

Socioekonomická klasifikace měst v ČSR

JIRÍ LINK

Výzkumný ústav výstavby a architektury, Praha

Úvod

Města na území ČSR prošla složitým ekonomickým, kulturně historickým a politickým vývojem. Procesy, charakterizující tento vývoj, se nedotkly všech měst stejnou měrou. V současné době jsou proto města ČSR značně diferencována. Tato mnohotvárná diferenciaci značně ztěžuje, ba i někdy znemožňuje zevšeobecňující pohled na strukturu městského osídlení, zvláště uvažujeme-li několik hledisek současně.

Vhodnou metodou k popisu a analýze existujícího stavu je mnohorozměrná klasifikace. Exaktním způsobem a s předem danou přesností kondenzuje velké množství informací o městech a seskupuje je do víceméně malého počtu skupin relativně homogenních co do všech uvažovaných hledisek.

V našem příspěvku uvedeme výsledky mnohorozměrné klasifikace měst z hlediska socioekonomických, demografických a bytových charakteristik. Tato klasifikace byla zpracována ve *Výzkumném ústavu výstavby a architektury* v letech 1973 až 1975 jako součást širšího výzkumného úkolu zabývajícího se problémy urbanizace a osídlení. Jelikož v klasifikaci nemohly být pro nedostatek příslušných dat vzaty v úvahu urbanistické charakteristiky, jde v podstatě o socioekonomickou klasifikaci opírající se o data o obyvatelstvu a bytovém fondu. Přes tato omezení přináší však řadu zajímavých a cenných poznatků a odpovídá nárokům kladeným na klasifikace měst, u nichž se předpokládá přínos mimo jiného v oblasti

- *pojmového zpřesnění*
- *vytvoření podkladů k nástrojům pro další výzkumy*
- *rozboru faktorů, které působí na urbanizaci a růst měst*
- *dalšího plánování sídel*
- *ujasnění vztahů mezi jednotlivými vlast-*

nostmi měst (v jaké míře a jak souvisí např. velikost města se strukturou obyvatelstva)

- *tvorba signalizačního systému na základě klasifikace, která poukazuje na nevyrovnanost vývoje určitého města.*

Výsledky klasifikace mohou být užitečné i pro další analogicky orientované studie a mají též nesporný metodologický význam.

Klasifikovaná města a indikátory

Předmětem klasifikace byly obce na území ČSR, které vykazovaly k datu posledního sčítání lidu, domů a bytů, 1. 12. 1970, 5000 a více trvale bydlících obyvatel. Tyto obce budeme dále nazývat městy (bez ohledu na jejich právní statut), obyvatelstvem budeme rozumět pouze trvale bydlící obyvatelstvo (pokud nebude výslovně uvedeno jinak).

V r. 1970 bylo na území ČSR 215 měst s 5000 a více obyvateli. Každé z těchto měst bylo charakterizováno více či méně rozsáhlým systémem charakteristik, které budeme nadále nazývat indikátory socioekonomické, demografické a bytové situace a úrovně města. Výběr indikátorů byl podmíněn těmito požadavky: Indikátory musí rovnoměrně pokrývat danou problémovou oblast a dobře vystihovat nejdůležitější stránky problému. Musí být ve všech městech dostupné a stejně spolehlivé. Z tohoto hlediska byl proveden předběžný výběr 82 indikátorů, jež byly vypočteny na základě exerce definitivních výsledků sčítání lidu, domů a bytů z r. 1970. Indikátory mají tvar procentuálních podílů či indexů a opírají se o absolutní údaje o obyvatelstvu a bytech přepočtené na územní organizaci platnou k 1. 1. 1972. Z pracovních důvodů a pro snadnou orientaci byly rozděleny do tří skupin — do skupiny indikátorů socioekonomické povahy (S), demografické povahy (D) a do skupiny indikátorů

bytového fondu (B). Pro každý indikátor bylo vypočteno jeho statistické rozložení a indikátor byl obvyklým způsobem standardizován na nulový průměr a jednotkovou směrodatnou odchylku.

Popis metodiky klasifikace

Vzhledem k tomu, že statistické rozložení některých indikátorů bylo většinou značně asymetrické, přikročili jsme z formálně matematických důvodů k transformaci proměnných, která zaručovala, aby distribuce znaků měla přibližně Gaussovo normální rozložení.

Statistická teorie totiž ukazuje, že charakter závislosti mezi dvěma normálně rozloženými proměnnými je lineární. Dále z ní vyplývá, že multinormalitou, tj. normalitou rozložení všech indikátorů v matici dat X , zajistíme ve velmi přijatelných mezích linearitu závislosti mezi všemi dvojicemi indikátorů matice X . Normalizací indikátorů splníme předpoklad adekvátního použití koeficientu lineární korelace (Bravais-Pearsonova), a tím i použití metody hlavních komponent. Rovněž mnohonásobná diskriminační analýza, které jsme použili, dává v případě normálně rozložených indikátorů přesnější výsledky.

Jsou známy návody, podle nichž na základě šikmosti a špičatosti rozložení určíme vhodnou transformaci. My jsme použili jiné metody. Pro každý indikátor bylo samočinným

počítačem určeno jeho statistické rozložení, dále pak aritmetický průměr, rozptyl, šikmost a špičatost. Každá proměnná pak byla transformována pomocí následujících funkcí: $1/X$, $X^{1/3}$, X^2 , X^3 , $\text{LOG}(X)$ a $\text{LOG}/\text{LOG}(X)$. Pro každý typ transformace byl zjištěn průběh rozložení a příslušné charakteristiky. Smirnov-Kolmogorovým testem jsme posléze testovali shodu s normálním rozložením. K další analýze jsme použili to rozložení, jež vykazovalo nejmenší odchylky od normálního rozložení. Oprávněnost postupu byla ověřována testováním odchylek měr šikmosti a špičatosti od očekávaných hodnot za předpokladu normálního rozložení (šikmost = 0, špičatost = 3). Dále byly z původního výběru indikátorů vyřazeny ty indikátory, jejichž komunalita (viz dále) byla nižší než 0,50. Zbylých 57 indikátorů bylo podrobeno další analýze. Zastoupení jednotlivých skupin indikátorů bylo rovnoměrné.

Zvolené indikátory by bylo možno již bezprostředně použít ke klasifikaci. Tento postup by však narazil na značné obtíže. Každou klasifikaci je třeba interpretovat a konfrontovat s odlišnými klasifikacemi. Interpretace založená na 57 takových klasifikacích by byla prakticky nemožná a sama o sobě značně nepřehledná. Dále by byla v důsledku zkořelovanosti indikátorů nutně poznamenána redundantní informací. Proto jsme ještě před vlastní klasifikací přikročili k analýze hlavních komponent korelační matice¹ založené na 57 indikátorech.

