



UNIVERSIDAD AUTÓNOMA CHAPINGO

CENTRO DE INVESTIGACIONES ECONÓMICAS, SOCIALES Y
TECNOLÓGICAS DE LA AGROINDUSTRIA Y LA AGRICULTURA
MUNDIAL

TESIS

EL USO DEL BIG DATA DE LA WEB 2.0 PARA EL
POSICIONAMIENTO DE PRODUCTOS AGROINDUSTRIALES:
EL CASO DEL MEZCAL

QUE COMO REQUISITO PARCIAL PARA OBTENER EL GRADO
DE:

MAESTRO EN ESTRATEGIA AGROEMPRESARIAL

PRESENTA:

ARMANDO VALDÉS LÓPEZ

BAJO LA SUPERVISIÓN DE:

JORGE AGUILAR ÁVILA DR.

Y NORMAN AGUILAR GALLEGOS DR.



DIRECCION GENERAL ACADEMICA
DEPTO. DE SERVICIOS ESCOLARES
OFICINA DE EXAMENES PROFESIONALES



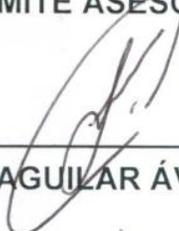
Chapingo, Estado de México Mayo 2017

Tesis realizada por **Armando Valdés López**, bajo la dirección del Comité Asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN ESTRATEGIA AGROEMPRESARIAL

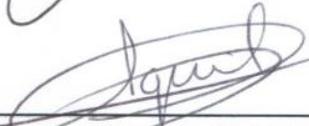
COMITÉ ASESOR

DIRECTOR:



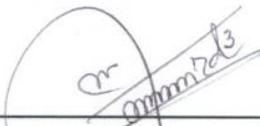
DR. JORGE AGUILAR ÁVILA

CO-DIRECTOR



DR. NORMAN AGUILAR GALLEGOS

ASESOR:



DR. MANRRUBIO MUÑOZ RODRÍGUEZ

ASESOR



DR. VINICIO HORACIO SANTOYO CORTÉS

Dedicatoria

Esta pequeña y sencilla publicación es dedicada con mucho cariño a Dios por haberme dado la oportunidad de cumplir una meta más en mi vida, a mi madre Alejandra y a mi padre Armando que, con su ejemplo, su comprensión, consejos y su apoyo incondicional me permitieron terminar mi Carrera Profesional y han guiado mi camino...

A mis hermanas Sayra y Lizbeth, a pesar de todos los problemas que se han presentado, me han enseñado a nunca rendirse...

A mi hermanito Jesús que a pesar de que ya no está con nosotros, siempre sigue en mi corazón...

A mi hermano Ramón, por ser pilar y soporte en momentos difíciles además porque es y será siempre mi mejor amigo.

Muchas gracias por todo...

Agradecimientos

La consecución de esta Tesis ha conllevado mucho trabajo y no poco esfuerzo personal. Difícilmente la hubiese culminado de no ser por la incondicionalidad de otras vidas que me han prestado su ayuda, apoyo y en las más de las veces su afecto. Sin duda, en los malos momentos, en lo personal y profesional, han colaborado a otorgar significado a este reto profesional.

Una vez, llegado el momento de “cerrar”, deseo manifestarles mi agradecimiento más sincero

A los doctores Jorge Aguilar Ávila y Norman Aguilar Gallegos, por su valiosa y acertada dirección en la realización del presente trabajo. Además, por sus atinadas observaciones e invaluable sugerencias que enriquecieron y mejoraron el trabajo.

Al Dr. Manrribio Muñoz Rodríguez, por su esencial apoyo y colaboración para la culminación del presente trabajo.

Al Dr. Vinicio Horacio Santoyo Cortés, por el invaluable apoyo prestado para que este trabajo concluyera.

Mis agradecimientos han de extenderse tanto a las personas que de una forma u otra han colaborado en las diferentes fases de esta investigación, como a aquellas que a posteriori apoyen y continúen esta iniciativa. De igual manera al CONACYT por el apoyo brindado para la realización de este trabajo. Estas páginas pertenecen a todos los que participamos en el significado semántico y empírico de las mismas.

Datos biográficos



Datos Personales

Armando Valdés López nació el 28 de Agosto de 1988 en Texcoco de Mora Estado de México. Cursó la Preparatoria Agrícola en la Universidad Autónoma Chapingo de donde también se graduó en el año 2011 como Ingeniero Agrónomo Especialista en Zootecnia con la tesis titulada “Uso de Tecnologías de la Información y Comunicación en un Modelo de Gestión de la Innovación” Tiene experiencia laboral como Prestador de Servicio Profesionales en las cadenas caprinos, conejos y bovinos leche, así como de enlace regional de la Unidad Técnica Especializada en Gestión de la Innovación en los estados de Guerrero y Puebla. En materia digital cuenta con las certificaciones de Google en Analítica Web, Cloud Computing y Google Adwords.

CURP: VALA880828HMCLPR07

Cédula profesional:9689881

Resumen general

EL USO DEL BIG DATA DE LA WEB 2.0 PARA EL POSICIONAMIENTO DE PRODUCTOS

AGROINDUSTRIALES: EL CASO DEL MEZCAL ¹

THE USE OF BIG DATA ON THE WEB 2.0 FOR THE POSITIONING OF AGRO-INDUSTRIAL PRODUCTS: THE CASE OF “MEZCAL”.

RESUMEN

El presente trabajo se orientó a utilizar las herramientas Big Data para definir los puntos críticos en el proceso de transformación de datos (extraídos de plataformas web) en información sobre el público objetivo y contenido que este público consume en internet de un tema o producto específico (en este caso mezcal), para contribuir al conocimiento en la implementación de una estrategia de posicionamiento web basada en este producto. Para lograr esto, se obtuvieron datos del motor de búsqueda de Google y redes sociales Twitter y Facebook. Después se utilizaron herramientas de minería de datos y análisis de redes sociales para obtener el perfil de usuarios que están interesados en el mezcal. Se encontró que hay un nicho de mercado en consumidores con un rango de edad de 18 a 34 años, preferentemente masculinos y del centro de México. También se identificaron las interacciones web que los usuarios tienen entre sí y con las empresas mezcaleras, identificando líderes de opinión y clientes potenciales para el mezcal. En cuanto a los temas de interés, se encontró una preferencia por cuestiones relacionadas con la información del producto, tales como el significado de mezcal, sus formas de preparación o sus diferencias con el tequila. Finalmente, se descubrió una preferencia por aquellas publicaciones con imágenes relacionadas al producto, tales como agaves o botellas de mezcal. Se concluye que la transformación de los datos en conocimiento útil para una empresa agroindustrial requiere una intervención lógica y asertiva en los procesos de categorización y filtrado de datos, así como una serie de conocimientos y habilidades en el uso de herramientas Big Data.

Palabras clave: Analítica web, Ciencia de los datos, Análisis de redes sociales, Mercadotecnia digital

ABSTRACT

This work was oriented to use the tools of Big Data in order to define the critical points in the transformation process from data (extracted from web platforms) to information about the target audience and contents that they consume on internet and which are related to a specific topic or product (in this case “mezcal”), in order to contribute to the knowledge about the implementation of a web-based positioning strategy on this product. To achieve this, data from Google’s search engine and the social networks Twitter and Facebook were gathered. Then, data mining tools and social network analysis were used to obtain the users’ profile who are interested in the *mezcal*. It was found that there is a niche market with consumers between 18 to 34 years old, preferably male from the centre of Mexico. It was also identified the web-interactions that users have among themselves and the *mezcal* distiller companies, opinion leaders and potential clients who are interested on the *mezcal* were identified. Regarding the topics of interest, it was found a preference for issues related to information about the product, such as the meaning of *mezcal*, its preparation ways or its differences with tequila. Finally, it was discovered a preference for those publications with images related to the product, such as *agaves* or bottles of *mezcal*. It is concluded that the transformation of the data into useful knowledge for an agro-industrial company requires a logical and assertive intervention in the processes of categorization, and data filtering, as well as a set of knowledge and skills in the use of Big Data tools.

Keywords: Web Analytics, Data Science, Social Network Analysis, Marketing

¹Tesis de Maestría en Estrategia Agroempresarial, Universidad Autónoma Chapingo

Autor: Armando Valdés López

Director de Tesis: Dr. Jorge Aguilar-Ávila

ABREVIATURAS

ARS	Análisis de redes Sociales
AMIPCI	Asociación Mexicana de Internet
API	Application Programming Interface
INEGI	Instituto Nacional de Estadística y Geografía
HT	Hashtag
RT	Retweet
MN	Menciones
TIC	Tecnologías de la Información y la comunicación
SEO	Search Engine Optimization
SEM	Search Engine Marketing
SMM	Social Media Mining

CONTENIDO

1.	INTRODUCCIÓN.....	1
1.1.	Justificación.....	2
1.2.	Preguntas de investigación.....	5
1.3.	Objetivos de investigación.....	5
1.3.1.	Objetivo general.....	6
1.3.2.	Objetivos particulares.....	6
1.4.	Hipótesis.....	7
1.5.	Estructura de la tesis.....	7
2.	REVISIÓN DE LITERATURA.....	9
2.1.	Penetración del internet en México.....	9
2.2.	Web 2.0 y la nueva mercadotecnia.....	12
2.3.	Redes sociales.....	14
2.4.	Google.....	16
2.5.	Facebook.....	19
2.6.	Twitter.....	23
2.7.	Los “ <i>influencers</i> ”.....	27
2.8.	Conceptos de Big Data.....	30
2.9.	Definición de Big Data.....	31
2.10.	Tipos de datos en el Big Data.....	33
2.11.	Características del Big Data (las 3V).....	34
2.12.	El desafío del análisis de datos.....	36
2.13.	La minería de datos en el Big Data.....	38
2.14.	Importancia del análisis de redes sociales y conceptos de ARS.....	39

3.	METODOLOGÍA	45
3.1.	La pirámide de la información.....	45
3.2.	Extracción depuración y análisis de datos para Google	48
3.3.	Extracción depuración y análisis de datos para Facebook	52
3.4.	Extracción depuración y análisis de datos para Twitter	55
3.5.	Importancia del mezcal como producto para el estudio de caso	58
4.	RESULTADOS Y DISCUSIÓN.....	61
4.1.	Google	61
4.2.	Facebook.....	79
4.3.	Twitter.....	93
	4.3.1. Identificación del contenido de interés (análisis de hashtag y primera forma de retweets 1).....	95
	4.3.2. Acercamiento al público meta (ubicación, análisis de menciones y segundo análisis de retweets 2)	105
5.	CONCLUSIONES.....	114
6.	LIMITACIONES Y RECOMENDACIONES	117
7.	LITERATURA CITADA.....	119
8.	APÉNDICES.....	130

Lista de cuadros

Cuadro 1.	Indicadores de red del ejemplo	43
Cuadro 2.	Estadísticos descriptivos y ejemplificación de las categorías de palabras clave.....	63

Cuadro 3. Descriptivos con respecto a 28 Fanpage que abordan temáticas sobre mezcal	79
Cuadro 4. Análisis de correlaciones entre variables de Facebook.....	83
Cuadro 5. Tipificación de las publicaciones	85
Cuadro 6. Publicaciones clasificadas por hora	86

Lista de figuras

Figura 1. Usuarios de internet en 2015 clasificados por edad	10
Figura 2. Porcentaje de usos en México para el internet en 2015	10
Figura 3. Porcentaje usuarios en México por entidad federativa	11
Figura 4. Ejemplo de la posición del botón “me gusta de página” en una Fanpage	21
Figura 5. Ejemplo de la posición del botón “me gusta” en una publicación	22
Figura 6. Visualización de un tweet con los hashtag #Oaxaca y #mezcal	24
Figura 7. Ejemplo de retweet, mensaje compartido de la cuenta @Turibuscdmx	24
Figura 8. Ejemplo de mención, mensaje en el que se menciona a la cuenta @armandoideas1	25
Figura 9. Mensaje publicitario de Samsung Galaxy enviado desde iPhone.....	29
Figura 10. Ejemplo de red con 22 nodos y 23 aristas para la representación de indicadores	43
Figura 11. Pirámide de la información de las plataformas digitales	47
Figura 12. Proceso de análisis de datos para Google	48
Figura 13. Función de autocompletado del buscador	49
Figura 14. Captura de pantalla de la herramienta de Google Adwords	50
Figura 15. Proceso de análisis para publicaciones de Facebook	53
Figura 16. Metodología para el análisis de tweets	56

Figura 17. Estructura del grafo de redes por análisis.....	58
Figura 18. Porcentaje de las categorías por el número de apariciones	67
Figura 19. Promedio de búsquedas mensuales por categoría.....	67
Figura 20. Usuarios que buscan los tópicos del mezcal seleccionados, distribuidos por estado.....	69
Figura 21. Usuarios que buscan los tópicos del mezcal seleccionados, distribuidos por estado.....	70
Figura 22. Clasificación de palabras clave por cuadrante (resultados/búsquedas)	72
Figura 23. Combinaciones de palabras elegidas por una agroindustria mezcalera	73
Figura 24. Fluctuación de búsquedas mensuales por combinación de palabras	75
Figura 25. Palabras clave que muestran fluctuaciones por temporada	76
Figura 26. Pasos que se siguieron para la generación de conocimiento en cuanto a palabras clave.....	78
Figura 27. Interacción de likes de usuarios con publicaciones de Fanpage	84
Figura 28. Porcentaje de publicaciones emitidas distribuidas por hora a lo largo del día.....	87
Figura 29. Publicaciones emitidas y publicaciones con comentarios, clasificadas por categoría.....	90
Figura 30. Representación del análisis de correspondencias entre las Fanpage y las categorías de las características de las publicaciones.....	92
Figura 31. Pasos que se siguieron para la generación de conocimiento en cuanto a Fanpage de Facebook.....	93
Figura 32. Red de hashtags (HT) representada al 1.98% de nodos visibles y 5.18% de aristas visibles para su mejor visualización.	96

Figura 33. Tweet viral emitido por el periodista mexicano Joaquín López Dóriga	98
Figura 34. Red de Retweets 1 con la palabra mezcal, filtro al 4.7% de nodos para su mejor visualización.....	99
Figura 35. Tweet viral emitido por la cuenta @Roycampos.....	100
Figura 36. Tweet que hace referencia a una canción de música regional mexicana	101
Figura 37. Tweet que hace referencia al mezcal con un hecho noticioso.....	102
Figura 38. Tweet acerca de un post sobre la cultura del mezcal en la ciudad de México	103
Figura 39. Tweet del gobierno de Oaxaca que habla sobre la tradición del mezcal	104
Figura 40. Distribución de tweets sobre mezcal a nivel mundial.....	106
Figura 41. Distribución de tweets sobre mezcal en México	106
Figura 42. Distribución de tweets sobre mezcal en la Ciudad de México	107
Figura 43. Comunidades más grandes de la red de menciones.....	110
Figura 44. Comunidades más grandes de la red de retweets.....	111

1. INTRODUCCIÓN

Con el auge de la comunicación en la era digital, ahora es más fácil la interacción entre usuarios de la red. Nos encontramos en un periodo de tiempo en el que los obstáculos para la transferencia de mensajes e ideas se redujeron notablemente, una era de constante transformación en un mundo cada vez más globalizado. Las tecnologías de la información y la comunicación han ayudado a vencer las barreras espacio temporales para el envío, el almacenamiento, la manipulación y difusión de información.

El uso de la tecnología ha traído consigo cambios sustanciales en la vida cotidiana. Así lo señalan autores como Acosta V.(2012), quien dice que un lugar común de las reflexiones en las ciencias sociales y humanas es el reconocimiento del impacto y transformaciones de las tecnologías de la información y comunicación (TIC) en el ámbito humano. Del Fresno García, Daly, & Supovitz, (2015) señalan que dentro de estos cambios se encuentran la estructura de las relaciones interpersonales, la generación, difusión e intercambio de información de manera colectiva.

Iglesias Rodríguez & Herrero Gutiérrez (2016) hacen énfasis en la transformación de lo denominado Web 1.0 a una nueva versión (Web 2.0) que ha hecho que el usuario pase de ser un sujeto pasivo que recibe información a ser un individuo capaz de interactuar o aportar contenido en espacios masivos. Alonso González (2016) añade que en la Web 2.0 se incorporan nuevos paradigmas para las empresas y su interacción con el usuario, convirtiendo a éste en el sujeto más importante de la relación ya que, tal y como menciona Fernández Burgueño (2012), su actividad puede construir o destruir la reputación online de aquellos que quieren forjar su identidad.

En este contexto, Sánchez Casado & Mostaza-Ucedo (2014) explican el surgimiento, en internet, de nuevas formas y tendencias de comunicación denominadas medios sociales, a través de las cuales las empresas pueden establecer relaciones con sus clientes y definir nuevas estrategias de marketing.

Entre los medios más populares se encuentran las denominadas redes sociales, Marín, Duque, & Manruffo (2014) indican que son medios de una altísima conectividad entre las personas y su capacidad de interacción supera lo que otro medio de comunicación había presentado antes. Mir Bernal, Fondevila Gascón, & Gutiérrez Aragon (2016) añaden que, debido a su elevada popularidad, las empresas desean utilizarlas como medio para incrementar su visibilidad, popularidad y ventas. Tal como mencionan Grandi Bustillos, González Ladrón de Guevara, & Segarra Oña (2014), esto se puede hacer sin invertir demasiados recursos pues ofrecen una forma atractiva de marketing, además de una interacción constante con una audiencia específica. Sin embargo, Viñarás Abad & Cabezuelo Lorenzo (2012) señalan que existe el riesgo que si la estrategia es equivocada el efecto boomerang será nefasto.

En esta nueva “era de la información”, donde el internet juega un rol primordial en la comunicación, se hace presente como un axioma la frase “Lo que no está en internet no existe”. Un autor que refuerza esta idea es Sixto García (2015) al decir que el internet es el mayor centro comercial del mundo, ofreciendo de todo y a todas horas y que a pesar que los usuarios todavía no han olvidado cómo se hacen las compras físicamente (la mayoría las siguen haciendo), sí lo usan para buscar información sobre lo que se quiere comprar; el autor cierra con la reflexión de que ahora ya no compramos sin informarnos y sin antes comparar los productos, para lo cual exigimos tiendas abiertas 24 horas todos los días del año.

1.1. Justificación

En la actualidad la generación de datos a través de los medios digitales es inmenso. De acuerdo con López García (2012), cada minuto que pasa, los 2,700 millones de personas con acceso a internet (que se calculó que había en el mundo en 2012), envían más de 200 millones de correos electrónicos, realizan 2 millones de consultas a Google, suben 48 horas de vídeo a YouTube, escriben más de 100,000 mensajes en Twitter, publican casi 30,000 nuevos artículos en

WordPress, suben más de 6,000 fotografías a Instagram y se descargan 47,000 aplicaciones del sistema operativo iOS. Congosto & Aragón (2012) señalan que el volumen de información generada supera con creces el volumen de cualquier sondeo y tiene la ventaja de que el muestreo sea continuo. Lo anterior ha hecho que autores como Goyzueta Rivera (2011) señalen que el comportamiento del consumidor “online” ha merecido que los mercadólogos se esfuercen en demasía para tratar de comprenderlo. Según Del Fresno García et al. (2015), este tipo de datos ofrece oportunidades y desafíos a los investigadores, donde lo más interesante radica, no en las cantidades de datos, sino en qué podemos hacer con los datos.

Sin embargo, Congosto & Aragón (2012) destacan también que la información es desestructurada, y para el caso de Twitter, por ejemplo, es limitada a 140 caracteres, además de presentar numerosas dificultades al momento de segmentar por edad, sexo, grado de estudios o nivel económico. Por tal motivo para obtener resultados concluyentes es preciso aplicar nuevos métodos de análisis que permitan comprender la reacción de los ciudadanos tanto en eventos importantes como en el día a día.

Por todo lo anterior, y retomando lo dicho por Oviedo García, Muñoz Expósito, & Castellanos-Verdugo (2015), podemos decir que se abre así, un campo incógnito de oportunidades y amenazas que deben ser afrontadas por los gestores de marketing y para ello, el primer paso ineludible es comprender la tecnología que subyace en los medios sociales y las peculiaridades, fines y capacidades de cada plataforma social para, más adelante, poder establecer objetivos de comunicación en cada una de ellas, gestionar su comunicación (controlando la información que se distribuye en ellas y participando en las mismas) y, finalmente, evaluar los resultados en función de los objetivos perseguidos.

En el marco de este gran desafío, provocado por las nuevas herramientas digitales, entra en juego la analítica del Big Data; la cual según Ortiz Morales, Joyanes Aguilar, & Giraldo Marín (2016) plantea el análisis de grandes volúmenes de datos para detectar relaciones entre ellos que puedan proveer

información útil a las empresas, facilitando la toma de decisiones en todos los procesos y las áreas de la organización basadas en datos de los clientes.

Ríos (2014) indica que hoy en día las empresas necesitan procesar una gran cantidad de datos, lo que en el mundo de la informática se engloba bajo el concepto de Big Data, lo cual da como resultado una mejor gestión del negocio. En el mismo sentido, Ortiz Morales et al. (2016) resaltan que la implementación del Big Data y la analítica de datos más que un reto supone una gran oportunidad para las empresas y para sus departamentos de marketing, pues con su utilización se puede llegar a obtener información relevante del cliente, sus gustos, sus criterios de compra y otros datos. Los mismos autores antes citados destacan la importancia de inquietarse por el Big Data y sus principales características (volumen, velocidad, variedad, veracidad, valor y visualización) que día a día llevan a una inestabilidad de la información y, por consiguiente, a un cambio constante en la toma de decisiones.

El poder de contextualización que nos dan los datos alojados en las diferentes plataformas digitales, también sirve de referente para inferir comportamientos futuros y tendencias en el entorno. Ejemplo de lo anterior son los trabajos de Gómez Martínez & Prado Román (2014), Tello-Díaz (2013), Barberá & Rivero (2012) que hablan del poder predictivo del buscador de Google, Facebook y Twitter, respectivamente, acerca de sus preferencias y su reacción a los mensajes y eventos principalmente políticos.

En el libro electrónico de Lecinski (2013) publicado por la empresa Google se resalta la frase: “Estoy convencida de que si escucháramos a los clientes todos los días, nos dirían lo que desean y necesitan”, dicha frase es un reflejo de lo que se puede obtener de las plataformas digitales, ya que dentro de ellas se vierte diariamente una cantidad inmensa de datos con respecto a gustos y preferencias. Dichos datos se pueden traducir en información y posteriormente en conocimiento, que nos permita conocer más a fondo tanto las características del público meta así como los temas y el tipo de contenido con mayor interés por parte de este público meta. No obstante, es necesario identificar y definir los

puntos críticos en el camino desde la extracción de datos brutos de las plataformas web hasta la generación del conocimiento que permita a una empresa de carácter agropecuario posicionarse dentro del entorno digital.

Por tal motivo, el presente trabajo de Big Data de plataformas web aplicado al producto mezcal busca contribuir a una aproximación con respecto al público objetivo y al contenido que éste busca en dichas plataformas. Para ello se usan herramientas relacionadas al procesamiento del Big Data, su transformación de simples datos a conocimiento para ser utilizado en una estrategia de posicionamiento web.

1.2. Preguntas de investigación

En correspondencia con lo anterior, se plantean las siguientes preguntas de investigación:

1. ¿Qué tipo de información, con respecto al público objetivo, de una empresa agropecuaria se puede generar, a partir de datos obtenidos de plataformas web 2.0?
2. ¿Qué tipo de información, con respecto al contenido de interés de los internautas y en torno a un producto agropecuario se puede generar, a partir de datos obtenidos de plataformas web 2.0?
3. ¿Cuáles son y cómo abordar los puntos críticos que se consideran en el proceso de obtención de conocimiento, a partir de datos obtenidos de plataformas web 2.0?

1.3. Objetivos de investigación

El propósito de la presente investigación se puede desglosar en un objetivo general y tres objetivos particulares que a continuación son descritos:

1.3.1. Objetivo general

Definir los puntos críticos que permitan la transformación de datos obtenidos de plataformas web en información acerca del público meta y características del contenido que éste consume en internet, mediante herramientas de Big Data con el fin de contribuir al conocimiento para la implementación de una estrategia de posicionamiento web en torno a un producto agropecuario.

1.3.2. Objetivos particulares

Mediante herramientas de Big Data identificar, describir, y ejemplificar el tipo de información con respecto a usuarios en internet que se pueden obtener a partir de datos extraído de la web 2.0 con la finalidad de conocer el segmento de mercado en el cual implementar una estrategia de posicionamiento digital, focalizando en una empresa agropecuaria.

Con el apoyo de herramientas de Big Data identificar, describir, y ejemplificar el tipo de información que se puede obtener con respecto a los temas tendencia en internet en torno a un producto, que permita orientar a una empresa agropecuaria para la generación de contenido en una estrategia de posicionamiento digital.

Identificar y describir los puntos críticos en el proceso de transformación de datos a conocimiento, con la finalidad de desarrollar una metodología replicable que permita obtener información de utilidad para las empresas agropecuarias, lo anterior mediante la aplicación de herramientas de Big Data para un producto agropecuario.

1.4. Hipótesis

1. Los parámetros de los internautas que se logran identificar son bastos y de fácil obtención, y con una estructura amigable que facilita su análisis; además, que con ellos se obtiene de manera puntual la segmentación del mercado deseada.
2. La información acerca de los temas tópicos están referidos exclusivamente a marcas, productos y promoción de los mismos, dado esto por el *branding* que se genera *offline* por parte de las principales empresas incursionadas en las plataformas web analizadas.
3. Un análisis cuantitativo en base a los datos otorgados por las diferentes plataformas, nos ofrece un amplio conocimiento con respecto a la audiencia a la cual se puede incursionar y se puede obtener de una manera sencilla con las herramientas y datos que existen actualmente sin la necesidad de hacer un exhaustivo análisis.

1.5. Estructura de la tesis

La presente tesis está dividida en ocho apartados. El primero corresponde a la introducción, donde se describe en un primer momento los cambios que han traído consigo las nuevas tecnologías de la información y la comunicación y que esos cambios incluyen la interacción que tienen las empresas con su audiencia en internet; se destaca el poder predictivo de los datos en las plataformas web y el problema de investigación que representa el desafío de cómo obtener valor a partir de la ingente producción de datos junto con sus correspondientes preguntas, objetivos e hipótesis.

En el apartado dos se presentan una revisión de literatura iniciando con una contextualización de la penetración de internet en México, la web 2.0 y las nuevas maneras de hacer mercadotecnia. Se explica y describen las tres diferentes

plataformas web utilizadas en la investigación, los conceptos de Big Data y el análisis de redes sociales.

En el apartado tres se describe la metodología realizada para alcanzar los objetivos de la investigación. La metodología es desglosada para cada una de las tres plataformas utilizadas y que se basó en el proceso del triángulo de la información que pasa de los datos a la generación de conocimiento; también se describe brevemente la importancia del producto utilizado para el estudio de caso.

En el apartado cuatro se presentan los resultados obtenidos y se discuten en el marco de otras investigaciones. El proceso se realiza desglosando los hallazgos para cada una de las tres plataformas web utilizadas.

El apartado cinco corresponde a las conclusiones, esto se hace tomando como base las hipótesis planteadas en la investigación y, por ende, de los objetivos mismos. En el apartado seis se bosquejan algunas limitantes derivadas del trabajo de investigación y que servirán para contextualizar los hallazgos obtenidos, pero también para futuros emprendimientos.

Por último, en los apartados siete y ocho se presentan la literatura citada y los apéndices, respectivamente.

2. REVISIÓN DE LITERATURA

2.1. Penetración del internet en México

El uso de internet en México ha venido en aumento. De acuerdo a datos de la Asociación Mexicana de Internet (AMIPCI, 2016), la población internauta en México aumento un 15.7% del 2014 a 2015 posicionándose en un 59%; destaca también que el perfil del usuario de internet en su mayoría (con un 20%) se encuentra en edades de 25 a 34 años , seguido con un 19% por internautas con edades de 13 a 18 años.

Del mismo estudio señala que el tiempo promedio diario de conexión es de siete horas con 14 minutos, destacando como uso principal el acceso a redes sociales con un 79% y en cuarto lugar con un 64% se encuentra la búsqueda de información.

Por su parte el Instituto Nacional de Estadística y Geografía (INEGI,2015) coloca en primer lugar en el uso de internet a los usuarios de entre 25 a 34 años con un 20%, empatado con el mismo porcentaje a los usuarios con rango de edad de 18 a 24 años (Figura 1). También destaca que el principal uso del internet con un 87% corresponde a la obtención de información y que el uso de internet para acceder a redes sociales se encuentra en cuarta posición con un 71% (Figura 2).

El INEGI (2015) también nos da una idea de cómo se encuentra el porcentaje de usuarios en internet por entidad federativa. En la Figura 3 se pudo destacar que el estado más rezagado (con 33.7%) es Chiapas, de ahí en adelante todos los estados superan el 40% destacando las dos Baja California , Nuevo León y la Ciudad de México como los estados punteros en usuarios de internet.

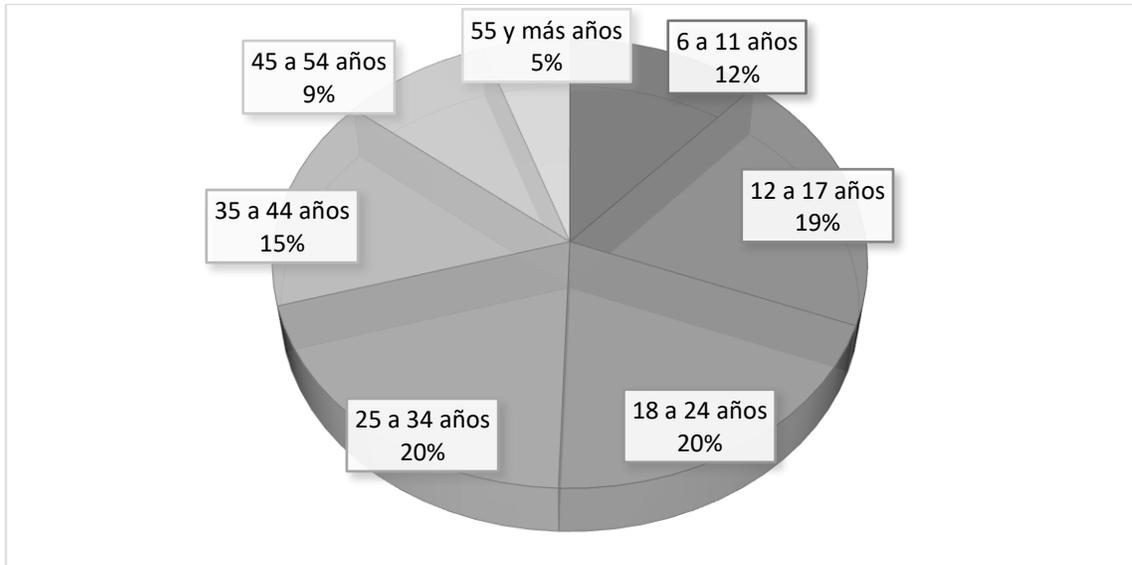


Figura 1. Usuarios de internet en 2015 clasificados por edad

Fuente: el Instituto Nacional de Estadística y Geografía (INEGI) 2015.

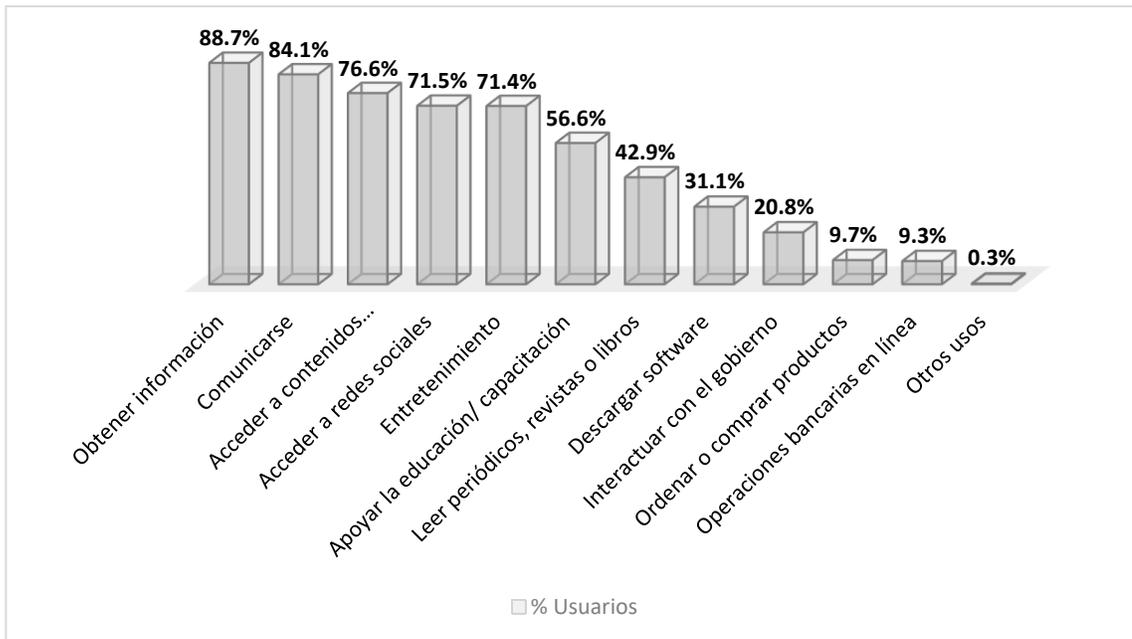


Figura 2. Porcentaje de usos en México para el internet en 2015

Fuente: el Instituto Nacional de Estadística y Geografía (INEGI) 2015.

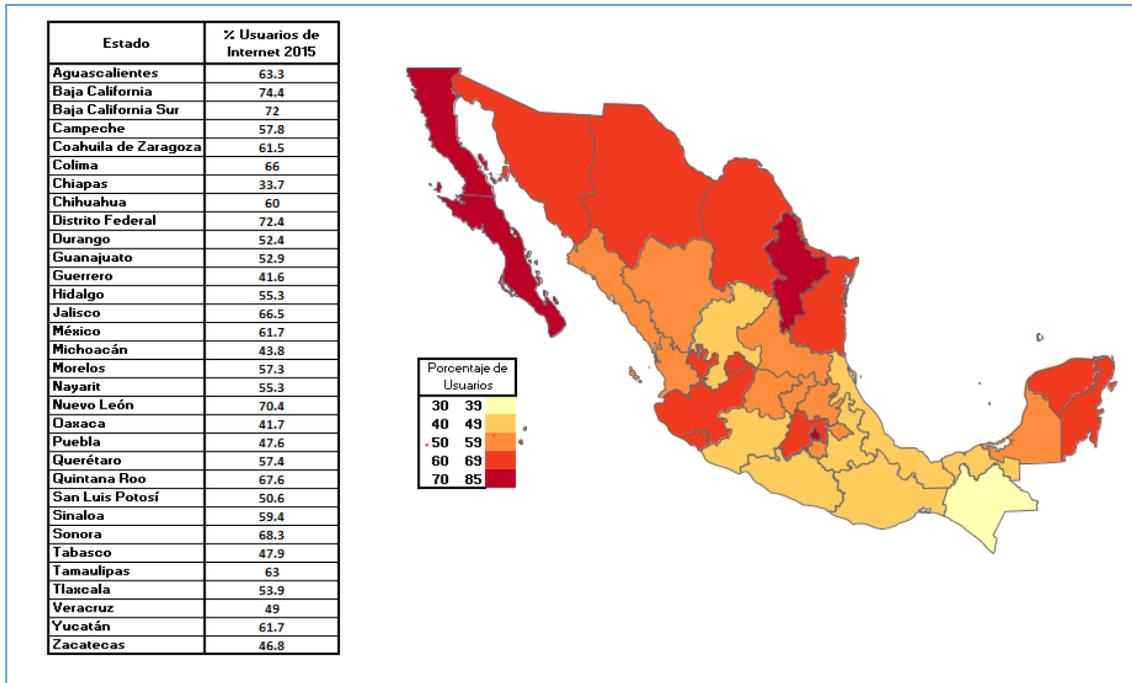


Figura 3. Porcentaje usuarios en México por entidad federativa

Fuente: el Instituto Nacional de Estadística y Geografía (INEGI) 2015.

