

## Project : BIG BASKET PROJECT

**Project Overview :** The Big Basket dataset is the database of various products being sold by Big Basket. It gives detailed information about the products, their category, sub category and type of products being sold apart from the market and sale price of the product alongwith their rating as given by the consumers. The dataset also contains a column giving detailed description of various products.

```
# importing libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.io as pio
```

## Step 1: Load Dataset

```
# Read data into Python

# The data contains 27555 records of products being sold by Big Basket
# arranged in 10 columns.
# The 'product' column gives the name of the product itself, whereas
# the 'category', 'sub_category'
# and 'type' columns describe in detail about the product. The
# 'market_price' and 'sale_price' give
# information about the market and sale price of various products. The
# 'rating' column is the rating
# of various products as given by consumers.

bb = pd.read_csv("//content//BigBasket Products.csv")
bb

{
  "summary": {
    "name": "bb",
    "rows": 27555,
    "fields": [
      {
        "column": "index",
        "properties": {
          "dtype": "number",
          "std": 7954,
          "min": 1,
          "max": 27555,
          "num_unique_values": 27555,
          "samples": [16202, 24963, 18965],
          "semantic_type": "\",
          "description": "\n"
        }
      },
      {
        "column": "product",
        "properties": {
          "dtype": "string",
          "num_unique_values": 23540,
          "samples": [
            "Pure Impact Body Wash With Purifying Micro Particles - Shower Gel For Body, Face & Hair",
            "Premium Care XXL - 30 Diaper Pants",
            "Tomato - Green"
          ],
          "semantic_type": "\",
          "description": "\n"
        }
      },
      {
        "column": "category",
        "properties": {
          "dtype": "category",
          "num_unique_values": 11,
          "samples": [
            "Snacks & Branded Foods"
          ]
        }
      }
    ]
  }
}
```

```
\"Beauty & Hygiene\", \n          \"Fruits & Vegetables\" \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\\n      \"column\": \"sub_category\", \n      \"properties\": { \n          \"dtype\": \"category\", \n          \"num_unique_values\": 90, \n          \"samples\": [ \n              \"Crockery & Cutlery\", \n              \"Fragrances & Deos\", \n              \"Energy & Soft Drinks\" \n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\\n      \"column\": \"brand\", \n          \"properties\": { \n              \"dtype\": \"category\", \n              \"num_unique_values\": 2313, \n              \"samples\": [ \n                  \"Red Label\", \n                  \"Lomani\", \n                  \"Happily Unmarried\" \n              ], \n              \"semantic_type\": \"\", \n              \"description\": \"\\n      \"column\": \"sale_price\", \n              \"properties\": { \n                  \"dtype\": \"number\", \n                  \"std\": 1202.102113074331, \n                  \"min\": 2.45, \n                  \"max\": 112475.0, \n                  \"num_unique_values\": 3261, \n                  \"samples\": [ \n                      24.95, \n                      40.0, \n                      5040.0 \n                  ], \n                  \"semantic_type\": \"\", \n                  \"description\": \"\\n      \"column\": \"market_price\", \n                  \"properties\": { \n                      \"dtype\": \"number\", \n                      \"std\": 581.7307174034938, \n                      \"min\": 3.0, \n                      \"max\": 12500.0, \n                      \"num_unique_values\": 1348, \n                      \"samples\": [ \n                          192.0, \n                          1164.0, \n                          74.38 \n                      ], \n                      \"semantic_type\": \"\", \n                      \"description\": \"\\n      \"column\": \"type\", \n                      \"properties\": { \n                          \"dtype\": \"category\", \n                          \"num_unique_values\": 426, \n                          \"samples\": [ \n                              \"Nappies & Rash Cream\", \n                              \"Hair Removal\", \n                              \"Crackers & Digestive\" \n                          ], \n                          \"semantic_type\": \"\", \n                          \"description\": \"\\n      \"column\": \"rating\", \n                          \"properties\": { \n                              \"dtype\": \"number\", \n                              \"std\": 0.7392170543515166, \n                              \"min\": 1.0, \n                              \"max\": 5.0, \n                              \"num_unique_values\": 40, \n                              \"samples\": [ \n                                  3.0, \n                                  1.5, \n                                  1.5, \n                                  3.9 \n                              ], \n                              \"semantic_type\": \"\", \n                              \"description\": \"\\n      \"column\": \"description\", \n                              \"properties\": { \n                                  \"dtype\": \"string\", \n                                  \"num_unique_values\": 21944, \n                                  \"samples\": [ \n                                      \"Premium Day Face Cream:Lightweight formula that will give your skin \n                                      a daily boost of radiance and hydration. Packed with the goodness of \n                                      skin brightening, Vitamin C Why Is It Special?- It contains the Most \n                                      Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is \n                                      the next generation Vitamin C that really works in improving skin \n                                      health.- Vitamin C is a proven skin-brightening agent known for its \n                                      antioxidant properties for decades. It boosts collagen formation and \n                                      effectively targets visible signs of ageing, giving skin the much- \n                                      needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3 \n                                      helps awaken and clarify your complexion, revealing a natural healthy- \n                                      looking glow. - SPF 30 & PA+++ provide the highest level of protection \n                                  ] \n                              } \n                          } \n                      } \n                  } \n              } \n          ] \n      } \n  } \n} \n
```

```

against UVA & UVB rays of the sun. The mineral-based sun protection agents shield the skin against the harsh solar radiation, minimizing skin dryness and dullness. - Multiple natural oils are combined for intense hydration and nourishment. Antioxidant-rich botanical extracts are added to keep skin moisturized, smooth & soft. - St.Botanica is a PETA-certified cruelty-free brand. The product does not contain any harmful chemicals & is free from parabens, silicones or mineral oils.\n          \"Kari Kari Japanese premium snacks are the exciting irresistible new taste of crunch. Lite and crispy, the savoury rice crackers are mixed with peanuts to make hearty, wholesome snack with delicious, bold and zesty taste with every bite.\",\n\"Safely store your food in this pink lunch box. The lunch box is made plastic which gives the container durability and sturdiness. The rectangular containers come with an airtight lid to keep food fresh for long. The wide mouth opening of the lunch makes it easy to fill food and easy to clean. The set makes a great gift option for special days.\n      ],\n      \"semantic_type\": \"\",\n\"description\": \"\"\n    }\n}\n,\"type\":\"dataframe\",\"variable_name\":\"bb\"}

```

## Step 2: Use Head function to look for first 12 rows

```
bb.head(12)
```

```

{"summary": {"name": "bb", "rows": 27555, "fields": [
  {"column": "index", "properties": {
    "dtype": "number", "std": 7954, "min": 1, "max": 27555, "num_unique_values": 27555, "samples": [16202, 24963, 18965], "semantic_type": "", "description": ""}, {"column": "product", "properties": {
      "dtype": "string", "num_unique_values": 23540, "samples": ["Pure Impact Body Wash With Purifying Micro Particles - Shower Gel For Body, Face & Hair", "Premium Care XXL - 30 Diaper Pants", "Tomato - Green"], "semantic_type": "", "description": ""}, {"column": "category", "properties": {
      "dtype": "category", "num_unique_values": 11, "samples": ["Snacks & Branded Foods", "Beauty & Hygiene", "Fruits & Vegetables"], "semantic_type": "", "description": ""}, {"column": "sub_category", "properties": {
      "dtype": "category", "num_unique_values": 90, "samples": ["Fragrances & Deos", "Crockery & Cutlery", "Energy & Soft Drinks"], "semantic_type": "", "description": ""}, {"column": "brand", "properties": {
      "dtype": "category", "num_unique_values": 1, "samples": ["\""], "semantic_type": "", "description": ""}}]
}
```

```
\"category\", \n      \"num_unique_values\": 2313, \n\n  \"samples\": [\n    {\"Red Label\", \"Lomani\"},\n    {\"Happily Unmarried\", \"semantic_type\": \"\", \"column\": \"\"},\n    {\"description\", \"\", \"semantic_type\": \"\", \"column\": \"\"},\n    {\"sale_price\", \"\", \"properties\": {\"dtype\": \"number\", \"std\": 1202.102113074331, \"min\": 2.45, \"max\": 112475.0, \"num_unique_values\": 3261, \"samples\": [24.95, 40.0, 5040.0]}, \"semantic_type\": \"\", \"column\": \"\"},\n    {\"market_price\", \"\", \"properties\": {\"dtype\": \"number\", \"std\": 581.7307174034938, \"min\": 3.0, \"max\": 12500.0, \"num_unique_values\": 1348, \"samples\": [192.0, 1164.0, 74.38]}, \"semantic_type\": \"\", \"column\": \"\"},\n    {\"type\", \"\", \"properties\": {\"dtype\": \"category\", \"num_unique_values\": 426, \"samples\": [\"Nappies & Rash Cream\", \"Hair Removal\", \"Crackers & Digestive\"]}, \"semantic_type\": \"\", \"column\": \"\"},\n    {\"rating\", \"\", \"properties\": {\"dtype\": \"number\", \"std\": 0.7392170543515166, \"min\": 1.0, \"max\": 5.0, \"num_unique_values\": 40, \"samples\": [3.0, 1.5, 1.5, 3.9]}, \"semantic_type\": \"\", \"description\": \"\", \"column\": \"\"},\n    {\"description\", \"\", \"semantic_type\": \"\", \"column\": \"\"},\n    {\"properties\": {\"dtype\": \"string\", \"num_unique_values\": 21944, \"samples\": [\"Premium Day Face Cream:Lightweight formula that will give your skin a daily boost of radiance and hydration. Packed with the goodness of skin brightening, Vitamin C Why Is It Special?- It contains the Most Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is the next generation Vitamin C that really works in improving skin health.- Vitamin C is a proven skin-brightening agent known for its antioxidant properties for decades. It boosts collagen formation and effectively targets visible signs of ageing, giving skin the much-needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3 helps awaken and clarify your complexion, revealing a natural healthy-looking glow. - SPF 30 & PA+++ provide the highest level of protection against UVA & UVB rays of the sun. The mineral-based sun protection agents shield the skin against the harsh solar radiation, minimizing skin dryness and dullness. - Multiple natural oils are combined for intense hydration and nourishment. Antioxidant-rich botanical extracts are added to keep skin moisturized, smooth & soft. - St.Botanica is a PETA-certified cruelty-free brand. The product does not contain any harmful chemicals & is free from parabens, silicones or mineral oils.\", \"Kari Kari Japanese premium snacks are the exciting irresistible new taste of crunch. Lite and crispy, the\"}, \"semantic_type\": \"\", \"column\": \"\"}\n  ]\n}
```

```

savoury rice crackers are mixed with peanuts to make hearty, wholesome
snack with delicious, bold and zesty taste with every bite.\",\n
\"Safely store your food in this pink lunch box. The lunch box is made
plastic which gives the container durability and sturdiness. The
rectangular containers come with an airtight lid to keep food fresh
for long. The wide mouth opening of the lunch makes it easy to fill
food and easy to clean. The set makes a great gift option for special
days.\n      ],\n      \"semantic_type\": \"\",\n      \"description\": \"\"\n    }\n  },\n  \"type\":\"dataframe\", \"variable_name\":\"bb\"}

```

## Step 3: Get description of the data in the Data Frame

```

bb.describe(include='all')

{
  "summary": {
    "name": "bb",
    "rows": 11,
    "fields": [
      {
        "column": "index",
        "properties": {
          "dtype": "number",
          "std": 9930.650518628143,
          "min": 1.0,
          "max": 27555.0,
          "num_unique_values": 6,
          "samples": [27555.0, 27555.0, 20666.5, 20666.5, 13778.0, 13778.0]
        },
        "semantic_type": "",
        "description": {
          "column": "product",
          "properties": {
            "dtype": "category",
            "num_unique_values": 26,
            "samples": [23540, 23540, 23540, 23540, 23540, 23540]
          }
        }
      },
      {
        "column": "category",
        "properties": {
          "dtype": "category",
          "num_unique_values": 4,
          "samples": [27555, 27555, 27555, 27555]
        },
        "semantic_type": "",
        "description": {
          "column": "sub_category",
          "properties": {
            "dtype": "category",
            "num_unique_values": 4,
            "samples": [2294, 2294, 2294, 2294]
          }
        }
      },
      {
        "column": "brand",
        "properties": {
          "dtype": "category",
          "num_unique_values": 4,
          "samples": [2313, 2313, 2313, 2313]
        },
        "semantic_type": "",
        "description": {
          "column": "sale_price",
          "properties": {
            "dtype": "number",
            "std": 39430.64768451627,
            "min": 2.45,
            "max": 112475.0,
            "num_unique_values": 8,
            "samples": [334.6483912301717, 334.6483912301717, 190.32, 190.32, 27549.0, 27549.0, 27549.0, 27549.0]
          }
        }
      },
      {
        "column": "market_price",
        "properties": {
          "dtype": "number",
          "std": 39430.64768451627,
          "min": 2.45,
          "max": 112475.0,
          "num_unique_values": 8,
          "samples": [334.6483912301717, 334.6483912301717, 190.32, 190.32, 27549.0, 27549.0, 27549.0, 27549.0]
        }
      }
    ]
  }
}

```

```

9987.087169346814,\n          \\"min\\": 3.0,\n          \\"max\\": 27555.0,\n\\\"num_unique_values\\": 8,\n          \\"samples\\": [\n            220.0,\n            27555.0\n          ],\n\\\"semantic_type\\": \"\\\",\\n          \\"description\\": \"\\\"\\n          }\\n        },\\n        {\n          \\"column\\": \"type\\\",\\n          \\"properties\\": {\n            \\"dtype\\": \"category\\\",\\n            \\"num_unique_values\\": 4,\n            \\"samples\\": [\n              426,\n              \"1508\\\",\\n              \"27555\\n            ],\\n            \\"semantic_type\\": \"\\\",\\n            \\"description\\": \"\\\"\\n          },\\n          {\n            \\"column\\": \"rating\\\",\\n            \\"properties\\": {\n              \\"dtype\\": \"number\\\",\\n              \\"std\\": 6687.726086899768,\n              \\"min\\": 0.7392170543515166,\n              \\"max\\": 18919.0,\n              \\"num_unique_values\\": 8,\n              \\"samples\\": [\n                3.9432951001638563,\n                4.1,\n                18919.0\\n              ],\\n              \\"semantic_type\\": \"\\\",\\n              \\"description\\": \"\\\"\\n            },\\n            {\n              \\"column\\": \"category\\\",\\n              \\"properties\\": {\n                \\"dtype\\": \"category\\\",\\n                \\"num_unique_values\\": 4,\n                \\"samples\\": [\n                  21944,\n                  \"47\\\",\\n                  \"27440\\n                ],\\n                \\"semantic_type\\": \"\\\",\\n                \\"description\\": \"\\\"\\n              },\\n              {\n                \\"column\\": \"\\n              }\n              ]\\n            }\\n          },\\n          \\"type\\": \"dataframe\"\n        }

```

The .describe() function gives detailed information about the dataset. It gives columnwise mean, median, Std Deviation apart from minimum & maximum values. It is a very useful function to understand the dataset using a single command.

## Step 4: Find information about the Data Frame

```

bb.loc[50:60, ['product', 'sale_price']]

{"summary": "{\\n  \\\"name\\\": \"bb\\\",\\n  \\\"rows\\\": 11,\\n  \\\"fields\\\": [\n    {\n      \\\"column\\\": \"product\\\",\\n      \\"properties\\\": {\n        \\"dtype\\\": \"string\\\",\\n        \\"num_unique_values\\\": 11,\n        \\"samples\\\": [\n          \"Soothing Cucumber Facial Scrub With\n          Apricot Seeds\\\",\\n          \"Pani Puri Mix Paste\\\",\\n          \"Corporate Planner Diary With Premium PU Leather Cover With Card\n          Holder\\n        ],\\n        \\"semantic_type\\\": \"\\\",\\n        \\"description\\\": \"\\\"\\n        },\\n        {\n          \\"column\\\": \"sale_price\\\",\\n          \\"properties\\\": {\n            \\"dtype\\\": \"number\\\",\\n            \\"std\\\": 232.6853584681566,\n            \\"min\\\": 48.95,\n            \\"max\\\": 850.0,\n            \\"num_unique_values\\\": 11,\n            \\"samples\\\": [\n              299.4,\n              48.95,\n              399.0\\n            ],\\n            \\"semantic_type\\\": \"\\\",\\n            \\"description\\\": \"\\\"\\n          },\\n          {\n            \\"column\\\": \"rating\\\",\\n            \\"properties\\\": {\n              \\"dtype\\\": \"number\\\",\\n              \\"std\\\": 6687.726086899768,\n              \\"min\\\": 0.7392170543515166,\n              \\"max\\\": 18919.0,\n              \\"num_unique_values\\\": 8,\n              \\"samples\\\": [\n                3.9432951001638563,\n                4.1,\n                18919.0\\n              ],\\n              \\"semantic_type\\\": \"\\\",\\n              \\"description\\\": \"\\\"\\n            },\\n            {\n              \\"column\\\": \"category\\\",\\n              \\"properties\\\": {\n                \\"dtype\\\": \"category\\\",\\n                \\"num_unique_values\\\": 4,\n                \\"samples\\\": [\n                  21944,\n                  \"47\\\",\\n                  \"27440\\n                ],\\n                \\"semantic_type\\\": \"\\\",\\n                \\"description\\\": \"\\\"\\n              },\\n              {\n                \\"column\\\": \"\\n              }\n              ]\\n            }\\n          },\\n          \\"type\\": \"dataframe\"\n        }

```

```
bb.loc[25:45, ['product', 'rating']]
```

```
{"summary": {"name": "bb", "rows": 21, "fields": [{}], "column": "product", "properties": {"dtype": "string", "num_unique_values": 21, "samples": ["Insulated Hot Fresh Casserole For Roti/Chapati - White", "Flavoured Cream Wafer Roll - Strawberry", "Eau De Toilette - Homme Green"], "semantic_type": "\\", "description": "\\n\\n"}, {"column": "rating", "properties": {"dtype": "number", "std": 0.747089591108628, "min": 1.5, "max": 5.0, "num_unique_values": 14, "samples": [3.9, 4.4, 3.3]}, "semantic_type": "\\", "description": "\\n\\n"}], "type": "dataframe"}}

bb.iloc[10:20, [3,4]]

{"summary": {"name": "bb", "rows": 10, "fields": [{}], "column": "sub_category", "properties": {"dtype": "string", "num_unique_values": 8, "samples": ["Chocolates & Biscuits", "Bins & Bathroom Ware", "Cooking & Baking Needs"], "semantic_type": "\\", "description": "\\n\\n"}, {"column": "brand", "properties": {"dtype": "string", "num_unique_values": 10, "samples": ["True Elements", "Sapphire", "Nakoda"]}, "semantic_type": "\\", "description": "\\n\\n"}], "type": "dataframe"}}

bb.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
```

The .info() function gives detailed information about the type of data in the dataset. It gives information about the type of data values in the dataset. It is a very useful function to understand the dataset using a single command.

