meta : calcular, con CTEs, cuántos usuarios alcanzan cada etapa del Embudo General, desde la primera visita hasta la compra

```
        WHERE event_name = 'begin_checkout'
            AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
    ),

    add_shipping_info AS (
        SELECT DISTINCT user_id, country
        FROM mercadolibre_funnel
        WHERE event_name = 'add_shipping_info'
            AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
    ),

    add_payment_info AS (
        SELECT DISTINCT user_id, country
        FROM mercadolibre_funnel
        WHERE event_name = 'add_payment_info'
            AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
    ),

    purchase AS (
        SELECT DISTINCT user_id, country
        FROM mercadolibre_funnel
        WHERE event_name = 'purchase'
            AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
    ),
```

```
 1  WITH first_visits AS (
 2      SELECT DISTINCT user_id, country
 3      FROM mercadolibre_funnel
 4      WHERE event_name = 'first_visit'
 5          AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
 6  ),
 7
 8  select_item AS (
 9      SELECT DISTINCT user_id, country
10      FROM mercadolibre_funnel
11      WHERE event_name IN ('select_item', 'select_promotion')
12          AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
13  ),
14
15  add_to_cart AS (
16      SELECT DISTINCT user_id, country
17      FROM mercadolibre_funnel
18      WHERE event_name = 'add_to_cart'
19          AND event_date BETWEEN '2025-01-01' AND '2025-08-31'
20  ),
21
22  begin_checkout AS (
23      SELECT DISTINCT user_id, country
24      FROM mercadolibre_funnel
```

```
    funnel_counts AS (
        SELECT
            fv.country,
            COUNT(fv.user_id) AS usuarios_first_visit,
            COUNT(si.user_id) AS usuarios_select_item,
            COUNT(a.user_id)  AS usuarios_add_to_cart,
            COUNT(bc.user_id) AS usuarios_begin_checkout,
            COUNT(asi.user_id) AS usuarios_add_shipping_info,
            COUNT(api.user_id) AS usuarios_add_payment_info,
            COUNT(p.user_id)   AS usuarios_purchase
        FROM first_visits fv
        LEFT JOIN select_item si
            ON fv.user_id = si.user_id AND fv.country = si.country
        LEFT JOIN add_to_cart a
            ON fv.user_id = a.user_id AND fv.country = a.country
        LEFT JOIN begin_checkout bc
            ON fv.user_id = bc.user_id AND fv.country = bc.country
        LEFT JOIN add_shipping_info asi
            ON fv.user_id = asi.user_id AND fv.country = asi.country
        LEFT JOIN add_payment_info api
            ON fv.user_id = api.user_id AND fv.country = api.country
        LEFT JOIN purchase p
            ON fv.user_id = p.user_id AND fv.country = p.country
        GROUP BY fv.country
)

SELECT
    country,
    usuarios_select_item * 100.0 / NULLIF(usuarios_first_visit, 0) AS conversion_select_item,
    usuarios_add_to_cart * 100.0 / NULLIF(usuarios_first_visit, 0) AS conversion_add_to_cart,
    usuarios_begin_checkout * 100.0 / NULLIF(usuarios_first_visit, 0) AS conversion_begin_checkout,
    usuarios_add_shipping_info * 100.0 / NULLIF(usuarios_first_visit, 0) AS conversion_add_shipping_info,
    usuarios_add_payment_info * 100.0 / NULLIF(usuarios_first_visit, 0) AS conversion_add_payment_info,
    usuarios_purchase * 100.0 / NULLIF(usuarios_first_visit, 0) AS conversion_purchase
FROM funnel_counts
ORDER BY conversion_purchase DESC;
```

