SPURIOUS CORRELATIONS
SEEING THE FOREST FROM THE TREES

Mike Even, Chairman of Man Numeric
Shanta Puchtler, President and CEO of Man Numeric

It seems that some of the key tenets of educational theory may need to be radically revised. Data researchers at The Economist have recently discovered a strong meaningful relationship between PISA scores (the OECD's benchmark for academic attainment) and ice-cream consumption per capita. Whilst Peru, Kazakhstan and Brazil all have both relatively meager appetites and comparatively low PISA ratings, the US, Finland and Australia are all elevated on both counts (the latter eats a heart-stopping 14 quarts a year per person). Across all countries the regression has an R^2 of 0.49[1], indicating that almost half the variance of a nation's academic ability can be explained by its proclivity for frozen milk.

There are many other examples of similarly suspect relationships. Golf has been singled out for particular attention, with recent studies suggesting significant relationships between CEO compensation and golfing ability[2], as well as institutional investor returns with geographic proximity to prestigious clubs[3]. At the heart of most of these studies is some type of correlation or regression analysis. These form one of the core statistical tools used by firms like Man Numeric for analysis of stock market returns and relationships. With the volume of data being produced today (around 8.7 quadrillion bytes will have been created in the five minutes it takes you to read this article), there will certainly be examples of correlations which are purely coincidental. It is therefore a critical part of the Man Numeric research process to carefully determine when an observed past pattern has the potential to persist in the future, and when the concept deserves to be thrown into the circular file.

For investors, such relationships can sing like Virgil's Sirens, presenting an opportunity which seems too good to be true but which, after the shipwreck, is revealed to be just that. Take the now famous 'Super Bowl indicator', for example. This states that if the Super Bowl is won by a team that played in the original NFL conference pre-1970, then the Dow Jones should be up on the year, and vice versa for an AFL franchise. Imagine you were an investor seeking meaningful correlations on New Year's Eve 1997. Looking back over 31 years of data since the first Super Bowl in 1967, you would observe that this 'strategy' had compounded at 14%, with only three down years and a maximum annual drawdown of just 4.8%. This particular song, however, though alluring would have proved just as risky as it did for Ulysses. Over the next three years you would have lost 41% of your portfolio. Even if you survived that, your fortunes would remain bleak, with a series of losses (including a massive 34% fall in 2008) meaning that, by the end of 2015, your CAGR since investment would stand at -3.6%.

The human mind can make the problem more acute by entrenching the tendency to see these chance patterns. The most dangerous behavioural biases in this regard include the following:

1. **Overconfidence bias** – the analyst overestimates her predictive abilities and thus does not treat regression results with enough scepticism
2. **Clustering illusion** – one result can be repeated across a certain time period which the analyst fails to place within its wider context. A football team can have a strong winning streak over the short term, for example, whilst having a long term record that is much less glittering
3. **Survivorship bias** – the analysis ignores failures that would otherwise influence the perception of the strength of the relationship being modelled
4. **Confirmation bias** – an erroneous way of looking at new information where the analyst gives outsized weight to data which confirms a pre-determined thesis
5. **Recency bias** – similar to confirmation bias, but in this instance the analyst puts undue emphasis on the most recently presented information

There is no one antidote to these problems, but we have found a number of practical steps which can potentially help minimise the damage. Firstly, you should always maintain a healthy scepticism towards your own abilities. A thorough and unflinching statistical examination of one's track record will inevitably turn up a litany of failures which should be kept at the forefront of the mind. This is not just an exercise in masochism, but instead represents a way of keeping overconfidence in check. Particular suspicion should be reserved for quantitative backtesting as even tiny look-ahead biases or misinterpretations of timelines can lead to incredible paper returns which in real life would have been impossible to achieve.

For example, the Super Bowl indicator that we have already discussed categorises teams according to their pre-1970 conference, NFL or AFL. Any franchise founded since the two merged is placed into the NFL bracket if it originally played in the NFC, and AFL if AFC. The Pittsburgh Steelers, for instance are classified as NFL, despite being a current member of the AFC because they were founded in 1933 and were a member of the original NFL prior to the 1970 merger. Given that they took the trophy home four times in the backtesting period (including in 1975 when the market rose 38%), this one classification decision (with which some would no doubt argue) would have had a major effect on the perceived attractiveness of the strategy.

