

FIAP GRADUAÇÃO

DATA SCIENCE: BIG DATA, BI & DATA ENGINEERING

BUILDING RELATIONAL DATABASE

PROF. TADEU KANASHIRO proftadeu.kanashiro@fiap.com.br
PROFa. RITA DE CÁSSIA rita@fiap.com.br

PROFESSOR: PERÍODO NOTURNO



- **EXPERIÊNCIA PROFISSIONAL:** Mais de 15 anos atuando na área de tecnologia, focado em análise e estruturação de dados;
- **ÁREAS DE ATUAÇÃO:** Big Data, Analytics Engineering, Business Intelligence e Database Marketing;
- **INDÚSTRIAS:** Saúde, setor imobiliário farmacêutica, fintech, financeiro, Internet, telecomunicações, educação e filantropia.
- **MBA:** Big Data (Data Science);
- **GRADUAÇÃO:** Sistemas de Informações e Gestão Financeira.

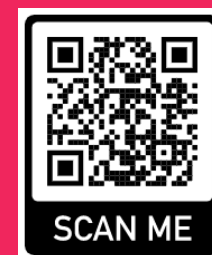
Professor: Tadeu Kanashiro



proftadeu.kanashiro@fiap.com.br



<https://www.linkedin.com/in/tadeukanashiro/>



■ AGENDA

- Desafio da Aula Anterior
- Banco de Dados
- Tipos de Ambientes de Sistemas com BD
- Características de um SGBD Relacional

DESAFIO ANTERIOR

	A	B	C	D
1	CODIGO	NOME	E_MAIL	DATA_CADASTRO
2	1010	Novak Djokovic	NovakDjokovic@email.com	2017-08-06
3	1011	Rafael Nadal	RafaelNadal@email.com	2012-02-25
4	1012	Daniil Medvedev	DaniilMedvedev@email.com	2018-02-14
5	1013	Dominic Thiem	DominicThiem@email.com	2019-10-20
6	1014	Stefanos Tsitsipas	StefanosTsitsipas@email.com	2019-12-11
7	1015	Roger Federer	RogerFederer@email.com	2020-07-06
8	1016	Alexander Zverev	AlexanderZverev@email.com	2017-04-01
9	1017	Andrey Rublev	AndreyRublev@email.com	2020-10-29
10	1018	Diego Schwartzman	DiegoSchwartzman@email.com	2015-04-30
11	1019	Matteo Berrettini	MatteoBerrettini@email.com	2014-09-15

	A	B	C
1	COD_CLIENTE	COD_TELEFONE	NUMERO
2	1010	1	+1-202-555-0167
3	1011	2	+1-202-555-0158
4	1012	3	+1-202-555-0111
5	1013	4	+1-202-555-0190
6	1014	5	+1-202-555-0161
7	1015	6	+1-202-555-0172
8	1016	7	+1-202-555-0190
9	1017	8	+1-202-555-0321
10	1018	9	+1-202-555-0432
11	1019	10	+1-202-555-0897
12	1010	11	+1-202-555-3472

BANCO DE DADOS

I BANCO DE DADOS

- Definição: **Coleção de dados armazenados** e **organizados** de modo a atender as necessidades integradas de seus usuários. **Possibilita** a **consulta** e a **manipulação de dados**, podendo ser manual ou computadorizado;
- Banco de Dados **Manual**: **Prontuário médico** em fichas de **papel**.
- Banco de Dados **Computadorizado**: Envolve o armazenamento de dados em **estruturas organizadas** e implementadas **em um software**. Esses softwares são conhecidos como SGBD.

