





Buy-side usage of Level 3 data analytics for algorithmic performance

WBR Insights & BMLL Technologies White Paper

Contents

Key Findings	3
Executive Summary	4
Chapter One: The State of Play	5
Chapter Two: Data and Analytics Generation	7
Chapter Three: The Way Forward	12
Conclusion	15
Methodology	16
About BMLL Technologies & WBR Insights	17

Contributors



Paul Humphrey CEO BMLL Technologies



Elliot Banks CPO BMLL Technologies



William L'Heveder COO BMLL Technologies

Our Typical Respondent

A Futures and Equities focused data scientist from a US based company with \$10-50bn in AUM, whose fund strategy was typically 'long only fundamental' with a holding period of 1-2 days for an average sized position.

Key Findings

82% of our respondents stated that between 20 - 75% of their organization's AUM is managed quantitatively.

54% of firms using in-house quantitative research in Equities are deploying these technologies to generate alpha (21%) and calibrate risk parameters (33%).

High quality data has become a commonly utilized commodity by most market participants - 74% of respondents say they use Level 3 data in their research program.

Of the respondents not using Level 3 data in their research programs – nearly 75% said the main reason was due to current vendors not providing that level of data.

Over 80% of respondents said they are already are or very likely to embrace cloud for their data and analytics generation and processing over the next 12 - 18 months.

64% of respondents said that at least 50% of their investment in new data and analytics capabilities will be from buying-in these capabilities.

41% of respondents saying that they will increase their budget allocations significantly for third-party data as a key element of their quantitative research.

The need for a secure service is a primary concern for many participants, when using a third-party data science platform for their alpha-generation and model back-testing.

Buy-side usage of Level 3 data analytics for algorithmic performance - Executive Summary

The brave new world

Global capital markets have always created huge quantities of data however much of this data was historically discarded as the tools to handle, sort and triage that data weren't available. The advent of electronic markets in the mid-nineties brought with it a new frontier, exponentially increasing the quantum of data being captured and stored by trading firms and venues. This created a new problem for many firms, how to leverage the latest computing technologies to make sense of this data and outperform the market. Utilizing the vast data to generate alpha has historically been both labor and capitally intensive, requiring specialist super computers, physical libraries of data sheets and teams of data scientists and data engineers.

Today, data and analytics play an increasingly important role in the workflows of all capital market participants. Supported by technological advances made in big data management and machine learning/ AI, powered by ever cheaper cloud, financial firms have created a new playing field of data and analytics. The rise in processing power has enabled more firms to start looking towards granular data to uncover hidden patterns in the data that can be used to predict the future. The more granular the data, the more powerful the prediction. Whereas historically, granular order by order (Level 3) data was the preserve of the most sophisticated hedge funds, it is now more widely available to all market participants thanks in part to efforts of exchanges & trading venues in supplying clean data, albeit in a raw unprocessed format and data distributors delivering that data to end clients.

Participants still face an assortment of challenges preventing them from utilizing this Level 3 data, including; lack of in-house resource to manipulate large data sets as well as the cost of acquiring and analyzing the data. Many participants have found that they have struggled to keep pace in the data and analytics race as their current vendors do not supply granular Level 3 market data. In some cases, when a vendor does supply the data its usability is contentious, requiring substantive effort to augment data quality to a usable level.

The hunt for alpha

The buy-side community is responding to the availability and predictability of this powerful new data set by investing heavily in resources, capabilities and partnerships to drive their never-ending quest for alpha. The next 12-18 months will be critical for all firms looking to embrace data and analytics to help drive increased risk-adjusted returns. Faced with increased pressures as well as competition from costeffective alternatives. It should come as no surprise that many investment firms are looking to partner with third parties to outsource data and analytics generation.

Having access to clean harmonized data is critical and when combined with a sophisticated analytics platform can not only enhance existing data infrastructures, it can offer firms a costeffective alternative to building, at great expense, a comprehensive data science capability in house. Firms that understand the benefit of partnering with a firm that can provide such data and analytics will be best place to leverage a more efficient quantitative trading division and generate more repeatable and scalable alpha. BMLL supports clients across every aspect of their quantitative market research and analysis with access to a harmonized Level 3 data lake, a data science platform, analytics feeds a suite of integrated visualization tools.

