

REGIONÁLNÍ DIFERENCIACE REPRODUKČNÍHO CHOVÁNÍ V ČESKU V LETECH 2015, 2019 A 2023: DESKRIPTIVNÍ A SHLUKOVÁ ANALÝZA

Kristýna Dvořáková¹⁾ – Filip Hon²⁾

REGIONAL DIFFERENTIATION OF REPRODUCTIVE BEHAVIOUR IN CZECHIA IN 2015, 2019, AND 2023: DESCRIPTIVE AND CLUSTER ANALYSIS

Abstract

This article examines regional differences in reproductive behaviour across Czech districts in 2015, 2019, and 2023. Its aim is to describe spatial disparities in selected indicators, assess their temporal stability, and identify groups of districts with similar reproductive patterns. The analysis draws on data from the Demographic Yearbook of Districts published by the Czech Statistical Office and uses descriptive statistics, correlation analysis, and hierarchical cluster analysis based on components derived through principal component analysis. The results confirm there is substantial spatial variability in the observed indicators as well as stable relationships between them. The cluster analysis identified five groups of districts, with the Northwestern Region forming the most internally homogeneous cluster, characterised by a low maternal age at first birth, a high share of nonmarital births, elevated levels of induced abortions for non-medical reasons, and a higher proportion of third- and higher-order births. The remaining clusters represent distinct types of reproductive behaviour, collectively demonstrating the considerable diversity of reproductive patterns across Czechia.

Keywords: Reproductive behaviour, Regionalisation, Cluster analysis

Demografie, 2026, 68(1): 52–71

DOI: <https://doi.org/10.54694/dem.0379>

ÚVOD

Reprodukční chování obyvatel Česka v posledních letech vykazuje výrazné výkyvy. Například při pohledu

na celkovou úroveň plodnosti v období celosvětové pandemie onemocnění Covid-19 dosáhla úhrnná plodnost v roce 2021 hodnoty 1,83, již v roce 2024 však úhrnná plodnost poklesla na 1,37 (ČSÚ, 2025).

1) Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky. Kontakt: dvok17@vse.cz.

2) Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky. Kontakt: filip.hon@vse.cz.

Tento výrazný propad se stal předmětem nejen odborných diskusí, ale také mediálního zájmu, a téma demografického vývoje tak rezonuje stále výrazněji i v širší veřejnosti.

Pokles úrovně plodnosti je často vysvětlován kombinací faktorů, mezi něž patří zejména ekonomická nejistota, vliv mezinárodních konfliktů či zhoršující se environmentální situace (Paloncyová a kol., 2021; Waldaufová – Štátná, 2025). Tyto faktory se promítají do rozhodovacího procesu jednotlivců a domácností o zakládání rodiny a počtu dětí.

Reprodukční chování v tomto článku chápeme jako soubor demografických procesů a rozhodování souvisejících s realizací rodičovství, který zahrnuje nejen intenzitu plodnosti, ale také načasování vstupu do mateřství, strukturu porodů podle pořadí, rodinný kontext početí a narození dítěte a způsoby reprodukční regulace, včetně vyskytu umělých přerušení těhotenství (zkratka UPT). Úhrnná plodnost přitom představuje pouze jeden syntetický indikátor, který nepostihuje zásadní změny probíhající v české společnosti. Současné reprodukční chování je výrazně utvářeno i dalšími procesy.

Časování plodnosti v Česku se dlouhodobě posouvá do vyššího věku, mezi zamýšleným a skutečným věkem při narození prvního dítěte se vytváří výrazná mezera. Odklad prvního porodu významně snižuje prostor pro další děti. Tento rozdíl je způsoben nejen individuálními rozhodnutími, ale také strukturálními faktory, jako je prodlužující se vzdělávání a nestabilita partnerských vztahů. Většina české populace stále preferuje dvoudětný model rodiny a za ideální věk pro vstup do rodičovství považuje dobu před třicátými narozeninami. Reálné reprodukční plány však zůstávají skromnější. Důvodem jsou přetrvávající ekonomické nejistoty, vysoké ceny bydlení a nedostatečná institucionální podpora rodin. Mladší generace zároveň čelí řadě překážek, mezi něž patří obtížnost najít stabilního partnera, časová náročnost péče o děti a také širší nabídka možností osobního rozvoje, které často soutěží s rodičovstvím jako životní prioritou (Slabá a kol., 2024; Waldaufová – Štátná, 2025).

Dále je utvářeno například mimomanželským rodičovstvím, jehož nárůst je jedním z klíčových rysů post-socialistické transformace. Odborné studie ukazují, že podíl dětí narozených mimo manželství v Česku vzrostl v posledních desetiletích několikanásobně,

především v souvislosti s posilující individualizací a změnami hodnotových orientací (Stípková, 2015).

Podobně významnou součástí reprodukčního chování je i potratovost, zejména intenzita záměrně vyvolaných přerušení těhotenství, která doplňuje informaci o reprodukčních preferencích, normách a rodinných strategiích. Přehledové analýzy ukazují, že česká potratovost se v posledních letech stabilizovala na nízké úrovni (Kocourková a kol., 2023).

Důležitým faktorem při hodnocení reprodukčního chování je také to, jak časté jsou v populaci rodiny se třemi a více dětmi a jakým způsobem se promítají do reprodukčních drah jednotlivých generací. Dostupné analýzy ukazují, že míry plodnosti třetího pořadí se v posledních dekadách sice mírně snížily, avšak zůstávají stabilnější než plodnost nižších pořadí a neprošly šokovým propadem v 90. letech. Navíc se posouvají do vyššího věku žen klidně až k 35 letům, zatímco v roce 1990 vrcholila plodnost třetího pořadí kolem 28 let (Hon, 2020).

Všechny tyto souvislosti ukazují, že pro pochopení regionální diferenciace reprodukčního chování je nezbytné zahrnout nejen úhrnnou plodnost, ale i ukazatele mimomanželské plodnosti, plodnosti vyšších pořadí, časování vstupu do rodičovství a také intenzitu UPT. Jedná se o ukazatele, které jsou na okresní úrovni dostupné a jejich kombinace umožňuje zachytit komplexní a v posledních letech rychle se měnící podobu reprodukčních procesů v českých okresech podstatně lépe, než pohled pouze na některý konkrétní ukazatel.

Cílem tohoto článku je rozšířit známé skutečnosti o nedávném vývoji reprodukčního chování v Česku o detailní regionální analýzu na úrovni okresů, a to s využitím moderních vícerozměrných statistických metod. Podle našich znalostí podobně podrobná aktuální regionální studie dosud v odborné literatuře není dostupná.

Konkrétněji lze definovat tři hlavní cíle tohoto článku. Prvním cílem je detailně popsat regionální diferenciaci reprodukčního chování v okresech Česka na základě vybraných demografických ukazatelů. Druhým cílem je analyzovat, jak se tyto regionální vzorce změnily mezi lety 2015, 2019 a 2023, a posoudit stabilitu či dynamiku sledovaných charakteristik v čase. Třetím cílem je identifikovat skupiny okresů

s obdobnými reprodukčními vzorci pomocí shlukové analýzy a charakterizovat jejich specifické rysy. Studie tak přispívá k hlubšímu pochopení prostorového rozložení reprodukčních vzorců v Česku a vytváří rámec pro následnou diskusi o jejich potenciálních socio-ekonomických a kulturních souvislostech.

Výzkumné otázky jsou na základě uvedených cílů definovány následovně. Článek zkoumá, jaké regionální rozdíly v reprodukčním chování okresů Česka lze pozorovat v letech 2015, 2019 a 2023 na základě vybraných ukazatelů. Text dále porovnává, jak se tyto rozdíly měnily ve sledovaném čase a do jaké míry vykazují sledované ukazatele stabilitu či dynamiku mezi uvedenými roky. Vícerozměrná analýza následně řeší problematiku seskupení okresů pomocí shlukové analýzy, a zároveň na základě jakých konkrétních zákonitostí z hlediska reprodukčního chování jsou okresy členěny do právě takových shluků.

Příspěvek je rozčleněn do několika částí, nejprve je zde uvedena rešerše seznamující zejména se staršími regionálními analýzami, jejich zjištěními a obecným přehledem využití vícerozměrné metody shlukové analýzy k regionalizaci v demografii. Druhá část článku je věnována metodologii ohledně dat i následně popisu konkrétní realizace shlukové analýzy, která samozřejmě právě zdrojová data využívá při výpočtu. Nejrozsáhlejší jsou následně analytické části, kde první zahrnuje deskriptivní statistiku využitých indikátorů reprodukčního chování a druhá se již věnuje výsledkům shlukové analýzy.

REŠERŠE

Deskriptivní i shluková analýza reprodukčních charakteristik okresů Česka není v odborné literatuře novým postupem, naopak představuje etablovanou součást regionální demografie. Shluková analýza dokonce umožňuje identifikovat územní jednotky s obdobnými demografickými rysy a vytvářet tak skupiny s vnitřní homogenitou a vnější odlišností. Z metodologického hlediska lze tento postup chápat jako specifický typ regionalizace, tedy procesu, při němž jsou územní celky členěny do regionů na základě sdílených charakteristik či vzorců chování (Fialová, 2018). Regionalizace v tomto pojetí nepředstavuje pouhé geografické vymezení, ale analytický nástroj zachycující prostorovou strukturu demografických jevů.

Regionalizace pomocí zvolených metod v tomto článku je pak běžná v demografii obecně, nejen v případě okresů. Shluková analýza je využívána nejen v kontextu Česka, ale i na mezinárodní úrovni, například příspěvek s názvem *Time Series Clustering Of Fertility in European Union* se zabývá shlukováním členských zemí Evropské unie na základě vývoje úhrnné plodnosti mezi lety 1965–2021 (Bakuncová – Löster, 2023).

K tvorbě shlukové analýzy jsou kromě indikátorů plodnosti, průměrného věku matek při porodu, případně i potratovosti, využívány také jiné demografické charakteristiky. Například Ondřej Šimpach a Marie Šimpachová Pechrová v příspěvku *Cluster Analysis Of EU's Regions According to Demographics Criteria* pomocí dalších ukazatelů, mezi které patří mimo jiné střední délky života, počty zemřelých či populace 85letých a starších, sestavili shlukovou analýzu napříč administrativními jednotkami NUTS2 členských států EU na základě dat z let 2004 a 2012. To dokládá, že vícerozměrné metody jsou standardním nástrojem studia regionální demografie a dále také, že dlouhodobé regionální disparity jsou doložené i v jiných demografických procesech (Šimpach – Šimpachová Pechrová, 2016).