RESULTADO

Resultado

country	conversion_select_item	conversion_add_to_cart	conversion_begin_checkout	conversion_add_shipping_info	conversion_add_payment_info	conversion_purchase
Uruguay	81.8182	22.7273	4.54545	4.54545	4.54545	4.54545
Bolivia	80.6452	9.67742	3.22581	3.22581	3.22581	3.22581
Mexico	79.7521	13.2231	4.13223	3.30579	2.89256	2.47934
Peru	84.5455	10	2.72727	2.72727	1.81818	1.81818
Argentina	75	8.75	4.375	1.875	1.875	1.25
Chile	78.3505	17.5258	8.24742	3.09278	2.06186	1.03093
Brazil	72.6027	8.90411	2.39726	1.36986	1.36986	0.684932
Ecuador	74.5763	10.1695	5.08475	1.69492	1.69492	0
Colombia	76.3636	9.69697	4.84848	3.0303	2.42424	0
Paraguay	71.4286	9.52381	0	0	0	0

Guía general de los pasos

Aquí medirás la retención real de usuarios con la tabla mercadolibre_retention.

Cada cohorte representa el grupo de usuarios que se registró en un mes específico.

La tabla 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).

1.

Contar usuarios activos acumulados por país (D7, D14, D21, D28)

Objetivo: Para cada país, contar usuarios activos acumulados desde su registro, en el rango 2025-01-01 → 2025-08-31, al día 7, día 14, día 21 y día 28.

```sql
1   SELECT
2       country,
3
4       COUNT(DISTINCT CASE
5           WHEN active = 1 AND day_after_signup >= 7
6           THEN user_id
7       END) AS users_d7,
8
9       COUNT(DISTINCT CASE
10          WHEN active = 1 AND day_after_signup >= 14
11          THEN user_id
12      END) AS users_d14,
13
14      COUNT(DISTINCT CASE
15          WHEN active = 1 AND day_after_signup >= 21
16          THEN user_id
17      END) AS users_d21,
18
19      COUNT(DISTINCT CASE
20          WHEN active = 1 AND day_after_signup >= 28
21          THEN user_id
22      END) AS users_d28
```

```sql
FROM mercadolibre_retention
WHERE activity_date BETWEEN '2025-01-01' AND '2025-08-31'
GROUP BY country
ORDER BY country;
```

RESULTADO

country	users_d7	users_d14	users_d21	users_d28
Argentina	1122	690	297	24
Bolivia	387	224	92	12
Brazil	2641	1649	740	76
Chile	782	484	206	16
Colombia	1384	851	357	27
Ecuador	473	299	123	15
Mexico	2075	1345	615	75
Paraguay	380	231	104	10
Peru	739	448	201	28
Uruguay	210	119	56	6

.

Convertir conteos a % de retención por país

Objetivo: Convertir los conteos del Task 1 en porcentajes de retención por país al día 7, día 14, día 21 y día 28.

```sql
SELECT
    country,

    ROUND(
        COUNT(DISTINCT CASE
            WHEN active = 1 AND day_after_signup >= 7
            THEN user_id
        END) * 100.0
        / NULLIF(COUNT(DISTINCT user_id), 0)
    , 1) AS retention_d7_pct,

    ROUND(
        COUNT(DISTINCT CASE
            WHEN active = 1 AND day_after_signup >= 14
            THEN user_id
        END) * 100.0
        / NULLIF(COUNT(DISTINCT user_id), 0)
    , 1) AS retention_d14_pct,

    ROUND(
        COUNT(DISTINCT CASE
            WHEN active = 1 AND day_after_signup >= 21
            THEN user_id
        END) * 100.0
```

```
25          / NULLIF(COUNT(DISTINCT user_id), 0)
26       , 1) AS retention_d21_pct,
27
28      ROUND(
29          COUNT(DISTINCT CASE
30              WHEN active = 1 AND day_after_signup >= 28
32              THEN user_id
32          END) * 100.0
33          / NULLIF(COUNT(DISTINCT user_id), 0)
34       , 1) AS retention_d28_pct
35
36   FROM mercadolibre_retention
37   WHERE activity_date BETWEEN '2025-01-01' AND '2025-08-31'
38   GROUP BY country
39   ORDER BY country;
```