I IMPORTÂNCIA DO BANCO DE DADOS

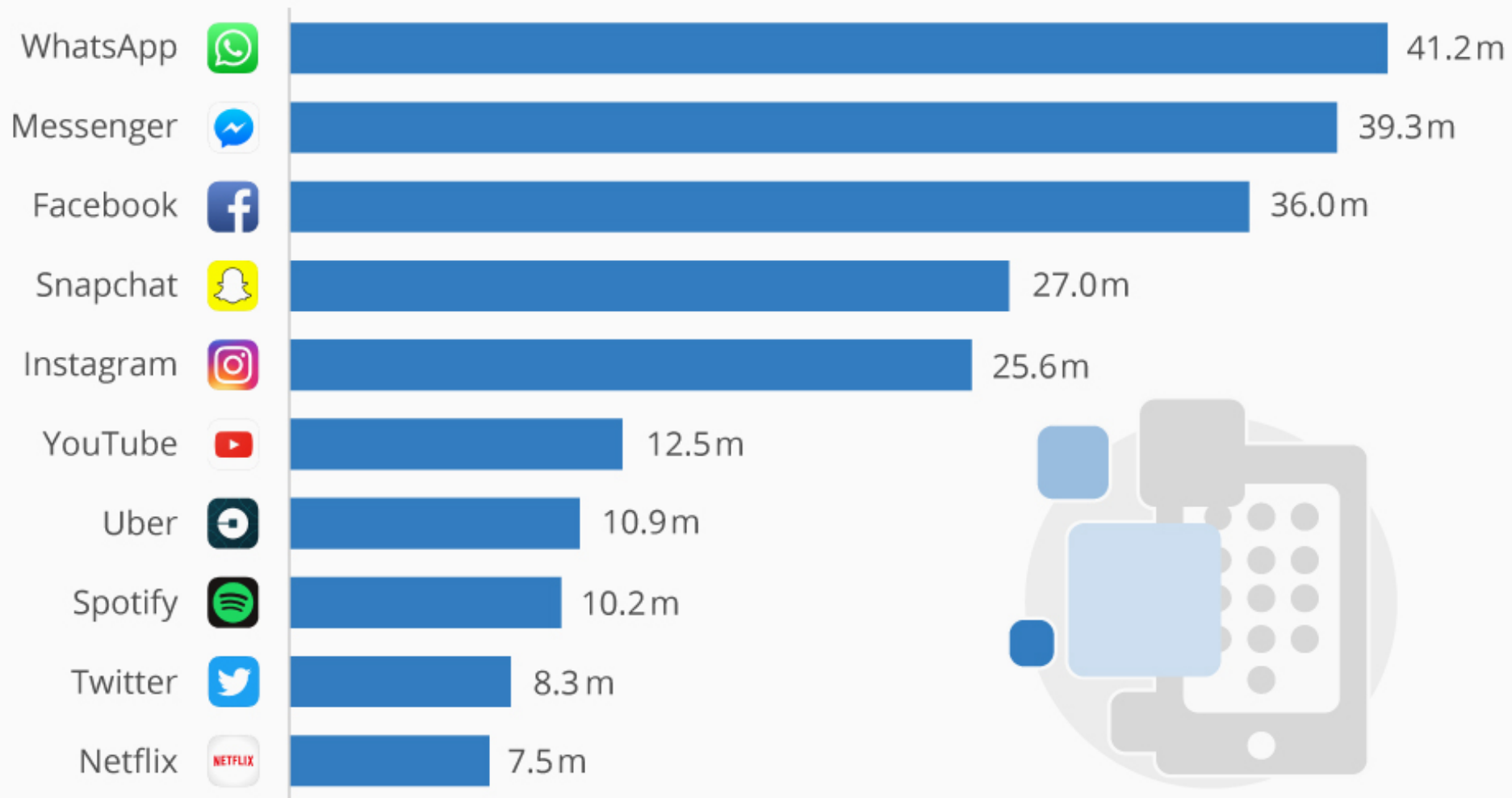
- Cada vez mais, a **informação** é considerada um dos **ativos estratégicos** mais importantes de uma empresa;
- Sendo assim, **um dos recursos** mais **importantes** para qualquer empresa **é a sua coleção de dados**;
- Um banco de dados pode ajudar na:
 - **Organização** de uma quantidade crescente de informações;
 - **Geração de informações** que contribuam aos negócios da empresa.

