

Website Analysis SQL Script

Description: This script performs various analyses on website data to extract insights and generate reports.

Author: Rupsa Chaudhuri

ANALYZING TRAFFIC SOURCES

TRAFFIC SOURCE ANALYSIS

Traffic source analysis is about understanding where your customers are coming from, and which channels are driving the highest quality traffic.

1. FINDING TOP TRAFFIC SOURCES

```
SELECT
    utm_source,
    utm_campaign,
    http_referer,
    COUNT(DISTINCT website_session_id) AS sessions
FROM
    website_sessions
WHERE YEAR(created_at) != '2015'
GROUP BY 1,2,3
ORDER BY 4 DESC;
```

	utm_source	utm_campaign	http_referer	sessions
▶	gsearch	nonbrand	https://www.gsearch.com	247564
	bsearch	nonbrand	https://www.bsearch.com	48072
	NULL	NULL	NULL	32793
	NULL	NULL	https://www.gsearch.com	28913
	gsearch	brand	https://www.gsearch.com	27395
	NULL	NULL	https://www.bsearch.com	6706
	bsearch	brand	https://www.bsearch.com	6545
	socialbook	desktop_targeted	https://www.socialbook.com	5590
	socialbook	pilot	https://www.socialbook.com	5095

gsearch non-brand seems to be a major traffic driver (247,564 sessions), followed by traffic from bsearch/non-brand (48,072 sessions). Direct traffic comes at third (32793).

2. MONTHLY TRENDS FOR GSEARCH, ALONGSIDE MONTHLY TRENDS FOR EACH OF OUR OTHER CHANNELS .

```
SELECT
    YEAR(created_at) AS Year,
    COUNT(DISTINCT case when utm_source='gsearch' then website_session_id
else null end) AS gsearch_sessions,
    COUNT(DISTINCT case when utm_source='bsearch' then website_session_id
else null end) AS bsearch_sessions,
```

```

COUNT(DISTINCT case when utm_source is null and http_referer is not
null then website_session_id else null end) AS organic_search_sessions,
COUNT(DISTINCT case when utm_source is null and http_referer is null
then website_session_id else null end) AS direct_search_sessions
FROM
    website_sessions
WHERE YEAR(created_at) != '2015'
GROUP BY 1;

```

	Year	gsearch_sessions	bsearch_sessions	organic_search_sessions	direct_search_sessions
►	2012	47888	8829	2973	2780
	2013	79672	16059	8914	8136
	2014	147399	29729	23732	21877

Gsearch traffic seems dominant with consistent increase in sessions throughout the year. Traffic from other channels (bsearch, organic direct) also increased consistently but way behind gsearch.

3. YEARLY TREND FOR GSEARCH, BUT THIS TIME SPLITTING OUT NONBRAND AND BRAND CAMPAIGNS SEPARATELY.

```

SELECT
    YEAR(ws.created_at) AS Year,
    COUNT(DISTINCT case when utm_campaign='nonbrand' then o.order_id else
null end)
    /COUNT(DISTINCT case when utm_campaign='nonbrand' then
ws.website_session_id else null end)*100 as nonbrand_conv_rate,
    COUNT(DISTINCT case when utm_campaign='brand' then o.order_id else
null end)
    /COUNT(DISTINCT case when utm_campaign='brand' then
ws.website_session_id else null end)*100 as brand_conv_rate
FROM
    website_sessions ws
    LEFT JOIN
    orders o ON o.website_session_id = ws.website_session_id
WHERE
    YEAR(ws.created_at) != '2015'
    AND ws.utm_source = 'gsearch'
GROUP BY 1 ;

```

	Year	nonbrand_conv_rate	brand_conv_rate
►	2012	3.8286	5.4729
	2013	6.4129	7.2094
	2014	7.2642	7.7062

For gsearch, brand traffic converts at a higher rate (7.7% in 2014) than non-brand traffic (7.26% in 2014) over all the three years.

Traffic Conversion Rates

4. WE NEED TO CHECK HOW MUCH OUR TRAFFIC SOURCE IS GENERATING SALES AND WHAT IS THE SESSION TO ORDER CONVERSION RATE.

```
SELECT
    utm_source,
    utm_campaign,
    http_referer,
    COUNT(DISTINCT w.website_session_id) AS sessions,
    COUNT(DISTINCT o.order_id) AS orders,
    COUNT(DISTINCT o.order_id) / COUNT(DISTINCT w.website_session_id)*100
AS session_to_order_conv_rate
FROM
    website_sessions w
    LEFT JOIN
    orders o ON o.website_session_id = w.website_session_id
WHERE YEAR(w.created_at) != '2015'
GROUP BY 1,2,3
ORDER BY 4 DESC;
```

	utm_source	utm_campaign	http_referer	sessions	orders	session_to_order_conv_rate
▶	gsearch	nonbrand	https://www.gsearch.com	247564	15797	6.3810
	bsearch	nonbrand	https://www.bsearch.com	48072	3237	6.7336
	NULL	NULL	NULL	32793	2304	7.0259
	NULL	NULL	https://www.gsearch.com	28913	2042	7.0626
	gsearch	brand	https://www.gsearch.com	27395	2026	7.3955
	NULL	NULL	https://www.bsearch.com	6706	580	8.6490
	bsearch	brand	https://www.bsearch.com	6545	564	8.6173
	socialbook	desktop_targeted	https://www.socialbook.com	5590	288	5.1521
	socialbook	pilot	https://www.socialbook.com	5095	55	1.0795

The conversion rate (CVR) is good for organic search and direct search at 8.6% and 7.02% respectively. Traffic from socialbook has a much lower CVR at 5.15%. Here we can bid down on socialbook.

