

## IBM SPSS CATEGORIES

### Odhalte skryté vazby v kategorizovaných datech

Využijte veškeré informace skryté ve vašich datech. Odhalit tyto informace vám pomohou percepční mapy, optimální škálování a techniky redukce počtu dimenzí. Modul IBM SPSS Categories poskytuje ucelený systém metod, které potřebujete pro podrobnou analýzu, přesnou interpretaci a hluboký vhled do vícerozměrných kategorizovaných dat a do vztahů mezi nimi. IBM SPSS Categories využijete například ke zjištění, které vlastnosti si zákazníci nejvíce spojují s vaším výrobkem či značkou, nebo k porozumění toho, jak zákazníci vnímají vaše produkty v porovnání s jinými.

IBM SPSS Categories nabízí regresní analýzy v případě, kdy je závislá i nezávislá proměnná numerická, ordinální nebo nominální. IBM SPSS Categories dále nabízí vizuální výstupy pro interpretaci dat a sledování asociací řádkových a sloupcových kategorií v rozsáhlých nebo přehledných dvou i vícerozměrných tabulkách četností. To vám dovoluje:

- zjistit vztahy v ordinálních a nominálních datech pomocí obdobných procedur jako je klasická lineární regrese, metoda hlavních komponent či metoda kanonických korelací,
- pracovat s rezidui s nenormálním rozdělením nebo s nelineárními vztahy mezi závislou a nezávislou proměnnou, nabízejí se možnosti využít hřebenovou regresí, metodu Lasso a Elastic Net, možnosti výběru proměnných a modelu pro číselná i kategorizovaná data,
- využít biploty a triploty pro zjištění vztahu mezi případy, kategoriemi a proměnnými nebo jejich sadami,
- znázornit podobnost mezi dvěma skupinami případů pomocí vzdáleností v procentuálních mapách.

## Nejdůležitější:

IBM SPSS Data Preparation umožňuje:

- znázorníte a objevíte komplexní kategorizovaná a číselná data, stejně jako vícerozměrná data
- porozumíte informacím pomocí dvou či vícerozměrných tabulek
- využijte biploty, triploty a percepční mapy pro zjištění vztahu mezi vašimi daty

## Přeměňte svá kvalitativní data na kvantitativní

Pokročilé procedury, které jsou k dispozici v IBM SPSS Categories, vám dovolí provádět komplikované statistické operace na kategorizovaných datech.

Lepší interpretace mnohorozměrných dat a hlubší analýza vztahů v datech s možnostmi IBM SPSS Categories

- **regresní procedury pro kategorická data** odhalují vztah nominální, ordinální i číselné proměnné na kombinaci číselných i nominálních proměnných
- **optimální techniky pro škálování** kvantifikují proměnné pro dosažení maximálního indexu determinace
- **techniky pro redukci dimenzí** pomáhají odhalovat vztahy v datech s využitím přehledných preceptuálních map a biplotů
- **souhrnné grafy** zobrazují vztahy mezi proměnnými nebo kategoriemi pro snadnější nalezení vztahu mezi dvěma i více proměnnými

Optimální škálování přiřadí kategoriím nominální či ordinální proměnné hodnoty na kvantitativní škále (určí její měřicí jednotky a nulovou hodnotu). S takovou kvantifikací se otevírají cesty k použití dalších statistických funkcí, metod a modelů, které předpokládají číselné proměnné.

Regresní metoda s optimálním škálováním rozšíří možnosti klasické regrese na případ, kdy jsou závislé i nezávislé proměnné směsí číselných, ordinálních a nominálních dat. Současná verze modulu obsahuje nástroje pro výběr modelu a jeho vyladění. Jednoduchá a vícerozměrná korespondenční analýza vám pomůže číselně ověřit podobnosti a asociace mezi dvěma nebo více nominálními proměnnými a jejich kategoriemi. Jinou možností je analyzovat libovolnou tabulku s nezápornými čísly pomocí korespondenční analýzy.

