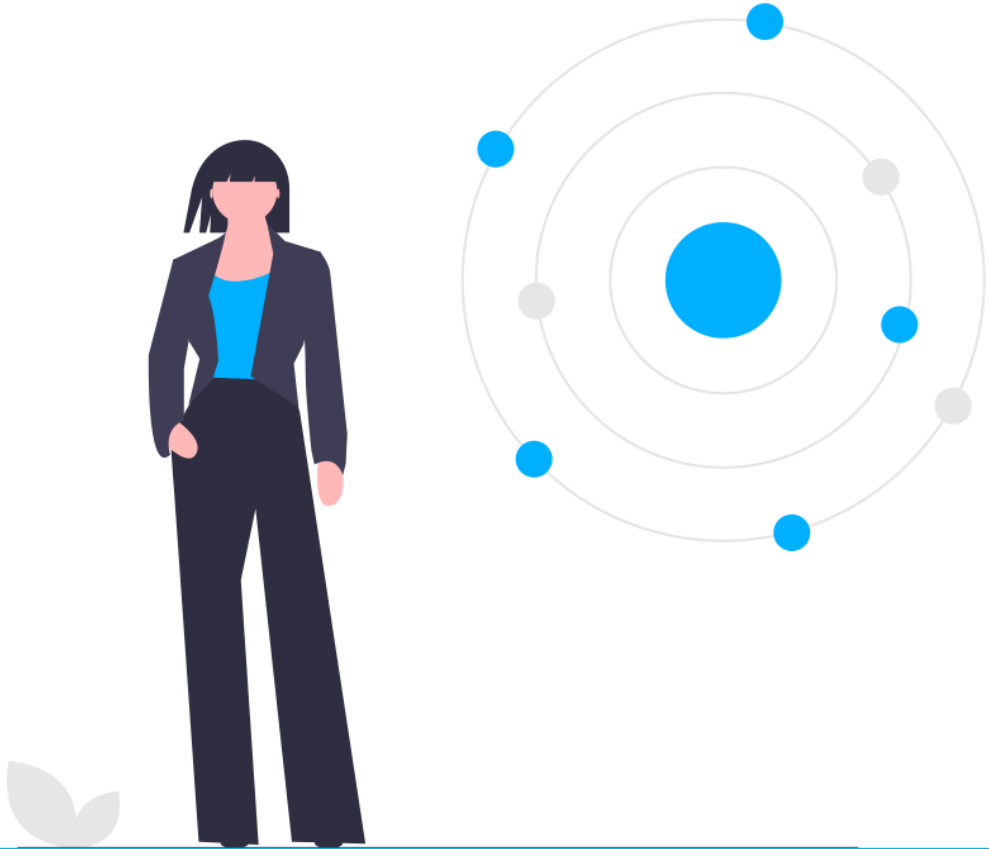


WHITE PAPER
SERIES

2025

Definitive Guide to Market Basket



Introduction

What products tend to be bought together? What products may benefit from promotion?

Market basket analysis is the tool in your cabinet that can help you answer both these questions at once. It works by analysing your sales data to discover association rules: certain pairs of products may be bought together more often than others because they complement each other, so it would make sense to recommend them as bundles to future customers.

“Bread and butter” is a common example of an association rule that market basket analysis may recommend to you and market basket analysis is the bread and butter of data science.

It’s results are easy to understand, but we need to go a bit deeper to fully appreciate this technique.

In the next few chapters we will explain the main idea behind association rules, look into different types of market basket analysis and off-the-shelf algorithms you can use.

We will also discuss the advantages and limitations you should anticipate on your way to effective market basket analysis.

David Foster
Partner
Applied Data Science Partners
(ADSP)



Focus Areas

This whitepaper series is a practical guide to developing your understanding of Market Basket Analysis.

It is organised around 8 key topics:



Chapter One:

What is Market Basket Analysis?

Chapter Two:

What Are The Types of Market Basket Analysis?

Chapter Three:

How does Market Basket Analysis Work?

Chapter Four:

What Different Algorithms Are Used in Market Basket Analysis?

Chapter Five:

Benefits and Advantages of Market Basket Analysis

Chapter Six:

What Are The Limitations of Market Basket Analysis

Chapter Seven:

Market Basket Analysis Examples

Chapter Eight:

Market Basket Analysis FAQs

1 What is Market Basket Analysis?

Market basket analysis is a technique used by retailers to increase sales by better understanding customer purchasing patterns. It involves analysing large datasets, such as purchase history, to reveal products that are likely to be purchased together.

It is a type of recommendation engine that recommends pairs of products bought together, so often that their association is statistically significant (we will see what this means in practice in Chapter Five).

