



Facultad de Ciencias Económicas y Empresariales
Universidad de León

Grado en Marketing e Investigación de Mercados

Curso 2021/2022

EXPERIENCIA TURÍSTICA MEMORABLE: ANÁLISIS DE
SENTIMIENTO SOBRE EL RECURSO TURÍSTICO ICÓNICO DE LA
CIUDAD DE LEÓN

MEMORABLE TOURIST EXPERIENCE: SENTIMENT ANALYSIS OF
THE ICONIC TOURIST RESOURCE OF THE CITY OF LEÓN

Realizado por la Alumna Dña. Elena Díaz Martín

Tutelado por la Profesora Dña. Ana María González Fernández

Cotutelado por la Profesora Dña. Sofía Blanco Moreno

León, marzo de 2022

MODALIDAD DE DEFENSA PÚBLICA:

Tribunal

Póster

ÍNDICE

RESUMEN	5
ABSTRACT	6
1. INTRODUCCIÓN	7
2. METODOLOGÍA	10
2.1. FUENTES DE INFORMACIÓN SECUNDARIA	10
2.2. FUENTES DE INFORMACIÓN PRIMARIA	10
CAPÍTULO I: LA EXPERIENCIA TURÍSTICA MEMORABLE Y LA IMPORTANCIA DEL USO DE INSTAGRAM EN EL TURISMO	12
1. LA EXPERIENCIA TURÍSTICA	12
2. EXPERIENCIA TURÍSTICA MEMORABLE (MTE)	13
2.1. Concepto	13
2.2. Efecto de las MTE sobre el comportamiento de los turistas	14
2.3. Dimensiones de las MTE	14
2.3.1. Hedonismo	15
2.3.2. Revitalización.....	15
2.3.3. Interacción social y cultura local.....	15
2.3.4. Significación.....	15
2.3.5. Conocimiento	16
2.3.6. Implicación.....	16
2.3.7. Novedad	16
2.4. Atributos del destino turístico	16
3. RELEVANCIA DE INSTAGRAM EN EL TURISMO	17
3.1. Crecimiento de las redes sociales en nuestra sociedad	17
3.2. Instagram como medio de comunicación	20
3.3. Importancia de Instagram para el turismo	22
CAPÍTULO II: ANÁLISIS DE COMENTARIOS MEDIANTE EL ANÁLISIS DE SENTIMIENTO	24
1. CONTENIDO GENERADO POR EL USUARIO	24
2. ANÁLISIS DE SENTIMIENTO	24
2.1. Concepto	24
2.2. Los tres niveles de aplicación en el Análisis de Sentimiento	26

2.3. Enfoques del Análisis de Sentimiento	27
2.4. Análisis de Sentimiento en la industria turística	28
CAPÍTULO III: ANÁLISIS DE SENTIMIENTO EN RSTUDIO SOBRE LA CIUDAD DE LEÓN	31
1. LEÓN COMO DESTINO DE TURISMO CULTURAL Y SMART DESTINATION	31
2. PROGRAMA ESTADÍSTICO RSTUDIO	33
3. METODOLOGÍA DEL ESTUDIO	34
3.1. Recogida de datos	34
3.2. Tratamiento de datos	34
4. ANÁLISIS ESTADÍSTICO EN RSTUDIO	35
4.1. Análisis de frecuencias de términos	35
4.2. Análisis de Sentimiento	43
5. ANÁLISIS ESTADÍSTICO EN RSTUDIO (MTE)	52
5.1. Análisis de frecuencias de términos (MTE)	52
5.2. Análisis de Sentimiento (MTE)	57
CONCLUSIONES	67
REFERENCIAS	70
ANEXOS	78
ANEXO I: CÓDIGO ANÁLISIS	78
ANEXO II: LOCALIZACIONES DE INSTAGRAM	82

ÍNDICE DE FIGURAS

Figura 1. Objetivos del estudio	9
Figura 2. Metodología primaria empleada	11
Figura 1.1. Historia de las redes sociales: cuándo nacen.....	18
Figura 2.1. Proceso de Análisis de Sentimiento	26
Figura 2.2. Enfoques del Análisis de Sentimiento	28
Figura 3.1. Importación base de datos a RStudio	35
Figura 3.2. Nube de palabras.....	38
Figura 3.3. Nube de palabras (MTE)	53

ÍNDICE DE GRÁFICOS

Gráfico 1.1. Penetración de las redes sociales en España de 2010 a 2020.....	19
Gráfico 1.2. Redes sociales con mayor porcentaje de usuarios en España en 2020	21
Gráfico 3.1. Frecuencia de los términos	42
Gráfico 3.2. Análisis de la polaridad de sentimientos	46
Gráfico 3.3. Análisis de Sentimiento	47
Gráfico 3.4. Clasificación de palabras según el sentimiento	49
Gráfico 3.5. Clasificación de palabras según las emociones	51
Gráfico 3.6. Frecuencia de los términos (MTE)	56
Gráfico 3.7. Análisis de la polaridad de sentimientos (MTE)	60
Gráfico 3.8. Análisis de Sentimiento (MTE)	60
Gráfico 3.9. Clasificación de palabras según el sentimiento (MTE)	61
Gráfico 3.10. Clasificación de palabras según las emociones (MTE)	63

ÍNDICE DE TABLAS

Tabla 2.2. Ejemplos de estudios de Análisis de Sentimiento en turismo	30
Tabla 3.1. Matriz término-documento (TDM).....	39
Tabla 3.2. Matriz de frecuencias por término (TFM)	41
Tabla 3.3. Clasificación de comentarios según las emociones	44
Tabla 3.4. Suma total de emociones del texto	45
Tabla 3.5. Matriz término-documento (MTE)	54
Tabla 3.6. Matriz de frecuencias por término (MTE)	55
Tabla 3.7. Clasificación de comentarios según las emociones (MTE).....	58
Tabla 3.8. Suma total de emociones del texto (MTE)	59
Tabla 3.9. Comparativa entre experiencias ordinarias y memorables	65

RESUMEN

El turismo representa un eje fundamental para la economía del país, por ello las empresas de la industria deben adaptarse a las nuevas necesidades de los turistas, aumentando su satisfacción y promoviendo sus emociones. Aquí nace la experiencia turística memorable.

Las redes sociales son aquellos lugares donde los turistas comparten dichas experiencias con la audiencia, especialmente en Instagram, por lo que es interesante indagar en ellas para poder descubrirlo.

Por ambas razones, en este estudio se ha realizado un Análisis de Sentimiento en Instagram para conocer de primera mano qué emociones son las predominantes en los viajeros, determinar la existencia o no de experiencias memorables y cómo, gracias a su recuerdo, se nutren de ellas. El análisis tendrá lugar sobre la ciudad de León, concretamente sobre la catedral.

Con este análisis, se concluye la existencia de experiencias turísticas memorables observándose una alta memorabilidad y satisfacción. Por otra parte, se afirma la presencia de sentimientos con polaridad positiva sobre el mencionado monumento, destacando las emociones que expresan alegría y confianza. Así mismo, aquellos visitantes que han experimentado experiencias memorables tras la visita han obtenido las mismas sensaciones.

Palabras clave: Experiencia turística memorable, Análisis de Sentimiento, turismo, redes sociales, Instagram, emociones.

ABSTRACT

Tourism represents a fundamental axis for the country's economy, which is why companies in the industry must adapt to the new needs of tourists, increasing their satisfaction and promoting their emotions. This is where the memorable tourist experience is born.

Social networks are those places where tourists share these experiences with the audience, especially on Instagram, so it is interesting to dig into them in order to discover it.

For both reasons, in this study a Sentiment Analysis has been carried out on Instagram to find out first-hand which emotions are predominant in travellers, to determine the existence or not of memorable experiences and how, thanks to their memory, they are nourished by them. The analysis will take place on the city of León, specifically on the cathedral.

With this analysis, the existence of memorable tourist experiences is concluded, observing a high memorability and satisfaction. On the other hand, the present of feelings with positive polarity about the monument is affirmed, highlighting the emotions that express joy and trust. Likewise, those visitors who have experienced memorable experiences after the visit have obtained the same feelings.

Keywords: Memorable tourist experience, Sentiment Analysis, tourism, social networks, Instagram, emotions.

1. INTRODUCCIÓN

El turismo en España es uno de los principales motores que sustenta la economía, tal y como avalan los datos que nos dejan algunos informes. Según recoge Dataestur (2021), en su estudio “Movimientos turísticos en fronteras”, el número de visitantes procedentes de otros países del mundo que visitaron España asciende hasta los 5.070.105. Por otra parte, en la “Encuesta de turismo de residentes”, Dataestur (2021a) muestra el número de viajes realizados por los residentes de España dentro de nuestras fronteras, el cual alcanza la cifra de 106.549.208. No obstante, estos datos son inferiores al resto de estudios realizados en años atrás, ya que se han visto altamente influenciados por la pandemia provocada a causa del virus COVID-19 y sus severas restricciones. A pesar de esto, el turismo es considerado un eje muy importante.

Dada la importancia de la industria turística, las empresas pertenecientes al sector deben cambiar la manera de ofrecer sus servicios turísticos, cambiando el tradicional turismo de masas por viajes que reporten auténticas experiencias de turismo completamente satisfactorias para el viajero (Fusté Forné y Nava Jiménez, 2015). Debido a esta satisfacción del turista y a las emociones positivas provocadas en él nace una nueva corriente que recibe el nombre de experiencia turística memorable, donde la memoria y el recuerdo juegan un papel determinante (Yu et al., 2021).

En la actualidad, las personas dan un alto nivel de importancia a las emociones y todo lo que producen en nuestra vida cotidiana. El ser humano es un ser racional, capaz de analizar, observar y estudiar exhaustivamente todo lo que sucede a su alrededor, para después, actuar en consecuencia y tomar una serie de decisiones. No obstante, esta no es la parte más importante, ya que el lado emocional de las personas puede proporcionar mucha más información acerca del comportamiento y las actuaciones de estos. Es por esto, que a la hora de llevar a cabo investigaciones de cualquier índole es muy necesario prestar especial atención a esta parte. Una manera muy eficaz de medir las emociones y sentimientos de las personas es a través del Análisis de Sentimiento, ya que nos permite extraer a raíz del contenido generado por el usuario las connotaciones del texto, las actitudes y los aspectos emocionales que el autor siente (Bazzaz Abkenar et al., 2021).

El lugar idóneo para poder efectuar este análisis en cuestión son las redes sociales, ya que suponen una nueva forma de comunicación donde multitud de personas se relacionan de forma inmediata entre sí, compartiendo diariamente sus opiniones y

pensamientos sobre diferentes áreas, eliminando con ello las fronteras culturales. En concreto Instagram es la más adecuada, ya que es considerada una de las redes sociales más importantes actualmente, en la cual se pueden compartir diferentes fotografías y vídeos gracias a las múltiples extensiones de las que dispone (Lavagna, 2020).

Por todo lo expresado, en este Trabajo Fin de Grado se va a realizar un Análisis de Sentimiento en la plataforma Instagram sobre una ciudad *Smart Destination* como León, teniendo como objetivo final detectar las emociones y las experiencias de turismo memorables de los turistas.

OBJETIVOS

Con el presente trabajo, se pretenden perseguir diversos objetivos, tanto desde el punto de vista teórico como práctico. Así bien, el objetivo principal a abordar es averiguar los sentimientos y emociones que pueden experimentar los turistas al visitar un destino de turismo cultural, como la ciudad de León y su catedral, así como conocer las experiencias turísticas memorables provocadas.

Siguiendo la estela de este objetivo principal, se pueden definir los siguientes objetivos secundarios, atendiendo a la siguiente clasificación:

- Objetivos secundarios con base en fuentes de información secundaria:
 - Demostrar la importancia de las redes sociales, más concretamente de la plataforma Instagram, para el turismo y el marketing en lo que a experiencia del consumidor se refiere.
 - Demostrar cómo el uso de las redes sociales en general y de Instagram en particular permite conocer el sentimiento del turista y la memorabilidad de la experiencia.
 - Evidenciar la importancia de la información que arrojan las redes sociales sobre el comportamiento de los turistas, y cómo su uso y análisis permite obtener ventajas competitivas a las empresas y organismos públicos.
 - Escuchar y analizar al consumidor a través de las opiniones y reseñas que redacta en las redes sociales sobre la mencionada capital, así como de todo el contenido que genera y que es de especial interés para los expertos en marketing.

- Objetivos secundarios con base en fuentes de información primaria:
 - Realizar un Análisis de Sentimiento a través del contenido generado por los turistas y que es obtenido de las redes sociales.
 - Conocer la polaridad de los comentarios y sentimientos de los turistas sobre la ciudad, es decir, determinar si las valoraciones tienen un carácter positivo o negativo.
 - Averiguar cuáles son aquellas cualidades que logran llamar más la atención de los visitantes sobre dicho monumento cultural.

El problema planteado que se pretende resolver en este trabajo es tomar conciencia de qué tipo de sentimientos pueden aflorar en los turistas sobre la ciudad de León, los cuales serán de gran utilidad para comprender mejor sus experiencias turísticas memorables.

Para conseguirlo, se ha realizado un Análisis de Sentimiento, en el cual cobra especial relevancia las redes sociales, especialmente la plataforma en línea Instagram por su diversidad de funciones y su gran auge entre los internautas. De dicha aplicación se han extraído los comentarios y textos oportunos escritos por los diferentes turistas que han visitado la ciudad y que, posteriormente, son objeto del citado análisis.

Figura 1: Objetivos del estudio



2. METODOLOGÍA

El presente trabajo consta de dos partes diferenciadas. La primera de ellas, el marco teórico, que ocupa los dos primeros capítulos y la segunda, la parte práctica, que está desarrollada en el tercer capítulo.

Para la realización de dichas partes se ha recurrido a fuentes de información tanto primarias como secundarias, cumpliendo así los objetivos establecidos previos a la ejecución del estudio.

2.1. FUENTES DE INFORMACIÓN SECUNDARIA

Para elaborar el marco teórico se ha llevado a cabo una revisión bibliográfica sobre la experiencia turística memorable y ordinaria, así como sobre información referente al Análisis de Sentimiento, utilizando para ello fuentes de información de carácter secundario.

Entre estas fuentes consultadas destacan las siguientes:

- Páginas webs y blogs especializados en turismo, marketing, redes sociales y Análisis de Sentimiento.
- Bases de datos, como *Dialnet* y *ScienceDirect*, donde hay multitud de artículos de revistas y textos que hablan sobre estas temáticas.
- Diversos proyectos, como tesis doctorales y trabajos de fin de grado y de máster.
- Revistas y periódicos online.
- Portales estadísticos y de opinión, como *Statista* y *Tripadvisor*.
- El sitio web oficial de Instagram.

2.2. FUENTES DE INFORMACIÓN PRIMARIA

Para realizar la parte práctica se ha elaborado un Análisis de Sentimiento, y para ello se han empleado diversos programas y plataformas. En primer lugar, *PhantomBuster* mediante el cual se ha llevado a cabo la recogida y descarga de la base de datos desde la red social Instagram. En segundo lugar, Excel para el tratamiento de estos datos. Y, por último, RStudio para la elaboración del análisis, así como para la extracción y visualización de los resultados y gráficos correspondientes.

La base de datos sobre la que se ha trabajado está compuesta por comentarios procedentes de los usuarios de la aplicación social Instagram acerca de León y su catedral, alcanzado un total de 2825 posts entre los años 2011 y 2021.

En la Figura 2 se puede observar el proceso anteriormente descrito.

Figura 2. Metodología primaria empleada



CAPÍTULO I: LA EXPERIENCIA TURÍSTICA MEMORABLE Y LA IMPORTANCIA DEL USO DE INSTAGRAM EN EL TURISMO

1. LA EXPERIENCIA TURÍSTICA

El turismo y, a su vez, viajar está cobrando cada vez más importancia en nuestra sociedad. Las personas recurrimos y realizamos viajes para conocer mundo, escapar del estrés del día a día, cambiar la rutina y lo cotidiano por aventuras nuevas que nos resulten satisfactorias, y hagan que escapemos de nuestros problemas. Es tal la relevancia de este concepto, que para satisfacer las necesidades de estos consumidores ya no basta con ofrecer productos y servicios de turismo de masas, sino que es muy importante dar forma y crear experiencias personalizadas para cada uno de los turistas que visiten un destino determinado (Fusté Forné y Nava Jiménez, 2015).

Una experiencia turística se puede definir como las sensaciones y emociones que un turista experimenta mientras es receptor de una serie de estímulos durante todas las fases de contacto con el servicio o destino turístico (Alonso y Sánchez, 2020).

Según Tung y Ritchie (2011), estas experiencias turísticas son evaluaciones exclusivamente subjetivas e individuales y son experimentadas por el turista durante tres momentos de tiempo diferentes:

- Antes de iniciar el viaje, lo que sería la planificación y preparación de este.
- Mientras se desarrollan las actividades turísticas en el destino.
- Una vez finalizado el viaje, es decir todos los recuerdos que pueden aflorar en la mente y memoria del turista.

Las experiencias turísticas serán más auténticas y reales en el momento en el cual el turista pase de ser un consumidor exclusivamente pasivo a uno activo, interactuando con el lugar en el que se encuentren o con la actividad que estén desempeñando en ese momento, lo que supondrá una mayor afinidad con el lugar y desarrollarán unas determinadas emociones con aquello que se les está ofreciendo (Fusté Forné y Nava Jiménez, 2015). Dicho esto, se puede decir que el servicio o el destino turístico pasa a ser la propia experiencia (Gonçalves Silveira y Barretto, 2010).

2. EXPERIENCIA TURÍSTICA MEMORABLE (MTE)

La experiencia turística queda en la memoria de los consumidores durante un largo tiempo y, a raíz de ella resulta la satisfacción o insatisfacción del turista. Esta experiencia también es capaz de reflejar la valoración del viajero ante un determinado viaje después de haber interactuado con el destino (Coelho et al., 2018). Como se puede observar, mediante el análisis y estudio en profundidad de las experiencias turísticas, los gestores e investigadores turísticos pueden obtener mucha información valiosa sobre la satisfacción del turista, su *engagement* y su lealtad con el destino. Una corriente muy extendida que permite descomponer la experiencia turística y analizar su recuerdo en el tiempo es la experiencia turística memorable (Yu et al., 2021).

2.1. Concepto

Las experiencias turísticas memorables (en adelante MTE) son aquellas experiencias que se construyen a partir de múltiples experiencias turísticas que son recordadas en el tiempo, después de haber ocurrido el viaje o visita. No obstante, no todas las experiencias turísticas pueden considerarse MTE, ya que va a depender de la memoria y su capacidad para recordar (Melón et al., 2018; Yu et al., 2021). Ahora bien, que una experiencia turística sea catalogada como memorable implica que el cliente resulte realmente satisfecho con el servicio turístico en cuestión, que se haya conseguido despertar ciertas emociones positivas en él, que exista la posibilidad de que repita en el futuro y que comparta sus impresiones y emociones con familiares, amigos o personas cercanas, lo que en marketing se denomina WOM (*Word of mouth* o boca a boca) (Wong y Lai, 2021).

Dicho esto, según Coelho et al. (2018) las MTE van a depender de dos componentes:

- La experiencia turística en un lugar y tiempo concreto.
- Una sucesión de recuerdos relacionados con la experiencia.

Las MTE muestran una reflexión centrada en el consumidor que captura sus emociones y todas sus respuestas subjetivas a los lugares que visita, por lo que están compuestas por todos los momentos vividos por los turistas: qué hacían, cómo se sentían y qué estaban pensando cuando realizaban determinadas actividades o acudían a determinados eventos (Yu et al., 2021).

2.2. Efecto de las MTE sobre el comportamiento de los turistas

Este concepto es muy importante a la hora de analizar los viajes realizados por los turistas porque proporciona información muy relevante sobre su comportamiento en el lugar, pero también sobre su comportamiento futuro. Según Melón et al. (2018), estas son las tres razones de tal importancia:

- Cuando la información extraída por las empresas turísticas proviene de las experiencias pasadas de los consumidores, la motivación y predisposición a comprar es alta.
- Los consumidores tienden a percibir estas experiencias pasadas como fuentes de información valiosas y muy confiables, ya que proceden de otros consumidores que ya han disfrutado de ese servicio turístico.
- La experiencia pasada tiene un gran poder de influencia sobre el comportamiento futuro.

A la vista de estas razones, se puede afirmar que las MTE no influyen solo sobre la conducta futura del propio turista, sino que también lo hacen sobre la del resto de turistas potenciales.

Como se ha mencionado antes, las MTE nos otorgan información importante sobre el comportamiento en el destino y el comportamiento futuro, la causa de esto último es la alta capacidad que poseen para influir en la toma de decisiones futuras. Estas decisiones pueden estar relacionadas con la intención de volver a visitar el lugar o la intención de recomendarlo a otras personas, sea cual sea la decisión final que tome el consumidor se concluye que está directamente relacionado con su comportamiento futuro.

2.3. Dimensiones de las MTE

Existe una escala de medición de carácter cuantitativo apta para ayudar a comprender el concepto, la evaluación y la gestión de la experiencia memorable. Esta escala consta de siete dimensiones que afectan a la memorabilidad de un individuo y que se desarrollan, según el autor Kim (2014), a continuación y son las siguientes: hedonismo, revitalización, interacción social y cultura local, significación, conocimiento, implicación y novedad.

2.3.1. Hedonismo

Cuando cualquier persona toma la decisión de planificar y realizar un viaje, evento o cualquier actividad relacionada con el ámbito del turismo, lo que realmente está buscando es vivir una experiencia agradable de disfrute y placer que le colme de emociones y sentimientos positivos. Es decir, está buscando el hedonismo y vivir una experiencia hedónica. Por este motivo, se realiza la premisa de que las experiencias hedónicas permiten construir experiencias memorables satisfactorias para el turista.

