EMBA Program
MB-511

Data Science for Managerial Decisions (MB 511)
Introduction to Data Science

Instructor
Anant Prakash Awasthi

Data Science for Managerial Decisions (MB 511)

EMBA Program
MB-511

Program Overview

- Introduction to Data Science
- Information Technology: An Overview
- Applications of Data Science in various fields
- MIS and Control Systems
- Data Collection and Data Pre-Processing
- Building Information Systems
- Support Systems for Management Decisions

Information Technology: An Overview

EMBA Program
MB-511

- Foundations of Information Technology
- Information Systems and Management
- Cybersecurity and IT Governance
- Future Trends in Information Technology

Information Technology: An Overview

Basic Operations in RDBMS

EMBA Program
MB-511

CRUD Operations:

- Create: Adding new data (e.g., adding a new customer).
- Read: Retrieving data (e.g., finding the price of a product).
- Update: Modifying existing data (e.g., updating a customer's address).
- Delete: Removing data (e.g., deleting an outdated product).

Benefits and Limitations:

Benefits:

Efficiency: Helps manage large data sets.

Security: Ensures only authorized people access sensitive data.

Accuracy: Reduces human error in managing data.

Limitations:

Complexity: Requires initial setup and learning.

Cost: High-end DBMS solutions can be expensive.

Information Technology: An Overview

Properties of RDBM(S)

EMBA Program
MB-511

ACID Property	Definition	Explanation	Example
A - Atomicity	All or nothing	Ensures that all parts of a transaction are completed. If one part fails, the entire transaction is rolled back, preventing partial updates.	If you transfer money between two accounts, both the debit and credit must succeed, or neither happens.
C - Consistency	Guarantees data integrity	Ensures that the database transitions from one valid state to another, maintaining all defined rules and constraints.	After a transaction, the balance of a bank account must meet the business rules (e.g., no negative balances).
I - Isolation	Transactions are independent	Ensures that the execution of one transaction is isolated from others, preventing concurrent transactions from interfering with each other.	Two users updating the same account will not see intermediate changes made by each other until completion.
D - Durability	Changes are permanent	Guarantees that once a transaction is committed, it will persist even in the event of a system crash or failure.	Once a payment transaction is confirmed, it remains saved in the system even if the server crashes afterward.

Information Technology: An Overview

Open Source RDBMS

EMBA Program
MB-511

Database Name	Key Features	Common Use Cases	Website
MySQL	Widely used, supports ACID transactions, scalability, and reliability.	Web applications, data warehousing	https://www.mysql.com/
PostgreSQL	Advanced SQL features, highly extensible, supports JSON, XML, and custom data types.	Business applications, analytics, research	https://www.postgresql.org/
MariaDB	Fork of MySQL, improved performance, advanced clustering.	Enterprise systems, cloud applications	https://mariadb.org/
SQLite	Lightweight, serverless, used for embedded databases.	Mobile applications, browsers, local databases	https://www.sqlite.org/
Firebird	Lightweight, cross-platform, supports stored procedures and triggers.	Embedded systems, small to medium databases	https://firebirdsql.org/
H2 Database	Lightweight, Java-based, embedded or server modes.	Java applications, testing databases	https://h2database.com/
Percona Server	Fork of MySQL, high performance, scalability, enterprise features.	High-traffic applications, e-commerce, finance	https://www.percona.com/
CockroachDB	Distributed SQL, designed for scalability and strong consistency.	Distributed systems, fault-tolerant applications	https://www.cockroachlabs.com/
Apache Derby	Lightweight, embedded, pure Java.	Embedded systems, testing, lightweight applications	https://db.apache.org/derby/

