

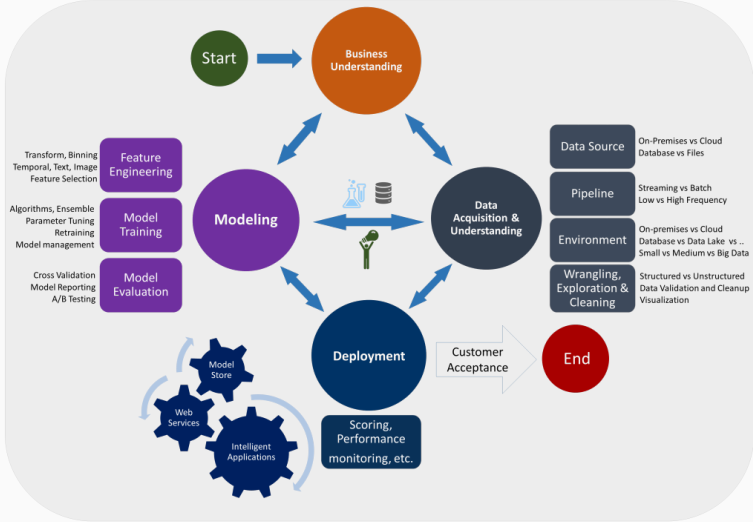
Bancos de Dados

Prof. Leandro Correia

Novembro, 2020

Importância dos Bancos de Dados para a Ciência de Dados

Data Science Lifecycle



Importância dos Bancos de Dados para o Cientista de Dados

Flexibilidade de atuação:

- Os dados estão muito além das planilhas
- Possibilidade de atuação em cenários de Big Data

Capacidade crítica:

- Conhecer o ciclo de vida da informação
- Analisar o processo de aquisição de dados
- Participar ativamente da aquisição e preparação de dados

Sistemas de Bancos de Dados

Desafios da Administração de Dados

- Redundância (Inconsistência)
- Desafios de acesso aos dados
- Necessidade de integridade
- Necessidade de atomicidade
- Anomalias de concorrência
- Restrições de segurança

Situação Problema: Organizar o churrasco via Zap

"Pessoal, estou organizando um churrasco para comemorar nossos 20 anos de formados. O modelo será cada um leva um item do churrasco. Nós somos 10 pessoas, logo, gostaria que vocês assumissem um (somente um) dos itens a seguir."

- Maminha
- Picanha
- Coração de Galinha
- Frango
- Calabresa
- Vinagrete
- Farofa
- Cerveja
- Refrigerante
- Água

Situação Problema: Organizar o churrasco via Zap

Potenciais problemas:

- Sobreposição da lista de itens (problema de concorrência)
- Indicação de um item que já foi indicado por outra pessoa (problema de integridade)
- Indicação de um item que não esteja na lista (problema de integridade)

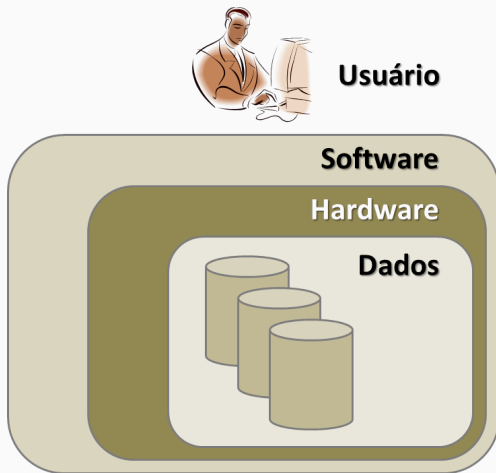
Situação Problema: Organizar o churrasco via Zap

Depois de migrar a lista para uma planilha:

- O grupo não consegue mais visualizar a situação atual da lista (problema de concorrência)
- O dono da planilha demora de responder as perguntas do grupo (problema de performance)
- O dono da planilha perde a planilha (problema de segurança)

Sistemas Gerenciadores de Bancos de Dados (SGBDs)

Elementos de um SGBD



Elementos de um Sistema de Bancos de Dados

Dados:

- Qualquer informação que faça sentido para um indivíduo ou uma organização.
- Devem ser integrados e compartilhados.

Hardware:

- Volumes de armazenamento (discos, controladoras, dispositivos de E/S) responsáveis pelo armazenamento dos dados.
- CPU e Memória, para permitir a execução do software SGBD.
- Dispositivos de comunicação, para permitir que os usuários tenham acesso aos dados.

Elementos de um Sistema de Bancos de Dados

Software:

- Conjunto de programas que formam o sistema gerenciador de banco de dados.
- Responsável pelo gerenciamento e armazenamento de informações e pelo atendimento das solicitações dos usuários para acesso aos dados.
- Permite uma visão individual para cada usuário, isolando-os dos detalhes de implementação internos ao SGBD.

Usuários:

- Usuários finais: acessam os dados através de uma aplicação ou por uma interface provida pelo SGBD.
- Programadores de aplicação: responsáveis por escrever aplicações que acessam dados armazenados em bancos de dados.
- Administradores de Bancos de Dados (DBAs): responsáveis pela instalação, configuração e administração do SGBD.

Benefícios oferecidos pelos SGBDs

- Controle da redundância
- Facilidade de acesso aos dados
- Implementação de garantias de integridade
- Viabilização da atomicidade (através das transações)
- Acesso compartilhado (controle de concorrência)
- Mecanismos da segurança