White Paper

Qlik's Associative Model

See the Whole Story that Lives Within Your Data

August, 2015

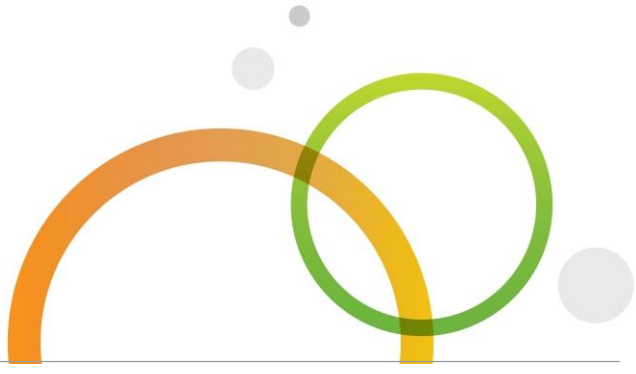


Table of Contents

Introduction

This white paper describes the unique associative model underpinning Qlik® products. It provides an overview of this highly interactive user experience, outlines the differences between Qlik's technology and query-based visualization tools, and explains how people can fundamentally improve the way they conduct analysis to achieve deeper, more accurate insights. The core questions this paper answers are: "What is Qlik's associative model?", "How is it different?", and "How does it help people make better decisions?" The target audience for this paper includes business and technology decision makers who are evaluating and comparing visual analytics software products.

Qlik's associative model

Qlik's associative model is the exploratory capability in our products that enables users of all skill levels across an organization to *see the whole story* that lives within their data.

Powered by our patented, in-memory associative data indexing (QIX) engine, Qlik products allow users to probe all the possible associations that exist in their data, across all of their data sources. People can freely explore using simple searches and selections, asking questions in any direction without restrictions or boundaries. Unlike query-based visualization tools and traditional BI, Qlik products do not limit users to predefined hierarchies or preconceived notions of how data should be related, allowing them to fully explore and understand how it truly is related.

Creation is not exploration

Most query-based visualization tools allow skilled users to conduct analysis through the assembly of new visualizations and queries. But even with claims of "self-service" simplicity, this approach requires a thorough understanding of underlying data models and how to construct analytics. So what happens *after* the experts build the visualizations? What happens when large communities of business users without creation skills simply want to search and explore data further?

Of course, Qlik products offer an outstanding experience for creating visualizations. But while this is the end of the road for other tools, it's only the beginning of the journey with Qlik. Qlik's associative model provides everyone with a simple and powerful way to explore analytics and refine context, discover new data relationships, and gain the right insights to make the best decisions. This fundamental, interactive behavior is unique to Qlik products, made possible by our innovative associative model and underlying QIX engine.

Qlik at a Glance

Founded in Lund, Sweden in 1993

Approximately 36,000 customers in more than 100 countries

Solutions driving leadership in the visual analytics market

- **Qlik Sense®**
  Next-generation visual analytics
- **QlikView®**
  Guided analytics and dashboards
- **Qlik® Cloud**
  Web-based services for analytics and data
- **Qlik® Analytics Platform**
  Developer platform for building visual analytic apps

