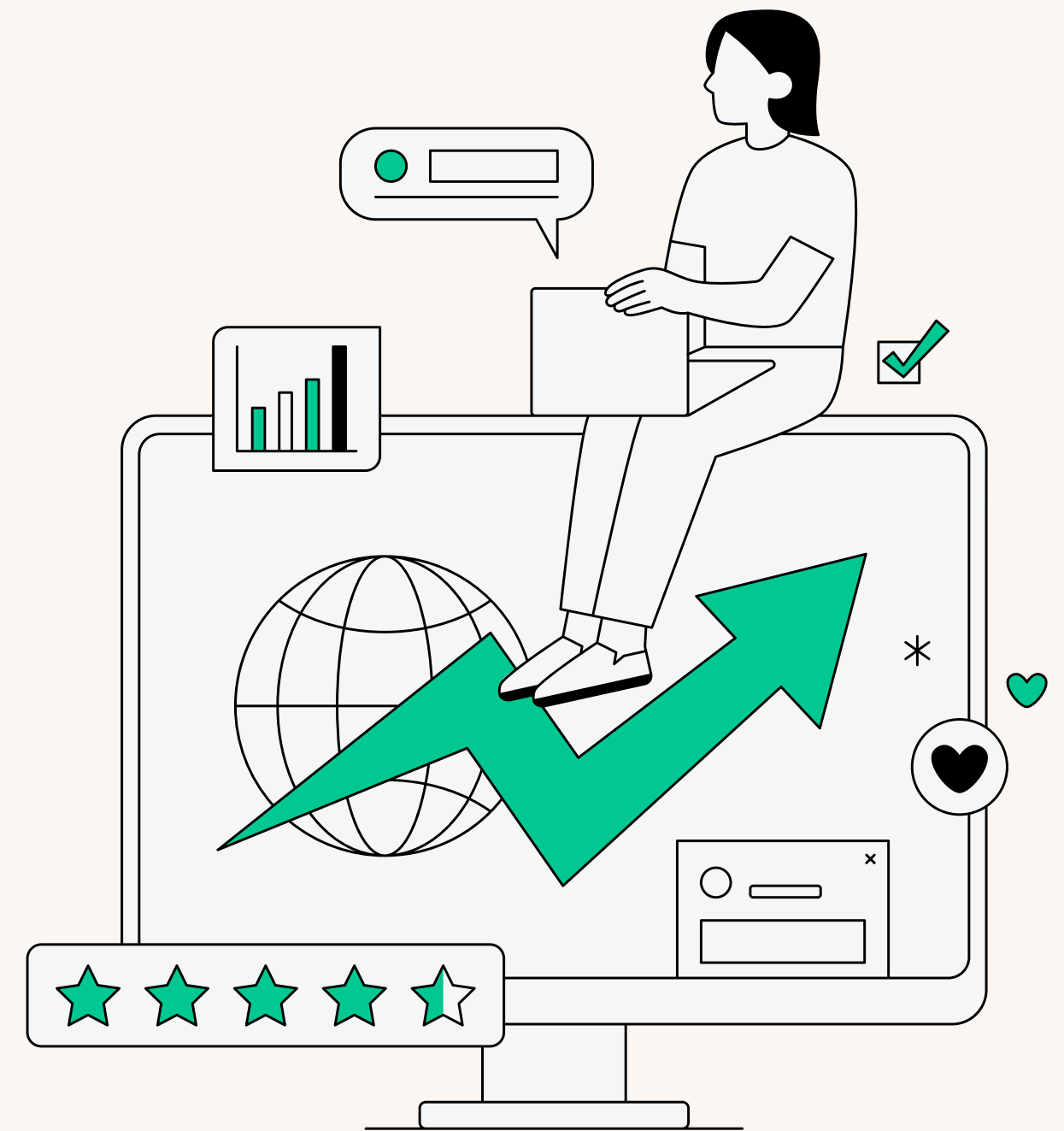


Análisis de Tweets

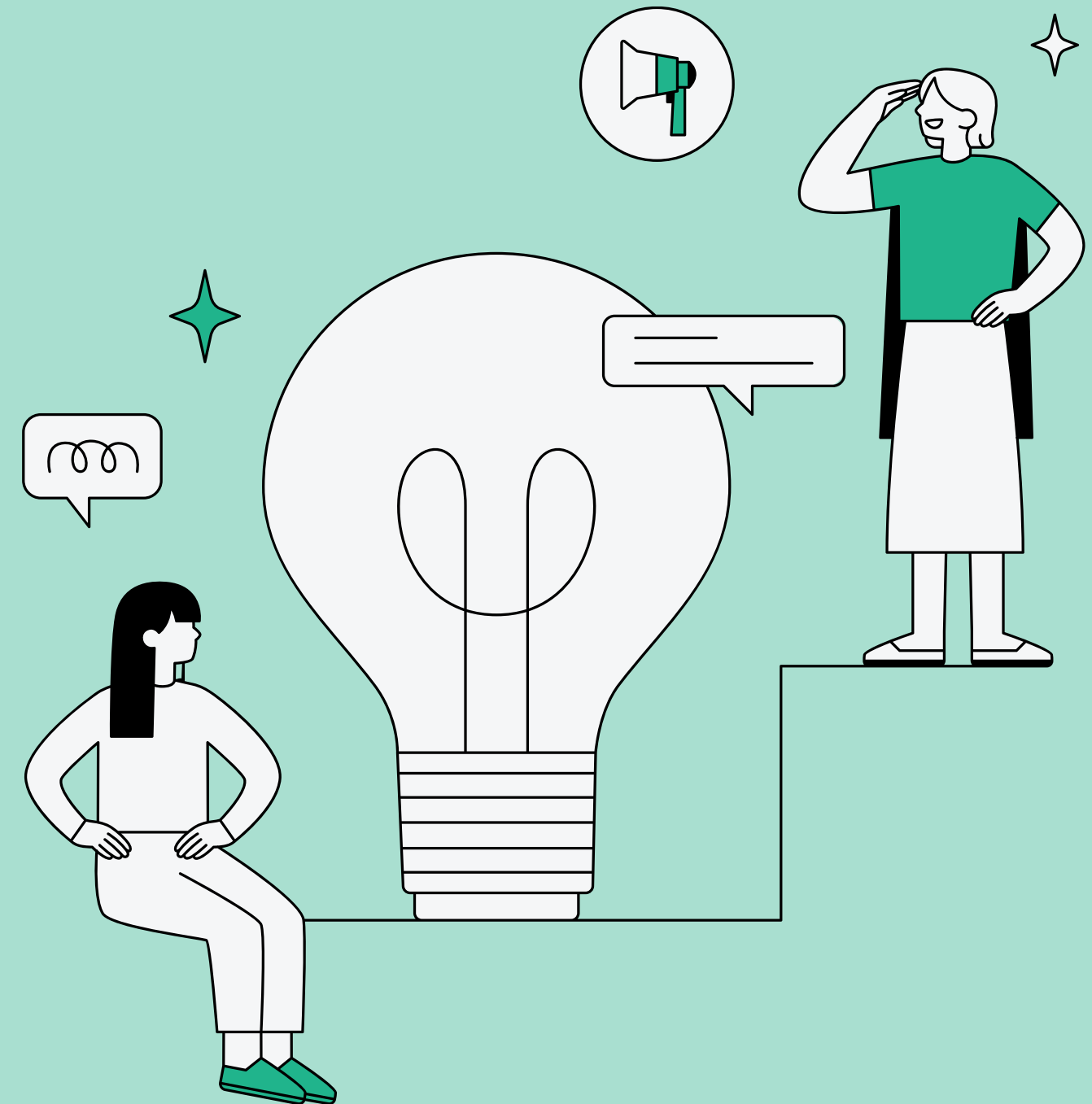
Realizado por:

- Álvaro Champi Huallpa



Introducción

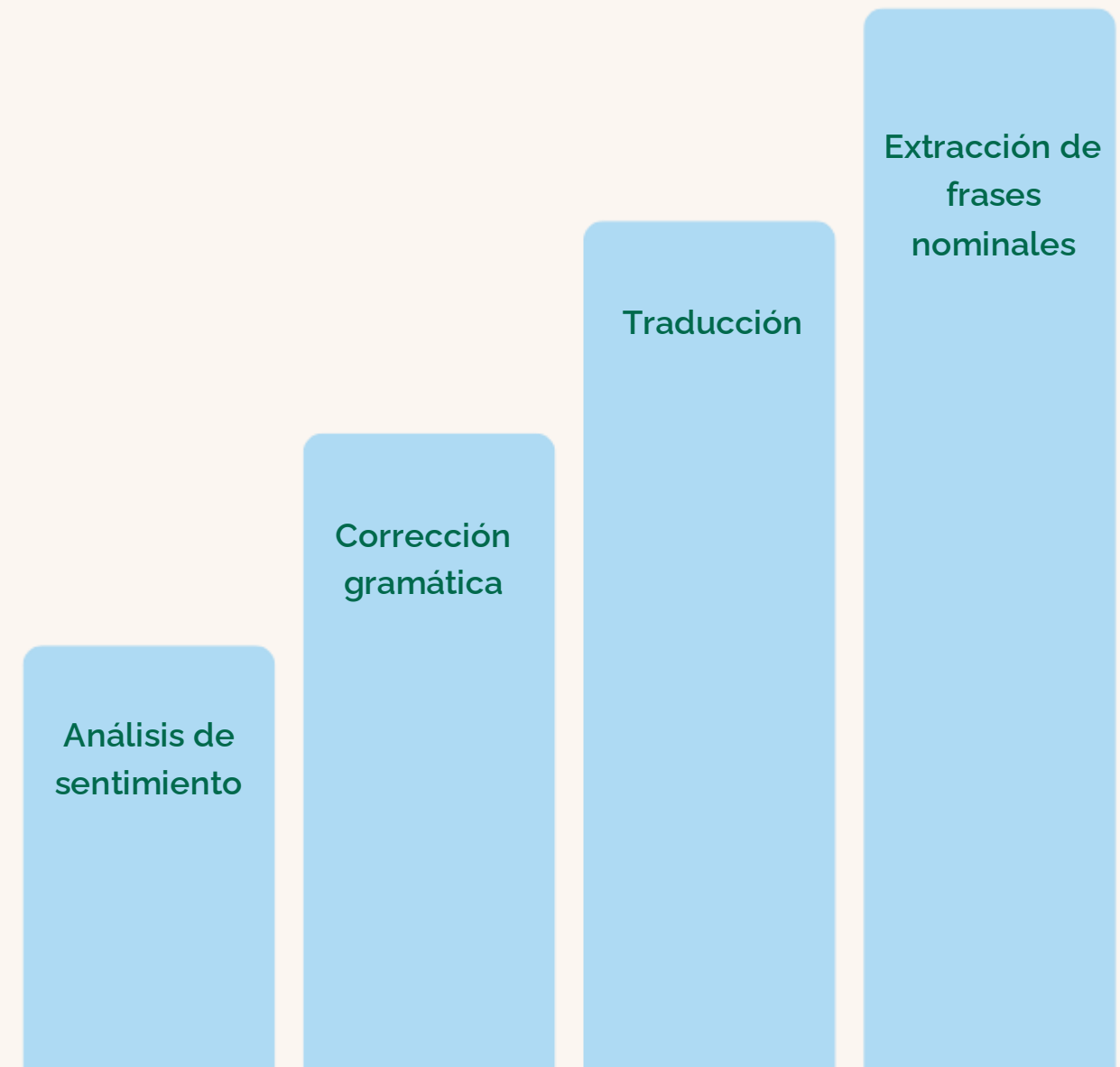
Twitter se ha convertido en un espacio clave donde profesionales de la salud, instituciones, medios y ciudadanos comparten noticias, opiniones y experiencias sobre temas de salud, como campañas de prevención, avances científicos, enfermedades emergentes y políticas sanitarias. Por ello, realizamos un análisis de las palabras más usadas en tweets relacionados con salud, procesando esta información para obtener resultados relevantes, los cuales se presentaron mediante dashboards que permiten visualizar las tendencias y preocupaciones más frecuentes.



