



## Assignment 2

# KELOMPOK 02

ANALISIS ASSOCIATION RULE TERHADAP  
DATASET ULASAN PRODUK KOSMETIK DAN  
KECANTIKAN

162112133038

R. Firdaus Dharmawan Akbar

162112133062

Ridho Pandhu Afrianto

162112133085

Janetha Maesha Purba

162112133104

Syahlla Muthi Huwaida





## Introduction



# Cosmetics & Beauty

Kosmetik dan produk kecantikan adalah sebuah produk yang biasanya digunakan oleh perempuan untuk merawat bahkan memperbaiki kesehatan kulit, rambut, dan lain-lain.

## KOSMETIK PADA ERA MODERN

Namun, pada zaman ini hal tersebut bukan lagi menjadi hal yang dikhkususkan untuk perempuan. Sehingga pria, wanita, remaja, bahkan orang tua pun bisa memakai produk-produk kecantikan.



## Introduction

# About Dataset

Berdasarkan dataset yang didapatkan dari website [Kaggle](#) yang berjudul,

### *Cosmetics and Beauty Products Reviews - Top Brands*

Dilakukanlah analisis untuk merekomendasikan produk-produk yang ada kepada orang-orang yang **bingung** dan **baru** ingin menggunakan kosmetik atau produk kecantikan.



# kaggle

	review_date	author	brand_name	product_title	price	review_rating	product_rating
0	2021-01-23 15:17:18	Ashton Dsouza	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	5.0	4.1
1	2020-09-07 15:30:42	Amrit Neelam	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	5.0	4.1
2	2020-11-13 12:24:14	Sanchi Gupta	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	4.0	4.1
3	2020-06-14 11:56:50	Ruchi Shah	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	3.0	4.1
4	2020-12-22 15:24:35	Sukanya Sarkar	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	2.0	4.1
...	...	...	...	...	...	...	...
61279	2021-04-12 13:28:41	Radhesh	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61280	2021-04-12 11:47:07	Shalini Sirohi	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61281	2021-04-12 10:58:34	Sana kanted	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61282	2021-04-11 21:44:55	neeru chouhan	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61283	2022-06-20 17:41:52	Mssuhv	Colorbar	Colorbar Mogra Mist - Scent Of An India Temple...	417	2.0	1.5

Dataset ini memiliki **61.284** observasi dan **18** atribut. Setelah melakukan data preprocessing, variabel yang digunakan ada sebanyak 7 variabel, yaitu review date, author, brand name, product title, price, review rating, dan product rating.





## Karena,

data mentah yang didapatkan masih memiliki **banyak kekurangan**.

Data mentah yang tersedia biasanya masih memiliki banyak missing data, data duplikat, bahkan outlier yang menyebabkan **sulitnya** menemukan pola atau insight baru dari data tersebut.

Dengan melakukan data pre-processing, data yang kita gunakan untuk analisis akan menjadi lebih **berkualitas**, **akurat**, dan **dapat dipercaya**. Ini akan membantu kita dalam membuat keputusan yang lebih baik berdasarkan analisis yang dilakukan.

# Why Data Pre-Processing?

Berikut ini adalah tahapan data preprocessing yang dilakukan:

Datetime  
Modification



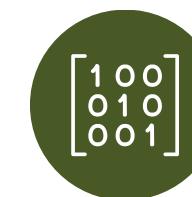
Hapus  
Missing Value



Hapus beberapa  
Author



Code  
*customer review*





# Data Pre-Processing

Step 1 - Datetime

# Datetime Modification

K E L O M P O K O 2 | D A T A M I N I N G

	review_date	author	brand_name	product_title	price	review_rating	product_rating	Jam	Waktu	Hari	Bulan	Tahun
0	2021-01-23 15:17:18	Ashton Dsouza	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	5.0	4.1	15	Siang	Sabtu	Januari	2021
1	2020-09-07 15:30:42	Amrit Neelam	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	5.0	4.1	15	Siang	Senin	September	2020
2	2020-11-13 12:24:14	Sanchi Gupta	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	4.0	4.1	12	Pagi	Jumat	November	2020
3	2020-06-14 11:56:50	Ruchi Shah	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	3.0	4.1	11	Pagi	Minggu	Juni	2020
4	2020-12-22 15:24:35	Sukanya Sarkar	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	2.0	4.1	15	Siang	Selasa	Desember	2020