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

More than 2,000 employees worldwide

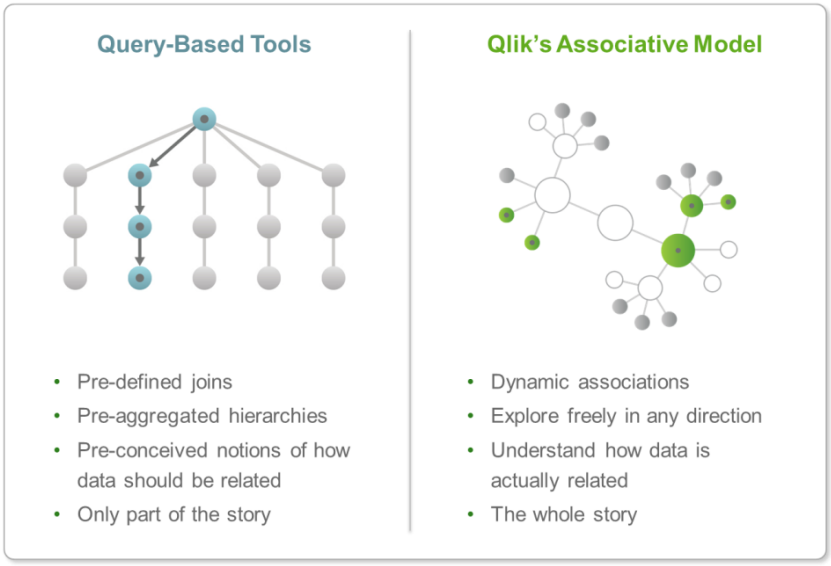
Recognized leader by industry analysts including Gartner, BARC, Forrester, Butler Analytics and Ventana Research

Headquartered in Radnor, PA, USA

NASDAQ: QLIK

Query-based visualization tools – only part of the story

Visualization tools that rely on SQL to query, join, and analyze data ultimately create blind spots in understanding. These tools use a structured, linear approach to visualize partial subsets of data, instead of offering the freeform exploration and search across all data that is offered by Qlik®'s associative model. Often these tools attempt to simulate Qlik's exploratory experience, but they quickly hit limits in terms of flexibility and performance. The only way to really ask new questions in a query-based tool is to build new visualizations and queries, which is inefficient and difficult for the average business user.



The drawbacks of query-based tools

The following are some of the drawbacks of query-based tools:

> *"Qlik's in-memory-driven associative experience continues to be a differentiator in the BI market, in particular in comparison to other visual discovery vendors…"*
>
> - Ovum, SWOT Assessment: Qlik QlikView 11.2, March 2014

> *"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."*
>
> - VP Information Systems

Partial views

Query-based tools extract a subset of data from the main dataset and return it in the form of a query result set. This result set is tied to a specific visualization and is divorced from the full dataset, making it difficult to get a complete picture of how things are related. In addition, the result set only includes associated values, not unrelated values, which often convey the most critical insights.

Limited flexibility

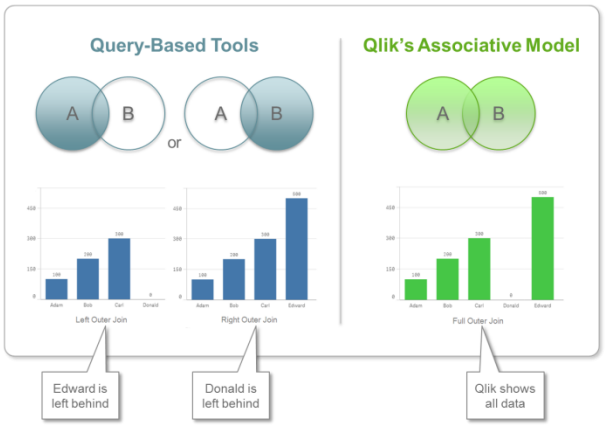
Visualizations and queries must typically be built in advance by power users, limiting business users to the pre-conceived notions the developers thought were the most relevant. If a business user wants to ask new questions or move in a direction beyond simple filtering, the only choice is to build a new query or visualization.

No context

As a user interacts with a visualization to apply filters, other visualizations and objects are not automatically kept in context. This is because each object is fed by a discrete query result set, divorced from the others. Some tools attempt to wire objects together into dashboards at the application level, but this approach cannot support more than a few objects without suffering performance issues.

Data loss

If information is loaded from multiple data sources, query-based tools may not be able to achieve the equivalent of a full outer join. This results in data loss from one source or the other, which a business user may not even be aware of.



