

Data Visualization

Name of Project:

 Final Project Presentation - Sephora
Product Analysis

Presented By:

Rumeysa Mercimek

Presented To:

Dina Deifallah

Overview

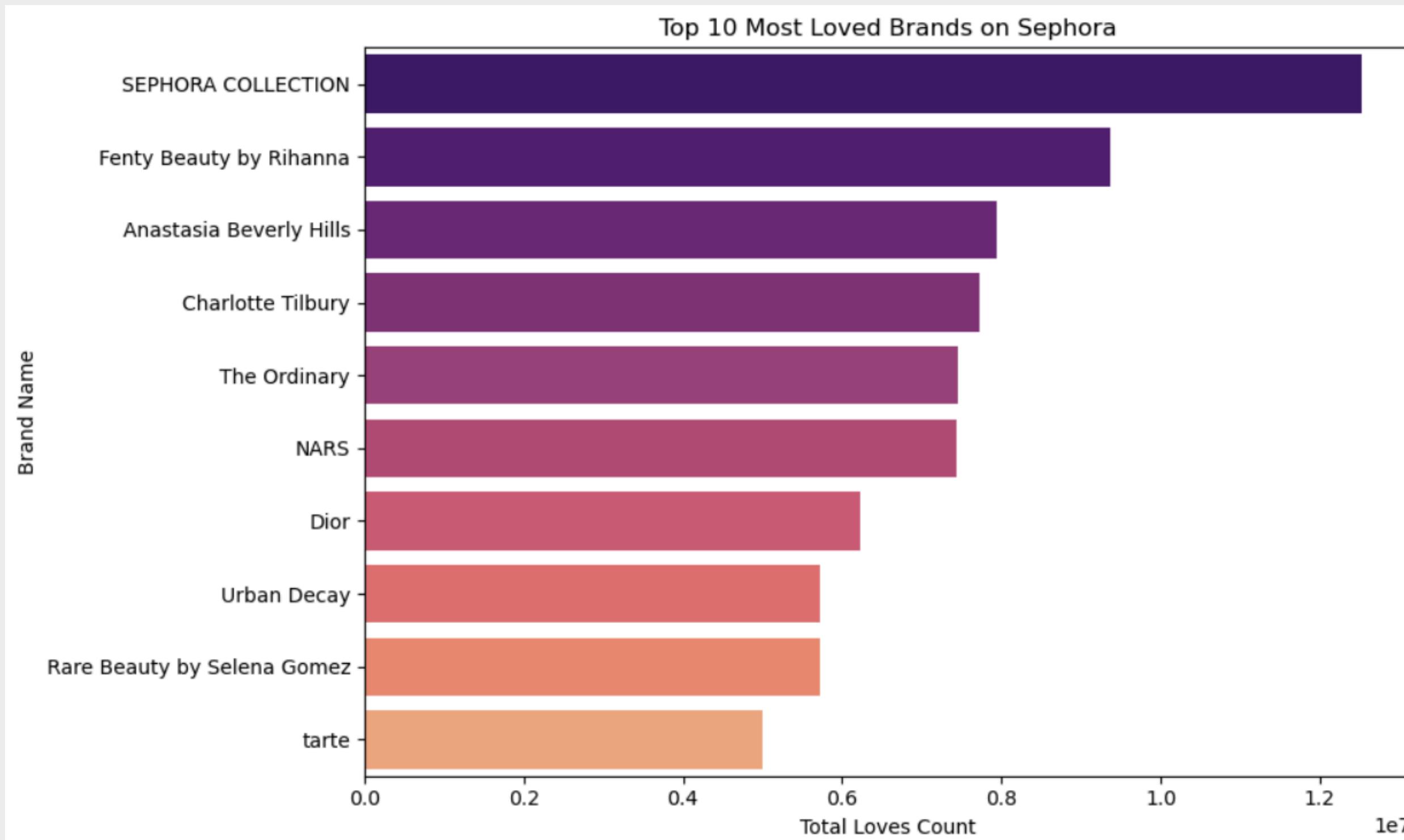
Our analysis of Sephora products reveals clear trends in customer preferences: Consumers are not necessarily drawn to the most expensive items. In fact, mid-range products (\$20-\$40) tend to receive higher ratings.

Certain brands consistently outperform others in terms of customer love, with some standing out across multiple categories.

Makeup and skincare dominate both in terms of product volume and average ratings. Overall, the data suggests that affordability, brand trust, and product type significantly influence customer satisfaction.

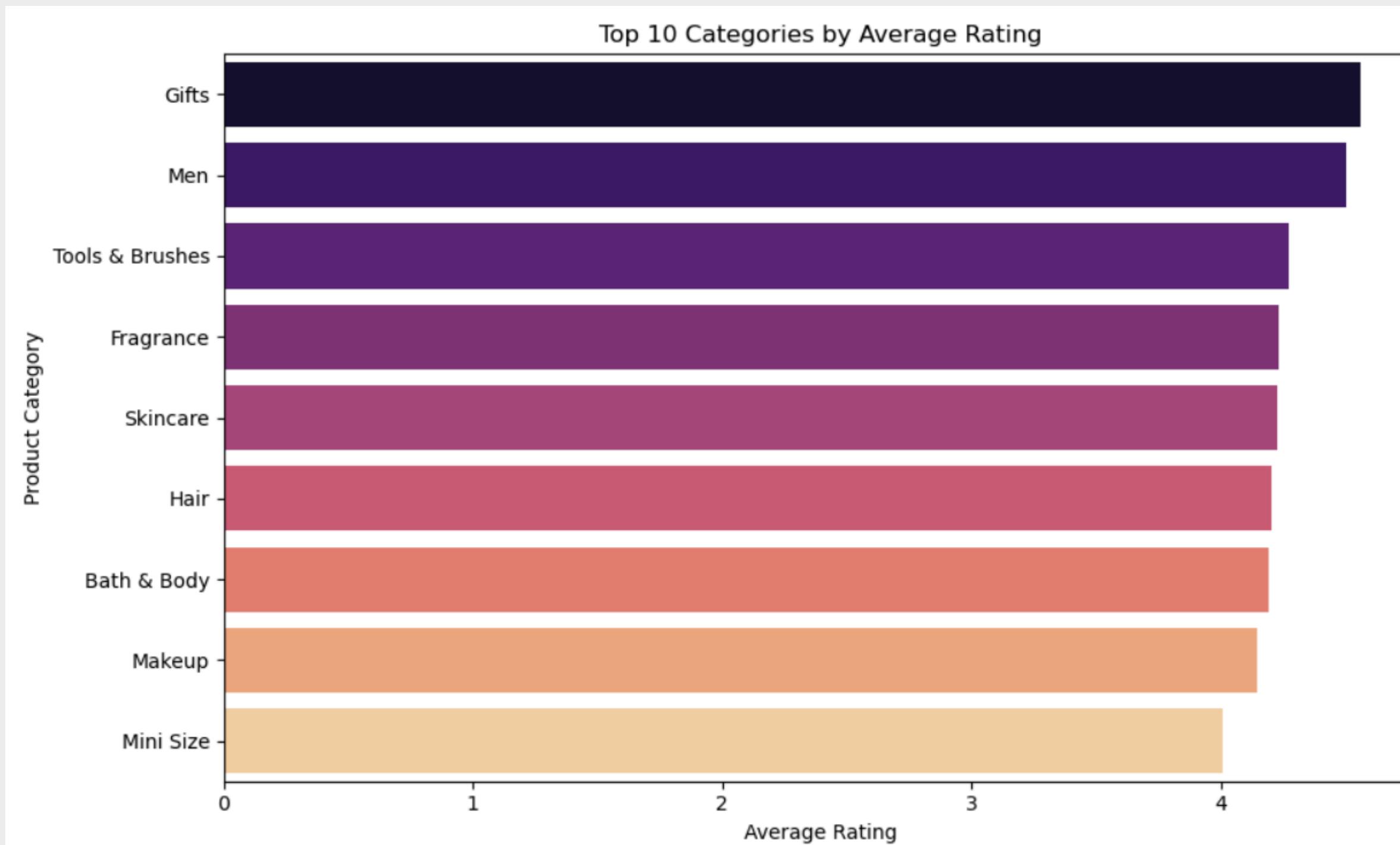
These insights can guide marketing strategies and inventory focus, helping beauty retailers better align with what customers truly value.

