An investigation into who has more influence on consumer purchasing decisions: celebrities or social media influencers?

Eleanor Zhu and Ore Bajela

Abstract

In an ever changing marketing landscape, discerning the impact of promotional strategies on consumer behavior is crucial. This exploration delves into the contrasting influences of celebrity endorsements and influencer marketing on consumer purchasing decisions, focusing specifically on skincare products. Leveraging Google Trends data and Instagram demographics, our research aimed to provide insights into these promotional strategies' effectiveness. We discovered that while influencer-centric promotions exhibited limited effectiveness, a synergy of both influencers and celebrities significantly influenced consumer behavior. Specifically, Cerave's combined approach emerged as the most impactful, outperforming both exclusive influencer and celebrity promotions. Our findings highlight the intricate dynamics of promotional strategies and their implications for consumer choices. This study not only highlights the importance of influencer selection but also encourages further exploration into consumer decision-making across diverse markets.

1. Introduction

In today's ever changing landscape of advertising and marketing, understanding the factors that influence consumer purchasing decisions has become increasingly important. As the digital realm continues to reshape the way brands connect with their audiences, it becomes increasingly essential to investigate the efficacy of different promotional strategies. This research project, therefore, aims to shed light on the pivotal question: Between celebrity promotion and influencer marketing, which has more influence on consumer purchasing decisions?

This research problem unfolds in the context of a rapidly changing marketing environment. Traditional advertising methods, such as television and print media, have witnessed a substantial shift towards digital platforms and social media. In this digital age, the power of endorsements is no longer restricted to celebrities who command a broad, but often distant, audience. It has now trickled down to a new class of influencers—individuals who may not have global fame but boast a deep and intimate connection with their followers, sometimes numbering in the millions. Understanding the relative impact of these two prominent promotional avenues, celebrities and influencers, is essential for marketers. Within this context, several gaps in knowledge emerge. While prior research has explored the impact of celebrity endorsements and influencer marketing independently, a comprehensive comparison of these two promotional strategies remains limited. Furthermore, most existing studies rely on self-reported data or small-scale surveys, which may not fully capture the nuanced decision-making processes of consumers in the real world.

To address these gaps, our research aims to employ the various methods we have learnt over the course of the semester to create a Python-based data analysis that will allow us to give a comprehensive answer to our research question. Our methodology will involve the use of various API’s as well as visualization techniques and statistical inference. By integrating these diverse methods, we aim to provide a holistic perspective on the influence of celebrity promotion and influencer marketing on consumer purchasing decisions. In doing so, we hope to deliver insights that can guide marketing strategies, inform decision-makers, and contribute to the discussion on the evolving world of advertising practices.

1.1 Data

Our data was sourced from Google Trends and the Instagram API. Google Trends provided insights into the temporal alignment of brand campaigns with changes in search volumes. Additionally, we analyzed related search queries to discern associations between brand names and influencers or celebrities. The Instagram API offered demographic details of brand followers, allowing us to compare these demographics with influencer and celebrity profiles. This comparison revealed potential consumer connections and gauged the impact of endorsements on search trends. Moreover, we identified co-occurrences of brand and personality names in searches, highlighting the digital relevance of influencers and celebrities. This comprehensive dataset allowed us to create visualizations, and employ statistical inference to facilitate nuanced analysis and informed insights.

1.2 Approach

This study employed a multifaceted approach, incorporating data visualizations that complemented statistical inference techniques such as p-tests, t-tests, and chi-squared hypothesis testing. These statistical analyses formed the cornerstone of our investigation, highlighting the relationship between various promotional strategies and subsequent fluctuations in search trends.

Our data visualizations included a line graph illustrating the time-related trends in Google searches. The shading of periods coinciding with celebrity promotions provided a visual depiction of the correlation between these promotions and fluctuations in search volumes. This visualization facilitated a clear understanding of the impact of promotional activities on search trends over time.

The utilization of hypothesis testing methodologies allowed us to assess the impact of promotional strategies on search trends. P-tests and t-tests enabled us to evaluate the significance of observed differences in search volumes during specific campaign periods. The chi-squared test aided in discerning associations between influencer and celebrity promotions and search trend variations, offering insights into the efficacy of these marketing approaches.