Svá data lépe pochopíte užitím kategorické metody hlavních komponent. Sumarizujte data užitím hlavních komponent odvozené na proměnných různých typů (nominální, ordinální, číselné). Využijete biploty a triploty pro zobrazení vztahů mezi případy, kategoriemi a proměnnými. Tyto možnosti jsou k dispozici i pro číselná data.

Optimální škálování vytvoří korelační matici i pro kvantifikovaná nominální a ordinální data. Kromě toho můžete také hledat vztahy mezi skupinami proměnných (různých typů) a analyzovat je metodou nelineární kanonické korelační analýzy.

## Odhalte graficky skryté vztahy

Ať již jsou vaše data jakéhokoli původu - segmentace trhu, subpopulace, politické názory a volební preference či biologické experimenty - metody optimálního škálování modulu IBM SPSS Categories vás osvobodí od obvyklého omezeného pohledu na data, založeného na separovaných dvourozměrných tabulkách. Modul nabízí nový komplexní pohled na vztahy v datech a rychlý vhled do skrytých strukturních vazeb. K dispozici máte nejen statistický přehled číselných výsledků metod, ale především korespondenční (percepční, konkurenční, vztahové) mapy svých mnohorozměrných dat.

Techniky snížení dimenzionality dovolují překročit obvyklé možnosti studia velkého počtu složitých a poněkud nepraktických tabulek. Vztahy ve svých datech uvidíte pomocí percepčních map a biplotů jasně a názorně.

Percepční mapy jsou sumarizační grafy, zobrazující blízko sebe asociované (podobně se chovající) proměnné a kategorie. Graf představuje jedinečný náhled na vztahy mezi více než dvěma kategorizovanými proměnnými.

- Biploty dovolují sledovat vztahy mezi případy, proměnnými a kategoriemi. Například můžete zkoumat zákazníky, značky a demografické charakteristiky.

S procedurou preferenčního škálování jsou k dispozici ještě širší možnosti vizualizace vztahů mezi objekty. Průlomový algoritmus, na kterém je založena tato procedura, dovoluje provádět nemetrické analýzy ordinálních dat se smysluplnými výsledky. Procedura pro škálování proximit mezi případy sleduje podobnost případů s využitím informace o jejich vlastnostech.

## Jak můžete využít modul IBM SPSS Categories

Následující procedury jsou k dispozici, aby odhalily význam vztahů a podpořily hlubší pohled vašich analýz.

**Kategorická regresní analýza (CATREG)** poskytuje modely predikce kategorizované závislé proměnné na základě kategorií směsi nezávislých nominálních, ordinálních nebo i numerických proměnných.

Procedura kvantifikuje kategorie pomocí technik optimálního škálování tak, aby byla dosažena maximální korelovanost mezi závislou a nezávislými proměnnými.

Regresi s optimálním škálováním můžete použít například pro zjištění, jak spokojenost s prací závisí na zaměstnanecké kategorii, geografickém regionu a na nutnosti dojíždět do práce.

Optimální škálování kvantifikuje proměnné tak, aby bylo dosaženo co nejvyššího koeficientu determinace. Optimální škálování může být použito i na číselné proměnné v případě, že rezidua neodpovídají normálnímu rozdělení, nebo pokud je vztah mezi závislou a nezávislou proměnnou nelineární. K dispozici jsou tři nové metody: hřebenová regrese, metoda Lasso a Elastic Net. Metody zvyšují přesnost a stabilitu odhadu parametrů. Automatický výběr prediktorů dovolí analyzovat rozsáhlé datové soubory, i ty co mají více proměnných než případů. Možnost škálovat i číselná data rozšiřuje použití metod Lasso a Elastic net i na tento typ dat.

V rámci kategorické regresní analýzy je k dispozici i zobecněný aditivní model (GAM) pro kategorizovaná i číselná data.

