Quantifying Strategy
BUILDING DATA STRATEGIES FOR THE NEXT DECADE
Quantifying Strategy
A DEEP DIVE INTO DATA STRATEGY IMPLEMENTATION

Building Data Strategies for the Next Decade

In June 2017, we released *Quantifying Intuition: Mapping the Data Science Landscape at Hedge Funds*, classifying hedge funds along a spectrum of building data science efforts.

The universe largely fell into four buckets: i) Early Adopters, ii) Early Mainstream, iii) Mainstream and iv) Nascent Adopters.

On the back of *Quantifying Intuition*, we received numerous requests to conduct a deeper dive into how firms are building out data strategies in practice. While “data” is perhaps one of the biggest buzzwords of 2017, specific information about data-related hiring, budgeting, structuring, expenses and allocation has been more limited. The Jefferies Data Science Working Group (listed below) has spent the last 18 months meeting with nearly 100 investment managers about their data strategies, priorities, infrastructure and staffing. Here, we focus on a sampling of 45 firms actively developing data strategies for the next decade.

Quantifying Strategy is a practical deep dive into how investment firms are addressing the explosive growth of data. In July and August 2017, Jefferies conducted nearly 50 in-depth interviews with investment managers, lawyers, accountants and data providers to get a more holistic and specific understanding of the emerging data ecosystem. Quantifying Strategy details the current range of approaches to institutionalizing data strategies, and explores initial efforts to build long term solutions.

The majority of firms we spoke to fall into the Early Mainstream bucket outlined in *Quantifying Intuition*, as these firms are among the most active in building and developing data efforts, and the vast majority were discretionary managers. Each has a Data Lead – a person for whom data is a part of their focus, though not necessarily full time – and are tasked with growing data efforts. But these organizations are in different stages of build vis-à-vis each other – and vis-a-vis their own stated goals. This diversity allowed for a broad discussion of what works, what doesn’t, and what some of the biggest challenges are for the long term.

There is broad agreement that the information available to shape investment, research and business decisions has changed – and it’s important to build platforms to address this. Funds of all sizes, strategies and maturities were nearly unanimous in reporting that their organizations are in the “earlier stages” of what they anticipate an end state data strategy will look like. While we can debate the future magnitude of data adoption by discretionary managers – what’s clear is that the game has changed, and data will be a part of the investment process going forward.

As such, the timing of our study allowed us to take a snapshot of this dynamic and rapidly changing corner of investment management – much like the early days of digitization. We hope *Quantifying Strategy* helps inform strategic business decision making as we collectively prepare for our industry’s next chapter.

CONTACTS
Shannon Murphy
+1 (212) 336-1139
shannon.murphy@jefferies.com

Dan Furstenberg
+1 (212) 323-3336
dfurstenberg@jefferies.com

Richard Carbone
+1 (212) 336-7472
rcarbone1@jefferies.com
EXECUTIVE SUMMARY: DATA AT A GLANCE

A deep dive into how investment managers are building data efforts could not be more timely. The convergence of various factors including the explosion of available data, the evolution of advanced statistical processes for mining and understanding data, and the growth of available infrastructure to store this data, means that data strategy is now a C-suite issue.

Quantifying Strategy puts structure around the diverse ways managers are institutionalizing their efforts. While there simply is no one size fits all approach (or even most), a number of things stand out from our survey.

BY THE NUMBERS

80%  Percentage of firms who consider data “critical” to their business in the next 5 years

42%  Percentage of firms with a Dedicated Data Lead

18  Number of different organizational titles held by those leading data initiatives

80%  Percentage of firms who leverage data to enhance risk management and their investment mosaic

37%  Percentage of firms who leverage data for idea generation

40%  Percentage of firms leveraging 30 or more discrete alternative data sets*

15%  Percentage of firms who include data and analytics in their broker vote

70%  Percentage of firms who have passed on a using a data set because of compliance or regulatory concerns

50%  Percentage of firms who have felt frustrated with data vendors and contract structures

15%  Percentage of firms who currently leverage artificial intelligence

10%  Percentage of firms who see data resources as a strategic tool for recruiting and retaining top talent

0  Number of firms who have the same data strategy

*We asked about discrete data sets, not the number of vendors firms work with
LOOKING AHEAD

We are in the early stages of investment managers’ building out institutional data strategies. Even large organizations with dedicated headcount and resources consider themselves in the beginning of a multi-year secular shift. We believe that much like this year, 2018 will be a period of continued transition and growth for data adoption among investment managers - and we expect that by 2019, the landscape will have been redefined for data adoption within organizations.

