

Managing Sustainable Competitive Advantage II

Market Basket Analysis

Dr. Ashutosh Singh

World Ranked – Triple Accredited – Award Winning





Shopping basket




Shopping basket recommended

- Market Basket Analysis, or Affinity Analysis, uncovers meaningful correlations between different entities according to their co-occurrence in a data set.
- It generates association rules that satisfy predefined criteria to identify the most important relationships in the frequent item sets, and helps uncover the hidden patterns in a big data.


Benefits of Market Basket Analysis

- Up-sell and cross-selling




Leather Bag
Lorem ipsum dolor sit amet, consectetur adipi blandit erat, at bla.
\$136 ~~\$250~~
[+ ADD TO BASKET](#)


Others Also Bought




Leather Bag
\$130.00



Classic Sunglasses
\$230.00



Woman Shoe
~~\$209.00~~ ~~\$250.00~~



Bag
\$135.00

Benefits of Market Basket Analysis



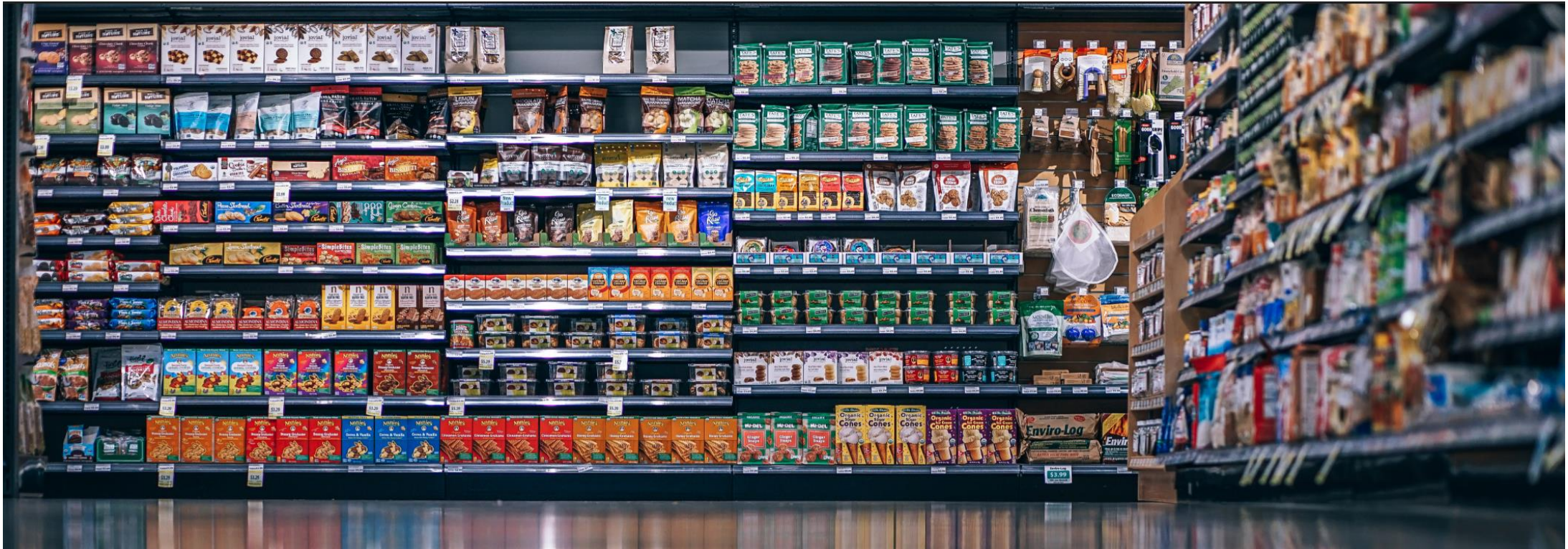
UNIVERSITY OF LEEDS

- Sales promotions, loyalty programs
 - When you understand product relationships and purchase sequences, you can identify and track customers who have bought a given product and deliver tailored messages to them.
 - With personalization, you're also able to create more effective marketing campaigns.
- Set Price, discount plans
 - Market basket analysis might tell a retailer that customers often purchase shampoo and conditioner together, so putting both items on promotion at the same time would not create a significant increase in revenue, while a promotion involving just one of the items would likely drive sales of the other.