**Korespondenční analýza (CORRESPONDENCE)** analyzuje dvourozměrné tabulky, které obsahují libovolnou míru korespondence mezi řádky a sloupci a zobrazuje tyto závislosti v grafu. Nejběžnějším typem takové tabulky je kontingenční tabulka zobrazující četnosti.

Korespondenční analýza popisuje vztahy mezi dvěma nominálními proměnnými v prostoru co nejnižší dimenze (nejčastěji v rovině). V percepční mapě potom graficky zobrazí současně vztahy mezi těmito proměnnými i vztahy mezi jejich kategoriemi.

**Mnohorozměrná korespondenční analýza (MULTIPLE CORRESPONDENCE)** je určena pro analýzu vícerozměrných kategorizovaných dat. Od korespondenční analýzy se liší tím, že dovoluje analyzovat současně více než dvě proměnné. Procedura považuje všechny analyzované proměnné za nominální (neuspořádané kategorie).

Mnohorozměrnou korespondenční analýzu můžete využít například pro znázornění vztahů mezi oblíbenými televizními pořady, věkovými kategoriemi a pohlavím. V grafickém zobrazení potom uvidíte, které skupiny diváků přitahují jednotlivé pořady, a současně pochopíme, které pořady jsou si z hlediska těchto kategorií nejpodobnější.

**Kategorická analýza hlavních komponent (CATPCA)** užívá optimální škálování ke zobecněním analýzy hlavních komponent pro skupinu proměnných různého typu. Na rozdíl od mnohorozměrné korespondenční analýzy je možno specifikovat typ každé proměnné (nominální, ordinální, číselná), který je pak v analýze respektován.

Zobrazte například vztah mezi různými značkami automobilů (případy) a jejich charakteristikami (cena, váha, spotřeba atd.). Současně můžete automobily popsat podle jejich třídy (nižší, vyšší, rodinné vozy atd.) a CATPCA použije tuto klasifikaci pro seskupení bodů reprezentujících jednotlivé automobily. Vyšší váha přiřazená proměnné definující třídu automobilu zajistí, že automobily jsou úzce soustředěny kolem bodů reprezentující jejich třídu. IBM SPSS Categories nakreslí komplexní vztahy mezi případy, jejich skupinami a proměnnými do přehledné mapy, které snadno porozumíte.

**Nelineární kanonická korelační analýza (OVERALS)** používá optimální škálování a je zobecněním kanonické korelační analýzy pro uspořádání proměnných různých typů. Tento typ analýz dovoluje porovnávat několik skupin proměnných v jednom grafu mezi sebou poté, co jsou odstraněny korelace mezi skupinami.

Analýzovat můžete například vlastnosti různých produktů (např. polévek) ve studii chuti. Hodnocení porotců reprezentují jednotlivé proměnné data matice, resp. skupiny hodnotících kritérií (skupiny proměnných), jednotlivé polévky jsou uvedeny jako řádky (případy). Po odstranění korelací u každého posuzovatele, OVERALS zprůměruje proměnné a kombinuje různé charakteristiky pro zobrazení vztahu mezi polévkami.

Alternativně může každý porotce hodnotit podle vlastní množiny charakteristik (proměnných), čímž pak každý zvlášť vytváří skupinu. V tomto případě OVERALS provede průměr napříč charakteristikami po odstranění závislostí a potom zkombinuje skóry jednotlivých porotců. Analýza také zobecňuje regresní metodu pro případ, že závislých proměnných je více. Proměnné jsou pak souhrnně předpovídány skupinou nezávislých proměnných.

**Mnohorozměrné škálování (PROXSCAL)** provádí analýzu jedné nebo více matic obsahujících míry proximity (míry podobnosti nebo nepodobnosti/vzdálenosti). Vstupy jsou určeny uživatelem nebo spočteny jako proximity mezi případy v matici vícerozměrných dat.

