

# Databázové systémy a SQL

---

Daniel Klimeš

### Daniel Klimeš

- Vzdělání: Obecná biologie
- PGS: onkologie
- Specializace: klinické database, Národní zdravotní registry
- Databáze ORACLE, PostgreSQL, MS SQL
  
- [klimes@iba.muni.cz](mailto:klimes@iba.muni.cz)

Každou středu od 10:00 – do 11:40

Teoretická přednáška → navazující praktické cvičení

Praktická část : PostgreSQL,  
Domácí cvičení: PostgreSQL

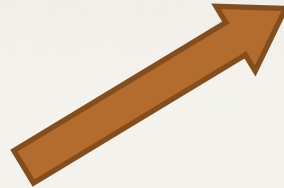
Zakončení: zápočet – domácí úkol  
zkouška – praktický test, pomůcky bez omezení, časový limit

1. Data jsou primárně v databázi uložena
2. Zpracováváme objemná data v řádu sto tisíc záznamů a více
3. Zpracování dat plánujeme provádět opakovaně
4. S daty bude pracovat více uživatelů

**Relační databáze** x NoSQL databáze

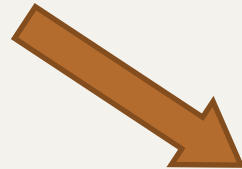
## Databáze SQL

- Předzpracování dat
- Čištění dat
- Popisná analýza,
- Filtrování
- Agregace dat



## Statistický SW

Statistika for Windows, R,  
SPSS, SAS, MS Excel



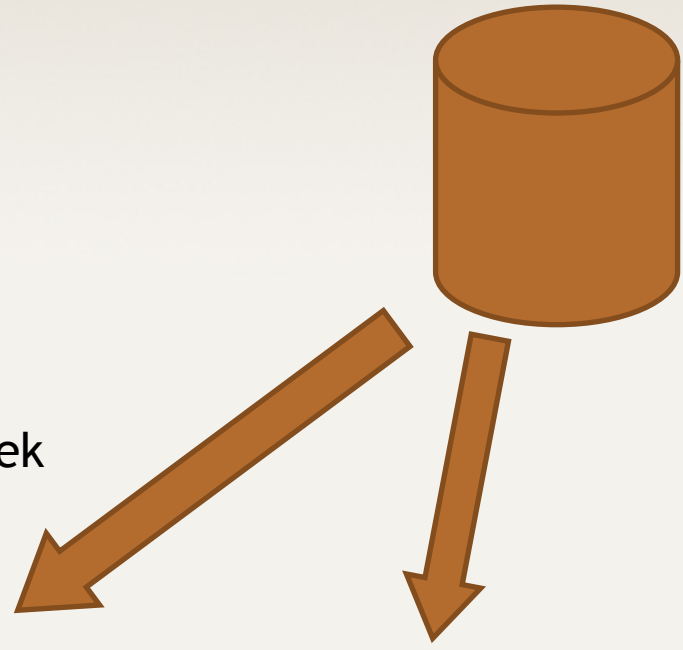
## Matematicko - statistický skriptový SW

R, Matlab, Maple, programovací jazyk

Relační databáze (RDBMS)  
 Relace - termín z relační algebry

Základ: tabulka  
 - sloupec = atribut/parametr  
 - řádek = popsáný objekt

Databáze = systém provázaných tabulek



ID pacienta	Jmeno	Prijmeni
1	Jan	Novák
2	Jana	Nová
3	Karel	Starý

ID pacienta	Datum vysetreni	Vysledek vysetreni
1	12.1.2011	39,5
1	15.3.2011	36,8
2	2.2.2011	37,5

## Dle dostupnosti

### Komerční

ORACLE - databáze\*  
MS SQL server\*  
DB2  
MS ACCESS  
FOX PRO

### Freeware

MySQL  
PostgreSQL  
Firebird

\* Okleštěné verze jsou k dispozici zdarma

## Dle počtu uživatelů

### Jednouživatelské

MS ACCESS  
FOX PRO

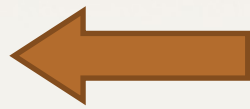
### Víceuživatelské

ORACLE  
MS SQL  
DB2  
MySQL  
PostgreSQL  
Firebird

Definovaná struktura, do které se vkládají záznamy

Definují se sloupce

- jméno
- datový typ
  - text
  - číslo
  - datum
  - BLOB
- doplňující vlastnosti