While it has been around for decades, its applicability and strength have increased dramatically with the rapid increase of digital point-of-sale systems in retail stores.

2 What Are The Types of Market Basket Analysis?

There are three types of market basket analysis:

1. Descriptive Market Basket Analysis

Why? The sole objective here is to derive insights from past data related to associations between products. These insights do not offer any predictions for future customers.

How? Machine learning algorithms belonging to the family of unsupervised learning can be used for this type of analysis. These algorithms are very handy as they do not require labels: they can learn simply on data of past purchases.

2. Predictive Market Basket Analysis

Why? When a customer has made a series of purchases and is waiting at the checkout, you would like to recommend products that they are likely to add to their existing basket. Predictive Market Basket Analysis can help you achieve this.

How? To make predictions, you need to use classification or regression machine learning algorithms. These will build a predictive model, that will an existing basket to products that will most likely be purchased by the customer.

3. Differential Market Basket Analysis

Why? Differential market basket analysis is a great tool for competitive analysis: why do consumers prefer a particular platform over others even when the products and prices are the same? Decisions of consumers can be based on multiple factors, such as delivery time and user experience.

How? As most machine learning algorithms detect correlation rather than causation, identifying which factors differentiate customers can be challenging. For practitioners of this form of market basket study, understanding the data and collecting more of it before making a policy change is critical.


3 How does Market Basket Analysis Work?

Let us assume that a store sells a set of four products $P = \{p_1, p_2, p_3, p_4\}$ and that it has a database that indicates which items were bought by a customer that arrived at time t_1 , a customer at time t_2 and a customer at time t_3 .

The database may look like this:

	p_1	p_2	p_3	p_4
t_1	0	1	1	1
t_2	1	0	0	1
t_3	0	1	1	0

Bought together



In market basket analysis, association rules are used to predict the likelihood of products being purchased together. Association rules count the frequency of items that occur together, seeking to find associations that occur far more often than expected.

Associations have the form $A \rightarrow B$: a customer that bought product A is likely to also buy product B. A is called the antecedent and B the consequent. Imagine we want to decide whether to recommend the association: $p_2 \rightarrow p_4$. Then we need to calculate the following metrics:

Support: This metric indicates how frequently two items are bought together.

Confidence: The confidence of a rule indicates the probability of both the antecedent and the consequent appearing in the same transaction.

In our example: confidence = $\text{Support}/P(p_4) = 0.33/0.77 = 0.42$ To decide whether to select an association we have to set a minimum requirement for support = s and a minimum requirement for confidence = c and examine all associations to see if they satisfy these requirements.

If s and c are too low, market basket analysis will return too many recommendations. If they are too high, then we may miss some useful recommendations. We therefore need to carefully choose these values for our dataset. This is where the different algorithms for market basket analysis come into play: by adopting different statistical techniques they differ in sensitivity and offer a wide breadth of alternatives for picking what works best for you.

4 What Algorithms Are Used in Market Basket Analysis?

Apriori, SETM and FP Growth are examples of algorithms that use association rules:

Apriori Algorithm: The Apriori Algorithm is a popular market basket analysis algorithm that finds association rules following the process we described in Chapter Three. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. Its drawback is its sensitivity: it may generate too many recommendations. Because of the large database, it must scan the database multiple times, which adds time and reduces performance.

SETM Algorithm: This algorithm generates candidate item sets on the spot. Its disadvantage is that it results in unnecessarily generating and counting too many candidate item sets.

FP Growth: The frequent pattern-growth (FP Growth), is based on the idea of representing data as a tree and is an advancement to the Apriori Algorithm.

5 Benefits and Advantages of Market Basket Analysis

The benefits of market basket analysis go both ways: retailers can increase sales, while consumers gain a more productive and valuable customer experience that generates brand loyalty.

Here are some specific effects you should expect from adopting market basket analysis:

- **Customer engagement is increased.** If the recommendations are on target, customers do not need to expend additional effort.
- **Sales are boosted.** Additional sales are achieved by cross-selling and bundling.
- **Customer experience is improved.** This can in turn increase brand loyalty.
- **Marketing strategies are optimised.** Descriptive market basket analysis can lend numerous insights that may prove valuable across marketing strategies beyond just cross-selling and bundling.
- **Customers feel better understood.** This creates a positive feedback loop where effective marketing strategies and engaged customers benefit each other.

The efficiency of supply chain and demand forecasting is increased. This is just one example of how insights generated by market basket analysis data can boost other business functions.

