
...
...			
27498	1071.00	1071.0	Kitchen Tools & Other Accessories
2.0			
27505	875.00	875.0	Other Edible Oils
NaN			
27514	1615.50	1795.0	Perfume
NaN			
27538	4500.00	4500.0	Supplements & Proteins
4.0			
27542	1525.75	1795.0	Face Care
4.2			

	description	discount
8	An exclusive blend with Vitamin B7 Biotin, Hyd...	998.0
47	1. Prevents Colour Run Accidents Colours that ...	699.0
51	Pekers peach syrup takes you on a historical t...	750.0
69	These are soft padded harness for your active ...	740.0
86	Fresho Signature brings to you one of the most...	683.0
...
27498	This juicer comes in various attractive colour...	971.0
27505	Merkera Extra Virgin Coconut Oil 100% natural,...	775.0
27514	VERGE for men paints a picture of a classy out...	1705.0
27538	Quista Pro is a whey protein blend that helps ...	4400.0
27542	Achieve an everlasting illuminated skin by inc...	1710.0

[2570 rows x 11 columns]

df[df["market_price"]>805.1416612230087]

	index	product \
8	9	Biotin & Collagen Volumizing Hair Shampoo + Bi...
25	26	Insulated Hot Fresh Casserole For Roti/Chapati...
34	35	Pet Solitaire Container Set - Silver
51	52	Peach Syrup
69	70	Padded Harness - 3/4 inch, Grey Colour
...

27498	27499	Juicer - Fruit & Vegetable, Light Green
27505	27506	Virgin Coconut Oil
27514	27515	Verge & Sheer Perfume For Pair
27538	27539	Quista Pro Advanced Whey Protein Formula forti...
27542	27543	Embellish Skin Lightening Serum

	category	sub_category
brand \		
8	Beauty & Hygiene	Hair Care
StBotanica		
25	Kitchen, Garden & Pets	Flask & Casserole
Cello		
34	Kitchen, Garden & Pets	Storage & Accessories
Steele		
51	Gourmet & World Food	Drinks & Beverages
Pekers		
69	Kitchen, Garden & Pets	Pet Food & Accessories
Glenand		
...
..		
27498	Kitchen, Garden & Pets	Kitchen Accessories
Ganesh		
27505	Foodgrains, Oil & Masala	Edible Oils & Ghee
Merker		
27514	Beauty & Hygiene	Fragrances & Deos
Titan		
27538	Beauty & Hygiene	Health & Medicine
Himalaya		
27542	Beauty & Hygiene	Skin Care
Harvest		Organic

	sale_price	market_price	type
rating \			
8	1098.00	1098.0	Shampoo & Conditioner
3.5			
25	659.00	895.0	Casserole
3.3			
34	499.00	969.0	Containers Sets
3.9			
51	850.00	850.0	Gourmet Juices & Drinks
4.2			
69	840.00	840.0	Pet Collars & Leashes
NaN			
...
...			
27498	1071.00	1071.0	Kitchen Tools & Other Accessories
2.0			
27505	875.00	875.0	Other Edible Oils
NaN			

27514	1615.50	1795.0	Perfume
NaN			
27538	4500.00	4500.0	Supplements & Proteins
4.0			
27542	1525.75	1795.0	Face Care
4.2			

		description	discount
8		An exclusive blend with Vitamin B7 Biotin, Hyd...	998.00
25		Stop your worries about keeping your food warm...	821.37
34		The Steelo plastic PET Container Set, is made ...	917.50
51		Pekers peach syrup takes you on a historical t...	750.00
69		These are soft padded harness for your active ...	740.00
...	
27498		This juicer comes in various attractive colour...	971.00
27505		Merkera Extra Virgin Coconut Oil 100% natural,...	775.00
27514		VERGE for men paints a picture of a classy out...	1705.00
27538		Quista Pro is a whey protein blend that helps ...	4400.00
27542		Achieve an everlasting illuminated skin by inc...	1710.00

[2521 rows x 11 columns]

```
df[df["discount"]> 704.1894592632917]
```

	index	product \
8	9	Biotin & Collagen Volumizing Hair Shampoo + Bi...
25	26	Insulated Hot Fresh Casserole For Roti/Chapati...
34	35	Pet Solitaire Container Set - Silver
51	52	Peach Syrup
69	70	Padded Harness - 3/4 inch, Grey Colour
...
27498	27499	Juicer - Fruit & Vegetable, Light Green
27505	27506	Virgin Coconut Oil
27514	27515	Verge & Sheer Perfume For Pair
27538	27539	Quista Pro Advanced Whey Protein Formula forti...
27542	27543	Embellish Skin Lightening Serum

	category	sub_category
brand \		
8	Beauty & Hygiene	Hair Care
StBotanica		
25	Kitchen, Garden & Pets	Flask & Casserole
Cello		
34	Kitchen, Garden & Pets	Storage & Accessories
Steelo		
51	Gourmet & World Food	Drinks & Beverages
Pekers		
69	Kitchen, Garden & Pets	Pet Food & Accessories
Glenand		
...

```

..
27498 Kitchen, Garden & Pets Kitchen Accessories
Ganesh
27505 Foodgrains, Oil & Masala Edible Oils & Ghee
Merkera
27514 Beauty & Hygiene Fragrances & Deos Skinn by
Titan
27538 Beauty & Hygiene Health & Medicine
Himalaya
27542 Beauty & Hygiene Skin Care Organic
Harvest