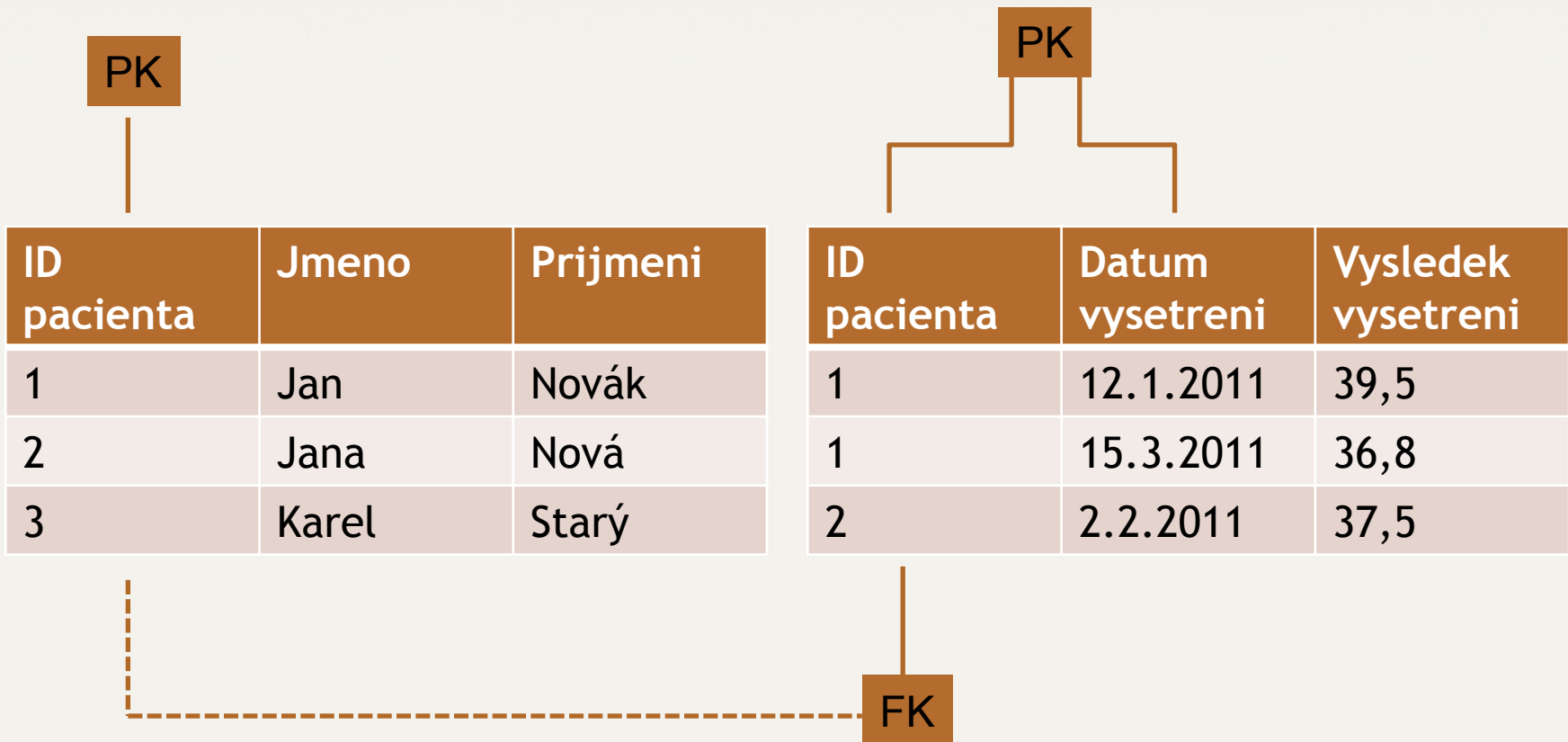
ID pacienta	Datum vyšetření	Vysledek vyšetření
1	12.1.2011	39,5
1	15.3.2011	36,8
2	2.2.2011	37,5