Information Technology: An Overview

Perpetual RDBMS

EMBA Program
MB-511

Database Name	Type	License	Key Features	Common Use Cases	Website
Ingres	Relational (SQL)	Perpetual License	Enterprise-level, highly secure, mission-critical applications.	Banking, telecommunications, government systems	https://www.actian.com/
Microsoft SQL Server	Relational (SQL)	Perpetual License	Enterprise-grade, highly scalable, robust analytics and business intelligence.	Enterprise systems, financial services, big data	https://www.microsoft.com/sqlserver/
Oracle Database	Relational (SQL)	Perpetual License	Comprehensive RDBMS with built-in AI and machine learning.	Large enterprises, mission-critical applications	https://www.oracle.com/database/
IBM Db2	Relational (SQL)	Perpetual License	AI-powered, highly scalable, with cloud, hybrid, and on-premise deployment.	Enterprise solutions, data lakes, and analytics	https://www.ibm.com/products/db2

Information Technology: An Overview

Types of NoSQL databases

EMBA Program
MB-511

- Document databases (like MongoDB, CouchDB) store data in a flexible, JSON-like format, allowing for varied data structures.
- Wide Column stores (like Cassandra, HBase) are excellent for handling large-scale data, especially for time series or real-time applications.
- Key-Value stores (like Redis, DynamoDB) provide simple key-value pairs and are very fast for operations like caching and session management.
- Graph databases (like Neo4j, Titan) are optimized for managing and querying relationships between entities, useful in social networks or fraud detection.
- Multi-model databases (like ArangoDB, Cosmos DB) support more than one data model, making them versatile for complex use cases.

Information Technology: An Overview

Business Applications of NoSQL databases

EMBA Program
MB-511

Database Name	Type	Key Features	Common Use Cases	Website
MongoDB	Document	Flexible schema, high scalability, supports JSON-like documents, ACID transactions.	Content management, real-time analytics, big data	https://www.mongodb.com/
CouchDB	Document	Easy replication, supports JSON, fault-tolerant, RESTful HTTP API.	Web applications, distributed data storage	https://couchdb.apache.org/
RavenDB	Document	ACID transactions, fully distributed, multi-model, integrated indexing.	Business applications, document management	https://ravendb.net/
MarkLogic	Document, Graph	High performance, integrates with enterprise systems, ACID transactions.	Government, healthcare, media, and publishing	https://www.marklogic.com/
Fauna	Document, Graph	Global distribution, ACID compliance, serverless, supports GraphQL.	SaaS applications, financial services	https://fauna.com/
Couchbase	Document, Key-Value	High performance, easy scalability, flexible schema, supports SQL-like queries.	Content management, mobile, and IoT applications	https://www.couchbase.com/
Elasticsearch	Document, Search	Full-text search engine, distributed, powerful query DSL, near real-time search.	Log/metric analysis, search engines, analytics	https://www.elastic.co/

Information Technology: An Overview

Business Applications of NoSQL databases

EMBA Program
MB-511

Database Name	Type	Key Features	Common Use Cases	Website
Neo4j	Graph	Optimized for relationships, powerful query language (Cypher), ACID compliance.	Social networks, fraud detection, recommendation engines	https://neo4j.com/
Amazon Neptune	Graph	Managed by AWS, supports both property graphs and RDF, highly available.	Knowledge graphs, social networking, fraud detection	https://aws.amazon.com/neptune/
Titan	Graph	Distributed, highly scalable, designed to work with large graphs.	Social networks, real-time recommendation systems	https://thinkaurelius.github.io/titan/
Redis	Key-Value	In-memory data store, extremely fast, supports complex data types (lists, sets).	Caching, real-time messaging, leaderboards	https://redis.io/
DynamoDB	Key-Value, Document	Managed by AWS, fully serverless, auto-scaling, low-latency performance.	IoT, real-time bidding, mobile backends	https://aws.amazon.com/dynamodb/
Tarantool	Key-Value, Document	In-memory computing, high performance, supports Lua scripting.	High-performance applications, caching	https://www.tarantool.io/
HyperDex	Key-Value, Document	High availability, distributed, consistent hashing.	Scalable applications, big data	https://hyperdex.org/

Information Technology: An Overview

Introduction to cloud computing

EMBA Program
MB-511

Cloud computing is a revolutionary paradigm in the field of information technology that has transformed the way individuals and organizations access, store, and process data. At its core, cloud computing refers to the delivery of computing services over the internet, providing users with on-demand access to a wide range of resources without the need for direct management or ownership of physical infrastructure.

