

# The Data Revolution: From Volume to Value

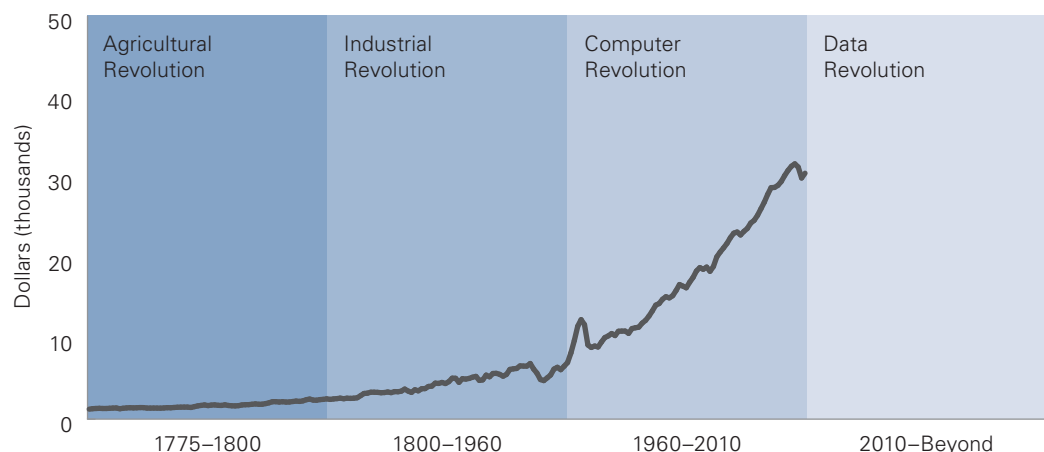
Enough data is already being created and stored today to stack DVDs from Earth to the moon and back—with some estimating this stack would reach halfway to Mars by the end of the decade.<sup>1</sup>

The world of data is exploding. Not long ago, technology’s “wow” statistic was the doubling of computer processing power every 18 months. Today, it’s the exponential rise of data. Each year, enough data is already being created and stored to stack DVDs from Earth to the moon and back—with some estimating this stack would reach halfway to Mars by the end of the decade.<sup>1</sup> We believe the integration of data analysis into industry and commerce represents a multi-decade transformation, one with vast potential to change the economy, as well as investors’ opportunities.

The Data Revolution is not the first time technology promises to transform our world. It may, however, be the fastest. In the long view of economic history, the first major technological shift forward was the Agricultural Revolution, which benefited humankind with more efficient farming methods and a higher standard of living. Next, in the Industrial Revolution of the late eighteenth and early nineteenth centuries, economic productivity doubled in about 150–200 years. Productivity skyrocketed in the Computer Revolution, with a three-fold gain in the half century starting in the 1960s.<sup>2</sup>

Now, at the onset of what we believe to be a fourth revolution, the world may be entering a period of even more rapid productivity gains thanks to the remarkable power which can be gained from data.

**United States GDP Per Capita: 1775–2010<sup>2</sup>**



<sup>1</sup> Source: Golub, Ben, “Enough data to fill a stack of DVDs to the moon (and back),” Computerworld, May 31, 2011.

<sup>2</sup> Source: “Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD,” Angus Maddison, University of Groningen and World Bank, Databank.

This information discusses general market activity, industry or sector trends, or other broad-based economic, market or political conditions and should not be construed as research or investment advice. Please see additional disclosures.

For illustrative purposes only. Source: GSAM. As of December 31, 2014.

## Big Data, Big Potential

Let's begin by splitting an oft-mentioned term—"Big Data"—into two parts. The "big" in "Big Data" refers to the huge quantity of information being produced. "Data," meanwhile, is simply a modern way of describing the age-old art of recording information, which has been taking place since the time of cave paintings. What makes Big Data different is the world of opportunity opened by the digitization of data.