An Introduction to BMLL

BMLL is an award-winning data and analytics company operating at the cutting edge of capital markets. Our mission is mission is to unlock the predictive power of pricing data and offer clients the insight they need to better understand how markets behave and make more informed decisions.

As a cloud-native managed service with unlimited compute power, we deliver AI/ML driven analytics to clients' applications, either for internal use or to enhance their client-facing products. Clients access our platform and data lake via 3 applications they need to answer their toughest trading related problems; A data science lab connecting quants with the scalable research environment, a data feed API allowing direct plug-ins to the data and analytics lake and a data visualization layer which brings the data to life.

BMLL serves some of the most sophisticated clients in the capital market space. From banks and brokers to hedge funds and the buy-side firms, to exchanges and trading venues as well as data re-distributors and academic institutions. Insights are delivered via our three cost-effective and consumable mechanisms.



Chapter One The State of Play

The operating environment for investment management firms is seeing a sustained transformation. Traditional sources of differentiation in investment management are becoming increasingly commoditized. But AI is bringing new opportunities that extend far beyond cost reduction and efficiency of operations.

Big data is being harnessed to source and mine new sources of information in the quest for alpha generating ideas. User generated data is being put to work to measure investor sentiment which can be mapped to asset prices.

Alpha produced by the 'old', simpler, linear strategies has ebbed with the advent of more sophisticated strategies. Now, a strong frontier for investment is the creation of quant strategies that embrace data and analytics.

It is why buy-side firms are refining their quant strategies through AI to complement their portfolios. Why? Because they offer two coveted features conventional investments lack. These are diversification and specific risk factorbased returns.

Across-the-board, recognition that the value of data and analytics in financial markets is increasing. This is clear from our survey findings. The majority (68%) of our respondents said that between 20 - 50% of their organization's AUM is now managed quantitatively. A further 14% cited AUM as high as 50 – 75%.

The survey also highlights the extent to which big data analytics is changing the world of capital markets and global banking. Analytics is becoming a cutting edge industry tool due to the sheer momentum of data flowing through the market every day. Indeed all our survey respondents said that they rely 100% on real-time data feeds for equity trading.

Firms deploying new analytics platforms now have the wherewithal to calculate enterprise credit and market risk in minutes versus hours, achieving close to real-time transaction cost analysis. Importantly, they can also observe and be ahead of fraud patterns in near real time too, by introducing data sets and techniques previously not possible in the ongoing search for alpha.

Risk management is a field of quantitative analysis that has grown exponentially in demand and perceived importance since the financial crisis of 2008. So it's not surprising to see the priority that firms are placing upon it in 2020.

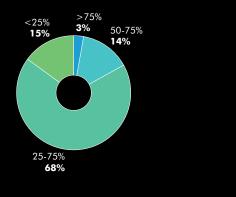
We saw 33% of our respondents stating that their organizations are currently undertaking quantitative research in-house in their risk parameters divisions. 21% (equally) said they are using it for alpha generation and performance benchmarking.

Front-office quants are growing in influence in a world that is witnessing fast shifts in market regimes. Role responsibilities are blurring too as the landscape of financial securities has become increasingly complex.

Quant analysts – who not only understand the complex mathematical models that price securities, but who are able to select the right data and parameters to generate profits and reduce risk – are increasingly indispensable assets.

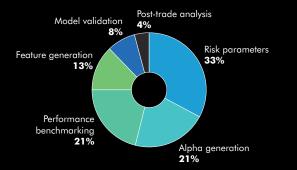
Nearly half (46%) of our respondents said that it was the quant analyst who was responsible for the integration and onboarding of new data sets and analytics within their organization. 68% of respondents said that between 20 - 50% of their organization's AUM is managed quantitatively.

Two thirds of the firms we surveyed have a balanced approach to managing quantitative AUM.



33% said that their organizations are currently undertaking their quantitative research in-house in their risk parameters division.

Which division are you currently undertaking quantitative research in-house?

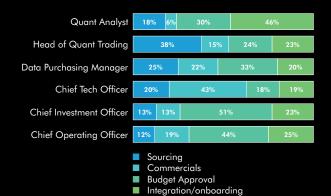


In an industry that is forever adapting and looking for ways to improve their models, it follows that 42% of respondents use Level 3 data for strategy backtesting (e.g., feature generation, alpha generation and model validation) which are all a core part of that process.



