

# *Segmentación de Usuarios en la Aplicación Móvil Trash to Treasure*

**PROYECTO FINAL**

Bárbara Ortiz



# CONTENIDOS

1

Metodología y Descripción de  
los Datos  
pág - 3

2

Principales Hallazgos del  
Análisis Exploratorio  
pág - 5

3

Resultados de la Segmentación  
de Usuarios  
pág - 7

4

Pruebas de Hipótesis  
pág - 8

5

Análisis de Métricas de Negocio  
pág - 11

6

Conclusiones y Recomendaciones  
pág - 15

7

Fuentes Utilizadas  
pág - 16

2

Dashboard en Tableau  
pág - 17



# SEGMENTACIÓN DE USUARIOS EN LA APLICACIÓN MÓVIL TRASH TO TREASURE

## Objetivo del Proyecto:

Segmentar a los usuarios y usuarias de la aplicación móvil Trash to Treasure en función de su comportamiento y analizar diferentes métricas del producto, como la tasa de retención, tiempo dedicado a la aplicación, frecuencia de eventos y conversión en el evento `contacts_show`.



# Metodología y Descripción de los Datos





# PRINCIPALES HALLAZGOS DEL ANÁLISIS EXPLORATORIO

## Distribución de Eventos por Tipo

El evento más común es **tips\_show**, seguido por **photos\_show** y **advert\_open**.

## Distribución Temporal de los Eventos:

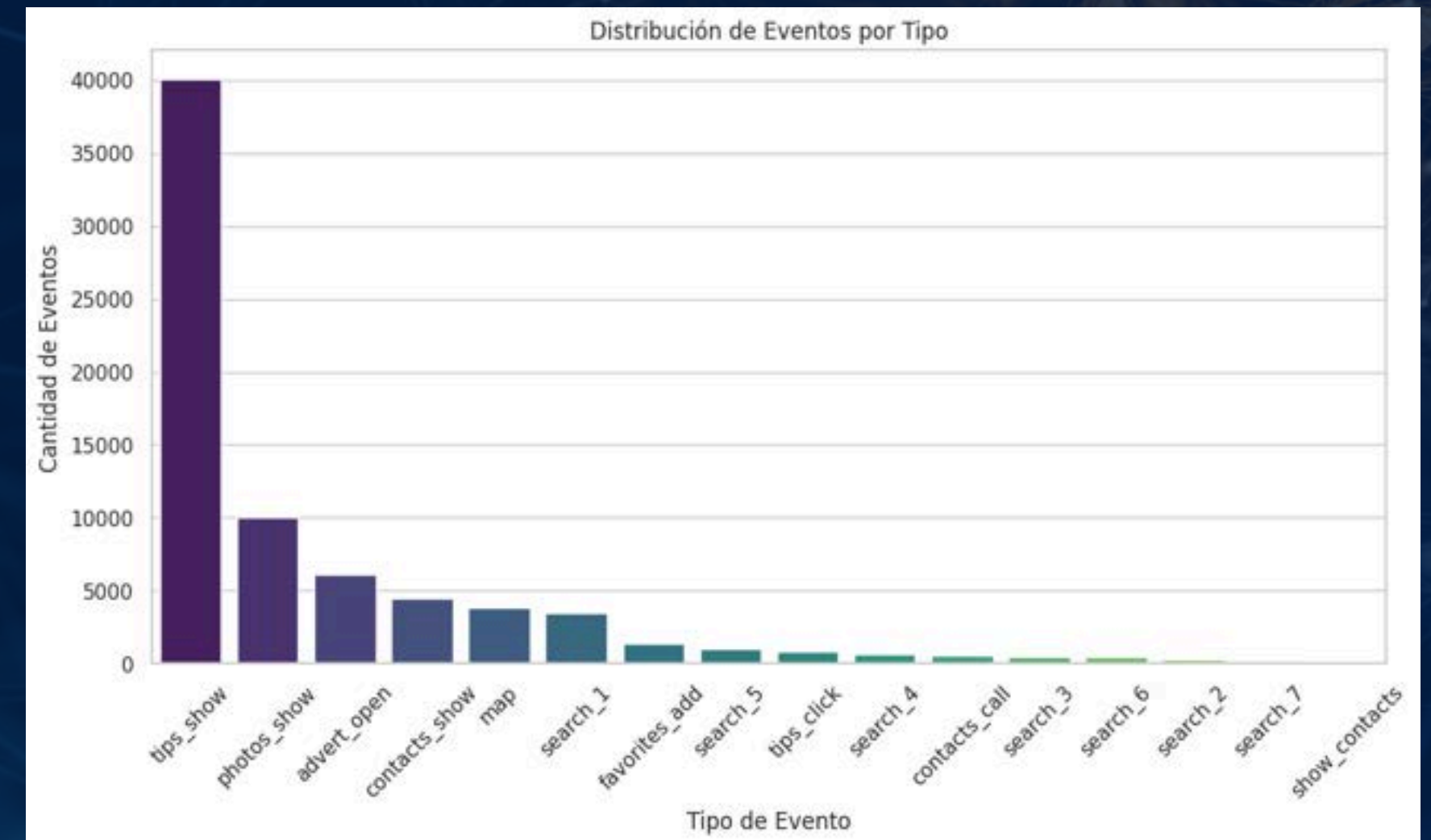
Los eventos están distribuidos de manera constante a lo largo del tiempo.

