

KARTOGRAM

verze 1.1

autoři listu: Jakub Jaroš, Jakub Lysák

Cíle

V tomto pracovním listu se student:

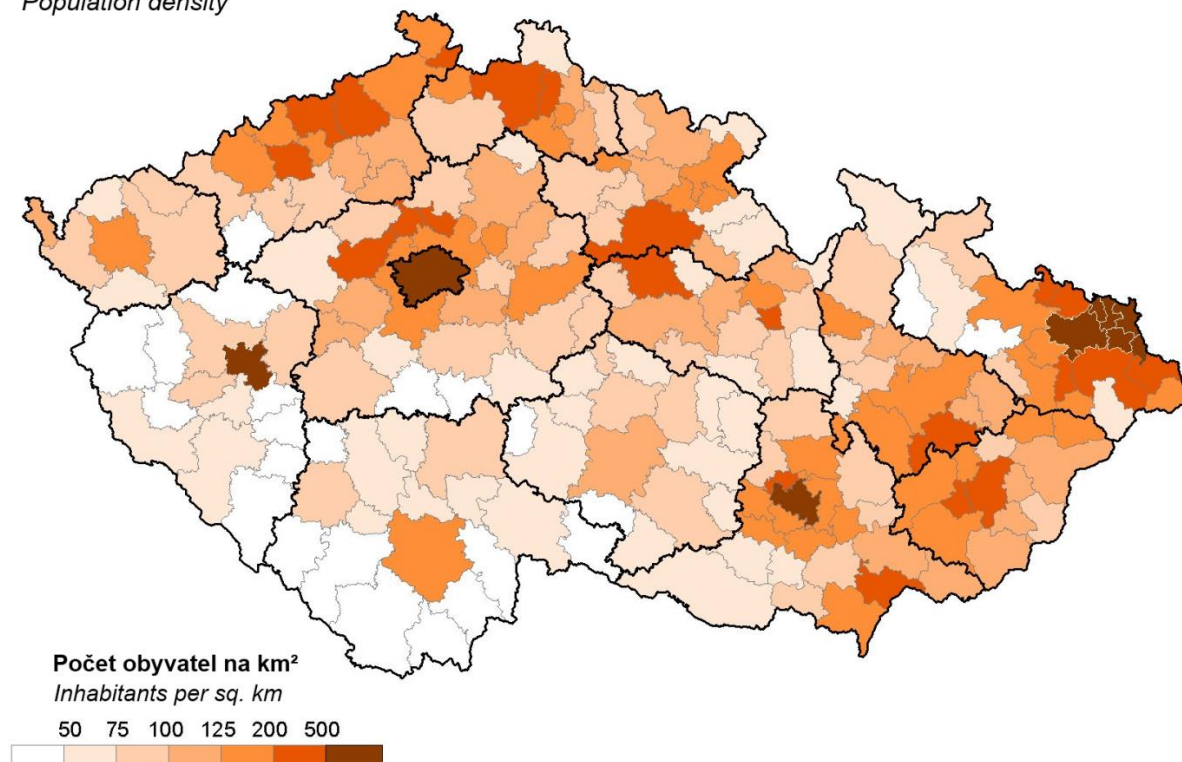
- dozví, co je kartogram a na jaká data je vhodné ho použít,
- seznámí s jednotlivými kroky tvorby kartogramu (zejména rozdělení dat do intervalů, tvorba stupnice) a hlouběji porozumí postupům, které se za těmito kroky skrývají,
- naučí se aplikovat tyto postupy na vlastní data,
- naučí se vytvořit kartogram v prostředí ArcGIS for Desktop.

Teorie

Podstatou kartogramu (angl. *choropleth map*) je znázornění jevu vyjádřeného relativními hodnotami, zachyceného za dílčí územní celky. Ukázka kartogramu je na obr. 1. Kartogram pracuje s kvantitativními daty. Správně použitá metoda kartogramu umožní srovnání jednotlivých územních celků mezi sebou a znázorní tak prostorovou proměnlivost jevu v rámci zpracovávaného území. Pro správné srovnání je klíčové, aby data byla relativní, tj. v ideálním případě přepočtená na plochu územní jednotky (tzv. pravý kartogram), akceptovatelné je i přepočítání s využitím jiné charakteristiky územní jednotky, např. na počet obyvatel (tzv. nepravý kartogram). Častou a zásadní chybou je použití této kartografické vyjadřovací metody na absolutní data. Pro ta využijeme kartodiagram. Chybná je také představa, že kartogram slouží pro zjišťování „přesných“ hodnot jevu v jednotlivých územních celcích. Tím, že jsou znázorněny intervaly hodnot a nikoliv hodnoty samotné, můžeme pouze srovnávat jednotlivé územní celky mezi sebou. Kartogram nám tedy dokáže odpovědět na otázky, např. „Je hustota zalidnění v Plzeňském kraji větší nebo menší než v kraji Jihočeském?“ či „Jaký je podíl nezaměstnaných v Pardubickém kraji oproti Kraji Vysočina?“. V žádném případě ale nejsme schopni přesně odpovědět na dotaz „Kolikaprocentní nezaměstnanost je v Ústeckém kraji?“ či „Jaký je podíl lesů na celkové ploše kraje v Libereckém kraji?“.

HUSTOTA ZALIDNĚNÍ

Population density



Obr. 1. Ukázka kartogramu. Zdroj: Český statistický úřad.

Příklady jevů, pro jejichž vyjádření na mapě je vhodné použít kartogram: hustota zalidnění za obce v Česku, podíl zemědělské půdy na rozloze SO ORP, podíl obyvatel bez náboženského vyznání na celkovém počtu obyvatel okresů, počet nakažených chřipkou na 10 000 obyvatel za kraje, volební zisk (vyjádřený v procentech) strany X v komunálních volbách za obce.

Příklady, kdy by bylo chybné použít kartogram: počet obyvatel za obce v Česku, rozloha zemědělské půdy za SO ORP, počet obyvatel bez vyznání za okresy, počet nakažených chřipkou za kraje, volební zisk (vyjádřený absolutním počtem hlasů) strany X v komunálních volbách za obce.

Proč je absolutní vyjádření dat kartogramem chybné? Je to z toho důvodu, že interpretace mapy může vést k zavádějícím výsledkům. Předpokládejme, že nás bude zajímat míra využití krajiny pro rekreaci, kterou budeme hodnotit počtem chat. Máme-li územní jednotku A o rozloze 100 km² a v ní 50 chat, je jistě míra využití krajiny pro rekreaci menší než v jiné územní jednotce B, která má 10 km² a 45 chat. Pravděpodobnost, že v jednotce B narazíme na chatu, je 9× větší (v A 0,5 chaty/km² vs. v B 4,5 chaty/km²), než že na ni narazíme v jednotce B. Pokud bychom data zobrazovali absolutně, vyšlo by nám z mapy, že jev je významnější v jednotce A. Podobně, vytváříme-li mapu kriminality za obce, měli bychom považovat za bezpečnější obec s 10 000 obyvateli, v níž došlo k 500 trestným činům, než obec s 500 obyvateli, v níž došlo k 100 trestným činům. Klíčové pro použití kartogramu je tedy rozhodnout, zda informace, kterou má mapa poskytnout, má či nemá vztah k ploše, počtu obyvatel či jiné souhrnné charakteristice daného území. Podrobnější zdůvodnění lze nalézt v [1], str. 97–98. Absolutní data může být rovněž užitečné vyjádřit, ale vyjadřují trochu jinou informaci a proto by se měla vyjadřovat jinak, např. kartodiagramem. Použití absolutních dat v kartogramu by z pohledu informace nevedlo v případě, že by všechny územní jednotky byly srovnatelné (ve smyslu údaje, kterým se jev relativizuje).

Existuje řada druhů kartogramů, ucelený přehled lze nalézt v [1] na str. 132–136. Zde se omezíme pouze na nejběžnější a nejčastěji používané druhy:

- **jednoduchý kartogram:** zobrazuje pouze jeden relativní jev, a to změnou barvy nebo rastru,
- **složený kartogram:** zobrazuje hodnoty dvou nebo více jevů, umožňuje jejich vzájemné srovnání, typicky je jeden jev vyjádřen barvou, druhý rastrem,
- **kvalifikační kartogram:** znázorňuje rozdíl jevu od zvolené střední hodnoty S , pro oblasti s hodnotou jevu větší než S se volí odstíny barvy opačného charakteru než pro jevy s hodnotu menší než S (např. odstíny červené pro jevy nad S , odstíny modré pro jevy pod S).

Při tvorbě kartogramu je třeba vyřešit tři hlavní úkoly:

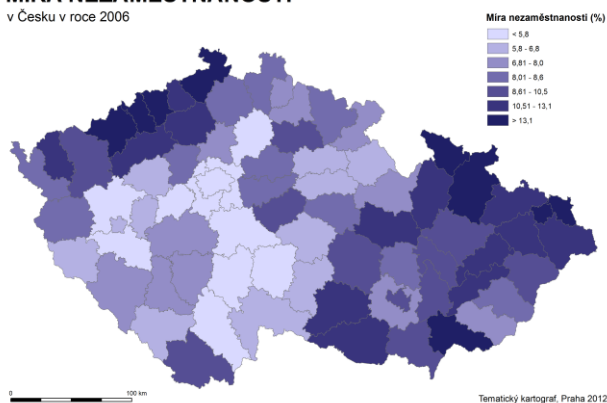
1. vytvořit intervalovou stupnici, tj. rozdělit vstupní data do intervalů (ty budou tvořit jednotlivé kategorie znázorněné v mapě),
2. graficky vyřešit jejich znázornění v mapě (obvykle pomocí barevné stupnice či rastru),
3. navrhnout správnou legendu.

Při **tvorbě intervalové stupnice** (klasifikaci vstupních dat) dat je třeba stanovit počet intervalů a jejich meze (hranice mezi intervaly). V kartografické literatuře (např. [1], str. 100) lze nalézt doporučené vzorce pro počet intervalů vycházející pouze z počtu územních jednotek, např. $m \approx y^{0,5}$, $m \leq 5 \log y$ či $m \approx 1 + 3,3 \log y$, kde m je počet intervalů a y je počet statistických jednotek v souboru (počet územních celků). V praxi se doporučuje použít 4–10 intervalů. Malé množství znamená velmi hrubou informaci o rozložení jevu v území, velké množství naopak činí mapu nepřehlednou, zejména z toho důvodu, že je pak pro čtenáře mapy obtížné rozlišit jednotlivé kategorie, mezi jejichž barvou či rastrem jsou pak jen velmi nepatrné rozdíly.

Obvykle se ovšem počet intervalů stanovuje zároveň s jejich mezemi na základě hlubší analýzy dat. Klíčovou roli při tom hraje histogram vytvořený z dat, obvykle reprezentovaný jako graf, kde na ose x jsou hodnoty jevu a na ose y počet výskytu jevu dané hodnoty (proložením křivky dostaneme rozložení četností jevu). Meze intervalů se stanovují na základě analýzy histogramu. V kartografické literatuře (např. v [1] na str. 100) lze nalézt zásady pro vhodné meze na základě typu statistického rozdělení, které data mají (a které je histogramem vyjádřeno). Geografická data mají často charakter tzv. vícevrcholového rozdělení. V tomto případě se hranice intervalů stanovují tak, že sledují lokální minima histogramu. Speciálně je pak třeba přistoupit ke znázornění extrémních hodnot, které obvykle vytvoří samostatnou kategorii. Více o extrémech lze nalézt v [1] na str. 102–109. V praxi se pak „teoreticky“ spočítané hodnoty mezí (ve smyslu správné klasifikace do skupin z hlediska zachování informace) ještě drobně upravují, typicky zaokrouhlením na dekadické hodnoty ($1 \cdot 10^x$, $5 \cdot 10^x$ apod., kde x je celé číslo). To je ale třeba dělat poměrně opatrně, protože každá větší změna meze obvykle znamená zmenšení množství informace reprezentované mapou. Zcela chybný je postup, kdy se stanoví meze pouze na základě znalosti minima a maxima a rozdělení je dáno „hezkými“ okrouhlými hodnotami v tomto rozmezí bez analýzy histogramu.