BID OPTIMIZATION

Analyzing for bid optimization is about understanding the value of various segments of paid traffic, so that you can optimize your marketing budget

5. TRAFFIC CONVERSION RATE BY DEVICE TYPE.

```
SELECT
    YEAR(ws.created_at) AS Year,
    COUNT(DISTINCT case when ws.device_type='desktop' then
ws.website_session_id else null end) AS desktop_sessions,
    COUNT(DISTINCT case when ws.device_type='desktop' then o.order_id else
null end) AS desktop_orders,
    COUNT(DISTINCT case when ws.device_type='mobile' then
ws.website_session_id else null end) AS mobile_sessions,
    COUNT(DISTINCT case when ws.device_type='mobile' then o.order_id else
null end) AS mobile_orders,
    COUNT(DISTINCT case when ws.device_type='desktop' then o.order_id else
null end)
/COUNT(DISTINCT case when ws.device_type='desktop' then
ws.website_session_id else null end)*100 as desktop_conv_rate,
    COUNT(DISTINCT case when ws.device_type='mobile' then o.order_id else
```

```

null end)
/COUNT(DISTINCT case when ws.device_type='mobile' then
ws.website_session_id else null end)*100 as mobile_conv_rate
FROM
    website_sessions ws
    LEFT JOIN
    orders o ON o.website_session_id = ws.website_session_id
WHERE YEAR(ws.created_at) != '2015'
GROUP BY 1 ;

```

	Year	desktop_sessions	desktop_orders	mobile_sessions	mobile_orders	desktop_conv_rate	mobile_conv_rate
▶	2012	46575	2353	15895	233	5.0521	1.4659
	2013	80186	6431	32595	1016	8.0201	3.1170
	2014	155524	14281	77898	2579	9.1825	3.3107

The conversion rate for desktop(9.18% in 2014) is way higher than mobile(3.3% in 2014).Based on this, it could be beneficial to increase bids for desktop users.

ANALYZING WEBSITE PERFORMANCE

ANALYZING TOP WEBSITE CONTENT

Website content analysis is about understanding which pages are seen the most by your users, to identify where to focus on improving your business

- Finding the most-viewed pages that customers view on your site
- Identifying the most common entry pages to your website - the first thing a user sees

6. IDENTIFYING TOP WEBSITE PAGES

```

SELECT
    pageview_url, COUNT(DISTINCT website_pageview_id) AS sessions
FROM
    website_pageviews
WHERE YEAR(created_at) != '2015'
GROUP BY 1
ORDER BY 2 DESC;

```

	pageview_url	sessions
▶	/products	222619
	/the-original-mr-fuzzy	141247
	/lander-2	131170
	/home	115357
	/cart	79550
	/lander-3	67644
	/shipping	54011
	/lander-1	47574
	/billing-2	40006
	/lander-5	37543
	/thank-you-for-your-...	26893
	/the-forever-love-bear	20585
	/the-birthday-sugar-...	14769
	/lander-4	9385
	/billing	3617
	/the-hudson-river-mi...	711

The /products page is the most visited page on the website with 222,619 sessions, followed by /the-original-mr-fuzzy with 141,247 sessions, followed by /lander-2 and /home with 131,170 & 115,357 respectively.

7. IDENTIFYING TOP ENTRY PAGES

```
WITH first_pageview AS(
SELECT
    website_session_id, MIN(website_pageview_id) AS min_pv_id
FROM
    website_pageviews
GROUP BY website_session_id)

SELECT
    wp.pageview_url AS landing_page,
    COUNT(distinct fp.website_session_id) AS sessions_hitting_this_lander
FROM
    first_pageview fp
    LEFT JOIN
    website_pageviews wp ON wp.website_pageview_id = fp.min_pv_id
WHERE YEAR(wp.created_at) != '2015'
GROUP BY wp.pageview_url
ORDER BY 2 DESC;
```

	landing_page	sessions_hitting_this_lander
▶	/lander-2	131170
	/home	115357
	/lander-3	67644
	/lander-1	47574
	/lander-5	37543
	/lander-4	9385

/lander-2 has the most sessions at 131,170 followed by /home at 115,357 sessions. This suggests that users are finding the /lander-2 page more engaging than other landing pages.

LANDING PAGE PERFORMANCE & TESTING

Landing page analysis and testing is about understanding the performance of your key landing pages and then testing to improve your results

8. CALCULATING BOUNCE RATE

```
SELECT b.pageview_url,
COUNT(DISTINCT b.website_session_id) as sessions,
COUNT(DISTINCT CASE WHEN b.pageview_count=1 THEN b.website_session_id ELSE
NULL END) as bounced_sessions,
COUNT(DISTINCT CASE WHEN b.pageview_count=1 THEN b.website_session_id ELSE
NULL END)/COUNT(DISTINCT b.website_session_id) * 100 as bounce_percentage
FROM
(
```

```

SELECT COUNT(DISTINCT website_pageview_id) as pageview_count,
pageview_url,website_session_id,created_at
FROM website_pageviews
GROUP BY website_session_id) b
WHERE YEAR(b.created_at) != '2015'
GROUP BY b.pageview_url
ORDER BY 4 DESC;

```

	pageview_url	sessions	bounced_sessions	bounce_percentage
▶	/lander-1	47574	25330	53.2434
	/lander-4	9385	4851	51.6889
	/lander-3	67644	34227	50.5987
	/lander-2	131170	59249	45.1696
	/home	115357	48546	42.0833
	/lander-5	37543	13851	36.8937

/lander-1 & /lander-4 have the highest bounce rate at 53.24% & 51.69%. While /lander-2 has 42.1% bounce rate. High bounce rates suggest visitors aren't finding what they expected, or the pages aren't engaging enough. This could be hurtful conversions.

