

## **50 Years of Data Science**

Data Science is a combination of collecting, managing, analyzing, visualizing a small or huge amount of Data, and Data Scientist is someone who is responsible for professionally executing the process of collecting, analyzing, and interpreting any kind of structured or unstructured data. Earlier statistics was used to manage data but it could only be used to a limited extent as it needs mathematical formulations to interpret quantitative data, unlike Data Science which uses scientific methods. Historically statistics focuses on very small quantities of data using some regression techniques and would interpret if the data is consistent. Every new invention gives us hope to enhance yourself that is the reason why data science was developed to infuse statistics and database administrators to work on a single platform. So, Data Science is just nothing but an advanced version of statistics including information technology to solve huge data sets. The Job's Meme shows how a hiring bump has occurred in the leading companies like Google and Amazon in hiring people who are skilled in the database and as well as statistics. The Master's degree in Data Science will definitely add an advantage to any organization as an individual will have knowledge of both statistics as well as databases. This is an added advantage of having special information technology skills. As mentioned in one of the examples how in 2010 it was told the life in villages in India has not changed in thousands of years, but when visited in the year 2015 over 8 billion people are using smartphones and are connected to the network. This shows the advancement of technology and also how we can use the data provided by daily activities and preferences of people effectively. For 50 years the usage of traditional statistics has improved a lot where now we are in a stage to use advanced techniques to solve any problem effectively. Here I would like to mention an example of how the data was stored in floppy disks, Digital Versatile Disks(DVD's) decades ago whereas now we are using Cloud technology to store any kind of data and access it from anywhere in the world. In the same way, Statistics would draw a conclusion based on uncertainty whereas Data Analysis would provide us with the specific results. Ultimately both Statistics and DataScience aims to extract knowledge from data however, they differ in modeling processes and the size of data.