El AMIPCI (2016) también destaca que, en cuanto a la penetración de redes sociales, Facebook sigue siendo la red referencia (con un 92%), seguida por Whatsapp (con el 79%) y Youtube (con 66%); Twitter, por su parte, se ubica en la cuarta posición con un 55%.

Otro estudio del AMIPCI (2014) menciona que un 79% de las empresas se encuentran presentes en Facebook (con un promedio de publicaciones de siete a la semana) y un 80% se encuentran en Twitter emitiendo 21 tweets por semana.

También se destaca que el 55% de las publicaciones que hacen en Facebook contienen imágenes, por un 21% del contenido en Twitter en donde predomina el texto puro con un 39%.

2.2. Web 2.0 y la nueva mercadotecnia

Como definición de lo denominado Web 2.0 se retoma lo expuesto por Paladines Galarza & Granda Tancazo (2013), quién señala que encierra aquellas plataformas que implican la interacción y la personalización, el intercambio y la colaboración. Estos autores añaden también que son espacios comunicativos relacionales, en los que se elabora y se coopera como redes sociales, blogs, wikis, etc. De acuerdo con Martínez & Monserrat Gauchi (2012), una de sus características principales es la transición de la comunicación unidireccional con la multidireccional. Por su parte Rodríguez Rangel & Sánchez Rivero (2015) abonan al concepto diciendo que se pasa de la mera consulta de contenidos creados por otros, principalmente empresas, a la posibilidad de que sean los propios consumidores los creadores de contenidos.

En este contexto, es inminente tomar en cuenta cuatro grandes aspectos que implica este nuevo entorno: la generación de vínculos con la audiencia, el conocimiento de las características de los tipos de usuario, la viralidad de los mensajes transmitidos a partir de boca a oreja y la metodología de medición de cada una de estas particularidades.

Con respecto a la primera idea, se retoma lo mencionado por Túñez-López, Sixto García, & Guevara-Castillo (2011) quienes dicen que la práctica del marketing ya no solo es sólo una transacción, sino el establecimiento de vínculos estables y duraderos con los usuarios que sean provechosos para las dos partes implicadas. Lo anterior, dicen, es parte del nuevo marketing de relaciones, el cual implica el establecimiento de vínculos estables y duraderos con los usuarios en la búsqueda de su satisfacción y lealtad.

Pasando a la segunda idea, se parafrasea lo dicho por Paladines Galarza & Granda Tancazo (2013) los cuales sostienen que se requiere de un mayor conocimiento tanto de los hábitos de consumo como de los soportes a través de los cuales se realiza dicho contacto; recomiendan también que los anunciantes tengan una actualización constante sobre dichos hábitos. Oviedo García et al.

(2015) abonan a esta idea señalando que el conocer a los usuarios permitirá escoger aquellos más apropiados a los objetivos de comunicación propios de la empresa y, a la vez, descartar los otros; de ese modo será posible evitar embarcarse en formas de comunicación que pasarán desapercibidas o provocarán rechazo en los usuarios.

La tercera idea gira en torno a la importancia del boca a oreja, concepto que siempre ha sido parte de la mercadotecnia pero que sin embargo al digitalizarse en la era de la web 2.0 se ha potencializado llegando a límites inimaginables como el caso de los mensajes virales. Aguado Guadalupe & García García (2009) destacan que ésto ha traído consigo el marketing viral, el cual permite que crezca exponencialmente la cartera de clientes, provocando en cada cliente satisfecho un número determinado de recomendaciones, dependiendo del círculo de amistades o contactos que tenga nuestro receptor-emisor. Los mismos autores señalan que el término “viral” es un adjetivo que una campaña de marketing se gana cuando los consumidores la hacen suya, participan, la comentan y la amplifican reforzándose cuando se multiplica, y esto constituye una excelente barrera de entrada contra los competidores.

La creación del marketing viral da pie a otro concepto importante: el “marketing de influencia”. De acuerdo con Castelló Martínez & del Pino Romero (2015), en éste son los líderes de opinión o los personajes famosos, como prescriptores o “*influencers*”, personas influyentes a los que las marcas dirigen sus esfuerzos comunicativos para, a través de ellos, llegar a más consumidores potenciales. Los mismos autores destacan la capacidad de identificar estos líderes de opinión, quienes pueden ayudar a una marca a conectar de forma natural y espontánea, no invasiva, con su público objetivo.

Como cuarta idea y englobando los apartados anteriores se destaca la necesidad de la medición de indicadores que nos permitan diagnosticar, definir y proponer una estrategia sólida de posicionamiento de marca. Al respecto, Oviedo García et al. (2015) por ejemplo mencionan que un paso para que las empresas puedan rentabilizar sus esfuerzos de marketing en los medios sociales es tratar de medir

lo que están obteniendo de ellos. Así, los gestores de marketing deben contar con indicadores que les permitan cuantificar los resultados de sus inversiones en medios sociales, además de proporcionarles criterios para segmentar a la audiencia y optimizar, por tanto, los esfuerzos de marketing.

2.3. Redes sociales

En palabras de Pérez Salazar & Aguilar Edwards (2012), las redes sociales es un término que ha sido socialmente construido de manera reciente al menos desde dos dimensiones: 1) Como categoría de análisis socio-relacional y 2) Para referirse al conjunto de herramientas informáticas en línea que permiten la administración de contactos (entre las que destacan por su popularidad actual Facebook y Twitter).

La segunda dimensión (que es la que aborda en este capítulo) hace referencia a lo que autores como Litche Fragoso & Sánchez Salinas (2014) denominan un servicio en línea entre personas que, por ejemplo, comparten intereses y actividades, además de utilizarse para transmitir experiencias y sensaciones que viven, así como para obtener tres beneficios principalmente: i) adquisición de información; ii) identificación o pertenencia a una comunidad; iii) y diversión o entretenimiento. Los autores antes citados señalan que un servicio de red social consiste en una representación de cada usuario (perfil), sus vínculos sociales, y una variedad de servicios adicionales.

Autores como Barradas Gurruchaga & Siano (2014) explican que la ascendencia de las interacciones en red procede por la mismas características del hombre al decir que el ser humano es social por naturaleza, busca la integración social al medio que lo rodea, es decir la aceptación como una persona capaz de expresar sus ideas, y aportar algo importante a los demás.

La importancia del uso de estas redes sociales ha entrado en apogeo en los últimos años. Así lo sostienen autores como Fainholc (2011) o Ferreas Rodríguez

(2011). El primer autor al decir que las redes sociales en internet son el fenómeno socio-electrónico contemporáneo que replantea la vida diaria en sus formas básicas de comunicación, interacción y producción de saber, mediado por TIC. Y el segundo autor al señalar que redes sociales tales como Facebook y Twitter han experimentado un crecimiento sostenido desde su creación, se han consolidado como fuentes de información, espacios de interacción y para la publicación de contenidos; además, los flujos publicitarios y los tiempos de consumo también aumentan.

Entrando en materia de las propiedades de las redes sociales y de acuerdo con lo mencionado por Iglesias Rodríguez & Herrero Gutiérrez (2016), podemos mencionar que las principales funcionalidades que hacen a los usuarios valorar positivamente la experiencia en las redes sociales son:

1. Posibilidad de comunicación con personas de todo el mundo.
2. Posibilidad de compartir contenidos audiovisuales.
3. Fácil localización de información sobre eventos de todo tipo (culturales, ocio entre otros) y productos.
4. Conocer los contenidos que ofrecen los medios de entretenimiento.
5. Conocer la actualidad informativa.

Además, habría que destacar la inmediatez de la que se caracterizan todas estas actividades dentro de las plataformas.

Hay agencias de comunicación o empresas que utilizan los medios sociales con criterios comerciales, y lo hacen bien de cara a sus intereses, el método es el mismo en la base: segmentar el público objetivo y hacerle llegar la información; pero un uso inoportuno o no calculado puede errar el tiro, y dar en otra parte (Beitia Vallés, 2012).

Cabe señalar que son varios los autores (Aguado Guadalupe & García García, 2009; Fernández Burgueño, 2012; Tello-Díaz, 2013) que nos invitan a reflexionar acerca de la privacidad prácticamente nula de los datos vertidos diariamente en

las plataformas virtuales, ya que en estos espacios las fronteras entre el espacio de lo privado y lo público se mueven; pues los datos personales sobre los individuos se convierten en públicos y globales, a iniciativa propia, disponible la información de una manera y en unas cantidades sin precedentes, a pesar que se insista sobre la naturaleza privada de estas plataformas. Se mencionan que el negocio de las redes es en donde el usuario paga el acceso con sus datos personales (perfil de usuario), colabora enviando fotos y videos además de generar audiencia para la venta de publicidad.

2.4. Google

Uno de los más importantes usos del internet es la obtención de información (Ver Figura 2) y varios son los autores (Aguillo, Ortega, & Granadino, 2006; Arbildi-Larreina, 2005; Grau Moracho & Guallar Delgado, 2004; Soltero Domingo & Bodas Sagi, 2005) que señalan Google (buscador fundado en 1996 por Larry Page y Sergei Brin) como el número uno, tan grande que se ha convertido en un sinónimo de buscar en la red y es reconocido de forma universal como la más completa, potente y eficiente herramienta de recuperación de información en internet.

La importancia de las visitas a una página web radica en gran medida en la posición que este ocupe en las respuestas dadas por el buscador; esto radica en gran medida por el volumen de información existente y los hábitos de los usuarios, que acostumbran a consultar los diez o veinte primeros registros de cada consulta (Morato, Sánchez, & Valiente, 2005).

Moráguez Bergues & Perurena Cancio (2014) definen el posicionamiento web como el proceso que permite mejorar la posición que ocupa un sitio web con respecto a su competencia en las páginas de resultados de los buscadores, mediante la utilización de palabras claves y de factores o técnicas de optimización que se aplican a los sitios web. La aplicación de estas técnicas de

posicionamiento se traduce en un incremento del tráfico y como consecuencia en una mejora de la visibilidad de los sitios web.

Es de suma importancia destacar que tal y como menciona Morato et al. (2005) los algoritmos de posicionamiento utilizan con frecuencia un enfoque multivalente. En este caso, la combinación de distintos criterios se utiliza para calcular el peso por el que serán ordenadas las páginas frente al resto de resultados. Moráquez Bergues & Perurena Cancio (2014) destacan aptitudes la página como lo son la visibilidad, la usabilidad y la accesibilidad multiplataforma como factores que intervienen en la penetración de una página; a estos se les añade también el algoritmo de prestigio, es decir que el posicionamiento de una página no depende únicamente de la asertividad de las palabras clave que use, pero éstas se consideran uno de los principales factores de posicionamiento.

Para explicar más acerca de palabras clave tomamos el trabajo de Amunátegui Perelló (2006), quien nos dice que las páginas que componen la World Wide Web tienen una doble cara: por una parte está aquel conjunto de fotos, sonidos y textos que aparecen proyectadas en el ordenador del usuario de internet al acceder a ella, y por otra está su “código fuente”, que corresponde a la programación misma que permite al computador interpretar como un hipertexto la información a la que accede a través de la red; dicho “código fuente” consiste en una cantidad de órdenes o tags contruidos en un lenguaje sencillo que normalmente aparece oculto a la vista del usuario y de esta manera, mientras el usuario “ve” en su pantalla un determinado conjunto de información, el verdadero contenido de la página permanece escondido y es “comprendido” solo por el ordenador.

El autor antes citado menciona también que algunos de estos tag suelen reunir las palabras clave de la información más relevante en la página visitada y su finalidad se encuentra estrechamente vinculada con los “buscadores” de páginas en la red, haciendo que cuando el usuario desee ubicar información sobre un tema determinado al introducir en su buscador ciertas palabras o frases relacionadas a su interés el buscador da una mirada a los tag de las páginas de

la red a fin de encontrar qué lugares contienen alguna designación relativa a la solicitada, formando de esta manera una lista que responda a las necesidades del usuario; actuando de manera similar a un faro dentro de la red, atrayendo la atención de los buscadores hacia ellos a fin de señalar qué información contiene la página en cuestión. Por tal motivo, el autor ya citado resalta que las palabras clave constituyen un elemento de alta importancia en la red, pues contribuyen a su organización interna haciéndola accesible a los usuarios y útil como medio de búsqueda de información.

Para la elección de las palabras clave idóneas son varios los autores que nos resaltan las variables a tomar en cuenta en nuestra decisión, Arbildi-Larreina (2005) por ejemplo nos señala tres principales:

1. La especificidad. Mide el interés intrínseco de la expresión para los objetivos del sitio. Si optimizamos un portal bibliotecario, no nos interesan usuarios que quieran “comprar libros”, sino quizá “libros gratis”, “libros gratuitos”, “lectura de libros”, “buscador de libros”, “libros digitales” o “libros infantiles”, entre otras, todas ellas búsquedas comunes en internet.
2. Volumen de búsquedas. Nos indica si se busca mucho o poco esa expresión. La herramienta de sugerencias de Google Adwords nos permite partir el estudio de lo que los usuarios buscan y clican por país/países y lenguas distintas. Indica la cantidad media de clics (veces que se pincha en un enlace), nos da una idea de lo mucho o poco que se busca una palabra.
3. Competitividad. Debemos saber si una expresión o palabra clave está siendo objeto de buen posicionamiento por muchos sitios web, y por tanto si merece la pena el esfuerzo. Dos parámetros nos lo indican: el número de links a las páginas que están más arriba en el ranking y la cantidad de páginas totales que aparecen como respuesta a esa búsqueda (la cifra se indica en la parte superior de la pantalla de respuestas).

Por último, es importante destacar lo mencionado en el trabajo de Amunátegui Perelló (2006), quien nos habla acerca de utilizar como palabras clave las de los competidores potenciales más relevantes, esto con el fin de atraer a los clientes de éstos a sus propias páginas. Según el autor antes citado, éste creó una serie de conflictos jurídicos en los Estados Unidos de Norteamérica, pues al encontrar diferentes páginas en los resultados de la búsqueda, el consumidor no sabe exactamente cuál de ellos corresponde efectivamente a la empresa que busca; esto a pesar de que con un somero análisis el usuario dará con la página que se proponía encontrar y no obstante no haber riesgo de confusión final (puesto que las páginas se diferencian lo suficiente como para que sepa en qué tienda compra), lo cierto es que ha logrado acceder a las páginas de los competidores de la compañía en cuestión a través del uso del nombre de la misma. En otras palabras, el autor argumenta que las empresas competidoras, en su afán de acceder al mercado, se están “colgando” del buen nombre de la empresa dominante. Consecuentemente, el esfuerzo realizado en publicidad, en mercadeo y, en general, en la obtención de una reputación por parte de la empresa afectada, redundará en beneficio de sus competidores.

2.5. Facebook

Facebook tal y como describe Acosta V. (2012) es un sitio web gratuito que oficia de mediador para la creación de redes sociales. Este sitio es uno de los más concurridos en la web debido a que ofrece una serie de servicios (lista de amigos, grupos y páginas, muro, fotos, regalos, aplicaciones y juegos) que permiten realizar una serie de actividades a sus miembros, tales como colgar y compartir fotos y videos; conversar en tiempo real con una o más personas (chat); encontrar y agregar amigos; establecer vínculos; enterarse de las actividades que realizan los amigos; crear o hacerse miembro de un grupo, distraerse con una gama de aplicaciones y de juegos, etc.

La apertura de una cuenta en Facebook incluye tres modalidades, la primera es la cuenta individual, la cual se abre con una cuenta de correo electrónico y

permite a un individuo agregar conocidos “amigos” con los cuales compartir información y contenidos relevantes para el usuario. La segunda modalidad “grupo” permite la conglomeración de diferentes cuentas individuales en un simil de foro digital, en donde comparten contenido relacionado a un solo tema. Por último, el tercer tipo de cuenta es la llamada Fanpage, la cual es principalmente para empresas o marcas que están interesadas en difundir los beneficios de sus productos o servicios; la asociación de este tipo de cuentas con las cuentas individuales es principalmente mediante el “like de página”, ya que las Fanpage no incluyen la función de agregar amigos como las cuentas individuales y su interacción depende de los usuarios que otorguen los derechos de recibir información de Fanpage y deseen interactuar con las mismas.

Túñez-López et al. (2011) señala como los efectos beneficiosos del uso de Facebook para las organizaciones las siguientes:

1. Gestionar su reputación digital.
2. Reforzar su comunicación corporativa.
3. Comunicar sus productos y/o eventos.
4. Incrementar su visibilidad.
5. Prevenir las crisis.
6. Atender a sus clientes y/o usuarios.
7. Gestionar el conocimiento mediante el intercambio de opiniones.
8. Comunicar de manera directa.
9. Captar el talento.

Uno de los principales recursos de estudios realizados sobre Facebook radica en el permiso que conceden los usuarios para acceder a los datos de sus perfiles, ofreciendo un nuevo abanico de posibilidades a la hora de realizar investigación. Dentro de esta información, los “me gusta” permiten obtener datos de ocurrencia natural, sobre hábitos y patrones de vida de las personas. Su particularidad radica en la menor reflexividad que plantean en términos generales si lo comparamos

con los datos de perfil. Aun así, este tipo de análisis es complejo debido al gran volumen y la dispersión de los materiales y las dificultades técnicas que entraña (Parra, Gordo, & D'Antonio, 2014).

Es importante señalar que existen dos tipos de “me gusta” o también llamados “like”, el primero al que le llamaremos “like de página” es con referencia al que se da en la parte principal de ésta (Figura 4), al ejecutar la función me gusta de página, aparte de que el usuario indica su preferencia por el tipo de contenido que le puede ofrecer la Fanpage, le otorga el derecho de que las publicaciones emitidas por esta página le sean presentadas en su muro de noticias; es decir, que al cliclear el me gusta de página recibimos cada una de las publicaciones de ésta sin necesidad de entrar a la página principal.

EL segundo tipo de “like” es con referente al que se realiza por publicación (Figura 5), éste representa nuestro agrado por ese mensaje en particular. Cabe resaltar que no es necesario dar “me gusta de página” para que podamos acceder e interactuar con las publicaciones de una Fanpage.



Figura 4. Ejemplo de la posición del botón “me gusta de página” en una Fanpage

Fuente: Facebook site



Figura 5. Ejemplo de la posición del botón “me gusta” en una publicación

Fuente: Facebook site

En el presente trabajo se abordará para el caso de Facebook lo denominado como comunidad de marca, la cual es definida por varios autores. En palabras de García-Ruiz (2005), las comunidades de marca son grupos sociales cuyos miembros comparten una lealtad intensa a una determinada marca comercial, su vinculación no se limita a la defensa de intereses comunes, sino que lo que une a los miembros de este tipo de comunidad es, más bien, la implicación con los valores y actitudes encarnados por la marca de su elección. El mismo autor menciona también que esta implicación une a los usuarios no sólo con la marca, sino también entre sí, según lazos de una solidaridad propia de quienes se reconocen como semejantes.

Palazón, Silicia, & Delgado (2014) por su parte comentan que, en las redes sociales, como en las comunidades de marca tradicionales, los consumidores son los que controlan su interacción con las marcas en la medida en que se unen a una comunidad de marca voluntariamente y una vez que son miembros o seguidores, quedan expuestos a los mensajes de esa marca. La empresa puede

comunicarse con la periodicidad deseada con sus seguidores, lo que le permite construir una relación marca-consumidor más fuerte, sin embargo, desde que se hacen miembros de la comunidad, los seguidores de marcas en redes sociales son algo más que meros receptores pasivos, ya que pueden adoptar una actitud activa en cualquier momento iniciando una conversación con la marca o con otros consumidores y el que una página tenga varios seguidores (“me gusta de página”) no implica necesariamente que alguien le guste la marca o vaya a comprarla es necesario considerar otras variables como la valoración y el grado de afecto que muestren hacia ella.

2.6. Twitter

Twitter es una red social global gratuita online. Desde sus inicios ha ido combinando elementos de blog, mensajes de texto y emisión de contenidos multimedia. Los usuarios escriben mensajes de texto limitados a 140 caracteres, llamados tweets, enviados a cualquiera que haya elegido recibir los tweets de otros perfiles; en cada tweet es posible enlazar con otros medios e incorporar enlaces de vídeo, imágenes y hashtags (Del Fresno García, Daly, & Segado Sánchez-Cabezudo, 2016).

A continuación se describen cada una de las tres principales funcionalidades de Twitter para emitir un mensaje. La primera funcionalidad corresponde a los hashtag, los cuales se pueden ser definidos como etiquetas otorgadas a un tweet o mensaje que ayudan a la búsqueda de mensajes similares o bajo la misma temática, estas etiquetas se caracterizan por que les antecede el símbolo “#”, un ejemplo se puede ver en la Figura 6 en donde se observa un mensaje usando dos hashtag representativos del mezcal.



Figura 6. Visualización de un tweet con los hashtag #Oaxaca y #mezcal

Fuente: Twitter

Las otras dos vertientes son las de retweets y menciones (Figura 7 y Figura 8), el primero se refiere básicamente a compartir en mi cuenta el mensaje de otro usuario, y el segundo como su nombre lo infiere se trata de nombrar a uno o más usuarios con el objetivo de que se percate(n) del mensaje que se está publicando.



Figura 7. Ejemplo de retweet, mensaje compartido de la cuenta @Turibuscdmx

Fuente: Twitter



Figura 8. Ejemplo de mención, mensaje en el que se menciona a la cuenta @armandoideas1

Fuente: Twitter

Algunas de las virtudes de Twitter son citadas por autores como Gamir Ríos (2016), quien dentro de los beneficios para las organizaciones de esta red social destaca seis:

1. Reduce la distancia y favorece el encuentro con su público potencia
2. Fomenta el diálogo con los usuarios pertenecientes a dicho target
3. Permite medir en tiempo real sus actitudes y opiniones
4. Facilita la segmentación de dichos usuarios
5. Aporta credibilidad e influencia mediante la participación en conversaciones
6. Permite alcanzar de forma directa a un gran número de personas interesadas en la propia organización.

Curioso & Carnero (2011) le otorga dos características únicas que lo distinguen de otras aplicaciones Web 2.0, una es la brevedad de los tweets, y la forma cómo estos son compartidos con el resto de usuarios, indica que la naturaleza compacta de los tweets hace que sea muy fácil enviar información (más aún con el reenvío de mensajes y el envío de enlaces web), permitiendo al usuario

exponerse a una gran cantidad de información, para rápidamente seleccionar aquella que le interesa permitiendo a Twitter proveer información casi en tiempo real, a diferencia de otros medios. La otra cualidad es que a diferencia de otras redes sociales Twitter comparte estos mensajes públicamente con el resto del mundo sin necesidad de estar relacionado por una solicitud de amigo (caso Facebook), se destaca que en Twitter uno sigue a las personas de acuerdo a las ideas expresadas en sus tweets, sin necesidad de conocerlos y la relación entre usuarios puede ser asimétrica.

Algunos de los autores que recomiendan el uso de Twitter para la investigación son Barberá & Rivero (2012), quienes mencionan que Twitter ofrece una oportunidad excepcional para el análisis de la opinión pública: los mensajes que intercambian los usuarios pueden contener información valiosa acerca de sus preferencias y su reacción a los mensajes y eventos políticos en un entorno que es accesible para el investigador. A esta idea se anexa los trabajos realizados por Carlos de la Paz, Gómez Masjuán, & Pérez Alonso (2016) y Litche Fragoso & Sánchez Salinas (2014), el primero al decir que Twitter puede funcionar como herramienta para monitorear tendencias, descubrir noticias de última hora o actualizaciones minuto a minuto y el segundo en donde se menciona que el uso de ésta herramienta posibilita varios beneficios para el analista: costos bajos para un volumen mayor de información; acceso rápido y fácil de un número considerable de usuarios; y con información contextual, es decir, se considera la región geográfica de la extracción de los datos.

Para dar paso al siguiente capítulo se retoma lo dicho por López & Sicilia (2016), quienes hacen hincapié en que la red social Twitter ofrece información que permite a las empresas identificar a posibles líderes de opinión y que cualquier individuo con un perfil en Twitter puede “seguir” a otro o a una marca en esta red social. Si un consumidor sigue a una marca implica que todo lo que escriba la marca podrá ser visto por ese consumidor y que está interesado en lo que esta persona pública.

2.7. Los “*influencers*”

Del Fresno García et al. (2016) define a los “*influencers*” como un nuevo tipo de actores sociales independientes, con capacidad de influir en las actitudes de audiencias desde los medios sociales de internet en competencia y cohabitación con los medios de comunicación profesionales. Otra definición es la dada por la ADECEC (2016) quien menciona puede ser cualquier persona que, por su experiencia y credibilidad, tiene el poder de influir en otras porque se identifican con ella, la admiran, confían en ella y respetan su opinión, independientemente de que lo haga a través de los medios on u off line y sin que la plataforma sea condicionante, estos puede generar para las marcas, reputación, notoriedad, y una llamada a la acción.

Castelló Martínez & del Pino Romero (2015) destacan que el fenómeno del influenciador se da a causa de que como espectadores, bajamos la guardia racional ante alguien a quien admiramos, cuya presencia ante un producto nos coloca en un estado de vulnerabilidad tal que de forma no consciente se activan los mecanismos emocionales en los que transferimos todo tipo de bondades al producto, por el mero hecho de una asociación positiva y que de forma ilusoria y guiados por el subconsciente el discurso permea y es retenido en la mente del consumidor.

Los autores antes citados resaltan también la gran oportunidad que se extiende para sacar el máximo partido de este poder de influencia de estos personajes. Sostienen que identificar a los líderes de opinión que pueden ayudar a conectar a una marca de forma natural y espontánea con su público objetivo es algo de valor incalculable en este momento de saturación publicitaria; desde celebridades en un extremo de la escala a amigos en la otra, pasando por gente que deja su impronta en diferentes canales de Internet, desde YouTube a Twitter, el abanico de opciones es amplio, y controlar ese poder de influencia ha dejado de ser una opción para las marcas para pasar a ser casi una asignatura obligatoria.

No obstante no es fácil, ya que tal como lo menciona la ADECEC (2016), saber con quién hay que contar y seleccionar de entre los miles de posibles

colaboradores a los que de verdad estén alineados con los mensajes de la marca, no es una tarea fácil y mientras crece el número de herramientas centradas en automatizar los procesos de identificación de “*influencers*”, no podemos olvidarnos que trabajamos con personas y que el mejor modo de conocer sus intereses y establecer relaciones duraderas con ellos es a través de una comunicación fluida one to one. Resalta que el trabajo con los “*influencers*” requiere tiempo, dedicación y esfuerzo y la clave del éxito estará en que seamos capaces de aportar contenidos de calidad, alineados con sus intereses y ajustados a los mensajes de la marca y que solo de esta manera se podrán establecer vínculos que generen beneficios para todos, desde la marca al “*influencer*”, pasando por la agencia.

La ADECEC (2016) señala como principales atributos a cumplir de un “*influencer*” su capacidad para movilizar las opiniones, crear reacciones ante un tema concreto y mover a la acción. Por ello, menciona, es importante analizar la audiencia potencial a la que pueden llegar cuando hablan sobre la materia que nos ocupa y buscar un alto nivel de participación en la conversación que se genere sobre dicho tema. La misma fuente sostiene también que en la búsqueda de “*influencers*” se utilizarán herramientas que nos permitan conocer los datos cuantitativos como el número de seguidores, la frecuencia de publicación o la viralización de contenidos en otras redes sociales, pero son los valores cualitativos como la compatibilidad con la marca, la identificación o el tono de los mensajes los que nos proporcionarán una selección eficaz.

Castelló Martínez & del Pino Romero (2015) por su parte destaca la importancia de la credibilidad, ya que la comunicación en las redes sociales tiene una naturaleza que dista de la publicidad convencional en las que es el consumidor es el que de forma voluntaria se acerca al contenido por el atractivo de éste, el lenguaje adoptado por el “*influencer*” no debe denotar agresividad comercial, de lo contrario, la tendencia es la de rechazar el mensaje, el cual debe formar parte del contexto de una historia con entidad, con peso, que invite a la interacción y debe cuidarse al máximo la acción realizada para que no revierta negativamente

tanto en la empresa como en el “*influencer*”. El ejemplo claro se muestra en la Figura 9 en donde un mensaje publicitario para una marca de móviles se ve afectada por un descuido de la figura pública que lo emitió desde un telefono de la competencia.



Figura 9. Mensaje publicitario de Samsung Galaxy enviado desde iPhone

Fuente: Twitter

Haciendo referencia a la compra de Tweets se cita lo mencionado por Castelló Martínez & del Pino Romero (2015) quienes dicen que la compra de tweets es una práctica muy empleada por las marcas, ya que permite generar mucho ruido en Twitter y asegura un gran alcance y difusión gracias a los miles o millones de usuarios que siguen a un personaje conocido. Los autores señalan que el precio medio de un tweet está entre 20 y 30 mil euros, pero resulta de vital importancia que el mensaje del “*influencer*” encaje con el target así como adoptar en el tweet el tono habitual del influenciador.

Por todo lo anterior, y para cerrar este apartado se citan las ideas de Serrano-Puche (2012), quien sostiene que es deseable que desde el mundo académico surjan propuestas teórico empíricas de medición de la influencia personal digital, acordes con el rigor y seriedad exigibles a las investigaciones científicas.

2.8. Conceptos de Big Data

En palabras de Tabares & Hernández (2014) el constante uso de las tecnologías ha traído consigo un crecimiento explosivo en la cantidad de datos, los cuales son generados a grandes velocidades y en distintos formatos. Según los autores antes citados, a partir de este aumento de información se da la necesidad de extraer de ella, patrones, tendencias y/o conocimiento, de forma rápida y eficiente, para lo cual, los métodos tradicionales han tenido que evolucionar en busca de rendimiento y escalabilidad. Los mismos autores añaden que el gran contenido de valor que genera este tipo de información está permitiendo a las organizaciones una mejora en la toma de sus decisiones, lo que conlleva a la obtención de ventajas competitivas en los diferentes campos de acción.

Secundando esta idea se encuentra Maté Jiménez (2014), quién menciona que en nuestros días, es un hecho la ingente cantidad de cuestionable información que se genera cada segundo en nuestro planeta y que dicha información que puede ser estructurada, semiestructurada o no estructurada también puede aportar enorme valor a cualquier entidad o puede suponer un consumo excesivo de recursos humanos, informáticos, etc. Ortiz Morales et al. (2016) por su parte añaden que en la actualidad no solo representa una ingente cantidad, variedad y volumen de información, sino el tema de “moda” que aparece diariamente en los periódicos y revistas; igualmente, los sectores económicos, las empresas y consultoras más importantes intentan mostrar sus posibles aplicaciones y generan informes frecuentes al respecto.

Es en este contexto que entra en juego lo denominado como Big Data que tal como menciona Tabares & Hernández (2014) se sitúa como el nuevo paradigma de análisis de datos, añadiendo que a pesar de sus tres características

principales (velocidad, variedad y volumen) también incluye un amplio espectro de técnicas, tecnologías, métodos y paradigmas no convencionales que apoyan la solución de problemas relacionados con datos de una forma diferente y, generalmente, más adecuada que los métodos tradicionales.

De acuerdo con Ríos (2014), lo que consigue el Big Data es generar valor agregado mediante la generación de información nueva para tomar decisiones, con lo cual es posible reducir costos o aumentar ingresos, señala también que hoy día, esta nueva información incluso genera nuevos negocios. De tal manera que como menciona Díaz, Osorio, Amadeo, & Romero (2013), esta tecnología se encuentra en pleno desarrollo, encontrando soluciones “open source” que es necesario estudiar e investigar en forma sistemática para obtener resultados comparativos que sean de utilidad.

Como señaló hace poco el diario español ABC, uno de los sectores que más interés está suscitando en las empresas es el estudio y procesamiento de grandes e ingentes cantidades de datos, el uso del Big Data en las compañías será cada vez más importante en los próximos años (Colle, 2013).

2.9. Definición de Big Data

En el año 2010 el término Big Data era prácticamente desconocido. A mediados de 2011 se convertía en una palabra que aparecía con frecuencia entre las últimas tendencias. Lo que parecía iba a ser una palabra de moda y, por ende pasajera, se ha convertido en todo un área de interés enorme para las empresas, organizaciones y administraciones públicas; generando un mercado profesional emergente, que es el de los científicos de datos, y abriendo nuevas líneas de investigación y nuevas revistas específicas en el contexto académico como Big Data o Big Data Research (Maté Jiménez, 2014).

Del Fresno García et al.(2015) explican que el Big Data son los datos que se producen de manera incesante ubicua y masiva y que este tipo de datos ofrece

oportunidades y desafíos a los investigadores, donde lo más interesante reside, no en las cantidades de datos, sino en qué podemos hacer con estas grandes cantidades de datos que no se puede hacer con pequeñas cantidades.

Ortiz Morales et al. (2016) por su parte describen al Big data como la estrategia organizacional, tecnológica y táctica que facilita capturar, almacenar, procesar y analizar los grandes volúmenes de datos generados en toda la cadena de valor de la empresa, que varía según el sector, la industria y las necesidades de cada compañía. Agregan también que muchas empresas requieren almacenar todos los datos (volumen) para hacer uso de ellos en cualquier momento del tiempo; incluso, pueden requerirlos incorporando a su operativa poco a poco y otras los utilizan para la toma de decisiones en tiempo real (velocidad) y así ofertar a sus clientes productos según los intereses presentados en el momento, con el fin de incitar a la compra inmediata. Igualmente, otras organizaciones guardan los diferentes tipos de datos para manejarlos en su estructura de negocio (variedad).

Con el término Big Data se hace referencia a la tendencia del avance de las tecnologías que han abierto las puertas hacia un nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos que llevaría demasiado tiempo cargarlos en una base de datos relacional para su posterior análisis. Por lo tanto, el Big Data se aplicará para toda aquella información que no pueda ser procesada por los métodos tradicionales (López García, 2012).

Tabares & Hernández (2014) sin embargo señala que el Big Data no son sólo datos, ellos mencionan que Big Data también comprende todo el espectro de técnicas, métodos, herramientas y tecnologías alternativas, que permiten resolver problemas que involucran cierta complejidad, de una forma más eficaz y eficiente que los métodos tradicionales.

Para finalizar retomamos lo dicho por Ferrer-Sapena & Sánchez-Perez (2013) quienes dicen que Big Data se refiere a grandes conjuntos de datos que tienen algún valor para las empresas, mencionan que en algunos casos han sido

generados por los científicos, aunque los más conocidos proceden de movimientos de Facebook, Twitter y Google y que el tratamiento suele realizarse en tiempo real utilizando tecnologías cada vez más potentes para la gestión de bases de datos, la visualización y el análisis.

Los mismos autores ejemplifican el uso de esta herramienta con el seguimiento de la campaña de marketing de la película Amanecer 2 y la campaña del candidato Barak Obama en las elecciones presidenciales de EUA de 2012.

2.10. Tipos de datos en el Big Data

Los datos generados en grandes volúmenes por fuentes como web y medios sociales, de máquinas, de transacciones o generados por interacción humana, se pueden clasificar básicamente en tres grupos: estructurados, semiestructurados y no estructurados.

Ortiz Morales et al. (2016) define a los datos estructurados como aquellos con formato o esquema fijo que poseen campos fijos, ejemplifica a esta categoría con hojas de cálculo y ficheros. Para el caso de datos semiestructurados los describe como los que no tienen formatos fijos, pero contienen etiquetas y otros marcadores, y los ejemplifica como texto de etiquetas XML y HTML. El mismo autor destaca que los datos no estructurados son los más abundantes y los que más contenido informativo pueden tener para las organizaciones, este tipo de datos los describe como datos sin tipos definidos que se almacenan principalmente como documentos u objetos sin estructura uniforme, como ejemplos de este tipo de datos nos señala a audio, vídeo, fotografía, formatos de texto libre (e-mails; SMS, artículos; libros; mensajería de tipo WhatsApp, Viber; etc.).

Es de suma importancia destacar que en esta categoría encontramos el tipo de datos que se analizarán en la presente investigación. López García (2012) resalta que el 80% de los datos de una empresa son no estructurados, y que gracias al

Big Data ahora se pueden analizar y obtener información útil para las organizaciones. No obstante, cabe recalcar que el Big Data no solo se centra en los datos no estructurados, sino que los hace también con los no estructurados y los semiestructurados.