I BANCO DE DADOS X SGBD

- A **diferença entre** um **BD** e o **SGBD** está em possuir ou não uma camada de **software** para **gerenciar** os **dados**;
- **Importante:** Embora Microsoft **SQL Server** e **Oracle** , por exemplo, sejam **SGBDs**, é comum, no dia a dia de trabalho, as pessoas/ profissionais se referirem a eles como Banco de Dados. **Conceitualmente** dizer isto está errado, ambos os softwares são SGBDs.

The Most Popular Apps in the World

Estimated number of global app downloads in May 2016

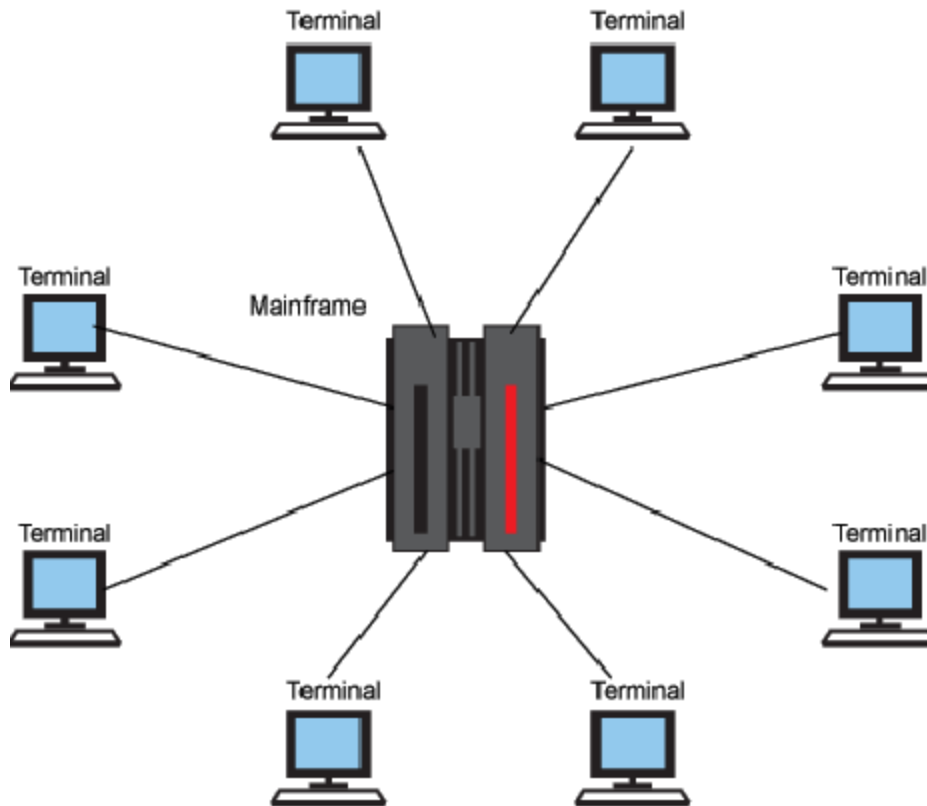


TIPO DE AMBIENTES DE SISTEMAS COM BD

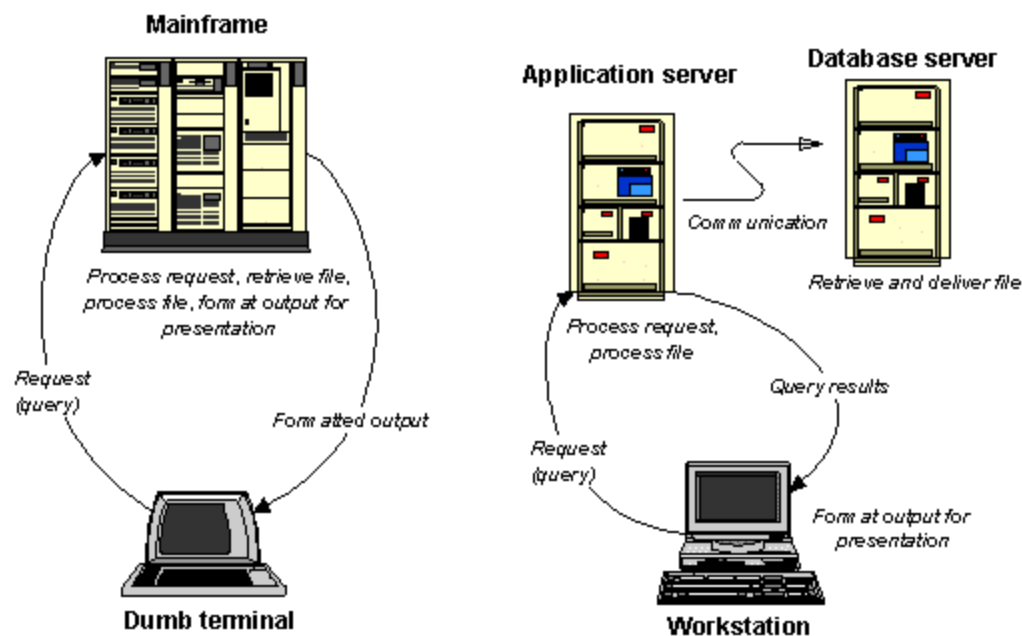
TIPOS DE AMBIENTES

- **MAINFRAME** (1960)
 - Vantagens: alto poder de processamento, segurança e disponibilidade
 - Desvantagens: alto custo, pouca flexibilidade e interface não amigável
- **CLIENTE/ SERVIDOR** (1980)
 - Vantagens: melhor organização
 - Desvantagens: investimentos em mais de um servidor e licenças, mais ambientes para manutenções
- **CLOUD COMPUTING** (COMPUTAÇÃO EM NUVEM) (2012)
 - Vantagens: processamento escalável e custo flexível (pague pelo uso)
 - Desvantagens: necessidade de conexão com internet e leis podem variar de acordo com o país

MAINFRAME



CLIENTE/ SERVIDOR



■ CLOUD COMPUTING



Google Cloud Platform



“Cloud Computing: entrega sob demanda de poder computacional, banco de dados, armazenamento, aplicativos e outros recursos de tecnologia pela internet sob definição de preço conforme o uso”



