

MYSQL PROJECT  
MAVEN FUZZY FACTORY ECOMMERCE  
DATABASE ANALYST PROJECT

BY SHUBHAM MITTAL

# PROJECT OBJECTIVE

Utilize SQL to optimize marketing channels, website performance, and product portfolio.

## Data Exploration with SQL

- Access and explore Maven Fuzzy Factory's database.
- Uncover valuable insights to support decision-making.

## Becoming the Data Expert

- Serve as the go-to person for critical analyses.
- Drive strategic growth for the company.

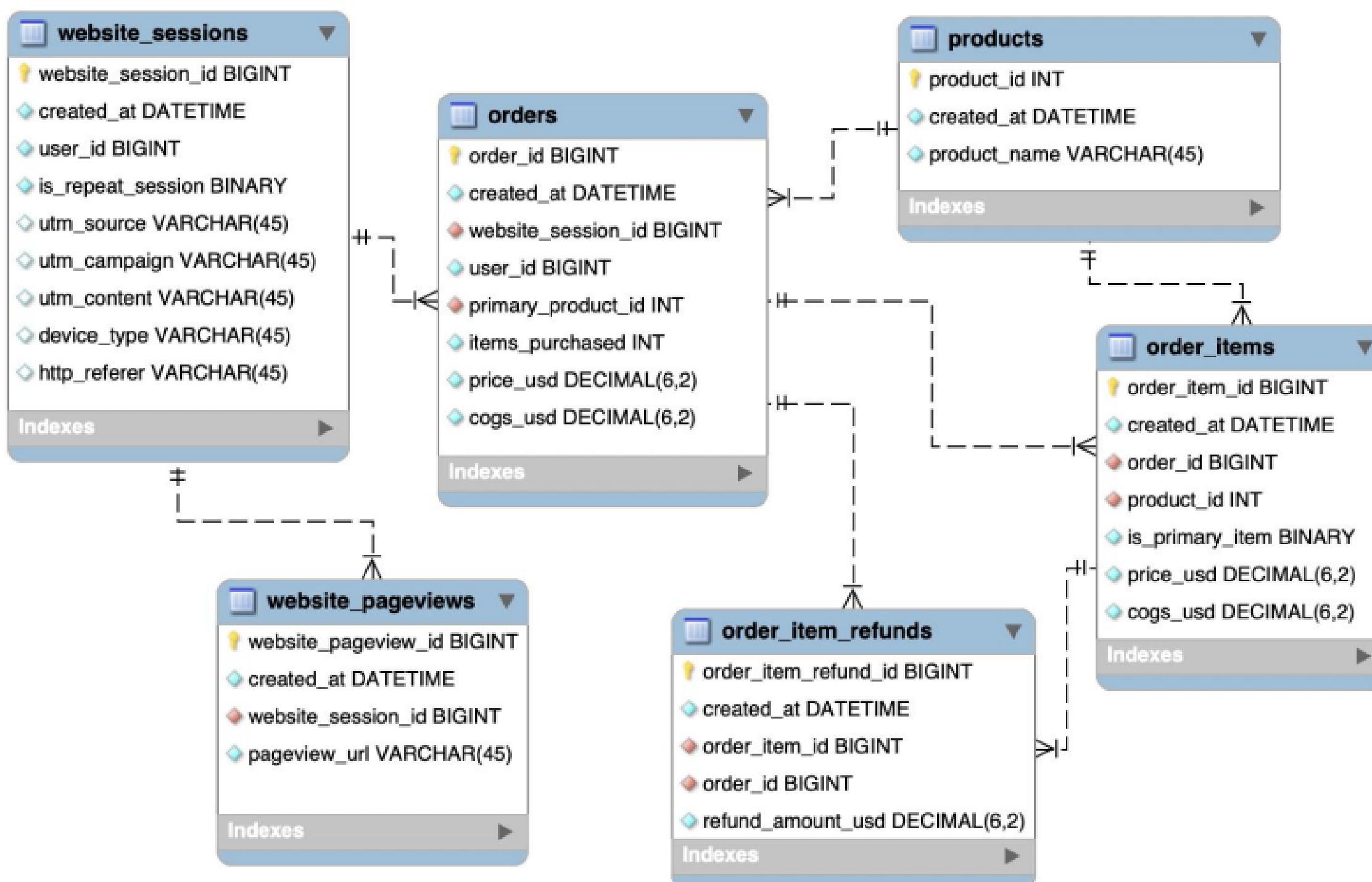
## Optimizing Marketing Channels

- Identify effective channels and allocate resources wisely.
- Refine marketing strategies for optimal results.