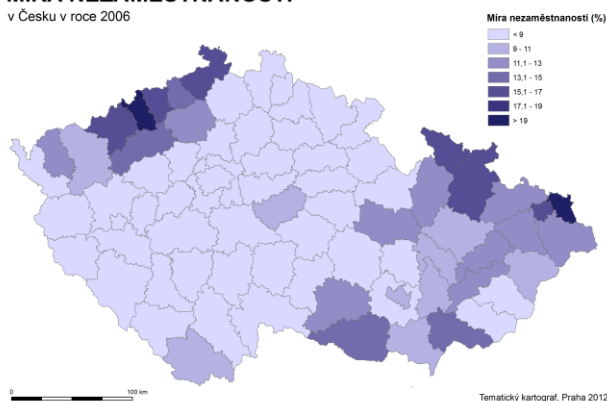
Pokud provedeme rozdělení dat do skupin chybně, může dojít k zavádějící až mylné interpretaci obsahu mapy, viz ukázky na obr. 2a – 2d. Všechny tyto mapy byly vytvořeny na základě stejných dat.

MÍRA NEZAMĚSTNANOSTI v Česku v roce 2006



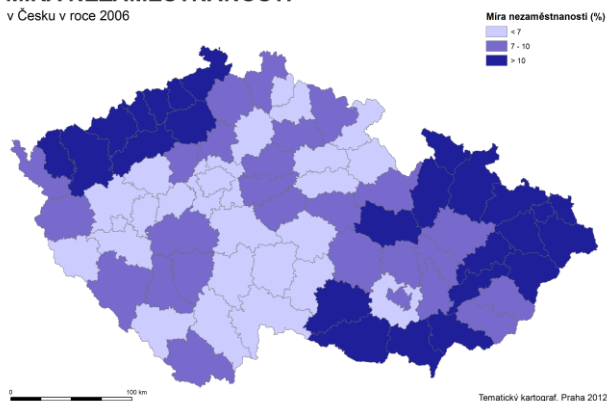
Obr. 2a. Správné řešení: dostatečný počet intervalů, správné rozdělení dat do intervalů.

MÍRA NEZAMĚSTNANOSTI v Česku v roce 2006



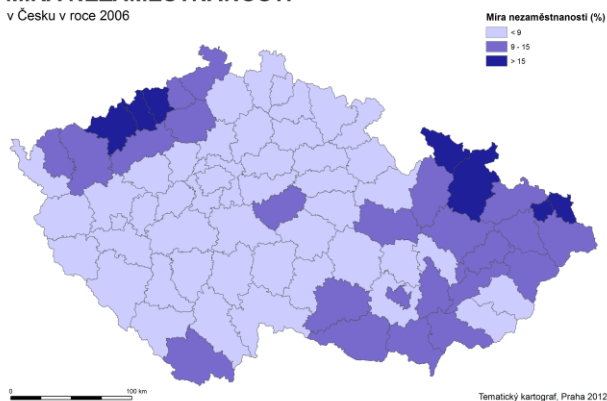
Obr. 2b. Chybné řešení: dostatečný počet intervalů, ale chybné rozdělení dat do intervalů.

MÍRA NEZAMĚSTNANOSTI v Česku v roce 2006



Obr. 2c. Chybné řešení: příliš nízký počet intervalů. Rozdělení dat do intervalů je relativně správné. Dva, případně tři intervaly, lze akceptovat pouze ve zvláštních případech, kdy chceme znázornit např. odchylky od průměrné hodnoty (podprůměrné, průměrné a nadprůměrné regiony).

MÍRA NEZAMĚSTNANOSTI v Česku v roce 2006



Obr. 2d. Chybné řešení: jednak malý počet intervalů, jednak chybné rozdělení dat do intervalů.

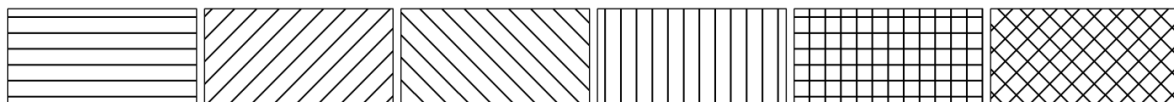
Máme-li data rozdělena do skupin, můžeme přistoupit k vlastnímu **kartografickému zpracování**. To znamená, že příslušné územní jednotky, které patří do stejné skupiny, vyplníme stejnou barvou či stejným rastrem. Z mapy by měla být dobře patrná prostorová diferenciací (prostorový vzorec) jevu v tom smyslu, že je na první pohled vidět, kde je jevu nejméně a kde nejvíce (tomu odpovídá intenzita barvy a rastru). Obecně musí platit, že jednotlivé kategorie by měly být dobře rozlišitelné, tj. mezi barvami, resp. rastry pro jednotlivé kategorie by měl být dostatečný rozdíl.

Při použití barvy pro tvorbu kartogramu bychom měli dodržet následující zásady:

- vybrat vhodnou základní barvu nebo barvy (tón barvy), který je asociativní ve vztahu ke znázorňovanému jevu,
- na základě tohoto tónu vytvořit barevnou stupnici. V případě jedné barvy měníme jas a sytost, v případě více barev se tvoří systematický barevný přechod (např. žlutá – hnědá),
- správně vytvořená barevná stupnice netrpí propadáním barev, tj. poklesu jasu nebo sytosti při současném růstu intenzity jevu,
- u kvalifikačního kartogramu platí totéž s tím, že stupnice jsou dvě, jedna pro jevy nad zvolenou střední hodnotu, druhá pro jevy pod zvolenou střední hodnotou, zvolené základní barvy jsou opačného charakteru.

Při použití liniového rastru („výšrafování“ ploch) pracujeme se šířkou linií, jejich rozestupem, orientací a strukturou. Měly by se respektovat tyto zásady:

- žádná z ploch by neměla zůstat nevyplněná ani zcela vyplněná,
- s narůstající intenzitou jevu roste šířka linie,
- s narůstající intenzitou jevu se rozestup mezi liniemi zmenšuje,
- existuje doporučené pořadí použití směru linií v závislosti na narůstající intenzitě jevu, které je uvedeno na obrázku 3, více viz [3] str. 69. U šikmých linií je vhodné použít sklon $\pm 45^\circ$. Je to z toho důvodu, aby byl rastr jednotný a u složeného rastru byly linie na sebe kolmé. Nakolik je reprezentování hodnot pouhou změnou směru dostatečně názorné i pro uživatele-nekartografa, je značně diskutabilní, a proto je vhodnější zároveň se směrem měnit i jiný, názornější parametr,
- s narůstající intenzitou jevu se struktura linie mění z přerušované na plnou, struktura linií rastru by se měla používat opatrně a až v situaci, kdy změnou ostatních parametrů nelze dosáhnout kýženého efektu.
- v případě složeného rastru je vhodné, aby rozestup linií v obou směrech byl stejný (linie tedy rozdělují plochu na čtverce, nikoliv obdélníky). Obdélníky se totiž používají pro vztahový (srovnávací) kartogram, ve kterém se jeden jev zachycuje liniovým rastrem orientovaným jedním směrem, druhý jev liniovým rastrem kolmým tento směr a výsledná interakce rastrů (dělicí plochu na obdélníky) vypovídá o souvislosti jevů.

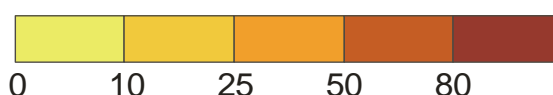


Obr. 3. Doporučené pořadí směru linií v liniovém rastru podle [3]. Intenzita jevu narůstá zleva doprava.

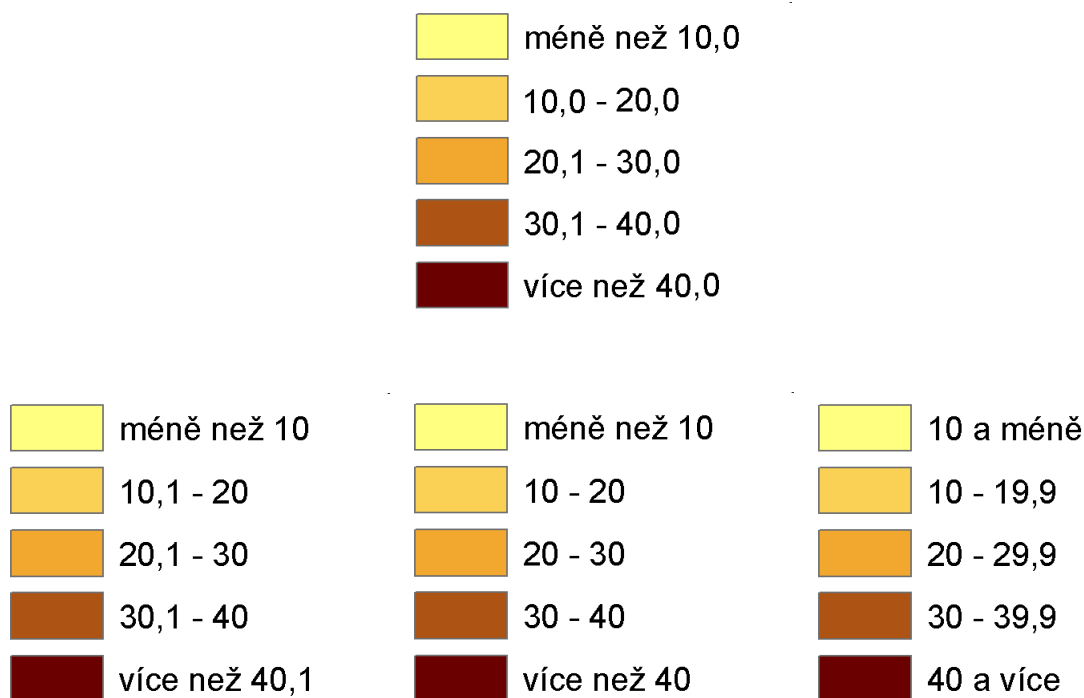
V případě kvalifikačního kartogramu má smysl pracovat i s barvou linií tvořících rastr (linie vyplňující oblasti s hodnotou jevu nad zvolenou střední hodnotou budou jednou barvou, linie vyplňující oblasti s hodnotou jevu pod zvolenou střední hodnotou budou druhou barvou, přičemž zvolené barvy jsou opačného charakteru).

Důležitou fází je rovněž **zpracování legendy**. Protože v případě správného použití kartogramu na kvantitativní relativní data zobrazovaný jev představuje reálné (tj. neceločíselné) hodnoty, je přirozeným řešením znázornit stupnici a meze intervalů [2], viz obr. 4. V praxi se lze ovšem setkat s tím, že se jevy rozepisují podle intervalů (viz obrázek 5). Při rozepisování jevů podle intervalů je třeba dbát na to, aby legenda byla jednoznačná. To znamená, aby libovolná hodnota patřila právě do jednoho z intervalů a zároveň byl těmito hodnotami pokryt celý obor hodnot jevu. Hlavní nevýhodou tohoto přístupu je nutnost uvádět intervaly ve tvaru *A až B – epsilon*, *B až C – epsilon* atd., kde *epsilon* je malá hodnota (např. 0,01). V případě použití rastru je u legendy důležité zohlednit to, že v každé poloze legendy (a zároveň i v každém územním celku zobrazeném v mapě) musí být alespoň dvě linie rastru (aby byl patrný jejich rozestup). Pokud tato podmínka není splněna, je třeba položku u legendy rozšířit, v případě mapy zvětšit měřítko nebo nastavit menší rozestup linií rastru.

Podíl orné půdy na celkové rozloze regionu [%]



Obr. 4. Doporučená podoba legendy kartogramu pro spojité jevy.



Obr. 5. Správné (v prvním řádku) a tři chybné (ve druhém řádku) legendy ke kartogramu.

Předpokládáme, že jev na obr. 5 má hodnotu uvedenou na jedno desetinné místo; i s ohledem na to je vhodné důsledně uvádět v legendě všechny hodnoty na stejný počet desetinných míst (a číslice za desetinnou čárkou důsledně uvádět). Protože jde v legendě o výrazy určující rozsah, je typograficky správné použít pomlčku (10–20) a nikoliv spojovník (10-20). V případě velmi dlouhých čísel lze zvážit přidání mezer okolo pomlčky (1200,000 – 1500,000) pro lepší přehlednost (standardně se mezery okolo pomlčky ve významu od–do nepíší). V prvním případě hodnota 10 nepatří do žádného z intervalů (šlo by opravit formulací 10 a méně), ve druhém případě se intervaly překrývají (hodnoty 20 a 30 patří do dvou intervalů zároveň), ve třetím případě hodnota 10 patří do dvou intervalů (šlo by opravit formulací méně než 10), uvedení hodnot typu $x9,9$ představuje z hlediska názornosti horší (byť formálně zcela správné) řešení než hodnot typu $x0,1$. Více viz vysvětlující text níže.

