

DAUGIAMAČIŲ SKALIŲ TAIKYMAS RINKOS TYRIMUOSE: PRIVALUMAI IR TRŪKUMAI

Doc. dr. Angelė Kėdaitienė
Europos humanitarinio universiteto
Politikos mokslų katedra
Valakampių g. 5, 10222 Vilnius
El. p. angele.kedaitiene@ucdconnect.ie

Doc. dr. Vytautas Kėdaitis
Vilniaus universiteto Ekonomikos fakulteto
Kiekybinių metodų ir modeliavimo katedra
Saulėtekio al. 9, 10222 Vilnius
Tel. (8 5) 236 6143
El. p. vytautas.kedaitis@ef.vu.lt

IVADAS

Kas yra daugiamačių skalių modelis (daugiamačis vertinimas, angl. *multidimensional scaling*, MDS), išduoda pats pavadinimas. Iš jo galima numanyti, jog tai – tikslingas objektų išdėstymas daugiamačiame erdvėje. Egzistuoja keli daugiamačių skalių (DS) apibrėžimai: daugiamačės skalės – tai matematinių procedūrų visuma, skirta atskleisti „nematomą duomenų struktūrą“ (Kruskal & Wish, 1978), daugiamačės skalės – tai duomenų analizės modelių rinkinys, leidžiantis pateikti atstumais išreikštų duomenų grafinį vaizdą (F. W. Youngas). Atliekant daugiamačių skalių analizę gaunamas grafinis vaizdas, kuriame objektai, pavyzdžiui, prekės, išdėstomi *n*-matėje erdvėje pagal jų panašumus. Panašumai tarp objektų vertinami pagal daugybę savybių. Sakoma, kad daugiamačių skalių modelis atskleidžia „nematomą duomenų struktūrą“. Grafiniame vaizde išryškėja savitos dimensijos, kurios dažnai nesutampa su pradinėmis savybėmis, pasirinktomis objektų panašumui vertinti.

Pagrindinis DS uždavinys – atvaizduoti objektus taip, kad tiriamoje erdvėje atstumai tarp jų būtų kuo mažesni. Kaip tuos objektus atvaizduoti tiriamoje erdvėje? Galima analizuoti metrinių ir nemetrinių artumų matricas. Be to, galima išskirti klasikinį modelį, kuriame sprendinys sudaromas naudojant matricų algebrą, o duomenys pateikiami metrinių artumų matrica. Vėliau buvo sukurti modernūs modeliai, leidžiantys analizuoti ne tik klasikinio modelio, bet ir kitokio pobūdžio artumų matricas. Prie tokių matricų būtų galima priskirti Pirsono koreliacijos koeficientą, euklidinį atstumą (nepanašumo atveju), Minkovskio (metriniais duomenimis) ir pan. Svarbiausi DS modeliai rinkos tyrimuose galėtų būti šie – klasikinis metrinis ir modernusis nemetrinis.

DS buvo sukurta ir pirmiausia taikoma psichometrikų. Netrukus jas imta taikyti psichologijoje ir sociologijoje, siekiant suprasti žmonių veiksmus ir tai, kaip jie vertina vienus ar kitus dalykus. Kadangi viena pagrindinių rinkodaros užduočių yra suprasti žmonių poreikius ir geriau juos tenkinti (siekiant pelno ar kitos naudos), DS susilaukė pasisekimo rinkos tyrimuose.

Galima išskirti daugybę DS panaudojimo rinkos tyrimuose sričių: naujų produktų ar idėjų generavimas, rinkos potencialo nustatymas, tinkamos reklamos kampanijos parinkimas, rinkos segmentavimas. DS gali būti naudojamos ten, kur reikia suprasti respondentų pasirinkimus ir jų motyvus, tai yra beveik visose rinkodaros ir rinkos tyrimų srityse. Dažniausiai naudojant DS siekiama atsakyti į vieną ar kelis iš šių klausimų: 1) Kokia pagrindinė produktų ar jų grupių suvokimo ir vertinimo esmė? 2) Kokios prekės laikomos panašiomis viena į kitą ir kodėl? 3) Kokie pagrindiniai vartotojų požiūriai? 4) Kokias naujų produktų galimybes palieka jau rinkoje egzistuojančios prekės? 5) Kaip išsidėstę vartotojų idealūs produktai tam tikrų savybių atžvilgiu? 6) Kaip įvairios reklamos priemonės veikia vartotojų suvokimą arba kiek reklamos kampanija atitinka prekės suvokimą?

DAUGIAMAČIŲ SKALIŲ ANALIZĖS ISTORIJA

DS buvo sukurtos ir ilgą laiką vystomos psichometrikų (psichometrija – mokslas apie matavimus psichologijoje), daugiausia prie to prisidėjo R. N. Shepardas („The analysis of proximities: multidimensional scaling with unknown distance function. I.“ *Psychometrika*, 1962) ir Josephas Bernardas Kruskalis („Nonmetric multidimensional scaling: a numerical method“, *Psychometrika* 29, 1964). Vis dėlto DS ištakos siekia XX a. ketvirtojo dešimtmečio pabaigą, kai buvo sukurti ir panaudoti pirmieji nedimensiniai modeliai. Šie modeliai nebuvo itin naudingi ir iki pat šeštojo XX a. dešimtmečio DS buvo primirštos. Paskutinais XX a. dešimtmečiais DS tobulinimą paskatino keli veiksniai, pagrindinis jų – naujų kompiuterių pasirodymas. Sukurti nauji modeliai ir algoritmai bei nauja kompiuterinė technika leido vis plačiau taikyti DS psichologijoje, sociologijoje, istorijoje ir kitose mokslo šakose.

Pirmą kartą, XX a. šeštojo dešimtmečio pabaigoje, DS rinkos tyrimuose pritaikė W. S. Torgersonas, sprendamas su vartotojų nuostatomis susijusią problemą, Naujosios Anglijos kompanijai, gaminančiai sidabro gaminius, pristatant naujus modelius. Kitas tyrėjas, ėmęs taikyti DS savo darbuose, buvo V. J. Stefflre'is. Jis pirmasis ėmė sistemškai taikyti DS

rinkos tyrimuose. Šis mokslininkas DS rezultatus naudojo kaip grafinę priemonę vartotojų supratimui apie prekių panašumus parodyti. Jo naudoti trimačiai modeliai, kuriuos jis vadino *tinkertoys* (vaikiškas erdvinis konstruktorius), pasirodė puiki priemonė suprantamai pateikti tyrimų rezultatus. V. J. Steffle'io darbuose buvo pabrėžiama viena pagrindinių DS savybių – sudėtingomis procedūromis apdorojus duomenis, išryškunami svarbiausi aspektai ir duomenis tampa lengviau analizuoti.

