**Statistická analýza dat II. – PSY 252**

**Ondřej Sedlák (414830), Lenka Tarabíková (414635)**

**Seminární úkol č. 5**

Pro faktorovou analýzu jsme si zvolili soubor s názvem dataprofa.xlxs. Zvolili jsme několik relevantních proměnných: obohacování, užitečnost, nasazení, příprava, budoucnost, uplatnění, časová náročnost a známka.

Tabulka 1: popisné statistiky proměnných použitých v analýze

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | N | Min | Max | M | SD |
| Obohacuje | 24 | 1 | 4,8 | 2,24 | 0,82 |
| Užitečný | 24 | 1 | 4,4 | 2,05 | 0,81 |
| Nasazení | 24 | 1,2 | 4,8 | 2,76 | 1,07 |
| Příprava | 24 | 1,2 | 5,3 | 3,14 | 1,21 |
| Budoucnost | 24 | 1 | 3,7 | 1,71 | 0,74 |
| Uplatnění | 24 | 1,2 | 4,1 | ,28 | 0,71 |
| Časová náročnost | 24 | 0,5 | 4,4 | 2,03 | 1,17 |
| Znalosti | 24 | 0,0% | 100% | 54% | 36% |

Provedli jsme kontrolu předpokladů pro faktorovou analýzu. Proměnné jsou měřeny alespoň na ordinální úrovni měření, mají normální rozložení a u každé máme odpovědi od více než 15 respondentů.

Na základě faktorové analýzy těchto proměnných jsme nechali vykreslit scree plot, abychom mohli upravit počet faktorů, který je ideální pro naši analýzu. Oblast zlomu byla u trojky, a proto jsme zvolili počet faktorů 3.

Graf 1: Scree plot



 Jejich počet byl ale snížen o 1 faktor, jelikož pouze 2 faktory mají hodnotu eigenvalue vyšší než 1. Výsledné 2 faktory redukují velmi dobře původní proměnné, protože dohromady tvoří 89,8 % rozptylu. První faktor je sycen proměnnými zaměřující se na přínosnost předmětu: obohacuje, užitečný, budoucnost, příprava a známka. Druhý faktor je sycen proměnnými týkající se celkové náročnosti předmětu: nasazení, příprava, časová náročnost.

Tabulka 2: Faktorová matice

|  |  |  |
| --- | --- | --- |
|  | Přínosnost předmětu | Celková náročnost na přípravu |
| Obohacuje | 0,97 |  |
| Užitečný | 0,98 |  |
| Nasazení |  | 0,94 |
| Příprava |  | 0,96 |
| Budoucnost | 0,92 |  |
| Uplatnění | 0,88 |  |
| Časová náročnost (průměrně kolik hodin za týden) |  | -0,92 |
| Známka předmětu (celkový průměr) | 0,97 |  |
| Eigenvalues | 4,49 | 2,69 |
| % of variance | 56,12 | 33,64 |

Podle Kaiser-Meyer-Olkin Measure of Sampling Adequacy, který určuje míru vhodnosti vybraných dat s hodnotou 0,77, jsme zjistili, že naše data jsou středně vhodná pro naši analýzu. Podle Bartlett's Test of Sphericity (224,2; 28) který byl signifikantní (p < 0,01) jsme zjistili, že proměnné na sobě v základním modelu závisí.

Pomocí faktorové analýzy jsme tedy zjistili, že 8 výše zmíněných proměnných lze nahradit proměnnými přínosnost předmětu a celková náročnost na přípravu.