Na okresní úrovni existuje řada významných studií, které se v minulosti zabývaly prostorovou diferenciací plodnosti. Následující odstavce shrnují klíčová zjištění, jež k tomuto tématu dosavadní odborná literatura přináší.

Vývoj regionální diferenciace plodnosti v devadesátých letech zkoumá například Bartoňová (1999) ve článku *Vývoj regionální diferenciace věkové struktury se zřetelem k územním rozdílům ve vývoji reprodukce v České republice* rozlišuje čtyři typy okresů v Česku podle vývoje úhrnné plodnosti na přelomu 80. a 90. let. První typ zahrnuje okresy s dlouhodobě podprůměrnou plodností, kde po roce 1992 dochází k výraznému poklesu; typické jsou pro velká města. Druhý typ odpovídá celostátnímu průměru, s pozvolným poklesem plodnosti do začátku 90. let a jejím rychlým snižováním po roce 1993. Třetí typ vykazuje nadprůměrnou úhrnnou plodnost, přestože jeho vývoj jinak odpovídá typu druhému; zahrnuje zázemí měst, průmyslové pohraničí i zemědělské okresy. Čtvrtý typ tvoří převážně zemědělské okresy s vysokou plodností přetrvávající ještě v roce 1991, která se výrazněji snižuje až po roce

1993. Studie dále upozorňuje na podstatně častější mimomanželskou plodnost v severovýchodní části republiky již od roku 1990 (Bartoňová, 1999). Obecně lze nicméně konstatovat, že změny reprodukčního chování v devadesátých letech nevedly k výraznějšímu rozvířání regionálních rozdílů v úrovni plodnosti. Spíše než prohlubování diferenciace docházelo k postupné proměně a přeskupování prostorového rozložení plodnosti, tedy k určitému vyrovnávání rozdílů mezi regiony (Bartoňová, 2001). Tento vývoj potvrzují také analýzy zaměřené na generační perspektivu, které poukazují na útlum regionálních kontrastů i při sledování rozdílů mezi kohortami (Rychtaříková, 2007).

Šídlo (2008) ve svém článku *Faktory ovlivňující regionální diferenciaci plodnosti v Česku na počátku 21. století* zdůrazňuje několik klíčových prostorových vzorců. Článek poskytuje důležitý kontext pro porozumění regionálním rozdílům plodnosti v Česku před nástupem výraznější rekuperace plodnosti a diverzifikace reprodukčního chování, kterou autor popisují v pozdějších pracích. Podle autora se tehdejší výsledky analýz shodují s „všeobecně vnímaným pohledem“ na regionální uspořádání plodnosti v Česku. Studie konstatuje postavení městských okresů, které se v rámci zkoumaného období jasně vymezují jako specifická skupina z hlediska reprodukčního chování. Současně se jako problematický region ukazuje severozápadní část Česka. Dalším výrazným zjištěním studie je zachování tradičních kulturních norem v určitých částech země, tyto tradiční vzorce jsou typické zejména pro Vysočinu a střední Moravu, které se tak odlišují od jiných regionů svým konzervativnějším charakterem reprodukčního chování (Šídlo, 2008).

Podrobněji je následně vhodné podívat se na některé studie vydané blíže současnosti. Kurkin ve své disertační práci *Faktory plodnosti a regionální diferenciace plodnosti v České republice po roce 1991* pomocí shlukové analýzy potvrzuje, že vývoj úhrnné plodnosti v českých okresech není homogenní, ale vytváří několik relativně stabilních regionálních vzorců. Výrazně se vyčleňuje severozápadní část Česka, která tvoří specifický shluk se zvýšenou intenzitou plodnosti. Shluk s okresy na jižní Moravě prošel přesunem z nadprůměrné plodnosti k podprůměrné. Další výrazný vzorec představují regiony v zázemí Prahy, které přešly z průměrných hodnot k nadprůměrné plodnosti, a rovněž některé okresy jihozápadních a západních

Čech, jež se z původně podprůměrných hodnot postupně konvergovaly směrem k průměru. Tyto výsledky ukazují, že český prostor se v období 1991–2014 diferencuje do několika jasně profilovaných trajektorií (Kurkin, 2015).

Ve studii *Regionálne rozdiely v charaktere plodnosti v Česku a na Slovensku* autoři analyzují prostorovou diferenciaci rekuperace plodnosti a ukazují, že právě časování mateřských a rodičovských startů je klíčovým faktorem rostoucí variability plodnosti mezi regiony. Zatímco samotná intenzita plodnosti v Česku vykazuje trend prostorového sblížování, časování porodů se naopak stále více rozchází. Tento proces vede k výrazné diverzifikaci reprodukčních modelů a ke vzniku pluralitních vzorců reprodukčního chování. Autoři tak identifikují silnou prostorovou divergenci v načasování rodičovství, která pravděpodobně bude ovlivňovat podobu plodnosti i v budoucích letech. Regiony se tak stále jasněji odlišují podle toho, jak rychle či pomalu přistoupily k posunu rodičovství do vyššího věku (Šprocha – Šídlo, 2016a).

Uvedené skutečnosti dále rozšiřuje článek s názvem *Spatial Differentiation and Fertility Postponement Transformation in Czechia* členění okresy Česka do čtyř shluků. Z výsledků vyplynulo, že nízká míra rekuperace po výrazném odkladu plodnosti je hlavní příčinou toho, že většina moravských okresů zaostávala za celostátním průměrem. Transformace plodnosti v Česku zásadně změnila prostorové vzorce reprodukčního chování, i když některé dlouhodobé charakteristiky – zejména ve velkých městských okresech, ekonomických centrech a západních a severozápadních pohraničních regionech – zůstávají stabilní. V západních a severozápadních regionech je celková plodnost vyšší, přičemž výrazněji se to projevuje u mladších žen. Článek ukazuje, že mnoho moravských okresů patří k těm, které se nejhůře přizpůsobily novým reprodukčním podmínkám. Původní jednotný model časného rození dětí však již neplatí a byl převážně opuštěn. Tempo a věková struktura plodnosti, úzce spojené s odkladem a následnou rekuperací, se nyní staly klíčovými prostorovými faktory, které rozlišují reprodukční chování mezi českými okresy. Dochází také k určité konvergenci okresů v kontextu rekuperace (Šprocha – Šídlo, 2016b).

V článku *Odkládání mateřství a regionální diferenciace plodnosti v Česku a na Slovensku* je doku-

mentována výrazná proměna prostorového rozložení plodnosti v Česku. Zatímco začátkem 90. let se vyšší plodnost soustřeďovala zejména v moravských okresech, nejnižší hodnoty vykazovala Praha, její zázemí a další hospodářská centra, následně dochází k inverzi tohoto vzorce. Regiony s původně nejnižší plodností se vyznačovaly nejvýraznějším oživením plodnosti ve vyšším věku a patřily k okresům s nejvyšší celkovou plodností. Tento obrat autoři interpretují jako důsledek rychlého přijetí odloženého rodičovství v ekonomicky silných a urbanizovaných oblastech.

Naopak mnohé moravské regiony, dříve charakteristické relativně vysokou plodností, vykazovaly perzistenci velmi nízké plodnosti, především kvůli absenci významnějšího zvýšení reprodukce ve vyšším reprodukčním věku. Mezi oblastmi s nízkou mírou oživení plodnosti patří zejména socioekonomicky problémové regiony. Jde především o příhraniční okresy na jihu a jihovýchodě při hranici se Slovenskem a Rakouskem a dále o příhraniční území na západě a severovýchodě Česka. Autoři upozorňují, že tato území dlouhodobě vykazují slabší socioekonomické ukazatele, což se projevuje také v opožděné či omezené adaptaci na nové reprodukční vzorce (Šídlo – Šprocha, 2018).

METODOLOGIE – ZDROJE DAT

Tento článek vychází z Demografické ročenky okresů, která je pravidelně publikována Českým statistickým úřadem (ČSÚ, 2024). Analýza se zaměřuje na demografické údaje za roky 2015, 2019 a 2023.

S ohledem na specifické postavení hlavního města Prahy, které funguje zároveň jako kraj i okres, nebyla tato jednotka do analýzy zahrnuta. Zahrnutí Prahy by vzhledem k její demografické výjimečnosti mohlo zkreslit výsledky jako statistické odlehle pozorování. Tato skutečnost byla prokázána výpočtem Mahalanobisovy vzdálenosti, její hodnoty a podrobnější zdůvodnění jsou následně zobrazeny dále v první analytické části článku.

Analýza tak zahrnuje 76 okresů Česka.

Pro analýzu byly pro všechny okresy Česka (vyjma Prahy) vybrány následující proměnné:

- úhrnná plodnost,
- podíl živě narozených dětí ve třetím a vyšším pořadí,
- podíl živě narozených dětí mimo manželství,

- index uměle přerušných těhotenství (nevynucených zdravotními důvody) a počtu živě narozených (dále jen index UPT),
- průměrný věk matky při narození prvního živého dítěte.

Výběr těchto proměnných byl veden dvěma hlavními kritérii. Důležitá byla jejich dostupnost, tedy skutečnost, že všechny tyto ukazatele jsou ve sledovaných letech k dispozici na úrovni okresů.

Dále byly zařazeny ukazatele rozšiřující kontext reprodukčních návyků obyvatel. Například kromě celkové úrovně aktuální plodnosti charakterizované syntetickým ukazatelem úhrnné plodnosti je tak zahrnuto procento dětí narozených mimo manželství a podobně.

Význam jednotlivých proměnných pro analýzu, jejímž cílem je charakterizovat reprodukční chování, a tedy důvod zařazení těchto proměnných, lze shrnout následovně:

Úhrnná plodnost – charakterizuje tendenci populace mít potomky a celkovou úroveň intenzity plodnosti.

Podíl dětí živě narozených ve třetím a vyšším pořadí – ukazuje, jak často dochází k překročení hranice reprodukce (více než dvě děti na ženu), respektive plodnosti vyšších pořadí.

