

Why is using data and data science both interesting and important within financial markets?

Annabelle Shrieves

Executive Director
Data, Valuations & Analytics
S&P Global Market Intelligence

Using data as part of the investment management process is far from new – but why is using data and data science both interesting and important within financial markets? And why is using alternative data so valuable?

We explored these topics at our recent [S&P Global Quantitative Investment Management Forum](#) in London.

Firstly, how do we define alternative data? Is it defined by asset class, or geography – anything outside of Equities, or the US? We observe for example the use of our proprietary capital flow data and systematic credit strategies rising. Is alternative data defined by public vs private markets, or by time-frequency – such as tick data? Does *reliability* mark the boundary between traditional and alternative data sets? For example, if a truly unique data set disappears there is nothing to replace it, whereas traditional data could be defined as having multiple sources. Is alternative data defined as data used within a financial market context but created for or originating from a different purpose? Or is alternative data defined by human and emotional bias – data so differentiated it is deemed to incur more risk, to generate alpha?

Ultimately, data is a spectrum between traditional and alternative, it is not binary – and importantly, and as straightforward as it may seem, data conveys *information*. Here are the thoughts of our panel on why and how best to leverage data and to optimize data-driven strategies:

- Data, and alternative data, provides a better connection with the real economy - a more effective ability to calculate and price risk, and to reconcile the price-formation process with the changing state of the economy.

- Data science is now required at every level of the value chain – from data collection, alpha signal extraction and portfolio construction, to trade/strategy execution and risk management. Data science at scale is now a requirement to gain a competitive advantage, especially when moving into new/alternative asset classes.
- As such - collect as much data as possible, from as close to the source as possible, understand its specificities and remove the data bias for each desired usage. Noise removal is imperative to understand the true signals offered by the data.
- However, data must not be blindly injected into systems. Data must be used with intentionality and purpose to fulfil a specific information gap. This approach will prove more additive over time than purely focusing on returns.
- Combining data sets is more powerful than using data in isolation – but contextualization is important. Combining market data + traditional data + alternative data and overlaying this with domain expertise, will enable you to build a complete picture of the assets you are interested in.
- Mapping is key – alternative data needs to be mapped to become meaningful, and mapping between datasets and between equity and fixed income assets remains a significant challenge. Efficient and accurate mapping of data relationships is essential to extract more signal from the ‘noise’, rather than building a picture manually. Building more complex data relationships will provide an information advantage. Our S&P Global Cross Reference suite allows seamless linking of reference data by company, security, and industry to better manage data integration and minimize manual processes:
[Cross Reference Services](#).
- Cross-Asset signals: To further overcome the challenge of mapping between asset classes as part of a data-driven cross-asset strategy, S&P Global's [Alpha Signals](#) factor library provides a suite of credit-derived equity factors, our [Bond-Linked Equity Signals](#), built on top of our proprietary [CDS Pricing Dataset](#) and [Bond Pricing Data](#), using a robust mapping algorithm. The quantitative feed provides a daily view of how credit markets impact equities, and these

“Data must be used with intentionality and purpose to fulfil a specific information gap”

“We help customers make the connection between credit and equity markets”

indicators have proven to provide unique information that has low commonality with both fundamental and alternative datasets. Our innovation has enabled us to create alpha-generative signals originated from our mapping capabilities, helping our customers make the connection between credit and equity markets.

Furthermore, our quantitative analysts have extended this research to combine our [Bond-Linked Equity Factors with Equity Short Interest](#) for an enhanced security selection process in large-cap equities in the US and Developed Europe.

- Returning to the thematic of traditional vs alternative data, there are still significant insights to be gained from traditional data, which remains the foundation of investment management, as demonstrated by our pioneering [S&P Capital IQ Financials Dataset](#), point in time [Compustat® Financials Dataset](#) and comprehensive [S&P Capital IQ Estimates Dataset](#).
- Moreover, our product innovation enables us to provide new *alternative* insights for our customers by leveraging our *traditional* data sets, for example with our forthcoming [S&P Global Company Connections: Detailed Estimates](#) product based on sell-side analyst estimates. Sell-side analyst coverage data provides a new and rich source of establishing connections between firms, as analysts (given their industry expertise) are likely to cover fundamentally related firms. Company to company connections are derived from the number of shared sell-side analysts between a pair of companies. As our panel highlights, enhanced data signals can be generated from the linking of data through company relationships.
- The extraction of textual and tonal data by machine learning is an important step forward, and with increased focus by the industry on LLMs, the importance of textual data is significantly growing. Text is not only used for model-building, but is also increasingly used for better timing advantage in execution algorithms and for improved risk management, as well as to enhance alpha generation.
- Text is also a key area of growth for our own S&P Global products: most unstructured data is in textual form from sources such as emails, transcripts, articles and documents. These text files are usually difficult, time-consuming and expensive to analyze and utilize. S&P Global identifies primary sources of textual information

