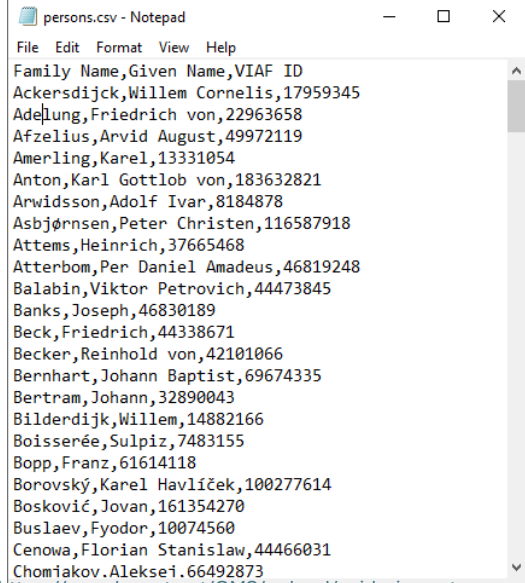


Databázové systémy

Úvod do databázových systémů



https://e.nodegoat.net/CMS/upload/guide-import_person_csv_notepad.png

```
<?xml version="1.0"?>
- <Dataset xmlns="http://www.safe.com">
  - <Building id="Surrey Head Office">
    <Address>"7445 132 St."</Address>
    <City>Surrey</City>
    <Province>BC</Province>
    <Country>Canada</Country>
  - <Location>
    <Longitude>-122.860</Longitude>
    <Latitude>49.138</Latitude>
  </Location>
  <Reference>https://www.google.ca/maps/3m1!4b1!4m5!3m4!1s0x5485dbd520cc122.8574636?hl=en</Reference>
  - <Room id="Admin_100">
    <Name>Reception</Name>
    <Category>Admin</Category>
    <Area units="m2">12</Area>
  - <Wall id="Admin_100_WallA">
    <material>wood</material>
  </Wall>
</Dataset>
```

https://support.safe.com/hc/article_attachments/25408415566093

```
{
  "type": "message",
  "channel": "C2147483705",
  "user": "U2147483697",
  "text": "Hello world",
  "ts": "1355517523.000005",
  "is_starred": true,
  "pinned_to": ["C024BE7LT", ...],
  "reactions": [
    {
      "name": "astonished",
      "count": 3,
      "users": [ "U1", "U2", "U3" ]
    },
    {
      "name": "facepalm",
      "count": 1034,
      "users": [ "U1", "U2", "U3", "U4", "U5" ]
    }
  ]
}
```

- Práca s databázovými systémami?
- Načo mi to bude?
- Nepostačuje mi textový súbor, CSV, XML, JSON alebo excel?

<https://blog.pagefreezer.com/hs-fs/hubfs/Screenshot%20Shot%202021-12-08%20at%2019.43%20AM.png?width=558&name=Screenshot%20Shot%202021-12-08%20at%2019.43%20AM.png>

"txt" typically refers to "plain text files".

What is a TXT file?

- **File Extension:** ".txt"

- **Definition:** A TXT file is a standard text document that contains unformatted plain text. It does not include any special fonts, colors, images, or other formatting elements.

- **Content:** Just simple characters, letters, numbers, and symbols.

- **Usage:** Commonly used for notes, readme files, logs, scripts, or any information that doesn't require formatting.

- **Compatibility:** TXT files can be opened and edited by almost any text editor or word processor, such as Notepad (Windows), TextEdit (Mac), or more advanced editors like Visual Studio Code or Sublime Text.