1.3 Summary and insights

Our findings shed light on the varying degrees of effectiveness of different promotional strategies within the skincare market. Notably, influencer-centric promotions, exemplified by Curology, exhibited marginal influence on increased Google searches, as indicated by high p values and low t-test values. Contrarily, Neutrogena's celebrity-focused approach displayed a statistically significant correlation with heightened search volumes, supported by low p-test values and high t-test results. However, Cerave's combination of influencer and celebrity promotions emerged as the most potent strategy, evident from significantly low p-values and high t-test outcomes. This research underscores the differential impact of influencer and celebrity promotions on consumer behavior within the skincare market. While influencer-centric campaigns showcased limited effectiveness, a synergistic blend of both influencer and celebrity endorsements emerged as the most impactful strategy. However because our data was limited to google searches rather than actual sales it is difficult to conclude that these correlate with consumer purchasing decisions. However, when corroborated with amazon bestseller lists, we can see that Cerave products appear the most in bestseller lists out of our three brands, followed by neutrogena and then Curology which does support our statistical conclusion.

2. Methods

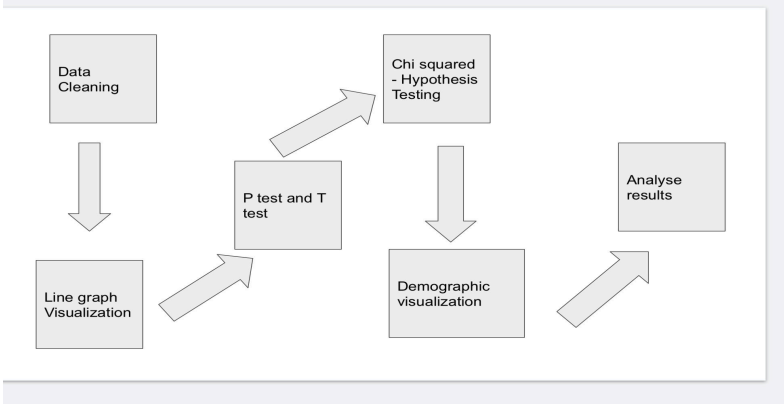

Figure 1 : Project workflow

2.1 Data Cleaning

To prepare our dataset, we carefully constructed three distinct Google Trends timelines, one dedicated to each skincare brand under analysis. The selection of timeline durations were specifically curated to encapsulate periods before, during, and after significant promotional activities associated with each brand.

For our analyses of Cerave and Curology, we opted for a comprehensive five-year timeline. This extended duration allowed us to gather weekly data points, totaling 260 points each. This breadth of data offered a large foundation to explore the fluctuations in search trends during various promotional phases associated with these brands.

Contrarily, our dataset for Neutrogena covered solely the most recent year. This focused timeline decision was influenced by historical controversies that led to product recalls in preceding years, which skewed search results. By limiting our dataset to the past year, consisting of 52 data points, we aimed to ensure the reliability and accuracy of our analysis.

In leveraging the Instagram API for demographic insights, we obtained data reflecting the current follower counts of the top three influencers or celebrities endorsing each brand. This data enabled direct comparisons across brands and provided a glimpse into the influence wielded by these personalities. Additionally, we collected data on related search queries during specific promotional periods. This allowed us to gauge the inclusivity of particular influencer or celebrity promotions within searches associated with the respective brand names.

2.2 Search trend line graphs

To visualize the relationship between search trends and campaign periods we employed a line graph with vertical shaded lines to indicate the points in our data that were in between promotional campaign periods. A line graph effectively illustrates temporal trends, showing how search volumes fluctuate over time. The vertical lines indicating periods of promotion highlight crucial intervals that allow us to visually segregate these periods , making them easily identifiable within the timeline. This allows us to create context for the observed fluctuations.

2.3.1 P test and T test

The p-test and t-test are statistical methods used to assess the significance of observed differences between groups or conditions in a dataset. They are used in research to determine if

there is a meaningful difference between groups or if the observed results could have occurred by chance.

We used the t-test to determine if there is a significant difference between the means of two groups- search volume during periods of promotion and search volumes during periods without promotion. It was computed by using python code to take the difference between the means of two groups and dividing it by the standard error of the difference. The resulting t-values were then compared against critical values from the t-distribution to determine its significance. If the calculated t-value exceeds the critical values, it suggests a significant difference between the paired means.