Podíl dětí živě narozených mimo manželství – odráží proměny rodinných modelů a postoj společnosti k manželství jako předpokladu rodičovství.

Index UPT – tento ukazatel byl vypočten jako podíl, kde v čitateli je počet UPT (s odečtením UPT ze zdravotních důvodů) a ve jmenovateli počet živě narozených dětí. Reprezentuje přerušená těhotenství s vyloučením umělých přerušení ze zdravotních důvodů. Pro lepší interpretaci byl vynásoben stem.

Průměrný věk matky při narození prvního živého dítěte – vypovídá o načasování rodičovství, které souvisí mimo jiné se vzděláním a profesními ambicemi žen.

Struktura použitých ukazatelů zahrnuje jeden syntetický indikátor celkové úrovně reprodukčního chování v podobě úhrnné plodnosti, ukazatel zachycující časování mateřství (průměrný věk prvorodičky) a tři podílové charakteristiky, které doplňují informace o reprodukčních vzorcích v jednotlivých regionech (podíl živě narozených mimo manželství, podíl živě narozených ve třetím a vyšším pořadí a index UPT).

Tyto ukazatele společně poskytují komplexní pohled na intenzitu, načasování i strukturu reprodukčního chování v rámci okresů.

METODOLOGIE – KVANTITATIVNÍ METODY

V úvodu bylo nutné učinit rozhodnutí o (ne)vyřazení některých pozorování (okresů) z analýzy.

Mahalanobisova vzdálenost byla použita k identifikaci okresů, které se ve vícerozměrném prostoru sledovaných demografických ukazatelů odlišují od zbytku souboru. Pro každý okres byla vypočtena vzdálenost:

$$D^2 = (x - \mu)^T S^{-1} (x - \mu), \quad (1)$$

kde x je vektor hodnot daného okresu, μ je vektor průměrů všech okresů a S^{-1} je inverzní kovarianční matice, což zohledňuje nejen rozdílné měřítka jednotlivých ukazatelů, ale také jejich vzájemné korelace, a proto umožňuje posoudit odlehlost okresu v kontextu celého vícerozměrného prostoru (Pecáková, 2011). Statistika D^2 je porovnána s chí-kvadrát rozdělením o p stupních volnosti, kde p představuje počet použitých ukazatelů. To umožňuje stanovit hranici, nad kterou lze okres považovat za statisticky signifikantně odlehlý. Kritická hodnota je dána:

$$D^2 > \chi^2_{p, 1-\alpha}. \quad (2)$$

V tomto článku byla použita hladina významnosti . Okresy, jejichž Mahalanobisova vzdálenost překročila odpovídající kritickou hodnotu chíkvadrát rozdělení, byly označeny za vícerozměrně odlehle pozorování (outliers). Tyto jednotky se neodlišují pouze v jedné proměnné, ale v kombinaci všech sledovaných ukazatelů a nelze takto identifikované odlehle pozorování využít s ohledem na výpočetní algoritmus vícerozměrné metody.

Z hlediska popisné statistiky článek pozoruje kvantily, konkrétně kvartily, jednotlivých proměnných během sledovaných let. Pro účely grafického zobrazení výsledků byly vytvořeny, jak je v podobných případech obvyklé, krabičkové grafy (boxploty), které kromě kvartilů datových souborů zobrazují minimální a maximální neodlehle hodnoty (whiskers), či jednorozměrná odlehle pozorování. Odlehle pozo-

rování jsou taková pozorování, pro která platí jedna z následujících podmínek:

hodnota < první kvartil - $(1,5 \times \text{mezikvartilové rozpětí})$,

nebo:

hodnota > první kvartil + $(1,5 \times \text{mezikvartilové rozpětí})$,

kde mezikvartilové rozpětí je rovno rozdílu třetího a prvního kvartilu datového souboru (Řezanková, 2024).

Dalším zkoumaným rozměrem byla jednorozměrná variabilita vstupních proměnných a její vývoj napříč lety. K tomu byly pro analýzu zkonstruovány variační koeficienty, jejichž výhodou je jejich bezrozměrnost, a tedy možnost srovnání nejenom stejných proměnných v různých časech, ale také různých proměnných (s různými škálami). Jedná se o poměr směrodatné odchylky a průměru vynásobený 100 (díky vynásobení jsou hodnoty uváděny v procentech) (Hindls a kol., 2007).

Deskriptivní analytická část je následně doplněna o přehled korelací zvolených proměnných a jejich vývoj v čase. K výpočtům monotónních závislostí byla použita metoda Spearmanova korelačního koeficientu, jehož výpočet vychází z transformace proměnných na ordinální typ a následného porovnání pořadí dvou sledovaných proměnných pro dané pozorování. Výsledky jsou zobrazeny ve vizualizované podobě korelační matice.

Shluková analýza je jednou z vícerozměrných metod, která se využívá k identifikování struktury v datech, konkrétně k seskupení objektů do skupin, shluků, na základě podobnosti proměnných. Z toho vyplývá, že objekty vyskytující se ve stejném shluku jsou homogenní, naopak objekty z různých shluků jsou heterogenní.

U shlukování není zařazení do skupin předem známé, metoda vytváří nové shluky. Přístupy ke shlukové analýze se z hlediska základního členění dělí na hierarchické a nehierarchické přístupy, u hierarchických přístupů autoři dopředu neurčují počet shluků, u nehierarchických ano.

Hierarchické přístupy se dále dělí na aglomerativní, kdy každý objekt nejprve tvoří samostatný shluk, a později dochází na základě podobností k seskupování

objektů ve vícečlenné shluky (v některých případech shluk může stále zůstat jednočlenný), a divizivní, přičemž tyto přístupy uvažují, že všechny objekty zpočátku tvoří jeden shluk, a později dochází k jejich rozdělování na méně početné shluky (Řezanková, 2024: 78).

V následující shlukové analýze bylo použito aglomerativní hierarchické shlukování, konkrétně Wardova metoda. Vzhledem k povaze metody byla jako míra podobnosti zvolena čtvercová euklidovská vzdálenost. Wardova metoda patří mezi nejčastěji využívané postupy v rámci hierarchických metod, protože, zjednodušeně řečeno, minimalizuje vnitroshlukovou variabilitu, je poměrně robustní a vytváří kompaktní shluky (Hebák a kol., 2015).

Vzhledem k silným závislostem mezi některými z dvojic proměnných nebyla shluková analýza konstruována z těchto dat, nýbrž z komponent. Komponenty byly vytvořeny pomocí metody hlavních komponent, a to na základě jejich hlavní vlastnosti, tedy vzájemné kolmosti, resp. nezávislosti. Tím se liší od jiných přístupů, jako je např. výpočet z-skórů, kdy takový přístup sice znormalizuje vstupní data, avšak proměnné mohou i nadále být korelované. Komponenty vycházejí z původních dat výše zmiňovaných proměnných za 3 roky – 2015, 2019 a 2023 (dohromady je analýza tvořena 15 proměnnými). Vzhledem k rozdílným škálám i rozptýlům původních proměnných, transformace vstupních dat metodou hlavních komponent vycházela z korelační matice. Byť tato metoda bývá často využívána jako redukční algoritmus, pro současnou úlohu nebylo potřeba snižovat počet vstupujících pro-

měnných a redukovat tak vstupní informaci. Motivací využití metody hlavních komponent je skutečnost, že tvorbou nových nekorelovaných latentních proměnných pomáhá metoda s fungováním shlukových algoritmů (Jolliffe, 2002).

Výpočty a vizualizace analýzy byly provedeny pomocí softwarů MS Excel (Microsoft Corporation, 2024), SPSS (SPSS Inc., 2009) a programovacího jazyku R (R Core Team, 2024). Konkrétně byly využity následující balíčky: *corrplot* (Wei – Simko, 2024), *dplyr* (Wickham a kol., 2023), *ggplot2* (Wickham et al., 2025), *RCzechia* (Lacko, 2024), *readxl* (Wickham – Bryan, 2025), *sf* (Pebesma, 2018), *tidyverse* (Wickham a kol., 2024).

DESKRIPTIVNÍ ANALÝZA VSTUPNÍCH DAT

Cílem kapitoly je poskytnout detailní přehled o základních rysech regionální diferenciacie reprodukčního chování a identifikovat hlavní okruhy, ve kterých jsou rozdíly mezi administrativními jednotkami nejvýraznější. Kapitola se zaměřuje na to, jak se jednotlivé sledované proměnné v okresech liší a jak se jejich prostorové rozložení proměňovalo mezi lety 2015, 2019 a 2023.

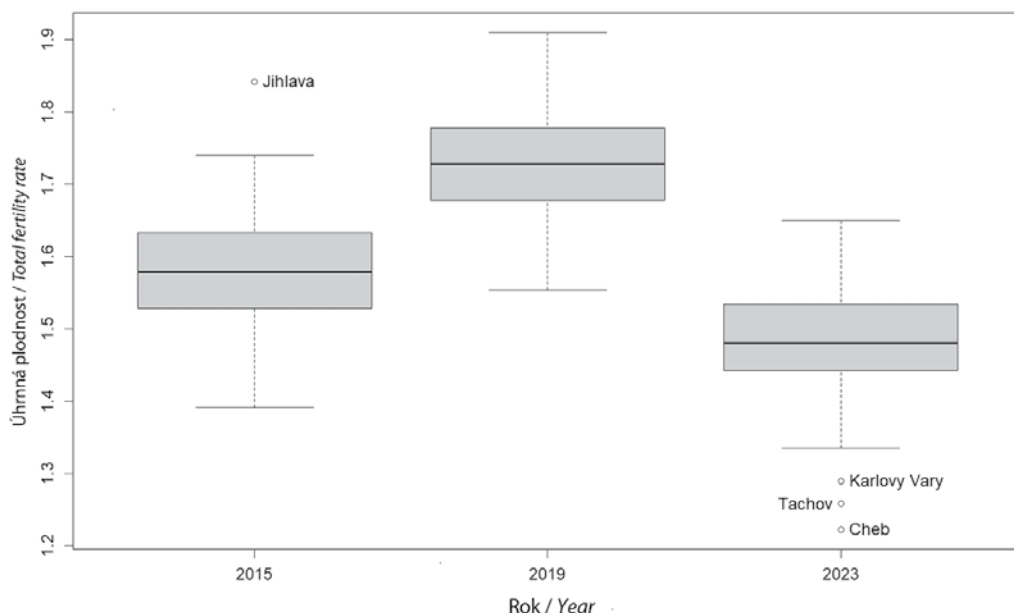
Tato část tak vytváří analytický základ pro následné komplexnější porovnání okresů a pro vysvětlení podobnosti jejich reprodukčních profilů pomocí více-rozměrné statistiky. Deskriptivní statistika a přehled monotónních vztahů mezi ukazateli zde slouží jako nástroj k pochopení toho, jaké charakteristiky českých

