Case Study 7

Lip Products in India and Indonesia

A Comparative Study

Arka Mukhopadhyay	B23120
Pranab Ray	B23169
Kamal Yadav	B23209
Arani Ghosh	B23119
Ayuj Aryan	B23198
Kunal Sharma	B23079

April 10, 2024

Contents

List of Tables

List of Figures

1 Introduction

1.1 Background

Lips are often the unsung heroes of our skincare routine. With thinner skin, fewer oil glands, and no natural protection from the elements, lips are prone to dryness, chapping, and cracking. To keep them looking and feeling their best, it's essential to incorporate lip care into your daily routine. Lip products are one of the most important parts of makeup. It is even considered as the most used beauty product in the world.

With a wide array of lip beauty products to choose from, finding the perfect one for you can be quite challenging. Dermatologists advise protecting your lips from the sun with lipsticks with at least SPF 15.

There are a lot of Lip products and Brands available all over the world. We would like to analyze the specification of the Lip products that are available in Indonesia's shade, brands, type and the prices that they are offering.

1.2 Objective

This research aims to determine several indicators of lip products in Indonesia based on population and sample gathered from secondary data. These incude:

1. nature of data provided for lip products.
2. mean, max and min price of a popular lip products.
3. popular brands and type of lip products.
4. mean shade diversity available for lip products.

1.3 Evolution of Lip Beauty

Lip products have a rich and fascinating history, tracing back thousands of years to ancient civilizations. The ancient Mesopotamian women used crushed gemstones to decorate their lips. The iconic Egyptian queet, Cleopatra later embraced vibrant red lips with crushed carmine beetles. Queen Elizabeth I of England popularized lip color in the 16th century, using a mixture of beeswax and mercuric sulphide. However, societal attitudes were not always favorable; in the 1770s, British Parliament associated lipstick wearers with witchcraft. Despite this, by the 20th century, lipsticks became symbols of empowerment. In 1930, Vogue encouraged artistic lip painting. From ancient ingenuity to modern self-expression, lip products continue to enchant and evolve, reflecting human creativity and beauty ideals through the ages.

1.4 Types of Lip Products

The following types of lip products were identified from the collected data:

1. **Lipsticks:**
   (a) **Bullet Lipsticks:** It provides convenience and ease of application, featuring a solid form crafted from a blend of waxes, pigments, and oils. It seamlessly delivers color, texture, and protection to the lips with a simple swipe.
   (b) **Liquid Lipsticks:** It is characterized by its liquid form and application with a small wand, delivering bold color and long-lasting wear. Ideal for those seeking intense pigmentation and extended durability.
2. **Lip Balm:** It is designed for medicinal and soothing purposes, offering hydration and long-lasting protection to the lips, keeping them moisturized and nourished.
3. **Lip Gloss:** It provides lips with a lustrous, glossy texture, coming in liquid or soft solid forms and offering a range of finishes, including clear, translucent, frosted, glittered, and metallic. It not only enhances the lips' appearance but also hydrates them for a soft, glossy finish.
4. **Lip Lacquer:** Combining the high shine of gloss with the long-lasting wear of liquid lipstick, lip lacquer is transfer-proof and can endure day and night without needing frequent reapplication. It delivers intense shine while maintaining color vibrancy.
5. **Lip Liner:** Also known as a lip pencil, lip liner is used to fill uneven areas on the outer edges of the lips after applying lipstick. It prevents bleeding and ensures a defined lip shape throughout the day.
6. **Lip Tint:** It provides a wash of color that can be easily washed away, offering lightweight, sheer coverage for a natural look.
7. **Lip Stain:** Available in liquid or gel form, lip stain imparts color to the lips that lasts longer than traditional lipsticks or tints by leaving a stain on the lips. However, it may cause dryness and is not recommended for use in winter months.
8. **Lip Polish:** It serves as an exfoliant, buffing away dead cells and flakes on the lips, leaving them soft, supple, and smooth. Its natural humectant properties help to maintain lip hydration.

2 Data and Statistics

2.1 What is Data?

Merriam Webster describes **Data** as factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation.

Before a problem is analyzed, all the information available must be converted into data. **Measurement** in the systemic process of assigning numbers to objects and their properties to facilitate the use of mathematics in studying and describing objects and their relationships.

2.2 Types of Data

2.2.1 Based on Source of Data

1. **Primary Data**: This type of data is collected firsthand by the researcher or investigator directly from the source. It involves gathering data through methods like surveys, interviews, observations, experiments, etc. Primary data is original and specific to the research or study at hand.

2. **Secondary Data**: Secondary data refers to data that has already been collected by someone else for a different purpose. This data is obtained from sources such as books, journals, government publications, websites, databases, etc. Secondary data analysis involves using existing data to derive insights or conclusions.

3. **Tertiary Data**: Tertiary data is derived from primary and secondary sources. It involves the aggregation, compilation, and analysis of primary and secondary data to create new datasets or information. Tertiary data is often used for market research, trend analysis, and decision-making processes.

There are several benefits of using secondary data:

- It is cost-effective being readily available and accessible at lower costs/for free.
- Using existing data eliminates the need for conducting new research, allowing researchers to analyze data immediately.
- It contributes to a better understanding of the problem.
- It serves as a foundation for comparing the data gathered by the researcher across different time periods and geographies; thereby facilitating trend analysis and benchmarking.

However, there are also disadvantages of using secondary data:

- Secondary data rarely fits within the framework of marketing research factors since researchers have limited control over methods of collection, processing and categorization of data.
- The quality, precision and reliability of secondary data is unknown.
- Data may be out of date.

2.2.2 Based on Levels of Measurement

Quantitative data consists of numerical or measurable values. It is typically collected through structured methods such as surveys, experiments, or measurements. Quantitative data can be analyzed statistically to identify patterns, trends, and relationships. [1]

Subcategories of Quantitative Data include:

1. **Discrete Data**: It comprises distinct, separate values that can be counted individually. Examples include the number of students in a class, the number of cars in a parking lot, etc.
2. **Continuous Data**: Continuous data represents measurements that can take any value within a range. It is typically obtained through instruments like scales, thermometers, or rulers. Examples include height, weight, temperature, and time.
3. **Interval Data**: Data that can be added or subtracted but not multiplied/divided. They do not have a true zero point. For example: temperature, year, etc.
4. **Ratio Data**: Data that can be added, subtracted, multiplied or divided. They have a true zero point. For example: height, weight, age and so on.

Data need not be inherently numeric to be useful in an analysis. For instance, male and female both are commonly used in almost any statistics report involving population but there is nothing numeric about these categories. This category of data is known as **Qualitative Data**.

Statisticians commonly distinguish two types of Qualitative Data:

1. **Nominal Data**: Categorical data without any inherent order or hierarchy. The categories are purely distinct labels or names. For example: types of fruits, colors, types of transportation, etc.
2. **Ordinal Data**: Categorical data with a natural order or hierarchy. While the categories have a meaningful sequence, the differences between them may not be uniform. Examples include education levels (e.g., high school, bachelor's degree, master's degree) or ratings.

2.3 What is Statistics?

Statistics is the science of data. This involves collecting, classifying, summarizing, organizing, analyzing, and interpreting data. It involves methods for designing experiments and surveys, gathering data, and drawing conclusions from that data. Statistics is widely used in various fields such as science, business, economics, engineering, social sciences, and many others. It helps in making informed decisions, predicting outcomes, testing hypotheses, and understanding patterns and trends in data.
There are two kinds of Statistics:

1. **Inferential Statistics**: Statistical inference is the science of characterizing or making decisions about a population by using information from a sample drawn from that population. This includes hypothesis testing, confidence intervals, and regression analysis.
2. **Descriptive Statistics**: Descriptive statistics uses data that provides a description of the population either through numerical calculated graphs or tables. It provides a graphical summary of data. It includes:
    - Measures of Central Tendency (Mean, Median and Mode)
    - Measures of Variability (Range, Variance, Dispersion, and so on)

3 Sampling

3.1 Key Terminologies

1. **Population**: The population refers to the entire group of individuals, objects, or events who represent a characteristic. A *census* study involves the entire population.
2. **Sample**: A sample is a subset of the population selected for observation or measurement.
3. **Sampling**: Sampling is the process of selecting a *sample* from the *population* to make statistical inferences and estimate population characteristics.
4. **Sampling Frame**: A sampling frame is a list of all the individuals, objects, or events in the population from which the sample will be selected.

3.2 Sampling Schemes

A good sample must reflect all the characteristics (of importance) of the population. A sample that accurately reflects its population characteristics is called a *representative sample*. A sample that is not representative of the population characteristics

is called a *biased sample*. The reliability or accuracy of conclusions drawn concerning a population depends on whether or not the sample is properly chosen so as to represent the population sufficiently well. [2]
The selection of a sampling method depends on factors such as the nature of the investigation, the availability of sampling frames (lists of population members), financial resources, desired accuracy level, and data collection method (e.g., questionnaires or interviews).
Common sampling techniques include:

1. **Simple Random Sample**: A sample selected in such a way that every element of the population has an equal chance of being chosen is called a simple random sample.

2. **Systematic Sampling**: A systematic sample is a sample in which every k^{th} element in the sampling frame is selected after a suitable random start for the first element with the population listed in some defined order.

