



Pontificia Universidad
Católica del Ecuador

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE INGENIERÍA

**ESTUDIO DE MERCADO UTILIZANDO LA RED SOCIAL TWITTER, PARA
HOTELES DECAMERON**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL GRADO DE
MAGISTER EN SISTEMAS DE INFORMACIÓN MENCIÓN CIENCIA DE DATOS**

LENNIS VERONICA MANOSALVAS PORRAS

Director: ALFREDO CALDERON Msc.

Quito, diciembre 2022

ÍNDICE DE CONTENIDO

1	Introducción	11
1.1	Planteamiento del Problema.	11
1.2	Objetivos	11
1.2.1	Objetivo General	11
1.2.2	Objetivos Específicos.....	11
1.3	Marco Teórico	11
1.3.1	Hoteles Decameron	12
1.3.2	Hoteles Decameron En Ecuador.	12
1.3.3	Multivacaciones Ecuador.	12
1.3.4	Valores De Hoteles Decameron	13
1.3.5	Estudio De Mercado.....	13
1.3.6	Fuentes de Datos	14
2	Procesamiento del Lenguaje Natural.....	15
2.1	Minería De Datos	15
2.2	Procesamiento del lenguaje natural (PLN)	17
2.2.1	Aplicaciones del Procesamiento del Lenguaje Natural.....	17
2.3	Análisis de Sentimientos	18
2.3.1	Tipos de análisis de sentimiento:.....	19
2.3.2	Aplicaciones del análisis de sentimiento	19
2.3.3	Desafíos con el análisis de sentimientos.....	20
2.3.4	Computación Emotiva.....	20
2.3.5	Nube de Palabras:	20
3	Pre-procesamiento de Datos	21
3.1	Extracción de los datos	21
3.1.1	Creación de una Cuenta de Desarrollador	21
4	Procesamiento	28
4.1	Análisis para Decameron Mompiche	28
4.2	Análisis para el Hotel Decameron Punta Centinela	40

4.3	Análisis para Multivacaciones	53
4.4	Análisis de la Encuesta	67
4.4.1	¿Cuál es tu edad?	67
4.4.2	¿Cuál es tu género?.....	67
4.4.3	¿En qué ciudad vives?	67
4.4.4	¿Te gusta las instalaciones del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.	67
4.4.5	¿Te gusta las instalaciones del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.	72
4.4.6	¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.....	78
4.4.7	¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.....	84
4.4.8	Qué opinas de multivacaciones?	90
4.4.9	Tiene algún comentario adicional?	110
5	Conclusiones y Recomendaciones	117
5.1	Conclusiones	117
5.2	Recomendaciones	117
6	Referencias Bibliográficas	118

ÍNDICE DE FIGURAS

Figura 1: Crear aplicación con acceso Elevate	22
Figura 2: Generar Keys	22
Figura 3: Accesos de Investigador Académico para Twitter	24
Figura 4: Accesos Cuenta de Investigador Académico.....	24
Figura 5: Código para conectar a la API de Twitter.....	25
Figura 6: Análisis de sentimientos en español para la búsqueda “Decameron Mompiche”	29
Figura 7: Análisis de sentimientos en inglés para la búsqueda “Decameron Mompiche”	29
Figura 8: Sentimientos por año para “Decameron Mompiche” (a)	30
Figura 9: Sentimientos por año “Decameron Mompiche” (b)	30
Figura 10: Nube de Palabras para la búsqueda “Decameron Mompiche”	31
Figura 11: Nube de Palabras Positivas para la búsqueda “Decameron Mompiche”	31
Figura 12: Nube de Palabras Neutras para la búsqueda “Decameron Mompiche”	32
Figura 13: Nube de Palabras Negativas para la búsqueda “Decameron Mompiche”	32
Figura 14: Conteo de palabras más usadas para “Decameron Mompiche”	33
Figura 15: Conteo de Bigramas más usadas “Decameron Mompiche”	34
Figura 16: Conteo de Trigramas más usados para “Decameron Mompiche”	35
Figura 17: Frecuencia de aplicaciones desde dónde se generaron los tweets para “Decameron Mompiche”	36
Figura 18: Frecuencia localidades desde dónde se generaron los tweets para “Decameron Mompiche”	37
Figura 19: Usuarios con más amigos.....	37
Figura 20: Descripción de Usuarios con sus seguidores y amigos para “Decameron Mompiche”	38
Figura 21: Sentimientos por localidad (a)	39
Figura 22: Sentimientos por localidad (b)	39
Figura 23: Sentimientos en idioma español.....	41
Figura 24: Sentimientos en idioma inglés	41
Figura 25: Sentimientos por año para “Decameron Punta Centinela” (a).....	42
Figura 26: Sentimientos por año para “Decameron Punta Centinela” (b)	42
Figura 27: Nube de Palabras para “Decameron Punta Centinela”	43
Figura 28: Nube de Palabras positivas para “Decameron Punta Centinela”	43
Figura 29: Nube de Palabras neutras para “Decameron Punta Centinela”	44
Figura 30: Nube de Palabras negativas para “Decameron Punta Centinela”	44
Figura 31: Promedio de Palabras en tweets para “Decameron Punta Centinela”	45
Figura 32: Conteo de Palabras más usadas para “Decameron Punta Centinela”	46
Figura 33: Conteo de Bigramas más usadas para “Decameron Punta Centinela”	47
Figura 34: Número de Trigramas más usados para “Decameron Punta Centinela”	48
Figura 35: Frecuencia de dispositivos desde dónde se publicaron los tweets para “Decameron Punta Centinela”	49
Figura 36: Frecuencia de las localidades para tweets generados para “Decameron Punta Centinela”	49
Figura 37: Sentimientos por localidades para “Decameron Punta Centinela” (a).....	50
Figura 38: Sentimientos por localidades para “Decameron Punta Centinela” (b)	51
Figura 39: Usuarios con más amigos para “Decameron Punta Centinela”	52

Figura 40: Descripción de Usuarios con sus seguidores y amigos para “Decameron Punta Centinela”	52
Figura 41: Análisis en español para los resultados de la búsqueda Multivacaciones.....	54
Figura 42: Análisis en inglés para los resultados de la búsqueda Multivacaciones	54
Figura 43: Análisis de sentimientos por años para “Multivacaciones” (a)	55
Figura 44: Análisis de sentimientos por años para “Multivacaciones” (b)	55
Figura 45: Nube de palabras para “Multivacaciones”	56
Figura 46: Nube de palabras positivas para “Multivacaciones”	57
Figura 47: Nube de palabras neutras para “Multivacaciones”	57
Figura 48: Nube de palabras negativas para “Multivacaciones”	58
Figura 49: Longitud promedio del texto del análisis de sentimientos para “Multivacaciones”	59
Figura 50: Promedio de palabras del análisis de sentimientos para “Multivacaciones”	59
Figura 51: Frecuencia de palabras más usadas para “Multivacaciones”	60
Figura 52: Frecuencia de bigramas más usados para “Multivacaciones”	61
Figura 53: Frecuencia de trigramas más usados para “Multivacaciones”	62
Figura 54: Frecuencia de dispositivos desde dónde se publicaron los tweets para “Multivacaciones”	63
Figura 55: Localidades desde dónde se publicaron los tweets para “Multivacaciones”	63
Figura 56: Análisis de sentimientos por localidades para “Multivacaciones” (a)	64
Figura 57: Análisis de sentimientos por localidades para “Multivacaciones”	65
Figura 58: Usuarios con más amigos para “Multivacaciones”	66
Figura 59: Descripción de Usuarios con sus seguidores y amigos para “Multivacaciones”	66
Figura 60: Análisis de sentimientos para la Pregunta 4, de la encuesta.....	67
Figura 61: Análisis de sentimientos por localidades para la Pregunta 4 de la encuesta (a).	68
Figura 62: Análisis de sentimientos por localidades para la Pregunta 4 de la encuesta (b).....	69
Figura 63: Análisis de sentimientos por género para la Pregunta 4 de la encuesta.	70
Figura 64: Nube de palabras para la Pregunta 4 de la encuesta.	70
Figura 65: Nube de palabras positivas para la Pregunta 4 de la encuesta.	71
Figura 66: Nube de palabras neutras para la Pregunta 4 de la encuesta.	71
Figura 67: Nube de palabras negativas para la Pregunta 4 de la encuesta.	72
Figura 68: Análisis de sentimientos para la Pregunta 5 de la encuesta.....	72
Figura 69: Análisis de sentimientos por edades para la Pregunta 5 de la encuesta (a).	73
Figura 70: Análisis de sentimientos por edades para la Pregunta 5 de la encuesta. (b)	74
Figura 71: Análisis de sentimientos por localidad para la Pregunta 5 de la encuesta. (a)	75
Figura 72: Análisis de sentimientos por localidad para la Pregunta 5 de la encuesta. (b)	76
Figura 73: Análisis de sentimientos por genero para la Pregunta 5 de la encuesta.....	76
Figura 74: Nube de palabras para la Pregunta 5 de la encuesta.	77
Figura 75: Nube de palabras positivas para la Pregunta 5 de la encuesta.	77
Figura 76: Nube de palabras neutras para la Pregunta 5 de la encuesta.	78
Figura 77: Nube de palabras negativas para la Pregunta 5 de la encuesta.	78
Figura 78: Análisis de sentimientos para la Pregunta 6, de la encuesta.....	79
Figura 79: Análisis de sentimientos por edades para la Pregunta 6 de la encuesta. (a)	80
Figura 80: Análisis de sentimientos por edades para la Pregunta 6 de la encuesta. (b)	80
Figura 81: Análisis de sentimientos por localidad para la Pregunta 6 de la encuesta. (a)	81
Figura 82: Análisis de sentimientos por localidad para la Pregunta 6 de la encuesta. (b)	82
Figura 83: Análisis de sentimientos por genero para la Pregunta 6 de la encuesta.....	82

Figura 84: Nube de palabras para la Pregunta 6 de la encuesta.	83
Figura 85: Nube de palabras positivas para la Pregunta 6 de la encuesta.	83
Figura 86: Nube de palabras neutras para la Pregunta 6 de la encuesta.	84
Figura 87: Nube de palabras negativas para la Pregunta 6 de la encuesta.	84
Figura 88: Análisis de sentimientos para la Pregunta 7 de la encuesta.	85
Figura 89: Análisis de sentimientos por edades para la Pregunta 7 de la encuesta. (a)	85
Figura 90: Análisis de sentimientos por edades para la Pregunta 7 de la encuesta. (b)	86
Figura 91: Análisis de sentimientos por localidad para la Pregunta 7 de la encuesta. (a)	87
Figura 92: Análisis de sentimientos por localidad para la Pregunta 7 de la encuesta. (b)	88
Figura 93: Análisis de sentimientos por genero para la Pregunta 7 de la encuesta.	88
Figura 94: Nube de palabras para la Pregunta 7 de la encuesta.	89
Figura 95: Nube de palabras positivas para la Pregunta 7 de la encuesta.	89
Figura 96: Nube de palabras neutras para la Pregunta 7 de la encuesta.	90
Figura 97: Nube de palabras negativas para la Pregunta 7 de la encuesta.	90
Figura 98: Análisis de sentimientos para la Pregunta 8 de la encuesta.	91
Figura 99: Análisis de sentimientos por edad para la Pregunta 8 de la encuesta. (a)	91
Figura 100: Análisis de sentimientos por edad para la Pregunta 8, de la encuesta. (b)	92
Figura 101: Análisis de sentimientos por localidad para la Pregunta 8 de la encuesta. (a)	93
Figura 102: Análisis de sentimientos por localidad para la Pregunta 8 de la encuesta. (b)	94
Figura 103: Análisis de sentimientos por genero para la Pregunta 8 de la encuesta.	95
Figura 104: Nube de palabras para la Pregunta 8, de la encuesta.	95
Figura 105: Nube de palabras positivas para la Pregunta 8 de la encuesta.	96
Figura 106: Nube de palabras neutras para la Pregunta 8 de la encuesta.	96
Figura 107: Nube de palabras negativas para la Pregunta 8 de la encuesta.	97
Figura 108: Análisis de sentimientos para la Pregunta 9 de la encuesta.	97
Figura 109: Análisis de sentimientos por edades para la Pregunta 9 de la encuesta. (a)	98
Figura 110: Análisis de sentimientos por edades para la Pregunta 9 de la encuesta. (b)	99
Figura 111: Análisis de sentimientos por localidad para la Pregunta 9 de la encuesta.	100
Figura 112: Análisis de sentimientos por localidad para la Pregunta 9 de la encuesta.	101
Figura 113: Análisis de sentimientos por genero para la Pregunta 9 de la encuesta.	102
Figura 114: Nube de palabras para la Pregunta 9 de la encuesta.	102
Figura 115: Nube de palabras positivas para la Pregunta 9 de la encuesta.	103
Figura 116: Nube de palabras neutras para la Pregunta 9 de la encuesta.	103
Figura 117: Nube de palabras negativas para la Pregunta 9 de la encuesta.	104
Figura 118: Análisis de sentimientos para la Pregunta 10 de la encuesta.	104
Figura 119: Análisis de sentimientos por edades para la Pregunta 10 de la encuesta. (a)	105
Figura 120: Análisis de sentimientos por edades para la Pregunta 10 de la encuesta. (b)	105
Figura 121: Análisis de sentimientos por localidad para la Pregunta 10 de la encuesta. (a)	106
Figura 122: Análisis de sentimientos por localidad para la Pregunta 10 de la encuesta. (b)	107
Figura 123: Análisis de sentimientos por genero para la Pregunta 10 de la encuesta.	108
Figura 124: Nube de palabras para la Pregunta 10 de la encuesta.	108
Figura 125: Nube de palabras positivas para la Pregunta 10 de la encuesta.	109
Figura 126: Nube de palabras neutras para la Pregunta 10 de la encuesta.	109
Figura 127: Nube de palabras negativas para la Pregunta 10 de la encuesta.	110

Figura 128: Análisis de sentimientos para la Pregunta 11 de la encuesta.....	110
Figura 129: Análisis de sentimientos por edades para la Pregunta 11 de la encuesta. (a)	111
Figura 130: Análisis de sentimientos por edades para la Pregunta 11 de la encuesta. (b)	112
Figura 131: Análisis de sentimientos por localidades para la Pregunta 11 de la encuesta. (a)	113
Figura 132: Análisis de sentimientos por localidades para la Pregunta 11 de la encuesta. (b).....	114
Figura 133: Análisis de sentimientos por genero para la Pregunta 11, de la encuesta.	114
Figura 134: Nube de palabras para la Pregunta 11 de la encuesta.	115
Figura 135: Nube de palabras positivas para la Pregunta 11de la encuesta.	115
Figura 136: Nube de palabras neutras para la Pregunta 11 de la encuesta.	116
Figura 137: Nube de palabras negativas para la Pregunta 11 de la encuesta.	116

ÍNDICE DE TABLAS

Tabla 1: Número de tweets por año para la búsqueda “Decameron Mompiche”	28
Tabla 2: Número de palabras más usadas para “Decameron Mompiche”	33
Tabla 3: Bigramas más usadas “Decameron Mompiche”	34
Tabla 4: Trigramas más usados para “Decameron Mompiche”	35
Tabla 5: Número de tweets por año para “Decameron Punta Centinela”	40
Tabla 6: Longitud promedio de los tweets para “Decameron Punta Centinela”	44
Tabla 7: Promedio de palabras de los tweets para “Decameron Punta Centinela”	45
Tabla 8: Palabras más nombradas para “Decameron Punta Centinela”	46
Tabla 9: Número de Bigramas más usados para “Decameron Punta Centinela”	47
Tabla 10: Número de Trigramas más usados para “Decameron Punta Centinela”	48
Tabla 11: Tweets por años para “Multivacaciones”	53
Tabla 12: Longitud promedio de los tweets para “Multivacaciones”	58
Tabla 13: Promedio de palabras de los tweets para “Multivacaciones”	58
Tabla 14: Conteo de Palabras más usadas para “Multivacaciones”	60
Tabla 15: Conteo de Bigramas más usadas para “Multivacaciones”	61
Tabla 16: Número de Trigramas más usados para “Multivacaciones”	62

RESUMEN

En la actualidad, las empresas están adoptando nuevos enfoques, pues buscan llegar a sus clientes de forma personalizada para mejorar la calidad de sus servicios. Con el avance de la tecnología, cada día se genera grandes volúmenes de datos de forma digital. En este contexto, las redes sociales están adquiriendo un papel significativo, adoptándose como medios ampliamente usados para compartir diversas opiniones sobre un sin número de temas. Por tanto, estos datos se han convertido en una fuente valiosa para conocer los niveles de satisfacción de los usuarios.

En el presente trabajo de titulación se realizará un estudio de mercado a través del análisis de opiniones expresadas vía Twitter sobre los servicios que ofrece Hoteles Decameron, mediante la aplicación de Minería de Datos y procesamiento de lenguaje natural (PLN), para determinar el grado de aceptación o rechazo hacia los servicios que ofrece la empresa.

A través del análisis de sentimientos, se identificará si las opiniones son positivas, negativas o neutras. Además, se buscará palabras claves, que permitirán determinar niveles de acogida de los servicios ofrecidos por los Hoteles Decameron, dependiendo de los años de emisión de los Tweets y los segmentos geográficos de las personas que emitieron las opiniones.

Asimismo, se realizará una encuesta de forma digital, con preguntas personalizadas sobre los servicios que ofrece Hoteles Decameron Mompiche, Hoteles Decameron Punta Centinela, Multivacaciones, estas opiniones permitirán complementar el análisis.

Palabras clave: Estudio de Mercado, Twitter, Minería de Datos, Procesamiento de Lenguaje Natural (PLN), Análisis de Sentimientos, Decameron.

ABSTRACT

Currently, companies are adopting new approaches, in so far as to reach their customers in a personalized way to improve the quality of their services. As technology advances, large volumes of data are generated digitally every day. In this context, social networks are acquiring a significant role, being adopted as widely used means to share various opinions on a number of topics. This data has become a valuable source for determining levels of user satisfaction.

In the present titling work, a market study will be carried out through the analysis of opinions expressed via Twitter about the services offered by Decameron Hotels, through the application of Data Mining and Natural Language Processing (NLP) to determine the degree of acceptance or rejection of the services offered by the company.

Sentiment Analysis will allow to identify whether the opinions are positive, negative or neutral. In addition, keywords will be searched, which will allow determining levels of reception of the services offered by Decameron Hotels, depending on the years of emission of the Tweets and the geographic segments of the people who issued the opinions.

Furthermore, a survey will be carried out digitally, with personalized questions about the services offered by Decameron Mompiche Hotels, Decameron Punta Centinela Hotels, Multivacaciones. These opinions will complement the analysis

Keywords: Market Research, Twitter, Data Mining, Natural Language Processing (NLP), Sentiment Analysis, Decameron.

1 Introducción

1.1 Planteamiento del Problema.

La necesidad constante de las empresas por mejorar sus productos y servicios ha permitido incursionar en nuevos métodos de análisis de datos. Hoy en día, los datos digitales forman pieza clave para el desarrollo de estudios de mercado.

Es de gran utilidad para las empresas conocer qué opinan los consumidores con respecto a sus productos o servicios en redes sociales. El presente estudio de mercado se realizará sobre los datos de la red social Twitter para Hoteles Decameron.

Hoteles Decameron podrá usar la información proveída por el estudio de mercado, para tomar decisiones estratégicas informadas y de esta manera mejorar las áreas donde se evidencie inconformidad por parte de los usuarios y así llegar a una población más amplia.

En los últimos tiempos han surgido herramientas de uso libre que apoyan el desarrollo de proyectos de análisis de datos. En el presente trabajo, se utilizará el lenguaje de programación Python.

1.2 Objetivos

1.2.1 Objetivo General

Desarrollar un estudio de mercado utilizando la red social twitter, para hoteles Decameron.

1.2.2 Objetivos Específicos

Determinar el nivel de satisfacción de los usuarios de los servicios ofrecidos por Hoteles Decameron Monpiche.

Evaluar el porcentaje de aceptación y rechazo a los servicios ofrecidos por Hoteles Decameron Punta Centinela.

Analizar el porcentaje de aceptación y rechazo a los servicios ofrecidos por Multivacaciones.

Establecer las localidades donde los servicios de los Hoteles Decameron tienen menor aceptación.

1.3 Marco Teórico

En esta sección se detalla los conceptos y términos necesarios para entender el presente trabajo.

1.3.1 Hoteles Decameron

Hoteles Decameron es una empresa multinacional con presencia en Colombia, Perú, Ecuador, México, Panamá, Jamaica, Haití, El Salvador y Costa Rica. Se enfoca en la calidad de los productos y servicios que ofrece, y en generar experiencias únicas en sus huéspedes.

Hoteles Decameron inauguró su primer hotel en 1987, con la visión de ofrecer al viajero local como internacional, un servicio enfocado en la experiencia, gastronomía, bebidas, entretenimiento. Así se crea el primer hotel Decameron Don Blas en Cartagena. Unos meses después, en la isla de San Andrés, se abre otro hotel, y otro más en 1991 en Santa Marta. A partir de ese momento, comenzó la expansión de la marca a otras regiones de Colombia y países de la región.

1.3.2 Hoteles Decameron En Ecuador.

Ecuador cuenta con el Hotel Royal Decameron Mompiche y el Hotel Royal Decameron Punta Centinela.

1.3.2.1 Royal Decameron Mompiche

El Hotel Royal Decameron Mompiche, se encuentra en la provincia de Esmeraldas, por la carretera del Aeropuerto Coronel Carlos Concha Torres a 140 km del Centro de Esmeraldas, es decir, a un trayecto de aproximadamente 2 horas.

Royal Decameron Mompiche ofrece un servicio todo incluido que permite al viajero adquirir experiencias en cuanto a servicios, instalaciones, entretenimiento y naturaleza. Cuenta con 282 habitaciones, restaurantes especializados y tipo buffet, 6 bares, discoteca, 5 piscinas, y un centro de convenciones con capacidad para 600 personas.

1.3.2.2 Royal Decameron Punta Centinela

Ubicado en la Provincia de Santa Elena, desde el Aeropuerto Internacional José Joaquín de Olmedo de la ciudad de Guayaquil, por vía terrestre, a un trayecto de aproximadamente 2 horas.

Las instalaciones del complejo consisten en 197 habitaciones, 1 restaurante tipo buffet y 2 restaurantes a la carta, 3 bares, 2 piscinas, canchas deportivas y un centro de convenciones con capacidad para 400 personas con tecnología para el desarrollo de eventos.

1.3.3 Multivacaciones Ecuador.

Es el Programa Vacacional de la cadena hotelera Decameron, ofrece a sus afiliados paquetes de “Decas”, para reservar alojamiento en todos los Hoteles Decameron con sede en 9 países, así como en hoteles afiliados.

Las decas tienen vigencia conforme a los años acordados en el contrato vacacional. Aparte del costo de adquisición de las decas se debe realizar un pago anual por el mantenimiento del Programa.

1.3.4 Valores De Hoteles Decameron

Hoteles Decameron busca llegar a sus huéspedes a través de los siguientes valores:

Frescura:

Busca en la experiencia de los viajeros la naturalidad de unas vacaciones sin complicaciones con las que se olviden del mundo externo y puedan ser ellos mismos

Calidez:

Desea transmitir sensaciones únicas, como paz y tranquilidad.

Generosidad:

Proporciona todo para complacer y sorprender a los viajeros, ofreciendo un paquete lleno de posibilidades, comodidades, sorpresas y mucha diversión.

Cercanía:

Busca conocer las necesidades y expectativas de los viajeros para que puedan vivir momentos llenos de calidez que los haga sentir mejor que en casa.

Libertad:

Dar rienda suelta a la imaginación haciendo realidad todo lo que más ha deseado para convertir sus vacaciones en una experiencia única.

Intimidad:

Ofrece una experiencia inolvidable que hará vibrar todos tus sentidos a través de una atención casi personalizada. (Decameron, s.f.)

El presente trabajo se centra en el valor “Cercanía”; se requiere determinar el nivel de satisfacción de los huéspedes con respecto a los servicios ofrecidos por Hoteles Decameron Mompiche, Decameron Punta Centinela y Multivacaciones, a través del análisis de sentimientos de las opiniones que los usuarios publican en la red social Twitter.

1.3.5 Estudio De Mercado

Se llama estudio de mercado a aquella investigación usada por varias industrias para afianzar la toma de decisiones y tener una mejor comprensión de sus operaciones y los retos a los que se enfrentan en determinado panorama empresarial. En nuestro caso, queremos determinar la manera

de comportarse de los usuarios en las redes sociales, relacionado con los servicios que ofrece Hoteles Decameron.

Para la empresa, su principal objetivo es satisfacer las necesidades de los turistas, por ese motivo es indispensable identificar exactamente cuáles son esas necesidades.

Nos centraremos en conocer que expectativas, motivaciones, opiniones de satisfacción o rechazo son expresadas por los turistas en la red social Twitter acerca de los Hoteles Decameron y Multivaciones.

1.3.6 Fuentes de Datos

1.3.6.1 Twitter

Es uno de los medios sociales gratuitos más importantes. Permite que diversos ámbitos sociales, como el familiar, laboral o amistoso, se comuniquen y se mantengan conectados entre sí mediante el intercambio de mensajes rápidos y frecuentes. La red consiste en publicar y leer mensajes de toda índole, como pueden ser noticias de última hora, tendencias actuales, entretenimiento, deportes, intereses cotidianos, opiniones. Los mensajes pueden contener fotos, videos, enlaces y texto; y se publican en los perfiles, se envían a sus seguidores y se encuentran en un sistema de búsqueda.

Para el presente estudio se buscará y recuperará tweets o mensajes relacionados con Decameron Ecuador en la plataforma de Twitter, a través de las siguientes palabras claves:

Decameron Mompiche
Decameron Punta Centinela
Multivaciones.

Adicional, para complementar el estudio y enfocarnos en opiniones puntuales sobre los servicios de Decameron Ecuador, se creó una encuesta, diseñada en el programa Twtpoll, y publicada en la red social Twitter.

Twtpoll, es una herramienta que tiene una versión gratuita y versiones pago que permite configurar encuestas y publicarlas en distintas redes sociales, incluyendo Twitter.

1.3.6.2 Encuesta:

Las preguntas incluidas en la encuesta son de opinión, se procura obtener respuestas abiertas.

¿Cuál es tu edad?

¿Cuál es tu género?

¿En qué ciudad vives?

¿Te gusta las instalaciones del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.

¿Te gusta las instalaciones del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.

¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.

¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel

¿Qué opinas sobre multivacaciones?

¿Regresarías a los Hoteles de Decameron?, si no los ha visitado, ¿te gustaría conocerlos?

¿Recomendarías a familiares, amigos o colegas Hoteles Decameron?

¿Tienes algún comentario adicional?

2 Procesamiento del Lenguaje Natural

2.1 Minería De Datos

La minería de datos es el proceso de obtener información valiosa mediante el análisis de grandes volúmenes de datos, este proceso busca descubrir patrones de interés, clasificar datos, predecir las salidas, es decir, los resultados, y hacer explícitas conexiones entre los datos. Este proceso usa herramientas automáticas que incorpora algoritmos sofisticados, sobre todo para encontrar patrones ocultos, asociaciones, anomalías, y estructuras desde el gran volumen de datos recolectados.

Una aplicación de la minería de datos es el análisis de sentimientos. Consiste en tratar de identificar y extraer opiniones de un texto. En este sentido, es un subcampo del procesamiento del Lenguaje Natural (PLN). La finalidad de esta aplicación es medir las actitudes y emociones basándose en el tratamiento computacional de la subjetividad contenida en los textos donde se refleja el habla cotidiana de los seres humanos.

Una de las herramientas que ofrecen más bondades para la minería de datos es el lenguaje de programación Python. Python, presenta una alta gama de herramientas que permiten descargar, transformar, almacenar y realizar análisis de sentimientos de los datos publicados en twitter.

Para nuestro caso de estudio se utilizará las librerías Tweepy, Textblob, Matplotlib, Seaborn, Plotly, Wordcloud, Sklearn, Psycpg2, Twarc y Pandas de Python. A continuación, se describe el detalle de cada una:

Tweepy: Tweepy es una Librería de código abierto que proporciona una manera muy fácil, de acceder a la API de Twitter. Es excelente para realizar automatizaciones básicas y para la creación de bots de Twitter. Tweepy permite consultar tweets en la línea de tiempo, así como crear y eliminar Tweets.

TextBlob: TextBlob es una librería de Python para procesar texto. Permite realizar tareas comunes de procesamiento de lenguaje natural mediante una API simple. Entre estas tareas se encuentra el etiquetado de partes del discurso, la extracción de frases nominales, el análisis de sentimientos, la clasificación y traducción.

Matplotlib: Matplotlib es una librería que permite generar visualizaciones estáticas, animadas e interactivas en Python. Algunas de sus visualizaciones más importantes son histogramas, diagramas de barras, diagramas de caja y bigote, diagramas de dispersión, diagramas de líneas, mapas de calor, entre otros.

Seaborn: Basada en Matplotlib, Seaborn es una librería de visualización de datos de Python. Brinda una interfaz para realizar gráficos estadísticos de alto nivel.

Plotly: Librería de gráficos de Python permite crear gráficos interactivos con calidad de publicación, cómo diagramas de líneas, diagramas de dispersión, gráficos de área, gráficos de barras, barras de error, diagramas de caja, histogramas, mapas de calor, subgráficos, gráficos de ejes múltiples, gráficos polares y gráficos de burbujas

Wordcloud: Esta librería permite representar los datos de acuerdo con su frecuencia, las palabras que tengan mayores repeticiones aparecerán con mayor tamaño en la visualización

Sklearn: Esta librería permite la implementación del aprendizaje automático en Python. Facilita una serie de herramientas para el aprendizaje automático y el modelado estadístico, como modelos de clasificación, regresión, agrupación, reducción de dimensionalidad.

Nltk: Librería que tiene funcionalidades para elaborar procesamiento de texto para clasificación, tokenización, lematización, etiquetado, análisis y razonamiento semántico.

Psycpg2: Psycpg es el adaptador para base de datos PostgreSQL, posibilita conectar Python con la base de Datos PostgreSQL y ejecutar instrucciones DML y DDL, como "INSERT" o "UPDATE".

2.2 Procesamiento del lenguaje natural (PLN)

El procesamiento del lenguaje natural trata de lograr que las máquinas comprendan datos de texto o voz y respondan con su propio texto o voz, tal como lo hacen los humanos.

El procesamiento del lenguaje natural, uno de los pilares de la ciencia computacional, especialmente en la rama de la inteligencia artificial (IA), que se ocupa de permitir que las computadoras entiendan el texto y el lenguaje hablado tal como es comprendido por los humanos cotidianamente.

Esta rama de la computación, combinando las reglas del lenguaje humano (campo estudiado por la lingüística computacional) con modelos estadísticos, aprendizaje automático, y modelos de redes neuronales, procesa el lenguaje natural, ya sea en forma hablada (datos de voz) o escrita (texto), posibilitando que las computadoras comprendan completamente su significado, incluyendo las intenciones y emociones de quien habla o escribe. Esto lo logra automatizando el proceso de extraer, clasificar y etiquetar los elementos de texto y datos de voz, designando una probabilidad estadística a cada posible significado de esos elementos.

En principio, la computarización del lenguaje natural humano debería implicar su formalización, es decir, su conversión a un lenguaje comprensible por las máquinas a través de un criterio lógico-formal. Esto es notablemente difícil porque el lenguaje natural está impregnado de innumerables irregularidades y ambigüedades. Por eso, un software que procese lenguaje natural humano debe tomar en cuenta homónimos, homófonos, sarcasmo, modismos, metáforas, excepciones gramaticales y de uso, variaciones en la estructura de las oraciones, para realizar su cometido con precisión desde el principio.

2.2.1 Aplicaciones del Procesamiento del Lenguaje Natural

Detección de Spam: Las compañías que brindan servicios de correo electrónico usan el Procesamiento de lenguaje natural para detectar correos maliciosos o SPAM. Esto se realiza escaneando correos electrónicos que contengan lenguaje que a menudo indica spam o phishing. Este tipo de lenguaje usualmente incluye el uso excesivo de términos financieros, mala gramática, lenguaje amenazador, urgencia desmedida, nombres de grandes corporaciones mal escritos, entre otros.

Traducción automática: Google Translate es un ejemplo de tecnología NLP ampliamente disponible en el trabajo. La traducción automática verdaderamente útil implica más que reemplazar palabras en un idioma con palabras de otro. La traducción efectiva tiene que capturar con precisión el significado y el tono del idioma de entrada y traducirlo a texto con el mismo significado e impacto deseado en el idioma de salida

Agentes virtuales y chatbots: Sistemas GPS operados por mando de voz, software de dictado de voz a texto, los agentes virtuales como Siri de Apple y Alexa de Amazon utilizan el reconocimiento de voz para reconocer patrones en los comandos de voz y la generación de lenguaje natural para responder con la acción adecuada o comentarios útiles. Los chatbots realizan el mismo trabajo en respuesta a las entradas de texto escritas. Los mejores también aprenden a reconocer

pistas contextuales sobre solicitudes humanas y las utilizan para proporcionar respuestas u opciones aún mejores con el tiempo. La próxima mejora para estas aplicaciones es la respuesta a preguntas, la capacidad de responder a nuestras preguntas, anticipadas o no, con respuestas útiles y relevantes en sus propias palabras.

Resumen de texto: El resumen de texto utiliza técnicas de PLN para interpretar grandes volúmenes de texto digital y crear resúmenes y sinopsis para índices, bases de datos de investigación o lectores ocupados que no tienen tiempo para leer el texto completo. Las mejores aplicaciones de resumen de texto utilizan razonamiento semántico y generación de lenguaje natural (GLN) para agregar contexto útil y conclusiones a los resúmenes.

Soluciones Empresariales: El procesamiento de lenguaje natural también se enfoca en soluciones empresariales que sirven para optimizar las operaciones comerciales, aumentar la productividad de los empleados y simplificar los procesos comerciales de misión crítica.

Medicina: Otra aplicación se enfoca al ámbito de la medicina, a través del análisis de datos, incluyen historia clínica electrónica, relatos de los pacientes, resúmenes de alta, protocolos de ensayos clínicos y datos bibliográficos, se busca encontrar diagnósticos tempranos a enfermedades.