¹ Korelační matici $R = (r_{ik})$ vypočteme ze vzorce

$$R = ZZ^t/N$$

kde Z je matice standardizovaných hodnot v jednotlivých indikátorech a N je rozsah souboru. Analýzou hlavních komponent matice R určíme nové proměnné jako lineární kombinace standardizovaných hodnot původních proměnných z_{ij} . Tyto nové proměnné jsou jednak lineární nezávislé, jednak jejich statistická důležitost měřena podílem jejich rozptylu na celkovém rozptylu všech proměnných (rovném n) klesá s pořadovým číslem. Analýza hlavních komponent vychází z modelu

$$Z = AF$$

přičemž $R = A^t A$; $F = (A^t A)^{-1} A^t Z = A^{-1} A^t Z$.

Matice $A = (a_{ip})$ se nazývá maticí zátěží hlavních komponent, kde a_{ip} má význam koeficientu lineární korelace mezi i -tou proměnnou a p -tou hlavní komponentou matice $F = (f_{pj})$ je maticí skóre p -té hlavní komponenty pro j -tý objekt (např. město) ve standardizovaném tvaru. Konečně matice A je diagonální matice vlastních čísel korelační matice R . Vlastní čísla matice R , obvykle značená symbolem λ_p , jsou rovna rozptylu vyčerpaného p -tou hlavní komponentou, přičemž platí

$$\lambda_p = \sum_i a_{ip}^2 \quad p = 1, 2, \dots, s$$

Suma

$$h_i^2 = \sum_p a_{ip}^2$$

$$h_i^2 = 1 \quad \text{pro } s = n$$

$$h_i^2 < 1 \quad \text{pro } s < n$$

nazýváme komunalitou i -té proměnné. Má význam čtverce koeficientu mnohonásobné korelace i -té proměnné všemi hlavními komponentami. Vzhledem k tomu, že množství vyčerpaného rozptylu daného komponenty klesá s jejím pořadovým číslem, postačí obvykle několik málo hlavních komponent k tomu, aby vyčerpalu podstatnou část celkového rozptylu proměnných. V tomto případě je komunalita každé z proměnných menší než 1. Součet všech vlastních čísel udává rozptyl vyčerpaný zvolenými hlavními komponentami. Otázkou kolik hlavních komponent extrahovat, tj. jakou část celkového rozptylu proměnných vyčerpat, aby byl popis pomocí hlavních komponent adekvátní, lze odpovědět testováním statistické významnosti komponent. V praxi považujeme na významné ty hlavní komponenty, jejichž vlastní číslo je větší než 1 (Harman 1967). Tohoto kritéria jsme použili i my.

Analýza ukázala 12 hlavních komponent s hodnotou λ větší než 1. Tyto komponenty vyčerpaly přes 82 % celkového rozptylu. Byly dále metodou VARIMAX rotovány a 10 z nich interpretováno. Rotované hlavní komponenty představují skryté popisné dimenze sledované problematiky, tj. socio-ekonomické, demografické a bytové diferenciace měst v ČSR. Pro každé město a každou z 12 rotovaných komponent bylo vypočteno skóre, které měří polohu města na té které popisné dimenzi. Skóre byla vypočtena tak že sigma rovná se 1 a podle vzorce v poznámce 1.

Poslední typ skóre byl použit jako vstupní data pro klasifikaci. Skóre o jednotkové směrodatné odchylce nejsou pro účely klasifikace příliš vhodná, neboť ruší vazbu skóre na původní hodnoty proměnných.

Interpretace rotovaných hlavních komponent

Interpretace rotovaných hlavních komponent vychází především z numerických hodnot zátěží. Přihlíželi jsme především k těm zátěžím, jejichž absolutní hodnota byla větší nebo rovna 0,45, což znamená, že daná komponenta vysvětluje cca 20 % rozptylu daného indikátoru. Druhým pomocným kritériem při interpretaci byla komponentní skóre. V zájmu stručnosti uvádíme název komponenty, procento jí vyčerpaného celkového rozptylu a pouze charakter závislosti na proměnné pro zátěže v absolutní hodnotě větší než 0,45.

I. Demograficko-ekonomický potenciál města (dále označovaný jen vitalita) — 18,2 %

Komponenta kladně koreluje s velikostí společně hospodařící domácnosti (měřené několika indexy), s podílem věkové skupiny 0—14, se Sauvyho indexem, s podílem bytů postavených po r. 1945, s růstem počtu obyvatel v r. 1961—1970 a s maskulinitou obyvatelstva. Záporná korelace byla zjištěna s podílem věkové skupiny 60+ a s podílem starých a méně kvalitních bytů.

Interpretace hlavních komponent, pokud je to třeba, vychází z numerických hodnot zátěží. Aby byla interpretace snazší, transformujeme původní matici zátěží A na matici B (musí platit $BB^T = R$) tak, aby zhruba řečeno v každém sloupci matice B odpovídajícímu zátěží jednotlivých transformovaných komponent bylo co nejvíce čísel zanedbatelně malých a pak čísel v ab-

II. Sociokulturní úroveň obyvatel města — 11,5 %

Kladně koreluje s úrovní vzdělání, s podílem zaměstnanců, s podílem domácností vlastních auto, chatu, ledničku a vysavač, dále s podílem obyvatel české národnosti. Záporná korelace se projevuje u podílů dělníků a obyvatel slovenské národnosti.

III. Stupeň urbanizace města — 10,6 %

Byla zjištěna kladná korelace s podílem bytů vybavených vodovodem, elektřinou a plynem, s podílem rodinných domků, s hustotou zalidnění, velikostí města měřenou počtem jeho obyvatel a velikostí domu měřenou počtem bytů. Zápornou korelaci pozorujeme pak u podílu bytů s elektřinou a vodovodem, u podílu bytových domů, podílu obyvatel zaměstnaných v zemědělství a lesnictví a nakonec u podílu bytů III. a IV. kategorie.

IV. Významnost terciéru v ekonomické aktivitě obyvatelstva — 6,4 %

Komponenta kladně koreluje s podílem obyvatel zaměstnaných v terciárním sektoru a s podílem zaměstnaných ve stavebnictví, záporně pak s podílem obyvatel zaměstnaných v průmyslu.

V. Celková úroveň ekonomické aktivity obyvatelstva — 6,2 %

Kladně koreluje s podílem ekonomicky aktivních obyvatel, s podílem ekonomicky aktivních žen a s podílem obyvatel v produktivním věku (věková skupina 15—59).

