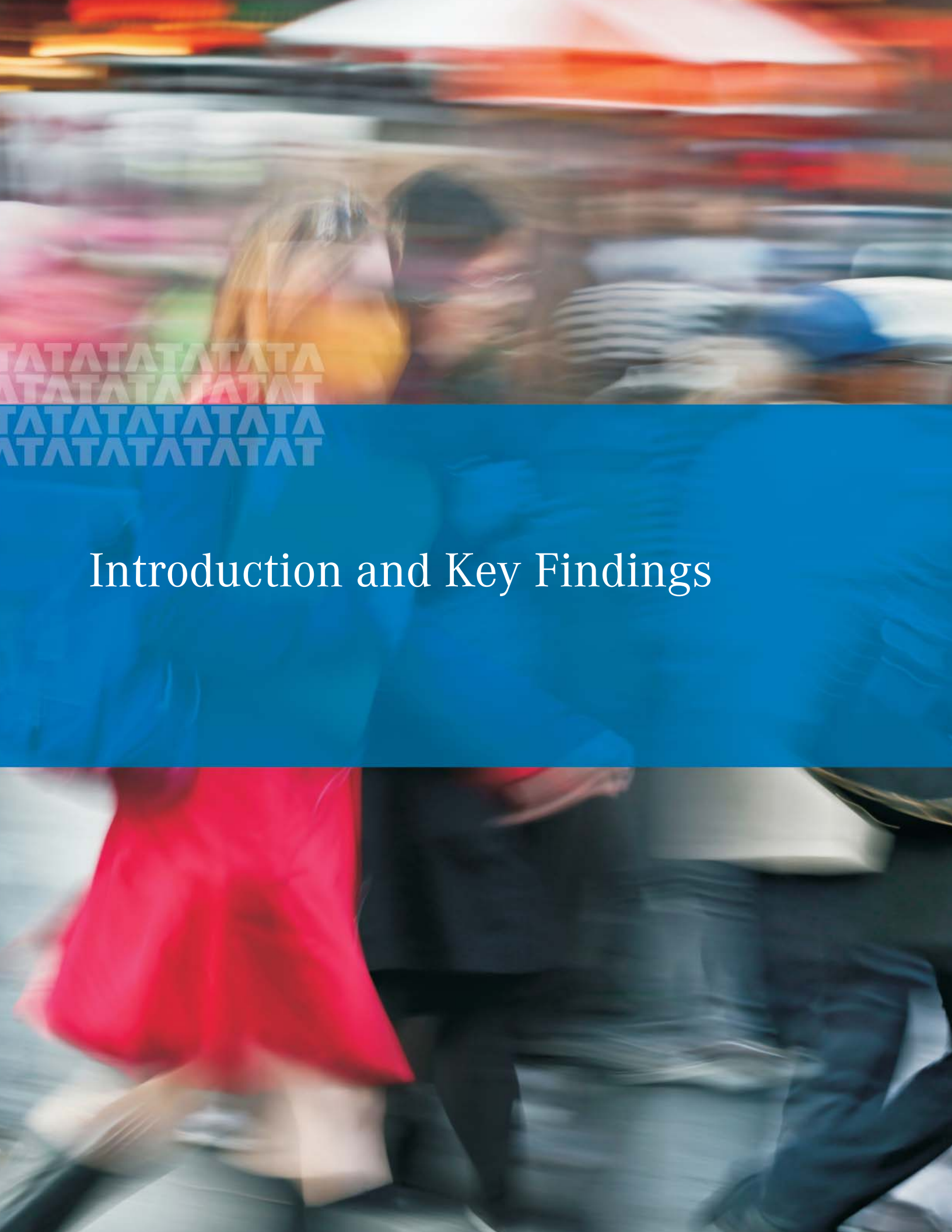
TATA CONSULTANCY SERVICES
Experience certainty.

The Emerging Big Returns on Big Data

A TCS 2013 Global Trend Study


643 COMPANIES
12 INDUSTRIES
83% > $ 1B REVENUE
IN-DEPTH INTERVIEWS
EXPERT OPINIONS
A TCS 2013
GLOBAL TREND STUDY

Contents



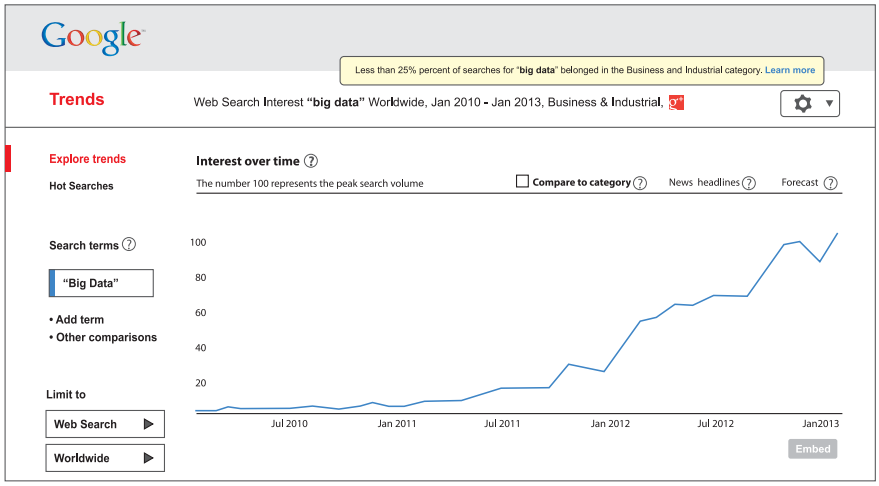
Introduction and Key Findings

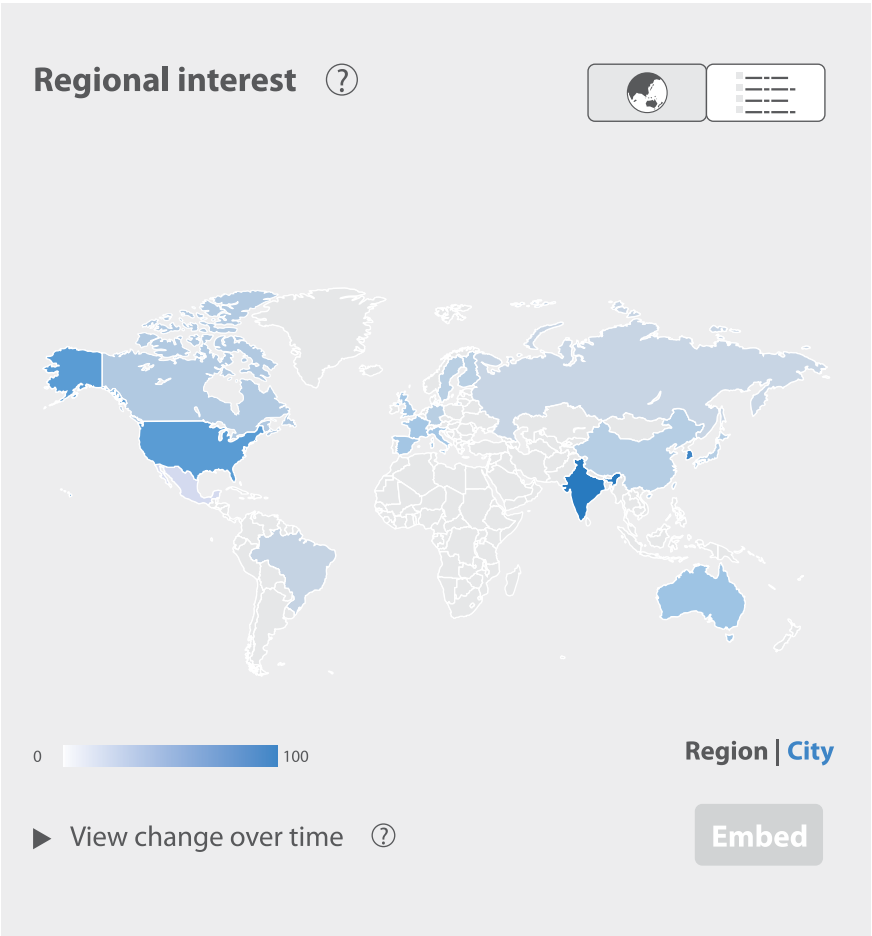
The Spotlight Continues to Shine on Big Data

Big Data has become big news almost overnight, and there are no signs that interest is waning. In fact, several indicators suggest executive attention will climb even higher. Over the last three years, few business topics have been mentioned in the media and researched as extensively as Big Data. Hundreds of articles have appeared in the general business press (for example, *Forbes, Fortune, Bloomberg BusinessWeek, The Wall Street Journal, The Economist*), technology publications and industry journals, and more seem to be written by the day. A March 2013 search on Amazon.com surfaces more than 250 books, articles and e-books on the topic, most of them published in the last three years.

Dozens of studies have been conducted on Big Data as well, and every week another one appears. Most of the big consulting firms and IT services companies have weighed in, as well as (of course) the technology research community: Gartner, Forrester, IDC and many of the rest. And the number of times the seven letters 'Big Data' have been clicked into a Google search has exploded in the last three years. (See Exhibit I-1) The term saw relatively little usage in online searches in the first half of 2010. And this Google Trend information shows 'Big Data' was of particularly keen interest in certain countries (Exhibit I-2): India, South Korea, the U.S., Australia, Canada, Western Europe, and Brazil.

Exhibit I-1: A Billowing Number of Big Data Online Searches





Late in 2012, we launched our own study on Big Data. We designed it to shed insights on six core issues, ones on which we felt the marketplace was looking for greater clarity:

1 How much are companies investing in Big Data, and what kinds of returns are they achieving on their spending?

2 What are companies in 12 industries doing with Big Data? That is, in which business functions and specific activities are they focusing their investments?

3 What kinds of digitized data are they finding to be the most important?

4 How are they organizing the professionals who process and analyze Big Data (e.g., embedded in business functions, in a central analytics group, etc.), and what are the pros and cons of those reporting relationships?

5 What are the biggest challenges of turning Big Data into insights that enable the company to make far better and faster decisions?

6 What is the current state of the technology, and where is it going?

Our 10 Key Findings

TCS surveyed 1,217 companies in nine countries in four regions of the world (U.S., Europe, Asia-Pacific and Latin America) in late December 2012 and January 2013. Of these companies, a little more than half (643) said they had undertaken Big Data initiatives in 2012. We also conducted in-depth interviews with more than a dozen executives across industries about their Big Data initiatives between December 2012 and March 2013. In addition, we interviewed two experts in the fast-evolving technologies of Big Data. This data, as well as our consultants' growing experience in helping large companies leverage Big Data, provide the basis for the findings in this report.

While our findings are numerous, we believe the following 10 are the most important ones:

- **About half of the firms surveyed are using Big Data, and many of them projected big returns for 2012.** 53% of the 1,217 firms surveyed had undertaken Big Data initiatives in 2012, and of those 643 companies, 43% predicted a return on investment (ROI) of more than 25%. About a quarter (24%) either had a negative return or didn't know what the return was. (See Section II)

- **There's a polarity in spending on Big Data, with a minority of companies spending massive amounts and a larger number spending very little.** Some 15% of the surveyed companies with Big Data initiatives spent at least $100 million each on them in 2012, and 7% invested at least $500 million. In contrast, nearly one-quarter (24%) spent less than $2.5 million apiece. This has resulted in a big spread between median ($10 million) and mean spending per company ($88 million). Industries spending the most are telecommunications, travel-related, high tech, and banking; life sciences, retail, and energy/resources companies spend the least. (See Section II)

- **Investments are geared toward generating and maintaining revenue**. 55% of the spending goes to four business functions that generate and maintain revenue: sales (15.2%), marketing (15.0%), customer service (13.3%) and R&D/new product development (11.3%). Less than half that amount (24%) goes to three non-revenue-producing functions: IT (11.1%), finance (7.7%), and HR (5.0%). (See Section II)

- **The business functions expecting the greatest ROI on Big Data are not the ones you may think.** Although sales and marketing garner the largest shares (a combined 30.2%) of the Big Data budget, the logistics and finance functions (which together get only 14.4% of the budget) expected much greater ROI on their Big Data investments. Furthermore, when asked to rate 75 activities in eight business functions on their potential to benefit from Big Data, companies around the world ranked just as many logistics activities as they did sales activities in the top 25. (See Section II)

- **The biggest challenges to getting business value from Big Data are as much cultural as they are technological.** When asked to rate a list of 16 challenges, companies placed an organizational challenge at the top of the list: getting business units to share information across organizational silos. A close second was a technological issue: dealing with what has become known as the three 'V's' of Big Data: data volume, velocity and variety. The third challenge was determining which data to use for different business decisions. (See Section II)

- **Nearly half the data (49%) is unstructured or semi-structured, while 51% is structured. The heavy use of unstructured data is remarkable given that just a few years ago it was nearly zero in most companies.** On another dimension of comparison, about 70% of the data is from internal sources rather than external. However, using external and unstructured data has outsized impacts. Companies that expect much bigger ROI on Big Data use more external and unstructured data than do companies expecting lower or no ROI. (See Section II)

- **The companies with the biggest projected 2012 returns on Big Data saw those returns coming from places that the laggards don't value as much.** To use a gold miner's analogy, the leaders pan for gold in different places – most of all in marketing, sales and service. The two activities where leaders see much greater potential than laggards are: improving customers' *offline* experience and marketing to consumers based on their physical location. ROI leaders also see much greater potential than do laggards in using Big Data to size and structure sales territories. And in customer service, leaders envision greater potential benefits in monitoring product usage to detect manufacturing and design problems. (See Section IV)

- **Companies that do more business on the Internet spend much more on Big Data and project greater ROI**. Companies that generate more than 75% of their revenue over the Internet spend about six times more on Big Data than do companies whose Internet business is 25% or less of total revenue. These Internet-centric companies also projected an ROI on Big Data (88%) that was nearly three times that of the less Internet-centric companies. Furthermore, the depth of the behavioral data that Internet-centric companies gather on their online customers gives them proprietary insights for developing superior new products and services, as companies such as Procter & Gamble Co. and Netflix Inc. have found. (See Section II)

- **Monitoring how customers use their products to detect product and design flaws is seen as a critical application for Big Data**, especially by heavy manufacturing companies whose customers depend on their products. (See Section IV)

- **Organizing a core unit of Big Data analysts in a separate function appears to be important to success.** Companies that expected the highest ROI on Big Data in 2012 are more likely to have a separate department of professionals who process and analyze Big Data than are companies expecting the least ROI (or no ROI). (See Section VII)

In the sections that follow, we explore these findings as we discuss how the survey results compared across regions of the world, by global industries and by business function. Lastly, we discuss the implications of our research and provide advice for companies that want to get more out of Big Data and need to know where and how to begin.

We base our prescriptions in Section VII on two sources: our analysis of what leading companies at the Big Data game (those with the greatest ROI) are doing differently than the rest, and insights from TCS consultants who are helping our clients capitalize on Big Data.

Note: *Graphs in this report representing percentage figures may not add up to 100% due to rounding off.*

The U.S. is Ahead on Big Data: Comparing Results Across Four Regions of the World

Highlights:

- U.S. companies are in the lead in using Big Data
- Median per-company spending on Big Data in 2012 was $10 million
- 80% of the companies say they've improved business decisions as a result
- Most of their data is structured (versus non-structured) and from internal sources (versus external)
- Only 27% of the companies sell their digital data, but those that did generated an average of $22 million from it in 2012
- About 24% of the companies either expected a negative return or haven't tried to quantify one for 2012
- Companies that projected an ROI averaged a 46% ROI, which includes those projecting a negative return
- The most difficult challenge in generating benefits from Big Data is organizational: getting organizational silos to share their data

Majority of the Companies Have Big Data Initiatives

Across all four regions of the world that we surveyed, 53% of the 1,217 companies said they had undertaken at least one Big Data initiative in 2012. The U.S. was the leader among regions in Big Data use. Of the four regions surveyed, this region had the highest percentage of companies that reported at least one Big Data initiative in 2012: 68%. Only a third of the U.S. companies said they didn't have even one Big Data initiative. (See Exhibit II-1)

Of the 53% of companies with Big Data initiatives in 2012 (643 in all), the Asia-Pacific region had the lowest percentage (particularly Australia and Japan) at 39%. A slight minority (45%) of European countries said they undertook Big Data initiatives. And about half of the Latin American survey respondents were in this game. A country-by-country analysis revealed a number of differences. A high percentage of Indian (70%) and Mexican (68%) companies reported Big Data initiatives. (See Exhibit II-2)

Exhibit II-1: North American Companies are More Likely to Have Big Data Initiatives

Q6_1 : Percentage of Companies (by Region) With Big Data Initiatives in 2012

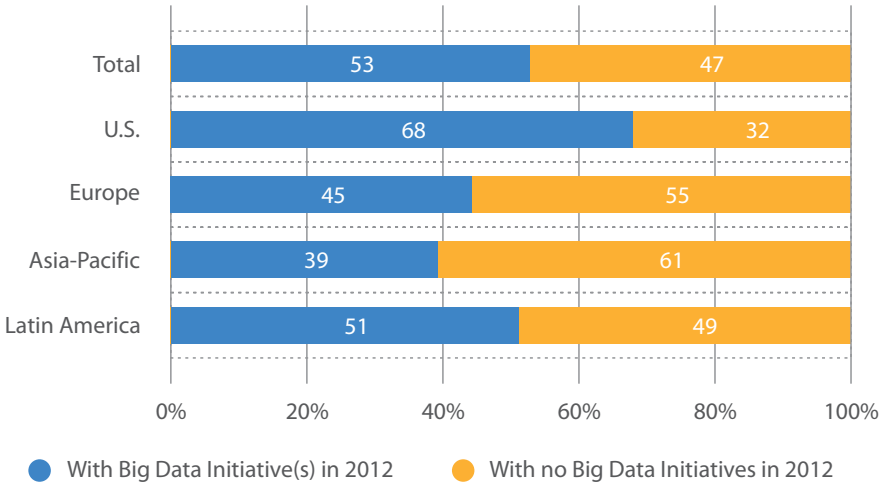
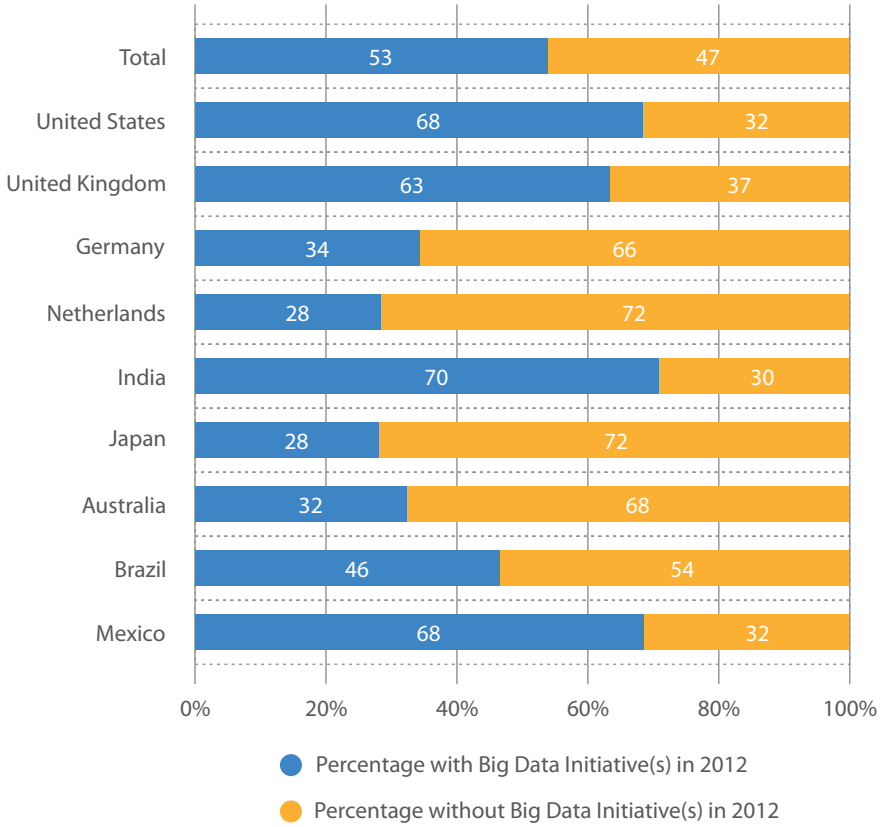


Exhibit II-2: US, India, UK and Mexican Companies are More Likely to Have Big Data Initiatives

Q6_1: Percentage of Companies by Country with Big Data Initiatives in 2012



They're Spending a Lot on Big Data

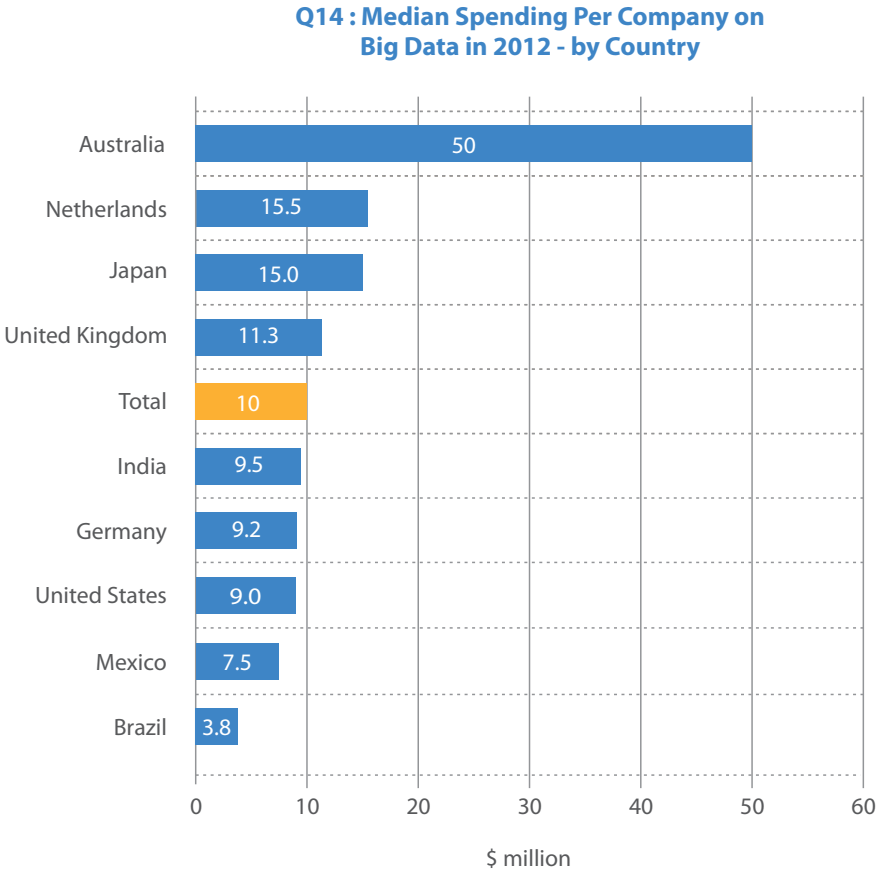
The investments these companies made in Big Data were sizable. We measure those investments in two ways: by the median and the average survey respondent:

- Median spending on Big Data was $10 million, which was 0.14% of revenue (based on median revenue of survey respondents: $6.9 billion). We believe the median spending numbers provide a more accurate picture of spending on Big Data than the mean (or average) numbers here since the mean was skewed because of a number of respondents (7% of the ones we asked for spending data) who spent more than $500 million on Big Data in 2012.
- The average survey respondent spending on Big Data was $88 million in 2012 , which was 0.5% of average revenue (of $19 billion). Again, we believe this is a less reliable indicator of what companies are spending on Big Data.

The large spread between median and mean spending on Big Data shows a big polarity in investment. In 2012, 15% of the companies invested at least $100 million apiece on Big Data initiatives. About half of them (7%) invested at least $500 million each. However, on the other end of the spectrum were the 24% of companies that spent relatively little on Big Data - less than $2.5 million each. This demonstrates that a distinct minority of companies – less than one in seven – have initiatives around Big Data and are investing heavily in it.

Using the median numbers, Australian companies were far in the lead, at $50 million in spending per company while US companies were close to the median ($9 million). (See Exhibit II-3.)