Análisis del sentimiento de las redes sociales: el PLN ha cobrado una gran importancia como instrumento comercial de obtención de información oculta en los canales de las redes sociales. Esta herramienta analiza el lenguaje natural humano expresado en diferentes formatos en las diversas redes sociales, de donde extrae los datos referentes a las respuestas que emiten los usuarios a diferentes estímulos como productos, promociones y eventos. Esta información puede ser utilizada por las empresas para sus distintas operaciones de diseño, marketing, etc.

2.3 Análisis de Sentimientos

Las tendencias a gran escala en las opiniones de los usuarios de Twitter pueden ser detectadas a través de los mensajes que estos publican en la red social. Esto, debido a que las opiniones se hacen públicas a través de los tuits, lo cual propicia cierta conversación informal generalizada entre los usuarios. Así, el autor Jungherr señala que “En el contexto de la comunicación política, el análisis de contenido y los estudios cuantitativos de los mensajes de Twitter permiten identificar patrones de comportamiento entre los usuarios y puntos de inflexión en las corrientes de opinión” (2015).

Uno de los enfoques del procesamiento del lenguaje natural es el análisis de sentimientos, que se conoce también como minería de opiniones. Sirve para identificar el tono emocional implícito en determinado texto. Las organizaciones usan extensamente este enfoque para determinar y categorizar las opiniones que los usuarios tienen respecto a productos, servicios o ideas. Conlleva el uso de minería de datos, aprendizaje automático e inteligencia artificial para identificar sentimiento e información subjetiva en el texto.

Las organizaciones usan los sistemas de análisis de opinión para extraer información que se encuentra en texto no organizado, no estructurado, y que proviene de fuentes de internet como

canales de redes sociales, foros, chats web, comentarios, correos electrónicos, publicaciones de blog o tickets de soporte. Por tanto, los algoritmos son una manera de reemplazar el procesamiento manual de datos, pues utilizan métodos basados en reglas, ya sean híbridos o automáticos. Los sistemas basados en reglas, para realizar un análisis de sentimientos de manera exitosa, se fundamentan en reglas léxicas predefinidas, mientras que los sistemas automáticos utilizan técnicas de aprendizaje automático. El método híbrido combina ambos enfoques.

Además de identificar sentimientos, la minería de opiniones determina la polaridad, esto es, qué tanta positividad o negatividad se manifiesta en el texto. Asimismo, el análisis de sentimientos se puede aplicar a diversos medios escritos, como documentos, párrafos, oraciones, noticias.

2.3.1 Tipos de análisis de sentimiento:

1. El análisis de sentimiento detallado proporciona un nivel más preciso de polaridad al dividirlo en más categorías, generalmente de muy positivo a muy negativo. Esto puede considerarse el equivalente de opinión de las calificaciones en una escala de 5 estrellas.
2. La detección de emociones identifica emociones específicas en lugar de positividad y negatividad. Los ejemplos podrían incluir felicidad, frustración, conmoción, ira y tristeza.
3. El análisis basado en la intención reconoce las acciones detrás de un texto además de la opinión. Por ejemplo, un usuario podría dejar un comentario en línea que exprese su intención de dejar un servicio por la demora en atención de una sucursal en específico, con esta información se podría tomar la decisión de comunicarse con el usuario y brindarle opciones.
4. El análisis basado en aspectos recopila el componente específico que se menciona positiva o negativamente. Un cliente podría dejar una reseña sobre un producto diciendo que la duración de las llantas fue demasiado corta. Luego, el sistema devolverá que el sentimiento negativo no se trata del producto en su conjunto, sino de la duración las llantas.

2.3.2 Aplicaciones del análisis de sentimiento

Las organizaciones pueden utilizar las herramientas de análisis de opinión para una variedad de aplicaciones, que incluyen:

- Identificar el conocimiento de la marca, la reputación y la popularidad en un momento específico o a lo largo del tiempo.
- Seguimiento de la recepción de los consumidores de nuevos productos o características.
- Evaluación del éxito de una campaña de marketing.
- Identificar el público objetivo o la demografía.
- Recopilación de comentarios de usuarios de sitios web, redes sociales o formularios en línea.
- Realización de estudios de mercado.
- Categorización de las solicitudes de atención al cliente.

2.3.3 Desafíos con el análisis de sentimientos

Los desafíos asociados con el análisis de sentimientos generalmente giran en torno a imprecisiones en los modelos de entrenamiento. La objetividad o los comentarios con un sentimiento neutral tienden a plantear un problema para los sistemas y, a menudo, se identifican erróneamente. Si un cliente recibió un artículo de color incorrecto y envió un comentario "El producto era negro", esto se identificaría como neutral cuando en realidad debería ser negativo.

El sentimiento también puede ser un desafío para identificar cuando los sistemas no pueden entender el contexto o el tono. Las respuestas a encuestas o preguntas de encuestas como nada o todo son difíciles de categorizar cuando no se proporciona el contexto, ya que podrían etiquetarse como positivas o negativas según la pregunta. De manera similar, la ironía y el sarcasmo a menudo no se pueden entrenar explícitamente y conducen a sentimientos etiquetados falsamente.

Los algoritmos también tienen problemas cuando se encuentran con emojis e información irrelevante. Se debe prestar especial atención a los modelos de entrenamiento con emojis y datos neutrales para no marcar incorrectamente los textos.

Por último, las personas pueden ser contradictorias en sus declaraciones. La mayoría de las reseñas tendrán comentarios tanto positivos como negativos, lo cual es manejable analizando las oraciones una a la vez. Sin embargo, cuanto más informal sea el medio, es más probable que las personas combinen diferentes opiniones en la misma oración y más difícil será para una computadora analizarlas.

2.3.4 Computación Emotiva

Tal como señala Baldassarri Santalucía, S. (2016) “Actualmente, la Computación Afectiva es un área de investigación emergente cuyo objetivo es el desarrollo de dispositivos y sistemas capaces de reconocer, interpretar, procesar y/o simular las emociones humanas para mejorar la interacción entre el usuario y la computadora.” Según la autora, estos sistemas deben cumplir tres requisitos fundamentales. El primero es realizar mediciones en las facciones de la cara, en la voz, el cuerpo, etc. para determinar con certeza el estado mental referente a una emoción el cliente. En segundo lugar, debe ser capaz de someter esta información a un cierto procesamiento que consiste en organizarla en una clasificación, que permita el aprendizaje a través de algoritmos. Finalmente, debe poder generar las respuestas oportunas a las emociones correspondientes, que se expresen a través de colores, sonidos, robots, personajes virtuales que expresen emociones por medio de gestos, caras, etc.

Con los avances de la tecnología en un futuro se conseguirán grandes resultados fusionando el análisis de sentimientos con la computación afectiva

2.3.5 Nube de Palabras:

Las nubes de palabras son una técnica de visualización de información, donde el tamaño de las palabras está correlacionado con la frecuencia de estas en el texto que se evaluado. Es decir, las palabras que tengan más repeticiones aparecerán con un mayor tamaño, captando así la atención a primera vista.

Los autores Viégas, F. y Wattenberg, M. (2008) indican que

Las Nubes de Palabras pueden ser utilizadas para fines analíticos, comunicar patrones de textos y presentar retratos de personas, por ejemplo. Esta Técnica de Visualización es flexible, pudiendo presentar grandes cantidades de textos, teniendo rápidamente una visión general de los términos más relevantes. La creciente demanda de las Nubes de Palabras indica la necesidad del usuario de visualizar una importante clase de información, información de tipo texto.

3 Pre-procesamiento de Datos

3.1 Extracción de los datos

3.1.1 Creación de una Cuenta de Desarrollador

Para iniciar con el proceso de extracción de los datos, se requiere que crear una Cuenta de Desarrollador en la plataforma Twitter, a través de los siguientes pasos:

1. Ingresar a <https://developer.twitter.com/en/portal/dashboard>, la red social solicitará el usuario de Twitter con al que va a asociar la cuenta de Desarrollador.

Twitter requerirá que responda una serie de preguntas, de cómo se considera usar la información para la cual estamos solicitando el acceso.

En el caso de que las respuestas no hayan sido claras, Twitter solicitará enviar nuevamente las justificaciones requeridas.

2. Una vez creada la cuenta de Desarrollador, adicional, Twitter asignará acceso de Elevate, con este acceso se procede a crear la aplicación.

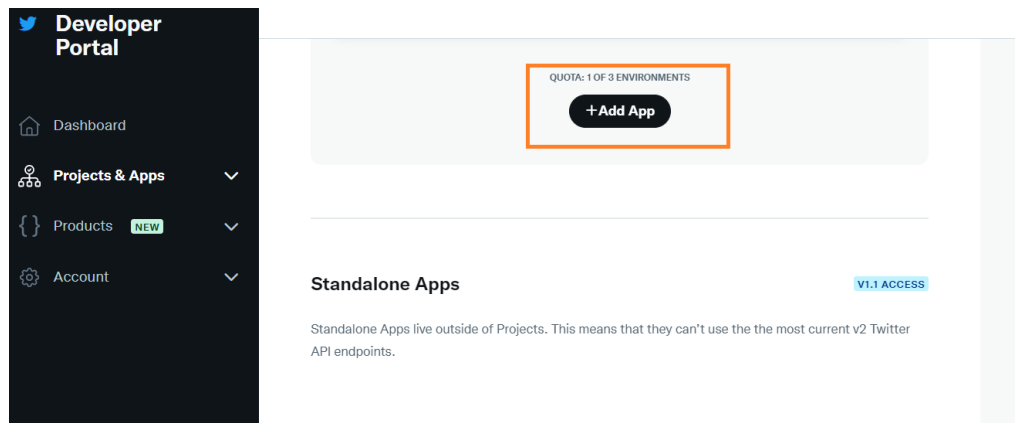


Figura 1: Crear aplicación con acceso Elevate

3. Se debe generar conjunto de claves, con las cuales se conectará al API de Twitter

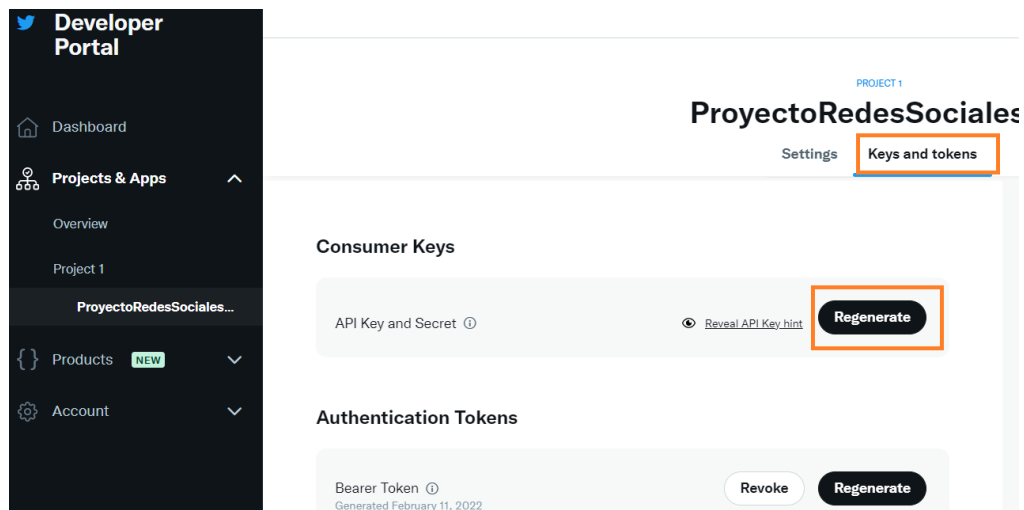


Figura 2: Generar Keys

Este nivel de acceso permite descargar 2.000.000 de tweets por mes, pero tiene la limitante que se recuperan los mensajes de los últimos 7 días únicamente.

Para el desarrollo del presente trabajo de titulación se requiere analizar el histórico de tweets publicados a partir del año 2012, con relación a los servicios que ofrece Hoteles Decamerón, por esa razón nos vimos en la necesidad de solicitar acceso de Investigador para la cuenta.

3.1.1.1 Crear acceso de Investigador Académico en Twitter

Se debe cumplir con los siguientes requisitos:

Comprobación del perfil académico

Nombre completo tal como aparece en la documentación de su institución

Enlaces a páginas web que ayudan a establecer identidad; debe proporcionar uno o más de los siguientes puntos:

- Un enlace a su perfil en el directorio de profesores o estudiantes de su institución
- Un enlace a su perfil de Google Scholar
- Un enlace al sitio web de su grupo de investigación, laboratorio o departamento donde aparece
- Información sobre su institución académica: su nombre, país, estado y ciudad
- El nombre de su departamento, escuela o laboratorio
- Su campo académico de estudio o disciplina en esta institución
- Su función actual como académico (ya sea estudiante de posgrado, candidato a doctorado, posdoctorado, profesor, científico investigador u otro miembro de la facultad)

Detalles del proyecto de Investigación

Esta sección ayuda a comprender cómo se pretende utilizar la API de Twitter y los datos obtenidos desde Twitter. Las respuestas a estas preguntas ilustran que tiene en mente un proyecto de investigación académica claramente definido.

Esta información también se usa para garantizar que los proyectos se adhieran al Acuerdo y política de desarrolladores de Twitter , incluidos términos para uso no comercial, y que se está protegiendo la privacidad y seguridad de quienes usan Twitter.

Twitter solicitará que describa en idioma inglés, cómo compartirá los resultados de su investigación (debe incluir herramientas, datos y/u otros recursos que espera construir y compartir).

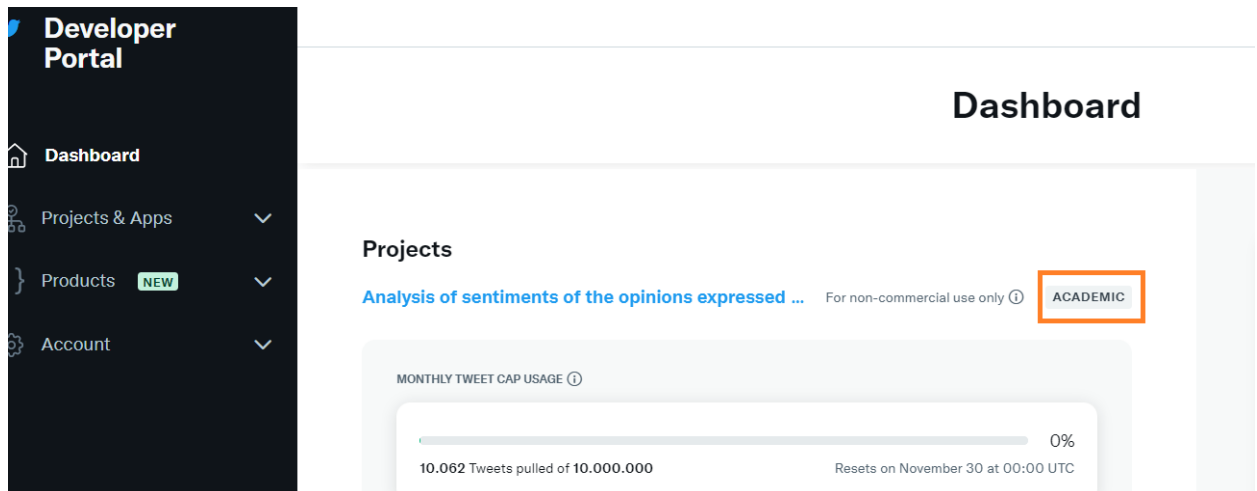


Figura 3: Accesos de Investigador Académico para Twitter

Los accesos de investigador permiten descargar 10 000 000 de tweets por mes sin límite de fecha histórica, sin ningún costo.

Access

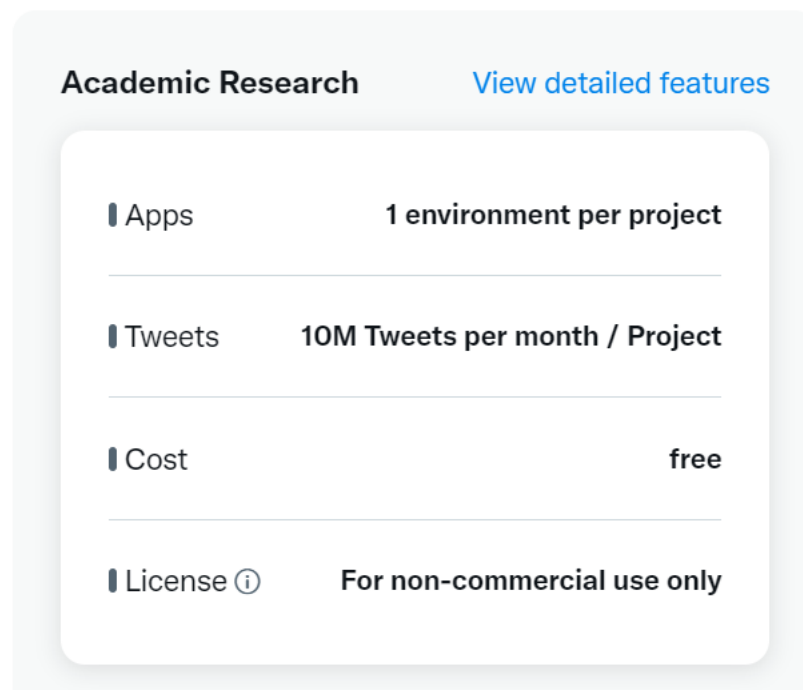


Figura 4: Accesos Cuenta de Investigador Académico

Para el desarrollo del presente proyecto se extrajo datos de tweets desde el primero de enero del 2012 hasta el 19 de noviembre del año 2022.

3.1.1.2 Conexión al API de Twitter

La conexión se realizó a través de la librería Tweepy:

```
!pip install tweepy
import tweepy
```

Posterior se procede a la autenticación con las claves generadas para la aplicación, esto si se está usando una cuenta de twitter con acceso Elevate.

```
consumer_key, consumer_secret, access_token, access_token_secret =
// Asignar las credenciales generadas para la aplicación a desarrollarse.
auth = tweepy.OAuthHandler(consumer_key,consumer_secret)
auth.set_access_token(access_token,access_token_secret)
api = tweepy.API(auth, wait_on_rate_limit=True)
```

Para conectarnos con una cuenta de Twitter con acceso de Investigador Académico, se requiere instalar la librería Twarc.

```
!pip install twarc
from twarc.client2 import Twarc2
from twarc.expansions import ensure_flattened
from twarc import Twarc2, expansions
from twarc.expansions import flatten
```

Para la autenticación usamos, el token generado desde el portal de desarrollador

Authentication Tokens

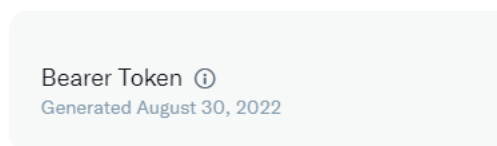


Figura 5: Código para conectar a la API de Twitter

```
T= Twarc2(bearer_token="xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx")
La búsqueda de tweets se la realiza a través de las siguientes instrucciones:
start_time = datetime.datetime(2012, 1, 1, 0, 0, 0, 0, datetime.timezone.utc)
end_time = datetime.datetime(2012, 12, 31, 0, 0, 0, 0, datetime.timezone.utc)
search_results = t.search_all(query="multivacaciones decameron lang:es -is:retweet",
start_time=start_time, end_time=end_time, max_results=100)
```

3.1.1.3 Almacenamiento de Datos

Se realizó el almacenamiento de datos en la base de Datos Postgres, mediante la librería psycopg2.

```

connection = psycopg2.connect(host="127.0.0.1",
                                database="xxxxx",
                                user="xxxxx",
                                password="xxxxxxx",
                                port="5432")

cur = connection.cursor()

cur.execute("""Insert into ...")

```

3.1.1.4 Limpieza de Datos

La limpieza de datos consiste en eliminar o reemplazar palabras que no son relevantes para el análisis, como, por ejemplo: se reemplazar letras con tildes por su igual, pero sin la tilde, transformar todo el texto a minúsculas, eliminar los emojis, eliminar las marcas de RT, eliminar números, eliminar stop words, La limpieza se la realizará a través de las siguientes instrucciones:

Eliminar Emojis: Consiste en reemplazar todos los símbolos que representen un emoticón o emoji por un carácter en blanco.

```
emoji.get_emoji_regexp().sub(r'', text.decode('utf8'))
```

Eliminar Números: Se reemplaza los números por un carácter en blanco

```
re.sub(r'[0-9]+', '')
```

Eliminar marcas de Retweet (RT): Elimina los caracteres “RT”

```
re.sub('RT @\w+: ', '', x)
```

Eliminar Stop Words: Implica eliminar artículos, pronombres y preposiciones de uso común como "y", "el" o "eso". Durante este proceso, algunas palabras muy comunes que parecen inadecuadas para los propósitos de la PNL se filtran y excluyen del texto procesado, lo que excluye términos comunes y frecuentes que no contienen información sobre la PNL.

Se Ignora las palabras que se pueden omitir de forma segura buscando en una lista predefinida de palabras clave, liberando espacio en la base de datos y reduciendo el tiempo de procesamiento.

No hay una lista general de palabras vacías. Se pueden preseleccionar o crear desde cero. Un enfoque posible es comenzar con palabras vacías predefinidas y agregar las palabras a la lista más tarde. Sin embargo, últimamente parece haber una tendencia general de pasar de usar una gran lista de palabras clave estándar a no usar ninguna lista de Stop words. El problema es que eliminar las stop words puede eliminar información importante y cambiar el contexto de una oración determinada. Por ejemplo, si estamos haciendo un análisis de

sentimientos, podemos restablecer nuestro algoritmo si eliminamos una Stop Word como "no". En estas condiciones, se puede elegir una lista mínima de palabras clave y agregar términos adicionales según sus objetivos específicos.

La librería nltk permite descargar un listado de stop Word predefinido.

```
nltk.corpus.stopwords.words('spanish')
```

Stemmer: Es un método que permite reducir una palabra a una raíz o morfema. Esto se hace para analizar las variaciones de las palabras en general. Por ejemplo, supongamos que tenemos un conjunto de palabras: modernizar, moderno, modernización, modernidad, moderna. Solo queremos que la raíz de estas palabras coincida en este caso obtendremos moderno, moderno, moderno, moderno, moderno, para el conjunto de palabras antes mencionado.

Actualmente, existen varios algoritmos que realizan un truncado del final o el principio de una palabra, teniendo en cuenta la presencia de prefijos y sufijos comunes. En muchos casos, este proceso suele tener éxito, pero no siempre.

Python proporciona el algoritmo de Porter. Según Panessi y Borgignon (2001), este algoritmo funciona bajo las siguientes normas:

- Un identificador de regla
- El sufijo a identificar
- El texto por el cual debe ser reemplazado al encontrar el sufijo
- El tamaño del sufijo
- El tamaño del texto de reemplazo
- El tamaño mínimo que debe tener la raíz resultante luego de aplicar la regla (esto es a los efectos de no procesar palabras demasiado pequeñas)
- Una función de validación (una función que verifica si se debe aplicar la regla una vez encontrado el sufijo)

La librería nltk proporciona el método PorterStemmer, para la aplicación de Stemmer

```
nltk.PorterStemmer()
```

3.1.1.5 Tokenización:

Tokenizar consiste en segmentar o dividir secuencias de texto en oraciones y palabras. Básicamente es el trabajo de cortar un fragmento de texto en piezas (tokens) a medida que se eliminan algunos caracteres como la puntuación. El proceso de

tokenización puede ser especialmente difícil para los dominios de texto biomédico que contienen muchos guiones, paréntesis y otros signos de puntuación.

Instrucción Python.

TextBlob(text_rc).words

4 Procesamiento

Como se había mencionado en secciones anteriores para realizar el estudio de Mercado se obtuvo los tweets desde el año 2012 bajo las palabras clave: Decameron Mompiche, Decameron Punta Centinela, Multivacaciones.

4.1 Análisis para Decameron Mompiche

Durante la extracción de datos asociada con Decameron Mompiche se obtuvo un total de 3664 tweets, distribuidos desde el año 2012 hasta noviembre del año 2022.

Nro.	Año	Nro. Tweets
1	2012	639
2	2013	795
3	2014	623
4	2015	457
5	2016	250
6	2017	216
7	2018	185
8	2019	238
9	2020	124
10	2021	76
11	2022	61
Total		3664

Tabla 1: Número de tweets por año para la búsqueda “Decameron Mompiche”

Una vez realizado el proceso de análisis de sentimientos sobre los tweets correspondientes al Hotel Decameron Mompiche, se pudo evidenciar que el modelo de clasificación no funciona de manera óptima para datos en español, por ese motivo se realizó la traducción del texto de los tweets al idioma inglés. En este idioma la clasificación de los tweets brindó mejores resultados.

Durante el desarrollo del presente trabajo se utilizará el texto u opiniones que se tradujo al idioma inglés para el análisis de sentimientos y para las nubes de datos se empleará las opiniones en idioma español.

En las Figuras 6 y 7 se puede apreciar diferencias en la clasificación de sentimientos para los tweets en español y los traducidos al inglés.

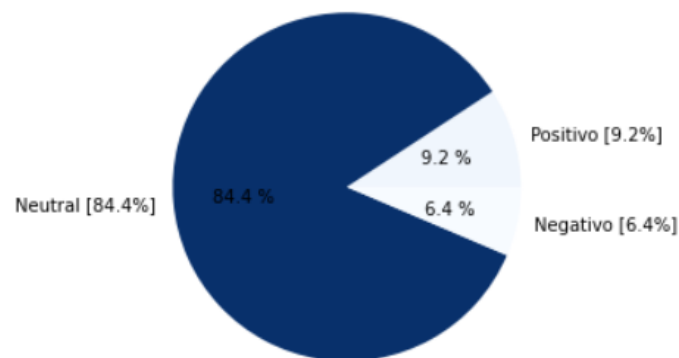


Figura 6: Análisis de sentimientos en español para la búsqueda “Decameron Mompiche”

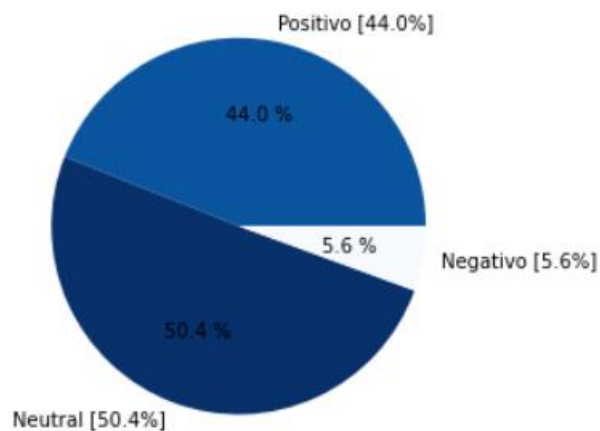


Figura 7: Análisis de sentimientos en inglés para la búsqueda “Decameron Mompiche”

En las figuras 8 y 9, se muestra cómo se distribuye los resultados del análisis de sentimientos por años, desde el año 2012 hasta el año 2022 de los tweets recuperados con las palabras claves “Decameron Mompiche” en idioma inglés.

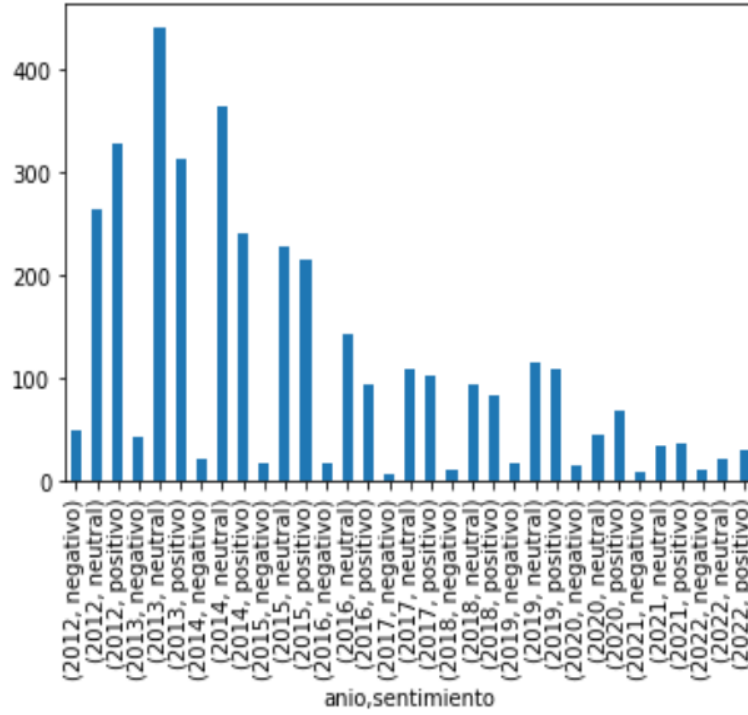


Figura 8: Sentimientos por año para “Decameron Mompiche” (a)

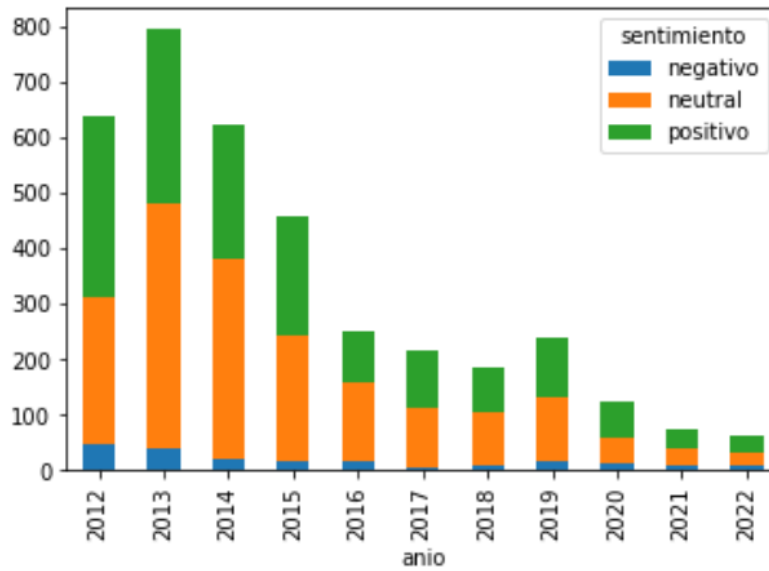


Figura 9: Sentimientos por año “Decameron Mompiche” (b)

Para tener un panorama global sobre las palabras que son más relevantes en las opiniones publicadas, se realizó una nube de palabras global que abarca todas las opiniones emitidas desde

el año 2012 hasta el 2022. Además, se elaboró una nube de palabras para las opiniones positivas, negativas y neutras

En la figura 10, se muestra con mayor tamaño las palabras que se presentaron con más frecuencia en las opiniones de los usuarios que compartieron sus mensajes en twitter bajo las palabras clave “Decameron Mompiche”.



Figura 10: Nube de Palabras para la búsqueda “Decameron Mompiche”

Las Figuras 11,12 y 13 representan las nubes de palabras para los Tweets positivos, neutros y negativos, podemos observar que una de las palabras positivas más relevantes es “disfruta”, en cuanto a la nube de palabras negativas se destaca la palabra “pésimo.”



Figura 11: Nube de Palabras Positivas para la búsqueda “Decameron Mompiche”



Nro.	Palabras	Frecuencia
1	decameron	3762
2	mompich	3758
3	royal	1549
4	ecuador	558
5	incluido	341

6	esmeralda	339
7	punta	326
8	hotel	315
9	Vacacion	285
10	Centinela	270
Total		11503

Tabla 2: Número de palabras más usadas para “Decameron Mompiche”

Conteo de palabras más usadas

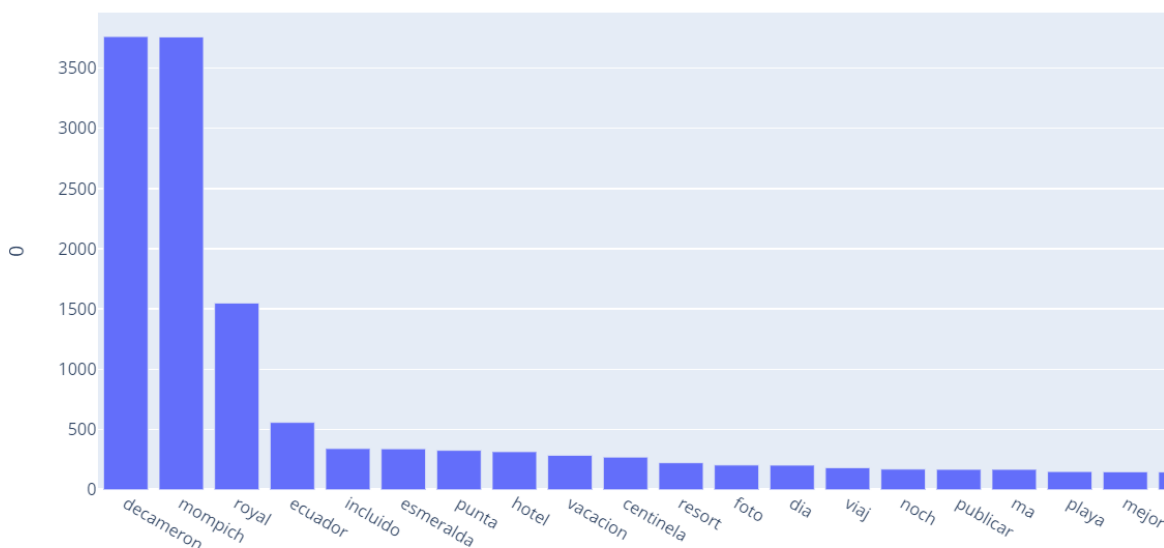


Figura 14: Conteo de palabras más usadas para “Decameron Mompiche”

Resulta de utilidad conocer los bigramas y trigramas que se mencionan con más frecuencia en los tweets de la búsqueda “Decameron Mompiche”, en la Tabla 3. y Figura 15., se indican los resultados de los bigramas encontrados. En la Tabla 4. y Figura 16. Se muestran los resultados de los trigramas identificados en el análisis.