ANALYZING & TESTING CONVERSION FUNNELS

Conversion funnel analysis is about understanding and optimizing each step of your user's experience on their journey toward purchasing your products
COMMON USE CASES:

- Identifying the most common paths customers take before purchasing your products
- Identifying how many of your users continue to each next step in your conversion flow, and how many users abandon at each step
- Optimizing critical pain points where users are abandoning, so that you can convert more users and sell more products

9. BUILDING CONVERSION FUNNELS

BUSINESS CONTEXT

- we want to build a mini conversion funnel, from /lander-2 to /cart
- we want to know how many people reach each step, and drop off rates
- we're looking at /lander-2 traffic only
- we're looking at customers who like Mr. Fuzzy only
- we're looking at gsearch nonbrand sessions only
- we're looking at /billing-2 page only

```

SELECT
COUNT(DISTINCT ws.website_session_id) AS total_sessions
, COUNT(CASE WHEN wpv.pageview_url = '/lander-2' THEN 1 ELSE NULL
END) / COUNT(DISTINCT ws.website_session_id) * 100 AS lander_2_ctr
, COUNT(CASE WHEN wpv.pageview_url = '/lander-2' THEN
ws.website_session_id ELSE NULL END) / COUNT(CASE WHEN wpv.pageview_url =
'/products' THEN ws.website_session_id ELSE NULL END) * 100 AS
lander_to_product_ctr

```

```

, COUNT(CASE WHEN wpv.pageview_url = '/the-original-mr-fuzzy' THEN

```

```

ws.website_session_id ELSE NULL END)/COUNT(CASE WHEN wpv.pageview_url =
'/products' THEN ws.website_session_id ELSE NULL END)*100 AS
product_to_mrfuzzy_ctr
, COUNT(CASE WHEN wpv.pageview_url = '/cart' THEN ws.website_session_id
ELSE NULL END)/COUNT(CASE WHEN wpv.pageview_url = '/the-original-mr-fuzzy'
THEN ws.website_session_id ELSE NULL END)*100 AS mrfuzzy_to_cart_ctr

, COUNT(CASE WHEN wpv.pageview_url = '/shipping' THEN
ws.website_session_id ELSE NULL END)/COUNT(CASE WHEN wpv.pageview_url =
'/cart' THEN ws.website_session_id ELSE NULL END)*100 AS
cart_to_shipping_ctr
, COUNT(CASE WHEN wpv.pageview_url = '/billing-2' THEN
ws.website_session_id ELSE NULL END)/COUNT(CASE WHEN wpv.pageview_url =
'/shipping' THEN ws.website_session_id ELSE NULL END)*100 AS
shipping_to_billing_ctr
, COUNT(CASE WHEN wpv.pageview_url = '/thank-you-for-your-order' THEN
ws.website_session_id ELSE NULL END)/COUNT(CASE WHEN wpv.pageview_url
= '/billing-2' THEN ws.website_session_id ELSE NULL END)*100 AS
billing_to_thankyou_ctr
FROM website_sessions AS ws
LEFT JOIN website_pageviews AS wpv
ON ws.website_session_id = wpv.website_session_id
AND utm_source = 'gsearch'
AND utm_campaign = 'nonbrand';

```

	total_sessions	lander_2_ctr	lander_to_product_ctr	product_to_mrfuzzy_ctr	mrfuzzy_to_cart_ctr	cart_to_shipping_ctr	shipping_to_billing_ctr	billing_to_thankyou_ctr
►	472871	21.3551	64.8056	62.5152	57.6391	67.6088	73.7731	67.2094

Summary of the conversion funnel for users who liked Mr Fuzzy and came from non-branded Google Search traffic:

64.8% of users who visit the /lander-2 page continue to the products page. 62.5% of users who view the products page visit the Mr. Fuzzy page.

57.6% of users who view the Mr. Fuzzy page add it to the cart. 67.6% of users who made it to the cart proceed to the shipping page. 73.8% of users proceed to billing pages.

67.2% of users successfully placed the order proceeding to thankyou page. The conversion funnel shows a good initial interest in Mr. Fuzzy (high click-through rates from lander-2 to product and product to Mr. Fuzzy pages).

ANALYSIS FOR CHANNEL MANAGEMENT

CHANNEL PORTFOLIO OPTIMIZATION

Analyzing a portfolio of marketing channels is about bidding efficiently and using data to maximize the effectiveness of your marketing budget

- Understanding which marketing channels are driving the most sessions and orders through your website