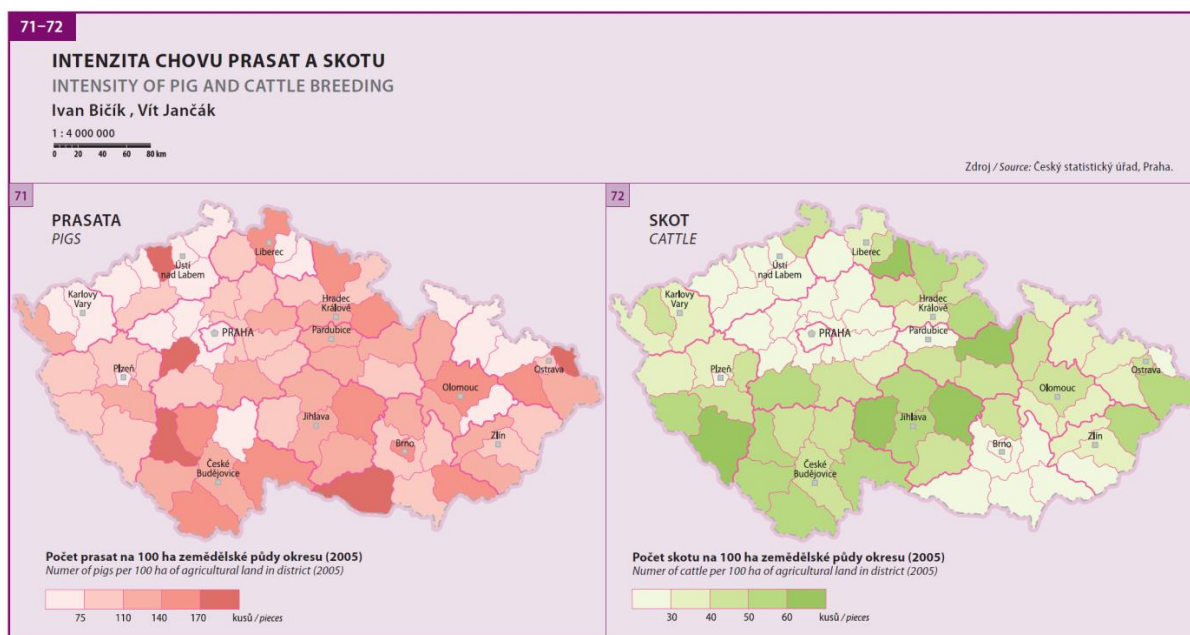
Jednoznačnost přiřazení do intervalů a požadavek na to, aby byl celý obor hodnot jevu intervaly pokryt, představuje pro jev, který může nabývat libovolné reálné hodnoty, problém. Ve smyslu značení z předchozího odstavce např. pro hodnoty v intervalu $(B - \epsilon, B)$ není pravda, že by patřily do některého z intervalů. To je nelogické z pohledu skutečnosti, že jev může být reálné číslo, v praxi to ovšem nevadí, neboť jsou sice mezi intervaly mezery, ale žádná ze znázorněných hodnot v žádné z nich neleží a další věcí je přesnost měření, kdy jevy znázorněné kartogramem dává smysl počítat na omezený počet desetinných míst (v tom případě už jde o diskrétní hodnoty a pokud je ϵ menší než počet desetinných míst, na které je měření prováděno, každé měření padne do nějakého intervalu). Jinou možností (byť v praxi nepříliš využívanou) je přímé uvedení intervalů s využitím symbolů pro otevřený/uzavřený. To ovšem klade určité nároky na čtenáře mapy. V některých oborech se sice toto předpokládá implicitně (např. intervaly uzavřené zleva), nicméně opakování meze v legendě je z pohledu kartografické správnosti chyba. Nejjednodušším způsobem, jak se těchto problémů vyvarovat, je použití již zmiňované stupnice.

Pro první a poslední položku v legendě můžeme buď použít výrazy typu „méně než X“, „X a méně“, „více než X“, „X a více“ apod., nebo uvést skutečný rozsah hodnot jevu. Má-li to smysl, pak dáváme přednost takovému stanovení mezí intervalů, které nese více informace: „10,0 a méně“ vs. „6,3–10,0“. Z rozsahu prvního intervalu vyplývá, že jev může nabývat hodnot i např. 3,1. Pokud ale nejnížší hodnota v tomtéž vstupním souboru dat je 6,3, pak použitím čísla 6,3 jako dolní meze intervalu upřesňujeme (snižujeme míru neurčitosti) rozmezí hodnot, kterých může jev nabývat. Obdobně lze postupovat při výskytu jediné odlehlé hodnoty ve vstupním datovém souboru, což je u geografických dat poměrně častý případ: např. „150,1–300,0“ a „300,1 a více“ vs. „150,1 – 300,0“ a „554,8“ (poslední hodnota není interval, ale jednoprvková množina).

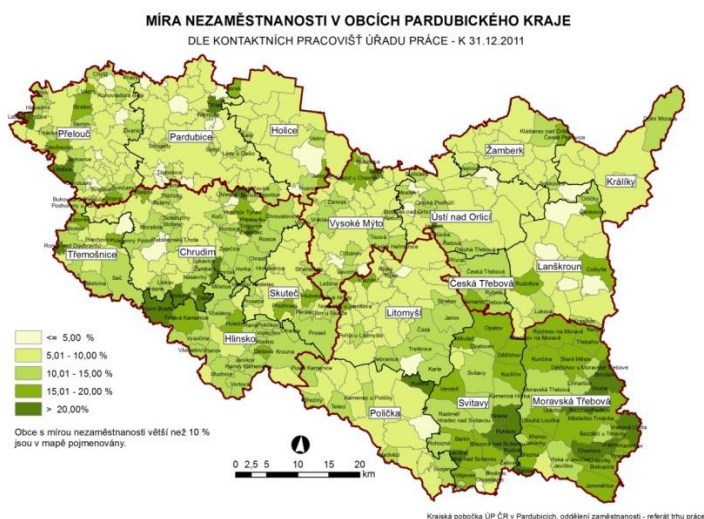
Jednotlivé položky legendy jsou jak v případě použití barvy, tak v případě použití rastru, ohraničeny linií s vlastnostmi (barva, šířka atd.), které odpovídají linii použité v mapě pro hranici mezi jednotlivými územními jednotkami.

Praktické aplikace

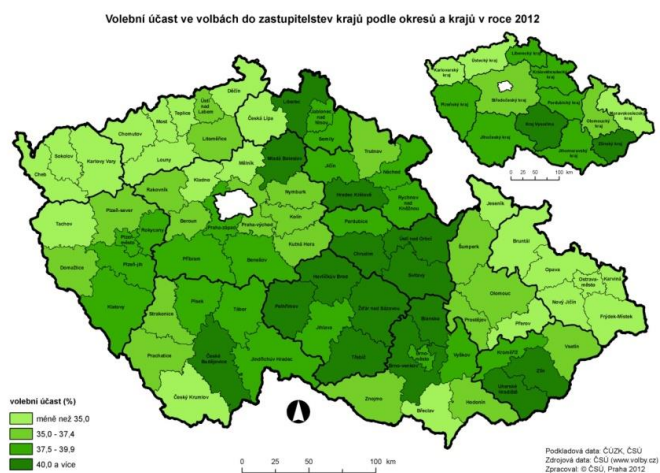
Kartogram je běžný a často používaný kartografický vyjadřovací prostředek, srozumitelný i laické veřejnosti. S kartogramy se lze setkat zejména na mapách socioekonomických jevů, jež jsou velmi často vázány či statisticky vyhodnocovány v rámci administrativních jednotek. V menší míře jsou kartogramy využívány i pro vyjadřování fyzicko-geografických jevů. Kartogramy se používají samostatně i v tematicky zaměřených atlasech. Z nepřeberného množství příkladů je několik uvedeno níže. Pozor však na to, že ne vždy je kvalita map zcela ideální.



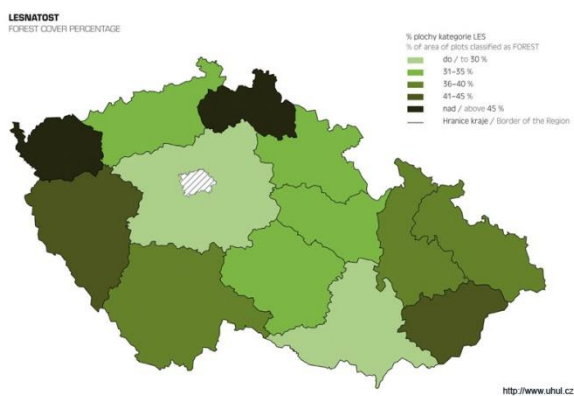
Zdroj: Atlas krajiny České republiky (2009). ISBN 978-80-85116-59-5.



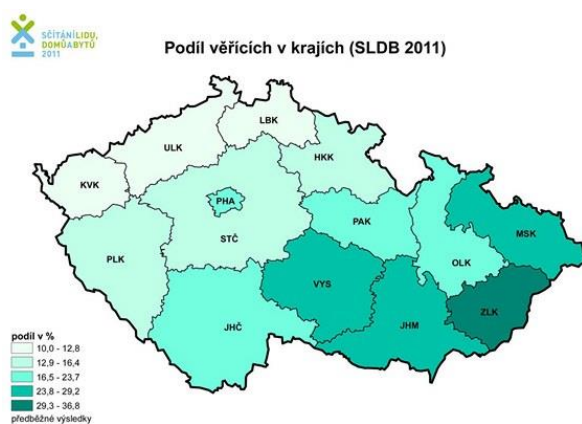
Zdroj: Ministerstvo práce a sociálních věcí.



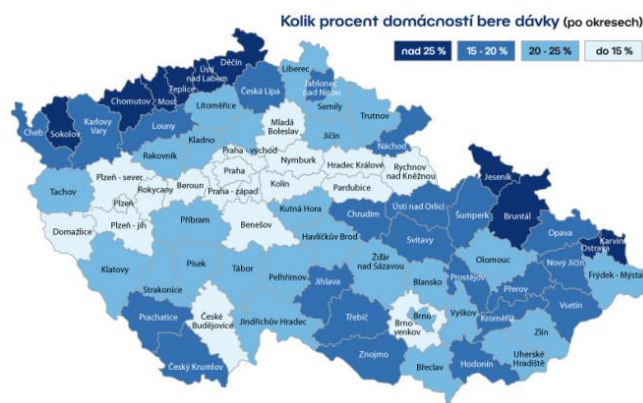
Zdroj: Český statistický úřad.



Zdroj: Ústav pro hospodářskou údržbu lesů (ÚHÚL).



Zdroj: zpravy.idnes.cz.



Zdroj: tn.nova.cz.

Praktická realizace v ArcGIS for Desktop

V následujícím textu si podrobně popíšeme tvorbu kartogramu v software ArcGIS for Desktop 10.2 (dále jen ArcGIS). Celý popis je zacílen pouze na tvorbu kartogramu a nikoli obecně na tvorbu mapy v programu ArcGIS. Dílčí kroky, které přímo nesouvisí s problematikou kartogramu (generalizace dat, tvorba kompozičních prvků apod.), zde proto nejsou podrobněji rozebrány.

Praktický postup je předveden na modelové úloze, jejímž cílem je vytvořit **složený kartogram** vyjadřující:

1. podíl lesní půdy na celkové ploše SO ORP v Moravskoslezském kraji,
2. podíl zemědělské půdy na celkové ploše SO ORP v Moravskoslezském kraji.

Použitá data *data_ms_orp* uložená v geodatabázi *kartogram.gdb* jsou výřezem z databáze ArcČR 500, verze 3.1. Atributy týkající se využití půdy (les_p, zem_p) byly doplněny z Městské a obecní statistiky ČSÚ.

Mapa bude tvořena v měřítku 1 : 550 000, které je pro ukázkový kartogram stále dostatečně podrobné a zároveň nám umožní tisk mapy na papír formátu A4. Měřítko i další vlastnosti datového rámce (*Data Frame*, v dalším textu budeme z důvodu jednoznačnosti používat zažitý anglický termín) si nastavíme ještě před započítím prací na kartogramu.