Benefits of Market Basket Analysis

- Store design, shelf layouts, aisle displays



Benefits of Market Basket Analysis

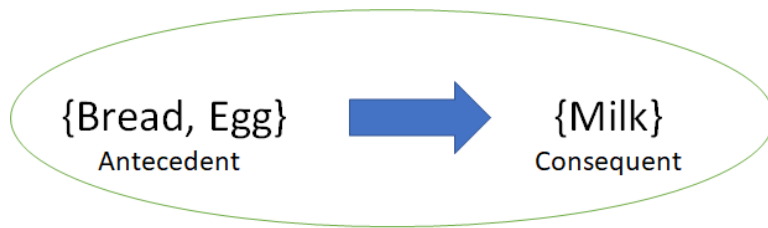
- Inventory management
 - Stocking the proper amount of the dependent product
- Refine marketing
 - Targeting segments based on their affinities

Key terms



UNIVERSITY OF LEEDS

ID	Items
1	{Bread, Milk}
2	{Bread, Diapers, Beer, Eggs}
3	{Milk, Diapers, Beer, Cola}
4	{Bread, Milk, Diapers, Beer}
5	{Bread, Milk, Diapers, Cola}
...	...



Itemset = {Bread, Egg, Milk}

An **itemset** is a representation of the list of all items which form the association rule

An **association** rule, {Bread, Eggs} => {Milk}, or more generally {X} => {Y}, indicates that if customers buy bread and eggs together, they are likely to also buy milk.

- **Support** is a measure of *absolute frequency*.
 - The support of 25% indicates that bread, eggs, and milk are purchased together in 25% of all transactions.
- **Confidence** is a measure of *correlative frequency*.
 - The confidence of 60% indicates that 60% of those who purchased bread and eggs also purchased milk.

- **Lift** is a measure of the strength of association between the products on the left and right hand side of the rule.
 - The probability of all of the items in a rule occurring together divided by the product of the probabilities of the items on the left and right hand side occurring as if there was no association between them.
 - For example, if bread, egg, and milk occurred together in 2.5% of all transactions, bread and egg in 10% of transactions and milk in 8% of transactions, then the lift would be: $0.025 / (0.1 * 0.08) = 3.125$.
 - The larger the lift the greater the link between the two products. A lift of more than 1 suggests that the presence of bread and eggs increases the probability that milk will also occur in the transaction.

Key terms: Summary



UNIVERSITY OF LEEDS

- Association Rule: $\{X \rightarrow Y\}$ is a representation of finding Y on the basket which has X on it
- Itemset: $\{X, Y\}$ is a representation of the list of all items which form the association rule
- Support: Fraction of transactions containing the itemset
- Confidence: Probability of occurrence of $\{Y\}$ given $\{X\}$ is present
- Lift: Ratio of *confidence* to baseline probability of occurrence of $\{Y\}$