2.11. Características del Big Data (las 3V)

El concepto de Big Data es caracterizado por tres términos o variables conocidas como las 3Vs: Volumen, Velocidad y Variedad, donde Volumen hace referencia a la cantidad de datos generados continuamente en un espacio de tiempo determinado; Velocidad, a la rapidez de entrada y salida con la cual éstos fluyen a través de distintos canales; y Variedad, a los diferentes formatos y fuentes en que éstos se encuentran. Las 3 Vs corresponden a la definición base de lo que representa Big Data para la mayoría de autores expertos en el tema (Tabares & Hernández, 2014). No obstante, los mismos autores antes citados señalan que existen otras V de las cuales se pueden echar mano para estudios o aplicaciones del Big Data, entre estas señalan a la Veracidad, Validez y Volatilidad, donde Veracidad, hace referencia a los datos libres de ruido, a través de los cuales se puede hacer minería y análisis, es uno de los retos más grandes en Big Data; Validez, a la correctitud y exactitud de los datos para el uso que se quiere dar a los mismos; y Volatilidad, al tiempo que los datos serán almacenados y por cuánto tiempo serán válidos. Y por último resalta una cuarta V, la cual se refiere al Valor que puede obtener a partir de los datos mediante procesos analíticos. Esta V dice, puede considerarse como una de las más importantes y se presenta como la oportunidad de extraer información valiosa mediante el procesamiento de datos complejos que las organizaciones generan en forma de datos estructurados y no-estructurados, con la finalidad de convertirlos en decisiones y/o conocimiento.

A pesar de que son varios los autores que aportan más características sobre la clasificación de las V del Big Data, en general se toman tres como las principales,

retomando el trabajo de López García (2012) se presenta a continuación una descripción de cada una:

1. **Volumen:** Suele utilizarse como sinónimo de Big Data. A pesar de ser uno de los aspectos más llamativos, no es el único. El reto relacionado con el volumen de datos se ha puesto de manifiesto recientemente, debido a la proliferación de los sistemas de información e inteligencia, el incremento del intercambio de datos entre sistemas y dispositivos nuevos, nuevas fuentes de datos, y el nivel creciente de digitalización de los medios de comunicación que antes sólo estaban disponibles en otros formatos, tales como texto, imágenes, videos y audio.
2. **Velocidad:** Se asocia con la proliferación de nuevas fuentes de datos, y la necesidad de utilizar estos datos más rápidamente. Los dispositivos que generan datos a intervalos más largos, tales como los teléfonos inteligentes, también terminan generando corrientes constantes de datos que necesitan ser ingeridos rápidamente. Por otro lado, todos estos datos tienen poco o ningún valor si no se convierten rápidamente en información útil.
3. **Variedad:** Los grandes volúmenes de datos incluyen cualquier tipo de datos, estructurados y no estructurados como texto, datos de sensores, audio, vídeo, secuencias de clic o archivos de registro, entre otros. Al analizar estos datos juntos se encuentra información nueva. Para explicar esta característica, la mejor opción es imaginar la creciente cantidad de información que almacena Facebook sobre sus usuarios y lo diversa que es esta. En su base de datos se puede encontrar la edad, el sexo o el país de millones de personas. Con Big Data esto es posible de hacer.

EL desafío del Big Data consiste en como manipular, administrar, almacenar, buscar y analizar grandes volúmenes de datos. Con el termino Big Data hacemos referencia a este gran desafío de las empresas consistente en el tratamiento y análisis de grandes repositorios de datos (López García, 2012).

Pare cerrar este apartado se retoma lo dicho por Fabrizio (2014) que menciona que para emprender este viaje al futuro con garantía de éxito hay, además de los datos por supuesto, tres ingredientes que son imprescindibles:

- a. Mentalidad 'Big Data' (creatividad y espíritu innovador)
- b. Talento analítico, tanto en profesionales internos como en socios (imprescindible para convertir datos masivos en bruto, del todo inertes, en información de alto valor)
- c. Tecnología no disruptiva (herramientas de productividad e integración con los activos informacionales existentes en el momento de abrir este nuevo frente)

2.12. El desafío del análisis de datos

Las características intrínsecas del Big Data (variedad, volumen, y velocidad), tal y como menciona Blázquez-Ochando (2016), hacen recurrente la cuestión del cómo hacer los métodos de gestión, procesamiento, y extracción de los datos. Aunado a esto, el mismo autor indica que el método de extracción y su procesamiento varía en función de su contexto de aplicación, género, formato e, incluso fuente de procedencia; es decir, no existe una metodología consensuada ni exclusiva para el manejo de este tipo de datos. Incluso IBM (2013) en un trabajo titulado "Las cinco mejores formas de abordar Big Data" hace referencia a una serie de preguntas clave a realizarse para el tratamiento del Big Data de las cuales podemos destacar: "¿Cómo explorar los datos en volúmenes de información grandes y complejos?, ¿Cómo separar el "ruido" del contenido útil? y ¿Cómo aplicar las perspectivas y llevar a cabo acciones?.

A pesar de que autores como Huertas Roig & Mariné Roig (2014) señalan que es necesario la creación de una metodología específica para el análisis comunicativo de los social media, adaptado a sus especificidades y sus necesidades evaluativas, que resulte práctico; otros autores como Colle (2013)

destacan que no es poca cosa integrar grandes cantidades de datos de una variedad de fuentes y es aún más difícil averiguar exactamente cómo traducir esa información en más visitas (en tiempo real), clientes por cliente. A esta idea le secunda Rodríguez Mareño (2014) quien dice que saber lo que los usuarios publican dentro del perfil de una determina marca es muy importante pero en marcas donde los seguidores son un gran número es muy difícil considerarlos a todos. Serrano-Puche (2012) por su parte señala que una de las paradojas de la comunicación digital es que tenemos acceso a multitud de datos, pero con frecuencia nos falta el contexto y la capacidad de interpretación.

En resumen, podemos decir que obtener la V de valor a partir de las otras V del Big Data es una tarea nada trivial que representa un gran reto para toda persona decidida a obtener conocimiento a partir de herramientas de Big Data. Litche Fragoso & Sánchez Salinas (2014) mencionan que los retos pueden agruparse en el desarrollo de metodologías de cálculo, problemas de privacidad y legalidad, desarrollo de capital humano, y desarrollo de asociaciones para el análisis; además, resaltan que el mayor reto no está en recoger los datos de internet, sino que está en la siguiente etapa, la del procesamiento que se desenvuelve en otras tres etapas: la extracción de la información deseada, depuración de la misma, y conversión en información útil para la elaboración de estadísticas.

Los mismos autores antes citados destacan que se deben considerar las siguientes limitaciones sobre el uso productivo de la información: los datos distan mucho de ser claros y contundentes, y no hay un registro sistemático y estable temporalmente. Sin embargo, agregan que los problemas referidos ordinariamente son parte de las batallas que enfrentan los investigadores del comportamiento social.

Colle (2013) por su parte dice que hay que saber primero cuáles son los datos importantes a acumular, investigar cuáles son las herramientas adecuadas para su análisis, juntar una primera cantidad (suficiente para un análisis) y sacar las lecciones de la experiencia; resalta que el pequeño secreto del Big Data es que

ningún algoritmo puede decirte lo que es relevante o develarte su significado, los datos entonces se convierten en otro problema que resolver.

2.13. La minería de datos en el Big Data

Una de las herramientas de las cuales podemos echar mano para el análisis de grandes volúmenes de información es la minería de datos o *Data Mining* que en palabras de Colle (2013) consiste esencialmente en aplicar técnicas que permiten descubrir y explicitar la información oculta en los megabytes de las bases de datos. López García (2012) por su parte la describe como una herramienta muy valiosa para las organizaciones que ofrece herramientas que ayudan a la toma de decisiones como pueden ser las agrupaciones o predicciones. No obstante, señala que a pesar de todas las ventajas que nos proporciona la minería de datos con la información que proporcionan las empresas, existen también una serie de inconvenientes como la falta de privacidad de los datos, errores en los modelos y patrones obtenidos, dificultades de escalabilidad, y manejo de software; todas estas limitantes hacen que la minería de datos tenga que mejorar.

Una variante de la minería de datos es la determinada por Romero Frías, Vaughan, & Rodríguez Ariza (2011) como webmetría (tipo de análisis que se utilizó para el análisis de palabras clave, ver 3.2) la cual describen como la gran amalgama de información disponible que nos permite considerar a la web como una enorme base de datos de naturaleza heterogénea para valorar la aplicación de procedimientos destinados a identificar patrones comunes en la información; ejemplo de ésta, sería la investigación basada en palabras clave: realizando análisis de impacto de ideas, conceptos, marcas, organizaciones, empresas, etc., a partir del recuento y análisis de las páginas web proporcionadas por los motores de búsqueda al formular una consulta tradicional.

Otra variante es la señalada por Del Fresno García et al. (2015) quienes hablan de la minería de datos en *Social Media Mining* (SMM), la cual se utilizó en la

presente investigación y en donde el Análisis de Redes Sociales (ARS) toma relevancia. Los autores utilizan el SMM en Twitter porque por medio de sus conexiones explícitas significativas permite hacer visible lo que hasta ahora permanecía invisible: la estructura en red de climas de opinión y la identificación de aquellos nodos que tienen una influencia social desproporcionada en su red. Todo lo anterior con el fin de descubrir patrones significativos a partir de las interacciones sociales en los medios sociales de internet.

2.14. Importancia del análisis de redes sociales y conceptos de ARS

El presente apartado tiene como finalidad el contextualizar la importancia de la herramienta usada sobre todo para el análisis de Twitter (ver sección 3.4), así como para definir los indicadores usados para la presente investigación. Se retoman en particular las ideas de dos trabajos (Del Fresno García et al., 2015; Garrido, 2001) para en un primer momento resaltar la importancia y la elección del Análisis de Redes Sociales como herramienta de análisis, que en sinergia con la minería de datos nos permitió abordar los retos del Big Data.

Garrido (2001) empieza resaltando que el ARS ha venido creando un aparato metodológico y técnico que constituye la base de un nuevo paradigma interdisciplinar que, a través, del uso de modelos teóricos de carácter formal expresados en términos matemáticos y del análisis sistemático de datos de carácter empírico, ha permitido hacer operativo el concepto de estructura social.

Por su parte Del Fresno García et al.(2015) describen al ARS como un método científico orientado a la optimización de la investigación de los medios sociales, sustituyendo a los enfoques anecdóticos (número de amigos, seguidores, gustos, etc.) y aportando resultados significativos en la identificación de los actores claves en la comunicación dentro de las plataformas sociales online. Le añade también el calificativo de una potente herramienta para la representación de los modelos de difusión de la información.

Los mismos autores destacan que para poder comprender a los “*influencer*” como nodos clave dentro de una red tenemos que resolver antes cómo pueden ser identificados entre miles de actores que están participando dentro de esas redes sociales.

Resaltan que esas densas redes de interacciones han permanecido invisibles, por lo que es difícil identificar los atributos generales de los “*influencer*”, sin embargo, el ARS ofrece ventajas singulares para la identificación de los “*influencer*” en comparación con otros métodos, entre estas ventajas se encuentran:

- a) En el ARS, el foco está en las relaciones y no en los atributos, pudiendo identificar la dirección de las relaciones.
- b) Determinar la fuerza de las relaciones.
- c) Segmentar e identificar los grupos, subgrupos y comunidades.
- d) Medir la posición estructural que los actores tienen dentro de una red.

Una red social consta de un conjunto finito de actores y de uno o más tipos de relaciones (Aguilar-Gallegos et al., 2016). El análisis de redes se centra en las relaciones entre los actores, y no en los actores individuales y sus atributos. Esto significa que los actores a menudo no son muestreados de forma independiente, como en muchos otros tipos de estudios (más frecuentemente en encuestas) (Hanneman, 2000a).

Para la presente investigación se utilizó un indicador de red (Modularidad) y dos indicadores de nodo; uno de primer orden, que en palabras de Aguilar-Gallegos et al. (2016), es aquel que vincula directamente a dos nodos a través de un solo paso y; uno de segundo orden, que según los mismos autores, son los que toman en cuenta los vínculos indirectos, que logran vincular a dos nodos a través de dos o más pasos. Los indicadores fueron la Centralidad de grado y el PageRank, respectivamente.

Aguilar-Gallegos et al. (2016) resaltan la importancia de no solo considerar los vínculos directos si no también los vínculos indirectos en un proceso de intercambio de información entre diversos actores.

La modularidad, es la fracción de los enlaces que caen dentro de grupos dados menos el valor esperado que dicha fracción hubiese recibido si los enlaces se hubiesen distribuido al azar. El valor de la modularidad se encuentra en el intervalo (-.5 a 1). Es positivo si el número de enlaces dentro de los grupos supera el número esperado sobre la base de pura casualidad. Para una determinada división de vértices de la red en algunos módulos, la modularidad refleja la concentración de los nodos dentro de los módulos en comparación con distribución al azar de los enlaces entre todos los nodos, independientemente de los módulos (Arenas, Fernández, & Gómez, 2008). Muff, Rao, & Caflisch (2005) destacan que muchas redes complejas tienen una estructura modular subyacente, es decir, subunidades estructurales (comunidades o grupos) caracterizada por nodos altamente interconectados. Describen que la modularidad se ha introducido como una medida para evaluar la calidad de clusterización que tiene una visión global, mientras que en muchas redes del mundo real racimos están vinculadas principalmente a nivel local entre sí (conectividad clúster local).

Por otro lado, se encuentra la centralidad de nodo de la cual Hanneman (2000b) destaca que los actores que tienen mayores vínculos con otros actores puede que tengan posiciones ventajosas. Además, dado que disponen de muchos vínculos, pueden tener acceso y pueden conseguir más del conjunto de los recursos de la red, tener más vínculos les posibilita a menudo ser terceros y permitir intercambios entre otros, pudiéndose beneficiar de esa posición. Así que señala que una muy simple, pero a menudo efectiva, medida de centralidad de un actor y de su poder potencial es su grado.

Hanneman (2000b) puntualiza que si un actor recibe muchos vínculos (grados de entrada) a menudo se dice que es prominente o de prestigio, o sea muchos otros actores buscan entablar vínculos con él, y esto puede indicar su importancia; por

otro lado, los actores que inusualmente tienen un alto grado de salida son actores que son capaces de intercambiar con muchos otros, o hacer a muchos otros conscientes de sus puntos de vista.

Cook (2012) por su parte nos describe que el PageRank de un nodo es la cantidad esperada de tiempo dedicado a visitar ese nodo en una caminata al azar en la red. Señala que el parámetro alfa añade una pequeña probabilidad de moverse entre dos pares de nodos, que permite que la métrica que se calcula incluso para redes que no están conectados firmemente y cuando alfa es igual a cero, PageRank es igual a la centralidad del vector propio estándar.

Tomando como base lo mencionado por Aguilar-Gallegos et al. (2016) sobre que durante el análisis de la red de intercambio de información y conocimiento para la innovación se podrían interpretar de manera incorrecta los resultados de cada indicador, es que es importante desarrollar un breve apartado en donde se pueda observar un ejemplo de la implementación de los indicadores propuestos para esta investigación.

En la Figura 10 el tamaño de los nodos está en función de su indicador de PageRank (Cuadro 1), por tal motivo se puede observar que a pesar de que el nodo B no es el nodo con mayor grado de entrada, sí se posiciona como el nodo más central ya que los nodos que lo conectan (A, C, y G) tienen prestigio a un segundo orden. Por el contrario, podemos ubicar al nodo A cuyo grado de entrada es de 5 (el mayor) pero los nodos que lo conectan no trascienden a un segundo nivel por tal motivo se encuentra en tercer lugar con respecto al PageRank antecedido por el nodo C, cuyo grado de entrada es de dos, pero lo conectan nodos (D y C1) que a su vez son conectados por otros, lo cual les da un mayor prestigio, en el PageRank no solo importa las conexiones si no el prestigio de éstas.

El algoritmo de modularidad, siendo calculado a una resolución de 1, nos dio como resultado una modularidad de 0.62 y la obtención de cuatro grupos (representados en escala de colores), en el primer grupo se encuentran los nodos

C1, C2, C3, C4, C5, y C6; el segundo grupo lo conformaron los nodos A, J, L, I, M, y K; el tercer grupo (más pequeño) estuvo compuesto por los nodos G, H, O, y Z; por último, el cuarto conglomerado abarcó a los nodos B, C, D, N, F, y P.

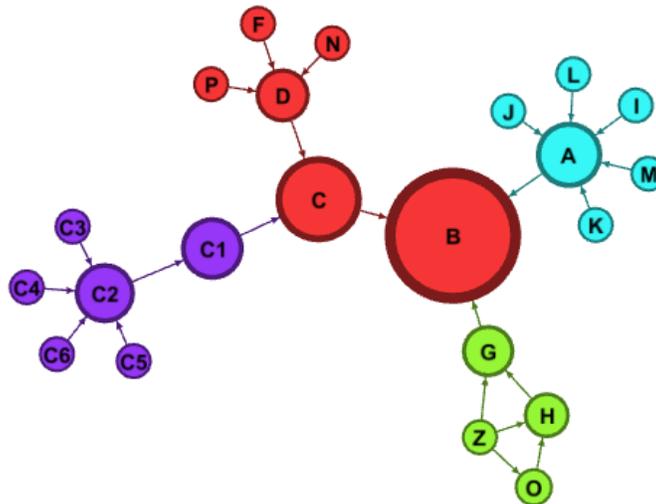


Figura 10. Ejemplo de red con 22 nodos y 23 aristas para la representación de indicadores

Fuente: Elaboración propia

Cuadro 1. Indicadores de red del ejemplo

Nodo	Grado de entrada	Grado de salida	Centralidad de Grado	Modularidad	PageRank
B	3	0	3	3	0.248
C	2	1	3	3	0.132
A	5	1	6	2	0.086
C1	1	1	2	1	0.078
C2	4	1	5	1	0.072
D	3	1	4	3	0.058

Nodo	Grado de entrada	Grado de salida	Centralidad de Grado	Modularidad	PageRank
G	2	1	3	4	0.054
H	2	1	3	4	0.039
O	1	1	2	4	0.021
C3	0	1	1	1	0.016
C4	0	1	1	1	0.016
C5	0	1	1	1	0.016
C6	0	1	1	1	0.016
I	0	1	1	2	0.016
J	0	1	1	2	0.016
K	0	1	1	2	0.016
L	0	1	1	2	0.016
M	0	1	1	2	0.016
F	0	1	1	3	0.016
N	0	1	1	3	0.016
P	0	1	1	3	0.016
Z	0	3	3	4	0.016

Fuente: Elaboración propia

3. METODOLOGÍA

3.1. La pirámide de la información

En este apartado se pretende puntualizar acerca de cada uno de los pasos que se siguieron para la obtención de los resultados. Es importante destacar que por la naturaleza de la investigación se dará en un primer momento una explicación general acerca de la orientación que tuvo y en apartados posteriores se especificará la metodología para cada objeto de estudio (Buscador de Google, Facebook y Twitter).

La metodología utilizada para cada caso busca transformar los datos brutos extraídos de las diferentes plataformas web en conocimiento que contribuya al posicionamiento de algún producto agropecuario, en nuestro caso el mezcal. Se sigue el camino de la pirámide de la información, la cual es descrita por García-Marco (2011), quien toma en cuenta que para llegar a la punta de la pirámide, en donde se encuentra la sabiduría, se debe pasar previamente por cuatro niveles.

El primer nivel son los datos, los cuales define el autor antes citado como la unidad de procesamiento que reclama o va a utilizar un esquema dado; menciona que un dato puede ser un texto o una palabra, un carácter, una fracción del carácter o incluso un punto, dependiendo de qué entidad es la que vaya a ser usada en el proceso en cuestión. El segundo nivel está compuesto por la información, la cual describe la relación entre datos, esto es al aspecto sistémico relacional de las estructuras cognitivas que llamamos conocimientos. Después sigue con el conocimiento, indicando que se caracteriza por la utilidad o sentido de la información, los por qué y para qué de la misma. La pirámide termina con la sabiduría, la cual describe como la nueva información creada dentro del sistema cognitivo a partir de los conocimientos adquiridos.

En la misma tesitura otro trabajo, de hecho previo al mencionado anteriormente, que hace uso del concepto piramidal de la información es el de Muñoz Rodríguez, Aguilar Ávila, Rendón Medel, & Altamirano Cárdenas (2007), quienes a diferencia

de García-Marco (2011) colocan en el último peldaño a la innovación la cual se refiere a los cambios que se adoptan con base al conocimiento adquirido o generado. Para llegar hasta este grado parten de los datos, los cuales describen como elementos discretos, huérfanos de contexto, mencionan también que cuando estos datos son contextualizados, se convierten en información y que sólo cuando esta información se configura según pautas más amplias y de mayor nivel y se analizan las relaciones causales, se llega a lo que pudiera denominarse conocimiento.

Otro autor que abona a los conceptos es López García (2012) quien los define de la siguiente manera: Dato, es un elemento primario de información que por sí solos son irrelevantes para la toma de decisiones. La manera más clara de verlo es con un ejemplo: un número de teléfono o un nombre de una persona, son datos, que sin un propósito o utilidad no sirven para nada. La información, es un conjunto de datos procesados y que tiene relevancia o propósito y que, por lo tanto, son de utilidad para las personas que la utilizan para la toma de decisiones. El conocimiento, es una mezcla de experiencias, valores, información y know-how que aplicarán los conocedores de éste para la toma de decisiones. El autor señala que donde realmente entra en juego el Big Data es en el proceso de encontrar la información la cual puede ser transformada en conocimiento entre esas grandes cantidades de datos recolectadas por las organizaciones y no en cómo se recolectan esos datos. La visión optimista de un Big Data perfecto será aquel en el que las empresas sean capaces de obtener datos de cualquier fuente, aprovechar esos datos y, obtener la información que se convertirá en conocimiento útil para la organización permitiendo incorporar todas las ventajas anteriormente mencionadas (López García, 2012).

Partiendo de las premisas anteriores antes citados se propone en la Figura 11 una imagen de la pirámide de la información; a la cual, a diferencia de lo descrito anteriormente, se le añade un paso extra, desde nuestro punto de vista necesario, el de la extracción de datos, el cual es un proceso complejo que se puede definir como los mecanismos y técnicas por los cuales se obtienen los

datos registrados en un periodo de tiempo de una plataforma web; dicho proceso es de suma importancia y no se puede dejar de lado. Si bien a pesar de que pareciera que solo los tres últimos escalones de la plataforma que tienen que ver con el conocimiento, la sabiduría y la aplicación de la sabiduría (innovación) son más relacionados al proceso cognitivo del “sentido humano” y que los pasos de la extracción a la información son más relacionados a procesos mecanizados, muy apoyados por herramientas tecnológicas, es importante señalar que, por las características del Big Data, para pasar de la extracción a la información los datos requieren de un proceso de edición, filtrado o depuración, y categorización de los mismos, procesos nada triviales que requieren tanto del uso y conocimiento de procesos y herramientas tecnológicas así como de la lógica y razonamiento cognitivo (sentido humano).

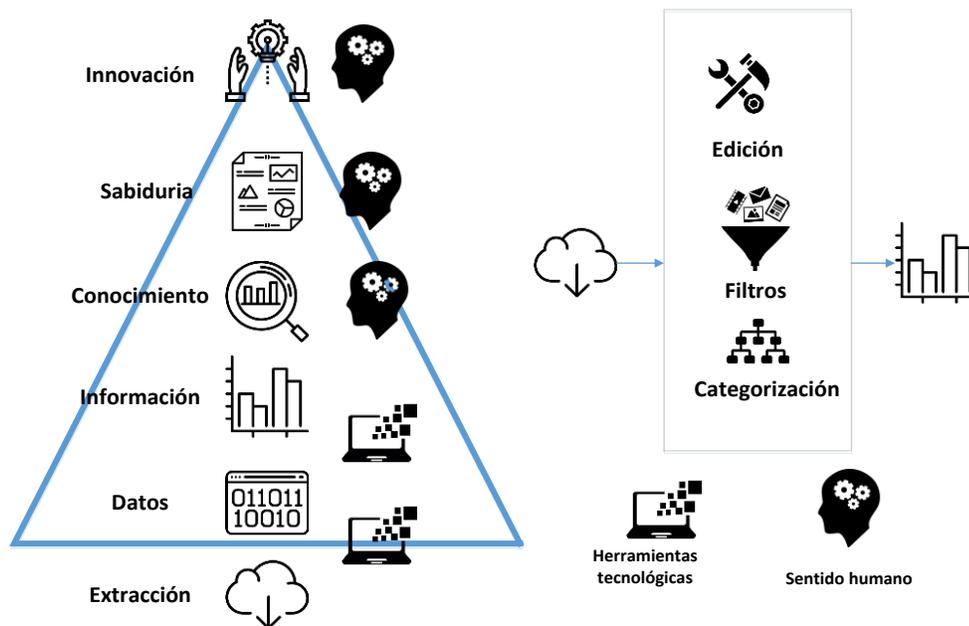


Figura 11. Pirámide de la información de las plataformas digitales

Fuente: elaboración propia a partir de Muñoz Rodríguez et al.(2007), García-Marco (2011), y (López García, 2012).

A continuación, se describen los procesos metodológicos que se siguieron para cada una de las plataformas. Cabe señalar que en este trabajo dichos procedimientos buscaron pasar desde la extracción de datos a la contextualización y comparación de los mismos.

3.2. Extracción depuración y análisis de datos para Google

En la Figura 12 se observa que para llegar a la generación de conocimiento a partir del análisis de palabra clave en Google se requirió recorrer ocho pasos, los cuales se describen a continuación.



Figura 12. Proceso de análisis de datos para Google

Fuente: elaboración propia

1. Obtención de combinaciones de palabras clave

En el caso de Google, el primer eslabón de la pirámide de la información concierne a las palabras clave (*keywords*) que la gente utiliza para hacer una búsqueda al igual que el número de búsquedas promedio mensual de cada combinación de palabras.

Como se muestra en la Figura 13, Google por defecto sugiere mediante su función de autocompletar diversas combinaciones de palabras a partir de un texto añadido en el buscador, estas sugerencias están basadas en búsquedas relacionadas que otros usuarios hayan hecho con el fin de agilizar la consulta. Haciendo uso de la herramienta en línea hypersuggest se obtuvieron la mayor cantidad de sugerencias posibles en torno a la palabra del producto (en este caso, mezcal).

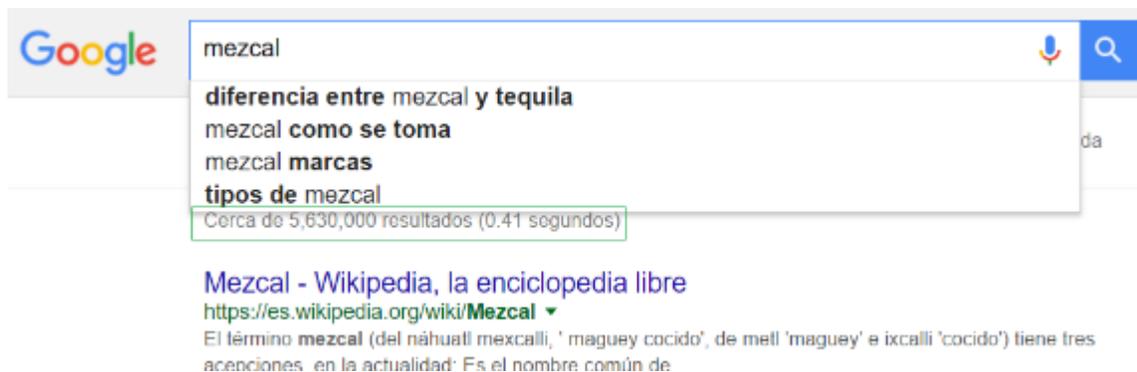


Figura 13. Función de autocompletado del buscador

Fuente: Buscador de Google

2. Obtención del número de búsquedas mensuales

Posterior a ello, las *keywords* extraídas fueron sometidas a la herramienta de Google Adwords para obtener el número de búsquedas mensuales hechas en México que se dieron en el periodo de enero a diciembre del 2015 (máximo

mostrado de un año). Cabe mencionar que en este proceso existió un primer filtro entre las *keywords* que tuvieron y no registro de búsquedas mensuales. Tal y como se muestra en la Figura 14 la herramienta Google Adwords nos permite obtener el promedio mensual de volumen de búsquedas de las palabras elegidas, también nos permite hacer filtros por idioma, escala de tiempo y región, además de otorgarnos una serie de sugerencias con relación a las palabras usadas en el análisis.

The screenshot shows the Google AdWords interface for the Keyword Planner tool. The main content is a table of keyword suggestions for 'mezcal'. The table has the following columns: Palabra clave (por relevancia), Prom. búsquedas mensuales, Competencia, Oferta sugerida, % de Impr. d anuncio, and Agregar al plan. The data rows are as follows:

Palabra clave (por relevancia)	Prom. búsquedas mensuales	Competencia	Oferta sugerida	% de Impr. d anuncio	Agregar al plan
mezcal tequila	100 – 1,000	Baja	MX\$8.25		»
mezcal worm	100 – 1,000	Baja	–		»
tequila mezcal	10 – 100	Baja	–		»
mezcal cocktails	100 – 1,000	Baja	–		»
mezcal brands	10 – 100	Baja	MX\$13.64		»
tequila with worm	10 – 100	Baja	–		»
el mezcal menu	10 – 100	–	–		»
mezcal restaurant	10 – 100	Baja	–		»
buy mezcal	10 – 100	Baja	MX\$5.90		»

Figura 14. Captura de pantalla de la herramienta de Google Adwords

Fuente: Google Adwords

3. Categorización de las palabras

Posteriormente las palabras se clasificaron en grupos arbitrarios basados en la apreciación de la afinidad semántica que mostraron las palabras, lo anterior con la finalidad de eliminar aquellos grupos que, a pesar de incluir la palabra “mezcal”, esta sólo apareció circunstancialmente o que no son opciones factibles de posicionamiento para una marca (ejemplo canciones o marcas establecidas).

4. Perfil de la audiencia

Una vez obtenidas la serie de palabras clave que se consideran importantes para el contenido, se hizo uso nuevamente de la herramienta de Google Adwords y se añadió la herramienta de Bing Ad Intelligence para obtener una aproximación del perfil del usuario de internet que hace uso de estas palabras para buscar contenido relacionado. Esta aproximación consistió en identificar por entidad federativa la procedencia de las búsquedas, al igual que la identificación por edades de los usuarios que las efectúan.

5. División de las palabras por cuadrante

Una vez que se tuvieron las *keywords* depuradas con su correspondiente número de búsquedas mensuales, a éstas se les añadió también el número de resultados mostrados por el buscador (el cual representa la competencia entre ellas); como ejemplo, en la Figura 13 aparecen al final (después de hacer la búsqueda) con la leyenda “Cerca de 5,630,000 resultados 0.41 segundos”. Teniendo las palabras con estos dos datos, se graficaron en un diagrama de dispersión tomando como ejes el promedio de búsquedas mensuales y el número de resultados mostrados; dicho gráfico se dividió en cuadrantes tomando la mediana de cada eje, esto con la finalidad de encontrar aquel cuadrante en donde se localizarán aquellas palabras con un mayor promedio de búsquedas mensuales y un menor número de resultados mostrados.

6. Selección de palabras por parte de la empresa

Las palabras identificadas en dicho cuadrante fueron presentadas ante una empresa de mezcal con el objetivo de destacar aquellas *keywords* de interés para

la empresa, tomado en cuenta la necesidad de generar contenido en base a esas palabras y utilizar específicamente las keywords elegidas.

7. Palabras seleccionadas por escala de tiempo

En este paso se examinaron las palabras elegidas a lo largo del tiempo para identificar la existencia de estacionalidad mediante las fluctuaciones en el promedio de búsquedas mensuales a lo largo del año.

8. Generación de conocimiento

Finalmente, con base en la información anterior se obtuvo lo planteado en la metodología ya que se encontró la edad, el género y la ubicación de las personas que indagan en internet en torno al producto elegido, además de la identificación del tipo de temas que más buscan en la red y se identificó si estas búsquedas fluctúan a lo largo del año.

3.3. Extracción depuración y análisis de datos para Facebook

En la Figura 15 se observa que para llegar a la generación de conocimiento a partir del análisis de datos en Facebook se requirió recorrer siete pasos, los cuales se describen a continuación.



Figura 15. Proceso de análisis para publicaciones de Facebook

Fuente: Elaboración propia

1. Obtención y captura de publicaciones

Esta se realizó mediante el uso de la API de Facebook la cual nos permitió obtener una gran variedad de datos no estructurados con respecto a publicaciones de distintas *Fanpage* relacionadas con el producto. Entre estos datos se encontraron el número de publicaciones por página, el número de “me gusta” por publicación al igual que las veces que se comentaron y las veces que fueron compartidas, de igual manera se obtuvo la red de usuarios que dieron me gusta a cada una de la publicaciones. Cabe mencionar que este tipo de datos tuvieron que ser codificados y ordenados para posteriormente ser llevados a una hoja de cálculo en donde se trabajaron.

2. Primer punto de corte (usuarios)

Para evitar la selección de usuarios que ya están ligados a una empresa, lo que varios autores (García-Ruiz, 2005; Marín Rives & Ruiz de Maya, 2007; Palazón et al., 2014) denominan como comunidad de marca, se obtuvo la lista de los usuarios que interactuaron con publicaciones de más de una Fanpage haciendo un corte, definido personalmente, de aquellos usuarios que hubieran interactuado con más de un cuarto de todas las páginas (esto equivalió a que al menos hubiera interactuado con siete Fanpage); considerando esto, se tomó a estos usuarios como los clientes potenciales.

3. Segundo punto de corte (publicaciones)

Una vez teniendo identificados a los clientes potenciales se decidió elegir a aquellas publicaciones que más gustaran a esta audiencia, para ello con el método de percentiles, mismo que fue ejecutado con el apoyo del paquete estadístico de SAS University Edition© se determinó un corte a aquellas publicaciones que tuvieran interacción, es decir un “me gusta”, de cuatro o más clientes potenciales.

4. Categorización de publicaciones

Una vez elegidas las publicaciones que más gustaban, éstas se categorizaron de acuerdo al tipo de contenido que presentaban, en imágenes y texto, además de si hacían alusión a algún evento o presentaban contenido de valor (notas informativas).

5. Análisis de publicaciones por tipo, hora y contenido

Se analizó el formato de las publicaciones, la hora en donde había una mayor interacción por parte de los usuarios y el tipo de contenido que generaba mayor intercomunicación.

6. Comparación entre empresas

Para este apartado se echó mano nuevamente del paquete estadístico de SAS University Edition© para realizar un análisis de correspondencia, el cual Lévy

Mangin & Varela Mallou (2003) describen como un método que representa gráficamente tablas de datos, constituyéndose en la generalización de un diagrama de dispersión y cuya principal ventaja radica en el principio geométrico de representación simultánea de individuos y variables en un mismo espacio, incluyendo su capacidad para tratar datos cualitativos. Goyzueta Rivera (2011) por su parte destaca que proporciona las herramientas para analizar la asociación entre las filas y columnas de una tabla de contingencias, con la idea de desarrollar índices simples que muestren la relación de las variables.

Dicho análisis nos ayudó a correlacionar de manera gráfica el tipo de publicación con la cual tenían más afinidad las empresas.

7. Generación de conocimiento

Finalmente, con base en la información recabada, se obtuvo una lista de clientes potenciales con los cuales se puede conectar; además de, los horarios en los que más interactúan, y el tipo de publicación más empática para los clientes potenciales y si este arquetipo ya está siendo usado por alguna empresa.

3.4. Extracción depuración y análisis de datos para Twitter

Para el análisis de las publicaciones en Twitter se siguieron cuatro pasos (Figura 16). El primero de ellos fue la extracción de tweets (por medio de la API) que incluyeran la mención del producto a analizar, una vez extraídos se procedió a la codificación para ser llevados a una hoja de cálculo. Posteriormente se geolocalizaron aquellos que tuvieran los datos de coordenadas y se desagregaron los tweets por hashtag (HT), retweets (RT) y menciones (MN) esto principalmente en atención a lo recomendado por Del Fresno García et al., (2016) quienes mencionan que la desagregación evita tener que hacer frente a una gran red que ocultaría relaciones y significados, además de que supone una inversión mucho menor en infraestructura tecnológica y capacidad computacional.

