



**universidad
de león**

Facultad de Ciencias
Económicas y Empresariales

Facultad de Ciencias Económicas y Empresariales
Universidad de León

Grado en Marketing e Investigación de Mercados

Curso 2020-2021

**ANÁLISIS DE SENTIMIENTO
DE LA MERCEDES-BENZ FASHION WEEK MADRID**

**SENTIMENT ANALYSIS
OF MERCEDES-BENZ FASHION WEEK MADRID**

Realizado por la alumna: Dña. Mireia Fernández Suárez

Tutelado por los Profesores: Dra. Dña. María Cristina Mendaña Cuervo
Dr. D. Enrique López González

León, a 16 de julio de 2021

MODALIDAD DE DEFENSA PÚBLICA:

Tribunal

Póster

ÍNDICE

INTRODUCCIÓN	1
METODOLOGÍA	3
CAPÍTULO I: INSTAGRAMIZACIÓN DE LOS DESFILES DE MODA	5
1.1. INFLUENCIA DE LAS REDES SOCIALES EN LA FORMA DE COMUNICARNOS.....	6
1.2. NUEVOS FORMATOS COMUNICATIVOS EN INSTAGRAM	7
1.3. IMPACTO DE INSTAGRAM EN EL MUNDO DE LA MODA.....	9
1.4. ORIENTACIÓN DE LOS DESFILES DE MODA A INSTAGRAM.....	11
CAPÍTULO II: ANÁLISIS DE SENTIMIENTO	14
2.1. ENFOQUES DEL ANÁLISIS DE SENTIMIENTO	14
2.1.1. Enfoque basado en Machine Learning.....	15
2.1.2. Enfoque basado en lexicon.....	16
2.2. APLICACIONES DEL ANÁLISIS DE SENTIMIENTO	17
2.3. LIMITACIONES DEL ANÁLISIS DE SENTIMIENTO.....	18
2.4. ANÁLISIS DE SENTIMIENTO EN RSTUDIO.....	19
CAPÍTULO III. ANÁLISIS DE SENTIMIENTO DE LA MERCEDES-BENZ FASHION WEEK A TRAVÉS DE INSTAGRAM	20
3.1. RECOLECTA DE DATOS	20
3.3. ANÁLISIS DE SENTIMIENTO.....	25
3.3.1. Análisis basado en el léxico <i>afinn</i>	25
3.3.2 Análisis basado en el léxico <i>nrc</i>	28
3.3.3 Análisis basado en la traducción del léxico <i>afinn</i>	30
3.3.4 Análisis basado en la traducción del léxico <i>nrc</i>	33
3.4. ANÁLISIS DE SENTIMIENTO A PARTIR DE EMOJIS.....	36
CONCLUSIONES	40
REFERENCIAS	42
ANEXO I: CÓDIGO PARA ANÁLISIS DE SENTIMIENTO	49

ÍNDICE DE FIGURAS

<i>Figura 1.1. Mapa del viaje del cliente (customer journey)</i>	7
<i>Figura 1.2. Instagram. Acceso directo a la función Compras</i>	11

ÍNDICE DE CUADROS

<i>Cuadro 3.1. Emojis más utilizados</i>	37
--	----

ÍNDICE DE IMÁGENES

<i>Imagen 1.1. Mapa del viaje del cliente (customer journey)</i>	7
<i>Imagen 1.2. Instagram. Acceso directo a la función Compras</i>	11
<i>Imagen 2.1. Enfoques del Análisis de Sentimiento</i>	15
<i>Imagen 3.1. Nube de palabras más comunes</i>	24
<i>Imagen 3.2. Nube comparativa según sentimiento</i>	26
<i>Imagen 3.3. Nube de palabras más comunes según emoción</i>	29
<i>Imagen 3.4. Nube de palabras según sentimiento</i>	32
<i>Imagen 3.5. Nube de palabras según emoción</i>	35

ÍNDICE DE GRÁFICOS

<i>Gráfico 3. 1. Palabras más comunes en los posts</i>	23
<i>Gráfico 3.2. Palabras más comunes en los posts (tras limpieza)</i>	24
<i>Gráfico 3.3. Clasificación de palabras más comunes según sentimiento</i>	26
<i>Gráfico 3.4. Proporción de sentimiento según la fecha</i>	27
<i>Gráfico 3.5. Palabras más comunes según emoción</i>	28
<i>Gráfico 3.6. Proporción de emoción según fecha</i>	30
<i>Gráfico 3.7. Palabras más utilizadas según sentimiento</i>	31
<i>Gráfico 3.8. Proporción del sentimiento según la fecha</i>	33
<i>Gráfico 3.9. Palabras más comunes según emoción</i>	34
<i>Gráfico 3.10. Proporción de emoción según la fecha</i>	36
<i>Gráfico 3.11. Evolución del sentimiento a lo largo del evento según emojis</i>	39

RESUMEN

Las redes sociales, y en concreto Instagram, han tenido un gran impacto en el mundo de la moda, tanto en la forma en que se consume y se habla de moda como en la forma en que las marcas se relacionan con sus clientes, lo que ha afectado a la forma de llevar a cabo el marketing para este tipo de marcas. Por este motivo, las marcas precisan extraer y analizar la información generada en las redes, siendo una de las herramientas válidas el Análisis de Sentimiento.

El objetivo de este Trabajo Fin de Grado es intentar conocer el sentimiento general del público o la audiencia de la Mercedes-Benz Fashion Week celebrada en la primavera/verano del año 2020 en Madrid, para cual se desarrolla un Análisis de Sentimiento en Rstudio, utilizando un enfoque basado en el uso de lexicons. Se extraerán los datos directamente del *hashtag* #MBFWM en Instagram, y se procederá a analizarlos con el fin de determinar el sentimiento general de quienes comentaron sobre el evento objeto de estudio. Se analizará tanto el texto escrito en dichos posts, como los *emojis* empleados para complementar el texto.

Finalmente, a partir de la información extraída de los análisis se establecerán una serie de conclusiones.

PALABRAS CLAVE

Análisis de Sentimiento, lexicon, redes sociales, Instagram, moda, Fashion Week, Madrid, MBFWM

ABSTRACT

Social networks, and specifically Instagram, have had a great impact on the world of fashion, both in the way in which fashion is consumed and talked about and in the way in which brands relate to their customers, which it has affected the way of carrying out marketing for these types of brands. For this reason, brands need to extract and analyze the information generated in the networks, being Sentiment Analysis one of the valid tools.

The objective of project is to try to know the general feeling of the public or the audience of the Mercedes-Benz Fashion Week held in the spring / summer of the year 2020 in

Madrid, for which a Sentiment Analysis is developed in Rstudio, using an approach based on the use of lexicons. The data will be extracted directly from the hashtag #MBFWM on Instagram and will be analyzed to determine the general sentiment of those who commented on the event under study. Both the text written in these posts and the emojis used to complement it will be analyzed.

Finally, based on the information extracted from the analyzes, a series of conclusions will be established.

KEYWORDS

Sentiment analysis, lexicon, social media, Instagram, fashion, Fashion Week, Madrid, MBFWM

INTRODUCCIÓN

Con el paso del tiempo y la evolución de las nuevas tecnologías las formas de comunicación entre los seres humanos han ido cambiando y evolucionando, y uno de los instrumentos de comunicación que más repercusión ha tenido sobre la forma en que los humanos interactuamos han sido las redes sociales.

Instagram es una red social que nació como una plataforma en la que compartir fotografía, pero ha ido evolucionando y desarrollando nuevos formatos comunicativos hasta tal punto que hoy en día permite a sus usuarios compartir vídeos cortos o largos, publicaciones efímeras (*Instagram Stories*), retransmisiones en directo e incluso hacer compras online.

Para cualquier tipo de evento, la comunicación entre las personas a través de las redes sociales, y especialmente Instagram, es un elemento muy importante a tener en cuenta debido a que es a través de estas donde el público general expresa sus sentimientos u opiniones. Realizar un monitoreo o un análisis de las conversaciones generadas en redes puede tener implicaciones muy positivas para las empresas y marcas, ya que permite conocer en mayor medida las necesidades, expectativas e impresiones de la audiencia.

El Análisis de Sentimiento consiste en el uso del procesamiento del lenguaje natural para la extracción de sentimientos y opiniones a partir de textos, tratando de conseguir conocer la connotación expresada en un texto, lo que permite conocer las emociones u opiniones del autor del texto. Este tipo de análisis es de especial utilidad para las empresas en el ámbito del análisis de datos en redes sociales por el gran valor de la información que consigue recoger con un coste relativamente bajo.

La Fashion Week, o Semana de la Moda, es un acontecimiento en el que, durante una semana aproximadamente, se organizan una serie de eventos y desfiles de moda. Se realiza dos veces al año en diferentes ciudades del mundo con el fin de que los diseñadores muestren sus colecciones de primavera/verano y otoño/invierno. Los desfiles de moda, que hoy en día son grandes eventos rodeados de considerable atención mediática no siempre han sido eventos tan extravagantes. En su origen en los siglos XX, XIX e incluso antes, el mayor escaparate para los diseñadores eran las mujeres que llevaban sus vestidos y accesorios cuando salían a pasear alrededor de los hipódromos (Fernández, 2016), donde serían fotografiadas y se hablaría de ellas y sus vestidos en los medios de comunicación de la época.

Con el paso del tiempo los desfiles de moda se han convertido en eventos sociales internacionales, y concretamente las Fashion Weeks se han convertido en una cita imperdible para diseñadores, *celebrities*, periodistas, profesionales y aficionados a la moda, etc.

La primera Fashion Week tuvo lugar en Nueva York en el año 1943, aunque en aquel momento fue bautizada como “*Press Week*” (García del Río, 2014). A partir de ese momento comenzó a extenderse el concepto y hoy en día se celebra dos veces al año en diferentes ciudades del mundo, siendo las más influyentes las celebradas en Nueva York, Londres, Milán, París y Miami (FashionWeekOnline [FWO], n.d.).

En el contexto de la moda española, la Semana de la Moda de Madrid o Mercedes-Benz Fashion Week Madrid (MBFWM), es uno de los eventos de moda más relevantes del año en nuestro país. Organizada por la Institución Ferial de Madrid (IFEMA) es el principal punto de encuentro para diseñadores, prensa, compradores, clientes y todo tipo de expertos y referentes de moda en España, y en este año celebró su 73 edición (IFEMA, 2020).

Con el aumento de la relevancia de Instagram en nuestra sociedad y sobre todo en el mundo de la moda, las pasarelas de moda se han tenido que adaptar a las demandas del público, por lo que en la actualidad los desfiles son planteados de forma que sean Instagrameables, posteables y aptos para ser retransmitidos en directo. Esto supone un aún mayor interés para la realización del Análisis de Sentimiento de este tipo de eventos a partir de datos extraídos de Instagram.

De acuerdo con lo anterior, el presente TFG se estructura como sigue: En el primer capítulo se realiza una conceptualización teórica basada en una breve revisión bibliográfica, con el objetivo de conocer el impacto que las redes sociales, y más en concreto Instagram, han tenido sobre el sector de la moda en los últimos años. El segundo capítulo se centra en la descripción del Análisis de Sentimiento, los enfoques que se pueden emplear para desarrollarlo y sus utilidades. Así, en el tercer capítulo se realiza una aplicación práctica de Análisis de Sentimiento al caso concreto del evento de la Mercedes-Benz Fashion Week de Madrid, exponiendo la recolección y tratamiento de datos, y la implementación realizada en RStudio. En último lugar, se expondrán las conclusiones alcanzadas como resultado del trabajo anterior.

Objetivos

De acuerdo con lo anterior, el objetivo principal de este Trabajo Fin de Grado (TFG) es determinar la polaridad en el sentimiento del público general con respecto a la Mercedes-Benz Fashion Week de Madrid (<https://www.ifema.es/mbfw-madrid>).

En relación este objetivo principal, se pueden establecer como objetivos secundarios los siguientes:

- Analizar el impacto de las redes sociales sobre el mundo de la moda.
- Comprender el verdadero valor y utilidad de los datos generados en de las redes sociales a través de su análisis.
- Obtener las palabras más características de los comentarios generados alrededor de la MBFWM.
- Conocer los emojis más característicos de los comentarios generados alrededor de la MBFW.
- Alcanzar una conclusión sobre el sentimiento del público de la MBFWM.

METODOLOGÍA

Para la elaboración del presente TFG se parte de un estado del arte basado en el interés por realizar un Análisis de Sentimiento sobre la Mercedes-Benz Fashion Week de Madrid de la temporada primavera/verano en el año 2020.

El trabajo se ha desarrollado dos partes, siendo la primera una parte teórica y la segunda una parte práctica basada y justificada por la teoría expuesta.

Para desarrollar los dos primeros capítulos (correspondientes a la parte teórica), se ha realizado una revisión bibliográfica del contenido publicado sobre los temas a tratar. Para ello, se ha recurrido a información secundaria obtenida de webs especializadas en las materias a tratar, artículos publicados en revistas, actas de conferencias, libros y estudios realizados por organizaciones como IAB Spain, de los que se han obtenido datos cuantitativos y cualitativos. También se ha recurrido a la red social Instagram, así como a su página web, y a blogs especializados tanto en temas de moda como en Análisis de

Sentimiento. Estas fuentes bibliográficas han sido consultadas tanto en castellano como en inglés.

Para la parte práctica se ha llevado a cabo, en primer lugar, un proceso de preparación a través de recursos online (vídeos de YouTube y blogs) para el manejo de las herramientas necesarias, que en este caso son R y la web utilizada para la extracción de datos. Por otra parte, dado que se ha realizado un análisis también sobre los *emojis* utilizados en el texto, ha sido necesario descargar un *diccionario de emojis*.

Una vez adquiridos los conocimientos necesarios, el siguiente paso ha sido recoger la información de Instagram para más tarde poder analizarla. Este proceso de extracción de datos se denomina *scraping* y se realizó a través de la herramienta PhantomBuster. A partir de lo anterior, se pusieron en práctica los conocimientos adquiridos sobre R, al ser aquí donde comienza en Análisis de Sentimiento como objetivo final del TFG.

CAPÍTULO I: INSTAGRAMIZACIÓN DE LOS DESFILES DE MODA

La moda, por definición, es un “uso, modo o costumbre que está en boga durante algún tiempo, en determinado país” así como un “gusto colectivo y cambiante en lo relativo a prendas de vestir y complementos” (Real Academia Española [RAE], 2020).

La moda, más allá de ser algo estético basado, como acabamos de definir, en el gusto colectivo es un reflejo de la sociedad en la que vivimos y, de la misma forma que con el paso de los años vamos evolucionando como sociedad, la moda evoluciona con nosotros. Este fenómeno es fácilmente observable si echamos la vista atrás en el tiempo sobre la moda (especialmente la femenina), y observamos como la forma de vestir de las mujeres ha ido cambiando a lo largo del tiempo de forma tan considerable que a veces, con solo ver una imagen, sabríamos determinar la época a la que pertenece. Un ejemplo sería, por ejemplo, en la época de los sesenta o setenta, la introducción y difusión dentro de la moda femenina del pantalón, lo que supuso una revolución y un cambio significativo en los roles de género que hasta entonces la sociedad tenía asumidos, es decir, un cambio en la sociedad a través de la moda (Collado, 2020).

De forma similar, las redes sociales están muy relacionadas con los cambios de nuestra sociedad. Esto puede analizarse desde dos puntos de vista; la forma en que los cambios sociales se reflejan y fomentan a través de las redes cuando los usuarios comparten publicaciones al respecto, y la forma en que las redes como tal han generado cambios en la sociedad y sus hábitos.

En el año 2020 se estimó que de media los usuarios dedican 145 minutos diarios al uso de redes sociales, casi una hora diaria más que hace 6 años, y la tendencia es creciente (Statista, 2021b). En España se estimó que el 71,2% de mujeres y el 67,5% de hombres entre 16 y 74 años utilizan internet para conectarse a redes sociales (Instituto Nacional de Estadística [INE], 2020), siendo Facebook la red social más utilizada, seguida de Instagram y YouTube (Statista, 2021a). Estos datos demuestran que, como sociedad, dedicamos una parte muy importante de nuestra vida al uso de redes sociales, lo cual puede llevar a preguntarnos cómo estas han influenciado nuestra forma de ser, actuar y comunicarnos.