VI. Významnost zemědělství v ekonomické aktivitě obyvatelstva — 5,8 %

Komponenta kladně koreluje s podílem obyvatel zaměstnaných v zemědělství, s podílem bytů postavených do r. 1869, a s podílem dočasně nepřítomného obyvatelstva na jedné

solutní hodnotě blízkých 1. Jde o transformaci typu rotace. Z řady metod rotace jsme použili metody VARIMAX (Kaiser 1958), která zachovává původní lineární nezávislost rotovaných komponent. Skóre rotovaných komponent vypočteme analogicky jako

$$G = (B^T B)^{-1} B^T Z,$$

kde G je matice skóre rotovaných komponent.

straně, na straně druhé koreluje záporně s podílem bytů postavených po r. 1945.

VII. *Integrovanost města ve vyšším sídelním útvaru — 5,5 %*

Komponenta kladně koreluje s indikátory podílů ekonomicky aktivního obyvatelstva vyjíždějícího za zaměstnáním mimo místo trvalého bydliště a s podílem obyvatel vyjíždějících jednou denně na jedné straně, na straně druhé záporně koreluje s podílem obyvatel, jimž cesta za zaměstnáním mimo místo trvalého bydliště trvá déle než 1 hodinu. Integrovanost je zde tedy chápána ve smyslu vyjíždky za zaměstnáním.

VIII. *Důležitost nestabilní složky obyvatelstva — 4,6 %*

Kladně koreluje s indikátory důležitosti nestabilní složky obyvatelstva (dočasně nepřítomného a dočasně přítomného obyvatelstva) vzhledem k celkovému počtu trvale bydlícího obyvatelstva.

IX. *Intenzita bytové nouze — 4,5 %*

Kladně koreluje s indikátory bytového nedostatku, konstruovanými z absolutních počtů společně hospodařících domácností a absolutních počtů bytů a s podílem obyvatelstva české národnosti.

X. *Sociální úroveň bydlení — 4,0 %*

Komponenta kladně koreluje s průměrnou velikostí obytné plochy bytu a s průměrnou obytnou plochou per capita na jedné straně, záporně s průměrným počtem osob na jednu obytnou místnost na straně druhé. Zbylé dvě rotované hlavní komponenty se nepodařilo spolehlivě interpretovat.

Skóre $\langle 0, \lambda \rangle$ všech 12 rotovaných hlavních komponent sloužila jako vstupní data pro klasifikaci, jejíž výsledky dále uvádíme.

Metoda klasifikace

Pro klasifikaci zkoumaného souboru 215 měst bylo použito modifikované metody wroclavské taxonomie [Kraus 1973] s následným použitím mnohonásobné diskriminační analýzy [Casetti 1964]. Za míru podob-

nosti objektů O_r a O_s charakterizovaných indikátory jsme zvolili vzdálenost d definovanou v n -rozměrném ortogonálním euklidovském prostoru. Klasifikované objekty považujeme za body v tomto prostoru

$$\begin{aligned} O_r &= (x_{1r}, x_{2r}, \dots, x_{nr}), \\ O_s &= (x_{1s}, x_{2s}, \dots, x_{ns}). \end{aligned}$$

Jejich vzdálenost d_{rs} je definována vztahem

$$d_{rs} = \sqrt{\sum_{i=1}^n (x_{ir} - x_{is})^2}$$

Lze dokázat, že vzdálenost bodů O_r a O_s definovaná v n -rozměrném standardizovaném euklidovském prostoru

$$d_{rs} = \sqrt{\sum_{i=1}^n (z_{ir} - z_{is})^2}$$

je totožná se vzdáleností d_{rs} definovanou v n -rozměrném euklidovském prostoru n hlavních komponent, jestliže skóre hlavních komponent (matice F) jsou standardizována na rozptyl rovný λ

$$d_{rs} = \sqrt{\sum_{p=1}^n (f_{pr} - f_{ps})^2}$$

Použijeme-li k výpočtu vzdálenosti pouze $m < n$ hlavních komponent, dopustíme se tím jisté chyby. Tato chyba bude tím menší, čím větší bude podíl rozptylu vyčerpaného m hlavními komponentami na celkovém rozptylu všech indikátorů rovném n . Praxe ukázala, že při podílu vyšším než 80 % je tato chyba nepodstatná.

Klasifikace tedy vycházela z výpočtu matice vzdálenosti D mezi všemi dvojicemi měst. Výsledná matice je trojúhelníková ($d_{rs} = d_{sr}$) a obsahuje $N(N-1)/2$ prvků. Před započítím klasifikace je důležitá volba vhodného kritéria podobnosti — vzdálenosti d . Čím bude d větší, tím bude počet skupin menší a klasifikace obecnější. Naopak při malém d obdržíme velký počet značně vnitřně homogenních skupin. Volba velikosti d je odvislá od účelu klasifikace. Existují sice exaktní metody k určení velikosti d , tyto metody však většinou nevyhovují praktickým účelům. Obvykle tedy postupujeme zkusmo. Zvolíme několik hodnot d a z výsledků zpravidla již snadno určíme vhodnou velikost míry podobnosti, která je v daném případě vhodná. Metoda wroclavské taxonomie vychází z matice vzdálenosti D . Její algoritmus je možno shrnout do následujících kroků: 1. Každému objektu určenému m indikátory

přičítáme číslo H_j vyjadřující hustotu objektů v místě j -tého objektu. H_j definujeme vztahem

$$H_j = \sqrt{\sum_{K=1}^n \frac{1}{(d_{jk} + 1)^2}}$$

Pro $j \neq k$, kde d_{jk} je vzdálenost v euklidovském prostoru. Hustota H_j roste s hustotou objektů v místě j -tého objektu.

2. Objekty seřadíme sestupně podle hodnoty H_j .
3. Zvolíme číslo d (kritérium podobnosti).
4. Z řady objektů vezmeme první objekt, který ještě nepatří do žádné skupiny. Tento bod považujeme za střed skupiny (centroid).
5. Do této skupiny zařadíme všechny objekty, jejichž vzdálenost od centroidu není větší než číslo d a které neleží v žádné skupině. Pokud takové objekty již neexistují a není-li vyčerpána jejich řada, pak se vrátíme na 4. Jinak je chod algoritmu skončen.

Výslednou klasifikaci je možno již použít. Obvykle se však snažíme řešení optimalizovat vzhledem k různým kritériím. Jedním z matematicky nejelegantnějších je optimalizace na základě maximální diskriminace skupin. V tomto případě maximalizujeme poměr meziskupinového a vnitroskupinového rozptylu.