Una vez desagregados los mensajes se procedió, con apoyo del software libre Gephi, al análisis de redes realizado en dos grupos: el primero fue enfocado a determinar la temática de interés en torno al producto, aquí se analizaron los hashtag y lo denominamos como retweets 1 y; el segundo análisis, fue orientado a la identificación de las comunidades y líderes de opinión de cada comunidad (análisis de menciones y de retweets 2).

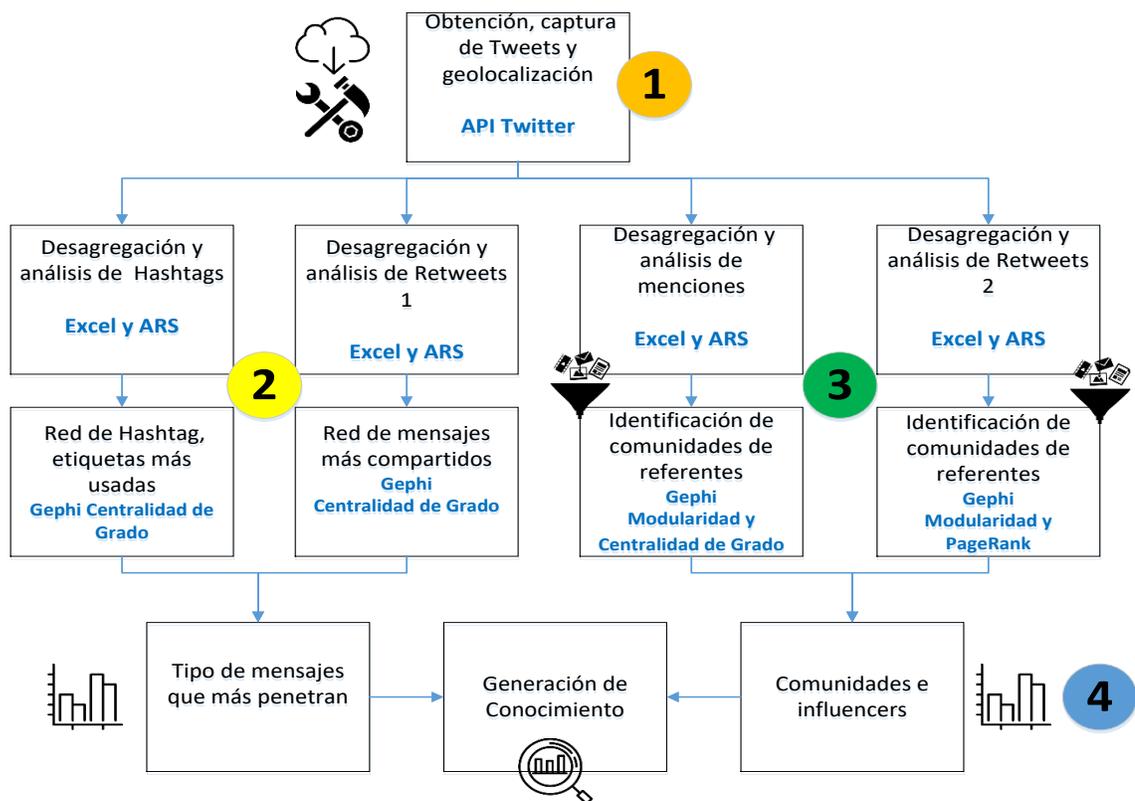


Figura 16. Metodología para el análisis de tweets

Fuente: Elaboración propia

En la Figura 17 se representa, en el inciso “A”, el modo en el que se constituyó el grafo de análisis de hashtag el cual va de un usuario a una etiqueta a la cual

le antecede el símbolo de número “#”; lo que significa es, que dicho usuario, al emitir un tweet usó la etiqueta a la cual señala. Es importante destacar que los nodos a los cuales les antecede el signo de número y que representan a los hashtag usados, sólo tendrán grados de entrada, por el contrario de los nodos que representan a los usuarios que solo tendrán grados de salida; por tal motivo, en el análisis de hashtag se buscó aquellos nodos que tuvieran el mayor número de grados de entrada.

El análisis de Retweets 1, representado en la Figura 17 con el inciso “B”, fue graficado de tal manera que se mostrara el vínculo de usuario hacia tweet. Esta representación tiene la finalidad de dar a entender que el usuario había compartido ese mensaje, la nomenclatura del tweet se conformó de tres partes, en primer lugar un código único por tweet seguido del número de veces que fue compartido ese tweet, seguido el nombre de la cuenta que lo emitió; por ejemplo, en la Figura 17B, el nodo llamado 3962-601.Lopezdoriga se interpreta como un tweet (con código único 3962) que fue compartido 601 veces y emitido por la cuenta @Lopezdoriga. Es importante señalar que, en este grafico al igual que para el caso del análisis de hashtag, los usuarios únicamente tuvieron grados de salida, mientras que los tweets sólo contaron con grados de entrada. En el análisis de retweets 1 se identificó y describió aquellos mensajes que fueron más compartidos.

Para el segundo bloque de análisis (menciones y retweets 2) se siguió con la estructura tipo “C” (Figura 17) en donde la interacción fue de cuenta con cuenta, para el análisis de las menciones se identificó primero, mediante el algoritmo de modularidad (Blondel, Jean-Loup, Renaud, & Etienne, 2008) de Gephi, a las comunidades que integraban la red, se seleccionaron las comunidades más grandes (abarcando más del 10% de la red) y dentro de las comunidades se eligieron a aquellos nodos con mayor centralidad de grado, los cuales fueron determinados como “*influencers*”. Un procedimiento parecido se realizó con el análisis de retweets 2, sólo que en lugar de tomar la centralidad de grado como

indicador principal de influencia, se tomó el algoritmo de Gephi para PageRank (Page, Brin, Motwani, & Winograd, 1998).

Finalmente, con base a la información se obtuvo la ubicación de los usuarios que hablan en torno al producto. Además, se hallaron comunidades e “*influencers*” dentro de estas comunidades, también el tipo de etiquetas, y contenido con las cuales relacionan al producto.

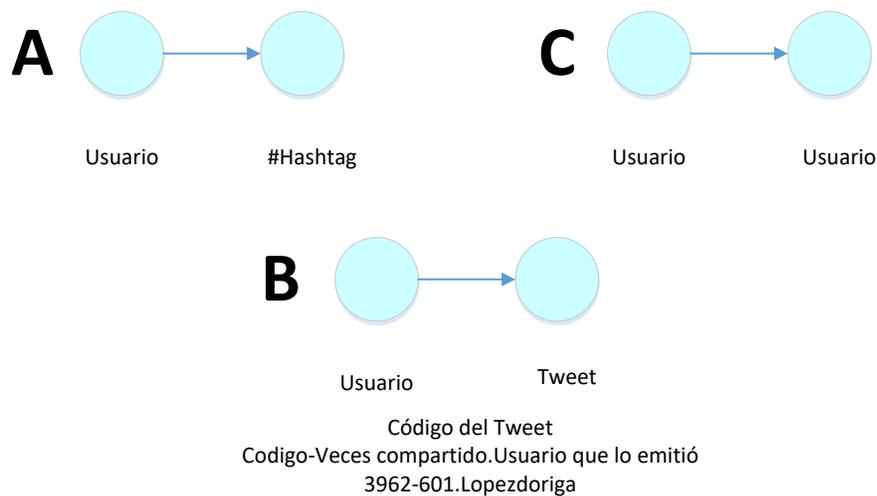


Figura 17. Estructura del grafo de redes por análisis

Fuente: elaboración propia

3.5. Importancia del mezcal como producto para el estudio de caso

Luisa & Hernández (2003) indican que las pequeñas empresas rurales juegan un papel importante en la economía de muchos países como México. Entre ellas se encuentran las de carácter artesanal (como el mezcal) y que por lo regular se producen dentro de los sectores marginados, pero por ello no dejan de estar inmersas en el contexto global de apertura económica y comercial mundial.

También señalan que se debe estar atento a la publicidad para que los compradores puedan percibir su producto como comparable y aceptable en cuanto a precio y calidad.

Se denomina mezcal a la bebida alcohólica regional obtenida por destilación y rectificación de mostos preparados directa y originalmente con los azúcares extraídos de las cabezas maduras de los agaves, a las que les llama localmente piñas, previamente hidrolizadas o cocidas, y sometidas a fermentación alcohólica con levaduras, cultivadas o no. El mezcal es un líquido de olor y sabor *sui géneris* de acuerdo a su tipo. Es incoloro o ligeramente amarillento cuando es reposado o añejado en recipientes de madera de roble blanco o encino, o cuando se aboque sin reposarlo o añejarlo. Es originario del estado de Oaxaca, México; pero también se produce en los estados de Durango, Guerrero, Guanajuato, San Luis Potosí, Tamaulipas, y Zacatecas, mismos que conforman la zona protegida por la denominación de origen “mezcal” (NOM 070,1994) (Hidalgo Reyes, Cruz Meza, & Perez Hernandez, 2010).

En el presente trabajo se opta por elegir la cadena mezcal como objeto de estudio ya que es un producto agropecuario representativo de la cultura mexicana y oaxaqueña, que cuenta además con un incipiente pero nutrido nivel de integración de las empresas en las plataformas digitales.

Hidalgo Reyes et al. (2010) mencionan que el número de marcas de mezcal ha aumentado notablemente, iniciando con 13 en 1994, 83 en 1997, en 2001 pasaron a 101, y para el 2010 se registraron más de 150 marcas.

En palabras de Colmenero-Robles, Rosas-Medina, & Vázquez-Sánchez (2011) lo sobresaliente en la producción del mezcal es su complejidad y diversidad cultural asociada a su producción artesanal que se da por tradición familiar, los mismos autores resaltan que la fabricación de mezcal a partir de varias especies de agave es una actividad cultural y económica que se desarrolla en más de diez estados de la República Mexicana. Secundando la idea de la importancia de la actividad mezcalera, Hidalgo Reyes et al.(2010) mencionan que la producción de

mezcal tiene un gran auge en el mercado nacional e internacional por ser un producto de excelente calidad, que se compara con las mejores bebidas del mundo dadas sus características organolépticas y por ser 100% orgánico.

Incluso en el trabajo de Bañuelos Flores & Salido Ataiza (2012) lo resalta como una bebida espirituosa que enorgullece a los sonorenses y menciona que forma parte esencial de las ceremonias y fiestas de algunos grupos y hasta aparece en canciones de corridos del estado.

Otras características son las extraídas del trabajo de Espinoza Ortega & López Rosas (2016) del cual se hace mención de tres ideas principales que nos apoyan a respaldar la importancia del mezcal, la primera señala que el mezcal es una de las bebidas más representativas de México, ya que cumple una importante función cultural en muchas comunidades productoras, al ser parte de su gastronomía, de las fiestas familiares, patronales y funerarias.

La segunda tiene que ver con el anclaje territorial y sus características endémicas ya que como mencionan los autores, su sabor, color, aroma, esencia y, todo lo que los caracteriza, se debe al espacio geográfico en el que fue creado, sin dejar de lado el saber hacer de las personas que lo producen. Señalan también que a pesar que el anclaje territorial es quien fundamenta su existencia, penosamente no ha tenido el papel protagónico que debería tener en los estudios sobre este tipo de productos.

Sin embargo, y pasando a la tercera idea, resaltan que a pesar de que en años anteriores no era bien aceptada la bebida por el consumidor nacional al identificarla como “una bebida de campesinos y fácil de adulterar”; hoy en día la dinámica es diferente; desde finales del siglo pasado se observa un resurgimiento productivo y una revalorización del consumidor, al grado que es la bebida nacional de mayor crecimiento en el mercado.

4. RESULTADOS Y DISCUSIÓN

En el presente apartado se muestran los resultados obtenidos una vez aplicada la metodología propuesta a las tres diferentes plataformas web. Se destacan también algunas de las particularidades encontradas durante el proceso de análisis. La estructura de este apartado se alinea con el orden en el que se estructuró la metodología, es decir primero se aborda lo relacionado al buscador de Google y el análisis de las palabras clave, en segundo lugar se encuentra la red social Facebook, en ésta se presenta lo relacionado al tratamiento que se llevó a cabo con las publicaciones emitidas por páginas de fans, y en tercer lugar se presenta Twitter en donde la materia prima fueron los tweets que estuvieran relacionados con el producto.

Es importante destacar que para cada apartado se fue mostrando lo determinado anteriormente como información, delimitando la información obtenida con respecto a usuarios de la información con respecto al contenido que éstos consumen, lo anterior en aras de dar respuesta a los dos primeros objetivos planteados de la investigación. Al final de cada apartado se presenta también la utilidad de la información obtenida, la cual es denominada en el presente trabajo como el conocimiento generado.

4.1. Google

Con el apoyo de la herramienta en línea Hyppersuggest se obtuvieron un total de 2,232 combinaciones de palabras con la característica de poseer la palabra “mezcal” antes, en medio y después de la frase; también se consideró su inclusión en forma de pregunta, básicamente el punto fue encontrar todo lo relacionado con “mezcal”.

Posteriormente, con el apoyo de la herramienta de Google Adwords, se procedió a la obtención del número de búsquedas mensuales de dichas combinaciones de palabras en un periodo de un año (enero del 2015-diciembre 2015), esto nos

permitió la identificación y eliminación de aquellas palabras clave que no presentaban registro de búsquedas (1,946) quedando 286 combinaciones, las cuales fueron clasificadas en 14 categorías:

1. Agave: Son combinaciones que incluyen la palabra agave.
2. Opiniones: Son combinaciones de palabras que muestran intención de compra o de búsqueda de opiniones acerca de los mejores mezcales.
3. Canciones: Son frases que hacen referencia a canciones que incluyen la palabra mezcal.
4. Frase: Son aquellas combinaciones de palabras que hacen alusión a la búsqueda de frases de mezcal, no necesariamente asociadas al tema anterior.
5. Gusano: Son aquellas búsquedas que incluyen la palabra gusano o en su caso *worm*.
6. Información: Son aquellas combinaciones que se enfocan más a la búsqueda de información (significado, historia, origen) acerca del mezcal.
7. Lugar: Estas palabras hacen referencia a la búsqueda de lugares como cantinas o restaurantes.
8. Marcas: En esta categoría se encuentran todas las combinaciones de palabras que hacen alusión a una marca en específico.
9. Mezcalería: Son las frases que incluyen la palabra mezcalería.
10. Mezcal y Oaxaca: Son combinaciones que aparte de la palabra mezcal incluyen también al estado de Oaxaca.
11. Mezcal y Tequila: Son combinaciones que además de la palabra mezcal también contienen la palabra tequila.
12. Mezcal de Pechuga: Son combinaciones que aparte de la palabra mezcal incluyen la palabra pechuga.
13. Preparación: Son combinaciones de palabras relacionadas a cómo se toma o cómo se puede preparar el mezcal.
14. Sólo Mezcal: En esta categoría aparecen palabras solitarias como mescal, mezcales, mezcals o la misma palabra mezcal únicamente.

Los estadísticos descriptivos y ejemplificación de cada categoría se pueden ver en el Cuadro 2.

Cuadro 2. Estadísticos descriptivos y ejemplificación de las categorías de palabras clave.

Categoría	Número de apariciones	Promedio de búsquedas mensuales	Mínimo	Máximo	Desviación estándar	Ejemplos
Agave	5	129.2	10	362	147.7	agave mezcal, mezcal agave, destilado de agave, agave espadín, agave liquor
Opiniones	21	91.0	7	262	76.7	el mejor mezcal, marcas de mezcal, los mejores mezcales, mezcal comprar, cual es el mejor mezcal, mezcal precio
Canciones	18	347.8	9	2,891	669.1	dos botellas de mezcal, jenni rivera dos botellas de mezcal, mezcal antonio aguilar, copitas de mezcal letra, 2 botellas de mezcal letra
Frase	3	257.8	159	370	106.5	frases de mezcal, para todo mal mezcal, para todo mal mezcal para todo bien también
Gusano	20	38.5	1	260	59.8	mezcal de gusano, mezcal with worm, gusano de mezcal, mezcal gusano, mezcal worm

Categoría	Número de apariciones	Promedio de búsquedas mensuales	Mínimo	Máximo	Desviación estándar	Ejemplos
Información	37	142.0	8	556	131.1	what is mezcal, historia del mezcal, de donde es el mezcal, de que esta hecho el mezcal, origen del mezcal, que significa mezcal
Lugar	13	33.5	2	235	70.3	el mezcal restaurant, mezcal cantina, mezcal bar, mezcal nyc, el mezcal mexican restaurant, cantina mescal
Marcas	107	187.4	8	1,200	229.2	alipus mezcal, mezcal los danzantes,mezcal convite, mezcal alacrán, mezcal beneva, mezcal oro de Oaxaca, mezcal 8 viboras, mezcal wahaka
Mezcalería	14	231.0	108	645	171.6	mezcalerías df, mezcalería coyoacan, el rey mezcalería, mezcalería merida, la mezcalería condesa
Mezcal y Oaxaca	7	211.8	75	447	145.2	mezcal de Oaxaca, oaxaca mezcal, mezcales de Oaxaca, mezcal oaxaqueño, mezcaloteca Oaxaca
Mezcal y Tequila	16	116.8	3	532	169.5	mezcal o tequila, diferencia entre tequila y mezcal, mezcal vs tequila, diferencia entre mezcal y tequila

Categoría	Número de apariciones	Promedio de búsquedas mensuales	Mínimo	Máximo	Desviación estándar	Ejemplos
Mezcal de Pechuga	3	286.4	26	720	378.1	mezcal de pechuga, mezcal pechuga, pechuga mezcal
Preparación	13	205.7	13	597	195.6	como se toma el mezcal, mezcal como se toma, mezcal cocktails, con que se toma el mezcal, bebidas preparadas con mezcal
Sólo Mezcal	9	2,086.1	10	15,950	5211.9	Mezcal, mescal, mezcales, mezcals, el mezcal, el mescal
Total	286	325.8	1	15,950	554.5	

Fuente: elaboración propia a partir de datos de Google Adwords

Del cuadro anterior es importante destacar que existen combinaciones de palabras que son muy parecidas cambiando sólo en algunas letras o inclusive en el orden en que están escritas, pero que en esencia conducen al mismo significado, tales son los casos de “como se toma el mezcal” y “mezcal como se toma” o también “mezcal de Oaxaca” con “mezcales de Oaxaca”, no obstante son combinaciones de palabras que se deben de analizar individualmente ya que cada una tiene su propio número de búsquedas mensuales y el número de resultados que muestra el buscador de Google también es diferente. Retomando el último ejemplo, el simple hecho de usar la terminación “es” a la palabra mezcal reduce el promedio de búsquedas mensuales pasando de 447.5 (mezcal de Oaxaca) a 225.8 (mezcales de Oaxaca).

Trasladando el Cuadro 2 a una gráfica de pastel por apariciones, podemos ver (Figura 18) que el mayor porcentaje lo representa la categoría de marcas lo que nos da una idea que existe una mayor variedad de palabras en torno a marcas específicas de mezcal, no obstante el intentar usar estas palabras para generar contenido en una página web sería tal y como lo señala Amunátegui Perelló (2006) colgarse del buen nombre de la empresa o las empresas dominantes y que de esta manera al hacer un usuario una búsqueda podría encontrar diferentes páginas en los resultados de la misma y por lo tanto el consumidor no sabría exactamente cuál de ellos corresponde efectivamente a la empresa que busca.

No obstante a pesar de la variedad de la que se dispone en la aparición de palabras por ejemplo de la categoría de “información” (ocupa el segundo puesto con el 13%) existen otras variables mucho más importantes, entre ellas está el promedio de búsquedas mensuales, que de igual manera al trasladarlo del Cuadro 2 a una gráfica de barras (Figura 19) podemos observar que la categoría “Sólo mezcal” es la que representa el promedio más alto, lo anterior se da porque la propia palabra mezcal tiene por si misma 15,950 búsquedas mensuales, sin embargo para tener una mejor selección de palabras clave útiles para posicionar nuestro sitio web es importante tomar otra variable como lo es la competencia, variable que se retoma más adelante.

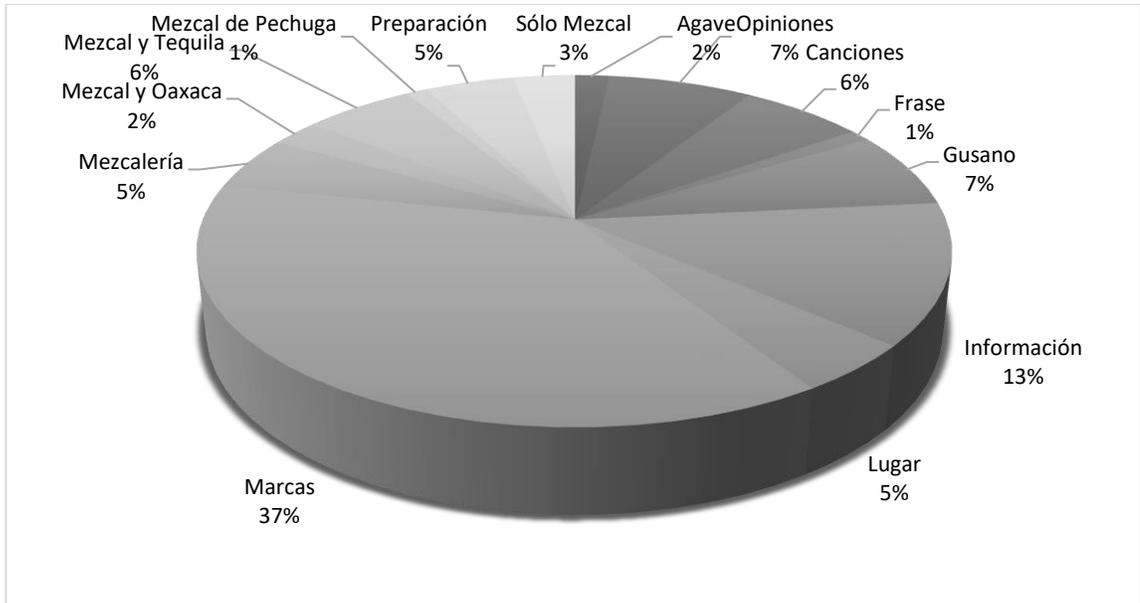


Figura 18. Porcentaje de las categorías por el número de apariciones

Fuente: Elaboración propia

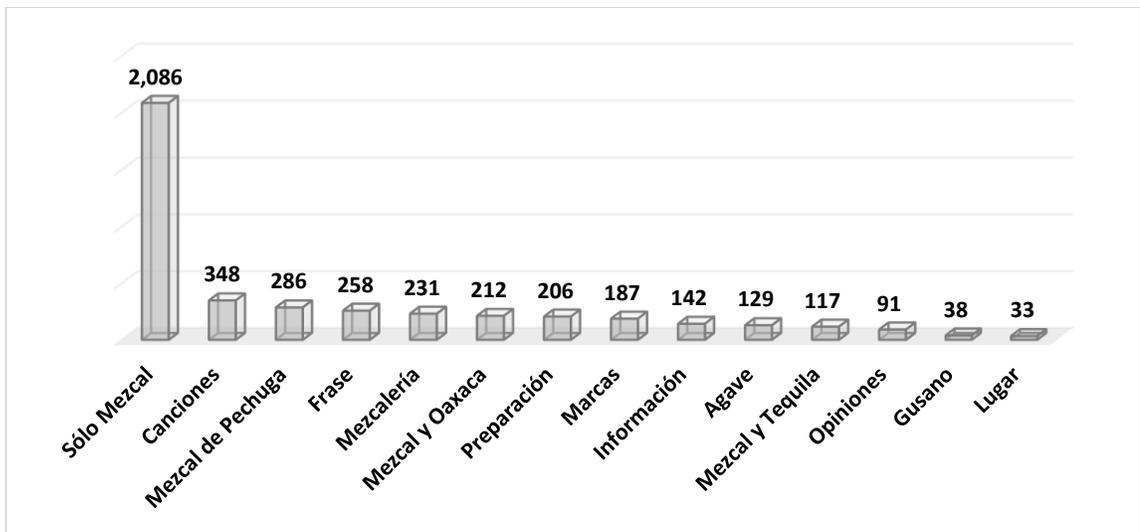


Figura 19. Promedio de búsquedas mensuales por categoría

Fuente: Elaboración propia a partir de datos de Google Adwords

Las figuras anteriores nos dan en un primer momento noción del tipo de tópicos de los que se habla en torno al mezcal y en cuáles de esos tópicos se puede incursionar como contenido en una página web y que ayude a atraer más visitantes, tales son los casos de las categorías de “Información” y “Opiniones”. La primera concierne a un sector de la audiencia que busca información acerca del mezcal, sus orígenes, su historia, su significado y procedencia, temas que se pueden abordar en un sitio web como enganchadores de visitantes.

También se detectaron aquellas categorías que no serían recomendables para generar contenido sobre ellas, éstas son “Canciones” y “Marcas”, así que estas categorías se retiraron de los análisis posteriores; la primera debido a que la búsqueda de alguna canción, a pesar de que incluya la palabra mezcal, es poco relevante para promocionar un marca (ya que el usuario sólo le interesa la melodía o el artista que la canta), y la segunda es porque si el usuario ya busca una marca en específica es resultado del *branding* de la empresa o marca y por tanto es más difícil lograr que cambie a otras opciones de compra además de ser poco honesto y antiético al colgarse del buen nombre de la empresa dominante (Amunátegui Perelló, 2006).

Una vez eliminadas las categorías de “Canciones” y “Marcas” se analizaron las 161 combinaciones de palabras restantes. En un primer momento para obtener una aproximación al tipo de audiencia en internet que podía ser consumidor potencial; en un segundo momento enfocado a identificar las combinaciones de palabras más pertinentes para lograr posicionarse.

Atendiendo al primer objetivo señalado (información acerca del perfil del usuario) se optó por someter las 161 combinaciones de palabras nuevamente a la herramienta de Google Adwords para visualizar cómo se distribuían los usuarios que indagaban sobre los tópicos de mezcal por entidad federativa (Figura 20).

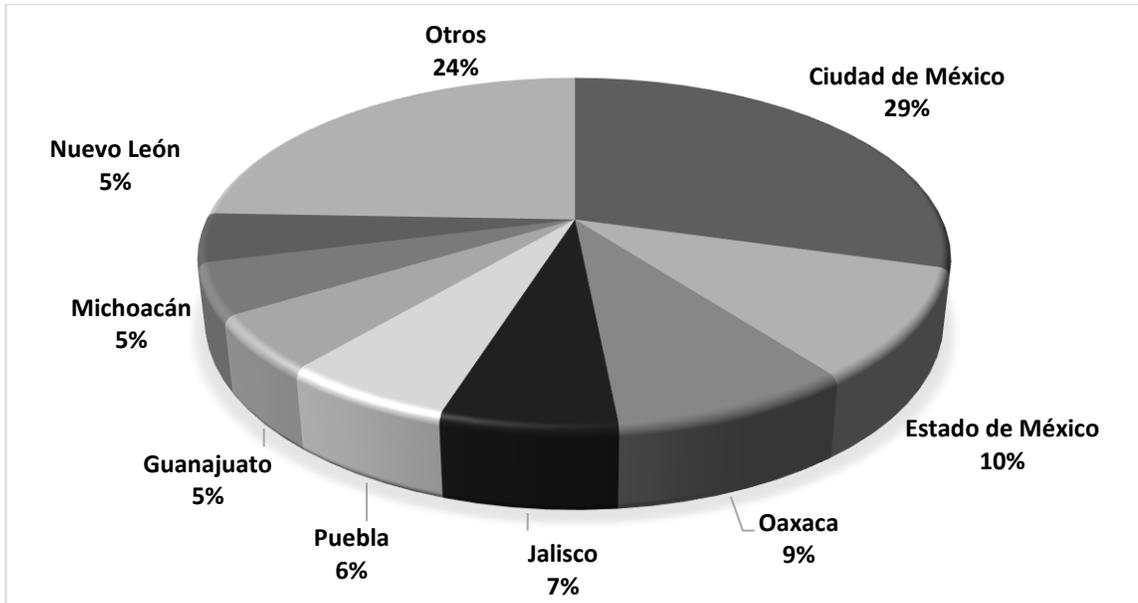


Figura 20. Usuarios que buscan los tópicos del mezcal seleccionados, distribuidos por estado

Fuente: Elaboración propia a partir de Google Adwords

De la figura anterior se resalta que casi un 40% de la población analizada, se encuentra en la parte centro del país. Sin embargo, es importante señalar que podría haber un sesgo en cuanto al número de internautas por entidad federativa ya que, retomando los datos de la Figura 3, el porcentaje de usuarios de internet de la Ciudad de México (72.4) se encuentra por encima de estados como Oaxaca (41.7), Puebla (47.6) o Michoacán (43.8) , no obstante los datos reflejados en la Figura 20 no representan con exactitud la distribución de internautas en México por estado, ejemplo de ello es la ausencia de estados como Baja California y Baja California Sur , los cuales ocupan las primeras posiciones de internautas con 74.4 y 70 por ciento respectivamente, además de la posición de Nuevo León que contrasta con la cuarta posición que ocupa a nivel nacional en cuanto a usuarios de internet con un 70.4%.

Las mismas 161 combinaciones de palabras se sometieron a la herramienta de Bing Ad Intelligence que es el estadista de Microsoft para palabras clave, esta

herramienta nos permitió analizar un total de 45 de las 161 combinaciones de palabras con respecto a las edades de los usuarios que realizaban búsqueda (Figura 21). De igual manera puede haber un sesgo por las edades de los usuarios de internet que se puede ver reflejada en que sólo un 1.6% de usuarios mayores a 65 años buscan contenido acerca del mezcal lo cual se alinea con los mostrado anteriormente en la Figura 1 en donde los usuarios de internet de mayor edad (55 y más años) son los que ocupan la última posición con un 5%, no obstante al igual que para el caso de las entidades federativas la Figura 21 no refleja exactamente la distribución de usuarios por edad que hay en México ya que de ser así los porcentajes de usuarios con edades de 18 a 34 años deberían reflejar lo que ocurre a nivel nacional (ver Figura 1) y ser mayores que él de la categoría de 35 a 49 años. Se resalta que la población que se realizó en el presente trabajo es la que se encuentra constantemente interactuando en la red.

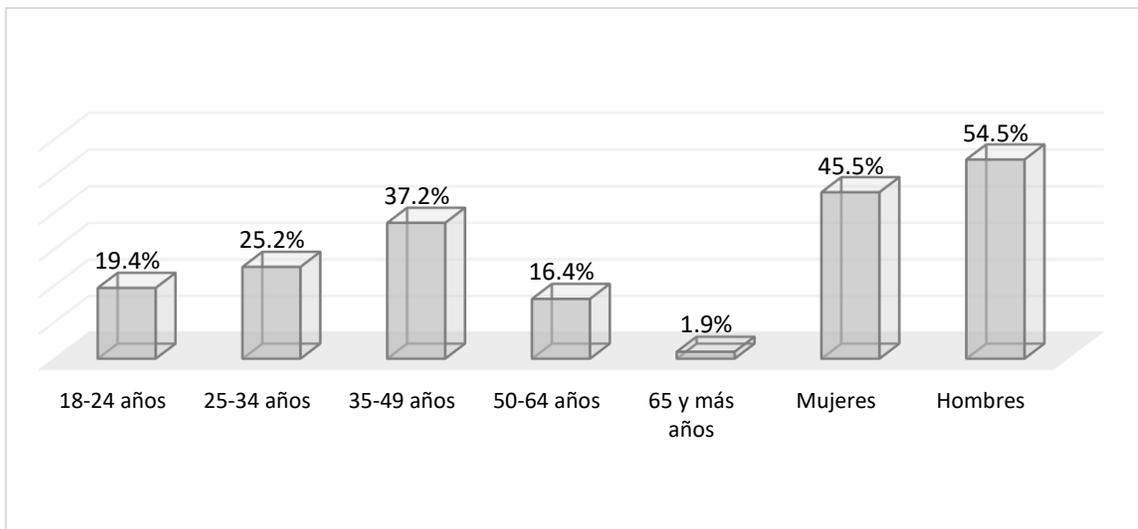


Figura 21. Usuarios que buscan los tópicos del mezcal seleccionados, distribuidos por estado

Fuente: Elaboración propia a partir de Bing Ad Intelligence

Por lo anterior y respondiendo a la primera incógnita de aproximación al tipo de audiencia, se puede afirmar que el gran potencial de una marca de mezcal en internet está en la parte centro del país seguido del estado de Oaxaca y Jalisco (estados productores de mezcal) con una audiencia en su mayoría hombres con edades de entre 35 y 49 años.

Con respecto a la segunda incógnita sobre la pertinencia de las palabras para posicionarse, es importante reiterar que la elección por una serie de palabras no está dada únicamente por el número de búsquedas que éstas tengan, sino también influye la cantidad de competencia en resultados que existe por frase (Amunátegui Perelló, 2006; Arbildi-Larreina, 2005; Grávalos Macho, 2013; Vález, 2011). De esta forma, se retoma el ejemplo de la palabra “mezcal” la cual tiene un total de 15,950 búsquedas promedio al mes, pero el número de resultados al hacer una búsqueda es de 59 millones, a comparación de la frase “mezcalerías de” que, si bien tiene un promedio de 404 búsquedas mensuales, los resultados que ofrece el buscador es de 6,640 lo que significa menos competencia entre palabras.

Siguiendo la lógica anterior se optó por graficar en un diagrama de dispersión las 161 combinaciones de palabras, tomando como ejes el número de resultados que se muestran al hacer la búsqueda y el promedio de búsquedas mensuales, se dividió en cuadrantes con respecto a la mediana de cada eje con el objetivo de encontrar el cuadrante en donde se encontrarán aquellas combinaciones de palabras que tuvieran el más alto número de búsquedas y el más bajo número de resultados (Figura 22).

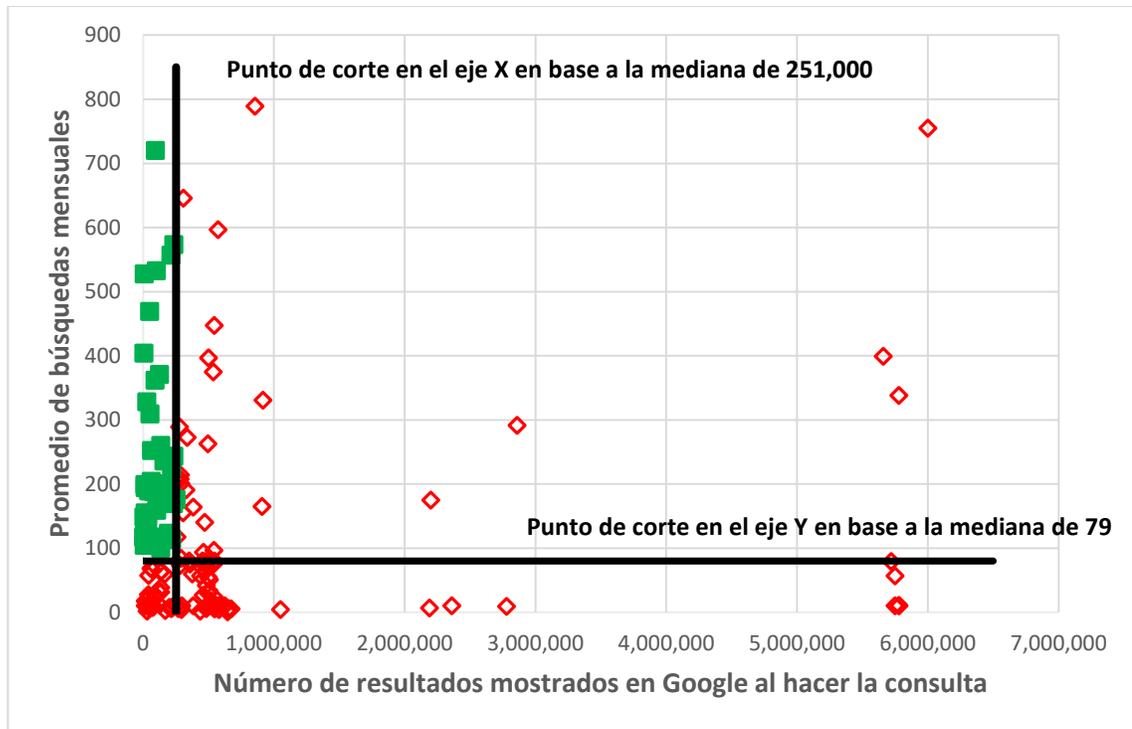


Figura 22. Clasificación de palabras clave por cuadrante (resultados/búsquedas)

Fuente: elaboración propia con datos de Google Adwords

En la Figura 22 se observa que existen puntos muy alejados del entrecruce de los dos cortes, pero que la mayor densidad se encuentra cercano a éste. También se puede observar que de las palabras que se ubican dentro del cuadrante superior derecho (mayor promedio de búsquedas mensuales, menor número de resultados mostrados) la mayor densidad se encuentra en menos de 200 búsquedas mensuales.

