



Összetett jelenségek, osztályozás

Összetett jelenségek mérésének igénye

- Területi kutatásokban gyakori az összetett társadalmi-gazdasági jelenségek mérésének igénye
 - Fejlettség
 - Urbanizáltság (városiasodottság) – ruralitás (vidékiiség)
 - Periferialitás – társadalmi vetülete: elszigeteltség, kirekesztettség, marginalizáltság
 - Életkörülmények, jóllét
 - Innovativitás
 - Versenyképesség
- Összetett jelenségek mérésének igénye más társadalomtudományban is jelentkezik csak a megfigyelési egységek változnak – pl: egyének, társadalmi csoportok,



Komplex mutatók

Komplex mutatók

- Több adatsor együttes figyelembevételének igénye teremti meg
- Gyakorik a fejlesztéspolitikában → tipizálás előfeltételeként
 - Település/térségkategóriák megállapítása
 - Centrum-periféria
 - Vidéki térségek lehatárolása
 - Egyes térségek fejlődési pályájának meghatározása (Izd. Csatári Bálint)
 - Kedvezményezettségi kategóriák meghatározására
 - 31 I/2007. (XI. 17.) Korm. Rendelet → besorolásról → 32 változó → pontozásos módszer
 - Komplex mutató változói: 3. számú melléklet a 67/2007. (VI. 28.) OGY határozathoz
 - 106/2015. (IV. 23.) Korm. Rendelet → 23 változó → normalizáláson alapuló

1. mutatócsoport: Társadalmi és demográfiai helyzet
 - Halálozási ráta (az 1000 lakosra jutó halálozások száma) az elmúlt öt év átlaga, ‰
 - Vándorlási különbözet ezer lakosra (az utolsó öt év átlaga), fő
 - Tízezer 0–2 éves állandó lakosra jutó bölcsődei és családi napközis férőhelyek száma, db
 - Rendszeres gyermekvédelmi kedvezményben részesítettek aránya a 0–24 éves állandó népességből, ‰
 - Aktív korúak ellátásában (rendszeres szociális segélyben és foglalkoztatást helyettesítő támogatásban) részesítettek ezer állandó lakosra jutó száma, fő
2. mutatócsoport: Lakás és életkörülmények
 - Használt lakások átlagos ára, Ft
 - Az utolsó öt év során épített lakások aránya az időszak végi lakásállományból, ‰
 - Komfort nélküli (lakott) lakások a lakott lakások %-ában
 - Az egy állandó lakosra jutó SZJA-alapot képező jövedelem, ezer Ft
 - A természetes személyek által üzemeltetett személygépkocsik kor szerint súlyozott ezer lakosra jutó száma, db
3. mutatócsoport: Helyi gazdaság és munkaerőpiac
 - A 18 éves és idősebbek között a legalább középiskolai érettségivel rendelkezők aránya, ‰
 - Nyilvántartott álláskeresők aránya a munkaképes korú állandó népességből, ‰
 - Tartósan – legalább 12 hónapja folyamatosan – nyilvántartott álláskeresők aránya a munkaképes korú állandó népességből, ‰
 - A legfeljebb általános iskolát végzett nyilvántartott álláskeresők aránya, ‰
 - A működő vállalkozások ezer lakosra jutó száma, db
 - Kiskereskedelmi üzletek ezer lakosra jutó száma, db
 - Az önkormányzatok helyi adóbevételeinek aránya a tárgyévi bevételekből, ‰
4. mutatócsoport: Infrastruktúra és környezet
 - Közüzemi szennyvízgyűjtő-hálózathoz kapcsolódott lakások aránya, ‰
 - A rendszeres hulladékgyűjtésbe bevont lakások aránya, ‰
 - Szélessávú internet előfizetők ezer lakosra jutó száma, db
 - Kiépített utak aránya az összes önkormányzati fenntartású közútból, ‰
 - A megyeszékhely elérésének mutatója, perc
 - A gyorsforgalmi csomópontok elérési mutatója, perc

Fejlettség mérése

- Összetett jelenség
- Több dimenzió
- Dimenzióként sok változó

Összetett jelenségekre általában jellemző a **sokdimenziós** és **sokmutatós** jelleg.

Komplex mutató kialakításának gyakorlati dilemmái

- Egyáltalán szükség van-e komplex mutatóra? – kiemelt mutatószámok alkalmazása
 - Pl: orvoslátta halottak aránya (XIX. Század utolsó évtizedeiben); -Beluszky P. a számítástechnikai foglalkozásúak, illetve a kézi anyagmozgatók és szállítómunkások aránya (1980 – technikai, modernizációs szakasz megyei polarizáltsága) – Nemes Nagy J.
- Milyen változókat válogassunk be a komplex mutatóba?
 - Témában már meglévő kutatásokra támaszkodhatunk
 - Néha az élet dönt!
 - Idősoros elemzésnél – mely változók érhetőek el idősorosán
 - Mi érhető el az adott területi szinten
 - Olyan mutatót ami széles körben elérhető – ellenőrizhetővé válnak megállapításaink
- **FONTOS!** Komplex mutatón belül nem keverhető az abszolút és fajlagos adatok!