Aštuntajame dešimtmetyje prie DS tobulinimo ir jo panaudojimo rinkos tyrimuose prisidėjo P. E. Greenas su kolegomis. Jie pradėjo taikyti naujus DS metodus ir algoritmus. Tarp jų minėtini J. D. Carrollo ir J. J. Chango sukurti INDISCAL (angl. *Individual Differences Scaling* – individualių skalių skirtumai), IDIOSCAL (angl. *Individual Differences in Orientation Scaling* – individualių skalių skirtumų orientavimas), MDPREF (angl. *Multidimensional Analysis Of Preference Data* – daugiamatė pirmenybių duomenų analizė), PREFMAP (angl. *Preference Mapping* – pirmenybių skalės / žemėlapis). Jie pradėjo naudoti kitus ir DS metodus. P. E. Greenas su kolegomis daugiau dėmesio skyrė grafinių modelių dimensijoms euklidinėje erdvėje, jų interpretavimui ir supratimui bei pradėjo naudoti sanaujų analizės ir DS junginį. Šiame dešimtmetyje DS panaudojimą rinkos tyrimuose skatino ir atsiradusi kombinuotos erdvės analizė. Visi šie nauji metodai leido DS tiksliau reprezentuoti objektus ir gauti reikiamas grafinių modelių interpretacijas. INDISCAL atsiradimas leido tyrėjams operuoti didesniu dimensijų skaičiumi. Iki tol egzistavę metodai leido išskirti tik nedidelį dimensijų skaičių, todėl DS buvo nelabai tinkamos atliekant sudėtingus tyrimus. Nors grafinis modelis, turintis daugiau nei tris dimensijas, yra sunkiai suprantamas, INDISCAL leidžia sumažinti dimensijų skaičių. Šiuo metodu kelios dimensijos gali būti sujungiamos į poerdvius, taip pagrindiniame modelyje galima pavaizduoti gerokai daugiau dimensijų, neprarandant galimybės suprasti ir interpretuoti rezultatus.

Šie metodai tapo klasikiniais. DS ir toliau buvo tobulinamos. Daugiausia buvo dirbama kuriant DS metodus, kurie naudotų Euklido atstumus su parinktais svoriais. Pastaraisiais metais atsiranda vis didesnė praraja tarp DS metodo taikymo ir rinkos tyrimų. Dažnai vietoje DS pasirenkamos kitos analizės rūšys, pavyzdžiui, regresinė analizė. Pagrindinė to priežastis yra ta, jog tyrėjų nebetenkina DS rezultatų interpretavimas. Atlikus DS gautas grafinis modelis turi įvairių trūkumų. Pagrindinis jų yra dimensijų nepatikimumas. Jas ne tik sunku pavadinti, bet ir norint gauti tinkamą sprendimą priimti modelį ašys neretai turi būti pasu-

kamos, tačiau nėra aišku, kaip tai atlikti. Kita problema yra ta, kad net gavus tinkamą grafinį modelį jis gali būti nenaudingas. Pavyzdžiui, kuriant naują produktą galima atrasti tašką, reprezentuojantį šį produktą, tačiau tiksliai jį apibūdinti sunku. Dalis problemų tenka ir rinkos tyrėjams, kurie sunkiai perpranta DS analizės rezultatus.

DAUGIAMAČIŲ SKALIŲ PROCEDŪROS

DS metodo tikslas – netiesiogiai sumažinti dimensijų, kurias respondentai naudoja vertindami alternatyvas, skaičių. Netiesioginis būdas buvo pasirinktas dėl to, kad daugeliu atvejų požymiai nėra žinomi, taip pat respondentas gali nenorėti (aktualiausia kalbant apie asmeninius dalykus, pavyzdžiui, santykius su artimaisiais) arba negalėti (pavyzdžiui, žmogui, neturinčiam išsilavinimo meno srityje, būtų sunku nurodyti požymius, dėl kurių jam patinka vienas ar kitas kūrinys) tiksliai nurodyti savo vertinimo motyvų. DS analizė pradedama nuo duomenų rinkimo, įvykdžius šią procedūrą, gaunamas norimo dimensijų skaičiaus grafinis modelis.

Pirminiai duomenys

DS analizei reikalinga objektų tarpusavio panašumų matrica. Kai tik buvo sukurtos DS, dažniausiai įvesties duomenys buvo paremti panašumo įverčiais. Respondentai, remdamiesi pateikta skale, vertino tiesioginį objektų panašumą. Surinkti tokius duomenis buvo sudėtinga, nes norint įvertinti n objektų panašumus, reikėjo gauti $\frac{n(n-1)}{2}$ įverčius. Tarkime, norima įvertinti 10 objektų panašumus, tam respondentas turėtų pateikti $\frac{10(10-1)}{2} = 45$ įverčius. Visgi šis duomenų gavimo būdas turi ir pranašumų: respondentams dažniausiai nesunku įvertinti dviejų objektų tarpusavio panašumą, taip jie neatskleidžia savo vertinimo motyvų.

Duomenis, reikalingus DS pirminių duomenų matricai, galima surinkti įvairiais būdais. Tam galima naudoti apklausas, stebėjimą, grupines diskusijas ir daugybę kitų metodų. Naudojant šiuos metodus taip pat galima pasirinkti ir skirtingas priemones duomenims gauti, pavyzdžiui: prašyti respondentų atsakyti į struktūrizuotus klausimus ar tiesiogiai įvertinti objektų tarpusavio panašumus. Dažniausiai renkant duomenis DS pasirenkami šie būdai: kortelių rūšiavimas, objektų reitingavimas, netiesioginis panašumo įvertinimo gavimas.

Kortelių rūšiavimas. Šis metodas dažnai naudojamas asmeninių apklausų metu. Respondentams išdalijamos kortelės, kuriose surašytos visos vertinamų objektų poros. Tuomet respondentų prašoma sudėlioti korteles į kelias krūveles (pavyzdžiui, į keturias porų krūveles: labai panašių, panašių, nepanašių ir labai nepanašių objektų) pagal tai, kiek objektai kortelėse yra tarpusavyje panašūs. Tada respondentų prašoma kiekvienos krūvelės korteles išdėlioti pagal kortelėse esančių objektų porų tarpusavio panašumus. Tokiu būdu gaunamas objektų tarpusavio panašumų įvertinimas. Šis rūšiavimas atliekamas dviem etapais, nes kortelių skaičius dažniausiai yra gana didelis, pavyzdžiui, vertinant dešimt objektų reikia 45 kortelių. Respondentams yra gerokai lengviau jas rūšiuoti į kelias krūveles, o vėliau atskirai vertinti mažesnį objektų porų skaičių.

