

ՀՀ ԿՐԹՈՒԹՅԱՆ ԵՎ ԳԻՏՈՒԹՅԱՆ ՆԱԽԱՐԱՐՈՒԹՅՈՒՆ
ԵՎՐՈՊԱԿԱՆ ԿՐԹԱԿԱՆ ՏԱՐԱԾԱՇՐՋԱՆԱՅԻՆ ԱԿԱԴԵՄԻԱ

ՀՐԱՆՏ ԱՇՈՏԻ ԴԱՎԹՅԱՆ

**ՄԵՃ ՏՎՅԱԼ ՆԵՐԻ ՏՆՏԵՍԱԳԻՏԱԿԱՆ ՎԵՐԼՈՒՅՈՒԹՅՈՒՆԸ՝ ԳՈՐԾԱՐԱՐ
ՈՐՈՇՈՒՄՆԵՐԻ ԿԱՅԱՑՄԱՆ ՀԻՄՆԱԽԴԻՐ**

ԱՏԵՆԱԽՈՍՈՒԹՅՈՒՆ

Ը.00.02 «Տնտեսություն, նրա ոլորտների տնտեսագիտություն և կառավարում» մասնագիտությամբ տնտեսագիտության թեկնածուի գիտական աստիճանի հայցման համար

Գիտական ղեկավար՝ տնտեսագիտության դոկտոր,
պրոֆեսոր ՀԱՐՈՒԹՅՈՒՆ ԱԼԵՔՍԱՆԻ ՄԱՐԶՊԱՆՅԱՆ

ԲՈՎԱՆԴԱԿՈՒ ԹՅՈՒՆ

ՆԵՐԱԾՈՒ ԹՅՈՒՆ	3
ԳԼՈՒԽ 1. ՄԵԾ ՏՎՅԱԼ ՆԵՐԻ ԲՈՎԱՆԴԱԿՈՒ ԹՅՈՒՆԸ ԵՎ ԿԻՐԱՌՄԱՆ ԱՆՅՐԱԺԵԾՏՈՒ ԹՅՈՒՆԸ ՏՆՏԵՍԱԳԻՏՈՒ ԹՅԱՆ ԵՎ ԿԱՌԱՎԱՐՄԱՆ ՈԼՈՐՏՈՒՄ	9
1.1 Մեծ տվյալ ներքն՝ որպես տնտեսագիտական հետազոտության առարկա.....	9
1.2 Մեծ տվյալ ներքի տնտեսագիտական վերլուծության գործնական կիրառությանը՝ միջազգային փորձի հաշվառմամբ	24
1.3 Ոլորտի հիմնական «խաղացողները» և զարգացման միտումները	42
ԳԼՈՒԽ 2. ՏՎՅԱԼ ՆԵՐԻ ԲԱԶԱՅԻ ՏՆՏԵՍԱԳԻՏԱԿԱՆ ՎԵՐԼՈՒԾՈՒ ԹՅԱՆ ՄԵԹՈԴԱԿԱՆ ՄՈՏԵՑՈՒ ՄՆԵՐԸ ԵՎ ԴՐԱՆՔ ԿԻՐԱՌՈՂ ՄԱՍՆԱԳԵՏՆԵՐԸ	49
2.1 Մեծ տվյալ ներքի և ավանդական տվյալ ներքի տնտեսագիտական վերլուծության առանձնահատկությունները և հիմնական տարբերությունները կառավարման ոլորտում	49
2.2 Տվյալ ագետների աշխատուժի առաջարկը ՀՀ աշխատանքային ընտրանքային հարցման արդյունքների վերլուծությանը	61
ԳԼՈՒԽ 3 – ԿԱՌԱՎԱՐՄԱՆ ՈԼՈՐՏՈՒՄ ՄԵԾ ՏՎՅԱԼ ՆԵՐԻ ՕԳՏԱԳՈՐԾՄԱՆ ԳՈՐԾԻՔԱԿԱԶՄԸ ԵՎ ԱՐԴՅՈՒՆԱՎԵՏՈՒ ԹՅԱՆ ԲԱՐՁՐԱՑՄԱՆ ՀԻՄՆԱԽՆԴԻՐՆԵՐԸ	79
3.1 Տվյալ ներքի վերլուծության դասական և տվյալ ագիտական մեթոդների համեմատական արդյունավետությանը գործարար կառավարման ոլորտում (վարկային տվյալ ներքի օրինակով)	79
3.2 Մեծ տվյալ ներքի օգտագործման արդյունավետության մատրիցը. ՀՀ կազմակերպության ներքի կողմից Մեծ տվյալ ներքի օգտագործման համեմատական արդյունավետության գնահատում ...	94
3.3 Մեծ տվյալ ներքի օգտագործման արդյունավետության բարձրացման հիմնախնդիրները և նախադրյալները ՀՀ-ում	120
ԵԶՐԱԿԱՑՈՒ ԹՅՈՒՆՆԵՐ	130
ՕԳՏԱԳՈՐԾՎԱԾ ԳՐԱԿԱՆՈՒ ԹՅԱՆ ՑԱՆԿ	135
ՀԱՊԱՎՈՒ ՄՆԵՐԻ ՑԱՆԿ	139
ՀԱՎԵԼՎԱԾ 1. ՔԱՆԱԿԱԿԱՆ ՀԱՐՑՄԱՆ ՀԱՐՑԱԹԵՐԹ	140
ՀԱՎԵԼՎԱԾ 2. ՈՐԱԿԱԿԱՆ ՀԱՐՑՄԱՆ ՀԱՐՑԱԹԵՐԹ	143

ՆԵՐԱՃՈՒ ԹՅՈՒՆ

Հետազոտության թեմայի արդիականությունը: Ժամանակակից տեխնոլոգիաները հնարավոր և հեշտ են դարձնում տվյալների անընդհատ և ինքնավար հավաքագրումը: Համաձայն հանրաճանաչ IBM կազմակերպության կայքի՝ վերջին երկու տարվա ընթացքում աշխարհում կուտակված տվյալներն իրենց ծավալով գերազանցում են մինչ այդ հավաքագրված տվյալներին¹: Նման խոշորածավալ տվյալների հավաքագրման պարագայում դրանց կառավարումն ու վերլուծությունը յուրահատուկ հետազոտության կարիք ունեն, այդ իսկ պատճառով ժամանակակից տնտեսագիտության և կառավարման մեջ նման տվյալներն առանձնացվում են և հայտնի են որպես «Մեծ տվյալներ»: Մեծ տվյալներն իրենց լայն կիրառությունն են գտել գործարար աշխարհում, սակայն վերջինիս տնտեսագիտական ուսումնասիրություններն ու վերլուծությունները դեռևս սակավաթիվ են: Այս աշխատանքը փորձ է կատարում ուսումնասիրել Մեծ տվյալների ոլորտը՝ Մեծ տվյալների տնտեսագիտական վերլուծությունը դիտարկելով որպես գործարար որոշումների կայացման հիմնախնդիր:

Սույն աշխատությունը ներկայացնում է Մեծ տվյալների կիրառության միջազգային փորձը, ցուցադրում է նրա կիրառության արդյունավետությունը: Հետազոտության արդյունքում մշակվել է Մեծ տվյալների օգտագործման գործիքակազմ, որը հնարավորություն կտա ձեռներեցին գնահատել տվյալների մեծությունը, օգտագործման մակարդակը և հարաբերական արդյունավետությունը: Գործիքակազմը կիրառվել է ՀՀ շուկայում գործող կազմակերպությունների կողմից իրենց հասանելի տվյալների օգտագործման արդյունավետության գնահատման նպատակով:

Համաձայն համաշխարհային տնտեսագիտական ֆորումի 2015 և 2016 թվականների հաշվետվության՝ Մեծ տվյալների վերլուծությամբ զբաղվող մասնագետների (տվյալագետների) աշխատանքը արդեն 2 անգամ անընդմեջ համարվել է տարվա

¹ [Petter Bae Brandtzæg, "Big Data, for better or worse: 90% of world's data generated over last two years." ScienceDaily. ScienceDaily, May 2013](http://www.sciencedaily.com/news/technology/130513_big_data_for_better_or_worse_90_percent_of_worlds_data_generated_over_last_two_years.htm)

ամենապահանջվածը: Սույն հետազոտությունը ներկայացնում է ՀՀ աշխատանքային առաջարկն ապահովող տվյալագետների ժողովրդագրական կազմը և բացահայտում է նրանց եկամտի վրա ազդող հիմնական փոփոխականները:

Մեծ տվյալների վերլուծությունը նոր զարգացող ճյուղ է, որն իր զարգացումն է ապրում ոչ ավանդական եղանակով: Ի տարբերություն դասական գիտությունների՝ Մեծ տվյալների ոլորտն իր կիրառությունն է գտնում շուկայում՝ դեռևս տեսական վերլուծության չենթարկված: Այդ իսկ պատճառով տվյալ ոլորտին առնչվող գրականությունն այդքան էլ հարուստ չէ: Չնայած վերոնշյալին՝ մի շարք հայրենական և արտասահմանյան հեղինակներ անդրադարձել են թեմային առնչվող առանձին խնդիրների ուսումնասիրությանը: Այս աշխատությունը փորձում է համապարփակ վերլուծության ենթարկել Մեծ տվյալների ոլորտը՝ ներկայացնելով վերջինիս զարգացումները, կարևորությունը, կիրառությունները և արդյունավետության գործիքակազմը՝ գործարարորոշումների կայացման տեսանկյունից:

Հետազոտության նպատակը և խնդիրները: Հետազոտության նպատակն է գործարար կառավարման ասպարեզում Մեծ տվյալների տնտեսագիտական վերլուծության անհրաժեշտության ուսումնասիրությունը, վերջինիս համեմատական և կիրառական արդյունավետության գնահատումը վիճակագրական և տեսական մոդելների օգնությամբ: Նպատակին հասնելու համար ատենախոսությունում քննարկվել, ձևակերպվել, առաջադրվել և լուծում են ստացել հետևյալ խնդիրները.

- ուսումնասիրել, ներկայացնել և տալ հստակ բնորոշում Մեծ տվյալների, վերջինիս տնտեսագիտական վերլուծությանն ու գործարար կառավարման ասպարեզում կիրառման անհրաժեշտությանը,
- վերհանել Մեծ տվյալների ոլորտի զարգացման միտումները և ներկայացնել ոլորտի հիմնական «խաղացողներին»,
- բացահայտել Մեծ տվյալների տնտեսագիտական վերլուծության մեթոդական առանձնահատկությունները,

- կատարել տվյալների դասական և տվյալագրիտական վերլուծության արդյունքների համեմատություն միևնույն տվյալների շտեմարանի օգտագործման պարագայում,
- վերլուծել ՀՀ աշխատանքային տվյալագրիտների աշխատուժի առաջարկը, վերջինիս ժողովրդագրական կազմը և համեմատել միջազգային ցուցանիշների հետ,
- բացահայտել ՀՀ աշխատանքային տվյալագրիտների եկամտային խմբի վրա ազդող փոփոխականները,
- մշակել Մեծ տվյալների օգտագործման գործիքակազմ և զարգացնել վերջինիս օգտագործման արդյունավետության մատրիցը,
- ուսումնասիրել ՀՀ-ում գործող կազմակերպությունների կողմից գործարար որոշումների կայացման գործընթացում Մեծ տվյալների կիրառության մակարդակն ու մատրիցում նրանց տեղադրությունը,
- բացահայտել Մեծ տվյալների օգտագործման արդյունավետության հիմնախնդիրները և ՀՀ-ում Մեծ տվյալների ոլորտի զարգացման անհրաժեշտ նախապայմանները:

Հետազոտության օբյեկտը և առարկան: Հետազոտության օբյեկտը Մեծ տվյալներն են, իսկ առարկան՝ գործարար կառավարման ասպարեզում վերջինիս տնտեսագրիտական վերլուծության հիմնախնդիրները:

Հետազոտության տեսական, մեթոդաբանական և տեղեկատվական հիմքերը: Ատենախոսության տեսամեթոդաբանական հիմք են հանդիսացել հայրենական և արտասահմանյան հեղինակների աշխատություններն ու զեկոլոյցները, հանրահայտ և Մեծ տվյալների ոլորտի զարգացման գործընթացում մեծ դեր ունեցող կազմակերպությունների հաշվետվությունները, միջազգային կառույցների հետազոտություններն ու ամփոփումները և այլ համացանցային ռեսուրսները, որոնք առնչվում են Մեծ տվյալների ոլորտին: Այստեղ կարելի է առանձնացնել Կ. Առոճուհի, Է. Բրինչոլ Ֆսոնի, Դ. Լանեյի, Գ. Վահանյանի, Ռ. Կապլանի, Դ. Նորթոնի, Պ. Դեննինգի, Ջ. Մաշեյի, Դ. Բլաքվելի և այլոց աշխատանքները, որտեղ

քննարկվել են ատենախոսության թեմային առնչվող առանձին հարցերի մեթոդական և գործնական խնդիրները:

Հետազոտության տեղեկատվական բազա են հանդիսացել ՀՀ ազգային վիճակագրական ծառայության (ԱՎԾ) պաշտոնական հրապարակումները, Ձեռնարկությունների ինկոնքատոր հիմնադրամի (ՁԻՀ), Համաշխարհային տնտեսագիտական ֆորումի (WEF), Տվյալների միջազգային կորպորացիայի (IDC) և այլ կառույցների տարեկան հրապարակումներն ու ուսումնասիրությունները, ինչպես նաև հեղինակի կողմից իրականացված որակական (խորքային) և քանակական (ընտրանքային) հարցման միջոցով հավաքագրված տվյալները:

Տեղեկատվության մշակման և վերլուծության համար օգտագործվել են տվյալների ամփոփման և խմբագրման, մշակման և գրաֆիկային պատկերման, կորելացիոն-ռեգրեսիոն վերլուծության, վիճակագրական մոդելավորման և խոր ուսուցման մեթոդներ, որոնք իրականացվել են MS Excel 2013 և IBM SPSS 22 ծրագրային փաթեթների ու R և Python ծրագրավորման լեզուների միջոցով:

Ատենախոսության հիմնական արդյունքները և գիտական

Նորույթը: Ատենախոսության գիտական նորույթ է հանդիսանում որոշումների կայացման գործընթացում Մեծ տվյալների տնտեսագիտական վերլուծության արդյունավետության հիմնախնդրի ուսումնասիրումը և գործնական լուծումների առաջարկումը: Գիտական նորույթը ամփոփված է հետևյալ հիմնական արդյունքներում.

- Վեր են հանվել տվյալների վերլուծության դասական և տվյալագիտական մոտեցումների առանձնահատկությունները, և գնահատվել է դրանց համեմատական արդյունավետությունը՝ որպես գործարար կառավարման մեթոդական հենք:
- Մշակվել են ընկերություններում տվյալների մեծության և կիրառելիության գնահատման կառուցակարգերը, որոնք հիմք են ստեղծում Մեծ տվյալները գործարար կառավարման հիմնախնդիրների

լ ու ծ ման մեջ կիրառել ու և գործիքակազմի ձևավորման համար:

- Մշակվել են կազմակերպության ներքին Մեծ տվյալների կիրառելիության նախադրյալներն ու պահանջները, ինչպես նաև Մեծ տվյալների օգտագործման արդյունավետության գնահատման մատրիցը, դրանում կազմակերպության ներքին տեղաշարժի հնարավորություններն ու հետևանքները:
- Վերլուծվել և բացահայտվել են ՀՀ-ում Մեծ տվյալների հավաքագրման և հասանելիության մեխանիզմների խոչընդոտները, մշակվել են դրանց ներդրման և օգտագործման ուղիներն ու հնարավորությունները, այդ թվում՝ ՀՀ աշխատաշուկայում տվյալագետների առաջարկի կառուցվածքը և դրավրաագդող հիմնական գործոնները:

Չեռագոտության տեսական և կիրառական նշանակությունը:

Ատենախոսության տեսական դրույթները կարող են հիմք հանդիսանալ Մեծ տվյալների ոլորտին նվիրված ապագա հեռագոտություն իրականացնելու համար: Չեռագոտության արդյունքները և եզրահանգումները, ներկայացված գործիքակազմն ու մեթոդները կարող են օգտագործվել ինչպես մասնավոր, այնպես էլ պետական և հանրային կազմակերպության ներքին կողմից Մեծ տվյալների հավաքագրման և կիրառության գնահատման, այդ ոլորտում ապագա ներդրումների պլանավորման, ռազմավարության մշակման նպատակով: Որպես մեթոդաբանական նյութ, ուսումնասիրության արդյունքները կարող են օգտագործվել (և արդեն իսկ օգտագործվում են) բուհերում և այլ ուսումնական հաստատություններում դասավանդման նպատակով:

Չեռագոտության արդյունքների փորձարկումը և հրապարակումները:

Ատենախոսության հիմնական արդյունքները հրապարակվել են հեղինակի 7 գիտական հոդվածներում, քննարկվել ու ներկայացվել են ԵԿՏԱ Տնտեսագիտության և կառավարման ամբիոնում, ՀԱՀ գործարարության և տնտեսագիտության ֆակուլտետում և առանձին թեմատիկ քննարկումների ժամանակ:

Հետազոտության արդյունքները ներկայացվել են և հավանության արժանացել «Տնտեսագիտության, գործարարության և կառավարման 4-րդ համաշխարհային կոնֆերանսի» և «Lindau տնտեսագիտական գիտությունների 6-րդ կոնֆերանսի» հանձնաժողովի կողմից:

Հետազոտության արդյունքների հիման վրա ձևավորվել, ներկայացվել, ընդունվել և ներկայումս ՀԱՀ գործարարության և տնտեսագիտության ֆակուլտետում դասավանդվում է «Բիզնես վերլուծություններ» առարկան, որը դիտարկում է գործարար կառավարման չորս առանցքներում որոշումների ընդունման գործընթացում Մեծ տվյալների տնտեսագիտական վերլուծության անհրաժեշտությունը, դերը, հնարավորություններն ու արդյունավետության գնահատման միջոցները:

ԵԿՏԱ Տնտեսագիտության և կառավարման ամբիոնի կողմից երաշխավորվել է աշխատության արդյունքների հիման վրա «Բիզնես վերլուծություններ» առարկայի ուսումնական ձեռնարկի ձևավորումը: Ձեռնարկին կից մշակման փուլում է գտնվում նաև առարկայի ուսուցման համար նախատեսված կայքը:

Ատենախոսության ծավալը և կառուցվածքը: Ատենախոսության կառուցվածքը և ծավալը պայմանավորված են հետազոտության նպատակով, առաջադրված խնդիրներով և մեթոդաբանական և վերլուծական տրամաբանությամբ: Աշխատանքը բաղկացած է ներածությունից, երեք գլուխներից, եզրակացություններից, հապավումների ցանկից, հավելվածից և օգտագործված գրականության ցանկից: Ատենախոսության ընդհանուր տեքստը (ներառյալ հավելվածները) կազմում է 134 էջ՝ ներառյալ 32 գծապատկեր, 16 աղյուսակ, 94 հղում և 2 հավելված՝ 5 էջ:

**ԳԼՈՒԽ 1. ՄԵՃ ՏՎՅԱԼ ՆԵՐԻ ԲՈՎԱՆԴԱԿՈՒ ԹՅՈՒՆԸ ԵՎ
ԿԻՐԱՄԱՆ ԱՆՐԱԺԵՇՏՈՒ ԹՅՈՒՆԸ
ՏՆՏԵՍԱԳԻՏՈՒ ԹՅԱՆ ԵՎ ԿԱՌԱՎԱՐՄԱՆ ՈԼՈՐՏՈՒՄ**

**1.1 Մեծ տվյալները՝ որպես տնտեսագիտական հետազոտության
առարկա**

Այն, ինչ չի չափվում, հնարավոր չէ կառավարել: Փիթեր Դրաքեր

Գլոբալացված և անընդհատ շարժման ու փոփոխությունների մեջ գտնվող տնտեսությունում ձեռնարկություններն ու կազմակերպությունները ցանկանում են ավելի լավ ճանաչել իրենց հաճախորդներին, ավելի լավ կառավարել ներքին միջավայրը, փոքրացնել արտաքին միջավայրի անորոշությունները, օգուտ քաղել հնարավորություններից, գտնել զարգացման նոր ուղիներ և այլն: Այդ նպատակներին հասնել ու համար նրանց պետք են տվյալներ և այդ տվյալների վերլուծությամբ զբաղվող մասնագետներ:

Ժամանակակից տեխնոլոգիաները հնարավոր և հեշտ են դարձնում տվյալների անընդհատ և ինքնավար հավաքագրումը: Երբ օգտատերերը որոնում են կատարում Google, Bing, Yandex և այլ ինտերնետային որոնման համակարգերում, նրանք հավաքագրում են տվյալներ յուրաքանչյուր օգտատիրոջ կողմից փնտրվող բառերի, գտնելուց հետո բացվող կայքերի, որոնումների հաճախականության և այլ գործողությունների վերաբերյալ: Երբ գնում է կատարվում Amazon կամ eBay առցանց խանութներից, վերջիններս ոչ միայն հավաքագրում և դիտարկում են բուն գնման մասին տվյալները, այլև մինչ գնումը և դրանից հետո հաճախորդի կողմից կայքում կատարվող բոլոր գործողությունները (այն է՝ այլ ապրանքների դիտում, համեմատություն, գույների կամ չափսերի ընտրություն և այլն):

Ժամանակակից տեխնոլոգիաների միջոցով տվյալների հավաքագրման գործընթացի փոփոխություններն ակնհայտ են ոչ միայն առցանց խանութներում, այլև սուպերմարկետներում և սովորական խանութներում: Մի քանի տասնյակ տարի առաջ սուպերմարկետները կարող էին հավաքագրել տվյալներ օրական վաճառքների մասին և դա կարող էր համարվել բարձրորակ տվյալ:

Ներկայումս հնարավոր է վաճառքի մասին տվյալը հավաքագրել իրական ժամանակահատվածում (անմիջապես վաճառքի պահին)՝ դրամարկղային սկաներների միջոցով: Ավելին, ժամանակակից սուպերմարկետները նաև հավաքագրում են տվյալներ սուպերմարկետի ներսում հաճախորդի շարժի, ապրանքների գնելու հերթականության մասին և այլն: Արդյունքում, այդ տվյալների հիման վրա, սուպերմարկետները որոշում են ապրանքների օպտիմալ դասավորվածությունը, սուպերմարկետի ներսում գովազդային վահանակների և ակցիաների արդյունավետ տեղադրությունը և այլն:

Չամաձայն հանրաճանաչ IBM կազմակերպության կայքի՝ վերջին երկու տարվա ընթացքում աշխարհում կուտակված տվյալներն իրենց ծավալով գերազանցում են մինչ այդ հավաքագրված տվյալներին²: Նման խոշորածավալ տվյալների հավաքագրման պարագայում դրանց կառավարումն ու վերլուծությունը յուրահատուկ հետազոտության կարիք ունի, այդ իսկ պատճառով ժամանակակից տնտեսագիտության և կառավարման մեջ նման տվյալներն առանձնացվում են և հայտնի են որպես «Մեծ տվյալներ» (անգլերեն՝ Big data)³:

Մեծ տվյալներն (այսուհետ՝ ՄՏ) իրենցից ներկայացնում են տվյալների բազմություն, որոնք այնքան խոշոր են և բարդ, որ հնարավոր չէ մշակել դասական տվյալների վերլուծության ծրագրերի, ալգորիթմների և մոդելների միջոցով: Առաջին անգամ տերմինը գործածության մեջ է դրվել տվյալների կառավարմամբ և վերլուծությամբ զբաղվող առաջատար կազմակերպություններից մեկի («SGI») ավագ հետազոտող, ձեռներեց Ջոն Մաշեյի կողմից՝ 1998 թվականին⁴: Չնայած դրան՝ «Մեծ տվյալներ» տերմինը լայնածավալ կիրառություն է ստացել միայն վերջին մի քանի տարիներին, ինչը բացատրվում է տեղեկատվական տեխնոլոգիաների ոլորտի բուռն զարգացմամբ:

² Տե՛ս [Petter Bae Brandtzæg. "Big Data, for better or worse: 90% of world's data generated over last two years." ScienceDaily. ScienceDaily, May 2013. https://www.sciencedaily.com/releases/2013/05/130522085217.htm](https://www.sciencedaily.com/releases/2013/05/130522085217.htm)

³ Բովանդակային տեսանկյունից այս տերմինի ճիշտ թարգմանությունն է հանդիսանում «Մեծ տվյալների բազա» տերմինը: Սակայն, ելնելով այն հանգամանքից, որ «Մեծ տվյալներ» տերմինը իր կիրառությունն է գտել առօրյա գործարար աշխարհում՝ վերջինս կօգտագործվի որպես «Big data» տերմինի հայերեն տարբերակ:

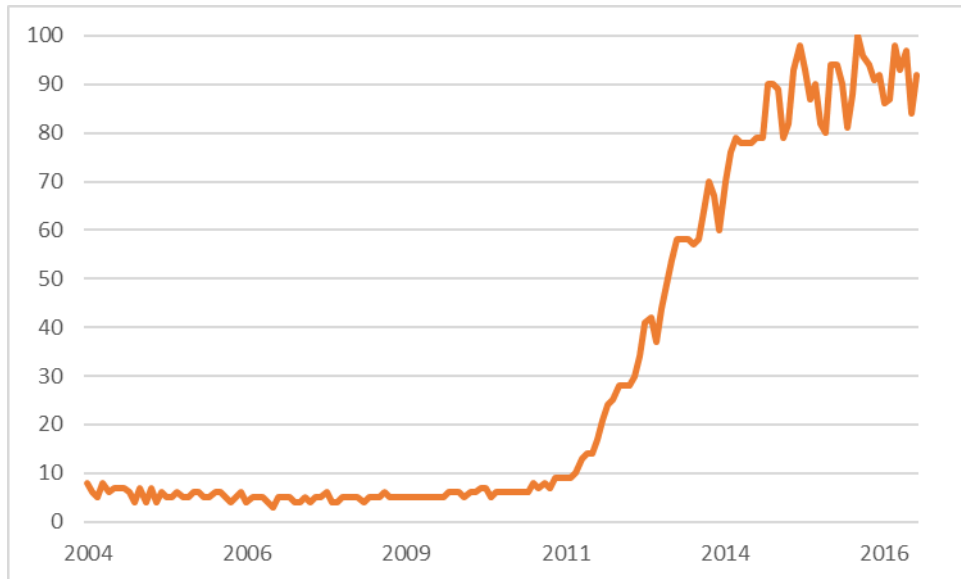
⁴ Տե՛ս [John R. Mashey. "Big Data ... and the Next Wave of InfraStress", Usenix, April 1998](#)

Հանրահայտ Google որոնման համակարգի տվյալներն օգտագործելով կարելի է տեսնել, թե ինչպես է «Մեծ տվյալներ» տերմինը ժամանակի ընթացքում ավելի ու ավելի հանրաճանաչ դառնում:

Գծապատկեր 1.

«Մեծ տվյալներ» տերմինի որոնումները Google որոնման համակարգում

**կառուցվել է մեր կողմից (աղբյուրը՝ Google trends⁵)*



Ինչպես երևում է գծապատկերից՝ «Մեծ տվյալներ» տերմինը միայն վերջին տարիներին է լայն տարածում գտել: Թերևս այդ պատճառով է, որ մինչ այժմ գոյություն չունի այդ տերմինի միանշանակ և համընդունելի սահմանում: Այնուամենայնիվ տարբեր հեղինակավոր կազմակերպություններ և հեղինակներ փորձել են իրենց սահմանումը տալ ՄՏ-ին: Բավականին հայտնի և ավանդական համարվող սահմանում է կառավարման ոլորտում խորհրդատվություն տրամադրող աշխարհի խոշորագույն կազմակերպության՝ McKinsey Global Institute-ի կողմից 2012 թվականին տրված հետևյալ սահմանումը.

«Մեծ տվյալներ տերմինը վերաբերում է տվյալների բազմություններին, որոնց չափերը գերազանցում են սովորական տվյալների շտեմարանների հետ աշխատող

⁵ Տե՛ս՝ <https://trends.google.com/trends/explore?date=all&q=big%20data>

գործիքների տվյալները կարդալու, պահպանելու, կառավարելու և վերլուծելու հնարավորությունները»:⁶

Այս սահմանումը, ըստ երևույթին, ընդգծում է այն խնդիրները, որոնք առաջանում են ՄՏ-ի վերլուծության պարագայում: Սույն հետազոտությանը գլուխ 1.2-ում անդրադարձ է կատարվում ՄՏ-ի տնտեսագիտական վերլուծության ինչպես ռավելություններին, այնպես էլ մարտահրավերներին՝ դիտարկելով նաև վերոնշյալ սահմանման մեջ նշված կետերը:

ՄՏ-ի վերլուծության շնորհիվ ծնվող հնարավորությունների տեսանկյունից է դիտարկում Տվյալների միջազգային կորպորացիայի փոխնախագահ (International Data Corporation) Ջոն Գանցի կողմից տրված սահմանումը.

«Մեծ տվյալների տեխնոլոգիաները իրենցից ներկայացնում են նոր սերնդի տեխնոլոգիաներ՝ ձևավորված մեծ ծավալներով բազմազան տվյալներից տնտեսագիտական արժեք ստանալու նպատակով:»⁷

Աշխարհահռչակ «Գարթներ» («Gartner») կազմակերպության տվյալների կառավարման և բիզնես վերլուծությունների բաժնի ղեկավար Դուգլաս Լանեյը 2001 թվականին հրատարակված իր հետազոտության մեջ տալիս է ՄՏ-ի սահմանումը՝ հիմնվելով 3 հիմնական բնութագրիչների վրա՝ խոշոր ծավալ, բարձր շարժունություն և ահռելի բազմազանություն:

«Մեծ տվյալներն իրենցից ներկայացնում են խոշոր ծավալով (volume), բարձր շարժունությամբ (velocity) և ահռելի բազմազանությամբ (variety) -ակտիվներ, որոնք պահանջում են մշակման նոր տարբերակներ՝ գործընթացների օպտիմալացման, տեղեկատվության բացահայտման և որոշումների կայացման ընդլայնման հնարավորությունն տալու համար:»⁸

Սույն հետազոտությունը հիմնվում է այս սահմանման վրա քանզի այն, թերևս, տալիս է համապարփակ պատկերացում ՄՏ-ի էության և կարևորության մասին՝ ընդգծելով այն 3 հիմնական

⁶ [McKinsey Global Institute \(MGI\), Big Data: The next frontier for innovation, competition, and productivity, Report, June, 2012, էջ 11](#)

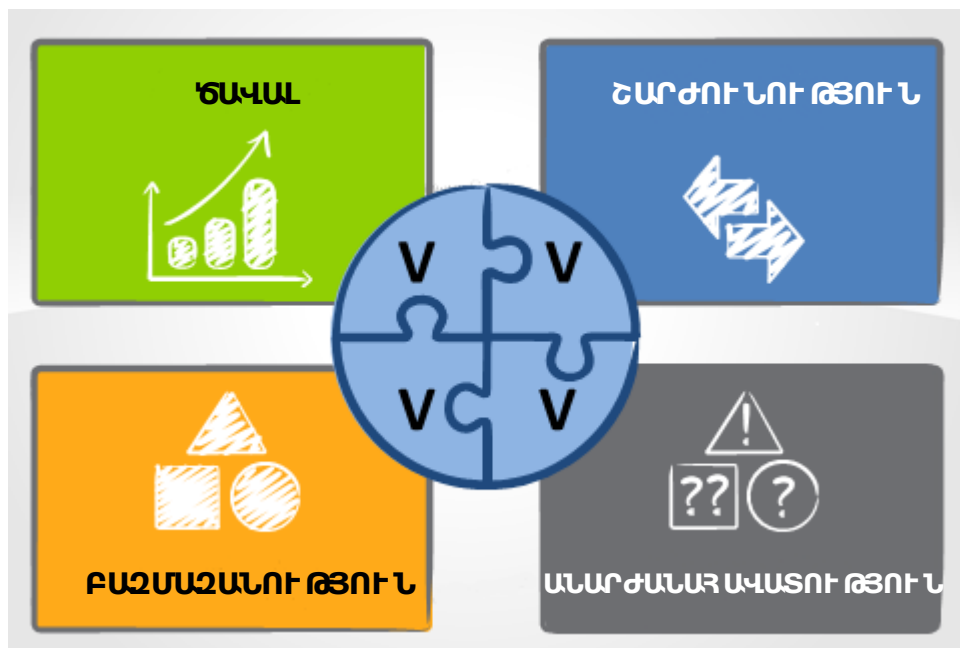
⁷ [Reinsel, D. Gantz, J. Extracting value from chaos. 2011, էջ 6](#)

⁸ [Laney, Douglas. "3D Data Management: Controlling Data Volume, Velocity and Variety". Gartner. February 2001](#)

բնութագրիչները, որոնցով ՄՏ-ն նկարագրվում են: Այդ իսկ պատճառով, վերջինս լայն տարածում է գտել որպես ՄՏ-ի 3 V-եր՝ ելնելով ծավալ, շարժունություն, բազմազանություն բառերի անգլերեն թարգմանությունների առաջին տառերից: Ժամանակակից գրականության մեջ հաճախակի տրվում է ՄՏ-ի ավելի ընդարձակ սահմանումը՝ գոյություն ունեցող երեք V-երին ավելացնելով չորրորդը՝ անարժանահավատություն (veracity):

Գծապատկեր 2.

Մեծ տվյալների 4 բնութագրիչները (4Vs)



ՄՏ-ի հետ առնչվող քննարկումներում հաճախ են հնչում հարցեր կապված դասական տվյալների վերլուծության և ՄՏ-ի վերլուծության տարբերության հետ: Այդ հարցերի պատասխանը թաքնված է 4 հիմնական բնութագրիչների մեջ, որոնք մանրամասն դիտակված են ստորև:

V1 – Ծավալ

Վերոնշյալ IBM ընկերության տվյալների համաձայն ամեն օր երկրում հավաքագրվում է մոտ 2,5 կվադրիտրիլիոն (10^{15}) մեգաբայթ ծավալ զբաղեցնող տվյալներ⁹: Համեմատության համար նշվում է, որ այդ քանակությամբ տվյալները լազերային սկավառակի (BLU-RAY) վրա պահպանելու դեպքում, անհրաժեշտ կլինեն այնպիսի քանակությամբ սկավառակներ, որ

⁹ Տե՛ս <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>

իրար վրա դնելով ավելի բարձր կլիներ քան 4 էյֆելյան աշտարակ:

Ագրեգացված ցուցանիշներից զատ ներկայացվում է նաև առանձին ձեռնարկությունների ցուցանիշներ՝ նշելով, որ ԱՄՆ-ում գործող ընկերությունների և կազմակերպությունների ճնշող մեծամասնությունը ունեն ավելի քան 100 տեռաբայթ (100000 գիգաբայթ) տվյալներ:

V2 – Շարժունություն

Շարժունությունը կարելի է ներկայացնել, որպես US-ի առաջացման արագություն: Ձեռնարկությունները, որոնք կարևորում են US-ի դերը գործարար կառավարման գործընթացում, ցանկանում են հնարավորինս արագ հավաքագրել տվյալները, որպեսզի մեկ քայլ առաջ լինեն իրենց մրցակիցներից: Արդյունքում տվյալների իրական ժամանակահատվածում վերլուծելու անհրաժեշտություն է առաջանում, որը թույլ կտա հաղթահարել ժամանակային լագը և խուսափել շարժունության հետևանքով հին տվյալների վրա որոշումների կայացման վտանգից:

V3 – Բազմազանություն

IBM կազմակերպության վերոնշյալ հրատարակությունում նշվում է, որ գոյություն ունեցող տվյալների ավելի քան 80%-ը չունեն ֆորմալ կառուցվածք: Դրանք բազմազան ու բազմատեսակ տվյալներ են՝ հավաքագրված տարբեր աղբյուրներից: Տեխնիկայի և տեխնոլոգիայի զարգացումը թույլ է տալիս հետազոտել բազմատեսակ տեղեկատվության աղբյուրներ, որոնք կարող են հանդես գալ ոչ միայն թվային տեսքով: Դրանց թվին են պատկանում որոնողական համակարգում որոնվող բանալի բառերը, սոցիալական ցանցերում կատարվող գրառումներն ու մեկնաբանությունները, համացանցում տեղադրվող նկարներն ու տեսաձայնագրությունները, աշխարհագրական դիրքի մասին հեռախոսակապի միջոցով անընդհատ ստացվող տվյալները և այլն:

V4 – Անարժանահավատություն

Եթե առաջին 3 բնութագրիչները տարբերակում են դնում US-ի և դասական տվյալների վերլուծության միջև, ապա 4-րդ բնութագրիչը հենց առաջին 3V-երի հետևանքն է: Մեծ ծավալով, անընդհատ շարժման մեջ գտնվող բազմազան տվյալները որոշումների կայացման համար անհրաժեշտ տեղեկատվության աղբյուր լինելուց բացի նաև անորոշության աղբյուր են հանդիսանում: KPMG կազմակերպության հարցման արդյունքում պարզվել է, որ հարցմանը մասնակցած ձեռներեցների 1/3-ը չի վստահում որոշումների կայացման գործընթացում իր կողմից օգտագործվող տեղեկատվությանը¹⁰: Այդուհանդերձ, ԱՄՆ տնտեսության տարեկան կորուստները անարժանահավատ տվյալների հիման վրա որոշում կայացնելու հետևանքով գնահատվում են մոտավորապես 3,1 տրիլիոն ԱՄՆ դոլար¹¹:

Ելնելով այն հանգամանքից, որ անարժանահավատությունը հանդիսանում է նախորդ 3 բնութագրիչների հետևանք, ավելին՝ այն չի բնութագրում մեծությունն՝ **սույն հետազոտության համար հիմք է ընդունվում 3 բնութագրիչների վրա հիմնված սահմանումը**, որը նաև ամենալայն տարածումն ունի մեզ հայտնի տնտեսագիտական հետազոտություններում:

Հիմնվելով այդ 3 բնութագրիչների վրա՝ աղյուսակ 1-ը ներկայացնում է US-ի և «ավանդական» տվյալների հիմնական տարբերությունները.

Աղյուսակ 1.

Երեք բնութագրիչները՝ Մեծ տվյալների և «ավանդական» տվյալների համար

**Կազմվել է մեր կողմից:*

Բնութագրիչ	Ավանդական տվյալներ	Մեծ տվյալներ
Ծավալ	Մեգաբայթ, Գիգաբայթ	Չե տտաբայթ

¹⁰ Տե՛ս [KPMG, “CEO Outlook 2016”, 2016, p. 20](#)

¹¹ Տե՛ս [Harvard Business Review, “Cost of bad data”, 2016, https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year](https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year)

Ծարժուկություն	Ստատիկ կամ դինամիկ ժամանակային շարք	Իրական կամ գրեթե իրական ժամանակահատված
Բազմազանություն	Մի քանի փոփոխականներ, ֆորմալ կառուցվածք	Բազմաթիվ փոփոխականներ, ոչ ֆորմալ կառուցվածք

ՄՏ-ի վերջին շրջանի հանրայնացումը պատահական է: Ելնելով 3 բնութագրիչներից՝ նախկին տեխնոլոգիական հզորությունների պայմաններում նման ահռելի տվյալների հավաքագրումը հնարավոր չէր լինի: Այսպիսով, կարելի է ասել, որ ՄՏ-ի ոլորտի զարգացման հիմքում ընկած են երեք հիմնական պատճառներ.

- I. Զեռահաղորդակցման տեխնոլոգիաների զարգացումն ու ցանցերի լայնատարածումը – վերջին տարիներին տեխնոլոգիաների զարգացումը ավելի ու ավելի շատ մարդկանց է թույլ տալիս օգտվել համացանցից և թվային սարքավորումներից, ինչը հնարավոր է դարձնում տվյալների ակնթարթային հավաքագրումը: Մեկ րոպեի ընթացքում Google որոնողական համակարգում տեղի է ունենում 2,3 միլիոն որոնում: Ստացվում է, որ յուրաքանչյուր րոպե Google-ը ստանում է 2300000 նոր տվյալ¹²:
- II. Մեծածավալ տվյալների պահպանման ամպային համակարգերի ստեղծումը – ՄՏ հավաքագրելու համար պետք է կարողանալ այդ տվյալները պահպանել ինչ-որ տեղ: Սովորական համակարգիչների հիշողության ծավալը բնականաբար չի բավարարեցնի: Այդ իսկ պատճառով ստեղծվեցին այսպես կոչված ամպային տեխնոլոգիաներ, որոնք հնարավորություն են տալիս վերբեռնել տվյալները համացանց և վերլուծությունն իրականացնել օնլայն ռեժիմով՝ առանց տվյալները ներբեռնելու: Նման ծառայություններ են առաջարկում մի շարք առաջատար կազմակերպություններ, ներառյալ՝ Google, Microsoft, Amazon ընկերությունները:

¹² Տե՛ս <https://www.google.com/insidesearch/howsearchworks/thestory/index.html>

III. Թվային տեխնոլոգիաների միջոցով հաշվողական գործընթացների զարգացումն ու օպտիմալացումը – չնայած այն հանգամանքին, որ ամպային տեխնոլոգիաները թույլ են տալիս իրականացնել վերլուծությունն ՄՏ-ի հետ օնլայն ռեժիմով՝ միևնույն է այդ տվյալների մեծ մասը դեռևս չի օգտագործվում: Պատճառն այն է, որ հաշվողական տեխնիկան դեռ չի հասցրել այնքան զարգանալ, որ հնարավորություն տա վերլուծել ամբողջական հավաքագրված տվյալները: Գործարարության, տնտեսագիտության և կառավարման 4-րդ միջազգային կոնֆերանսին ներկայացված հետազոտությունների համաձայն՝ ֆինանսական ոլորտի ձեռնարկությունները չեն օգտագործում իրենց հավաքագրված տվյալների 80-90%-ը¹³:

ՄՏ-ի վերլուծությունը նոր զարգացող ոլորտ է, որն իր զարգացումն է ապրում ոչ ավանդական եղանակով: Ի տարբերություն դասական գիտությունների՝ ՄՏ ոլորտն իր կիրառությունն է գտնում շուկայում՝ դեռևս տեսական վերլուծության չենթարկված: Մակեդոնիայի համալսարանի տնտեսագիտության դեպարտամենտի 3 պրոֆեսորների կողմից 2013 թվականին կատարվել է հետազոտություն, որի նպատակն է պարզել, թե որքան ակադեմիական հետազոտություն են կատարվում ՄՏ ոլորտում գիտության տարբեր ճյուղերում և տվյալների տարատեսակ աղբյուրների ուսումնասիրությամբ^{14,15}: Գծապատկեր 3-ում ցուցադրվում են հետազոտության արդյունքները:

Գծապատկեր 3.

