

Typologie zaměstnanců a její vliv na míru pracovní absence: Shluková analýza

Řízení lidských zdrojů (HRM) prošlo v posledních desetiletích významným vývojem a posunulo se od čistě administrativní funkce směrem ke strategickému partnerství v rámci podniku. Boxall a Purcell (2000) ve své klíčové práci upozorňují, že strategické řízení lidských zdrojů (SHRM) se nezabývá pouze izolovanými personálními praktikami, ale zkoumá, jak tyto systémy přispívají k celkové efektivitě organizace a její životaschopnosti. Moderní pojetí SHRM, jak uvádějí Boon et al. (2018), se snaží integrovat pohled na strategický lidský kapitál s konkrétními HR praktikami, přičemž klade důraz na to, že zaměstnanci představují unikátní zdroj, který může firmě zajistit konkurenční výhodu.

Základním předpokladem pro efektivní řízení tohoto kapitálu je pochopení, že pracovní síla není homogenní masou. Granillo-Velasquez et al. (2024) ve své studii o mapování moderní pracovní síly zdůrazňují, že dynamika trhu práce a měnící se požadavky na kompetence (KSAOs) vyžadují detailnější vhled do profilů zaměstnanců. Tradiční přístup „jedna velikost pro všechny“ (one-size-fits-all) se ukazuje jako nedostatečný, protože ignoruje variabilitu v potřebách, motivacích a životních situacích jednotlivých pracovníků.

Právě zde nachází své uplatnění shluková analýza (Cluster Analysis) jako metoda umožňující segmentaci zaměstnanců do smysluplných skupin (typologií). Na rozdíl od tradičních regresních modelů, které hledají lineární vztahy mezi proměnnými, shluková analýza umožňuje identifikovat přirozené skupiny osob, které sdílejí podobné charakteristiky. Tento přístup v kontextu HR úspěšně aplikovaly například Švarcová a Stavjaničková (2013), které pomocí shlukové analýzy zkoumaly profesní preference studentů a budoucích HR manažerů, čímž prokázaly užitečnost této metody pro identifikaci specifických profilů v rámci širší skupiny respondentů.

V kontextu této práce je klíčové zaměření nejen na pracovní výkon, ale i na zdraví a životní styl zaměstnanců (well-being). Výzkum Thøgersen-Ntoumani a Foxe (2005) představuje zásadní teoretické východisko pro tento přístup. Tito autoři využili shlukovou analýzu k vytvoření typologie zaměstnanců na základě jejich fyzické aktivity a duševní pohody. Identifikovali specifické skupiny, jako jsou „nešťastní zaměstnanci“ (unhappy employees) či „sportující šťastní zaměstnanci“ (exercising happy employees), a prokázali, že tato segmentace umožňuje organizacím lépe cílit intervence v oblasti zdraví a benefitů.

Cílem této analýzy je navázat na uvedené teoretické poznatky a vytvořit typologii zaměstnanců ve sledované organizaci. Analýza se zaměří na identifikaci vnitřně homogenních skupin na základě kombinace demografických údajů, logistických faktorů (dojezdová vzdálenost) a ukazatelů životního stylu (BMI, kouření). Tento postup umožní managementu opustit plošná opatření a přistoupit k diferencovanému řízení lidských zdrojů, které lépe reflektuje realitu heterogenní pracovní síly.

Popis datového souboru a proměnných

Empirická část této práce vychází z datového souboru „Absenteeism at work“, který je archivován v mezinárodní databázi UCI Machine Learning Repository. Původní sběr dat a jejich následné zpracování probíhalo v rámci akademického výzkumu na Universidade Nove de Julho (Postgraduate Program in Informatics and Knowledge Management).

Datová matice obsahuje celkem 740 záznamů, které poskytují detailní přehled o profilech zaměstnanců. Jedná se o vysoce variabilní dataset, který umožňuje modifikaci atributů (na kategoriální, celočíselné či reálné typy) a jejich vzájemnou kombinaci dle specifických výzkumných záměrů.

