



Facultad de Ciencias Económicas y Empresariales

CASO MAHOU:

**ESTRATEGIAS DE MARKETING Y LA
CONSTRUCCIÓN DE UNA LOVE BRAND A TRAVÉS
DEL MARKETING EMOCIONAL Y EL
ENGAGEMENT.**

Autor: Jimena Gómez Carreño

Tutor: Ana Isabel Jiménez Zarco

MADRID | Junio 2024

RESUMEN

El presente Trabajo tiene como objetivo principal el analizar si Mahou ha logrado posicionarse como una *Love Brand* en el competitivo mercado cervecero. Una *Love Brand* se caracteriza por generar una conexión emocional profunda y duradera con sus consumidores, similar a una relación de amor sentimental, lo que resulta en una lealtad y confianza sinigual. El análisis se basa en identificar los factores que contribuyen y las estrategias de marketing implementadas por la empresa para fortalecer su vínculo con el público. Para ello, se ha llevado a cabo un análisis empírico empleando técnicas de *Data Analytics*, concretamente el *Text Mining* y *Sentiment Analysis*, aplicadas a comentarios y opiniones de consumidores en redes sociales como Twitter, utilizando RStudio para procesar y analizar los datos.

PALABRAS CLAVE

Love brand; Mahou; amor; marketing; marca; emociones; sentimientos; text mining; Twitter.

ABSTRACT

The main objective is to analyze whether Mahou has managed to position itself as a Love Brand in the competitive beer market. A Love Brand is characterized by generating a deep and lasting emotional connection with its consumers, similar to a sentimental love relationship, which results in unparalleled loyalty and trust. The analysis is based on identifying the contributing factors and marketing strategies implemented by the company to strengthen its bond with the public. To the end, an empirical analysis has been carried out using Data Analytics techniques, specifically Text Mining and Sentiment Analysis, applied to consumer comments and opinions on social networks such as Twitter, using RStudio to process and analyze the data.

KEYWORDS

Love brand; Mahou; love; marketing; brand; emotions; sentiment; text mining; Twitter.

ÍNDICE DE CONTENIDOS

CAPÍTULO I. PLANTEAMIENTO INICIAL	5
1. INTRODUCCIÓN	5
1.1. Objetivo y relevancia del estudio	5
1.2. Estructura del Trabajo	6
1.3. Metodología de investigación	6
2. MARCO TEÓRICO: SOBRE LAS “ <i>LOVEBRANDS</i> ”	8
2.1. Conceptos y calificación	8
2.1.1. Origen y definición	8
2.1.2. Características de la lovebrand	12
2.1.3. Factores clave para el éxito de una Love Brand	12
2.2. Técnicas del marketing	14
2.2.1. Marketing Emocional.....	14
2.2.2. Estrategias de Engagement	15
3. MARCA MAHOU	18
3.1. Historia de Mahou	18
3.2. Sus estrategias de marketing	19
3.2.1. Campañas icónicas y marketing emocional.....	19
3.2.2. Colaboraciones y ediciones especiales	20
CAPÍTULO II. ANÁLISIS EMPÍRICO	21
1. OBJETIVOS	21
2. HERRAMIENTAS UTILIZADAS	22
2.1. Octoparse 8	22
2.2. RStudio y sus librerías	22
3. TEXT MINING	24

3.1. Introducción al <i>text mining</i>	24
3.2. Metodología	26
3.2.1. Extracción de datos con Octoparse	26
3.2.2. Limpieza de datos	28
3.2.3. Tokenización y creación de la Matriz DTM	28
3.2.4. Análisis de la frecuencia de palabras	29
4. SENTIMENT ANALYSIS	32
4.1. Introducción al Análisis de Sentimiento	32
4.1.1. Niveles de análisis de sentimiento	32
4.1.2. Límites en el análisis de sentimientos.....	33
4.2. Análisis en RStudio	34
CAPÍTULO III. CONCLUSIONES GENERALES DEL ANÁLISIS	38
BIBLIOGRAFÍA	41
ANEXO	45

CAPÍTULO I. PLANTEAMIENTO INICIAL

1. INTRODUCCIÓN

1.1. Objetivo y relevancia del estudio

De acuerdo a MONTALVO-ARROYARVE, el concepto de “*Love Brand*” se refiere a una marca que ha logrado generar una conexión emocional profunda y duradera con sus consumidores, similar a una relación de amor. Esta conexión se traduce en una lealtad y confianza excepcionales, donde los consumidores no solo eligen la marca por sus productos, sino también por la experiencia emocional que ofrecen¹.

Considerando lo anterior, este trabajo tiene como principal objetivo el análisis detallado del posicionamiento de Mahou en el mercado cervecero, identificando los factores que han contribuido a su éxito. Además, se busca desentrañar los patrones característicos de una “*Love Brand*” y las estrategias específicas que Mahou ha implementado para lograr esta distinción.

Otro aspecto crucial es medir el nivel de afecto que los consumidores tienen hacia la marca Mahou, relacionándolo con actitudes positivas, neutras o negativas. Se evaluará la efectividad de las estrategias de marketing de Mahou en la construcción de la lealtad del cliente y el *engagement* generado con su audiencia. Mediante este análisis, se buscará entender cómo las acciones de Mahou influyen en la percepción y comportamiento de quienes consumen sus productos.

Además, pretende determinarse si Mahou puede ser clasificada correctamente como una *Love Brand*, y cómo esta condición impacta en la empresa y el negocio a lo largo del tiempo. Finalmente, se explorará la importancia de las estrategias de marketing en la consolidación de Mahou como una *Love Brand* y su influencia en la relación empresa-cliente en un entorno en constante evolución. Mediante este estudio se pretende ofrecer una visión profunda de cómo Mahou ha logrado este estatus y sus implicaciones en el mercado cervecero.

De esta manera, se pretende responder las siguientes preguntas: ¿Podemos calificar a Mahou como una *Love Brand*? ¿Qué estrategias han llevado a la marca a posicionarse de esa manera en el mercado? ¿Qué atributos son los más relevantes a la hora de convertirse en una *Love Brand*?

¹ Arroyave, L. F. M. (2023). Las Lovebrands: Aproximación al concepto, arquitectura y mediciones. Cuadernos del Centro de Estudios en Diseño y Comunicación. Ensayos. Pp. 75-88.

1.2. Estructura del Trabajo

La estructura del presente Trabajo permite un análisis detallado del posicionamiento de Mahou en el mercado, su clasificación como una *lovebrand* y las estrategias que han contribuido a ello.

En un primer momento, se establece el contexto y los objetivos del estudio, que dan luz a la relevancia de analizar el posicionamiento de Mahou y su estatus en el mercado. Seguidamente, se lleva a cabo una revisión de la literatura sobre los conceptos desarrollados en el Trabajo, explotando definiciones, características y dimensiones de las *lovebrands*, incluyendo teorías y modelos de referencia que ayudan a entender el fondo del asunto. A continuación, se presenta un resumen del perfil de la marca, por medio de un recorrido histórico sobre la trayectoria de Mahou y un detalle de las principales estrategias llevadas a cabo por la marca a lo largo de los años.

Posteriormente, se lleva a cabo un análisis empírico aplicando técnicas de análisis de datos sobre las opiniones de los consumidores de la marca en las redes sociales, que buscan ilustrar que nivel de amor por la marca tienen los consumidores de Mahou, por medio de técnicas de *Text Mining* y *Sentiment Analysis*.

Finalmente, se concluye el estudio realizado, respondiendo a las preguntas de investigación planteadas y ofreciendo perspectivas a cerca de las estrategias de marketing de la marca.

1.3. Metodología de investigación

Primeramente, se busca realizar un análisis teórico utilizando técnicas cualitativas mediante la revisión exhaustiva de literatura académica y científica. Este análisis tiene como objetivo delimitar el marco conceptual sobre qué son las *lovebrands*, cómo calificarlas, qué atributos y dimensiones poseen y cuáles son los factores clave que contribuyen a su desarrollo. Esto permitirá establecer una base sólida que permita entender los componentes esenciales que hacen que una marca sea amada por sus consumidores.

Posteriormente, se realizará un estudio empírico del caso concreto de Mahou empleando técnicas de *Data Analytics*, que se centrará en analizar la percepción y el sentimiento de los usuarios de Twitter respecto a la marca Mahou. Para ello, se utilizan técnicas como el *Text Mining* empleando Octoparse 8 como vía para hacer *web scraping*, y *Sentiment Analysis* de

los comentarios encontrados en ellas por medio de RStudio, que permitirá buscar el máximo nivel de sentimiento positivo contenido en las opiniones de la comunidad.

La combinación de ambas vías permitirá obtener una visión holística del posicionamiento de la marca como una *lovebrand*, proporcionando desde el contexto y criterios necesarios para evaluarla, hasta datos concretos y cuantificables sobre la percepción que se tiene de Mahou por parte de la comunidad.

2. MARCO TEÓRICO: SOBRE LAS “LOVEBRANDS”

2.1. Conceptos y calificación

2.1.1. Origen y definición

El término “amor de marca” deriva del concepto de amor interpersonal en psicología. STENBERG² afirma que existen tres dimensiones de este: la intimidad, la pasión y el compromiso. Mientras que la intimidad está relacionada con el apego emocional a la relación, la pasión está relacionada con la psicología. El compromiso, por otro lado, tiene que ver con el lado cognitivo del amor y se refiere a la decisión, al reconocimiento de una relación a corto plazo y el deseo de preservar el vínculo a largo plazo³. Así, mientras el gusto está unido a un enamoramiento a corto plazo, el compromiso contribuye a uno a largo plazo construido mediante una repetición de compras a lo largo del tiempo.

Al ser un concepto relativamente nuevo (propuesto por Kevin Roberts en 2005), las definiciones del concepto de “Lovebrand” son diversas. CARROL y AHUVIA⁴, por ejemplo, lo definen como “*el grado de apego emocional apasionado que un consumidor satisfecho siente por un nombre comercial concreto*”. DE CÓRDOBA⁵ identifica este fenómeno como marcas con las que los consumidores establecen una relación que trasciende del mero acto de compra. Constituyen una identidad personal, proporcionando un sentimiento de pertenencia a un grupo o una sensación de diferenciación o exclusividad. También hay quienes consideran que una Love Brand es aquella que consigue generar una afinidad tan fuerte con el cliente que influye en su decisión de compra, impulsándolo a actuar de manera más emocional que racional. En esta línea, y cómo dice ROBERTS, “*se trata de marcas, productos o servicios que inspiran una lealtad más allá de la razón, para construir y fortalecer lazos emocionales entre marcas y consumidores*”. Según él, “*hablar de Lovebrands es hablar de amor puro por la marca, de un amor que sobrepasa la razón*”⁶.

Según varios autores, como ALBERT y MERUNKA⁷, la identificación con la marca y la confianza en ella se conciben como antecedentes del amor hacia una marca, pues cuando un

² Sternberg, R. (1986). A triangular theory of love. *Psychological Review*. Pp. 100 y ss

³ Montaguti, A. C., & Lelis, C. (2017). How to Make a Lovebrand: A Tentative Recipe.

⁴ Carroll, B. A. & Ahuvia, A. C. (2006). Some antecedents and outcomes of brand love.

⁵ De Córdoba, F. (2022). Los secretos de las marcas. Kailas Editorial.

⁶ Roberts, K. (2005). Lovemarks: El futuro más allá de las marcas. Ediciones Urano S.A., p.37 y ss

⁷ Albert, N., Merunka, D., & Valette-Florence, P. (2013). Brand passion: Antecedents and consequences. *Journal of Business Research*.

consumidor confía en una marca, es más probable que establezca cona relación más estrecha con ella. Así, cabe definir también estos términos originales:

a) Identificación de marca

El concepto se refiere al hecho de que los consumidores no compran productos sólo por su mera utilidad, sino que también lo hacen por lo que esos productos representan, por su beneficio simbólico⁸. De ahí que la identificación de marca esté asociada a la necesidad de identidad social y autodefinición de marca, definida por STERNBERG como “*el grado en el que el consumidor ve su propia autoimagen superpuesta a la imagen de marca*”.

En esta línea, CARROL y AHUVIA también afirman que es más probable que un consumidor llegue a amar una marca si se siente identificada con ella. Las marcas participan en la construcción de la autoimagen y la identidad de los consumidores, que las utilizan por su capacidad de reflejarse en a sí mismos en ellas, que las ayuda a construir sus propias identidades.

b) Confianza en la marca

La confianza en la marca se refiere a la percepción que tiene el consumidor de una marca sobre la misma como honesta y fiable, como que cumplirá sus expectativas y creencias⁹. Conlleva una actitud positiva de un consumidor hacia una marca, lo que le dispone a continuar la relación con ella. Esta actitud del consumidor influye positivamente en su comportamiento, pues conduce a un boca a boca positivo, a que las personas reincidan en la compra de la marca y a la resistencia a los competidores y, en última instancia, a la lealtad¹⁰.

