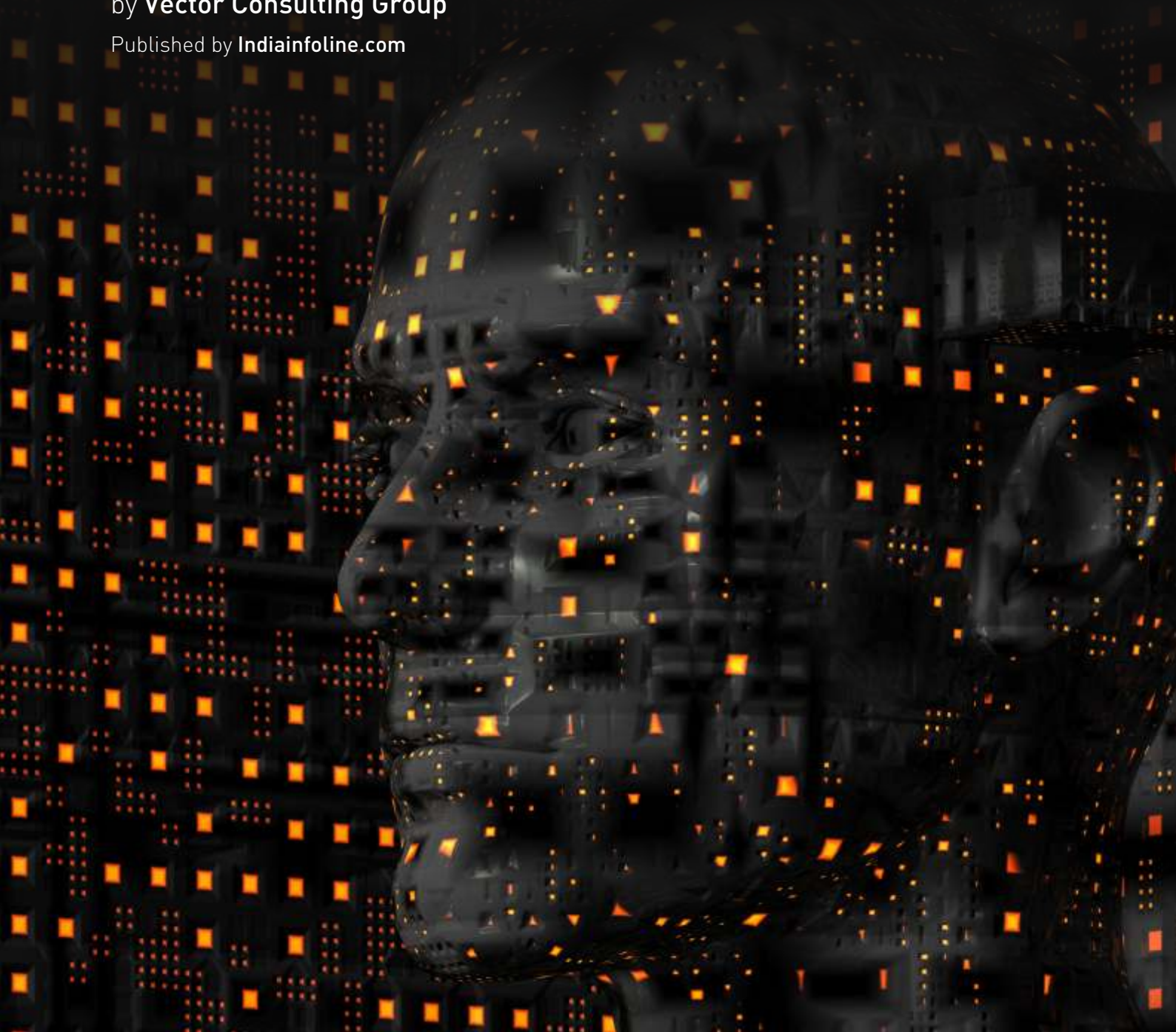# What enables effective decisions?

by **Vector Consulting Group**

Published by **Indiainfoline.com**

"Big Data" or "Big Intuition"

With cloud computing and ever reducing prices of storage systems, the available capacity to store data has gone up exponentially. At the same time, data sets, collated from varied sources, are also growing at an ever-increasing pace to the size of petabytes. As experts put it, this combined effect is a huge opportunity. With right technological support to store and analyse petabytes across sources, one can gain new insights in the field of business, medicine, e-commerce, intelligence gathering etc. There are many anecdotal evidences from data rich industries to back up the claim. This definitely sounds exciting.

But before jumping into the bandwagon, we need to ask a few critical questions –

- With increasing data over the years, have we been able to improve our decision-making in the field of social science?

- Can the ability of "big-data" technologies to provide new insights, eventually, replace the need to rely on human intuition?

