

# women's month

## The role of data science in money management

Shriya Roy, Quantitative Analyst

With the fourth industrial revolution underway, organisations that respond proactively and appropriately to technological changes will gain a competitive advantage. In the investment industry, data science, which relies on structured and unstructured data sources to inform decision making, promises to be one of the most exciting and value-adding areas of analysis.

In 2006, Clive Humpy coined the term “data is the new oil”. He said that like oil, data is valuable, but unusable if unrefined. With the abundance of available data, it would be a mistake to believe that the more data collected, the more valuable the outcome. However, the validity and integrity of the data determine the value of the results generated.

In a money management context, data science has provided us with various statistical techniques to extract valuable insights out of the data that we use. The first step is to clean and transform the data because substantial datasets frequently include missing values and/or anomalies. Data cleaning and transformation can take up to 80% of the project's duration.

Although time-consuming, cleaning the data ensures that any analysis based on the available data will generate reliable results that allow us to make conclusive decisions – as the saying goes, “garbage in, garbage out” is a fundamental truth in the field of data science.

There are other considerations that need to be taken into account when dealing with data. Firstly, the sample size needs to be sufficiently large to ensure that any conclusions drawn from the available information are free from inaccuracies or bias. Second, the source of the data needs to be relevant and reliable. For example, a time series dataset of share prices is regarded as a traditional data source. Another source of data that data scientists are increasingly tapping is “alternative data” because non-traditional sources of data enable them to enhance their views. For example, through natural language processing, text data from Twitter can generate valuable information on the state of, and trends in, sentiment. Other sources include sell-side data, web data, or satellite images. All of these datasets allow us to develop a more holistic view of what is happening in the world and helps us quantify it and use it in our investment decisions.

Data science methods have given us a better perspective on different asset classes and risk management tools. For example, a correlation matrix based on a set of assets with a historical time series provides insight into the relationships between each asset-pair. The more positive the correlation, the less desirable the pair because it implies that they behave similarly, therefore lacking the diversification we seek in the portfolio. The more negative the correlation between the asset-pairs, the more beneficial the combination as the behaviour of these assets differ, thereby supporting diversification. A simple indexed performance chart of a set of assets based on each of their historical time series data enables us to see how these assets performed over a specific time period in relation to one another.

We employ a technique known as scenario analysis to monitor the sensitivity of our portfolios. This analysis depicts how our portfolios behaved in periods of crisis and the information helps us adjust our portfolio positions proactively. What makes our monitoring and analysis valuable is that it is real-time and uses our current asset weights and compositions. Data science allows us to process large amounts of data timeously, enabling us to make quicker decisions and remain relevant.



Shriya Roy,  
Quantitative Analyst

An exciting part of incorporating data science in our processes is the range of visualisations they allow. By reviewing the data outputs visually, we can view the data outputs from various angles and quickly grasp underlying patterns or trends. As an asset manager, it allows us to graphically show our current strategic and tactical asset allocation positioning across all our funds, as well as compare them to peers and their respective benchmarks. Most of our visualisations are interactive, which allows the user to adjust the view to one that meets their needs.

We consolidate our data into four primary factor indicators, namely Value, Economics, Financial Conditions and Sentiment. These indicators are further broken down into about 20 sub-factors, which we quantify and analyse too. These factors are allocated to each of our asset classes: Equities,

Bonds, Credit, Income Money Market and Cash, Real Assets and Foreign Exchange, as well as their sub-categories. For each indicator, we calculate a score that determines the weighting of each factor within each asset class.

We have recently refined our scoring results to include “how” positive or “how” negative the score is, providing more depth to our views. Our factor modelling and scoring process takes into account around 500 time series’ and more than one million data points. The process is heavily data-driven, statistical and systematic. Systematic investing was quite challenging in the past, due to the deficit of valuable data and inefficient systems that could not support or process large datasets. But data science has evolved significantly since then and we have taken advantage of these advances in our own systematic investment process.

Constructing our portfolio asset weightings is also a systematic process. We determine our assets of choice through exploratory analysis. Then, based on the portfolio constraints and investment objectives, we use statistical optimisation to generate a spectrum of weights for our assets across a range of targeted risk buckets. Interestingly, as a

result of the sound logic of the algorithms used in optimisation, certain assets are rejected due to their high correlation with other assets. Some optimisation methods take considerable time to generate optimal results. However, the more repetitions the optimisation runs, the more accurate the results. There has always been a trade-off between the capital outlay for computational systems and their accuracy. At Prescient, we never compromise on accuracy, and for that reason we direct our investments into enhanced computing power and substantial data sources.

Data science has disrupted the investment industry in recent years, particularly traditional fundamental investing, and, as it becomes more sophisticated, its array of useful applications in the investment industry will expand and it will be essential to understand data-based analysis.

The benefits of systematic investing are clear. Relying on data science to inform systematic investing eliminates human bias, creates more efficient processes, enhances consistency, scalability and automation of investment decision making, which, in turn, enables swift, proactive and well-informed decision-making. ■