

Análisis de embudo y retención para MercadoLibre

Introducción

Eres analista de producto en MercadoLibre, dentro del equipo de Crecimiento y Retención.

El director de producto te plantea un reto:

“Necesitamos entender en qué etapa del proceso perdemos usuarios y cómo podemos mejorar su retención a lo largo del tiempo.”

Tu misión será usar SQL para mapear el embudo de conversión completo, identificar los principales puntos de fuga y evaluar la retención de usuarios por cohortes.

Finalmente, deberás proponer mejoras accionables basadas en los datos.

Objetivos del proyecto

Al finalizar este proyecto podrás:

1. Construir embudos multietapa en SQL usando CTEs.
2. Calcular tasas de conversión entre pasos y detectar caídas.
3. Analizar la retención de usuarios por cohortes.
4. Simular mejoras en conversión o retención.
5. Validar resultados y comunicar hallazgos ejecutivos.

Dataset del proyecto

Tablas disponibles

- **mercadolibre_funnel**

Registra eventos de usuarios durante el proceso de compra.

Columna	Tipo de Dato (Ejemplo)	Descripción
user_id	STRING (ej. "c786f5e6-a5b8-45fc-94ee-efeba1a209ce")	Identificador único del usuario. Permite rastrear el comportamiento de un mismo usuario a lo largo del tiempo.
session_id	STRING (ej. "sid_1129")	Identificador único de la sesión. Cada sesión agrupa los eventos

		generados durante una visita del usuario.
event_name	STRING (ej. "add_to_cart", "purchase")	Nombre del evento realizado por el usuario (por ejemplo: agregar al carrito, iniciar checkout, compra, etc.).
event_times	FLOAT o INTEGER (ej. 1.76001e+15)	Marca temporal del evento en formato Unix timestamp (milisegundos desde 1970-01-01). Permite ordenar cronológicamente los eventos.
country	STRING (ej. "Argentina")	País desde donde se genera la sesión o el evento.
device_category	STRING (ej. "mobile", "desktop", "tablet")	Tipo de dispositivo utilizado por el usuario.
platform	STRING (ej. "android", "iOS", "web")	Plataforma o sistema operativo desde donde se origina el evento.
product_cat	STRING (ej. "Electrónica", "Moda", "Hogar")	Categoría del producto asociado al evento (solo aplica para eventos de producto). Puede estar vacío para eventos no relacionados con productos.
price	FLOAT (ej. 1250.00)	Precio del producto en el momento del evento (si aplica). Vacío para eventos que no involucran un ítem.
currency	STRING (ej. "USD", "ARS")	Moneda en la que se registra el precio.

referral_source	STRING (ej. "organic", "paid_search", "social")	Fuente de tráfico que originó la sesión (orgánico, búsqueda pagada, redes sociales, etc.).
event_date	DATETIME (ej. "2025-08-09 11:03:54")	Fecha y hora legible del evento, derivada del event_times.
year	INTEGER (ej. 2025)	Año en que ocurrió el evento. Útil para filtrados y particionamiento.

- **mercadolibre_retention**

Mide actividad recurrente por usuario y periodo.

Columna	Tipo de Dato (Ejemplo)	Descripción
user_id	STRING (ej. "00198c1f-bc1e-403e-b46a-67b0b3ac7657")	Identificador único del usuario. Permite hacer seguimiento individual del comportamiento de retención.
signup_date	DATE (ej. "2025-05-01")	Fecha del registro del usuario (sin hora). Se usa como punto de referencia para calcular los días

		posteriores al registro.
signup_datetime	DATETIME (ej. "2025-05-01 18:02:25")	Fecha y hora exactas en que el usuario completó su registro o creó la cuenta.
country	STRING (ej. "Brazil")	País del usuario al momento del registro.
device_category	STRING (ej. "tablet", "mobile", "desktop")	Tipo de dispositivo utilizado en el registro o durante la actividad.
platform	STRING (ej. "web", "android", "iOS")	Plataforma o sistema operativo desde el cual el usuario accedió.
day_after_signup	INTEGER (ej. 8)	Número de días transcurridos desde la fecha de registro (signup_date) hasta la fecha de actividad (activity_date). Sirve para calcular cohortes de retención

		(D1, D7, D30, etc.).
activity_date	DATE (ej. "2025-05-09")	Fecha en que el usuario tuvo actividad (por ejemplo, sesión iniciada, compra, vista de producto).
active	INTEGER (ej. 0 o 1)	Indicador binario de actividad: 1 si el usuario estuvo activo en esa fecha, 0 si no.
prob_active	FLOAT (ej. 0.1697)	Probabilidad estimada de que el usuario esté activo en esa fecha (por ejemplo, calculada con un modelo predictivo de retención).

Definición del Macro Journey (Embudo General)

El negocio está interesado particularmente en el siguiente embudo de conversión.

Etapa del Embudo	Evento	Descripción	Métrica Clave
● Descubrimiento	first_visit	El usuario entra por primera vez al sitio o aplicación de Mercado Libre.	Número de nuevos usuarios / sesiones nuevas
● Interés / Consideración	select_item / select_promotion	El usuario selecciona un	Tasa de clics (CTR) en

		producto o una promoción específica, mostrando intención de compra.	productos o promociones
● Intención de Compra	add_to_cart	El usuario añade un producto al carrito.	Porcentaje de usuarios que agregan al carrito después de ver un producto
● Inicio de Compra	begin_checkout	El usuario inicia el proceso de pago.	Porcentaje de carritos que inician el checkout
● Información de Envío	add_shipping_info	El usuario ingresa su dirección o método de envío.	Porcentaje de checkouts que completan la información de envío
● Información de Pago	add_payment_info	El usuario agrega o selecciona un método de pago.	Porcentaje de checkouts que agregan información de pago
● Conversión / Compra	purchase	El usuario completa la compra.	Tasa de conversión final (compras / sesiones)

Contexto del negocio

La dirección de producto busca responder:

¿En qué etapa se pierden más usuarios?

- ◆ Entre el [01/01/2025] y el [08/31/2025], ¿cuál es la tasa de conversión entre cada etapa clave del embudo?.
- ◆ *¿En qué paso se observa la mayor caída porcentual de usuarios?*
- ◆ *¿Cómo varía esta pérdida por país (country)?*

Métricas clave:

- Tasa de conversión por etapa

- Agrupación de eventos por device_category y referral_source

¿Qué tan bien retenemos a los usuarios a lo largo del tiempo?

- ◆ Para los usuarios que se registraron entre el [01/01/2025] y el [06/01/2025], ¿cuál es la tasa de retención en D7, D14, D21, D28?
- ◆ *¿Cómo se comporta la retención por país (country)?*

Métrica clave: Retención por cohorte (D7, D14, D21, D28)

Proceso general

1. Explorar el esquema y los datos base
2. Construir el embudo de conversión
3. Calcular tasas y caídas
4. Analizar retención y cohortes
5. Simular mejoras y redactar el informe ejecutivo