Al realizar un zoom al cuadrante superior derecho (Figura 23) en donde se encuentran las palabras con mayor promedio de búsquedas mensuales y menor competencia, se hallaron un total de 48 “combinaciones recomendadas”, las cuales redujeron su número a 26 (palabras resaltadas en la Figura 23) al ser sometidas a la selección por parte de una empresa mezcalera con base en

mezcalerías (círculos verdes) ya que al igual que las marcas, existieron varias palabras de esta categoría que hablaban de lugares específicos o relacionados con alguna marca, el ejemplo se presenta con “mezcalería los amantes” que si bien es una de las palabras con menor número de resultados y mayor número de búsquedas mensuales, el mismo nombre hace referencia a una marca en específico , por tal motivo no es conveniente la inclusión de esta frase dentro del contenido.

Una vez identificadas las 26 palabras seleccionadas por la empresa como materia prima de generación de contenido se procedió a visualizar la fluctuación de dichas palabras a lo largo del periodo de recolección que fue de enero a diciembre del 2015 (Figura 24), en donde se puede apreciar que en su mayoría, las combinaciones de palabras elegidas, no superan las 200 búsquedas por mes, y este promedio de búsquedas se mantiene constante a lo largo del año, no obstante existen combinaciones que se muestran una clara estacionalidad en algunos meses.

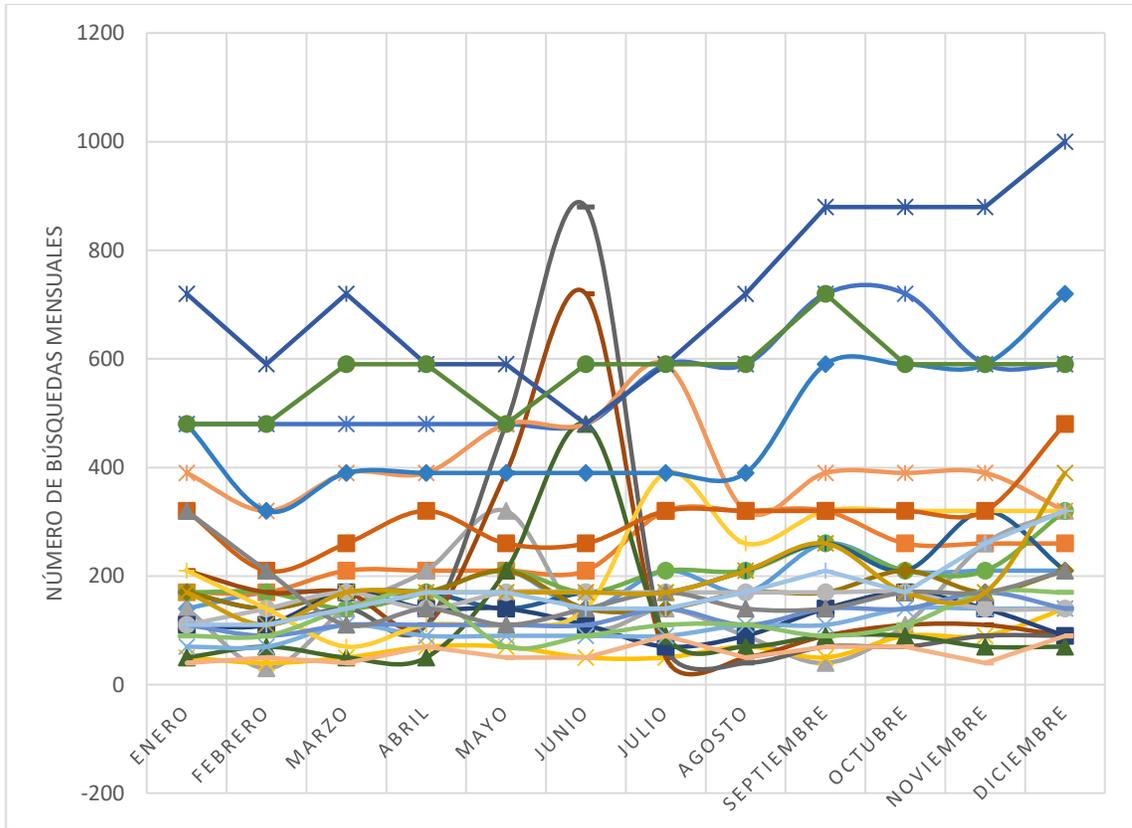


Figura 24. Fluctuación de búsquedas mensuales por combinación de palabras

Fuente: Google Adwords

Con el objetivo de tener una mejor visualización de las figuras anteriores, se aislaron las frases más fluctuantes o tendencias mayores a 200 búsquedas (Figura 25). De dichas palabras destacan “significado de mezcal”, “que significa mezcal” y “mezcal significado” con un pico en el mes de junio, es decir que en ese mes cobra relevancia la audiencia, esto puede ser causado porque en el mes de Julio es cuando se realiza el evento de la Guelaguetza Oaxaqueña. También se puede apreciar que en algunas palabras existe un pico en el mes de septiembre, tal es el caso de las palabras “diferencia entre mezcal y tequila”, “cocteles con mezcal” o el mismo “mezcal de pechuga”. Es importante resaltar que las temporadas donde se aprecian picos en los volúmenes de búsqueda son

los meses de Junio, Noviembre y Diciembre, lo anterior se puede explicar por la presencia en estos meses de eventos como lo son la Guelaguetza, día de muertos y fiestas de fin de año, esto implica la necesidad de las agroindustrias de este producto en poner especial atención y maximizar sus esfuerzos de promoción y vinculación con su audiencia por medio de medios digitales en fechas que se identifican como clave. .

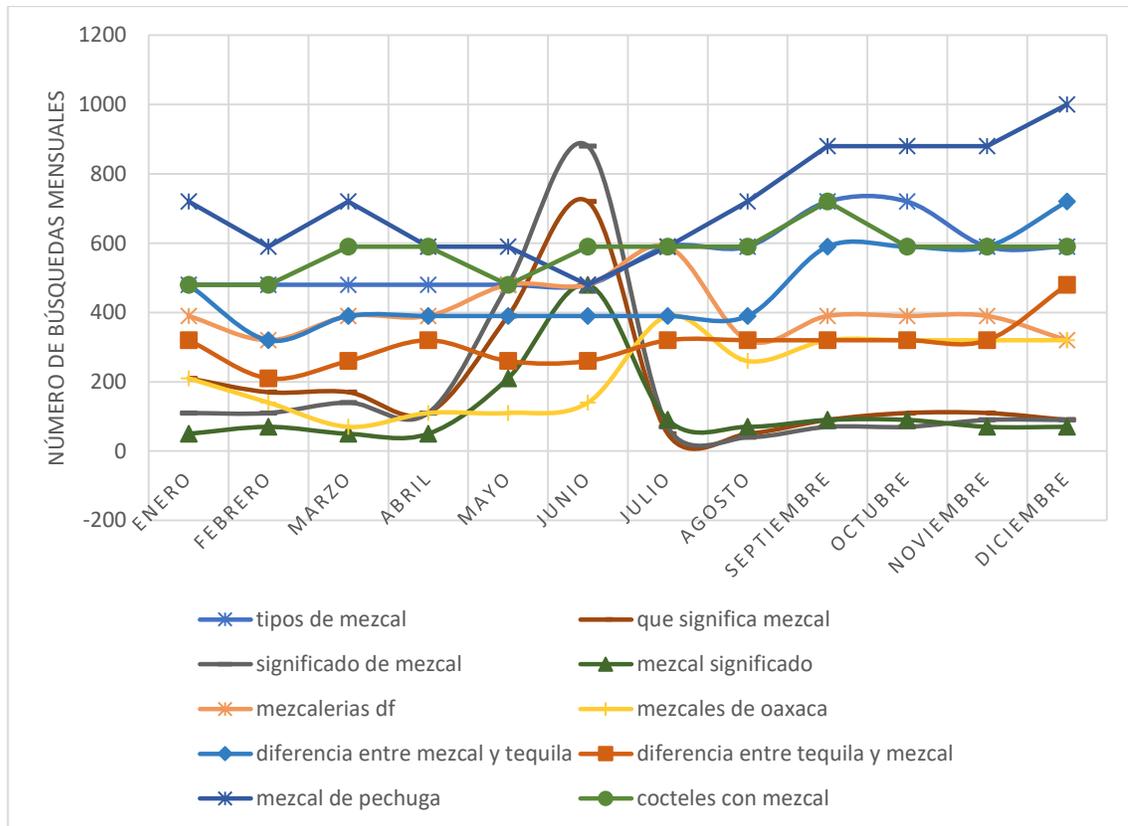


Figura 25. Palabras clave que muestran fluctuaciones por temporada

Fuente: Google Adwords

Hasta este punto se han venido desglosando y mostrando los resultados obtenidos durante el proceso desde la extracción de datos hasta la obtención de información de la metodología planteada (Figura 26), resta para el caso de Google, describir el paso número 8, el cual corresponde al conocimiento que se generó, refiriéndonos a conocimiento como el paso subsecuente a la obtención

de información (ver Figura 11) y en palabras de García-Marco (2011) caracterizado por la utilidad o sentido de ésta, los porqués y para qué de la misma.

Cabe señalar que durante el proceso de análisis se fue disminuyendo la cantidad de palabras clave pasando de 2,232 hasta llegar a un total de 26 palabras consideradas útiles para una empresa de mezcal (Figura 26).

Entrando en materia del conocimiento generado, se puede decir que aquella empresa de mezcal que quiera incurrir en los buscadores debe tener en cuenta lo siguiente:

En primer lugar, el potencial de una empresa con el temático mezcal se encuentra en un público con una edad de entre 35 y 49 años, que se encuentren principalmente en el centro del país (Ciudad de México, Estado de México, Puebla) y la vinculación con este segmento se debe hacer a través de contenido relacionado con las siguientes 26 combinaciones de palabras:

agave espadín, frases de mezcal, para todo mal mezcal para todo bien también, gusano rojo mezcal, tipos de mezcal, mezcal espadín, beneficios del mezcal, que significa mezcal, significado de mezcal, propiedades del mezcal, elaboración del mezcal, mezcal significado, mezcal denominación de origen, mezcalerías df, mezcalerías, mezcales de Oaxaca, mezcal oaxaqueño, ruta del mezcal Oaxaca, diferencia entre mezcal y tequila, diferencia entre tequila y mezcal ,mezcal vs tequila, mejores mezcales, mezcal de pechuga, cocteles con mezcal, como tomar mezcal, tragos con mezcal.

En segundo lugar, se destaca la importancia de incluir ya sea dentro del contenido o incluso en el título la combinación de palabras exacta, sin variar en el orden o en alguna letra ya que, como se ejemplificó anteriormente, el simple hecho de utilizar o no la terminación “es” en los ejemplos de “mezcal de Oaxaca”

y “mezcales de Oaxaca, redujo el promedio de búsquedas mensuales en 221.7 pasando de 447.5 (mezcal de Oaxaca) a 225.8 (mezcales de Oaxaca).

Otro punto importante es que es recomendable abordar temas acerca del significado del mezcal durante el mes de junio, que es en donde hay una mayor cantidad de búsquedas sobre información de este tema, también hacer énfasis en el mes de septiembre con respecto a las diferencias entre mezcal y tequila o preparación de bebidas con mezcal (cocteles) y finalizar el año con temas como mezcal de pechuga.

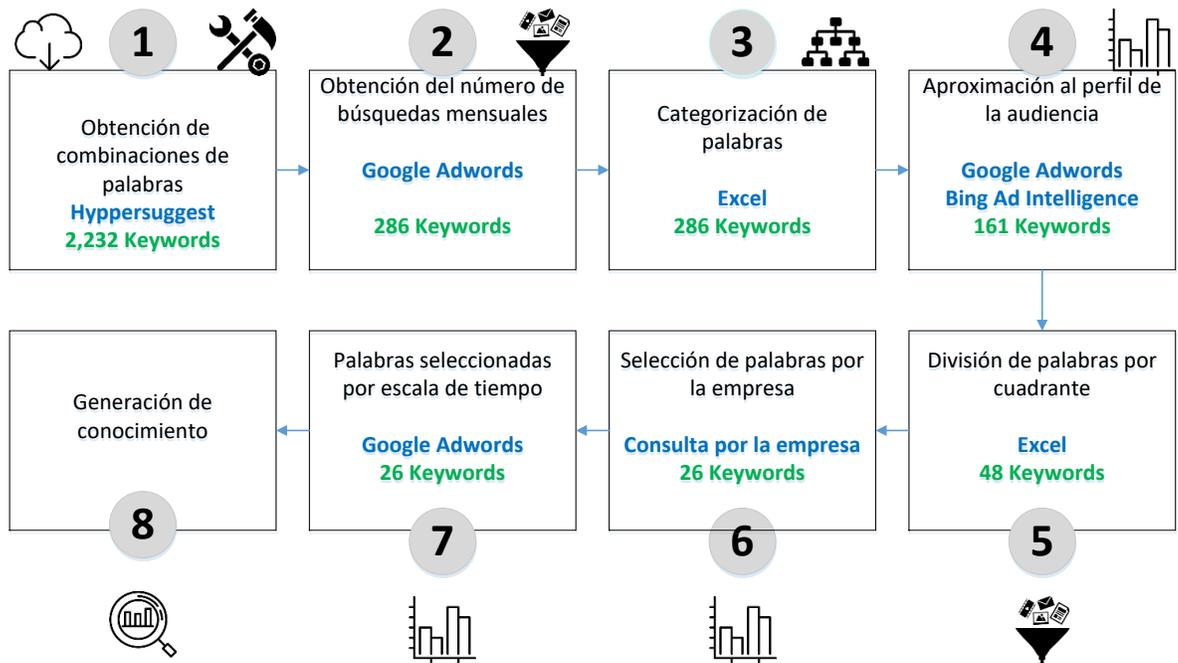


Figura 26. Pasos que se siguieron para la generación de conocimiento en cuanto a palabras clave

Fuente: Elaboración propia

4.2. Facebook

Para el caso de Facebook y siguiendo la metodología propuesta se inició con buscar Fanpage relacionadas con el tema de mezcal, principalmente marcas, se hallaron 28 páginas de las cuales se extrajeron mediante la API 11,851 publicaciones y las interacciones de éstas con 80,345 usuarios. En el Cuadro 3 se presentan los descriptivos con respecto a los números brutos de cada página, los cuales (salvo el promedio) se obtienen de manera directa de la API sin realizar algún tipo operación o análisis adicional.

Cuadro 3. Descriptivos con respecto a 28 Fanpage que abordan temáticas sobre mezcal

Cuenta de Facebook	Likes de página (Número de me gusta de página)	Publicaciones (Número de publicaciones emitidas)	Usuarios (Perfiles que han dado por lo menos un like a una publicación)	Interacción (total de likes en las publicaciones)	Promedio (likes por publicación)
Mezcal Buen Viaje	18,922	617	30,240	67,726	109.8
Mezcal Ilegal	13,907	999	7,242	19,390	19.4
Sabe Mezcal	6,432	999	6,752	35,082	35.1
Mezcal Los Siete Misterios	8,403	999	4,725	12,988	13.0
Mezcal oro de Oaxaca	13,181	999	4,375	17,396	17.4
Mezcal Convite	10,060	335	4,209	10,323	30.8
Mezcal Marca Negra	5,335	397	3,267	9,175	23.1
Wahaka Mezcal	6,556	999	2,903	16,318	16.3

Cuenta de Facebook	Likes de página (Número de me gusta de página)	Publicaciones (Número de publicaciones emitidas)	Usuarios (Perfiles que han dado por lo menos un like a una publicación)	Interacción (total de likes en las publicaciones)	Promedio (likes por publicación)
Mezcal Amores Spain	5,880	233	2,768	5,634	24.2
Mezcal El Jolgorio	2,417	618	2,259	7,027	11.4
Del Maguey Single Village Mezcal	4,317	679	1,954	7,199	10.6
Rey Campero	1,583	367	1,369	9,369	25.5
Mezcal Los Amantes	930	310	1,327	3,057	9.9
Por qué Mezcal	3,585	267	1,157	3,369	12.6
Mezcal La Bendición	1,327	213	967	2,308	10.8
Mezcalillera	2,414	999	822	5,546	5.6
Koch el Mezcal de Oaxaca	2,934	334	776	1,841	5.5
Scorpions Mezcal	3,045	58	743	1,900	32.8
Mezcal Rajabule	930	392	737	3,812	9.7
La Sabina Mezcaleria	1,422	186	595	1,255	6.8
Mezcal los danzantes	1,700	256	479	2,019	7.9
Consejo Regulador Del Mezcal	2,266	43	158	218	5.1

Cuenta de Facebook	Likes de página (Número de me gusta de página)	Publicaciones (Número de publicaciones emitidas)	Usuarios (Perfiles que han dado por lo menos un like a una publicación)	Interacción (total de likes en las publicaciones)	Promedio (likes por publicación)
Mezcal Dainzu	544	100	148	425	4.3
Mezcal los Javis	655	87	146	454	5.2
Expresiones del Mezcal	226	138	110	341	2.5
Proyecto Mezcal	320	158	55	162	1.0
Mezcal Enmascarado	1,297	25	39	60	2.4
Mezcal drinking Challenge around the World	164	44	23	49	1.1
Total	120,752	11,851	80,345	244,443	16.4

Fuente: Elaboración propia a partir de la API de Facebook

A pesar de contar con estos datos es necesario rescatar lo mencionado por Sánchez (2012) quien recomienda no caer en el clásico error de obsesionarnos con el número de seguidores que podemos tener en Facebook (like de página) o Twitter y que mucho más importante que esto, es la actividad que generamos (like en las publicaciones, comentarios, compartidos), contenido interesante que genere debate, que empuje a la gente a participar.

Al empezar a analizar la interacción de los usuarios en Facebook se encontraron comportamientos peculiares, entre éstos se encuentran la relación que guardan algunas variables, al hacer un análisis de correlaciones (Cuadro 4) se pudo observar que los likes de página (que representan el tamaño de la audiencia) se encuentran, de acuerdo la escala de Martínez Ortega (2009), con una buena

correlación con respecto a los usuarios, la interacción y al promedio. Lo anterior significa que si una página tiene un gran volumen de audiencia también tendrá una gran cantidad de usuarios que den like a alguna de sus publicaciones y viceversa, ya que estas dos variables se encuentran relacionadas positivamente, lo anterior igual aplica para la suma total de likes en sus publicaciones (Interacción) y por tanto en el promedio; es decir que una Fanpage con mayor cantidad de audiencia (likes de página) puede escribir contenido no relevante, incluso sin tener que ver con el mezcal, como lo podría ser un mensaje de “buenos días”, y este mensaje podría tener mayor interacción que una publicación de contenido relevante que sea emitido por una Fanpage con menor cantidad de audiencia. Por tal motivo el querer clasificar y seleccionar las publicaciones más relevantes por el número de likes (números brutos) que éstas tienen, daría una aproximación errónea. Por tanto, es indispensable aplicar una metodología de filtrado para la identificación de contenido relevante.

También se destaca que la correlación entre el número de publicaciones y el promedio de likes por publicación es regular (en la escala de Martínez Ortega, 2009), por tal motivo se puede inferir que no necesariamente entre más publicaciones emita una página, éstas también tendrán más likes, lo que nos da una idea de que existe contenido relevante y no relevante para los usuarios en Facebook. En la Figura 27 se hace una representación gráfica por medio de flechas de la peculiar interacción de likes emitidos de los usuarios (representados con las letras C y D) hacia las publicaciones (letras B y E) realizadas por una Fanpage (letra A)

Aquellos internautas representados por la letra C son los que pertenecen a la llamada comunidad de marca, es decir que sólo interactúan (dan like) con las publicaciones de una sola Fanpage, es importante señalar que este tipo de usuarios representan arriba del 95% y que el intentar vincularse con este tipo de usuarios que pertenecen a una comunidad de marca (García-Ruiz, 2005; Marín Rives & Ruiz de Maya, 2007; Palazón et al., 2014) mediante contenido sería más infructuoso que el hacerlo por medio de los usuarios representados con la letra

D (audiencia potencial), los cuales interactúan con las publicaciones de más de una Fanpage. Por lo anterior el primer reto está en encontrar a este tipo de usuarios que, si bien les interesa la temática en torno al mezcal, no están estrechamente vinculados con una marca en específico. El segundo reto está en que una vez teniendo la audiencia potencial, identificar clasificar y analizar el tipo de contenido que llama la atención a esta audiencia potencial (letra E) y determinar si este tipo de contenido está siendo utilizado prioritariamente por alguna de las Fanpage. Para tal efecto y siguiendo la metodología propuesta se realizaron dos puntos de corte el primero correspondiente al tipo de usuarios con el objetivo de encontrar a la audiencia potencial y el segundo con respecto a las publicaciones, obteniendo aquellas que presentaran mayor interacción con la audiencia potencial.

Con respecto a los usuarios se realizó un corte en función de aquellos usuarios que mínimamente interactuaran con un cuarto de las Fanpage analizadas, es decir que hayan dado like a publicaciones de mínimamente siete páginas diferentes. Se encontraron un total de 69 usuarios (Apéndice 2), a los cuales les llamaremos audiencia potencial.

Cuadro 4. Análisis de correlaciones entre variables de Facebook

	Likes de página	Publicaciones	Usuarios	Interacción
Likes de página	1			
Publicaciones	.699**	1		
Usuarios	.898**	.827**	1	
Interacción	.851**	.889**	.964**	1

Fuente: Elaboración propia

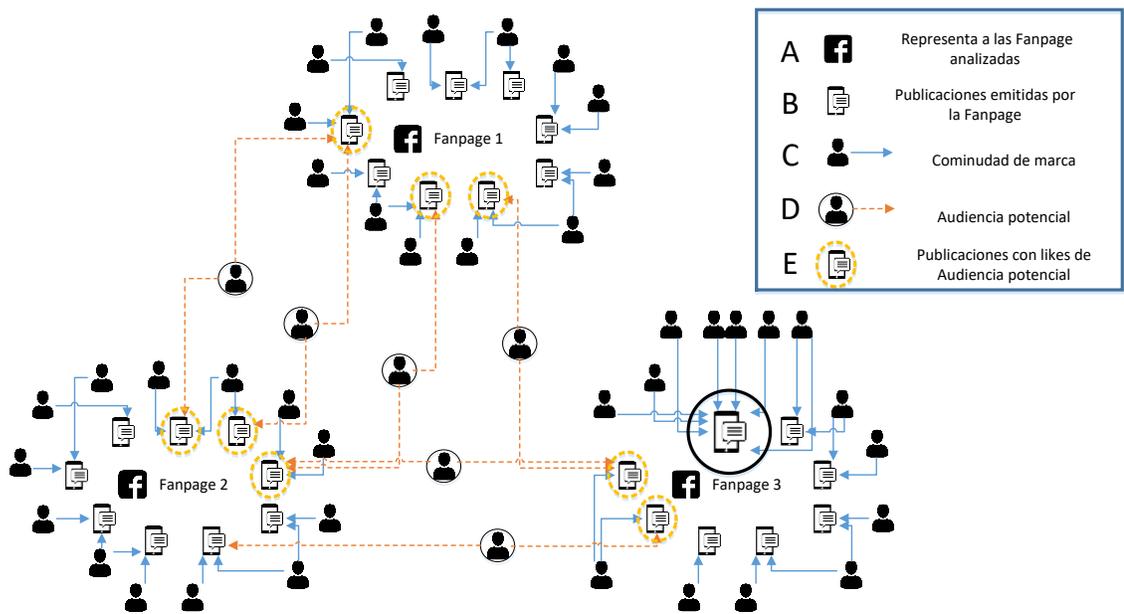


Figura 27. Interacción de likes de usuarios con publicaciones de Fanpage

Fuente: Elaboración propia

Para el segundo punto de corte con respecto a las publicaciones, se eligieron por medio del método de percentiles a aquellas que tuvieron interacción con mínimamente cuatro usuarios de la audiencia potencial, obteniéndose un total de 560 publicaciones, las cuales al clasificarse y analizarse con base al contenido que presentaban (links, imagen, sólo texto y video), se pudo observar (Cuadro 5) que las publicaciones con imágenes son el tipo de publicación favorita de la audiencia potencial además de que ésta genera en promedio mayor cantidad de comentarios, likes y compartidos, por tal motivo se puede asegurar que en Facebook el tipo de contenido recomendado a publicar debe contener imágenes.

Cuadro 5. Tipificación de las publicaciones

Tipo	Número de publicaciones	Likes	Comentarios	Compartidos
Links	42	37.7 b	0.7 b	7.3 b
Imagen	510	114.2 a	3.0 a	18.0 a
Sólo texto	4	26.3 b	2.0 a	2.0 c
Video	4	29.0 b	2.5 a	5.3 b

^{a b} Diferentes literales por columna indican diferencias estadísticamente significativas ($p \leq 0.05$) según prueba de Dunn's

Fuente: Elaboración propia

En referencia a lo dicho por Carlos de la Paz et al. (2016) en cuanto a la importancia de tomar en cuenta los horarios de las publicaciones. Se encontró (Cuadro 6) que a pesar de que el porcentaje de estas publicaciones es mayor a las 10:58 de la mañana, las publicaciones emitidas a las 7:59 de la noche son las que presentan mayor cantidad de comentarios, tienen mayor cantidad de likes y son más compartidas, esto se alinea con lo mencionado por Rodríguez Mareño (2014) pues indica que *“un aspecto importante son las horas en las que se hacen los posteos, no es lo mismo postear por la mañana que por la noche, una empresa que comercializa bebidas alcohólicas el target al cual se dirige son jóvenes, entonces el horario de posteos debe ser donde ellos estén más atentos, por las tardes o las noches en este caso (p.34)”*.

Cuadro 6. Publicaciones clasificadas por hora

Hora de publicación	Porcentaje	Apariciones	Likes	Comentarios	Compartidos
01:10:00 a. m.	9.1%	51	34.0	0.8	1.7
02:59:00 a. m.	0.2%	1	30.0	0.0	7.0
07:51:00 a. m.	6.8%	38	126.9	1.7	7.7
09:44:00 a. m.	0.2%	1	16.0	1.0	1.0
10:13:00 a. m.	0.2%	1	32.0	2.0	0.0
10:54:02 a. m.	0.4%	2	87.5	3.0	9.0
10:58:00 a. m.	32.1%	180	32.7	1.5	4.8
11:06:00 a. m.	0.2%	1	7.0	2.0	8.0
11:09:00 a. m.	1.4%	8	26.5	0.4	8.8
12:33:00 p. m.	0.2%	1	21.0	2.0	15.0
01:28:00 p. m.	8.2%	46	27.4	2.0	2.6
02:30:00 p. m.	0.4%	2	18.0	0.0	6.5
03:26:00 p. m.	2.9%	16	22.1	2.1	12.9
03:32:00 p. m.	3.6%	20	122.5	7.2	11.6
04:28:00 p. m.	0.2%	1	38.0	1.0	40.0
04:39:00 p. m.	0.5%	3	47.0	3.3	5.0
04:49:00 p. m.	7.1%	40	95.5	2.1	7.3
05:15:00 p. m.	11.4%	64	14.8	0.3	3.4
05:46:00 p. m.	0.4%	2	23.5	0.0	3.0
06:16:30 p. m.	8.6%	48	85.9	2.3	59.5

Hora de publicación	Porcentaje	Apariciones	Likes	Comentarios	Compartidos
07:06:46 p. m.	0.5%	3	52.7	3.3	2.0
07:59:00 p. m.	4.5%	25	1340.4	27.1	162.7
10:52:00 p. m.	0.2%	1	128.0	0.0	74.0
11:02:00 p. m.	0.2%	1	19.0	0.0	3.0
11:44:00 p. m.	0.7%	4	20.0	0.3	5.0

Fuente: Elaboración propia

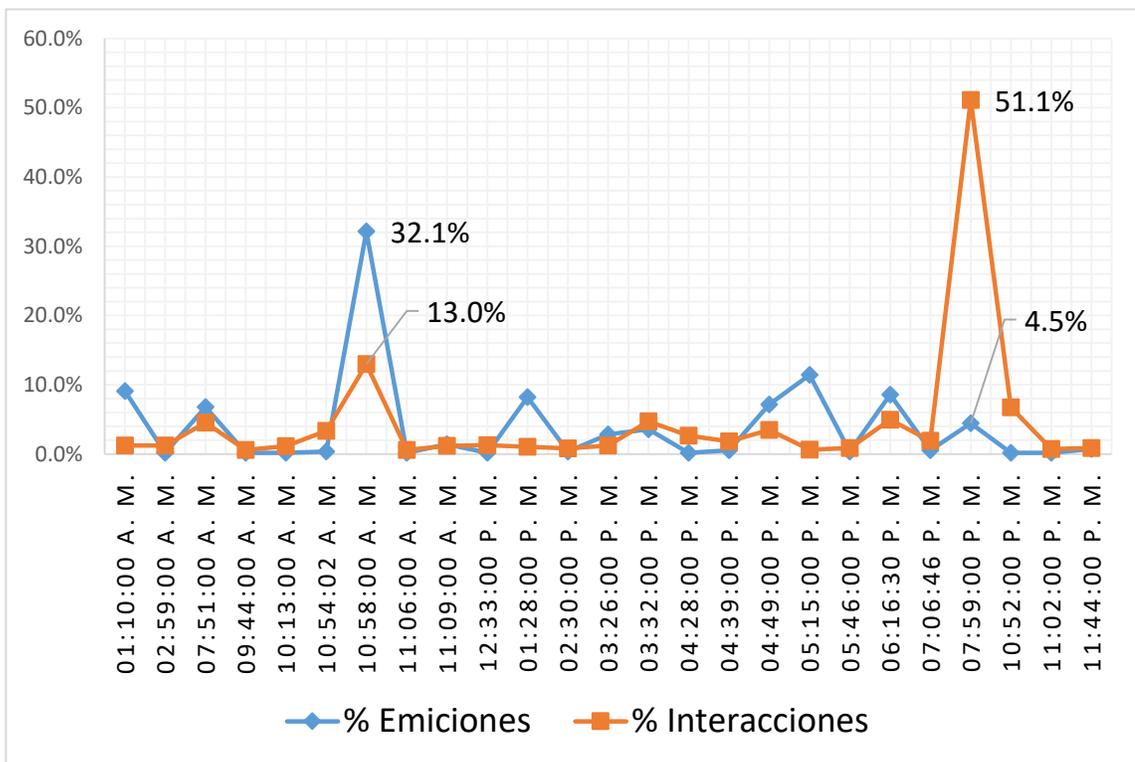


Figura 28. Porcentaje de publicaciones emitidas distribuidas por hora a lo largo del día

Fuente: Elaboración propia

Hasta este punto se cumple con el primer objetivo correspondiente a la información que se puede obtener con respecto al usuario, consiguiendo la lista de los 69 internautas considerados como la audiencia potencial (Apéndice 2) y detectando el horario preferido en el que más interactúa esta audiencia. En aras a solventar el segundo objetivo con respecto al tipo de contenido de interés, las publicaciones fueron clasificadas en 24 categorías no excluyentes entre sí, el primer grupo de estas categorías le antecede la letra “F” la cual significa que incluyen algún tipo de frase, ya sea cómica, motivacional, que invita a tomar o frase mezcalera. Al segundo grupo le antecede la letra “I” la cual significa la presencia de imágenes y éstas pueden ser acerca de paisajes, agaves, imágenes de copas, botellas, stock de botellas, de alguna parte del proceso productivo, de algún productor o de la empresa (instalaciones). La categoría a la cual le antecede la letra “N” corresponde a aquellas publicaciones que presentan notas, estas pueden ser informativas, notas del agave, notas de la empresa o notas acerca del mezcal. Para el caso de las categorías que incluyen eventos se incluyen aquellas que tienen que ver con el mes de septiembre, diciembre o día de muertos, también aquellos que son promocionados por la empresa (Evento), o los que incluyen días festivos en general (Festivo) como el día del padre o el día de la mujer; por último, se mencionan aquellas categorías que hablan de México del estado de Oaxaca o que incluyen publicaciones con nada que ver con el tema mezcal (Nada).

La aparición de las publicaciones clasificadas por categoría se muestra en la Figura 29 en donde se puede observar que aquellas publicaciones que tienen imágenes de agaves, botellas o stock de botellas son las más recurrentes pero las correspondientes a notas informativas, días festivos y frases cómicas generan un mayor porcentaje de diálogo entre los usuarios y la empresa, es decir que son publicaciones que generan más comentarios. Ejemplos de estas categorías se pueden ver en el Apéndice 1.

En comparación con estos resultados se alinean los mostrados por Maria Luz Congosto (2015) quien señala que los mensajes con contenido multimedia como fotos son mucho más virales que los que incluyen enlaces o sólo texto, a estos se abonan los del trabajo de Rodríguez Rangel & Sánchez Rivero (2015) quien en un estudio similar, pero para el caso de gimnasios de tipo fitness, halló que los contenidos multimedia (símil de las categorías con imágenes), información específica (símil a la categoría de Informativa) e imagen de marca eran los más publicados por las cadenas, coincidiendo con el mayor número de likes, siendo también más compartidos y comentados, en referencia a esta última categoría “imagen de marca” los autores la describen como los mensajes utilizados por la organización para crear ambiente positivo en torno a su imagen, mensajes de bienvenida, de ánimo, de agradecimiento y, de motivación algo similar a las frases cómicas o frases mezcateras.