Összevonhatóvá tétel

- Komplex mutató számítás alapja a változók összevonhatóvá tétele
 - Eltérő mértékegységek
 - Eltérő volumenek
- Ehhez az adatok átalakítása, új adatok létrehozása szükséges → annak érdekében, hogy összevonhatóvá váljanak → dimenziótlanító eljárások
- Dimenziótlanító eljárásoknak köszönhetően eltűnnek:
 - Mértékegységek
 - Nagyságrendi értékkülönbségek
- Dimenziótlanító eljárások
 - Rangsorolás
 - Standardizálás
 - Adatsorok jellegadó értékeihez történő viszonyítás
 - Normalizálás: maximum-minimum intervallumra vetítés
 - Pontozás
- Csak azonos módon transzformált adatokból számítsunk komplex mutatót.

Különböző adat-átalakítások következményei:

- Bármelyik transzformációt használjuk:
 - A megfigyelési egységek sorrendje megegyezik az eredeti adatsorral
 - Minden transzformált adatsorban új érték tartozik a megfigyelési egységekhez
 - Bár az átlag és szélsőértékek értéke változik, de nem változik a helyzetük – ugyan ahhoz a megfigyelési egységhez tartoznak, mint az eredeti adatsorban
 - (Az ordinális skálát kivéve) az átalakítások lineáris transzformációt jelentenek, hiszen konstansokkal történnek a műveletek (vagyis, ha $A - B = C - D$, akkor $A' - B' = C' - D'$)
- A transzformáció után változik:
 - Az adatsor átlaga, szórása, terjedelme
- Különböző módon transzformált adatok nem vonhatók össze

Egyirányúsítás problémája

- Általában a nagyobb értékhez kötődik a pozitív tartalom
 - Pl: egy főre jutó jövedelem, személygépkocsi ellátottság
- DE! pl: Munkanélküliek aránya → egyirányúsítani kell
 - 100-munkanélküliek aránya
 - Mutató reciprokát vesszük → egy orvosra jutó állandó népesség helyett ezer lakosra jutó orvosok számát
- Van olyan mutató, amelyről csak a kutatási kontextus ismeretében tudjuk eldönteni, hogy az negatív vagy pozitív értéktartalommal bír → pl: természetes szaporodás
 - Kutatói felelősség
 - Minden esetben jól indokoljuk meg döntésünket

Súlyozás

- Komplex mutatók előállításánál gyakran felmerül → egyes változók többet nyomnak a latba
 - pl: jövedelem többet nyom a latba, mint a kiépített utak hossza
- Ez megoldatlan probléma
- Nem igazán használjuk
- Megállapítása szakértői becsléssel

Komplex mutató kiszámításának módja transzformált adatsorokból

- Rangszámok összeadása → legfeljettebb a legkisebb értékű
- Transzformált adatok egyszerű számtani átlaga → legpraktikusabb
 - Hátrány: szélsőséges értékek kiegyenlítik egymást
 - Ha valamelyik változó nulla, akkor is van komplex mutató
- Mértani átlag
 - Ha valamely mutató értéke 0, akkor a komplex mutató is az lesz
 - Ha fontos hogy valamennyi kritérium teljesüljön

Sajátos komplex mutatók

- Sajátos naturáliákra történő transzformálás
- Egy ágazat, tevékenységi terület naturáliáit teszik összehasonlíthatóvá
 - Állatállomány esetén – számosállat egység
 - Egy haszonállat átlagos élősúlyra történő átszámítása
 - Különböző energiahordozók – kőszén egyenérték
 - Földminőség – aranykorona érték
 - Ökológiai lábnyom: gazdaság erőforrás-fogyasztási és hulladékfeldolgozási szükségleteinek megadása termékeny földterületben mérve
- Humán Fejlettségi Index - HDI - ENSz

A mérési skálák rendszere

Mérési skálák	Tulajdonság	Sajátosságok	Jellemző példák
Arány x_a / x_b	Megkülönböztetés, sorrend, különbség, arány	Van elméleti minimum, azonos előjelű	Népességszám, jövedelem, utasforgalom
Intervallum $x_a - x_b$	Megkülönböztetés, sorrend, különbség	Pozitív és negatív értékek	Vándorlási különbözlet
Ordinális (sorrendi) $x_a \geq x_b$	Megkülönböztetés, sorrend	Nehezen mérhető, csak sorrendbe állítható	Sorrendek, rangok, eltérő funkcionális szintek
Nominális $x_a \neq x_b$	Megkülönböztetés	Nem számszerű	Név, születési hely, nem

Mérési skálák hierarchiája

- Mindegyik mérési skála rendelkezik az alatt lévő tulajdonságaival
- A „hierarchia csúcsán” az arányskála áll
 - Legteljesebb összehasonlításra ad lehetőséget
- Mérési skála meghatározza a matematikai-statisztikai módszereket
 - Brazil válogatott nem 63X jobb mint a magyar
 - 0 átlagú adatsort nem lehet az átlag %-ában megadni
- Többváltozós vizsgálatoknál:
 - Többféle mérési skála, de azonos mérési skálájú adatokra van szükség → adat-transzformáció