Q1: Which brands are the most popular?



The most loved brands are not necessarily the most expensive or diverse in their offerings. This shows strong brand loyalty from customers. Certain brands consistently receive high engagement, which signals trust and satisfaction.

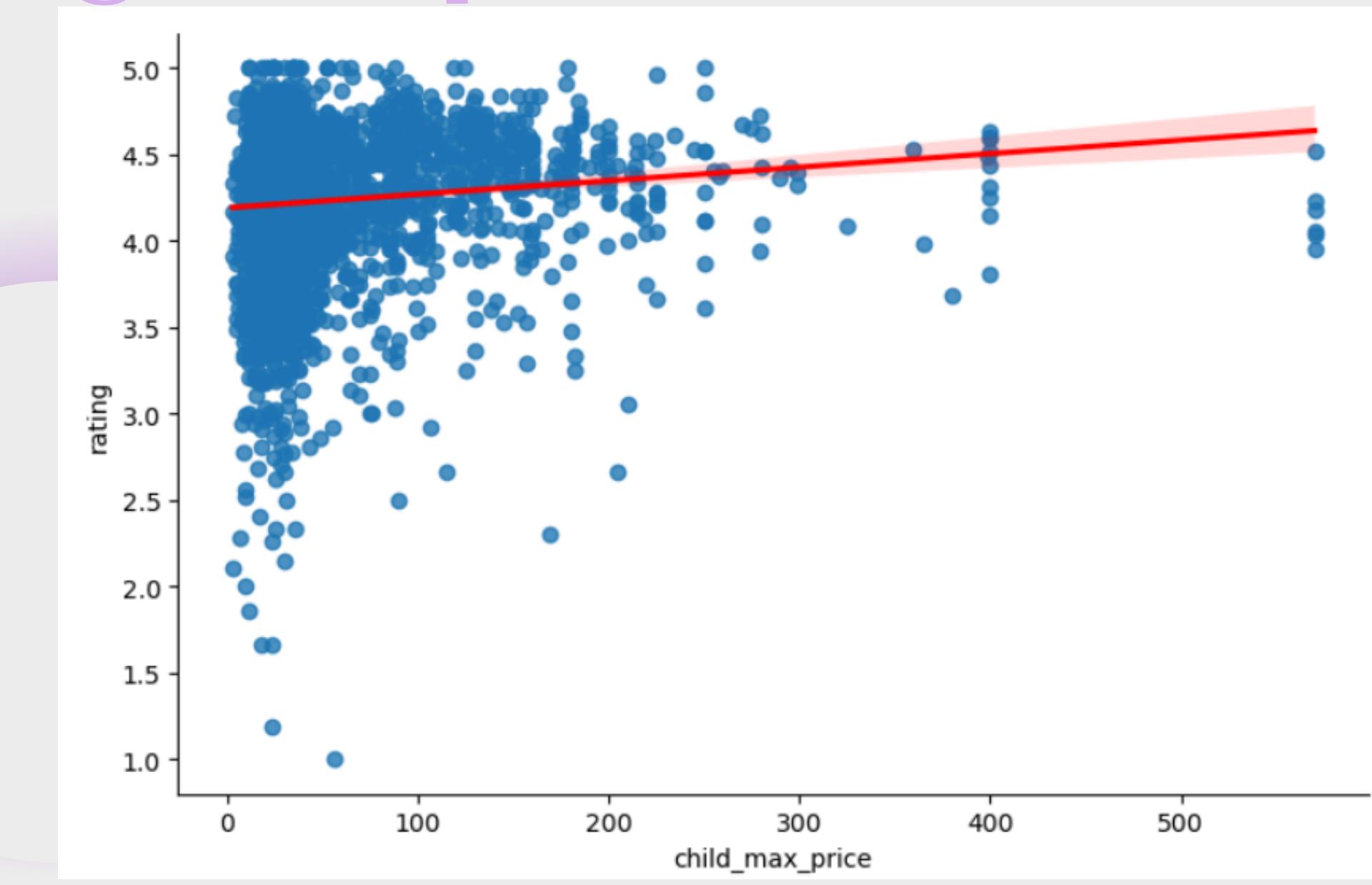
Q2: What is the average rating by product category?



Makeup and skincare products tend to receive higher average ratings compared to others like tools or fragrance. This suggests that users are more satisfied with core beauty products.

Categories like “Gifts” and “Men” receive the highest average ratings, while “Mini Size” and “Makeup” rank lower. This suggests that users might value practicality or curation in their beauty purchases.