En definitiva, se observa una conexión emocional entre la marca y el cliente que va más allá de la lógica, impulsando comportamientos de lealtad y de compra que son guiados por sentimientos arraigados en lugar de decisiones puramente cognitivas¹¹.

Como expresa DELGADO, la confianza en la marca es descrita como la sensación de seguridad que los consumidores experimentan en sus interacciones con las marcas. Es una confianza que se fundamenta en la percepción de los consumidores de que la marca no es sólo fiable, sino también comprometida con los intereses y el bienestar de sus clientes. Este autor lo

⁸ Albert, N., Merunka, D., & Valette-Florence, P. (2013). Brand passion: Antecedents and consequences. *Journal of Business Research*.

⁹ *Ibid*

¹⁰ Montaguti (2017)

¹¹ Arroyave, L. F. M. (2023). Las Lovebrands: Aproximación al concepto, arquitectura y mediciones. Cuadernos del Centro de Estudios en Diseño y Comunicación. Ensayos. Pp. 75-88.

relaciona con el concepto de *lovebrand* pues asume que el amor por una marca implica estar dispuesto a asumir riesgos al confiar en la promesa de valor que la marca representa. Luego la confianza en la marca implica una serie de expectativas generales que existen ante una posibilidad de error, pues los consumidores confían en una marca a pesar de ser conscientes de que puede haber fallos. Por tanto, hay una expectativa de que, a pesar de la posibilidad de resultados negativos, ocurrirán los positivos¹².

De esta idea principal expresada por el autor, nacen dos dimensiones correlacionadas:

i. La fiabilidad en la marca

Esta dimensión tiene una naturaleza técnica, ya que implica la percepción de que la marca puede cumplir o satisfacer las necesidades de los consumidores. Esto está relacionado con la creencia por parte del individuo de que la marca cumple con su promesa de valor. DELGADO considera que es una dimensión fundamental para que una persona deposite su confianza en la marca, puesto que, si los consumidores ven a la marca como una promesa de rendimiento futuro, y se fía en las vías para su cumplimiento, dicha fiabilidad guía al consumidor a confiar en la satisfacción futura del rendimiento prometido. Por tanto, la fiabilidad se reconoce como un punto de partida crucial para describir la confianza en la marca¹³.

ii. La intencionalidad

Esta segunda dimensión retrata la seguridad emocional de los individuos hacia la marca y se centra en la creencia de que la marca no se aprovechará de la vulnerabilidad de los consumidores. Es la idea de que la marca tiene la intención de actuar de manera ética y responsable, y no se aprovechará de la vulnerabilidad de los consumidores. En otras palabras, los consumidores consideran y creen que la marca se comportará de manera justa y honesta, incluso en situaciones desafiantes.

De esta manera, ambas dimensiones son fundamentales y complementarias para construir una relación de confianza marca-consumidor sólida y duradera, asegurando por un lado que la marca puede cumplir con sus promesas y satisfacer las necesidades de las personas,

¹² Delgado-Ballester, E. (2011). Development and validation of a brand trust scale.

¹³ *Ibid*

y por el otro, aportando la seguridad emocional necesaria para que los consumidores se sientan protegidos y valorados¹⁴.

c) Lealtad hacia la marca

La lealtad hacia una marca se define como el compromiso consciente de volver a consumir sus productos¹⁵. Según SHAMUGAN¹⁶, este compromiso hacia la marca es resistente a influencias externas. Es decir, los consumidores que han desarrollado una lealtad sólida hacia una marca no se ven influenciados por promociones de competidores o cambios en las estrategias de marketing, sino que es una relación profunda y estable, que va más allá de la simple satisfacción del consumidor en una compra individual.

Para LIMPASIRISUWAN y DONKWA¹⁷, este concepto es un factor que destaca el rendimiento de una empresa. La lealtad a la marca puede convertirse en una ventaja competitiva en el mercado, pues puede evaluarse tanto desde la perspectiva del comportamiento como de la actitud.

La lealtad comportamental se refiere a las acciones repetitivas de compra de un producto específico. Como señala KANG¹⁸, este concepto se ha convertido en un tema popular tanto en el marketing empresarial como en el marketing al consumidor, representando una repetición del comportamiento de compra a menudo por hábito o conveniencia. Es un tipo de lealtad que refleja constancia en la elección de la marca, influenciada más por la familiaridad y accesibilidad que por la evaluación consciente y deliberada de la marca.

Por otro lado, la lealtad actitudinal se define como el compromiso psicológico del consumidor hacia una marca. Como señala HUANG¹⁹, esta forma de lealtad conlleva una creencia firme de que la marca es única y diferente a las demás, reflejando no sólo una preferencia consciente, sino también un compromiso emocional profundo con la marca, lo que

¹⁴ Delgado-Ballester (2011)

¹⁵ Madeline, S., & Sihombing, S. O. (2019). The impacts of brand experiences on brand love, brand trust, and brand loyalty: an empirical study. *Jurnal Bisnis dan Manajemen*. Pp. 91-107

¹⁶ Kabiraj, S., & Shanmugan, J. (2010). Development of a conceptual framework for brand loyalty: A Euro Mediterranean perspective. *Journal of Brand Management*.

¹⁷ Limpasirisuwan, N., & Donkwa, K. (2017). A structural equation model for enhancing online brand research: an international journal.

¹⁸ Kang, S., (2015). Dual Dimensions of Attitudinal Loyalty and Behavioral Loyalty among Hotel Customers. *Journal of Tourism Research & Hospitality*.

¹⁹ Huang, C. C. (2017). The impacts of brand experiences on brand loyalty: mediators of brand love and trust. *Management Decision*.

puede llevar a los consumidores a defenderla y recomendarla activamente, incluso en ausencia de incentivos externos.

2.1.2. Características de la lovebrand

Las *lovebrands* se distinguen por ser empresas especialmente atractivas y apreciadas por sus consumidores. Son marcas que logran establecer conexiones emocionales de fuerte alcance y trascender la mera transacción comercial. Son marcas que, como norma general²⁰:

Primero. Son capaces de forjar relaciones emocionales profundas con sus consumidores: son conexiones que van más allá del producto o servicio que se ofrece, creando vínculos afectivos que influyen en la percepción y el comportamiento de los consumidores.

Segundo. Los clientes muestran una lealtad inquebrantable. Lo común es que se prefieran estas marcas sobre sus competidores, incluso cuando existen alternativas más convenientes. Es una lealtad basada en la confianza, en la satisfacción emocional que estas marcas proporcionan.

Tercero. Los consumidores perciben un gran valor en estas marcas, no sólo en términos de producto, sino también a nivel emocional y de experiencia. Es un valor agregado que contribuye a su preferencia y lealtad hacia la marca.

Cuarto. Tienen una identidad de marca clara y fuerte. Las *lovebrands* son fácilmente reconocibles y comunican consistentemente su esencia y valores, lo que las diferencia en el mercado.

Quinto. Generalmente poseen elementos icónicos fácilmente identificables. Son elementos que forman parte de la cultura y los recuerdos de los consumidores, y que consolidan su posición en el corazón y mente de las personas.

2.1.3. Factores clave para el éxito de una Love Brand

Una de las aspiraciones principales de las marcas es conseguir que los clientes desarrollen este sentimiento de amor hacia ellas. Sin embargo, se trata de un proceso sistemático que demanda tiempo y que implica comprender profundamente cómo los clientes viven la experiencia de la marca, que eventualmente culmina en la adopción de este sentimiento²¹.

²⁰ Montalvo-Arroyave, L. F. (2023). Las Lovebrands: Aproximación al concepto, arquitectura y mediciones.

²¹ Belk, R. (1988). Possessions as the Extended Self. *Journal of Consumer Research*.

Hay autores que, comparándolo con el amor sentimental entre personas, entienden que es lógico pensar que no es lo normal que los clientes se enamoren de la marca a “primera vista”, es más bien un proceso metódico que requiere tiempo, fortaleciéndose a lo largo de diversas experiencias y encuentros que permiten transformar sentimientos iniciales en una preferencia emocional más profunda²².

Según ROBERTS, para transformar una marca en una *lovebrand*, se debe proporcionar al consumidor tres elementos esenciales de experiencia: el misterio, la sensualidad y la intimidad. Son tres componentes que se consideran fundamentales para generar una marca emocionalmente atractiva y que contribuyen significativamente a crear una experiencia de *lovebrand*, fomentando a su vez que los consumidores desarrollen un fuerte apego y se conviertan en seguidores entusiastas de la marca ²³. El **Misterio**, por un lado, atrae a los consumidores al ofrecer una historia con la que pueden conectarse. La **Sensualidad**, sin embargo, se manifiesta en experiencias sensoriales agradables a través del *packaging*, de los colores, aromas, etc., lo que es crucial para la marca. Mientras que algunas marcas no utilizan todos los sentidos para cautivar a los consumidores, las verdaderas *lovebrands* los incorporan todos en su estrategia de branding. Finalmente, la **Intimidad**, que implica que los consumidores esperen ser influenciados a nivel personal por las marcas.

Por otro lado, así como dice MADELINE y SIHOMBING, el amor hacia una marca puede crecer a partir de experiencias directas con ella. Las experiencias positivas y memorables generan en el consumidor una conexión emocional profunda que refuerza el vínculo con la marca. Otra forma de hacerlo crecer es a través de la comunicación, ya sea controlada, mediante iniciativas de patrocinio por parte de la empresa, o publicidad y promociones, que permiten transmitir los valores y beneficios de la marca de manera directa; o incontrolada, como el boca a boca, que juega un rol imprescindible en la estrategia de marca. Las recomendaciones y opiniones de terceros pueden influir de manera significativa en la percepción y el deseo hacia la marca. Con esto, se entiende que un consumidor puede desear o anhelar una marca sin haberla experimentado directamente o consumido sus productos. Es un fenómeno que puede ocurrir dada la fuerte presencia o atractivo de la marca en el mercado, despertando interés y deseo entre los consumidores²⁴.

²² Sternberg (1986)

²³ Uribe-Torres, G. A., Romero-Jiménez, W., Rojas-Berrio, S., & Robayo-Pinzón, O. (2022). Lovemarks e investigaciones académicas en mercadotecnia: avance de un proceso de revisión sistemática de literatura. Clío América.

²⁴ Madeline, Sihombing (2019)

La personalidad individual también tiene un impacto de alto valor en el desarrollo de una *lovebrand*. Los consumidores tienden a sentirse atraídos por marcas que reflejan su autoconcepto y valores personales, así sea, cuando una marca proyecta una imagen que resuena con la identidad del consumidor incrementa la probabilidad de que este desarrolle un amor y lealtad superior hacia la marca²⁵.

2.2. Técnicas del marketing

En el ámbito del marketing, existen varias técnicas que las marcas pueden utilizar para fortalecer su relación con los consumidores y convertirse en *lovebrands*, buscando crear conexiones profundas y duraderas con los clientes. Entre ellas, destacamos el marketing emocional y las estrategias de *engagement*.

2.2.1. Marketing Emocional

El marketing emocional se centra en crear y cultivar una conexión emocional entre la marca y el consumidor. Es una técnica basada en la idea de que las decisiones de compra no son puramente racionales, sino que están fuertemente influenciadas por las emociones. Se consigue a través de diferentes vías, desde campañas publicitarias conmovedoras, experiencias memorables, hasta historias de marca auténticas. En definitiva, se permite la oportunidad de experimentar emociones en el mercado consumidor²⁶.

GOBÉ, reconocido por haber introducido este concepto, sostiene que los consumidores han transformado su manera de comprar: ahora quieren conocer más sobre lo que compran. La clave está en identificar las emociones específicas que se desea evocar en los consumidores para luego diseñar y ofrecer productos que satisfagan esas emociones. Al entender las necesidades emocionales de los clientes, las empresas pueden crear ofertas que no solo cumplan sus expectativas funcionales, sino que generen también una conexión emocional²⁷.

GÓMEZ desarrolla la idea del Marketing de Emociones en su artículo “La forma para lograr fidelidad de los clientes”, defendiendo que este tipo de estrategia se ha convertido en una herramienta clave para muchas empresas exitosas que han comprendido, desde sus inicios, la importancia de entender a sus clientes. Las empresas reconocen que las emociones son una parte integral de su manera de operar y que las relaciones a largo plazo son fundamentales para

²⁵ *Ibid.*

²⁶ Barragán Codina, J. N., Guerra Rodríguez, P., & Villalpando Cadena, P. (2017). La economía de la experiencia y el marketing emocional: estrategias contemporáneas de comercialización.