3. **Stratified Sample**: Here, a sample obtained by stratifying (dividing into non-overlapping groups) the sampling frame based on some factor(s) and then selecting some elements from each of the strata. A population with N elements is first divided into 's' sub-populations, then a sample is drawn from each subpopulation independently.

4. **Cluster or Area Sampling**: In cluster sampling, the sampling unit contains naturally existing groups of elements called clusters instead of individual elements of the population. A cluster is an intact group naturally available in the field.

3.3 Bias in Sampling

Sampling bias refers to the systematic error introduced into a sample as a result of the sampling method. It occurs when some members of a population are systematically more likely to be selected in a sample than others, leading to inaccurate or misleading conclusions and limits the generalizability of the findings. [1] [3]
The following are some common types of sampling biases:

1. **Selection Bias**: Selection bias exists if some potential subjects are more likely than others to be selected for the study sample; usually due to the sampling process.

2. **Volunteer Bias**: Volunteer bias refers to the fact that people who volunteer to be in studies are usually not representative of the population as a whole. For this reason, results from entirely volunteer samples might be considerably different from those who do not volunteer.

3. **Non-Response Bias**: Non-response bias occurs when individuals selected for the sample do not respond to the survey or study. This can lead to under-representation of certain groups in the sample, skewing the results.

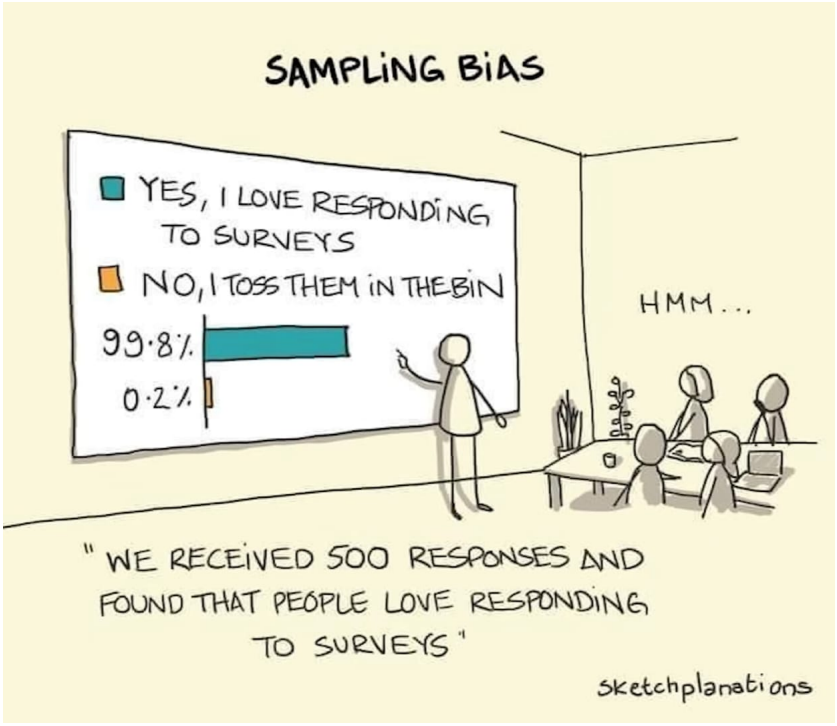

Figure 1: Example of Sampling Bias

4 Methodology

4.1 Data Collection - Indonesia

Considering the gigantic dataset pertaining to lip products in Indonesia, secondary data is used in our study in order to minimize time spent collecting the same otherwise.

The initial data for this study was obtained from the sample of 50 lip products given in a prior report [4]. The data given was sourced from 2022 reports, and hence, outdated. We used latest datasets to update 2 parameters - number of shades available in the market and the price of the lipsticks.

To find the latest and credible data, we have used the maximum retail price (MRP) of the product as mentioned on the company's official Indonesian Website. Note that there are some companies which don't have country specific websites (or don't have an operational official website at all). In such cases, we have used Lazada/Shoppee to collect the data. Why these websites? These two websites are leading e-commerce websites in southeast Asia and Taiwan and hence are most credible source that could have been used.

4.2 Data Collection - India

The Dataset for Indian Lip Brands was also obtained from secondary sources. Names of all popular Indian makeup brands were noted from various websites [5] [6] [7] and also from surveys among students. Then, an exhaustive list of products was prepared for each brand by visiting their official website, sorting all available products on the basis of popularity and noting the top products along with their MRP and number of shades available. In some cases where data was missing, we used Amazon and Nykaa to fill in the missing data.

4.3 Data Analytics

This study utilizes stratified sampling to determine the sample by dividing the population into subgroups based on factors like brands and product types. Also, the data that is used in this study is nominal in nature since it has categories without necessarily implying mathematical order. Libraries such as matplotlib, numpy, seaborn, and pandas provide easy means to analyze and visualize datasets.

The composition of lip products by type is examined, with details displayed on bar and donut charts. Brand distribution is analyzed using the 'Brand' column, showing the count and percentage of products per brand on a bar plot.

Statistical measures such as mean, median, mode, and quartiles are calculated for prices and the number of shades. Price distribution is visualized through a histogram and box plot, aiding in identifying price ranges and distribution. A bar graph displays the distribution of products based on the number of shades they offer.

Kernel Density Estimation (KDE) is used on histograms to create smooth curves representing the data distribution. This enhances visualization by revealing underlying patterns and providing a refined understanding of price and shade distribution across the dataset.

Color mapping is applied to visualize frequency levels, with lower frequencies represented in blue and higher frequencies in red. Lastly, data is grouped by companies to determine the total number of shades offered by each company, providing insights into shade diversity among different brands.

5 Analytics for Indonesia

5.1 Raw Data

Table 1: Raw Data for Study (2023)

Lip Product	Brand	Type of Lip Product	Shades	Price
NIVEA LIP BALM SOOTHE & PRTECT	Beiersdorf	Stick	2	50000
Extra lip tint	Bobbi Brown	Stick	10	711636
Perfect Matte Lip Coat	Dear Me Beauty	Liquid	6	129000
Creamytint	Emina	Liquid	5	46000
magic potion lip tint	Emina	Liquid	5	50000
Squeeze me up Lip Matte	Emina	Liquid	4	58000
Smoochies Lip balm	Emina	Solid	1	32000
Matte Lip Liquid	ESQA	Liquid	7	165000
Dear Darling Water gel tint	Etude House	Liquid	3	55000
Organic lip balm	Eucalie	Stick	1	79000
lip and cheek dual use liquid	Focallure	Liquid	10	38000
Melted Matte Lip	Goban Cosmetics	Liquid	6	130000
Sheen. Tinted lip balm + UV filter	HALE.	Stick	4	98000
Urban Lip Cream Matte	Implora	Liquid	20	25000
Beauty Lip & Cheeck Crayon	Indoganic	Crayon	2	129000
Vivid oil tint	Innisfree	Liquid	4	104000
Metallic Lip Cream	Inul Beauty	Liquid	5	89000
Infalible Pro Matte Lip Liquid	L'oreal	Liquid	9	150000
Rouge Signature Liquid Matte Lipstick	L'oreal	Liquid	14	151376
Color Riche Matte	L'oreal	Stick	8	354267
Intense Matte Lip Cream	Liquid	Liquid	12	119000
Longlasting Matte Lip Cream Metalic	LT Pro	Liquid	3	109900
Ultra Light Lip Stain	Luxcrime	Liquid	8	79000
Airy lip mousse	Luxcrime	Liquid	8	109000
Dew tinted 6hr lip moisturizer	Mad for Makeup	Stick	6	109000
magnifique lip tint	Madame Gie	Liquid	6	33000
Brilliant Glaze Lip Liquide	Madame Gie	Liquid	6	35000
Moist Velvet & Smooth Lip Liquide	Madame Gie	Liquid	6	15765
Hydrastay lip whip	Makeover	Liquid	12	119000
Powestay Transfer Proof Matte Lip Cream	Makeover	Liquid	12	135000
Sensational Liquid Matte	Maybelline	Liquid	19	66,023
color sensational lip tint	Maybelline	Liquid	19	45000
Super Stay Matte Ink	Maybelline	Liquid	19	239571
Color sensational the powder mattes	Maybelline	Stick	24	88900
Hydra Lip Cheek Tint	Mineral Botanica	Liquid	4	51900
the one A-Z lip balm SPF 25	Oriflame	Stick	2	149000
Lip Cream	PIXY	Liquid	16	55000
2 in 1 color tint	Purbasari	Liquid	3	51900
Lip Cream Series	Raiku	Liquid	13	118000
SUEDED! Lip & Cheek Cream	Rollover Reaction	Liquid	12	109000
Juicy Lip Balm	Rose All day	Stick	3	119000
Lip Color	Runa Beauty	Stick	5	138000
Lip Care	Sensatia Botanica	Liquid	5	80000
Coconut lip sleeping balm	Tiff Body	Liquid	1	88000
delight tony tint	Tony Moly	Liquid	3	49000
Exclusive Matte Lip Cream	Wardah	Liquid	24	66500
Colorfit Velvet Matte Lip Mousse	Wardah	Liquid	14	79000
Everyday Moisture Lip nutrition	Wardah	Stick	2	28500
Color Fit Ultralight Matte	Wardah	Stick	5	47500
The Simplicity Love You tint	Y.O.U	Liquid	4	45100

