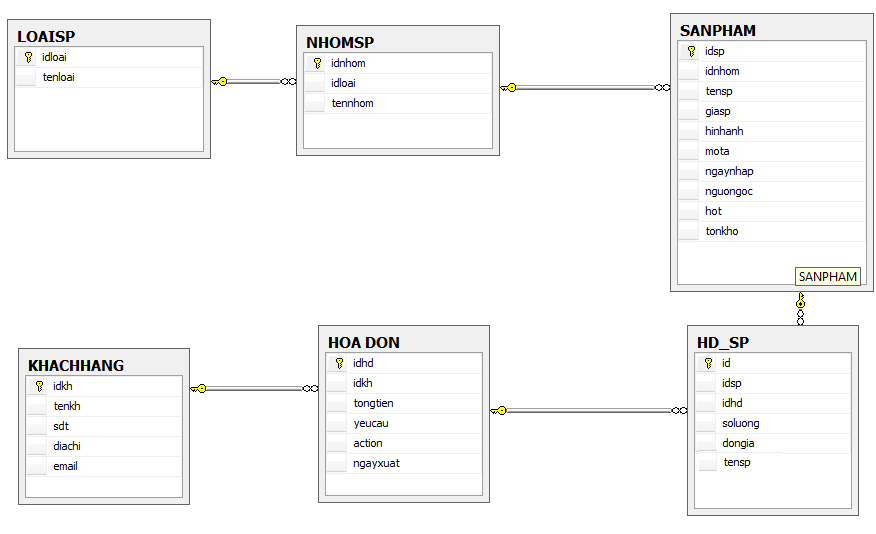
## Đây là database của anh ấy. Thuật toán kia cũng là của anh ấy hi



## Chức năng tìm kiếm sản phẩm mua kèm

Ý tưởng của thuật toán đã được nhắc ở phần giới thiệu, đó là quét các hóa đơn từ bảng hoadon-sanpham để xác định các sản phẩm nào hay đi kèm với nhau với đầu vào là sản phẩm đang được khách hàng chọn.

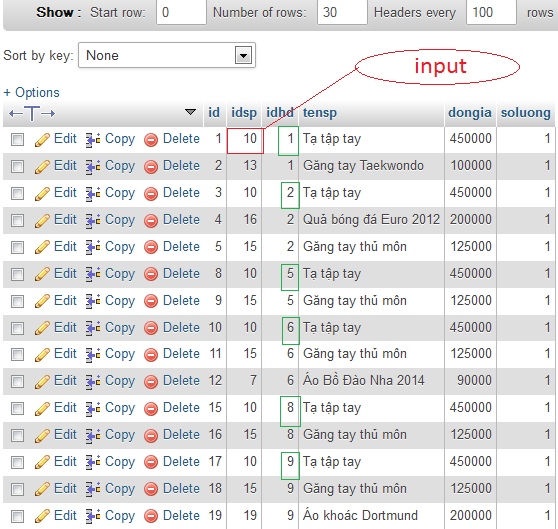
Các bước của thuật toán để tìm kiếm các sản phẩm đó được chia làm 3 phần chính là **đếm, sắp xếp** và **lấy ra.**

Trong đó bước **đếm** bao gồm đếm, phân loại là bước quan trọng nhất và ta sẽ phân tích kĩ phần này.

Bước **đếm** bao gồm các bước sau:

Input: Đầu vào của bước này là idsp (mã sản phẩm của mặt hàng khách hàng đang xem, kiểu int).

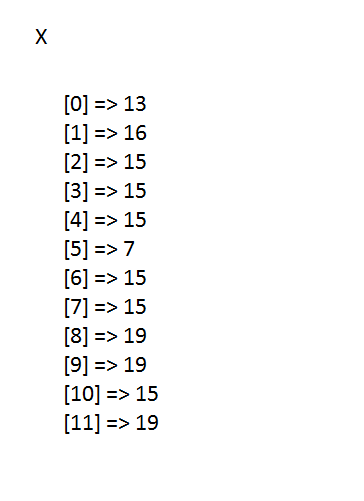
Sau khi có idsp ta quét ở bảng hoadon-sanpham để tìm các idhd (mã hóa đơn) mà sản phẩm đang xét thuộc vào nó.