Mengubah variabel review date yang berisikan tanggal dan waktu sehingga waktu, hari, bulan, dan tahunnya menjadi variabel yang baru. Hal ini nantinya akan divisualisasikan untuk mendapatkan insight baru.

```
••• Datetime
locale.setlocale(locale.LC_TIME, 'id_ID')
df['review_date'] = pd.to_datetime(df['review_date'])
bins1 = pd.to_datetime(['00:00:00', '12:00:00', '18:00:00', '23:59:59'])
df['Jam'] = df['review_date'].dt.hour
df['Waktu'] = pd.cut(df['Jam'], bins= [0, 12, 18, 24], labels=['Pagi', 'Siang', 'Malam'])
df['Hari'] = df['review_date'].dt.day_name(locale='id_ID')
df['Bulan'] = df['review_date'].dt.month_name(locale='id_ID')
df['Tahun'] = df['review_date'].dt.year
```



# Hapus Missing Value

Missing value bisa disebabkan oleh banyak faktor, namun karena data ini merupakan data sekunder, sehingga kita **tidak dapat** mencari nilai dari missing value tersebut.

Berdasarkan hal tersebut, diputuskan untuk membersihkan data dengan **menghapus** missing value yang ada. Penghapusan missing value juga dapat mempermudah dan mengefisienkan hasil yang ada.

**df = df.dropna()**





# Data Pre-Processing

## Step 3 – Author

# Hapus Beberapa Author

	review_date	author	brand_name	product_title	price	review_rating	product_rating
0	2021-01-23 15:17:18	Ashton Dsouza	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	5.0	4.1
1	2020-09-07 15:30:42	Amrit Neelam	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	5.0	4.1
2	2020-11-13 12:24:14	Sanchi Gupta	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	4.0	4.1
3	2020-06-14 11:56:50	Ruchi Shah	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	3.0	4.1
4	2020-12-22 15:24:35	Sukanya Sarkar	Olay	Olay Ultra Lightweight Moisturiser: Luminous W...	1599	2.0	4.1
...	...	...	...	...	...	...	...
61279	2021-04-12 13:28:41	Radhesh	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61280	2021-04-12 11:47:07	Shalini Sirohi	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61281	2021-04-12 10:58:34	Sana kanted	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61282	2021-04-11 21:44:55	neeru chouhan	Herbal Essences	Herbal Essences Aloe & Eucalyptus Shampoo For ...	525	5.0	4.3
61283	2022-06-20 17:41:52	Mssuhv	Colorbar	Colorbar Mogra Mist - Scent Of An India Temple...	417	2.0	1.5

Rekomendasi barang didapatkan dari review yang diberikan oleh author, oleh sebab itu diperlukan perhitungan {value count} author mereview beberapa barang. Author yang mereview sedikit barang akan dibuang karena hal ini dapat menyebabkan minimal supportnya menjadi rendah. Pada data ini, diputuskan bahwa author yang dibuang ketika mereview kurang dari 25 barang.

• • • Slicing Data  
`counts = df['author'].value_counts()  
to_remove = counts[counts < 25].index.tolist()  
df.drop(df[df['author'].isin(to_remove)].index, inplace=True)`



# Data Pre-Processing

## Step 4 – Binary Code

# Code

# Customer Review

product_title	Get Party Ready With Kay Beauty Lip Liner - Dance Off & Matte Lipstick - Premier	Herbal Essences Aloe & Bamboo Conditioner Soft Smooth Hair- No- Sulphates and Paraben(400gm)	Herbal Essences Aloe & Bamboo Shampoo + Conditioner For Soft Smooth Hair, No- Sulphates, Paraben and Sulphates & Paraben	Herbal Essences Aloe & Eucalyptus Shampoo + Conditioner For Soft Smooth Hair, No- Sulphates, Paraben and Sulphates & Paraben	Herbal Essences Aloe & Eucalyptus Shampoo For Soft Smooth Hair, No- Sulphates, Paraben and Sulphates & Paraben	Herbal Essences Argan Oil Of Moroccan Shampoo & Conditioner With Free Neem Comb	Herbal Essences Argan Oil Of Morocco Conditioner - For Frizz Free Hair - Paraben Free	Herbal Essences Argan Oil Of Morocco Shampoo - For Frizz Free Hair - Paraben Free		
author	Akriti Awasthi	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	Aliaaa	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	Ananya	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	Ankita	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0	Ankita Pawar	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