Right now, building out a data strategy is a differentiator among investment managers – but is still more a luxury item – as firms are at various levels of exploration, adoption and marketing this competency. Allocators, too, are more frequently asking investment managers about their data strategies. We believe by 2019, both of these themes will have converged, as formal adoption of data strategies becomes an integral part of investment managers’ investment processes, and allocators focus on a complete and explicit understanding of data strategies (and make this a formal part of the due diligence process). We also expect that those who do not have a clear and cogent data strategy by 2019 may run the risk of appearing as an outlier among peers.

Top Predictions:

Ninety percent of the available data in the world has been created in the last two years – investment managers will increasingly have to dedicate resources to source, digest and disseminate this available information.¹ We wanted to highlight what we believe will be some fundamental shifts on the horizon for investment managers to properly address the explosion of available information.

¹ IBM
BEYOND THE BUZZWORD

Data has become one of the buzzwords of 2017 – and many articles have chronicled the use of data across different industries. We will not retread this argument here, because when organizations as diverse as Coca-Cola, BP, Formula One racing and even The National Enquirer have embedded data efforts, a clear tipping point has been reached.²

The Harvard Business Review reflected on this move from back office to C-suite issue in May, “More than ever, the ability to manage torrents of data is critical to a company’s success… Data was once critical to only a few back-office processes such as payroll and accounting. Today it is central to any business, and the importance of managing it strategically is only growing.”³

While many of the headlines around data and investment management have focused on the hunt for frontier or pre-commercial data sets or the hiring of marquee Dedicated Data Leads, the reality is many of our managers are more focused on: i) articulating clear strategy, process and structure around common data use, ii) hiring the right individuals to manage and execute these strategies, and iii) building flexible infrastructure to digest data in the coming decade.

I. BUT DATA (ITSELF) IS NOT A STRATEGY

Of the 45 managers we surveyed who report currently having a data effort, 80% said one of their top priorities was to institutionalize the initiative broadly across the organization – whether further embedding across investment teams, improving onboarding and workflow processes or scaling nascent efforts. Only 18% said their current main focus is on discovering pre-commercial or frontier data sets.

CHART 1

Top Goals for Data Efforts in Next 12 – 18 Months

<table>
<thead>
<tr>
<th>Goal Description</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improve data processes, including: sourcing, onboarding and ingestion</td>
<td>80%</td>
</tr>
<tr>
<td>Enhance education and culture related to data usage</td>
<td>62%</td>
</tr>
<tr>
<td>Source frontier or pre-commercial data sets</td>
<td>18%</td>
</tr>
</tbody>
</table>

Source: Jefferies Prime Services

Research paradigms continuously evolve to incorporate new resources and approaches – new data and analytics are just the newest tools to follow channel checks, proprietary models or professional expert networks, in being added to the investment process. A small but growing number of firms we surveyed even report adding ‘data and analytics’ considerations to their broker votes.

Strategy matters. You cannot get where you’re going if you don’t know what your destination is. Which is why it’s important for any organization implementing a data strategy to identify specific goals from the start. A growing trend we see among investment managers is the formation of data committees or working groups to understand and optimize the process. These groups may include founders, CIOs, CROs, PMs, traders, or tech analysts, or, in larger organizations, comprise a broader group across multiple verticals, asset classes or products.

Starting on a new path without knowing where you want to end up can lead to distractions along the way. The right strategy is what allows businesses to evolve, adapt and grow without becoming victim to mission creep – in this case, spending millions of dollars on data sets, only to have subscriptions languish from lack of use or inability to derive accretive information from them. Without proper buy-in and vision from senior stakeholders, data initiatives are likely to alienate or confuse other employees, create obstacles for compliance, or be met with skepticism in the broader organization. Having a Data Lead alone is not enough – constant education among stakeholders is critical, and often success is a result of senior stakeholders making data initiatives an explicit priority with incentives for adoption.

Among discretionary investment managers, it’s still relatively rare for current data efforts to drive idea generation. Eventually, perhaps it will, but today, only one-third cited idea generation as a current objective of data initiatives; whereas more than 85% cited enhancement of pre-existing investment processes as a goal.