## Distribución de Usuarios por Fuente de Descarga:

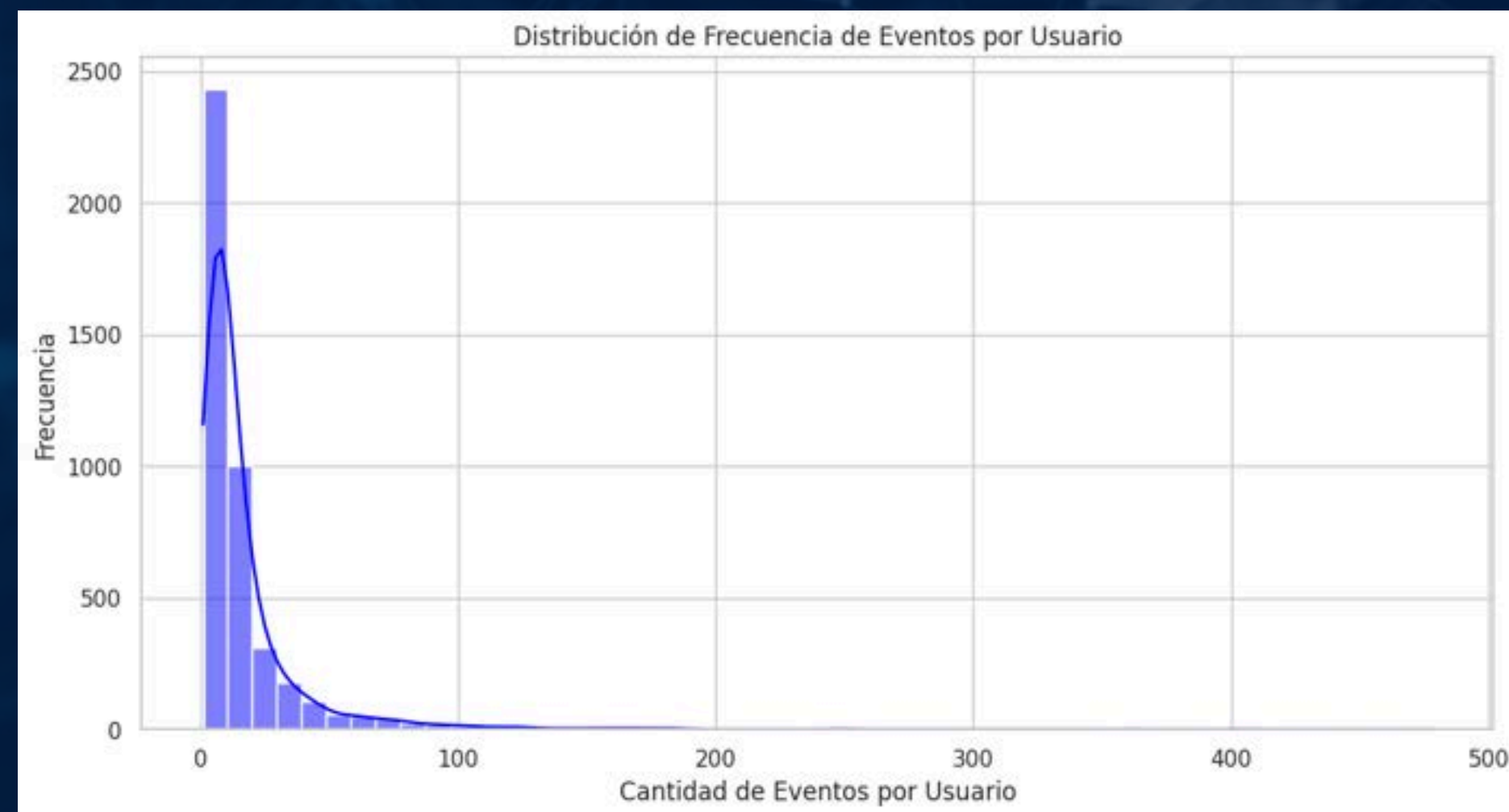
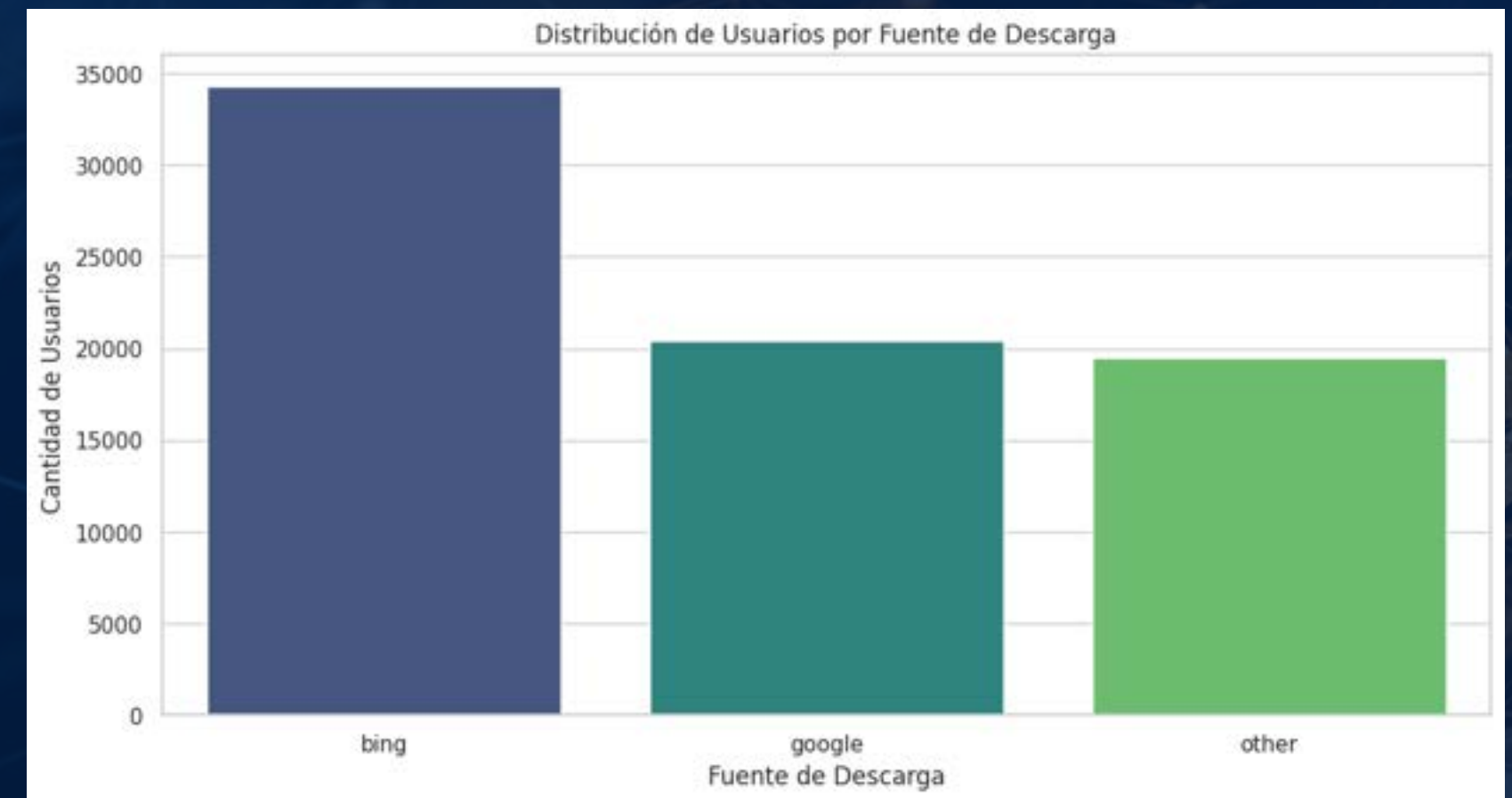