Pro analýzu byly vybrány proměnné, které pokrývají tři klíčové oblasti – sociodemografické faktory, pracovní podmínky a životní styl:

Sociodemografické a kvalifikační ukazatele:

- **Age** (věk): biologický věk zaměstnance, který může souviset s životní fází a prioritami
- **Education** (vzdělání): ordinální proměnná kategorizující dosaženou úroveň vzdělání (od základního po vysokoškolské)
- **Son** (počet dětí): počet nezaopatřených dětí v rodině. Tato proměnná je zásadní pro identifikaci skupiny zaměstnanců s vysokou rodinnou zátěží

Logistické a pracovní faktory:

- **Service time** (délka praxe): počet let odpracovaných u současného zaměstnavatele, indikující loajalitu a zkušenost
- **Distance from Residence to Work** (vzdálenost): geografický údaj v kilometrech měřící vzdálenost mezi bydlištěm a pracovištěm
- **Transportation expense** (náklady na dopravu): měsíční finanční náklady spojené s dojížděním
- **Work load Average/day** (pracovní zátěž): průměrná denní pracovní zátěž, která slouží jako indikátor tlaku na výkon

Zdraví a životní styl:

- **BMI** (Body mass index): indikátor tělesné hmotnosti, který může signalizovat zdravotní rizika
- **Social drinker** (konzumace alkoholu): binární proměnná (ano/ne) zachycující životní styl zaměstnance v oblasti konzumace alkoholu
- **Social smoker** (kouření): binární proměnná (ano/ne). Zahrnutí tohoto faktoru je klíčové pro oddělení zdravotně rizikových skupin od zaměstnanců se zdravým životním stylem.

Před samotným provedením shlukové analýzy prošel datový soubor procesem čištění a selekce proměnných. Cílem této fáze bylo odstranit atributy, které by mohly zkreslit výslednou typologii zaměstnanců nebo které nenesou relevantní informaci pro strategické řízení lidských zdrojů.

Pro samotnou shlukovou analýzu tak bylo využito 10 klíčových proměnných, které reprezentují stabilní sociodemografické, zdravotní a pracovní charakteristiky zaměstnance. Hodnoty těchto proměnných byly před výpočtem standardizovány (z-skóre), aby bylo zamezeno zkreslení vlivem rozdílných měřítek jednotlivých atributů.

Metodologie

K identifikaci specifických profilů (typologie) zaměstnanců byla zvolena metoda **shlukové analýzy**. Jedná se o vícerozměrnou statistickou techniku, jejímž cílem je roztřídit objekty (zaměstnance) do relativně homogenních skupin tak, aby vnitroskupinová podobnost byla maximalizována a meziskupinová odlišnost byla co největší. Výpočet byl proveden v softwaru TIBCO Statistica.

Vzhledem k tomu, že vstupní proměnné nabývaly rozdílných měrných jednotek a řádů, byla data v prvním kroku **standardizována** (transformace na z-skóre). Tento krok zajistil, že proměnné s vyšším rozptylem neúměrně neovlivnily výsledky shlukování.

Jako klasifikační algoritmus byla použita **nehierarchická metoda K-průměrů**. Tato metoda přiřazuje jednotlivé případy do předem stanoveného počtu shluků (k) na základě minimalizace vzdálenosti mezi objektem a středem shluku. Jako míra vzdálenosti byla zvolena **Eukleidovská metrika**.

Kvalita rozdělení a statistická významnost rozdílů mezi vytvořenými shluky byla pro jednotlivé proměnné posouzena **pomocí analýzy rozptylu** (ANOVA). Za statisticky významné rozdíly mezi profily jsou považovány hodnoty $p < 0,05$.

K interpretaci výsledných shluků a jejich pojmenování byly využity grafy průměrů. Tyto grafy vizualizují průměrné hodnoty standardizovaných proměnných pro každý shluk, což umožňuje identifikovat charakteristické rysy daného typu zaměstnance.

Analýza

Pro hlubší pochopení struktury zaměstnanců byla realizována nehierarchická shluková analýza (metoda k-průměrů). Na základě analýzy rozptylu (ANOVA) byla z dalšího zkoumání vyřazena proměnná *Work load Average/day (pracovní zátěž)*, která nevykazovala statistickou významnost ($p = 0,219$), a nesloužila tak jako relevantní diskriminátor. Finální model pracuje s **9 standardizovanými proměnnými**, u kterých byla potvrzena vysoká statistická významnost rozdělení ($p < 0,001$). Na základě matematických kritérií a interpretovatelnosti byl stanoven optimální počet **3 shluků**.