Các ô khoanh màu xanh là các mã hóa đơn chứa sản phẩm đang xét.

Sau khi có idhd ta vẫn quét lần nữa bảng hoadon-sanpham để tìm ra các idsp khác với idsp đang xét nhưng vẫn nằm trong các hóa đơn đã quét ở trên.

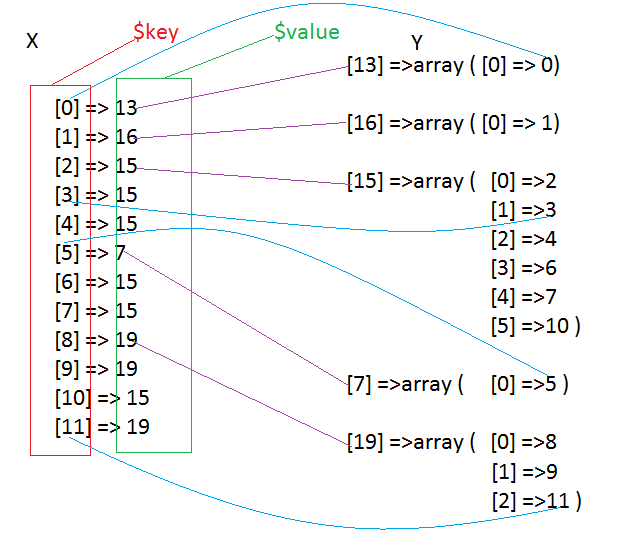
Các giá trị sau khi quét là các idsp lấy được có trong mỗi hóa đơn chứa sản phẩm đang xét, các giá trị này có thể trùng nhau, ta đưa tất cả chúng vào một mảng, mảng này kí hiệu là X, giả sử ta có:



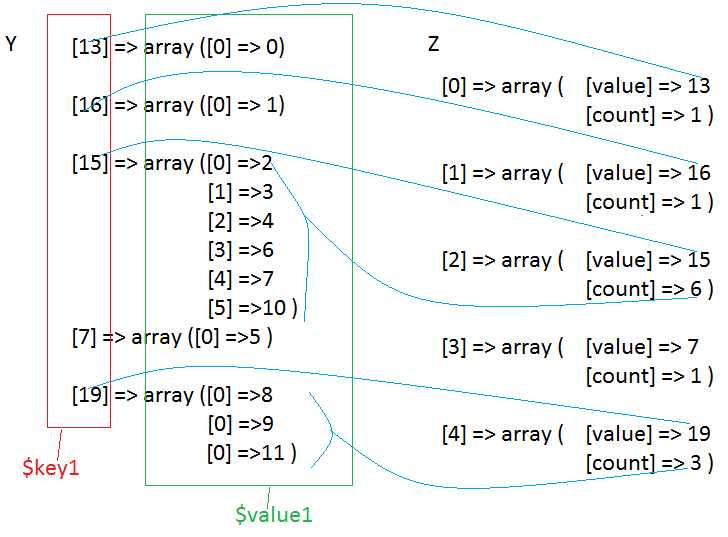
[0], [1]… là số thứ tự của mảng X, 13, 16…là giá trị tương ứng với số thứ tự của bản ghi, đó là các idsp lấy được từ trên.

Tiếp theo quét mảng X, gán số thứ tự bản ghi là $key, giá trị của bản ghi là $value.