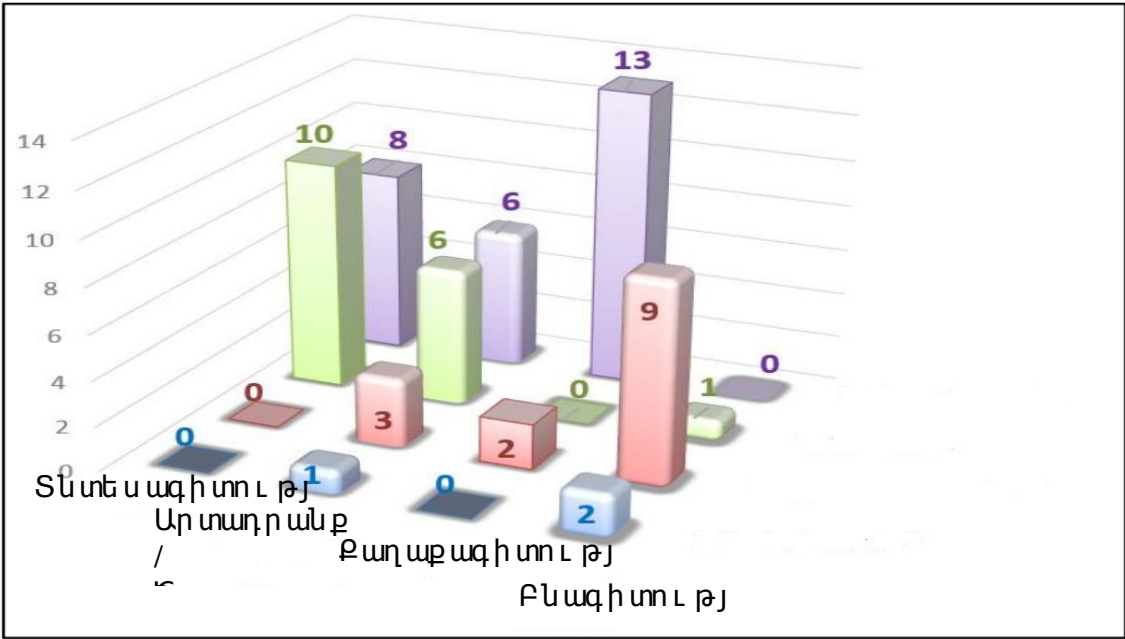
Մեծ տվյալների ոլորտում կատարված ակադեմիական հետազոտությունների քանակը՝ ըստ հետազոտության ոլորտների¹⁶

¹³ Տե՛ս [Paul Zikopoulos, “Understanding Big Data”, 2012](#)

¹⁴ Տե՛ս [Tarabanis K. Kalampokis E., Tambouris E., “Understanding the Predictive Power of Social Media”, Internet Research, Vol.23, September 2013](#)

¹⁵ 4 գոյություն ունեցող ներկայացնում են տվյալների 4 տարատեսակ աղբյուրներ, որոնք օգտագործվել են գծապատկերում նշված ոլորտներում հետազոտության նպատակով:

¹⁶ Տե՛ս նույն տեղում:



Չարկավոր է նշել, որ այնտեղ ընդգրկված հետազոտությունները հաշվի են առնում միայն անգլերեն, գիտական ժողովածուներում տպագրված և ՄՏ ոլորտն ուղղակիորեն հետազոտության ենթարկող աշխատությունները: Այն հետազոտությունները, որոնք կատարվել են մասնավոր կազմակերպությունների կողմից (նույնիսկ եթե գիտական բնույթի են), բայց չեն տպագրվել ժողովածուներում, կամ այն աշխատանքները, որոնք առնչվում են ՄՏ ոլորտին, սակայն ուղղակիորեն չեն հետազոտում այդ ոլորտում տեղի ունեցող տնտեսագիտական գործընթացները ընդգրկված չեն ստորև ներկայացված գծապատկերում:

Ինչպես երևում է գծապատկերից՝ ամենաշատ տպագրված ակադեմիական հետազոտությունները վերաբերում են տնտեսագիտությանը (18 հետազոտություն)¹⁷: Երկրորդ տեղն են զբաղեցնում արտադրանքին կամ ծառայությանը վերաբերող հետազոտություններ (16): Միանշանակապես սաբավականին փոքր թիվ է, սակայն այն պայմանավորված է ՄՏ ոլորտի երիտասարդությանը և վկայում է այդ ոլորտում նորանոր հետազոտական աշխատանքների անհրաժեշտության մասին:

Չնայած ոլորտի երիտասարդությանը՝ ՄՏ-ի զարգացումն ու հանրայնացումը անակնկալ երևույթ է և կանխատեսվել է մի շարք

¹⁷ 4 տարբեր տվյալների աղբյուրներից յուրաքանչյուրի օգտագործմամբ համապատասխանաբար 0,0,10 և 8 հետազոտություն

հեղինակների աշխատություններում: Այսպես, դեռևս 1990թ. Մասաչուսեթսի տեխնոլոգիական համալսարանի պրոֆեսոր Փիթեր Դենինգը կանխատեսում էր, որ համակարգչային տեխնոլոգիաների զարգացումը թույլ կտա հավաքագրել և վերլուծել մեծածավալ և բազմաբնույթ տվյալներ՝ կանխատեսումներ կատարելու և գործարար որոշումներ իրականացնելու նպատակով¹⁸: Յամակարգչային տեխնոլոգիաների և հատկապես մեքենայական գրաֆիկայի զարգացմանն են նվիրված նաև հայրենական հեղինակներից Գ. Վահանյանի մի շարք աշխատություններ: Մասնավորապես «Մեքենայական գրաֆիկա» աշխատության մեջ հեղինակը դեռևս 1985թ.-ին ընդգծում է պլանավորման և կառավարման գործընթացներում թվային տեխնոլոգիաների անընդհատ աճող դերի, գործարար խնդիրների տնտեսամաթեմատիկական մոդելավորման հնարավորությունների մասին՝ ներկայացնելով ինչպես ԽՍՀՄ և միջազգային փորձը, այնպես էլ կառավարման ոլորտում համակարգչային գրաֆիկայի կիրառության ավգործիմները, մեթոդներն ու մոտեցումները¹⁹: Յեղինակի մեկ այլ աշխատությունում («Виртуальные технологии менеджмента (системотехника менеджмента)», 2005թ.) ներկայացվում է տեղեկատվական տեխնոլոգիաների միջոցով կառավարման գործընթացի կատարելագործումը՝ առանձնացնելով նպատակին հասնելու գործընթացում խնդիրների մոդելավորման, ավգործիմների ձևավորման, ծրագրավորման, տեխնիկական օժանդակության, մտավոր կապիտալի կիրառման և տեղեկատվության ու գիտելիքի ստացման փուլերը²⁰:

Չնայած այն հանգամանքին, որ ՄՏ վերլուծության արժեքի մասին հետազոտությունները սակավ են (կամ չկան)՝ դասական տվյալների վերլուծության արժեքի հետկապված տեսությունները, միևնույն է սերտորեն կապված են այս աշխատանքի հետ: Տվյալների վերլուծության արժեքի հետ կապված ժամանակակից տեսությունները իրենց սկիզբն են առել տնտեսագետ-

¹⁸ Denning P., "Saving All The Bits", American Scientist, Volume 78, Issue 5, Oct 1990, էջեր 402-405

¹⁹ Ваганян ГА, "Машинная графика в управлении", Монография, "Айастан", Ереван, 1985, էջեր 19-21

²⁰ Ваганян ГА, Ваганян ОГ, "Виртуальные технологии менеджмента (системотехника менеджмента)", Монография, "Нжар", Ереван, 2005, էջեր 14-16

վիճակագրագետ Դեյվիդ Բլեքուելի 1953 թվականին հրատարակված աշխատությունից²¹: Աշխատության մեջ Բլեքուելը ներկայացնում է իր հայտնի թեորեմը, համաձայն որի տեղեկատվության վերլուծությունից ստացվող օգտակարության սպասված արժեքը միշտ դրական է: Վերջինս հիմք է հանդիսանում մի շարք ժամանակակից հետազոտությունների, որոնք ուսումնասիրում են տեղեկատվությունը ստանալու նպատակով իրականացվող ներդրումների արդյունավետությունը: Մասաչուսեթսի համալսարանի տնտեսագիտության և կառավարման բաժնի պրոֆեսոր Էրիկ Բրինջոլֆսոնն ունի ավելի քան 20 աշխատություն, որոնք հետազոտում են տեղեկատվությունը ստանալու վրա կատարվող ներդրումների բիզնես արժեքը²²: Նրա հետազոտությունները ցույց են տալիս, որ տեղեկատվությունը ստանալու համար կատարվող ներդրումները, որոնք հիմնականում տեղեկատվական տեխնոլոգիաների ձեռքբերման կամ թարմացման տեսքով են դրսևորվում, դրականորեն են ազդում աշխատողների արտադրողականության վրա²³: Բայց նայնպես, նշվում է, որ արտադրողականության աճը տարբեր է՝ կախված ընկերության հատուկ փոփոխականներից (այն է՝ շուկա, աշխարհագրական դիրք, ներքին միջավայր և այլն):

Մեկ այլ հետազոտություն, որն արդեն իսկ վերաբերում է ՄՏ-ի վերլուծության ներդրումային եկամտաբերությանը, իրականացվել է Տեխասի համալսարանի գործարարության ամբիոնի պրոֆեսորների կողմից²⁴: Նրանց հետազոտությունն ընդգծում է, որ մի շարք կազմակերպություններ կարողացել են ապահովել բարձր ֆինանսական ցուցանիշներ՝ ՄՏ-ի վերլուծության մեջ ներդրում անելու շնորհիվ: Աղյուսակ 2-ում ներկայացված են 5 կազմակերպություն, որոնք ՄՏ-ի վերլուծության միջոցով կարողացել են ունենալ գրավիչ ֆինանսական ձեռքբերումներ:

²¹ Տե՛ս Blackwell D., "Equivalent Comparisons of Experiments". The Annals of Mathematical Statistics, 1953, էջեր 265-272

²² Բրինջոլֆսոն Է., տվյալ ոլորտին հարող աշխատությունների ցանկ – <http://ebusiness.mit.edu/erik/ITandBusinessvalue.html>

²³ Տե՛ս Sinan Aral, Erik Brynjolfsson, and Marshall W. Van Alstyne, "Information, Technology, and Information Worker Productivity" Information Systems Research, March, 2012.

²⁴ Տե՛ս [Barua A., Manis D., Mukherjee R., Measuring the Business Impacts of Effective Data, University of Texas, Austin 2015](#)

Հատկանշական է, որ չնայած այդ 5 կազմակերպություններից 4-ը մասնավոր ընկերություն են՝ Բրազիլիայի Սան Պաուլո համայնքը նույնպես ընդգրկված է այդ շարքում: Սա ցույց է տալիս այն հանգամանքը, որ ՄՏ-ի վերլուծությունը կիրառելի և եկամտաբեր է ոչ միայն մասնավոր կազմակերպությունների համար, այլև պետական հատվածի:

Որպեսզի հնարավոր դառնա քանակապես նկարագրել ՄՏ-ի վերլուծության ազդեցությունը կազմակերպության գործունեության վրա՝ Տեխասի համալսարանի հետազոտողները հարցում են անցկացրել աշխարհի 150 խոշորագույն հասուն ընկերություններին կազմակերպություններում՝ հավաքագրելով \$իսանսական ցուցանիշների մասին տվյալներ: Հետազոտության ընթացքում պարզվել է, որ ՄՏ-ի վերլուծության կիրառության միջոցով կարելի է միջինում հասնել կապիտալի եկամտաբերության 16% և միջին ընդհանուր հասուն թի 2,01 միլիարդ ԱՄՆ դոլար աճի:

Այլ ուսակ 2.

Մեծ տվյալների վերլուծության ներդրումային եկամտաբերության օրինակներ²⁵

Կազմակերպություն	Ձեռքբերում
American Airlines	1 մլն ԱՄՆ դոլար արժողությամբ կեղծիքների հայտնաբերում
Սան Պաուլոյի համայնք, Բրազիլիա	100 մլն ԱՄՆ դոլար արժողությամբ չհարկված եկամուտների հայտնաբերում
Cell	Միայն մեկ պրոյեկտի վրա 20 մլն ԱՄՆ դոլար խնայողություն
AOK Hessen	3,2 մլն ԱՄՆ դոլար արժողությամբ կեղծիքների հայտնաբերում
HMV Japan	Տարատեսակ ապրանքների վաճառքի 3-15 անգամ աճ

Կազմակերպության արտադրանքի ծավալի և աշխատակիցների արտադրողականության վրա ՄՏ-ի ազդեցությանն է վերաբերում Մասաչուսեթսի տեխնոլոգիական համալսարանի և Ուարթոն քիզնես

²⁵ Տե՛ս նույն տեղում

դարոցի հետազոտողների համատեղ աշխատանքը²⁶: Վերջիններս ուսումնասիրում են 179 խոշոր բաց բաժնետիրական ընկերություններ և բաժանում նրանց երկու խմբի՝ ընկերություններ որոնք կիրառում են ՄՏ-ի վերլուծությունները որոշումների կայացման գործընթացում և ընկերություններ, որոնք չեն կիրառում: Համաձայն այդ հետազոտության՝ որոշումների կայացման գործընթացում ՄՏ-ի վերլուծությամբ զբաղվող ընկերությունները միջինում 5-6% ավելի բարձր արտադրողականություն ունեն, քան չօգտագործողները:

ՄՏ-ի վերլուծության ոլորտում ակադեմիական հետազոտությունները բացը լրացնում են մասնավոր կազմակերպությունները, որոնք փորձում են տարբեր քանակական և որակական հարցումների միջոցով նկարագրել ՄՏ-ի վերլուծության պահանջարկը, առաջարկը և վերջինիս կիրառությամբ զբաղվելու հիմնական խոչընդոտները: Արդեն իսկ նշված Գարթներ ընկերությունը 2013 թվականին հրատարակեց ոլորտի խոչընդոտների մասին իրենց հետազոտությունը²⁷: Առկա տվյալներից ամենաուշագրավը, թերևս, հետևյալ 3-ն են.

- Հարցված կազմակերպությունների 75%-ը հնարավորություն չունեն արդի տեխնոլոգիաների և նրանց համալրման միջոցով վերլուծել ՄՏ-ն: Սա նշանակում է, որ այդ կազմակերպությունները ստիպված կլինեն արմատապես փոխել օգտագործվող տեխնոլոգիան՝ ՄՏ-ի վերլուծությամբ զբաղվելու համար:
- Հարցված կազմակերպությունների 86%-ը չեն կարողանում տեղեկատվությունը ստանալ և՛ բարձր որակով (ճշգրիտ), և՛ իրական ժամանակահատվածում: Այդ երկուսի միջև առկա է հակադարձ համեմատական կապ, այն իմաստով, որ ճշգրիտ տվյալներ հավաքագրելու համար, կազմակերպությունների ճնշող մեծամասնությունից ժամանակ է պահանջվում: Եվ հակառակը, իրական

²⁶ Տե՛ս Brynjolfsson E., Heekyung K., Hitt L., Strength in Numbers: How Does Data-Driven Decision-making Affect Firm Performance?, 2011

²⁷ Տե՛ս [Gartner Business Intelligence & Information Management Summit 2013](http://www.gartner.com/imagesrv/summits/docs/apac/business-intelligence/BIIIM-Summit-2013-Brochure.pdf), <http://www.gartner.com/imagesrv/summits/docs/apac/business-intelligence/BIIIM-Summit-2013-Brochure.pdf>

Ժամանակահատվածում տվյալները ունենալը կոմպենսացվում է տվյալների ոչ ամբողջականությամբ և ոչ ճշգրտությամբ:

- 501-1000 աշխատող ունեցող կազմակերպությունների 79%-ը նշում են, որ US-ի վերլուծությամբ զբաղվելը խնդիր է առաջացնում իրենց աշխատակիցների մոտ՝ համապատասխան մոդելների չիմացության պատճառով: Պատկերն այլ է ավելի խոշոր կազմակերպություններում: 3000 և ավելի աշխատող ունեցող կազմակերպություններից միայն 55%-ի մոտ է US-ի վերլուծությունը դիտվում որպես խնդիր:

Ընկերության աշխատակիցների կողմից US-ի վերլուծությամբ զբաղվելու համար անհրաժեշտ գիտելիքների պակասը ընդգծվում է նաև Արկանզասի և Վերբերի Պետական Համալսարանի կառավարման ամբիոնի դասախոսների 2013 թվականին կատարված հետազոտության մեջ²⁸: Ըստ այդ հետազոտության՝ US-ի վերլուծությամբ զբաղվող հիմնական մասնագետները ունեն մաթեմատիկայի և ծրագրավորման ոլորտում բարձրագույն կրթություն, և չեն տիրապետում ոլորտին վերաբերող խնդիրներին (օրինակ՝ մաթեթիկագային քաղաքականություն): Մյուս կողմից, ոլորտի խնդիրների մասին քաջատեղյակ մասնագետները, ըստ վերոնշյալ հետազոտության, չունեն հմտություններ US-ի վերլուծությամբ զբաղվելու նպատակով: Հետազոտությունում նշվում է, որ US-ի հաջողակ վերլուծաբանները, որոնց ընդունված է նաև անվանել տվյալագետ (անգլերեն՝ Data Scientist) հանդիսանում են այն մասնագետները, որոնք կարողանում են տնտեսագիտությունը համակցել վիճակագրության և համակարգչային գիտելիքների հետ:

Վերոնշյալ մասնագետների, ինչպես նաև US-ի վերլուծության նկատմամբ գոյություն ունեցող առեւտրի պահանջարկը բացատրվում է կիրառելիության բազմազանությամբ ու բազմաբնագավառությամբ: Հետազոտության հաջորդ հատվածում ներկայացվում է US-ի տնտեսագիտական վերլուծության առավելություններն ու թերությունները, կիրառության

²⁸ Տե՛ս [Matthew A., Stanley E., "Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management", Journal of Business Logistics, 2013, 34\(2\), էջեր 77–84](#)

անհրաժեշտությունը, առանձնահատկությունները և միջազգային փորձը՝ գործարար որոշումների կայացման գործընթացում:

1.2 Մեծ տվյալների տնտեսագիտական վերլուծության գործնական կիրառությունը՝ միջազգային փորձի հաշվառմամբ

Կիրառությունն ու միջազգային փորձը

Մեծ տվյալներն իրենց լայն կիրառությունն են գտել գործունեության ամենատարբեր ոլորտներում գործող ձեռնարկությունների կողմից տարատեսակ որոշումների կայացման գործընթացում²⁹: Հետազոտության այս հատվածի նպատակն է դասակարգել ՄՏ-ի կիրառությունն ըստ.