Nro.	Bigramas	Frecuencia
1	decameron mompiche	2913
2	royal decameron	1528
3	mompiche decameron	258
4	punta centinela	234

5	mompiche ecuador	213
6	mompiche esmeraldas	210
7	publicar foto	162
8	esmeraldas ecuador	154
9	mompiche punta	149
10	mompiche royal	149
Total		5970

Tabla 3: Bigramas más usadas “Decameron Mompiche”

Conteo de Bigramas más usados

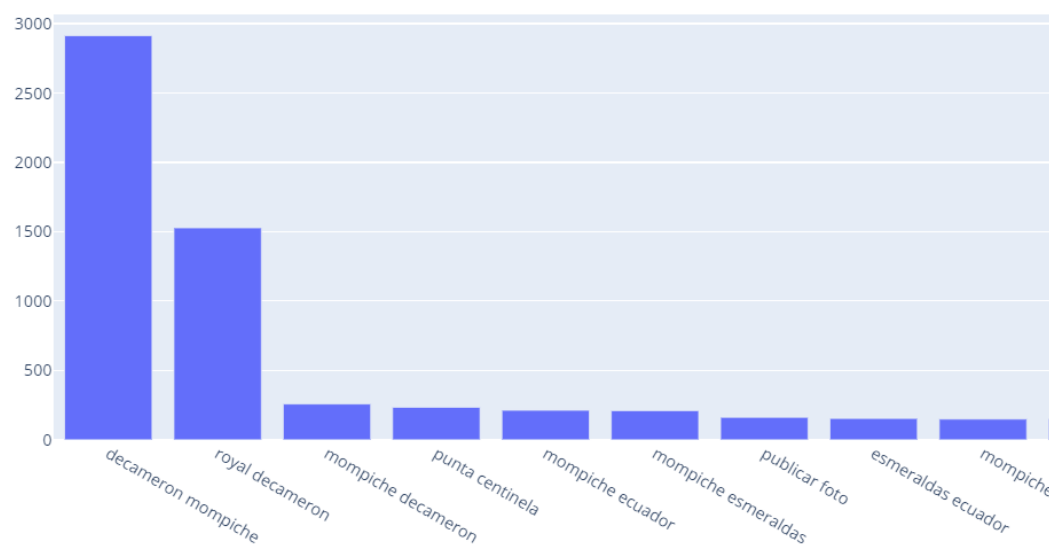


Figura 15: Conteo de Bigramas más usadas “Decameron Mompiche”

Nro.	Trigramas	Frecuencia
1	royal decameron mompiche	1363
2	decameron mompiche esmeraldas	192
3	mompiche royal decameron	148
4	decameron mompiche ecuador	135

5	mompiche punta centinela	131
6	mompiche esmeraldas ecuador	126
7	publicar foto royal	112
8	foto royal decameron	112
9	acabo publicar foto	104
10	royal decameron hotel	83
11	decameron mompiche punta	82
12	hotel amp resort	81
13	decameron hotel amp	81
14	decameron punta centinela	65
15	esmeraldas ecuador video	62
16	decameron mompiche espera	59
17	hotel decameron mompiche	58
18	acaba publicar foto	58
19	punta centinela punta	53
20	hotel royal decameron	52
Total		3157

Tabla 4: Trigramas más usados para “Decameron Mompiche”

Conteo de Trigramas más usados

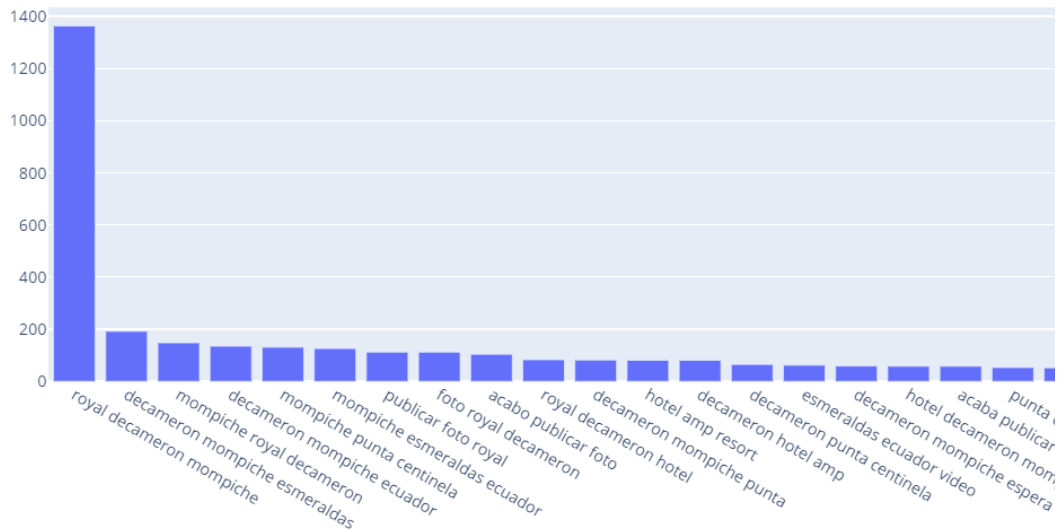


Figura 16: Conteo de Trigramas más usados para “Decameron Mompiche”

Los usuarios que emiten opiniones sobre Decameron Mompiche, utilizan distintas aplicaciones para publicar sus opiniones, en la Figura 17, se presenta la aplicación desde donde se emitieron los tweets, con su respectiva frecuencia.

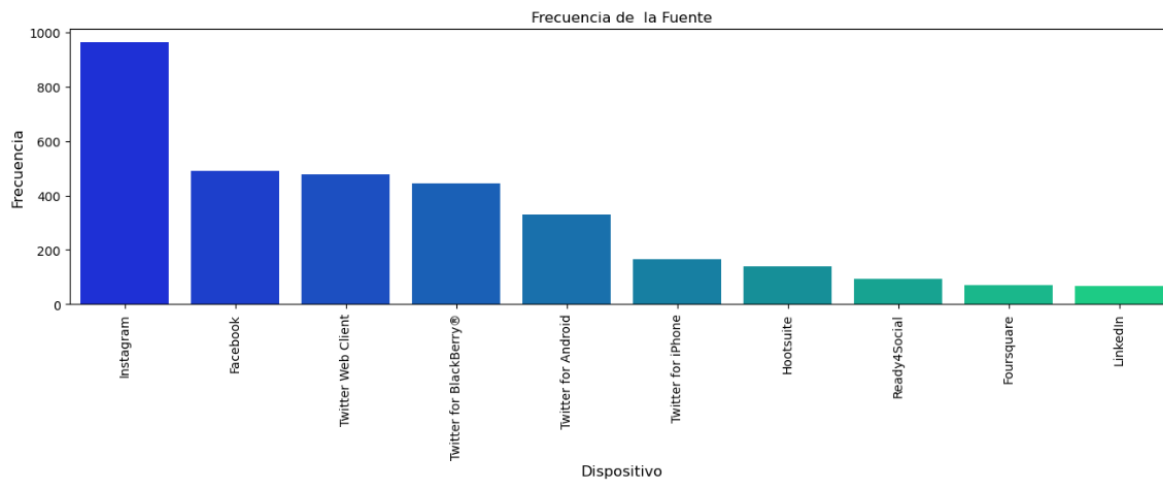


Figura 17: Frecuencia de aplicaciones desde dónde se generaron los tweets para “Decameron Mompiche”

La Figura 18. Permite visualizar desde que localidades se emitieron los tweets, esta información tiene fundamental relevancia, debido que nos permitirá identificar desde que ciudades o países existe interés por Decameron Mompiche

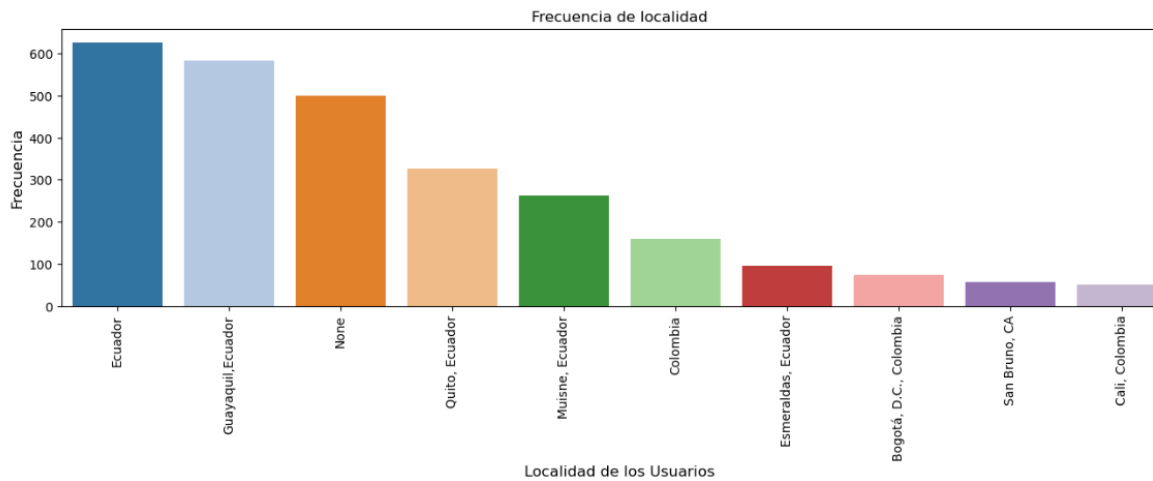


Figura 18: Frecuencia localidades desde dónde se generaron los tweets para “Decameron Mompiche”

Resulta de interés analizar a los usuarios que opinaron sobre Decameron Mompiche, En la Figura 19. Se visualiza los usuarios que tiene un mayor número de amigos, en la Figura 20. Se puede observar los usuarios que tienen mayor número de seguidores, lo que podría indicarnos el nivel de influencia que poseen estos usuarios.

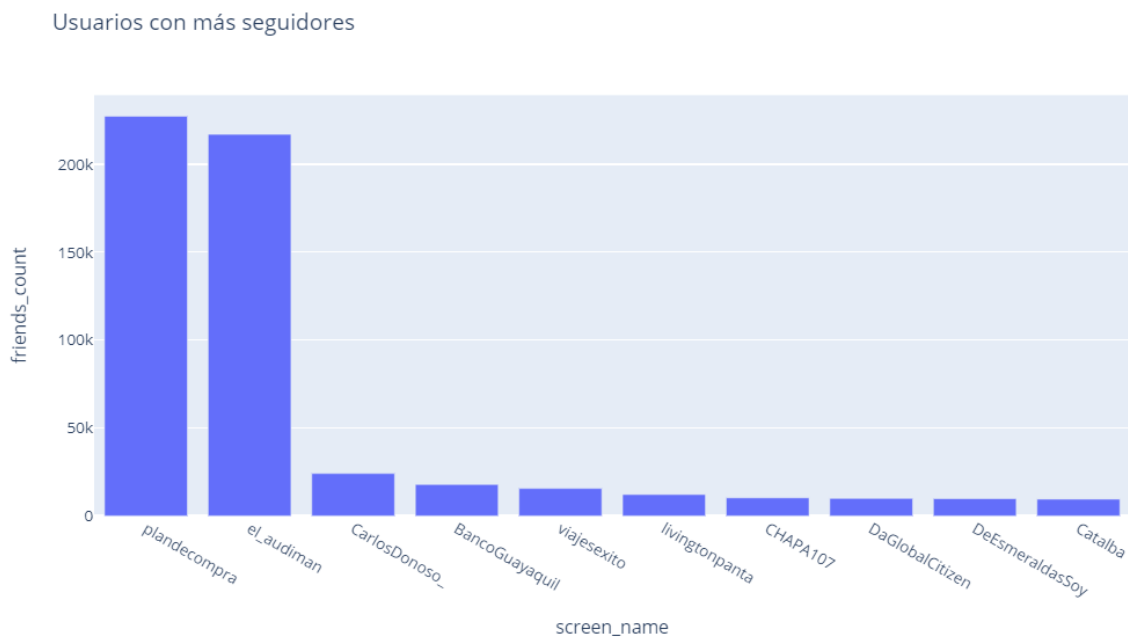


Figura 19: Usuarios con más amigos

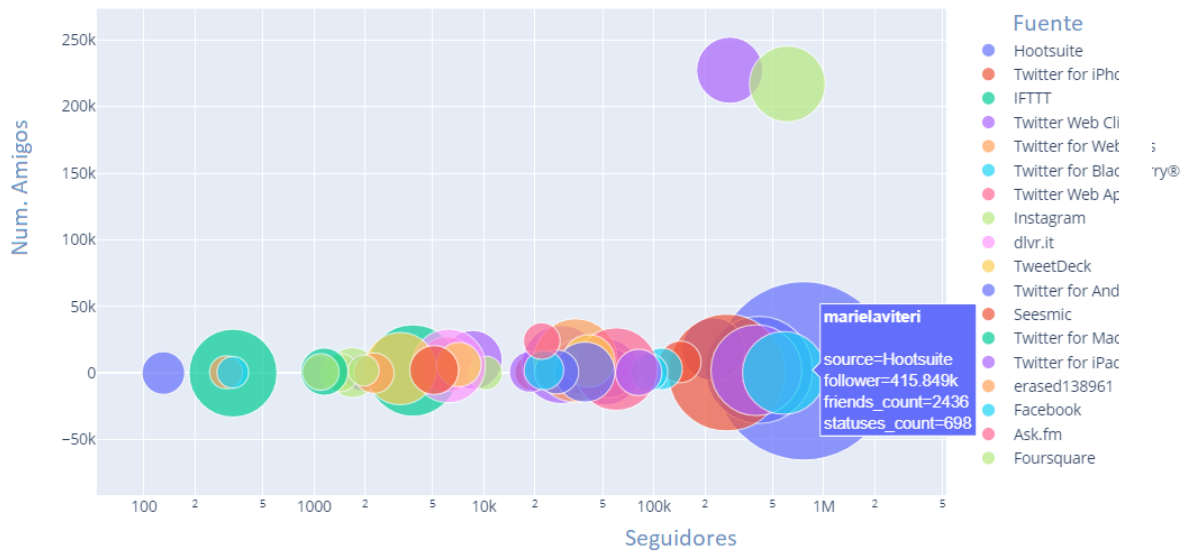


Figura 20: Descripción de Usuarios con sus seguidores y amigos para “Decameron Mompiche”

El poder conocer que apreciación tienen los viajeros sobre Decameron Mompiche por localidad, permitirá enfocar los esfuerzos, en localidades donde hay menor acogida. En las Figuras 21 y 22 se puede observar la frecuencia de tweets positivos, neutros y negativos por localidades.

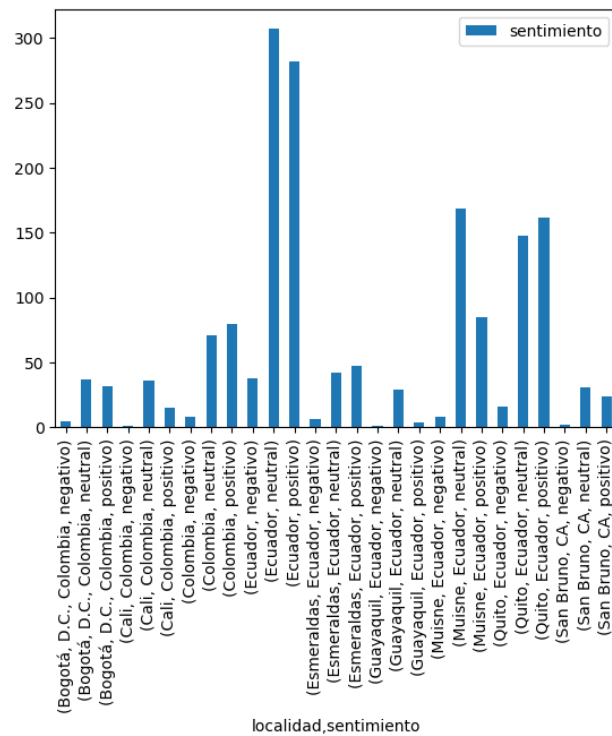


Figura 21: Sentimientos por localidad (a)

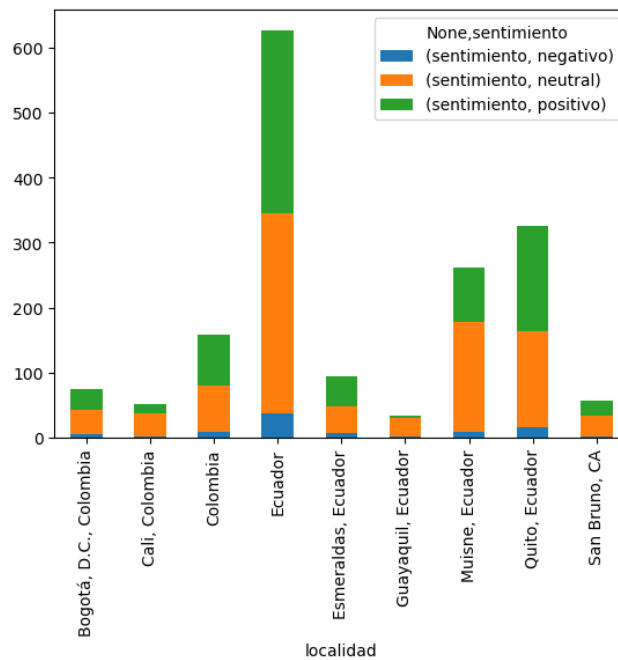


Figura 22: Sentimientos por localidad (b)

4.2 Análisis para el Hotel Decameron Punta Centinela

Durante la extracción de datos asociada con Decameron Punta Centinela: Se obtuvo un total de 3204 tweets, distribuidos desde el año 2012 hasta noviembre del año 2022.

Nro.	Año	Nro. Tweets
1	2012	374
2	2013	690
3	2014	629
4	2015	495
5	2016	258
6	2017	161
7	2018	195
8	2019	171
9	2020	109
10	2021	64
11	2022	58
Total		3204

Tabla 5: Número de tweets por año para “Decameron Punta Centinela”

Una vez realizado el proceso de análisis de sentimientos sobre los tweets correspondientes al Hotel Decameron Punta Centinela, se puede evidenciar un alto porcentaje de tweets con clasificación Neutra para el análisis realizado en español, por esta razón el análisis de sentimientos se lo realizará en idioma inglés.

En las Figuras 23 y 24 se puede apreciar diferencias en la clasificación de sentimientos para los tweets en español y los traducidos al inglés.

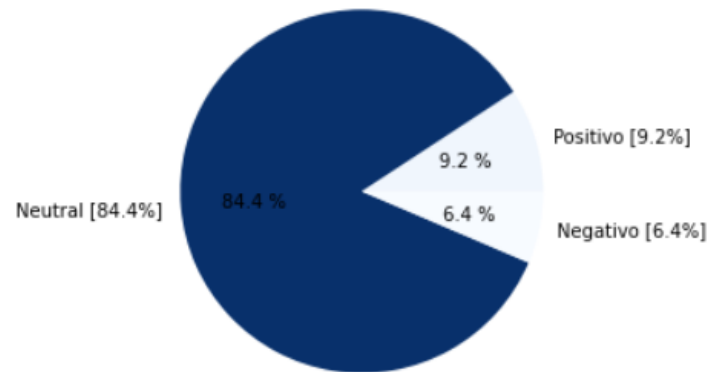


Figura 23: Sentimientos en idioma español

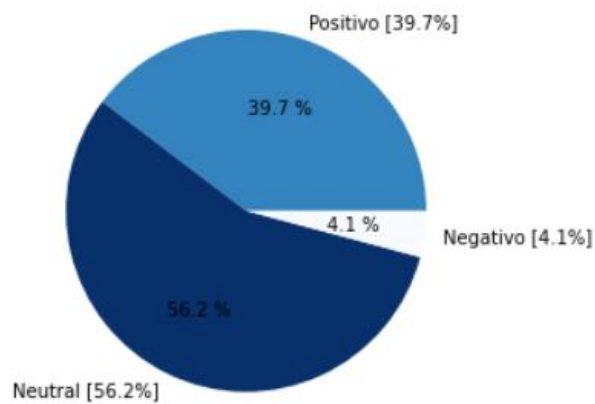


Figura 24: Sentimientos en idioma inglés

En las Figuras 25 y 26, se muestra cómo se distribuye los resultados del análisis de sentimientos por años, desde el año 2012 hasta el año 2022 de los tweets recuperados con las palabras claves “Decameron Punta Centinela”.

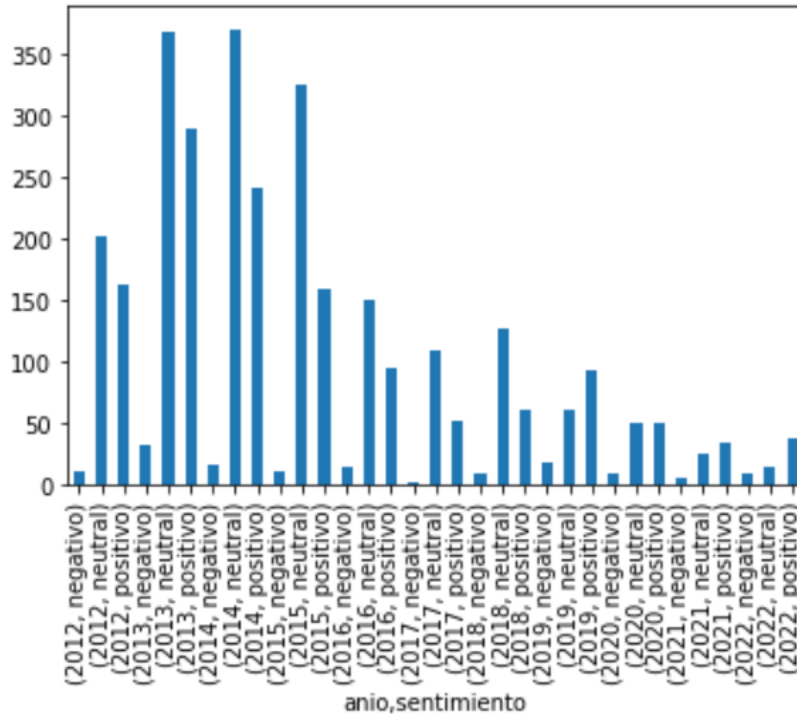


Figura 25: Sentimientos por año para “Decameron Punta Centinela” (a)

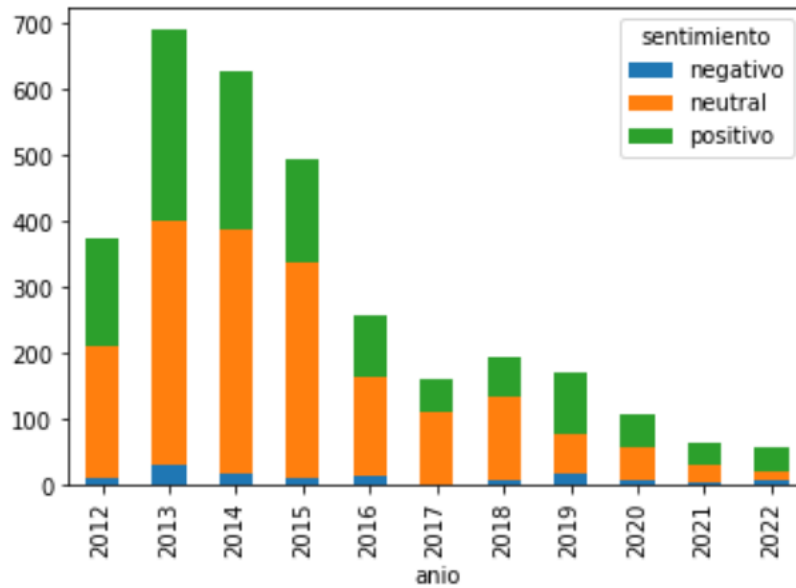


Figura 26: Sentimientos por año para “Decameron Punta Centinela” (b)

Con el propósito de analizar de forma visual las opiniones publicadas para el resultado de la búsqueda “Decameron Punta Centinela”, se realizó una nube de palabras global que abarca todas las opiniones emitidas desde el año 2012 hasta el 2022.

En la figura 27, se indican con tamaño aumentado las palabras que las opiniones de los usuarios presentan con más frecuencia.

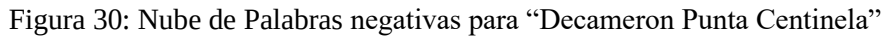
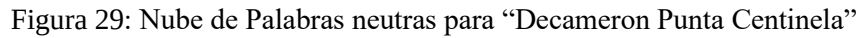


Figura 27: Nube de Palabras para “Decameron Punta Centinela”

Las Figuras 28, 29 y 30 representan las nubes de palabras para los Tweets positivos, neutros y negativos, podemos observar que una de las palabras positivas más relevantes es “lindo”, en cuanto a la nube de palabras negativas, “paga” es una de las palabras que se destaca.



Figura 28: Nube de Palabras positivas para “Decameron Punta Centinela”



Sentimiento	Longitud promedio de los tweets
negativo	123,08
neutral	77,93
positivo	103,65

Tabla 6: Longitud promedio de los tweets para “Decameron Punta Centinela”

Sentimiento	Promedio del número de palabras de los tweets
negativo	19.12
neutral	10.71
positivo	15.15

Tabla 7: Promedio de palabras de los tweets para “Decameron Punta Centinela”

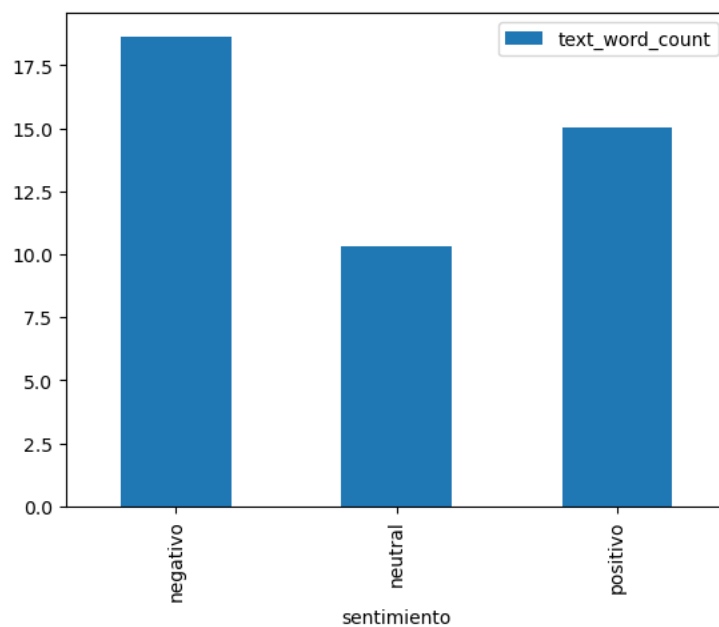


Figura 31: Promedio de Palabras en tweets para “Decameron Punta Centinela”

Una vez aplicado stemming se realizó un conteo de las palabras más usadas en los tuits. A continuación, se presenta en la Tabla 8. y la Figura 32. los resultados.

Nro.	Palabras	Frecuencia
1	punta	3313
2	decameron	3293
3	centinela	3218
4	royal	1768
5	ecuador	442
6	santa	372
7	elena	352

8	incluido	295
9	noch	259
10	mompich	237
Total		13549

Tabla 8: Palabras más nombradas para “Decameron Punta Centinela”

Conteo de palabras más usadas

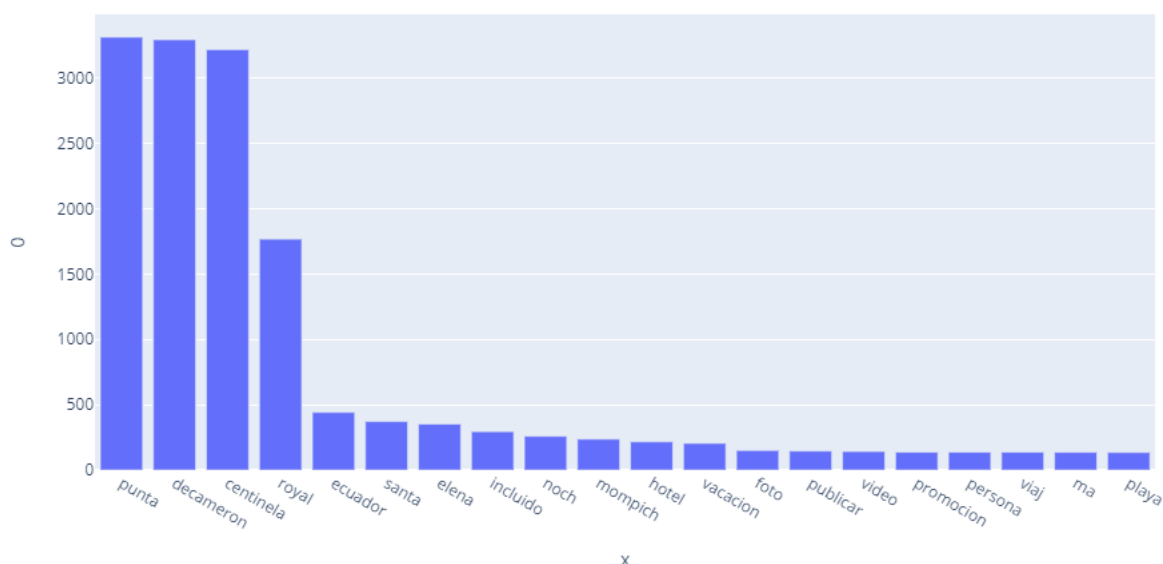


Figura 32: Conteo de Palabras más usadas para “Decameron Punta Centinela”

Resulta de utilidad conocer los bigramas y trigramas que se mencionan con más frecuencia en los tweets de la búsqueda “Decameron Mompiche”, en la Tabla 9. y Figura 33., se presentan los resultados de los bigramas encontrados. En la Tabla 10. y Figura 34. Se muestran los resultados de los trigramas identificados en el análisis.

Nro.	Bigrama	Frecuencia
1	punta centinela	3203
2	decameron punta	2677
3	royal decameron	1757
4	santa elena	338
5	centinela santa	286
6	centinela decameron	152
7	mompiche punta	145
8	publicar foto	134
9	decameron mompiche	120
10	elena ecu	116
Total		8928

Tabla 9: Número de Bigramas más usados para “Decameron Punta Centinela”

Conteo de Bigramas más usados

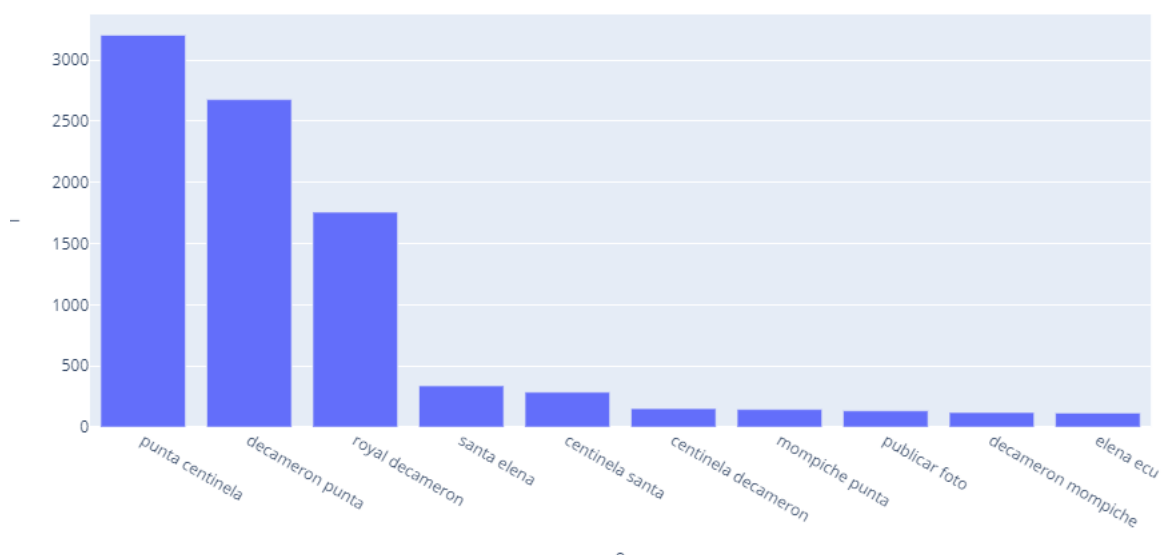


Figura 33: Conteo de Bigramas más usadas para “Decameron Punta Centinela”

Nro.	Trigrama	Frecuencia
1	decameron punta centinela	2662
2	royal decameron punta	1627
3	punta centinela santa	286
4	centinela santa elena	261
5	punta centinela decameron	151
6	mompiche punta centinela	131
7	santa elena ecu	116
8	foto royal decameron	108
9	publicar foto royal	107
10	santa elena ecuador	107
11	punta centinela ecuador	104
12	acabo publicar foto	83
13	elena ecuador video	83
14	punta centinela punta	79
15	decameron mompiche punta	78
16	experiencias teambuildingecuador royal	75
17	teambuildingecuador royal decameron	75
18	hotel royal decameron	71

19	centinela punta sal	63
20	lobby royal decameron	52
	Total	6319

Tabla 10: Número de Trigramas más usados para “Decameron Punta Centinela”

Conteo de Trigramas más usados

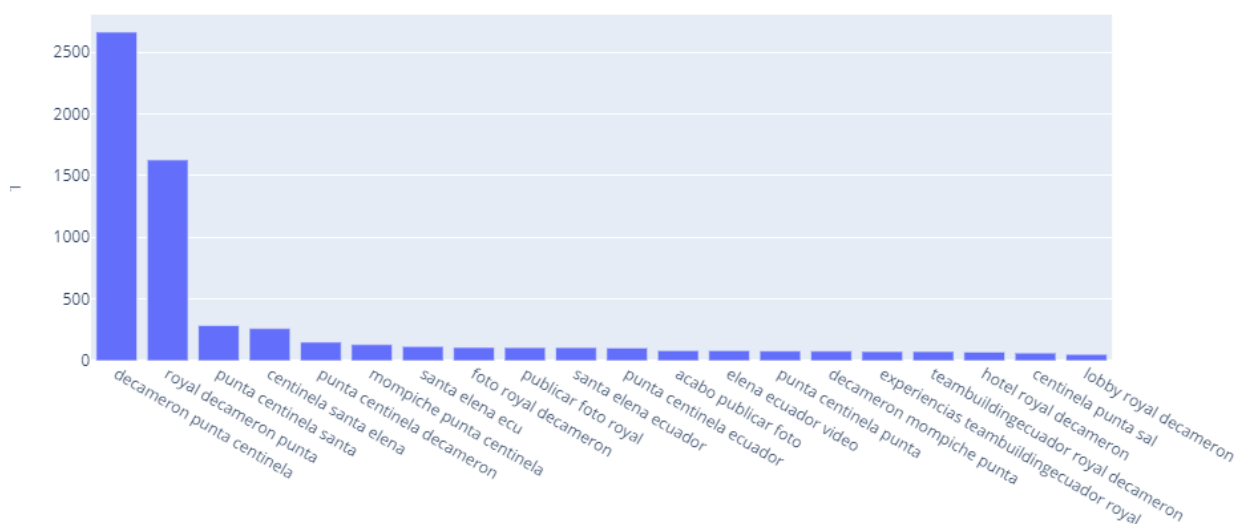


Figura 34: Número de Trigramas más usados para “Decameron Punta Centinela”

Los usuarios que emiten opiniones sobre Decameron Punta Centinela, utilizan distintas aplicaciones para publicar sus tweets, en la Figura 35, se muestra las aplicaciones desde donde se emitieron los tweets, con su respectiva frecuencia.

