

# The Qlik Associative Difference®

Your Unique Advantage for Uncovering Hidden Insights



## INTRODUCTION

---

Qlik's one-of-a-kind associative technology sets it apart from all other BI solutions on the market. This white paper provides an overview of that technology and the value that only Qlik® can deliver. It also outlines the differences between the Qlik Associative Engine and query-based BI tools, and it explains how Qlik enables users of all skill levels to improve the effectiveness of their analysis and achieve a deeper, more complete understanding of their data.

## The Associative Difference®

Almost all of the analytics tools on the market today share the same essential architecture. They're built on relational databases and use SQL and query-based approaches to support analysis. That's a major flaw. Because while SQL is required to pull data from many sources, it was never designed to support interactive analytics, and it places strict limits on how data can be accessed and explored.

In query-based BI tools, data sources have to be brought together using SQL joins, and assumptions must be made in advance about what types of questions users will have. All other data is left behind. If a user discovers something interesting and wants to pivot their thinking to an adjacent question or data set, they'll almost certainly have to re-build complex queries – and that means going back to the data experts. We call this the “ask, wait, answer cycle,” where every new type of question has a waiting period.

*“Moving to Qlik has been like night and day. We have all the data we need right at our fingertips – and we can just play. When we see something interesting, we can explore [further].”*

- VP, Information Systems

Qlik offers an entirely different technology for exploring your data: The Qlik Associative Engine. Designed specifically for interactive, free-form exploration and analysis, the Associative Engine fully combines large numbers of data sources and indexes them to understand the associations – without leaving any data behind. It offers powerful on-the-fly calculation and aggregation that instantly updates analytics and highlights relationships in the data, exposing both associated and unrelated values after each click.

This means that any user – at any skill level – can search and explore their data in any direction, following their curiosity wherever it leads. It's like having peripheral vision, removing blind spots and uncovering hidden insights that aren't available in query-based tools. That's why Qlik users consistently drive tremendous value by discovering completely new insights. We call it The Associative Difference – and only Qlik can deliver it.

## Query-Based Tools – Limitations and Blind Spots

Qlik's Associative Engine offers freeform exploration and search across all your data – capabilities that query-based tools simply can't match. Often, visualization tools that rely on queries for analysis attempt to offer an agile experience, but they quickly hit limits in terms of flexibility and performance. They're not designed to uncover the unexpected.

Query-based tools use a structured, linear approach to visualize partial subsets of data, ultimately creating blind spots in understanding. And the only way for these tools to answer unanticipated questions is to re-build complex SQL queries, which usually requires an experienced data expert.

### Qlik at a Glance

Founded in Lund, Sweden in 1993

Approximately 48,000 customers in more than 100 countries

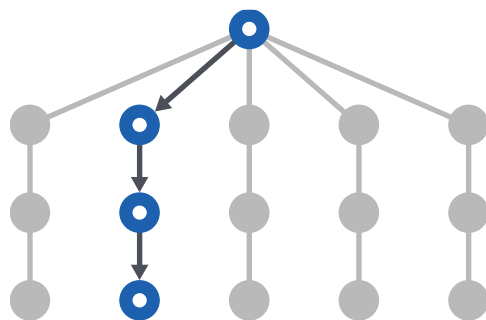
Solutions driving leadership in the visual analytics market

- **Qlik Sense®**  
Self-service data analytics
- **Qlik Analytics Platform®**  
Custom and embedded analytics
- **QlikView®**  
Guided analytics and dashboards
- **Qlik NPrinting®**  
Centralized reporting
- **Qlik GeoAnalytics®**  
Mapping and geographic analysis
- **Qlik Sense Cloud®**  
Visual analytics online
- **Qlik DataMarket®**  
Third party data-as-a-service
- **Qlik Connectors®**  
Broad data connectivity options

More than 1,700 technology, solution, OEM, consulting and system integrator partners

More than 2,000 employees worldwide  
Recognized leader by industry analysts including Gartner, BARC, Forrester, Butler Analytics, and Ventana Research

Headquartered in Pennsylvania, USA



VS



- x Partial subsets of data
- x Restricted linear exploration
- x Slow performance
- x “Ask, wait, answer cycle”

- ✓ All your data
- ✓ Explore without boundaries
- ✓ Speed of thought
- ✓ **Unexpected insights**

## Where Query-Based Tools Fall Short

### Partial views

Query-based tools either connect directly to underlying sources or use some type of data staging area or data warehouse. These databases must be fully modeled in advance before any analysis can be done, requiring significant effort and time. SQL joins are required for combining sources, risking data loss or inaccuracy – especially when a large number of sources are involved. And to support user analysis, queries are run – limiting flexibility and confining users to partial subsets of data (query result sets).

