

DISRUPTORES ENDOCRINOS EN COSMÉTICA

Silvia Di Bonaventura, Celia Manzano, Iván Simón

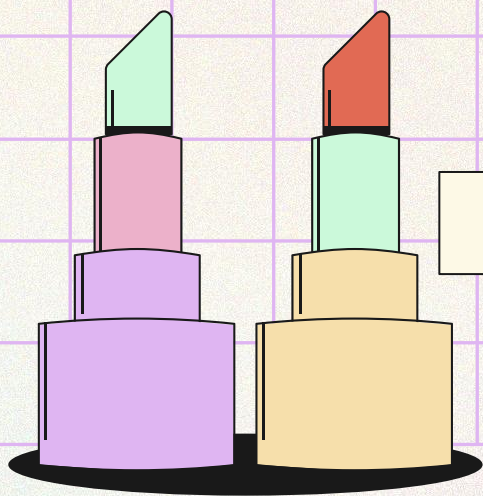


TABLE OF CONTENTS

#01

INTRODUCCIÓN

#02

METODOLOGÍA

#03

VISUALIZACIÓN

#04

CONCLUSIONES

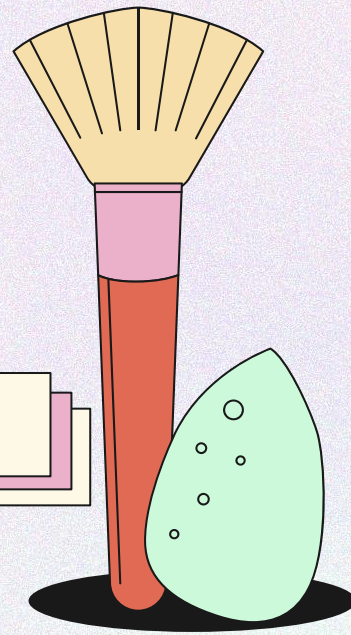
#05

PRÓXIMOS PASOS

#01

INTRODUCCIÓN

Planteamiento del proyecto



OBJETIVO DEL PROYECTO

Entender cómo de avanzada está la regulación respecto la presencia de disruptores endocrinos en diferentes productos de cosmética y maquillaje. Y estudiar la composición de las formulaciones para entender esto.



¿Qué son los disruptores endocrinos?

Los disruptores endocrinos son sustancias químicas que pueden interferir con el sistema hormonal y afectar a la salud a largo plazo. Dado el uso diario en productos cosméticos, evaluar su seguridad es fundamental.

