Data analytics – a promising career frontier

PK VISWANATHAN

Professor (Analytics), Director (AIMLA – Artificial Intelligence, Machine Learning and Analytics), Great Lakes Institute of Management, Chennai

For business houses, the discovery of patterns and trends from data paves the way for improving profitability

Data analytics is a set of tools that is used to discover patterns, trends and associations from a given data set. Like the *Annapakshi* (the mythical swan-like bird) that is said to be able to separate pure milk from a mixture of milk and water, and drink only the milk, data analytics separates insights (milk) from the other portion (water) of data and draws inferences for decisions.

While business analytics refers to the use of data and quantitative analysis to find solutions for business problems, data analytics has a much broader focus. For example, data analytics can help a not-for-profit making enterprise, such as a home for exceptional children, in identifying the driving causes of autism.

One can say business analytics is data analytics applied to business situations to arrive at the best possible solution to a given problem. Here, the ambit of data includes numerical and text data. Text mining involves unstructured natural language processing (NLP) to extract key insights. NLP is considered extremely important in performing sentiment analysis-based qualitative data that includes Web and social

media. For example, what people think of GST in various parts of India can be sifted from sentiment analysis which, in turn, helps policy makers in the government to fine-tune and modify GST for certain items.

Why has data analytics become important today?

Three significant trends have triggered the current meteoric growth in the use of data analytics in decision-making.

Trend 1: The revolution of internet of things (IoT), web and social media, and mobile phones produce large amounts of data from which insights can be extracted. For business houses, the discovery of patterns, trends and associations from such data will pave the way for improving profitability, understanding customer expectations, and pricing products appropriately, so that the company can gain a competitive advantage in the marketplace.

Trend 2: An enormous increase in the computing power available today facilitates efficient processing, and analysis of massive amounts of data. Therefore, sophisticated and faster algorithms are available today to solve complex problems.

Trend 3: Large data storage capability, parallel computing and cloud computing, coupled with better computer hardware, have enabled organisations to solve large-scale problems faster than ever before.

What are the pillars of data analytics?

There are four pillars of data analytics. They are:

1) Descriptive analytics
2) Diagnostic analytics
3) Predictive analytics
4) Prescriptive analytics

Descriptive analytics focuses on what the numbers say. Is there any pattern emanating from the data? Do the visuals display any patterns? Tools like histogram and box plot are used for identifying patterns.

Diagnostic analytics describes why the numbers behave in a particular way. What are the possible hypotheses? Can they be

proved or disproved? Statistical hypothesis are formulated and tested for inferences.

Predictive analytics looks at predicting the outcome. Artificial intelligence, machine learning and analytics (AIMLA) are used to predict certain outcomes, such as forecasting which party will win the election, whether a new product will succeed in the market place, and whether a customer will default on a loan taken.

Prescriptive analytics does the job of prescribing a particular course of action or a set of recommendations for decision making. In other words, it provides answers to the question: "What should we do?"

What is big data?

The conventional definition of big data is: "A set of data that cannot be managed, processed, or analysed with traditional software/algorithms within a reasonable amount of time." Big data revolves around volume, velocity, variety, value and veracity. It is important to realise that big data is relative because computing power keeps changing over time.

In a practical way, big data can be defined thus: "If the data set can challenge the computing power, both in terms of hardware and software, only then is it big data." One must remember analytics exists with and without big data. There are a number of situations where data size may not be very large and yet one needs analytics to make decisions.

Walmart handles over one million purchase transactions per hour, and Facebook processes more than 250 million picture uploads per day. These are examples of big data. Amazon's Product Recommendation Systems, Google's Advertisement Valuation Systems and Twitter's Trending Topics are other examples of big data impacting the consumer and social media space.

What job opportunities are available?

Job opportunities in analytics have been constantly evolving in India over the years. There is a steady rise in the overall demand for analytics professionals, with new jobs being advertised each day.