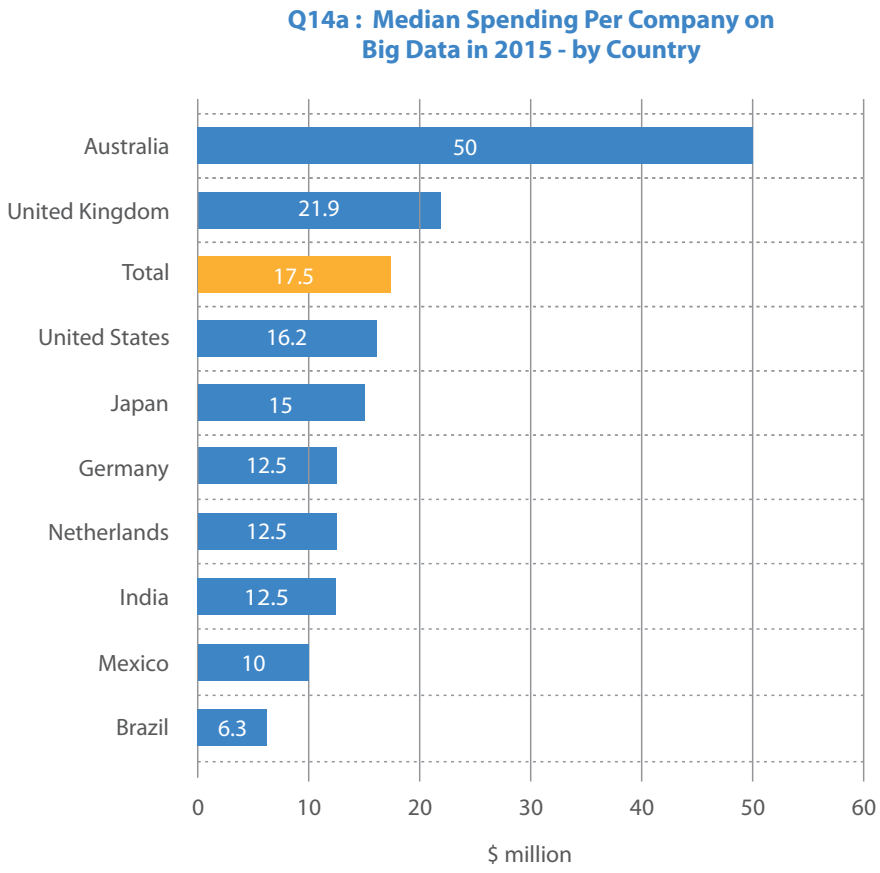
Exhibit II-3: Median Per-Company Big Data Spending in 2012 by Region



We frequently came across stories of outsized spending on Big Data in our interviews with leading companies. General Electric announced in 2012 that it would spend $1 billion over four years on a new San Ramon, California-based software and analytics center. One major telecommunications company is spending "tens of millions of dollars" on Big Data, especially in its mobile service unit. "Our customers get new phones every couple of years, and it's a highly competitive business with [lots of churn]," one executive told us. "The transactional data for those customers is a massive amount of information." Another company spoke about its 200-person centralized group of Big Data analysts. Although the company did not reveal specific numbers, figuring an average annual salary of $90,000 (a number from a firm called SiSense), we believe the firm has at least an $18 million annual expense in wages alone.[1]

By the year 2015, companies across the surveyed regions expect to spend 75% more on Big Data, with Australia and U.K. companies projecting the highest spending per company. Median spending across all countries is projected to increase by 75% to $17.5 million. (Exhibit II-4)

Exhibit II-4: 2015 Projected Per-Company Spending on Big Data



1 The $90,000 salary for the average data scientist in the U.S. comes from a study by SiSense, which can be read here in a TechCrunch article: http://techcrunch.com/2012/08/12/why-the-search-for-the-mystical-data-scientist-should-not-be-a-feat-of-magic/

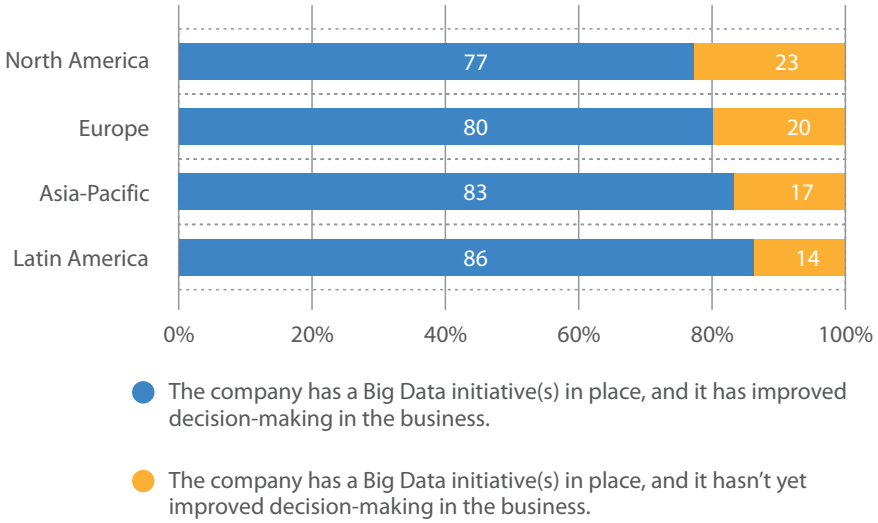
Big Data is Improving Decisions in Most Companies

Spending on Big Data, the technology that enables it and the people who analyze the information are, of course, not an end unto themselves. Slicing and dicing huge volumes and varieties of digital data can keep data scientists busy for days or even weeks. But if the insights they derive do not provide useful guidance – or if business managers don't utilize that guidance – all that spending is futile.

That is why we wanted to know whether Big Data was having beneficial impact. To find out, our survey first asked participants whether their initiatives had improved decision-making in the business. The answer for the clear majority – 80% -- was indeed yes. The lowest percentage of improvement in decision making was in the U.S. (77%) while the highest was in Latin America (86%). (See Exhibit II-5.)

Exhibit II-5: How Many Companies are Making Better Decisions?

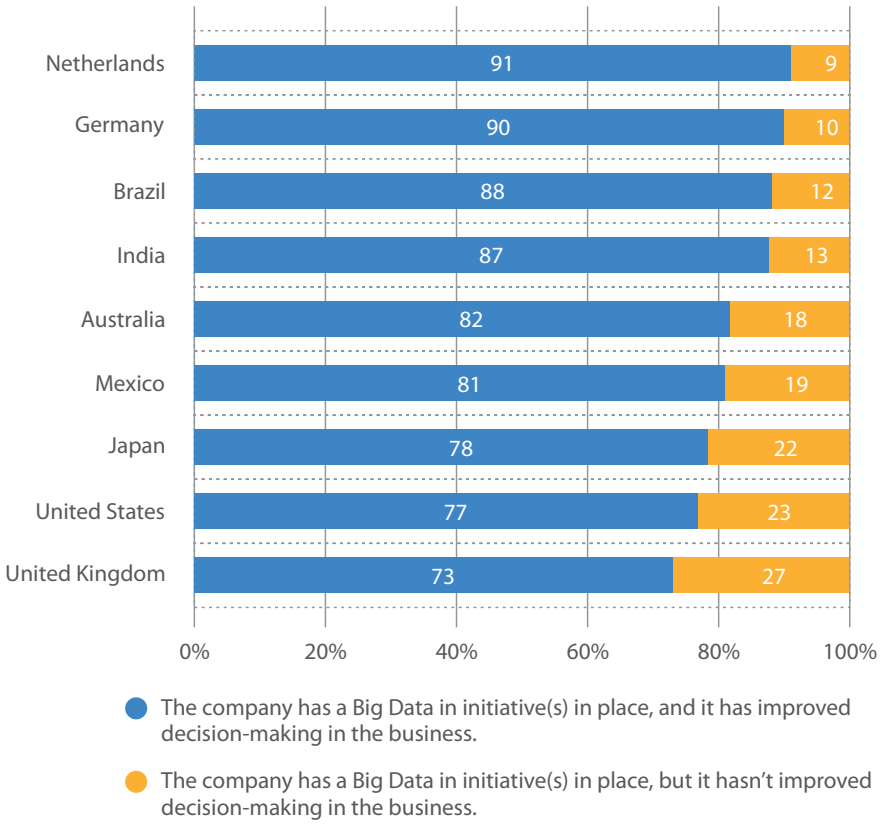
Q6 : Percentage of Companies Whose Big Data Initiatives Have Improved Decision-Making



On a country basis, U.K. companies reported a relatively lower percentage: 73%. Dutch companies reported the highest percentage: 91%. (See Exhibit II-6)

The countries with the highest percentage of Big Data initiatives were also the ones with the lowest percentage of improved decision-making. One interpretation: Companies that are newer to Big Data achieve greater impact.

Exhibit II-6: By Country, Percentage of Companies Where Big Data Improved Decision-Making



What Kinds of Digital Data are Companies Using?

One way that Big Data experts such as Tom Davenport distinguish between the eras of 'big' and 'little' data is on the type of data companies are using. Big Data is more associated with unstructured and external data. But what does this mean? While there are many ways to classify such data, the two most common are:

- **The degree to which the data is 'structured'**. Data that is numerical (financial, order, and other data) is regarded as structured – neatly able to fit in the columns and rows of modern database management software. 'Unstructured' data cannot so easily be compiled into older database formats. This data could be digital video, text (increasingly coming from comments on social media sites such as Twitter, Facebook and LinkedIn), digitized audio and other types. To analyze this data, the technology needs to process it in some manner. ('Sentiment analysis' is a hot trend in how to treat social media data – e.g., determining people's sentiments about a company and its products and practices.)
- **Whether the data is 'internal' or 'external' data**. Is data generated by the company or brought from the outside? For example, an increasing number of companies (particularly retailers and restaurant chains) are seeking external data from telecommunications firms that can track customers' locations through their mobile devices. The value of this data to retailers is the ability to intercept potential customers who are in the vicinity of their stores with targeted marketing offers that may convince them to walk in.

Defining Types and Sources of Digital Data

In our research, we defined data along two dimenions: structured versus unstructured and internal versus external. Given below are the definitions we used.

On the dimension of data structure:

- Structured – Data that resides in fixed fields (for example, data in relational databases or in spreadsheets)
- Unstructured – Data that does not reside in fixed fields (for example, free-form text from articles, email messages, untagged audio and video data, etc.)
- Semi-structured – Data that does not reside in fixed fields but uses tags or other markers to capture elements of the data (for example, XML, HTML-tagged text)

On the dimension of data source:

- Internal - from a company's sales, customer service, manufacturing, and employee records; from visits to the company's website, etc.
- External - from sources outside a company such as third-party data providers, public social media sites such as Facebook, Twitter and Google+, etc.

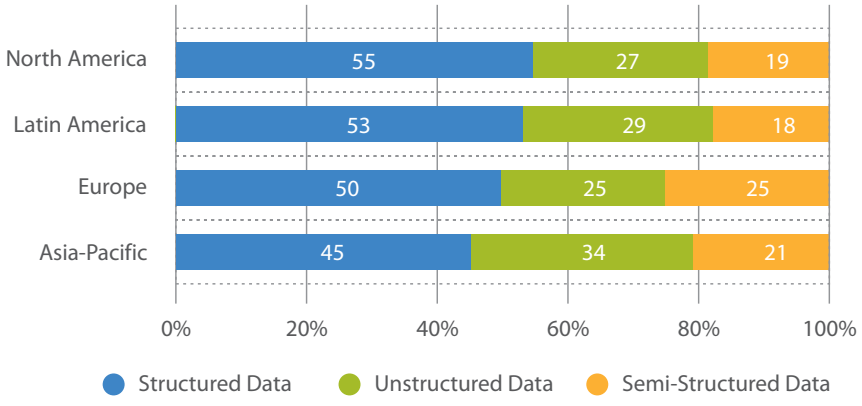
Classifying Big Data along these two dimensions, we then wanted to know how much of companies' data was structured versus unstructured, as well as how much was generated internally versus externally. We were surprised by the combined results across all four regions of the world that we surveyed:

- 51% of data is structured
- 27% of data is unstructured
- 21% of data is semi-structured

A much higher than anticipated percentage of data was not structured – either unstructured or 'semi-structured' (when combined, about half). (See Exhibit II-7) And a little less than a quarter of the data was external. (See Exhibit II-8)

Exhibit II-7: Percentage of Data that is Structured versus Unstructured

Q8 : Mean Estimated Percentage of Structured, Unstructured and Semi-Structured Data, Across all of the Company's Big Data Initiatives

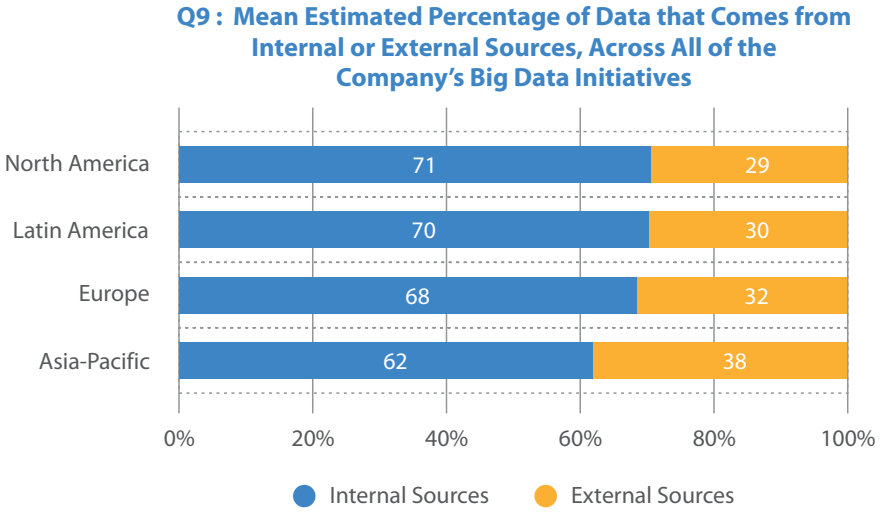


North American companies had the highest percentage of structured data; Asia-Pacific companies had the most unstructured data. North American companies also had the highest percentage of internal data; Asia-Pacific companies had the lowest.

To discover new patterns in Big Data, companies need highly efficient ways to aggregate data across data warehouses and other data stores. Since most data in these stores is structured, it is far easier for analysts to explore it. It is also not difficult to create structured data out of semi-structured data such as web activity.

However, unstructured data (for example, free-form text, video, audio, and image data where context needs to be derived from the data) is hard to discern. The most sought-after data right now, text as natural language processing (NLP), can be used to derive context that is beyond the typical sentiment analysis. Nonetheless, some text data (particularly Twitter tweets) are fairly semi-structured. Hashtags give some sense of context, while mentions, retweets, and @'s provide references to people. Facebook posts, blog posts, and other free-form text are more difficult to analyze, as noted above. However, tags and other meta-data can help narrow down the context of a comment.

Exhibit II-8: Percentage of Data that is Internal versus External



In the interviews that our research team conducted, many executives said their companies' usage of unstructured data is not only increasing but is also becoming essential. "Studies have been done on electronic records that show, on average, 80%-90% or more of data in records is unstructured data," one health care executive said. "That requires natural language processing to extract information." He said much of the health care industry is trying to improve capturing and analysis of unstructured data such as images, emails, physician and nurses' notes, etc.

Companies are increasingly looking to external data to get a fuller picture of activities that might affect them – particularly customer behavior. The soaring use of mobile devices now provides companies with data that, at least in theory, can help them track customer movements. This kind of external data is fully on the radar of global companies.

The head of Ford Motor Company's analytics group, John Ginder, put it this way to one trade magazine: "We recognize that the volumes of data we generate internally … as well as the universe of data that our customers live in and that exists on the Internet … are huge opportunities for us that will likely require some new specialized techniques or platforms to manage." Internet data that consumers provide appears to be of big interest. "The fundamental assumption of Big Data is the amount of that data is only going to grow and there's an opportunity for us to combine that external data with our own internal data in new ways. For better forecasting or better insights into product design, there are many, many opportunities."[2]

Who is Selling Their Big (Digitized) Data?

With companies capturing so much more digital data today to understand their operational performance moment-by-moment, the behavior of customers and suppliers, and other vital signs of the business, it's begun to raise eyebrows of both opportunity and concern. Executives are seeking data the organization has that might be of value to another organization, and from which the firm might be able to profit. That's the opportunity side.

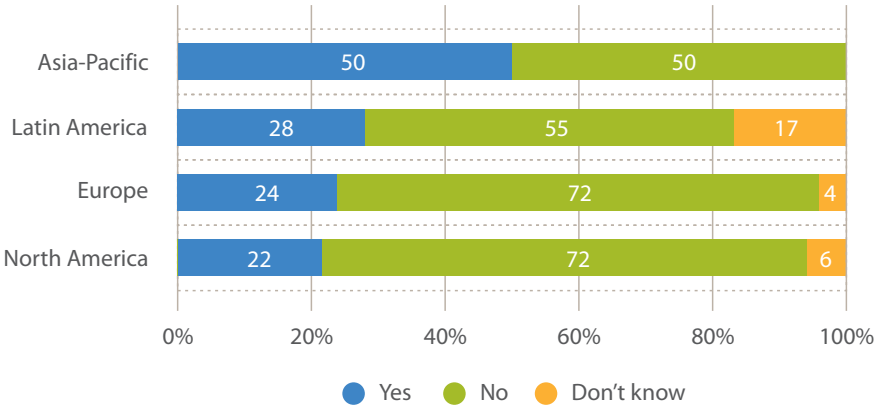
---

2 Jason Hiner, Zdnet, "Ford's Big Data chief sees massive possibilities, but the tools need work," July 5, 2012. http://www.zdnet.com/fords-big-data-chief-sees-massive-possibilities-but-the-tools-need-work-7000000322/

In 2012, about one-quarter of the companies we surveyed (27%) were capitalizing on this opportunity: selling their digital data. U.S. companies profited least from such data, with only 22% doing so. In contrast, half the Asia-Pacific companies we polled said they sell their digital data. About one-quarter of European and Latin American companies sold their digital data in 2012. (See Exhibit II-9)

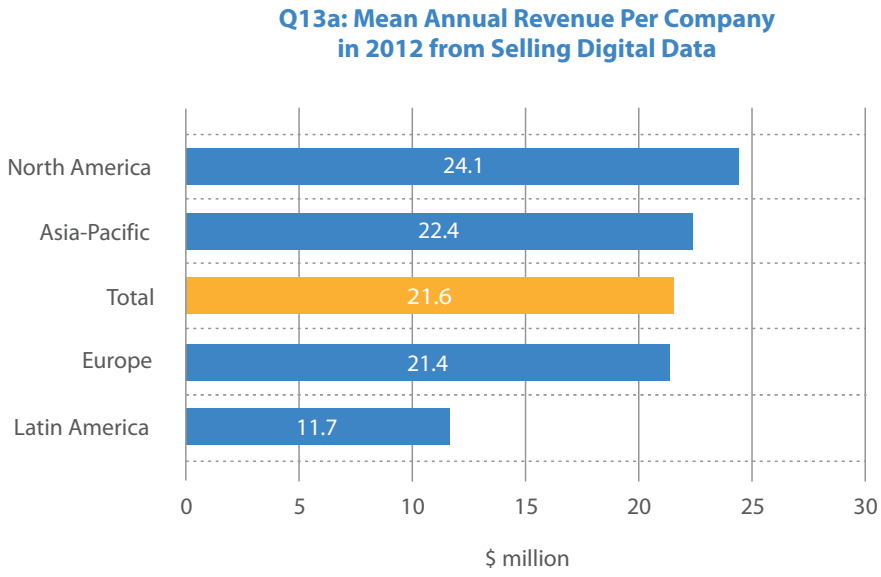
Exhibit II-9: Who's Selling Their Digital Data?

Q10: Percentage of Companies that Sell their Digital Data



For the approximately one-quarter of companies that sell their digital data, how lucrative is it? Our survey found that the annual revenue from selling such data was not trivial. In 2012, on an average, selling digital data contributed $21.6 million to the revenue of companies. (Exhibit II-10)

Exhibit II-10: How Much Money are Companies Generating from the Data They Sell?



So clearly, some companies are profiting from their data, albeit a distinct minority today. However, of the 73% of companies that did not sell such data, 22% said they do plan to sell such data by 2015; 55% don't; and 23% did not know. That means by 2015, 43% of companies will sell their digital data (the 27% that already do today, plus the 22% of the 73% that don't today).

What's the Return on Big Data?

One of the most important questions with regards to Big Data is, of course, whether companies are seeing a return on their hefty investments in it. As we mentioned earlier, one of our primary research goals was to begin to assess the return on investment on Big Data.

We asked survey respondents[3] to tell us what return they expected in 2012 on their Big Data investments.[4] To boost the number of people who would respond to this question and increase the chances that they calculated ROI the same way, we provided a range of returns with bands (e.g., "26% to 50%"), and furnished a formula for calculating percentage ROI:

$$\text{Percentage ROI} = \frac{\text{Gain from investment - Cost of investment}}{\text{Cost of investment}} * 100$$

Here's what we found on the ROI front: Only 8% said they expected a negative return, and 16% said they didn't know what the return would be. That could be read to mean that in about one-quarter of the companies, the return on Big Data was either negative or questionable. A larger percentage of respondents (33%) said the return would be low (between 0% and 25%). However, 20% expected the ROI to be between 26% and 50%. And nearly one-quarter (23%) projected a 2012 return of more than 50%. (See Exhibit II-11)

Exhibit II-11: Across Regions, the ROI on Big Data

Q17: Company's Expected Return in 2012 on Big Data Investments	
Negative return	8%
Between 0%-25%	33%
Between 26%-50%	20%
Between 51%-100%	16%
More than 100%	7%
Don't know	16%

3 For this calculation of Big Data ROI across a company, we asked the survey respondents who were either in the IT group or an analytics group, believing that since they supported the larger organization (and not just one business function), they would have a better view on overall enterprise ROI. When we present the data in our section on business functions, the managers of those functions calculated ROI numbers for their function.

4 Note: Since we began fielding the survey in December 2012, we consider these ROI 'expectations' more reliable indicators of real returns than if we had fielded the survey at the beginning of 2012 and asked companies to project their ROIs for 2012.

Using the above data , we then calculated the mean ROI overall across all regions of the world that we surveyed, as well as for each region. [5] Of those companies that reported ROI numbers (even if they were negative):

- The average ROI was 46% across all four regions
- Companies in Asia-Pacific (71% ROI) and Latin America (64% ROI) reported greater ROI than their counterparts in the U.S. (37%) and Europe (43%)

So what would an average 46% ROI be on, say, a $2 million investment in Big Data? It would mean that a company had a $2.9 million increase in revenue (or decrease in cost), but a $900,000 return after subtracting the cost of the investment.[6] That appears to be a decent return, and one likely to exceed many companies' internal 'hurdle' rates for investments. In addition, based on our 12 interviews with leading companies, this 46% mean ROI actually appears to be a very modest return. We heard about ROIs on Big Data of 25 to 1 or even more from senior executives.

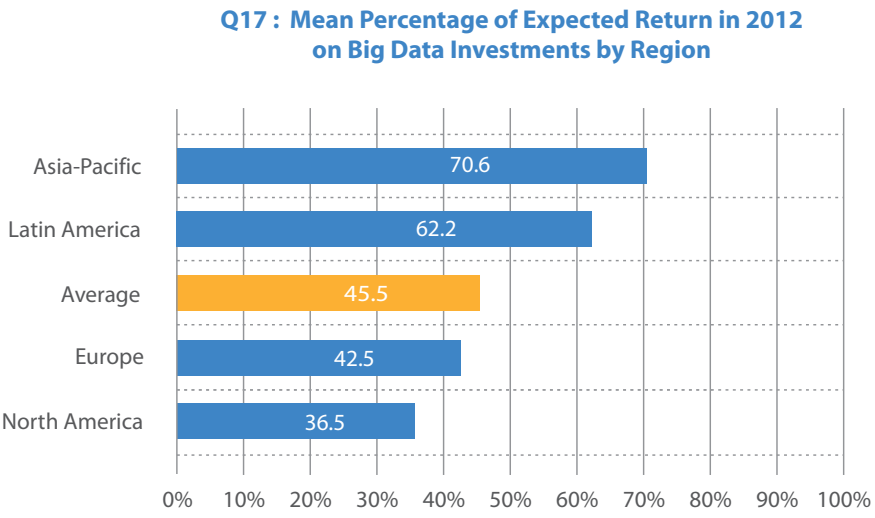
For example, an analytics director in one large Internet company indicated his group has helped increase firm revenue by hundreds of millions of dollars over the last five years. An executive at a major telecommunications company (which spends hundreds of millions of dollars annually on marketing) said analytics have boosted sales on promotional programs from 2% to 10%. After using analytics to determine what offer it should make to customers whose contracts were expiring, the company reduced the churn it could control by 20-30 basis points, a significant amount. And the analytics work also boosted sales of one product line to consumers by 10% to 15%, this executive estimated.

In a $2 billion Internet company, an analytics director says the ROI has been at least tenfold. "Big data has changed the way we do business fundamentally. We have a very big member base, which buys products that we create for them," he told us. "Analytics have helped us decide what products to build to serve our base. It's enabled us to segment them and understand what they need." The company appreciates the potential of the analytics team, which has grown from one person to 30 in two years. "Our CEO and the executive team understand the value we bring to the table," the executive said.