Mérési skálák transzformációja

- Leggyakrabban:
 - intervallum- vagy arányskálán mért jellemzők ordinális adatskálára átalakítása (pl. komplex mutatóknál: rangsorolás)
- Azonos értékek: rangszámok is azonosak
 - Páratlan számú (pl. 3) adat egyezése: középső rangszám (8., 9. és 10. helyett 9., 9. és 9.)
 - Páros számú (pl. 2) adat egyezése: rangszámok átlaga (4. és 5. helyett 4,5. és 4,5.)
- Nincs holtversenyben elsőség
 - 1. és 2. helyett 1,5. és 1,5 (1. és 1. helyett)

Dimenziótlanító és egyéb komplex mutatóhoz alkalmazott eljárások

1. Rangsorolás
2. Pontozáson alapuló módszer
3. Mértani átlag
4. Adatsor jellegadó (szélső-, közép- és szórás) értékeihez való viszonyítás
 - a. Átlaghoz viszonyítás
 - b. Maximumhoz viszonyítás (Bennett-féle komplex mutató)
 - c. Normalizálás
 - d. Standardizálás
5. Főkomponens- és a faktoranalízis
 - Adatsor fejlettséget-elmaradottságot milyen módon, milyen irányba fejezi ki?
 - Növekvő érték a fejlettséget, vagy elmaradottságot fejezi ki

I. Rangsorolás

- Más néven: ordinális skálára transzformálás
- Viszonylag egyszerű - adatok sorba rendezésén alapul
- Eljárás:
 - Adatsor értékeit sorba rendezzük
 - Rendezést követően 1-től n-ig sorszámot rendelünk az értékekhez
 - Azonos értékek esete:
 - Azonos „átlagolt” sorszámot használunk
 - Vesszük a rangsorban az azonos értékkel rendelkezők sorszámát, átlagoljuk és az átlagot rendeljük valamennyi azonos értékű elemhez

Komplex mutató képzés → ordinális skálára átalakított adatokat ált. egyszerű számtani átlag számítással szokták összevonni

Előnye: egyszerűsége, laikusok is könnyen használhatják

Hátránya: jelentős információvesztés → rangsorban nem tükröződik az elemek közötti különbség mértéke

2. Pontozáson alapuló módszer

- Eljárás:
 - Adatsor sorba rendezése (sok esetben az extrém értékek elhagyásával pl: főváros)
 - Terjedelem kiszámítása (max-min)
 - Terjedelmet egyenlő osztályközökre bontása (1-5; 1-10; 1-100;)
 - Területegységek pontozása attól függően melyik osztályközbe esnek
 - Területegységeket 1-től 5/10/100-ig pontozzák a változó növekvő értékétől függően
 - Fordított mutatóknál 1-től 5/10/100-ig pontozzák a változó csökkenő értékétől függően
- Komplex mutató számítása: különböző adatsorok pontozásával kapott értékek számtani átlaga (ritkábban összege) alapján számítják
- Előnye: egyszerűsége
- Hátránya: információveszteség

3. Mértani átlag

- *Önmaga adja a komplex mutató értékét*

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

4. Adatsor jellegadó értékeihez való viszonyítás

- Egyszerű
- Jellegadó értékek
 - Középértékek
 - Szélsőértékek
 - Szórás
- Legtöbbször az adatsor maximumához hasonlítják a többi értéket, esetleg az átlaghoz, minimumhoz vagy szóráshoz
- Eljárás:
 - Adatsor minden eredeti adatát a maximumérték (vagy más jellegadó érték) százalékában fejezünk ki
- Figyelemmel kell lenni az ellentétes előjelű mutatók megfordítására
→ kapott maximumra vetített érték 100%-ból történő kivonása

4.a. Átlagoz viszonyítás

- Eltérhet az értékek intervalluma
- Eltérhet az adatsorok szórása

$$Z_i = \frac{X_i}{\bar{X}} * 100$$

$$KM_j = \sum_{i=1}^n \frac{z_{ji}}{n}$$

4.b. Maximumhoz viszonyítás a Bennett-féle komplex mutatóval

- Bennett-féle komplex mutató:

$$Z_i = \frac{X_i}{X_{\max}} * 100$$

- Maximumra vetített jelzőszámok területegységenkénti átlagolására szolgál

- Maximum = 100

- Népszerű, mert az eljárás eredményeként a %-ra átalakított értékek, ill. azok átlagának értékkészlete a (0;100) intervallumba esik.

$$BEN_j = \sum_{i=1}^n \frac{z_{ji}}{n}$$

- (j=területegységek száma)

4.c Normalizálás

- Terjedelempre $(X_{max} - X_{min})$ történő vetítés
- Eljárás:

$$Z_i = \frac{X_i - X_{min}}{X_{max} - X_{min}} * 100$$

- Értéke $0 < Z_i < 100$
- Minimum = 0, maximum = 100
- Például a járások kedvezményezettségi kategóriáinak kialakításakor

4.d. Standardizálás

- Hátránya: Több lépésből áll, értéke negatív is lehet → kevésbé népszerű
- Előnye: szerényebb információveszteség az előbbiekhöz képest
- Standardizált változó kiszámítása:

$$S_x = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}$$

- Kiszámoljuk az adatsor szórását
- Majd az egyes változókból kivonjuk az adatsor átlagát
- A kivonatot elosztjuk az adatsor szórásával

$$Z_i = \frac{X_i - \bar{X}}{S_x}$$

- Standardizált változó átlaga 0, szórása 1.