“Our innovation provides new, alternative insights from our traditional data sets.”

that can be parsed and structured for ease of use, bypassing the entire process of sourcing, cleansing and maintaining the data, while enabling metadata tagging and linking to other datasets such as financials and estimates. Our [Textual Data Suite](#) includes machine-readable transcripts, filings and broker research, as well as our [Textual Data Analytics: Sentiment Scores & Behavioral Metrics Dataset](#).

- Combining our cross-asset and textual themes, new research from our Quantamental Research team combines our Textual Data Analytics & Credit Default Swap Pricing datasets to examine the effect of earnings call sentiment derived from our earnings call transcripts on CDS spreads: [Watch Your Language: Executives' Remarks on Earnings Calls Impact CDS Spreads](#)

In conclusion, data is core to the investment management process – increasingly so with LLMs becoming more mainstream - and alternative data is becoming more necessary. However, more data and bigger models are not necessarily better – the key is to understand what you are asking for, and to keep the intention and application of the data in focus.

And as for what defines traditional vs alternative data? In a world of AI, nothing will replace *human* understanding of data, and this is what ultimately unites all data.

For more information on how to access these data sets, please contact the sales team at: h-ihsm-global-equitysalesspecialists@spglobal.com or visit www.marketplace.spglobal.com.

Contact Us

Europe, Middle East & Africa

+44 (0) 134 432 8300

The Americas

+1 800 447 2273

Asia Pacific

+60 4 291 3600

Japan

+81 3 6262 1887

Copyright © 2023 by S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.

These materials, including any software, data, processing technology, index data, ratings, credit-related analysis, research, model, software or other application or output described herein, or any part thereof (collectively the "Property") constitute the proprietary and confidential information of S&P Global Market Intelligence or its affiliates (each and together "S&P Global") and/or its third-party provider licensors. S&P Global on behalf of itself and its third-party licensors reserves all rights in and to the Property. These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable.

Any copying, reproduction, reverse-engineering, modification, distribution, transmission or disclosure of the Property, in any form or by any means, is strictly prohibited without the prior written consent of S&P Global. The Property shall not be used for any unauthorized or unlawful purposes. S&P Global Market Intelligence's opinions, statements, estimates, projections, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security, and there is no obligation on S&P Global Market Intelligence to update the foregoing or any other element of the Property. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. The Property and its composition and content are subject to change without notice.

THE PROPERTY IS PROVIDED ON AN "AS IS" BASIS. NEITHER S&P GLOBAL NOR ANY THIRD PARTY PROVIDERS (TOGETHER, "S&P GLOBAL PARTIES") MAKE ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE PROPERTY'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE PROPERTY WILL OPERATE IN ANY SOFTWARE OR HARDWARE CONFIGURATION, NOR ANY WARRANTIES, EXPRESS OR IMPLIED, AS TO ITS ACCURACY, AVAILABILITY, COMPLETENESS OR TIMELINESS, OR TO THE RESULTS TO BE OBTAINED FROM THE USE OF THE PROPERTY. S&P GLOBAL PARTIES SHALL NOT IN ANY WAY BE LIABLE TO ANY RECIPIENT FOR ANY INACCURACIES, ERRORS OR OMISSIONS REGARDLESS OF THE CAUSE. Without limiting the foregoing, S&P Global Parties shall have no liability whatsoever to any recipient, whether in contract, in tort (including negligence), under warranty, under statute or otherwise, in respect of any loss or damage suffered by any recipient as a result of or in connection with the Property, or any course of action determined, by it or any third party, whether or not based on or relating to the Property. In no event shall S&P Global be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees or losses (including without limitation lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Property even if advised of the possibility of such damages. The Property should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions.

The S&P Global logo is a registered trademark of S&P Global, and the trademarks of S&P Global used within this document or materials are protected by international laws. Any other names may be trademarks of their respective owners.

The inclusion of a link to an external website by S&P Global should not be understood to be an endorsement of that website or the website's owners (or their products/services). S&P Global is not responsible for either the content or output of external websites. S&P Global keeps certain activities of its divisions separate from each other in order to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain nonpublic information received in connection with each analytical process. S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global Ratings' public ratings and analyses are made available on its sites, www.spglobal.com/ratings (free of charge) and www.capitaliq.com (subscription), and may be distributed through other means, including via S&P Global publications and third party redistributors.