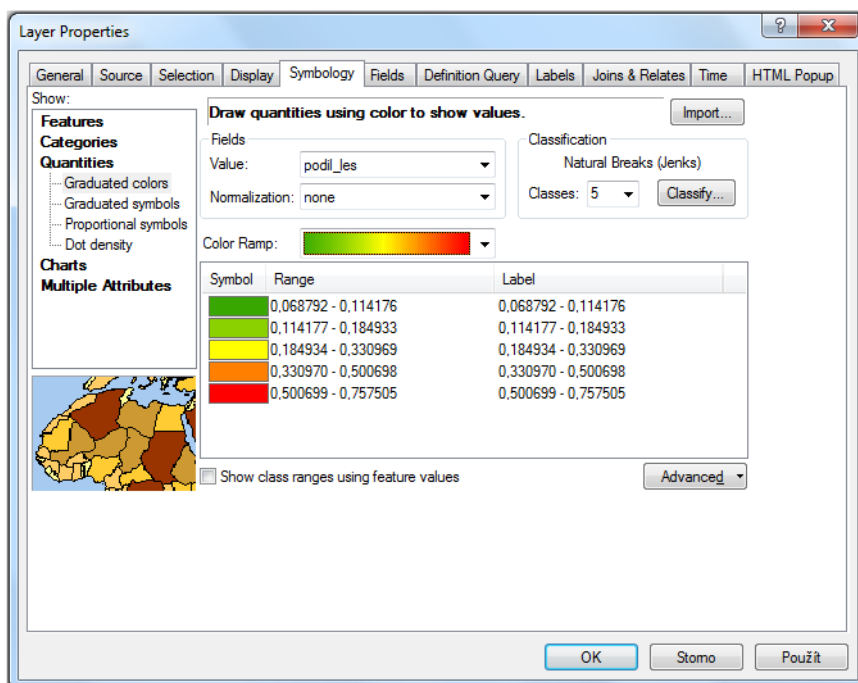
Nastavení vlastností *Data Frame*:

- měřítko: 1 : 550 000
- kartografické zobrazení: upravené Albersovo kuželové
- referenční měřítko: 1 : 550 000
- ohraničení: bez ohraničení (*Border: <None>*)

Dále nastavíme také velikost a orientaci papíru a pro usnadnění práce na více počítačích také relativní cesty k našim souborům.

Kartogram je z pohledu software ArcGIS způsob, jak vizualizovat kvantitativní data. Nástroj pro jeho tvorbu je zařazen na kartě *Symbology*, která je součástí dialogového okna vlastností datové vrstvy (*Layer Properties*). Protože mají být vizualizovaná kvantitativní data, zvolíme z nabídky „Show“ v levé části okna položku *Quantities* a následně podpoložku *Graduated Colors*. Tento název může

vzbuzovat dojem, že se touto cestou vytváří pouze kartogramy s využitím kvantitativních barev, pravdou ale je, že stejnými nástroji se vytváří taktéž rastrový kartogram. Kromě úvodních úprav dat realizovaných v atributové tabulce budou všechny další kroky prováděny skrze okno *Graduated Colors*. Veškerá nastavení, provedená v tomto okně se projeví na datech uložených v atributové tabulce ve sloupci, jehož název uvedeme do pole *Value*, viz obr. 6.



Obr. 6. Nabídka *Graduated Colors*.

Příprava dat

Modelová data *data_ms_orp* načteme do programu ArcGIS a otevřeme jejich atributovou tabulku. Data popisující využití půdy se nacházejí ve sloupcích *les_p* (rozloha lesní půdy) a *zem_p* (rozloha zemědělské půdy). V obou případech je rozloha vyjádřena v metrech čtverečních, jedná se tedy o absolutní data. Abychom mohli sestavit kartogram a také abychom splnili zadání modelové úlohy, přepočítáme data z obou sloupců na celkovou plochu SO ORP.

V atributové tabulce si nejprve vytvoříme nový sloupec „rozloha“ (datový typ Double) a do tohoto sloupce za pomoci nástroje *Calculate Geometry* vypočteme plochy jednotlivých SO ORP. Dále si vytvoříme dva nové sloupce *podil_les* a *podil_zem*, do kterých pomocí nástroje *Field Calculator* vypočteme podíly lesní a zemědělské půdy na celkových plochách SO ORP. Tímto postupem jsme získali relativní data vztažená k ploše SO ORP, která je již možné znázornit kartogramem.

Rozdělení dat do intervalů (klasifikace dat)

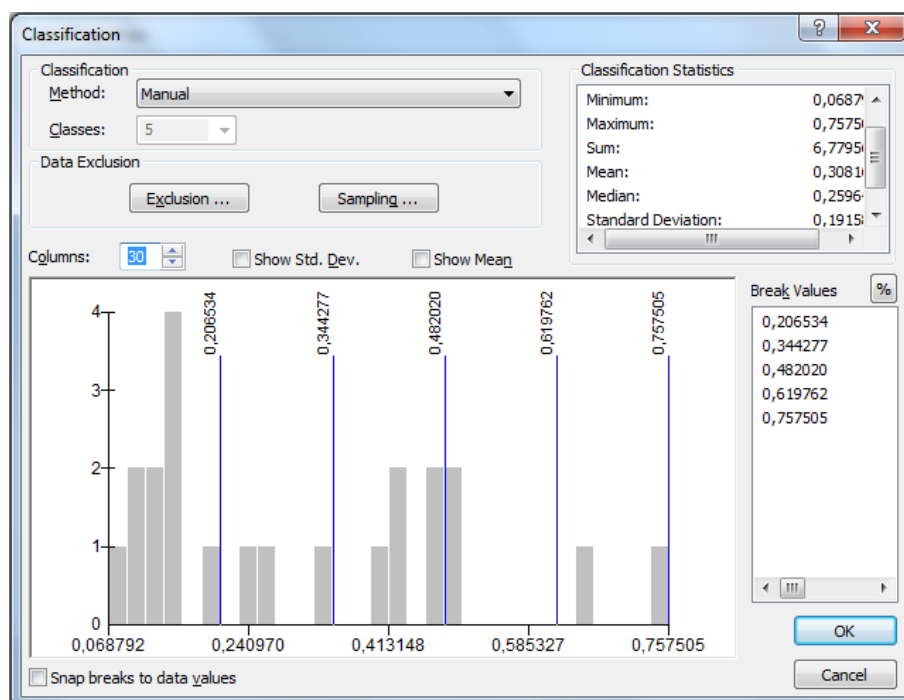
Dříve, než můžeme data rozdělit do intervalů, je třeba stanovit, jaký počet intervalů požadujeme. Tento počet úzce souvisí s celkovým počtem znázorňovaných územních jednotek, v našem případě SO ORP. Z atributové tabulky můžeme vyčíst, že celé naše zájmové území je složeno z 22 SO ORP.

Pokusíme se tedy do vzorců uvedených v teoretické části za y dosadit hodnotu 22. Výsledný počet intervalů je značen m .

| vzorec | hodnota m pro $y = 22$ |
|----------------------------|--------------------------|
| $m \approx y^{0,5}$ | 4,69 |
| $m \leq 5 \log y$ | 6,71 |
| $m \approx 1 + 3,3 \log y$ | 5,42 |

Ze vzorců je patrné, že vhodný počet intervalů by mohl být pět až sedm. Zde je na místě uvést, že počet intervalů získaný takovýmto výpočtem je třeba chápat pouze jako doporučení a nikoliv jako striktní předpis. Obzvláště budeme-li tvořit kartogram pro velké množství územních jednotek, je vhodnější držet se poučky nevytvářet více než deset intervalů, bez ohledu na výsledek výpočtů. Pro naší modelovou úlohu je přijatelný počet intervalů pět a šest. Finální rozhodnutí tedy záleží na účelu mapy, zda chceme mapu spíše podrobnější či spíše přehlednější a lépe čitelnou. Dále je důležité zvážit počet intervalů ve vztahu ke statistickému rozdělení znázorňovaných dat. Prozatím zvolme přehlednější variantu mapy s pěti intervaly.

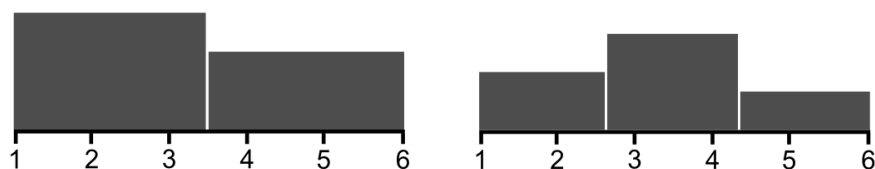
Jakým způsobem data do intervalů rozdělit, je dáno jejich statistickým rozdělením. V ArcGIS lze statistické rozdělení zkoumat prostřednictvím histogramu. Histogram je graf, který zachycuje četnosti výskytu jednotlivých hodnot, tj. na ose x jsou znázorněny hodnoty a na ose y počet výskytů dané hodnoty x . Histogram lze zobrazit stisknutím tlačítka *Classify* v okně nabídky *Graduated Colors*. Před zobrazením histogramu ještě zkontrolujeme, zda máme jako parametr *Value* zadanou hodnotu *podíl_les*, neboť pro tato data chceme histogram zobrazit. Histogram zachycující rozdělení dat *podíl_les* je zachycen na obrázku 7. Ve stejném okně se bude následně provádět také rozdělení dat do intervalů.



Obr. 7. Histogram v ArcGIS.

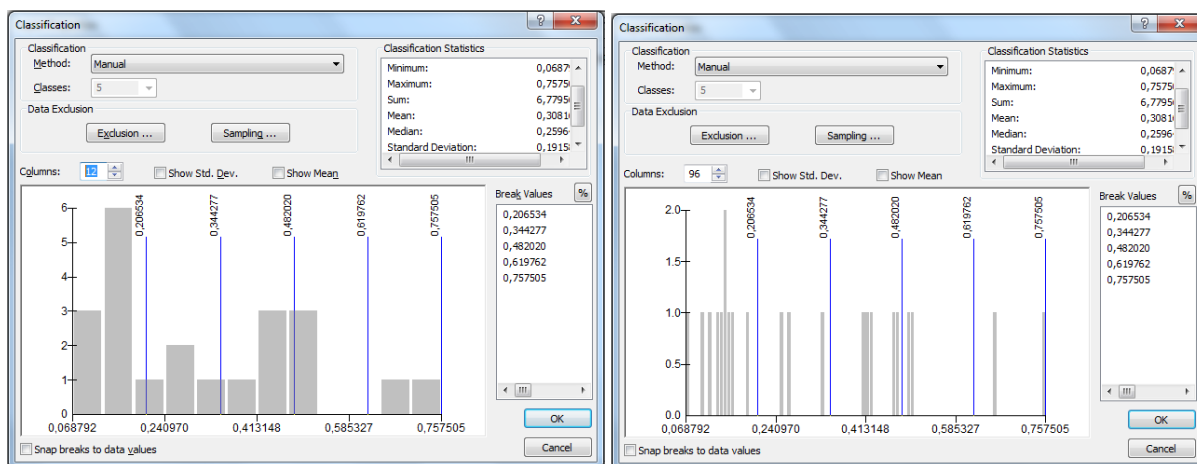
Podoba histogramu se odvíjí od statistického rozdělení dat a dále od zadaného počtu sloupců (*Columns*). Výška každého sloupce znázorňuje počet hodnot ze vstupních dat, které spadnou do rozmezí vymezeného šířkou sloupce na ose x. Jasnější představu lze získat z následujícího příkladu.

Opusťme na chvíli lesní půdu a představme si, že máme naprosto jednoduchá data, která popisují 10 hodů kostkou. Padla čísla 1, 1, 2, 3, 3, 3, 4, 4, 5 a 6. Nyní se podívejme na obrázek 8, jak by vypadal histogram těchto dat se dvěma a třemi sloupci. U prvního obrázku znázorňuje levý sloupec svoji výškou šest hodnot (1, 1, 2, 3, 3, 3) a pravý čtyři hodnoty (4, 4, 5, 6). Obdobě odpovídá výška sloupců četnostem konkrétních hodnot u obrázku druhého. A stejným způsobem je konstruován histogram také v ArcGIS.



Obr. 8. K ilustraci tvorby histogramu v ArcGIS.

Je tedy zřejmé, že pro korektní posouzení statistického rozdělení je žádoucí vyzkoušet více různých nastavení počtu sloupců. Na obrázku 9 jsou uvedeny histogramy ze stejných dat jako na obrázku 7, ale se 12 a 96 sloupci. Všimněte si, že při zadání extrémně vysokého počtu sloupců, je v podstatě každá hodnota vykreslena jedním sloupcem. Dalšími užitečnými ukazateli, které nám mohou pomoci při analýze dat, jsou průměr a medián, jež nalezneme vypočtené v okně *Classification Statistics*.



