

Análisis del impacto de publicaciones en redes sociales de alto tráfico para el desarrollo de campañas de marketing

Universidad del Rosario

Escuela de Ingeniería Ciencia y Tecnología

Juan Urbano
Andrés Bonilla
Juan Mora

Contenidos

Planteamiento	2
Reglas de Negocio	2
Entidades, Atributos y Relaciones	3
Entidades	3
Atributos	4
Relaciones	4
Modelo Entidad Relación	5
Escenarios de análisis	7
Análisis de rendimiento de publicaciones por plataforma	7
Análisis de la Audiencia	7
Análisis Temporal de las Publicaciones	7
Evaluación del impacto de las campañas de marketing	7
Análisis de tendencias de contenido	7
Desarrollo de Análisis	8
Conclusiones	15

Repositorio GitHub

<https://github.com/andre5fbonilla/Ingenieria-de-datos>

Planteamiento

Se plantea trabajar con una base de datos que contiene información amplia y detallada sobre una alta variedad de publicaciones en las principales redes sociales, las de mayor tráfico, que tienen un enfoque publicitario y de mercadeo principalmente.

Los datos permiten identificar y obtener gran variedad de precisiones sobre las publicaciones y a las interacciones que estas provocan; también incluyendo información sobre quienes publican y quienes reaccionan a dichas publicaciones. Lo anterior facilita el desarrollo de estrategias para alcanzar un mayor impacto sobre una cantidad de público objetivo más grande.

La base de datos a trabajar se obtuvo del sitio Web Kaggle, en la cual se realizó la última actualización el 13 de abril del presente año, lo que permite trabajar con datos sobre un periodo de tiempo reciente y, consecuentemente, se obtiene información aplicable en nuestra época actual, con oportunidad de uso en distintas áreas de comercio. Esta base de datos cuenta con 100.000 líneas y 24 columnas, es decir, con 100.000 registros y 24 categorías o características de análisis.

Link de la base de datos: <https://www.kaggle.com/datasets/aliredaelblgiyh/social-media-engagement-report/data>

Reglas de Negocio

1. Validación de la Plataforma: Se debe verificar que las plataformas registradas sean válidas.
2. Integridad del post_ID: Cada publicación debe tener un identificador único que no se repita en el conjunto de datos.
3. Formato de Fechas y Horas: Las fechas y horas de publicación deben estar en un formato estándar y ser válidas.
4. Consistencia en los Tipos de Publicación: Los tipos de publicación deben seguir un conjunto predefinido de categorías y estar correctamente etiquetados.
5. Precisión en las Métricas de Participación: Las métricas como likes, comentarios, compartidos, impresiones y alcance deben ser numéricas y no negativas.
6. Cálculo Correcto de la Tasa de Participación: La tasa de participación debe calcularse correctamente y estar dentro de un rango válido entre 0 y 1 (0% a 100%).
7. Datos Demográficos de la Audiencia: Se deben verificar los datos demográficos de la audiencia para garantizar su coherencia y validez.
8. Asociación Correcta con Campañas de Marketing e Influencers: Las publicaciones deben estar correctamente asociadas con las campañas de marketing y los influencers, si corresponde.
9. Análisis de Sentimientos: Se debe verificar que el análisis de sentimientos se haya realizado correctamente y que los resultados sean coherentes con el contenido de la publicación.
10. Privacidad y Cumplimiento Normativo: Se deben tomar medidas para proteger la privacidad de los datos, especialmente en lo que respecta a la información demográfica y de comportamiento de la audiencia.

Entidades, Atributos y Relaciones

Entidades

Entidad	Descripción	Atributos
Post	Representa toda la información relacionada con cada publicación	<ul style="list-style-type: none">• Plataforma• post_ID• post_type• post_content• post_timestamp• post_date• post_day• Time• post_period_time• campaign_ID• Sentiment• Influencer_ID
Indicator	Representa los indicadores sobre el impacto de la cada publicación	<ul style="list-style-type: none">• Likes• Num_Comments• Shares• impressions• Reach• engagement_rate
Audience	Representa toda la información de la audiencia o público objetivo	<ul style="list-style-type: none">• audience_age• age_group• audience_gender• audience_location• audience_continent• audience_interest