Tabulka 1: Konečné středy shluků (standardizovaná z-skóre)

	Shluk 1	Shluk 2	Shluk 3
Náklady na dopravu			0,78
Vzdálenost do práce	0,52		
Délka praxe		0,50	
Věk		0,52	
Vzdělání			0,95
Děti			0,79
Konzumace alkoholu	0,52		
Kouření			0,82
BMI		0,62	

Tabulka 2: Analýza rozptylu (ANOVA)

Proměnná	SS Between	SS Within	F	signif. p
Transportation expense	98,93	640,07	56,96	0,000000
Distance from work	304,48	434,52	258,22	0,000000
Service time	77,24	661,76	43,01	0,000000
Age	330,33	408,67	297,87	0,000000
Education	638,46	100,54	2340,16	0,000000
Son	53,47	685,53	28,74	0,000000
Social drinker	231,12	507,88	167,69	0,000000
Social smoker	68,03	670,97	37,36	0,000000
Body mass index	133,37	605,63	81,15	0,000000

Na základě analýzy konečných středů (viz Tabulka 1) byly identifikovány tři specifické skupiny, které se liší socio-demografickými charakteristikami i životním stylem:

- **Shluk 1: Lokálně orientovaní zaměstnanci s aktivním společenským životem** (N = 470)

Tento nejpočetnější segment zahrnuje pracovníky s mírně nadprůměrnou dojezdovou vzdáleností (0,52) a sklonem k příležitostné konzumaci alkoholu (0,52) Ostatní parametry, jako věk či rodinná zátěž, se u této skupiny pohybují kolem průměru souboru.

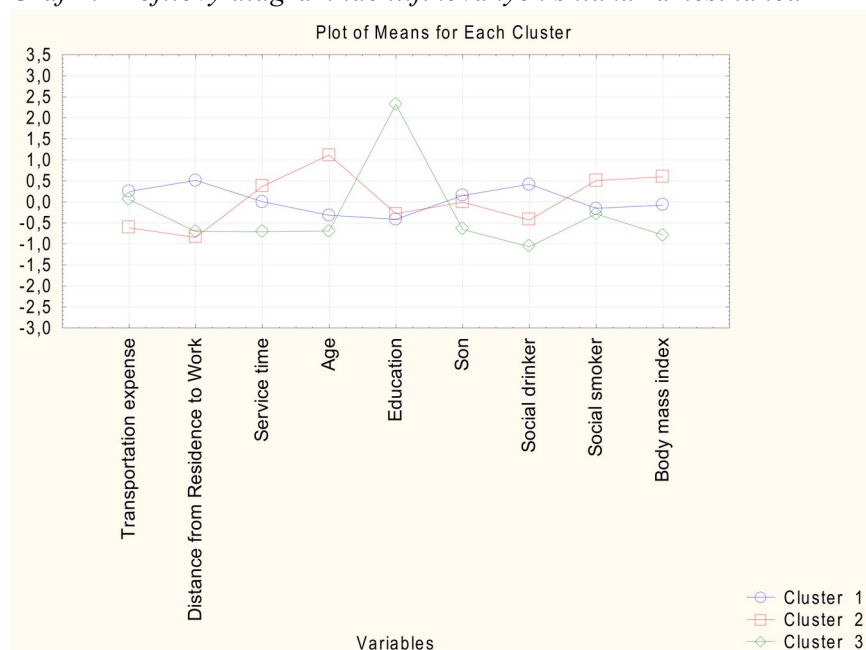
- **Shluk 2: Stabilní loajální jádro se zvýšeným zdravotním rizikem** (N = 120)

Skupina je definována nejvyšším průměrným věkem (0,52) a nejdelší dobou působení v podniku (0,50). Charakteristickým rysem je nejvyšší index tělesné hmotnosti – BMI (0,62) což v kombinaci s věkem naznačuje potenciální náchylnost k absencím ze zdravotních důvodů spojených s fyziologickými faktory stárnutí.

- **Shluk 3: Vysoce vzdělaní dojíždějící s rodinnou a logistickou zátěží** (N = 150)

Tento shluk vykazuje největší statistickou odchylku od průměru. Zahrnuje zaměstnance s nejvyšším vzděláním (0,95) a největším počtem dětí (0,79). Tato skupina čelí extrémním nákladům na dopravu (0,78) a vykazuje nejsilnější sklon ke kouření (0,82), což může sloužit jako nevhodný mechanismus zvládání stresu (coping mechanism) spojeného s logistickým a rodinným zatížením.