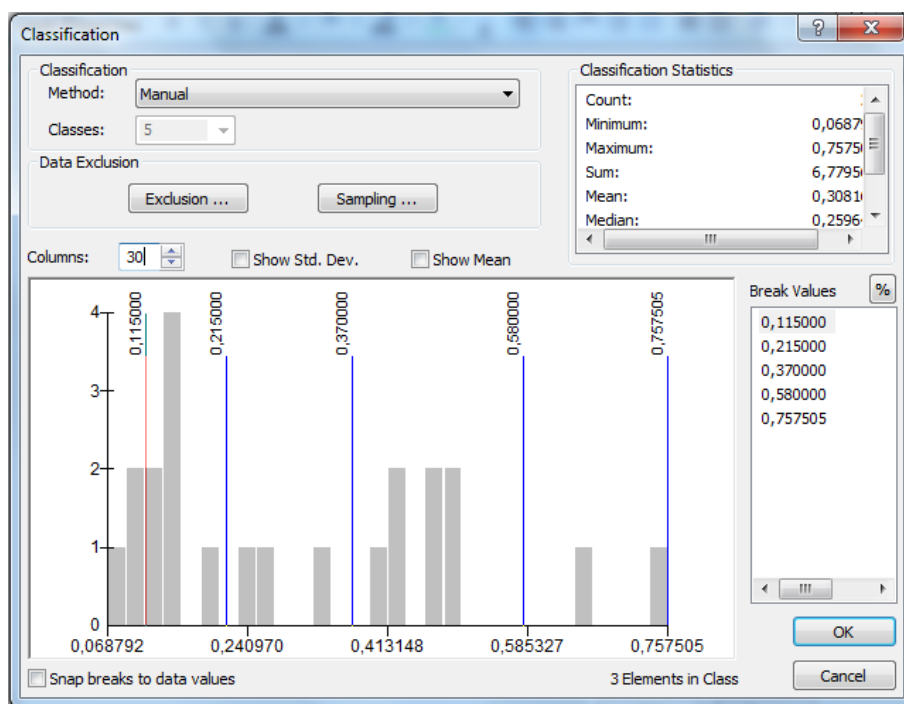
Obr. 9. K ilustraci vlivu počtu sloupců na výslednou podobu histogramu.

Nyní se podívejme, co lze z histogramu vyčíst. Rozdělení dat je na první pohled nerovnoměrné. Většina hodnot je soustředěna v levé části grafu, přičemž největší shluk je přibližně okolo hodnoty 0,15. Další shluk hodnot je mezi hodnotami 0,4 a 0,5 a dále se v datech objevují také dvě odlehlá pozorování, která jsou zachycena v pravé části histogramu. Pokud se pokusíme výše uvedené závěry interpretovat, můžeme říci, že v Moravskoslezském kraji existuje z hlediska lesní půdy několik skupin vzájemně si podobných SO ORP. První skupinou jsou SO ORP, kde lesy zaujímají přibližně

15 %. Dále je to skupina SO ORP, kde lesy tvoří 40–50 % celkové plochy. A poslední skupinu tvoří SO ORP s velmi vysokým podílem lesů.

Pokud vezmeme v potaz, že cílem kartogramu je poskytnout porovnání mapovaných územních jednotek, pak je zřejmé, že obce uvnitř výše uvedených skupin, by měly být znázorněny stejně (nebo podobně) a jednotlivé skupiny by od sebe měly být naopak odlišeny. A právě převedení této myšlenky do praxe, je cílem rozdělení dat do intervalů neboli klasifikace dat.

Klasifikaci lze v ArcGIS provést dvěma způsoby. Zaprvé ručně, na základě podobné analýzy dat, jakou jsme prováděli v předešlých odstavcích. Nebo automaticky za pomoci některého z připravených algoritmů, s tím, že výslednou klasifikaci lze dodatečně ručně poupravit. Metoda klasifikace se vybírá z rozbalovacího seznamu nazvaného *Method*. První položka seznamu je označena *Manual* a umožňuje ruční klasifikaci. V praxi postupujeme tak, že po zvolení možnosti *Manual* zadáme požadovaný počet intervalů (*Classes*) a do okna *Break values* zadáme zvolené meze intervalů. Pro naše data by ruční klasifikace na základě výše popsané úvahy mohla vypadat jako na obrázku 10. Všechny popsané skupiny jsou zařazené do samostatných intervalů a nejpočetnější skupina SO ORP s podílem lesů okolo 15 % je ještě rozdělena na dva intervaly, aby mapa měla větší informační potenciál. Dalo by se samozřejmě diskutovat o tom, zda má být druhá hranice zleva mezi pátým a šestým sloupcem a ne raději mezi čtvrtým a pátým. Tato diskuze by však nejspíš nedošla k jednoznačnému závěru, neboť vhodných způsobů klasifikace často existuje několik a žádný přitom nelze prohlásit za jednoznačně nejlepší.



Obr. 10. Nastavení klasifikace metodou Manual.

Klasifikační algoritmy v ArcGIS

V praxi se obvykle využívá některého z klasifikačních algoritmů, které jsou v ArcGIS implementovány a výsledek automatického zpracování se následně ručně upraví. Je proto vhodné rozumět tomu, jak jednotlivé algoritmy pracují.

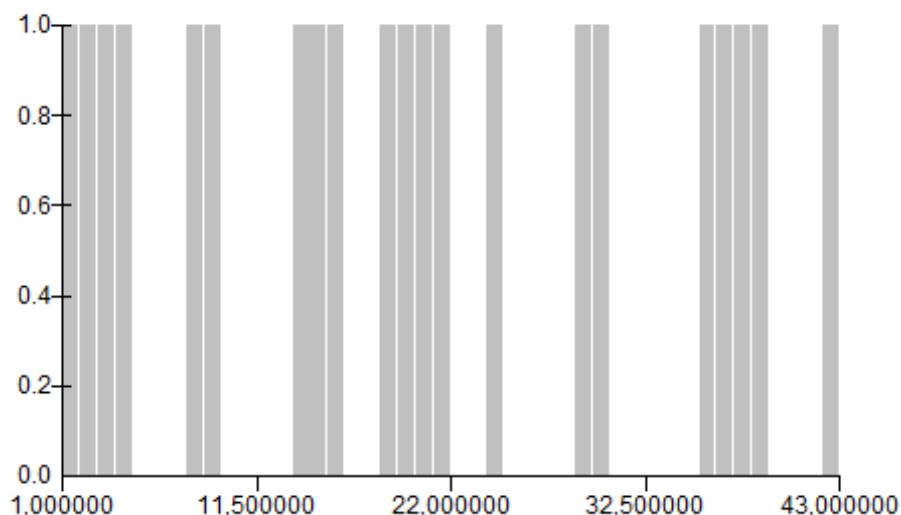
Manual představuje ruční klasifikaci, při které tvůrce mapy nastaví hranice intervalů dle analýzy dat a vlastního uvážení.

Algoritmus **Equal Interval** v datech nejprve nalezne minimum a maximum a poté rozdělí rozpětí mezi těmito hodnotami do stejně velkých intervalů. Pokud by bylo minimum např. 0, maximum 10 a požadovali bychom 5 intervalů, pak by šířka všech intervalů byla 2 a hranice mezi intervaly by odpovídaly hodnotám 2, 4, 6 a 8. Tento algoritmus nelze doporučit, neboť zcela ignoruje statistické rozdělení dat. Budeme-li mít data, která jsou rozdělena nerovnoměrně, může dojít k tomu, že v některých intervalech bude hodnot příliš mnoho a v jiných naopak příliš málo či dokonce žádné. Zároveň hrozí umístění hranic mezi hodnoty, které reprezentují podobné regiony, a tudíž by měly spadat do intervalu jednoho.

Opačný postup nabízí klasifikace **Defined Interval**. Zatímco v předchozím případě jsme zadávali počet intervalů a program dopočítával jejich šířku, zde zadáváme šířku intervalu a počet je závislý na tom, kolikrát se zadaný interval vejde mezi minimum a maximum. Použijeme stejný příklad avšak pro data s minimem 0 a maximem 10 budeme nyní požadovat intervaly s šířkou 3. Hranice intervalů budou v hodnotách 3, 6 a 9 a data tedy budou rozdělena do 4 intervalů. Také u tohoto postupu platí, že zde není žádná opora v analýze statistického rozdělení dat a proto ani tento algoritmus nelze doporučit. Výjimkou mohou být snad jen případy, kdy máme intervaly definovány nějakým specifickým zadáním, jež je třeba dodržet (vyhláška apod.).

První klasifikací, která zohledňuje rozdělení dat, je klasifikace **Quantile**. Hranice mezi intervaly odpovídají zvoleným kvantilům. V případě, že zvolíme počet intervalů 2, bude hranice mezi intervaly odpovídat mediánu. Pro počet intervalů 4 budou jako hranice použity kvartily, pro 5 intervalů kvintily atd. Jako příklad vezmeme data 1, 1, 2, 3, 7, 11, počet hodnot n je roven 6. V případě, že bychom chtěli pouze dva intervaly, bude každý z těchto intervalů obsahovat právě $n/2$ hodnot a hranice mezi intervaly bude rovna mediánu, tedy hodnotě 2,5. Pro tři intervaly by každý interval obsahoval právě $n/3$ hodnot a hranice by byly 1,5 a 5. Před výpočtem kvantilů (hranic) musí být data vždy seřazena. Tento algoritmus je jistě vhodnější než algoritmy předchozí, neboť eliminuje možnost, že by do některého z intervalů padlo příliš mnoho hodnot a do jiných téměř žádné či dokonce žádné. Stále zde však může docházet k nevhodnému umístění hranic tak, že územní celky s podobnými hodnotami budou nakonec v různých intervalech. Tento problém se týká dat, která nemají normální rozdělení hodnot a je patrný i z uvedeného příkladu, kdy při rozdělení do dvou intervalů hodnota 3 spadla do intervalu s hodnotami 7 a 11, přestože k hodnotám 1,1 a 2 má viditelně blíže.

Za velmi sofistikovaný algoritmus lze považovat **Natural Breaks**, v literatuře často označovaný jako **Jenks** či **Jenks Optimization**. Algoritmus při výpočtu hranic sleduje dva parametry. Prvním je vnitřní homogenita intervalů. Jinými slovy, snaží se nastavit hranice tak, aby mezi daty spadajícími do jednoho intervalu byly co nejmenší vzájemné rozdíly. Druhým rozhodujícím parametrem je vzájemná heterogenita mezi intervaly, která by měla být co nejvyšší. Představme si, že bychom měli do tří intervalů klasifikovat data 1, 2, 3, 4, 8, 9, 14, 15, 16, 19, 20, 21, 22, 24, 29, 30, 36, 37, 38, 39 a 43. Graficky jsou data znázorněna na obrázku 11. Záměrně byl zvolen vysoký počet sloupců v histogramu, aby se každá hodnota projevila jako jeden samostatný sloupec.

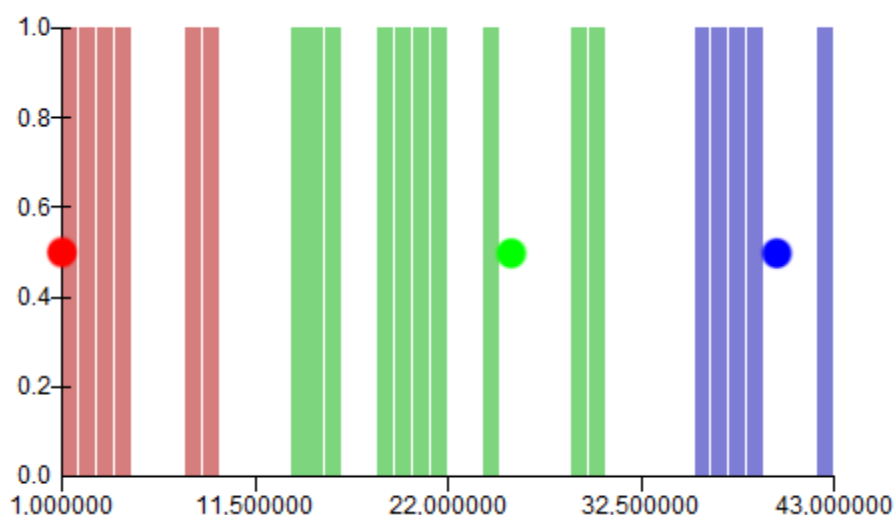


Obr. 11. Histogram pro uvedené hodnoty pro vysvětlení algoritmu *Natural Breaks*.

Nejprve se věnujme vnitřní homogenitě souborů. Algoritmus postupuje tak, že se nesnaží přímo stanovit hranice, ale definovat tzv. centroidy intervalů. Centroid je taková hodnota, pro níž platí, že leží uvnitř intervalu a suma kvadrátů vzdáleností ke všem ostatním hodnotám uvnitř intervalu je minimální. Aby byla splněna podmínka heterogenity mezi třídami, usiluje algoritmus zároveň o co největší rozdíly mezi hodnotami centroidů.