En los tres primeros meses desde el lanzamiento de Instagram, esta red social alcanzó el millón de usuarios, cifra que creció rápidamente hasta alcanzar los 10 millones antes de que se cumpliera el primer aniversario de su creación. Hasta el día de hoy esta cifra ha aumentado hasta los más de 1000 millones de usuarios, posicionándose como una de las redes sociales favoritas de los usuarios.

1.1. INFLUENCIA DE LAS REDES SOCIALES EN LA FORMA DE COMUNICARNOS

Las redes sociales han tenido un innegable impacto en la forma en la que los seres humanos nos comunicamos y creamos vínculos interpersonales con la gente de nuestro entorno. Y no solo eso, sino que ha posibilitado la comunicación instantánea con personas, conocidas o desconocidas, que se pueden encontrar en el otro lado del mundo. Dado el “excesivo” uso de las redes sociales, estas continúan integrándose cada vez más en nuestro día a día, cambiando cada vez más y de forma cada vez más irreversible nuestra forma y capacidad comunicativa.

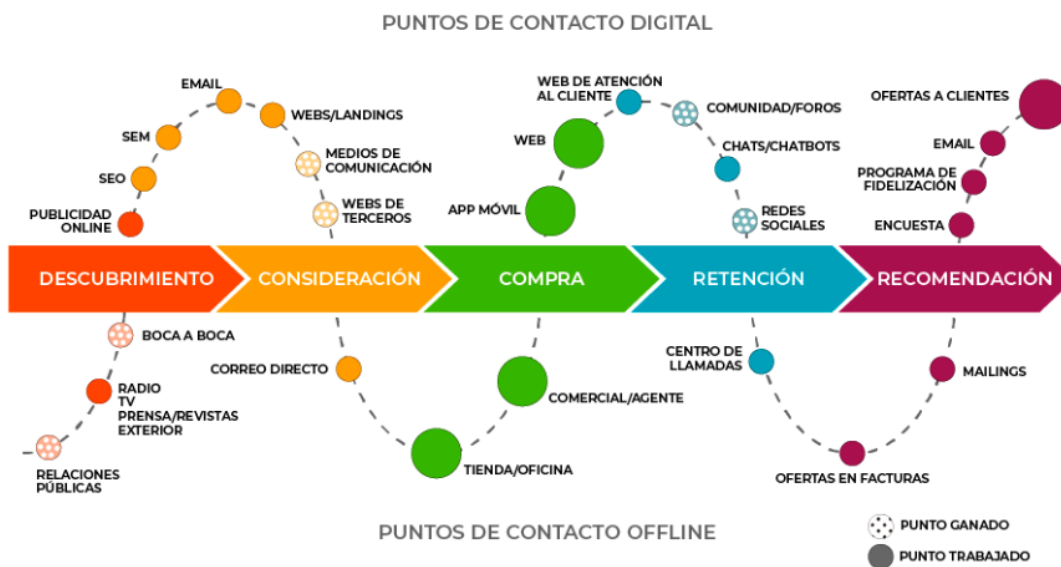
Las redes sociales han cambiado el número y tipo de personas con el que interactuamos a diario, así como la profundidad de nuestras relaciones con otras personas, al no basarse en interacciones cara a cara. Esto, en mayor o menor medida, ha supuesto una pérdida de las capacidades para la interacción social en persona. Además, nos permiten seguir y visualizar contenido de otros usuarios cuyos pensamientos van en cierto modo en línea con los nuestros, reduciendo por tanto la amplitud del espectro de opiniones y puntos de vista a los que estamos expuestos, facilitando la polarización de la sociedad (Subramanian, 2017).

Otro ejemplo de cambios producidos en nuestro modo de comunicarnos son las abreviaturas que han dado el salto de nuestras pantallas a nuestro lenguaje común (Edwards, 2015). Un ejemplo de esto último pueden ser términos como “LOL”, abreviatura en inglés para la expresión que significa “*Laughing Out Loud*”, y que hoy en día escuchamos a jóvenes de todas partes del mundo utilizar en su vocabulario habitual.

En relación con el mundo de los negocios y el marketing en concreto, las redes sociales, sobre todo YouTube e Instagram, han cambiado la experiencia del cliente, así como el conocido como viaje del cliente o *customer journey*, que es el proceso que atraviesa una

persona a la hora de realizar una compra, es decir, todos los puntos de contacto con la marca (Bel, 2019). Como se puede observar en la Imagen 1.1, durante ese proceso los consumidores tienen contacto con la marca en varias ocasiones antes de tomar la decisión final de compra o no compra, y con la evolución y crecimiento en la presencia e importancia de las redes sociales en nuestras vidas, estas han desarrollado roles muy importantes en este proceso, convirtiéndose en uno de los puntos de contacto con la marca más importante y dignos de tener en cuenta a la hora de *mapear el customer journey* para un producto o una marca. Las redes sociales han ganado un papel importantísimo lo largo de todo el proceso, desde el primer momento donde el consumidor tiene conciencia de la existencia de la marca o el producto, hasta el punto post-compra en el que se convierte en *defensor* de la marca (la recomienda), pasando por las fases de consideración, decisión, compra y regreso a la marca. Este papel de las redes sociales y el *e-commerce* ha generado cambios que afectan incluso a los procesos de compra en tiendas físicas, ya que cada vez es más común que los consumidores combinen ambos canales.

Imagen 1.1. Mapa del viaje del cliente (customer journey)



Fuente: (Territorio Digital, n.d.)

1.2. NUEVOS FORMATOS COMUNICATIVOS EN INSTAGRAM

Instagram, fundada en 2010, es una de las redes sociales más utilizadas en el mundo. En España tiene una audiencia potencial de 21 millones de usuarios, siendo el 53.7% de sus usuarios público femenino y el 46.3% público masculino (Hootsuite, 2021). Desde el

momento de su creación, donde surgió como una plataforma en la que compartir imágenes *artísticas*, hasta ahora, Instagram ha ido cambiando y evolucionando, tanto en sus funciones como en la relevancia que tiene dentro de nuestra sociedad. Con respecto a sus funciones, hay que destacar que Instagram ha pasado de ser una plataforma destinada simplemente a compartir fotografía y vídeo a convertirse en una plataforma en la que los usuarios crean comunidades, se mantienen al día de las últimas noticias y tendencias, hacen compras y comparten contenido en nuevos formatos comunicativos cada vez más innovadores que no hacen más que cambiar la forma en la que las personas nos comunicamos e interrelacionamos.

Estos nuevos formatos comunicativos permiten a los usuarios expresarse y conectar con sus seguidores, entre los que pueden encontrarse sus amigos, conocidos o *admiradores*. Estos formatos comunicativos son los siguientes (Instagram, 2021):

- *Instagram TV* (IGTV): vídeos de larga duración que permiten a los usuarios conectar mejor con su audiencia. Estos vídeos pueden durar hasta una hora y en cierto modo le hacen la competencia a YouTube, dado que permiten a los usuarios compartir contenido de diversos tipos.
- *Instagram Stories*: a través de las cuales los usuarios pueden compartir momentos del día a día de forma rápida y sencilla, teniendo la posibilidad de incluir GIF, *stickers*, música, texto e imágenes, así como efectos tanto de imagen como de vídeo y con una duración de 24 horas. La herramienta de los *stories* ha pasado a convertirse en una de las favoritas para marcas y empresas a la hora de conectar con sus clientes y audiencia.
- *Reels*: vídeos cortos de hasta 30 segundos en los que el usuario puede emplear filtros de realidad aumentada, que permiten dar rienda suelta a su imaginación. Este formato además permite la colaboración con otros usuarios a través de la función “remix” y tiene un alto potencial para la viralidad.
- *Compras*: la pestaña de Instagram en la que el usuario puede explorar, descubrir y adquirir productos sin necesidad de abandonar la plataforma para realizar la compra, convirtiéndolo por tanto en el lugar perfecto para las empresas que buscan conectar con sus compradores.

- *Messenger*: permite a los usuarios, además de enviar mensajes directos, compartir de forma privada vídeos y fotos, conversar a través de videochat e incluso compartir mensajes efímeros que desaparecen tras haber sido visualizados, etc.
- *Directos*: retransmisiones en directo que permiten a la audiencia participar a través de reacciones y comentarios. Esta herramienta ha sido y está siendo de gran importancia durante la crisis del Covid-19 para industrias como la música, que se han visto paralizadas desde principios de 2020. Los directos de Instagram hicieron posibles iniciativas como la de los festivales online que invadieron la plataforma durante el confinamiento fruto de la pandemia (Castro-Martínez et al., 2020).

Instagram es, de las redes sociales más usadas en España, en la que más *engagement* generan las cuentas de marcas, con una tasa media del 26% en el año 2019 y siendo el *retail* el sector que más interacciones genera dentro de las redes (IAB Spain, 2019). Esto, sin lugar a dudas, ha tenido un impacto significativo sobre el mundo de los negocios en general y sobre todo en el mundo de la moda.

1.3. IMPACTO DE INSTAGRAM EN EL MUNDO DE LA MODA

Los cambios y nuevos formatos descritos han tenido un claro impacto en el mundo de la moda, especialmente Instagram, que por sus características ha sido capaz de cambiar la forma en que se habla, se comparte y se consume la moda. (Ahmed, 2019).

El impacto de Instagram sobre el mundo de la moda tiene varias vertientes. Ha cambiado la forma en que los consumidores se comportan, así como la manera en que las marcas interactúan con sus clientes y la forma en que desarrollan sus estrategias de marketing. Estas nuevas estrategias de marketing suelen estar basadas en el *marketing de influencers*, que consiste en que personas con alta visibilidad y relevancia en Internet colaboren con las marcas (Gregorio, 2017). La idea detrás de esta estrategia es que los *influencers* cuentan con un público que confía en ellos y que, por tanto, está receptivo a sus recomendaciones, por lo que pueden acercar el mensaje de la marca a su público objetivo de forma más sencilla (Content Marketing Institute (CMI), n.d.).

Una clara consecuencia sobre el mundo de la moda de esta nueva forma de hacer comunicación de marca se ve en el sector del modelaje, donde Instagram y los *influencers* han generado un cambio significativo en la dinámica del sector. Sobre todo, en 2020,

cuando el mundo entero tuvo que parar debido a la pandemia del Covid-19, a las/los modelos se les empezó a pedir que crearan su propio contenido y lo publicaran en sus redes personales, del mismo modo que durante años venían haciendo los *influencers*. Incluso, en los casos en los que las/los modelos no tenían repercusión suficiente, las marcas pasaron directamente a utilizar *influencers*. Esta estrategia, para la cual no es necesario contar con fotógrafos o alquilar un lugar para el *Photoshop*, permite reducir costes, convirtiéndola así en una alternativa mucho más atractiva para las marcas (Criddle, 2021). Prueba de la notable evolución de este “fenómeno” es que, en el año 2013, la inversión a nivel mundial en *marketing de influencers* fue de 10 millones de dólares mientras que para el 2020 alcanzó la cifra de más de 8.000 millones de dólares (Statista, 2020). Además, dado que vivimos en una sociedad en la que estamos altamente condicionados por la opinión de los demás, los *influencers* y, en cierto modo también los propios consumidores, se han convertido en los responsables de crear y marcar las tendencias. Tanto es así, que alrededor de un 84% de los consumidores consultan al menos una red social antes de realizar una compra (Poncelin, 2021).

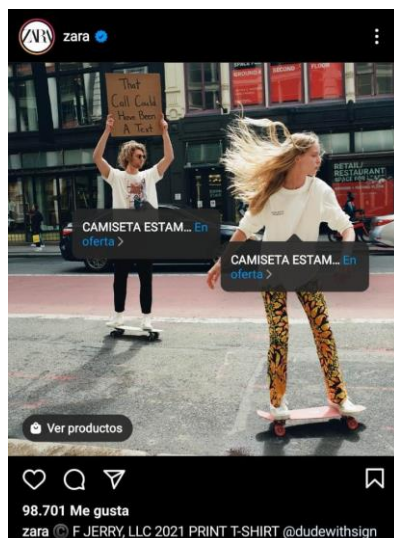
Por otra parte, dada la globalidad inherente a las redes sociales, estas tendencias son capaces de extenderse por todo el mundo a una velocidad mucho más rápida de lo que lo harían por la vía más tradicional. Al mismo tiempo, las redes sociales han hecho que los usuarios también abandonen las modas mucho más rápido, lo cual supone un reto para las marcas, que deben estar mucho más pendientes de sus consumidores y sus necesidades (Poncelin, 2021).

Otro cambio ocasionado en el mundo de la moda facilitado gracias a Instagram es lo que se conoce como “democratización de la moda”. Consiste en la eliminación de las barreras entre el público general y la moda, es decir, acercar la moda a todo el mundo, sin importar precios o tallas. Instagram ha abierto el mundo de la moda a todo tipo de usuarios y, a través de sus diferentes formatos comunicativos, permite a esos usuarios expresarse y dialogar de formas creativas (Taroy, 2015). Los más grandes diseñadores de la industria, a través de sus perfiles en redes, publican contenido con el que acercan su mundo a cualquier usuario que tenga un mínimo de curiosidad y les permiten al mismo tiempo participar en la industria (Pous, 2013).

Otra forma en que Instagram ha transformado la moda y ha logrado convertirse en su escaparate natural tiene que ver con la necesidad de la audiencia más joven de poder

adquirir los productos de forma instantánea (Imagen 1.2) algo que la red social permite a través de su función *Compras* (Valrie Moatti & Cline Abecassis-Moedas, 2018) con apenas un par de *clicks*.

Imagen 1.2. Instagram. Acceso directo a la función *Compras*



Fuente: (ZARA, 1 junio de 2021)

Además, se ha convertido también en un escaparate para marcas más pequeñas, permitiéndoles proliferar y expandirse a otros mercados geográficos a un ritmo que de otra forma resultaría casi imposible. Tanto estas como marcas más grandes tienen, además, gracias a las redes la posibilidad de conectar con los usuarios de forma más directa y personal, favoreciendo así tanto la imagen de marca como la fidelidad de los clientes.

1.4. ORIENTACIÓN DE LOS DESFILES DE MODA A INSTAGRAM

La influencia de Instagram en la moda llega hasta los desfiles, que cada vez más se organizan y orientan teniendo la red social en mente. Instagram permite que los desfiles sean accesibles para cualquier persona en el mundo en el mismo momento que están teniendo lugar (Espinosa Grau, 2020).

Esta funcionalidad de Instagram ha cambiado la concepción tradicional de las pasarelas que, antes de que entrásemos en esta era del directo en la que vivimos, eran eventos exclusivos al alcance de solo unos pocos. Hoy, cualquiera de nosotros podemos tener

acceso a desfiles y semanas de la moda de todas partes del mundo a través de nuestras pantallas. Prueba de ello es por ejemplo que durante el mes de la moda del año 2015, en Instagram se generaron 280 millones de interacciones provenientes de 42 millones de personas, todas ellas relacionadas con las semanas de la moda de Londres, Nueva York, París y Milán, que son las más relevantes a nivel global (Instagram, 2016).

Instagram supone una oportunidad para este tipo de eventos, sobre todo en relación a la repercusión que se genera a través de la red social. Por ello, los diseñadores y los organizadores de estos eventos están cada vez más centrados en crear decorados, puestas en escena e incluso prendas y accesorios que sean “*instagrammeables*”. El objetivo principalmente es crear un desfile que sea relevante en redes, que se convierta en viral, conseguir un elevado número de *likes*, comentarios, etc. (Espinosa Grau, 2020)

La orientación hacia Instagram es tan grande en el mundo de la moda que, incluso en el diseño de zapatillas de marcas como Reebok, se tienen en cuenta factores como la apariencia del zapato a la hora de sacarle una foto con el móvil en determinadas condiciones de iluminación (O’Flaherty, 2021). En este mismo sentido, Jack McCollough y Lázaro Hernández (diseñadores de la marca Proenza Schouler) afirmaron que los colores negros, por ejemplo, no permiten a las prendas destacar demasiado en las pantallas de nuestros dispositivos por lo que la solución es incluir más colores en las pasarelas. Además, es muy importante mantener la narrativa de los desfiles y espectáculos lo más interesante posible para el público. El foco ya no se pone solo en la ropa sino en cómo se traduce el desfile en la pantalla de nuestros móviles (Taroy, 2015).