2.3.2. Revitalización

Esta dimensión, también llamada relajación o renovación, hace énfasis en la importancia que le dan los consumidores al sentimiento de paz y tranquilidad cuando deciden viajar. Tal y como se apuntaba al inicio de este trabajo, las personas viajamos para liberarnos del estrés y los problemas de la rutina, y es por esto por lo que los destinos turísticos deben brindárselo a sus visitantes. El escapismo es una variable muy importante en cualquier experiencia de viaje. Esta sensación de renovación influye de manera muy positiva en los recuerdos de los viajeros, es decir, mejora la memorabilidad de las experiencias turísticas.

2.3.3. Interacción social y cultura local

La interacción con la cultura local es un factor de motivación cada vez más importante a la hora de viajar, ya que permite aprender sobre las costumbres y el estilo de vida del lugar, y es gracias a este sentimiento de conexión con los lugares lo que impulsa a los visitantes a involucrarse e interactuar con el destino y con sus residentes. De esta forma, los turistas pueden enriquecerse del lugar experimentando y aprender de una forma única. Por esta razón, la cultura local hace que el viaje sea más memorable.

2.3.4. Significación

Las personas cada vez más se esfuerzan por darle un sentido y encontrar un significado a sus vidas, ya que esto les reporta una sensación de bienestar consigo mismas. Al igual que con sus vidas personales, cada vez son más los turistas que buscan viajes únicos que les proporcionen un significado en sus experiencias de viaje. En el momento en el cual el visitante sienta que un determinado viaje le ha servido en su vida o le ha reportado un significado, ya sea emocional, espiritual o físico, dicho viaje y, por tanto, dicha experiencia pasará a ser una de las más memorables de su vida.

2.3.5. *Conocimiento*

Entre las múltiples razones que pueden existir para tomar la decisión de llevar a cabo un viaje, se encuentra la de aprender cosas nuevas y desarrollar nuevos conocimientos o habilidades. Logrando esto el consumidor turístico quedará altamente satisfecho con el viaje, alcanzando así la experiencia memorable. Por tanto, el desarrollo intelectual se ha convertido también en uno de los componentes más importantes de este tipo de experiencias.

2.3.6. *Implicación*

También llamada intervención o participación, esta dimensión defiende que cuando las personas deciden viajar y, con ello, sumergirse e implicarse en un determinado evento es más probable que logren alcanzar la experiencia memorable. La intervención activa en las actividades que conforman sus viajes les reportará una satisfacción y un sentimiento de felicidad que perdurarán en su recuerdo y en su memoria.

2.3.7. *Novedad*

Los consumidores turísticos se sienten motivados por lo original y lo nuevo, razón por la cual un individuo decide viajar. Dicho esto, se puede afirmar que tienden a escoger destinos con culturas y estilos de vida desconocidos para ellos, en definitiva, destinos que sean diferentes a sus lugares y países de origen. La causa de esto es el deseo de experimentar algo nuevo y explorar lo que para ellos es algo desconocido sobre el mundo. Estos aspectos novedosos pueden calar hondo en los turistas haciendo que los recuerden y, por tanto, almacenándolos en su memoria.

2.4. Atributos del destino turístico

Por definición, los atributos son “cada una de las cualidades o propiedades de un ser” (Real Academia Española, 2021). Por lo que, respecto al ámbito del turismo, podemos definirlo como las cualidades o propiedades de un destino o lugar turístico.

Todos los destinos turísticos están compuestos por una serie de atributos que son capaces de afectar e influir en la percepción de los visitantes sobre dicho destino. Por tanto, también tienen un gran poder sobre la decisión de elección de uno u otro cuando se plantea la necesidad de viajar. Pero, además, los atributos del destino van a determinar la satisfacción o insatisfacción del turista, así como el comportamiento que pueden adoptar en el futuro (repetición del viaje, recomendación a familiares o amigos, etc.). Por todos estos motivos, se afirma que los atributos influyen de manera

significativa en la formación de la imagen del destino en la mente de los consumidores turísticos. De la misma manera, también influyen en la creación de experiencias de turismo memorables, por lo que es un aspecto a tener en cuenta cuando queremos describir o mostrar las cualidades de nuestro servicio turístico a los visitantes potenciales (Kim, 2014).

Shen et al. (2021) y Savi Mondo y Gonçalves Silveira (2016) muestran en sus investigaciones algunos ejemplos de atributos de atractivos turísticos.

Con la investigación de los primeros autores citados, se pueden ver los atributos relacionados con puertos deportivos y destinos de navegación. Gracias a los resultados de sus análisis se observa como algunos atributos relevantes en este ámbito son: instalaciones sanitarias limpias, conexiones de agua y electricidad, el clima, seguridad personal, terrenos limpios, actitudes de los empleados, entre otros.

Savi Mondo y Gonçalves Silveira, por su parte, exponen los atributos que más destacan de ocho destinos turísticos brasileños y son: buena localización y fácil acceso, infraestructuras y equipamientos, estética, seguridad, limpieza, las condiciones climáticas, la atención por parte de los guías y empleados, el entretenimiento, entre muchos más.

3. RELEVANCIA DE INSTAGRAM EN EL TURISMO

En los apartados anteriores se ha descrito y explicado qué es una experiencia turística memorable y todos los factores que influyen y toman parte en ella. Ahora bien, para poder afirmar su existencia en los diferentes viajes que se llevan a cabo, tenemos que observar y analizar todo lo que dicen y comentan los turistas que han protagonizado en primera persona esos viajes. Una de las formas más extendidas para poder hacerlo es escuchando todo lo que comparten en las redes sociales, ya que es un espacio donde pueden mostrar su opinión libremente sin limitaciones, por lo que no existirá ningún tipo de sesgo (LetsLaw, 2021).

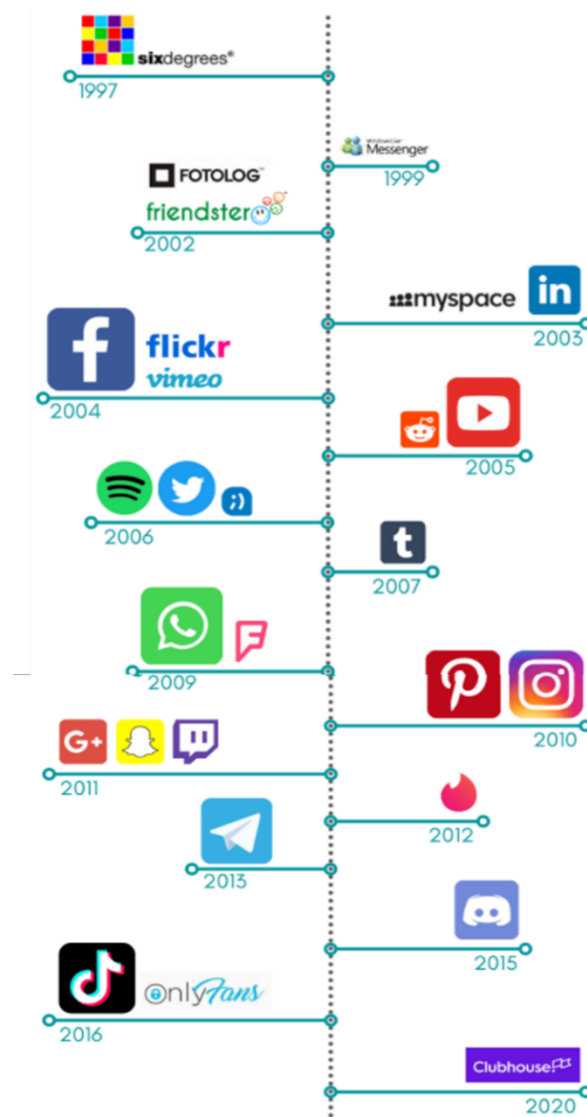
3.1. Crecimiento de las redes sociales en nuestra sociedad

Hace ya algunos años de la llegada de Internet a nuestros hogares, lo que generó una auténtica revolución, ya que nos encontrábamos ante una nueva forma de comunicación entre usuarios sin necesidad de estar cerca unos de otros. Gracias a Internet se pudieron eliminar gran cantidad de fronteras culturales y de idioma. Todo empezó con el correo

electrónico, las páginas web y los foros que poco a poco fueron tomando más fuerza y relevancia. A raíz de esto, se planteó un nuevo escenario, las redes sociales (RR.SS.), las cuales son sitios y plataformas webs donde personas, tanto físicas como jurídicas, interactúan entre sí creando relaciones de forma rápida y sin ningún tipo de limitación física o geográfica (Hera, 2021; Statista, 2021a).

Desde su llegada en 1997, han ido surgiendo múltiples y variadas redes sociales, hasta el 2020, año de lanzamiento al mercado de la última:

Figura 1.1. Historia de las redes sociales: cuándo nacen

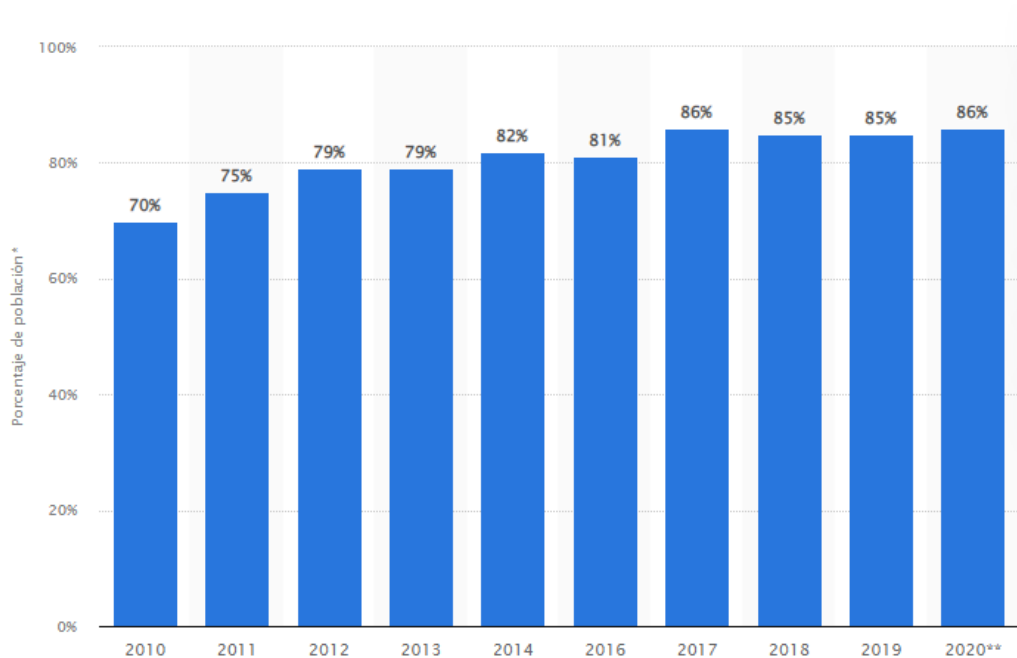


Fuente: (Hera, 2021)

Como se puede observar en la Figura 1.1., existen varios tipos de redes sociales, desde las centradas en contenidos audiovisuales y de ocio, hasta las enfocadas al ámbito profesional.

Desde su nacimiento, han experimentado un éxito indiscutible siendo cada vez más descargadas por los internautas. Una muestra de ello es el gráfico que se muestra continuación:

Gráfico 1.1. Penetración de las redes sociales en España de 2010 a 2020



Fuente: (Statista, 2021b)

A la vista del gráfico, se aprecia como en el periodo comprendido entre los años 2010 y 2020, el uso de las RR.SS. muestra una tendencia creciente alcanzando porcentajes del 80%, a excepción de un ligero descenso en el 2016 y 2018. El porcentaje actual de penetración se sitúa en el 86%.

En la actualidad, un total de 37,4 millones de españoles se consideran usuarios activos de las redes sociales a fecha de enero de 2021, lo que equivale al 80% de la población española total. Por si fuera poco, esta cifra se encuentra en crecimiento, y ha aumentado un 27,6%, lo que supone 8,1 millones de usuarios activos más en este último año en comparación con el pasado año 2020. Estos son datos de nuestro país, pero si ponemos la mirada en el resto del mundo los datos son abrumadores. Se calcula que hay 4.200 millones de usuarios de estas plataformas, es decir, un 53,6% del total de habitantes del planeta, encontrándose esta cifra también en crecimiento, concretamente un 13,2% más con respecto al año pasado (Europa Press, 2021).

Tomando como referencia estos datos, se afirma la importancia innegable que tienen las redes sociales en nuestra sociedad y el gran volumen de información que diariamente los usuarios de estas deciden escribir y compartir en ellas. Por lo tanto, es por esta razón que en el presente trabajo se va a hacer uso de estas plataformas multimedia para analizar las posibles experiencias de turismo memorables.

Como se ha podido ver en la Figura 1.1., se han lanzado multitud de redes sociales al mercado, de las cuales nos centraremos en Instagram, ya que se la considera una de las más importantes en la actualidad, así como una de las más versátiles (Statista, 2021c).

3.2. Instagram como medio de comunicación

Instagram es una red social y aplicación móvil que cuenta con más de 1.000 millones de usuarios en todo el mundo, siendo una de las que mayor crecimiento tiene. Empezó a estar disponible al público en el año 2010 y fue creada para publicar y compartir imágenes y vídeos con el resto de personas (Lavagna, 2020).

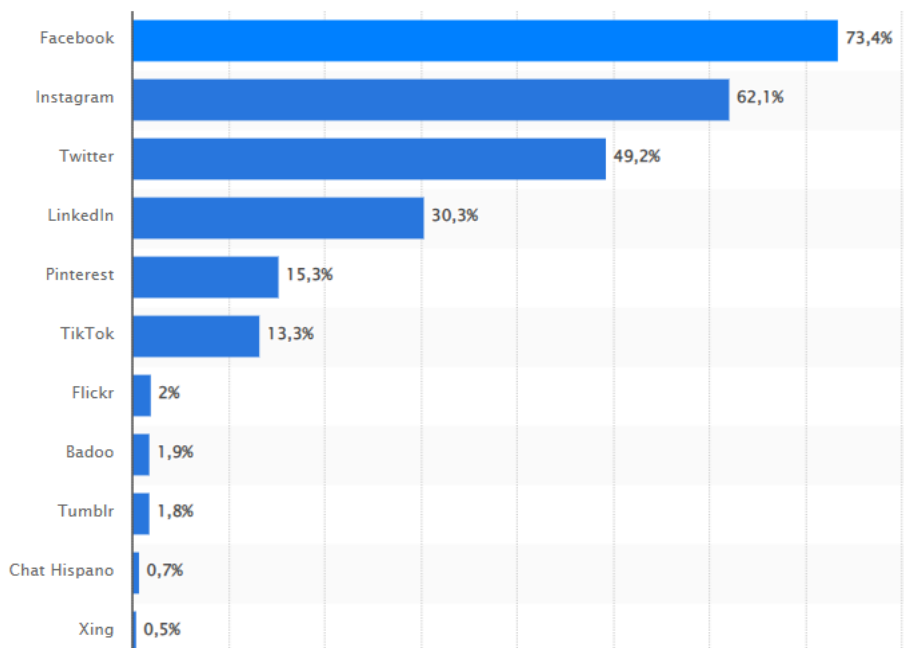
Esta plataforma es considerada una de las más versátiles de todas las disponibles en el mercado, ya que cuenta con multitud de funciones que ayudan a los usuarios a expresarse y conectarse con el resto de los internautas, y son las siguientes:

- *Instagram Stories*: con una duración limitada de 24 horas, las historias son una forma rápida y fácil de compartir a través de vídeos o fotos momentos y experiencias vividas, pudiendo añadirle texto, música, GIF y stickers, así como filtros de imagen y vídeo. Además, con esta función también se puede interactuar con amigos agregando la modalidad de preguntas y respuestas o encuestas, gracias a las cuales los seguidores pueden responder y participar. Fue una de las primeras funciones de la aplicación en activarse y, actualmente, es una de las más usadas por los usuarios pero también por las marcas (Instagram, 2021c).
- *Reels*: son vídeos cortos de hasta 30 segundos en los que puedes dar rienda suelta a la creatividad con distintos textos, filtros, efectos e incluso herramientas y clips de audio. Este formato permite crear como usuario tus propios vídeos, pero también existe la posibilidad de interactuar junto al de otra persona grabando tu vídeo junto al suyo, lo que Instagram denomina “remixar” un *reel* (Instagram, 2021d; Romero, 2021).

- *Instagram TV (IGTV)*: esta extensión de Instagram está destinada a la creación de vídeos de larga duración, teniendo el máximo establecido en 60 minutos. Son vídeos de mayor calidad que, como pasa con las funciones anteriores, también permite interactuar entre los usuarios y amigos (Francés, 2020).
- *Compras*: este apartado de Instagram ha sido uno de los últimos en habilitarse y está destinado a la compra de productos de distintas marcas. Se pueden efectuar muchas compras con tan solo un toque en la etiqueta del producto deseado y no es necesario salir de la aplicación (Instagram, 2021a).
- *Messenger*: esta opción permite al usuario mantener conversaciones privadas con sus amigos o seguidores. Se pueden enviar mensajes de voz, fotos y vídeos con infinidad de efectos y filtros, enviar publicaciones o historias y realizar videollamadas (Instagram, 2021b).
- *Live*: también llamados directos, esta función da la posibilidad de realizar retransmisiones en vivo para compartir cualquier tipo de contenido con los seguidores, y a su vez estos pueden enviar comentarios como *feedback* a ese directo (MOTT Marketing Digital, 2019).

Esta plataforma es una de las más descargadas, como muestra el siguiente gráfico:

Gráfico 1.2. Redes sociales con mayor porcentaje de usuarios en España en 2020



Fuente: (Statista, 2021c)

Solo por debajo de Facebook, con un 62,1% de usuarios, Instagram es una de las redes sociales preferidas por los internautas y esto es en parte por la gran versatilidad que posee. Gracias a las extensiones anteriormente mencionadas, los internautas pueden compartir mucha información en forma de fotografías y vídeos, tanto largos como cortos, y mostrar su vida cotidiana, así como sus opiniones y sentimientos acerca de temas de actualidad.

Ninguna otra aplicación tiene tantos formatos para poder comunicar, lo que la convierte en un gran pozo de información, esto sumado a su innegable popularidad hace que sea la plataforma ideal para analizar el sentimiento surgido a raíz de las experiencias memorables de los turistas.

3.3. Importancia de Instagram para el turismo

El nacimiento de Internet y las redes sociales marcaron un hito en la historia y fueron muchos los sectores que se reinventaron y decidieron sumarse a esta revolución, y uno de ellos fue el sector turístico.

Las redes sociales, especialmente Instagram, ha supuesto un antes y un después en el mundo de los negocios. Muchas empresas, marcas y pequeños comercios han decidido embarcarse en esta nueva vía de comunicación abriéndose una cuenta en la plataforma y promocionando sus productos. Gracias a esta nueva forma de vender ha surgido la figura de los *influencers*, personas con un volumen de seguidores elevado que tienen la capacidad de influir en ellos.

Al igual que otros sectores, el turismo también se ha unido a esta nueva forma de comunicar, una prueba de ello es la cantidad de empresas de la industria turística que han optado por convertirse en usuarios de esta red social, como por ejemplo agencias de viaje y oficinas de turismo, entre otros. No obstante, no solo las empresas son las únicas que han optado por este canal de comunicación, a ellos los acompañan los consumidores y turistas, y es que cada vez son más los que deciden descargarse la aplicación y compartir en ella sus vivencias personales.

Gracias a las RR.SS. las personas se pueden comunicar sobre cualquier tema con el resto del mundo, y más concretamente en el turismo, pueden intercambiarse información sobre el servicio turístico o sobre las características y opiniones que tienen acerca de los distintos destinos y países. El impacto de las redes sociales, concretamente

Instagram, en la industria del turismo se ha magnificado enormemente, una muestra de ello es la importancia que le dan los turistas a la información que leen en ella, considerándola una importante referencia para viajar. Así mismo, al compartir constantemente sus experiencias personales e interactuar con las del resto de usuarios, pueden conseguir más información que respalde su decisión de viaje, la cual condicionará su comportamiento futuro. Por todo esto, se reforzó la idea de que las experiencias turísticas se co-crean en las redes sociales. Dicho esto, estas plataformas se han considerado como una fuente de información y una herramienta significativa para la investigación en turismo y marketing (Fatanti y Suyadnya, 2015; Gon, 2021).

Instagram como plataforma multimedia y gracias a la posibilidad de realizar instantáneas y vídeos, ofrece a los usuarios un escaparate de destinos que consigue atraer a la audiencia, creando una percepción sobre el destino turístico. Como ya se ha comentado en el apartado anterior, esta red social dispone de diversas extensiones para comunicar, que cuando son utilizadas por el consumidor mostrando fotografías de sus viajes y publicando sus impresiones y sentimientos, se convierten en un símbolo y en un vehículo de transmisión muy importante de la experiencia del usuario como viajero. Con esta contribución de los internautas se está empezando a configurar la marca de muchos destinos turísticos. Dicha marca se encuentra influenciada por el eWOM (*electronic word of mouth* o el boca o boca electrónico), gracias a los comentarios que escriben los consumidores sobre los productos y servicios que observan en Instagram (Fatanti y Suyadnya, 2015).

Como se puede ver, esta plataforma es utilizada por multitud de sectores y consumidores y es de gran ayuda para la industria turística, ya que gracias a la interacción de los internautas se fomentan diferentes experiencias de turismo, a la vez que influyen en las marcas turísticas y, por lo tanto, en la imagen de los destinos.

CAPÍTULO II: ANÁLISIS DE COMENTARIOS MEDIANTE EL ANÁLISIS DE SENTIMIENTO

1. CONTENIDO GENERADO POR EL USUARIO

La proliferación de las redes sociales en esta era digital y, con ello, la gran cantidad de contenido que se genera ha sido uno de los impulsores de cambio más importantes en el sector turístico.