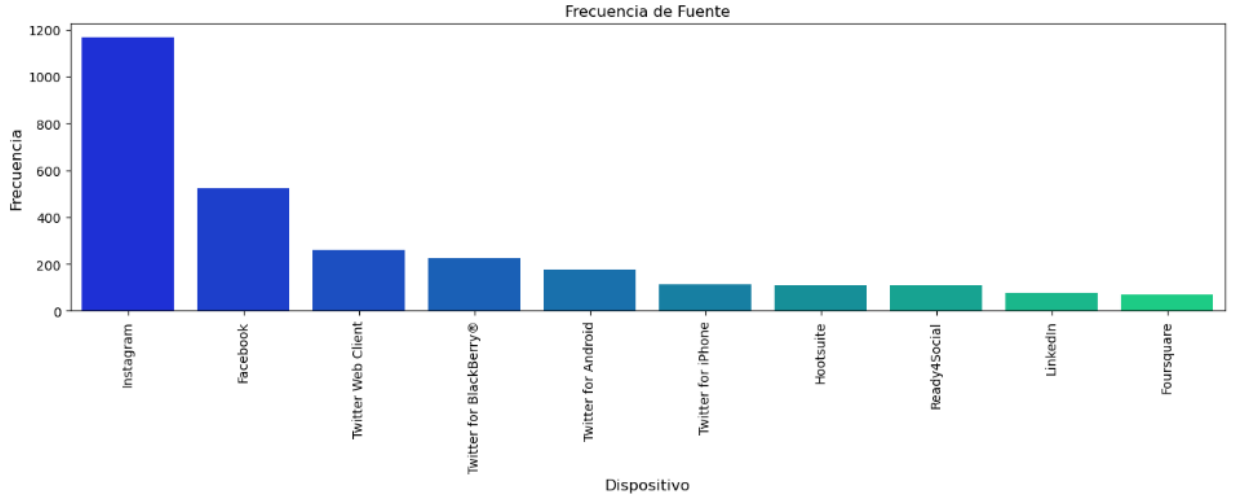


Figura 35: Frecuencia de dispositivos desde dónde se publicaron los tweets para “Decameron Punta Centinela”

La Figura 36. Permite visualizar desde que localidades se emitieron los tweets, esta información tiene fundamental relevancia, debido que nos permitirá identificar desde que ciudades o países existe interés por Decameron Punta Centinela

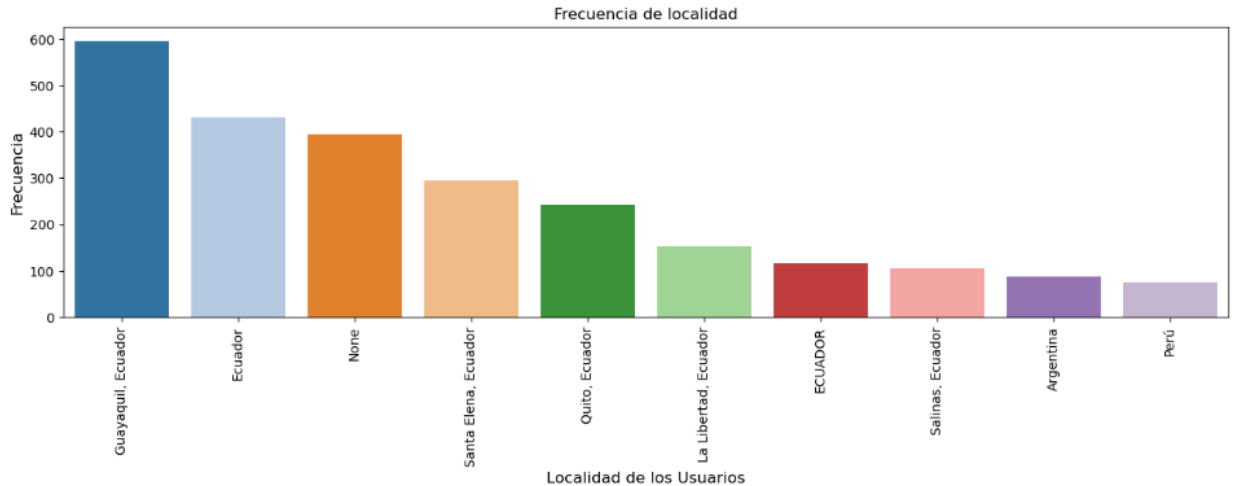


Figura 36: Frecuencia de las localidades para tweets generados para “Decameron Punta Centinela”

El poder conocer que apreciación tienen los viajeros sobre Decameron Punta Centinela por localidad, permitirá enfocar los esfuerzos, en localidades donde hay menor acogida. En las Figuras 37 y 38 se puede observar la frecuencia de tweests positivos, neutros y negativos por localidades.

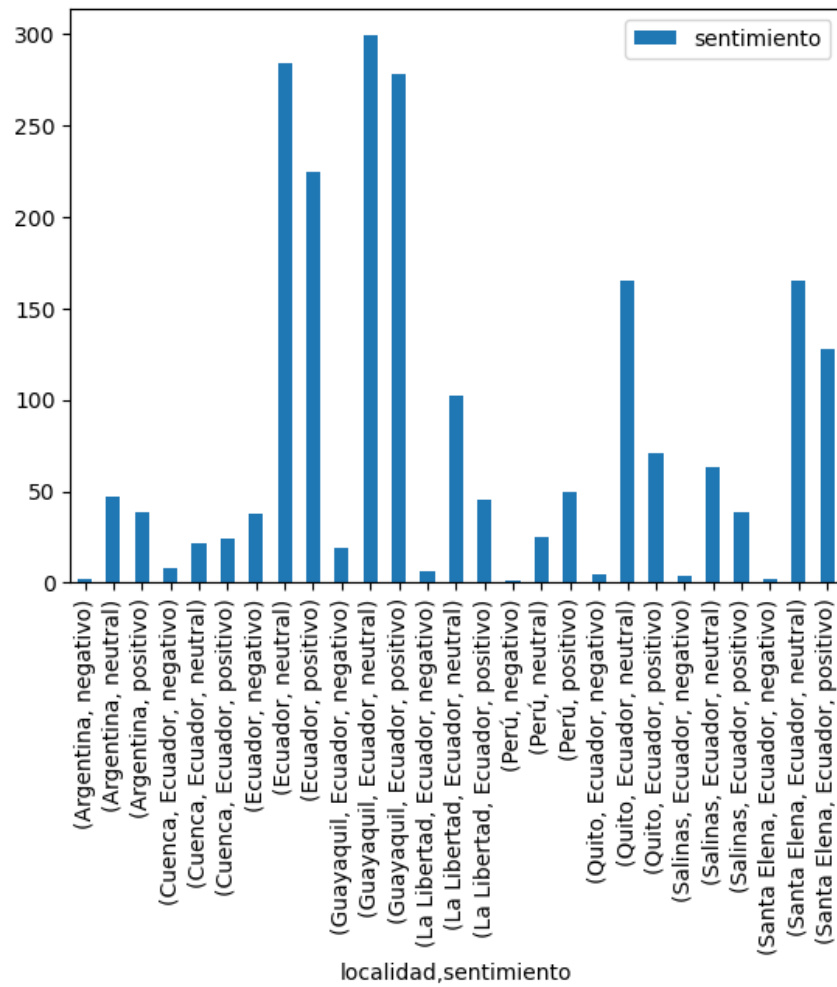


Figura 37: Sentimientos por localidades para “Decameron Punta Centinela” (a)

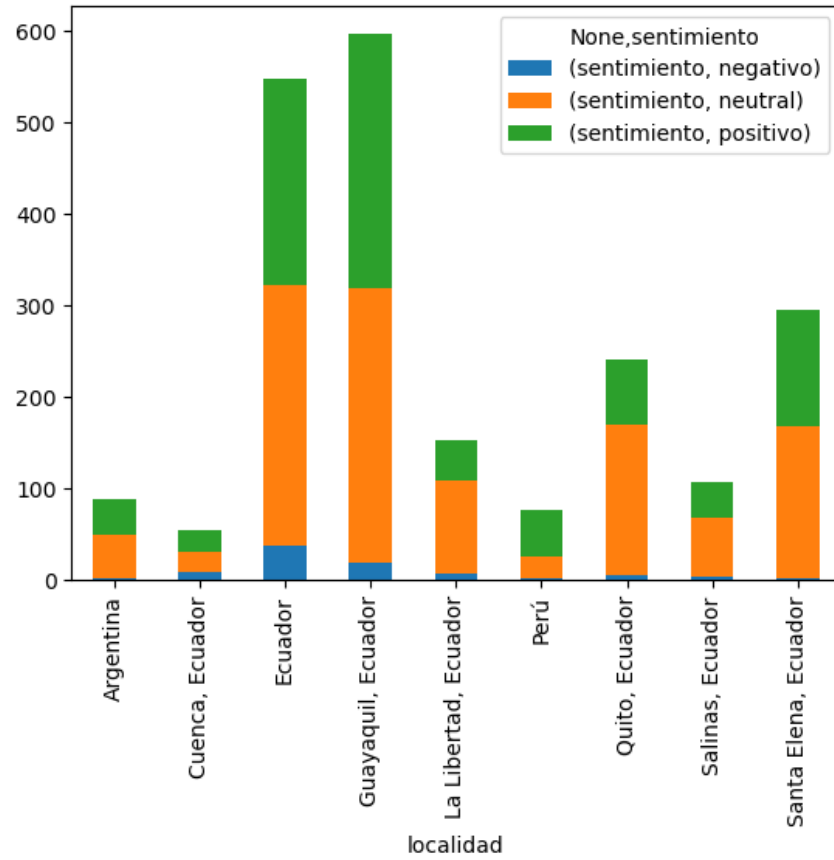


Figura 38: Sentimientos por localidades para "Decameron Punta Centinela" (b)

Resulta interesante analizar a los usuarios que opinaron sobre Decameron Punta Centinela, En la Figura 40. Se visualiza los usuarios que tiene un mayor número de amigos, en la Figura 39. Se puede observar los usuarios que tienen mayor número de seguidores, lo que podría indicarnos el nivel de influencia que poseen estos usuarios.

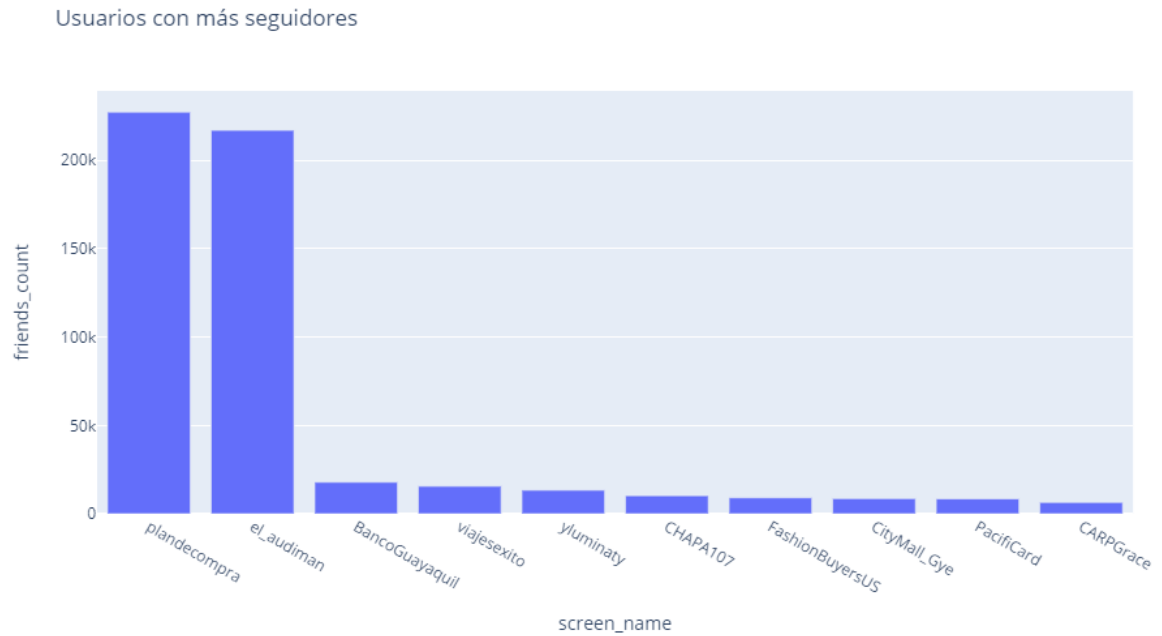


Figura 39: Usuarios con más amigos para “Decameron Punta Centinela”

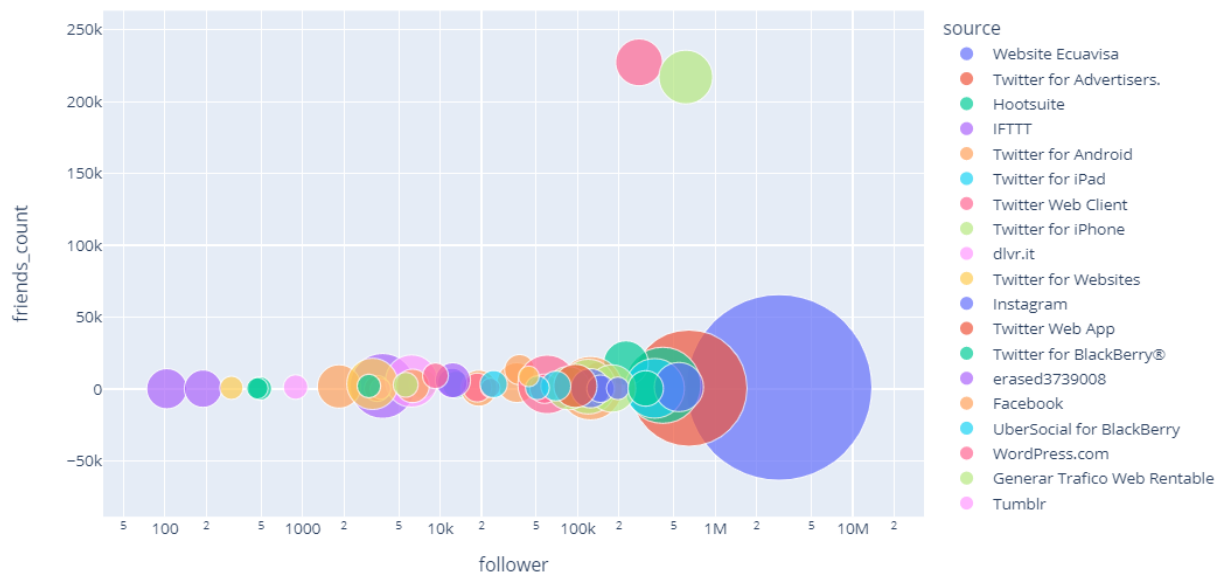


Figura 40: Descripción de Usuarios con sus seguidores y amigos para “Decameron Punta Centinela”

4.3 Análisis para Multivacaciones

Durante la extracción de datos asociada con Multivacaciones, se obtuvo un total de 581 tweets, distribuidos desde el año 2012 hasta noviembre del año 2022. Considerando que tanto Hoteles Decameron como Multivacaciones son multinacionales, para el caso de Multivacaciones se obtuvieron opiniones de diferentes países.

Nro.	Año	Nro. Tweets
1	2012	50
2	2013	82
3	2014	135
4	2015	56
5	2016	52
6	2017	30
7	2018	46
8	2019	29
9	2020	38
10	2021	34
11	2022	29
Total		581

Tabla 11: Tweets por años para “Multivacaciones”

Una vez realizado el proceso de análisis de sentimientos sobre los tweets correspondientes al Multivacaciones, se puede evidenciar un alto porcentaje de tweets con clasificación Neutra para el análisis realizado en español, por esta razón el análisis de sentimientos se lo realizará en idioma inglés.

En las Figuras 41 y 42 se puede apreciar diferencias en la clasificación de sentimientos para los tweets en español y los traducidos al inglés.

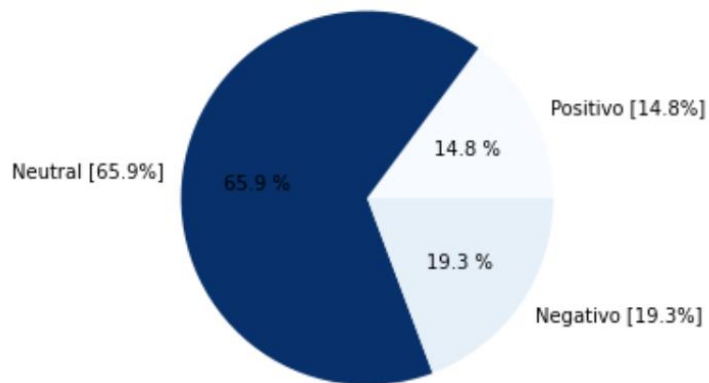


Figura 41: Análisis en español para los resultados de la búsqueda Multivacaciones

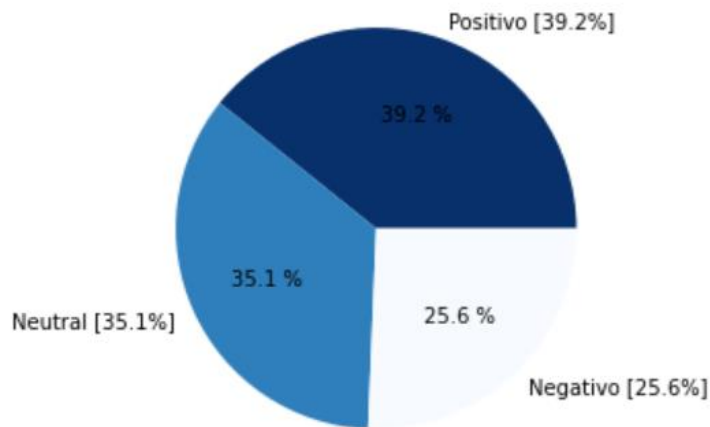


Figura 42: Análisis en inglés para los resultados de la búsqueda Multivacaciones

En las figuras 43 y 44, se muestra cómo se distribuye los resultados del análisis de sentimientos por años, desde el año 2012 hasta el año 2022 de los tweets recuperados con la palabra claves “Multivacaciones” en idioma inglés.

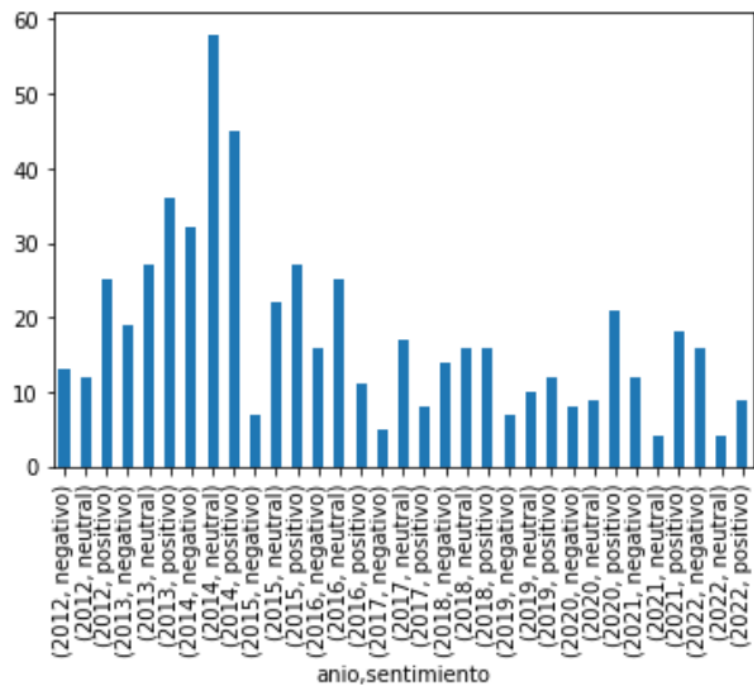


Figura 43: Análisis de sentimientos por años para “Multivacaciones” (a)

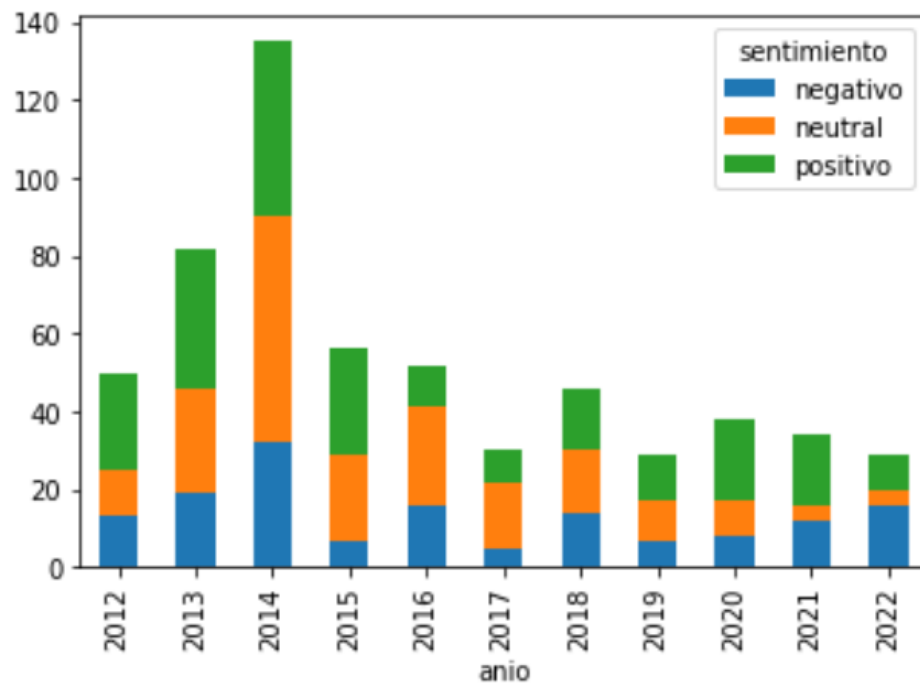


Figura 44: Análisis de sentimientos por años para “Multivacaciones” (b)

Para tener una visualización global sobre las palabras que son más relevantes en las opiniones publicadas, se realizó una nube de palabras global que abarca todas las opiniones emitidas desde



Figura 46: Nube de palabras positivas para “Multivacaciones”

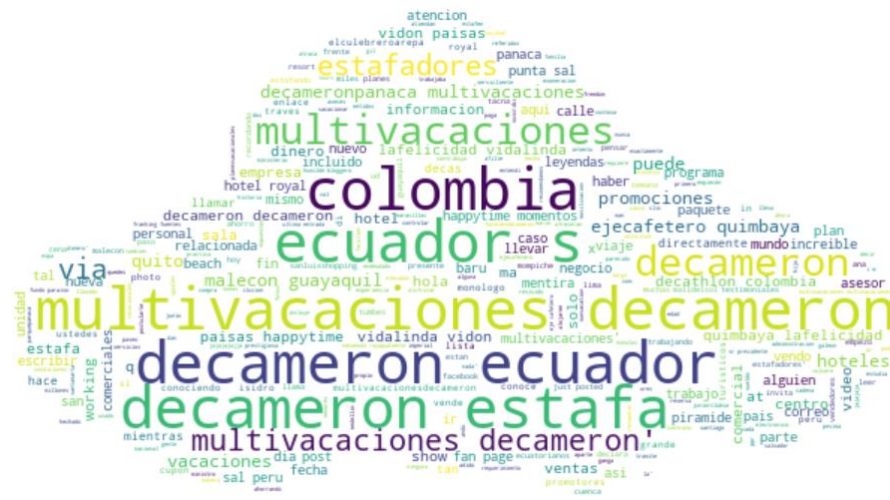


Figura 47: Nube de palabras neutras para “Multivacaciones”



Figura 48: Nube de palabras negativas para “Multivacaciones”

Durante el análisis se pudo encontrar que los tweets negativos tienen mayor longitud y mayor número de palabras en promedio, como se muestra en las Tablas 12, Tablas 13 y Figura 49 y Figura 50, respectivamente.

Sentimiento	Longitud promedio de los tweets
negativo	128.78
neutral	81.65
positivo	127.78

Tabla 12: Longitud promedio de los tweets para “Multivacaciones”

Sentimiento	Promedio del número de palabras de los tweets
negativo	19.58
neutral	10.46
positivo	19.14

Tabla 13: Promedio de palabras de los tweets para “Multivacaciones”

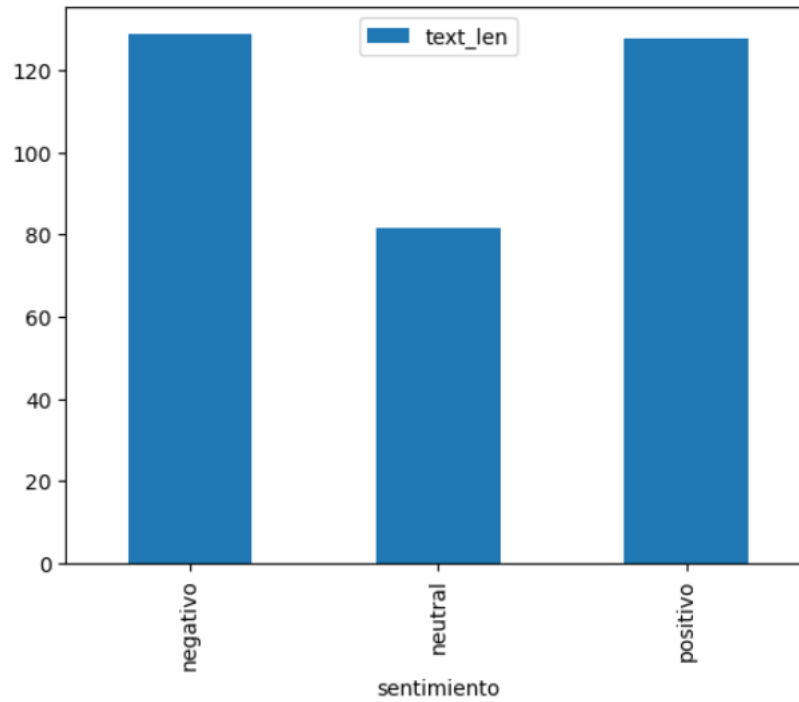


Figura 49: Longitud promedio del texto del análisis de sentimientos para “Multivacaciones”

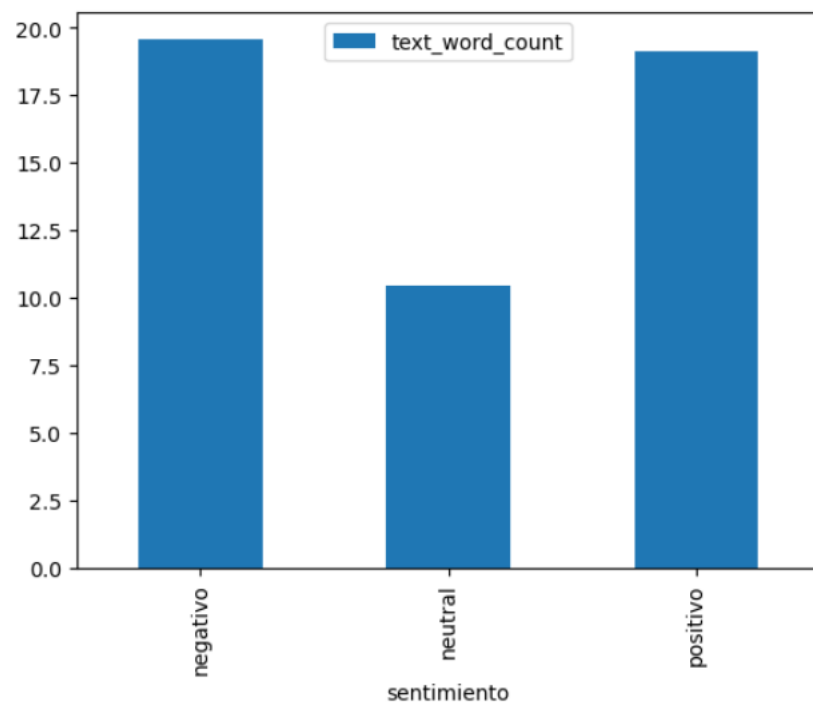


Figura 50: Promedio de palabras del análisis de sentimientos para “Multivacaciones”

Luego de aplicar stemmer se realizó un conteo de palabras más usadas en los tuits. A continuación, se presenta en la Tabla 14. y la Figura 51. los resultados.

Nro.	Palabras	Frecuencia
1	decameron	607
2	multivacacion	576
3	estafa	80
4	hotel	68
5	favor	39
6	empresa	37
7	si	32
8	q	32
9	vacacion	30
10	colombia	29
	Total	1530

Tabla 14: Conteo de Palabras más usadas para “Multivacaciones”

Conteo de palabras más usadas

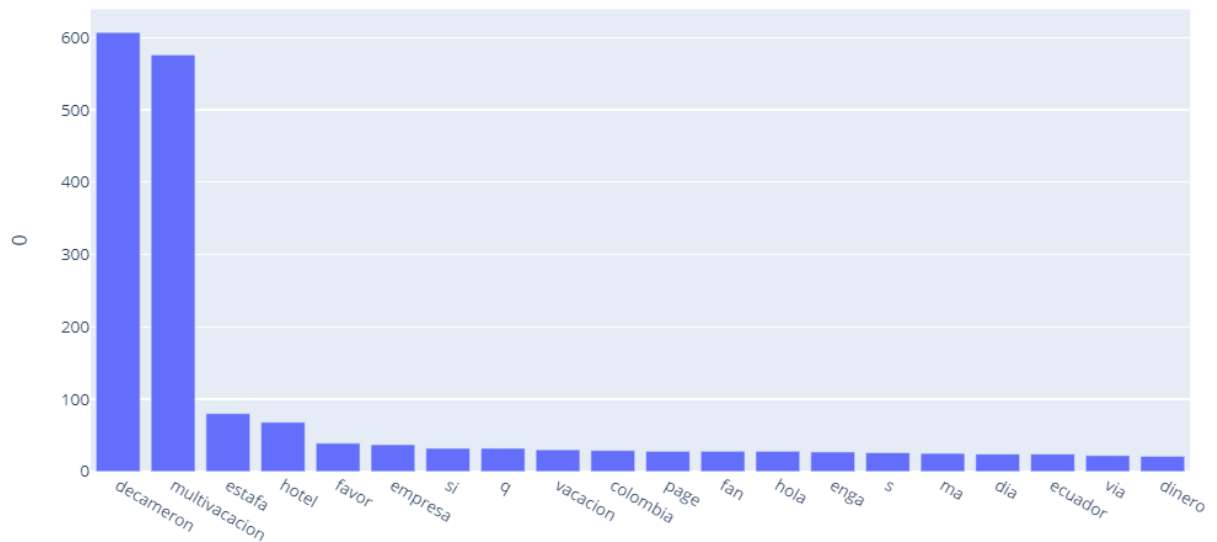


Figura 51: Frecuencia de palabras más usadas para “Multivacaciones”

Resulta beneficioso conocer los bigramas y trigramas que se mencionan con más frecuencia en los tweets de la búsqueda “Decameron Mompiche”, en la Tabla 15. y Figura 52., se presentan los resultados de los bigramas encontrados. En la Tabla 16. y Figura 53. Se muestran los resultados de los trigramas identificados en el análisis.

