

**UNIVERSITY OF ECONOMICS AND LAW
FACULTY OF INFORMATION SYSTEM**



MIDTERM PROJECT

Social Media Data Analytics
- 232MI6101-

Lecturer: PhD. Lê Hoành Sứ

Group: BAMOS

1. Nguyễn Thành Luân
- 2.
- 3.
- 4.
- 5.

Ho Chi Minh City, June 3rd, 2024

ACKNOWLEDGEMENT

We extend our sincere acknowledgment to PhD. Le Hoanh Su, whose unwavering support and guidance were instrumental in the successful realization of this project. Your expertise and mentorship have significantly shaped our project's direction and enriched its outcomes. We offer our deepest appreciation for your pivotal contributions, which have made this endeavor possible and fruitful.

Ho Chi Minh City, 2024

Group BAMOS

COMMITMENT

We solemnly commit to the project assigned to us under the instruction of PhD. Le Hoanh Su. We pledge that we will not copy or reproduce any materials, information, or content in any form, be it text, images, or any other media, without proper authorization or reference to the original source, ensuring full compliance with intellectual property rights and ethical standards. We are fully aware of the consequences of plagiarism and unauthorized copying, and we are dedicated to upholding the principles of academic and professional integrity in all aspects of our work. Our commitment to this endeavor is unwavering, and we will strive to deliver results that reflect the highest standards of excellence and ethics.

Ho Chi Minh City, 2024

GROUP MEMBERS

NO	FULL NAME	STUDENT ID	ROLE	CONTRIBUTION
1	Nguyễn Thành Luân	K214160991		100%
2				
3				
4				
5				

TABLE OF CONTENT

CHAPTER 1: INTRODUCTION ABOUT LIPOVITAN.....	8
1.1 Overview of Lipovitan.....	8
1.1.1 Overview.....	8
1.1.2 Market Share, Strategy, and Customers.....	10
1.1.1.1 Market Share.....	10
1.1.1.2 Strategy.....	10
1.1.1.3 Customers.....	11
1.1.3 Brand Positioning and Marketing Channels.....	12
1.1.4 Product and Competitive Advantage.....	14
1.1.5 Media Analysis.....	15
1.2 Who - Key Competitors.....	16
1.2.1 Red Bull.....	16
1.2.1.1 Overview.....	16
1.2.1.2 Market Share, Strategy, and Customers.....	17
1.2.1.3 Brand Positioning and Marketing Channels:.....	17
1.2.1.4 Product and Competitive Advantage:.....	18
1.2.1.5 Media Analysis:.....	19
1.2.2 Monster Energy.....	19
1.2.2.1 Overview.....	19
1.2.2.2 Market Share, Strategy, and Customers.....	20
1.2.2.3 Brand Positioning and Marketing Channels:.....	20
1.2.2.4 Product and Competitive Advantage:.....	20
1.2.2.5 Media Analysis:.....	21
1.2.3 Rockstar.....	21
1.2.3.1 Overview.....	21
1.2.3.2 Market Share, Strategy, and Customers.....	21
1.2.3.3 Brand Positioning and Marketing Channels:.....	22
1.2.3.4 Product and Competitive Advantage:.....	22
1.2.3.5 Media Analysis:.....	22
1.3 Lipovitan Media Channels Strategy.....	22
1.3.1 Owned Media.....	23
1.3.1.1 Lipovitan Vietnam YouTube Channel.....	23
1.3.1.2 Lipovitan Vietnam Facebook Channel.....	24
1.3.1.3 Lipovitan Vietnam Tiktok Channel.....	25

1.3.2 Paid Media - Advertising and Promotion Tactics.....	27
1.3.3 Earned Media - Public Relations and Customer Advocacy.....	28
1.4 Crawl Data.....	29
1.4.1 Related Research.....	29
1.4.2 Data Collection Focus - Prioritizing Owned Media.....	31
1.4.3 Overview Data thu được.....	34
Chapter 2: DATA PROCESSING.....	35
2.1 Data processing and EDA của Youtube Lipovitan.....	35
2.1.1 Data processing.....	35
2.1.2 EDA (Exploratory Data Analysis).....	37
2.2 Data processing and EDA của Facebook Lipovitan.....	41
2.2.1 Data processing.....	41
2.2.2 EDA (Exploratory Data Analysis).....	42
2.3 EDA của Tiktok Lipovitan.....	43
CHAPTER 3: SENTIMENT ANALYSIS.....	48
3.1 NLP Model.....	48
3.2 Facebook sentiment analysis.....	49
3.3 Tiktok sentiment analysis.....	53
3.4 Youtube sentiment analysis.....	60
Chapter 4: RECOMMENDATION AND CONCLUSION.....	64
4.1. Achieved result.....	64
4.2. Evaluation.....	64
4.3. Limitation.....	64
4.4. Future work.....	64
References.....	66

CHAPTER 1: INTRODUCTION ABOUT LIPOVITAN

1.1 Overview of Lipovitan

1.1.1 Overview

Lipovitan, known as Libogen and Livita in English-speaking markets, is an energy drink produced by Taisho Pharmaceutical Company and licensed by Osotspa. Widely popular in the Southeast Asian region, Lipovitan is available in a distinctive 100ml brown tinted bottle with bright yellow liquid. Marketed to alleviate both physical and mental fatigue, Lipovitan is sought after by consumers for its revitalizing effects.



Figure 1.1 Lipovitan in Japan

The main ingredient in Lipovitan is taurine. Its potent formula includes Lipovitan D, containing 1000 mg of taurine, 20 mg of nicotinic acid extract (Vitamin B3), 5 mg each of Vitamin B1, B2, and B6, and 50 mg of caffeine.[1] Lipovitan D Super contains 2000 mg of taurine and 300 mg of arginine, while MAXIO variant contains 3000 mg of taurine. The warning label on all Lipovitan products advises against consuming more than 100 ml per day.

Lipovitan is available in Japan and Asian specialty stores worldwide. It predates Red Bull and was first launched in 1962.[2] By 1963, Taiwan had nearly 1,400 convenience stores and supermarkets officially selling Lipovitan, marking the initial

expansion into the Asian market. Subsequently, Lipovitan entered fifteen countries and regions including Thailand (1965), Hong Kong (1969), Singapore and Malaysia (1972), the Philippines (1973), the United Arab Emirates (1977), Indonesia (1979), Bahrain (1981), the United States (1982), Vietnam (1996), the People's Republic of China (1998), Qatar (2000), Kuwait (2001), and Mexico (2003). Arginine, the primary ingredient in Red Bull, was altered for sale in Japan. The Japanese national rugby team is sponsored by Lipovitan D.

In 1999, Lipovitan officially entered the Vietnamese market, with a factory located in Khanh Hoa province. From the very beginning, the company established its mission as "Enriching lives by improving the health and beauty of everyone responsible to society."

Thanks to its quality and appealing taste, the product quickly gained the trust of Vietnamese consumers. To maintain its position as one of the most beloved energy drink brands in Vietnam, the company continuously conducts research to improve the quality and taste of its products to better suit the preferences of Vietnamese consumers.

In 2016, the product made a breakthrough by enhancing its flavor, resulting in a more palatable beverage that received positive feedback from consumers. With its delicious honey flavor and convenient canned packaging, the product is increasingly favored by Vietnamese consumers.



Figure 1.2 Lipovitan in Vietnam

1.1.2 Market Share, Strategy, and Customers

1.1.1.1 Market Share

Lipovitan's impressive market share extends beyond Vietnam, as the brand has also made significant strides on the global stage. With its distinctive formulation and effective marketing strategies, Lipovitan has secured a noteworthy position in the competitive energy drink market worldwide.

While specific market share data for Lipovitan on a global scale may vary, its reputation as a leading energy drink brand in Vietnam undoubtedly contributes to its recognition and adoption in other markets. The brand's expansion efforts, coupled with its consistent product quality and consumer appeal, have likely enabled it to capture a meaningful share of the global energy drink market.



Figure 1.3 Information of Lipovitan in Vietnam

According to Euromonitor International, the energy drink market in Vietnam is poised for substantial growth, with a projected value of 10,000 billion dong by 2025. This growth trajectory is fueled by a compounded annual growth rate (CAGR) of 5.5% from 2020 to 2025, indicating a steadily expanding market landscape.

Moreover, Lipovitan's success in Vietnam serves as a testament to its potential for growth and success in other regions. By leveraging its strong brand equity and understanding of consumer preferences, Lipovitan continues to expand its footprint and assert its presence in the increasingly competitive global energy drink market. As such, Lipovitan's market share both domestically and internationally underscores its status as a formidable player in the energy drink industry.

1.1.1.2 Strategy

To maintain and grow in a competitive market like the energy drink industry, Lipovitan has implemented a series of robust strategies. Firstly, Lipovitan has successfully diversified its product range by offering various types of energy drinks

such as Lipovitan D, Lipovitan Extra, Lipovitan C, and Lipovitan B. This diversification not only caters to the diverse needs of consumers but also provides choices, thereby increasing consumption opportunities and loyalty.

Secondly, Lipovitan consistently focuses on product quality. The brand is committed to delivering high-quality products with safe and effective ingredients. This commitment not only builds consumer trust but also creates long-term benefits for the brand by establishing a reputable and reliable image.

Thirdly, Lipovitan has heavily invested in building a strong brand through effective marketing campaigns and advertisements. Increasing brand awareness and attracting new customers are essential parts of this strategy. By continuously creating innovative and captivating advertising campaigns, Lipovitan not only garners attention from the target audience but also solidifies its position in the minds of consumers.

Lastly, Lipovitan has strengthened its distribution channels to ensure easy access to consumers. Building an extensive distribution network ensures that the product is available everywhere, from convenience stores to major supermarkets. This not only increases sales opportunities but also creates a convenient and comfortable experience for consumers when purchasing Lipovitan products.

In summary, the implementation of these strategies has helped Lipovitan maintain and expand its market share in the energy drink industry, while solidifying its position as one of the leading brands in the sector.

1.1.1.3 Customers

Lipovitan's customer base encompasses a diverse range of individuals who seek energy-boosting solutions to support their active lifestyles and demanding daily routines. The brand caters to consumers across various demographics, including young adults, professionals, athletes, and individuals seeking an energy lift to enhance their performance and productivity.

Young Adults: Lipovitan appeals to young adults who lead dynamic lifestyles and often require a quick energy boost to keep up with their busy schedules. Whether they're students juggling classes and extracurricular activities or young professionals balancing work and social life, Lipovitan offers a convenient solution to combat fatigue and stay energized throughout the day.

Professionals: Working professionals represent a significant portion of Lipovitan's customer base. Whether in corporate settings, creative industries, or entrepreneurial

ventures, professionals often rely on energy drinks like Lipovitan to maintain focus, alertness, and productivity during demanding workdays or long hours.

Athletes and Fitness Enthusiasts: Athletes and fitness enthusiasts form another important segment for Lipovitan. Whether engaging in intense workouts, endurance training, or competitive sports, these individuals rely on energy drinks to fuel their performance, enhance endurance, and speed up recovery.

On-the-Go Consumers: Lipovitan caters to individuals who lead fast-paced lifestyles and are constantly on the move. Whether commuting, traveling, or engaging in outdoor activities, these on-the-go consumers appreciate the convenience and portability of Lipovitan's energy drinks to provide a quick and effective energy boost whenever and wherever needed.

Health-Conscious Consumers: While primarily known for its energy-boosting properties, Lipovitan also attracts health-conscious consumers who seek products with safe and effective ingredients. Lipovitan's commitment to quality and safety appeals to individuals who prioritize their health and well-being while enjoying the benefits of an energy drink.

Overall, Lipovitan's customer base is characterized by individuals who value convenience, performance, and quality in their choice of energy-boosting beverages, making Lipovitan a preferred brand among those seeking to fuel their active lifestyles and achieve their daily goals with vitality and vigor.

1.1.3 Brand Positioning and Marketing Channels

Brand Positioning: Lipovitan is positioned as a premium energy drink that provides a quick and effective boost of energy for individuals who lead busy and active lifestyles. It is targeted towards young professionals, athletes, students, and anyone in need of a reliable energy boost to help them stay focused, alert, and productive throughout the day. Lipovitan's unique blend of ingredients is formulated to deliver not just energy, but also mental clarity and physical endurance, setting it apart from other energy drinks on the market.



Figure 1.4 Advertisement of Lipovitan

Marketing Channels: Lipovitan employs a comprehensive marketing strategy utilizing diverse channels to effectively reach its target market. Its products are readily available in supermarkets, convenience stores, and pharmacies, ensuring accessibility for consumers seeking energy drinks on the go. Additionally, the brand maintains a robust online presence, leveraging e-commerce platforms to cater to the growing trend of online shopping. Sponsorships of various sporting events, concerts, and high-energy activities bolster brand visibility and foster direct engagement with the target audience. Active engagement on social media platforms like Instagram, Facebook, and YouTube enables Lipovitan to share engaging content, run promotional campaigns, and interact with consumers to cultivate brand loyalty and awareness. Collaborations with influencers, athletes, and celebrities resonate with the brand's demographic, amplifying reach through authentic and relatable endorsements. Furthermore, sampling campaigns conducted at gyms, colleges, and other high-traffic locations provide consumers with firsthand product experiences, sparking word-of-mouth buzz and driving brand recognition.



Figure 1.5 Advertisement of Lipovitan in Tet Holiday

1.1.4 Product and Competitive Advantage

Lipovitan is known as an energy drink with the highest honey content along with vitamins and essential nutrients for the body, featuring a flavor specifically designed to suit Vietnamese consumers. Additionally, Lipovitan is the only energy drink that contains Vitamin B1 - a vitamin famously known as the "energy vitamin." To diversify its flavors and meet the diverse needs of consumers, Lipovitan recently launched a new product line called TONGKAT - the only energy drink in Vietnam containing Tongkat Ali herbal extract, along with high levels of vitamins and nutrients.



Figure 1.6 Products of Lipovitan in Vietnam

Lipovitan's competitive advantage lies in its commitment to delivering premium energy drink options supported by a robust distribution network and a strategic brand presence. Emphasizing natural ingredients and diverse product variants, Lipovitan ensures a distinct offering in the competitive energy drink market. Its dedication to

quality and health benefits, exemplified by formulations featuring honey and ginseng, resonates with health-conscious consumers seeking more than just a quick energy fix. Furthermore, Lipovitan's strategic marketing initiatives and sponsorships of high-energy events bolster brand visibility and engagement with its target audience, fostering brand loyalty and trust. Through these efforts, Lipovitan continues to carve out a unique position in the market, competing effectively with established leaders like Red Bull.

1.1.5 Media Analysis

"SỨC KHOẺ CỦA BẠN LÀ SỨ MỆNH CỦA CHÚNG TÔI" is the ethos driving Lipovitan's brand identity. Offering a line of energy drinks tailored for health-conscious consumers, Lipovitan goes beyond just providing a quick energy boost. Their products are formulated to enhance overall well-being, aligning perfectly with the lifestyle aspirations of Vietnamese consumers. With flavors crafted to suit local preferences, Lipovitan ensures a delightful experience with every sip, promoting vitality and wellness.



Figure 1.7 Slogan of Lipovitan

In the realm of marketing, Lipovitan adopts a strategy that resonates with its audience on a personal level. The brand's media content centers around themes of vitality, success, and the pursuit of wellness, tapping into the aspirations of its target market. By showcasing stories of individuals achieving their goals and embracing an active lifestyle, Lipovitan instills a sense of dynamism and confidence in its audience, inspiring them to live life to the fullest.

Lipovitan leverages a mix of media channels to effectively communicate its brand message. From engaging social media content to captivating influencer partnerships, the brand ensures maximum visibility and engagement with consumers. By fostering genuine connections and aligning with the values of its audience, Lipovitan strengthens its brand presence and solidifies its position as a trusted companion on the journey to a healthier and more energized life.

1.2 Who - Key Competitors

1.2.1 Red Bull

1.2.1.1 Overview

Red Bull is a product of Red Bull GmbH, an Austrian company. It made its debut in Austria in 1987, pioneering the global energy drink sector. The brand offers a variety of 8.4-ounce cans, including the original flavor, sugar-free, Total Zero (zero calories), and the Editions line, which features watermelon, blueberry, tropical, coconut, dragonfruit, peach, and strawberry-apricot flavors.

The key ingredients in Red Bull drinks are caffeine, taurine, B vitamins, and carbonated water, along with sucrose and glucose (excluding the sugar-free versions). As of the end of 2021, Red Bull is available in 172 countries and employs over 13,000 individuals. The brand is experiencing significant growth in emerging markets, with sales increasing by 37% in India, 30% in Brazil, and 25% in Africa. Moreover, Red Bull is actively focusing on expanding its presence in the U.S., Western Europe, and the Far East.

