S.SHANMUGA PRIYA
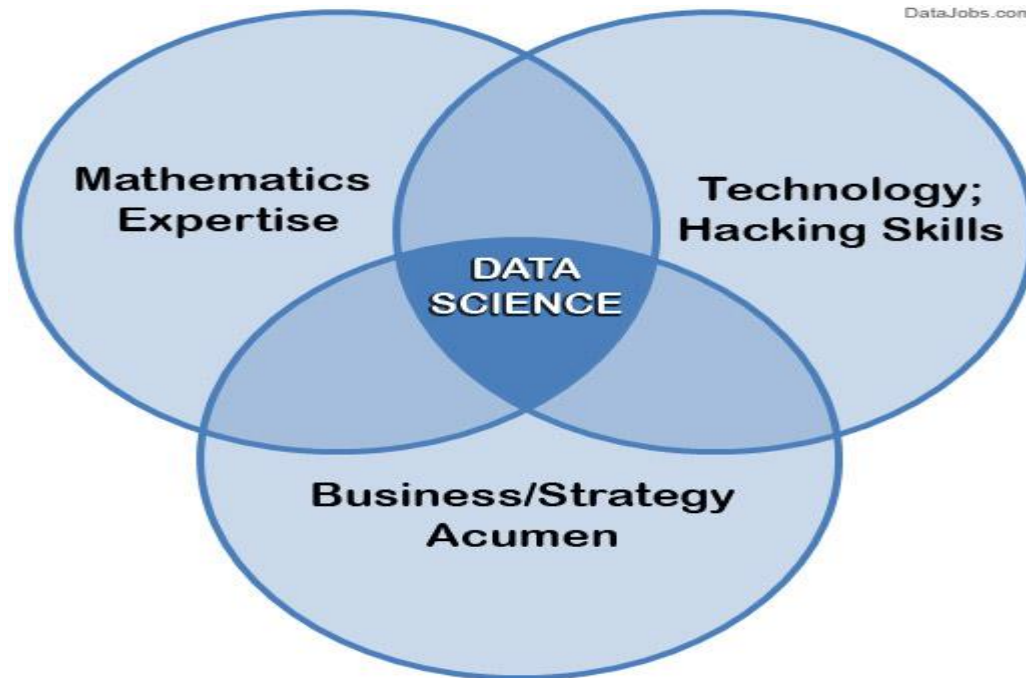# III-MCA
N8MCAL038

DATA SCIENCE

* Data Science is a multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data.

* Data science is the same concept as data mining and big data.

* It plays a vital role in solving analytically complex problems
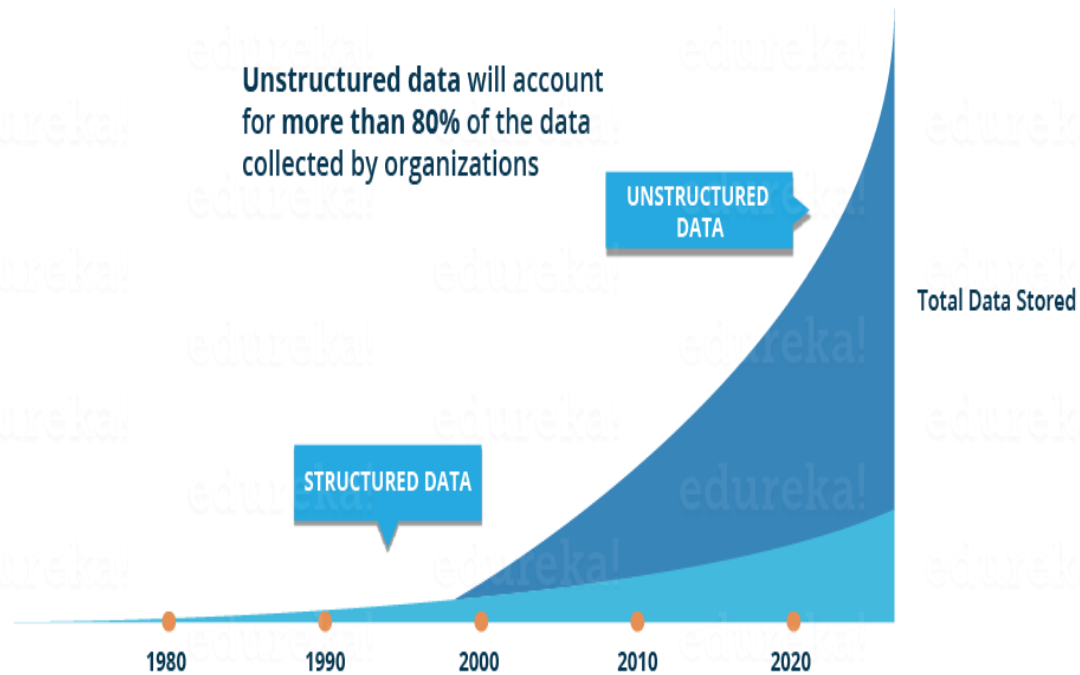
THE REQUESITE SKILL SET OF DATA SCIENCE

Data science is a blend of skills in three major areas.

WHY WE NEED DATA SCIENCE

* Traditionally, the data that we had was mostly structured and small in size, which could be analyzed by using the simple BI tools. Unlike data in the traditional systems which was mostly structured, today most of the data is unstructured or semi-structured.

Unstructured data will account for **more than 80%** of the data collected by organizations

UNSTRUCTURED DATA

Total Data Stored

STRUCTURED DATA

1980   1990   2000   2010   2020

* In the previous diagram, the data is generated from different sources like financial logs, text files, multimedia forms, sensors, and instruments. Simple BI tools are not capable of processing this huge volume and variety of data. This is why we need more complex and advanced analytical tools and algorithms for processing, analyzing and drawing meaningful insights out of it.

DOMAINS WHERE DATA SCIENCE IS CREATING IT'S IMPRESSION