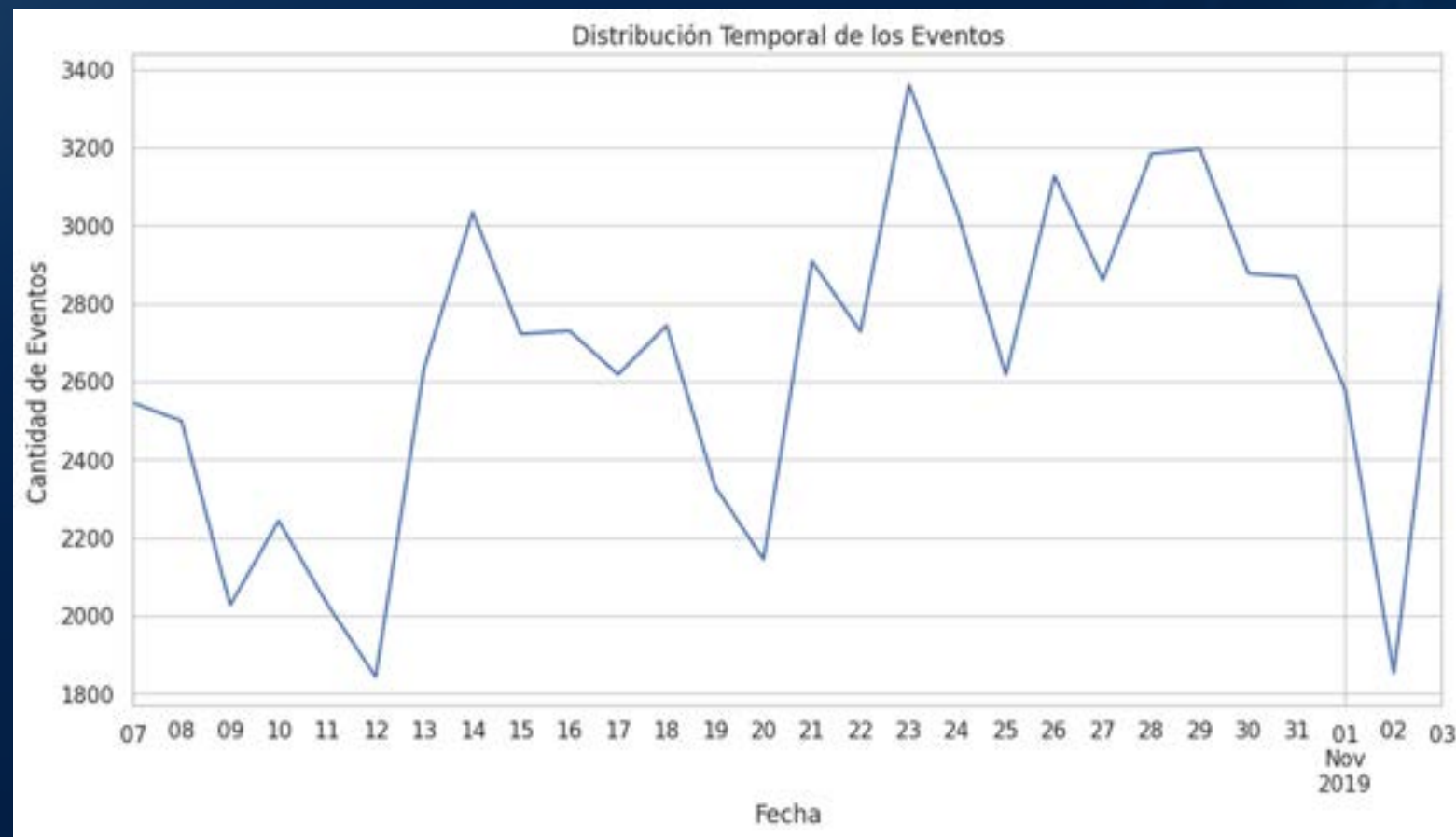
La mayoría de los usuarios descargaron la aplicación desde Bing, seguido por Other y Google.

