

The Potential of Big Data Analytics is Significant

Richard Arthurs | 21/07/2016



So why are so many businesses failing to find value in it?

By Richard Arthurs, CPA, CMA, MBA, CFE, CRMA, CIA, QIAL

Is value creation from Big Data analytics a contradiction? From my experience overseeing data analytic projects in the U.S. and Canada over the last 15 years, there are many critical things to consider to drive value from such analytics. This article identifies several reasons why Big Data analytics projects fail and why others succeed. In the near future the companies that get this right will leverage Big Data analytics to gain powerful competitive advantage.

To establish some perspective on how the quantity of Big Data is growing, consider the following statement by IBM in 2015: “Every day the world creates 2.5 quintillion bytes of data — so much, that 90% of the data on the planet has been created in the last two years alone.” This trend has been driven by factors such as the global usage of smart phones and tablets that allow video and photos to be instantly uploaded to social media, along with mass online video services such as YouTube. So how can we generate valuable return from Big Data analytics? The possibilities are unlimited, but many have failed at achieving the desired return. Here is a list of common reasons Big Data projects have failed and what must be considered to achieve a material return.

Not Knowing What Insight You Want to Gain and Why

I have learned that before you start any analytics project, you should always have a very specific hypothesis to test and verify an insight or conclusion you would like to achieve. If you do not have a value-focused hypothesis with a predicted potential outcome, you will often spend more time and money on the project than required. Analytic projects without a target hypothesis could also lead to a much higher percentage of false positives in the results. False positives are data analytic findings that look like valuable findings, but in actuality are not. These are often caused by anomalies in the data, such as two different customers with the exact same name.

A good example of a desired data analytics insight goal, is leveraging Big Data to optimize customer satisfaction. What if web sites could remember exactly what you like and do not like using on a site that sells goods or services, and the site automatically reconfigures the layout to optimize your personal experience when you log in. Data analytics can be used to capture this insight both on a macro level (comparing all customer segments) and the micro level (analyzing what is important within each customer segment). For example, airlines have for years tracked data on specific key performance indicators that have a high correlation with customer satisfaction issues, such as flights not on time, lost baggage, complaints, check-in

time, and benchmarking versus competitors. Frequent flier programs also collect demographic and loyalty data that can be integrated into this data analytic analysis. Airlines who utilize this type of Big Data analytics for monitoring customer satisfaction drivers are able to detect issues much faster than competitors who do not use analytics.

Not Having One of These Three Must-Have Capabilities

There are three critical capabilities that need to coexist to optimize data analytics success.

The first is having technical and data analytic systems expertise, with intimate knowledge of your data. There are several new systems that are much more user friendly than data analytic systems were a decade ago. These new systems have simplified how to use many sources and types of data. However, many companies do not really understand what data they have and whether it is accurate or has integrity. The best technical data analytics experts know how to get what data is sufficient to achieve the Big Data analytics target analysis, and then they will also know what will be the optimal system to achieve this.

The second critical capability is having enough relevant industry and business experience to connect the data to critical objectives, processes, and controls. Business acumen will help the data analytic leader identify the best data, to either verify that everything is correct, or identify issues that could be very costly to the business.

Lastly, and often an area that does not get enough consideration, is having an internal leadership sponsor who has influence with key decision makers. Most high-value data analytic projects can take time to verify the real findings and this usually results in the need for change management. This leadership champion must manage the expectations of key decision makers and maintain the trust and support when the change management required becomes hard to digest.

As these three capabilities improve with your data analytic knowledge, so will the ability to venture into higher levels of Big Data analytic maturity, such as predictive analytics. Some industries, like global consumer packaged goods companies, are using predictive analytics to optimize everything from where to put a product in a store to what are the optimal dates to promote a product. It can also be used to indicate what product or service innovation to focus on to give you the greatest return on investment. There may be 10 or more years of category management and consumer purchasing data by retail chain, by city, being utilized to support these analytics. There is so much data that it would be impossible to do this manually or without utilizing some form of Big Data analytics.

Companies like FED-EX use analytics to map out hot spots to map out physical

locations that either are bottlenecks in the delivery process, or when deliveries are late. This is a form of Big Data visual analytics for continuous monitoring that requires live links to multiple systems.

Not Understanding Your Data

You should not start your Big Data analytic project until you understand your data well. This requires an inventory of all data sources to be utilized, including awareness of the quality of the data. You should ask if the data is complete, accurate, and consistent. You should know if the data is structured or unstructured. If you are using data from various sources, you should know how to link the data sources, such as by using employee or project name and number. Also ask if the data fields are exactly the same in every source (July 4, 2016 vs 2016-06-04)? What format is the data stored in? How do you intend to access, filter, securely transfer and store the data? The more you learn, the easier it will be to know where to look for the greatest insight and determine how to complete the analytics without creating any data security issues.

Eventually companies that want to invest in Big Data analytics should have a data governance role to oversee data accuracy, completeness, security and sources. Every company has critical databases that must be accurate or they could corrupt the integrity of your Big Data analytic output. A few good examples of sacred databases are the employee master file and payroll data. When errors exist in these data files it can cause several issues, such as overpaying an employee, or paying a fictitious employee that was set up as part of a fraud scheme.

Putting Too Much Effort In to Having Perfect Data

It is often not possible to have perfect data and maybe not even cost effective to attempt to have perfect data. The required quality of data depends on what you need the data for. It is also possible to use data analytics to simplify the process of cleaning and reconciling data to achieve the required quality and integrity. For example, scanning data for inconsistencies and / or missing data. Analytics can even be utilized to make corrections. With a clear vision of what your highest priority data analytic initiatives should be, you can also determine what data will be critical to achieve this.