Digital information is more storable, more shareable and more accessible than the technologies which preceded it. Take the example of Wikipedia, the online encyclopedia. Wikipedia's contents surpassed 2,600 printed library encyclopedia volumes as of 2010—and the service has grown substantially since then.<sup>3</sup> Besides sheer size, this digital-age encyclopedia is more real-time in nature and more accessible than predecessors—and the entirety of its data is storable on a flash drive the size of a human thumb. Its size and shareable nature only increases the importance of how we interpret and use this information source.

Digitization by itself, of course, is not enough to understand Big Data's transformation of the modern economy. For that, we must think about advanced analytics. Increasingly, Big Data enables businesses, governments and investors to glean more insight about the surrounding world. In other words, Big Data is less about possessing reams of letters, numbers and pictures than about how human beings use them. Here are a few examples of the way data analysis is driving innovation and competitive advantages across the economy.

- **Health care:** McKinsey & Co. estimates that Big Data has the potential to create \$300–450 billion in cost savings in the health care industry by optimizing patient care—through detection and remediation of ailments sooner and more effectively than today's technologies allow.<sup>4</sup>
- **Auto insurance:** Many insurers now offer discounted rates to customers who install tracking devices in on-board diagnostic systems—providing that they drive with care.
- **Retail:** Customer loyalty programs collect consumer data which is used in predictive models to segment customers, drive pricing and determine which promotions to offer.
- **Home automation:** Homes can be equipped with "smart" appliances which belong to interconnected networks, improving efficiency, comfort and safety.
- **Agriculture:** "Precision agriculture" uses high-tech equipment, cloud services and data analytics in pursuit of increased crop yields, higher food production, lower environmental impact and reduced costs.
- **Transportation:** Ride-sharing services are crowd-sourcing the taxi industry through a network of drivers for hire. Users of these services download a mobile application to their smartphones—rides are priced based on time of day, location and demand.

---

<sup>3</sup> Source: Wikipedia.

<sup>4</sup> Source: "Kayyali, Basel, Knott, David and Van Kuiken, Steve, "The Big-Data Revolution in US Health Care: Accelerating Value and Innovation," McKinsey & Co., April 2013.

## The Data Revolution in Investment Management

Today, there exists an enormous amount of data about every company—data which potentially can influence stock prices and other investment opportunities. As humans, we struggle to keep up with the torrent of information from hundreds or even thousands of sources. New data-analysis tools can help find signals in the noise. For example, satellite images of US retailers’ parking lots today can be used to gauge foot traffic in stores—helping predict same-store-sales before official data is released. Pattern recognition tools and natural language processing can be used to comb through terabytes of data in search of new investment opportunities. (A single terabyte can hold 17,000 hours of music, 310,000 photos or nearly 86 million pages of Word documents.<sup>5</sup>)

Let’s examine the example of the investment research industry, where these tools can be applied to the research published by securities analysts.

## The Velocity of Data and Markets

The analysis of structured data—numerical data sets and tables—is well-understood and generally available. But the analysis of unstructured data—news stories, social media posts, and financial filings, to give a few examples—is where considerable energy is focused today. We believe a world of possibility would open up for businesses which succeed in quantifying the qualitative: the ability to react instantly when a news story shows that an earthquake threatens their supply chain or a tweet indicates that employees should be evacuated from a geopolitical hotspot.

Satellite images of US retailer parking lots today can be used to gauge foot traffic in stores—helping predict same-store-sales before official data is released.

Tweets per day	Bytes of data created every day, IBM estimate	Appliances, devices and other “things” connected to the Internet	Assets under management, exchange-traded funds
<b>500 million<sup>6</sup></b>	<b>2,500,000,000,000,000<sup>7</sup></b>	<b>28 billion (2020 forecast)<sup>8</sup></b>	<b>~\$2 trillion</b>

Source: Goldman Sachs Asset Management, IDC, Hedge Fund Research, Bloomberg

<sup>5</sup> Source: Brown, Kelly: “A Terabyte of Storage Space: How Much is Too Much?” The Information Umbrella, University of Oregon, July 8, 2014.

<sup>6</sup> Source: Twitter, April 2015.