## Enhancing Website Performance

- Measure and test website conversion rates.
- Implement improvements for better customer engagement.

# DATA BASE OVERVIEW



We will be working with six related tables, which contain eCommerce data about:

- Website Activity
- Products
- Orders and Refunds

We'll use MySQL to understand how customers access and interact with the site, analyze landing page performance and conversion, and explore product-level sales

# Analyzing Monthly Trends for Gsearch Sessions and Orders

## MySQL Code

```
SELECT
    YEAR(website_sessions.created_at) AS yr,
    MONTH(website_sessions.created_at) AS mo,
    COUNT(DISTINCT website_sessions.website_session_id) AS sessions,
    COUNT(DISTINCT orders.order_id) AS orders,
    COUNT(DISTINCT orders.order_id)/COUNT(DISTINCT website_sessions.website_session_id) AS conv_rate
FROM website_sessions
LEFT JOIN orders
    ON orders.website_session_id = website_sessions.website_session_id
WHERE website_sessions.utm_source = 'gsearch'
GROUP BY 1,2;
```

## Observation

- Steady growth in sessions from 2012 to 2014 indicates increased website traffic over time.
- Conversion rate remains relatively stable, ranging from 2.57% to 7.37%, with the highest rate in January 2014.
- Seasonal fluctuations in conversion rates suggest potential impact from seasonal factors and marketing strategies.

## Result

yr	mo	sessions	orders	conv_rate
2012	4	3574	92	0.0257
2012	5	3410	97	0.0284
2012	6	3578	121	0.0338
2012	7	3811	145	0.0380
2012	8	4877	184	0.0377
2012	9	4491	188	0.0419
2012	10	5534	234	0.0423
2012	11	9715	417	0.0429
2012	12	7038	333	0.0473
2013	1	4207	244	0.0580
2013	2	5124	341	0.0665
2013	3	4433	268	0.0605
2013	4	5741	398	0.0693
2013	5	5984	390	0.0652
2013	6	5873	422	0.0719
2013	7	6243	393	0.0630
2013	8	6457	424	0.0657
2013	9	6765	438	0.0647
2013	10	7520	504	0.0670
2013	11	10265	629	0.0613
2013	12	11060	713	0.0645
2014	1	8607	634	0.0737

# Analyzing Monthly Trends for Gsearch - Brand vs. Non-Brand Campaigns

## MySQL Code

```
SELECT
    YEAR(website_sessions.created_at) AS yr,
    MONTH(website_sessions.created_at) AS mo,
    COUNT(DISTINCT CASE WHEN utm_campaign = 'nonbrand' THEN website_sessions.website_session_id ELSE NULL END) AS nonbrand_sessions,
    COUNT(DISTINCT CASE WHEN utm_campaign = 'nonbrand' THEN orders.order_id ELSE NULL END) AS nonbrand_orders,
    COUNT(DISTINCT CASE WHEN utm_campaign = 'brand' THEN website_sessions.website_session_id ELSE NULL END) AS brand_sessions,
    COUNT(DISTINCT CASE WHEN utm_campaign = 'brand' THEN orders.order_id ELSE NULL END) AS brand_orders
FROM website_sessions
LEFT JOIN orders
    ON orders.website_session_id = website_sessions.website_session_id
    AND website_sessions.utm_source = 'gsearch'
GROUP BY 1,2;
```

## Observation

- Nonbrand sessions consistently exceed brand sessions, emphasizing the significance of nonbranded searches in driving traffic.
- Nonbrand conversion rates fluctuate between 2.5% to 6%, while brand sessions show negligible conversions.
- Seasonal patterns impact both nonbrand and brand sessions, with peaks during certain months. Brand orders remain minimal, suggesting limited impact on new orders.

