

Name: Ruth Asekhamhe(DA2).

rasekhamhe@gmail.com

Types of Databases:

1: NoSQL Databases: This type of data is kept as a collection of key-value pairs just like a dictionary or hash table. It provides incredibly quick lookups. They are used for profile storage, session management, and caching.

Examples; Redis, and Memcached.

2: Column Databases: Data is kept in columns rather than rows and can be grouped into “column families”. It is perfect for querying enormous datasets by specific columns. They are used for enormous-scale data processing, time-series data, and real-time analytics.

Examples; HBase, and Cassandra.

3: Graph Databases: It is used to represent and store data, use nodes (entities), and edges (relationships). They are extremely well-suited for swiftly navigating intricate connections. They are used for recommendation engines, fraud detection, and social networks.

Examples; Amazon Neptune, and Neo4j.

4: Time-Series Databases: They are designed to handle and query time-indexed data, frequently from IoT devices, monitoring tools, or sensors. It is used for tracking stock prices, application performance, and IoT device data.

Examples; InfluxDB, and Prometheus.

Different Method of Collecting Data Tools:

1: Surveys/Questionnaires: It uses a series of standardized questions to collect data from a sample of individuals.

Tools Used; Google Forms, Qualtrics, and SurveyMonkey.

2: Observation: Methodically observing and documenting actions, occurrences, or artifacts in a natural environment.

Tools Used; Cameras, Mobile observation apps, and Field notes.

3: Sensors/IoT: Automated data gathering from physical devices, such as activity trackers, flow meters, and temperature gauges.

Tools Used; AWS IoT, Azure IoT Hub, and Smart meters.

4: Interviews: Comprehensive, one-on-one or small-group discussions to extract specific details regarding a subject’s viewpoints or experiences.

Tools Used; Microsoft Teams, Zoom, and Notion.

5: Social Media: Gathering user-generated content from social platforms.

Tools Used; Twitter, BrandWatch, and Meta API.

Differences between Data Analysis, Data Science, and Data Engineering:

Data Analysis: It is used for business insights, and reporting. Dashboards, reports, and responses to certain business queries (What happened, Why?). The core skills of data analysis are SQL, Excel, and Visualization Tools.

Data Science: Predictive Modeling and Experimentation. Its predictive models include forecasts, classifiers, and statistical inferences (What will happen? How can we improve?) The core skills of data science are Python/R, Machine Learning, and Advanced Statistics.

Data Engineering: Infrastructure and data flow. Reliable, scalable data pipelines. Data engineering core skills are SQL, Data Warehousing (Snowflake), and Big Data Tools (Spark).

Differences and Similarities between Data Analysis and Business Intelligence Analysis:

Differences:

Data Analysis is more expansive, applicable to any data set (academic, scientific, business, e.t.c.). It frequently conducts ad hoc (one-off) analysis to test theories or provides answers to particular, in-depth questions. It entails more thorough statistical investigation while Business Intelligence Analysis is particularly concentrated on business data to assist in making operational decisions. It uses both past and present data to track KPIs, monitor performance, and create consistent reports.

Similarities:

Both make extensive use of SQL for data searching and exhibit results using data visualization programs like Tableau and Power BI. The ultimate objective for both is to draw conclusions from data that can be used to improve decision-making. Both employ exploratory data analysis, data purification, and descriptive statistics.

Applications of Data Analysis in Different Industries:

- 1: Finance: Fraud detection, risk management, portfolio analysis, credit scoring.
- 2: Education: Student performance tracking, curriculum improvement, dropout prediction, resource allocation.
- 3: Healthcare: Patient diagnosis support, hospital performance tracking, disease outbreak prediction, medical outreach.
- 4: Retail and E-Commerce: Inventory optimization, pricing strategy analysis, customer segmentation, product recommendation systems.
- 5: Manufacturing: Quality control analytics, production efficiency analysis, supply chain optimization, predictive maintenance.