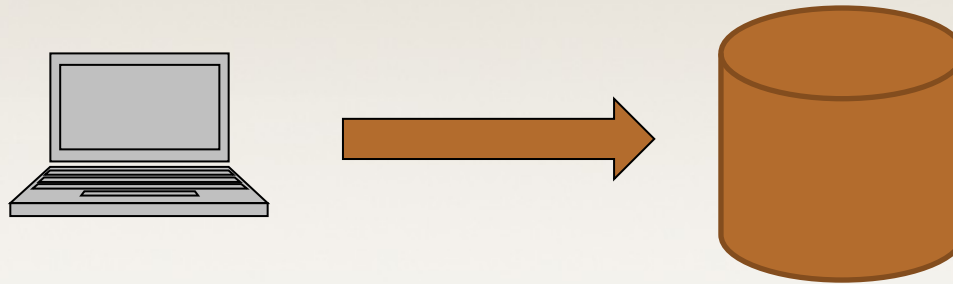


Vybrané sloupce se označují jako klíče (keys)

Primární klíč (primary key - PK)- 1 až n sloupců jednoznačně identifikující řádek

Cizí klíč (foreign key - FK) - identifikuje nadřazený řádek v rodičovské tabulce





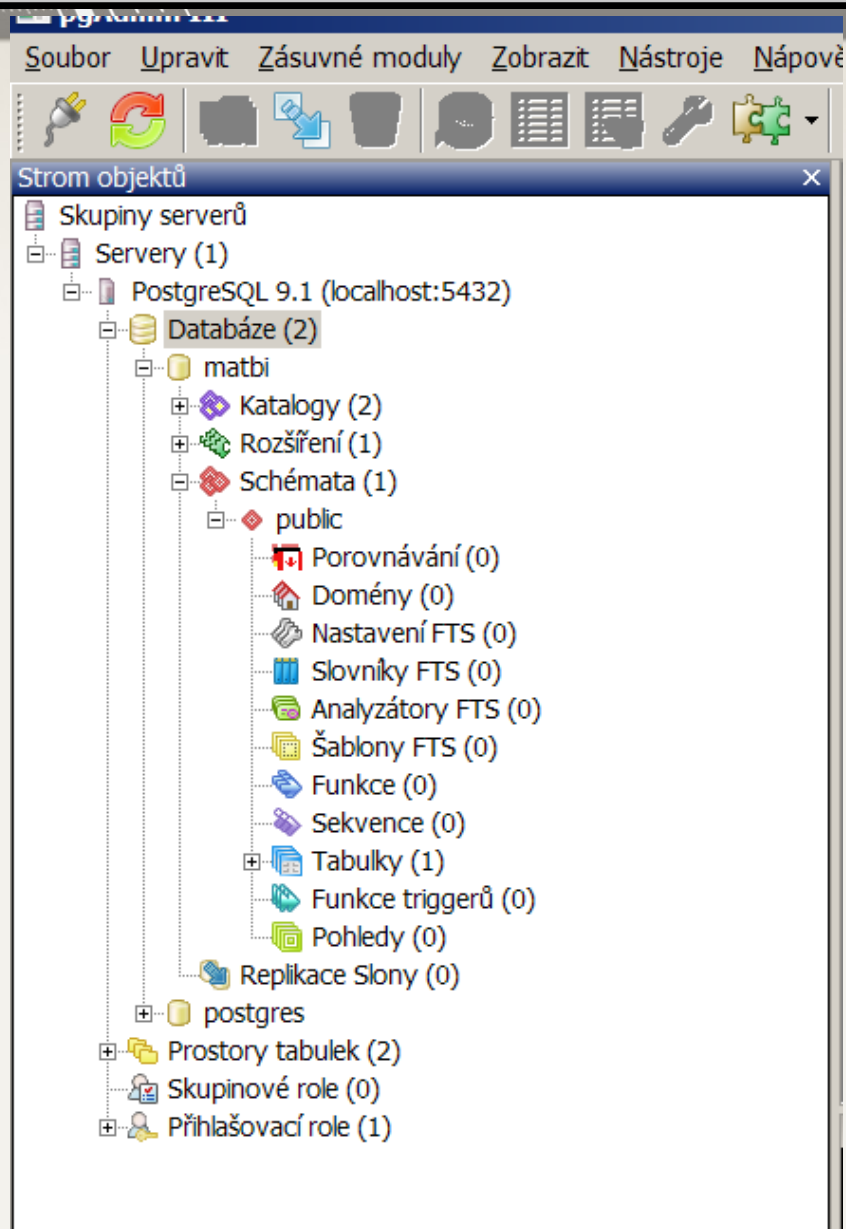
**Klient** = SW umožňující ověření uživatele a spuštění řídicích příkazů

