StarQube

INVESTMENT DECISION IN A BOX

StarQube - A solution for data-driven asset management

The rise of alternative and ESG data

The world of asset management has always been a large consumer of data (price data, accounting data, earnings consensus, macroeconomic data, etc.). Today, it is faced with an explosion in the number of data sets: nearly 2,000 non-traditional data sets are currently available and relevant to buy-side players, and conservative forecasts are for 5,000 data sets by 2024.

ESG is one of the drivers of this alternative dataset inflation. Bloomberg estimates that ESG assets under management will reach $53 trillion by 2025, representing more than a third of total assets under management (vs. ~25% at the end of 2020). The pandemic will have accelerated an already strong trend by revealing how ESG can help assess a new set of financial risks.

The pressure from consumers and investors for greater transparency of companies on their environmental, social and governance policies will therefore increase. For their part, asset managers are also under pressure from investors and regulators to demonstrate their commitment to socially responsible investment. They need to be able to collect raw ESG data on their large investment universes to build proprietary and composite ESG scores.

The number of ESG datasets will therefore continue to grow significantly over the next few years.

Opportunities and challenges for asset managers

This proliferation of alternative/ESG datasets is a tremendous opportunity for asset managers: it considerably broadens the scope of fundamental research, creates new areas of thematic differentiation and allows for the development of new discretionary or quantitative investment strategies.

Asset managers who know how to take advantage of this new data will be able to create innovative sources of alpha and build a competitive advantage. But the reverse is also true: those who are slow to adapt their governance and tools to the rise of alternative data will see their return prospects and margins erode.

The increase in the amount of data thus raises many challenges for asset managers in the years to come:

- Budgetary: they must be able to have the data they need while controlling their acquisition costs and avoiding redundancy (data paid for several times from different suppliers).

- Infrastructure: they must have database management systems to organize and handle large volumes of data efficiently.

- Data administration: they must have the IT teams and tools (APIs, connectors) to collect new data electronically, in real time.

- Governance: the adoption of a data-driven asset management model requires a strong managerial impetus and the establishment of a dedicated data team to ensure data quality.

As recommended by Accenture in its report "The power of data-driven asset management", transforming data into a long-term differentiating asset requires action at three levels:

1. Building a strong data foundation including data governance, data sourcing management, data quality control, data architecture and dedicated technology.

2. Harnessing advanced technologies to extract maximum information from data using artificial intelligence or machine learning techniques.

3. Spreading a real data culture within the teams.

CASE STUDY: ESG research teams faced with the multiplication of data sets

The proliferation of ESG data providers and datasets raises a number of practical difficulties for teams in charge of data collection and manipulation. Some of the most common difficulties are as follows:

- *At an instrument level*: difficulty in exploiting all the ESG data available on a single financial instrument, as each data provider uses a proprietary (and therefore different) identifier.

- *At the issuer level*: the same problem, an issuer may be identified by a different code from one database to another; it is, therefore, difficult to exploit all the information on the same issuer.

- *Issuer/instrument linkages*: it can be difficult to roll up ESG scores calculated on an issuer to all instruments (stocks, bonds) issued by that issuer—especially when the instruments are issued by several companies within the same group.

- *ESG histories*: where the ESG team has implemented an ESG scoring methodology, it can be very complex to recalculate reliable historical scores if the raw data has not been timestamped as soon as it is collected; it may have been subject to revisions by its providers.

- *ESG portfolio management*: this requires portfolio managers to be able to (1) visualize the ESG scores of their investment universe; (2) calculate the aggregate ESG scores of their portfolio and, where appropriate, compare them with those of their benchmark; (3) optimize their portfolio not only in terms of performance/risk, but also by taking into account the ESG objectives/constraints of their management.

StarQube: the solution that turns your data into an asset that can be easily valorized

StarQube offers front-office teams a modular solution for organizing their data and supporting their investment decision process.

The solution can integrate all types of data (financial or extra-financial, structured or unstructured, on liquid or illiquid assets, and for all frequencies up to intraday). It is not specifically dedicated to ESG data, but given the importance of socially responsible investment, it is used in particular for ESG research and investment purposes.

The StarQube solution has been developed on a "data" layer, on top of which an overlay consisting of portfolio management modules (optional) has been built.

The « data » layer

The StarQube platform's "data" layer is, itself, made up of three main tools:

1. A proprietary database (Qube)

2. A data update management module (SQ Data Loader)

3. A simplified data manipulation language (FQL)

The database has unique features:

- NoSQL type: it is highly compressed and optimized for heavy financial calculations (backtests, portfolio optimizations). It offers response times ~10 times faster than a traditional relational database.

- It can host any type of data and is perfectly scalable to accommodate and manage very large volumes of data.

- It is natively "point-in-time". All data is timestamped. When data is revised it is not overwritten in the database; instead, new data is created with its timestamp. It is, therefore, possible to launch backtests without look-ahead bias since, by construction, they only use the data that was available at the time.

- These characteristics allow only raw data to be stored in the database; the calculated data are stored "on the fly", as the response times allow them to be available almost instantaneously.

The database is fed by an industrial update management module (SQ Loader):

- SQ Loader is a task scheduler that manages the data update functions throughout the day.

- Updates are configured in files ("connectors") that communicate with data providers.

- Data sources can be external or internal to the asset management company. They can be financial/extra-financial data providers, but they can also be a position-keeping tool (PMS/IBOR) or portfolio position files from custodians.

- The connectors are set up to collect data at a certain frequency (e.g., once a day for extra-financial data, every two minutes for portfolio positions). They can also retrieve data automatically as soon as it has been made available by their provider (on an SFTP server).