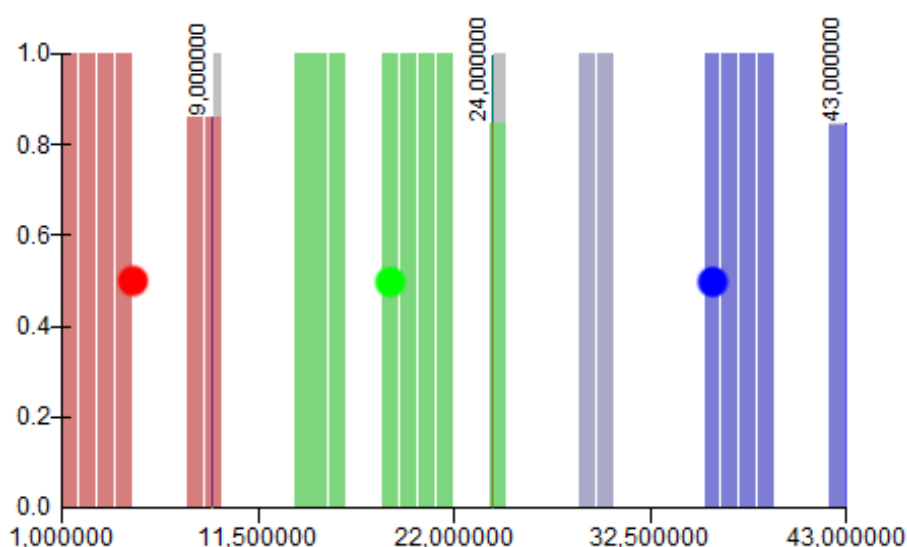
V prvním kroku jsou hodnoty centroidů zvoleny náhodně, následně se počítají jako průměry z hodnot, které jsou centroidu nejbližší. Počáteční a závěrečný stav je zachycen na obrázcích 12a a 12b.

Jako výchozí hodnoty centroidů zvolme např. 1, 25 a 40. Zvolit bychom však mohli jakékoliv jiné tři hodnoty ležící mezi minimální a maximální hodnotou v datech. Hodnoty, které jsou k centroidům nejbližší, jsou obarveny barvou příslušného centroidu.



Obr. 12a. K vysvětlení algoritmu *Natural Breaks* – stav po inicializaci.

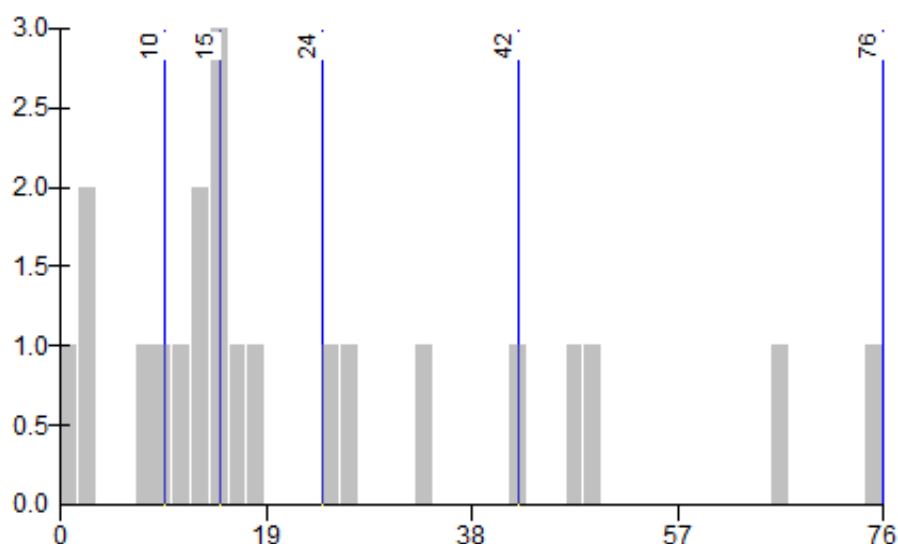
Následně se z hodnot obarvených stejnou barvou vypočte průměr a centroid se přesune na pozici průměru. Pro červenou skupinu bude centroid přesunut na hodnotu 4,5, pro zelenou na hodnotu 21 a pro modrou na hodnotu 38,6. Zároveň se zjišťuje vzájemná heterogenita mezi intervaly (vzdálenost mezi centroidy). Ta se tímto krokem sice zmenšila, nárůst homogenity uvnitř tříd je však tak výrazný, že snižující se heterogenitu převáží. Následně se celý postup opakuje a to až do doby, než se pozice centroidů ustálí a mezi dvěma iteracemi (opakováními) nedojde k přesunu žádného centroidu. Posledním krokem algoritmu je stanovení hranic mezi intervaly, což v případě, že známe zařazení hodnot do intervalů, nepředstavuje problém. Výsledná klasifikace algoritmem *Natural Breaks* bude vypadat jako na obrázku 12b. Hranice budou procházet mezi hodnotami 9–14 a 24–29. Program ArcMap se vždy snaží stanovit hranici u některé z hodnot. V tomto případě by byly hranice stanoveny v hodnotách 9 a 24. Vhodnější však je vést hranice skutečně mezi krajními hodnotami sousedících intervalů. V tomto případě by bylo lepší hranice manuálně přesunout do hodnot 11 či 12 a do hodnot 26 či 27.



Obr. 12b. K vysvětlení algoritmu *Natural Breaks* – finální stav.

Zcela zásadní výhodou tohoto algoritmu je skutečnost, že usiluje o rozdělení dat na základě stejných kritérií, jako jsme to činili my při manuální klasifikaci na základě analýzy statistického rozdělení dat. Tedy, aby hodnoty vzájemně si blízké spadaly do stejného intervalu. Z toho důvodu lze *Natural Breaks* považovat za velmi vhodný pro klasifikaci dat za účelem tvorby kartogramu. Přinejmenším může sloužit jako vodítko při našem rozhodování, přičemž navržené hranice můžeme vždy dodatečně manuálně upravit.

Algoritmus *Geometrical Intervals* byl vyvinut společností ESRI a původně byl určen pro nadstavbu *Statistical Analyst*. Princip algoritmu je obchodním tajemstvím a nebyl zveřejněn. Zdokumentovány jsou pouze vlastnosti výsledných klasifikací, které algoritmus generuje. Pro vytvořené intervaly platí, že jejich velikost vychází z intervalu nejmenšího, který je vymezen v oblasti s největší hustotou dat. Velikost (šířka) ostatních intervalů je vypočtena jako násobek šířky nejmenšího intervalu a koeficientu c , jehož hodnota roste/klesá geometrickou řadou.

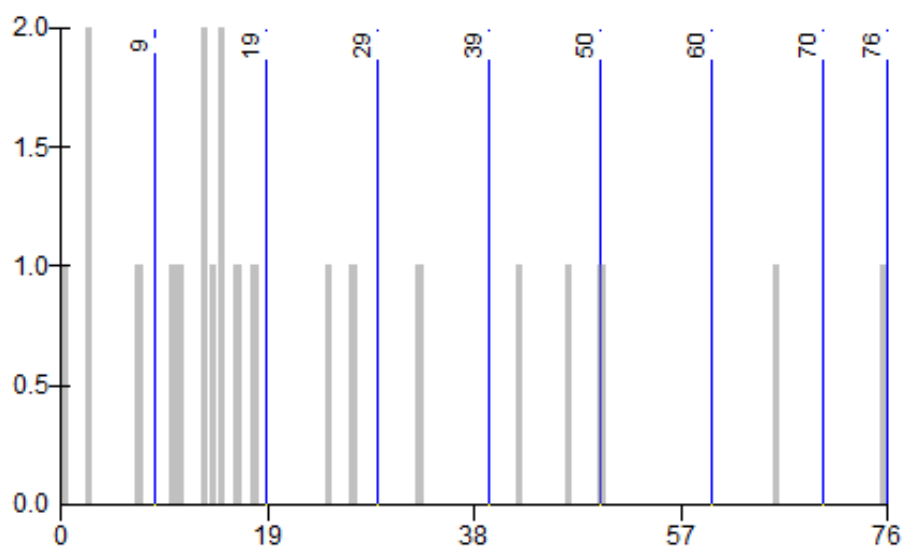


Obr. 13. Histogram pro ukázková data pro vysvětlení algoritmu *Geometrical Intervals*.

Rozdělení dat do intervalů je zachyceno na obr. 13. Nejmenší interval má šířku 5 a nachází se v místě, kde je v histogramu shluknuto nejvíce hodnot. Koeficient c byl algoritmem stanoven na hodnotu 1,89. Oba intervaly okolo nejmenšího intervalu jsou stejně velké a jejich šířka odpovídá šířce nejmenšího intervalu násobně první mocninou koeficientu c . Šířka intervalů tedy vychází $5 \cdot 1,89 = 9$. Protože nejmenší hodnota v datech je 1, nalevo od nejmenšího intervalu lze vytvořit pouze jeden jediný další interval. Napravo je obdobným způsobem vypočtena šířka dalšího intervalu, pouze se místo koeficientem c násobí šířka nejmenšího intervalu druhou mocninou c . Dokud nebudou všechna data zařazena do některého intervalu, budou se dopočítávat další intervaly, přičemž šířka intervalu neustále roste geometrickou řadou (s rostoucí mocninou c).

Jakým způsobem je na začátku výpočtu stanovena šířka nejmenšího intervalu, jeho poloha a hodnota koeficientu c , bohužel není v dokumentaci k algoritmu uvedeno. Společnost ESRI uvádí, že algoritmus produkuje kartograficky korektní výsledky, a to i v případě, že pracuje s daty, jež nemají normální rozdělení. Vzhledem k tomu, že princip algoritmu je utajen, nezbyvá nám, než tomuto popisu věřit. Pokud se však rozhodneme tento algoritmus použít, vždy bychom měli řádně zkontrolovat, zda výsledek skutečně odpovídá kartografickým požadavkům, tak jak autoři algoritmu deklarují.

Poslední možností, jak klasifikovat data, je využití algoritmu **Standard Deviation**. Tento algoritmus je velmi jednoduchý a do značné míry připomíná *Defined Intervals*. Oba algoritmy dělí data do intervalů o zadané velikosti. V případě *Standard Deviation* však nevolíme velikost intervalů libovolně, ale vybíráme z násobků směrodatné odchylky vypočtené z klasifikovaných dat. Hodnota směrodatné odchylky je mj. uvedena ve výčtu základních statistických ukazatelů *Classification statistics*. Na obr. 14 jsou data rozdělena do intervalů s šířkou rovnou polovině směrodatné odchylky. Z obrázku je také patrné, že poloha intervalů je dána polohou výchozího intervalu, který je umístěn vždy symetricky okolo průměru vypočteného ze vstupních dat. Podobnost s algoritmem *Defined Intervals* s sebou bohužel přináší podobné nevýhody, kvůli kterým nelze tento způsob třídění dat příliš doporučit. Opět se jedná především o nezohlednění statistického rozdělení původních dat, které může způsobit nevhodné roztřídění dat do intervalů.