5.2 Sampled Data

5.2.1 On Basis of Brands

Table 2: Products grouped by Brand

Brand	Frequency	Percentage	Cumulative Percentage
Beiersdorf	1	2.0	2.0
Bobbi Brown	1	2.0	4.0
Dear Me Beauty	1	2.0	6.0
Emina	4	8.0	14.0
ESQA	1	2.0	16.0
Etude House	1	2.0	18.0
Eucalie	1	2.0	20.0
Focallure	1	2.0	22.0
Goban Cosmetics	1	2.0	24.0
HALE.	1	2.0	26.0
Implora	1	2.0	28.0
Indoganic	1	2.0	30.0
Innisfree	1	2.0	32.0
Inul Beauty	1	2.0	34.0
L'oreal	3	6.0	40.0
Liquid	1	2.0	42.0
LT Pro	1	2.0	44.0
Luxcrime	2	4.0	48.0
Mad for Makeup	1	2.0	50.0
Madame Gie	3	6.0	56.0
Makeover	2	4.0	60.0
Maybelline	4	8.0	68.0
Mineral Botanica	1	2.0	70.0
Oriflame	1	2.0	72.0
PIXY	1	2.0	74.0
Purbasari	1	2.0	76.0
Raiku	1	2.0	78.0
Rollover Reaction	1	2.0	80.0
Rose All day	1	2.0	82.0
Runa Beauty	1	2.0	84.0
Sensatia Botanica	1	2.0	86.0
Tiff Body	1	2.0	88.0
Tony Moly	1	2.0	90.0
Wardah	4	8.0	98.0
Y.O.U	1	2.0	100.0

5.2.2 On Basis of Type

Table 3: Products grouped by Type

Type	Frequency	Percentage	Cumulative Percentage
Stick	12	24.0	24.0
Liquid	36	72.0	96.0
Solid	1	2.0	98.0
Crayon	1	2.0	100.0

5.2.3 On Basis of Shades

Table 4: Products grouped by Shades

Shades	Frequency	Percentage	Cumulative Percentage
1	3	6.0	6.0
2	4	8.0	14.0
3	5	10.0	24.0
4	5	10.0	34.0
5	6	12.0	46.0
6	6	12.0	58.0
7	1	2.0	60.0
8	3	6.0	66.0
9	1	2.0	68.0
10	2	4.0	72.0
12	4	8.0	80.0
13	1	2.0	82.0
14	2	4.0	86.0
16	1	2.0	88.0
19	3	6.0	94.0
20	1	2.0	96.0
24	2	4.0	100.0

5.3 Graphs and Stats

Table 5: Descriptive Statistics of Shades and Prices

Attribute	**Shades**	**Price**
Count	50	50
Mean	8.04	104456.76
Median	6	84000
Mode	5, 6	79000, 109000, 119000
Std Dev	6.11	105480.23
Variance	37.3	11126079540.23
Minimum	1	15765
Maximum	24	711636
Percentiles:		
- 0th	1	15765
- 25th	4	50000
- 50th	6	84000
- 75th	12	119000
- 100th	24	711636