El contenido generado por el usuario (UGC) hace referencia a todo aquel contenido que ha sido creado por los usuarios de Internet. Los consumidores confían cada vez más en este contenido en lugar de en el contenido generado por la marca, ya que al ser textos escritos en primera persona por los internautas se consideran más confiables (Morgan et al., 2021).

UGC es un factor externo que puede influir en los consumidores y, por ende, en su manera de actuar. En concreto en la industria turística, este tipo de contenido es considerado una fuente muy valiosa de información capaz de influir en la toma de decisiones futuras sobre qué destino elegir para las vacaciones, a dónde irse de viaje, etc. El contenido generado por el usuario es capaz de provocar expectativas en los futuros turistas, las cuales pueden considerarse, en algunas ocasiones, como las desencadenantes de las decisiones de realizar esos determinados viajes, movidos por el entusiasmo y las ganas de vivir aquello que han leído (Bigne et al., 2020).

Por estos motivos, muchas son las empresas que están aprovechando esta tendencia al alza del UGC invirtiendo en publicidad y posicionamiento en línea, es decir, en marketing de redes sociales. Según Cambria et al. (2012, p. 558), “el poder del marketing radica, no en llevar información a las masas, sino en aprovechar eficazmente a los individuos que ejercen influencia sobre los demás”.

2. ANÁLISIS DE SENTIMIENTO

2.1. Concepto

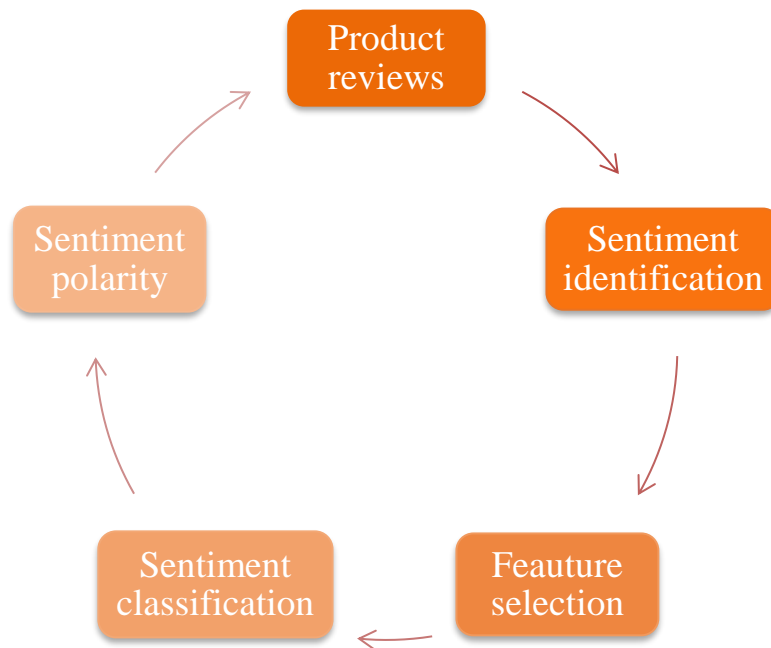
El Análisis de Sentimiento, también llamado *opinion mining*, es una nueva disciplina que se ha empezado a plantear recientemente, cobrando cada vez más interés. Consiste en la extracción e identificación de actitudes, emociones y sentimientos que puede experimentar una persona mediante el análisis de ciertos comentarios u opiniones, es decir, mediante el análisis del contenido generado por el usuario. Se encuentra dentro

del área del Procesamiento del Lenguaje Natural (PLN) y clasifica la polaridad de los sujetos en positivos, negativos o neutros determinando e identificando la percepción del público. Gracias a este análisis se pueden extraer opiniones sobre multitud de productos y servicios, marcas, política o cualquier otro tema considerado de interés. El Análisis de Sentimiento tiene varias aplicaciones como el análisis de reseñas y comentarios, atención al cliente, detección de emociones, análisis de intenciones, monitoreo de redes sociales y monitoreo de marca (Bazzaz Abkenar et al., 2021; Cambria et al., 2012). Los resultados que arrojen estos análisis serán un factor de consideración relevante para la toma de decisiones.

Esta técnica de Análisis de Sentimiento es considerada muy importante a la hora de proporcionar información crucial a las empresas sobre sus propios clientes. Las razones que avalan esta afirmación son la innegable necesidad que tienen los consumidores por interactuar e intercambiar información, sentimientos o pensamientos con el resto de los usuarios, todo ello a través de las redes sociales, lo que fomenta el eWOM. Siendo esto la causa directa y más influyente en la formación de opiniones de los internautas, las cuales al estar plasmadas en las redes sociales son de fácil acceso para las empresas y las pueden utilizar, mediante el mencionado análisis, para sacar conclusiones y mejorar la toma de decisiones (Bueno et al., 2022).

Según Medhat et al. (2014), el proceso que se sigue cuando se realiza un Análisis de Sentimiento es el siguiente, tal y como se muestra en la Figura 2.1.:

- Encontrar opiniones o reseñas.
- Identificar los sentimientos que se expresan en dichas opiniones.
- Seleccionar las características que se detecten en las reseñas.
- Clasificar el sentimiento.
- Determinar la polaridad del sentimiento.

Figura 2.1. Proceso de Análisis de Sentimiento

Fuente: (Medhat et al., 2014, p. 1094)

2.2. Los tres niveles de aplicación en el Análisis de Sentimiento

El Análisis de Sentimiento puede ser considerado como un proceso de clasificación de sentimientos, por lo que cabe destacar los tres niveles principales a los que se aplica: nivel de documento, nivel de oración y nivel de aspecto. Según Medhat et al. (2014) y Panico y Cesteros (2018):

- Nivel de documento o texto: su función consiste en estudiar documentos o textos completos y determinar cuál es la opinión o sentimiento que se manifiesta en ellos, es decir, clasificarlos en positivos, negativos o neutros. Gran parte de los estudios actuales se centran en este tipo.
- Nivel de oración: su fin es identificar los sentimientos de las oraciones. Para ello, antes se debe categorizar la oración como objetiva o subjetiva, sabiendo que se podrá llevar a cabo la clasificación de las emociones en positivas o negativas si se determina la oración como subjetiva. Las oraciones pueden considerarse textos cortos, por lo que no se encuentran diferencias significativas entre la clasificación a nivel de documento y a nivel de oración, exceptuando que, en las oraciones, normalmente, solo se expresa una opinión, mientras que en los textos se pueden interpretar varias.

- Nivel de aspecto o de características: lo que se pretende con este tipo de clasificación es identificar los aspectos relacionados con la entidad objeto de estudio (considerando entidad a una empresa y los aspectos a sus atributos, por ejemplo, entidad: empresa del sector de la restauración y aspectos: trato al cliente, limpieza, variedad de los platos, personal, etc.). Para después, concluir si hay algún tipo de opinión sobre ellos. Si se revela que sí, hay que clasificar la opinión en positiva, negativa o neutra. Sobre este determinado nivel existen muy pocos trabajos.

El Análisis de Sentimiento que se va a realizar en este estudio está desarrollado en base al nivel de oración.

2.3. Enfoques del Análisis de Sentimiento

Para poder llevar a cabo este tipo de análisis se va a depender de varios enfoques o métodos, los cuales son: enfoque con base en el aprendizaje automático, enfoque con base en el léxico y el enfoque híbrido. Se explican, según Medhat et al. (2014), a continuación.

En primer lugar, el enfoque de aprendizaje automático o *machine learning* está caracterizado por utilizar diferentes algoritmos, para poder así clasificar el texto en positivo o negativo. Dentro de este tipo podemos encontrar:

- Métodos de aprendizaje supervisados: son aquellos que hacen uso de una gran cantidad de documentos etiquetados, tomándolos como ejemplo.
- Métodos de aprendizaje no supervisados: en algunas ocasiones, es complicado encontrar estos documentos etiquetados, y es entonces cuando se emplea este tipo de método.

En segundo lugar, cuando se decide emplear el enfoque con base en el léxico, se utiliza una lista de palabras o términos de sentimientos conocidos para extraer la polaridad de las opiniones. Al igual que el enfoque anterior, también se subdivide en dos:

- Basado en diccionario: consiste en recopilar de forma manual un conjunto de palabras o términos de opinión etiquetadas previamente con el sentimiento que expresan. Posteriormente, se amplía este listado mediante *WordNet*, una base de datos léxica, y se buscan sus sinónimos y antónimos. Cada palabra encontrada se va añadiendo a la lista deteniéndose este proceso cuando no se

encuentren más términos. La gran desventaja de este tipo es que en ocasiones se puede perder cierta precisión, puesto que determinadas palabras pueden poseer distintos significados según el contexto en el que se encuentren y, por tanto, distinta polaridad. Por ejemplo, el calificativo “silencioso” tiene una connotación positiva cuando se refiere al ambiente nocturno en los alrededores de un hotel, pero se convierte en negativa si hace referencia al sonido de un dispositivo electrónico (Pauli y Soliani, 2019).

- Basado en corpus: este método nos va a ayudar a resolver el problema planteado anteriormente. Para ello, se utilizarán corpus compuestos por palabras y términos relacionados con un dominio o temática específica, para poder reconocer su significado y polaridad según el contexto en el que se encuentren.

Por último, el enfoque híbrido es aquel método que combina los dos enfoques anteriores, con base en aprendizaje automático y con base en el léxico.

De los tres enfoques explicados, el que se va a llevar a cabo en esta investigación es el enfoque con base en el léxico y, dentro de este, el basado en diccionario.

Figura 2.2. Enfoques del Análisis de Sentimiento



Fuente: Adaptado de Medhat et al., 2014

2.4. Análisis de Sentimiento en la industria turística

Obtener la visión que tienen los turistas sobre los productos y servicios turísticos es de vital importancia para las empresas y los destinos.

La cantidad de comentarios generados en las redes sociales por los usuarios sobre turismo, viajes y otros factores relacionados es enorme, razón por la cual esta industria

está altamente influenciada por las valoraciones de los visitantes sobre sus experiencias turísticas. Estas valoraciones abarcan desde la experiencia en hoteles y restaurantes, los eventos o actividades turísticas, monumentos o visitas culturales, etc. Todos estos comentarios que, como se ha visto, hacen referencia a distintos aspectos, deben ser aprovechados para optimizar la toma de decisiones de las compañías turísticas, para mejorar aquellos aspectos que no obtengan las valoraciones y los resultados esperados, seguir potenciando todo lo positivo y recalibrar las acciones de marketing lanzadas por las empresas. Además, realizar una escucha activa en las redes sociales puede ayudar a detectar las fortalezas del destino como marca turística, aprovechar las oportunidades de negocio que puedan quedar insatisfechas, mejorar servicios o redefinir campañas de marketing (HuffPost, 2018).

Estas opiniones no solo reportan beneficios a las compañías o a los destinos, sino que también sirven de gran ayuda a los turistas, puesto que al leer los comentarios de otros usuarios van a obtener de primera mano mucha información, y va a ser en base a dicha información que van a tomar las decisiones sobre su viaje. Por estas razones, las valoraciones de los turistas en los diferentes portales webs van a influir de manera muy directa en la reputación de los destinos y de las empresas de turismo (HuffPost, 2018).

Existen diversas investigaciones que utilizan el Análisis de Sentimiento en el turismo y Flores-Ruiz et al. (2021) nos muestra varios ejemplos de ellas llevadas a cabo por distintos autores. Por un lado, uno de estos estudios analiza en profundidad las características que los clientes de *Airbnb* priorizan de una ciudad, elaborando una tipología de ciudades desde la perspectiva del turista. Para ello, la metodología empleada es la recopilación de datos que, posteriormente, son tratados en R. Por otra parte, otra investigación analiza las percepciones que tienen los turistas sobre el País Vasco como destino turístico, concluyendo que la mayor influencia ejercida sobre la imagen de la región como atractivo turístico viene dada por los recursos naturales y culturales. Algunos ejemplos más de estas investigaciones que muestra este autor se sintetizan en la siguiente tabla:

Tabla 2.2. Ejemplos de estudios de Análisis de Sentimiento en turismo

AUTORES	AÑO	TEMAS	RESULTADOS
(Valdivia et al., 2018)	2018	Cotejar reseñas negativas sobre monumentos para detectar mejoras	Los usuarios son muy objetivos a la hora de dar su opinión
(Gascón et al., 2019)	2019	Análisis de opinión en RR.SS. para la comunicación digital de tres empresas turísticas	Priman los comentarios positivos frente a los negativos
(Gálvez-Rodríguez et al., 2020)	2020	Análisis de la participación en línea de turistas en Facebook	Tendencia positiva en la participación y compromiso en línea de los turistas
(Martínez-Torres y Toral, 2019)	2019	Identificación de reseñas engañosas en sector de la hostelería	Críticas engañosas enfatizan los inconvenientes físicos y las no engañosas se centran más en los sentimientos y en las cualidades de la ciudad

CAPÍTULO III: ANÁLISIS DE SENTIMIENTO EN RSTUDIO SOBRE LA CIUDAD DE LEÓN

En este último capítulo, se va a realizar el Análisis de Sentimiento con el programa estadístico RStudio sobre la ciudad de León y uno de sus monumentos culturales más importantes, la catedral.

1. LEÓN COMO DESTINO DE TURISMO CULTURAL Y *SMART DESTINATION*

La investigación va a estar centrada en la ciudad de León y, como se ha dicho con anterioridad, en la catedral.

Para esta elección nos basamos en un hecho y es en el gran patrimonio cultural del cual hace gala la capital leonesa, entre lo que podemos destacar la catedral que resalta por sus vidrieras, el convento y parador de San Marcos, la Casa Botines que es uno de los escasos ejemplos de las obras de Gaudí fuera de la capital catalana Barcelona y el Museo de Arte Contemporáneo MUSAC (Destino Turístico Inteligente, 2018).

Llegado este punto, cabe destacar que el turismo cultural está inspirado en tres causas concretas: la valoración en alza de la cultura y el patrimonio histórico por la sociedad, la tendencia a abandonar el modelo tradicional de sol y playa y, en último lugar, la mayor importancia que las personas otorgan a los momentos dedicados al ocio. Por estas tres causas, se puede determinar que el turismo cultural es aquel turismo en el que el patrimonio cultural e histórico se encuentra en el centro de la visita (Prada et al., 2016). Por consiguiente, teniendo en cuenta esta definición y a la vista de las razones descritas en el párrafo anterior, la ciudad de León se encasilla dentro de los lugares que son considerados destinos de turismo cultural. Este hecho ha sido uno de los factores clave para la elección de León como la ciudad objeto del análisis.

No obstante, existen otras razones por las cuales León se convierte es una ciudad de atractivo turístico. En primer lugar, en la consideración de la ciudad como uno de los destinos más populares de turismo de interior en España. En segundo lugar, el hecho de ser uno de los puntos del Camino de Santiago, considerado Patrimonio de la UNESCO. En tercer lugar, el galardón que le fue otorgado a la ciudad considerándola como Capital Gastronómica Española en el año 2018. Y en último lugar, la festividad de Semana

Santa, siendo una de sus fiestas más importantes, la cual fue declarada Fiesta de Interés Turístico Internacional (Destino Turístico Inteligente, 2018).

Así mismo, debido a una serie de iniciativas llevadas a cabo por la capital, ha tenido lugar su incorporación al proyecto Destinos Inteligentes (*Smart Destination*) en el año 2018. Dichas iniciativas son las siguientes (Destino Turístico Inteligente, 2018).

- Posesión de la certificación *Biosphere*, reconocimiento otorgado a los destinos y empresas que han hecho del turismo sostenible el eje central de sus actividades, siendo ejemplos de sostenibilidad en el sector turístico a nivel mundial (Biosphere Tourism, s. f.).
- Adhesión a la declaración de Shanghái de ciudades saludables, demostrando un gran compromiso con el desarrollo sostenible.
- Apoyo al emprendimiento y la innovación, una muestra de ello es el Centro de Empresas son Base Tecnológica.
- Búsqueda de la participación ciudadana.
- Proyecto europeo Simpla, iniciativa para apoyar a los municipios en las políticas de movilidad urbana sostenible.
- Defensa de los derechos de las personas con discapacidad de León, gracias al Consejo Municipal para la Discapacidad.

Este reconocimiento también ha sido un punto clave para que la investigación tenga lugar en torno a la ciudad de León.

La principal actividad que los turistas realizan cuando visitan Castilla y León es para visitar monumentos, concretamente un 68.2% de los visitantes, según un informe del Portal de Turismo de la Comunidad (Junta de Castilla y León, 2020). Por lo tanto, como se ha dicho previamente, esta investigación va a estar centrada en la catedral, puesto que, de todos los puntos turísticos que se encuentran a disposición de los visitantes durante su estancia en la capital, esta es la que mayor público atrae. Los turistas la consideran una de las más bonitas del mundo, así como el elemento más representativo de la ciudad, por lo que, dada su importancia, es la mejor opción sobre la cual llevar a cabo este estudio (Revista Catedral de León, 2021). Una muestra de esto son los comentarios depositados en *Tripadvisor*. “*El centro de León es la catedral*”, “*Uno de los monumentos que hay que ver en España*”, “*Si vas a León la visita obligada es la catedral, es preciosa por dentro y por fuera*”, “*No hay adjetivos para describirla, la*

más bonita del mundo”, estos son algunos ejemplos de las reseñas registradas por los turistas y que prueban la relevancia de este templo (Tripadvisor, 2021).

2. PROGRAMA ESTADÍSTICO RSTUDIO

RStudio es un programa estadístico que es capaz de dar sentido a los datos. El fin principal de esta herramienta es crear un software libre y de código abierto para la ciencia de datos, que sea útil para ayudar a las personas a comprender y mejorar el mundo a través de los datos. R viene equipado con múltiples funciones para la importación de datos, el ajuste y la evaluación de modelos estadísticos, representaciones gráficas, etc. Entre todas estas funciones encontramos los paquetes (*packages*), que son considerados una colección de funciones y datos que se encuentran almacenados en una carpeta. La lista de paquetes disponibles abarca hasta las 12.000, pudiendo utilizarse cada uno dependiendo de las áreas de aplicación en las que nos encontremos (RStudio, s. f.; Santana y Hernández, 2016)

RStudio es utilizado para realizar análisis de diversa índole, entre los que se encuentran el Análisis de Sentimiento.

De todos los paquetes disponibles en R, en esta investigación, se va a usar el paquete *syuzhet*, entre otros. Este paquete es uno de los más adecuados para el Análisis de Sentimiento, porque extrae las emociones y sentimientos derivados de la opinión del texto usando para ello una amplia variedad de diccionarios (Bueno et al., 2022; Jockers, s. f.). Concretamente utiliza cuatro y son: *Bing*, *Afinn*, *Stanford* y *NRC*. Este último está disponible en más de cien idiomas, entre ellos el castellano, ventaja por la cual es el elegido para emplear en esta investigación. Dicho diccionario contiene exactamente 14.182 palabras repartidas todas ellas atendiendo a estas dos clasificaciones (Isasi, 2021):

- El sentimiento, ordenándose como positivas o negativas.
- Las emociones, clasificando las palabras según ocho emociones principales, que son: enfado, anticipación, asco, miedo, alegría, tristeza, sorpresa y confianza.

Este paquete ha sido empleado en varias investigaciones, y en gran parte de ellas la red social más extendida para llevar a cabo el análisis fue Twitter frente al resto. Un ejemplo de ello es el estudio realizado por Karmegam y Mappillairaju (2022), en el que

analizan los *tuits* de la población de la India sobre la diabetes. La razón por la cual es la preferida por los investigadores para llevar a cabo este tipo de análisis, es por la dificultad de obtención de datos de otras redes sociales como Instagram o Facebook, que cada vez tienen más restringido el acceso a sus amplios volúmenes de datos por motivos de privacidad (Teles da Mota et al., 2022). No obstante, hay algunos investigadores que sí han tenido acceso a estos datos y han realizado estudios en Instagram con este mismo paquete *syuzhet*, como Ginzarly y Jordan Srour (2021) que llevaron a cabo un proyecto sobre el intercambio de contenido referente al patrimonio cultural durante la pandemia de COVID-19.

3. METODOLOGÍA DEL ESTUDIO

Para el desarrollo de esta investigación, la base de datos necesaria contiene información procedente de la red social Instagram y se han elegido aquellos comentarios en los cuales se difundan opiniones sobre la ciudad de León, pero, sobre todo, aquellos en los que se hable sobre la catedral. Con estos datos se realizará un Análisis de Sentimiento mediante el programa estadístico RStudio.

3.1. Recogida de datos

Para poder llevar a cabo el análisis el primer paso es descargar la base de datos a analizar, y el programa que se va a emplear es *PhantomBuster*. Esta herramienta es la responsable de automatizar y recopilar la información de las redes sociales, tales como Instagram, Twitter, etc. (Ranker, 2022).

La base de datos ha sido descargada en formato .csv y, posteriormente, se ha procedido con su importación en Excel. El fichero resultante tiene un total de 2825 comentarios, de los cuales 415 no tienen ningún tipo de comentario u opinión, 450 están compuestos exclusivamente de *emojis* o *hashtags* y 1552 poseen al menos un *hashtag*. Todos estos comentarios corresponden al período de tiempo comprendido entre los años 2011 hasta 2021.

3.2. Tratamiento de datos

Una vez descargado e importado a Excel el archivo se empieza con el tratamiento de los datos. En primer lugar, se han obtenido varias pestañas de Excel para comprender mejor el comportamiento de los turistas, seleccionando, finalmente, la pestaña denominada “sin dupli”, que es la que contiene los comentarios de los turistas. No obstante, dentro de esta pestaña hay diversas columnas con información de distinta índole: URL de la

imagen, ID del usuario, fecha de publicación del comentario, etc., siendo únicamente la columna H la necesaria para realizar el Análisis de Sentimiento, ya que es donde se encuentran los posts de los turistas en castellano (véase en la Figura 3.1.).

