

Bases de Dados

Engenharia Informática (2ºano)

Tecnologias Digitais e Gestão (2ºano)

Capítulo 2. Sistemas de Gestão de Bases de Dados (SGBD)

João Paulo Pereira | jprp@ipb.pt
Marisa Ortega | marisa.ortega@ipb.pt
David Dias | davide.dias@ipb.pt
Tiago Santos | tiago.santos@ipb.pt

2025



Índice

● Conteúdo

- ➔ 1. Introdução aos ambientes de Base de Dados
 - ⇒ - Noção de Sistema de Informação
 - ⇒ - A Informação nas Organizações
 - ⇒ - Tecnologias de Informação
 - ⇒ - Gestão de Informação
- ➔ 2. Sistemas de Gestão de Bases de Dados
 - ⇒ - Abordagem e Vantagens
 - ⇒ - Arquitetura de um SGBD
 - ⇒ - Tipos de Utilizadores num SGDB
- ➔ 3. Modelação e Normalização de Dados
 - ⇒ - Manutenção da Integridade
 - ⇒ - Redundância e Chaves
 - ⇒ - Diagramas E-R
 - ⇒ - Modelo Relacional
- ➔ 4. Álgebra Relacional e SQL - (MySQL)
 - ⇒ - Conceitos e aplicação de Álgebra Relacional
 - ⇒ - Ferramentas de Administração MySQL
 - ⇒ - Comandos DDL
 - ⇒ - Comandos DML
- ➔ 5. Introdução às Bases de Dados NoSQL
 - ⇒ Bases de Dados NoSQL,
 - ⇒ Tipos de Bases de Dados NoSQL,
 - ⇒ Bases de dados orientadas a documentos
- ➔ 6. MongoDB
 - ⇒ Estruturas JSON e BSON,
 - ⇒ Modelação de dados,
 - ⇒ Criação de coleções e documentos,
 - ⇒ Operações CRUD e agregação,
 - ⇒ Indexação e transações



Conceitos de Bases de Dados

- Base de Dados:
 - ➔ Conjunto estruturado de informação, organizada em Campos, Registos e Tabelas, que se inter-relacionam por intermédio de chaves e que permitem o acesso a diferentes perfis de utilizadores.
- Dados:
 - ➔ São elementos de Informação que permitem descrever as Entidades ou Eventos.
- Evento:
 - ➔ Acontecimento ou ocorrência relativa a uma entidade do qual se pretende guardar Dados e Informação.
- Informação:
 - ➔ Conjunto de dados que descrevem entidades ou eventos, organizados de uma forma significativa e relevante para um determinado contexto.
- Campo:
 - ➔ Componente elementar de um registo onde se armazenam os Dados que caracterizam as Entidades.
- Tabela:
 - ➔ Conjunto organizado de registos que armazenam a informação relacionada com as Entidades do sistema, segundo uma perspetiva matricial.
- Registo:
 - ➔ Corresponde a uma entrada individual na base de dados onde se armazena informação relacionada com as Entidades ou Eventos.



Conceitos de Bases de Dados

Exemplo

Tabela cliente

Num_Cliente	Localidade	E_mail
1000	Bragança	joao@mail.pt
1100	Lisboa	joana@mail.pt
1200	Bragança	maria@mail.pt

Tabela encomenda

Num_Cliente	Num_Encomenda	Data_Encomenda
1000	123	01/03/2022
1200	150	01/03/2022
1200	245	12/03/2022

Características das Bases de Dados

→ Nas Bases de Dados, a Informação pode ser visualizada sob a forma de Tabelas, com uma linha para cada Registo ou Entrada e uma série de colunas para cada Campo ou Propriedade, o que origina que a inserção e manipulação dos dados seja feita de forma:

- ⇒ Organizada...
- ⇒ Consistente...
- ⇒ Integra...
- ⇒ Não Redundante...
- ⇒ Segura...

Redundância:

→ Duplicação desnecessária dos dados com inconvenientes:

- ⇒ Maior custo com armazenamento dos dados;
- ⇒ Operações de atualização desnecessárias;
- ⇒ Aumento custos com Pessoal;
- ⇒ Maior possibilidade de Inconsistência dos dados.

Inconsistência:

→ A existência de redundância, pode levar a que determinados processos de atualização de dados, não atualizem os campos redundantes, penalizando-se assim a Integridade da Base de dados.

Integridade:

→ Existência de dados com conteúdos únicos.