Řídicí příkazy = Structured Query Language - SQL

- DDL - vytváření, změna, rušení objektů (tabulka, index, pohled, ...)
  - CREATE / ALTER / DROP
- DML
  - SELECT - získávání dat z databáze
  - INSERT - vkládání dat do databáze
  - DELETE - mazání dat v databázi
  - UPDATE - změna/aktualizace dat
- transakční příkazy
  - COMMIT - potvrzení transakce
  - ROLLBACK - odvolání transakce

## Klient

- **pgAdmin**
  - Přihlášení k serveru
  - Nová databáze - matbi
  - Schémata – public
- 
- **psql**
  - ve Windows nutné nastavení jazykové sady a fontu



## SQL jazyk (Structured Query Language)

- case insensitive
- klíčová slova – pro názornost VELKÝM písmem
- názvy objektů (tabulek, sloupců)
  - pouze alfanumerické znaky
  - první znak písmeno
  - omezená délka (ORACLE 32 znaků)
- operátory
- funkce
- různá rozšíření v jednotlivých DB produktech
- SQL příkazy – ve skriptu ukončeny defaultně středníkem (;)
- komentáře odděleny - -                   nebo v bloku /\* komentar \*/

SELECT \* FROM tabulka; -- všechny řádky i sloupce tabulky

/\* vybrané sloupce, všechny řádky \*/

SELECT sloupec1, sloupec2, sloupec1 + sloupec2 AS soucet FROM tabulka;

/\* všechny sloupce, vybrané řádky \*/

SELECT \* FROM tabulka WHERE sloupec1 = 1;

SELECT \* FROM tabulka WHERE sloupec2 = 'Jan' ; -- text do apostrofů

SELECT \* FROM tabulka

WHERE sloupec1 = 1 AND sloupec2 > 10 AND sloupec3 < sloupec 4

--Setřídění výstupu (ORDER BY)

SELECT \* FROM tabulka ORDER BY sloupec1; -- vzestupné třídění

SELECT \* FROM tabulka ORDER BY sloupec2 DESC; -- sestupné třídění

SELECT \* FROM tabulka ORDER BY sloupec1, sloupec2 DESC --kombinace

SELECT \* FROM tabulka; -- všechny řádky i sloupce tabulky

/\* vybrané sloupce, všechny řádky \*/

SELECT sloupec1, sloupec2, sloupec1 + sloupec2 AS soucet FROM tabulka;