Librería utilizada para el análisis

TextBlob es una librería de procesamiento de lenguaje natural (NLP) en Python que permite realizar análisis de texto de manera sencilla y eficiente.

TextBlob se utiliza principalmente para:





Recolección de datos

Utilizamos un dataset público para poder analizar los Tweets de salud:

Link: <https://www.kaggle.com/datasets/kazanova/sentiment140/data>



Análisis rápido

01.

Carga de data:

En el sistema cargamos la data de 150,000 tweets y se procede con la búsqueda de tweets referidos a temas de salud lo cual aplica la búsqueda de 183 tweets.

```
BUSCANDO TWEETS SOBRE 'HEALTH'...
Cargando dataset base...
Dataset cargado: 150,000 tweets
Tweets encontrados sobre salud: 183
```

02.

Procesamiento de datos:

- La información es procesada.
- Normalizamos las letras a minúsculas.
- Creamos una lista de términos clave
- Quitamos los URLs, menciones de usuario y hashtags.
- Aplicamos una Tokenización explícita para la eliminación de stopwords

03.

Resultado:

La información es procesada y validamos la opinión de la gente sobre la salud.

```
PROCESANDO TWEETS SOBRE SALUD...
Analizando sentimientos sobre salud...

ESTADÍSTICAS SOBRE TWEETS DE SALUD:
• Total de tweets analizados: 183
• Polaridad promedio: -0.022

¿Cómo habla la gente sobre salud? (Etiquetas originales):
• Negativo: 183 tweets (100.0%)

Análisis con TextBlob:
• Neutral: 80 tweets (43.7%)
• Positivo: 53 tweets (29.0%)
• Negativo: 50 tweets (27.3%)
```

Visualización de resultados

01.

Distribución de sentimientos originales

02.

Distribución de sentimiento con TextBlob

03.

Distribución de polaridad con TextBlob

04.

Análisis de sentimiento a lo largo del tiempo

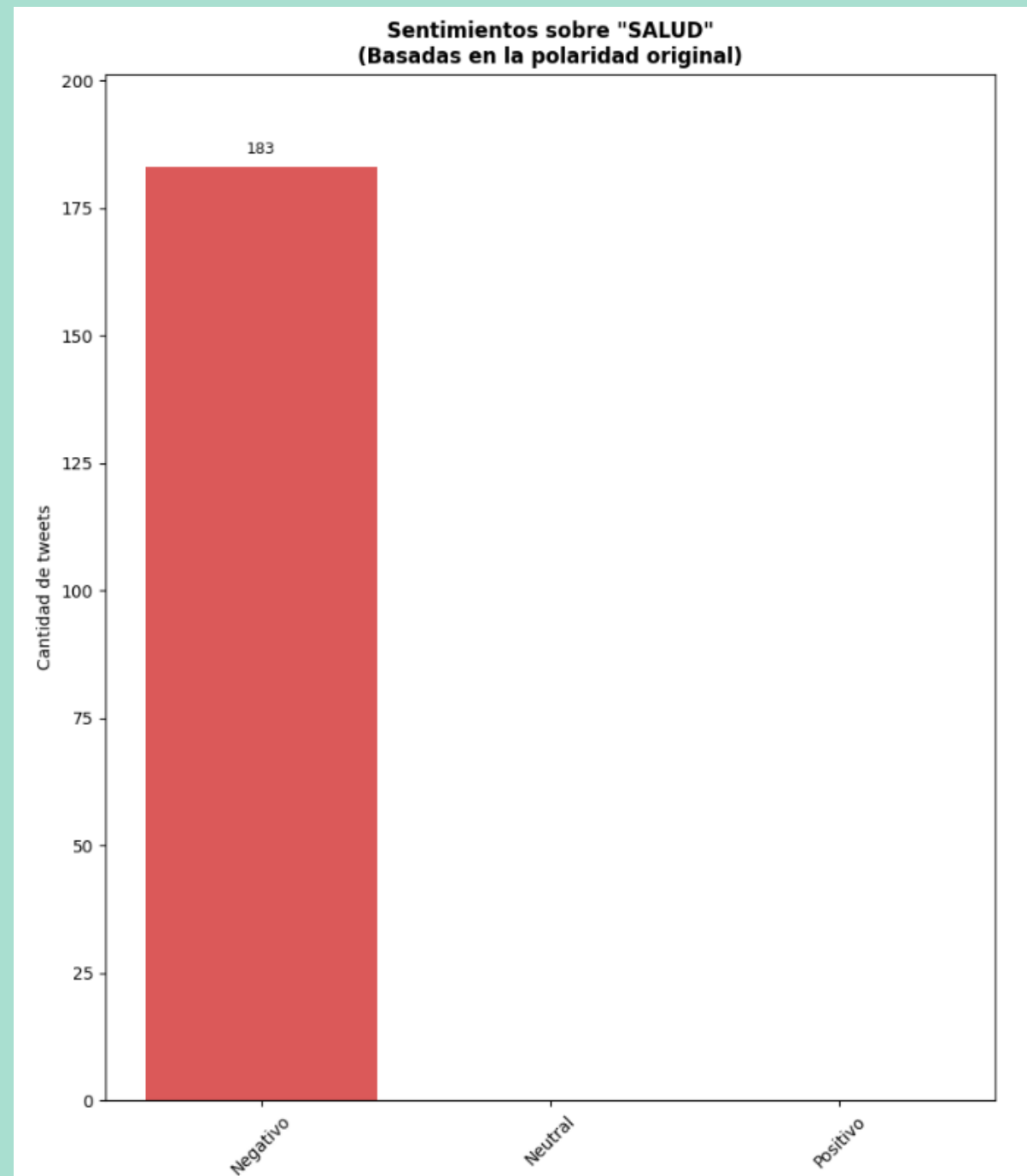
05.