```

```

      sale_price market_price type
rating \
8      1098.00      1098.0 Shampoo & Conditioner
3.5
25      659.00      895.0 Casserole
3.3
34      499.00      969.0 Containers Sets
3.9
51      850.00      850.0 Gourmet Juices & Drinks
4.2
69      840.00      840.0 Pet Collars & Leashes
NaN

```

```

...      ...      ...
...
27498      1071.00      1071.0 Kitchen Tools & Other Accessories
2.0
27505      875.00      875.0 Other Edible Oils
NaN
27514      1615.50      1795.0 Perfume
NaN
27538      4500.00      4500.0 Supplements & Proteins
4.0
27542      1525.75      1795.0 Face Care
4.2

```

```

      description discount
8      An exclusive blend with Vitamin B7 Biotin, Hyd... 998.00
25      Stop your worries about keeping your food warm... 821.37
34      The Steelo plastic PET Container Set, is made ... 917.50
51      Pekers peach syrup takes you on a historical t... 750.00
69      These are soft padded harness for your active ... 740.00
...      ...      ...
27498      This juicer comes in various attractive colour... 971.00
27505      Merkera Extra Virgin Coconut Oil 100% natural,... 775.00
27514      VERGE for men paints a picture of a classy out... 1705.00
27538      Quista Pro is a whey protein blend that helps ... 4400.00
27542      Achieve an everlasting illuminated skin by inc... 1710.00

```

```
[2665 rows x 11 columns]
```

```
df["sale_price"]=df["sale_price"].apply(lambda  
x:df["sale_price"].mean()if x >694.1209780754292 else x)  
df["sale_price"]
```

```
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
```

```
df["market_price"]=df["market_price"].apply(lambda  
x:df["market_price"].mean() if x >805.1416612230087 else x)  
df["market_price"]
```

```
0      220.0  
1      180.0  
2      250.0  
3      176.0  
4      162.0
```

```
...  
27550    249.0  
27551     75.0  
27552    200.0  
27553    495.0  
27554    390.0
```

```
Name: market_price, Length: 27555, dtype: float64
```

```
#df["discount"].info()  
#df["discount"].isnull().sum()  
df["discount"]=df["discount"].fillna(0)  
df["discount"] = pd.to_numeric(df["discount"], errors='coerce')  
  
# Apply the transformation  
df["discount"] = df["discount"].apply(lambda x: df["discount"].mean()  
if x < 0 else x)
```

```
df["discount"]
```

```
0      120.000000  
1       80.000000  
2     202.400000
```

```
3      91.340000
4      62.000000
```

```
...
```

```
27550    169.000000
27551    290.675784
27552    100.000000
27553    415.000000
27554    334.990000
```

```
Name: discount, Length: 27555, dtype: float64
```

```
numerical_columns = ['sale_price', 'rating', 'market_price',  
                      'discount']
```

```
# Create box plots for each numerical column
```

```
plt.figure(figsize=(5, 15))
```

```
for i, column in enumerate(numerical_columns, 1):
```

```
    plt.subplot(len(numerical_columns), 1, i) # Create subplots for each column
```

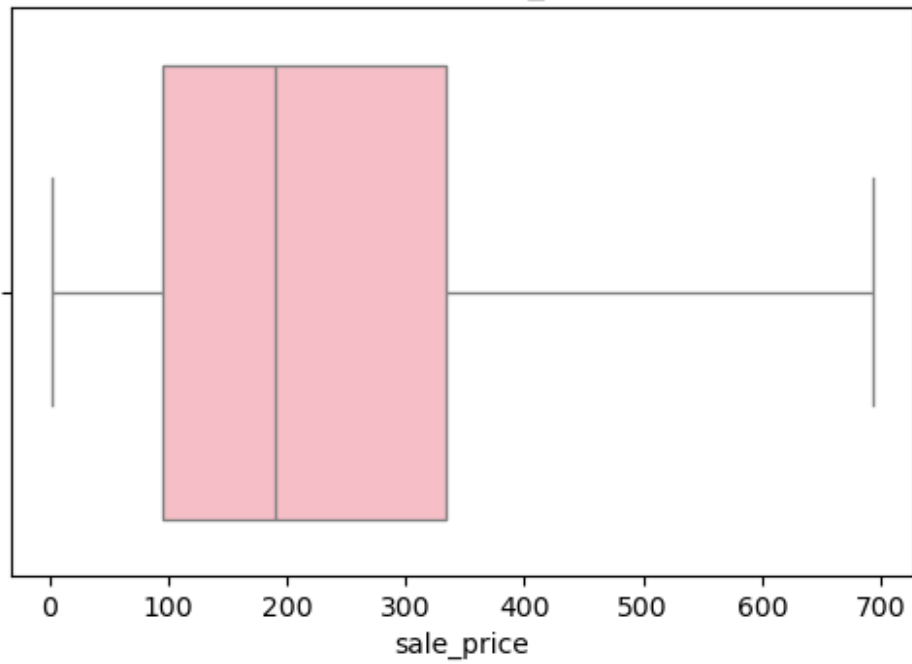
```
    sns.boxplot(df[column], orient='h', color='lightpink')
```

```
    plt.title(f'Box Plot - {column}')
```