²⁷ Gobé, M. (2005) Branding Emocional. Studio Divine Egg.

la consecución de sus logros. Esta rama clarifica los principios necesarios para crear relaciones duraderas²⁸.

Las empresas deben invertir continuamente en conocer y entender las necesidades y deseos de sus clientes, adaptando sus estrategias para mantener y fortalecer las relaciones con el tiempo. De esta manera, el marketing emocional necesita apoyo de los directivos, pues los líderes de la organización han de comprometerse a, por un lado, respaldar las iniciativas de marketing, y por el otro a promover una cultura que valore y priorice la conexión emocional con los clientes. Así, para que sea verdaderamente efectivo, debe haber un compromiso integral y continuo a lo largo de toda la organización, integrando esta idea en la cultura de cada empresa²⁹.

2.2.2. Estrategias de Engagement

Una vez establecido el vínculo emocional con los consumidores, se consideran *lovebrands* las marcas que han logrado establecer una conexión emocional profunda y duradera con sus consumidores. Para mantenerla y fortalecerla, existen estrategias de *engagement* efectivas que pueden ser clave para lograrlo. Como establece GAMBETTI, actualmente se considera el *engagement* como la fuerza motriz fundamental del comportamiento y la toma de decisiones del consumidor postmoderno³⁰.

A) Contenido generado por usuarios (UGC³¹)

El “*User Generated Content*” ha ganado protagonismo en la era digital en los últimos años, especialmente en el contexto de las redes sociales. Este tipo de estrategia se ha convertido en un pilar fundamental para la comunicación entre las empresas y sus audiencias, abarcando cualquier tipo de contenido (desde reseñas, vídeos o comentarios) creado y compartido por los usuarios de las marcas en plataformas digitales, “sustituyendo” al producido por las empresas o medios de comunicación tradicionales³².

Según NAAB y SEHL, el UGC ha de ser publicado y accesible al público o a un grupo para permitir una discusión generalizada, y ha de ser desarrollado fuera del ámbito profesional

²⁸ Gómez, H. D. C. (2004). Marketing de emociones. La forma para lograr fidelidad de los clientes. Semestre económico.

²⁹ Gómez (2004)

³⁰ Gambetti, R. C., & Graffigna, G. (2010). The concept of engagement: A systematic analysis of the ongoing marketing debate. *International Journal of Market Research*. Pp. 801-826.

³¹ Por sus siglas en inglés: “*User Generated Content*”

³² Christian David, F. G. (2023). El Impacto del Contenido Generado por Usuarios (UGC) en las Redes Sociales.

del periodismo. De esta manera, se entiende que el contenido debe estar disponible en plataformas abiertas que permita una audiencia amplia que participen en debates e interactúen entre sí, pero además ha de ser creado por individuos comunes, pues ha de reflejar las perspectivas y experiencias personales de los usuarios. Este tipo de contenido refleja la autenticidad y transparencia que los consumidores actuales buscan, haciendo que la marca se perciba como más accesible y cercana. El UGC tiene la capacidad de transformar a los consumidores en verdaderos embajadores de las marcas, promoviendo comunidades activas y comprometidas, lo que es fundamental para el éxito a largo plazo de cualquier *lovebrand*.³³

B) Interacción en redes sociales

Autores como KIETZMANN y HERMKENS destacan la creciente importancia de la presencia de las empresas en las redes sociales. Especialmente en el ámbito del *engagement*, las redes permiten generar conversaciones y obtener respuestas directas por parte de los consumidores³⁴.

Como dice FERNÁNDEZ, la era digital ofrece a las empresas nuevas oportunidades y las impulsa a reinventarse constantemente para conectar con el público y gestionar el incesante flujo de información *online*. En este escenario, las plataformas y redes sociales se han convertido en los mejores escaparates para que los medios interactúen con los usuarios, participando activamente este proceso de comunicación que fomenta el establecimiento de vínculos de compromiso y lealtad hacia quien comparte el contenido³⁵.

La difusión de contenido es, hoy en día, un factor estratégico clave para ofrecer valor añadido al usuario, siendo fundamental estar presente en las redes sociales para que, en el caso de que los consumidores no busquen el contenido, los contenidos se acerquen a ellos. Hoy en día, se entiende que estas plataformas ofrecen a sus usuarios herramientas más cercanas, accesibles y, en muchos casos, más confiables que los medios de comunicación tradicionales³⁶. El *engagement* se manifiesta a través de la cantidad de recomendaciones, referencias,

³³ Naab, T. K., & Sehl, A. (2017). Studies of user-generated content: A systematic review. *Journalism*.

³⁴ Moreno Fernández, Á., & Fuentes Lara, C. M. (2019). 'Engagement' y redes sociales. Análisis bibliométrico desde el ámbito científico de las relaciones públicas.

³⁵ Fernández, S. G. (2020). Estrategias comunicativas para informar y crear engagement en Instagram. El caso del periódico El País. *Razón y Palabra*,

³⁶ *Ibid.*

comentarios o sugerencias que realizan los consumidores, lo que refleja también el nivel de complacencia y agrado de los consumidores sobre la marca³⁷.

³⁷ Van Doorn, J. et al.(2010). Customer Engagement Behavior: Theoretical Foundations and Research Directions. *Journal of Service Research*, vol. 13 (3), Pp. 253-266.

3. MARCA MAHOU

3.1. Historia de Mahou

Mahou es una marca emblemática en el sector cervecero español cuya historia se remonta a más de un siglo. Fundada en 1890 en Madrid, Mahou tuvo sus inicios como una pequeña cervecería familiar bajo el nombre de “Hijos de Casimiro Mahou”, y desde entonces, la empresa fue ganándose rápidamente un lugar en el mercado de la época³⁸.

A lo largo del siglo XX, Mahou experimentó un crecimiento significativo, expandiéndose y modernizando sus instalaciones para satisfacer su cada vez más creciente demanda. En 1969 la compañía lanzó su producto estrella, Mahou Cinco Estrellas, una cerveza que se convertiría en el buque insignia de la marca y en un referente de calidad dentro del mercado español, consolidando a la empresa como una de las marcas de cerveza más reconocidas y apreciadas en España. Como cuenta el director general de marketing, con su campaña, Mahou Cinco Estrellas quiere conectar con sus consumidores haciéndoles protagonistas de ella. Ésta ya forma parte de una larga colección de spots y de acciones comunicativas que incitan a disfrutar de las pequeñas cosas como uno de los placeres de la vida, porque *“mantener el espíritu de sus marcas y no dejar nunca de lado la innovación son dos de los leitmotivos del Grupo”*³⁹. La historia de Mahou resulta, al final, la historia de un éxito basado en la tradición y la innovación⁴⁰.

Durante las últimas décadas, la marca ha logrado una notable presencia internacional, exportando sus productos a numerosos países y llevando a cabo una expansión global que ha permitido a la marca mantenerse relevante en el mercado internacional. Como señalan RUIZ y ROLDÁN, España es actualmente el octavo productor de cerveza en el mundo, y Mahou se ha convertido en la novena marca europea⁴¹.

Destaca también su fuerte presencia en lo relativo a la cultura y sociedad españolas. La marca ha patrocinado numerosos eventos deportivos, culturales e incluso musicales, fortaleciendo su conexión emocional con los consumidores y consolidando su imagen de marca como cercana y comprometida.

³⁸ Vijande, L. S., & Múgica, B. G. B. (2000). Estrategias de marketing para un crecimiento rentable: casos prácticos, 103.

³⁹ Rey, J. P. (2019). Mahou San Miguel: Maestros de la tradición y artesanos de la innovación.

⁴⁰ Raposo, N. P. (1999). José Luis García Ruiz y Constanza Laguna Roldán, Cervezas Mahou, 1890-1998. Un siglo de tradición e innovación. Revista de Historia Industrial. Pp. 235-238.

⁴¹ *Ibid.*

3.2. Sus estrategias de marketing

3.2.1. Campañas icónicas y marketing emocional

A lo largo de los años, se ha reconocido ampliamente a Mahou por su lanzamiento de varias campañas publicitarias que se han convertido en verdaderos hitos dentro del panorama publicitario español, todas ellas basadas fuertemente en los principios del marketing emocional.

Una de las campañas más emblemáticas y exitosas en la historia de la marca fue el caso de “Soy muy de Mahou” (2014)⁴², que fue lanzada con el objetivo de reforzar la identidad de la marca y su conexión emocional con los consumidores. Fue una campaña que, en su momento, hizo convertirse a la empresa como un referente dentro del marketing cervecero.

Esta campaña fue lanzada en un momento en el que las marcas de cerveza competían intensamente por captar la lealtad de los consumidores, y buscaba principalmente reforzar el sentimiento de pertenencia y orgullo entre los consumidores de Mahou, destacando las cualidades que los hacen únicos y que los unen a la marca. Busca posicionar a Mahou, no sólo como una cerveza, sino como un símbolo de identidad y estilo de vida para muchos españoles.

La estrategia fue principalmente basada en el uso del marketing emocional para conectar con los consumidores, mediante una serie de anuncios de televisión, radio y redes sociales, presentando historias emotivas que reflejan momentos cotidianos en la vida de las personas, utilizando testimonios reales de consumidores y figuras públicas que se identificaban con la marca. Enlazaba testimonios que enunciaban “soy muy de Mahou porque...”.

Esta campaña fue concebida como un referente en el marketing emocional español, que junto a otras muy conocidas (ej. Loterías), demostró el poder de las emociones en la construcción de marcas exitosas.

A esta, le siguieron muchas otras como “Un sabor muy grande” (2018)⁴³ o “Somos familia” (2020)⁴⁴, que consolidaron a la marca en este tipo de acciones.

⁴² ReasonWhy (2014)

⁴³ Marketing Directo (2018). Un Sabor Muy Grande, el homenaje de la nueva campaña de Mahou Cinco Estrellas al sabor que vive en las calles. Recuperado de: <https://www.marketingdirecto.com/marketing-general/publicidad/un-sabor-muy-grande-el-homenaje-de-la-nueva-campana-de-mahou-cinco-estrellas-al-sabor-que-se-vive-en-las-calles>

⁴⁴ PR Noticias (2021). Mahou San Miguel presenta: Reencuentros, la cuarta entrega de su campaña 'Somos Familia' para homenajear a los hosteleros. Recuperado de: <https://prnoticias.com/2021/07/16/mahou-san-miguel-presenta-reencuentros-la-cuarta-entrega-de-su-campana-somos-familia-para-homenajear-a-los-hosteleros/>

3.2.2. *Colaboraciones y ediciones especiales*

La marca Mahou ha implementado diversas colaboraciones y ediciones especiales que han resonado en la industria, y que les ha permitido reforzar su imagen de marca de manera generalizada, en diferentes sectores.

Uno de los ejemplos más destacados es la reciente colaboración de Mahou con el artista de música Duki⁴⁵. La marca lanzó una edición especial de sus botellas con diseños personalizados por el artista, y lo publicitó abiertamente por todo el territorio nacional con paneles y carteles publicitarios. Se dice que esto reforzó la imagen de la marca como actual y en sintonía de las tendencias culturales.

Como esta, hay numerosas otras colaboraciones, como la llevada a cabo con el chef David Muñoz, o la edición especial lanzada a razón del festival Mad Cool Festival. Destaca también la campaña de “Una Cerveza, Una Buena Causa”, cuyos beneficios fueron destinados a proyectos de sostenibilidad ambiental y apoyo a comunidades locales.

Todo ello forma parte de la estrategia de marketing de la marca que busca, en definitiva, fortalecer su conexión emocional con los consumidores, asegurando su lealtad y apoyo continuo.

⁴⁵ ARAL (2024). Mahou y Duki lanzan la nueva cerveza 'La Diabla'. Recuperado de: <https://www.revistaaral.com/texto-diario/mostrar/4862098/mahou-duki-lanzan-nueva-cerveza-diabla>

CAPÍTULO II. ANÁLISIS EMPÍRICO

1. OBJETIVOS

El objetivo primordial de esta parte del trabajo es evaluar y comprender la percepción de los consumidores de la marca a través del análisis de sus comentarios en redes sociales (concretamente, en Twitter).