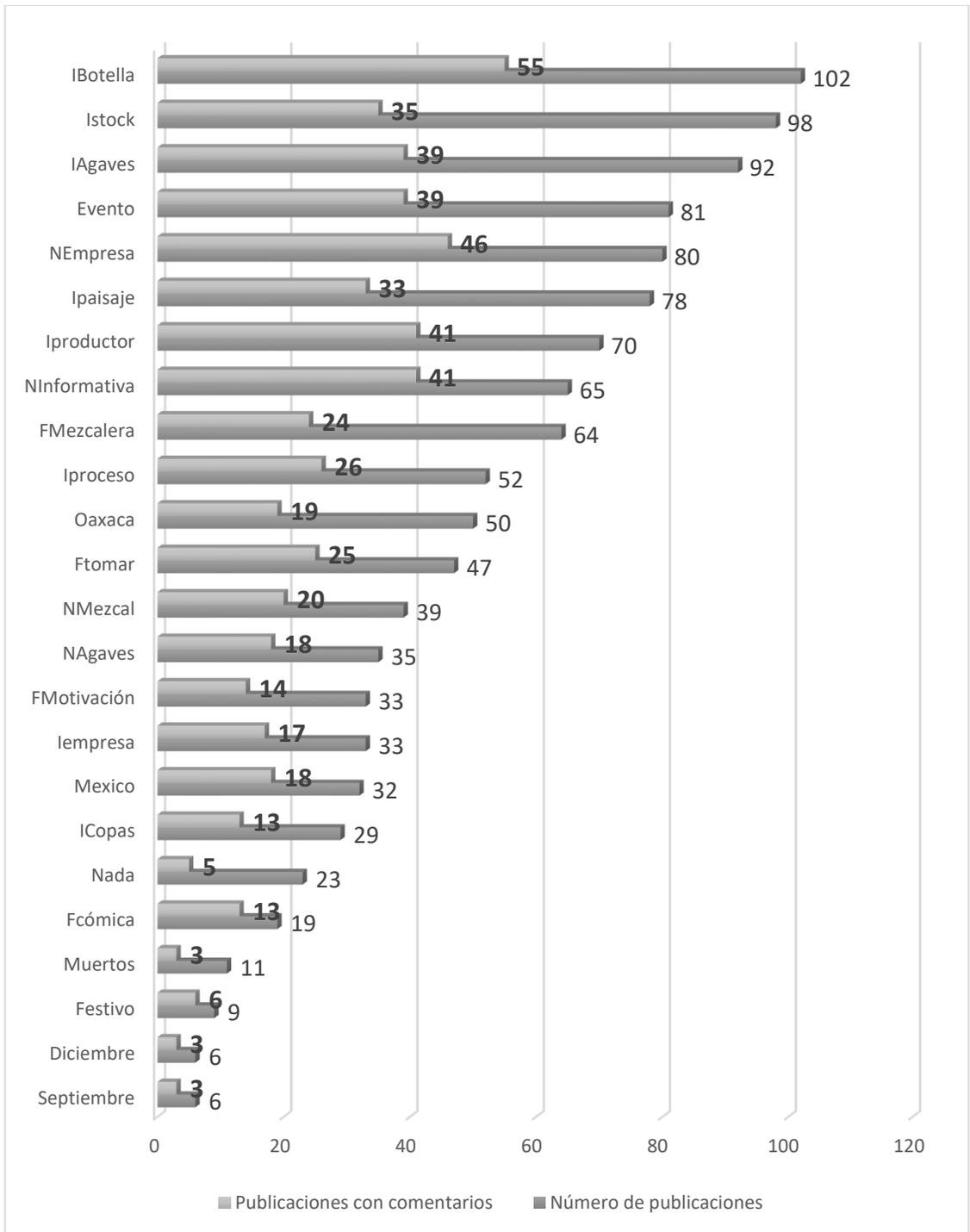


Figura 29. Publicaciones emitidas y publicaciones con comentarios, clasificadas por categoría

Fuente: Elaboración propia

Al realizar un análisis de correspondencia con el fin de identificar qué empresas tienen una mayor preferencia por alguna categoría, enfatizando en las seis categorías importantes por número de publicaciones emitidas y porcentaje de publicaciones con comentarios, podemos ver con apoyo de la Figura 30 que sólo las FanPage de las empresas Rey Campero y Rajabule (1 y 10) son las que están haciendo mayor uso de publicaciones con que contengan notas informativas (29), también que la categoría correspondiente a días festivos (25) no tiene una clara preferencia por alguna empresa, por el contrario las publicaciones relacionadas con frases cómicas (13) son de mayormente usadas por la página de la empresa Sabe Mezcal (4) y en cuanto a las tres categorías con más aparición que corresponden a las que tienen imágenes de botellas de stock de botellas y de agaves (números 19, 20 y 17), podemos decir que únicamente la categoría de publicaciones con imágenes de agaves (17) se relacionan con una empresa (Rey Campero), las otras dos no muestran un favoritismo con respecto a su uso por parte de las páginas de las empresas.

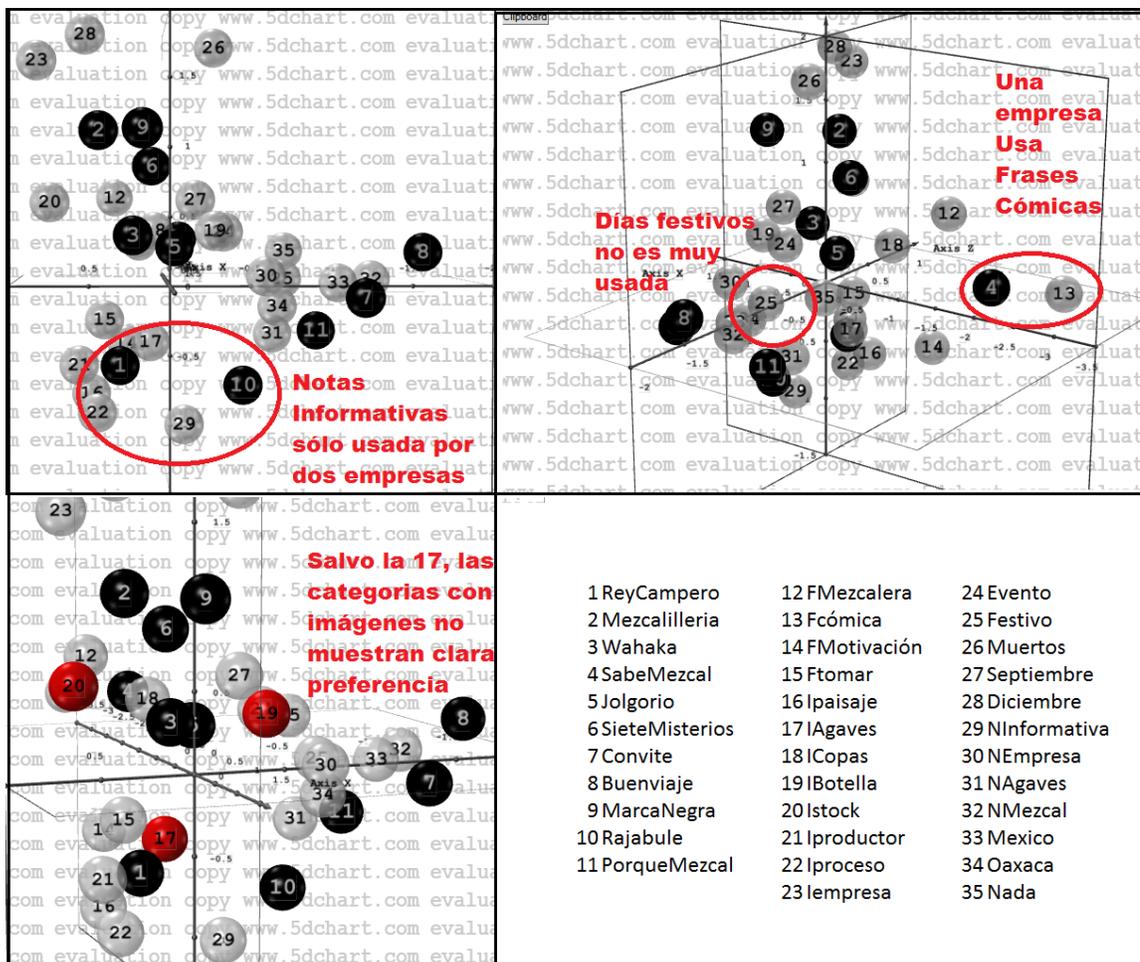


Figura 30. Representación del análisis de correspondencias entre las Fanpage y las categorías de las características de las publicaciones.

Fuente: Elaboración propia

Los análisis anteriores siguieron la metodología propuesta representada en la Figura 31 en donde también se representa disminución en cuanto a número de Fanpage publicaciones y usuarios a raíz del filtrado en busca de la audiencia potencial y los contenidos relevantes para estas audiencias.

Resta describir el paso número siete de la Figura 31 correspondiente a la generación de conocimiento a partir de la información de los pasos (2, 5 y 6); se recomienda en primer lugar interactuar con los 69 usuarios seleccionados en el

paso dos (Apéndice 2), también se recomienda que las publicaciones se hagan en dos horarios, el primero alrededor de las 11 hrs y el segundo alrededor de las 20:00 hrs; las publicaciones deberán incluir de preferencia imágenes de botellas, agaves o stock de botellas; en cuanto al contenido, es preferible presentar notas informativas, que hagan alusión a días festivos (no necesariamente de septiembre o diciembre) con frases cómicas para generar diálogo con la audiencia, ejemplo de este tipo de publicaciones se hallan en el Apéndice 1.

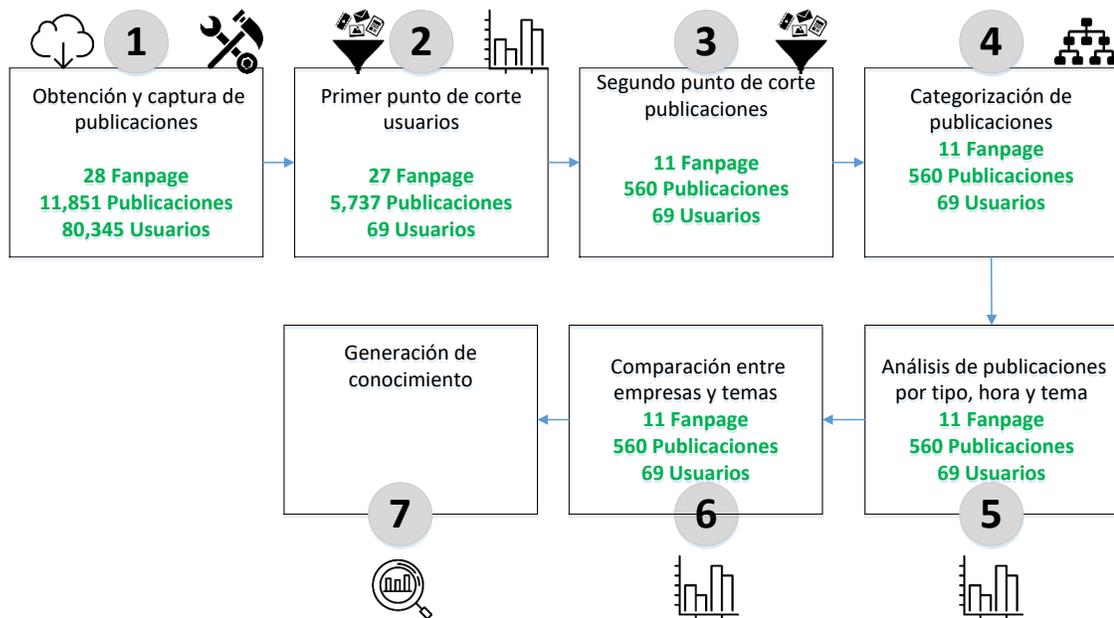


Figura 31. Pasos que se siguieron para la generación de conocimiento en cuanto a Fanpage de Facebook

Fuente Elaboración propia

4.3. Twitter

Se extrajeron y analizaron un total de 57,362 tweets que tuvieran dentro del texto la palabra mezcal. Estos mensajes fueron escritos por 33,468 usuarios diferentes; la extracción fue intermitente abarcado un periodo del 23/dic/2015 a las 11:46:06 a. m. hasta el día 12/jul/2016 a las 10:36:23 p. m. Se identificaron que 65% de los tweets fueron en español y 5 % en inglés.

De los 33,468 usuarios diferentes que escribieron un tweet (Mención, Retweet, Reply), el 79.7% participaron con solo un mensaje, el 11% escribió dos tweets, el 8.2% redactaron entre tres y nueve mensajes y el 1.2% (396 cuentas) mandaron arriba de diez tweets.

De esta forma se identificaron aquellas cuentas que tenían más participación en cuanto a la publicación de mensajes con la palabra mezcal agregando variables como número de cuentas siguiendo, número de seguidores, total de tweets escritos, etc. Las cuentas con más de 50 tweets escritos sobre mezcal se muestran en el (Apéndice 3). Sin embargo, es importante destacar que no se debe de tomar únicamente el número de tweets o el número de seguidores que posee cada cuenta ya que tal y como menciona Del Fresno García, Daly, & Segado Sánchez-Cabezudo (2016) identificar a los *“influencer”* solo a través de su actividad (independientemente del contenido) y/o a través del conteo de seguidores o amigos no es una aproximación válida y conducirá a resultados pobres. López & Sicilia (2016) apoyan esta idea al decir que se han considerado líderes de opinión a aquellos individuos que tienen muchos seguidores en Twitter, entendiendo que si una persona tiene muchos seguidores es porque a muchos consumidores les interesa lo que escribe esa persona, sin embargo, esta medida puede manipularse y puede no implicar influencia.

Con base en la argumentación anterior se decidió realizar un análisis más riguroso. Siguiendo las recomendaciones de Del Fresno García et al. (2016), se desagregaron el total de tweets en hashtags, retweets y menciones, a continuación se presenta en un primer momento el análisis que nos dan un acercamiento hacia los temas de interés de la audiencia en Twitter con respecto al producto mezcal, éste realizado por medio de el mapeo de redes de los hashtag y los retweets en su primer forma “retweets 1” (donde la interacción es de usuario a tweet). En un segundo momento se presentan los resultados de aquellos análisis que nos acercan al perfil del usuario con el cual es recomendable interactuar, mostrando en un primer momento la ubicación de la audiencia y posteriormente la identificación de las comunidades y los líderes de

opinión de cada comunidad , esto mediante un mapeo de redes de menciones de retweets en su segunda forma “retweets 2” (donde se muestra una interacción de usuario con usuario).

4.3.1. Identificación del contenido de interés (análisis de hashtag y primera forma de retweets 1)

Del total de tweets, un 43.7% (25,124 tweets) contienen al menos un hashtag (HT), un máximo de 18 y un promedio de 2.2. Se encontraron un total de 8,945 hashtags diferentes emitidos por 12,400 usuarios, el HT #mezcal fue el que más se repitió, apareciendo 13,414 veces (grado de entrada).

Al obtener los hashtags más importantes (sin contar el HT #mezcal) con sus grados de entrada, se destaca que las 10 etiquetas que más relacionan los usuarios de Twitter (Figura 32) con el mezcal son: #Oaxaca (3,518), #México (1,247), #Denver (947), #Tequila (734), #Agave (651), #Cultura (296), #Cocktails(280), #MezcalesdeOaxaca (256), #TwitterOax (237) y #Sellamamezcal (224)

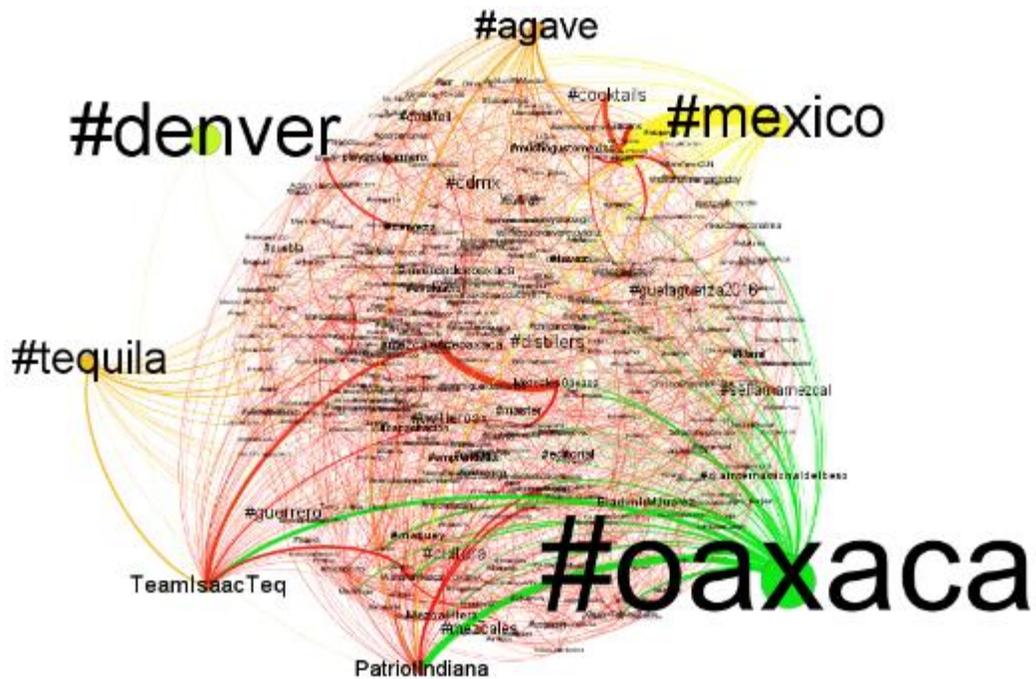


Figura 32. Red de hashtags (HT) representada al 1.98% de nodos visibles y 5.18% de aristas visibles para su mejor visualización.

Fuente: Elaboración propia

A pesar de existir 8,944 hashtag diferentes, únicamente sobresalieron cinco y se nota la gran diferencia en grados de entrada entre la posición cinco “#agave” con 651 y la posición seis “#cultura” con 296, lo que significa que existen pocas palabras dominantes que identifican a la bebida mezcal, las primeras dos en orden de dominancia #Oaxaca y #México, que precisamente son lugares con los cuales se identifica esta bebida, pues es emblemática a nivel nacional e internacional. Los hashtags correspondientes a la cuarta y quinta posición son #tequila y #agave, respectivamente; se pueden explicar por la asociación que el mezcal presenta con estas palabras y que se ha notado a lo largo de esta investigación.

Situaciones a destacar con respecto a las demás posiciones (se pueden consultar el detalle en el Apéndice 4) son que:

- a. El HT #cocktails ocupa la octava posición y da pie a pensar que se opta por alternativas en cuanto a la preparación del mezcal que tradicionalmente se toma sólo.
- b. El HT #sellamamezcal ocupa la décima posición en respuesta a la norma emitida por la secretaría de economía en donde mencionaban cambiar el nombre de mezcal a komil de algunas bebidas a las cuales consideraba apócrifas, por tal motivo el auge de este hashtag que en contraposición se encontró el #komil en la posición 45.
- c. Después de #Oaxaca los estados que le continuaron en asociación con el mezcal fueron #guerrero en la posición 11, #cdmx en la posición 13, #durango ocupando la posición 24 y #puebla con #michoacan en las posiciones 28 y 31, respectivamente. De esto se destaca que Guerrero es el segundo estado con el cual asocian a la bebida y la Ciudad de México, a pesar de no ser una entidad productora, también destaca posiblemente por tweets emitidos por consumidores en la Ciudad.
- d. En cuanto a las marcas, estas no aparecen en las primeras 20 posiciones sino hasta la posición 25 con #silenciomezcal seguido de #reycampero en la 34 y el mezcal #sacapalabras en la posición 49. Esto responde a lo mencionado por Aguado Guadalupe & García García (2009) que mencionan como elemento clave del marketing viral a la credibilidad y que cuando se añade el aspecto comercial, las personas lo notan y automáticamente pierde su interés y su credibilidad, por tal motivo es difícil asociar lo viral a una marca comercial.

En cuanto al hashtag #denver el cual ocupa el tercer lugar se explica como un caso inusual ya que está dado por un evento circunstancial. El hashtag se incrusta dentro de un tweet con una particular resonancia dada por la intervención de dos figuras públicas (periodistas) con un mensaje de desaprobación a una tercera (político) que conjugó entre otras cosas; la empatía de un sector grande de la audiencia con el mensaje, la inclusión de manera cómica de palabras altisonantes, que fue dado en el marco de un evento como son elecciones

presidenciales y el acoplamiento del nombre de la marca “ilegal “con el contexto que envolvía el mensaje.



Figura 33. Tweet viral emitido por el periodista mexicano Joaquín López Dóriga

Fuente: Extraído de la cuenta de Twitter @Lopezdoriga

Del total de tweets, 39.4% fueron retweets (22,603) de 14,678 usuarios diferentes a 7,872 tweets diferentes. Recordando la Figura 17, en este apartado se toman los vínculos que tuvieron las cuentas al compartir los diferentes tweets, en la se aprecia la red de tweets más compartidos (la lista de los 100 primeros se encuentra en el Apéndice 5).

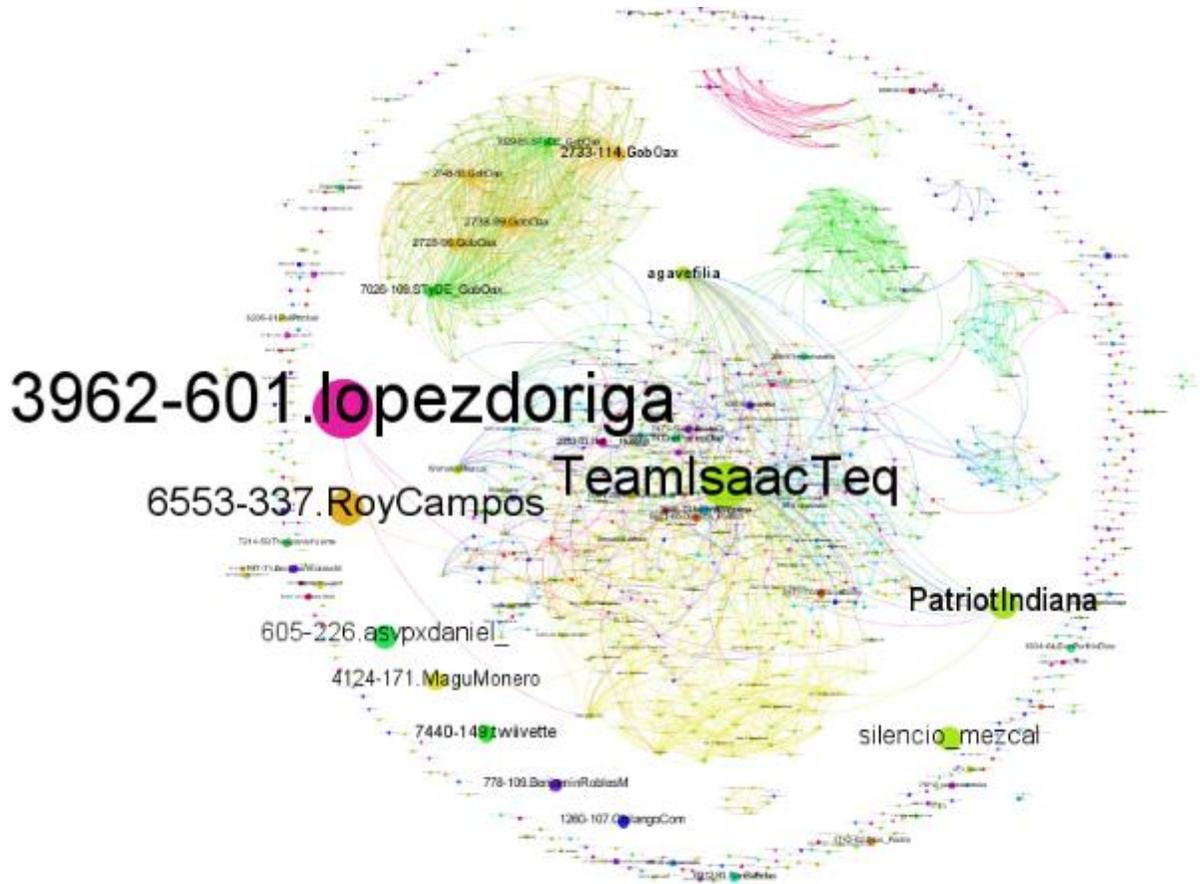


Figura 34. Red de Retweets 1 con la palabra mezcal, filtro al 4.7% de nodos para su mejor visualización

Fuente elaboración propia a partir de datos de la API

El tweet más compartido (601 veces) fue el emitido por el periodista Joaquín López Doriga secundado por el de la cuenta @Roycampos que igualmente es una figura pública y casualmente es una mención del mismo mensaje (Figura 35). Lo cual explica su efecto viral, a parte de la coyuntura de efectos mencionados anteriormente, por el hecho de que tal y como afirma Aguado Guadalupe & García García (2009) el mensaje utiliza la motivación humana como lo es el sentimiento de pertenencia hacia una ideología además de características como ser gracioso, controvertido o absurdo.



Figura 35. Tweet viral emitido por la cuenta @Roycampos

Fuente: Extraído de la cuenta de usuario de Roy Campos

Después de esos dos tweets, el siguiente en relevancia fue el que tiene el código **605-226.asvpxdaniel_** (Figura 36), este tweet es sobre el video de una canción de música regional mexicana que tiene en su título la palabra mezcal, la conjetura acerca del mérito de este tweet es que llega a un sector grande la población que le gusta este tipo de música y temática. Se destaca que este tipo de contenido con respecto a la música regional mexicana se encontró previamente en el análisis de las palabras clave realizadas en Google, por lo que abre la puerta a reflexionar acerca de la posibilidad de promocionarse a través de la asociación con este tipo de artistas, ya que como mencionan Castelló Martínez & del Pino Romero (2015) el uso de las celebridades para promocionar una marca se basa en una dimensión aspiracional de la audiencia y en mecanismos de identificación de ésta con el personaje que ayudan a la formación del autoconcepto ideal, presentándose al famoso como representante de las inclinaciones del target. En este caso la empatía se generaría con un segmento de la población que disfruta

de este tipo de música y toma como “representantes del pueblo” a los encargados de la interpretación de este género.



Figura 36. Tweet que hace referencia a una canción de música regional mexicana

Fuente: Extraída de la cuenta de @asvpxdaniel_

El tercer tweet en orden de relevancia por grado de entrada es el que tiene el código **4124-171.MAguMonero**, este tweet fue emitido por la cuenta de un caricaturista de un periódico en México, en donde en una imagen de naturaleza cómica, hace alusión al enfrentamiento de maestros con fuerzas policiacas del gobierno del estado de Oaxaca, invitando a la negociación con una mesa de diálogo llena de productos oaxaqueños, incluidos el mezcal. El posicionamiento de este mensaje debe gran parte de la pertinencia con la que se inserta ante un

evento noticioso, aunado a la popularidad del personaje que emitió el tweet y el sentido caricaturístico apelando al humor.



Figura 37. Tweet que hace referencia al mezcal con un hecho noticioso

Fuente: extraído de la cuenta de @MAGuMonero

El cuarto tweet con folio **7440-149.Twiivette** que fue más veces compartido fue emitido por una cuenta con número de seguidores mayor a 14,000, y que se dedica a hablar de moda y estilo en un formato de blogger. El tweet (Figura 38) es la conexión hacia un post de un blog que versa acerca de una exposición de mezcal en el museo de arte de Polanco (en la Ciudad de México) en donde se presentaron obras de arte de artistas Oaxaqueños; además de, presentar la historia del agave y el mezcal, su creación, sus usos, como olerlo y degustarlo, los años de añejamiento, como combinarlo, entre otras cosas. Este tweet responde hacia un sector de la población con alto nivel económico que presenta un interés hacia la cultura mexicana y oaxaqueña, dicho argumento concuerda con las geolocalizaciones encontradas anteriormente en zonas de alto poder económico de la ciudad de México, que nos reiteran el pensamiento de que el

mercado de mezcal se está posicionando en público meta que anteriormente no se consideraban.



Figura 38. Tweet acerca de un post sobre la cultura del mezcal en la ciudad de México

Fuete extraído de la cuenta de @ Twiivette

El siguiente ejemplo ocupó el sexto lugar en orden de trascendencia, es un tweet emitido por la cuenta del gobierno oaxaqueño (Figura 39) que presenta una serie de imágenes alusivas a la cultura del mezcal además de presentar un link hacia una página externa que contiene más información acerca del tema.



Figura 39. Tweet del gobierno de Oaxaca que habla sobre la tradición del mezcal

Fuente: extraído de la cuenta de @GobOax

Se puede concluir que, salvo los últimos dos casos en los tweets que han sido mayormente compartidos, el concepto de mezcal se presenta más como un factor adyacente al mensaje principal, lo anterior sugiere que una táctica de difusión de un mensaje se puede dar colateralmente a un mensaje principal de mayor impacto, aprovechando acontecimientos de la vida cotidiana (las manifestaciones de los maestros o elecciones gubernamentales) siempre cuidando la asertiva sinergia entre el acontecimiento, los estímulos sentimentales que se desean provocar y la marca pues, si bien como menciona Aguado Guadalupe & García García (2009) es difícil asociar la viralidad con lo comercial, el no tener un enfoque claro podría acabar por tener una gran cantidad de tráfico sin ningún valor y con costes, es decir la audiencia a la que llegue esta información .a pesar del volumen ,puede que no se ajusten a los objetivos de la campaña de mercadeo ya que no son el público meta, aquí se retoman las palabras de la ADECEC (2016) que dice que no se trata de llegar a muchos si no a los adecuados. Por otra parte los costos que implicarían abarcan desde la conexión con otros usuarios que difundieran el mensaje hasta el pago por mensajes publicitarios a

figuras públicas que según Castelló Martínez & del Pino Romero (2015) puede estar entre los 20 y 30 mil euros.

Los mismos autores señalan que el concepto de “viral” implica dejar la conversación en manos de los consumidores, por lo que la marca, lo más que puede hacer, es incentivar esta conversación, ofreciendo algo a sus clientes que merezca la pena comentar (contenidos, noticias, servicios exclusivos, una gran experiencia de uso, etc.) como es el caso de los últimos dos tweets analizados, los cuales ofrecen contenido a su audiencia.

4.3.2. Acercamiento al público meta (ubicación, análisis de menciones y segundo análisis de retweets 2)

Siguiendo con la parte de búsqueda de información acerca de la audiencia se comenzó con el posicionamiento de 1,987 tweets que contaban con coordenadas, lo anterior con el objetivo de identificar la ubicación de la audiencia que más escribía mensajes en torno al producto. Las geoposiciones se pueden observar en las siguientes figuras (Figura 40, Figura 41 y Figura 42) y cuyo mapa completo se puede consultar en el siguiente link <https://goo.gl/RqT9R7>



Figura 40. Distribución de tweets sobre mezcal a nivel mundial

Fuente: Elaboración propia



Figura 41. Distribución de tweets sobre mezcal en México

Fuente: Elaboración propia

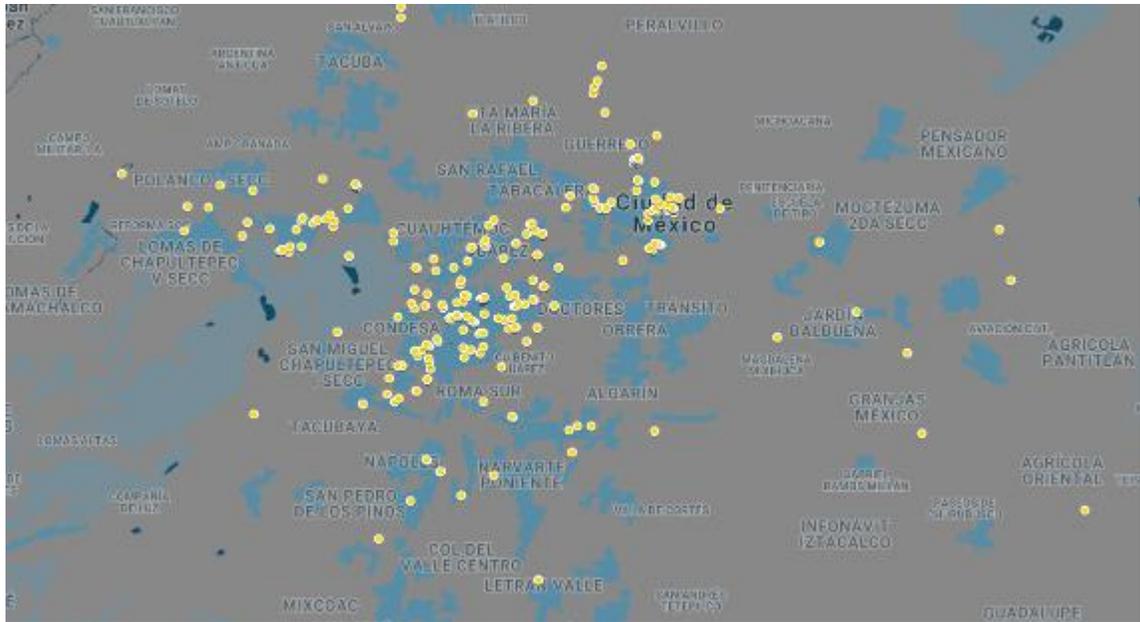


Figura 42. Distribución de tweets sobre mezcal en la Ciudad de México

Fuente: Elaboración propia

En las figuras anteriores se puede apreciar que los tweets acerca del mezcal se distribuyen en su mayoría en América del norte y Europa. Para el caso de México, las principales zonas son la Ciudad de México, Oaxaca y Guadalajara, esto concuerda con lo hallado con la identificación de palabras clave en el buscado de Google. En un zoom a la Ciudad de México se encuentra que, en cuanto a los tweets emitidos, la zona centro es la principal abarcando lugares con alto nivel económico como son las colonias Condesa, Roma, Polanco y Lomas de Chapultepec lo cual da pie a pensar en la revalorización del producto por parte del consumidor como señalan Espinoza Ortega & López Rosas (2016).

Una vez teniendo un panorama general de la ubicación de la audiencia se procedió a la búsqueda de comunidades y líderes de opinión dentro de éstas, para tal efecto se recurre al análisis de menciones y retweets de la forma interacción usuario con usuario.

Para el caso de menciones, se encontró que 16,801 tweets tienen menciones (29.2%), con máximo 11 y promedio de 1.5. Se hicieron 24,762 menciones a 8,194 cuentas diferentes. Se reitera que en esta parte los grafos de redes se basaron en la interacción de cuenta con cuenta (ver Figura 17) es decir cada nodo representa una cuenta de usuario y la dirección de la flecha desde un nodo a otro representa la interacción por medio de una mención por tal motivo cada nodo tiene la posibilidad de tener tantos grados de entrada (cuando otra cuenta los menciona) como grados de salida (cuando mencionan a otro usuario).

Al analizar la totalidad de la red de menciones y tomando como indicador de selección la centralidad de grado y más concretamente los grados de entrada de cada actor se identificaron como las 10 primeras posiciones (primeras 100 posiciones están en el Apéndice 6) a las cuentas: @lopezdoriga, @STyDE_GobOax, @YouTube, @GobOax, @Sabemezcal, @GabinoCue, @Change_Mex, @NacionMezcal, @LeyendaEterna y @SaintAnejo. No obstante retomando los argumentos de varios autores (Abejón, Sastre, & Linares, 2012; Aguado Guadalupe & García García, 2009; Carlos de la Paz et al., 2016; López & Sicilia, 2016) se plantea la necesidad de identificar comunidades ya que hay que comprender que estamos ante un modelo comunicativo en el que, para poder competir, las compañías deben comprender que no venden los productos a clientes individuales sino más bien a redes de clientes.

Al hacer el análisis de agrupación se halló una modularidad cercana a uno (0.917), el cual, recordando lo mencionado por Arenas et al. (2008), es el máximo índice de concentración de los nodos dentro de los módulos en comparación con una distribución al azar de los enlaces entre todos los nodos. También se hallaron un total de 2,444 comunidades, de las cuales se eligieron las primeras tres. Éstas abarcaron 10.20% del total de nodos y 16.07% del total de conexiones. Es importante destacar que dentro de las tres comunidades más grandes sólo se hallaron dos de las cuentas propuestas al principio únicamente por el grado de entrada (@NacionMezcal y @SaintAnejo.). Lo anterior se explica por situaciones como que los mensajes emitidos a pesar de tener un alto grado de interacción,

son esporádicos o por cuentas que no están relacionadas con el tema. Por tal motivo no generan comunidad, tal es el caso de @Change_Mex o la cuenta de @lopezdoriga las cuales surgieron de manera fortuita. Otro de los posibles motivos es que pertenecen a comunidades cerradas o pequeñas como podría ser el caso de las tres cuentas pertenecientes a instancias gubernamentales (@STyDE_GobOax, @GobOax, y @GabinoCue).

Al aislar las comunidades más grandes se eligieron como “*influencers*” a los que muestren un mayor grado de entrada lo anterior por la naturaleza de las menciones en donde no transita el mensaje entre nodos a más de un paso, más bien funciona como un indicador de prestigio, es decir si yo emito un tweet con mención hacia un usuario específico es porque me interesa que éste sea leído por el usuario al que menciono.

Entre las cuentas consideradas como “*influencers*”, se hallaron a: @NacionMezcal, @SaintAnejo, @PorQueMezcal, @lalanoguera, @SoyMezcalero, @liladowns y @sabormexico. Cabe señalar que, dentro de estas comunidades existe la presencia de una gran diversidad de perfiles incluyendo la presencia de marcas, esto implica que una vez obtenida la lista de “*influencers*” se requiere incursionar más en los posibles candidatos con los que se decida generar una interacción o un vínculo en función de la naturaleza y valores de la marca y los objetivos que ésta se plantee, la finalidad de esta metodología es la obtención de la lista de los “*influencers*” de esta comunidad y para este ejercicio la lista se encuentra en el Apéndice 7.

indicador de PageRank (épsilon 0.001, probabilidad de 0.85) como variable determinante para la identificación de “*influencers*”.