<sup>7</sup> Source: International Business Machines.

<sup>8</sup> Source: Goldman Sachs Global Investment Research.

## The Application of Data Analysis to Investment Research

The investment-research industry publishes hundreds of thousands of research reports each year—and a single investor generally can only read a limited number of them. Today, with more than 13,000 individual analysts active in the market, issuing an estimated 2.7 million monthly estimate revisions, deep and systematic analysis of these reports may be even farther than ever beyond the power of any single, unassisted human being (or group of human beings).<sup>9</sup>

Consider: We estimate that a person who reads about 200 words per minute<sup>10</sup> would need more than 865 hours to read through a single month's worth of reports.<sup>11</sup> This means he or she would be reading continuously for more than an entire month, without sleep.

In lieu of insomnia, data can be leveraged to analyze these reports at a deeper level—potentially uncovering new and previously hidden investment insights.

Here's an illustration. We believe investment research analysts may sometimes be reluctant to raise or lower a price target or rating too rapidly. Analysts may instead opt to reflect new views incrementally, by changing the tone and view of the text they write in their reports. Can there be a possibility of predicting some changes to price targets and analyst ratings? We believe the answer is yes—and the explosion in data analytics has created ways potentially to identify these cases.

One emerging application of data to sell-side analyst research is to use data analysis to spot a positive turn in the language an analyst uses before an estimate or ratings upgrade takes place. We believe that identifying an analyst's evolving views prior to the release of higher ratings potentially can provide investors advantages in the decision to buy or sell a stock.

In the reverse scenario, we believe analysts preparing to downgrade a stock sometimes may publish increasingly negative language. By utilizing a wide database of research reports, trends can be identified, potentially before they affect price returns.

We believe insight, not data, is what drives value. There isn't much value to data unless you can interpret it as information and act upon it.

## Big Data: More than Computers

As the investment research example shows, Big Data is not primarily about computers. The real benefit is what humans do with the insights Big Data can generate. Blindly relying on data analysis can mislead.

For instance, data can show that there exists a 99% correlation between per capita consumption of mozzarella cheese and civil engineering doctorates awarded over the course of a decade, or 99% correlation between divorces in Maine and the per-capita consumption of margarine.<sup>12</sup> But we know intuitively that these phenomena are not likely interrelated. We believe data is a tool to help investors—that is, humans, not computers—distinguish signals from noise. This task becomes more important in an era of vast data.

We believe insight, not data, is what drives value. There isn't much value to data unless you can interpret it as information and act upon it. Sometimes the analysis that turns data into insight is simple, but data users need to be careful: blindly relying on data analysis can be dangerous. We believe that effective use of data requires careful judgment and oversight by investment managers—computers and technology alone are not enough.

---

<sup>9</sup> Source: Thomson Reuters I/B/E/S.

<sup>10</sup> Source: Ziefle, Martina, "Effects of Display Resolution on Visual Performance," *Human Factors*, December 1998.

<sup>11</sup> Source: Estimate for single month of reports is sourced from Goldman Sachs Asset Management's analyst report database.

<sup>12</sup> Source: Vigen, Tyler, "Spurious Correlations," accessed April 20, 2015.

## Implications for Investment Management

What does the Data Revolution mean for investment management? We believe the implications are significant. Big Data can enable investment managers to see hidden connections and relationships between companies, including across industries. We believe data analytics can enable investment managers to see these hidden relationships faster and sooner than market participants who are not similarly equipped—leading to a potential advantage in selecting investments.

One example is the way that Big Data can be used to supplement or even supersede the study of supplier and customer relationships. Traditionally, a manager seeking to forecast a company's sales performance can watch suppliers for order surges, lags or bottlenecks in search of insight into current-quarter or future sales. The same manager can study consumer-sentiment indicators and same-store sales in attempt to understand sales prospects. Today, data has broadened the opportunity for information advantage substantially. Managers today can study hundreds or even thousands of other relationships which create the potential for new insights into company prospects.