Análisis de usuarios por sentimiento y subjetividad

06.

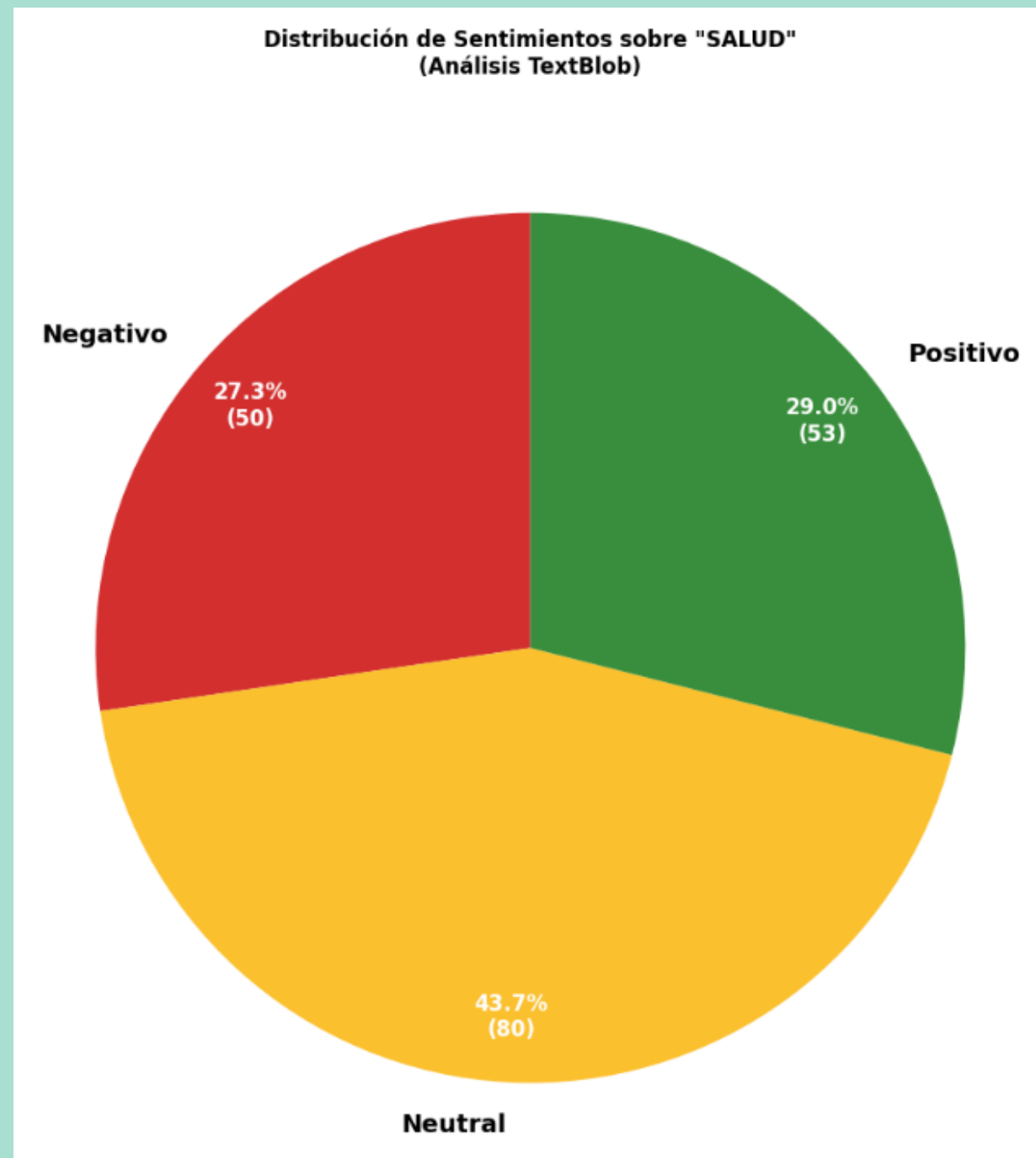
Comparación de métodos

01. Distribución de sentimientos originales



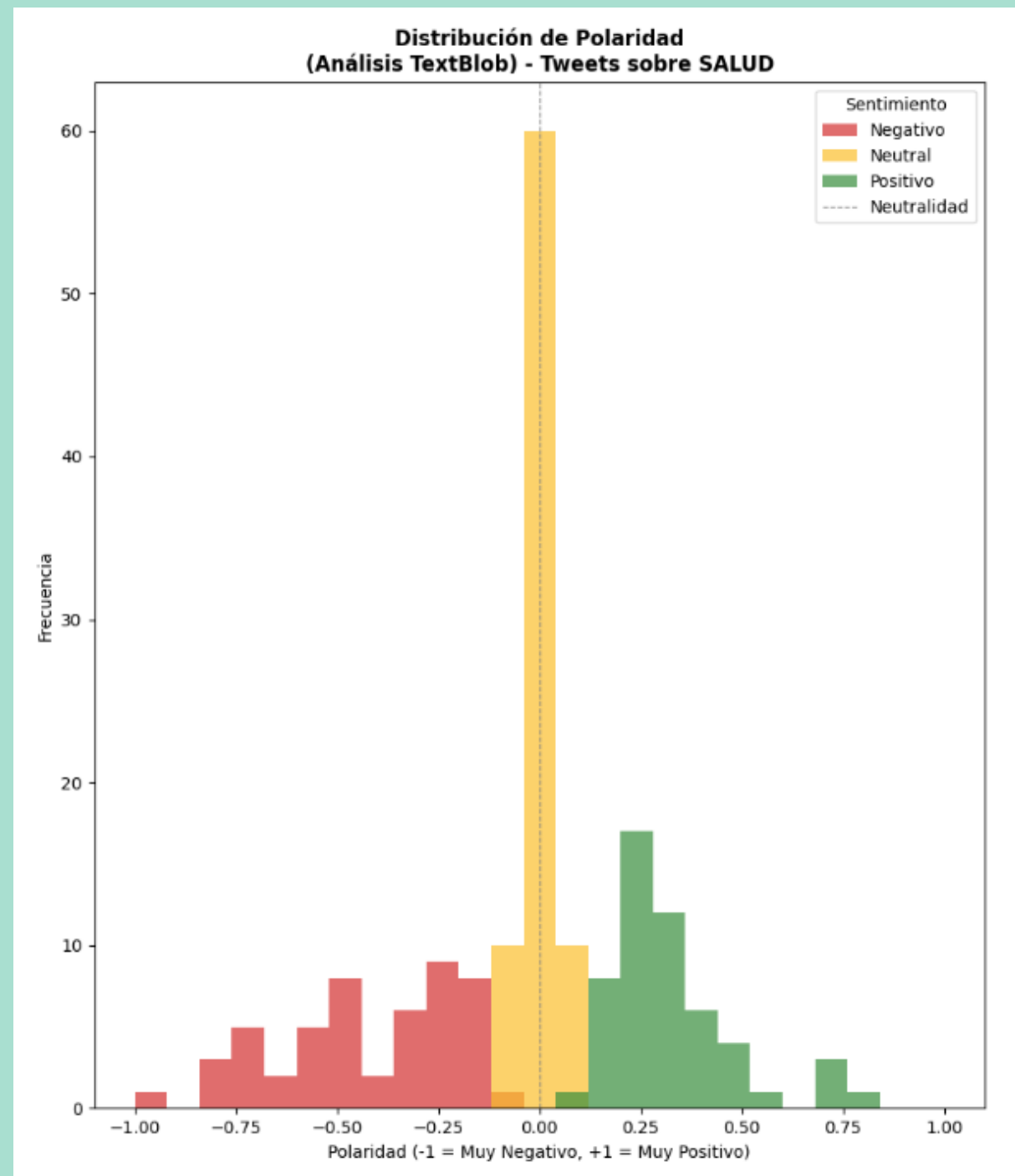
Este gráfico de barras indica que según la clasificación original del dataset **la mayoría de los tweets sobre "SALUD" son Negativos.**