Además, como se ha comentado anteriormente, la pandemia del Covid-19 ha ayudado a acelerar este proceso de *instagramización* en general y de los desfiles de moda en particular. Como consecuencia del confinamiento de las ciudades, la mayoría de los desfiles que tuvieron lugar durante 2020 y 2021 tuvieron que ser trasladados al formato digital. Por primera vez en su historia, desfiles como la Fashion Week de París no acogieron público. En su lugar, fueron retransmitidos a través de redes sociales como Instagram y otras (Arciniegas, 2020). Según la revista VOGUE, el British Fashion Council (BFC) fue el primero en anunciar que su semana de la moda masculina tendría lugar a través de un formato online (Browchuk, 2020), pero han sido muchos los eventos y semanas de la moda que se han desarrollado de esta forma durante el último año.

El formato íntegramente digital ha supuesto un gran reto creativo para todas las marcas, incluso las más grandes. Algunas de ellas, como Gucci o Tommy Hilfiger, decidieron ni siquiera participar en la semana de la moda y esperar a que la situación se estabilizara. En cambio, otras, como Moschino, optaron por explotar su creatividad al máximo creando desfiles sin modelos ni público y utilizando marionetas en su lugar.

Después de los esfuerzos realizados durante este año, hay expertos que creen que el formato digital ha llegado para quedarse, como explicó la CEO del BFC en una entrevista a la agencia de noticias Reuters (Gumuchian, 2021).

CAPÍTULO II: ANÁLISIS DE SENTIMIENTO

El Análisis de Sentimiento u *Opinion Mining* se basa en el análisis de opiniones, sentimientos, actitudes y emociones a partir del lenguaje escrito (Liu, 2012). Se enmarca dentro del campo del Procesamiento del Lenguaje Natural (PLN) y consiste, a grandes rasgos, en la clasificación automática de un volumen masivo de documentos en función de la connotación contenida en los mismos. A través de este análisis es posible extraer información de gran valor para negocios y marcas (Intelligent.es, 2017). Cobra importancia y relevancia especialmente con el desarrollo de las redes sociales, donde los usuarios comparten sus opiniones, generándose gran volumen de datos de forma constante.

El número de publicaciones respecto al Análisis de Sentimiento comienza a incrementarse de forma significativa alrededor del año 2006. A pesar de existir antecedentes previos, no es hasta entrado el siglo XXI que comienza a ganar popularidad, coincidiendo además con la aparición de la posibilidad de automatizar el proceso (Mäntylä et al., 2018). A día de hoy se muestra como una herramienta válida para la extracción de información con un coste relativamente bajo.

En relación con el mundo del marketing, el Análisis de Sentimiento ha sido utilizado en numerosas ocasiones para realizar estudios sobre temas, ámbitos y marcas muy diferentes. Así, por ejemplo, se han realizado análisis para determinar los principales desafíos que enfrenta el marketing en relación con el movimiento #MeToo (Reyes-Menendez et al., 2020), para conocer y comparar el sentimiento y la actitud del público con respecto a las marcas de ropa Nike y Adidas a través de comentarios publicados en Twitter (Rasool et al., 2019) o para conocer opiniones sobre las ofertas de #BlackFriday (Saura et al., 2018). En el mundo de la moda también son muchas las referencias, como por ejemplo análisis de imágenes en *Instagram* (AbdelFattah et al., 2017), o el mundo de la moda en Amazon (Lee et al., 2017), entre otros muchos.

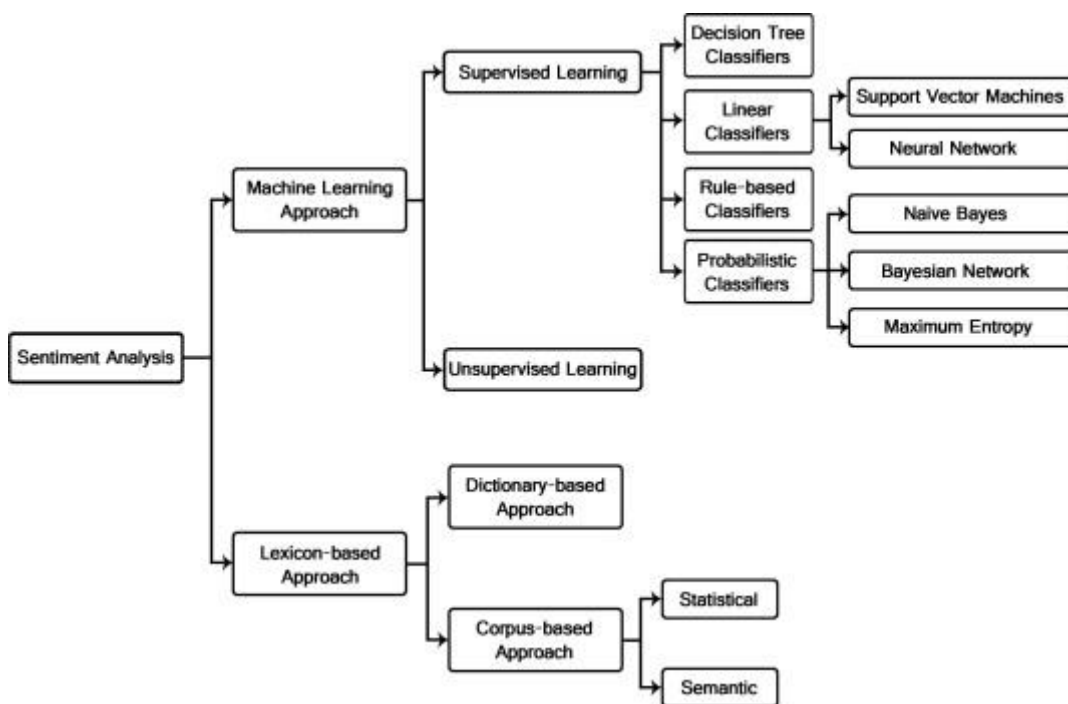
2.1. ENFOQUES DEL ANÁLISIS DE SENTIMIENTO

El Análisis de Sentimiento puede llevarse a cabo a partir de múltiples plataformas o aplicaciones como, por ejemplo, MonkeyLearn (<https://monkeylearn.com/>), MeaningCloud (<https://www.meaningcloud.com/>) o Cloud Natural Language (de

Google) (<https://cloud.google.com/natural-language>), entre otras muchas. Además, se puede realizar “manualmente” a partir de entornos y lenguajes de programación como Rstudio de R (<https://www.rstudio.com/>) o Python (<https://www.python.org/>).

Por otra parte, existen múltiples enfoques o técnicas que se pueden emplear desarrollarlo. Como se puede observar en la Imagen 2.1, en primer lugar, es factible optar entre utilizar un enfoque basado en el uso del Machine Learning o en la utilización de un Lexicon. Además, para conseguir mejores resultados, existe la posibilidad de utilizar un método híbrido que combine el enfoque del *Machine Learning* junto con el enfoque Lexicon.

Imagen 2.1. Enfoques del Análisis de Sentimiento



Fuente: (Walaa Medhat, Ahmed Hassan, 2014)

2.1.1. Enfoque basado en Machine Learning

El *Machine Learning* es “una rama de la *inteligencia artificial* basada en la idea de que los sistemas pueden aprender de datos, identificar patrones y tomar decisiones con mínima intervención humana” (SAS Institute, n.d.). Los enfoques basados en esta tecnología pueden ser supervisados o no supervisados. La diferencia se encuentra en que el aprendizaje supervisado consiste en entrenar al algoritmo a través de predicciones iteradas, necesitando intervención humana directa, mientras que el aprendizaje no

supervisado encuentra patrones escondidos sin necesidad de dicha intervención (Delua, 2021). Dentro de los métodos supervisados, existen diferentes tipos de clasificadores, dentro de los cuales destacan los clasificadores basados en Máquinas de Soporte Vectorial (*Support Vector Machines, SVM*); el clasificador *Bayesiano Ingenuo* y el clasificador de *Máxima Entropía*.

Los *clasificadores basados en SVM* son una herramienta que encuentra un hiperplano de separación que permite distinguir a qué clase pertenecen los datos analizados. Parten de *vectores de soporte*, que son los puntos de entrada que el sistema mapea (Betancourt, 2005). Para dibujar ese hiperplano, las SVM transforman los datos con la ayuda de funciones matemáticas denominadas *kernels*. El proceso para construir un modelo parte de la recolecta de datos para entrenar y testear la SVM. Luego, se vectorizan los datos y, por último, se crea un modelo, se entrena, y se realizan las predicciones (Vasista Reddy, 2018).

El *clasificador Bayesiano Ingenuo* consiste en utilizar la regla de Bayes como norma general para la clasificación de los documentos. Para entrenar al clasificador se analiza un conjunto de datos que han sido previamente clasificados de forma correcta (Elements of AI, n.d.).

Por su parte, el *clasificador de Máxima Entropía* busca mantener tanta incertidumbre como sea posible, y consiste en la selección del modelo con mayor entropía para realizar la clasificación.

2.1.2. Enfoque basado en lexicon

Los *enfoques basados en léxicos* se basan en el uso de léxicos de sentimiento prediseñados con los que calificar las palabras incluidas en los documentos analizados. Dicho léxico contiene puntuaciones para cada palabra, y la forma de calcular el sentimiento de los documentos analizados es agregando la puntuación asociada a cada palabra contenida dentro de la unidad analizada. Este enfoque presenta como ventaja sobre el enfoque del *Machine Learning* que no requiere entrenamiento de los datos.

Dentro del enfoque basado en léxicos, podemos distinguir entre un enfoque basado en *diccionarios* o un enfoque basado en *corpus*, que al mismo tiempo puede ser *semántico* o *estadístico*.

El subenfoque de *diccionarios* consiste en el uso de listados de términos para etiquetar palabras, mientras que basado en *corpus* emplea coocurrencias estadísticas o patrones sintácticos incorporados dentro de un *corpus*.

Para realizar el análisis mediante este enfoque, en primer lugar, es necesario que se divida el texto que se va a analizar en unidades únicas (palabras o frases) y, una vez hecho esto, se cruzan esas palabras con el léxico que se haya decidido emplear. De este modo, se obtiene, para cada palabra, una puntuación relativa a la polaridad de la palabra. Todas estas puntuaciones agrupadas darán lugar a la puntuación global del texto analizado (Pauli, 2019).

Existen varios léxicos diferentes y la elección entre uno u otro dependerá del investigador, así como del objetivo del análisis. Existen léxicos que puntúan las palabras en función de su polaridad (positiva/negativa), pero hay otros léxicos que puntúan, por ejemplo, las emociones expresadas por las palabras.

Así, el *léxico AFINN*, creado por Finn Årup Nielsen, contiene un listado de palabras con puntuaciones entre +5 y -5, siendo aquellas palabras con puntuación -5 las más negativas y aquellas con puntuación +5 las de connotación más positiva (Nielsen, 2015).

Por otra parte, el *léxico NRC* o *National Research Council - Emotion Lexicon*, es un listado de palabras asociadas con ocho emociones diferentes: enfado, miedo, anticipación, confianza, sorpresa, tristeza, felicidad y asco; así como dos sentimientos: positivo y negativo (National Research Council Canada, 2019).

Existen otros muchos léxicos como pueden ser el listado Bing (Hu & Liu, 2004), el listado SENTIWORDNET 3.0 (Baccianella et al., n.d.) o el listado ANEW (Bradley & Lang, 2017), que se pueden conseguir fácilmente en la red.

2.2. APLICACIONES DEL ANÁLISIS DE SENTIMIENTO

Uno de los motivos por el que el Análisis de Sentimiento ha ido ganando relevancia en los últimos años es porque se ha convertido en una herramienta que puede ser de utilidad

en multitud de áreas. En política, por ejemplo, se puede utilizar para analizar la opinión pública durante las campañas electorales y prever los resultados.

En el mundo de los negocios es de especial utilidad, sobre todo en el área de Marketing e Investigación de Mercados. Este tipo de análisis puede aportar gran valor a empresas y marcas, dado que permite analizar y monitorear no solo la opinión o sentimiento de los consumidores de la propia marca, sino también la opinión de los consumidores de la competencia (Fan & Gordon, 2014). También puede ser útil para llevar a cabo investigaciones para la optimización de productos y la experiencia del cliente (Setuserv, n.d.).

Además, puede ser utilizado también en otras áreas de la empresa, como por ejemplo Servicio al Cliente, donde el Análisis de Sentimiento y la Clasificación de Textos permite dirigir automáticamente las consultas al departamento adecuado para gestionarlas (MonkeyLearn, 2020).

Desde el punto de vista del consumidor se puede emplear para resumir reseñas online, de forma que el usuario podrá saber cuál es la opinión general de quienes han dejado una reseña sobre un servicio o producto sin necesidad de leer todas, y así tomar la decisión de compra o no compra más fácilmente (Ding et al., 2008).

2.3. LIMITACIONES DEL ANÁLISIS DE SENTIMIENTO

A pesar de que desde sus orígenes hasta el momento actual el Análisis de Sentimiento ha ido desarrollándose y mejorando, aún tiene limitaciones en cuanto a su efectividad y precisión. El sentimiento es algo subjetivo de cada persona, y las palabras tienden a cambiar su significado dependiendo de factores como el contexto o el tono en que los utilicemos. Por ello, en el lenguaje natural hay determinados elementos que aún suponen un reto para este tipo de análisis.

Al expresarnos, las personas a veces utilizamos falsos negativos. En ocasiones para dar una opinión negativa utilizamos palabras negativas y las precedemos con un término de negación, y viceversa. De este modo, si decimos “No me disgusta”, estamos utilizando un verbo que sería clasificado como negativo (disgustar) para expresar un sentimiento positivo (Jiménez Zafra et al., 2014).

En este sentido, decir palabras como “mierda” u otras malsonantes que, en principio serían negativas, pueden ser utilizadas con una connotación positiva dependiendo de la frase (Ochieng et al., 2016). De este modo, la frase hecha “mucho mierda” (expresión utilizada para desear suerte en el mundo del teatro) al contener la palabra “mierda” sería clasificada como negativa, cuando en realidad, prácticamente cualquier persona podría entender que tiene un significado positivo. Otra característica común del lenguaje natural que el Análisis de Sentimientos no es capaz de detectar es el sarcasmo o la ironía.

2.4. ANÁLISIS DE SENTIMIENTO EN RSTUDIO

Rstudio “es un entorno de desarrollo integrado para el lenguaje de programación R y Python” que cuenta con una consola y otras herramientas de trabajo para el usuario (Rstudio, n.d.) y es una de las herramientas a disposición de los usuarios para realizar el Análisis de Sentimiento. Por ese motivo y por su gratuidad es la herramienta por la que se ha optado para desarrollar la parte práctica de este TFG.

Dentro de Rstudio existen diferentes paquetes que se pueden emplear para la realización de este tipo de análisis. Un paquete de R es una colección de funciones y datos almacenados en una carpeta y que se pueden descargar de un repositorio de forma gratuita (Santana & Hernández, 2016). Para desarrollar la parte práctica del presente TFG, a cuyo detalle se dedica el siguiente capítulo, utilizaremos, entre otros, tres paquetes bastante conocidos: *tidytext*, *tidyverse* y *wordcloud*.

El paquete *tidytext* persigue el objetivo de ayudar al usuario a trabajar con datos ordenados o *tidy data*, ya que trabajar con datos de esta forma facilita mucho las tareas de la minería de datos (Silge & Robinson, 2021). El paquete incluye funciones que permiten la conversión de textos en formatos ordenados (CRAN-r, n.d.-a).

El paquete *tidyverse* es una colección de otros paquetes, que tiene la sencilla función de permitir instalar varios paquetes en un único paso (CRAN-r, n.d.-b). Esos otros paquetes serán necesarios para la realización del Análisis de Sentimiento.