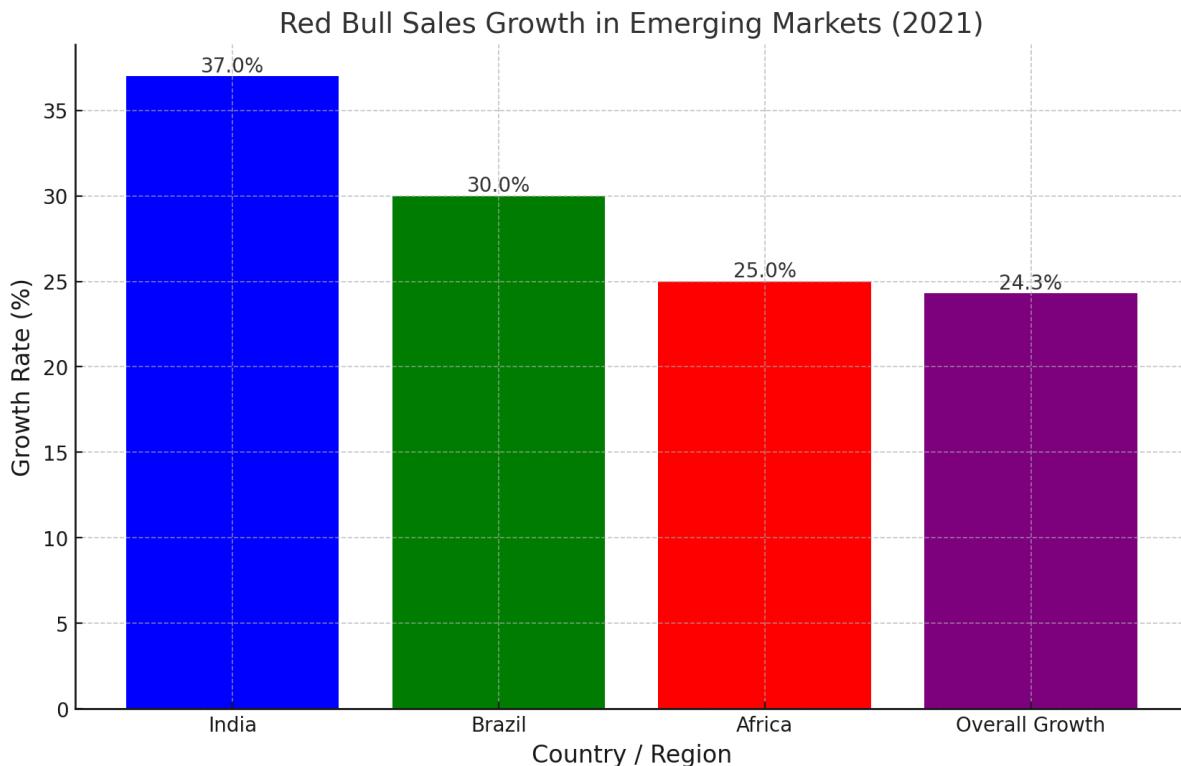


Figure 1.8: Red Bull Sales Growth in Emerging Markets (2021) Source: Own Research

Owned by CEO Dietrich Mateschitz and the Thai businessman Chalerm Yoovidhya and his family, Red Bull GmbH is a private entity. Due to its private status, detailed financial information is scarce. However, the company disclosed that it sold more than 9.8 billion cans in 2021, marking a 24.3% increase from the previous year and setting a new sales record for the company.

1.2.1.2 Market Share, Strategy, and Customers

Red Bull leads the global energy drink market with over 43% market share in 2022, demonstrating its dominance and widespread appeal. According to a 2021 statistic, Dietrich Mateschitz has a total assets of 24.5 billion USD. The brand's strategy is centered around sponsoring sports events, cultivating a dynamic brand image, and maintaining a premium product line. This approach has successfully attracted a young, energetic customer base that has a passion for sports and adventurous activities. Red Bull's targeted marketing efforts and strategic positioning have ensured it remains top-of-mind for consumers seeking an energy boost coupled with a lifestyle statement.

1.2.1.3 Brand Positioning and Marketing Channels:

Brand Positioning: Red Bull positions itself as a premium energy drink, symbolizing an active and dynamic lifestyle. This positioning is reflected in its marketing

messages, product design, and overall brand persona, appealing to consumers who aspire to a vibrant way of life.

Marketing Channels: Red Bull utilizes a diverse range of marketing channels including sports sponsorships, television advertising, social media, and influencer marketing. These efforts are designed to engage with the brand's target audience on multiple platforms, reinforcing its image as a premium, lifestyle-oriented product.



Figure 1.9: Red Bull in sports marketing source: Internet

1.2.1.4 Product and Competitive Advantage:

Products: The brand's product lineup includes the traditional Red Bull flavor, Red Bull Sugarfree, and Red Bull Editions, which offer various fruit flavors. This range caters to a wide variety of taste preferences while maintaining the core energy-boosting properties.



Figure 1.10 Red Bull's Flagship Product Source: Internet

Competitive Advantage: Red Bull's strong brand, premium product offerings, and extensive distribution network form the cornerstone of its competitive advantage. These elements, combined with effective marketing and a loyal customer base, have solidified Red Bull's leadership position in the energy drink market.

1.2.1.5 Media Analysis:

The "Gives you wings" campaign has become iconic, instilling a sense of dynamism and confidence among its audience. Red Bull's media content focuses on sports, adventure, and success, aligning with the aspirations of its target market. By leveraging a mix of media channels and incorporating influencer marketing, Red Bull effectively communicates its brand message and engages with consumers on a personal level, enhancing its visibility and appeal.



Figure 1.11 Social Campaign “RedBull Gives you Wings”

1.2.2 Monster Energy

1.2.2.1 Overview

Based in the USA, Monster Energy is one of Red Bull's major competitors in the global energy drink market. The brand was established in 2002 by Hansen Natural. Monster Energy is renowned for sponsoring many sports events, including motorcycle racing, BMX, skateboarding, as well as concerts and music events. Monster's products often have a high caffeine content and are advertised to provide "a big energy boost" to consumers. The brand has a diverse product line with many flavors and special formulas to cater to a wide range of customer needs.

1.2.2.2 Market Share, Strategy, and Customers

Monster Beverage, since its inception in 1985 and the launch of Monster Energy drinks in 2002, has ascended to become the second-largest energy drink brand globally by 2020, holding a 39% market share. The company's strategy focuses on sponsoring extreme sports, music events, and entertainment, targeting a young, vibrant customer base that gravitates towards novelty, personality, and the excitement of music and extreme sports.



Figure 1.12 Monster Energy Events

1.2.2.3 Brand Positioning and Marketing Channels:

Monster Energy positions itself as the energy drink of choice for an adventurous, dynamic lifestyle, appealing to those seeking thrills and action. Its marketing strategy leverages sponsorships in extreme sports, music, and entertainment events, as well as influencer marketing, to reinforce its brand identity. This approach not only broadens its appeal among the younger demographics but also cements its position in the energy drink market as a brand synonymous with adventure and excitement.

1.2.2.4 Product and Competitive Advantage:

With a diverse product range that includes over 40 varieties and 18 different brands such as the original Monster drink, Burn, NOS, and Predator, Monster Energy caters to a wide array of tastes and preferences. The drinks are known for their high caffeine content and inclusion of taurine, an amino acid recognized for its mental stimulation properties. This vast product diversity, combined with competitive pricing, positions Monster as a versatile and accessible option for consumers, contributing to its competitive edge in the market.

1.2.2.5 Media Analysis:

Monster Energy's robust media presence is characterized by its powerful "monster" brand imagery, which resonates with the youth for its boldness and energy. The company's media strategy focuses on music, extreme sports, and entertainment, consistently engaging with its audience through these channels. Collaborations with well-known artists and athletes further amplify its brand visibility, making Monster Energy a prominent figure in the cultural landscape of sports and music entertainment. This media strategy effectively amplifies Monster's brand recognition and solidifies its connection with its target audience.



Figure 1.13 Monster Energy collaborating with a popular athlete

1.2.3 Rockstar

1.2.3.1 Overview

Rockstar, another American energy drink brand, was founded in 2001 by Russell Weiner. Rockstar emphasizes creating products for people with active and adventurous lifestyles, with the slogan "Party like a rockstar." The brand sponsors and participates in many music and sports events, including music festivals and car races. Rockstar's products vary in flavor and functionality, including energy drinks that contain vitamins, amino acids, and caffeine to enhance energy and alertness for consumers.

1.2.3.2 Market Share, Strategy, and Customers

Rockstar holds approximately 17% of the energy drink market, ranking third after Red Bull and Monster Energy. Its marketing strategy is centered around embodying the "rockstar" persona, focusing on individuality and distinctive product flavors. The

brand targets young consumers who appreciate uniqueness, personality, and are enthusiasts of rock music and lifestyle. This demographic is drawn to Rockstar's engagement in action sports, motorsports, live music events, and collaborations with models, reflecting a vibrant and adventurous lifestyle.

1.2.3.3 Brand Positioning and Marketing Channels:

Positioned as the energy drink for the ultimate "rockstar" experience, Rockstar differentiates itself through its unique and bold flavors. Marketing channels include sponsorships in music, street art, extreme sports, influencer partnerships, and social media advertising. This approach not only enhances brand visibility but also establishes a deep connection with its core audience who lives and breathes the rockstar lifestyle, whether through music or action-packed activities.

1.2.3.4 Product and Competitive Advantage:

Rockstar's product lineup stands out with its unique flavors, ranging from the original Rockstar to sugar-free options, the Pure Zero line, and recovery drinks. Each product is designed to cater to the diverse needs and preferences of its consumers, offering added vitamins and minerals in its Recovery line and fruit blends in Rockstar Juiced. This diversity in product selection and the commitment to innovation serve as significant competitive advantages, enabling Rockstar to capture and retain a substantial segment of the youth market.

1.2.3.5 Media Analysis:

Rockstar's media presence is distinguished by its emphasis on music, street art, and extreme sports, fostering a strong sense of community among its consumers. The brand leverages this content to resonate with its audience, enhancing brand loyalty and encouraging consumer engagement through social media platforms. By maintaining a consistent theme of adventure, excitement, and nonconformity, Rockstar successfully amplifies its brand identity and fosters a dedicated "rockstar" community, further solidifying its position in the market.

1.3 Lipovitan Media Channels Strategy

Lipovitan has utilized a variety of communication channels to reach and engage with its audience. In today's marketing world, effectively using communication channels is a key factor in expanding brand presence and connecting with customers. In Lipovitan's communication channel strategy, they have built it based on leveraging three main types of media: owned, paid, and earned. Each of these media types follows separate principles, designed to optimize marketing effectiveness and enhance customer interaction.

1.3.1 Owned Media

Owned media, also known as owned media channels, are communication channels created and completely controlled by the business itself. These are channels that businesses can use to directly convey their messages to customers without relying on third parties such as the press or social media.

Social Media: In leveraging owned media, Lipovitan has built a diverse and integrated communication system, including an official website, social media pages, blogs, and mobile applications. On the official website, Lipovitan provides detailed information about its products, ingredients, usage, and benefits for consumers. They also create quality and engaging content on their blog, regularly sharing posts about healthy living, energy enhancement, and health tips.

Additionally, Lipovitan utilizes popular social media platforms such as Facebook, YouTube, and TikTok to interact with the online community. They share creative and compelling content, including images, videos, and user stories, to enhance interaction and create a vibrant community around the brand. Furthermore, Lipovitan develops mobile applications that allow users to access product information, receive the latest news updates, and participate in special promotions.

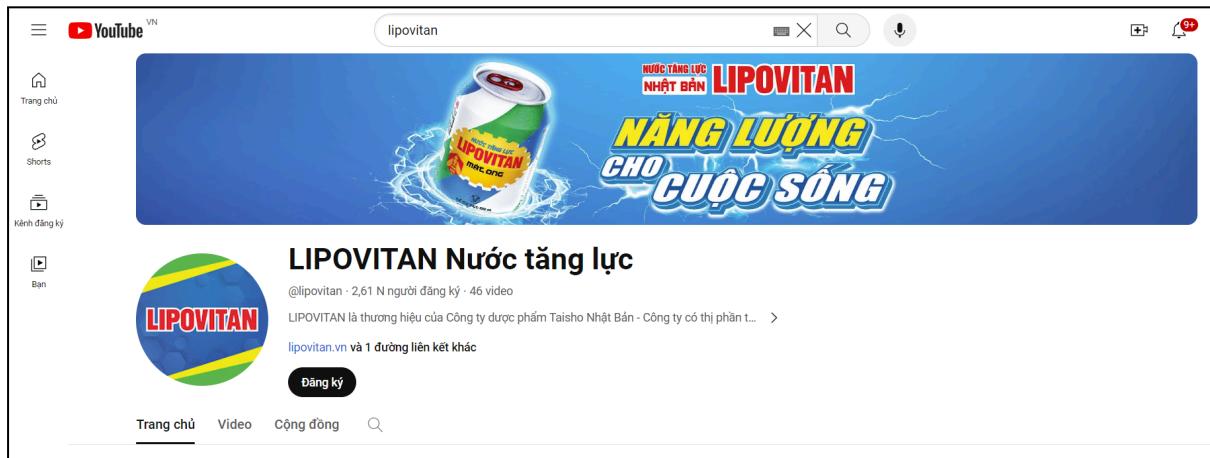


Figure 1.14 Lipovitan Vietnam YouTube Channel

1.3.1.1 Lipovitan Vietnam YouTube Channel

Lipovitan Vietnam's YouTube channel, titled "LIPOVITAN Nước tăng lực," has garnered a substantial following, amassing 2.61K subscribers since its establishment on March 27, 2014. Accessible through [this link](#), the channel serves as a vibrant platform for engaging content.

At the heart of the channel's content lies a diverse array of offerings tailored to captivate its audience. From product advertisements spotlighting Lipovitan's renowned products such as Lipovitan Original, Lipovitan C, Lipovitan D, and more, to

insightful glimpses into brand promotion endeavors, the channel leaves no stone unturned in its efforts to inform and entertain. Moreover, viewers are treated to a selection of amusing short videos infused with humor and vitality, alongside instructional guides aimed at enhancing energy levels and product usage comprehension. The channel further enhances interactivity through live stream sessions, facilitating direct engagement with its audience.

Lipovitan Vietnam's YouTube channel is driven by a multiplicity of objectives. Foremost among these is the aspiration to bolster brand recognition by delivering Lipovitan's brand and product offerings in an engaging and visually compelling manner, thus fostering meaningful connections with consumers. Moreover, the channel endeavors to spotlight promotional initiatives and exclusive offers, enticing viewers to explore and embrace Lipovitan's product range through compelling advertisement content. Additionally, the channel serves as a conduit for direct customer interaction, providing a platform for timely query resolution and proactive solicitation of customer feedback. Furthermore, the channel seeks to cultivate a vibrant community of Lipovitan enthusiasts on YouTube, facilitating mutual engagement and interaction among its followers. Lastly, by sharing inspirational narratives and spotlighting Lipovitan's impactful social endeavors, the channel endeavors to uplift the brand image and resonate authentically with its audience.

1.3.1.2 Lipovitan Vietnam Facebook Channel

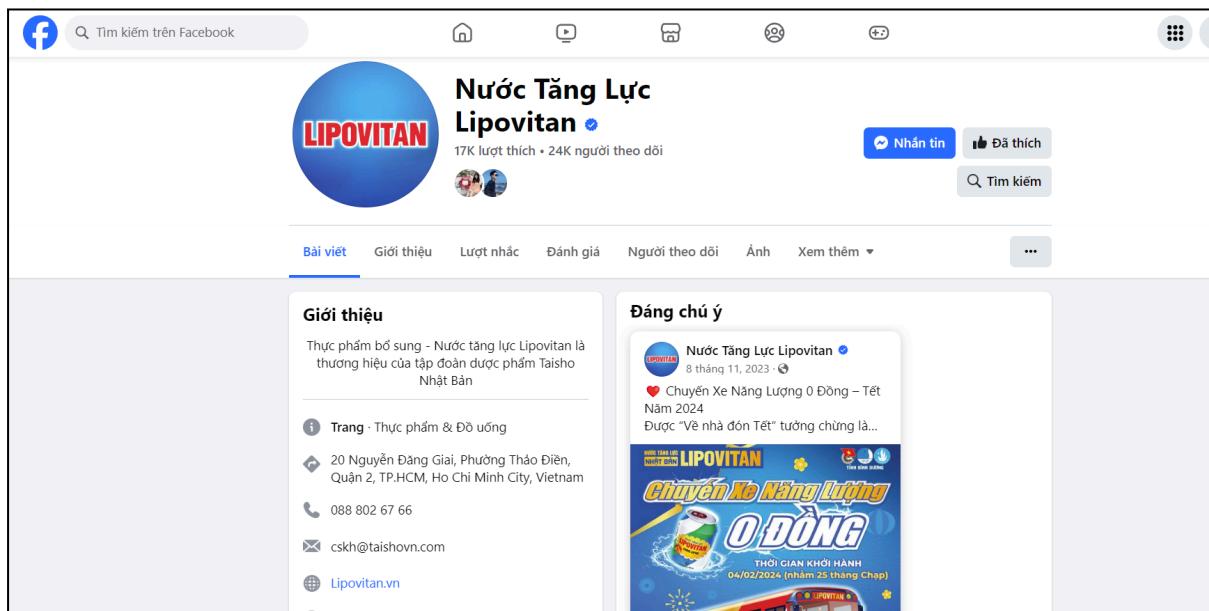


Figure 1.15 Lipovitan Vietnam Facebook Channel

The LIPOVITAN Energy Drink channel was established on May 16, 2011, and to date, it has garnered over 1.4 million followers on its Facebook page. The channel

serves not only as a platform to share detailed information about Lipovitan products such as Lipovitan Original, Lipovitan C, Lipovitan D, and more but also as a hub for engaging in attractive promotions, minigames, and giveaways, adding diversity and appeal to its followers.

The content uploaded on the channel goes beyond mere entertainment; it fosters positive interaction between the brand and its users. Humorous posts, images, and videos are posted to stimulate interest and interaction from the community of enthusiasts. Moreover, the channel regularly engages directly with users by responding to comments and messages, organizing contests, and conducting surveys to gather feedback and opinions.