Comparing channel Characteristics

10. COMPARING CHANNEL CHARACTERISTICS OF GSEARCH AND BSEARCH BRAND AND NONBRAND CAMPAIGN AND THE PERCENTAGE OF TRAFFIC COMING FROM DIFFERENT DEVICES.

```
SELECT
    utm_source,
    utm_campaign,
    COUNT(DISTINCT website_session_id) AS sessions,
    COUNT(DISTINCT CASE WHEN device_type = 'mobile' THEN
website_session_id END) * 100 / COUNT(DISTINCT website_session_id) AS
pct_mobile,
    COUNT(DISTINCT CASE WHEN device_type = 'desktop' THEN
website_session_id END) * 100 / COUNT(DISTINCT website_session_id) AS
pct_desktop
FROM
    website_sessions
WHERE
    utm_source IN ('gsearch' , 'bsearch')
    AND YEAR(created_at) != '2015'
GROUP BY 1,2;
```

	utm_source	utm_campaign	sessions	pct_mobile	pct_desktop
▶	bsearch	brand	6545	13.2315	86.7685
	bsearch	nonbrand	48072	13.8293	86.1707
	gsearch	brand	27395	40.2847	59.7153
	gsearch	nonbrand	247564	31.1277	68.8723

GSearch drives most traffic (gsearch >> bsearch). Non-brand campaigns dominate for both search engines. Mobile traffic is significant, but desktop still leads. Interestingly, bsearch has a higher desktop share than gsearch. Consider prioritizing mobile optimization and investigate lower bsearchmobile traffic.

Cross Channel Bid Optimization

11. NONBRAND CONVERSION RATES FROM SESSION TO ORDER FOR GSEARCH AND BSEARCH, AND SLICE THEM BY DEVICE TYPE.

```
SELECT
    ws.device_type,
    ws.utm_source,
    #ws.utm_campaign,
    COUNT(DISTINCT ws.website_session_id) AS sessions,
    count(distinct o.order_id) as orders,
    count(distinct o.order_id)/COUNT(DISTINCT ws.website_session_id)*100
as conv_rate
FROM
    website_sessions ws
    LEFT JOIN
    orders o ON ws.website_session_id = o.website_session_id
WHERE YEAR(ws.created_at) != '2015' AND ws.utm_campaign='nonbrand'
group by 1,2;
```


	device_type	utm_source	sessions	orders	conv_rate
▶	desktop	bsearch	41424	3031	7.3170
	desktop	gsearch	170503	13404	7.8614
	mobile	bsearch	6648	206	3.0987
	mobile	gsearch	77061	2393	3.1053

Here's a conclusion on non-brand conversion rates:

Desktop conversions are higher for both search engines (gsearch: 7.86%, bsearch: 7.32%).

Mobile conversions are lower (gsearch: 3.10%, bsearch: 3.09%).

Bid up both the channels gsearch and bsearch for desktop as they perform identically.

Focus on improving mobile user experience or investigate lower mobile conversion rates

ANALYZING DIRECT TRAFFIC

Analyzing your branded or direct traffic is about keeping a pulse on how well your brand is doing with consumers, and how well your brand drives business

- Identifying how much revenue you are generating from direct traffic - this is high margin revenue without a direct cost of customer acquisition
- Understanding whether your paid traffic is generating a "halo" effect, and promoting additional direct traffic
- Assessing the impact of various initiatives on how many customers seek out your business

12. ANALYZING DIRECT TRAFFIC

```
SELECT
  CASE
    WHEN http_referer IS NULL THEN 'direct_type_in'
    WHEN http_referer='https://www.gsearch.com' and utm_source is null
  THEN 'gsearch_organic'
    WHEN http_referer='https://www.bsearch.com' and utm_source is null
  THEN 'bsearch_organic'
    ELSE 'others'
  END AS search_type,
  count(distinct website_session_id) as sessions
FROM
  website_sessions
GROUP BY 1
ORDER BY 2 DESC;
```

	search_type	sessions
▶	others	389543
	direct_type_in	39917
	gsearch_organic	35202
	bsearch_organic	8209

Direct traffic isn't the largest source (39,917 sessions).

"Others" category is largest (389,543 sessions), likely including paidmarketing.

Paid traffic might still influence brand awareness even if users come directly later.

Analyzing Free Channels

13. HOW THE ORGANIC SEARCH ,DIRECT TYPE IN AND PAID BRAND SEARCH ARE PERFORMING WITH RESPECT TO PAID NONBRAND SEARCH?

```
SELECT
    year(created_at) as yr,
    COUNT(DISTINCT case when utm_campaign='nonbrand' then
website_session_id end) AS nonbrand,
    COUNT(DISTINCT case when utm_campaign='brand' then website_session_id
end) AS brand,
    COUNT(DISTINCT case when utm_campaign='brand' then website_session_id
end)/
    COUNT(DISTINCT case when utm_campaign='nonbrand' then
website_session_id end)*100 AS brand_pct_of_nonbrand,
    COUNT(DISTINCT CASE WHEN http_referer IS NULL THEN website_session_id
end) AS direct,
    COUNT(DISTINCT CASE WHEN http_referer IS NULL THEN website_session_id
end)/
    COUNT(DISTINCT case when utm_campaign='nonbrand' then
website_session_id end)*100 AS direct_pct_of_nonbrand,
    COUNT(DISTINCT case when utm_source is null and http_referer is not
null then website_session_id end) as organic,
    COUNT(DISTINCT case when utm_source is null and http_referer is not
null then website_session_id end)/
    COUNT(DISTINCT case when utm_campaign='nonbrand' then
website_session_id end)*100 as organic_pct_of_nonbrand
FROM
    website_sessions
WHERE YEAR(created_at) != '2015'
GROUP BY YEAR(created_at);
```

	yr	nonbrand	brand	brand_pct_of_nonbrand	direct	direct_pct_of_nonbrand	organic	organic_pct_of_nonbrand
►	2012	53873	2844	5.2791	2780	5.1603	2973	5.5185
	2013	87242	8489	9.7304	8136	9.3258	8914	10.2176
	2014	154521	22607	14.6304	21877	14.1579	23732	15.3584

Paid non-brand drives the most traffic (highest non-brand counts). Organic search & direct traffic are like paid brand search. Focus on organic search optimization and brand building.