We used the p test to help us determine the significance of the observed differences given to us by the t test. It indicates the probability of observing a result as extreme as the one obtained in the data, assuming that the null hypothesis is true. A low p-value indicates that the observed results are unlikely to have occurred by chance. It was calculated using python code as with the t test. The p value was derived from the t-distribution and the calculated t-statistic. It represents the probability of observing the given t-statistic (or a more extreme value) under the assumption that there is a correlation between increased search volume during promotional periods.

2.3.2 Chi squared Hypothesis testing

Chi-squared (χ^2) hypothesis testing is a statistical method used for categorical data analysis to determine if there is a significant association between two categorical variables. For our analysis of influencer/celebrity promotions and search trends, where variables are categorical -influencer/ celebrity promotion included in search queries or not the chi-squared test is fitting. It allowed us to assess independence or association between specific influencer/ celebrity promotions and the occurrence of brand-related search queries during campaigns. The calculation involved comparing observed frequencies of categories in a contingency table with the frequencies we would expect if there were no association between the variables.

Formula: $x^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$

Where O_i represents the observed frequency in each category and E_i represents the expected frequency in each category if there were no association between variables.The sum ($\sum$) is taken across all categories.

In this context, our null hypothesis:

H_0 = There is no relationship between periods of higher search volume and ad campaigns for the brand, and alternative hypothesis:

H_A= There is a relationship between periods of higher search volume and ad campaigns for the brand.

To decide whether to reject the null hypothesis, we set a significance level of 0.15. If the p-value obtained from the Chi-squared test is less than the chosen significance level (0.15), we reject the null hypothesis, and conclude that there was evidence for a relationship between the variables. In this case we chose 0.15 rather than the usual 0.05 because it creates higher sensitivity to the observed differences which is more appropriate for exploratory analysis like this. It allows us to detect potential differences that might have been overlooked with stricter thresholds.

2.4 Demographic visualization pie chart

We were able to create pie charts that showed what percentage of a chosen influencer, celebrity or brand's following fit into each demographic. We then visually compared each brand's following to their chosen influencer or celebrity to see how these align or differ.

3. Results and Discussion

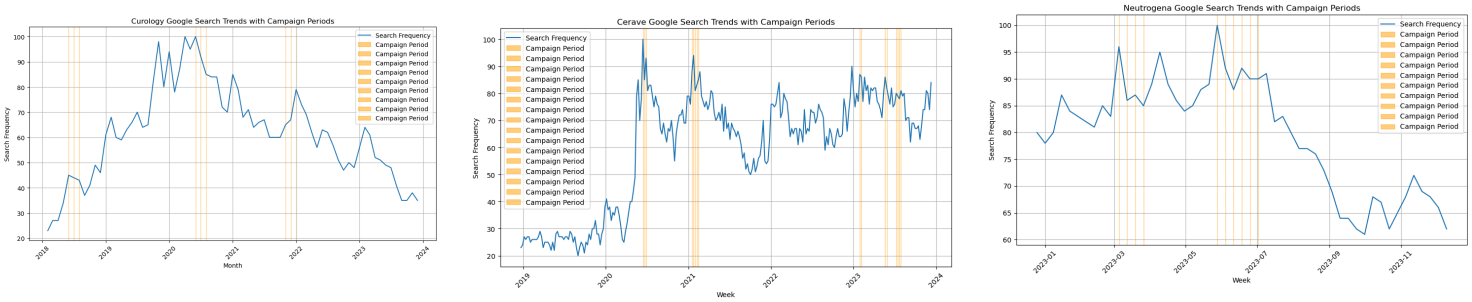

Figure 2. Google Searches of Curology, Cerave, and Neutrogena Over Last (5, 5, 1) Year

3.1 Line graph visualizations

As we can observe in these graphs, for curology it appears that search volume fluctuated significantly. There were some relative peeks that corresponded with periods of promotion, but overall the highest peaks were not during periods of influencer promotion. These observations suggest that other factors might have influenced the highest surges in interest. Other marketing strategies, perhaps seasonal variations or price reductions.