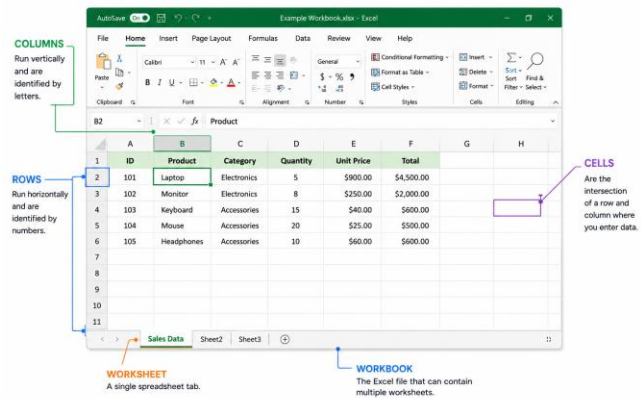
Key Characteristics

Feature	Description
Format	Plain text (ASCII or Unicode)
File Size	Usually very small
Editable By	Any text editor
Formatting	None (no bold, italic, fonts, etc.)
Common Use Cases	Notes, logs, scripts, configuration files

Summary

A "TXT file" is the simplest type of text file, containing only plain, unformatted text. It's widely used because of its simplicity and universal compatibility across all devices and operating systems.

<https://img2-2-f.schubassers.com/img/document/66178771/original/6dafbd1f13/1?v=1>



<https://baserow-backend-production2024052812452433900000001.s3.amazonaws.com/pagedown-uploads/cdd2fc5f-7a94-40aa-aa31-6962e26de55d/what%20is%20a%20spreadsheet.png>

Excel vs. Databázové systémy

Vlastnosť	Microsoft Excel	Databázové systémy (DBMS)
Primárny účel	Analýza dát, grafy a rýchle výpočty	Bezpečné ukladanie a správa veľkých dát
Maximálna kapacita	Obmedzená (~1 milión riadkov na hárok)	Takmer neobmedzená (milióny až miliardy)
Práca viacerých ľudí	Komplikovaná, hrozia konflikty verzií	Plne podporovaná pre tisíce používateľov naraz
Integrita a čistota dát	Nízka (bunka znesie akýkoľvek text či formát)	Vysoká (striktné pravidlá pre typy dát a väzby)
Prepojenie dát (Vzťahy)	Manuálne a náchylné na chyby (XLOOKUP)	Automatické pomocou relačných väzieb (SQL)
Bezpečnosť a práva	Slabá (zamknutie súboru alebo hárku heslom)	Pokročilá (šifrovanie, detailné prístupové práva)
Rýchlosť pri veľkých dátach	Výrazne spomaľuje a zamrzá	Optimalizovaná a stabilne rýchla (indexovanie)

Kedy použiť Excel?

- Na jednorazové výpočty, grafy a finančné modelovanie.
- Pre menšie množstvá dát (do niekoľko tisíc riadkov).
- Na rýchlu osobnú analýzu, kde netreba zdieľaný prístup tímu.

