



# Základní pojmy světa databází pro datové analytiky

Jan Kammermayer

[www.linkedin.com/in/kammermayerj](https://www.linkedin.com/in/kammermayerj)

# Relační databáze

- Místo pro uložení dat do tabulek.
  - Sofistikovanější než například pouhý textový soubor.
  - Mohou obsahovat miliony/miliardy záznamů.
- 
- Často velice robustní a využívané ve velkých aplikacích.
  - Relační: tvoříme vztahy mezi záznamy.
  - NoSQL = Not only SQL - nejčastěji key-value nebo dokumentová databáze.

# Relační databázové systémy (RDBMS)

- “Program”, se kterým komunikujeme a jsou v něm uloženy databáze.
- Máme mnoho variant: PostgreSQL, MySQL, MSSQL, Oracle, ...
- V různých nastaveních a možnostech rozšíření se liší, ale v základu s nimi můžeme vždy pracovat stejně. (přenositelnost)

# Schéma databáze

- Každá relační databáze má své schéma.
- Vytváříme jej během návrhové fáze.
- Nejedná se o samotná data, ale o jejich strukturu.
- Pomáhá nám logicky seskupovat data a zbavovat se redundancí.

# SQL

- Structured Query Language
  - Je to strukturovaný jazyk pro dotazování.
- Jeho prostřednictvím komunikujeme s databázemi.
- Základy stejné pro různé RDBMS.
- Dělí se na více podskupin.

# Druhy SQL

- DDL – Data Definition Language
  - Definice struktury - CREATE, ALTER, DROP, ...
- DML – Data Manipulation Language
  - Manipulace s daty - INSERT, UPDATE, DELETE
  - Podskupina: **DQL – Data Query Language (SELECT)**
- DCL – Data Control Language
  - Přidělování oprávnění - GRANT, REVOKE
- DTL - Data Transaction Language
  - Správa transakcí - START TRANSACTION, COMMIT, ROLLBACK

# Integritní omezení

- Omezení, která nám pomáhají mít data konzistentní.
- Definujeme na úrovni tabulek, sloupců i vazeb.
- Databáze nás poté hlídá během operací.
- Ideální na ně myslet už při návrhu schématu.

# Datové typy

- Ve své podstatě taky integritní omezení.
- Každý sloupec tabulky má svůj datový typ a právě takový do něj musíme ukládat.
- Příklady datových typů:
  - Celé číslo – integer, bigint
  - Řetězec – char, varchar X nchar, nvarchar - s podporou unicode
  - Logická hodnota – boolean
  - Čas – date, time, datetime
  - [dokumentace](#)