The primary purpose of the channel is not only to effectively promote the Lipovitan brand and its products on the Facebook platform but also to build and grow a community of Lipovitan enthusiasts. The channel aims to create a positive interactive environment where users can feel connected and engage in conversations about energy and lifestyle.

The results of the channel have been remarkably successful, with over 1.4 million followers on the Facebook page. The consistent updates of fresh content, positive interaction, and multidimensional information sharing have contributed significantly to building and developing the Lipovitan brand in the Vietnamese market.

1.3.1.3 Lipovitan Vietnam Tiktok Channel

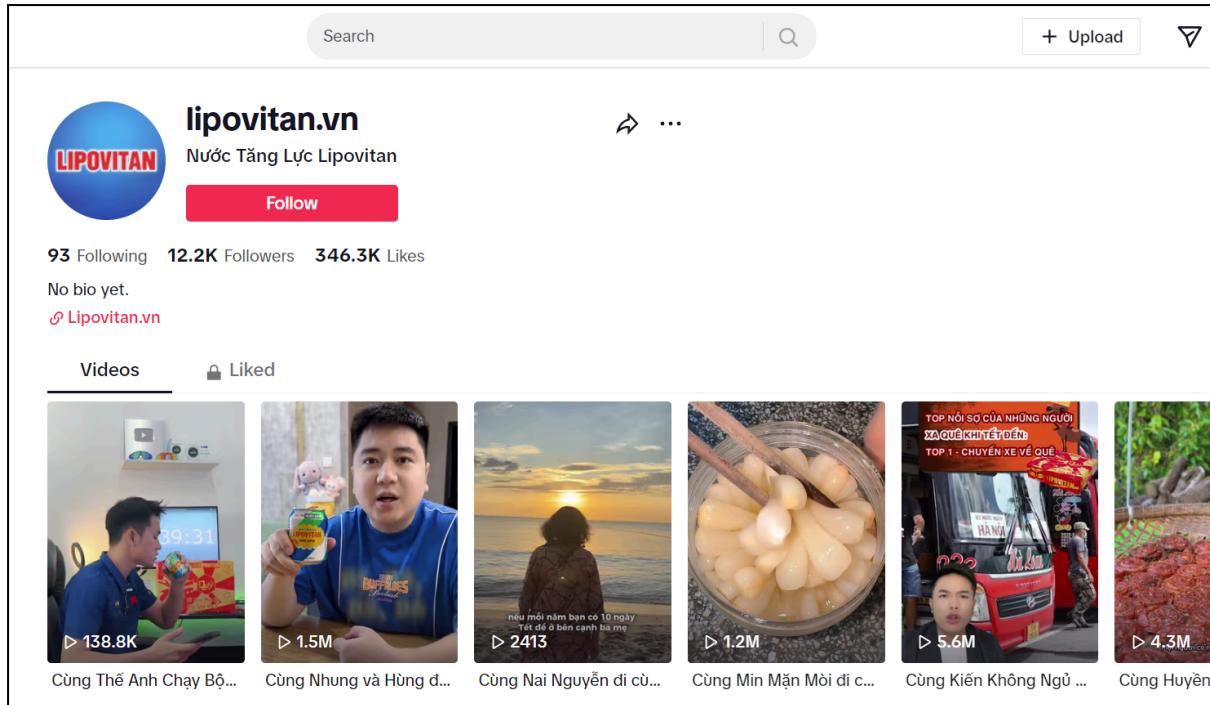


Figure 1.16 Lipovitan Vietnam Tiktok Channel

Lipovitan Nước tăng lực, a channel dedicated to promoting Lipovitan energy water, was established on August 11, 2020. Over the years, it has garnered a significant following of 71.8K subscribers, demonstrating the channel's effectiveness in engaging with its audience. Although the specific URL for the channel has been omitted, its content remains accessible to viewers, showcasing Lipovitan's commitment to reaching out to its target demographic.

The channel features an array of engaging content designed to captivate its audience. From entertaining videos to informative product showcases, Lipovitan Nước tăng lực offers a diverse range of content to cater to various interests. Its video lineup includes humorous sketches, trending TikTok parodies, lively music videos, and informative product introductions. By providing a mix of entertainment and education, Lipovitan ensures that its content resonates with its audience and keeps them coming back for more.

The primary objective of Lipovitan's TikTok channel is to enhance brand awareness and promote its products to a younger, more dynamic audience. Through creative and engaging videos, Lipovitan aims to introduce its brand and products in a way that appeals to the TikTok community's preferences. Furthermore, the channel seeks to drive sales by showcasing the features and benefits of Lipovitan products and enticing viewers to make purchases. Additionally, Lipovitan prioritizes direct interaction with its audience, using TikTok as a platform to engage with users, address their inquiries,

and gather feedback. By fostering a sense of community and dialogue, Lipovitan aims to build stronger connections with its target demographic.

The Lipovitan Vietnam TikTok channel has achieved notable success since its inception. With a steady following of 71.8K subscribers and numerous videos amassing hundreds of thousands of views, the channel has proven to be an effective platform for promoting the Lipovitan brand. By leveraging TikTok's popularity and engaging content, Lipovitan has successfully built brand recognition, strengthened connections with its audience, and established a vibrant and dynamic brand image. Overall, the channel's impact has been instrumental in driving brand growth and fostering a loyal community of Lipovitan enthusiasts on TikTok.

By utilizing these proprietary media assets, Lipovitan can fully control its content and messaging while creating a consistent and positive experience for customers. This helps them build a strong brand and enhance their recognition in consumers' minds.

Email Marketing: Sending regular emails with information about new products, special offers, and brand news helps Lipovitan maintain customer relationships and encourages repeat purchases.

1.3.2 Paid Media - Advertising and Promotion Tactics

Paid Media, also known as paid advertising, encompasses communication channels through which businesses pay to reach potential customers and promote their products or services. This includes online advertising on platforms such as Google Ads, Facebook Ads, Instagram Ads, YouTube Ads, and TikTok Ads. It also encompasses traditional advertising on TV, radio, newspapers, magazines, and billboards, as well as affiliate marketing, where businesses collaborate with content publishers to promote their products or services on websites or blogs. Influencer marketing, where businesses partner with social media influencers to reach their target audience, is another key component.

Lipovitan has established a diverse and detailed advertising and promotional strategy to reach its target audience. They utilize various channels within paid advertising to ensure their message is effectively disseminated and reaches the intended audience.

One popular paid media channel Lipovitan employs is advertising on television and radio platforms. This approach allows them to reach a large audience, particularly through channels that their target audience frequently watches or listens to. Lipovitan often selects programs, events, or genres that resonate strongly with their consumer targets, such as sports programs, tournaments, or top entertainment shows.

In addition to traditional media, Lipovitan also utilizes online advertising, including websites, mobile apps, and social media networks. This enables them to reach users while they browse the web, use apps, or interact on social media. Lipovitan typically selects platforms with relevant user demographics and employs audience targeting to maximize the effectiveness of their advertising.

Beyond advertising, Lipovitan frequently organizes promotional programs to generate interest and stimulate consumer purchases of their products. These programs may include discounts, gift vouchers, bundled products, or other incentives. They also organize special events such as exhibitions, trade shows, or community engagement programs to enhance customer engagement and attract attention to their brand.

Overall, Lipovitan's advertising and promotional strategy focuses not only on reaching a large consumer audience but also on generating interest and fostering positive interactions. This approach helps promote brand recognition and strengthens Lipovitan's connection with their target audience.

1.3.3 Earned Media - Public Relations and Customer Advocacy

Earned media, in the context of Public Relations (PR) and Customer Advocacy, refers to the recognition and coverage that a brand receives without paying for advertising. Unlike paid media, where a business pays for advertising space, earned media results from PR activities, positive evaluations from customers, articles in the press, and other natural interactions on social media or online platforms. Earned Media typically appears in the following forms:

- Articles: Articles, blog posts, or news website features about the business, its products/services.
- Reviews: Customer reviews of products/services on e-commerce websites, social media, forums, etc.
- Recommendations: Word-of-mouth recommendations from customers to friends, family via social media, email, messaging, etc.
- Social media shares: Likes, comments, shares on business posts on social media platforms.
- Review videos: User-generated videos about their experience with the business's products/services shared on YouTube, Facebook, etc.

Lipovitan has successfully garnered community interest through PR strategies and customer advocacy. One way they achieve this is by organizing events or community interaction programs. Lipovitan often hosts meetups, seminars, or exhibitions on health and healthy lifestyles, where they share information about their products and create opportunities for direct interaction with customers.

Additionally, Lipovitan regularly participates in community activities and charity events. They may sponsor health, sports, or environmental programs and even organize volunteer activities to support the local community. In doing so, Lipovitan creates a positive image and strengthens ties with the community, demonstrating their commitment to social values and the environment.

Regarding earned media through customer advocacy, Lipovitan often encourages consumers to share their positive experiences with their products. They may organize promotions for current users and potential customers, and even offer special incentives for users who share about the product on social media or online forums. In this way, Lipovitan generates a viral marketing effect, enhancing the brand's credibility and trust within the consumer community.

In summary, through PR activities and customer advocacy, Lipovitan has built a strong relationship with the community and established trust and credibility for their brand. This not only attracts interest from the target audience but also creates a dedicated community of supporters for their products and brand.

1.4 Crawl Data

1.4.1 Related Research

Before delving into data collection, let's review previous studies on social media data analysis for energy drink brands. We will focus on 5 key aspects: research objectives, data collection methods, types of data collected, and finally, analysis methods along with their outcomes.

(1) For the article "The extent of energy drink marketing on Canadian social media" (<https://doi.org/10.1186/s12889-023-15437-w>):

- Research Objectives: The study aims to describe the social media marketing of caffeinated energy drinks (CEDs) by estimating the frequency of user-generated and company-generated marketing on social media and analyzing the marketing techniques used by CED brands in Canada.
- Data Collection Methods: Data on the frequency, reach, and engagement of posts related to CEDs created by users and Canadian CED brands on Facebook, Instagram, Twitter, Reddit, Tumblr, and YouTube were licensed from Brandwatch for the years 2020–2021.
- Types of Data Collected: The types of data included the number of mentions of CED products by users and businesses on social media platforms, estimated total user reach, and engagement with posts, including likes, shares, comments, and views.

- Analysis Methods: Content analysis was conducted on company-generated social media posts using a coding guide to assess marketing techniques. The analysis included both user-generated and company-generated content.
- Outcomes: There were a total of 222,119 user-level mentions of CED products, reaching an estimated total of 351,707,901 users across platforms. The most common marketing techniques used by Canadian CED brands included viral marketing strategies and themes targeting teenagers. The results indicate that CED companies are promoting their products on social media using viral marketing strategies and themes that appeal to teenagers.

(2) For the article "Exposure to digital marketing enhances young adults' interest in energy drinks: An exploratory investigation":

- Research Objectives: This study aims to explore the impact of digital marketing on young adults' attitudes and intentions to purchase and consume energy drinks, using a pre-and post-test experimental control model.
- Data Collection Methods: The experiment involved randomly dividing participants (aged 18-24) into two groups: a control group and an experimental group. The experimental group was exposed to the websites and social media of two popular energy drink brands (Red Bull and V Energy) for 8 minutes.
- Types of Data Collected: Data included participants' attitudes towards energy drinks before and after exposure to marketing content, their intentions to purchase and consume energy drinks, along with responses from post-exposure interviews.
- Analysis Methods: Content analysis was used based on field notes taken during interviews to identify key factors leading to attitude changes post-experiment. Quantitative data were analyzed using statistical methods such as independent t-tests and the Mann-Whitney U test.
- Outcomes: Exposure to digital marketing content of energy drinks improved the experimental group's attitudes, intentions to purchase, and consumption of energy drinks. The study indicates the strong influence of online marketing on young adults, even when they have developed awareness of the persuasive intent of marketers.

(3) For the article "A thematic content analysis of how marketers promote energy drinks on digital platforms to young Australians":

Research Objectives: The goal is to explore the nature, scale, and user engagement with the appealing strategies used by the food industry to advertise energy drinks on digital platforms.

Data Collection Methods: The study collected data from the three most popular social media platforms (Facebook, Twitter, and YouTube), advergames on the websites of nine popular energy drink brands in Australia, and official content on these brands' websites during June 2015. Data from websites and advergames were collected over two days in July 2015.

Types of Data Collected: The data included 624 textual and visual elements extracted from Facebook posts, Twitter tweets, and YouTube videos. Each post/tweet/video generated one textual and one visual element, as each post included an image or video as well as a caption. Additionally, 15 advergames were identified from the websites.

Analysis Methods: Thematic content analysis was conducted to code the textual and visual elements. The analysis included both quantitative and qualitative methods to provide descriptive statistics on the frequency of advertising techniques and thematic analysis of the messages contained in the advertising content.

Outcomes: Four main themes were identified from the digital advertising content of energy drink brands: online social connection, desired social identity, performance enhancement, and mood improvement. Energy drinks were prominently advertised on digital platforms, especially on social media sites, as evidenced by high user engagement (e.g., 'likes' and 'comments').

1.4.2 Data Collection Focus - Prioritizing Owned Media

Based on the analyses from the three articles, a proposed topic for analyzing social media data regarding Lipovitan energy drinks that our team has discussed and decided upon is "Analyzing Lipovitan's Digital Marketing Strategies on Social Media and Their Impact on Young Consumers' Attitudes and Behaviors." This topic aims to explore how Lipovitan uses social media platforms to engage young consumers and the impact of these marketing strategies on their perceptions and consumption behaviors.

Platform	Crawl Tool	Advantages	Disadvantages	Conclusion
Youtube	Python (library: googleapiclient)	Can crawl via API. Data collected includes two files: YouTube data (157 records) and YouTube	Lipovitan does not heavily focus on advertising on this platform, resulting in limited and less diverse data.	Data collected is limited in quantity, but it's structured and usable.

		comments (24-30 records).		
Facebook	Python (library: Selenium)	Can crawl automatically like a real user. Very useful for extracting data since Facebook requires login, button clicks, and other actions.	Time-consuming and requires monitoring. Crawling may fail, requiring a restart if errors occur.	Data collected is relatively comprehensive but the overall value is moderate.
Tiktok	Web Apify	Efficient interaction with TikTok APIs for structured and rapid data collection.	Data might be incomplete due to challenges in expanding comment sections and other interactions.	The data collected is abundant and valuable. TikTok appears to be Lipovitan's most engaging channel.

1.4.3 Overview about dataset

YouTube Data:

The dataset from YouTube crawling contains detailed information about 157 videos related to the Lipovitan brand. Each row represents a specific video and includes the following columns:

1. video_id: Unique ID of the video on YouTube.
2. channel_id: ID of the YouTube channel posting the video.
3. published date: Date and time the video was published.
4. video title: Title of the video.
5. video description: Detailed description of the video content.
6. likes: Number of likes the video received.
7. views: Number of views of the video.
8. comment_count: Number of comments on the video.

The data includes detailed information about viewer interaction levels with the videos such as likes, views, and comment counts. These details can help analyze the popularity and marketing effectiveness of Lipovitan-related videos on the YouTube platform. Additionally, the group also obtained data about comments on the crawled videos, although the number of commenters is relatively low.

Facebook Data:

The dataset obtained contains information about Facebook posts related to the Lipovitan brand. Each row represents a specific post and includes the following columns:

1. Time: Time the post was published.
2. Content: Content of the post.
3. Like Num: Number of likes the post received.
4. Comment Num: Number of comments on the post.
5. Share Num: Number of shares of the post.
6. URL: Link to the post on Facebook.

This data provides an overview of user interaction levels with the posts, including likes, comments, and shares, as well as detailed content of each post. Additionally, the group obtained data about comments on the crawled posts, which helps the group understand user feedback and opinions through comments on Facebook posts.

TikTok Data:

The data collected from TikTok using Apify includes detailed information about a series of TikTok videos. Each row in the dataset represents a separate TikTok video and includes attributes such as video quality, format, size, and various URLs for cover images, downloads, and subtitles. Additionally, the data includes social interaction metrics such as shares, comments, views, and digg counts, providing deep insights into the popularity and interaction of each video. Text descriptions of the videos are also provided, often containing hashtags and promotional content. Furthermore, the data records the creation time of each video in ISO format. Moreover, the group can easily crawl TikTok comments with Apify. The data from the cmt_apify.csv file contains information about comments on TikTok videos. Each row in the dataset represents a single comment, including the comment content and a link to the TikTok video where the comment appears. The comment content reflects viewer responses and opinions about the video, while the link allows direct access to the corresponding video. This data provides deep insights into the level of interaction and reactions of viewers to videos on TikTok.

CHAPTER 2: DATA PROCESSING

2.1 Data processing and EDA của Youtube Lipovitan

2.1.1 Data processing

This section details the data processing workflow for YouTube data related to Lipovitan. The process involves collecting, cleaning, and transforming data to prepare it for Exploratory Data Analysis (EDA).

Data collection starts with using the YouTube API to gather information related to Lipovitan videos. Key information collected includes video ID, title, description, published date, view count, like count, comment count, and share count. This helps create a comprehensive and rich database of activities and interactions surrounding Lipovitan videos on YouTube.

After data collection, data cleaning is a crucial step to ensure the accuracy and consistency of the data. First, the published dates of the videos (`published_date`) are converted into datetime format for easier processing and analysis later on. Next, the data is checked for duplicate entries. Specifically, the rows are examined based on `video_id` to identify and remove duplicate videos, ensuring that each video appears only once in the dataset. The number of rows before and after removing duplicates is recorded to confirm the effectiveness of the cleaning process.