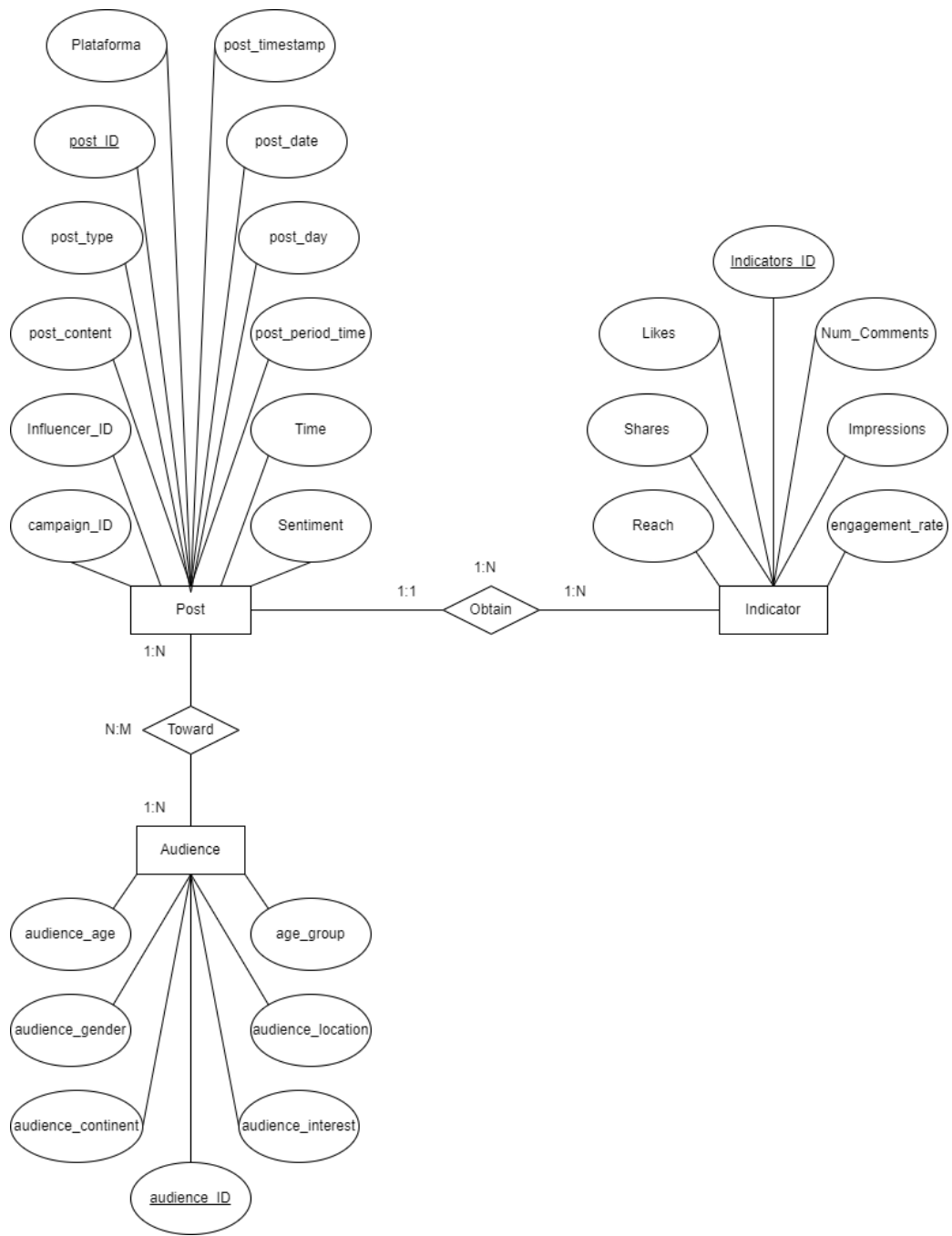
Atributos

Columna	Descripción
Plataforma	Red social en la que se publicó.
post_ID	Código identificador de la publicación.
post_type	Tipo de la publicación.
post_content	Contenido de la publicación.
post_timestamp	Fecha y hora en que se realizó la publicación.
post_date	Fecha en que se realizó la publicación.
post_day	Día en que se realizó la publicación.
Time	Hora en que se realizó la publicación.
post_period_time	Periodo del día en que se realizó la publicación.
Likes	Cantidad de likes recibidos.
Num_Comments	Cantidad de comentarios recibidos.
Shares	Cantidad de veces que se compartió la publicación.
impressions	Cantidad de veces que se mostró la publicación.
Reach	Cantidad de usuarios únicos que vieron la publicación.
engagement_rate	Tasa de participación en la publicación.
audience_age	Rango de edad de la audiencia objetivo.
age_group	Grupo etario de la audiencia objetivo.
audience_gender	Genero de la audiencia objetivo.
audience_location	Ubicación geográfica de la audiencia objetivo.
audience_continent	Continente donde se ubica la audiencia objetivo.
audience_interest	Intereses o datos demográficos de la audiencia objetivo.
campaign_ID	Código identificador de la campaña de marketing asociada a la publicación.
Sentiment	Sentimiento expresado en la publicación.
influencer_ID	Código identificador del influencer asociado a la publicación, si corresponde.

