Name: Oguejiofor Viola

Email: violapurisima@gmail.com

ASSIGNMENT

1. Describe the other types of databases with examples

- DocumentOriented Database: Stores data as JSONlike documents, ideal for hierarchical data.

  Examples: MongoDB, CouchDB.

- KeyValue Database: Uses a simple keyvalue pair model for fast retrieval.

  Example: Redis, DynamoDB.

- ColumnFamily / WideColumn Database: Organizes data by columns rather than rows, good for analytical queries.

  Example: Apache Cassandra, HBase.

- ObjectOriented Database: Stores data as object with methods and properties, aligning with OOP languages.

  Example: db4o, ObjectDB.

- Data Warehouse: Central repository for analytical and reporting workloads, often denormalized.

  Example: Amazon Redshift, Google BigQuery.

- InMemory Database: Keeps data in RAM for ultrafast access and low latency.

  Example: Redis, MemSQL.

- Distributed Database: Spreads data across multiple nodes or locations for scalability and fault tolerance.

  Example: CockroachDB, Google Cloud Spanner.

- Cloud Database: Hosted and managed by a cloud provider, offering scalability and automatic maintenance.

Example: Amazon Aurora, Microsoft Azure Cosmos DB.

2. Describe the different methods of collecting data and tools used

*Methods:

- Surveys/Questionnaires– structured or   unstructured data from respondents.

- Observations– recording behavior directly.

- Interviews– In-depth, oneonone   conversations.

- Experiments– controlled tests to observe outcomes.

- Secondary sources– existing data (reports, databases).

*Tools:

- Google Forms, SurveyMonkey (surveys).

- SPSS, SAS (statistical analysis).

- Cameras, sensors (observation).

- Lab equipment, online experiment platforms (experiments).

- Excel, CSV files, APIs, web scraping libraries (PythonBeautifulSoup, Scrapy).

3. Difference between Data Analysis, Data Science and Data Engineering

:Data Analysis

Data analysis focuses on inspecting, cleaning, and transforming data to extract useful information and support decisionmaking. It usually answers specific, often retrospective questions ("What happened?" or "Why did it happen?"). Analysts work with structured datasets, apply descriptive and sometimes diagnostic statistics, and present results through reports or visualizations. Typical skills include SQL, Excel, Tableau, and basic statistical knowledge.

:Data Science

Data science goes beyond analysis by using advanced techniques to build predictive or prescriptive models and uncover hidden patterns. It combines expertise in mathematics, statistics, programming, and domain knowledge to solve complex, forwardlooking problems ("What could happen?" or "What should we do?"). Data scientists often work with large, unstructured data, employ machinelearning algorithms, and develop datadriven solutions. Common tools are Python, R, TensorFlow, and Spark.

:Data Engineering

Data engineering centers on designing, constructing, and maintaining the infrastructure that stores and processes data. It ensures data is available, reliable, and scalable for downstream use by analysts and scientists. Engineers build data pipelines, manage databases, and handle ETL (extract, transform, load) processes. Their skill set typically includes SQL, Hadoop, cloud platforms, and programming languages such as Java or Python for data architecture.

In Summary;

- Data Analysis– interpret existing data to answer concrete questions.

- Data Science– create models and predictions using sophisticated methods.

- Data Engineering– build and maintain the data pipelines and storage systems that enable the former two.

4. Difference and similarities between Data Analysis and Business Intelligence (BI) Analysis

*Differences:

-Focus: Data Analysis; uncovers patterns/insights for any purpose; BI Analysis: supports business decisionmaking and reporting.

-Tools: Data Analysis; often R, Python, Excel, While; BI: Power BI, Tableau, Looker.

*Similarities:

- Both involve querying, cleaning, and visualizing data.

- Both aim to turn data into actionable information.

- Both require understanding of business context and stakeholders.

5. Outline some applications of Data Analysis in different industries

- Healthcare– patient outcome prediction, resource allocation.

- Finance– fraud detection, credit scoring, risk management.

-Retail– customer segmentation, inventory optimization.

-Marketing– campaign performance tracking, churn analysis.

-Manufacturing– predictive maintenance, qualitycontrol monitoring.

-Transportation– route optimization, demand forecasting.