**CHART 2**

Top Objectives for Acquiring Data Sets

<table>
<thead>
<tr>
<th>Objective</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhance Investment Processes</td>
<td>98%</td>
</tr>
<tr>
<td>Enhance Risk Management and Audit</td>
<td>78%</td>
</tr>
<tr>
<td>Facilitate Idea Generation</td>
<td>37%</td>
</tr>
<tr>
<td>Optimize Business Processes</td>
<td>33%</td>
</tr>
<tr>
<td>Other</td>
<td>26%</td>
</tr>
</tbody>
</table>

Source: Jefferies Prime Services
II. DATA LEADERSHIP COMES IN MANY FORMS

One of the most interesting findings of our discussions focused on who has emerged to lead data efforts at different organizations – whether from hiring externally or leveraging internal expertise. Of the 45 firms we focused on, Data Leads held nearly 20 different organizational titles. For the purposes of this paper, “Data Lead,” is the person who holds primary ownership of data efforts at their firm.* Data Leads may be bucketed into two groups: Dedicated Data Leads (those for whom data strategy is their sole responsibility) and Shared Data Leads (those who have multiple additional responsibilities, ranging across verticals).

In both categories, these individuals’ titles, where they sit in their organizations and to whom they report vary widely. This variation suggests how nascent and situational data strategies are among investment managers. Though most headlines focus on Chief Data Scientist hires (often from Silicon Valley), this is the exception rather than the rule, and in reality, most organizations initially look within and subsequently expand and specialize from there.

In fact, slightly more than 50% of funds we spoke with have a Shared Data Lead that – at least currently – assumes the role alongside other responsibilities. Reporting lines vary widely, and in some cases, teams have split or dual reporting lines into both investment and non-investment professionals. About two-thirds of the time, Data Leads are currently reporting into non-investment professionals, whether in Risk (often the CRO), the C-suite, Operations or Finance.

CHART 3
Dedicated vs. Shared Data Lead

Source: Jefferies Prime Services

CHART 4
For Those with a Shared Data Lead...Intent to Hire Dedicated Headcount in Next 12 Months

Of those who currently don’t have a Dedicated Data Lead, more than one-third are actively trying to hire one. One of the most frequent questions we get from those looking to hire an initial Dedicated Data Lead, is whether their primary expertise should come from technical skills (computer science and programming fluency) or finance and investing, and whether the individual should be a senior hire, or more mid-level and grown from within. There is no right answer. Given the idiosyncratic nature of investment managers, it is situationally dependent on the fund, based on factors including firm strategy, size, maturity, organizational structure and investment style.

*Glossary available in Appendix II
TOOL FOR TALENT

More than one organization noted that the build out of data efforts serves more than the current purposes of enhancing investment, risk or business processes. In fact, firms across the size, strategy and maturity spectrum see these initiatives as strongly accretive to long term talent attraction and retention. A number of firms reported seeing a build out of enhanced data capabilities as an additional tool for recruiting top talent and staying competitive as a destination for cutting edge investment research, process and culture.

How do organizations structurally approach data headcount and team building? Given the diversity of organizational titles of these leads, it is extremely variable.

At this stage, responsibilities associated with data efforts are fairly fungible and amorphous, as firms continue to iterate and focus on specific objectives. Top questions emerging as organizations continue to build their teams include:

- Where should these individuals sit in our organization? And to whom should they report?
- How do we incentivize cultural adoption across the firm?
- What is the best way to approach quantifying a return on investment for these efforts?
- Should we hire advanced statistical analyses (machine learning/artificial intelligence) specialists?
- At what point would a cloud specialist be accretive?
- Should our long term build out include data specialists by vertical or product (i.e. – satellite/geolocation vs. transaction vs. sensor or IoT data) or sector of investment?

Source: Jefferies Prime Services
III. BUILDING FOUNDATIONS


Few of the words firms used to describe their own data strategies convey mature development. Nearly two-thirds of firms surveyed feel they are in the “first half” of building and deploying their data strategies. Indeed, even the small handful of firms with systematic strategies that we interviewed reported feeling they “still had quite a ways to go” with regards to taking advantage of the full slate of technologies, analytics and data sets they feel will be regularly available in the years to come.