- I. Ցուցանիշների հաշվեկշռված համակարգի (ՑՀՀ) 4 առանցքների,
- II. Տնտեսության ոլորտների,
- III. Ըստ տվյալների հավաքագրման աղբյուրների:

Մեծ տվյալների կիրառության դասակարգումն ըստ ՑՀՀ 4 առանցքների

Հաջողության հասնելու համար ընկերությունները ցանկանում են մշակել ճիշտ և հստակ ռազմավարություն և, ինչպես Փիթեր Դրաբերն է նշում, ունենալ չափելի նպատակներ: Ժամանակակից տնտեսագիտության և կառավարման տեսությունում բավականին լայն կիրառությունն է գտել Ցուցանիշների հաշվեկշռված համակարգի մոդելը (Balanced ScoreCard), որը ընկերության գործունեությունը դիտարկում է չորս հիմնական առանցքներից՝ հաճախորդներ, անձնակազմ, գործառնություններ, ֆինանսներ³⁰: Այս մոտեցումը չափելի է դարձնում կազմակերպության նպատակները և օգնում է մշակել անհրաժեշտ ռազմավարություն՝ նրանց հասնելու համար:

²⁹ Դավթյան, Յ., Big data in different sectors of the economy: application scenarios. «Ակունք» ուսումնամեթոդական հոդվածների ժողովածու, №2 (14), Երևան, 2016 թ., էջեր 122-125
³⁰ Kaplan R., Norton D., "The Balanced Scorecard: Translating Strategy into Action", Harvard Business School Press, August 1996, էջեր 43-46

Յետագոտության սույն հատվածում փորձ է արվում ներկայացնել ՄՏ-ի գործնական կիրառությունը գործարար կառավարման ոլորտում ելնելով Յուգանիշների հաշվեկշռված համակարգի չորս առանցքներից: Յուրաքանչյուր առանցքում դիտարկվում են տվյալ դասին պատկանող խնդիրներ, որոնք կարելի է լուծել ՄՏ-ի վերլուծության միջոցով՝ տնտեսության տարբեր ոլորտներում:

Գծապատկեր 4.

Ընկերության գործունեության 4 առանցքները՝ ըստ

Յուգանիշների հաշվեկշռված համակարգի

1) Հաճախորդների վերլուծություն



Ընկերության ռազմավարական նպատակներին հասնելու համար կրիտիկական գործոնների թվում է դասվում հաճախորդների հետ ամուր, շահավետ և երկարաժամկետ հարաբերությունների կառուցումը: Ժամանակակից գործարար աշխարհում հաճախորդներին միանման ագրեգացված մոտեցման ցուցաբերումը մրցակցային բարձր դիրքերի հանգեցնել չի կարող: Այդ իսկ պատճառով կազմակերպությունները ցանկանում են ճանաչել իրենց հաճախորդներին և յուրաքանչյուրին տրամադրել անհատականացված մոտեցում: Հաճախորդներին ճանաչելու համար ընկերությունները հավաքագրում են տվյալներ տարբեր ուղիներով՝ կախված գործունեության ոլորտից, ինչը հնարավորություն է տալիս բարձրացնել հաճախորդների բավարարվածությունը մատուցված ծառայությունից կամ արտադրված/վաճառված ապրանքից, շահել

հաճախորդի վստահությունը, հաստատել երկարաժամկետ հարաբերություններ:

Հաճախորդների մասին հավաքագրված ՄՏ-ի վերլուծության վառ օրինակ է թիրախային զեղչի կիրառությունը: Ընկերություններն իրենց կողմից վաճառվող ապրանքների և ծառայությունների գները սովորաբար զեղչում են ելնելով սեզոնայնությունից կամ այլ գործոններից, որոնք անկախ են հաճախորդներից յուրաքանչյուրի անհատական գործոններից: Սակայն տարբեր հաճախորդներ նույն ապրանքի կամ ծառայության համար ունեն վճարման տարբեր պատրաստակամություններ (անգլերեն՝ willingness to pay): Վճարման տարբեր պատրաստակամություն ունեցող հաճախորդներին միանման հավասարակշիռ գին առաջարկելու հետևանքով առաջանում է միկրոտնտեսագիտության մեջ հայտնի սպառողական ավելցուկը:

ՄՏ-ի վերլուծությամբ կարելի է պարզել յուրաքանչյուր հաճախորդի վճարման պատրաստակամությունը և վերջինիս առաջարկել համապատասխան գին (զեղչ): Այս մոտեցումը թույլ կտա կիրառել գնային խորականություն և արտադրողի/վաճառողի կողմից յուրացնել սպառողական ավելցուկը:

2) Անձնակազմի կառավարում

Անձնակազմի կառավարումը ՄՏ-ի վերլուծության կիրառման (և ընդհանրապես տվյալների կիրառման) թերևս առավել քիչ զարգացած ոլորտն է: Ծառ ժամանակ կառավարիչները ցուցադրում են ընդհանրացված մոտեցում իրենց աշխատակիցների նկատմամբ՝ վերջիններիս մոտ աշխատանքային պայմաններից անբավարարվածություն առաջացնելով: Եթե այդ խնդիրը ցանկության դեպքում լուծելի է հարցումների կամ քննարկումների միջոցով փոքր կազմակերպություններում, ապա խոշոր կազմակերպություններում, անկախ կազմակերպական կառուցվածքից, առաջանում է աշխատակիցների անհատականության ճանաչման խնդիր:

ՄՏ-ի վերլուծությամբ կարելի է պարզել, թե որ բաժինների աշխատակիցներն են միմյանց հետ ավելի շատ առնչություն ունենում և, այդ տվյալներն օգտագործելով, իրականացնել կազմակերպության ներսում պահանջարկի պլանավորում,

կազմակերպել բաժինների աշխարհագրական վերադասավորումը, առկա թափուր աշխատատեղերի մասին հայտարարությունները նախօրոք տարածել ներքին միջավայրում՝ գործունեության ամենամոտ աշխատակիցների շրջանում և այլն:

3) Գործառույթներ կառավարում

Նոր ապրանք արտադրելու կամ նոր ծառայություն մատուցելու որոշումը դեռ վաղուց հիմնված է եղել շուկայի ուսումնասիրության, առավելությունների և թերությունների, հնարավորությունների և վտանգների վերլուծության վրա: ՄՏ-ի հավաքագրումն ու վերլուծությունը թույլ է տալիս ավելի արդյունավետ դարձնել գործառույթների կառավարումը՝ շուկայի վերլուծությունը հանելով նոր մակարդակի վրա: Վերջինիս վառ օրինակ է ընկերության կողմից հին ապրանքի նոր տարբերակը շուկա հանելու որոշումները: Ժամանակակից մի շարք կայքեր և ծառայություններ հնարավորություն են տալիս կարծիք հայտնել գնված ապրանքի կամ ծառայության որակի մասին: Ծառ ընկերություններ (օր.՝ հեռախոսային ծրագրերի արտադրողներ) հավաքագրում են այդ տեքստային տվյալները և ավտոմատացված վերլուծության ենթարկում՝ պարզելով, թե ապրանքի կամ ծառայության, որ գործուն է ենթակափոփոխման, որը՝ ոչ:

4) Ֆինանսների կառավարում

Կազմակերպության ֆինանսները միշտ էլ վերլուծության առարկա են եղել: Հաշվապահություն վարելը գալիս է դեռ անտիկ ժամանակաշրջանից, երբ Եգիպտոսում և Բաբելոնում հաշվառում էին թագավորության ունեցվածքը, վերջինիս պարզ փոփոխությունները ժամանակի ընթացքում³¹: Այսօր հաշվապահություն վարելը և նշված ժամանակահատվածում հաշվետվություններ կայացնելը պարտադիր պայման է հանդիսանում ընկերությունների համար: Սակայն դասական հաշվապահությունը միայն նկարագրողական դեր է կրում՝ ներկայացնելով կազմակերպության ֆինանսական վիճակը, դրամական հոսքերը, դրանց շարժը ժամանակի ընթացքում:

³¹ [A History of Accountancy, New York State Society of CPAs, November 2003](#)

Այսօր, մի շարք կազմակերպություններ հավաքագրում և վերլուծում են ՄՏ՝ կեղծիքների բացահայտման և կազմակերպության ապագա ֆինանսական վիճակի մոդելավորման նպատակով: Իհարկե, դասական դարձած մի շարք մոդելներ ինչպիսիք են Բենֆորդի օրենքը և ARIMA մոդելները թույլ են տալիս հայտնաբերել կեղծիքները և կանխատեսել կազմակերպության ֆինանսական վիճակը համապատասխանաբար: Սակայն հարկ է նշել, որ ՄՏ-ի դերը որոշումների կայացման գործընթացում կայանում է ոչ միայն նոր մոտեցումների մշակման մեջ, այլև հին մոտեցումների արդյունավետության բարձրացման մեջ:

Մեծ տվյալների կիրառության դասակարգումն ըստ տնտեսության ճյուղերի

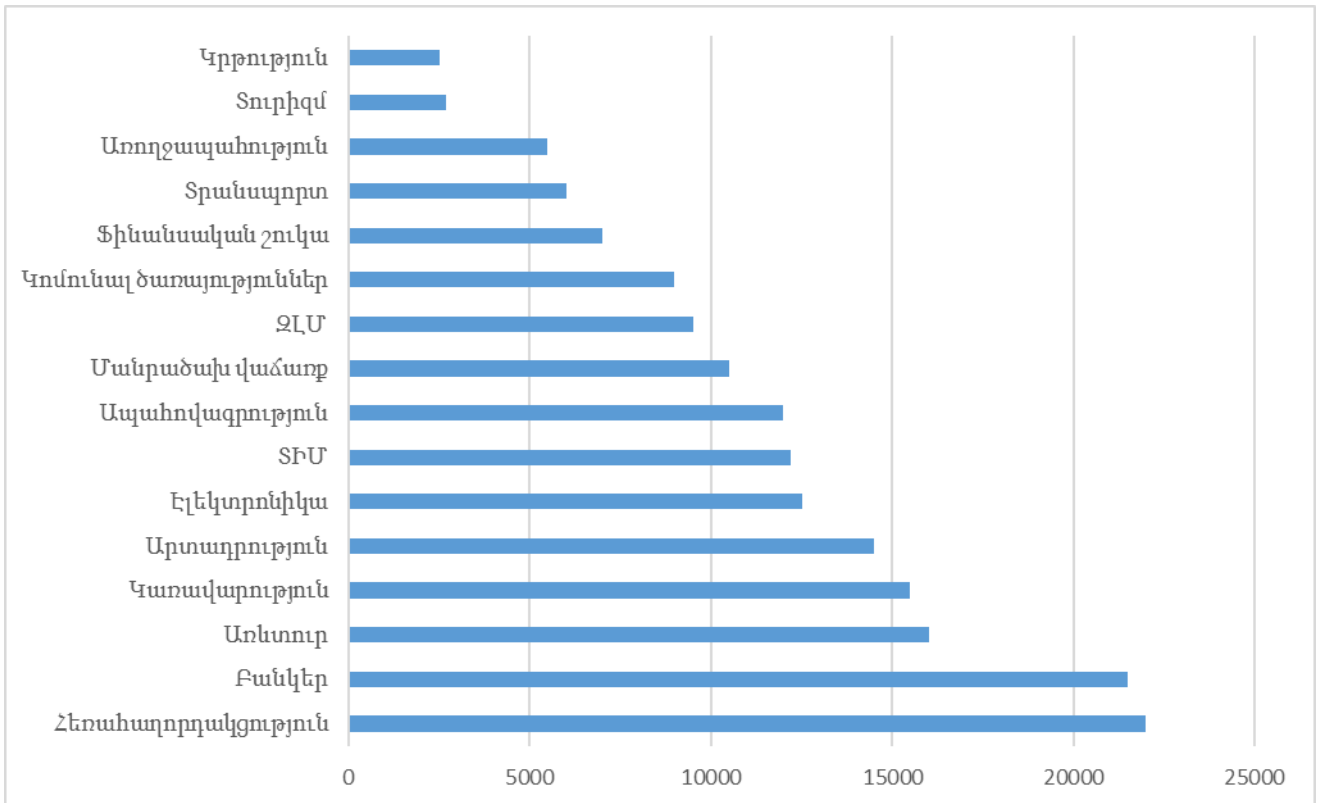
Մեծ տվյալների կիրառության դասակարգման մեկ այլ մոտեցում է դասակարգումն ըստ այն կիրառող կազմակերպության գործունեության բնույթի և տնտեսության ճյուղային պատկանելիության: IBM ընկերությունը յուրաքանչյուր տարի հրապարակում է հաշվետվություն, որտեղ նշվում է ինչպես ընկերության ֆինանսական ցուցանիշները, այնպես էլ այդ տարվա ընթացքում ունեցած մեծագույն ձեռքբերումներն ու հիմնական հետազոտությունների արդյունքները: Համաձայն հաշվետվության՝ IBM ընկերությունը 2015թ. հետազոտություն է անցկացրել պարզելու, թե տնտեսության որ ոլորտներում են ավելի շատ ներկայումս օգտագործվում ՄՏ-ն: Հետազոտության արդյունքները ներկայացված են Գծապատկեր 6-ում, որի օրդինատների առանցքը պարունակում է ճյուղի անվանումը, իսկ աքսիսներից՝ այդ տարվա ընթացքում տվյալ ճյուղում ներդրումների ծավալը՝ ՄՏ ոլորտում (մլրդ ԱՄՆ դոլար):

Հետաքրքրաշարժ է այն փաստը, որ հետազոտված կազմակերպությունների միասնական ներդրումները կազմում են 179 մլրդ ԱՄՆ դոլար, որից 102 մլրդ ԱՄՆ դոլարը բաժին է ընկնում առաջին 6 ճյուղերին (պատկերված գծապատկերի ստորին հատվածում): Այսպիսով, հետազոտված ճյուղերում կատարված ընդամենը ներդրումների ավելի քան կեսը (57%) բաժին են ընկնում

հեռահաղորդակցության, բանկային, առևտրի, կառավարության, արտադրության և էլեկտրոնիկայի ոլորտներին:

Ինչ վերաբերվում է վերջին տեղը զբաղեցրած կրթության ոլորտին, ապա հարկավոր է նշել, որ այստեղ խոսքը գնում է որոշումների կայացման գործընթացում անմիջապես ՄՏ-ի վերլուծության կիրառման, այլ ոչ թե ՄՏ-ի վերլուծությանը առնչվող կրթական ծրագրերի մշակման և ներդրման մասին:

Գործարար կառավարման գործընթացում Մեծ տվյալների վերլուծությունն կիրառող տնտեսության հիմնական ճյուղերը³²



ՄՏ-ի հավաքագրման և պահպանման համար ամպային լուծումներ առաջարկող խոշորագույն ընկերություններից մեկը՝ Cloudera-ն /Բլաուդերա/, տնտեսության տարբեր ոլորտներում գործող և որոշումների կայացման գործընթացում ՄՏ օգտագործող ընկերությունների շրջանում ընտրանքային հարցում է անցկացրել և պարզել, թե ինչ խնդիրների լուծման նպատակով են օգտագործվում այդ տվյալները³³: Հետազոտության արդյունքում ստացված ձեռներեցական խնդիրների բազմությունը դասակարգվել է 2 խմբի՝ տնտեսության 8 տարբեր ոլորտներից յուրաքանչյուրի համար: Այդ ուսակում 3-ում ներկայացված են ոլորտներից յուրաքանչյուրում ՄՏ-ի օգտագործմամբ լուծված ամենահաճախ հանդիպող ձեռներեցական խնդիրները:

³² Ռազմավարական լուծումների գնահատականներ, IBM ընկերության տարեկան հաշվետվություն, 2015, <https://www.ibm.com/annualreport/2015/bin/assets/IBM-Annual-Report-2015.pdf>

³³ Տե՛ս Awadallah A., “How Hadoop Revolutionized Business Intelligence and Advanced Data Analytics”, November 2011, <http://assets.en.oreilly.com/1/event/63/How%20Hadoop%20is%20Revolutionizing%20Business%20Intelligence%20and%20Advanced%20Data%20Analytics%20Presentation.pdf>

**Մեծ տվյալների օգտագործմամբ լուծված ամենահաճախահանդիսարդ
խնդիրները³⁴**

Ոլորտ	Խնդիր 1	Խնդիր 2
Էլեկտրոնային կոմերցիա	Սոցիալական ցանցերի վերլուծություն	Այցելուի կողմից կայքում կատարվող գործողությունների ուսումնասիրություն
ՉԼՄ	Լուրերի օպտիմիզացում	Օգտատերերի ընդգրկում
Հեռահաղորդակցություն	Ցանցերի վերլուծություն	Տեղանքային արդյունավետ սակագների որոշում
Մանրածախ վաճառք	Հավատարմության գնահատում	Մատակարարման շղթաների կառավարում
Ֆինանսներ	Կեղծումների հայտնաբերում	Գործարքների հսկողություն
Պեստություն	Կառույցների վերլուծություն	Կիրառանվտանգություն
Կենսաբանություն	Հաջորդականության վերլուծություն	Գենոմի քարտեզագրում
Արտադրություն	Որակի կառավարում	Լոգիստիկայի օպտիմիզացում

Մեծ տվյալների կիրառությունն դասակարգումն րստիավարարման աղբյուրների

Մեծ տվյալների 3 հիմնական բնութագրիչներից մեկը բազմազանությունն է, որն ընդգծում է վերջիններիս հավաքագրման աղբյուրների տարատեսակայնությունն ու

³⁴ Տե՛ս նույն տեղում, էջ 15

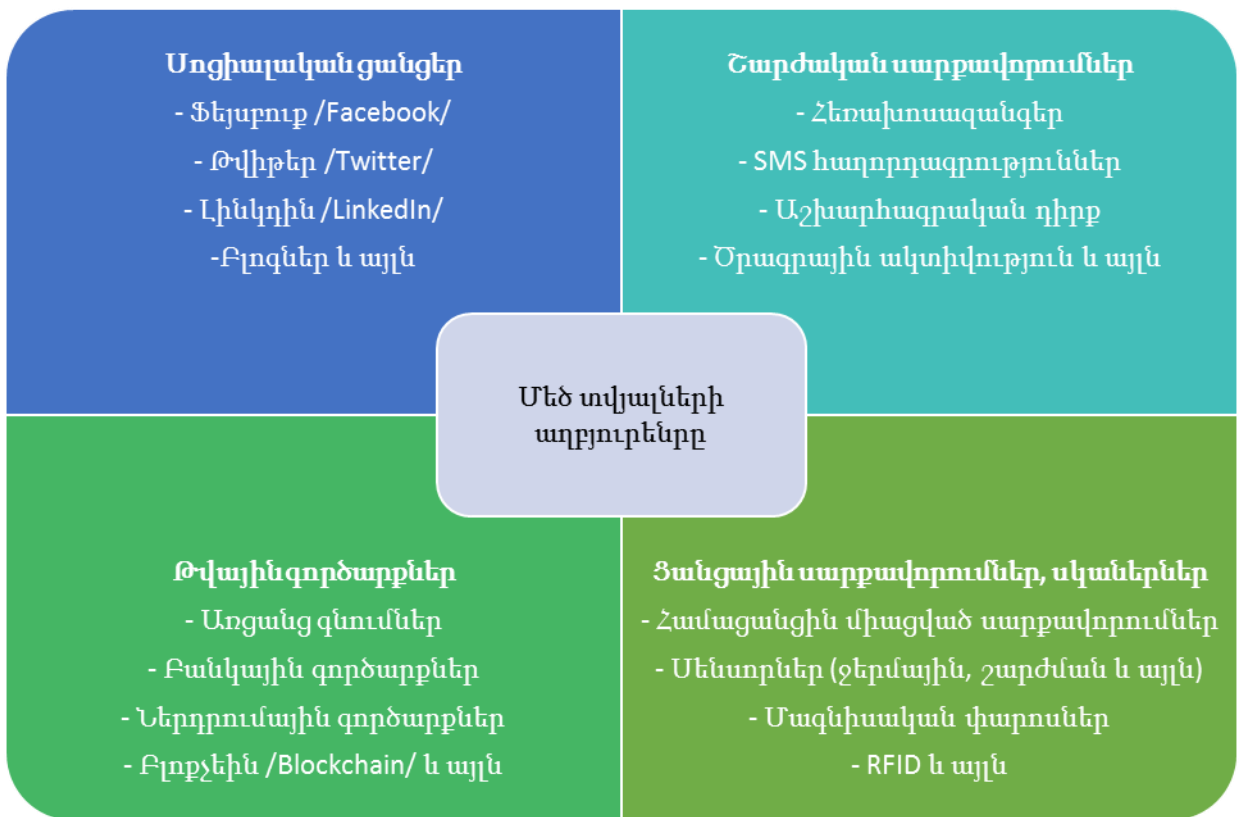
հավաքագրվող տվյալների ամենատարբեր ձևաչափերը: Ըստ աղբյուրի դասակարգման տեսանկյունից կարելի է առանձնացնել հետևյալ խմբերը.

- I. Սոցիալական ցանցեր,
- II. Շարժական սարքավորումներ,
- III. Թվային գործարքներ,
- IV. Ցանցային սարքավորումներ, սկաներներ:

Գծապատկեր 8.