### Context

Query-based visualizations are discrete, disconnected entities that don't stay in context with each other. This means that filtering a single visualization won't inherently show the relationship or impact that selection has on the other objects within an app or dashboard. Some tools try to create a unified context by allowing objects to be “wired” together using front-end code, but that results in a lot of queries running at the same time, leading to major performance issues. The approach simply doesn't scale.

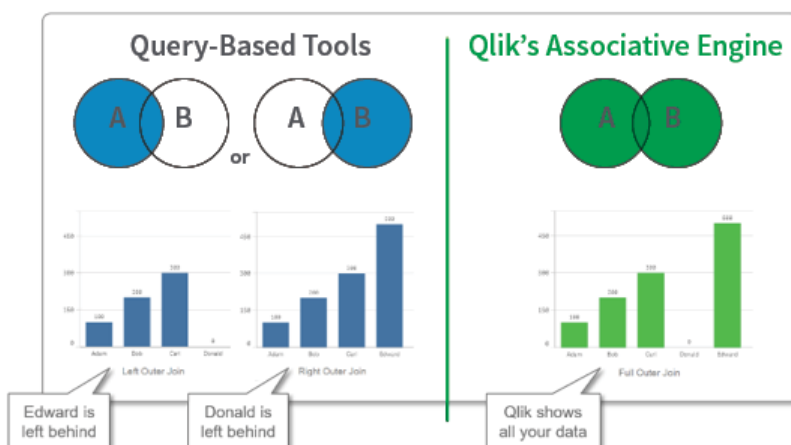
## Limited flexibility

Typically, only power users can build visualizations and queries, which limits everyone else to the preconceived questions the developers coded into the tool. There's no way for business users go beyond simple filtering within an object to search, explore, or ask new questions. The only option is to wait for a data expert to build a new query or visualization.

*Global tech giant Cisco brought together 500 million records of complex customer data from multiple systems and allowed their large sales force to freely explore customer portfolios. Users uncovered the best cross-sell opportunities for maintenance and additional products, helping the company generate \$100 million in support renewals and \$4 million in cost savings – and greatly improving customer satisfaction.*

## Data Loss

If information is loaded from multiple sources, SQL joins executed at load time can result in lost data. This approach typically requires a primary data source to be defined. That means users can only bring in subsets of data from secondary sources that match the values in the primary source. Anything else is lost, and business users might not even know.



## Incorrect results

Query-based tools also run the risk of inaccurate calculations. Values can be double- or even triple-counted if queries and joins aren't defined correctly and data is denormalized. Getting things right typically requires users to have a strong familiarity with the data model and the skills to properly structure queries.

## Poor performance

Query-based tools are at the mercy of the databases that support them. And no matter how robust the database, queries take time to execute. The problem is further compounded when wired dashboards execute multiple queries at the same time. As more and more concurrent users try to execute more and more queries, underlying databases simply can't keep up. A slow and unresponsive system disrupts a person's train of thought, and they lose their ability to connect ideas and take next steps.

## An Analogy

Let's say you want to understand how an internal combustion engine works. With a query-based tool, you would see some of the individual parts of the engine in isolation. You'd be able to evaluate one part at a time, but you'd have to guess at the relationships (or associations) between the parts, and how they fit together as a cohesive whole to create a working engine.



On the other hand, with Qlik's associative technology, you have the ability to analyze a complete working engine, to see each part and its relationship to all of the other parts. You can tweak the throttle (make a selection) and see how this affects the fuel intake, carburetor, and exhaust. You can watch the pistons pump and turn the crankshaft. You can deconstruct the engine at your leisure and look at each part in the context of the parts next to it. This is the power of Qlik's Associative Engine.

With Qlik, when users look at information, they know precisely how it's related. If they want to narrow the context down to a particular set of selections, they can immediately see how the rest of the data across an entire application responds. The Qlik Associative Difference means users aren't limited to seeing only a subset of the data contained in a query result set – and aren't restricted in how they explore it.

