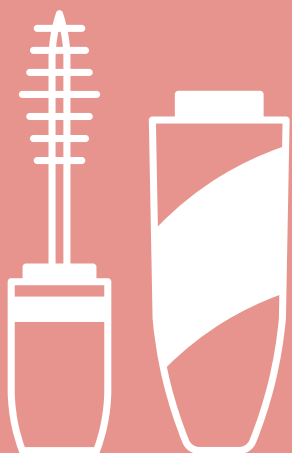


BIG DATA: HOW ACCESSIBLE IS CRUELTY FREE MAKEUP?



Contents:

1. Abstract
2. Initial Ideas
3. Research Contexts
6. Web Scraping: Logical Harmony
7. Web Scraping: Peta
8. Brand Data
9. Accuracy
10. Dissecting Data
11. Overall Results
12. Initial Design
13. Modelling
14. Final Product
- 145 Prototype
16. Limitations/Challenges
17. References



ABSTRACT



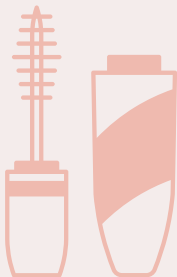
This project investigates the accessibility of cruelty-free makeup on top cosmetic and beauty websites and compares the brands on these websites against cruelty-free product databases. By comparing the cruelty-free products from verified cosmetic brands with the total number of brands listed on cruelty-free websites, this project aims to evaluate the accuracy of cruelty-free databases. In order to conduct this research, web scraping was required to collect data from databases such as PETA and Logical Harmony. This data was then cross-referenced with information from popular cosmetic and beauty websites.

The analysis focuses on creating visual representations of this data, including bar graphs and pie charts, to effectively present the findings. These visualisations allow for a comprehensive comparison between the number of cruelty-free brands and the total brand count on each website. Additionally, the study highlights discrepancies between the databases and the collected website data, providing insights into the reliability of the information.

Findings from this research highlight the beauty industry's commitment to animal testing. The study also discusses the limitations of web scraping and potential sources of error, aiming to provide recommendations for improving the accuracy of cruelty-free databases.

Building upon this research, the project's final design aims to create a unique cosmetic product that visually displays and compares data related to cruelty-free brands. By using colour and logos, the final product provides a clear representation of which beauty brands are making an effort to sell cruelty-free makeup. This integration of cosmetics and data visualisation aims to raise awareness about cruelty-free options and empower consumers to make more informed, ethical choices.

Throughout the project, documentation will showcase the research process, data collection and analysis methods, visual representations of the findings, and reflections on the implications for the beauty industry and animal welfare.



INITIAL IDEAS

For my formative assignment I originally set out to do a completely different project, and wanted to create a a projection surrounding data on the moon. Although this topic is something I find interesting, I quickly found the module limiting and wasn't sure if the final projects was something I wanted to do.

I liked the idea of using an API as I have used them previously and I wasn't too confident in web scraping. However, I realised I needed to push myself and try something different or I was not going to achieve something I was interested in.

Big Data Idea: A Visual Exploration of the Lunar Cycle in relation to Social and Cultural practice

By Shaan Rehshi



Project Premise

- The project aims to explore the lunar cycle by integrating big data from the moon and its various phases. The project will involve collecting and analysing large amounts of data, including the phase, zodiac sign, age, illumination, angle, distance, and visibility of the moon, obtained through moon and weather APIs. The data will be processed and visualised using projection mapping techniques to create an immersive and interactive experience. An additional idea would be to integrate constellations into the project relating to the visibility of the sky.
- The end product will be a web application that allows users to input any date over a long set of years such as their birthdate and generate the corresponding moon phase projection, along with detailed information about the moon's properties for that date.
- This project meets the brief by integrating big data from natural events to create a unique and engaging installation that educates and informs users about the moon's cycle. The use of data scraping and projection mapping techniques also showcases technical skills that align with the goals of the brief. Additionally, the interactive web application provides an accessible way for users to explore and learn about the moon's cycle, making it a highly engaging and informative project.

I then moved onto an idea of sentiment analysis and looked at using Reddit API to gather data on peoples opinions. I thought about looking at feminism, film and music as these were all topic that I had an interest in. I managed to scrape data on hot posts and certain sub redדים, however, I wasn't sure how I would present this information.



```
from textblob import TextBlob
import pandas as pd
from textblob.sentiments import NaiveBayesAnalyzer
from afinn import Affinn
```

```
In [2]: # download stopwords if not already downloaded
# nltk.download('stopwords')
```

```
In [3]: # Load text data
text = "This is a sample text to perform sentiment analysis."
```

```
In [4]: # preprocess text data
stop_words = set(stopwords.words('english'))
words = nltk.word_tokenize(text.lower())
filtered_words = [word for word in words if word not in stop_words and word.isalpha()]
```

```
In [5]: # perform sentiment analysis
blob = TextBlob(' '.join(filtered_words))
polarity = blob.sentiment.polarity
subjectivity = blob.sentiment.subjectivity
```

```
# print sentiment scores
print(f"Polarity score: {polarity:.2f}")
print(f"Subjectivity score: {subjectivity:.2f}")
```

```
Polarity score: 0.00
Subjectivity score: 0.00
```

```
In [6]: reddit = praw.Reddit(client_id='KfelTKUsMTd0s0vSqBvmw', client_secret='STqurafuzAlfyi1X9G9H9snnqvMp_A', user_agent='test')
```

```
In [9]: hot_posts = reddit.subreddit('crueltyfree').hot(limit=10)
for post in hot_posts:
    print(post.title)
```

```
I just thought of something. I've always thought of animal testing as outdated science. But why is it even still an option?
```

```
In [30]: url = "https://sephora.p.rapidapi.com/products/list"
```

```
In [31]: querystring = {"categoryId": "cat150006", "pageSize": "70", "currentPage": "1"}
```

```
In [32]: headers = {
    "X-RapidAPI-Key": "9c2ded9108msha5090905dfe96dp152c1ajsnc4c239548328",
    "X-RapidAPI-Host": "sephora.p.rapidapi.com"
}
```

```
In [33]: response = requests.get(url, headers=headers, params=querystring)
```

```
In [34]: # Create DataFrame from JSON response
df = pd.DataFrame.from_records(response.json()[ 'products' ])
df.to_csv('sephora_products.csv', index=False)
# Display first 5 rows of DataFrame
```

```
{ 'categoryId': 'cat150006', 'content': { 'bestsellersTargetURL': '/best-selling-skin-care', 'justArrivedTargetURL': '/new-skin-care-products', 'region3': { { 'componentName': 'Sephora Link Group Component', 'componentType': 9, 'enableTesting': False, 'links': [ { { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Skin care Quiz', 'name': 'app_link_skincare_quiz', 'targetScreen': { 'targetScreen': 'contentstore', 'targetValue': '50300057' }, 'testEnabled': False }, { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Luxury Skincare', 'name': 'app_link_luxe_skincare', 'targetScreen': { 'targetScreen': 'contentstore', 'targetValue': '102100018' }, 'testEnabled': False }, { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Skincare for Life', 'name': 'app_link_skincareforlife', 'targetScreen': { 'targetScreen': 'contentstore', 'targetValue': '78900020' }, 'testEnabled': False }, { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Cleanical Skincare', 'name': 'app_link_cleanicalskincare_020923', 'targetScreen': { 'targetScreen': 'contentstore', 'targetValue': '103300019' }, 'testEnabled': False }, { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Skin Routine Quiz', 'name': 'app_link_skin_routine', 'targetScreen': { 'targetScreen': 'contentstore', 'targetValue': '87300017' }, 'testEnabled': False }, { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Best Skincare Under $35', 'name': 'app_link_skin_care_under_35', 'targetScreen': { 'targetScreen': 'contentstore', 'targetValue': '86800020' }, 'testEnabled': False }, { 'componentName': 'Sephora Link Component', 'componentType': 8, 'enableTesting': False, 'linkText': 'Clean Skincare Guide', 'n
```

Finally, after deciding I wanted to do something to do with current events, I started looking at beauty websites and trying to scrape them. I originally just set out to scrape Sephora and found an API to try and so. It gave me some data but it didn't seem accurate and the API didn't give me many options.