Tipizálási módszerek

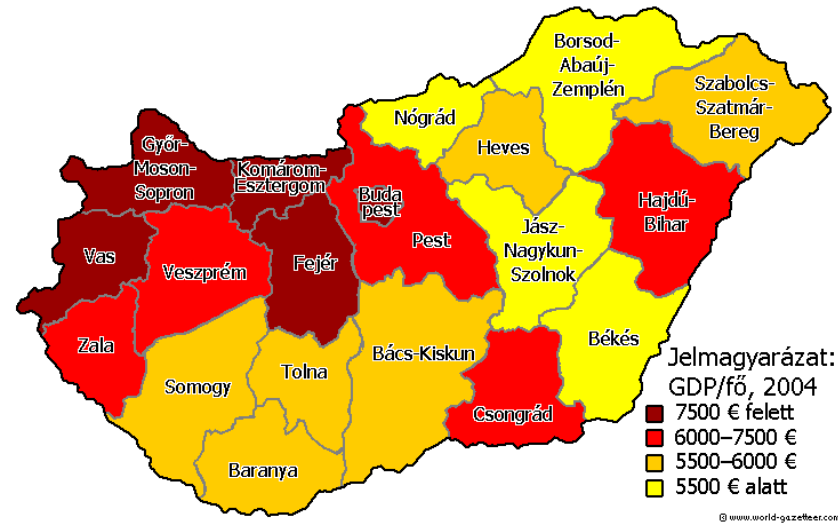
Egydimenziós osztályozás

- Azonos elemszámú csoportok (kvantilisek) kijelölése
 - Egyenlő elemszámú csoportosítás
- A jelzőszám értékei szerinti, egyenközű osztályozás
 - Egyenlő osztályközű csoportosítás
- Jellemző értékek szerinti csoportok kijelölése
 - Kitüntetett értékekhez igazodó csoportosítás (pl. átlag)
 - Kerek számokhoz igazodó csoportosítás
 - Adatsor jelentősebb töréseihez igazodó csoportosítás

Kategória határok megállapításának szempontjai

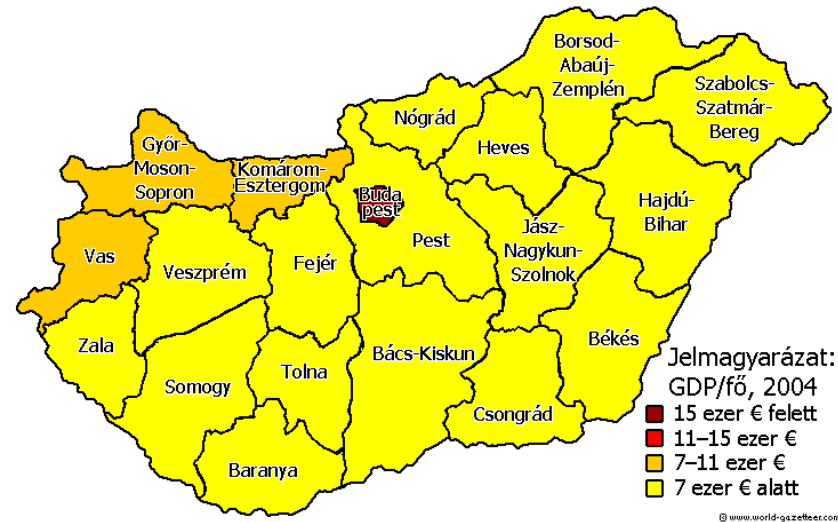
1. A csoportok nagyjából hasonló elemszámúak legyenek (ne legyen üres kategória) → egyenlő elemszámú csoportosítás (medián, tercilis stb.)
2. A csoportok nagyjából azonos értéktartományként következzenek → egyenlő osztályközű csoportosítás
3. Jelenjen meg bizonyos értelmezhető értéktartalom (pl. átlagfelettség) → kitüntetett értékekhez (pl. átlag) igazodó csoportosítás
4. Könnyen értelmezhető kategória határok → kerek számokhoz igazodó csoportosítás
5. Hasonló értékű régiókat ne válasszon ketté kategória határ → adatsor jelentősebb töréseihez igazodó csoportosítás
 - Nincs általános elfogadható recept → a fenti szempontok kombinálása szükséges