6 What Are The Limitations of Market Basket Analysis

As any marketing strategy deriving power from data, market basket analysis entails limitations that one needs to take into account before designing their own solution:

Spurious correlations: When we derive association rules from data, we are basically detecting correlations between products. A common mistake is to interpret these correlations causally. Datasets often hide confounding factors: two items may appear to be bought together not because they complement each other, but because there is a hidden factor that is causing them both.

No clear calls to action: Even if there's a true link between the two products, it takes time and skill to figure out how to act on this information. For example, let's say you learn that 50% of customers who buy bread, buy rice within two weeks. You could use your market basket analysis tool to have a recommendation sent to each online shopper who bought bread within two weeks of their purchase. But when do you send it to ensure you capitalise on the correlation? The tool can't tell you.

Time and cost of testing: Because there are no clear calls to action, retailers who use these solutions must dedicate time to A/B test any actions they take as a result of the data. But how do you decide which correlation to A/B test? Testing even just the strongest correlations takes enormous time and effort. Let's say you manage to pick a product pairing that you think will generate the most revenue. There's a lot of work to make the test happen.

For instance, if you're going to run an in-store cross-promotion, you'll need to re-organise your shelves, rework your planograms, and send directives down to all stores. From there, you'll need to train staff on the new locations and make them aware of the promotion. You'll then need an adequate amount of time (weeks? months?) to test whether or not you were right. Costs aside, from learning the correlations to making changes and then testing their efficacy and using what you learned moving forward, this entire process is slow.

7 Market Basket Analysis Examples

Market basket analysis is a versatile tool that can be useful in many industries:

Retail: Amazon.com is maybe the most well-known Market Basket Analysis case study. When you go to Amazon to look at an item, the product description will immediately suggest "Items purchased together regularly." It's the simplest and most direct illustration of Market Basket Analysis cross-selling strategies. In addition to e-commerce, physical retail stores can also benefit from market basket analysis through visual merchandising and shelf optimisation. At the supermarket, for example, you'll often see lactose-free products grouped together.

Telecom: With the telecom industry's fierce rivalry, it is important to offer competitive packages that cover users' diverse needs. Here, it is common to combine TV, phone and internet bundles to decrease migration.

Banking: Having access to clients' transactions made by credit and debit cards, banks can build an in-depth understanding of one's consuming behaviour. This type of knowledge can be useful in many ways. For example, banks can optimise their loyalty programs.

Healthcare: These last years have taught us that, when it comes to dealing with healthcare emergencies, efficiency and proactiveness are essential. Knowing that certain diseases occur together can help people take precautionary measures on time and can help healthcare practitioners anticipate a need for medications and facilities.

8 Market Basket Analysis FAQs

Brief and succinct answers to these frequently asked market basket analysis questions:

What is Market Basket Analysis?

Market basket analysis is a technique used by retailers to increase sales by better understanding customer purchasing patterns. It involves analysing large datasets, such as purchase history, to reveal products that are likely to be purchased together. Its output are association rules of the form $A \rightarrow B$, where A is called the antecedent and B the consequent.

What are the metric used in Market Basket Analysis?

There are two important metrics for characterising the association rule between two products. Support indicates the percentage of the total transactions where the two products were bought together. Confidence is the probability that the consequent is bought considering that the antecedent has been bought.

Who uses Market Basket Analysis?

Market basket analysis is a diverse tool that can be useful in many industries, such as retail, telecom, banking, and healthcare.

What algorithms are used in Market Basket Analysis?

Some popular algorithms are the Apriori Algorithm, SETM Algorithm, and FP Growth.

What is an example of Market Basket Analysis?

Amazon.com is maybe the most well-known Market Basket Analysis case study. When you go to Amazon to look at an item, the product description will immediately suggest "Items purchased together regularly." It's the simplest and most direct illustration of Market Basket Analysis cross-selling strategies.

What are the benefits of Market Basket Analysis?

Benefits of market basket analysis can be found both on the retailer's side (sales are boosted, marketing strategies are optimised, efficiency of supply chain and demand forecasting is increased) and the consumer's side (customers get an improved shopping experience, are more engaged and feel better understood).

What are the limitations of Market Basket Analysis?

The limitations that one needs to anticipate when employing this technique are: spurious correlations in the data that may lead to wrong insights, a lack of a clear path from insight to actions, and testing periods for new policies that require time and incur costs.

9 Conclusion

Market basket analysis is a simple yet effective tool for increasing sales and improving customer experience.

Its main objective is to discover products that are often bought together in order to advise future recommendations.

A firm grasp of limitations of statistical analysis, such as spurious correlations, and a clear understanding of the cost and time it takes to implement policy changes advised by market basket analysis, are necessary to derive value from this technique.

For most industries, the question that decision-makers need to answer is not whether they need market basket analysis, but how they can apply it most effectively to outcompete the competition.