02. Distribución de sentimiento con TextBlob



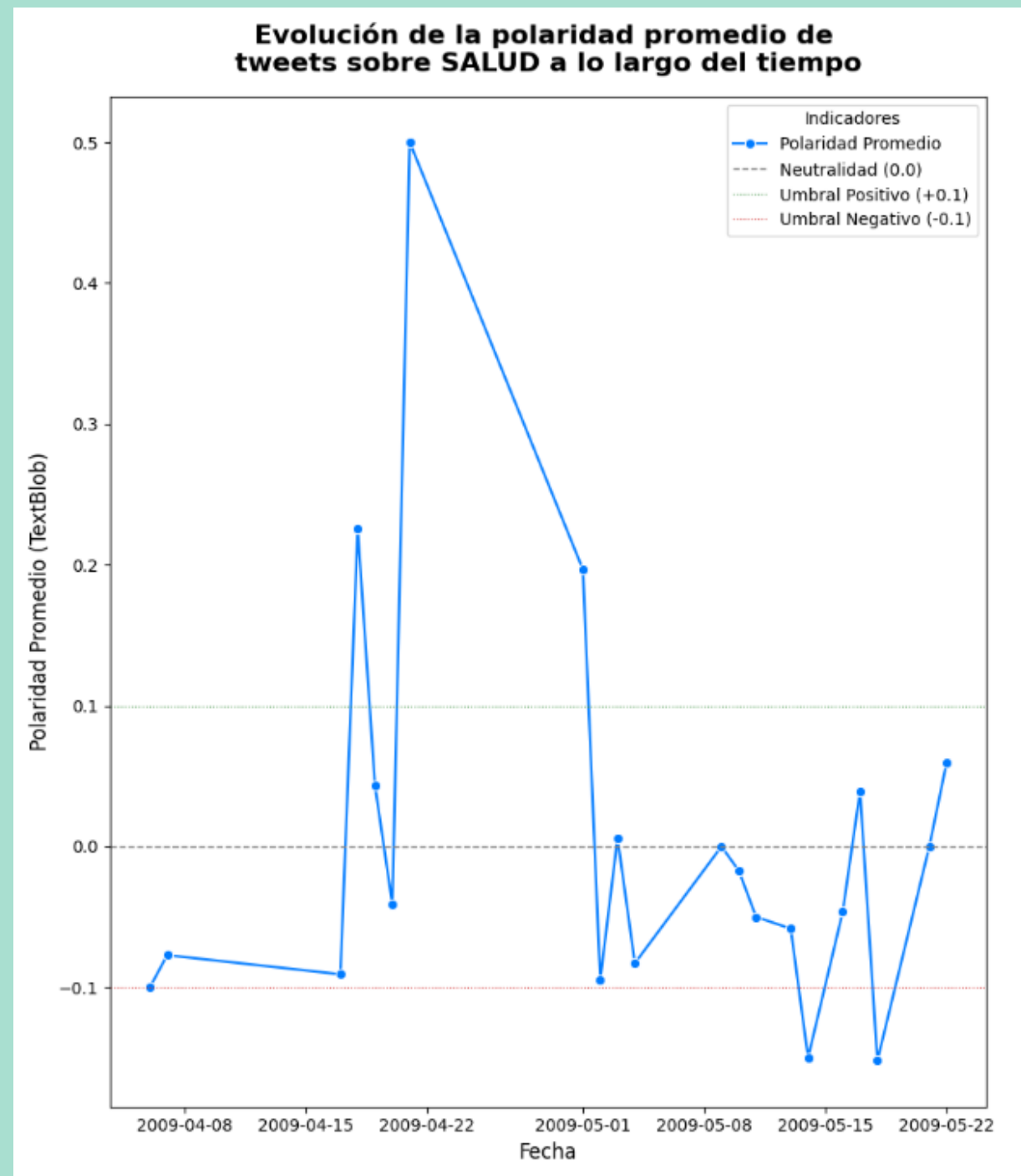
Este gráfico de pastel muestra que la **mayoría de los tweets sobre "SALUD" son Neutrales (63.7%)**, seguidos por Positivos (29.0%) y una menor proporción de Negativos (7.3%).

03. Distribución de polaridad con TextBlob



