WHAT'S IN YOUR BASKET?

BASKET ANALYSIS

How can we make it happen using DAX? →

Shared by LOSC

What is Basket Analysis?

Suppose you're running an online store and you want to understand which products are your customers buying together.

Basket Analysis will help you in achieving this.

For example, you can achieve analysis such as 'Toothbrush' & 'Toothpaste' or 'Milk' & 'Cereal' are bought frequently.

3 Key Concepts

There are 3 key concepts to understand for Basket Analysis:

SUPPORT

CONFIDENCE

LIFT

SUPPORT

Support is the percentage of transactions that contain products as a product basket.

- Having the higher support means that product basket is occurring more frequently.
- Higher support also indicates that such product baskets are applicable for a higher future transaction.

$$\text{Support} = \frac{\text{No. of transactions of one or more products}}{\text{Total no. of transactions}}$$

CONFIDENCE

Confidence is the probability that the second product or group is in the basket despite any preconditions.

- The higher the confidence, the greater the probability that the second product will be present in the basket if the first product is already there.

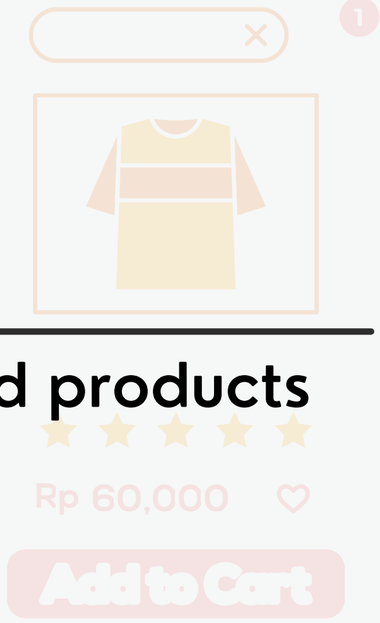
$$\text{Confidence} = \frac{\text{No. of transactions including both products}}{\text{No. of transactions including first product}}$$

LIFT

Lift demonstrates the solidity between two product baskets.

- A lift close to 1 implies that there is no important relationship between the products.
- Lift higher than 1 (positive relationship) implies that the buyers purchase the basket more repeatedly than it would occur by chance
- Lift lower than 1 (negative relationship) implies that the buyers purchase the basket less often than it would occur by chance

$$\text{Lift} = \frac{\text{Support of product basket}}{\text{Supporrt of first product} * \text{Support of second products}}$$



Let's assume you have a table of products purchased.

This table has 4 columns named Product 1, Product 2, Product 3, and Product 4.

In the Query Editor,

1. Add an Index Column.
2. Select the Index Column and rename it as TransactionID. Then right-click on the column and select Unpivot Other Columns.
3. You now have 3 columns: TransactionID, Attributes & Value. Remove the Attributes Column and Rename Value Column to Items.
4. Filter the Items Column and unselect the (blank) value. Now, close and apply.

Let's create a Basket Analysis Table in Power BI

```
Basket Analysis =
FILTER (
    CROSSJOIN (
        VALUES ( basket[Item] ),
        SELECTCOLUMNS (
            VALUES ( basket[Item] ),
            "Item2", basket[Item]
        )
    ),
    basket[Item] > [Item2]
)
```

Item2	Item
pastry	tidbits
salty snack	tidbits
sausage	tidbits
semi-finished bread	tidbits
soda	tidbits
pickled vegetables	tidbits
canned beer	tidbits
misc. beverages	tidbits
hygiene articles	tidbits
rolls/buns	tidbits
frankfurter	tidbits
curd	tidbits
beef	tidbits

The logic consists of two parts, 'FILTER' & 'CROSSJOIN'. FILTER is used to remove any duplicate rows occurring. CROSSJOIN is used to create one row for each couple of items i.e., it's generating a table with rows like 'tidbits' & 'pip fruit', 'tidbits' & 'other vegetables', etc.

Let's Create a Product Basket

```
Basket =
CONCATENATE (
    CONCATENATE ( 'Basket Analysis'[Item], " - " ),
    'Basket Analysis'[Item2]
)
```

Item2	Item	Basket
pastry	tidbits	tidbits - pastry
salty snack	tidbits	tidbits - salty snack
sausage	tidbits	tidbits - sausage
semi-finished bread	tidbits	tidbits - semi-finished bread

Concatenating both Item1 and Item2 to create a product basket.

Let's create a SUPPORT column

```
Support Basket =
VAR Item1 = 'Basket Analysis'[Item]
VAR Item2 = 'Basket Analysis'[Item2]
VAR TransactItem1 =
    SELECTCOLUMNS (
        FILTER ( basket, basket[Item] = Item1 ),
        "transactionID", basket[TransactionID]
    )
VAR TransactItem2 =
    SELECTCOLUMNS (
        FILTER ( basket, basket[Item] = Item2 ),
        "transactionID", basket[TransactionID]
    )
VAR TransactItem12 =
    INTERSECT ( TransactItem1, TransactItem2 )
RETURN
    COUNTROWS ( TransactItem12 ) / DISTINCTCOUNT ( basket[TransactionID] )
```

In the above logic, we've created 4 variables. Item1 & Item2 are referring both Item Columns. TransactItem1 will return a table with all the transactions which include the first item, TransactItem2 will return a table with all the transactions which include the second item, and TransactItem12 will return a table with all the transactions which include both items. We want the function to return the ratio between all the transactions which include both items and the total number of transactions.

