Chapter 7

1. Cloud Computing:

- **Definition:**
    - Cloud computing is a technology that enables access to a pool of computing resources over the internet, allowing users to use and pay for resources as needed.
- **Key Characteristics:**
    - *On-demand self-service:* Users can provision and manage computing resources as needed.
    - *Broad network access:* Services are available over the network and can be accessed by various devices.
    - *Resource pooling:* Resources are shared among multiple users to maximize efficiency.
    - *Rapid elasticity:* Resources can be rapidly scaled up or down based on demand.
    - *Measured service:* Usage is monitored, and users pay for the resources they consume.
- **Service Models:**
    - *Infrastructure as a Service (IaaS):* Provides virtualized computing resources over the internet.
    - *Platform as a Service (PaaS):* Offers a platform allowing customers to develop, run, and manage applications without dealing with the complexity of infrastructure.
    - *Software as a Service (SaaS):* Delivers software applications over the internet on a subscription basis.
- **Deployment Models:**
    - *Public Cloud:* Services are provided over the internet and are available to the general public.
    - *Private Cloud:* Cloud infrastructure is exclusively used by a single organization.
    - *Hybrid Cloud:* Combination of public and private clouds.

2. Data Management:

- **Definition:**
    - Data management involves the planning, execution, and supervision of activities related to the collection, storage, retrieval, and use of data.
- **Key Components:**
    - *Data Governance:* Establishing policies and procedures for managing data assets.
    - *Data Quality:* Ensuring the accuracy, completeness, and reliability of data.
    - *Master Data Management (MDM):* Managing and maintaining consistent and accurate business data.
    - *Data Integration:* Combining data from different sources to provide a unified view.
    - *Data Security:* Protecting data from unauthorized access and ensuring confidentiality.
    - *Data Warehousing:* Storing and managing data from different sources in a centralized repository.
- **Challenges:**
    - *Data Silos:* Isolated storage of data that hinders collaboration and data access.
    - *Data Security:* Protecting sensitive data from breaches and unauthorized access.
    - *Data Quality Issues:* Inaccurate or incomplete data impacting decision-making.
    - *Compliance and Regulations:* Adhering to data protection and privacy laws.

3. Mobile Databases:

- **Definition:**
    - Mobile databases are designed to operate in mobile and sometimes offline environments, allowing applications to store and retrieve data locally.
- **Characteristics:**
    - *Lightweight:* Optimized for mobile devices with limited resources.
    - *Synchronization:* Capabilities to synchronize data with a central server.
    - *Offline Support:* Ability to operate without a constant network connection.
- **Examples:**

- *SQLite:* A self-contained, serverless, and zero-configuration relational database engine.
- *Realm:* A mobile database that is easy to integrate and provides real-time synchronization.
- *Firebase Realtime Database:* A NoSQL database that enables real-time data synchronization across devices.

4. Hadoop:

- **Definition:**
  - Hadoop is an open-source framework for distributed storage and processing of large datasets using a cluster of commodity hardware.
- **Components:**
  - *Hadoop Distributed File System (HDFS):* A distributed file system for storage.
  - *MapReduce:* A programming model for processing and generating large datasets.
- **Advantages:**
  - *Scalability:* Easily scales by adding more nodes to the Hadoop cluster.
  - *Fault Tolerance:* Can recover from hardware or software failures.
  - *Cost-Effectiveness:* Uses inexpensive, commodity hardware.
- **Use Cases:**
  - Big data processing, analytics, and storage of large datasets.

5. SQLite Database:

- **Definition:**
  - SQLite is a self-contained, serverless, and zero-configuration relational database engine.
- **Characteristics:**
  - *Embedded Database:* No separate server process and is embedded directly into the application.
  - *Lightweight:* Consumes minimal memory and disk space.
  - *ACID Properties:* Guarantees the reliability of transactions.
- **Use Cases:**
  - Embedded systems, mobile applications, small to medium-sized websites.

6. SQL-MongoDB:

- **SQL (Relational Database):**
    - *Structured Query Language:* A language used for managing and querying relational databases.
    - *Tables with Predefined Schema:* Data is organized into tables with a fixed structure.
    - *ACID Properties:* Ensures reliable processing of database transactions.
- **MongoDB (NoSQL Database):**
    - *JSON-like Documents (BSON):* Data is stored in flexible, JSON-like documents in BSON format.
    - *No Fixed Schema:* Each document in a collection can have a different structure.
    - *Horizontal Scalability:* Scales horizontally by adding more servers.
- **Use Cases:**
    - SQL is suitable for applications with structured and tabular data.
    - MongoDB is suitable for applications with unstructured or semi-structured data, or where flexible schema design is required.