## Step 5: Find out Top & least sold products

```
# Count occurrences of each product
product_counts = bb['product'].value_counts()

# Find the maximum & minimum count
max_count = product_counts.max()
min_count = product_counts.min()

# Filter products with the most & least sales
most_sold_products = product_counts[product_counts == max_count]
least_sold_products = product_counts[product_counts == min_count]

# Display the results
print("Product Counts")
print(product_counts)

print("Most Sold Products:")
print(most_sold_products)

print("Least Sold Products:")
print(least_sold_products)

print("Maximum count", max_count)
print("Minimum count", min_count)

Product Counts
product
Turmeric Powder/Arisina Pudi          26
Extra Virgin Olive Oil                15
Cow Ghee/Tuppa                        14
Soft Drink                            12
Colorsilk Hair Colour With Keratin   12
                                         ..
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, Length: 23540, dtype: int64
Most Sold Products:
product
Turmeric Powder/Arisina Pudi      26
Name: count, dtype: int64
```

```

Least Sold Products:
product
Chips - Mediterranean 1
Smoked Cheese - Black Pepper, Sliced 1
Variety Protein Bars Cranberry Orange-Kesar Almond Fudge-Coconut 1
Namkeen - Mini Samosa 1
Baked Cake Decorator - With Icing Tube 1
..
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, Length: 20566, dtype: int64
Maximum count 26
Minimum count 1

# Count occurrences of each category
category_counts = bb['category'].value_counts()

# Find the maximum & minimum count
max_count = category_counts.max()
min_count = category_counts.min()

# Filter products with the most & least sales
most_sold_category = category_counts[category_counts == max_count]
least_sold_category = category_counts[category_counts == min_count]

# Display the results
print("Most Sold category:")
print(most_sold_category)

print("Least Sold category:")
print(least_sold_category)

print("Maximum count", max_count)
print("Minimum count", min_count)

Most Sold category:
category
Beauty & Hygiene    7867
Name: count, dtype: int64
Least Sold category:
category
Eggs, Meat & Fish    350
Name: count, dtype: int64
Maximum count 7867
Minimum count 350

category_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

```

## Step 6: Measuring discount on a certain Item

```

bb['Discount']=bb['market_price']-bb['sale_price']
bb['Discount%']=bb['Discount']/bb['market_price']*100
bb

{
  "summary": {
    "name": "bb",
    "rows": 27555,
    "fields": [
      {
        "column": "index",
        "properties": {
          "dtype": "number",
          "std": 7954,
          "min": 1,
          "max": 27555,
          "num_unique_values": 27555,
          "samples": [
            16202,
            24963,
            18965
          ],
          "semantic_type": "\",
          "description": "\"
        }
      },
      {
        "column": "product",
        "properties": {
          "dtype": "string",
          "num_unique_values": 23540,
          "samples": [
            "Pure Impact Body Wash With Purifying Micro Particles - Shower Gel For Body, Face & Hair",
            "Premium Care XXL - 30 Diaper Pants",
            "Tomato - Green",
            "\n",
            "category",
            "properties": {
              "dtype": "category",
              "num_unique_values": 11,
              "samples": [
                "Snacks & Branded Foods",
                "Beauty & Hygiene",
                "Fruits & Vegetables",
                "\n",
                "sub_category",
                "properties": {
                  "dtype": "category",
                  "num_unique_values": 90,
                  "samples": [
                    "Crockery & Cutlery",
                    "Fragrances & Deos",
                    "Energy & Soft Drinks",
                    "\n",
                    "semantic_type": "\",
                    "description": "\n",
                    "column": "brand",
                    "properties": {
                      "dtype": "category",
                      "category": "\n",
                      "num_unique_values": 2313,
                      "samples": [
                        "Red Label",
                        "Lomani",
                        "Happily Unmarried",
                        "\n",
                        "semantic_type": "\",
                        "description": "\n"
                      ],
                      "column": "\n"
                    }
                  ]
                }
              ]
            }
          ]
        }
      }
    ]
  }
}

```

```
\"sale_price\", \n      \"properties\": {\n          \"dtype\":\n\"number\", \n          \"std\": 1201.9769904246652, \n          \"min\":\n2.45, \n          \"max\": 112475.0, \n          \"num_unique_values\":\n3261, \n          \"samples\": [\n              24.95, \n              40.0, \n5040.0\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\"\n      }, \n      \"column\":\n\"market_price\", \n      \"properties\": {\n          \"dtype\":\n\"number\", \n          \"std\": 581.7307174034938, \n          \"min\":\n3.0, \n          \"max\": 12500.0, \n          \"num_unique_values\": 1348,\n          \"samples\": [\n              192.0, \n              1164.0, \n74.38\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\"\n      }, \n      \"column\":\n\"type\", \n      \"properties\": {\n          \"dtype\": \"category\", \n          \"num_unique_values\": 426, \n          \"samples\": [\n              \"Nappies & Rash Cream\", \n              \"Hair Removal\", \n              \"Crackers & Digestive\"\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\"\n      }, \n      \"column\": \"rating\", \n      \"properties\": {\n          \"dtype\":\n\"number\", \n          \"std\": 0.6841238205552671, \n          \"min\":\n1.0, \n          \"max\": 5.0, \n          \"num_unique_values\": 486,\n          \"samples\": [\n              3.9428571428571426, \n              2.2666666666666666, \n              4.485714285714286\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\"\n      }, \n      \"column\": \"description\", \n      \"properties\": {\n          \"dtype\": \"string\", \n          \"num_unique_values\": 21944, \n          \"samples\": [\n              \"Premium Day Face Cream:Lightweight formula that will give your skin\na daily boost of radiance and hydration. Packed with the goodness of\nskin brightening, Vitamin C Why Is It Special?- It contains the Most\nStable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is\nthe next generation Vitamin C that really works in improving skin\nhealth.- Vitamin C is a proven skin-brightening agent known for its\nantioxidant properties for decades. It boosts collagen formation and\neffectively targets visible signs of ageing, giving skin the much-\nneeded glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3\nhelps awaken and clarify your complexion, revealing a natural healthy-\nlooking glow. - SPF 30 & PA+++ provide the highest level of protection\nagainst UVA & UVB rays of the sun. The mineral-based sun protection\nagents shield the skin against the harsh solar radiation, minimizing\nskin dryness and dullness. - Multiple natural oils are combined for\nintense hydration and nourishment. Antioxidant-rich botanical extracts\nare added to keep skin moisturized, smooth & soft. - St.Botanica is a\nPETA-certified cruelty-free brand. The product does not contain any\nharmful chemicals & is free from parabens, silicones or mineral\noils.\", \n              \"Kari Kari Japanese premium snacks are the\nexciting irresistible new taste of crunch. Lite and crispy, the\nsavoury rice crackers are mixed with peanuts to make hearty, wholesome\nsnack with delicious, bold and zesty taste with every bite.\", \n              \"Safely store your food in this pink lunch box. The lunch box is made
```

```

plastic which gives the container durability and sturdiness. The
rectangular containers come with an airtight lid to keep food fresh
for long. The wide mouth opening of the lunch makes it easy to fill
food and easy to clean. The set makes a great gift option for special
days.\n      ],\n      \\"semantic_type\\": \"\",\\n
\\\"description\\\": \"\"\n      }\\n    },\\n    {\\n      \\\"column\\\":
\\\"Discount\\\",\\n      \\\"properties\\\": {\\n        \\\"dtype\\\":
\\\"number\\\",\\n        \\\"std\\\": 1099.1979133300101,\\n        \\\"min\\\": -
110360.0,\\n        \\\"max\\\": 4320.0,\\n        \\\"num_unique_values\\\":
2517,\\n        \\\"samples\\\": [\\n          950.0,\\n          1464.99,\\n
30.120000000000005\\n        ],\\n        \\"semantic_type\\\": \"\",\\n
\\\"description\\\": \"\"\n      }\\n    },\\n    {\\n      \\\"column\\\":
\\\"Discount%\\\",\\n      \\\"properties\\\": {\\n        \\\"dtype\\\":
\\\"number\\\",\\n        \\\"std\\\": 420.5648399253825,\\n        \\\"min\\\": -
68813.95348837209,\\n        \\\"max\\\": 83.66666666666667,\\n
\\\"num_unique_values\\\": 3821,\\n        \\\"samples\\\": [\\n
36.07214428857716,\\n          30.05128205128205,\\n
22.142857142857142\\n        ],\\n        \\"semantic_type\\\": \"\",\\n
\\\"description\\\": \"\"\n      }\\n    }\\n  ]\\n}\",\\n  \"type\":\"dataframe\",\\n  \"variable_name\":\"bb\"}

```

The above function calculates the discount and discount % values on the products using sale\_price and market\_price values and displays them in the respective columns.

```

bb['Discount%']=bb['Discount%'].round(2)
bb

{
  "summary": {
    "name": "bb",
    "rows": 27555,
    "fields": [
      {
        "column": "index",
        "properties": {},
        "dtype": "number",
        "std": 7954,
        "min": 1,
        "max": 27555,
        "num_unique_values": 27555,
        "samples": [
          16202,
          24963,
          18965
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "product",
        "properties": {
          "dtype": "string"
        },
        "num_unique_values": 23540,
        "samples": [
          "Pure Impact Body Wash With Purifying Micro Particles - Shower Gel For Body, Face & Hair",
          "Premium Care XXL - 30 Diaper Pants",
          "Tomato - Green",
          "category",
          "category",
          "category",
          "category",
          "category",
          "category",
          "category",
          "category",
          "category"
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "category",
        "properties": {
          "dtype": "category"
        },
        "num_unique_values": 11,
        "samples": [
          "Snacks & Branded Foods",
          "Fruits & Vegetables",
          "Fragrances & Deos",
          "Crockery & Cutlery",
          "Beauty & Hygiene",
          "Sub Category"
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "sub_category",
        "properties": {
          "dtype": "category"
        },
        "num_unique_values": 90,
        "samples": [
          "Fragrances & Deos"
        ],
        "semantic_type": "",
        "description": ""
      }
    ]
  }
}

```

```

        ],\n      \"semantic_type\":\n    \"\",\\n      \"description\": \\"\\n      }\\n    },\\n      {\\n\n    \"column\": \"brand\",\\n      \"properties\": {\\n        \"dtype\":\n      \"category\",\\n      \"num_unique_values\": 2313,\n    \"samples\": [\n      \"Red Label\",\\n      \"Lomani\",\\n\n      \"Happily Unmarried\"\\n    ],\\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n    },\\n      {\\n\n      \"column\":\n      \"sale_price\",\\n      \"properties\": {\\n        \"dtype\":\n      \"number\",\\n        \"std\": 1201.9769904246652,\n        \"min\": 2.45,\n        \"max\": 112475.0,\n        \"num_unique_values\": 3261,\n      \"samples\": [\n        24.95,\n        40.0,\n      5040.0\\n    ],\\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n    },\\n      {\\n\n      \"column\":\n      \"market_price\",\\n      \"properties\": {\\n        \"dtype\":\n      \"number\",\\n        \"std\": 581.7307174034938,\n        \"min\": 3.0,\n        \"max\": 12500.0,\n        \"num_unique_values\": 1348,\n      \"samples\": [\n        192.0,\n        1164.0,\n      74.38\\n    ],\\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n    },\\n      {\\n\n      \"column\": \"type\",\\n      \"properties\": {\\n        \"dtype\": \"category\",\\n      \"num_unique_values\": 426,\n      \"samples\": [\n        \"Nappies & Rash Cream\",\\n        \"Hair Removal\",\\n\n        \"Crackers & Digestive\"\\n      ],\\n      \"semantic_type\":\n      \"\",\\n      \"description\": \"\\n      }\\n    },\\n      {\\n\n      \"column\": \"rating\",\\n      \"properties\": {\\n        \"dtype\": \"category\",\\n      \"number\",\\n        \"std\": 0.6841238205552671,\n        \"min\": 1.0,\n        \"max\": 5.0,\n        \"num_unique_values\": 486,\n      \"samples\": [\n        3.9428571428571426,\n      2.2666666666666666,\n        4.485714285714286\\n    ],\\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n    },\\n      {\\n\n      \"column\": \"description\",\\n      \"properties\": {\\n        \"dtype\": \"string\",\\n      \"num_unique_values\": 21944,\n      \"samples\": [\n\n        \"Premium Day Face Cream:Lightweight formula that will give your skin\n        a daily boost of radiance and hydration. Packed with the goodness of\n        skin brightening, Vitamin C Why Is It Special?- It contains the Most\n        Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is\n        the next generation Vitamin C that really works in improving skin\n        health.- Vitamin C is a proven skin-brightening agent known for its\n        antioxidant properties for decades. It boosts collagen formation and\n        effectively targets visible signs of ageing, giving skin the much-\n        needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3\n        helps awaken and clarify your complexion, revealing a natural healthy-\n        looking glow. - SPF 30 & PA+++ provide the highest level of protection\n        against UVA & UVB rays of the sun. The mineral-based sun protection\n        agents shield the skin against the harsh solar radiation, minimizing\n        skin dryness and dullness. - Multiple natural oils are combined for\n        intense hydration and nourishment. Antioxidant-rich botanical extracts\n        are added to keep skin moisturized, smooth & soft. - St.Botanica is a\n\n      ]\n    }\n  }\n}\n
```

```

PETA-certified cruelty-free brand. The product does not contain any
harmful chemicals & is free from parabens, silicones or mineral
oils.\n          \"Kari Kari Japanese premium snacks are the
exciting irresistible new taste of crunch. Lite and crispy, the
savoury rice crackers are mixed with peanuts to make hearty, wholesome
snack with delicious, bold and zesty taste with every bite.\",\n
\"Safely store your food in this pink lunch box. The lunch box is made
plastic which gives the container durability and sturdiness. The
rectangular containers come with an airtight lid to keep food fresh
for long. The wide mouth opening of the lunch makes it easy to fill
food and easy to clean. The set makes a great gift option for special
days.\n      ],\n      \"semantic_type\": \"\",\\n
\"description\": \"\"\n      }\n    },\\n    {\n      \"column\": \"Discount\",\\n      \"properties\": {\n        \"dtype\": \"number\",\\n        \"std\": 1099.1979133300101,\n        \"min\": -110360.0,\n        \"max\": 4320.0,\n        \"num_unique_values\": 2517,\n        \"samples\": [\n          950.0,\n          1464.99,\n          30.12000000000005\n        ],\n        \"semantic_type\": \"\",\\n\n        \"description\": \"\"\n      }\n    },\\n    {\n      \"column\": \"Discount%\",\\n      \"properties\": {\n        \"dtype\": \"number\",\\n        \"std\": 420.56482233441704,\n        \"min\": -68813.95,\n        \"max\": 83.67,\n        \"num_unique_values\": 2666,\n        \"samples\": [\n          10.66,\n          41.67,\n          8.84\n        ],\n        \"semantic_type\": \"\",\\n\n        \"description\": \"\"\n      }\n    }\n  ]\n},\"type\":\"dataframe\",\"variable_name\":\"bb\"}

```

## Step 7: Find out missing values from the Dataset

The .round(2) function restricts the display of digits upto two decimal points after the decimal in discount% column.

```

bb.isna().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
Discount%       6
dtype: int64

```

The .isna().sum() function gives the sum total of the records which have missing or null values. E.g. the above table shows that information about 1 no. product, 1 no. brand, 6 no. sale\_price, 8636 no. rating values, 115 no. description and 6 no. discount values are missing in the database.

```
bb.isnull()

{"summary": {"\n    \"name\": \"bb\", \n    \"rows\": 27555, \n    \"fields\": [\n        {\n            \"column\": \"index\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 1, \n                \"samples\": [\n                    false\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"product\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 2, \n                \"samples\": [\n                    true\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"category\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 1, \n                \"samples\": [\n                    false\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"sub_category\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 1, \n                \"samples\": [\n                    false\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"brand\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 2, \n                \"samples\": [\n                    true\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"sale_price\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 2, \n                \"samples\": [\n                    true\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"market_price\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 1, \n                \"samples\": [\n                    false\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"type\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 1, \n                \"samples\": [\n                    false\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"rating\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 2, \n                \"samples\": [\n                    true\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"description\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 2, \n                \"samples\": [\n                    true\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\\n\" \n            }\n        }, \n        {\n            \"column\": \"Discount\", \n            \"properties\": {\n                \"dtype\": \"boolean\", \n                \"num_unique_values\": 2, \n                \"samples\": [\n                    true\n                ]\n            }\n        }\n    ]\n}
```

```
],\n      \"semantic_type\": \"\",\n      \"description\": \"\"\n    },\n    {\n      \"column\": \"Discount%\",\n      \"properties\": {\n        \"dtype\": \"boolean\",\n        \"num_unique_values\": 2,\n        \"samples\": [\n          true\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    }\n  ]\n},\"type\":\"dataframe\"}
```

The .isnull() function displays the whole dataset as False (no null value) and True (having null value)

```
bb['product'].isnull().sum()\n1
```

The above .isnull().sum() function counts the no. of records in products column having null values and displays the result.

```
missing_product = bb[bb['product'].isnull()]\nmissing_product\n{"repr_error":"0","type":"dataframe","variable_name":"missing_product"}
```

The above code displays the rows having missing product value in the database

```
missing_product.nunique() # Check how many unique values exist in\n# each column for missing product rows\n\nindex      1\nproduct     0\ncategory    1\nsub_category 1\nbrand       1\nsale_price   1\nmarket_price 1\ntype         1\nrating       1\ndescription  1\nDiscount     1\nDiscount%    1\ndtype: int64
```

The .nunique() function displays the no. of unique values in other columns of the database for rows which have missing product value. E.g. every column has a single unique value (non null) corresponding to the row which has missing product information.

```
bb['category'].isnull().sum()\n0
```

```

bb['sub_category'].isnull().sum()

0

bb['brand'].isnull().sum()

1

missing_brand = bb[bb['brand'].isnull()]
missing_brand

{"repr_error": "0", "type": "dataframe", "variable_name": "missing_brand"}

missing_brand.unique() # Check how many unique values exist in each
# column for missing brand rows

index      1
product    1
category   1
sub_category 1
brand      0
sale_price 1
market_price 1
type       1
rating     0
description 0
Discount   1
Discount%  1
dtype: int64

```

The .nunique() function displays the no. of unique values in other columns of the database for rows which have missing brand value. E.g. every column has a single unique value (non null) other than rating and description columns which have null values corresponding to the row which has missing brand information.

```

bb['sale_price'].isna().sum()

0

missing_sale_price = bb[bb['sale_price'].isnull()]
missing_sale_price

{"repr_error": "Out of range float values are not JSON compliant:
nan", "type": "dataframe", "variable_name": "missing_sale_price"}

missing_sale_price.unique() # Check how many unique values exist in
# each column for missing sale_price rows

index      0
product    0
category   0
sub_category 0

```

```

brand          0
sale_price     0
market_price   0
type           0
rating          0
description    0
Discount        0
Discount%      0
dtype: int64

```

The .nunique() function displays the no. of unique values in other columns of the database for rows which have missing sale\_price value. E.g. corresponding to 6 missing sale\_price rows, category has 4 unique values (4 different category values), rating has 3 unique values (3 different rating values) whereas the discount and discount% columns have null values simply because the corresponding sale\_price column has null values.

```

# Fill missing values in sale_price with corresponding market_price
values
bb['sale_price'] = bb['sale_price'].fillna(bb['market_price'])

bb['sale_price'].isna().sum()

0

bb['market_price'].isna().sum()

0

bb['type'].isna().sum()

0

bb['rating'].isna().sum()

8636

missing_rating = bb[bb['rating'].isnull()]
missing_rating

{
  "summary": {
    "name": "missing_rating",
    "rows": 8636,
    "fields": [
      {
        "column": "index",
        "properties": {
          "dtype": "number",
          "std": 7957,
          "min": 56,
          "max": 27532,
          "num_unique_values": 8636,
          "samples": [5108, 8520, 19439]
        }
      },
      {
        "column": "product",
        "properties": {
          "dtype": "string",
          "num_unique_values": 7646,
          "samples": [
            "Coconut Cookies",
            "Juice - Guava",
            "Bright I - Anti Dark Circle Cream",
            "Aloeevera"
          ]
        }
      }
    ],
    "semantic_type": []
  }
}