PROXSCAL zobrazuje vztahy mezi objekty jako vzdálenosti v rovině nebo vícerozměrné mapě tak, abychom snadno pochopili a mohli pohodlně a správně vizuálně interpretovat mnoho vztahů současně. V případě násobných paralelních matic proximít, PROXSCAL analyzuje to, co mají společné a graficky zobrazuje rozdíly mezi nimi.

Například PROXSCAL lze použít pro zobrazení podobností různých chutí Cola nápojů preferovaných zákazníky různých věkových skupin. Mladší lidé jsou citlivější na rozdíly mezi tradiční a novou chutí, zatímco starší výrazněji vnímají rozdíl mezi dietním a nediетním nápojem.

**Preferenční škálování (PREFSCAL)** vizuálně vyšetřuje vztahy mezi proměnnými. Preferenční škálování provádí mnohorozměrné rozprostření dvou skupin objektů za účelem nalezení společné kvantitativní škály.

Výsledky jsou zaznamenány do grafu, kde je vztah mezi oběma skupinami objektů zachycen vzdálenostmi mezi body v grafu. To připravuje prostor pro nalezení klastrů mezi proměnnými.

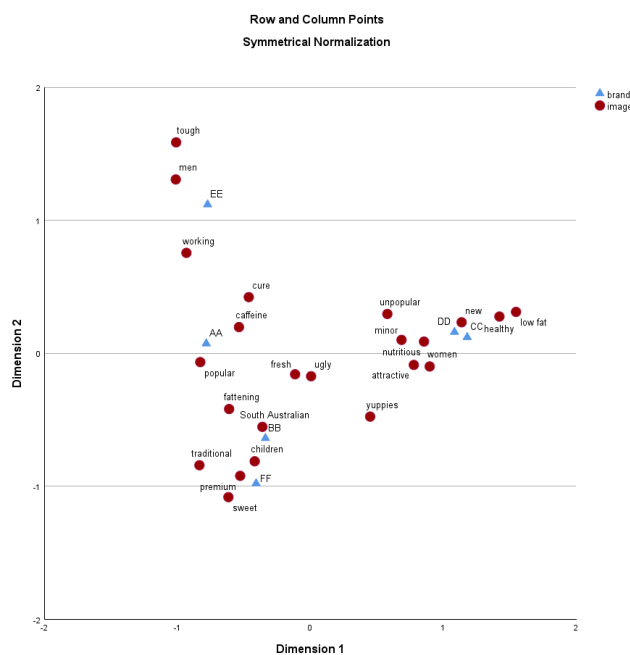
Například skupina řidičů hodnotí 26 modelů aut, přičemž hodnotí 10 vlastností na šestibodové stupnici. Vytváří se graf klastrů, který zachycuje, jak jsou si modely podobné a které vlastnosti jsou s nimi asociovány. Graf je vytvořen na datech ze všech deseti atributů. Další graf deseti atributů ukazuje váhu, s jakou ovlivňují jednotlivé atributy polohy bodů v grafu s klastry.

## Jak lépe porozumět vnímání zákazníků

Výzkumníci v jižní Austrálii se snažili lépe pochopit, jak zákazníci vnímají šest značek kávy. Požádali proto zákazníky, aby ohodnotili každou z těchto šesti značek kávy (označených AA až FF) pomocí dvaceti tří různých atributů. Z výsledné tabulky se 138 buňkami by však bylo pro analytiku obtížné nalézt vztahy mezi značkami a vnímanými atributy.

Výzkumníci proto použili korespondenční analýzu v IBM SPSS Statistics, aby identifikovali nejsilnější charakteristiky vztahů mezi značkami a atributy. Tím, že bylo každé značce i atributu přiřazeno specifické číslo v obou dimenzích, mohla být informace snadno zobrazena v jednoduchém a srozumitelném grafu, který se nazývá percepční mapa.

Z grafu například vyplývá, že značka AA je trhem nejvíce identifikovaná jako „populární“. Podobně výzkumníci snadno zjistí, že zákazníci, kteří se zajímají o zdravé a nízkotučné produkty si více považují značky CC a DD, zatímco FF je vnímána jako kvalitní a sladká káva.