I. Egyenlő elemszámú csoportosítás



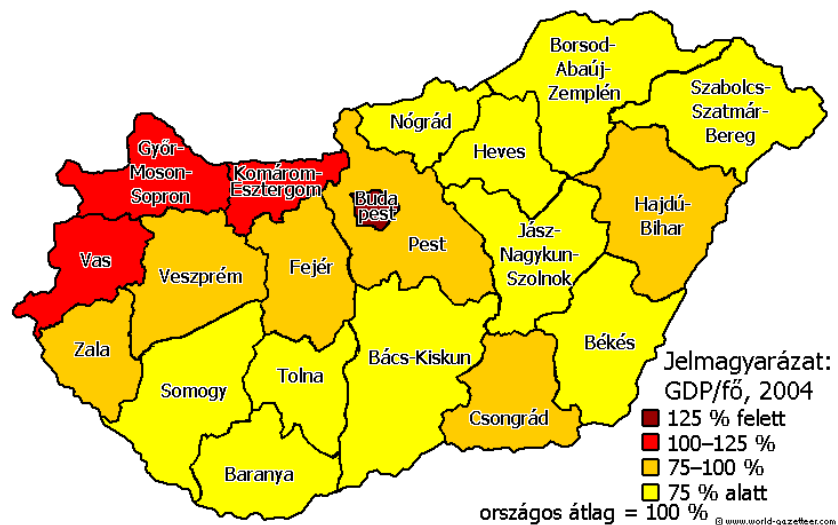
- **Előny:**
 - Minden kategóriához tartozik régió (még hozzá arányosan) → inkább nagy elemszámnál
- **Hátrány:**
 - Olykor erőltetett határok (pl. Bács-K és Borsod-A-Z között)
 - Eltérő értéktartományú kategóriák
 - Nem jelenik meg az átlag

2. Egyenlő osztályközű csoportosítás



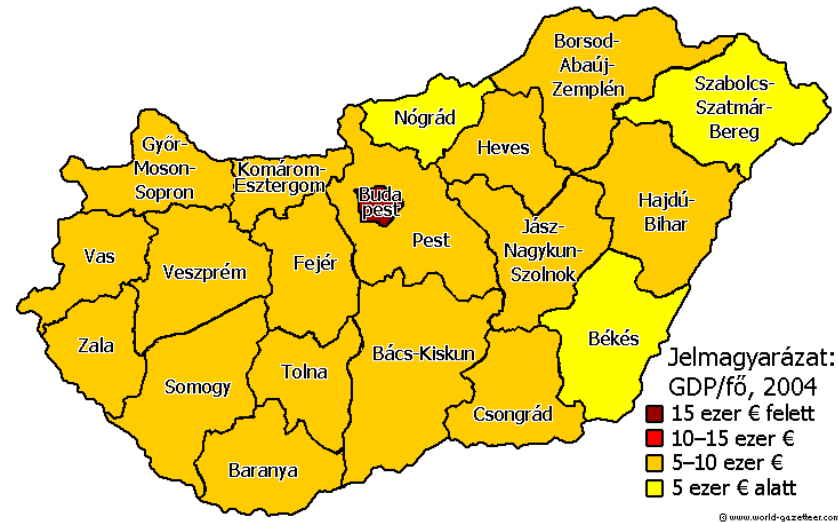
- **Előny:**
 - Kategóriák azonos intervallumonként következnek
 - Többféleképpen is meg lehet határozni a határokat
- **Hátrány:**
 - Nem minden kategóriához tartozik régió (pl. 11-15 ezer €/fő)
 - Nem jelenik meg az átlag
 - Olykor erőltetett határok

3. Kitüntetett értékekhez igazodó csoportosítás



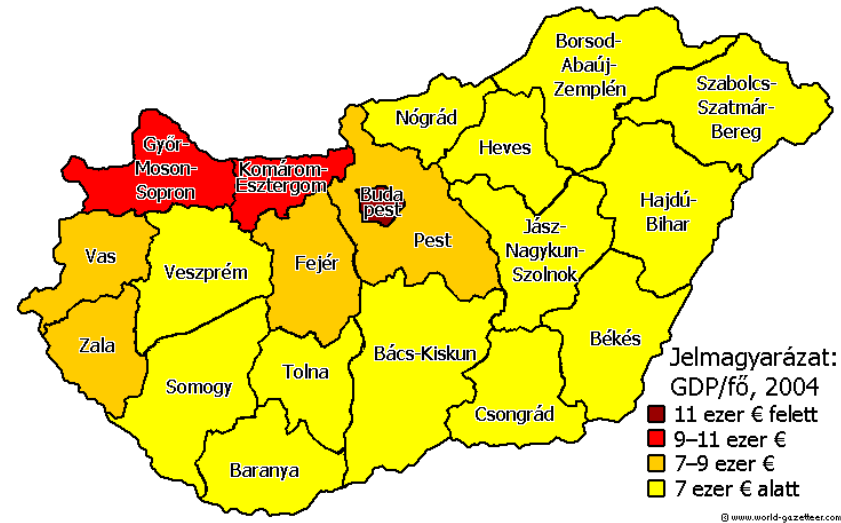
- Előny:
 - Megjelenik bizonyos értéktartalom
 - Kategóriák általában azonos intervallumonként következnek
 - Többféleképpen is meg lehet határozni a határokat
- Hátrány:
 - Olykor erőltetett határok (pl. Hajdú-Bihar és Baranya között)