Conversely, for neutrogena, we can see that while search volume fluctuated as well the highest peaks were during periods of celebrity promotion. We can see that in March and April, when Jenna Ortega, a celebrity was first announced as a brand ambassador we have the second highest peak and then in June and July when an advertisement featuring Jenna Ortega and Jennifer Garner is aired we see the highest peak. According to google trends related queries, both these celebrity names were searched alongside the neutrogena brand name. These observations underscore the significant influence of celebrity endorsements, indicating that the presence of these personalities in Neutrogena's promotions contributed notably to heightened consumer interest and search activity

Similarly, with cerave, we can see that even though search volume fluctuated, almost all relative peaks coincided with influencer and celebrity promotions. According to our related queries data, the highest peak search volume was paired with searches of the cerave brand influencer Hyram who is a skincare influencer. This consistent alignment strongly suggests a correlation between Cerave's search surges and the timing of influencer and celebrity endorsements. It indicates that these promotional activities consistently stimulated consumer interest and subsequent search behavior. Direct Influence of specific influencers like Hyram, underscores their direct influence, in driving heightened interest and search activity for Cerave. It

highlights the potency of targeted influencer collaborations in engaging consumers and driving brand visibility.

3.2. 1 P tests an t test

For Curology, we arrived at a T-statistic approximately equal to 1.23, this relatively low t value indicates that there was a small difference between the mean search volume during periods of promotion and periods of no promotion.
The p value is approximately equal to 0.219, this relatively high p value indicates that any observed increase in search volume during periods of influencer promotion were likely by chance and are not significant evidence of a correlation.

For Neutrogena, we arrived at a T-statistic approximately equal to 4.33 this t value is relatively higher than that of curology indicates that there was a significant difference between the mean search volume during periods of promotion and periods of no promotion.
The p value is approximately equal to 7.29e-5, this very low p value indicates that any observed increase in search volume during periods of influencer promotion likely did not occur by chance and offers significant evidence of a correlation.

For Cerave, we arrived at a T-statistic approximately equal to 5.26 this t value is higher than that of both curology and neutrogena and indicates that there was a significant difference between the mean search volume during periods of promotion and periods of no promotion.
The p value is approximately equal to 3.03e-7, this very low p value indicates that any observed increase in search volume during periods of influencer promotion likely did not occur by chance and offers significant evidence of a correlation.

3.2.2 Chi squared hypothesis test

For curology our chi squared statistic p value was approximately equal to 0.337, this is greater than our chosen significance level of 0.15. Therefore we fail to reject the null hypothesis and conclude that there is no sufficient evidence to claim a relationship between periods of influencer promotion and higher search volumes.

For neutrogena our chi squared statistic p value was approximately equal to 0.113, this is less than our chosen significance level of 0.15. Therefore we reject the null hypothesis and conclude that there is sufficient evidence to claim a relationship between periods of celebrity promotion and higher search volumes.

For cerave our chi squared statistic p value was approximately equal to 1.34e-6, this is significantly less than our chosen significance level of 0.15. Therefore we reject the null hypothesis and conclude that there is sufficient evidence to claim a relationship between periods of celebrity promotion and higher search volumes.