Types of Cloud Services :

Public Cloud:

Definition: In a public cloud, cloud services are provided by third-party service providers over the internet. These services are made available to the general public, and multiple organizations or users can share the same infrastructure.

Advantages: Cost-effective, scalable, and easy to access. Users don't need to invest in or manage their own infrastructure.

Private Cloud:

Definition: A private cloud is operated solely for a single organization. It may be managed internally or by a third-party provider. The infrastructure is dedicated to the organization, offering greater control and customization.

Advantages: Enhanced security, control, and customization. Ideal for organizations with specific compliance or data privacy requirements.

Microsoft Azure amazon web services ORACLE Google Cloud

Information Technology: An Overview

Introduction to cloud computing

EMBA Program
MB-511

Types of Cloud Services :

Hybrid Cloud:

Definition: Hybrid cloud combines elements of both public and private clouds. It allows data and applications to be shared between them, offering greater flexibility. Organizations can use public cloud resources for non-sensitive functions and maintain critical workloads in a private cloud.

Advantages: Balances the benefits of both public and private clouds, providing flexibility, scalability, and the ability to meet specific security and compliance needs.

Multi-Cloud:

Definition: Multi-cloud involves using services from multiple cloud providers. Organizations may choose different cloud providers for different services or applications, avoiding vendor lock-in and optimizing for specific features.

Advantages: Redundancy, flexibility, and the ability to leverage the strengths of different cloud providers for various use cases.

Microsoft Azure amazon web services ORACLE Google Cloud

Information Technology: An Overview

Introduction to cloud computing

EMBA Program
MB-511

Key components and characteristics of cloud computing include:

On-Demand Self-Service: Users can provision and manage computing resources, such as server instances, storage, and applications, as needed without requiring human intervention from service providers.

Broad Network Access: Cloud services are accessible over the internet from a variety of devices, including laptops, smartphones, and tablets. This accessibility ensures that users can connect to cloud resources from virtually anywhere.

Resource Pooling: Cloud providers pool computing resources to serve multiple customers. Resources, such as storage and processing power, are dynamically assigned and reassigned based on demand, optimizing efficiency.

Rapid Elasticity: Cloud resources can be quickly scaled up or down to accommodate changes in demand. This elasticity allows users to scale resources seamlessly, ensuring optimal performance and cost-effectiveness.

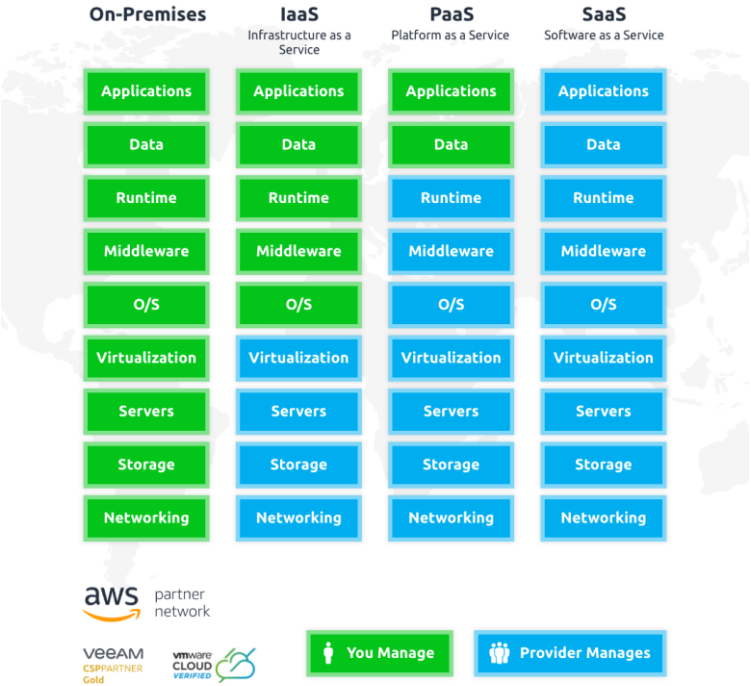
Measured Service: Cloud computing resources are metered, and users are billed based on their actual usage. This pay-as-you-go model provides cost efficiency, as users only pay for the resources they consume.