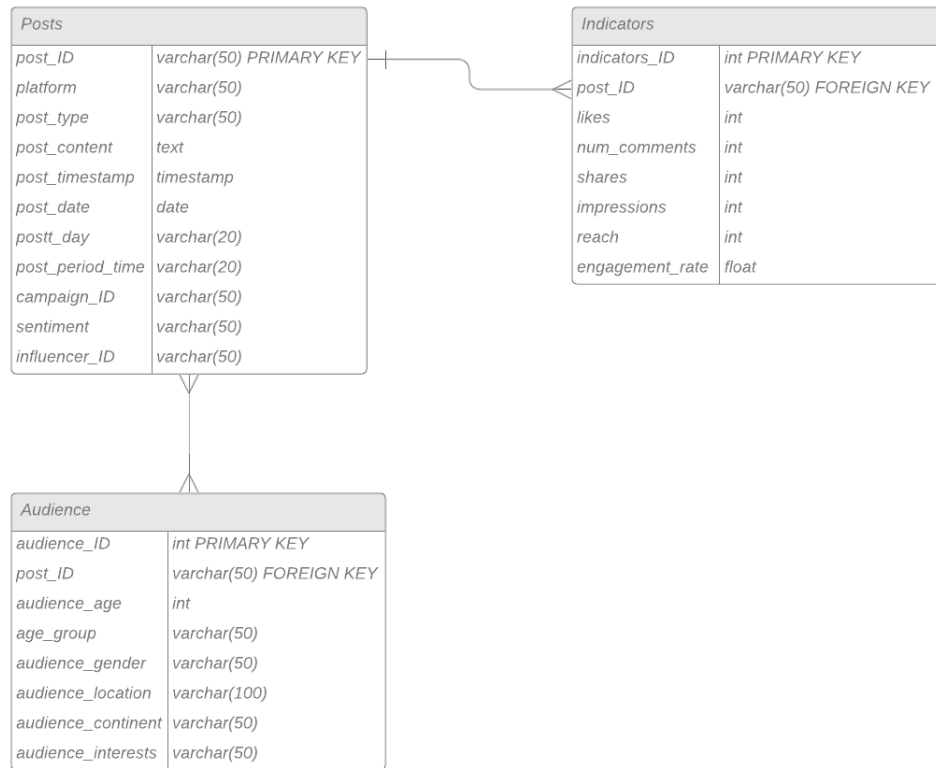
Relaciones

Entidades	Relación
Post - Indicator	Obtain
Post - Audience	Toward

Modelo Entidad Relación



Modelo Relacional



Escenarios de análisis

Análisis de rendimiento de publicaciones por plataforma

Se evaluará el impacto de las publicaciones en las diferentes plataformas de redes sociales. Al comparar métricas como likes, comentarios, compartidos, impresiones y el alcance entre las plataformas en la tabla *Indicator*; se propone identificar qué plataformas son más efectivas para alcanzar objetivos de marketing y engagement para una campaña. Este análisis permitirá enfocar las campañas de publicidad en las plataformas más rentables y optimizar la estrategia de contenido para cada una de ellas.

Análisis de la Audiencia

Se hará un enfoque de análisis intencional a los datos de la tabla *Audience* para entender mejor al público objetivo de las diferentes publicaciones. Este análisis permitirá adaptar mejor el contenido a las preferencias de la audiencia, analizando variables como la edad, el género, la ubicación, los intereses de la audiencia, e inclusive parámetros como lo es la plataforma, hora de publicación y demás características de las cuales puede depender el alcance de dicho post.

Análisis Temporal de las Publicaciones

Se propone un análisis temporal de las publicaciones, utilizando los datos de tiempo en la tabla *Post*. Este análisis nos permitirá entender cuándo es el mejor momento para publicar, teniendo en cuenta tanto el día de la semana como el momento del día (mañana, tarde o noche). Se evaluará si las publicaciones realizadas en diferentes días y momentos tienen diferentes niveles de impacto.