Incorrect results

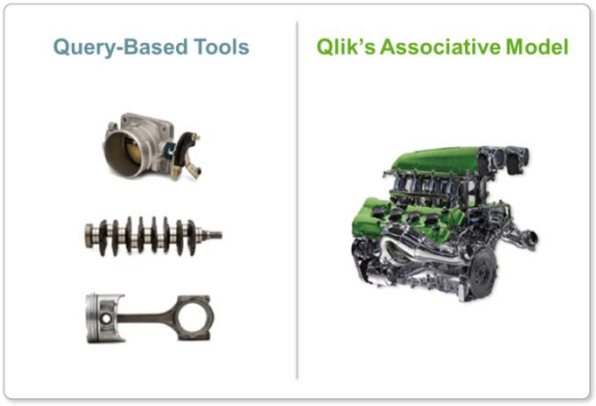
Query based tools require users to have strong familiarity with the data model in order to ensure correct results. There is a significant risk for incorrect calculations, as values can be double or even triple counted if queries and joins are not defined correctly.

Poor performance

Query-based tools are at the mercy of the database(s) 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 explore information, underlying databases simply can't keep up.

An analogy

Let's say the goal is 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 would be able to evaluate one part at a time, but would be left on your own to attempt to understand the relationships (associations) between the parts, and how they fit together as a cohesive whole to create a working engine.



With Qlik's associative model, however, you have the ability to analyze a complete working engine, with each part having a 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 the associative model.

With Qlik's associative model, when users look at information, they know precisely how it is 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 the entire data set responds. They are not limited or restricted to seeing only a subset of the data relationships contained in a query result set.

The associative model A to Z

Qlik®'s associative model allows users of any skill level to explore and refine context through simple searches and selections, starting anywhere and going anywhere. Every time the user clicks, the QIX engine instantly responds, dynamically recalculating all analytics based on the new context and highlighting associations in the data across all dimensions and sources. This instant feedback encourages users to think of new questions and continue their individual path of exploration, imposing no limits on the user and prompting them to ask and answer the next question.

Understanding associations

An association is simply a relationship between a value in one field and a value in another. For example, a product may have been sold in a certain country but not others. This product would be associated with the country it was sold in, and unrelated to the others. In Qlik's associative model, 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 in nature — a value that has a relationship to another, or negative — a value that is unrelated to another. It is important to note that the unrelated (gray) values provide as much insight as the positive (white) ones – often indicating new opportunities or areas of risk.

Every unique value can be associated with many others in the data set. A product could have many associations in addition to countries, including the customers that 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 the set of selected products and countries.

Users won't always know what associations exist or which are important in advance. A query-based tool could easily report that a product was sold in several regions. However, noticing that it and other products were not sold through certain channels in a given region may not be as easy. And by missing critical insights like this, users are only getting part of the story.

All your data

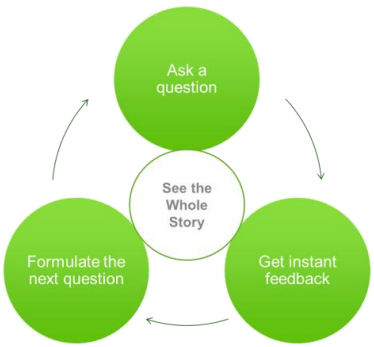
Qlik's associative model fully integrates data from multiple sources, without suffering data loss from joining or filtering information at load time. Having access to all the data allows the QIX engine to dynamically expose associations as needed after every user interaction. This means the user gets a complete understanding without blind spots or incorrect insights due to missing information.

For more insight on bringing together data from multiple sources, check out the blog post entitled "Equal Rights For Your Data" on the Qlik blog.