Customer bisa mereview banyak barang sehingga menghasilkan data yang banyak, oleh sebab itu customer diubah menjadi unique code. Unique code diubah dengan pivot sehingga jika customer mereview produk tersebut akan dilabelling menjadi 1 dan jika customer tidak mereview produk tersebut akan dilabelling menjadi 0.

```
Code Customer Review

item_count_pivot = df.pivot_table(index='author',columns='product_title', values='review_rating',aggfunc='sum').fillna(0)
item_count_pivot[item_count_pivot > 0] =1
item_count_pivot.head()
```



1

## EXPLORATORY DATA ANALYSIS

Dataset akan diidentifikasi untuk menemukan **pola** pada setiap item yang direview oleh author. Identifikasi dilakukan dengan cara memvisualisasikan pola-pola pada masing-masing kategori.

3

## MENENTUKAN NILAI ASOSIASIF MINIMUM

Pada analisis ini kami menentukan nilai minimum confident sebesar 1 atau 100%. Nilai confident pada masing-masing asosiasi yang kurang dari 1 tidak termasuk rekomendasi review itemset.

2

## MENENTUKAN NILAI FREKUENSI SUPPORT MINIMUM

Pada analisis ini kami menentukan nilai minimum support sebesar **0,15** atau **15%**. Maka dari itu, nilai support pada masing-masing itemset yang kurang dari 0,15 tidak akan dilanjutkan ke proses selanjutnya.

4

## MENENTUKAN NILAI LIFT

Nilai lift berfungsi untuk memberikan informasi seberapa kuatnya asosiasi tersebut. Asosiasi dikatakan kuat apabila nilai lift pada asosiasi tersebut lebih dari 1. Semakin besar nilai lift pada asosiasi tersebut maka tingkat keakuratannya dapat digunakan untuk rekomendasi reviewset selanjutnya.