Objektų reitingavimas. Šis būdas nedaug skiriasi nuo pirmojo. Kartais neįmanoma ar nepraktiška respondentams dalyti korteles, pavyzdžiui, atliekant apklausas paštu ar internetu. Tuomet respondentų prašoma įvertinti, kiek objektų poros yra tarpusavyje panašios pagal pateiktą skalę. Skalė gali būti labai įvairi, priklausomai nuo to, kaip tiksliai reikia atskirti objektus, pavyzdžiui, šešių ar dešimties padalų skale nuo „labai panašūs“ iki „visiškai nepanašūs“. Taip pat renkant duomenis šiuo būdu galima prašyti respondentų įvertinti kiekvieno objekto panašumą su vienu objektu. Vėliau šis objektas pasirenkamas kaip atskaitos taškas ir pagal įvertintus panašumus gaunami objektų porų panašumai.

Netiesioginis panašumo įvertinimų gavimas. Dažnai tyrimuose objektų tarpusavio panašumai gaunami netiesiogiai, iš kitų vertinimų. Objektai vertinami atskirai vienas nuo kito, pagal pasirinktas savybes. Pagal šiuos vertinimus randamas ir pasirenkamas panašumo matas. Tam naudojami įvairūs panašumą vertinantys rodikliai, pvz., koreliacijos koeficientai, atstumo matai, tikimybiniai panašumo koeficientai ir t. t.

Dažnai gautų rezultatų patikimumas priklauso nuo pasirinktų objektų tipo, jų informatyvumo. Vertinant kino filmų panašumus gali išaiškėti veiksmo kiekio (daug veiksmo – mažai veiksmo) ar garso takelio kokybės (gera kokybė – bloga kokybė) dimensijos, tačiau sunku tikėtis, jog tyrimas išryškins pavojaus susižeisti (didelis pavojus – mažas pavojus) ar poveikio sveikatai (teigiamas poveikis – neigiamas poveikis) dimensijas. Vertinant objektų panašumus, dažniausiai remiamasi vienomis savybėmis ir mažai dėmesio kreipiama į kitas. Pasirinkti objektai yra svarbūs ne tik numatant, kokios savybės gali išryškėti atlikus DS analizę. Grafiniame modelyje išryškėjančios dimensijos tiesiogiai priklauso nuo objektų struktūros (pavyzdžiui,

gaiviųjų gėrimų). Jei tarp DS analizei pasirinktų objektų neįtrauksime svarbių objektų (pavyzdžiui, atlikdami gaiviųjų gėrimų analizę neįtrauksime energetinių gėrimų), greičiausiai negausime vienos ar kelių svarbių dimensijų (pavyzdžiui, tonizuojančio poveikio). Kita vertus, negalima ir persistengti renkantis objektus DS analizei. Jei bus pasirinkti itin skirtingi objektai ar jų grupės (pavyzdžiui, atliekant gėrimų DS analizę bus pasirinkti alkoholiniai, karštieji ir gaivieji, nealkoholiniai gėrimai), greičiausiai rezultatas pateiks kelias objektų grupes (šio pavyzdžio atveju – alkoholinių, karštųjų ir gaiviųjų nealkoholinių gėrimų grupes). Toks DS rezultatas nepateiks naujų duomenų apie tiriamus objektus ar reiškinius ir bus bevertis. Todėl svarbu atidžiai rinkti objektus, kurie bus naudojami DS analizei atlikti.

Kaip ir patys objektai (jų savybės), tinkamam DS rezultatui gauti svarbus ir jų skaičius. Paprastai objektų, reikalingų tinkamam DS rezultatui gauti, skaičius 2–3 kartus viršija gaunamų dimensijų skaičių. Kaip ir bet kokiame kitame tyrime, naudojant DS negalima tikėtis gerų rezultatų, jei duomenų kokybė yra prasta, t. y. duomenų patikimumas nėra patikrintas papildomu tyrimu arba juose yra netipinių, tiriamajai visumai nebūdingų duomenų. Renkant duomenis DS, svarbu atsižvelgti į respondentų galimybes ir norą rinkti patikimą informaciją. Didėjant pasirinktų objektų skaičiui, objektų porų, kurių panašumus reikia įvertinti, skaičius sparčiai didėja. Pavyzdžiui, norint gauti dvidešimties objektų tarpusavio panašumų įvertinimus, reikės gauti net 190 objektų porų panašumo įvertinimus. Šiais laikais, kai ryškėja žmonių nenoras dalyvauti apklausose, būtų itin sunku rasti pakankamą skaičių respondentų, kurie norėtų atsakyti į tokį klausimų skaičių. Net ir radus jų pakankamai, būtų sunku išlaikyti jų dėmesį visos apklausos metu. Šią problemą galima lengvai išspręsti, jei respondentai yra homogeniški. Tuomet galima suskirstyti klausimą ir respondentų imtį į kelias dalis. Tokiu atveju respondentai turėtų įvertinti tik dalies objektų porų panašumus.

Daugiamačių skalių (grafinio modelio) gavimas

Standartiniai DS modeliai pateikia grafinę objektų reprezentaciją vektorinėje erdvėje. Toks modelis, kuriame yra įvertinti objektų tarpusavio panašumai, paprastai atliekamas dviem etapais.

Pirminis sprendimas (objektų išdėstymas). Pirmasis DS etapas yra pirminio sprendimo gavimas arba pirminis objektų išdėstymas erdvėje. Vienas būdų tai atlikti yra atsitiktine tvarka išdėstyti objektus norimo dimensijų skaičiaus erdvėje. Nors toks metodas yra paprastas, tačiau dažnai jis būna neefektyvus ir gerokai apsunkina kitą DS etapą, todėl yra retai naudojamas.

Dažniausiai pirminiam objektų išdėstymui gauti naudojamas trianguliacijos metodas. Pirmiausia tyrėjo nuožiūra yra pasirenkami du objektai. Šie objektai pavaizduojami erdvėje kaip taškai. Atstumas tarp šių taškų erdvėje turi atitikti pagal pasirinktą panašumo (derėtų kalbėti ne apie panašumo, o apie nepanašumo matus, nes kuo didesnis atstumas tarp taškų, tuo objektai mažiau panašūs tarpusavyje) matą nustatytą atstumą. Toliau iš duomenų parenkamas trečias objektas, randami atstumai tarp šio objekto ir jau pavaizduotų erdvėje objektų. Anksčiau atidėti taškai pasirenkami kaip apskritimų centrai ir iš šių centrų brėžiami apskritimai, kurių spinduliai atitinka atstumus iki norimo atidėti taško. Jei nubrėžti apskritimai liečia vienas kitą, jų lietimosi vietoje atidedamas taškas, vaizduojantis trečiąjį objektą. Po šio žingsnio erdvėje būna atidėti trys taškai. Analogiška procedūra kartojama su trimis apskritimais. Jei duomenų matricos duomenys yra nepriekaištingi, tai visi trys apskritimai susikirs viename taške. Šis sankirtos taškas bus pasirinktas ketvirtajam objektui atvaizduoti. Tačiau tai ne visada pavyksta – apskritimai nesusikerta viename taške, o jų sankirta sudaro kreivinį trikampį. Tokiu atveju šio trikampio centras pasirenkamas ketvirtajam objektui atvaizduoti. Analogiška procedūra kartojama kaskart pridėdant po vieną tašką, kol galiausiai erdvėje pavaizduojami visi objektai.