Segurança:

→ A atribuição de determinadas permissões de acordo com os perfis de utilização, limita os acessos e processamentos a efetuar à base de dados, garantindo-se assim um maior nível de segurança e proteção dos dados.

Cliente

Num_Cliente	Nome_Cliente	Localidade	E_mail
1000	Joana Ferreira	Bragança	joana@mail.pt
1001	António Ribeiro	Lisboa	ant@mail.pt
1003	Vera Gaspar	Bragança	vera@mail.pt
1004	João Ferreira	Porto	joao@mail.pt

Dados → (points to row 1001)

Registos → (bracket on the right side of the table)

Campos → (bracket on the bottom of the table)



Conceitos de Bases de Dados

● Redundância

- ➔ A redundância ocorre sempre que o mesmo facto é armazenado mais do que uma vez na mesma base de dados, ou quando factos derivados são armazenados numa base de dados.
- ➔ Os problemas que a redundância pode trazer a um sistema de bases de dados podem ser resumidos nos pontos seguintes:
 - ⇒ Problemas de manutenção:
 - Redundância significa o armazenamento repetido dos mesmos dados.
 - Como resultado, uma simples alteração ou remoção pode implicar o acesso a várias partes da base de dados, tornando-se difícil manter a coerência dos dados armazenados.
 - ⇒ Custos de espaço de armazenamento:
 - Apesar de, atualmente, não ser um problema significativo (o preço das memórias secundárias desce cada vez mais), não deixa de ser importante evitar ocupar espaço de armazenamento com algo que não acrescenta nada ao que já existe armazenado.
 - ⇒ Problemas de desempenho:
 - Apesar de não ser tão evidente, a redundância pode também trazer problemas de desempenho ao sistema.
 - A razão para isso é que a razão mais importante para um mau desempenho de um SGBD são os acessos aos discos onde as bases de dados se encontram armazenadas.
 - Se a redundância for significativa, isso implicará mais acessos aos discos para trazer os mesmos dados.
- ➔ Como exemplo de redundância, considere-se a relação Empregados e Projetos, a qual mostra os relacionamentos existentes entre projetos e os empregados que nele participam.

Código de Empregado	Número do Projeto	Orçamento	Tempo Total
E1	PROJ1	20	20
E3	PROJ1	20	16
E2	PROJ2	17	35
E2	PROJ3	84	42
E3	PROJ2	17	17
E2	PROJ1	20	83
E4	PROJ3	84	41
	PROJ4	90	

A coluna “Orçamento” do projeto é armazenado várias vezes, tantas quanto o número de pessoas (coluna Código de Empregado) participantes no projeto.

As desvantagens deste tipo de situações são várias:

- 1) Se o orçamento de um projeto muda, então teremos de alterar mais do que uma linha da tabela;
- 2) De cada vez que um pessoa é incluída num projeto, é necessário incluir também o orçamento do projeto;
- 3) Um projeto sem pessoas atribuídas, como PROJ4, só terá a coluna Orçamento preenchida; as restantes ficarão vazias. Isto significa que uma simples operação de introdução de uma pessoa num projeto será realizada de modos diferentes, de acordo com o estado atual da base de dados. Por exemplo, para inserir uma pessoa no projeto PROJ2 resume-se a inserir um nova linha na tabela Empregados e Projetos; para acrescentar uma pessoa ao projeto PROJ4, deve-se alterar o valor das colunas que estão vazias na linha correspondente a esse projeto.