If efficacy of data driven decision-making is proven by anecdotes, let us try to look for counterexamples. Nassim Taleb, in his latest book 'AntiFragile' highlights how US government was unable to predict the Arab Spring revolutions or even the financial crisis of 2008 despite investing billions in predictive analytics.

He argues that, in a physical world, we may be able to predict the trajectory of a rocket's flight but it is difficult to predict the rare events (he calls it the black swan types) in a non-linear complex system (where cause may not have proportionate effect due to feedback loops). The mathematical models will fail regardless of the sophistication or multiplicity of data used in the model.
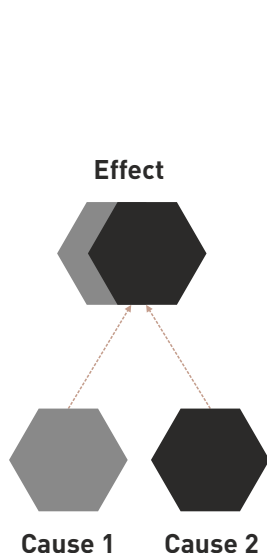
Linear Relationship between Cause and Effect
- Cause has a defined
  proportionate effect
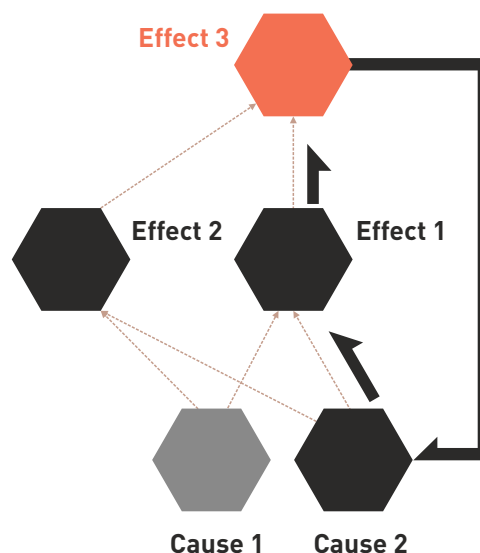
Non-Linear Relationship between Cause and Effect, with a feedback loop
- Cause can have a disproportionate
  or a multiplier effect

The way to test this claim is to use the predictive models in retrospective i.e., predict a past social event with information obtained from periods preceding the event. Most demand forecasting tools fail this test.

If this is the case, why do we feel so sure about our ability to predict which movie will be a super hit or which product will be the next hit in the market?
This is because of the hindsight bias. Our ability to create a perfect narrative story of cause and effect in the hindsight for an observed major event or a crisis, after it has happened, makes us believe that with more information, collated from different sources, we can easily predict it in future. If we analyze every major terrorist attack, in hindsight, the indirect signals leading unto the attack seem to be obviously predictable and signalling the attack. Hence we feel a sense of frustration with "incompetency" of the people in charge.

But if we look at data as it is arriving, much of it is contradictory and full of noise. As the historian Roberta Wohlstetter once remarked "After the event, of course, a signal is always crystal clear; we can see what disaster it was signalling. But before the event it is obscure and pregnant with conflicting meanings". In his book "The Drunkard's Walk" (regarded as one of the 10 top Science books of 2008), physicist Leonard Mlodinow, remarks, "The crystal ball of events is possible only when the event has happened. So we believe we know why a film did well, a candidate won an election, a new product failed or a disease turned worse. But such expertise is empty in the sense as it is of little use in predicting when a film will do well, a new product will fail or a team will lose." Randomness, contradiction, irrelevance in the data makes it difficult to pick up signals. At the same time our biases and prejudices in thinking can act as another blinding force for detecting the signals, when they are distinctly present in the data. We can, at times, ignore what does not fit our thinking paradigms – The confirmation bias!

In 2001, Cisco, one of the most "wired" supply chains announced to the stock market about writing off $2.5 billion of excess raw material. Was it the problem of huge errors in the forecasting software? The answer is No as the number involved was almost half of typical quarterly sales. The real problem was suppliers of Cisco producing in anticipation of future consumption. So when demand dropped with recession, the suppliers kept on producing at the older rate leading to gradual build-up of excess components over a period of 18 months leading to an eventual catastrophe of write-offs. Was the data of increasing inventory at suppliers not visible to planners in Cisco? Or is it a trap of local optima paradigms.

Not many supply chain managers would bother about inventory levels of suppliers; when they are more driven by the need to meet their local need for fast supplies. It is a paradigm by which they look at data around them. If local optimum is the predominant paradigm, one is blinded to signals of potential problems at global level until the mess hits the global picture. In India, almost the entire Auto supply and Consumer Goods supply chain go through a similar "bullwhip" effect at monthly horizon – a heavy month-end skew followed by a dip in first two weeks even though actual end consumer demand variation has no such trend. This way, working has a major havoc on working capital and stock availability at point of sale as space and capital is locked up in slow or non-moving items, while others are stocked out.