CARACTERÍSTICAS DE UM SGBD RELACIONAL

I AS 12 REGRAS DE “Codd”



- O Dr. Edgard Frank Codd (1923-2003), “Ted **Codd**”, pesquisador da IBM, criador da teoria de banco de dados relacional (1970), **propôs** um conjunto de **12 regras** (1985), às quais um **banco** de **dados relacional** “ideal” deveria atender;
- Na **prática** de **mercado**, mesmo os mais avançados SGBDRs **não** **chegam** a estar 100% em conformidade com **todas** as 12 **regras**.

AS 12 REGRAS DE “CODD”

1. **Dados em tabelas:** Dados em “relações” bidimensionais, ou seja, organizados em **linhas** e **colunas**.
Devem estar em conformidade com o Modelo Relacional, incluindo o dicionário de dados do BD;
2. **Garantia de acesso:** Todos os **dados** devem ser **acessíveis**. **Através** da **referência** de nome da tabela + valor da chave primária + nomes das colunas;
3. **Tratamento de valores nulos:** “nulos” (**ausência** ou indefinição de **valor**) **devem ser suportados**, independente dos domínios (tipos de dados) definidos para os atributos;
4. **Dicionário de dados:** O “**catálogo**”, ou **metadados** (dados sobre os dados), do BD **deve** também **estar disponível** em **tabelas**;
5. **Inserção, atualização e exclusão em alto nível:** possibilidade de **trabalhar** as **operações** em **conjunto**.