En primer lugar, se busca identificar las palabras más comunes utilizadas por los consumidores en los comentarios que publican en la red social sobre Mahou. Esto se ha llevado a cabo utilizando la herramienta de Octoparse 8 para hacer *web scraping* de los tweets que, mediante una serie de búsquedas específicas, están relacionados con la marca. Posteriormente, se ha llevado a cabo un acto de procesamiento y limpieza de la base de datos para eliminar el ruido y asegurar que las palabras clave identificadas son relevantes y significativas para el estudio. Así, se procede a transformar estos comentarios en una estructura que permita analizar la frecuencia de las palabras más utilizadas.

A continuación, el siguiente objetivo se basa en evaluar cómo los consumidores perciben la marca Mahou a escala general. Este estudio se lleva a cabo mediante el análisis de los sentimientos captados en los comentarios. Para ello, se procede a aplicar técnicas de análisis de sentimientos para clasificar los comentarios como positivos, negativos o neutros, para después visualizar la distribución de éstos y llegar a comprender la percepción general de la marca.

Finalmente, más allá de los sentimientos, es de interés entender las emociones específicas que los consumidores asocian con la marca. Esto se lleva a cabo utilizando léxicos especializados como el NRC Emotion Lexicon, que permite identificar y analizar las emociones presentes en las palabras de los consumidores. De la misma manera, se crean visualizaciones que permiten mostrar la prevalencia de las diferentes emociones frente a otras (alegría, tristeza, enfado, etc.).

Con todo ello, y como objetivo principal, se busca comparar la percepción y las emociones que se asocian con Mahou frente a los criterios de una *lovebrand*, lo que conlleva aplicar los criterios ya definidos en la parte teórica para evaluar si Mahou cumple con ellos basándonos en los resultados del análisis, para posteriormente discutir las implicaciones de los mismos de cara a la estrategia de marketing de la marca.

2. HERRAMIENTAS UTILIZADAS

2.1. Octoparse 8

La metodología que se ha empleado para extraer los comentarios de las redes sociales es la plataforma de “Octoparse 8”. De cara a analizar si Mahou puede ser considerada o no una *lovebrand*, se ha usado esta herramienta de *web scraping* y análisis de datos que se define como “*el software de web scraping que permite extraer datos de sitios web de manera eficiente y estructurada*”. En definitiva, es una plataforma que facilita la recolección de datos de diversas páginas web sin necesidad de utilizar los métodos clásicos de programación⁴⁶.

Esta plataforma permite a los usuarios configurar tareas semiautomatizadas de *scraping* sin necesidad de confeccionar códigos de programación, sino más bien por una combinación de clics y secciones de la interfaz que permite definir las reglas de extracción de datos y su posterior descarga. Una de las características más destacadas de Octoparse 8 es la capacidad que tiene para manejar estructuras de datos complejas (extracciones de tablas, listas, contenido dinámico) cargado mediante JavaScript. Es una capacidad altamente valorada a la hora de trabajar en sitios web con contenido altamente estructurado o dinámico, como es el caso de las redes sociales, y específicamente Twitter.

Octoparse 8 ha permitido programar la tarea de *scraping* para que se ejecute automáticamente en un intervalo específico que se selecciona manualmente en un inicio. Esto permite recolectar datos de manera continua y mantener la base de datos actualizada sin necesidad de intervención constante, lo que ha permitido proporcionar un flujo de datos consistente para el análisis.

Los datos extraídos pueden ser exportados en diversos formatos, incluyendo CSV, Excel, JSON, etc. Con ello, se permite su posterior procesamiento en otras herramientas de análisis, como RStudio, que se detalla a continuación.

2.2. RStudio y sus librerías

RStudio es una plataforma integrada para el desarrollo en R⁴⁷, un lenguaje de programación y entorno de software libre diseñado para el análisis estadístico y visualización de datos. Se ha tomado la decisión de llevar a cabo la parte empírica de este Trabajo por medio de esta

⁴⁶ Pérez Molano, G. (2023). An Overview of Web Scraping: Technical Aspects and Exercises. Computer Science.

⁴⁷ Stander, J., & Dalla Valle, L. (2017). On enthusing students about big data and social media visualization and analysis using R, RStudio, and RMarkdown. Journal of Statistics Education.

herramienta por varias razones, empezando por su robustez y flexibilidad en el análisis de datos textuales y *sentiment analysis*. R es conocido por su amplia gama de paquetes y librerías que facilitan en análisis estadístico y la manipulación de datos, por lo que lo convierte en la herramienta ideal para proyectos como éste, de investigación académica.

Por otro lado, como dice STANDER⁴⁸, las capacidades de visualización de datos en RStudio, a través de paquetes como “*ggplot2*” y “*wordcloud*” permiten crear representaciones gráficas y efectivas de los resultados, muy útiles para el caso concreto de este análisis.

Es una herramienta que ofrece un entorno de desarrollo integrado que facilita, tanto la escritura, como la ejecución y depuración de códigos, mejorando la productividad y eficiencia. Para estas tareas, las librerías que se han utilizado son las siguientes:

<p>Readxl</p>	<p>Permite leer archivos de Excel en R. Es particularmente útil para este caso porque permite importar los datos extraídos desde Octoparse en formato Excel. Se decidió importar la base en este formato para facilitar una manipulación inicial de los datos antes de pasar a etapas más complejas del análisis, ya que los <i>outputs</i> eran bastante manejables y permitió descartar filas que, claramente a simple vista, no aplicaban para el estudio (Ejs.: texto con símbolos y sin contenido, textos en otros idiomas ilegibles, etc.)⁴⁹.</p>
<p>Tidyverse</p>	<p>Conjunto de paquetes (“<i>dyplr</i>”, “<i>ggplot2</i>”, “<i>tidyr</i>” y “<i>readr</i>”, entre otros) que permite utilizar herramientas de manipulación y limpieza de datos, empleado para procesar y limpiar los comentarios, así como crear visualizaciones preliminares⁵⁰.</p>
<p>Quanteda</p>	<p>Librería potente para el análisis cuantitativo de datos textuales, que proporciona herramientas para la tokenización, creación de matrices de términos y documentos (DTM) y análisis de frecuencias de palabras⁵¹.</p>

⁴⁸ Stander (2017)

⁴⁹ Wickham, H., Bryan, J., Kalicinski, M., Valery, K., Leittenne, C., Colbert, B., ... & Bryan, M. J. (2019). Package ‘readxl’

⁵⁰ Kabacoff, R. (2022). R in action: Data analysis and graphics with R and Tidyverse. Simon and Schuster.

⁵¹ Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. Journal of Open Source Software.

Styuzhet	Utilizada para el análisis de sentimientos. Contiene funciones que permiten extraer sentimientos y emociones de los textos utilizando diferentes léxicos, como el NRC Lexicon, que permite determinar la polaridad de los comentarios y extraer las emociones de los mismos ⁵² .
Wordcloud	Se emplea para crear nubes de palabras, como forma visual altamente efectiva de mostrar la frecuencia de palabras en un corpus de texto, y poder visualizar las palabras más comunes contenidas en los tweets ⁵³ .
RColorBrewer	Proporciona paletas de colores para usar en gráficos y visualizaciones, mejorando la estética de las gráficas creadas por <i>ggplot2</i> y <i>wordcloud</i> ⁵⁴ .
Stringi	Ofrece funciones avanzadas para la manipulación de cadenas de texto, incluyendo la decodificación de secuencias de escape y la normalización de texto, asegurando una correcta limpieza de los comentarios extraídos ⁵⁵ .

3. TEXT MINING

3.1. Introducción al *text mining*

La capacidad de almacenar datos ha ido creciendo exponencialmente en los últimos años, pero nuestra habilidad para procesarlos y utilizarlos no ha avanzado al mismo ritmo. En esta línea, la minería de datos surge como una tecnología esencial para explorar, analizar, comprender y aplicar el conocimiento obtenido de grandes volúmenes de datos⁵⁶.

Hoy en día, las redes sociales se han convertido en una parte esencial de la rutina diaria de las personas, donde expresamos y damos nuestra opinión desde cualquier lado y bajo cualquier circunstancia. Empleamos las redes para diversos fines, como la publicidad, la difusión de opiniones políticas y tendencias financieras, la obtención de opiniones de los consumidores sobre los productos y la divulgación de notificaciones, entre otros muchos usos. Es una manera de

⁵² Kim, H. (2022). Sentiment Analysis: Limits and Progress of the Syuzhet Package and Its Lexicons. *DHQ: Digital Humanities Quarterly*.

⁵³ Fellows, I., Fellows, M. I., Repp, L., & Repp, L. (2018). Package 'wordcloud'. R package version 2.

⁵⁴ Neuwirth, E., & Neuwirth, M. E. (2014). Package 'RColorBrewer'. ColorBrewer Palettes.

⁵⁵ Gagolewski, M. (2022). stringi: Fast and portable character string processing in R. *Journal of Statistical Software*, 103. Pp. 1-59

⁵⁶ Ballesteros, H. F. V., Iñiguez, E. G., & Velasco, S. R. M. (2018). Minería de datos. *RECIMUNDO: Revista Científica de la Investigación y el Conocimiento*. Pp. 339-349.

permitir a las personas compartir sus pensamientos, sentimientos y opiniones con otras personas de forma instantánea y sencilla.

Twitter (ahora “X”) es una de las redes sociales más populares. Los usuarios publican una media de 500 millones de tweets diarios⁵⁷, proporcionando una plataforma mediática que permite compartir opiniones por medio de diferentes formas de contenido (texto, imágenes, enlaces, etc.), lo que convierte esta plataforma en una vía adecuada para extraer opiniones a gran escala y casi a tiempo real, utilizando el *text mining*⁵⁸.

Pero, ¿qué es la minería de datos o *text mining*? Es un proceso que, mediante el uso de programas de ordenador, permite descubrir información, previamente desconocida, extrayendo automáticamente datos de diversas fuentes escritas. A diferencia de la búsqueda web, donde los usuarios buscan algo ya conocido y escrito por otros, el *text mining* tiene como principal objetivo descubrir información desconocida hasta el momento, algo que nadie sabe y que aún no ha sido escrito. Es una variación del *data mining*, que busca patrones interesantes en grandes bases de datos, pero que se centra en extraer patrones de texto en lenguaje natural en lugar de bases de datos estructurados. Como dice HEARST, las bases de datos están diseñadas para ser procesadas automáticamente por programas, mientras que el texto está escrito para ser leído por personas⁵⁹.

El *text mining* implica cuatro procesos principales⁶⁰:

- (1) Definición del contexto y bases de la extracción
- (2) Recolección de datos
- (3) Construcción de diccionarios
- (4) Análisis de datos

Según, LEE y HO, el *text mining* explora datos en archivos de texto para establecer patrones y reglas valiosas que indican tendencias y características significativas sobre temas específicos. El internet genera una gran cantidad de información, y este proceso representa una reserva de inteligencia empresarial que puede ser altamente valiosa para muchas empresas, ya que las

⁵⁷ Cranell, W.C. (2016). A pattern-matched twitter analysis of US cancer-patient sentiments

⁵⁸ Öztürk, N., & Ayvaz, S. (2018). Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis.

⁵⁹ Hearst, M. (2003). What is text mining. SIMS, UC Berkeley, 5.

⁶⁰ Gémar, G., & Jiménez-Quintero, J. A. (2015). Text mining social media for competitive analysis. *Tourism & Management Studies*. Pp. 84-90.

empresas pueden utilizar herramientas de *text mining* que las permita convertir datos online en información competitiva⁶¹.

Como dice BALLERSTEROS, el principal objetivo de la minería de datos es descubrir patrones y tendencias relevantes en grandes bases de datos, permitiendo estructurar la información de manera comprensible y útil para la toma de decisiones empresariales. Permite a las empresas transformar grandes volúmenes de datos en información relevante y conocimiento, lo que facilita la identificación de oportunidades para incrementar los ingresos, mejorar los márgenes operativos y desarrollar planes estratégicos basados en datos. Para él, los procesos de *text mining* implica la separación de distintas etapas⁶²:

Primera. Entendimiento del problema: es decir, definir claramente la pregunta a resolver y establecer objetivos específicos.

Segunda. Selección de datos: determinar qué bases de datos se utilizarán para el análisis.

Tercera. Procesamiento y limpieza: implica manejar datos faltantes y valores atípicos para asegurar la calidad del análisis.

Cuarta. Transformación de datos: generar nuevos indicadores que representen mejor el fenómeno analizado.