Por último, el paquete *wordcloud* tiene la finalidad de permitir al usuario crear nubes de palabras (CRAN-r, n.d.-c), que son muy útiles para la visualización de los datos en este tipo de análisis.

CAPÍTULO III: ANÁLISIS DE SENTIMIENTO DE LA MERCEDES-BENZ FASHION WEEK A TRAVÉS DE INSTAGRAM

3.1. RECOLECTA DE DATOS

Dado el carácter visual y tan dado a la fotografía de un evento de moda como es la Mercedes-Benz Fashion Week de Madrid, hemos optado por realizar este análisis a partir de datos recogidos de la red social Instagram.

Para la obtención de dichos datos se ha recurrido a la plataforma PhantomBuster (<https://phantombuster.com/>), que es un software online a partir del cual es posible extraer datos de varias plataformas como Instagram, Twitter o LinkedIn.

Hemos optado por descargar todas aquellas publicaciones que incluyan el hashtag #MBFWMADRID, ya que es el hashtag oficial del evento. Dichos posts se han descargado en un fichero formato .csv a partir del cual podremos trabajar posteriormente en Rstudio (<https://www.rstudio.com/>).

3.2. TRATAMIENTO DE DATOS

Una vez recogidos los datos, procedemos a la limpieza y tratamiento de estos.

En primer lugar, utilizando Excel, vamos a filtrar la base de datos en función de la fecha. Dado que desde PhantomBuster no es posible elegir el rango de fechas para las publicaciones, hemos descargado un volumen de datos en bruto que abarca más del periodo en el que estamos interesados. Por eso, filtramos la columna correspondiente a la fecha de publicación de los posts y elegimos mantener solo las publicaciones correspondientes a las fechas entre el 8 de abril de 2021 y el 11 de abril de 2021 (fechas oficiales del evento). Además, al descargar los datos, PhantomBuster recoge una serie de variables que no nos interesan para nuestro análisis y que, por tanto, procedemos a eliminar, conservando únicamente la fecha de publicación de los datos y el texto de los *posts*, que vienen recogidas en las columnas "pubDate" y "description", respectivamente.

Con la base de datos obtenida, ya podemos proceder con Rstudio, para lo cual instalamos todos los paquetes que vamos a utilizar para el análisis, que son principalmente, como se ha comentado en el capítulo anterior, *tidytext*, *tidyverse* y *wordcloud*, además de los siguientes:

```
library(tidytext)
library(dplyr)
library(tidyverse)
library(wordcloud)
library(reshape2)
library(RColorBrewer)
library(lubridate)
library(zoo)
library(scales)
```

El siguiente paso será proceder a limpiar los datos. El objetivo de esta limpieza será eliminar del texto símbolos como “@” o “/”, así como símbolos de puntuación y números. Para ello, procedemos con las siguientes líneas de código, en las que se incluye la función ***gsub***, que se encargará de reemplazar en la tabla todas aquellas ocurrencias de un determinado argumento por otro argumento indicado. Dado que el fin es eliminar esas ocurrencias para dejar de trabajar con ellas, le indicaremos que el nuevo argumento estará vacío:

```
datosplit <- data.frame(datos$pubDate)
datosplitdesc <- data.frame(datos$description)
cleandescrip <- map_df(datosplitdesc, ~gsub("[:punct:]", "", .x))
cleandescrip <- map_df(cleandescrip, ~gsub("[0-9]", "", .x))
cleandescrip <- map_df(cleandescrip, ~gsub("https", "", .x))
datosclean <- data.frame(datosplit, cleandescrip)
```

Una vez realizada la limpieza de los datos contenidos en la tabla, vamos a proceder a dividir las publicaciones en palabras únicas, utilizando la función ***unnest_tokens***:

```
datos_token <- unnest_tokens(tbl = datosclean, output = "word", input =
  "description", token= "words", to_lower = TRUE)
```

De esta forma, hemos obtenido un nuevo *dataframe*, en el que tenemos 47326 filas, correspondientes al número total de palabras incluidas en las publicaciones.

A mayores, con el fin de poder comparar el sentimiento por fechas más adelante, vamos a proceder a utilizar la función ***separate*** para dividir la variable `datos.pubDate` en dos variables nuevas que serán Fecha y Hora:

```
datos_token <- datos_token %>% separate(datos.pubDate, into = c("Fecha",
  "Hora"), sep = " ")
```

Utilizamos la función `count()` para calcular la frecuencia de las palabras y ordenarlas en orden de mayor a menor incidencia:

```
count(datos_token, word, sort = TRUE)
```

Observamos que las palabras que representan una mayor frecuencia son palabras irrelevantes para el estudio, como preposiciones, artículos, etc. Estas palabras se denominan *stopwords*, y son palabras vacías de significado relevante, por lo que vamos a filtrarlas y eliminarlas antes de continuar con el análisis de los datos. Lo haremos con ayuda de la función `anti_join()` y siguiendo el siguiente proceso:

1. En primer lugar, vamos a instalar los paquetes **quanteda** y **tm** para poder obtener las *stopwords* en español, ya que en el paquete *tidytext* solo vienen incluidas *stopwords* para el inglés:

```
install.packages("quanteda")
install.packages("tm")
library(quanteda)
library(tm)
```

2. A continuación, utilizamos la siguiente línea para crear la lista de *stopwords* en español:

```
mystopwords <- bind_rows(stop_words, data_frame(word =
tm::stopwords("spanish"), lexicon = "custom"))
```

3. Por último, procedemos a eliminar las *stopwords* del *dataframe*, quedándonos ahora con 30041 palabras para proceder con el análisis:

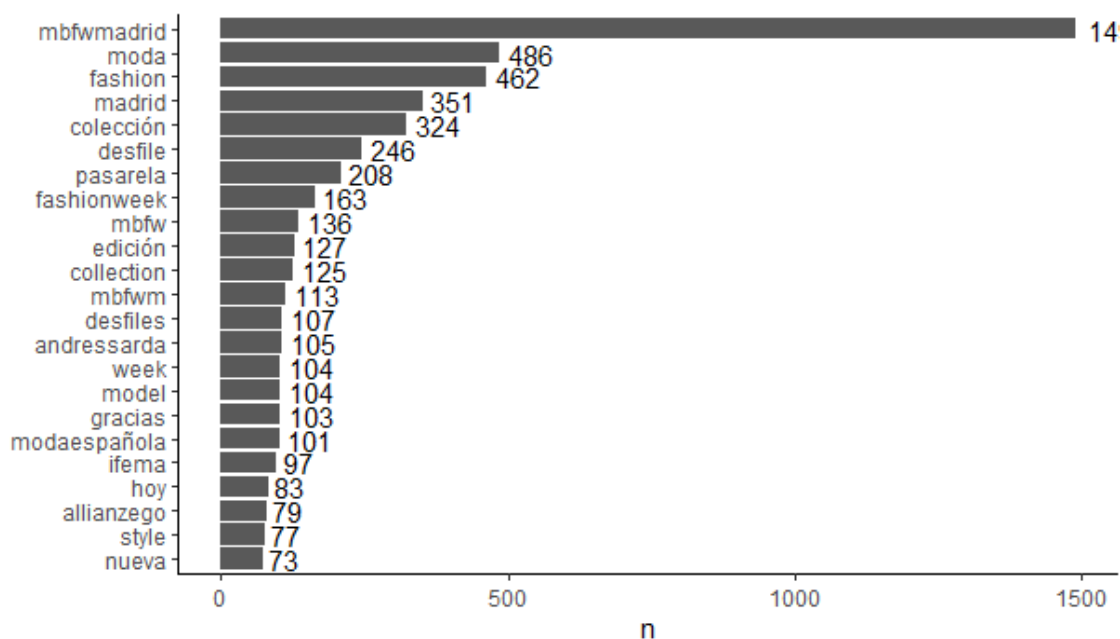
```
datos_token <- anti_join(x=datos_token, y= mystopwords, by="word")
```

De esta forma hemos eliminado *stopwords* tanto del vocabulario español como del inglés, ya que estamos trabajando con texto en ambos idiomas.

Con el fin de observar aquellas palabras que han sido utilizadas con mayor frecuencia (una vez limpiado los datos) hemos elaborado el Gráfico 3.1., utilizando el paquete *ggplot* y la siguiente línea de código:

```
datos_token %>% count(word, sort = TRUE) %>% filter(n > 70) %>%
mutate(word = reorder(word, n)) %>% ggplot(aes(word, n)) +
geom_text(aes(label=n), hjust= -0.2) + geom_col() + xlab(NULL) +
coord_flip()+ theme_classic()
```

Gráfico 3.1. Palabras más comunes en los posts



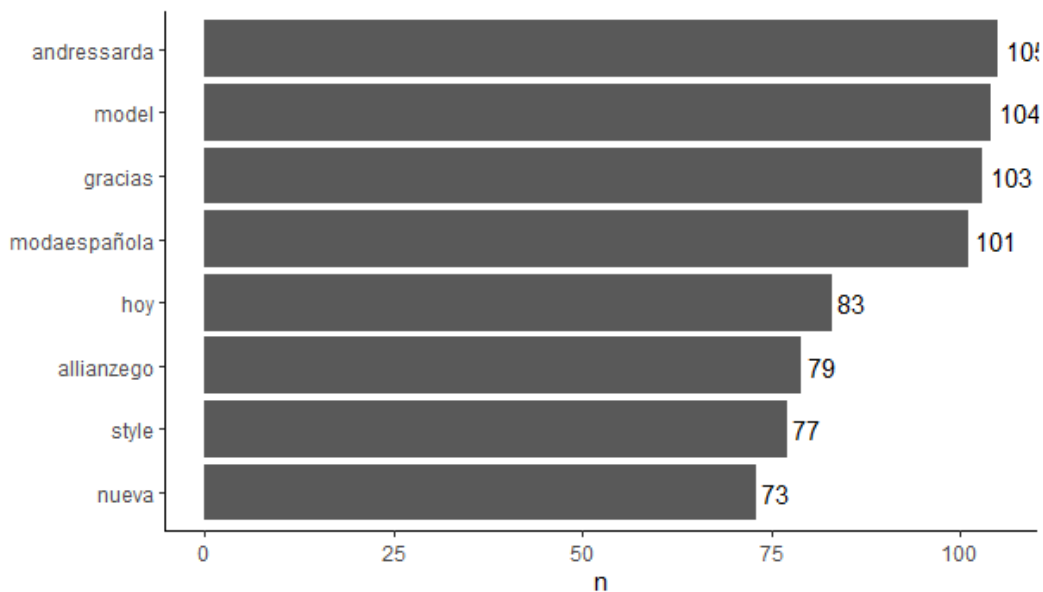
Fuente: Elaboración propia con Rstudio

Como se puede observar, las palabras utilizadas con mayor frecuencia son palabras como “mbfwmadrid”, “moda”, “fashion”, “madrid”, “colección”, “desfile”, “pasarela”, “fashionweek”, “mbfw”. Estas palabras pertenecen al vocabulario común relativo al evento y no denotan ningún tipo de sentimiento, por lo que las consideramos irrelevantes y procederemos a eliminarlas para continuar con el análisis. Para eliminar estas palabras usaremos la siguiente línea de código:

```
datos_token <- filter(datos_token, word!="mbfwmadrid" & word!="moda" &
word!="fashion" & word!="madrid" & word!="colección" & word!="desfile" &
word!="pasarela" & word!="fashionweek" & word!="mbfw" &
word!="collection" & word!="edición" & word!="mbfwm" & word!="2021" &
word!="desfiles" & word!="week" & word!="2021" & word!="ifema" &
word!="mercedes" & word!="benz")
```

Una vez eliminadas las palabras que hemos considerado irrelevantes, volvemos a ejecutar el gráfico, y obtenemos el resultado que se muestra en el Gráfico 3.2., en el que se puede observar que la palabra con mayor incidencia es "andressarda", que hace referencia a Andrés Sardá, un diseñador de moda de baño y lencería.

Gráfico 3.2. Palabras más comunes en los posts (tras limpieza)



Fuente: Elaboración propia con Rstudio

A continuación, hemos procedido a la elaboración de nubes de palabras para una mejor visualización de los datos, para lo cual utilizando la función *wordcloud* y con el código siguiente hemos realizado la primera nube de palabras (Imagen 3.1) en la que tenemos todas las palabras contenidas en la base de datos.

```
datos_token %>% count(word) %>% with(wordcloud(words=word, freq=n,
max.words = 250, scale = c(3,1), rot.per = 0.3, random.order = FALSE,
colors=brewer.pal(6, "Dark2")))
```

Imagen 3.1. Nube de palabras más comunes



Fuente: Elaboración propia con Rstudio

3.3. ANÁLISIS DE SENTIMIENTO

En este apartado se describe el proceso realizado del Análisis de Sentimientos del evento objeto de estudio. Dado que estamos utilizando el paquete *tidytext*, el cual contiene varios léxicos de los cuales podremos obtener los sentimientos de nuestras palabras, pero todos ellos en inglés, vamos a proceder en dos pasos. En primer lugar, vamos a analizar los datos en inglés y, en segundo lugar, procederemos con los datos en español, ya que el proceso tendrá que ser ligeramente diferente.

Para ello, vamos a utilizar dos de los léxicos incluidos dentro del paquete *tidytext*, que son *afinn* y *nrc*. Para descargar los léxicos, usaremos la siguiente función y después procederemos a realizar primero el análisis basándonos en el léxico *afinn* y después basándonos en el léxico *nrc*.

```
get_sentiments("nrc")
get_sentiments("afinn")
```

3.3.1. Análisis basado en el léxico *afinn*

Con la ayuda de la función *inner_join()* procedemos a relacionar cada palabra de nuestra base con las de los léxicos, y de este modo les asociamos su correspondiente sentimiento:

```
datos1_afinn <- datos_token %>% inner_join(get_sentiments("afinn"))
datos1_afinn %>% count(word, value, sort=TRUE)
```

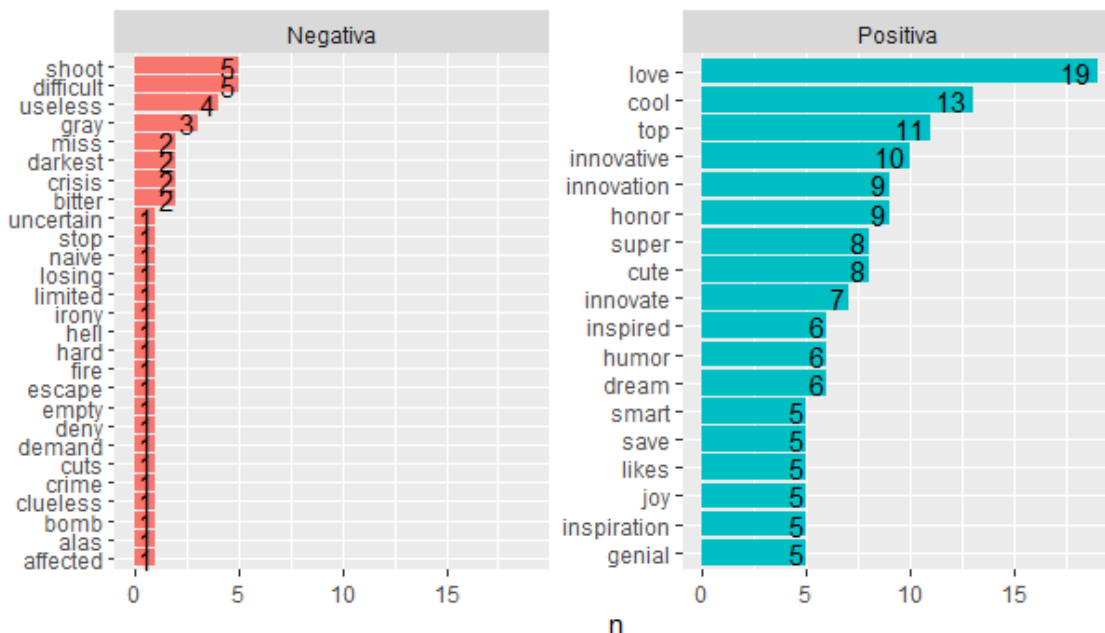
Con la siguiente línea, vamos a crear una nueva columna que determinará si el sentimiento expresado por la palabra es positivo o negativo, en función del valor (entre -4 y 4) asignado a cada palabra:

```
datos1_afinn <- datos1_afinn %>% mutate(Tipo=ifelse (value >0,
"Positiva", "Negativa"))
```