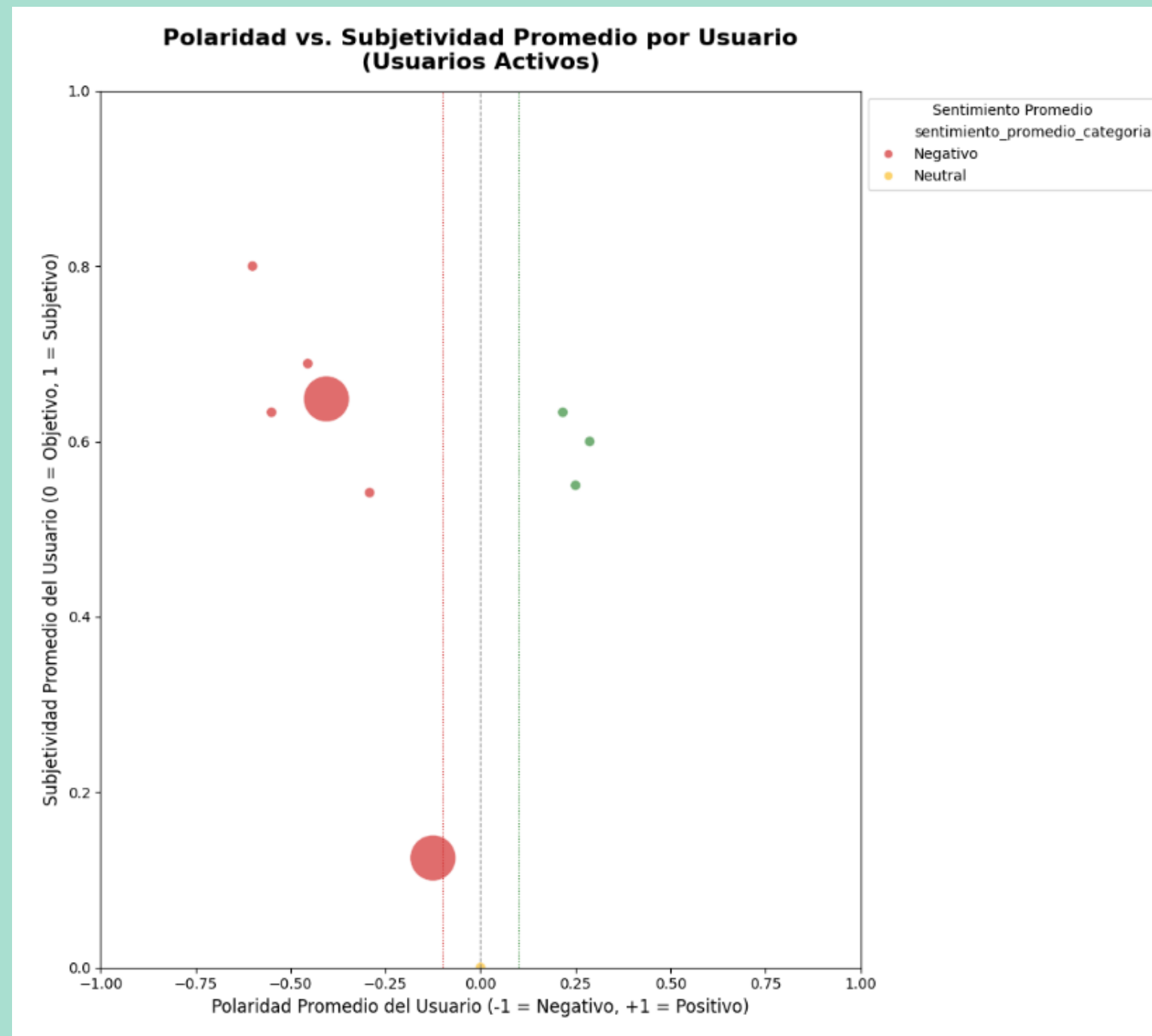
Esta es una distribución de frecuencias de la polaridad de los tweets analizados por TextBlob. **La mayoría de los tweets se concentran alrededor de la polaridad cero (neutral)**, lo que refuerza la gran proporción de tweets neutrales observada en el gráfico de pastel.

