



BIG BASKET MINI PROJECT

PRESENTATION

AGENDA

Introduction

Understanding the dataset

Hypothesis

EDA

Observation

Recommendation

Conclusion

PROJECT INTRODUCTION

INTRODUCTION

- This project explores the e-commerce domain, specifically focusing on the analysis of data from Big Basket, India's largest online grocery platform.
- Launched in 2011, Big Basket has consistently maintained its position as a market leader despite the emergence of competitors like Blinkit, thanks to its widespread popularity and a strong online presence.

E-commerce Overview:

- Buying and selling goods online.
- Key technologies: mobile commerce, electronic funds transfer, supply chain management, online transaction processing, etc.
- Driven by advancements in the semiconductor industry.

WHAT TO EXPECT ?

- An overview of e-commerce and its technological backbone.
- Exploratory Data Analysis (EDA) on Big Basket's dataset, featuring 10 key attributes, including product categories, pricing, brands, and customer ratings.
- Insights derived from the dataset using Python libraries and data visualization tools.



UNDERSTANDING THE DATASET

DATASET DESCRIPTION

Dataset Description:

- Contains 10 attributes.
- Captures essential aspects of e-commerce data.

Attributes:

- Index: Row identifier.
- Product: Product title.
- Category: Main category of the product.
- Sub-category: Subdivision under the main category.
- Brand: Product brand.
- Sale Price: Discounted price.
- Market Price: Original market price.
- Type: Classification type.
- Rating: Customer rating.
- Description: Textual data details.



KEY STEPS TO PERFORM

KEY STEPS TO PERFORM

1. Load Dataset.
2. Use head function to look for first 12 rows.
3. Get Description of the data in the Data Frame.
4. Find Information about the Data Frame.
5. Find out the Missing Values from the Dataset and treat them.
6. Find out the outliers from the dataset and treat them.
7. Find out Top & least sold products.
8. Measuring discount on a certain item.
9. Create Plots or visualizations
10. Hypothesis

HYPOTHESIS



HYPOTHESIS 1

**Beverages Are the Most Popular
and Highly Rated Category Apart
From Fresh Groceries Category.**

Rating based on category Beverages

– 4.08



HYPOTHESIS 2

Higher Discounts Lead to Lower Product Ratings

- **Based on correlation between Discount% and Rating, we found that the correlation between them is around 0.**



HYPOTHESIS 3

Premium Brands Offer Fewer Discounts.

Validation - We've analyzed the average discount percentage across different brands.

