

Sephora Product Analysis Report

Overview

This report provides an analysis of Sephora's product dataset with a focus on understanding product ratings, price distribution, and consumer engagement as measured by the number of favourites (loves_count). Additionally, an analysis was conducted on high-priced products within the skincare category that received positive reviews. Insights into top-rated products and brands, as well as category-wise average ratings, are presented to identify key trends in product preferences and value perception.

Data Overview

The dataset contains detailed information about a variety of products listed on the platform, containing attributes related to product identification, pricing, categories, and customer engagement. Below is a description of the key columns in the dataset:

- **product_id**: A unique identifier assigned to each product.
- **product_name**: The name of the product.
- **brand_id**: A unique identifier assigned to each brand.
- **brand_name**: The name of the brand associated with the product.
- **price_usd**: The price of the product in US dollars (USD).
- **rating**: The customer rating of the product, typically ranging from 0 to 5.
- **loves_count**: The number of times users have added the product to their favourites (or "loves").
- **reviews**: The total number of reviews submitted for the product by customers.
- **primary_category**: The main category to which the product belongs (e.g., Skincare, Makeup).
- **secondary_category**: A sub-category of the product, providing further classification (e.g., Moisturizers, Lipstick).
- **size**: The size of the product, which may be in various units such as ounces (oz), millilitres (mL), grams (g), or units (packs).
- **ingredients**: A list of ingredients in the product, which may include multiple variations of the product (e.g., different formulas).
- **value_price_usd**: The potential savings or discounted value of the product as displayed on the site.
- **tertiary_category**: A third-level category providing further granularity in classification.

Note: While this analysis focuses on the columns mentioned above, the dataset includes additional columns that are not utilized in this specific analysis.

Key Statistics

Average Rating: 4.2 (ranging from 1.0 to 5.0)

Average Price: \$51.66 (ranging from \$3.00 to \$1900.00)

Most products are rated between 3.9 and 4.5, and the majority of products are priced between \$25.00 and \$58.00.

Data Cleaning and Transformation

Before proceeding with the analysis, data cleaning and transformation steps were applied to ensure accuracy and consistency. Initial data processing involved checking for missing values and duplicates, followed by data cleaning such as correcting improperly formatted values and removing rows with missing ratings. The subset of the dataset relevant for rating and price analysis was further refined to exclude unnecessary data.

Handling Missing Values

Missing values were identified in the dataset, particularly in the 'rating' and 'reviews' columns. Given that the proportion of missing values was relatively small, rows containing missing values in these key columns were removed from the dataset.

Duplicate Records

The dataset was checked for duplicate records, and no significant duplication issues were found. The number of duplicate rows in the subset was minimal, and no further actions were required.

Data Formatting

To maintain consistency in the presentation, product and brand names were formatted to title case, ensuring uniformity across the dataset.

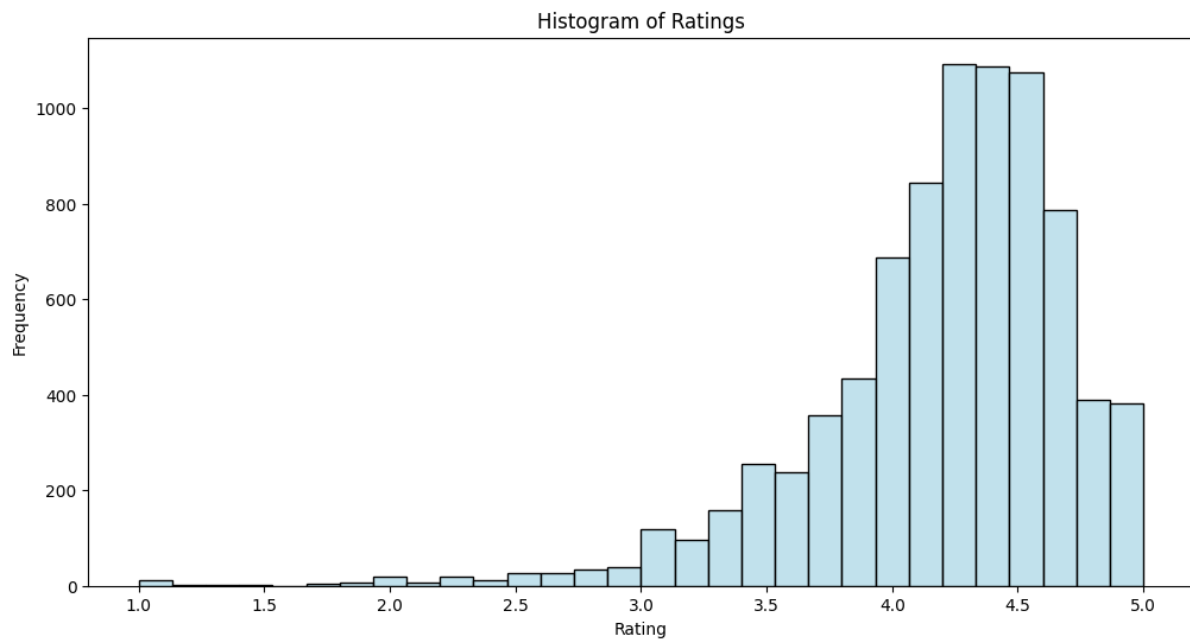
Data Sub setting

Several data subsets were created to facilitate specific analyses. For example, a subset focusing on skincare products was used to analyse high-priced products with positive reviews. This subset excluded high-tech tools, as they were not relevant to the analysis of skincare products.