Postupujeme tak, že vyjdeme z původní klasifikace N objektů do m skupin. Vypočteme kritérium R . Na základě dále popsaného algoritmu změníme obsazení skupin tak, že R vzroste. Dostaneme tak novou klasifikaci, spočítáme opět R , upravíme obsazení skupin a tento postup opakujeme tak dlouho, dokud R nedosáhne maxima. Tomuto postupu říkáme mnohonásobné diskriminační iterace. Jak ukázaly četné pokusy (Casetti 1964), dospějeme z poměrně libovolné původní klasifikace N objektů do m skupin po konečném počtu kroků vždy ke stejné optimální klasifikaci (ve smyslu maximalizace R) N objektů do m skupin. Původní klasifikace nemusí být tedy příliš kvalitní. Je však výhodné, není-li od optimální klasifikace příliš odlišná — uspoříme tím mnoho hodin strojového času na samočinném počítači. Metoda mnohonásobné diskriminační analýzy spočívá v těchto základních operacích:

Budiž:

X) matice původních dat rozměru $N \times p$

s elementy X_{htj} , kde

$h = 1, 2, \dots, m$ počet skupin

$j = 1, 2, \dots, p$ počet parametrů

$i = 1, 2, \dots, n_h$ počet objektů v h -té skupině

X_{ht} řádkový vektor = $(X_{ht1}, X_{ht2}, \dots, X_{htp})$

$\bar{X}_{h..j}$ skalár — průměr j -tého parametru v h -té skupině

$\bar{X}_{..j}$ skalár — průměr j -tého parametru ze všech N objektů

($\bar{X}$) matice, kterou dostaneme z matice (X), jestliže dosadíme za každé X_{htj} příslušný průměr $\bar{X}_{h..j}$

($\bar{X}$) matice, kterou dostaneme z matice (X), jestliže za každé X_{htj} dosadíme příslušný průměr $\bar{X}_{..j}$

(V) je rozdíl matic ($\bar{X}$) — ($\bar{X}$)

(W) je rozdíl matic (X) — ($\bar{X}$)

Odtud plyne, že v původním p -rozměrném prostoru je každý objekt reprezentován vektorem X_{ht} , každá skupina vektorem $\bar{X}_{ht} = (\bar{X}_{h..1}, \bar{X}_{h..2}, \dots, \bar{X}_{h..p})$ a celková klasifikace vektorem $X_{..} = (X_{..1}, X_{..2}, \dots, X_{..p})$.

Nyní budeme hledat lineární kombinace parametrů ($X_1 \dots X_p$), které budou vyhovovat optimální diskriminaci mezi skupinami v novém diskriminačním prostoru. Definujeme nyní y_{htz} jako hodnotu z -té diskriminační funkce pro i -tý objekt v h -té skupině

$y_{htz} = b_{z1}X_{ht1} + b_{z2}X_{ht2} + \dots + b_{zp}X_{htp}$ čili $(Y_z) = (B_z) \cdot (X)$, kde b jsou koeficienty, které chceme odvodit. Dále definujeme matici skóre (Y_z) jako

$$(Y_z) = (X) \cdot (B_z)$$

Průměr y_{htz} pro h -tou skupinu je tedy

$$\bar{y}_{h..z} = \sum_i y_{htz} / n_h$$

a všeobecný průměr všech y_{htz} na z -té funkci

$$\bar{y}_{..z} = \sum_h \sum_i y_{htz} / N$$

dále definujeme ($\bar{Y}_z$) jako vektor, který dostaneme z (Y_z), když za každé y_{htz} dosadíme příslušný průměr $\bar{y}_{h..z}$. Možno dokázat, že

$$(\bar{Y}_z) = (X_{..}) \cdot (B_z)$$

Analogicky bude ($\bar{Y}_z$) vektor, který dostaneme z (Y_z), když dosadíme $\bar{y}_{..z}$ za každé y_{htz} ; tento vektor je roven ($\bar{X}$) $\cdot$ (B_z). Těmito operacemi jsme původní p -rozměrný prostor transformovali pomocí lineárních diskriminačních funkcí do z -rozměrného diskriminačního prostoru.

Vlastní diskriminační procedura se může lišit v závislosti na:

- typu transformace — v našem případě jde o lineární transformaci, transformace však může být i kvadratická atd.;
- pravidla, jímž určujeme reprezentativní bod skupiny — obvykle za něj považujeme centroid skupiny, tj. bod daný aritmetickým průměrem jednotlivých souřadnic;
- způsobu měření relativní polohy bodů — v našem případě použijeme vzdálenost v euklidovském mnohorozměrném prostoru.

Řešení mnohonásobných diskriminačních funkcí vyžaduje maximalizaci podílu R

$$R = \frac{\frac{1}{N} \sum_h n_h (\bar{y}_{h..} - \bar{y}_{..z})^2}{\frac{1}{N} \sum_h \sum_i (y_{hiz} - \bar{y}_{..z})^2}$$

Čitatel ve výrazu se vztahuje k průměrné vzdálenosti od centroidu ke středu (těžišti) všech bodů. Jmenovatel pak vyjadřuje průměrnou vzdálenost od bodů (představujících klasifikované objekty) k těžišti. Maximalizujeme-li tedy podíl R , maximalizujeme tím variabilitu mezi skupinami ve srovnání s celkovou variabilitou.

Vyjádříme-li R pomocí výše definovaných vektorů, dostaneme

$$R = \frac{(\bar{Y}_{..z} - \bar{Y}_z)^{tr} (\bar{Y}_{..z} - \bar{Y}_z)}{(Y_z - \bar{Y}_z)^{tr} (Y_z - \bar{Y}_z)} = \frac{B_z^{tr} V^{tr} V B_z}{B_z^{tr} W^{tr} W B_z}$$

Nyní hledáme řešení pro (B_z) , které maximalizuje R . Řešení je analogické řešení hlavních komponent

$$(G - \lambda I) = 0$$

kde

$$G = (W^{tr} W)^{-1} V^{tr} V \quad \text{a} \quad \lambda = \frac{B_z^{tr} V^{tr} V B_z}{B_z^{tr} W^{tr} W B_z}$$

Řešení dává p vlastních čísel a jim odpovídajících vlastních vektorů. Výsledkem je matice zátěží lineárních diskriminačních funkcí a příslušné vlastní číslo lze považovat za míru diskriminační síly dané funkce. Podíl $\lambda_j / \sum \lambda_r$ dává pak relativní část rozptylu mezi skupinami vyčerpaného z -tou funkcí.

Takto vypočtené lineární diskriminační funkce mají následující vlastnosti:

- Jsou ortogonální.
- Počet lineárně nezávislých diskriminačních funkcí je maximálně roven p .