Մեծ տվյալների հիմնական աղբյուրները

**կազմվել է մեր կողմից*



Սոցիալական ցանցեր

Համաձայն Գլոբալ վեբ ինդեքսի տվյալների՝ միջին վիճակագրական ամերիկացին օրական 4,3 ժամ է անցկացնում սոցիալական կայքերում (որից 30%-ը՝ Ֆեյսբուքում)³⁵: Նույն ցուցանիշի համաշխարհային միջինը օրական կազմում է 3,6 ժամ: Ֆեյսբուքյան պաշտոնական վիճակագրության համաձայն՝ մեկ րոպեի ընթացքում այդ սոցիալական կայքում հավանում են մոտ

³⁵ [GlobalWebIndex, Q3 Report, 2016](#)

3125000 գրառում, վերբեռնում են 243055 նկար և այլն: Նման մեծածավալ տվյալների ստեղծման և հավաքագրման վիճակագրությունն էլ հենց պատճառ է հանդիսանում այն հանգամանքի, որ սոցիալական ցանցերը առաջին տեղում են մեր՝ ըստ աղբյուրների դասակարգման աղյուսակում: Իրական ժամանակահատվածում մարդկանց կարծիքների, նախապատվությունների և մեկնաբանությունների մասին հավաքագրվող այս տվյալները հիմք են հանդիսանում մի շարք կազմակերպությունների կողմից գովազդային արշավի պլանավորման, օպտիմալացման և թիրախավորման համար: Սոցիալական ցանցերի օգտատերերի կողմից տարատեսակ ապրանքների և ծառայությունների մասին կատարվող գրառումները կարևոր տեղեկատվության աղբյուր են հանդիսանում ընկերությունների համար ոչ միայն սեփական արտադրանքի որակի կառավարման նպատակով, այլև մրցակիցների ուժեղ և թուլլ կողմերի վերլուծության տեսանկյունից:

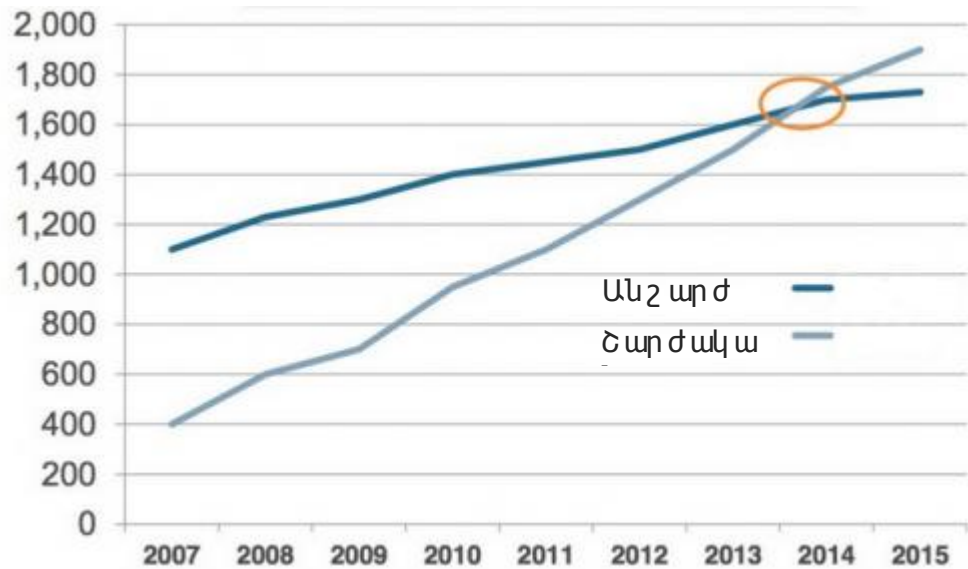
Շարժական սարքավորումներ

Մարքեթինգային վիճակագրություն և հավաքագրող ինտերնետային կայք Statistica-ի համաձայն՝ աշխարհում շարժական հեռախոսակապից օգտվողների թիվը 2016 թվականին գնահատվում է 4,6 մլրդ, այն դեպքում, երբ օգտագործվող հեռախոսների քանակը արդեն իսկ գերազանցում է աշխարհի բնակչության թիվը³⁶:

Գծապատկեր 9.

Շարժական և անշարժ միջոցներից ինտերնետի օգտագործումը (աղբյուրը՝ www.statistica.com)³⁷

³⁶ <https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/>, (09/11/2016)
³⁷ Տե՛ս նոյն տեղում



Ավելի ին՝ աշխարհահռչակ Մորգան Սթենլիի միջազգային ֆինասական ծառայություններ մատուցող կորպորացիայի հետազոտական կենտրոնի տվյալներով աշխարհում ներկայումս կա ավելի շատ մարդ, ով ինտերնետից օգտվում է հեռախոսի միջոցով, քան թե համակարգչի³⁸:

Նման պարագայում, անկասկած, կարևորություն է ստանում շարժական հեռախոսակապի օգտագործման հետևանքով հավաքագրվող տվյալների բազմությունը: Հատկանշական է, որ շարժական հեռախոսակապի միջոցները միակ շարժական սարքավորումները չեն, որ կարող են տվյալների փոխանցել օգտագործողի մասին: Նման ֆունկցիա կարող են իրականացնել նաև Գլոբալ տեղորոշման համակարգերը (անգլերեն՝ GPS), էլեկտրոնային ժամացույցները, սարտային էլեկտրոնային թանոցները և այլն: Այդ բոլոր սարքավորումները տեղեկատվություն են փոխանցում սարքավորումն արտադրող կամ վերջինիս լրացնող միջոցներ արտադրող կազմակերպություններին՝ օգտագործողի մասին: Օրինակ, սարտային էլեկտրոնային թանոցները տվյալներ են հավաքագրում այն կրողի կենսաֆիզիկական վիճակի (օր.՝ սրտի աշխատանքի), վերջինիս փոփոխությունների մասին: Այս տվյալները իրական ժամանակահատվածում վերլուծվում են՝ օգտագործողին ճիշտ խորհուրդներ տալու նպատակով (օր.՝ քայլել ավելի արագ կամ դանդաղ և այլն):

³⁸ Տե՛ս [Morgan Stanley Research Center, "Mobile Internet Report 2016"](#)

Ի տարբերություն վերոնշյալ թևերի, երբ տվյալները ստանում և վերլուծում են անմիջապես տվյալ ապրանքի արտադրող ընկերությունը, շարժական հեռախոսակապի ծառայություններ մատուցողները օգտատիրոջ աշխարհագրական դիրքի և այլ տվյալները ստանում են լրացնող ապրանքի՝ հեռախոսի միջոցով: Աշխարհագրական դիրքի մասին տվյալները շատ հաճախ օգտագործվում են ճիշտ գնային քաղաքականության մշակելու նպատակով: Սակայն շարժական հեռախոսակապի ծառայություններ մատուցողները, երբեմն, գտնում են այդ տվյալների այլ ընտրանքային օգտագործման հնարավորություններ: Օր.՝ ԱՄՆ մի քանի խոշոր ընկերություններ հիմնել են դուստր տուրիստական կազմակերպություններ, որոնք օգտագործում են շարժական հեռախոսակապի մասին տվյալները ամենաայցելվող տեղանքների մասին տեղեկատվություն ստանալու նպատակով և համապատասխանաբար օպտիմալացնում իրենց տուրիստական փաթեթները*:

Թվային գործարքներ

2015 թվականին իրականացվել է ավելի քան 4263 մլրդ թվային գործարք³⁹: Այս թիվը ներառում է ինչպես էլեկտրոնային կոմերցիոն գործարքները, այնպես էլ բանկային, ներդրումային և այլ տեսակի ֆինանսական գործարքները: Նման գործարքների թվային լիները թույլ է տալիս հեշտորեն և իրական ժամանակահատվածում հավաքագրել գործարքի հետ փոխկապակցված տվյալները (սեզոնայնություն, միտում, գործարքային հաջորդականություն և հաճախականություն և այլն) և օգտագործել նկարագրողական, կանխատեսողական և կարգադրողական վերլուծության նպատակով:

Ցանցային սարքավորումներ, սկաներներ

* դիտարկվել են Huawei, BBVA, BCG և մի շարք այլ կազմակերպությունների օրինակներ

** դիտարկվել են Amazon, Walmart և մի շարք այլ կազմակերպությունների օրինակներ

³⁹ Տե՛ս [Global Digital Payment Transactions Set to Reach 10 Percent Growth for First Time Finds World Payments Report 2016](https://www.capgemini.com/sites/default/files/global_digital_payment_transactions_set_to_reach_10_percent_growth_for_first_time_finds_world_payments_report_2016.pdf), https://www.capgemini.com/sites/default/files/global_digital_payment_transactions_set_to_reach_10_percent_growth_for_first_time_finds_world_payments_report_2016.pdf

Ոչ համացանցային աշխարհում ՄՏ-ի հավաքագրման, թերևս, խոշորագույն աղ-բյուրը ցանցային, հաճախականային, սենսորային, սկաներային սարքավորումներն են: Մի շարք սուպերմարկետներ օգտագործում են Ռադիոհաճախականային նույնակա-նացման չիպեր (անգլերեն՝ RFID), որպեսզի հետևեն սայլակի տեղաշարժը սուպերմարկետի ներսում*: Այդ տվյալները հետագայում (գրեթե իրական ժամանակահատվածում) վերլուծվում են, և արդյունքում ստացվում է սուպերմարկետի գուճավորված քարտեզը՝ ավելի մոզ նշելով այն կետերը, որտեղ հաճախորդները (և, հետևաբար, նրանց սայլակները) ավելի երկար են կանգ առել: Դա թույլ է տալիս որոշել այն օպտիմալ տարածքը, որտեղ կարելի է տեղադրել օրինակ գովազդի առավել կարիք ունեցող ապրանքը:

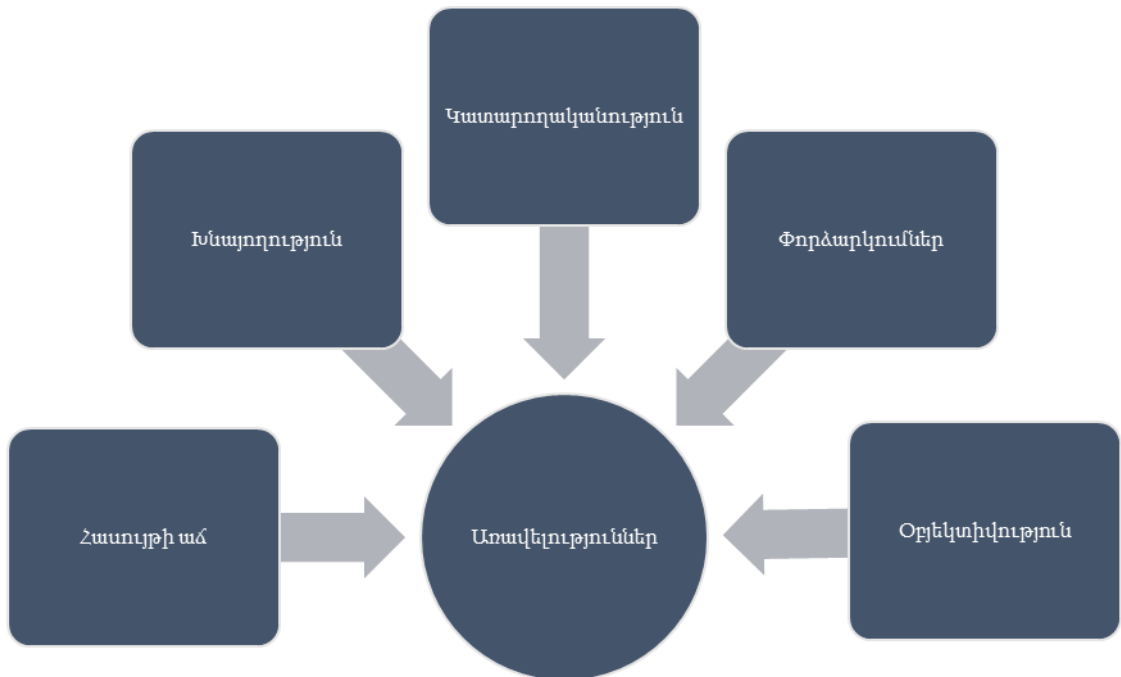
Առավել ություններն ու թերությունները

Վերոնշյալ կիրառություններն իրենց տարածումն են գտել մեկ պարզ պատճառով. ՄՏ-ի տնտեսագիտական վերլուծությունը կարող է գործարար առավել ությունների հիմք հանսիանալ: Ստորև ներկայացված և մանրամասնորեն նկարագրված են 5 առավել ություններ, որոնք հնարավոր է ձեռք բերել ՄՏ-ի տնտեսագիտական վերլուծության միջոցով: Այդ առավել ությունները հիմնված են մեր կողմից հարցված կազմակերպություններից ստացված որակական տվյալների, ինչպես նաև ուսումնասիրված միջազգային փորձի վրա:

Գծապատկեր 10.1

Մեծ տվյալների տնտեսագիտական վերլուծության առավել ությունները

Հասույթի աճ: Հասույթի աճը կարելի է դիտարկել երկու տեսանկյունից: Մի կողմից ընկերությունները, վերլուծելով իրենց հաճախորդներին, աշխատակիցներին, գործառույթները և ֆինանսները կարող են ապահովել հասույթի աճ՝ տվյալ անենք որոշումների կայացման միջոցով: Մյուս կողմից, ՄՏ-ի



վերլուծությունը հնարավորություն է տալիս պարզել, թե որ բաղադրիչի ներգործությունն է պատասխանատու հասույթի աճի համար: Ի տարբերություն «ավանդական» տվյալների՝ ՄՏ-ն հնարավորություն է տալիս վերլուծվող սուբյեկտին դիտարկել 360° մոտեցմամբ, այն է՝ դիտարկել հետազոտվող սուբյեկտիվ տարբեր կողմերից՝ հիմնվելով ՄՏ-ի բազմազանության վրա: Շարժունությունն իր հերթին ապահովում է որոշումների կայացման համար անհրաժեշտ տեղեկատվության ճիշտ ժամանակին ստացումը, որը մրցակցային առավելության աղբյուր կարող է հանդիսանալ:

Ինսայդություն: Գործարար կառավարման ոլորտում կայացվում են ինչպես ճիշտ այնպես էլ սխալ որոշումներ: Այդուհանդերձ, ընկերությունները կարող են կորցնել մեծ գումարներ սխալ որոշումների կայացման հետևանքով: ՄՏ-ի գոյությունը և նրանց վերլուծությունը թույլ է տալիս նվազեցնել անորոշությունը և

հնարավորինս ճշգրիտ կանխատեսել հետևանքները՝ մինչ որոշման կայացումը: Բանն այն է, որ այստեղ գործում է Մեծ թվերի օրենքը, և ընտրանքային շեղվածության հետ կապված խնդիրները, որոնք հնարավոր են առաջանալ «ավանդական» տվյալների դեպքում, այլևս սպառնալիք չեն հանդիսանում:

Կատարողականություն: Արտաքին միջավայրի անորոշության մեջ գործող ընկերությունները միշտ չէ, որ գնում են ռիսկի և շուկա են դուրս բերում լրիվ նոր արտադրանք: Պատճառն այն է, որ տնտեսվարող սուբյեկտները ավելի շատ օգտակարություն են կորցնում գումարի կորստից, քան միևնույն չափի գումարի ձեռքբերումից են ստանում⁴⁰: ՄՏ-ի վերլուծության միջոցով հնարավոր է դառնում նվազեցնել գործող անորոշությունը և բարձրացնել կատարողականությունը, որը հնարավորություն է տալիս արտադրվող ապրանքների և մատուցվող ծառայությունների բազմազանության ընդլայնման:

Փորձարկումներ: Ծառ ժամանակ ակնհայտ չէ թե՛ առկա այլ ընտրանքներից որն է կազմակերպության համար հանդիսանում լավագույնը: Վառ օրինակ կարող են հանդիսանալ մարքեթինգային քաղաքականության մշակումը կամ վեբկայքի պատրաստումը: Ձեռներեցի առջև են դրված բազում տարատեսակ մարքեթինգային հնարավորություններ (արտաքին գովազդ, հեռուստագովազդ, գովազդ սոցիալական ցանցերից մեկում կամ մի քանիսում և այլն), զանազան և տարբեր ճաշակների համապատասխան կայքեր և այլն (պարզ և հարմարավետ, բարդ բայց բազմաֆունկցիոնալ և այլն): ՄՏ-ի հավաքագրումը թույլ է տալիս փորձարկել մի քանի այլ ընտրանք և պարզել, թե որն է կազմակերպության համար ավելի ցանկալի իրավիճակի հանգեցնում: Օրինակ, ընկերությունները կարող են հրապարակել նույն վեբկայքի 2 տարբերակ և հավաքագրել այցելուների տվյալները (այս մեթոդը հայտնի է որպես A/B testing, որտեղ A-ը և B-ն կայքի համապատասխանաբար առաջին և երկրորդ տարբերակներն են): Ընկերությունը, որը վաճառում է միանգամյա օգտագործման ապրանք կամ ծառայություն, կընտրի կայքէջի այն տարբերակը, որը առավել ազույն այցելություն կապահովի: Սակայն

⁴⁰ Տե՛ս Arrow, K. J. (1965). "Aspects of the Theory of Risk Bearing".

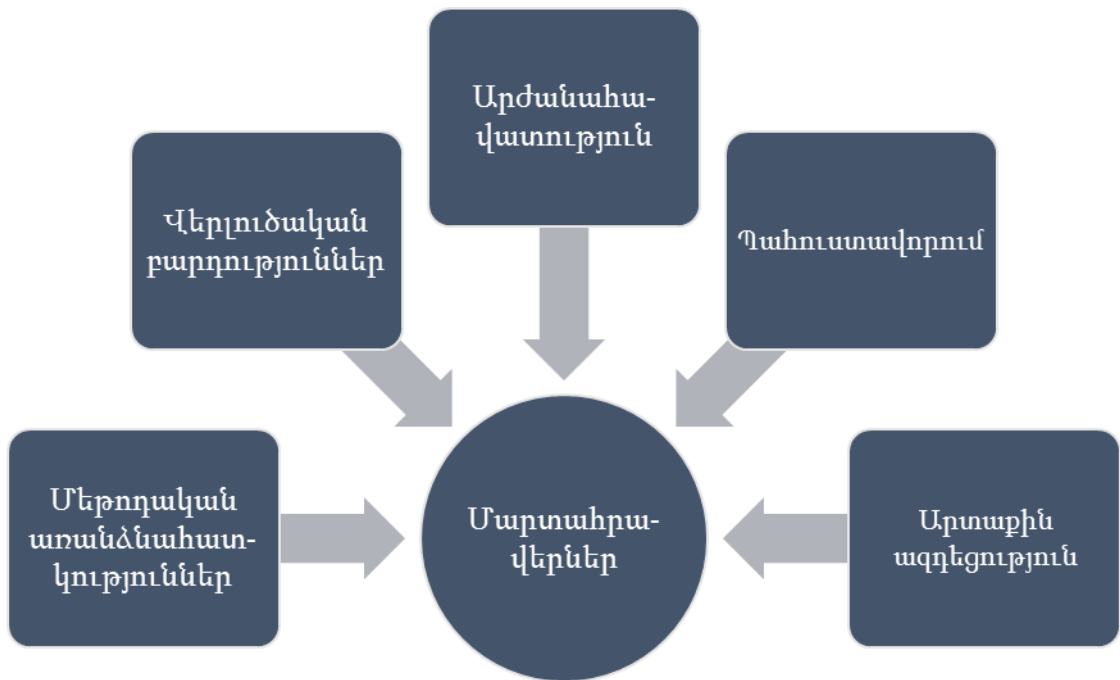
այն ընկերությունը, որի նպատակն է հաճախորդների հետ երկարաժամկետ, հստակ և շահավետ հարաբերությունների ապահովը, կընտրի վեբկայքի այն տարբերակը, որում այցելուներն ավելի երկար ակտիվություն են ցուցաբերել:

Օբյեկտիվություն: Ցանկացած որոշում, որը կայացվում է այս կամ այն անձի կողմից պարունակում է որոշակի սուբյեկտիվություն՝ հիմնվելով տվյալ անհատի գիտելիքների, փորձի և/կամ տվյալ հարցի մոտեցման վրա: Պատմության ընթացքում շատ են եղել դեպքեր, երբ մենեջերի փոփոխության հետևանքով ընկերության ոչ միայն աշխատաճը, այլև ընդունած ռազմավարությունը շրջադարձ է ապրել: Հաջողության հասնելու համար ընկերությունները ցանկանում են ունենալ ոչ թե անձի վրա հիմնված սուբյեկտիվ ռազմավարություն, այլ տվյալ ահենք օբյեկտիվ և չափելի նպատակներ և վերջիններիս հասնելու պլանավորված ուղիներ: Վերլուծությունը հնարավորություն է տալիս հիմնավորել կայացրած որոշումները կազմակերպության ներքին և արտաքին միջավայրից ելնելով:

Վերոնշյալ առավելություններն էլ հենց պայմանավորում են ՄՏ-ի այս հատվածում նշված բազմազան ու բազմաբնույթ կիրառությունները գործարար կառավարման տարբեր ոլորտներում և ձեռներեցական տարատեսակ խնդիրների վերաբերյալ որոշումների կայացման գործընթացում: Այնուամենայնիվ, հանդիսանալով երիտասարդ և հետազոտության ցածր աստիճան ունեցող ոլորտ, ՄՏ-ի տնտեսագիտական վերլուծությունը դեռևս ունի մի շարք անպատասխան մարտահրավերներ: Ստորև ներկայացված են այն 5 առանձնահատկությունները, որոնք խոչընդոտեն ծնում ՄՏ-ի տնտեսագիտական վերլուծության գործընթացում:

Մեթոդական առանձնահատկություններ: ՄՏ-ի ծավալը, շարժունությունն ու բազմազանությունը խոչընդոտներ են ծնում վերջինիս տնտեսագիտական վերլուծության գործընթացում: Այն տնտեսագիտական, վիճակագրական և մաթեմատիկական մոդելներն ու մեթոդները, որոնք սովորաբար կիրառվում են «ավանդական» տվյալների վերլուծության պարագայում այլևս կիրառելի չեն ՄՏ-ի համար: Այդ իսկ պատճառով, որոշումների կայացման

գործընթացում US-ի կիրառության համար կազմակերպությունները ստիպված են աշխատանքի ընդունել մասնագետների, որոնք տիրապետում են այդ նոր մոդելներին և մեթոդներին: Վերջինս ընկերություններին հավելյալ ներդրումներ է պահանջում, որոնց իրականացումը պայմանավորված է US-ի տնտեսագիտական վերլուծության կիրառության համեմատական արդյունավետությամբ: Նշված մեթոդական առանձնահատկություններն ավելի մանրամասնորեն ներկայացված



են գլուխ 2.1-ում, իսկ դրանց համեմատական արդյունավետությունը «ավանդական» տվյալների վերլուծության նկատմամբ՝ գլուխ 3.1-ում:

Գծապատկեր 10.2

Մեծ տվյալների տնտեսագիտական վերլուծության մարտահրավերները

Վերլուծական բարդություններ: Վերոնշյալ մեթոդական առանձնահատկությունների հետևանքով US-ի տնտեսագիտական վերլուծությունը դասական տվյալների վերլուծության փաթեթների և ծրագրերի միջոցով իրականացնելն անհնար է դառնում: Պատճառն այն է, որ այդ նոր մոդելներն ու մեթոդները համատեղելի չեն Excel, SPSS, Stata, Eviews կամ նման այլ ծրագրային փաթեթների հետ: Այդ իսկ պատճառով, US-ի վերլուծությունը

պահանջում է առավել խորը համակարգչային գիտելիքներ, որոնք թույլ կտան վերլուծաբաններին կիրառել առավել բարդ գործիքներ՝ վերլուծության գործընթացում: ՀՀ աշխատանքային US-ի վերլուծաբանների կողմից այդ գործիքներին տիրապետելը, նրանց հմտություններն ու այլ մասնագիտական և ժողովրդագրական տվյալները վերլուծվում են գլոբալ 2.2-ում:

Արժանահավատություն: Գլոբալ 1.1-ում ներկայացված US-ի սահմանումներից մեկում դասական 3 բնութագրիչներին ավելացվում էր 4-րդը՝ արժանահավատությունը: Բանն այն է, որ US-ի ծավալը, շարժունությունը և բազմազանությունը մեծացնում են անարժանահավատ տվյալների հավաքագրելու ռիսկերը: Արդյունքում անարժանահավատ աղբյուրներից տվյալների հավաքագրումը, կարող է մեծ դրամական կորուստների հանգեցնել: Մի կողմից US-ի հավաքագրումն ինքնին ծախսատար և մեծ ներդրումներ պահանջող գործընթաց է, ինչը նշանակում է, որ անարժանահավատության դեպքում այդ գումարները ջուրն են ընկնում: Մյուս կողմից ոչ ճշգրիտ տվյալների հիման վրա կայացվող որոշումները կարող են կրիտիկական ազդեցություն ունենալ կազմակերպության գործունեության հետագա ընթացքի վրա:

Պահուստավորում: US-ի ծավալը անհնարին է դարձնում համակարգչի կամ կոշտ սկավառակի մեջ նրանց պահպանումը: Այդ խնդիրը, ինչպես արդեն նշվել է, լուծվում է ամպային տեխնոլոգիաների միջոցով: Այնուամենայնիվ, վերջինիս հետևանքով ծնվում են պահուստավորման խնդիրներ: Ցանկացած կազմակերպության համար տվյալները հանդիսանում են չափազանց կարևոր և, շատ դեպքերում, կոնֆիդենցիալ ակտիվ: Այդ իսկ պատճառով անհրաժեշտություն է առաջանում ռիսկերի նազեցման համար պահպանել այդ ակտիվը ոչ թե մեկ, այլ մի քանի տարբեր վայրերում: US-ի առեւտրի ծավալը թույլ է չի տալ իսապահովության համար երկրորդային պահուստավորում իրականացնել, քանի որ վերջինս ուղղակիորեն կրկնապատկում է ամպային համակարգում տվյալների պահպանման ծախսերը:

Արտաքին ազդեցություն: Չնայած գործարար որոշումների կայացման գործընթացում US-ի տնտեսագիտական վերլուծության

կիրառու թյան դրական ազդեցությանը՝ վերջինս կարող է բացասական արտաքին ազդեցություն ունենալ հասարակության վրա: Պատճառն այն է, որ ՄՏ-ի գոյությունը կարող է անհատական գաղտնիության խնդիրներ առաջացնել: Կազմակերպությունները հավաքագրում են տվյալներ օգտատերերի աշխարհագրական տեղադրության, սոցիալական ցանցերում մեկնաբանությունների, այցելած կյանքերի հաջորդականության և այլ քայլերի մասին: Նման մանրամասն տվյալները ցանկության դեպքում կարող են օգտագործվել ինչպես հաճախորդի օգտին՝ բավարարվածությունը բարձրացնելով, այնպես էլ նրանց դեմ՝ խախտելով գաղտնիության սկզբունքները:

Վերոնշյալ պայմանները մարտահրավերներ են ծնում ՄՏ-ի տնտեսագիտական վերլուծության համար, որոնք առավել մանրամասն ուսումնասիրվել են հաջորդ գլուխներում: Հետազոտության հաջորդ հատվածը նվիրված է սույն ոլորտի հիմնական դերակատարներին, ինչպես նաև ներկայացնում է ոլորտում նկատվող և սպասվող զարգացման միտումները:

1.3 Ոլորտի հիմնական «խաղացողները» և զարգացման միտումները

Մեծ տվյալների ոլորտի զարգացումը հնարավորություն է տալիս ընկերություններին այն կիրառել տարբեր նպատակներով, որոնք կարելի է դասակարգել հետևյալ 3 խմբերի.

- Շարունակել իրենց ներկայիս գործունեությունը՝ գրանցելով ավելի և ավարդյունքներ,
- Լայնացնել իրենց ներկայիս գործունեության շրջանակները,
- Արտադրել նոր ապրանք կամ մատուցել նոր ծառայություն:

Ելնելով վերոնշյալ երեք կետերից՝ գնալով ավելի ու ավելի է տարածվում ՄՏ-ի կիրառությունը տարատեսակ կազմակերպությունների և ընկերությունների կողմից: Ներկայումս, շուկայում առանձնանում են որոշ ընկերություններ, որոնց կարելի է համարել ՄՏ-ի ոլորտի հիմնական «խաղացողները»: Այդ ընկերություններից է

աշխարհահռչակ IBM կազմակերպությունը, որը ոչ միայն դասական տվյալների վերլուծության SPSS ծրագրային փաթեթի սեփականատերն է, այլև հեղինակ է մի շարք ՄՏ ոլորտի ուսումնասիրությունների, որոնք արդեն իսկ հիմք են հանդիսացել այս հետազոտության համար: IBM-ը հանդիսանում է Watson արհեստական ինտելեկտի ստեղծողը, որը տարիներ առաջ հաղթեց հռչակավոր շախմատի ստոպարի Կասպարովին: Այդ համակարգն իրականում միայն շախմատ խաղալու համար չէ նախատեսված: Watson արհեստական ինտելեկտը ակտիվ տվյալներ է հավաքագրում համացանցից և մի շարք կազմակերպությունների կողմից հանդիսանում հաճախորդների և աշխատակիցների վերլուծման հիմնական միջոց: Այն կարողանում է ավտոմատաբար դասակարգել հաճախորդներին տարբեր խմբերի և ձեռներեցին թույլ տալ յուրաքանչյուր խմբի վերաբերյալ կայացնել անհատականացված որոշում: Հատկանշական է, որ նման խոշոր կազմակերպությունը ունի նաև իր ներկայացուցչությունը Հայաստանում: IBM տեղեկատվական տեխնոլոգիաների կենտրոնը գրեթե մեկ տարի է ինչ դասավանդում է տվյալ ազիտություն՝ ուսանողներին ներկայացնելով IBM Watson-ի հնարավորությունները հաճախորդների վերլուծության գործընթացում: Հիմնական դասավանդվող առարկաներն են վիճակագրությունը, համակարգչազիտությունը և գործարարությունը:

IBM կազմակերպության, ինչպես նաև վերոնշյալ տվյալներ վերլուծող արհեստական ինտելեկտի հիմնական մրցակիցը թերևս Microsoft /Մայքրոսոֆթ/ կազմակերպությունն է իր Cortana /Կորտանա/ կոչվող արհեստական ինտելեկտով: Կցելով Cortana-ն վեբկայքին կամ հեռախոսային ծրագրին՝ այն կարող է պարզել, մարդկանց գրավոր արտահայտված կարիքների դրական կամ բացասական լինելը և համապատասխան խորհուրդներ տա ձեռներեցին արտադրության կամ ծառայությունների մատուցման գործընթացում անհրաժեշտ բարեփոխումների մասին: Չնայած այն հանգամանքին, որ ՀՀ-ում կա Microsoft կազմակերպության մասնաճյուղ՝ նրանք դասավանդում են միայն ծրագրավորում և չեն զբաղվում տվյալ ազիտության ուսուցմամբ:

ՄՏ-ի դասակարգման սահմանման հեղինակ Գարթները կազմակերպությունը յուրաքանչյուր տարի իրականացնում է հարցում՝ ՄՏ-ի վերլուծության ոլորտում ներդրումային քաղաքականության մասին տեղեկատվություն հավաքելու նպատակով: Հարցմանը մշտապես մասնակցում են տնտեսության տարբեր ոլորտների ներկայացնող կազմակերպություններ: Համաձայն 2015 թվականի տվյալների՝ հարցված կազմակերպությունների 75%-ը արդեն իսկ կատարել է կամ պլանավորում է կատարել ներդրում ՄՏ-ի վերլուծության ոլորտում⁴¹: Սույն հարցման տվյալների համաձայն՝ ներդրումները ներդրումներն իրականացվելու են հետևյալ նպատակներով.

- Բարձրացնել հաճախորդների բավարարվածությունը (64%),
- Իրականացնել անհատականացված մարքեթինգ (47%),
- Ապահովել ավելի որակյալ անվտանգություն (23%):

Հատկանշական է, որ հարցված կազմակերպությունների 25%-ը, որոնք դեռևս պլաններ չունեն ՄՏ-ի վերլուծության ոլորտում կատարել ներդրումներ, այդ որոշման հիմքում ընկած հիմնական պատճառն են համարել ներդրումային արդյունավետության հաշվարկման մեթոդի բացակայությունը: Ինչպես այլ ոչ նյութական ակտիվների պարագայում, այնպես էլ ՄՏ-ի վերլուծության ոլորտում կատարվող ներդրումները ուղղակիորեն չափելը հնարավոր չէ: Վերջինիս ազդեցությունը կազմակերպության արդատարողականության, ֆինանսական և այլ ցուցանիշների վրա անուղղակի է և միջնորդված է մեկ կամ մի քանի այլ գործոններով: Այդ իսկ պատճառով, մինչ օրս գոյություն չունի ՄՏ-ի ոլորտում իրականացվող ներդրումների արդյունավետության հաշվարկման համար օգտագործվող համընդհանուր մեթոդ: Արդյունքում 43%-ը այն կազմակերպությունների, որոնք կատարել են և 38%-ը նրանց, ովքեր պլանավորում են կատարել ներդրումներ այդ ոլորտում ընդգծում են, որ չգիտեն արդյո՞ք ներդրումները արդյունավետ են/կլինեն թե ոչ: Այս փաստն ընդգծում է ՄՏ-ի վերլուծության ոլորտում

⁴¹ Տե՛ս [Gartner, “Gartner Symposium, survey summary” 2016, http://www.gartner.com/newsroom/id/3130817](http://www.gartner.com/newsroom/id/3130817)

ներդրումային արդյունավետության հաշվարկման համընդհանուր մեթոդոլոգիայի ստեղծման անհրաժեշտությունը:

Ի տարբերություն Գարթների կողմից 2015 թվականին հրատարակված հարցման 2014 թվականին ՄՏ-ի վերլուծության ոլորտում ներդրումային պլաններ չունեցող կազմակերպությունների նման որոշման հիմնական պատճառը ձեռներեցի կողմից ՄՏ վերլուծությանը ռազմավարական կարևորություն չտալն է: Այդ ցուցանիշը մեկ տարվա ընթացքում ընկել է առաջինից 5-րդ հորիզոնականը՝ ընդգծելով այն հանգամանքը, որ ձեռներեցները գնալով ավելի ու ավելի են կարևորում ՄՏ-ի ռազմավարական նշանակությունը:

ՄՏ-ի ոլորտում մատուցված ծառայությունների և արտադրված արտադրանքի դիմաց ստացված հասույթը 2015 թվականին Միջազգային տվյալների կորպորացիայի կողմից (անգլերեն՝ IDC) գնահատվում է 122 մլրդ ԱՄՆ դոլար⁴²: Այստեղ առյուծի բաժինը հասնում է ՄՏ-ի վերլուծական ծառայություններ մատուցող կազմակերպություններին: Բանն այն է, որ ոլորտի դեռևս զարգացած չլինելու հետևանքով ՄՏ-ի վերլուծաբաններն այդքան էլ շատ չեն: Արդյունքում ստեղծվել են կազմակերպություններ, որոնք իրենց շուրջն են հավաքել մասնագետներին և վերլուծական ծառայություններ են մատուցում իրենց հաճախորդ հանդիսացող ընկերություններին: **Միջազգային տվյալների կորպորացիայի կանխատեսումների համաձայն՝ մինչև 2019 թվականը վերոնշյալ հասույթը կաճի ավելի քան 50%-ով՝ կազմելով 187 մլրդ ԱՄՆ դոլար:**

Ոլորտում ձևավորվող նման հասույթի և պլանավորվող մեծաքանակ ներդրումների հետևանքով ձևավորվում է այդ ոլորտի մասնագետների մեծ պահանջարկ: Համաձայն համաշխարհային տնտեսագիտական ֆորումի 2015 և 2016 թվականների հաշվետվության՝ տվյալագետը (և ունի նաև ՄՏ-ի վերլուծաբանը) արդեն 2 անգամ անընդմեջ համարվել է տարվա ավագ ունի աշխատանքը⁴³: Այդ ուսակ 4-ում ցուցադրված են համաշխարհային տնտեսագիտական ֆորումի կողմից հրատարակված 20 ավագ ունի աշխատանքները:

⁴² Տե՛ս [International Data Corporation \(IDC\), "Worldwide Semiannual Big Data and Analytics Spending Guide" 2015, http://www.idc.com/getdoc.jsp?containerId=IDC_P33195](http://www.idc.com/getdoc.jsp?containerId=IDC_P33195)

⁴³ Տե՛ս <https://www.weforum.org/agenda/2015/10/worlds-best-job-the-answer-might-surprise-you/>

Ինչպես երևում է աղյուսակից՝ տվյալագետները ոչ միայն զբաղեցնում են առաջին հորիզոնականը, այլև միջին տարեկան աշխատավարձով զգալիորեն գերազանցում են ցուցակում հայտնված մնացած մասնագիտություններին: Չնայած այն հանգամանքին, որ տվյալագիտությունը պահանջում է նաև համակարգչային գիտելիքների իմացությունն՝ ներկայումս ԱՄՆ-ում հաշվարկվում է, որ ծրագրավորողների միայն 4%-ն է տիրապետում տվյալագիտական հմտություններին: Համաձայն McKinsey /Մաքկինսի/ խորհրդատվական ընկերության կանխատեսումների՝ 2018 թվականին միայն ԱՄՆ-ում կլինեն 4 միլիոն աշխատանք կապված ՄՏ-ի վերլուծության հետ⁴⁴: Նման մեծ թվով հաստիքների առկայության պարագայում ակնհայտ է ակնկալել պահանջարկի և առաջարկի ճեղքվածք, որը վերոնշյալ ընկերությունը գնահատում է մոտ 190,000 (այն է՝ 2018 թվականին միայն ԱՄՆ-ում տվյալագիտությանը հարող մոտ 190,000 թափուր աշխատատեղ կլինի):

⁴⁴ Տե՛ս [McKinsey & Co. "Big data: The next frontier for innovation, competition, and productivity".
\[http://www.mckinsey.com/~media/McKinsey/Business Functions/McKinsey Digital/Our Insights/Big data The next frontier for innovation/MGI_big_data_full_report.aspx\]\(http://www.mckinsey.com/~media/McKinsey/Business Functions/McKinsey Digital/Our Insights/Big data The next frontier for innovation/MGI_big_data_full_report.aspx\)](http://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.aspx)

2015թ. լ ալ ա գ ու լ յ ն 20 ա ջ խ ա ս տ ան ք ն եր ը ⁴⁵
(աղբյ ու ը ը՝ Զ ա մ ա ջ խ ա ր հ ա յ ի ն տ ն տ Ե ս ա ջ հ տ ա կ ան ֆ ո ղ ու մ ⁴⁶)

N	Հաստիքի անվանումը՝ անգլերեն	Աշխատավարձ (միջին տարեկան, ԱՄՆ դոլար)
1	Data Scientist	114,808
2	SEO Manager	45,720
3	Talent Acquisition Specialist	63,504
4	Social Media Manager	40,000
5	Substitute Teacher	24,380
6	Recruiting Coordinator	44,700
7	UX Designer	91,440
8	Digital Marketing Manager	70,052
9	Marketing Assistant	32,512
10	Web Developer	66,040
11	Risk Analyst	69,088
12	Civil Engineer	65,532
13	Client Manager	71,120
14	Instructional Designer	66,040
15	Marketing Analyst	60,000
16	Software QA Engineer	91,440
17	Web Designer	53,848
18	Research Technician	36,525
19	Program Analyst	71,120
20	Data Analyst	58,928

Տվյալ ազատների վերոնշյալ սակավորյունը, և նրանց բարձր աշխատավարձը միշտ կազմակերպչությունների համար գործարար տեսանկյունից ծախսատար է դարձնում տվյալ ազատներին աշխատանքի ընդունելը: Նման պարագայում այդ կազմակերպչությունների համար տվյալ ազատի փոխարինողն է հանդիսանում տվյալների վերլուծաբանը, որը վերոնշյալ լալագույն 20 աշխատանքներում զբաղեցնում է վերջին հորիզոնականը և գրեթե կրկնակի ցածր միջին աշխատավարձ ունի, քան տվյալ ազատներին է: Սակայն հարկավոր է նշել, որ դասական վիճակագրական և տնտեսաչափական մոդելների կիրառությամբ տվյալների վերլուծաբանը չի կարող նույն լուծել այն խնդիրները, որոնք սովորաբար դրված են US վերլուծաբանի առջև: US-ի վերլուծությունն ունի միշտ արանձնատկություններ, որոնք ծնում են տվյալ ազատության և ավանդական տվյալների

⁴⁵ Հաստիքները ներկայացված են անգլերեն տարբերակով, քանզի նրանցից որոշը դեռևս չունեն իրենց հայերեն համարժեքը
⁴⁶ Տե՛ս <https://www.weforum.org/agenda/2015/10/worlds-best-job-the-answer-might-surprise-you/>

վերլուծության մեթոդական տարբերությունները, որոնք կարող են նշանակալի դեր խաղալ ՄՏ-ի վերլուծության գործընթացում՝ նշանակալիորեն ազդելով վերլուծության ճշգրտության ցուցանիշի վրա:

Այսպիսով, ելնելով գործարար կառավարման ասպարեզում որոշումների կայացման գործընթացում ՄՏ-ի վերլուծության կիրառության մակարդակից, տվյալ ագետների նկատմամբ առկա պահանջարկից և նրանց աշխատանքի ընդունելի ծախսատարությունից՝ սույն հետազոտության հաջորդ հատվածում կներկայացնենք տվյալ ագիտության համեմատական արդյունավետությունը, ինչպես նաև կվերլուծենք ՀՀ աշխատանքային տվյալ ագիտությամբ զբաղվող մասնագետների ժողովրդագրական կազմը: Հաջորդող գլխի վերջին հատվածում կդիտարկենք նաև ՀՀ տվյալ ագետների եկամտային խմբի վրա ազդող հիմնական փոփոխականներն ու նրանց ազդեցության ուղղությունը:

ԳԼՈՒԽ 2. ՏՎՅԱԼՆԵՐԻ ԲԱԶԱՅԻ ՏՆՏԵՍԱԳԻՏԱԿԱՆ ՎԵՐԼՈՒՆՈՒԹՅԱՆ ՄԵԹՈԴԱԿԱՆ ՄՈՏԵՑՈՒՄՆԵՐԸ ԵՎ ԴՐԱՆՔ ԿԻՐԱՌՈՂ ՄԱՆԱԳԵՏՆԵՐԸ

2.1 Մեծ տվյալների և ավանդական տվյալների տնտեսագիտական վերլուծության առանձնահատկությունները և հիմնական տարբերությունները կառավարման ոլորտում

Տեղեկատվությունն ստանալու նպատակով տվյալների խորը վերլուծությունը միշտ եղել է և կա տնտեսագիտության և գործարար կառավարման հիմնախնդիր: Այդ տեսանկյունից տնտեսագիտության մեջ առանձնանում է տնտեսաչափությունը (Էկոնոմետրիկա), որն իրենից ներկայացնում է տնտեսամաթեմատիկական և վիճակագրական մոդելների կիրառմամբ փոփոխականների միջև կորելացիոն կապերի բացահայտման, ազդեցության գնահատման համար օգտագործվող գիտություն⁴⁷: Տնտեսաչափության կիրառությունը լայն տարածված է ինչպես ակադեմիայում, այնպես էլ գործարար աշխարհում՝ կառավարման որոշումներ ընդունելիս:

Այնուամենայնիվ, մեծ տվյալների դարաշրջանում ավանդական դարձած տնտեսաչափական միջարք մոդելներ և մոտեցումներ այլևս կիրառելի և նպատակահարմար չեն գործարար որոշումների կայացման գործընթացում⁴⁸: Այն դեպքում, երբ տնտեսաչափությունը նպատակադրում է կախյալ փոփոխականի վրա անկախ փոփոխականների ազդեցության ուսումնասիրությունը որոշակի ենթադրությունների և պայմանների բավարարման պարագայում, ՄՏ-ի վերլուծության պարագայում այդ պայմանները բավարարելը կամ դրանց ստուգելը այլևս առաջնային խնդրոն առարկա չէ: Բանն այն է, որ ձեռնարկությունները ցանկանում են տվյալները վերլուծել իրական կամ գրեթե իրական ժամանակահատվածում, որն անհնար է դարձնում միջարք տարատեսակ հիպոթեզների թեստավորումը: Այսպիսով, ՄՏ-ի վերլուծության

⁴⁷ Տե՛ս [M. Hashem Pesaran, "Econometrics, The New Palgrave: A Dictionary of Economics", v. 2, 1987, էջեր 8-22](#)

⁴⁸ Դավթյան, Ջ., Մեծ Տվյալների տնտեսագիտական վերլուծության առանձնահատկությունները գործարար կառավարման ասպարեզում: «Ֆինանսներ և Էկոնոմիկա» հանդես, 1-2 (197-198), 2017 թ., էջեր 12-15

հիմնական նպատակն է լավագույնս կանխատեսում իրականացնել անորոշ միջավայրում: Հիմնական նպատակների տարբերությունը ծնում է մի շարք առանձնահատկություններ, որոնք բնորոշ են ՄՏ-ի վերլուծությանը և մի շարք դեպքերում հակասում են դասական տնտեսաչափական մեթոդների տրամաբանությանը: Չնայած այն հանգամանքին, որ Google որոնման համակարգի պատմության ամբողջ ընթացքում «Էկոնոմետրիկա» տերմինի որոնումները միջինում գրեթե 3 անգամ գերազանցում են «Մեծ տվյալներ» տերմինի որոնումներին, կարող ենք տեսնել, որ վերջին տարիներին ՄՏ-ի ոլորտի առեւելի աճի հետևանքով, վերջինիս որոնումները զգալիորեն գերազանցում են «Էկոնոմետրիկա» տերմինի որոնումները: Այնուամենայնիվ, հարկավոր է նշել, որ «Էկոնոմետրիկա» տերմինը, ինչպես երևում է գծապատկերից, պահպանում է իր հաստատուն դիրքը երկարաժամկետ հատվածում:

Գծապատկեր 11.

«Մեծ տվյալներ» (կարույտ) և «Էկոնոմետրիկա» (կարմիր) տերմինների որոնումները Google որոնման համակարգում

**կառուցվել է մեր կողմից (աղբյուրը՝ Google trends⁴⁹)*



Ավանդական տվյալների վերլուծության և տվյալագիտության կիրառման մեթոդաբանական առանձնահատկությունները բազմազան ու բազմաբնույթ են: Յուրաքանչյուր առանձնահատկությունն բնորոշ է որոշակի խնդրի լուծմանը և/կամ որոշակի ոլորտում որոշման կայացմանը: Այս հետազոտությունը փորձում է առանձնացնել այն 3 հիմնական տարբերակիչ

⁴⁹ Տե՛ս <https://trends.google.com/trends/explore?q=big%20data,econometrics>, (07/01/2017)

յ ու ր ա հ ա տ կ ու թ յ ու ն ն եր ը , ո Ր ո ն ք առ ա վ ե լ հ ա ճ ա խ ե ն հ ա ն դ ի պ ու մ գ ո Ր ծ ա ր ա ր ո Ր ո շ ու մ ն եր ի կ ա յ ա ց մ ա ն գ ո Ր ծ ը ն թ ա ց ու մ : Ա յ դ 3 առ ա ն ձ ն ա հ ա տ կ ու թ յ ու ն ն եր ը հ ե տ ն յ ա լ ն ե ն .

ա) Բ ա ց ա կ ա յ ո ղ տ լ յ ա լ ն եր ի խ ն դ ի ի ու ծ ու մ ,

բ) Վ եր լ ու ծ ու թ յ ա ն հ ա մ ար օ գ տ ա գ ո Ր ծ վ ո ղ տ լ յ ա լ ն եր ի ամ ք ո ղ ջ ա կ ա ն ու թ յ ու ն .

գ) Մ ա թ ե մ ա տ ի կ ա կ ա ն /վ ի ճ ա կ ա գ ր ա կ ա ն մ ո ղ ե լ ն եր :

ա) Բ ա ց ա կ ա յ ո ղ տ լ յ ա լ ն եր ի խ ն դ ի ի ու ծ ու մ

Ի ն չ ա լ ե ս ա ր դ ե ն ն շ վ ե լ է հ ե տ ա գ ո տ ու թ յ ա ն ս կ գ բ ն ա մ ա ս ու մ` US-ն իր ա կ ա ն կ ա մ գ Ր ե թ ե իր ա կ ա ն ժ ա մ ա ն ա կ ա հ ա տ վ ա ծ ու մ բ ա գ մ ա գ ա ն գ ո Ր ծ ո ն ն եր ի մ ա ս ի ն հ ա վ ա ք ա գ ր վ ո ղ և մ ե ծ ծ ա վ ա լ ու ն ե ց ո ղ տ լ յ ա լ ն եր ն ե ն : Տր ա մ ա բ ա ն ա կ ա ն է ե ն թ ա դ Ր ե լ , ո Ր ո Ր ո շ գ ո Ր ծ ո ն ն եր ի մ ա ս ի ն տ լ յ ա լ ն եր ը կ ա ր ո ղ ե ն թ եր ի լ ի ն ե լ կ ա մ բ ա ց ա կ ա յ ե լ ի ն չ -ո Ր ք ա ն ա կ ի հ ե տ ա գ ո տ ու թ յ ա ն ս ու բ յ ե կ տ ն եր ի հ ա մ ար : Ա յ դ ի ս կ պ ա տ ճ առ ո վ տ լ յ ա լ ն եր ի վ եր լ ու ծ ու թ յ ա ն ը ն ա խ ո Ր դ ո ղ փ ու լ էր ի ց մ ե կ ը տ լ յ ա լ ն եր ի շ տ ե մ ար ա ն ի մ ա ք Ր ու մ ն է , ո Ր ը թ ու յ լ է տ ա լ ի ս լ ու ծ ե լ բ ա ց ա կ ա յ ո ղ ա Ր ժ ե ք ն եր ի խ ն դ ի Ր ը :

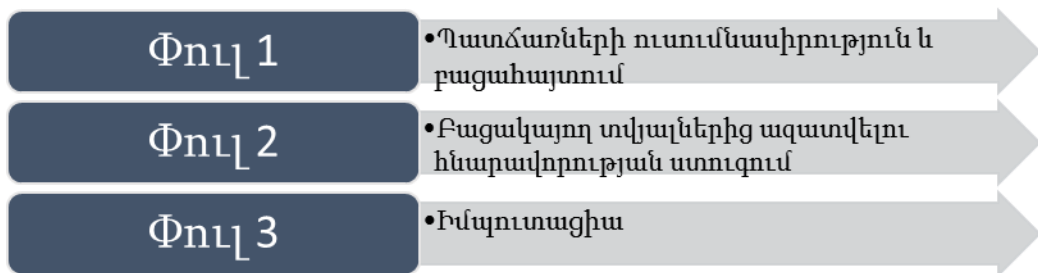
Տ ն տ ե ս ա չ ա փ ա կ ա ն դ ա ս ա կ ա ն վ եր լ ու ծ ու թ յ ա ն պ ար ա գ ա յ ու մ բ ա ց ա կ ա յ ո ղ ա Ր ժ ե ք ն եր ու ն ե ց ո ղ հ ե տ ա գ ո տ ու թ յ ու ն ն եր ը (ս ու բ յ ե կ տ ն եր ը) դ ի տ վ ու մ ե ն ո Ր պ ե ս խ ո չ ը ն դ ո տ և ա գ ա տ վ ու մ ե ն ն ր ա ն ց ի ց` դ ե ն ն ե տ ե լ ո վ շ տ ե մ ար ա ն ի ց : Գ ո Ր ծ ն ա կ ա ն ու մ շ ա տ հ ա ճ ա խ ա ն հ Ր ա ժ ե շ տ է լ ի ն ու մ կ ա ն խ ա տ ե ս ե լ ի ն չ -ո Ր հ ա ճ ա խ ո Ր դ ի գ ն ու մ կ ա տ ար ե լ ու հ ա վ ա ն ա կ ա ն ու թ յ ու ն ը կ ա մ ի ն չ -ո Ր ա շ խ ա տ ա կ ց ի ա շ խ ա տ ա ն ք ը լ ք ե լ ու կ ա մ փ ո խ ե լ ու հ ա վ ա ն ա կ ա ն ու թ յ ու ն ը , ե Ր բ վ եր ջ ի ն ի ս մ ա ս ի ն ու ն ե ց ա ծ տ լ յ ա լ ն եր ը ու ն ե ն բ ա ց ա կ ա յ ո ղ ա Ր ժ ե ք ն եր : Ա վ ա ն դ ա կ ա ն մ ո տ ե ց մ ա մ ք հ ն ար ա վ ո Ր չ է իր ա կ ա ն ա ց ն ե լ ա յ դ կ ա ն խ ա տ ե ս ու մ ը , ք ա ն ի ո Ր ն մ ա ն հ ա ճ ա խ ո Ր դ ի կ ա մ ա շ խ ա տ ա կ ց ի մ ա ս ի ն տ լ յ ա լ ն եր ի ու ղ ղ ա կ ի կ ջ ն ջ վ ե ն շ տ ե մ ար ա ն ի ց⁵⁰ : Ե վ դ ա հ ե տ ա գ ո տ ու թ յ ա ն տ ե ս ա ն կ յ ու ն ի ց խ ն դ ի Ր չ ի առ ա ջ ա ց ն ի , ք ա ն գ ի , ի ն չ ա լ ե ս ն շ վ ա ծ է Ր վ եր ն ու մ , տ ն տ ե ս ա չ ա փ ու թ յ ա ն ն պ ա տ ա կ ն է բ ա ց ա հ ա յ տ ե լ կ ո Ր ե լ ա ց ի ո ն կ ա պ ը փ ո փ ո խ ա կ ա ն ն եր ի մ ի ջ և , և ո չ թ ե իր ա կ ա ն ա ց ն ե լ կ ա ն խ ա տ ե ս ու մ ա ն հ ա տ ա կ ա ն հ ե տ ա գ ո տ ա կ ա ն

⁵⁰ Տե խ ն ի կ ա պ ե ս կ ա ն խ ա տ ե ս ու մ իր ա կ ա ց ն ե լ ը ա յ դ պ ար ա գ ա յ ու մ ն ու յ ն ա լ ե ս հ ն ար ա վ ո Ր է , ս ա կ ա յ ն կ ա ն խ ա տ ե ս մ ա ն հ ա մ ար չ ի օ գ տ ա գ ո Ր ծ վ ի բ ա ց ա կ ա յ ո ղ ա Ր ժ ե ք ու ն ե ց ո ղ փ ո փ ո խ ա կ ա ն ը և ա Ր դ յ ու ն ք ու մ կ ա ն խ ա տ ե ս մ ա ն ճ շ գ Ր տ ու թ յ ու ն ը կ տ ու ժ ի :

սուբյեկտների վերաբերյալ: Այդ իսկ պատճառով ՄՏ-ի վերլուծության պարագայում, որի առաջնային նպատակն է հենց կանխատեսում կատարելը, բացակայող տվյալները նույնպես դիտվում են որպես արժեքավոր տվյալներ և հատուկ ուշադրության են արժանանում: Այդ գործընթացը, այնպես ինչպես է կոնոմետրիկայում, կոչվում է տվյալների մաքրման գործընթաց և հիմնականում բաղկացած է 3 փուլից (տես՝ գծապատկեր 2):

Գծապատկեր 12.

Բացակայող տվյալների խնդրի տվյալ ագիտական և ուժման փուլերը



Փուլ 1 – Պատճառների ուսումնասիրություն և բացահայտում

1-ին փուլում բացակայող արժեքները ենթարկվում են ուսումնասիրության, որպեսզի պարզ դառնա վերջիններիս առաջացման պատճառները: Շատ դեպքերում, բացակայող արժեքներն ինքնին տեղեկատվություն են պարունակում հետազոտվող սուբյեկտների որևէ կատեգորիայի մասին: Ուսումնասիրության արդյունքում պատճառների բացահայտման դեպքում բացակայող արժեքներին վերագրվում է որևէ նոր կատեգորիա և այդքանով մաքրման գործընթացն ավարտվում է: Պատճառների բացահայտման անհնարի նության պարագայում (որը, ելնելով ՄՏ-ի մեծությանից, առավել հավանական է) անցում է կատարվում 2-րդ փուլ:

Փուլ 2 – Բացակայող տվյալներից ազատվելու հնարավորության ստուգում

2-րդ փուլում ստուգվում է թե արդյոք հնարավոր է ազատվել բացակայող տվյալներից՝ ինչպես ավանդականորեն հաճախ արվում է: Այդ պատճառով ուսումնասիրվում է կանխատեսման ենթակա փոփոխականի արժեքների բաշխումը առանց բացակայող տվյալների և բացակայող տվյալների ներկայության: Եթե երկու դեպքում առկա է միևնույն բաշխումը, ապա բացակայող տվյալները դեն են նետվում

և այս քանով տվյալների մաքրման գործընթացն ավարտվում է: Այս մոտեցումը բացատրվում է այն հանգամանքով, որ բաշխման անփոփոխ մնալու պարագայում բացակայող արժեք ունեցող հետազոտությանը կամ քիչ են և կարելի է անտեսել առանց գործնական արժեք կորցնելու կամ, որ նրանք այնքան նման են ոչ բացակայող տվյալներ ունեցող հետազոտությանը, որ վերջիններիս կանխատեսման արդյունքները կարելի է կիրառել նաև բացակայող արժեք ունեցող սուբյեկտների վա:

Սակայն շատ հաճախ կանխատեսվող փոփոխականի բաշխվածությունը փոխվում է՝ անհնար դարձնելով բացակայող տվյալներից ազատվելու հնարավորությունը: Նման խնդրի առաջ հաճախ են կանգնում բանկային համակարգում: Այսօր գրեթե բոլոր առևտրային բանկերը ունեն վարկառուի դիմում հայտի գնահատման ավտոմատացված համակարգ, որը կանխատեսումն իրականացնում է գոյություն ունեցող տվյալների հիման վրա: Իհարկե, առևտրային բանկերը բազում տվյալներ ունեն այն հաճախորդների մասին, որոնք դիմում են ներկայացրել վարկ ստանալու համար և ստացել դրական պատասխան: Սակայն կան հաճախորդներ, որոնց դիմումը մերժվել է և այդպիսով, առևտրային բանկը հնարավորություն չի ունեցել հավաքել տվյալներ այդ հայտատուների մասին (օր.՝ ինչպիսի հաճախականության մեջ և պարտականության մեջ հայտատու կիրականացնել վճարումները, եթե ստանար վարկը): Այնհայտ է, որ նման պարագայում բացակայող արժեքներից ազատվելուց հետո առևտրային բանկի շտեմարանում կմնան տվյալներ միայն վարկը ստացած հաճախորդների մասին (այն է՝ կանխատեսվող փոփոխականի բաշխումը կփոխվի), որը կխեղաթյուրի ավտոմատացված համակարգի աշխատանքը: Այդ դեպքում, անցում է կատարվում 3-րդ փուլ, որտեղ բացակայող արժեքները ենթարկվում են վիճակագրական իմպուտացիայի:

Փուլ 3 – Իմպուտացիա

Իմպուտացիան իրանից ներկայացնում է բացակայող արժեքների ավտոմատացված վերագրում ինչ-որ նոր արժեքով: Տեսականորեն իմպուտացիան կարելի է բաժանել 2 տեսակի.