- The connectors can incorporate data quality control instructions to issue an alert in case of missing or inconsistent data. It is then up to the team in charge of data administration to verify the anomaly, have the erroneous data corrected by its supplier and restart the file download.

- SQ Loader can parallelize tasks and handle millions of tasks each day, providing an industrial solution for updating the asset manager data.

The data can then be manipulated using a simplified language (FQL):

- FQL is a proprietary language developed by StarQube whose difficulty level does not exceed that of Excel functions.

- All common mathematical/statistical/financial formulas have a function defined in FQL.

- FQL thus allows easy manipulation of data and objects stored on StarQube, for example, to build proprietary ESG scores (with pillars and sub-pillars) from the raw data collected.

StarQube's integration into the asset manager's ecosystem

StarQube does not provide any data but integrates naturally into the asset manager ecosystem. When an asset manager adopts the platform, the StarQube team takes care of:

1. The creation of a unique instrument referential: all financial instruments in the asset manager's universe are linked to a unique identifier so that all data sets, regardless of their source, communicate with each other.

2. The creation of a unique issuer referential: all issuers in the asset manager's universe are linked to a unique identifier and the hierarchy of company cascades is created so that each company is connected to its Group Head. It is therefore easy to construct the ESG scores of a holding company and to cascade them down to all the financial instruments issued by the group's companies.

3. The creation of connectors with data providers: the StarQube team builds connectors with all the financial/extra-financial data providers to which the asset manager subscribes, and if necessary with (1) the internal databases, (2) the position keeping tool (PMS), and (3) the order execution tool (OMS).

In addition, StarQube can take over the administration of its clients' data, based on a Data-as-a-Service model. The StarQube team then manages the recurrent and time-consuming tasks of data update (collection, cleaning). Customers who opt for this service can then focus on the higher value-added tasks of analyzing, transforming and exploiting their data.

Portfolio management modules

On top of the "data" layer described above, StarQube has also built (optional) portfolio management modules. The objective is to offer an integrated investment decision platform that can assist front-office teams from data collection to order execution.

The different modules allow:

- Backtesting of an investment strategy (SQ Backtester)

- Parameterization of elaborate portfolio risk models (SQ Risk Model Builder)

- Setting up of complex sets of objectives and constraints to optimize a portfolio (SQ Optimizer)

- Visualization of all the key indicators on a portfolio, line by line and in aggregate, triggering portfolio optimization and generating and transmitting orders to be executed to the OMS (SQ Analyzer)

- Communication with the user's usual work tools via APIs (SQ API): Python, Matlab for data scientists, C, C++, .NET/C#, Java, REST for developers, Excel for all

Collaborative and secure, the platform also allows all data and objects (backtests, risk models, optimization models, portfolios, etc.) to be shared and be granted permissions (read-only or read/write). The ESG team can thus have sole access rights to raw ESG data from external suppliers, then build composite indicators and open up read-only rights to the portfolio management teams.

CASE STUDY: Building and using a proprietary ESG score with StarQube

The ESG team subscribes to data from five external suppliers. It thus has 500 fields of information on a universe of 5,000 issuers representing several tens of thousands of financial instruments (shares, bonds).

With StarQube, all this data will be structured around a single issuer-instrument referential, stored in a highly compressed NoSQL database, automatically updated every day via connectors built with the five providers, and timestamped.

With a few FQL functions, the ESG team will be able to easily apply its proprietary ESG methodology: build composite ESG scores that can use different fields from one sector to another, cascade the ESG scores of groups' leading companies down to all the instruments issued by the groups' companies, calculate ESG scores on the fly from raw data, and have unbiased historical scores because they are calculated from reliable, validated and timestamped data.

It will be able to open access to its composite scores (read-only) to employees of its choice in the quantitative research or portfolio management teams.

The quantitative research team will be able to backtest investment strategies based on the composite ESG scores, with no look-ahead bias in backtests.

The portfolio management team will be able to view the ESG score of its portfolio, optimize its investment strategy by integrating an ESG constraint, and transmit its orders for execution.

Summary

The era of Big Data is still in its infancy in asset management. The number of alternative data sets available (ESG in particular) will increase significantly over the next few years.

Few players—including hedge funds—have transformed their organizations to adapt to this new situation. The adoption of a data-driven asset management model requires a strong managerial impetus to build a solid foundation for data and to instill a true data culture within the teams.

However, everyone understands the importance of the challenge: those who know how to transform their data into an asset will be able to build innovative sources of alpha and sustainable competitive advantages.

StarQube offers an integrated platform for organizing asset managers' data and helping them make investment decisions. Data is structured around a single referential, hosted in a highly compressed database, automatically updated and monitored throughout the day via connectors with the asset manager's data providers, and timestamped.

The datasets communicate naturally with each other and can therefore be easily manipulated to build "augmented data" (e.g., proprietary ESG scores) from the raw data.

Connectors with custodians, position-keeping (IBOR/PMS) and order-execution (OMS) tools, APIs and portfolio management modules allow front-office teams to be supported throughout the investment decision process—from data collection to order transmission.

The platform is collaborative and allows data and objects to be shared securely between different teams.

StarQube offers many advantages to asset managers so that they can exploit the full wealth of their data.

About StarQube

Established in 2013, StarQube develops a suite of data organization and front-office solutions designed to streamline investment processes, starting from data acquisition, through fast backtesting of client-defined strategies and risk management, all the way to portfolio rebalancing, dispatching of orders and reporting. StarQube provides asset managers with a nimble way to automate each of their clients' bespoke portfolio management processes while saving on structural and data costs. StarQube is particularly well-liked by systematic strategies and socially responsible investments.

Contact

StarQube
Rue des Corps-Saints 4
1201 Genève
Phone : +41 22 535 3993
E-mail : info@starqube.com
Web : www.starqube.com