## Result

yr	mo	nonbrand_sessions	nonbrand_orders	brand_sessions	brand_orders
2012	3	1852	60	10	0
2012	4	3509	86	76	6
2012	5	3295	91	140	6
2012	6	3439	114	164	7
2012	7	3660	136	195	9
2012	8	5318	174	264	10
2012	9	5591	172	339	16
2012	10	6883	219	432	15
2012	11	12260	396	556	21
2012	12	8066	298	668	35
2013	1	4502	213	630	31
2013	2	5702	302	468	39
2013	3	4950	251	438	17
2013	4	6480	366	496	32
2013	5	6654	364	605	26
2013	6	6550	384	579	38
2013	7	6817	359	718	34
2013	8	7023	381	722	43
2013	9	7381	392	736	46
2013	10	8126	451	856	53
2013	11	11157	574	960	55
2013	12	11900	632	1281	81
2014	1	9114	543	1369	91
2014	2	9869	556	1360	84

# Analyzing Gsearch Nonbrand Traffic - Monthly Sessions and Orders by Device Type

## MYSQL Code

```
SELECT
    YEAR(website_sessions.created_at) AS yr,
    MONTH(website_sessions.created_at) AS mo,
    COUNT(DISTINCT CASE WHEN device_type = 'desktop' THEN website_sessions.website_session_id ELSE NULL END) AS desktop_sessions,
    COUNT(DISTINCT CASE WHEN device_type = 'desktop' THEN orders.order_id ELSE NULL END) AS desktop_orders,
    COUNT(DISTINCT CASE WHEN device_type = 'mobile' THEN website_sessions.website_session_id ELSE NULL END) AS mobile_sessions,
    COUNT(DISTINCT CASE WHEN device_type = 'mobile' THEN orders.order_id ELSE NULL END) AS mobile_orders
FROM website_sessions
LEFT JOIN orders
    ON orders.website_session_id = website_sessions.website_session_id
WHERE website_sessions.utm_source = 'gsearch'
    AND website_sessions.utm_campaign = 'nonbrand'
GROUP BY 1,2;
```

## Observation

- Desktop sessions consistently exceed mobile sessions, indicating a higher desktop user base.
- Mobile conversions show gradual improvement, suggesting mobile optimization opportunities.
- Seasonal variations impact both device types, with higher activity during certain months.

## Result

yr	mo	desktop_sessions	desktop_orders	mobile_sessions	mobile_orders
2012	3	1128	50	724	10
2012	4	2139	75	1370	11
2012	5	2276	83	1019	8
2012	6	2673	106	766	8
2012	7	2774	122	886	14
2012	8	3515	165	1158	9
2012	9	3171	155	1056	17
2012	10	3934	201	1263	18
2012	11	7007	354	2250	42
2012	12	4947	282	1548	16
2013	1	2795	190	896	23
2013	2	3603	271	1139	31
2013	3	3082	224	997	27
2013	4	4058	335	1283	31
2013	5	4247	321	1261	43
2013	6	4093	351	1309	33
2013	7	4328	316	1326	43
2013	8	4174	321	1703	60
2013	9	3991	329	2183	63
2013	10	4482	385	2346	66
2013	11	6124	461	3364	113
2013	12	6552	505	3470	127
2014	1	4948	451	2552	92
2014	2	5393	454	2830	102

# Analyzing Monthly Traffic Trends for Gsearch and Other Marketing Channels

## MySQL Code

```
SELECT DISTINCT
    utm_source,
    utm_campaign,
    http_referer
FROM website_sessions;
SELECT
    YEAR(website_sessions.created_at) AS yr,
    MONTH(website_sessions.created_at) AS mo,
    COUNT(DISTINCT CASE WHEN utm_source = 'gsearch' THEN website_sessions.website_session_id ELSE NULL END) AS gsearch_paid_sessions,
    COUNT(DISTINCT CASE WHEN utm_source = 'bsearch' THEN website_sessions.website_session_id ELSE NULL END) AS bsearch_paid_sessions,
    COUNT(DISTINCT CASE WHEN utm_source IS NULL AND http_referer IS NOT NULL THEN website_sessions.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_sessions.website_session_id ELSE NULL END) AS direct_type_in_sessions
FROM website_sessions
LEFT JOIN orders
    ON orders.website_session_id = website_sessions.website_session_id
GROUP BY 1,2;
```