Conceitos de Bases de Dados

● Manutenção da Integridade

- A manutenção da integridade de uma base de dados relacional é fundamental para o seu bom funcionamento, devendo ser assegurada pelo SGBD (Sistema de Gestão de Bases de Dados) ou pelas aplicações que interagem com os dados.
- Assim, uma base de dados relacional deve assegurar três tipos de integridade:
 - ⇒ Integridade de Domínio
 - Esta é a forma mais elementar de manutenção de integridade. Consiste em que o SGBD faça respeitar as restrições que determinados atributos impõem sobre os valores que lhes podem ser atribuídos.
 - Exemplo:
 - Numa relação que guarde informação sobre pessoas, o **atributo Sexo apenas poderá aceitar um de dois valores possíveis**: um valor que represente o sexo masculino, e outro valor que represente o sexo feminino.
 - Nesta mesma relação, o atributo Data de Nascimento apenas poderá aceitar valores que representem datas válidas. Por exemplo, a **data 30 de Fevereiro de 2005 é inválida** e, portanto, não poderá ser aceite pelo atributo Data de Nascimento.
 - ⇒ Integridade de Entidade
 - O problema da integridade de entidade coloca-se em relação à possível **existência de valores nulos** como valores de um atributo de uma relação.
 - Na terminologia das bases de dados, um valor nulo não significa um valor igual a zero; antes, significa que o valor não existe, ou que é desconhecido, ou que simplesmente não se aplica nessa ocorrência da relação.
 - O princípio da integridade de entidade estipula que nenhum atributo pertencente à **chave primária de uma relação pode conter valores nulos**.
 - Uma vez que a chave primária constitui o identificador dos tuplos da relação, a violação deste princípio equivaleria a admitir que a base de dados poderia registar informações de algo não identificável. Esta possibilidade é rejeitada pelo modelo relacional.
 - ⇒ Integridade Referencial
 - O princípio da integridade referencial estabelece que um **valor de uma chave estrangeira tem obrigatoriamente de existir como valor de um atributo da chave primária da tabela relacionada com aquela chave estrangeira**.
 - Ou seja, um valor que não exista como chave primária na sua tabela não pode ser usado como chave estrangeira de uma outra tabela.
 - Em rigor, nenhum atributo que represente datas poderá aceitar qualquer valor que não represente uma data válida.

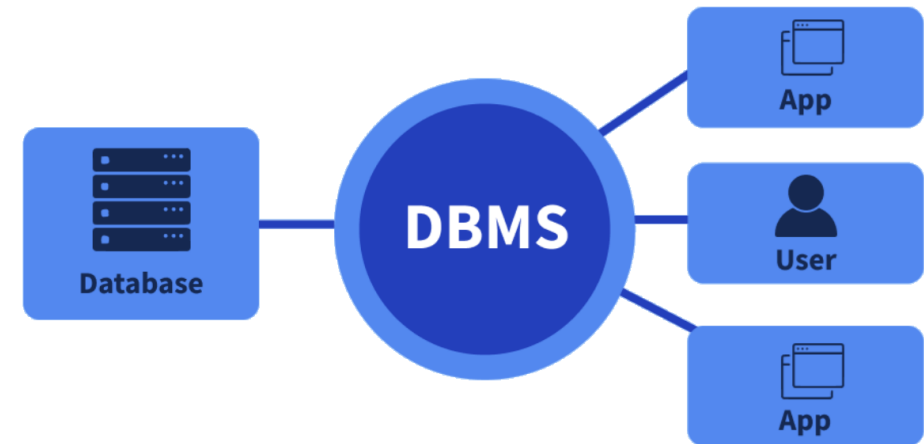
Conceitos de Bases de Dados

- Objetivos das Bases de Dados:
 - ➔ Facilitar o tratamento de grandes volumes de dados;
 - ➔ Processar informação de forma centralizada;
 - ➔ Assegurar a independência dos dados face aos processamentos;
 - ➔ Evitar a redundância no registo dos dados;
 - ➔ Evitar inconsistências ao nível dos processamentos de dados;
 - ➔ Garantir integridade dos dados;
 - ➔ Facilitar a partilha dos dados;
 - ➔ Garantir a confidencialidade e a segurança dos dados.



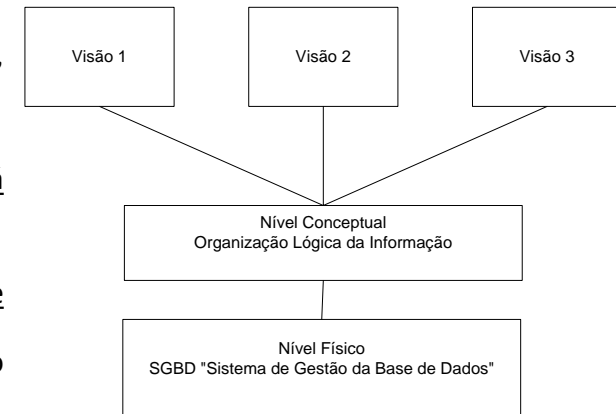
SGBD (Sistema de Gestão de Bases de Dados)

- O que é um SGBD (Sistema de Gestão de Bases de Dados)
 - ➔ O SGBD é uma porção de software que faz a ligação entre os utilizadores, ou programas de aplicação, e a base de dados propriamente dita.
 - ➔ É tarefa do SGBD conciliar os pedidos de acesso à base de dados originados pelos vários utilizadores.
 - ➔ Esse software deve proporcionar aos utilizadores um ambiente flexível, eficiente e de fácil utilização para a consulta, introdução, alteração e remoção de informação na base de dados.
- Os SGBDs são usados geralmente para:
 - ➔ criar ou gerir utilizadores
 - ➔ consultar dados
 - ➔ alterar dados e tabelas
 - ➔ excluir dados
 - ➔ relacionar tabelas
 - ➔ importar e exportar dados