Given the unique human ability to self-deceive, we feel third party involvement is a crucial component of this process of self-criticism – both to give a realistic assessment of past records, and to critique present findings. Indeed, whether an unconnected analyst can review the same research to come separately to the same conclusion is a good litmus test of the model's strength in our view.

1. See http://www.economist.com/blogs/graphicdetail/2016/04/daily-chart
2. See http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1374239
3. See http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2638577

This is becoming an even more useful practice as both the volume of data available, and the levels of computing power available to analyse it, continue to advance apace. This progression continues to make datamining easier. Machine learning can potentially exacerbate this by reaching elegant conclusions with increasingly complex algorithmic structuring, making the logic behind the solution more and more difficult to discern.

Secondly, we believe quantitative investors should avoid the temptation of rejecting qualitative thought entirely. We believe in the value of spending considerable time thinking philosophically about whether a model's output has a credible rationale behind it. At Man Numeric we saw a good example of this when it came to researching Emerging Markets strategies in 2008. At the time, Quality of Earnings factors did not seem to be a major driver of returns in backtested results. We were sceptical of this quantitative conclusion, however, as the research time period covered the late 1990's through to the end of 2008, a timeframe marked by a huge surge in growth related to China's industrial development specifically, and maturation of the emerging market economies more generally. Because, from a qualitative perspective, we strongly believed that China's growth rate was unsustainable and that, when it did stall investors would focus more on how companies were managing their balance sheets, we decided to include the Quality model despite the statistical information suggesting this was the wrong path.

We believe an important consideration, therefore, is whether the model does more than simply describe stock market returns, which can be extremely whimsical, and instead says something about a company or market's fundamental state – its earnings, its margins, its asset efficiency or other underlying performance metrics. For instance, a model that both explains future potential returns but also seeks to predict which stocks may have positive earnings surprises is, in our view, likely to be more reliable than a model that simply fits the return pattern, but has no real linkage to the fundamental prospects of the underlying companies in which it invests.

Finally, it is important to think about the integrity of the data itself. Very little data in the world is 'as was'; in other words, it is very rarely observed and recorded in real time. Most is back-filled, mapped after the fact or reconstructed from other information using a variety of manual or automated business logic, rules, and sometimes even whim. Sometimes, if you find a correlation that looks interesting, it may be driven by contaminated data; a deep understanding of how information was collected, rather than a simple Bloomberg drag-and-drop, is in our view an essential component of finding potentially real, as opposed to ephemeral or suspect, relationships.

These principles we have described will obviously not remove all the errors of human thought. We believe, however, that they do provide a good platform from which to build the framework of a quantitative investment process which has the potential to effectively deploy correlation analysis (and all of its statistical brethren) as an analytical tool, whilst being all the while cognisant of the fact that it can prove duplicitous.

Michael Even, CFA
Chairman of Man Numeric

Mike Even is the Chairman of Man Numeric ('Numeric'). He joined Numeric in 2006 and served as President and CEO until January 2016. Prior to joining Numeric, Mike worked at Citigroup for nine years building a global asset management organization. After Citi merged with Travelers, Mike became Global CIO of the Private Bank and eventually, Global CIO for Citigroup Asset Management, overseeing investment teams running nearly USD500bn in client assets. Prior to joining Citigroup Mike spent 13 years at Independence Investment Associates in Boston. Mike holds a B.A. in Economics and a B.S. in Operations Research from Cornell University and an M.S. in Business from the Massachusetts Institute of Technology. Mike is a CFA charterholder and currently serves on the investment committees of the Massachusetts Pension Reserves Management Board (PRIM) and the Trustees of the Reservation.

Shanta Puchtler, CFA
President and CEO of Man Numeric

Shanta is President and CEO of Man Numeric ('Numeric'). He was named CEO in January 2017 and President in January 2016. He is also a member of the Man Group Executive Committee. Formerly, he was Chief Investment Officer and Head of Research at Numeric, directing research efforts focused on new alpha sources, product design, and risk modeling. He joined Numeric in 1999 as a research analyst. Prior to joining Numeric, Shanta was an electronic commerce technology analyst at Forrester Research, a Cambridge-based market research firm. He also co-founded an electronic commerce company which focused on the analysis of on-line buying behavior. Prior to that, Shanta worked as a management consultant and taught at private schools in the US and Asia. Shanta received a B.A. in Computer Science from Dartmouth College, graduating Summa Cum Laude, and is a CFA charterholder.