In addition to checking and removing duplicate rows, handling missing values is another important part of data cleaning. Missing values can be replaced with mean values, default values, or the rows with missing values can be removed, depending on the importance of the missing data.

The next step is data transformation to standardize and prepare the data for deeper analysis. This process includes normalizing the values in the columns to ensure consistency, such as converting text values into a standard form (lowercase, removing special characters, etc.). Additionally, creating new columns from the original data is crucial. For instance, creating columns that represent the day, month, and year from the `published_date` column makes it easier to analyze trends over time. Time processing might also involve converting the publishing times into appropriate time zones or specific time formats to fit subsequent analyses.

video_id	channel_id	published_date	video_title	video_description	likes	views	comment_count
0	fogz9JpsU8	UC-NVmcQ2iV7c_F_vg92uyMg	2021-03-24 LIPOVITAN - Nước Tăng Lực Mang Thương Hiệu Nhâ...	LIPOVITAN - Một trong những thương hiệu nước t...	22	7673	0
1	JYztFHCrxsU	UC-NVmcQ2iV7c_F_vg92uyMg	2023-12-08 ✨ Tết Rồng Phúc Quý – Năng Lượng Như Ý 🌟	Tết đến cận kề, bầu không khí rộn ràng khắp ph...	11	2514	1
2	Fbdtn7S02jc	UC-NVmcQ2iV7c_F_vg92uyMg	2023-12-13 Tết 2024 Việc Gi Khó Có Lipo	⭐ Tết 2024 Việc Gi Khó Có Lipo 🙏\n\nTết Nguy...	10	3431949	0
3	klHx2mfqTRk	UC-NVmcQ2iV7c_F_vg92uyMg	2021-03-24 LIPOVITAN - Nước Tăng Lực Từ Nhật Bản	👉 LIPOVITAN – CHỊNH PHỤC BẠN SAU 3 LON\n👉 Bạn c...	39	15313	10
4	MIHc5U4Ki80	UC-NVmcQ2iV7c_F_vg92uyMg	2023-12-25 TẾT 2024 TẾT KHÔNG SỢ KHÓ, ĐÃ CÓ LIPO!	Tết đến nơi rồi, ôi dzời ôi các anh minh o!!!!....	1	601	0

Figure 2.1 Data after processing

By performing these steps, the data is fully prepared and ready for the following Exploratory Data Analysis (EDA). This data processing workflow not only ensures the accuracy and consistency of the data but also optimizes the analysis and helps derive meaningful insights about the activities and interactions surrounding Lipovitan videos on YouTube.

2.1.2 EDA (Exploratory Data Analysis)

The Exploratory Data Analysis (EDA) of YouTube data for Lipovitan involves examining and visualizing the data to uncover patterns, trends, and insights that can inform strategic decisions. The following sections highlight the key aspects of the EDA process.

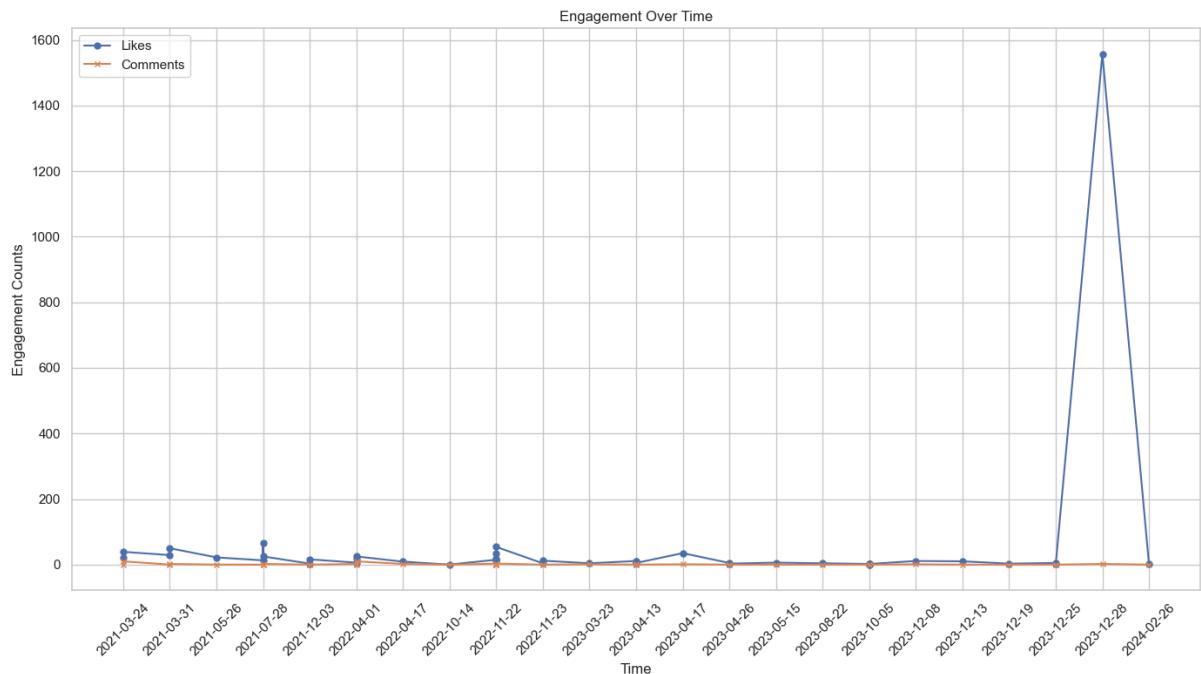


Figure 2.2 Engagement over time

The first visualization (Figure 2.2) shows the engagement metrics, specifically likes and comments, over time. The line plot indicates the fluctuation in user engagement with the videos across different dates. From the graph, it is evident that most videos received relatively low engagement, with a notable spike around the end of December 2023. This significant peak could be attributed to a special event, campaign, or viral content that resonated with the audience during that period. Analyzing such spikes helps in understanding the factors that drive high engagement and can guide future content planning to replicate such success.

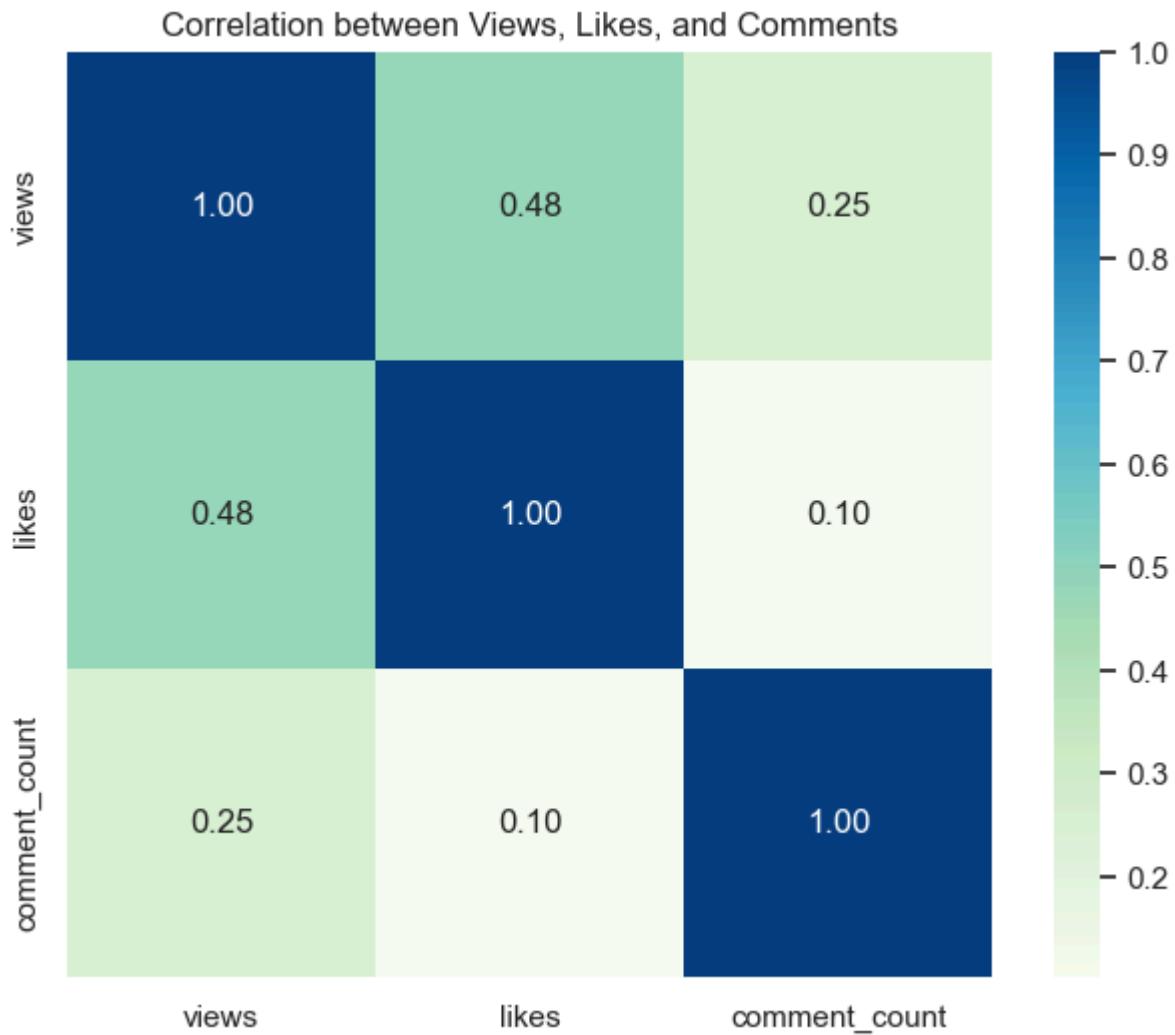


Figure 2.3 Heatmap of data

The second visualization (Figure 2.3) is a heatmap displaying the correlation between views, likes, and comments. The correlation matrix shows that there is a moderate positive correlation (0.48) between views and likes, suggesting that videos with more views tend to receive more likes. However, the correlation between views and comments (0.25) and between likes and comments (0.10) is relatively weak. This indicates that while viewership and likes are somewhat related, comments do not necessarily follow the same pattern. Understanding these relationships helps in identifying which metrics are most interconnected and can guide the focus of future content strategies.



Figure 2.4 WorldCloud of title

The third visualization (Figure 2.4) is a word cloud generated from the titles of Lipovitan videos. The word cloud highlights the most frequently used words in the video titles, with larger words indicating higher frequency. Prominent words such as "Tết," "Việc," "Khó," "Gì," and "Có" suggest common themes or topics addressed in the videos. This analysis provides insights into the content focus and can help in identifying popular themes that attract viewers. It also aids in optimizing video titles to align with audience interests and improve discoverability.



Figure 2.5 WordCloud of description

The fourth visualization (Figure 2.5) is a word cloud generated from the descriptions of Lipovitan videos. This word cloud shows the frequent terms used in the video descriptions, such as "Lipovitan," "Tết," "Việc," "năng lượng," and "khó." These terms indicate the emphasis on energy, work, and the Tết holiday, reflecting the brand's messaging and target audience interests. Understanding the common keywords in descriptions helps in crafting more engaging and relevant content that resonates with viewers.

The exploratory data analysis (EDA) of YouTube data for Lipovitan reveals several key insights that can significantly inform the brand's future content strategies. Firstly, the engagement trends over time provide a clear picture of how viewer interaction fluctuates across different periods. By identifying specific time frames where there is a noticeable spike in likes and comments, such as the significant increase observed in late December 2023, Lipovitan can delve deeper into the factors that contributed to these peaks. This could involve examining the content type, themes, or promotional activities that coincided with these spikes. Understanding these drivers allows the brand to replicate successful strategies in future campaigns.

Secondly, the correlation analysis between views, likes, and comments uncovers important relationships among these metrics. A moderate positive correlation between views and likes suggests that higher viewership is generally accompanied by an increase in likes, indicating a level of viewer appreciation and satisfaction. However, the weaker correlations between views and comments and between likes and comments highlight that viewer engagement through comments does not necessarily follow the same pattern as views and likes. This insight suggests that while viewers may watch and like videos, they are less inclined to leave comments, which could guide Lipovitan to explore strategies to encourage more interactive viewer participation through comments.

Lastly, the word cloud visualizations for video titles and descriptions offer a comprehensive look into the prevalent themes and keywords associated with Lipovitan's content. Words such as "Tết," "Việc," "năng lượng," and "khó" frequently appear, indicating the brand's focus on topics related to energy, work, and the Tết holiday. This thematic consistency helps Lipovitan understand what resonates most with their audience, allowing them to tailor their content to match viewer interests and preferences more closely. By leveraging these insights, Lipovitan can craft video titles and descriptions that not only attract more viewers but also align with the audience's expectations and interests, ultimately enhancing viewer engagement and satisfaction.

2.2 Data processing and EDA của Facebook Lipovitan

2.2.1 Data processing

In this section, we focus on the data processing procedures applied to the Facebook data for Lipovitan. The primary goal is to clean and transform the data to ensure consistency and readiness for exploratory data analysis (EDA).

Firstly, we addressed the 'Time' column, which initially contained both dates and specific times in a mixed format. We removed specific times from the entries and standardized the dates. This involved manually correcting the first few entries to include the year '2024' where it was missing. Next, we mapped month names to their corresponding numerical values to facilitate easier date parsing. A function, `transform_month`, was created to replace month names with numbers. The 'Time' column was then converted into a datetime format and sorted chronologically to ensure proper sequence of posts.

After transforming the 'Time' column, we focused on numerical values representing engagement metrics. The 'Like Num' column contained likes represented in both numerical and abbreviated 'K' formats. To standardize this, a transformation function was employed to convert all values into integers, multiplying where necessary to account for the 'K' abbreviation.

Similarly, 'Comment Num' and 'Share Num' columns required cleaning. These columns included values suffixed with textual indicators like 'comments' and 'shares'. A transformation function was utilized to strip these suffixes and convert the values into integers.

The final dataset, with cleaned and transformed columns, provided a structured and consistent format ready for further analysis. This preprocessing step was crucial in ensuring the accuracy and reliability of subsequent exploratory data analysis, laying a solid foundation for deriving meaningful insights from the Facebook data for Lipovitan.

	Time	Content	Like Num	Comment Num	Share Num	URL
0	2022-08-08	MỚI - NƯỚC TĂNG LỰC ĐẦU TIÊN CÓ CHIẾT XUẤT CÁ...	23.0	6	4	https://www.facebook.com/Lipovitan.vn/posts/pf...
1	2022-08-08	SINH LỰC DỒI DÀO – CHINH PHỤC ĐỈNH CAO\nSau n...	502.0	38	17	https://www.facebook.com/Lipovitan.vn/posts/pf...
2	2022-08-08	NĂNG CHỐI CHANG – TONGKAT ALI SẴN SÀNG\nVào n...	25.0	13	13	https://www.facebook.com/Lipovitan.vn/posts/pf...
3	2022-08-16	Món Quà Từ "1 Lon Nước Nhỏ - 1 Cộng Đồng To"\...	31.0	2	7	https://www.facebook.com/Lipovitan.vn/posts/pf...
4	2022-08-22	Một Sự Đóng Góp Bằng Vạn Điều Sé Chia\nĐi học ...	69.0	4	2	https://www.facebook.com/Lipovitan.vn/posts/pf...
5	2022-08-22	Món Quà Từ "1 Lon Nước Nhỏ - 1 Cộng Đồng To"\...	175.0	11	2	https://www.facebook.com/Lipovitan.vn/posts/pf...
6	2022-08-22	Tạo Năng Lực – Tiếp Sức Đến Trường\nVới ni...	91.0	11	5	https://www.facebook.com/Lipovitan.vn/posts/pf...
7	2022-08-22	Hành Trình Trao Giác Mơ Đến Trường\nChương tr...	30.0	2	5	https://www.facebook.com/Lipovitan.vn/posts/pf...
8	2022-09-06	Tongkat Ali Và Những Công Dụng Bạn Không Thể ...	12.0	4	8	https://www.facebook.com/Lipovitan.vn/posts/pf...
9	2022-09-06	Bạn Có Biết Đến Loại Nước Tăng Lực làm từ mật...	23.0	2	9	https://www.facebook.com/Lipovitan.vn/posts/pf...

Figure 2.6 Data after processing

2.2.2 EDA (Exploratory Data Analysis)

Exploratory Data Analysis (EDA) is an essential step in the data analysis process to gain a better understanding of the data through statistical techniques and visualization. For Lipovitan's Facebook data, EDA was performed to explore user engagement trends and patterns over the months and days of the week.