Rule: $X \Rightarrow Y$

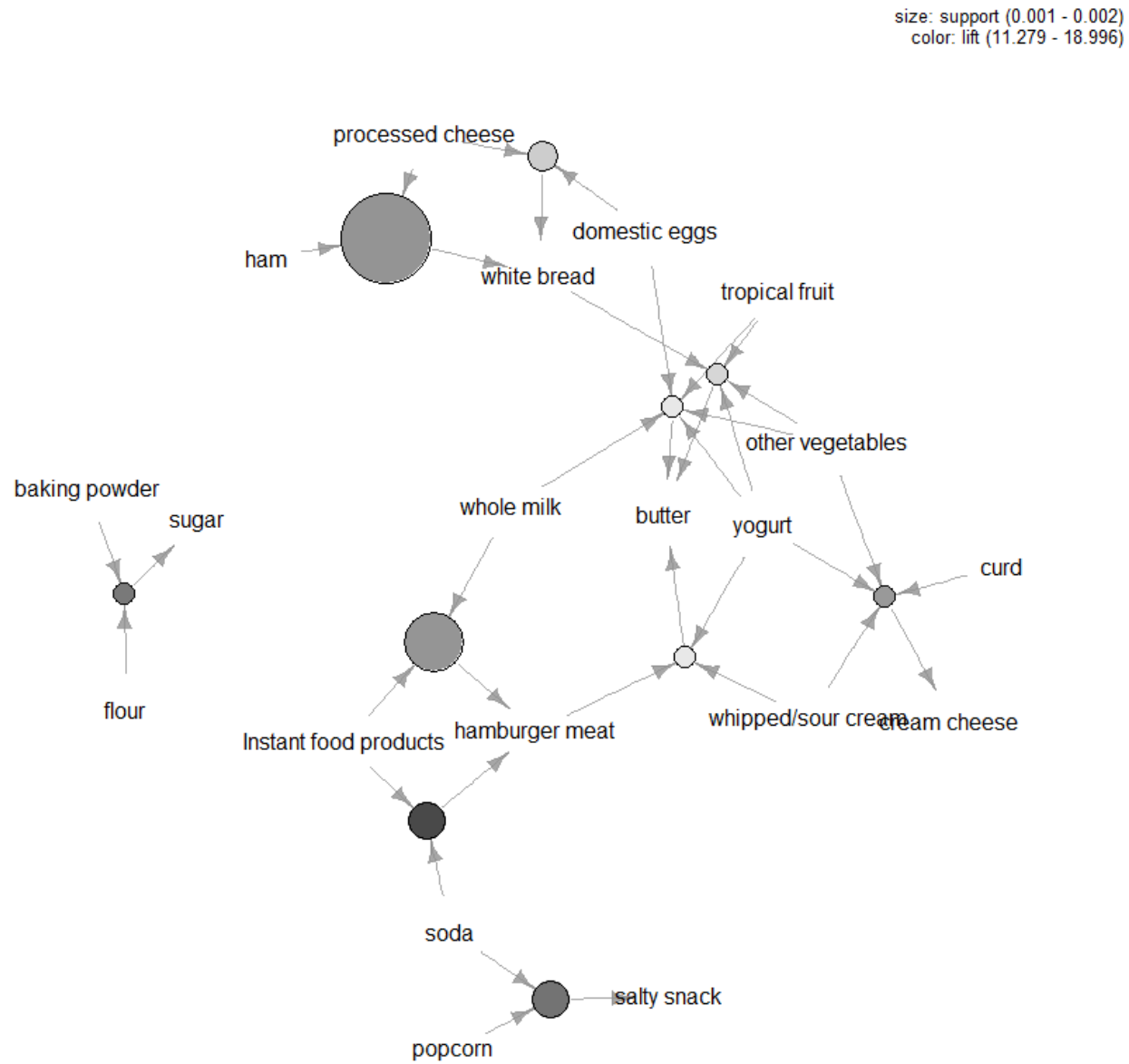
$$\text{Support}(\{X\} \rightarrow \{Y\}) = \frac{\text{Transactions containing both } X \text{ and } Y}{\text{Total number of transactions}}$$

$$\text{Confidence}(\{X\} \rightarrow \{Y\}) = \frac{\text{Transactions containing both } X \text{ and } Y}{\text{Transactions containing } X} = \frac{\text{support}(X \cup Y)}{\text{support}(X)}$$

$$\text{Lift}(\{X\} \rightarrow \{Y\}) = \frac{(\text{Transactions containing both } X \text{ and } Y) / (\text{Transactions containing } X)}{\text{Fraction of transactions containing } Y} = \frac{\text{support}(X \cup Y)}{\text{support}(X) \times \text{support}(Y)}$$

Rule	Support	Confidence	Lift
{instant food products, soda}=>{hamburger meat}	0.001	0.632	19.00
{soda, popcorn}=>{salty snacks}	0.001	0.632	16.70
{flour, baking powder}=>{sugar}	0.001	0.556	16.41
{ham, processed cheese}=>{white bread}	0.002	0.633	15.05
{whole milk, instant food products}=>{hamburger meat}	0.002	0.500	15.04

Results



Graph-based visualisation of the top ten rules in terms of lift