# Qlik's Associative Technology: How It Works

With Qlik, users of any skill level can explore and refine context through simple searches and selections, starting anywhere and going anywhere. Every time a user clicks, the Associative Engine instantly responds, dynamically recalculating all analyses based on the new context and highlighting associations in the data across all data sources. This “speed-of-thought” feedback encourages users to think of new questions and continue their individual paths of exploration and discovery.

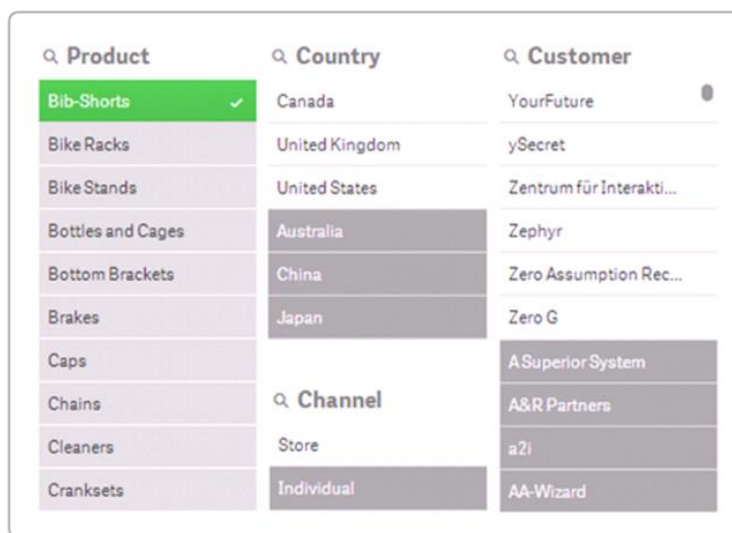
## Understanding Associations

An association is simply a relationship between one data value and another. For example, a product may have been sold in a certain country but not in others. This product would be associated with the country it was sold in and unrelated to the others. With Qlik's Associative Technology, if you select the product, it will appear in green; the associated country will appear in white; and the unrelated countries will appear in gray.

Associations can be:

- ✓ Positive:  
a value is related to another
- ✓ Negative:  
a value is unrelated to another

It's important to note that the unrelated (gray) values provide as much insight as the positive (white) ones – and can often indicate new opportunities or areas of risk.



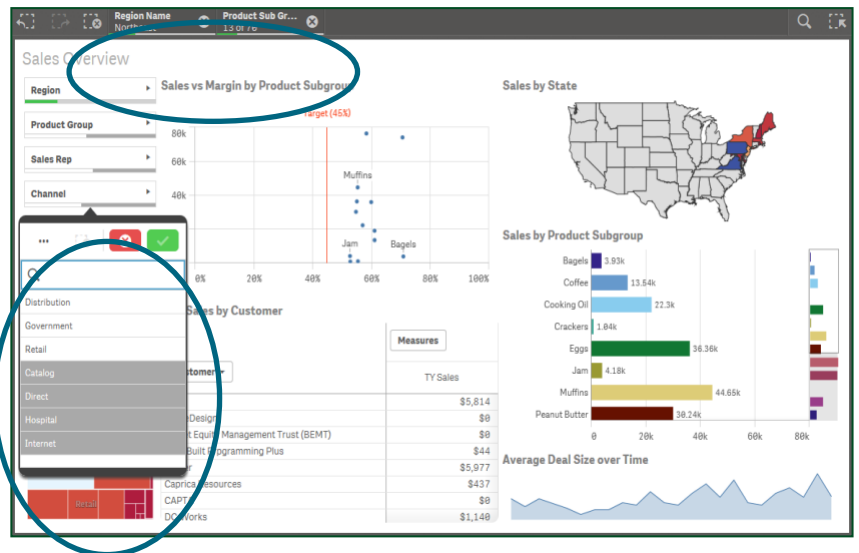
The screenshot shows three data tables: Product, Country, and Customer. The Product table has a search bar and a list of items. 'Bib-Shorts' is selected and highlighted in green. The Country table shows 'Canada' in white, indicating it is associated with the selected product, while other countries like 'United Kingdom', 'United States', 'Australia', 'China', and 'Japan' are in gray, indicating they are not associated. The Customer table shows various customer names, all of which are in gray, indicating they are not associated with the selected product.