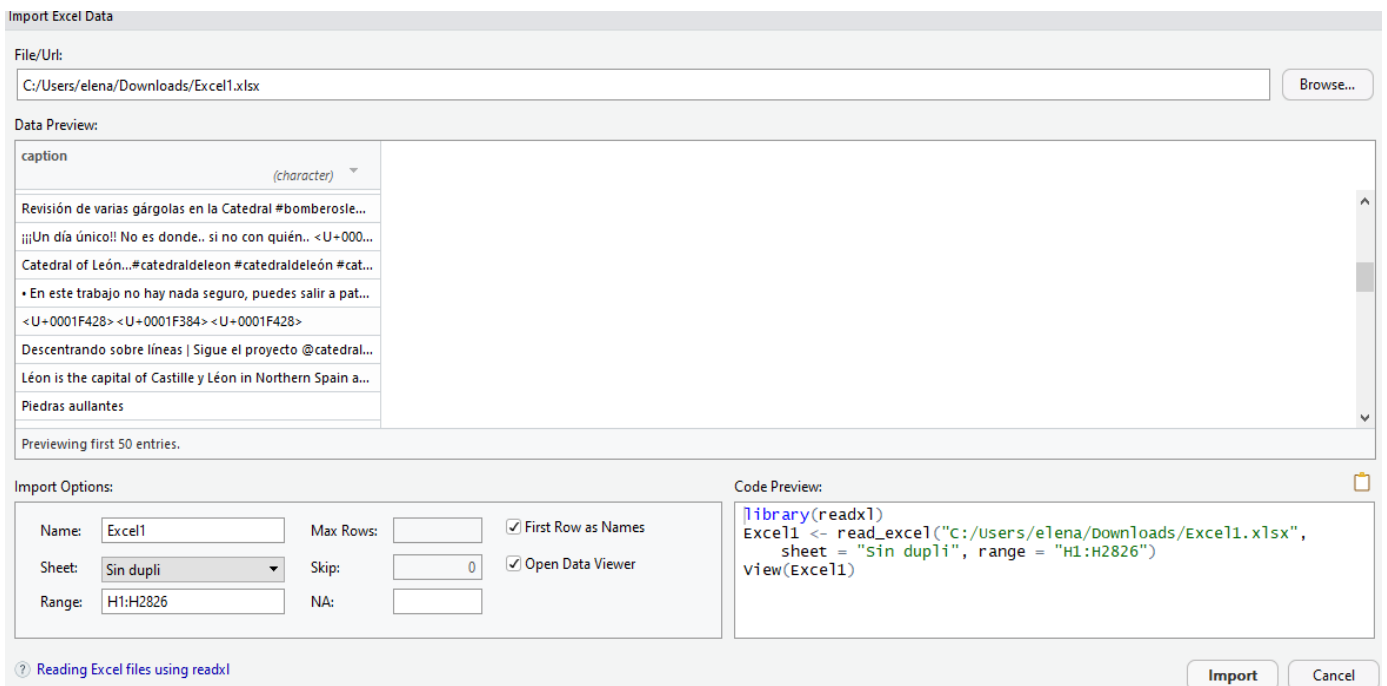
4. ANÁLISIS ESTADÍSTICO EN RSTUDIO

4.1. Análisis de frecuencias de términos

En este estudio se van a realizar dos Análisis de Sentimiento. En primer lugar, se va a llevar a cabo un primer análisis del que van a formar parte todos los comentarios extraídos y que componen la base de datos. Y, en segundo lugar, se desarrollará otro análisis con los comentarios que impliquen memorabilidad y, por tanto MTE, para poder observar las diferencias entre ambos.

Empezando con el primer análisis, una vez que la base de datos está en formato Excel, se procede a la importación de esta en el programa estadístico RStudio. Como se ha mencionado en el punto anterior, únicamente se va a necesitar la pestaña llamada “sin dupli” y, dentro de esta, la columna de los posts en castellano, es decir la H.

Figura 3.1. Importación base de datos a RStudio



Fuente: RStudio

Como resultado de esta importación RStudio reporta este comando:

```
library(readxl)
Excel1 <- read_excel("C:/Users/elena/Downloads/Excel1.xlsx",
sheet = "Sin dupli", range = "H1:H2826")
```

Una vez finalizada la importación del archivo, se descargan e instalan los paquetes, si fuese necesario, para poder llevar a cabo el análisis. Como se ha mencionado en el apartado anterior, el paquete principal *syuzhet* será uno de los elegidos, siendo su función realizar el Análisis de Sentimiento. No obstante, no solo se va a emplear ese paquete, sino que también se hará uso de seis más, y son:

- *Tidyverse*: está formado por otros paquetes de R, todos ellos diseñados para la manipulación y visualización de los datos (Santana y Hernández, 2020).
- *Tm*: es el encargado de organizar, transformar y analizar los datos textuales (Feinerer et al., 2008).
- *Snowballc*: es utilizado para hacer *stemming*, es decir para reducir una palabra a su raíz. Gracias a ello se pueden analizar variaciones de una palabra. Por ejemplo, de la palabra conectar se pueden obtener las variaciones conexión, conexiones, conectado, entre otras (Snowballstem, s. f.).
- *Wordcloud*: ayuda en la visualización de los datos creando nubes de palabras (CRAN.R, 2018).
- *RColorBrewer*: su función reside en proporcionar esquemas y paletas de colores para los diferentes gráficos (CRAN.R, 2014).
- *Tidytext*: está compuesto por distintas funciones y conjuntos de datos las cuales son útiles para manipular el texto en beneficio del análisis (CRAN.R, 2021).

```
if (!require("tidyverse")) install.packages("tidyverse")
if (!require("tm")) install.packages("tm")
if (!require("Snowballc")) install.packages("Snowballc")
if (!require("wordcloud")) install.packages("wordcloud")
if (!require("RColorBrewer")) install.packages("RColorBrewer")
if (!require("syuzhet")) install.packages("syuzhet")
```

Es imprescindible cargar los paquetes necesarios para el desarrollo del análisis, sino el programa detectará error y no se podrá continuar. Los comandos para ejecutar son los siguientes:

```
library("tidyverse")
library("tm")
library("Snowballc")
library("wordcloud")
library("RColorBrewer")
library("syuzhet")
library("tidytext")
```

Una vez que ya están cargados los paquetes se puede empezar a desarrollar el análisis. Lo primero que hay que hacer es limpiar el documento, y para ello se ejecuta este código:

```
Excel1$texto_limpio <- Excel1$caption %>% tolower() %>%
removePunctuation() %>% removewords(words = stopwords(kind
="spanish")) %>% removewords(words = stopwords(kind ="english"))
%>% removewords(words =c("buen" , "buenos" , "buenas" , "bueno",
"buena" , "follow" , "architecture" , "art" , "gothic" , "night"
, "ciudad" , "city", "león", "leon" , "catedral" , "cathedral" ,
"spain" , "españa" ,"leonesp" , "mundo" , "días" ,
"catedraldeleon" , "igersleonesp" , "castilla y león" ,
"castillayleon" , "travel", "estaesleon" , "semana" , "foto" ,
"fotos" , "estaescastillayleon" , "estaescastillaleon" , "hoy" ,
"día" , "dia" , "day" , "picoftheday" , "estaesespania" ,
"siempre" , "catedraldeleón" , "igersspain" , "photography" ,
"igers" , "igersleon" , "catedrales" , "tag" , "photo" ,
"photooftheday" , "travelgram", "travelphotography" , "photos" ,
"okspain" , "gran" , "san" , "ser")) %>% stripwhitespace()
```

Se han utilizado diferentes funciones para limpiar los datos. En primer lugar, la función *tolower* es la encargada de transformar todo en minúsculas. En segundo lugar, con *removePunctuation* se suprimen los signos de puntuación. En tercer lugar, gracias a la función *removeWords* se pueden eliminar las *stopwords* en español y en inglés, que son palabras carentes de contenido y significado (por ejemplo, preposiciones o artículos), así como las palabras que no aporten nada a la investigación por su poca relevancia. En este caso, se han suprimido una lista de 45 términos carentes de importancia para el fin del estudio. Y, por último, *stripWhitespace* nos ayuda a eliminar los espacios en blanco.

A continuación, se convierten los documentos en un corpus. Un corpus es una serie de documentos que contienen texto y que serán objeto de un análisis determinado (RDocumentation, s. f.).

```
Corpus <- VCorpus(VectorSource(Excel1$texto_limpio))
```

Una vez que el corpus ya está formado, se procede a realizar una nube de palabras. Esta representación gráfica muestra las palabras que se repiten con más frecuencia en los comentarios, siendo representadas con un tamaño proporcional a dicha frecuencia (Rodríguez, 2019). Para ello, se debe ejecutar la siguiente línea de código:

```
wordcloud(Corpus, scale = c(2, 0.1), min.freq = 5,
max.words = 30, colors = brewer.pal(n = 4 , name ="Dark2"))
```

En este comando se hace uso de diferentes funciones. Por un lado, con la función *scale* se fija el tamaño máximo y mínimo de letra, en este caso el tamaño máximo es 2 y el mínimo 0.1. Por otra parte, la función *min.freq* representa la frecuencia mínima que debe tener una palabra para ser representada, es decir el número mínimo de veces que tiene que aparecer, en este caso 5. *Max.words* lo que hace es establecer el máximo número de palabras que pueden ser representadas en la nube, en este estudio se ha fijado el límite en 30. Y, finalmente, gracias a la función *colors* se puede dar color a las palabras representadas, eligiendo la paleta deseada. Teniendo en cuenta todo esto, se ha obtenido la siguiente representación gráfica:

Figura 3.2. Nube de palabras



Como se puede observar, las dos palabras que más utilizan los turistas para referirse a León y a su catedral son “caminodesantiago” y “camino”, ya que son las representadas en mayor tamaño. “Gotico”, “arquitectura”, “vidrieras” y “santa” son palabras también bastante recurridas. En cambio, “plaza”, “arte”, “turismo”, “beautiful”, entre varias más, son empleadas en menor medida.

Para poder procesar de manera correcta los datos, el siguiente paso a seguir va a ser convertir el corpus en una matriz término-documento, también llamada *term-document matrix* o TDM. El código que hay que introducir para conseguirlo es el siguiente:

```
TDM <- TermDocumentMatrix(Corpus)
```

Para poder observar sus características hay que escribir el sencillo comando que se muestra abajo, obteniéndose la siguiente matriz:

```
inspect(TDM)
```

Tabla 3.1. Matriz término-documento (TDM)

TÉRMINOS	DOCUMENTOS									
	118	1426	1809	1926	2053	2136	2154	2264	2707	323
Arquitectura	0	1	0	0	0	0	0	0	0	0
Arte	0	0	0	0	0	0	0	0	0	0
Beautiful	0	0	0	0	0	0	0	0	0	0
Camino	0	0	0	0	2	0	0	0	0	1
Camino de Santiago	0	0	0	0	0	0	0	0	0	1
Gotico	0	0	0	0	0	0	0	0	0	0
Luz	0	2	0	0	0	0	0	0	0	0
Pulchra	0	1	0	0	0	0	0	0	0	0
Santa	0	0	0	0	0	1	0	0	0	0
Vidrieras	0	3	0	0	1	0	0	0	0	0

Esta TDM tan solo es una pequeña muestra que es proporcionada por RStudio de toda la matriz. Se puede ver cómo las filas representan algunos términos que están dentro del corpus y las columnas algunos documentos que contiene el corpus. Todos los dígitos que están representados determinan el número de veces o la frecuencia de los términos en cada uno de los documentos, en este caso, el número de veces que un término aparece en cada uno de los comentarios. Por ejemplo, se aprecia como el término “vidrieras” se ha mencionado tres veces en el comentario número 1426, mientras que ninguna vez en el 323.

Posteriormente, una vez visto en la matriz una pequeña muestra de la frecuencia de las palabras en cada comentario, se han extraído los términos de la TDM con más de 50 y más de 30 apariciones en todos los comentarios que componen el documento. Para ello, se ha elaborado la siguiente instrucción:

```
findFreqTerms(TDM, lowfreq = 50)
findFreqTerms(TDM, lowfreq = 30)
```

Respecto a la primera línea de código, algunas de las palabras resultantes con una frecuencia igual o superior a 50 son las que se muestran a continuación:

```
"luz"           "arquitectura"  "vidrieras"
"arte"          "beautiful"     "camino"
"caminodesantiago"
```

Por el contrario, algunos términos que se han obtenido con apariciones superiores o iguales a 30 correspondientes al segundo comando son los siguientes:

```
"history"       "gracias"       "santiago"
"monumentalspain" "friends"       "noche"
"plaza"
```

Como se puede ver, todas estas palabras están directamente relacionadas con la Catedral de León, dejando ver cómo algunos de los turistas aprecian su arquitectura y arte, la historia que alberga, las espectaculares vidrieras de la fachada y la luz que estas desprenden, etc.

A continuación, se va a convertir la TDM actual en una matriz, la cual recibirá el nombre de `tdm`. Esta nueva matriz sí que va a proporcionarnos la tabla entera, a diferencia de la matriz término-documento (TDM) extraída anteriormente, que solo nos sacaba una pequeña muestra. Se puede obtener utilizando la siguiente instrucción:

```
tdm <- as.matrix(TDM)
```

La tabla resultante es demasiado extensa, tanto de filas como de columnas. Además, al ser exactamente igual a la Tabla 3.1. solo que con mayor extensión y volumen de datos se va a omitir su representación.

El siguiente paso, una vez creadas las matrices que nos aportan el número de veces que aparece cada palabra en cada comentario, va a ser elaborar una matriz de frecuencias por término, también denominada *term-frequency matrix* o TFM. Con esta nueva tabla, se podrá ver la frecuencia o recuento total de cada uno de los términos existentes. Se empleará para ello el siguiente código:

```
TFM <- data.frame(Frecuencia = rowSums(tdm)) %>%
  rownames_to_column("término") %>% arrange(desc(Frecuencia))
```


Tabla 3.2. Matriz de frecuencias por término (TFM)

TÉRMINO	FRECUENCIA		
Camino	127	Regla	44
Camino de Santiago	116	Love	43
Gotico	81	Santiago	43
Arquitectura	80	Vida	42
Vidrieras	72	María	39
Santa	69	Turismo	39
Pulchra	61	Bonita	38
Arte	56	Noche	36
Beautiful	56	Visitar	36
Luz	52	Amigos	33
Plaza	51	History	33
Leonina	50		
Monumentalspain	50		
Gracias	48		
Instagood	45		
Pulchrleonina	45		
Gótico	44		

Al igual que pasaba con la tdm, esta matriz es muy extensa, por lo que se he optado por exponer tan solo una pequeña parte de toda la tabla.

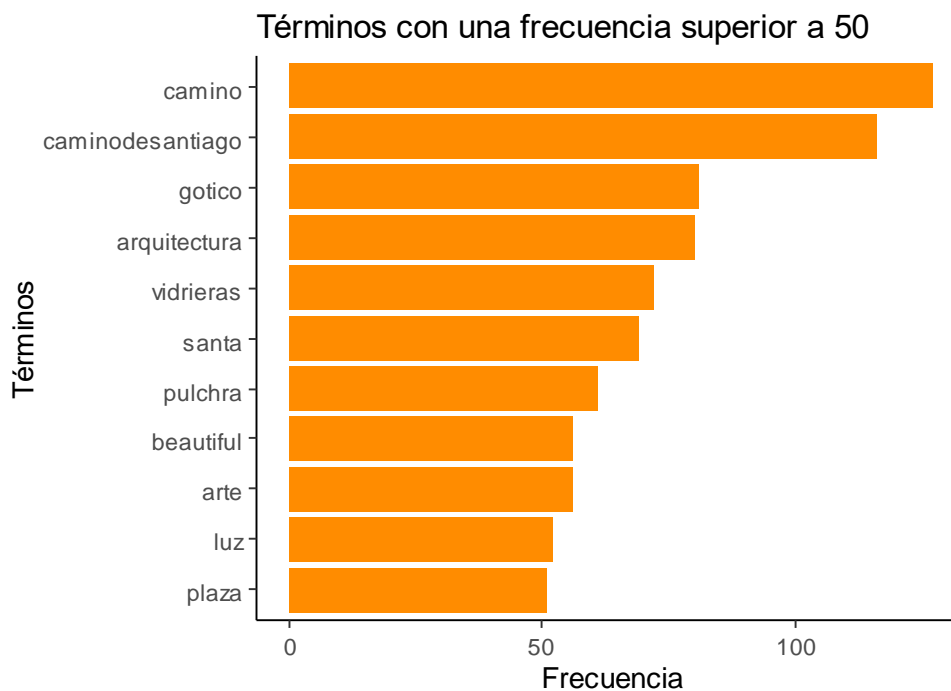
Por lo tanto, como se puede observar en la Tabla 3.2., el término más empleado por los turistas para referirse a la Catedral de León es “camino” con 127 repeticiones, seguido muy de cerca por “caminodesantiago” con 116. “Gotico” y “arquitectura” también son palabras bastantes recurridas alcanzando un valor de 81 y 80 repeticiones respectivamente. Por el contrario, “bonita”, “history” y “noche” son menos empleadas, teniendo alrededor de 30 repeticiones.

A continuación, para poder representar gráficamente la TFM mostrada hay que introducir la siguiente instrucción:

```
TFM %>% filter(Frecuencia > 50) %>% arrange(desc(Frecuencia))
%>% ggplot(mapping = aes(x= reorder(Término, Frecuencia),
y = Frecuencia)) + geom_col(fill = "Dark Orange") + coord_flip()
+ labs(title = "Términos con una frecuencia superior a 50",
x = "Términos",) + theme_classic()
```

Se han seleccionado aquellos términos con una frecuencia mayor a 50 (*frecuencia*>50). La función *reorder* lo que hace es ordenar los términos según su frecuencia, *geom_col* ayuda en el diseño y color del gráfico, *coord_flip* es la encargada de transponer el gráfico para que las palabras se lean mejor, con *labs* se pone el título deseado al gráfico y a los ejes que lo componen y *theme_classic* le aporta un diseño clásico a la representación. Por lo tanto, el gráfico que ha resultado es el siguiente:

Gráfico 3.1. Frecuencia de los términos



Al igual que se ha visto en la Tabla 3.2., se puede ver en este gráfico que la palabra “camino” y “caminodesantiago” son las más empleadas por los visitantes que acuden a León, haciendo referencia a que están visitando la ciudad y, en concreto, la catedral porque se encuentran en pleno Camino de Santiago. El término “gotico” también es de los más utilizados, y demuestra el aprecio de los turistas hacia el estilo artístico en el que fue construida la catedral. “Arquitectura”, por su parte, hace referencia a la buena construcción y diseño que aprecian del edificio. Con el término “vidrieras” y “luz”, se puede interpretar que a los viajeros les gusta la iluminación que desprende la

composición de estas cristalerías. “Santa” y “plaza”, siendo este último el término menos utilizado, son dos palabras que hacen referencia al lugar en el cual está situada la catedral, ya que la plaza donde se encuentra recibe el nombre de Santa María de Regla. Esto refleja el interés que tienen los visitantes por profundizar sobre el lugar y querer conocer más de él, en vez de limitarse a observar el templo. “Pulchra”, que significa bella, es una parte del segundo nombre que se le conoce a la Catedral de León, siendo el nombre completo Pulchra Leonina, que quiere decir Bella Leonesa. Con esto, al igual que pasaba con los dos términos mencionados con anterioridad, se demuestra el conocimiento sobre el edificio lo que es un indicador del buen resultado de la visita. Por último, dos de las palabras menos frecuentes de todas las que han sido utilizadas en más de 50 ocasiones son “Arte”, la cual da una señal de la buena sensación que sienten los turistas, ya que han podido disfrutar y apreciar el arte que esconde el monumento, y “*beautiful*”, palabra que significa bonito, demuestra la buena opinión e impresión que tienen los viajeros.

4.2. Análisis de Sentimiento

Después de analizar la frecuencia de repetición de los términos en los comentarios compartidos, se va a realizar el análisis de comentarios mediante el Análisis de Sentimiento.

En primer lugar, para poder realizar el análisis se debe ejecutar la siguiente línea de código:

```
sentimientos <- get_nrc_sentiment(Excel1$texto_limpio,  
language = "spanish")  
Excel1 <- cbind(Excel1, sentimientos)
```

Con la función *get_nrc_sentiment* se obtienen los sentimientos, tanto positivos como negativos, presentes en los comentarios. Más específicamente, busca la presencia en cada uno de los comentarios de las ocho emociones diferentes que esta función es capaz de detectar.

Tabla 3.3. Clasificación de comentarios según las emociones

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Negative	Positive
1	0	0	0	0	0	0	0	0	0	0
2	2	1	1	2	0	3	0	2	6	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	1
7	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	2
14	0	0	0	0	0	0	0	0	0	0

Esta tabla es un pequeño ejemplo de la tabla más grande que ha proporcionado el programa RStudio. En las columnas se pueden ver los ocho sentimientos que se pretenden detectar en los comentarios: *anger* (enfado), *anticipation* (anticipación), *disgust* (asco), *fear* (miedo), *joy* (alegría), *sadness* (tristeza), *surprise* (sorpresa) y *trust* (confianza). Las dos últimas columnas nos muestran el sentimiento positivo o negativo del comentario en cuestión y las filas representan el número de comentario. Además, los números que componen la tabla significan el número de palabras que han sido relacionadas con las emociones. Por ejemplo, en el segundo comentario se puede ver cómo hay dos palabras que son relacionadas con la emoción que hace referencia al enfado, una relacionada con anticipación y otra con asco, dos que hacen referencia al miedo, tres con relación a la tristeza, dos que tienen que ver con la confianza y, finalmente, 6 son las palabras que clasifican este comentario como negativo. Por su parte, el comentario trece tiene dos palabras que lo clasifican como positivo. Los comentarios que tienen ceros en todas las emociones probablemente no tengan ninguna

palabra que se pueda clasificar en alguna de las emociones propuestas, por ejemplo, por su corta extensión.

