

Affinity Analysis One Pager

Market-based Analysis is a Data Mining technique that identifies customer buying patterns to plan store layout and place similar products on shelves/aisles. By putting related (bought simultaneously) products close together, tempts customers to buy further items (even off the Shopping List!) It may be used to organize and provide discounts on certain products and boost sales, such as 20% off honey with a Green Tea pack!

Support: It helps to determine how popular is any itemset/item, purchased during all transactions.

Confidence: is the measure of likelihood for certain items/itemset to be purchased together. It is the conditional probability that range is $[0,1]$. 1 indicates they are purchased together.

Lift: is measure of how likely itemset is purchased if antecedent is bought as compared to likelihood of itemset purchased if antecedent is not bought. Range is $[0, +\infty]$. Greater value indicates more likely to occur together than expected by chance.

Conviction: It measures the degree of dependence/strength between antecedent and the consequent. Range is $[0, +\infty]$. Infinity indicates rule is reliable and they are strongly dependent on each other.

For instance, 4 transactions recorded from a grocery store during early morning hours.

Transaction 1	Eggs, Bread, Milk
Transaction 2	Bread, Honey, Milk, Cereals
Transaction 3	Eggs, Bread, Butter, Cardamom, Cereals
Transaction 4	Sugar, Milk, Milkmaid, Cereals, Yogurt

Support (Eggs):

2 out 4 transactions had eggs in the market basket = $2/4$ ($2/4 * 100 = 50\%$)

Out of the all the transactions done at the grocery store, 50% were eggs. (Quite a famous item at the store!)

Confidence (Eggs => Bread)

No. of transactions with both eggs & bread = 2

No. of transaction with eggs = 2
 $2/2 = 1$ (100 %)

Certain that Bread will be purchased every time when eggs are brought.

Lift = Confidence (Eggs => Bread) / Support (Eggs) = $1 / (1/2) = 2$

Value 2 i.e., greater than 1. Suggesting that Buying bread with eggs is more likely to occur than expected by chance.

Conviction (Eggs => Bread)

$$(1 - \frac{2}{4}) / (1 - 1) = \infty$$

Infinity indicates strong dependence between both and that eggs will always occur with bread and vice-versa.

*Support and Confidence are non-symmetric whereas Lift and Conviction are symmetric.