

## MEANINGFUL APPLICATION OF QUANTITATIVE METHODS

While we make heavy use of statistical analysis and computer models, we want to emphasize what we don't do. We do not indulge in data mining. Data driven methods can uncover useful relationships but if used carelessly, they can just as easily “discover” meaningless, spurious patterns. A million dollar jackpot winner, when asked by the press how he chose the winning number, said he had a “scientific reason” for choosing the number 48. The number 7 was his lucky number, he explained, and two of them would be luckier still – and as  $7 \times 7$  is 48, he chose it and won! The story is almost certainly false, but it is highly illustrative of bad assumptions, bad calculations, and a cognitive bias that results in ex-post justification of a fluke win.

A contrived investment rule made popular by the investment website Motley Fools was: “select the 10 highest yielding Dow Stocks, pick the 5 lowest priced stocks of these, exclude the bottom one, and give double weight to the second one.” Further data mining attempts by the Motley Fools to “improve the performance” included dividing the dividend by the square root of the price.

In an article titled *Stupid Data Miner Tricks*, Berkeley professor David J. Leinweber showed, tongue in cheek, that butter production in Bangladesh was the best predictor for the S&P 500 Index from 1983-'93. Over that long period, his regression model had a R-squared statistics of 0.75; which the author gleefully reported he could increase to ever higher numbers by adding more and more meaningless variables.

Harry S. Dent used the number of 46 ½ year olds in the US population as an input into a model that would predict the level of Dow Jones Index. He came out with a book in 1998 titled *The Roaring 2000's: Building the Wealth and Life Style you Desire in the Greatest Boom in History* and followed this up in 1999 with a sequel *The Roaring 2000's Investor*; which predicted that the stock market would continue to boom throughout the decade (2000's) and crash in 2009 – all this based on the number of people in the age bracket of 45 to 50.

In a light hearted comment, John von Neumann said that given enough degrees of freedom, he could make an elephant pass through the eye of a needle. Taking a cue from this great scientist, we steer clear of all approaches which come even remotely close to data mining.



*The Baya weaver uses only the simplest of materials – grass, twigs and clay – yet weaves sturdy nests with a complex structure. The best trade ideas are rooted in simplicity.*