Výzkumníci studovali zákazníky a jejich vnímání šesti značek kávy. Značky byly označeny AA až FF a jsou charakterizovány různými kategorizovanými atributy, jako například „kvalitní“. Korespondenční analýza v IBM Statistics vytvořila tuto percepční mapu.

## SPECIFIKACE

### CATREG

- *Kategorická regresní analýza metodou optimálního škálování*
- *úroveň optimálního škálování pro analýzu jednotlivých proměnných: spline ordinal, spline nominal, ordinal, nominal, multiple nominal nebo numeric*
- *diskretizace spojitých proměnných nebo konverze textové proměnné na celočíselnou násobením, seřazením nebo seskupením hodnot do předem daného počtu kategorií podle volitelné distribuce (normální nebo rovnoměrné) nebo seskupením hodnot předem daného intervalu do kategorií; uspořádání a seskupení lze použít také pro rekódování kategorizovaných dat*
- *chybějící hodnoty: nahrazení chybějících hodnot modelem proměnné, vytvoření zvláštní kategorie nebo vynechání celého řádku datové matice (listwise deletion)*
- *výběr metody pro každou proměnnou zvlášť*
- *specifikace objektů, které budou považovány za suplementární*
- *volba metody pro výpočet výchozí konfigurace*
- *kontrola počtu iterací*
- *volba konvergenčních kritérií*
- *grafy výsledků*
  - *transformační grafy (optimální kvantifikátory versus původní hodnoty v jednotlivých kategoriích)*
  - *grafy reziduí*
- *tisk výsledků*
  - *vícenásobné R, R2, adjustované R2*
  - *standardizované regresní koeficienty, standardní chyby, přímé korelace, částečné korelace (part correlation), parciální korelace, Prattova relativní míra důležitosti pro transformované prediktory, tolerance před a po transformaci, F statistika*
  - *tabulka popisných statistik zahrnující marginální četnosti, druh transformace, počet chybějících hodnot, modus*
  - *iterační historie*
  - *tabulky statistik modelů a parametrů: tabulka ANOVA se stupni volnosti podle úrovně optimálního škálování, statistiky modelů s adjustovaným koeficientem determinace, hodnoty t statistiky, signifikance, samostatná tabulka s přímými, parciálními a částečnými korelacemi, tolerance a důležitost před i po transformaci*
  - *korelace transformovaných prediktorů a vlastní čísla korelační matice*
  - *korelace původních prediktorů a vlastní čísla korelační*

*matice*

- *kvantifikace kategorií*
- *uložení transformovaných proměnných, predikovaných hodnot a reziduí do aktivního souboru*
- *zápis diskretizovaných a transformovaných proměnných do externího datového souboru*
- *odhady predikovaných hodnot doplňkových objektů pomocí regrese a optimálního škálování odvozených u aktivních objektů*
- *tři nové odhadové metody: hřebenová regrese, Lasso a Elastic net*
  - *zvýšená přesnost a stabilita odhadu parametrů*
  - *analýza rozsáhlých dat (více proměnných než případů)*
  - *automatický výběr prediktorů*
  - *zápis odhadnutých modelů a jejich koeficientů do zvláštního souboru pro další použití*
- *dvě nové metody pro výběr modelu a dosažení požadované přesnosti modelu: 632 bootstrap a křížová validace (Cross Validation CV)*
  - *nalezení optimálního modelu pro predikci s možnostmi algoritmu 632(+) a křížové validace*
  - *neparametrické odhady standardních chyb koeficientů metodou bootstrap*
- *použití více počátečních řešení*
  - *nalezení globálního optima i v případě, že jsou v modelu použity nemonotónní transformace*
  - *zápis odhadnutých regresních koeficientů do nového souboru pro další použití*