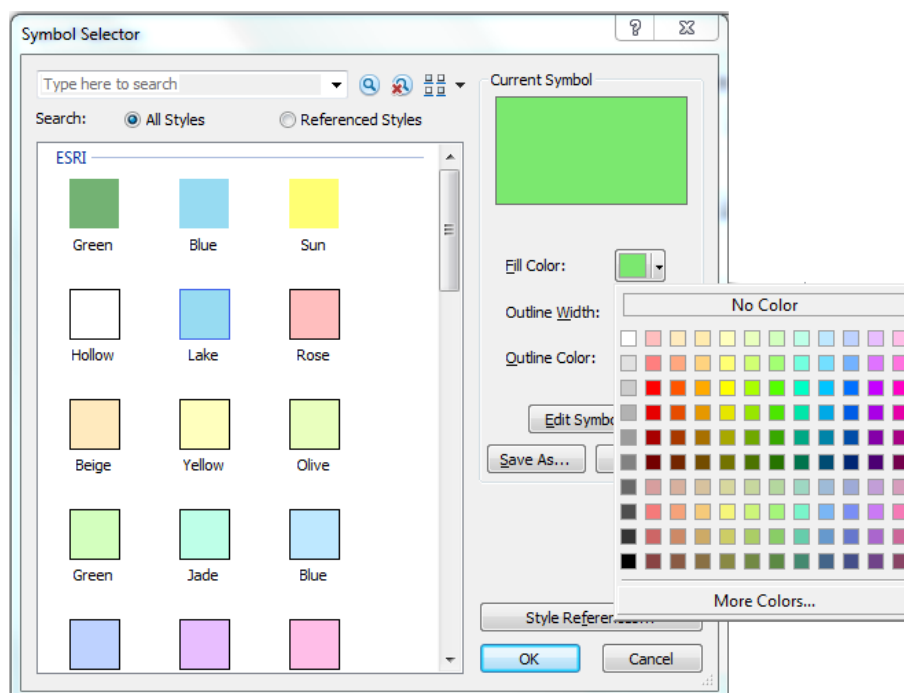
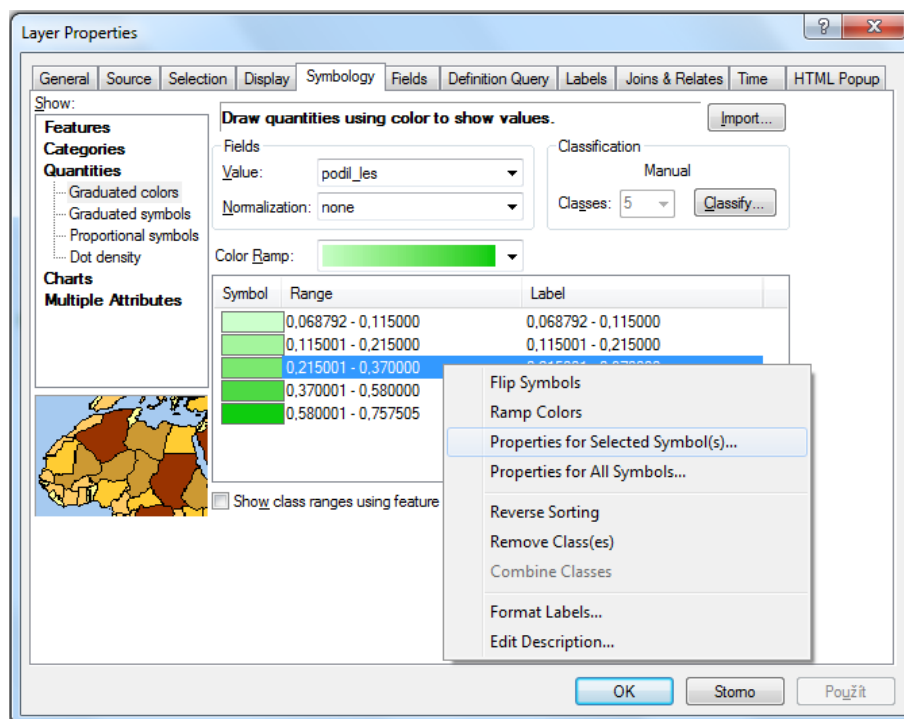
Obr. 14. Histogram pro ukázková data pro vysvětlení algoritmu *Standard Deviation*.

Vizualizace dat

Poté, co jsme data rozdělili do intervalů, musíme každému z nich přiřadit vhodnou grafickou reprezentaci, jež bude daný interval zastupovat v mapě. Z kartografických vyjadřovacích prostředků máme k dispozici barvu či rastr. Vždy je důležité mít na paměti, že s rostoucí intenzitou jevu (vyšší hodnoty) musí růst také intenzita barvy či rastru. V případě barvy je dále třeba omezit použití různých barevných tónů a naopak pracovat převážně s jasnem a sytostí.

V ArcGIS jsou znaky pro jednotlivé intervaly definovány v okně *Graduated Colors*. Ve spodní části okna se nachází tabulka, ve které každý řádek představuje jeden interval. Ve sloupcích jsou postupně uvedeny kartografické znaky pro intervaly (*Symbol*), meze intervalů (*Ranges*) a popis, který se následně objeví v legendě (*Label*). Meze intervalů by měly odpovídat mezím, které jsme nastavili při klasifikaci dat (*Break values* v okně *Classify*).

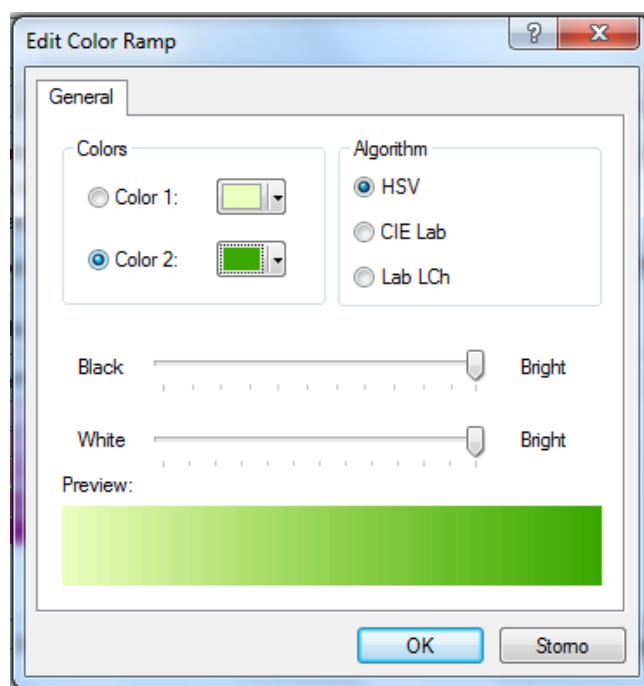
Nejdříve se pokusme vytvořit kartogram s využitím barev. Pravým tlačítkem myši klikneme na některý řádek v uvedené tabulce a z kontextové nabídky zvolíme položku *Properties for Selected Symbol(s)*... . Otevřeme tak okno *Symbol Selector*, ve kterém můžeme definovat barvu, jež bude reprezentovat příslušný interval. Pokud nám žádná z nabízených barev nevyhovuje, můžeme kliknout na položku *More Colors*... a definovat si barvu vlastní, viz obr. 15. Z nabízených barevných modelů je v tuto chvíli nejvhodnější HSV, neboť potřebujeme pracovat především s barevným jasnem ($V = value$ neboli jas).



Obr. 15. Manuální nastavení barvy pro vybraný symbol.

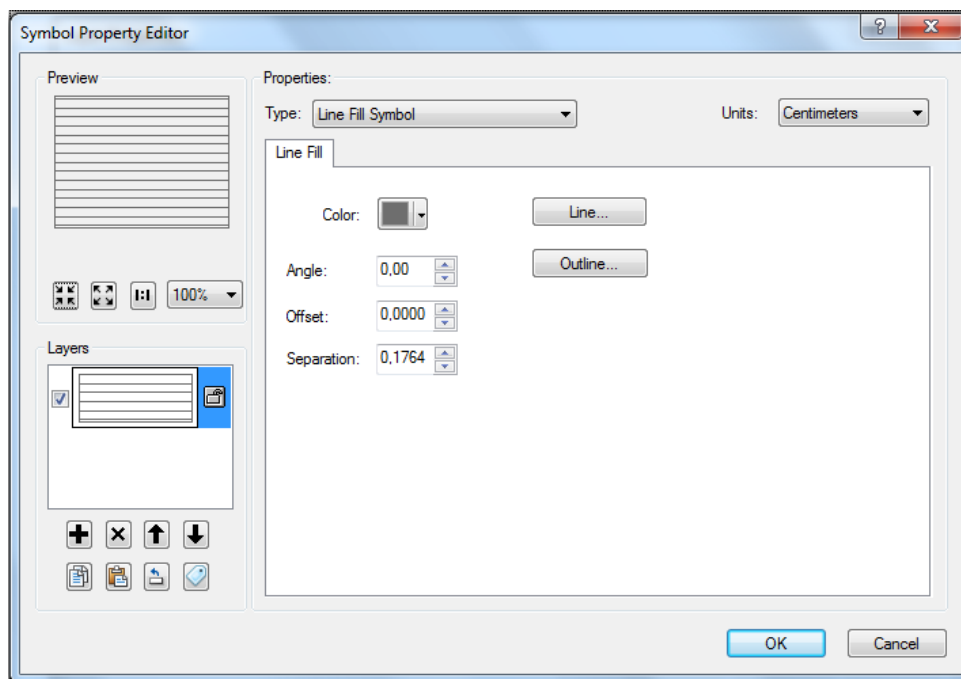
Výhodou manuálního definování jednotlivých barev je přímá kontrola nad barevnou podobou kartogramu. Zároveň je ale tento způsob značně pracný a dává největší prostor pro vznik chyb (propadání barev ve vytvořené stupnici). Proto je často vhodnější zvolit některou z připravených barevných stupnic či si vytvořit stupnici vlastní definováním nejméně a nejvíce intenzivní barvy s tím, že ostatní barvy dopočítá program za nás. Paletu s barevnými stupnicemi získáme kliknutím na barevný pruh nad tabulkou (nabídka *Color Ramp*). Pozor, ne všechny stupnice jsou vhodné pro kartogram!

V případě, že nám žádná ze stupnic nevyhovuje, můžeme kliknout pravým tlačítkem na barevný pruh, zvolit položku *Properties...* a sestavit si stupnici dle vlastních představ (viz obr. 16).



Obr. 16. Definice vlastní barevné stupnice.

Kromě barev lze jednotlivé intervaly rozlišit také rastrem. Použití kvantitativního rastru je v ArcMap nepatrně složitější, neboť program neumí automaticky generovat rastrové stupnice a tudíž musíme znaky pro všechny intervaly definovat manuálně. Klikneme tedy pravým tlačítkem na první řádek v tabulce s intervaly, zvolíme položku *Properties for Selected Symbol(s)...*. Následně v okně *Symbol Selector* klikneme na tlačítko *Edit Symbol...*, čímž otevřeme okno *Symbol Property Editor* umožňující pokročilá nastavení kartografických znaků.



Obr. 17. Dialogové okno pro tvorbu liniového rastru.

Abychom získali rastr, vybereme z rozbalovacího seznamu *Type*: položku *Line Fill Color*, viz obr. 17. U rastru můžeme nastavit natočení linií (*Angle*:), posunutí linií (*Offset*:) a rozestup linií (*Separation*:). Kromě toho lze po stisknutí tlačítka *Line...* definovat barvu, šířku a případně i strukturu linií. Tlačítko *Outline...* slouží pro nastavení parametrů linie ohraničující polygon vyplněný rastrem. Hranice linií je však výhodnější nastavit hromadně pro všechny intervaly (viz dále). Pokud pracujeme s větším počtem intervalů a nevystačíme s jednosměrným rastrem, můžeme také využít vícesměrný rastr. Další vrstvu linií přidáme stisknutím tlačítka se symbolem „plus“ v levé dolní části okna. V podokně *Layers* se objeví další vrstva linií, kterou po označení můžeme opět libovolně editovat. Zda je navržená stupnice rastrů správná, lze snadno poznat při pohledu do mapového pole s tím, že i bez legendy musí být na první pohled zřejmé, ve kterém polygonu rastr je nejméně intenzivní, který je druhý v pořadí atd. V případě pochybností je vhodné parametry rastru upravit.

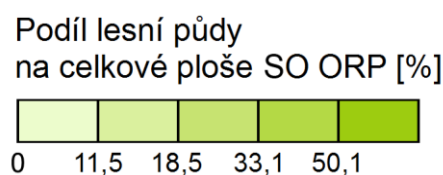
Doposud jsme pro polygony tvořící kartogram řešili pouze jejich výplň. Je však třeba mít na paměti, že polygonu mají také hranice, a i tyto je třeba náležitě znázornit. Hranice by měly být pro všechny polygony jednotné. Stejně jako u výplně můžeme pracovat postupně a nastavit všem polygonům hranice ručně (v okně *Symbol Selector* položky *Outline Width*: a *Outline Color*:. Praktičtější však je, nastavit hranice všem polygonům najednou. Toho docílíme tak, že po kliknutí pravým tlačítkem na některý z intervalů v okně *Graduated Colors* vybereme položku *Properties for All Symbols...* a parametry hranice nastavíme zde.

Konstrukce legendy kartogramu

Jak budou jednotlivé intervaly v legendě popsány, je dáno popisky uvedenými v tabulce intervalů v okně *Graduated Colors*. Pro zadávání popisků je v tabulce připraven sloupec *Label*. Popisky často není potřeba zcela přepisovat a stačí upravit jejich formátování. Okno pro formátování popisků vyvoláme kliknutím levým tlačítkem na záhlaví sloupce *Label*.