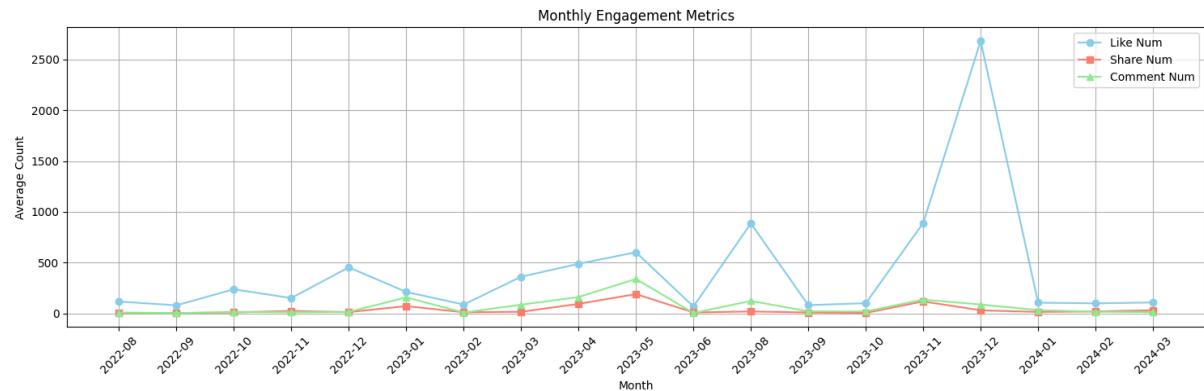


Figure 2.7 Engagement over time

The first chart illustrates the average number of likes (Like Num), shares (Share Num), and comments (Comment Num) on Lipovitan's Facebook posts from August 2022 to March 2024. The number of likes shows fluctuations across the months, particularly in December 2023, where likes spiked to over 2500, possibly due to promotional campaigns or special events during this month. Shares also fluctuated but were less volatile compared to likes, with no significant spikes in any month. Similarly, the number of comments had minor fluctuations and no notable peaks. Overall, likes dominated over shares and comments, indicating that Lipovitan's Facebook users tended to express their appreciation for posts more than sharing or commenting on them.

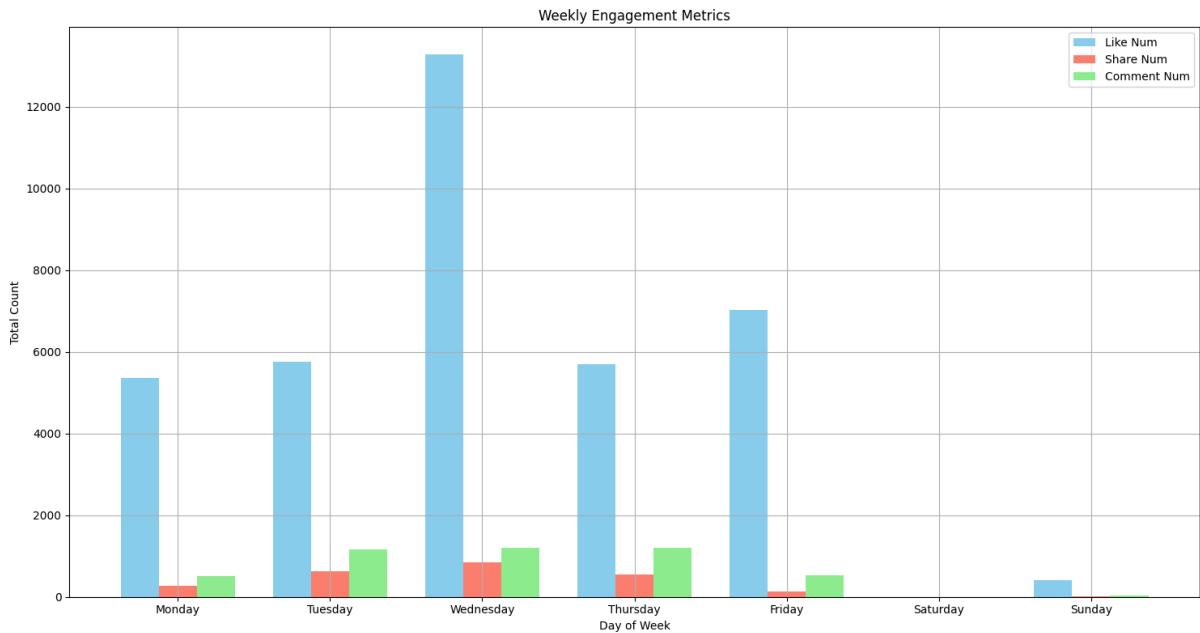


Figure 2.8 Weekly engagement

The second chart shows the total number of likes, shares, and comments on Lipovitan's Facebook posts by day of the week. Likes peaked on Wednesday with over 12,000, followed by Tuesday and Friday, while Sunday had the lowest number of likes. Shares did not show significant variations by day, but Tuesday and Wednesday tended to be higher than other days. Comments were also highest on Wednesday and Tuesday, similar to the trend for likes, with Sunday having the fewest comments. This analysis suggests that user engagement is highest mid-week (Tuesday, Wednesday) and decreases towards the weekend (Saturday, Sunday).

From the EDA, we identified several important trends in user engagement on Lipovitan's Facebook, which can help guide content strategies and marketing activities in the future.

2.3 EDA của Tiktok Lipovitan

In the Exploratory Data Analysis (EDA) of comments on Tiktok related to the Lipovitan product, we examined three main aspects: the distribution of comment lengths, the number of words in each comment, and the number of comments per post.

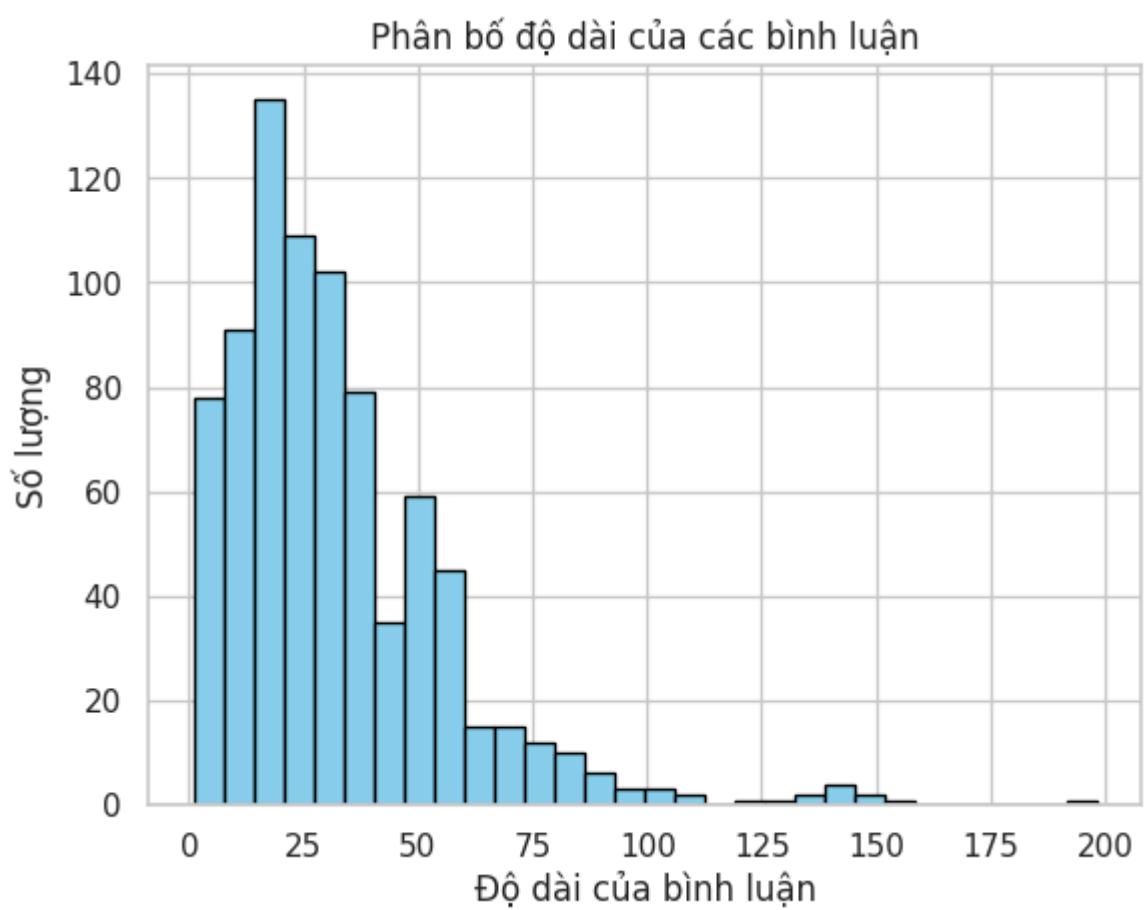


Figure 2.9 Comment length distribution

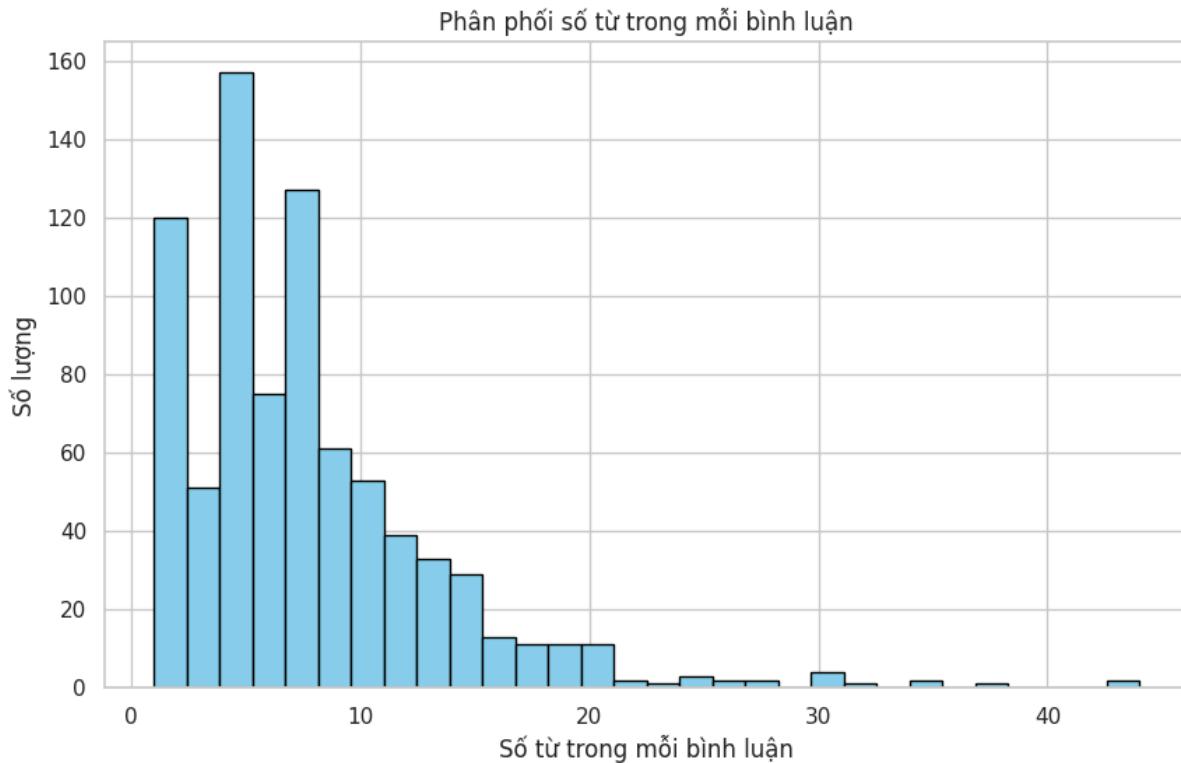


Figure 2.10 Number distribution from each comment

Firstly, from the distribution of comment lengths, it can be seen that the majority of comments are less than 50 characters long. Specifically, the number of comments with lengths between 0 and 50 characters is predominant, then gradually decreases as the comment length increases. This indicates that users tend to write short and concise comments. Next, the distribution of the number of words per comment shows a similar trend. Most comments have between 0 and 10 words, with the highest proportion being those with 0 to 5 words, then gradually decreasing as the word count increases. This is consistent with the distribution of comment lengths, confirming that users often leave short comments.

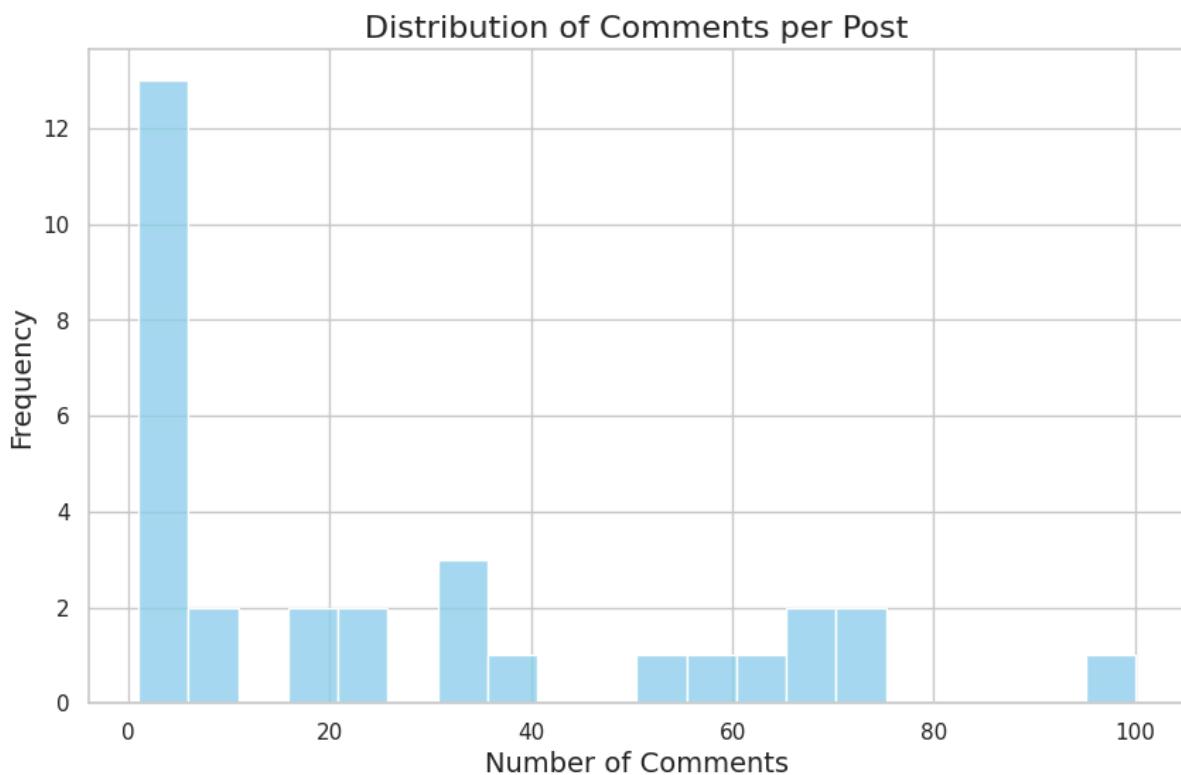


Figure 2.11 Distribution of comments per post

Finally, the distribution of the number of comments per post shows that most posts receive fewer than 20 comments. There are a few posts with a higher number of comments, but these are quite rare. This indicates that the level of engagement on the posts is dispersed, with some posts receiving significantly more comments than others.

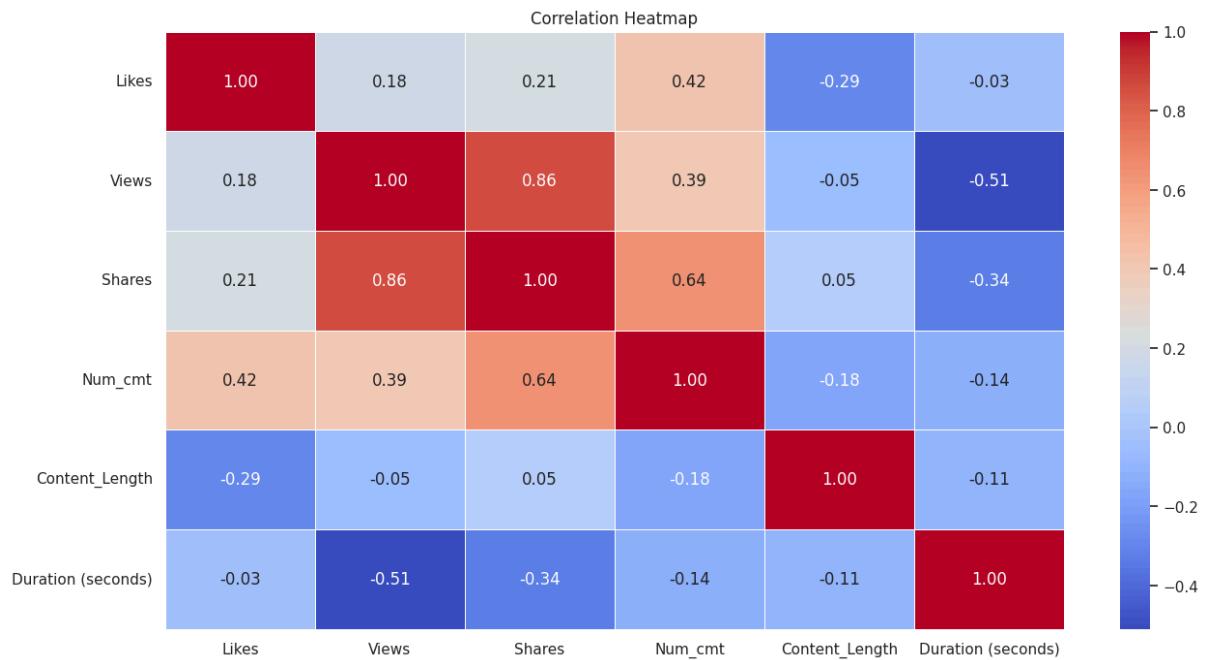


Figure 2.12 Correlation heatmap

Combined with the heatmap, we can better understand the relationships between variables related to likes (Likes), views (Views), shares (Shares), content length (Content_Length), and video duration (Duration). The heatmap indicates a strong positive correlation (0.86) between Likes and Views, Likes and Shares, and Views and Shares. This suggests that posts with many likes typically have many views and shares. Conversely, there is a moderate negative correlation (-0.29) between Content_Length and Likes, as well as Content_Length and Views, suggesting that longer content may not necessarily receive more likes or views. Finally, Duration has a low correlation with the other variables, indicating that video length has little impact on the other metrics.