Čia pirminio taškų išdėstymo radimas aprašytas dvimatėje erdvėje. N -matėje erdvėje (kai $n > 2$) procedūra nedaug skiriasi nuo aprašytosios. Pirmieji taškai dedami taip pat, kaip ir dvimatės erdvės atveju. Trečiasis taškas randamas brėžiant ne apskritimus, o sferas (trimatėje erdvėje) arba hipersferas (n -matėje erdvėje, kai $n > 3$). Šis taškas pasirenkamas imant bet kurį tašką iš sankirtos sferos ar hiperplokštumos. Jei sankirtos taškų nėra, ketvirtasis ir visi likę taškai atidedami sferų arba hipersferų sankirtos sudarytos erdvės centre.

Nesvarbu, kelių dimensijų erdvėje atidedami taškai, pirminio taškų išdėstymo procedūra yra beveik identiška. Tačiau jei dimensijų yra trys, atidėti taškus nenaudojant kompiuterinių programų paketų tampa sudėtinga, o jei dimensijų yra keturios ir daugiau – labai sunku. Todėl jei pasirinktos erdvės dimensijų yra daugiau nei keturios, vienintelis būdas išdėstyti taškus yra naudoti specializuotas kompiuterines programas.

Antrasis žingsnis. Antrajame DS analizės etape siekiama optimizuoti pradinį taškų išdėstymą. Tai atliekama naudojant gradiento procedūrą (gradientas – greičiausio funkcijos augimo kryptį nurodantis vektorius), kartais šis būdas dar vadinamas „kopimo į

kalną“ būdu (funkcija maksimizuojama). Kai funkcija minimizuojama, tai reiškia, kad leidžiamės į daubą priešinga gradientui kryptimi. Gradientinis optimizavimo metodas yra bendra optimizacijos procedūra. Nesvarbu, kaip šis metodas būtų vadinamas, jo esmė – kryptingas taškų išdėstymo keitimas, siekiant optimalaus jų išdėstymo pasirinkto dimensijų skaičiaus erdvėje.

Dažniausiai taškų išdėstymas erdvėje optimizuojamas, skaičiuojant atstumus tarp atidėtų taškų ir lyginant juos su pradiniais duomenimis. Jei atstumai duomenų matricoje nesutampa su apskaičiuotais atstumais, reikia apskaičiuoti paklaidas ir koreguoti taškų koordinates, siekiant šias paklaidas minimizuoti. Siekiant patikrinti, kokią naudą davė koordinatinių pakeitimais, naudojamas paklaidos indeksas, dar vadinamas STRESS funkcija. Skirtingiems DS modeliams yra sukurta daugybė skirtingų STRESS funkcijų, tačiau viena populiariausių yra Džozefo Bernardo Kruskalio STRESS₁ funkcija (Kruskal, Joseph B. „Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis“, *Psychometrika*, 1964-03-29). Pažymėkime atstumą tarp objektų i ir j , gautų iš duomenų matricos d_{ij} , atstumą tarp taškų, reprezentuojančių objektus i ir j , atidėtų n dydžio dimensijos erdvėje \hat{d}_{ij} . Tuomet Kruskalio STRESS₁:

$$\sqrt{\frac{\sum (d_{ij} - \hat{d}_{ij})^2}{\sum d_{ij}^2}} = \sqrt{\frac{\sum \text{paklaidos}^2}{\sum \text{atstumai}^2}}$$

Pagal šio indekso reikšmes galima nustatyti, ar koordinatinių pakeitimais išdėlijo taškus erdvėje taip, kad jie geriau reprezentuotų objektus, ar pablogino rezultatą (kuo mažesnė STRESS₁ reikšmė, tuo tiksliau atidėti taškai reprezentuoja atstumus tarp objektų). Parinktų taškų koordinatės keičiamos tol, kol STRESS₁ reikšmė yra didesnė nei nustatyta tyrėjo arba kol pasiekiamas didžiausias nustatytas iteracijų skaičius (jį nustato tyrėjas arba kompiuterinė programa). Neretai sunku nuspręsti, kokią didžiausią leistiną šio indekso reikšmę pasirinkti. Tai priklauso nuo to, kiek tikslūs turi būti tyrimo rezultatai. Tyrėjų nuomonės apie tai, kokia funkcijos reikšmė leidžia gautą grafinį sprendimą laikyti tikslu, skiriasi. D. B. Kruskalis šios funkcijos taikymo pradžioje pasiūlė modelio tikslumą vertinti taip: STRESS₁ reikšmė 1) iki 0,05 – puikus, 2) 0,05–0,1 – geras, 3) 0,1–0,2 – patenkinamas, 4) 0,2 ir daugiau – blogas. Tai vienas priimtinausių vertinimų, kuriuo derėtų naudotis, jei tyrimo specifika nereikalauja ko kito. Atliekant DS analizę, reikia pasirinkti, kiek dimensijų turės ieškomas grafinis modelis. Dimensijų skaičiaus nustatymas yra sudėtinga užduotis. Yra daugybė teorijų, apibūdinančių reikiamą dimensijų skaičių ir būdus jį nustatyti.

Vis dėlto dažniausiai tyrėjai jomis nesinaudoja, o pasirenka praktinį, bandymais pagrįstą kelią. Šiuolaikiniai kompiuterinių programų paketai leidžia atlikti DS analizę labai greitai, todėl ji atliekama pasirinkus skirtingus dimensijų skaičius ir iš rezultatų išrenkant tinkamiausius. Tinkamam dimensijų skaičiui nustatyti naudojami du būdai. Pirmasis – iš anksto nustatyti paklaidos indekso (pavyzdžiui, STRESS) reikšmę. Tyrėjas iš anksto pasirenka indekso reikšmę, kurios negalima viršyti, ir pagal bandymų rezultatus išsirenka dimensijų skaičių, kuriam esant ši reikšmė neviršijama. Kitas būdas yra nubraižyti grafiką, kurio skirtingose ašyse atidedamos STRESS funkcijos reikšmės ir dimensijų skaičius. Tuomet grafike ieškoma taško, kuriame pridėjus papildomą dimensiją STRESS reikšmė pradeda po truputį mažėti. Pirmame paveiksle pateikiamas tokio grafiko pavyzdys. Šiame paveiksle matome, kad kreivės nuolydis stipriai sumažėja, kai dimensijų yra dvi, o kai jų yra trys, kreivė tampa beveik horizontali. Iš šio grafiko negalima tiksliai pasakyti, dvi ar tris dimensijas pasirinkti. Tai priklauso nuo tyrėjo nuomonės – to, ką jis laiko nereikšmingu STRESS funkcijos pakitimu, ir tyrimo specifikos. Šiuo atveju greičiausiai būtų pasirinkta atlikti DS trimatėje erdvėje, nes pridėjus ketvirtąją dimensiją, grafinio modelio tikslumas pakistų nedaug, o interpretuoti keturmatį modelį yra gerokai sudėtingiau nei trimatį. Vieta, kurioje STRESS funkcijos reikšmė, pridėjus papildomą dimensiją, pradeda mažėti nereikšmingu laipsniu, vadinama „alkūne“.