Product	Country	Customer
Bib-Shorts ✓	Canada	YourFuture
Bike Racks	United Kingdom	ySecret
Bike Stands	United States	Zentrum für Interakti...
Bottles and Cages	Australia	Zephyr
Bottom Brackets	China	Zero Assumption Rec...
Brakes	Japan	Zero G
Caps		ASuperiorSystem
Chains	Channel	A&R Partners
Cleaners	Store	a2i
Cranksets	Individual	AA-Wizard

*Selections appear in green, associated values in white, and unrelated values in gray*

All data values have associations with others in the data set. For example, a product could be associated to the customers who bought it, the channels it was sold through, and individual transactions it was included in. Associations can work in any direction, extend across multiple relationships, and apply to combinations of values — such as regions associated with a set of selected products and countries.

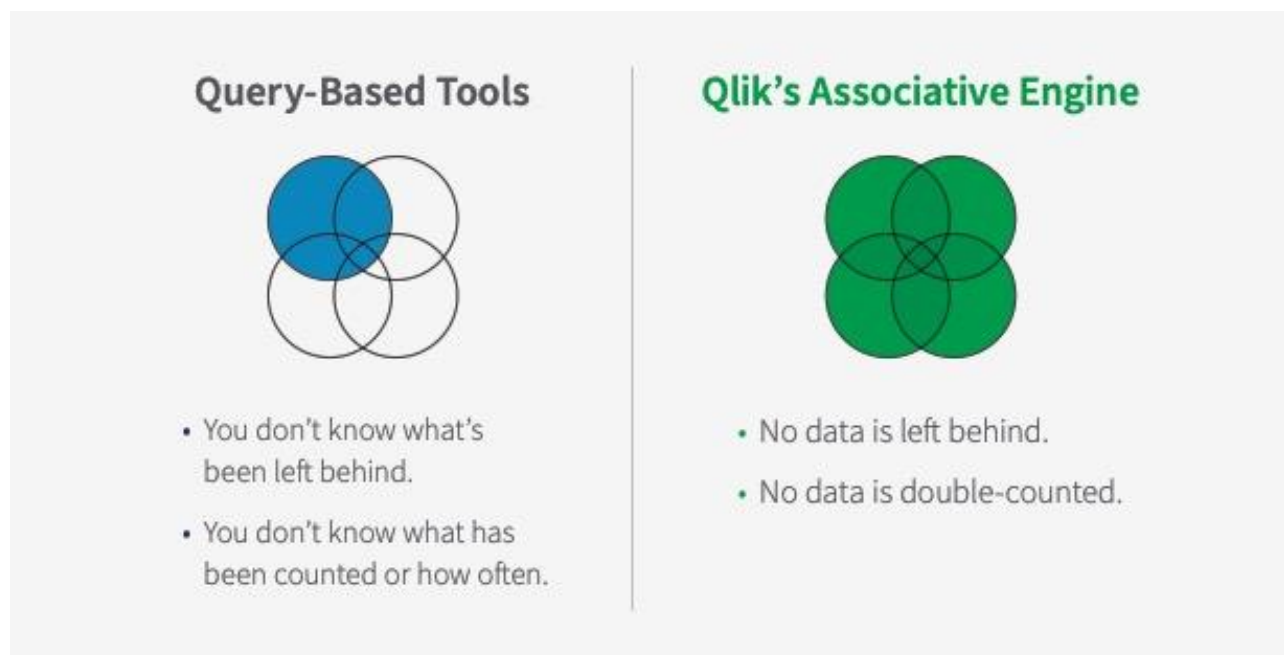
Users won't always know which associations exist – or which ones are important. A query-based tool could easily report that a product was sold in a region. But noticing that this product wasn't sold through certain channels in the geography may not be as easy. And by missing critical insights like this, users only get part of the story.



## All Your Data

Qlik's Associative Engine fully integrates data from multiple sources, without the data loss problems that happen with SQL joins at load time. This is the equivalent of what is technically known as a "many-to-many full outer join," which means users have access to *all* their data from *all* their sources, in any combinations they need – without leaving any data behind or double-counting data points. This allows people to gain a complete understanding without blind spots or inaccurate insights.

For more about integrating data from multiple sources, check out the Qlik blog post "[Equal Rights for Your Data.](#)"





# Explore Without Boundaries

## Associative exploration

With Qlik, people can explore freely, across all their analytics, without restrictions or boundaries. Users can interact anywhere, within visualizations, charts, graphs, filter panes, or even a global selections interface. After each click, the Associative Engine instantly recalculates all analyses and highlights associations in the data. And because the engine dynamically calculates based on a full set of record-level data, users can pivot in their analysis to new ideas or data, making any selections they want – at any level of detail – without the limits of predefined questions or hierarchies.