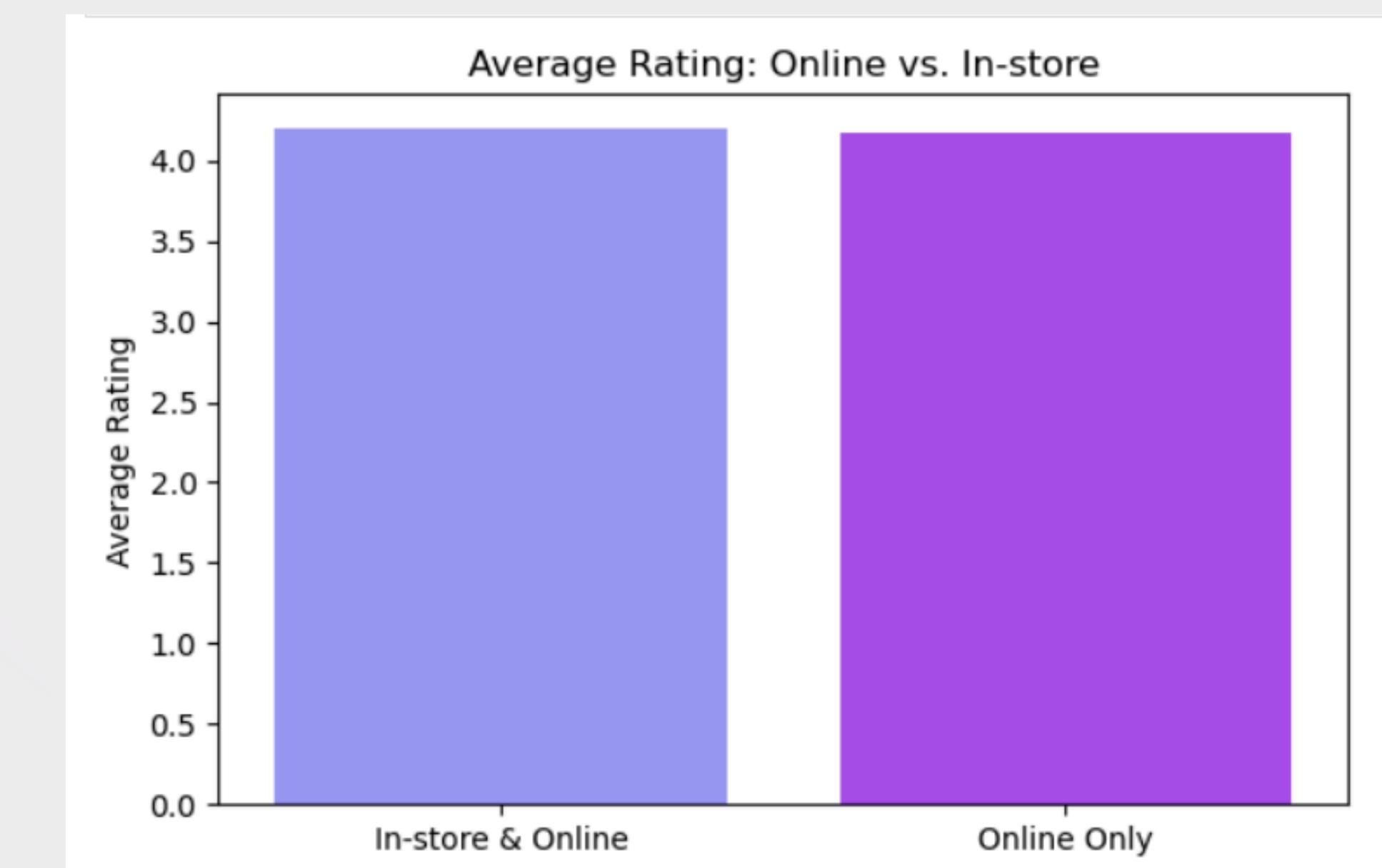
Q3: Is there a relationship between product rating and price?



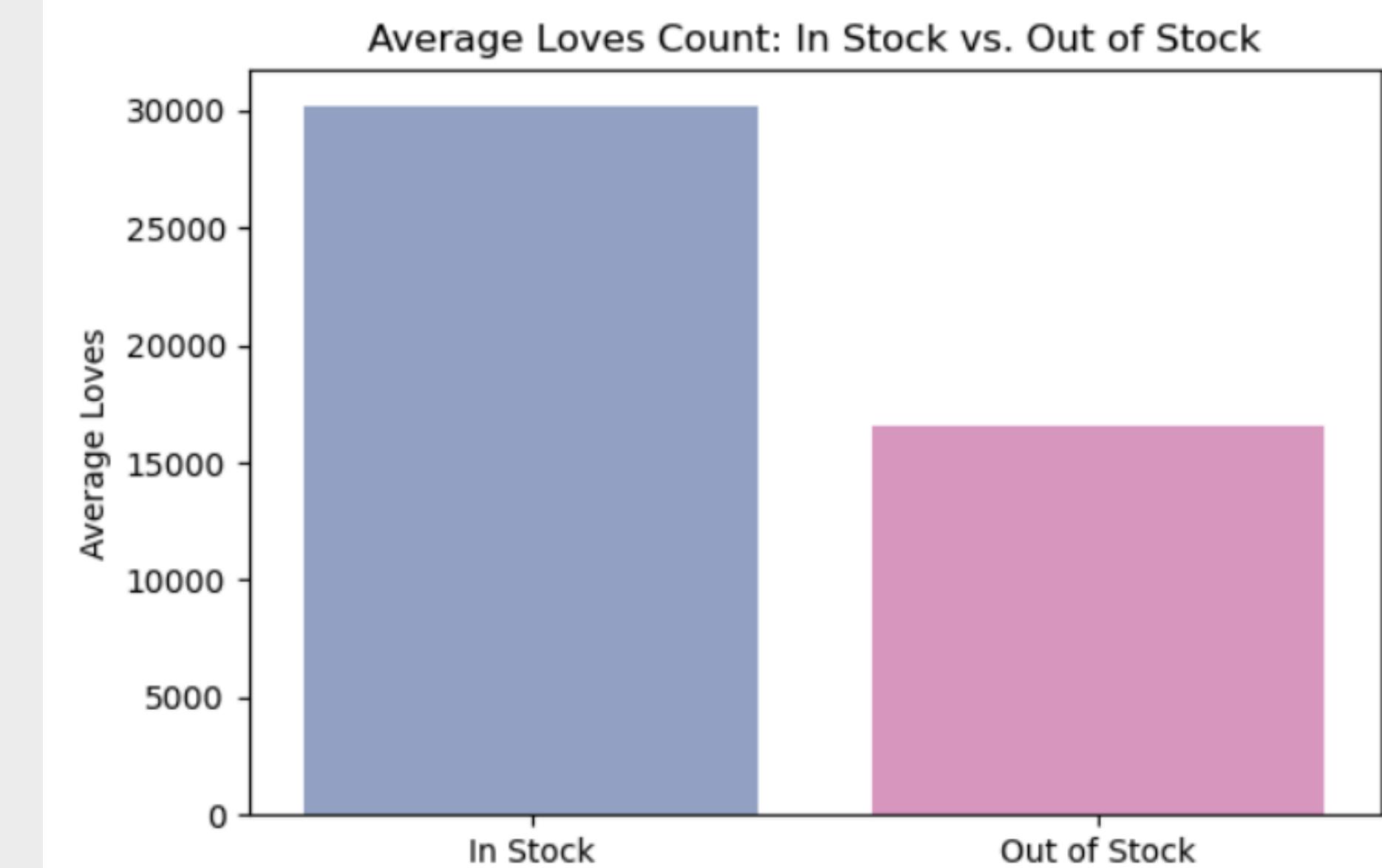
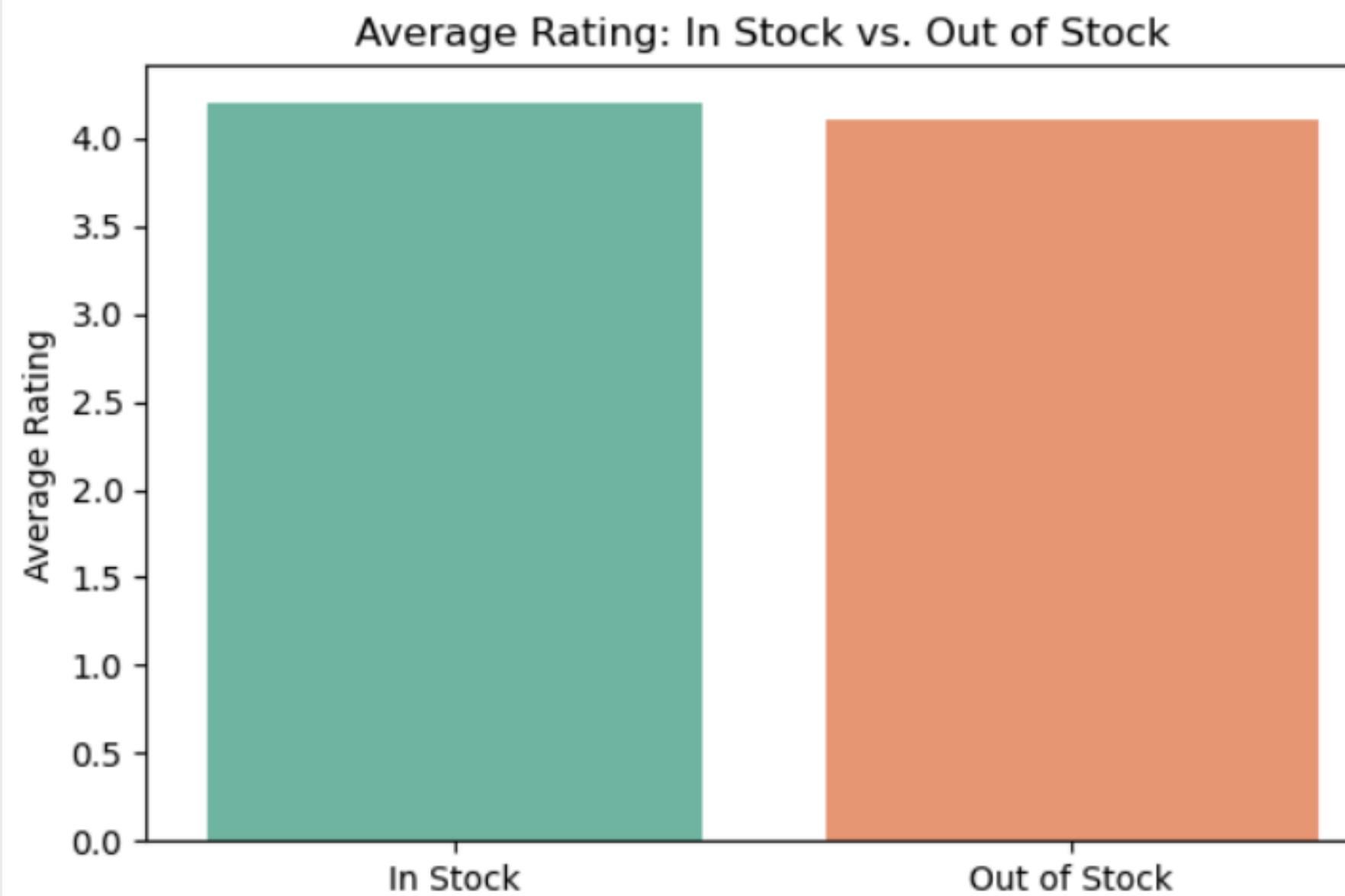
Products in the \$20–40 range often get high ratings and many likes. This means that customers care about both price and quality when choosing what to buy.