46% said that the Quant Analyst role holds the responsibility for the integration and onboarding of new data sets and analytics within their organization.

Which role within your organization is mostly responsible for each area of the onboarding process for new data sets/analytics?



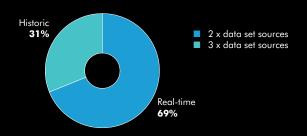
The data follows of the rise of the data purchasing manager, a role that 5 years ago barely existed in the industry. The rise of such a role and their importance in collecting huge data sets is testament to the advancement of the industry. Gone are the days of narrow data focus, now it's all about collecting, curating and consuming enormous lakes of harmonized granular data.



Paul Humphrey CEO, BMLL Technologies

100% of respondents are using real-time data sets for their equities trading.

How many data sources do you use for equity trading? (Real-time number of data sources)



Chapter Two Data and Analytics Generation

Buy-side firms are facing increasing client pressure to acquire larger number of data sets, many requiring customization. It's why today's firms are employing a growing range of published indexes. But as benchmark and reference data sources increase, so does complexity.

Data access and data centricity are key to results today. However, legacy infrastructure and insufficient in-house analytics resources mean this. Not all firms have been able to put in place the necessary in-house 'plumbing' to access this data wealth, let alone manipulate it.

In an ideal world, use of cleansed data would be as easy as turning on a tap. In reality, some firms are outsourcing their data analytics requirements. But they are then finding they can't bridge the gaps between their external service provider and their own internal resources.

As the granularity of data increases, so does firms' need for external buy-in. Our survey showed that 40% of respondents said they are generating 50% of their Level 1 order statistics to data analytics in-house. This compared to 100% buy-in of their order statistics Level 2 to generate data analytics.

The majority (74%) of our respondents, however, are now going beyond looking at patterns at the top of the order book. They are leveraging Level 3 data (all price levels of a limit order book with individual orders visible) in their research programs. The granularity of this data reveals all the individual messages in the limit order book. Providing traders and researches with a deep view across the workings of the market.

These datasets are large and complex and not all firms have the capabilities to derive meaningful analytics from them. But for those that do, risk management is again seen as another key motivator. Respondents commented, for example, that they "rely on tick-by-tick data to increase analytics accuracy for risk assessment". Fortunately, evolving technology and advanced analytics are enabling this. They are also de-biasing approaches to improve decision making and help risk managers make better choices.

Our survey gave a clear picture of firms being 'haves' and 'have-nots'. Of the 'have-nots' – the respondents not using Level 3 data in their research programs – nearly 75% said the main reason was its inaccessibility.

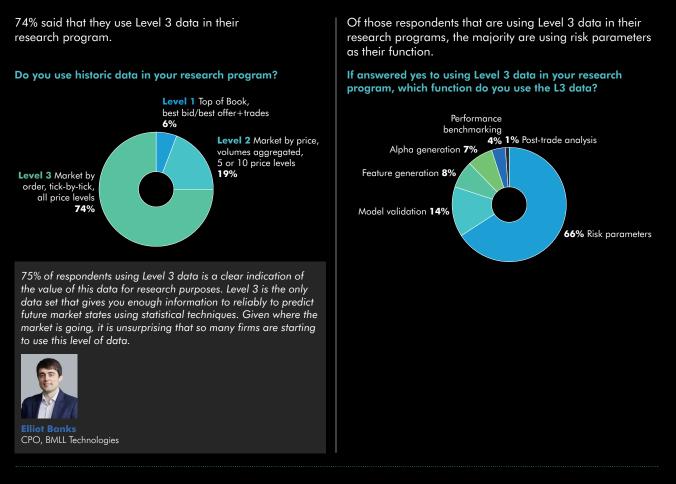
Their key hurdle to generating data analytics in-house was that their current market data vendors do not supply Level 3 data. Other constraints described were: concern about data cleanliness (18%) and cost (20%). But the main obstacles was lack of in-house resource to manipulate such large data sets (23%).

Further comments showed a fear of embracing big data and advanced analytics with a third party. On top of this was an unwillingness to move away from more traditional processes.

