

Lab 2 – Phân tích lưu lượng Website

Trong bài học này, các bạn sẽ đóng vai trò là một nhà phân tích dữ liệu để thực hiện phân tích lưu lượng website từ đó làm gia tăng hiệu quả tiếp thị.

1. Yêu cầu 1 – Khảo sát nguồn lưu lượng

Với vai trò là một nhà phân tích dữ liệu bạn nhận được yêu cầu giúp Giám đốc kinh doanh biết được các session trên trang web bán hàng của công ty đến từ đâu tính đến trước ngày 12-04-2012.

Vị giám đốc này mong muốn sẽ xem được thông tin về `utm_source`, `utm_campaign`, `http_referer` và tổng số lượng các `website_session_id`. Hãy viết một truy vấn để tìm được các thông tin như vị giám đốc muốn xem. Sau khi viết truy vấn, hãy dựa vào kết quả trả về để đưa ra nhận xét

Gợi ý:

- Sử dụng bảng `website_sessions`
- Sử dụng Group by
- Kết quả trả về như hình dưới:

utm_source	utm_campaign	http_referer	number_session
gsearch	nonbrand	https://www.gsearch.com	3485
NULL	NULL	NULL	26
NULL	NULL	https://www.gsearch.com	25
gsearch	brand	https://www.gsearch.com	23
NULL	NULL	https://www.bsearch.com	7
bsearch	brand	https://www.bsearch.com	7

Một số thông tin mở rộng:

- `utm_source`: Đây sẽ là trường để lưu giữ nguồn thông tin session, ví dụ như các bạn chạy quảng cáo qua facebook, sau đó người dùng kích vào quảng cáo và được điều hướng đến website của bạn thì lúc này `utm_source=facebook`. Nếu các bạn thực hiện phân tích lưu lượng mà thấy trong database có trường là `utm_source` thì hãy hiểu ý nghĩa của trường này.
- `utm_campaign`: Người sở hữu website thường sẽ có nhiều chiến dịch quảng cáo, mỗi chiến dịch như vậy sẽ coi là một campaign.
- `http_referer`: Đây là tên miền đầy đủ của mỗi `utm_source`

2. Yêu cầu 2 - Tỷ lệ chuyển đổi từ nguồn lưu lượng truy cập

Nhìn vào kết quả của yêu cầu 1, bạn nhận thấy gsearch và nonbrand có nguồn lưu lượng truy cập lớn, nhưng bạn cần xác thực lại liệu những phiên từ gsearch và nonbrand có giúp thúc đẩy doanh số bán hàng hay không. Để tìm được điều này, bạn sẽ tính toán tỷ lệ chuyển đổi (calculate the conversion rate - CVR) từ các đơn đặt hàng (order_id). Yêu cầu đặt ra là tỷ lệ CVR phải lớn hơn 4%. Nếu bạn tính ra CVR thấp hơn 4% thì có nghĩa bạn phải giảm bid. Nếu bạn tính ra CVR cao hơn thì có nghĩa bạn có thể tăng bid.

Hãy viết câu truy vấn để tìm ra CVR trong các bản ghi có utm_source là gsearch và utm_campaign là nonbrand. Lưu ý là các thông tin cần được lấy trước ngày 12-04-2012. Sau khi viết truy vấn, hãy dựa vào kết quả trả về để đưa ra nhận xét

Gợi ý:

- Sử dụng join để kết hợp bảng webste_sessions và orders
- $$CVR = \frac{\text{tổng số order_id}}{\text{tổng số website_session_id}}$$
 (lưu ý là bạn sẽ chỉ tìm trong các hàng chứa gsearch và nonbrand)
- Kết quả trả về như hình dưới:

sessions	orders	session_to_order_conv_rt
3816	111	0.0291

Một số thông tin mở rộng:

- **Conversion rate** là khái niệm nhằm chỉ ra độ HIỆU QUẢ của website thông qua các hành vi mà khách hàng thực hiện.
 - o Ví dụ: website của bạn nhận được 1000 khách hàng truy cập trong một tháng và có 90 khách hàng đăng ký nhận tư vấn. Tỷ lệ chuyển đổi sẽ là 90/1000 hay 9%.
 - o Sự chuyển đổi có thể là bất kì hành động nào của khách hàng mà doanh nghiệp mong muốn họ thực hiện. Nó có thể là tỉ lệ người đăng ký nhận tư vấn/người truy cập website. Đó cũng có thể là tỉ lệ người bấm đăng ký dịch vụ/người truy cập web. Mỗi mục tiêu mà doanh nghiệp đặt ra sẽ có một conversion rate riêng.
- Theo dõi **tỷ lệ chuyển đổi** cho phép các bạn đo hiệu suất của website trong việc:

- o Nắm bắt được đang có bao nhiêu phần trăm người dùng đang hoàn thành các mục tiêu các bạn đặt ra.
- o Đánh giá sự thành công của một website và xác định các phần nên được cải thiện.
- **Bid:** Trong quá trình marketing, việc tối ưu hoá ngân sách tiếp thị là việc rất quan trọng. Các chiến dịch marketing sẽ có nhiều cách để tính phí. Ví dụ như khi chạy quảng cáo trên facebook, bạn sẽ phải đặt giá cho các chiến dịch quảng cáo. Như vậy giá ở đây sẽ được hiểu là bid. Giá chiến dịch càng cao thì đồng nghĩa với việc chiến dịch sẽ được tiếp cận đến nhiều khách hàng hơn (tránh nhầm lẫn giữa việc tiếp cận đến nhiều khách hàng hơn là sẽ chắc chắn tăng tỷ lệ chuyển đổi).

3. Yêu cầu 3 - Xu hướng nguồn lưu lượng truy cập

Dựa trên phân tích tỷ lệ chuyển đổi của bạn ở yêu cầu 2, công ty đã giảm giá những mặt hàng thuộc về nonbrand của gsearch vào ngày 15-04-2012. Công việc của bạn là viết truy vấn để tính số lượng phiên theo xu hướng (nonbrand và gsearch) theo tuần để xem liệu những thay đổi về giá có khiến số lượng phiên giảm chút nào không. Các số liệu sẽ được tính đến trước ngày 15-05-2012.

Gợi ý:

- Sử dụng bảng website_sessions
- Sử dụng group by theo week(created_at)
- Kết quả trả về như hình dưới:

week_started_at	sessions
2012-03-19	854
2012-03-25	949
2012-04-01	1147
2012-04-08	984
2012-04-15	645
2012-04-22	598
2012-04-29	684
2012-05-06	595

4. Yêu cầu 4 – Tối ưu hoá bid cho các lưu lượng có trả tiền

Bạn nhận được một phản hồi từ giám đốc tiếp thị về việc trải nghiệm không tốt khi sử dụng trang web trên mobile trước ngày 11-05-2012. Và vị giám đốc này muốn tính CVR trên các thiết bị. Nếu CVR trên Desktop tốt hơn trên mobile thì có thể tăng bid trên các session truy cập bằng Desktop.. Hãy viết câu truy vấn để thực hiện yêu cầu của vị giám đốc, sau đó đưa ra nhận xét về kết quả trả về.

Gợi ý:

- Sử dụng join để kết hợp bảng webste_sessions và orders
- Group by theo device_type
- Kết quả trả về như hình dưới:

device_type	orders	conv_rt
desktop	144	0.0372
mobile	24	0.0098

5. Yêu cầu 5 – Kiểm tra kết quả khi tăng bid

Sau khi phân tích tỷ lệ chuyển đổi theo device, các bạn sẽ nhận thấy Desktop cho CVR cao hơn. Vì vậy, công ty đã tăng bid cho các chiến dịch Desktop (nonbrand và gsearch) vào ngày 19-05-2012 Công việc của bạn là viết một truy vấn để theo dõi các xu hướng hàng tuần cho cả các sessions Desktop và mobile để công ty có thể thấy tác động đối với tổng số lượng session, sau đó đưa ra nhận xét đối với kết quả trả về.

(khoảng thời gian tìm kiếm sẽ bắt đầu từ 15-04-2012)

Gợi ý:

- Group by theo created_at
- Kết quả trả về như hình dưới:

week_start_date	desktop_sessions	mobile_sessions
2012-04-15	397	248
2012-04-22	364	234
2012-04-29	429	255
2012-05-06	414	276
2012-05-13	410	228
2012-05-20	656	185
2012-05-27	583	184
2012-06-03	553	152

Tổng kết: Trong bài lab này, các bạn đã đóng vai trò là một nhà phân tích dữ liệu để đi phân tích lưu lượng truy cập website. Đầu tiên các bạn sẽ khảo sát để tối ưu hoá bid, sau đó sẽ nhận định để xem việc điều chỉnh bid có đem lại hiệu quả hay không. Không chỉ về

cách viết truy vấn mà các bạn hãy hiểu về mindset phân tích trong bài lab (bắt đầu từ đâu, các bước thực hiện phân tích). Việc này sẽ giúp các bạn có một tư duy về phân tích dữ liệu, cụ thể trong trường hợp này là phân tích lưu lượng website