Analyze reasons behind lower conversion rate for paid brand search.

BUSINESS PATTERNS & SEASONALITY

ANALYZING SEASONALITY & BUSINESS PATTERN

Analyzing business patterns is about generating insights to help you maximize efficiency and anticipate future trends

- Day-parting analysis to understand how much support staff you should have at different times of day or days of the week
- Analyzing seasonality to better prepare for upcoming spikes or slowdowns in demand

14.ANALYZING SEASONALITY BY QUARTER

```
SELECT
    YEAR(ws.created_at) AS yr,
    QUARTER(ws.created_at) AS
    qtr,
    COUNT(DISTINCT ws.website_session_id) AS sessions,
    COUNT(DISTINCT o.order_id) AS orders,
    COUNT(DISTINCT ws.website_session_id)/COUNT(DISTINCT o.order_id) AS
    cvr
FROM
    website_sessions ws
    LEFT JOIN
    orders o ON ws.website_session_id = o.website_session_id
WHERE YEAR(ws.created_at) != '2015'
GROUP BY 1,2;
```

	yr	qtr	sessions	orders	cvr
►	2012	1	1879	60	31.3167
	2012	2	11433	347	32.9481
	2012	3	16892	684	24.6959
	2012	4	32266	1495	21.5826
	2013	1	19833	1273	15.5797
	2013	2	24745	1718	14.4034
	2013	3	27663	1840	15.0342
	2013	4	40540	2616	15.4969
	2014	1	46779	3069	15.2424
	2014	2	53129	3848	13.8069
	2014	3	57141	4035	14.1613
	2014	4	76373	5908	12.9270

Conversion rate (CVR) has decreased since launch (from 31.3 to 12.9), but sessions and orders have increased (2012-2014).

More users are coming to the site despite lower CVR. Holiday season (4th quarter) sees a spike in sessions.

Investigate CVR decline and optimize marketing for peak seasons.

Focus on attracting more users while improving conversion efficiency.

Analyzing Business Patterns

15. AVERAGE WEBSITE SESSION VOLUME, BY HOUR OF DAY AND DAY WEEK, TO UNDERSTAND HOW MUCH SUPPORT STAFF IS NEEDED DURING DIFFERENT HOURS.

```

SELECT
    hr,
    round(avg(website_session_id),1) as avg_sessions,
    round(avg(case when wkday=0 then website_session_id end),1) as mon,
    round(avg(case when wkday=1 then website_session_id end),1) as tue,
    round(avg(case when wkday=2 then website_session_id end),1) as wed,
    round(avg(case when wkday=3 then website_session_id end),1) as thu,
    round(avg(case when wkday=4 then website_session_id end),1) as fri,
    round(avg(case when wkday=5 then website_session_id end),1) as sat,
    round(avg(case when wkday=6 then website_session_id end),1) as sun
FROM
    (SELECT
        date(created_at) as created_date,
        weekday(created_at) as wkday,
        hour(created_at) as hr,
        count(distinct website_session_id) as website_session_id
    from website_sessions
    where YEAR(created_at)!='2015'
    group by 1,2,3) as daily_hourly_sessions
GROUP BY 1
order by 1;

```

	hr	avg_sessions	mon	tue	wed	thu	fri	sat	sun
▶	0	10.1	13.0	9.4	10.5	11.0	12.0	8.4	6.5
	1	8.5	11.6	8.6	8.1	8.1	10.0	7.1	5.6
	2	7.6	9.8	6.3	7.9	8.9	9.1	5.5	5.3
	3	6.8	9.2	6.3	7.5	6.6	7.6	5.2	4.9
	4	7.0	8.1	7.6	7.0	7.2	8.7	5.3	4.7
	5	7.5	9.9	7.5	8.0	7.6	9.0	5.2	4.8
	6	8.7	9.4	8.7	9.9	9.5	12.2	5.5	5.2
	7	10.6	13.1	10.9	11.6	12.3	13.5	7.2	5.2
	8	15.4	19.4	15.7	17.6	18.8	21.0	7.7	7.4
	9	21.9	27.1	22.8	27.7	25.9	30.5	10.5	8.7
	10	25.9	32.4	28.7	33.8	30.1	34.0	11.8	10.6
	11	26.9	32.7	32.8	34.2	30.4	36.0	11.7	10.1
	12	27.1	34.4	31.7	33.7	33.1	33.5	12.8	10.1
	13	26.1	32.7	30.5	32.6	31.9	32.3	11.7	10.9
	14	26.1	31.8	30.5	32.0	30.4	33.7	13.0	11.2
	15	26.4	32.4	30.3	33.8	31.0	34.8	11.0	11.8
	16	26.1	30.7	31.7	33.3	30.5	34.3	10.6	11.4
	17	23.0	31.6	26.0	28.1	26.9	26.0	9.9	11.7
	18	19.0	23.7	22.3	23.2	22.2	22.2	9.0	10.0
	19	16.5	21.2	19.4	19.0	18.2	18.9	8.6	10.1
	20	15.3	17.8	17.9	19.2	16.1	15.2	8.6	12.3
	21	15.0	18.4	16.3	17.1	16.8	13.9	7.8	14.7
	22	13.8	16.0	14.7	16.8	14.4	12.2	8.8	13.6
	23	12.2	13.4	13.0	12.6	13.0	11.9	8.1	13.3

*Weekday traffic peaks during business hours (8 AM - 5 PM),
with a lunchtime rush (12 PM - 1 PM).*

Weekends see lower traffic than weekdays.