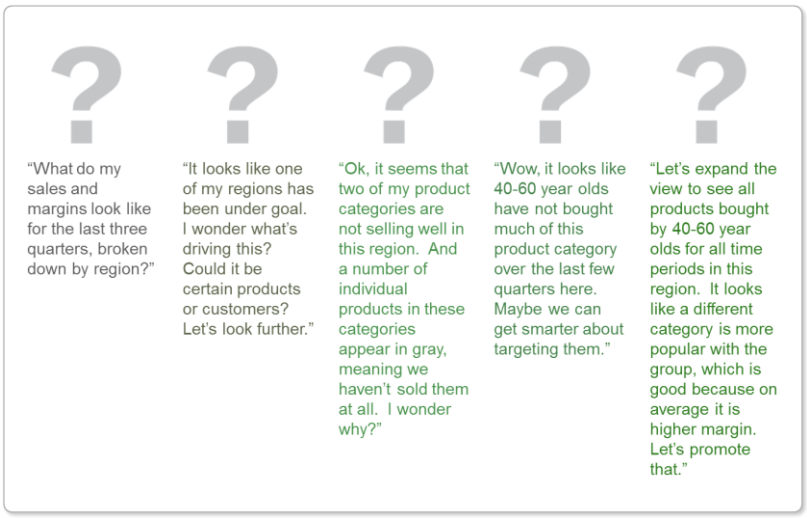
Explore and search

With Qlik®'s associative model, users can explore analytics and associations through a step-by-step process of questioning and understanding. They can start with open ended questions, not knowing where they will lead, and ultimately take their own paths to insight. The process builds on itself as users ask more questions, add more context, and become more informed at every step along the way.

1 Ask a question
   - Use simple selections and natural language searches
   - Ask any questions, not just pre-defined ones
   - Interact with any visualization or chart
2 Get instant feedback
   - Associations revealed to the user (green, white, gray)
   - All analytics and KPIs dynamically recalculated
   - New context (selection state) applies across the entire app
3 Evaluate results, make discoveries, and formulate 'the next question'
   - Spot insights across all visualizations and objects in an app
   - Understand how data is associated relative to current selections
   - Formulate follow-up questions based on these insights



An example of exploration and discovery



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. A key insight about a particular customer segment in a certain region is ultimately revealed, prompting the user to take action. The next user might start with a different question and end up taking a different path, but either way, the 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 allows for a simplified approach to asking questions, when a person does not necessarily know where to look for the right information. People can search across all their data using keywords, and will get immediate, ranked feedback on where their values match. Furthermore, if multiple values are entered, results will include not only the matching dimensions but also the associations that exist between the values, sorted intelligently based on strength of association. This powerful capability provides an accessible way for people to ask questions and get immediate insight.

Key benefits

The following are some key benefits of Qlik®'s associative model:

AGGREGATE INDUSTRIES

"We discovered that some of our perceptions of who were our best customers conflicted with the information we started to see in the Qlik applications. Now, we can see who is paying consistently."

- CIO

The power of gray

With the associative model, users can see not only the values that are associated with their selections, but also those that are unrelated. These values appear in gray, and often convey the most impactful insights such as areas of new opportunity or risk. For example, if you select a set of products, you may notice there were certain customer segments these products were not sold to. This capability is unique to Qlik and a great way to find unexpected insights.

BUTLER ANALYTICS

"… What really distinguishes Qlik® Sense is its extensibility and the associative data engine, which gives new meaning to the word 'discovery.'"

- Butler Analytics, Qlik Sense and Tableau Positioning, June 2015

Context is key

Qlik's associative model maintains a single context for all analytics across an entire application. When a user interacts with an object, making a selection or search 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', to spot potential areas of interest and get a sense of where to go next.

bics

"With Qlik's Associative Search capabilities, our customers are now able to intuitively analyze over 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

No restrictions

Qlik's associative model does not place any restrictions or boundaries on the user. 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 when looking at information from many angles, and at many different levels, to ultimately get a better understanding of the situation as a whole.