### CORRESPONDENCE

*Korespondenční analýza*

- *vstupní data mohou být ve formě souboru nebo tabulky*
- *volba suplementárních kategorií*
- *volba počtu dimenzí pro řešení*
- *výběr míry vzdálenosti*
  - *chí-kvadrát v korespondenční analýze*
  - *euklidovská vzdálenost pro analýzy typu biplot*
- *pět typů standardizace*
  - *odečtení řádkového průměru*
  - *odečtení sloupcového průměru*
  - *odečtení řádkového a sloupcového průměru*
  - *převedení řádkových součtů na stejnou hodnotu*
  - *převedení sloupcových součtů na stejnou hodnotu*
- *pět typů normalizace*
  - *symetrická*
  - *principal pro vzdálenosti mezi řádky a sloupci*
  - *principal pro vzdálenosti řádků*
  - *principal pro vzdálenosti sloupců*
  - *uživatelská (uživatelsky definovaná hodnota dovolující*

- volbu mezi *principal* pro vzdálenosti řádků a *principal* pro vzdálenosti sloupců)
- tiskové výstupy
  - tabulka
  - sumarizace: singulární hodnoty, inercie, podíl inercie pro každou dimenzi řešení a kumulativní podíly inercie, konfidenční statistika pro maximální počet dimenzí
  - řádkové i sloupcové statistiky pro spolehlivost: směrodatné odchylky a korelace aktivních řádkových bodů
  - permutovaná tabulka: řádky a sloupce seřazené podle řádkových a sloupcových skóre v dané dimenzi
  - řádkové profily
  - sloupcové profily
  - přehled řádkových a sloupcových bodů: relativní četnosti (*mass*), skóry, inercie, příspěvek bodů k inercii dimenze, příspěvek dimenze k inercii bodů
- výstupní grafy
  - transformace hodnot řádkových kategorií na řádkové skóry
  - transformace hodnot sloupcových kategorií na sloupcové skóry
  - řádkové skóry
  - sloupcové skóry
  - biplot - graf řádkových a sloupcových skóre
  - výstupní soubory
  - uložení řádkových a sloupcových skóre, rozptylů a kovariancí do souboru ve formě maticového datového souboru
- vhodné proměnné ve fixní konfiguraci
- výběr z pěti možností normalizace
  - *principal* pro proměnné (optimalizace asociace mezi proměnnými)
  - *principal* pro objekty (optimalizace vzdáleností mezi objekty)
  - symetrická (optimalizace vztahu mezi objekty a proměnnými)
  - nezávislá
  - uživatelská (uživatelsky definovaná hodnota dovolující volbu mezi normalizací proměnných a objektů)
- tisk výsledků
  - souhrnné informace o modelu
  - iterační statistiky a historie
  - popisné statistiky (četnosti, chybějící hodnoty, modus)
  - diskriminační míry podle proměnných a dimenzí
  - kvantifikace kategorií (centroidy), *mass*, inercie, příspěvek kategorií k inercii dimenzí, příspěvek dimenzí k inercii kategorií
  - korelační matice transformovaných proměnných a její vlastní čísla pro každou dimenzi
  - korelační matice původních proměnných a její vlastní čísla
  - skóry pro jednotlivé objekty
  - příspěvky objektů: *mass*, inercie, příspěvek kategorií k inercii dimenzí, příspěvek dimenzí k inercii kategorií
- grafické výstupy
  - grafy pro kategorie: body kategorií, transformace (optimální kvantifikace proti indikátorům kategorií), rezidua pro vybrané proměnné, a spojené zobrazení bodů kategorií vybraných proměnných
  - graf skóre pro jednotlivé objekty
  - graf měř diskriminace
  - biplot objektů a centroidů vybraných proměnných
- připojení transformovaných proměnných a skóre objektů do pracovního datového souboru
- zápis diskretizovaných dat, transformovaných dat a skóre objektů do externího datového souboru