Initially, the problem was attributed to lack of easy access to important data like the actual sales at different levels of distribution. Over the last few decades, supply chains are more connected than ever before. Lot of investment has gone into various enterprise software and connectivity to gather

all possible data points. But the problem of monthly "bullwhip" effect (also called the hockey stick effect) remains at the same level over the last many decades without any decisive improvement. This is because the paradigm of management has remained unchanged. The entities in the distribution chains continue to work towards meeting their planned target numbers, which are obviously very static over a year (and ambitious) in nature, which results in push of inventory even when actual consumption trends at the point-of -sale is different. The data point of actual consumption information of the end consumer, even when it is available, will not be of any use to anyone in the organization who is driven by the target-driven behavioural problem or the paradigm of "push".

Erroneous paradigms are the biggest blinders for us to even recognize signals from data. Paradigms in our mind controls the way we look at data and convert it into information. The hypotheses (or paradigms) are formed in our minds based on experiences around us and the way we perceive it.

To solve the problem of erroneous paradigms, one can approach data with blank mind – use good computing power with statistical tools to search for correlations from vast data collected across various sources – the big data approach.

This approach also has a problem; two pieces of data can be accidentally correlated, but assuming them to be cause and effect would be grossly wrong. The data of cancer-related deaths could be highly correlated with the fact that most of them who died also paid their taxes on time. But, our intuition tells us that this is not cause and effect.

Initially it was thought that the disease malaria was associated with damp air of night. Two data sources highly correlated. Later on it was found that malarial cases were found in dry environments and also absent in places where air was very damp. The control on the disease was not greatly enhanced by this correlation knowledge. Our ability to control malaria was enhanced when we understood, after many controlled experiments, to eliminate irrelevant associations (or correlations) to identify the root cause of how malaria spreads – when a female anopheles mosquito bites an infected person, it becomes a carrier of the bacteria and the same bacteria is transferred when the mosquito bites a healthy person. This insight not only opened up new possibilities to control malaria, it also helped understand the correlation of damp air with malaria. Every deeper effect-cause- effect understanding opens up new possibilities for more inventions and better understanding of existing knowledge. A correlation between two variables (random associations) provides limited help, unless one is in the business of classifying information like Google.

Erroneous paradigms are the biggest blinders for us to even recognize signals from data.

Even when we are in an era of petabytes, many times it is impossible to "observe" and count the cause directly. In organizations, it may not be possible to get direct data on entities like the "disinterested distributor" or "not so aggressive sales man" or "committed employee". So many times, we use other "observable" data points (effects) to validate the existence of causes. But this translation can be erroneous. It is not uncommon for consultants to show data to clients to provide a "surprising insight", and in the end clarity of information from clients surprises the consultant. In one case, consultants evaluated products of a fashion company on various parameters of market share and market growth and suggested some products to be trimmed, as they were the "strangers or the slow runners". It so happened that in the next season, the products, which were on the trimming list suddenly, became high sellers. The field people provided the intuition for the erroneous analysis – the data of "sales" was being used to depict demand, but since placements of the product was not good in the first season, the sales data is not reflective of true demand. The resultant effect was a reversal of decisions. Clearly we need help of intuition, as some entities can never be observed directly with data.

But we also discussed that looking at data with intuition can be a signal blinder (because of biases), but at the same time, looking at data without any intuition can lead to a blunder of assuming signals when there is none at all. We seem to be caught between the rock and the hard place. Clearly, we cannot do without intuition but how do we train ourselves to not fall into the trap of our biases?

The typical approach of proving a hypothesis is to look for instances, which support the hypothesis – the process of inductive logic where the mind generalizes, based on specifics instances. I have seen many white swans – so all swans are white. (See that white swan, I had told you before!!) This way of inductive reasoning has been widely accepted by some as a scientific approach. A hypothesis is formed and proved by data of observation.