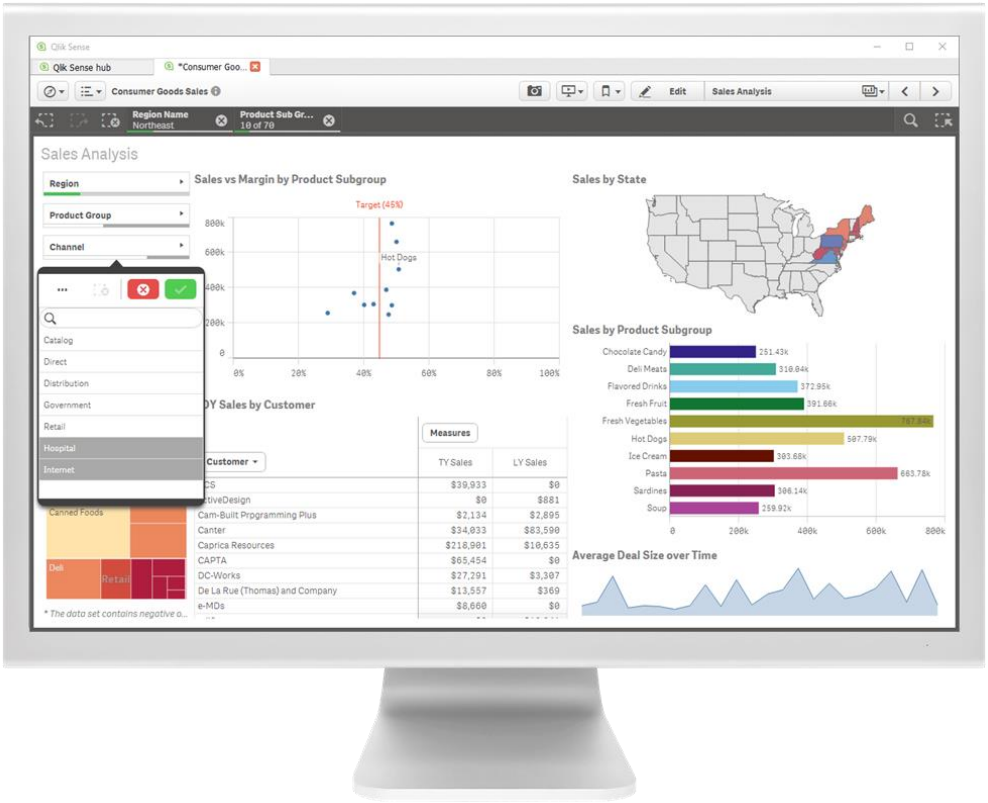
Qlik's associative model vs. query-based tools

The following table summarizes the key differences between Qlik's Associative Model and query-based visualization tools:

	Query-based tools	Qlik's Associative Model
Exploration	Limited to drill/filter within a visualization and query result set	Unrestricted ability to explore across all objects and multi-source data
Interactivity	Visualizations support predefined drill/filter only	All objects support full set of user interactions
Responsiveness	A single visualization responds to the filters applied within it	All objects respond to show impact of interactions anywhere in the app
Asking a new question	Must create or modify visualizations and queries to ask new questions	Ask new questions through simple selections and searches
Direction of questioning	Limited exploration within query result sets	Start anywhere and go anywhere, no restrictions or boundaries
Data association	Not available or very limited, queries return related values only	Complete set of associations after each click, including unrelated data
Keyword search	Keyword search typically not available or very limited	Search across all data using combinations of keywords
Context	Objects 'wired together' into dashboards at application level	Engine automatically keeps all visualizations and objects in context
Dynamic calculation	Calculation for a single object when a user applies a filter	Calculation for all visualizations and objects after every click
Table joins	Defined at application level, executed at load time	Known to the engine, executed as needed when user clicks
Multiple data sources	Limited data integration, SQL joins at load time can result in data loss	Robust integration of multiple sources with no loss of data
Accuracy	Users must know data model to ensure accuracy	Engine ensures the right data is used for calculations
Scalability and performance	Dependent on underlying database, large numbers of concurrent users risk poor performance	Flexibility and performance for high numbers of concurrent users and large data sets

Additional resources

- For more information and to download Qlik® products, head to qlik.com.
- To experience the associative model in action, check out our online demos.
- To create your own analytics in the cloud, register to use Qlik® Sense® Cloud for free.