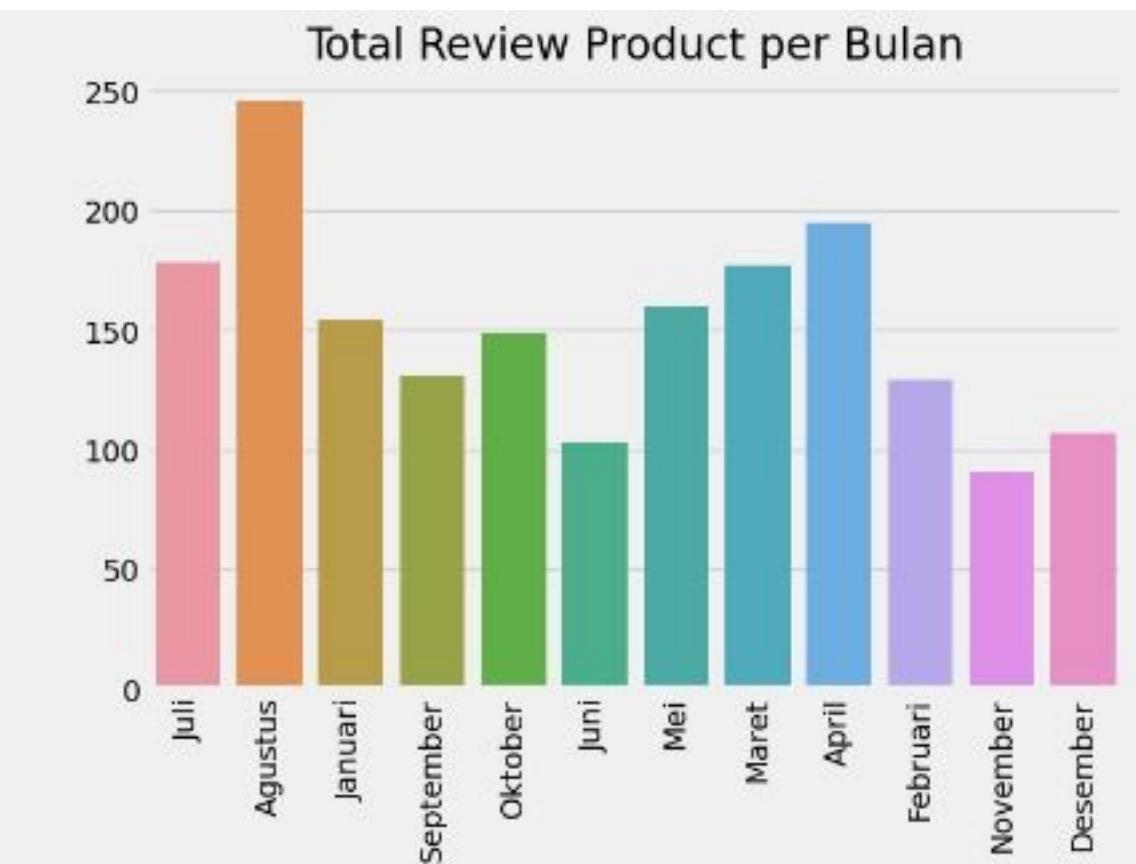
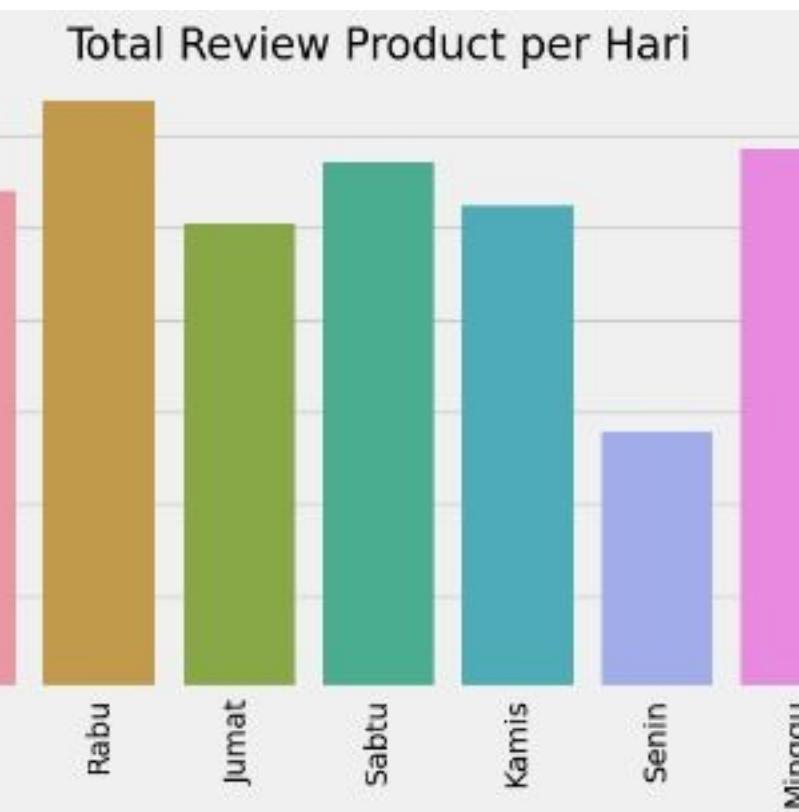