AS 12 REGRAS DE “CODD”

6. **Linguagem** de **tratamento** de **dados** que permita:
 - **Definição** de **estruturas** e **manipulação** de **dados**, com possibilidade de utilização em programas aplicativos;
 - Definição de visões;
 - Restrições de integridade;
 - Definições de acesso;
 - Controle de transações.
7. **Independência física: alterações** na forma de armazenamento, ou métodos de acesso, físico dos dados **não devem implicar** mudanças nas **aplicações**;
8. **Independência lógica: alterações** nas estruturas do BD (**tabelas**, **campos**), que **não envolvam remoção** de **elementos**, não devem implicar mudanças nas aplicações.

I AS 12 REGRAS DE “CODD”

9. **Capacidade** de **atualizar visões**: de forma **transparente** nas respectivas **tabelas**;
10. **Integridade** de **dados**: regras de integridade definidas através de uma linguagem suportada e armazenadas no dicionário de dados. **Garantir integridade independente** dos programas de **aplicação**;
11. **Independência** de **distribuição**: **arquiteturas** de rede, **plataformas** e distribuição ou centralização do banco de dados **não devem afetar** a **manipulação** de **dados**;
12. **Segurança** e **controle** de **acesso**: o sistema não deverá permitir ultrapassar as regras de integridade.

OBJETOS DE UM SGBDR

OBJETOS	DESCRIÇÃO
TABELAS	São os objetos que contém os tipos de dados e os dados reais.
COLUNAS OU CAMPOS	São as partes das tabelas que armazenam os dados. Devem receber um tipo de dados e ter um nome único.
TIPO DE DADOS	Há vários tipos de dados para serem utilizados como: caractere , número , data . Um único tipo de dados é atribuído a uma coluna dentro de uma tabela.
STORED PROCEDURES (PROCEDIMENTOS ARMAZENADOS)	São como macros em que o código SQL pode ser escrito e armazenado sob um nome.
TRIGGERS (GATILHOS)	São como storeds procedures que são automaticamente ativados quando os dados são inseridos, alterados ou apagados. Asseguram que regras de negócio e de integridade sejam impostas ao banco de dados.
REGRAS (RULES)	São atribuídas a colunas de modo que os dados que estão sendo inseridos devem se adaptar aos padrões definidos. Por exemplo, pode-se utilizar regras para permitir que um campo que irá armazenar a UF contenha somente Estados válidos.

OBJETOS DE UM SGBDR

OBJETOS	DESCRIÇÃO
CHAVES PRIMÁRIAS (PK)	Embora não sejam objetos em si, as chaves são essenciais para os bancos de dados relacionais. Promove a característica de unicidade das linhas , proporcionando uma maneira de identificar de forma única cada item que você queira armazenar.
CHAVES ESTRANGEIRAS (FK)	Novamente, não são objetos em si, as chaves estrangeiras são colunas que fazem referências as chaves primárias de outras tabelas .
PADRÕES (DEFAULTS)	Podem ser configurados em campos de modo que, se nenhum dado for inserido durante uma operação de Insert, os valores padrão serão utilizados.
VIEWS (VISUALIZAÇÕES)	Consistem basicamente em consultas armazenadas nos bancos de dados que podem fazer referência a uma ou muitas tabelas. Você pode criar e salvar views e utiliza-las no futuro. Normalmente excluem certas colunas de uma tabela e vinculam duas ou mais tabelas entre si. Podem ser utilizadas também como mecanismo de segurança.
ÍNDICES	Podem ajudar os dados de modo que as consultas executem mais rápido .

■ DESAFIO

- Continuando o exercício da aula anterior:
 - Na(s) tabela(s) que você criou, há um campo que pode ser eleito como chave primária? Se não, como você o criaria?
 - Classificar todos os campos em: texto, numérico e data.
 - Há algum campo que você classificaria como índice? Por que?

Copyright © 2024 Prof. Tadeu Kanashiro e Prof. André Santos

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).