

## Analysis Data File:

Dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380.csv dataset

### Unit of Observation

Each row in this dataset represents a specific Sephora brand's release of a lip product and its color, meaning that each row corresponds to a specific beauty brand collected from Apify scraping.

---

### Variable Definitions and Summaries

#### 1. Brand

- **Definition:** The name of the product brand company from Sephora that made a specific lip product.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summary Statistics:**

"Summer Fridays"	"Tower 28 Beauty"	"LANEIGE"
"DIOR"	"Glossier"	"Gisou"
"Fenty Beauty by Rihanna"	"SEPHORA COLLECTION"	"Charlotte Tilbury"
"OLEHENRIKSEN"	"MILK MAKEUP"	"REFY"
"Rare Beauty by Selena Gomez"	"Saie"	"MAKE UP FOR EVER"
"EADAM"	"Clarins"	"Benefit Cosmetics"
"CLINIQUE"	"Yves Saint Laurent"	"MAKEUP BY MARIO"
"tarte"	"Topicals"	"Prada Beauty"
"Kulfi"	"Valentino"	"MERIT"
"ILIA"	"LYS Beauty"	"Too Faced"
○ "HUDA BEAUTY"	"HAUS LABS BY LADY GAGA"	"Hourglass"
"Lancôme"	"Anastasia Beverly Hills"	"NARS"
"LAWLESS"	"Kaja"	"VIOLETTE_FR"
"Tatcha"	"BASMA"	"Kosas"
"Sol de Janeiro"	"Sephora Favorites"	"Paula's Choice"
"Gucci"	"Givenchy"	"Drunk Elephant"
"Urban Decay"	"Armani Beauty"	"ONE/SIZE by Patrick Starr"
"Mario Badescu"	"iNNBEAUTY PROJECT"	"Bobbi Brown"
"Hello Sunday"	"Jack Black"	"Freck Beauty"
"PATRICK TA"	"GXVE BY GWEN STEFANI"	"Dr. Jart+"
"PAT MCGRATH LABS"	"Sarah Creal"	"Iris&Romeo"
○ "TOM FORD"	"Westman Atelier"	"CAY SKIN"

"ISAMAYA"	"Supergoop!"	"Ami Colé"
"Iconic London"	"La Mer"	"Danessa Myricks Beauty"
"Laura Mercier"	"NUDESTIX"	"Grande Cosmetics"
"Biossance"	"Dr. Idriss"	"ROSE INC"
"GUERLAIN"	"Mango People"	"Viséart"
"fresh"	"SIMIHAZE BEAUTY"	"The INKEY List"
"Farmacy"	"First Aid Beauty"	"Fara Homidi"
"ALPYN"	"Buxom"	"NATASHA DENONA"
"MACRENE actives"	"Herbivore"	"caliray"
"Saint Jane Beauty"	"Dr. Dennis Gross Skincare"	"Then I Met You"
○ "Fashion Fair"	"goop"	"Augustinus Bader"
○ "MARA"	"BeautyBio"	"Kiehl's Since 1851"

## 2. detail\_link

- **Definition:** The product address link. The link takes you to product's website.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summary Statistics:**
  - NA

## 3. id

- **Definition:** The specific product's identifier. Starts with P and followed by unique set of numbers.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summary Statistics:**

"P455936"	"P448854"	"P515527"	"P453814"	"P512013"	"P505233"	"P67988453"	"P443563"
"P511572"	"P281411"	"P458268"	"P503878"	"P509439"	"P420652"	"P509414"	"P483124"
"P417985"	"P433516"	"P505568"	"P507492"	"P514572"	"P430969"	"P511921"	"P515492"
"P1272"	"P186805"	"P500636"	"P122751"	"P509397"	"P500744"	"P506548"	"P480000"
"P481114"	"P514464"	"P515611"	"P501496"	"P477838"	"P467977"	"P515038"	"P506489"
"P514666"	"P403322"	"P515971"	"P377710"	"P515552"	"P509711"	"P509710"	"P473708"
"P515227"	"P512014"	"P504041"	"P512637"	"P481969"	"P515082"	"P482750"	"P449342"
"P516036"	"P512259"	"P433531"	"P511431"	"P510654"	"P510932"	"P513624"	"P515775"
"P465757"	"P511969"	"P515638"	"P515107"	"P480303"	"P509300"	"P471243"	"P505330"
"P510741"	"P509674"	"P467449"	"P483662"	"P512007"	"P509681"	"P483060"	"P13046537"
"P503692"	"P504609"	"P508041"	"P509660"	"P509474"	"P516006"	"P515465"	"P513177"