Hipótesis

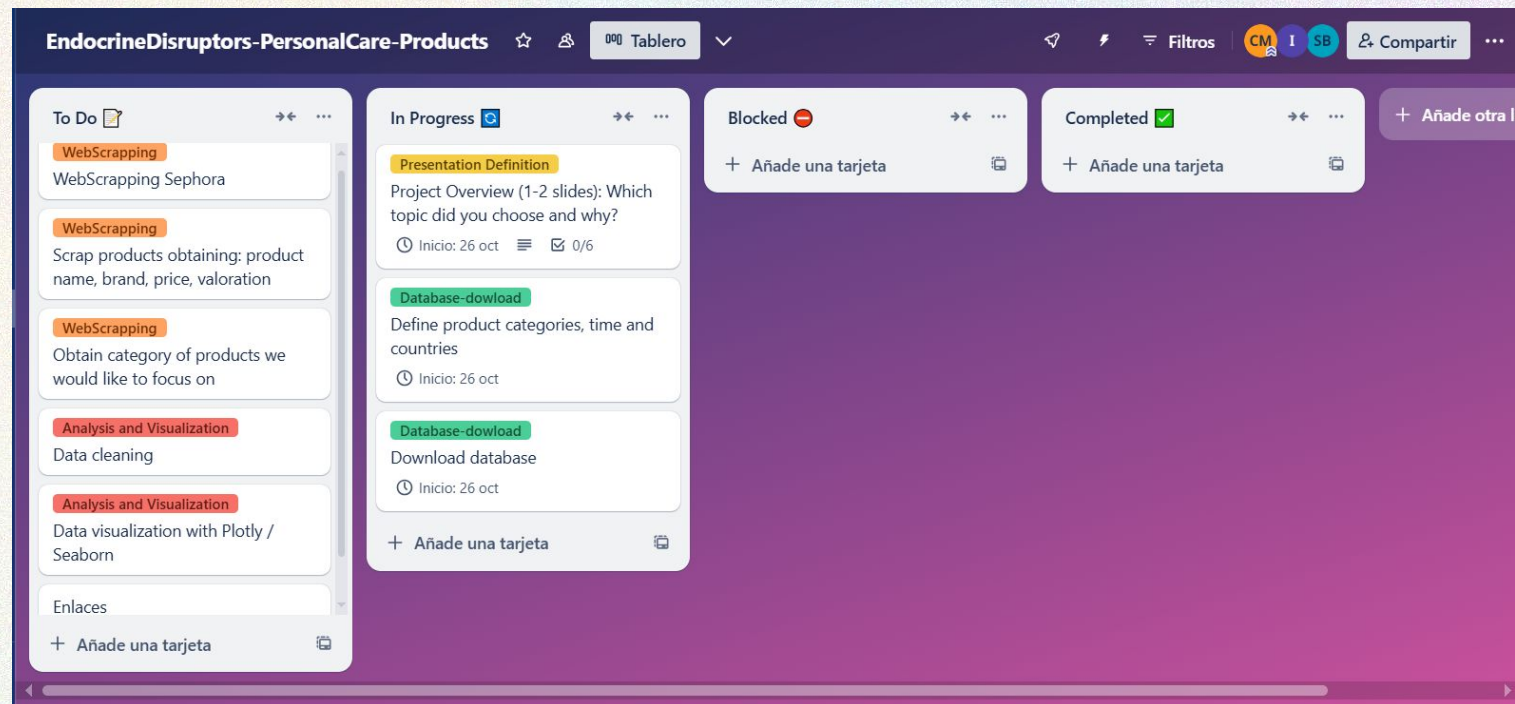
Hipótesis

1. La cantidad de productos con disruptores endocrinos no varía dependiendo del tipo de producto.
2. Las categorías y marcas con mayor uso o mayores aditivos (Ej:make-up será la categoría con más frecuencia) presentan una mayor frecuencia de productos que contienen disruptores endocrinos en comparación con otras.
3. Existe una relación inversa entre el precio del producto y la cantidad de disruptores endocrinos que contiene; los productos más caros contienen menos disruptores que los más económicos.
4. Los consumidores no son conscientes de la presencia de disruptores endocrinos en los productos.
5. Los disruptores endocrinos más comunes difieren entre categorías, existen característicos de cada categoría.

Pregunta

1. ¿Cual es N° de productos por categoría y subcategoría que contienen disruptores endocrinos?
2. ¿Cuáles son las categorías y marcas con más disruptores endocrinos?
3. ¿Tiene alguna relación el valor económico con con el n° de disruptores endocrinos?¿Las marcas más caras contienen menos disruptores y las más baratas?
4. ¿Es consciente el consumidor de los disruptores endocrinos?¿Lo valoran en las opiniones?¿Teniendo en cuenta la media de opiniones y a partir de esta media para arriba en productos con más de 100 valoraciones, contienen todos disruptores endocrinos y por ende podemos asumir que no tienen conocimiento los usuarios de este tema?
5. ¿Cuál es el top 5 de disruptores endocrinos en esta muestra en cada categoría?¿Hay alguna diferencia?

Punto inicial



Estructuración del proyecto

#01

Obtener los datos

Hacer webscrapping de la pagina de Sephora en españa para obtener los datos y usar la lista oficial de disruptores endocrinos de la ECHA

#02

Limpiar los datos

Limpiar bien la base de datos para poder representar y trabajar cómodamente. Sobre todo foco en la columna de composición de la formulación donde se realizará un "match" con los disruptores endocrinos

#03

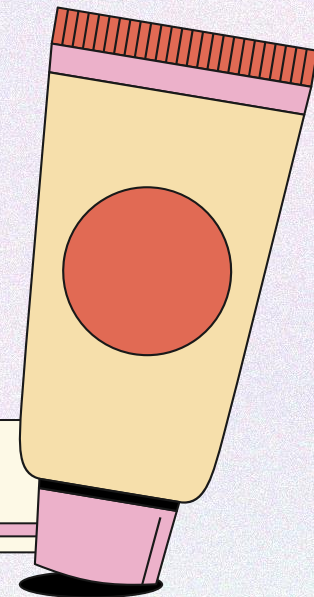
Visualizar los datos

Visualizar datos para extraer conclusiones y verificar nuestras hipótesis

#02

METODOLOGÍA

Explicación paso a paso del proyecto



Extracción de datos-primer obstáculo

1

Web scraping

Delimitación de categorías y web.

2

Beautifulsoup

Primeros intentos con BeautifulSoup

3

Selenium

Cambio a selenium para la obtención de listados de URL's

4

**Beautifulsoup
& request**

Scrapeo del listado de URL's.

+5000 url's

De las cuales analizamos 4318 tras la limpieza

Nuestra base de datos conseguida

#1

Columnas y valores

```
#Check unique values per each column and also column naming  
ingredients_df.nunique()
```

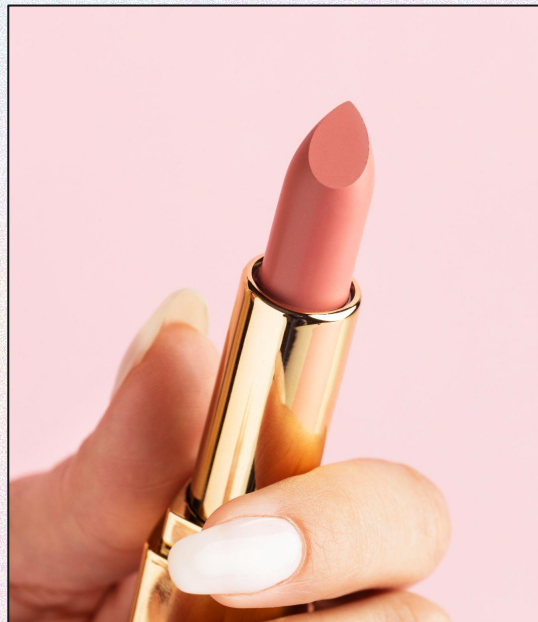
✓ 0.0s

subcategory1	1
subcategory2	7
subcategory3	47
product_name	4535
brand_name	142
rating	33
review_count	1002
price	419
ingredients	4223
product_url	4657
dtype: int64	

142 marcas

7 categorías

47 subcategorías



¡Ahora a limpiar!