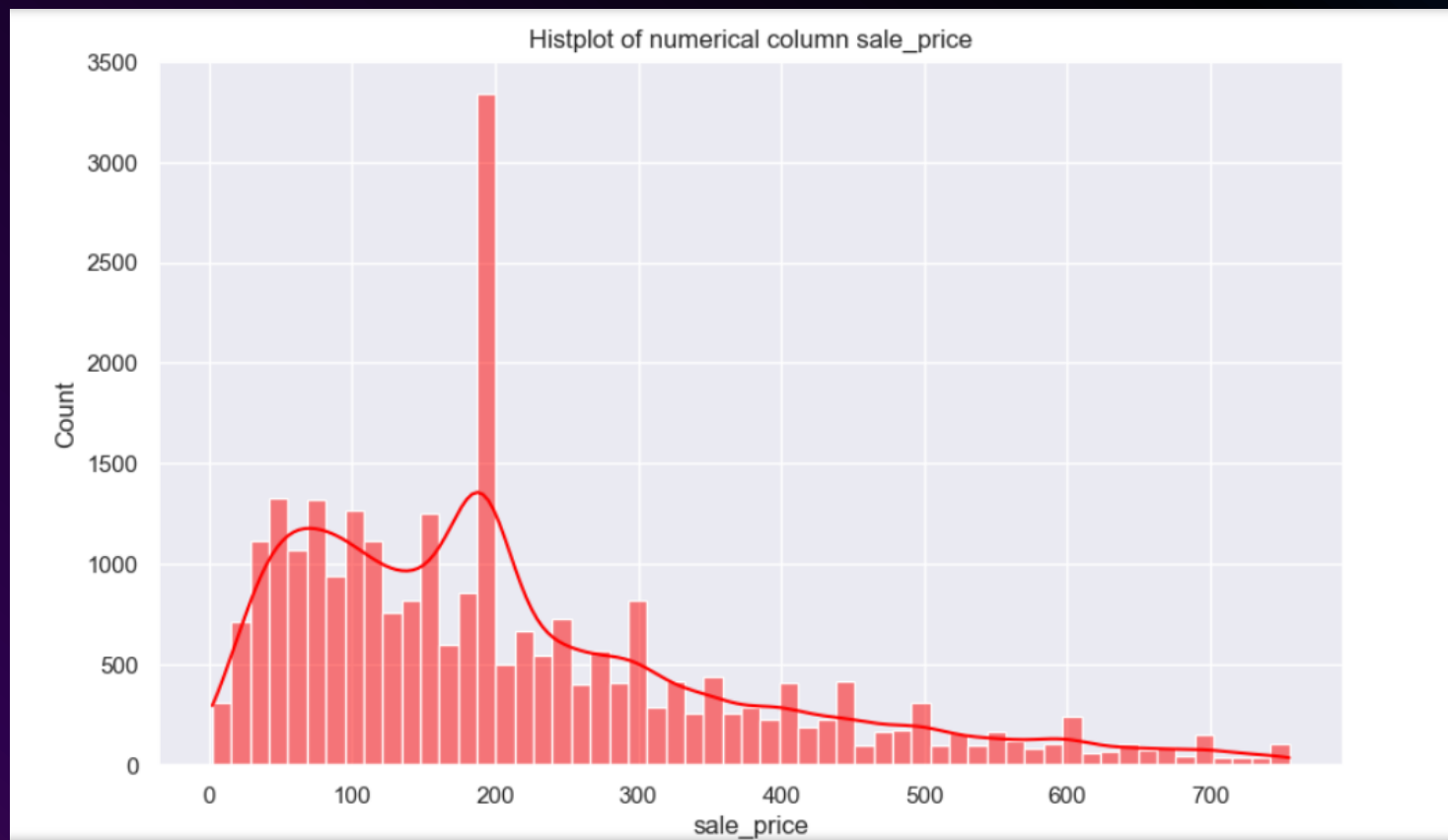


EXPLORATORY DATA ANALYSIS

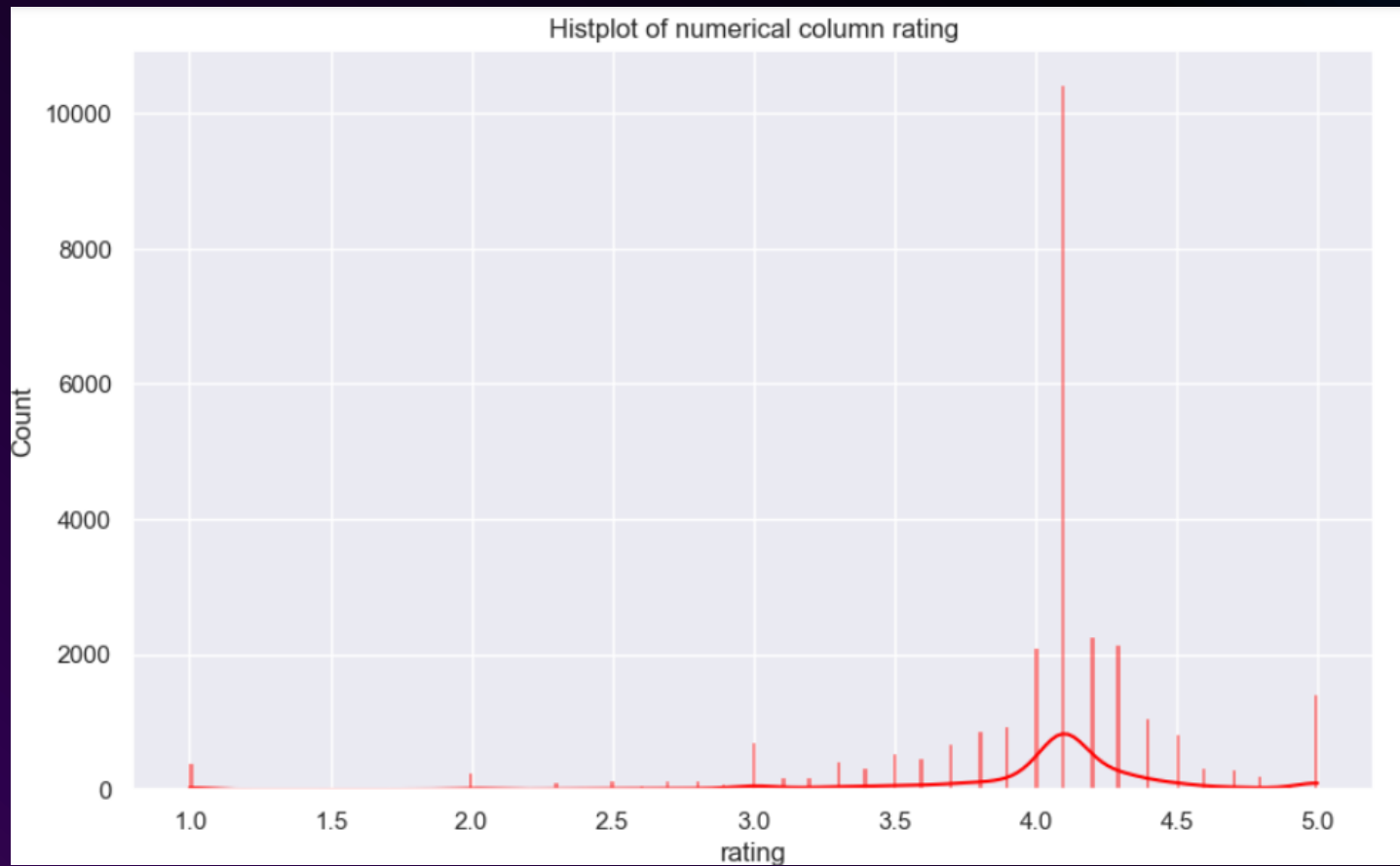
EDA

- Top Selling Product -- Turmeric Powder/Arisina Pudi
- Least Selling Product -- Green Tea - Pure Original