Tab. 1: Vícerozměrná odlehlá pozorování v letech 2015, 2019 a 2023 pomocí výpočtu Mahalanobisovy vzdálenosti / Multivariate outlier detection in the years 2015, 2019, and 2023 calculated using Mahalanobis distance

2015		2019		2023	
Okres / District	Mahalanobisova vzdálenost Mahalanobis distance	Okres / District	Mahalanobisova vzdálenost Mahalanobis distance	Okres / District	Mahalanobisova vzdálenost Mahalanobis distance
Hlavní město Praha Capital City of Prague	28,7	Hlavní město Praha Capital City of Prague	25,3	Hlavní město Praha Capital City of Prague	27,5
Praha-západ Prague-West	15,5	Jeseník	21,2	Jeseník	20,7
Jihlava	13,3	Ústí nad Labem	14,1	Cheb	14,4
		Žďár nad Sázavou	12,9	Klatovy	13,3
				Praha-západ Prague-West	13,1

Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: Authors' calculation based on CZSO data.

Graf 1: Regionální diferenciace úhrnné plodnosti v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabičkový graf / Regional differentiation of the total fertility rate in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

okresů se v čase sblíží, které se naopak rozcházejí a v jakých demografických procesech se regionální kontrasty nejvíce prohlubují.

Nejprve zde na začátku analytických částí článku, je nicméně vhodné podpořit konkrétními čísly nemožnost zahrnutí Prahy do regionálního srovnání. Pro tento účel je využita Mahalanobisova vzdálenost.

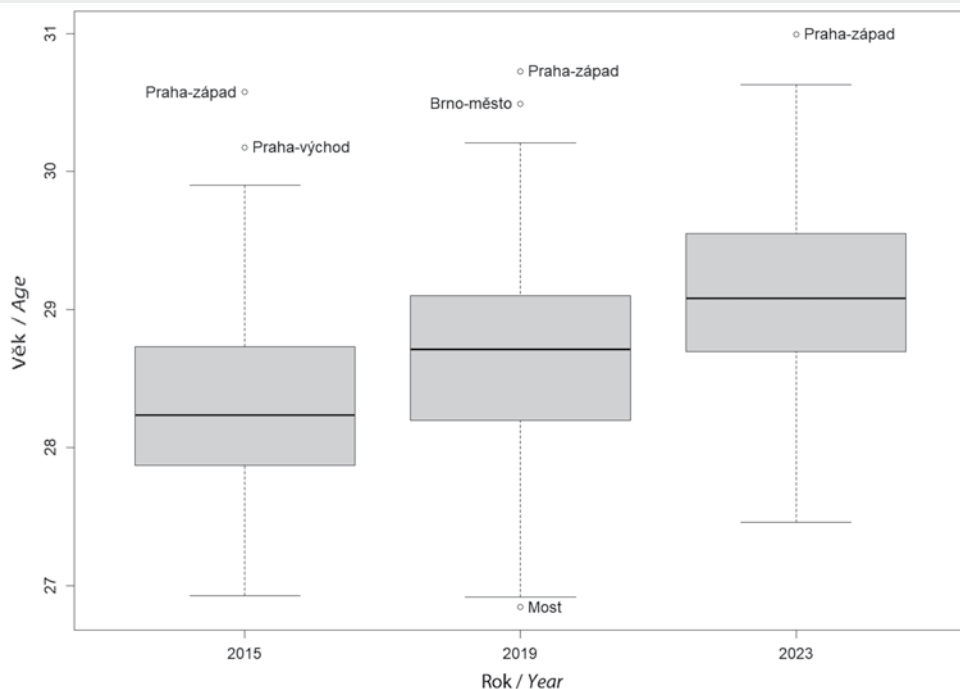
Jediným okresem, který byl ve všech sledovaných letech dle Mahalanobisovy vzdálenosti označen za vícerozměrné odlehlé pozorování, je Hlavní město Praha. Zároveň je jeho odlehlost patrná i z komparace s jinými odlehlými okresy. Z toho důvodu se tento okres v analýze (deskriptivní analýze ani shlukové analýze) nadále již nevyskytuje tak, aby nezkrášloval výsledky a současně byly všechny provedené analýzy vzájemně porovnatelné. Pro možnost komparace Prahy s výsledky ostatních okresů, v případě zájmu čtenáře, jsou následně konkrétní hodnoty za Prahu zařazeny do přílohy 1.

Z grafu 1 vyplývá, že mezi lety 2015 a 2019 došlo v okresech Česka k nárůstu úhrnné plodnosti, v roce 2023 naopak došlo ke znatelnému poklesu. Zajíma-

vým výsledkem této analýzy je skutečnost, že v roce 2019 byla naměřena minimální úhrnná plodnost (v okrese Plzeň-sever) vyšší než mediánová úhrnná plodnost pro všechny okresy ČR v roce 2023. Úhrnná plodnost ve zmíněném okrese byla dokonce i vyšší než třetí kvartil úhrnné plodnosti z roku 2023. Maximální zjištěná hodnota úhrnné plodnosti v roce 2023 (v okrese Kroměříž) byla nižší než první kvartil sledující stejnou charakteristiku v roce 2019. Přestože se naměřené hodnoty úhrnné plodnosti v čase výrazně mění, difference mezi okresy, charakterizovaná mezikvartilovým rozpětím, se ve všech sledovaných letech pohybovala kolem jedné desetiny, jak je patrné z krabičkových grafů.

Ve třech okresech v roce 2023 úhrnná plodnost nepřekročila ani hranici 1,3, jedná se o okresy Cheb, Tachov a Karlovy Vary (okres Cheb geograficky sousedí se zbylými jmenovanými okresy a společně s Karlovými Vary se nachází v Karlovarském kraji, okres Tachov je součástí Plzeňského kraje). Hodnoty těchto územních celků jsou považovány za odlehlé.

Graf 2: Regionální diferenciace průměrného věku matky při narození prvního živého dítěte v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabíčkový graf / Regional differentiation of the mean age of mothers at first live birth in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

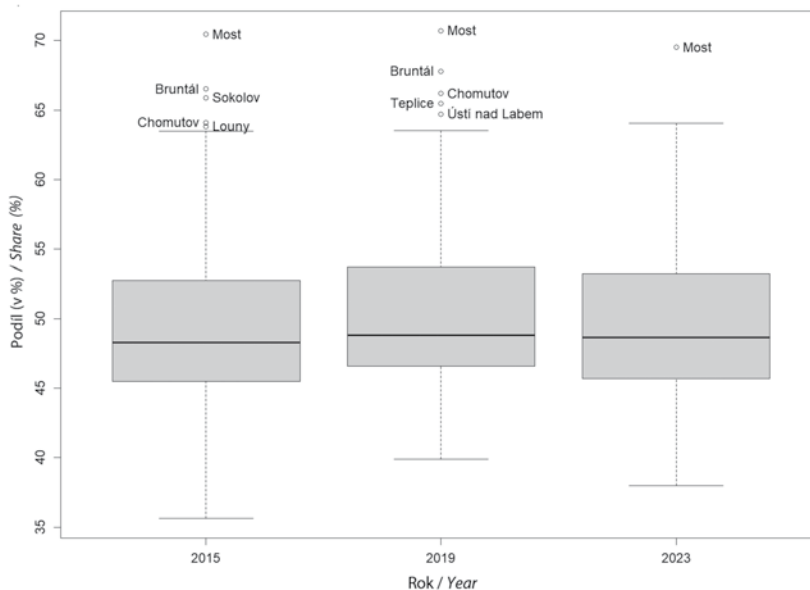
Ve všech sledovaných letech na grafu 2 byl zaznamenán nejvyšší průměrný věk prvorodiček v okrese Praha-západ. Zároveň ve všech letech údaj figuroval jako odlehle pozorování. V roce 2015 byl také odlehlelým pozorováním okres Praha-východ, v roce 2019 Brno-město. Ve všech uvedených případech byly okresy označeny za odlehlá pozorování z důvodu příliš vysokého průměrného věku prvorodiček, což poukazuje na odlišný životní styl a přístup k plánování rodinného života žen. V roce 2019 byl kromě již výše zmíněných okresů za odlehle pozorování považován také okres Most, kde však průměrný věk prvorodiček byl považován za odlehle pozorování z důvodu nízké hodnoty sledovaného ukazatele. Také v roce 2015 byl v tomto okrese v rámci celorepublikového srovnání nejnižší průměrný věk prvorodiček. Naopak v roce 2023 okres Most dosahoval až sedmé nejnižší hodnoty. Nejnižší a druhý nejnižší průměrný věk prvorodiček byl v okresech Děčín a Chomutov, tedy v okresech z Ústeckého kraje. V průběhu času lze na základě všech

tří kvartilů konstatovat narůstající průměrný věk prvorodiček.

Z grafu 3 lze vyčíst, že v roce 2015 bylo v porovnání s dalšími sledovanými roky napozorováno nejvyšší variační rozpětí podílu živě narozených dětí mimo manželství. Rozdíl mezi nejvyšší a nejnižší sledovanou hodnotou činil v tomto roce při zohledňování i odlehklých hodnot po zaokrouhlení 34,83 procentního bodu. Je zajímavé, že na rozdíl od předchozích sledovaných ukazatelů je zde poměrně stabilní trend, kdy medián ani kvartily se v čase v podstatě nemění.

V okrese Most se pravidelně rodilo nejvyšší procento dětí mimo manželský svazek, přičemž v letech 2015 a 2019 byla hodnota vyšší než 70 % (v roce 2023 69,5 %). Nejenom v okrese Most, ale i v jiných okresech Ústeckého kraje se narodil výrazný podíl dětí mimo manželství. Naopak okres Brno-venkov patřil během všech tří let mezi tři okresy s nejnižším procentem živě narozených dětí mimo manželství.