This 'no can-do' attitude was mainly reflected in concerns about in-house agility. Respondents referenced: "The infrastructure within which they would operate"; "Diversification takes a lot longer and we don't have that much time on our hands" and "Challenges arise when framework and systems aren't the same internally and externally".

99% said that their organizations were currently consuming third-party historic data and analytics. This was via traditional data vendors such as Refinitiv and Bloomberg Respondents' comments showed that a psychological barrier to innovation could be that their long-time use has now become ingrained in their firms' processes and legacy systems.

These included: "Our systems are more aligned to this data and so are our policies" and "They have been our primary source from quite some time now and [it] would be difficult to change them".



We asked our respondents to explain their answer further, here is what they told us:

"Our risk management capabilities are the best because of the level of analytics we follow."

"We rely highly on analytics to provide accurate risk assessment. Tick-by-tick data makes our analytics effectively accurate."

"This process needs this kind of data. Risk doesn't always carry a pattern and in order to respect that we need to keep on researching about different types of risk."

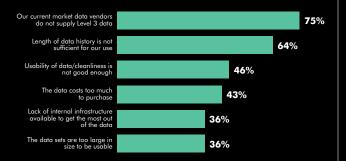
"It is complicated and needs more enforcement strategies than any other process. This makes data imperative for this process."

"We need this data to keep developing our strategies across the platforms that we use for alpha generation."

"Our analytical framework needs Level 3 data for certain customized concepts that we follow."

Of those respondents that aren't using Level 3 data in their research programs, they said the main reason for not doing so, is that their current market data vendors do not supply Level 3 data.

If you do not use Level 3, what are the key reasons for not using this data within your organization?



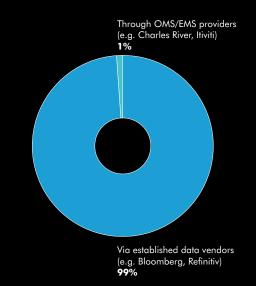
When you consider the size of the data sets and the infrastructure required to deliver these data sets it's unsurprising that the traditional vendors aren't able to supply this data. However having Level 3 alone isn't sufficient, the data needs to harmonized in a consistent information rich format and have sufficient history to make it useful.



COO, BMLL Technologies

99% said that their organizations are currently consuming third-party historic data and analytics via established data vendors.

If you currently consume 3rd party historic data and analytics, what is your primary source?



We asked our respondents to explain their answer further and here is what they told us:

"They have been our primary source from quite some time now and would be difficult to change them because of their continuous improvements."

"Their service and data solutions are magnificent and all that we need."

"Our systems are more aligned to this data and so are our policies."

"We can be assured of receiving quality data, which is key for us."

"It's the alignment of using this data for all our analytics requirements."

The surveyed firms prefer - and are also set up to - ingest traditional methods of data distribution rather than adopt nascent data marketplaces. This is driven by the trust traditional players have instilled in market participants, regardless of the often lower data quality. However, given the increasing reliance on alternative pricing data for alpha generation, I expect these firms will struggle to get the value they need out of existing data pipelines.



CEO, BMLL Technologies

23% said that their current market data vendors do not supply Level 3 data which is the key challenges their organizations are facing when generating data analytics in-house.



Most firms looking to generate analytics in house will first have a have a good source of Level 3 data, these data sets are large and sizable. Most other vendors simply capture their real-time feed without making any efforts at rebuilding the order book into Level 3.



We asked our respondents what key concerns their organizations have when outsourcing their data analytics generation to a third-party. Here is what they told us:

"Quality of our data does become the main concern."

"The level of security of data and operations."

"When we have an internal analytics team that is so strong we don't quite need external analytics."

"The immediate thing that comes to my mind is the fact that we cannot test accuracy of the data if it is from a third party."

"Safety is always the first and biggest concern."

"Over dependence on them in times of crisis."

"Our main concern is how much of the data is furnished and how much of it is true."

"Their skill set and costs are the biggest challenge. Both don't meet our expectation sometimes."

"Diversification takes a lot longer and we don't have that much time on our hands in certain situations."

"In today's business, data security is the biggest concern for almost every organization that uses it."

"Cost is usually a concern because these activities are not cheap."

"The infrastructure within which they would operate and whether or not they will be able to handle the tasks designated."