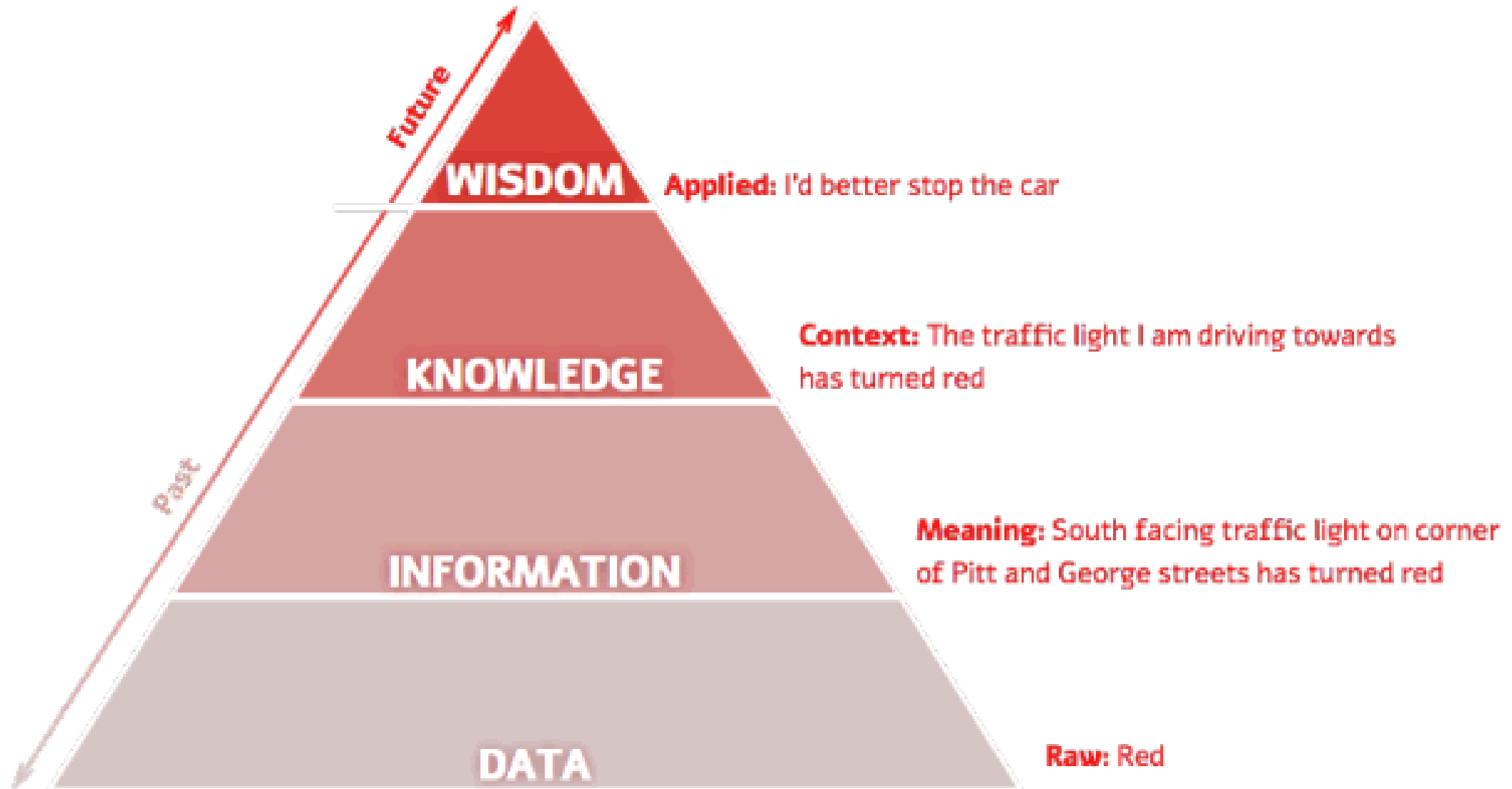
Kedy použiť databázové systémy

- Keď dáta neustále rastú a využívajú ich celé tímy.
- Ak potrebujete prísne prístupové práva, auditné záznamy a vysokú bezpečnosť.
- Pri prepojení s webovými aplikáciami alebo firemnými systémami cez API

- Dáta
- Databáza
- Systém Riadenia Bazy Dát
- Databázový Systém
- ...

Ako sa v tých pojmoch vyznať?

DIKW pyramid



- **Dáta**/údaje sú správy, ktoré vyjadrujú určité fakty o procesoch, alebo prvkoch reálneho sveta.
 - Údajmi môžu byť písmená, čísla, slová, znaky, text, číslo, obrázok, zvuk, video,
- **Informácia** je zaznamenaný a overený údaj (správa), ktorý v rámci kontextu
 - vedie k zvýšeniu znalosti a zníženiu neistoty a
 - ovplyvňuje rozhodovanie, správanie sa.
- **Báza dát, databáza (DB)** je súhrn dát (údajov), ktoré podľa možnosti sú príbuzné a štruktúrované (papierové kartotéky, clay tablets summer, banka dát,...).
- **Počítačová (digitálna) databáza** je štruktúrovaná kolekcia dát a metadát, ktoré sú uložené v počítačovom systéme.