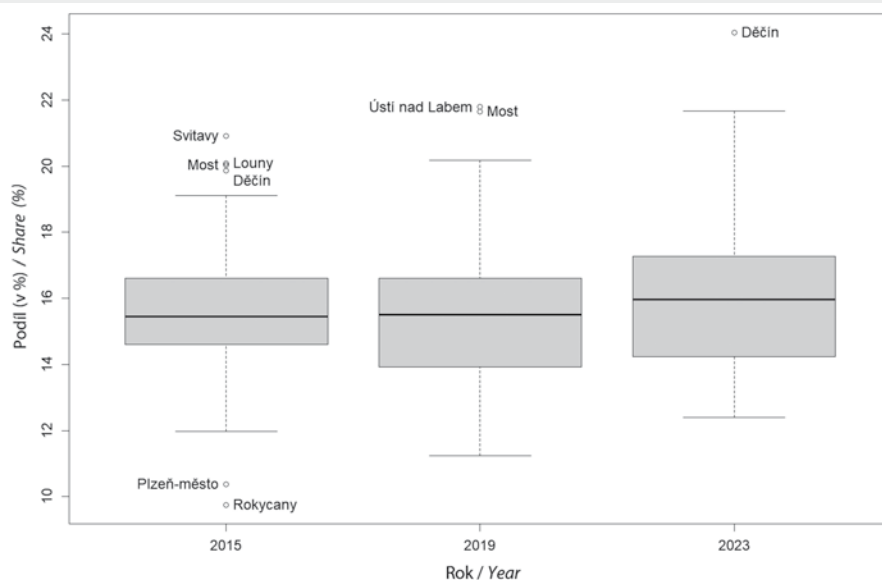
Graf 3: Regionální diferenciace podílu živě narozených dětí mimo manželství v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabčkový graf / Regional differentiation of the percentage of live births outside marriage in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

Graf 4: Regionální diferenciace podílu živě narozených dětí třetího a vyššího pořadí v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabčkový graf / Regional differentiation of the percentage of third- and higher-order live births in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

V grafu 4 je patrný nárůst mediánu podílu narozených dětí třetího a vyššího pořadí. Vrátime-li se ke grafu 1, medián úhrnné plodnosti v roce 2023 byl ze tří sledovaných let nejnižší, medián třetího a vyššího pořadí byl však za tři sledované roky nejvyšší, tím pádem lze konstatovat, že došlo z relativního hlediska primárně k poklesu plodnosti potenciálních prvorodiček a matek, které by porodily druhé dítě, a v menší míře matek, které již živě porodily alespoň 2 děti.

V roce 2015 bylo, v porovnání s následujícími sledovanými roky, zaznamenáno nejnižší mezikvartilové rozpětí. Odlehle hodnoty však způsobily, že variační rozpětí v tomto roce bylo naopak oproti rokům 2019 a 2023 nejvyšší. Za cenu určitého zjednodušení lze konstatovat, že diferenciací okresů z hlediska plodnosti třetího a vyššího pořadí ve sledovaném období výrazně vzrostla.

Výrazný podíl dětí narozených ve třetím a vyšším pořadí byl opět zaznamenán primárně v okresech Ústeckého kraje, v roce 2015 naopak dva okresy z Plzeňského kraje dosáhly v porovnání se zbylými

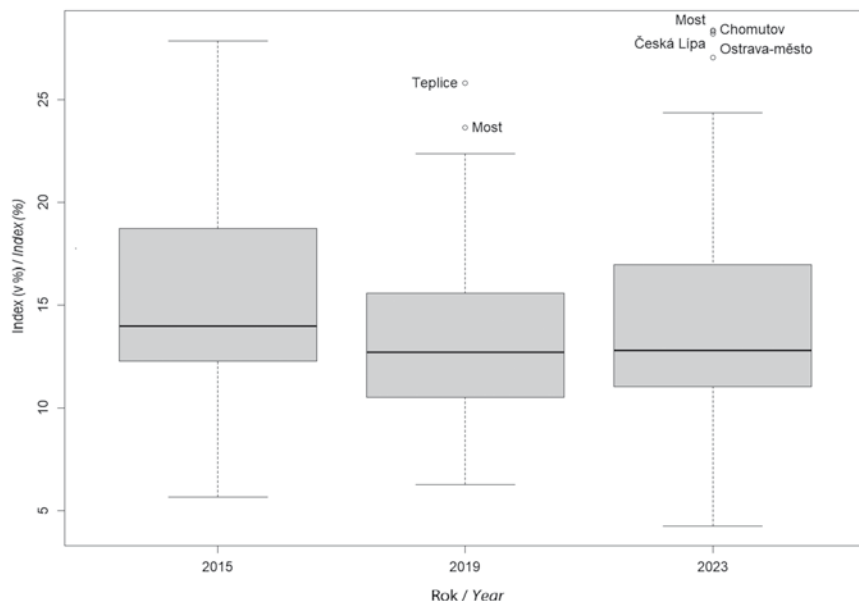
sledovanými okresy příliš nízkých hodnot, a tak byly označeny za odlehlá pozorování.

Taktéž tento ukazatel poukazuje na skutečnost, že vysokých (odlehklých) hodnot nabývaly okresy primárně z Ústeckého kraje, okres Most byl dokonce jako jediný v letech 2019 a zároveň 2023 odlehlým pozorováním. Největší variační rozpětí bylo zaznamenáno v roce 2023, kdy v okrese Most tento ukazatel nabýval hodnoty 28,37 %, zatímco v okrese Klatovy 4,27 %.

Pro sledování variability napříč různými proměnnými i různým časovým obdobím byl doplněn výpočet variačních koeficientů.

Nejvyšší variabilita byla z hlediska typu proměnných napozorována u indexu UPT. Naopak nejnižší hodnoty variačního koeficientu byly napozorovány pro proměnnou průměrný věk prvorodiček, a to i přesto, že v rámci proměnné úhrnná plodnost se v datovém souboru vyskytuje méně odlehlých pozorování. Tato proměnná se od ostatních liší tím, že nejvyšší variabilita byla napozorována v roce 2019. Variabilita podílu živě narozených dětí mimo manželství v čase klesá.

Graf 5: Regionální diferenciací indexu UPT v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabíkový graf / Regional differentiation of the index of induced abortions for non-medical reasons per number of live births in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

Tab. 2: Vývoj variačních koeficientů (v %) v letech 2015, 2019 a 2023 (soubor okresů bez Prahy)

Change in the coefficients of variation (in %) in 2015, 2019, and 2023 (a set of districts excluding Prague)

Ukazatel / Indicator	Rok / Year		
	2015	2019	2023
Podíl živě narozených dětí mimo manželství (v %) Share of live births outside marriage (%)	14,84	13,59	13,73
Úhrnná plodnost / Total fertility rate	5,34	4,49	5,40
Podíl živě narozených třetího a vyššího pořadí (v %) Share of third- and higher-order live births (%)	13,53	14,42	15,10
Index UPT / Index of induced abortions	29,58	30,66	36,24
Průměrný věk matky při narození prvního živého dítěte Average age of mother at first live birth	2,62	2,66	2,33

Zdroj: Vlastní zpracování na základě dat ČSÚ.**Source:** Authors' calculation based on CZSO data.

Pro zbylé proměnné platí, že variační koeficient nabývá nejvyšších hodnot pro data z roku 2023.

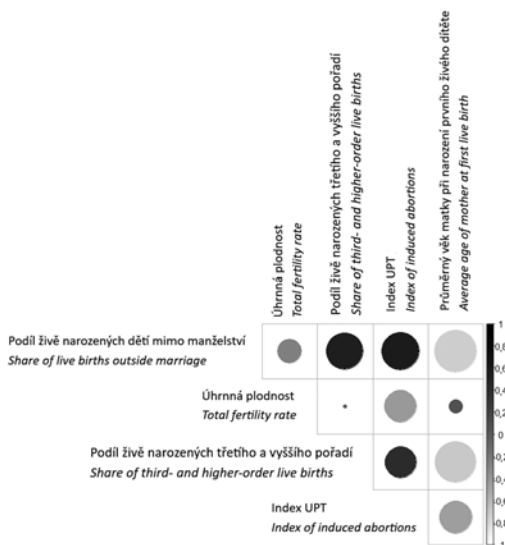
Pro doplnění deskriptivního pohledu byly také pro každý rok (2015, 2019, 2023) vytvořeny korelační matice, které porovnávají závislost mezi sledovanými proměnnými. Vzhledem ke skutečnosti, že během výše zmíněných let nedošlo k výrazným a interpretačně příliš zajímavým změnám v rámci hodnot monotónních závislostí, obsahuje tato část pouze grafický výstup korelační matice za rok 2023. Korelační matice vyobrazující závislosti v letech 2015 a 2019 byly zařazeny

jako přílohy 2 a 3. Pro provedení výpočtů byl využit Spearmanův korelační koeficient.

Hodnota Spearmanova korelačního koeficientu nabývá hodnot od -1 do 1 . Čím větší je kruh mezi dvěma pozorovanými proměnnými, tím větší je mezi proměnnými závislost, tmavá barva kruhu znázorňuje pozitivní závislost (korelační koeficient nabývá kladné hodnoty), světlé šedý kruh naopak vyobrazuje negativní závislost (korelační koeficient nabývá záporné hodnoty). Malý šedý kruh znázorňuje nízkou (téměř nulovou) závislost.

Graf 6: Korelační matice sledující závislost mezi sledovanými proměnnými v roce 2023

Correlation matrix monitoring the dependence between selected variables in 2023

**Zdroj:** Vlastní zpracování na základě dat ČSÚ.**Source:** Authors' calculation based on CZSO data.

Z grafu 6 je patrných několik poměrně očekávaných závislostí, jako je při vyšším procentu mimomanželsky narozených nižší průměrný věk prvorodiček a podobně, jak představí následný výčet.