The process builds on itself as users ask more questions, add more context, and become more informed at every step along the way. Here's how it works:

### 1. Ask a question

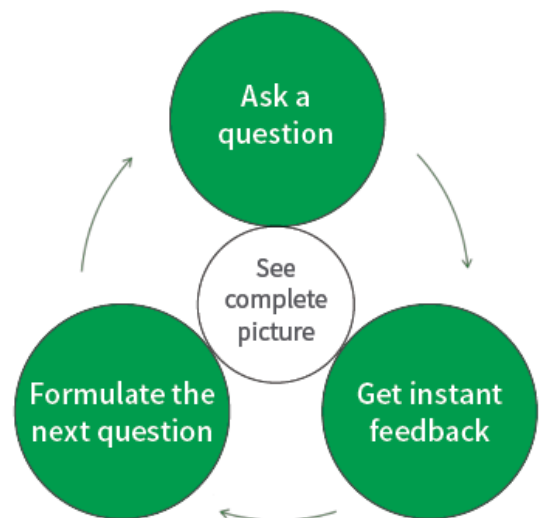
- Use interactive selections and keyword searches
- Ask any questions, not just pre-defined ones
- Interact with any visualization or chart, in any order

### 2. Get instant feedback

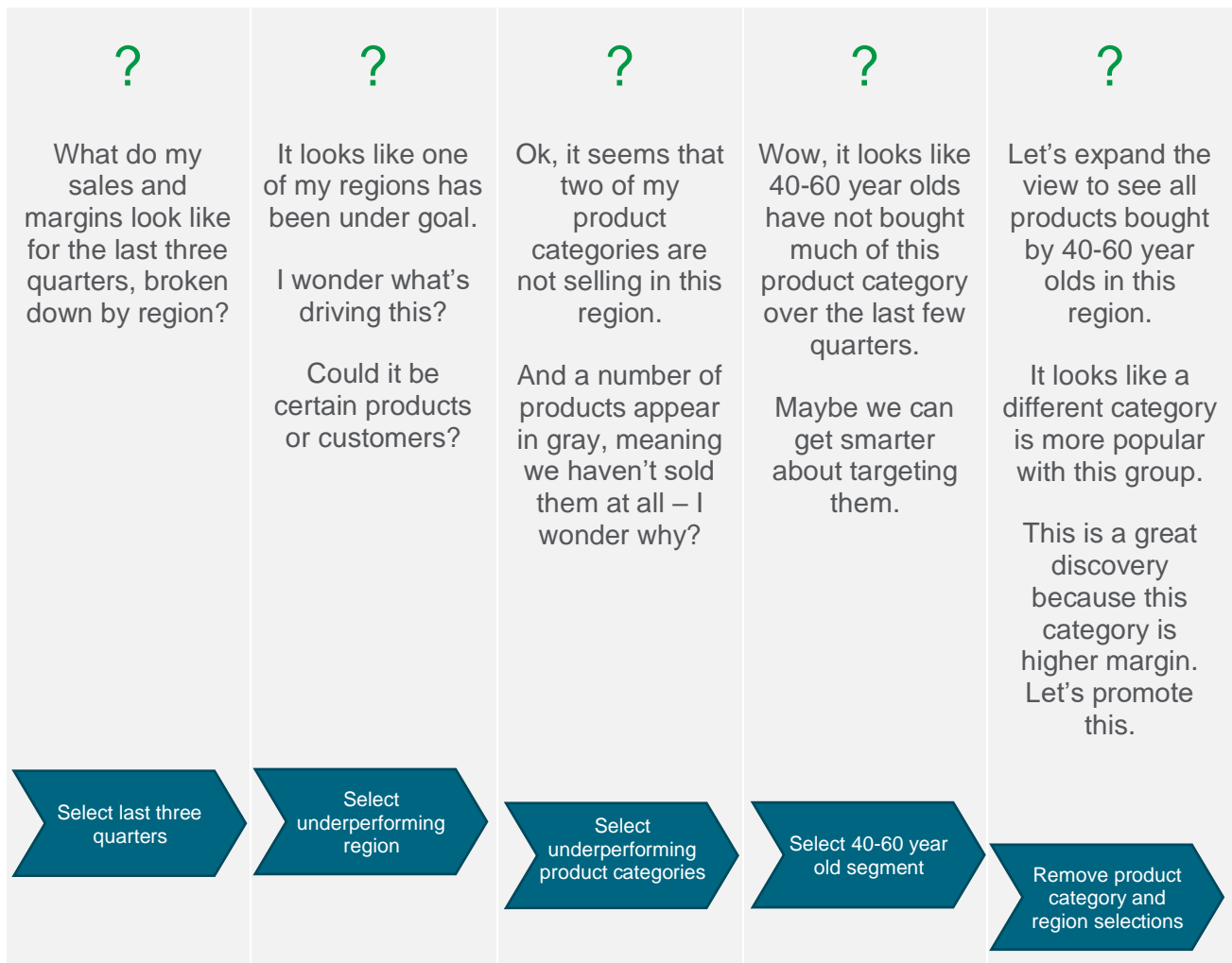
- All analytics and KPIs are dynamically recalculated
- Associations are revealed in green, white, and gray
- New context (selection state) applies across the entire app