Despite concern that many investment managers are leveraging hundreds of data sets, nearly 80% of those we surveyed leverage less than 100. Since no two firms have the same data strategy, how these discrete sets are being used varies widely.

The range of number of alternative sets currently leveraged varies widely across our sample set, from 5 to more than 200.

**CHART 7**

Approximate Number of Alternative Data Sets Leveraged by Investment Managers

For those leveraging alternative data sets, numerous issues arise, including:

- How to source rare, frontier, pre-commercial or bespoke data sets
- How to vet and streamline these sources from a compliance perspective – and evidence the processes
- How to analyze the value of exclusivity or semi-exclusivity
- How to efficiently and effectively source, onboard, clean, digest and disseminate multiple data sets
- How to negotiate with vendors for terms of data subscriptions that better match manager needs rather than more common structures and tenors
- How to distribute data internally across the organization

A small but growing number of respondents reported interest in sourcing frontier or pre-commercial data sets. Some investment managers may work with third parties to identify and unearth information that isn’t yet being broadly disseminated. But challenges remain, as some respondents noted that in some cases, data owners don’t necessarily understand the value of their data, and need assistance in delivering useful information sets. These data sets may require additional vetting from a compliance perspective, but may be of higher value since they aren’t broadly commercially available.
There is an additional misperception from some managers with nascent data strategies that they “can’t keep up with those who have multi-million dollar data budgets, or find data sets that have any alpha left.” One of the strongest themes that emerged from our discussions with organizations that are actively building data efforts is that no two data strategies are alike, and no two organizations are looking at data in the same way for the same purpose.

Data isn’t being used monolithically, and firms are utilizing the same data sets for often entirely different purposes. Firms rarely (if ever) express the same idea, in the same size, for the same duration, at identical entry and exit points. There are infinite ways to layer and combine data – and we anecdotally heard that a number of firms are leveraging traditional data in its raw form to combine with other traditional or alternative data sets. It is increasingly likely that multiple data sets will be leveraged to build different mosaics – rather than the perception that data acquisition is solely a race to unearth new data sets to be consumed on a standalone basis. A number of firms spoke to us about their efforts to build systems that allow for the overlay of multiple data sets for a more thorough understanding of a topic.

Cost & Allocation

With regards to individual data set costs, current spend is clustered far more tightly in the mid-five figures than often thought. In fact, the majority of firms surveyed estimated that their spend for individual data sets ranges from $20,000 – 75,000 per set but with considerable variability between subscriptions. Data set costs range from de minimus, if the data set is pre-commercial and the firm could be working with the data provider to better gauge its usefulness, to seven figures, used by a small number of firms.

At this point, most firms are splitting the costs of third party data sources between the management company and the fund. 13% allocate at the management company level, 35% allocate the cost to the fund level and 26% split between the two. Another 26% were unsure or unwilling to share allocation structure for data efforts.

CHART 8
Where Data Costs Are Allocated

<table>
<thead>
<tr>
<th>Allocation</th>
<th>0%</th>
<th>5%</th>
<th>10%</th>
<th>15%</th>
<th>20%</th>
<th>25%</th>
<th>30%</th>
<th>35%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund Expense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35%</td>
<td></td>
</tr>
<tr>
<td>Management Co. &amp; Fund</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26%</td>
<td></td>
</tr>
<tr>
<td>Doesn’t Know/Would Not Respond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26%</td>
<td></td>
</tr>
<tr>
<td>Management Company</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Jefferies Prime Services

A small – but growing - number of firms report having created “R&D” budgets specifically focused on building their data strategies, internalizing these costs at the management company level.
Infrastructure

As firms continue to build out data strategies, a critical question centers on how much storage will be needed for these efforts. Twenty-eight percent of managers currently leverage strictly on prem (on premises) solutions, more than half (52%) have embraced some type of cloud based solutions, with 24% strictly utilizing cloud solutions and 28% leveraging both on prem and hybrid, and a small number (around 10%) are currently transitioning from on prem to cloud-based or a combination of the two.

Though storage and infrastructure are often less discussed than sourcing a new or “hot” alternative data set, storage becomes increasingly critical as more firms bring more data in house. One respondent referred to their cloud infrastructure specialist as the “unsung hero” of their data strategy.