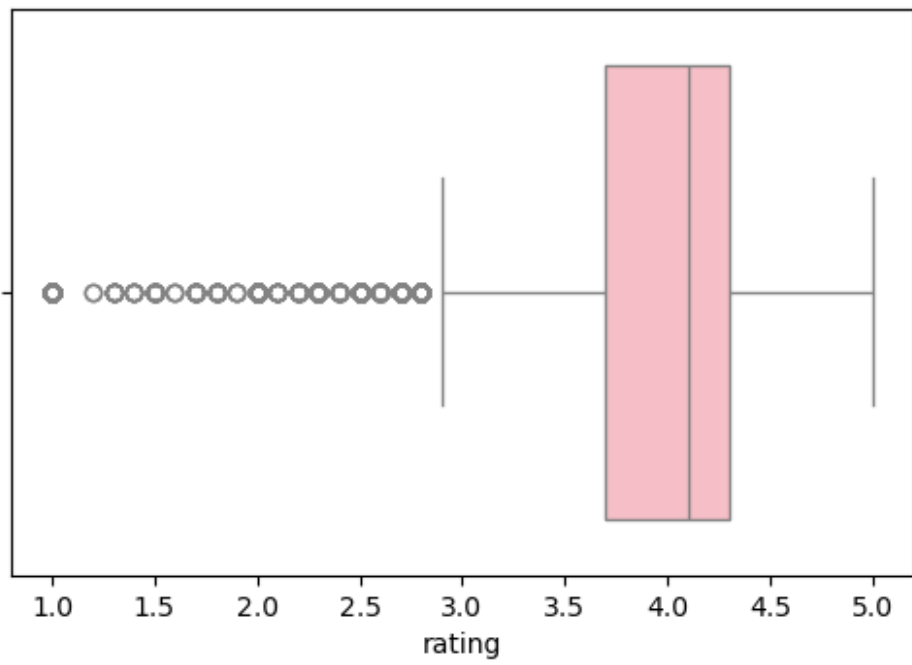
```
    plt.xlabel(column)
```

```
    plt.tight_layout()
```

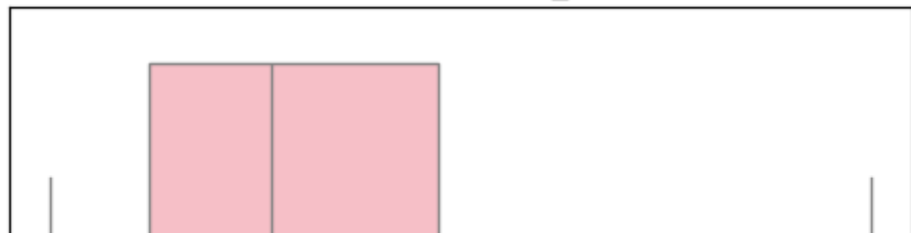
Box Plot - sale_price



Box Plot - rating



Box Plot - market_price



Step 9: Create Plots or visualizations.

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 27555 entries, 0 to 27554
Data columns (total 11 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  
10  discount               27555 non-null  float64  
dtypes: float64(4), int64(1), object(6)
memory usage: 2.3+ MB
```