- Vektory (Y_z) jsou ortogonální jen tehdy, jsou-li vstupní data ortogonální (např. skóre hlavních komponent).
- Zátěže diskriminačních funkcí a jejich skóre je možno interpretovat podobně jako hlavní komponenty.
- Některé diskriminační funkce je možno vynechat vzhledem k jejich malému příspěvku k celkové diskriminaci (analogicky jako u hlavních komponent).

Po tomto kroku možno přistoupit k optimalizaci původní klasifikace. Mějme tedy původní klasifikaci N objektů do m skupin, tj.

$$A_s = (a_s, t_i, i),$$

kde s jest daná klasifikace a i objekt, který je v této klasifikaci zařazen do skupiny t_i . Nyní aplikujeme diskriminační proceduru D , aby-chom dostali novou klasifikaci. Toto možno zapsat jako

$$DA_s = (DA_s, t_i, i) = A_{s+1} = (a_{s+1}, t_i, i)$$

Tím jsme zároveň ukázali, že dvě klasifikace jsou ekvivalentní, jestliže

$$a_s, t_i, i = a_{s+1}, t_i, i$$

pro všechna

$$i = 1, 2, \dots, N$$

Z toho plyne, že opakované použití diskriminační procedury může být vyjádřeno jako mocnina D :

$$\begin{aligned} DA_1 &= A_2 \\ DA_2 &= A_3 = D(DA_1) = D^2 A_1 \\ DA_{N-1} &= A_N = D^{N-1} A_1 \end{aligned}$$

Tato řada kroků se nazývá diskriminačními iteracemi. Nyní se naskytá otázka, kdy bude dosažena limitní klasifikace. V posloupnosti klasifikace

$$A_1, A_2, \dots, A_v, \dots, A_w, A_{w+1}$$

jest A_w limitní klasifikací, když

$$A_z = A_w \quad \text{pro} \quad z \geq w$$

a

$$A_z \neq A_w \quad \text{pro} \quad z < w$$

(w je konečné kladné celé číslo)

Bylo dokázáno, že pro konečný počet objektů původní klasifikace bude konečným počtem iterací dosaženo klasifikace limitní.

Z toho plyne algoritmus pro dosažení limitní, tj. optimální klasifikace z klasifikace původní:

- Z původní klasifikace vypočteme lineární diskriminační funkce a skóre objektů na těchto funkcích.
- Vypočteme vzdálenosti každého objektu

od každého centroidu skupiny v diskriminačním prostoru.

3. Pro každý objekt vyhledáme nejkratší z těchto vzdáleností.
4. Na základě toho přiřadíme daný objekt do nejbližší skupiny, tím data přetřídíme a hned vypočteme nové lineární diskriminační funkce. Tyto kroky opakujeme tak dlouho, dokud nedosáhneme limitní — optimální klasifikace.

Kromě toho je možno právě popsané metody diskriminačních iterací použít například k těmto účelům:

- a) *Zhodnocení relativní kvality klasifikace.* Kvalitu možno hodnotit například podle počtu iteračních kroků, které vedou k dosažení limitní klasifikace nebo podle rozdílu mezi původní a limitní klasifikací vzhledem k indexům.

$$R^0 = \frac{\text{variabilita uvnitř skupin}}{\text{variabilita celková}} = \\ = \frac{1}{1 + \sum_r \lambda_r^0} ; \quad 0 < R^0 < 1$$

$$R = \frac{\text{variabilita mezi skupinami}}{\text{variabilita uvnitř skupin}} = \\ = \sum_r \lambda_r ; \quad 0 < R < \infty$$

Je-li R^0 malé, pak variabilita uvnitř skupin je relativně nízká, tudíž rozdíly mezi skupinami musí být vysoké. Naproti tomu nízké hodnoty R odpovídají špatné klasifikaci a vysoké hodnoty R klasifikaci dobré. Konečně další mírou kvality je srovnání počtu objektů v různých skupinách klasifikace.

- b) *Identifikaci jader skupin.* Jádra představují ty objekty, které nezměnily svou příslušnost k dané skupině po celou dobu iterací.
- c) *Testování validity klasifikace* vzhledem ke vstupním údajům získaných v různých časových a prostorových bodech.

Aplikace metody diskriminačních iterací na naše údaje vedla k podstatnému zpřesnění původní typologie založené na wrocławské metodice. Rozdělení měst do skupin bylo po aplikaci diskriminační analýzy více v souladu i s běžným povědomím o charakteru jednotlivých měst. Tento výsledek potvrzuje zkušenosti získané i jinými autory [Linhart, Rak, Vozenílek 1977].

Klasifikace měst do skupin

V této části podáváme výsledky jedné z výsledných klasifikací souboru českých měst. Zvolili jsme typologii vycházející z míry podobnosti $d = 11$. Tato typologie nepatří k nejpodrobnějším. Nepatří však ani k těm, v nichž vysoký stupeň generalizace již činí výsledky málo zajímavými a přínosnými, zejména pro plánovací praxi. Zvolená klasifikace sice stírá některá specifika, například nepostihuje skupinu lázeňských měst, která se objevuje v některých podrobnějších typologiích, skupinu regionálních metropolí apod., poskytuje však poměrně dobrý přehled.

Skupina 1 — 27 měst — střední a velká města — obslužná a administrativní centra

Obsazení:

Praha, Kladno, Slaný, Kolín, Kralupy, Mladá Boleslav, Karlovy Vary, Klatovy, Plzeň, Brno, Prostějov, Ústí n. L., Olomouc, Ostrava, Jičín, Turnov, Přerov, Rakovník, České Budějovice, Nymburk, Liberec, Česká Třebová, Louny, Jablonec n. N., Český Těšín, Hradec Králové, Pardubice.

Skupina 2 — 33 měst — malá a střední rozvíjející se průmyslová města

Obsazení:

Vlašim, Neratovice, Ostrov, Stará Role, Hodonín, Blansko, Kuřim, Gottwaldov, Otrokovice, Napajedla, Chotěboř, Kopřivnice, Mohelnice, Vsetín, Rožnov p. R., Uherské Hradiště, Uherský Brod, Žďár n. S., Nové Město n. M., Klášterec n. O., Milevsko, Prachatice, Strakonice, Sezimovo Ústí, Dobříš, Štětí, Slavkov, Zábřeh, Veselí n. M., Pelhřimov, Valašské Meziříčí, Hulín, Hlinsko.

Skupina 3 — 20 měst — malá, převážně stagnující průmyslová města

Obsazení:

Švermov, Nýřany, Dubí, Duchcov, Krupka, Nový Bydžov, Třebachovice, Petřvald, Náchod, Broumov, Červený Kostelec, Hronov, Jaroměř, Semily, Dvůr Králové, Úpice, Tanvald, Železný Brod, Lomnice n. P.