## Frecuencia de Eventos por Usuario:

La mayoría de los usuarios realizan una cantidad moderada de eventos.









# RESULTADOS DE LA SEGMENTACIÓN DE USUARIOS

## Número Óptimo de Clusters:

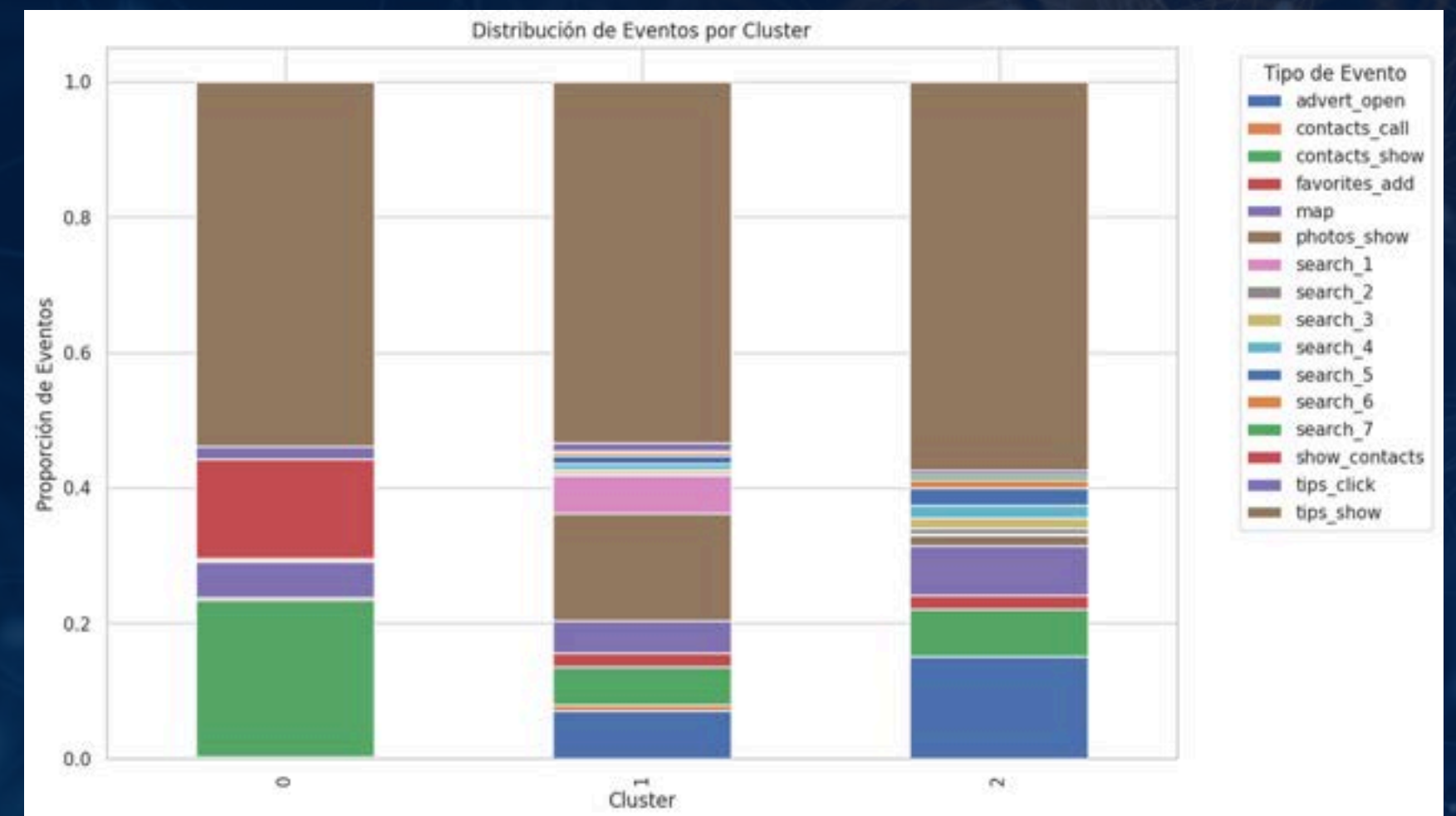
Determinado usando el método del codo y el método de la silueta (3 clusters).

## Clusters Identificados:

Cluster 1: Mayor cantidad de usuarios, mayor proporción de **tips\_show** y **photos\_show**.

Cluster 2: Alta interacción, mayor proporción de **contacts\_show**.

Cluster 3: Comportamiento menos común, menor cantidad de usuarios.





# PRUEBAS DE HIPÓTESIS

En este proyecto, formulamos y probamos dos hipótesis estadísticas para evaluar diferencias clave en el comportamiento de los usuarios. A continuación se describen las hipótesis y los métodos utilizados para probarlas.

