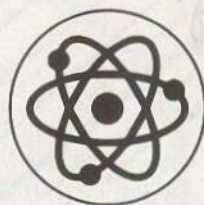## P K Viswanathan
Professor (Analytics)

He is the Programme Director of Business Analytics and Business Intelligence, at Great Lakes Institute of Management, Chennai.

SECTION

4

THE SCIENCES & MATHEMATICS

Proof of the Pudding...

... is in the data. Knowing who will win an election before it happens, identifying customers who will default on a loan before they do – sounds like magic? No, that's data science.

Quick Scan

■ Experience required in Python, R, SAS, Hadoop, and similar applications and/or languages

■ Companies with the most analytics openings in 2017 include Amazon, HCL, IBM, CitibankCiti, and Goldman Sachs, among others

■ Job titles: data scientist, statistician, data analyst, analytics manager, and analytics consultant

Data science is multi-disciplinary and involves transforming hypotheses and data into actionable insights and predictions. Statistician and computer scientist William Cleveland considers data science to be an inter-disciplinary field larger than statistics itself. It involves acquisition of data, managing data, selection of appropriate models, writing programs, and validating results.

Predicting which party will win the election, whether a new product will succeed in the market place, which customer will default on a loan taken, which advertisement will be a success, what characteristics sharply differentiate market segments in terms of consumer behaviour, and what the drivers of a landing page in a website are, are just some examples of data science, among many.

Big data impacts data science in an emphatic manner. Amazon's product recommendation systems, Google's advertisement valuation systems, and Twitter's trending

topics are well-known examples of the big data flavour in action.

Data science permeates multiple disciplines and sectors of industry. Analytic firms are not the only ones who apply data science to solve business problems. Examples of its applications include marketing analytics for FMCGs, financial analytics for banks, and operations analytic for engineering and manufacturing units. This field involves feedback and interaction between the data scientist and all other stakeholders.

The life cycle of a data science project is succinctly captured by the following visual. Please note that the boundaries between stages are fluid and one has to loop back and forth before moving forward in the process.
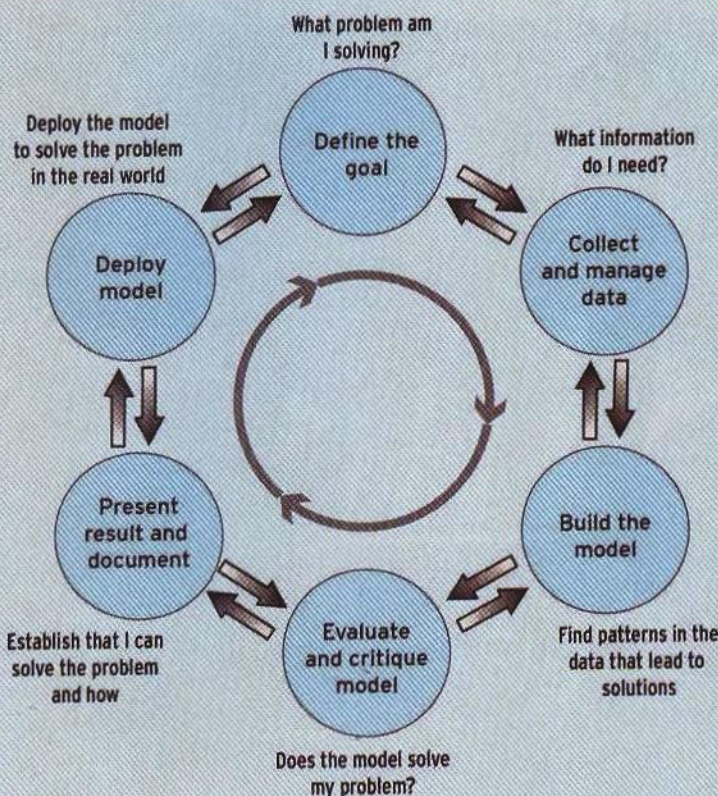
> ...Google's advertisement valuation systems, and Twitter's trending topics are well-known examples of the big data flavour in action.

Job opportunities

Three decades ago, computer science attracted many students because of lucrative jobs and great opportunities to occupy top positions. The same phenomenon is happening today for data science.

The *Analytics and Data Science India Jobs Study 2017*, by Edvancer Eduventures and Analytics India Magazine (AIM), reveals the following interesting findings:

- The number of analytics jobs increased by 52% in the timeframe of of April 2015 to April 2016.
- While it is difficult to give precise number, it is estimated that at present 50,000 positions related to analytics are available in India, cutting across all functions, type of industry, and skill-sets.
- India contributes just 12% of job openings in this field currently. But the number of jobs in India is slated to increase at a much faster rate compared to the rest of the world. More analytics projects are likely to be outsourced to India due to paucity of skilled professionals across the world.
- Ten corporates with the most number of analytics openings in 2017 were: Amazon, Citibank, HCL, Goldman Sachs, IBM,

Source: *Practical Data Science with R by Nina Zumel and John Mount, Manning Publications Co., U.S.*

JPMorgan Chase, Accenture, KPMG, E&Y, and Capgemini.

- The average salary for an analytics professional in India for the year 2017, was ₹11.7 lakh (a 22% increase from 2016).

Roles and skills required

The most attractive job titles that are associated with data science are data scientist, statistician, data analyst, analytics manager, and analytics consultant. All these roles require sound knowledge of statistical modelling, besides technical and software skills. The technical side requires data architecture and Hadoop applications for big data. Software skills include expertise in R and Python programming, and/or SAS programming. Defining the problem, the art of modelling the problem, finding the best solution, implementing the solution, exceptional communication, and report writing skills are expected of data science professionals.

