

FDA UNIT -1

1. What is Data Analytics?

- Data analytics is the process of **collecting, organizing, cleaning, and analyzing raw data** to find useful information, draw conclusions, and help in decision-making.
- It helps organizations to **spot patterns, trends, and relationships** in data, which can be used to improve performance and solve problems.
- Data analytics is used in almost every field today, like business, healthcare, finance, education, and more.

2. Why is Data Analytics Important?

- **Better Decision Making:** Gives facts and insights, so decisions are based on data, not just guesses.
- **Problem Solving:** Helps find the root cause of issues by analyzing data deeply.
- **Efficiency:** Identifies areas where time, money, or resources are being wasted, so processes can be improved.
- **Customer Understanding:** Reveals what customers like, dislike, or need, so services/products can be improved.
- **Competitive Advantage:** Companies using data analytics can react faster to market changes and stay ahead of competitors.

3. Steps in the Data Analytics Process

1. Data Collection:

- Gather data from different sources like databases, websites, sensors, surveys, or files.
- Data can be structured (tables, spreadsheets) or unstructured (text, images).

2. Data Management/Storage:

- Store and organize data in databases (like SQL databases) for easy access and management.

3. Data Cleaning:

- Remove errors, duplicates, and fix missing or inconsistent values to make data reliable for analysis.

4. Data Analysis:

- Use statistical methods, algorithms, or machine learning to find patterns, trends, and relationships in the data.
- Techniques include regression, clustering, classification, and time-series analysis.

5. Data Visualization:

- Present findings using charts, graphs, dashboards, or reports so they are easy to understand.

6. Interpretation & Reporting:

- Draw conclusions from the analysis and share insights with decision-makers for action.

➤ Types of Data Analytics

Data analytics is divided into four main types.

1. Descriptive Analytics

- **Definition:** Descriptive analytics summarizes and explains what has happened in the past using data.
- **Purpose:** Gives a clear picture of past events, trends, and patterns.
- **How it works:** Uses tools like tables, charts, graphs, averages, and totals to present data in a simple way.

Advantages:

- **Easy to understand:** Summarizes complex data into simple charts, graphs, and tables, making it accessible for everyone.
- **Cost-effective:** Requires basic tools and less technical expertise, so it's cheaper than advanced analytics.
- **Supports decision-making:** Provides a clear view of past events, helping organizations make informed decisions.

Disadvantages:

- **Limited to past data:** Only shows what happened, not why or what will happen next.
- **No cause analysis:** Does not explain reasons behind trends or patterns.
- **No predictions:** Cannot forecast future outcomes or challenges.
- **Depends on data quality:** If data is wrong or incomplete, results can be misleading.

Examples:

- **Monthly sales reports showing sales growth or decline.**
- **Website analytics showing number of visitors each week.**
- **Product review summaries.**
- **Keywords: Summary, trends, patterns, reporting, visualization, past data.**

2. Diagnostic Analytics

- **Definition:** Diagnostic analytics focuses on why something happened by finding the root cause of events or changes.
- **Purpose:** Helps organizations understand the reasons behind outcomes or problems.
- **How it works:** Uses methods like correlation, regression analysis, drill-down, and data mining to identify relationships and causes.

Advantages:

- **Finds root causes:** Helps understand why something happened by digging deeper into data.
- **Improves problem-solving:** Identifies reasons for issues, so organizations can fix problems effectively.
- **Supports better decisions:** Knowing the cause helps make more targeted and useful decisions.

Disadvantages:

- **Needs skilled analysts:** Requires more technical knowledge and expertise than descriptive analytics.

- **Time-consuming:** Deep analysis can take longer, especially with large or complex datasets.
- **Depends on data quality:** Wrong or incomplete data can lead to wrong conclusions about causes.

Examples:

- Analyzing why sales dropped in a particular month.
- Finding out why a machine failed in a factory.
- Understanding the cause of a sudden increase in website traffic.

Keywords: Root cause, correlation, regression, drill-down, explanation, why.