### 3. Evaluate results, make discoveries, and formulate the next question

- Spot insights in all surrounding visualizations
- Understand which values are associated with current selections
- Ask follow-up questions based on what you see



## Exploration and Discovery In Action

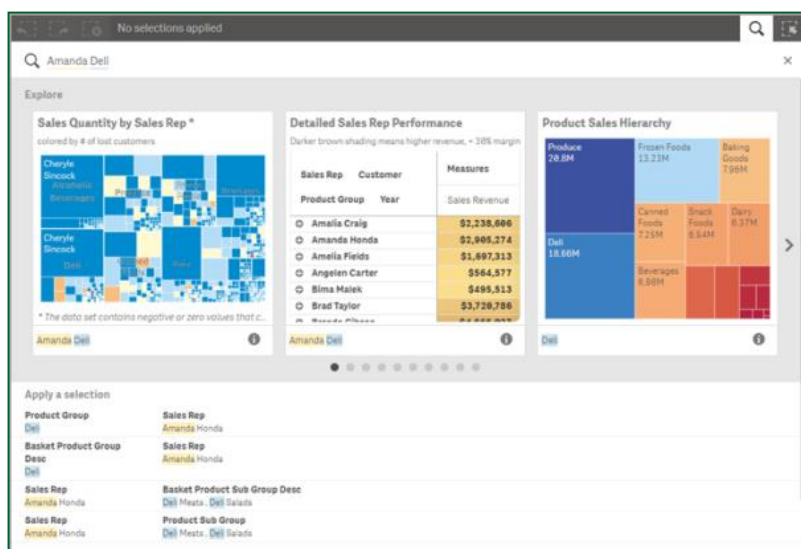


In this example, the user starts with an open-ended question and ends up asking a series of questions that lead to a better understanding of the business as a whole. Insight is generated at every step along the way, and key insights about a particular customer segment in a certain region ultimately reveal a path to more value. Now the user can take action that can directly affect the business.

The next user might start with a different question and end up taking a different path. Either way, questions are answered without the need to build additional queries or visualizations. The value adds up as more users make more discoveries across the business.

## Smart search

Smart search offers a simplified approach to asking questions when a person doesn't necessarily know where to look for the right information. Users can search across all their data using keywords and will get immediate, ranked feedback on where their values match. And if multiple values are entered, results will include matching dimensions – as well as associations that exist between the values – ranked intelligently based on strength of association.



## The power of gray

Qlik has a unique and powerful capability not available in query-based tools; we call it the “power of gray.” It’s the ability for users to see unrelated values in their analysis in addition to related data, relative to their selections. This information often conveys the most impactful insights – products that didn’t sell, for example, or customers who didn’t buy – giving users peripheral vision and helping them discover hidden areas of opportunity or risk. With query-based tools, these values are filtered out, leaving people with only a partial data set and an incomplete perspective.

*A leading global bank and investment firm discovered nearly \$20 million in mortgage pipeline that wasn't associated with any particular loan processor. These mortgages showed up “in the gray.” The firm immediately pursued the lost pipeline, generating tremendous value from a single discovery that wouldn't have been possible with query-based tools.*

## Context is key

Qlik's associative technology maintains a single context for all analytics across an entire application. This means that when a user interacts with an object – like making a selection or searching to refine context – all analytics and data relationships are instantly updated to reflect this new context. This allows a user to explore across all their visualizations, at different levels of detail, at the speed of thought. They can spot potential areas of interest and get a sense of where to go next.