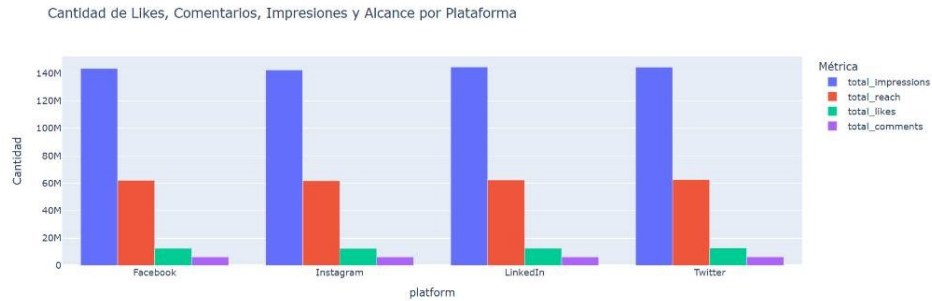
Evaluación del impacto de las campañas de marketing

Mediante este análisis, se evaluará el rendimiento de las publicaciones asociadas a campañas de marketing específicas y/o que fueron realizadas por influencers. Al comparar métricas de *Indicator*, se podrá determinar el impacto de cada una en términos de participación de la audiencia y percepción de la marca. Estos hallazgos ayudarán a evaluar si es rentable realizar una campaña de marketing involucrando influencers.

Análisis de tendencias de contenido

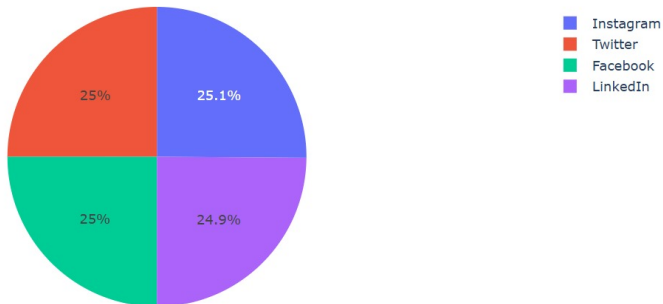
Se busca identificar los temas y tipos de contenido que generan más engagement y participación por parte de la audiencia; analizar qué tipo de publicaciones o temas específicos son más populares entre tus seguidores. Esto te ayudaría a ajustar la estrategia de contenido para crear publicaciones más relevantes y atractivas.

Desarrollo de Análisis



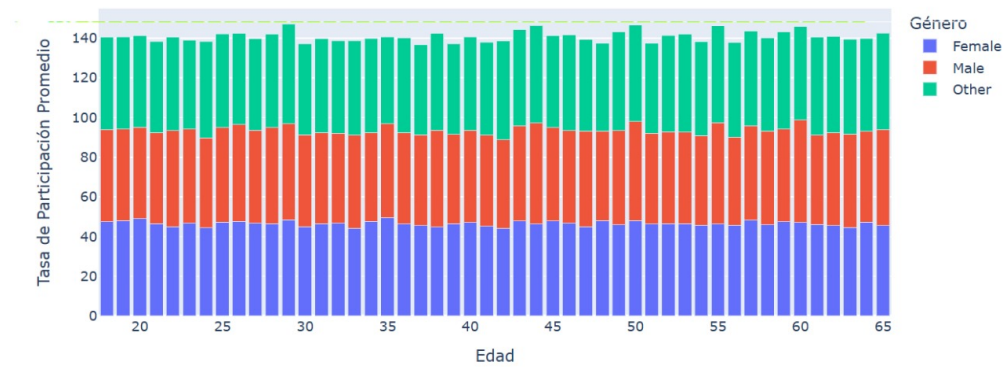
Respecto a las redes sociales en las que se desarrollaron las publicaciones, se puede apreciar que cuentan todas con indicadores muy poco dispersos, en cuanto a la cantidad de reacciones e interacciones del público objetivo sobre las publicaciones; lo cual no es necesariamente negativo, ya que sugiere que el rendimiento es consistente en cada una de las plataformas.

