ĐẠI HỌC QUỐC GIA TPHCM – TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

KHOA CÔNG NGHỆ THÔNG TIN – BỘ MÔN HỆ THỐNG THÔNG TIN

A picture containing logo

Description automatically generated

**BÀI TẬP NHÓM**

**(HW#4)**

**Môn học: Hệ thống thông tin phục vụ trí tuệ kinh doanh**

*GVHD: Hồ Thị Hoàng Vy*

|  |  |  |
| --- | --- | --- |
| **Mã nhóm** | **MSSV** | **Họ và tên** |
| **TTKD-24** | 18120227 | Phạm Văn Minh Phương |
| 18120299 | Trương Công Quốc Cường |
| 18120035 | Đoàn Nguyễn Tấn Hưng |
| 18120534 | Hoàng Công Sơn |

Mục lục

[I. Giải thích ý nghĩa thuộc tính nguồn dữ liệu 4](#_Toc85060512)

# **Phân tích yêu cầu**

Công ty Amadeus Entertainment cần phân tích dữ liệu campaign delivery (chiến dịch tiếp thị sản phẩm thông qua tin nhắn email đến khách hàng). Cần biết được phản ứng của khách hàng tới email nhận được từ chiến dịch.

Sự kiện: Khi người dùng nhận được và đọc email offer từ chiến dịch tiếp thị

Bối cảnh sự kiện:

* Ai: Khách hàng
* Ở đâu: Kênh gửi - nhận email
* Cái gì: Trạng thái email tiếp thị, chiến dịch tiếp thị, cộng đồng triển khai chiến dịch
* Khi nào: Ngày email được gửi đến khách hàng

Đo lường (dữ kiện): Tiến trình của email tiếp thị

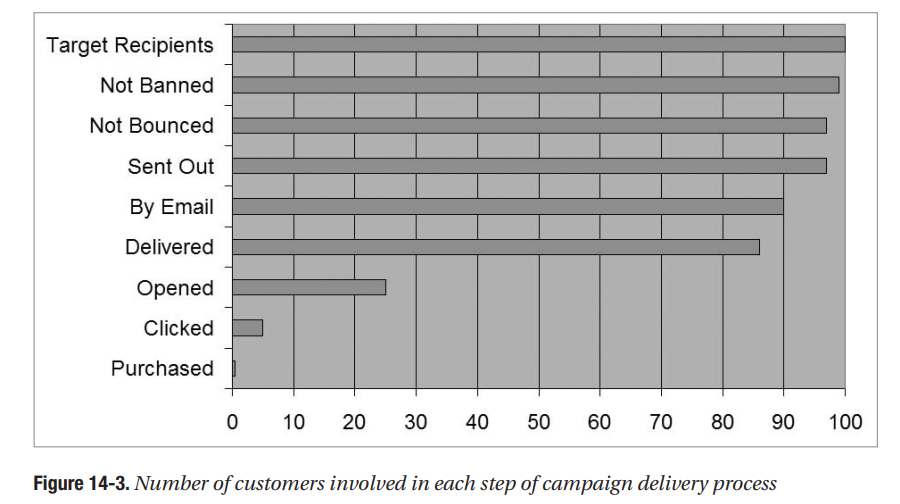
Fact table:

* Các giá trị có sẵn: Trạng thái mỗi tiến trình(đã gửi, đã giao, đã mở, ...)
* Các giá trị cần tính toán: trap hit
* Cấp chi tiết của dữ liệu: mỗi dòng trong bảng Delivery and Response Fact tương ứng với mỗi phản hồi từ người nhận dự kiến

Thiết kế chiều:

* Các chiều liên quan sự kiện phân tích: Chiều thời gian, chiều khách hàng, Chiều kên email, chiều chiến dịch, chiều trạng thái email, chiều cộng đồng

Giả sử khi chiến dịch được thực hiện, đã có 100.000 người nhận mục tiêu. Giả sử 1.000 trong số những người nhận mục tiêu này đã bị lọc ra do bị ban; 2.000 bị lọc ra vì địa chỉ e-mail của người nhận đã bị trả lại bốn lần. Vì vậy, 97.000 quảng cáo đã thực sự được gửi đi: 90.000 qua email; 4.000 qua đường bưu điện; 2.000 bởi RSS; và 1.000 bằng tin nhắn văn bản. Trong số 90.000 email, có 86.000 thư đến được người nhận dự định và 4.000 thư chưa được gửi (vì miền email không tồn tại, vì email không hợp lệ, vì hộp thư đã đầy, v.v.). Trong số 86.000 e-mail đã được gửi đi, 25.000 đã được mở / đọc và 5.000 được nhấp; nghĩa là, khách hàng đã truy cập vào các trang web được chỉ định. Cụ thể xem hình biểu diễn tỉ lệ dưới đây:



Số lượng tin nhắn e-mail đã được gửi đi, số lượng tin nhắn đã gửi, số lượng tin nhắn không được gửi (bị trả lại) và số lượng e-mail không được gửi đi (trong danh sách bị cấm hoặc danh sách bị trả lại) được gọi là dữ liệu phân phối chiến dịch và số lượng tin nhắn e-mail đã mở / đọc, số lượng tin nhắn được khách hàng nhấp vào và số lượng khách hàng đã mua một số sản phẩm được gọi dữ liệu phản hồi của chiến dịch. Cả dữ liệu phân phối chiến dịch và dữ liệu phản hồi của chiến dịch đều quan trọng vì chúng cho biết sự thành công hay thất bại của chiến dịch; ví dụ: nếu tỷ lệ mở và tỷ lệ nhấp (CTR) cao nghĩa là chiến dịch thành công.

* Cần thiết phải lưu thông tin về trạng thái của các email (đã gửi, đã nhận, bị trả lại, đã xem, được click, bị phàn nàn, bị đánh dấu spam…) và các thông tin liên quan tới email đó, đặc biệt là các thông tin phục vụ nghiên cứu thói quen mua sắm của khách hàng (ai mua cái gì ở đâu, khi nào, qua đâu, số tiền bỏ ra, hình thức…) và phân chia các thông tin đó theo từng danh mục khác nhau tương ứng để phục vụ việc thống kê, báo cáo, vẽ biểu đồ, hỗ trợ ra quyết định và đáp ứng các nhu cầu kinh doanh khác.

Chart, bar chart

Description automatically generated

Chart, bubble chart

Description automatically generated

# **Thiết kế**

Dimension Data Schema (DDS)

Diagram

Description automatically generated