*Mondays have slightly lower traffic than other weekdays.
Fridays might have slightly higher traffic than other
weekdays.*

*Allocate more support staff weekdays (especially lunchtime
& business hours).*

PRODUCT ANALYSIS

PRODUCT SALES ANALYSIS

Analyzing product sales helps you understand how each product contributes to your business, and how product launches impact the overall portfolio

- Analyzing sales and revenue by product
- Monitoring the impact of adding a new product to your product portfolio
- Watching product sales trends to understand the overall health of your business

Product level sales analysis

16.SHOW QUARTERLY FIGURES SINCE THE LAUNCH, FOR SESSION-TO-ORDER CONVERSION RATE, REVENUE & MARGIN PER ORDER, AND REVENUE & MARGIN PER SESSION.

```
SELECT
    YEAR(website_sessions.created_at) AS yr,
    QUARTER(website_sessions.created_at) AS qtr,
    COUNT(DISTINCT orders.order_id)/COUNT(DISTINCT
website_sessions.website_session_id) AS session_to_order_conv_rate,
    SUM(price_usd)/COUNT(DISTINCT orders.order_id) AS revenue_per_order,
    SUM(price_usd)/COUNT(DISTINCT website_sessions.website_session_id) AS
revenue_per_session,
    SUM(price_usd-cogs_usd)/COUNT(DISTINCT orders.order_id) AS
margin_per_order,
    SUM(price_usd-cogs_usd)/COUNT(DISTINCT
website_sessions.website_session_id) AS margin_per_session
FROM website_sessions
    LEFT JOIN orders
        ON website_sessions.website_session_id =
orders.website_session_id
WHERE YEAR(website_sessions.created_at)!='2015'
GROUP BY 1,2
ORDER BY 1,2;
```

	yr	qtr	session_to_order_conv_rate	revenue_per_order	revenue_per_session	margin_per_order	margin_per_session
►	2012	1	0.0319	49.990000	1.596275	30.500000	0.973922
	2012	2	0.0304	49.990000	1.517233	30.500000	0.925698
	2012	3	0.0405	49.990000	2.024222	30.500000	1.235022
	2012	4	0.0463	49.990000	2.316217	30.500000	1.413175
	2013	1	0.0642	52.142396	3.346809	32.006677	2.054379
	2013	2	0.0694	51.538312	3.578211	31.583818	2.192807
	2013	3	0.0665	51.734533	3.441114	31.706522	2.108954
	2013	4	0.0645	54.715688	3.530741	33.735092	2.176887
	2014	1	0.0656	62.160684	4.078136	39.164875	2.569465
	2014	2	0.0724	64.374207	4.662462	40.665670	2.945312
	2014	3	0.0706	64.494949	4.554298	40.757125	2.878056
	2014	4	0.0774	63.793497	4.934885	40.384056	3.123997

Revenue, margin, and session-to-order conversion rate (conv_rate) have all increased consistently over the years (2012-2014).

Both higher conversion rate and more website traffic (sessions) contributed to revenue and margin growth.

The website shows strong performance with increasing customer acquisition and revenue generation.

PRODUCT LEVEL WEBSITE ANALYSIS

Product-focused website analysis is about learning how customers interact with each of your products, and how well each product converts customers

Product conversion Funnels

17. COMPARING CONVERSION FUNNELS FROM EACH PRODUCT PAGE TO CART.

```

WITH products_sessions AS(
SELECT
    CASE WHEN pageview_url='/the-original-mr-fuzzy' THEN 'mrfuzzy'
    WHEN pageview_url='/the-forever-love-bear' THEN 'lovebear'
    WHEN pageview_url='/the-birthday-sugar-panda' THEN 'sugarpanda'
    WHEN pageview_url='/the-hudson-river-mini-bear' THEN 'minibear'
    ELSE 'error' END AS product_seen,
    website_session_id, website_pageview_id
FROM website_pageviews
WHERE pageview_url IN ('/the-original-mr-fuzzy', '/the-forever-love-bear', '/the-birthday-sugar-panda', '/the-hudson-river-mini-bear')
AND YEAR(created_at) != '2015')

SELECT product_seen,
sessions,
to_cart/sessions AS product_page_click_rt,
    to_shipping/to_cart AS cart_click_rt,
    to_billing/to_shipping AS shipping_click_rt,
    to_thankyou/to_billing AS billing_click_rt

FROM
    (SELECT product_seen,
COUNT(DISTINCT ps.website_session_id) AS sessions,
COUNT(CASE WHEN wp.pageview_url='/cart' THEN '1' ELSE NULL END) AS
to_cart,
COUNT(CASE WHEN wp.pageview_url='/shipping' THEN '1' ELSE NULL END) AS

```

```

to_shipping,
COUNT(CASE WHEN wp.pageview_url='/billing-2' THEN '1' ELSE NULL END)
AS to_billing,
COUNT(CASE WHEN wp.pageview_url='/thank-you-for-your-order' THEN '1'
ELSE NULL END) AS to_thankyou
FROM products_sessions ps
LEFT JOIN website_pageviews wp
ON ps.website_session_id = wp.website_session_id
AND wp.website_pageview_id >= ps.website_pageview_id
GROUP BY 1) AS funnel_sessions
ORDER BY 2 DESC;

```

	product_seen	sessions	product_page_click_rt	cart_click_rt	shipping_click_rt	billing_click_rt
▶	mrfuzzy	141247	0.4303	0.6812	0.7233	0.6879
	lovebear	20585	0.5560	0.6680	0.7939	0.6183
	sugarpanda	14769	0.4647	0.6794	0.8031	0.6385
	minibear	711	0.6498	0.6515	0.8206	0.6113

Mr. Fuzzy has the highest brand awareness (sessions) but lower click-through rates in the funnel.