SGBD (Sistema de Gestão de Bases de Dados)

- O modo como a informação está fisicamente armazenada não é relevante para o utilizador da base de dados...
 - pelo contrário, esse conhecimento apenas aumentaria a complexidade da operação.
- Nestas condições, uma função importante do SGBD é **apresentar ao utilizador uma visão abstrata da informação, escondendo os detalhes sobre como a informação está fisicamente armazenada** e, assim, simplificar a ação do utilizador.
- Normalmente, distinguem-se três níveis de abstração da informação, definidos como:
 - **Nível Físico:**
 - ⇒ É o mais baixo nível de abstração e descreve o modo como a informação está fisicamente armazenada nos dispositivos físicos apropriados (exemplo: HD, SSD, etc).
 - **Nível Conceptual (ou lógico):**
 - ⇒ Este nível é intermédio e descreve qual a informação que está armazenada na base de dados, e as várias relações internas a essa informação.
 - ⇒ Este nível é ainda relativamente complexo devido à inclusão de toda a informação sobre a organização dos dados existentes na base de dados.
 - ⇒ No entanto, o nível conceptual já não inclui a forma como a informação está fisicamente armazenada.
 - **Nível de Visualização (ou aplicação):**
 - ⇒ Também designado por nível de interação ou nível externo, é o nível com maior abstração da informação e contempla apenas uma parte da informação contida na base de dados.
 - ⇒ A informação disponível a este nível é ajustada às necessidades e competências de cada utilizador específico.
 - Em virtude da diversidade de utilizadores, existem normalmente várias visualizações para a mesma base de dados.



SGBD (Sistema de Gestão de Bases de Dados)

● Níveis de abstração da informação

→ É frequente que a estrutura de uma base de dados seja alterada ao longo do tempo para melhor responder aos requisitos que lhe são endereçados, e assim aumentar a sua capacidade e eficiência.

⇒ A possibilidade de alterar a definição de um nível de abstração de informação sem afetar o nível seguinte designa-se por **independência da informação**.

→ De acordo com os níveis de abstração considerados, existem dois níveis de independência da informação:

⇒ Independência da informação física:

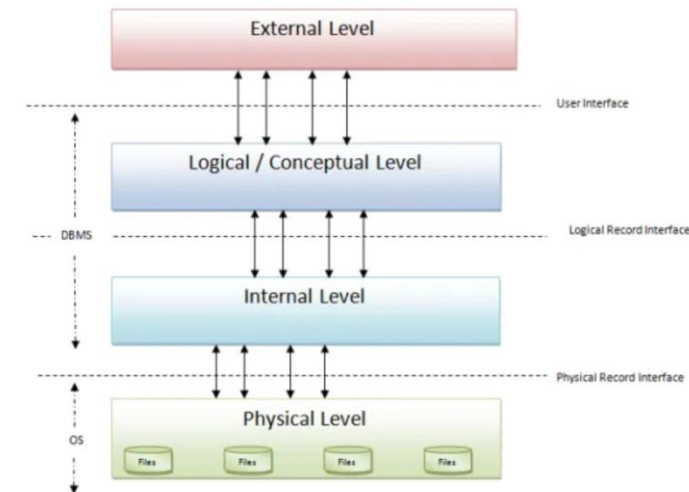
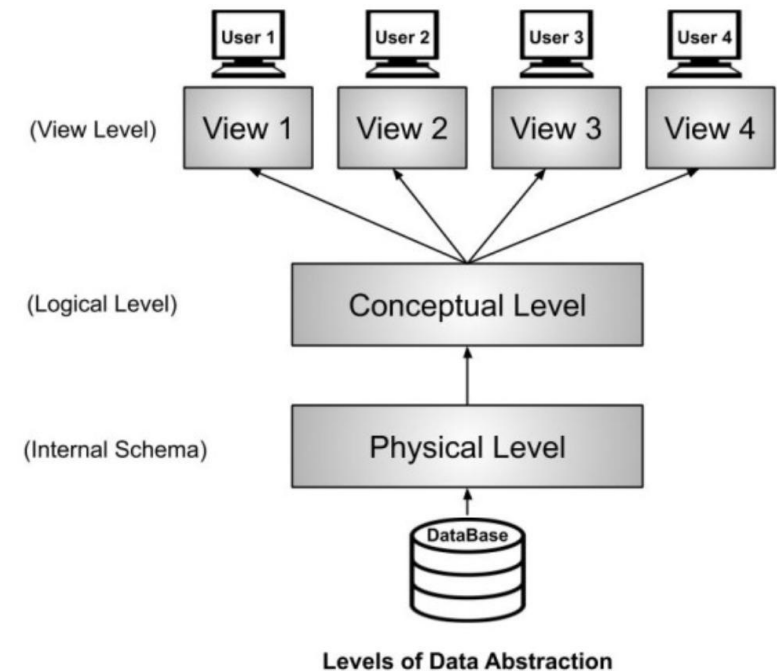
→ É a capacidade de alterar a estrutura de armazenamento físico da informação sem alterar a estrutura ao nível conceptual, e por conseguinte, sem alterar os programas de aplicação.