Pokud máme popisky upraveny, můžeme vygenerovat legendu (*Insert* → *Legend*). Automaticky vygenerovaná legenda se v mnohém rozchází s kartografickými pravidly, a proto je nutné ji náležitě upravit (více o formátování legendy lze nalézt v nápovědě k programu ArcGIS). Z hlediska kartogramu vyžaduje zvláštní pozornost zejména legenda k rastru. Znak rastru musí být v legendě reprezentován vždy aspoň dvěma liniemi, aby byl zřejmý rozestup linií. Pokud se dvě linie do políčka v legendě nevejdou, je třeba buď upravit rastr či zvětšit políčka legendy. V případě, že je rozestup linií rastru menší než velikost políčka v legendě (a dvě linie se tam mohou vejít), může pomoci posunutí legendy či nastavení parametru *Offset*: v okně *Symbol Property Editor* (viz výše).

Především na barevných kartogramech se někdy můžeme setkat s legendou zpracovanou do podoby barevného pruhu (viz obr. 4, obr. 18). Tato varianta je vhodná pro spojitě jevy a eliminuje riziko překrývajících se intervalů. Na druhou stranu ji nelze v ArcGIS vygenerovat automaticky a je třeba ji sestavit manuálně po převedení automaticky vytvořené legendy do grafiky.

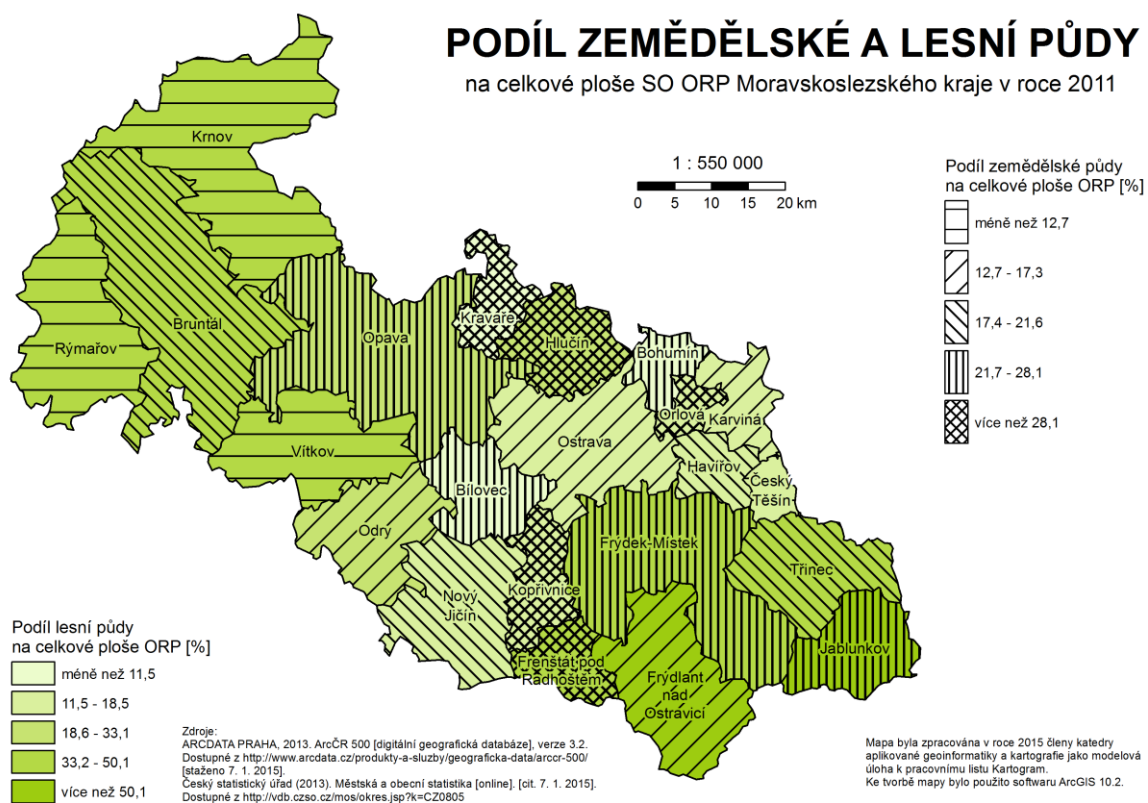


Obr. 18. Legenda v podobě barevného pruhu vhodná pro spojitě jevy.

Tvorba složeného kartogramu

Zadáním modelové úlohy bylo vytvořit složený kartogram vyjadřující podíl lesní půdy na celkové ploše SO ORP a podíl zemědělské půdy na celkové ploše SO ORP. Složený kartogram vznikne kombinací dvou jednoduchých kartogramů. Pokud chceme kombinovat více kartogramů, je vhodné použít kartogramy rastrové, či jeden kartogram barevný a ostatní rastrové. Kombinace více barevných kartogramů je značně problematická a v praxi se příliš nevyužívá.

Pomineme-li v tuto chvíli pokročilé techniky s využitím kartografických reprezentací, platí, že každé datové vrstvě v ArcGIS lze přiřadit pouze jeden způsob vizualizace. Protože budeme kombinovat dva kartogramy, načteme si naše data *data_ms_orp* dvakrát (tj. do dvou vrstev). Spodní vrstvu použijeme pro kartogram lesní půdy horní pro kartogram zemědělské půdy. Data z obou vrstev postupně klasifikujeme do intervalů a následně spodní vrstvu upravíme do podoby barevného kartogramu a horní vrstvu do podoby rastrového kartogramu. Pokud rastr zpracujeme správně, bude pod ním ve všech polygonech prosvítat barevný kartogram. Po dokončení všech kompozičních prvků může výsledek vypadat podobně jako na obrázku 19.



Obr. 19. Výsledný složený kartogram.

Cvičení

Úkol 1

Diskutujte, zda je, či naopak není, metoda kartogramu vhodná pro znázornění následujících témat? Uvažujeme pouze použití výše uvedených typů kartogramů (jednoduchý homogenní, jednoduchý kvalifikační a složený).

- Vyjádření etnického složení obyvatelstva v krajích Česka.
- Znázornění, do jaké míry se podílí obnovitelné zdroje energie na celkové produkci elektřiny ve státech Evropy.
- Zachycení vztahu mezi průměrnou mírou nezaměstnanosti, porodností a podílem vysokoškolsky vzdělaného obyvatelstva v SO ORP v Plzeňském kraji.
- Vyjádření počtu pivovarů ve spolkových zemích Německa.
- Zmapování hustoty železniční a silniční sítě v okresech Česka.
- Mapa znázorňující celkovou kapacitu mateřských škol na 10 km² v obcích okresu Písek.
- Mapa průměrných ročních cen nafty ve státech USA.
- Znázornění ročního srážkového úhrnu (přestože se srážky uvádějí v mm, měří se v litrech na m²) ve státech světa.

Řešení

- Z uvedených typů kartogramů není žádný vhodný pro vyjadřování struktury mapovaného jevu. Použit by šel strukturní kartogram. Nejlepším řešením by však bylo použití kartodiagramu, neboť ten se na rozdíl od strukturního kartogramu výrazně lépe čte a interpretuje.
- Zde je použití kartogramu na místě. Data o produkci energie z obnovitelných zdrojů jsou prezentována jako podíl na celkové produkci energie. Jedná se tedy o data relativní a tudíž vhodná pro znázornění kartogramem.
- Také v tomto případě je metoda kartogramu vhodnou volbou. Všechny tři jevy jsou popsány relativními daty. Abychom byli schopni zachytit více jevů do jedné mapy, musíme použít složený kartogram. Dva jevy znázorníme rastrem a jeden barvou.
- Kartogram použít nelze. Pro mapování absolutních hodnot (počtů pivovarů) není metoda kartogramu určena.
- Obecně platí, že pro mapování hustot (v geografickém smyslu) je kartogram vhodný, neboť hustota vždy vyjadřuje počet jednotek přepočtený na určitou plochu. Taková data jsou vždy relativní.
- Po stránce kartografické by bylo použití kartogramu v pořádku. Nicméně je třeba zamyslet se také nad smyslem mapy. Pokud si položíme otázku, zda má smysl přepočítávat kapacitu mateřských škol na plochu územních jednotek, nejspíš dojdeme k závěru, že nikoliv. Lze předpokládat, že kapacita škol bude ovlivněna spíše počtem dětí v oblasti či obecněji počtem obyvatel. Správnější by proto bylo přepočítat počet mateřských škol právě k těmto jevům.
- Výpočet průměru za časové období nezpůsobuje relativizaci dat. Kartogram proto obdobně jako tomu v příkladu d) nelze použít.
- Zde nemá příliš smysl použití kartogramu zvažovat, neboť srážky jsou typickou ukázkou jevu, jež se v území mění spojitě a není nijak vázán na administrativní hranice vymezené uměle. Kromě toho jsou srážky měřeny lokálně na srážkoměrných stanicích a tudíž ani nelze přesně určit, kolik srážek spadlo na území určitého státu. Problém by také nastal při interpretaci mapy. Na hranicích mezi státy by docházelo k ostrým přechodům mezi intervaly, což by mohlo evokovat, že na jedné

straně hranice příší výrazně více než na druhé, což je samozřejmě silně nepravděpodobné. Tato data je vhodné nejprve interpolovat a do mapy znázornit pomocí areálové metody či izolinií.

Úkol 2

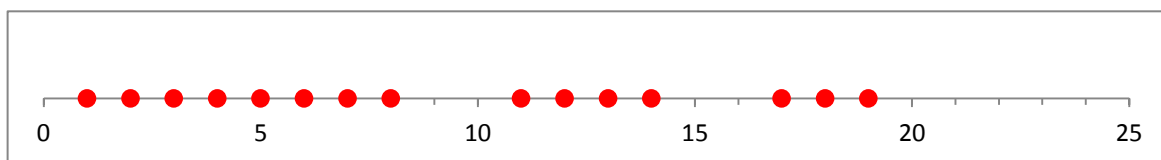
Na základě počtu územních jednotek navrhnete vhodný počet intervalů pro mapy znázorňující hustotu zalidnění v krajích, okresech, SO ORP a obcích Česka.

Řešení

- kraje (14): 4–5 intervalů,
- okresy (77): 7–9 intervalů,
- SO ORP (205): teoreticky 9–14 intervalů, v praxi lze za realizovatelné maximum považovat 10 intervalů,
- obce (6253): teoreticky 14–80 intervalů, v praxi lze za realizovatelné maximum považovat 10 intervalů.

Úkol 3

Pro 15 hodnot zaznamenaných na grafu níže navrhnete klasifikaci do pěti intervalů. Zdůvodněte, proč je Vaše klasifikace vhodná. Pro níže uvedená data si můžete také vyzkoušet vytvořit histogramy se třemi, šesti a devíti sloupci.



Řešení

Cílem klasifikace by mělo být sdružit do intervalů hodnoty navzájem si podobné. Na první pohled jsou v datech tři shluky hodnot, mezi nimiž by se zajisté měly nacházet hranice intervalů. Dále můžeme rozdělit na tři menší části největší shluk hodnot (1–8) nebo rozdělit tento shluk pouze na dvě části a na dvě části rozdělit také shluk hodnot (11–14). Zde by záleželo na charakteru dat a účelu mapy.

Úkol 4

Navrhnete rastrovou stupnici pro 6 intervalů, aniž byste měnili rozestup linií.

Řešení

Varianta, v níž se mění pouze směr rastru, není příliš vhodná (viz teoretická část). V ukázce níže dochází jednak ke změně směru a dále ke změně tloušťky linií, což čtení rastru značně usnadňuje.



Souhrn

- Metoda kartogramu je určena pro relativní data. Máme-li data absolutní, musíme je nejprve zrelativizovat (tj. vztáhnout k rozloze územní jednotky, jejím počtu obyvatel atd.).
- Data je třeba klasifikovat do vhodného počtu intervalů. Algoritmy dostupné v ArcGIS nám tuto práci usnadní, ale vždy bychom měli výsledné rozdělení zkontrolovat na základě histogramu.
- Při volbě barev či rastru je třeba mít na paměti, že jednotlivé kategorie musí být dobře rozlišitelné a s narůstající hodnotou jevu musí narůstat intenzita barvy či rastru.
- Při tvorbě legendy je třeba se vyvarovat překrývání jednotlivých intervalů, tj. libovolná hodnota ze vstupního souboru musí patřit do právě jednoho intervalu.

Literatura

- [1] Voženílek V., Kaňok J. a kol. (2011): Metody tematické kartografie – vizualizace prostorových jevů. Olomouc: Univerzita Palackého, 2011.
- [2] Bláha J. D. (2013): Tvorba map ve věku geoinformačních systémů (3. část): Vyjadřovací prostředky mapy. Geografické rozhledy, 3/12-13, ročník 22, s. 10–11.
- [3] Voženílek V. (2001): Aplikovaná kartografie I.: tematické mapy, 2. vydání. Olomouc: Univerzita Palackého, 2001.