RESEARCH CONTEXTS



In the UK, animal testing for cosmetics and their ingredients has been banned since 1998. The primary legislation that governs animal testing in the country is the Animals (Scientific Procedures) Act 1986 ((Rowan, A. N. (2007))). This Act specifically regulates the use of animals for scientific or experimental purposes, which includes testing. In 2023 Cruelty Free International lost a high court challenge regarding allegations that the UK government secretly abandoned the ban on testing cosmetic product ingredients on animals. However, documents revealed that the Home Office had secretly abandoned the ban in 2019, citing a law originating from the European Union that requires testing certain cosmetic ingredients on animals for worker safety (Smith, S. (2023)). Cruelty Free International intends to appeal the high court's decision and urges the government to reinstate the policy ban, emphasizing the need to fight against animal testing in UK laboratories (Smith, S. (2023)).

According to a study conducted by Orbis Research (2018), the global cosmetics industry is valued at least \$805 billion as of 2023 (Cahill, E. (2018)). Despite its explosive growth, some consumers would prefer for beauty brands to cease animal testing and employ ethical labour practices (Baker, 2015).

There has been an increase in “conscious consumerism,” as consumers are becoming more concerned with the origins of their products, which has led to newer beauty brands creating more vegan and/ or cruelty-free products. Consumer behaviour refers to the actions individuals take when acquiring, consuming, and disposing of products and services (Cahill, E. (2018)). It involves decision-making processes regarding when, where, and why consumers make purchases, as well as their post-purchase evaluations. Regarding brand engagement, the research suggests that companies can enhance their engagement and sales by redesigning their business practices to cater to the preferences of Millennials and Generation Z. Key factors include building early loyalty, promoting transparency, and adopting sustainable approaches (Leon, D. I. B. (2020)).

Vegan makeup brand launches alone grew by 175% from 2013 to 2018 and have become more successful with younger consumers (McDougall, 2018). In 2017, the cruelty-free makeup market has been expected to grow by 6.1% by the end of 2023 (Whitehouse, 2017). Despite the progress that has been made, many makeup companies still choose to sell their products in China, one of the most profitable global makeup markets, whose laws require cosmetic products to be tested on animals to be sold (Chitrakorn, 2016). As larger international beauty brands choose between ethics and profits, smaller makeup brands in the United States have arisen to attempt to meet the needs of conscious consumers by going cruelty-free. Smaller makeup brands in the United States have emerged to cater to conscious consumers by prioritizing ethics over profits, while larger international beauty brands face the challenge of balancing ethical considerations with financial gain. These smaller brands have opted to go cruelty-free in order to meet the growing demand for ethically produced cosmetics (Kulkarni, N.).

A study conducted in 2020 examined whether cosmetic companies can increase their brand engagement when implementing a cruelty-free approach to designing their products for the 18 to 24 age group . The findings suggest that businesses can definitely profit by implementing a sustainable and cruelty-free approach. Companies that integrate sustainability into their core values, support environmental causes, use organic products, and engage with local communities have managed to target Millennials and Generation Z successfully, resulting in increased profits. Examples of such successful companies include Lush Cosmetics, Patagonia, Unilever, and IKEA (Leon, D. I. B. (2020)). Additionally, the research indicates that cosmetic companies can increase brand engagement among the 18-24 age group by adopting a cruelty-free approach. Overall, the research concludes that there is a significant growth in the availability of cruelty-free and sustainable products for the 18-24 age range. Millennials and Generation Z are challenging traditional business practices, demanding sustainability, and driving changes in the global economy, therefore, companies that prioritise sustainability, cruelty-free practices, organic products, and community support will successfully target these two generations, resulting in increased profitability (Leon, D. I. B. (2020)).

Nath (2016) and Morgan (2018) highlight that consumers are becoming more conscious of animal testing and also care about how employees are treated within brands. Fromm (2016) suggests that some brands manipulate consumers through their strategies, providing only partial information about their products' production, distribution, and testing. Rose (2019) adds that certain brands operate ethically in some regions but still engage in animal testing in countries where it is required. As a result, consumers have become more conscious of regulations and are demanding transparency from companies.

RESEARCH CONTEXTS



The Cruelty-Free Pseudo-Event

Cosmetic brands can claim to be “cruelty-free,” but there are no official laws in the US regarding how to label cruelty-free products (“Cruelty-free labeling, n.d.), meaning the term cruelty-free can vary significantly across c brands, and brands are able to determine what cruelty-free means without consumers know what they are purchasing. According to "Cruelty-free labeling" (n.d.), the term "cruelty-free" often carries the expectation among consumers that no animals were harmed in any stage of the supply or manufacturing chain. However, this term can be misleading as even brands claiming to be cruelty-free may still source ingredients that have been tested on animals or rely on previous animal testing results, as well as engage in animal testing overseas. Moreover, the labelling of a brand as "cruelty-free" does not necessarily guarantee ethical practices in other areas, such as labour conditions involving sweatshops or child labor, or the creation of harmful environmental impacts through factory operations and shipping practices ("Cruelty-free labeling," n.d.). The term “cruelty-free” is a pseudo-environment (Lippmann, 2004) used to create a positive image around a brand that encourages consumers to buy the products. Because of this, brands can capitalise on a singular aspect of cruelty-free, such as no animal testing, to gain popularity and loyal customers even if the rest of the brand is unethical.

What Drives Cruelty-Free Choices?

A study from 2019 surveyed 108 female participants, many of whom expressed increased awareness and interest in the topic of cruelty-free cosmetics. It was noted that the participants ranked price as the second most important factor when purchasing cosmetics, indicating that financial considerations were more significant to them than ethical values. However, 60% of the participants still had a positive likelihood of purchasing cruelty-free cosmetics in the future, indicating that the concept is still being embraced by consumers who are learning about it (Alaouir, T., Gustavsson, R., & Schmidt, N. (2019)). The study found a strong positive correlation between attitude and purchase intention. Participants with a positive attitude towards cruelty-free cosmetics were more likely to have the intention to purchase them. It is important to note that a positive attitude does not guarantee actual purchase behavior, as there may be an attitude-behaviour gap. Additionally, the study aimed to explore the influence of altruism on purchase intention. However, the analysis did not support a positive effect of altruism on purchase intention. Nevertheless, respondents expressed concerns for animal welfare and the environment as motivations for purchasing cruelty-free cosmetics, supporting the idea that ethical considerations play a role in consumer behavior (Alaouir, T., Gustavsson, R., & Schmidt, N. (2019)).

Solutions

In order for consumers to combat challenges surrounding cruelty-free labeling, combining reflexive law and false advertising law to standardise the labeling of cosmetics is a solution (Winders, D. J. (2006)). This approach aims to address the lack of official guidelines and varying interpretations of cruelty-free claims in the industry. By incorporating reflexive law, which encourages self-regulation and transparency, along with false advertising law, which focuses on preventing deceptive marketing, this approach seeks to establish consistent and reliable labeling practices for cruelty-free cosmetics (Winders, D. J. (2006)). This proposal aligns with the concerns raised about the ambiguity surrounding the meaning of "cruelty-free" and the potential differences between claims and actual practices. Implementing standardised labeling practices would provide consumers with the necessary information to make informed choices, supporting their desire to purchase products aligned with their ethical values.

My Findings

In my research and final project, I intend to explore the various factors and studies discussed and examine if similar challenges exist regarding the accessibility and reliability of brands claiming to be cruelty-free. Moreover, my project aims to assist consumers in becoming more aware and encouraging them to conduct their own research. By investigating the issues surrounding cruelty-free labelling and the potential inconsistencies between claims and actual practices, I hope to contribute to the overall understanding of this topic.

