Дослідження деталей рекламних кампаній Google і Facebook і підготовка даних про події користувачів

Гамова Катерина

ДАНІ ТА ІНСТРУМЕНТИ

ads_analysis_goit_course
DBeaver аналітика про щоденну ефективність рекламних кампаній на платформах Facebook і Google

публічний датасет GA4 в Google BigQuery робота з подієвими даними інтернет-магазину

Завдяння з даними
ads_analysis_goit_course y DBeaver

- Знайти середнє, максимум та мінімум spend окремо для платформ Google та Facebook в одній таблиці
- Знайти 5 днів з найбільшим загальним ROMI, враховуючи дані з платформ Google та Facebook разом
- Знайти кампанію з найвищим рівнем загального тижневого value
- Знайти кампанію з найбільшим приростом охоплення (reach) місяць-до-місяця, обчислений як абсолютна різниця між охопленням поточного та попереднього місяця.
- Знайти найдовший безперервний (щоденний) показ adset_name (разом з Google та Facebook)

- Середнє, максимум та мінімум spend окремо для платформ Google та Facebook в одній таблиці

```
with Facebook_Google_ads as(select fabd.ad_date,'Facebook' as media_source, fc.campaign_name,
fa.adset_name, fabd.spend, fabd.impressions, fabd.reach, fabd.clicks, fabd.leads, fabd.value
from facebook_ads_basic_daily fabd
left join facebook_campaign fc on fc.campaign_id=fabd.campaign_id
left join facebook_adset fa on fa.adset_id=fabd.adset_id
union all
select ad_date,'Google' as media_source, campaign_name, adset_name, spend, impressions,
reach, clicks, leads, value
from google_ads_basic_daily)
select ad_date, media_source,
avg(spend) as avg_spend,
min(spend) as min_spend,
max(spend) as max_spend
from Facebook_Google_ads
where ad_date is not null
group by 1,2
order by ad_date;
```

ad_date	media_source	avg_spend	min_spend	max_spend
2020-11-11	Facebook	189	189	189
2020-11-12	Facebook	1,150	34	2,266
2020-11-12	Google	199	66	332
2020-11-13	Google	1,054	1,054	1,054
2020-11-13	Facebook	318	14	622
2020-11-14	Facebook	727	727	727
2020-11-14	Google	1	1	1
2020-11-15	Google	682	682	682
2020-11-15	Facebook	698	698	698

- 5 днів з найбільшим загальним ROMI, враховуючи дані з платформ Google та Facebook разом

```sql
with Facebook_Google_ads as(select fabd.ad_date,'Facebook' as media_source,fc.campaign_name,
fa.adset_name,fabd.spend,fabd.impressions,fabd.reach,fabd.clicks,fabd.leads,fabd.value
from facebook_ads_basic_daily fabd
left join facebook_campaign fc on fc.campaign_id=fabd.campaign_id
left join facebook_adset fa on fa.adset_id=fabd.adset_id
union all
select ad_date,'Google' as media_source,campaign_name,adset_name,spend,impressions,
reach,clicks,leads,value
from google_ads_basic_daily)
select ad_date,media_source,
case
    when sum(spend)=0 then '0' else round(cast((sum(value))as numeric)/cast(sum(spend) as numeric),1)
end as romi
from Facebook_Google_ads
where ad_date is not null
group by 1,2
order by romi desc
limit 5;
```

ad_date	media_source	romi
2022-01-11	Facebook	2.5
2022-01-07	Facebook	2.5
2022-01-16	Facebook	2.4
2022-02-12	Facebook	2.4
2022-02-13	Facebook	2.4

- **Кампанія з найвищим рівнем загального тижневого value**

```sql
with Facebook_Google_ads as(select fabd.ad_date,'Facebook' as media_source,fc.campaign_name,
fa.adset_name,fabd.spend,fabd.impressions,fabd.reach,fabd.clicks,fabd.leads,fabd.value
from facebook_ads_basic_daily fabd
left join facebook_campaign fc on fc.campaign_id=fabd.campaign_id
left join facebook_adset fa on fa.adset_id=fabd.adset_id
union all
select ad_date,'Google' as media_source,campaign_name,adset_name,spend,
impressions,reach,clicks,leads,value
from google_ads_basic_daily)
select date(date_trunc('week', ad_date)) as week_start,media_source,campaign_name,
sum(value) as total_value
from Facebook_Google_ads
where ad_date is not null
group by 1,2,3
order by total_value desc
limit 1;
```

	week_start	media_source	campaign_name	total_value
1	2022-04-11	Facebook	Expansion	1,254,027

