

Modern Data Ecosystem and the Role of Data Analytics:

Modern Data Ecosystem:

- The constant increase in data processing speeds and bandwidth, the nonstop invention of new tools for creating, sharing and consuming data, and the steady addition of new data creators and consumers around the world, ensure that data growth continues unabated. Data begets more data in a constant virtuous cycle.
- Modern Data Ecosystem includes **interconnected, independent, continually evolving** entities.

Key Players in Data Ecosystem:

- **Data Engineer-** develops and maintains data architecture and makes data available for business operations and analysis. They extract, integrate and organize data from different sources, cleans, transforms and prepares data and designs, stores and manages data in data repositories.
- **Data Analyst-** translates data and numbers into plain language so that organizations can make decisions. They inspect and clean data for deriving insights, identify correlations, find patterns and apply statistical methods to analyze and mine data and visualize data to interpret and present the findings of data analysis.
- **Data Scientist-** analyze data for actionable insights and create predictive models using ML and DL that train on past data to create predictive models.

Defining Data Analysis:

- Data Analysis is the process of gathering, cleaning, analyzing and mining data, interpreting results and reporting the findings.
- Data Analysis helps businesses understand their past performance and informs their decision-making. Using data analysis, businesses can validate a course of action before committing to it.
- There are different types of data analysis, namely,
 - (i) **Descriptive Analysis-** helps in answering questions about what happened over a given period of time by summarizing past data.
 - (ii) **Diagnostic Analysis-** helps in answering why did it happen. It takes insights from descriptive analysis to dig deeper to find the cause of outcome.
 - (iii) **Predictive Analysis-** helps in answering what might happen next. Here historical data and trends are used to predict future outcomes.
 - (iv) **Prescriptive Analysis-** helps in answering what should be done about it. It analyzes past decisions and events to estimate the likelihood of different outcomes on the basis of which a course of action is decided.
- The data analysis process is as follows-
 - (i) **Understanding the problem and desired result-** defines where we are and where we wanna be.
 - (ii) **Setting a clear metric-** deciding what will be measured and how.
 - (iii) **Gathering Data-** identifying the data we require, its sources and best tools for this job.
 - (iv) **Cleaning Data-** Fixing quality issues in data that could affect the analysis and standardizing the data coming from multiple sources.
 - (v) **Analyzing and Mining Data-** extracting, analyzing and manipulating data.

- (vi) **Interpreting Results-** interpreting results, evaluating defendability of analysis under circumstances under which analysis may not hold true.
- (vii) **Presenting Findings-** communicating and presenting our findings in clear, impactful and convincing ways.