5 We then calculated mean ROI for each region by using the halfway points in our ranges as the multiplier.

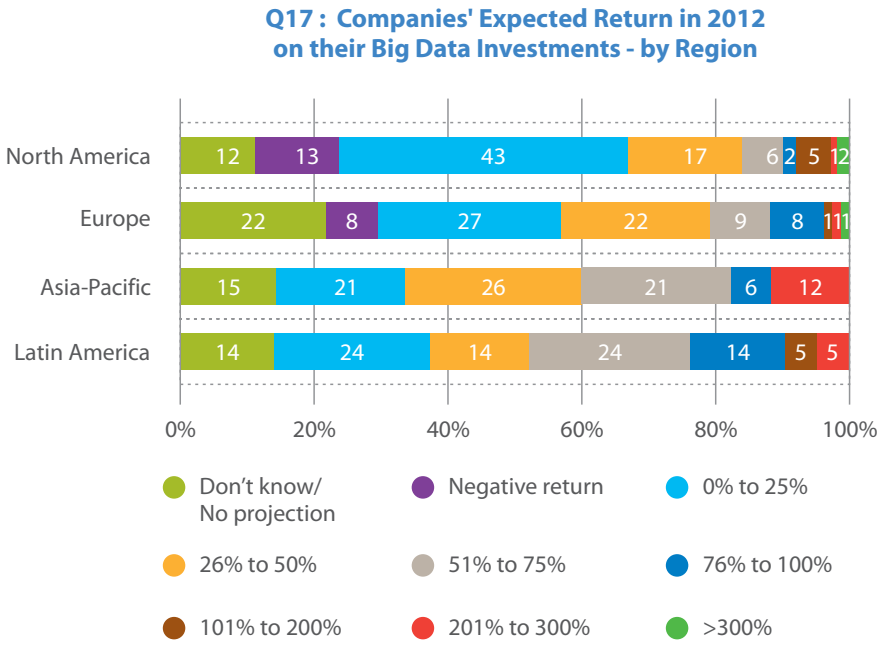
6 This data is self-reported, meaning that we haven't verified the ROI.

Exhibit II-12: Mean Projected ROI on Big Data by Regions of the World



For a detailed breakout of ROI by region and by the 'bands' of ROI we used, see Exhibit II-13.

Exhibit II-13: Detailed Breakdown of Mean ROI on Big Data by Regions of The World



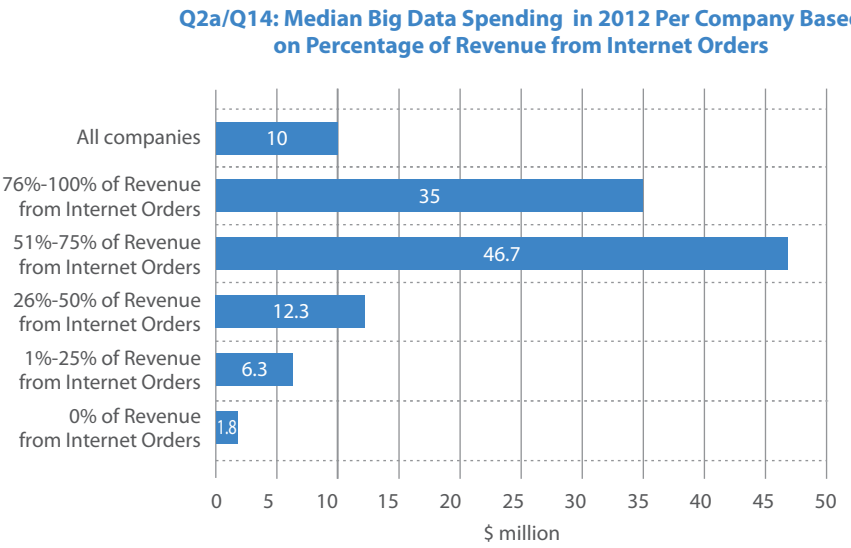
The Companies That are Furthest Ahead in Big Data are More Internet-Centric

If a company is doing more and more of its business through the Internet, it of course must constantly improve its online experience. That requires continually collecting data about customers (and non-customers –viewers who don't buy) who connect with it using online channels (the corporate website, ecommerce site, social media sites, mobile apps, email, and many more) and analyzing their behavior across all those digital channels (increasingly, their comments in social media channels). And this implies investing in technologies and analysts who constantly monitor that activity and suggest changes to the website, ecommerce site, mobile apps and more.

In fact, our survey showed that the more business a company does on the Internet, the more it invests in Big Data. We asked respondents to tell us how much of their revenue came from Internet orders. Those who said they generated 76% to 100% of revenue via Internet orders (a group comprising 11 companies) spent a median of $35 million per company on Big Data – about six times the amount spent by companies that generated from 0% to 25% of revenue from Internet orders. Median spending on Big Data by companies generating 51% to 75% of revenue from the Internet (there were 32 such companies) was even greater: $47 million per company.[7]

Adjusted for company size, companies amassing a higher percent of revenue via the Internet spent more on Big Data. Spending, in companies with 76% to 100% revenue from Internet orders, was higher on a percentage of revenue basis (1.0% of median annual revenue) than it was for companies with 51% to 75% of revenue from the Internet (0.4% of median annual revenue). The more digitally-centric a company is, the more it invests in Big Data. (See Exhibit II-14)

Exhibit II-14: Median Spending Levels on Big Data in 2012 by Percentage of Company Revenue that Came Via the Internet



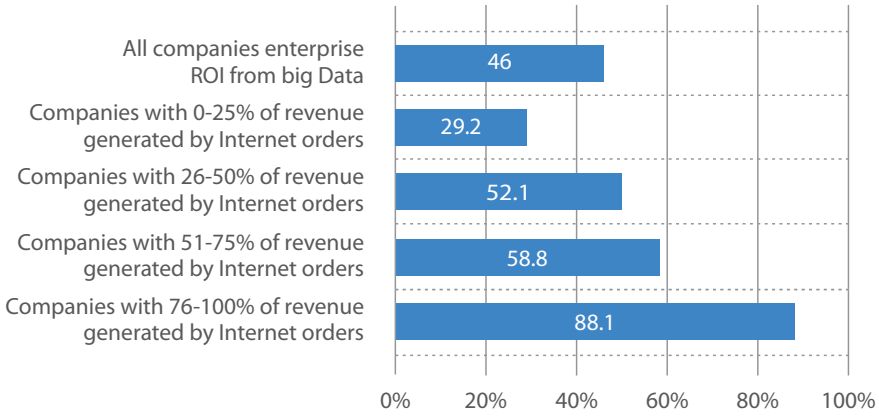
7 Using mean rather than median data, spending on Big Data for companies that generated 76%-100% of sales through the Internet was $304 million (3% of revenue) vs. $89 million for all companies (or 0.49% of revenue). For companies generating 51-75% of revenue through Internet orders, mean spending was $132 million (or 0.5% of revenue). For companies generating 1%-25% of revenue through the Internet, mean spending on Big Data was $47 million (0.2% of revenue).

Given how much Internet-centric companies depend on Big Data, it is no coincidence that many of the early technologies of Big Data (such as the Hadoop database management system) emerged from these Internet companies. To manage all their Big Data, companies had to build many of the tools from scratch. These companies continue to spend hugely on Big Data and data scientists.

The ROI data also supports our premise that digitally-centric companies are more advanced with Big Data. (See Exhibit II-15) Companies with the largest percentage of revenue from Internet orders expected an average 88% ROI on their Big Data investments in 2012 – nearly double the 46% average ROI for all companies. And it was *three* times the average ROI of companies whose Internet orders were 0% to 25% of revenue.

Exhibit II-15: Internet-Centric Companies Projected the Highest Returns on Big Data in 2012

Q17 : Expected 2012 Enterprise ROI on Big Data by Percentage of Revenue Generated from Internet Orders



How Have Companies Organized Their Big Data Activities?

From our interviews with executives, one of the biggest areas of concern and interest was identifying how to optimally organize Big Data activities in a company. We used the online survey to shed light on this question, too. (See Exhibit II-16) We asked respondents to tell us in which function their Big Data professionals resided[8]:

- **Those who processed the data:** The most frequent response (by 45% of the respondents) was in the IT function. Some 28% said analysts were in the business functions that used the data to make decisions. And the smallest percentage (24%) said they were organized in a separate Big Data group.

- **Those who analyzed the data**: Here the most frequent response was in the business functions that used the data (42%). In 30% of the companies, the analysts were in the IT function; and in 25% of the companies, they were in a separate Big Data group.

However, there were some large regional differences. Most notably, Asia-Pacific companies were more likely to have Big Data professionals (for both processing and analysis of data) in a separate group while U.S. companies were least likely to do this. Latin American companies were more likely to outsource Big Data processing and analytics to external companies, although only 7% said this was where their Big Data expertise was located.

---

8 We must stress that these percentages are percentages of companies that put their Big Data professionals in one of the three functions (IT, business functions or an analytics group). The percentages are not how the average company divides up its Big Data staff in these three areas.

Exhibit II-16: Where Big Data Professionals Reside in Companies

Q19 & 20: Where Companies Place Their Big Data Professionals (By Percentage of Companies)					
	Total	U.S.	Europe	Asia-Pacific	Latin America
Professionals Who Analyze Big Data					
In IT function	**29.5%**	28.6%	27.0%	30.7%	38.0%
In business functions that use the data	**42.1%**	49.4%	42.9%	28.7%	28.2%
In a separate Big Data group	**25.0%**	19.2%	26.4%	39.6%	26.8%
With a third-party provider	**2.3%**	1.3%	3.1%	1.0%	7.0%
Other	**0.9%**	1.6%	0.6%	0%	0%
Professionals Who Process Big Data					
In IT function	**45.4%**	45.5%	49.1%	38.6%	46.5%
In business functions that use the data	**27.7%**	30.5%	25.8%	23.8%	25.4%
In a separate Big Data group	**23.5%**	20.8%	21.5%	36.6%	21.1%
With a third-party provider	**2.5%**	1.3%	3.7%	1.0%	7.0%
Other	**0.9%**	1.9%	0%	0%	0%

From our interviews, we heard that a number of analytics groups were located in the IT function. A $2 billion industrial manufacturer's three full-time analytics professionals were part of the IT group, reporting to the Chief Information Officer. The company also has analysts in each business unit (at least one per unit).

A large hospitality company has two advanced analytics groups: one that works for the Chief Marketing Officer, and the other for a group that reports to the Chief Operating Officer and focuses on hotel operations. In addition, the company has small analytics teams in business units and functions. However, these professionals do not work on advanced analytics.

What are the Biggest Challenges to Getting Returns?

The media has written extensively about the challenges companies face in collecting, processing, analyzing and using Big Data in their businesses. Much public discussion has focused on the three 'V's' – handling the volume, variety and velocity of the data. Another frequently mentioned challenge is finding people who know how to analyze the data – data scientist has become a hot profession. Driving business decision-makers to actually use the data and abandon making decisions on intuition is also a challenging task.

But which of these challenges are the greatest? To gain greater insights on this, we asked respondents to rate 16 challenges that we found mentioned most frequently in press articles, public speaking presentations, and our client work. We provided a scale of 1 to 5 (1 = not at all a challenge, 2= minor challenge, 3= moderate challenge, 4= high challenge, and 5= very high challenge). The results can be seen in Exhibit II-17.

Exhibit II-17: Key Challenges of Big Data Across Regions of World

Q23: Mean Rating of 16 Challenges in Getting Business Value from Big Data (Scale of 1-5)		
Rank	**Challenge**	**Score**
1	Getting business units to share information across organizational silos	3.37
2	Being able to handle the large volume, velocity and variety of Big Data	3.35
3	Determining what data (both structured and unstructured, and internal and external) to use for different business decisions	3.34
4	Building high levels of trust between the data scientists who present insights on Big Data and the functional managers.	3.26
5	Finding and hiring data scientists who can manage large amounts of structured and unstructured data and create insights	3.23
6	Getting top management in the company to approve investments in Big Data and its related investments (e.g., training, etc.)	3.22
7	Putting our analysis of Big Data in a presentable form for making decisions (e.g., visualization/ visual models)	3.21
8	Finding the optimal way to organize Big Data activities in our company	3.20
9	Understanding where in the company we should focus our Big Data investments	3.18
10	Determining what to do with the insights that are created from Big Data	3.18
11	Reskilling the IT function to be able to use the new tools and technologies of Big Data	3.15
12	Getting functional managers to make decisions based on Big Data, rather than on intuition	3.12
13	Determining which Big Data technologies to use	3.10
14	Keeping the data in Big Data initiatives secure from external parties	3.08
15	Getting the IT function to recognize that Big Data requires new technologies and new skills	3.05
16	Keeping the data in Big Data initiatives secure from internal parties	2.95
Scale of 1-5, 1= not at all a challenge, 2=minor challenge, 3= moderate challenge, 4=high challenge, 5= very high challenge		

We have several observations regarding our findings on this question, combined from the four regions around the world:

- **None of the 16 challenges stands far above the others**. All but one of the 16 challenges received mean ratings between 3.0 and 3.4. In other words, across all 643 companies, they were slightly more than 'moderate' challenges. That surprised us; we expected to see a few rated higher than 4.0.
- **The highest-rated challenge is not a technological issue**. It is an organizational or cultural issue: getting business units in a company to share information across the organizational silos (divisions, business functions, etc.). This issue has plagued companies for decades, long before the phrase 'Big Data' was coined. Business functions become protective of their data and often don't have any incentive to share it internally. However, many business decisions can be dramatically improved when decision makers have access to the bigger picture: what's happening with certain customers, problematic products, and service issues. Xerox Corp., for example, sees a 'big play' with Big Data when it can combine data on how its 1.2 million machines are operating in the field with claims information. The devices themselves report data on their operating condition back to Xerox every day. But the company wants to aggregate this data with claims and other data kept by a company division that provides business process outsourcing services.[9]
- **Finishing a close second was a technological issue:** dealing with the three V's of Big Data.
- **Companies see getting managers to make decisions based on Big Data rather than on intuition – relative to the other 15 challenges – to be a smaller challenge**. It ranked 12th on the list, just higher than figuring out which technologies to use.

We explore some of the top-rated challenges below:

Getting Business Units to Share Information Across Organizational Silos

Getting business units and functions to share data is a huge cultural challenge at most companies, where functional or divisional measures and performance incentives trump organization-wide metrics and rewards. In our consulting work, this issue is vastly under-appreciated amidst all the hype about Big Data. It has prevented many companies from tapping the greater potential of Big Data.

A senior executive at a large company who spoke to our research team said employees in his company are very protective of data because of privacy and security issues. (Customer data privacy is a major concern in this company's industry.) He said most data sources in the company are locked up. "The easiest part of Big Data is the technology part. The hard part is practice standards, legal ramifications, regulatory issues, all these kind of others things."

Handling the Three V's

The three V's of Big Data – volume, variety and velocity continue to be a key challenge. The telecommunications company we spoke with said handling the velocity of the data they get is of paramount importance, especially in the firm's wireless communications unit. Customer turnover is higher there than in other businesses, and consumers get new phones every couple of years. "I think speed is probably the important thing for us because if we don't quickly get that information, it's not that useful to us," he said.

9 Xerox vice president and global chief information officer Carol Zierhoffer mentioned this to The Wall Street Journal recently. See story here: http://blogs.wsj.com/cio/2013/02/21/the-morning-download-xeroxs-big-play-with-big-data/

Determining What Data to Use: Building Trust Between Data Scientists and Functional Managers

In our interviews, we often heard that if functional managers don't trust the data scientists who bring fact-based recommendations for improvements, those managers are not likely to accept the advice. This is especially the case if those managers are insecure about their current strategy.

An executive who runs marketing analytics at one large insurance company said it is important to get functional managers involved upfront in an initiative. "One of the key challenges is people who question the data. They don't like the results, so they'll find problems with the data. … You get people to believe it's good data by getting them in the design of the project at the beginning. We work with folks and say, "Here's what we want to do and measure, and here's why we think it's going to work." If we get them involved upfront, they're less likely to have a problem with it." The executive said that this approach has resulted in building trust across business functions. "When I first came on board several years ago, you had to beg to get a seat at the table. We had to push our ideas toward them. Now it's shifted to more of a pull. They know what we can do with data, and now they are pulling information from us – contacting us rather than us contacting them."

An analytics manager at one large company said one of the keys to building strong relationships with functional managers is having analysts with 'storytelling skills'. They must be able to translate analytics into something that busy executives understand and care about. If they can't, the likelihood of having executives do things differently is small. "If you have generated lots of data, the first key to success is being able to be persuasive with it," this manager said. The company provides training to its analysts on how to make data more visual and give succinct answers to pointed questions.

We heard many times that, at the end of the day, executives must be willing to listen to the often-contrary conclusions of what the Big Data analysts are telling them. This was nicely summarized by Bank of America's head of technology and operations, Catherine Bessant, in a conversation with *The Wall Street Journal*: Big Data pays off when "the manufacturing of brilliant data and making brilliant use of it [is accompanied by] a drive for abject purity in listening."[10]

Hiring Data Scientists

Data scientists who are experts in quantitative skills and highly effective at communicating their findings in language that functional managers understand are very rare. This does not bode well for companies, given that the need for professionals who can process and analyze Big Data is large and growing quickly.

Ford Motor's analytics chief John Ginder indicated to a reporter that the $134 billion automaker doesn't have nearly enough highly skilled people who know how to use Big Data tools for managing huge sets of data (like Hadoop) and analysis. "We have our own specialists who are working with the tools and developing some of their own in some cases,

10 From an article by Michael Hickins, Banks Using Big Data to Discover 'New Silks Roads', The Wall Street Journal, Feb. 6, 2013. http://blogs.wsj.com/cio/2013/02/06/banks-using-big-data-to-discover-new-silk-roads/ Note: Bessant has also said she dislikes the term "Big Data.""It implies something monumental out of something that should be fundamental, and should be basic, which is the creation of accurate, timely data on a reliable basis," she explained to CIO magazine in January 2013: http://www.cio.com/article/726169/Bank_of_America_CIO_Says_Simplification_and_Risk_Reduction_Are_Keys_to_IT_Success?page=5&taxonomyId=3172

and applying them to specific problems," he told ZD Net.[11] "But there is this future state where we'd like to be where all that data would be exposed – where data specialists (not computer scientists) could go in and interrogate it and look for correlations that they might have not been able to look at before. That's a beautiful future state, but we're not there yet."

A 2011 McKinsey Global Institute study predicted that by 2018, the demand for analytics and Big Data people in the U.S. alone will exceed supply by as much as 190,000.[12] Many companies are already seeing a shortage. A July 2012 survey of 108 business technology professionals by the trade publication *InformationWeek* found only 17% believe they'll 'easily' fill Big Data jobs. More than half (53%) said the skill set may be hard to find. And 23% said the salary demands of many data scientists may be more than they can afford.

A senior executive at a major insurance company said he sees many more resumes than real candidates for the analytics jobs in his company. "In addition to the job skills, our people have to be able to communicate and build relationships with [functional and business unit] managers. It can be very, very hard to get all of those things in one person."

He recruits Big Data analysts largely through personal connections – friends of friends. "The good news is that there are tons of people out there looking for jobs. The bad part is that there are tons of people out there who I would never think of hiring," he told us. "That cuts down the population big time. Applicants have to have credibility to tell me what we need to do. Then they need the ability to do it and communicate and build relationships."

Optimally Organizing Big Data Activities

Should a company's Big Data professionals operate outside the business functions so they might offer more unbiased information? But won't that mean they don't understand the business functions they are asked to help as deeply as they should? And if they operate within a business function, how do they avoid finding data and making recommendations that help confirm the functional head's long-held beliefs and strategy?

And if analytics is centralized, should it be part of a central IT group so that it can work more closely with the technologists who maintain many of the firm's core data sources? Or should analytics be centralized but report somewhere else?

Questions like these were on the minds of the executives who spoke to our research team. A telecommunications company executive said his firm's three business units had their own analytics groups focused on their own customer segments. But within each business unit, functions (such as marketing) also had their own analytics professionals. Another large company's analytics function (with more than 75 employees) resides within IT. Nonetheless, other analysts are embedded within business units across the organization.

A health care company has pushed down its Big Data activities into business units and functions. In recent years, the IT organization launched an internal advisory group that collects best practices and tries to standardize usage and licensing of analytics tools. However, this has been difficult to do because of the company's decentralized approach to Big Data.

These and the other eight challenges vary, sometimes significantly, by industry and by business function, as we'll explain in the corresponding sections of the report. Nonetheless, combined across all regions, industries and business functions, none of the challenges stands head and shoulders above the others.

---

11 Zdnet article. http://www.zdnet.com/fords-big-data-chief-sees-massive-possibilities-but-the-tools-need-work-7000000322/

12 The McKinsey report can be found here: http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

Telecom, Travel, High Tech and Banking are Ahead of the Pack on Big Data: Findings by Global Industries

Highlights:

- Telecom, travel, high tech and banking firms spend the most on Big Data
- However, utilities and energy and resources companies expected the biggest ROI
- Media and entertainment companies use the most unstructured data; high tech and telecom companies use the most external data
- Utilities and telecom companies are most likely to sell their digital data, but insurance companies make the most from that data

We went into the research believing that Big Data practices would vary significantly in each of the 12 industries that we focused on (see box). While, of course, all these industries have invested heavily in information technology over many decades, they have different levels of data intensity – that is, the volume, variety, and velocity with which digital data courses through their information systems.

Retail chains, for instance, must process data on purchases of hundreds of millions of items from their stores. And they must do so rapidly to better match pricing and product availability to regional or local demand. In contrast, auto manufacturers sell thousands of cars to independent dealers every week, and thus have a less onerous task of understanding sales trends.

With this in mind, we wanted to know whether certain industries spent more than others; whether they achieved higher ROI; how they focused their investments; how their key challenges compared; and other issues.

This section summarizes these differences.

The Industries That We Studied

- Banking/financial services
- High Tech
- Retailing
- Consumer products manufacturing
- Energy and resources
- Travel, hospitality and airlines
- Insurance
- Heavy manufacturing
- Telecommunications services
- Pharmaceuticals/life sciences
- Media and entertainment
- Utilities

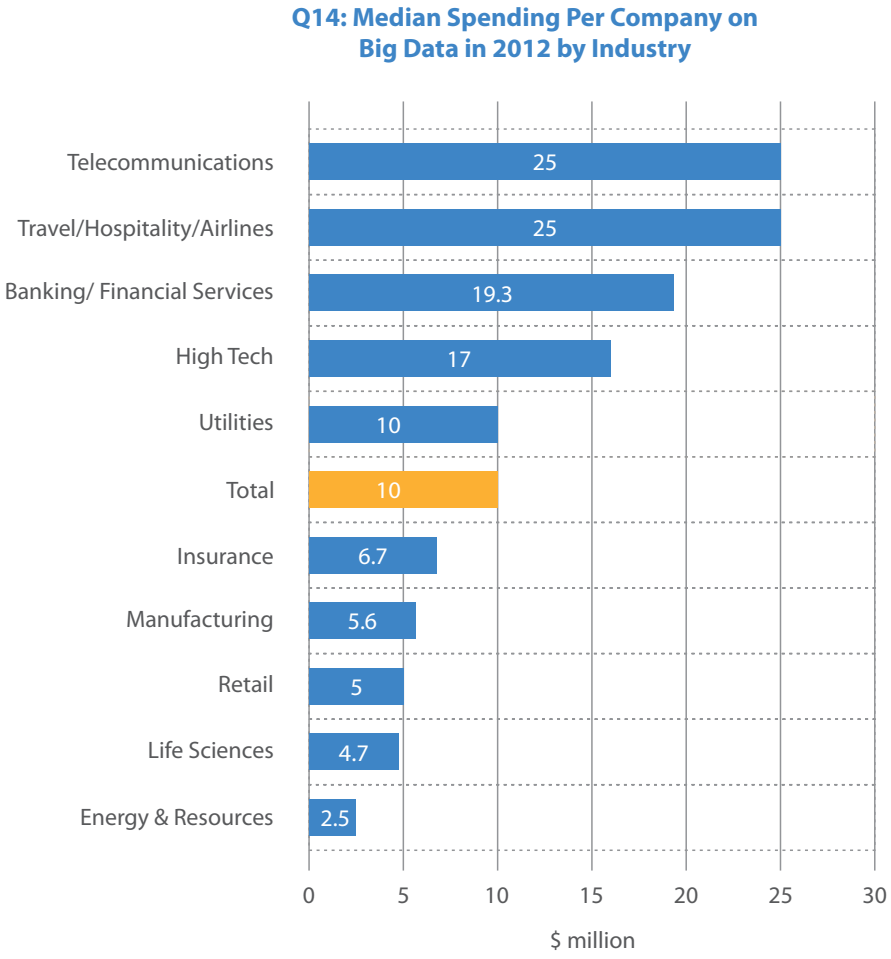
Spending: Which Industries are in the Lead?

Travel/hospitality/airlines, telecommunications, banking/financial services and high tech companies spent more on Big Data in 2012 than the other sectors that we surveyed (by median spending), as shown in Exhibit III-1[13].

Travel/hospitality/airlines companies spent a median $25 million/company, as did telecom companies. High tech companies' median spending per company was $17 million. Banking/financial services companies' median spending was $19.3 million per company.

On the opposite end of the spectrum were life sciences companies ($4.7 million) and energy & resources firms ($2.5 million).

Exhibit III-1: Per-Company Spending on Big Data by Global Industry



13 In two of our 12 industries (media & entertainment and consumer goods), we had low numbers of IT or analytics managers who answered this question. We have, therefore, not reported their figures.

