

Nykaa Product Review Analysis 2021

Introduction

When purchasing products from Nykaa, three key factors often influence decision-making: **price, rating, and review count**. These metrics provide valuable insights into the product's affordability, quality, and popularity among consumers. The motivation behind this analysis was to explore these factors to understand their impact on consumer choices and brand performance.

In this analysis, we sought to answer the following key questions:

- Which are the highest-rated brands?
Identifying the top-rated brands helps consumers make informed choices based on overall satisfaction.
- Which are the costliest brands?
Understanding price distribution across brands can offer insights into market positioning and the presence of luxury products.
- Is there a relationship between price and rating?
We explored whether there is a trend where cheaper products receive higher ratings, which could indicate consumer preference for value-for-money options.
- Do cheaper products tend to receive higher ratings?
Analysing the review count can provide insights into the product's popularity and whether it correlates with higher ratings.

Data Overview:

The dataset was taken from kaggle

<https://www.kaggle.com/datasets/susant4learning/nykaacosmeticsproductsreview2021>

Which has 625 rows and 18 columns, including key columns like Product Id, Product Name, Product Brand, Product Price, and Product Rating

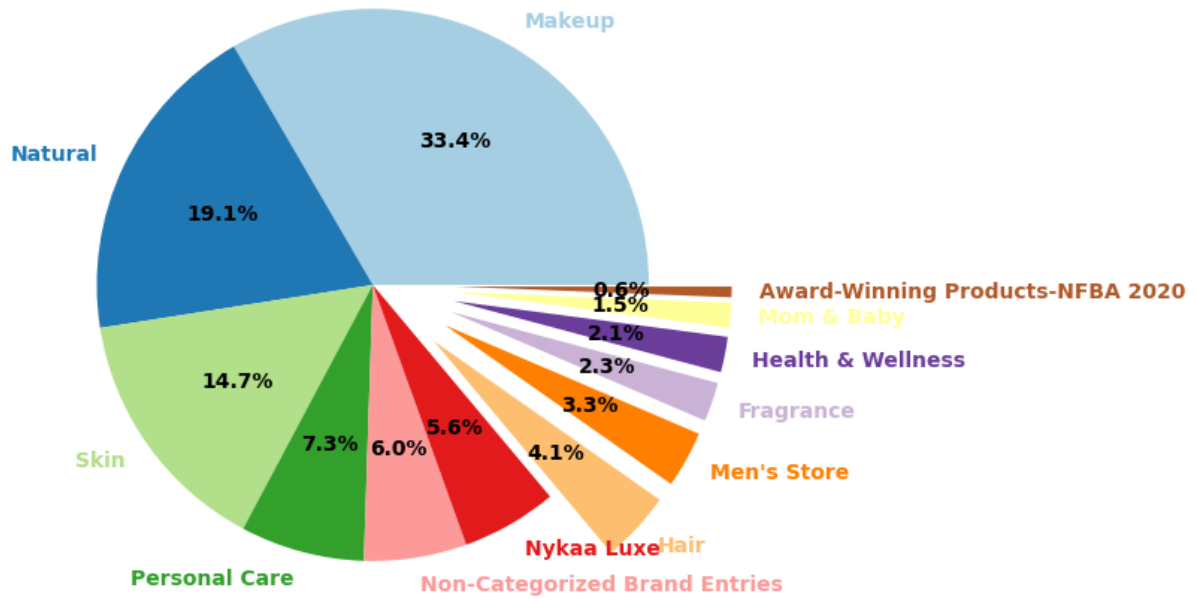
Steps taken for Data Wrangling:

- Filtering the required columns
- Renaming the columns in proper convention
- Replacing irrelevant entries
- Removing missing values
- Removing duplicates
- Assigning proper data types to the column value
- Removing trail spaces from column value