WEB SCRAPING: LOGICAL HARMONY



```
In [2]: # import Libraries
import requests
from bs4 import BeautifulSoup

import functools as ft
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import re
```

My first step was to scrape a website with information on cruelty free and vegan products, I decided to use Logical Harmony to begin with. I used BeautifulSoup and Requests and looped through the website every time there was an instance of entry-content, as this was where the brands were listed in the cruelty free brand list. Once I gathered the amount of brands returned, which was 934, I stored the results into a csv data file.

```
In [3]: r = requests.get("https://logicalharmony.net/cruelty-free-brand-list/")
soup = BeautifulSoup(r.text, 'html.parser')

# create an empty List that will store brands
brands = []

# create another List that will contain each List item still with the tags
rows = []

# get all List tags in entry-content part of the page (found by inspecting site)
brand_section1 = soup.find(class_='entry-content')
for brand in brand_section1.find_all('li'):
    # make sure brands are being extracted correctly
    # print(brand.text)
    # append each brand to the List of brands
    brands.append(brand.text)
    # keep the tags info in this List
    rows.append(brand)

len(brands)

Out[3]: 934
```

```
In [20]: brand_df.loc[brand_df.Partial_vegan, 'Brand'] = partial_veg_brands
brand_df[brand_df['Brand'].str.contains(re.escape('('))]
```

Out[20]:

	Brand	All_vegan	Partial_vegan	Bad_parent_company	Black_owned
84	Balm, The (theBalm)	False	False	False	False
139	Bossy Cosmetics (Bossy Lipstick)	False	False	False	True
246	Earth Friendly Products (ECOS)	False	False	False	False
256	ELF (Eyes Lips Face/E.L.F.)	True	False	False	False
267	EOS (Evolution of Smooth)	False	False	False	False
363	Hero Cosmetics (also known as Mighty Patch)	False	False	False	False
465	La Couleur Couture (LACC)	True	False	False	False
485	LH Cosmetics (Linda Hallberg Cosmetics)	False	False	False	False
570	MOMMA (Ministry of Mineral Makeup Australia)	True	False	False	False
577	MOX (Moxe Aromatherapy)	False	False	False	False
639	OOO Polish (Triple O Polish)	True	False	False	True
648	Orchard View Naturals (OV Naturals)	True	False	False	False
678	Personna (shaving razors)	False	False	False	False
715	R + Co (R and Co)	True	False	False	False

```
In [3]: # create dataframe - start with just brands list, then add more info
df = pd.DataFrame({'Brand_info': brands})
df.to_csv('C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/merged_brands.csv', index=False)
df.head()
```

Out[3]:

	Brand_info
0	If a brand is completely vegan "100% Vegan" ne...
1	It is noted if a brand is cruelty-free but is ...
2	To help make shopping easier, I try to link to...
3	100% Pure
4	100BON - 100% vegan

Logical Harmony also tell users other information such as: vegan, partially vegan, Black owned, and products with bad parent company - this tell us that the brand itself is cruelty free but the brand that company isn't. I then updated the data frame and also created pie charts on my findings.

My original plan, was to focus primarily on comparing the brands available on Sephora, given its status as the second most popular cosmetic website globally and the fourth most popular in the UK according to similarweb.com. However, to ensure a more comprehensive analysis and broaden the scope of the research, I decided to expand my data collection to include three additional websites: Cult Beauty, Beauty Bay, and Ulta. The inclusion of these websites was motivated by Logical Harmony, which provides data specifically for these platforms.

```
# merge them all with merge_ordered
merge2 = ft.reduce(lambda left, right: pd.merge_ordered(left, right, on=['Brand', 'All_vegan', 'Partial_vegan', 'Bad_parent_company', 'Black_owned', 'Ultr', 'Sephora', 'Beauty_Bay', 'Cult_Beauty'], index=False))
merge2.to_csv('C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/merged_brands.csv', index=False)
merge2.head(30)
```

Out[3]:

	Brand	All_vegan	Partial_vegan	Bad_parent_company	Black_owned	Ultr	Sephora	Beauty_Bay	Cult_Beauty
0	100% Pure	False	False	False	False	NaN	NaN	NaN	NaN
1	100BON	True	False	False	False	NaN	NaN	NaN	NaN
2	1121 Apothecary	True	False	False	False	NaN	NaN	NaN	NaN
3	21 Drops	True	False	False	False	NaN	NaN	NaN	NaN
4	2B	False	False	False	False	NaN	NaN	NaN	NaN
5	2nd KIND	True	False	False	False	NaN	NaN	NaN	NaN
6	7 Virtues, The	True	False	False	False	NaN	1.0	NaN	NaN
7	8 Greens	False	False	False	False	NaN	1.0	NaN	NaN

I used the same method to gather all the brands, instead looping through each page dedicated to a brand, and search for words such as vegan. If a brand is available on a website it is marked at 1.0 in the data frame. I had now gathered all data required on this website.

WEB SCRAPING: PETA

During the research process, I discovered that the information on Logical Harmony's website had not been updated since 2019. Recognising the importance of reliable and up-to-date data, I decided to incorporate a second database to ensure the accuracy of my findings. To achieve this, I chose to scrape data from PETA, which explicitly states its database is updated twice a year.

By including PETA's database and comparing the information collected from both sources, I aimed to minimise anomalies and discrepancies in the results. Utilising the same web scraping methodology employed for Logical Harmony, I conducted the data collection process across the same websites.

This approach not only allowed for a more robust evaluation of the accessibility of cruelty-free makeup but also provided an opportunity to validate and cross-reference the information between the two databases. By considering multiple reliable sources, I aimed to enhance the overall reliability and accuracy of my analysis.

I found it harder to scrape the contents of PETA, I had many failed attempts, however, I managed to find a way also using a loop as seen in the code on the right. Overall the length of the brands scraped was over 6000, proving there was more data to use. Once I collected all the brands, I cleaned the data and put it into a data frame as I had done previously.

```
# Send a GET request to the PETA page listing brands that don't test on animals
url = 'https://crueltyfree.peta.org/companies-dont-test/?per_page=all'
response = requests.get(url)

# Create a BeautifulSoup object to parse the HTML content
soup = BeautifulSoup(response.text, 'html.parser')

# Find the unordered list containing the brands
list_container = soup.find('ul', {'class': 'search-results'})

# Initialize an empty list to store the brands
brands = []

# Iterate over the list items
for item in list_container.find_all('li'):
    # Extract the brand name
    brand = item.text.strip()
    brands.append(brand)

# Print the list of brands that don't test on animals
print(brands)
```

[illegible]

```
all_peta_df = pd.concat([vegan_peta_df, peta_df], ignore_index=True)

# Drop duplicates based on the 'Brand_info' column
all_peta_df = all_peta_df.drop_duplicates(subset='Brand_info')

# Save the merged DataFrame as all_peta.csv
all_peta_df.to_csv("C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/all_peta.csv", index=False)

# Print the merged DataFrame
print(all_peta_df)
```

```

      Brand_info
0      OrganicaLibra
1      1121 Apothecary
2      L
3      180 Innovations
4      SanMelix Laboratories
...      ...
9639      Zoo On Yoo
9640      Zorah biocosmétiques
9641      Zosimos Botanicals LLC
9643      Zuii Organic
9644      Zuzu Cosmetics