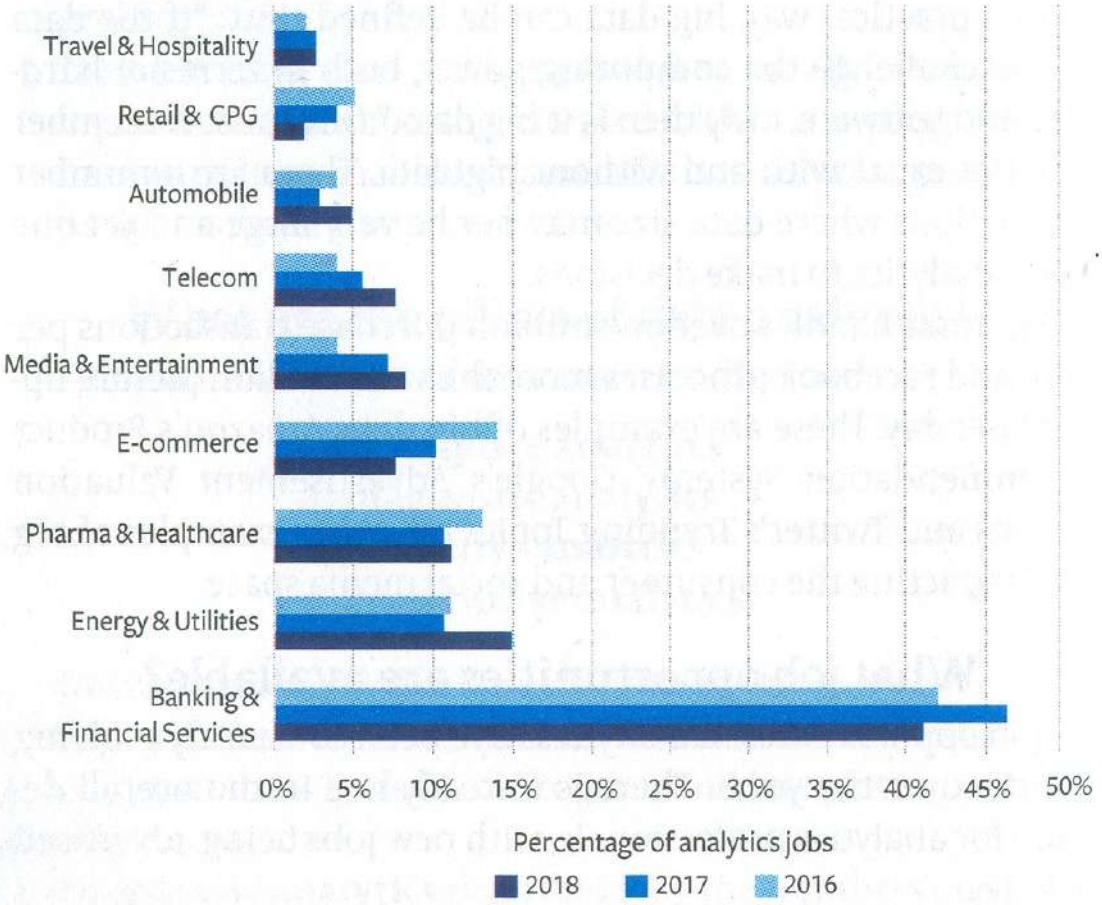
A research study conducted in 2019 reveals the following interesting findings:

- The median salary for an analytics professional in India for the year 2018-19 remained at ₹12.6 lakh per annum cutting across experience levels and skill sets.
- Data analytics professionals are currently benefiting from the big data wave and are earning 26% higher salary than a typical software engineer in India.
- Analytics salaries continue to exceed other software engineering roles in India with earnings that outperform Java counterparts by almost 50%.
- Big data and data analytics professionals who work on unstructured data continue to earn more than normal analytics professionals.

Source: ANALYTICS INDIA SALARY STUDY 2019 - Analytics India Magazine & AnalytixLabs

Forty-one per cent of all jobs in analytics were from the banking and financial services sector and 11 per cent were from the energy and utilities sector for the year 2018 *(See figure below)*

Percentage of analytics jobs across industries



Source: Analytics and Data Science India Jobs Study 2018-EDVANCER and Analytics India Magazine (AIM)

What does a typical data analytics course entail?

Ideally the curriculum should include statistical methods for decision making (SMDM), advanced statistical modelling, data visualisation, predictive analytics, data mining, AIMLA, and NLP. Some useful software skills-sets would be: R (mainly used for statistical analysis) Python (that provides a general approach to data science), SAS, Tableau, and Power BI (a business analytics service by Microsoft).

Eligibility criteria

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

Aptitude needed

Strong numerical ability, aptitude for quantitative analysis, problem-solving skills, the ability to communicate numbers with a good business sense and fine report writing skills are the hallmarks of a data analytics professional.

Data analytics programmes in India

In India, all the IIMs, ISB, IITs, IISC, Great Lakes Institute of Management, and IIITs offer programmes in business analytics and data analytics via classroom as well as online and hybrid blended learning modes. There are good training institutes that offer analytics courses for executives. Some of these are Jigsaw Academy (Bengaluru), AnalytixLabs (Delhi), and UpX Academy (Hyderabad).