Limpieza del dataframe- 3 funciones

#01

LIMPIEZA

```
def cleaning_columns(df):  
    """  
    Cleans and standardizes various columns in the DataFrame:  
    - Removes irrelevant columns and specific values in subcategories.  
    - Translates and normalizes values in category columns.  
    - Cleans and converts prices to numeric format.  
    - Normalizes and standardizes the ingredients column.  
    """
```

#02

MANEJO NULOS

```
# Fill NaN values in certain columns  
df['rating'].fillna(0, inplace=True) # Fill ratings with 'Not Rated'  
df['review_count'].fillna(0, inplace=True) # Fill review count with 0  
df['price'].fillna(df['price'].mean(), inplace=True) # Fill price with mean (or median) value  
df['category'].fillna('Unknown', inplace=True)  
df['subcategory'].fillna('Unknown', inplace=True)  
  
# Replace null values in 'product_name' based on 'brand_name'  
df['product_name'] = np.where(  
    df['product_name'].isna() & df['brand_name'].notna(),  
    'Unknown product from ' + df['brand_name'],  
    df['product_name']  
)  
  
# Drop rows where 'ingredients' is NaN  
df = df.dropna(subset=['ingredients'])
```

#03

MANEJO DUPLICADOS

```
# Remove duplicates and reset index  
df_cleaned = df.drop_duplicates(keep='first').reset_index(drop=True)
```


Momento clave- búsqueda de disruptores

**Búsqueda de disruptores,
identificación y número de
disruptores por producto**

Oh no! La lista de
ingredientes que son
disruptores para la ECHA no
están presente en nuestros
productos...

**Solución: Crear nuestra propia lista basada en
estudios con evidencia científica sobre los
disruptores endocrinos**

Endocrine disruptor assessment list - ECHA

Substance name 	EC / List no 	CAS no 	Authority 	Status 	Outcome 
Terephthalic acid	202-830-0	100-21-0	France	Concluded	not ED
Tributyl O-acetyl citrate	201-067-0	77-90-7	France	Concluded	not ED
p-cresol	203-398-6	106-44-5	United Kingdom	Concluded	not ED
Thiram	205-286-2	137-26-8	Sweden	Concluded	not ED
(1,3,4,5,6,7-hexahydro-1,3-dioxo-2H-isoindol-2-yl)methyl (1R-trans)-2,2-dimethyl-3-(2-methylprop-1-enyl)cyclopropanecarboxylate	214-619-0	1166-46-7	Germany	Under development	Under development (BPR)
1,3,4,6,7,8-hexahydro-4,6,6,7,8,8-hexamethylindeno[5,6-c]pyran	214-946-9	1222-05-5	France	Under development	Under development (other)
1,7,7-trimethyl-3-(phenylmethylene)bicyclo[2.2.1]heptan-2-one	239-139-9	15087-24-8	Germany	Concluded	



Nuestra lista + función buscar E.D.

```
# Function to check if the ingredients contain any endocrine disruptor with fuzzy matching
def find_endocrine_disruptors(ingredient_list):
    """
    Check if the ingredient list contains any endocrine disruptors using fuzzy matching.

    Args:
    | ingredient_list (str): A string of ingredients separated by commas.

    Returns:
    | list: A list of found endocrine disruptors or ["none"] if none are found.
    """
    found_disruptors = [] # List to store found disruptors
    for disruptor in potential_endocrine_disruptors_cosmetics:
        # For each disruptor, look for approximate matches
        if any(fuzz.partial_ratio(disruptor, ingredient) > 95 for ingredient in ingredient_list.split(", ")):
            found_disruptors.append(disruptor) # Add to list if found
    return found_disruptors if found_disruptors else ["none"] # Return list or 'none' if empty

# Apply the function and create the new column
final_ingredients_df['endocrine_disruptors_found'] = final_ingredients_df['ingredients'].apply(find_endocrine_disruptors)

# Create a new column to count the number of disruptors
def count_disruptors(disruptors_list):
    """
    Count the number of endocrine disruptors in the list.

    Args:
    | disruptors_list (list): A list of found endocrine disruptors.

    Returns:
    | int: The number of disruptors found, 0 if "none".
    """
    return 0 if disruptors_list == ["none"] else len(disruptors_list)

# Add the new column with the count of disruptors
final_ingredients_df['number_disruptors'] = final_ingredients_df['endocrine_disruptors_found'].apply(count_disruptors)

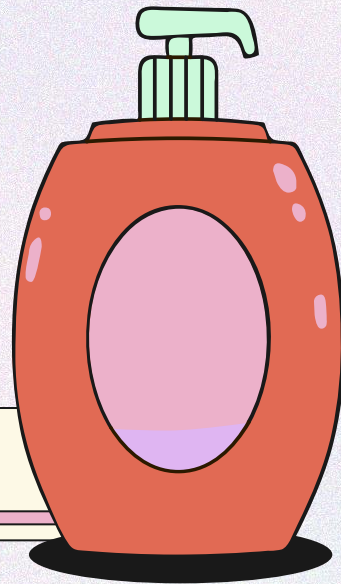
# Display the complete dataframe
final_ingredients_df
```