Nro.	Bigrama	Frecuencia
1	multivacaciones decameron	463
2	decameron estafa	33
3	fan page	28
4	decameron multivacaciones	24
5	hoteles decameron	22
6	decameron ecuador	21
7	empresa multivacaciones	16
8	multivacaciones hotel	13
9	estafadores multivacaciones	13
10	royal decameron	13
	Total	646

Tabla 15: Conteo de Bigramas más usadas para “Multivacaciones”

Conteo de Bigramas más usados

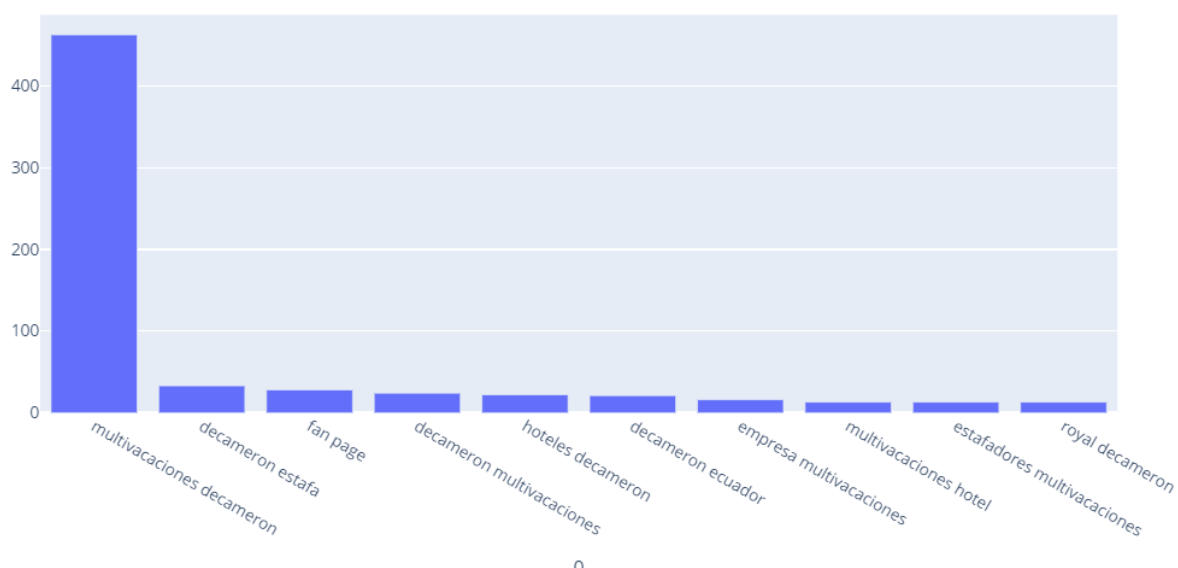


Figura 52: Frecuencia de bigramas más usados para “Multivacaciones”

Nro.	Bigramas	Frecuencia
1	multivacaciones decameron estafa	29
2	multivacaciones decameron ecuador	20
3	empresa multivacaciones decameron	15

4	estafadores multivacaciones decameron	13
5	fan page facebook	11
6	decameronpanaca multivacaciones decameron	10
7	decameron estafa malecon	10
8	estafa malecon 2000	10
9	guayaquil colombia 2014	10
10	publicidad enga osa	9
11	recomendamos escribir fan	9
12	escribir fan page	9
13	fan page unidad	9
14	page unidad negocios	9
15	unidad negocios multivacaciones	9
16	negocios multivacaciones decameron	9
17	multivacaciones decameron gt	9
18	escribir directamente correo	9
19	directamente correo electronico	9
20	correo electronico serviciente	9
Total		227

Tabla 16: Número de Trigramas más usados para “Multivacaciones”

Conteo de Trigramas más usados

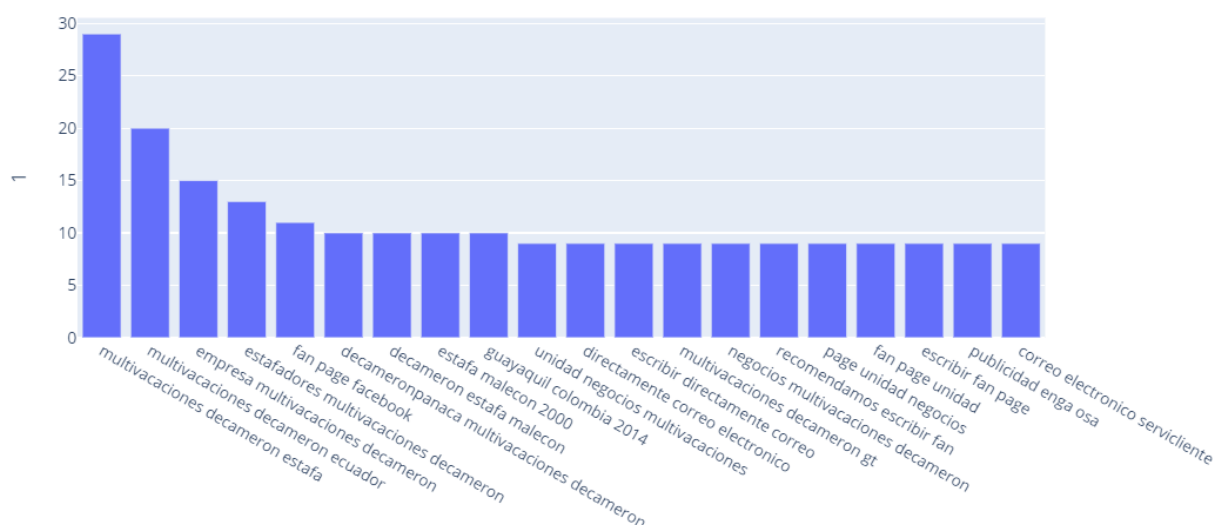


Figura 53: Frecuencia de trigramas más usados para “Multivacaciones”

Los usuarios que emiten opiniones sobre Multivacaciones, utilizan distintas aplicaciones para publicar sus opiniones, en la Figura 54, se presenta la aplicación desde donde se emitieron los tweets, con su respectiva frecuencia.

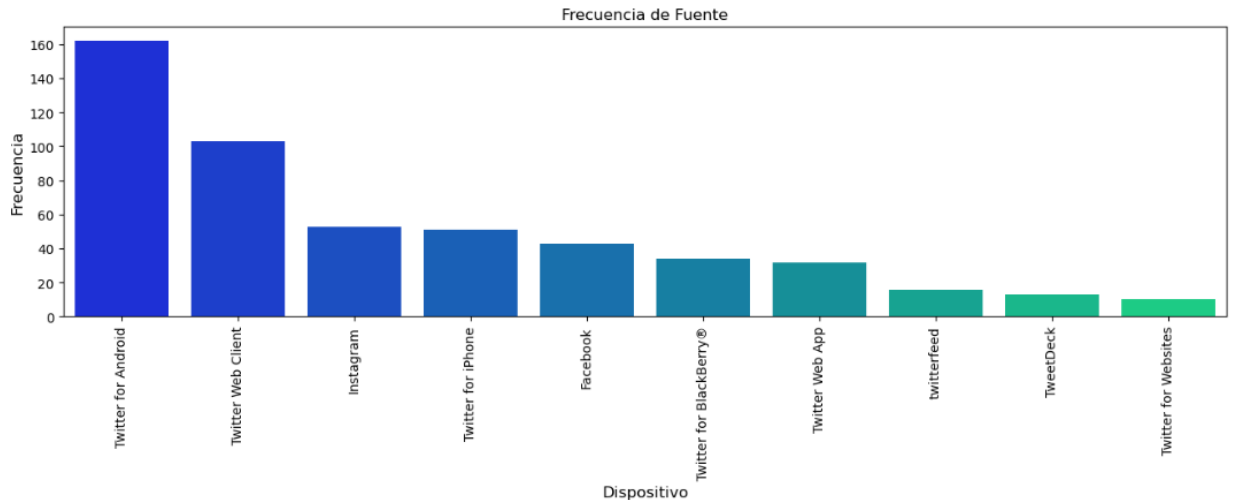


Figura 54: Frecuencia de dispositivos desde dónde se publicaron los tweets para “Multivacaciones”

La Figura 55. Permite visualizar desde que localidades se emitieron los tweets, esta información tiene fundamental relevancia, debido que nos permitirá identificar desde que ciudades o países existe mayor número de opiniones sobre Multivacaciones

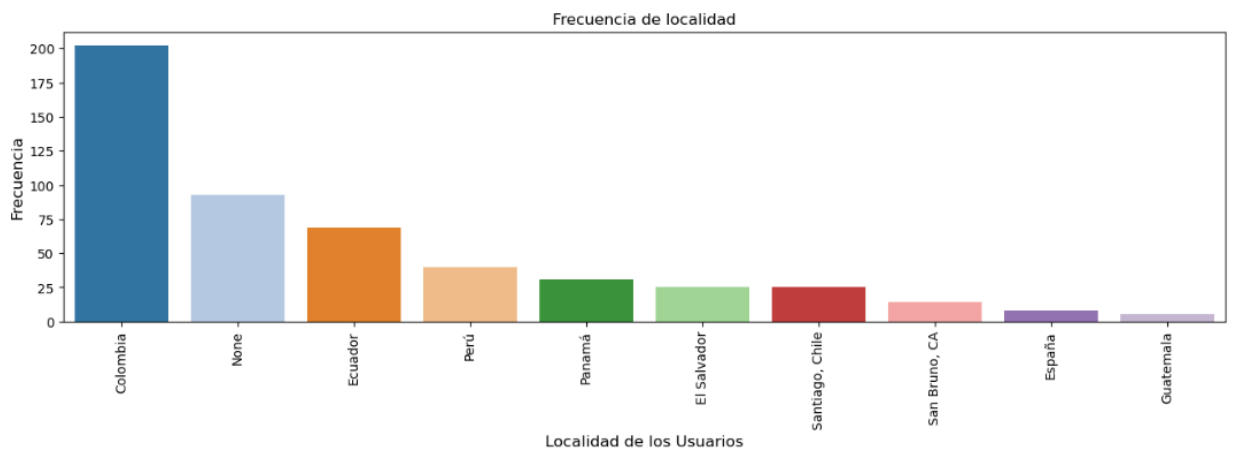


Figura 55: Localidades desde dónde se publicaron los tweets para “Multivacaciones”

El poder conocer que apreciación tienen los usuarios sobre Multivacaciones por localidad, permitirá enfocar los esfuerzos, en localidades donde hay menor acogida. En las Figura 56 y Figura 57. se puede observar la frecuencia de tweests positivos, neutros y negativos por localidades.

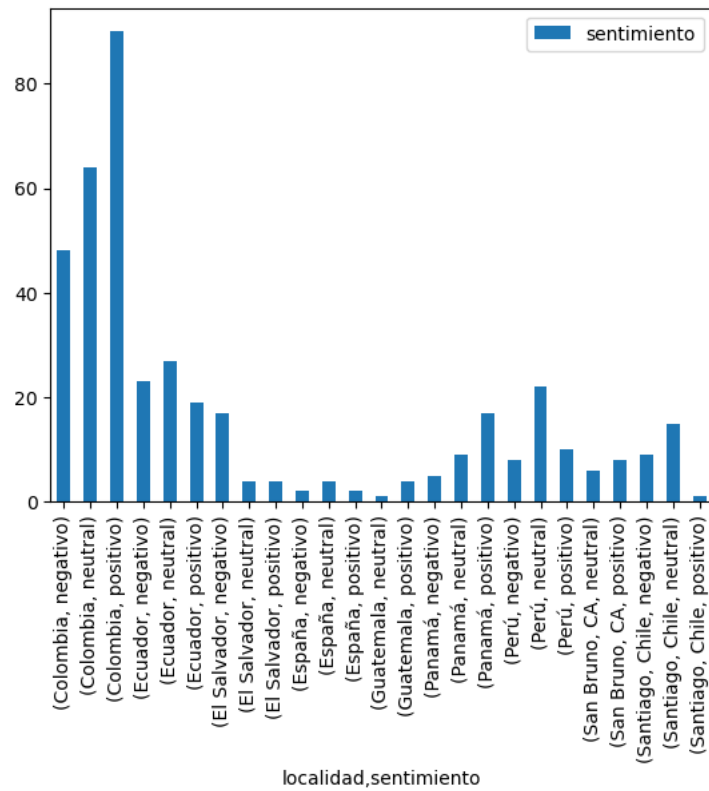


Figura 56: Análisis de sentimientos por localidades para “Multivacaciones” (a)

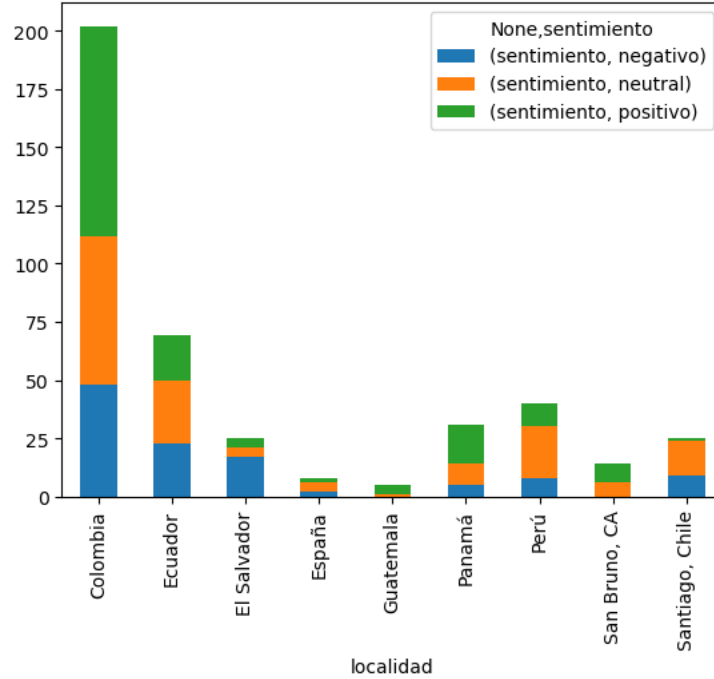


Figura 57: Análisis de sentimientos por localidades para “Multivacaciones”

Resulta de interés analizar a los usuarios que opinaron sobre Multivacaciones, En la Figura 58. Se visualiza los usuarios que tiene un mayor número de amigos, en la Figura 59. Se puede observar los usuarios que tienen mayor número de seguidores, lo que podría indicarnos el nivel de influencia que poseen estos usuarios.

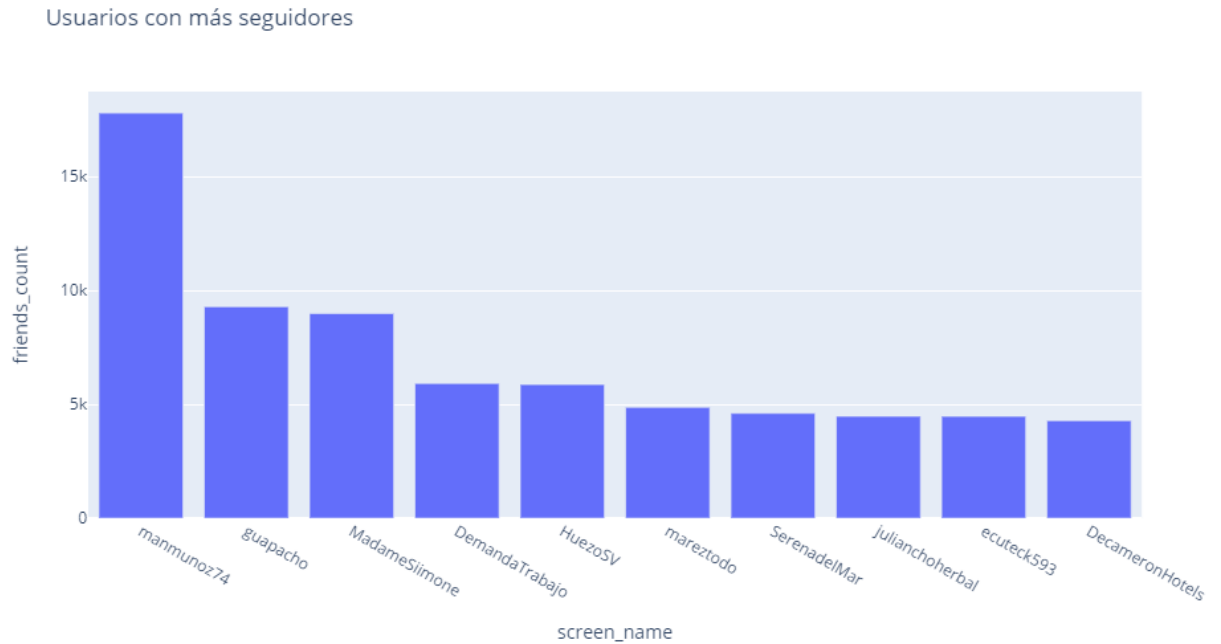


Figura 58: Usuarios con más amigos para “Multivacaciones”

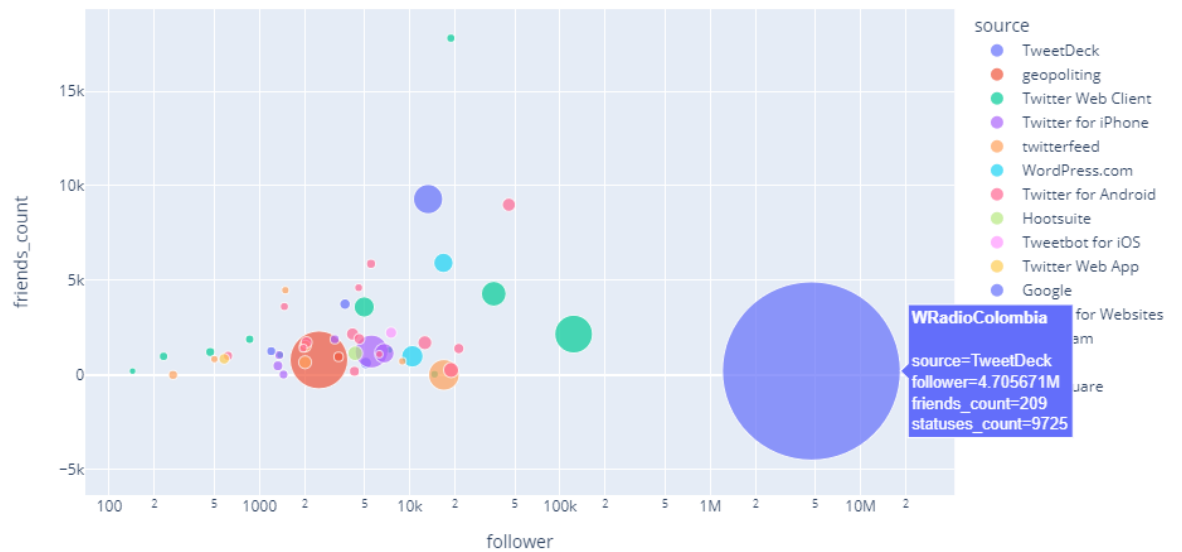


Figura 59: Descripción de Usuarios con sus seguidores y amigos para “Multivacaciones”

4.4 Análisis de la Encuesta

La encuesta se diseñó con 11 preguntas abiertas, la encuesta fue publicada en la red social Twitter. Se recolectaron datos de 95 encuestas.

Las preguntas 1, 2 y 3 nos servirán de apoyo para el análisis de las preguntas 4 a 11.

Pregunta 1:

4.4.1 ¿Cuál es tu edad?

Pregunta 2:

4.4.2 ¿Cuál es tu género?

Pregunta 3:

4.4.3 ¿En qué ciudad vives?

Pregunta 4:

4.4.4 ¿Te gusta las instalaciones del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.

La figura 60, indica de forma macro el porcentaje de opiniones positivas, negativas y neutras para las instalaciones del Hotel Decameron Mompiche.

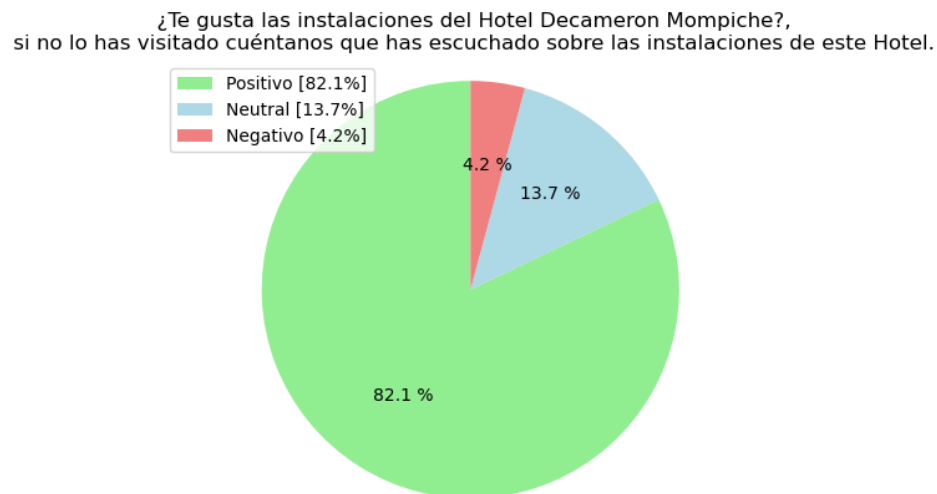


Figura 60: Análisis de sentimientos para la Pregunta 4, de la encuesta.

Las Figuras 61 y 62, muestran el análisis sentimientos sobre las opiniones concernientes a las instalaciones del Hotel Decameron Mompiche por localidad, este resultado permitirá determinar en qué localidades el Hotel tiene mayor acogida.

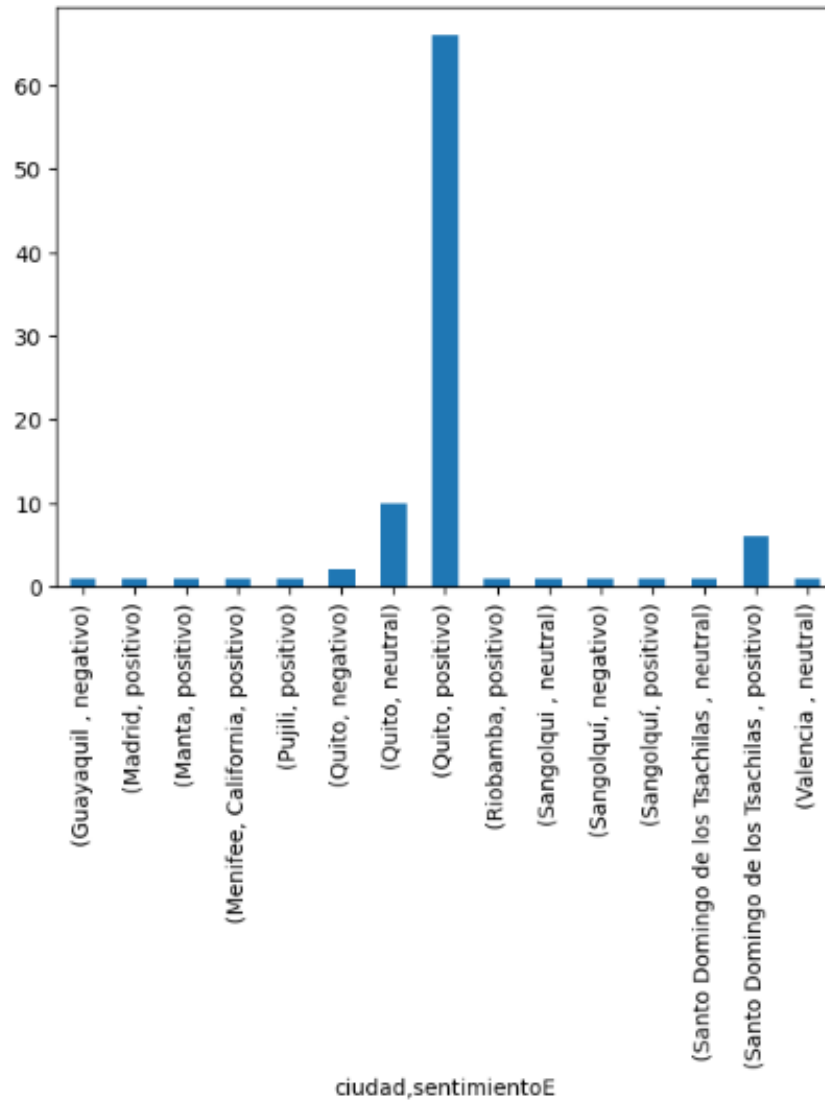


Figura 61: Análisis de sentimientos por localidades para la Pregunta 4 de la encuesta (a).

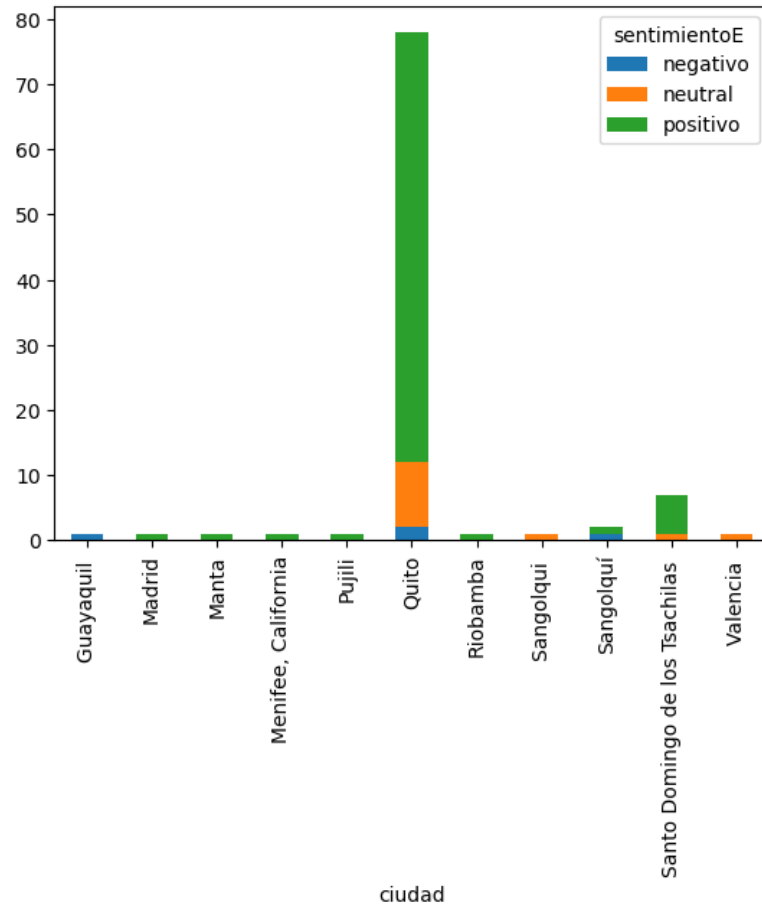


Figura 62: Análisis de sentimientos por localidades para la Pregunta 4 de la encuesta (b).

La Figura 63. permite visualizar el análisis de sentimientos por genero, podemos notar que el genero femenino expreso menor cantidad de opiniones negativas.

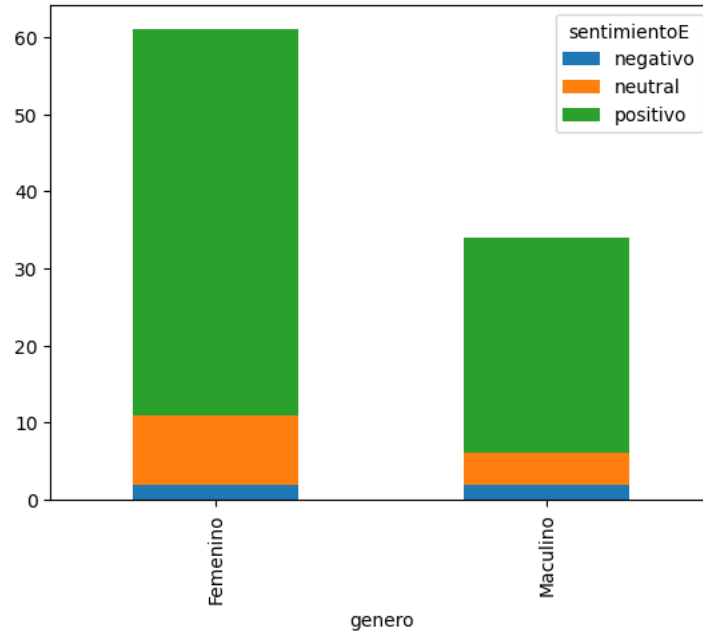


Figura 63: Análisis de sentimientos por género para la Pregunta 4 de la encuesta.

Para tener una visión global de las palabras usadas en las opiniones sobre la pregunta 4 de la encuesta, se elaboró una nube de palabras global, como se muestra en la Figura 64.



Figura 64: Nube de palabras para la Pregunta 4 de la encuesta.

Las Figuras 65, 66 y 67 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 4 de la encuesta, podemos observar que una de las palabras positivas

más relevantes es “bonita”, en cuanto a la nube de palabras negativas, “pandemia” es una de las palabras que se destaca.



Figura 65: Nube de palabras positivas para la Pregunta 4 de la encuesta.



Figura 66: Nube de palabras neutras para la Pregunta 4 de la encuesta.



Figura 67: Nube de palabras negativas para la Pregunta 4 de la encuesta.

Pregunta 5:

4.4.5 ¿Te gusta las instalaciones del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.

La figura 68, indica de forma general el porcentaje de opiniones positivas, negativas y neutras para las instalaciones del Hotel Decameron Punta Centinela.

¿Te gusta las instalaciones del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre las instalaciones de este Hotel.

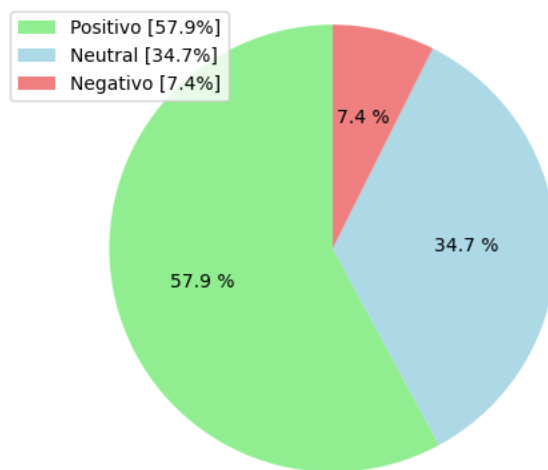


Figura 68: Análisis de sentimientos para la Pregunta 5 de la encuesta.

Las Figuras 69 y 70, permiten visualizar el análisis sentimientos sobre las opiniones concernientes a las instalaciones del Hotel Decameron Punta Centinela por edad, este resultado permitirá determinar el rango de edades donde las instalaciones del Hotel tienen mayor acogida.

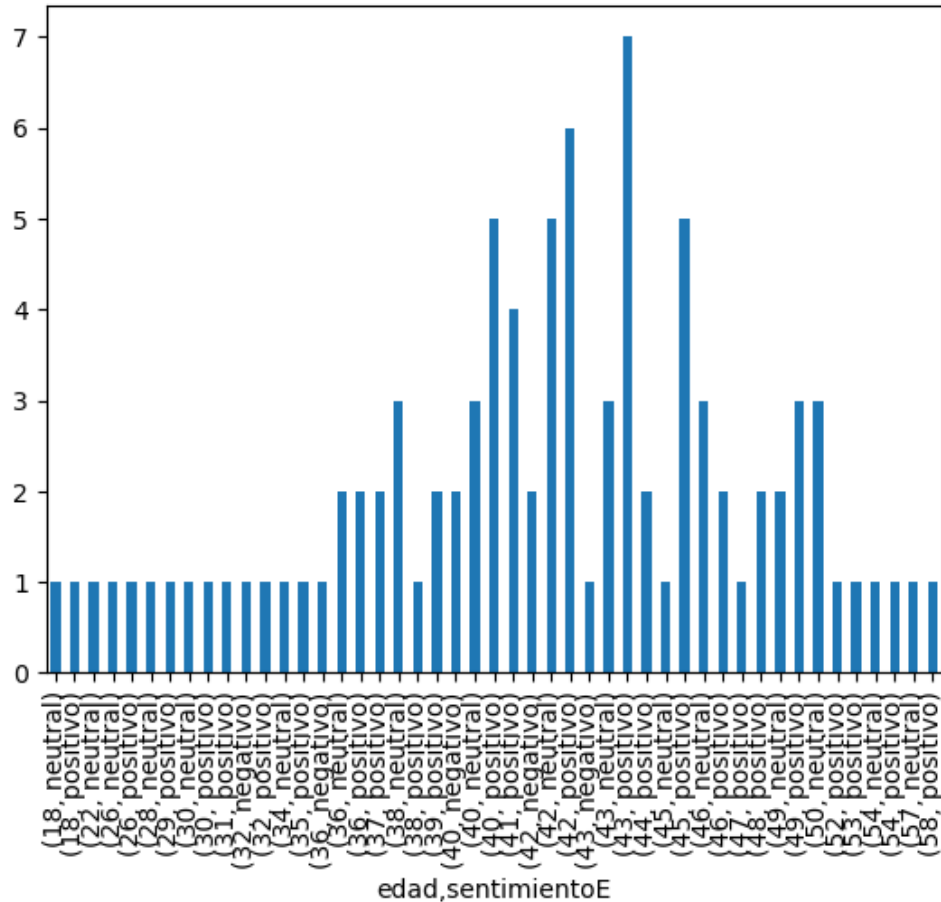


Figura 69: Análisis de sentimientos por edades para la Pregunta 5 de la encuesta (a).

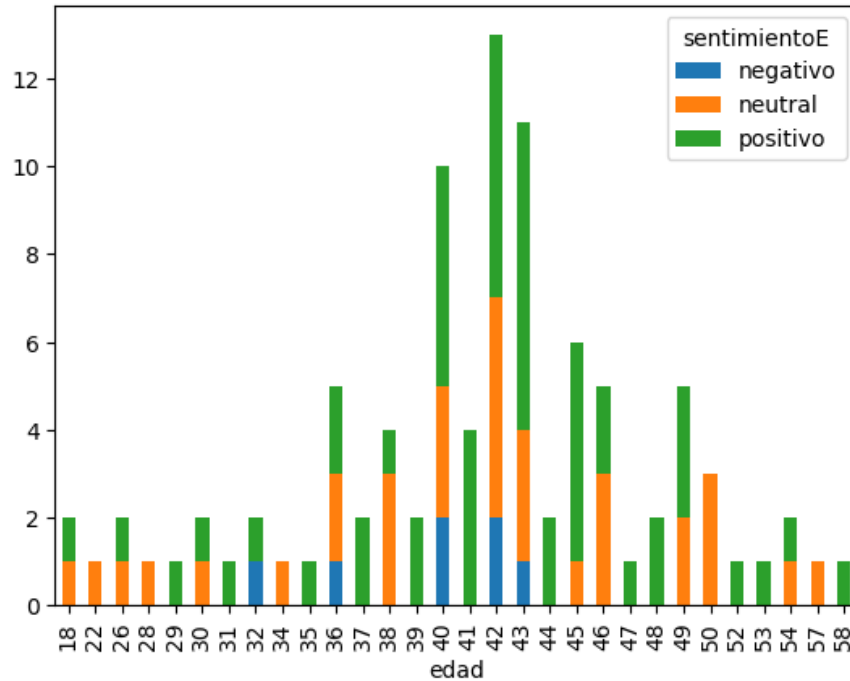


Figura 70: Análisis de sentimientos por edades para la Pregunta 5 de la encuesta. (b)

Las figuras 71 y 72 exponen el análisis de sentimientos relacionado a la pregunta 5 de la encuesta, por localidad, lo que permitirá determinar las localidades donde existe menor acogida de las instalaciones del Hotel Decameron Punta Centinela.