Para sintetizar y organizar la información de la Tabla 3.3., se ha creado un *data.frame* y para ello se ha empleado el siguiente código:

```
sentimientos_df <- data.frame(valor
=rowSums(t(sentimientos)))%>% rownames_to_column("Emoción")
%>% arrange(desc(valor))
```

Tabla 3.4. Suma total de emociones del texto

EMOCIÓN	VALOR
Positive	1442
Trust	707
Joy	649
Anticipation	631
Negative	533
Surprise	340
Fear	320
Sadness	308
Anger	165
Disgust	124

En esta tabla se puede ver que la emoción principal en los comentarios de los turistas es la confianza (*trust*) con un valor de 707 palabras, la segunda emoción más predominante es la alegría (*joy*) con un valor de 649. Por el contrario, la emoción compartida menos común es el asco (*disgust*) con tan solo 124 palabras, seguida muy de cerca por el enfado (*anger*) con 165. Esto quiere decir que los comentarios analizados reportan un mayor volumen de emociones satisfactorias.

Respecto a los sentimientos, se han detectado 1442 palabras positivas mientras que tan solo 533 negativas, lo que se traduce en un mayor número de comentarios positivos.

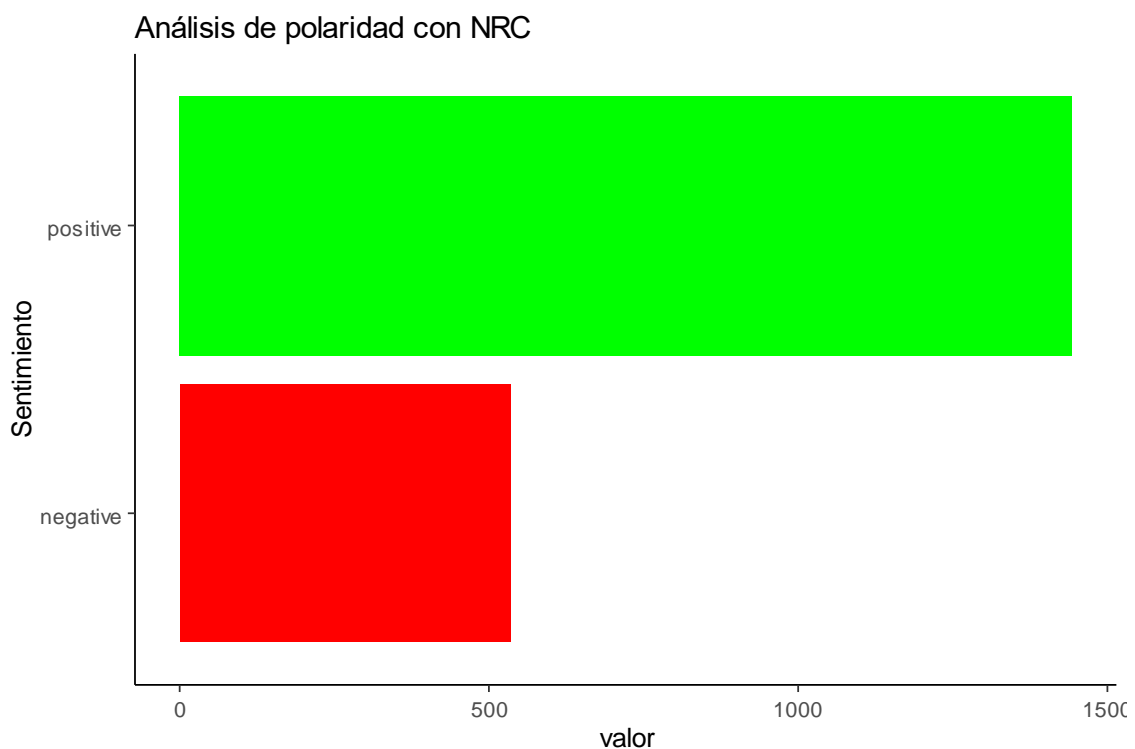
Se puede concluir, a la vista de esta tabla, que los comentarios dejados por los turistas en la red social Instagram son mayoritariamente positivos (polaridad positiva), siendo la emoción más destacada la confianza y la alegría.

Para representar gráficamente los datos recogidos en la Tabla 3.4, se deben ejecutar los comandos que se indican a continuación.

En primer lugar, para observar la polaridad, tanto positiva como negativa, de los comentarios analizados:

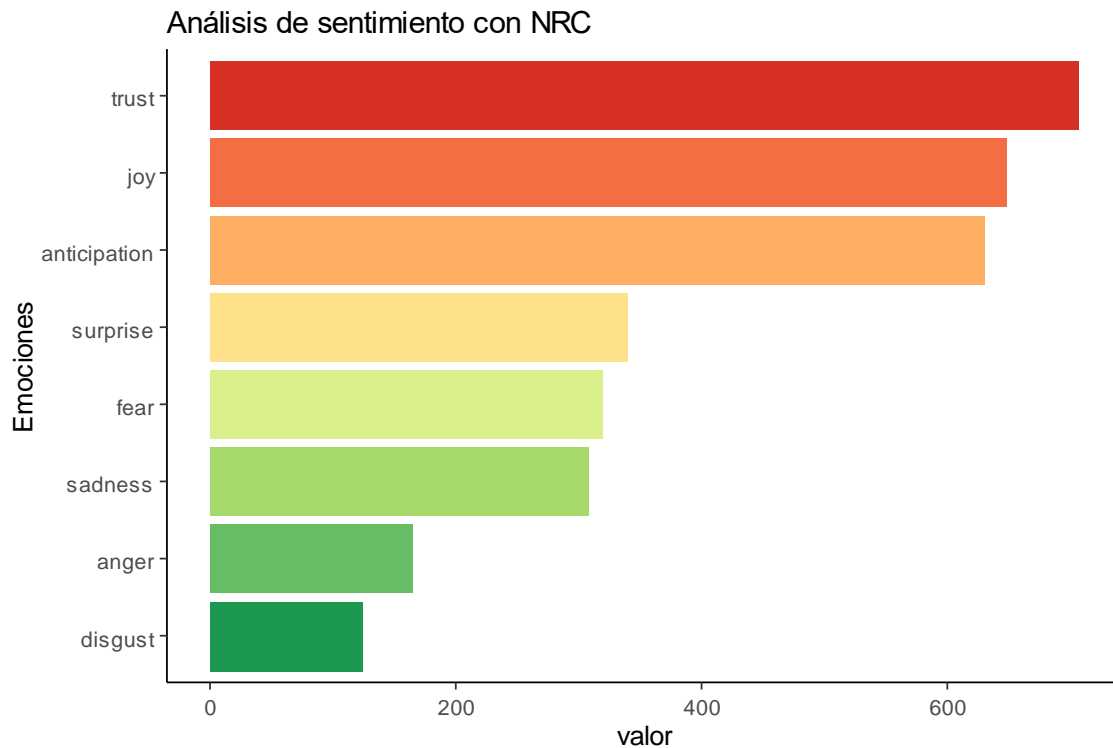
```
sentimientos_df %>% filter(Emoción %in% c("positive",  
"negative")) %>% ggplot(mapping = aes(x= reorder(Emoción,  
valor), y = valor)) + geom_col(fill = c("Green", "Red")) +  
coord_flip() + labs(title = "Análisis de polaridad con NRC",  
x = "Sentimiento") + theme_classic()
```

Gráfico 3.2. Análisis de la polaridad de sentimientos



En segundo y último lugar, para examinar las emociones que más predominan en los comentarios:

Gráfico 3.3. Análisis de Sentimiento



Una vez vista tanto la polaridad como las emociones predominantes en los comentarios, se van a catalogar las palabras atendiendo a dicha clasificación para descubrir cuáles son las más utilizadas en cada sentimiento y en cada emoción.

Para ello, se van a necesitar dos léxicos (hasta ahora solo se había empleado uno, el léxico *nrc*). No obstante, como los datos que los componen se encuentran en inglés, se va a utilizar una traducción de estos, para que los resultados obtenidos figuren en castellano. El primero de ellos, el léxico *Afinn*, se empleará para la separación de palabras según su sentimiento. Este léxico está formado por un conjunto de palabras puntuadas entre -5 a 5 de acuerdo con la percepción de su significado, pudiendo encasillarse como positivas, más cerca del cinco, o negativas, más cerca del -5. Por otro lado, el léxico *nrc*, se usará para clasificarlas teniendo en cuenta las emociones. Al igual que el anterior, también está compuesto por una sucesión de palabras, en este caso, distribuidas en función de la emoción que trasmite su definición.

En primer lugar, para extraer los términos más populares asociados a cada sentimiento (positivo o negativo) se debe ejecutar el siguiente código:

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

Con este primer comando, se va a descargar la traducción del léxico *Afinn* para después poder hacer uso de él.

A continuación, hay que introducir la siguiente línea de código:

```
Excell_afinn <- Excell %>% unnest_tokens(input = "texto_limpio",
output = "Palabra") %>% inner_join(afinn, ., by = "Palabra") %>%
mutate(Tipo = ifelse(Puntuacion > 0, "Positiva", "Negativa"))
```

Con la función *unnest_tokens* se va a conseguir separar cada uno de los comentarios en palabras independientes para después, con la función *inner_join*, asignar a cada una de ellas el sentimiento que le corresponde. *Mutate*, en este caso, lo que hace es crear una nueva columna llamada “tipo” donde clasifica a las palabras, según la puntuación que tengan, como positivas o negativas.

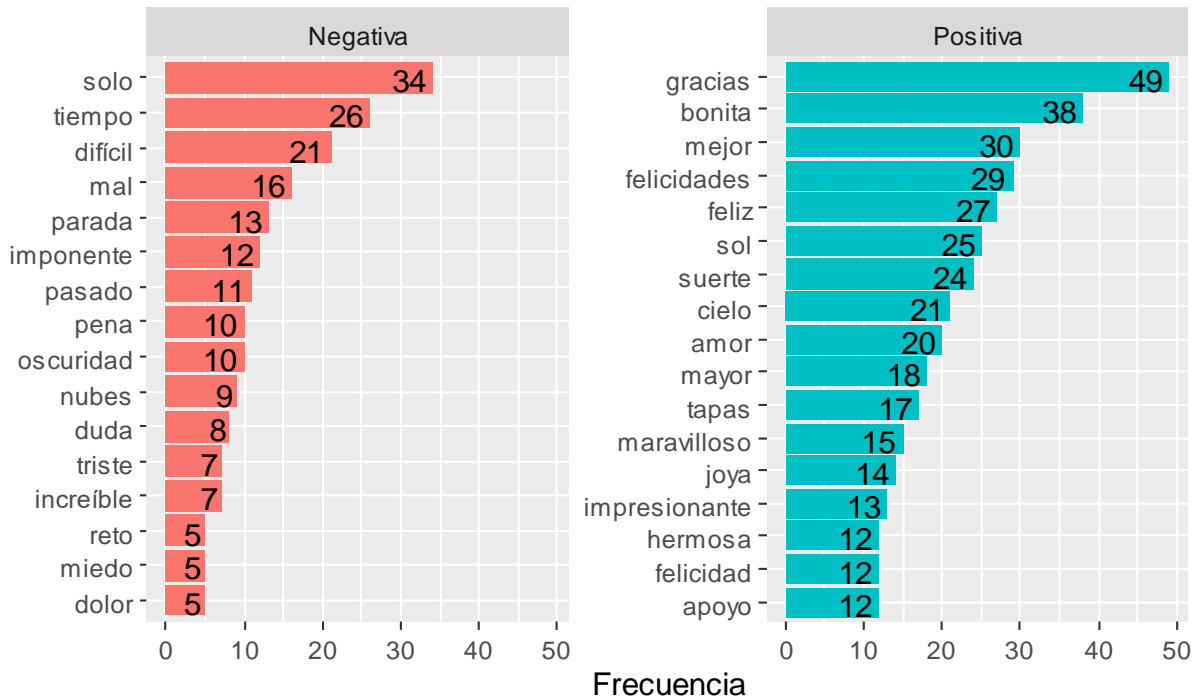
Finalmente, para poder representar gráficamente esta distinción hay que ejecutar el siguiente comando:

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

Se emplean múltiples funciones. En esta instrucción, *Count* cuenta las palabras según el sentimiento, *group_by* se encarga de agruparlas según la columna “tipo”, es decir según el sentimiento. Con *top_n* se selecciona el número de palabras deseado para representar. *Mutate* reordena mediante la función *reorder* la variable “Palabra” según la variable “n”, en otras palabras, ordena las palabras según la frecuencia. *Ggplot* es la encargada de asignar a cada sentimiento un color. Gracias a *geom_col* y *geom_text* se puede ocultar la leyenda y agregar una etiqueta a los valores del eje, respectivamente. Mediante *facet_wrap* se puede descomponer un gráfico en varios, en este caso en dos según el tipo de sentimiento, positivo y negativo. Por último, con *xlab* se puede ocultar

el nombre del eje x para una mejor visión del gráfico. Por lo tanto, el gráfico resultante es el que se muestra a continuación:

Gráfico 3.4. Clasificación de palabras según el sentimiento



Se puede observar cómo “solo” es la palabra negativa más empleada con 34 repeticiones, seguida por “tiempo” y “difícil” con una frecuencia de aparición de 26 y 21 cada una. En cambio “reto”, “miedo” y “dolor” son las menos usadas con tan solo 5 repeticiones. A la vista de estas palabras, se puede decir que algunos puntos negativos que los turistas sienten al visitar León y la catedral son el estado meteorológico, de ahí las palabras “nubes” y “tiempo”, aunque esta última podría hacer referencia también a la tristeza por una estancia corta en la ciudad.

En contraposición, “gracias” es la palabra positiva más empleada con 49 repeticiones, seguida por “bonita” con 38. Por el contrario, “hermosa”, “felicidad” y “apoyo” son menos usadas con únicamente 12 apariciones. En este caso, se puede interpretar que los visitantes resaltan positivamente la belleza de la catedral por palabras como “bonita”, “joya”, “hermosa” e “impresionante”. También destacan la comida, como se puede apreciar gracias a la aparición del término “tapas”, y se puede ver la satisfacción que les ha reportado el viaje según los conceptos “felicidad”, “gracias”, “feliz” y “felicidades”.

En segundo lugar, para extraer los términos más populares asociados a cada una de las ocho emociones se debe ejecutar la siguiente instrucción:

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

Al introducir esto se descargará la traducción del léxico NRC, para poder empezar con el análisis.

Para modificar la composición de este léxico, en concreto el nombre de columna dos, se debe incorporar el siguiente código:

```
colnames(nrc)[2]<- "Emoción"
```

Este paso es importante para evitar confusiones, ya que por defecto el nombre establecido era “sentimiento” y tal denominación hace referencia a la clasificación analizada anteriormente. Por lo cual, se ha creído conveniente modificarla y llamarla “Emoción”, como indicativo de las ocho emociones que van a ser analizadas.

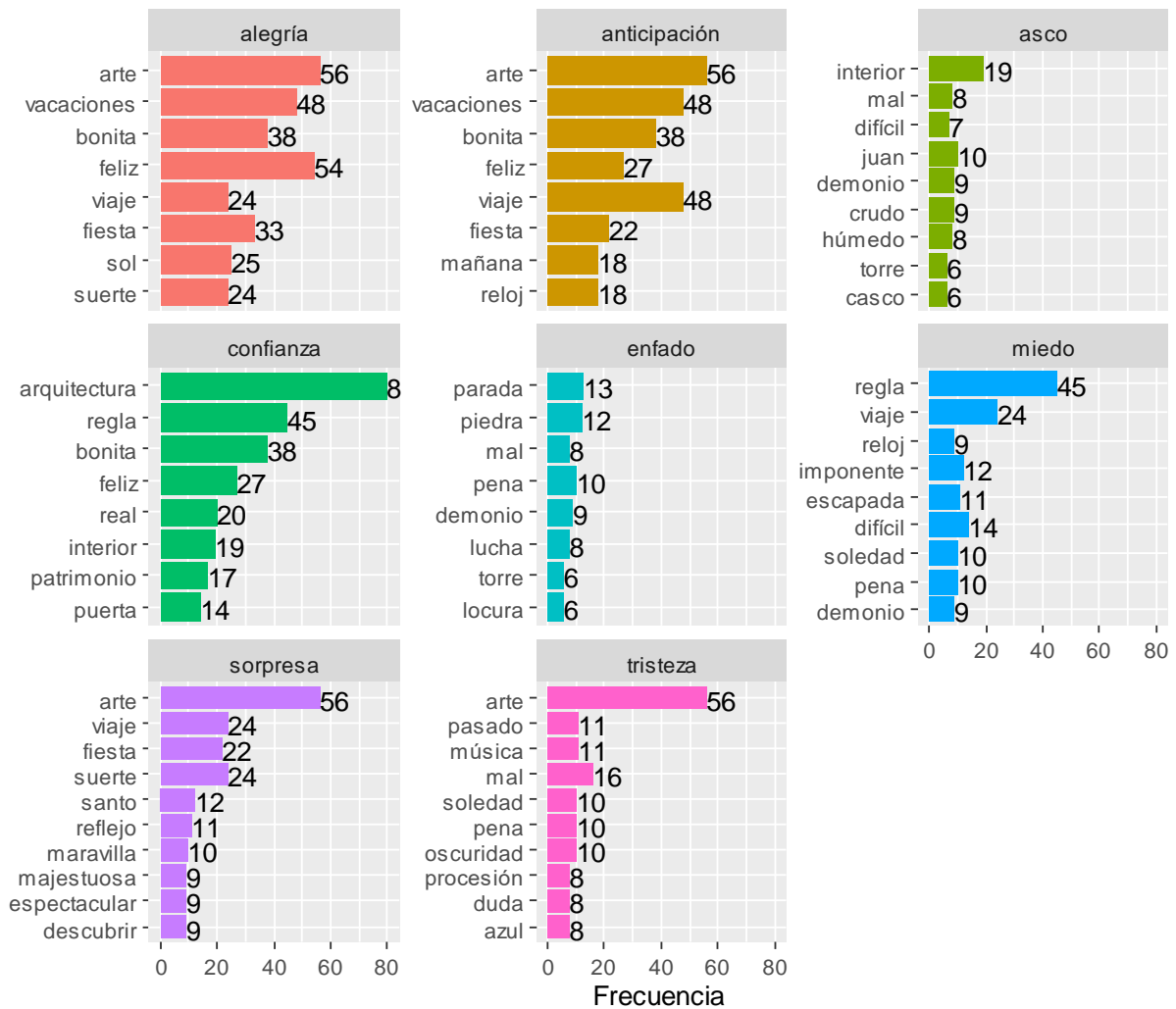
A continuación, para poder obtener la representación gráfica oportuna se introducirán las siguientes líneas de código:

```
Excell_nrc <- Excell %>% unnest_tokens(input = "texto_limpio" ,
  output = "palabra") %>% inner_join(nrc, by = "palabra")

Excell_nrc %>% filter(Emoción!="negativo" & Emoción!="positivo")
%>% count(palabra, Emoción, sort=TRUE) %>% group_by(Emoción) %>%
top_n(8) %>% ungroup() %>% mutate(palabra=reorder(palabra, n))
%>% ggplot(aes(palabra,n,fill=Emoción))+ geom_col(show.legend =
FALSE) + geom_text(aes(label=n), hjust= 0) +
facet_wrap(~Emoción, scales = "free_y") + coord_flip() +
labs(y="Frecuencia") + xlab(NULL)
```

En este caso y a diferencia del gráfico que clasifica las palabras según su polaridad, con la función *filter*, se van a filtrar los sentimientos “negativo” y “positivo” para omitirlos en la representación. De este modo, solo aparecerán en el gráfico las ocho emociones objeto del análisis. La función *group_by* va a distribuir las palabras según las emociones. Y, mediante *facet_wrap*, el gráfico va a descomponerse en ocho, uno por cada emoción. Finalmente, la representación obtenida es la siguiente:

Gráfico 3.5. Clasificación de palabras según las emociones



Este gráfico muestra la clasificación de las palabras según las emociones predominantes. Se observa que el término “arte” se cataloga como el más común dentro de las emociones alegría, anticipación, sorpresa y tristeza con 56 repeticiones. En cambio, en la emoción que expresa confianza las más destacada es “arquitectura” con 80 apariciones. La lectura que se puede hacer a la vista de estos resultados es la buena impresión que han manifestado los turistas al contemplar la Catedral de León. Esto se puede apreciar en los términos “bonita”, “interior”, “patrimonio”, “maravilla”, “majestuosa” y “espectacular”, palabras clasificadas dentro de las categorías alegría, anticipación, confianza y sorpresa, respectivamente. “Fiesta” y “feliz” encasilladas en estas mismas emociones, también son comúnmente utilizadas y pueden expresar el buen ambiente que se respira en la ciudad y el bienestar que les ha reportado el viaje.

5. ANÁLISIS ESTADÍSTICO EN RSTUDIO (MTE)

5.1. Análisis de frecuencias de términos (MTE)

Con el fin de poder realizar comparaciones entre los datos resultantes del análisis efectuado con los comentarios que implican una experiencia turística ordinaria y los datos de aquellos que contienen una experiencia turística memorable, se va a realizar un nuevo Análisis de Sentimiento con estos últimos comentarios.

El análisis va a estar desarrollado de igual forma que el mostrado en el punto 3, es decir los paquetes y los comandos empleados son los mismos, exceptuando alguna pequeña modificación que se indicará en cada caso. La única gran diferencia va a ser la composición de la base de datos.