	INCI name
Parabens	Parabens
	Isopropylparaben
	Isobutylparaben
	Phenylparaben
	Benzylparaben
	Pentylparaben
Benzophenone	Triclosan
	Benzophenone-3
Phthalates	Octocrylene
	Dibutyl Phthalate (DBP)
	Diethyl Phthalate (DEP)
	Dimethyl Phthalate (DMP)
	Diisononyl Phthalate (DINP)
PFAs	"flouro"
	"perflouro"
	"polyflouro"
	"perfluorinated"
Heavy Metals	Lead
	Mercury
	Cadmium
	Arsenic
	Nickel
	Copper
	Zinc
	Barium
	Manganese
Formaldehydes or formaldehyde releasers	Formaldehyde
	Formalin
	Methylene Glycol
	DMDM Hydantoin
	Imidazolidinyl Urea
	Quaternium-15
	Diazolidinyl Urea

#03

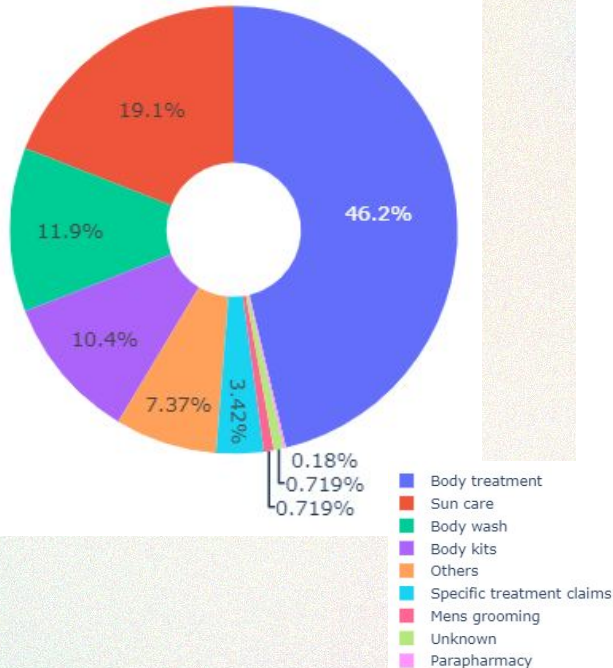
VALIDACIÓN HIPÓTESIS

Respuestas a preguntas clave y visualización de datos

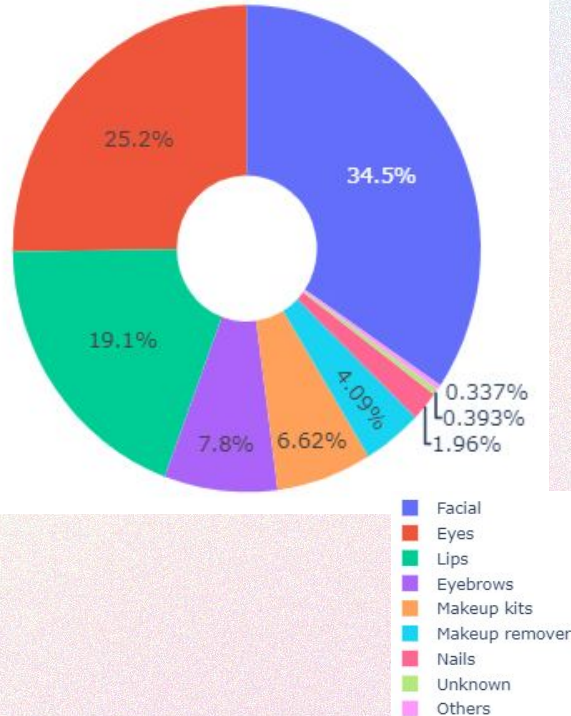


Estructura base de datos por categoría y subcategoría

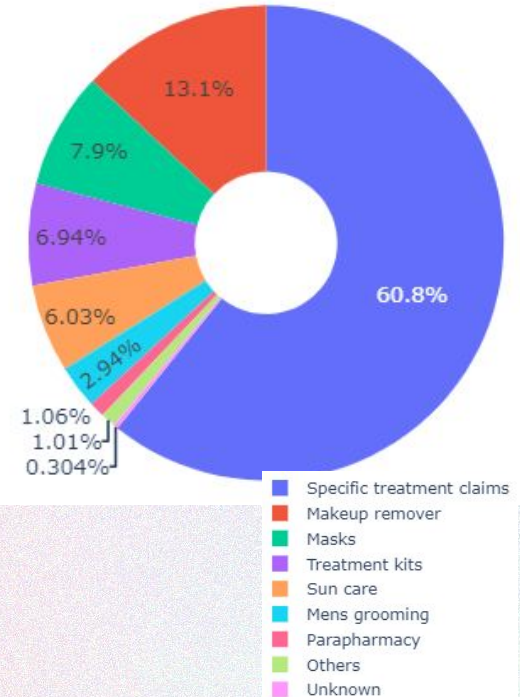
Distribution of Products in Body



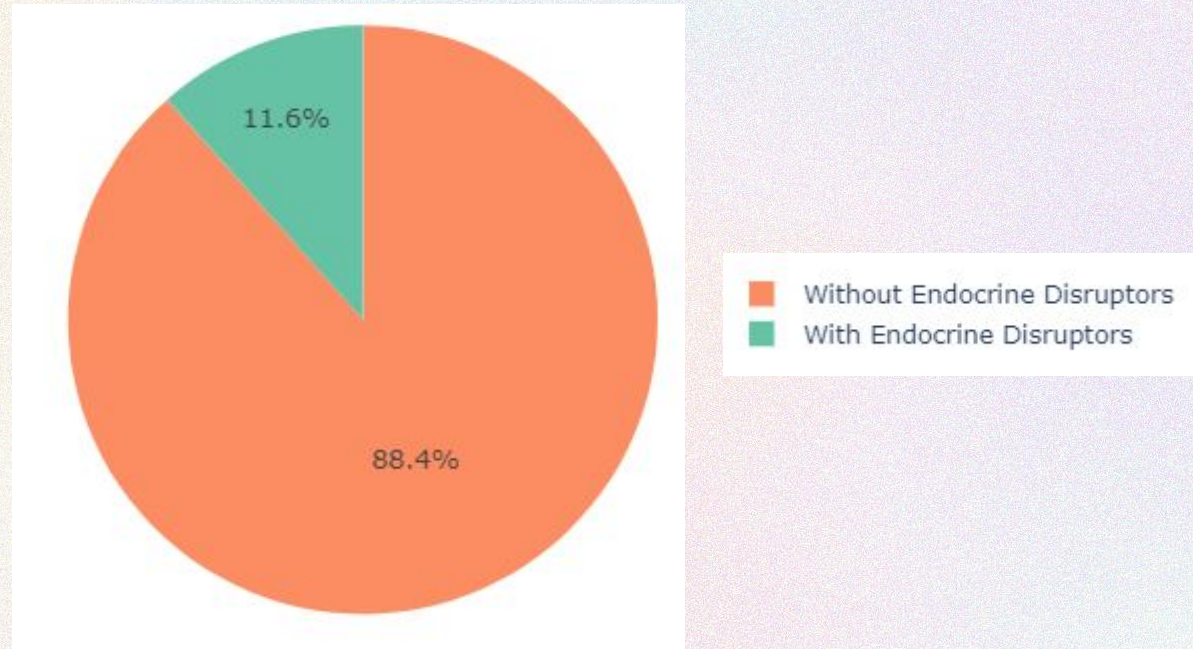
Distribution of Products in Makeup



Distribution of Products in Treatment



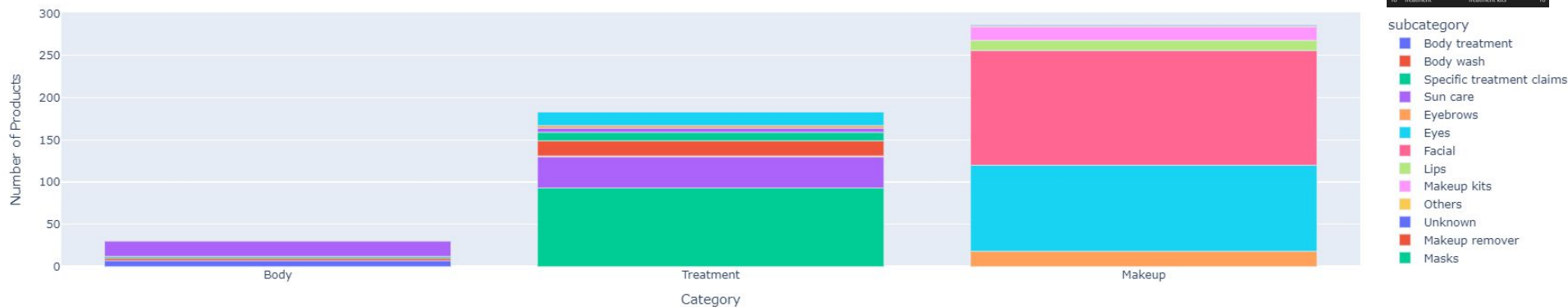
¿Cuántos productos de la muestra tienen disruptores endocrinos?



