

**MUNI  
MED**

# **Informatika a statistika ve zdravotnictví - cvičení**

Tomáš Pavlík ([pavlik@iba.muni.cz](mailto:pavlik@iba.muni.cz))

Daniel Klimeš ([klimes@iba.muni.cz](mailto:klimes@iba.muni.cz))

jaro 2022

# Data management

*Tato prezentace je autorským dílem vytvořeným zaměstnanci Masarykovy univerzity. Studenti předmětu mají právo pořídit si kopii prezentace pro potřeby vlastního studia. Jakékoliv další šíření prezentace nebo její části bez svolení Masarykovy univerzity je v rozporu se zákonem.*

# Kvalitní data

- Chybná data => chybná statistika => chybné závěry
- Bezchybná data neexistují
- Všichni chtějí data, ale nikdo je nechce sbírat
- 80% času analýzy dat zabere čištění dat
- Pro sběr dat lze použít sofistikované nástroje – jenže jsou dost drahé
- ... a mnohé datové soubory tak vznikají v excelu

# **Pokročilé funkce MS Excel**

# Základní ovládání

- Označení bloku
- CTRL – šipka
- CTRL – Shift – šipka
- Kopírování hodnot – pravý dolní roh
- Datové typy sloupců – Pravé tlačítko – Formát buněk
  - Obecný
  - Číslo
  - Datum
  - Text
  - Automatické formátování a jak mu zabránit

# Filtry

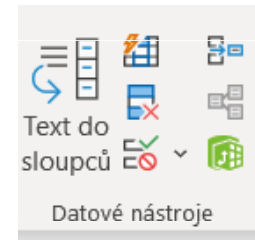
- Označit => Data => Filtr
- Rychlá kontrola unikátních hodnot
- Filtrování a následné odstranění chyb

# Vzorce

- Výběr v menu nebo „=„
- Odkaz na buňku A1 x \$A1 x \$A\$1
- & = spojení/řetězení
- VYČISTIT, PROČISTIT, NAHRADIT
- SVYHLEDAT

# Odstranění duplicit, text do sloupců

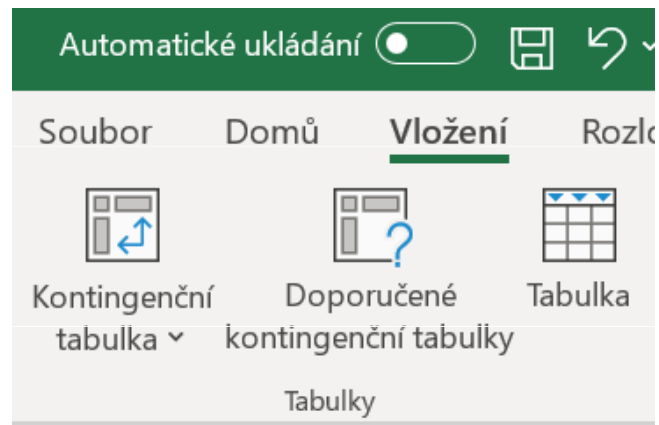
- Označení bloku
- Data => Odebrat duplicity
- Data => Text do sloupců





# Kontingenční tabulka

- Rychlý přehled
- Dohledání duplicit

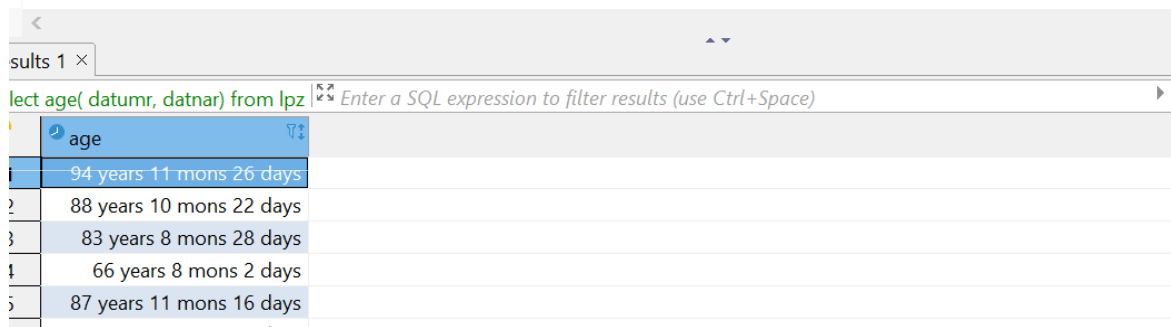


# Cvičení

- Soubor ockovani.xlsx
  - Je zde cislopacienta jednou nebo opakovaně?
  - Nahrazení NULL za „prázdkno“
  - Hl. m. Praha => Praha
  - Počet dnů mezi vakcinací a vykázáním
  - Doplnit ICO na 8 míst
  - Doplnit datum narození
  - Výpočet přibližného věku
  - Počty po krajích, po zařízeních
  - Filtr na Smluvní zařízení, Kraj

# Ukázka práce s relační databází

```
select age( datumr, datnar) from lpz
```



The screenshot shows a SQL query execution interface. The query is `select age( datumr, datnar) from lpz`. The result is displayed in a table with one column named 'age'. The results are:

age
94 years 11 mons 26 days
88 years 10 mons 22 days
83 years 8 mons 28 days
66 years 8 mons 2 days
87 years 11 mons 16 days

**Podrobně:**

**Bi5447 Databázové systémy v biomedicíně**

**MUNI  
MED**

**Děkuji za  
pozornost**