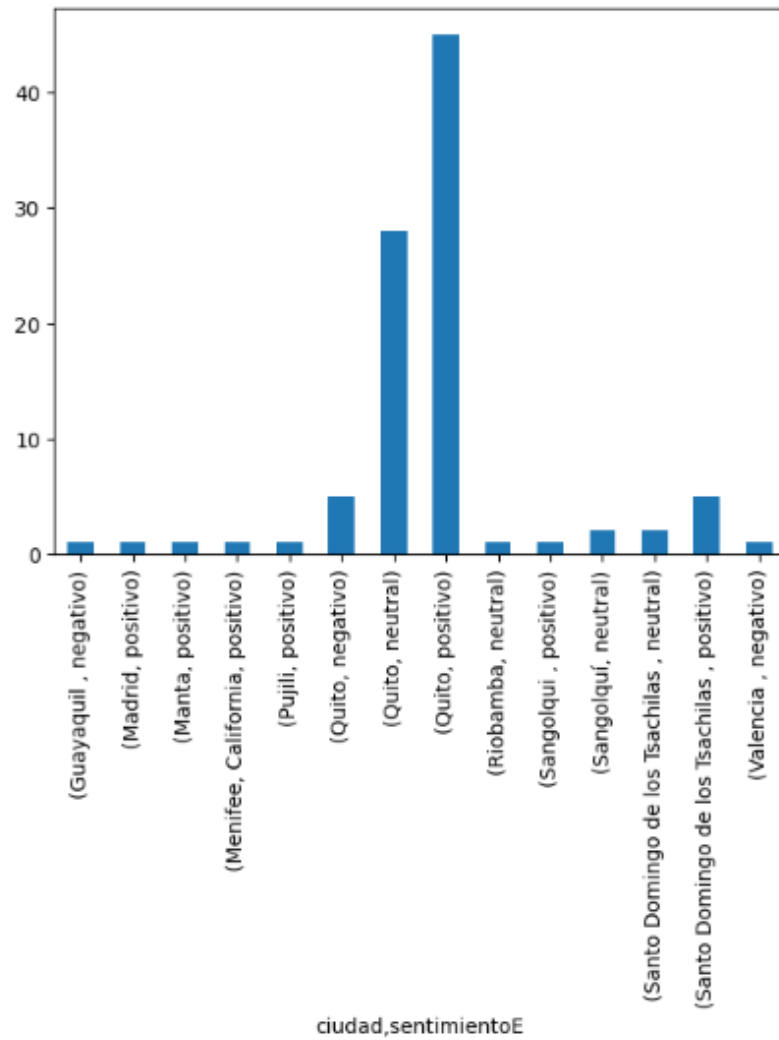


Figura 71: Análisis de sentimientos por localidad para la Pregunta 5 de la encuesta. (a)

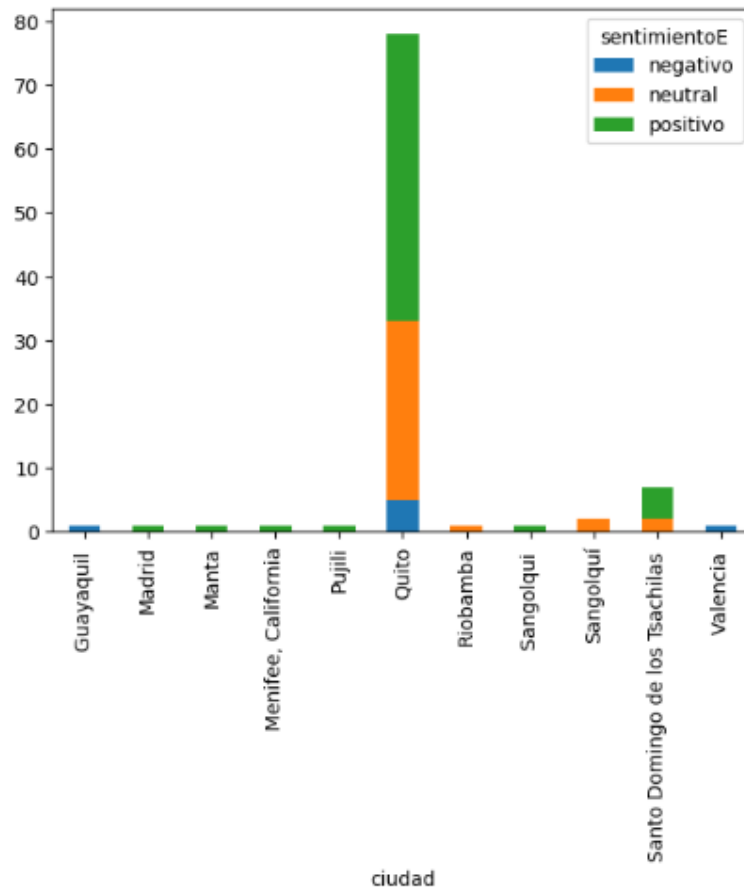


Figura 72: Análisis de sentimientos por localidad para la Pregunta 5 de la encuesta. (b)

La Figura 73 permite visualizar el análisis de sentimientos por genero, podemos notar que el genero femenino tiene un número mayor de opiniones.

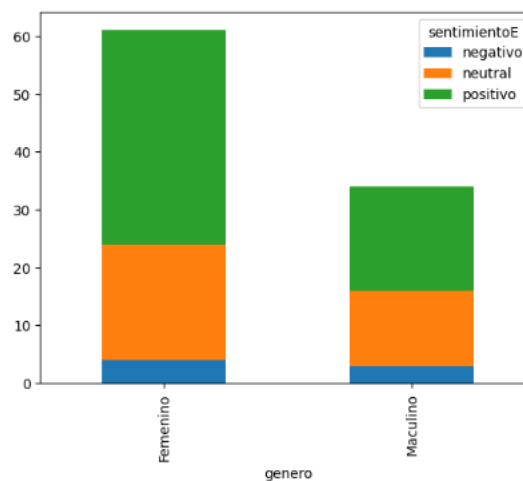


Figura 73: Análisis de sentimientos por genero para la Pregunta 5 de la encuesta.

Para tener una visión general de las palabras usadas en las opiniones sobre la pregunta 5 de la encuesta, se elaboró una nube de palabras global, como se muestra en la Figura 74.

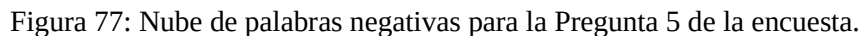


Figura 74: Nube de palabras para la Pregunta 5 de la encuesta.

Las Figuras 75, 76 y 77 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 5 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “lindas”, en cuanto a la nube de palabras negativas, “comida” es una de las palabras que se destaca.



Figura 75: Nube de palabras positivas para la Pregunta 5 de la encuesta.



4.4.6 ¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.

En figura 78, podemos observar la el porcentaje de opiniones positivas, negativas y neutras asociadas a los servicios que brindan Hotel Decameron Mompiche.

¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Mompiche?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.

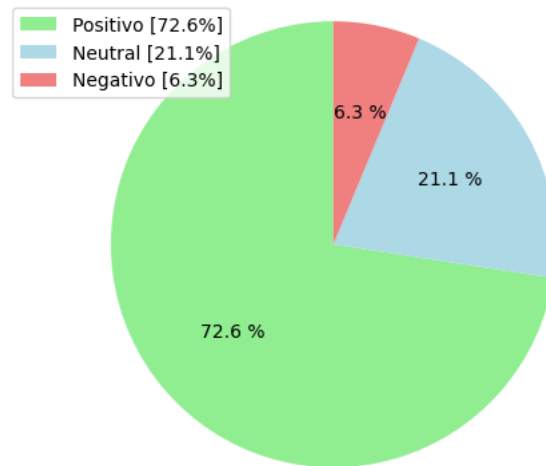


Figura 78: Análisis de sentimientos para la Pregunta 6, de la encuesta.

Las figuras 79 y 80, permiten visualizar el analisis de sentimientos por edades asociado a la pregunta 6 de la encuesta, con esta información se podrá determinar el rango de edades que en las cuales los servicios que ofrece Hoteles Decameron tiene mayor acogida.

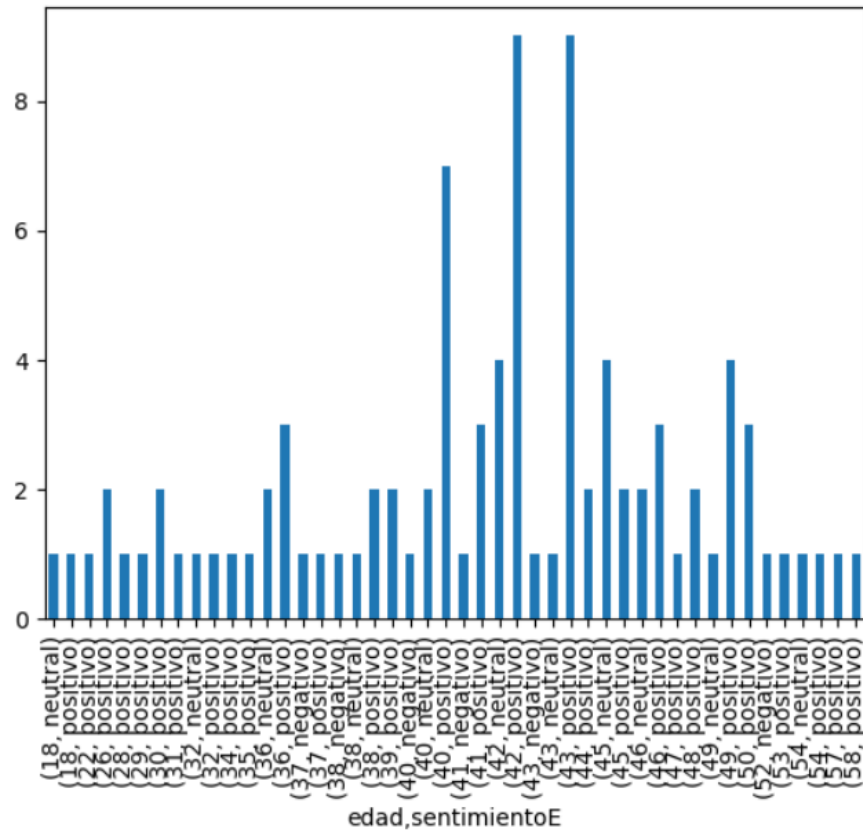


Figura 79: Análisis de sentimientos por edades para la Pregunta 6 de la encuesta. (a)

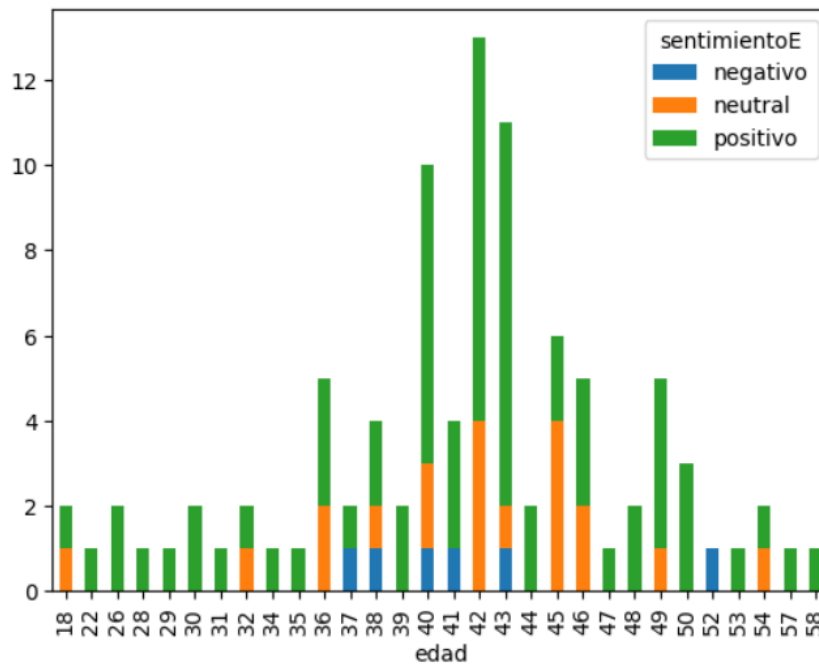


Figura 80: Análisis de sentimientos por edades para la Pregunta 6 de la encuesta. (b)

Las Figuras 81 y 82, muestran el análisis sentimientos sobre las opiniones concernientes a los servicios del Hotel Decameron Mompiche por localidad.

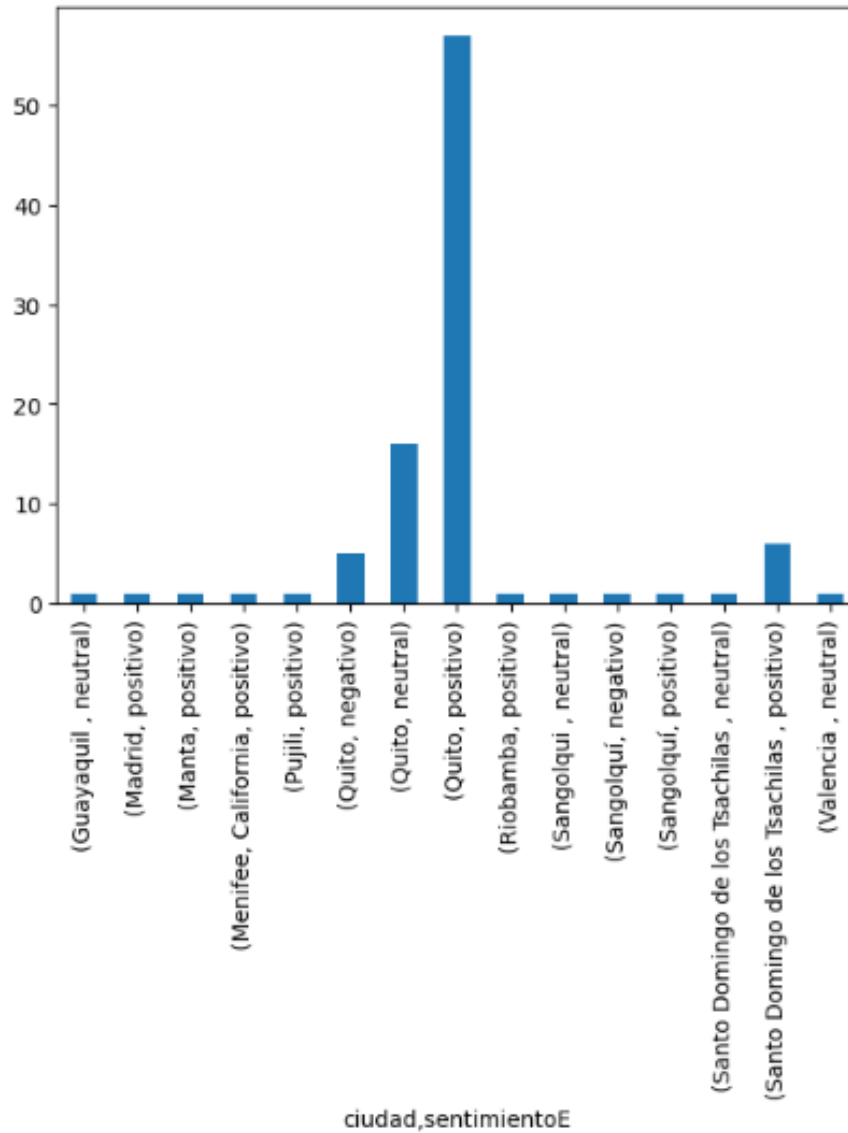


Figura 81: Análisis de sentimientos por localidad para la Pregunta 6 de la encuesta. (a)

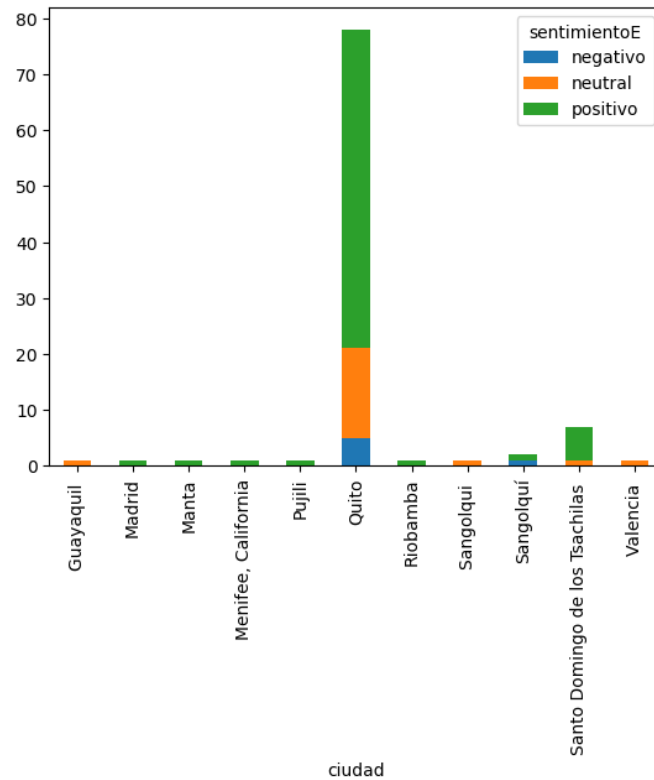


Figura 82: Análisis de sentimientos por localidad para la Pregunta 6 de la encuesta. (b)

La Figura 83 permite visualizar el análisis de sentimientos por genero.

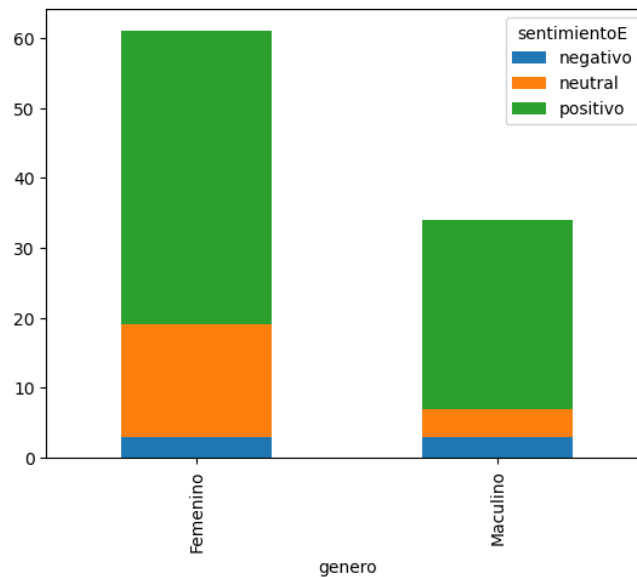


Figura 83: Análisis de sentimientos por genero para la Pregunta 6 de la encuesta.

Para tener una visión integral de las palabras usadas en las opiniones sobre la pregunta 6 de la encuesta, se elaboró una nube de palabras, como se muestra en la Figura 84.



Figura 84: Nube de palabras para la Pregunta 6 de la encuesta.

Las Figuras 85, 86 y 87 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 6 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “buena”, en cuanto a la nube de palabras negativas, “mala” es una de las palabras que se destaca.

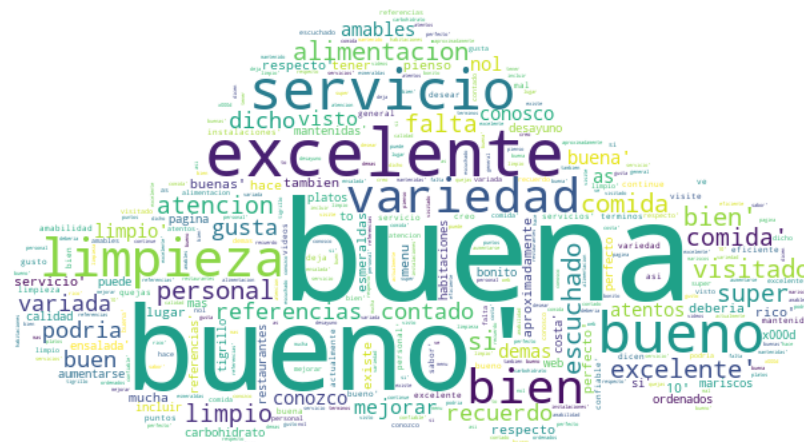


Figura 85: Nube de palabras positivas para la Pregunta 6 de la encuesta.



Figura 86: Nube de palabras neutras para la Pregunta 6 de la encuesta.



Figura 87: Nube de palabras negativas para la Pregunta 6 de la encuesta.

Pregunta 7:

4.4.7 ¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.

La figura 88, indica de forma general el porcentaje de opiniones positivas, negativas y neutras para las instalaciones del Hotel Decameron Punta Centinela.

¿Qué opinas del servicio (alimentación, limpieza, personal) del Hotel Decameron Punta Centinela?, si no lo has visitado cuéntanos que has escuchado sobre el servicio de este Hotel.

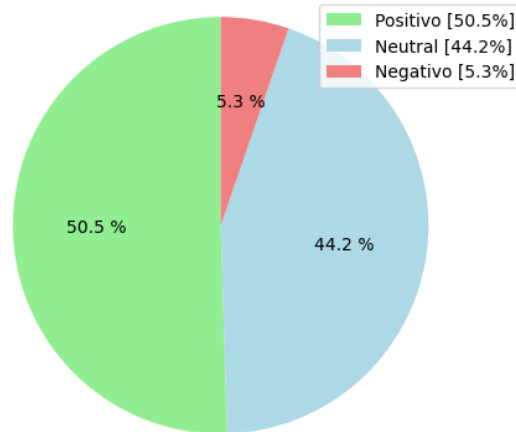


Figura 88: Análisis de sentimientos para la Pregunta 7 de la encuesta.

Las Figuras 89 y 90, muestran el análisis sentimientos sobre las opiniones concernientes a las instalaciones del Hotel Decameron Punta Centinela por edad.

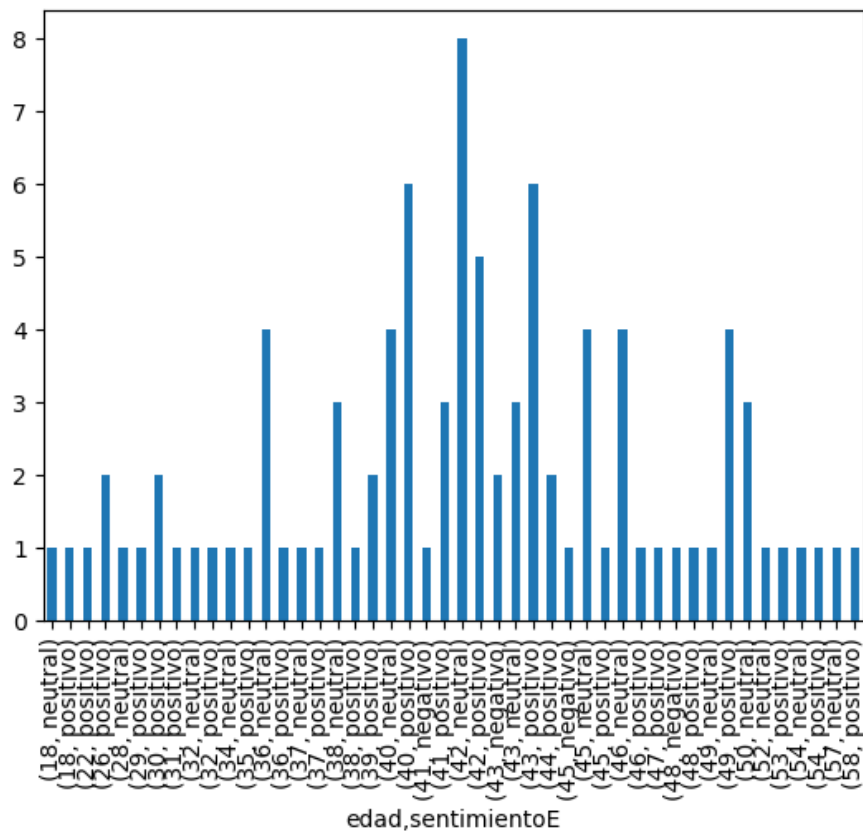


Figura 89: Análisis de sentimientos por edades para la Pregunta 7 de la encuesta. (a)

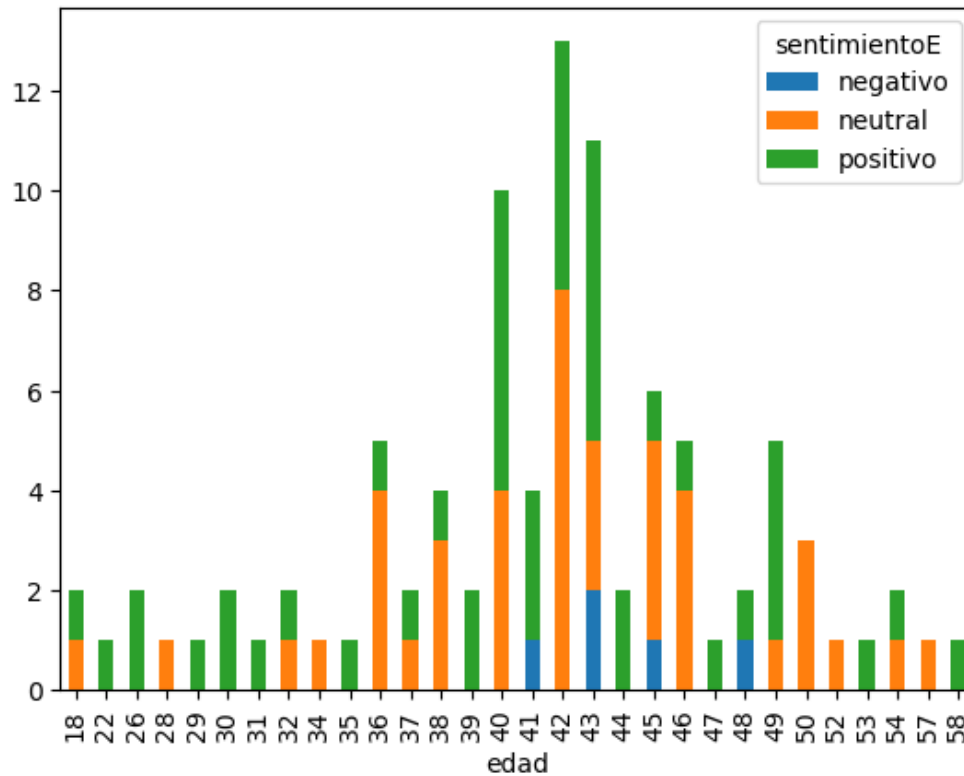


Figura 90: Análisis de sentimientos por edades para la Pregunta 7 de la encuesta. (b)

En las Figuras 91 y 92, se puede visualizar el análisis de sentimientos relacionado a los servicios ofrecidos por el Hotel Decameron Punta Centinela por localidades.

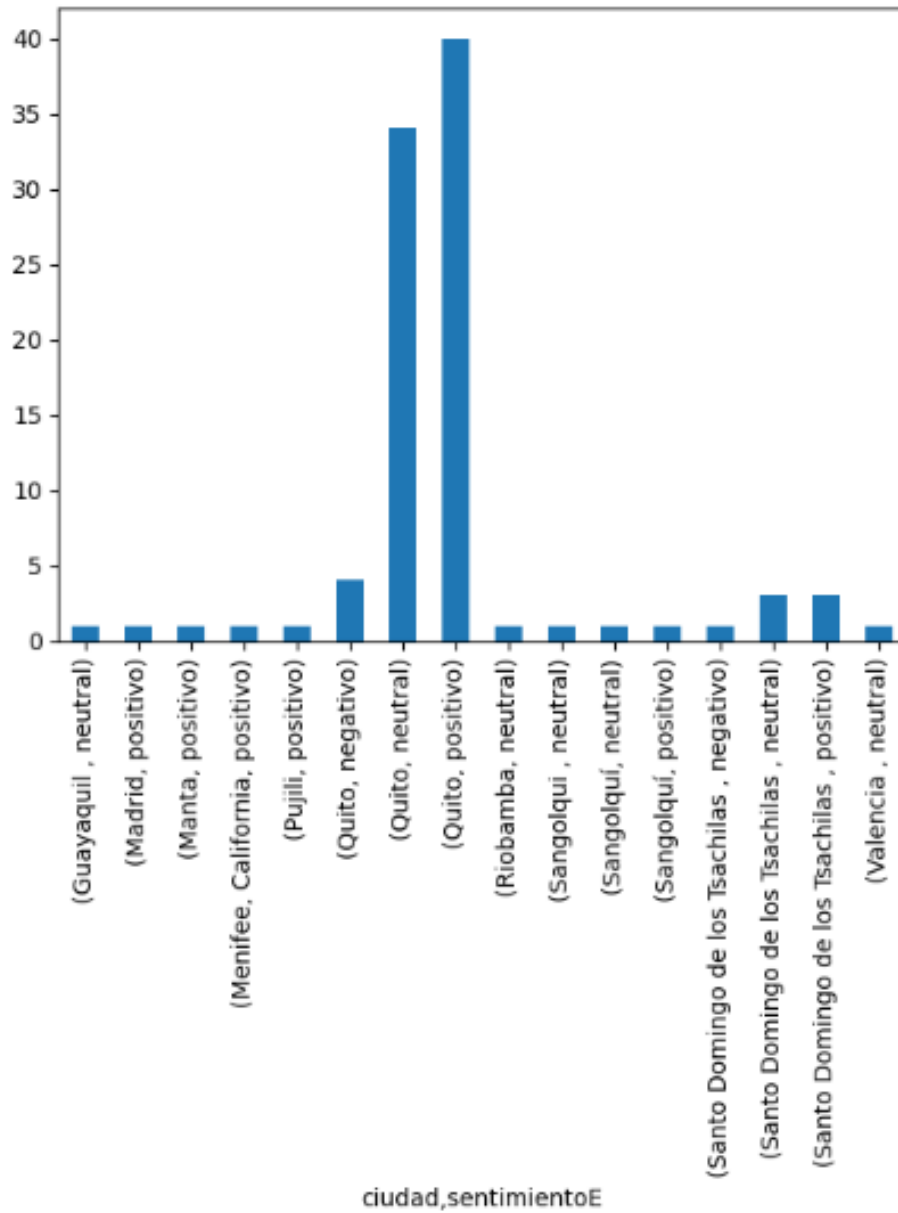


Figura 91: Análisis de sentimientos por localidad para la Pregunta 7 de la encuesta. (a)

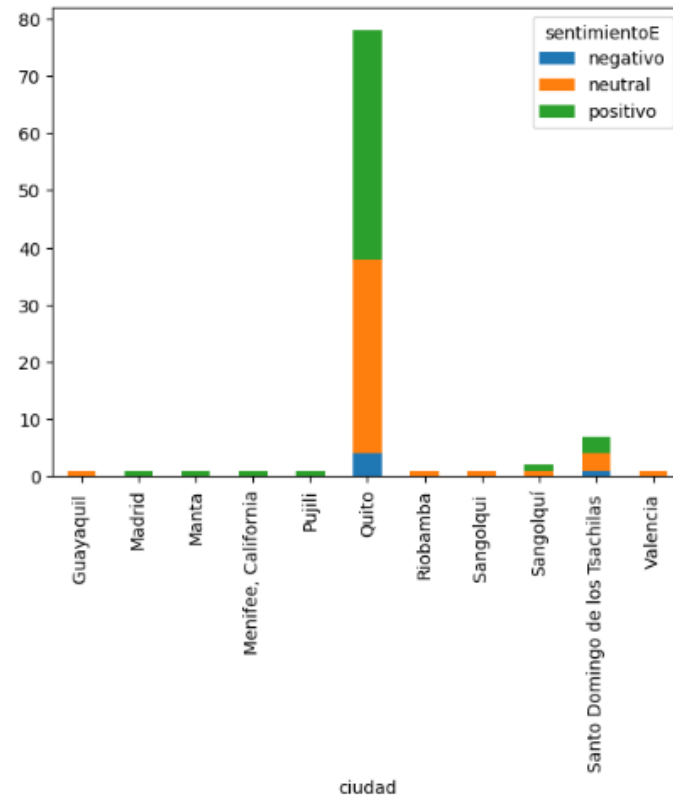


Figura 92: Análisis de sentimientos por localidad para la Pregunta 7 de la encuesta. (b)

La Figura 93 permite visualizar el análisis de sentimientos por genero, podemos notar que el genero femenino expreso menor cantidad de opiniones negativas.

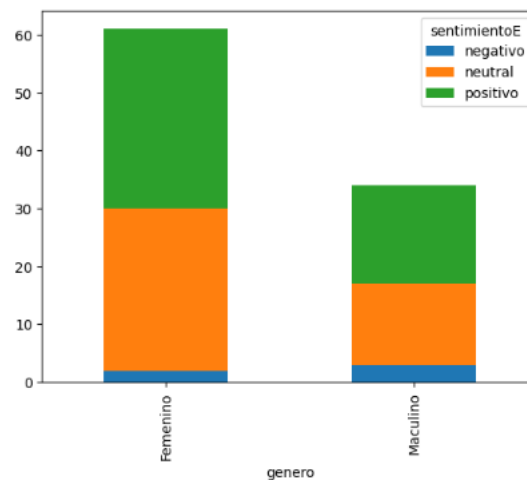


Figura 93: Análisis de sentimientos por genero para la Pregunta 7 de la encuesta.

Para tener una visión global de las palabras usadas en las opiniones sobre la pregunta 7 de la encuesta, se elaboró una nube de palabras general, como se muestra en la Figura 94.



Figura 94: Nube de palabras para la Pregunta 7 de la encuesta.

Las Figuras 95, 96 y 97 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 7 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “respuesta”, en cuanto a la nube de palabras negativas, “pandemia” es una de las palabras que se destaca



Figura 95: Nube de palabras positivas para la Pregunta 7 de la encuesta.



Figura 96: Nube de palabras neutras para la Pregunta 7 de la encuesta.



Figura 97: Nube de palabras negativas para la Pregunta 7 de la encuesta.

Pregunta 8:

4.4.8 Qué opinas de multivacaciones?

En la Figura 98, podemos observar los porcentajes de opiniones positivas, negativas y neutras de las opiniones sobre Multivacaciones.

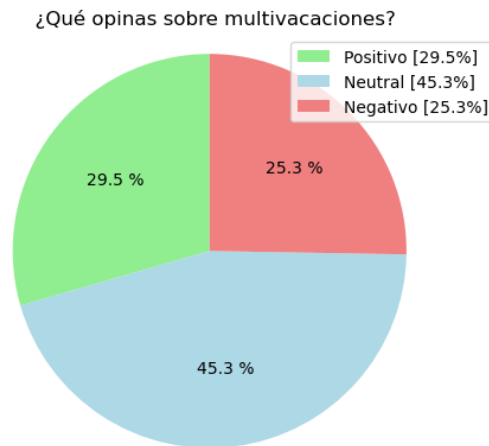


Figura 98: Análisis de sentimientos para la Pregunta 8 de la encuesta.

Las Figuras 99 y 100, muestran el análisis sentimientos sobre las opiniones concernientes a Multivacaciones por edad.