Often a small percentage of your data will be considered critical for your project. Ultimately, you never want to be in a position where data analytic conclusions are later found to be incorrect because of a lack of data integrity. In general companies have experienced a significant improvement in data quality and data analytic tools. It is becoming easier for end users to develop or modify visual analytic dashboards,

while only utilizing data you know has integrity. Eventually data integrity will just be the norm and every computer will have various data analytic tools being used on a regular basis. Eventually it will replace the usage of spreadsheets and maybe even PowerPoint.

Allowing the Scope of the Big Data Analytics Project Become Too Complex

The integrity of a data analytics project can be compromised even when the data quality and integrity is high. This can be caused by trying to correlate too many things to gather insight. This may cause complexity that can create misleading insight. It is always best to brainstorm on what are the most critical root causes or drivers and then ask what data would be most critical to focus on. Always start with the most-simple correlations between data sets that have a high probability to provide the insight you are looking for. A good example would be to ask lost customers for the reason why they have left and to continuously monitor the most common reason why customers leave using data analytics. This may provide an early warning indicator if a business is at risk of losing more customers.

Healthcare analytics can often become very complex. Everything from age, weight, body mass index, family history, habits, diet, amount of exercise, personal history and current prescriptions can be factored into one analysis. The permutations and combinations can often become complex and maybe even hard to place reliance on, not to mention the need to meet data privacy requirements. For example, it may be hard to blame a health condition on diet when that condition is pervasive in your family history. The main point is you should always try to keep your analysis simple unless you are confident the data correlation is strong enough to provide insight you can trust. Healthcare analytics is already exploring advanced Big Data analytics, such as using both predictive and prescriptive insight to optimize your personal health in the future.

Failing to Prepare for Potential Results

For many years, corporate internal audit departments have been seen as progressive leaders in the use of analytics for forensic investigation and assurance. Those who have used internal analytics for investigations have often exposed more than they expected in the analysis, such as fraud involving a senior executive. This is always a possibility and mature data analytic teams know how to manage these situations to minimize the risk of mismanaging such things as communication or evidence collection. What you want to avoid is being in a situation where you did not include the right people up front (human resources, legal, etc.) and ensured a well thought-out protocol would be followed to optimize confidentiality and protecting all parties until the evidence is verified.

Not Assigning Formal Accountability for Data Management

Never assume a data analytic visual dashboard will continue to maintain the integrity of insights after it has been built and tested. We have all heard the term “garbage in – garbage out” and this applies to live links to data analytics dashboards. When anything changes in relation to source data it can corrupt the data analytics visualization. It is very important that roles and responsibilities are defined to hold someone accountable to maintain the integrity of the data and the data analytics. For example, when material changes are made in the configuration of the source data systems, it can corrupt the integrity of the continuous monitoring analytics. Or, if an ERP system like SAP is reconfigured to add in new lines of business or an acquisition, it may change the source data. Even though data analytics systems have improved, they are only as good as the integrity of both the data and the quality in the design of the analytic script or tool.

Not Realizing Big Data Analytics Can Be Disruptive

Big Data analytics has already matured to see great work being done in predictive and prescriptive insight that has already eliminated jobs and will disrupt the security of current business models. Any repetitive project or task that exists today, that was created due to difficulties consolidating or cross referencing data because of legacy system limitations, may be an opportunity for data analytics to replace this work. Many internal audit functions are asked to repeat the same audits every year; over the next 10 years we will see data analytics be used to develop continuous auditing tools that will be handed over to management to self-audit. A good example of this is monitoring employee expense reporting. Eventually red flags will be sent to management in an exception report or visual analytic tool that will force management to address an issue when it happens. As internal audit hands these tools back to management, it will free up time for internal audit to focus resources on even higher-value audit work.

Systems will eventually take over sales departments, report creators, older less agile systems and business models. Businesses must embrace technology proactively to not become a victim of disruption.

For example, every December I receive a system-generated email reminding me to send flowers to my mother for her birthday. The email shows me a photo of the arrangements I sent in the past and even my mother’s feedback on each arrangement. It has a customer loyalty discount identified beside a button to quickly advance to a fully completed sale order based on personal data I provided in the past. For many years I received multiple phone calls instead of an email. I suspect the email and web based system may have eliminated one or more jobs at this flower shop. Is it possible flower growers may cut out the flower shops and deal directly with consumers? Could a business model be created for flower providers

that allows people to prepare and deliver arrangements from their homes (sort of an UBER for want to be florists)?

The Future of Big Data Analytics In Canada

Even though Big Data analytics can be challenging to execute, it has great potential for powerful insight. It will continue to grow in popularity as leaders see how it can enhance high-quality decision making. We will continue to see data analytic technology evolve and transform the way we live and do business. Many businesses will be able to create sustainable competitive advantage from Big Data. The insight gained from this will feed the development of valuable artificial intelligence and ways to become more accurate at predicting what will happen next. The financial and utility industries, in particular, have recently promoted how Big Data is being used to develop a much better understanding of their customer segmentation, demographics, customer satisfaction drivers, and even trends that may be able to predict future demand for goods and services. This is taking customer relationship management to a whole new level of sophistication. When integrated with artificial intelligence, systems may be able to know what people want even before they know what they want.

Albert Einstein once stated, “Imagination is more important than knowledge.” I think this quote applies to this new world of Big Data; it is not about waiting for the data to tell you what to do, it is more about using your imagination to develop new ways of using the data to gain value added insight.

About the Author: Richard Arthurs is MNP’s National Governance and Risk Management Leader. He brings more than 25 years of experience assisting complex global organizations with their governance, risk management, internal audit, IT audit, data analytics, ethics and compliance needs. To find out more about MNP’s Enterprise Risk Services contact Richard at richard.arthurs@mnp.ca