CHAPTER 3: SENTIMENT ANALYSIS

3.1 NLP Model

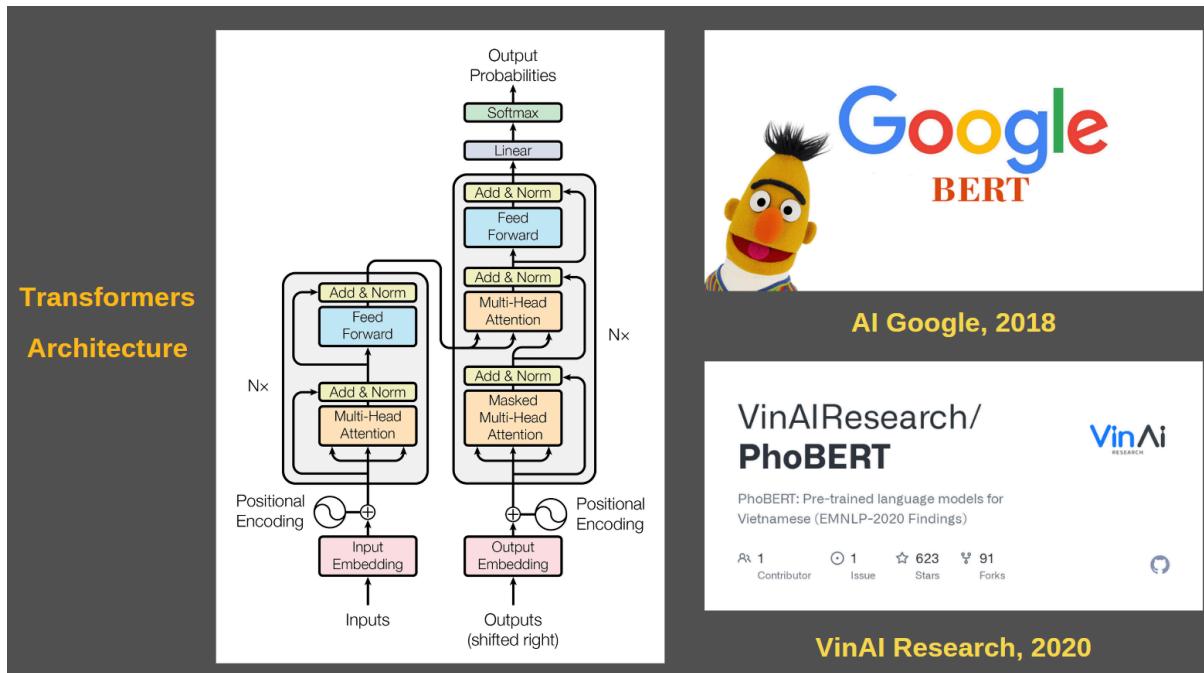


Figure 3.1 Transformers Architecture and Bert Model

Choosing a model to analyze comments on social media platforms like Facebook, YouTube, and TikTok is crucial for achieving accurate results in sentiment analysis, topic categorization, and understanding user engagement. For Lipovitan, a well-known energy drink brand, analyzing customer comments can provide valuable insights into consumer sentiments and preferences.

BERT (Bidirectional Encoder Representations from Transformers) is one of the most prominent models widely used in the field of Natural Language Processing (NLP). Developed by Google AI in 2018, BERT is based on the Transformers architecture, a structure that has revolutionized NLP by allowing models to handle long-term dependencies in text. The Transformers architecture consists of an encoder and a decoder, each with multiple layers of multi-head self-attention mechanisms and feed-forward neural networks. This allows BERT to understand the context of words within the entire sentence, thus capturing meaning more accurately than traditional models.

PhoBERT is a variant of BERT specifically designed for Vietnamese. Developed by VinAI Research in 2020, PhoBERT leverages the architecture and advantages of BERT but is fine-tuned on Vietnamese texts. This makes PhoBERT highly suitable for understanding the nuances and expressions specific to Vietnamese comments. The

pre-trained versions of PhoBERT are available on the Hugging Face website, making it easy to implement—simply input the data to get the analysis results.

To analyze Lipovitan's comments on social media platforms, the following steps are needed: Collect comments from Lipovitan's official pages and videos on Facebook, YouTube, and TikTok; clean the data to remove noise such as emojis, special characters, and URLs; use PhoBERT for all Vietnamese comments to analyze sentiment and extract key topics; aggregate the results to derive useful insights about consumer sentiments, common discussion topics, and potential areas for product improvement or marketing strategies.

By leveraging advanced NLP models like PhoBERT, Lipovitan can gain a deeper understanding of customer feedback across different platforms, thereby enhancing their interaction strategies and product offerings.

3.2 Facebook sentiment analysis

The diagram illustrates the flow of data analysis. It starts with a box labeled "Raw Data" with an arrow pointing to a box labeled "Data Info". From "Data Info", an arrow points to a code snippet showing a pandas DataFrame summary. Below this is a table visualization of the data.

	Comment	Link
0	Nguyễn Hương	https://www.facebook.com/Lipovitan.vn/posts/pf...
1	Revu Vietnam	https://www.facebook.com/Lipovitan.vn/posts/pf...
2	Top fan\nTrần Nam	https://www.facebook.com/Lipovitan.vn/posts/pf...
3	Nguyễn Hiền	https://www.facebook.com/Lipovitan.vn/posts/pf...
4	Thanh Lê\nQuá tuyệt vời, vừa chất lượng vừa đẹ...	https://www.facebook.com/Lipovitan.vn/posts/pf...
5	Thanh Lê	https://www.facebook.com/Lipovitan.vn/posts/pf...

Figure 3.2 Overview data comment on Facebook

This report presents an analysis of data from comments on Lipovitan's Facebook page, collected from the beginning of 2021 to the present. The data is organized in a DataFrame using the Python pandas library, allowing us to gain a deeper understanding of user interactions with the brand. The dataset includes 1436 posts with 1405 comments, each reflecting user responses and linked directly to the original Facebook post via an attached URL.

The purpose of this analysis is to assess customer response trends, consumer sentiment, and interaction levels over different periods. This information supports the

development of marketing and promotional strategies, while helping Lipovitan adjust its business strategies to better meet customer needs and expectations. By thoroughly analyzing the comments, we can gain insights into how users interact with the brand, thereby contributing to the improvement and optimization of Lipovitan's marketing activities.

#	Column	Non-Null Count	Dtype
0	Comment	1380	non-null object
1	Link	1380	non-null object
dtypes: object(2)			
memory usage: 32.3+ KB			

	Comment	Link
0	Nguyễn Hương	https://www.facebook.com/Lipovitan.vn/posts/pf...
1	Revu Vietnam	https://www.facebook.com/Lipovitan.vn/posts/pf...
2	Top fan\nTrần Nam	https://www.facebook.com/Lipovitan.vn/posts/pf...
3	Nguyễn Hiền	https://www.facebook.com/Lipovitan.vn/posts/pf...
4	Thanh Lê\nQuá tuyệt vời, vừa chất lượng vừa đẹ...	https://www.facebook.com/Lipovitan.vn/posts/pf...

#	Column	Non-Null Count	Dtype
0	Time	70	non-null datetime64[ns]
1	Content	70	non-null object
2	Like Num	70	non-null float64
3	Share Num	70	non-null float64
4	Comment Num	70	non-null int64
5	Link	70	non-null object
dtypes: datetime64[ns](1), float64(2), int64(1), object(2)			
memory usage: 3.8+ KB			

	Time	Content	Like Num	Share Num	Comment Num	Link
16	2022-08-08	SINH LỰC ĐỔI DÀO – CHINH PHỤC ĐỈNH CAO\nSau n...	502.0	17.0	32	https://www.facebook.com/Lipovitan.vn/posts/pf...
65	2022-08-08	MỚI - NUỐC TĂNG LỰC ĐẦU TIÊN CÓ CHIẾT XUẤT CÂ...	23.0	4.0	3	https://www.facebook.com/Lipovitan.vn/posts/pf...
72	2022-08-16	Món Quà Từ "1 Lon Nước Nhỏ - 1 Công Đồng To"\...	31.0	7.0	2	https://www.facebook.com/Lipovitan.vn/posts/pf...
68	2022-08-22	Hành Trình Trao Giác Mơ Đến Trường\nChương tr...	30.0	5.0	2	https://www.facebook.com/Lipovitan.vn/posts/pf...
45	2022-08-22	Tạo Năng Lượng – Tiếp Sức Đến Trường\nVới ni...	91.0	5.0	11	https://www.facebook.com/Lipovitan.vn/posts/pf...

Figure 3.3 Data Preprocessing

After collecting user comment data, the team began preprocessing the data for a project analyzing comments on Lipovitan's Facebook page. Initially, the dataset consisted of only two columns: "Comment" (the content of the comments) and "Link" (the URL to the original post). From these URLs, the research team was able to identify the total number of unique posts and proceeded to calculate the number of comments, likes, and shares for each post.

Through the preprocessing process, the data table was expanded to include new columns such as "Time" (the time of posting), "Content" (post content), "Like Num" (number of likes), "Share Num" (number of shares), and "Comment Num" (number of comments). To refine the data, the team also addressed null values, ensuring higher integrity and accuracy of the dataset.

The outcome of the preprocessing is a new dataset containing 1380 comments from 70 posts, stored in a DataFrame using 3.8 KB of memory. This data is then used for further analysis to better understand user interactions on Lipovitan's Facebook posts.

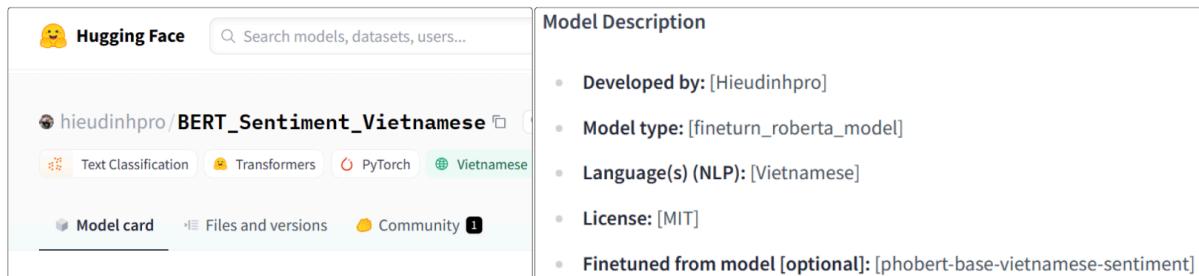


Figure 3.4 Model to Sentiment Analysis

Continuing from the data preprocessing stage, the team utilized a Vietnamese sentiment analysis model from Hugging Face to analyze the collected comments. This model, developed under the name hieudinhpro/BERT_Sentiment_Vietnamese, is a fine-tuned version of phobert-base-vietnamese-sentiment, specially designed to identify emotions in Vietnamese text with three main categories: negative, neutral, and positive.

Modern NLP (Natural Language Processing) technology allows this model to capture overall sentiment trends from the Facebook comment data, highlighting the predominant emotions of consumers towards the posts. The model is highly efficient and flexible, ensuring accuracy in classification and enabling the team to gain deeper insights into user interactions.

The results of this analysis not only provide detailed insights into user reactions on social media but also support Lipovitan in adjusting their marketing strategies and content management to optimize engagement and customer satisfaction.

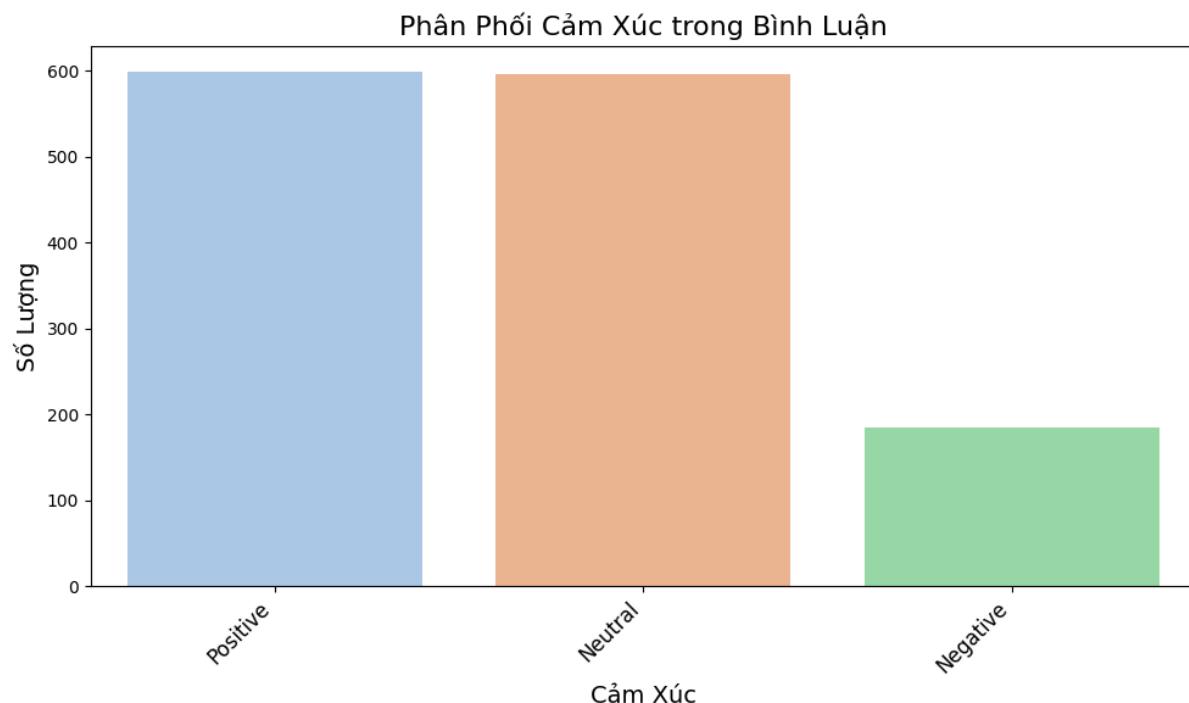


Figure 3.5 Result Sentiment Analysis

The image above displays the results of sentiment analysis from comments on Lipovitan's Facebook page, based on a Vietnamese sentiment analysis model. The bar chart shows the distribution of comments across three emotional categories: positive, neutral, and negative. The majority of comments fall into the positive category, indicating a predominantly positive response from users. The number of neutral comments is also quite high, reflecting objective opinions that do not carry clear emotional tones, while only a small number of comments exhibit negative sentiments. This suggests that overall, the Lipovitan brand is well-received by consumers, with the majority of online community feedback being positive or neutral, and only a small group expressing dissatisfaction. This information can assist Lipovitan in refining their marketing strategies to enhance customer engagement and satisfaction.

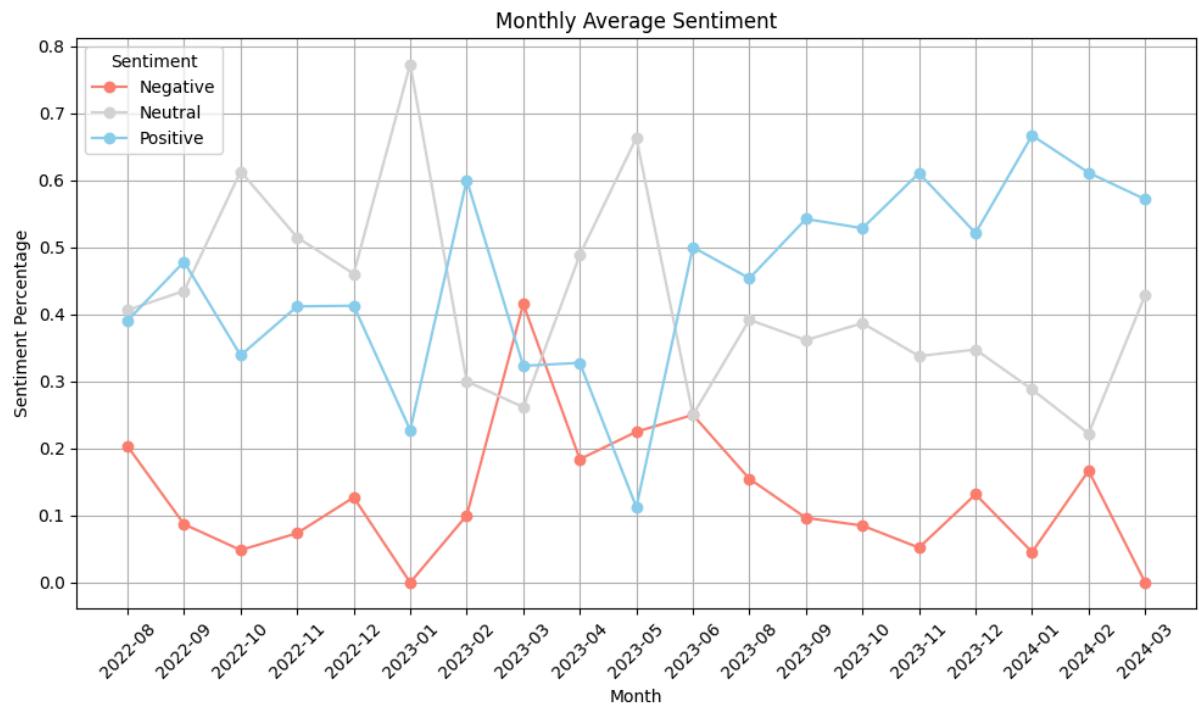


Figure 3.6 Monthly Average Sentiment

The chart above displays the sentiment trends of Facebook comments over time, categorized into three emotional states: positive, neutral, and negative, from August 2020 to March 2023. It is evident that the number of positive comments has significantly increased, particularly starting from early 2022, indicating that users are becoming increasingly satisfied with Lipovitan's content or products. Meanwhile, the level of neutral sentiment remains stable, reflecting objective opinions that are not strongly emotional. Notably, the amount of negative feedback has steadily decreased since early 2021, suggesting that Lipovitan may have addressed previous issues causing customer dissatisfaction or has been successful in effectively resolving negative feedback. These results are a positive indication that Lipovitan is making progress in enhancing product and service quality and improving interactions with

customers on social media, thereby strengthening the brand image and customer relationships.