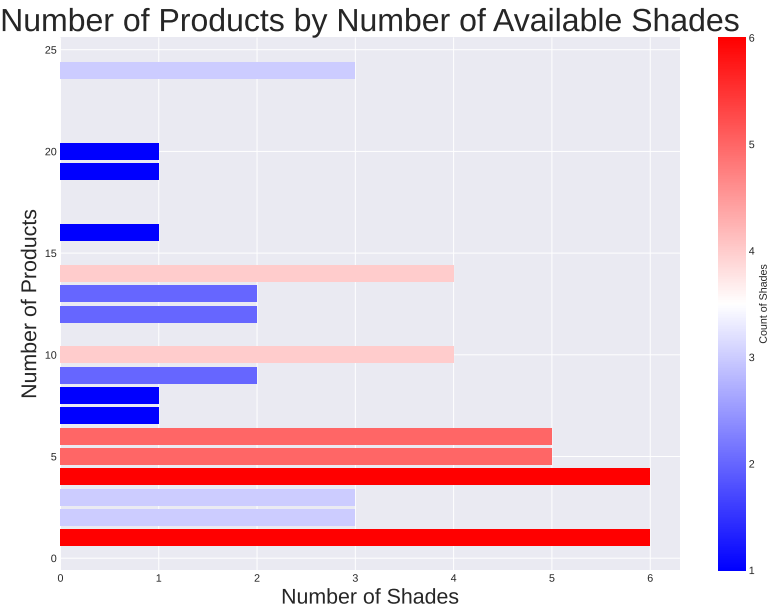

Figure 2: Total Products Grouped by Shades

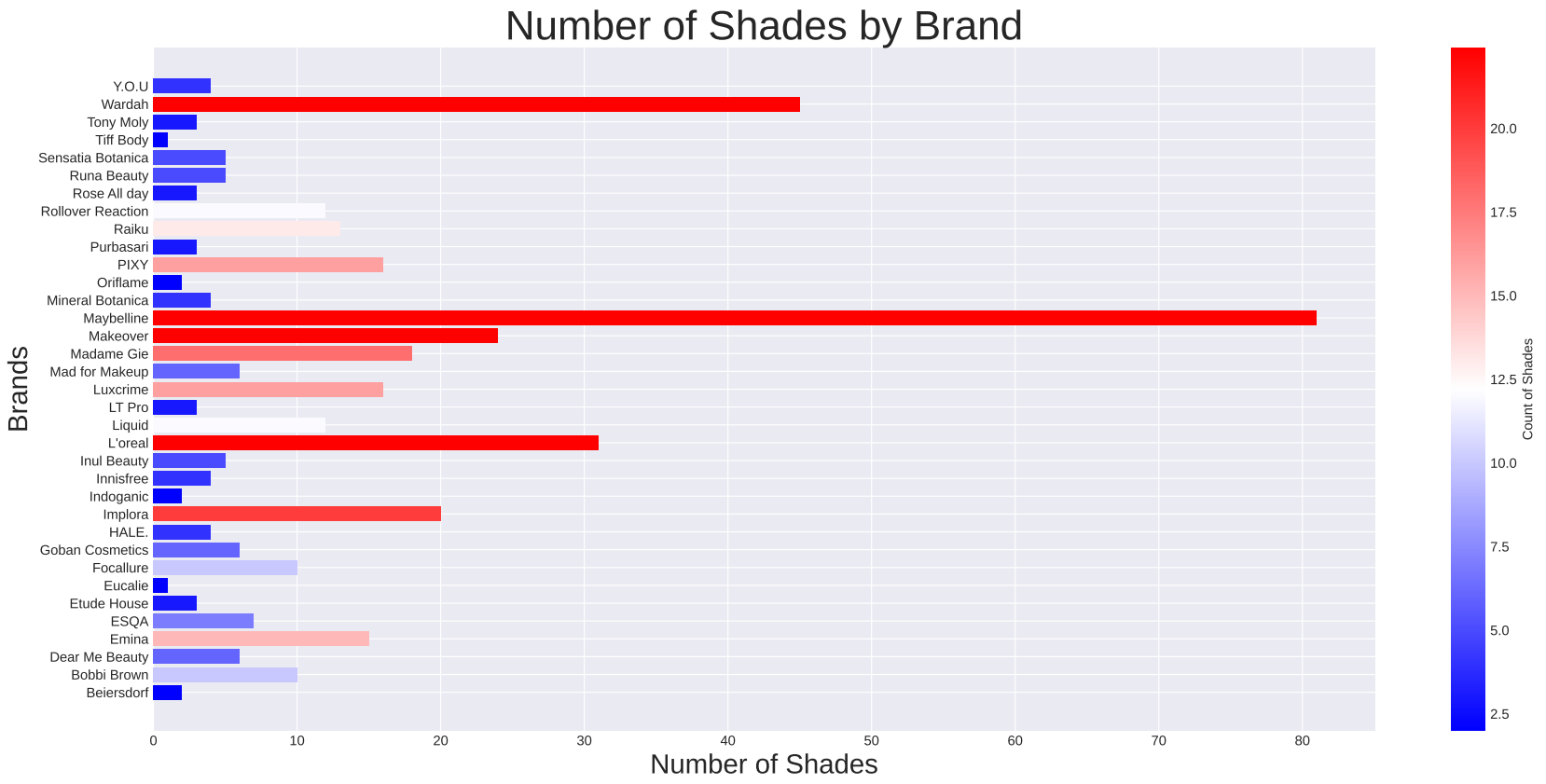

Figure 3: Total Shades Grouped by Brands

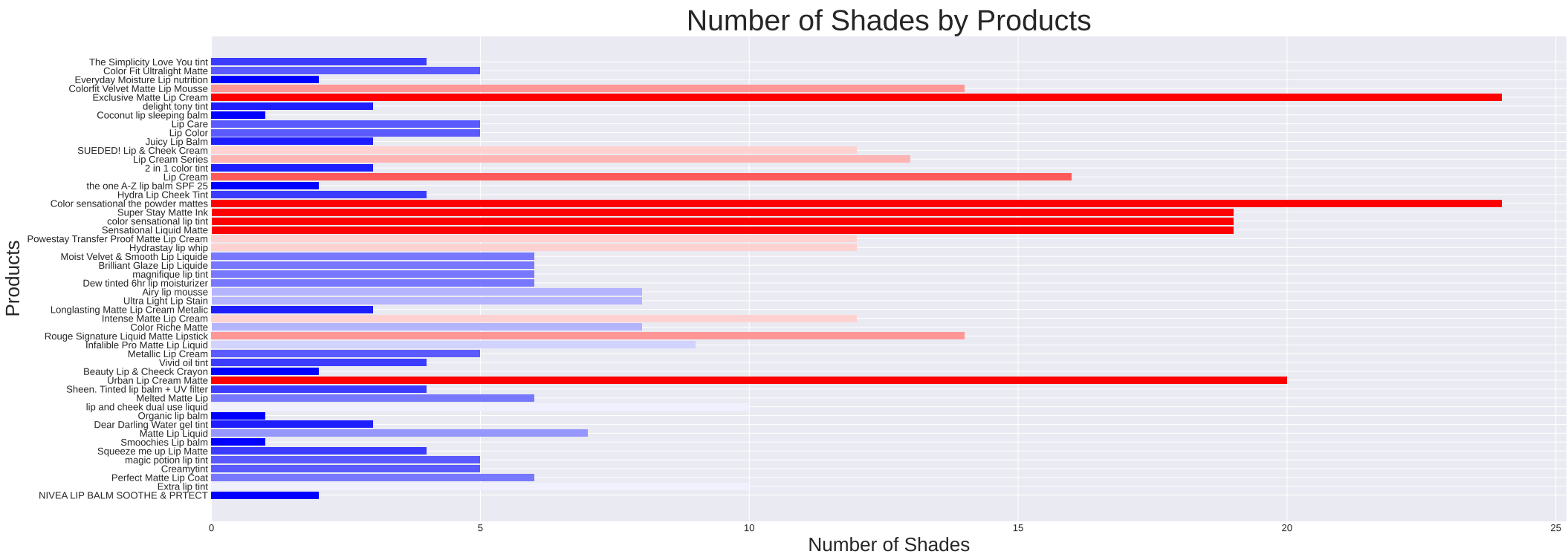

Figure 4: Total Shades Grouped by Product

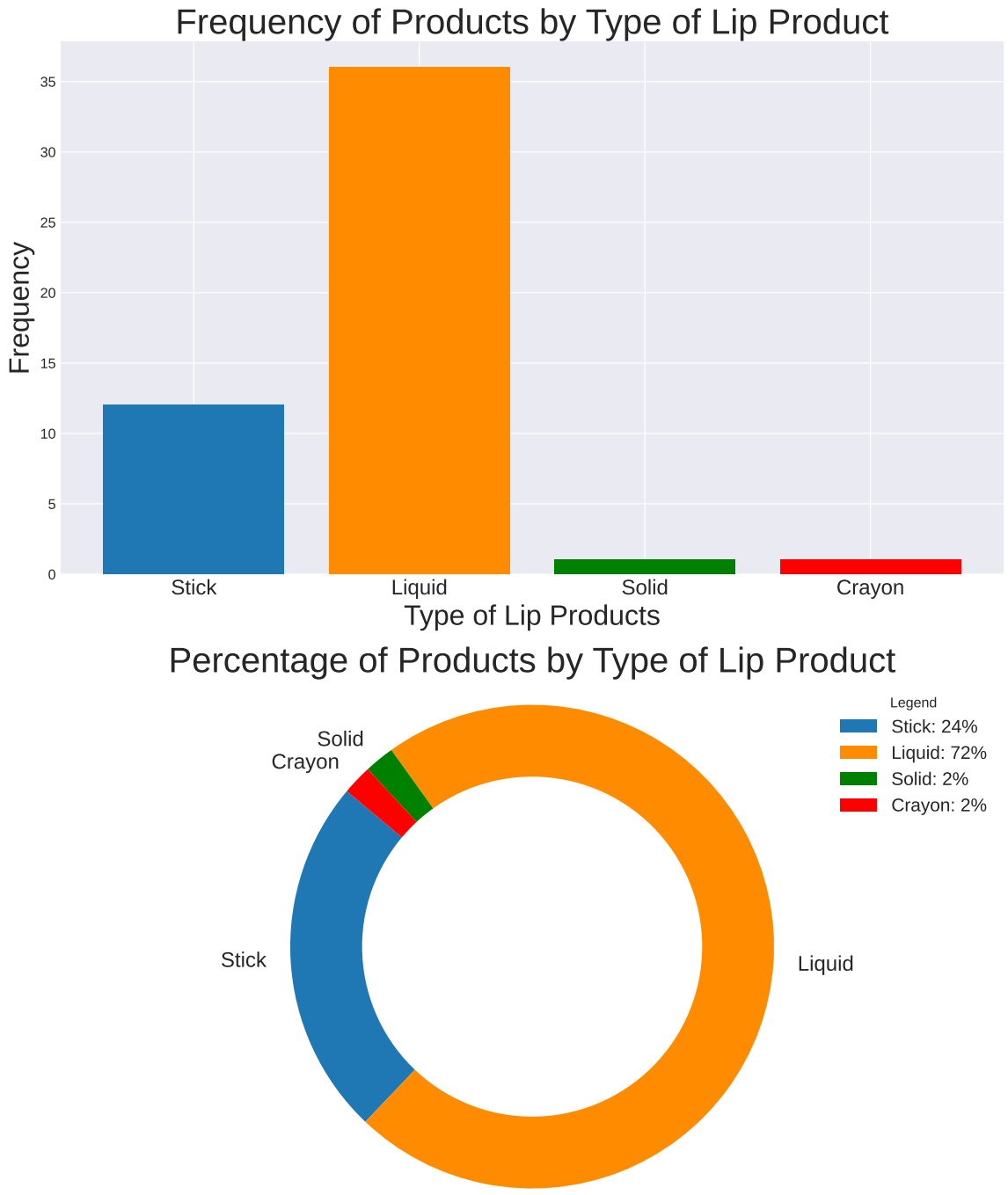

Figure 5: Total Products Grouped by Type

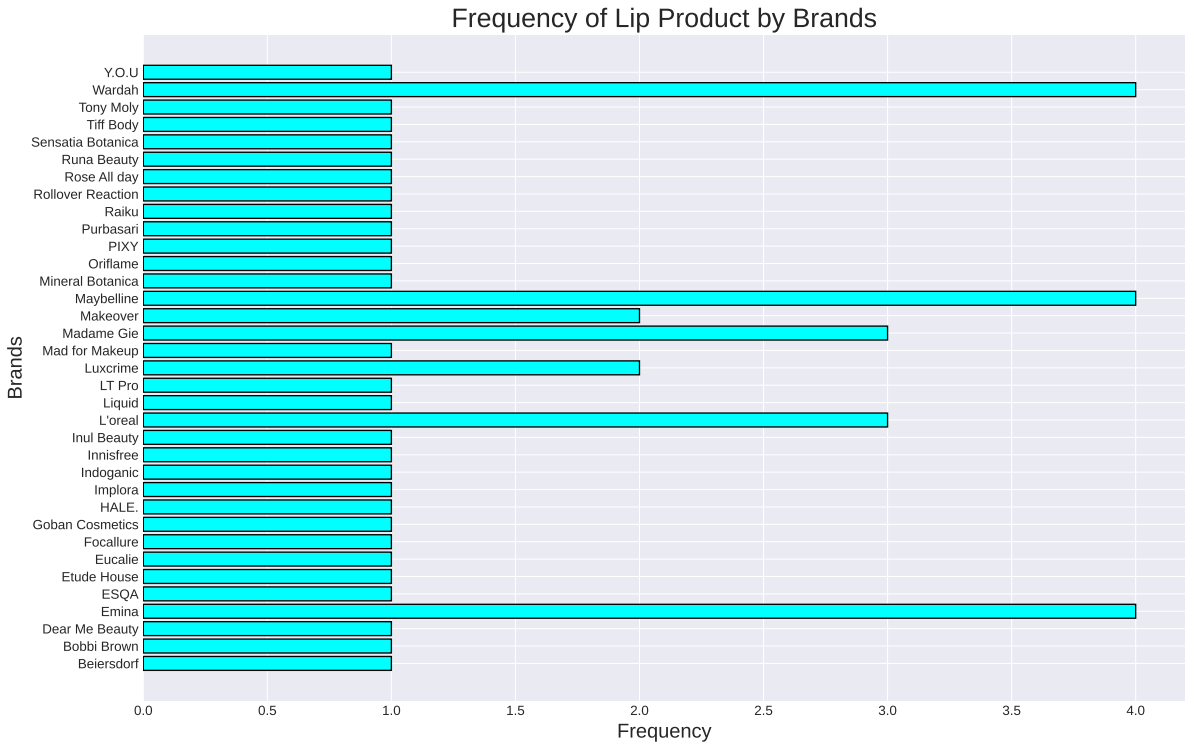

Figure 6: Frequency of Products by Brands

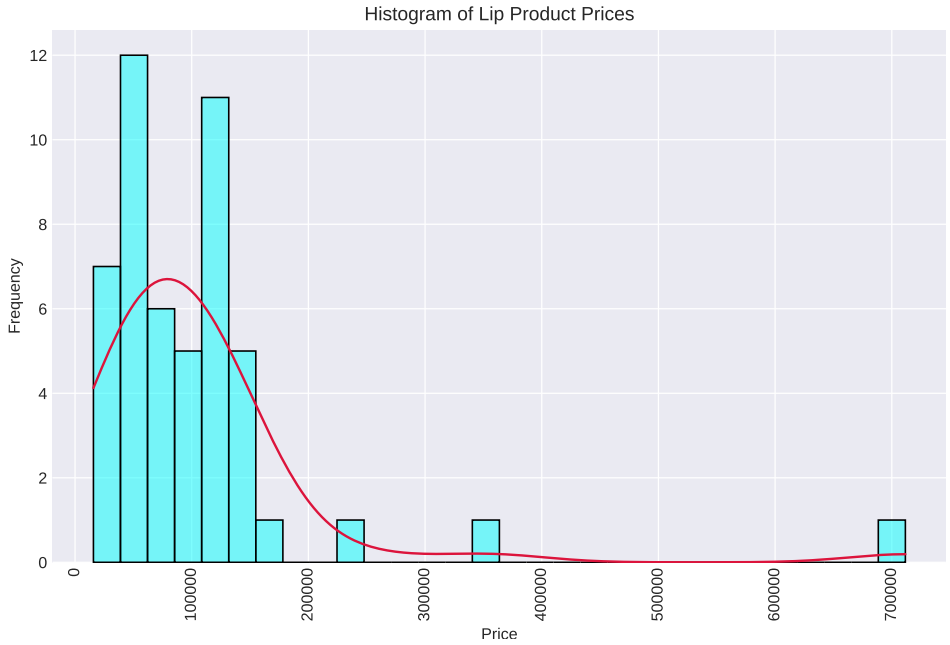

Figure 7: Product - Price Distribution - with KDE

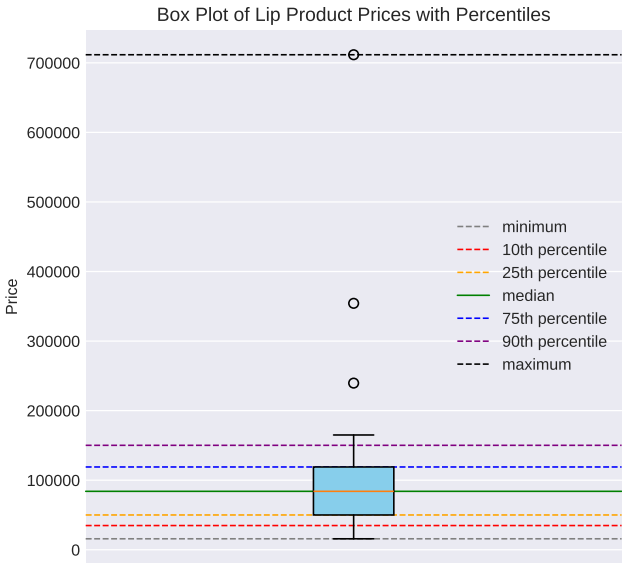

Figure 8: Product - Price Distribution - BoxPlot

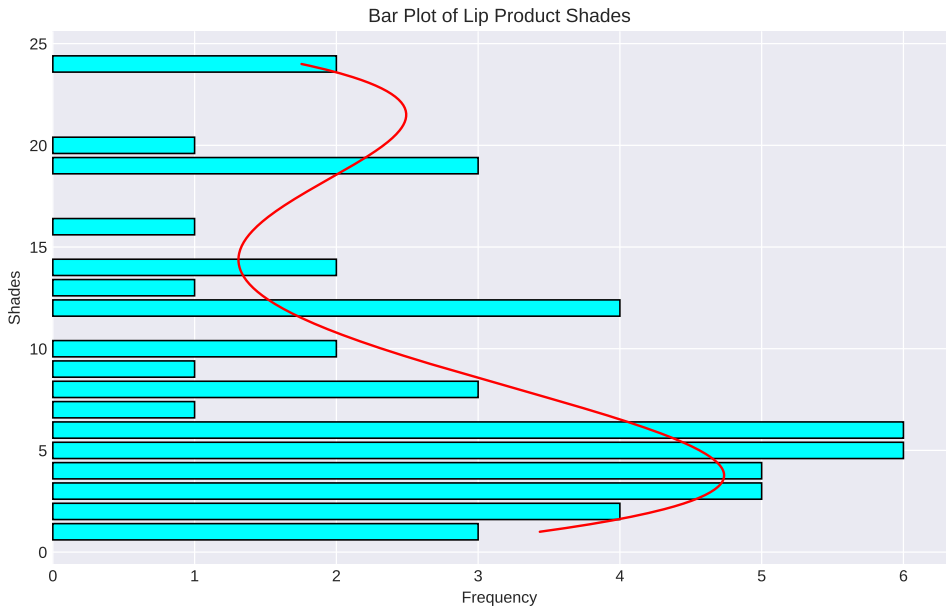

Figure 9: Product - Shades Distribution - with KDE

6 Analytics for India

6.1 Raw Data

Table 6: Raw Data for Study (2023)

Lip Product	Brand	Type	Shades	Price
Color Sensational Creamy Matte Lipstick	Maybelline	Lipstick	44	299
SuperStay Matte Ink Liquid Lipstick	Maybelline	Liquid Lipstick	30	699
SuperStay Matte Ink Crayon Lipstick	Maybelline	Lip Crayon	10	599
Lifter Gloss	Maybelline	Lip Gloss	4	799
Baby Lips Loves Color Lip Balm	Maybelline	Lip Balm	3	175
Baby Lips Bloom Lip Balm	Maybelline	Lip Balm	2	180
9 to 5 Powerplay Priming Matte Lip Color	Lakmé	Lipstick	35	599
Absolute Matte Melt Liquid Lip Color	Lakmé	Liquid Lipstick	15	800
Absolute Plump and Shine Lip Gloss	Lakmé	Lip Gloss	8	800
Lip Love Chapstick	Lakmé	Lip Balm	7	160
Xtraordin-airy Lip Mousse	Lakmé	Liquid Lipstick	14	700