```

```

  "description": """
    },
    {
      "column": "category",
      "properties": {
        "dtype": "category",
        "num_unique_values": 11,
        "samples": [
          "Fruits & Vegetables",
          "Snacks & Branded Foods",
          "Beauty & Hygiene",
          "Snacks & Branded Foods"
        ],
        "semantic_type": "category",
        "description": """
          },
          {
            "column": "sub_category",
            "properties": {
              "dtype": "category",
              "num_unique_values": 90,
              "samples": [
                "Men's Grooming",
                "Drinks & Beverages",
                "Organic Staples"
              ],
              "semantic_type": "category",
              "description": """
                },
                {
                  "column": "brand",
                  "properties": {
                    "dtype": "category",
                    "num_unique_values": 1183,
                    "samples": [
                      "Daucy",
                      "Mama",
                      "Iyers"
                    ],
                    "semantic_type": "category",
                    "description": """
                      },
                      {
                        "column": "sale_price",
                        "properties": {
                          "dtype": "number",
                          "min": 2.45,
                          "max": 1824.0898131519548,
                          "num_unique_values": 1766,
                          "samples": [
                            548.8,
                            65.0,
                            824.98
                          ],
                          "semantic_type": "number",
                          "description": """
                            },
                            {
                              "column": "market_price",
                              "properties": {
                                "dtype": "number",
                                "min": 5.0,
                                "max": 768.8466234961703,
                                "num_unique_values": 967,
                                "samples": [
                                  134.0,
                                  658.0
                                ],
                                "semantic_type": "number",
                                "description": """
                                  },
                                  {
                                    "column": "type",
                                    "properties": {
                                      "dtype": "category",
                                      "num_unique_values": 379,
                                      "samples": [
                                        "Bakeware Moulds, Cutters",
                                        "Choco & Nut Spread",
                                        "Gift Boxes"
                                      ],
                                      "semantic_type": "category",
                                      "description": """
                                        },
                                        {
                                          "column": "rating",
                                          "properties": {
                                            "dtype": "number",
                                            "std": null,
                                            "min": null,
                                            "max": null,
                                            "num_unique_values": 0,
                                            "samples": [],
                                            "semantic_type": "number",
                                            "description": """
                                              },
                                              {
                                                "column": "discount",
                                                "properties": {
                                                  "dtype": "string",
                                                  "num_unique_values": 7058,
                                                  "samples": [],
                                                  "semantic_type": "string",
                                                  "description": """
                                                    },
                                                    {
                                                      "column": "discount",
                                                      "properties": {
                                                        "dtype": "number",
                                                        "std": 1683.3106958243072,
                                                        "min": -110360.0,
                                                        "max": 4320.0,
                                                        "num_unique_values": 1252,
                                                        "samples": [],
                                                        "semantic_type": "number",
                                                        "description": """
                                                          },
                                                          {
                                                            "column": "discount_percent",
                                                            "properties": {
                                                              "dtype": "number",
                                                              "std": 100.04631048507072,
                                                              "min": -1
                                                            }
                                                          }
                                                        }
                                                      }
                                                    }
                                                  }
                                                }
                                              }
                                            }
                                          }
                                        }
                                      }
                                    }
                                  }
                                }
                              }
                            }
                          }
                        }
                      }
                    }
                  }
                }
              }
            }
          }
        }
      }
    }
  }
}

```

```

8561.68,\n      \\"max\\": 83.67,\n      \\"num_unique_values\\":\n1172,\n      \\"samples\\": [],\n      \\"semantic_type\\": \"\",\\n\n      \\"description\\": \"\\\"\\n      }\\n    }\\n  ]\\n}\",\"type\":\"dataframe\",\"variable_name\":\"missing_rating\"}

missing_rating.nunique()

index          8636
product        7646
category        11
sub_category     90
brand          1183
sale_price       1766
market_price      967
type            379
rating           0
description      7058
Discount         1252
Discount%        1172
dtype: int64

```

The useful information that can be derived from the above code is that corresponding to 8636 null rating values, there are 7646 unique products among 11 unique categories, 90 unique sub categories and 1183 unique brands.

```

missing_rating_products = bb[bb['rating'].isnull()]['product',
'brand', 'sale_price']
missing_rating_products

{"summary":{"\n  \"name\": \"missing_rating_products\",\n  \"rows\":\n8636,\n  \"fields\": [\n    {\n      \"column\": \"product\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 7646,\n        \"samples\": [\n          \"Coconut Cookies\",\n          \"Bright I - Anti Dark Circle\nCream\",\n          \"Juice - Guava, Aloevera\"\n        ],\n        \"semantic_type\": \"\",\\n        \"description\": \"\\\"\\n      }\n      },\n      {\n        \"column\": \"brand\",\n        \"properties\": {\n          \"dtype\": \"category\",\n          \"num_unique_values\":\n1183,\n          \"samples\": [\n            \"Daucy\",\n            \"Mama\",\n            \"Iyers\"\n          ],\n          \"semantic_type\": \"\",\\n          \"description\": \"\\\"\\n      }\n          },\n          {\n            \"column\": \"sale_price\",\n            \"properties\": {\n              \"dtype\": \"number\",\n              \"std\":\n1824.0898131519548,\n              \"min\": 2.45,\n              \"max\":\n112475.0,\n              \"num_unique_values\": 1766,\n              \"samples\": [\n                548.8,\n                65.0,\n                824.98\n              ],\n              \"semantic_type\": \"\",\\n              \"description\": \"\\\"\\n      }\n                }\n              ],\n              \"type\":\"dataframe\",\"variable_name\":\"missing_rating_products\"}

```

The above function displays the information about products, brands and sale price of products having missing rating values.

```
missing_rating_products_total_sale_price =  
missing_rating_products.groupby('product')['sale_price'].sum()  
missing_rating_products_total_sale_price =  
missing_rating_products_total_sale_price.sort_values(ascending=False)  
print(missing_rating_products_total_sale_price)  
  
print("The total sale price of products with missing rating is : ",  
missing_rating_products_total_sale_price.sum().round(2))  
  
product  
Beard Kit 112475.0  
4mm Aluminium Induction Base Chapati Roti Tawa - Silver 112178.0  
Extra Virgin Olive Oil 15574.5  
Olive Oil - Extra Virgin 14824.0  
Bravura Clipper 12500.0  
...  
Instant Popcorn - Seven Spices 9.0  
Tomato - Local 9.0  
Marigold Flower - Orange 6.0  
Tomato - Green 6.0  
Exam Standard Scale 5.0  
Name: sale_price, Length: 7646, dtype: float64  
The total sale price of products with missing rating is : 4041097.14
```

The above function displays the information about the sale price of products having missing rating values and the total sale price of these products.

```
missing_rating_brands_total_sale_price =  
missing_rating_products.groupby('brand')['sale_price'].sum()  
missing_rating_brands_total_sale_price =  
missing_rating_brands_total_sale_price.sort_values(ascending=False)  
missing_rating_brands_total_sale_price  
  
brand  
HAZEL 143751.00  
Uncle Tony 121393.25  
Dkny 110875.00  
bb Combo 79146.79  
bb Royal 73821.51  
...  
THIRUMALA 22.50  
7 Up Nimbooz 20.00  
Appy Fizz 20.00  
Jaji 20.00  
Bindu 17.00  
Name: sale_price, Length: 1183, dtype: float64
```

```
# Replace missing rating values using interpolation (linear)
bb['rating'] = bb['rating'].interpolate(method='linear')

bb['rating'].isna().sum()

0

bb['description'].isna().sum()

115

missing_description = bb[bb['description'].isnull()]
missing_description

{"repr_error": "0", "type": "dataframe", "variable_name": "missing_description"}

bb['Discount'].isna().sum()

0

missing_Discount = bb[bb['Discount'].isnull()]
missing_Discount

{"repr_error": "Out of range float values are not JSON compliant: nan", "type": "dataframe", "variable_name": "missing_Discount"}

bb

{"summary": "{\n    \"name\": \"bb\", \n    \"rows\": 27555, \n    \"fields\": [\n        {\n            \"column\": \"index\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 7954, \n                \"min\": 1, \n                \"max\": 27555, \n                \"num_unique_values\": 27555, \n                \"samples\": [\n                    16202, \n                    24963, \n                    18965\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"product\", \n            \"properties\": {\n                \"dtype\": \"string\", \n                \"num_unique_values\": 23540, \n                \"samples\": [\n                    \"Pure Impact Body Wash With Purifying Micro Particles - Shower Gel\nFor Body, Face & Hair\", \n                    \"Premium Care XXL - 30 Diaper\nPants\", \n                    \"Tomato - Green\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        }, \n        {\n            \"column\": \"category\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 11, \n                \"samples\": [\n                    \"Snacks & Branded Foods\", \n                    \"Beauty & Hygiene\", \n                    \"Fruits & Vegetables\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        }, \n        {\n            \"column\": \"sub_category\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 90, \n                \"samples\": [\n                    \"Crockery & Cutlery\", \n                    \"Fragrances & Deos\", \n                    \"Energy & Soft Drinks\"\n                ], \n                \"semantic_type\": \"\"\n            }\n        }\n    ]\n}"}
```

```
\\",\n      "description": "\\n      }\n    },\n    {\n      \"column\": \"brand\",\n        \"properties\": {\n          \"num_unique_values\": 2313,\n          \"samples\": [\n            \"Red Label\",\n            \"Lomani\",\n            \"Happily Unmarried\"],\n            \"semantic_type\": \"\",\n            \"description\": \"\\n      }\n    },\n    {\n      \"column\": \"sale_price\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 1201.9769904246652,\n          \"min\": 2.45,\n          \"max\": 112475.0,\n          \"num_unique_values\": 3261,\n          \"samples\": [\n            24.95,\n            40.0,\n            5040.0\n          ],\n            \"semantic_type\": \"\",\n            \"description\": \"\\n      }\n    },\n    {\n      \"column\": \"market_price\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 581.7307174034938,\n          \"min\": 3.0,\n          \"max\": 12500.0,\n          \"num_unique_values\": 1348,\n          \"samples\": [\n            192.0,\n            1164.0,\n            74.38\n          ],\n            \"semantic_type\": \"\",,\n            \"description\": \"\\n      }\n    },\n    {\n      \"column\": \"type\",\n        \"properties\": {\n          \"dtype\": \"category\",\n          \"num_unique_values\": 426,\n          \"samples\": [\n            \"Nappies & Rash Cream\",\n            \"Hair Removal\",\n            \"Crackers & Digestive\"\n          ],\n            \"semantic_type\": \"\",,\n            \"description\": \"\\n      }\n    },\n    {\n      \"column\": \"rating\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 0.6841238205552671,\n          \"min\": 1.0,\n          \"max\": 5.0,\n          \"num_unique_values\": 486,\n          \"samples\": [\n            3.9428571428571426,\n            2.2666666666666666,\n            4.485714285714286\n          ],\n            \"semantic_type\": \"\",,\n            \"description\": \"\\n      }\n    },\n    {\n      \"column\": \"description\",\n        \"properties\": {\n          \"dtype\": \"string\",\n          \"num_unique_values\": 21944,\n          \"samples\": [\n            \"Premium Day Face Cream:Lightweight formula that will give your skin a daily boost of radiance and hydration. Packed with the goodness of skin brightening, Vitamin C Why Is It Special?- It contains the Most Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is the next generation Vitamin C that really works in improving skin health.- Vitamin C is a proven skin-brightening agent known for its antioxidant properties for decades. It boosts collagen formation and effectively targets visible signs of ageing, giving skin the much-needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3 helps awaken and clarify your complexion, revealing a natural healthy-looking glow. - SPF 30 & PA+++ provide the highest level of protection against UVA & UVB rays of the sun. The mineral-based sun protection agents shield the skin against the harsh solar radiation, minimizing skin dryness and dullness. - Multiple natural oils are combined for intense hydration and nourishment. Antioxidant-rich botanical extracts are added to keep skin moisturized, smooth & soft. - St.Botanica is a PETA-certified cruelty-free brand. The product does not contain any\n          }\n        }\n      }\n    }\n  }\n}\n
```

```

harmful chemicals & is free from parabens, silicones or mineral
oils.\n          \"Kari Kari Japanese premium snacks are the
exciting irresistible new taste of crunch. Lite and crispy, the
savoury rice crackers are mixed with peanuts to make hearty, wholesome
snack with delicious, bold and zesty taste with every bite.\",\n
\"Safely store your food in this pink lunch box. The lunch box is made
plastic which gives the container durability and sturdiness. The
rectangular containers come with an airtight lid to keep food fresh
for long. The wide mouth opening of the lunch makes it easy to fill
food and easy to clean. The set makes a great gift option for special
days.\n      ],\n      \"semantic_type\": \"\",\\n
\"description\": \"\"\n      },\\n      {\n        \"column\":
\"Discount\",\\n        \"properties\": {\n          \"dtype\":
\"number\",\\n          \"std\": 1099.1979133300101,\n          \"min\": -110360.0,\n          \"max\": 4320.0,\n          \"num_unique_values\": 2517,\n          \"samples\": [\n            950.0,\n            1464.99,\n            30.12000000000005\n          ],\n          \"semantic_type\": \"\",\\n
\"description\": \"\"\n      },\\n      {\n        \"column\":
\"Discount%\",\\n        \"properties\": {\n          \"dtype\":
\"number\",\\n          \"std\": 420.56482233441704,\n          \"min\": -68813.95,\n          \"max\": 83.67,\n          \"num_unique_values\": 2666,\n          \"samples\": [\n            10.66,\n            41.67,\n            8.84\n          ],\n          \"semantic_type\": \"\",\\n
\"description\": \"\"\n      }\n    }\n  }\n}, \"type\": \"dataframe\", \"variable_name\": \"bb\"}

len(bb)

27555

bb.shape[0]

27555

bb.isnull().sum()

index          0
product         1
category        0
sub_category    0
brand           1
sale_price      0
market_price    0
type            0
rating          0
description     115
Discount         0
Discount%        0
dtype: int64

```

## Step 8: Find out Outliers from the Dataset according to the columns and fill them with the mean

```
bb.iloc[0:2]
```

```
{"summary": {"\n    \"name\": \"bb\",\n    \"rows\": 2,\n    \"fields\": [\n        {\n            \"column\": \"index\",\n            \"properties\": {\n                \"dtype\": \"number\",\n                \"std\": 0,\n                \"min\": 1,\n                \"max\": 2,\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    2,\n                    1\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"product\",\n            \"properties\": {\n                \"dtype\": \"string\",\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    \"Water Bottle - Orange\",\n                    \"Garlic Oil -\nVegetarian Capsule 500 mg\"\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"category\",\n            \"properties\": {\n                \"dtype\": \"string\",\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    \"Kitchen, Garden & Pets\",\n                    \"Beauty &\nHygiene\"\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"sub_category\",\n            \"properties\": {\n                \"dtype\": \"string\",\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    \"Storage & Accessories\",\n                    \"Hair Care\"\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"brand\",\n            \"properties\": {\n                \"dtype\": \"string\",\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    \"Mastercook\",\n                    \"Sri Sri\nAyurveda\"\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"market_price\",\n            \"properties\": {\n                \"dtype\": \"number\",\n                \"std\": 28.284271247461902,\n                \"min\": 180.0,\n                \"max\": 220.0,\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    180.0,\n                    220.0\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"type\",\n            \"properties\": {\n                \"dtype\": \"string\",\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    \"Water & Fridge Bottles\",\n                    \"Hair Oil &\nSerum\"\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        },\n        {\n            \"column\": \"rating\",\n            \"properties\": {\n                \"dtype\": \"number\",\n                \"std\": 1.2727922061357855,\n                \"min\": 2.3,\n                \"max\": 4.1,\n                \"num_unique_values\": 2,\n                \"samples\": [\n                    2,\n                    1\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\\n            }\n        }\n    ]\n}
```

```

2.3,\n        4.1\n            ],\n            \"semantic_type\": \"\",\\n
\"description\": \"\\n            }\\n        },\\n        {\\n            \"column\":\\n            \"description\":,\\n            \"properties\": {\\n                \"dtype\":\\n                \"string\",\\n                \"num_unique_values\": 2,\\n                \"samples\":\\n                [\\n                    \"Each product is microwave safe (without lid),\\nrefrigerator safe, dishwasher safe and can also be used for re-heating\\nfood and not for cooking. All containers come with airtight lids and a\\nwide variety of attractive colours. Stack these stylish and colourful\\ncontainers in your kitchen with ease and for a look-good factor.\",\\n                \"This Product contains Garlic Oil that is known to help proper\\ndigestion, maintain proper cholesterol levels, support cardiovascular\\nand also build immunity. For Beauty tips, tricks & more visit\\n                https://bigbasket.blog/\\n            ],\\n            \"semantic_type\": \"\",\\n            \"description\": \"\\n            }\\n        },\\n        {\\n            \"column\":\\n            \"Discount\":,\\n            \"properties\": {\\n                \"dtype\":\\n                \"number\",\\n                \"std\": 0.0,\\n                \"min\": 0.0,\\n                \"max\": 0.0,\\n                \"num_unique_values\": 1,\\n                \"samples\":\\n                [\\n                    0.0\\n                ],\\n                \"semantic_type\": \"\",\\n                \"description\": \"\\n            }\\n        },\\n        {\\n            \"column\":\\n            \"Discount%\":,\\n            \"properties\": {\\n                \"dtype\":\\n                \"number\",\\n                \"std\": 0.0,\\n                \"min\": 0.0,\\n                \"max\": 0.0,\\n                \"num_unique_values\": 1,\\n                \"samples\":\\n                [\\n                    0.0\\n                ],\\n                \"semantic_type\": \"\",\\n                \"description\": \"\\n            }\\n        }\\n    ]\\n},\\n    \"type\": \"dataframe\"\n}
bb.loc[:, ["brand","type"]].head(10)

{
    "summary": {
        "name": "bb",
        "rows": 10,
        "fields": [
            {
                "column": "brand",
                "properties": {
                    "dtype": "string",
                    "num_unique_values": 10,
                    "samples": [
                        "StBotanica",
                        "Mastercook",
                        "Nature Protect",
                        "Shampoo & Conditioner",
                        "Water & Fridge Bottles",
                        "Disinfectant Spray & Cleaners"
                    ],
                    "semantic_type": "",
                    "description": ""
                }
            }
        ]
    },
    "type": "dataframe"
}

bb.loc[50 : 100]

{
    "summary": {
        "name": "bb",
        "rows": 51,
        "fields": [
            {
                "column": "index",
                "properties": {
                    "dtype": "number",
                    "std": 14,
                    "min": 51,
                    "max": 101,
                    "num_unique_values": 51,
                    "samples": [
                        94,
                        91,
                        97
                    ],
                    "semantic_type": "",
                    "description": ""
                }
            }
        ]
    },
    "type": "dataframe"
}

```

```

n      \\"num_unique_values\\": 51,\n          \\"samples\\": [\n\\\"Namkeen - Madras Mixture\\\",\\n          \\"Ashwagandha Tablets\\\",\\n\\\"Hair Remover Spray - Foam Lemon\\\",\\n          ],\\n\\\"semantic_type\\\": \\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"category\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"category\\\",\\n              \\"num_unique_values\\\":\n9,\n              \\"samples\\\": [\n                  \\"Fruits & Vegetables\\\",\\n                  \\"Gourmet & World Food\\\",\\n                  \\"Cleaning & Household\\\"\n              ],\\n          \\"semantic_type\\\": \\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"sub_category\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"string\\\",\\n              \\"num_unique_values\\\": 34,\n              \\"samples\\\": [\n                  \\"Snacks, Dry Fruits, Nuts\\\",\\n                  \\"Pasta, Soup & Noodles\\\",\\n                  \\"Cookware & Non Stick\\\"\n              ],\\n          \\"semantic_type\\\":\n\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"brand\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"string\\\",\\n              \\"num_unique_values\\\": 45,\n              \\"samples\\\": [\n                  \\"Brut\\\",\\n                  \\"Fresho Signature\\\",\\n                  \\"Sanjay \\n          \"],\\n          \\"semantic_type\\\": \\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"sale_price\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"number\\\",\\n              \\"std\\\": 292.4805648233273,\n              \\"min\\\": 10.0,\n              \\"max\\\": 1266.38,\n              \\"num_unique_values\\\": 49,\n              \\"samples\\\": [\n                  35.0,\n                  32.55,\n                  1266.38\n              ],\\n          \\"semantic_type\\\": \\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"market_price\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"number\\\",\\n              \\"std\\\": 308.1217724645025,\n              \\"min\\\": 10.0,\n              \\"max\\\": 1398.0,\n              \\"num_unique_values\\\": 38,\n              \\"samples\\\": [\n                  249.0,\n                  1398.0,\n                  225.0\n              ],\\n          \\"semantic_type\\\":\n\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"type\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"string\\\",\\n              \\"num_unique_values\\\": 44,\n              \\"samples\\\": [\n                  \\"Namkeen & Savoury Snacks\\\",\\n                  \\"Toothbrush\\\",\\n                  \\"Chips & Corn Snacks\\\"\n              ],\\n          \\"semantic_type\\\": \\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"rating\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"number\\\",\\n              \\"std\\\": 0.6911131181275106,\n              \\"min\\\": 1.0,\n              \\"max\\\": 5.0,\n              \\"num_unique_values\\\": 25,\n              \\"samples\\\": [\n                  3.5,\n                  4.04,\n                  4.1\n              ],\\n          \\"semantic_type\\\": \\"\\\",\\n          \\"description\\\": \\"\\n          \"},\\n      {\n          \\"column\\\": \\"description\\\",\\n          \\"properties\\\": {\n              \\"dtype\\\": \\"string\\\",\\n              \\"num_unique_values\\\": 50,\n              \\"samples\\\": [\n                  \\"Bisk Farm Sugar Free Petit Beurre is one of the best\nsugar free petit beurre for diabetic patients and for other health\nissues. They have a good flavor and these biscuits are made of salty,\nspice and sweet. It is hygienically baked and well packed in wrappers\nthat makes them long lasting and remain fresh giving crispy taste. It\n"