- Hipótesis 1: Diferencia en la Conversión en Vistas de Información de Contacto entre Usuarios de Bing y Google
- Hipótesis 2: Diferencia en el Tiempo Promedio Dedicado a la Aplicación entre Usuarios de Diferentes Fuentes





# HIPÓTESIS 1: DIFERENCIA EN LA CONVERSIÓN EN VISTAS DE INFORMACIÓN DE CONTACTO ENTRE USUARIOS DE BING Y GOOGLE

- **Pregunta:** ¿Existe una diferencia significativa en la tasa de conversión en vistas de información de contacto entre usuarios que descargaron la aplicación desde Bing y aquellos que la descargaron desde Google?
- **Hipótesis Nula (H0):** No hay diferencia en la tasa de conversión entre usuarios de Bing y Google.
- **Hipótesis Alternativa (H1):** Hay una diferencia en la tasa de conversión entre usuarios de Bing y Google.
- **Método:** Utilizamos una prueba de proporciones Z para comparar las tasas de conversión. Esta prueba es adecuada para comparar proporciones entre dos grupos independientes.
- **Resultados:** El P-valor obtenido fue 0.8244, lo cual es mayor que el nivel de significancia de 0.05. Esto indica que no hay evidencia suficiente para rechazar la hipótesis nula. Concluimos que no hay una diferencia significativa en la tasa de conversión en vistas de información de contacto entre usuarios de Bing y Google.



## HIPÓTESIS 2: DIFERENCIA EN EL TIEMPO PROMEDIO DEDICADO A LA APLICACIÓN ENTRE USUARIOS DE DIFERENTES FUENTES

- **Pregunta:** ¿Existe una diferencia significativa en el tiempo promedio dedicado a la aplicación entre usuarios que descargaron la aplicación desde diferentes fuentes (Bing, Google y otras)?
- **Hipótesis Nula ( $H_0$ ):** No hay diferencia en el tiempo promedio dedicado a la aplicación entre usuarios de diferentes fuentes.
- **Hipótesis Alternativa ( $H_1$ ):** Hay una diferencia en el tiempo promedio dedicado a la aplicación entre usuarios de diferentes fuentes.
- **Método:** Realizamos pruebas t de Student para comparar los tiempos promedio dedicados a la aplicación entre los grupos de usuarios de Bing y Google, y entre los usuarios de Bing y otras fuentes. La prueba t de Student es adecuada para comparar las medias de dos grupos independientes.
- **Resultados:**
  - Comparación entre Bing y Google: El P-valor fue 0.5200, indicando que no hay evidencia suficiente para rechazar la hipótesis nula. No hay una diferencia significativa en el tiempo promedio dedicado a la aplicación entre usuarios de Bing y Google.
  - Comparación entre Bing y Other: El P-valor fue 0.2651, indicando que no hay evidencia suficiente para rechazar la hipótesis nula. No hay una diferencia significativa en el tiempo promedio dedicado a la aplicación entre usuarios de Bing y otras fuentes.