Q4: What is the average rating of online-only products?

- The average ratings of online-only products and those available both in-store and online are very close. This suggests that the shopping channel does not strongly affect how customers rate products.

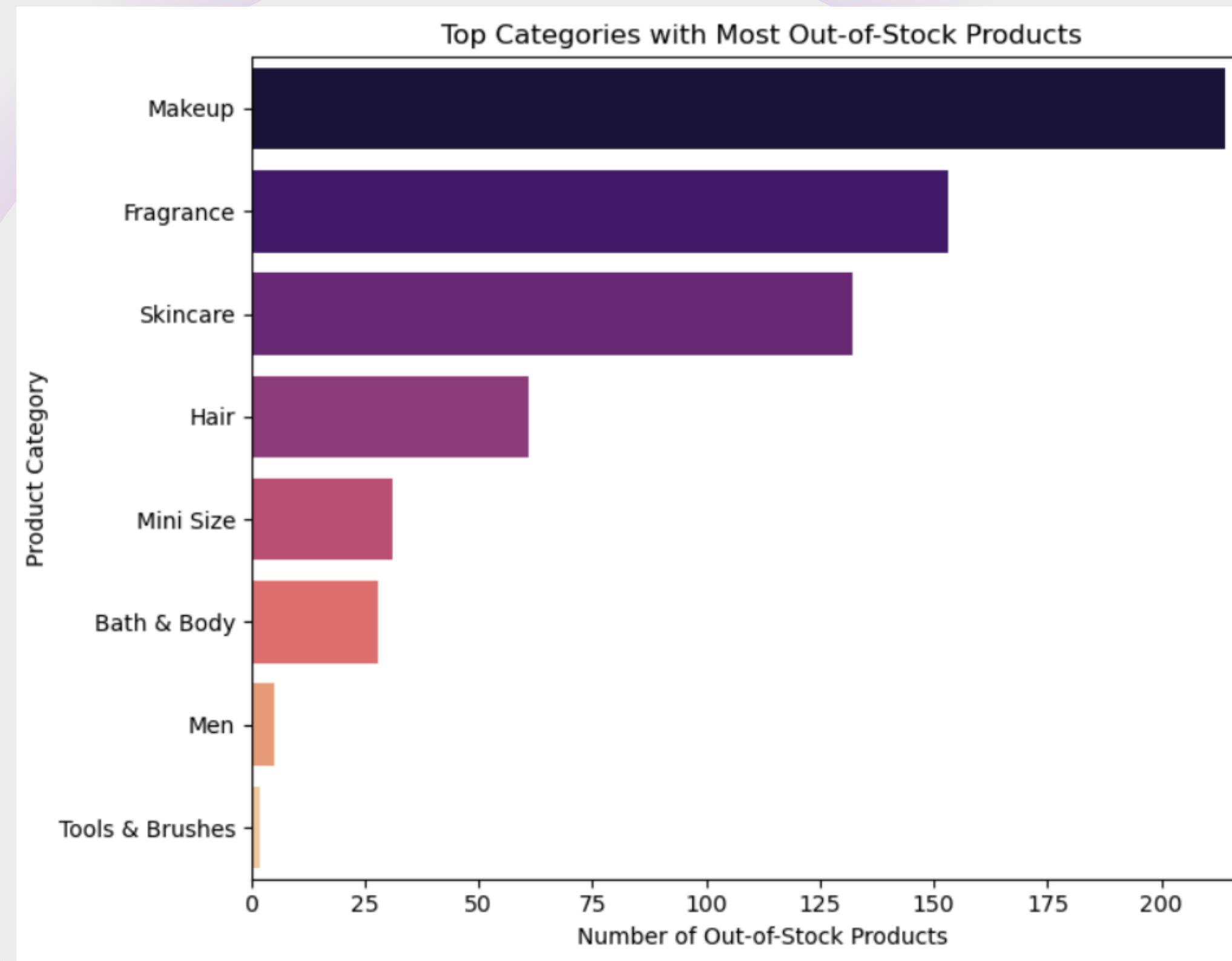


Q5: What are the common features of out-of-stock products?