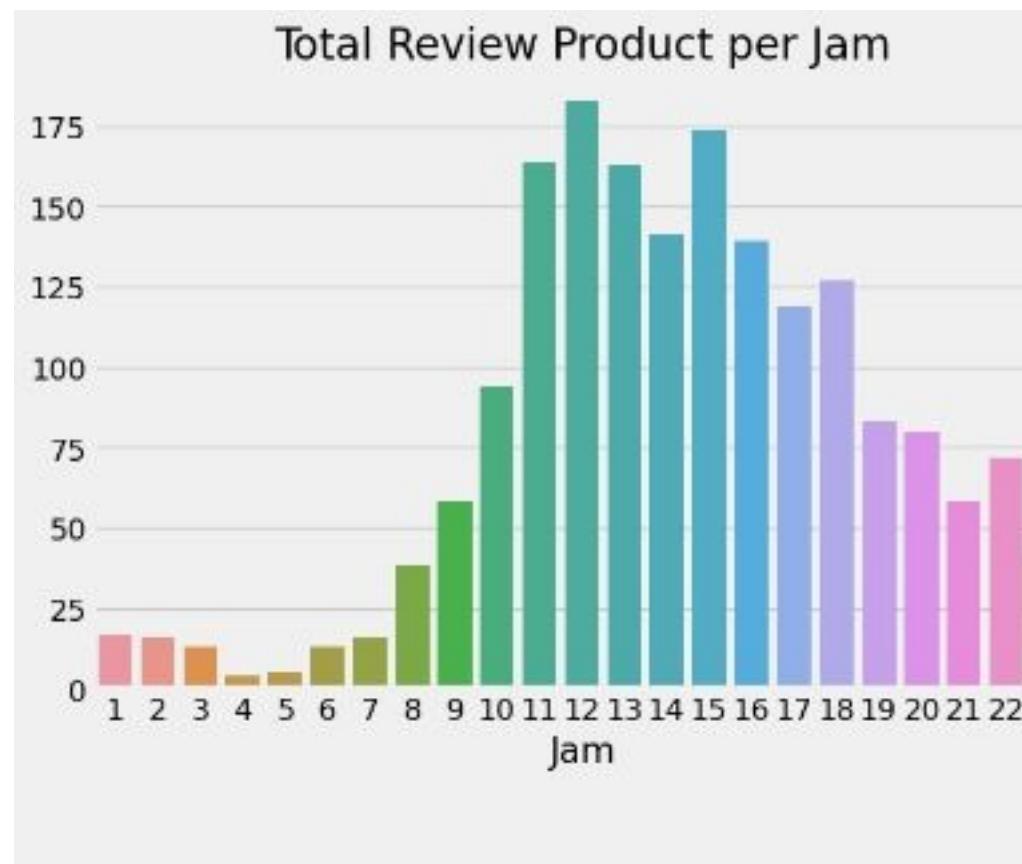