Analinis daugiamačių skalių vertinimas

Atlikus DS analizę, gaunamas objektų išdėstymas pagal tam tikras savybes. Rezultatai pateikiami norimo dimensijų skaičiaus erdvėje, dažniausiai dviejų arba trijų, nes tokius grafikus lengva interpretuoti. Nors gautasis modelis leidžia susidaryti vaizdą, kaip objektai išsidėstę dimensijose, dimensijų pavadinimai nėra žinomi. Parinkti tinkamus pavadinimus dimensijoms – sudėtingas uždavinys, nes požymių, pagal kuriuos išdėstomi objektai, dažniausiai yra daugiau nei gautų dimensijų. Viena dimensija gautame grafiniame modelyje gali apimti kelias savybes ar tam tikrą jų derinį, todėl dimensijoms pavadinti reikia patirties ir išradingumo. Paprastesniuose DS rezultatuose dimensijų pavadinimus galima nustatyti ieškant bendrų bruožų ir skirtumų tarp kraštinių objektų kiekvienoje dimensijoje. Norint tinkamai pavadinti gautas dimensijas, galima surinkti įvairių savybių, kurias turi skalių objektai, panašumų vertinimus. Tam tikras savybes reprezentuojantys taš-

kai paprastai „parodo“, kurioms dimensijoms šios savybės yra būdingos, tuo pat metu tai palengvina ir pačių dimensijų supratimą. Kitas būdas, leidžiantis patikslinti dimensijų pavadinimus, susijęs su papildomos informacijos analizavimu. Tyrimo metu galima surinkti objektų vertinimus pagal tam tikras savybes. Vėliau reikėtų ieškoti koreliacijos tarp šių įvertinimų ir objektų padėties gautame grafiniame modelyje. Jei randamas tvirtas ryšys tarp objektų pozicijų ir jų įvertinimų, šis ryšys tyrėjui yra užuomina, suteikianti galimybę tiksliau pavadinti dimensijas.

Vis dėlto nederėtų pamiršti, jog gauti rezultatai dažnai yra preliminarūs. Patartina DS analizę naudoti sudarant hipotezes, o ne kaip galutinį rinkos ar problemos sprendimo modelį. Gavus grafinį modelį, reikia pasirinkti naują imtį ir patikrinti rezultatus kitu metodu. Tik gavus tokį patį rezultatą, juo galima pasitikėti.

DAUGIAMAČIŲ SKALIŲ TAIKymo PRAKTIKOJE YPATUMAI

Siekiant aiškiau apibūdinti DS, pateikiamas praktinis pavyzdys. Jo tikslas – padėti suprasti DS procedūras, remiantis konkrečiu pavyzdžiu.

Duomenų rinkimas ir panašumų matricos gavimas

Šiam DS pavyzdžiui buvo pasirinkti devyni objektai. Siekiama įvertinti uždarnosios akcinės bendrovės „Švyturys-Utenos alus“ gaminamų alaus rūšių panašumą. Tyrimui buvo pasirinktos devynios „Švyturio“ alaus daryklos gaminamos alaus rūšys: „Ekstra“, „Ekstra Draught“, „Švyturio“, „Gintarinis“, „Originalusis“, „Stipriausias“, „Baltijos“, „Baltas“ ir „Degintas“. Tyrimo objektų pasirinkimą lėmė tyrėjo asmeninė nuomonė ir žinios apie šiuos produktus. Norint tinkamai atlikti DS analizę, tyrėjui reikia nemažai žinių apie pasirinktus objektus, kitaip būtų sunku įvertinti ir interpretuoti gautą modelį.

Duomenims rinkti buvo pasirinktas tiesioginis objektų porų panašumo vertinimas. Toks duomenų rinkimo būdas pasirinktas dėl jo paprastumo ir operatyvumo. Ne visos alaus rūšys respondentams buvo vienodai pažįstamos, todėl tiesioginis objektų porų panašumų vertinimas leido respondentams juos vertinti pagal tas savybes, kurios jiems yra žinomos. Norint įvertinti devynių objektų tarpusavio panašumus iš viso reikėjo įvertinti 36 objektų poras.

Duomenims apie objektų porų panašumus gauti buvo parengta anketa. Anketą sudarė 36 standartizuoti skalės tipo klausimai. Respondentų buvo prašoma

įvertinti 36 objektų porų panašumus 10 padalų skalėje, nuo 1 – visiškai nepanašios, iki 10 – identiškos. Pirmiausia apklausa buvo atliekama internetu. Anketa buvo paskelbta apklausų portale [apklausa.lt](http://www.apklausa.lt/answerform.php?form=10639) (<http://www.apklausa.lt/answerform.php?form=10639>). Toks duomenų rinkimo būdas pasirodė visiškai neefektyvus. Buvo gautos tik septynios anketos, iš jų trys užpildytos iki galo, keturios apklausos nutrūko neužpildžius anketos iki galo. Pasireiškė vienas pagrindinių tiesioginio objektų panašumo vertinimo trūkumų – dėl didelio klausimų skaičiaus respondentai prarado susidomėjimą apklausa. Dėl mažo internetinės apklausos efektyvumo, šio duomenų rinkimo būdo teko atsisakyti.

Nepavykus surinkti duomenų internetine apklausa, teko rinktis kitą duomenų gavimo būdą – dalijant anketas. Iš viso buvo išdalyta 30 anketų. Toks duomenų rinkimo būdas pasirodė gerokai efektyvesnis, sugrįžo visos anketos. Vis dėlto net aštuonias iš sugrįžusių anketų teko atmesti. Pasireiškė nesąmoningos respondentų klaidos dėl išsiblašymo ir nuovargio. Tai pasireiškė tuo, jog respondentai, įpusėję apklausą, toliau rinkosi vienodus panašumus visiems likusiems objektams. Tokiais atsakymais nebuvo galima pasitikėti ir teko

juos atmesti, todėl galiausiai naudoti tik 22 respondentų pateikti duomenys.