```

```

is crispy and gives a superior taste.\",\n          \"Muniyal
Ayurveda\u00a0Ashwagandha Tablet contains the Ayurveda herb,
Ashwagandha, which has worked wonders with patients of cardiac and
neural disorders for centuries. The tablet strengthens the nervous
system and enables its users to handle stress better by elevating
their stress threshold considerably. It is effective against
hypertension and is also beneficial in renal complaints because of the
diuretic properties of its ingredients. Its anti-inflammatory
properties reduce the risk of fever, swelling and tenderness.\",\n
\"Batchelor's Cup a Soup with Croutons Cream of Asparagus: Serving
Suggestion: 1. Empty a Sachet into a cup or mug, 2. Add 230ml of
boiling water, 3. Stir well, wait a few moments, 4. Sit back and
enjoy. Ingredients: As Served (greatest first) Water, Asparagus (5%),
Maize Starch, Glucose Syrup, Vegetable Oil, Croutons (1.3), (Wheat
Flour, Vegetable Oil, Salt, Yeast, Rosemary Extract), Lactose, Potato
Starch, Salt, Onion Powder, Milk Proteins, Flavouring, Flavours
Enhancers (Monosodium Glutamate, Disodium 5-ribonucleotides),
Emulsifier (Mono-and Diglycerides of Fatty Acids), Acidity Regulator
(Dipotassium Phosphate), Yeast Extract (Contain Barley), Turmeric,
Natural Flavouring (Contains Celery).\n      ],\n      \"semantic_type\": \"\",\n      \"description\": \"\\n      \",\n      {\n        \"column\": \"Discount\",\\n        \"properties\": {\n          \"dtype\": \"number\",\\n          \"std\": 57.74555566770961,\n          \"min\": 0.0,\n          \"max\": 199.6000000000002,\n          \"num_unique_values\": 32,\n          \"samples\": [\n            2.45000000000003,\n            51.0,\n            44.84999999999994,\n            15.403400711862172,\n            0.0,\n            54.85,\n            \"samples\": [\n              11.0,\n              48.19,\n              23.85\n            ],\n            \"semantic_type\": \"\",\\n            \"description\": \"\\n            \",\n            {\n              \"column\": \"Discount%\",\\n              \"properties\": {\n                \"dtype\": \"number\",\\n                \"std\": 15.403400711862172,\n                \"min\": 0.0,\n                \"max\": 54.85,\n                \"num_unique_values\": 20,\n                \"samples\": [\n                  11.0,\n                  48.19,\n                  23.85\n                ],\n                \"semantic_type\": \"\",\\n                \"description\": \"\\n                \",\n                {\n                  \"column\": \"product\",\\n                  \"properties\": {\n                    \"dtype\": \"string\",\\n                    \"num_unique_values\": 13087,\n                    \"samples\": [\n                      \"Hair Conditioner - Colour Protection\",\\n                      \"Organic -\n                      Poha\",\\n                      \"Baby Shampoo - Gentle Cleansing, 0-5 years\"\n                    ],\n                    \"semantic_type\": \"\",\\n                    \"description\": \"\\n                    \",\n                    {\n                      \"column\": \"category\",\\n

```

```

  "properties": {
    "category": {
      "dtype": "category",
      "samples": [
        "Beauty & Hygiene",
        "Organic Staples"
      ],
      "semantic_type": "\\",
      "column": "category"
    },
    "sub_category": {
      "properties": {
        "category": {
          "dtype": "category",
          "num_unique_values": 89,
          "samples": [
            "Ready To Cook & Eat",
            "Sausages, Bacon & Salami"
          ],
          "semantic_type": "\\",
          "column": "category"
        },
        "brand": {
          "properties": {
            "Vanish": {
              "Delfi Twister"
            }
          },
          "semantic_type": "\\",
          "column": "brand"
        },
        "sale_price": {
          "properties": {
            "std": 1053.1468166876825,
            "min": 5.0,
            "max": 111649.0,
            "num_unique_values": 2504,
            "samples": [
              6.0,
              205.0,
              150.1
            ],
            "semantic_type": "\\",
            "column": "sale_price"
          },
          "semantic_type": "\\",
          "column": "sale_price"
        },
        "market_price": {
          "properties": {
            "std": 584.4746584407657,
            "min": 5.0,
            "max": 12500.0,
            "num_unique_values": 1068,
            "samples": [
              74.0,
              855.0,
              2495.0
            ],
            "semantic_type": "\\",
            "column": "market_price"
          },
          "semantic_type": "\\",
          "column": "market_price"
        },
        "type": {
          "properties": {
            "Jam, Conserve, Marmalade": {
              "Banana, Sapota & Papaya"
            },
            "Floor & Other Cleaners": {
              "rating"
            }
          },
          "semantic_type": "\\",
          "column": "type"
        },
        "rating": {
          "properties": {
            "std": 0.2728832189242142,
            "min": 4.014285714285714,
            "max": 5.0,
            "num_unique_values": 173,
            "samples": [
              4.183333333333333,
              4.22,
              4.52
            ],
            "semantic_type": "\\",
            "column": "rating"
          },
          "semantic_type": "\\",
          "column": "rating"
        },
        "description": {
          "properties": {
            "string": {
              "This hairspray made out of mist formula is specially formulated to offer the combination of a fine controlled mist with super setting properties. It gives the ultimate hold without stickiness.",
              "Giving your skin an oil overhaul doesn't have to be a drag. Oil stuck in your pores is what makes your skin feel oily again after a wash! POND'S Clay Foam and Mask is the most fun way to say goodbye oil and hello to an all-day matte glow. Made with 100% natural Moroccan"
            }
          },
          "semantic_type": "\\",
          "column": "description"
        }
      }
    }
  }
}

```

clay that has 4x oil absorption power, it sucks out dirt and oil stuck deep within your pores. What's left behind? Skin that's glowing and matte all day long! Pond's Clay Foam is the most enjoyable and effective way to keep your skin oil-free for longer. Revolutionise face washing with the enriching power of Mineral Clay. One of the most efficacious ingredients in deep cleansing. Its enriched with skin-loving minerals to give you a bouncy glow. So, step up your deep cleansing regimen for an oil-free glow. The clay range comes in two exciting formats. The Pond's white beauty mineral clay foam brightens and smoothens your skin for an oil-free glow! Combined with Vitamin B3+ technology, it stimulates the appearance of millions of pristine skin cells to reveal your bright true shade. The Pond's pure white mineral clay foam is enriched with activated charcoal which gently exfoliates and removes dead cells. Detoxes your skin for an oil-free glow! Top that up with the Clay masks that can absorb the oil from deep within, leaving you feeling matte and glowing!\",\n

\"Now with the best ever Pantene, get hair that\u2019s stronger inside, shinier outside. Pantene anti-dandruff shampoo gently cleanses hair without drying the scalp, with weightless moisturisers that leave hair healthy and shiny. Enhanced with Zinc Pyrithione, it helps prevent reoccurrence of dandruff. His formula cleanses hair without drying the scalp. Prevents reoccurrence of dandruff.\u00a0Hair feels healthier and shinier with regular Pantene system use vs non-conditioning shampoo based on test results under laboratory conditions. For Beauty tips, tricks & more visit

```
https://bigbasket.blog/"\n      ],\n      \"semantic_type\":\n      \"\",\\n      \"description\": \"\"\n      },\\n      {\n      \"column\": \"Discount\",\\n      \"properties\": {\n      \"number\",\\n      \"std\": 936.4749714952011,\\n      \"min\": -110360.0,\\n      \"max\": 4320.0,\\n      \"num_unique_values\": 1887,\\n      \"samples\": [\n      16.400000000000006,\\n      30.18000000000007,\\n      610.0\n      ],\\n      \"semantic_type\": \"\",\\n      \"description\": \"\"\n      },\\n      {\n      \"column\": \"Discount%\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\": 90.83377096864153,\\n      \"min\": -8561.68,\\n      \"max\": 79.24,\\n      \"num_unique_values\": 1782,\\n      \"samples\": [\n      9.46,\\n      13.16,\\n      20.11\n      ],\\n      \"semantic_type\": \"\",\\n      \"description\": \"\"\n      }\n      }\n    ],\"type\":\"dataframe\"}
```

```
bb.loc[bb['rating'] > 4, ['product',\n'sale_price']].sort_values(by='sale_price', ascending=False)
```

```
{"summary": "{\n  \"name\": \"bb\",\\n  \"rows\": 14880,\\n  \"fields\": [\n    {\n      \"column\": \"product\",\\n      \"properties\": {\n        \"dtype\": \"string\",\\n        \"num_unique_values\": 13087,\\n        \"samples\": [\n          \"Energise Assam Tea\",\\n          \"Face Wash - Blush & Glow Strawberry Gel\",\\n          \"The Protein Bar - Cocoa 35 g + The Fibre Bar - Mixed Fruits 35 g\"\n        ]\n      }\n    }\n  ]\n}
```

```

    \"semantic_type\": \"\",\\n          \"description\": \"\"\n      }\\n    },\\n    {\\n      \"column\": \"sale_price\",\\n      \"properties\": {\\n        \"dtype\": \"number\",\\n        \"std\": 1053.1468166876823,\\n        \"min\": 5.0,\\n        \"max\": 111649.0,\\n        \"num_unique_values\": 2504,\\n        \"samples\": [\\n          141.0,\\n          1420.0,\\n          199.0\\n        ],\\n      }\\n    }\\n  ]\\n},\"type\":\"dataframe\"}

bb.loc[bb['rating'] > np.average(bb['rating'])], ['product', 'sale_price', 'rating']).sort_values(by='rating', ascending=False)

{ "summary": {\\n  \"name\": \"bb\",\\n  \"rows\": 17648,\\n  \"fields\": [\\n    {\\n      \"column\": \"product\",\\n      \"properties\": {\\n        \"dtype\": \"string\",\\n        \"num_unique_values\": 15359,\\n        \"samples\": [\\n          \"Care & Protect Moisturising Cream Beauty Bathing Bar\",\\n          \"Beard Colour - Brown Black, For Men, 102\",\\n          \"Tampons Easy Twist - Regular\\n        ],\\n        \"semantic_type\": \"\",\\n        \"description\": \"\"\n      }\\n    },\\n    {\\n      \"column\": \"sale_price\",\\n      \"properties\": {\\n        \"dtype\": \"number\",\\n        \"std\": 986.8317439758428,\\n        \"min\": 5.0,\\n        \"max\": 111649.0,\\n        \"num_unique_values\": 2700,\\n        \"samples\": [\\n          183.0,\\n          160.2,\\n          498.0\\n        ],\\n        \"semantic_type\": \"\",\\n        \"description\": \"\"\n      }\\n    },\\n    {\\n      \"column\": \"rating\",\\n      \"properties\": {\\n        \"dtype\": \"number\",\\n        \"std\": 0.2851851688885017,\\n        \"min\": 3.9428571428571426,\\n        \"max\": 5.0,\\n        \"num_unique_values\": 190,\\n        \"samples\": [\\n          3.9857142857142853,\\n          3.9750000000000005,\\n          4.279999999999999\\n        ],\\n        \"semantic_type\": \"\",\\n        \"description\": \"\"\n      }\\n    }\\n  ]\\n},\"type\":\"dataframe\"}

def find_outliers_iqr(bb):
    outliers_dict = {} # Store outliers separately for each column
    outliers_bb = pd.DataFrame() # DataFrame to store all outlier rows

    # Select only numeric columns
    numeric_cols = bb.select_dtypes(include=['int', 'float']).columns

    for col in numeric_cols:
        Q1 = bb[col].quantile(0.25)
        Q3 = bb[col].quantile(0.75)
        IQR = Q3 - Q1
        lower_limit = Q1 - 1.5 * IQR
        upper_limit = Q3 + 1.5 * IQR

        # Store column-wise outliers in dictionary

```

```

        outliers_dict[col] = bb[(bb[col] < lower_limit) | (bb[col] >
upper_limit)][col].tolist()

    # Append outlier rows to outliers_df
    outliers_bb = pd.concat([outliers_bb, bb[(bb[col] <
lower_limit) | (bb[col] > upper_limit)]])

    # Print IQR details
    print(f"Column: {col}")
    print(f"  Lower Limit: {lower_limit}")
    print(f"  Upper Limit: {upper_limit}")
    print(f"  Outliers Detected: {len(outliers_dict[col])}\n")

return outliers_bb

outliers = find_outliers_iqr(bb)
outliers

Column: index
Lower Limit: -13776.0
Upper Limit: 41332.0
Outliers Detected: 0

Column: sale_price
Lower Limit: -301.0
Upper Limit: 755.0
Outliers Detected: 2205

Column: market_price
Lower Limit: -387.5
Upper Limit: 912.5
Outliers Detected: 2147

Column: rating
Lower Limit: 2.8000000000000007
Upper Limit: 5.199999999999999
Outliers Detected: 1751

Column: Discount
Lower Limit: -76.5
Upper Limit: 127.5
Outliers Detected: 3360

Column: Discount%
Lower Limit: -30.0
Upper Limit: 50.0
Outliers Detected: 438

{"summary": "\n  \"name\": \"outliers\", \n  \"rows\": 9901,\n  \"fields\": [\n    {\n      \"column\": \"index\", \n

```

```

  "properties": {
    "id": {
      "dtype": "number",
      "min": 7940,
      "max": 27555,
      "num_unique_values": 5897,
      "samples": [
        25836, 25002, 22367
      ],
      "semantic_type": "\",
      "description": "\n      \"\n    }",
      "column": "product",
      "properties": {
        "category": {
          "dtype": "string",
          "min": "Sliced Jalapenos",
          "max": "Stainless Steel Dish Cover - Mesh, 25 cm",
          "num_unique_values": 5379,
          "samples": [
            "Tempting Tuna Flavour",
            "Cat Treat - Snacks & Branded Foods",
            "Beauty & Hygiene",
            "Bakery, Cakes & Dairy"
          ],
          "semantic_type": "\",
          "description": "\n      \"\n    }",
          "column": "category",
          "properties": {
            "brand": {
              "dtype": "category",
              "min": "Bins & Bathroom Ware",
              "max": "Hair Care",
              "num_unique_values": 11,
              "samples": [
                "Baby Bath & Hygiene",
                "Raju Candles",
                "Appitas"
              ],
              "semantic_type": "\",
              "description": "\n      \"\n    }",
              "column": "brand",
              "properties": {
                "sale_price": {
                  "dtype": "number",
                  "min": 1021,
                  "max": 112475.0,
                  "num_unique_values": 83,
                  "samples": [
                    "Ramson",
                    "Raju Candles",
                    "Appitas"
                  ],
                  "semantic_type": "\",
                  "description": "\n      \"\n    }",
                  "column": "sale_price",
                  "properties": {
                    "market_price": {
                      "dtype": "number",
                      "min": 3657.7349005305614,
                      "max": 112475.0,
                      "num_unique_values": 1688,
                      "samples": [
                        47.0, 2625.0, 814.0,
                        1249.0, 1134.0, 4000.0
                      ],
                      "semantic_type": "\",
                      "description": "\n      \"\n    }",
                      "column": "market_price",
                      "properties": {
                        "type": {
                          "dtype": "category",
                          "min": 1168.060311740716,
                          "max": 12500.0,
                          "num_unique_values": 978,
                          "samples": [
                            "Organic Millet & Flours",
                            "Flavoured & Other Oils",
                            "Luxury Chocolates, Gifts"
                          ],
                          "semantic_type": "\",
                          "description": "\n      \"\n    }",
                          "column": "type",
                          "properties": {
                            "rating": {
                              "dtype": "number",
                              "min": 1.0497938825599078,
                              "max": 5.0,
                              "num_unique_values": 354,
                              "samples": [
                                2.3666666666666667,
                                3.0999999999999996
                              ],
                              "semantic_type": "\",
                              "description": "\n      \"\n    }",
                              "column": "rating",
                              "properties": {}
                            }
                          }
                        }
                      }
                    }
                  }
                }
              }
            }
          }
        }
      }
    }
  }
}

