**Organização Física dos Dados**

**Estruturas primitivas – arquivos/índices**

Primeiras estruturas pensadas em termo de armazenamento de dados.

Registro de tamanho fixo: sempre passa a existir em uma estrutura sempre que se pensa em uma variável com um determinado tamanho fixo,

Registro de tamanho variado: presentes em estruturas que possuem uma variável com seu respectivo tamanho variado.

Um char não possui tamanho variado, ele aloca em sua memória o tamanho estabelecido, independente do quando você ocupa desse tamanho. Diferente do varchar, que utiliza da memória somente o conteúdo utilizado até o limite estabelecido.

**Tipos de arquivos**

* Arquivo sequencial:
  + registros dispostos ordenadamente
  + possuem chave de ordenação, uma espécie de identificador único, ordenada de forma ascendente
  + Por seu código de identificador ser de forma ordenada, a cada inserção de dados onde o código não fosse maior que o último, o algoritmo teria que varrer elemento por elemento até achar se lugar. Caso contrário, somente insere no final
  + Em uma alteração, o registro é encontrado a partir do seu código executando uma pesquisa binária e os dados são alterados e alocados no mesmo espaço de memória
  + Temos dois tipos de exclusão:
    - lógica: todo registro terá involuntariamente um campo de 1 byte que será oculto, onde será utilizado uma flag que servirá para indicar se tal registro está excluído ou não e essa linha não será exibida para o usuário
    - física: quando é feito a reorganização do arquivo, onde os registros com o flag excluído são removidos fisicamente
* Arquivo sequencial indexado:
  + Arquivo sequencial acrescido de um índice
  + Índice não pode existir sem um dado, mas um dado não precisa de um índice
    - Índice esparso: quando tenho uma referência para um conjunto de registros
  + Em uma pesquisa, por lógica, é encontrado o índice que faz mais sentido para o determinado código e então faz uma pesquisa sequencia a partir da linha do índice estabelecido
  + É recomendado quando temos um conteúdo mais expressivo de dados, porém, não é possível ter dois índices na mesma estrutura de armazenamento
* Arquivo indexado:
  + Arquivo acessado através de um índice
  + Índice denso: cada linha tem uma entrada no índice, assim, pode ser inserido um dado em qualquer posição no banco
  + Quanto mais índice eu tenho em uma tabela, mais tempo o banco levara para atualizar e gerenciar um índice, isso porque eles são ordenados em ordem crescente. Ou seja, quando um dado é inserido, a lista de índices é reordenada novamente
  + Inserção é vaga ou no fim com atualização do índice
  + Na exclusão liberasse a área alocada
  + E na alteração é feita a localização e modificação de um ou mais atributos
* Arquivo invertido:
  + Identificam-se os registros que agregam determinado valor – lista invertida
  + As entradas dos índices são horizontais
  + Acesso sequencia ou aleatório
  + Inserção em um endereço vago ou no fim
  + Exclusão libera a área alocada
  + Localização e modificação de um ou mais campos