Appendix: Let's geek out on the technology

The associative model in Qlik® products is only possible because of the unique capabilities of the patented QIX Associative Data Indexing Engine. This powerful calculation and data indexing engine has powered Qlik products for the last fifteen years. It is built to respond instantly to the user, offering high-speed dynamic calculation and 'speed of thought' response after each click. In this appendix we will cover the key aspects of how our engine works and what makes it different.

Data indexed and highly optimized, in-memory

Robust data integration

The QIX engine integrates a full set of record level data from multiple sources into its in-memory engine. Qlik provides a robust set of data integration capabilities for transforming and bringing together disparate data sources, including graphical interfaces for loading data, data relationship profiling, and powerful scripting for complex transformations. And because table joins are not executed until user interaction, all data from all sources is retained. The QIX engine achieves the equivalent of a full outer join, and does not suffer the data loss associated with one-sided SQL joins executed at load time.

For more insight on bringing together data from multiple sources, check out the blog post entitled "Equal Rights For Your Data" on the Qlik blog.

Compressed binary storage

Data is stored in a highly optimized, in-memory compressed binary format that optimizes performance and drives the associative model of exploration. Data relationships are known to the engine, defined through common column names in the multi-table model. Instead of storing values repetitively, the engine creates binary pointers for each unique value, storing the actual values only once. Analytics are not pre-calculated, meaning that any calculation can be performed on demand — leading to near limitless flexibility. And because data relationships are managed by the engine, calculations will always occur on the correct tables, avoiding the risk of incorrect results.

This unique columnar, binary indexing capability is the foundation for the QIX engine supporting interactive, associative data exploration and on-demand calculation, across high numbers of concurrent users and large data sets.

Dynamic calculation for 'speed of thought' exploration

As high numbers of users explore information, it is 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 cannot be predicted in advance. The QIX engine handles this difficult task through a two-step process every time a user interacts with an application.

Logical inference

Step one is logical inference — which essentially means determining data associations. Before a user selects anything, the assumption is all that data is in play. However, each time a user makes a selection, the engine immediately calculates what distinct values in all related tables are associated to the new

context. By doing this, the engine can highlight for the user the important relationships in the data, revealing both positive and negative associations. In addition, this process allows the engine to minimize the data needed to support subsequent calculations, maximizing performance. Logical inference means the QIX engine always knows which data in which tables to use, and always takes the fastest path to the right data.

Dynamic calculation

Step two is dynamic calculation. Once the associated data set has been determined by logical inference, the QIX engine performs all calculations and aggregations on-demand — storing the results in a cached hypercube. Calculations are broken into pieces and performed on individual tables as needed to maximize performance. Every user interaction triggers the engine to recalculate the values stored in the hypercube, providing the user with answers specific to their unique questions. And because data is highly compressed and optimized in-memory, and data sets are minimized by logical inference, the calculation is extremely fast.

For more information on Qlik's patented engine technology, see the white paper entitled "Interactive Data Exploration With An In-Memory Analytics Engine", written by Mike Ferguson from Intelligent Business Strategies.

Queries can't keep up

Achieving this combination of flexibility and performance is extremely difficult if not impossible with query-based tools. There is simply no way to provide 'speed of thought' response for high numbers of concurrent users, analyzing large, complex data sets, who are asking questions that are not pre-defined. 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.

The unique combination of compressed binary storage of data, logical inference, and dynamic calculation allows the QIX engine to make a highly complex task incredibly simple and transparent for users. Our engine has evolved over fifteen years of innovation and investment, delivering unprecedented value to over 36,000 customers.