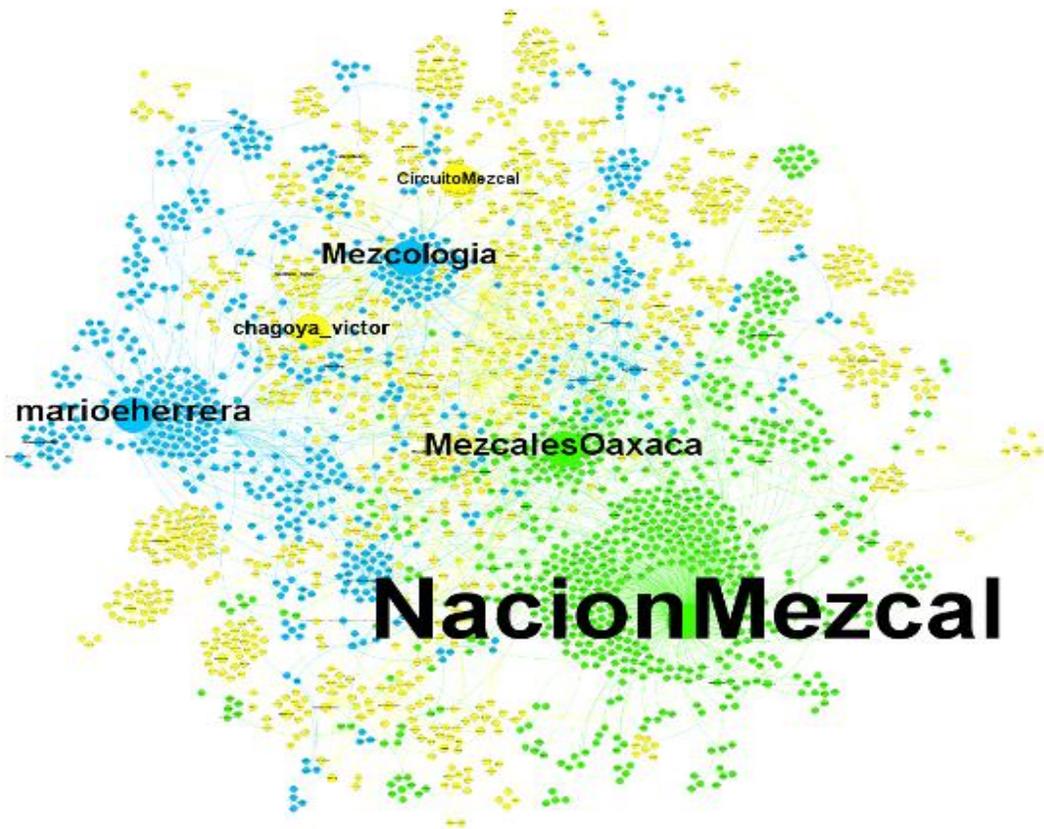


Figura 44. Comunidades más grandes de la red de retweets

Fuente: Elaboración propia

Entre las cuentas influyentes se hallaron a @NacionMezcal, @MezcalesOaxaca, @Marioherrera, @animalgourmet, @VisitGuerreroMx y @Sectur_Guerrero, al igual que en el caso anterior queda en función de la empresa la elección de las

cuentas con las que quiera generar una mayor relación, el listado completo de “*influencers*” se puede hallar en el Apéndice 8

A grandes rasgos esta metodología no pretende determinar específicamente los actores con los cuales se necesita generar una conexión, la decisión de con quien interactuar queda con base en la naturaleza de la empresa, sus valores misión y visión y los objetivos que busque: Los resultados de la metodología para este caso más bien propone, una lista seleccionada (Apéndice 7 y Apéndice 8) de entre los miles de posibles candidatos de usuarios con los cuales se puede relacionar, lo anterior en palabras de la ADECEC (2016) no es tarea fácil requiere tiempo, dedicación y esfuerzo.

El mérito de los usuarios propuestos es que no caen en el error de ser seleccionados únicamente por los números brutos como el número de seguidores, mencionado por Sánchez (2012), o el número de tweets escritos, más bien se indaga en su importancia en prestigio (como las menciones) y alcance del mensaje (retweets) no tanto de la red en general si no dentro de comunidades preexistentes, siendo conscientes de los nuevos tipos de marketing, el marketing viral (Túñez-López et al., 2011) y el marketing de influencia (Castelló Martínez & del Pino Romero, 2015) y tomando en cuenta la importancia descrita por varios autores (Abejón et al., 2012; Carlos de la Paz et al., 2016; López & Sicilia, 2016) acerca de las comunidades dentro del entorno virtual.

A pesar de que se intentó hacer una tipología acerca del perfil de algunas cuentas (Apéndice 3), se considera que ésta es difícil dado que la información que se obtiene en redes sociales sobre el perfil de los usuarios se limita a su nombre de usuario y fotografía y, al no haber variables de corte, la clasificación es subjetiva de quien la realiza y cuestionable.

Para finalizar, se aborda la generación de conocimiento a partir de la información obtenida, es decir responder a la pregunta de ¿el para qué de esta información?, encausando lo encontrado hacia una recomendación útil para el posicionamiento

de una empresa de mezcal en la red social Twitter. Se destaca en primer lugar que se pudieron encontrar a aquellos usuarios que en teoría son agentes potenciales para viralizar un mensaje dentro de comunidades donde el mezcal es el tema principal; estos agentes se encuentran enlistados en los anexos (Apéndice 7 y Apéndice 8) y son con los que se aconseja empezar a interactuar ya que son líderes de opinión dentro de sus comunidades y el que ellos mencionen o compartan un tweet de la empresa mezcalera ayudará al maximizar el alcance de ese mensaje. Se debe considerar que para esta red social al igual que con las otras plataformas, el público meta se halla en la parte centro de la república mexicana. Ahora respondiendo a la incógnita con respecto al tipo de contenido por el cual la empresa deba apostar en esta red social, se recomienda que se aventure con contenidos que involucren al mezcal con Oaxaca, México, tequila y agave. Sin embargo se reitera lo mencionado anteriormente, que una táctica de difusión de un mensaje en twitter se puede dar colateralmente a un mensaje principal de mayor impacto, aprovechando acontecimientos de la vida cotidiana, siempre cuidando la asertiva sinergia entre el acontecimiento y los estímulos sentimentales que se desean provocar. En otras palabras, que se aprovechen sucesos políticos, deportivos, irónicos etc. que sean virales en internet para empatizar con el sentimiento del usuario con respecto a estos eventos y relacionarlo con el producto y la marca, como lo hicieron con el tweet de Donald Trump simpatizando con el sentimiento de enfado y en un contexto actual e importante como lo era un proceso electoral dado en ese momento.

5. CONCLUSIONES

En el presente trabajo de investigación se destaca la importancia del tratamiento de datos provenientes de la web para la generación de conocimiento que permita el desarrollo de una estrategia de posicionamiento para una empresa agropecuaria en entornos digitales.

En correspondencia con las preguntas de investigación y los objetivos planteados, se puede decir que se logró identificar, describir y ejemplificar el tipo de información que se puede obtener tanto de usuarios como de temas de interés, de los cuales se concluye lo siguiente:

Con respecto a la información de los internautas que se relacionan en la web con algún producto, el caso del buscador de Google es el más equilibrado, ya que ofrece datos duros más asimilables como son: edad, género, ubicación, etc. Por el contrario, la obtención de estos tipos de datos en las redes sociales es prácticamente imposible, ya que al manejar identidades, los mecanismos de privacidad en estas plataformas impiden la incursión en la información del usuario, facultándonos muchas veces en ver únicamente la foto y el nombre de los usuarios y no se encontró herramienta que permita acceder a datos privados de los usuarios; por tal motivo, y al no haber variables de corte, se optó por omitir cualquier tipo de clasificación con respecto al perfil del usuario en redes sociales, ya que su tipificación resultaría ser subjetiva, laboriosa y cuestionable. Por ejemplo, en el caso del mezcal se pudo identificar el perfil de edades, género, y localización de los usuarios que más hacían búsquedas con respecto a este producto.

En las redes sociales por su parte, si bien no se puede acceder a información exacta del perfil, la información de la actividad que los internautas realizan es prácticamente pública y se mantiene en el tiempo, permitiéndonos así identificar y analizar de manera periódica cada una de las reacciones de esta audiencia a y monitorear las relaciones que establecen con otros internautas con la ventaja de

obtener listados puntuales con la clave de identificación de cada uno. Para el caso del mezcal se logró identificar los líderes de opinión y los clientes potenciales.

En la red social Facebook particularmente, se pudo obtener la lista de aquellos usuarios (interesados en un tema o producto determinado, mezcal por ejemplo, y que no necesariamente presentan afinidad por alguna marca) con los cuales se puede empezar a conectar para ir formando comunidad de marca.

En Twitter, por su parte, se pudo obtener igualmente de manera puntual el listado de usuarios líderes de opinión dentro de sus comunidades en el tema mezcal encontrándose entre estos a periodistas, chefs de cocina y promotores de la cultura oaxaqueña, con los cuales se puede conectar y así permitir posteriormente catalizar la difusión de un mensaje publicitario

En relación al segundo objetivo de la investigación, el cual buscaba indagar acerca del tipo de información que se puede obtener con respecto a los temas tendencia en internet en torno a un producto, se concluye que la identificación de los temas de interés es más laboriosa, ya que este procedimiento no se puede automatizar o siquiera estandarizar dada la inminente y marcada intervención lógica en la categorización de los datos, el cual depende de la naturaleza del contenido y de los objetivos que se planteen. No obstante, una vez terminada la categorización y elegido el método de análisis, la identificación de los temas sobresalientes se facilita. Así, en el caso del mezcal se decidió agrupar las tendencias en el buscador de Google categorías como información sobre el mezcal, formas de preparar mezcal, diferencias entre mezcal y tequila etc. y en redes sociales por las etiquetas como Oaxaca, tequila, agave entre otras.

Lo anterior nos encamina a dar respuesta a la tercera incógnita planteada en la investigación, de la cual se concluye que el proceso de transformación de datos a conocimiento se dificulta debido a las características intrínsecas del propio Big Data, ya que al ser datos no estructurados, la variabilidad de éstos demanda una

asertiva clasificación; además, el volumen de los datos precisa que se tenga un sistema de depuración para quitar el “ruido” y dejar materia prima de utilidad para el análisis. Por tal motivo, se puede decir que estos dos pasos en el proceso son importantes, ya que son muy difíciles de categorizar y requieren de una intervención lógica por parte del investigador que sea a fin a las metas que desee conseguir con estos datos.

Además, es importante señalar que aparte de contar con datos depurados y categorizados como materia prima, es indispensable contar con una serie de pericias en manejo de software y herramientas metodológicas de minería y análisis de datos. Lo anterior plantea la necesidad de personal calificado en este ámbito ya que algunas de las empresas delegan la parte de atención en plataformas digitales a gente joven o becarios sin experiencia en el ramo que erróneamente, por su edad y afinidad al uso de estas plataformas, parecen ser la mejor opción disponible. Por otra parte, las empresas que dejan el manejo de sus redes a profesionales requieren tener en cuenta la existencia de supuestos expertos en esta área (como se notó en la revisión de literatura) que basan muchos de sus indicadores en datos brutos (sin tratamiento), como lo son el número de seguidores, número de likes, número de publicaciones o tweets emitidos, y que al no hacer un análisis más minucioso llegan a conjeturas erróneas.

6. LIMITACIONES Y RECOMENDACIONES

Como cualquier proyecto, esta investigación tiene áreas de mejora y limitantes, algunas de las cuales se puntualizan a continuación.

Para el proyecto de extracción de datos se echó mano de la interfaz de programación de aplicaciones tanto de Facebook como de Twitter por la comodidad de extraerlos a nuestra preferencia y de manera gratuita; sin embargo, existen un sin número de aplicaciones gratuitas que se encargan de extraer los datos de una manera sencilla y amigable que no exigen más que dar una serie de clics, incluso algunos son capaces de generar informes poco robustos, pero que pueden servir de orientación.

Los procesos de categorización y filtrado podrían ser cuestionables o mejorables, ya que dependen mucho de la lógica del analista, al igual que las metodologías y los indicadores usados para cada proceso.

Se requiere complementar este tipo de investigación con datos propios generados dentro de cualquier empresa; además de analizar sus ventas y su cartera de clientes.

Se requiere validar lo encontrado de manera real con la implementación de una estrategia de posicionamiento y ver tanto los aciertos como las áreas de mejora. Lo cual llama a la necesidad de investigaciones empíricas basadas en algunas de las plataformas mencionadas.

El análisis que se realizó corresponde sólo a una fracción de tiempo, dada la velocidad con la que se generan los datos. Por tanto, se recomienda que este análisis se haga periódicamente para observar la evolución de los tópicos y los líderes de opinión, la desaparición de unos y la incursión de otros.

Es recomendable también incluir en otras investigaciones el uso y análisis de otras plataformas virtuales como los son YouTube, LinkedIn, WhatsApp, entre otros posibles; ya que los medios de comunicación de los internautas también se

van modificando, algunos van quedando desplazados y otros van tomando terreno, como el caso de WhatsApp que se posicionó hasta el segundo puesto como la red social con mayor penetración, o YouTube que es el segundo buscador después de Google. Es tarea de las empresas estar en donde se encuentra su audiencia.

El uso de los buscadores y de las redes sociales no es la panacea de las empresas, requieren tomar en serio otros aspectos como la innovación y mejora continua en el desarrollo del producto o servicio.

7. LITERATURA CITADA

- Abejón, P., Sastre, A., & Linares, V. (2012). Facebook y twitter en campañas electorales en España. *Disertaciones: Anuario Electrónico de Estudios En Comunicación Social*, 5(1), 129–159. Retrieved from <http://dialnet.unirioja.es/descarga/articulo/4041937.pdf>
<http://dialnet.unirioja.es/servlet/extart?codigo=4041937>
- Acosta V., G. L. (2012). Gestión de pasiones y polarización en las redes sociales. Un análisis del aplicativo Grupos en Facebook. *Discurso & Sociedad*, 6(4), 684–719.
- ADECEC. (2016). Trabajar con influencers. Fórmulas para una relación eficaz. Retrieved from http://www.adecec.com/pdf/adecec_presenta_la_guia_“trabajar_con_influencers._formulas_para_una_relacion_eficaz”_.pdf
- Aguado Guadalupe, G., & García García, A. (2009). Del Word-of-mouth al marketing viral: aspectos claves de la comunicación a través de redes sociales. *Comunicación Y Hombre*, (5), 41–51. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=fua&AN=49126603&lang=es&site=eds-live>
- Aguilar-Gallegos, N., Martínez-González, E. G., Aguilar-Ávila, J., Santoyo-Cortés, H., Muñoz-Rodríguez, M., & García-Sánchez, E. I. (2016). Análisis de redes sociales para catalizar la innovación agrícola: de los vínculos directos a la integración y radialidad. *Estudios Gerenciales*, 32(140), 197–207. <http://doi.org/10.1016/j.estger.2016.06.006>
- Aguillo, I. F., Ortega, J. L., & Granadino, J. (2006). Contenidos del buscador Google . Distribución por países , dominios e idiomas. *El Profesional de La Información*, 15(5), 1–6.
- Alonso González, M. (2016). Comunicación empresarial de las empresas del IBEX

35 en Twitter. *Revista Mediterránea de Comunicación*, 7(1), 1–25.
<http://doi.org/10.14198/MEDCOM2016.7.1.8>

AMIPCI. (2014). *Estudio de Marketing Digital y Social Media 2014*. Asociación Mexicana de Internet AMIPCI. Retrieved from https://amipci.org.mx/images/Estudio_Marketing_2014.pdf

AMIPCI. (2016). *12º Estudio sobre los Hábitos de los Usuarios de Internet en México 2016*. Asociación Mexicana de Internet AMIPCI. Retrieved from https://www.amipci.org.mx/images/Estudio_Habitosdel_Usuario_2016.pdf

Amunátegui Perelló, C. (2006). Los metatags en el comercio electrónico un análisis de la jurisprudencia norteamericana y la doctrina española. *Revista Chilena de Derecho*, 33(2), 245–257.

Arbildi-Larreina, I. (2005). Posicionamiento en buscadores: una metodología práctica de optimización de sitios web. *El Profesional de La Información*, 14(2), 108–124. <http://doi.org/10.3145/epi.2005.mar.03>

Arenas, A., Fernández, A., & Gómez, S. (2008). Analysis of the structure of complex networks at different resolution levels. *New Journal of Physics*, 10, 21. <http://doi.org/10.1088/1367-2630/10/5/053039>

Bañuelos Flores, N., & Salido Ataiza, P. L. (2012). El mezcal en Sonora , México , más que una bebida espirituosa. *Estudios Sociales*, (2), 173–197.

Barberá, P., & Rivero, G. (2012). ¿Un tweet, un voto? Desigualdad en la discusión política en Twitter. In *I Congreso Internacional en Comunicación Política y Estrategias de Campaña* (pp. 1–22).

Barradas Gurruchaga, A. R., & Siano, M. (2014). Las redes sociales y su efecto de posicionamiento a nivel empresarial, caso México y Estados Unidos. *Razón Y Palabra*, (88), 17. Retrieved from http://www.razonypalabra.org.mx/N/N88/Varia/43_BarradasSiano_V88.pdf

- Beitia Vallés, R. (2012). Los social media y el factor humano. *adComunica: revista Científica de Estrategias, Tendencias E Innovación En Comunicación*, 3, 223–225. <http://doi.org/10.6035/2174-0992.2012.3.17>
- Blázquez-Ochando, M. (2016). Diseño de un agregador para la gestión de los big data informativos. *El Profesional de La Información*, 25(4), 671–683. <http://doi.org/10.3145/epi.2016.jul.17>
- Blondel, V. D., Jean-Loup, G., Renaud, L., & Etienne, L. (2008). Fast unfolding of communities in large networks, in Journal of Statistical Mechanics: Theory and Experiment. *Journal of Statistical Mechanics: Theory and Experiment*, 10, 1000.
- Carlos de la Paz, J., Gómez Masjuán, M. E., & Pérez Alonso, B. (2016). Estrategia de comunicación en redes sociales. Escenarios teórico prácticos para los medios de comunicación en Cuba. *RAZÓN Y PALABRA Primera Revista Electrónica En Iberoamérica Especializada En Comunicación*, 92(92), 1–40.
- Castelló Martínez, A., & del Pino Romero, C. (2015). La comunicación publicitaria con influencers. *Redmarka: Revista Académica de Marketing Aplicado*, 8(14), 21–50. Retrieved from <http://dialnet.unirioja.es/servlet/articulo?codigo=5159613&info=resumen&idoma=SPA>
- Colle, R. (2013). Prensa y “Big Data”: El desafío de la acumulación y análisis de datos. *Revista Mediterránea de Comunicación*, 4(1).
- Colmenero-Robles, A., Rosas-Medina, I., & Vázquez-Sánchez, E. (2011). Algunas consideraciones del mezcal en el municipio de Nombre de Dios, Durango, México. *Vidsupra, Instituto Politécnico Nacional*, 3(2), 48–52. Retrieved from <http://www.repositoriodigital.ipn.mx/handle/123456789/8171>
- Congosto, M. L. (2015). Elecciones europeas 2014: viralidad de los mensajes en Twitter. *Redes. Revista Hispana Para El Análisis de Redes Sociales*, 26(1),

23–52. <http://doi.org/10.5565/rev/redes.529>

Congosto, M. L., & Aragón, P. (2012). Twitter, del sondeo a la sonda: nuevos canales de opinión, nuevos métodos de análisis. *Más Poder Local. Especial Estudios de Caso*, 50–56. Retrieved from <http://dialnet.unirioja.es/descarga/articulo/4013889.pdf>
<http://dialnet.unirioja.es/servlet/extart?codigo=4013889>

Cook, S. (2012). Tutorial 4 - Centrality Measures. Retrieved from <http://www.fna.fi/blog/2012/11/05/tutorial-4-centrality-measures/>

Curioso, W. H., & Carnero, A. M. (2011). Promoviendo la investigación en salud con Twitter. *Revista Médica Herediana*, 22(3), 121–130.

Del Fresno García, M., Daly, A. J., & Segado Sánchez-Cabezudo, S. (2016). Identificando a los nuevos influyentes en tiempos de Internet: medios sociales y análisis de redes sociales. *Revista Española de Investigaciones Sociológicas*, 153, 23–42. <http://doi.org/10.54777/cis/reis.153.23>

Del Fresno García, M., Daly, A. J., & Supovitz, J. (2015). Desvelando climas de opinión por medio del social media mining y análisis de redes sociales en Twitter. El caso de los Common Core State Standards. *Redes. Revista Hispana Para El Análisis de Redes Sociales*, 26(1), 53–75. <http://doi.org/10.5565/rev/redes.531>

Díaz, F. J., Osorio, M. A., Amadeo, A. P., & Romero, D. L. (2013). Aplicando estrategias y tecnologías de Inteligencia de Negocio en sistemas de gestión académica. In *XV Workshop de Investigadores en Ciencias de la Computación* (pp. 225–229). Universidad de la plata ,Argentina.

Espinoza Ortega, A., & López Rosas, C. A. (2016). El consumidor de mezcal y su preferencia y aceptabilidad por el mezcal tradicional. Retrieved from <http://mezcologia.mx/los-consumidores-de-mezcal-conocedores/>

Fabrizio, S. (2014). Big Data :¿la ruta o el destino? *Advanced Series*, (3), 26.

- Fainholc, B. (2011). Un análisis contemporáneo del Twitter. *Revista de Educación a Distancia*, (26), 1–12.
- Fernández Burgueño, P. (2012). Aspectos jurídicos de la identidad digital y la reputación online. *adComunica: Revista Científica de Estrategias, Tendencias E Innovación En Comunicación*, 3(3), 125–142. <http://doi.org/10.6035/2174-0992.2012.3.8>
- Ferreas Rodríguez, E. M. (2011). La estrategia de la corporación EITB (Euskal Irratia Telebista) en Facebook y Twitter. *Razon Y Palabra*, 16(76). Retrieved from <http://www.redalyc.org/articulo.oa?id=199519981018%0ACómo>
- Ferrer-Sapena, A., & Sánchez-Perez, E. A. (2013). Open data, big data: ¿hacia dónde nos dirigimos? *Anuario ThinkEPI*, 7, 150–156.
- Gamir Ríos, J. (2016). Blogs, Facebook y Twitter en las Elecciones Generales de 2011. Estudio cuantitativo del uso de la web 2.0 por parte de los cabezas de lista del PP y del PSOE. *Revista DÍGITOS*, (2), 101–120.
- García-Marco, F.-J. (2011). La Pirámide de la Información Revisitada: Enriqueciendo el Modelo Desde la Ciencia Cognitiva. *El Profesional de La Informacion*, 20(1), 11–24. <http://doi.org/10.3145/epi.2011.ene.02>
- García-Ruiz, P. (2005). Comunidades de marca. El consumo como relación social. *Política Y Sociedad*, 42(1), 257–272.
- Garrido, F. J. (2001). El análisis de redes en el desarrollo local. In *Prácticas locales de creatividad social* (pp. 67–89).
- Gómez Martínez, R., & Prado Román, C. (2014). La actividad de búsquedas en Google anticipa los resultados electorales. *IDP. Revista de Internet, Derecho Y Política*, (18), 2–15. Retrieved from <http://www.redalyc.org/articulo.oa?id=78832841002>
- Goyzueta Rivera, S. I. (2011). Segmentación de la red social facebook: una

oportunidad para la empresa. *Perspectivas*, (28), 33–61.

Grandi Bustillos, C. V., González Ladrón de Guevara, F., & Segarra Oña, M. del V. (2014). ¿Cómo potenciar el uso de las redes sociales como estrategia de marketing en empresas del sector hostelero? El re-descubrimiento de “Pinterest.” *Tec Empresarial*, 8(3), 7–18.

Grau Moracho, J., & Guallar Delgado, J. (2004). El negocio de buscar en internet . Análisis del mercado de los buscadores en 2003. *El Profesional de La Informacion*, 13(4), 292–300.

Grávalos Macho, D. (2013). La calidad de una página web como herramienta de comunicación. *Estudios Sobre El Mensaje Periodístico*, 19, 253–261. http://doi.org/10.5209/rev_ESMP.2013.v19.42032

Hanneman, R. A. (2000a). *Introducción a los métodos del análisis de las redes sociales Capítulo I*. Retrieved from <http://wizard.ucr.edu/~rhannema/networks/text/textindex.html>

Hanneman, R. A. (2000b). *Introducción a los métodos del análisis de las redes sociales Capítulo VI*.

Hidalgo Reyes, M., Cruz Meza, P., & Perez Hernandez, L. M. (2010). Tecnificación del proceso de producción de mezcal en el estado de Oaxaca, México. In *XIV INTERNATIONAL CONGRESS ON PROJECT ENGINEERING* (pp. 2137–2149).

Huertas Roig, A., & Mariné Roig, E. (2014). Uso y utilidades de las herramientas de análisis online para la evaluación de la comunicación de las marcas de destino a través de los social media. *Sphera Publica*, (Numero Especial), 117–134. Retrieved from <http://repositori.udl.cat/handle/10459.1/48177>

IBM. (2013). Las cinco mejores formas de abordar big data. Nueva York. Retrieved from http://www-01.ibm.com/software/es/events/doc/pdfs/Las_cinco_mejores_formas_de_a

bordar_big_data.pdf

- Iglesias Rodríguez, C., & Herrero Gutiérrez, F. J. (2016). Los programas deportivos de la radio española en las redes sociales : estudio comparado del periodo 2010-2015. *Área Abierta*, 16(1), 39–59. http://doi.org/10.5209/rev_ARAB.2016.v16.n1.49853
- INEGI. (2015). Ciencia y Tecnología: Usuarios. Retrieved March 3, 2016, from <http://www3.inegi.org.mx/sistemas/temas/default.aspx?s=est&c=19007>
- Lecinski, J. (2013). *Ganando el momento cero de la verdad* (Vol. 1). Retrieved from <http://v1.zeromomentoftruth.com/google-zmot-es.pdf>
- Lévy Mangin, J. P., & Varela Mallou, J. (2003). *Análisis Multivariable para las ciencias Sociales*.
- Litche Fragoso, P. M., & Sánchez Salinas, J. C. (2014). *Uso productivo de Big Data y redes sociales en el sector turismo. Secretaría de Turismo Subsecretaría de Planeación y Política Turística Documentos de Investigación Estadística y Económica*.
- López, M., & Sicilia, M. (2016). Identificación de líderes de opinión leales en Twitter. *Cuadernos de Gestión*, 1, 1–20. <http://doi.org/10.5295/cdg.140508ml>
- López García, D. (2012). *Análisis de las posibilidades de uso de Big Data en las organizaciones*.
- Luisa, M., & Hernández, D. (2003). Estrategias de mercadotecnia y los negocios de mezcal 1. *Reflexión Política*, (31), 187–203.
- Marín, F., Duque, M., & Manruffo, R. (2014). Redes sociales como plataforma de posicionamiento de un programa radial. *Universidad Privada Dr. Rafael Belloso Chacín*, 12(1), 194–206.
- Marín Rives, L., & Ruiz de Maya, S. (2007). La identificación del consumidor con la

empresa : más allá del marketing de relaciones. *Universia Business Review*, 1, 62–75.

Martínez Ortega, R. M. (2009). El coeficiente de correlación de los rangos de Spearman caracterización. *Rev. Haban Cienc Méd La Habana*, 8(2), 1–19. Retrieved from <http://scielo.sld.cu/pdf/rhcm/v8n2/rhcm17209.pdf>

Maté Jiménez, C. (2014). Big Data. Un nuevo paradigma de análisis de datos. *Anales de Mecánica Y Electricidad*, 10–16.

Mir Bernal, P., Fondevila Gascón, J. F., & Gutiérrez Aragon, Ó. (2016). El uso de las brand fan pages de Facebook entre jóvenes y profesionales : análisis empírico. *Enlace, Revista Venezolana de Información, Tecnología Y Comocimiento*, 13(1), 28–44.

Moráguez Bergues, M., & Perurena Cancio, L. (2014). Propuesta de factores a considerar en el posicionamiento de los sitios web de salud. *GECONTEC: Revista Internacional de Gestión Del Conocimiento Y La Tecnología*, 2(1), 10–30. Retrieved from <https://rio.upo.es:443/xmlui/handle/10433/1095>

Morato, J., Sánchez, S., & Valiente, M. C. (2005). Análisis de estrategias de posicionamiento en relación con la relevancia documental. *El Profesional de La Información*, 14(1), 21–29. <http://doi.org/10.3145/epi.2005.feb.03>

Muff, S., Rao, F., & Caflisch, A. (2005). Local modularity measure for network clusterizations. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 72(5), 1–5. <http://doi.org/10.1103/PhysRevE.72.056107>

Muñoz Rodríguez, M., Aguilar Ávila, J., Rendón Medel, R., & Altamirano Cárdenas, J. R. (2007). *Análisis de la dinámica de innovación en cadenas agroalimentarias. CIESTAAM - Universidad Autónoma Chapingo*. <http://doi.org/10.1017/CBO9781107415324.004>

Ortiz Morales, M. D., Joyanes Aguilar, L., & Giraldo Marín, L. M. (2016). Los desafíos del marketing en la era del big data. *E-Ciencias de La Información*,

6(1), 30. Retrieved from <http://www.gartner.com/it-glossary/big-data/>
https://books.google.com/books?id=p_glBgAAQBAJ&pgis=1%5Cnhttp://www-05.ibm.com/services/es/bcs/pdf/Big_Data_ES.PDF%5Cnhttp://ezproxy.si.unav.es:2090/eds/detail/detail?sid=c6a9a4d2-3622-4ca0-bd6c-b78efdc16f5f

Oviedo García, M. de los Á., Muñoz Expósito, M., & Castellanos-Verdugo, M. (2015). La expansión de las redes sociales. Un reto para la gestión de marketing. *Contabilidad Y Negocios: Revista Del Departamento Académico de Ciencias Administrativas*, 10(20), 59–69. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/5399422.pdf%5Cnhttps://dialnet.unirioja.es/servlet/extart?codigo=5399422>

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The PageRank Citation Ranking: Bringing Order to the Web. *World Wide Web Internet And Web Information Systems*, 54(1999–66), 1–17. <http://doi.org/10.1.1.31.1768>

Paladines Galarza, F. Y., & Granda Tancazo, C. V. (2013). La Marca Online Como Parte De La Comunicación Integral . Manejo De Marcas Ecuatorianas En Las. *REDMARKA UIMA-Universidad de A Coruña - CIECID*, 6(6), 95–116.

Palazón, M., Silicia, M., & Delgado, E. (2014). El papel de las redes sociales como generadoras de “ amor a la marca .” *Universia, BusinessReview, Primer Tri*, 23.

Parra, P., Gordo, A. J., & D’Antonio, S. A. (2014). La investigación social aplicada en redes sociales. Una innovación metodológica para el análisis de los «Me gusta» en Facebook Social. *Revista Latina de Comunicacion Social*, (69), 195–212. <http://doi.org/10.4185/RLCS-2014-1008>

Pérez Salazar, G., & Aguilar Edwards, A. (2012). Reflexiones Conceptuales en torno a las Redes Sociales en Las Redes Sociales: Un Recorrido De La teoría a las prácticas comunicativas en Facebook, Twitter Y Google+. *Razón Y Palabra*, 17(79). Retrieved from

<http://www.redalyc.org/articulo.oa?id=199524411018%0ACómo>

Ríos, S. (2014). Big Data : Información de calidad para mejores decisiones. *Revista ISCI*, 14(4), 17–19.

Rodriguez Mareño, N. A. (2014). Proceso publicitario de le empresa PIL Andina en la red social Facebook. *Con-Sciencias Sociales*, 6(11), 31–34.

Rodríguez Rangel, C., & Sánchez Rivero, M. (2015). Presencia de alojamientos extremeños en redes sociales. *Gran Tour: Revista de Investigaciones Turísticas*, (12), 4–22.

Romero Frías, E., Vaughan, L., & Rodríguez Ariza, L. (2011). El empleo de la Webmetría para el análisis de los indicadores de desempeño y posición financiera de la empresa: un análisis exploratorio en diversos sectores económicos de los Estados Unidos. *Revista Innovar Journal*, 25(56), 99–111. <http://doi.org/10.15446/innovar.v25n56.48993>.http

Sánchez, E. (2012). El social media en la estrategia de comunicación. *adComunica: Revista Científica de Estrategias, Tendencias E Innovación En Comunicación*, (3), 221–222. <http://doi.org/10.6035/2174-0992.2012.3.16>

Sánchez Casado, N., & Mostaza-Ucedo, M. F. (2014). Uso de las redes sociales como herramienta de marketing : una aplicación a las entidades de crédito. *Anuario de Jóvenes Investigadores*, 7, 93–95.