Jau pirmoje tyrimo dalyje išryškėjo pagrindinės problemos, su kuriomis susiduriama renkant duomenis DS analizei. Dažnai apklausos duomenys gali būti gana nepatikimi. Darbe aprašomi keli būdai, kaip galima sumažinti duomenų nepatikimumą. Vis dėlto svarbiausia yra tinkamai motyvuoti respondentus, tam galima nurodyti tyrimo reikšmingumą ar naudoti įvairius paskatinimus. Taip pat derėtų rinktis asmeninę apklausą, taip apklausėjas galėtų geriau motyvuoti ir skatinti respondentus.

Surinkus duomenis iš 22 respondentų reikia nuspręsti, kaip apibendrinti jų pateiktus duomenis ir sudaryti atstumų matricą. Kadangi surinkti duomenys yra eilės skalės tipo (ranginiai), reikia rinktis vieną iš struktūrinių vidurkių. Iš jų buvo pasirinkta mediana, taip sumažinant kraštutinių reikšmių įtaką. Gaunama surinktų duomenų matrica (1 lent.).

Gautoje duomenų matricoje yra pateikiami ne atstumai tarp duomenų, o jų panašumai. DS analizei atlikti dažniausiai naudojami nepanašumo matai – atstumai. Paversti duomenis atstumais nėra sudėtinga, tačiau visai nebūtina, nes daugelis statistinių programų paketų gali dirbti tiek su panašumo, tiek su nepanašumo matais.

1 lentelė. Duomenų matrica

	„Ekstra“	„Ekstra Draught“	„Švyturio“	„Gintarinis“	„Originalusis“	„Stipriausias“	„Baltijos“	„Baltas“	„Degintas“
Ekstra	10	7	7	5	5	3	4	3	3
Ekstra Draught		10	6	5	4	2	4	4	3
Švyturio			10	8	7	4	5	3,5	3
Gintarinis				10	7	4	5	4	3
Originalusis					10	3	5	3	3
Stipriausias						10	4	2	3,5
Baltijos							10	4	4
Baltas								10	2
Degintas									10

Grafinio modelio gavimas

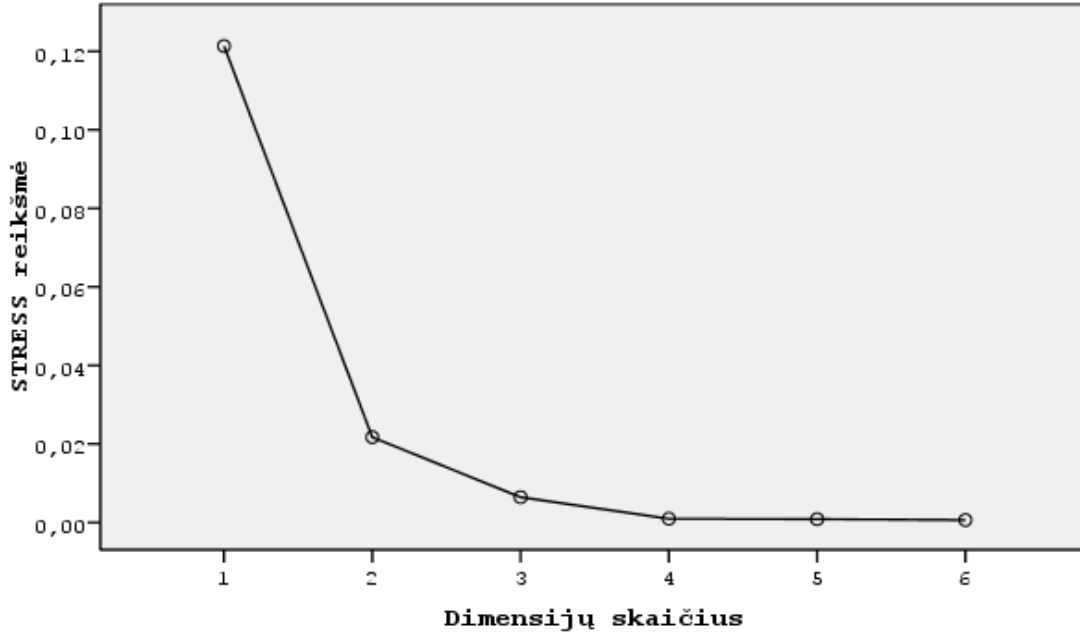
Darbe aprašytas grafinio modelio gavimas padeda suvokti, kaip tai atliekama, tačiau dažniausiai šį darbą atlieka statistinės programos. Turint trikampę duomenų matricą galima pereiti prie grafinio modelio gavimo. Šiame darbe grafinis modelis gaunamas naudojant statistinių programų paketą SPSS 15. Iš jo

DS analizei buvo pasirinkta PROXSCAL (Proximity Scaling) procedūra. Ši procedūra skirta atlikti DS analizei, minimizuojant STRESS reikšmę. Programoje nurodoma, jog duomenys yra panašumai, išdėstyti trikampės duomenų matricos viršutinėje dalyje. Taip pat nurodoma, jog reikia pateikti duomenis nuo vienos iki šešių dimensijų. Programa pateikia STRESS reikšmės mažėjimo grafiką (1 pav.).

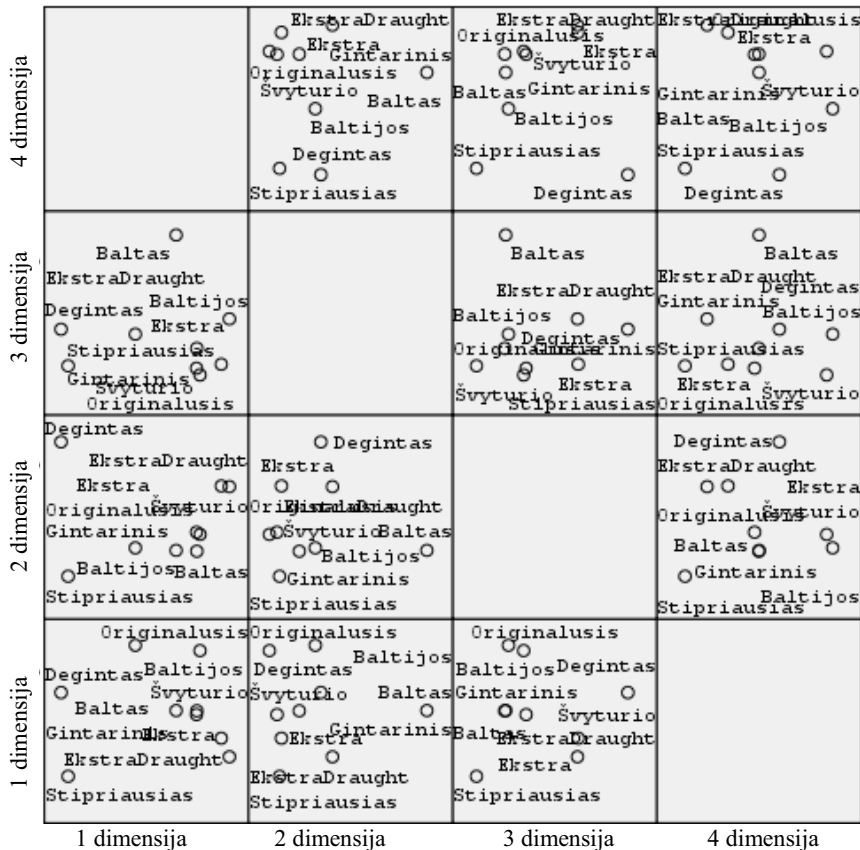
Grafike ieškoma „alkūnės“, pagal kurią pasirinkamas dimensijų skaičius. Iš grafiko matyti, kad esant dviem dimensijoms, STRESS reikšmė labai stipriai, ji taip pat mažėja, kai dimensijų skaičius padidėja iki trijų ir keturių. Esant daugiau nei keturioms dimensijoms, STRESS reikšmės mažėjimo beveik