Tasa de Participación Promedio por Plataforma



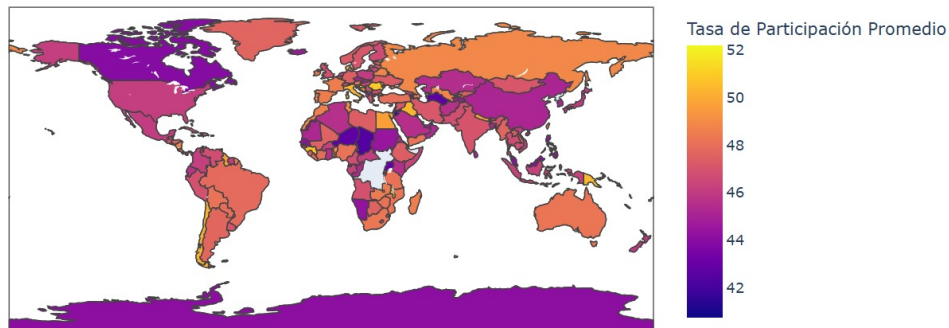
Facebook e Instagram lideran en términos de impresiones totales y alcance total, lo que indica su eficacia para llegar a un público amplio. Facebook, con su vasta base de usuarios y herramientas publicitarias, genera un mayor número de impresiones. Sin embargo, Instagram se destaca por tener la tasa de compromiso promedio más alta, lo que nos dice que, aunque su alcance puede ser menor, los usuarios interactúan más con el contenido en esta plataforma. Esto se refleja en sus indicadores, que son significativamente más altos en Instagram. Esta diferencia puede deberse a la naturaleza visual del contenido de Instagram, que estimula una mayor interacción por parte de los usuarios. Por lo que, mientras que Facebook se caracteriza por la visibilidad masiva, Instagram cuenta con una mayor cantidad de interacciones y conexión con el público objetivo.

Tasa de Participación Promedio por Edad y Género del Público Objetivo



En cuanto a la edad del público objetivo, se puede concluir que dentro de un rango de edad, a partir de los 30 años hasta los 60 años, se ve mayor enganche y compromiso en las publicaciones, por parte de hombres; con un promedio entre los distintos grupos de edad del 50.9% de participación.

Tasa de Participación Promedio por País



Según lo observado, el país con la tasa de participación más baja en las publicaciones de redes sociales es Martinique, con una tasa de participación promedio de aproximadamente 40.76%. Este porcentaje es considerablemente inferior al de otros países incluidos en el análisis, lo que sugiere la necesidad de adaptar o mejorar las estrategias de publicidad en redes sociales para este público específico y de otras regiones con porcentajes cercanos. Para aumentar la efectividad de las campañas publicitarias, se podría considerar investigar las preferencias y comportamientos del público objetivo en esta región; tal vez incluyendo enfoques más personalizados.

Tasa de Participación Promedio por Hora de Publicación

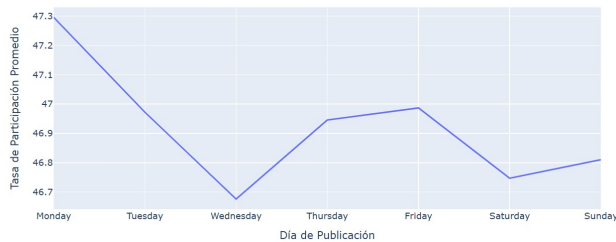


Tasa de Participación Promedio por Periodo de Tiempo del Post



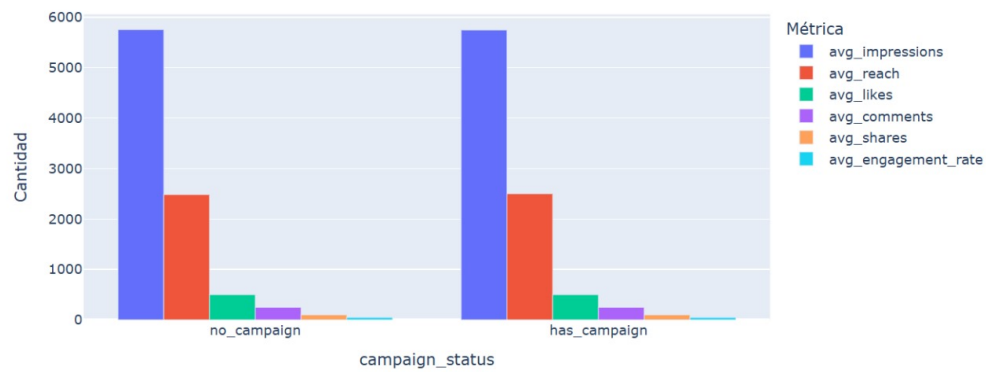
Sobre un rango horario, es posible evidenciar que, el que beneficia con mayor interacción es desde las 11:00, hasta las 16:00, ya que es susceptible al enfoque y la manipulación para optimizar la cantidad de vistas e impresiones. Según la gráfica, las horas con la tasa de participación más baja en redes sociales son las 6:00 y las 18:00, con tasas de 45.9% y 46.1% respectivamente. Por lo que estos horarios podrían no ser los más adecuados para publicar. La baja participación en estas horas del día podrían deberse a que a las 6:00 muchas personas están ocupadas con su rutina matutina, y a las 18:00, pueden estar terminando su jornada laboral o en tránsito. Sería óptimo evitar estos horarios y considerar aquellos entre las 11:00 y las 16:00, donde se encuentran pausas laborales o la hora del almuerzo.