## Result

yr	mo	gsearch_paid_sessions	bsearch_paid_sessions	organic_search_sessions	direct_type_in_sessions
2012	3	1860	2	8	9
2012	4	3574	11	78	71
2012	5	3410	25	150	151
2012	6	3578	25	190	170
2012	7	3811	44	207	187
2012	8	4877	705	265	250
2012	9	4491	1439	331	285
2012	10	5534	1781	428	440
2012	11	9715	3101	624	571
2012	12	7038	1696	692	646
2013	1	4207	925	662	607
2013	2	5124	1046	528	470
2013	3	4433	955	471	405
2013	4	5741	1235	512	483
2013	5	5984	1275	626	564
2013	6	5873	1256	625	571
2013	7	6243	1292	728	640
2013	8	6457	1288	773	662
2013	9	6765	1352	775	688
2013	10	7520	1462	921	870
2013	11	10265	1852	969	946
2013	12	11060	2121	1324	1230
2014	1	8607	1876	1404	1320
2014	2	9311	1016	1506	1313

## Observation

- Gsearch consistently dominates as the primary traffic source, contributing to the majority of sessions each month.
- Gsearch\_paid and bsearch\_paid sessions indicate growing impact from paid search efforts, while organic\_paid and direct\_paid sessions show stable contributions from organic and direct traffic sources.
- Analyzing conversion rates for each channel is crucial to understand their effectiveness in driving conversions.