```

```

1.833333333333333\n      ],\n      \\"semantic_type\\": \"\",\\n
\\\"description\\\": \"\\n      }\\n    },\\n    {\\n      \\\"column\\\":\n\\\"description\\\",\\n      \\\"properties\\\": {\\n        \\\"dtype\\\":\n\\\"category\\\",\\n        \\\"num_unique_values\\\": 4696,\\n\n\\\"samples\\\": [\n      \\\"Soul flower bath & body products provide\nyou with your very own spa at home. The products range in our gift set\nare formulated by combining precious essential oils with the blend of\nnatural and organic ingredients.\\u00a0The idea behind the whole\ncreation is to provide you and your dear ones the natural and vegan\nskin care products which are healthy, regenerating and maintain body\nand skin balance.\",\\n      \\\"Easily pack lunch for your loved\nones with this green lunch box by Tedemei. Made of high-quality\nstainless steel, the lunch box is sturdy, durable, and easy to clean.\nThe lunch box is airtight which helps in keeping the food fresh for\nlong. It is a single layer lunch box which lets you pack solid and\nliquid food separately. The lunch box features flap and lock design\nwhich makes it easy to open and carry the lunch box. The modish\nlooking lunch boxes catches the eye.\",\\n      \\\"This man is tough\nand mysterious, he has the power of seduction at his fingertips. On\nthis man, Whisky Vintage is dangerously hot. Try this long-lasting\nfragrance to stay fresh and noticeable all day.\\"\\n    ],\\n\n\\\"semantic_type\\\": \"\",\\n      \\\"description\\\": \"\\n      }\n\n      {\\n        \\\"column\\\": \\\"Discount\\\",\\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"number\\\",\\n          \\\"std\\\":\n3536.667435148134,\\n          \\\"min\\\": -110360.0,\\n          \\\"max\\\":\n4320.0,\\n          \\\"num_unique_values\\\": 1286,\\n          \\\"samples\\\": [\n            158.06,\\n            911.0,\\n            278.0\\n          ],\\n\n          \\\"semantic_type\\\": \"\",\\n          \\\"description\\\": \"\\n          }\n\n          {\\n            \\\"column\\\": \\\"Discount%\\\",\\n            \\\"properties\\\": {\n              \\\"dtype\\\": \\\"number\\\",\\n              \\\"std\\\":\n1218.1665113057945,\\n              \\\"min\\\": -68813.95,\\n              \\\"max\\\":\n83.67,\\n              \\\"num_unique_values\\\": 1492,\\n              \\\"samples\\\": [\n                30.8,\\n                27.05,\\n                58.05\\n              ],\\n\n              \\\"semantic_type\\\": \"\",\\n              \\\"description\\\": \"\\n              }\n\n            }\n          ]\\n        },\\n        \\\"type\\\": \"dataframe\",\\n        \\\"variable_name\\\": \"outliers\"\n      }\n\noutliers.loc[outliers['rating'] > np.average(outliers['rating']),\n['product', 'sale_price', 'rating']].sort_values(by='rating',\nascending=False)\n\n{\\n  \\\"summary\\\": {\\n    \\\"name\\\": \\\"outliers\\\",\\n    \\\"rows\\\": 6031,\\n\n    \\\"fields\\\": [\n      {\\n        \\\"column\\\": \\\"product\\\",\\n\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"string\\\",\\n\n          \\\"num_unique_values\\\": 3116,\\n          \\\"samples\\\": [\n            \\\"Steel Dinner Plate/Thali - No. 12, Beggi China\\\",\\n\n            \\\"Basmati Rice/Basmati Akki - Biryani Special\\\",\\n            \\\"Face\nCare Kit - Face Cream, Face Wash & Face Toner\\n          ],\\n\n          \\\"semantic_type\\\": \"\",\\n          \\\"description\\\": \"\\n          }\n\n        }\n      }\n\n      {\\n        \\\"column\\\": \\\"sale_price\\\",\\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"number\\\",\\n          \\\"std\\\":\n\n          \\\"min\\\": 10.0,\\n          \\\"max\\\": 1000.0,\\n          \\\"mean\\\": 100.0,\\n          \\\"std\\\": 100.0,\\n          \\\"count\\\": 6031,\\n          \\\"sum\\\": 603100.0,\\n          \\\"variance\\\": 1000000.0,\\n          \\\"median\\\": 100.0,\\n          \\\"q1\\\": 50.0,\\n          \\\"q3\\\": 200.0,\\n          \\\"iqr\\\": 150.0,\\n          \\\"outlier\\\": false\n        }\n      }\n\n      {\\n        \\\"column\\\": \\\"rating\\\",\\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"number\\\",\\n          \\\"std\\\":\n\n          \\\"min\\\": 0.0,\\n          \\\"max\\\": 5.0,\\n          \\\"mean\\\": 3.0,\\n          \\\"std\\\": 1.0,\\n          \\\"count\\\": 6031,\\n          \\\"sum\\\": 18093.0,\\n          \\\"variance\\\": 10000.0,\\n          \\\"median\\\": 3.0,\\n          \\\"q1\\\": 1.0,\\n          \\\"q3\\\": 5.0,\\n          \\\"iqr\\\": 4.0,\\n          \\\"outlier\\\": false\n        }\n      }\n\n      {\\n        \\\"column\\\": \\\"category\\\",\\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"category\\\",\\n          \\\"std\\\":\n\n          \\\"min\\\": 0.0,\\n          \\\"max\\\": 1.0,\\n          \\\"mean\\\": 0.5,\\n          \\\"std\\\": 0.25,\\n          \\\"count\\\": 6031,\\n          \\\"sum\\\": 3015.5,\\n          \\\"variance\\\": 0.25,\\n          \\\"median\\\": 0.5,\\n          \\\"q1\\\": 0.0,\\n          \\\"q3\\\": 1.0,\\n          \\\"iqr\\\": 1.0,\\n          \\\"outlier\\\": false\n        }\n      }\n\n      {\\n        \\\"column\\\": \\\"discount\\\",\\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"number\\\",\\n          \\\"std\\\":\n\n          \\\"min\\\": -110360.0,\\n          \\\"max\\\": 4320.0,\\n          \\\"mean\\\": -108020.0,\\n          \\\"std\\\": 3536.667435148134,\\n          \\\"count\\\": 6031,\\n          \\\"sum\\\": -648000.0,\\n          \\\"variance\\\": 1590000000.0,\\n          \\\"median\\\": -108020.0,\\n          \\\"q1\\\": -110360.0,\\n          \\\"q3\\\": 4320.0,\\n          \\\"iqr\\\": 114560.0,\\n          \\\"outlier\\\": false\n        }\n      }\n\n      {\\n        \\\"column\\\": \\\"discount%\\\",\\n        \\\"properties\\\": {\n          \\\"dtype\\\": \\\"number\\\",\\n          \\\"std\\\":\n\n          \\\"min\\\": -68813.95,\\n          \\\"max\\\": 83.67,\\n          \\\"mean\\\": -10798.0,\\n          \\\"std\\\": 1218.1665113057945,\\n          \\\"count\\\": 6031,\\n          \\\"sum\\\": -648000.0,\\n          \\\"variance\\\": 1590000000.0,\\n          \\\"median\\\": -10798.0,\\n          \\\"q1\\\": -68813.95,\\n          \\\"q3\\\": 83.67,\\n          \\\"iqr\\\": 75.82,\\n          \\\"outlier\\\": false
        }
      }
    ]
  }
}

```

```

3640.9922424252118,\n          \"min\": 2.45,\n          \"max\":\n111649.0,\n          \"num_unique_values\": 1258,\n          \"samples\":\n[\"1946.25,\n          1299.0,\n          850.0\"],\n],\n      },\n      {\n        \"column\": \"rating\",\n        \"description\": \"\n\n\",\"properties\":\n{\n          \"dtype\": \"number\",\n          \"std\":\n0.38282320507800455,\n          \"min\": 3.52,\n          \"max\": 5.0,\n          \"num_unique_values\": 197,\n          \"samples\": [\n            3.925,\n            4.0500000000000001,\n            4.76\n          ],\n        },\n        \"semantic_type\": \"\",\"description\": \"\n\n\"}\n    }\n  ]\n},\"type\":\"dataframe\"}

```

bb.head(1)

```

{"summary":{\n  \"name\": \"bb\", \"rows\": 27555, \"fields\":\n[{\n    \"column\": \"index\", \"properties\": {\n      \"dtype\": \"number\", \"std\": 7954, \"min\": 1,\n      \"max\": 27555, \"num_unique_values\": 27555,\n      \"samples\": [\n        16202,\n        24963,\n        18965\n      ],\n      \"semantic_type\": \"\", \"description\": \"\n\n\"}\n    },\n    {\n      \"column\": \"product\", \"properties\": {\n        \"dtype\": \"string\", \"num_unique_values\": 23540,\n        \"samples\": [\n          \"Pure Impact Body Wash With Purifying Micro Particles - Shower Gel\nFor Body, Face & Hair\", \"Premium Care XXL - 30 Diaper\nPants\", \"Tomato - Green\", \"\n        \",\n        \"\n      ],\n      \"semantic_type\": \"\", \"description\": \"\n\n\"}\n    },\n    {\n      \"column\": \"category\", \"properties\": {\n        \"dtype\": \"category\", \"num_unique_values\": 11,\n        \"samples\": [\n          \"Snacks & Branded Foods\", \"\n          \"Beauty & Hygiene\", \"\n          \"Fruits & Vegetables\"\n        ],\n        \"semantic_type\": \"\", \"description\": \"\n\n\"}\n    },\n    {\n      \"column\": \"sub_category\", \"properties\": {\n        \"dtype\": \"category\", \"num_unique_values\": 90,\n        \"samples\": [\n          \"\n          \"Crockery & Cutlery\", \"\n          \"Fragrances & Deos\", \"\n          \"Energy & Soft Drinks\"\n        ],\n        \"semantic_type\": \"\", \"description\": \"\n\n\"}\n    },\n    {\n      \"column\": \"brand\", \"properties\": {\n        \"dtype\": \"category\", \"num_unique_values\": 2313,\n        \"samples\": [\n          \"Red Label\", \"\n          \"Lomani\", \"\n          \"Happily Unmarried\", \"\n          \"\n        ],\n        \"semantic_type\": \"\", \"description\": \"\n\n\"}\n    },\n    {\n      \"column\": \"sale_price\", \"properties\": {\n        \"dtype\": \"number\", \"std\": 1201.9769904246652,\n        \"min\": 2.45,\n        \"max\": 112475.0,\n        \"num_unique_values\": 3261,\n        \"samples\": [\n          24.95,\n          40.0,\n          5040.0\n        ],\n        \"semantic_type\": \"\", \"description\": \"\n\n\"}\n    },\n    {\n      \"column\": \"market_price\", \"properties\": {\n        \"dtype\": \"\n\n\"}\n    }\n  ]\n},\"type\":\"dataframe\"}

```

```
"number",\n          "std": 581.7307174034938,\n          "min":\n            3.0,\n          "max": 12500.0,\n          "num_unique_values": 1348,\n          "samples": [\n            192.0,\n            1164.0,\n            74.38\n          ],\n          "semantic_type": "\",\n          "description": \"\\n      }\\n    },\\n    {\n      \"column\": \"type\",\n      \"properties\": {\n        \"dtype\": \"category\",\n        \"num_unique_values\": 426,\n        \"samples\": [\n          \"Nappies & Rash Cream\",\n          \"Hair Removal\",\n          \"Crackers & Digestive\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\\n      }\\n    },\\n    {\n      \"column\": \"rating\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.6841238205552671,\n        \"min\":\n          1.0,\n        \"max\": 5.0,\n        \"num_unique_values\": 486,\n        \"samples\": [\n          3.9428571428571426,\n          2.2666666666666666,\n          4.485714285714286\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\\n      }\\n    },\n    {\n      \"column\": \"description\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 21944,\n        \"samples\": [\n          \"Premium Day Face Cream:Lightweight formula that will give your skin a daily boost of radiance and hydration. Packed with the goodness of skin brightening, Vitamin C Why Is It Special?- It contains the Most Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is the next generation Vitamin C that really works in improving skin health.- Vitamin C is a proven skin-brightening agent known for its antioxidant properties for decades. It boosts collagen formation and effectively targets visible signs of ageing, giving skin the much-needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3 helps awaken and clarify your complexion, revealing a natural healthy-looking glow. - SPF 30 & PA+++ provide the highest level of protection against UVA & UVB rays of the sun. The mineral-based sun protection agents shield the skin against the harsh solar radiation, minimizing skin dryness and dullness. - Multiple natural oils are combined for intense hydration and nourishment. Antioxidant-rich botanical extracts are added to keep skin moisturized, smooth & soft. - St.Botanica is a PETA-certified cruelty-free brand. The product does not contain any harmful chemicals & is free from parabens, silicones or mineral oils.\",\n          \"Kari Kari Japanese premium snacks are the exciting irresistible new taste of crunch. Lite and crispy, the savoury rice crackers are mixed with peanuts to make hearty, wholesome snack with delicious, bold and zesty taste with every bite.\",\n          \"Safely store your food in this pink lunch box. The lunch box is made plastic which gives the container durability and sturdiness. The rectangular containers come with an airtight lid to keep food fresh for long. The wide mouth opening of the lunch makes it easy to fill food and easy to clean. The set makes a great gift option for special days.\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\\n      }\\n    },\n    {\n      \"column\": \"Discount\",\n      \"properties\": {\n        \"dtype\": \"
```

```

\"number\", \n          \"std\": 1099.1979133300101, \n          \"min\": -\n110360.0, \n          \"max\": 4320.0, \n          \"num_unique_values\":\n2517, \n          \"samples\": [\n            950.0, \n            1464.99, \n30.12000000000005\n          ], \n          \"semantic_type\": \"\", \n          \"column\":\n          \"Discount%\", \n          \"properties\": {\n            \"dtype\":\n            \"number\", \n            \"std\": 420.56482233441704, \n            \"min\": -\n68813.95, \n            \"max\": 83.67, \n            \"num_unique_values\":\n2666, \n            \"samples\": [\n              10.66, \n              41.67, \n8.84\n            ], \n            \"semantic_type\": \"\", \n          }\n        }\n      },\n      \"type\": \"dataframe\", \"variable_name\": \"bb\"}\n\ndef replace_outliers_iqr(bb):\n    outliers_dict = {}\n    # Select only numeric columns\n    numeric_cols = bb.select_dtypes(include=['int', 'float']).columns\n\n    for col in numeric_cols:\n        Q1 = bb[col].quantile(0.25)\n        Q3 = bb[col].quantile(0.75)\n        IQR = Q3 - Q1\n        lower_limit = Q1 - 1.5 * IQR\n        upper_limit = Q3 + 1.5 * IQR\n        mean_value = bb[col].mean()\n\n        # Create a mask for outliers\n        outliers = (bb[col] < lower_limit) | (bb[col] > upper_limit)\n        outliers_dict[col] = bb.loc[outliers, col].tolist()\n\n        # Replace outliers with mean\n        bb.loc[outliers, col] = mean_value\n\n    return bb\n\nreplace_outliers_iqr(bb)\n\n{ \"summary\": { \n    \"name\": \"bb\", \n    \"rows\": 27555, \n    \"fields\": [\n      { \n        \"column\": \"index\", \n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 7954, \n          \"min\": 1, \n          \"max\": 27555, \n          \"num_unique_values\": 27555, \n          \"samples\": [\n            16202, \n            24963, \n            18965\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      }, \n      { \n        \"column\": \"product\", \n        \"properties\": {\n          \"dtype\": \"string\", \n          \"num_unique_values\": 23540, \n          \"samples\": [\n            \"Pure Impact Body Wash With Purifying Micro Particles - Shower Gel\nFor Body, Face & Hair\", \n            \"Premium Care XXL - 30 Diaper\nPants\", \n            \"Tomato - Green\"\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      }\n    ]\n  }\n}
```

```
n      },\n      {\n        \"column\": \"category\",\\n        \"properties\":\n          {\n            \"dtype\": \"category\",\\n            \"num_unique_values\":\n              11,\n            \"samples\": [\n              {\\n                \"description\": \"Snacks & Branded Foods\",\\n                \"semantic_type\": \"\",\\n                \"samples\": [\n                  {\\n                    \"description\": \"Fruits & Vegetables\\n                  ]},\\n                  {\\n                    \"column\": \"sub_category\",\\n                    \"properties\": {\n                      \"dtype\": \"category\",\\n                      \"num_unique_values\": 90,\n                      \"samples\": [\n                        {\\n                          \"description\": \"Crockery & Cutlery\",\\n                          \"semantic_type\": \"\",\\n                          \"samples\": [\n                            {\\n                              \"description\": \"Fragrances & Deos\",\\n                              \"semantic_type\": \"\",\\n                              \"samples\": [\n                                {\\n                                  \"brand\": \"\",\\n                                  \"properties\": {\n                                    \"category\": \"\",\\n                                    \"num_unique_values\": 2313,\n                                    \"samples\": [\n                                      {\\n                                        \"Red Label\",\\n                                        \"Lomani\",\\n                                        \"Happily Unmarried\",\\n                                        \"semantic_type\": \"\",\\n                                        \"description\": \"\",\\n                                        \"column\": \"sale_price\",\\n                                        \"properties\": {\n                                          \"dtype\": \"number\",\\n                                          \"std\": 151.88289560995017,\n                                          \"min\": 2.45,\n                                          \"max\": 693.0,\n                                          \"num_unique_values\": 2360,\n                                          \"samples\": [\n                                            {\\n                                              \"35.2\",\\n                                              \"116.25\",\\n                                              \"383.04\",\\n                                              \"semantic_type\": \"\",\\n                                              \"description\": \"\",\\n                                              \"column\": \"market_price\",\\n                                              \"properties\": {\n                                                \"dtype\": \"number\",\\n                                                \"std\": 181.78174014888793,\n                                                \"min\": 3.0,\n                                                \"max\": 805.0,\n                                                \"num_unique_values\": 786,\n                                                \"samples\": [\n                                                  {\\n                                                    \"60.0\",\\n                                                    \"774.0\",\\n                                                    \"504.0\",\\n                                                    \"semantic_type\": \"\",\\n                                                    \"description\": \"\",\\n                                                    \"column\": \"type\",\\n                                                    \"properties\": {\n                                                      \"dtype\": \"category\",\\n                                                      \"num_unique_values\": 426,\n                                                      \"samples\": [\n                                                        {\\n                                                          \"Nappies & Rash Cream\",\\n                                                          \"Hair Removal\",\\n                                                          \"Crackers & Digestive\",\\n                                                          \"semantic_type\": \"\",\\n                                                          \"description\": \"\",\\n                                                          \"column\": \"rating\",\\n                                                          \"properties\": {\n                                                            \"dtype\": \"number\",\\n                                                            \"std\": 0.17748463134947098,\n                                                            \"min\": 3.6999999999999997,\n                                                            \"max\": 4.5,\n                                                            \"num_unique_values\": 190,\n                                                            \"samples\": [\n                                                              {\\n                                                                \"4.3428571428571425\",\\n                                                                \"3.757142857142857\",\\n                                                                \"3.8666666666666663\",\\n                                                                \"semantic_type\": \"\",\\n                                                                \"description\": \"\",\\n                                                                \"column\": \"description\",\\n                                                                \"properties\": {\n                                                                  \"dtype\": \"string\",\\n                                                                  \"num_unique_values\": 21944,\n                                                                  \"samples\": [\n                        {\\n                          \"Premium Day Face Cream:Lightweight formula that will give your skin\n                          a daily boost of radiance and hydration. Packed with the goodness of\n                          skin brightening, Vitamin C Why Is It Special?- It contains the Most\n                          Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is\n                          the next generation Vitamin C that really works in improving skin\n                          health.- Vitamin C is a proven skin-brightening agent known for its\n                          antioxidant properties for decades. It boosts collagen formation and\n                        ]},\\n                      ]},\\n                    ]},\\n                  ]},\\n                ]},\\n              ]},\\n            ]},\\n          ]},\\n        ]},\\n      ]},\\n    ]},\\n  ]},\\n}
```

effectively targets visible signs of ageing, giving skin the much-needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3 helps awaken and clarify your complexion, revealing a natural healthy-looking glow. - SPF 30 & PA+++ provide the highest level of protection against UVA & UVB rays of the sun. The mineral-based sun protection agents shield the skin against the harsh solar radiation, minimizing skin dryness and dullness. - Multiple natural oils are combined for intense hydration and nourishment. Antioxidant-rich botanical extracts are added to keep skin moisturized, smooth & soft. - St.Botanica is a PETA-certified cruelty-free brand. The product does not contain any harmful chemicals & is free from parabens, silicones or mineral oils.",\n \"Kari Kari Japanese premium snacks are the exciting irresistible new taste of crunch. Lite and crispy, the savoury rice crackers are mixed with peanuts to make hearty, wholesome snack with delicious, bold and zesty taste with every bite.\",\n \"Safely store your food in this pink lunch box. The lunch box is made plastic which gives the container durability and sturdiness. The rectangular containers come with an airtight lid to keep food fresh for long. The wide mouth opening of the lunch makes it easy to fill food and easy to clean. The set makes a great gift option for special days.\",\n ],\n \"semantic\_type\": \"\",\n \"description\": \"\"\\n }\\n },\\n {\\n \"column\": \"Discount\",\\n \"properties\": {\\n \"dtype\": \"number\",\\n \"std\": 27.719445762316436,\\n \"min\": 0.0,\\n \"max\": 118.5,\\n \"num\_unique\_values\": 1508,\\n \"samples\": [\\n 65.45000000000005,\\n 4.349999999999994,\\n 9.280000000000001\\n ],\\n \"semantic\_type\": \"\",\\n \"description\": \"\"\\n }\\n },\\n {\\n \"column\": \"Discount%\",\\n \"properties\": {\\n \"dtype\": \"number\",\\n \"std\": 13.379322214696623,\\n \"min\": 0.0,\\n \"max\": 50.0,\\n \"num\_unique\_values\": 2406,\\n \"samples\": [\\n 22.74,\\n 30.06,\\n 21.51\\n ],\\n \"semantic\_type\": \"\",\\n \"description\": \"\"\\n }\\n }\n ],\n \"type\": \"dataframe\", \"variable\_name\": \"bb\"}\n\ndef replace\_outliers\_iqr(bb):\n numeric\_cols = bb.select\_dtypes(include=['int', 'float']).columns\n for col in numeric\_cols:\n Q1 = bb[col].quantile(0.25)\n Q3 = bb[col].quantile(0.75)\n IQR = Q3 - Q1\n lower\_limit = Q1 - 1.5 \* IQR\n upper\_limit = Q3 + 1.5 \* IQR\n mean\_value = bb[col].mean()\n\n for i in range(len(bb[col])):\n if bb[col].iloc[i] < lower\_limit or bb[col].iloc[i] >\nupper\_limit:\n bb.at[i, col] = mean\_value

```

    return bb

replace_outliers_iqr(bb)

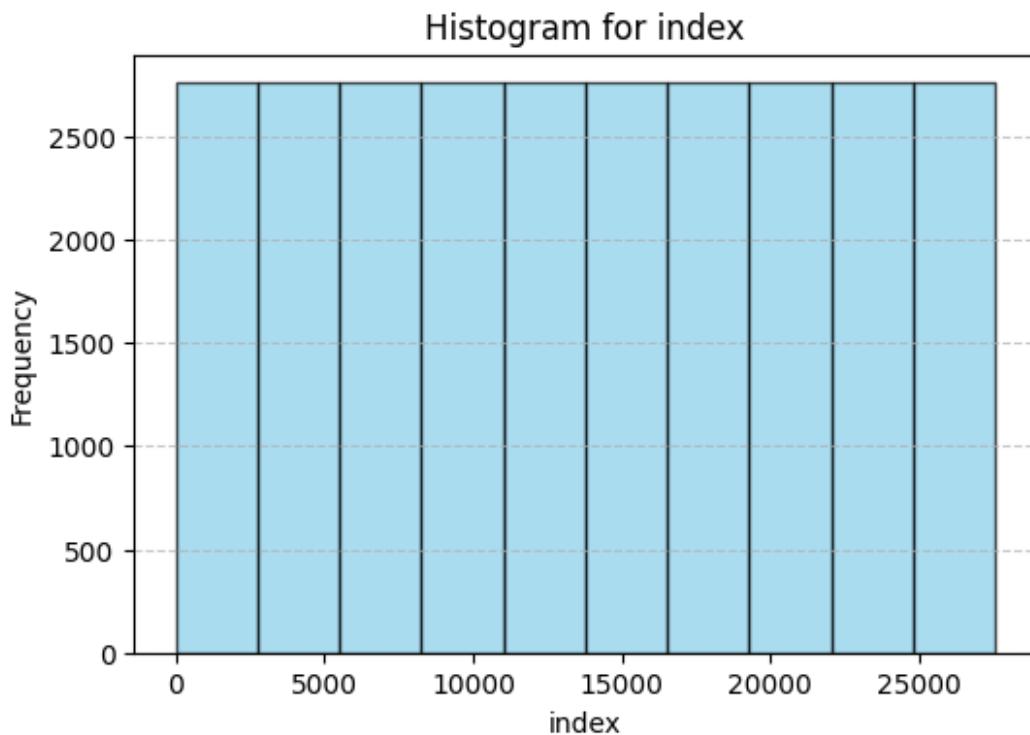
{"summary": {"\n    \"name\": \"bb\", \n    \"rows\": 27555, \n    \"fields\": [\n        {\n            \"column\": \"index\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 7954, \n                \"min\": 1, \n                \"max\": 27555, \n                \"num_unique_values\": 27555, \n                \"samples\": [\n                    16202, \n                    24963, \n                    18965\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"product\", \n            \"properties\": {\n                \"dtype\": \"string\", \n                \"num_unique_values\": 23540, \n                \"samples\": [\n                    \"Pure Impact Body Wash With Purifying Micro Particles - Shower Gel\nFor Body, Face & Hair\", \n                    \"Premium Care XXL - 30 Diaper\nPants\", \n                    \"Tomato - Green\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"category\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 11, \n                \"samples\": [\n                    \"Snacks & Branded Foods\", \n                    \"Beauty & Hygiene\", \n                    \"Fruits & Vegetables\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"sub_category\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 90, \n                \"samples\": [\n                    \"Crockery & Cutlery\", \n                    \"Fragrances & Deos\", \n                    \"Energy & Soft Drinks\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"brand\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 2313, \n                \"samples\": [\n                    \"Red Label\", \n                    \"Lomani\", \n                    \"Happily Unmarried\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"sale_price\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 162.32593459912707, \n                \"min\": 2.45, \n                \"max\": 755.0, \n                \"num_unique_values\": 2432, \n                \"samples\": [\n                    131.75, \n                    469.64, \n                    166.25\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"market_price\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 194.98763384063898, \n                \"min\": 3.0, \n                \"max\": 910.0, \n                \"num_unique_values\": 835, \n                \"samples\": [\n                    402.0, \n                    633.0, \n                    477.0\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }, \n            \"column\": \"type\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 426, \n                \"samples\": [\n                    \"Nappies & Rash Cream\", \n                    \"Hair Removal\", \n                    \"Crackers & Digestive\"\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        }\n    ]\n}
```