Microsoft Azure amazon web services ORACLE Google Cloud

Information Technology: An Overview

Introduction to cloud computing – Service Models

EMBA Program
MB-511



1. On-Premise: All components are managed by the business.

2. Infrastructure as a Service (IaaS): Provides virtualized computing resources over the internet. Users can rent virtual machines, storage, and networking components.

3. Platform as a Service (PaaS): Offers a platform that includes infrastructure and tools to develop, test, and deploy applications. Users can focus on application development without managing the underlying infrastructure.

4. Software as a Service (SaaS): Delivers software applications over the internet, eliminating the need for users to install, maintain, and update software locally.

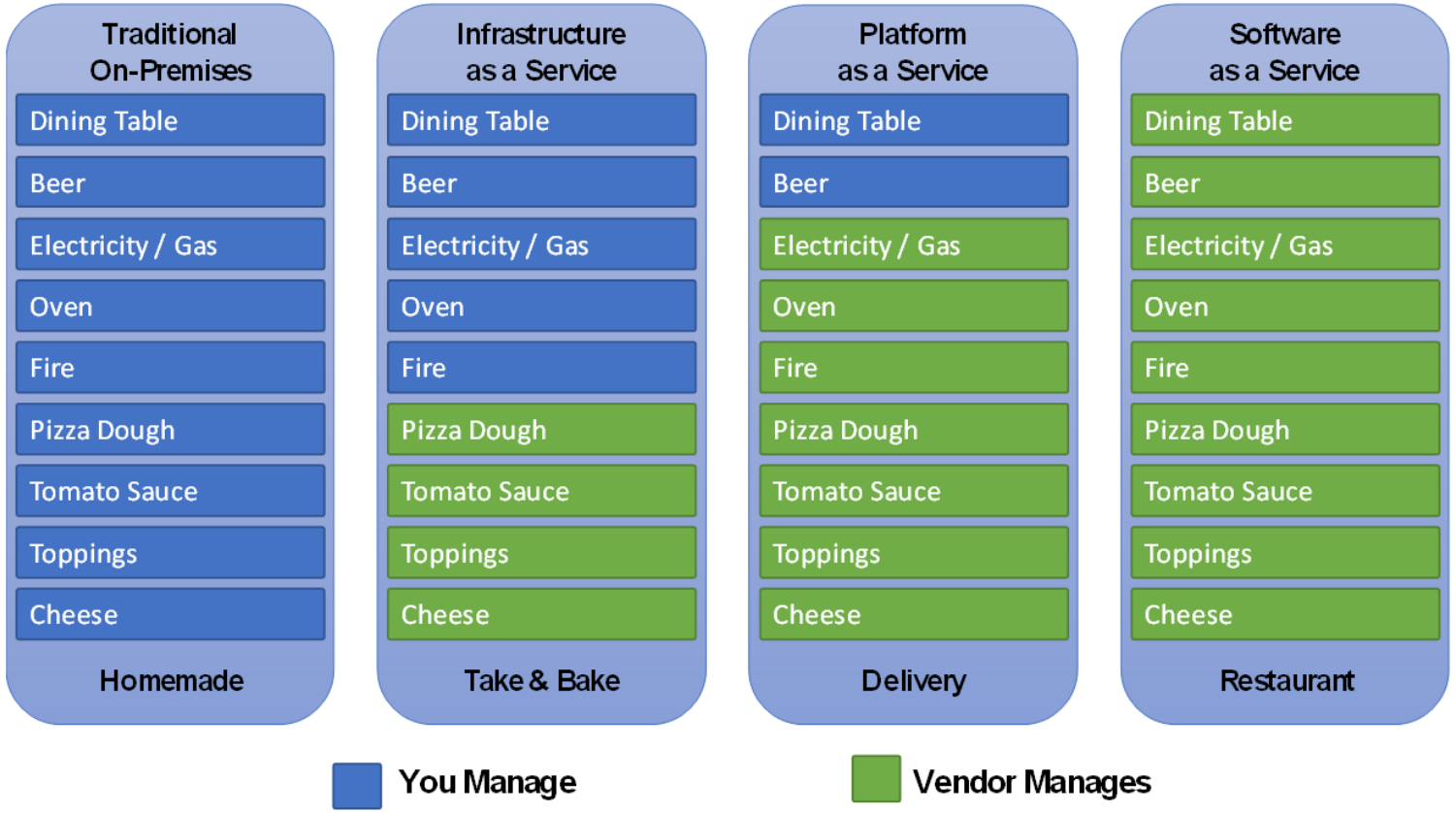
Microsoft Azure amazon web services ORACLE Google Cloud