Para poder empezar con el nuevo análisis, el primer paso es preparar la base de datos. Como se ha dicho, los comentarios objeto de estudio son aquellos en los cuales está presente la memorabilidad, por lo tanto, hay que filtrar la base de datos inicial, en concreto la columna en la que se encuentran los posts, para eliminar aquellos que no sean necesarios.

Como muestra Bigne et al. (2020), las experiencias turísticas memorables pueden ser identificadas mediante palabras. Por lo tanto, para filtrar el archivo hay que seleccionar la columna deseada e introducir los términos oportunos para que aparezcan los comentarios, en este caso, las palabras clave y frases introducidas han sido: “memorable”, “memori”, “destacable”, “recue”, “recor”, “olvid”, “en la memoria”, “no olvidar”, “nunca olvidar”, “para ser recordado”, “experiencia duradera”, “digno de recordar”, “recordado” y “durader”. Algunas palabras se han introducido cortadas porque así los resultados abarcan un mayor número de posts. De todos los términos citados algunos han arrojado resultados y otros no. No obstante, de los 2825 comentarios existentes en la base de datos original, 117 han sido identificados como posts que implican memorabilidad, es decir, experiencias de turismo memorables.

Con la base de datos filtrada se procede a realizar el análisis. Como la mecánica es la misma que la llevada a cabo anteriormente se mostrarán los resultados directamente.

En primer lugar, para conseguir una nube de palabras adecuada se han modificado algunos aspectos del comando, siendo el definitivo el siguiente:

```
wordcloud(Corpus, scale = c(2, 0.1), min.freq = 5,
max.words = 15, colors = brewer.pal(n = 4 , name ="Dark2"))
```

Dada la disminución considerable de comentarios y, por ende, de palabras, se ha decidido rebajar el número de términos máximos para ser representados, siendo el nuevo límite 15. De esta forma, se facilita la lectura de la figura resultante.

Figura 3.3. Nube de palabras (MTE)



Como se puede apreciar, la palabra más repetida es “recuerda”. Aunque los términos “gracias”, “recomendamos” y “amigas” también tienen una alta frecuencia. En cambio, las menos populares entre los turistas son “mundo”, “gótico” y “plaza” entre otras.

Figura 3.1. Nube de palabras



A diferencia de la Figura 3.1. en la cual los términos más frecuentes hacían referencia a las características del propio monumento, en esta nueva imagen hay palabras que son

claros indicadores de la existencia de experiencias turísticas memorables por parte de los turistas, como por ejemplo “recuerda”, “recomendamos” y “gracias”.

Los resultados que arroja la nueva TDM o matriz término-documento se recogen en la siguiente tabla:

Tabla 3.5. Matriz término-documento (MTE)

TÉRMINOS	DOCUMENTOS									
	103	108	110	25	34	35	53	58	7	99
Amigas	0	0	0	0	0	0	0	0	0	0
Congratulations	0	0	0	0	0	0	0	0	0	0
Gracias	0	0	0	0	0	0	0	0	0	0
Internet	0	0	0	0	0	0	0	0	0	0
Lugar	0	0	0	0	0	0	0	0	1	0
Plaza	0	0	0	0	0	0	0	0	0	0
Recomendamos	0	0	0	0	0	0	0	0	0	0
Recuerda	0	0	0	0	1	0	0	0	0	0
Recuerdos	0	0	0	0	0	0	0	0	2	0
Santa	0	1	0	1	1	1	0	1	0	0

Se puede ver que el término “recuerdos” aparece dos veces en el comentario número 7 pero ninguna vez en el 58. En relación con la TDM inicial, se observa cómo la frecuencia de aparición de las palabras es mayor que en esta nueva matriz, es decir, las palabras aparecían más veces en diferentes comentarios y dentro de cada uno la frecuencia también era mayor. Esto puede deberse a la diferencia entre el número de comentarios de ambas bases de datos, puesto que a mayor cantidad de posts más posibilidades hay de que las palabras se repitan.

Ahora bien, para extraer los términos de la TDM que más se repiten en todos los comentarios, se ha tenido que hacer una modificación en la frecuencia:

```
findFreqTerms(TDM, lowfreq = 10)
```

```
findFreqTerms(TDM, lowfreq = 5)
```

Originalmente la frecuencia estaba establecida en 50 y 30, pero dado la ausencia de resultados se ha optado por rebajarlo a 10 y 5, siendo los resultados los siguientes:

"amigas"	"gracias"	"recuerdos"
"recomendamos"		
"caminodesantiago"	"arte"	"amor"
"discover"	"familia"	

Estos términos aparecen en los comentarios más de 10 y 5 veces, respectivamente. A diferencia de los resultados salidos en el primer análisis, estos nuevos términos, principalmente, hacen referencia a los recuerdos, a la memoria y a las sensaciones causadas. Aunque también se pueden ver dos términos propios del análisis inicial como “caminodesantiago” o “arte”, los cuales no guardan relación con la memorabilidad de la experiencia.

Respecto a la TFM o matriz de frecuencias por término, los nuevos resultados son los que se muestran en la siguiente tabla:

Tabla 3.6. Matriz de frecuencias por término (MTE)

TÉRMINO	FRECUENCIA
Recuerda	18
Gracias	14
Amigas	13
Recomendamos	12
Santa	12
Recuerdos	10
Internet	9
Lugar	9
Plaza	9
Congratulations	8
Gótico	8
Conocer	7

En esta nueva matriz se puede ver que el término “recuerda” ha sido el más empleado por los turistas al hacer balance de su viaje alcanzando la cifra de 18 repeticiones, seguido por “gracias” y “amigas” con 14 y 13 respectivamente. En contraposición, las menos empleadas han sido “congratulations”, “gótico” y “conocer” que no alcanzan las 10 repeticiones.

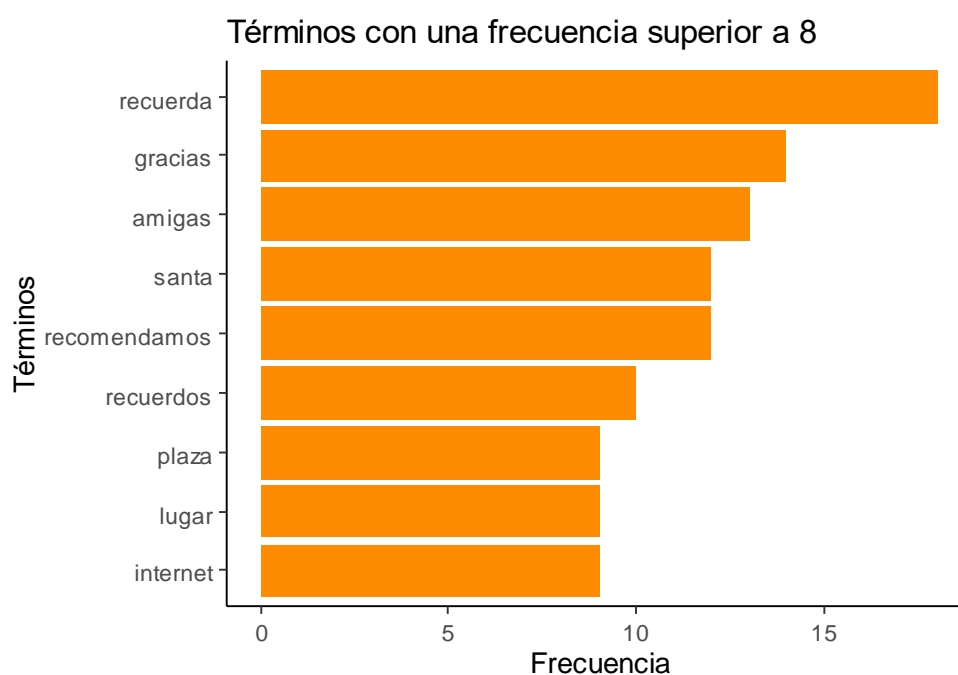
Si lo comparamos con la TFM original, la palabra más repetida era “camino” y “caminodesantiago”, términos referentes a los motivos de la visita a la capital, en cambio, con este nuevo análisis la palabra más utilizada apela al recuerdo que ha forjado el viaje, considerado un signo de memorabilidad.

Para obtener la nueva representación gráfica de esta TFM también se ha tenido que modificar el comando, siendo el código actual:

```
TFM %>% filter(Frecuencia > 8) %>% arrange(desc(Frecuencia))
%>% ggplot(mapping = aes(x= reorder(Término, Frecuencia),
y = Frecuencia)) + geom_col(fill = "Dark Orange") + coord_flip()
+ labs(title = "Términos con una frecuencia superior a 8",
x = "Términos",) + theme_classic()
```

El cambio se ha realizado en la frecuencia, en el gráfico actual se ha fijado en 8 y no en 50, dado que, al trabajar con muchos menos comentarios, ninguna palabra alcanza ese nivel de repetición.

Gráfico 3.6. Frecuencia de los términos (MTE)



Exceptuando algunas palabras, el resto son claros indicadores de memorabilidad. En primer lugar, “recuerda” y “recuerdos”, enumeradas de mayor a menor frecuencia de aparición, siendo “recuerda” la más empleada con 18 repeticiones y “recuerdos” la que menos con tan solo 10, son términos que están totalmente relacionados con la memoria del turista, así como con los recuerdos del viaje en cuestión. Como se ha visto en el Capítulo I de este trabajo, una experiencia turística se cataloga como memorable cuando es recordada en el tiempo una vez vivida, y dados los términos más descritos por los visitantes, se puede afirmar que este viaje en particular lo recuerdan y lo recordarán. En segundo lugar, “recomendamos”, con una frecuencia de repetición de 12, también deja entrever la existencia de memorabilidad, ya que un viaje es definido como memorable cuando el turista recomienda ese destino a sus amigos o familiares, lo que se ha descrito en capítulos anteriores como WOM. Por estos motivos, a la vista de este término estamos ante experiencias turísticas memorables. Y, en tercer lugar, la palabra “gracias”, la segunda más empleada, también podría indicarnos algún signo de memorabilidad, ya que es evidente que el turista se siente agradecido por el trascurso del viaje lo que fomenta su satisfacción, siendo esto un signo más de experiencias de turismo memorables.

Por otra parte, “amigas”, “santa” y “plaza” son expresiones referentes a las características del viaje. El primero haciendo referencia a la compañía y las restantes a una de las cualidades del monumento en cuestión.

Si comparamos este gráfico con el Gráfico 3.1. inicial, se observa cómo el primero de ellos representa términos que indican las características de la ciudad y de la catedral que le llaman la atención al turista, como “arquitectura” o “arte”. En cambio, en este nuevo gráfico, se ve cómo el visitante otorga más importancia a su memoria y a todos los recuerdos de la experiencia. Dicho esto, se puede apreciar la diferencia entre una experiencia turística ordinaria (primer análisis) y una experiencia turística memorable (segundo análisis).

5.2. Análisis de Sentimiento (MTE)

En este apartado se va a realizar el análisis de comentarios a través del Análisis de Sentimiento y se realizarán las comparaciones necesarias con el primer análisis.

En primer lugar, la tabla resultante referente a la clasificación de los comentarios según las emociones es la mostrada a continuación:

Tabla 3.7. Clasificación de comentarios según las emociones (MTE)

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Negative	Positive
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	1	0	0
5	0	1	0	0	1	1	1	0	1	3
6	0	0	0	0	0	0	0	0	0	0
7	2	5	0	0	8	5	3	7	6	16
8	0	0	0	0	0	0	0	0	1	1
9	0	0	0	0	1	0	1	1	0	1
10	0	0	0	0	0	0	0	0	0	0

Se ve cómo todos los comentarios mostrados (hay que recordar que solo se ha representado una pequeña parte de toda la tabla general) se clasifican en una o varias emociones. Por ejemplo, el comentario 7 tiene dos palabras que expresan enfado, cinco relacionadas con la anticipación, ocho más que manifiestan alegría, cinco tristeza, tres sorpresa, siete confianza y, finalmente, seis palabras con connotación negativa y dieciséis con positiva.

En la mayoría de los comentarios se han detectado palabras que expresan emociones relacionadas con lo positivo y otras con todo lo contrario, pero dentro del mismo post. Esto puede deberse al uso de palabras catalogadas de entrada como negativas pero que, dentro del contexto, pueden significar todo lo contrario. Un ejemplo puede ser usar la palabra “pena” para referirse a que da pena irse del lugar porque estás muy cómodo, y en este caso, “pena” es una palabra negativa que será clasificada como tal, pero realmente con su uso se quiere manifestar algo bueno. Esto puede verse reflejado en el comentario 8, por ejemplo.

Sintetizando toda la información de la tabla anterior en un *data.frame* se obtiene el siguiente resultado:

Tabla 3.8. Suma total de emociones del texto (MTE)

EMOCIÓN	VALOR
Positive	216
Negative	111
Trust	103
Joy	83
Anticipation	74
Sadness	64
Fear	49
Surprise	44
Anger	34
Disgust	25

Se puede ver que la emoción que más predomina en este segundo análisis es la confianza (*trust*) y la alegría (*joy*) con 103 y 83 palabras respectivamente. Por el contrario, las menos populares son asco (*disgust*) y enfado (*anger*) con 25 y 34 palabras cada una. Con estos datos se verifica que los comentarios analizados manifiestan mayoritariamente emociones placenteras.

En cuanto a los sentimientos, se han encontrado un total de 216 palabras que expresan sentimientos positivos y 111 que están relacionadas con aspectos negativos. Por lo tanto, los comentarios, en este caso, manifiestan un mayor volumen de sentimientos con polaridad positiva.

Comparando la Tabla 3.4 del análisis inicial con esta nueva tabla, no se observan diferencias. En ambos análisis las dos emociones predominantes son la confianza (*trust*) y la alegría (*joy*), detectándose en ambas una clara polaridad positiva.

Las representaciones oportunas para exponer gráficamente esta tabla son:

Gráfico 3.7. Análisis de la polaridad de sentimientos (MTE)

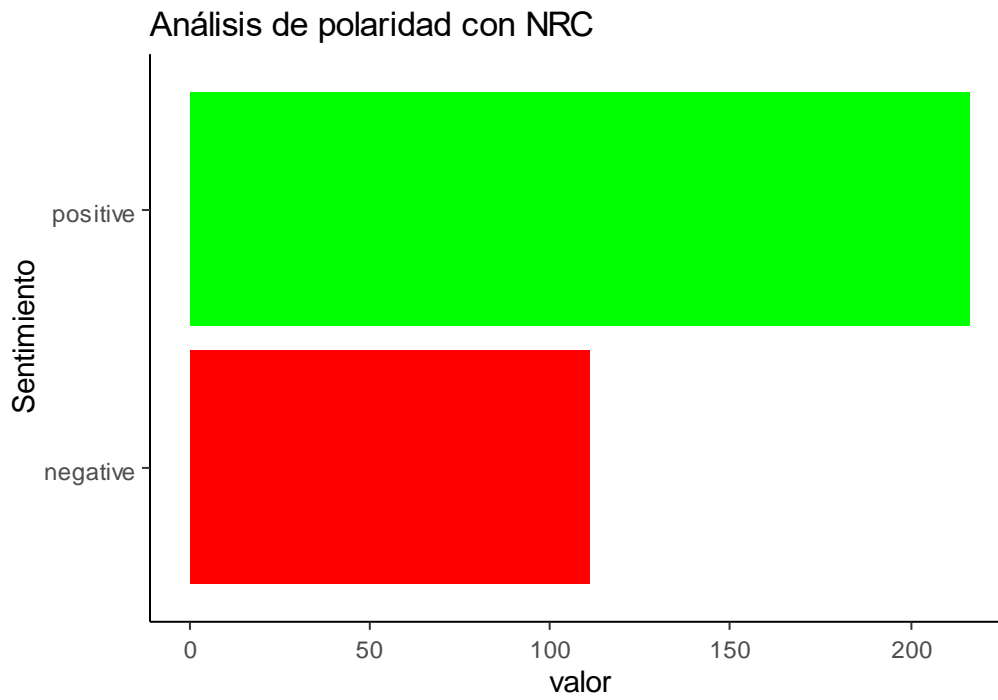
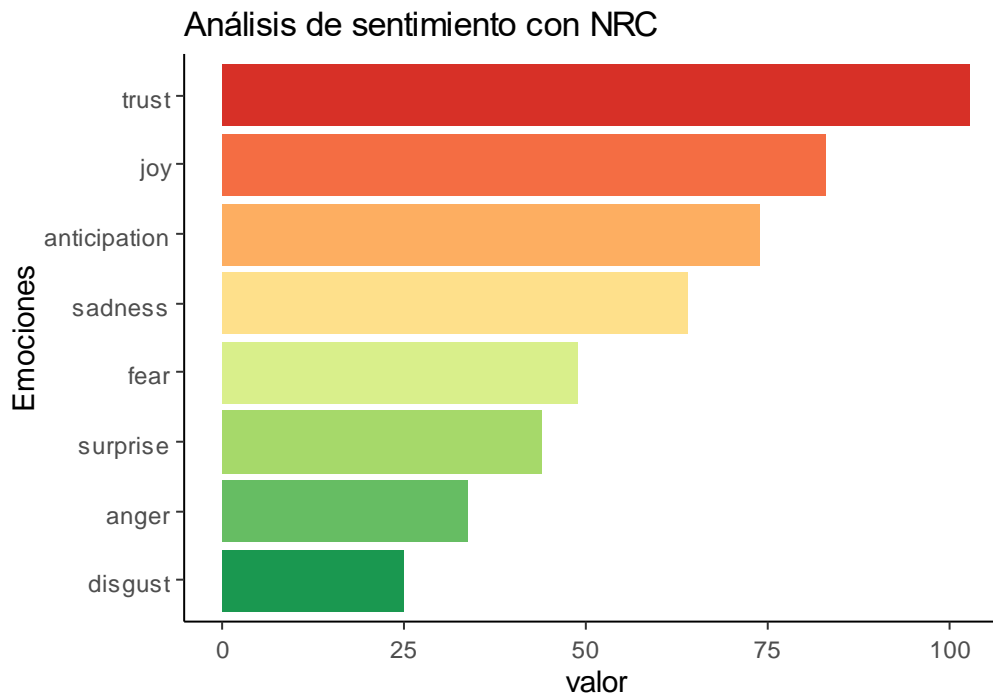
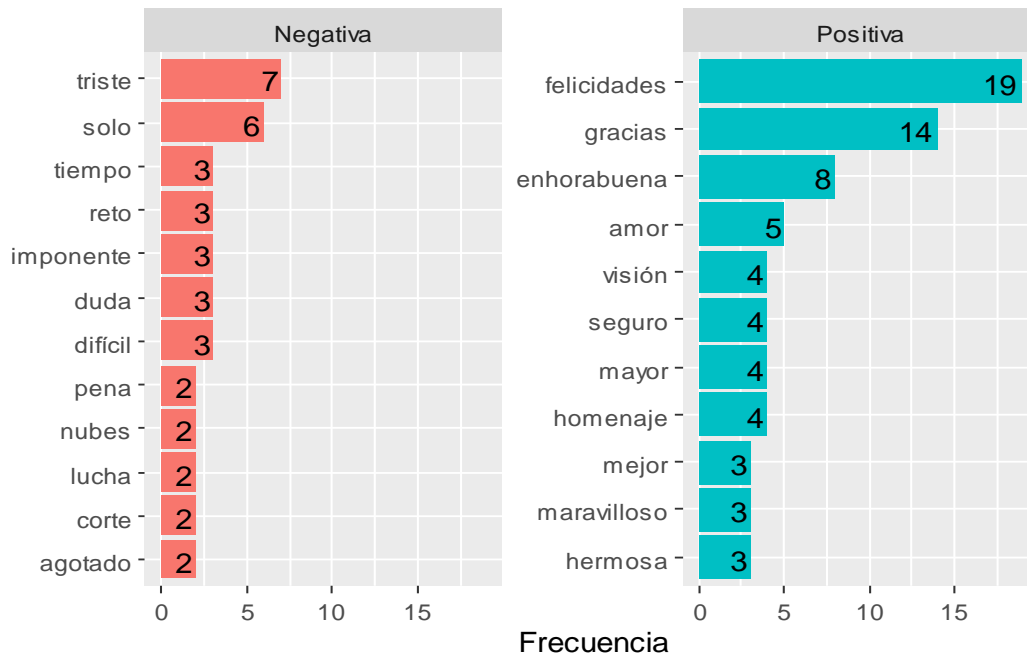


Gráfico 3.8. Análisis de Sentimiento (MTE)



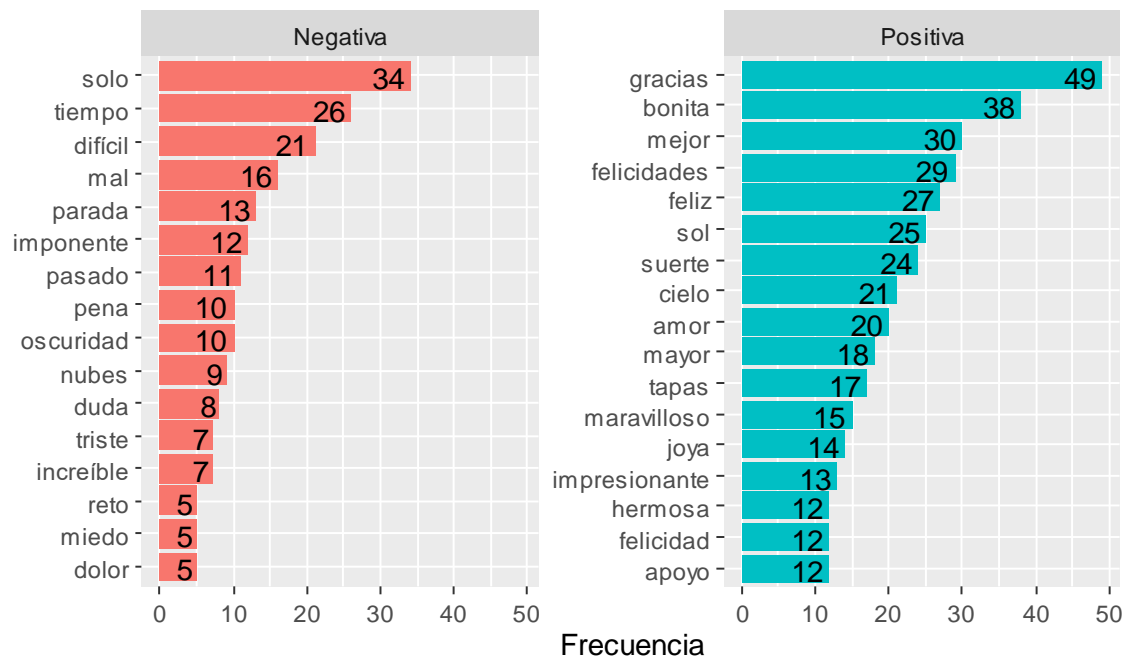
La clasificación de las palabras según su sentimiento ha resultado de la siguiente manera:

Gráfico 3.9. Clasificación de palabras según el sentimiento (MTE)



A la vista de esta representación gráfica, se puede observar que las palabras negativas más comunes son “triste” con 7 repeticiones y “solo” con 6, mientras que algunas de las menos empleadas son “agotado”, “corte” y “nubes” con 2 apariciones cada una.