```
df["category"].value_counts()
```

```
category
Beauty & Hygiene          7867
Gourmet & World Food      4690
Kitchen, Garden & Pets    3580
Snacks & Branded Foods    2814
Foodgrains, Oil & Masala  2676
Cleaning & Household      2675
Beverages                 885
Bakery, Cakes & Dairy     851
Baby Care                 610
Fruits & Vegetables       557
Eggs, Meat & Fish         350
Name: count, dtype: int64
```

A bar plot showing the count of products within each category

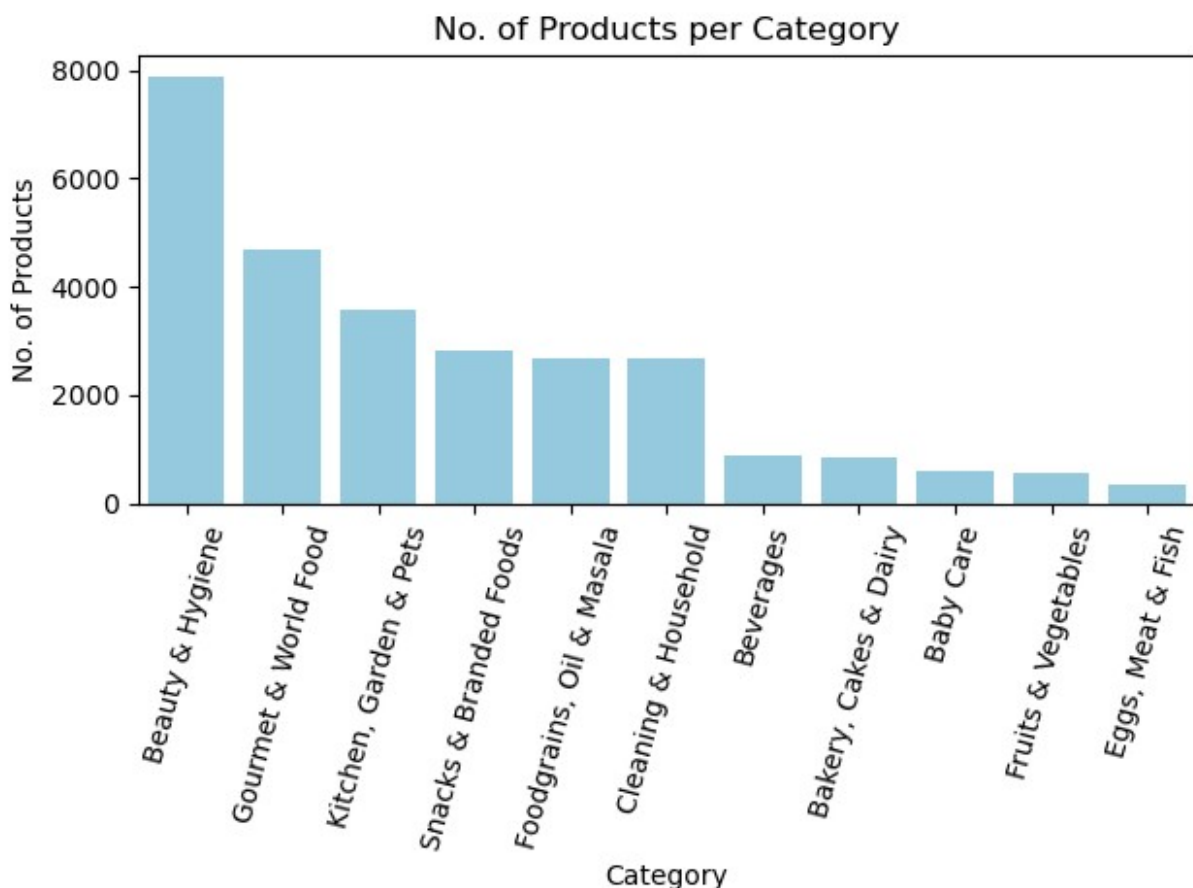
```
# Data
data = {'category': ["Beauty & Hygiene", "Gourmet & World Food",
                    "Kitchen, Garden & Pets",
                    "Snacks & Branded Foods", "Foodgrains, Oil &
Masala", "Cleaning & Household",
                    "Beverages", "Bakery, Cakes & Dairy", "Baby
Care", "Fruits & Vegetables",
                    "Eggs, Meat & Fish"]}
```

```

df2 = pd.DataFrame(data)
counts = df['category'].value_counts()

sns.barplot(x=counts.index, y=counts.values, color='skyblue')
plt.xticks(rotation=75)
plt.xlabel('Category')
plt.ylabel('No. of Products')
plt.title('No. of Products per Category')
plt.tight_layout()
plt.show()

```

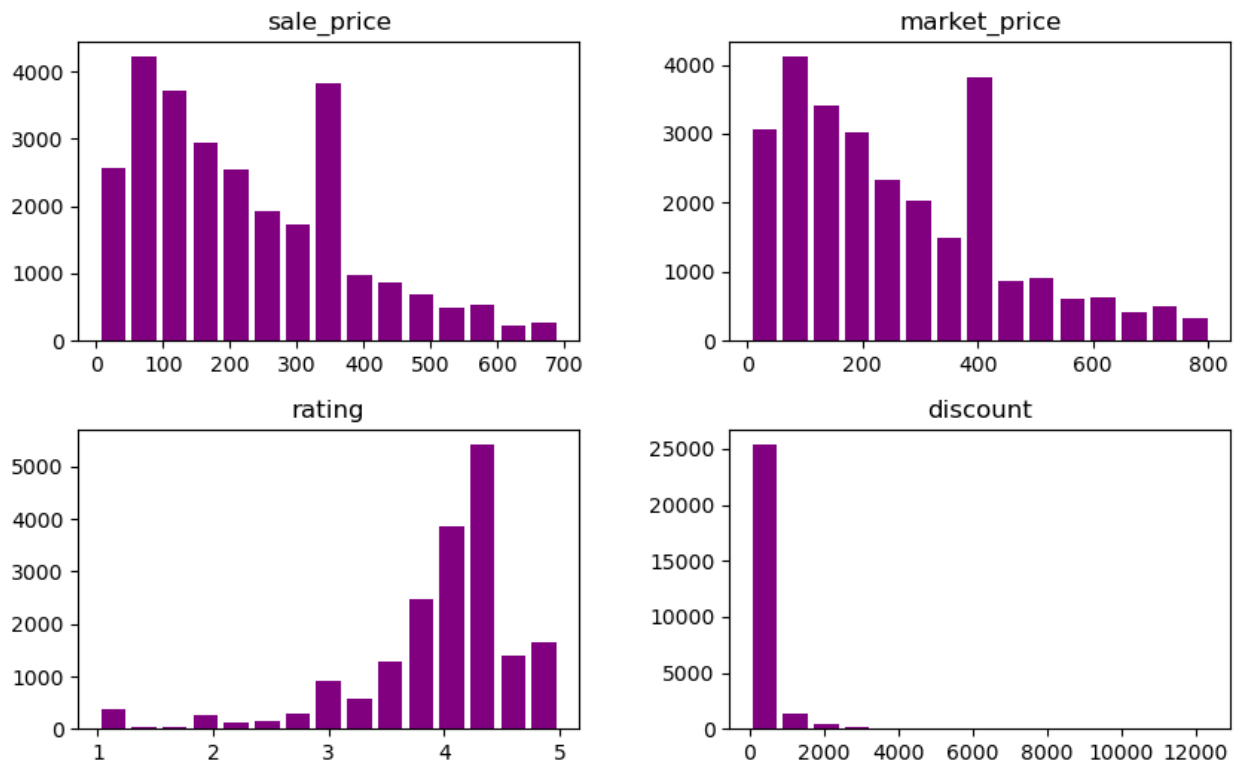


Key Takeaways: ##### • Beauty & Hygiene dominates with the highest number of products (5458). ##### • Kitchen, Garden & Pets follows closely with 2493 products. ##### • Snacks & Branded Foods and Gourmet & World Food have a similar product count, around 2400. ##### • Foodgrains, Oil & Masala and Cleaning & 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.

outliers removed from data