3.3 Pie chart Demographics

Curology

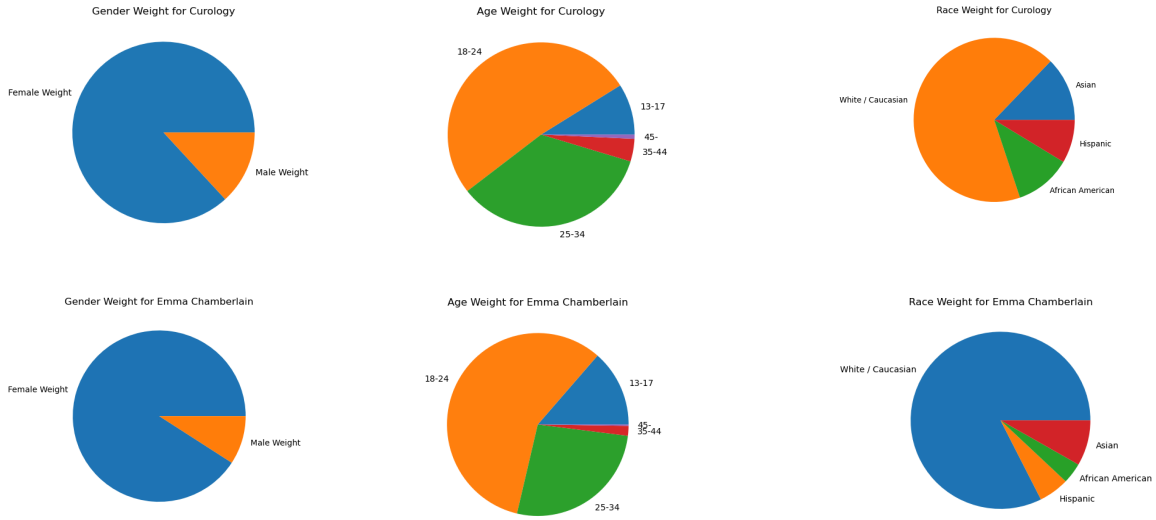

Figure 3. Comparing Curology and Emma Chamberlain Instagram Demographics

Neutrogena

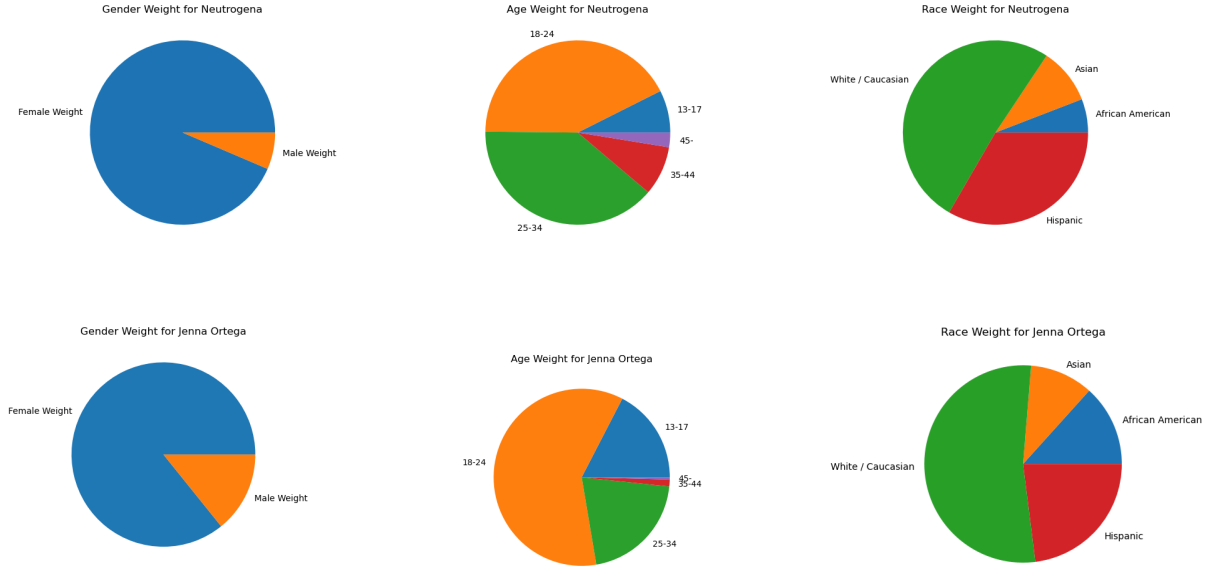

Figure 4. Comparing Neutrogena and Jenna Ortega Instagram Demographics

Cerave

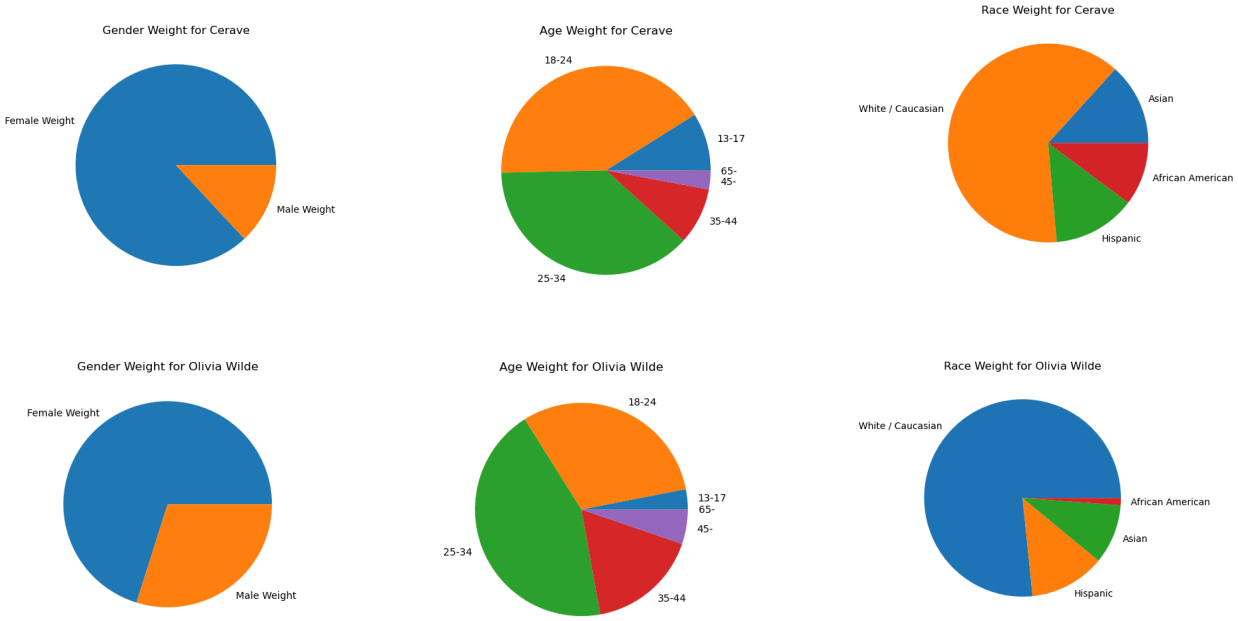

Figure 5. Comparing Cerave and Olivia Wilde Instagram Demographics

From the pie charts, we can see that for all the brands, their chosen influencers following is very similar to theirs and cerave has the most diverse following while curology has the least diverse following.
This alignment between brand and influencer/ celebrity following suggests that the influencers/ celebrities selected by each brand resonate well with the brand's target audience. It signifies a strategic alignment in choosing influencers//celebrities whose follower demographics align with the brand's target consumer base.

Cerave's diverse following indicates a broader appeal across various demographics, potentially catering to a wider consumer base. On the other hand, Curology's less diverse following might signify a more focused or niche audience. The similarity between influencers' following and the respective brands' following reflects a strategic approach in influencer selection. It implies that these brands have chosen influencers whose audience aligns well with their own, maximizing the potential impact of influencer endorsements. Cerave's diverse following also suggests a brand that engages with a broad spectrum of consumers, potentially employing versatile marketing strategies. Conversely, Curology's less diverse following might indicate a more targeted approach in engaging with a specific segment of the market.

3.4 Further Analysis

Overall, we can conclude from all our results that curology's influencer promotion was the least effective in stimulating consumer interest, while Cerave's combination of influencer promotion and celebrity promotion was the most effective. This was surprising because we expected that influencer promotion would be at least more significant than celebrity promotion because influencers on the surface seem to be able to build more personal connections with their audience as they make a career out of opening up their lives but according to our data, even though curology had some of the biggest influencers promoting their products, it did not seem to have as much of an effect.

We also observed that with Cerave's results, the highest peaks in search volumes corresponded with their influencer promotions rather than their celebrity promotions which further contrasted our results. However, we were able to come up with an explanation for this. When comparing the influencers that promote curology compared to those who promote cerave there is a significant difference. While curology's influencers had far larger followings, cerave's influencers were perhaps more effective because they are professionals in skincare. Ceraves influencers are all either licensed estheticians or dermatologists compared to Curology's more general lifestyle influencers. This indicates that consumer behavior does not respond as much to influencers who they feel are not knowledgeable in the field. This conclusion was supported by research in the field form Determ- a company that focuses on AI media monitoring and analytics. They found that it is best to leverage influencers with targeted audiences that align with the brand's speciality.