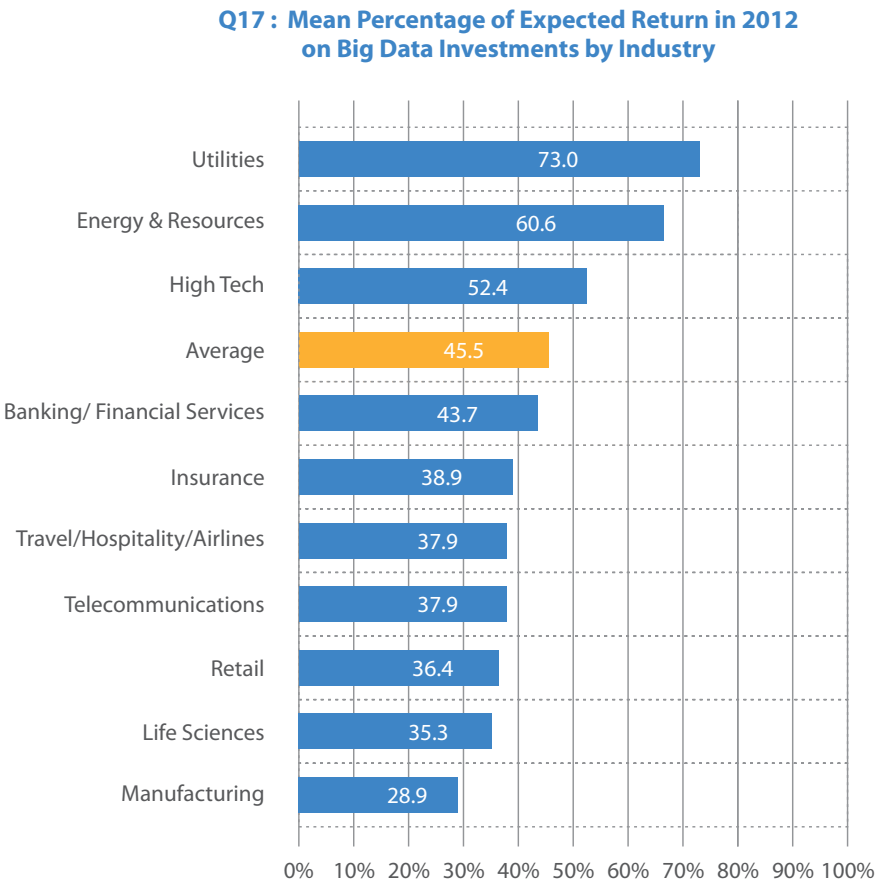
Return on Investments: Utility and Energy Companies Expected the Most

We mentioned earlier in this report that the IT and analytics managers we surveyed estimated their expected 2012 return on Big Data investments to average 45.5%. So which industries were on the high and low side of that mean? Two may surprise you.

Utilities and energy & resources companies had the highest expectations for generating returns on their investments, even though they spent far less than average per company on Big Data in 2012. The average return per utility company was 73%; for the average energy & resources company, it was 61%. High tech companies' estimated 2012 ROI was also higher than average, at 52%, while banks/financial services companies were just lower than average (44%). (See Exhibit III-2)

Companies in the heavy manufacturing (29% ROI), life sciences (35%), retailing (36%), travel/hospitality/airlines (38%), and telecom sectors (38%) had the lowest expected returns on Big Data in 2012.

Exhibit III-2: Mean Expected ROI on Big Data by Industry



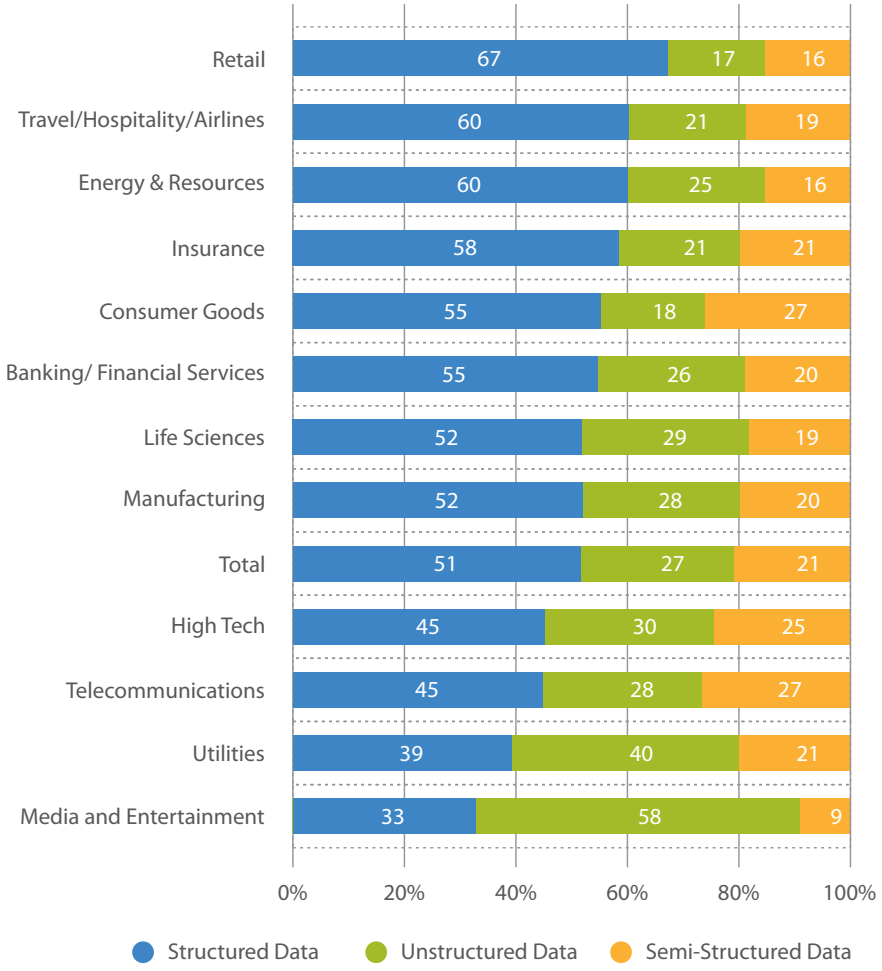
Which Industries Use More Unstructured and External Data?

We found that some industries are much heavier users of the unstructured and externally sourced data that has more recently flooded into the digital veins of corporate IT networks.

When it came to the use of structured and unstructured data, media and utility companies reported using the highest percentages of unstructured or semi-structured data (and thus the lowest percentage of structured data). In fact, both industries indicated that about two-thirds of their data was unstructured or semi-structured. In comparison, retailers, travel/hospitality/airlines and energy & resources companies had the highest percentage of structured data – all at least 60%. (See Exhibit III-3)

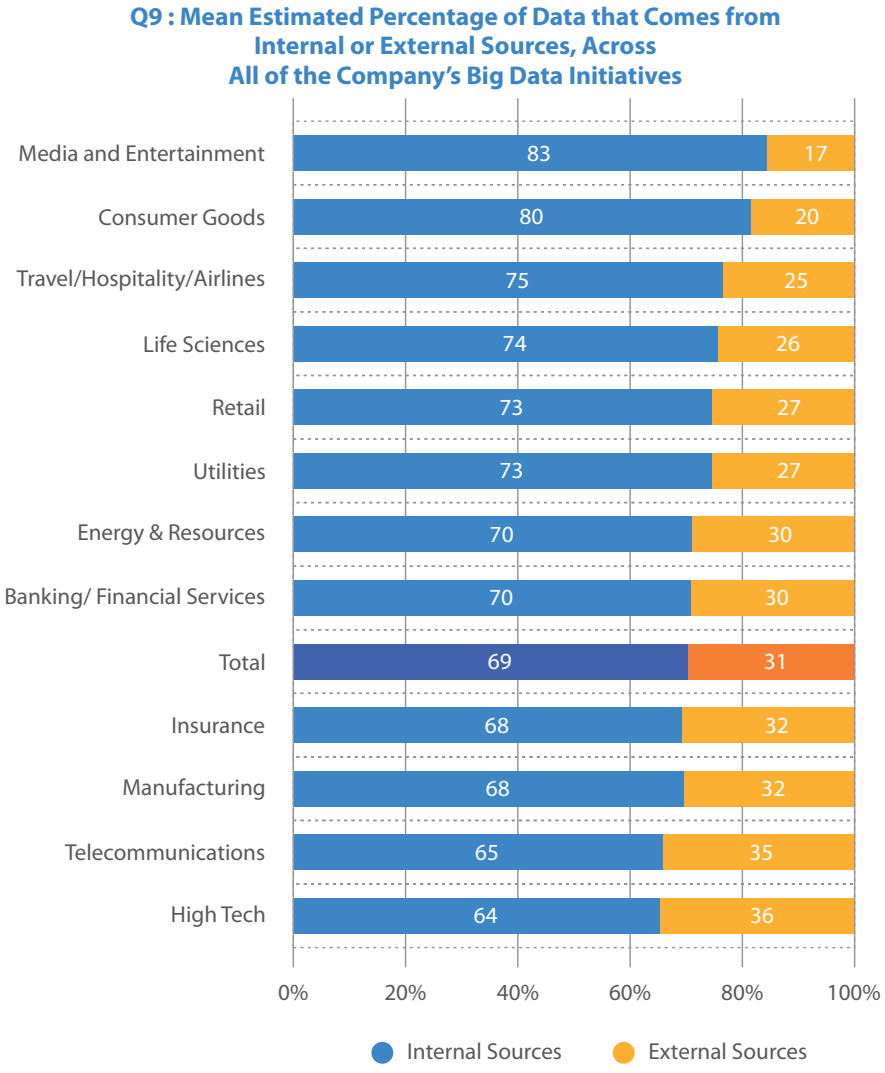
Exhibit III-3: Industry Comparisons on Structure of Data

Q8 : Mean Estimated Percentage of Structured, Unstructured and Semi-Structured Data, Across All of the Company's Big Data Initiatives



When it came to the percentage of data sourced internally compared to the percentage of data sourced externally, the companies with the greatest mix of external data were from high tech (36%), telecom (35%), heavy manufacturing (32%), and insurance (32%). At the other end were the media & entertainment industry (only 17% of data was sourced externally) and consumer goods (with 20% of data sourced externally). (See Exhibit III-4)

Exhibit III-4: Industry Comparisons on Sources of Data (Internal vs. External)

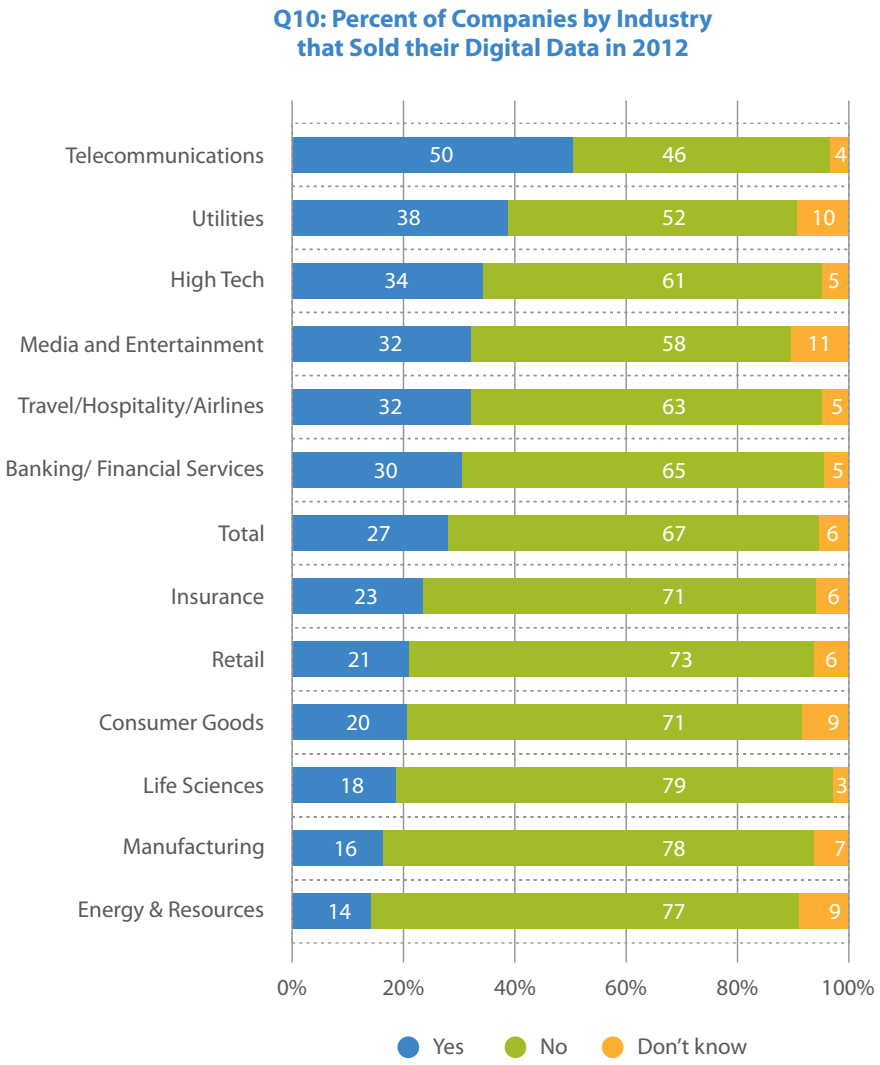


Which Industries Sell Their Data and Make Money From It?

Not only do these industries differ widely by how much they spend on Big Data, the returns they generate, and the type of data they use, they also vary greatly in their data-selling practices. Companies in certain industries – telecom and utilities in particular – are far more likely to sell digital data externally. Half the telecom companies we surveyed do so, along with 38% of the utility companies. (See Exhibit III-5)

Industries that are least likely to sell digital data are energy & resources (only 14% of companies do so), heavy manufacturing (16%), and life sciences (18%).

Exhibit III-5: Which Industries are More Likely to Sell Their Digital Data?

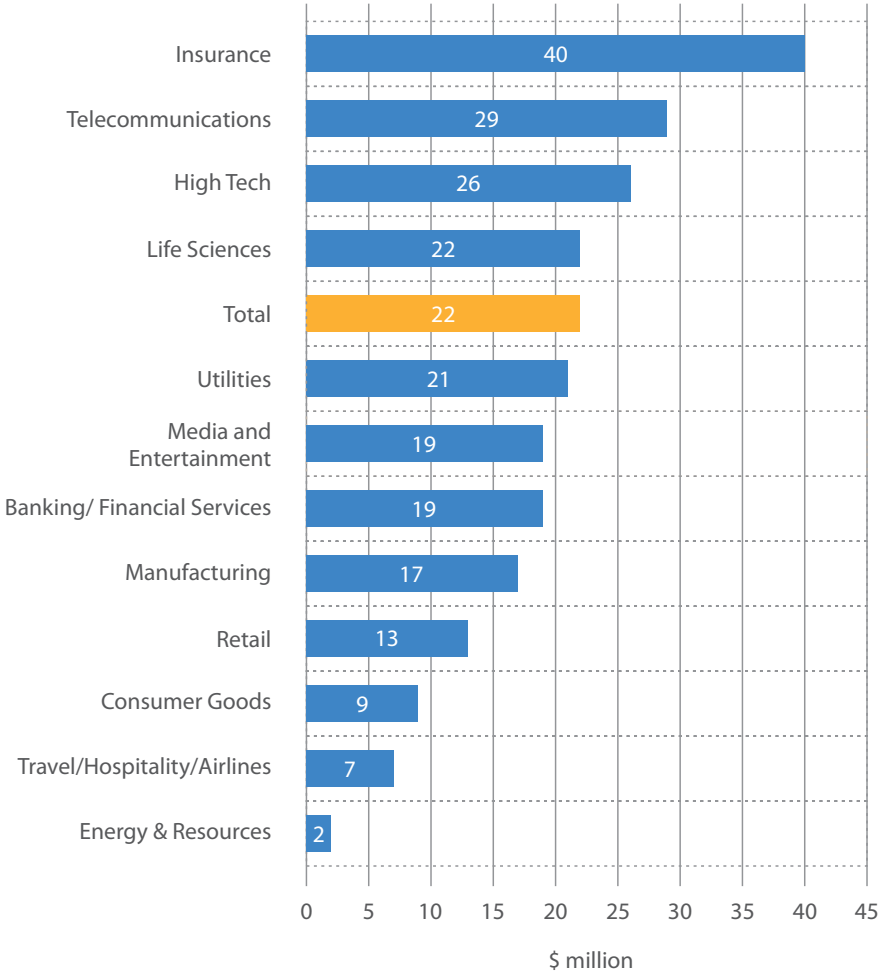


So how do these industries compare in terms of how much revenue they generate from selling digital data? One industry – insurance – finishes far above the others. This is especially noteworthy given that only 23% of the insurance companies surveyed sell their digital data to third parties. The average insurance company we surveyed had $40 million in revenue in 2012 from the digital data that it sold. Telecom ($29 million per company) and high tech ($26 million) companies were a distant second and third. (See Exhibit III-6)

At the other end of the spectrum were energy & resources companies (averaging only $2 million in revenue for data sold) and travel/hospitality/airline companies ($7 million).

Exhibit III-6: Insurers Generate the Most Revenue from Selling their Data

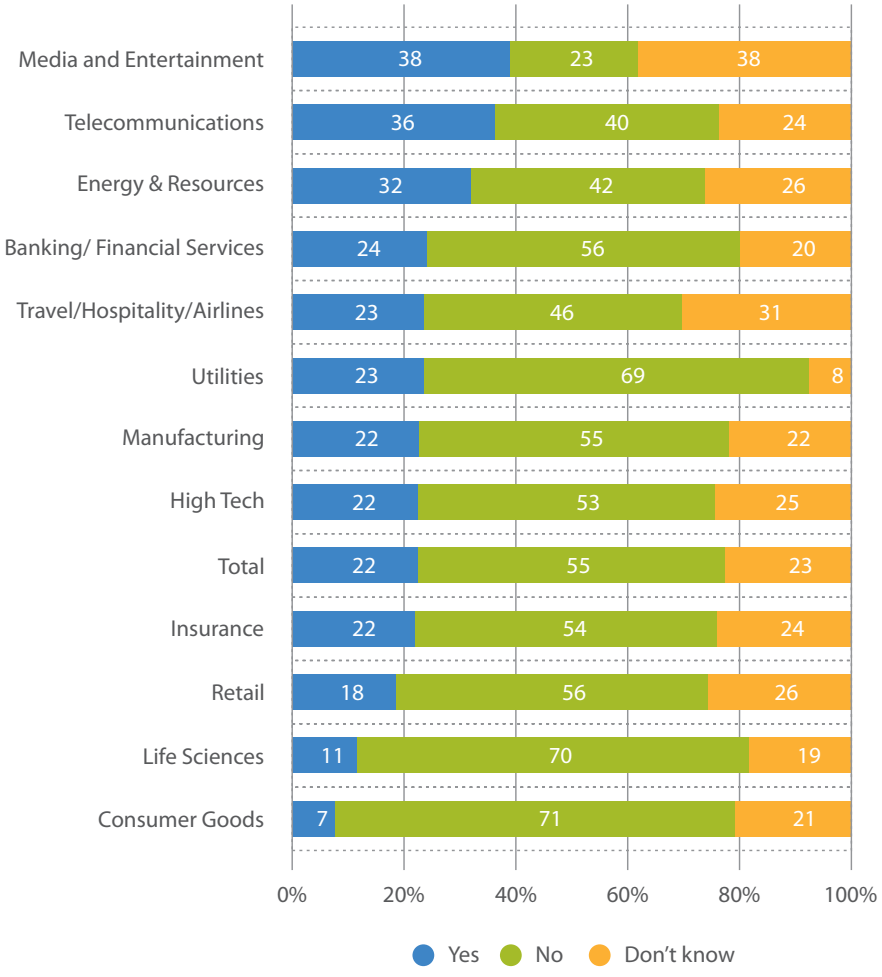
Q13a: Mean Annual Revenue Per Company by Industry in 2012 from Selling Digital Data



So what about the companies that didn't sell their digital data in 2012? Which industries were most likely to change their minds by 2015 and sell their data? Media and entertainment (38% of those that don't sell their data today plan to do so by 2015), telecom (36%), and energy & resources (32%), in that order. Only 22% of insurers that don't sell data today plan to do so in three years – despite being the industry with the highest per-company revenue from selling digital data. (See Exhibit III-7)

Exhibit III-7: Who Plans to Sell their Data by 2015?

Q13b: Percent of Companies by Industry that Expect to Sell their Digital Data by 2015



Sales and Marketing Get More of the Big Data Budget: Findings by Business Function

Highlights:

- Sales and marketing get the biggest shares of the Big Data pie
- However, finance and logistics expect the highest ROI on Big Data
- Eight business functions vary significantly in where they see the benefits from Big Data – and the biggest challenges they face in gaining those benefits

Cutting the Big Data Pie by Function

In addition to asking the IT and analytics functional managers to estimate how much their companies spent on Big Data in 2012, we asked them how the pie was divided. Which functions were getting the greatest and least shares?

The three business functions that are most directly related to generating revenue in any company account for 42% of total Big Data spending: sales (15.2%), marketing (15%), and R&D/product development/product engineering (11.3%). Customer service, another function that involves daily customer interaction, accounts for 13.3% of the Big Data budget.

"No matter which industry you talk about, I see analytics having the greatest impact on marketing first and customer service, second," says Venkatesh Ravirala, executive director of business analytics at MGM Resorts International, a $9 billion gaming and hotel company. "You are going to see companies address those two areas with Big Data because of real-time service and real-time offers. The decision-making capability and the response time to do this have shortened significantly."

Marketing and sales is where one multibillion auto insurance company has begun its Big Data initiative. The head of marketing analytics told us, "It was a good subset of data, and it wasn't too big to tackle."

The Eight Business Functions That We Explored for Big Data Practices

In addition to surveying IT and analytics executives, we also wanted to collect the experiences of senior managers in eight core business functions:

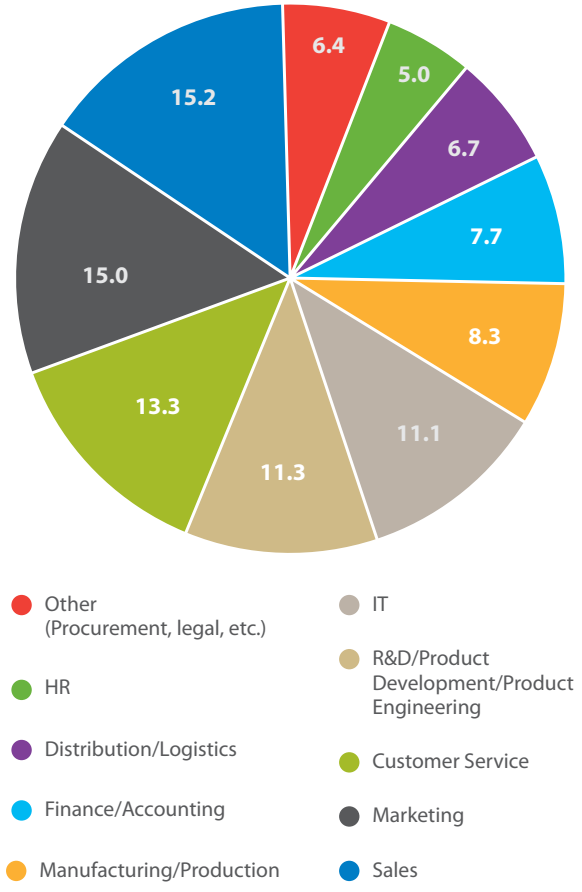
- Marketing
- Customer service (post-sale)
- R&D/product development/product engineering
- Human resources
- Sales
- Manufacturing (or production in services companies)
- Logistics/distribution
- Finance/accounting

These managers accounted for 62% of the total survey population.

In contrast, four functions competed for less than 30% of the budget: manufacturing (or, in a services company, operations, which commanded 8.3% of the budget), finance/accounting (7.7%), distribution/logistics (6.7%) and human resources/personnel (5%). Of course, these functions are further removed from the revenue generation process. However, if this is the way companies decide how to allocate their Big Data dollars, our research shows it may be misguided, as we explain in the next part of this section.