3.2. Metodología

3.2.1. Extracción de datos con Octoparse

A) Búsqueda avanzada en Twitter

Para empezar, y tras la instalación de la versión 8 de Octoparse en el ordenador, se procedió a definir la búsqueda en Twitter que permitiese obtener los comentarios más afines al contenido que se buscaba. Para ello, se utilizó la funcionalidad de “búsqueda avanzada” en Twitter que te permite imponer una serie de palabras clave, lapsos de tiempo o incluso menciones y hashtags a la búsqueda para así filtrar la clase de contenido que quieres obtener de una manera mucho más precisa y personalizada. En dicha configuración de búsqueda, se buscó que los comentarios de los tweets contuviesen obligatoriamente la palabra “Mahou” y, además, alguna de las siguientes palabras: “cerveza”, “una” o “la”. Esto aseguró que los tweets no solo

⁶¹ Lau, K. N., Lee, K. H., & Ho, Y. (2005). Text mining for the hotelindustry. Cornell Hotel and Restaurant Administration Quarterly

⁶² Ballesteros (2018)

mencionaran la marca, sino que también estuvieran en el contexto de la marca y que, independientemente de eso, estuviesen en español (esto fue necesario porque, aparentemente, hay un personaje de anime japonés con el mismo nombre, y se distorsionaba en gran manera los resultados de búsqueda).

Una vez establecidas las palabras clave, se generó un URL de búsqueda avanzada de Twitter que reflejara estos criterios, y fue el URL que se introdujo en Octoparse para comenzar el proceso de *web scraping*.

B) Configuración de la plataforma

En primer lugar, la URL introducida te reconduce directamente a la interfaz de Twitter de manera directa desde la plataforma, y se procede a dar acceso a la cuenta personal de Twitter para que realice la extracción sin interrupciones de ningún tipo durante el proceso. Esto se hace fijando las cookies con los datos de la cuenta introducidos, dando acceso a la información directamente desde el perfil.

A continuación, se procede a crear los *loops* de extracción. El primero se crea para que Octoparse extraiga el texto principal de cada tweet, especificando que sólo se extraiga en contenido en formato texto del tweet, sin incluir el número de likes, comentarios, u otra clase de métricas que no interesan para el estudio. A continuación, para poder capturar una gran cantidad de tweets, se configura un segundo *loop* en modo *scroll* para realizar un total de 1,000 *scrolls* permitiendo que la plataforma refrescara los tweets de manera continua y fuera extrayendo los datos de las publicaciones automáticamente.

C) Ejecución de la tarea y obtención de base de datos

Una vez configuradas todas las opciones, Octoparse ejecuta el *scraper* para comenzar la recolección de datos. La plataforma navega de manera automática por las páginas, extrayendo datos relevantes y guardándolos en formato estructurado. Estos datos recopilados, posteriormente son guardados y exportados en formatos que permitan ser analizados posteriormente (CSV o Excel), lo que facilita la importación de los datos en herramientas de análisis de texto y sentimiento, como es nuestro caso, con RStudio.

Así, después de configurar el sistema y asegurar que todo está correctamente establecido, se procedió a ejecutar la tarea, que duró aproximadamente 15 minutos en completarse. Al finalizar, se proporcionó un total de 4,851 líneas de datos que, tras eliminar duplicados,

quedaron reducidas a c.840 líneas de tweets, listas para ser procesadas y limpiadas en la siguiente etapa del análisis.

Octoparse exportó los datos recolectados en un fichero de Excel y, con estos datos exportados, se procedió a la siguiente fase del proyecto que sería realizar un análisis de sentimiento que evalúe la percepción de los consumidores sobre Mahou. Este proceso se basa en identificar patrones y determinar si los comentarios reflejan una alta lealtad y afecto hacia la marca.

3.2.2. Limpieza de datos

Limpiar los datos es un proceso esencial para preparar el texto extraído para el análisis de la manera más precisa y detallada posible.

Tras ser incorporada la extracción en RStudio (utilizando la librería “*readxl*”), el siguiente paso fue decodificar las secuencias de escape en los tweets para asegurarse que los caracteres especiales quedaban representados correctamente (mediante la librería “*stringi*”).

La limpieza del texto conllevó:

- (1) Convertir a minúsculas todo el texto, lo que asegura uniformidad y evita problemas de sensibilidad a los cambios de mayúsculas y minúsculas durante el análisis.
- (2) Eliminar menciones de usuario (@), ya que no aportan información relevante para el análisis de sentimientos.
- (3) Se eliminaron las URLs para evitar ruido en los datos, puesto que tampoco contribuyen al estudio en cuestión.
- (4) Se prescindieron de los signos de puntuación, puesto que ayuda a centrarse en las palabras clave y el contenido real de los tweets.
- (5) Eliminación de dígitos, ya que generalmente no añaden valor tampoco al análisis.

Tras la limpieza, se verifican los datos para asegurarse de que se ha realizado correctamente y que el contenido estaba en formato adecuado para el resto de fases del análisis.

3.2.3. Tokenización y creación de la Matriz DTM

Una vez que los datos han sido limpiados, el siguiente paso es la tokenización del texto y la creación de la Matriz Document-Term (DTM), que representa la frecuencia de palabras en

cada documento⁶³. Es una parte del proceso fundamental para convertir el texto en una estructura que pueda ser analizada cuantitativamente.

El primer paso es crear un corpus⁶⁴ de texto a partir de los tweets limpios. Posteriormente, se “tokeniza”⁶⁵ el texto, para obtener tokens que sean palabras individuales eliminando números y signos de puntuación. Fue este el momento en el que se eliminaron las *stopwords*, que son palabras comunes que generalmente no aportan mucho valor al análisis, como “de”, “la”, “y”. Además, se incluyeron palabras que, más tarde durante la visualización de datos, se detectaron con una presencia elevada y con poca relevancia en el caso (ej. “alguien”), y se eliminaron palabras obvias como “Mahou” o “cerveza”.

Para la creación de la matriz DTM, donde las filas eran los tweets y las columnas los tokens. Cada celda de la matriz indica la frecuencia de un término en cada documento.

3.2.4. Análisis de la frecuencia de palabras

Antes de aplicar el *stemming*, se realizó un análisis de frecuencia de palabras que identificase las 60 palabras más comunes del corpus (Figura 1 y 2).

madrid	hoy	ser	estrella	bien	duki	puedo	toda	tiempo	mierda
350	257	250	188	166	96	93	92	92	85
hace	mejor	vida	hacer	anuncio	dice	noche	sabe	tarde	momento
164	162	159	158	157	84	83	82	82	82
solo	semana	gente	quiero	casa	mientras	menos	cosas	hola	veo
156	150	149	147	145	81	80	80	79	79
ahora	cruzcampo	gusta	bar	puta	encanta	cosa	mal	verde	temporada
135	135	133	133	129	79	77	75	74	73
galicia	siempre	canción	acabo	años	viendo	mañana	forma	ayuso	abril
128	128	120	115	111	73	72	71	71	71
nunca	visto	buena	dios	vibra	mundo	ir	serie	mala	tomar
109	108	102	101	99	70	69	69	68	68

Figura 1.

⁶³ Alaminos-Fernández, A. F. (2023). Introducción a la minería de texto y análisis de sentimiento con R.

⁶⁴ Colección estructurada de textos que facilita el análisis textual, donde las filas representan los documentos y las columnas los términos *Ibid.*

⁶⁵ Proceso de dividir el texto en unidades más pequeñas, llamadas tokens. *Ibid.*

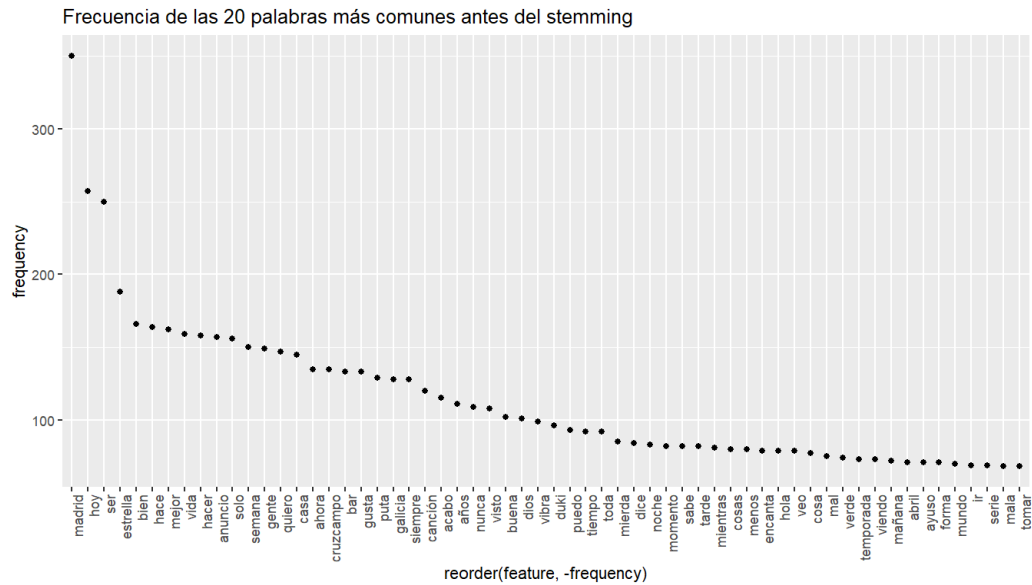


Figura 2.

Posteriormente, se generó una nube de palabras para visualizar las más comunes de manera gráfica (Figura 3). Este tipo de gráfica es una herramienta visual que muestra las palabras más frecuentes en tamaños proporcionales a su frecuencia.



Figura 3.

La palabra “Madrid” es la más prominente en la nube, lo que sugiere que gran cantidad de tweets menciona la ciudad, lo cual puede estar relacionado con la conexión que se le hace a la cerveza con la ciudad⁶⁶. “Estrella” y “Galicia”, por otro lado, sugiere que hay comparaciones y menciones a esta otra popular marca de cerveza española. “Buena” y “mejor” sugieren una percepción positiva sobre la calidad de la cerveza entre los usuarios. Son percepciones que, más tarde, se comprobarán en la siguiente fase del proceso.

Finalmente, se procede a realizar la táctica de *stemming*. Es una técnica utilizada en el procesamiento del lenguaje natural (NLP) y minería de textos que reduce las palabras a su raíz o forma base, eliminando los sufijos o prefijos⁶⁷. Se utiliza principalmente para agrupar

⁶⁶ Como dice Mahou: “Mahou no se entiende sin Madrid, y Madrid no sería lo mismo sin Mahou. Una ciudad que conocemos como nadie y que te animamos a visitar de nuestra mano”. En su web <https://www.mahou.com/this-is-madrid/>

⁶⁷ Kathiravan, P., & Haridoss, N. (2018). Preprocessing for Mining the Textual data-A Review. International Journal of Scientific Research in Computer Science Applications and Management Studies.

diferentes formas de una palabra para que puedan ser analizadas como una única entidad en el análisis de texto.

Permite⁶⁸, por un lado, reducir la dimensionalidad a su raíz, disminuyendo la cantidad de términos únicos del corpus y simplificando la matriz DTM, haciendo en análisis más manejable y eficiente. Además, mejora la precisión del análisis, agrupando diferentes formas de una palabra bajo una única raíz, evitando así la fragmentación de la información. Por último, las palabras derivadas que comparten la misma raíz son tratadas como una sola unidad, proporcionando una visión más coherente y precisa de los datos textuales.

Se empleó la librería “*quanteda*”, concretamente la función “*dfm_wordstem*”, para reducir las palabras a sus raíces, consolidando diferentes formas de la misma palabra bajo una misma entrada. Se vuelve a analizar posteriormente la frecuencia de palabras para identificar las raíces más comunes y se genera una nueva nube de palabras para visualizarlas de manera gráfica (Figura 4.)



Figura 4.

⁶⁸ Jivani, A. G. (2011). A comparative study of stemming algorithms.

4. SENTIMENT ANALYSIS

4.1. Introducción al Análisis de Sentimiento

El análisis del sentimiento o *sentiment analysis*, según DEVIKA, es un proceso intelectual de extracción de sentimientos y emociones de las personas, considerado como un campo de procesamiento de lenguaje natural (NPL). Permite clasificar textos basados en la orientación del sentimiento de las opiniones que contienen, determinando si un texto es positivo, negativo o neutral⁶⁹.

4.1.1. Niveles de análisis de sentimiento

Para JOSHI y ITKAT, el *sentiment analysis* se investiga a tres niveles principales⁷⁰:

A) Análisis a nivel documento

Clasifica si un documento de opinión expresa un sentimiento positivo o negativo, asumiendo que cada documento expresa opiniones sobre una sola entidad. Es un nivel que analiza el sentimiento general de un documento como un todo, clasificándolo como positivo, negativo o neutral, o utilizando otro sistema de clasificación. Se asume que el documento evalúa una sola entidad (como un servicio o producto), por lo que no es adecuado para documentos que traten sobre múltiples entidades a la vez⁷¹.