Conversely, because we still observed correlation between celebrity promotions and increases in search volume, we may be able to conclude that unlike influencers, we respond to celebrities even when they are not experts in the field.

3.5 Limitations

Unfortunately, there are some limits to the reliability of our observed results. For example, correlation doesn't necessarily imply causation. A high correlation between influencer and celebrity promotion periods and search interest doesn't confirm that one causes the other. There may also be external market events, trends, or competitor actions that might influence consumer behavior, affecting the reliability of our conclusions.
High follower counts also don't guarantee high engagement or influence. The quality of the influencer's interaction with their audience matters. Similarly perception of celebrities may change due to unrelated events or controversies, affecting their influence on consumer decisions. There might also be critical factors not considered in our analysis, for example curology is a much newer company than cerave and neutrogena so this affects the popularity of the product.

Our findings could also be specific to the skincare industry and might not extrapolate to other industries or markets.

However, the most significant limitation was our lack of sales figures. Most companies do not publish their sales which makes sales data impossible to retrieve. It's hard to directly correlate online activity with actual purchases, impacting our assessment of marketing influence on purchasing decisions.

In an effort to mitigate this we employed the amazon bestsellers list, specifically the skincare section. Amazon is one of the top companies where consumers purchase their skincare products and the bestsellers list reveals which products are the highest selling. As our data predicted, cerave had the most products in the top hundred, boasting 12 products more than any other brand. Neutrogena had 3 products in the bestsellers list and curology had none.

Interestingly, the best selling skincare item- the CosrX snail mucin- is a product which has received very little promotion by the brand that owns it in the USA. It is a korean skincare product which trended on the internet but was pushed not by influencers or celebrities, but regular people who tried to product and then posted on the internet detailing how well it worked for them. It seems influencers are just far less relatable than we would expect, especially when compared to regular people.

4. Conclusion

In this exploration, we aimed to compare the impact of influencer marketing versus celebrity endorsements on consumer purchasing decisions within the skincare market and through rigorous analysis of Google Trends data and Instagram demographics, we uncover compelling insights. Overall we were able to conclude that a combination of influencer and celebrity promotion is the most effective strategy and exclusively influencer promotion is the least effective.

These results hint at the nuanced influence of promotional strategies and the varying impacts of influencers versus celebrities on consumer behavior. However, the absence of direct sales figures limited our ability to conclusively tie online activity to purchasing decisions. Mitigating this limitation, our comparison with Amazon bestseller lists affirmed Cerave's heightened presence, reinforcing our statistical conclusions.

Our findings emphasize the importance of influencer selection. Cerave's influencers, mostly skincare professionals, seemed to resonate more effectively than Curology's lifestyle-focused influencers, indicating that expertise in the field might significantly influence consumer responses.

While our study sheds light on skincare marketing dynamics, it's essential to acknowledge several limitations. Correlation does not imply causation, and external market forces or unaccounted factors may impact consumer behavior. Future research could delve deeper into understanding the nature of influencer and celebrity impact, considering more industries or evaluating specific consumer engagement metrics beyond search volumes. In conclusion, our study underscores the complexity of marketing influences within the skincare industry. It highlights the potency of tailored influencer collaborations, offering valuable insights for marketers. However, it also prompts further exploration into the multifaceted realm of consumer decision-making and the role of influencers and celebrities across diverse product markets.

5. Roles

Eleanor: Researched and collected data from Google Trends and Influencer Marketing AI API. Then extracted and cleaned data from two sources. Used data from Influencer API to create pie charts on how company demographics compare to one influencer or celebrity's instagram demographic. Created final slideshow for class presentation.

Ore: Came up with the initial idea. Collected campaign date data through looking at Google search peaks. Created line graph of Google searches overtime compared to campaign dates. Performed t-test and chi-squared test. Wrote Final Report.

References

1. Determ, 5 Ways Influencer Marketing Can Improve Your PR Campaigns, June 23, 2023 URL:https://www.determ.com/blog/5-ways-influencer-marketing-can-improve-your-pr-campaigns/

Link to google drive containing all data used as well as code

- https://drive.google.com/drive/u/1/folders/13QvV2M9Q7Wf8A4FQepk_aERoqfEAcfGy