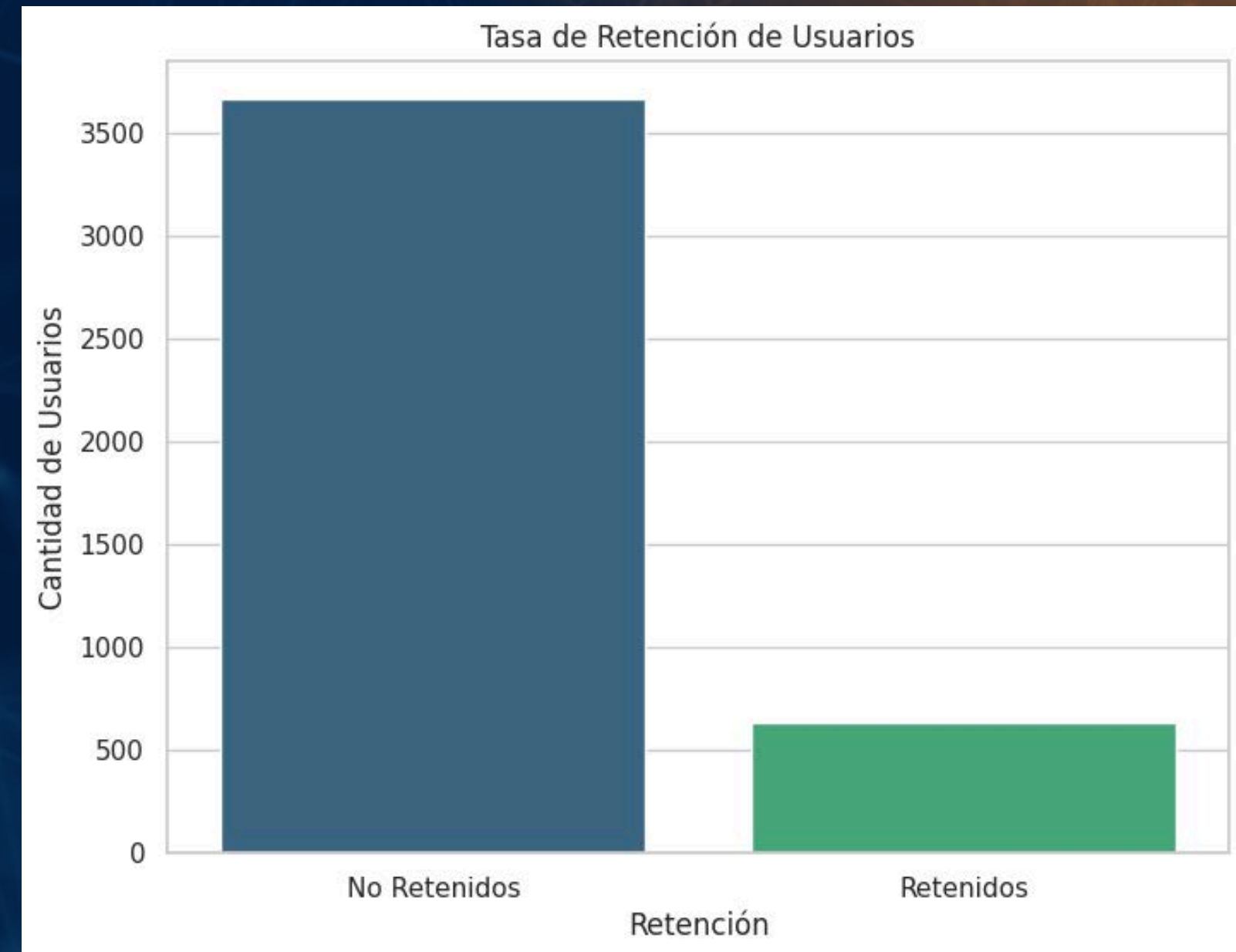


# ANÁLISIS DE MÉTRICAS DE NEGOCIO

1

## RETENCIÓN

La tasa de retención muestra que aproximadamente el 14.70% de los usuarios regresan a la aplicación después de una semana.



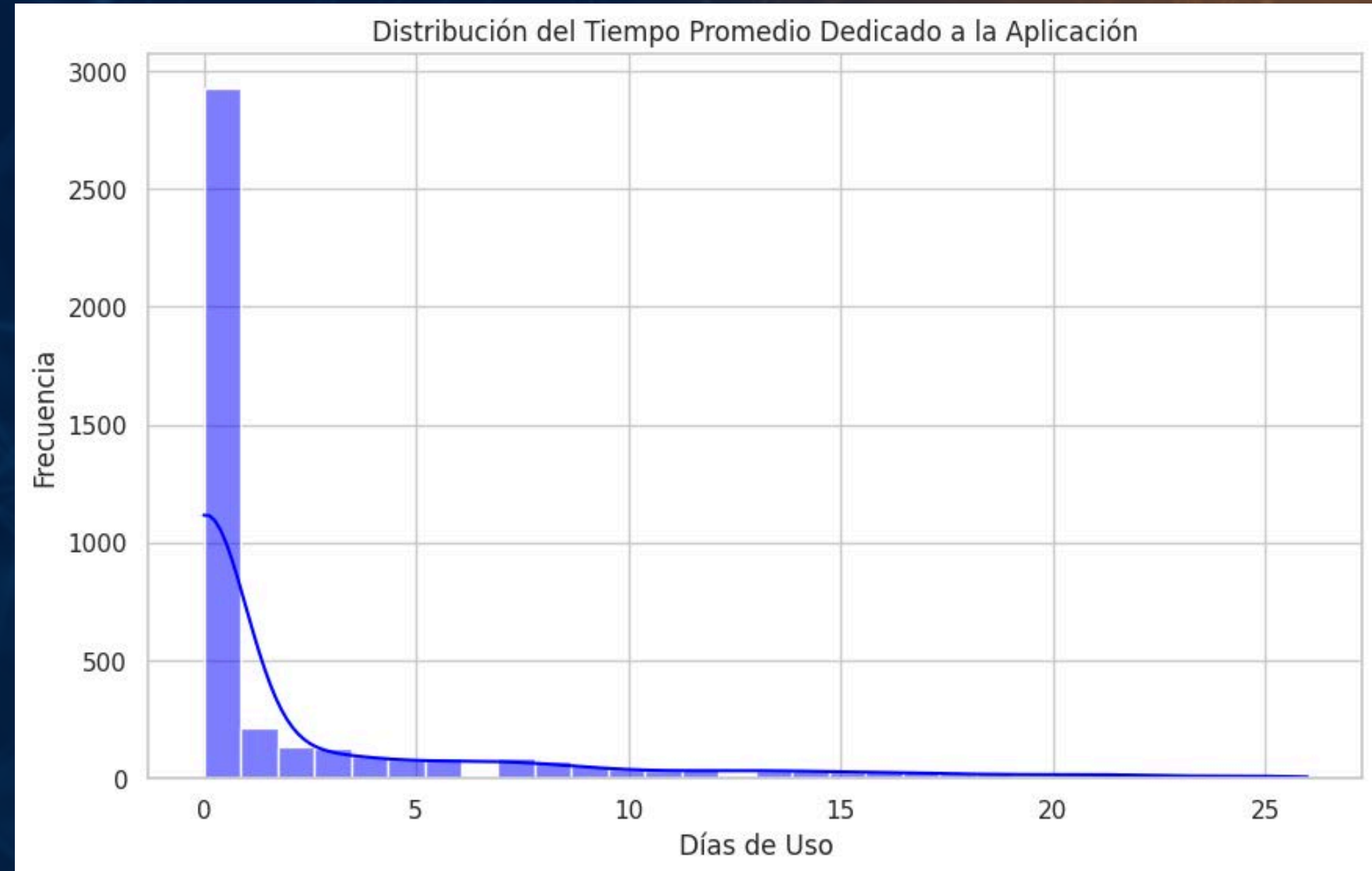


# ANÁLISIS DE MÉTRICAS DE NEGOCIO

2

## TIEMPO

El tiempo promedio dedicado por los usuarios es de 2.44 días, mostrando una distribución donde la mayoría de los usuarios utilizan la aplicación por pocos días.