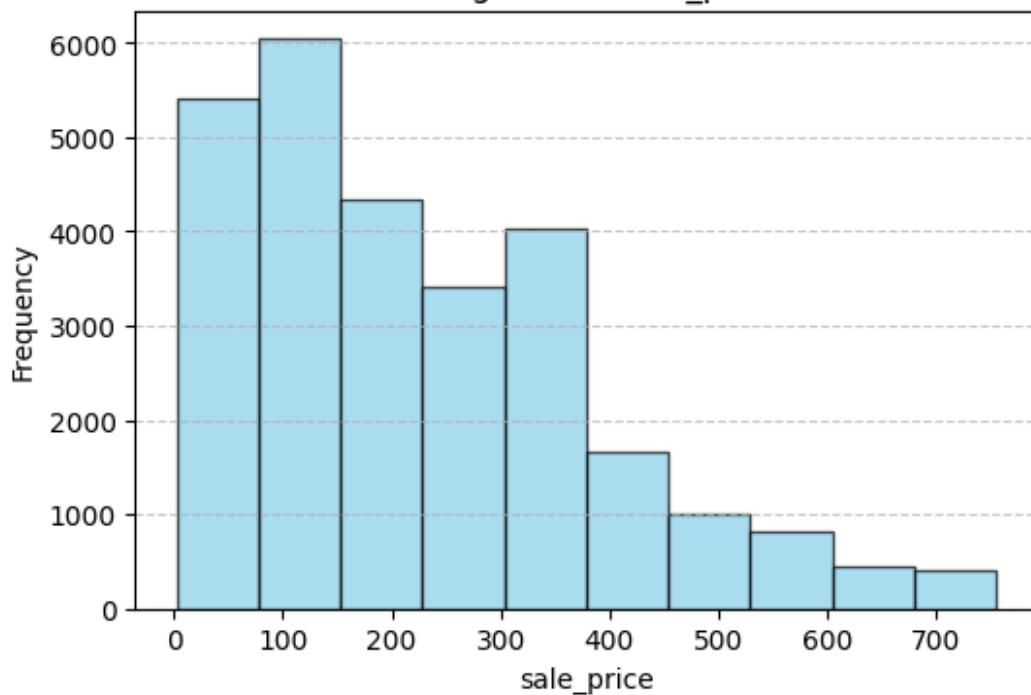
```
\"semantic_type\": \"\",  
\"description\": \"\"\n    }  
}\n  ]\n},  
\"type\":\"dataframe\",  
\"variable_name\":\"bb\"}
```

## Step 9: Create Plots or Visualisations

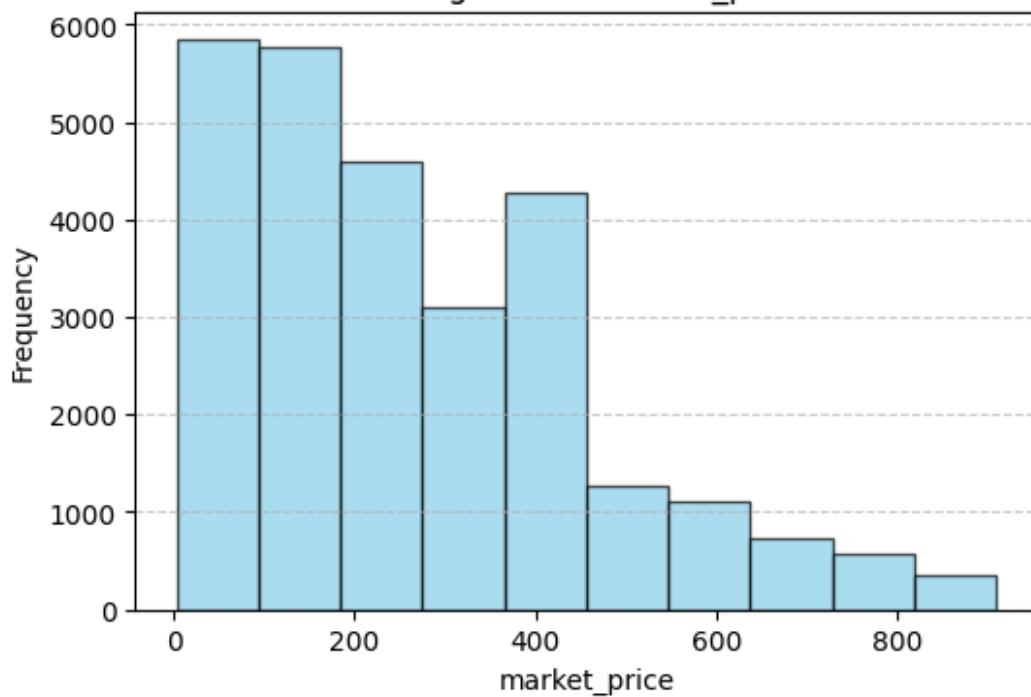
```
# Select numeric columns  
numeric_cols = bb.select_dtypes(include=['int', 'float']).columns  
  
# Plot histograms for each numeric column  
for col in numeric_cols:  
    plt.figure(figsize=(6, 4))  
    plt.hist(bb[col], bins=10, color='skyblue', edgecolor='black',  
alpha=0.7)  
    plt.title(f'Histogram for {col}')  
    plt.xlabel(col)  
    plt.ylabel('Frequency')  
    plt.grid(axis='y', linestyle='--', alpha=0.7)  
    plt.show()
```



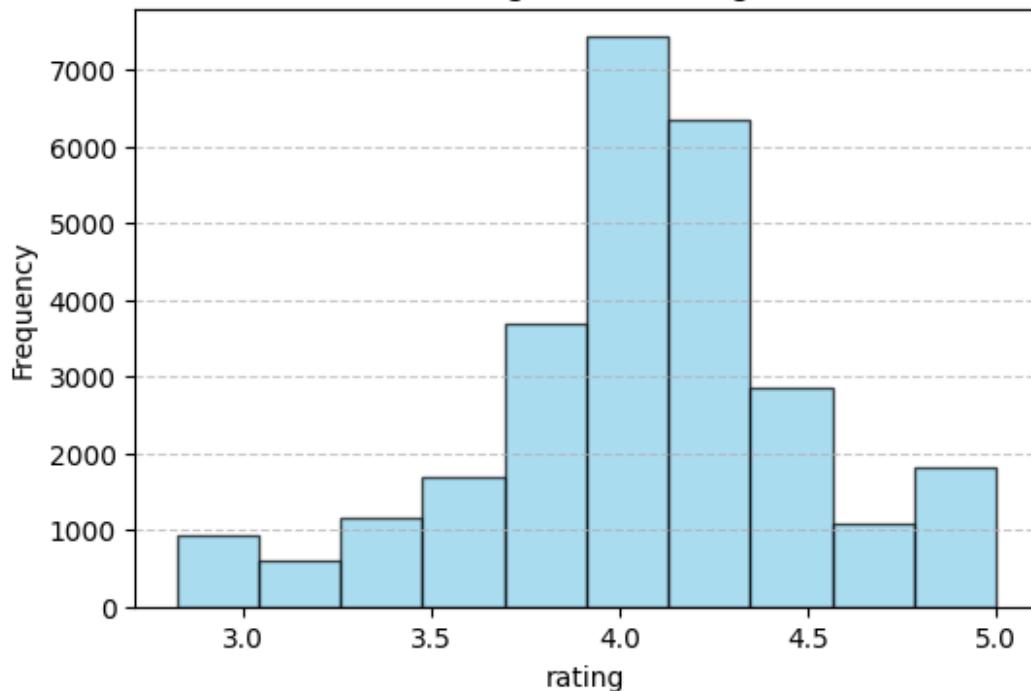
Histogram for sale\_price



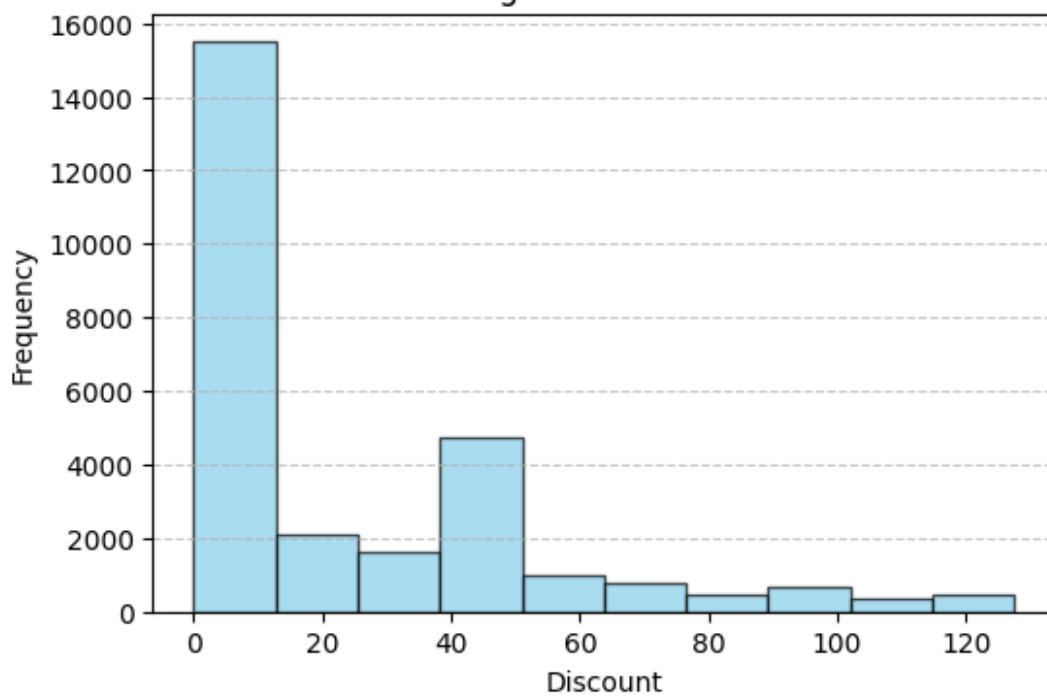
Histogram for market\_price

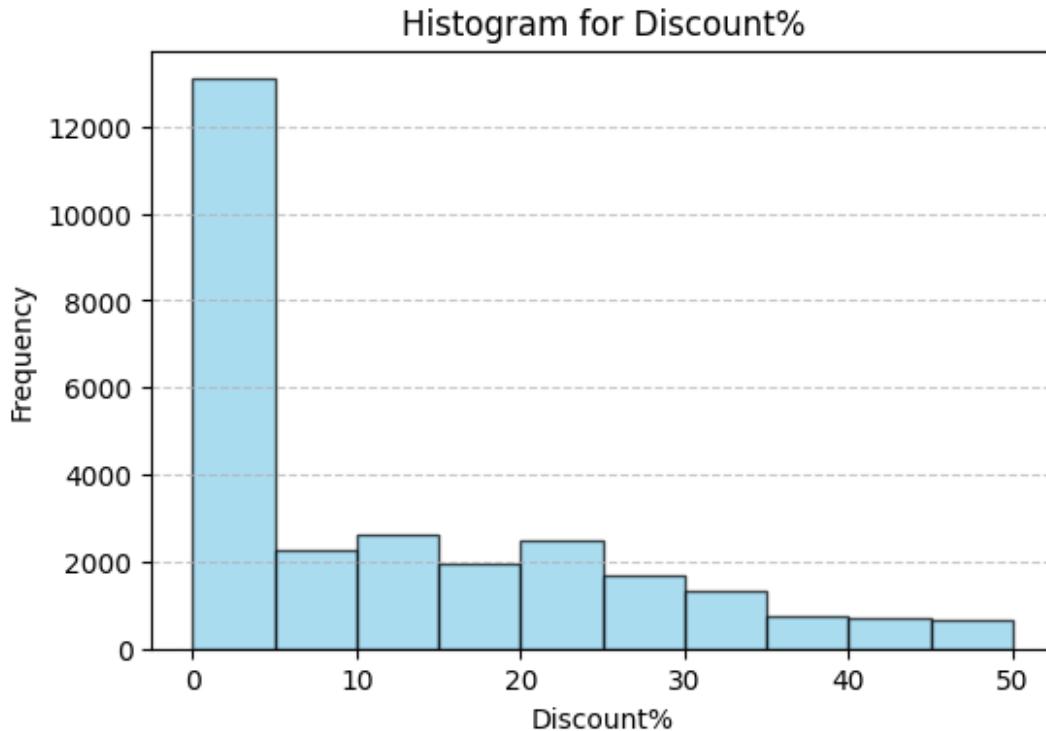


Histogram for rating



Histogram for Discount

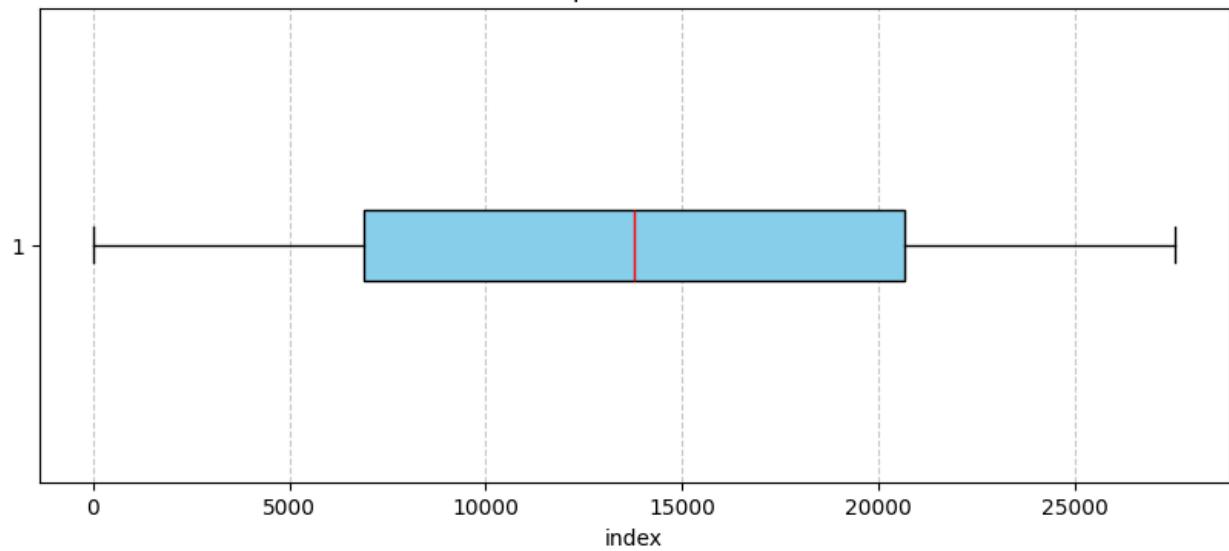




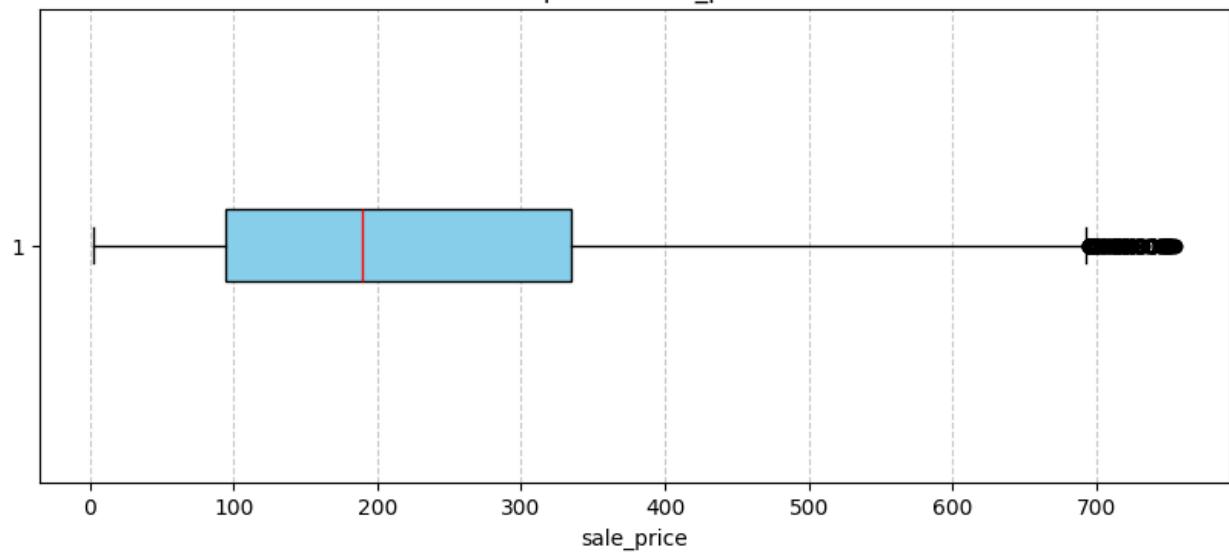
```
# Select numeric columns
numeric_cols = bb.select_dtypes(include=['int', 'float']).columns

# Plot boxplots for each numeric column
for col in numeric_cols:
    plt.figure(figsize=(10, 4))
    plt.boxplot(bb[col], vert=False, patch_artist=True,
                boxprops=dict(facecolor='skyblue', color='black'),
                whiskerprops=dict(color='black'),
                capprops=dict(color='black'),
                medianprops=dict(color='red'))
    plt.title(f'Boxplot for {col}')
    plt.xlabel(col)
    plt.grid(axis='x', linestyle='--', alpha=0.7)
    plt.show()
```

