**PHƯƠNG PHÁP LUẬN KHOA HỌC DỮ LIỆU**

**Assignment 1**

**SUPERMARKET BUSINESS GROWTH**

**Thách thức rõ ràng, mục tiêu chưa rõ ràng**

Với đề bài là “**Khách hàng của bạn là một siêu thị lớn và hội đồng quản trị đặt ra cho bạn một thách thức để giúp tăng trưởng việc kinh doanh của họ**”, ta có thể thấy mục tiêu chưa được xác định rõ ràng. Tăng trưởng kinh doanh có thể là tăng độ nhận diện thương hiệu hoặc tăng lượng khách hàng mới hoặc tăng trưởng doanh thu bán hàng nói chung. Do đó, việc đầu tiên cần phải xét đến là xác định mục tiêu của việc tăng trưởng kinh doanh là gì.

**Xác định mục tiêu (goal)**

Với những mục tiêu được nêu bên trên, yếu tố **khách hàng (người mua hàng)** luôn là thành phần không thể bỏ qua. Việc thấu hiểu khách hàng cũng như trở thành doanh nghiệp với mô hình khách hàng là trọng tâm sẽ giúp cho doanh nghiệp có được lòng tin (loyalty) của khách hàng hiện tại. Khi đã có được tín nhiệm từ khách hàng cũ, họ sẽ mang khách hàng mới đến với doanh nghiệp mà không cần sử dụng quá nhiều chi phí cho marketing bằng phương tiện truyền thông đại chúng do đã có phương tiện truyền thông đáng tin nhất là truyền miệng. Điều này sẽ dẫn đến tăng trưởng số lượng khách hàng cơ sở và theo đó sẽ là tăng trưởng việc kinh doanh.

Việc thấu hiểu khách hàng có thể đến từ nhiều hướng:

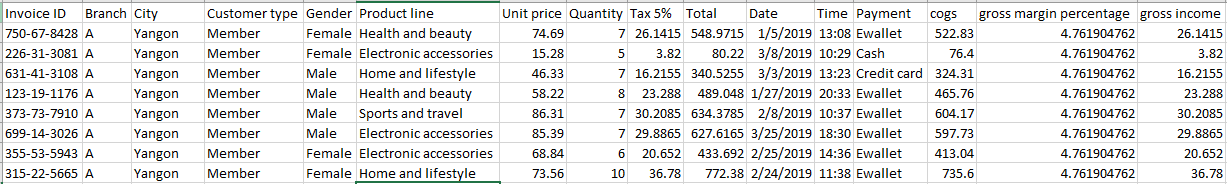
* Xác định nhu cầu (needs and wants) của khách hàng tiềm năng
* Phân nhóm khách hàng (customer segmentation)

Trong phần này chúng ta sẽ chỉ nói đến việc **phân nhóm khách hàng hiện tại** của siêu thị do đây là tập khách hàng đã có độ nhận diện với thương hiệu mà không cần mất thêm chi phí acquisition và thời gian thu thập thông tin cá nhân.

Vậy tập khách hàng hiện tại được định nghĩa như nào? Giả sử siêu thị lớn này đã có Member Card Program, khách hàng đăng ký làm khách hàng thân thiết của siêu thị và được nhận thẻ này. Tập khách hàng được coi là khách hàng hiện tại là những người sử dụng Member Card. Tập còn lại đã từng mua sắm tại siêu thị nhưng không sử dụng Member Card là khách hàng vãng lai.

Do không có dữ liệu sample nên trong bài này ta sẽ tạm thời lấy data từ Kaggle (<https://www.kaggle.com/aungpyaeap/supermarket-sales>) và giả sử đây là giao dịch của những khách hàng sử dụng Member Card.

Dữ liệu sẽ bao gồm giới tính, sản phẩm, giá sản phẩm, thời điểm mua hàng và cách thức thanh toán.



**Xác định phương pháp giải bài toán dữ liệu**

Với việc có được dữ liệu cùng với vấn đề được đặt ra - Phân nhóm khách hàng hiện tại, ta hiểu được kết quả cần có được là mô tả được chân dung hiện tại của nhóm khách hàng sử dụng Member Card. Chân dung sẽ dựa vào dữ liệu đang có và chỉ ra được nhóm khách hàng hiện tại có tỷ lệ nam/nữ như nào, họ quan tâm đến sản phẩm gì, thời điểm nào họ sẽ thực hiện việc mua sắm, …

Do đó, phương pháp giải bài toán này sẽ là thống kê mô tả (descriptive analytic).

Giả sử, mục tiêu được đề ra của siêu thị này là tăng doanh thu trong na

Tăng trưởng việc kinh doanh bao gồm những gì:

Được phân tách theo từng field:

Customer là trọng tâm

* Marketing: communication ra sao?
* Operation (sắp xếp, vận chuyển, thanh toán, utilities, …)
* Warehouse: Kho hàng, nhập xuất tồn ra sao
* Product: sản phẩm có phù hợp với nhu cầu của khu vực đó hay không, có đáp ứng thị hiếu không?
* Pricing: giá cả như thế nào so với mặt bằng chung của thị trường, khu vực lân cận?
* Location: có ở gần khu dân cư hay ở xa? Nếu gần thì oke, xa thì cần có biện pháp khác (ecom, shipping, …)
* Những vấn đề xung quanh liên quan gián tiếp đến trải nghiệm mua sắm (đỗ xe, úng nước, nắng nóng, nghỉ chân, vệ sinh, …)

Khách hàng hiện tại là ai?

Khách hàng mong muốn là ai?

Họ nghĩ gì về sau khi mua hàng tại siêu thị?

Những điều có thể có liên quan:

* Giảm chi phí cố định
* Giảm chi phí nhân công
* Tự phục vụ
* Tăng lợi nhuận biên ở từng sản phẩm
* Bán kèm bán chéo
* Sự gắn kết của khách hàng với siêu thị như nào? (thẻ hội viên, tích điểm, quà tặng, voucher, …)