```

```
[6263 rows x 1 columns]
```



BRAND DATA

After scraping the previous two websites, I started to scrape each individual cosmetic website in order to scrape their total brands in comparison to the verified cruelty-free sites on their websites.

```
In [23]: # Send a GET request to the website and retrieve the HTML content
url = "https://www.cultbeauty.com/brands.list"
response = requests.get(url)
html_content = response.text

# Create a BeautifulSoup object to parse the HTML content
soup = BeautifulSoup(html_content, "html.parser")

# Find all the 'a' tags within the HTML content that contain the brand names
brand_tags = soup.find_all("a", attrs={"class": "responsiveBrandsPageScroll_brand"})

# Extract the brand names from the 'a' tags
brands = [tag.text.strip() for tag in brand_tags]

# Print the brand names
print(brands)

# Create a DataFrame from the extracted data
data = {"Brand_info": brands}
all_brands_cultbeauty_df = pd.DataFrame(data)

# Save the DataFrame as a CSV file
all_brands_cultbeauty_df.to_csv("C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/all_brands_cultbeauty.csv", index=False)

# Print the DataFrame
print(all_brands_cultbeauty_df)

['111SKIN', '19/99 Beauty', 'ABOUT-FACE', 'Advanced Nutrition Programme', 'Aesop', 'Alfresco', 'Algenist', 'Allies of Skin', 'Alpha-H', 'Ameliorate', 'Amika', 'Anastasia Beverly Hills', 'Anthony', 'Augustinus Bader', 'Aveda', 'BAD NORWEGIAN', 'Bala Bangles', 'bareMinerals', 'Baübo', 'Baxter of California', 'BeautyBio', 'BEAUTYBLENDER', 'BeautyStat Cosmetics', 'Bed of Nails', 'Bella Freud', 'Benefit', 'BeYou', 'Bio:Ionix', 'BIOEFFECT', 'Biossance', 'BLEACH London', 'Bondi Sands', 'Bond No.9', 'Bouclème', 'Boy Smells', 'BREAD BEAUTY SUPPLY', 'Briogeo', 'Buttah Skin', 'BYOMA', 'BYREDO', 'BY TERRY', 'Caroline Hirons',
```

After collecting the necessary data, I compared the brand databases and the data frames from Logical Harmony and PETA databases. The aim was to identify brands that appeared in both data frames, indicating their verification as cruelty-free. Subsequently, a new column was added to the data frame, marking these brands as "True."

```
# Load data from CSV files
all_brands_cultbeauty_df = pd.read_csv('C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/cruelty_free_cultbeauty_df')
cruelty_free_brands = pd.read_csv('C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSet

# Extract brand names from each DataFrame and clean them
all_brands = all_brands_cultbeauty_df['Brand_info'].str.strip().str.lower().tolist()
cruelty_free_brands = cruelty_free_cultbeauty_df['Brand_info'].str.strip().str.lower().tolist()

# Create a DataFrame to store the results
results_df = pd.DataFrame({'Brand_info': all_brands})
results_df['cruelty_free'] = results_df['Brand_info'].str.lower().isin(cruelty_free_brands)

# Print the results
print(results_df.head(40))

Brand_info  cruelty_free
0          111skin         False
1      19/99 beauty         False
2      about-face         False
3  advanced nutrition programme  False
4          aesop         False
5        alfresco         False
6        algenist         False
7  allies of skin         False
8        alpha-h         False
9      ameliorate         False
```

```
# Load data from CSV files
all_peta_df = pd.read_csv('C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/all_peta.csv')
all_brands_cultbeauty_df = pd.read_csv('C:/Users/shaan/Desktop/big_data/bsc-big-data/Project/DataSets/all_brands_cultbeauty.c

# Convert brand names to lowercase for case-insensitive comparison
all_peta_df['Brand_info'] = all_peta_df['Brand_info'].str.lower()
all_brands_cultbeauty_df['Brand_info'] = all_brands_cultbeauty_df['Brand_info'].str.lower()

# Add a column to store the result of cruelty-free check
all_brands_cultbeauty_df['cruelty_free'] = all_brands_cultbeauty_df['Brand_info'].isin(all_peta_df['Brand_info'])

# Calculate the number of cruelty-free products
num_cruelty_free = all_brands_cultbeauty_df['cruelty_free'].sum()

# Calculate the percentage of cruelty-free products out of all products
percentage_cruelty_free = (num_cruelty_free / len(all_brands_cultbeauty_df)) * 100

# Calculate the percentage of non-cruelty-free products out of all products
percentage_non_cruelty_free = 100 - percentage_cruelty_free

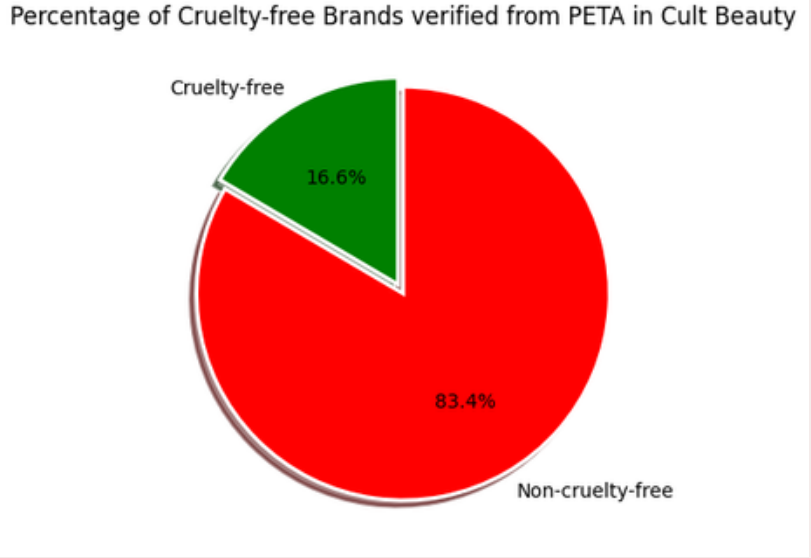
# Print the updated DataFrame
print(all_brands_cultbeauty_df.head(50))

# Print the number of cruelty-free products and their percentages
print(f"Number of cruelty-free products: {num_cruelty_free}")
print(f"Percentage of cruelty-free products: {percentage_cruelty_free:.2f}%")
print(f"Percentage of non-cruelty-free products: {percentage_non_cruelty_free:.2f}%")

Brand_info  cruelty_free
0          111skin         False
1      19/99 beauty         False
2      about-face         False
3  advanced nutrition programme  False
```

The final data frame represents the most accurate results regarding the percentage of cruelty-free brands on each website. To visually present the findings, I created pie charts to depict the distribution of cruelty-free brands among the websites.

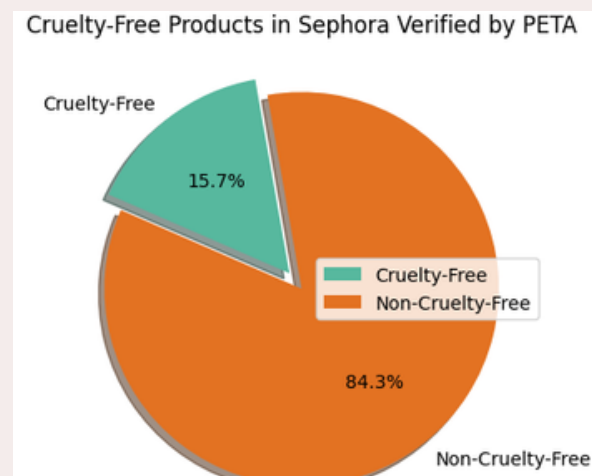
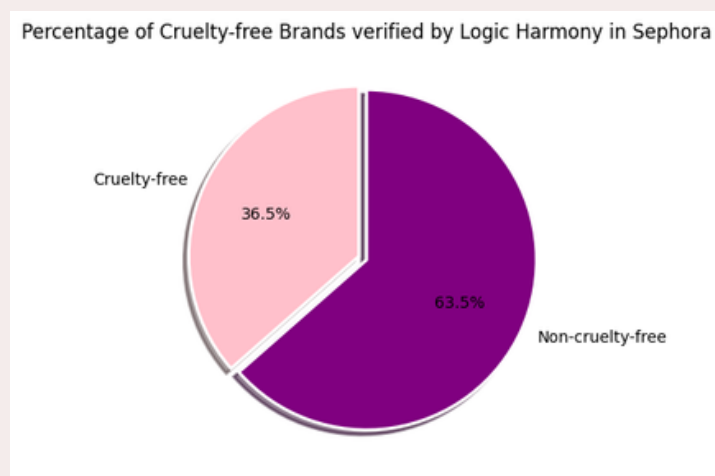
This approach ensures a comprehensive evaluation of brand verification across multiple sources and provides a clear representation of the proportion of cruelty-free brands on each website. The pie charts offer an easily interpretable visual summary of the data, enabling effective communication of the research outcomes. I continued this process for the other three websites.