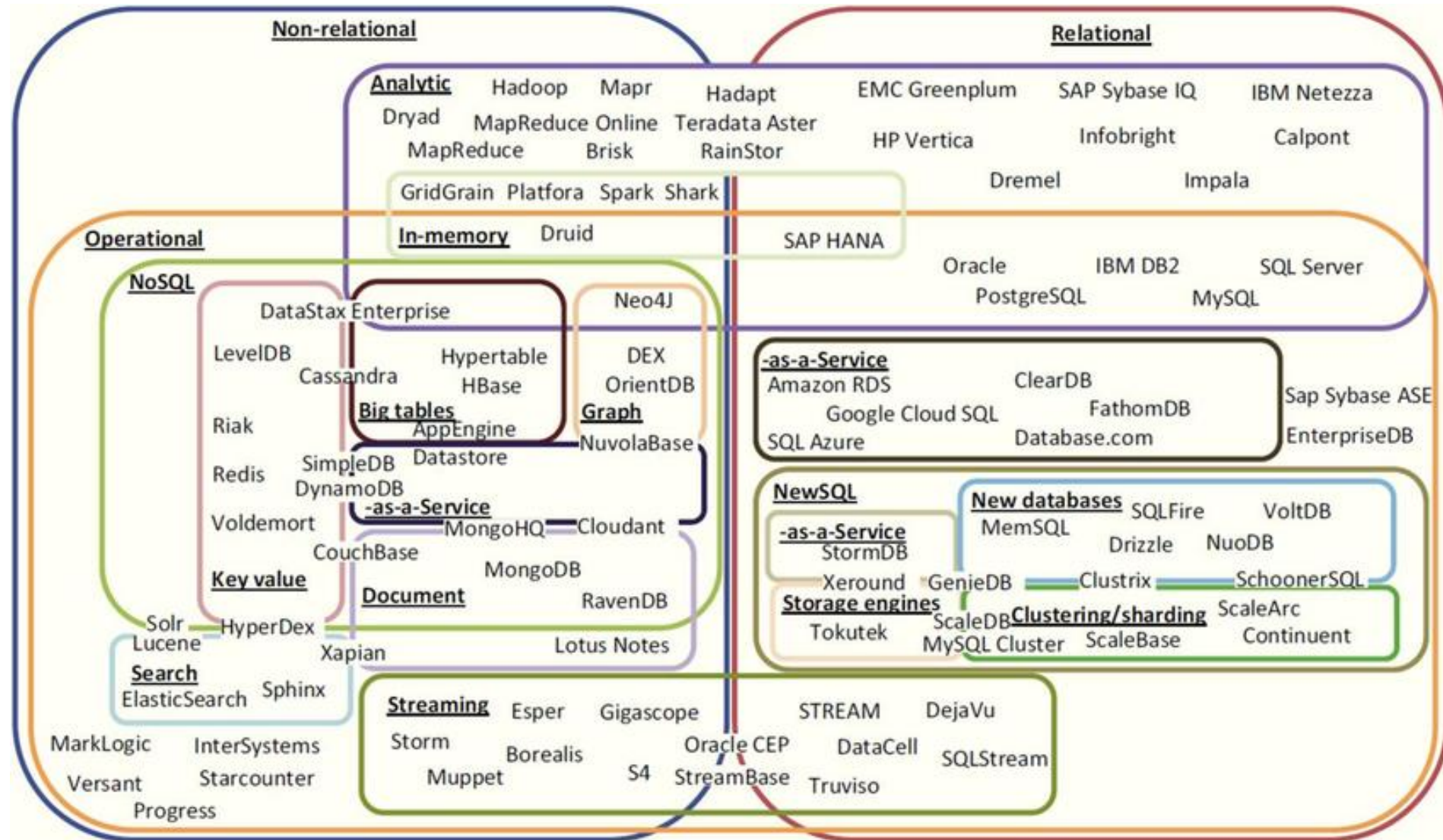
- **Systém Riadenia Bázy Údajov/Dát = SRBD** (angl. **DataBase Management System = DBMS**) je softvérový program, ktorý slúži na vytváranie, správu, úpravu a bezpečné ukladanie údajov v databáze.
- Funkcie DBMS
 - Správa dát a ich ukladanie (Data Storage)
 - Manipulácia s dátami (CRUD operácie)
 - Zabezpečenie a riadenie prístupu (Security)
 - Správa súbežného prístupu (Concurrency Control)
 - Zálohovanie a obnova dát (Backup & Recovery)
 - Udržiavanie integrity dát (Data Integrity)
- **Databázový systém (DBS)** je komplexný systém na spracovanie dát, ktorý je tvorený integráciou **štyroch** základných zložiek: samotnej bázy dát (databáza), systému riadenia bázy dát (softvéru), technických prostriedkov (hardvéru) a používateľov (vrátane administrátorov a koncových aplikácií).
- DBS = DB + DBMS + Hardvér + Používatelia
- Ak firma povie, že im zlyhal databázový systém, môže to znamenať, že spadol server (hardvér), vyskytla sa chyba v softvéri (DBMS) alebo sa poškodili samotné súbory (databáza).