04. Análisis de sentimiento a lo largo del tiempo



Este gráfico de líneas muestra cómo el sentimiento promedio de los tweets sobre "SALUD" (según TextBlob) ha variado día a día. Se observa que **la polaridad promedio se mantiene mayormente en el rango negativo a neutral.**

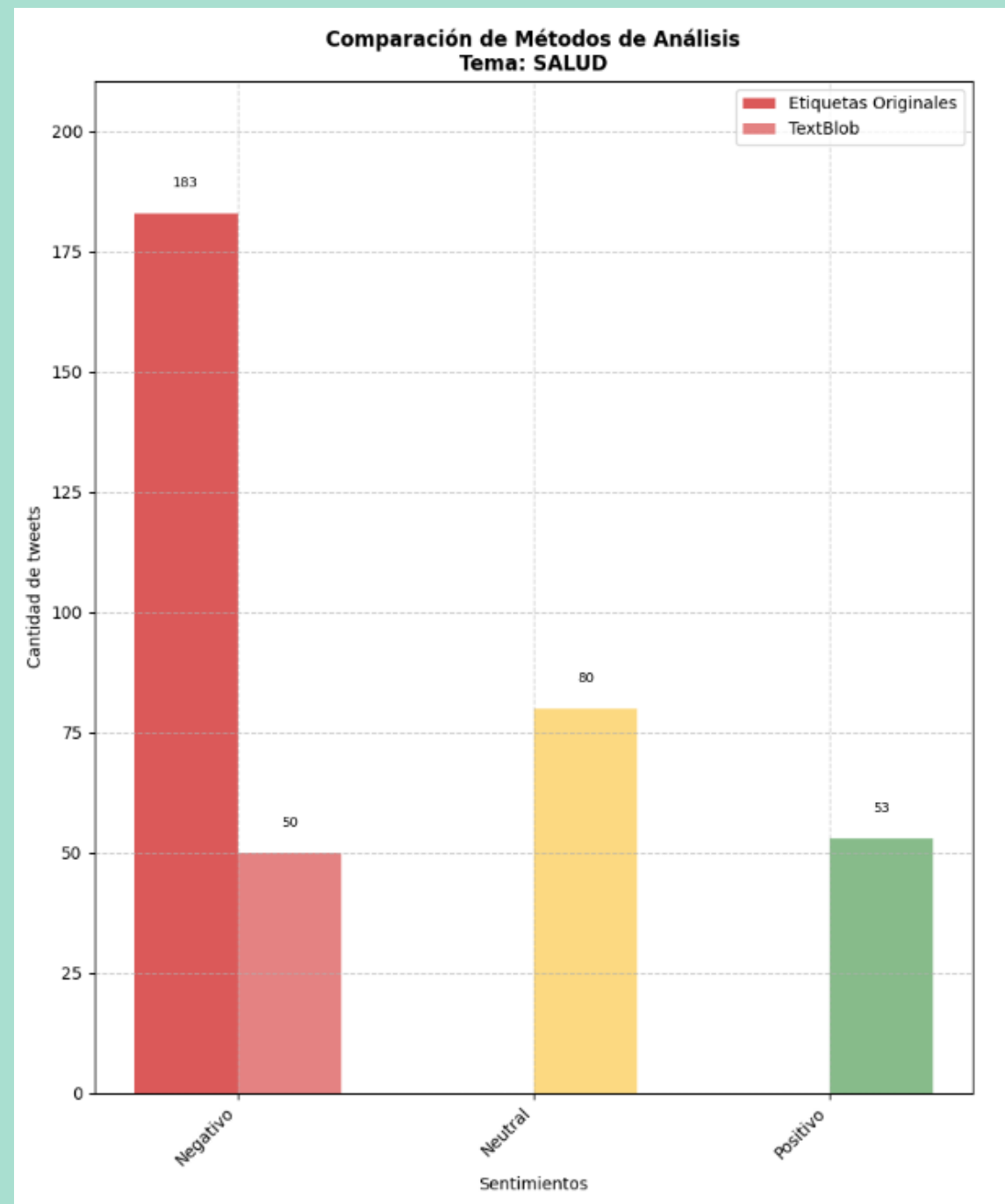
05. Análisis de usuarios por sentimiento y subjetividad



Este gráfico de dispersión **revela la relación entre la polaridad promedio y la subjetividad promedio de los usuarios** más activos. Los puntos grandes indican usuarios con más tweets.