Skupina 4 — 9 měst — malá stagnující moravská města

Obsazení:

Ivančice, Jablunkov, Šenov, Frýdlant n. O., Kravaře, Strážnice, Rychvald, Kojetín, Orlová.

Skupina 5 — 54 měst — malá a střední smíšená města

Obsazení:

Čáslav, Mnichovo Hradiště, Kutná Hora, Rokycany, Kroměříž, Chrudim, Příbor, Hořice, Nová Paka, Litomyšl, Choceň, Vysoké Mýto, Vyškov, Tábor, Domažlice, Mariánské Lázně, Jindřichův Hradec, Třeboň, Poděbrady, Roudnice n. L., Kostelec n. O., Třešť, Ústí n. O., Havlíčkův Brod, Telč, Český Krumlov, Písek, Benešov, Tišnov, Kyjov, Beroun, Český Brod, Hořovice, Mělník, Rychnov n. K., Boskovice, Litovel, Žamberk, Lipník n. B., Hranice, Frenštát p. R., Holešov, Moravské Budějovice, Nové Město n. Met., Soběslav, Sušice, Polička, Přelouč, Třebíč, Velké Meziříčí, Humpolec, Bystřice p. H., Vimperk.

Skupina 6 — 23 měst — malá a střední rozvíjející se průmyslová města typická velkým podílem nové zástavby sídlištního typu

Obsazení:

Kynšperk, Most, Bílina, Frýdek-Místek, Trinec, Havířov, Karviná, Nový Bohumín, Uničov, Hlučín, Jirkov, Habartov, Meziboří, Vratimov, Bystřice n. P., Stochov, Dubňany, Sokolov, Chodov, Studénka, Kadaň, Příbram.

Skupina 7 — 11 měst — malá města: satelity

Obsazení:

Roztoky, Zbraslav, Holice, Lysá n. L., Čelákovice, Dobruška, Říčany, Šlapanice, Brandýs n. L. a Stará Boleslav, Horní Počernice, Radotín.

Zbývající tři skupiny měst jsou lokalizovány převážně v pohraničí, zahrnují však i některá města z centrální části Čech.

Skupina 8 — 11 měst — malá a střední města v pohraničí s dílčí průmyslovou výrobou

Obsazení:

Mimoň, Kraslice, Vrbno p. Pr., Varnsdorf, Nový Bor, Aš, Nejdek, Rumburk, Hrádek n. Nisou, Frýdlant.

Skupina 9 — 6 měst — menší střední města — centra převážně zemědělského okolí

Obsazení:

Bruntál, Vítkov, Tachov, Rýmařov, Stříbro, Mikulov.

Skupina 10 — 21 měst — střední a větší města v pohraničí převážně s komplexní strukturou socioekonomických funkcí

Obsazení:

Žatec, Cheb, Lanškroun, Lovosice, Teplice, Břeclav, Moravská Třebová, Krnov, Jeseník, Chomutov, Jihlava, Děčín, Šternberk, Trutnov, Nový Jičín, Opava, Šumperk, Znojmo, Svitavy, Česká Lípa, Litoměřice.

Některá města sice poněkud méně odpovídají charakteru skupiny, tato závada však tkví v samé podstatě jakékoli generalizace.

Přejdeme nyní k podrobnější charakteristice jednotlivých skupin a pokusme se o interpretaci jejich socioekonomických, demografických struktur.

Podrobná charakteristika a interpretace skupin

Skupiny lze podrobně charakterizovat pomocí průměrných hodnot skóre rotovaných komponent. Skóre dané komponenty udává polohu daného města na stupnici vyjadřující intenzitu vlastnosti, kterou tato komponenta vyjadřuje, tedy například vitalitu města.

Tabulka 1 udává průměrné hodnoty skóre rotovaných hlavních komponent, standardizovaná na jednotkovou směrodatnou odchylku. Uvádíme pouze skóre prvních 10 rotovaných komponent, tj. těch komponent, které byly interpretovány a o něž se budeme při interpretaci opírat. Hodnoty skóre uvádíme pouze schematicky a v následující formě: symbol +++ odpovídá hodnotě skóre menší než +1,5, symbol ++ volíme pro skóre v intervalu od +1,5 do +1,0, symbol + vyhrazuje pro skóre v intervalu od +0,5 do -0,5. Symboly — — —, — — a — mají analogický význam. Skóre označené symboly +++ nebo — — — považujeme za silné nadprůměrná (podprůměrná), skóre označená symboly ++ eventuálně — — považujeme za středně nadprůměrná (podprůměrná), skóre označená symboly + eventuálně — hodnotíme jako mírně nadprůměrná (podprůměrná) a konečně skóre označená symbolem 0 představují skóre průměrná. Vzhledem ke standardizaci skóre na (0,1) by za předpokladu normálního rozložení mělo zhruba 95 % hodnot ležet v intervalu <-2; +2>.

Skupina 1 — střední a velká města, obslužná a administrativní centra

Skupina je typická nadprůměrnou urbanitou, mírně nadprůměrnou terciárností ekonomických aktivit obyvatel a mírně nadprůměrnou ekonomickou aktivitou obyvatel na jedné straně, na druhé straně mírně podprůměrnou vitalitou a mírně podprůměrnou integrovaností ve vyšším sídelním útvaru. Jinými slovy, skupinu tvoří města s dobrou technikou vybaveností, vyšší hustotou zalidnění, většími bytovými domy a nízkým podílem obyvatel zaměstnaných v zemědělství. V městech této skupiny je vyšší podíl lidí se středoškolským a vysokoškolským vzděláním, vyšší podíl lidí zaměstnaných v nevýrobní sféře, dále vyšší podíl domácností vlastních auto, chatu, ledničku a vysavač. Je zde poměrně rozvinutá terciární sféra. Jedná se o skupinu měst, která se v posledních letech vyvíjela obvykle relativně pomaleji a některá z nich nemají pro další vývoj ani výraznější předpoklady demografické, ekonomické a bytové povahy. Nabízejí dostatek pracovních příležitostí, a proto intenzita vyjíždky za prací je nutně omezena. Je zde vesměs silně vyvinutá obslužná a administrativní funkce. Všechna města této skupiny byla a ve velké většině dosud jsou krajskými nebo okresními městy. Vysoce rozvinutá obslužná a administrativní funkce jim vtiskuje řadu charakteristických rysů, které jsme jen zčásti postihli pomocí námi zvolených indikátorů.