Exhibit IV-1: How Companies Cut the Big Data Pie by Functional Area

Q16: Where Companies Across Industries Focused Their 2012 Big Data Investments (as a Percentage of Total Big Data Investments)



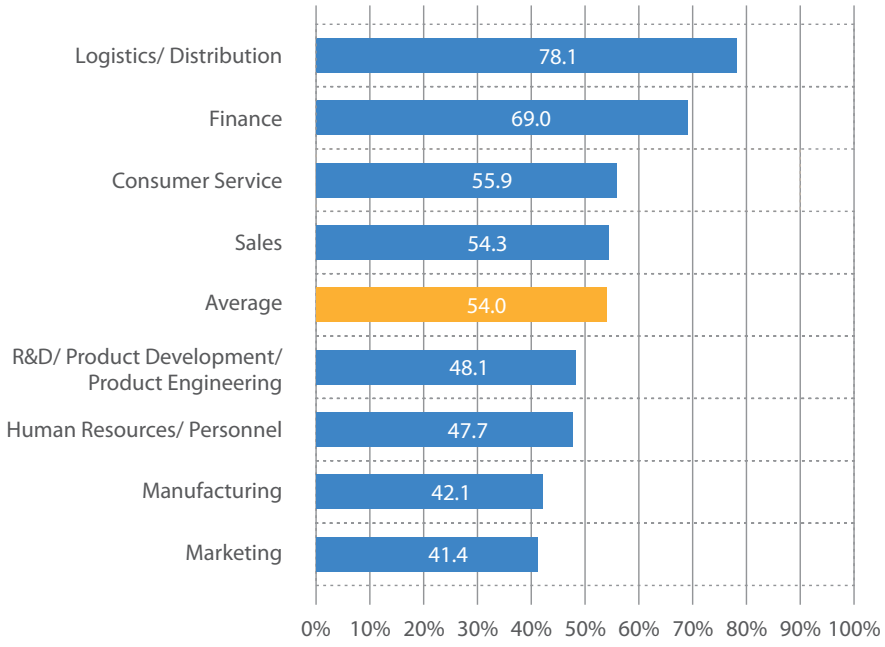
Logistics and Finance Expected the Biggest ROI in 2012

In addition to asking IT and analytics managers to estimate their company's total returns on Big Data in 2012, we asked managers of business functions to calculate the ROI they expected in 2012 within their functional area.

Since sales and marketing had, by far, the largest shares of the Big Data pie, it would be logical to expect them to get the largest returns. However, that is not necessarily the case. In fact, two of the functions with the smallest pieces of the pie – logistics and finance – expected the highest ROI on Big Data in 2012. Logistics managers expected Big Data to generate a 78% return, while finance managers expected a return of 69%. Both ROIs were far higher than the average ROI predicted by marketing executives (41%). (See Exhibit IV-2)

Exhibit IV-2: Which Functions Expect the Biggest ROI?

Q17b: Mean Percentage of Expected Returns on Big Data Investments by Function in 2012



We must note that the overall ROI average across all eight business functions that we surveyed was eight points higher than what the IT/analytics managers estimated across the company (54% versus 46%). We noted a similar pattern in the percentage of respondents who expected negative returns: Only about half the functional managers expected a negative return of 4.5% for 2012, compared to 8% for the IT/analytics managers. Therefore, depending on how you look at it, there may have been a 'halo' effect among functional managers in projecting ROI for the last year. Conversely, they may have had a more accurate estimate because of their greater knowledge of how their function used Big Data than the managers in a central IT or analytics group. In that case, the 46% ROI expected by the IT/analytics managers may have been understated.

Where Business Functions See the Gold in Big Data – and the Barriers in Mining It

We wanted to scratch below the surface to see what other studies had found with Big Data and most importantly, understand which business activities in these eight core business functions were thought to be the biggest beneficiaries of Big Data.

We asked managers in each business function to rate a number of activities for the potential benefits Big Data could bring to their function, using a five-point scale, with 1 denoting 'no benefits' and 5 denoting 'very high benefits'. We created this list after conducting extensive secondary research, combing through articles about companies and their Big Data initiatives, as well as public speaking presentations by those firms.[14] In total, we asked for ratings on 75 functional activities.

So where do these 643 companies see the gold to be in Big Data? Exhibit IV-3 shows the functional activities rated highest in terms of potential benefits. The eight most highly rated functional activities (and the function in which they operate) are:

- Identifying customers with the most potential or actual value (sales)
- Monitoring product quality (R&D)
- Monitoring product shipments (logistics)
- Identifying customer needs for new products and product enhancements (R&D)
- Identifying customers at risk of defecting (customer service)
- Determining marketing campaign effectiveness (marketing)
- Pinpointing where inventory is disappearing (logistics)
- Identifying spikes in logistics costs (logistics)

14 The list of activities for each of the eight business functions is by no means a complete one. As a result, we do not claim that this data represents the way large companies view all their opportunities for using Big Data in their businesses.

Exhibit IV-3: Where Companies See the Gold in Big Data

Q18a-h: Functional Activities Rated as Having Highest Potential Benefits from Big Data (Scale of 1-5, 1=no benefits, 2=minor benefits, 3= moderate benefits, 4=high benefits, 5=very high benefits)			
Overall Rank	**Functional Activity**	**Function**	**Potential Benefit (1-5)**
1	Identifying customers with the most value/potential value	Sales	**4.05**
2	Monitoring product quality	R&D	**4.02**
3	Monitoring product shipments	Logistics	**4.00**
4	Identifying customer needs for new products and enhancements to existing products	R&D	**3.95**
5	Identifying customers at risk of dropping our products/services	Customer service	**3.94**
6 (tie)	Determining marketing campaign effectiveness	Marketing	**3.90**
6 (tie)	Determining locations of inventory shrinkage	Logistics	**3.90**
6 (tie)	Identifying spikes in logistics costs and where and why they are occurring	Logistics	**3.90**
7	Analyzing customer behavior on the website to see which pages are most and least useful	Customer service	**3.89**
8	Improving employee retention by determining who is most likely to leave	Human resources	**3.85**
9 (tie)	Identifying cross-selling opportunities	Sales	**3.80**
9 (tie)	Determining marketing channel effectiveness	Marketing	**3.80**
10 (tie)	Determining optimal sales approaches	Sales	**3.78**
10 (tie)	Testing new product designs	R&D	**3.78**
11	Identifying patterns in customer complaints (both internal in the call center, and external)	Customer service	**3.77**
12	Determining optimal sales offers	Sales	**3.75**
13 (tie)	Tailoring marketing campaigns and promotional offers	Marketing	**3.73**
13 (tie)	Getting continuous customer feedback on products already in the market	R&D	**3.73**
14	Determining optimal sales messages (for example, from sales wins and losses)	Sales	**3.72**
15 (tie)	Determining customer value	Marketing	**3.71**
15 (tie)	Identifying trends in customer inquiries (for example, to better allocate service personnel to peak times)	Customer service	**3.71**
16 (tie)	Boosting energy efficiency	Logistics	**3.70**
16 (tie)	Identifying supply chain bottlenecks to speed flow of goods, materials	Logistics	**3.70**
16 (tie)	Determining appropriate inventory levels	Logistics	**3.70**
17	Pricing	Sales	**3.68**

This data suggests that activities which companies believe have the greatest potential to benefit from Big Data go far beyond marketing and sales. In fact, of the 25 highest-rated activities, there are an equal number in logistics and sales (six). In addition, marketing and customer service had four each. In other words, opportunities to capitalize on Big Data exist in numerous corners of a large, global company.

In each business function, Exhibit IV-4 below shows the two activities with the highest expected benefit, as well as the activity with the lowest expected benefit. We'll go into the details of each function in the next section of this chapter.

Exhibit IV-4: Functional Activities With the Highest Expected and the Least Expected Benefits from Big Data

Q18: Two Highest and One Lowest-Rated Functional Activities in Potential Benefits from Big Data (Scale of 1-5, 1=no benefits, 3=moderate benefits, 4= high benefits, 5=very high benefits)				
Business Function	*Activity with Highest Expected Benefit*	*Activity with the Second Highest Expected Benefit*	*Activity with the Lowest Expected Benefit*	*Expected 2012 ROI from Big Data*
Marketing	Determining campaign effectiveness (3.9)	Determining channel effectiveness (3.8)	Monitoring and improving the 'offline' customer experience (3.14)	**41%**
Sales	Identifying the highest-value customers (4.05)	Identifying cross-selling opportunities (3.8)	Determining which customers to avoid (3.43)	**54%**
Customer Service	Identifying customers at risk of dropping the product/service (3.94)	Analyzing website customer behavior to improve site (3.89)	Monitoring customer usage of products to detect manufacturing or design problems (3.40)	**56%**
Manufacturing/ Operations	Product quality/defect tracking (3.37)	Supply planning (3.34)	Enabling mass-customization (2.75)	**42%**
R&D/Product Development/Product Engineering	Monitoring product quality (4.02)	Identifying customer needs for new products and enhancements (3.95)	Conducting 'open innovation' – soliciting new products ideas through the web (3.51)	**48%**
Distribution/Logistics	Monitoring product shipments (4.0)	Identifying spikes in logistics costs (3.9)	Determining appropriate inventory levels (3.65)	**78%**
Human Resources/ Personnel	Improving employee retention by identifying most likely to leave (3.85)	Determining recruiting campaign effectiveness (3.5)	Identifying potential recruits who work outside the company (3.23)	**48%**
Finance/Accounting	Measuring risk (3.55)	Budgeting/forecasting/planning (3.55)	Identifying areas of external theft (2.87)	**69%**

Furthermore, we wanted to understand what managers directing these business functions perceived as the greatest barriers to achieving such benefits from Big Data. Thus, we processed our data on key challenges by business function. Exhibit IV-5 below shows the five biggest challenges for two of the eight business functions and compares them to what IT managers and analytics professionals said.

Exhibit IV-5: Comparing Key Challenges of Sales, Marketing, IT and Analytics Managers

A Stakeholder View of Big Data's Biggest Challenges				
	Functional Managers (Two Examples)		IT Management	Big Data/Analytics Professionals
	Sales	Marketing		
Challenge No. 1	Getting business units to share information across organizational silos	Being able to handle the large volume, velocity and variety of Big Data	Being able to handle the large volume, velocity and variety of Big Data	Being able to handle the large volume, velocity and variety of Big Data
Challenge No. 2	Being able to handle the large volume, velocity and variety of Big Data	Finding the optimal way to organize Big Data activities in one's company	Determining what data (both structured and unstructured, and internal and external) to use for different business decisions	Getting business units to share information across organizational silos
Challenge No. 3	Putting our analysis of Big Data in a presentable form for making decisions (for example, use of visualization/visual models)	Getting business units to share information across organizational silos	Getting business units to share information across organizational silos	Finding and hiring data scientists who can manage large amounts of structured and unstructured data and create insights
Challenge No. 4	Determining what to do with the insights that are created from Big Data	Reskilling the IT function to be able to use the new tools and technologies of Big Data	Getting top management in the company to approve investments in Big Data and its related investments (for example, training, etc.)	Determining what to do with the insights that are created from Big Data
Challenge No. 5	Building high levels of trust between the data scientists who present insights on Big Data and the functional managers	Understanding where Big Data investments in the company should be focused	Building high levels of trust between the data scientists who present insights on Big Data and the functional managers	Understanding where Big Data investments in the company should be focused

We next present the main areas of benefit (where the gold is) and the key challenges (what's preventing managers from getting that gold), function by function. Let's start with marketing.

Where Marketing Sees the Benefits – and the Challenges to Getting Them

The marketing managers we surveyed identified two areas with the highest benefits: determining campaign effectiveness and channel effectiveness. Close behind were tailoring marketing campaigns and promotion offers, and determining customer value.

A large telecommunications company that spoke to us illustrated just how important Big Data had become in determining customer value, pricing and understanding customer behavior. The firm has a database with a sizable sample of customers in one of its consumer business units. It has collected data on every customer transaction, starting from the time someone becomes a customer and including the offers they choose. The company monitors that data to sort out high-profit from low-profit customers, and how to convert the latter to the former.

Tailoring marketing campaigns can begin with figuring out which prospects should be targeted in the first place. An auto insurance company that we spoke with found that over 80% of the over 100,000 households it sent direct mails to had the wrong demographics and would never buy insurance from the firm. In fact, only 1% bought policies as a result of these campaigns. After using analytics to determine the right demographics, the company conducted a more targeted direct mail campaign and increased its response rate tenfold. That boosted revenue and cut marketing costs.

The company also used analytics to laser focus a marketing campaign following Hurricane Sandy in 2012, that let customers in the Northeast know the company was standing by to help them. The firm said analytics enabled it to create newspaper and broadcast ads for a more targeted set of media outlets, which reduced its campaign cost by 75% (as opposed to running a more scatter-shot marketing campaign). "Needless to say we couldn't get to every single customer. But with the money we spent, we probably hit 90% of our members with an ad for a much lower cost," said an executive at the firm.

Companies such as retail giant Sears Holdings Corp. have discussed publicly how they're using Big Data to set prices in close to real time. The $42 billion retailer has millions of products, about 4,000 stores and over 100 million customers. Three years ago, it took Sears eight weeks to run its pricing algorithms. In addition, the company was only able to tap a tenth of the potential of its pricing data because of the limitations of its aging mainframe computers. Since then, the company began using open source database Hadoop. By last summer, the firm was running its pricing algorithms in a week, and some in less than a day.[15]

At the opposite end of the scale (although still termed a 'moderate' benefit) were monitoring and improving customers' offline experience and discerning competitors' moves beyond pricing.

15 Rachael King, 'How Sears Uses Big Data to Get a Handle on Pricing', The Wall Street Journal, June 14, 2012. http://blogs.wsj.com/cio/2012/06/14/how-sears-uses-big-data-to-get-a-handle-on-pricing/

Exhibit IV-6: Areas of Greatest Benefits for Marketing

Q18a: Degree of Potential Benefits Big Data Could Generate for the Company: Marketing - Mean Summary

So what do marketers believe holds them back the most from leveraging Big Data? They felt the greatest challenge was technological: handling the volume, variety and velocity of the data. This isn't surprising, since marketing has usually been the least automated function in many companies. And marketing is far less worried about keeping data secure from internal or external parties who shouldn't have access to it.

We spoke with a number of executives – especially from the IT and analytics group – who said that marketing managers were daunted by the thought of having to deal with volumes, types and data speeds of a size they were not familiar with. The head of analytics at an insurance company told us that the marketers in his firm weren't initially prepared to create and execute marketing campaigns more rapidly. "It usually takes about three to four months to get a big campaign out the door," he said. "When I tell them they could do that in three or four weeks, they can't do it yet. They don't have the processes in place. They don't have the mindset to move that quickly. And, in some cases, they don't have the resources, or the experience of trusting the data, or being able to move that quickly."

Exhibit IV-7: Greatest Big Data Challenges for Marketing

Q23: Greatest Challenges to Getting Value from Big Data: Mean Summary - Marketing

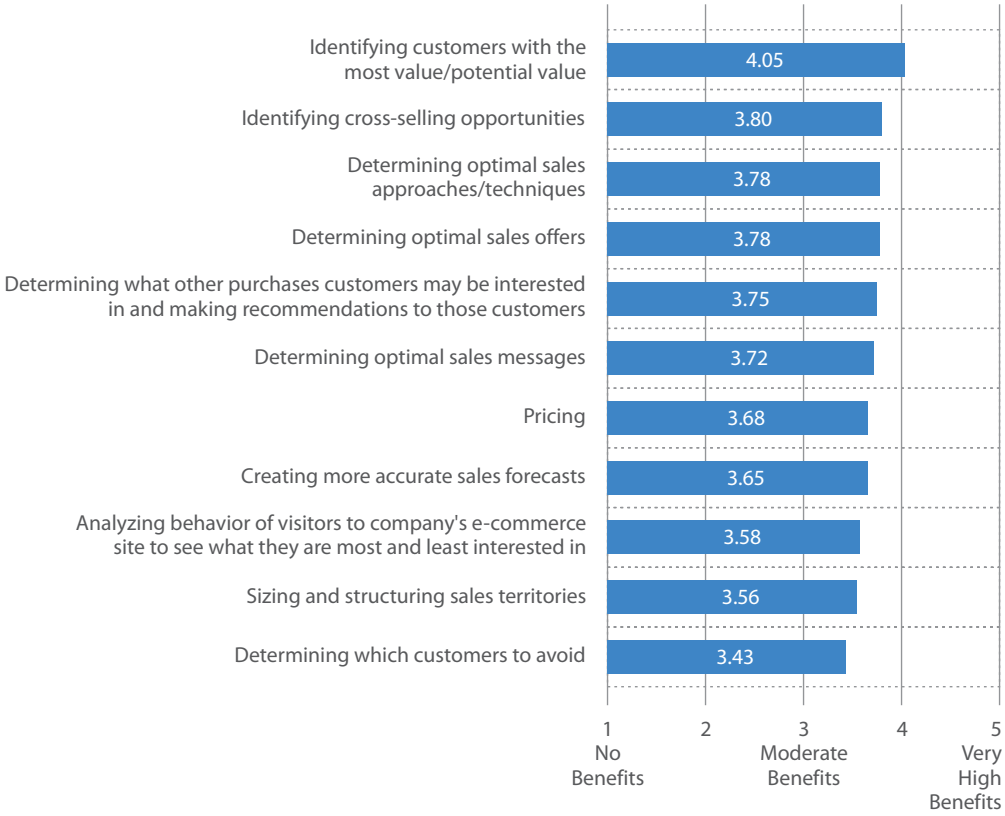
Where Sales Sees the Benefits and the Challenges

So where do sales executives see the gold in Big Data? The most popular choice was using it to identify the most valuable customers for the organization. That was rated much higher than any other activity. Finding opportunities to cross-sell products and services came in next. Identifying which customers not to sell to finished lowest on the list.

The case of technology maker and services provider Dell Inc. illustrates how companies are using Big Data to better identify their best prospective customers. In 2007, the company began working with a provider of cloud-based analytics for software that would help Dell's sales force pinpoint which prospects were better than others, according to a press article last December.[16] The software identifies the types of customer behavior most likely to lead to the purchase of different products (for example, signing an office lease). An executive from Dell told a reporter that the company's sales productivity, efficiency and revenue in Europe nearly doubled. This was achieved in part by halving the number of leads that marketers forwarded to the sales organization, which enabled sales to focus on the best prospects.

Exhibit IV-8: Areas of Greatest Benefits for Sales

Q18b: Degree of Potential Benefits Big Data Could Generate - Sales – Mean Summary



16 Rachael King, 'How Dell Predicts Which Customers are Most Likely to Buy', The Wall Street Journal, Dec. 5, 2012. http://blogs.wsj.com/cio/2012/12/05/how-dell-predicts-which-customers-are-most-likely-to-buy/

Sales executives say that their biggest obstacle to making the most of Big Data is the silo problem: getting business units (divisions, other functions, etc.) to share information with the sales function that could prove valuable in identifying trends and making decisions.

Similar to their colleagues in marketing, sales managers didn't rate keeping data secure very highly in their list.

Exhibit IV-9: Greatest Big Data Challenges for Sales

Q23: Greatest Challenges to Getting Value from Big Data: Mean Summary - Sales

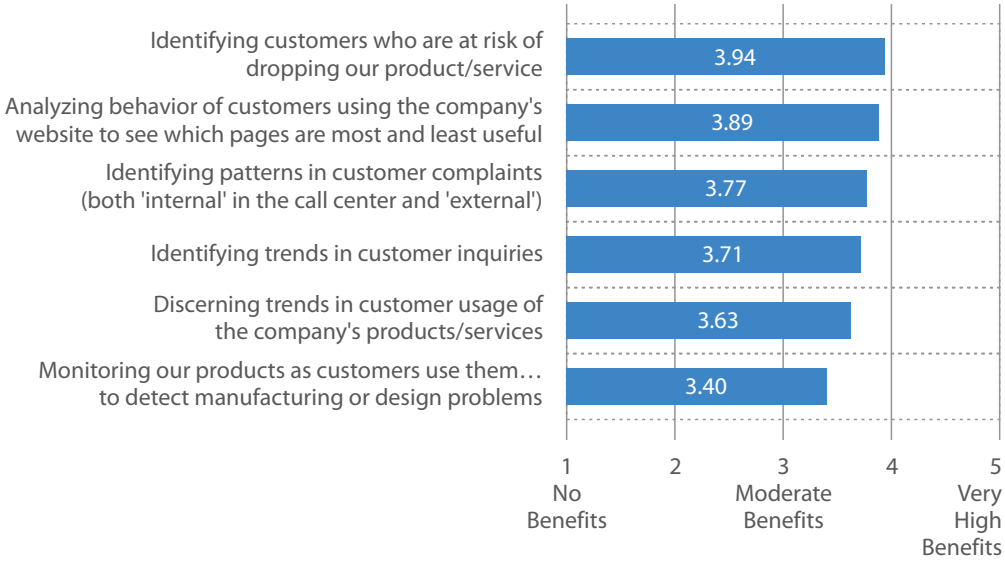
Where Customer Service Sees the Benefits and the Challenges

Customer service executives indicated that revenue was also an important potential benefit of Big Data for their function – even if it didn't directly generate revenue. They said the greatest benefit from Big Data was to identify customers at risk of discontinuing use of the company's offerings. Analyzing the way customers used the firm's website came next.

The activity rate lowest on this list of six – monitoring products that customers use to detect manufacturing or design problems – was nevertheless rated as one with more than moderate potential benefits (at 3.40). An excellent example of this is what Xerox Corp. does with the 1.2 million devices (copiers, printers, etc.) in use at customer sites. The company told *The Wall Street Journal* recently that it gathers terabytes of data from those machines every day and uses predictive algorithms to determine which ones may stop operating.[17]

Exhibit IV-10: Areas of Greatest Benefits for Customer Service

Q18c: Degree of Potential Benefits Big Data Could Generate - Customer Service Mean Summary



17 M. Hickins, 'The Morning Download: Xerox's 'Big Play with Big Data', Feb. 21, 2013. http://blogs.wsj.com/cio/2013/02/21/the-morning-download-xeroxs-big-play-with-big-data/

Customer service managers are in sync with sales executives on the top challenge of Big Data: getting data that exists in other business functions or divisions. Determining what data was useful to make different service decisions, and finding data scientists were also rated difficult challenges.

And like their colleagues in sales and marketing, service managers did not show great concern for data security issues.

Exhibit IV-11: Greatest Big Data Challenges for Customer Service

Q23: Greatest Challenges to Getting Value from Big Data: Mean Summary - Customer Service

Where Manufacturing/Operations Sees the Benefits and the Challenges

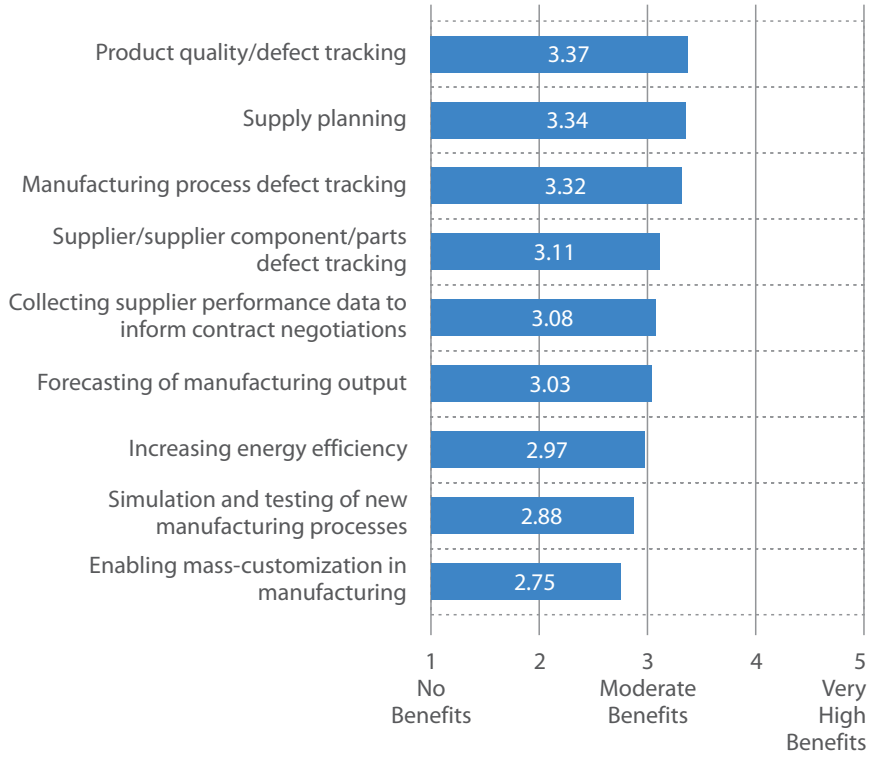
Manufacturing and production managers believe the greatest opportunities of Big Data for their function are to detect product defects and boost quality, and to improve supply planning. Better detection of defects in the manufacturing/production processes is next on the list.

A $2 billion industrial manufacturer told us that analyzing sales trends to keep its manufacturing efficient was the main focus of its Big Data investments. The company's products are largely engineered to order. Understanding the behavior of repeat customers is critical to delivering in a timely and profitable manner. Most of its profitability analysis is to make sure that the company has good contracts in place. The company says its adoption of analytics has facilitated its shift to lean manufacturing, and has helped it determine which products and processes should be scrapped.

They see far less opportunity in using Big Data for mass customization, simulating new manufacturing processes, and increasing energy efficiency.

Exhibit IV-12: Areas of Greatest Benefits for Manufacturing/Operations

Q18d: Degree Of Potential Benefits Big Data Could Generate - Manufacturing Mean Summary



So what holds them back the most from seizing these opportunities? At the top of the list (and we didn't see this ranked as highly by most functional managers) was building strong levels of trust between the data scientists and themselves. Perhaps the quants have a harder time relating to managers on the factory floor than they do with other, more 'white-collar' functions. Next on the list was figuring out what data to use for different business decisions.