Analysis

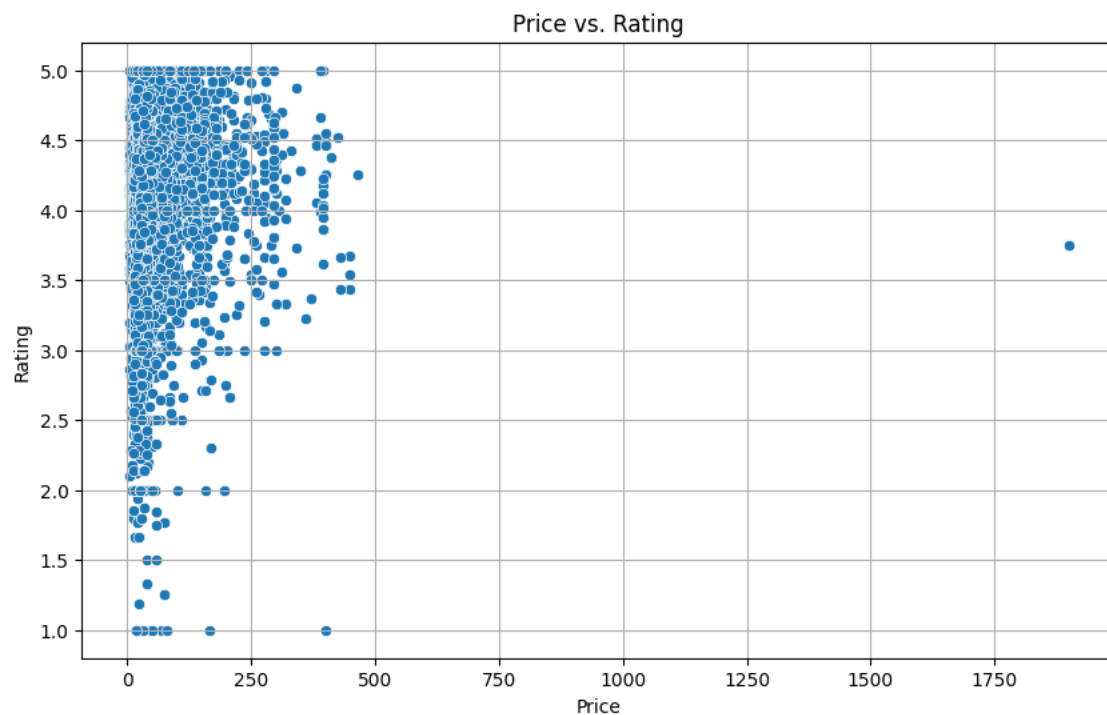
Analysis 1: Distribution of Product Ratings

A histogram analysis of the product ratings shows a right-skewed distribution, indicating that the majority of products are highly rated. The bulk of the ratings fall between 4.0 and 4.5, with very few products receiving ratings below 3.0. This suggests that most products in the dataset receive favourable feedback from consumers.



Analysis 2: Price vs. Rating

A scatter plot of product price versus rating was created to examine if there is any correlation between a product's price and its rating. The analysis shows that most products are priced below \$250, and there is no clear trend indicating that higher-priced products tend to have higher ratings. One outlier was identified, a product priced at approximately \$1750, which does not conform to the overall distribution. This suggests that while product pricing may influence consumer perceptions, it does not necessarily dictate product satisfaction.



Analysis 3: Top 10 Highest-Rated Products

Using a subset of the data, the top 10 highest-rated products were identified. This analysis extracted the products with the highest ratings for each brand and ranked them accordingly. The following are some of the highest-rated products based on this analysis:

ProductName	BrandName	Rating
Delikate Try Me Kit	Kate Somerville	5.0
Forest Stone Solid Refillable Body Moisturizer	Kate Mcleod	5.0
Good Girl Gone Bad Eau Fraîche	Kilian Paris	5.0
The Sun Show Glowly Warmth Talc-Free Baked Bronzer	Kosas	5.0
Mini Resistance Heat Protecting Leave-In Treat...	Kérastase	5.0
Mini Advanced Génifique Radiance Boosting Face...	Lancôme	5.0
Satin Luxe Classic Cream Lipstick - Baby	Lawless	5.0
By The Fireplace And Beach Walk Perfume Duo Set	Maison Margiela	5.0
E 3 Makeup Brush	Makeup By Mario	5.0
Mini Perfect Eau De Parfum Set	Marc Jacobs Fragrances	5.0

These findings can be useful for identifying the most favoured products in the marketplace.

