

# Rozdělení zemí na základě zdravotních, sociálních a ekonomických parametrů

---

## Teoretické ukotvení

Při mezinárodním srovnávání zemí je běžné usilovat o jejich systematické rozdělení do skupin na základě vybraných charakteristik, které odrážejí jejich socioekonomickou úroveň a celkové životní podmínky. Tento přístup vychází z předpokladu, že mezi státy existují strukturální rozdíly v oblasti zdraví, sociální situace a ekonomické výkonnosti, které se vzájemně ovlivňují a společně utvářejí jejich postavení v globálním systému (Delahoz-Domínguez et al., 2025).

Způsoby, jakými jsou země v rámci těchto srovnání kategorizovány, přitom historicky vznikaly v kontextu nerovnoměrných mocenských vztahů, které se promítly i do současného chápání globálních rozdílů mezi státy (Khan et al., 2022).

V současném výzkumu se stále více prosazuje vícerozměrný přístup, který umožňuje komplexně zachytit rozdíly mezi jednotlivými státy a zároveň lépe porozumět jejich vzájemným podobnostem i odlišnostem, jelikož socioekonomický rozvoj nelze adekvátně popsat pomocí jediného ukazatele, ale vyžaduje kombinaci ekonomických a sociálních faktorů (Saisana & Tarantola, 2002; Milićević et al., 2014).

## Popis dat

Pro účely této analýzy byla použita cvičná data zahrnující 167 zemí, která obsahují informace o ekonomických a zdravotních ukazatelích.

### Konkrétně se jedná o proměnné:

- **Dětská úmrtnost** – počet úmrtí dětí do 5 let na 1000 živě narozených
- **Export** – export zboží a služeb na obyvatele
- **Zdraví** – celkové výdaje na zdravotnictví na obyvatele
- **Import** – import zboží a služeb na obyvatele
- **Příjem** – čistý průměrný příjem na osobu
- **Růst HDP** – roční růst HDP v procentech

- **Střední délka života** – průměrná délka dožití
- **Úhrnná porodnost** – průměrný počet dětí na jednu ženu
- **HDP na obyvatele** – hrubý domácí produkt dělený počtem obyvatel

## Analýza dat

Data byla zpracována v programu Statistica. Nejprve jsem použila metodu **hierarchické shlukové analýzy**, která mi umožnila vizualizovat strukturu dat a odhadnout vhodný počet skupin. Následně jsem aplikovala **metodu K-Means**, která data rozděluje do předem určeného počtu shluků (v mém případě 3).

Princip metody spočívá v určení středních bodů (centroidů) jednotlivých skupin, ke kterým se následně přiřazují jednotlivé záznamy podle nejbližší vzdálenosti. Po přiřazení všech dat se centroidy aktualizují na základě průměrných hodnot v daném shluku a celý proces se opakuje, dokud se jejich pozice nezastabilizuje.

Vzhledem k rozdílným jednotkám a rozsahům hodnot jsem data před analýzou **standardizovala na z-skóry**. Tato transformace zajistila, že každá proměnná měla na výsledné shlukování **rovnoměrný a srovnatelný vliv**.

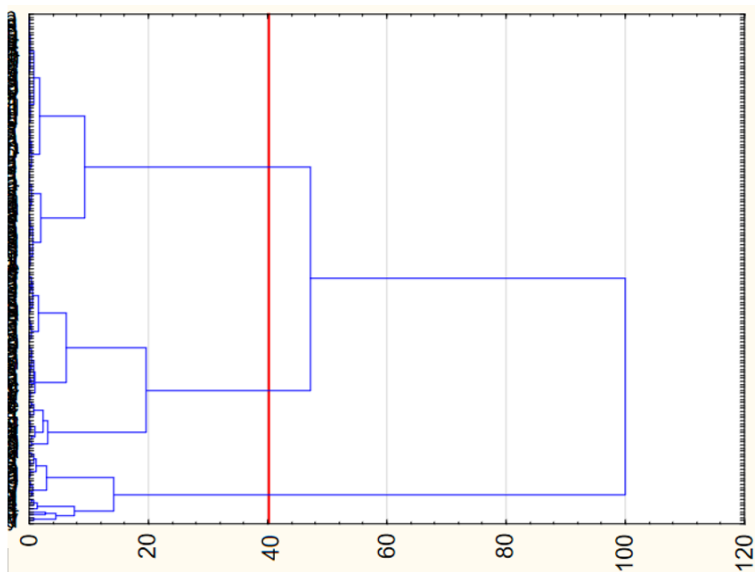
## Výsledky analýzy

### **Hierarchická shluková analýza**

Pro spojování shluků lze využít různých metod. Já jsem zvolila Wardovu metodu, která při každém kroku slučování minimalizuje součet čtverců odchylek uvnitř shluků, což umožňuje minimalizovat vnitřní variabilitu a vytvářet homogenní skupiny.

Jako míru vzdálenosti jsem použila **Euklidovskou vzdálenost**. Obrázek č. 1 zobrazuje dendrogram shlukové analýzy, na jehož základě jsem rozdělila data do 3 klastrů.

