



# Market Basket analysis.



# Introduction

Market Basket Analysis is used to analyze retail basket or transaction data, which helps to identify strong association rules in a transaction dataset.

It is one of the key techniques used by large retailers to uncover associations between items. It works by looking for combinations of items that occur together frequently in transactions. To put it another way, it allows retailers to identify relationships between the items that people buy.



# Applications

1. Grouping products that co-occur in the design of a store's layout to increase the chance of cross-selling
2. Driving online recommendation engines ("customers who purchased this product also viewed this product")
3. Targeting marketing campaigns by sending out promotional coupons to customers for products related to items they recently purchased.



## Example

Assume there are 100 customers. 10 of them bought milk, 8 bought butter and 6 bought both of them.

bought milk  $\Rightarrow$  bought butter.

In the example milk is the antecedent and butter is the consequent. We are trying to identify the relationship between milk and butter.



# Support

$$\text{support} = P(\text{Milk \& Butter}) = 6/100 = 0.06$$

The percentage of transactions that contain all of the items in an itemset (e.g., pencil, paper and rubber). The higher the support the more frequently the itemset occurs. Rules with a high support are preferred since they are likely to be applicable to a large number of future transactions.

I.e 6 customers out of a hundred bought milk and butter.



# Confidence

$$\text{confidence} = \text{support}/P(\text{Butter}) = 0.06/0.08 = 0.75$$

The probability that a transaction that contains the items on the left hand side of the rule (in our example, milk) also contains the item on the right hand side (Butter). The higher the confidence, the greater the likelihood that the item on the right hand side will be purchased.



# Lift

Lift summarises the strength of association between the products on the left and right hand side of the rule; the larger the lift the greater the link between the two products.

$$\text{lift} = \text{confidence} / P(\text{Milk}) = 0.75 / 0.10 = 7.5$$