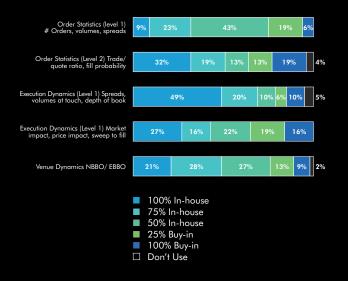
BMLL has spent a lot of time on our data processes and operations, from quality, scalability, security, resilience and cost as these are the most important factors for firms looking at outsourcing that function. The fact we have addressed these so comprehensively is what differentiates BMLL from other data vendors.





40% of respondents said they are generating 50% of their order statistics Level 1, orders, volumes and spreads to data analytics in-house vs 100% of their order statistics Level 2 to generate data analytics via buy-in.

Considering the analytics you use in your trading, which of the following analytics do you generate in-house vs. buy-in?



The main challenges our respondents organizations have are: reduction in quality of data from third party vendors, latency issues/difficulty due to lack of harmonisation/normalisation, little cost/benefit with third party vendors, and security issues.

"There comes a point when there is a reduction in quality."

"Data and analytics don't flow at the desired speed always."

"Missed inputs can cause challenges in the output and that could be disastrous."

"Their ideas, concepts and operations models don't align to ours. This puts everything off."

"Speed and accuracy are the biggest challenges here for us."

"Contracts cost way too much and don't provide better value when these activities are conducted internally."

"Challenges arise when framework and systems aren't the same internally and externally."

"Managing two different sets of analytical processes and teams isn't cost effective."

"Data security is something that cannot be trusted with any other organization apart from the one that we work with."

"Service customizations aren't always possible."

"Sometimes situations do get difficult because there are two organizations involved and the understanding does hit road blocks."

"There are more dead-ends than opportunities because of external compliance rules."

Firms focus are turning towards downstream integration and how firms match existing workflows with new providers of data and analytics in a way that protects key IP whilst scaling resources is a key concern for many players. We at BMLL have APIs & FTPs that integrate into clients workflows allowing firms to seamlessly combine a scalable research environment with their production environments.





Chapter Three The Way Forward

Regulation, client demands and electronic markets are transforming data management requirements at firms. The need to integrate functions such as risk and performance is a driver here. These factors are impacting how buy-side firms operate in the new world of data management and how they interact with vendors.

For many buy-side firms, the rise in data complexity is complicating operational efficiency.

Having skilled traders engaged in data management functions for alpha generation and model back testing is not efficient. Lack of alternative in-house capabilities exacerbates this. Few firms, likewise, have the storage or skills capacity to maintain and update their own data lakes.

Data enablement is essential to an agile asset management process. Firms today have two options to ensure they can access the right data, at the right time and in the right format.

They can buy ever more technology and increase headcount to manage growing volumes of data. Or they can opt for custom managed data services to help them with the heavy lifting.

Some firms are now turning to specialist vendors as partners to supply data and support the stream of their data workflow. This is borne out by our survey findings. 61% of respondents said they planned to increase their quantitative research budget allocations over the next 12-18 months for third-party data.

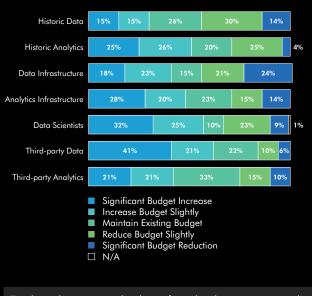
Outsourcing data management needs is the choice of most, with some companies opting for a hybrid approach. 64% of respondents said that at least 50% of their investment in new data and analytics capabilities will be from buying-in. It was clear too that there was recognition that for this to happen, only the Cloud can offer the critical storage and scalability. Over 80% of respondents said they were likely to or are already embracing Cloud for their data and analytics generation and processing over the next 12-18 months. The need "to secure and speed up the analytical process and diversify the data stream" was cited.

Regulators have made it clear that they understand the need for outsourcing certain functions to third parties, including data management. But that responsibility ultimately rests with the firm itself.

We asked respondents what it would take for them to use a third-party cloud data science platform for alpha generation and model back testing? It was plain that security and data protection were their primary concerns.

Comments included that the service provider "should have the safest data protection eco-system in the market". Another major consideration would be "how clear and defined [were] their cyber security measures". 60% said that they will increase their budget allocations significantly for third-party data and analytics as a key element of their quantitative research.