- **Кампанія з найбільшим приростом охоплення (reach) місяць-до-місяця, обчислений як абсолютна різниця між охопленням поточного та попереднього місяця.**

```sql
with Facebook_Google_ads as(select fabd.ad_date,'Facebook' as media_source,
fc.campaign_name,fa.adset_name,fabd.spend,fabd.impressions,fabd.reach,
fabd.clicks,fabd.leads,fabd.value
from facebook_ads_basic_daily fabd
left join facebook_campaign fc on fc.campaign_id=fabd.campaign_id
left join facebook_adset fa on fa.adset_id=fabd.adset_id
union all
select ad_date,'Google' as media_source,campaign_name,adset_name,spend,
impressions,reach,clicks,leads,value
from google_ads_basic_daily),
monthly_reach as
(select to_char(ad_date,'YYYY-MM-01') as ad_month,
campaign_name,sum(reach) as total_reach
from Facebook_Google_ads
group by 1,2),
prev_month_reach as (select ad_month,campaign_name,total_reach,
lag(total_reach,1) over(partition by campaign_name order by ad_month) as prev_month_reach
from monthly_reach)
select ad_month,campaign_name,total_reach,prev_month_reach,
round((total_reach::numeric-prev_month_reach::numeric),2) as growth_reach
from prev_month_reach
where ad_month is not null
order by growth_reach desc nulls last
limit 1;
```

A-Z ad_month	A-Z campaign_name	123 total_reach	123 prev_month_reach	123 growth_reach
2022-04-01	Hobbies	5,011,659	745,084	4,266,575

- **Найдовший безперервний (щоденний) показ adset_name (разом з Google та Facebook)**

```sql
with Facebook_Google_ads as(select fabd.ad_date,'Facebook' as media_source,
fc.campaign_name,fa.adset_name,fabd.spend,fabd.impressions,fabd.reach,
fabd.clicks,fabd.leads,fabd.value
from facebook_ads_basic_daily fabd
left join facebook_campaign fc on fc.campaign_id=fabd.campaign_id
left join facebook_adset fa on fa.adset_id=fabd.adset_id
union all
select ad_date,'Google' as media_source,campaign_name,adset_name,
spend,impressions,reach,clicks,leads,value
from google_ads_basic_daily),
numbered_date as(select adset_name,ad_date::date,
row_number() over (partition by adset_name order by ad_date) as date_number
from Facebook_Google_ads
where ad_date is not null
group by 1,2),
grouped_number as(select adset_name,ad_date::date,
ad_date::date - date_number * interval '1 day' as grouped_number
from numbered_date),
intervals as(select adset_name,
min (ad_date::date) as start_date,
max (ad_date::date) as end_date,
count(*) as duration_ad
from grouped_number
group by adset_name, grouped_number
order by adset_name, start_date)
select *
from intervals
order by duration_ad desc
limit 1;
```

	A-Z adset_name	start_date	end_date	123 duration_ad
1	Narrow	2021-05-17	2021-09-01	108

Завдяння з даними

GA4 в Google BigQuery

- Дістати інформацію про події, користувачів та сесії в GA4
- Дістати інформацію про конверсії від початку сесії до покупки
- Порахувати конверсію між різними посадковими сторінками
- Порахувати кореляціюміж залученістю користувачів та здійсненням покупок

- **Інформація про події, користувачів та сесії в GA4**

```
SELECT event_date,timestamp_micros(event_timestamp) as event_timestemp, user_pseudo_id, event_name, (select ee.value.int_value
from e.event_params ee where ee.key = 'ga_session_id') as session_id,
geo.country as country, device.category, traffic_source.name as campaign, traffic_source.source as source, traffic_source.medium
as medium
FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*` e
where _table_suffix >= '20210101' and _table_suffix <= '20211231' and
event_name in ('session_start', 'view_item', 'add_to_cart', 'begin_checkout', 'add_shipping_info', 'add_payment_info', 'purchase')
```

Row	event_date	event_timestemp	user_pseudo_id	event_name	session_id	country	category	campa...	source	medium
1	20210103	2021-01-03 07:19:3...	1160488.23759...	session_start	3672590...	Portugal	desktop	(data ...	(data dele...	(data dele...
2	20210103	2021-01-03 08:55:1...	1272248.35274...	session_start	1844478...	Canada	desktop	(data ...	(data dele...	(data dele...
3	20210103	2021-01-03 01:45:5...	1412894.31573...	session_start	6264365...	Brazil	mobile	(data ...	(data dele...	(data dele...
4	20210103	2021-01-03 12:51:1...	1561053.39117...	session_start	22759485	United Stat...	mobile	(data ...	(data dele...	(data dele...
5	20210103	2021-01-03 12:24:4...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
6	20210103	2021-01-03 12:30:0...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
7	20210103	2021-01-03 12:15:5...	2051660.36997...	session_start	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
8	20210103	2021-01-03 12:19:2...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
9	20210103	2021-01-03 12:26:3...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
10	20210103	2021-01-03 12:25:2...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
11	20210103	2021-01-03 12:32:3...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
12	20210103	2021-01-03 12:26:3...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...
13	20210103	2021-01-03 12:27:0...	2051660.36997...	view_item	2660893...	Canada	desktop	(data ...	(data dele...	(data dele...