CHART 9

Infrastructure Approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>On Prem only</td>
<td>28%</td>
</tr>
<tr>
<td>On Prem and Cloud</td>
<td>28%</td>
</tr>
<tr>
<td>Cloud only</td>
<td>24%</td>
</tr>
<tr>
<td>Shifting from On Prem to Cloud</td>
<td>12%</td>
</tr>
<tr>
<td>Did not respond</td>
<td>8%</td>
</tr>
</tbody>
</table>

Source: Jefferies Prime Services

Advanced Statistical Processing: Machine Learning, Artificial Intelligence, Natural Language Processing

A small percentage of firms report pursuing machine learning in house as a core competency. Slightly less than 10% of discretionary firms we spoke to report conducting any machine learning or artificial intelligence in house to assist (or drive) data efforts right now. A very small number report currently trying to actively hire individuals with this as a core competency in the next 12 months.

Given the explosive rise in amount of available data, it’s likely investment managers will need advanced statistical processing to keep up and process it all. However, at this point, firms also need to balance the additional costs associated with artificial intelligence, including hiring talent with the appropriate skills, and investing in the computing and storage needed. As with most other technologies, the costs will likely contract, but for now, a number of firms reported the time and expense associated with artificial intelligence as too high to implement.
IV. THE ROAD AHEAD

Available data continues its exponential growth march – with estimates of 44 zettabytes (or 44 trillion gigabytes) of data points generated annually by 2020. This compares with just 4.4 zettabytes that the digital universe generated in 2013 – a 10-fold growth in just seven years.

Our firms report the following issues in focus for the next 12 – 24 months as they continue to develop data efforts:

- **An intense focus on education and culture**: nearly two-thirds report culture as a top 2 issue in more effectively leveraging data science across their firms. Many Data Leads report additional buy-in from stakeholders across the firm as one of the most critical issues in determining the ultimate success or failure of their data efforts.
- **Advanced Statistical Processing**: ~15% report wanting to bring machine learning capabilities in house in the future to more rapidly and effectively process large amounts of data.
- Working with **third parties to source pre-commercial data sets**; in some cases collaborating to determine viability and best use.
- **Better alignment** of terms of data contracts – more than 50% report being frustrated with data vendors, seeing a disconnect between often rigid subscription terms versus the shifting timeline of positioning and investments. Data Leads report a misalignment of matching duration of interest of various data sets with the tenor of a data contract.
- **Staying current** on legal and regulatory issues, especially across jurisdictions – 70% report having passed on potential data sets because of legal or compliance questions.

How Jefferies Can Help

The scope of the data science universe is enormous, and growing – further complicating things is that there simply is not a one size fits all approach to adopting these solutions effectively. The idiosyncratic nature of hedge funds and their investment processes means that leaders, PMs, analysts, traders and risk managers need to carefully consider the right approach for their organizations.

Jefferies has centralized its data science expertise, resources and solutions to better help clients approach this issue strategically, to maximize the return on your invested time covering the issue. Please do not hesitate to let us know how we can be of further assistance as you learn more about the ever-changing landscape of data science, integrate alternative data into your investment process, or if we may follow up on anything discussed herein.

The most common question from hedge funds that have not begun exploring data science is: “Where do I even start?”

We offered the following in *Quantifying Intuition* as a basic five-step framework for starting to understand and leverage data science at your organization and felt it bore repeating here (please see Appendix I).

---

APPENDIX I

Basic Initial Framework to Understand and Incorporate Data Science into Hedge Funds

<table>
<thead>
<tr>
<th>Identify</th>
<th>What do I want to get out of this exercise?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Enhance investment processes</td>
</tr>
<tr>
<td></td>
<td>• Improve risk management approach</td>
</tr>
<tr>
<td></td>
<td>• Streamline business processes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Determine</th>
<th>What kind of resources are needed?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Appoint a data science expert or competency?</td>
</tr>
<tr>
<td></td>
<td>• Would an OQIO (Outsourced Quant Information Officer) be accretive?</td>
</tr>
<tr>
<td></td>
<td>• How much are we willing to spend on a go-forward basis?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Explore</th>
<th>Start with Big Data; broaden to ML</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Explore what big/non-traditional data sets exist for your strategy</td>
</tr>
<tr>
<td></td>
<td>• Determine whether you want structured or unstructured data</td>
</tr>
<tr>
<td></td>
<td>• Explore machine learning solutions for down the road</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Educate</th>
<th>Employees about benefits and drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Take the time to build consensus internally about implementation</td>
</tr>
</tbody>
</table>

| Implement | Strategically incorporate or build solutions into processes |

Source: Jefferies Prime Services

Questions to help build a focused and effective data strategy

**Identify**

*What do I hope to get out of data science? Why am I even exploring different solutions for my firm?*

- Improving investment theses. Expanding inputs or ideas. Enhancing risk management. Identifying unintended bets. Running a more efficient organization. Staying current with industry themes.