B) Análisis de oración

Evalúa si cada oración expresa una opinión positiva, negativa o neutral. El documento se divide en oraciones individuales para extraer la opinión de cada una. La opinión de cada oración puede clasificarse como positiva, negativa o neutral, o asignarse un valor según otro tipo de medida.

C) Análisis de aspecto

Es un análisis más minucioso que se centra directamente en la opinión misma en lugar de en los constructos lingüísticos. Este nivel de análisis, el más detallado posible, examina una

⁶⁹ Devika, M. D., Sunitha, C., & Ganesh, A. (2016). Sentiment analysis: a comparative study on different approaches. *Procedia Computer Science*.

⁷⁰ Joshi, Neha S., Itkat, Suhasini A. (2014). "A Survey on Feature Level Sentiment Analysis" (*IJCSIT International Journal of Computer Science and Information Technologies*, Vol. 5).

⁷¹ Sobrino Sande, J. C. (2018). Análisis de sentimientos en Twitter.

entidad compuesta por varios elementos o aspectos, y evalúa la opinión sobre cada uno de ellos, pudiendo tener polaridades diferentes en cada caso⁷².

Para TAMISELVI⁷³, este proceso requiere el uso de un conjunto de entrenamiento y su calidad juega un papel muy importante en la evaluación final del texto.

4.1.2. Límites en el análisis de sentimientos

Como dice SOBRINO⁷⁴, este procedimiento presenta una serie de problemas y obstáculos que han de tenerse en cuenta a la hora de obtener los resultados. Estos son, en síntesis:

A) El contexto de las palabras

La interpretación del sentimiento en las palabras a menudo depende del contexto en el que se encuentran. De esta manera, y a modo de ejemplo, la palabra “cabeza” puede tener una connotación positiva en la frase “tener la cabeza bien amueblada”, pero negativa en “perder la cabeza”. Es una dependencia del contexto que hace que sea especialmente complicado para los sistemas de análisis de sentimientos el determinar la polaridad exacta sin un contexto adecuado.

B) Negación

La presencia de palabras de negación puede cambiar completamente el sentimiento de las frases. Por ejemplo, en “la película no es buena”, el sentimiento es negativo a pesar de la presencia de la palabra “buena”, que generalmente tiene una connotación positiva. Manejar adecuadamente las negaciones es un problema recurrente en este tipo de análisis.

C) Sarcasmo o ironía

Detectar estos tonos del lenguaje es especialmente difícil para las máquinas. Frases como “¡Sólo me ha durado dos meses!” utilizan palabras que pueden parecer positivas fuera de contexto, pero en realidad expresan un sentimiento negativo. La sutileza del lenguaje sarcástico o irónico requiere una comprensión profunda y matizada del lenguaje natural por parte de los sistemas.

D) Intensificadores

⁷² Sobrino (2018)

⁷³ Tamilselvi, A. ParveenTaj, M. “Sentiment Analysis of Micro blogs using Opinion Mining Classification Algorithm “ International Journal of Science and Research (IJSR) Vol. 2 (10).

⁷⁴ Sobrino (2018)

Los intensificadores como “muy” en “muy malo” o “apenas” en “apenas aceptable”, modifican el grado de polaridad de una palabra. La interpretación de estos modificadores es extremadamente importante para una evaluación precisa del sentimiento, ya que pueden amplificar o atenuar la connotación de la palabra que acompañan.

E) Ambigüedad

La ambigüedad en las oraciones es uno de los desafíos más destacados en este ámbito, pues frases como “Juan vio un niño con un telescopio en la ventana” puede interpretarse de múltiples maneras dependiendo de quién posee el telescopio. Este tipo de ambigüedades pueden llevar a errores en la clasificación del sentimiento si no se resuelven de manera adecuada mediante el contexto y el conocimiento previo.

En definitiva, el *sentiment analysis* es una herramienta poderosa que, a pesar de sus limitaciones, puede ofrecer un profundo entendimiento de las opiniones expresadas en los textos, pero requiere un desarrollo continuo de los métodos de análisis y la integración de técnicas de contextualización que mejoren la precisión y la efectividad de los sistemas ⁷⁵.

4.2. Análisis en RStudio

Para llevar a cabo el análisis de sentimientos de los tweets, se han empleado diversas herramientas y léxicos especializados en el procesamiento del lenguaje natural (NLP) dentro del entorno de RStudio. Las librerías empleadas han sido “*quanteda*” y “*syuzhet*”, junto con el léxico NRC.

El léxico NRC (National Research Council Canada) Emoticon Lexicon es un recurso que permite asociar palabras con 8 emociones básicas y 2 polaridades de sentimiento (positivo y negativo), siendo las emociones las siguientes: alegría, tristeza, enfado, miedo, confianza, desagrado, sorpresa y anticipación ⁷⁶. Es un diccionario bastante útil a la hora de matizar las emociones presentes en los textos más allá de “positivo” o “negativo”, permitiendo no solo determinar la polaridad general del contenido, sino también identificar emociones específicas que los consumidores asocian con la marca Mahou.

Mediante la función “*get_nrc_sentiment*” se evalúa cada palabra en el texto contrastándolo con el NRC lexicón, que asigna una puntuación de sentimiento y emoción. Posteriormente, la

⁷⁵ Sobrino (2018)

⁷⁶ Tasente, T., Alina Carataş, M., & Alabdullah, T. T. Y. (2024). Análisis del sentimiento en la actividad del Banco Central Europeo en las redes sociales durante la pandemia del Covid-19 y la guerra de Ucrania: una comunicación de crisis navegante. *Doxa Comunicación*, (38).

polaridad se calcula como la diferencia entre la suma de palabras positivas y la de las negativas, donde un valor cercano al cero indica neutralidad. Es decir, se suman las palabras positivas y restan las negativas, para obtener un valor de polaridad para cada comentario, dando el resultado final las siguientes cifras:

```
positivos neutros negativos
41.10879 35.9205 22.97071
```

Todo esto se ve reflejado en la visualización de la distribución de los sentimientos mediante el histograma creado por “*ggplot2*”. Como se puede observar en la figura 4, el eje horizontal representa la polaridad de los sentimientos, que varía de positivo a negativo, y el eje vertical muestra la frecuencia de los tweets para cada valor de polaridad, es decir, cuántos tweets tienen una determinada connotación positiva o negativa.

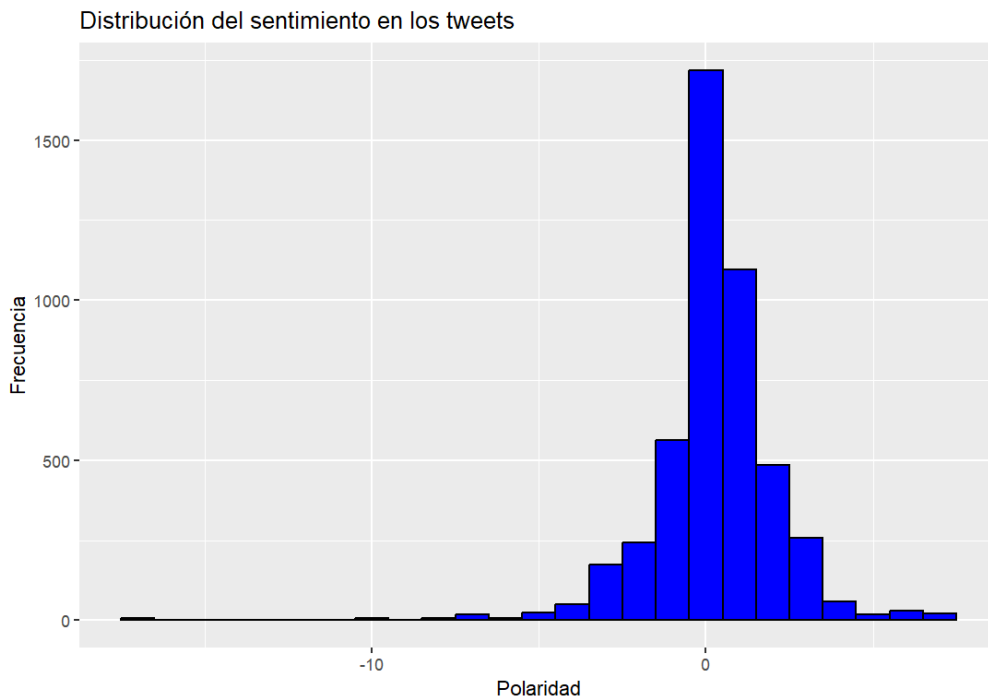


Figura 5.

El gráfico muestra cómo hay un pico principal en el valor de la polaridad cero. Esto indica que una gran cantidad de tweets son neutros, es decir, que no expresan una fuerte emoción ni positiva ni negativa hacia la marca. Por otro lado, se observa una distribución asimétrica, con una mayor concentración de tweets en el rango positivo en comparación con el negativo, lo que nos sugiere que los consumidores tienden a tener una percepción más positiva que negativa sobre Mahou. Finalmente, hay una frecuencia significativa de tweets con valores de polaridad de entre 1 y 3 puntos de positividad, lo que podría indicar que gran parte de los

usuarios tienen comentarios moderadamente positivos sobre la marca. Los valores de negatividad extrema (-10 a -5) tienen muy poca frecuencia, lo que da a entender que pocos son los tweets extremadamente negativos hacia la marca.

Por otro lado, además de la polaridad general, este análisis de emociones proporciona una visión más detallada de las reacciones emocionales de los consumidores hacia la marca. Con el diccionario NRC Lexicon, se puede identificar y cuantificar las emociones específicas presentes en los comentarios, como ya se ha mencionado anteriormente. Así, se vuelve a utilizar “*get_nrc_sentiment*” para extraer las palabras asociadas con las ocho emociones básicas, cuantificando la presencia de las mismas en los comentarios.

Con la función “*barplot*” se procede a representar las emociones identificadas durante este proceso, utilizando la paleta de colores de “*set3*” de “*RColorBrewer*” para mejorar el atractivo visual del gráfico (Figura 6):

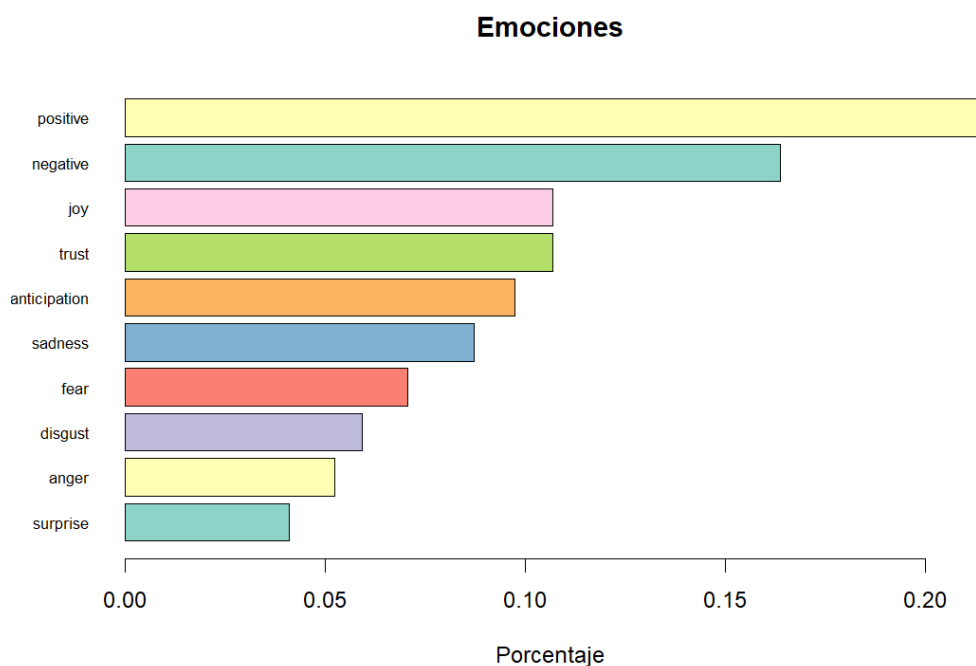


Figura 6.