Například mezi okresy existuje silná pozitivní závislost mezi proměnnými:

- Procento živě narozených dětí mimo manželství a index uměle přerušovaných těhotenství nevynucených zdravotními důvody na počet živě narozených (0,594), jelikož obě proměnné poukazují na liberálnější postoje k manželství a rodičovství i patrnou větší nestabilitu partnerství v regionu.
- Procento narozených třetího a vyššího pořadí a procento živě narozených dětí mimo manželství (0,555), kdy v některých okresech může být rozšířenější kohabitace i při vyšším počtu dětí v rodině, opět to může souviset se socioekonomickým statusem části populace.
- Procento narozených třetího a vyššího pořadí a index UPT (0,413), tento vztah poukazuje na výskyt oblastí, ve kterých je běžnější mít více dětí, a kde se zároveň častěji objevují situace, kdy je těhotenství neplánované nebo časově nevhodné, což může vést k vyšší míře záměrně přerušovaných koncepcí.

Naopak silná negativní závislost se vyskytuje mezi proměnnými:

- Průměrný věk matek při narození prvního živého dítěte a procento živě narozených dětí mimo manželství (−0,737), platí, že mladší ženy častěji mají děti v kohabitaci, nebo ještě před vstupem do manželství. V regionech s nižším věkem při narození prvního dítěte je proto běžnější rodičovství mimo manželství, zatímco v oblastech s odkládaným vstupem do mateřství bývá rodičovství častěji spojeno s manželským svazkem.
- Průměrný věk matek při narození prvního živého dítěte a procento narozených třetího a vyššího pořadí (−0,703). V regionech, kde ženy vstupují do mateřství v mladším věku, je obvykle více času i prostoru pro pokračování reprodukce, což vede k vyššímu podílu dětí vyššího pořadí. Naopak tam, kde dochází k výraznému odkládání prvního porodu do vyššího věku, se snižuje pravděpodobnost dosažení tří

a více dětí, což se promítá do nižšího podílu porodů vyššího pořadí.

- Úhrnná plodnost a index UPT (−0,429), z logické podstaty věci proměnných je patrné, že signifikantní poměr chtěných uměle přerušovaných těhotenství a počtu živě narozených negativně ovlivňuje úhrnnou plodnost.

SHLUKOVÁ ANALÝZA

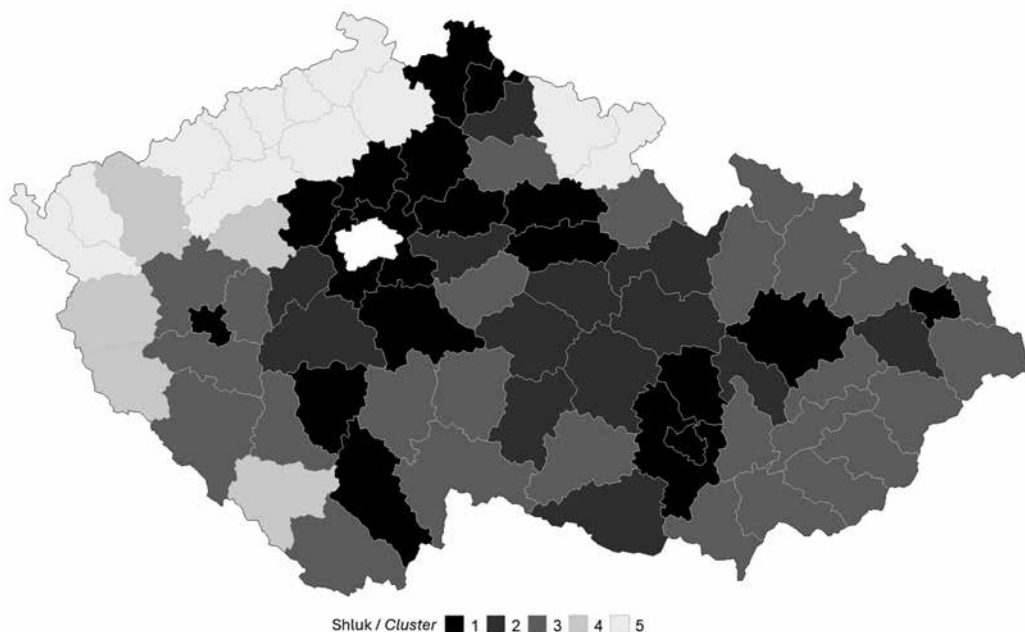
Regionální rozdíly v reprodukčním chování okresů jsou výsledkem dlouhodobě působících sociálních, kulturních i ekonomických procesů, které vytvářejí specifické prostorové vzorce. Zatímco deskriptivní část ukázala, jak se jednotlivé ukazatele v okresech liší, tato kapitola se zabývá otázkou, zda lze na základě kombinace sledovaných charakteristik rozpoznat skupiny okresů s obdobným reprodukčním profilem. Tato strukturace českého prostoru umožňuje zachytit vývoj komplexních vztahů mezi ukazateli v letech 2015, 2019 a 2023, které nejsou patrné při izolovaném pohledu na jednotlivé charakteristiky.

Shluková analýza je zde využita jako nástroj k odhalení těchto vzorců a výsledné shluky okresů poskytují ucelený obraz o tom, jak se různé podoby reprodukčního chování v českých regionech koncentrují, které oblasti se sobě nejvíce podobají a v čem se naopak odlišují. Kapitola tak přináší interpretaci regionálních typů reprodukčního chování a ukazuje, jak stabilní nebo proměnlivé tyto typy jsou. S ohledem na srozumitelnou interpretaci i dobrou čitelnost v černobílém zobrazení je v kartogramu zobrazena varianta s pěti shluky.

Z obrázku 1 lze na první pohled rozklíčovat patrné zákonitosti. Shlukovou analýzou byly rozděleny okresy do celkem 5 shluků. Je zřejmé, že obdobným způsobem se proměnné ve sledovaných letech vyvíjely v příměstských regionech (okolí Prahy a Brna), v okresech Plzeň-město a dalších regionech s významným vlivem městského centra (shluk 1). Pro tyto okresy je typické odkládání založení rodin (nejvyšší průměrný věk matky při narození prvního živého dítěte), což může také mít vliv na skutečnost, že pouze nízké procento dětí se zde narodí ve 3. a vyšším pořadí. Tyto děti se primárně rodí sezdáným pářím. Obecně by se dal tento shluk z hlediska reprodukčního chování pojmut jako okresy, kde rodiče zakládají stabilní

Obrázek 1: Kartogram zobrazující shlukování okresů Česka (bez Prahy) do 5 shluků v letech 2015, 2019 a 2023

Choropleth map showing the clustering of Czech districts (excluding Prague) into five clusters in 2015, 2019 and 2023



Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

rodiny ve vyšším věku. Jedná se o poměrně početný shluk, který je tvořen 19 okresy.

Jiné reprodukční chování bylo zjištěno na přelomu Čech a Moravy (shluk 2). Pro tuto oblast je typická nejvyšší úhrnná plodnost, která je podporována nejnižším indexem UPT, tedy nejnižším poměrem UPT, o kterých nebylo rozhodnuto na základě zdravotního stavu, a počtem živě narozených dětí. Do druhého shluku je seskupeno 13 okresů.

Typická je také oblast hraničního východního přechodu se Slovenskem, tedy oblast Moravy a Slezska, zároveň byl obdobný vývoj charakteristik sledován v jižní a jihozápadní oblasti republiky (shluk 3). Tato oblast nabývá z hlediska celorepublikového srovnání v rámci sledovaných proměnných primárně průměrných hodnot, v ničem nevybočuje. Jedná se o nejpočetnější shluk vytvořený touto analýzou, obsahuje 27 okresů.

Obdobné vzorce chování byly napozorovány rovněž v Ústeckém a Karlovarském kraji (shluk 5), jediným okresem z těchto dvou krajů, který se od zbylých lišil, byl okres Karlovy Vary (shluk 4). Součástí stej-

ného shluku jako je většina okresů ze Severozápadu jsou také okresy Trutnov a Náchod. Pro tyto okresy je typický vysoký podíl dětí narozených nesezdaným párům, nejnižší průměrný věk matky při narození prvního živého dítěte, nejvyšší index UPT, což souvisí s nižšími hodnotami ukazatele úhrnné plodnosti, či nejvyšší podíl dětí narozených ve třetím a vyšším pořadí. Tento shluk působí jako kontrast shluku 1, kdy na rozdíl od prvního zmiňovaného shluku jsou zde zakládány rodiny sice mladšími lidmi (resp. mladšími prvoroďičkami), nicméně samotné počety a jeho budoucnost můžou a často i jsou nestabilní, o čemž svědčí vysoký podíl dětí narozených nesezdaným párům a vysoký index UPT. Shluk je tvořen 12 okresy.

Nejméně početný shluk (shluk 4) figuruje jako určitý mezičlánek mezi shluky 3 a 5, a to nejenom z hlediska geografické polohy okresů, ale také z hlediska hodnot sledovaných proměnných. Tento shluk obsahuje pouze 5 okresů.

Výsledky klasifikace okresů je možné dále podrobněji popsat opět pomocí deskriptivní statistiky.

Tab. 3: Průměrné hodnoty a variační koeficienty sledovaných proměnných napříč shluky v roce 2015

Mean values and coefficients of variation of the monitored variables across clusters in 2015

Shluk Cluster	Podíl živě narozených dětí mimo manželství (v %) <i>Share of live births outside marriage (%)</i>		Úhrnná plodnost <i>Total fertility rate</i>		Podíl živě narozených třetího a vyššího pořadí (v %) <i>Share of third- and higher-order live births (%)</i>		Index UPT (v %) <i>Index of induced abortions (%)</i>		Průměrný věk matky při narození prvního živého dítěte <i>Average age of mother at first live birth</i>	
	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)
1	46,17	11,68	1,61	3,71	13,76	11,06	14,01	23,56	29,01	2,27
2	46,72	11,55	1,61	3,45	16,08	9,41	12,32	26,73	28,39	2,29
3	48,28	11,03	1,54	3,62	15,27	9,97	14,48	20,06	28,28	2,28
4	54,76	9,49	1,59	3,48	15,72	9,43	15,20	17,70	27,69	2,31
5	60,97	8,53	1,57	3,37	18,21	7,82	22,59	11,80	27,45	2,36

Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

Následující oddíl detailně popisuje vývoj sledovaných ukazatelů ve stanovených shlucích ve sledovaných letech.