Over the next 12-18 months how will you change you budget allocations for key elements of quantitative research?

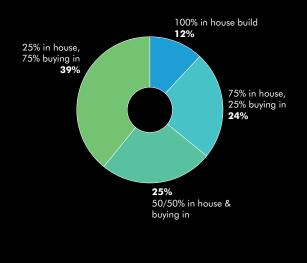


Two key takeaways can be drawn from this data - our surveyed funds are slowing investment in their own data infrastructure , potentially driven by major migration plans to the cloud, and investing heavily in third party data and analytics.



39% said that 25% of their investment in new data analytics capabilities will be in-house development and 75% would be from buying-in.

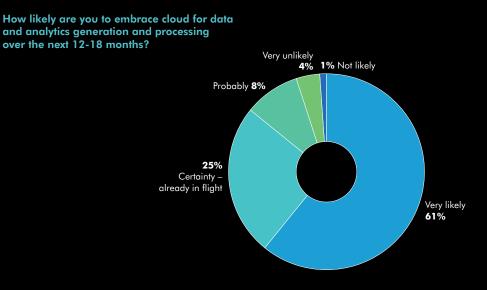
How much of your investment will be in house development vs. buying in new analytics capabilities?



The growing trend of partnering with 3rd party firms to scale their data and analytics operations rather than building in house is clearly visible here with only 12% of firms spending all their investment on in-house build. In many cases our clients don't want to be seen to be marking their own homework when it comes to validating models, especially for regulatory purposes.



61% of respondents said they are very likely to embrace cloud for their data and analytics generation and processing over the next 12-18 months.



We asked our respondents to explain their answer further Cloud will speed up and secure analytical process.

"This is something that will secure and speed up the analytical process and diversify the data stream."

"Cloud an help us improve our analytics and take it to the next level."

"The current system suits us quite effectively."

"The demand for cloud analytics has risen considerably in the past 12 months."

"We have other upgrades to be taken care of and cloud solutions will assist with those."

We asked our respondents what it would take for them to use a third-party cloud data science platform for their alpha generation and model back-testing. Here is what they told us:

"It could be one of the most difficult things to do given the current quality and talent we hold internally."

"It's difficult to understand our necessity of a third-party analyst. AS an organization we are quite selfsufficient in terms of analytics."

"How clear and defined are their cyber security measures because along with safety even intrusions are getting sophisticated."

"The service provider should have the safest data protection eco-system in the market."

"Per-organization policy alpha generation analytics is highly confidential and would never be associated with an external organization."

"Not a lot of firms would consider outsourcing alpha generation analytics."

"I would personally like to see more intelligence packed in the service, but I would rather not allow alpha generation analytics to leave the premises."

"A completely safe platform that could guarantee that no data leakages would be possible at any point in time."

"A disruptive solutions provider who has a legacy in the market with data security."

"A more efficient solution."

Conclusion

The investment management industry is, at heart, and always has been, a data processing industry. But the ways in which that data is now collected and analysed is being transformed.

Our survey shows the extent to which AI is now used to find more patterns and relationships between asset prices and data from other alternative data sources.

Its end result is that portfolio managers and researchers are more effective and better equipped to deliver predictive alphas, faster and at a lower price point than before. Speed, in the old sense, for asset management is over. No longer is it enough for investors to be slick in teasing trends and investible insights from traditional information sources of trading figures, market shares and economic updates.

Today's success relies on a different kind of speed (and scalablity) and a different way of doing things. And neither length of experience, brand prestige or even the quality of client relationships will insulate them from this.

The new speed means demonstrating value to existing and future clients by keeping pace with digital opportunity. But adapting to data-driven decision-making brings challenges. Our survey highlighted many of these.

Among them is the glaring need for data science talent, new business processes, data governance and the organizational muscle to capture real value from analytics. Predictability only comes from a deep understanding of how the market behaves. The quest for alternative sources of alpha generation 'twas ever thus... This fact has not changed – the route to it has.