3. Predictive Analytics

- **Definition:** Predictive analytics uses current and historical data to forecast what might happen in the future.
- **Purpose:** Helps organizations anticipate trends, risks, and opportunities.
- **How it works:** Uses statistical models, machine learning, and forecasting techniques to make predictions.

Advantages:

- **Forecasts future trends:** Uses historical data to predict what might happen next, helping with planning and risk management.
- **Supports proactive decisions:** Organizations can prepare for future challenges and opportunities.
- **Uses machine learning:** Can handle large and complex data for more accurate predictions.

Disadvantages:

- **Needs high-quality data:** Predictions are only as good as the data used; poor data leads to poor forecasts.
- **Complex and costly:** Requires advanced tools, algorithms, and skilled professionals, which can be expensive.

- **Not always accurate:** Predictions are based on patterns, but unexpected changes can make them wrong.

Examples:

- Predicting customer churn (who might stop using a service).
- Forecasting demand for a product.
- Credit scoring in banks.

Keywords: Forecasting, trends, machine learning, prediction, future, modeling.

4. Prescriptive Analytics

- **Definition:** Prescriptive analytics suggests what actions to take for the best possible outcome, based on data and predictions.
- **Purpose:** Helps organizations decide the best course of action among different options.
- **How it works:** Uses optimization, simulation, and advanced algorithms (often with AI and machine learning) to recommend solutions.

Advantages:

- **Recommends best actions:** Suggests what to do for the best results, helping organizations make optimal decisions.
- **Handles complex scenarios:** Can analyze many options and constraints to find the best solution.
- **Supports automation:** Can be used in automated systems for real-time decision-making.

Disadvantages:

- **Very complex:** Needs advanced algorithms, lots of data, and expert knowledge to set up and maintain.
- **Expensive:** High cost for tools, technology, and skilled staff.

- Depends on accurate predictions: If predictive analytics is wrong, prescriptive recommendations may also be wrong.
- Difficult to implement: Integrating into existing systems and processes can be challenging.

Examples:

- Recommending the best marketing strategy to increase sales.
- Suggesting inventory levels to avoid shortages or excess.
- Automated loan approval systems.

Keywords: Recommendation, optimization, action, solution, decision, best outcome.

Type	Main Question	Main Use	Example
Descriptive	What happened?	Summarize past data	Sales reports, web analytics
Diagnostic	Why did it happen?	Find root causes	Analyzing sales drop
Predictive	What might happen?	Forecast future trends	Demand forecasting, credit risk
Prescriptive	What should we do?	Recommend actions	Best pricing, inventory levels

➤ Applications of Data Analytics

Business & Marketing

- Product recommendation: Suggests products based on past user behavior (e.g., Amazon, Netflix).
- Sales forecasting: Predicts future sales using historical data.

Healthcare

- Disease prediction: Identifies high-risk patients by analyzing medical data.
- Medical image analysis: AI helps doctors read X-rays, MRIs.

- **Operational efficiency:** Reduces patient wait times, manages resources.

Finance & Banking

- **Fraud detection:** Spots suspicious transactions in real time.
- **Credit scoring:** Evaluates loan risk using customer data.
- **Risk management:** Identifies and manages financial risks.

Education

- **Student performance analysis:** Tracks attendance, grades, and progress.

➤ **Benefits of Data Analytics**

- **Improved Decision Making:** Accurate, timely decisions using data.
- **Operational Efficiency:** Finds inefficiencies in processes and supply chains.
- **Cost Reduction:** Detects overspending and resource wastage.
- **Customer Satisfaction:** Understands and meets customer needs.
- **Early Problem Detection:** Identifies risks, fraud, or issues before they grow.
- **Product/Service Improvement:** Uses feedback to make things better.
- **Business Growth:** Helps plan for the future and expand smartly.

➤ **DIFFERENTIATIONS**

○ **DATA ANALYSIS AND DATA ANALYTICS**

Aspect	Data Analysis	Data Analytics
Definition	Inspecting, cleaning, interpreting data to find patterns or insights.	Collecting, cleaning, transforming, and modeling data to support decision making.