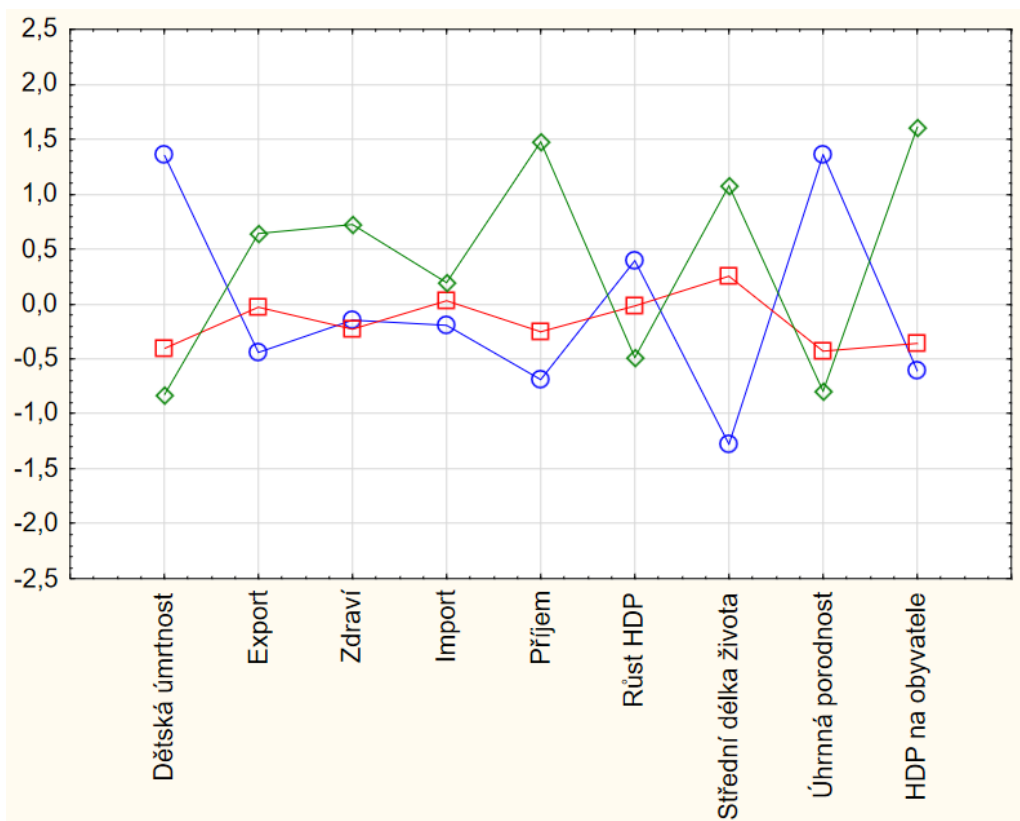
Obrázek č. 1- Stromový graf hierarchické shlukové analýzy



## Metoda K-means

Na základě výsledků hierarchické analýzy jsem použila metodu **K-průměrů** s předem určeným počtem 3 klastrů.

Obrázek č. 2- Graf průměrných skóre jednotlivých klastrů



Z grafu je patrné, že **Cluster č. 1** se vyznačuje nadprůměrnou dětskou úmrtností, nízkou střední délkou života a podprůměrnými hodnotami HDP na obyvatele, což naznačuje nižší úroveň rozvoje.

Naopak **Cluster č. 3** dosahuje opačných charakteristik, tedy nízké dětské úmrtnosti, vysoké střední délky života, vysokého příjmu a nadprůměrného HDP na obyvatele, což odpovídá vyspělým ekonomikám s vyšší životní úrovní.

**Cluster č. 2** se pohybuje převážně kolem průměrných hodnot sledovaných proměnných a představuje tak střední skupinu bez výrazných extrémů.

## Závěr

Na základě provedené shlukové analýzy byly země rozděleny do tří skupin, které se mezi sebou liší především v úrovni ekonomických a zdravotních ukazatelů. První skupina zahrnuje země s horšími hodnotami sledovaných ukazatelů, zejména s vyšší dětskou úmrtností, nižší střední délkou života a nižší ekonomickou úrovní. Naopak třetí skupina je charakteristická příznivými hodnotami, tedy nízkou úmrtností, vyšší délkou života a vyššími hodnotami příjmu a HDP na obyvatele. Druhá skupina představuje přechod mezi těmito dvěma extrémy a vyznačuje se převážně průměrnými hodnotami.

## Literatura

Delahoz-Domínguez, E., Mendoza-Mendoza, A., & Visbal-Cadavid, D. (2025). Clustering of countries through UMAP and K-means: A multidimensional analysis of development, governance, and logistics. *Logistics*, 9(3), Article 108. <https://doi.org/10.3390/logistics9030108>

Khan, T., Abimbola, S., Kyobutungi, C., & Pai, M. (2022). How we classify countries and people—and why it matters. *BMJ Global Health*, 7(6). <https://doi.org/10.1136/bmjgh-2022-009704>

Milićević, M., Dobrota, M., & Savić, G. (2014). A multivariate approach in measuring socio-economic development of MENA countries. *Economic Modelling*, 38, 604–608. <https://doi.org/10.1016/j.econmod.2014.02.011>

Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development*. European Commission