```
df[["sale_price", "market_price", "rating", "discount"]].hist(bins=15, fig
size=(10,6), grid=False, rwidth=0.8, color="purple")
plt.show()
```



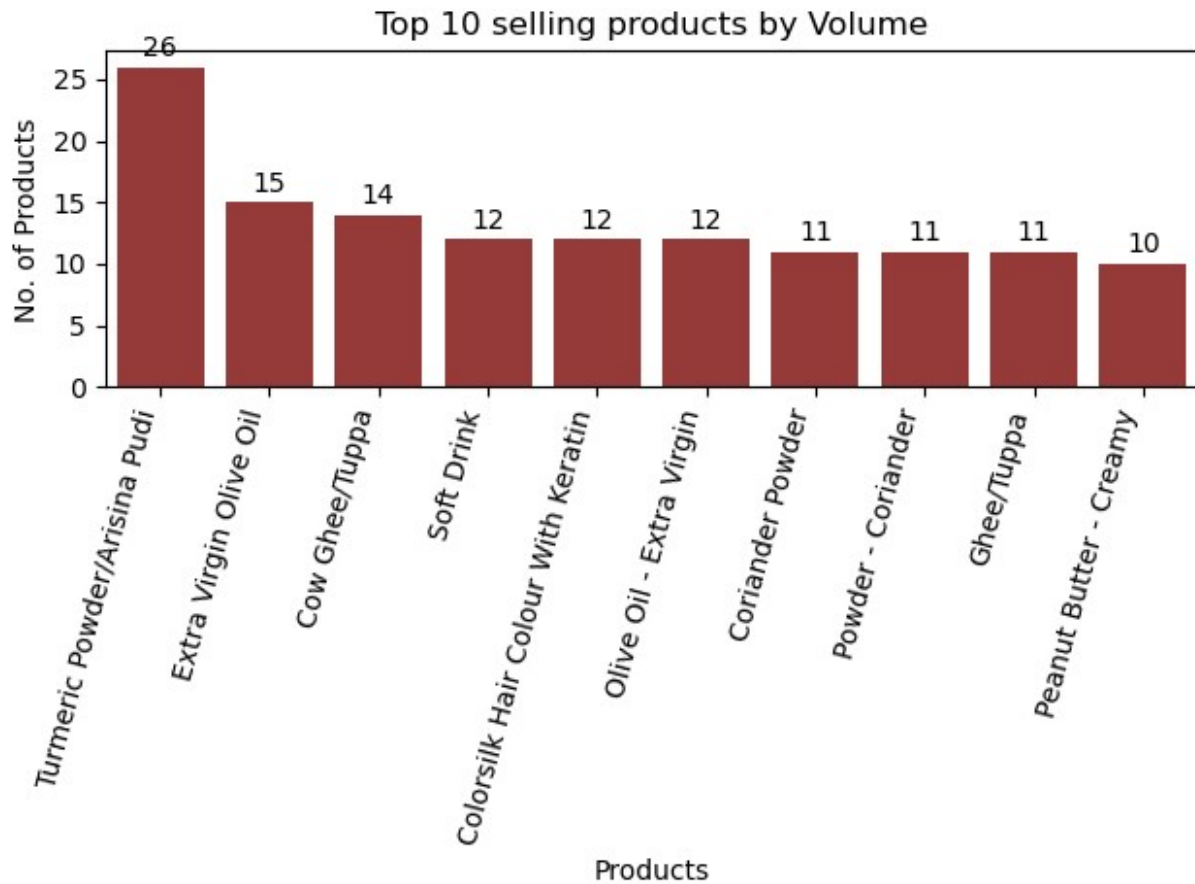
Top 10 Selling Products by Volume