## MULTIPLE CORRESPONDENCE

Mnohorozměrná korespondenční analýza (nahrazuje HOMALS z verzí starších než SPSS Categories 13.0)

- volba vah proměnných
- diskretizace spojitých proměnných nebo konverze textové proměnné na celočíselnou násobením, seřazením nebo seskupením hodnot do předem daného počtu kategorií podle volitelné distribuce (normální nebo rovnoměrné) nebo seskupením do kategorií předem daného intervalu. Uspořádání a seskupení lze užít také pro rekódování kategorizovaných dat
- volba objektů a proměnných, které mají v modelu suplementární roli
- volba počtu dimenzí v řešení
- specifikace způsobu zacházení s chybějícími hodnotami: vyloučení buněk datové matice, kde není platná hodnota, nahrazení chybějících hodnot modem proměnné nebo zvláštní kategorií, metoda listwise
- volba maximálního počtu iterací
- určení konvergenčních kritérií
- volba souborů obsahujících koordináty konfigurace a

## CATPCA

- analýza hlavních komponent metodou optimálního škálování
- úroveň optimálního škálování, podle které budou analyzovány jednotlivé proměnné: *spline ordinal*, *spline nominal*, *ordinal*, *nominal*, *multiple nominal* nebo *numeric*
- specifikace vah proměnných
- chybějící data: vynechávání *pairwise*, *listwise*, nahrazení modem proměnné nebo speciální kategorií
- určení objektů a proměnných se suplementární rolí
- určení počtu dimenzí
- vstup souborů obsahujících souřadnice proměnných,

kteře tvořív fixní konfiguraci a k nimž se datové proměnné porovnávají

- diskretizace spojitých proměnných nebo konverze textové proměnné na celočíselnou násobením, seřazením nebo seskupením hodnot do předem daného počtu kategorií podle volitelné distribuce (normální nebo rovnoměrné) nebo seskupením hodnot předem daného intervalu do kategorií. Uspořádaní a seskupení lze užít také pro rekódování kategorizovaných dat.
- výběr z pěti možností normalizace
  - proměnné (optimalizace asociace mezi proměnnými)
  - objekty (optimalizace vzdáleností mezi objekty)
  - symetrická (optimalizace vztahu mezi objekty a proměnnými)
  - nezávislá
  - uživatelská (uživatelsky definovaná hodnota dovolující volbu mezi proměnnou a objektovou normalizací)
- určení maximálního počtu iterací
- určení kritérií konvergence
- tisk výsledků:
  - souhrnné informace o modelu
  - iterační statistiky a historie
  - popisné statistiky (četnosti, chybějící hodnoty, modus)
  - variance vysvětlená proměnnou a dimenzí
  - koeficienty zátěže hlavních komponent pro proměnné s optimální úrovní vektorové kvantifikace
  - kvantifikace a souřadnice kategorií pro každou dimenzi (vektory a/nebo centroidy)
- objektové body a zátěže hlavních komponent
- objektové body a centroidy
- výběr proměnných pro biplot a triplot
- zátěže a centroidy
  - triplot objektových bodů, zátěží komponent pro proměnné s optimálními hodnotami vektorové kvantifikace a centroidy pro proměnné s optimálními škálovými hodnotami
  - graf centroidů s projekcí na každou proměnnou ze seznamu
- výstupní grafy
  - grafy objektových bodů a bodů kategorií
  - graf zátěží
  - transformace bodů (optimální kvantifikace kategorií proti indikátorům kategorií)
  - rezidua pro vybrané proměnné
  - objekty a proměnné
  - biplot
- maximální flexibilita v biplotu užitím dvou seznamů proměnných (jeden pro zátěže a jeden pro centroidy) u klíčového slova LOADINGS ve specifikačním příkaze PLOT
- přidání proměnných do pracovního datového souboru nebo zápis proměnných do externího datového souboru; transformované proměnné, skórování objektů a přibližná

hodnota pro proměnné, které nemají optimální škálovací hodnoty

- úplný výstup je tištěn i pro kategorie, které se vyskytují jen u dodatečných (suplementárních) objektů, speciálně centroidy, které klasifikují suplementární objekty do skupin