Information Technology: An Overview

Introduction to cloud computing – Service Models

EMBA Program
MB-511

Pizza as a Service



Microsoft Azure amazon web services ORACLE Google Cloud

Information Technology: An Overview

Introduction to cloud computing

EMBA Program
MB-511

Your Cloud Journey:

AI Engineer

AZ900 >> DP900 >> AI900 >> AI100 >> AI101

Data Analyst

AZ900 >> DP900 >> AI900 >> PL300 >> DA100 >> 70-778 >> 70-779

Data Engineer

AZ900 >> DP900 >> AI900 >> DP201 >> DP203 >> DP420

Data Scientist

AZ900 >> DP900 >> AI900 >> DP100

Microsoft Azure amazon web services ORACLE Google Cloud

Information Technology: An Overview

Information Systems and Management

EMBA Program
MB-511

What are information systems:

Information Systems (IS) in business refer to the organized, coordinated, and integrated systems that collect, process, store, and distribute information within an organization. These systems play a crucial role in supporting business operations, decision-making processes, and strategic planning. Information Systems in business can be categorized into various types, each serving specific functions.

What is importance information systems:

The importance of information systems in various aspects of modern life and business cannot be overstated. Some key aspects include:

- Enhanced Decision-Making
- Improved Efficiency
- Increased Productivity
- Strategic Planning
- Competitive Advantage
- Customer Relationship Management (CRM)
- Global Connectivity:
- Risk Management
- Innovation and Adaptability
- Compliance and Security

Information Technology: An Overview

Information Systems and Management

EMBA Program
MB-511

Examples of information systems:

- Transaction Processing Systems (TPS): Record and process day-to-day transactions, supporting routine business operations.
- Management Information Systems (MIS): Provide summarized, structured data for middle management to facilitate decision-making.
- Decision Support Systems (DSS): Assist in complex decision-making by providing analytical tools and interactive capabilities.
- Executive Information Systems (EIS): Offer top-level executives strategic information for long-term planning and decision-making.
- Enterprise Resource Planning (ERP): Integrate and manage various business processes and functions across an organization.
- Supply Chain Management Systems (SCM): Manage the flow of goods, services, and information throughout the supply chain.
- Customer Relationship Management (CRM): Focus on managing interactions with customers to enhance relationships and satisfaction.
- Knowledge Management Systems (KMS): Capture, organize, and disseminate organizational knowledge to improve decision-making.
- Business Intelligence Systems (BI): Collect, analyze, and present business data to support informed decision-making.
- Collaboration Systems: Facilitate communication and collaboration among individuals and teams within an organization.

EMBA Program
MB-511

Have a question?

Feel Free to Reach out at

- +91-88846-52929 (WhatsApp)
- anant.awasthi@outlook.com (E-Mail)