

## Modern Data Ecosystem :-

- Modern data ecosystem consists of Interconnected, Independent and Continually evolving entities.
- Data formats → Texts, videos, images, social media posts, data from IOT devices, real-time data & data provided by Agencies.
- When working with data, 1<sup>st</sup> step is :-  
Pull a copy of the data from original source; into a <sup>new</sup> data repository. → Acquiring the needed data in correct format, through interfaces.  
**Challenges at this stage :-** Reliability, Security, Integrity of data.
- As 2<sup>nd</sup> stage :-  
Raw data is cleaned up, organised and optimised for access by end users.  
Data also needs to conform with compliance and standards enforced in the organisation.
  - Ex:- conforming to guidelines that regulate the storage & use of personal data, such as biometrics, health.**Challenges at this stage :-** Data management, working with data repo that provide high availability, flexibility, & security, & accessibility.

⇒ In the 3<sup>rd</sup> stage:-

The Business stakeholders, applications, programmers, data analysts & scientists pull these data for use.

Challenges at this stage:- Interfaces, API's, applications, used appropriately, for specific cases.

## •) Emerging Tech. shaping Modern Data Ecosystem:

- ⇒ Cloud Technologies,
- ⇒ Machine learning,
- ⇒ Big Data

## •) Key Players in Data Ecosystem :-

### 1) Data Engineers :-

⇒ People who develop and maintain data architectures and make data available for analysis.

⇒ Works in Data Ecosystem to Extract, Integrate, Organise data coming from various sources.

⇒ Clean, store, manage, transform, prepare data, & design & manage data repositories.

⇒ Presents the data for business intelligence & analysts.

⇒ Skills :- Programming, System & technology architecture, Relational & non-relational data storage.

## ii) Data Analyst :-

- ⇒ Translates data & numbers into plain language, that enables decision making.
- ⇒ Inspect & clean data for deriving insights.
- ⇒ Find correlations & patterns, and apply statistical methods to analyse & mine data.
- ⇒ Visualize data to interpret & present the results of data analysis.
- ⇒ Skills :-
  - ⇒ Knowledge of spreadsheets, writing queries, using statistical tools to create charts & dashboards
  - ⇒ Analytical & story-telling skills.

## iii) Data Scientists :-

- ⇒ Analyze data for actionable insights, & build machine learning & deep learning models, trained on past data.
- ⇒ Produce future predictive results, and answers to questions such as:-  
Ex:-
  - ⇒ "Is a financial transaction unusual?"
  - ⇒ "What % age of customers will terminate product usage next month?"

Skills:- Mathematics, statistics, domain knowledge, understanding of databases & building of data models.

## iv) Business Analyst / BI Analyst :-

- Uses the results given by Data Analysts & Scientists, to look at possible implications (conclusions) for business, and actions that needs to be taken.
- Main focus:- Market forces & external influence that shape their business.
- Explore data to extract insights that improve business.