The cornerstone of data science curricula is machine learning that could be dichotomised into supervised and unsupervised learning. The subjects that are covered in-depth in this programme are data extraction, exploratory data analysis, big data analytics, statistical modelling, machine learning algorithms, R, Python, natural language processing (NLP), and deep learning using artificial intelligence, besides many others. The defining technique of data science in the coming years will be the use of neural networks for automation of machine learning. In fact, the neural network black box model will become explainable.

A bachelor's degree in engineering or computer science, and a master's degree in mathematics, statistics, or econometrics with a consistent academic record will be ideal for entry into this course. Two years' work experience in the domain of analytics will be an added advantage.

Strong numerical ability and aptitude for quantitative analysis of numbers is a must. Problem-solving skills with an algorithmic approach, the ability to communicate numbers with business sense, and fine report writing skills are expected from the participants pursuing data science course.

Courses
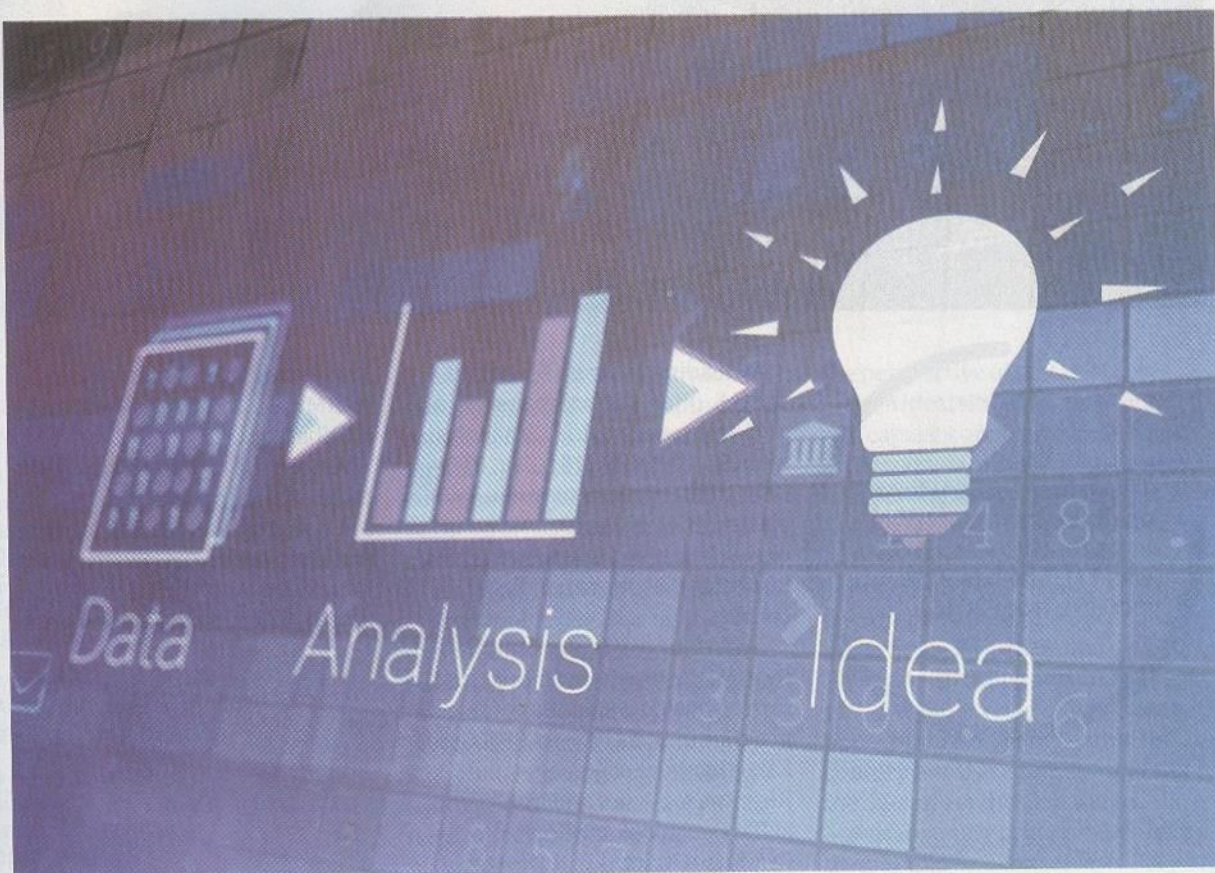
> The average salary for an analytics professional in India for the year 2017, was ₹11.7 lakh (a 22% increase from 2016).

All leading business schools in the U.S. offer M.S. programmes in data science or business analytics. In India, IIMs, ISB, IITs, IISc, IIITs offer similar programmesvia classroom or online delivery. For executives, there are good training institutes that offer analytics courses.

According to the recent survey by AIM, the top three among the top ten business analytics programmes are PGP-BABI at Great Lakes Institute of Management, Business Analytics and Intelligence at IIM Bangalore, and Post-Graduate Diploma in Business Analytics (PGDBA) offered jointly by IIM Calcutta, ISI, Kolkata, and IIT Kharagpur.

Among the training institutes in India offering analytics/data science programmes, the top three of the top ten are Jigsaw Academy, Bengaluru, AnalytixLabs, Delhi, and UpX Academy, Hyderabad. ∎