For instance, investment managers can use data to identify “clusters” of companies which share overlooked relationships. These relationships could center on the way their businesses are affected by a rise or fall in oil prices, or a change in weather conditions. They could concern the manner in which new regulations affect industries besides the ones mentioned in news headlines, or the way that important implications for a firm in one country may be buried in a patent filing in another.

Each of these examples entails the linkage of data which cannot be downloaded from standard market-data terminals. We believe this “unstructured” data has the most potential to transform investment opportunities. Asset managers who master this unstructured data by investing in state-of-the-art technology and deploying the most extensive data sources, potentially can enjoy a competitive advantage over their peers.

## Conclusion: Here to Stay

We believe the Data Revolution is here to stay and that investors should recognize its potential to reshape the economic landscape. We believe the changes wrought by the Data Revolution will continue to ripple across industries—separating winners from losers, based on those who can best use data as an advantage—including in the world of investment management.

**Important Disclosures**

Equities may decline in value due to both real and perceived general market, economic and industry conditions. Bonds are subject to interest rate, price and credit risks. Prices tend to be inversely affected by changes in interest rates. Typically, when interest rates rise, there is a corresponding decline in the market value of bonds. Credit risk refers to the risk that the issuer of the bond will not be able to make principal and interest payments.

This material is provided for informational purposes only. It is not an offer or solicitation to buy or sell any securities.

THIS MATERIAL DOES NOT CONSTITUTE AN OFFER OR SOLICITATION IN ANY JURISDICTION WHERE OR TO ANY PERSON TO WHOM IT WOULD BE UNAUTHORIZED OR UNLAWFUL TO DO SO.

Views and opinions expressed are for informational purposes only and do not constitute a recommendation by GSAM to buy, sell, or hold any security. Views and opinions are current as of the date of this presentation and may be subject to change, they should not be construed as investment advice. Although certain information has been obtained from sources believed to be reliable, we do not guarantee its accuracy, completeness or fairness. We have relied upon and assumed without independent verification, the accuracy and completeness of all information available from public sources.

Views are as of April 20, 2015 and subject to change in the future.

This information discusses general market activity, industry or sector trends, or other broad-based economic, market or political conditions and should not be construed as research or investment advice. This material has been prepared by GSAM and is not financial research nor a product of Goldman Sachs Global Investment Research (GIR). It was not prepared in compliance with applicable provisions of law designed to promote the independence of financial analysis and is not subject to a prohibition on trading following the distribution of financial research. The views and opinions expressed may differ from those of Goldman Sachs Global Investment Research or other departments or divisions of Goldman Sachs and its affiliates. Investors are urged to consult with their financial advisors before buying or selling any securities. This information may not be current and GSAM has no obligation to provide any updates or changes.

Economic and market forecasts presented herein reflect our judgment as of the date of this presentation and are subject to change without notice. These forecasts do not take into account the specific investment objectives, restrictions, tax and financial situation or other needs of any specific client. Actual data will vary and may not be reflected here. These forecasts are subject to high levels of uncertainty that may affect actual performance. Accordingly, these forecasts should be viewed as merely representative of a broad range of possible outcomes. These forecasts are estimated, based on assumptions and are subject to significant revision and may change materially as economic and market conditions change. Goldman Sachs has no obligation to provide updates or changes to these forecasts. Case studies and examples are for illustrative purposes only.

**Past performance does not guarantee future results, which may vary. The value of investments and the income derived from investments will fluctuate and can go down as well as up. A loss of principal may occur**

Goldman Sachs does not provide legal, tax or accounting advice to its clients. All investors are strongly urged to consult with their legal, tax, or accounting advisors regarding any potential transactions or investments. There is no assurance that the tax status or treatment of a proposed transaction or investment will continue in the future. Tax treatment or status may be changed by law or government action in the future or on a retroactive basis.

© 2015 Goldman Sachs. All rights reserved.

Date of first use: September 18, 2015.

Compliance code: 169659.MF.MED.OTU.