## Result & Analysis

### EDA - Time Chart

# Hasil Analisis EDA

K E L O M P O K O 2 | D A T A M I N I N G





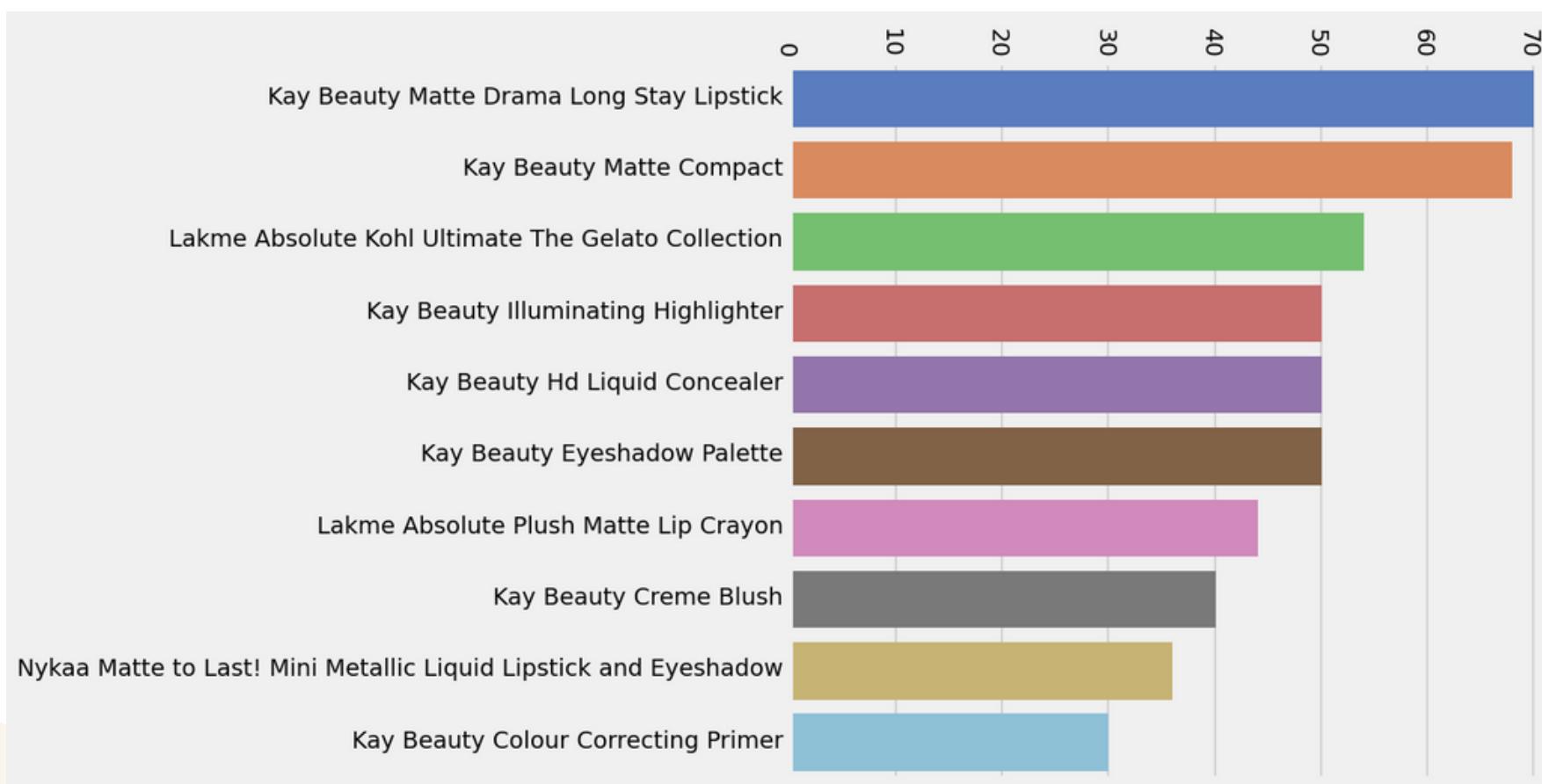
# Result & Analysis

## EDA - Item Chart

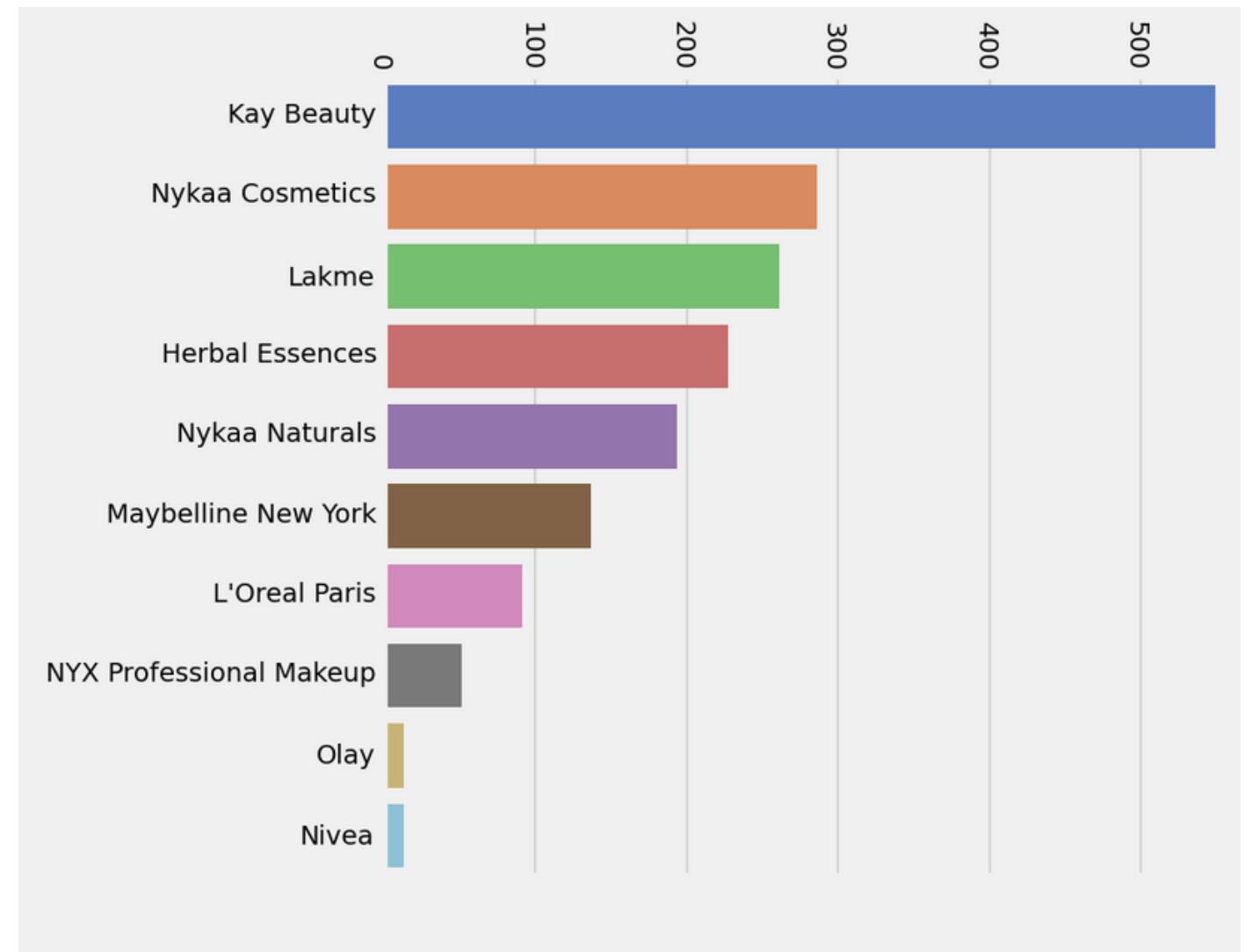
# Hasil Analisis

## EDA

### Top Products



### Top Brands





Pada proses preparation ini kami menggunakan tabel pivot dimana informasi pada database akan diringkas atau dirangkum agar data yang besar dapat mudah dibaca. Pada tabel pivot di bawah ini, kami mengidentifikasi dimensi dataset, jumlah reviewer item, dan jumlah item.

Dimensi Dataset	Jumlah Reviewer Item	Jumlah Item
(41,220)	41	220



## Result & Analysis

DM - Finding Frequent Itemsets

# Hasil Analisis Data Modelling

K E L O M P O K 0 2 | D A T A M I N I N G

No.	Itemsets	Support
1	Kay Beauty Matte Compact	0.51
2	Kay Beauty Eyeshadow Palette	0.49
3	Kay Beauty Illuminating Highlighter	0.44
4	Lakme Absolute Kohl Ultimate The Gelato Collection	0.44
5	Lakme Kajal - Black	0.44
6	Nykaa Matte to Last! Mini Metallic Liquid Lipstick and Eyeshadow	0.44

No.	Itemsets	Support
7	Kay Beauty Ultra Black Volume & Length Mascara - Midnight	0.41
8	'Kay Beauty Matte Compact', 'Kay Beauty Eyeshadow Palette'	0.41
9	Kay Beauty Colour Corrector Stick	0.40
10	Kay Beauty Creme Blush	0.40
...	...	...
400	Kay Beauty Eyeshadow Palette, Lakme Absolute Kohl Ultimate The Gelato Collection, Kay Beauty Hd Liquid Concealer, Kay Beauty Ultra Black Volume & Length Mascara - Midnight, Kay Beauty Matte Compact	0.17