3.3 Tiktok sentiment analysis

#	Column	Non-Null Count	Dtype
0	Content	31 non-null	object
1	Published Date	31 non-null	object
2	Likes	31 non-null	int64
3	Views	31 non-null	int64
4	Shares	31 non-null	int64
5	Num_cmt	31 non-null	int64
6	Duration (seconds)	31 non-null	int64
7	URL	31 non-null	object

	Content	Published Date	Likes	Views	Shares	Num_cmt	Duration (seconds)	URL
0	Cùng Thé Anh Chạy Bộ đi cùng Lipovitan #LIPOVITAN...	2024-01-02T06:05:21.000Z	186	138900	2	2	80	https://www.tiktok.com/@lipovitan.vn/video/731...
1	Cùng Nhung và Hùng đi cùng Lipovitan #LIPOVITAN...	2024-01-02T06:04:34.000Z	4517	1500000	32	101	57	https://www.tiktok.com/@lipovitan.vn/video/731...
2	Cùng Nai Nguyễn đi cùng Lipovitan #LIPOVITAN #...	2024-01-02T06:03:46.000Z	39	2460	3	1	87	https://www.tiktok.com/@lipovitan.vn/video/731...
3	Cùng Min Mèo đi cùng Lipovitan #LIPOVITAN ...	2024-01-02T05:39:25.000Z	3258	1200000	17	7	201	https://www.tiktok.com/@lipovitan.vn/video/731...
4	Cùng Kiến Không Ngủ đi cùng Lipovitan #LIPOVITAN...	2024-01-02T05:38:38.000Z	22600	5600000	143	326	85	https://www.tiktok.com/@lipovitan.vn/video/731...

Figure 3.7 Overview data of TikTok

The research team continues to analyze user interactions on the TikTok platform, specifically through the videos posted on Lipovitan's profile. Data from 31 videos were collected and systematically organized to assess interaction effectiveness. This allows us to gain deeper insights into user engagement with video content. The dataset includes details for each video such as the posting date, number of likes, views, shares, comments, video duration, and direct URL to the video.

The purpose of this analysis is to evaluate customer response trends, consumer sentiment, and engagement levels over different periods. This information supports the development of marketing and promotional strategies, as well as helps adjust business strategies to better meet customer needs and expectations. By thoroughly analyzing video data, we can gather insights into how users interact with the content, thereby contributing to the improvement and optimization of the brand's marketing activities on TikTok.

Figure 3.8 Overview comment data of Tiktok

The results shown provide an overview of the comment data collected from TikTok videos posted on Lipovitan's profile. The dataset includes a total of 811 comments gathered from 31 videos. The data structure indicates that each entry consists of two columns: "Comment" and "Link." The "Comment" column contains the text of user comments, while the "Link" column provides direct URLs to the respective TikTok videos. The dataset comprises 811 non-null entries for both comments and links. Examples of comments range from user reactions like "quá đă" (awesome) to product feedback and other user interactions. This slide visually supports the earlier discussion on user engagement analysis by highlighting the volume and variety of comments collected, which will be further analyzed to understand user sentiment and interaction patterns.

		Link	Num_Cmt
0	https://www.tiktok.com/@lipovitan.vn/video/731...		2
1	https://www.tiktok.com/@lipovitan.vn/video/731...		1
2	https://www.tiktok.com/@lipovitan.vn/video/731...		68
3	https://www.tiktok.com/@lipovitan.vn/video/731...		33
4	https://www.tiktok.com/@lipovitan.vn/video/731...		100

Figure 3.9 Statistic Comment Tiktok

The statistical analysis of the comments shows that from 31 videos, a total of 811 comments were collected. On average, each video received 26 comments. The video with the highest number of comments garnered 100 comments, while the video with the least number of comments received just 1 comment. This detailed breakdown helps in understanding the distribution and engagement level of each video, providing insights for optimizing future content and marketing strategies on TikTok.

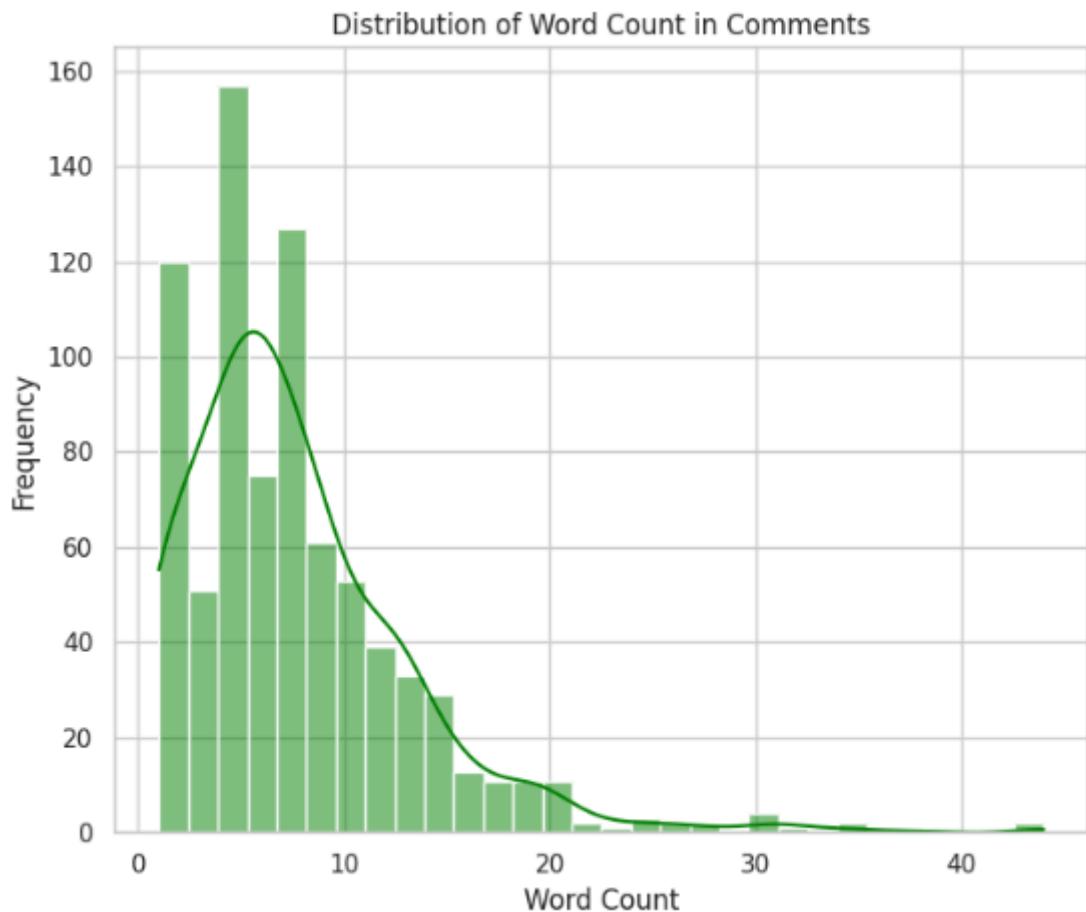


Figure 3.10 Distribution of Word Count in Comments

Further analysis of the word count in the comments reveals additional insights. The distribution chart shows that most comments consist of 1 to 10 words, with the majority concentrated around 5 to 7 words. Short comments could indicate quick and subjective reactions, possibly driven by prompts from Lipovitan, thus appearing more subjective. Longer comments, however, suggest that consumers are expressing their opinions in more detail, indicating a more objective engagement. This analysis helps in understanding the nature and depth of user interactions, allowing for better content strategy planning to encourage more meaningful and engaged comments.

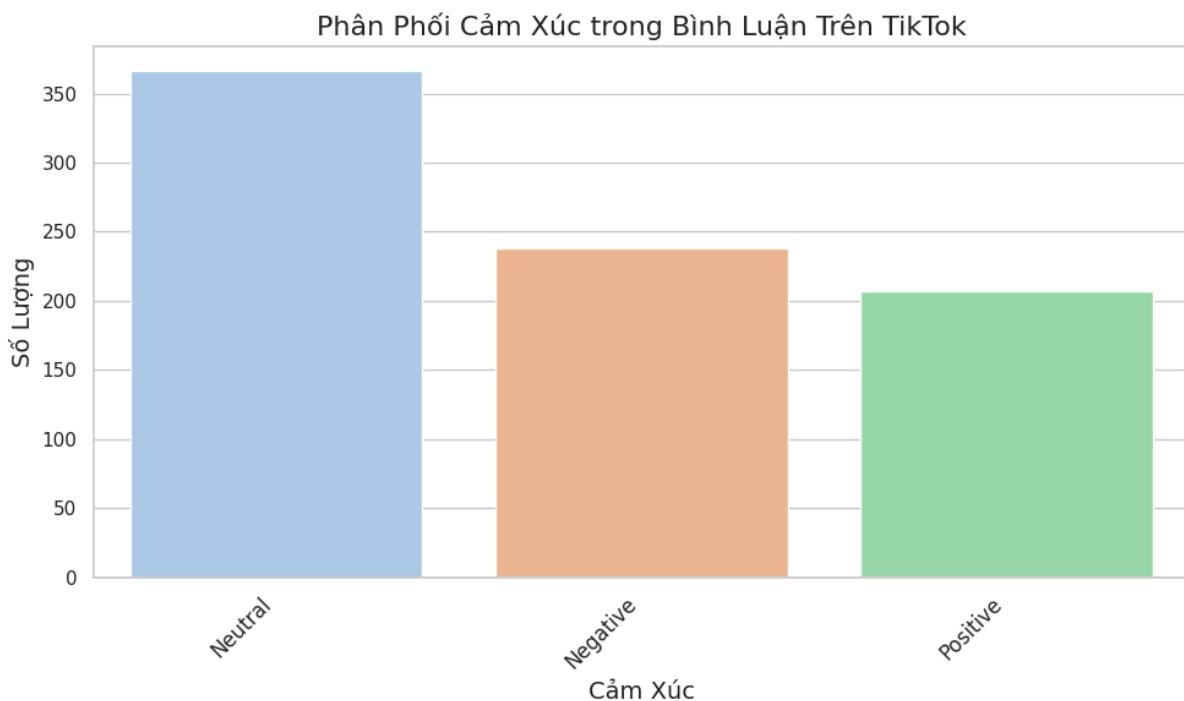


Figure 3.11 Result Sentiment Analysis Tiktok

The chart above shows the distribution of sentiments in TikTok comments, divided into three categories: Neutral, Negative, and Positive.

Detailed analysis:

- Neutral Sentiment: The number of neutral comments is the highest, with over 350 comments. This indicates that many users tend to leave comments that do not express strong emotions or clear opinions about the product or video content.
- Negative Sentiment: The number of negative comments is second highest, with around 250 comments. This is quite a significant number and could be a sign that there are issues or shortcomings that users are dissatisfied with.
- Positive Sentiment: The number of positive comments is the lowest, with fewer than 200 comments. This shows that the level of satisfaction and positive feedback from users is not very high.

Most comments are neutral, suggesting that many users do not have strong opinions or do not want to express their personal views clearly. The number of negative comments being higher than positive ones might be due to an inappropriate sentiment analysis model, leading to incorrect or inaccurate classification. However, if the data is accurate, this is a warning sign that product quality or content needs improvement to enhance user satisfaction.

Improvements needed for the future:

- Review and improve the sentiment analysis model: Ensure that the model being used has high accuracy in classifying comments.
- Analyze the causes of negative comments: Look into the details of negative comments to understand specific reasons and implement corrective measures.
- Encourage positive comments: Create campaigns, events, or content that stimulate users to leave more positive comments.

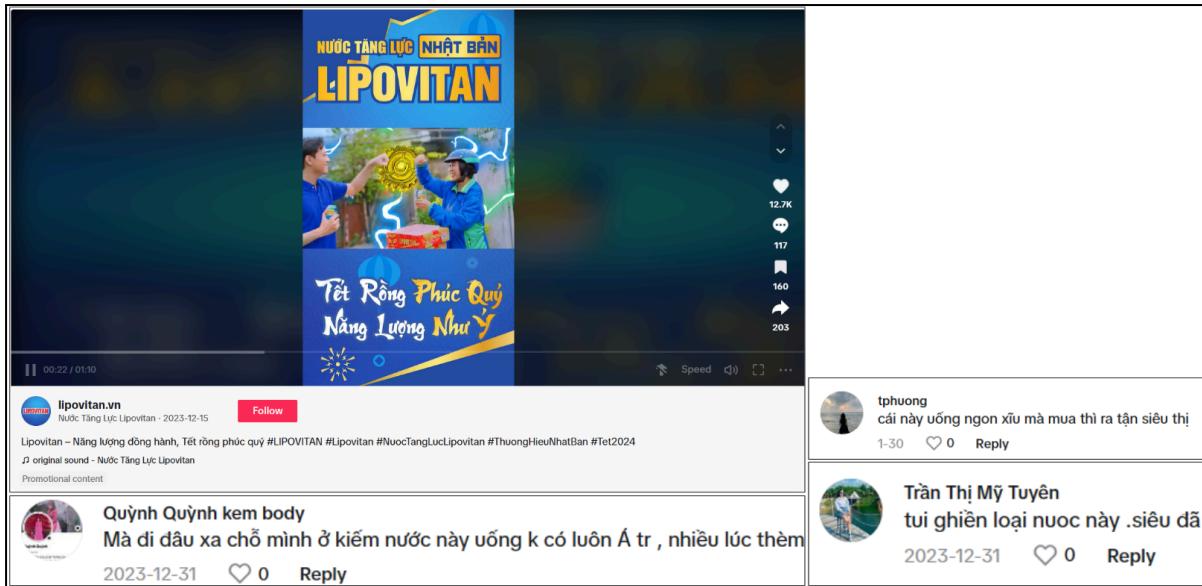


Figure 3.12 The post in top 1 positive

The TikTok post by Lipovitan has garnered significant attention from the user community, with 12.7K views, 117 likes, 160 comments, and 203 shares. Comments from users like 'tphuong' and 'Trần Thị Mỹ Tuyén' express a strong liking for the product, but there are also complaints about the difficulty of finding the product near their residence. This indicates that while the product is highly rated, its accessibility does not meet consumer demand. The high engagement and positive feedback from users show that the product has great market potential but also highlight the challenges related to distribution.

In response to this feedback, Lipovitan should consider improving its distribution network by expanding sales points or enhancing collaboration with online retailers to make the product more accessible to customers. Additionally, using the positive comments from this post in advertising campaigns will not only raise product awareness but also encourage users to share their positive experiences on social platforms, thereby strengthening the brand image and customer satisfaction.



Figure 3.13 The post in top 1 negative

Lipovitan's TikTok post, featuring a bus image and a headline related to the fear of rural dwellers, has garnered significant attention with 22.6K likes and 327 comments. However, the high number of negative comments suggests that the video content may have touched on unpleasant experiences of viewers related to rural trips. This may include inconvenience, overcrowding, or safety issues during long journeys. The negative reaction from the online community reflects a disagreement with the presentation or sensitive issues addressed in the video, requiring careful consideration of content.

In light of this situation, Lipovitan needs to reconsider and adjust its content to ensure it does not evoke negative emotions or bad experiences that the audience may have had. Choosing to address sensitive issues discreetly and thoughtfully can help avoid unwanted reactions and promote a more positive brand image. Additionally, Lipovitan could focus on creating content that encourages optimism, especially in contexts related to homesickness or challenges when away from home, aiming to enhance positive interaction and strengthen customer relationships.

