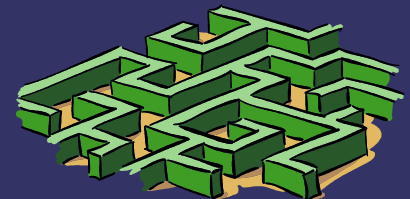


# *Statistika*

Statistické funkce v tabulkových kalkulátorech  
MSO Excel a OO.o Calc



# Základní pojmy tabulkových kalkulátorů

- ➔ Cílem není vyložit pojmy tabulkových kalkulátorů, ale jen definovat pojmy vyskytující se v této prezentaci
- *zadaný rozsah buněk:*
  - oblast buněk vymezená levou horní a pravou spodní buňkou, př. A1:C13
  - různé spojené oblasti, př. A1:A13,B1:B12
  - jednotlivé údaje nebo odkazy na buňky odděleny čárkou
- *Alfa:* číslo v rozsahu 0 - 1

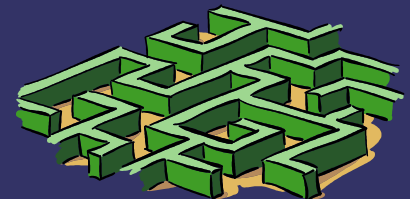


# Průměry

⇒ Aritmetický průměr – součet čísel vydělený jejich počtem

$$\bar{x} = \frac{1}{n} \cdot (x_1 + x_2 + \dots + x_n) = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

- Excel: =PRŮMĚR(*zadaný rozsah buněk*)
- Calc: =AVERAGE(*zadaný rozsah buněk*)

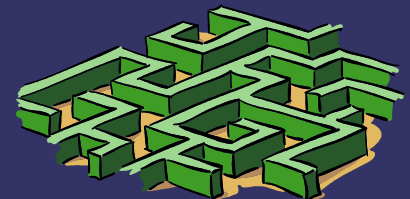


# Vážený průměr

- ➔ **Vážený průměr** – u každé veličiny je znám počet výskytů

$$\bar{x} = \frac{w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i}$$

- Excel: =SOUČIN.SKALÁRNÍ(rozsah buněk s počtem výskytu; rozsah buněk s hodnotami)/SUMA(rozsah buněk s počtem výskytu)
- Calc: =SUMPRODUCT(rozsah buněk s počtem výskytu; rozsah buněk s hodnotami)/SUM(rozsah buněk s počtem výskytu)

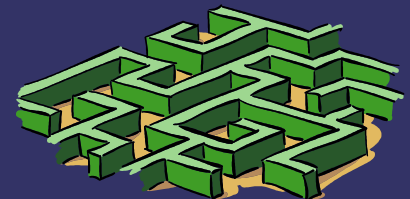


# Průměry

- ➔ Geometrický průměr – n-tá odmocnina ze součinu n čísel

$$\bar{x}_G = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

- Oba tabulkové kalkulátory  
=GEOMEAN(*zadaný rozsah buněk*)

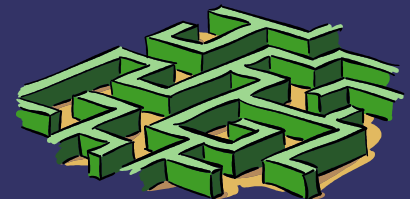


# Průměry

- ➔ Harmonický průměr – využívá se tam, kde má smysl součet převrácených hodnot proměnné

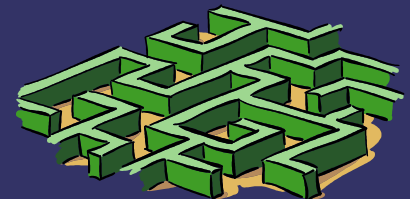
$$\bar{x}_H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

- Oba tabulkové kalkulátory  
=HARMEAN(*zadaný rozsah buněk*)



# Výjimečné (extrémní) hodnoty

- ⇒ Maximum – největší hodnota
- ⇒ Minimum – nejmenší hodnota
  - Oba tabulkové kalkulátory
    - =MAX(*zadaný rozsah buněk*)
    - =MIN(*zadaný rozsah buněk*)
- ⇒ K-tá největší hodnota
  - Oba tabulkové kalkulátory
    - =LARGE(*zadaný rozsah buněk; pořadí*)
- ⇒ K-tá nejmenší hodnota
  - Oba tabulkové kalkulátory
    - =SMALL(*zadaný rozsah buněk; pořadí*)

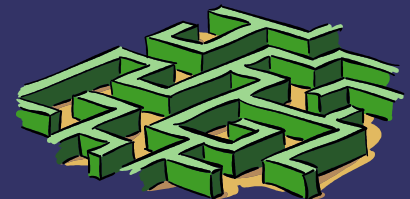


# Střední hodnoty

- ➔ Modus – hodnota nejčastěji se vyskytujícího znaku v souboru

$\hat{x}$

- Oba tabulkové kalkulátory  
=MODE(*zadaný rozsah buněk*)