Exhibit IV-13: Greatest Big Data Challenges for Manufacturing/Production

Where R&D Sees the Benefits and the Challenges

Surprisingly, R&D, product development and product engineering managers see monitoring product quality as the biggest potential benefit from Big Data. That indicates they may be more interested in using Big Data to protect established products than in inventing new ones. However, identifying customer needs for new products and enhancements to current products featured next on the list.

Consumer product companies such as Procter & Gamble, Starbucks Corp. and Kraft Foods Group Inc. have become known for using their websites to crowdsource new product ideas from consumers, university researchers and others. One of the pioneers of such 'open innovation' has been P&G. The company's 'Connect+Develop' program receives more than 4,000 ideas a year for new products and product enhancements, from around the world.[18] P&G says 50% of its new products have elements based on such external ideas, compared with 15% in 2000. R&D expenses have fallen, from 4.8% of revenue to 3.4%. Additionally, R&D productivity has skyrocketed by 60%.[19] P&G CEO Bob McDonald believes Big Data and analytics are instrumental tools in improving R&D productivity at the company.[20]

A very different illustration of how companies are using Big Data to propel product development comes from Netflix Inc., the online provider of movies and TV programming. The $3.6 billion subscription service for movie and TV fans has created a huge database of the viewing preferences of 27 million U.S. customers (and 33 million worldwide) since it began distributing movies through the mail in 1999. In going beyond the film distribution business to the business of providing original content, the company used its Big Data capabilities to anticipate what subscribers might want. The company collects enormous volumes and varieties of data: 30 million 'plays' daily (what customers are watching); 4 million consumer ratings of its content per day; billions of hours of streamed video, device and device location data, social media data, and other digital information, according to one technology publication's summary of a Netflix data scientist's conference presentation in June 2012.[21]

Such data told Netflix there was a large market for a remake of the 1990 British Broadcasting Corporation miniseries called 'House of Cards'. Netflix's chief content officer, Ted Sarandos, remarked at an investor conference that the company knows how many of its subscribers liked the star of the series (Kevin Spacey) as well as how many have rented the original BBC series. "You get a very addressable audience," he told the conference

18 From a P&G press release dated Feb. 7, 2013, 'P&G Connect+Develop Launches New Open Innovation Website', http://news.pg.com/press-release/pg-corporate-announcements/pg-connectdevelop-launches-new-open-innovation-website

19 P&G numbers cited in the McKinsey Global Institute study on Big Data, p. 69. http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

20 P&G CEO Bob McDonald's affinity for analytics was mentioned by Big Data guru Tom Davenport in his Wall Street Journal column, Feb. 13, 2013, 'P&G Finds a 'Goldmine' in Analytics'.

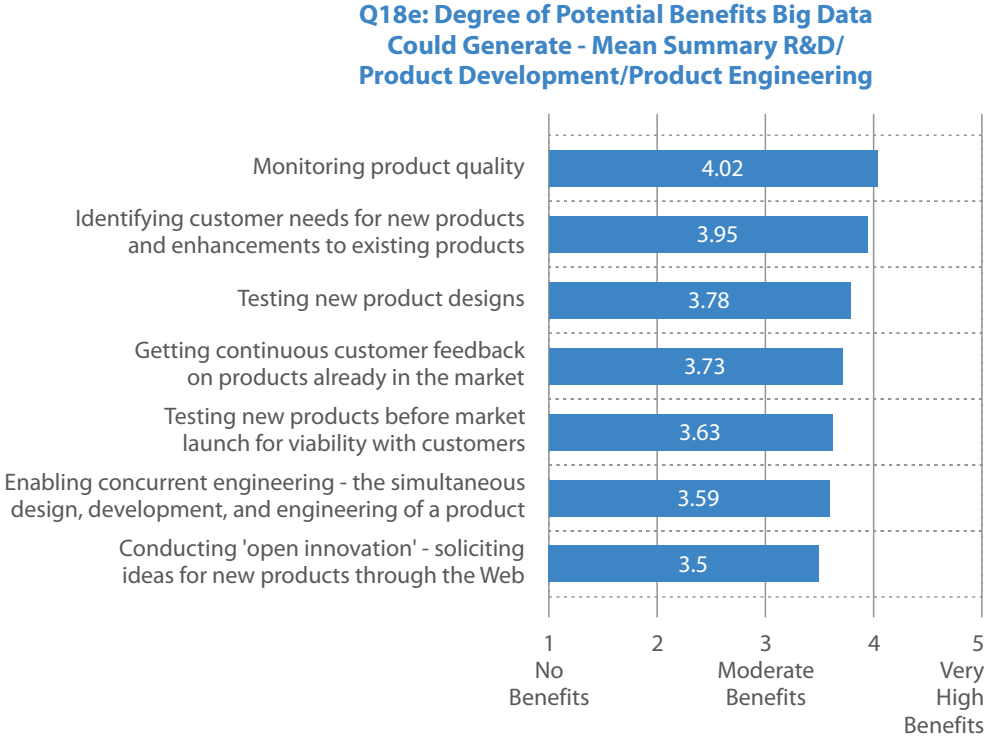
21 Derrick Harris, 'Netflix analyzes a lot of data about your viewing habits', GigaOM, June 14, 2012. http://gigaom.com/2012/06/14/netflix-analyzes-a-lot-of-data-about-your-viewing-habits/

participants. "Better than that, I know exactly who they are."[22] (One magazine said the company bid more than $100 million for the licensing rights to the BBC show for two years.) "We know what people watch on Netflix and we're able, with a high degree of confidence, to understand how big a likely audience is for a given show based on people's viewing habits," one Netflix executive told a journalist.[23] Netflix's ability to market the production to its customer base in a well-targeted manner was, most likely, another consideration in licensing the series.

In a short span, the series has become a hit. Netflix has said it is the most streamed Web content in the US as well as 40 other nations.[24]

Clearly, R&D managers see the potential in using Big Data to identify new product opportunities.

Exhibit IV-14: Areas of Greatest Benefits for R&D



22 From the Website Deadline, an article by David Lieberman, 'Netflix Unfazed by Growing Competition from Amazon', Sept. 13, 2012. http://www.deadline.com/2012/09/netflix-competition-amazon-ted-sarandos/#more-335796

23 From a Wired magazine story in November 2012: http://www.wired.com/gadgetlab/2012/11/netflix-data-gamble/

24 David Carr, 'Giving Viewers What They Want', The New York Times, Feb. 24, 2013. http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html?emc=eta1&_r=0&pagewanted=all

R&D managers rated finding data scientists to be their biggest Big Data challenge. Next on the list was getting managers in other functions to share data and handling the three V's of Big Data.

They perceived getting R&D managers to make decisions based on Big Data (rather than on gut) to be less of a challenge (though still moderate).

Exhibit IV-15: Greatest Big Data Challenges for R&D

Q23: Greatest Challenges to Getting Value from Big Data:
Mean Summary - R&D/Product Development/Product Engineering

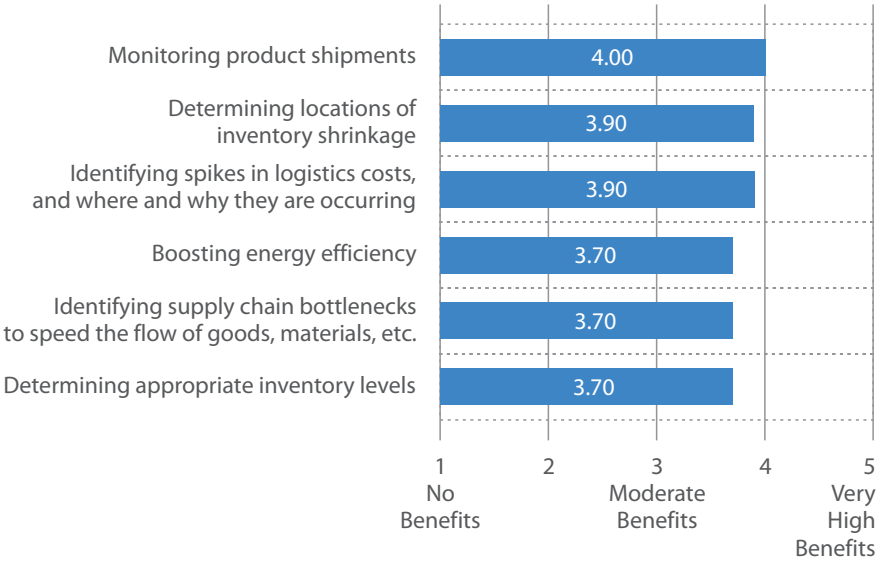
Where Logistics/Distribution Sees the Benefits and the Challenges

Logistics managers saw high value in all six areas that we asked them about. They perceived the highest value in monitoring product shipments (which is increasingly important in a world of terrorism, increasing theft, and increasingly demanding customers who want products to be available at the stores they frequent).

Also highly rated were finding places where inventory is disappearing and identifying spikes in costs.

Deere & Co., the $34 billion global manufacturer of agricultural and construction equipment, is experimenting with SAP technology that will help it assess its vast inventory of parts and other components in real time. That no doubt will help Deere send parts more quickly to places around the world where they are needed.[25]

Exhibit IV-16: Areas of Greatest Benefits for Logistics/Distribution



25 Wall Street Journal article, 'Deere's Technology Use Solves 'Insolvable' Problems', by Michael Hickins, Jan. 10, 2013. http://blogs.wsj.com/cio/2013/01/10/deeres-technology-use-solves-insolvable-problems/

Unlike their colleagues outside the logistics function, logistics managers are quite concerned about the theft of data. In fact, they're most worried about keeping it secure from internal parties. Second on their list of top challenges is figuring how to use the data to make decisions.

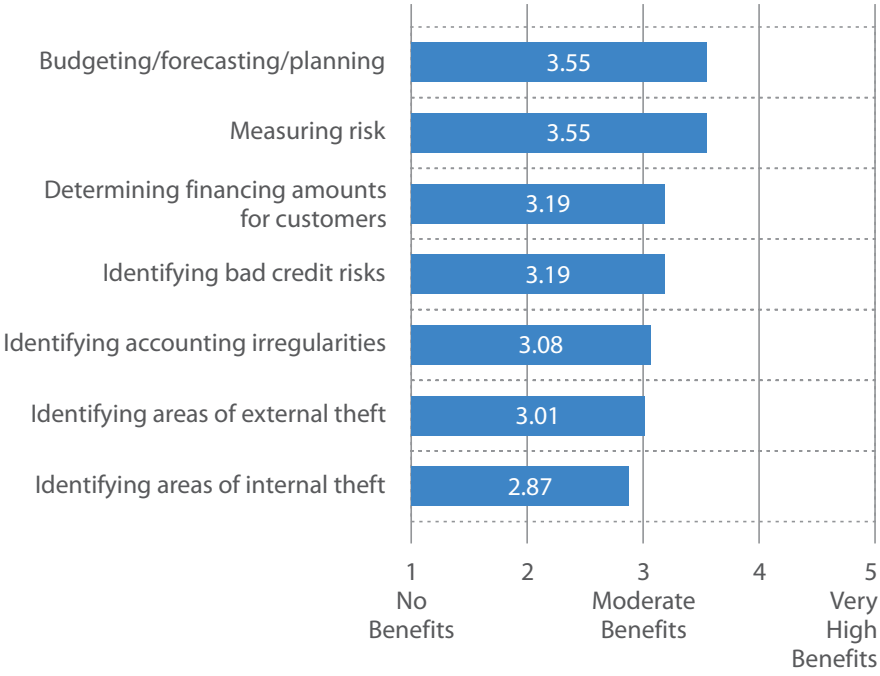
Exhibit IV-17: Greatest Big Data Challenges for Logistics/Distribution

Where Finance/Accounting Sees the Benefits and the Challenges

Finance and accounting managers see the most value in Big Data for two activities: measuring risk and improving budgeting and forecasting. They are less interested in the benefits of Big Data to reduce internal or external theft.

Exhibit IV-18: Areas of Greatest Benefits for Finance

Q18h: Degree of Potential Benefits Big Data Could Generate - Mean Summary Finance



What holds finance managers back from generating value from Big Data? Most of all, determining what data they need to make different business decisions. And next, getting other functions and units to share information.

Exhibit IV-19: Greatest Big Data Challenges for Finance

Q23: Greatest Challenges to Getting Value from Big Data: Mean Summary - Finance

Where Human Resources Sees the Benefits and the Challenges

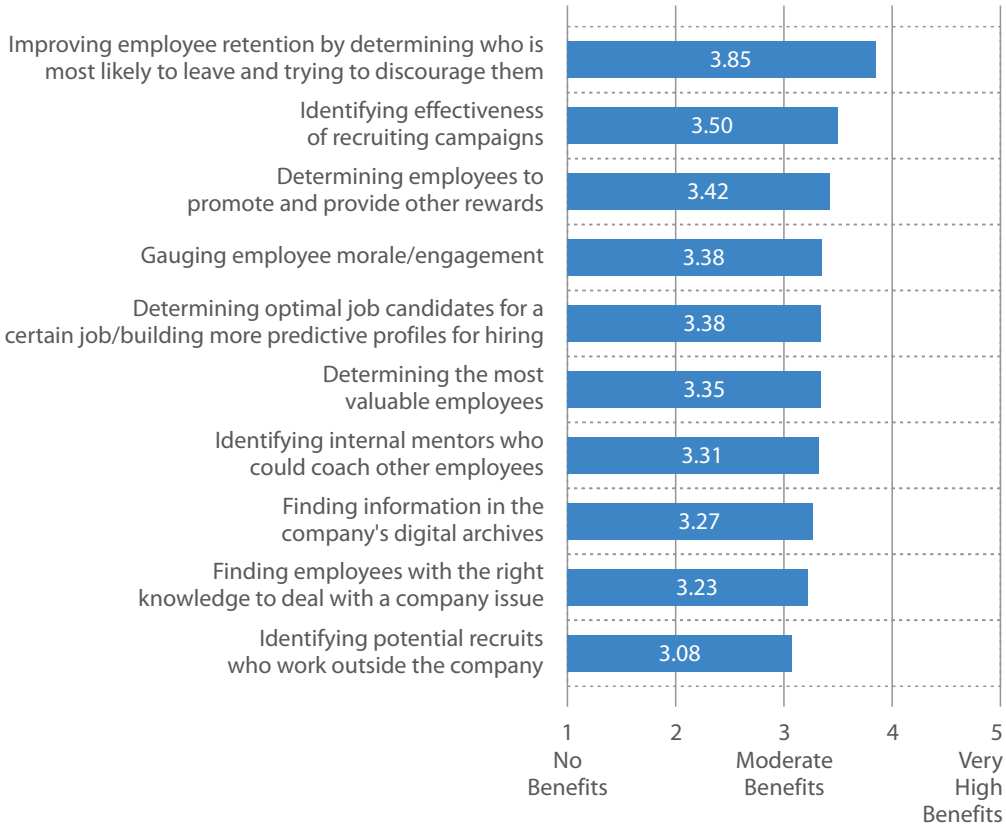
HR managers see Big Data's greatest potential in improving employee retention – determining which employees are most likely to leave and discouraging them from doing so.

Rated second was understanding the effectiveness of recruiting campaigns, followed by gauging of employee morale or engagement. This is an issue rated close to No. 1.

HR sees fewer benefits in using Big Data to identify people outside the company who could become valuable employees.

Exhibit IV-20: Areas of Greatest Benefits for Human Resources

Q18g: Degree of Potential Benefits Big Data Could Generate - Mean Summary – Human Resources/Personnel



HR managers said their biggest challenge with Big Data was understanding where to focus the investments in the HR function. Next on the list was finding data scientists.

Exhibit IV-21: Greatest Big Data Challenges for Human Resources

Q23: Greatest Challenges to Getting Value from Big Data: Mean Summary - Human Resources/Personnel

The Big Data Story of General Electric

How GE is Building Big Data, Software and Analytics Capabilities for an 'Industrial Internet'

General Electric Co. (GE) is known for making what it refers to as 'big swings' –large bets to grab the lead in emerging markets with enormous potential. The $145 billion manufacturer of jet engines, power plants, and locomotives believes Big Data and analytics is one of these markets. How big an opportunity GE sees with Big Data, how much and where it is investing, and how it's marshaling its expertise provide lessons for many companies, even much smaller ones.

In 2011, the Fairfield, Connecticut-based industrial giant announced the launch of a global software center and a $1 billion investment to build software and expertise for GE's version of Big Data analytics. "What we've done is centralized what I would call the real rocket science of Big Data," says William Ruh, vice president and corporate officer of the center. "These are people with deep expertise and experience. We use them to develop the very most complex capabilities and reusable components."

GE brought in Ruh from Cisco Systems in 2011 to lead the center, located in San Ramon, California. That's a short drive from Silicon Valley, an emerging epicenter of Big Data startups and established technology companies. Ruh's charter is to build a center that powers up software development and data science capabilities in GE's Big Data domain of interest – a niche it refers to as 'the industrial Internet'.

By this, the company means the torrent of digital data emanating from sensors and other digital devices embedded in machines such as GE's jet engines, turbines, trains, and hospital MRI equipment. Revenue of those GE businesses totalled $94 billion in 2011, nearly two-thirds of total sales. Harnessing such data would enable GE to help customers identify maintenance problems before they occur, improve fuel efficiency, and make other operational improvements that could add up to trillions of dollars in savings. To put it succinctly, it's about "making machines more intelligent and getting data to right people in real-time," as Ruh explains.

The digital data that GE could collect through such sensors would be gigantic. Ruh says a typical GE gas turbine generates 500 gigabytes of data daily. With 12,000 of them in service, that "certainly becomes Big Data," he explains. GE's gas turbines and other utility power equipment power a quarter of the world's electricity. But that isn't the only sensor data the company has in mind. GE is designing its next wave of airline jet engines to eventually capture information on engine performance for every flight. This will result in GE gathering a huge amount of aircraft engine information – in fact, more in one year than in the 96-year history of its aircraft engine business.[26] With this and other data it plans to collect and analyze, the word 'Big' may be an understatement when GE says "Big Data".

26 Jessica Leber, 'General Electric Pitches an Industrial Internet,'MIT Technology Review, Nov. 28, 2012. The estimate came from a GE researcher she talked to in the San Ramon center. http://www.technologyreview.com/news/507831/general-electric-pitches-an-industrial-internet/

Where GE Sees the Payback

In 2011, GE CEO Jeffrey Immelt announced the company would commit $1 billion to its analytics and software center over four years. That would put the company in the top 9% of our survey sample in Big Data and analytics investments.[27]

While a sizable amount, it's a small downpayment on what GE envisions as a $30 trillion opportunity by 2030. Using what it believes to be a conservative 1% savings in five sectors that buy its machinery (aviation, power, healthcare, rail, and oil and gas), a GE report[28] estimates the savings from an industrial Internet for these sectors alone could be nearly $300 billion in the next 15 years. Take the aviation industry: A 1% boost in fuel efficiency would put $2 billion a year into airlines' coffers.

A growing percentage of GE's business is services that support its industrial products – services that help customers use GE's machines more effectively and efficiently. Providing insights based on Big Data will be one more service offering.

"We believe this is a foundational change, in the same way that the consumer Internet has remade consumer industries in the past decade," Ruh told us. "Who does Walmart view as its competitor? Amazon? Who does American Express see as its competitor? PayPal. In the next 10 years, the changes that we saw in the consumer Internet will happen in the industrial world."

Why GE Has Centralized Big Data and Analytics

Today the center has a staff of about 300, up from just two people (Ruh and his executive assistant) in late 2011. Not all of them are in San Ramon, however. Ruh has employees located around the globe – Bangalore, New York and Cambridge–who report into the center. The company plans to staff up to about 1,000 people at the San Ramon facility, which is between San Francisco and Silicon Valley. Their work will support the efforts of another 9,000 GE software engineers who operate in its various product businesses globally.

Of the 300 currently in San Ramon, there are a number of 'hardcore data scientists' as Ruh refers to them. Why centralize these people? Ruh said it came down to three factors:

- **An acute shortage of talent.** "The first reason for centralization is that there is only a limited amount of talent – and I actually mean extraordinarily limited," said Ruh, who added that his center hires less than 5% of job applicants. "You can find a lot of people who are subject matter experts, basically analysts. You can find a lot of people who do business intelligence. You can find programmers. But the fact is that you're still probably not getting a whole lot out of your data other than reporting. The in-depth data science and deep analytics capabilities are held by such a small number of people."

- **Employee retention.** Data scientists at GE will need a career path if they're to stay for long. "They need to feel coupled and placed into a leadership program where they can get promoted based on their capabilities. When you put them in the businesses, their ability to grow, get promoted, and take on increasingly bigger roles is limited in many ways by the hierarchal structure of those businesses."

---

27 9% of the companies surveyed online (and which provided enterprise data) said they spent at least $250 million on Big Data in 2012.

28 GE research report 'Industrial Internet: Pushing the Boundaries of Minds and Machines,' by Peter C. Evans and Marco Annunziata, published Nov. 26, 2012. http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf

- **Reusability in technology.** "We need to build high-end capabilities [in solving deep technical problems]," Ruh said. "That cannot be built by each group over and over again. The reason is you can't find the talent, you can't maintain it, and so on. We believe this idea of reuse is going to differentiate the winners from the losers."

The center has begun to organize employees into disciplines such as machine learning, statistics and operations research. "They are very different approaches, and there isn't one approach that solves every problem," Ruh said.

The keys to success for GE software and analytics? Ruh boils it down to this: continually bringing a new portfolio of compelling service offerings. To be attractive to the market, those services must help airlines, electric utilities, hospitals and other customers tap GE's Big Data expertise and generate big savings and other improvements.

"In the end, our service offerings must foundationally improve how our customers manage, operate and maintain these big machines," Ruh concluded. "If we do that, we will be a leader in helping bring about this industrial Internet."

Views of the Visionaries: Learning from Leading Experts about the Technologies of Big Data

What companies are doing today with Big Data is not only limited by the acute shortage of analytical manpower – that rare beast now known as the 'data scientist' – they are also constrained by the capabilities of today's technologies in collecting, organizing, making sense of, processing, and presenting digital data in its many forms.

Nonetheless, billions of dollars in venture capital have been flowing to Big Data startups. That funding can be found in nascent companies like Cloudera, Palantir, Mu Sigma, Opera Solutions, VoltDB, and 10gen. A number of large, established technology companies have also turned the spending spigot on Big Data wide open: IBM, Oracle, SAS, and SAP, to name just a few. Companies like these are investing deeply in Big Data technologies. Their initiatives are aimed at enabling their existing software and hardware to take on the industrial-strength duties of Big Data and analytics – and create new software and hardware.

Thus, the technology of Big Data is evolving rapidly. To get some insights into what the technology makes possible today and what it may make possible in the near future, TCS interviewed two leading pioneers of Big Data technologies: Joseph Hellerstein of the University of California at Berkeley, and VS Subrahmanian of the University of Maryland.

Here are the highlights of those discussons.

"We're in the Early Days of Big Data – Like the Early 1900s' Era Before Washing Machines"

Joseph Hellerstein, Chancellor's Professor of Computer Science, UC Berkeley, EECS Computer Science Division

Joseph Hellerstein likens today's times for Big Data to the early 1900s before the advent of the washing machine. (The first electric washing machines began appearing in the first decade of that century.) Back then, women spent an average 60 hours a week manually washing clothes.

Cleansing Big Data is in a similar state, Hellerstein believes. He and several colleagues interviewed 35 analysts in companies across industries. They told them they spent 60% to 80% of their time on data preparation. "We're getting data from all over the place and it's not prepared for analysis or to be integrated with other data and analysis tools," he says. "The tools available are not designed for analysts."

Hellerstein sees a big opportunity in bringing data cleansing into the modern-day equivalent of the electric washing machine. He is founder and CEO of a data analysis tools startup called Trifacta.

"The Amount of Unstructured Data You Will Need to Have Will Be Vastly Larger Than Your Structured Data"

VS Subrahmanian, Professor of Computer Science and Director, Center for Digital International Government, University of Maryland

Subrahmanian has done extensive research and has developed technology on databases, artificial intelligence and optimization methods to track and forecast behaviors of terrorist groups, socio-cultural groups, health care and other areas. Much of this data is unstructured. He believes such unstructured data will be equally important to business – in fact, more important in the future than structured data. "The amount of unstructured data you will need to have will be vastly larger than your structured data," he says.

To make good decisions, managers will need both unstructured and structued data. But the problem today is that the accuracy with which software can make sense of unstructured data such as text is far lower than it is for structured data such as point of sale information. "People shoot for about 80% accuracy in text analytics. To go from 80% to 90% is a very steep curve," he says. "And you can spend a lot of time and money in trying to get there, but you might not."

Despite that, he believes the ability to make predictions on reading text with 80% accuracy is very good – especially if such data comes from several sources about the same phenomenon. For example, if a company knows a prospect likes jogging with 80% certainty, biking with 80% certainty, and soccer with 70% certainty, it can be pretty confident that the person likes sports. "That person is a good prospect for Nike," he says.