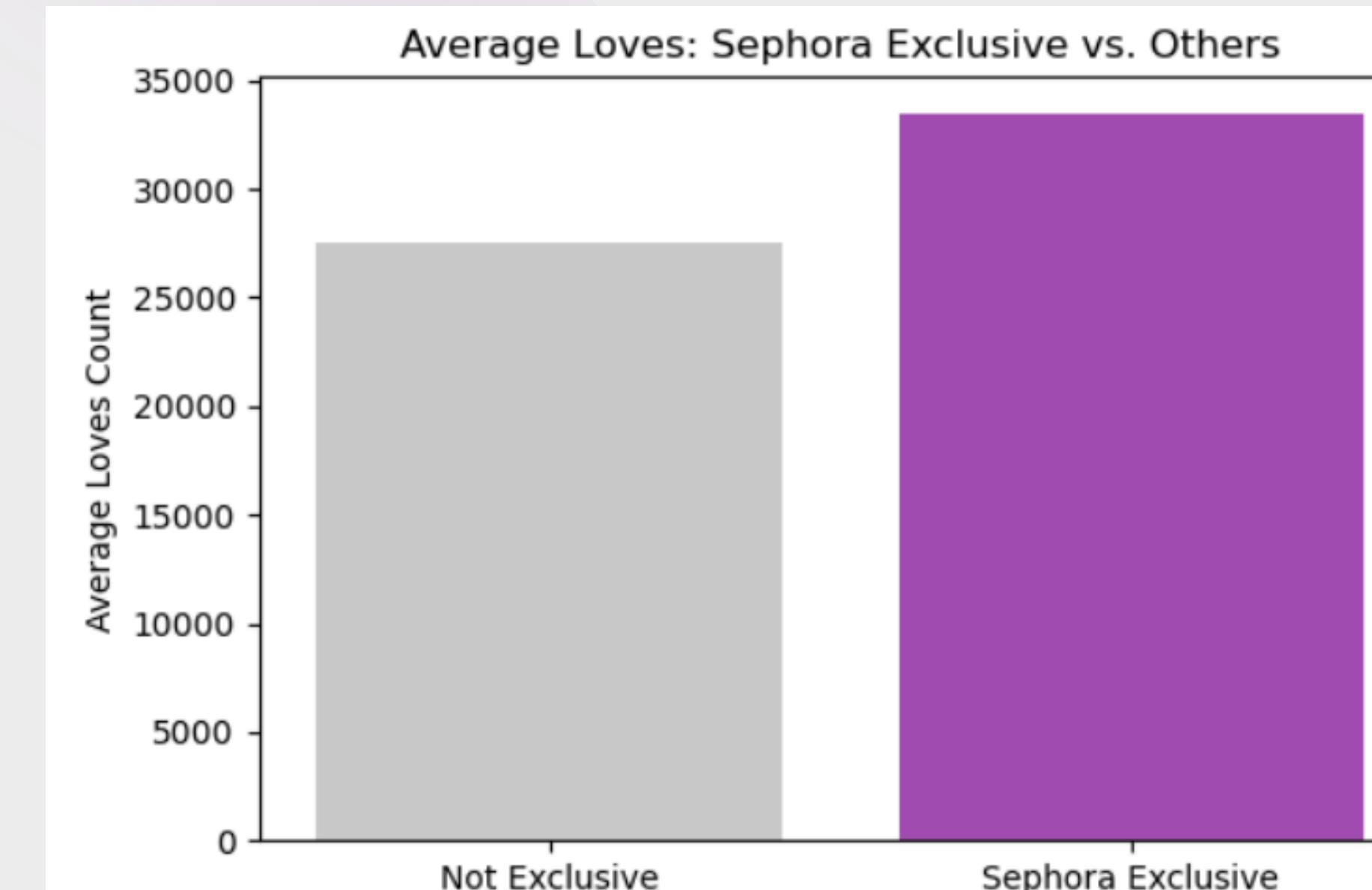
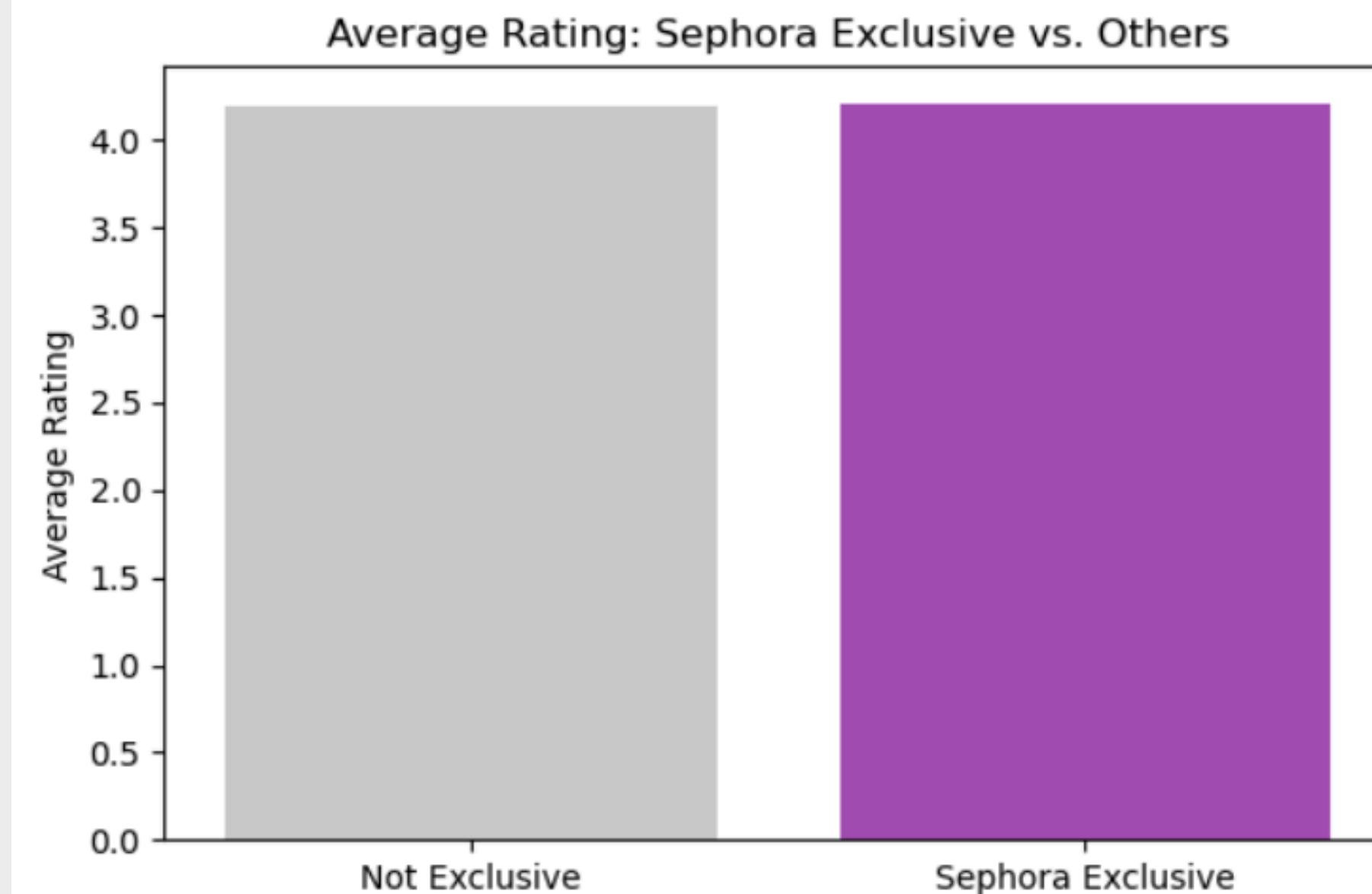
Products that are in stock tend to have slightly higher ratings and receive more likes. This could mean that more popular or better-rated products are restocked more quickly.

