

WHAT IS MARKET BASKET ANALYSIS (MBA)

- Also known as affinity analysis which helps identification of co-occurring relationships
- This study can be performed on items, instances, activities, etc.
- Typically this affinity analysis i.e. technique of identification of association is widely used on sales data where it examines various orders/baskets to conclude and hence also commonly known as 'Market Basket Analysis'
- This studies a. Frequent item set (items occurring together frequently), b. association (strength of relationship)
- Data requirement: Data elements are expected in categorical form, not suitable on numeric data

WHY MBA?

- Huge product portfolio
- Difficult to identify different purchase combinations & preferences manually
- Reliance on human intuitions is too big of a risk
- Business knowledge for emerging products and customer response is difficult to gauge

Computer aided intelligence can provide a much reliable, faster and comprehensive inputs

MBA APPLICATIONS



Store Layout

- Product placement, Catalogue optimization
- Inventory management
- Product Bundling
- Price Decisions (Discounts, loyalty programs)



Buying Behavior

- Cross Selling
- Up-selling
- Personalization (offers, prices)
- Recommendations



Online shopping

- Target marketing (emails, push notifications)
- Product recommendations
- Bundle discounts
- Offer design

KEY TERMINOLOGIES IN MBA

- **Items** : target objects (e.g. products in basket)
- **Transaction** : A unique identifier for group of item purchase
- **Frequent Item set** : Items occurring at high frequency.
- **Support** : % transactions containing item/s under study in overall transactions
- **Association Rule** : An association rule is a pattern that states when X occurs,Y occurs with a certain probability

$X \rightarrow Y$, If milk then Sugar, LHS \rightarrow RHS

- **Confidence** : likelihood/probability that item on LHS & RHS co-occur.This is a conditional probability that randomly selected transaction will include item X given item Y
- **Lift**

ILLUSTRATION

- In a store, there are 1000 transactions overall. Item A appears in 80 transactions and Item B occurs in 100 transactions. Items A and B appear in 20 transactions together.
- **Items** : A, B, C, ..
- **Transaction** : 1000 transactions
- **Frequent Item set** : A+B
- **Support** : Support of A = $\Pr(A) = 80/1000 = 8\%$ and Support of B = $\Pr(B) = 100/1000 = 10\%$.
- **Confidence** : Confidence of $B \Rightarrow A = \Pr(A/B) = 20/100 = 20\%$.
- **Lift** : Indicates how efficient in the rule is in finding consequences, compared to random selection of transaction

Can be expressed as the ratio of the probability of Items A and B occurring together to the multiple of the two individual probabilities for Item A and Item B.

$$\text{Lift} = \Pr(A, B) / \Pr(A) \cdot \Pr(B) = (20/1000) / ((80/1000) \times (100/1000)) = 2.5$$

- **Lift > 1** : Item/s on LHS has **increased** probability that transactions also contain item/s on RHS
- **Lift < 1** :**lowered**

APPROACH FOR MBA

- Support
 - It is recommended to select items/rules with high support
 - As these will be applicable to a larger number of transactions
- Confidence
 - Prioritize rules with higher confidence (typically >80%)
 - As this gives higher confidence in the recommendation
- Lift
 - Select rules with lift greater than 1
 - **Lift > 1** : Item/s on LHS has **increased** probability that transactions also contain item/s on RHS
 - **Lift < 1** :**lowered**