

## **Bài toán phân tích giỏ hàng : Market-Basket Analysis**

Mục đích tìm các cặp sản phẩm hay được mua cùng nhau (VD. Mua bia kèm lạc, mua bia kèm bìm,...)

Dữ liệu các giỏ hàng được lấy từ link sau

<https://www.kaggle.com/irfanasrullah/groceries>

Yêu cầu: Duyệt và đưa ra các cặp sản phẩm xuất hiện cùng nhau nhiều nhất trong các giỏ hàng (mỗi dòng = 1 giỏ hàng mua sắm).

Cách làm

- + Đơn giản nhất duyệt list tất cả các cặp sản phẩm có thể và đếm tần số của mỗi cặp xem xuất hiện trong bao nhiêu đơn hàng ( giỏ hàng)
- + Cải tiến bằng thuật toán tối ưu hơn nếu có thể (trong 1 siêu thị, số lượng mặt hàng có thể lên tới  $n=10000 - 1000000$  sản phẩm, số lượng cặp có thể tới  $n^2$ )