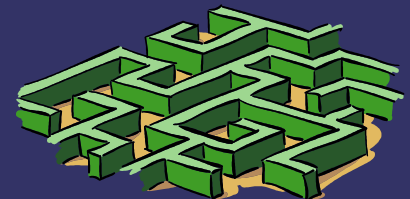


# Střední hodnoty

- ⇒ Medián – prostřední hodnota v souboru uspořádaném podle velikosti při lichém počtu; průměr dvou prostředních hodnot v souboru uspořádaném podle velikosti při sudém počtu

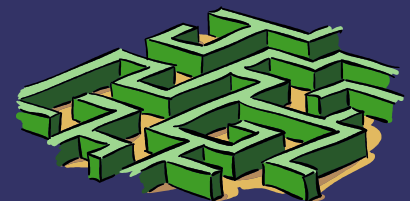
$\tilde{x}$

- Oba tabulkové kalkulátory  
=MEDIAN(*zadaný rozsah buněk*)



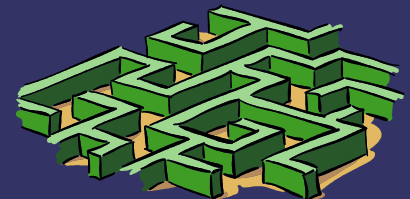
# *Kvantily*

- ➔ Kvantil – hodnota proměnné, kdy hodnoty, které jsou menší (a stejné), tvoří určitou stanovenou část rozsahu statistického souboru, např. 1, 25, 50, 90 % apod., kdežto hodnoty, které jsou větší (a stejné), tvoří zbývající část rozsahu souboru, tj. např. 99, 75, 50, 10 % atd. 50% kvantil je medián.
- ➔ Kvartily – tři hodnoty proměnné, které rozdělují neklesající řadu hodnot proměnné na čtyři stejně četné části. První – dolní kvartil je 25% kvantil (odděluje čtvrtinu statistických jednotek s nejnižší hodnotou proměnné  $x$  od tří čtvrtin jednotek s vyšší, popř. stejnou hodnotou proměnné  $x$ ). Prostřední kvartil = medián, třetí – horní kvartil je 75% kvantil.



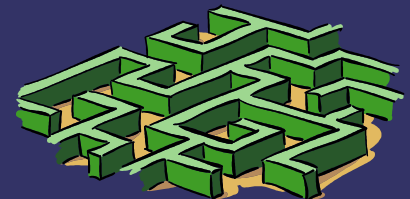
# *Kvantily*

- EXCEL:      =QUARTIL(*zadaný rozsah buněk;Typ*)
- CALC:      =QUARTILE(*zadaný rozsah buněk;Typ*)
  
- Typ – 0 .. minimum, 1 .. první (25%) kvartil, 2 .. medián, 3 .. třetí (75%) kvartil, 4 .. maximum



# Kvantily

- ⇒ Decily tvoří devět hodnot proměnné, které rozdělují neklesající řadu hodnot proměnné na deset stejně četných částí. První decil je 10% kvantil, druhý 20% kvantil, ..., devátý decil 90% kvantil.
- ⇒ Percentily tvoří 99 hodnot proměnné, které rozdělují neklesající řadu hodnot proměnné na 100 stejně četných částí. První percentil je 1% kvantil, druhý percentil 2% kvantil, ..., 99. percentil je 99% kvantil.
- EXCEL:        =PERCENTIL(*zadaný rozsah buněk;Alfa*)
- CALC:         =PERCENTILE(*zadaný rozsah buněk;Alfa*)

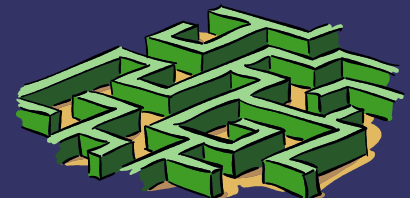


# Variabilita

- ➔ Průměrná absolutní odchylka – aritmetický průměr absolutních hodnot znaku všech prvků souboru od aritmetického průměru souboru

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

- EXCEL: =PRŮMODCHYLKA(*zadaný rozsah buněk*)
- CALC: =AVEDEV(*zadaný rozsah buněk*)

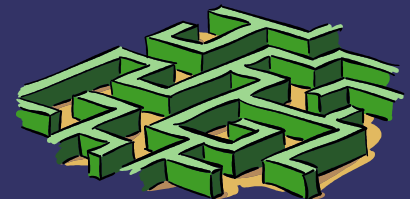


# Variabilita

- ➔ Rozptyl – aritmetický průměr druhých mocnin odchylek hodnot znaku od aritmetického průměru

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

- EXCEL: =VAR(*zadaný rozsah buněk*)
- CALC: =VARP(*zadaný rozsah buněk*)  
kromě toho jen součet čtverců odchylek oba kalkulátory
- =DEVSQ(*zadaný rozsah buněk*)  
odmocnina z rozptylu (směrodatná odchylka)
- EXCEL: =SMODCH(*zadaný rozsah buněk*)
- CALC: =STDEVP(*zadaný rozsah buněk*)



# Variabilita

⇒ Variační koeficient – poměr směrodatné odchylky a aritmetického průměru

- EXCEL:      =SMODCH(*zadaný rozsah buněk*)/PRŮMĚR(*zadaný rozsah buněk*)
- CALC:      =STDEVP(*zadaný rozsah buněk*)/AVERAGE(*zadaný rozsah buněk*)