Serrano-Puche, J. (2012). Herramientas web para la medición de la influencia digital: análisis de klout y PeerIndex. *El Profesional de La Informacion*, 21(3), 298–303. <http://doi.org/10.3145/epi.2012.may.11>

Sixto García, J. (2015). Desarrollo de las redes sociales como herramienta de marketing. Estado de la cuestión hasta 2015. *Anagramas: Rumbos Y Sentidos de La Comunicación*, 14(26), 179–196. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/5191802.pdf%5Cnhttps://dialnet.unirioja.es/servlet/extart?codigo=5191802>

- Soltero Domingo, F. J., & Bodas Sagi, D. J. (2005). Clasificadores inductivos para el posicionamiento web. *El Profesional de La Información*, 14(1), 4–13.
- Tabares, L. F., & Hernández, J. F. (2014). Big Data Analytics : Oportunidades , Retos y Tendencias. *Universidad de San Buenaventura*, 20.
- Tello-Díaz, L. (2013). Intimidad y «extimidad» en las redes sociales. Las demarcaciones éticas de Facebook. *Comunicar*, 21(41), 205–213. <http://doi.org/10.3916/C41-2013-20>
- Túñez-López, M., Sixto García, J., & Guevara-Castillo, M. (2011). Redes sociales y marketing viral: repercusión e incidencia en la construcción de la agenda mediática. *Palabra Clave*, 14(1), 53–65. <http://doi.org/10.5294/1872>
- Vállez, M. (2011). Keyword Research : métodos y herramientas para identificar palabras clave. *BID Textos Universitaris de Biblioteconomia I Documentació*, (27), 16.
- Viñarás Abad, M., & Cabezuelo Lorenzo, F. (2012). Claves para la participación y generación de contenido en las redes sociales: estudio de caso del Museo del Prado en Facebook. *adComunica: Revista Científica de Estrategias, Tendencias E Innovación En Comunicación*, 3, 87–103. <http://doi.org/10.6035/2174-0992.2012.3.6>

8. APÉNDICES

Apéndice 1. Ejemplos de publicaciones dentro las categorías destacada

Imágenes de Botella



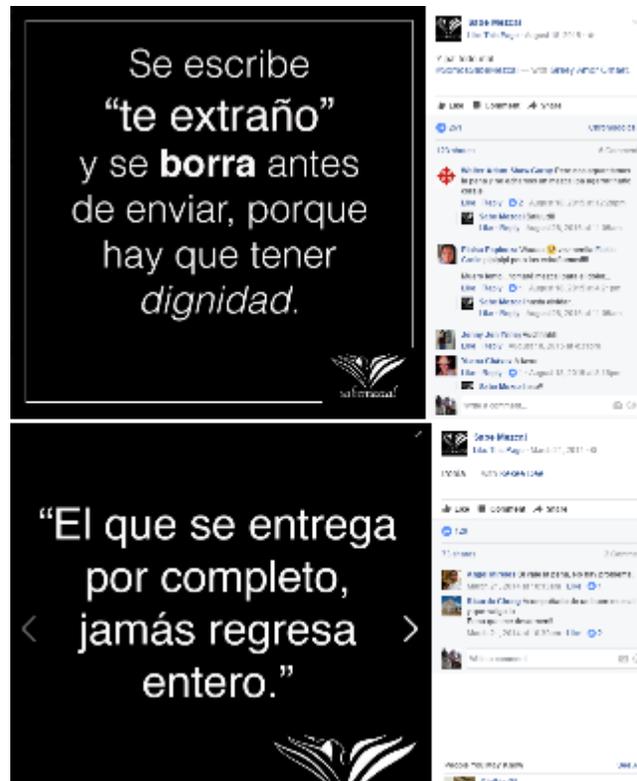
Imágenes de Stock de botellas



Imágenes de Agaves



Frases cómicas



Notas Informativas



Días festivos (día del padre y la mujer)



Apéndice 2. Lista de clientes potenciales con los cuales se puede conectar

Id Usuarios para conectar en Facebook	Número de empresas con las que interactuó	Código de usuario
ef1b981a9d6d0b2d40b7277ca9fe3d2b62617bb9	21	user_841387419269944
f50b91684f4ba562ea0d346b2f6aff70cde19c37	20	user_10154187427082524
22ee44b5be614d77e23c7b632d8c926b31de340c	19	user_10206526865425820
e3a9a7c58cbc73c687ebfcfa5b301aeb3da3b82b	15	user_1046832115367953
d323cec7f9385c7a6f3a597b16fba1ab13e706c2	14	user_1622042571398994
796d7101a5d1cbdac1bd25589bcd373e2cc2ecfc	13	user_10204745664713244
211ae33eea49eb263439007731293e2092a79445	12	user_869927319723541
48a77abc2441b49011b80064cc571d31608bec3a	11	user_717285311717070
81f4056db32c180c10014957bb898b699516bf1	11	user_1690708127823939
3fd997d185fb7c97e07f752808467cb5d4a11dd	10	user_10153594504095042

Id Usuarios para conectar en Facebook	Número de empresas con las que interactuó	Código de usuario
7c3d7109315a18ae1c42deac4d0509840cb7ad40	10	user_10203202472342882
c4dc528d77e24882f1db0209edecee2b13b4a7c3	10	user_887012898043974
df9aeee395a48ce40d93ebe260bc96836a0e21d3	10	user_451414838267592
f0fda5d0b87fc8eeb39df7cbfc771491cd638e25	10	user_1045079968843441
17d1cb3eae54fc3c11e6a6944d7ab1fba0f6d715	9	user_10208094439381092
1961f24781ddc15847bc52f742f3ac0733cdf594	9	user_10206958436028836
304c7e2036009a6aae8075dcc4ada747057f3f13	9	user_10153357048822641
3537a6d5c8922a8c6882db2a13c46cd8f323c5d0	9	user_10153407050843413
5afb5dd49afed281a536c36e5e87bdcbe4842a989	9	user_716563771709336
5f15daee1f2660cbafea0e0c8fc5351c3aaad658	9	user_1155851227775225
6ba970134dfe90aa065e6bc93e8654ff0b111e0f	9	user_10205630677140718
71ef365edf3f936cc6f73f1293911d315aec0a22	9	user_396244847252103
73b99e652c7d7fb33c58596536bf90bb00a8692f	9	user_10207089029732467
7b02935d4674c8e1497207a292fe737204b55dba	9	user_887683197947684
8b4d95ce02281848b3298b1c48e4180987f22b49	9	user_950931781616531
8eee0ecc4835552bbf01d607d260b3d56c3bd993	9	user_10156095049060580
c064f47fc3e41748e90b06ec4d6132fe60e16cea	9	user_10203880102762231
fe9b543294a7198ba79accede77cd93cb95ea25a9	9	user_10155620683190074
03161ed5f210c37ce38d83655db17794a7496f92	8	user_10153460177109038
570468399a673e1e48725160f4ee5a1953b6b36b	8	user_10153190184003137
6e3058ee027af2c259966ce07c49e5382d5d075f	8	user_10202906070465517
75cb2ca69b0a1e660a89426cd0e71638a370d6c9	8	user_1431962403709009
816cf8f75fe6e6ef5ffe7e32b2c58ba4bdc1b883	8	user_1091438500869039
93378e1a8e514fafa33b640579861972e62c2ac4	8	user_10153417195507594
bbfd6123c8891225d0321fea6fa50c707168d97e	8	user_10153549843077505
bc10ad56d3ef03ac70ba13ea658d629f5f8025cb	8	user_10152744754111851
d9495e710b77d4d080e8c90eadfe631583367fd4	8	user_10201117578177124
f621214533bd1fa2d8baf4e503c930873fcf9d7a	8	user_10153932470039816
fdbe347e21512b3cee6bb139169ba848ecf80af9	8	user_10204890025918037
02b874a7f10fbab069da0269157c666aa9ff8bf3	7	user_10154037432401754

Id Usuarios para conectar en Facebook	Número de empresas con las que interactuó	Código de usuario
0937cade0203e00a03a99b800d2c10a65b010985	7	user_10207840403030629
11910e767f315e3dd6cf1cb4520a14f587e73fb5	7	user_884697611561050
1274f42f62e6afb5d1280722a683cf833288cf54	7	user_916268591789703
15d2011edb849a15b94eb3f0b89a441530b231e8	7	user_10205372463291224
1691209479f7dc085053efe7d93a19caca74d204	7	user_761202447366
2c19381f15455f20a70cffb1b0ad500c96142dc1	7	user_10153177562265218
2df8da23e9293f6cc59ea6196000e58479c36e3f	7	user_10205888910116528
46db7eca1f5cab66a20994d5b4ce87eb08be455e	7	user_1247683658592151
567d5cc14a3d937272f36860618673bfeda26764	7	user_10207438447936651
59891ea324b6a7eb48f7b4ceb6c6266624568a8e	7	user_10153309073046306
67107ac19a4d56c7c6fd989d04aff213b8cad3e3	7	user_10153513747803629
6936c55dad57c969dffa73c50bae3ac5d114929	7	user_481603672046637
6aaf1f8bb83e00cde027a5480a3addb241ce6522	7	user_10152887698721612
7b488c307380225ffb34a5b63ad383d7bfa55de0	7	user_995959543794945
7cd8f4fe59b5d651faefff3f4839a79107adae0c	7	user_10153836280237847
7f50b12caf448d7ddbce965d72d92888c7b13bf5	7	user_10206768485995258
9992a262380089bd45610ebb9d17b2f982315171	7	user_998566670188217
a210588b57735253109b80b8a58d832f0db2f103	7	user_10153346988989176
af2bf330674315fc29a5fb737f874b4474e59fa5	7	user_10153528511117152
b3b04feb86d71bc55907722bf3f930189e61dbd2	7	user_1508283736162133
b4247ed8b38f5d3e319f80767b1b6ddcc6e813a6	7	user_283920615076415
b4da15ceabf67a32ac618a34dab32b0eafbbed981	7	user_10155507195555594
b9d3aff03898ce07b6bd19980ec629ee29950a20	7	user_10153765707165739
c6c08ca8dc8d625b4f8c4436003d0ad6f6c00aeb	7	user_10155529093955029
d4307fd3315eedffd9ab962fbbe4b9f7762208bf	7	user_1116401468390039
e089bb037c02cfda63d150f5ed2eae06b4bd6028	7	user_712579018873001
f25ade93ca04c406ca6c940729c544d2015e9fac	7	user_10152326194855428
f400426a2b42b48bc8ee4f59374a050c5e228baf	7	user_136692534355
fa83f6b23ce451825a8e49929413bd24a4ed4337	7	user_10152859364297680

Apéndice 3. Características de las cuentas con mayor número de tweets escritos con mezcal

Usuario	Tweets con “mezcal”	Total de tweets	Siguiendo	Seguidores	% Tweets	Clasificación
TeamIsaacTeq	459	24,766	1,875	1,256	1.9	Habla de la cultura mexicana
ZacMezcal	329	7,990	3,576	3,575	4.1	Marca de mezcal
PatriotIndiana	293	131,703	4,738	2,002	0.2	Habla de la cultura del mezcal
silencio_mezcal	226	1,438	95	19	15.7	Habla de la cultura del mezcal
NacionMezcal	158	943	113	13,394	16.8	Cuenta que promueve la cultural del mezcal
agavefilia	141	2,915	819	454	4.8	Cuenta que promueve la cultural del mezcal
marioeherrera	116	88,458	6,802	15,429	0.1	Habla de la cultura Oaxaqueña
MezcalesOaxaca	114	1,099	235	1,445	10.4	Cuenta que promueve la cultural del mezcal
WahakaMezcal	111	11,516	5,023	7,123	1.0	Marca de mezcal
MezcalFresno	106	2,527	2	404	4.2	Marca de mezcal
DyS_Mezcal	101	2,805	755	1,957	3.6	Marca de mezcal
mezcalesconalma	98	841	891	653	11.7	Marca de mezcal
GonzoGourmands	94	89,434	1,131	1,571	0.1	Marca de mezcal
mimezcal_	94	2,106	1,523	821	4.5	Otro: restaurantes, lugares, gastronomía
LaMezcalCentro	93	8,347	667	1,508	1.1	Mezcalería
BladimirMJuarez	93	2,128	1,170	833	4.4	Habla de la cultura del mezcal
LaChoperia59	86	21,859	1,276	5,642	0.4	Otro: restaurantes, lugares, gastronomía

Usuario	Tweets con "mezcal"	Total de tweets	Siguiendo	Seguidores	% Tweets	Clasificación
Enmezcalado	85	3,382	3,593	14,108	2.5	Habla de la cultura del mezcal
MezcalSilencio	79	6,540	0	26,097	1.2	Marca de mezcal
MezcalEsLaMeditid	78	1,186	555	354	6.6	Marca de mezcal
Mezcologia	77	5,078	1,058	3,979	1.5	Cuenta que promueve la cultural del mezcal
Mezcal8Viboras	69	5,228	845	3,861	1.3	Marca de mezcal
VancouverGuero	68	18,875	1,536	2,191	0.4	Habla de bebidas de agave
mezcalmusical	67	897	1,994	532	7.5	Otro: restaurantes, lugares, gastronomía
NaaLiuu	66	866	3,101	1,434	7.6	Marca de mezcal
Ms_Terree	65	195,782	12,496	22,586	0.0	Otro: restaurantes, lugares, gastronomía
AI_3_Mezcal	65	616	643	649	10.6	Marca de mezcal
LeyendaEterna	65	318	360	417	20.4	Marca de mezcal
timberelk	64	55,234	5,325	5,537	0.1	Habla de bebidas de agave
MezcalZacbe	62	2,144	1,055	1,752	2.9	Marca de mezcal
CortijoLaMezca	61	869	180	557	7.0	Mezcalería
lalanoguera	58	1,483	937	906	3.9	Habla de bebidas de agave
Mezcal_3P_Km70	57	3,265	1,965	1,552	1.7	Marca de mezcal
julunamx	56	45,421	717	1,666	0.1	Mezcalería
Insitumezcal	56	9,558	2,311	2,832	0.6	Habla de la cultura mexicana
MezcalJoyaAzul	55	275	616	267	20.0	Marca de mezcal
playadelcarmenx	54	156,638	172	659	0.0	Otro: restaurantes, lugares, gastronomía

Usuario	Tweets con "mezcal"	Total de tweets	Siguiendo	Seguidores	% Tweets	Clasificación
agavache	54	819	407	365	6.6	Cuenta que promueve la cultural del mezcal
WildAgave	53	1,657	339	372	3.2	Marca de mezcal
MezcalCampero	51	1,250	636	2,736	4.1	Marca de mezcal
Perla_Sanjuan	49	24,682	1,446	1,249	0.2	Habla de la cultura Oaxaqueña
mezcal	49	9,729	896	533	0.5	Habla de la cultura del mezcal
MezcalElCortijo	49	1,177	657	3,151	4.2	Marca de mezcal

Fuente: Twitter

Las 43 cuentas que escribieron más de 50 tweets con la palabra mezcal se clasificaron en ocho tipos. El primer tipo es el perteneciente a las marcas de mezcal el cual abarca 18 cuentas, este grupo como su nombre lo indica son cuentas de empresas mezcaleras y por lo tanto se pueden descartar como posibles partidarios para promocionar un mensaje publicitario de otra marca. En segundo lugar, se encuentran 5 cuentas cuya finalidad es exclusivamente la promoción de la cultura del mezcal y si bien escriben tweets sobre eventos y notas informativas, también comparten mensajes de marcas mezcaleras que consideran importantes. También se encontraron cuatro tipos de cuentas que tocaban temas como cultura del mezcal, bebidas de agave, cultura mexicana o cultura oaxaqueña, con una frecuencia de 5, 4, 2 y 2 respectivamente, a este tipo de cuentas se les denominó con el prefijo «Habla de» ya que, si bien tocan estos temas, su contenido no versa única y exclusivamente sobre esa materia. Dentro de los usuarios se encontraron 3 cuentas de mezcalerías y 5 que se clasificaron como «otros», dentro de estos últimos se encontraron que tenían que ver con restaurantes, bares, cuentas de chefs, y lugares que a pesar que su temática principal no es el mezcal, sí lo usan en diferentes tweets.

Apéndice 4. Primeras 100 posiciones de hashtag con sus grados de entrada

Posición	Hashtag	Grados de entrada
1	#oaxaca	3,518
2	#mexico	1,247
3	#denver	947
4	#tequila	734
5	#agave	651
6	#cultura	296
7	#cocktails	280
8	#mezcalesdeoaxaca	256
9	#twitteroax	237
10	#sellamamezcal	224
11	#guerrero	215
12	#master	213
13	#cdmx	211
14	#cheers	189
15	#distillers	185
16	#guelaguetza2016	185
17	#laverdaddeoaxaca	174
18	#mezcaleria	165
19	#mezcales	165
20	#cocktail	162
21	#cerveza	159
22	#nashville	143
23	#maguey	135
24	#durango	134
25	#elsilencio	129
26	#editorial	126
27	#repost	125
28	#puebla	119
29	#diainternacionaldelbeso	117
30	#drink	116
31	#michoacan	116
32	#ixtapazihuatanejo	114
33	#tuvoz	110
34	#reycampero	107

Posición	Hashtag	Grados de entrada
35	#chilpancingo	101
36	#muchogustomexico	101
37	#playadelcarmen	101
38	#sanmigueldeallende	100
39	#beer	99
40	#drinks	97
41	#yosoymezcalesdeoaxaca	97
42	#mezcalito	95
43	#rutacaminosdelmezcal	95
44	#espadin	91
45	#komil	91
46	#mezcalleyendaeterna	89
47	#capacitacion	87
48	#empreendedor	87
49	#sacapalabras	87
50	#viveloparacreerlo	87
51	#evaluacion	86
52	#quieromialipus	86
53	#salud	86
54	#oaxac	84
55	#gastronomia	82
56	#mezcalzine	82
57	#nationalmargaritaday	82
58	#instagram	80
59	#bar	79
60	#guanajuatosisabe	79
61	#packaging	79
62	#acapulco	78
63	#pulque	78
64	#sma	77
65	#duoimperial	75
66	#tacos	73
67	#nowplaying	71
68	#rivieramaya	70
69	#travel	70
70	#amor	68

Posición	Hashtag	Grados de entrada
71	#delmaguey	68
72	#viernes	68
73	#sabiasque	67
74	#buenviernes	66
75	#chela	66
76	#enterate	66
77	#locales	66
78	#simequieresvermuyfeliz	66
79	#atreveteoaxaca	65
80	#diadelingeniero	65
81	#mezcalamores	65
82	#mezcaleros	65
83	#food	62
84	#friends	62
85	#quality	62
86	#video	62
87	#arte	61
88	#aestahoramefalta	60
89	#turismo	60
90	#golazo	59
91	#art	57
92	#puntomexico	57
93	#felizlunes	56
94	#singlevillage	56
95	#organic	54
96	#periscope	54
97	#sotol	54
98	#artesanal	53
99	#cantina	53

Apéndice 5. Primeras 100 posiciones de mensajes compartidos retweets

Posición	Etiqueta	Cuenta que lo emitió	Grado de entrada
1	3962-601.lopezdoriga	lopezdoriga	601
2	6553-337.RoyCampos	RoyCampos	337

Posición	Etiqueta	Cuenta que lo emitió	Grado de entrada
3	605-226.asvpxdaniel_	asvpxdaniel_	225
4	4124-171.MaguMonero	MaguMonero	170
5	7440-149.twiivette	twiivette	149
6	2733-114.GobOax	GobOax	114
7	778-109.BenjaminRoblesM	BenjaminRoblesM	109
8	7026-109.STyDE_GobOax	STyDE_GobOax	109
9	2738-99.GobOax	GobOax	99
10	2728-96.GobOax	GobOax	96
11	2748-86.GobOax	GobOax	86
12	6912-83.SonBebidas	SonBebidas	83
13	2893-83.Hell___Huesos	Hell___Huesos	82
14	7029-81.STyDE_GobOax	STyDE_GobOax	81
15	1833-78.DonPorfirioDiaz	DonPorfirioDiaz	78
16	3646-73.lajornadaonline	lajornadaonline	73
17	777-71.BenjaminRoblesM	BenjaminRoblesM	70
18	5471-70.MrScottEddy	MrScottEddy	70
19	5871-65.Oaxaca_Politico	Oaxaca_Politico	65
20	1834-64.DonPorfirioDiaz	DonPorfirioDiaz	64
21	1742-62.Dios_Padre	Dios_Padre	62
22	6205-61.PoliPocket	PoliPocket	61
23	5591-59.NacionMezcal	NacionMezcal	59
24	7214-59.TheSpanishJams	TheSpanishJams	59
25	7473-59.UnPoetaEs	UnPoetaEs	59
26	2886-57.HectorAstudillo	HectorAstudillo	57
27	6260-56.ppcalzada	ppcalzada	56
28	6918-54.sopitas	sopitas	54
29	6596-53.S0NCELOSAS	S0NCELOSAS	53
30	1901-52.DunjaNT	DunjaNT	52
31	7744-50.Xabiani	Xabiani	50
32	639-48.avenaamador	avenaamador	48
33	5346-48.mikeliturriaga	mikeliturriaga	48
34	7787-47.yomuevoamexico	yomuevoamexico	47
35	6909-45.SonBebidas	SonBebidas	45
36	3270-44.jimenezmerino	jimenezmerino	44
37	3001-43.iCamelopardalys	iCamelopardalys	43
38	6597-43.S0NCELOSAS	S0NCELOSAS	43

Posición	Etiqueta	Cuenta que lo emitió	Grado de entrada
39	3643-42.LaJornada	LaJornada	42
40	7394-42.Trapitta	Trapitta	42
41	2142-41.EIMalditoHuesos	EIMalditoHuesos	41
42	6640-40.SAGARPA_mx	SAGARPA_mx	40
43	7855-40.zero_twitts	zero_twitts	40
44	6842-39.Silvano_A	Silvano_A	39
45	521-38.animalgourmet	animalgourmet	38
46	5263-38.Mezcologia	Mezcologia	37
47	608-49.AtreveteOaxaca	AtreveteOaxaca	36
48	964-35.business	business	35
49	1010-35.CanalOnceTV	CanalOnceTV	35
50	4225-35.marianabenitez	marianabenitez	35
51	5627-35.NacionMezcal	NacionMezcal	35
52	179-34.Aeromexico	Aeromexico	34
53	965-33.business	business	33
54	5687-34.NacionMezcal	NacionMezcal	33
55	3167-32.j_medina37	j_medina37	32
56	7621-32.VivaAerobus	VivaAerobus	32
57	3166-31.j_medina37	j_medina37	31
58	3726-31.LaPavlovich	LaPavlovich	31
59	5632-30.NacionMezcal	NacionMezcal	30
60	6920-30.sopitas	sopitas	30
61	7607-30.VisitGuerreroMx	VisitGuerreroMx	30
62	2192-32.emeequis	emeequis	29
63	6778-29.SergioOchoaV	SergioOchoaV	29
64	6056-28.paulaarangoh	paulaarangoh	28
65	6742-28.SECTUR_mx	SECTUR_mx	28
66	237-27.AIGAdesign	AIGAdesign	27
67	2913-27.hiletrado	hiletrado	27
68	4501-27.mexdesconocido	mexdesconocido	27
69	6911-27.SonBebidas	SonBebidas	27
70	1743-26.DIOS_SARCASMO	DIOS_SARCASMO	26
71	6843-26.Silvano_A	Silvano_A	26
72	518-25.animalgourmet	animalgourmet	25
73	2661-25.Gastrobites	Gastrobites	25
74	2750-25.GobOax	GobOax	25

Posición	Etiqueta	Cuenta que lo emitió	Grado de entrada
75	3647-25.lajornadaonline	lajornadaonline	25
76	6163-25.Piniaexpress	Piniaexpress	25
77	6492-25.rix	rix	25
78	6623-25.sacapalabras	sacapalabras	25
79	609-25.AtreveteOaxaca	AtreveteOaxaca	24
80	2950-25.Huesud0	Huesud0	24
81	5611-24.NacionMezcal	NacionMezcal	24
82	5636-24.NacionMezcal	NacionMezcal	24
83	6739-24.Sectur_Guerrero	Sectur_Guerrero	24
84	6919-24.sopitas	sopitas	24
85	4499-23.mexdesconocido	mexdesconocido	23
86	5608-23.NacionMezcal	NacionMezcal	23
87	5619-23.NacionMezcal	NacionMezcal	23
88	2753-22.GodinezOficina	GodinezOficina	22
89	4500-22.mexdesconocido	mexdesconocido	22
90	6324-22.qhacermx	qhacermx	22
91	7445-21.tyler_durdend	tyler_durdend	21
92	1961-20.Eater	Eater	20
93	2145-20.ElMonitou	ElMonitou	20
94	2582-20.foodandwine	foodandwine	20
95	3304-20.JornalOGlobo	JornalOGlobo	20
96	3886-20.LeyendaEterna	LeyendaEterna	20
97	5665-20.NacionMezcal	NacionMezcal	20
98	6162-20.Piniaexpress	Piniaexpress	20
99	2315-20.Pnamen	Pnamen	20
100	1212-19.Socialmex	Socialmex	20

Apéndice 6. Primeras 100 posiciones en la red de menciones con respecto al grado de entrada

Posición	label	indeg ee	outdeg ee	degr ee	modularity_cl ass	Posición de su comunidad
1	lopezdoriga	342	0	342	2989	7
2	STyDE_GobOax	216	8	224	1511	6

Posición	label	indegreee	outdegreee	degreee	modularity_cl ass	Posición de su comunidad
3	YouTube	213	0	213	1872	10
4	GobOax	188	13	201	1511	6
5	Sabemezcal	152	0	152	1875	5
6	GabinoCue	146	0	146	1511	6
7	Change_Mex	139	1	140	1665	17
8	NacionMezcal	108	29	137	2775	2
9	LeyendaEterna	71	7	78	2723	8
10	SaintAnejo	71	0	71	2771	1
11	SAGARPA_mx	67	1	68	1171	19
12	SECTUR_mx	63	3	66	3040	21
13	PorQueMezcal	62	3	65	3006	3
14	UMSNHOficial	56	0	56	1189	28
15	ferolivera09	53	6	59	1670	20
16	marioeherrera	51	39	90	2723	8
17	RodrigoUrbano	51	1	52	2980	31
18	VisitMex	49	0	49	889	16
19	mezcalamores	48	6	54	889	16
20	abcradiopuebla 1	46	0	46	1171	19
21	MezcalesOaxaca	45	23	68	1511	6
22	GtoSiSabe	44	13	57	1670	20
23	Oaxtravel	42	1	43	1511	6
24	munchies	40	2	42	1706	26
25	SECTURGTO	40	9	49	1670	20
26	EspitaDC	39	4	43	2771	1
27	lalanoguera	38	59	97	3006	3
28	DelMagueyMezcal	37	12	49	797	9
29	PrrosPerdidos MX	37	4	41	1534	29
30	SE_mx	37	0	37	1927	11
31	MezcalSilencio	36	14	50	1875	5
32	400conejosMX	35	4	39	889	16
33	guanajuato	35	8	43	1670	20

Posición	label	indegreee	outdegreee	degreee	modularity_cl ass	Posición de su comunidad
34	llegalMezcal	34	3	37	1875	5
35	FoodandTravel MX	33	4	37	3006	3
36	ppcalzada	33	0	33	1171	19
37	SoyMezcalero	33	5	38	2775	2
38	VisitZ	33	19	52	3006	3
39	FENADUoficial	32	4	36	2806	15
40	Guillermogenge l	32	6	38	1670	20
41	WahakaMezcal	32	34	66	2771	1
42	GaelGarciaB	31	0	31	1525	34
43	totc	31	3	34	2771	1
44	turismo_sma	31	3	34	1670	20
45	escollo_mx	30	2	32	1927	11
46	SieteMisterios	30	6	36	2771	1
47	GarciaPosti	29	2	31	2071	32
48	Karime_AcaSh ore	29	0	29	2379	40
49	liladowns	29	0	29	3006	3
50	ThinkTequila	29	0	29	1178	14
51	CheechMarin	28	1	29	1706	26
52	GobMichoacan	28	2	30	1189	28
53	MezcalUnion	28	1	29	889	16
54	mutemgaribaldi	28	0	28	2873	13
55	mezcal	27	9	36	3006	3
56	Montelobos	27	7	34	2771	1
57	oyameldc	27	0	27	1898	35
58	realminero	27	1	28	2815	12
59	RegeneracionM x	27	1	28	2862	4
60	munchiesES	26	4	30	1706	26
61	PierdeAlmasM X	26	5	31	797	9
62	Rebmangas	26	3	29	2815	12
63	AmathusDrinks	25	28	53	797	9

Posición	label	indegr ee	outdegr ee	degr ee	modularity_cl ass	Posición de su comunidad
64	kochelMezcal	25	2	27	2815	12
65	CDI_mx	24	0	24	1738	36
66	Expendio_Oax	24	3	27	2775	2
67	mezcalesconal ma	24	10	34	2862	4
68	AtreveteOaxac a	23	0	23	3040	21
69	CdSatelite	23	1	24	2723	8
70	ciewtcmexico	23	0	23	3006	3
71	COFEMER	23	0	23	1927	11
72	Liquor	23	3	26	1178	14
73	secturslp	23	2	25	3040	21
74	ChapulínRest	22	2	24	889	16
75	Insitumezcal	22	9	31	2873	13
76	LosAmantesMz c	22	3	25	2815	12
77	mezcal400voce s	22	2	24	1927	11
78	OaxacaSabe	22	2	24	2815	12
79	ZignumMezcal USA	22	1	23	1875	5
80	lagusanaciega	21	1	22	719	27
81	MezcalAlipus	21	4	25	797	9
82	punch_drink	21	1	22	1178	14
83	sabormexico	21	18	39	3006	3
84	TresPapalote	21	2	23	1706	26
85	AgroBaja	20	0	20	1171	19
86	c_insurgentes	20	21	41	2775	2
87	Dssdufoo	20	19	39	3006	3
88	Lakatrinamx	20	23	43	3006	3
89	Txalaparta_Bar	20	0	20	2873	13
90	alipuscondesa	19	2	21	797	9
91	bullterrier105	19	3	22	719	27
92	HectorAstudillo	19	0	19	1738	36
93	hllerena	19	0	19	1178	14

Posición	label	indegree	outdegree	deegree	modularity_class	Posición de su comunidad
94	Indocumentado	19	16	35	1927	11
95	j_medina37	19	3	22	2973	30
96	mezcal_oaxaca	19	9	28	2775	2
97	MezcalCampero	19	3	22	2862	4
98	Mezcalillera_	19	6	25	2862	4
99	pedrocastelan	19	20	39	3006	3
100	sidraSN	19	17	36	2775	2

Apéndice 7. Lista de “influencers” en la red de menciones

label	modularity_class	indegree	outdegree	deegree
NacionMezcal	2775	108	29	137
SaintAnejo	2771	71	0	71
PorQueMezcal	3006	62	3	65
EspitaDC	2771	39	4	43
lalanoguera	3006	38	59	97
FoodandTravelMX	3006	33	4	37
SoyMezcalero	2775	33	5	38
VisitIZ	3006	33	19	52
WahakaMezcal	2771	32	34	66
totc	2771	31	3	34
SieteMisterios	2771	30	6	36
liladowns	3006	29	0	29
mezcal	3006	27	9	36
Montelobos	2771	27	7	34
Expendio_Oax	2775	24	3	27
ciecwtcmexico	3006	23	0	23
sabormexico	3006	21	18	39

label	modularity_class	indegree	outdegree	degree
c_insurgentes	2775	20	21	41
Dssdufoo	3006	20	19	39
Lakatrinamx	3006	20	23	43
mezcal_oaxaca	2775	19	9	28
pedrocastelan	3006	19	20	39
sidraSN	2775	19	17	36
df_condesa	2775	18	10	28
Eater	2771	18	1	19
meneseselayo	3006	18	41	59
Angustina	3006	17	27	44
BRUXOMEZCAL	2771	17	7	24
Dios_Padre	2775	17	0	17
NayaaMezcal	3006	17	0	17
raychagoya	2775	17	4	21
RosmarRestauran	3006	17	10	27
crucrumx	2775	16	0	16
LaEuropeaMexico	2771	16	2	18
licgallegos	3006	16	12	28
mezcaloro	2771	16	2	18
MezcalVago	2771	15	4	19
Mezcologia	2775	15	14	29
tipsdeviajero	3006	15	0	15
Carlos_Borboa	2771	14	2	16
df_soy	2775	14	18	32
dollanag	3006	14	11	25
mx_df	2775	14	5	19
PuntoMX_Madrid	3006	14	7	21
SabraDios1	2775	14	1	15
ShareThis	2771	14	0	14
ViceMexico	2775	14	0	14
AMHIZAC	3006	13	10	23
CervBalaPerdida	2775	13	0	13
Chef_Reynada	3006	13	14	27
ComeAcapulco	3006	13	0	13
DONDEIRweb	2775	13	0	13
Luciferina34	2771	13	6	19

label	modularity_class	indegree	outdegree	degree
Maria_LeonPL	2775	13	0	13
mezcalbuenviaje	3006	13	1	14
Porfiriamezcal	3006	13	4	17
VisitGuerrero_	3006	13	1	14
yosedooonde	2775	13	15	28
CcgmGuerrero	3006	12	3	15
FiftyMils	2771	12	5	17
magnimty	3006	12	0	12
SalonCascabel	3006	12	4	16
CORTV_Radio	2775	11	0	11
CulpaHola	2775	11	3	14
JamesESchroeder	2771	11	0	11
JavPlascencia	3006	11	5	16
mayodurand	3006	11	21	32
MezcaleriaOax	2771	11	5	16
Mistercono	2775	11	10	21
SaboresGuerrero	3006	11	2	13
el_cineclub	2775	10	2	12
LaParejaViajera	3006	10	0	10
lwoodhouse	2771	10	0	10
MacCocktail	2771	10	24	34
monicadionne	3006	10	2	12
RicardoRocha_MX	3006	10	0	10
AristeguiOnline	2775	9	0	9
EIAserin	2775	9	1	10
IMEXEsp	3006	9	2	11
laloplascencia	3006	9	2	11
Lat20Playa	2775	9	0	9
max_cerecedo	3006	9	0	9
PoliPocket	2775	9	2	11
rahelA2guzman	3006	9	3	12
siembraspirits	2771	9	1	10
amayamexico	2771	8	2	10
beautifulbooze	2771	8	5	13
conAntojomx	2775	8	0	8
CulturaColectiv	2775	8	0	8

label	modularity_class	indegree	outdegree	degree
ItzelAriadnaOca	3006	8	3	11
lera_jorge	2775	8	3	11
Mezcal_BenditAl	2775	8	4	12
MezcalTinieblo	2771	8	0	8
PasionMezcal	2775	8	0	8
shorba	2771	8	2	10
USBGDC	2771	8	6	14
cantinaypunto	3006	7	0	7
CarlsJrMx	3006	7	0	7
Cuauhtemoc_1521	2775	7	1	8

Apéndice 8. Lista de “*influencers*” de la red de retweets 2

label	Indegree	outdegree	degree	pageranks	modularity_class
NacionMezcal	474	24	498	0.07347366	36
marioeherrera	148	21	169	0.02355766	6
MezcalesOaxaca	128	21	149	0.02354857	36
Mezcologia	119	22	141	0.02262017	6
chagoya_victor	27	2	29	0.01365663	139
CircuitoMezcal	17	1	18	0.01050056	139
agavache	20	7	27	0.01016613	6
madame_bijoux	31	0	31	0.00999479	139
CBagatech	22	5	27	0.00945323	139
Oaxacking	14	1	15	0.00818593	6
vernemexico	17	0	17	0.00847572	881
mezcalesconalma	37	3	40	0.00762605	385
Lakatrinamx	18	5	23	0.00739483	434
MezcalConvite	18	3	21	0.0073715	385
eduardocervants	22	1	23	0.0069879	881
ortizromeroc	8	3	11	0.00626887	881
MiMuseoIndigena	18	0	18	0.00588506	434
Oaxaca_Digital	22	0	22	0.00586821	385
BladimirMJuarez	14	14	28	0.00572652	385
CcgmGuerrero	16	4	20	0.00551535	434
SoyMezcalero	22	11	33	0.00514855	385

label	Indegree	outdegree	degree	pageranks	modularity_class
VisitTaxco	10	0	10	0.00514845	434
CDI_mx	25	0	25	0.0051479	434
CortijoLaMezca	8	2	10	0.00512297	385
GranMittla	12	4	16	0.00497991	385
sanzvill	5	1	6	0.0049378	434
IBAmezcal	25	3	28	0.00458299	385
aleq_andro	6	0	6	0.00449193	881
OrodeOaxacaCdMx	7	0	7	0.00443407	881
drewbie_g	10	1	11	0.00439946	434
Chio_Vilchis	4	0	4	0.00437359	881
Agavenews	4	0	4	0.00436971	881
LaFabulosaNita	5	0	5	0.00430369	881
MezcalEICortijo	17	0	17	0.00429777	385
Angustina	23	7	30	0.00428706	434
PorqueSoyRaquel	5	1	6	0.0042635	881
mardmarcelina	3	0	3	0.00400228	881
Mezcalillera_	22	3	25	0.0039893	385
Vive_oaxaca	3	1	4	0.00398901	881
toledomariana	4	1	5	0.00387137	881
El_Oriente	3	1	4	0.0038658	881
jose_barros	2	0	2	0.00382599	881
Isma7903	1	2	3	0.00380324	881
RenatoGaliciaM	1	1	2	0.00380324	881
debrayando	1	0	1	0.00380324	881
maheva_13	1	0	1	0.00380324	881
jesuscanseco	2	1	3	0.00371435	881
Nuvia_Mayorga	23	0	23	0.00354449	434
Chef_Reynada	16	7	23	0.00347893	434
LaDamajuanaCabo	7	1	8	0.00340043	385
MezcalJoyaAzul	9	7	16	0.00321724	881
Porfiriamezcal	9	2	11	0.00310013	434
LaEuropeaMexico	22	0	22	0.00304247	385
humbertodiaz22	10	2	12	0.00293888	434
anasaldana	3	0	3	0.00280196	385
odontsoto	3	1	4	0.0027077	385
conchmurillo	17	4	21	0.00269658	434

label	Indegree	outdegree	degree	pageranks	modularity_class
IMCINE	15	0	15	0.00268186	881
IDConline	2	0	2	0.00265462	385
BonitoOaxaca	3	0	3	0.00263837	385
myjuicytaco	2	0	2	0.00256742	385
EffingTastyBev	2	0	2	0.00251529	385
menesespelayo	10	7	17	0.0025068	434
NineAcapulco	9	0	9	0.00248507	434
MareGorjon	14	5	19	0.00245291	434
adaircesar	2	0	2	0.00238736	385
Jclicona12	11	1	12	0.00237849	434
CentroPublicoMX	4	0	4	0.00230474	881
rojasector	1	0	1	0.00227594	385
chatero	13	0	13	0.00221803	434
mimezcal_	6	3	9	0.00217078	434
VogueMexico	13	0	13	0.00215704	881
RutadelMezcal	4	0	4	0.00211637	434
Rayadita_Bleu77	2	0	2	0.00198984	385
sandriuri	2	0	2	0.00198984	385
pedrocastelan	5	6	11	0.0019852	434
VisitIZ	12	2	14	0.001977	434
MenuAcapulco	12	2	14	0.0019744	434
yuyulopez8	2	1	3	0.00191022	385
VickyAladro	2	0	2	0.00191022	385
soymareli	2	0	2	0.00191022	385
NaaLiuu	15	3	18	0.00189042	881
MeliMarroquin16	11	2	13	0.00184229	434
LmedinaN	2	0	2	0.00183491	385
ladyli_929	1	0	1	0.00183061	385
dmt_hanna	1	0	1	0.00183061	385
tiempodemi	1	0	1	0.00183061	385
Gelosnivon	1	0	1	0.00183061	385
khersank	1	0	1	0.00183061	385
vicman310	3	1	4	0.00179412	434
DigitalGuerrero	12	0	12	0.00179293	434
javefire	2	0	2	0.00164509	385
sdd	2	0	2	0.00164509	385

label	Indegree	outdegree	degree	pageranks	modularity_class
Mistercono	9	3	12	0.00163831	881
CapitalOax	8	0	8	0.0015277	434
CasadelCineMx	9	0	9	0.00145522	881
RomanYeah	1	0	1	0.00144442	434
PedroHacesSordo	9	0	9	0.00142383	434
CostanzoOficial	8	0	8	0.00138158	434