nematome. „Alkūnės“ radimas tyrėjui, kuris neturi daug patirties, yra pats paprasčiausias ir efektyviausias būdas pasirinkti dimensijų skaičių. Pagal šį metodą, tikslinga būtų pasirinkti keturias dimensijas. Vis dėlto tokį modelį (2 pav.) būtų itin sunku interpretuoti ir analizuoti, ypač jei tyrėjui trūksta patirties.

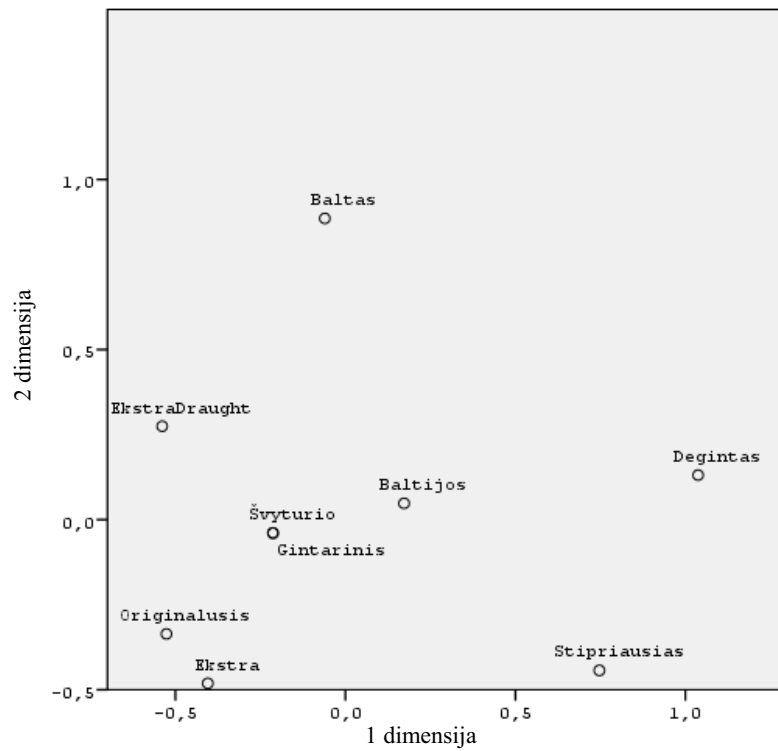
1 pav. STRESS funkcijos reikšmės mažėjimas



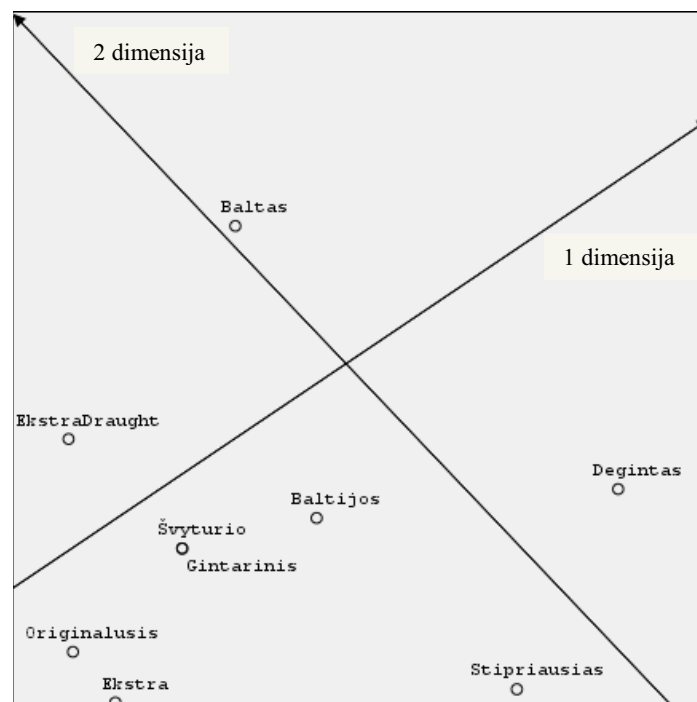
2 pav. Keturių dimensijų modelis



3 pav. Dviejų dimensijų modelis



4 pav. Dviejų dimensijų modelis su paslinktomis ir pasuktomis ašimis



Todėl interpretacijai pasirenkamas dviejų dimensijų modelis (3 pav.). Tokį modelį lengviau interpretuoti, nors jis yra mažiau tikslus.

Daugiamačių skalių analizės rezultatai ir jos interpretavimas

Gavus grafinį modelį, reikia jį išanalizuoti ir interpretuoti. Ši tyrimo dalis yra labiau kūrybinė nei matematinė. Interpretacijai pasirinktas dviejų dimensijų modelis (3 pav.). Pirmiausia reikia pavadinti dimensijas. Šiame pavyzdyje dimensijos gali būti įvairios: kaina, prestižas, etiketės spalva, skonis (saldus – kartus), stiprumas, populiarumas. Kaip pavadinti dimensijas turi nuspręsti tyrėjas. Vienas paprasčiausių būdų yra palyginti kraštutinius objektus. Šiuo atveju pirmoje dimensijoje kraštutiniai objektai yra „Ekstra Draught“ ir „Degintas“. Šios alaus rūšys labiausiai skiriasi spalva ir stiprumu, todėl pirmoji dimensija gali būti spalva (šviesus – tamsus) arba stiprumas (lengvas – stiprus). Antroje dimensijoje kraštutiniai objektai yra „Ekstra“ ir „Baltas“. Skirtumą tarp šių alaus rūšių atrasti yra sudėtinga, tačiau labiausiai jie skiriasi populiarumu, todėl antroji dimensija yra populiarumas (populiarus – nepopuliarus). Taigi pasirenkamos dimensijos: pirmoji yra spalva, antroji – populiarumas. Toks įvertinimas yra subjektyvus ir tik nuo tyrėjo priklauso, kiek jis yra tikslus.