→ Alterações ao nível físico são esporadicamente efetuadas para aumentar a eficiência do armazenamento da informação e, consequentemente, o desempenho dos programas de aplicação que acedem a uma determinada base de dados.

⇒ Independência da informação lógica:

→ É a possibilidade de alterar a base de dados ao nível conceptual sem alterar os programas de aplicação.

→ Esta propriedade é difícil de conseguir, dado que os programas de aplicação baseiam-se na descrição da informação ao nível conceptual.



SGBD (Sistema de Gestão de Bases de Dados)

- Níveis de abstração da informação

- ➔ A interação das aplicações e dos utilizadores com o SGBD faz-se à custa de linguagens apropriadas.
- ➔ Ao nível conceptual, a estrutura de uma base de dados é especificada por um conjunto de instruções numa Linguagem de Definição de Dados (DDL) a qual deve permitir que o esquema de uma base de dados possa ser definido.
 - ⇒ O resultado dessa definição é um conjunto de tabelas armazenadas naquilo a que se chama dicionário de dados, o qual pode ser visto como um depósito central que descreve e define o significado de toda a informação usada no esquema da base de dados.
- ➔ Os SGBDs modernos também disponibilizam aos utilizadores uma Linguagem de Manipulação de Dados (DML), que permite manipular (isto é, consultar, inserir, apagar ou alterar) os dados armazenados numa base de dados.
 - ⇒ A linguagem padrão que permite definir e manipular os dados armazenados numa base de dados chama-se SQL (“Structured Query Language”, isto é, “Linguagem de Perguntas Estruturadas”).



SGBD (Sistema de Gestão de Bases de Dados)

- Os dois principais modelos de SGBD são os relacionais e os não relacionais, e a diferença básica entre eles é a estrutura utilizada para organizar os dados.
- Modelos de bases de dados:
 - ➔ 1- Relacional (SQL)
 - ⇒ São o tipo mais comum de SGBD. Expressam os dados em tabelas e relacionamentos através de chaves, e utilizam a linguagem SQL para relacionar os dados.
 - ⇒ Aplicações: Oracle SQL, Microsoft SQL Server, PostgreSQL, MySQL, MariaDB.
 - ➔ 2- Não relacional (NoSQL)
 - ⇒ São bases de dados que não usam somente SQL ou o modelo relacional. São ideais para ambientes onde o modelo de dados é muito volátil (por exemplo: eventos em um website).
 - ⇒ Aplicações: Cassandra, Redis, MongoDB, CouchDB, Apache Hbase, Riak.
 - ➔ 3- Graph
 - ⇒ É um tipo de NoSQL que representa os dados em um grafo (graph) de vértices e arestas. Ideal para dados que estão altamente relacionados entre si (por exemplo, modelo de recomendação).
 - ⇒ Aplicações: Neo4j, Azure Cosmos DB, SAP HANA, Sparksee, Oracle Spatial and Graph, OrientDB



Utilizadores de Bases de Dados

- Relativamente ao tipo de utilizadores que interagem com os SGBDs, é possível distinguir duas classes de utilizadores com interesses e posturas diferentes perante estes sistemas:
 - ➔ O administrador da base de dados e os utilizadores propriamente ditos:
 - ⇒ Administrador
 - ⇒ Utilizadores



