

→ Um Sistema de Gestão de Bases de Dados é uma aplicação que interage com o utilizador, outras aplicações, e a própria base de dados para capturar e analisar informação. Uma SGBD genérica é desenhada para permitir a definição, criação, consulta, actualização e administração de bases de dados.

Vantagens:

- Controlo de redundância de dados
- Garantia de integridade de dados
- Partilha de dados entre utilizadores
- Aplicações de standards
- Satisfação rápida de queries
- Gestão e segurança integradas
- Cópias de segurança e recuperações de dados

→ Uma vista de um utilizador, ou vista de utilização, é o conjunto de dados requerido por um determinado utilizador para tomar uma decisão ou realizar uma tarefa - tipicamente é uma área funcional de uma organização.

Podemos identificar as vistas dos utilizadores durante a fase de levantamento de requisitos, através de métodos como:

- Análise de diagramas de fluxos de dados
- Realização de entrevistas aos utilizadores
- Análise e estudo de procedimentos, relatórios, formulários
- Observar o modo de funcionamento da organização

→ As restrições de integridade garantem a correção e consistência da base de dados, verificando que qualquer operação permitida não a pode deixar num estado incorreto. Os principais tipos de integridade são:

- integridade de domínio - qualquer atributo tem um domínio cujas restrições devem ser identificadas no momento da escolha dos domínios
- integridade de entidades - a chave primária de uma relação não pode admitir valores nulos; este tipo de restrição deve ser considerado no momento em que se faz a definição das chaves primárias para cada tipo de entidade
- integridade referencial - se uma chave estrangeira contém um valor, esse valor deve existir na relação por para que seja aceite na chave estrangeira.
- restrições de organização - regras da própria organização que são aplicadas para regulamentar as suas próprias actividades

→ O administrador é responsável por:

- realização física da base de dados, ou seja, pelo design e implementação da mesma
- segurança e controle de integridade
- manutenção do sistema operativo
- garantir um desempenho satisfatório das aplicações para os utilizadores.

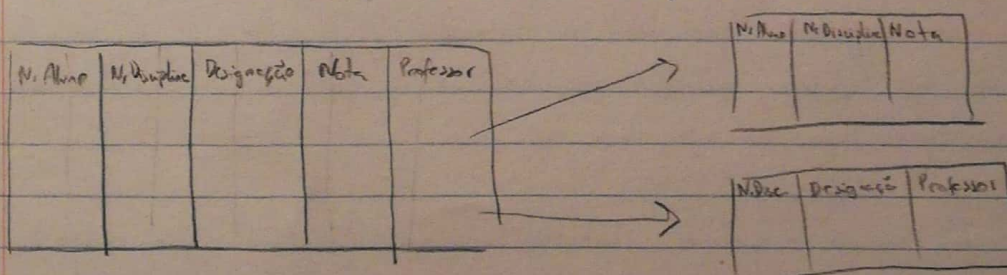
→ A normalização de bases de dados é um processo de remoção de dados redundantes das tabelas que a constituem, com o intuito de aumentar a eficiência de armazenamento, a integridade dos dados e as suas características de estabilidade.

1ª Forma Normal - se todos os valores de todos os atributos forem atômicos, isto é, se não for possível decompor-los.

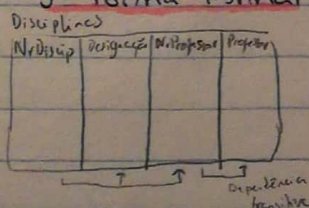
- cada entrada apenas tem um valor
- o domínio de cada atributo de uma relação apenas pode conter valores atômicos, e o valor de cada atributo é constituído apenas por um simples valor do seu domínio.

2ª Forma Normal - os atributos não-primos (que não pertencem à chave candidata) não pode ser funcionalmente dependentes de um sub-conjunto próprio de qualquer chave candidata

- se todos os seus atributos não-primos forem totalmente dependentes da sua chave primária
- se estiver em 1ª forma normal.



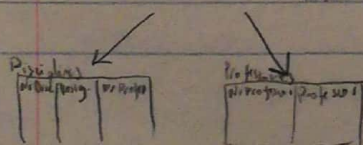
3ª Forma Normal - se estiver em 2ª Forma Normal



- nenhum atributo não primo depende de atributos não primos

- permite evitar dados que não são "desejados" numa dada tabela

- não são permitidas dependências transitivas entre os atributos de uma relação



→ "uncommitted dependency problem" - o problema da dependência de não confirmação surge quando um processo altera temporariamente uma variável (desfazendo essa alteração posteriormente) e outro processo concorrente usa essa variável com o valor alterado, submetendo um dado errado na base de dados.

→ Uma chave candidata é um atributo, ou um mínimo conjunto, de uma entidade que identifica de forma única uma instância dessa entidade. Podemos identificar mais do que uma chave candidata, mas devemos escolher apenas uma chave primária.

A chave primária não pode admitir valores nulos.

No processo de escolha de uma chave primária devemos escolher a chave candidata:

- com o menor conjunto de atributos
- com menos possibilidades de ter os seus valores modificados
- com menor probabilidade de perder a sua unicidade no futuro
- de menor tamanho
- mais "conveniente", que seja mais fácil de utilizar.

→ Uma entidade é um objeto ou um conceito que está claramente identificada numa organização como tendo uma existência independente.

Um dos métodos para identificar as entidades é através da análise da especificação de requisitos de uma dada função de um utilizador dentro da organização.

A caracterização de uma entidade deverá incluir os seguintes elementos:

- Nome da entidade
- Descrição
- Sinónimos e outras designações conhecidas.
- Ocorrência.

→ Dois tipos de intervenção que um administrador de um sistema de bases de dados poderá realizar para melhorar o desempenho de uma dada query sobre uma base de dados são a indexação e a desnormalização. Ambas diminuem eficiência espacial com vista a melhorar eficiência temporal e a escolha do método é uma análise caso por caso.

Se tivermos um atributo muito utilizado indexá-lo vai aumentar a performance das queries, por outro lado, se tivermos queries com joins que possam ser evitados através de desnormalização, esta ação será benéfica para a B.D.

→ O termo **NoSQL** significa "não só SQL". ~~As bases de dados~~ e é usado para referir bases de dados não relacionais. As bases de dados NoSQL são bases de dados sem tabelas, armazenando a informação de forma diferente das tabelas relacionais, sendo um dos tipos utilizados é o documento. Estas fornecem esquemas flexíveis e crescem mais facilmente com grandes quantidade de informações e utilizadores.

→ O **Teorema de CAP** diz que um sistema distribuído apenas consegue assegurar duas das três propriedades seguintes:

- Partição: o sistema continua operacional independentemente de falhas de "network"
- Consistência: qualquer leitura recebe a informação escrita mais recente ou um erro
- Disponibilidade: (Availability) qualquer leitura recebe informação mas sem garantia que seja a mais recente

→ Bases de dados **orientadas por documentos** são BD's NoSQL, que em vez de armazenarem informação em colunas e linhas fixas, usam documentos flexíveis. Esses documentos são organizados em coleções.

→ **Motivos** para um administrador de um sistema de bases de dados a optar por uma solução **NoSQL (MongoDB)**

- ter dados inconsistentes o que é mais fácil de suportar devido à flexibilidade dos modelos do MongoDB
- grandes quantidades de dados devido à maior facilidade em escalar a base de dados em mongo
- se for necessário guardar uma grande quantidade de dados numa só entrada, devido ao limite de 16MB que supera o limite dos servidores de SQL que normalmente permitem no máximo dados na ordem dos KB.