Por su parte, las palabras positivas más utilizadas son “felicidades” y “gracias” con 19 y 14 repeticiones respectivamente. Por el contrario, con tan solo 3 apariciones se encuentran “hermosa”, “maravilloso” y “mejor”. Se puede apreciar que las dos palabras más recurridas están relacionadas con la memorabilidad, ya que ambas expresan sentimientos de satisfacción con la visita al lugar.

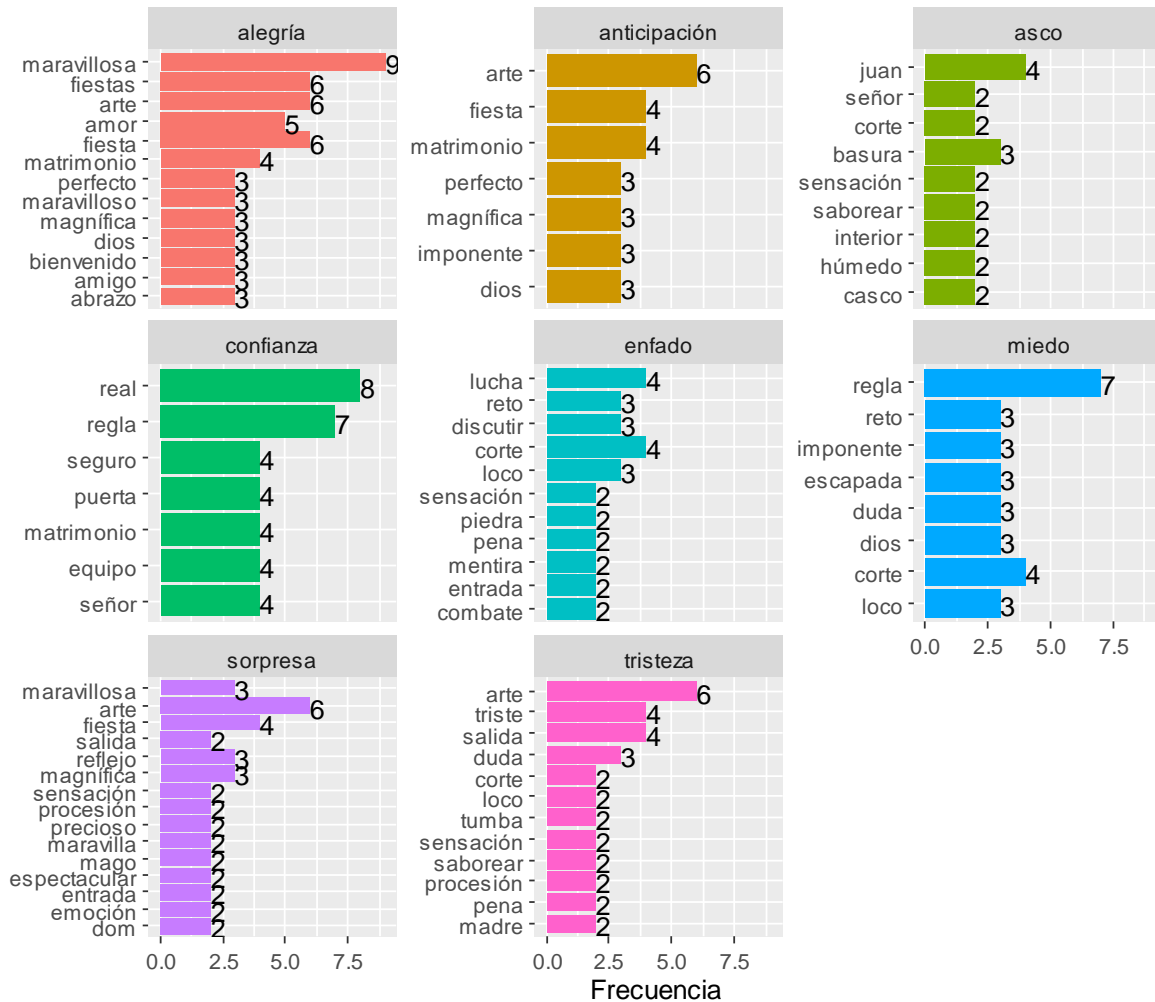
Gráfico 3.4. Clasificación de palabras según el sentimiento

Si se realiza una comparación con el Gráfico 3.4. inicial, se puede ver que la clasificación acorde al sentimiento negativo está compuesta, en su mayoría, por las mismas palabras. “Agotado” es uno de los nuevos términos que aparecen pudiendo hacer referencia a la inexistencia de entradas o reservas a los diferentes monumentos o restaurantes. En este caso, ambas representaciones siguen la misma línea, con escasas diferencias.

Respecto a la clasificación de palabras positivas, algunas de las nuevas que aparecen respecto al gráfico original son “mayor”, que puede estar relacionada con la Plaza Mayor de León; “seguro”, término que nos dice que los turistas se han sentido seguros durante su estancia; y “enhorabuena”, palabra que indica la buena sensación que le ha reportado al visitante el viaje. También se puede apreciar cómo en la nueva clasificación las palabras con connotaciones memorables están mejor posicionadas que en la clasificación inicial, concluyendo, una vez más, la existencia de experiencias turísticas memorables.

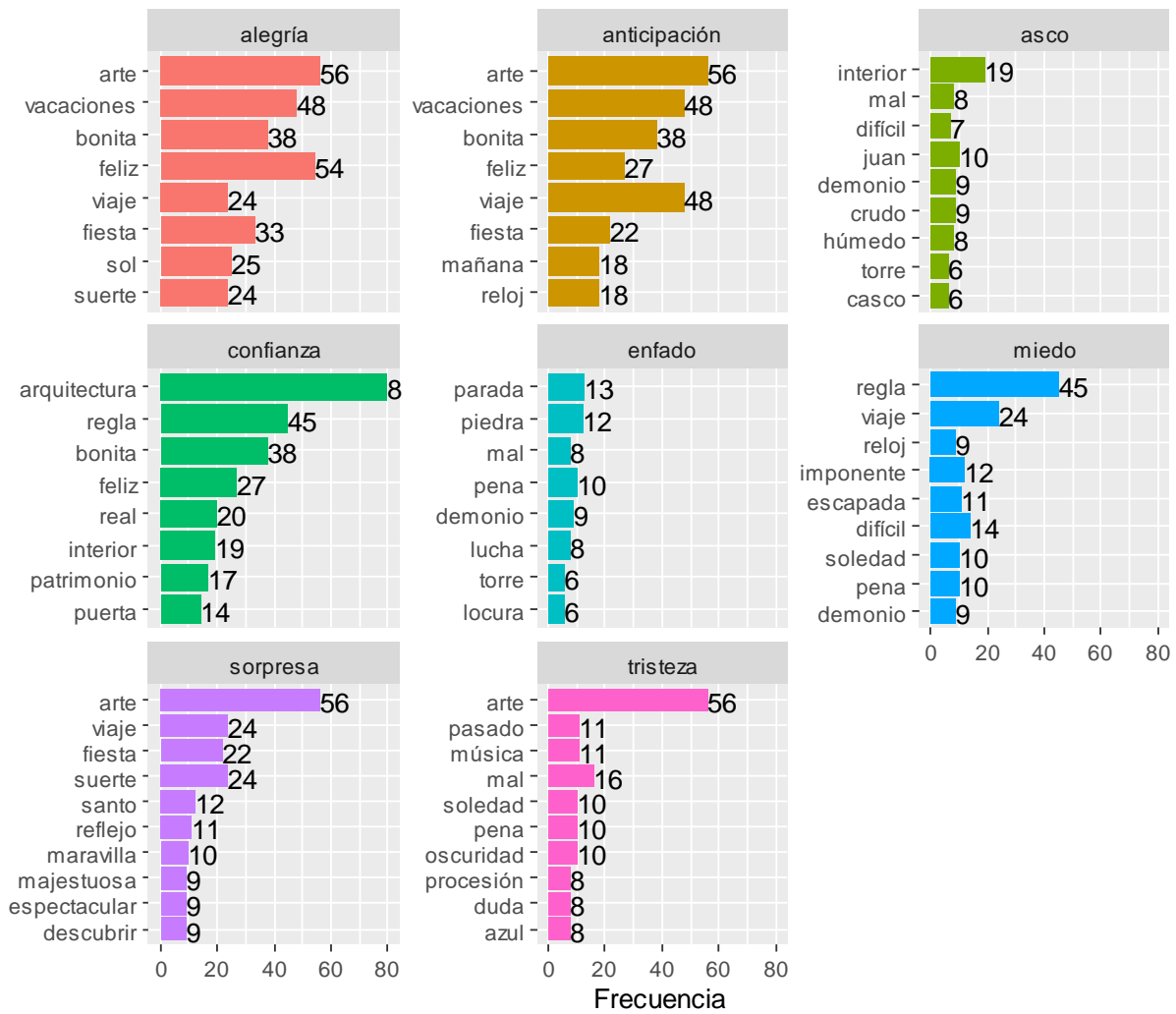
En relación con la clasificación de palabras según las emociones, se ha obtenido la siguiente representación gráfica:

Gráfico 3.10. Clasificación de palabras según las emociones (MTE)



A la vista de este gráfico se aprecia que el término “maravillosa” se encuentra en primer lugar en las emociones que indican alegría y sorpresa alcanzando 9 y 3 apariciones en cada una. Por otra parte, la palabra más empleada en anticipación y tristeza es “arte” con 6 repeticiones. “Entrada”, en la emoción de enfado y sorpresa, es una palabra que puede transmitir el rechazo que sienten los visitantes ante la posible escasez de entradas para contemplar el monumento. Respecto a la emoción de sorpresa hay dos términos interesantes que implican experiencias memorables y estos son: “sensación” y “emoción”. Esto puede ser un indicador de que la visita ha calado en los sentimientos y las emociones de los turistas, lo que la convertirá en una experiencia difícil de olvidar.

Gráfico 3.5. Clasificación de palabras según las emociones



Si se realiza una comparación con el Gráfico 3.5 inicial, no se observan diferencias significativas. En ambas representaciones se encuentran palabras similares catalogadas en las mismas emociones, haciendo todas ellas referencia al bienestar y la satisfacción de los turistas después de la visita. Un ejemplo puede ser “maravilloso” y “magnífica” en la emoción de alegría, “precioso” y “espectacular” en la emoción que indica sorpresa, “perfecto” en anticipación. Por lo tanto, en este caso, ambos gráficos se encuentran en la misma línea. La única diferencia notable puede encontrarse en los términos “emoción” y “sensación” propios de la emoción referente a la sorpresa, ya que en la nueva clasificación sí que aparecen representadas mientras que, en el gráfico original no, siendo esto un indicador de memorabilidad.

A continuación se va a insertar una tabla comparativa donde se pueden observar las diferencias entre el número de palabras que hacen referencia a los sentimientos y a las emociones, con sus correspondientes proporciones, en las experiencias turísticas ordinarias y memorables.

Tabla 3.9. Comparativa entre experiencias ordinarias y memorables

	ORDINARIO	MEMORABLE	% del total	
Post totales	2.825	117	4%	
Palabras encontradas				
	ORDINARIO	%	MEMORABLE	%
Palabras positivas	1442	73%	216	66%
Palabras negativas	533	27%	111	34%
Total palabras encontradas	1975	100%	327	100%
Emociones encontradas				
Emociones de enfado	165	5%	34	7%
Emociones de anticipación	631	19%	74	16%
Emociones de asco	124	4%	25	5%
Emociones de miedo	320	10%	49	10%
Emociones de alegría	649	20%	83	17%
Emociones de tristeza	308	9%	64	13%
Emociones de sorpresa	340	10%	44	9%
Emociones de confianza	707	22%	103	22%
Total emociones encontradas	3244	100%	476	100%

Se puede observar cómo existe mayor proporción de palabras positivas en las experiencias ordinarias, alcanzando un 73%, que en las experiencias memorables con un 66%. Por el contrario, el porcentaje de palabras negativas es mayor en las experiencias memorables, con un 34%, que en las ordinarias, llegando al 27%. Respecto a las proporciones calculadas en las emociones de ambas experiencias, hay ligeras diferencias. Se puede observar que la proporción de emociones negativas es ligeramente superior en las experiencias que implican memorabilidad que en las ordinarias. Esto se debe a que la memorabilidad puede ser tanto negativa como positiva, es decir las

experiencias pueden ser recordadas también por lo malas que resultaron, guardando un recuerdo con connotaciones negativas. En esta investigación, solamente se ha estudiado la parte positiva, al igual que el estudio de Rasoolimanesh et al. (2021).

CONCLUSIONES

Gracias a la revisión bibliográfica llevada a cabo en el presente trabajo, así como a la investigación empírica desarrollada, se han podido extraer una serie de conclusiones determinantes.

Como se ha podido ver en el marco teórico, el turismo representa un eje muy importante en la economía del país, y es por este motivo que se está adaptando y reinventando a las nuevas necesidades que demandan los turistas. Cumplir con las expectativas que estos tienen preestablecidas antes de empezar el viaje, es un punto importante para alcanzar porque se verá cumplido su deseo y la satisfacción será elevada. Además, los destinos y empresas turísticas tienen que sacarles el máximo partido a los servicios turísticos, actividades o eventos que ponen a disposición de los visitantes, haciendo que se involucren, interactúen de forma directa y aprendan sobre la cultura y las costumbres del lugar que visiten. Cumpliendo estos principios básicos el consumidor pasará de vivir una experiencia turística ordinaria, como cualquier otra, a experimentar una experiencia turística memorable que va a perdurar en su recuerdo y memoria una vez transcurrido el viaje.

Al destino como motor turístico le va a suponer un gran beneficio la existencia de este tipo de experiencias por varios motivos. En primer lugar, el turista la compartirá con su entorno cercano haciendo que aumenten las probabilidades de recibir nuevos visitantes. En segundo lugar, al cumplir con sus expectativas es posible que repita la experiencia en el futuro. Y, finalmente, lograr que alcance un gran nivel de satisfacción despertará ciertas emociones positivas en él, lo que fomentará una buena imagen del destino para ese turista en concreto. Por lo tanto, se puede concluir, que fomentar la experiencia de turismo memorable reportará beneficios tanto al turista como al destino.

Este tipo de experiencias se encuentran muy ligadas a las redes sociales. Debido a la importancia que estas han recibido en los últimos años, son consideradas una fuente valiosa de información, donde los turistas narran y dejan constancia de sus viajes y experiencias para que otros usuarios se nutran de ellas. De todas las redes, Instagram es una de las más populares dada la gran diversidad de funciones que posee para poder documentar los viajes. De este modo, las empresas y los destinos de turismo deben estar activos en las redes sociales para conocer de primera mano todo lo que se está gestando entorno a ellos, pero también entorno a su competencia. Una excelente forma para

conseguirlo y para aprovechar estos datos es mediante la **realización de un Análisis de Sentimiento mediante UGC en Instagram**, dado, además, el papel fundamental que juegan las emociones en las personas.

Respecto a la parte práctica se han podido extrapolar varias conclusiones. En primer lugar, se han **conseguido conocer los sentimientos y las emociones** más representativas de los turistas ante un monumento cultural como es la Catedral de León. Tanto si se habla de experiencias ordinarias como si se habla de experiencias memorables los sentimientos destacan por la **polaridad ampliamente positiva**, siendo las **emociones predominantes aquellas que implican alegría y confianza** con el monumento. Así mismo, de todos los **términos empleados por los viajeros** cuando visitan este lugar, y que han sido analizados, tienen un nivel de frecuencia mayor **“caminodesantiago”, “arquitectura”, “gotico” y “vidrieras”**, respecto a las experiencias ordinarias. No obstante, los términos son diferentes si analizamos las experiencias memorables, ya que **“recuerda” y “gracias”** son los que más se repiten.

Se ha podido afirmar la existencia de experiencias de turismo memorables por parte de los turistas, destacando la alta satisfacción y el cumplimiento de sus expectativas, lo que hará que en un futuro recuerden la experiencia y se vea incrementada la memorabilidad del viaje. Esto se puede ver reflejado en palabras como **“recuerda”, “gracias”, “enhorabuena”, “recomendamos” y “recuerdos”**, las cuales son las más empleadas por los turistas para describir las vivencias de su viaje.

Toda esta información ha sido obtenida gracias a la red social **Instagram**, por lo que se puede **afirmar que es una buena plataforma para realizar este tipo de análisis** por la amplia variedad de datos.

Los **datos extraídos en el presente trabajo pueden ser utilizados para generar una ventaja competitiva**. Una vez que se conoce aquello que más llama la atención por parte de los visitantes, puede ser incluido, posteriormente, en la estrategia de marketing. Por ejemplo, se pueden destacar cualidades como la arquitectura, el estilo artístico, las vidrieras de la catedral, el arte que desprende, etc. De esta forma, la información recabada supondrá una ventaja y una nueva manera de promocionarse.

A lo largo del trabajo se ha comentado que, de todos los post analizados, 415 no contenían ningún tipo de comentario, lo que supone el 15% de la muestra. Por

consiguiente, sería recomendable analizar también otro tipo de contenido que aporta Instagram, como son las fotografías e imágenes. Además, 450 posts únicamente contienen *emojis* o *hashtags* y 1552 emplean al menos una de estas etiquetas, por lo tanto se puede concluir que los *hashtags* son muy utilizados por los usuarios para comunicarse en las redes sociales y suponen una vía muy importante de promoción para las empresas y los destinos. De este modo, sería conveniente que el destino crease un *hashtag* propio para poder agrupar todo el contenido que generan los usuarios entorno a ellos y poder ver qué y cómo se habla. Un buen ejemplo de esto es Castilla y León bajo la etiqueta #cylesvida o Madrid que ha creado #MeQuedoContigoMadrid. León como destino de turismo debe reforzar su estrategia y crear y fomentar el uso de estas herramientas. Por otra parte, en este trabajo se ha empleado únicamente la escala desarrollada por Kim, analizando solo las MTE positivas. Por consiguiente, sería oportuno que en futuras investigaciones se realicen estudios que también tengan en cuenta las dimensiones negativas de las experiencias turísticas.

Una vez realizado este estudio, se puede concluir que el **Análisis de Sentimiento sobre el contenido de Instagram ha permitido escuchar y analizar a los turistas** obteniendo con ello conclusiones relevantes para los gestores turísticos.

Algunas de las limitaciones que se han encontrado a la hora de desarrollar el estudio son, en primer lugar, el complejo manejo del programa estadístico RStudio. Y en segundo lugar, la posible clasificación de algunas palabras en una determinada emoción o sentimiento pero que, al leerse dentro del contexto, no existe relación entre la emoción en la que se ha clasificado y el significado que expresa dicho término. Por este motivo, sería conveniente realizar de nuevo el estudio con léxicos específicos compuestos por palabras referentes a la temática en cuestión.