Color Riche Lipstick	L'Oréal Paris	Lipstick	43	799
Rouge Signature Matte Liquid Lipstick	L'Oréal Paris	Liquid Lipstick	13	899
Matte Luxe Lipstick	Nykaa	Lipstick	12	849
So Matte Lipstick	Nykaa	Lipstick	48	450
So Creme! Creamy Matte Lipstick	Nykaa	Lipstick	18	329
Matte To Last! Liquid Lipstick	Nykaa	Liquid Lipstick	24	675
Gloss It Up! High Shine Lip Gloss	Nykaa	Lip Gloss	4	550
Macaron Lip Balm	Nykaa	Lip Balm	7	299
Lips Don't Lie! Line & Fill Lip Liner	Nykaa	Lip Liner	16	399
Ultra Matte Lipstick	Nykaa	Lipstick	17	649
Matte-ilicious lip crayon lipstick	Nykaa	Lip Crayon	8	799
Serial Kisser Moisturising Tinted Lip Balm	Nykaa	Lip Balm	5	179
Paintstix! Waterproof Matte Lipstick	Nykaa	Lipstick	9	525
All day matte liquid lipstick	Nykaa	Liquid Lipstick	19	399
Gloss it up pH Lip gloss oil	Nykaa	Lip gloss	1	575
Get set matte! Demi matte lip cream liquid lipstick	Nykaa	Liquid Lipstick	12	499
8 hour lasting full cover matte gloss	Nykaa	Lip gloss	8	599
Matte tattoo liquid lipstick	Nykaa	Liquid Lipstick	10	899
Matte As Hell Crayon Lipstick	Sugar Cosmetics	Lipstick	34	899
Smudge Me Not Liquid Lipstick	Sugar Cosmetics	Liquid Lipstick	47	499
Time To Shine Lip Gloss	Sugar Cosmetics	Lip Gloss	8	499
Nothing Else Matter Longwear Lipstick	Sugar Cosmetics	Lip Balm	21	599
Lipping On The Edge Lip Liner	Sugar Cosmetics	Lip Liner	6	525
Soft Matte Lip Cream	Miss Claire	Lipstick	52	225
Long Lasting Matte Lipstick	Miss Claire	Liquid Lipstick	21	325