Love Bear has strong click-through rates throughout the funnel, suggesting a well-optimized experience.

Sugar Panda shows good click-through rates like Love Bear. Mini Bear has lower click-through rates, indicating room for funnel optimization.

Analyze user behavior and optimize the funnel for each product to improve conversion rates.

PRODUCT REFUND ANALYSIS

Analyzing product refund rates is about controlling for quality and understanding where you might have problems to address.

18. PRODUCT REFUND RATES

```

SELECT
YEAR(oi.created_at) AS yr,
QUARTER(oi.created_at) AS qtr,

COUNT(DISTINCT CASE WHEN product_id=1 THEN oi.order_item_id ELSE NULL
END) AS p1_orders,
COUNT(DISTINCT CASE WHEN product_id=1 THEN oir.order_item_id ELSE NULL
END) /
COUNT(DISTINCT CASE WHEN product_id=1 THEN oi.order_item_id ELSE NULL
END)*100 AS p1_refund_rt,
COUNT(DISTINCT CASE WHEN product_id=2 THEN oi.order_item_id ELSE NULL
END) AS p2_orders,
COUNT(DISTINCT CASE WHEN product_id=2 THEN oir.order_item_id ELSE
NULL)END) /
COUNT(DISTINCT CASE WHEN product_id=2 THEN oi.order_item_id ELSE NULL

```

```

END)*100 AS p2_refund_rt,
    COUNT(DISTINCT CASE WHEN product_id=3 THEN oi.order_item_id ELSE NULL
END) AS p3_orders,
    COUNT(DISTINCT CASE WHEN product_id=3 THEN oir.order_item_id ELSE NULL
END)/
    COUNT(DISTINCT CASE WHEN product_id=3 THEN oi.order_item_id ELSE NULL
END)*100 AS p3_refund_rt,
    COUNT(DISTINCT CASE WHEN product_id=4 THEN oi.order_item_id ELSE NULL
END) AS p4_orders,
    COUNT(DISTINCT CASE WHEN product_id=4 THEN oir.order_item_id ELSE NULL
END)/
    COUNT(DISTINCT CASE WHEN product_id=4 THEN oi.order_item_id ELSE NULL
END)*100 AS p4_refund_rt
FROM
    order_items oi
    LEFT JOIN
        order_item_refunds oir ON oir.order_item_id = oi.order_item_id
WHERE YEAR(oi.created_at) != '2015'
GROUP BY 1,2;

```

	yr	qtr	p1_orders	p1_refund_rt	p2_orders	p2_refund_rt	p3_orders	p3_refund_rt	p4_orders	p4_refund_rt
►	2012	1	60	1.6667	0	NULL	0	NULL	0	NULL
	2012	2	347	4.8991	0	NULL	0	NULL	0	NULL
	2012	3	684	8.3333	0	NULL	0	NULL	0	NULL
	2012	4	1495	6.8896	0	NULL	0	NULL	0	NULL
	2013	1	999	5.9059	274	2.1898	0	NULL	0	NULL
	2013	2	1451	6.1337	266	3.0075	0	NULL	0	NULL
	2013	3	1556	5.6555	291	1.7182	0	NULL	0	NULL
	2013	4	2145	2.8438	492	2.0325	139	7.1942	0	NULL
	2014	1	2097	3.7196	727	1.7882	655	6.7176	407	0.7371
	2014	2	2840	3.9789	705	2.4113	854	5.9719	806	1.4888
	2014	3	2975	10.5546	732	2.8689	887	5.8625	894	1.1186
	2014	4	4208	3.2795	1048	2.0992	1332	5.6306	1414	1.5559

Product 1: Highest refund rate, especially in Q3 of 2012 and 2014. Investigate customer reviews and product changes for those quarters. Product 2: Consistently low refund rates, indicating good performance. Product 3: High refund rates initially but decreasing over time. Possible improvements made to the product or customer service.

Product 4: Low refund rates, like Product 2.

Analyze customer feedback, product changes, and marketing campaigns to understand refund rate drivers and take targeted actions to potentially reduce them, especially for Product 1. This can improve customer satisfaction and potentially increase revenue.

USER ANALYSIS

ANALYZE REPEAT BEHAVIOR

Analyzing repeat visits helps you understand user behavior and identify some of your most valuable customers

- Analyzing repeat activity to see how often customers are coming back to visit your site
- Understanding which channels, they use when they come back, and whether you are paying for them again through paid channels

19. IDENTIFYING REPEAT VISITORS

```
SELECT
    repeats,
    COUNT(DISTINCT user_id) AS users
FROM
    (SELECT
        user_id,
        sum(is_repeat_session) as repeats
    FROM
        website_sessions

    WHERE YEAR(created_at) != '2015'
    GROUP BY 1
    HAVING MIN(is_repeat_session) = 0
    ) AS repeat_session
GROUP BY 1
ORDER BY 1;
```

	repeats	users
▶	0	301979
	1	31076
	2	466
	3	10786

"repeats = 0" represents new visitors, with a significantly high number compared to repeat visitors.