En el Gráfico 3.3 podemos observar, en función de su clasificación como palabras positivas o negativas, cuáles han sido las más repetidas:

```
datos1_afinn %>% count(word, Tipo, sort = TRUE) %>% group_by(Tipo) %>%
top_n(15) %>% ungroup () %>% mutate(word=reorder(word, n)) %>%
ggplot(aes(word, n, fill=Tipo)) + geom_col(show.legend = FALSE) +
geom_text(aes(label=n), hjust=1.2) + facet_wrap(~Tipo, scales =
"free_y")+ coord_flip() + xlab(NULL)
```

Gráfico 3.3. Clasificación de palabras más comunes según sentimiento

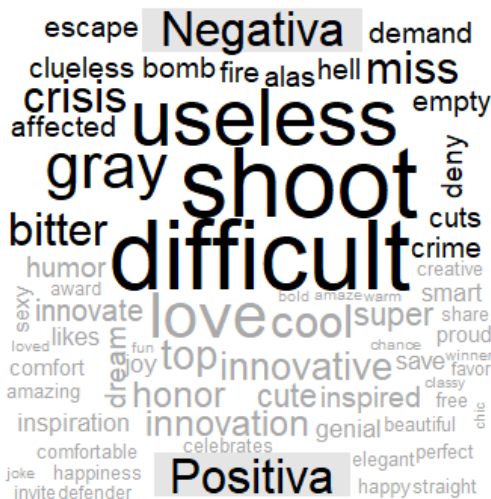


Fuente: Elaboración propia con Rstudio

A continuación, hemos procedido a elaborar una nube de palabras (Imagen 3.2) en la que clasificaremos las palabras positivas y negativas según el léxico afinn:

```
datos1_afinn %>% count(word, Tipo, sort=TRUE) %>% acast(word~Tipo, value.var = "n", fill = 0) %>% comparison.cloud(colors = c("black", "dark grey"), max.words = 150, title.size = 2)
```

Imagen 3.2. Nube comparativa según sentimiento



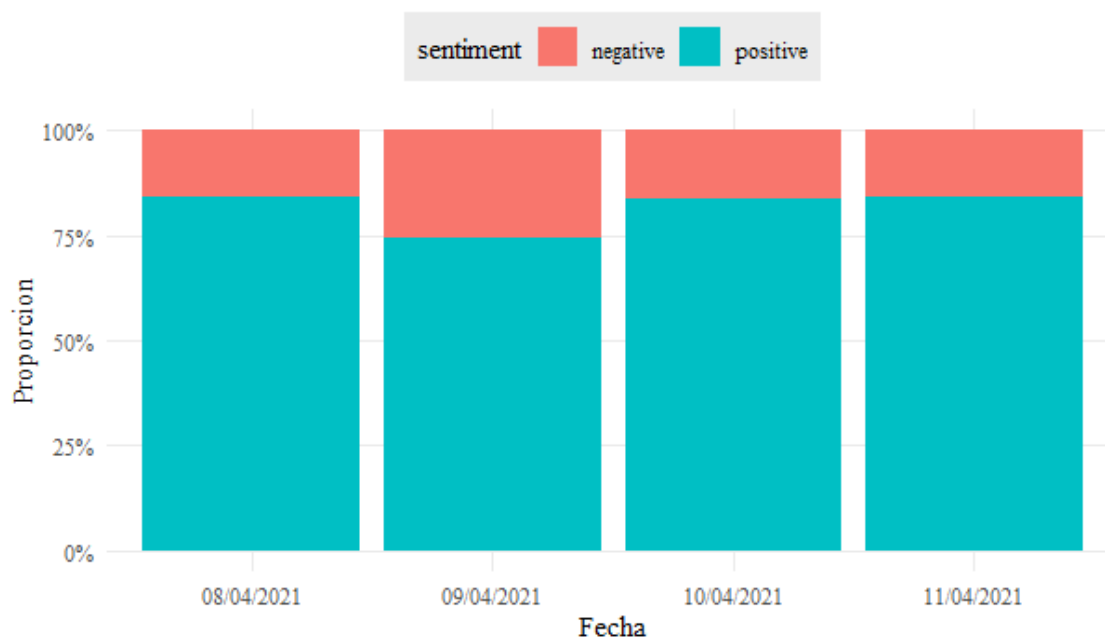
Fuente: Elaboración propia con Rstudio

Como podemos observar, las palabras negativas más comúnmente empleadas han sido "difficult", "useless", "bitter". Vemos también que se han incluido bajo la denominación de palabras negativas términos como "shoot" o "gray" que son palabras que mientras que en otro contexto podrían ser negativas, en el ámbito de la moda no lo son. Entre las palabras positivas encontramos términos como "love", "cool", "innovative".

En último lugar, hemos elaborado el Gráfico 3.4, en el que podremos observar la proporción de positividad y negatividad durante el evento en función de la fecha. Para ello utilizamos la siguiente línea de código:

```
datos1_afinn %>% count(Fecha, Tipo) %>% group_by(Fecha) %>%
mutate(Proporcion = n / sum(n)) %>% ggplot() + aes(Fecha, Proporcion,
fill = Tipo) +geom_col() +scale_y_continuous(labels = percent_format())
+tema_graf +theme(legend.position = "top")
```

Gráfico 3.4. Proporción de sentimiento según la fecha



Fuente: Elaboración propia con Rstudio

Como podemos observar, el sentimiento de positividad es más grande que el de negatividad a lo largo de todo el evento, destacando el segundo día, en el que los usuarios expresaron sentimientos más negativos.

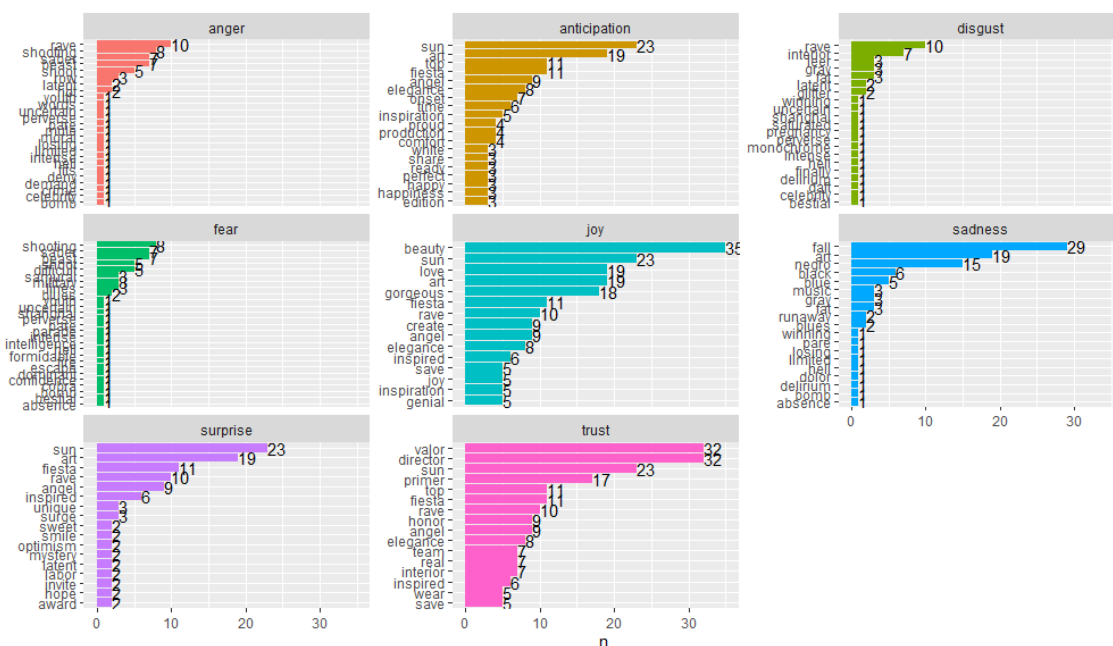
3.3.2 Análisis basado en el léxico *nrc*

Procedemos ahora con el análisis a partir del léxico *nrc*, para lo cual comenzamos asociando a cada palabra su emoción correspondiente:

```
datos_nrc <- datos_token %>% inner_join(get_sentiments("nrc"))
datos_nrc %>% count(word, sentiment, sort = TRUE)
```

En el Gráfico 3.5 podemos observar las palabras más empleadas según la emoción que expresan:

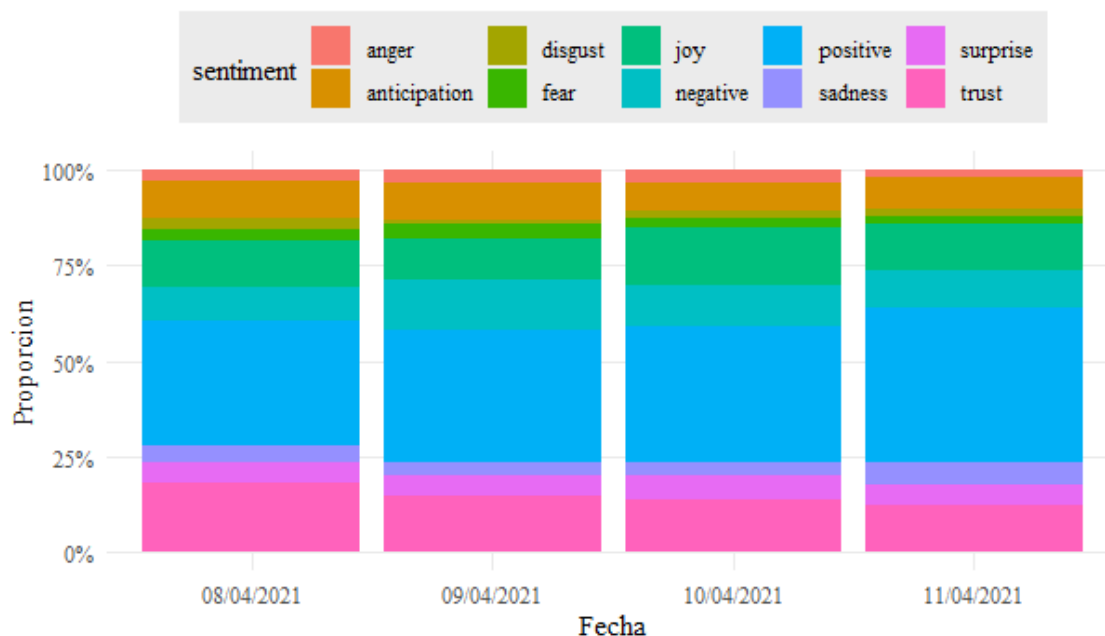
Gráfico 3.5. Palabras más comunes según emoción



Fuente: Elaboración propia con Rstudio

```
datos_nrc %>% filter(sentiment!="negative" & sentiment!="positive") %>%
count(word,sentiment,sort=TRUE) %>%
group_by(sentiment) %>%
top_n(15) %>%
ungroup() %>%
mutate(word=reorder(word,n)) %>%
ggplot(aes(word,n,fill=sentiment))+
geom_col(show.legend = FALSE) + geom_text(aes(label=n), hjust= 0) +
facet_wrap(~sentiment,scales = "free_y")+
coord_flip() + xlab(NULL)
```


Gráfico 3.6. Proporción de emoción según fecha



Fuente: Elaboración propia con Rstudio

3.3.3 Análisis basado en la traducción del léxico *afinn*

A continuación, procedemos con el léxico *afinn*, para el cual disponemos de una traducción al español, de forma que podremos utilizar el paquete *tidytext* para el análisis.

Importamos el léxico y realizamos los cambios necesarios sobre el *dataframe* para poder operar con él más adelante:

```
affin <- read.csv("lexico_afinn.en.es.csv")
colnames(affin)[1]<- "word" colnames(affin)[3] <- "wordenglish"
```

Para poder trabajar sin cometer errores creamos una copia del objeto que contiene los datos tokenizados.

```
datos_token2 <- datos_token
colnames(datos_token2)[2]<-"word"
```

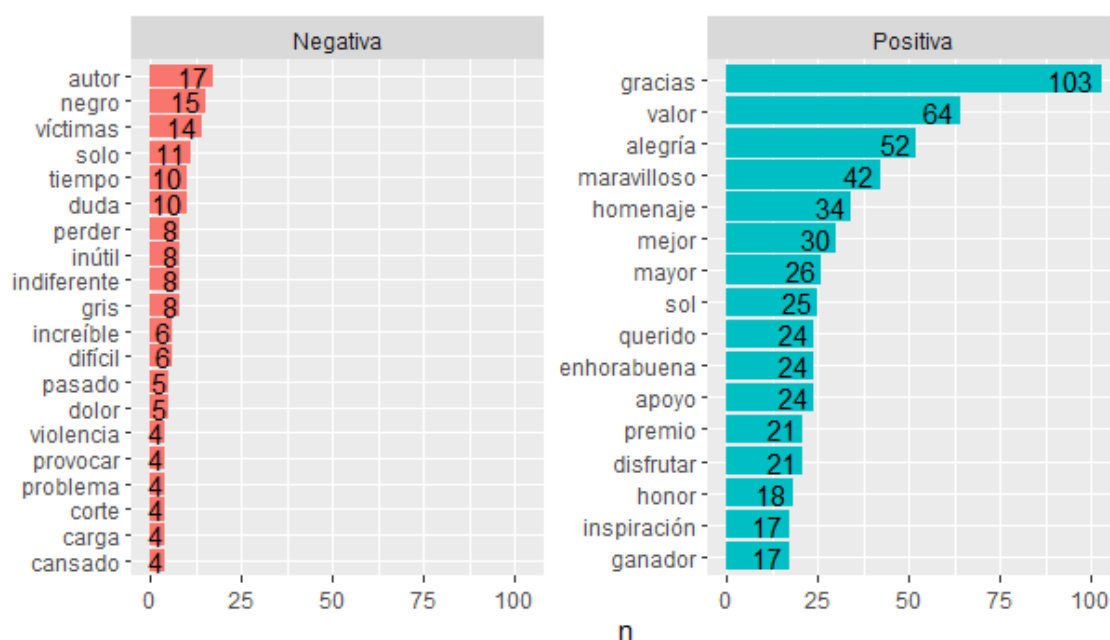
El léxico *afinn* contiene el sentimiento positivo o negativo de cada palabra incluida en la lista, por lo que al cruzarlo con nuestros datos y utilizando la función *mutate* obtendremos el sentimiento de cada palabra:

```
datos_affin <- datos_token2 %>% inner_join(affin, ., by="word") %>%
mutate(Tipo=ifelse (Puntuacion >0, "Positiva", "Negativa"))
```

Una vez hemos asociado a cada palabra su correspondiente sentimiento, elaboramos el Gráfico 3.7 en el que poder visualizar las palabras más utilizadas según su sentimiento:

```
datos_affin %>% count(word, Tipo, sort = TRUE) %>% group_by(Tipo) %>%
top_n(15) %>% ungroup () %>% mutate(word=reorder(word, n)) %>%
ggplot(aes(word, n, fill=Tipo)) + geom_col(show.legend = FALSE) +
geom_text(aes(label=n), hjust=1.2) + facet_wrap(~Tipo, scales =
"free_y")+ coord_flip() + xlab(NULL)
```