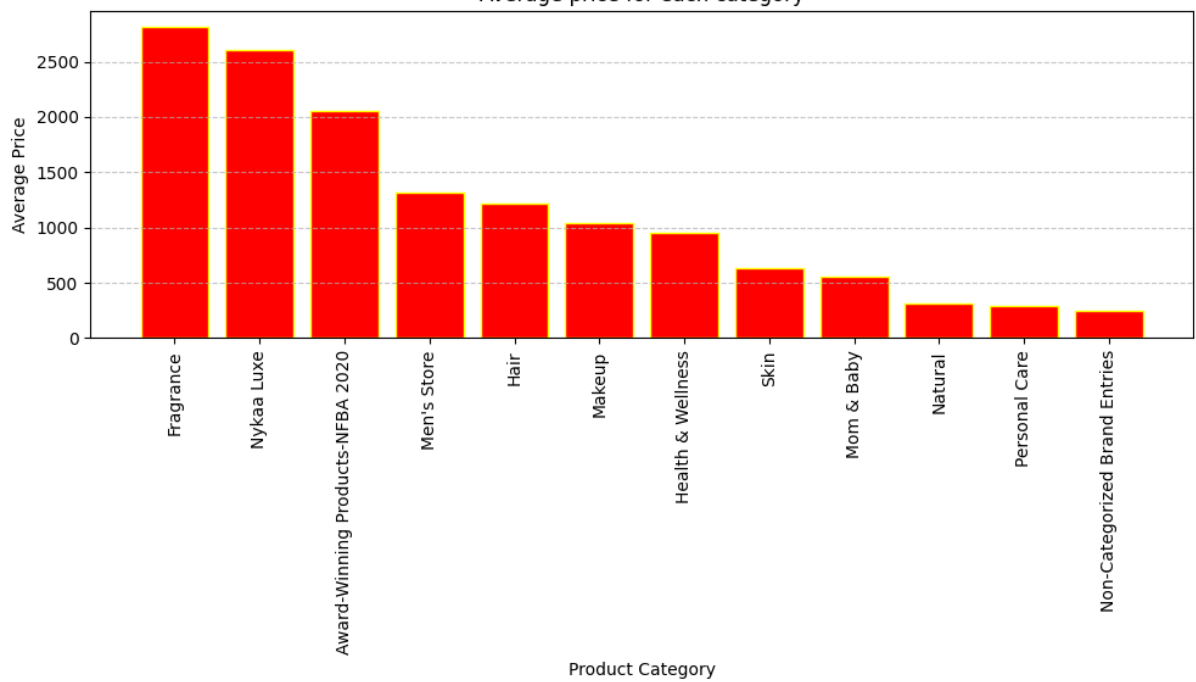
Analysis and Findings:

- Different types of Categories that are available and the one that dominates the market:

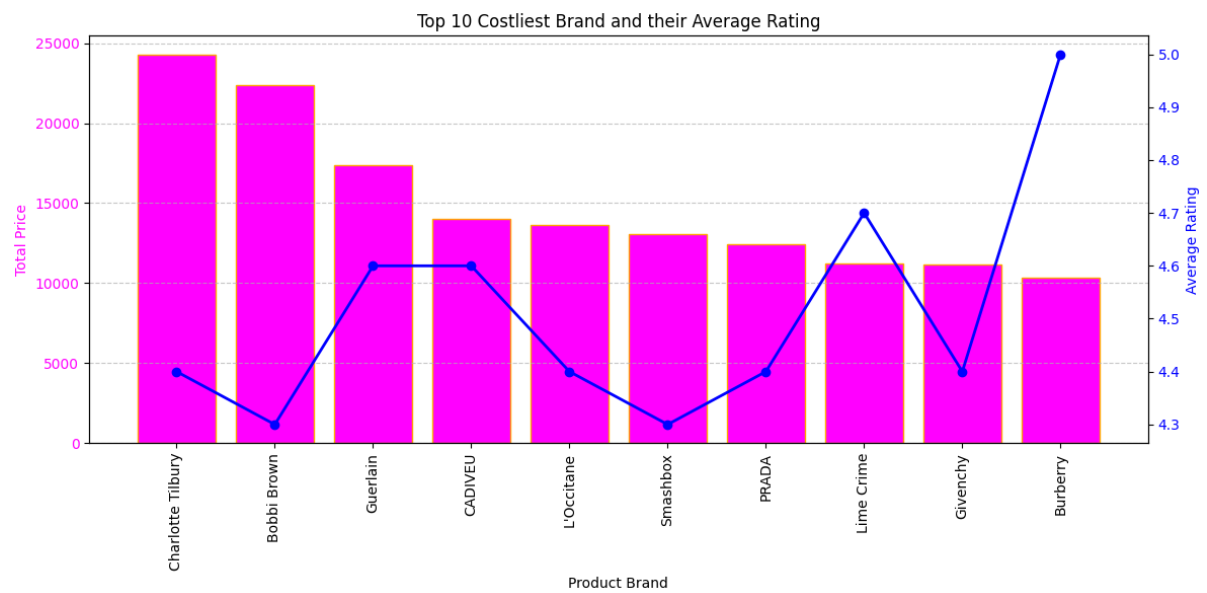
Product Category Distribution



Average price for each category

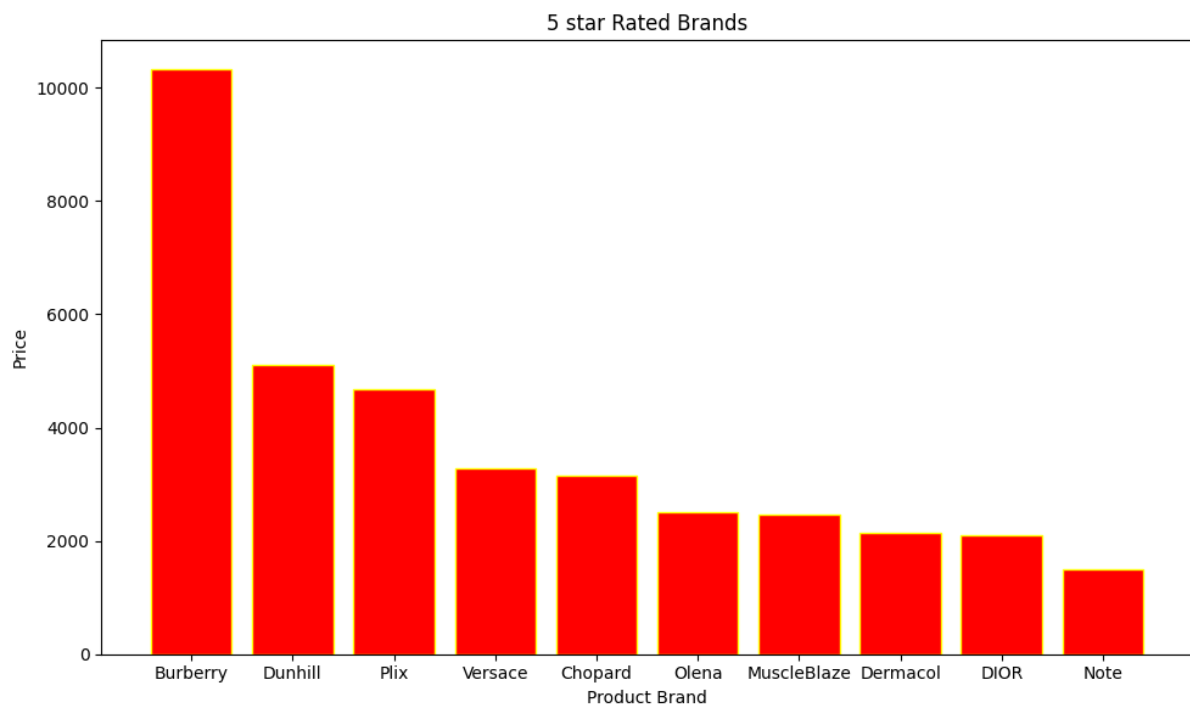


- Costliest Brands:



These Brands are usually considered as luxury brands. Here we can see that the costliest brand has the lower rating and the cheapest out of the 10 brands has the highest rating.

- Top10 5star Rated brands:



- Relationship between Price and Rating and Review Count:

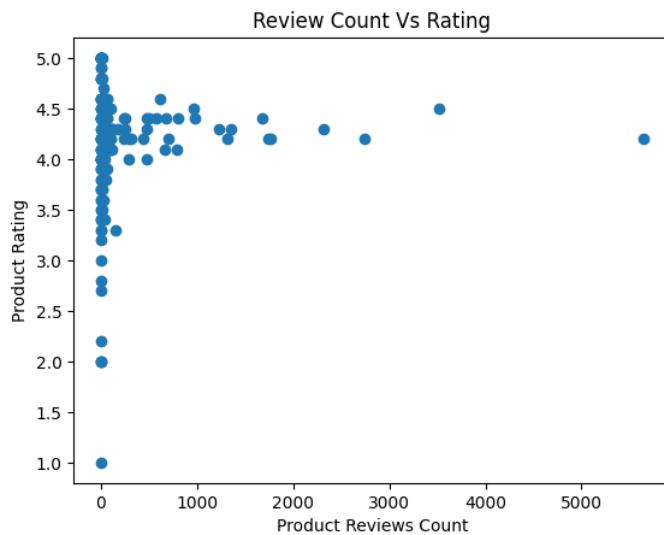
From the given dataset we were not able to establish any relationship, which can be explained by their correlation co-efficient value.

The correlation co-efficient value is always between -1 and 1 , closer to -1 indicates a strong negative correlation and value closer to +1 indicates strong positive correlation , value closer to 0 indicates no relation between the variables.

In our case correlation coefficient table indicates no relation between the two variables:

	Total Price	Average Rating	Average Review Count
Total Price	1.000000	0.135464	0.011916
Average Rating	0.135464	1.000000	0.033559
Average Review Count	0.011916	0.033559	1.000000

Which is again confirmed by the scatter plots



Conclusion:

The analysis of Nykaa product reviews has provided several valuable insights into consumer behavior and brand performance.

- We found that certain brands consistently receive high ratings, indicating strong consumer satisfaction with their products.
- The costliest brands tend to be associated with luxury, reflecting their market positioning.
- Interestingly, the analysis revealed that there is no strong correlation between product price and rating, challenging the assumption that cheaper products are more likely to receive higher ratings. This suggests that while price is an important factor, it is not the sole determinant of consumer satisfaction.
- Products with a higher review count often have a balanced rating, indicating that while they are popular, the quality may vary.

Through this project, I gained deeper insights into data analysis techniques, including data cleaning, exploratory data analysis, and correlation analysis. I also learned the importance of visualizing data to uncover patterns and trends.