Tasa de Participación Promedio por Día de Publicación



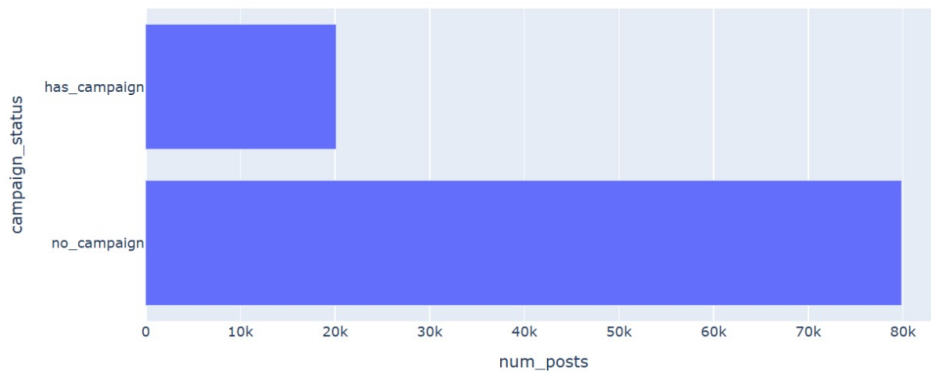
Respecto a los días de la semana en los que se logra mayor enganche, los lunes y los viernes destacan. Por el contrario, el miércoles presenta el peor potencial de compromiso, con una tasa de participación de aproximadamente 46.7%. Esto nos dice que las publicaciones realizadas los miércoles podrían no ser tan efectivas en términos de interacción de la audiencia en comparación con las realizadas los lunes y viernes. Para una mejor participación de la audiencia, sería mejor enfocar las publicaciones más importantes para los lunes y viernes, evitando los miércoles o dejándolos para contenido menos crítico.

Promedio de Indicadores de Posts con y sin Campaña de Publicidad



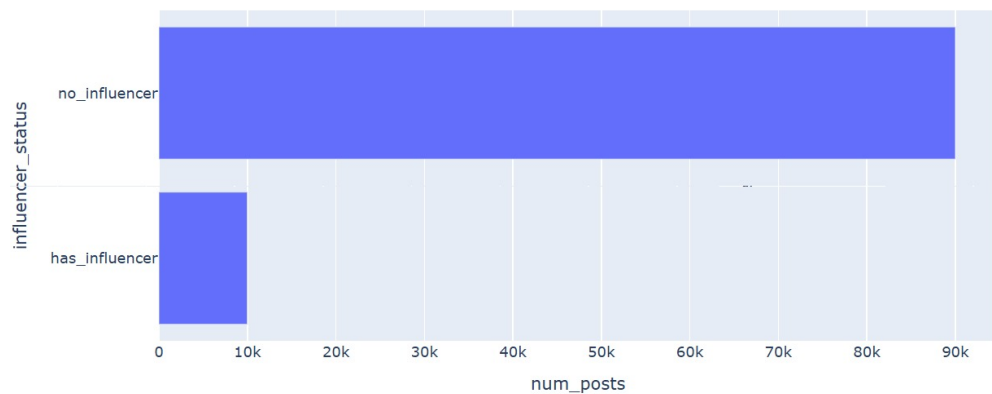
Sorprendentemente, las publicaciones que no contaban con una campaña, fueron los que contaron con mayor enganche posteriormente. Lo anterior podría sugerir que proponer un contenido más orgánico y espontáneo es de mayor estímulo o gusto para el público.