Utilizadores de Bases de Dados

● Administrador

- ➔ O administrador de uma base de dados é o responsável máximo pelo bom funcionamento de todo o SGBD. As suas funções mais importantes são:
 - ⇒ Definição da estrutura: A estrutura da base de dados é definida por um conjunto de instruções que são transformadas pelo compilador da linguagem de definição de dados do SGBD num conjunto de definições que são armazenadas no dicionário de dados.
 - ⇒ Alteração da estrutura: Às vezes, torna-se necessário alterar a estrutura de uma base de dados já existente. Esta ação é também definida por um conjunto de instruções utilizando a linguagem de definição de dados disponibilizada pelo SGBD.
 - ⇒ Autorização de acesso aos dados: É tarefa do administrador especificar quais os dados que podem ser consultados ou alterados por cada tipo de utilizador que necessita de aceder à base de dados.
 - ⇒ Condições de integridade: Cabe ao administrador definir quais as condições que devem ser satisfeitas pela informação existente na base de dados. Por outras palavras, o administrador da base de dados deve definir todas as restrições que garantem a integridade dos dados armazenados.



Utilizadores de Bases de Dados

● Utilizadores

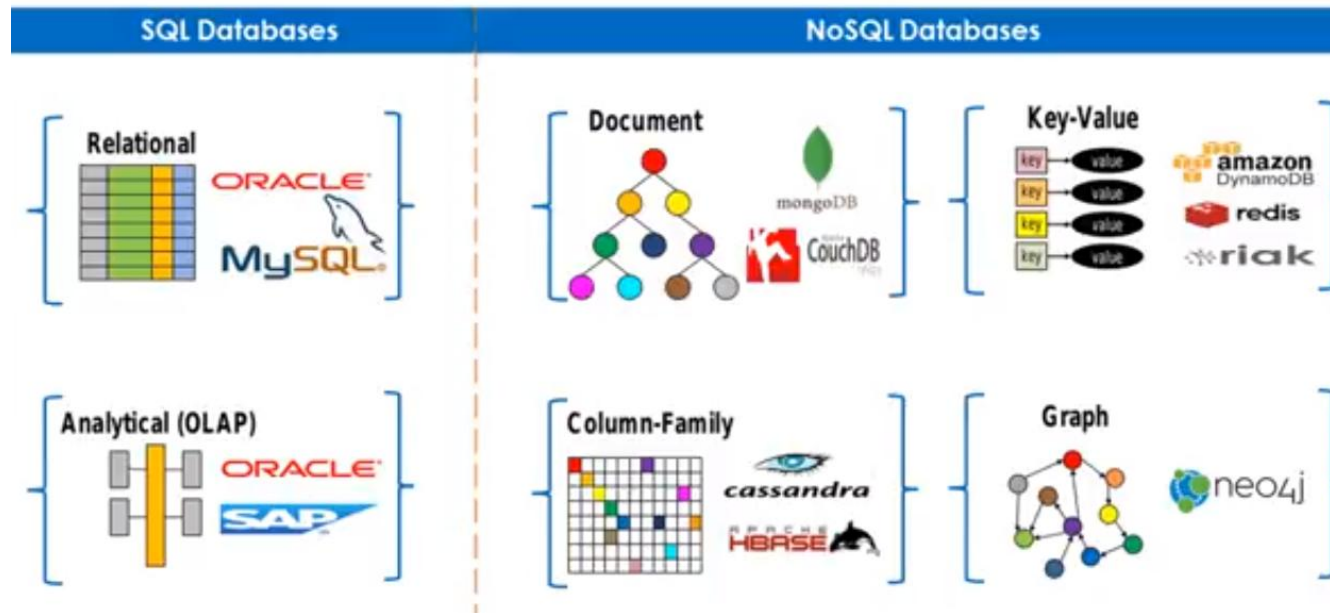
- ➔ O perfil de um utilizador de um SGBD é definido consoante o tipo de interação que este necessita de ter com o sistema.
- ➔ Basicamente, existem dois tipos de utilizadores:
 - ⇒ Programadores de aplicações:
 - São profissionais de informática que, através de linguagens de programação de alto nível (por exemplo, Pascal, C, Visual Basic, etc.) e através do esquema conceptual da base de dados, desenvolvem as aplicações informáticas que os utilizadores finais irão depois utilizar para aceder à base de dados.
 - ⇒ Utilizadores finais:
 - São a razão de existir da base de dados. Acedem ao sistema para consultar, alterar, adicionar ou remover dados, sob as condições de acesso definidas e impostas pelo administrador da base de dados. Estes utilizadores não poderão nunca alterar a estrutura (isto é, o esquema conceptual) da base de dados.



Tipos de SGBD mais usados

- SQL e NoSql

SQL and NoSQL



- SGBD relacionais mais populares



- SGBD orientados a documento mais populares