## OVERALS

Kanonická korelační analýza dvou nebo více skupin proměnných volitelnou metodou nejmenších čtverců

- možnost výběru metody pro určení výchozí konfigurace
- tvorba transformačního typu každé proměnné
  - násobné nominální
  - jednoduché nominální
  - ordinální
  - kardinální
- volba počtu skupin proměnných a určení proměnných v jednotlivých skupinách
- volba počtu dimenzí
- volba maximálního počtu iterací
- určení konvergenčních kritérií
- tisk výsledků
  - marginální četnosti analyzovaných proměnných
  - iterační historie
  - multiple fit
  - single fit
  - single loss per variable
  - pskóry pro kvantifikaci kategorií
  - projekce centroidů
  - centroidy
  - skóry pro jednotlivé objekty
  - kvantifikace kategorií
  - jednotlivé a násobné souřadnice
  - váhy
  - faktorové zátěže
- grafické výstupy
  - kvantifikace kategorií
  - komponenty kategorií
  - centroidy kategorií
- ukládání objektových skórů do proměnných v aktivním/ pracovním souboru

## PROXSCAL

Mnohorozměrné škálování proximitních dat o vzdálenostech a blízkostech objektů - hledání reprezentace objektů v rovině nebo prostoru s malým počtem dimenzí pomocí metody nejmenších čtverců.

- vstupy
  - matice dat
  - matice proximit: vstup jedné nebo několika čtvercových proximitních matic, buď symetrických

*nebo nesymetrických, nebo čtení proximit ve specifikovaném formátu*

- *čtení vah, iniciačních konfigurací, pevných souřadnic a nezávislých proměnných*
- *počáteční konfigurace: simplex, Torgerson, jednoduchý náhodný start, násobný náhodný start, vlastní konfigurace z vnějšího souboru*
- *proximity jako ordinální (nemetrická) nebo numerická (metrická) informace; ordinální transformace mohou zpracovávat spojené hodnoty jako diskrétní nebo spojité*
- *tři modely individuálních diferencí, model identity*
- *určení pevných souřadnic nebo nezávislých proměnných, které omezí konfiguraci*
- *určení transformací (numerické, nominální, ordinální a spline) pro nezávislé proměnné*
- *výstup zahrnuje vstupní a transformované proximity, historii iterací, váhy v individuálních prostorech,*

*vzdálenosti a dekompozici stresu*

## **PREFSCAL**

Preferenční škálování vizuálně vyšetřuje vztahy mezi proměnnými dvou množin objektů za účelem nalezení společné kvantitativní škály.

- *vstup jedné nebo více obdélníkových matic*
- *vstup vah, prvotní konfigurace a pevných souřadnic*
- *transformace proximit: lineární, ordinální, smooth ordinal, spliny*
- *specifikace mnohorozměrného rozprostření: identický, vážený euklidovský, nebo obecný euklidovský model*
- *specifikace pevných řádkových nebo sloupcových souřadnic z důvodu omezení konfigurace*
- *vstupní konfigurace: klasická trojúhelníková, klasická Spearmanova, Ross-Cliff, korespondence, centroidy, náhodný start, uživatelská konfigurace; iterační kriteria a penalizační parametry*
- *grafy pro násobné starty, počáteční společný prostor, stres na dimenzi, konečný společný prostor, váhy prostoru, individuální prostory, bodový graf kvality modelu, grafy reziduí, transformační grafy, Shepardovy grafy*
- *výstupy - vstupní data, násobné starty, počáteční společný prostor, iterační historie, míry kvality modelu, dekompozice stresu, konečný společný prostor, váhy prostoru, přiřazené vzdálenosti a transformované*
- *uložení společných souřadnic prostoru, individuálních vah, vzdáleností a transformovaných proximit do souboru*