4. Kerek számokhoz igazodó csoportosítás



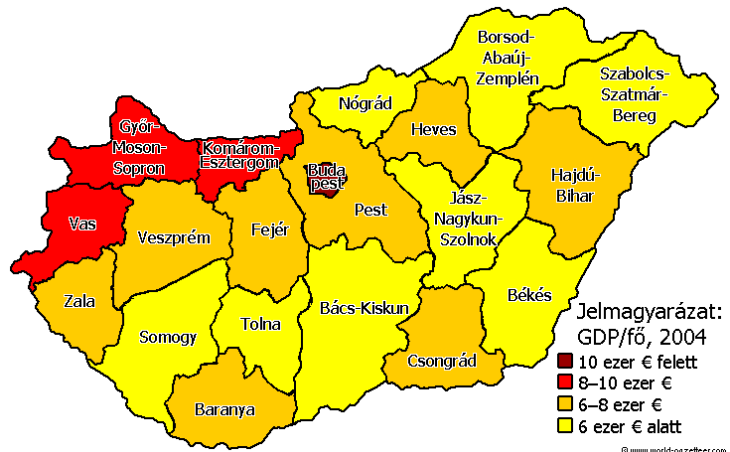
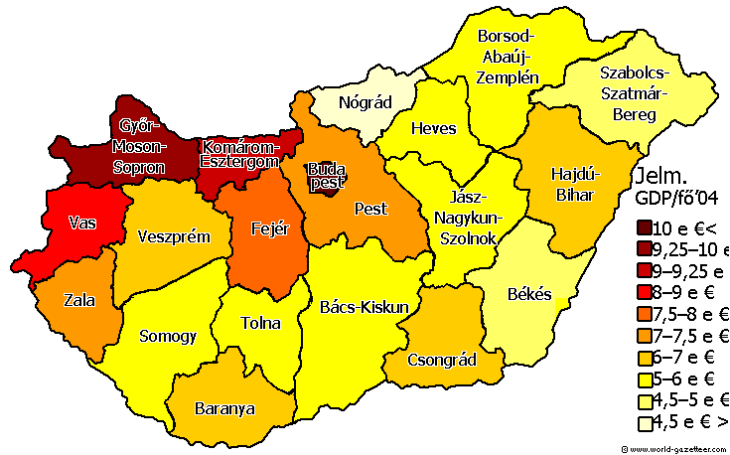
- **Előny:**
 - Kategóriák általában azonos intervallumonként következnek
 - Többféleképpen is meg lehet határozni a határokat
- **Hátrány:**
 - Nem minden kategóriához tartozik régió (pl. 10-15 ezer €/fő)
 - Olykor erőltetett határok

5. Adatsor jelentősebb töréseihez igazodó csoportosítás



- **Előny:**
 - Minden kategóriához tartozik régió (ha nem is egyenletes megoszlással)
- **Hátrány:**
 - Eltérő értéktartományú kategóriák
 - Nem jelenik meg az átlag

Csoportok számának meghatározása



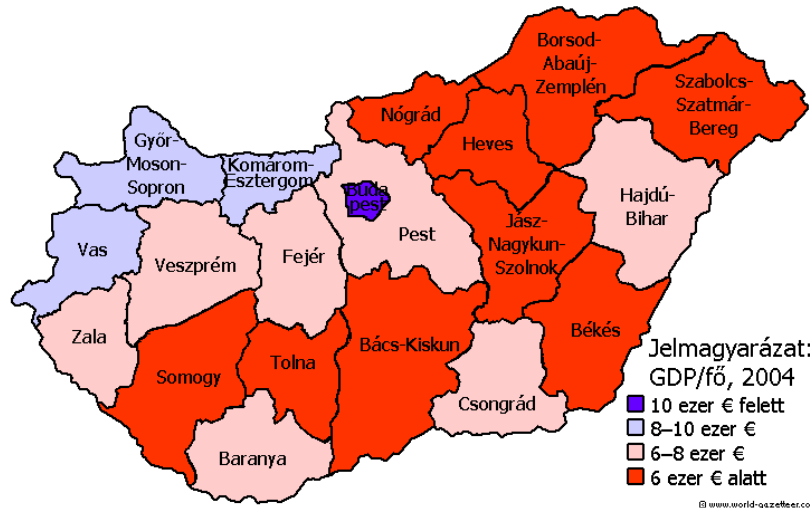
- Ne legyen túl sok csoport
 - Kevesebb színekategória lehet
 - 6–7-nél több csoport nehezen átlátható
 - Sok csoport egy elemű csoportok veszélye (főleg a nagy értékeknél)
 - Cél a modellalkotás → csoportosítás
- Csoportok száma függ az elemszámtól (területi egységek számától)
 - Csoportok száma $\approx (\text{elemszám} + 7) / 5$
 - De maximum 7 csoport!