1. N° productos por categoría y subcategoría que contienen E.D.

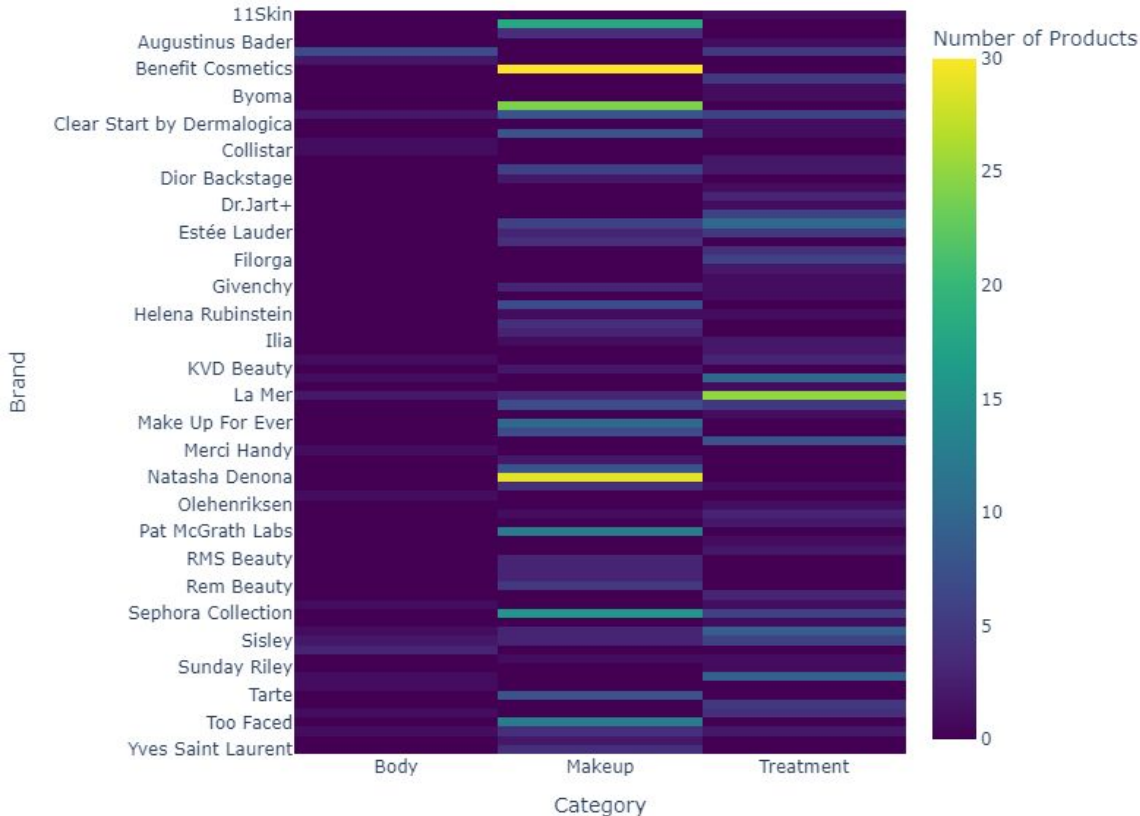
	category	subcategory	n_products
0	Body	Body treatment	7
1	Body	Body wash	3
2	Body	Specific treatment claims	2
3	Body	Sun care	18
4	Makeup	Eyebrows	18
5	Makeup	Eyes	102
6	Makeup	Facial	136
7	Makeup	Lips	12
8	Makeup	Makeup kits	16
9	Makeup	Others	1
10	Makeup	Unknown	1
11	Treatment	Makeup remover	18
12	Treatment	Masks	10
13	Treatment	Mens grooming	5
14	Treatment	Others	1
15	Treatment	Parapharmacy	3
16	Treatment	Specific treatment claims	93
17	Treatment	Sun care	37
18	Treatment	Treatment kits	16

Number of Products with Endocrine Disruptors by Category and Subcategory



2. Categorías y marcas con más disruptores endocrinos

Heatmap of Products with Endocrine Disruptors by Brand and Category



category	brand_name	product_count
Makeup	Benefit Cosmetics	30
Makeup	Natasha Denona	29
Treatment	La Mer	25
Makeup	Charlotte Tilbury	24
Makeup	Anastasia Beverly Hills	18
Makeup	Sephora Collection	16
Makeup	Too Faced	13
Makeup	Pat McGrath Labs	13
Treatment	Erborian	10
Treatment	Kiehl's Since 1851	10