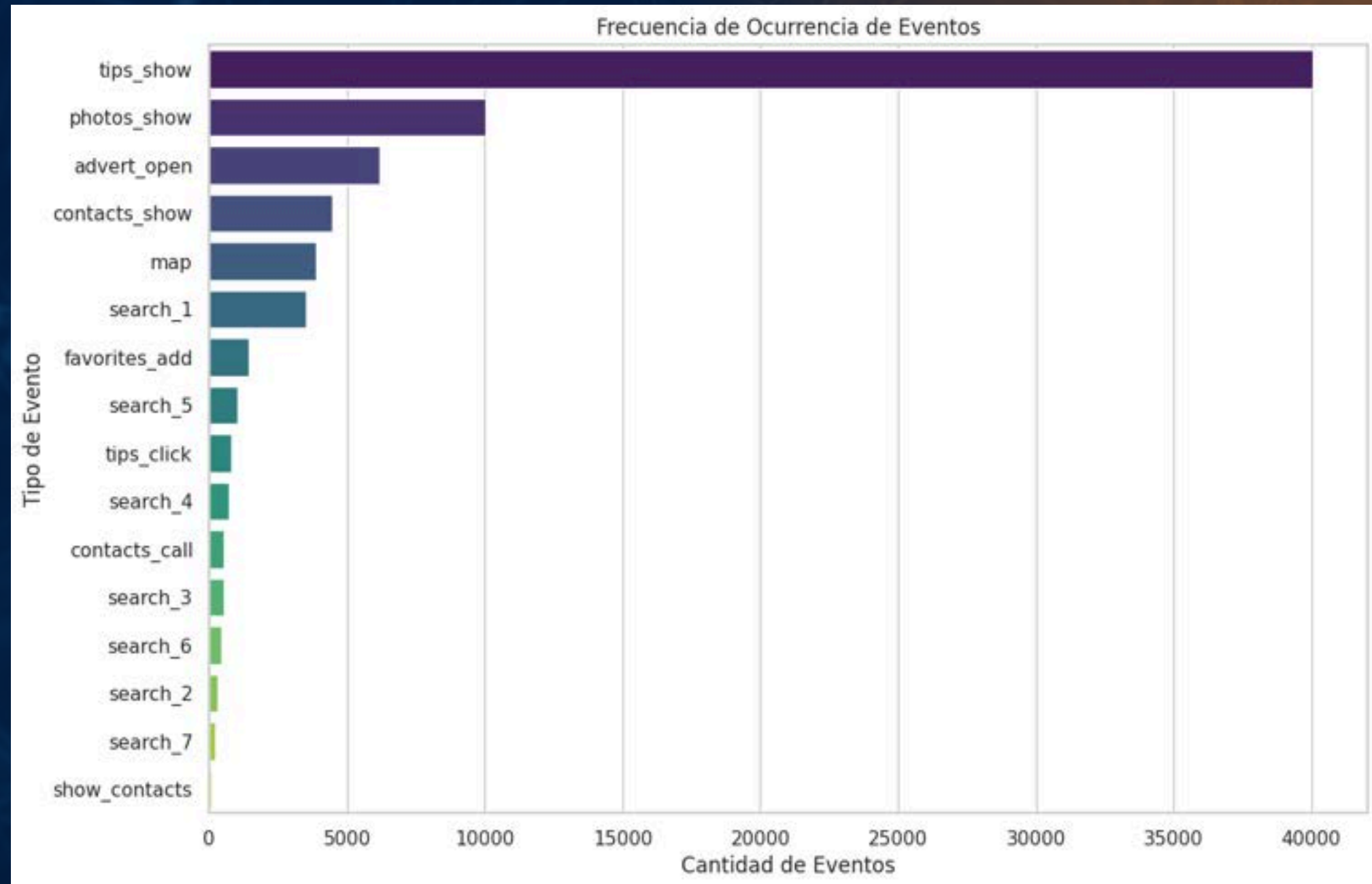
# ANÁLISIS DE MÉTRICAS DE NEGOCIO

3

## FRECUENCIA

Los eventos más comunes en la aplicación son:

- tips\_show
- photos\_show
- advert\_open.