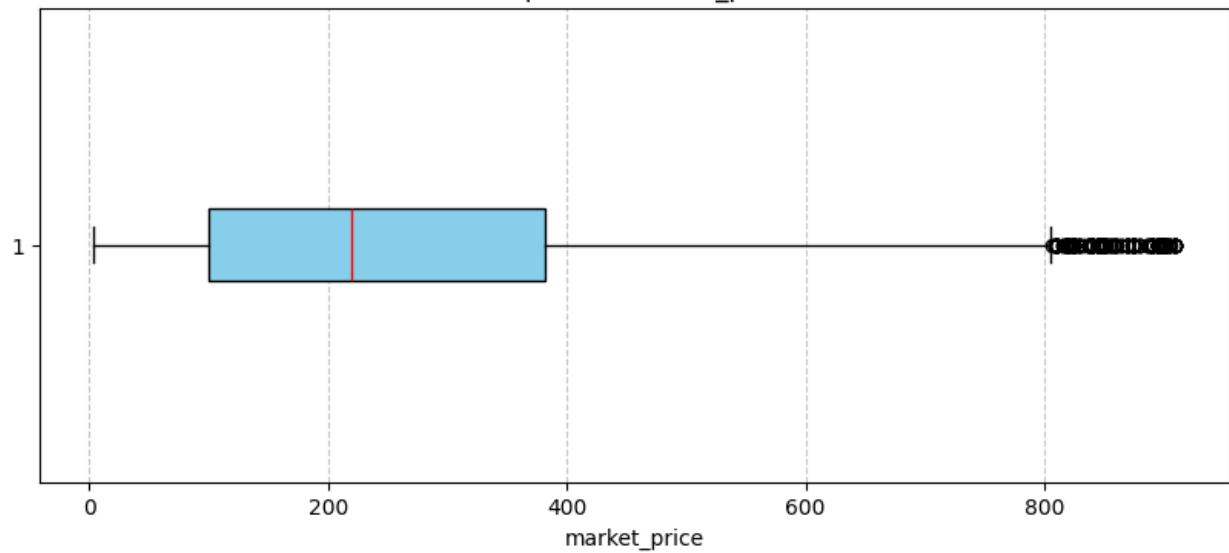
Boxplot for index



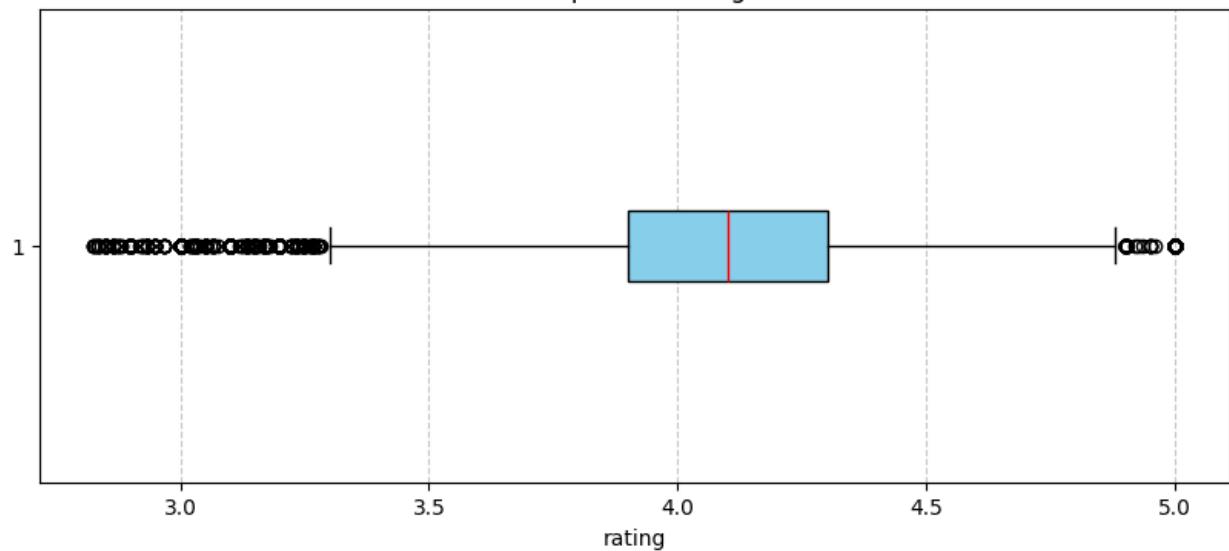
Boxplot for sale\_price



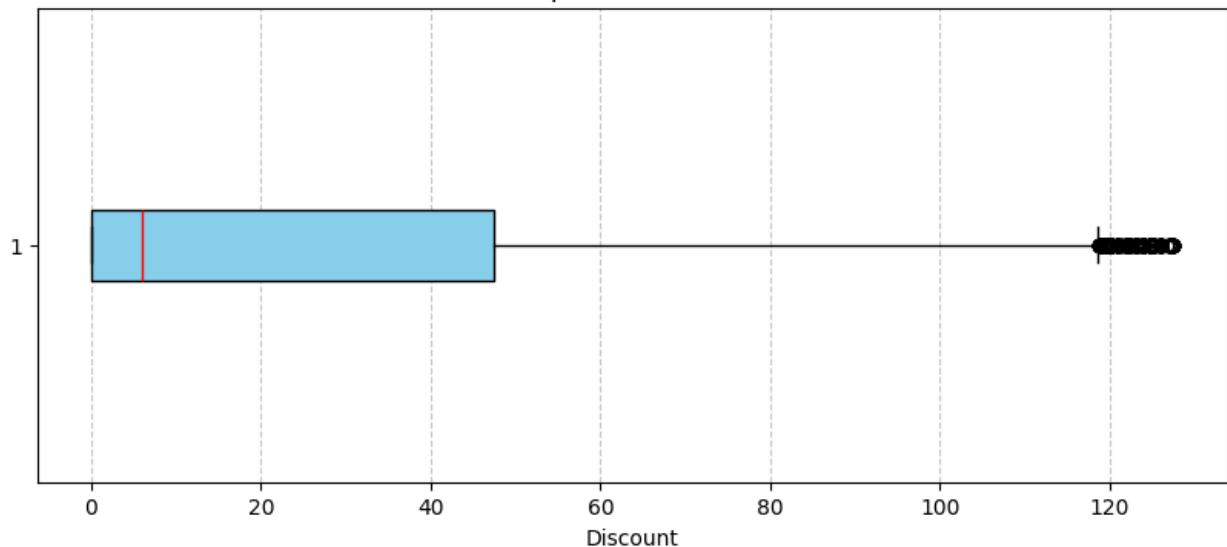
Boxplot for market\_price



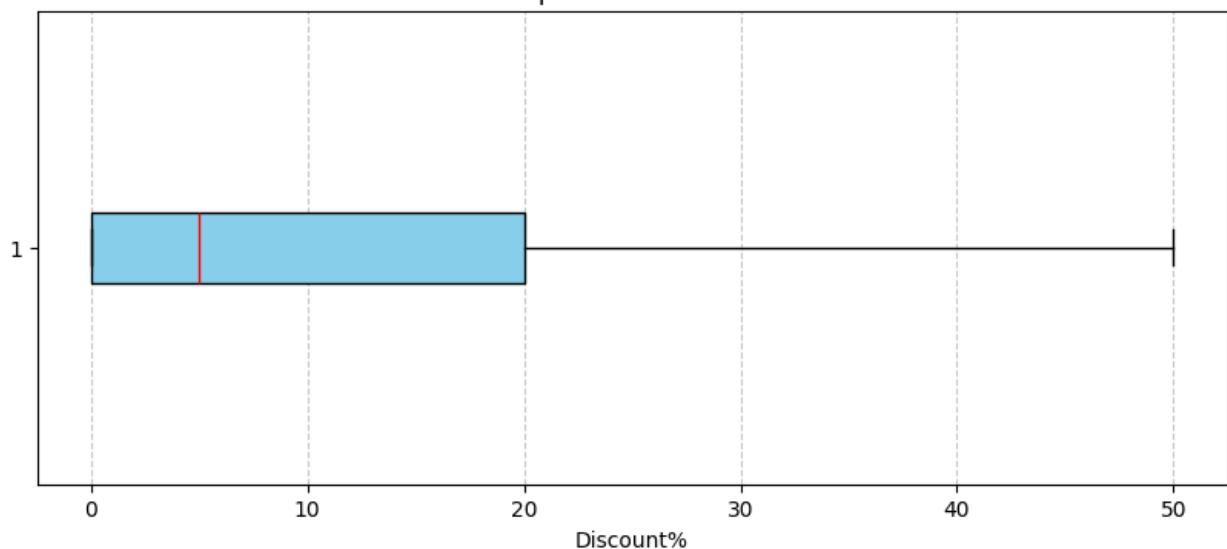
Boxplot for rating



Boxplot for Discount



Boxplot for Discount%



```
# Plot Bar Chart
plt.figure(figsize=(10, 6))
plt.bar(bb['product'].astype(str), bb['sale_price'], color='skyblue')
plt.tight_layout()
plt.xticks(rotation=45, ha='right')
plt.xlabel('product')
plt.ylabel('sale_price')
plt.title('Bar Chart of Sale Price by Product')
plt.show()
```

-----  
KeyboardInterrupt

Traceback (most recent call

```
last)
<ipython-input-52-d1f4cc0acc01> in <cell line: 0>()
    1 # Plot Bar Chart
    2 plt.figure(figsize=(10, 6))
----> 3 plt.bar(bb['product'].astype(str), bb['sale_price'],
color='skyblue')
    4 plt.tight_layout()
    5 plt.xticks(rotation=45, ha='right')

/usr/local/lib/python3.11/dist-packages/matplotlib/pyplot.py in bar(x,
height, width, bottom, align, data, **kwargs)
    2979     **kwargs,
    2980 ) -> BarContainer:
-> 2981     return gca().bar(
    2982         x,
    2983         height,

/usr/local/lib/python3.11/dist-packages/matplotlib/__init__.py in
inner(ax, data, *args, **kwargs)
    1519     def inner(ax, *args, data=None, **kwargs):
    1520         if data is None:
-> 1521             return func(
    1522                 ax,
    1523                 *map(cbook.sanitize_sequence, args),

/usr/local/lib/python3.11/dist-packages/matplotlib/axes/_axes.py in
bar(self, x, height, width, bottom, align, **kwargs)
    2647         else: # horizontal
    2648             r.sticky_edges.x.append(l)
-> 2649             self.add_patch(r)
    2650             patches.append(r)
    2651

/usr/local/lib/python3.11/dist-packages/matplotlib/axes/_base.py in
add_patch(self, p)
    2444         if p.get_clip_path() is None:
    2445             p.set_clip_path(self.patch)
-> 2446         self._update_patch_limits(p)
    2447         self._children.append(p)
    2448         p._remove_method = self._children.remove

/usr/local/lib/python3.11/dist-packages/matplotlib/axes/_base.py in
_update_patch_limits(self, patch)
    2466         # Loop through each segment to get extrema for Bezier
curve sections
    2467         vertices = []
-> 2468         for curve, code in p.iter_bezier(simplify=False):
    2469             # Get distance along the curve of any extrema
    2470             _, dzeros = curve.axis_aligned_extrema()
```

```
/usr/local/lib/python3.11/dist-packages/matplotlib/path.py in
iter_bezier(self, **kwargs)
    443                 yield BezierSegment(np.array([first_vert])),
code
    444             elif code == Path.LINETO: # "CURVE2"
--> 445                 yield BezierSegment(np.array([prev_vert,
verts]), code
    446             elif code == Path.CURVE3:
    447                 yield BezierSegment(np.array([prev_vert,
verts[:2],
```

KeyboardInterrupt:



```
# Count occurrences of each product
product_counts = bb['product'].value_counts().head(5)

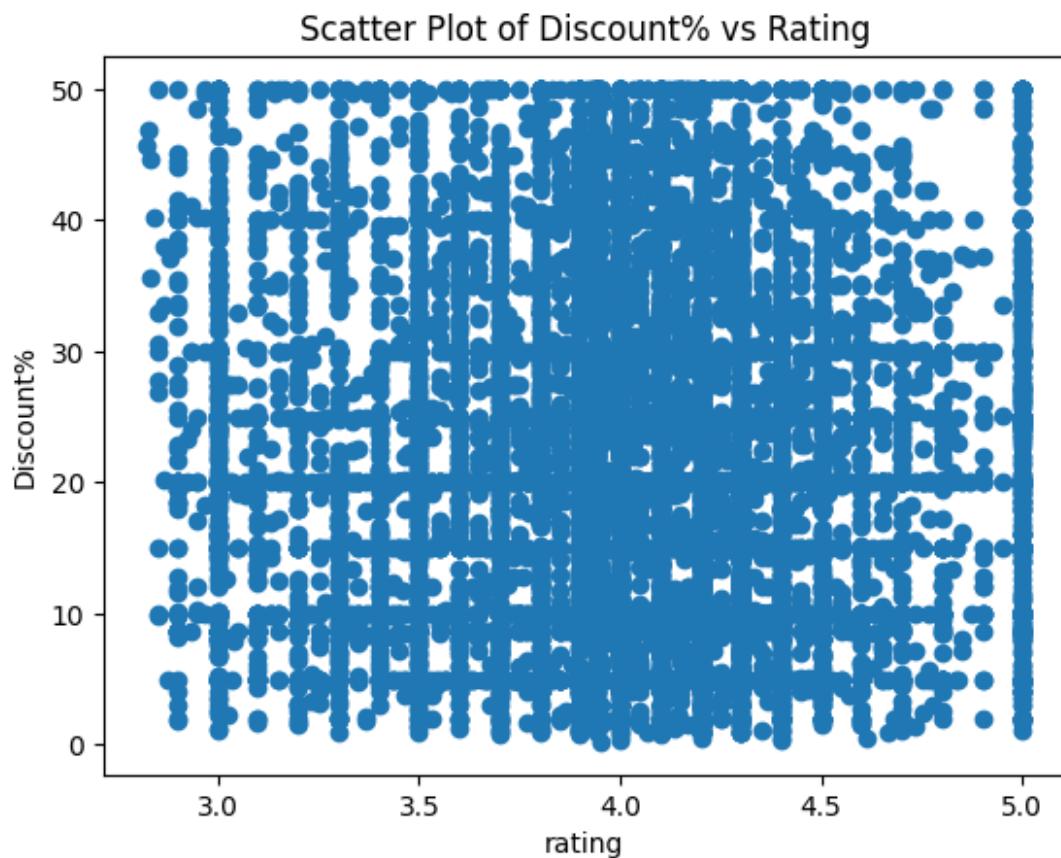
# Pie chart for 'Units Sold'
plt.figure(figsize=(8, 6))
plt.pie(bb['product_counts'], labels=product_counts.index,
autopct='%.1f%%',
        colors=['skyblue', 'orange', 'lightgreen', 'pink'],
startangle=90)
plt.title('Percentage of Units Sold by Product')
plt.show()
```

```

plt.scatter(bb['rating'], bb['sale_price'].where(bb['sale_price'] < 20000))
plt.xlabel('rating')
plt.ylabel('sale_price')
plt.title('Scatter Plot of Sale Price vs Rating')
plt.show()

plt.scatter(bb['rating'], bb['Discount%'].where(bb['Discount%'] > 0))
plt.xlabel('rating')
plt.ylabel('Discount%')
plt.title('Scatter Plot of Discount% vs Rating')
plt.show()

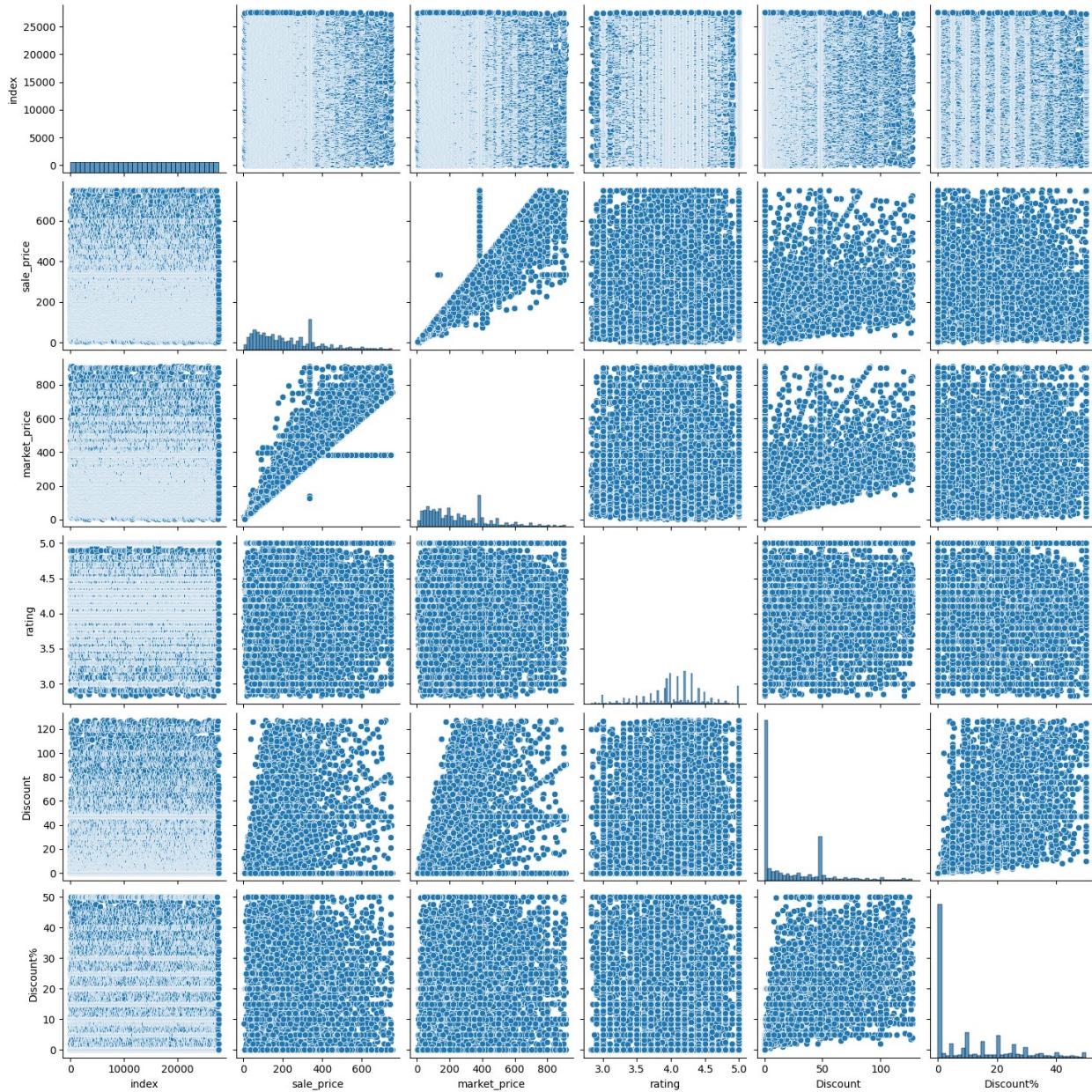
```



```

sns.pairplot(bb)
plt.show()