REFERENCIAS

- Alonso, M., y Sánchez, J. (2020, enero 15). Diseño de experiencias turísticas. *Otra Piel*.
<https://www.otrapieldesign.com/post/experiencias-turisticas>
- Bazzaz Abkenar, S., Haghi Kashani, M., Mahdipour, E., y Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics*, 57(September 2020), 101517.
<https://doi.org/10.1016/j.tele.2020.101517>
- Bigne, E., Fuentes-Medina, M. L., y Morini-Marrero, S. (2020). Memorable tourist experiences versus ordinary tourist experiences analysed through user-generated content. *Journal of Hospitality and Tourism Management*, 45(April), 309-318.
<https://doi.org/10.1016/j.jhtm.2020.08.019>
- Biosphere Tourism. (s. f.). *Certificación de sostenibilidad / Biosphere Responsible Tourism*. Recuperado 17 de febrero de 2022, de
<https://www.biospheretourism.com/es>
- Bueno, I., Carrasco, R. A., Ureña, R., y Herrera-Viedma, E. (2022). A business context aware decision-making approach for selecting the most appropriate sentiment analysis technique in e-marketing situations. *Information Sciences*, 589, 300-320.
<https://doi.org/10.1016/j.ins.2021.12.080>
- Cambria, E., Grassi, M., Hussain, A., y Havasi, C. (2012). Sentic Computing for social media marketing. *Multimedia Tools and Applications*, 59(2), 557-577.
<https://doi.org/10.1007/s11042-011-0815-0>
- Coelho, M. de F., Gosling, M. de S., y Almeida, A. S. A. de. (2018). Tourism experiences: Core processes of memorable trips. *Journal of Hospitality and Tourism Management*, 37(May 2017), 11-22.
<https://doi.org/10.1016/j.jhtm.2018.08.004>
- CRAN.R. (2014, diciembre). *CRAN - Package RColorBrewer*. <https://cran.r-project.org/web/packages/RColorBrewer/index.html>
- CRAN.R. (2018, agosto). *CRAN - Package Wordcloud*. <https://cran.r-project.org/web/packages/wordcloud/index.html>

- CRAN.R. (2021). *CRAN - Package tidytext*. <https://cran.r-project.org/web/packages/tidytext/index.html>
- Dataestur. (2021a). *Encuesta de turismo de residentes (ETR) - DATAESTUR*. <https://www.dataestur.es/general/etr/>
- Dataestur. (2021b). *Movimientos turísticos en fronteras (FRONTUR) - DATAESTUR*. <https://www.dataestur.es/general/frontur/>
- Destino Turístico Inteligente. (2018). *León (Castilla y León) - Red de Destinos Turísticos Inteligentes*. <https://www.destinosinteligentes.es/destinos/leon-castilla-y-leon/>
- Europa Press. (2021). *El 80% de españoles usa las redes sociales, con 37,4 millones de usuarios activos*. <https://www.europapress.es/portaltic/socialmedia/noticia-80-espanoles-usa-redes-sociales-374-millones-usuarios-activos-20210210070041.html>
- Fatanti, M. N., y Suyadnya, I. W. (2015). Beyond User Gaze: How Instagram Creates Tourism Destination Brand? *Procedia - Social and Behavioral Sciences*, 211(September), 1089-1095. <https://doi.org/10.1016/j.sbspro.2015.11.145>
- Feinerer, I., Hornik, K., y Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5), 1-54. <https://doi.org/10.18637/jss.v025.i05>
- Flores-Ruiz, D., Elizondo-Salto, A., y Barroso-González, M. D. L. O. (2021). Using social media in tourist sentiment analysis: A case study of andalusia during the Covid-19 pandemic. *Sustainability (Switzerland)*, 13(7), 1-19. <https://doi.org/10.3390/su13073836>
- Francés, J. (2020, abril 7). *Instagram TV (IGTV): qué es y cómo implementarlo en tu estrategia de marketing*. InboundCycle. <https://www.inboundcycle.com/blog-de-inbound-marketing/instagram-tv-igtv-que-es-como-implementarlo-estrategia-de-marketing>
- Fusté Forné, F., y Nava Jiménez, C. (2015). La Experiencia Turística y su Crítica Intercultural. *Revista Turismo em Análise*, 26(4), 843-858. <https://doi.org/10.11606/issn.1984-4867.v26i4p843-858>

- Gálvez-Rodríguez, M. del M., Alonso-Cañadas, J., Haro-de-Rosario, A., y Caba-Pérez, C. (2020). Exploring best practices for online engagement via Facebook with local destination management organisations (DMOs) in Europe: A longitudinal analysis. *Tourism Management Perspectives*, 34(February 2019), 100636. <https://doi.org/10.1016/j.tmp.2020.100636>
- Gascón, J. F. F., Ormaechea, S. L., y Aragón, Ó. L. G. (2019). Análisis semántico en medios sociales para la comunicación digital turística. *Comunicació: revista de recerca i d'anàlisi*, 0(0), 71-94. <https://doi.org/10.2436/20.3008.01.179>
- Ginzarly, M., y Jordan Srour, F. (2021). Cultural heritage through the lens of COVID-19. *Poetics*, February, 101622. <https://doi.org/10.1016/j.poetic.2021.101622>
- Gon, M. (2021). Local experiences on Instagram: Social media data as source of evidence for experience design. *Journal of Destination Marketing and Management*, 19(February 2020), 100435. <https://doi.org/10.1016/j.jdmm.2020.100435>
- Gonçalves Silveira, F., y Barretto, M. (2010). Turismo de transformación. Ecoturismo en reservas de la biosfera dentro del marco de la economía de la experiencia. *Estudios y perspectivas en turismo*, 19(2), 315-329. <https://dialnet.unirioja.es/servlet/articulo?codigo=3351496>
- Hera, C. de la. (2021, junio 30). *Historia de las Redes Sociales: cómo nacieron y cuál fue su evolución* -. Marketing 4 Ecommerce. <https://marketing4ecommerce.net/historia-de-las-redes-sociales-evolucion/>
- HuffPost. (2018, febrero 15). *Análisis del sentimiento sobre turismo en redes sociales* / *El HuffPost*. https://www.huffingtonpost.es/instituto-de-ingenieria-del-conocimiento/analisis-del-sentimiento-sobre-turismo-en-redes-sociales_a_23357222/
- Instagram. (2021a). *Comprar en Instagram | Descubre y compra productos de las marcas y los creadores que te encantan*. <https://about.instagram.com/es-la/features/shopping>

- Instagram. (2021b). *Direct Messenger en Instagram | Usa videochat, comparte contenido y envía mensajes directos*. <https://about.instagram.com/es-la/features/direct>
- Instagram. (2021c). *Instagram | Cómo usar Instagram Stories | Sitio oficial*. <https://about.instagram.com/es-la/features/stories>
- Instagram. (2021d). *Instagram Reels - Crea y comparte videos cortos | Sitio oficial*. <https://about.instagram.com/es-la/features/reels>
- Isasi, J. (2021, marzo 23). *Análisis de sentimientos en R con «syuzhet»*. Programming Historian en español; University of Sussex. <https://doi.org/10.46430/PHES0051>
- Jockers, M. (s. f.). *Syuzhet: extrae sentimientos y arcos de tramas derivados de sentimientos de texto*. Recuperado 15 de enero de 2022, de <https://rdr.io/cran/syuzhet/>
- Junta de Castilla y León. (2020). *Boletín de Coyuntura Turística de Castilla y León. Junio 2020*. <https://www.turismocastillayleon.com/es/espacio-profesionales/boletines-coyuntura>
- Karmegam, D., y Mappillairaju, B. (2022). Social media analytics and reachability evaluation - #Diabetes. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 16(1), 102359. <https://doi.org/10.1016/j.dsx.2021.102359>
- Kim, J. H. (2014). The antecedents of memorable tourism experiences: The development of a scale to measure the destination attributes associated with memorable experiences. *Tourism Management*, 44, 34-45. <https://doi.org/10.1016/j.tourman.2014.02.007>
- Lavagna, E. (2020, julio 23). ¿Qué es Instagram y para qué sirve esta red social? *Web Escuela*. <https://webescuela.com/que-es-instagram-para-que-sirve/>
- LetsLaw. (2021). *Libertad de expresión en redes sociales - LetsLaw*. <https://letslaw.es/libertad-de-expresion-en-redes-sociales/>

- Martinez-Torres, M. R., y Toral, S. L. (2019). A machine learning approach for the identification of the deceptive reviews in the hospitality sector using unique attributes and sentiment orientation. *Tourism Management*, 75(June), 393-403. <https://doi.org/10.1016/j.tourman.2019.06.003>
- Medhat, W., Hassan, A., y Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Melón, M. P. A., Herrera, C. F., y Sarasa, R. G. (2018). *Las experiencias turísticas memorables (MTE): Análisis de sus antecedentes y consecuencias* [Trabajo Fin de Máster]. Universidad de Zaragoza. Facultad de Economía y Empresa.
- Morgan, A., Wilk, V., Sibson, R., y Willson, G. (2021). Sport event and destination co-branding: Analysis of social media sentiment in an international, professional sport event crisis. *Tourism Management Perspectives*, 39(September 2020), 100848. <https://doi.org/10.1016/j.tmp.2021.100848>
- MOTT Marketing Digital. (2019). *Historia de Instagram : Servicios y características - MOTT.Marketing*. <https://mott.marketing/informacion-sobre-la-historia-servicios-caracteristicas-de-la-red-social-instagram/#comments>
- Panico, C., y Cesteros, A. F.-P. (2018). *La eficacia del análisis de sentimientos para la empresa: el caso de estudio Dell Technologies Inc.* [Trabajo Fin de Grado]. Universidad Complutense de Madrid. Facultad de Filología.
- Pauli, P. A., y Soliani, V. I. (2019). Análisis de sentimiento, comparación de algoritmos predictivos y métodos utilizando un lexicon español [Trabajo Final]. En *Instituto Tecnológico de Buenos Aires-ITBA*. Buenos Aires.
- Prada, J., Armijos, D., Torres, L., y Crespo, A. (2016). Ciudades Patrimoniales, Turismo Cultural Y Perfiles De Los Visitantes: Algunas Consideraciones a Partir Del Caso De Estudio De Cuenca (Ecuador). *Ingeba.Org*, 39, 199-216. <http://www.ingeba.org/lurralde/lurranet/lur39/39prada.pdf>
- Ranker. (2022). *Phantombuster: Herramienta de automatización para redes*. <https://rankerstudio.com/phantombuster-que-es/>

- Rasoolimanesh, S. M., Seyfi, S., Hall, C. M., y Hatamifar, P. (2021). Understanding memorable tourism experiences and behavioural intentions of heritage tourists. *Journal of Destination Marketing and Management*, 21(May), 100621. <https://doi.org/10.1016/j.jdmm.2021.100621>
- RDocumentation. (s. f.). *Corpus function - RDocumentation*. Recuperado 23 de enero de 2022, de <https://www.rdocumentation.org/packages/tm/versions/0.7-8/topics/Corpus>
- Real Academia Española. (2021). *atributo | Definición | Diccionario de la lengua española | RAE - ASALE*. <https://dle.rae.es/atributo>
- Revista Catedral de León. (2021). *Revista Catedral de León | RÉCORD HISTÓRICO DE TURISTAS QUE DESTACAN LA CATEDRAL Y LA REVISTA CATEDRAL DE LEÓN*. <https://www.revistacatedraldeleon.es/record-historico-turistas-leon-destacando-la-catedral-la-nueva-revista-catedral-leon/>
- Rodríguez, C. E. (2019, agosto). *Clasificación de documentos de texto con R*. <http://sigma.iimas.unam.mx/carloserwin/clasificacion.html>
- Romero, I. (2021, septiembre 15). *Guía Reels Instagram - Todo lo que necesitas saber*. Metricool. <https://metricool.com/es/reels-instagram/>
- RStudio. (s. f.). *¿Qué hace que RStudio sea diferente? - RStudio*. Recuperado 15 de enero de 2022, de <https://www.rstudio.com/about/what-makes-rstudio-different/>
- Santana, A., y Hernández, C. N. (2016). *Librerías en R*. Departamento de Matemáticas. <https://estadistica-dma.ulpgc.es/cursosR4ULPGC/5-librerias.html>
- Santana, A., y Hernández, C. N. (2020, enero). *Librerías en R: Tidyverse*. Departamento de Matemáticas. <https://estadistica-dma.ulpgc.es/cursosR4ULPGC/5b-Tidyverse.html>
- Savi Mondo, T., y Gonçalves Silveira, G. (2016). ATRIBUTOS DE LA CALIDAD DE SERVICIOS EN ATRACTIVOS TURÍSTICOS. Un estudio netnográfico en el uso del protocolo TOURQUAL. *Estudios y Perspectivas en Turismo*, 25(2), 124-144. <http://estudiosenturismo.com.ar/PDF/V25/N02/v25n2a02.pdf>

- Shen, Y., Kokkranikal, J., Christensen, C. P., y Morrison, A. M. (2021). Perceived importance of and satisfaction with marina attributes in sailing tourism experiences: A kano model approach. *Journal of Outdoor Recreation and Tourism*, 35, 100402. <https://doi.org/10.1016/j.jort.2021.100402>
- Snowballstem. (s. f.). *Snowball*. Recuperado 22 de enero de 2022, de <https://snowballstem.org/>
- Statista. (2021a). *Las redes sociales en España - Datos estadísticos* | Statista. <https://es.statista.com/temas/6566/las-redes-sociales-en-espana/#dossierKeyfigures>
- Statista. (2021b). *Redes sociales: penetración España 2010-2020* | Statista. <https://es.statista.com/estadisticas/474955/redes-sociales-tasa-penetracion-usuarios-prevision-espana/>
- Statista. (2021c). *Redes sociales con más usuarios España 2020* | Statista. <https://es.statista.com/estadisticas/489153/porcentaje-de-internautas-en-las-redes-sociales-en-espana/>
- Teles da Mota, V., Pickering, C., y Chauvenet, A. (2022). Popularity of Australian beaches: Insights from social media images for coastal management. *Ocean and Coastal Management*, 217(January), 106018. <https://doi.org/10.1016/j.ocecoaman.2021.106018>
- Tripadvisor. (2021). *Santa María de León Cathedral* - Tripadvisor. https://www.tripadvisor.es/Attraction_Review-g187492-d244192-Reviews-Santa_Maria_de_Leon_Cathedral-Leon_Province_of_Leon_Castile_and_Leon.html#REVIEWS
- Tung, V. W. S., y Ritchie, J. R. B. (2011). Exploring the essence of memorable tourism experiences. *Annals of Tourism Research*, 38(4), 1367-1386. <https://doi.org/10.1016/j.annals.2011.03.009>
- Valdivia, A., Martínez-Cámara, E., Chaturvedi, I., Luzón, M. V., Cambria, E., Ong, Y. S., y Herrera, F. (2018). What do people think about this monument? Understanding negative reviews via deep learning, clustering and descriptive rules. *Journal of Ambient Intelligence and Humanized Computing*, 11(1), 39-52. <https://doi.org/10.1007/s12652-018-1150-3>

Wong, J. W. C., y Lai, I. K. W. (2021). Gaming and non-gaming memorable tourism experiences: How do they influence young and mature tourists' behavioural intentions? *Journal of Destination Marketing and Management*, 21(July), 100642. <https://doi.org/10.1016/j.jdmm.2021.100642>

Yu, Q., Pickering, S., Geng, R., y Yen, D. A. (2021). Thanks for the memories: Exploring city tourism experiences via social media reviews. *Tourism Management Perspectives*, 40(June), 100851. <https://doi.org/10.1016/j.tmp.2021.100851>

ANEXOS

ANEXO I: CÓDIGO ANÁLISIS

Importamos las bases de datos

```
Excel1 <- read_excel("C:/Users/elena/Downloads/Excel1.xlsx",
sheet = "Sin dupli", range = "H1:H2826")
```

Instalamos paquetes, si fuera necesario

```
if (!require("tidyverse")) install.packages("tidyverse")
if (!require("tm")) install.packages("tm")
if (!require("SnowballC")) install.packages("SnowballC")
if (!require("wordcloud")) install.packages("wordcloud")
if (!require("RColorBrewer")) install.packages("RColorBrewer")
if (!require("syuzhet")) install.packages("syuzhet")
```

Cargamos los paquetes

```
library("tidyverse")
library("tm")
library("SnowballC")
library("wordcloud")
library("RColorBrewer")
library("syuzhet")
library("tidytext")
```

Se limpia el documento

```
Excel1$texto_limpio <- Excel1$caption %>% tolower() %>%
removePunctuation() %>% removewords(words = stopwords(kind
="spanish")) %>% removewords(words = stopwords(kind ="english")) %>%
removewords(words =c("buen" , "buenos" , "buenas" , "bueno" , "buena",
"follow" , "architecture" , "art" , "gothic" , "night" , "ciudad" ,
"city", "león", "leon" , "catedral" , "cathedral" , "spain" ,
"españa", "leonesp" , "mundo" , "días" , "catedraldeleon" ,
"igersleonesp" , "castilla y león" , "castillayleon" , "travel",
"estaesleon" , "semana" , "foto" , "fotos" , "estaescastillayleon",
"estaescastillaleon" , "hoy" , "día" , "dia" , "day" , "picoftheday" ,
"estaesespania" , "siempre" , "catedraldeleón" ,"igersspain" ,
"photography" , "igers" , "igersleon" , "catedrales" , "tag" ,
"photo", "photooftheday" , "travelgram" , "travelphotography" ,
"photos" , "okspain" , "gran" , "san" , "ser")) %>% stripwhitespace()
```

Se convierten los documentos en un corpus

```
Corpus <- vCorpus(VectorSource(Excel1$texto_limpio))
```

Representación en una nube de palabras (ambos análisis)

```
wordcloud(Corpus, scale = c(2, 0.1), min.freq = 5, max.words = 30,
  colors = brewer.pal(n = 4 , name = "Dark2"))
```

```
wordcloud(Corpus, scale = c(2, 0.1), min.freq = 5, max.words = 15,
  colors = brewer.pal(n = 4 , name = "Dark2"))
```

Se convierte el corpus en una TDM y se observan las características

```
TDM <- TermDocumentMatrix(Corpus)
```

```
inspect(TDM)
```

Se buscan los términos con más de 50 y 30, 10 y 5 apariciones (ambos análisis)

```
findFreqTerms(TDM, lowfreq = 50)
```

```
findFreqTerms(TDM, lowfreq = 30)
```

```
findFreqTerms(TDM, lowfreq = 10)
```

```
findFreqTerms(TDM, lowfreq = 5)
```

Conversión en una matriz

```
tdm <- as.matrix(TDM)
```

Se crea una matriz de frecuencias de términos en orden decreciente

```
TFM <- data.frame(Frecuencia = rowSums(tdm)) %>%
  rownames_to_column("Término") %>% arrange(desc(Frecuencia))
```

```
view(TFM)
```

Representación gráfica de frecuencias (ambos análisis)

```
TFM %>% filter(Frecuencia > 50) %>% arrange(desc(Frecuencia)) %>%
  ggplot(mapping = aes(x= reorder(Término, Frecuencia), y = Frecuencia))
  + geom_col(fill = "Dark Orange") + coord_flip() + labs(title =
  "Términos con una frecuencia superior a 50", x = "Términos",) +
  theme_classic()
```

```
TFM %>% filter(Frecuencia > 8) %>% arrange(desc(Frecuencia))%>%
  ggplot(mapping = aes(x= reorder(Término, Frecuencia), y = Frecuencia))
  + geom_col(fill = "Dark Orange") + coord_flip() + labs(title =
  "Términos con una frecuencia superior a 8", x = "Términos",) +
  theme_classic()
```

```

# ANÁLISIS DE SENTIMIENTO

# Detección de las emociones y sentimientos
sentimientos <- get_nrc_sentiment(Excel1$texto_limpio, language =
"spanish")

Excel1 <- cbind(Excel1, sentimientos)

# Se crea un data.frame con la suma total de emociones
sentimientos_df <- data.frame(valor = rowSums(t(sentimientos)))%>%
rownames_to_column("Emoción") %>% arrange(desc(valor))

view(sentimientos_df)

# Representación gráfica de la polaridad
sentimientos_df %>% filter(Emoción %in% c("positive", "negative")) %>%
ggplot(mapping = aes(x= reorder(Emoción, valor), y = valor)) +
geom_col(fill = c("Green", "Red")) + coord_flip() + labs(title =
"Análisis de polaridad con NRC", x = "Sentimiento") + theme_classic()

# Representación gráfica de las emociones
sentimientos_df %>% filter(!Emoción %in% c("positive", "negative"))
%>% ggplot(mapping = aes(x= reorder(Emoción, valor), y = valor)) +
geom_col(fill = brewer.pal(n = 8, name = "RdYlGn")) + coord_flip() +
labs(title = "Análisis de sentimiento con NRC", x = "Emociones") +
theme_classic()

# Clasificación y representación de palabras según los sentimientos
afinn <- read.csv("lexico_afinn.en.es.csv")

Excel1_afinn <- Excel1 %>% unnest_tokens(input = "texto_limpio" ,
output = "Palabra") %>% inner_join(afinn, ., by = "Palabra") %>%
mutate(Tipo = ifelse(Puntuacion > 0, "Positiva", "Negativa"))

Excel1_afinn %>% count(Palabra, Tipo, sort = TRUE) %>% group_by(Tipo)
%>% top_n(15) %>% ungroup () %>% mutate(Palabra=reorder(Palabra, n))
%>% ggplot(aes(Palabra, n, fill=Tipo)) + geom_col(show.legend = FALSE)
+ geom_text(aes(label=n), hjust=1.2) + facet_wrap(~Tipo, scales =
"free_y")+ coord_flip() + labs(y="Frecuencia") + xlab(NULL)

# Clasificación y representación de palabras según emociones
nrc <- read.csv("lexico_nrc.csv", encoding = "UTF-8")

colnames(nrc)[2]<- "Emoción"

```



```
Exce11_nrc <- Exce11 %>% unnest_tokens(input = "texto_limpio" , output  
= "palabra") %>% inner_join(nrc, by = "palabra")
```

```
Exce11_nrc %>% filter(Emoción!="negativo" & Emoción!="positivo") %>%  
count(palabra,Emoción, sort=TRUE) %>% group_by(Emoción) %>% top_n(8)  
%>% ungroup() %>% mutate(palabra=reorder(palabra,n)) %>%  
ggplot(aes(palabra,n,fill=Emoción))+ geom_col(show.legend = FALSE) +  
geom_text(aes(label=n), hjust= 0) + facet_wrap(~Emoción,scales =  
"free_y") + coord_flip() + labs(y="Frecuencia") + xlab(NULL)
```

ANEXO II: LOCALIZACIONES DE INSTAGRAM

ETIQUETAS DE FILA	CUENTA DE LOCATION_ID
catedral-de-leao	94
catedral-de-leon	407
catedral-de-leon-espana	301
catedral-de-leon-espana-leon-cathedral-spain	200
catedral-de-leon-leon-y-castilla-espana	136
catedral-de-santa-maria-de-leon	484
cathedrale-de-leon	129
leon-cathedral	400
plaza-catedral-de-leon	50
plaza-de-la-catedral-leon	132
plaza-de-regla	211
plaza-de-regla-catedral-sta-maria-de-leon-espana	281
TOTAL GENERAL	2825