ACCURACY

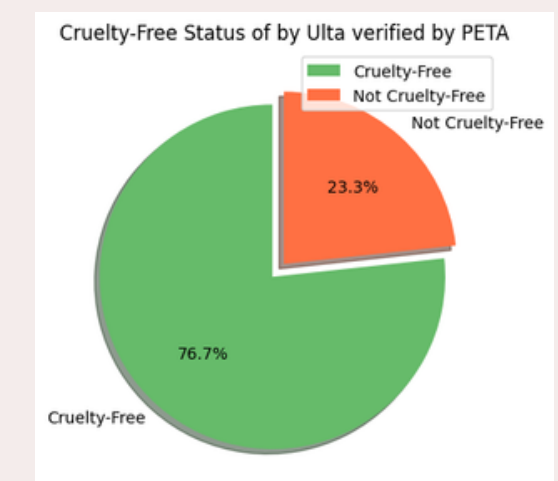
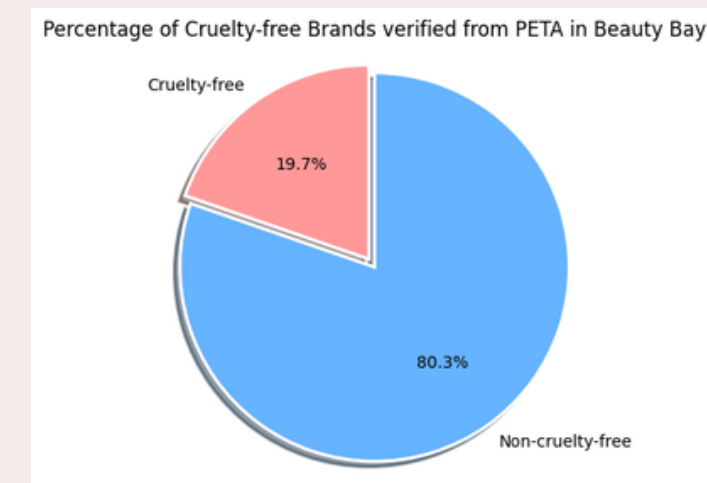
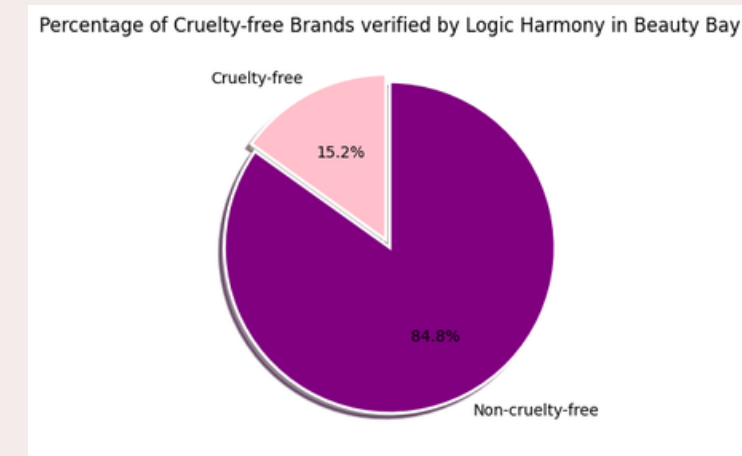
Upon conducting a thorough comparison and analysis of the collected data, I proceeded to examine the findings for any discernible trends or patterns. One notable discovery was the difference in the data between the cruelty-free databases for the brand data on Sephora. Logical Harmony's data indicated a higher number of cruelty-free products compared to PETA's data. This finding raises the possibility that Sephora may have removed several cruelty-free brands from their offerings since 2019, resulting in a decline in their accessibility over the past four years. Another potential explanation could be that certain brands have undergone changes, and newer databases might not have updated information on these specific brands.

Interestingly, the results obtained from PETA's database showcased relatively consistent outcomes across all brands. This suggests that PETA's database may be more accurate and up-to-date in the present context. Overall, the large contrast in cruelty-free results between Sephora and the other databases indicates that the data for the other websites align more closely with each other, showing greater credibility and reliability to their accuracy.



Here you can see the large contrast in Sephora's data

The data in these graphs seem relatively similar, Cult Beauty has a fairly small set of data from Logical Harmony, However, the results on PETA's databases are fairly small too, Overall you can see there isn't much change in the graphs

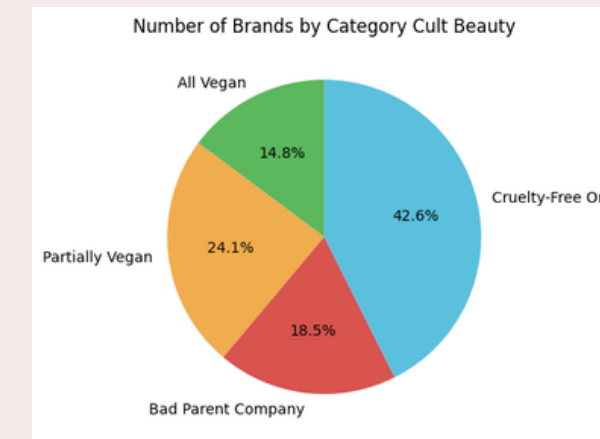


DISSECTING DATA

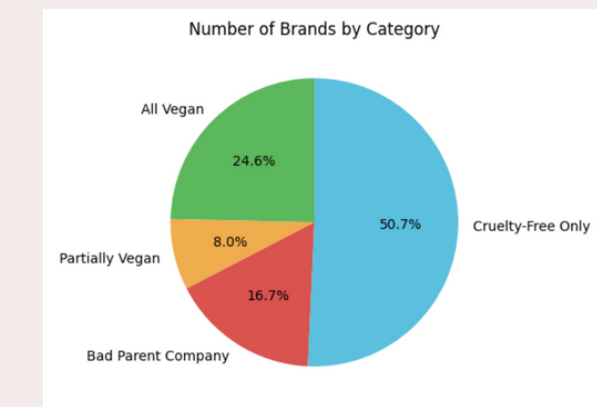
During my research, I discovered that Logical Harmony provides additional categories such as "vegan products" and "bad parent companies". To make my cosmetic product more engaging and informative, I decided to compare as many graphs as possible in order to present a comprehensive view of the data.

One of the intriguing aspects I explored was the comparison between cruelty-free products and products that were not only cruelty-free but also vegan. The results were surprising, as they revealed a relatively low percentage of products that met both criteria. Despite many brands avoiding animal testing, less than a quarter of all four of these websites' brands actually contained no animal-derived ingredients.

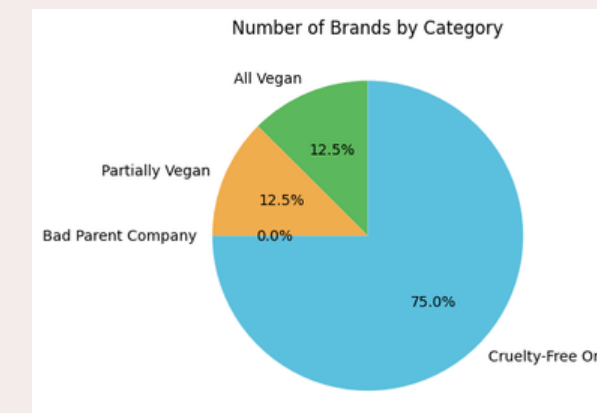
These findings shed light on an important aspect of the cosmetics industry, highlighting the distinction between cruelty-free and truly vegan products. While companies may abstain from testing on animals, the incorporation of animal-free ingredients remains limited. By visualising this information through pie charts, my cosmetic product will convey this crucial message to users, encouraging them to make more informed choices and supporting brands that genuinely align with their values.