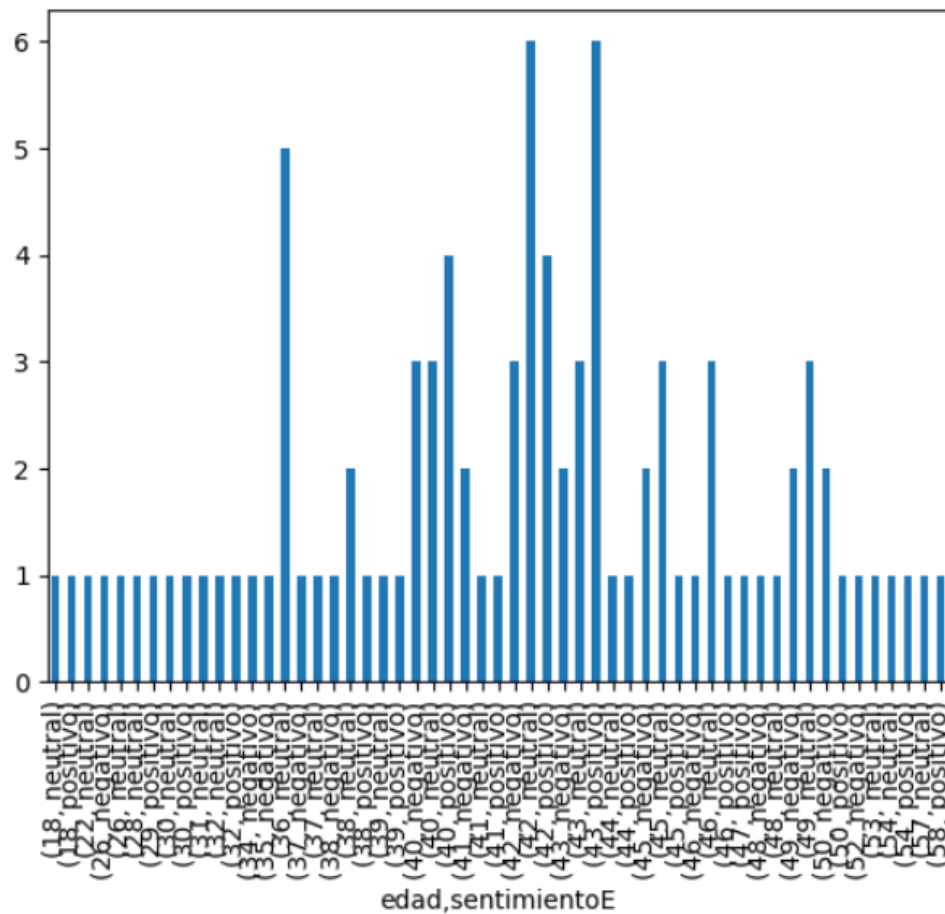


Figura 99: Análisis de sentimientos por edad para la Pregunta 8 de la encuesta. (a)

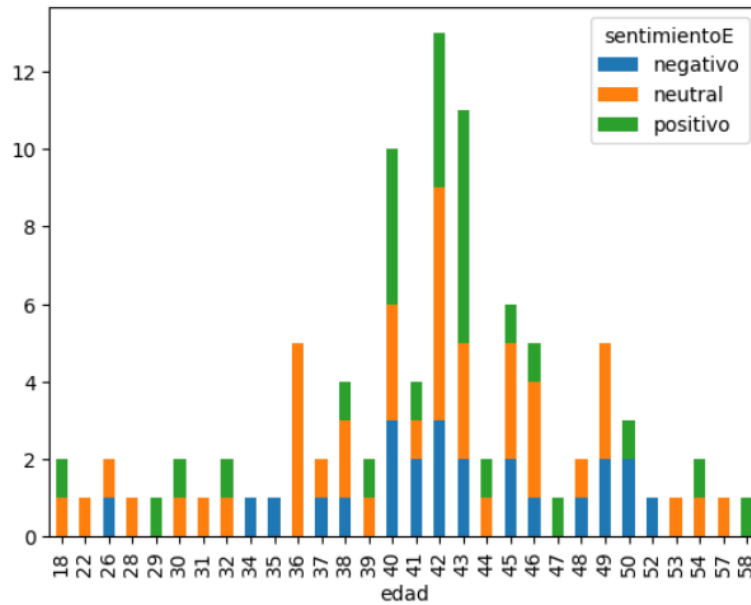


Figura 100: Análisis de sentimientos por edad para la Pregunta 8, de la encuesta. (b)

Las Figuras 101 y 102, muestran el análisis sentimientos sobre las opiniones relacionadas a Multivacaciones, por localidad, este resultado permitirá determinar en qué localidades el Hotel tiene menor acogida.

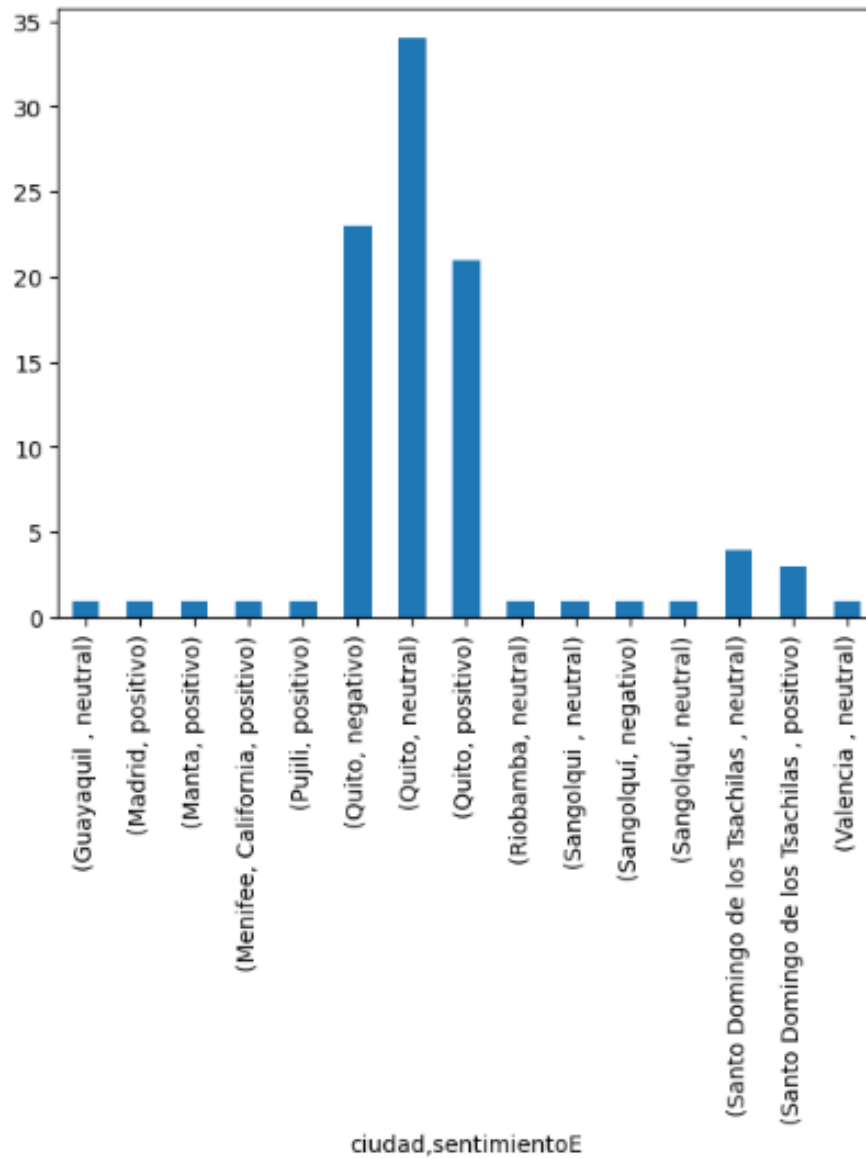


Figura 101: Análisis de sentimientos por localidad para la Pregunta 8 de la encuesta. (a)

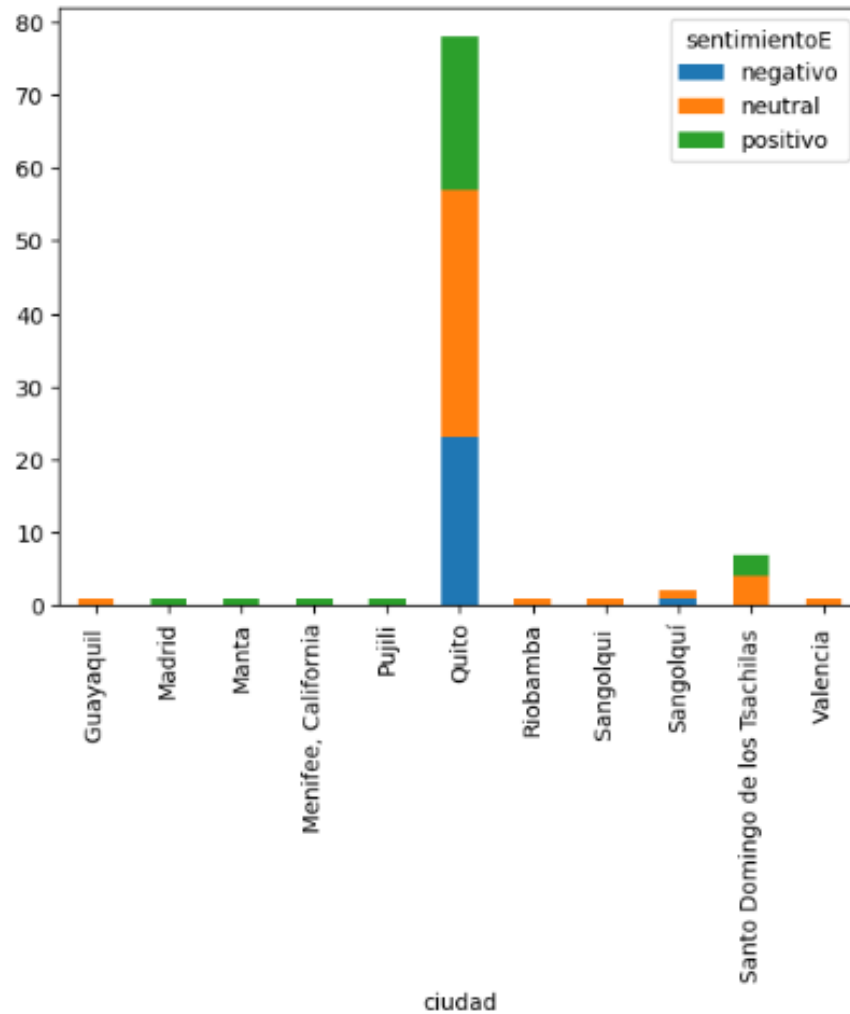


Figura 102: Análisis de sentimientos por localidad para la Pregunta 8 de la encuesta. (b)

La Figura 103 permite visualizar el análisis de sentimientos por genero.

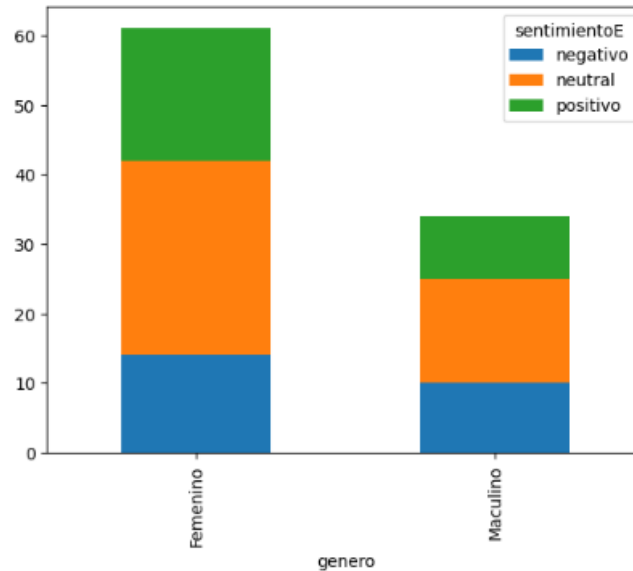


Figura 103: Análisis de sentimientos por genero para la Pregunta 8 de la encuesta.

Para tener una visión global de las palabras usadas en las opiniones sobre la pregunta 8 de la encuesta, se elaboró una nube de palabras, como se muestra en la Figura 104.



Figura 104: Nube de palabras para la Pregunta 8, de la encuesta.

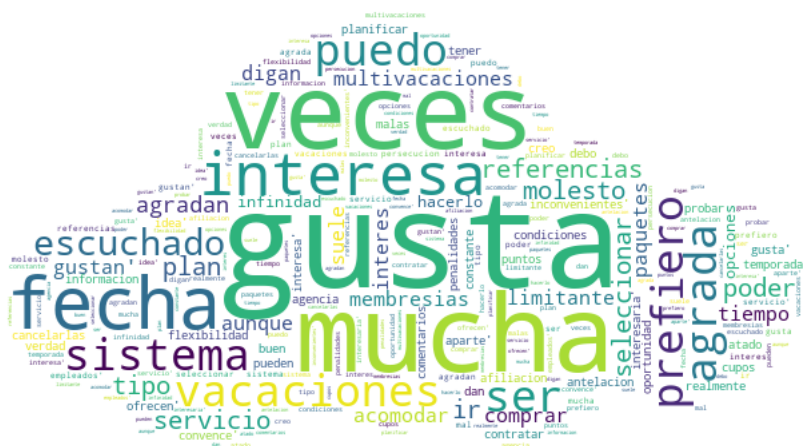
Las Figuras 105, 106 y 107 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 8 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “excelente”, en cuanto a la nube de palabras negativas, “escuchado” es una de las palabras que se destaca.



Figura 105: Nube de palabras positivas para la Pregunta 8 de la encuesta.



Figura 106: Nube de palabras neutras para la Pregunta 8 de la encuesta.



Las Figuras 109 y 110, muestran el análisis sentimientos sobre las opiniones emitidas para la pregunta 9.

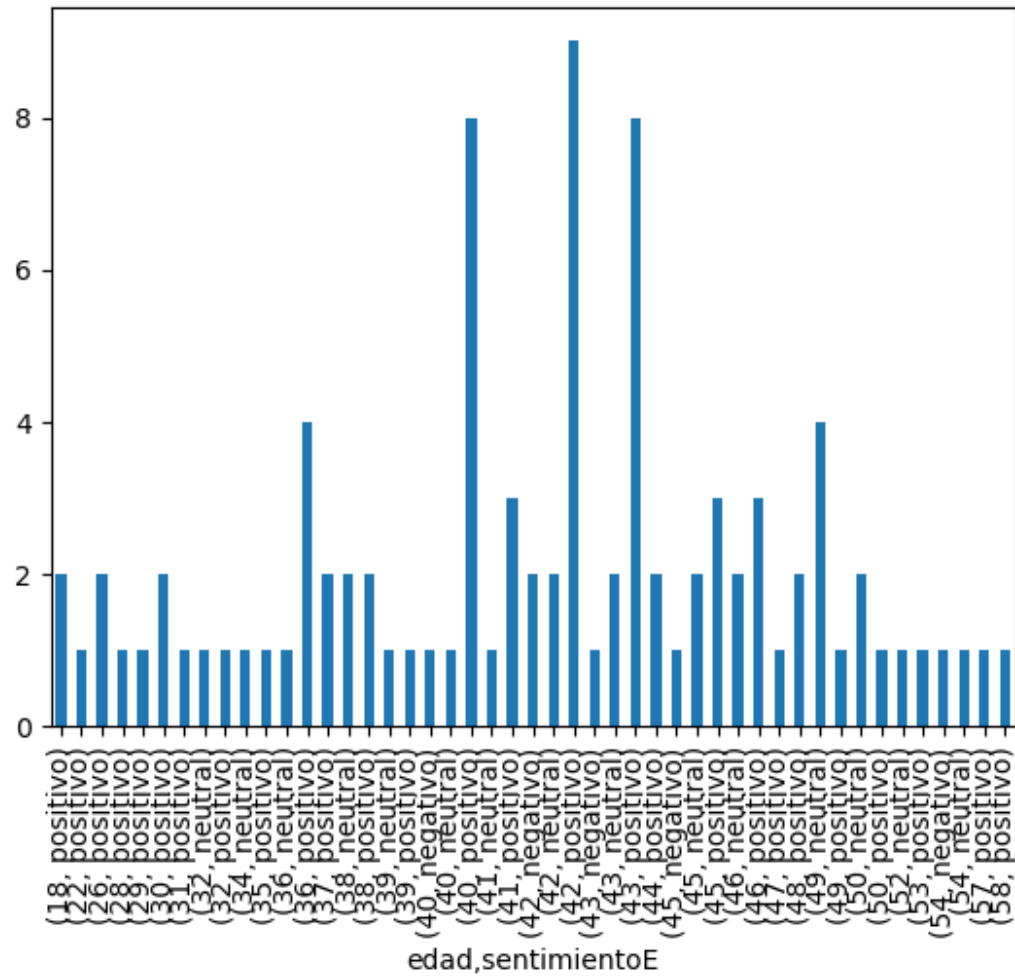


Figura 109: Análisis de sentimientos por edades para la Pregunta 9 de la encuesta. (a)

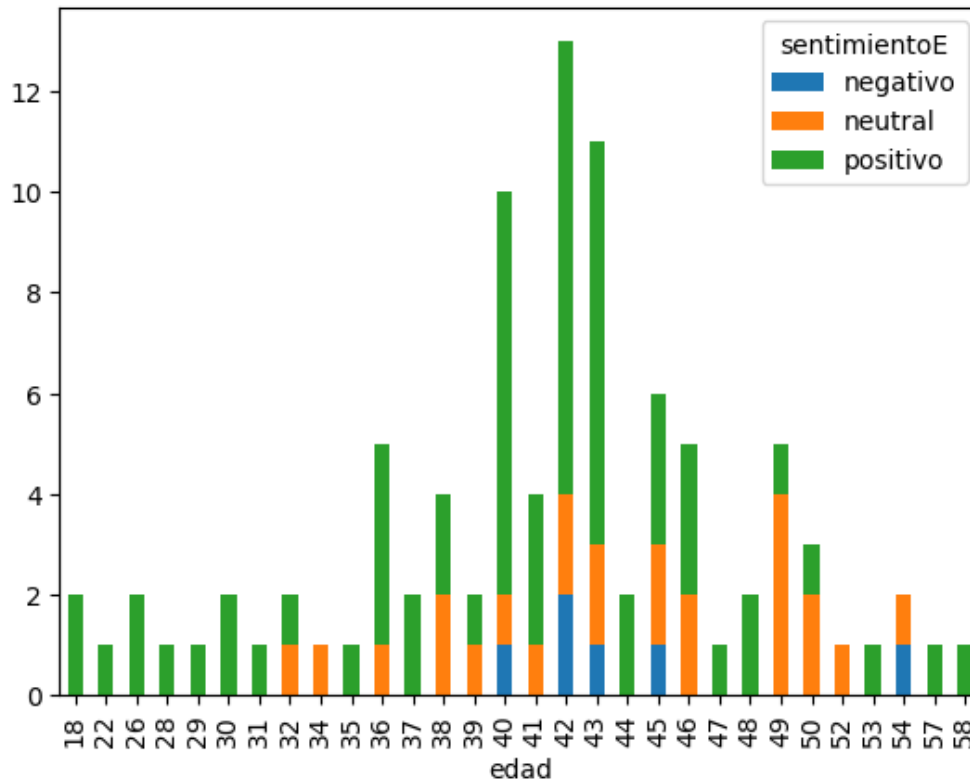


Figura 110: Análisis de sentimientos por edades para la Pregunta 9 de la encuesta. (b)

Las Figuras 111 y 112, permiten visualizar el análisis sentimientos sobre las opiniones sobre si Regresaría a Hotel Decameron Mompiche por localidad, este resultado permitirá determinar en qué localidades el Hotel tiene mayor acogida.

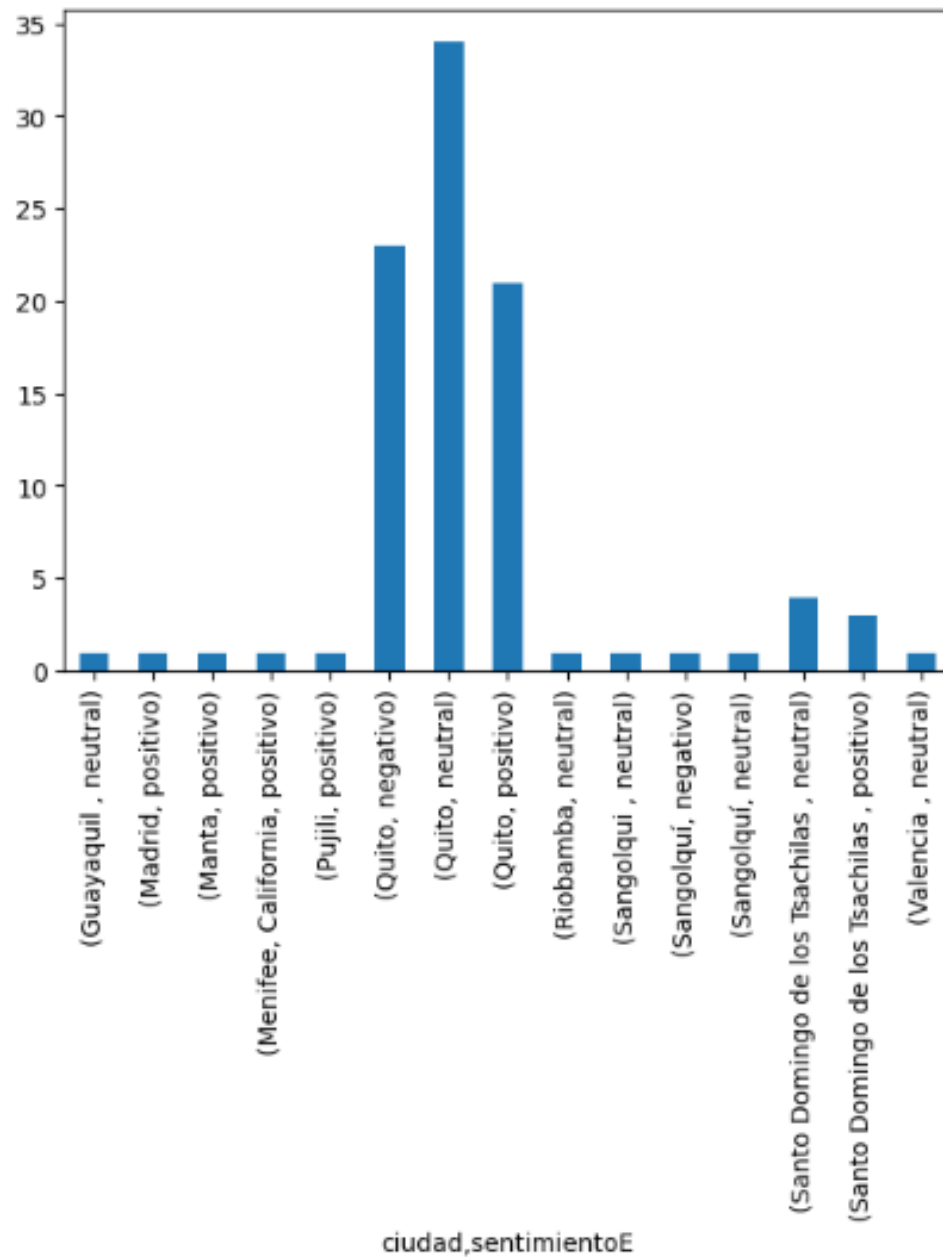


Figura 111: Análisis de sentimientos por localidad para la Pregunta 9 de la encuesta.

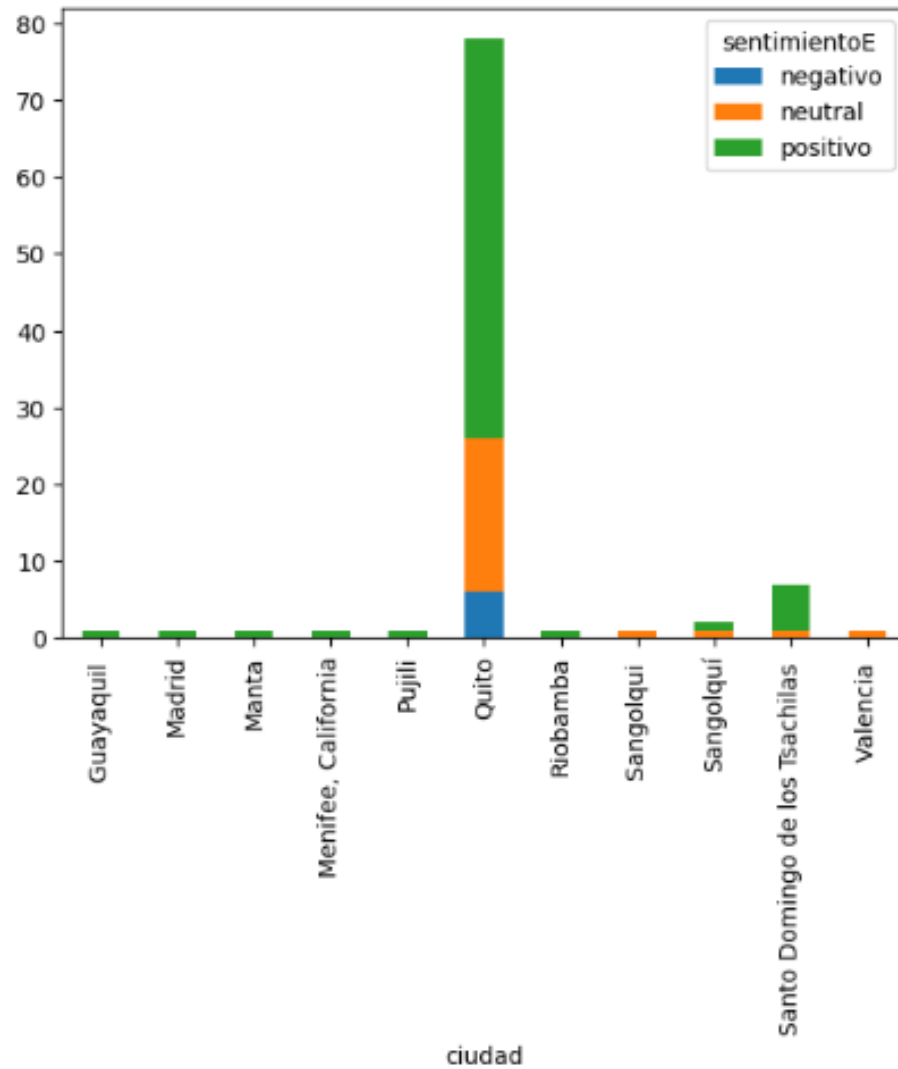


Figura 112: Análisis de sentimientos por localidad para la Pregunta 9 de la encuesta.

La Figura 113 permite visualizar el análisis de sentimientos por genero.

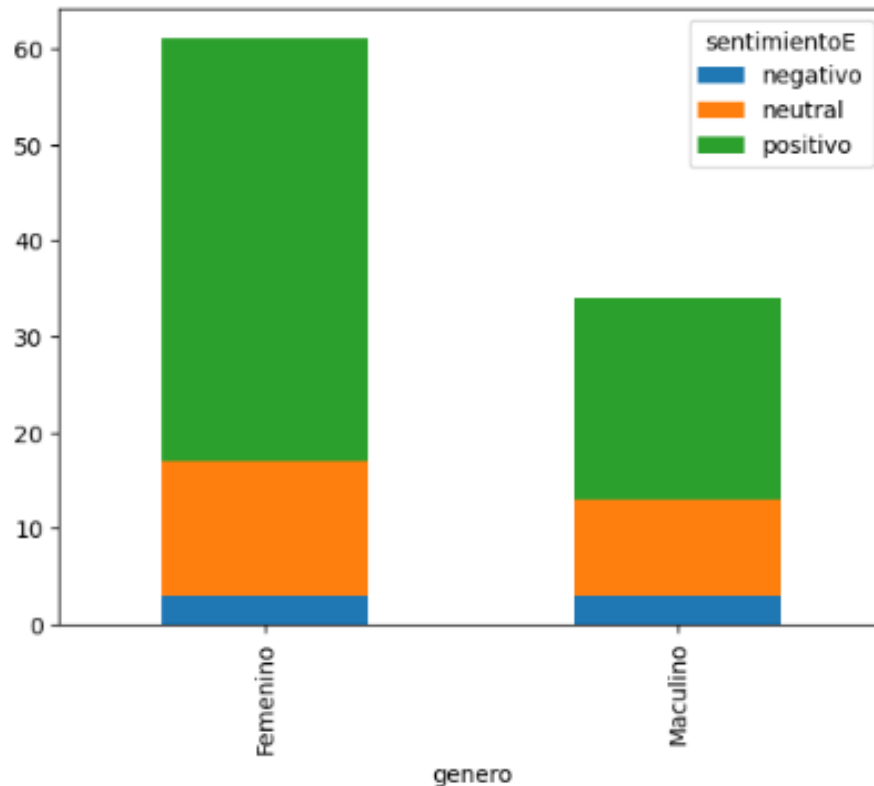


Figura 113: Análisis de sentimientos por genero para la Pregunta 9 de la encuesta.

La figura 114 muestra una visión global de las palabras usadas en las opiniones de la pregunta 9.



Figura 114: Nube de palabras para la Pregunta 9 de la encuesta.

Las Figuras 115, 116 y 117 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 9 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “gustaría”, en cuanto a la nube de palabras negativas, “servicio” es una de las palabras que se destaca.



Figura 115: Nube de palabras positivas para la Pregunta 9 de la encuesta.

Neutral



Figura 116: Nube de palabras neutras para la Pregunta 9 de la encuesta.

Negativo



Figura 117: Nube de palabras negativas para la Pregunta 9 de la encuesta.

Pregunta 10:

4.4.8.2 Recomendarías a familiares, amigos o Colegas Hoteles Decameron?

La figura 118, indica de forma global el porcentaje de opiniones positivas, negativas y neutras para las pregunta.

¿Recomendarías a familiares, amigos o colegas Hoteles Decameron?

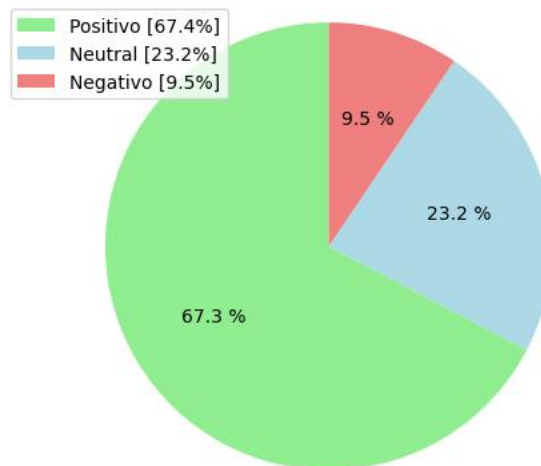


Figura 118: Análisis de sentimientos para la Pregunta 10 de la encuesta.

Las Figuras 119 y 120, muestran el análisis sentimientos sobre las opiniones concernientes a la pregunta 10 por edad, este resultado permitirá determinar los rangos de edades de las personas que recomendarían a Hoteles Decameron.

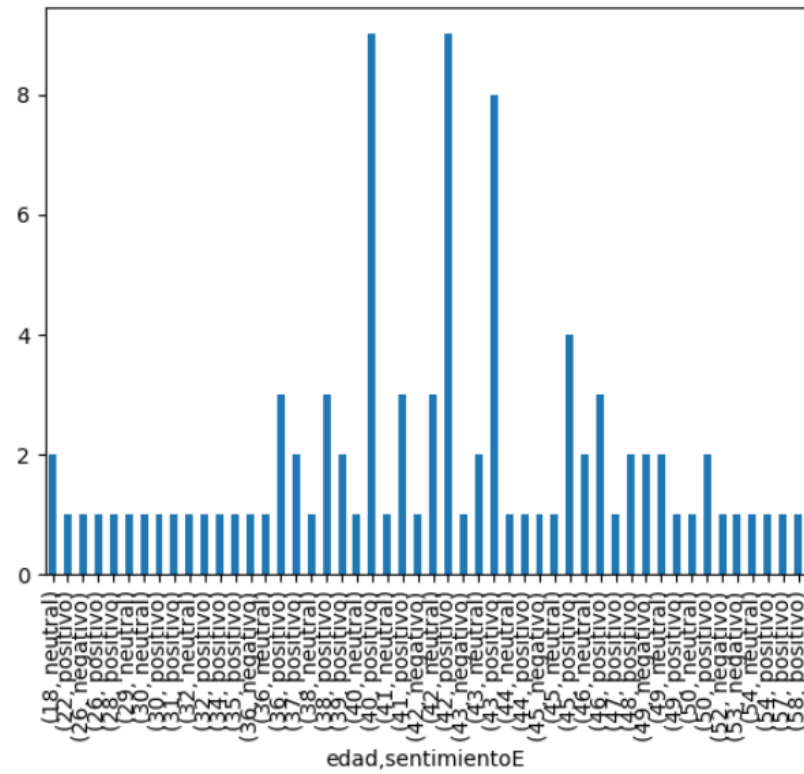


Figura 119: Análisis de sentimientos por edades para la Pregunta 10 de la encuesta. (a)

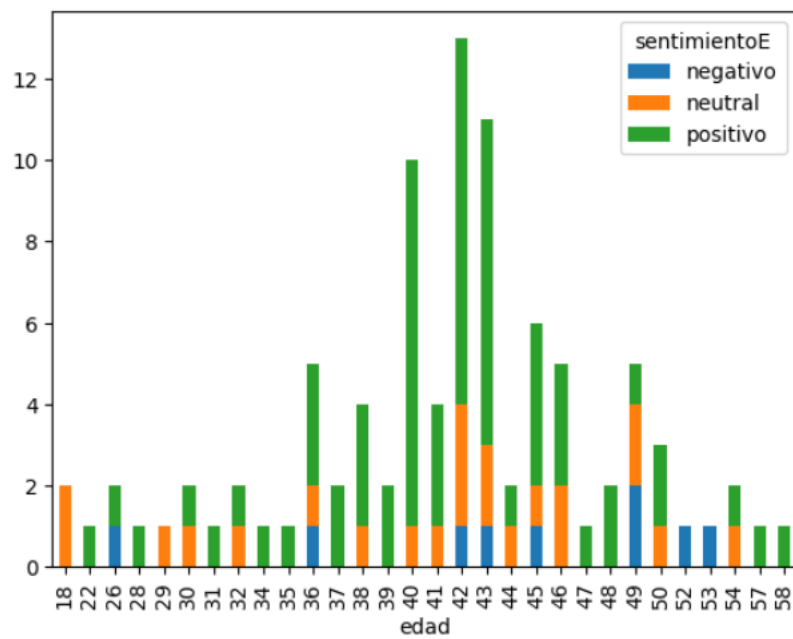


Figura 120: Análisis de sentimientos por edades para la Pregunta 10 de la encuesta. (b)

Las Figuras 121 y 122, muestran el análisis sentimientos sobre las opiniones relacionadas a la pregunta 9 por localidad, este resultado permitirá determinar en qué localidades están ubicadas las personas que recomendarían Hoteles Decameron.