Gráfico 3.7. Palabras más utilizadas según sentimiento

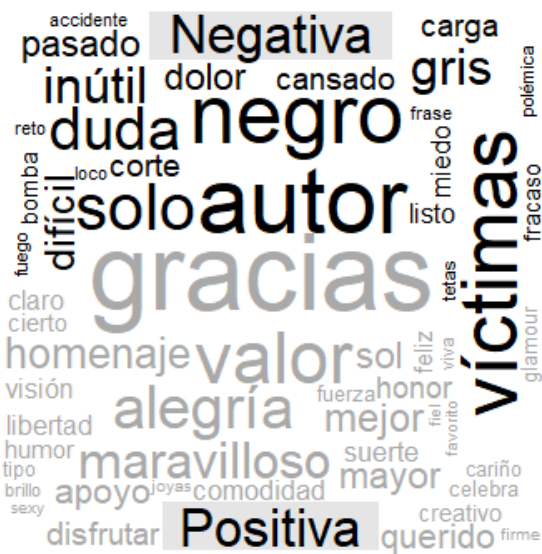


Fuente: Elaboración propia con Rstudio

Ahora, como hicimos anteriormente, vamos a representar esta información en forma de nube (Imagen 3.4), con la siguiente línea:

```
datos_affin %>% count(word,Tipo,sort=TRUE) %>% acast(word~Tipo,value.var
= "n", fill = 0) %>% comparison.cloud(colors = c("black","dark grey"),
max.words = 150, title.size = 2)
```

Imagen 3.4. Nube de palabras según sentimiento



Fuente: Elaboración propia con Rstudio

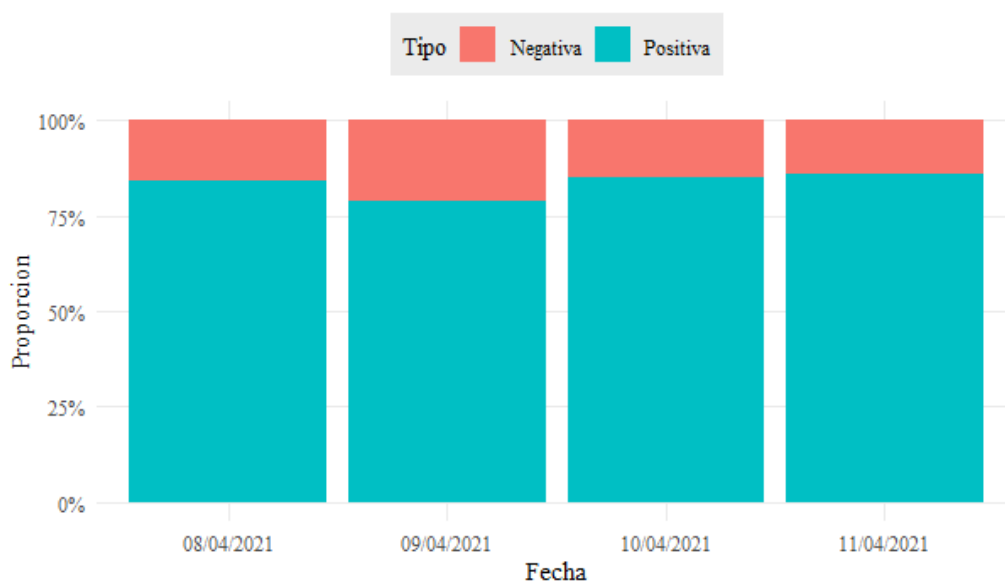
De esta nueva nube de palabras comparativa podemos extraer como palabras negativas términos como "víctimas" pero también "negro", "gris" o "corte", que como ya hemos mencionado anteriormente son términos propios del vocabulario del mundo de la moda y que no pueden ser clasificados como positivos o negativos.

En último lugar, hemos elaborado el Gráfico 3.8 en el que podremos observar la proporción de positividad y negatividad durante el evento en función de la fecha. Para ello usamos la siguiente línea de código:

```
datos_affin %>% count(Fecha, Tipo) %>% group_by(Fecha) %>%
mutate(Proporcion = n / sum(n)) %>% ggplot() + aes(Fecha, Proporcion,
fill = Tipo) +geom_col() +scale_y_continuous(labels = percent_format())
+tema_graf +theme(legend.position = "top")
```

De forma análoga al caso de los datos en inglés, podemos observar sobre este gráfico que el día donde más negatividad expresaron los usuarios fue el segundo día del evento y que, en general, el sentimiento con respecto al evento es positivo.

Gráfico 3.8. Proporción del sentimiento según la fecha



Fuente: Elaboración propia con Rstudio

3.3.4 Análisis basado en la traducción del léxico *nrc*

En este apartado se muestra el análisis con el léxico *nrc*, utilizando igual que hemos hecho con el léxico *afinn*, una traducción de este.

Procedemos de la misma forma que anteriormente, comenzando por importar el léxico y realizando todos los cambios necesarios sobre la tabla para proceder con el análisis:

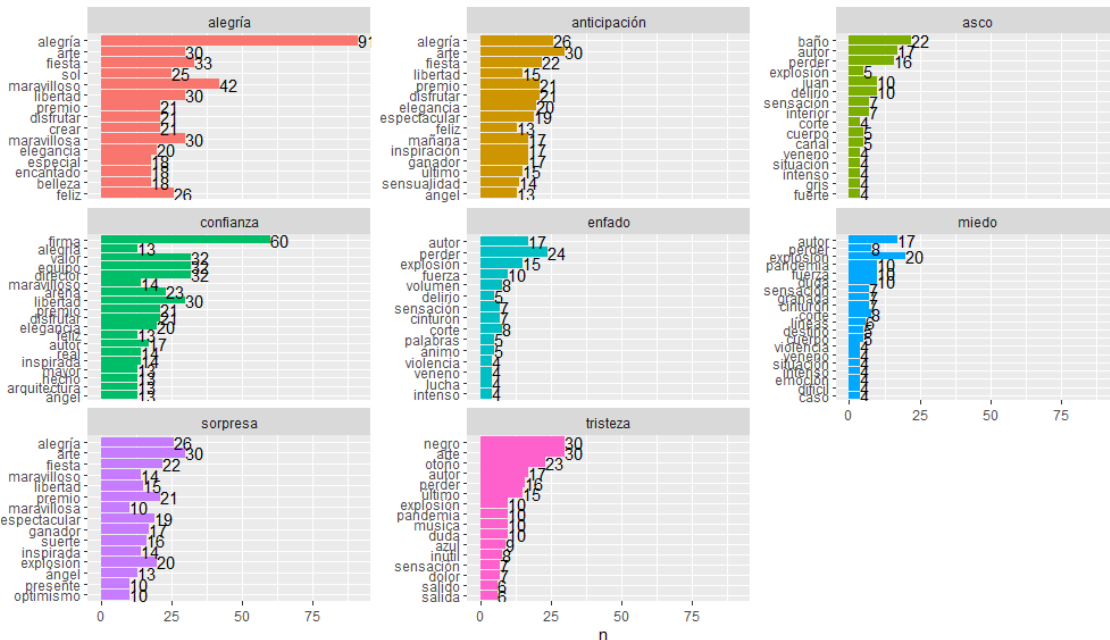
```
nrc <- read.csv("lexico_nrc.csv", encoding = "UTF-8")
colnames(nrc)[1] <- "word" colnames(nrc)[1] <- "englishword"
datos_token3 <- datos_token
datos_nrcspanish <- datos_token3 %>% inner_join(nrc) datos_nrcspanish %>%
count(word, sentimiento, sort = TRUE)
```

En el Gráfico 3.9 podemos observar, en función de su clasificación según la emoción que expresan las palabras, cuales han sido las más repetidas:

```
datos_nrcspanish %>% filter(sentimiento!="negativo" &
sentimiento!="positivo") %>% count(word,sentimiento,sort=TRUE) %>%
group_by(sentimiento) %>% top_n(15) %>%
ungroup() %>%
mutate(word=reorder(word,n)) %>%
ggplot(aes(word,n,fill=sentimiento))+
```

```
geom_col(show.legend = FALSE) + geom_text(aes(label=n), hjust= 0) +
facet_wrap(~sentimiento,scales = "free_y")+
coord_flip() + xlab(NULL)
```

Gráfico 3.9. Palabras más comunes según emoción

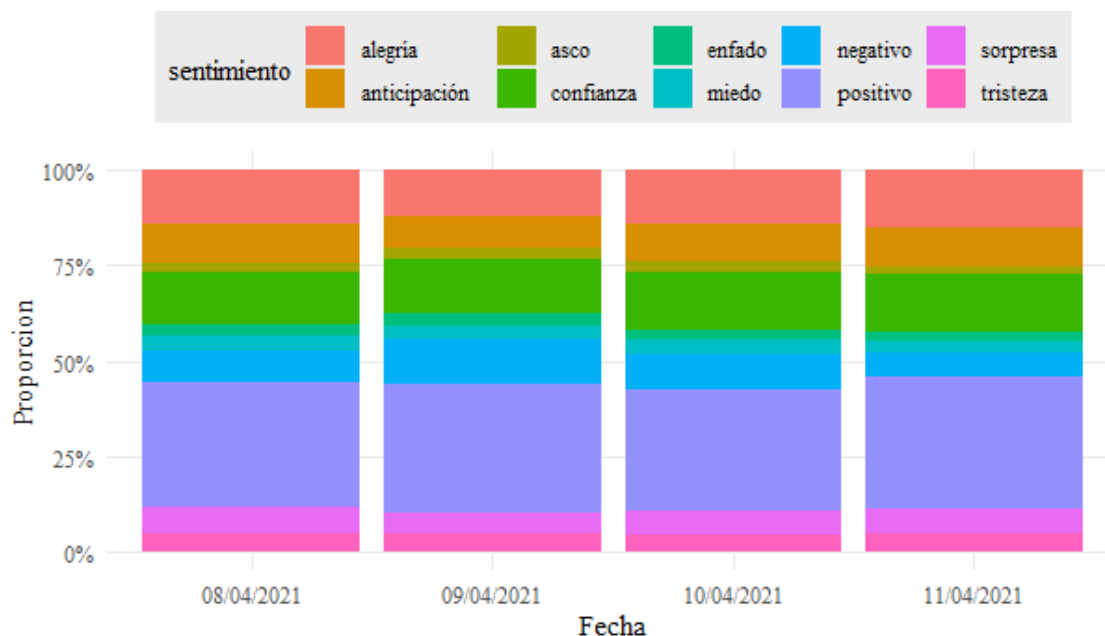


Fuente: Elaboración propia con Rstudio

La representación en forma de nubes de palabras de esta información se muestra en la Imagen 3.5. En esta nube de palabras podemos observar que la palabra más comúnmente empleada ha sido "baño" que, si tenemos en cuenta que la palabra más utilizada a rasgos generales, ha sido "Andrés Sardá", podemos deducir que están relacionadas y, por tanto, concluir que el tema más comentado durante la Fashion Week ha sido la colección de ropa de baño presentada por Andrés Sardá.

```
datos_nrcspanish %>% count(word,sentimiento,sort=TRUE) %>%
filter(sentimiento!="positivo" & sentimiento!="negativo") %>%
acast(word~sentimiento,value.var = "n", fill = 0) %>%
comparison.cloud(title.size = 1.0)
```


Gráfico 3.10. Proporción de emoción según la fecha



Fuente: Elaboración propia con Rstudio

3.4. ANÁLISIS DE SENTIMIENTO A PARTIR DE EMOJIS

Para la realización del análisis de emojis emplearemos paquetes de R que ya hemos usado para el análisis de datos textuales, como son *tidyverse* y *tm*, y a mayores utilizaremos otros paquetes como *rvest* y *unicode*.

Para poder realizar este análisis vamos a utilizar dos funciones previamente prediseñadas que nos permitirán, en primer lugar, extraer los emojis que se encuentran en una cadena y todas sus ocurrencias y, en segundo lugar, una función derivada de la primera que nos permitirá extraer esas ocurrencias en un *dataframe*. Hemos denominado a las funciones *cont_coincidencias* y *emojis_coincidencias*, y se puede encontrar su detalle en el Anexo.

El siguiente paso será disponer de un diccionario de emojis de la web. En este caso lo hemos descargado de la siguiente url: <https://raw.githubusercontent.com/today-is-a-good-day/emojis/master/emojis.csv>, y hemos procedido a importarlo a R. Dado que el diccionario está tanto en inglés como en español, hemos optado por importar a R solo la columna que contiene la descripción de los emojis en español:

```
emojisdi <- read.csv2("emojis.csv") %>% select(description = EN,
r_encoding = ftu8, unicode)
```

A continuación, procedemos a importar la base de datos con la que hemos estado trabajando a lo largo de todo este trabajo, para extraer de ella los emojis utilizados en los posts analizados:

```
datosig <- read.csv("public.csv", sep = ";", encoding = "UTF-8")
```

Convertimos los datos a un formato a partir del cual podamos trabajar con ellos, en concreto cambiamos la codificación a ASCII para poder ver el código de los emojis:

```
datosigclean <- datosig %>% mutate(text = iconv(description, from = "UTF-8", to = "ascii", sub = "byte"))
```

El siguiente paso, y a partir de la siguiente codificación, es obtener los emojis más utilizados en las publicaciones analizadas:

```
rank <- emojis_coincidencias(datosigclean$text, matchto, description) %>%
group_by(description) %>% summarise(n = sum(count, na.rm = TRUE)) %>%
arrange(-n)
```

Al ejecutar la siguiente línea obtendremos una lista con el top 10 de los emojis más utilizados y su frecuencia.

```
head(rank, 10)
```

Observamos de esta forma que los emojis más comúnmente empleados han sido los que se muestran en el Cuadro 3.1.

Cuadro 3.1. Emojis más utilizados

Descripción	Emoji
Corazón rojo	❤️
Chispea	💫
Cuadrado negro (grande)	■
Alto voltaje	⚡
Flecha hacia abajo	⬇️
Flecha hacia derecha	➡️
Avión	✈️
Doble exclamación	!!
Cuadrado negro (mediano-pequeño)	▪️
Símbolo femenino	♀️

El siguiente paso es obtener el sentimiento de los emojis. Para ello, utilizamos el diccionario mencionado anteriormente, ya importado a R, como sigue:

```
url <- "http://kt.ijs.si/data/Emoji_sentiment_ranking/index.html"

emojisent <- url %>% read_html() %>% html_table() %>% data.frame() %>%
select(-Image.twemoji., -Sentiment.bar.c.i..95..)
```

La última línea del código anterior indica que vamos a mantener fuera dos columnas de la tabla, dado que contienen imágenes en lugar de información útil o necesaria para el análisis. Así, cambiamos la codificación de los emojis para que coincida con lo anterior y poder cruzar la tabla de los emojis con *emojisdi*:

```
emojisentclean <- emojisent %>% mutate(unicode = as.u_char(unicode)) %>%
mutate(description = tolower(description)) emojis_merged <-
emojisentclean %>% merge(emojisdiclean, by = "unicode")
```

Con la siguiente línea vamos a obtener un dataframe en el que visualizaremos la puntuación para cada post en función de los emojis empleados en dicho post:

```
sentiments <- emojis_coincidencias(datosigclean$text, new_matchto,
new_description, sentiment) %>% mutate(sentiment = count *
as.numeric(sentiment)) %>% group_by(text) %>% summarise(sentiment_score =
sum(sentiment, na.rm = TRUE))

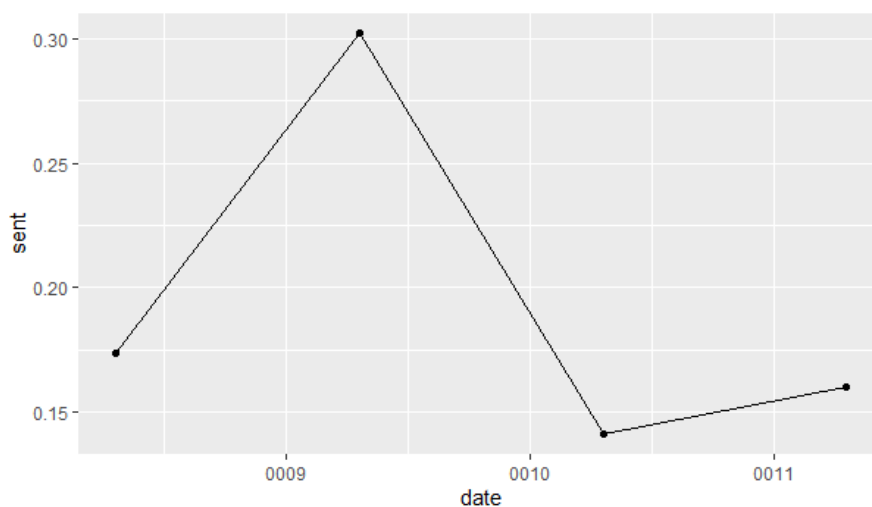
datosig_merged <- datosigclean%>% select(text, pubDate) %>%
merge(sentiments, by = "text", all.x = TRUE)
```

En dicha tabla algunos de los posts no tienen puntuación, lo que se debe a que no contienen ningún emoji. Dado que la tabla es demasiado grande para incluirla entera, hemos extraído solo un pequeño fragmento en el cual se visualizan 15 publicaciones con su puntuación.