Skupina 2 — malá a střední rozvíjející se průmyslová města

Města ve skupině jsou charakteristická mírně nadprůměrnou vitalitou a mírně nadprůměrnou sociokulturní úrovní obyvatel. Urbanitu, „terciárnost“ a „primárnost“ ekonomických aktivit obyvatel možno považovat za mírně podprůměrnou. Naznačuje to, že zmíněná města se v posledních letech dynamicky rozvíjela a že mají předpoklady k dalšímu rozvoji. Jde o města průmyslová s vyšším podílem domácností s vyššími příjmy (vlastnictví auta, chaty, ledničky a vysavače). Existuje zde již většinou tradice určitého speciálního oboru průmyslové výroby, která zajišťuje dostatečnou nabídku pracovních příležitostí a pozitivně ovlivňuje demografickou a bytovou situaci. Technické vybavení

měst této skupiny, hustota zalidnění a velikost bytových domů je na nižší než průměrné úrovni.

Skupina 3 — malá stagnující průmyslová města

Skupina se vyznačuje silně podprůměrnou vitalitou, mírně podprůměrným socioekonomickým statutem, mírně podprůměrnou terciárností ekonomických aktivit obyvatel a sociální úrovní bydlení na jedné straně. Na straně druhé mírně nadprůměrnou integrovaností ve vyšším sídelním útvaru.

Města této skupiny se v minulých letech velmi pomalu rozvíjela. Je zde nedostatečná nabídka pracovních příležitostí, která se projevuje relativně intenzivní vyjíždkou za zaměstnáním. Nižší procento lidí pracujících v terciární sféře a intenzivní vyjíždka za prací svědčí o slabě vyvinuté administrativní a obslužné funkci. Tyto funkce plní jiná města v okolí. Relativní stagnace těchto měst je v mnoha případech podmíněna nepříznivou polohou.

Skupina 4 — malá stagnující moravská města

Tato města se vyznačují podprůměrnou ekonomickou aktivitou obyvatelstva, mírně podprůměrným zastoupením zemědělského obyvatelstva a mírně podprůměrnou důležitostí nestabilní složky obyvatelstva. V městech bydlí převážně lidé s nižší úrovní vzdělání, je zde větší podíl domácností s nižšími příjmy a větší zastoupení národnostních menšin. Ekonomická aktivita obyvatel je nižší v důsledku nízké ekonomické aktivity žen a nedostatečné nabídky pracovních příležitostí.

Skupina 5 — malá a střední smíšená města

Smíšenými městy rozumíme města, v nichž jsou dostatečně významně zastoupeny všechny tři hospodářské sektory (primární, sekundární a terciární). Počet obyvatel pracujících v terciéru je v těchto městech na průměru, naproti tomu počet obyvatel spjatých s primárním sektorem je mírně nad průměrem. Mírně nadprůměrná je též sociokulturní úroveň obyvatel. Ekonomická aktivita obyvatel je nižší než odpovídá průměru. Mírně podprůměrná je též zaměstnanost žen. Města této skupiny jsou v podstatě jakýmsi středním typem měst v podmínkách osídlení ČR.

Tabulka 1. Průměrné hodnoty skóre rotovaných komponent

Zkrácený název komponenty	Skupina									
	1	2	3	4	5	6	7	8	9	10
I. Vitalita	—	+	—	0	0	0	++	—	0	+++
II. Sociokulturní úroveň obyvatel	+	+	—	—	+	—	—	+	0	—
III. Urbanita	++	—	0	—	0	0	+	—	+	—
IV. Terciérnost	+	—	—	—	0	—	0	+	+	+++
V. Ekonomická aktivita	0	0	0	—	—	+++	—	0	+	+
VI. Primárnost	0	—	0	—	+	0	0	—	+	+
VII. Integrovanost	—	0	+	0	0	—	+	+	0	—
VIII. Nestabilní obyvatelstvo	0	0	0	—	+	0	0	—	—	—
IX. Nedostatek bytů	0	0	0	0	0	0	0	0	0	++
X. Soc. úroveň bydlení	0	0	—	0	0	+++	0	0	0	0

Skupina 6 — malá a střední rozvíjející se průmyslová města typická velkými obytnými soubor

Skupina je typická především silně nadprůměrnou vitalitou, dále mírně nadprůměrnou urbanitou a mírně nadprůměrnou integrovaností ve vyšším sídelním útvaru na jedné straně, na straně druhé mírně podprůměrnou ekonomickou aktivitou a mírně podprůměrným socioekonomickým statusem. Jde vesměs o města, na jejichž katastru byla postavena velká sídliště pro pracující v hornictví, metalurgickém a strojírenském průmyslu. Tyto druhy průmyslu jsou zde převažující, což značnou měrou ovlivňuje demografickou a sociální strukturu obyvatelstva. Velký národohospodářský význam měst této skupiny vyvolává intenzivní bytovou výstavbu, což ještě umocňuje jejich dynamický rozvoj, podporovaný také příznivou demografickou strukturou a již založenými růsto-

vými trendy. Mírně nadprůměrná urbanita ukazuje na dobrou technickou vybavenost, relativně husté zalidnění a nadprůměrnou velikost bytových domů. Vzhledem k úzké specializaci měst (pokud se týče jejich ekonomické základny) je zde nutná poněkud vyšší vyjízdka za zaměstnáním mimo město. Podprůměrná ekonomická aktivita obyvatel vyplývá zejména z nízké zaměstnanosti žen, která je typická pro domácnosti horníků.

Skupina se jeví jako dynamická a hospodářsky vysoce významná; i v budoucnosti se bude intenzivně rozvíjet.

Skupina 7 — malá města — satelity

Města s mírně nadprůměrnou sociokulturní úrovní obyvatelstva, mírně nadprůměrnou integrovaností ve vyšším sídelním útvaru, středně podprůměrnou důležitostí nestabilní složky obyvatelstva a mírně podprůměrným zastoupením zemědělců. Jde o malá města tvořící zázemí velkých měst — především

Prahy. Tato města složí téměř výhradně k bydlení obyvatel, kteří vyjíždějí za prací do blízké metropole. Tím je determinován jejich dosavadní i budoucí rozvoj. Vykazují nízkou technickou vybavenost, zástavba má mnohdy charakter rodinných domků a činností vil, to vše se promítá do podprůměrné hodnoty urbanity. Úzkou závislostí na metropoli vysvětlujeme též nadprůměrnou intenzitu terciárních ekonomických aktivit obyvatel. Města slouží především k bydlení obyvatel pracujících v metropoli a nemají prakticky jiné funkce. Úzká závislost na metropoli nutí města-satelity sledovat její vývoj. Některá města byla již metropolí pohlcena (Zbraslav, Horní Počernice a Radotín Praha). Mírně nadprůměrná primárnost ekonomických aktivit dokumentuje vyšší podíl zemědělského obyvatelstva a odráží minulý vývoj a lokalizaci těchto měst.