- **Los usuarios negativos** (puntos rojos) tienden a ser predominantemente subjetivos (con valores Y mayores a 0.5).
- **Los usuarios positivos** (puntos verdes) también muestran una tendencia a ser subjetivos.

06. Comparación de métodos



Este gráfico de barras agrupadas compara directamente el conteo de tweets por sentimiento entre las etiquetas originales del dataset y el análisis de TextBlob.

- Las **etiquetas originales** muestran un **conteo muy alto de tweets Negativos** y prácticamente nada en otras categorías.
- **TextBlob**, en cambio, asigna un número **significativamente menor a "Negativo"**, **una gran cantidad a "Neutral"** y una cantidad notable a "Positivo".

Nube de palabras por tipo de sentimiento

01.

Palabras en Tweets positivos sobre SALUD

02.

Palabras en Tweets negativos sobre SALUD

03.

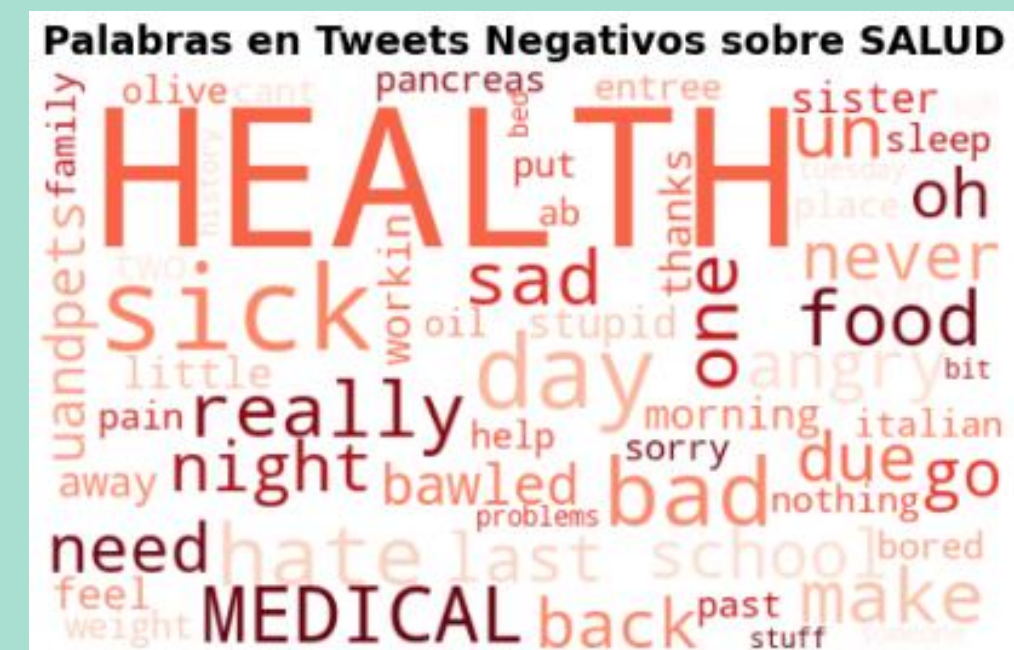
Palabras en Tweets neutrales sobre SALUD

- ♦ TEMAS DE SALUD ASOCIADOS CON SENTIMIENTO POSITIVO:
 - health: 50 menciones
 - medical: 4 menciones
 - care: 2 menciones
 - painting: 1 menciones
 - scared: 1 menciones
 - careful: 1 menciones
 - painful: 1 menciones
- ♦ TEMAS DE SALUD ASOCIADOS CON SENTIMIENTO NEGATIVO:
 - health: 48 menciones
 - sick: 9 menciones
 - medical: 4 menciones
 - pain: 2 menciones
 - care: 1 menciones
- ♦ TEMAS DE SALUD ASOCIADOS CON SENTIMIENTO NEUTRAL:
 - health: 72 menciones
 - medical: 11 menciones
 - care: 4 menciones

Palabras en Tweets positivos sobre SALUD

Palabras en Tweets negativos sobre SALUD

Palabras en Tweets neutrales sobre SALUD



Conclusiones y preguntas

- ¿Cuál es el sentimiento general sobre el tema?

La percepción general sobre SALUD en Twitter es: NEUTRAL ya que la interpretación de las opiniones sobre salud están balanceadas entre positivas y negativas

- ¿Qué palabras se repiten con frecuencia en los tweets negativos/positivos?

Se encuentra entre las palabras repetidas health, work, really.

- ¿Qué otros insights pudieron identificar?

La salud genera 183 conversaciones en la muestra analizada

El 45% de las conversaciones son neutrales

29.5% de tweets expresan optimismo sobre salud

25.7% de tweets expresan preocupación sobre salud

DATOS CLAVE:

- Total de tweets sobre salud analizados: 183
- Sentimiento predominante: Neutral (43.7%)
- Polaridad promedio: -0.022

CONCLUSIÓN PRINCIPAL:

La percepción general sobre SALUD en Twitter es: NEUTRAL/MIXTA

Interpretación: Las opiniones sobre salud están balanceadas entre positivas y negativas

INSIGHTS:

- La salud genera 183 conversaciones en la muestra analizada
- El 44% de las conversaciones son neutrals
- 29.0% de tweets expresan optimismo sobre salud
- 27.3% de tweets expresan preocupación sobre salud

Muchas gracias