RESULTADOS

country	retention_d7_pct	retention_d14_pct	retention_d21_pct	retention_d28_pct
Argentina	85.1	52.3	22.5	1.8
Bolivia	80.8	46.8	19.2	2.5
Brazil	87.2	54.4	24.4	2.5
Chile	83.7	51.8	22.1	1.7
Colombia	84.5	52	21.8	1.6
Ecuador	79.1	50	20.6	2.5
Mexico	86.1	55.8	25.5	3.1
Paraguay	80.9	49.1	22.1	2.1
Peru	84.3	51.1	22.9	3.2
Uruguay	86.1	48.8	23	2.5

Definir la cohorte de registro

Objetivo: Ahora vamos a analizar la retención por cohort. El primer paso es crear una consulta SQL que asigne el cohort en formato YYYY-MM a cada usuario (usando su primera fecha de registro).

```
 1   SELECT
 2       user_id,
 3       MIN(signup_date) AS signup_date,
 4       TO_CHAR(
 5           DATE_TRUNC('month', MIN(signup_date)),
 6           'YYYY-MM'
 7       ) AS cohort
 8   FROM mercadolibre_retention
 9   GROUP BY user_id
l0   LIMIT 5;
l1
```

RESULTADO

Resultado

user_id	signup_date	cohort
0002b1ba-9c7f-4989-87cb-54109b84c2cb	2025-05-02	2025-05
0011c921-8b74-4984-9f90-b50daae0442b	2025-02-19	2025-02
00147274-7efe-42fd-aaaf-a1b7aefb834f	2025-02-01	2025-02
0017b94f-3a9f-4850-a011-4b3b3141c201	2025-08-05	2025-08
00198c1f-bc1e-403e-b46a-67b0b3ac7657	2025-05-01	2025-05

Calcular retención por cohorte y periodo D7, D14, D21, D28

Objetivo: Ahora, para cada cohorte mensual (**YYYY-MM**), vas a calcular el % de usuarios activos al día 7, 14, 21, y 28 desde su registro.

```sql
WITH cohort AS (
    SELECT
        user_id,
        MIN(signup_date) AS signup_date,
        TO_CHAR(
            DATE_TRUNC('month', MIN(signup_date)),
            'YYYY-MM'
        ) AS cohort
    FROM mercadolibre_retention
    GROUP BY user_id
),

activity AS (
    SELECT
        r.user_id,
        c.cohort,
        r.day_after_signup,
        r.active
    FROM mercadolibre_retention r
    LEFT JOIN cohort c
        ON r.user_id = c.user_id
    WHERE r.activity_date BETWEEN '2025-01-01' AND '2025-08-31'
)
```

```sql
SELECT
    cohort,

    ROUND(
        100.0 * COUNT(DISTINCT CASE
            WHEN day_after_signup >= 7 AND active = 1
            THEN user_id
        END) / NULLIF(COUNT(DISTINCT user_id), 0)
    , 1) AS retention_d7_pct,

    ROUND(
        100.0 * COUNT(DISTINCT CASE
            WHEN day_after_signup >= 14 AND active = 1
            THEN user_id
        END) / NULLIF(COUNT(DISTINCT user_id), 0)
    , 1) AS retention_d14_pct,

    ROUND(
        100.0 * COUNT(DISTINCT CASE
            WHEN day_after_signup >= 21 AND active = 1
            THEN user_id
        END) / NULLIF(COUNT(DISTINCT user_id), 0)
    , 1) AS retention_d21_pct,

    ROUND(
        100.0 * COUNT(DISTINCT CASE
            WHEN day_after_signup >= 28 AND active = 1
            THEN user_id
        END) / NULLIF(COUNT(DISTINCT user_id), 0)
    , 1) AS retention_d28_pct

FROM activity
GROUP BY cohort
ORDER BY cohort;
```

RESULTADO

Resultado

cohort	retention_d7_pct	retention_d14_pct	retention_d21_pct	retention_d28_pct
2025-01	86.2	56.2	24.1	3
2025-02	86.8	56	24.6	2.7
2025-03	87.7	56.8	26.6	3
2025-04	87.2	53.9	23	2
2025-05	86	54.5	26.2	3
2025-06	85.9	55.1	25.2	2.1
2025-07	86.4	56.4	25.9	2.7
2025-08	70.8	29.7	7.5	0.2