Para finalizar, con la siguiente línea creamos el Gráfico 3.11, en el cual podremos observar la evolución con el tiempo del sentimiento expresado por los emojis utilizados en los posts analizados:

```
datosig_merged %>% mutate(date = as.Date(pubDate)) %>% group_by(date) %>%
summarise(sent = mean(sentiment_score, na.rm = TRUE)) %>% ggplot + aes(x
= date, y = sent) + geom_point() + geom_line()
```

Gráfico 3.11. Evolución del sentimiento a lo largo del evento según emojis



Fuente: Elaboración propia con Rstudio

Como podemos observar, desde el comienzo del evento hasta el segundo día la positividad de los sentimientos expresados a través de los emojis se incrementa, pero para el tercer día desciende hasta un nivel inferior al del primer día, siendo este día el día en que menos positividad expresan los emojis, ya que el último día vuelven a incrementar ligeramente. Basándonos solo en los emojis podríamos decir que el día donde el sentimiento ha sido más positivo ha sido el día 9.

Por tanto, y de acuerdo con lo anterior, el sentimiento expresado por el texto y el sentimiento expresado por los emojis se contradice, ya que a través del análisis de datos textuales hemos obtenido que el día con mayor sentimiento negativo ha sido el día nueve y a través del análisis de emojis obtenemos el resultado contrario. Esta variación en los resultados puede deberse a motivos como la subjetividad que representa el concepto de un emoji, ya que cada persona puede interpretarlos de diferente forma y, por tanto, emplearlos para expresar emociones distintas a otros usuarios. Además, a pesar de que el emoji es un elemento cada día más importante a la hora de comunicarnos a través de redes sociales, hemos observado en la base de datos que había un número alto de publicaciones que no incluían emojis, pudiendo resultar esto también en una desviación de los resultados.

CONCLUSIONES

La realización del presente TFG ha permitido extraer una serie de conclusiones a partir de los resultados obtenidos del Análisis de Sentimiento y la información recogida de fuentes bibliográficas.

El mundo de la moda ha experimentado un cambio significativo especialmente en la última década debido a la gran influencia que tienen las redes sociales en la vida de los consumidores. Instagram en concreto se ha convertido en el mayor escaparate para las marcas de moda, reportándoles numerosos beneficios. Además, esta red social ha tenido gran protagonismo en el proceso de democratización que está viviendo toda la industria. Para las marcas más pequeñas, Instagram se ha convertido en la mejor oportunidad para crecer, alcanzar la audiencia objetivo y expandirse geográficamente en nuevos mercados.

Gracias a la forma en la que la sociedad está *atada* a las redes sociales, el Análisis de Sentimiento presenta gran utilidad para las empresas a nivel de marketing, al permitir recabar información clave para el desarrollo de estrategias y con un coste razonable. Además, teniendo en cuenta que la información publicada se encuentra al alcance de todos de forma gratuita y que la herramienta de Rstudio también es gratuita para cualquier usuario, se trata de un análisis con un coste muy bajo para las empresas, que reporta unos resultados que, si son bien utilizados, pueden reportar grandes beneficios económicos para las marcas. Si bien es cierto que para realizar el análisis con Rstudio se requiere de cierta formación especializada, existen también herramientas (aunque la mayoría de pago) que permiten automatizar el proceso aún más y, por tanto, simplificarlo, permitiendo que cualquier persona sin una formación profunda en el área sea capaz de beneficiarse de las ventajas y posibilidades que ofrece este análisis. Aún así, es importante tener en cuenta que el Análisis de Sentimiento presenta una serie de limitaciones a la hora de procesar el lenguaje que aún requerirán de trabajo por parte de los investigadores para superarlas y que es muy importante tenerlas en cuenta al aplicar este tipo de análisis para poder obtener los mejores resultados.

En cuanto al estudio realizado sobre la Mercedes-Benz Fashion Week de Madrid en 2020, las principales conclusiones que hemos podido obtener han sido en primer lugar, que el diseñador sobre el que más se ha comentado ha sido el diseñador de ropa de baño Andrés Sardá. Además, hemos obtenido que el sentimiento global del público

del evento ha sido positivo, por lo que podemos concluir que el desfile y la colección presentada por Andrés Sardá ha sido del agrado del público.

El análisis de emojis, por otra parte, nos ha proporcionado un resultado significativamente diferente. Esto, se ha concluido que se debe a la subjetividad inherente a los *emojis* así como al hecho de que no todos los *posts* analizados incluían *emojis*, lo cual también puede generar diferencias en los resultados. De ahí se concluye que el análisis de *emojis* no debería ser la herramienta única para el Análisis de Sentimiento en datos publicados en redes sociales, sino que debe utilizarse en combinación con el análisis de datos textuales más tradicional como una forma para aportar más información y valor.

Limitaciones

Para llevar a cabo este estudio, las principales limitaciones encontradas han sido:

En primer lugar, la utilización de Rstudio sin ser una experta en su manejo ha supuesto ciertas limitaciones en cuanto al tiempo que ha requerido hacer el análisis y a la *calidad* del código utilizado.

Por otra parte, como se ha explicado en el presente trabajo, el lenguaje es algo complejo de analizar para una máquina. Existen palabras que, al ser utilizadas dentro de un contexto concreto, pueden tener un significado u otro, por lo que sería interesante repetir el estudio utilizando un léxico que incluya términos del mundo de la moda. De esta manera, palabras como “negro” o “corte” que tiene un significado neutro dentro del mundo de la moda pero que en el análisis fueron clasificadas como negativas, habrían sido interpretadas de forma correcta por el programa.

Lecciones aprendidas

Llevar a cabo este trabajo ha supuesto un proceso de aprendizaje interno mediante el cual he descubierto la importancia real de Instagram para el mundo de la moda, que va mucho más allá de lo que hubiera pensado en un principio. Además, he comprendido en mayor profundidad la utilidad del Análisis de Sentimiento y los resultados que genera para las empresas o los departamentos de marketing.

Por otra parte, mis capacidades de manejo de Rstudio han experimentado una significativa mejora gracias a toda la práctica que este trabajo ha requerido.

REFERENCIAS

- (INE), I. N. de E. (2020). *Porcentaje de usuarios de Internet en los últimos tres meses por tipo de actividad realizada. 2020.* https://www.ine.es/jaxi/Datos.htm?path=/t00/mujeres_hombres/tablas_1/10/&file=c04003.px
- AbdelFattah, M., Galal, D., Hassan, N., Elzanfaly, D. S., & Tallent, G. (2017). A sentiment analysis tool for determining the promotional success of fashion images on instagram. *International Journal of Interactive Mobile Technologies*, 11(2), 66–73. <https://doi.org/10.3991/ijim.v11i2.6563>
- Ahmed, O. (2019). How Instagram transformed the fashion industry - i-D. *ID - The Homegrown Issue*, 355. https://i-d.vice.com/en_uk/article/bj9nkz/how-instagram-transformed-the-fashion-industry
- Arciniegas, Y. (2020, July 7). *Empezó la primera Semana de la Moda virtual de París.* France24. <https://www.france24.com/es/20200707-empezó-la-primera-semana-de-la-moda-virtual-de-parís-en-su-historia-por-la-pandemia>
- Baccianella, S., Esuli, A., & Sebastiani, F. (n.d.). *SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining.* Retrieved June 27, 2021, from <http://wordnetcode.princeton>.
- Bel, O. (2019). *Customer journey: qué es y cómo definirlo en tu estrategia.* <https://www.inboundcycle.com/blog-de-inbound-marketing/customer-journey>
- Betancourt, G. A. (2005, April 27). *LAS MÁQUINAS DE SOPORTE VECTORIAL (SVMs).* Scientia Et Technica. https://www.researchgate.net/publication/49588125_LAS_MAQUINAS_DE_SOPORTE_VECTORIAL_SVMs
- Bradley, M. M., & Lang, P. J. (2017). *Center for the Study of Emotion and Attention.* University of Florida. <https://csea.php.ufl.edu/media/anevmessage.html>
- Browchuk, E. (2020, June 29). *Digital Fashion Weeks and Virtual Shows: A Rundown of Fashion's New Rhythm | Vogue.* <https://www.vogue.com/article/digital-fashion-weeks-2020>
- Castro-Martínez, A., Pérez-Ordóñez, C., & Torres-Martín, J. L. (2020). Eventos musicales online durante la crisis de la COVID-19 en España. Análisis de festivales en redes sociales y de sus estrategias de comunicación. *Hipertext.Net*, 21, 41–56. <https://doi.org/10.31009/hipertext.net.2020.i21.04>

- Collado, V. (2020). *Pantalón de mujer: su historia y evolución a lo largo de la historia / Vogue México y Latinoamérica*. <https://www.vogue.mx/moda/articulo/pantalon-de-mujer-su-historia-y-evolucion-en-la-moda-femenina>
- Content Marketing Institute (CMI). (n.d.). *The Complete Guide to Influencer Marketing: Strategies, Templates & Tools*.
- CRAN-r. (n.d.-a). *CRAN - Package tidytext*. Retrieved June 29, 2021, from <https://cran.r-project.org/web/packages/tidytext/index.html>
- CRAN-r. (n.d.-b). *CRAN - Package tidyverse*. Retrieved June 29, 2021, from <https://cran.r-project.org/web/packages/tidyverse/index.html>
- CRAN-r. (n.d.-c). *CRAN - Package wordcloud*. Retrieved June 29, 2021, from <https://cran.r-project.org/web/packages/wordcloud/index.html>
- Criddle, C. (2021, April 3). *How Instagram's influencers changed the model industry - BBC News*. <https://www.bbc.com/news/technology-56592913>
- Delua, J. (2021, March 12). *Supervised vs. Unsupervised Learning: What's the Difference? / IBM*. IBM Analytics. <https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning>
- Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. *WSDM'08 - Proceedings of the 2008 International Conference on Web Search and Data Mining*, 231–239. <https://doi.org/10.1145/1341531.1341561>
- Edwards, M. (2015). *How Social Media Has Changed How We Communicate - Future Of Work*. <https://fowmedia.com/social-media-changed-communicate/>
- Elements of AI. (n.d.). *Clasificación bayesiana ingenua - Elements of AI*. Retrieved June 27, 2021, from <https://course.elementsofai.com/es/3/3>
- Espinosa Grau, A. (2020). La instagramización de los desfiles de moda en tiempos de hipermodernidad: era del directo y última temporada 2020. *Redmarka. Revista de Marketing Aplicado*, 24(2), 110–128. <https://doi.org/10.17979/redma.2020.24.2.7073>
- Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, 57(6), 74–81. <https://doi.org/10.1145/2602574>
- FashionWeekOnline. (n.d.). *When Is Fashion Week? / Fashion Week Online®*. Retrieved July 1, 2021, from <https://fashionweekonline.com/when-is-fashion-week>
- Fernández, D. (2016, October 15). *EL HIPÓDROMO DE LONGCHAMPS. Centro de modas e inicios del Street-Style | VESTUARIO ESCÉNICO*.

- <https://vestuarioescenico.wordpress.com/2016/10/15/el-hipodromo-de-longchamps-centro-de-modas-e-inicios-del-street-style/>
- García del Río, L. (2014, February 16). *La primera New York Fashion Week | Actualidad, Moda | S Moda EL PAÍS*. <https://smoda.elpais.com/moda/la-primera-new-york-fashion-week/>
- Gregorio, M. (2017). *¿Qué es el marketing de influencers?* <https://www.inboundcycle.com/blog-de-inbound-marketing/que-es-marketing-influencers>
- Gumuchian, M.-L. (2021, March 3). *Digital shows here to stay as fashion week gets a makeover | Reuters*. <https://www.reuters.com/business/media-telecom/digital-shows-here-stay-fashion-week-gets-makeover-2021-03-03/>
- Hootsuite. (2021, January). *Digital Trends 2021*. <https://www.hootsuite.com/es/pages/digital-trends-2021>
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. <https://doi.org/10.1145/1014052.1014073>
- IAB Spain. (2019). *VII Observatorio de marcas en redes sociales*. https://iabspain.es/wp-content/uploads/2019/06/observatorio_rrss_2019_vreducidafinal.pdf
- IFEMA. (2020). *¿Qué es Mercedes Benz Fashion Week Madrid? | MBFWMADRID*. https://www.ifema.es/mbfw-madrid/que-es?_gl=1*12k296a*_up*MQ..*_ga*MjA0NjM3NzE0MC4xNjI0OTgzNjMw*_ga_Y63TW790MY*MTYyNTE1NDYxNS4xLjAuMTYyNTE1NDYxNS42MA..
- Instagram. (2016, September 8). *Una pasarela para darse a conocer: a la moda con Instagram | Blog de Instagram para empresas | Instagram for Business*. https://business.instagram.com/blog/fashion-week-on-instagram-2016?locale=es_ES
- Instagram. (2021). *Instagram | Información | Sitio oficial*. <https://about.instagram.com/es-la>
- Intelligent.es. (2017, July 19). *Análisis de sentimiento, ¿qué es, cómo funciona y para qué sirve?* <https://itelligent.es/es/analisis-de-sentimiento/>
- Jiménez Zafra, S. M., Martínez Cámara, E., Martín Valdivia, M. T., & Ureña López, L. A. (2014). Desafíos del Análisis de Sentimientos. *V Jornadas TIMM*, 16. https://www.researchgate.net/publication/266673842_Desafios_del_Analisis_de_Sentimientos

- Lee, D., Jo, J.-C., & Lim, H.-S. (2017). User Sentiment Analysis on Amazon Fashion Product Review Using Word Embedding. *Journal of the Korea Convergence Society*, 8(4), 1–8. <https://doi.org/10.15207/jkcs.2017.8.4.001>
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–184. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Mäntylä, M. V., Graziotin, D., & Kuuttila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. In *Computer Science Review* (Vol. 27, pp. 19–21). Elsevier Ireland Ltd. <https://doi.org/10.1016/j.cosrev.2017.10.002>
- MonkeyLearn. (2020). *Sentiment Analysis – The Go-To Guide*. Monkeylearn.Com. <https://monkeylearn.com/sentiment-analysis/>
- National Research Council Canada. (2019, June 3). *Sentiment and emotion lexicons*. Canada.Ca. <https://nrc.canada.ca/en/research-development/products-services/technical-advisory-services/sentiment-emotion-lexicons>
- Nielsen, F. Å. (2015). *AFINN: A new word list for sentiment analysis on Twitter* « Finn Årup Nielsen's blog. <https://finnaarupnielsen.wordpress.com/2011/03/16/afinn-a-new-word-list-for-sentiment-analysis/>
- O'Flaherty, M. C. (2021, June 18). The Secret Psychology of Sneaker Colors - The New York Times. *New York Times*. <https://www.nytimes.com/2021/05/24/style/sneaker-color-psychology.html>
- Ochieng, S. B. O., Loki, M., & Sambuli, N. (2016). LIMITATIONS OF SENTIMENT ANALYSIS ON FACEBOOK DATA. *International Journal of Social Sciences and Information Technology*, 2(4). https://www.researchgate.net/publication/304024274_LIMITATIONS_OF_SENTIMENT_ANALYSIS_ON_FACEBOOK_DATA
- Pauli, P. A. (2019). *Análisis de sentimiento*. Instituto Tecnológico de Buenos Aires.
- Poncelin, C. (2021). *Harness the power of social media to anticipate fashion trends*. Heuritech.Com. <https://www.heuritech.com/blog/articles/trends-on-social-media/>
- Pous, T. (2013, February 6). *The Democratization of Fashion: A Brief History* / *TIME.com*. <https://style.time.com/2013/02/06/the-democratization-of-fashion-a-brief-history/>
- Rasool, A., Tao, R., Marjan, K., & Naveed, T. (2019). Twitter Sentiment Analysis: A Case Study for Apparel Brands. *Journal of Physics: Conference Series*, 1176(2).