```

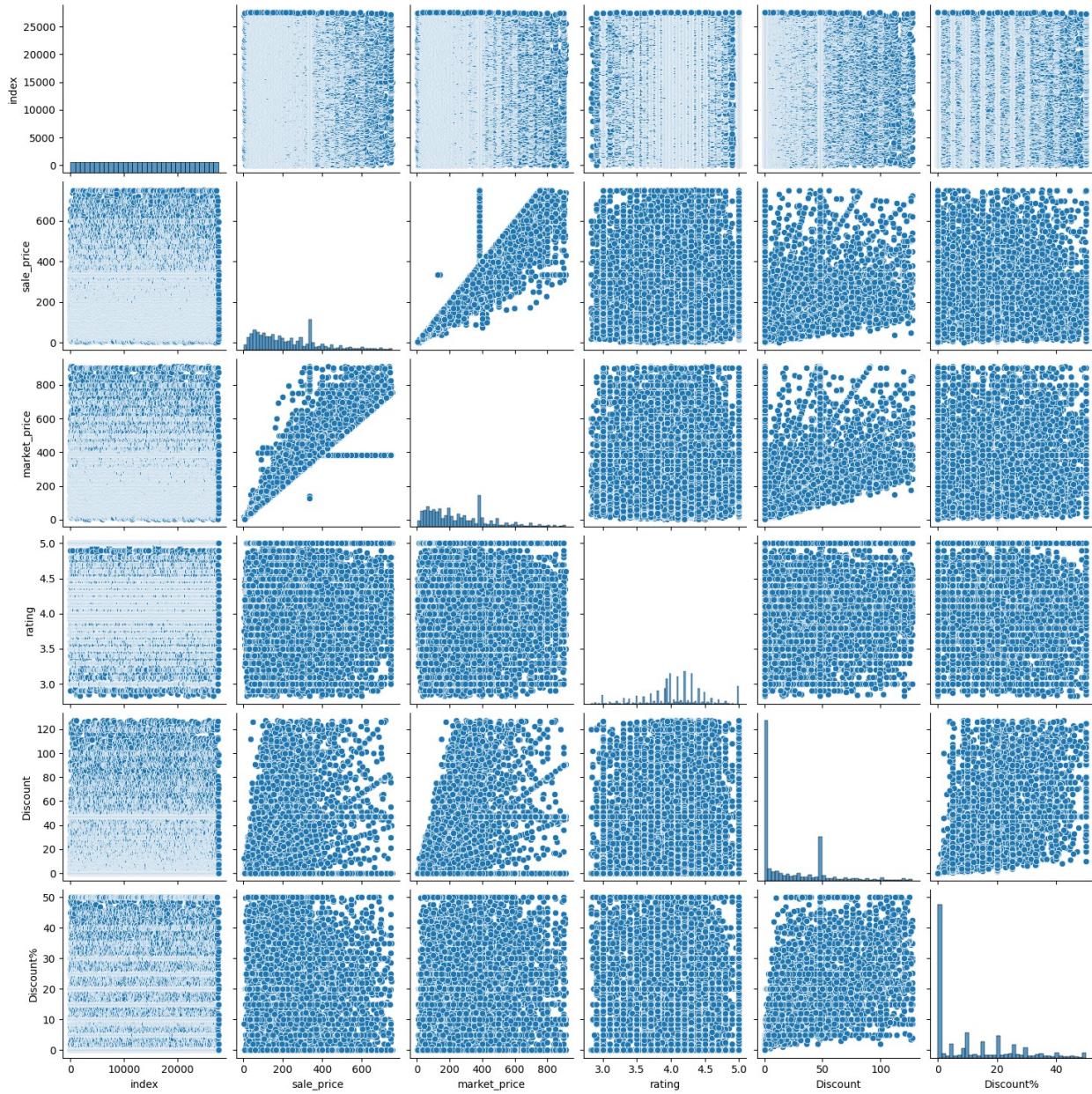


```
# Replace negative discount% values to zero
bb['Discount%'] = bb['Discount%'].where(bb['Discount%'] >= 0, 0)

discount_negative = bb[bb['Discount%']<0]
discount_negative

{"repr_error": "Out of range float values are not JSON compliant: nan", "type": "dataframe", "variable_name": "discount_negative"}

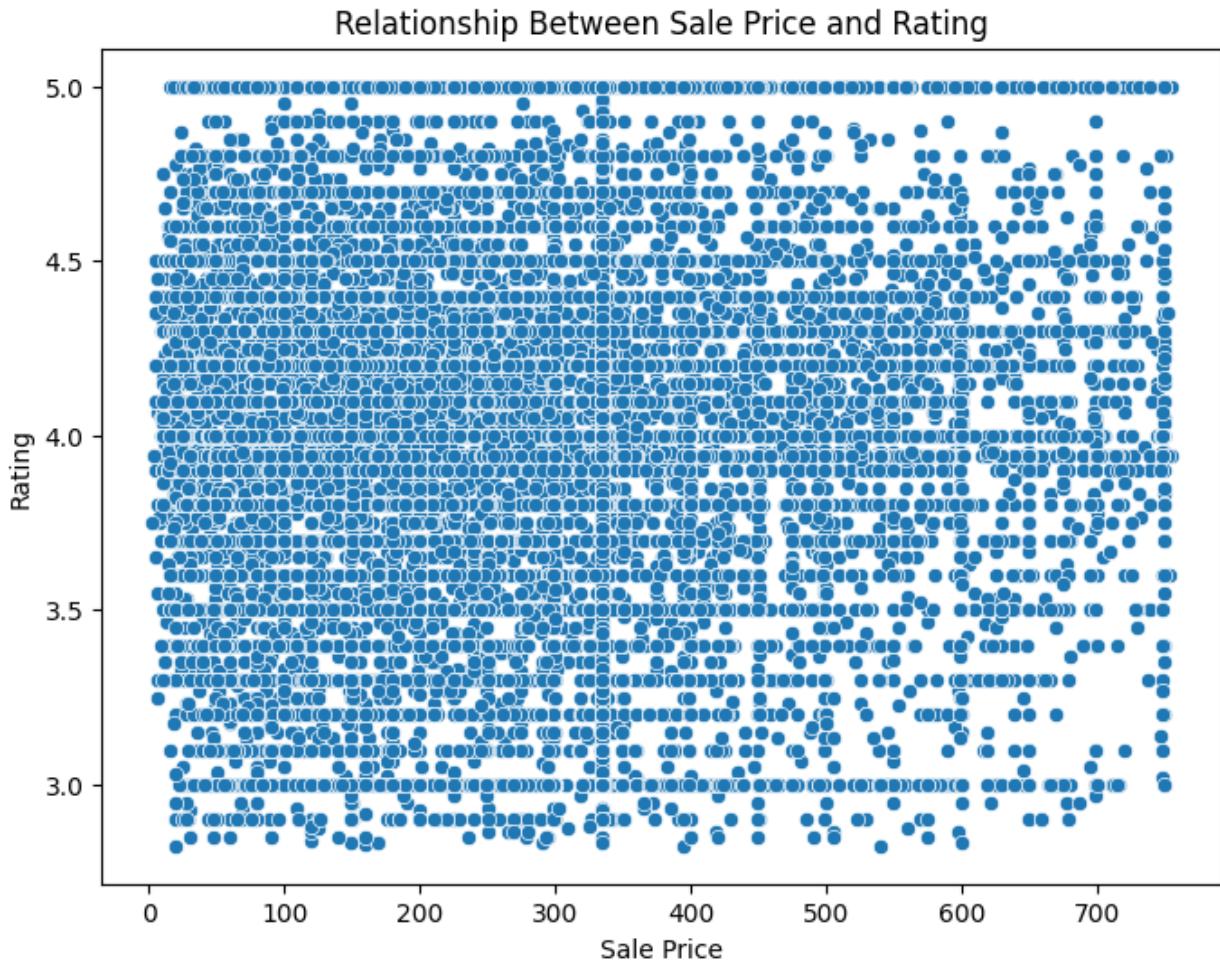
sns.pairplot(bb)
plt.show()
```



```
# Create scatter plot using seaborn
import matplotlib.pyplot as plt
import seaborn as sns
plt.figure(figsize=(8, 6))
sns.scatterplot(x='sale_price', y='rating', data=bb[bb['sale_price'] < 20000])

plt.title('Relationship Between Sale Price and Rating')
plt.xlabel('Sale Price')
plt.ylabel('Rating')
```

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



bb

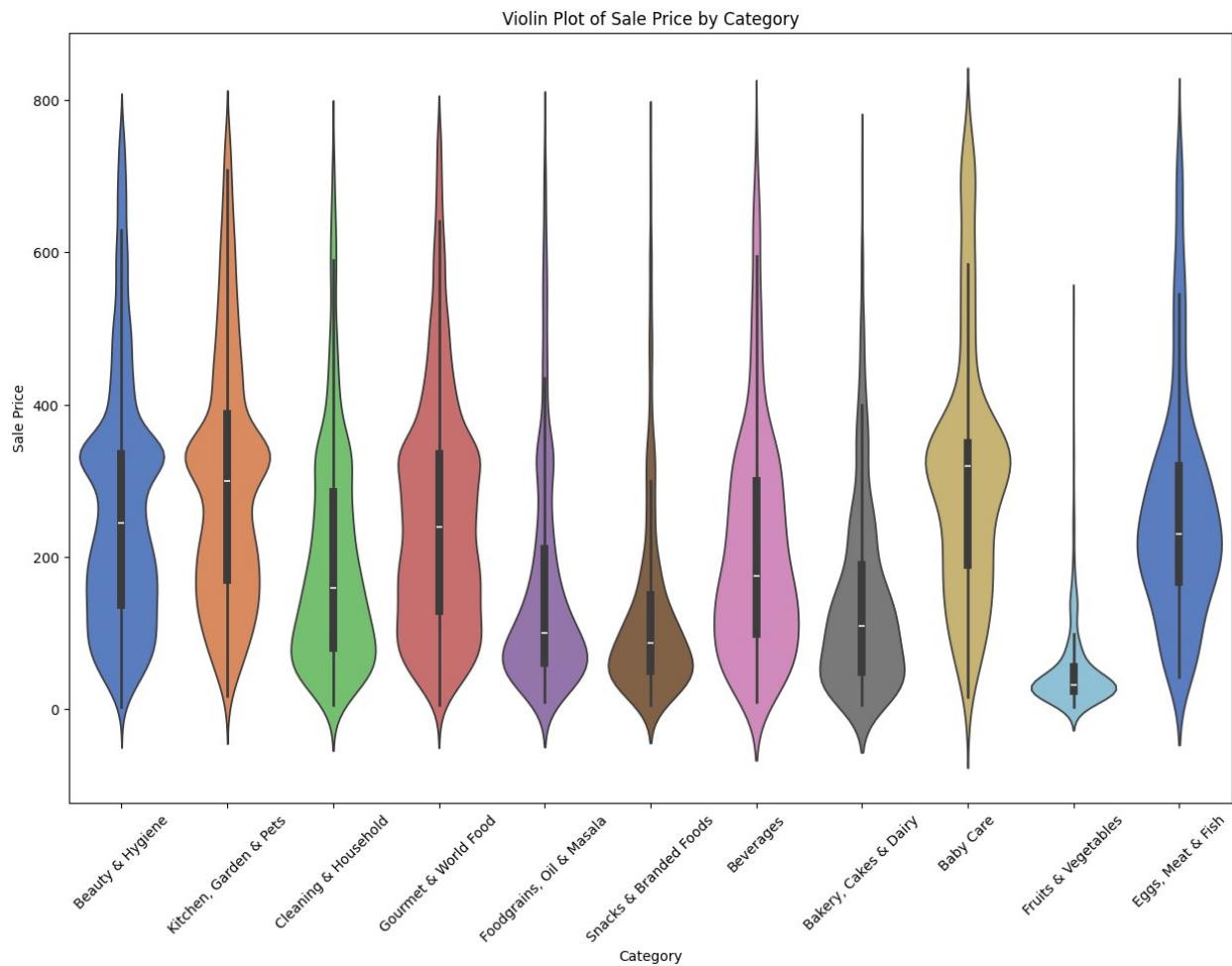
```
{"summary":{  
  "name": "bb",  
  "rows": 27555,  
  "fields": [  
    {  
      "column": "index",  
      "properties": {  
        "dtype": "number",  
        "std": 7954,  
        "min": 1,  
        "max": 27555,  
        "num_unique_values": 27555,  
        "samples": [16202, 24963, 18965]  
      },  
      "semantic_type": "",  
      "description": ""  
    },  
    {  
      "column": "product",  
      "properties": {  
        "dtype": "string",  
        "num_unique_values": 23540,  
        "samples": [  
          "Pure Impact Body Wash With Purifying Micro Particles - Shower Gel  
          For Body, Face & Hair",  
          "Premium Care XXL - 30 Diaper  
          Pants",  
          "Tomato - Green"],  
        "semantic_type": "",  
        "description": ""  
      },  
      "category",  
      "properties": {}  
    }  
  ]  
},  
"rows": 27555,  
"fields": [  
  {  
    "column": "index",  
    "properties": {  
      "dtype": "number",  
      "std": 7954,  
      "min": 1,  
      "max": 27555,  
      "num_unique_values": 27555,  
      "samples": [16202, 24963, 18965]  
    },  
    "semantic_type": "",  
    "description": ""  
  },  
  {  
    "column": "product",  
    "properties": {  
      "dtype": "string",  
      "num_unique_values": 23540,  
      "samples": [  
        "Pure Impact Body Wash With Purifying Micro Particles - Shower Gel  
        For Body, Face & Hair",  
        "Premium Care XXL - 30 Diaper  
        Pants",  
        "Tomato - Green"],  
      "semantic_type": "",  
      "description": ""  
    },  
    "category",  
    "properties": {}  
  }  
],  
"version": 1}
```

```
{\n      "dtype": "category",\n      "num_unique_values":\n11,\n      "samples": [\n          {"Snacks & Branded Foods",\n          "Beauty & Hygiene",\n          "Fruits & Vegetables"},\n          {"description": "\n          }\n      },\n      {"column": "sub_category",\n      "semantic_type": "\",\n      "properties": {\n          "dtype": "category",\n          "num_unique_values": 90,\n          "samples": [\n              "Crockery & Cutlery",\n              "Fragrances & Deos",\n              "Energy & Soft Drinks"],\n          "semantic_type": "\",\n          "description": "\n          }\n      },\n      {"column": "brand",\n      "properties": {\n          "dtype": "category",\n          "num_unique_values": 2313,\n          "samples": [\n              "Red Label",\n              "Lomani",\n              "Happily Unmarried"],\n          "semantic_type": "\",\n          "description": "\n          }\n      },\n      {"column": "sale_price",\n      "properties": {\n          "dtype": "number",\n          "std": 162.32593459912707,\n          "min": 2.45,\n          "max": 755.0,\n          "num_unique_values": 2432,\n          "samples": [\n              131.75,\n              469.64,\n              166.25\n          ],\n          "semantic_type": "\",\n          "description": "\n          }\n      },\n      {"column": "market_price",\n      "properties": {\n          "dtype": "number",\n          "std": 194.98763384063898,\n          "min": 3.0,\n          "max": 910.0,\n          "num_unique_values": 835,\n          "samples": [\n              402.0,\n              633.0,\n              477.0\n          ],\n          "semantic_type": "\",\n          "description": "\n          }\n      },\n      {"column": "type",\n      "properties": {\n          "dtype": "category",\n          "num_unique_values": 426,\n          "samples": [\n              "Nappies & Rash Cream",\n              "Hair Removal",\n              "Crackers & Digestive"],\n          "semantic_type": "\",\n          "description": "\n          }\n      },\n      {"column": "rating",\n      "properties": {\n          "dtype": "number",\n          "std": 0.44247678020336517,\n          "min": 2.825,\n          "max": 5.0,\n          "num_unique_values": 384,\n          "samples": [\n              3.4749999999999996,\n              4.214285714285714,\n              2.84\n          ],\n          "semantic_type": "\",\n          "description": "\n          }\n      },\n      {"column": "description",\n      "properties": {\n          "dtype": "string",\n          "num_unique_values": 21944,\n          "samples": [\n              "Premium Day Face Cream:Lightweight formula that will give your skin a daily boost of radiance and hydration. Packed with the goodness of skin brightening, Vitamin C Why Is It Special?- It contains the Most Stable & Effective form of Vitamin C, 3-O Ethyl Ascorbic Acid. It is the next generation Vitamin C that really works in improving skin health.- Vitamin C is a proven skin-brightening agent known for its antioxidant properties for decades. It boosts collagen formation and effectively targets visible signs of ageing, giving skin the much-needed glow. - The presence of Vitamin E, Pro-Vitamin B5 & Vitamin B3\n          }\n      }\n}
```

helps awaken and clarify your complexion, revealing a natural healthy-looking glow. - SPF 30 & PA+++ provide the highest level of protection against UVA & UVB rays of the sun. The mineral-based sun protection agents shield the skin against the harsh solar radiation, minimizing skin dryness and dullness. - Multiple natural oils are combined for intense hydration and nourishment. Antioxidant-rich botanical extracts are added to keep skin moisturized, smooth & soft. - St.Botanica is a PETA-certified cruelty-free brand. The product does not contain any harmful chemicals & is free from parabens, silicones or mineral oils.\n",\n "Kari Kari Japanese premium snacks are the exciting irresistible new taste of crunch. Lite and crispy, the savoury rice crackers are mixed with peanuts to make hearty, wholesome snack with delicious, bold and zesty taste with every bite.\n",\n "Safely store your food in this pink lunch box. The lunch box is made plastic which gives the container durability and sturdiness. The rectangular containers come with an airtight lid to keep food fresh for long. The wide mouth opening of the lunch makes it easy to fill food and easy to clean. The set makes a great gift option for special days.\n\n ],\n {\n \"semantic\_type\": \"\",\n \"description\": \"\n \",\n {\n \"column\": \"Discount\",\n \"properties\": {\n \"dtype\": \"number\",\n \"std\": 30.04561889478489,\n \"min\": 0.0,\n \"max\": 127.5,\n \"num\_unique\_values\": 1555,\n \"samples\": [\n 74.57999999999998,\n 4.40000000000006,\n 5.34000000000003\n ],\n \"semantic\_type\": \"\",\n \"description\": \"\n \",\n {\n \"column\": \"Discount%\",\n \"properties\": {\n \"dtype\": \"number\",\n \"std\": 13.379322214696623,\n \"min\": 0.0,\n \"max\": 50.0,\n \"num\_unique\_values\": 2406,\n \"samples\": [\n 22.74,\n 30.06,\n 21.51\n ],\n \"semantic\_type\": \"\",\n \"description\": \"\n \",\n }\n ]\n },\n \"type\": \"dataframe\", \"variable\_name\": \"bb\"\n }\n # Create a violin plot\n plt.figure(figsize=(15, 10))\n sns.violinplot(x='category', y='sale\_price', data=bb, palette='muted')\n\n # Add title and labels\n plt.title('Violin Plot of Sale Price by Category')\n plt.xlabel('Category')\n plt.ylabel('Sale Price')\n plt.xticks(rotation=45)\n\n # Show the plot\n plt.show()\n\n <ipython-input-60-cdf39ea01b30>:3: FutureWarning:\n Passing `palette` without assigning `hue` is deprecated and will be\n\n

```
removed in v0.14.0. Assign the `x` variable to `hue` and set  
`legend=False` for the same effect.
```

```
sns.violinplot(x='category', y='sale_price', data=bb,  
palette='muted')
```



```
bb.describe(include="all")  
  
{"summary": {"name": "bb",  
 "rows": 11,  
 "fields": [  
 {"column": "index",  
 "properties": {},  
 "dtype": "number",  
 "std": 9930.650518628143,  
 "min": 1.0,  
 "max": 27555.0,  
 "num_unique_values": 6,  
 "samples": [13778.0, 20666.5]  
 ],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "product",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 4,  
 "samples": [23540, 26],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "category",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "brand",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "type",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "sub_type",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "sub_sub_type",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "unit",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "unit_size",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "unit_weight",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "unit_length",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "unit_width",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"},  
 {"column": "unit_height",  
 "properties": {},  
 "dtype": "category",  
 "num_unique_values": 11,  
 "samples": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11],  
 "semantic_type": "",  
 "description": "\n"}]
```



```
\"number\",  
    \"std\": 9737.155343867656,  
    \"min\":  
0.0,  
    \"max\": 27555.0,  
    \"num_unique_values\": 7,  
\"samples\": [  
    27555.0,  
    11.044069853147665,  
20.0  
],  
    \"semantic_type\": \"\",  
\"description\": \"\"\n        }  
    ]\n}","type":"dataframe"}
```