Cantidad de posts con y sin campaña publicitaria



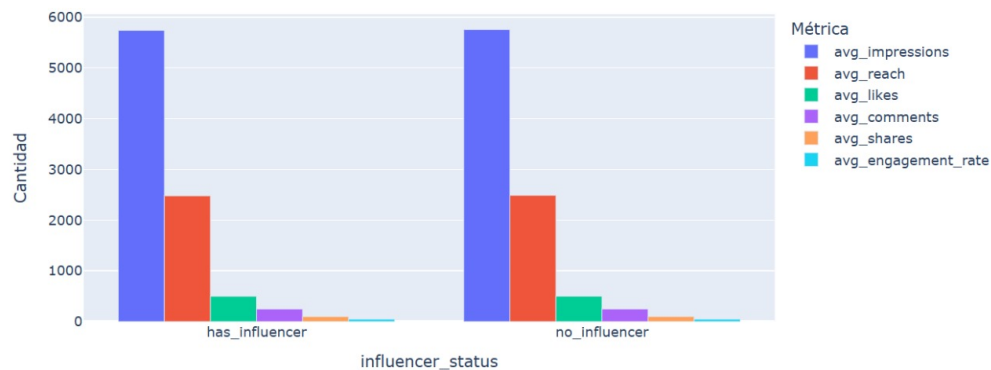
La presencia de una campaña publicitaria no tiene un impacto significativo en los indicadores de rendimiento de las publicaciones. Sin embargo, es importante destacar que hay una cantidad considerablemente mayor de publicaciones sin campaña (79,868) en comparación con aquellas que sí tienen campaña (20,132). Esto sugiere que, aunque las campañas publicitarias pueden no ser cruciales para mejorar la tasa de participación, la mayoría de las publicaciones se realizan sin el apoyo de estas campañas.

Cantidad de posts con y sin Influencer asociado



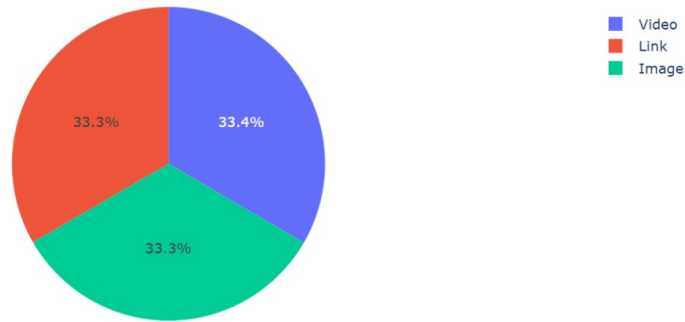
En cuanto al uso o el apoyo de un influencer para el desarrollo de una publicación, aquellas publicaciones que no contaban con influencer relacionaron obtuvieron mejores resultados, lo que podría indicar que la audiencia prefiere contenido auténtico, contrario a contenido genérico.

Promedio de Indicadores de Posts con y sin Influencer asociado



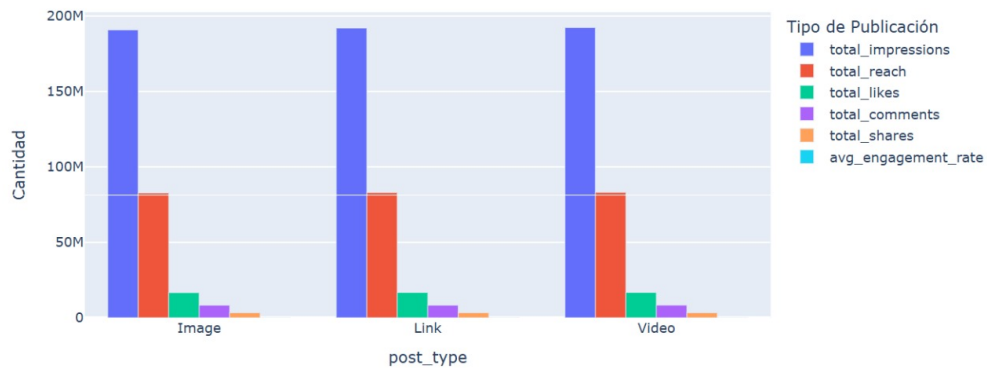
Al igual que con la campaña publicitaria, la presencia de un influencer asociado a la publicación no parece ser determinante para una variación significativa en los indicadores. Se puede evidenciar que la cantidad de publicaciones sin influencer es mucho mayor, con aproximadamente 90,006 publicaciones, en comparación con alrededor de 9,994 posts que sí están asociados con un influencer. Esto indica que, aunque la colaboración con influencers es una práctica común, no garantiza una mejora notable en la participación o enganche de las publicaciones.