# Result & Analysis

## DM - Generating Association Rules

# Hasil Analisis Data Modelling

No	Antecedents	Consequents	Sup	Conf	Lift	No	Antecedents	Consequents	Sup	Conf	Lift
1	{'Kay Beauty Hd Liquid Concealer', 'Lakme 9to5 Primer + Matte Liquid Lip Color'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	11	{'Nykaa Naturals Apple Cider Vinegar & Ginger Shampoo & Conditioner- Anti-Dandruff Combo', 'Kay Beauty Gel Eye Pencil'}	{'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	0.17	1	31.54
2	{'Kay Beauty Hd Liquid Concealer', 'Nykaa Naturals Apple Cider Vinegar & Ginger Shampoo & Conditioner- Anti-Dandruff Combo'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	12	{'Kay Beauty Matte Compact', 'Kay Beauty Ultra Black Volume & Length Mascara - Midnight', 'Kay Beauty Illuminating Highlighter'}	{'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	0.17	1	31.54
3	{'Kay Beauty Creme Blush', 'Kay Beauty Eyeshadow Palette', 'Kay Beauty Hd Liquid Concealer'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	13	{'Kay Beauty Matte Compact', 'Lakme Absolute Kohl Ultimate The Gelato Collection', 'Kay Beauty Illuminating Highlighter'}	{'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	0.17	1	31.54
4	{'Kay Beauty Creme Blush', 'Kay Beauty Eyeshadow Palette', 'Kay Beauty Ultra Black Volume & Length Mascara - Midnight'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	14	{'Lakme 9to5 Primer + Matte Liquid Lip Color', 'Kay Beauty Gel Eye Pencil'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28
5	{'Lakme 9to5 Primer + Matte Liquid Lip Color', 'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	15	{'Nykaa Naturals Apple Cider Vinegar & Ginger Shampoo & Conditioner- Anti-Dandruff Combo', 'Kay Beauty Gel Eye Pencil'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28
6	{'Nykaa Naturals Apple Cider Vinegar & Ginger Shampoo & Conditioner- Anti-Dandruff Combo', 'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	16	{'Kay Beauty HD Translucent Loose Powder - White', 'Kay Beauty Illuminating Highlighter'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28
7	{'Kay Beauty Creme Blush', 'Kay Beauty Matte Compact', 'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	17	{'Kay Beauty Colour Correcting Primer', 'Kay Beauty Eyeshadow Palette', 'Kay Beauty Illuminating Highlighter'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28
8	{'Kay Beauty Matte Compact', 'Kay Beauty Creme Blush', 'Kay Beauty Eyeshadow Palette', 'Kay Beauty Ultra Black Volume & Length Mascara - Midnight'}	{'Kay Beauty Gel Eye Pencil',}	0.17	1	41	18	{'Kay Beauty Colour Corrector Stick', 'Kay Beauty Eyeshadow Palette', 'Kay Beauty Illuminating Highlighter'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28
9	{'Herbal Essences Bio:Renew Strength Whipped Cocoa Butter Conditioner'}	{'Herbal Essences Strawberry & Mint Shampoo - For Cleansing & Volume - Paraben Free',}	0.17	1	31.54	19	{'Kay Beauty Eyeshadow Palette', 'Lakme 9to5 Primer + Matte Liquid Lip Color', 'Kay Beauty Gel Eye Pencil'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28
10	{'Lakme 9to5 Primer + Matte Liquid Lip Color', 'Kay Beauty Gel Eye Pencil'}	{'Kay Beauty Hd Liquid Concealer', 'Kay Beauty Eyeshadow Palette'}	0.17	1	31.54	20	{'Kay Beauty Eyeshadow Palette', 'Nykaa Naturals Apple Cider Vinegar & Ginger Shampoo & Conditioner- Anti-Dandruff Combo', 'Kay Beauty Gel Eye Pencil'}	{'Kay Beauty Hd Liquid Concealer',}	0.17	1	29.28

## Antecedents

K	3	2	4	1
Value Counts	81	67	20	1

## Consequents

K	1	2	3
Value Counts	140	28	1



## Result & Analysis

DM - Further Generating Association Rules

# Hasil Analisis Data Modelling

K E L O M P O K 0 2 | D A T A M I N I N G

**Analisis Produk Skincare**

Berdasarkan Review Pelanggan  
Menggunakan Algoritma Apriori

---

Filter Brand Name :

Choose an option

Item 1

None

Item 2

None

Item 3

None

Item 4

None

---

<https://rfdharma-association-rule--skincare-review2app-selectbox-19z2m7.streamlit.app/>

# Kesimpulan Analisis

Berdasarkan hasil dari penerapan Association Rule dengan algortima Apriori pada dataset review produk scincare menggunakan minimum support sebesar 15% dan didapatkan item frequent sebanyak 400.

Kemudian menggunakan minimum confidence sebesar 100% dan didapatkan 169 Association Rule yang mempunyai 20 asosiasi antar produknya.

Pada penelitian ini ditemukan asosiasi yang paling tinggi yaitu 'Kay Beauty Hd Liquid Concealer', 'Lakme 9to5 Primer + Matte Liquid Lip' Color direview, maka 'Kay Beauty Gel Eye' Pencil juga akan direview dengan nilai support sebesar 0.17, nilai confidence sebesar 1, dan nilai lift sebesar 41.

Dengan adanya hasil penelitian ini diharapkan dapat membantu merekomendasikan para Author untuk memilih itemset yang akan diriview kedepannya.





**Cosmetics**  
& Beauty



K E L O M P O K 0 2

# THANK YOU

ASSIGNMENT 2 | DATA MINING

Analisis Association Rule terhadap dataset Ulasan  
Produk Kosmetik dan Kecantikan

End Of Slide