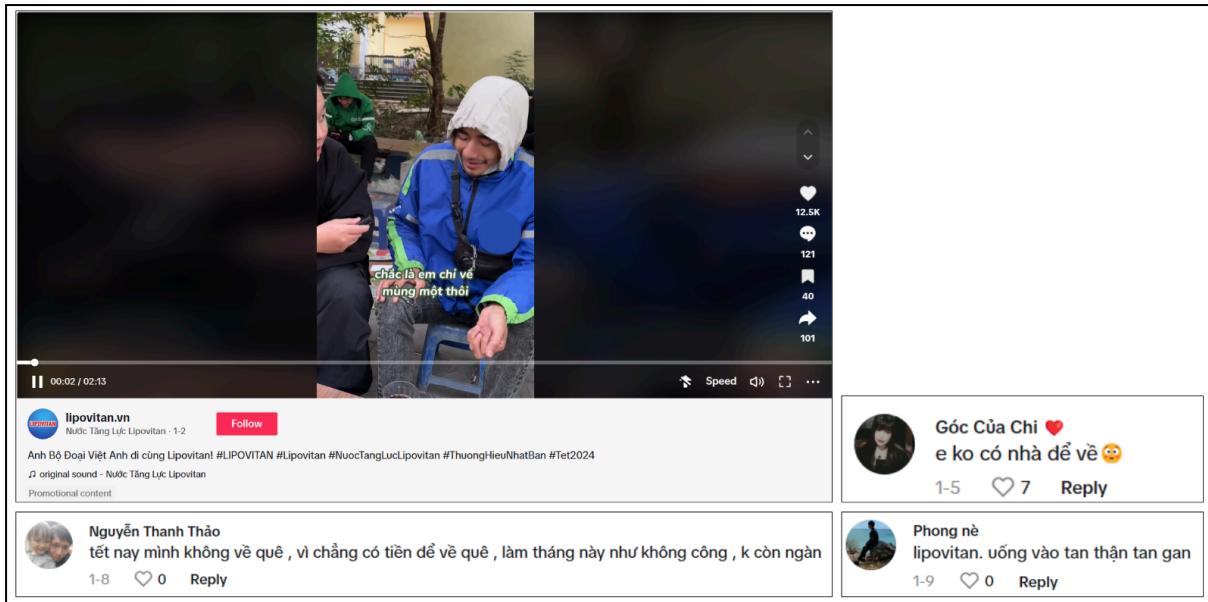
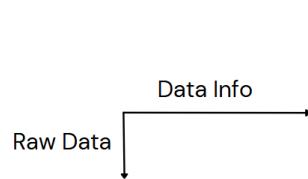


Figure 3.14 The post in top 2 negative

The figure 3.14 depicts another TikTok post by Lipovitan, noted to have some negative comments. However, what stands out is that the negative comments seem to relate more to viewers' personal stories and emotions rather than directly to Lipovitan's product. For instance, one comment mentions that the user has no home to go to during the holidays, while another says they don't have money to return to their hometown. These comments reflect the personal difficulties viewers are experiencing, which may evoke negative emotions when viewing content related to topics such as homeland or family.

Based on this context, Lipovitan could consider strategies to minimize unnecessary negativity and enhance emotional connection with the audience. Firstly, Lipovitan should carefully consider the themes and content of the video to ensure they are appropriate and do not inadvertently trigger negative emotions in particularly sensitive contexts. Secondly, the brand could develop campaigns aimed at creating a positive impact, such as promoting unity and community support, or even providing support initiatives for those in difficult circumstances during the holidays. In doing so, Lipovitan not only improves its image but also contributes to building a stronger relationship with consumers through meaningful actions and positive interactions on social media platforms.

3.4 Youtube sentiment analysis



#	Column	Non-Null Count	Dtype
0	video_id	46 non-null	object
1	published_date	46 non-null	datetime64[ns]
2	video_title	46 non-null	object
3	video_description	46 non-null	object
4	likes	46 non-null	int64
5	views	46 non-null	int64
6	comment_count	46 non-null	int64

video_id	published_date	video_title	video_description	likes	views	comment_count
0	fozg9rJpsU8	2021-03-24 LIPOVITAN - Nước Tăng Lực Mang Thương Hiệu Nhâ...	LIPOVITAN - Một trong những thương hiệu nước t...	22	7673	0
1	JYztFHCrxsU	2023-12-08 🔥 Tết Rồng Phúc Quý – Năng Lượng Như Ý 🌺	Tết đến cận kề, bầu không khí rộn ràng khắp ph...	11	2514	1
2	Fbdtn7S02jc	2023-12-13 Tết 2024 Việc Gi Khô Có Lipo	⚡ Tết 2024 Việc Gi Khô Có Lipo 🎉 \n\nTết Nguy...	10	3431949	0
3	kIHx2mfqTRk	2021-03-24 LIPOVITAN - Nước Tăng Lực Từ Nhật Bản	LIPOVITAN – CHINH PHUC BAN SAU 3 LON!n 🔥 Bạn c...	39	15313	10
4	MIHc5U4Ki80	2023-12-25 TẾT 2024 TẾT KHÔNG SỢ KHÓ, ĐÃ CÓ LIPO!	Tết đến nơi rồi, ôi dzời ôi các anh minh o!!!...	1	601	0
5	GEOa7jzhbmo	2023-04-17 LIPOVITAN - Năng Lượng Cho Cuộc Sống	Với Vitamin nhóm B, Taurine, Inositol cùng vị ...	35	4521852	1
6	5pkEHAoERpY	2021-03-31 LIPOVITAN - BẬT NẮP TRÚNG VÀNG - UỐNG LÀ TRÚNG!	LIPOVITAN – BẬT NẮP TRÚNG VÀNG\nUỐNG LÀ TRÚNG!...	50	6046472	2
7	Afe5ZF7Z22c	2023-03-23 🔥 Lipovitan Bật Nắp Trúng Vàng 2023 *	🔥 Lipovitan Bật Nắp Trúng Vàng 2023 * \nChươn...	4	3393	1
8	HLVs4lolx0k	2021-03-31 LIPOVITAN - Bật nắp trúng vàng - Uống là trúng !	#LIPOVITAN – BẬT NẮP TRÚNG VÀNG \nUU ĐÃI NGÀ...	29	10156	0
9	UGJastVG01g	2023-04-13 🔥 Lipovitan Uống Càng Nhiều Trúng Càng Đã *	🔥 Lipovitan Uống Càng Nhiều Trúng Càng Đã * ...	11	4324220	0

Figure 3.15 Overview data of Video on Youtube

The image provides an overview of the data for videos uploaded by Lipovitan on YouTube. The data table displays a total of 46 videos with columns including video ID (video_id), published date (published_date), video title (video_title), video description (video_description), number of likes (likes), number of views (views), and comment count (comment_count). Each row in the table represents a specific video, allowing for the study of user interaction with each video.

The purpose of this analysis is to evaluate customer feedback trends, consumer sentiment, and engagement levels over different periods. This information supports the development of marketing and advertising strategies, as well as helps adjust business strategies to better meet customer needs and expectations. By thoroughly analyzing the video data, the research team can gain valuable insights into how users interact with the content, contributing to the improvement and optimization of the brand's marketing activities on YouTube.

	likes	views	comment_count
count	46.00	46.00	46.00
mean	47.28	1186218.85	0.89
std	228.07	2096566.81	2.15
min	0.00	60.00	0.00
25%	3.00	816.25	0.00
50%	9.50	2953.50	0.00
75%	21.25	1476146.50	1.00
max	1557.00	7423692.00	10.00

Figure 3.16 Statistic video on Youtube

The statistics table shows significant discrepancies in likes, views, and comments for each Lipovitan video on YouTube. With an average of only about 47 likes and a high standard deviation of 228, it indicates inconsistent viewer interest across different videos. Similarly, views average over 1 million, but some videos have as few as 60 views, showing that certain videos fail to attract attention. Additionally, the low engagement through comments, averaging less than 1 comment per video, suggests that the content may not be compelling enough to encourage viewer feedback.

Based on this analysis, Lipovitan needs to investigate the factors that make some videos more successful than others. Analyzing themes, video quality, and specific marketing strategies for each video could provide insights into optimizing content to attract more views and engagement. Moreover, Lipovitan should seek ways to enhance user interaction, possibly through direct engagement campaigns like contests or polls, to encourage viewers not only to watch but also to leave feedback. Optimizing posting times and using effective keywords will also improve visibility and attractiveness on YouTube, thereby strengthening the brand's online marketing efforts.



Figure 3.17 Top 1 Interaction Score

This slide provides a detailed look at one of Lipovitan's most successful videos on YouTube, measured by the Interaction Score, which is calculated based on the total number of views, likes, and comments. This video, titled "Tết 2024 | Việc Gì Khó Có Lipo!", achieved over 7 million views and 1.5 thousand likes, but also received at least one negative comment. The video was posted on December 28, 2023, right before the Lunar New Year, a significant time in Vietnamese culture when people prepare for the festival and tend to seek positive energy and support.

The success of this video can be attributed to several factors. First, the timing of the release was crucial; posting the video just before Tết likely attracted special attention as people were looking for holiday-related content and preparation tips. Second, the video's content resonated with the cultural context and psychology of the viewers, addressing overcoming difficulties during Tết, which likely built strong empathy and connection with the audience.

To build on the success of this video, Lipovitan should continue to focus on releasing videos during special occasions, tying them to cultural events and significant times. Additionally, thoroughly analyzing viewer feedback, including negative comments, will help the brand better understand audience expectations and potential issues. Finally, enhancing engagement with the audience through contests or direct interactions on social media will strengthen relationships with viewers and increase the appeal of the content.

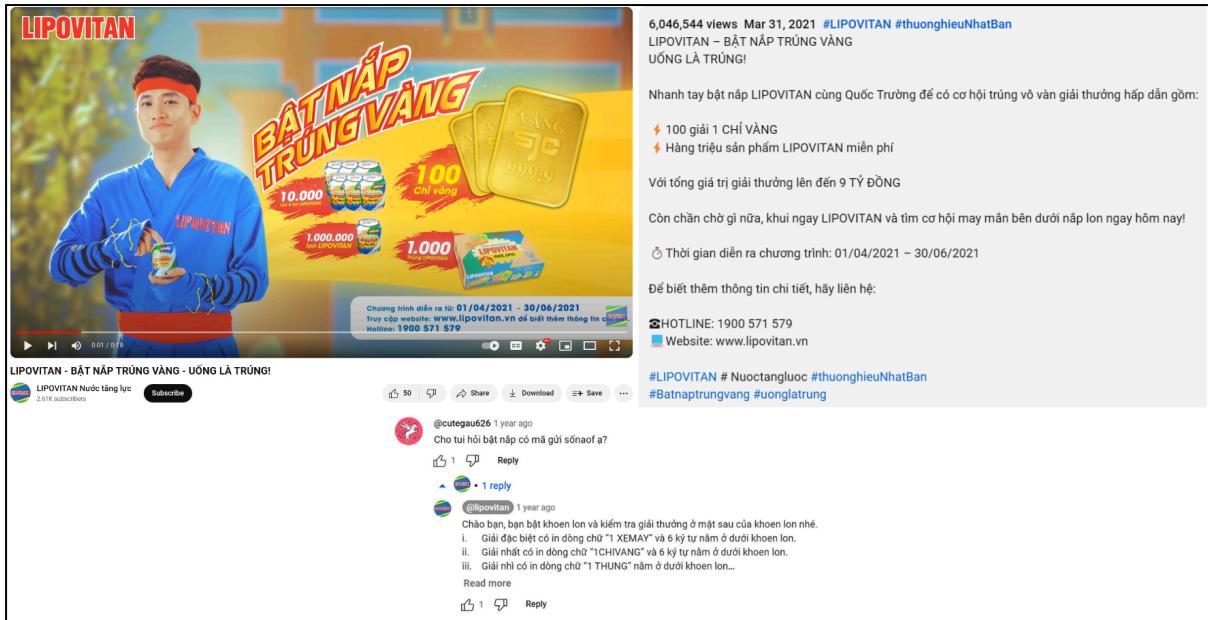


Figure 3.18 Top 2 Interaction Score

This Lipovitan video, titled "BẮT NẮP TRÙNG VÀNG - UỐNG LÀ TRÙNG!", has the second highest Interaction Score among the brand's videos. It garnered over 6 million views, 50 likes, and presumably some comments that were not listed on the slide. The standout feature of this video is its association with the "Bắt nắp trùng vàng" campaign, a promotion where users have the chance to win gold by using the product.

The significant attention this video received is largely due to the attractive promotional campaign, with the grand prize being gold, a highly valuable asset. The combination of an appealing prize and the video's timely release (right before the promotional campaign started) were likely key factors driving views and engagement. Additionally, user feedback indicated interest in the details of the promotion, demonstrating the effectiveness of using valuable prizes to attract attention.

Based on the success of this video, Lipovitan could consider applying a similar strategy for other products or special occasions. Continuing to use high-value rewards in promotional campaigns not only increases views but also encourages community interaction, such as comments and shares. To further improve, Lipovitan could enhance two-way communication by discussing and addressing consumer questions directly on YouTube, thereby strengthening the relationship between the brand and its customers and fostering a loyal product community.

CHAPTER 4: RECOMMENDATION AND CONCLUSION

4.1. Achieved result

The project successfully met its initial goals, demonstrating effective strategies in social media analytics for Lipovitan. Key achievements include the detailed sentiment analysis across multiple platforms—Facebook, YouTube, and TikTok—which revealed that the majority of interactions were neutral to positive, indicating a favorable consumer reception. The data-driven approach enabled us to quantify and visualize user engagement metrics effectively, showing peaks in activity that correlated with specific marketing initiatives. The project also innovated in the application of PhoBERT, a Vietnamese NLP model, which enhanced the accuracy of sentiment analysis for the regional language, providing deeper insights into local consumer behavior.

4.2. Evaluation

The evaluation of the project outcomes shows that the methodologies employed were highly effective in achieving detailed insights into consumer sentiment and engagement. The use of advanced data analytics tools allowed for a comprehensive analysis of complex datasets. The integration of sentiment analysis with NLP technology provided nuanced understandings of consumer responses, which is critical for refining marketing strategies. However, there were challenges in correlating sentiment analysis results directly with sales data, indicating a need for more integrated analytics approaches that can link social media dynamics with actual business outcomes.

4.3. Limitation

Despite its successes, the project faced several limitations. The reliance on available social media data meant that the analysis could have missed sentiments expressed outside these platforms. Additionally, the scope of sentiment analysis was confined to detected expressions, which may not capture implicit sentiments or the full spectrum of consumer emotions. The project also encountered limitations in data collection due to API restrictions, which could affect the completeness of the data.

4.4. Future work

Future projects should aim to incorporate a wider range of data sources, including offline feedback and direct consumer surveys, to develop a more holistic view of consumer sentiments. Integrating real-time data analytics would allow for more agile marketing responses. Additionally, exploring machine learning models that can predict consumer behavior changes based on sentiment trends could significantly benefit strategic planning. Further research is also recommended to refine the NLP

models used, ensuring they better capture and interpret the subtleties of consumer language, especially in a multilingual context.

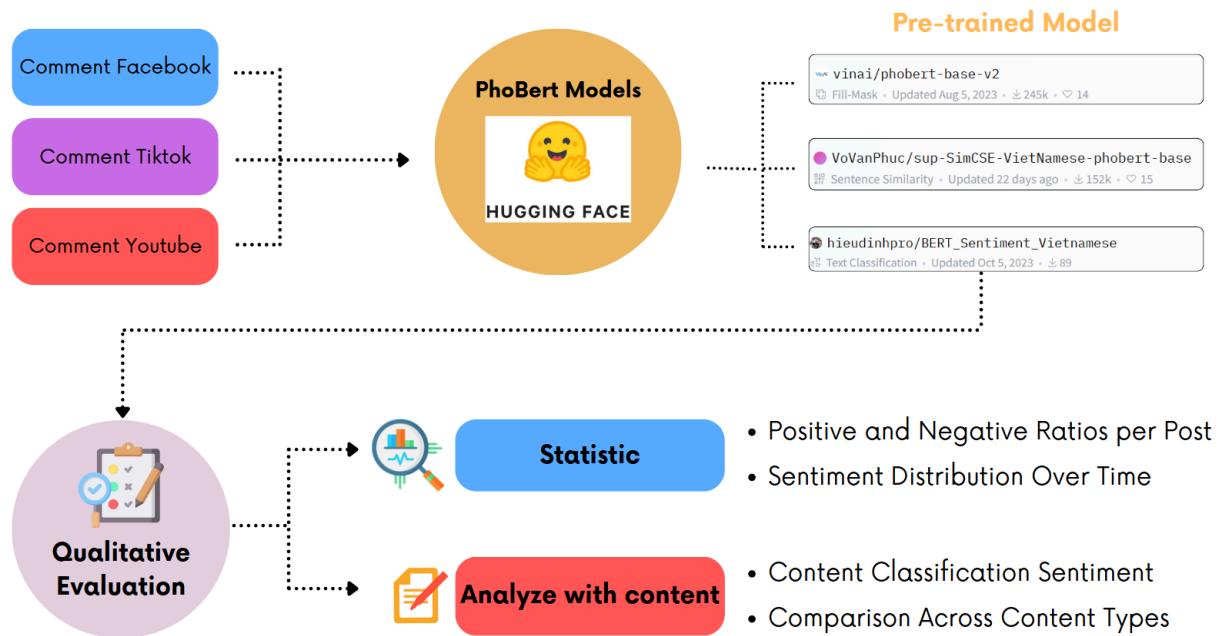


Figure 4.1 Next Analysis Process

In the current context, sentiment analysis of comments on social media platforms like Facebook, TikTok, and YouTube using the PhoBERT model has shown potential in better understanding consumer behavior. However, to enhance analytical capabilities and gain deeper insights, it is essential to expand the existing dataset. Collecting data from more platforms and over longer periods will help track and analyze sentiment trends over time, as well as evaluate the effectiveness of marketing campaigns based on user reactions.

To achieve this, we propose increasing the collection of comments not only from major social networks but also from forums and blogs, integrating data from specific events and campaigns for more detailed analysis. Additionally, developing a system to monitor sentiment trends will help detect changes in user behavior and content effectiveness. Moreover, applying advanced machine learning and deep learning techniques to refine and develop analysis models will improve the accuracy of sentiment classification and provide insights into the relationship between content and viewer reactions. These efforts will ensure that Lipovitan can optimize its content and marketing strategies, thereby achieving sustainable growth and efficiency in business operations.