Continued on next page

Table 6 – Continued from previous page

Lip Product	Brand	Type	Shades	Price
Matte & Pearly Gloss	Miss Claire	Lip Gloss	22	75
Butter Lip Balm	Miss Claire	Lip Balm	9	150
Glimmersticks Lip Liner	Miss Claire	Lip Liner	25	65
Ecostay matte lip lacquer	Lotus Herbals	Lip Lacquer	16	616.25
Proedit lip plumper + gloss	Lotus Herbals	Lip Plumper	10	505.75
Lotus herbals LIP BALM	Lotus Herbals	Lip Balm	2	169.28
Ecostay butter matte lip color	Lotus Herbals	Lip Color	22	531.25
Colorkick exfoliating & Hydrating Lip sugar	Lotus Herbals	Lip Color	2	293.25
Proedit liquid matte lip color	Lotus Herbals	Lip Color	12	633.25
Proedit silk touch Matte Lip Color	Lotus Herbals	Lip Color	2	590.75
Super lustrous lipstick	Revlon	Lipstick	38	799
Super lustrous(bold matte)	Revlon	Lipstick	12	799
Colorstay satin Ink Lip Color	Revlon	Lipstick	22	999
Colorstay overtime lip color	Revlon	Lipstick	10	1300
Colorstay Longwear Lip Liner	Revlon	Lip Liner	4	810
Super Lustrous The Luscious Mattes Lipstick	Revlon	Lipstick	12	799
Colorstay matte lite crayon	Revlon	Lip Crayon	10	999
Ultra HD vinyl Lip Polish	Revlon	Lip Polish	8	1250
Super lustrous lipstick	Revlon	Lipstick	12	799
Colorstay Matte lite crayon	Revlon	Lip Crayon	10	999
Colorstay suede ink	Revlon	Lipstick	10	1199
Nourishing Tinted 100% Natural Lip Balm	Mamaearth	Lip Balm	1	199
Nourishing 100% Natural Lip Balm	Mamaearth	Lip Balm	1	149
Tinted 100% Natural Lip Balm	Mamaearth	Lip Balm	4	299
Soft Matte Long Stay Lipstick	Mamaearth	Lipstick	9	399
Moisture Matte Long Stay Lipstick	Mamaearth	Lipstick	17	499
Creamy Matte Long Stay Lipstick	Mamaearth	Lipstick	9	399
Feather Light Liquid Matte Lipstick	Mamaearth	Liquid Lipstick	4	249
Color Pop Matte Lip Color	Elle 18	Lipstick	24	100
Color Pops Silk Lipstick	Elle 18	Lipstick	3	100
Liquid Lip Color	Elle 18	Liquid Lipstick	40	150
Lit Lip Stack Liquid Lipstick	Elle 18	Liquid Lipstick	6	275
OMG Lip Gloss	Elle 18	Lip Gloss	7	175
2 Timing Lip & Cheek Tint	Elle 18	Lip Tint	2	150

6.2 Sampled Data

6.2.1 On Basis of Brands

Table 7: Products grouped by Brand

Brand	Frequency	Percentage	Cumulative Percentage
Nykaa	16	22.85	22.86
Revlon	11	15.71	38.58
Mamaearth	7	10.00	48.58
Lotus Herbals	7	10.00	58.58
Elle 18	6	8.57	67.15
Maybelline	6	8.57	75.72
Lakmé	5	7.14	82.86
Sugar Cosmetics	5	7.14	90.00
Miss Claire	5	7.14	97.14
L'Oréal Paris	2	2.86	100.0

6.2.2 On Basis of Type

Table 8: Products grouped by Type

Type	Frequency	Percentage	Cumulative Percentage
Lipstick	22	31.44	31.44
Liquid Lipstick	13	18.57	50.01
Lip Crayon	4	5.71	55.72
Lip Gloss	6	8.57	64.29
Lip Balm	11	15.71	80.00
Lip Liner	4	5.71	85.71
Lip gloss	2	2.86	88.57
Lip Lacquer	1	1.43	90.00
Lip Plumper	1	1.43	91.43
Lip Color	4	5.71	97.14
Lip Polish	1	1.43	98.57
Lip Tint	1	1.43	100.0

6.2.3 On Basis of Shades

Table 9: Products grouped by Shades

Shades	Frequency	Percentage	Cumulative Percentage
1	3	4.28	4.28
2	5	7.14	11.42
3	2	2.86	14.28
4	5	7.14	21.42
5	1	1.43	22.85
6	2	2.86	25.71
7	3	4.28	29.99
8	5	7.14	37.13
9	4	5.71	42.84
10	7	10.0	52.84
12	6	8.56	61.40
13	1	1.43	62.83
14	1	1.43	64.26
15	1	1.43	65.69
16	2	2.86	68.55
17	2	2.86	71.41
18	1	1.43	72.84
19	1	1.43	74.27
21	2	2.86	77.13
22	3	4.28	81.41
24	2	2.86	84.27
25	1	1.43	85.70
30	1	1.43	87.13
34	1	1.43	88.56
35	1	1.43	89.99
38	1	1.43	91.42
40	1	1.43	92.85
43	1	1.43	94.28
44	1	1.43	95.71
47	1	1.43	97.14
48	1	1.43	98.57
52	1	1.43	100.0

6.3 Graphs and Stats

Table 10: Descriptive Statistics of Shades and Prices

Attribute	**Shades**	**Price**
Count	70.0	70.0
Mean	14.74	532.54
Median	10.0	525.0
Mode	10.0	799.0
Std Dev	12.72	305.88
Variance	161.76	93565.36
Minimum	1.0	65.0
Maximum	52.0	1300.0
Percentiles:		
- 0th	1.0	65.0
- 25th	6.25	279.56
- 50th	10.0	525.0
- 75th	20.5	799.0
- 100th	52.0	1300.0