Zbývající tři skupiny měst se vyznačují poněkud odlišnou demografickou strukturou obyvatelstva (vesměs nadprůměrně zastoupení mladších věkových kategorií) a příznivější bytovou situací z hlediska zalidněnosti bytů a plošné výměry bytů. Z hlediska sociální struktury a charakteru bydlení jsou tato města velmi různorodá, jak ostatně vyplývá z průměrných komponentních skóre uvedených v tabulce 1.

Nejvýznamnější je zde skupina středních a větších pohraničních měst. V těchto městech, podobně jako v některých dalších pohraničních městech zařazených do skupiny 1, se demografická struktura, ovlivněná v některých případech dosti podstatně odsunem německého obyvatelstva po roce 1945, dnes podstatně přiblížila situaci v ostatních městech. Jde vesměs o významná centra průmyslové výroby s vyvinutou administrativní a obslužnou funkcí. Komplexnost těchto městských funkcí se odráží ve zvýšeném podílu osob se středoškolským a vysokoškolským vzděláním a ve vyšším zastoupení zaměstnaneckých složek obyvatelstva s dobrou životní úrovní. Města nabízejí svým obyvatelům dostatek pracovních příležitostí, a proto je i rozsah dojížděky za prací mimo město poměrně omezený.

Charakteristiky zbývajících dvou kategorií měst si čtenář odvodí z tabulky 1 snadno samostatně, popřípadě se může obrátit na původní práci kde jsou výsledky publikovány in extenso [Link 1975].

Závěr

Všech deset skupin, zahrnujících všech 215 měst s více než 5000 obyvateli v ČSR, představuje základní typy měst, která ohraničují socioekonomickou diferenciaci českého osídlení v roce 1970. Klasifikace slouží spíše k získání kvalifikovaného přehledu než k hlubší analýze příčin a souvislostí, jichž jsme se ve stati pouze dotkli.

Literatura

- Anderson, T. W.: *An Introduction to Multivariate Statistical Analysis*. New York, John Wiley, 1958.
- Bartlett, M. S.: *Tests of Significance in Factor Analysis*. BJP Stat. Sec., 3, 1950.
- Casetti, E.: *Multiple Discriminant functions*. Technical Report No. 11, Computer Applications in the Earth Sciences Project, Department of Geography, Northwestern University, 1964.
- Casetti, E.: *Classificatory and Regional Analysis by Discriminant Iterations*. Technical Report No. 12, Computer Applications in the Earth Sciences Project, Department of Geography, Northwestern University, 1964.
- Harman, H.: *Modern Factor Analysis*, Chicago and London, The University of Chicago Press 1967.
- Kaiser, H. F.: *The Varimax Criterion for Analytic Rotation in Factor Analysis*. Psychometrika, 23, 1958.
- Kraus, P.: *Obecný taxonomický program*. Praha VÚVA, 1973 (rotaprint).
- Link, J.: *Klasifikace měst z hlediska jejich socioekonomických, demografických a bytových charakteristik v ČSR r. 1970*. Praha VÚVA, 1975.
- L'nhart, J. — Rak, V. — Voženilek, J.: *Sociální aspekty ekologické zónace hlavního města Prahy*. Sociologický časopis 13, 1977, č. 1, s. 94.
- Moser, C. A., Scott, W.: *British Towns: A Statistical Study of their Social and Economic Differencies*. London, Oliver and Boyd, 1964.
- Racine, J. B., Cavalier, M.: *Ecologie factorielle et attributs géographiques*. Cahiers de géographie de Québec, 16, 1972.

Резюме

Линк, Й.: Социо-экономическая классификация городов в ЧСР

В статье приводятся результаты классификации 215 городов ЧСР, численность населения которых к 1970 г. превысила 5000 человек. Классификация основана на 57 показателях, отражающих социальную структуру населения, демографический состав, оснащенность

квартир и общее состояние жилфонда в городах. Показатели были выведены на основании переписи населения 1970 г.

На основании анализа главных компонентов исходные показатели были сведены к 12 синтетическим показателям, которые охватывают 82% рассеяния всех наблюдаемых переменных. Синтетические показатели выводились на основании использования главных компонентов, для уточнения теоретической модели простой структуры была использована максимизация критерия VARIMAX.

Десять наиболее важных главных компонентов были сформулированы следующим образом:

1. Демографический и экономический потенциал города
2. Социо-культурная зрелость населения
3. Степень урбанизации города
4. Место «третичного» сектора в экономической активности населения
5. Общий уровень экономической активности населения
6. Место сельского хозяйства в занятости населения
7. Включение города в агломерацию
8. Место, занимаемое нестабильным компонентом населения
9. Интенсивность жилищного кризиса
10. Социальный уровень жилфонда

Эти компоненты и два следующих компонента, которые не удалось сформулировать однозначно, были использованы как исходные данные для нумерической таксономии. Классификация основана на модифицированном методе так наз. Вроцлавской таксономии, последовательно уточняемой при помощи множественного дискриминационного анализа. Далее в статье описывается разделение городов на 10 основных групп и рассматриваются связи этой классификации с развитием социально-экономической системы ЧСР в последние годы.

Summary

Link J.: Socio-economic Classification of Towns in the Czech Socialist Republic (CSR)

The paper presents the results of the classification of 215 towns in the CSR whose number of inhabitants has, by 1970, surpassed

five thousand. The classification is based on 57 indicators covering the social structure of the population, the demographic structure, household equipment and the general housing situation in the towns. The indicators were determined in consideration of the 1970-census results.

By applying the principal component analysis, the original indicators collapsed into twelve synthetic indicators (component scores) exhausting 82% of the total dispersion of the battery of variables. The scores were derived from rotated principal components; an approximation to the theoretical model of the simple structure was attained by maximizing the VARIMAX criterion.

Ten most relevant principal components were interpreted in the following way:

1. demographic and economic potential of the town (vitality),
2. socio-cultural level of the inhabitants,
3. urbanization level of the town,
4. significance of the tertiary sector in the population's economic activity,
5. general level of the population's economic activity,
6. significance of agriculture in the population's economic activity,
7. integration of the town into a higher settlement unit,
8. relevance of the unstable population component,
9. intensity of housing shortage,
10. social level of habitation.

These and the two remaining components (which defied all unambiguous interpretation) were used as input data for numerical taxonomy. The classification is based on the modified Wroclaw-taxonomy method; iteratively it is made more precise with the aid of a multiple discrimination analysis. The paper goes on to describe the breakdown of the towns into ten principal groups and to discuss the relations of this classification to the most recent developments of the CSR's socio-economic system.