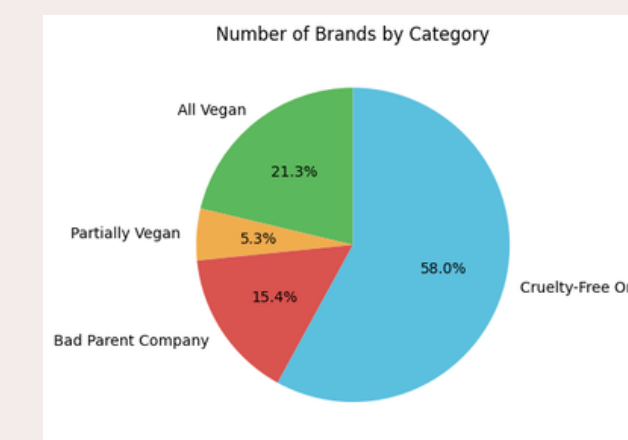
CULT BEAUTY



SEPHORA



BEAUTY BAY



ULTA

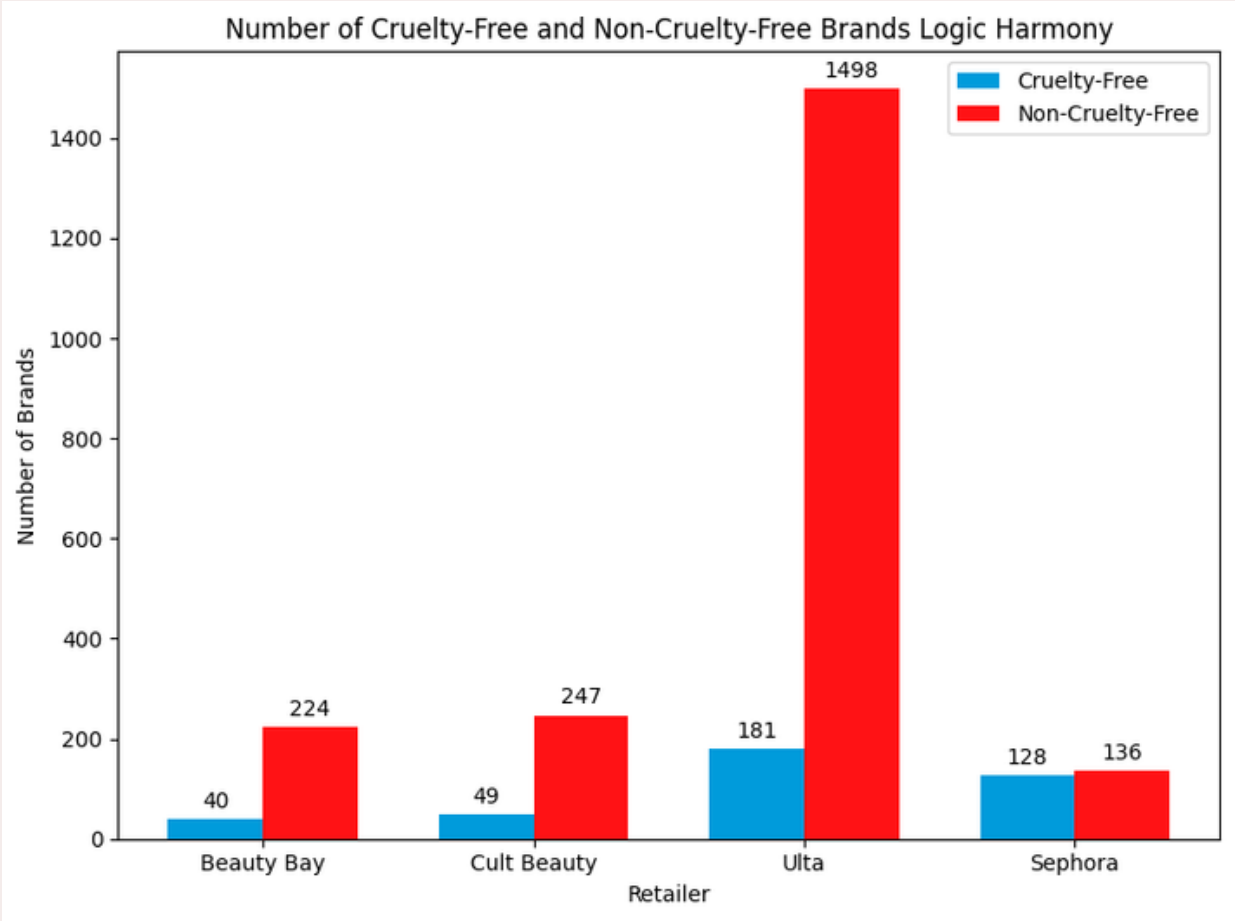
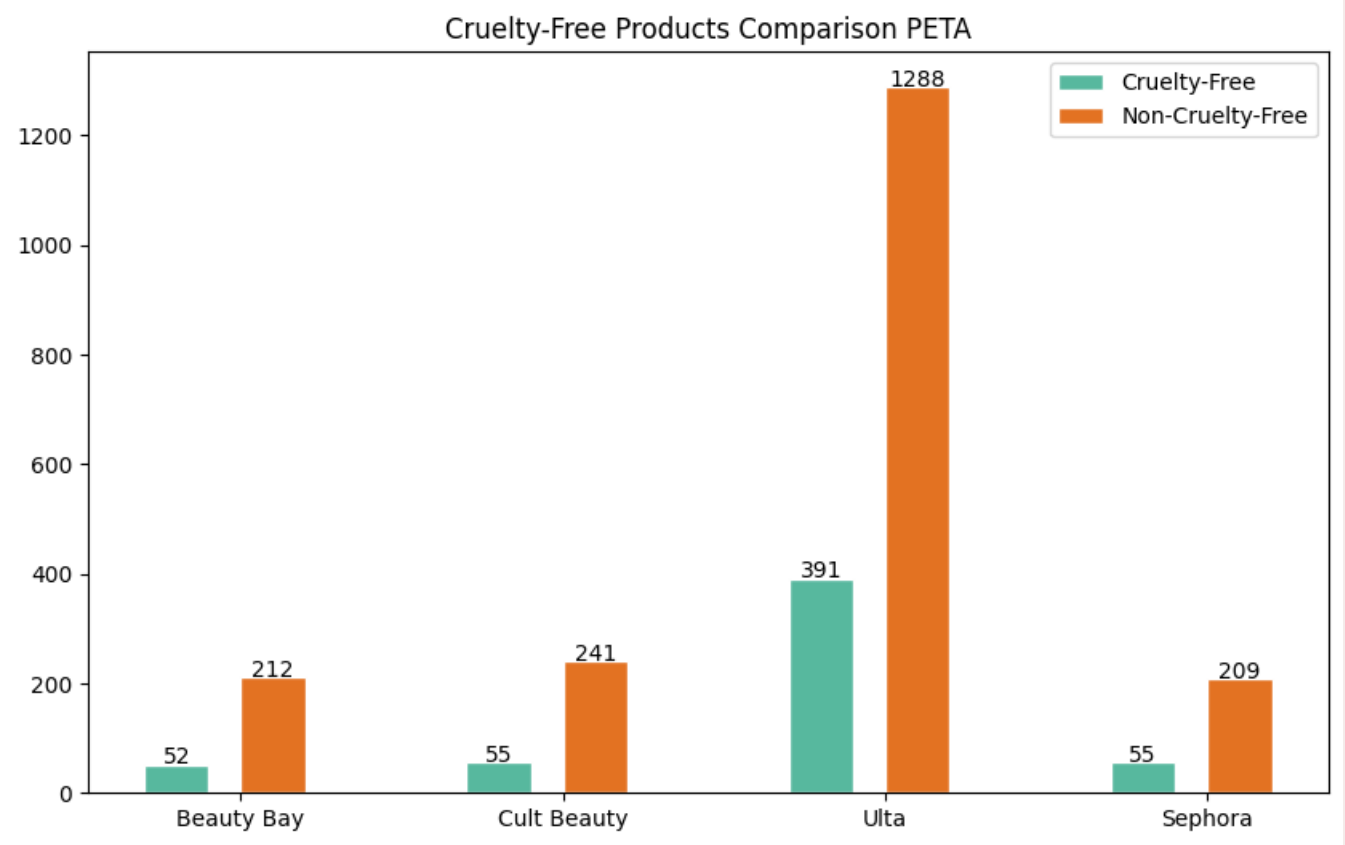
OVERALL RESULTS

When I initially started this project, I had uncertainties about what the results would reveal. It is important to acknowledge that not all brands are included in cruelty-free databases, as websites are constantly changing. Additionally, brands often modify their policies over time. Despite these challenges, the findings are surprising.