Analysis 4: Top 10 Brands by Average Rating

The top 10 brands with the highest average ratings were identified by calculating the mean rating for each brand. The analysis revealed the following brands as the highest-rated:

BrandName	AvgRating
Erno Laszlo	5.000000
Aquis	4.904800
Macrene Actives	4.889420
Mara	4.823860
Canopy	4.813733
Dominique Cosmetics	4.785700
Bondiboost	4.776950
Damdam	4.752775
Kate Mcleod	4.751583
Maude	4.740100

This analysis highlights the brands that consistently receive high ratings across their product offerings.

Analysis 5: High-Priced Products with Positive Reviews

A subset of high-priced products in the skincare category was analysed to identify those that are considered "worth the money" based on their high ratings and positive reviews. Products priced above \$300.00 with ratings higher than 4.0 were examined. The top 10 most expensive and positively reviewed products are listed below:

ProductName	BrandName	PriceUsd	Rating	Reviews
Crème de la Mer Moisturizer	La Mer	380.0	4.0566	1377.0
The Concentrate Serum	La Mer	425.0	4.5187	563.0
The Regenerating Serum	La Mer	410.0	4.3750	152.0
The Moisturizing Soft Cream Moisturizer	La Mer	380.0	4.5146	103.0
The Moisturizing Soft Lotion	La Mer	320.0	4.0704	71.0
The Moisturizing Matte Lotion	La Mer	320.0	4.2258	62.0
Ultimate Revival Cream	SK-II	400.0	4.4667	60.0
Night Serum	Dr. Barbara Sturm	310.0	4.7021	47.0
Crème de La Mer Duet	La Mer	380.0	4.4615	13.0
Ultimate Revival Essence	SK-II	315.0	4.5455	11.0

This analysis reveals that even within the high-end product segment, certain products receive a significant amount of consumer approval.

Analysis 6: Average Ratings by Product Category

An analysis of average ratings by product category revealed the following insights:

PrimaryCategory	AvgRatings
Gifts	4.563450
Men	4.504992
Tools & Brushes	4.271458
Fragrance	4.230889
Skincare	4.228890
Hair	4.201113
Bath & Body	4.195015
Makeup	4.146845
Mini Size	4.005665

The highest-rated category was Gifts, while the lowest-rated category was Mini Size. This information could be leveraged to identify areas where product quality and satisfaction could be improved.

Conclusion

The analysis shows that most products in Sephora's dataset receive high ratings, and there is no strong correlation between price and rating. Top-rated products and brands were identified, and high-end skincare products were shown to be highly regarded by consumers. The insights gained from this analysis could be useful for identifying consumer preferences and potential areas of improvement for product development and marketing strategies.

Impact on Product Development:

By leveraging these findings, Sephora can focus its product development efforts on enhancing product quality in underperforming categories. Particularly, the Mini Size category shows room for improvement, and addressing the concerns in this space could prevent customer dissatisfaction and reduce product returns. Additionally, continuous innovation in the highest-rated segments, such as Gifts, can further solidify Sephora's reputation for delivering high-value offerings that align with consumer desires.

Impact on Marketing Strategies:

The analysis suggests that price alone does not dictate customer satisfaction. This insight can guide marketing strategies to emphasize the value and unique attributes of products rather than just their cost. Sephora can highlight customer reviews, testimonials, and product performance in its marketing campaigns to strengthen consumer trust. Focusing promotional efforts on high-rated products in categories like Skincare and Men's Grooming will allow Sephora to drive higher engagement and sales by targeting consumers' trust in these segments.

Impact on Customer Experience:

Improving customer experience is crucial for sustaining Sephora's competitive edge. The findings reveal opportunities for gathering consumer feedback on lower-rated categories, such as Mini Size, to identify specific pain points. In addition, Sephora can use this data to improve product features, packaging, or size offerings based on customer preferences. Additionally, enhancing customer education on high-end product value (e.g., explaining the science behind premium skincare ingredients) could drive greater satisfaction and loyalty among discerning consumers.

In conclusion, this analysis provides actionable insights that Sephora can use to refine its product portfolio, optimize marketing efforts, and enhance the overall customer experience, ensuring it continues to thrive in a competitive retail landscape.

Business Recommendations Based on Analysis:

- 1. Focus marketing and promotion efforts on the top-performing category:** The category with the highest average rating is **Gifts**. Concentrating on this segment can help leverage customer satisfaction to boost sales further.

2. **Improve product quality or re-evaluate inventory for the lowest-rated category:** The category with the lowest rating, **Mini Size** , requires attention. Consider improving product quality or reducing inventory to avoid customer dissatisfaction.

3. **Conduct customer feedback surveys:** Gathering insights from customers about the 'Mini Size' category can help identify specific pain points and areas for improvement.

4. **New products and cross-category bundles:** Launch new products in the high-performing category (Gifts) or offer cross-category bundles, which may attract customers and boost overall sales.