En este caso, el eje horizontal representa el porcentaje de cada emoción presente en los tweets, como una proporción de comentarios que contienen palabras asociadas con cada emoción. El eje vertical, por otro lado, lista las diferentes emociones identificadas en los tweets, parte del diccionario NRC Lexicon. Se identifica, en este caso, un porcentaje de aproximadamente un 20% de emociones positivas, siendo la polaridad más prevalente en todo el conjunto. Esto sugiere que, probablemente, muchos comentarios sobre Mahou contengan

términos que reflejan una perspectiva positiva acerca de la marca, lo cual es un indicador favorable para la misma. Aproximadamente el 15% son comentarios considerados negativos. Aunque hay presencia de esta polaridad, es notablemente menor que la cantidad de términos positivos.

En cuanto a las emociones específicas, cerca del 10% es alegría, el 9% es confianza (emoción importante en la construcción de una *lovebrand*) y aproximadamente el 8% es anticipación, lo que refleja la expectativa y el interés positivo hacia futuras experiencias de la marca (lo que puede ser crucial para mantener el interés de los consumidores a largo plazo). En la parte negativa, aproximadamente el 6% representa tristeza y un 5% corresponde al miedo. El resto de emociones tienen una presencia menos relevante que el resto.

CAPÍTULO III. CONCLUSIONES GENERALES DEL ANÁLISIS

El análisis empírico realizado a lo largo de este Trabajo ha permitido obtener una percepción sobre el sentimiento de los consumidores de Mahou hacia la marca que, junto a las bases teóricas enunciadas en este estudio, ha facilitado la obtención de las siguientes conclusiones:

1º.- Mahou ha conseguido establecer una conexión emocional significativa con sus usuarios. La frecuencia de palabras como “Madrid”, “Siempre”, “Ser” o “Mejor” puede sugerirnos que los consumidores asocian la marca no sólo con un producto, sino también con otra serie de valores que, posiblemente a raíz de las estrategias y diferentes campañas llevadas a cabo, se encuentran arraigados en la sociedad, ya sea por experiencias o por cuestiones culturales. Ello refleja un nivel de confianza e identificación de marca que, como se ha expuesto previamente, es fundamental para ser considerada una *lovebrand*.

2º.- El análisis de la polaridad de los sentimientos muestra una tendencia general hacia sentimientos positivos. Como muestran los resultados obtenidos en el análisis, los tweets presentan una polaridad con incidencia mayoritariamente positiva, lo que indicaría que los consumidores tienen una percepción favorable de la marca. Si bien es cierto que la presencia de polaridad negativa es también relevante, si nos dejamos guiar por resultados cuantitativos, prevalece la positividad, lo que nos lleva a concluir que hay afecto hacia la marca.

3º.- Concluimos que, más allá de la polaridad general, encontramos diversidad emocional. El análisis de emociones utilizando el léxico NRC revela que las emociones de “alegría” y “confianza” son las más prevalentes. Como hemos visto, estas emociones son esenciales para perfilar una *lovebrand*, ya que la alegría refuerza la satisfacción del consumidor, mientras que la confianza fortalece la lealtad y disposición a seguir eligiendo la marca.

4º.- A pesar de la predominancia de sentimientos y emociones positivas, también se han identificado algunas negativas como “tristeza”, “miedo” y “enfado”. Mientras que el primero podría justificarse, tras analizar las campañas llevadas a cabo (que, como se ha constatado, buscaba “remover” emocionalmente al público), precisamente como reacciones a dichas acciones que buscan suscitar sentimientos, los sentimientos de miedo y enfado suelen estar vinculados a problemas con el producto, o expectativas no cumplidas. Sin embargo, la presencia de críticas no descalifica automáticamente a una marca de ser una *lovebrand*, pero sí indica áreas de mejora necesarias para fortalecer la relación con sus consumidores.

5°.- La alta frecuencia de menciones relacionadas con campañas específicas de Mahou demuestra la efectividad de sus estrategias de marketing emocional. Las campañas han conseguido, no solo captar la atención de los consumidores, sino también provocar interacciones emocionales y generar *engagement* en redes, lo que es un componente clave para el perfil de una *lovebrand*.

6°.- Las menciones recurrentes a otras marcas como Estrella Galicia en los comentarios sugiere que, aunque Mahou tiene un fuerte seguimiento, existen competidores que también tienen una percepción positiva significativa, lo que subraya la necesidad de que la marca siga innovando y mejorando para mantener su posición en un mercado tan competitivo.

7°.- Los resultados indican que Mahou ha conseguido obtener la confianza y lealtad por parte de sus consumidores. Es una confianza que se ve reflejada en la disposición de sus usuarios a recomendar y defender la marca frente a críticas. Sin embargo, para consolidarse plenamente como una *lovebrand* de manera generalizada ante el público, Mahou ha de continuar construyendo sobre esta base de confianza y lealtad, para responder adecuadamente a las críticas y conseguir alcanzar las expectativas de sus consumidores.

Con todo esto, podemos concluir que los resultados del análisis expuesto indican que Mahou tiene muchas características de una *lovebrand*, especialmente en términos de conexión emocional, sentimientos positivos y confianza del consumidor. Sin embargo, también revela algunas áreas a perfeccionar, especialmente respecto a la competencia con otras marcas.

Con ello, se recomienda sacarle partido a aspectos en los que sí destaca favorablemente respecto a sus competidores: la marca ha conseguido establecer una conexión emocional fuerte con sus consumidores, especialmente en Madrid, donde se asocia fuertemente con su cultura y el estilo de vida local. Además, sus campañas de marketing han sido exitosas en generar un arraigo positivo y humanizar la marca, generando emociones positivas que resuenan con su público. Los consumidores muestran altos niveles de confianza y lealtad, lo que debería ser aprovechado y maximizado a la hora de recomendar y defender la marca frente a sus críticas. La adaptabilidad e innovación de Mahou le permiten mantenerse relevante en el mercado, mientras que su integración en la vida y cultura de sus consumidores refuerza su identidad como una marca cercana y querida.

Si Mahou continúa construyendo sobre estos pilares, abordando las áreas de mejora expuestas y manteniendo su atención en la conexión emocional con su público, es muy probable que se consolide como una *lovebrand* a largo plazo en el mercado cervecero.

*** Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado**

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Jimena Gómez Carreño, estudiante de 5º de E-3 Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado " CASO MAHOU: ESTRATEGIAS DE MARKETING Y LA CONSTRUCCIÓN DE UNA LOVE BRAND A TRAVÉS DEL MARKETING EMOCIONAL Y EL ENGAGEMENT.", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Interpretador de código:** Para realizar análisis de datos preliminares.
3. **Estudios multidisciplinares:** Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
4. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
5. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
6. **Generador de problemas de ejemplo:** Para ilustrar conceptos y técnicas.
7. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 19 de Junio de 2024

Firma: Jimena Gómez Carreño

BIBLIOGRAFÍA

Artículos científicos y de opinión

- Alaminos-Fernández, A. F. (2023). Introducción a la minería de texto y análisis de sentimiento con R.
- Albert, N., Merunka, D., & Valette-Florence, P. (2013). Brand passion: Antecedents and consequences. *Journal of Business Research*.
- Arroyave, L. F. M. (2023). Las Lovebrands: Aproximación al concepto, arquitectura y mediciones. *Cuadernos del Centro de Estudios en Diseño y Comunicación. Ensayos*. Pp. 75-88.
- Ballesteros, H. F. V., Iñiguez, E. G., & Velasco, S. R. M. (2018). Minería de datos. *RECIMUNDO: Revista Científica de la Investigación y el Conocimiento*. Pp. 339-349.
- Barragán Codina, J. N., Guerra Rodríguez, P., & Villalpando Cadena, P. (2017). La economía de la experiencia y el marketing emocional: estrategias contemporáneas de comercialización.
- Belk, R. (1988). Possessions as the Extended Self. *Journal of Consumer Research*.
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*.
- Biagoni, R (2016). The SenticNet sentiment lexicon: exploring semantic richness in multi-word concepts. Vol. 4.
- Carroll, B. A. & Ahuvia, A. C. (2006). Some antecedents and outcomes of brand love.
- Cranell, W.C. (2016). A pattern-matched twitter analysis of US cancer-patient sentiments
- Christian David, F. G. (2023). El Impacto del Contenido Generado por Usuarios (UGC) en las Redes Sociales.
- Dam, T. C. (2020). The effect of brand image, brand love on brand commitment and positive word-of-mouth. *The Journal of Asian Finance, Economics and Business*. Pp. 449-457.
- De Córdoba, F. (2022). Los secretos de las marcas. Kailas Editorial.
- Delgado-Ballester, E. (2011). Development and validation of a brand trust scale.
- Devika, M. D., Sunitha, C., & Ganesh, A. (2016). Sentiment analysis: a comparative study on different approaches. *Procedia Computer Science*.
- Fellows, I., Fellows, M. I., Rcpp, L., & Rcpp, L. (2018). Package 'wordcloud'. R package version 2.
- Fernández, S. G. (2020). Estrategias comunicativas para informar y crear engagement en Instagram. El caso del periódico El País. *Razón y Palabra*,
- Gagolewski, M. (2022). stringi: Fast and portable character string processing in R. *Journal of Statistical Software*, 103. Pp. 1-59
- Gambetti, R. C., & Graffigna, G. (2010). The concept of engagement: A systematic analysis of the ongoing marketing debate. *International Journal of Market Research*. Pp. 801-826.
- Gémar, G., & Jiménez-Quintero, J. A. (2015). Text mining social media for competitive analysis. *Tourism & Management Studies*. Pp. 84-90.
- Gómez, H. D. C. (2004). Marketing de emociones. La forma para lograr fidelidad de los clientes. *Semestre económico*.
- González Fernández, M. (2021). La estrategia de marketing internacional en las empresas cerveceras españolas.

- Hearst, M. (2003). What is text mining. SIMS, UC Berkeley, 5.
- Huang, C. C. (2017). The impacts of brand experiences on brand loyalty: mediators of brand love and trust. *Management Decision*.
- Jivani, A. G. (2011). A comparative study of stemming algorithms.
- Jo, T. (2019). Text mining. *Studies in Big Data*.
- Joshi, Neha S., Itkat, Suhasini A. (2014). "A Survey on Feature Level Sentiment Analysis" (IJCSIT) *International Journal of Computer Science and Information Technologies*, Vol. 5.
- Joshi, R., & Garg, P. (2022). Assessing brand love, brand sacredness and brand fidelity towards halal brands. *Journal of Islamic Marketing*. Pp. 807-823.
- Kabacoff, R. (2022). *R in action: Data analysis and graphics with R and Tidyverse*. Simon and Schuster.
- Kabiraj, S., & Shanmugan, J. (2010). Development of a conceptual framework for brand loyalty: A Euro Mediterranean perspective. *Journal of Brand Management*.
- Kang, S., (2015). Dual Dimensions of Attitudinal Loyalty and Behavioral Loyalty among Hotel Customers. *Journal of Tourism Research & Hospitality*.
- Kathiravan, P., & Haridoss, N. (2018). Preprocessing for Mining the Textual data-A Review. *International Journal of Scientific Research in Computer Science Applications and Management Studies*.
- Kazmi, S. H. A., & Khalique, M. (2019). Brand experience and mediating roles of brand love, brand prestige and brand trust.
- Kim, H. (2022). Sentiment Analysis: Limits and Progress of the Syuzhet Package and Its Lexicons. *DHQ: Digital Humanities Quarterly*.
- Lau, K. N., Lee, K. H., & Ho, Y. (2005). Text mining for the hotel industry. *Cornell Hotel and Restaurant Administration Quarterly*
- Limpasirisuwan, N., & Donkwa, K. (2017). A structural equation model for enhancing online brand research: an international journal.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*.
- Madeline, S., & Sihombing, S. O. (2019). The impacts of brand experiences on brand love, brand trust, and brand loyalty: an empirical study. *Jurnal Bisnis dan Manajemen*. Pp. 91-107.
- Montaguti, A. C., & Lelis, C. (2017). How to Make a Lovebrand: A Tentative Recipe.
- Montalvo-Arroyave, L. F. (2023). Las Lovebrands: Aproximación al concepto, arquitectura y mediciones.
- Montoya-Restrepo, I. A., Sánchez-Torres, J. A., Rojas-Berrio, S. P., & Montoya-Restrepo, A. (2020). Lovemark effect: analysis of the differences between students and graduates in a love brand study at a public university. *Innovar*. Pp. 43-56.
- Moreno Fernández, Á., & Fuentes Lara, C. M. (2019). 'Engagement' y redes sociales. Análisis bibliométrico desde el ámbito científico de las relaciones públicas.
- Naab, T. K., & Sehl, A. (2017). Studies of user-generated content: A systematic review. *Journalism*.
- Neuwirth, E., & Neuwirth, M. E. (2014). Package 'RColorBrewer'. *ColorBrewer Palettes*.
- Niyomsart, S., & Khamwon, A. (2015). Brand Love, Brand Loyalty, and Word of Mouth: A Case of AirAsia.