Cantidad de Publicaciones de cada Tipo



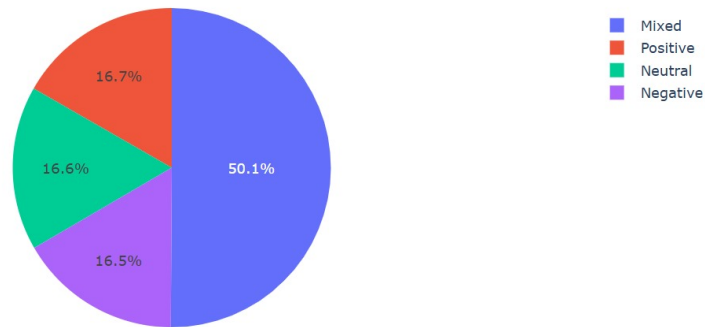
Un aprovechamiento del contenido multimedia como lo son los videos, la combinación entre audio e imagen principalmente, permitirá la más amplia implementación de estilos que sean los más adecuados para cada caso.

Indicadores por Tipo de Publicación

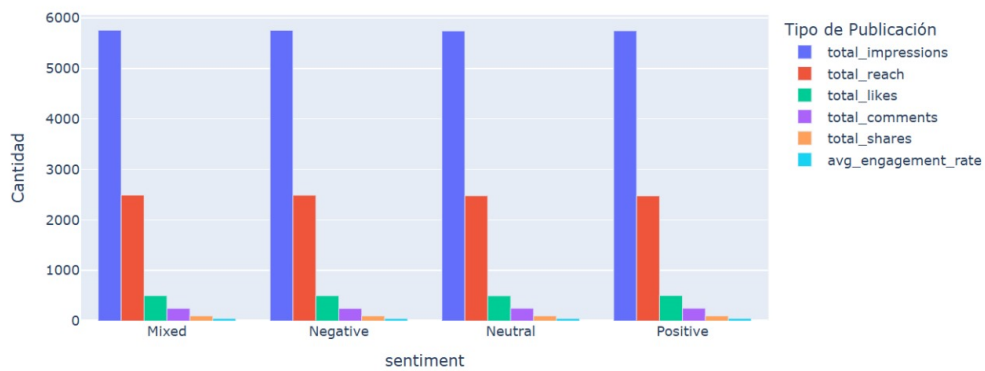


Los tipos de publicaciones son variados y tienen una cantidad muy similar, lo que indica que no hay un tipo de publicación que resalte significativamente sobre otro. Sin embargo, es importante mencionar que los posts en formato de video tienen una tasa de participación levemente mayor, con aproximadamente un 47.1%, lo que sugiere que este tipo de contenido podría ser ligeramente más efectivo para generar interacción con la audiencia. es un medio versátil y útil para todo tipo de publicaciones.

Cantidad de Publicaciones por Sentimiento



Indicadores por Sentimiento de Publicación



Los posts con un sentimiento expresado mixto son más comunes, representando el 50.1% del total, es decir, 50,100 posts. En comparación, los posts con sentimientos positivos constituyen el 16.7%, los neutrales el 16.6%, y los negativos el 16.5%. Esta distribución indica que los sentimientos mixtos son más prevalentes en las expresiones de las redes sociales en comparación con los sentimientos puramente positivos, neutrales o negativos. Esta observación sugiere que se debe de ser consciente de la complejidad emocional del público objetivo en las redes sociales para poder adaptar las estrategias de contenido. Esto podría implicar mayor amplitud en matices y una comprensión más profunda de las emociones detrás de las interacciones y reacciones.

Conclusiones

El planteamiento del proyecto fue diseñado de manera óptima, centrándose en evaluar el impacto de las publicaciones en redes sociales de alto tráfico para el desarrollo de campañas de marketing. Se realizó un análisis exhaustivo de métricas clave, como likes, comentarios, compartidos, impresiones y alcance por plataforma, así como de factores como la audiencia, el tiempo de publicación, el impacto de campañas de marketing y tendencias de contenido. Esta aproximación permitió obtener una visión integral y detallada de la interacción de los usuarios con el contenido en las redes sociales.

La utilización de una base de datos con 100,000 registros y 24 categorías, junto con la aplicación de rigurosas reglas de negocio para garantizar la integridad y validez de los datos, facilitó un enfoque estratégico y optimizado para mejorar las estrategias de marketing en redes sociales. Además, la creación adecuada de tablas en la base de datos optimizó el manejo de los datos al realizar consultas, lo que contribuyó significativamente a la eficiencia y precisión en la visualización en Dash. Esto a su vez facilitó la lectura de los datos y su respectivo análisis, permitiendo tomar decisiones informadas y estratégicas.

Para el desarrollo de este proyecto se pudieron también evidenciar oportunidades de mejora para un proceso más riguroso. Se fueron haciendo correcciones sobre cada aspecto, aprovechando al máximo una base de datos tan completa y rica en conocimientos que son en realidad útiles.