V roce 2015 byl, jak ukazuje tabulka 3, napříč všemi shluky zaznamenán nejnižší průměrný věk matky při narození prvního živého dítěte, tento ukazatel ve všech shlucích v čase nabýval vyšších hodnot, což potvrzuje zjištění patrné z grafu 2. Nejvíce heterogenní z hlediska porovnání variačních koeficientů je shluk 1, což je primárně dáno vysokým počtem

v něm sdružených okresů. Naopak nejvíce homogenní jsou okresy ve shluku 5.

Z tabulky 4 v porovnání s předchozím sledovaným obdobím, s rokem 2015, došlo kromě nárůstu průměrného věku prvorodičky také k výraznému nárůstu úhrnné plodnosti (což navazuje na komentář grafu 1) a zároveň téměř ve všech shlucích (kromě shluku 4) došlo k poklesu indexu UPT a nárůstu podílu živě narozených dětí mimo manželství. Okresy členěné do shluku 2 jsou v porovnání s předchozím

Tab. 4: Průměrné hodnoty a variační koeficienty sledovaných proměnných napříč shluky v roce 2019

Mean values and coefficients of variation of the monitored variables across clusters in 2019

Shluk Cluster	Podíl živě narozených dětí mimo manželství (v %) <i>Share of live births outside marriage (%)</i>		Úhrnná plodnost <i>Total fertility rate</i>		Podíl živě narozených třetího a vyššího pořadí (v %) <i>Share of third- and higher-order live births</i>		Index UPT (v %) <i>Index of induced abortions (%)</i>		Průměrný věk matky při narození prvního živého dítěte <i>Average age of mother at first live birth</i>	
	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)	Průměr Mean	Var. koef. (v %) Var. coeff. (%)
1	46,73	9,82	1,75	3,30	13,96	11,19	11,86	22,32	29,42	1,93
2	47,49	5,20	1,81	3,65	16,00	8,29	11,46	18,83	28,71	1,64
3	49,70	11,95	1,69	4,18	15,24	12,61	11,87	22,37	28,56	1,93
4	53,18	5,98	1,73	4,66	15,39	6,61	17,29	19,32	28,36	0,95
5	61,21	8,11	1,72	3,34	18,49	11,30	20,02	13,84	27,72	2,10

Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

Tab. 5: Průměrné hodnoty a variační koeficienty sledovaných proměnných napříč shluky v roce 2023

Mean values and coefficients of variation of the monitored variables across clusters in 2023

Shluk Cluster	Podíl živě narozených dětí mimo manželství (v %) <i>Share of live births outside marriage (%)</i>		Úhrnná plodnost <i>Total fertility rate</i>		Podíl živě narozených třetího a vyššího pořadí (v %) <i>Share of third- and higher-order live births (%)</i>		Index UPT (v %) <i>Index of induced abortions (%)</i>		Průměrný věk matky při narození prvního živého dítěte <i>Average age of mother at first live birth</i>	
	Průměr <i>Mean</i>	Var. koef. (v %) <i>Var. coeff. (%)</i>	Průměr <i>Mean</i>	Var. koef. (v %) <i>Var. coeff. (%)</i>	Průměr <i>Mean</i>	Var. koef. (v %) <i>Var. coeff. (%)</i>	Průměr <i>Mean</i>	Var. koef. (v %) <i>Var. coeff. (%)</i>	Průměr <i>Mean</i>	Var. koef. (v %) <i>Var. coeff. (%)</i>
1	45,49	9,07	1,46	3,56	14,57	9,37	14,63	30,23	29,71	1,88
2	47,52	7,20	1,56	2,89	16,57	7,58	11,87	20,53	29,00	1,18
3	48,92	11,68	1,50	3,97	15,09	11,51	11,87	25,77	29,13	1,23
4	51,00	4,78	1,39	7,30	17,28	7,36	13,57	31,21	28,90	1,97
5	61,07	6,75	1,43	5,62	19,92	10,03	22,33	20,11	28,08	1,66

Zdroj: Vlastní zpracování na základě dat ČSÚ.

Source: Authors' calculation based on CZSO data.

obdobím více homogenní. Zároveň pro shluk 1 již neplatí, že by okresy jej tvořící byly nejvíce heterogenní, nejvyšší variabilita napříč sledovanými proměnnými byla napozorována v nejpočetnějším shluku 3. Naopak nejvíce homogenním shlukem pro tento rok byl nejméně početný shluk 4.

Z tabulky 5 je patrné, že v porovnání s rokem 2019 opět došlo ve všech shlucích v průměru k nárůstu průměrného věku matky při narození prvního živého dítěte a k poklesu úhrnné plodnosti, a i přes tento pokles, kromě shluku 3 došlo k nárůstu podílu dětí narozených ve třetím a vyšším pořadí, což poukazuje na nižší tendenci zakládání rodin či rozšíření rodin s 1 dítětem na rodiny se 2 dětmi. Nejvíce homogenní v tomto roce byly okresy tvořící shluk 2. V porovnání se zbylými shluky jsou o něco více heterogenní shluky 3 a 4.

ZÁVĚR

Analýza reprodukčního chování okresů Česka v letech 2015, 2019 a 2023 ukázala, že prostorové rozdíly mezi regiony jsou výrazné a přetrvávají napříč sledovaným obdobím. Deskriptivní statistika potvrdila, že sledované ukazatele se mezi okresy liší nejen úrovní, ale i rozsahem variability. Analýza ukázala, že zatímco distribuce úhrnné plodnosti vykazovala mezi roky 2019 a 2023 výrazný pokles v naměřených hodnotách,

průměrný věk matky při narození prvního dítěte dlouhodobě roste. Poměrně výrazná variabilita byla napozorována u indexu UPT, mimomanželská plodnost zůstávala z hlediska měř polohy poměrně konstantní, a to i přes častý výskyt extrémních hodnot v Severozápadním regionu Česka.

Shluková analýza rozdělila okresy do pěti skupin s odlišnými reprodukčními profily. Jeden shluk zahrnuje okresy s nízkým průměrným věkem matek při prvním narození živého dítěte, vysokým podílem mimomanželsky narozených a vyšším indexem UPT. Další shluk tvoří okresy s opačnými charakteristikami – vyšším věkem prvorodiček a nízkými hodnotami ostatních ukazatelů. Třetí shluk sdružuje okresy s průměrnými hodnotami většiny ukazatelů, čtvrtý shluk se vyznačuje vyšším zastoupením dětí třetího a vyššího pořadí a pátý shluk zahrnuje okresy se specificky vyšší intenzitou UPT. Tato struktura potvrzuje, že české okresy lze rozdělit do několika jasně odlišitelných typů reprodukčního chování, které nejsou dány pouze geografickou polohou, ale kombinací více demografických charakteristik.

Stanovené cíle a související výzkumné otázky byly ve všech sledovaných krocích naplněny a zodpovězeny. První cíl byl splněn detailním popisem regionální diferenciace reprodukčního chování na základě pěti klíčových ukazatelů, což potvrdilo výraznou a dlou-

hodobě přítomnou prostorovou variabilitu. Druhý cíl byl naplněn porovnáním vývoje těchto ukazatelů mezi lety 2015, 2019 a 2023, komparace odhalila jak stabilní rysy, tak dílčí změny ve variabilitě ukazatelů. Třetí cíl byl splněn identifikací pěti shluků okresů s odlišnými reprodukčními profily, které poskytují ucelený obraz

o tom, jak se jednotlivé kombinace ukazatelů v českém prostoru sdružují.

Výsledky studie tak přinášejí komplexní a aktuální pohled na prostorovou strukturu reprodukčního chování v Česku a vytvářejí základ pro interpretaci jeho socioekonomických a kulturních souvislostí.

Literatura

- Bakuncová, K., – Löster, T. 2023. Time Series Clustering Of Fertility in European Union. *International Days of Statistics and Economics*, s. 19–35. Praha: VŠE. https://msed.vse.cz/msed_2023/sbornik/toc.html.
- Bartoňová, D. 1999. Vývoj regionální diferenciace věkové struktury se zřetelem k územním rozdílným ve vývoji reprodukce v České republice. *Geografie – Sborník ČGS*, 104(1), s. 13–23. ISSN 1212-0014. <https://doi.org/10.37040/geografie1999104010013>.
- Bartoňová, D. 2001. Demografické chování populace České republiky v regionálním a evropském kontextu. In Hampl, M. (Ed.) *Regionální vývoj: specifika české transformace, evropská integrace a obecná teorie*, s. 45–73, Praha: DemoArt.
- ČSÚ. 2025. *Narození*. <https://csu.gov.cz/narozeni>.
- Fialová, L. 2018. *Sociologická encyklopedie*. Dostupné z: <https://encyklopedie.soc.cas.cz/w/Regionalizace>.
- Hindls, R. – Hronová, S. – Seger, J. – Fischer, J. 2007. *Statistika pro ekonomy*. Praha: Professional Publishing. 415 s. ISBN 978-80-86946-43-6.
- Hon, F. 2020. Současná úroveň plodnosti třetího pořadí v České republice. In *Sborník příspěvků z konference RELIK*, s. 129–139. Praha: VŠE. <https://relik.vse.cz/2020/download/pdf/363-Hon-Filip-paper.pdf>.
- Kocourková, J. – Štašná, A., Idbeková, B. – Slabá, J. 2023. Trendy plodnosti a potratovosti v Česku. *Časopis lékařů českých*, 162, s. 299–306. <https://www.prolekare.cz/casopisy/casopis-lekaru-ceskych/2023-7-8-1/trendy-plodnosti-a-potratovosti-v-cesku-136675>.
- Kurkin, R. 2015. *Faktory plodnosti a regionální diferenciace plodnosti v České republice po roce 1991* [Disertační práce, Univerzita Karlova]. <https://explorer.cuni.cz/publication/502597>.
- Lacko, J. 2024. *RCzechia: Spatial Objects of the Czech Republic*. <https://rczechia.jla-data.net>. <https://doi.org/10.21105/joss.05082>.
- Microsoft Corporation. 2024. *Microsoft Excel* [Software]. <https://office.microsoft.com/excel>.
- Paloncyová, J. – Hašková, H. – Höhne, S. – Kuchařová, V. – Žáčková, L., – Lukić, A. 2021. *Bezdětní a postoje k rodičovství*. VÚPSV. 170 s. https://katalog.vupsv.cz/fulltext/vz_487.pdf.
- Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1), s. 439–446. <https://doi.org/10.32614/RJ-2018-009>.
- Pecáková, I. 2011. *Statistika v terénních průzkumech*. Praha: Professional Publishing. 236 s. ISBN 978-80-7431-039-3.
- R Core Team. 2024. *R: A language and environment for statistical computing* [Software]. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rychtaříková, J. 2007. Regionální diferenciace plodnosti v průřezové a kohortní perspektivě. In Kučera, T. – Polásek, V. (Eds.) *Sborník příspěvků XXXVII. Výroční demografické konference ČDS*, s. 92–103, Olomouc: ČDS.
- Řezanková, H. 2024. *Statistické metody se zaměřením na kategoriální data*. Praha: Oeconomica. 210 s. <https://oeconomica.vse.cz/publikace/statisticke-metody-se-zamerenim-na-kategoriální-data-e-kniha/>.
- Slabá, J. – Kocourková, J. – Štašná, A. 2024. The fertility timing gap: the intended and real timing of childbirth. *Journal of Biosocial Science*, 56(3), s. 504–517. doi:10.1017/S002193202400004X.
- SPSS Inc. 2009. *PASW Statistics for Windows, Version 18.0* [Software]. SPSS Inc.
- Stípková, M. 2015. Ideational and Economic Causes of the Rise in Non-marital Childbearing in the Czech Republic. *European Journal of Population*, 31(5), s. 473–494. <https://doi.org/10.1007/s10680-015-9350-4>.
- Šídlo, L. 2008. Faktory ovlivňující regionální diferenciaci plodnosti v Česku na počátku 21. století. *Demografie, revue pro výzkum populačního vývoje*, 50(3), s. 186–198.