Vis dėlto dimensijos nėra unikalios, DS analizėje tik atstumai tarp objektų yra unikalūs ir nekintantys. Todėl pagal tyrėjo nuomonę galima naudoti ir modelio veidrodinį atspindį ar apverstą modelį. Taip pat interpretuojant modelį galima paslinkti ar pasukti ašis. Tai atlikus, atstumai tarp objektų nepakis, tačiau dimensijas bus galima interpretuoti kitaip. Modelio su paslinktomis ir pasuktomis ašimis pavyzdys yra pateikiamas 4 pav.

Taip pasukus ašis, pirmoje dimensijoje kraštutiniaisiais objektais tampa „Originalusis“ ir „Degintas“, o antroje – „Baltas“ ir „Stipriausias“. Pagal šių alaus rūšių skirtumus, pirmąją dimensiją galima pavadinti *prestižas* (įprastas – išskirtinis), o antrąją – *stiprumas* (stiprus – silpnas). Pasukus ar paslinkus ašis modelis lieka tikslus, tačiau jo interpretacija gali visiškai skirtis. Tokia DS analizės savybė rodo jos pagrindinį trūkumą – rezultatų nevienareikšmiškumą. Modelio naudingumas ir tikslumas priklauso nuo subjektyvaus tyrėjo vertinimo.

Iš gauto DS grafinio modelio galima matyti objektų sancaupas. Jos nekinta slenkant ar sukant ašis, todėl yra vienareikšmės. Iš šio pavyzdžio matyti, jog vieną sancaupą sudaro „Švyturio“, „Baltijos“ ir „Gintarinis“ alus, o kitą – „Originalusis“ ir „Ekstra“.

Dvi alaus rūšys, „Gintarinis“ ir „Švyturio“, yra atvaizduotos viename taške, vartotojai nemato tarp jų jokio skirtumo, todėl vieną šių rūšių galima panaikinti ir sumažinti tiek produkcijos, tiek reklamos išlaidas.

Pateiktas DS analizės pavyzdys leidžia pamatyti, kaip šis metodas veikia praktikoje. Taip pat išryškunami pagrindiniai DS trūkumai. Dažnai atlikus DS analizę gautas rezultatas yra nevienareikšmis ir subjektyvus. Tyrėjui reikia nemažai pastangų ir kūrybiškumo, kad DS rezultatai būtų naudingi.

DS yra svarbi priemonė, plačiai taikoma įvairiose mokslo srityse. Šis straipsnis turi padėti susidaryti bendrą DS raidos vaizdą. Psichometrikų sukurtos DS yra plačiai naudojamos įvairiose mokslo srityse, nes objektų tarpusavio panašumų išsidėstymas savybių erdvėje yra vaizdinga priemonė, leidžianti susidaryti bendrą problemos vaizdą. Apžvelgtos DS vystymosi tendencijos jos atsiradimo pradžioje leidžia geriau suprasti dabartines DS analizės galimybes ir perspektyvas.

Atlikti DS analizę nėra sudėtinga, šiais laikais tai daro įvairūs kompiuterinių programų paketai. Vis dėlto egzistuoja daugybė subtilybių atliekant pačią DS procedūrą, o padaryti teisingas išvadas taip pat nėra lengva. Tik pasirinkus tinkamus duomenų rinkimo būdus ir tinkamą DS analizės modelį galima gauti konkretaus uždavinio sprendimui tinkamą grafinį modelį. Gautas grafinis modelis gali būti pakankamas, tačiau jo panaudojimas konkrečiam uždaviniui spręsti priklauso nuo tyrėjo nuovokos ir supratimo. Derėtų nepamiršti, jog DS analizę geriausia naudoti ne kaip galutinę sprendimo paieškos priemonę, o kaip priemonę hipotezėms formuluoti.

IŠVADOS

Viena svarbiausių DS panaudojimo sričių yra rinkos tyrimai. Rinkodaros specialistams labai svarbu žinoti, kaip vartotojai vertina įvairius produktus, kokie yra vartotojų poreikiai ir jų pirmenybės. Visa tai padeda atskleisti DS analizę, todėl ji plačiai naudojama rinkos tyrimuose (rinkos potencialui nustatyti, tinkamai reklamos kampanijai parinkti, rinkai segmentuoti ir t. t.). Šiuo metu daugiausia dėmesio kreipiama į atliekant tęstines apklausas ir stebint gautų duomenų daugiamatę analizę ir DS panaudojimą kuriant naujus produktus bei konkurencingą jų rinkodaros strategiją. Tai itin perspektyvios sritys, tinkamas jų supratimas gali padėti rinkodaros specialistams siekti tikslų. Nepaisant DS naudingumo rinkos tyrimams, jo panaudojimas šioje srityje mažėja. Jį išstumia kitos analizės rūšys, o naujų metodų praktinio pritaikymo galimybių stygius jaučiamas vis labiau.



LITERATŪRA

1. Bernatavičienė, J. 2004: *Daugiamačių duomenų vizualizavimo Sammono algoritmu tyrimas* (magistro diplominis darbas). Vilnius: Vilniaus pedagoginio universiteto Matematikos ir informatikos fakulteto Informatikos katedra, p. 17–18.
2. Carroll, J. D., Green, P. E. 1997: Psychometric methods in marketing research: part II, multidimensional scaling, *Journal of marketing research*, Nr. XXXIV, p. 193–204.
3. Green, P. E. 1975: Marketing applications of MDS: assessment and outlook, *Journal of marketing*, 1975, Nr. 39, p. 24–31.
4. Green, P. E. 1975: On the robustness of multidimensional scaling techniques, *Journal of marketing research*, Nr. XII, p. 73–81.
5. Kėdaitis, V. 2009: *Matavimai ir duomenų rinkimo metodai*, Vilnius: Vilniaus universitetas.
6. Kėdaitis, V. 2005: *Sankaupų analizė (paskaitų konspektas)*, Vilnius: Vilniaus universitetas.
7. Lehmann, D., Gupta, S., Steckel, J. 1998: *Marketing Research*, Massachusetts: Addison-Wesley Publishing company.
8. Young, Forrest W. *Multidimensional scaling*. Interneto prieiga: <http://forrest.psych.unc.edu/teaching/p208a/mds/mds.html> (žiūrėta 2009-11-24).