

## Processo de normalização através das formas

O processo de normalização vai seguindo o caminho de “filtro”, onde a cada forma fazemos uma filtragem para melhorar a estrutura das tabelas. Então eu passo pela primeira forma, se houver condições eu passo para a segunda forma e assim sucessivamente.

Claro que uma relação pode não passar da terceira forma ou da forma boyce could, ou até mesmo chegar na quinta forma, e está tudo bem.

### Primeira forma normal.

- Cada tupla contém exatamente um valor para cada atributo
- Os atributos devem ser atômicos e indivisíveis
- Atributos compostos e multivalorados devem ser representados por novas linhas ou novas tabelas

Essas regras estão relacionadas ao fato de termos um atributo que possa ser composto por vários valores. Por exemplo, um **Departamento** pode ter várias localizações, então no meu atributo **Localização** da entidade **Departamento** eu poderia ter **UM VALOR** de atributo = **{SP, MG, RS}**.

Essa *prática não é correta* pois isso torna minha tupla **não atômica**, já que eu teria vários valores dentro de um **atributo**, o ideal aqui seria passar a **Localização** para uma outra tabela e relaciona-la com o **Departamento**.

*Aliás, essa solução acima é chamada de Decomposição de Relação (Que será feita com frequência)*

### Segunda forma normal.

A segunda forma normal é fazer todas as dependências funcionais serem TOTAIS, ou seja, não posso ter dependências funcionais PARCIAIS.

**Dependência funcional:** O conceito de **dependência funcional** é quando existe em uma relação à dependência de atributos necessários para que ela possa existir. Por exemplo, **relação X** depende dos atributos da **relação Y** (**X -> Y**).

**Dependência total e parcial:** A **dependência total** ocorre quando **se tirarmos um atributo** do lado **X** da relação então os “dados” são comprometidos **impossibilitando a relação**, por exemplo:

Como saber a quantidade de horas que um funcionário trabalhou em um projeto, precisaríamos dos seguintes atributos no lado **X** (**CPF, Numero\_Projeto**) **que resultaria em horas no lado Y -> Y(Horas)**. Se eu tirar o **CPF, ou o Numero\_Projeto** do lado **X** eu não consigo chegar nas **HORAS**.

A parcial por outro lado me permite tirar qualquer um dos atributos do lado X e ainda assim chegar no resultado esperado no lado Y. Por exemplo, se eu quisesse saber o nome do funcionário e no lado X eu tenho **X (CPF, Numero\_Projeto)** **para chegar no lado Y -> Y(Nome\_Funcionario)**, eu poderia tirar o **número do projeto** e usar apenas o **CPF** e conseguiria chegar **no nome do funcionário**, por outro lado se eu **tirasse o CPF** e usasse apenas o **numero\_projeto** **não conseguiria**. Isso é uma Relação Parcial

## Terceira forma normal.

**Dependência transitiva:** Ocorre quando um atributo além de depender da chave primária, também depende de outro atributo não chave.

A terceira forma normal diz que não devemos ter essas dependências transitivas nos modelos relacionais. Para fazer isso precisamos **decompor a relação original** em **outra que contenha as colunas** que *dependem transitivamente da outra coluna, que não faz parte da chave primária.*