A tematikus térképkészítés egyéb elvi szempontjai

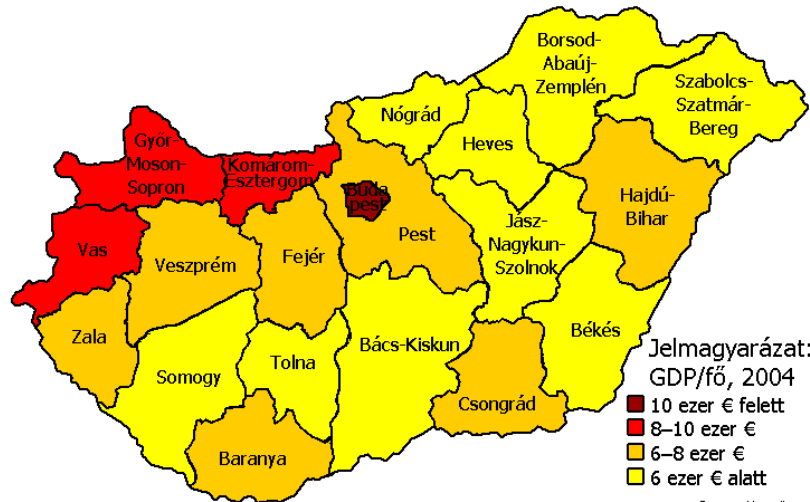
Színválasztás

- Ne legyen benne fehér, fekete
 - Fehér: nincs adat, vizsgálaton kívüli, fekete feliratok miatt
- Sokféle szín helyett egy vagy két szín színárnyalatai
- Irányultság legyen a színárnyalatok között
- Egy szín színárnyalatai:
 - Vagy sötét→világos vagy világos→sötét
 - (Sötét figyelemfelhívóbb → válság- vagy a sikertérségekre akarunk fókuszálni?)
- Két szín színárnyalatai:
 - Vagy hideg→meleg vagy meleg→hideg
 - Két dimenzió: intenzitás (sötétség), irány (szín)
 - (Csak színesben látszik → ritkábban használjuk inkább + és – értékeknél, pl. migráció)
- Jelmagyarázat mindig fel legyen tüntetve

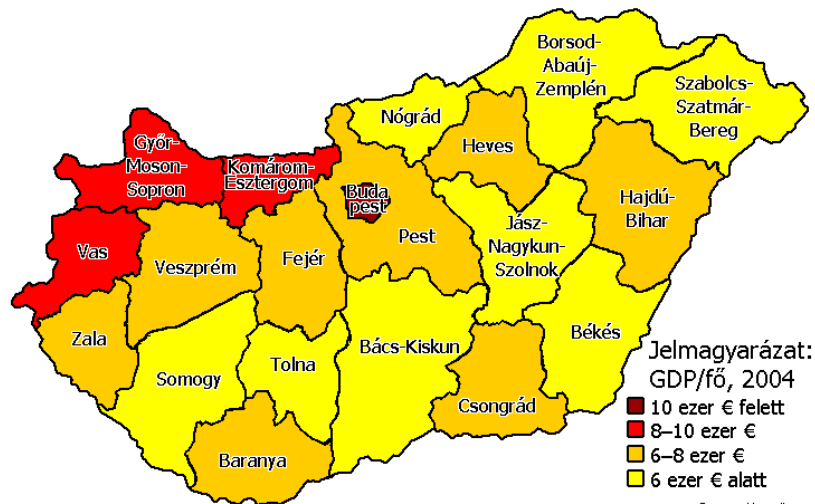
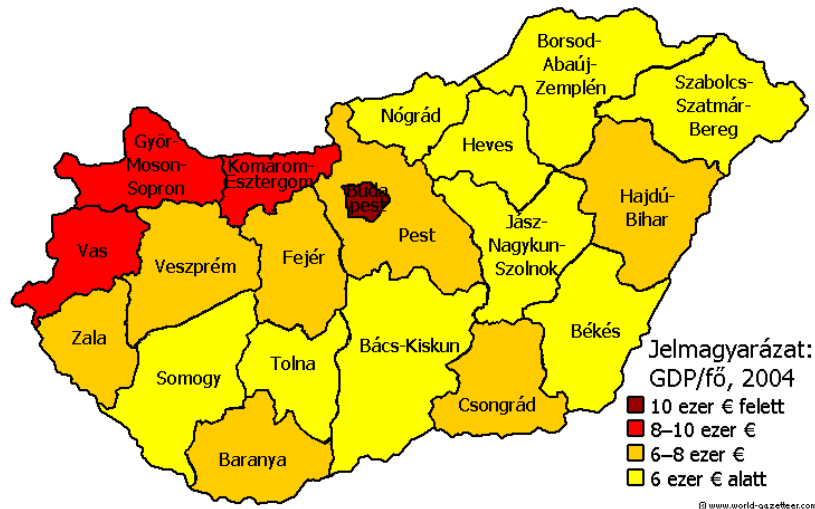
A hideg–meleg színezés hátránya