	"P506965"	"P505364"	"P514837"	"P509453"	"P453225"	"P512207"	"P511102"	"P515735"
	"P481403"	"P472977"	"P467760"	"P511209"	"P511113"	"P506541"	"P515726"	"P441883"
	"P514675"	"P511952"	"P516070"	"P515012"	"P500307"	"P500637"	"P515456"	"P415747"
	"P514610"	"P515923"	"P468175"	"P87987897"	"P504823"	"P511183"	"P516072"	"P467652"
	"P502222"	"P509396"	"P433530"	"P483654"	"P479855"	"P467795"	"P509305"	"P514614"
	"P504829"	"P516100"	"P515059"	"P514674"	"P453273"	"P512469"	"P505694"	"P39787544"
	"P510442"	"P480576"	"P515119"	"P512254"	"P403818"	"P506279"	"P409523"	"P510926"
	"P507545"	"P501181"	"P479501"	"P510283"	"P509390"	"P504605"	"P510516"	"P440503"
	"P508880"	"P457812"	"P501583"	"P471038"	"P422224"	"P504365"	"P511648"	"P506812"
	"P436153"	"P12573"	"P469819"	"P472978"	"P509899"	"P451920"	"P481729"	"P454116"
○	"P515680"	"P515938"	"P480326"	"P500061"	"P509658"	"P400701"	"P514562"	"P502461"
	"P513294"	"P482714"	"P513257"	"P516192"	"P504872"	"P514289"	"P513666"	"P510431"
	"P507174"	"P481367"	"P510377"	"P510999"	"P472700"	"P479843"	"P502850"	"P504802"
	"P511724"	"P515457"	"P468691"	"P500491"	"P516018"	"P504216"	"P506182"	"P468347"
	"P504779"	"P512173"	"P506917"	"P473850"	"P468398"	"P512684"	"P479337"	"P404241"
	"P510861"	"P505562"	"P503938"	"P411848"	"P452737"	"P469873"	"P500329"	"P463452"
	"P444964"	"P471221"	"P515459"	"P511548"	"P515917"	"P505704"	"P468353"	"P501504"
	"P515902"	"P441880"	"P509299"	"P404831"	"P468343"	"P509877"	"P511701"	"P514910"
	"P506268"	"P407471"	"P480575"	"P513592"	"P506239"	"P434009"	"P471239"	"P421485"
	"P515446"	"P446609"	"P514919"	"P515374"	"P482730"	"P468863"	"P468385"	"P443764"
	"P511215"	"P475178"	"P515093"	"P500284"	"P501008"	"P514911"	"P510448"	"P512272"
○	"P449372"	"P455322"	"P515455"	"P500283"	"P405231"	"P506509"	"P511733"	"P511241"
	"P475179"	"P461456"	"P507343"	"P516077"	"P474358"	"P477837"	"P455558"	"P512525"
	"P516166"	"P504210"	"P483470"	"P515575"	"P480308"	"P509656"	"P481098"	"P500777"
	"P503350"	"P483448"	"P515447"	"P457245"	"P507664"	"P507592"	"P513264"	"P443623"
	"P512468"	"P505757"	"P509060"	"P506281"	"P512645"	"P474948"	"P482002"	"P436270"
	"P500992"	"P515690"	"P500997"	"P504127"	"P512149"	"P452736"	"P219001"	"P513832"
	"P293008"	"P510391"	"P428662"	"P511961"	"P511650"	"P514912"	"P515047"	"P433864"
	"P483643"	"P512183"	"P513466"	"P305132"	"P512544"	"P510564"	"P421813"	"P378111"
	"P480450"	"P511748"	"P458749"	"P483702"	"P513674"	"P515996"	"P451938"	"P458748"
	"P400625"	"P515380"	"P174213"	"P510159"	"P510553"	"P511549"	"P514891"	"P506750"
	"P429018"	"P480599"	"P504363"	"P511230"	"P422073"	"P502303"	"P503639"	"P505140"
○	"P481996"	"P411531"	"P512545"	"P471584"	"P468455"	"P446610"	"P244911"	"P509462"
	"P505699"	"P514805"	"P183916"	"P420163"	"P2866"	"P502223"	"P514240"	"P516580"
	"P395733"	"P507457"	"P442003"	"P483435"	"P509778"	"P471224"	"P516548"	"P477805"
	"P506901"	"P461218"	"P400708"	"P432239"	"P502353"	"P74520"	"P503304"	"P474832"
○	"P467765"	"P457649"						

#### 4. image

- **Definition:** The image address link that shows the product image.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summary Statistics:**
  - NA

#### 5. name

- **Definition:** The lip product name.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify

- **Missing Observations:** N/A
- **Summary Statistics:**
  - NA

## 6. price

- **Definition:** The lip product price.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summary Statistics:**

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
○	3.00	22.50	27.00	29.64	35.00	95.00

## 7.quantity\_rating

- **Definition:** The amount of ratings for the specific lip product.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summarized Data:**

	min_rating	max_rating	mean_rating	median_rating	sd_rating	count
○	0	20953	758.0502	309.5	1651.864	458

## 8. rating

- **Definition:** The average rating for a specific lip product.
- **Source Data:**
  - Website: Sephora.com
    - Dataset: dataset\_sephora-com-scraper\_2025-04-03\_04-22-12-380
- **Processing Steps:** Scrapped by Apify
- **Missing Observations:** N/A
- **Summarized Data:**

	min_rating	max_rating	mean_rating	median_rating	sd_rating	count
○	1	0	5	4.169607	4.34	0.7248356 458