## No restrictions

With Qlik, users won't experience any restrictions or boundaries. People are free to explore and search all their data, in any direction they want, probing possible data relationships and key areas of concern as they follow their own paths to insight. This flexibility is critical – because looking at information from many angles, and at many different levels, gives people a better understanding of the situation as a whole.

*“We discovered that some of our perceptions about our best customers conflicted with the information we saw in the Qlik applications. Now, we can see who is paying consistently.”*

- CIO

*“What really distinguishes Qlik Sense is its extensibility and the Associative Engine, which gives new meaning to the word ‘discovery.’”*

- Butler Analytics

*“With Qlik's associative search capabilities, our customers are now able to intuitively analyze more than a terabyte of data with speed-of-thought response – without being confined to a limited data set or a defined path of questions.”*

- Data Warehouse Manager

## Speed of Thought

Query-based tools may be sufficient for creating visualizations, but only a handful of skilled users can or want to do that. And even fewer people can create the queries to support these visualizations. So what happens to the majority of your business users? They end up with slow, static visualizations and reports – which are restricted to linear exploration on partial subsets of data. Unless the exact questions are predicted in advance, and the queries are built appropriately, people have to go back to the data experts. Once again, the “ask, wait, answer” cycle puts the brakes on getting insights.

In contrast, the Qlik Associative Engine dynamically calculates analytics and highlights associations as quickly as users can think of questions. With instant responses to any question, people can get the answers they need without having to wait – for the system or for the experts. This means faster time-to-value and hundreds of more informed decisions every day.

The Associative Engine achieves an unmatched combination of speed and flexibility that simply isn't possible with query-based tools. How? By combining several unique, patented, in-memory innovations – including compressed binary indexing, logical inference, and dynamic calculation. And Qlik delivers this unprecedented performance on big and small data sets, in combination with each other. As a result, large communities of concurrent users can ask unanticipated questions and get instant feedback, leading to new questions and new discoveries.

This engine is our core technology advantage – and the brains and brawn behind Qlik's Associative Difference.

## Qlik Versus Query-Based Tools

The Associative Difference is unique to Qlik. Nobody else can provide it.

Query-Based Tools	The Qlik Associative Engine
Can't combine large numbers of different data sources	Combine any number of disparate data sources
Must fully model data and hierarchies in advance	No need to fully model data or hierarchies
SQL joins at load time double-count data or leave data behind	No joins at load time, and no data left behind
Can't handle dirty, incomplete, or sparse data	Dirty, incomplete, and sparse data is loaded and exposed
Often require pre-aggregation for large data sets	No pre-aggregation needed, and transaction details are always available
Can't support non-linear exploration or new questions	Built for interactive, freeform exploration
No global search or search-based analysis	Global search for data, associations, and analytics
Multiple objects don't stay in context without wiring	All objects stay in context together automatically
Loss of unrelated values	Unrelated (gray) values are exposed in analysis
Only offer static hierarchies and no on-the-fly calculations	Delivers dynamic hierarchies with on-the-fly calculations
Can't keep up with a user's thinking and derails analysis	Provides speed-of-thought analysis
Slow performance, especially with many users and queries	High-performance dynamic calculation for large numbers of concurrent users
Need to query production systems or build DW	No effect on production systems and no need for DW
Risk exposing sensitive data if database is not secure	Offers dynamic data reduction for data security
Limited capabilities for big data sources	Scalability for big – and small – data sources

## Additional Resources

For more information, head to [qlik.com](https://qlik.com).

To experience Qlik's associative technology in action, check out our [online demos](#).

To get started, register to use [Qlik Sense Cloud®](#) for free.

## Appendix: Let's Geek Out on the Technology

The Associative Difference is possible because of the unique capabilities built into the patented Qlik Associative Engine. Forged in 15+ years of innovation and investment, this powerful calculation and data indexing engine is Qlik's core advantage. Designed specifically to support interactive, freeform exploration and analysis, the Associative Engine offers high-speed dynamic calculation and speed-of-thought response for high numbers of users – and both big and small data. In this appendix we'll cover the key technical aspects of how our engine works and what makes it different.