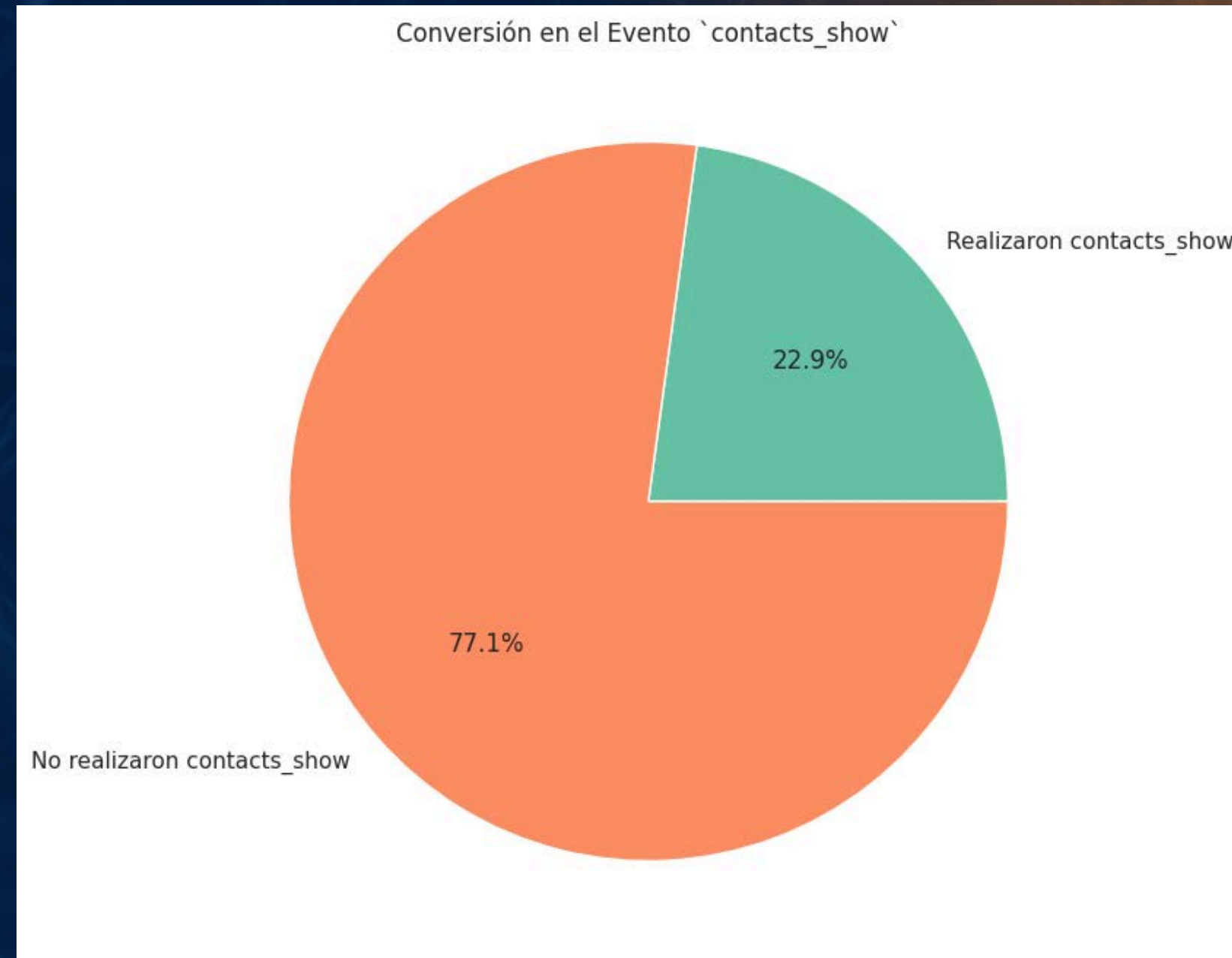


# ANÁLISIS DE MÉTRICAS DE NEGOCIO

4

## CONVERSIÓN

- Aproximadamente el 22.85% de los usuarios realizaron el evento `contacts_show` al menos una vez.





# CONCLUSIONES Y RECOMENDACIONES



La segmentación de usuarios revela tres clusters distintos basados en la frecuencia de eventos.



No hay diferencias significativas en la tasa de conversión en vistas de información de contacto entre usuarios que descargaron la aplicación de Bing y Google.



No hay diferencias significativas en el tiempo promedio dedicado a la aplicación entre usuarios de diferentes fuentes de descarga.



Implementar estrategias de retención personalizadas para cada cluster de usuarios.



Continuar monitoreando las métricas clave para identificar posibles cambios en el comportamiento de los usuarios.



Realizar estudios adicionales para entender mejor las necesidades y preferencias de los usuarios en cada cluster.



# FUENTES UTILIZADAS

- pandas. (n.d.). Documentación de pandas. <https://pandas.pydata.org/pandas-docs/stable/>
  - Ayuda en la manipulación y análisis de datos tabulares, así como en la limpieza y procesamiento de datos.
- scikit-learn. (n.d.). Documentación de scikit-learn. <https://scikit-learn.org/stable/>
  - Proporciona herramientas para la segmentación de usuarios utilizando algoritmos de clustering como K-means y para la normalización de datos.
- matplotlib. (n.d.). Documentación de matplotlib. <https://matplotlib.org/stable/contents.html>
  - Utilizada para crear visualizaciones de datos, como gráficos de barras y gráficos de línea.
- seaborn. (n.d.). Documentación de seaborn. <https://seaborn.pydata.org/>
  - Proporciona una interfaz de alto nivel para dibujar atractivas visualizaciones estadísticas.
- scipy.stats. (n.d.). Documentación de scipy.stats. <https://docs.scipy.org/doc/scipy/reference/stats.html>
  - Utilizada para realizar pruebas estadísticas, como la prueba t de Student.
- statsmodels. (n.d.). Documentación de statsmodels. <https://www.statsmodels.org/stable/index.html>
  - Proporciona herramientas para realizar pruebas de proporciones Z y otras pruebas estadísticas avanzadas.
- Kaggle. (n.d.). <https://www.kaggle.com/>
  - Fuente de inspiración para técnicas de análisis de datos y segmentación de usuarios a partir de competencias y notebooks públicos.
- Medium. (n.d.). Artículos sobre análisis de datos y segmentación. <https://medium.com/>
  - Proporcionan ejemplos prácticos y casos de estudio sobre segmentación de usuarios y análisis de métricas de negocio.



# DASHBOARDS TABLEAU



Dashboard:  
"Análisis de Eventos y Usuarios"



CLICK



Dashboard:  
"Análisis de Eventos Diarios y  
Usuarios por Fuente"





*Por su atención*  
**MUCHAS GRACIAS**