```
#This bar plot displays the top 10 products with the
highest sales volume, indicating the most popular
products among customers.
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='brown')
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=(0,4),
    textcoords='offset points')
plt.tight_layout()

```



Key Takeaways:

.Turmeric Powder/Arisina Pudi is the top-selling product with a significantly higher volume (26 units) compared to others.

.Extra Virgin Olive Oil is the second-highest selling product with 15 units, followed closely by Cow Ghee/Tuppa with 14 units.

.Soft Drink, Olive Oil - Extra Virgin, and Colorsilk Hair Colour With Keratin share the same volume (12 units), showcasing moderate sales performance.

.Coriander Powder, Powder - Coriander, and Ghee/Tuppa each have 11 units sold, indicating steady demand.

.Casting Creme Gloss Hair Color rounds out the top 10 with 10 units sold, slightly lower than the rest.

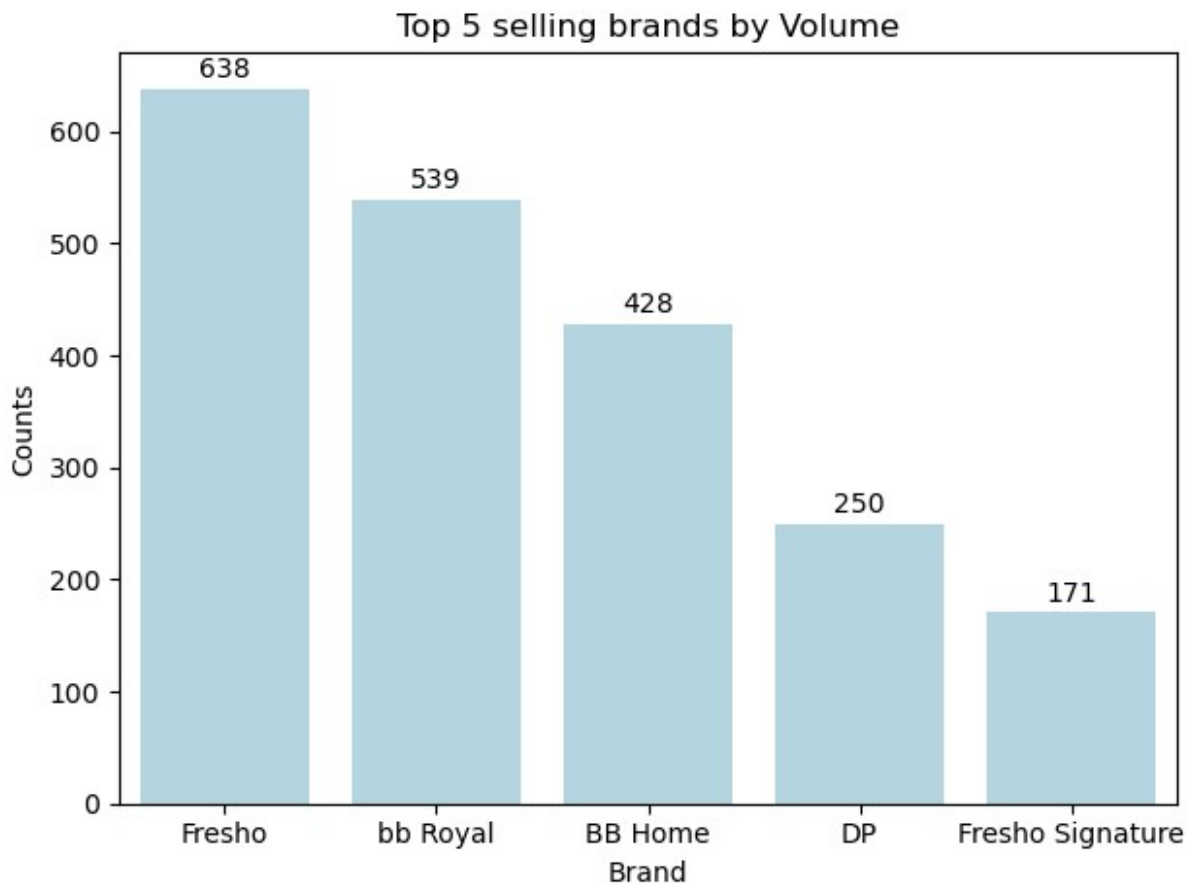
Recommendations:

.Leverage Top-Selling Products: ##### .Increase promotions and visibility for Turmeric Powder/Arisina Pudi and Extra Virgin Olive Oil to further capitalize on their popularity. ##### .Bundle these products with related items to boost cross-selling opportunities. ### Optimize Moderate-Performing Products: ##### .Analyze factors driving the performance of products like Soft Drink and Colorsilk Hair Colour With Keratin to replicate success in similar categories. ##### .Explore opportunities to enhance their visibility through targeted campaigns. #### .Enhance Sales for Lower-Performing Products: ##### .Investigate the lower sales volume of Casting Creme Gloss Hair Color and develop strategies to increase its appeal, such as offering discounts or showcasing customer testimonials. ##### .Evaluate the product placement and pricing strategies to drive higher sales

Top 5 Selling Brands by Volume