But there is a problem in this approach as highlighted by philosopher David Hume – how does one distinguish between incorrect inductions? One can always find a way to prove one's point. This opens a Pandora's box of clear demarcation between what can be considered as Science and non-science. Palmistry can also be proven by inductive method and so are Newton's Laws of Motion. The problem of induction is at source of the confirmation bias.
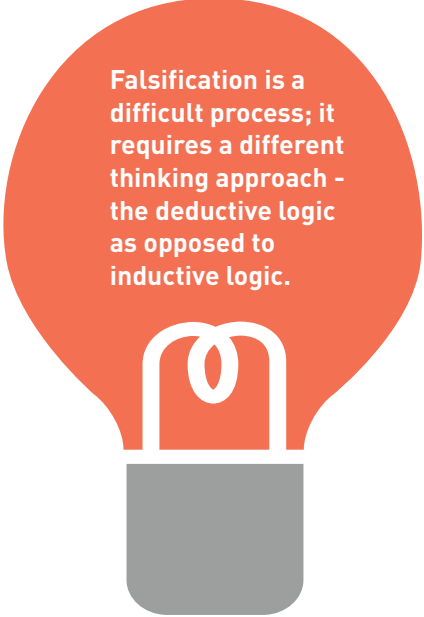
Karl Popper, one of the greatest (and most controversial) philosophers gave a way out of this problem of induction. He argued that Science has progressed by "falsification" or test of failures. So any subject can be called as scientific only if it is "falsifiable" or in other words, one can set up an objective test in which the hypothesis can fail. So by definition, a scientific statement should clearly indicate what it clearly debars. For example, Law of Thermodynamics states that it is impossible to have a perpetual motion machine. It is a scientific statement because it clearly sets up a close case where it can be falsified – if somebody invents a perpetual machine, the theory stands falsified. Only when a hypothesis stands tests of failure by subjecting it to different testing scenarios, we can say that theory is "corroborated" (not true). If it fails a test, it is still a scientific statement. But we have to either drop the hypothesis or modify it for further testing.

Clearly, we cannot do without intuition but how do we train ourselves to not fall into the trap of our biases?

Let us check if palmistry is falsifiable. Let us check the hypothesis; if a lifeline is short, a person will die a premature death. The test of failure would be to look at all cases of premature death and see if most of them had long lifelines. If data shows that actually most of them had normal lifelines, the palmists will not agree to the falsification of the hypothesis. Post facto, they will bring about other hypothesis to support what they are observing and if you further test and find contradictions with new modified hypothesis, it will be further modified post facto and so on. Hence a test of failure cannot be decided upfront – every observed case can be "explained" post facto.

This makes palmistry a non-science, because there is no way to objectively set up a test upfront. Same is the case of a Management Theory, which describes 10 prescriptive ways to be a great company. The theory is built around "research" of great companies. But if one shows many cases to the guru (of the above theory) where companies followed all 10 prescriptive ways but died in the process, the guru will try and "explain" the cases, which do not match his hypothesis. So like palmistry, there is no objective way to agree on a falsification test upfront with the proponent of the theory. "Falsifiability" of a theory is decided if a theory can be called scientific because that lays the foundation for two experts to argue and build on the theory. Inventors love their inventions. They like to see confirmations. Falsification is a difficult process; it requires a different thinking approach – the deductive logic as opposed to inductive logic.

Falsification is a difficult process; it requires a different thinking approach - the deductive logic as opposed to inductive logic.

Deductive logic is a way to challenge the generic statement by observing the specifics. All swans are white (a generic hypothesis) can be falsified by one observation of a black swan. Such deductive logic can be used to set up test of failures under different conditions – does it stand good under different conditions?

Does it contradict with another observation of an existing theory? Does it contradict with itself under a different scenario?

In an interconnected system, a cause is bound to have multiple, different effects. A problem in one organ will lead to effects on other organs due to connectivity between them. So any actual effect, which is contradictory to the predicted effect, will force us to either drop or define the limiting conditions for hypothesis.

This approach gives us a way forward. If there is no way to observe without bias (we will collect data to prove our point) then the best way is to expose one's hypothesis to repeated test of failure by arguing with another person who has exactly an opposite point of view around the same topic – a process of Collective Confrontation of Intuitions (CCI).

For example, if the head of production claims that he has improved his operational performance significantly (or even has data to prove his point); the best person to falsify it is the sales team. Are the sales experiencing an improved delivery? Has the number of expediting requests from customers come down dramatically? Have the last moment emergency shipments come down? If the answer is No to any of them, then hypothesis needs to be questioned. Such an approach requires one to think like a scientist and focus on erroneous assumptions, which falsifies the hypothesis without feeling insulted or falling into an acrimonious debate of blaming each other. This is how quantum physics evolved. When there was no apparatus to directly observe the quantum particles, scientists resorted to thought experiments (imaginary experiments) to test the predicted effects. The arguments between Einstein and Bohr around thought experiments to disprove each other's hypothesis created significant progress for quantum physics. The way we think has a significant impact on the way we visualize the world around us.

Our ability to see reality as it is (or see it objectively) is more constrained by our mental models than data availability. We have an ability to see data the way it suits our opinion. So the way out is an objective confrontation (checking for erroneous assumptions) with contradictory intuitions about reality (process of collective confrontation of intuitions). Success of the hard sciences is because of this method. This requires not Big Data, but a Big Change in the thought process. As Einstein once remarked, "Not everything that counts can be counted. Not everything that can be counted, counts."