Prehľadné porovnanie pojmov

Pojem	Čo to zjednodušene znamená?	Analógia z reálneho sveta
Dáta (Údaje)	Samotné surové fakty a hodnoty bez kontextu.	Jednotlivé papiere, zmluvy a bločky rozhádzané na stole.
Databáza (Báza dát)	Usporiadané a štruktúrované úložisko týchto dát.	Šanón v archíve, kde sú tie papiere presne roztriedené.
Systém riadenia bázy dát (SRBD / DBMS)	Softvér, ktorý umožňuje s databázou pracovať (hľadať, meniť, mazať).	Knihovník alebo sekretárka, ktorá vie, kde čo je, a dokumenty spravuje.
Databázový systém	Kompletné riešenie: Dáta + Databáza + Softvér (SRBD) + Používatelia.	Celá fungujúca firemná kartotéka vrátane personálu a počítačov.

Aké sú najpoužívanéjšie DBMS s ktorými sa uplatníť na svetovom trhu?

DataBase Management System



OnLine Transaction Processing = OLTP

- Vykonávanie transakcií v reálnom čase
- **Transakcia** je logická jednotka práce v databáze, ktorá zahŕňa jeden alebo viac príkazov vykonaných ako **nedeliteľný logický celok**
- **Štyri** základné vlastnosti (**ACID**):
 - **A – Atomicity (Atomickosť/Nedeliteľnosť)**: Transakcia sa vykoná buď celá, alebo vôbec. Neexistuje nič medzi tým.
 - **C – Consistency (Konzistencia)**: Transakcia prenesie databázu z jedného platného stavu do druhého platného stavu, pričom neporuší žiadne pravidlá ani obmedzenia.
 - **I – Isolation (Izolovanosť)**: Operácie vnútri transakcie sú skryté pred ostatnými súbežne prebiehajúcimi transakciami, kým sa transakcia nedokončí.
 - **D – Durability (Trvácnosť/Odolnosť)**: Akonáhle je transakcia úspešne dokončená, jej zmeny sú trvalé a nestratia sa ani pri následnom výpadku systému.
- Využitie: Dátové sklady
 - Spotify personalizuje domovské stránky pomocou vlastných skladieb a playlistov na základe preferencií používateľa.
 - Systém odporúčania filmov na platforme Netflix.

OnLine Transaction Processing = OLTP

- **Výhody OLTP**

- **Umožňujú používateľom** rýchlo vykonávať operácie čítania, zápisu a mazania dát.
- **Podporujú nárast** počtu používateľov a transakcií pre prístup k dátam v reálnom čase.
- **Poskytujú lepšiu ochranu** dát prostredníctvom viacerých bezpečnostných funkcií.
- **Pomáhajú pri rozhodovaní** vďaka presným a aktuálnym dátam.
- **Zaisťujú integritu**, konzistenciu a vysokú dostupnosť dát.

- **Nevýhody OLTP**

- **Obmedzená analytická schopnosť** – nie sú vhodné na komplexnú analýzu alebo reportovanie.
- **Vysoké náklady na údržbu** z dôvodu častých aktualizácií, zálohovania a obnovy dát.
- **Náchylnosť na výpadky** počas zlyhania hardvéru, čo priamo ovplyvňuje online transakcie.
- **Sklon k problémom**, ako sú duplicitné alebo nekonzistentné dáta.

Online Analytical Processing = OLAP

- Určená na rýchlu, viacrozmernú analýzu veľkého objemu firemných dát (Big Data).
- Používa sa najmä na reporting, plánovanie, prognózy a manažérske rozhodovanie.