There are also significantly fewer visitors on their third visit ("repeats = 2") compared to both second-time visitors and potentially even fourth-time visitors ("repeats = 3").

This data suggests a need to focus on strategies to convert new visitors into repeat customers, with particular attention to keeping users engaged after their second visit.

Analyzing Repeat Behavior

20. TO UNDERSTAND THE BEHAVIOR OF THESE REPEAT CUSTOMERS, LET'S CHECK THE MINIMUM, THE MAXIMUM AND AVERAGE TIME BETWEEN THE FIRST AND THE SECOND SESSION FOR CUSTOMERS WHO DO COME BACK.

```
with user_next_session_date as(
SELECT user_id, website_session_id,
    is_repeat_session,
    created_at,
    lead(created_at) over(partition by user_id order by
website_session_id) as next_session_date
FROM
    website_sessions
where YEAR(created_at) != '2015')

select
    avg(datediff(next_session_date, created_at)) as avg_diff,
```

```

min(datediff(next_session_date,created_at)) as min_diff,
max(datediff(next_session_date,created_at)) as max_diff
  from user_next_session_date
where is_repeat_session=0
and next_session_date is not null;

```

	avg_diff	min_diff	max_diff
▶	33.8225	1	69

Some users revisit almost immediately the next day.

On average, repeat visitors return after about a month (avg 33.82 days). The maximum time between the first and second visit is 69 days.

Analyze user behavior and segment repeat visitors to understand their browsing patterns and tailor strategies to keep them engaged and potentially increase customer loyalty and revenue.

21. ANALYZING REPEAT CHANNEL BEHAVIOR TO FIND OUT THE CHANNELS OUR REPEAT VISITORS COME BACK THROUGH.

```

SELECT
  case when utm_source is null and http_referer in
    ('https://www.gsearch.com','https://www.bsearch.com') then
    'organic_search'
    when utm_campaign='nonbrand' then 'paid_nonbrand'
    when utm_campaign='brand' then 'paid_brand'
    when utm_source is null and http_referer is null then
    'direct_type_in'
    when utm_source ='socialbook' then 'paid_social'
  end as channel_group,
  count(case when is_repeat_session=0 then website_session_id end) as
new_sessions,
  count(case when is_repeat_session=1 then website_session_id end) as
repeat_sessions
FROM
  website_sessions
where YEAR(created_at) != '2015'
group by 1
order by 3 desc;

```

	channel_group	new_sessions	repeat_sessions
▶	organic_search	13376	22243
	paid_brand	12463	21477
	direct_type_in	12147	20646
	paid_nonbrand	295636	0
	paid_social	10685	0

Organic search is the top channel for repeat visitors, followed by paid brand and direct type-in.

Paid non-brand and paid social have lower repeat session

counts compared to others.

This suggests users from organic search, brand awareness, or direct visits are more likely to return.

New vs. Repeat performance

22. LET'S COMPARE THE CONVERSION RATES AND REVENUE PER SESSION FOR REPEAT SESSIONS VS NEW SESSIONS.

```
SELECT
    is_repeat_session,
    COUNT(DISTINCT ws.website_session_id) AS sessions,
    COUNT(DISTINCT order_id)/COUNT(DISTINCT ws.website_session_id) AS
conv_rate,
    SUM(price_usd)/COUNT(DISTINCT ws.website_session_id) AS
revenue_per_session

FROM
    website_sessions ws
    LEFT JOIN
    orders o ON ws.website_session_id = o.website_session_id
WHERE
    YEAR(ws.created_at) != '2015'
GROUP BY 1;
```

	is_repeat_session	sessions	conv_rate	revenue_per_session
▶	0	344307	0.0637	3.780333
	1	64366	0.0768	4.607075

Repeat visitors (is_repeat_session = 1) have higher conversion rates(conv_rate) and revenue per session compared to new visitors.

This suggests repeat visitors are more likely to buy and spend more per purchase.

Possible reasons include brand familiarity, higher engagement, and trust built from past experiences.

23. LET'S CHECK THE CUSTOMER LIFETIME VALUE (CLTV)

```
WITH AOV AS (
SELECT
    user_id, (items_purchased * price_usd * COUNT(*)) AS
AOVPurchase_frequency
FROM
    orders
GROUP BY 1),
customer_lifespan AS
(SELECT
    user_id, AVG(datediff(next_session_date, created_at)) as avg_diff
FROM
```

```

(
SELECT user_id,website_session_id,
       is_repeat_session,
       created_at,
       lead(created_at) over(partition by user_id order by
website_session_id) as next_session_date
FROM
       website_sessions
where YEAR(created_at)!='2015') as user_next_session_date
       where is_repeat_session=0
       and next_session_date is not null
GROUP BY 1),
CLTV AS
(SELECT a.user_id,AOVPurchase_frequency*avg_diff AS CLTV_per_user
FROM customer_lifespan cl
JOIN AOV a  ON a.user_id=cl.user_id )

SELECT AVG(CLTV_per_user)/2 AS Avg_CLTV
FROM CLTV;

```

	Avg_CLTV
▶	1433.90639267229638

Average CLTV of USD 1433.91 per year per user indicates that customers who purchase from us tend to generate significant revenue over time, reflecting strong customer relationships and the potential for recurring revenue.