Graf 1: Profilový diagram identifikovaných shluků zaměstnanců



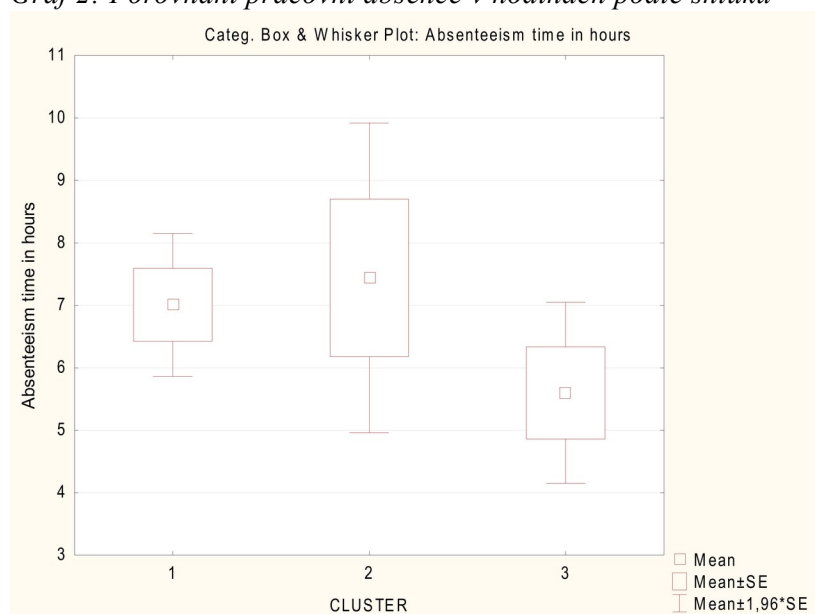
Následně byla testována závislost mezi identifikovanou typologií a reálnou pracovní absencí měřenou v hodinách (viz Tabulka 3). Výsledky analýzy průměrů (Breakdown) prokázaly existenci významných rozdílů v docházce jednotlivých skupin.

Z naměřených dat vyplývá, že nejvyšší průměrnou absencí vykazuje **Shluk 3** (7,57 hod). Vysoká hodnota směrodatné odchylky v této skupině (18,63) však naznačuje přítomnost extrémních případů s velmi vysokou neúčastí. Těsně v závěsu se nachází **Shluk 1** (7,43 hod). Naopak statisticky nejspolehlivějším segmentem je **Shluk 2** (4,15 hod), tedy starší loajální pracovníci, kteří i přes vyšší BMI vykazují nejnižší a nejvíce vyrovnanou míru absence (SD = 4,87).

Tabulka 3: Deskriptivní statistika pracovní absence podle příslušnosti ke shluku

Shluk (CLUSTER)	Mean (průměrná absence/hodiny)	N (počet osob v shluku)	Std.Dev. (Směrodatná odchylka)
1. Lokální skupina	7,43 hod.	470	12,71
2. Loajální senioři	4,15 hod.	120	4,87
3. Vzdělání dojíždějící	7,57 hod.	150	18,63
Celý soubor	6,92 hod.	740	13,33

Graf 2: Porovnání pracovní absence v hodinách podle shluků



Získané výsledky naznačují, že v rámci sledovaného podniku není primárním zdrojem absence věk či zdravotní stav (reprezentovaný BMI), ale spíše kombinace rodinné zátěže, logistické náročnosti dojíždění a s tím spojeného životního stylu (kouření).

Managementu podniku lze doporučit:

1. Cílenou podporu Shluku 3: Implementace flexibility (home-office, úprava začátku pracovní doby) pro snížení tlaku na logistiku u rodičů s vysokou kvalifikací.
2. Prevenici u Shluku 2: Udržení loajality seniorních pracovníků skrze programy podpory zdraví a ergonomie, které zmírní rizika plynoucí z vyššího BMI.
3. Audit životního stylu: Zaměření na antistresové programy a podporu zdravého životního stylu u skupin s vysokým podílem kouření a konzumace alkoholu.

Zdroje

Boon, C., Eckardt, R., Lepak, D. P., & Boselie, P. (2018). Integrating strategic human capital and strategic human resource management. *The International Journal of Human Resource Management*, 29(1), 34-67.

Boxall, P., & Purcell, J. (2000). Strategic human resource management: where have we come from and where should we be going?. *International Journal of Management Reviews*, 2(2), 183-203.

Granillo-Velasquez, K. E., Reyes, D. L., & Musemeche, N. P. (2024). Mapping the modern workforce: An overview of workforce competency demands. *The Career Development Quarterly*.

Švarcová, J., & Stavjaníčková, I. (2013). Cluster Analysis of Professional Focus of Future HR Managers. *WSEAS Transactions on Business and Economics*, 10(3), 249-256.

Thøgersen-Ntoumani, C., & Fox, K. R. (2005). Physical activity and mental well-being typologies in corporate employees: A mixed methods approach. *Work & Stress*, 19(1), 50-67.