• Hibás ábra



Feliratok körüli kontúrok jelentősége



- Felirat színe
 - Feliratok színe elütő legyen minden régió esetében
 - Ne legyen közbülső színárnyalat
 - Feliratokat érdemes a színétől eltérő kontúrral körbevenni („stroke”)
 - Egyes vektoros programoknál ennek már neve van, pl. halo
- Nem fontos a felirat

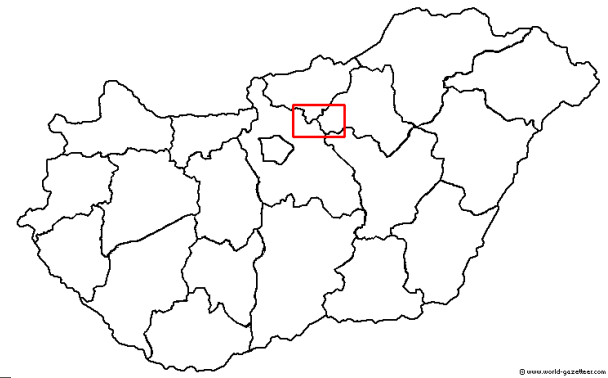
Ideális fájl típus raszteres tematikus térképekhez



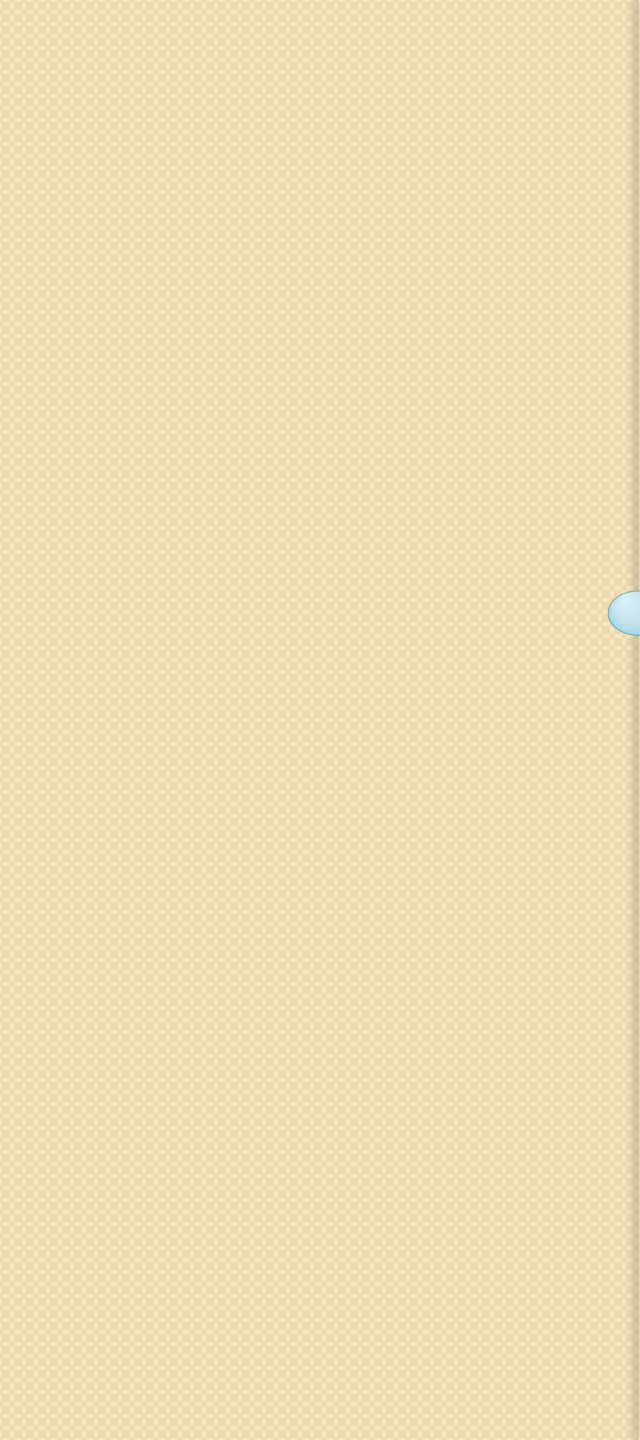
Kiterjesztés	Fájlméret
.gif	6982 Byte
.png	19145 Byte
.jpg	60766 Byte
.bmp	1016804 Byte

- Legjobb: .png, .gif
- Jó: .bmp (jó, de túl nagy méretű a fájl)
- Nem jó: .jpg (kicsi fájl, de „lekerekíti” a színhatárokat → nem alkalmas a színezésre)

A .jpg fájl hátránya



	Alaptérkép	Színezés után
.png		
.jpg		
	.jpg → sokféle fehér színárnyalat	.jpg → megoldás: színcserénél a „tolerancia” erősítése





Köszönöm a megtisztelő figyelmet!

varga.agil4@gmail.com