Bonus Q: Which categories are most often out of stock?



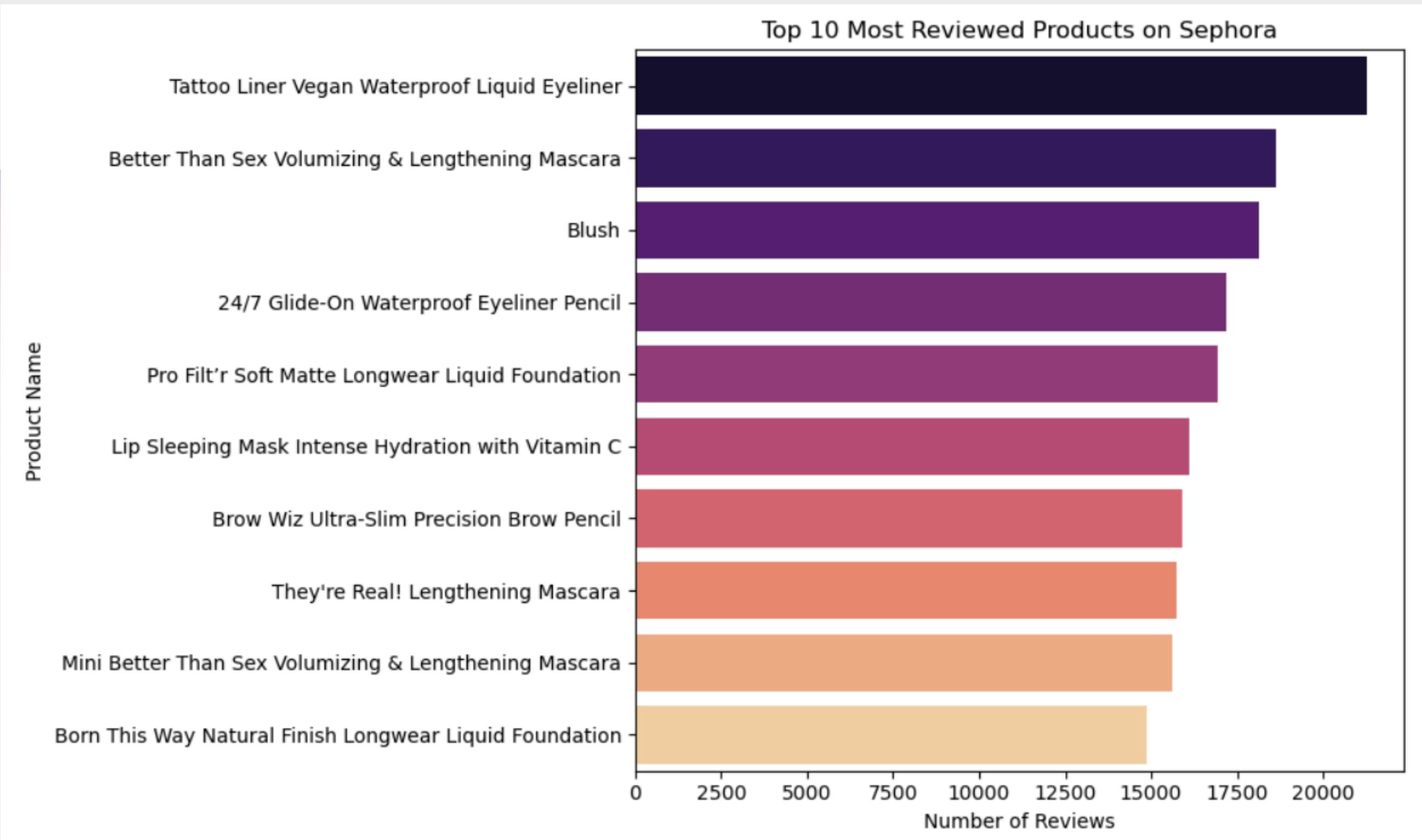
Makeup, fragrance, and skincare are the top categories with the highest number of out-of-stock products. This suggests exceptionally high consumer demand or insufficient supply in these segments. These categories may require improved inventory management or prioritization in restocking strategies.

Q6: Are Sephora-exclusive products rated differently?



Sephora-exclusive products receive slightly more likes and have similar ratings to non-exclusive ones. This suggests that exclusivity may help with visibility and popularity, but not necessarily with perceived quality.

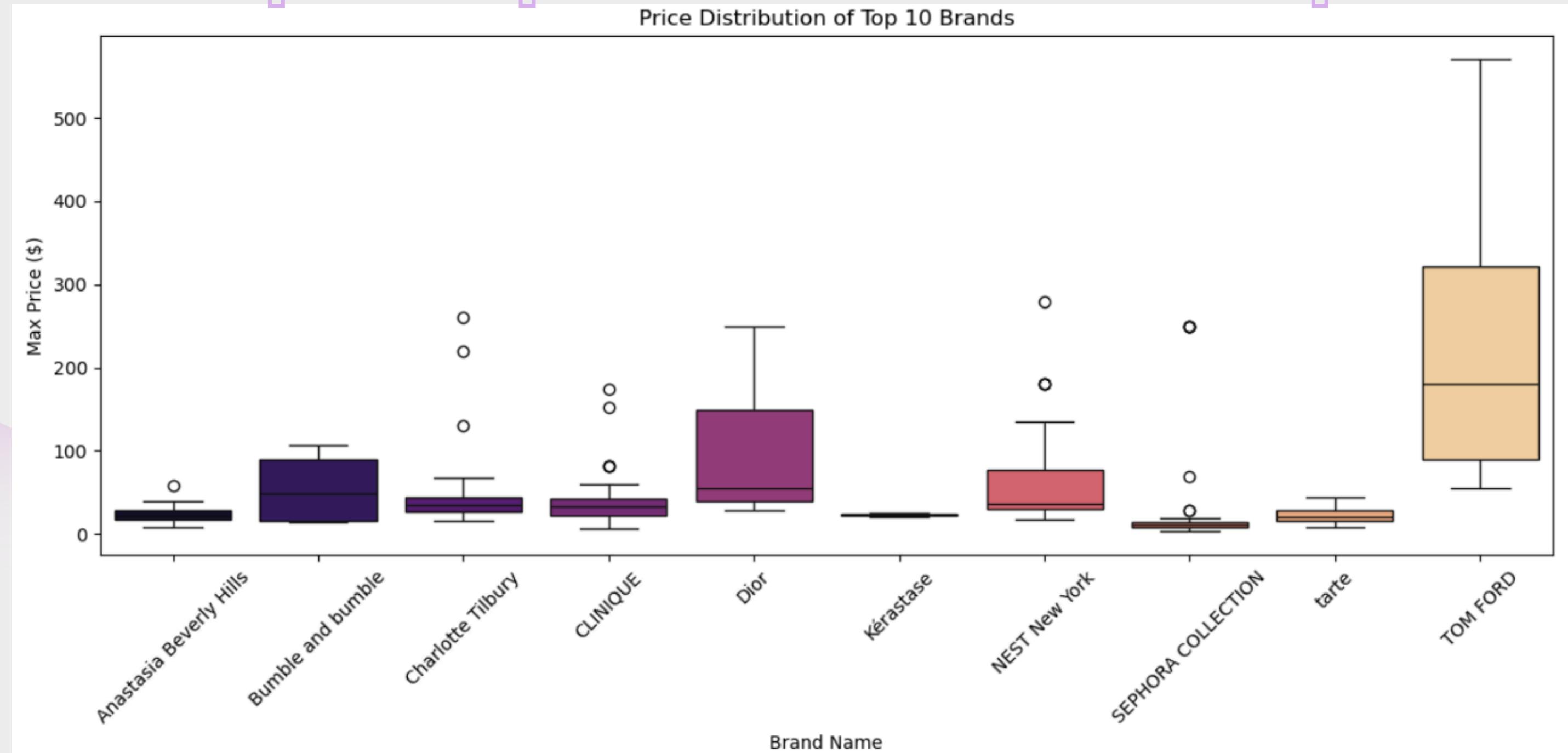
Q7: Which products have the most reviews?