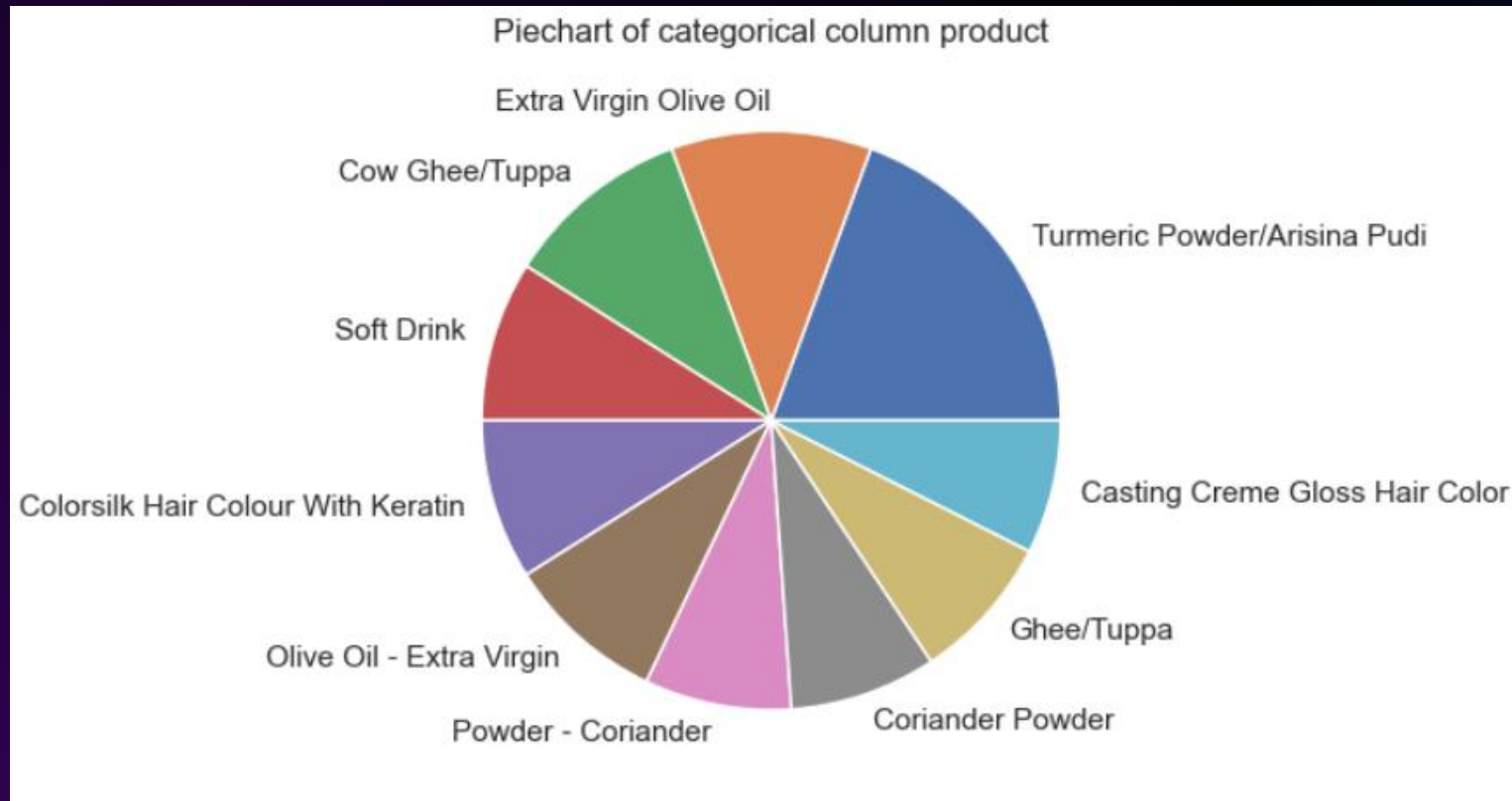
EDA - GRAPHICS



EDA - GRAPHICS

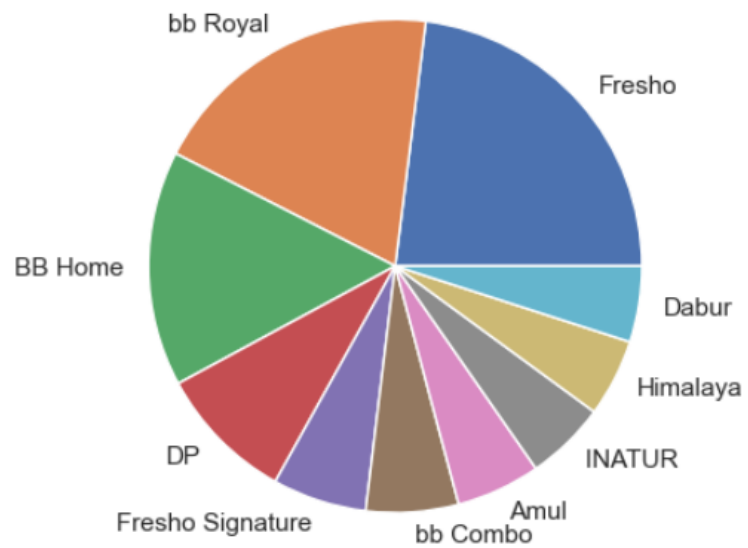


EDA - GRAPHICS

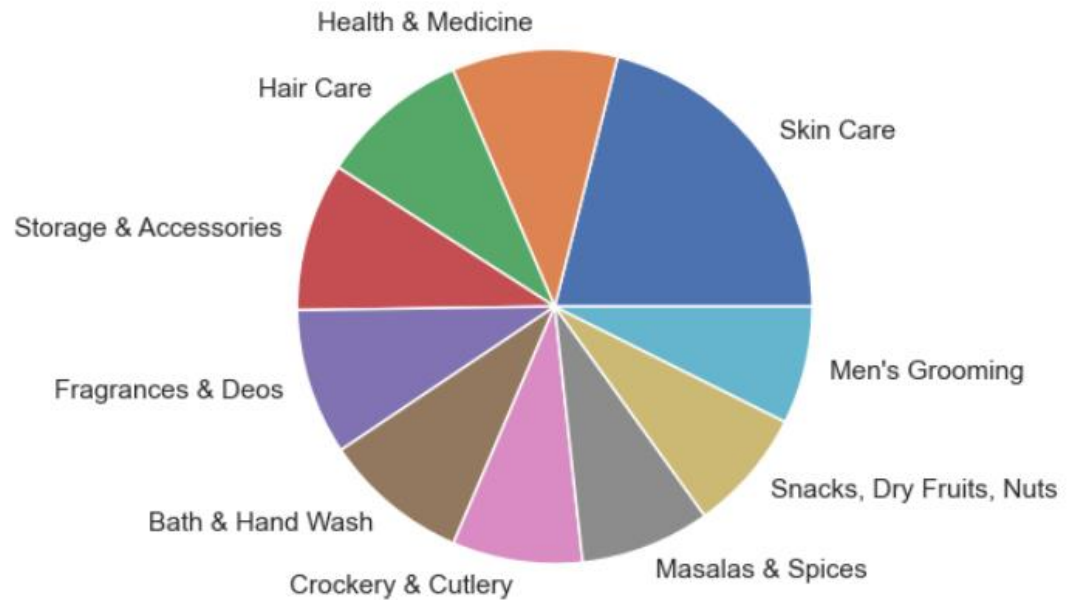


EDA - GRAPHICS

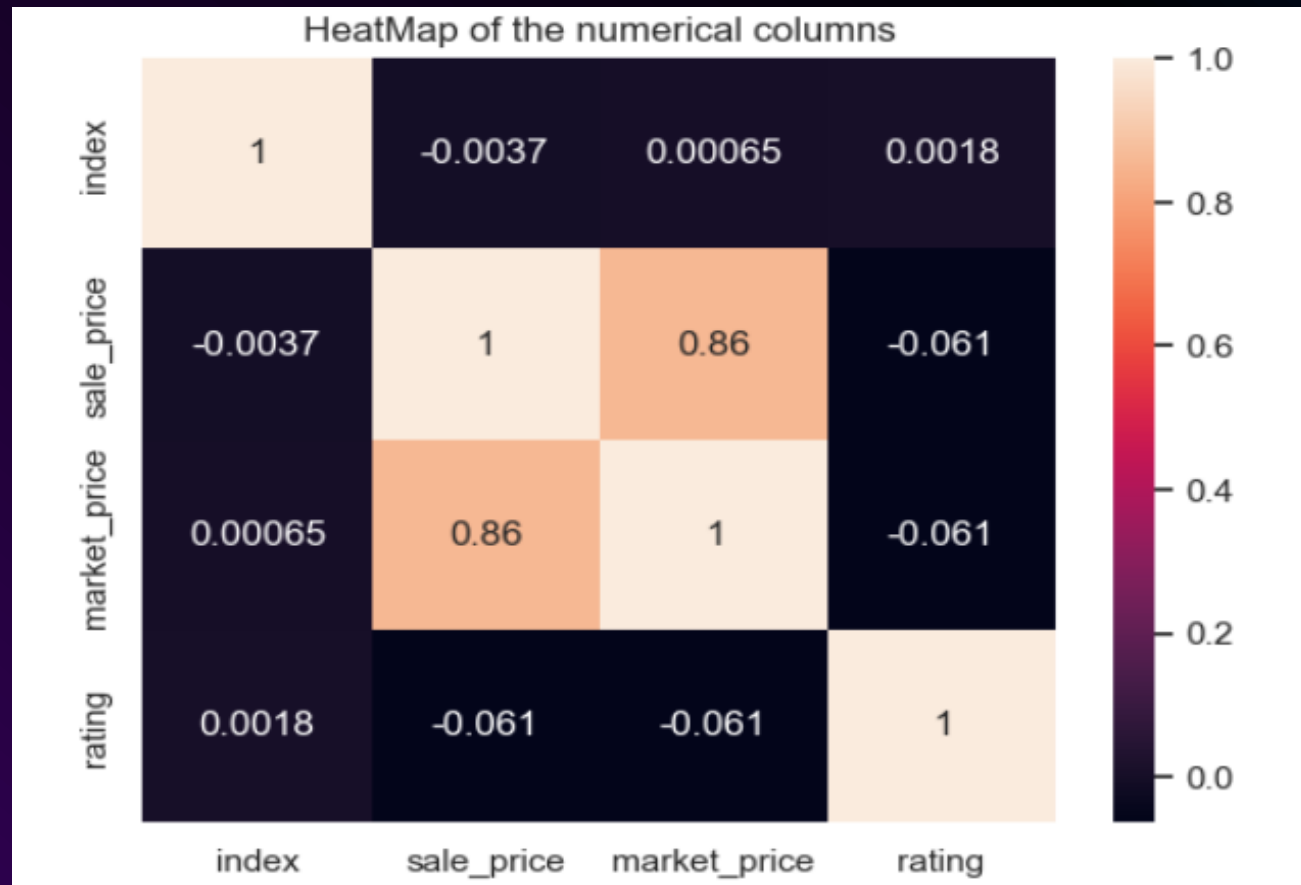
Piechart of categorical column brand



Piechart of categorical column sub_category



EDA - GRAPHICS



OBSERVATIONS

- **Beverages** is the **highest-rated category** with an average rating of **4.08**, indicating strong customer satisfaction.
- There is **no significant correlation** between higher discounts and lower ratings (-0.0027 correlation).
- Some brands provide **zero discounts**
- **Foodgrains, Oil & Masala (4.06)** and **Baby Care (4.02)** are also among the top-rated categories.

RECOMMENDATION

Leverage High-Rated Categories:

- Promote and expand product offerings in **Beverages, Foodgrains, and Baby Care**, as these categories have the highest customer satisfaction.

Refine Discount Strategies:

- Since discounts do not significantly impact ratings, focus on **value-driven pricing** rather than excessive discounts.

Improve Ratings for Low-Rated Categories:

- Conduct customer surveys for **Kitchen, Garden & Pets (3.73)** and **Beauty & Hygiene (3.93)** to identify areas for improvement.

CONCLUSION

- **Beverages** leads in customer satisfaction, making them a high-priority category for marketing and sales strategies.
- **Discounting does not significantly impact ratings**, so brands should focus on quality and customer experience instead.
- Optimizing the product mix based on customer ratings can **enhance sales and brand reputation**.

THANK YOU