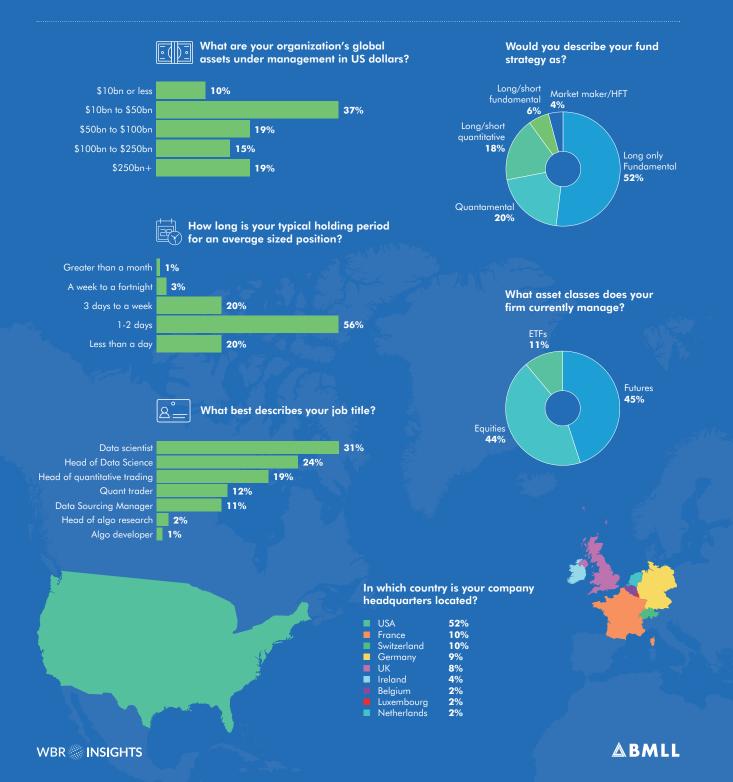


Methodology

In Q4 of 2020, WBR Insights surveyed 100 Heads of Data, Chief Data Officers, Data Scientists and similar from across medium-sized buy side firms in the EMEA, and North American regions to find out how capital markets participants use predictive data and analytics in their daily trading decisions.

The report aims to gain a greater understanding of how hedge funds and asset managers are challenged by back-testing alternative data and their ability to derive meaningful insights from new alternative data, such as Level 3 data, which in the past five years has only become widely available.

The survey was conducted by appointment over the telephone. The results were compiled and anonymized by WBR Insights and are presented here with analysis and commentary from BMLL's contributors.



About BMLL Technologies & WBR Insights



BMLL is an award-winning historic data and analytics company operating at the cutting edge of capital markets. Our mission is to unlock the predictive power of pricing data and offer clients the insight they need to better understand how markets behave and make more informed decisions.

As a cloud-native managed service with unlimited compute power, we deliver AI/ML driven analytics to clients' applications, either for internal use or to enhance their client-facing products. Clients access our platform and historic data lake via 3 applications they need to answer their toughest trading related problems; A data science lab connecting quants with the scalable research environment, a historic data feed API allowing direct plugins to the data and analytics lake and a data visualisation layer which brings the data to life

BMLL serves some of the most sophisticated clients in the capital market space. From banks and brokers to hedge funds and the buy-side firms, to exchanges and trading venues as well as data redistributors and academic institutions. Insights are delivered via our three cost-effective and consumable mechanisms.

For more information, please visit our website, bmlltech.com or visit our Twitter @bmlltech.

WBR INSIGHTS

We use research-based content to drive conversations, share insights and deliver results. Connect with our audience of high-level decision-makers in Europe and Asia from industries including: Retail & eCommerce, Supply Chain & Procurement, Finance, as well as many more. From whitepapers focused on your priorities, to benchmarking reports, infographics and webinars, we can help you to inform and educate your readers and reach your marketing goal at the same time.

Contact us to find out how your business could benefit from:

- Year-round access to our network of decision-makers and industry-leaders
- In-depth research on current fast-moving issues and future trends
- Lead generation campaigns that fit your priorities
- Promoting your organization as an authority in your industry

To find out more, visit: **wbr.co.uk**

WBR INSIGHTS

© BMLL Technologies & WBR Insights 2021. All rights reserved. This document and its content are proprietary to BMLL Technologies & WBR Insights and may not be reproduced, republished or resold. The information contained within is provided on an "AS IS" basis for information purposes only and BMLL Technologies make no warranties of any kind including in relation to the content or suitability. The name BMLL Technologies, their logo and any associated brand names are all trademarks of BMLL Technologies.