The most reviewed products tend to be iconic items like eyeliners and mascaras. Their popularity suggests a strong and consistent user base willing to share feedback, making them standout performers in the catalog.

These products may benefit from high visibility, strong branding, or long-standing reputation.

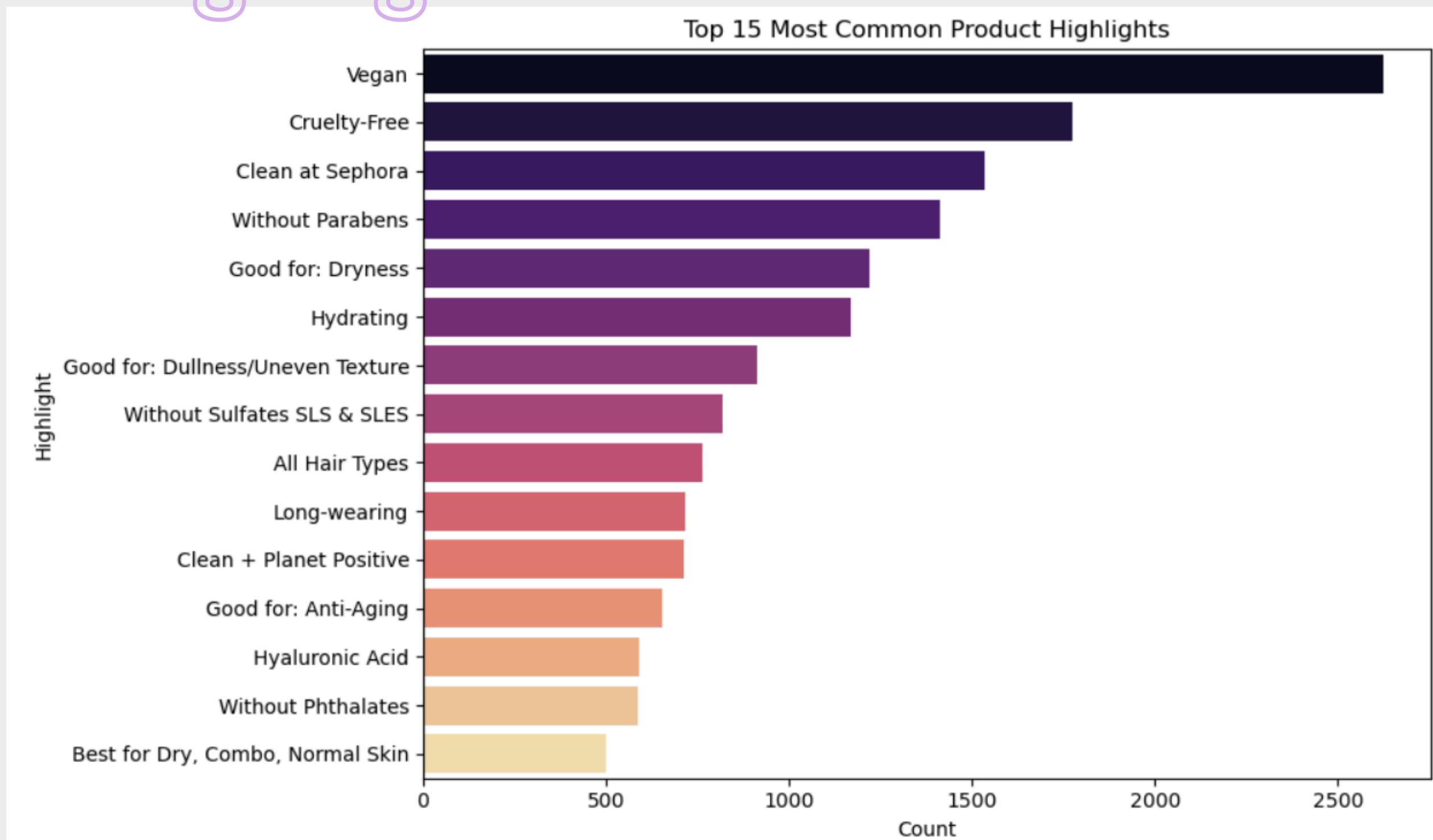
Q8: Compare prices across top brands



Luxury brands like Tom Ford and Dior have significantly higher price ranges, while brands like Sephora Collection and Kérastase offer more affordable options. This variation suggests diverse pricing strategies catering to different consumer segments.

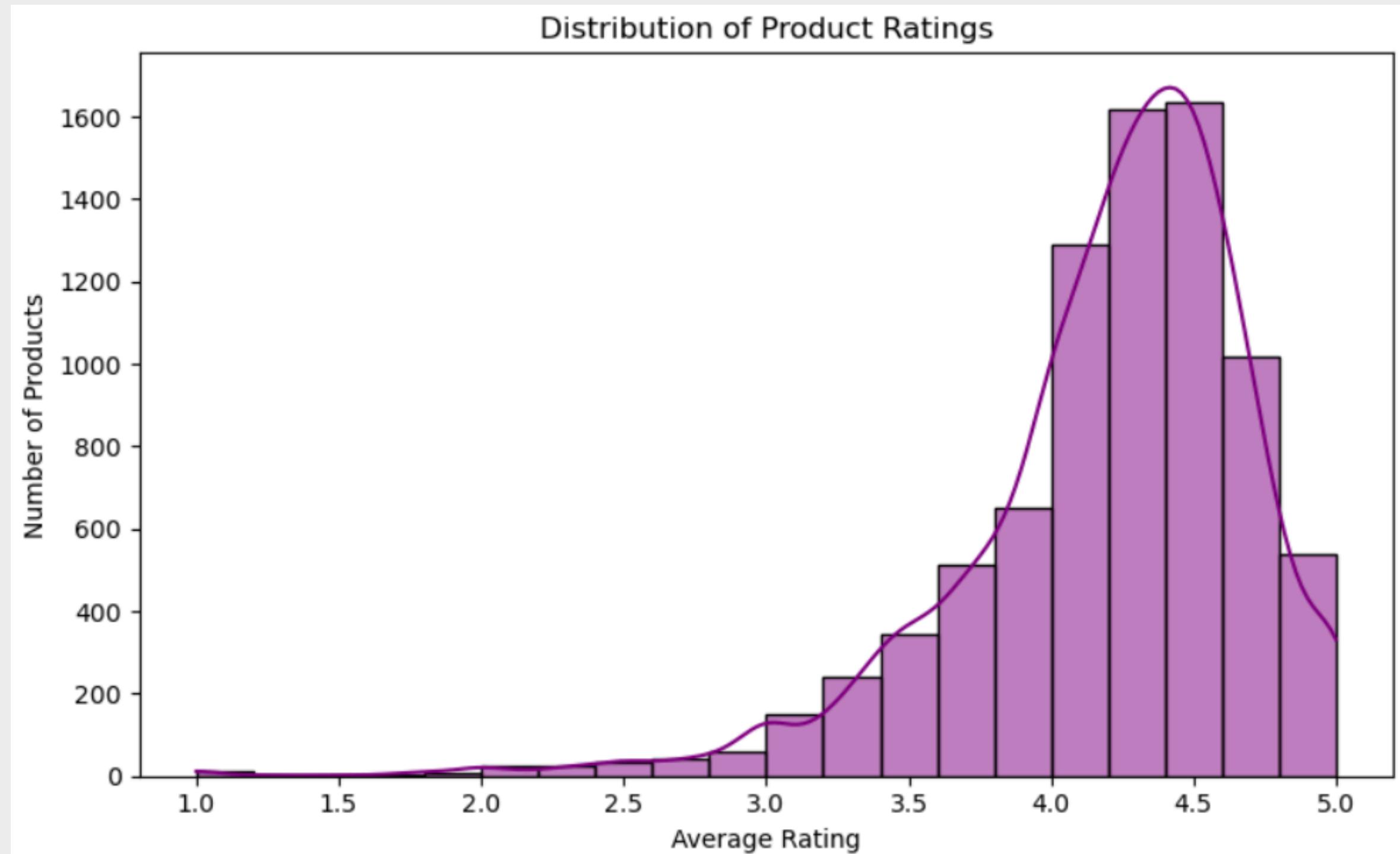
Q9: What are the most common product highlights?