- Šídlo, L. – Šprocha, B. 2018. Odkládání mateřství a regionální diferenciace plodnosti v Česku a na Slovensku. *Geografie*, 123(3), s. 407–436. <https://doi.org/10.37040/geografie2018123030407>.
- Šimpach, O. – Šimpachová Pechrová, M. 2016. Cluster Analysis Of EU's Regions According to Demographics Criteria. *International Days of Statistics and Economics*, s. 1806–1815. Praha: VŠE. https://msed.vse.cz/msed_2016/article/166-Simpach-Ondrej-paper.pdf.
- Šprocha, B. – Šídlo, L. 2016a. Regionálne rozdiely v charaktere plodnosti v Česku a na Slovensku. *Sborník příspěvků z konference RELIK*, s. 559–571. Dostupné z: <https://relik.vse.cz/2016/download/pdf/57-Sprocha-Branislav-paper.pdf>.
- Šprocha, B. – Šídlo, L. 2016b. Spatial Differentiation and Fertility Postponement Transformation in Czechia. *AUC Geographica*. Dostupné z: https://karolinum.cz/data/clanek/3189/Geogr_2_2016_08_Sprocha.pdf. <https://doi.org/10.14712/23361980.2016.18>.
- Waldaufová, E. – Šťastná, A. 2025. Mezi ideálem a realitou: současný pokles porodnosti očima české populace. *Demografie, revue pro výzkum populačního vývoje*, 67(2), s. 83–101. <https://doi.org/10.54694/dem.0359>.
- Wei, T. – Simko, V. 2024. *R package „corrplot“: Visualization of a Correlation Matrix*. <https://github.com/taiyun/corrplot>.
- Wickham, H. – Bryan, J. 2025. *readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- Wickham, H. – François, R. – Henry, L. – Müller, K. – Vaughan, D. 2023. *dplyr: A Grammar of Data Manipulation* [Software]. R package version 1.1.4. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, H. – Chang, W. – Henry, L. – Pedersen, T. L. – Takahashi, K. – Wilke, C. – Woo, K. – Yutani, H. – Dunnington, D. – Brand, T. van den. 2025. *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://ggplot2.tidyverse.org>.
- Wickham, H. – Vaughan, D. – Girlich, M. 2024. *tidyr: Tidy Messy Data*. <https://tidyr.tidyverse.org>.

KRISTÝNA DVOŘÁKOVÁ

Je studentkou magisterského programu Ekonomická demografie na Fakultě informatiky a statistiky VŠE v Praze, dále působí jako výzkumný pracovník na Ústavu informatiky AV ČR. Věnuje se kvantitativním metodám a modelování demografických procesů.

FILIP HON

Je absolventem Fakulty informatiky a statistiky VŠE v Praze. Od roku 2018 je členem Katedry demografie VŠE v Praze a výzkumným pracovníkem RILSA v Praze. Zabývá se demografickým prognózováním, výzkumem plodnosti a sociální ochranou.

PŘÍLOHA / APPENDIX

Příloha 1: Vývoj sledovaných ukazatelů v letech 2015, 2019 a 2023 pro Prahu

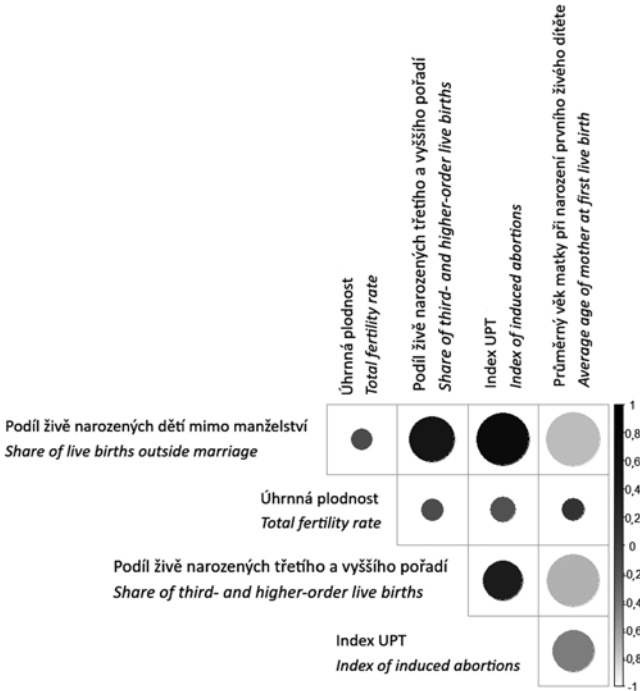
Changes in the selected indicators in 2015, 2019, and 2023 for the district of the Capital City of Prague

Ukazatel / Indicator	Rok / Year		
	2015	2019	2023
Podíl živě narozených dětí mimo manželství (v %) Share of live births outside marriage (%)	61,08	56,35	60,39
Úhrnná plodnost / Total fertility rate	1,50	1,55	1,33
Podíl živě narozených třetího a vyššího pořadí (v %) Share of third- and higher-order live births (%)	10,73	10,42	9,82
Index UPT / Index of induced abortions	13,14	10,46	11,45
Průměrný věk matky při narození prvního živého dítěte Average age of mother at first live birth	30,09	30,24	30,58

Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: Authors' calculation based on CZSO data.

Příloha 2: Korelační matice sledující závislost mezi proměnnými v roce 2015

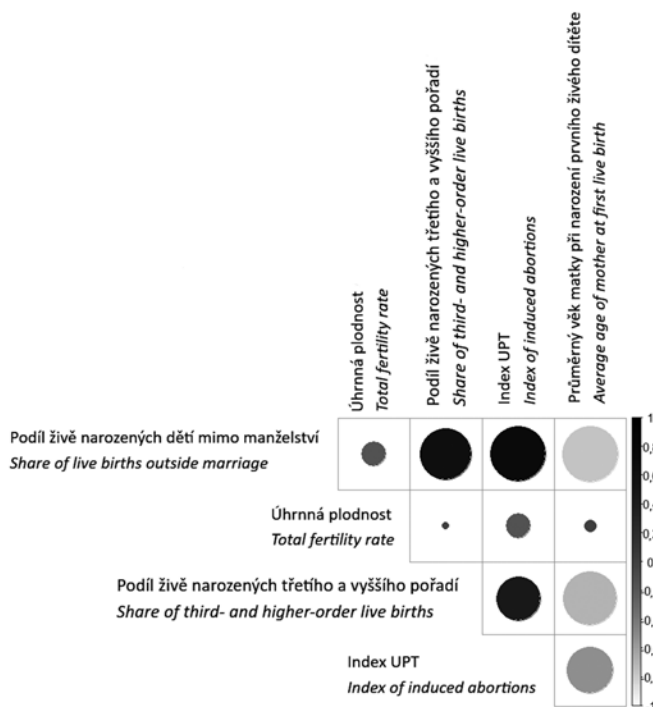
Correlation matrix monitoring the dependence between variables in 2015



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: Authors' calculation based on CZSO data.

Příloha 3: Korelační matice sledující závislost mezi proměnnými v roce 2019

Correlation matrix monitoring the dependence between variables in 2019



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: Authors' calculation based on CZSO data.

SUMMARY

Our analysis of reproductive behaviour across Czech districts in 2015, 2019, and 2023 shows persistent and substantial spatial differences. Descriptive statistics confirm variation in both the levels and variability of key indicators, while also revealing shifts in their distribution and identifying districts that consistently diverge from national patterns. Correlation analysis indicates stable relationships between variables, reflecting the distinct regional reproductive strategies that remain visible across the entire period.

Cluster analysis divided the districts into five groups with clearly differentiated profiles. One cluster is characterised by a low mean age at first birth, a high share of nonmarital births, and a higher index of induced abortions for nonmedical reasons. Another cluster shows the opposite pattern, with older firsttime mothers and low values for the remaining indicators.

A third cluster groups districts with average values, while a fourth is marked by share of third- and higher-order live births and a fifth by a distinctively high index of induced abortions for nonmedical reasons. These patterns demonstrate that reproductive behaviour in Czech districts can be differentiated into several identifiable types that are shaped by multiple demographic characteristics rather than geography alone.

The study achieved its objectives of describing regional differentiation, comparing developments across the three years, and identifying five clusters that capture how reproductive patterns are grouped across Czech districts. The findings provide an up-to-date and comprehensive view of the spatial structure of reproductive behaviour and offer a solid basis for further socio-economic and cultural interpretation, especially in relation to long-term demographic change.