- Is it possible that effective implementation of data science could attract incremental investors, new talent, or new enthusiasm about the research and development (R&D) happening at our firm?

**Determine**

*What expertise can I leverage in better understanding the data science landscape? What resources am I willing to dedicate to these efforts?*

- Not everyone needs to hire a Dedicated Data Lead. Expertise comes in many forms and may be sourced from current employees with an interest in the topic (typically a CTO or head of risk), third party consultants, or even experts at counterparties.

- In some cases, an “OQIO” (Outsourced Quant Information Officers) might be of help. Just as Outsourced Chief Investment Officers (OCIOs) have become a solution for some asset owners to manage their investment portfolios, it’s possible we may increasingly hear of OQIOs serving as consultants to assist fundamental managers in understanding, selecting and managing the data solutions that make the most sense for their organizations.

- Should I attend any events hosted by data science providers or my counterparties that focus on these issues?
What is my anticipated spend for product, system and implementation? Is it a recurring or one-time cost – and what is my appetite for that spend growing on a go-forward basis?

- Data science solution costs cross a broad spectrum; beyond recurring subscription costs, one time implementation or project costs can add up. It’s critical to budget both base and bull cases for bringing these innovations on board, but costs can scale quite linearly.
- Where will these allocations live in our budget? Under IT? Under research costs? Do we want to syndicate the costs across investment teams (if organizationally appropriate)?

Explore

What types of alternative data solutions exist? How should I be thinking about machine learning down the line – if at all?

- There are already an endless – and growing – number of alternative data brokers. It’s important to understand which will best serve the goals of your firm and which are least likely to reflect data decay in the short term (i.e. – that the sets get so commoditized, their signaling benefits are minimized)
- Identify the data sets, sectors or themes that are most efficient and effective for your processes
- Start to learn about machine learning and other methods of digesting these big data sets for a basic understanding of these capabilities, even if you don’t incorporate them in the short term

Educate

There is often a cultural component to trying to incorporate data science into our organization

- John Maynard Keynes said, “Difficulty lies not so much in developing new ideas as in escaping from old ones.” PMs, analysts, traders and risk managers from across the organization need to buy in and fully understand both the opportunities and pitfalls of bringing new solutions on board. As with most cultural issues, data science needs to be embraced at the top. No group can advance to efficiently and effectively onboard these solutions without buy in from the principal. Implementation may be vertical (the principal buys in, and pushes the issue down across the organization), or horizontally shared across the C-suite, potentially including the CTO, COO, Head of Risk or Trading or senior technology analysts.
- Sometimes a few training or information sessions are required to fully educate employees and minimize disruption or conflict as new solutions are brought on.

Implement

- Roll out solutions on a strategic timeline, allowing sufficient space for incorporating and iterating as needed.
- Education of these solutions is now an ongoing effort; PMs, analysts or risk managers should work to stay abreast of emerging innovations.

Other Considerations

Do I really need to spend time on this, or is it just another fad?

- The exponential growth of available data sets and processing solutions has meant that many people don’t even know what they don’t know. The universe is sufficiently vast and novel that at a minimum, it’s worth exploring what might be of use to enhance pre-existing approaches or improve productivity.

Could new data science solutions become part of your marketing or investor relations approach?

- We are aware of a number of managers who have started incorporating more quantitative analyses in their pitches, tear sheets and marketing materials to show how they’re leveraging new data sets or machine learning solutions. They’re expanding the number of firm employees who meet with potential allocators to include data scientists or quantitative analysts, where appropriate. This is currently a differentiated approach; but it’s likely by 2019, it will be common. There is considerable opportunity in turning the black box into a grey box and helping inform investors how new quantamental approaches are supplementing more traditional processes

What are the legal implications of incorporating data science into my firm?