It is evident that the popular cosmetic websites I examined offer a surprisingly limited number of verified cruelty-free options compared to the overall number of brands available. Despite the increased awareness of cruelty-free cosmetics in the past decade, progress in terms of providing a wide range of options has been remarkably slow.

The final charts I have created on the right side clearly depict this. They demonstrate the stark contrast between the relatively small percentage of brands that are verified as cruelty-free and the rest of the brands available on these websites. The visual representations serve as compelling evidence of the current state of the industry and highlight the need for continued efforts to promote and support cruelty-free alternatives.

It is crucial to use these findings as motivation for further discussions and actions within the cosmetics industry to accelerate the rise of cruelty-free practices and ensure that ethical choices become more available to consumers.



INITIAL DESIGN

After successfully gathering the required data, I proceeded to plan the presentation of my findings. Considering my aim to create a cosmetic product, I embarked on designing a visual representation that would effectively convey the conclusion of my project. Opting for an eye shadow palette, I found it to be an intriguing choice as it allowed me to explore a range of vibrant colours and accommodate multiple design variations.

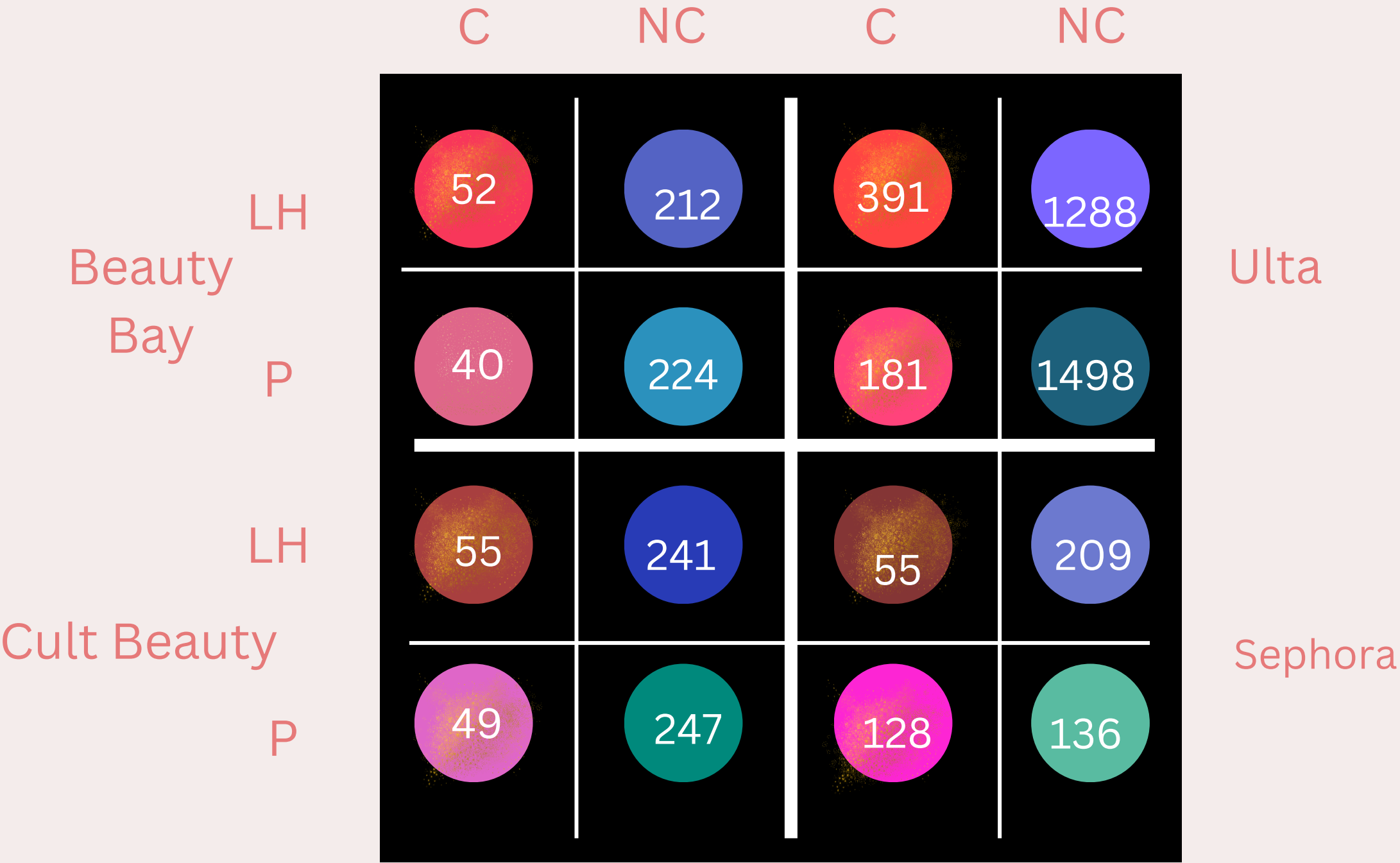
Furthermore, I recognised the packaging of the product as an opportunity to provide an explanation of the concept and design, ensuring that users would grasp the underlying idea. By incorporating informative elements into the packaging, I aimed to create a cohesive and engaging experience that would enable users to comprehend the purpose of the product.

For this palette to work and be strong data visualisation, every shade has to have a specific data element and be carefully colour code to make sense alongside this project. I began to to layout the data using columns and grids to categorise where certain colours should go.

I created a 4x4 grid where each brand would have their own 2x2 grid containing their data. Following I categorised each row with a specific cruelty free database. The columns cruelty are alternating cruelty free and non cruelty free. Next to each shade you can see the corresponding data figure, which will be displayed on the bottom of the pallet in the packaging for further context.

Colour Pallete:

Shimmers/Warm tones: Cruelty Free
Matte shades/ Blues / green : Non cruelty free



MODELLING



Figure 1: <https://www.allure.com/gallery/best-eyeshadow-palette>

I envisioned a product that would be marketed to people, encompassing both physical and digital components. To get inspiration for the layout of my products, I began examining cosmetic advertisements, like the one depicted in Figure 1. It was important for me to ensure that the digital aspect of the product would seamlessly complement the physical one, as my ultimate goal was to create a single, cohesive 3D product.



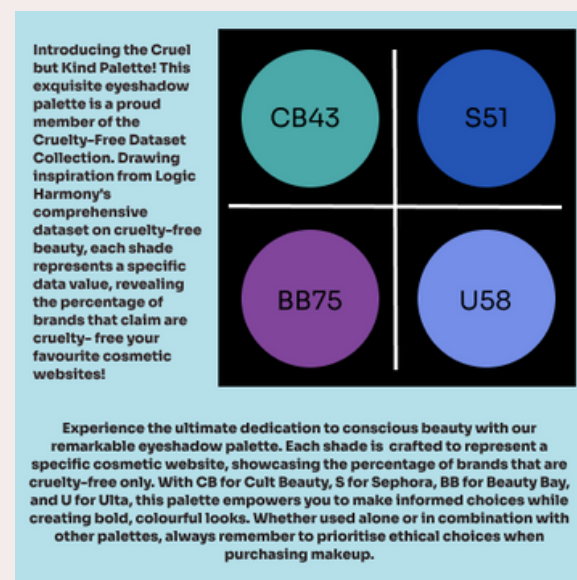
I started prototyping my palette using Blender to visualise its appearance. My goal was to create a makeup collection that reflects my research findings, resulting in a total of five palettes. The four mini quad palettes were inspired by the cruelty-free breakdown pie charts from Logical Harmony. Each quad represents the data I gathered, showcasing whether the brands were cruelty-free only, partially vegan, 100% vegan, or had a questionable parent company. The palette colours were inspired by the pie charts, and each palette has dedicated names related to the data. The larger palette incorporates data gathered from PETA and Logical Harmony, comparing the number of cruelty-free and non-cruelty-free brands, along with their respective cosmetic websites. Here is a mock-up of the final advertisement billboard, showcasing the palette collection.