3. Relación valor económico y disruptores endocrinos

Correlation Between Price Range and Average Number of Endocrine Disruptors



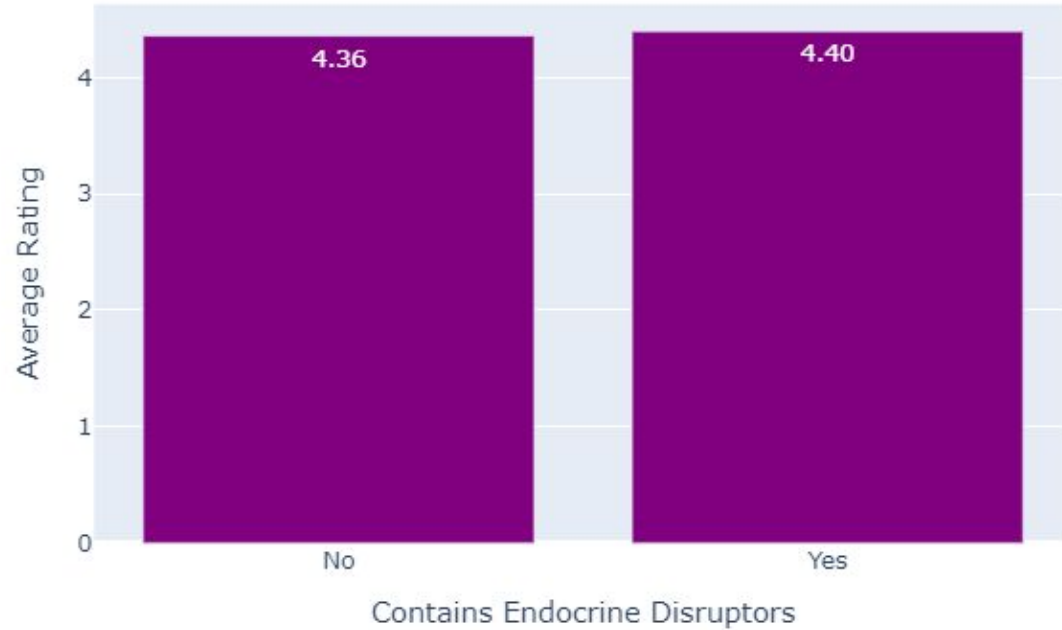
Average number of endocrine disruptors by price range:

price_range	
Low	1.033333
Medium	1.059289
High	1.056000
Very High	1.000000
Luxury	1.000000

4. Conciencia del consumidor sobre disruptores endocrinos

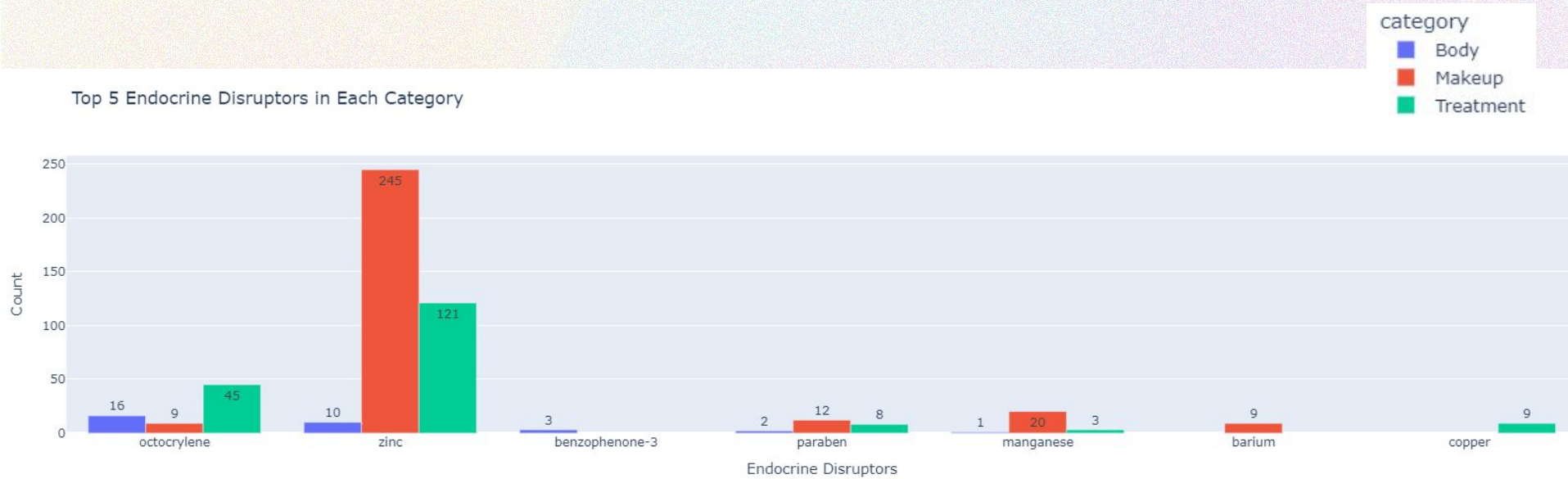
Valorando sólo productos con más de 100 opiniones

Average Ratings of Products with and without Endocrine Disruptors



5. Top 5 disruptores endocrinos por categorías

Top 5 Endocrine Disruptors in Each Category



#04



CONCLUSIONES

¿Qué podemos extraer de todo esto?

Evaluación gestión proyecto y colaboración

The image shows a Kanban board with four columns: To Do, In Progress, Blocked, and Completed. Each column contains task cards with progress bars, titles, and due dates. The 'In Progress' column has a white border around the first card.