Online Analytical Processing = OLAP

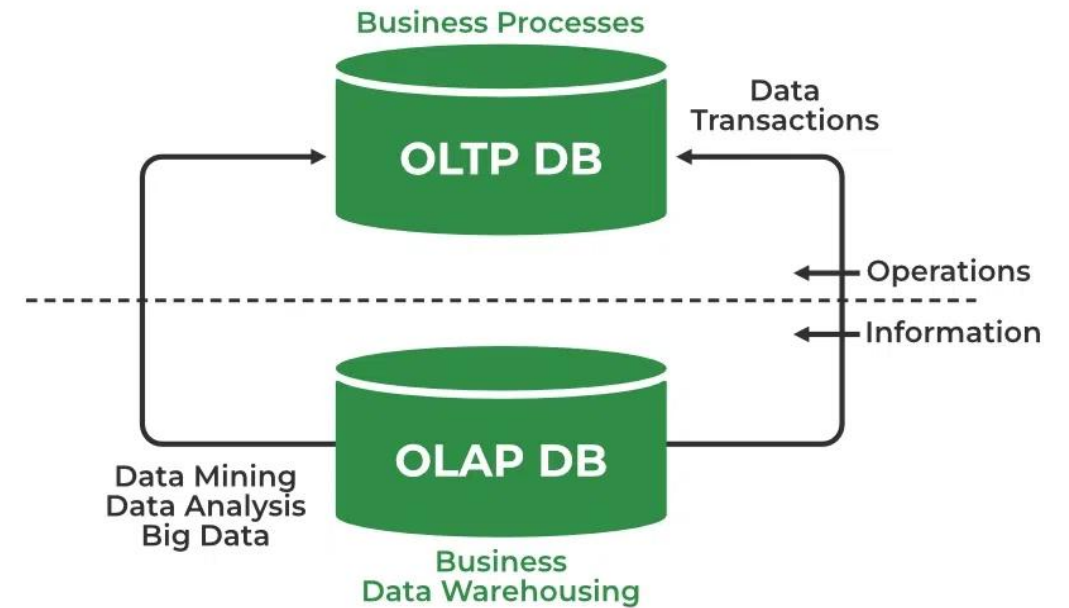
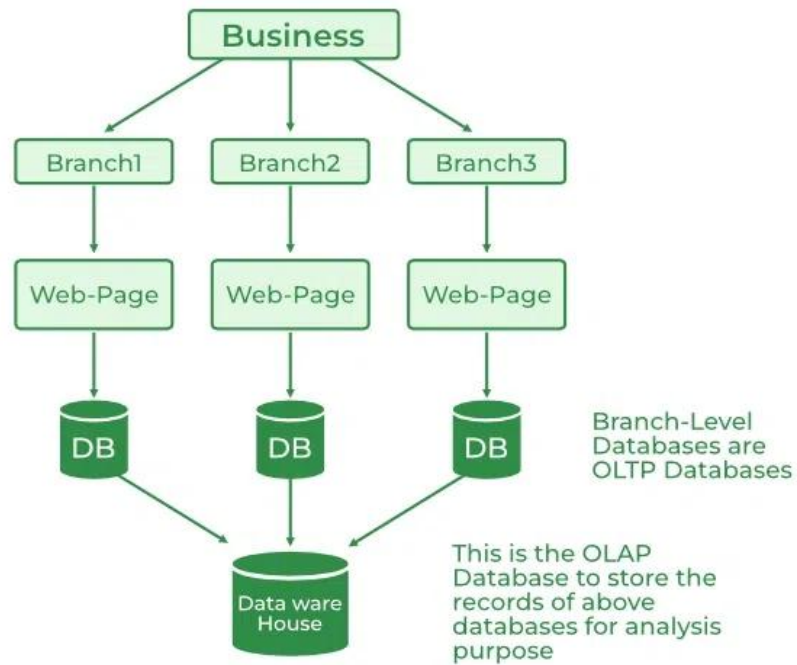
- **Výhody OLAP**

- **Pomáhajú udržiavať** konzistenciu a vykonávať výpočty nad dátami.
- **Dokážu uchovávať** plánovanie, analýzu a tvorbu rozpočtov pre podnikovú analytiku v rámci jednej platformy.
- **Efektívne spracovávajú** veľké objemy dát, vďaka čomu sú vhodné pre podnikové aplikácie na firemnej úrovni.
- **Pomáhajú pri uplatňovaní** bezpečnostných obmedzení na ochranu dát.
- **Poskytujú multidimenzionálny pohľad** na dáta, čo pomáha aplikovať operácie na dáta rôznymi spôsobmi.