Let's create a CONFIDENCE column

```
Confidence Item 1 -> Basket =
VAR Item1 = [Item]
VAR TransactFreq =
    DISTINCTCOUNT ( basket[TransactionID] )
VAR SupportItem1 =
    COUNTROWS ( FILTER ( basket, basket[Item] = item1 ) ) / TransactFreq
RETURN
    [Support basket] / SupportItem1
```

In the above logic, we've created 3 variables. Item1 refers to the first Item Column. TransactFreq will count the total number of transactions and SupportItem1 will return the support value for the first Item. We want the function to return the ratio between the Support Basket that we've calculated earlier and the Support of Item1.

The same logic will be used for the second Item.

Let's create a LIFT column

```
Lift =
VAR Item1 = [Item]
VAR Item2 = [Item2]
VAR TransactFreq =
    DISTINCTCOUNT ( basket[TransactionID] )
VAR SupportItem1 =
    COUNTROWS ( FILTER ( basket, basket[Item] = Item1 ) ) / TransactFreq
VAR SupportItem2 =
    COUNTROWS ( FILTER ( basket, basket[Item] = Item2 ) ) / TransactFreq
RETURN
    DIVIDE ( [Support basket], ( SupportItem1 * SupportItem2 ), 0 )
```

In the above logic, we've created 5 variables. Item1 & Item2 both are referring to the Item Columns. TransactFreq will count the total number of transactions, SupportItem1 will return the support value for the first Item, and SupportItem2 will return the support value for the second Item. We want the function to return the ratio of the Support of Basket (including all the items) that we've calculated earlier with the product of Support of Item1 and Support of Item2

Let's Visualize

Market Basket Analysis

Product	Support Basket	Confidence Item 1 -> Basket	Confidence Item 2 -> Basket	Lift
whole milk	0.18%	1.09%	12.56%	0.76
other vegetables	0.15%	1.23%	10.70%	0.85
soda	0.13%	1.34%	9.30%	0.93
rolls/buns	0.12%	1.06%	8.37%	0.74
root vegetables	0.10%	1.42%	6.98%	0.98
sausage	0.10%	1.66%	6.98%	1.15
yogurt	0.08%	0.91%	5.58%	0.63
tropical fruit	0.07%	1.08%	5.12%	0.75
pastry	0.07%	1.28%	4.65%	0.89
whipped/sour cream	0.07%	1.52%	4.65%	1.06
napkins	0.05%	2.10%	3.26%	1.45
pip fruit	0.05%	0.95%	3.26%	0.66
onions	0.04%	1.99%	2.79%	1.38
pork	0.04%	1.07%	2.79%	0.74
shopping bags	0.04%	0.83%	2.79%	0.58
herbs	0.03%	3.16%	2.33%	2.19
ice cream	0.03%	2.21%	2.33%	1.53
hygiene articles	0.03%	1.96%	1.86%	1.36
pot plants	0.03%	3.39%	1.86%	2.35
hamburger meat	0.02%	0.93%	1.40%	0.64
long life bakery product	0.02%	1.12%	1.40%	0.77
meat	0.02%	1.19%	1.40%	0.83
misc. beverages	0.02%	1.28%	1.40%	0.89
photo/film	0.02%	3.80%	1.40%	2.63
waffles	0.02%	1.08%	1.40%	0.75
hard cheese	0.01%	0.93%	0.93%	0.64
ketchup	0.01%	6.25%	0.93%	4.33
liver loaf	0.01%	4.26%	0.93%	2.95
newspapers	0.01%	0.34%	0.93%	0.23
pickled vegetables	0.01%	1.50%	0.93%	1.04
processed cheese	0.01%	1.34%	0.93%	0.93
semi-finished bread	0.01%	1.41%	0.93%	0.98
sliced cheese	0.01%	0.95%	0.93%	0.66
sugar	0.01%	0.77%	0.93%	0.53

Product Selection

- [ ] decalcifier
- [ ] dental care
- [ ] dessert
- [ ] detergent
- [ ] dish cleaner
- [ ] dishes
- [ ] dog food
- [ ] domestic eggs
- [ ] female sanitary products
- [ ] finished products
- [ ] fish
- [ ] flour
- [ ] flower (seeds)
- [ ] flower soil/fertilizer
- [ ] frankfurter
- [ ] frozen chicken
- [ ] frozen dessert
- [ ] frozen fish
- [ ] frozen fruits
- [ ] frozen meals
- [ ] frozen potato products
- [ ] frozen vegetables
- [ ] fruit/vegetable juice
- [x] grapes
- [ ] hair spray
- [ ] ham
- [ ] hamburger meat
- [ ] hard cheese
- [ ] herbs
- [ ] honey
- [ ] house keeping products
- [ ] hygiene articles
- [ ] ice cream
- [ ] instant coffee
- [ ] Instant food products
- [ ] jam