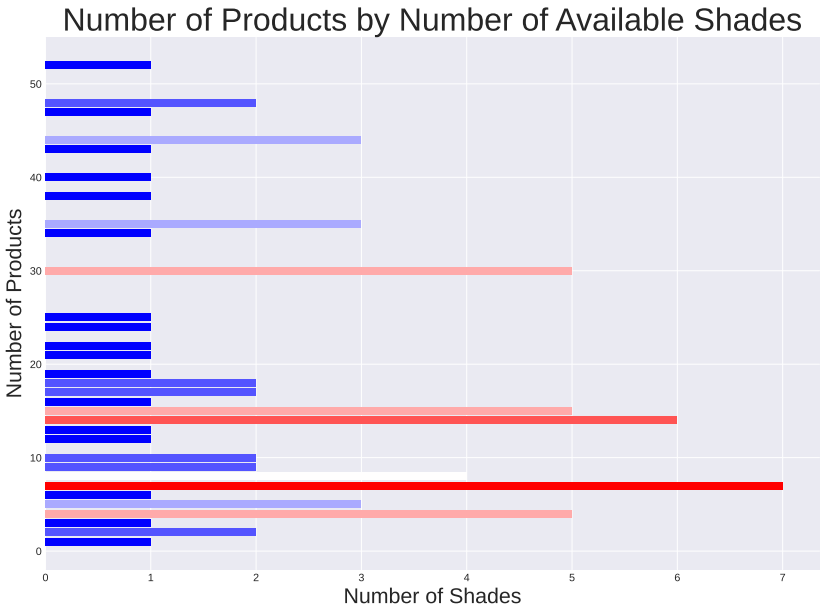

Figure 10: Total Products Grouped by Shades

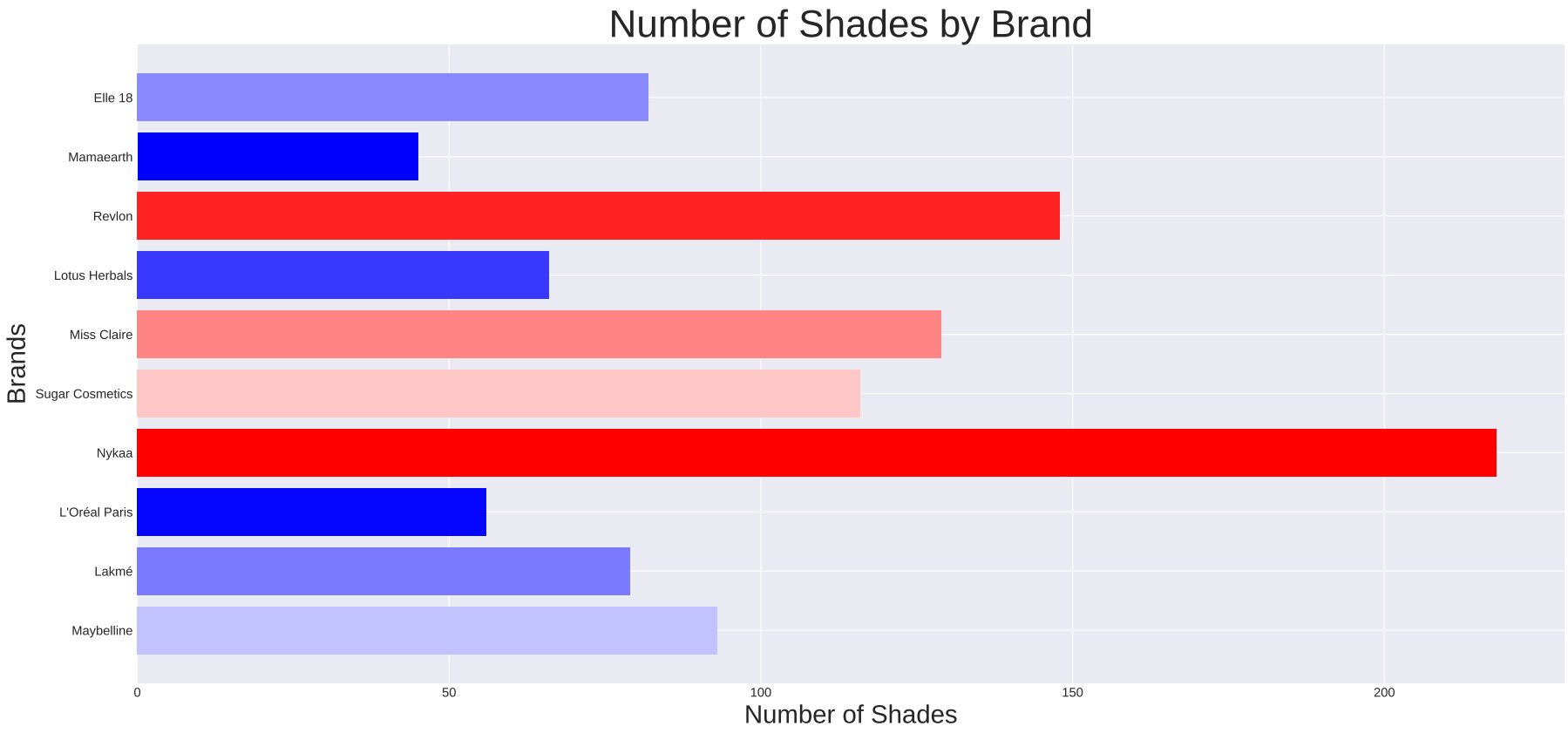

Figure 11: Total Shades Grouped by Brands

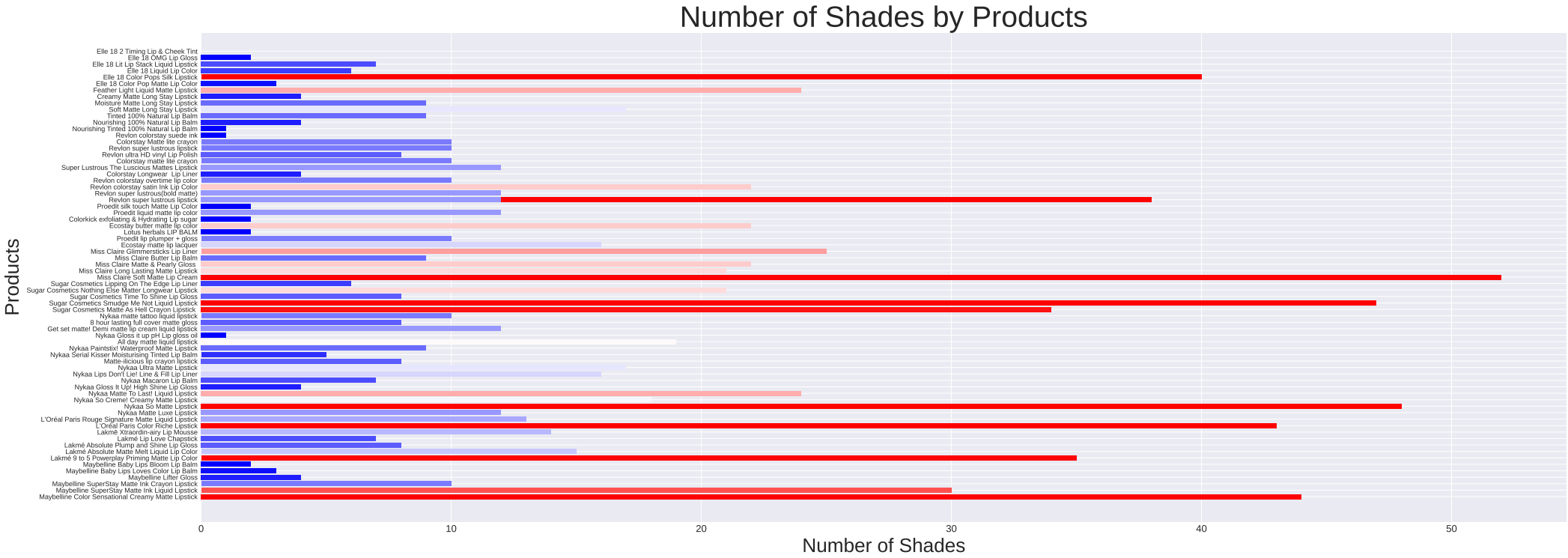

Figure 12: Total Shades Grouped by Product

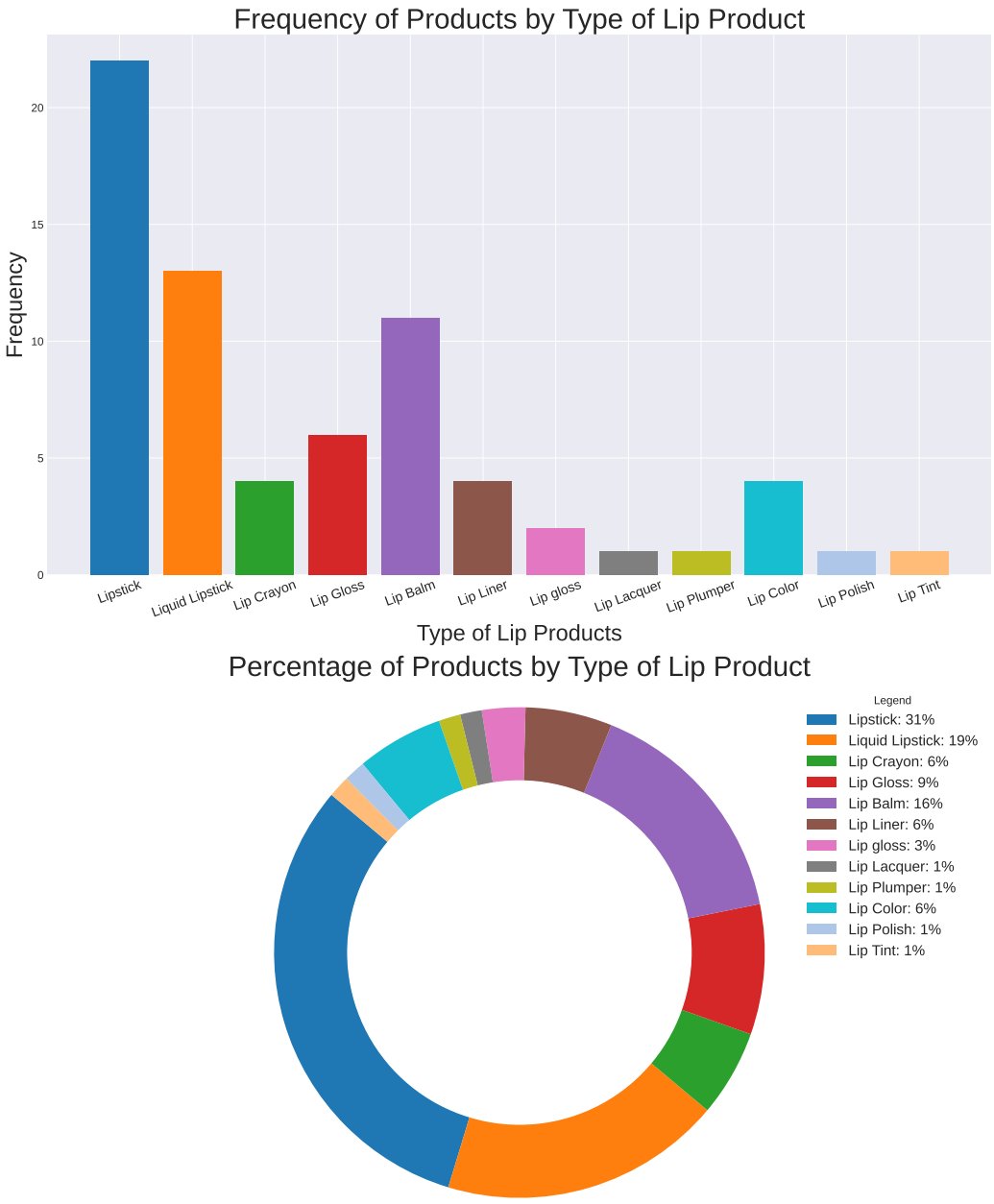

Figure 13: Total Products Grouped by Type

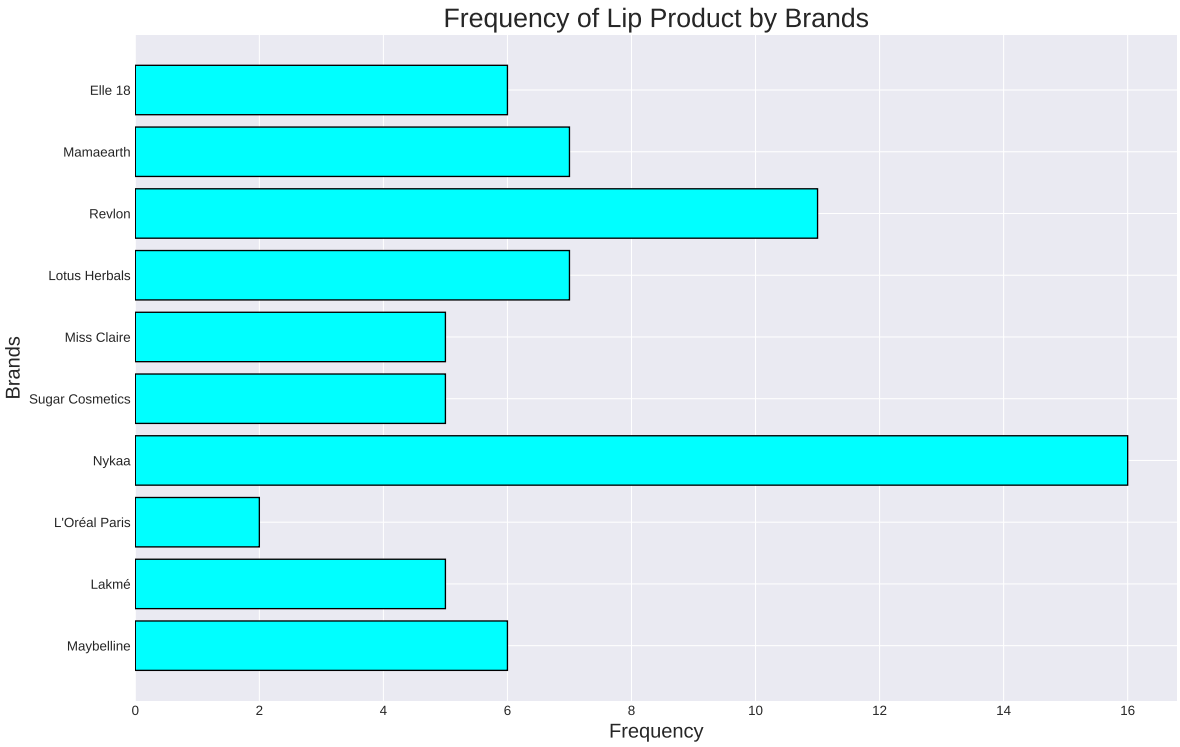

Figure 14: Frequency of Products by Brands

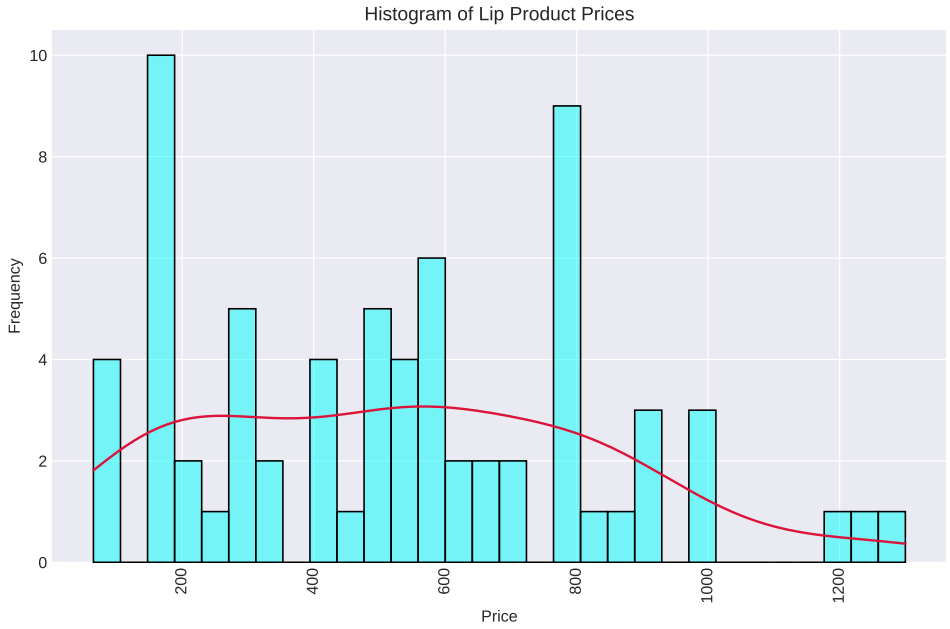

Figure 15: Product - Price Distribution - with KDE

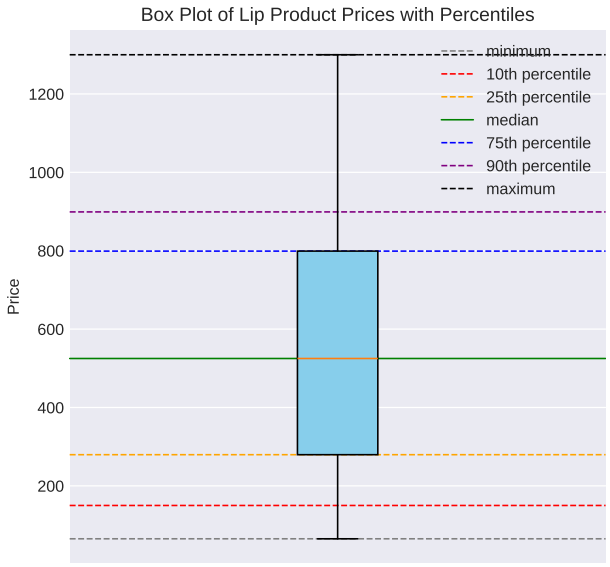

Figure 16: Product - Price Distribution - BoxPlot

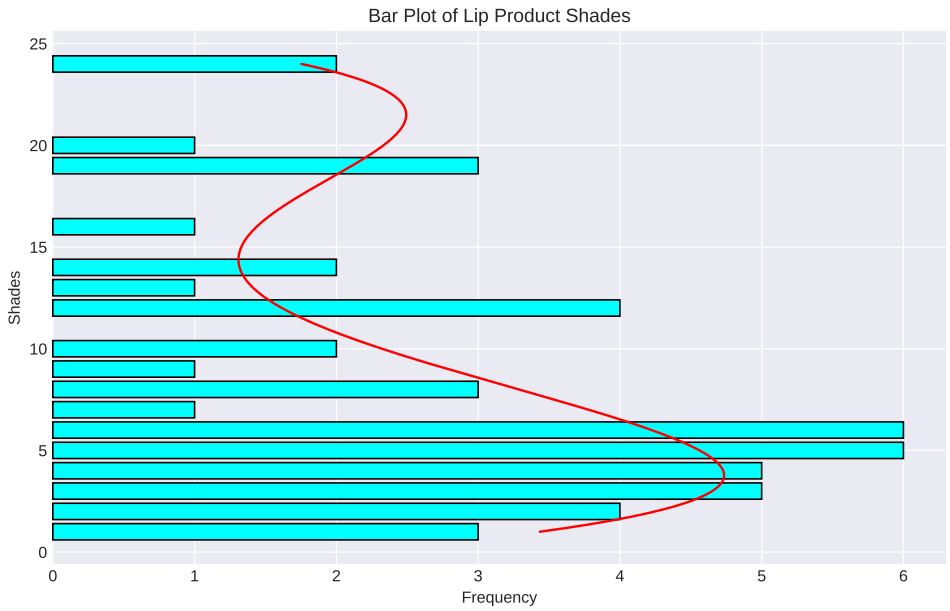

Figure 17: Product - Shades Distribution - with KDE

7 Inference

Most of the products are present in the price range of 15000 to 175000. Some of the products lie outside this range and can be seen as occasional peaks in the bar plot. The maximum price of any product is 711636 and the minimum is 15765. The mean price of all products is 104456 and the median is 84000. The data also shows a standard deviation of around 105480.

The number of shades per product varies from 1 to 24, with the mean being 8.04 and the median being 6. Also, companies like Wardah, Maybelline, Makeover, L'oreal and Implora boast a large number of shades in their products while other companies do not have as much variety.

Both variables have relatively high standard deviations compared to their means, indicating wide variability when it comes to prices and shade counts around their respective averages.

Outliers are present in both variables, especially considering the large difference between the 75th percentile and the maximum values for both price and shades.

Most of the lip products produced by these brands fall under the 'liquid' category, as evident from the donut chart. It accounts for 72 percent of the total products. From the remaining, 24 percent are 'stick' type and solid and crayon account for 2 percent each.