- **Nevýhody OLAP**

- **Vyžadujú profesionálov** na správu dát kvôli zložitému postupu modelovania.
- **Sú nákladné** na implementáciu a údržbu v prípadoch, keď ide o veľké súbory dát.
- **Analýza dát prebieha** až po ich extrakcii a transformácii, čo vedie k oneskoreniam systému.
- **Nie sú dostatočne efektívne** pre operatívne rozhodovanie, keďže sa aktualizujú periodicky (v pravidelných intervaloch).

OLAP vs OLTP



Rozdiely OLAP vs OLTP

Kategória	OLAP (Online Analytical Processing)	OLTP (Online Transaction Processing)
Zdroj dát	Historické dáta z viacerých databáz.	Aktuálne prevádzkové (operačné) dáta.
Účel	Používa sa na analýzu a rozhodovanie.	Používa sa na každodenné transakcie.
Použitá metóda	Využíva dátový sklad (Data Warehouse).	Využíva štandardný systém riadenia bázy dát (DBMS).
Normalizácia	Tabuľky nie sú normalizované.	Tabuľky sú normalizované (3NF).
Typ dopytu	Komplexné dopyty s dôrazom na čítanie (pomalé).	Jednoduché dopyty na čítanie a zápis (rýchle).
Objem dát	Veľký (TB – PB).	Malý (MB – GB).
Aktualizácia	Aktualizované periodicky v dávkach (batches).	Často aktualizované samotnými používateľmi.
Záloha a obnova	Periodické zálohovanie.	Nepretržité (kontinuálne) a prísne zálohovanie.
Používatelia	Analytici, manažéri a vedúci pracovníci.	Úradníci, predajcovia a prevádzkoví pracovníci.
Zameranie	Orientované na tému/subjekt (analýza).	Orientované na aplikáciu (prevádzka).

Relačné vs nerelačné databázy

Vlastnosť	Relational (SQL)	NoSQL
Dátový model	Tabuľky s pevne danými riadkami a stĺpcami.	Dokumenty (JSON), kľúč-hodnota, grafy, široké stĺpce.
Schéma	Striktná a vopred definovaná (pred zápisom dát).	Dynamická a flexibilná (zmeny za behu).
Škálovanie	Vertikálne (zvyšovanie výkonu jedného servera).	Horizontálne (pridávanie ďalších serverov do siete).
Dopytovací jazyk	Štandardizovaný SQL (podpora zložitých spojení / JOINov).	Špecifické API podľa typu DB (Unql, CQL, JSON dopyty).
Záruky (Transakcie)	ACID: • Atomicity (Všetko alebo nič) • Consistency (Okamžitá konzistencia) • Isolation (Nezávislosť transakcií) • Durability (Trvalé uloženie)	BASE: • Basically Available (Vždy dostupná odpoveď) • Soft state (Meniaci sa stav synchronizácie) • Eventual consistency (Postupná konzistencia)
Podiel na trhu & Popularita	Dominantný na trhu (cca 75% - 80%): Tvorí drvivú väčšinu trhu. Preferovaná voľba pre kritické podnikové systémy.	Rýchlo rastúci (cca 20% - 25%): Jeho dynamický rast je stimulovaný masívnym nárastom neštruktúrovaných dát, rozvojom umelej inteligencie (AI), IoT a cloudových riešení.
Hlavná výhoda	Spoľahlivosť a integrita dát.	Rýchlosť, flexibilita a škálovateľnosť.
Príklady	PostgreSQL, MySQL, SQLite, Oracle.	MongoDB, Redis, Cassandra, Neo4j.