FINAL PRODUCT



After allowing my shadows to dry, I carefully adhered them to the designated slots, aligning with my design. To complete the packaging, I printed out the labels as stickers and added them to the palette, giving it its final touch. The eyeshadows are fully functional and ready to be used. Here is the final product:

Here is what the final packaging looks like. 'Data Dissect' is the name of the physical large palette. On the back of the palette would be an explanation of the meaning behind the palette as well as the colour layout. Below is the packaging design for one of the four mini palletes. All the designs are similar to represent a cohesive collection.

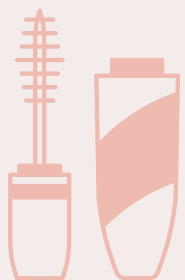


In order to create a 3D model of my data, I purchased an empty makeup palette as my canvas. I took ready made individual eye shadows, and crushed them up. I mixed different eyeshadows together to create my desired colours and then mixed them with rubbing alcohol to create a paste. I then placed the new colours into an empty eyeshadow container and let them dry. I continued this process until every shade was completed.



PROTOTYPE

Here is the final prototype of the whole collection, using all the data I was able to scrape:



LIMITATIONS/CHALLENGES

1. **Data Accuracy:** The accuracy of the results heavily relies on the accuracy of the data sources used. The data sources' errors or inconsistencies can lead to incorrect conclusions.
2. **Data Coverage:** The analysis is limited to the specific brands and retailers included in the provided datasets. If certain brands or retailers are not included, the analysis may not provide a comprehensive view of the cruelty-free status across the beauty industry.
3. **Definition of Cruelty-Free:** The determination of cruelty-free status can vary among different organizations and certifications. The analysis assumes the provided datasets accurately represent cruelty-free brands, but there might be variations in the criteria used to classify a brand as cruelty-free.
4. **Data Collection Method:** The analysis relies on publicly available data sources. It's possible that some brands or retailers may have updated their cruelty-free status since the data was collected. Regular updates and verification of data sources are necessary to maintain accuracy.
5. **Brand Ownership:** Due some brands being owned by parent companies, these parent companies may also change their policies. This complexity can affect the accuracy of determining the overall cruelty-free status of a brand or retailer.
6. **Changing Landscape:** The beauty industry is constantly changing, and the cruelty-free status of brands and retailers can change over time. New brands may emerge as cruelty-free, while existing brands may change their policies. Regular updates and continuous monitoring are essential to keep the analysis up to date.



IMPROVEMENTS

Overall I am very happy with my final product, as I was able to complete a digital and physical outcome using the data I scraped. I am happy that I managed to scrape 6 websites and research a topic that I am passionate about. However, if I had more time the improvement I would make would be:

- **Expanded Data Coverage:** Expand the scope of data collection to include a wider range of brands and retailers in the beauty industry. This will provide a more comprehensive view of the cruelty-free landscape and enable a more accurate analysis.
- **Real-time Data Updates:** Implement a system to regularly update the data to reflect any changes in brand cruelty-free status. This could involve monitoring brand websites, official certifications, or using web scraping techniques to collect real-time data.
- **Data Validation:** Use more cruelty-free datasets to validate the data obtained from different sources to ensure its accuracy and consistency. Cross-reference these data sets and verify the cruelty-free status of brands.



REFERENCES



- Alaouir, T., Gustavsson, R. and Schmidt, N., 2019. Factors driving purchase intention for cruelty-free cosmetics: A study of female millennials in Jönköping, Sweden.
- Bajaj, P., Harris, J.F., Huang, J.H., Nath, P. and Iyer, R., 2016. Advances and challenges in recapitulating human pulmonary systems: at the cusp of biology and materials. *ACS Biomaterials Science & Engineering*, 2(4), pp.473-488.
- Baker, J. (2015, April 02). The rise of the conscious consumer: Why businesses need to open up. Retrieved June 10, 2020, from <https://www.theguardian.com/women-in-leadership/2015/apr/02/the-rise-of-the-conscious-consumer-why-businesses-need-to-open-up>
- Cahill, E., 2018. An investigation into the extent to which a consumer's perspectives and loyalties are impacted by cosmetic brands testing their products on animals (Doctoral dissertation, Dublin, National College of Ireland).
- Chitrakorn, K. (2016, January 13). Is the global cosmetics market moving towards a cruelty-free future? Retrieved June 10, 2020, from <https://www.businessoffashion.com/articles/inside-beauty/is-the-global-cosmetics-market-moving-towards-a-cruelty-free-future>
- Cruelty-free labelling (2022) MSPCA. Available at: https://www.mspca.org/animal_protection/cruelty-free-labeling/ (Accessed: 13 June 2023).
- Fromm, J., (2016). How to Market To Millennials in The Experience Economy. *Forbes*. Available at: <https://www.forbes.com/sites/jefffromm/2016/05/07/how-to-market-to-millennials-in-the-experience-economy/#7f9333165ef7>
- Gangai, D.K.N. and Agrawal, R., 2016. The influence of personality traits on consumer impulsive buying behaviour. *International Journal of Marketing and Business Communication*, 5(1).
- Kulkarni, N., The Cruelty-Free Pseudo-Event.
- Lau, M., Chang, M., Moon, K. & Cruelty fr Liu, W. (2006). Influence of brand loyalty on consumer sportswear. *International Journal of Economics and Management*, [online] 5(1), pp.221-236. Available at: https://textiles.ncsu.edu/tatm/wpcontent/uploads/sites/4/2017/11/Lau_Full_169_05.pdf
- Leon, D.I.B., 2020. An examination of the growth of cruelty free products available for the 18-24 age range.
- Lippmann, W. (2004). *Public opinion*. Mineola, NY: Dover Publications.
- McDougall, A. (2018, August 20). How beauty brands can leverage veganism's popularity. Retrieved June 10, 2020, from <http://www.mintel.com/blog/beauty-market-news/how-can-beauty-companies-make-the-most-of-veganisms-rising-popularity>
- Rose, S., (2019). How L'Oréal is Misleading Customers About Being Cruelty Free. *Cruelty Free Kitty*. Available at: <https://www.crueltyfreekitty.com/news/loreal-animal-testing-notcruelty-free>
- Rowan, A. N. (2007). A brief history of the animal research debate and the place of alternatives. *Alternatives to Animal Testing and Experimentation*, 12(3), 203-211.
- Seidle, T. (2013). *Humane Society International's Global Campaign to End Animal Testing*. [ebook] Toronto: Humane society international, pp.1-3. Available at: http://altweb.jhsph.edu/pubs/journals/atla/41_6/seidle.pdf
- Smith, S. (2023) What does the reversal of the ban on animal tests for cosmetics mean for UK beauty, *TheIndustry.beauty*. Available at: <https://theindustry.beauty/what-does-the-reversal-of-the-ban-on-animal-tests-for-cosmetics-mean-for-uk-beauty/> (Accessed: 13 June 2023).
- Whitehouse, L. (2017, July 18). 'Cruelty-free' cosmetics market to grow by over 6% in coming years. Retrieved June 10, 2020, from <https://www.cosmeticsdesign-europe.com/Article/2017/07/18/Cruelty-free-cosmetics-market-to-grow-by-over-6-in-coming-years>
- Winders, D.J., 2006. Combining reflexive law and false advertising law to standardize cruelty-free labeling of cosmetics. *NYUL Rev.*, 81, p.454