```
# This bar plot illustrates the five brands with the
highest sales volume, highlighting the most popular
brands.
counts=df['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='lightblue')
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=(0,4),textc
oords='offset points')
```

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

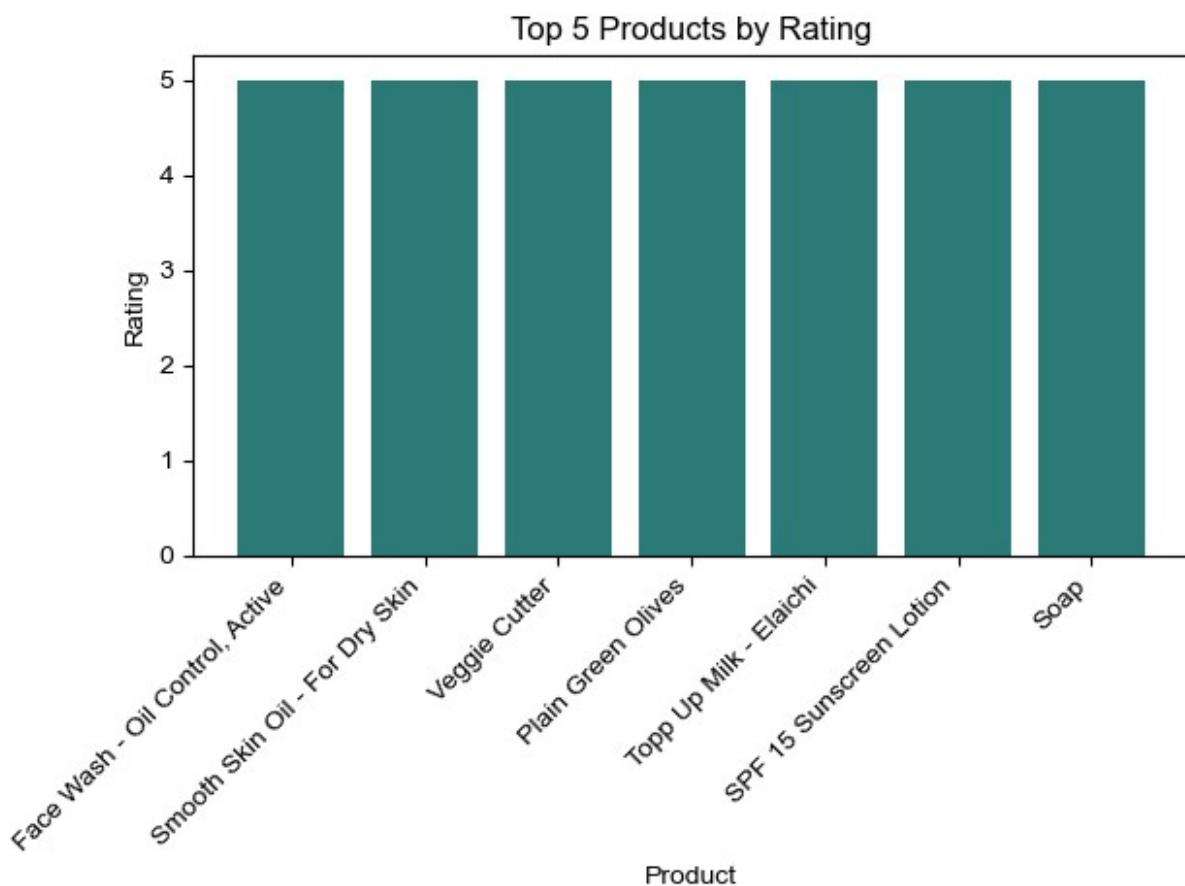


Key Insights: ##### Fresho is the top-selling brand with a volume of 638 units, significantly outperforming others. ##### bb Royal follows with 539 units, showing strong market presence. ##### BB Home, DP, and Fresho Signature occupy the remaining top 5 positions, with sales volumes of 428, 250, and 171, respectively. ##### The volume difference between the top-performing brands (Fresho and bb Royal) and the lower performers (DP and Fresho Signature) is notable, indicating possible brand preference disparities. ## Recommendations: ##### Focus on Fresho and bb Royal: Strengthen marketing strategies and maintain inventory for these high-demand brands. ##### Improve Mid-Tier Brands: Investigate the factors limiting the sales of DP and Fresho Signature, such as pricing, placement, or promotions. ##### Optimize Brand Performance: Perform a detailed analysis on BB Home to identify opportunities to increase its market share further. ##### Customer Insights: Conduct surveys or feedback sessions to understand customer preferences for the top-selling brands. ##### Expand Offerings: Consider introducing new products under the Fresho and bb Royal brands to capture additional market share.

Top 5 Products by Rating:

```
top_5_products =df.nlargest(7, 'rating')[['product','rating']]