- To Do** (Icon: 📄)
 - LINKS (Icon: 🔗)
 - ☰
 - QUESTIONS (Icon: 📝)
 - ☑ 0/5
 - + Añade una tarjeta (Icon: 📄)
- In Progress** (Icon: 🔄)
 - Project Overview (1-2 slides): Which topic did you choose and why? (Icon: ✎)
 - 🕒 Inicio: 26 oct ☰ ☑ 0/6
 - Data cleaning
 - 🕒 Inicio: 4 nov
 - Data visualization with Plotly / Seaborn
 - 🕒 Inicio: 29 oct
 - + Añade una tarjeta (Icon: 📄)
- Blocked** (Icon: 🛑)
 - Scrap products obtaining: product name, brand, price, valoration, number of reviews, ingredient list
 - WebScrapping Sephora
 - + Añade una tarjeta (Icon: 📄)
- Completed** (Icon: ✅)
 - Define product categories, time and countries
 - 🕒 Inicio: 26 oct
 - Obtain category of products we would like to focus on
 - ☑ 4/4
 - Download database
 - 🕒 Inicio: 26 oct
 - + Añade una tarjeta (Icon: 📄)

Mayores Retos



Bloqueo del Scrapeo

Extraer todas las URLs por categoría, ya que el programa debía detectar y activar el botón de "mostrar más productos" sin ser detectado como un bot



Comparación con la tabla oficial de la ECHA

La comparación con el listado oficial de disruptores endocrinos fue un gran bache puesto que los datos obtenidos no hacían alusión a disruptores

Conclusiones finales



- De momento, no encontramos una gran cantidad de disruptores endocrinos en la muestra que tomamos. **Sólo un 11.6% → Europa tiene las legislaciones más estrictas en cuanto a químicos, y algunos de los ingredientes listados ya estaban limitados o prohibidos por la regulación REACH**
- En la categoría de **make-up** hemos encontrado **más posibles futuros disruptores endocrinos**, pues analizamos metales pesados utilizados en pigmentos.
- **El precio no se relaciona con la presencia de disruptores endocrinos.** Hay que tener en cuenta que si que influirá en la calidad de los activos, perfume...Pero esto no quiere decir, que el riesgo químico sea menor
- **El consumidor no tiene el conocimiento** sobre la composición química y el riesgo de muchos ingredientes. Y más si hablamos de los disruptores endocrinos, ya que ni siquiera existe aún una clasificación para cosmética→ **Entrará en vigor una nueva CLP regulation en la que sí se va a considerar los disruptores endocrinos como una clase de peligro**
- **Disruptores endocrinos identificados: Octocrylene, Benzophenone-3, Parabenos.** Todos ellos han mostrado alguna evidencia de efectos hormonales, aunque la magnitud y la relevancia de estos efectos pueden variar. **No considerados disruptores: Zinc, Manganeso, Bario, Cobre,** aunque su exposición en exceso puede conllevar otros riesgos para la salud.

#05

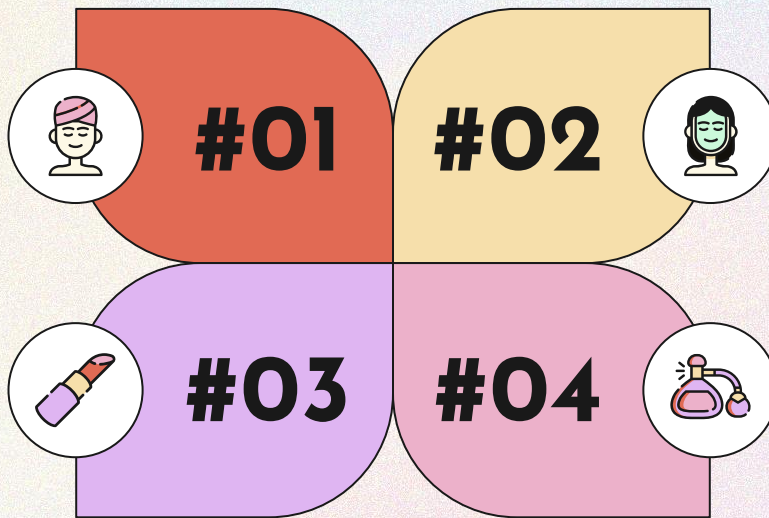
PRÓXIMOS PASOS

Más investigación sobre d.e.

Buscar, informarse y crear una lista de posibles disruptores endocrinos más fundamentada y respaldada por evidencia científica

Más info para el consumidor

Sería interesante crear un webpage o rrss para hablar y divulgar sobre temas relacionados con cosmetica



Realizar revisiones

Próximamente, se buscará establecer un acuerdo en Europa sobre la metodología para identificar las sustancias como disruptores endocrinos. Esto resultará en un incremento de la lista de la ECHA, lo que nos permitirá realizar una revisión y también validar nuestras posibles opciones.

Otros informes

Se podría usar esta misma database de Sephora para otros informes o investigaciones: como alergenicos en fragancia, ver si cumplen REACH.

#GRACIAS

¿Alguna pregunta?

¡Sabemos que puede sonar algo complejo!