Aspect	Data Analysis	Data Analytics
Scope	focuses on examining datasets.	covers the entire cycle: collection, cleaning, analysis, visualization, interpretation.
Goal	Understand and interpret data.	Gain insights and guide business strategies or actions.
Approach	Manual or statistical examination.	Uses statistical, and technology-based methods.
Example	Checking why sales dropped last month.	Using multiple datasets and predictive models to forecast and improve future sales.

- DATA ANALYST AND DATA SCIENTIST

Aspect	Data Analyst	Data Scientist
Main Role	Examines historical data to find trends, patterns, and insights.	Builds models and algorithms to predict future outcomes and uncover complex patterns.
Goal	Explain what happened and why.	Predict what will happen and how to optimize outcomes.
Skills	Data cleaning, descriptive statistics, visualization, business reporting.	Machine learning, advanced statistics, , big data handling.
Programming	Moderate (mostly for querying and visualization).	High (for building models, automation, data pipelines).

Aspect	Data Analyst	Data Scientist
Tools	Excel, SQL, Tableau, Power BI, basic Python/R.	Python, TensorFlow, Py Torch, advanced ML libraries.
Data Handling	Mostly works with structured data (tables, spreadsheets, databases).	Works with structured, semi-structured, and unstructured data (images, text, audio, predictive models, AI).
End Deliverables	Reports and data summaries.	Solutions and decision-support systems.

- **BUSINESS ANALYST AND BUSINESS INTELLIGENCE ANALYST**

Aspect	Business Analyst	Business Intelligence Analyst
Main Focus	Understand business needs, improve processes, define requirements for solutions.	Analyze business data and create reports/dashboards to guide decision-making.
Goal	Bridge the gap between business and technical teams to deliver solutions that meet business objectives.	Provide data-driven insights to help the business track performance and plan strategies.
Work Nature	More process and requirement focused.	More data and analysis focused.
Skills	Process mapping, stakeholder management, documentation, requirements gathering.	Data analysis, visualization, KPI tracking.

Aspect	Business Analyst	Business Intelligence Analyst
Tools	Confluence, Lucidchart, Visio, Jira, Microsoft Office suite.	SQL, PowerBI, Tableau, Excel, Looker, Qlik.
End Deliverables	Requirement docs, business cases.	Interactive dashboards, data visualization, trend insights.

➤ **Role and Responsibilities of a Data Analyst**

- **Data Collection:** Gather data from databases, files, or online platforms.
- **Data Cleaning & Preparation:** Remove duplicates, handle missing/incomplete data, correct inconsistent formats, normalize data for analysis.
- **Data Analysis & Interpretation:** Study data to find patterns, trends, or useful information.
- **Data Visualization & Reporting:** Create charts, graphs, and reports to present findings clearly.
- **Database Querying:** Use tools like SQL to pull specific information from databases.
- **Collaboration & Communication:** Share results and work with teams to make better decisions.

➤ **Skills Required for Data Analyst**

- **Technical Skills:**
 - **Statistical knowledge:** Descriptive (mean, median, mode, standard deviation), inferential (hypothesis testing, correlation, regression), probability concepts.
 - **Data analysis tools:** Excel (pivot tables, formulas), SQL (queries, joins, aggregations), Python (numpy, pandas, matplotlib, seaborn), R (tidyverse, ggplot2).
 - **Data visualization tools:** PowerBI, Tableau, Data Studio, Qlik Sense.
 - **Database management:** MySQL, PostgreSQL, SQL Server.
 - **Data cleaning techniques:** Handling outliers, missing values, inconsistent entries.

- **Soft Skills:**
 - **Critical thinking:** Analyze information objectively and make reasoned judgments.
 - **Problem solving:** Find practical solutions using data insights.
 - **Communication:** Share ideas clearly.
 - **Collaboration:** Work well with others.
 - **Attention to detail:** Notice small mistakes.
 - **Time management:** Use time wisely.