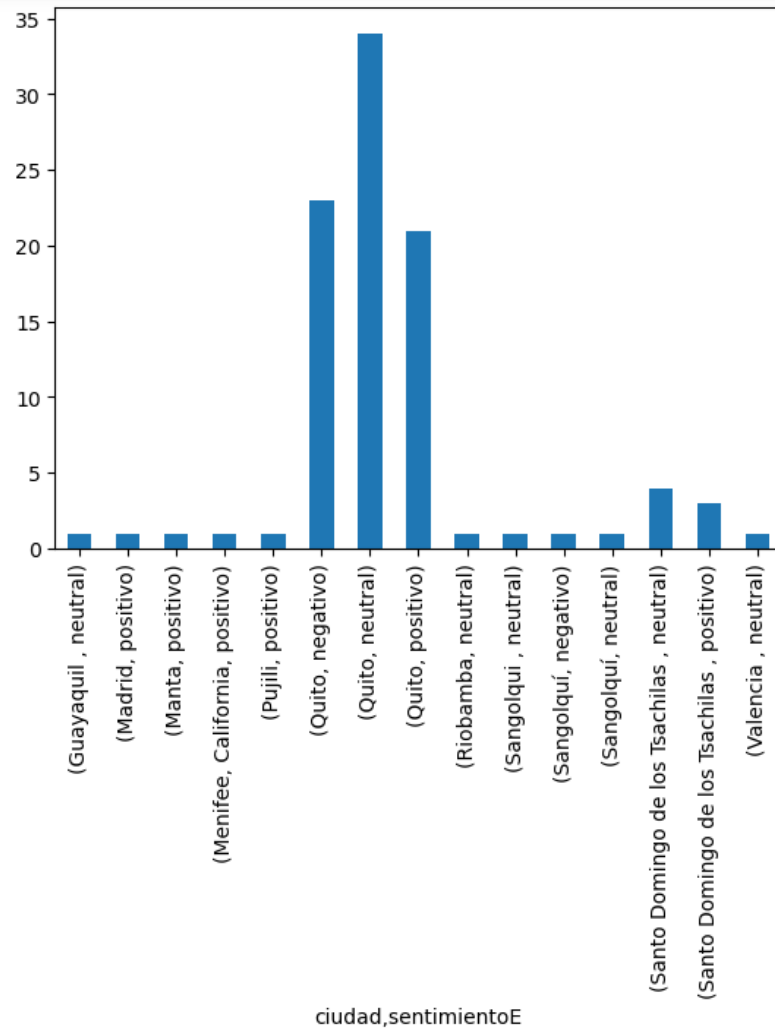


Figura 121: Análisis de sentimientos por localidad para la Pregunta 10 de la encuesta. (a)

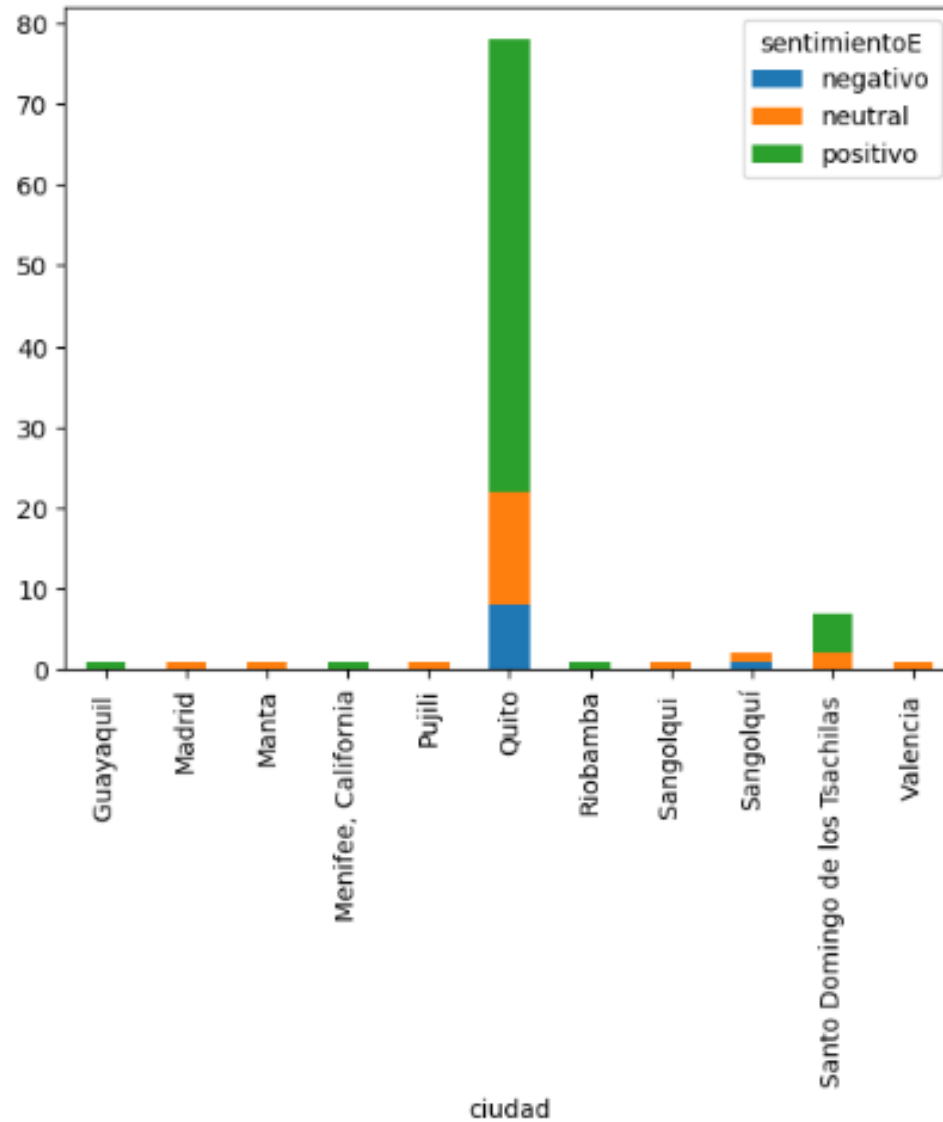


Figura 122: Análisis de sentimientos por localidad para la Pregunta 10 de la encuesta. (b)

La Figura 123 permite visualizar el análisis de sentimientos por genero.

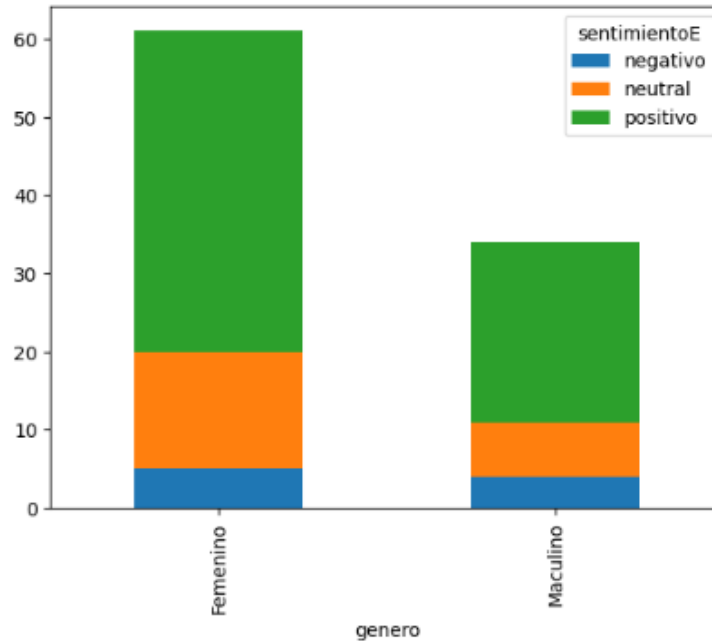


Figura 123: Análisis de sentimientos por genero para la Pregunta 10 de la encuesta.

Para tener una visión global de las palabras usadas en las opiniones sobre la pregunta 9 de la encuesta, se elaboró una nube de palabras global, como se muestra en la Figura 124.

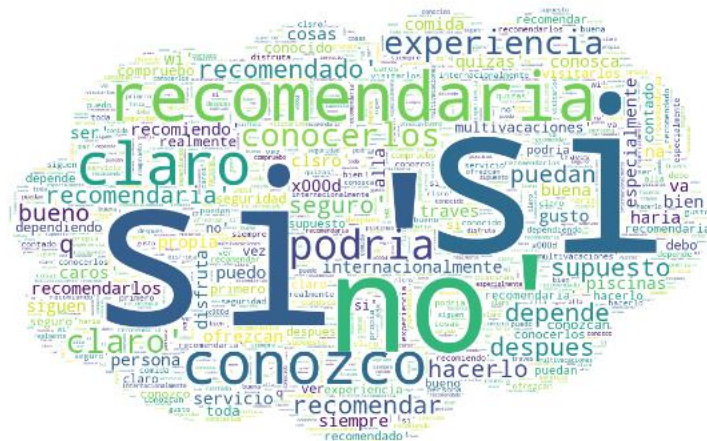


Figura 124: Nube de palabras para la Pregunta 10 de la encuesta.

Las Figuras 125, 126 y 127 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 9 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “sí”, en cuanto a la nube de palabras negativas, “conozco” es una de las palabras que se destaca.



Figura 125: Nube de palabras positivas para la Pregunta 10 de la encuesta.

Figura 96:



Figura 126: Nube de palabras neutras para la Pregunta 10 de la encuesta.

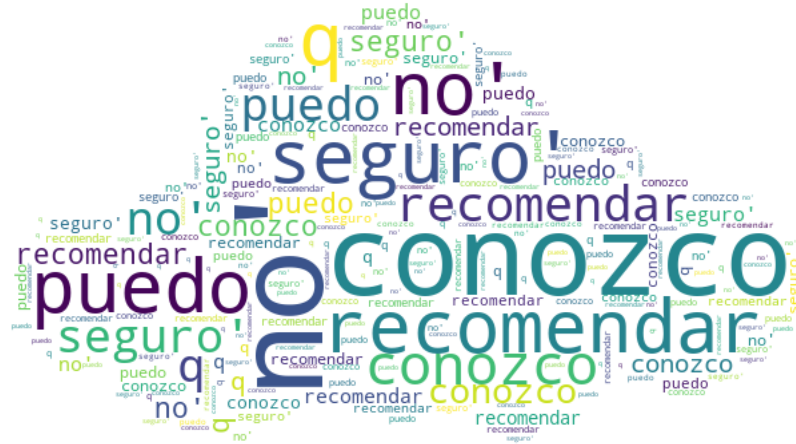


Figura 127: Nube de palabras negativas para la Pregunta 10 de la encuesta.

Pregunta 11:

4.4.9 Tiene algún comentario adicional?

En la figura 128, podemos observar los porcentajes de opiniones positivas, negativas y neutras para la pregunta 11.

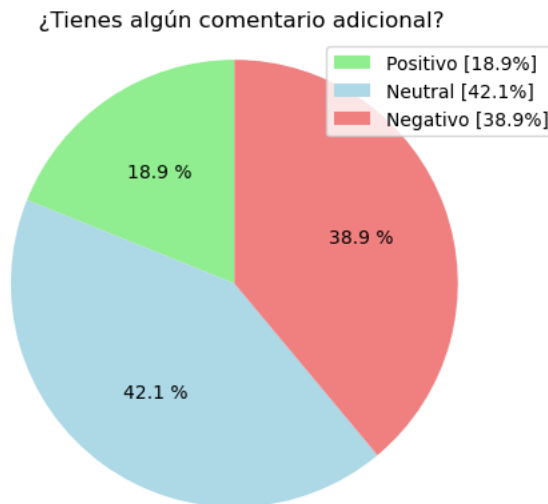


Figura 128: Análisis de sentimientos para la Pregunta 11 de la encuesta.

Las Figuras 129 y 130, muestran el análisis de sentimientos sobre las opiniones concernientes a la pregunta 11 por edad.

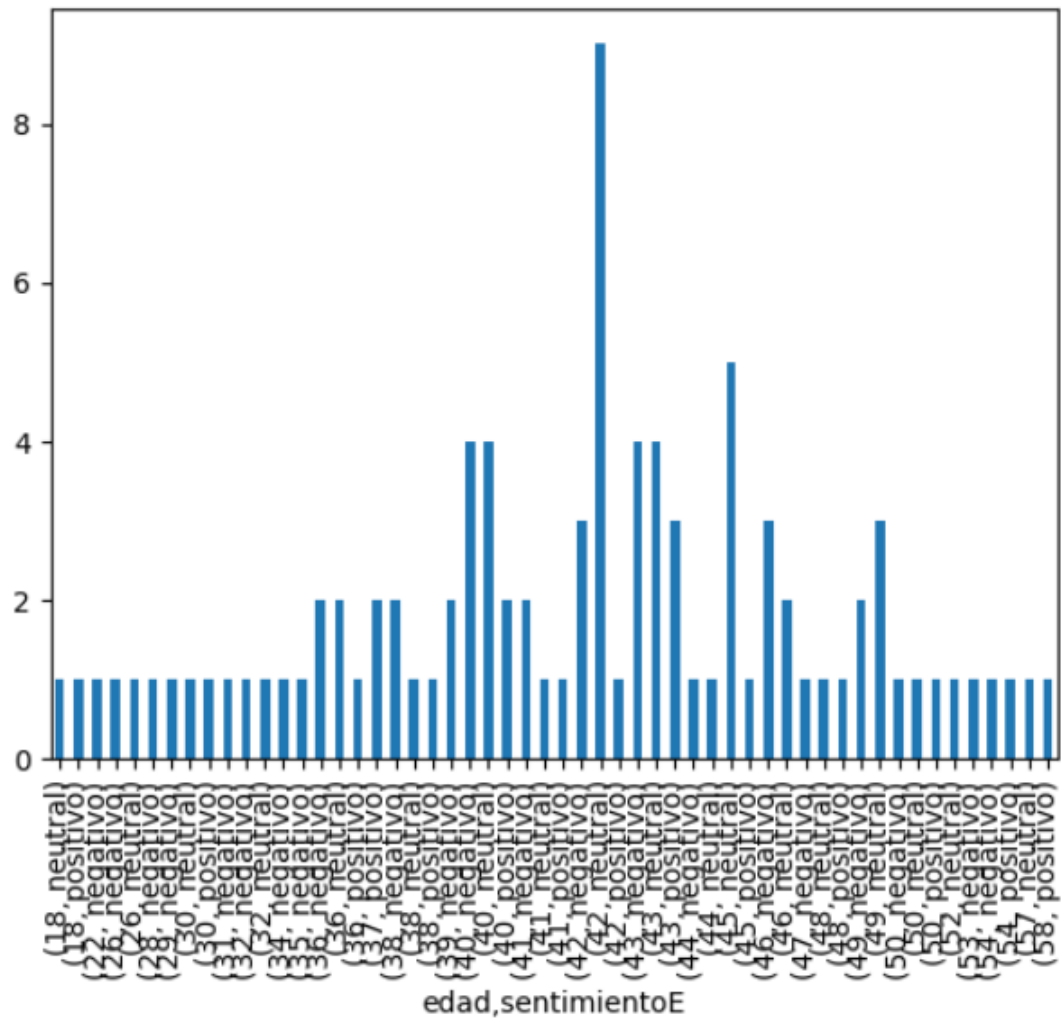


Figura 129: Análisis de sentimientos por edades para la Pregunta 11 de la encuesta. (a)

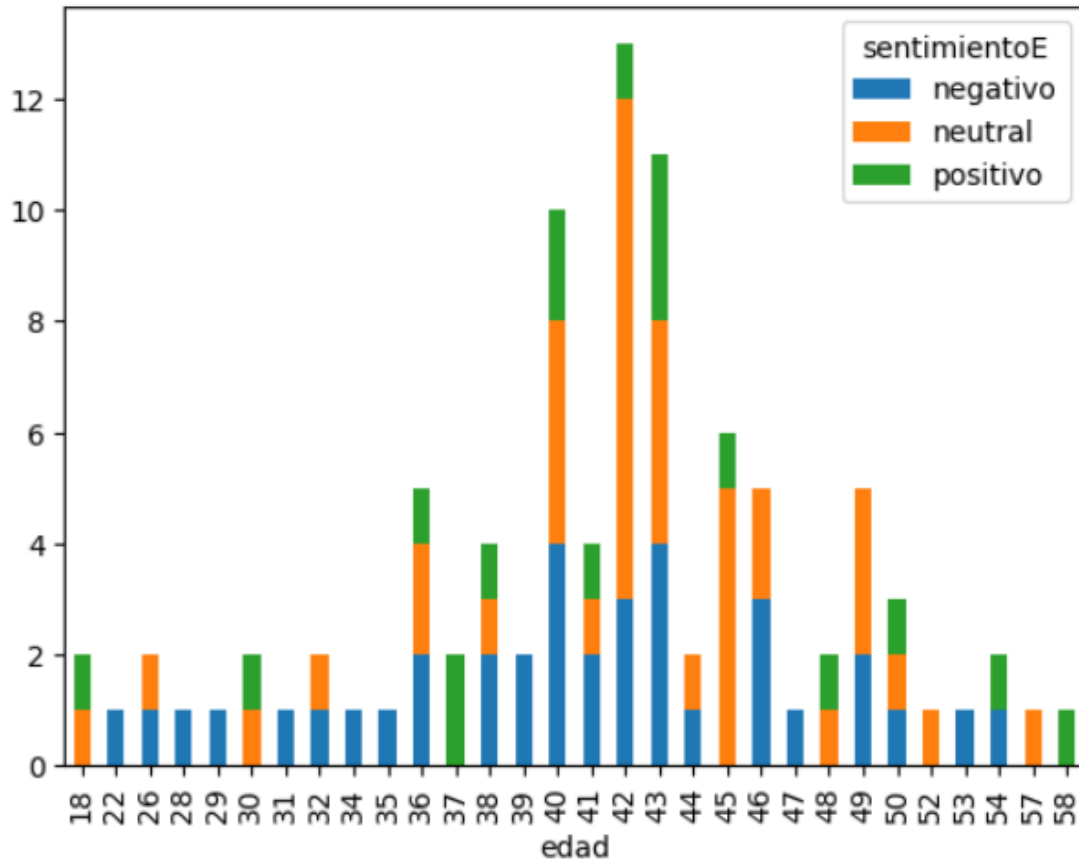


Figura 130: Análisis de sentimientos por edades para la Pregunta 11 de la encuesta. (b)

Las Figuras 131 y 132, muestran el análisis sentimientos sobre las opiniones concernientes a la pregunta 10 por localidad.

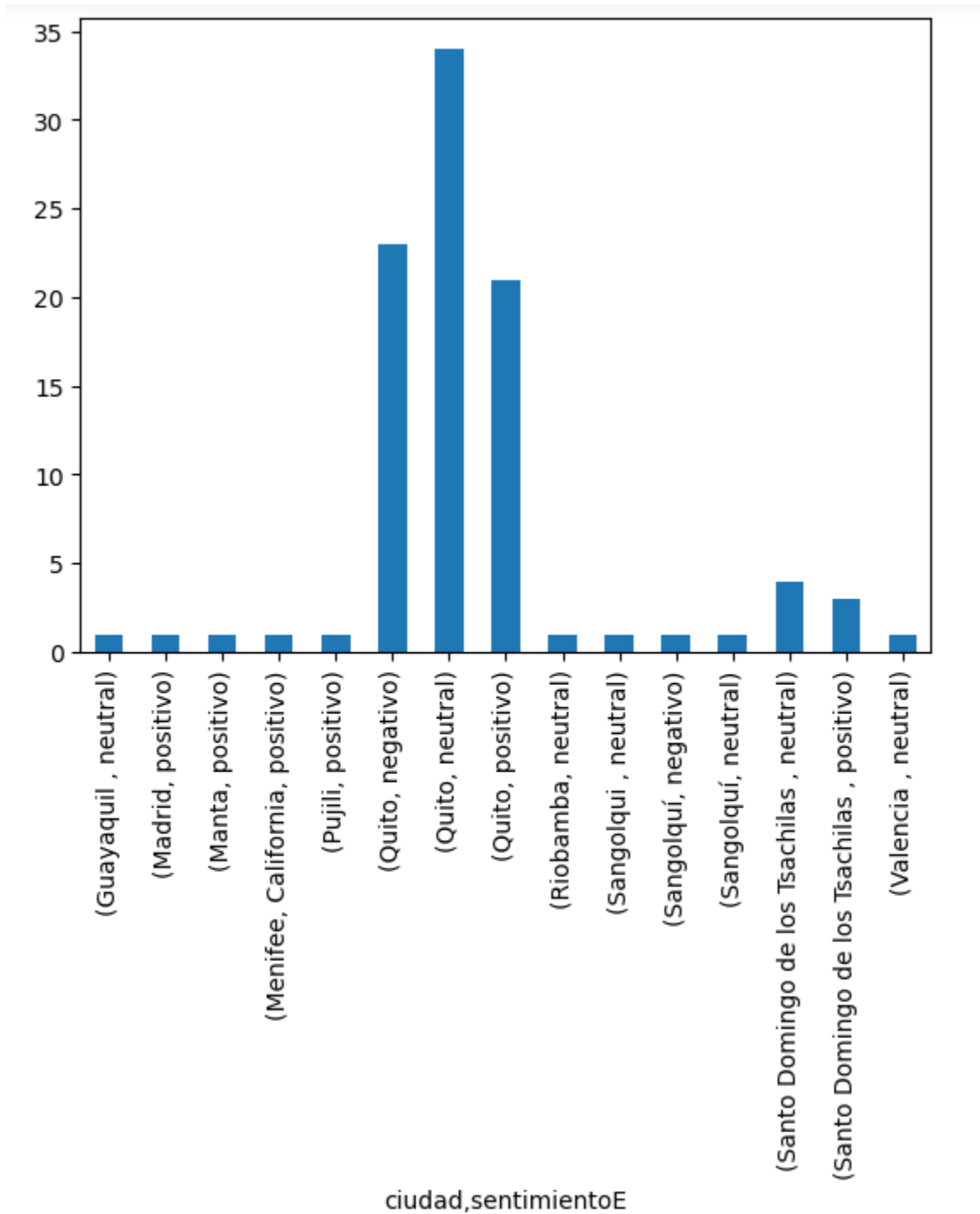


Figura 131: Análisis de sentimientos por localidades para la Pregunta 11 de la encuesta. (a)

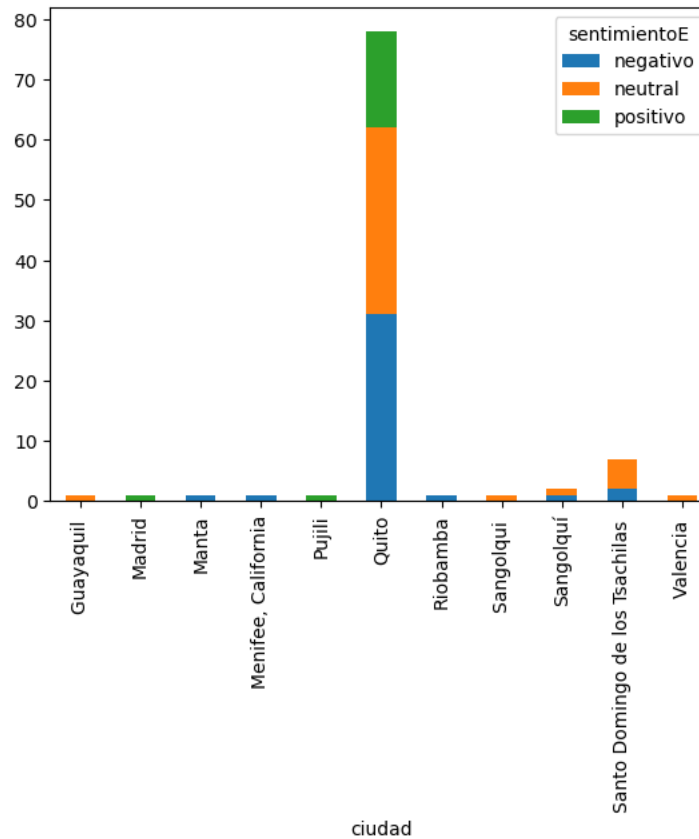


Figura 132: Análisis de sentimientos por localidades para la Pregunta 11 de la encuesta. (b)

La Figura 133 permite visualizar el análisis de sentimientos por genero.

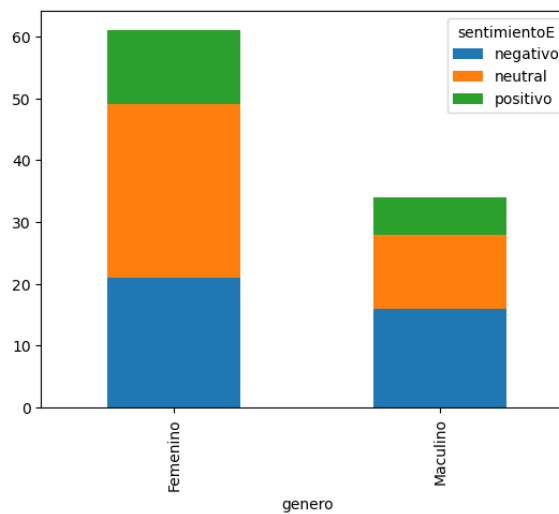


Figura 133: Análisis de sentimientos por genero para la Pregunta 11, de la encuesta.

Para tener una visión global de las palabras usadas en las opiniones sobre la pregunta 11 de la encuesta, se elaboró una nube de palabras global, como se muestra en la Figura 134.



Figura 134: Nube de palabras para la Pregunta 11 de la encuesta.

Las Figuras 135, 136 y 137 representan las nubes de palabras para las opiniones positivas, neutras y negativas, para la pregunta 11 de la encuesta, podemos observar que una de las palabras positivas más relevantes es “bueno”, en cuanto a la nube de palabras negativas, “ninguna” es una de las palabras que se destaca.

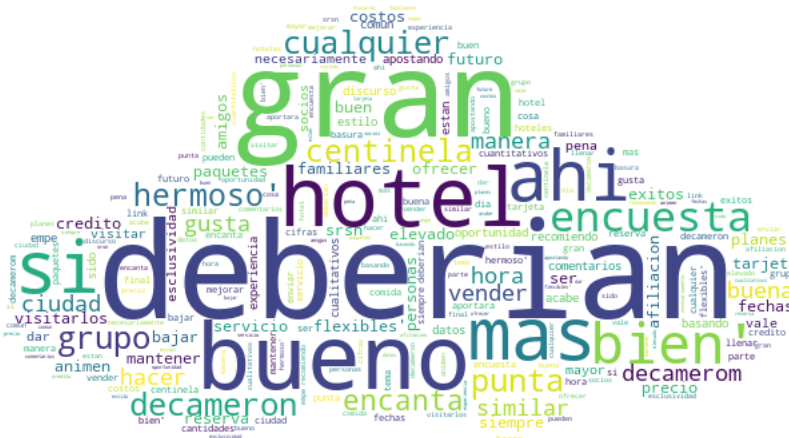


Figura 135: Nube de palabras positivas para la Pregunta 11de la encuesta.

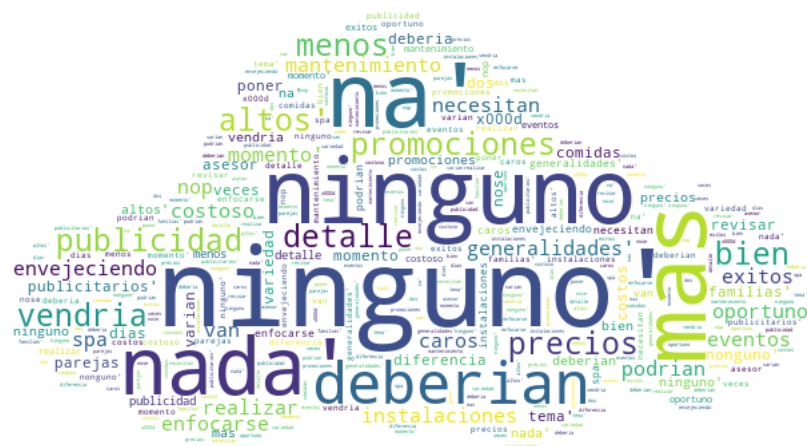


Figura 136: Nube de palabras neutras para la Pregunta 11 de la encuesta.



Figura 137: Nube de palabras negativas para la Pregunta 11 de la encuesta.

5 Conclusiones y Recomendaciones

5.1 Conclusiones

La aplicación del proceso de análisis de sentimientos desarrollado en el presente trabajo, sobre los datos de tweets y encuesta, permitió identificar información relevante sobre el nivel de satisfacción de los usuarios que usan los servicios de los Hoteles Decameron.

Se pudo evidenciar un alto porcentaje de aceptación de los servicios e instalaciones que ofrece el Hotel Decameron Mompiche. Además, durante los años 2012 al 2015, es el periodo donde hubo mayor número de opiniones positivas en Twitter, a partir del año 2016 el número de opiniones publicadas en Twitter tuvo una disminución paulatina notable.

El Hotel Decameron Punta Centinela, cuenta con amplia acogida por los huéspedes de Guayaquil y Santa Elena. El número máximo de comentarios positivos en Twitter se produjo en los años 2013, 2014, 2015.

Multivaciones tiene mayor porcentaje de opiniones negativas comparado con los porcentajes de opiniones negativas atribuidas al Hotel Decameron Mompiche y Decameron Punta Centinela de acuerdo con el análisis de sentimientos realizado sobre las opiniones publicadas en Twitter. En la encuesta se evidenció que muchas personas no conocen los servicios que proporcionan la institución.

Las provincias de la Amazonía Ecuatoriana no muestran presencia en las opiniones publicadas en Twitter. Al contrario de Quito y Guayaquil muestran alto protagonismo para los Hoteles Decameron Mompiche y Punta Centinela, respectivamente.

5.2 Recomendaciones

Resultaría beneficioso para Hoteles Decameron la ejecución periódica de un estudio de mercado a través del análisis de sentimientos.

Al encontrar el mayor porcentaje de opiniones negativas para Multivaciones, se sugiere, cambiar el enfoque en la presentación de la empresa.

Al identificarse los usuarios que tienen mayor relevancia en cuanto a opiniones sobre Hoteles Decameron y Multivaciones, se recomienda hacer un acercamiento con estas personas o empresas para procurar obtener una influencia positiva.

6 Referencias Bibliográficas

Aballay, L., Aciar, S., & Reategui, E. (2017). Método para detección de emociones desde foros utilizando Text Mining. *Campus Virtuales*, 6(1), 89-98.

Alfredo Bordignon, F. R., & Panessi, W. (2011). PROCESAMIENTO DE VARIANTES MORFOLÓGICAS EN BÚSQUEDAS DE TEXTOS EN CASTELLANO. *Revista Interamericana De Bibliotecología*, 24(1). Recuperado a partir de <https://revistas.udea.edu.co/index.php/RIB/article/view/7843>

Baldassarri Santalucía, S. (2016). Computación afectiva: tecnología y emociones para mejorar la experiencia del usuario. *Bit & Byte*, 2. <http://sedici.unlp.edu.ar/handle/10915/53441>

Baviera, T. (2017). Técnicas para el análisis del sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength. *Dígitos*, 1(3), 33-50.

Comm, J., & BURGE, K. (2009). *O poder do Twitter*. São Paulo: Gente.

Cortez Vásquez, A., Vega huerta, H., Pariona Quispe, J., & Huayna, A. M. (2009). Procesamiento de lenguaje natural. *Revista De investigación De Sistemas E Informática*, 6(2), 45–54. Recuperado a partir de <https://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/view/5923>

De Battista, A., Retamar, S., Cristaldo, P., Schab, E., Casanova, C., Ramos, L., ... & Herrera, N. E. (2021). Aplicaciones de procesamiento de lenguaje natural y Ciencia de Datos. In *XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja)*.

Decameron. (s.f.). CO Site - Decameron. Obtenido de Decameron: <https://www.decameron.com/es/98-espanol/co-site>

Decameron: 2022. Página Oficial. <https://www.decameron.com/es/ec-conocenos>

Dubiau, L., & Ale, J. M. (2013). Análisis de Sentimientos sobre un Corpus en Español: Experimentación con un Caso de Estudio. In *XIV Argentine Symposium on Artificial Intelligence (ASAI)-JAIIO 42 (2013)*..

Ekos, 2014. Hoteles Decameron tienen nuevo dueño. <https://www.ekosnegocios.com/articulo/hoteles-decameron-tienen-nuevo-dueno>

Fernandez, F. J. (2017). Estudio de mercado. Lulu. com.

Hernández, M. B., & Gómez, J. M. (2013). Aplicaciones de procesamiento de lenguaje natural. *Revista Politécnica*, 32.

Jungherr, A. (2015): *Analyzing Political Communication with Digital Trace Data: The Role of Twitter Messages in Social Science Research*. Cham: Springer International Publishing Switzerland.

Macias, A. C. R., Cobeña, L. E. S., & Valle, J. E. P. (2021). Análisis de sentimientos de las Elecciones Públicas del Ecuador basado en la red social Twitter. *Informática y Sistemas: Revista de Tecnologías de la Informática y las Comunicaciones*, 5(1), 7-16.

Orjuela Córdova, S., & Sandoval Medina, P. (2002). Guía del estudio de mercado para la evaluación de proyectos.

Twitter. (2022). Página Oficial Twitter. <https://developer.twitter.com/en/docs/twitter-api>

Vásquez, A. C., Quispe, J. P., & Huayna, A. M. (2009). Procesamiento de lenguaje natural. *Revista de investigación de Sistemas e Informática*, 6(2), 45-54.

Viégas, F.; Wattenberg, M. Timelines tag clouds and the case for vernacular visualization. *Interactions*, 2008, vol. 15, n. 4, pp. 49-52.