- There are numerous legal and regulatory questions that arise for these innovations, including: prohibitions around data scraping (is your data “clean?”), privacy laws, trespassing, misappropriation, and confidentiality. At this point, there is little clarity among many of these issues, and little standardization across jurisdictions.
APPENDIX II Glossary

In *Quantifying Intuition*, we defined some of the terms used, given the sometimes confusing or oversimplified ways in which they are used. It is worth doing the same here.

**Data Lead:** the person who holds primary responsibility for a firm’s data strategy – may be dedicated or shared.

**Dedicated Data Lead:** those for whom data strategy is their sole responsibility.

**Shared Data Lead:** those who have multiple responsibilities in addition to managing the firm’s data strategy, ranging across verticals.

**Data Set:** any collection of information. Examples range from number of consumers who walk into a Wal-mart the day after Thanksgiving (or Thanksgiving dinner!) to search terms entered into Google to earnings reports from various sectors.

**Big Data:** a data set so large traditional processing methods have a hard time digesting it – typically a terabyte or larger (a terabyte is one thousand bytes…which is equivalent to 1,000 copies of Encyclopedia Britannica).

**Alternative Data:** information gathered from non-traditional sources, such as: mapping the most popular run routes in the U.S. via fitness trackers or leveraging satellite data to measure mining activity in real-time. Not all Big Data is Alternative, and not all Alternative Data is Big.

**Exclusive/Semi-exclusive:** a somewhat amorphous topic in the still nascent data industry, but exclusivity or semi-exclusive typically connotes the idea that the distribution of a certain data set is limited to a sole or select group of firms.

**Pre-commercial:** a data set that may be sold down the road, but whose owner is still in the early stages of determining scope and viability of the set as a commercial asset.

**Structured Data:** data sets that have gone through processes resulting in a higher degree of organization that its original/raw state.

**Unstructured/Raw Data:** data that has not yet gone through processes resulting in a higher degree of organization than its original state.

**Data Exhaust:** data that is generated as a byproduct of other actions like online activity or business transactions.

**Web Scraping:** a technique whereby information is pulled from websites (with the assumption the website’s terms of use permit this).
IMPORTANT DISCLAIMER

THIS MESSAGE CONTAINS INSUFFICIENT INFORMATION TO MAKE AN INVESTMENT DECISION.

Please contact your Jefferies representative for copies of the most recent research reports on individual companies.

This is not a product of Jefferies' Research Department, and it should not be regarded as research or a research report. This material is a product of Jefferies Equity Sales and Trading department, and intended for Institutional Use. Unless otherwise specifically stated, any views or opinions expressed herein are solely those of the individual author and may differ from the views and opinions expressed by the Firm's Research Department or other departments or divisions of the Firm and its affiliates. Clients should assume that this material is not independent of the Firm’s proprietary interests or the author's interests. For example: (i) Jefferies may trade for its own account or make markets in the securities referenced in this communication (and such trading may be entered into in advance of this communication); (ii) Jefferies may engage in securities transactions that are contrary to or inconsistent with this communication and may have long or short positions in such securities; and (iii) the author of this communication may have a financial interest in the referenced securities.

The information and any opinions contained herein are as of the date of this material and the Firm does not undertake any obligation to update them. All market prices, data and other information are not warranted as to the completeness or accuracy and are subject to change without notice. Past performance is not indicative of future results, and no representation or warranty, express or implied, is made regarding future performance. The Firm is not providing investment advice through this material. This material does not take into account individual client circumstances, objectives, or needs and is not intended as a recommendation to particular clients. Securities, financial instruments, products or strategies mentioned in this material may not be suitable for all investors. Jefferies does not provide tax advice. As such, any information contained in Equity Sales and Trading department communications relating to tax matters were neither written nor intended by Jefferies to be used for tax reporting purposes. Recipients should seek tax advice based on their particular circumstances from an independent tax advisor. In reaching a determination as to the appropriateness of any proposed transaction or strategy, clients should undertake a thorough independent review of the legal, regulatory, credit, accounting and economic consequences of such transaction in relation to their particular circumstances and make their own independent decisions.

OPTIONS ARE NOT SUITABLE FOR ALL INVESTORS. Please ensure that you have read and understand the current options risk disclosure document before entering into any option transaction. The options disclosure document can be accessed at the following web address: http://optionsclearing.com/publications/risks/riskchap1.jsp. © 2017 Jefferies LLC