In the next 2-3 years, Subrahmanian sees companies having technology for developing effective metrics on the effectiveness of social media, software that will tie it "somewhat closely to ROI." He says search engines already have good technology to discern the impact of advertising on their own revenue. "They are less forthcoming in providing companies with data on the impact of search engine advertising on their revenue."

Nonetheless, he predicts this will change over the next two years. "By then, you will have the ability to see the impact of search engine and social media marketing on your revenue."

Implications and Recommendations

In the preceding sections, we described the landscape of Big Data initiatives in large companies in the world's four largest economic regions. In 12 industries, the data reflects a hotbed of activity in more than half the big companies in these regions. The median spending is not at all a trivial amount: $10 million per company. And our survey and interviews show that many companies are spending far more than that -- e.g., GE (which says it has committed $1 billion over the next four years). In fact, 23 (4%) of the 643 companies we surveyed said they spent at least $250 million on Big Data in 2012.

With this as the backdrop, it's no surprise that many companies have been asking us fundamental questions about Big Data and how best to organize for it. We hear four questions repeatedly:

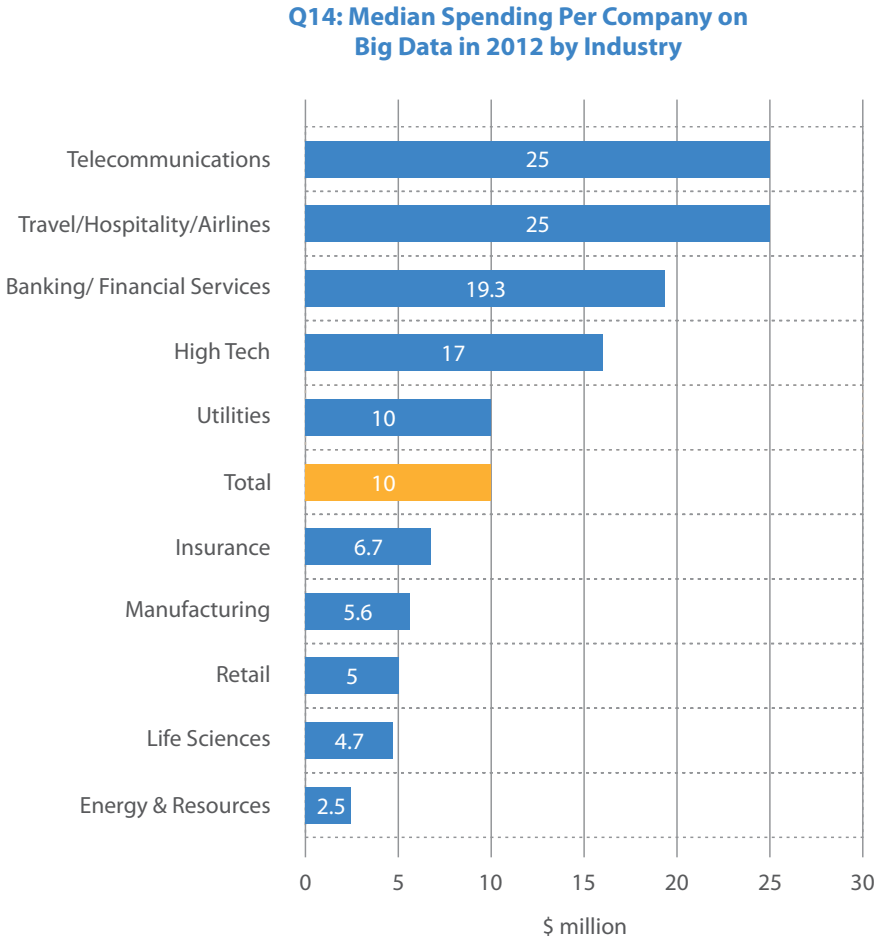
- Which industries should be investing more in Big Data?
- Where should companies focus their Big Data resources?
- What kinds of data should they collect and analyze to get the highest returns on their spending?
- What is the optimal way to organize Big Data activities?

In this section, we address these questions based on this research and our client experience. We draw especially on the survey data that compares 'leaders' (>50% ROI in 2012) and 'laggards' (25% or lower ROI).

Which Industries Should Be Investing More in Big Data?

There are at least two aspects to this question: a) which industries should consider making greater-than-average investments in Big Data and analytics?, and b) which other industries have smaller numbers of early adopters of Big Data and thus have many companies that need a wake-up call? To address the first question, we revisit the survey results on median industry spending on Big Data (Exhibit VII-1).

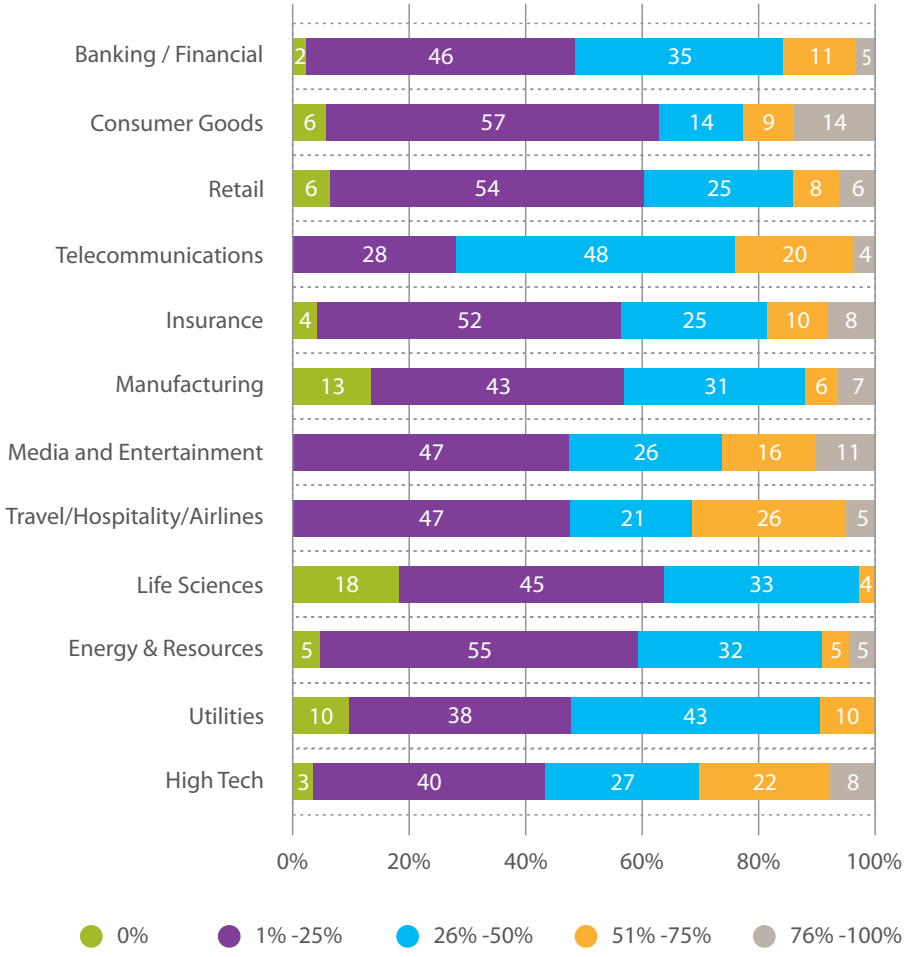
Exhibit VII-1: Which Industries Spend the Most on Big Data?



Four industries -- telecom, travel-related, high tech, and banking/financial services – told us they spend much more than the median on Big Data. So what do these industries have in common? For one, they have high numbers of customer interactions (especially online). In addition, according to our data, three of them generate higher than average percentages of revenue from Internet orders. (See Exhibit VII-2.)

Exhibit VII-2: By Industry, Percentage of Companies With Revenue From Internet Orders

Q2a : Percent of 2012 Revenue from Customers Who Order Products Over the Internet by Industry



While it may be easy to believe that if your company is not an Internet company that Big Data is far less important to you, it's not the case. The reason is that the way customers are interacting with companies online today goes far beyond ordering merchandise from their websites. The points of digital interaction abound: social media sites such as Facebook, Twitter, and Pinterest; mobile applications that consumers use on their smart phones; sensors on machines that report their vital signs back to the manufacturer; and many more.

If you simply look at the early adopters of Big Data – Internet companies that continually tinkered with their websites and mined their web viewer data to sell customers more of what they want – you're likely to miss the next wave of Big Data adoption. These adopters, exemplified by companies such as General Electric, will use it to pinpoint fraud, predict machine failures, and make a wide range of other product and process improvements in their bricks-and-mortar businesses.

Now let's look at the second question we posed: In which industries are a small minority of companies starting to pull ahead, thus industries where the majority need a wake-up call? Through our survey, we found a small minority of companies with projected ROIs on Big Data of more than 50% in four sectors:

- Consumer goods (in only 9% of these companies did functional managers report ROI >50%)
- Utilities (15%)
- Insurance (17%)
- Media and entertainment (19%)

That's a sign that the clear majority of companies in these industries have a long way to go to catch up to the leaders.

Another way to shed light on which industries may have some of the biggest opportunities with Big Data is to understand which may be betting the 'most bang from the buck' -- the sectors with the highest projected returns on Big Data[29] *despite* the lowest investments (i.e., Big Data spending as a percent of revenue). Exhibit VII-3 shows these to be energy & resources, life sciences, travel-related, banking, insurance and heavy manufacturing.

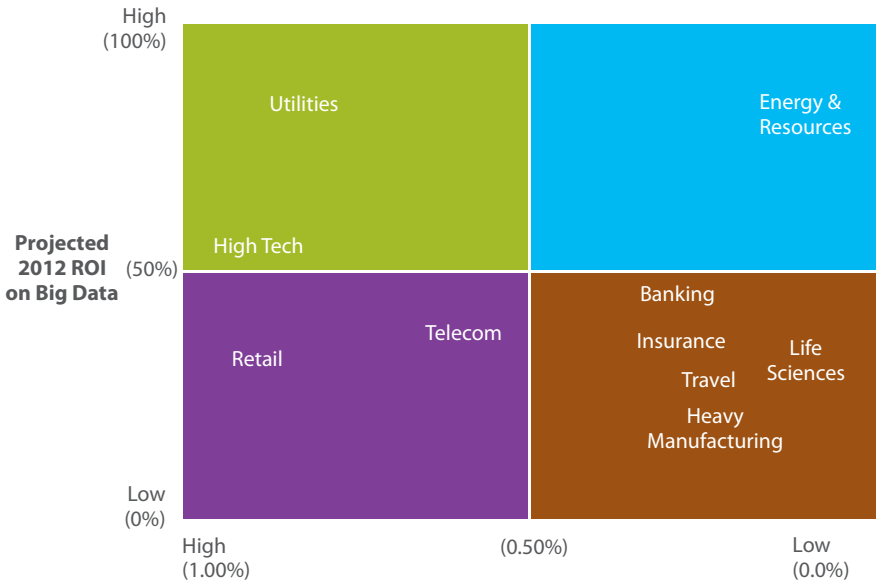
29 In this calculation, ROI is enterprise ROI, not functional ROI, as reported by the IT or analytics managers who answered the survey on behalf of the entire company, not a single business function.

Exhibit VII-3: Who's Getting the Biggest Bang from the Buck – and Should Strive for More?

Sectors with Highest Projected Returns and Lowest Spending on Big Data			
Industry	Mean Expected ROI	Big Data spending as % of mean industry revenue	Expected ROI divided by spending as % of mean industry revenue
Energy & Resources	**60.6%**	**0.07%**	**866**
Life Sciences	**35.3%**	**0.10%**	**353**
Travel/Hospitality/Airlines	**37.9%**	**0.24%**	**158**
Banking/Financial Services	**43.7%**	**0.28%**	**156**
Insurance	**38.9%**	**0.25%**	**156**
Heavy Manufacturing	**28.9%**	**0.23%**	**126**
Utilities	73.0%	0.81%	90
Telecommunications	37.9%	0.59%	64
High Tech	52.4%	0.83%	63
Retail	36.4%	0.85%	43
Media & Entertainment	NA	NA	NA
Consumer Goods	NA	NA	NA

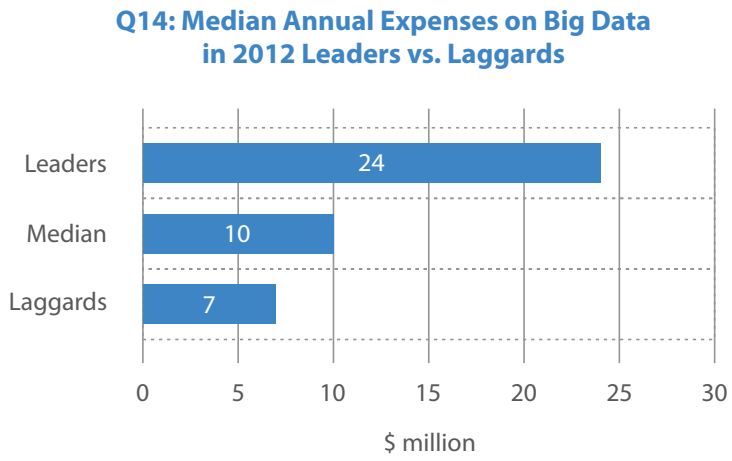
Plotted on a 2x2 chart, the data looks like this:

Exhibit VII-4: Plotting the Industries by ROI and Spending on Big Data



Opinions may vary on which industries need to take the lead in Big Data. But one fact is undeniable for all 12 industries that we surveyed: The leaders spent more than three times what the laggards spent on Big Data in 2012. The median spending for leaders was $24 million; for laggards, it was $7 million. (See Exhibit VII-5.) With Big Data, there appears to be a certain level of foundational investments necessary to play the game at a high level.

Exhibit VII-5: Leaders Far Outspend Laggards on Big Data



Where Should Companies Focus?

There is no one-size-fits-all answer to this question. It will largely depend on the industry, a company's competitive dynamics in that industry, and (most of all) where it needs to improve its performance with customers, among other factors.

Nonetheless, our data on functional activities rated highest in potential benefits provides guidance. Perhaps even more so does a comparison of the ROI leaders vs. ROI laggards on this data (see Exhibit VII-6). Six functional activities that ROI leaders rated much higher than laggards did are in marketing. Of the top 15, all but two were in marketing, sales and service.

Exhibit VII-6: Where ROI Leaders Pan for Gold (and Laggards Don't as Much)

Q18a-h : 15 Key Differences Between Leaders and Laggards in Functional Activities with Highest Potential Benefits from Big Data

Function	Activity and Average Score for All Companies	Leaders	Point Differential over Laggards
Marketing	Monitoring and improving the customer experience in offline channels (3.14)	3.78	0.95
Marketing	Marketing to consumers based on their physical location (3.24)	3.89	0.95
Sales	Sizing and structuring sales territories (3.56)	4.05	0.94
Marketing	Understanding competitors moves (beyond pricing) (3.22)	3.78	0.90
Marketing	Monitoring customer and market perceptions of the company (e.g. analyzing customer sentiment) (3.37)	4.00	0.83
Marketing	Identifying new geographic markets for existing products (3.24)	3.78	0.78
Customer Service	Monitoring products customers use them to detect manufacturing or design problem (3.40)	4.00	0.77
Sales	Determining optional sales approaches (3.78)	4.09	0.68
Marketing	Comparing prices with competitors prices (3.24)	3.56	0.61
Sales	Creating more accurate sales forecasts (3.65)	3.91	0.60
R&D	Testing new products before market launch for viability with customers (3.63)	4.00	0.60
Sales	Analyzing behavior of e-commerce site to determine what they are most and least interested in (3.58)	3.77	0.53
Marketing	Discerning customer needs for enhancements to existing products/services (3.27)	3.56	0.50
Marketing	Pricing (3.29)	3.78	0.49
Manufacturing	Simulation and testing of new manufacturing process (2.88)	3.07	0.46
Operations Customer	Identifying customers at risk of dropping our products/services (3.94)	4.44	0.44
Service Customer	Identifying patterns in customer complaints (both internal in the call center and external) (3.77)	4.00	0.44
Marketing	Monitoring and improving the customer experience on the web or mobile devices (3.33)	3.56	0.44
Marketing	Determining customer value (3.71)	4.00	0.41

In marketing, the leaders see the potential benefits of Big Data as being much higher than the laggards see them- in marketing to consumers based on their physical location, improving their experience of the company in offline channels, discerning competitors' moves and monitoring customer and market brand perceptions.

In sales, ROI leaders see much greater potential in using Big Data to size and structure sales territories. And in customer service, leaders see much more potential in using Big Data to monitor customer usage of products to detect manufacturing and design flaws. This is where companies such as GE and Xerox see the greatest potential of Big Data.

Again, where a particular company should focus investments depends on the dynamics of its customer and competitive situation. We have found it helps to think about the opportunities in three broad categories:

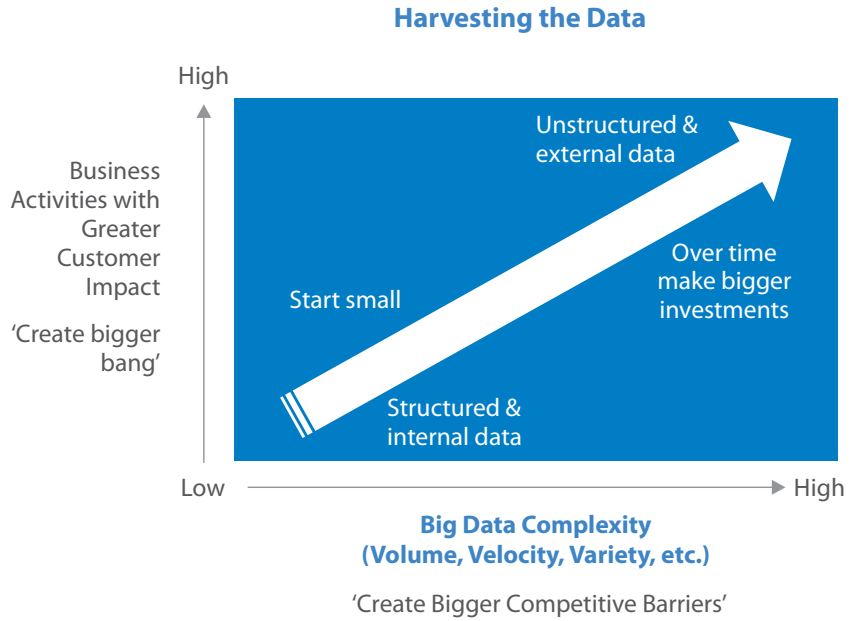
- **Revenue**: These investments will most often fall in marketing, sales, service, and R&D/product development
- **Risk reduction**: Finance and supply chain operations are full of opportunities to detect risk, from spotting customers whose credit rating is likely to suffer to transportation routes and locations that face heightened theft risks
- **Operational effectiveness**: These investments enable companies to improve the way they engage with customers and make every interaction more effective. This camp includes activities that may not directly lead to more revenue but which are important to keeping customers satisfied. Consider investments such as getting to the root of customer service problems, manufacturing process flaws, and supply chain bottlenecks.

Companies that create a portfolio of initiatives in these three categories and then implement them over time will reduce the chances of overlooking good opportunities, especially less obvious ones that don't directly contribute to revenue. (Don't forget that the functions with the highest Big Data ROI were finance and logistics – back-office operations that might be easy for the data scientists to overlook.)

To prove that Big Data has value, companies that are just starting to make serious investments should focus their portfolio on acute pain points. Those that can be solved by analyzing structured and internal data are particularly good to focus on. This is because structured data is typically much easier to gather and process; internal data mostly likely already resides in a company's information systems. Having data that is easy to manage allows a company to cut its teeth in a more controlled environment. Gaining these skills will help it deal later with more complex types and sources of data.

As the returns on early Big Data initiatives become clear, it becomes easier to get funding for more elaborate Big Data initiatives. (See Exhibit VII-7.) These may require unstructured or semi-structured data, as well as data from outside the company.

Exhibit VII-7: Where to Begin (and End) with Big Data



Over time, as a company's Big Data initiatives prove their value, it will be far easier to secure funding to address bigger, more costly problems -- ones that may very well require new technologies and new people. These initiatives have the opportunity to produce a 'bigger bang' – a higher top line or larger cost reduction. But they could also require larger investments.

At the same time, these initiatives can create bigger competitive barriers: data scientists who have built company-proprietary algorithms for crunching data; home-grown processing and analytics tools designed for industry-idiosyncratic data challenges (For example, in velocity, variety and volume); and possibly data that competitors have not collected. GE sees a $30 trillion opportunity in the industrial markets it serves for companies that can use Big Data and analytics to boost the efficiency of aircraft, power turbines, trains, and other costly industrial equipment. The Internet company whose website improvements have lifted revenue by hundreds of millions of dollars over the last few years provides an excellent example of this. Employing 70 Big Data analysts to recommend changes to the website is, of course, a major expense. However, the return appears to be one of far greater magnitude.

The path over time is for Big Data investments to provide both 'bigger bang' and 'bigger competitive barriers'. We'll discuss more about how to provide the latter in the next two sections.

What Data Should Companies Collect?

Just like the question about where a company should focus its Big Data investments, the question about which data it should collect depends on the company and the problems it wants to solve. Again, no one-size-fits-all approach exists. Nonetheless, two inescapable trends in Big Data provide guidelines for the overall types of data that companies should be looking to gather:

- **Unstructured data**: Many great insights to be derived from Big Data are likely to come from such sources as digitized video and audio, sensor data, email, documents, and the unformed text that fills Facebook, Twitter and other social media websites. Digital data from sensors and other remote devices attached to products – a GE aircraft engine or a Xerox copier, for example – enable companies to track their products long after they've been delivered to customers. This provides a major opportunity for companies that can collect unstructured sensor and other data from their products. Consider the airline industry. An estimated $284 billion is wasted annually in the airline industry by inefficient fuel management, unscheduled aircraft maintenance, flight delays and other issues. Imagine how a company like GE (whose engines are bolted onto 53% of the world's wide-bodied planes[30]) that can help customers cut their fuel and maintenance costs, will, no doubt, boost market share and increase revenue.

- **External data**: This data exists outside the information systems of a company. It's in the hands of customers, third-party data providers, suppliers, social media sites and other sources. Getting a fuller picture of customers requires companies to collect external data. Imagine a chain retailer that can get real-time data on customers who are in motion (based on their cellphone location) and within five miles of their stores (the telecom companies have this data); who are about to make a major purchase (such as replacing a refrigerator, the records of which are owned by the manufacturer); and who are most influenced by reviews in *Consumer Reports* (which are, of course, in *Consumer Reports'* online archives). For that chain retailer to get the customer to visit their nearest store and make a $2,000 purchase, it needs to tap external data held by the telecom company, the appliance manufacturer, and Consumer Reports – and this, within minutes or perhaps even seconds, before the customer moves to another store. The retailer's internal data is not nearly enough to influence the consumer's purchasing decision.

The leaders that we identified from our survey respondents – the companies with the greatest expected ROI in 2012 on their Big Data initiatives – are far more likely to recognize the value of both unstructured and external data than are the laggards. (See Exhibits VII-8 and VII-9.) On the dimension of structure, 55% of leaders' data is unstructured or semi-structured vs. 45% of laggards. And 37% of leaders' data is external vs. internal. For laggards, that percentage is 26%.

---

30 GE market share of the aircraft engine business is from market researcher Ascend, as cited here: http://www.bloomberg.com/news/2011-06-15/ge-sees-record-jet-engine-output-on-fuel-efficient-model-demand.html

Exhibit VII-8: How Leaders Differ From Laggards in Usage of Structured and Unstructured Data

Q8 : Mean Estimated Percentage of Structured, Unstructured and Semi-Structured Data, Across All of the Company's Big Data Initiatives Leaders v Laggards - IT / Big Data

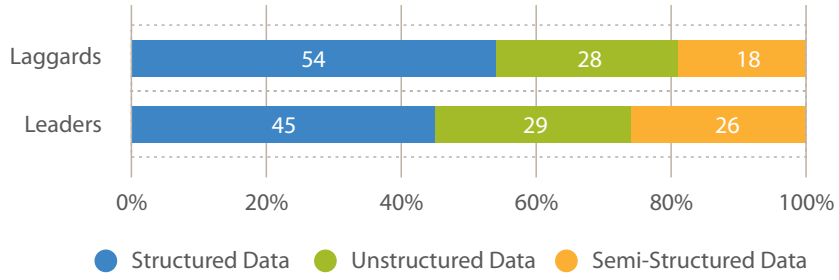
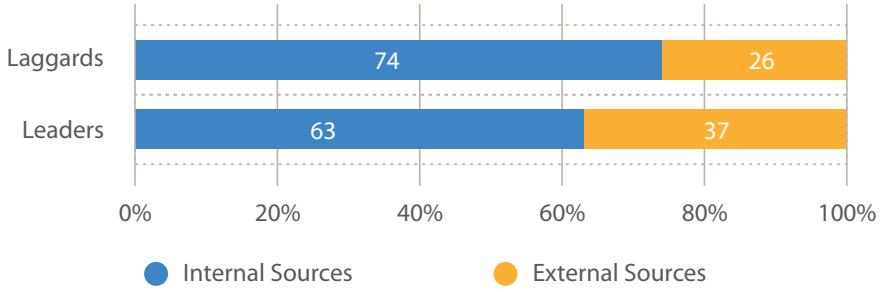


Exhibit VII-9: How Leaders Differ From Laggards in Usage of Internal and External Data

Q9 : Mean Estimated Percentage of Data that Comes from Internal or External Sources, Across All of the Company's Big Data Initiatives Leaders v Laggards - IT / Big Data



These differences can be seen in the importance that leaders and laggards place on internal vs. external and structured vs. unstructured data. (Exhibit VII-10.)