plt.bar(top_5_products['product'],top_5_products['rating'],color='#2B7A78')
plt.xlabel('Product')
plt.ylabel('Rating')
plt.title('Top 5 Products by Rating')
plt.xticks(rotation=45,ha='right')
plt.tight_layout()
plt.show()
```



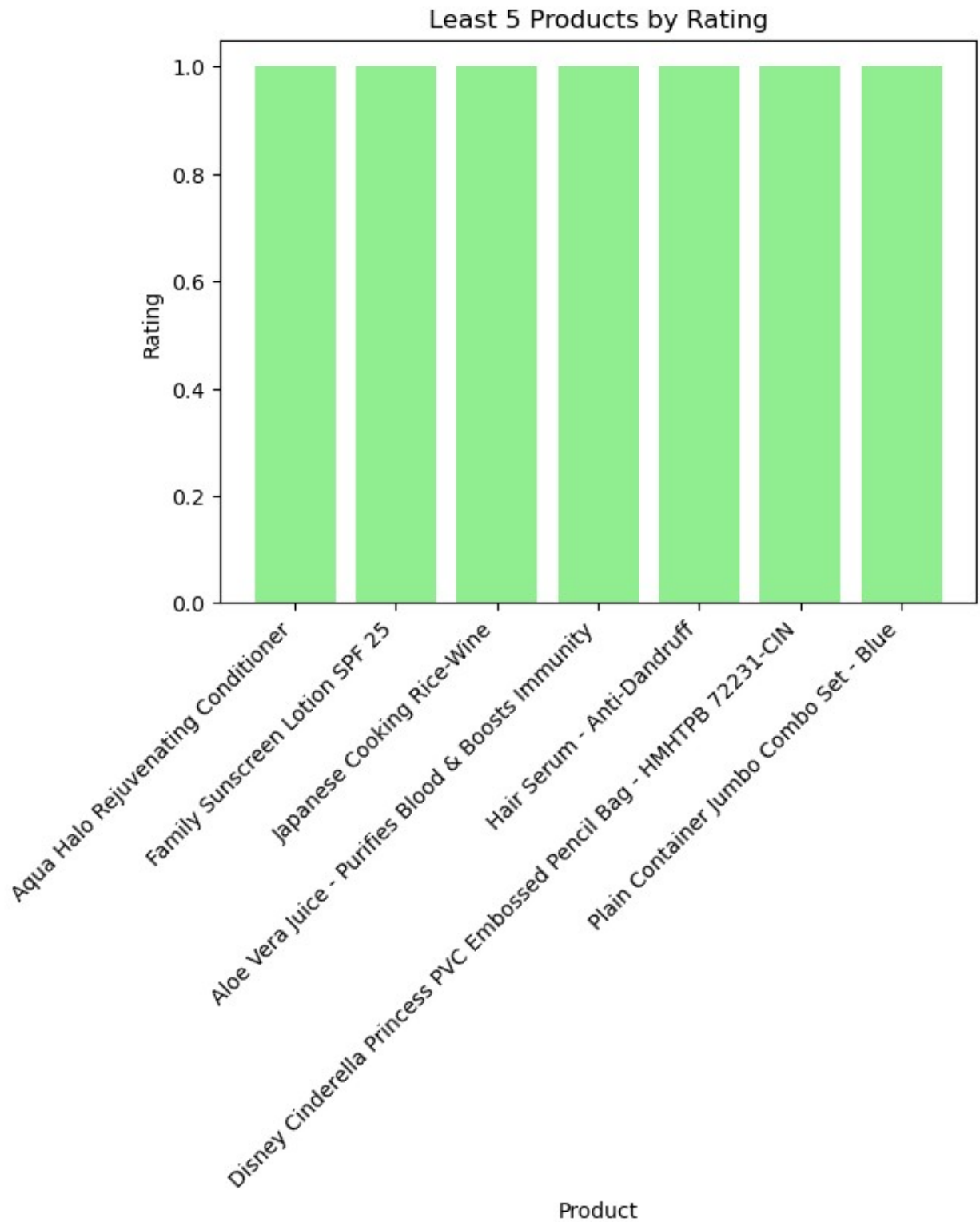
Key Insights:

Top 5 Products by Rating: ##### All five products have achieved a high rating of 4+, indicating consistent customer satisfaction: ##### 1. Face Wash - Oil Control, Active ##### 2. Smooth Skin Oil - For Dry Skin ##### 3. Salted Pumpkin ##### 4. Veggie Cutter ##### 5. Plain Green Olives ### Recommendations: ##### Maintain High Standards: ##### Continue to focus on quality and performance for these products to sustain high customer ratings. ##### Gather and showcase customer reviews for these top-rated products to build trust and attract

new buyers. ### Expand Related Offerings: ##### Introduce variations or complementary products in the same categories (e.g., more types of skincare products or kitchen tools) to capture broader interest. ##### Use these high-rated products as anchors to promote bundles or combos. ### Leverage Marketing: ##### Highlight these products in promotional campaigns or feature them in "Top Rated" sections on websites or stores. ##### Offer limited-time discounts or loyalty rewards for these products to boost sales further

Least 5 Products by Rating

```
#A bar plot of the least-rated products, showing the top 5 products with the least ratings.
top_5_products=df.nsmallest(7,'rating')[['product','rating']]
plt.bar(top_5_products['product'],top_5_products['rating'],color='lightgreen')
plt.xlabel('Product')
plt.ylabel('Rating')
plt.title('Least 5 Products by Rating')
plt.xticks(rotation=45, ha='right')
plt.show()
```



Key Takeaways:

Water Bottle - Orange ##### Brass Angle Deep - Plain, No.2 ##### Germ - Removal Multipurpose Wipes ##### Multani Mati ##### Biotin & Collagen Volumizing Hair Shampoo + Biotin & Collagen Hair Conditioner ##### Butter Cookies Gold Collection ##### Insulated Hot Fresh Casserole For Roti/Chapati - White ##### All of these products are the lowest-rated, each 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 ratings. ## Overall Conclusion

Big Basket's strategy effectively balances high-frequency essentials, strategic discounting, and in-house brand development to build customer loyalty.

Improvements could include: ##### Addressing quality issues for lower-rated products. ##### Expanding the product mix in underperforming categories