### Compressed Binary Indexing

#### Data assembly

The Qlik Associative Engine integrates a full set of record-level data from multiple sources into its in-memory engine. Qlik provides a robust set of data preparation and integration capabilities for transforming and bringing together disparate data sources, including:

- Visual interfaces for loading and transforming data
- Smart data profiling for data relationships and values
- Powerful scripting for complex data integration scenarios

And because table joins are performed dynamically as users explore, all data from all sources is retained. This is the equivalent of a “many-to-many full outer join,” which eliminates the data loss associated with one-sided SQL joins executed at load time.

For more about integrating data from multiple sources, check out the Qlik blog post [“Equal Rights for Your Data.”](#)

## Binary indexing

The Associative Engine indexes and stores data in a highly optimized, in-memory, compressed binary format that maximizes performance and drives the associative exploratory experience. There are four main pillars to this columnar approach:

- **Data relationships:** These are managed by the engine and defined through common column names in the multi-table model.
- **Data storage:** Instead of storing values repetitively, the engine creates binary pointers for each unique value, storing the actual values only once and ensuring that data is never double-counted.
- **Calculations:** Analytics are not pre-calculated, meaning that any calculation can be performed on demand – leading to near limitless flexibility in terms of both exploration and aggregation.
- **Table joins:** Because table joins are executed by the engine dynamically as the user explores, calculations will always occur on the correct tables and the correct set of values, avoiding the risk of incorrect results.

This unique columnar, binary indexing capability sets the foundation for the Qlik Associative Engine to support interactive, associative data exploration and on-demand calculation across high numbers of concurrent users and large data sets.

## Logical inference and calculation

As large numbers of users explore information, it's a challenge to provide each of them with the flexibility and “speed-of-thought” response they need. Users constantly interact, making selections on the fly, searching for new information, and executing complex calculations on different subsets of data that can't be predicted in advance. The Qlik Associative Engine handles this difficult task through a two-step process every time a user interacts with an application.



## Logical inference

The first step is logical inference, which means determining data associations relative to the current context:

1. Before a user selects anything, all data is in play.
2. Each time a user makes a selection, the engine immediately calculates what distinct values – in all related tables – are associated to the new context.
3. Then the engine highlights the important relationships in the data, revealing both associated and unrelated values to the user.

This logical inference step allows the engine to minimize the data needed to support subsequent calculations while maximizing performance. It means that the Associative Engine always knows which data in which tables to use, and always take the fastest path to the right data.

## Dynamic calculation

The second step is dynamic calculation:

1. Once the associated data set has been determined by logical inference, the Associative Engine calculates all analytics and aggregations in real-time – storing the results in a cached hypercube.
2. Calculations are broken into pieces and performed on various tables as needed, and a number of advanced techniques are used to maximize performance.
3. Every user interaction triggers the engine to recalculate the values stored in the hypercube, providing the user with answers specific to their unique questions.

Because the data is highly compressed and optimized in-memory, and because data sets for calculations are minimized by logical inference, this dynamic calculation happens extremely quickly – giving users the answers they need when they need them, and increasing time-to-value.

To learn more about Qlik's patented Associative Engine technology, read the white paper "Interactive Data Exploration With an In-Memory Analytics Engine."

## Queries Can't Keep Up

This combination of flexibility and performance is difficult – if not impossible – for query-based tools. There's simply no way for them to provide speed-of-thought responses for high numbers of concurrent users who are analyzing large, complex data sets and asking questions that aren't predefined. Even query-based tools that claim to have in-memory technologies for supporting ad-hoc queries are still limited by the very same query structure they rely on.

Compressed binary data storage, logical inference, and dynamic calculation: This unique combination is what enables the Qlik Associative Engine to make highly complex tasks incredibly simple and transparent for the user. Our engine has evolved over 20 years of innovation and investment, delivering unprecedented value to more than 48,000 customers.

## About Qlik

Qlik's vision is a data-literate world, one where everyone can use data to improve decision-making and solve their most challenging problems. Only Qlik offers end-to-end, real-time data integration and analytics solutions that help organizations access and transform all their data into value. Qlik helps companies lead with data to see more deeply into customer behavior, reinvent business processes, discover new revenue streams, and balance risk and reward. Qlik does business in more than 100 countries and serves over 50,000 customers around the world.

[qlik.com](https://qlik.com)



© 2020 QlikTech International AB. All rights reserved. All company and/or product names may be trade names, trademarks and/or registered trademarks of the respective owners with which they are associated.

ASSOCDIFF0427