Ethical and clean beauty attributes dominate product highlights, with terms like "Vegan," "Cruelty-Free," and "Clean at Sephora" appearing most frequently. This reflects a strong market shift toward sustainability, transparency, and health-conscious consumer preferences.



Q10: What is the overall rating distribution across all products?

This chart shows that most products receive ratings between 4.0 and 4.7, suggesting high overall customer satisfaction. The distribution is right-skewed, indicating that low-rated products are relatively rare. This may reflect customer filtering behavior or the removal of poorly performing items from the platform.



16.07.2025 15:20

streamlit_app

```

{
  "cells": [
    {
      "cell_type": "code",
      "execution_count": 5,
      "id": "86725496-29cf-4ee8-8ad2-bfd6f53026a7",
      "metadata": {},
      "outputs": [
        {
          "data": {
            "text/plain": [
              "0:\"DeltaGenerator()\""
            ]
          }
        }
      ],
      "execution_count": 5,
      "metadata": {},
      "output_type": "execute_result"
    }
  ],
  "source": [
    "0:import streamlit as st\n",
    "\n",
    "1:import pandas as pd\n",
    "\n",
    "2:import seaborn as sns\n",
    "\n",
    "3:import matplotlib.pyplot as plt\n",
    "\n",
    "4 : \"\n",
    "\n",
    "5:# Load data\n",
    "6:df = pd.read_csv(\"product_info.csv\")\n",
    "7 : \"\n",
    "\n",
    "8:# Drop missing\n",
    "9:df = df.dropna(subset=[\"brand_name\", \"primary_category\", \"rating\", \"child_max_price\"])\n",
    "10:\"\n",
    "\n",
    "11:# Sidebar filters\n",
    "12:st.sidebar.title(\"Filters\")\n",
    "13:brands = st.sidebar.multiselect(\"Select Brands\", options=df[\"brand_name\"].unique(), default=df[\"brand\n",
    "14:categories = st.sidebar.multiselect(\"Select Categories\", options=df[\"primary_category\"].unique(), def\n",
    "15:price_range = st.sidebar.slider(\"Price Range\", float(df[\"child_max_price\"].min()), float(df[\"child_ma\n",
    "16:\""
    "\n",
    "17:# Filter data\n",
    "18:filtered_df = df["
  ]
}

```

localhost:8503

16.07.2025 15:20

streamlit_app

```

19:"      (df[\"brand_name\"].isin(brands)) &
  "
20:"      (df[\"primary_category\"].isin(categories)) &
  "
21:"      (df[\"child_max_price\"].between(price_range[0], price_range[1]))\n",
  "
22:"]\n  "
23:"\n  "
24:# Metrics
  "
25:st.title(\"Sephora Product Explorer\")\n  "
26:st.metric(\"Average Rating\", round(filtered_df[\"rating\"].mean(), 2))\n  "
27:st.metric(\"Average Price\", f\"${round(filtered_df[\"child_max_price\"].mean(), 2)}\")\n  "
28:\n  "
29:# Scatter plot
  "
30:st.subheader(\"Rating vs. Price\")\n  "
31:fig1, ax1 = plt.subplots()\n  "
32:sns.scatterplot(data=filtered_df, x=\"child_max_price\", y=\"rating\", ax=ax1, color=\"purple\")\n  "
33:st.pyplot(fig1)\n  "
34:\n  "
35:# Rating histogram
  "
36:st.subheader(\"Rating Distribution\")\n  "
37:fig2, ax2 = plt.subplots()\n  "
38:sns.histplot(filtered_df[\"rating\"], bins=20, kde=True, ax=ax2, color=\"orange\")\n  "
39:st.pyplot(fig2)\n  "
40:\n  "
41:# Show filtered data
  "
42:st.subheader(\"Filtered Products\")\n  "
43:st.dataframe(filtered_df[\"product_name\", \"brand_name\", \"rating\", \"child_max_price\"]).reset_index(dro
  "
44:\n  "
]
  "
1 : {
  "cell_type": "code",
  "id": "5d2eb2ad-a19b-4ef2-b448-c1f55c315a68",
  "outputs": [],
  "source": []
}

```

localhost:8503

16.07.2025 15:20

streamlit_app

```

"metadata": {
  "kernelspec": {
    "display_name": "Python [conda env:base]",
    "language": "python",
    "name": "conda-base-py"
  }
},
"language_info": {
  "codemirror_mode": {
    "name": "ipython",
    "version": 3
  },
  "file_extension": ".py",
  "mimetype": "text/x-python",
  "name": "python",
  "nbconvert_exporter": "python",
  "pygments_lexer": "ipython3",
  "version": "3.12.7"
},
"nbformat": 4,
"nbformat_minor": 5
}

```

localhost:8503

1/3 localhost:8503

Key Values

- **Mid-range pricing wins:** Products priced between \$20–\$40 tend to receive higher ratings, showing that affordability matters more than luxury pricing.
- **Brand loyalty is real:** A few standout brands consistently earn more likes and higher ratings, indicating strong customer trust.
- **Category dominance:** Makeup and skincare products lead both in quantity and average rating, highlighting their central role in customer demand.
- **Value over price:** Expensive doesn't mean better users favor products that balance quality and value.
- **Strategic insight:** These trends can inform marketing, pricing, and inventory decisions for beauty retailers aiming to align with real customer preferences.

Sites

01

<https://www.kaggle.com/code/melissamonfared/skincare-products-eda-sentiment-analysis/input?scriptVersionId=191669866>

02

<https://github.com/RumeysaMercimek/sephora-dashboard>