Also, most companies only have one product present in the data, with some exceptions being Emina, L'oreal, Luxcrime, Madame Gie, Makeover, Maybelline and Wardah which have 2 or more products each. Emina, Wardah and Maybelline form the most popular brands in Indonesia closely followed by Madame Gie and L'oreal.

References

[1] Sarah Boslaugh. *Statistics in a Nutshell*. O'Reilly Media, 2nd edition, 2012.

[2] Chris P. Tsokos and Kandethody M. Ramachandran. *Mathematical Statistics with Applications*. Academic Press, 1st edition, March 2009.

[3] Pritha Bhandari. Sampling bias and how to avoid it | types & examples, 2020.

[4] Dety Nurfadilah and Yulita F Susanti. *Case Study for Descriptive Statistics*. Ipmi Press, 2022.

[5] Best lipstick brands in india. Online. Accessed on April 10, 2024.

[6] 12 best lipsticks in 2023: India's top brands & shades. Online. Accessed on April 10, 2024.

[7] 10 best makeup brands in india - top indian cosmetic brands. Online. Accessed on April 10, 2024.

Contributions

Table 11: Contributions of the authors

Name	**Roll No.**	**Contribution in Report Writing**	**Contribution in Analysis**	**Details of use of web resources/ Codes/AI tools, etc.**	**Overall Contribution to work done**
Arka Mukhopadhyay	B23120	Remaining	Grouping products by brands and type	CSV to LaTeX table converters	16.8
Pranab Ray	B23169	Data Analysis Inference	Remaining	ChatGPT and Gemini to refine writeups	16.8
Kamal Yadav	B23209	Data Collection	Collecting Indian Data	-	16.6
Arani Ghosh	B23119	-	KDE and Descriptive Statistics	GitHub Copilot to beautify graphs	16.6
Ayuj Aryan	B23198	Data Collection	Collecting Indonesian Data	-	16.6
Kunal Sharma	B23079	Data Collection	Collecting Indian Data	-	16.6