# Analyzing Conversion Funnel for Landing Page Test

## MYSQL Code

```
SELECT
    website_sessions.website_session_id,
    website_pageviews.pageview_url,
    -- website_pageviews.created_at AS pageview_created_at,
    CASE WHEN pageview_url = '/home' THEN 1 ELSE 0 END AS homepage,
    CASE WHEN pageview_url = '/lander-1' THEN 1 ELSE 0 END AS custom_lander,
    CASE WHEN pageview_url = '/products' THEN 1 ELSE 0 END AS products_page,
    CASE WHEN pageview_url = '/the-original-mr-fuzzy' THEN 1 ELSE 0 END AS mrfuzzy_page,
    CASE WHEN pageview_url = '/cart' THEN 1 ELSE 0 END AS cart_page,
    CASE WHEN pageview_url = '/shipping' THEN 1 ELSE 0 END AS shipping_page,
    CASE WHEN pageview_url = '/billing' THEN 1 ELSE 0 END AS billing_page,
    CASE WHEN pageview_url = '/thank-you-for-your-order' THEN 1 ELSE 0 END AS thankyou_page
FROM website_sessions
LEFT JOIN website_pageviews
    ON website_sessions.website_session_id = website_pageviews.website_session_id
WHERE website_sessions.utm_source = 'gsearch'
    AND website_sessions.utm_campaign = 'nonbrand'
    AND website_sessions.created_at < '2012-07-28'
    AND website_sessions.created_at > '2012-06-19'
ORDER BY
    website_sessions.website_session_id,
    website_pageviews.created_at;
```

```
-- then this would produce the final output, part 1
SELECT
    CASE
        WHEN saw_homepage = 1 THEN 'saw_homepage'
        WHEN saw_custom_lander = 1 THEN 'saw_custom_lander'
        ELSE 'uh oh... check logic'
    END AS segment,
    COUNT(DISTINCT website_session_id) AS sessions,
    COUNT(DISTINCT CASE WHEN product_made_it = 1 THEN website_session_id ELSE NULL END) AS to_products,
    COUNT(DISTINCT CASE WHEN mrfuzzy_made_it = 1 THEN website_session_id ELSE NULL END) AS to_mrfuzzy,
    COUNT(DISTINCT CASE WHEN cart_made_it = 1 THEN website_session_id ELSE NULL END) AS to_cart,
    COUNT(DISTINCT CASE WHEN shipping_made_it = 1 THEN website_session_id ELSE NULL END) AS to_shipping,
    COUNT(DISTINCT CASE WHEN billing_made_it = 1 THEN website_session_id ELSE NULL END) AS to_billing,
    COUNT(DISTINCT CASE WHEN thankyou_made_it = 1 THEN website_session_id ELSE NULL END) AS to_thankyou
FROM session_level_made_it_flagged
GROUP BY 1;
```

```
1 CREATE TEMPORARY TABLE session_level_made_it_flagged
SELECT
    website_session_id,
    MAX(homepage) AS saw_homepage, MAX(custom_lander) AS saw_custom_lander, MAX(products_page) AS product_made_it,
    MAX(mrfuzzy_page) AS mrfuzzy_made_it, MAX(cart_page) AS cart_made_it, MAX(shipping_page) AS shipping_made_it,
    MAX(billing_page) AS billing_made_it, MAX(thankyou_page) AS thankyou_made_it
) AS pageview_level GROUP BY website_session_id;
```

```
2
    FROM(SELECT
        website_sessions.website_session_id,
        website_pageviews.pageview_url,
        -- website_pageviews.created_at AS pageview_created_at,
        CASE WHEN pageview_url = '/home' THEN 1 ELSE 0 END AS homepage, CASE WHEN pageview_url = '/lander-1' THEN 1 ELSE 0 END AS custom_lander,
        CASE WHEN pageview_url = '/products' THEN 1 ELSE 0 END AS products_page, CASE WHEN pageview_url = '/the-original-mr-fuzzy' THEN 1 ELSE 0 END AS mrfuzzy_page,
        CASE WHEN pageview_url = '/cart' THEN 1 ELSE 0 END AS cart_page, CASE WHEN pageview_url = '/shipping' THEN 1 ELSE 0 END AS shipping_page,
        CASE WHEN pageview_url = '/billing' THEN 1 ELSE 0 END AS billing_page, CASE WHEN pageview_url = '/thank-you-for-your-order' THEN 1 ELSE 0 END AS thankyou_page
    FROM website_sessions
    LEFT JOIN website_pageviews
        ON website_sessions.website_session_id = website_pageviews.website_session_id
    WHERE website_sessions.utm_source = 'gsearch' AND website_sessions.utm_campaign = 'nonbrand' AND website_sessions.created_at < '2012-07-28'
    AND website_sessions.created_at > '2012-06-19'
    ORDER BY
        website_sessions.website_session_id, website_pageviews.created_at
    ) AS pageview_level GROUP BY website_session_id;
```

```
3
    SELECT
        CASE
            WHEN saw_homepage = 1 THEN 'saw_homepage'
            WHEN saw_custom_lander = 1 THEN 'saw_custom_lander'
            ELSE 'uh oh... check logic'
        END AS segment,
        COUNT(DISTINCT CASE WHEN product_made_it = 1 THEN website_session_id ELSE NULL END)/COUNT(DISTINCT website_session_id) AS lander_click_rt,
        COUNT(DISTINCT CASE WHEN mrfuzzy_made_it = 1 THEN website_session_id ELSE NULL END)/COUNT(DISTINCT CASE WHEN product_made_it = 1 THEN website_session_id ELSE NULL END) AS products_click_rt,
        COUNT(DISTINCT CASE WHEN cart_made_it = 1 THEN website_session_id ELSE NULL END)/COUNT(DISTINCT CASE WHEN mrfuzzy_made_it = 1 THEN website_session_id ELSE NULL END) AS mrfuzzy_click_rt,
        COUNT(DISTINCT CASE WHEN shipping_made_it = 1 THEN website_session_id ELSE NULL END)/COUNT(DISTINCT CASE WHEN cart_made_it = 1 THEN website_session_id ELSE NULL END) AS cart_click_rt,
        COUNT(DISTINCT CASE WHEN billing_made_it = 1 THEN website_session_id ELSE NULL END)/COUNT(DISTINCT CASE WHEN shipping_made_it = 1 THEN website_session_id ELSE NULL END) AS shipping_click_rt,
        COUNT(DISTINCT CASE WHEN thankyou_made_it = 1 THEN website_session_id ELSE NULL END)/COUNT(DISTINCT CASE WHEN billing_made_it = 1 THEN website_session_id ELSE NULL END) AS billing_click_rt
    FROM session_level_made_it_flagged
    GROUP BY 1;
```

```
4
```

# Data Exploration with SQL: Analyzing Conversion Funnel for Landing Page Test

## Result

	segment	lander_click_rt	products_click_rt	mrfuzzy_click_rt	cart_click_rt	shipping_click_rt	billing_click_rt
▶	saw_custom_lander	0.4676	0.7128	0.4508	0.6638	0.8528	0.4772
	saw_homepage	0.4166	0.7261	0.4327	0.6757	0.8400	0.4286

## Observation

- Click-through rates vary across funnel stages, with "billing\_click\_rt" showing the highest engagement.
- Users who saw the custom landing page maintain consistent engagement, while "saw\_homepage" users exhibit higher rates in certain stages.
- These insights suggest opportunities for refining user experience and optimizing conversion rates in the funnel.

# THANK YOU

FOR YOUR ATTENTION