/\* všechny sloupce, vybrané řádky \*/

SELECT \* FROM tabulka

WHERE sloupec1 = 1 AND sloupec2 > 10 AND sloupec3 < sloupec 4

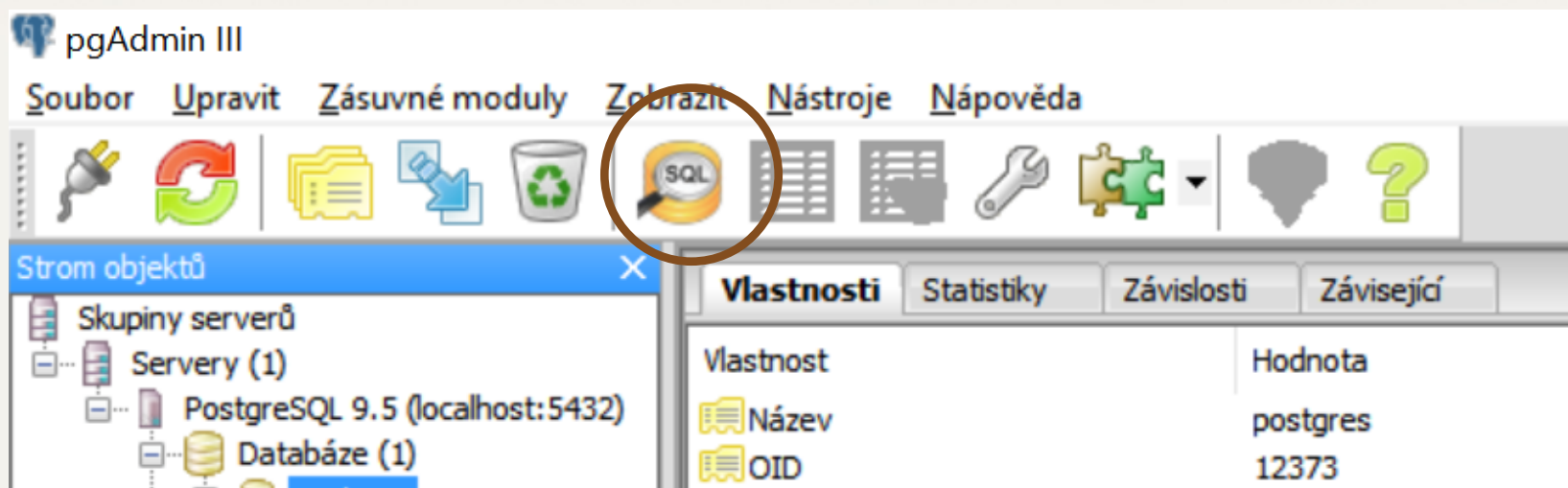
--Sumární výstupy/Agregační funkce

SELECT COUNT(\*) FROM tabulka -- počet řádků v tabulce

SELECT SUM(sloupec1), AVG(sloupec2), MIN(sloupec3), MAX(sloupec4)  
FROM tabulka

SELECT COUNT(\*), sloupec1 FROM tabulka - nelze

## PgAdmin - spuštění, přihlášení, otevření databáze

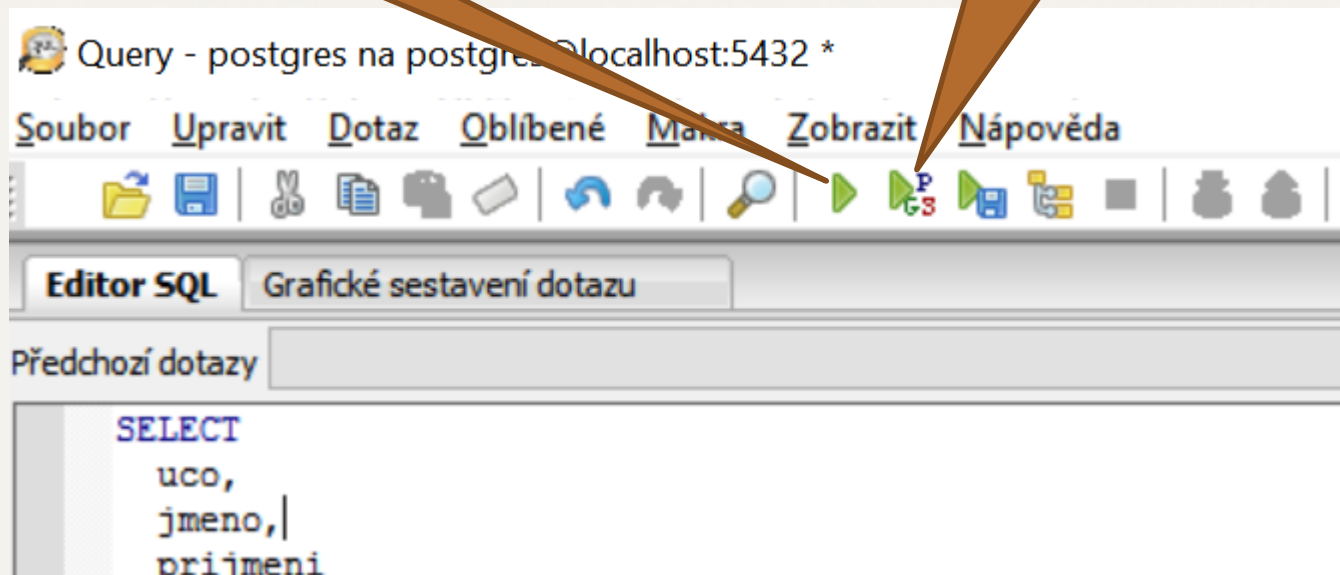


Příkazové okno – CTRL-E

# Příkazový řádek

Spuštění  
jednoho příkazu

Spuštění skriptu





Spuštění prvního skriptu (skript1.sql)

Tabulka **STUDENT**

- Zobrazte celý obsah tabulky
- Zobrazte jen jméno a příjmení
- Setříd'te výstup podle studia, příjmení
- Kolik má tabulka řádků?
- Vyberte pouze svůj záznam (své UČO)
- Minimum, maximum a průměrná hodnota sloupce *UCO*?
- Minimum, maximum z příjmení
- Kolik máme v seznamu Terez?

## 1. Přečíst kapitolu 1 ve skriptech

<http://portal.matematickabiologie.cz/index.php?pg=zaklady-informatiky-pro-biology--databazove-systemy-v-biomedicine>

## 2. Instalace Postgresql