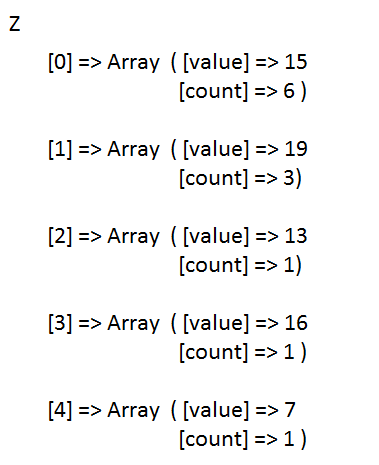
Ta xây dựng tiếp một mảng mới là Y (mảng 2 cấp), mảng này có số thứ tự bản ghi là $value, với mỗi giá trị của bản ghi $value ta đẩy chúng vào trong một mảng con với giá trị là $key, tiếp ta có:



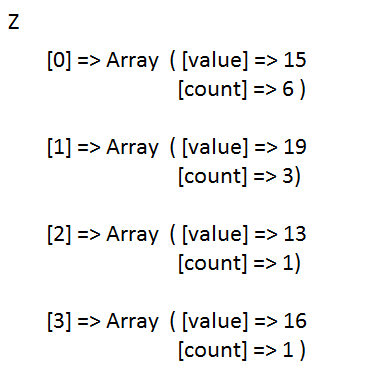
Tiếp nữa, quét mảng Y, gán số thự tự bản ghi là $key1, giá trị của mỗi bản ghi là $value2, ta có mảng Z là mảng 2 cấp, với mỗi bản ghi Z ta đẩy vào các giá trị là ’value’=$key1, ’count’=count($value1) (đếm số phần tử của ($value1). Hai trường này thuộc về một mảng con nằm trong Z.



Sau khi có mảng Z ta dùng thuật toán sắp xếp nổi bọt để sắp xếp thứ tự của mảng từ lớn đến nhỏ theo giá trị count. Đây là bước **sắp xếp** kết quả sẽ là:



Bước cuối là **lấy ra** sẽ thực hiện lấy số phần tử tùy chọn tùy vào yêu cầu thực tế, kết quả cuối có thể là một mảng nhỏ hơn mảng Z ban đầu.Kết quả sau khi lấy ra (giả sử lấy 4 bản ghi):



Nếu tồn tại các sản phẩm có count (số lần cùng trong một hóa đơn) thì thuật toán sẽ ưu tiên sản phẩm nào có trong hóa đơn trước đó.

Ví dụ: mã sản phẩm 11 và 15 có cùng số lần thuộc một hóa đơn với mã sản phẩm 10 nhưng khi quét bảng hoadon-sanpham sản phẩm 15 cùng với sản phẩm 10 trong hóa đơn 3 còn sản phẩm 12 cùng với sản phẩm 10 nhưng lại trong hóa đơn sau là 9 thì mã sản phẩm 15 sẽ được ưu tiên trước.

Các kĩ thuật có thể cải tiến chức năng có thể ở khâu **sắp xếp** vì khâu này sử dụng thuật toán sắp xếp nổi bọt, thuật toán này tuy đơn giản nhưng không phải là tối ưu trong sắp xếp, khi danh sách hóa đơn nhiều lên và số lượng đầu vào là lớn thì có thể ảnh hưởng đến quá trình hiển thị sản phẩm. Ngoài cải tiến thuật toán sắp xếp thì việc quét số hóa đơn giới hạn số lượng và được sắp xếp theo thời gian cũng là một cách giúp ta cải thiện chức năng vì theo thời gian thì những kết quả sau cùng mới là kết quả phản ánh thực tại nhất.

Thực hiện đầy đủ các bước trên là đã đầy đủ cách lấy các idsp thỏa mãn là những sản phẩm hay đi cùng với sản phẩm đang xét nhiều nhất, quét mảng Z lần cuối, gán giá trị của bản ghi là một biến (giả sử là $value2), gán một biến chỉ lấy trường value của mảng con Z (giả sử là $id, $value2 là mảng 2 chiều trong Z) ta được một dãy các idsp phù hợp với yêu cầu bài toán.