- <https://doi.org/10.1088/1742-6596/1176/2/022015>
- Real Academia Española (RAE). (2020). *moda* | Definición | Diccionario de la lengua española | RAE - ASALE. <https://dle.rae.es/moda>
- Reyes-Menendez, A., Saura, J. R., & Filipe, F. (2020). Marketing challenges in the #MeToo era: gaining business insights using an exploratory sentiment analysis. *Heliyon*, 6(3). <https://doi.org/10.1016/j.heliyon.2020.e03626>
- Rstudio. (n.d.). *RStudio* | Open source & professional software for data science teams - RStudio. Retrieved June 29, 2021, from <https://www.rstudio.com/>
- Santana, A., & Hernández, C. N. (2016). *Librerías en R*. <https://estadistica-dma.ulpgc.es/cursoR4ULPGC/5-librerias.html>
- SAS Institute. (n.d.). *Aprendizaje automático: Qué es y por qué es importante* | SAS. Retrieved June 25, 2021, from https://www.sas.com/es_mx/insights/analytics/machine-learning.html
- Saura, J., Reyes-Menendez, A., & Palos-Sanchez, P. (2018). Un Análisis de Sentimiento en Twitter con Machine Learning: Identificando el sentimiento sobre las ofertas de #BlackFriday. *Espacios*, 39(2010), 16. <http://www.revistaespacios.com/a18v39n42/a18v39n42p16.pdf>
- Setuserv. (n.d.). *Role of Sentiment Analysis in Market Research* - Setuserv. Retrieved June 27, 2021, from <https://www.setuserv.com/role-of-sentiment-analysis-in-market-research/>
- Silge, J., & Robinson, D. (2021). *Introduction to tidytext*. <https://cran.r-project.org/web/packages/tidytext/vignettes/tidytext.html>
- Statista. (2020, October 30). • *Instagram: gasto mundial en marketing de influencers 2013-2020* | Statista. <https://es.statista.com/estadisticas/1038369/gasto-mundial-en-marketing-de-influencers-en-instagram/>
- Statista. (2021a). *¿Qué redes sociales usamos más?* <https://es.statista.com/grafico/24720/personas-que-usan-regularmente-las-siguientes-plataformas-de-redes-sociales/>
- Statista. (2021b). *Tiempo medio diario empleado en redes sociales a nivel mundial 2012-2020*. <https://es.statista.com/estadisticas/513084/cantidad-tiempo-uso-diario-redes-sociales/>
- Subramanian, K. R. (2017). Influence of Social Media in Interpersonal Communication Separate project View project SUSTAINING GLOBALIZATION IN AN ENVIRONMENT OF RISING NATIONALISM View project Influence of Social

- Media in Interpersonal Communication Senior Consultant & Professor Of M. *Issue*, 109(02), 71–72. www.ijsspr.com
- Taroy, D. (2015, October 11). *How Instagram is Democratizing Fashion*. <https://www.fastcompany.com/3053382/how-instagram-is-democratizing-fashion>
- Territorio Digital. (n.d.). *Customer Journey*. https://territorioidigital.net/wp-content/uploads/2018/12/customer_experience_touchpoints_territorio-digital.png
- Valrie Moatti, & Cline Abecassis-Moedas. (2018, June 24). *How Instagram became the natural showcase for the fashion world | The Independent | The Independent*. <https://www.independent.co.uk/life-style/fashion/features/instagram-fashion-industry-digital-technology-a8412156.html>
- Vasista Reddy. (2018, November 12). *Sentiment Analysis using SVM. Sentiment Analysis is the NLP technique... | by Vasista Reddy | Medium*. <https://medium.com/@vasista/sentiment-analysis-using-svm-338d418e3ff1>
- Walaa Medhat, Ahmed Hassan, H. K. (2014). *Sentiment classification techniques* (p. 1095). <https://reader.elsevier.com/reader/sd/pii/S2090447914000550?token=3194252F6A5D09BE8CF30634112992D46CE061626FAD156B0D41C5CF90BAB4CD6C52F3FA3541D47C9757A5AB19065692&originRegion=eu-west-1&originCreation=20210625115947>
- Webb, J. (2011, March 31). *With sentiment analysis, context always matters - Matthew Russel on the limitations and applications of sentiment analysis*. <http://radar.oreilly.com/2011/03/sentiment-analysis-context.html>
- ZARA. (2021). *WhatsApp Image 2019-04-01 at 15*. https://www.instagram.com/p/CPIREfmBouW/?utm_medium=copy_link

ANEXO I: CÓDIGO PARA ANÁLISIS DE SENTIMIENTO

#ANÁLISIS DE SENTIMIENTO#

```
setwd("C:/Users/mirei/Desktop")

#Importamos los paquetes#

library(tidytext)
library(dplyr)
library(tidyverse)
library(wordcloud)
library(reshape2)
library(RColorBrewer)
library(lubridate)
library(zoo)
library(scales)

#Importamos los posts#

datos <- read.csv("public.csv", sep = ";", encoding = "UTF-8")

#Limpiamos los datos#

datossplit <- data.frame(datos$pubDate)
datossplitdesc <- data.frame(datos$description)
cleandescrip <- map_df(datossplitdesc, ~gsub("[:punct:]", "", .x))
cleandescrip <- map_df(cleandescrip, ~gsub("[0-9]", "", .x))
cleandescrip <- map_df(cleandescrip, ~gsub("https", "", .x))
datosclean <- data.frame(datossplit, cleandescrip)

datos_token <- unnest_tokens(tbl = datosclean, output = "word", input =
"description", token= "words", to_lower = TRUE)

datos_token <- datos_token %>% separate(datos.pubDate, into = c("Fecha",
"Hora"), sep = " ")

#Ordenamos las palabras por frecuencia#

count(datos_token, word, sort = TRUE)

#Tratamiento de stopwords#

install.packages("quanteda")
install.packages("tm")
library(quanteda)
library(tm)
```

```

mystopwords <- bind_rows(stop_words, data_frame(word =
tm::stopwords("spanish"), lexicon = "custom"))

datos_token <- anti_join(x=datos_token, y= mystopwords, by="word")

#Comenzamos con el análisis#

datos_token %>% count(word, sort = TRUE) %>% filter(n > 70) %>% mutate(word =
reorder(word, n)) %>% ggplot(aes(word, n)) + geom_text(aes(label=n), hjust= -
0.2) + geom_col() + xlab(NULL) + coord_flip()+ theme_classic()

#Gráfico eliminando palabras poco relevantes#

datos_token <- filter(datos_token, word!="mbfwmadrid" & word!="moda" &
word!="fashion" & word!="madrid" & word!="colección" & word!="desfile" &
word!="pasarela" & word!="fashionweek" & word!="mbfw" & word!="collection" &
word!="edición" & word!="mbfwm" & word!="2021" & word!="desfiles" &
word!="week" & word!="2021" & word!="ifema" & word!="mercedes" &word!="benz")

#Wordcloud#

datos_token %>% count(word) %>% with(wordcloud(words=word, freq=n, max.words =
250, scale = c(3,1), rot.per = 0.3, random.order = FALSE,
colors=brewer.pal(6,"Dark2"))

#Obtenemos los léxicos#

get_sentiments("nrc")
get_sentiments("afinn")

#Análisis con léxico afinn#

datos1_afinn <-datos_token %>% inner_join(get_sentiments("afinn"))

datos1_afinn %>% count(word,value,sort=TRUE)

datos1_afinn <- datos1_afinn %>% mutate(Tipo=ifelse (value >0, "Positiva",
"Negativa"))

datos1_afinn %>% count(word, Tipo, sort = TRUE) %>% group_by(Tipo) %>%
top_n(15) %>% ungroup ()%>%mutate(word=reorder(word, n)) %>% ggplot(aes(word,
n, fill=Tipo)) + geom_col(show.legend = FALSE) + geom_text(aes(label=n),
hjust=1.2) + facet_wrap(~Tipo, scales = "free_y")+ coord_flip() + xlab(NULL)

datos1_afinn %>% count(word,Tipo,sort=TRUE) %>% acast(word~Tipo,value.var =
"n", fill = 0) %>% comparison.cloud(colors = c("black","dark grey"), max.words
= 150, title.size = 2)

```

```

datos1_afinn %>% count(Fecha, Tipo) %>% group_by(Fecha) %>% mutate(Proporcion
= n / sum(n)) %>% ggplot() + aes(Fecha, Proporcion, fill = Tipo) +geom_col()
+scale_y_continuous(labels = percent_format()) +tema_graf
+theme(legend.position = "top")

```

#Análisis con léxico nrc#

```
datos_nrc <- datos_token %>% inner_join(get_sentiments("nrc"))
```

```
datos_nrc %>% count(word, sentiment, sort = TRUE)
```

```

datos_nrc %>% filter(sentiment!="negative" & sentiment!="positive") %>%
count(word,sentiment,sort=TRUE) %>% group_by(sentiment) %>% top_n(15) %>%
ungroup() %>% mutate(word=reorder(word,n)) %>%
ggplot(aes(word,n,fill=sentiment))+ geom_col(show.legend = FALSE) +
geom_text(aes(label=n), hjust= 0) + facet_wrap(~sentiment,scales = "free_y")+
coord_flip() + xlab(NULL)

```

```

datos_nrc %>% count(word,sentiment,sort=TRUE) %>% filter(sentiment!="positive"
& sentiment!="negative") %>% acast(word~sentiment,value.var = "n", fill = 0)
%>% comparison.cloud(title.size = 1.0)

```

```

datos_nrc %>% count(Fecha, sentiment) %>% group_by(Fecha) %>%
mutate(Proporcion = n / sum(n)) %>% ggplot() + aes(Fecha, Proporcion, fill =
sentiment) +geom_col() +scale_y_continuous(labels = percent_format())
+tema_graf +theme(legend.position = "top")

```

#Análisis con traducción del léxico afinn#

```
affin <- read.csv("lexico_afinn.en.es.csv")
```

```
colnames(affin)[1]<- "word" colnames(affin)[3] <- "wordenglish"
```

```
datos_token2 <- datos_token
```

```
colnames(datos_token2)[2]<-"word"
```

```

datos_affin <- datos_token2 %>% inner_join(affin, ., by="word") %>%
mutate(Tipo=ifelse (Puntuacion >0, "Positiva", "Negativa"))

```

```

datos_affin %>% count(word, Tipo, sort = TRUE) %>% group_by(Tipo) %>%
top_n(15) %>% ungroup ()%>%mutate(word=reorder(word, n)) %>% ggplot(aes(word,
n, fill=Tipo)) + geom_col(show.legend = FALSE) + geom_text(aes(label=n),
hjust=1.2) + facet_wrap(~Tipo, scales = "free_y")+ coord_flip() + xlab(NULL)

```



```
datos_affin %>% count(word,Tipo,sort=TRUE) %>% acast(word~Tipo,value.var =
"n", fill = 0) %>% comparison.cloud(colors = c("black","dark grey"), max.words
= 150, title.size = 2)
```

```
datos_affin %>% count(Fecha, Tipo) %>% group_by(Fecha) %>% mutate(Proporcion =
n / sum(n)) %>% ggplot() + aes(Fecha, Proporcion, fill = Tipo) +geom_col()
+scale_y_continuous(labels = percent_format()) +tema_graf
+theme(legend.position = "top")
```

#Análisis con traducción del léxico nrc#

```
nrc <- read.csv("lexico_nrc.csv", encoding = "UTF-8")
```

```
colnames(nrc)[1] <- "word" colnames(nrc)[1] <- "englishword"
```

```
datos_token3 <- datos_token
```

```
datos_nrcspanish <- datos_token3 %>% inner_join(nrc) datos_nrcspanish %>%
count(word, sentimiento, sort = TRUE)
```

```
datos_nrcspanish %>% filter(sentimiento!="negativo" & sentimiento!="positivo")
%>% count(word,sentimiento,sort=TRUE) %>% group_by(sentimiento) %>% top_n(15)
%>% ungroup() %>% mutate(word=reorder(word,n)) %>%
ggplot(aes(word,n,fill=sentimiento))+ geom_col(show.legend = FALSE) +
geom_text(aes(label=n), hjust= 0) + facet_wrap(~sentimiento,scales = free_y")+
coord_flip() + xlab(NULL)
```

```
datos_nrcspanish %>% count(word,sentimiento,sort=TRUE) %>%
filter(sentimiento!="positivo" & sentimiento!="negativo") %>%
acast(word~sentimiento,value.var = "n", fill = 0) %>%
comparison.cloud(title.size = 1.0)
```

```
datos_nrcspanish %>% count(Fecha, sentimiento) %>% group_by(Fecha) %>%
mutate(Proporcion = n / sum(n)) %>% ggplot() + aes(Fecha, Proporcion, fill =
sentimiento) +geom_col() +scale_y_continuous(labels = percent_format())
+tema_graf +theme(legend.position = "top")
```

```
#ANÁLISIS DE EMOJIS#

options(stringsAsFactors = FALSE)

#Importamos los paquetes#

library(tidyverse)

library(rvest)

library(Unicode)

library(tm)

#Creamos las funciones prediseñadas#

cont_coincidencias <- function(string, matchto, description, sentiment = NA)
{
  vec <- str_count(string, matchto)

  matches <- which(vec != 0)

  descr <- NA

  cnt <- NA

  if (length(matches) != 0) {
    descr <- description[matches]

    cnt <- vec[matches]

    df <- data.frame(text = string, description = descr, count = cnt, sentiment
= NA)

    if (!is.na(sentiment) && length(sentiment[matches]) != 0) {

      df$sentiment <- sentiment[matches]

    }

    return(df)
  }

}

emojis_coincidencias <- function(texts, matchto, description, sentiment = NA)
{

  texts %>%

    map_df(cont_coincidencias,

           matchto = matchto,

           description = description,

           sentiment = sentiment)}

}
```

```

#Comenzamos con el análisis#
emojisdi <- read.csv2("emojis.csv") %>% select(description = EN, r_encoding =
ftu8, unicode)

emojisdiclean <- emojisdi %>% mutate(description = tolower(description))
%>%mutate(unicode = as.u_char(unicode))

matchto <- emojisdiclean$r_encoding

description <- emojisdiclean$description

datosig <- read.csv("public.csv", sep = ";", encoding = "UTF-8")

datosigclean <- datosig %>% mutate(text = iconv(description, from = "UTF-8",
to = "ascii", sub = "byte"))

rank <- emojis_coincidencias(datosigclean$text, matchto, description)
%>% group_by(description) %>% summarise(n = sum(count, na.rm = TRUE))
%>%arrange(-n)

head(rank, 10)

# A tibble: 10 x 2
  description          n
  <chr>                <int>
1 red heart            103
2 sparkles             72
3 black small square  64
4 high voltage        36
5 down arrow          34
6 right arrow         34
7 airplane            21
8 double exclamation mark 16
9 black medium-small square 12
10 female sign        8

url <- "http://kt.ijs.si/data/Emoji_sentiment_ranking/index.html"

emojisent <- url %>% read_html() %>% html_table() %>% data.frame() %> select(-
Image.twemoji., -Sentiment.bar.c.i..95..)

names(emojisent) <- c("char","unicode","occurrences","position",
"negative", "neutral", "positive", "sentiment_score", "description", "block")

```

```
emojisentclean <- emojisent %>% mutate(unicode = as.u_char(unicode))
%>% mutate(description = tolower(description))

emojis_merged <- emojisentclean %>% merge(emojisdiclean, by = "unicode")

new_matchto <- emojis_merged$r_encoding

new_description <- emojis_merged$description.x

sentiment <- emojis_merged$sentiment_score

sentiments <- emojis_coincidencias(datosigclean$text, new_matchto,
new_description, sentiment) %>% mutate(sentiment = count *
as.numeric(sentiment)) %>% group_by(text) %>% summarise(sentiment_score =
sum(sentiment, na.rm = TRUE))

datosig_merged <- datosigclean%>% select(text, pubDate) %>% merge(sentiments,
by = "text", all.x = TRUE)
```