- Öztürk, N., & Ayvaz, S. (2018). Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis.
- Pansari, A., & Kumar, V. (2018). Customer engagement marketing.
- Pérez Molano, G. (2023). An Overview of Web Scraping: Technical Aspects and Exercises. Computer Science.
- Ramírez Fonseca, P. A. (2022). LoveBrand.
- Ramírez, S. A. O., Veloutsou, C., & Morgan-Thomas, A. (2019). I hate what you love: brand polarization and negativity towards brands as an opportunity for brand management. *Journal of Product & Brand Management*. Pp. 614-632.
- Raposo, N. P. (1999). José Luis García Ruiz y Constanza Laguna Roldán, Cervezas Mahou, 1890-1998. Un siglo de tradición e innovación. *Revista de Historia Industrial*. Pp. 235-238.
- Rey, J. P. (2019). Mahou San Miguel: Maestros de la tradición y artesanos de la innovación.
- Roberts, K. (2005). Lovemarks: El futuro más allá de las marcas. Ediciones Urano S.A.
- Rodrigues, P., Reis, R., & Cantista, I. (2015). Consumer behavior: How the “brand love” affects you. In *Ideas in Marketing: Finding the New and Polishing the Old: Proceedings of the 2013 Academy of Marketing Science (AMS) Annual Conference*. Springer International Publishing. Pp. 583-588.
- Sobrino Sande, J. C. (2018). Análisis de sentimientos en Twitter.
- Stander, J., & Dalla Valle, L. (2017). On enthusing students about big data and social media visualization and analysis using R, RStudio, and RMarkdown. *Journal of Statistics Education*.
- Sternberg, R. (1986). A triangular theory of love. *Psychological Review*. Pp. 119–135.
- Tamilselvi, A. ParveenTaj, M. “Sentiment Analysis of Micro blogs using Opinion Mining Classification Algorithm “ *International Journal of Science and Research (IJSR) Vol. 2 (10)*.
- Tasente, T., Alina Carataş, M., & Alabdullah, T. T. Y. (2024). Análisis del sentimiento en la actividad del Banco Central Europeo en las redes sociales durante la pandemia del Covid-19 y la guerra de Ucrania: una comunicación de crisis navegante. *Doxa Comunicación*, (38).
- Uribe-Torres, G. A., Romero-Jiménez, W., Rojas-Berrio, S., & Robayo-Pinzón, O. (2022). Lovemarks e investigaciones académicas en mercadotecnia: avance de un proceso de revisión sistemática de literatura. *Clfo América*.
- Van Doorn, J. et al.(2010). Customer Engagement Behavior: Theoretical Foundations and Research Directions. *Journal of Service Research*, vol. 13 (3), Pp. 253-266.
- Vijande, L. S., & Múgica, B. G. B. (2000). Estrategias de marketing para un crecimiento rentable: casos prácticos, 103.
- Wickham, H., Bryan, J., Kalicinski, M., Valery, K., Leitiene, C., Colbert, B., ... & Bryan, M. J. (2019). Package ‘readxl’.
- Wijaya, N. H. S., & Anjasari, B. A. (2022). Brand experience and WOM: the mediating effects of brand love, brand image, and brand loyalty. *Asia Pacific Journal of Management and Education*. Pp. 48-57.
- Wong, A. (2023). Understanding consumer brand love, Brand Commitment, and brand loyalty. *Journal of Relationship Marketing*. Pp. 87-114.

Noticias

ARAL (2024). Mahou y Duki lanzan la nueva cerveza 'La Diabla'. Recuperado de: <https://www.revistaaral.com/texto-diario/mostrar/4862098/mahou-duki-lanzan-nueva-cerveza-diabla>

Marketing Directo (2018). Un Sabor Muy Grande, el homenaje de la nueva campaña de Mahou Cinco Estrellas al sabor que vive en las calles. Recuperado de: <https://www.marketingdirecto.com/marketing-general/publicidad/un-sabor-muy-grande-el-homenaje-de-la-nueva-campana-de-mahou-cinco-estrellas-al-sabor-que-se-vive-en-las-calles>

PR Noticias (2021). Mahou San Miguel presenta: Reencuentros, la cuarta entrega de su campaña 'Somos Familia' para homenajear a los hosteleros. Recuperado de: <https://prnoticias.com/2021/07/16/mahou-san-miguel-presenta-reencuentros-la-cuarta-entrega-de-su-campana-somos-familia-para-homenajear-a-los-hosteleros/>

ReasonWhy (2014). Mahou estrena su anuncio de navidad. Recuperado de: <https://www.reasonwhy.es/actualidad/campanas/mahou-estrena-su-anuncio-de-navidad-2014-12-02>

ANEXO

CODIGO RSCRIPT

1. Instalamos y cargamos las librerías

```
install.packages("readxl")  
install.packages("tidyverse")  
install.packages("quanteda")  
install.packages("syuzhet")  
install.packages("wordcloud")  
install.packages("RColorBrewer")  
install.packages("stringi")
```

```
library(readxl)  
library(tidyverse)  
library(quanteda)  
library(syuzhet)  
library(wordcloud)  
library(RColorBrewer)  
library(stringi)
```

2. Importamos y preparamos los datos

2.1. Cargamos el Excel que contiene los comentarios extraídos con Octoparse

```
ruta <- "C:/Users/Asus/Desktop/fichero_tfg/DATOS_final.xlsx"  
tweets <- read_excel(ruta)
```

2.2. Verificamos las primeras filas del archivo cargado para asegurarnos de que los datos se han importado correctamente

```
print(head(tweets))
```

2.3. Renombramos la columna para facilitar el manejo de datos

```
colnames(tweets) <- "tweet"
```

2.4. Decodificamos para asegurar que los caracteres especiales se representen correctamente

```
tweets <- tweets %>%  
  mutate(tweet = stri_unescape_unicode(tweet))
```

2.5. Verificamos las primeras filas tras la decodificación

```
print(head(tweets))
```

2.6. Limpiamos los datos para preparar el texto para el análisis

```
tweets_limpio <- tweets %>%  
  mutate(tweet = as.character(tweet)) %>%  
  mutate(tweet = tolower(tweet)) %>%  
  mutate(tweet = gsub("@\\w+", "", tweet)) %>%  
  mutate(tweet = gsub("http\\S*", "", tweet)) %>%  
  mutate(tweet = gsub("[[:punct:]]", "", tweet)) %>%  
  mutate(tweet = gsub("[[:digit:]]", "", tweet))
```

3. Text Mining

3.1. Creamos el corpus de texto a partir de los comentarios limpios

```
micorpus <- corpus(tweets_limpio$tweet)
```

3.2. Tokenizamos y eliminamos stopwords para preparar los datos para el análisis de frecuencia

```
tokens <- tokens(micorpus,  
  remove_numbers = TRUE,  
  remove_punct = TRUE) %>%  
  tokens_remove(c(stopwords("spanish"), "da", "creo", "alguien", "año", "vez", "día",  
  "cerveza", "Akogarete", "shoujo", "anime", "voy", "va", "lata", "pues", "cómo", "Mahou", "tan",  
  "así", "ver", "xq", "dos", "de", "a", "es", "una", "con", "en", "me", "del", "para", "o", "si",  
  "se", "que", "la", "el", "y", "q", "d", "x", "m"))
```

aquí hay varios términos que se han introducido a mano que, más que stopwords, son términos que se han observado que tienen relevancia en cuanto a frecuencia, pero que no aportan al estudio (ej. Akogarete, día, creo...)

3.3. Construimos la matriz DTM (Document-Term Matrix) a partir de los tokens

```
midfm <- dfm(tokens)
```

3.4. Verificamos la estructura antes del stemming

```
print(str(midfm))
```

4. Análisis de Frecuencia (Sin Stemming)

4.1. Visualizamos los términos más frecuentes antes del stemming

```
terminos_freq <- topfeatures(midfm, n = 60)
```

```
print(terminos_freq)
```

4.2. Convertimos la lista de frecuencias en dataframe

```
terminos_freq_df <- data.frame(  
  feature = names(terminos_freq),  
  frequency = unname(terminos_freq)  
)
```

4.3. Visualizamos la frecuencia con ggplot

```
ggplot(terminos_freq_df, aes(x = reorder(feature, -frequency), y = frequency)) +  
  geom_point() +  
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +  
  labs(title = "Frecuencia de las 20 palabras más comunes antes del stemming")
```

4.4. Visualizamos con una nube de palabras

```
frecuencia_palabras <- convert(midfm, to = "data.frame") %>%  
  gather(term, count, -doc_id) %>%  
  group_by(term) %>%  
  summarise(count = sum(count)) %>%  
  arrange(desc(count))
```

```
wordcloud(words = frecuencia_palabras$term,  
  freq = frecuencia_palabras$count,  
  min.freq = 4,
```

```
max.words = 40,  
colors = c("Green", "Red", "Purple", "Blue", "Orange", "Black"),  
random.order = F,  
random.color = F,  
scale = c(3, 0.1),  
rot.per = 0.3)
```

5. Análisis de Frecuencia (Stemming)

5.1. Aplicamos stemming a la matriz DTM para reducir las palabras a su raíz

```
midfm_stem <- dfm_wordstem(midfm, language = "spanish")
```

5.2. Verificamos la estructura después del stemming

```
print(str(midfm_stem))
```

5.3. Visualizamos los términos más frecuentes después del stemming

```
terminos_freq_stem <- topfeatures(midfm_stem, n = 20)  
print(terminos_freq_stem)
```

5.4. Convertimos la lista de frecuencias en un dataframe

```
terminos_freq_df_stem <- data.frame(  
  feature = names(terminos_freq_stem),  
  frequency = unname(terminos_freq_stem)  
)
```

5.5. Visualizamos la frecuencia tras realizar stemming con ggplot

```
ggplot(terminos_freq_df_stem, aes(x = reorder(feature, -frequency), y = frequency)) +  
  geom_point() +  
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +  
  labs(title = "Frecuencia de las 20 palabras más comunes después del stemming")
```

5.6. Visualización con una nube de palabras tras stemming

```
Frecuencia_palabras_stem <- convert(midfm_stem, to = "data.frame") %>%
```



```

gather(term, count, -doc_id) %>%
group_by(term) %>%
summarise(count = sum(count)) %>%
arrange(desc(count))
wordcloud(words = frecuencia_palabras_stem$term,
          freq = frecuencia_palabras_stem$count,
          min.freq = 4,
          max.words = 40,
          colors = c("Green", "Red", "Purple", "Blue", "Orange", "Black"),
          random.order = F,
          random.color = F,
          scale = c(3, 0.1),
          rot.per = 0.3)

```

6. Sentiment Analysis

6.1. Analizamos el sentimiento utilizando el paquete syuzhet

```
sentimientos <- get_nrc_sentiment(tweets_limpio$tweet, language = "spanish")
```

6.2. Incorporamos estos resultados al dataframe original

```
tweets_sentimientos <- cbind(tweets_limpio, sentimientos)
```

6.3. Calculamos la polaridad del contenido de los tweets

```
tweets_sentimientos <- tweets_sentimientos %>%
mutate(nrc_polarity = positive - negative)
```

6.4. Visualización de la distribución del sentimiento que contienen los tweets

```
ggplot(tweets_sentimientos, aes(x = nrc_polarity)) +
geom_histogram(binwidth = 1, fill = "blue", color = "black") +
labs(title = "Distribución del sentimiento en los tweets", x = "Polaridad", y =
"Frecuencia")
```

6.5. Resumen de sentimientos

```
summary_sentimientos <- tweets_sentimientos %>%
  summarise(
    positivos = 100 * sum(nrc_polarity > 0) / n(),
    neutros = 100 * sum(nrc_polarity == 0) / n(),
    negativos = 100 * sum(nrc_polarity < 0) / n()
  )
print(summary_sentimientos)
```

6.6. Analizamos las emociones de Plutchik

```
emociones <- get_nrc_sentiment(tweets_limpio$tweet, language = "spanish")
```

6.7. Añadimos estas emociones al dataframe original

```
tweets_sentimientos <- cbind(tweets_sentimientos, emociones)
```

6.8. Visualizamos las emociones

```
barplot(
  sort(colSums(prop.table(as.matrix(emociones)))),
  horiz = TRUE,
  cex.names = 0.7,
  col = brewer.pal(n = 8, name = "Set3"),
  las = 1,
  main = "Emociones",
  xlab = "Porcentaje"
)
```