- Інформація про конверсії від початку сесії до покупки

```
with user_session_id_info as (SELECT timestamp_micros(event_timestamp) as event_timestemp,
user_pseudo_id || (select ee.value.int_value from e.event_params ee where ee.key = 'ga_session_id')
as user_session_id, event_name,
traffic_source.name as campaign, traffic_source.source as source, traffic_source.medium as medium
FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*` e
where event_name in ('session_start', 'add_to_cart', 'begin_checkout', 'purchase')),
users_count as (select date(event_timestemp) as event_date,source,medium,campaign,
count (distinct case when event_name ='session_start' then user_session_id
end) as count_session_start,
count (distinct case when event_name ='add_to_cart' then user_session_id
end) as count_add_to_cart,
count (distinct case when event_name ='begin_checkout' then user_session_id
end) as count_begin_checkout,
count (distinct case when event_name ='purchase' then user_session_id
end) as count_purchase
from user_session_id_info
group by 1,2,3,4)
select event_date,source,medium,campaign, count_session_start,
round(count_add_to_cart/count_session_start,2) as visit_to_cart,
round(count_begin_checkout/count_session_start,2) as visit_to_checkout,
round(count_purchase/count_session_start,2) as visit_to_purchase
from users_count;
```

Row	event_date	source	medium	campaign	count_session_start	visit_to_cart	visit_to_checkout	visit_to_purchase
1	2020-12-31	<Other>	<Other>	<Other>	388	0.03	0.02	0.01
2	2020-12-31	(data dele...	(data deleted)	(data deleted)	112	0.09	0.06	0.04
3	2020-12-31	google	organic	(organic)	862	0.03	0.01	0.01
4	2020-12-31	(direct)	(none)	(direct)	612	0.03	0.01	0.0
5	2020-12-31	google	cpc	<Other>	118	0.05	0.02	0.01
6	2020-12-31	shop.goo...	referral	(referral)	175	0.07	0.03	0.02
7	2020-12-31	<Other>	referral	(referral)	228	0.05	0.03	0.02
8	2020-12-31	<Other>	organic	(organic)	90	0.04	0.02	0.0

- **Конверсія між різними посадковими сторінками**

```
with user_sessions as (select user_pseudo_id,
cast((select value.int_value from unnest(event_params) where key = 'ga_session_id') as string) as session_id,
regexp_extract((select value.string_value from unnest(event_params) where key = 'page_location'),
r'(?:\w+\:\/\/)?[^\/]+\/([^\?#]*)') as page_path,
(select value.string_value from unnest(event_params) where key = 'page_location') as page_location
from `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*` e
where _table_suffix between '20200101' and '20201231'
and event_name = 'session_start'),
purchases as (
select user_pseudo_id,
cast((select value.int_value from e.event_params where key = 'ga_session_id') as string)
as session_id
from `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*` e
where _table_suffix between '20200101' and '20201231'
and event_name = 'purchase'),
count_unic_purchase as (select page_path,
count(distinct us.user_pseudo_id ||us.session_id) as unic_user_sessions,
count(distinct p.user_pseudo_id||p.session_id) as purchases
from user_sessions us
left join purchases p
on us.user_pseudo_id = p.user_pseudo_id
and us.session_id = p.session_id
group by 1)
select page_path,unic_user_sessions,purchases,
round(purchases/unic_user_sessions,2) as conversion
from count_unic_purchase
group by 1,2,3
order by conversion desc;
```

Row	page_path	unic_user_sessions	purchases	conversion
1	Google+Redesign/Apparel/Google+Kirkland+Campus+Unisex+Tee	1	1	1.0
2	Google+Redesign/Accessories/Google+Large+Pet+Leash+Blue+Green	1	1	1.0
3	Google+Redesign/Apparel/Google+Chicago+Campus+Unisex+Tee	2	2	1.0
4	Google+Redesign/Apparel/Google+Cambridge+Campus+Ladies+Tee	1	1	1.0
5	Google+Redesign/Accessories/Google+Felt+Strap+Keyring	2	1	0.5
6	Google+Redesign/Apparel/Google+Toddler+Hero+Tee+Black	2	1	0.5
7	Google+Redesign/Campus+Collection/Google+Sunnyvale+Campus+La	2	1	0.5

- **Кореляція між залученістю користувачів та здійсненням покупок**

```sql
with user_session_id_info as (SELECT
user_pseudo_id || (select ee.value.int_value from e.event_params ee where ee.key = 'ga_session_id') as user_session_id,
sum(coalesce((select ee.value.int_value from e.event_params ee where ee.key = 'engagement_time_msec'),0)) as user_eng_time,
max(coalesce((select ee.value.int_value from e.event_params ee where ee.key = 'session_engaged'),
safe_cast ((select ee.value.string_value from e.event_params ee where ee.key = 'session_engaged')as integer),0)) as
session_engaged,
max(case
when event_name='purchase' then 1 else 0 end) as purchase_check
FROM `bigquery-public-data.ga4_obfuscated_sample_ecommerce.events_*` e
group by 1)
select
corr (user_eng_time, purchase_check) as corr_eng_time_purchase,
corr(session_engaged,purchase_check) as corr_engaged_purchase
from user_session_id_info;
```

Row	corr_eng_time_purchase	corr_engaged_purchase
1	0.31798788665534389	0.041310862178407105

Дякую за увагу!