Exhibit VII-10: The Importance of Types of Data for Leaders and Laggards

How Leaders and Laggards Rate Importance of Data (Scale of 1-5, 1= no importance, 5= highest importance)		
	Leaders	**Laggards**
Internal data	3.84	3.73
External data	**3.50**	**3.12**
Structured data	3.67	3.83
Unstructured data	**3.19**	**2.65**

What is the Best Way to Organize Big Data Activities (and the People Who Do Them)?

The history of how companies have adopted information technology shows time and again how they had to first centralize a new technology and the skills to master it. Big Data skills – particularly data-driven analytics abilities found in the best data scientists – are in acute shortage today.

Over time, however, as a new technology becomes easier to use and people learn how to use it, the technology and the people who use it are shifted from the corporate IT function to business functions.

Yet Big Data is fundamentally different from other technologies. Companies today need a single view of each customer – not multiple versions (one held by sales, a slightly contradictory one held by marketing, a third one by customer service, and a fourth by the accounting department). It won't be easy to break down the organizational silos that prevent a company from using all its information about customers (or suppliers, for that matter).

We see two separate but related aspects to the 'organizational silos' issue:

- How to get divisions and business functions to share data that they often fiercely protect. This is an issue that likely requires a mandate from the top of a company, as well as a CTO, CIO or chief analytics officer with the political skills to carry it out.
- How to knit the data together technologically since it will be in different kinds of database management systems and be beset by other technical disparities that can make it a nightmare to integrate. There has been a shortage of technologies that can extract and merge big volumes, varieties and velocities of Big Data from different sources. However, this is changing. For example, Sears has broken down its organizational data silos in an Analytics Center of Excellence and is using the Hadoop open source data process platform to house huge volumes of data. It plans to sell analytics services to other companies.[31]

While a company must be able to tap and knit together siloed data to get a more comprehensive picture of its business problems, it still may not know the answer to the question of how to organize the people who analyze that information. Should they reside in the business functions that need the insights? Should the analysts operate in the IT function, where much of the data often exists? Should they live in a separate analytics department – a 'center of excellence' of the type that GE, a large Internet company, and other firms have launched?

The companies that have created large and well-funded analytics centers of excellence gave us of several reasons for doing so. One was preserving the data scientists' independence -- their ability to provide unbiased advice to functional managers about how to run their businesses. The manager of 70 analysts at a large Internet company told us that the key to helping the company increase revenue by hundreds of millions of dollars through numerous (and ongoing) tweaks of its website was extracting the analysts from the company's product units and centralizing them. "There was a heavy bias back then for analytics to confirm what the product units were doing," he said. Taking Big Data analysts out of the product units has led to better insights – ones that product unit managers might not have wanted to hear from their underlings.

Centralizing a company's Big Data analysts has another big benefit: giving them an attractive home within a big company. That was one reason why General Electric set up its $1 billion analytics center of excellence. The analytics manager from the Internet company said the same thing: having a center of excellence provides a more attractive career path. Even more important, he said, was to get analysts sharing methods far more freely than they did in the past. "Centralization got us to talk together," he said. "That was the special sauce. It didn't happen overnight. But within a year, we really had improved our analytics skills."

Our survey findings of leaders and laggards points to the need for centralizing at least some of the analytics staff. Some 37% of leaders put their analysts in a separate Big Data group compared with 23% of laggards. And 37% of leaders put people who process the data in a central Big Data group vs. 19% of laggards. (See Exhibit VII-11.)

31 From an April 2012 article in the Chicago Tribune: http://articles.chicagotribune.com/2012-04-24/business/chi-sears-looks-to-market-its-data-management-service-20120424_1_generation-of-database-technology-data-management-sears-holdings

Exhibit VII-11: How Leaders and Laggards Organize Their Big Data Staff

Where Leaders and Laggards Put Their Big Data Professionals		
	Leaders	**Laggards**
Those Who Analyze Big Data		
In IT function	42%	45%
In business functions that use the data	**21%**	**32%**
In a separate Big Data group	**37%**	**23%**
Those Who Process Big Data		
In IT function	**47%**	**69%**
In business functions that use the data	14%	11%
In a separate Big Data group	**37%**	**19%**

From our interviews with executives, we find that the optimal organizational structure must accomplish the following:

- It enables data scientists/analysts across a company to collaborate closely, share and improve one another's methods and technologies
- It provides them with a clear internal career path (thus, not forcing them to move into a business function to move up in the company)
- It allows them to provide objective insights on decisions that business unit and functional managers need to make – insights that are not watered down or censored by the business before they are delivered
- It helps analysts learn enough about the business operations that increasingly need data-driven insights without becoming trapped by traditional thinking
- It helps data scientists/analysts build trusting relationships with business unit and functional managers to increase the chances that they act on Big Data

We believe the optimal structure for a company's Big Data analysts is to create a center of excellence -- and have other analysts operating within divisions and business functions. The task of the central analytics group would be to build the deepest, most rigorous data science capabilities – skills that enable the company to get to the truth of what's really happening in its business and what it should do about it. The analysts in the business functions may not need such deep quantitative skills. But they need to be experienced enough to know how to put in place the recommendations that the central group makes for their functions.

Our research suggests that the optimal way to accomplsh this is to create a center of excellence for a good portion of a company's Big Data analytics. The center must report to or near the top of the company (to be objective and not be influenced by a particular business unit or function). However, to help data scientists better understand the business functions that need their insights, regular interactions with functional managers and tours of duty in those functions or units will do much to strengthen relationships -- even if the reporting lines are to the analytics center.

Taking the Big Data Game to the Next Level

While Internet companies may have jumped to the lead on Big Data, many other industries need to follow quickly, even those whose business is not threatened by Internet firms (present and future). The growing number of companies whose customers purchase their products and services over the Internet (and increasingly from mobile devices) will have a distinct competitive advantage if they can incisively analyze customer behavior on their sites and other data – and act on it quickly. Amazon.com Inc.'s ascension to a $61 billion in revenue over 20 years and Netflix's decimation of Blockbuster Entertainment show how companies with superb Web data and analytics capabilities can elbow aside traditional players that operate too much on intuition.

Yet Amazon, Netflix and other companies are also showing what can happen when a company possesses much deeper insights on customers based on their digital habits: it can get into the product business itself (e.g., Netflix's "House of Cards" TV series). Companies that don't use their analytics to see the next great product or service opportunity run the risk of letting analytics-savvy competitors trump them in product innovation.

But bricks-and-mortar companies that don't compete against Internet businesses such as GE's aircraft engines and turbines divisions believe they have an immense opportunity to use the Internet and Big Data (especially unstructured data) to keep improving their products and help customers get more value from them. Companies that sell big-ticket purchases (to consumers *or* businesses), whose products' performance must be frequently monitored to ensure they work, have a great opportunity. They can turn data that had been 'external' (collected by customers) into 'internal'.

By applying Big Data in the right places in the organization, centralizing and nurturing talent, and building bridges to functional managers who need data-driven insights to make superior decisions, companies will greatly raise the odds of keeping up in a world in which digital data-driven decisions become the norm, not the exception.

Research Approach and Survey Demographics

The topic of Big Data is a fast-moving one. So for years to come, there will always be something new to explore with Big Data. Consider also that the technologies for Big Data -- for collecting, storing, processing, analyzing and presenting enormous volumes and types of digitized data in formats that inform – are evolving rapidly. That's in no small part due to the venture capital that has rushed to startup companies, which has totaled $3.5 billion since 2010 in the U.S. alone.[1] And the subject is constantly evolving as companies become increasingly sophisticated with Big Data, using it to inform a growing number of business decisions: from who they should and shouldn't sell to, to uncovering the sources of supply chain bottlenecks.

The challenge for us in embarking on this study was saying something that hadn't already been said in the dozens of studies that appeared in 2012 alone. To explore new ground, we pored through these studies in November 2012, looking for the 'white space' – ground that other research had covered insufficiently or not at all, in our view. We had identified six overriding issues that we believe remained fertile and important ground for exploration:

- How much are companies investing in Big Data? How big a financial return are they expecting from their investments? Where in the organizations do they see the greatest returns on investments coming from?
- Just how advanced are they with Big Data? More precisely, what are companies and industries doing with their Big Data investments – that is, in which business activities and decisions are they focusing those investments?
- What kinds of data are they finding to be the most important (including structured vs. unstructured, and internal vs. external)?
- How are they organizing the professionals who process and analyze Big Data (especially by function), and what are the pros and cons of those reporting relationships?
- What are the biggest challenges of turning all this data (along with the technology and the people who use it) into information that changes the way the company makes decisions?
- What is the state of the technology today, where is it heading, and what opportunities will that bring to businesses?

In December 2012, we designed the instruments through which we collected our data:

- **Online survey**: We developed an extensive 23-question online survey of the Big Data practices at hundreds of companies around the world.
- **Best-practice interviews**: We constructed an interview guide to structure one-hour phone discussions with executives who are leading Big Data initiatives in their companies (12 in all), as well as executives at technology companies and professors who are at the leading edge of Big Data technology.
- **Secondary research**: We created a document that guided our secondary researchers on where to look for examples of companies using Big Data. Specifically, our literature

searchers looked in business, industry, and technology publications for companies that had been profiled for their Big Data initiatives; we also looked at conference presentations found online that were delivered by managers who are leading Big Data projects in their organizations.

By December 2012 and January 2013, our quantitative data collection, the online survey, was fielded in four regions and nine countries around the world: North America (the U.S.), Europe (UK, Germany, and Netherlands), Asia-Pacific (India, Australia and Japan) and Latin America (Mexico and Brazil). Research Now, a major research panel provider reaching 6.5 million people (including executives across numerous industries) in 38 countries, fielded the survey. Research Now closed the survey by the end of January after collecting more than 1,200 total responses.

Our research partner Bloom Group LLC secured and conducted 11 one-hour phone interviews with companies it determined were 'best practice' examples. Bloom Group found most of these companies through articles that had been published about them or conference presentations the companies had delivered. Three of the 11 interviews were with executives who participated in the online survey and were willing to participate in followup phone discussions. Two other interviews were secured by TCS.

To obtain these interviews and encourage these executives to share frank insights, most of these interviewees' and their companies' names are not mentioned. Additionally, we mention the names of companies found through secondary research. Throughout this document, we provide the sources of this publicly available information.

Defining the Phenomenon: What is 'Big' and What is Not?

Every Big Data study we have read has struggled with defining exactly what constitutes 'Big Data'. Is a chain retailer that crunches petabytes of structured transactional data from its point of sales systems processing 'Big Data'? What about a consumer products company that tries to discern consumer sentiments from the unstructured comments of millions of Twitter followers and Facebook fans?

Trying to define what is and is not big data is a slippery slope. A major research report from the McKinsey Global Institute in 2011 addressed the problem of determining what is 'big'. It said the term Big Data is "intentionally subjective and incorporates a moving definition of how big a dataset needs to be to be considered "Big Data." Thomas H. Davenport, a Babson College professor and acknowledged Big Data expert who has written prolifically on business analytics, views today's big data activities and technologies as the evolution of technologies and techniques for analyzing computerized information that began in the mid-1950s.[32] Davenport calls the period from this genesis to about 2009 'Analytics 1.0', and notes that it was characterized by a small number of data sources, structured data that came from within the company, and analytics activities that reported on prior events. Davenport says the 'Analytics 2.0' period began in 2010 (when the term Big Data was coined) and that it differed from the prior era in several fundamental ways:

32 Thomas H. Davenport, 'Preparing for Analytics 3.0', The Wall Street Journal, Feb. 20, 2013. http://blogs.wsj.com/cio/2013/02/20/preparing-for-analytics-3-0/?mod=wsj_ciohome_cioreport

companies used a lot of external data, the volumes of that data were much larger, and they were often unstructured (not filling neatly into the columns and rows of a database). He believes companies just three years later are moving into what he calls "Analytics 3.0" – a period in which companies use huge amounts of structured and unstructured data, sourced both internally and externally, but especially to provide *predictive* insights.

We decided that the task of deciding what is and is not 'Big Data' was highly subjective. As a result, we let the 1,217 survey respondents decide whether the initiatives they conducted in processing digitized data were 'Big Data' or not. We did offer a definition, however, from which they could decide whether they qualified to take the survey:

"The collection, processing and usage of large volumes of digitized data to improve how companies make important decisions and operate the business."

It is with this definition that nearly half the survey takers (47%) dropped out. Our research focused on the 53% of companies that remained – 643 in four regions around the world – that said they had Big Data initiatives.

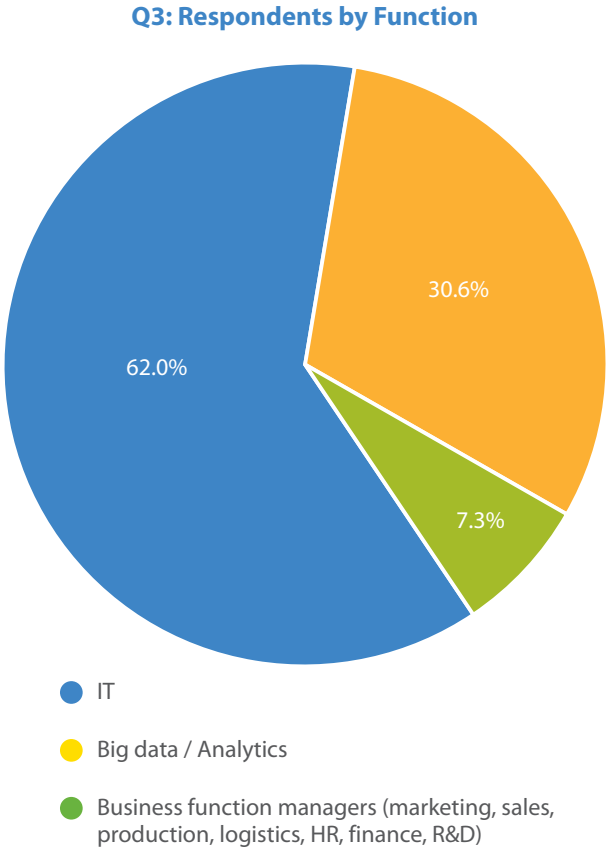
Survey Demographics: Getting a 360-Degree View on Big Data

Trying to understand what a large company is doing with Big Data depends upon whom you ask in that firm. Corporate IT management is likely to have a broader but more superficial view of what the company is doing with the technology. Managers of marketing, sales, service, and other business functions are likely to have greater knowledge of how Big Data is used in their areas. And managers running analytics groups (apparently a growing breed) have perhaps the best view of how an organization is using Big Data.

To get a better picture of how companies are using Big Data, we designed the study to collect data from IT, business functions, and analytics managers. Nearly one third were IT managers; 62% were from eight business functions (marketing, sales, service, production/manufacturing, logistics, research & development, finance, and human resources). And the remainder (7%) operated in analytics groups. (See Exhibit VIII-1.) In all, 88% either headed one of those functions or reported to the head of it.

We also wanted people in these functions who had intimate knowledge of their company's Big Data activities. The majority (58%) said they played supporting roles in this endeavor, and 23% played leading roles. The rest (19%) said they had no role but substantial knowledge about what their company was doing with Big Data.

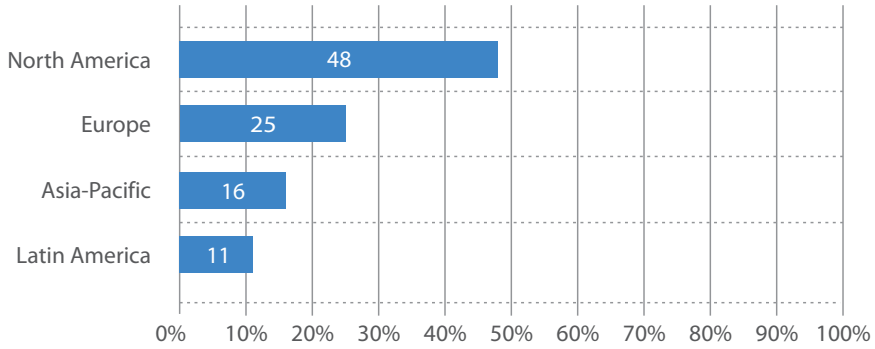
Exhibit VIII-1: Survey Respondents by Functional Role



We also wanted to limit ourselves to large companies. The clear majority of companies – 83% -- had revenue of more than $1 billion. In fact, the median revenue was $6.9 billion (while the average revenue was much higher, at $19 billion). Of the 643 companies that reported at least one Big Data initiative, the percent by regions can be seen in Exhibit VIII-2. Nearly half were from North America (that is, the U.S.). One quarter were from Europe, 16% were from Asia-Pacific and 11% were from Latin America.

Exhibit VIII-2: Survey Respondents by Region of World

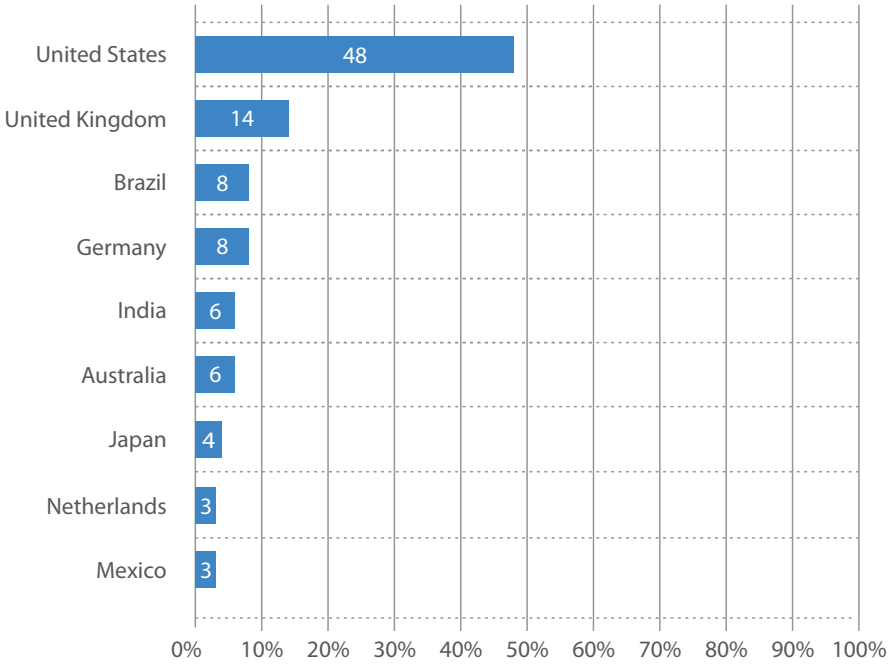
Q5-a: Where Companies are Headquartered (Region of World)



The breakout by the nine countries we surveyed can be seen below.

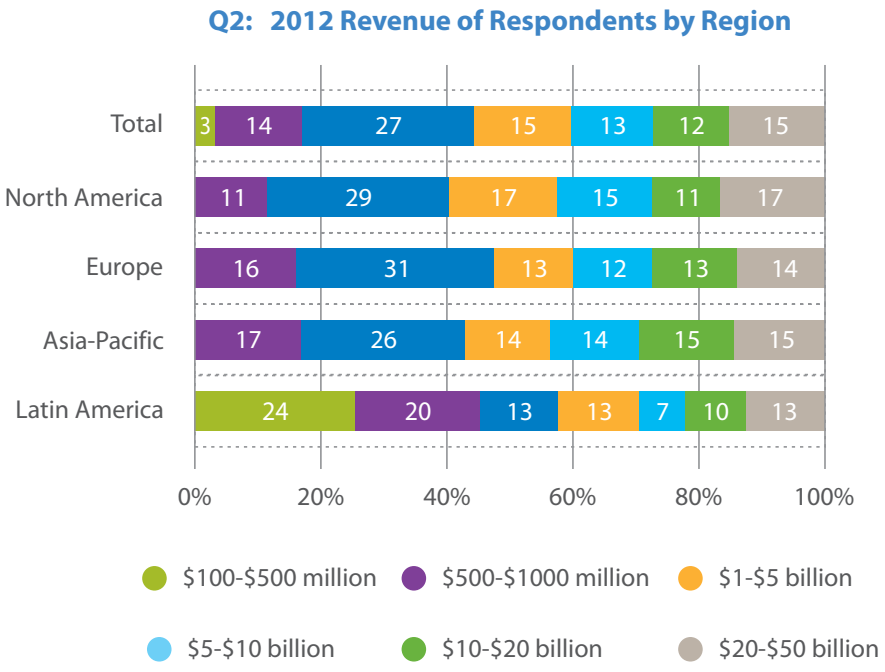
Exhibit VIII-3: Survey Respondents by Country

Q5-b: Where Companies are Headquartered (Country)



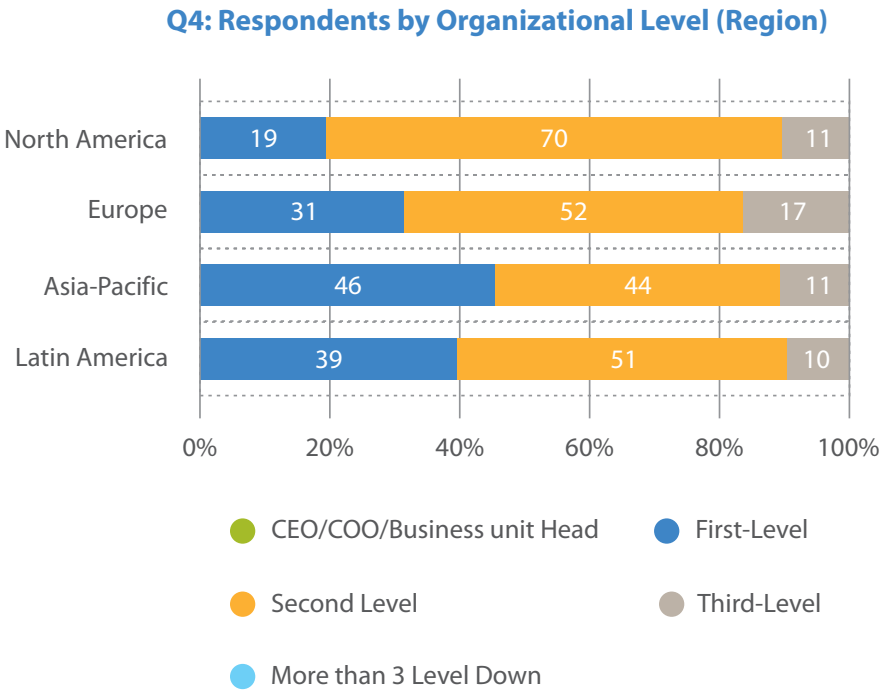
The survey population was tilted toward large companies. The majority had revenue of at least $1 billion in all four regions (see Exhibit VIII-4).

Exhibit VIII-4: Survey Respondents by Revenue



Most executives surveyed either ran a business function (marketing, sales, service, manufacturing/production, R&D, logistics, finance or HR), the IT function, or an analytics function – or reported to the person who ran one of those functions. (See Exhibit VIII-5).

Exhibit VIII-5: Survey Respondents by Organizational Level



Acknowledgements

We would like to thank many TCS subject matter experts for contributing their time and insights to this study: Ananth Krishnan, Vice President & Chief Technology Officer; Satya Ramaswamy, Vice President and Global Head of TCS' Mobility and Next Gen Solutions Unit; Frank Diana, Principal, Digital Enterprise Solutions; Kevin Mulcahy, Head of Business Analytics Solutions, John Propper, Head Digital Enterprise Solutions Architecture, and Mark Conzone, Solutions Manager.

We also thank John Lenzen, TCS Global Head of Marketing, for sponsoring and guiding the study, and Serge Perignon, Head of Marketing -TCS Global Consulting Practice, for overseeing and directing it. And lastly we thank the Bloom Group, our research partner.

i CB Insights and Orrick Herrington & Sutcliffe say that venture capital investments in big data-related companies has totaled $3.5 billion between 2010 and 2012 (including $1.4 billion last year), according to an article in PE Hub: http://www.pehub.com/187503/sv-angel-sequoia-ia-are-ventures-top-three-big-data-investors/


Kowloon

TATA CONSULTANCY SERVICES

Experience certainty.

About Tata Consultancy Services Ltd (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model™, recognized as the benchmark of excellence in software development. A part of the Tata Group, India's largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at **www.tcs.com**

IT Services
Business Solutions
Outsourcing