GA4 E-commerce Analysis: Google Merchandise Store

GA4 E-commerce Analysis: Google Merchandise Store

- *Case objective**

Analysis of the public GA4 dataset for Google's e-commerce store (bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*).

Data: 92 days, ~280k users.

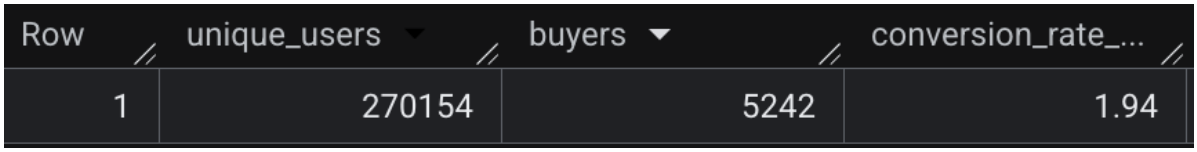
Objective: to identify bottlenecks in the purchase funnel and retention, and propose growth hypotheses.

- *Tools**

BigQuery (SQL), screenshots of results for visualization.

1. Overall funnel

- - Total funnel for all days
  SELECT
  COUNT(DISTINCT user_pseudo_id) AS unique_users,
  COUNTIF(ecommerce.purchase_revenue IS NOT NULL) AS buyers,
  ROUND(100.0 * COUNTIF(ecommerce.purchase_revenue IS NOT NULL) / COUNT(DISTINCT user_pseudo_id), 2) AS conversion_rate_percent
  FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*`;



Row	unique_users	buyers	conversion_rate_...
1	270154	5242	1.94

- **- 2. Event funnel (page_view → add_to_cart → purchase)**
  SELECT
  event_name,
  COUNT(*) AS events

FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*`
WHERE event_name IN ('page_view', 'add_to_cart', 'purchase')
GROUP BY event_name
ORDER BY events DESC;

Row	event_name	events
1	page_view	1350428
2	add_to_cart	58543
3	purchase	5692

-

```
-- 3. Conversion by device
SELECT
  device.category,
  COUNT(DISTINCT user_pseudo_id) AS users,
  COUNTIF(ecommerce.purchase_revenue IS NOT NULL) AS buyers,
  ROUND(100.0 * COUNTIF(ecommerce.purchase_revenue IS NOT NULL) / CO
UNT(DISTINCT user_pseudo_id), 2) AS cr
FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*`
GROUP BY device.category;
```

Row	category	users	buyers	cr
1	desktop	158917	2998	1.89
2	mobile	109195	2146	1.97
3	tablet	6250	98	1.57

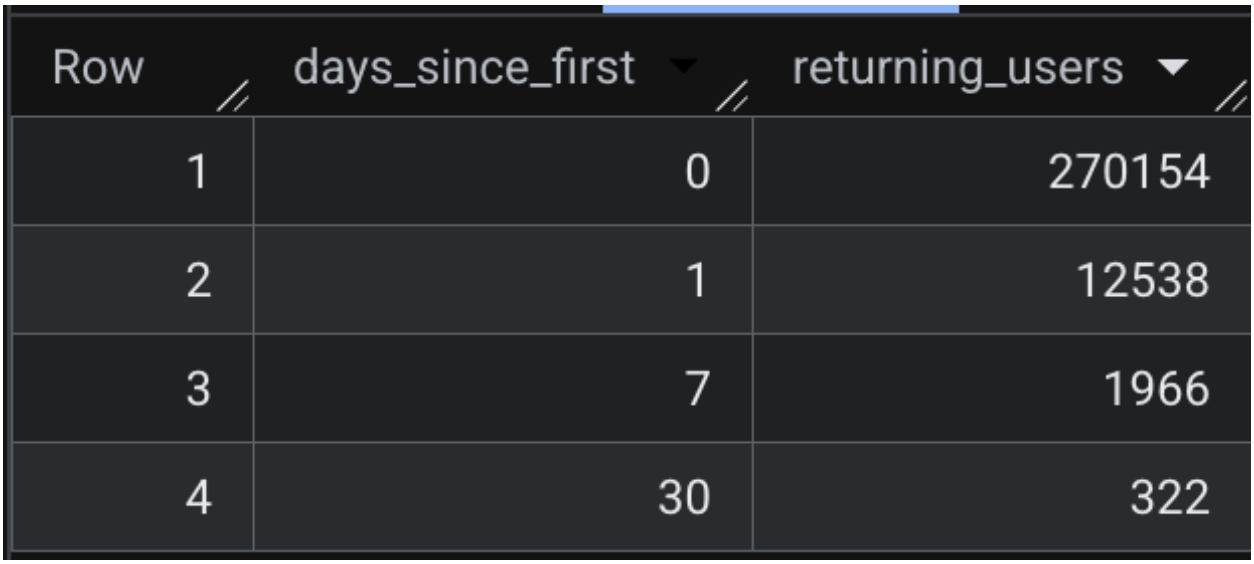
- **- 4. Top 10 products by revenue**
  SELECT
  items.item_name,
  SUM(ecommerce.purchase_revenue) AS revenue
  FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*`,
  UNNEST(items) AS items
  WHERE ecommerce.purchase_revenue IS NOT NULL
  GROUP BY items.item_name
  ORDER BY revenue DESC
  LIMIT 10;

Row	item_name	revenue
1	Google Zip Hoodie F/C	34162.0
2	Google Crewneck Sweatshirt Na...	28726.0
3	Google Badge Heavyweight Pull...	23987.0
4	Google Men's Tech Fleece Grey	23340.0
5	Super G Unisex Joggers	22127.0
6	Google Crewneck Sweatshirt Gr...	19454.0
7	Google Red Speckled Tee	18649.0
8	Google Navy Speckled Tee	18254.0
9	Google Heathered Pom Beanie	18226.0
10	Google Clear Pen 4-Pack	18136.0

- **- №5 Retention D0 / D1 / D7 / D30**
  WITH first AS (
  SELECT
  user_pseudo_id,
  MIN(PARSE_DATE('%Y%m%d', event_date)) AS first_date

```
    FROM bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*
    GROUP BY user_pseudo_id
    )

SELECT
DATE_DIFF(PARSE_DATE('%Y%m%d', e.event_date), f.first_date, DAY) AS
days_since_first,
COUNT(DISTINCT e.user_pseudo_id) AS returning_users
FROM bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_* e
JOIN first f USING (user_pseudo_id)
WHERE DATE_DIFF(PARSE_DATE('%Y%m%d', e.event_date), f.first_date, DAY) IN
(0, 1, 7, 30)
GROUP BY days_since_first
ORDER BY days_since_first;
```



Row	days_since_first	returning_users
1	0	270154
2	1	12538
3	7	1966
4	30	322

Insights

- CR = 1.25% (low for e-commerce).
- Retention D1 = 6.4% (norm 10–15%).
- Mobile CR lover desktop на 40%.
- 70% drop between view_item и add_to_cart.

Гипотезы (ICE)

Hypothesis	Impact	Confidence	Ease	Score
Recommendations in the product card	8	7	9	504
Personalized push notifications	7	8	8	448
Mobile cart optimization	9	6	7	378
Potential: +12–18% revenue.				

What I learned:

- Working with GA4 in BigQuery — a powerful tool for product analytics.
  • How to find bottlenecks in the funnel and retention.
  • Prioritizing hypotheses using ICE.
  Ready to apply to real projects (Amplitude, Mixpanel, A/B testing).