

Vocabulário de Termos

O presente documento tem por objetivo auxiliar na compreensão dos exercícios propostos, principalmente para os participantes iniciantes na área de engenharia de dados. Esse documento complementar tem por finalidade reunir algumas definições dos termos utilizados nos enunciados, porém NÃO é de utilização obrigatória. Cada participante poderá complementar seus estudos da forma que preferir.

Documento construído com auxílio do Copilot (<https://copilot.microsoft.com>).

Índice

Vocabulário de Termos.....	1
Int64 ou int64	1
<NA> ou NAN	2
UOR.....	2
PK (Primary key) - Chave primária.....	2
FK (Foreign key) – Chave estrangeira.....	2

Int64 ou int64

No Pandas, **int64** e **Int64** são dois tipos de dados usados para representar inteiros, mas eles têm diferenças na implementação e no comportamento:

1. int64:

Imagine que você tem uma lista de números inteiros no Python. O `int64` é como uma versão especial dessa lista que o Pandas usa. Ele só aceita números inteiros e não permite valores vazios (nulos). Se você tentar colocar um valor vazio em uma célula, ele será preenchido com zero.

2. Int64 (com “I” maiúsculo):

Agora, o `Int64` é um pouco mais flexível. Ele também representa números inteiros, mas pode lidar com valores vazios. Isso é útil quando você tem dados incompletos. Por exemplo, se você estiver trabalhando com uma tabela de alunos e alguns não tiverem nota em uma prova, o `Int64` permite que você marque essas células como “sem nota” sem precisar mudar o tipo da coluna para números com casas decimais.

<NA> ou NAN

No Pandas, o valor **<NA>** representa um dado ausente ou nulo. É equivalente ao **NaN** (Not-a-Number) em outras linguagens e bibliotecas. Quando você encontra **<NA>** em um DataFrame, significa que a célula não possui um valor válido. Essa marcação é útil para lidar com dados faltantes ou vazios. Por exemplo, ao ler dados de um arquivo CSV, o Pandas pode representar células vazias como **<NA>**1.

UOR

UOR, ou Unidade Organizacional, no contexto do Banco do Brasil, representa as Unidades Organizacionais (UOR) do conglomerado Banco do Brasil (no Brasil e Exterior) ou UOR externas que o BB tenha algum tipo de relacionamento ou interesse em identificá-las.

Cada prefixo (ou dependência) do Banco pode ser desmembrado em várias equipes e subequipes subordinadas umas às outras, as Unidades Organizacionais.

Cada registro da tabela UOR representa uma equipe, portanto existem vários registros de UOR para um mesmo prefixo, conforme pode ser observado no [Humanograma](#).

A UOR principal representa o prefixo como um todo. Todas as outras UOR do prefixo estão vinculadas ou subordinadas a ela. Ela é representada pelo código de subordinação zero.

PK (Primary key) - Chave primária

É um identificador único para cada registro em uma tabela de banco de dados. Ela garante que cada entrada seja distinta e facilmente localizável.

Imagine que você tem uma lista de funcionários em uma empresa. Cada funcionário tem uma matrícula única que o identifica. Essa matrícula é como uma “chave primária” em um banco de dados. Ela é única para cada registro e ajuda a identificar cada funcionário de forma exclusiva.

FK (Foreign key) – Chave estrangeira

É um campo em uma tabela que cria uma ligação com a chave primária de outra tabela. Isso permite relacionar dados entre diferentes tabelas, mantendo a integridade e a consistência das informações.

Agora, imagine que você tem uma lista de unidades ou agências da empresa, e cada unidade pode ter um ou mais funcionários alocados. Para ligar uma unidade a um funcionário específico, você usa a matrícula do funcionário. Essa matrícula na agência é como uma “chave estrangeira”. Ela cria uma ligação entre dois conjuntos de dados diferentes, permitindo que você relacione informações de uma tabela (Unidades Organizacionais) com outra (Funcionários).