➤ Կենտրոնական հակվածության գնահատականների միջոցով
իմպուլտացիա

Կենտրոնական հակվածության գնահատման համար
առօրյայում օգտագործվում են 3 գնահատականներ՝
միջին, մեդիան և մոդա: Միջինն իրենից ներկայացնում է
հասարակ թվաքանակի միջինը: Այն դեպքում, երբ առկա է
բազմարժեք փոփոխական՝ բացակայող արժեքներով,
օգտագործվում է բացակայող արժեքների միջին արժեքով
իմպուլտացիայի մեթոդը: Այդ մեթոդի առավելությունը
կայանում է նրանում, որ իմպուլտացիայից հետո
փոփոխականի արժեքը միջինում մնում է անփոփոխ:
Սակայն մոտեցումն ունի մի մեծ թերությունն ունյնպես.
ծայրահեղ արժեքների առկայությունը խեղաթյուրում է
միջին ցուցանիշը: Այդ իսկ պատճառով, եթե հետազոտվող
փոփոխականը ունի ծայրահեղ արժեքներ (քիչ քանակով
շատ մեծ և/կամ շատ փոքր արժեքներ), ապա իմպուլտացիայի
համար օգտագործվում է մեդիան: Մեդիան արժեքների
աճման կամ նվազման կարգով դասավորվածության
դեպքում մեջտեղում հանդիպող թիվն է (զույգ քանակով
հետազոտություններ ունենալու դեպքում՝ մեջտեղի 2
թվերի միջինը): Այս պարագայում քիչ թվով ծայրահեղ
արժեքները ազդեցություն չեն ունենում գնահատականի
վրա:

Միջին ցուցանիշի մեկ այլ թերությունն է հանդես գալիս
կատեգորիկ փոփոխականների հետ աշխատելու
պարագայում: Օրինակ, եթե դիտարկենք հետազոտվող
սուբյեկտի սեռը, ապա այդ փոփոխականի արժեքը կարող է
լինել միայն 1 կամ 0: Միջինի հաշվարկման դեպքում
ստացվելու են արժեքներ, որոնք 0-ից 1 միջակայքում
ռացիոնալ թվեր են չլինելու՝ այլ ոչ թե ամբողջ թիվ: Այդ
պարագայում, ավելի նպատակահարմար է կիրառել մոդա
գնահատականը, որն իրենից ներկայացնում է ամենահաճախ
հանդիպող թիվը:

➤ Բացակայող արժեքների էքստրապոլյացիա

Այն դեպքերում, երբ բացակայում են ոչ մեծ թվով, բայց գործարար կառավարման համար մեծ կարևորություններ կայացնող արժեքներ, այդ տվյալներին առանձնահատուկ ուշադրություն է ներկայացվում և բոլոր բացակա արժեքների գնահատման համար նույն գնահատականը չի օգտագործվում: Այդ պարագայում բավականին կիրառելի է k ամենամոտ հարևաններ (այսուհետ՝ kNN – k Nearest Neighbors) մեթոդը: Այս մեթոդը յուրաքանչյուր բացակայող արժեք ունեցող սուբյեկտին համադրում է k քանակով ոչ բացակայող արժեք ունեցող սուբյեկտներ, որոնք ամենամոտ են գտնվում իրեն՝ ըստ որևէ չափման միավորի: Այնուհետև այդ k ամենամոտ սուբյեկտների հետազոտության ենթակա փոփոխականների արժեքների թվաքանական միջինը վերագրվում է բացակայող արժեքին:

Մոդելի կիրառման համար անհրաժեշտ է նշել 2 պարամետր՝ ամենամոտ հարևանների թիվը (k) և հարևանների համեմատության համար հիմք հանդիսացող չափման միավորը: Առաջինի որոշումը կատարվում է էմպիրիկ սկզբունքով. օգտագործվում են մի քանի արժեքների k -ի համար, և ստուգվում է, թե որի դեպքում է կանխատեսումը լավագույնը: Երկրորդի ընտրության համար տարբերակների թիվը նույնպես բավականին լայն է, այդ իսկ պատճառով գործնականում կիրառվում է դասական էվկլիդեսյան հեռավորության չափման միջոցը (կոորդինատային տարածության մեջ հետազոտվող սուբյեկտների միջև եղած հեռավորությունը):

բ) Վերլուծության համար օգտագործվող տվյալների ամբողջականությունը

Դասական տվյալների վերլուծության և ՄՏ-ի վերլուծության հիմնական տարբերիչ առանձնահատկություններից մյուսը վերլուծության համար օգտագործվող տվյալների ամբողջականությունն է: Քանի որ տնտեսաչափության նպատակը փոփոխականների միջև կախվածության բացահայտումն ու

ագդեցու թյան գնահատումն է, այն վերլուծությունն իրականացնում է հասանելի տվյալների ամբողջական շտեմարանի վրա:

ՄՏ-ի գոյությունը թույլ է տալիս վերլուծությունն իրականացնել միայն տվյալների մի մասի վրա, քանզի մեծության շնորհիվ վիճակագրական նշանակալիությունը միևնույն է ապահովվում է: Իսկ ահա տվյալների մյուս մասը, որը վերլուծության համար չէր օգտագործվել, կիրառվում է որպես թեստավորման միջոց՝ կանխատեսման արդյունքները ստուգելու և վավերացնելու համար: Եվս մեկ անգամ սահեռանալ է հանդիսանում այն հանգամանքի, որ ՄՏ-ի վերլուծության նպատակը կանխատեսում անելն է, ինչը յուրահատուկ կարևորություն է ուղղորդում կանխատեսված արդյունքների թեստավորմանն ու գնահատմանը: Ընդունված արակտիկա է վերլուծությունն իրականացնել ՄՏ-ի շտեմարանի 50-70%-ի վրա, և արդյունքները թեստավորել և վավերացնել մնացյալ 50-30%-ի վրա: Վերջինս հնարավորություն է տալիս միաժամանակ կիրառել մի քանի մոդել և թեստավորման ու վավերացման միջոցով պարզել, թե այդ մոդելներից որն է լավագույնս կանխատեսում նպատակադրված փոփոխականը: Չատկանշական է, որ ՄՏ-ի այդ մասնատումը տեղի է ունենում պատահականության սկզբունքով, որպեսզի ժամանակային կամ մեկ այլ գործոնից կախված տվյալների մասնաբաժինների միջև տարբերություններ չլինեն:

Տվյալների մասնաբաժինների պատահական ընտրությունն իր թերությունն ունի: Վերջինիս հետևանքով, երբեմն, հնարավոր է ստանալ կանխատեսման ճշգրտության տարբեր արդյունքներ՝ կախված պատահականորեն ընտրված տարբեր թեստավորման և վավերացման մասնաբաժիններից: Այդ խնդրի լուծման համար կիրառվում է խաչաձև վավերացման տեխնիկան /cross validation/: Այս տեխնիկան ենթադրում է տվյալների մասնաբաժանումը ավելի քան 2 խմբի (օրինակ՝ 10 խմբի): Այնուհետև, ՄՏ-ի վերլուծությունն իրականացվում է այդ խմբերից մի քանիսի վրա (օրինակ 9-ի) և թեստավորվում ու վավերացվում վերջինի վրա: Չափորդ քայլում վերլուծությունն իրականացվում է այդ խմբերից ընտրված մի

քանի այլ խմբի վրա և թեստավորվում և վավերացվում մեկ ուրիշ խմբի հիման վրա (օրինակ՝ վերլուծության համար 10 խմբից ընտրվում են մեկ այլ 9 խումբ, իսկ թեստավորման համար՝ մնացյալ տասներորդը): Այս գործողությունն այնքան է շարունակվում մինչև բոլոր հնարավոր տարբերակների համար թեստավորումն ու վավերացումն իրականացված են լինում (10 խմբի դեպքում այդ գործողությունը կիրականացվի 10 անգամ): Արդյունքում ստացված բոլոր ճշգրտության արդյունքների միջինը ընդունվում է որպես ճշգրտության արդյունք և հանդիսանում տարատեսակ մոդելների համեմատության համար հիմք:

գ) Վիճակագրական մոդելներ

ՄՏ-ի մեծությունը (ծավալ, շարժունություն և բազմազանություն) անհնարին է դարձնում դասական տնտեսաչափական մոդելների կիրառությունը՝ կանխատեսման նպատակով: Այդ իսկ պատճառով զարգացել են նոր վիճակագրական և մաթեմատիկական մոդելներ, որոնց հիմնական նպատակը հենց ՄՏ-ի հետ աշխատանքն է: Հարկավոր է նշել, որ այդ մոդելներից ոմանք այդքան էլ նոր չեն: Իրականում դրանք վաղուց արդեն գոյություն ունեն, սակայն դեռևս տեսական հետազոտությունից արակտիկ կիրառության անցում չէր կատարվել: Նման մոդելների թվին է պատկանում Արհեստական նեյրոնային ցանցը (Artificial neural network), որի արմատները ձգվում են մինչև 1943 թ.⁵¹: Այնուամենայնիվ վերջինս իր կիրառությունն է գտել կառավարման ոլորտում միայն վերջին տարիներին: Արհեստական նեյրոնային ցանցը իրենից ներկայացնում է ներդրված ֆունկցիաների ամբողջություն, որոնցից յուրաքանչյուրը ունի իր կշիռը ցանցի ներսում: Տնտեսաչափական հայտնի լոգիստիկ ռեգրեսիայի մոդելը (և առավելագույն ճշմարտանմանության գնահատականը) կարելի է համարել արհեստական նեյրոնային ցանցի պարզ օրինակ: Այս մոտեցումը թույլ է տալիս ստանալ կանխատեսման առավել ճշգրիտ արդյունքներ՝ փոփոխականների միջև կախվածությունը դիտարկելով ինչպես գծային, այնպես էլ ոչ գծային մակարդակում: Նման առանձնահատկությունը բավականին բարձր է գնահատվում

⁵¹ Տե՛ս [McCulloch, Warren; Walter Pitts \(1943\). "A Logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathematical Biophysics](#)

անընդհատ փոփոխություն մեջ գտնվող գործարար աշխարհում, քանզի փոփոխականների միջև մշտապես կայուն և միատեսակ կախվածություն ենթադրելը իրատեսական չէ (օրինակ՝ կայուն գծային կախվածության մասին ենթադրությունը):

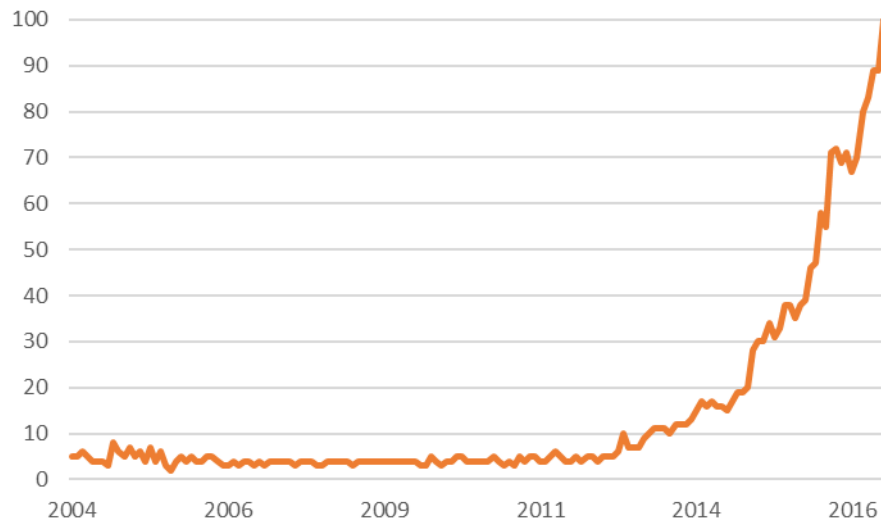
Արհեստական նեյրոնային ցանցերի վրա հիմնված գիտության ճյուղը, որը նաև հանդիսանում է արհեստական ինտելեկտի հիմքը, կոչվում է «խորը ուսուցում» /Deep learning/: Օգտագործելով Google որոնողական համակարգի տվյալները՝ կարելի է տեսնել թե ինչպես է խորը ուսուցումը գնալով ավելի հանրաճանաչ դառնում:

Խորը ուսուցման գիտության հանդեպ եղած հետաքրքրությունն այսպիսի աստիճանային աճը պայմանավորված է ՄՏ-ի վերլուծության գործընթացում վերջիններիս կիրառման անհրաժեշտությամբ: Համարելով «Մեծ տվյալներ» և «խորը ուսուցում» տերմինների որոնման մասին տվյալները կարելի է տեսնել, որ նրանց մեջ առկա է խոշոր կորելացիոն կապ:

Գծապատկեր 13.

«Խորը ուսուցում» տերմինի որոնումների միտումը

**կառուցվել է մեր կողմից (տվյալների աղբյուրը՝ Google trends⁵²)*

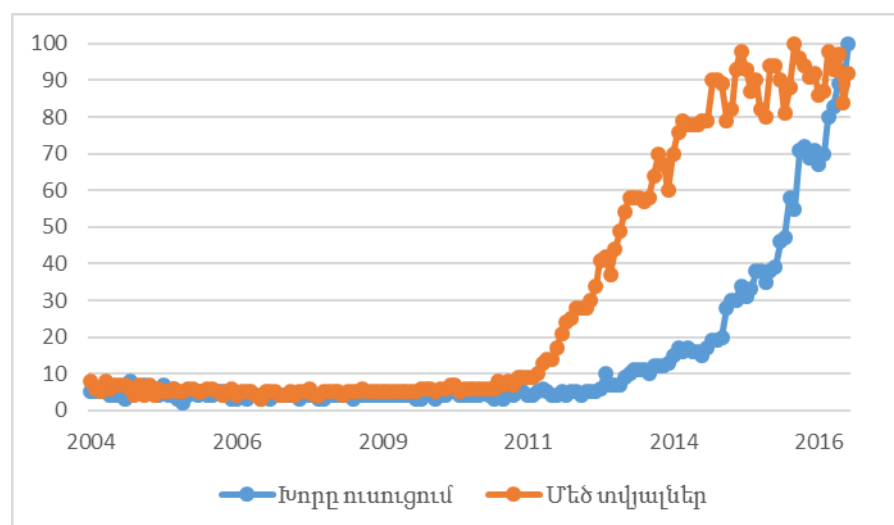


Գծապատկեր 14-ը ցույց է տալիս, որ սկզբնական հատվածում վերոնշյալ տերմիններից ոչ մեկը և ոչ մյուսը չեն գրանցել բարձր որոնման արդյունքներ: Այնուամենայնիվ, վերջին 5 տարիների ընթացքում և՛ «Մեծ տվյալներ», և՛ «խորը ուսուցում» տերմինի նկատմամբ որոնողական համակարգում գրանցվել է հետաքրքրության զգայուն աճ:

Գծապատկեր 14.

«Մեծ տվյալներ» և «խորը ուսուցում» տերմինների որոնումների համադրություն

**կառուցվել է մեր կողմից (տվյալների աղբյուրը՝ Google trends⁵³)*



⁵² Տես՝ <https://trends.google.com/trends/explore?date=all&q=deep%20learning>, (07/01/2017)

⁵³ Տես՝ <https://trends.google.com/trends/explore?date=all&q=deep%20learning,big%20data>, (07/01/2017)

Գծապատկերից երևում է, որ սկզբնական հատվածում US-ի աճը գերակայում է խորը ուսուցման որոնումների աճին: Սա բացատրվում է այդ երկու ոլորտների միջև գոյություն ունեցող պատճառահետևանքային կապով. սկզբից զարգացել է ՏՀՏ ոլորտը, որն էլ հնարավորություն է տվել US-ի հավաքագրման և հանրաճանաչության: Իսկ ահավերջինս պատճառ է հանդիսացել խոր ուսուցման գիտության ձևավորման և զարգացման համար: Վերոնշյալ տերմինների որոնումների համար հաշվարկվել է Փիրսոնի (Pearson) կորելացիայի գործակիցը, որի արժեքն է 0.79: Նման բարձր կորելացիոն կապի առկայությունը հիմնավորում է այն ենթադրությունը, որ այս երկու ոլորտները միմյանց փոխկապակցված են:

Ինչպես տեսնում ենք տվյալագիտությունն ու ավանդական տվյալների վերլուծությունը ունեն մի շարք առանձնահատկություններ, որոնք կարող են որոշիչ դեր խաղալ որոշումների կայացման գործընթացում: Իհարկե, չի կարելի համեմատականորեն գերադասել մեկը մյուսից, քանզի այդ գիտություններից յուրաքանչյուրն իր տեղն ունի: Սակայն ձեռներեցի տեսանկյունից կարևոր է հասկանալ խնդիրը, որը լուծվում է տվյալների վերլուծության միջոցով և որոշել, թե որ մոտեցումն է ավելի նպատակահարմար:

Դասական տվյալների վերլուծության և տվյալագիտության հիմնական տարբերությունները հիմք են հանդիսանում այդ վերլուծությամբ զբաղվող մասնագետների տարբերության համար: Դասական վերլուծությամբ զբաղվող մասնագետները սովորաբար տիրապետում են SPSS, Stata, EViews, Excel և նման այլ ծրագրային փաթեթների, որոնք, ինչպես նշվեց առաջին գլխում, հնարավոր է կիրառել US-ի վերլուծության պարագայում: Հատկանշական է, որ դասական տվյալների վերլուծության և տվյալագիտության վերոնշյալ 3 հիմնական տարբերակիչ առանձնահատկություններից միայն բացակայող արժեքների խնդիրն է, որ մասամբ լուծելի է նմանատիպ ծրագրային փաթեթների միջոցով: Մյուս երկու առանձնահատկությունները՝ խաչաձև վավերացման տեխնիկայի կիրառությունն ու US-ի վերլուծության պարագայում

օգտագործվող մաթեմատիկական/վիճակագրական մոդելների առկայությունը, բավարարված չեն այդ ծրագրային փաթեթների կողմից: Այդ իսկ պատճառով տվյալ ազետները հանդիսանում են US-ի վերլուծությանը զբաղվող մասնագետներ, որոնք առօրյայում կիրառում են տվյալ ազիտության համար նախատեսված Python, R և այլ ծրագրային փաթեթներ /ինչը ցույց են տալիս մեր կողմից իրականացված հարցման տվյալները/: Նշված փաթեթները պահանջում են ավելի խորը համակարգչային գիտելիքներ և ավելի բարդ են, քան դասական տվյալների վերլուծության համար կիրառվող ծրագրերը, ուստի տվյալ ազետներից բացի տնտեսագիտական/վերլուծական ունակություններից պահանջվում է նաև խորը համակարգչային գիտելիքներ: ՀՀ աշխատանքային տվյալ ազիտությանը զբաղվող մասնագետների համապատասխանությունը վերոնշյալ պահանջներին ստուգելու, նրանց ժողովրդագրական կազմը բացատրելու և եկամտային խմբի վրա ազդող հիմնական փոփոխականները բացահայտելու նպատակով՝ հետազոտության հաջորդ հատվածում ներկայացվում է ՀՀ աշխատանքային տվյալ ազետների աշխատուժի առաջարկի ուսումնասիրությունը՝ արդյունքները համեմատելով միջազգային ցուցանիշների հետ:

2.2 Տվյալ ազետների աշխատուժի առաջարկը ՀՀ աշխատանքային տնտեսությանը

Տվյալ ազիտության ամենձևահատկությունները պահանջում են տվյալ ազետից տիրապետել որոշակի մաթեմատիկական/վիճակագրական մոդելների, US-ի հետ աշխատելու համար մշակված տեխնիկաների և համապատասխան ծրագրային փաթեթների, որոնք դեռևս ամբողջությամբ չեն դասավանդվում \$որմալ կրթական համակարգում: Հետազոտության սույն հատվածում ներկայացվում է ՀՀ աշխատանքային տվյալ ազիտությանը զբաղվող մասնագետների առաջարկի կառուցվածքը՝ ընդգծելով ծրագրային փաթեթների տիրապետումը, գիտելիքի աղբյուրները, ժողովրդագրական կազմը և միջարքայլ փոփոխականներ:

Տվյալ ազետների աշխատուժի առաջարկն ուսումնասիրելու և գնահատելու նպատակով մեր կողմից իրականացվել է առցանց

հարցում⁵⁴: Հարցումն ուղղվել է US վերլուծությանը՝ զբաղվողների թիրախային խմբին, ովքեր հայտնաբերվել են սոցիալական ցանցերում համապատասխան խմբերին անդամագրման, տվյալագիտական միջոցառումներին և մրցույթներին մասնակցության փաստերով: Ելնելով այն հանգամանքից, որ տվյալագիտական տերմինաբանությանը հայերենում հիմնականում դեռևս չի ձևավորվել, հարցումն անց է կացվել անգլերեն լեզվով: Հարցումը եղել է անանուն, և մասնակիցները տեղեկացված են եղել, որ տվյալները հրապարակվելու են միայն ամբողջականացված տարբերակով՝ ապահովելով առանձին մասնակիցներին բնութագրող անհատական տեղեկատվության գաղտնիությունը:

Հարցաշարը կազմված է եղել 12 հարցից (այդ թվում՝ և՛ բաց, և՛ փակ հարցեր), որոնց մի մասը նպատակաուղղված է եղել պարզորոշել տվյալագետի անձը բնակչության ժողովրդագրական կազմի առումով, իսկ մյուս մասը՝ ստանալ համապատասխան տեղեկատվություն տվյալագետների գիտելիքների ծավալի և դրանց գործնական կիրառության վերաբերյալ⁵⁵: Հետազոտության արդյունքում պարզ է դարձել, որ ՀՀ-ում տվյալագիտական խմբերում, միջոցառումներում և մրցույթներում կամոտ 300 անդամագրված անձ: Սույն թվում հաշվի է առնվել այն հանգամանքը, որ միևնույն անձը կարող էր մասնակցել 2 կամ ավելի մրցույթների: Աշխատանքային ճիշտ պատկերը ստանալու նպատակով նման անձինք հարցվողների ցուցակում ընդգրկվել են միայն մեկ անգամ: Անհրաժեշտ է նշել, որ այս 300-ի մեծ մասը հանդիսանում են ուսումնառության նպատակով անդամագրված անհատներ, որոնք համապատասխան գործունեությանը դեռևս չեն զբաղվում: Արդյունքում՝ գոտ տվյալագետների համախումբը գնահատվել է ընդամենը մոտ 120 հոգի: Հետևաբար, ռեսպոնդենտների նման թվի պարագայում նշանակալի արդյունքների ստացման համար անհրաժեշտ էր ձևավորել առնվազն 54 հոգուց բաղկացած ընտրանք, որը կապահովեր 10% սխալ անք՝ 95% վստահելիության միջակայքում:

⁵⁴ Մարգարյան Յ., Դավթյան Յ., Տվյալագետների աշխատելի առաջարկը ՀՀ աշխատանքային հարցման արդյունքների վերլուծությանը: «Ակունք» ուսումնամեթոդական հոդվածների ժողովածու, №1 (15), Երևան, 2017թ.

⁵⁵ Որոշ հարցերը բաղկացած են եղել նաև ենթահարցերից:

Հարցմանը մասնակցել են 76 հոգի՝ գերազանցելով վերոնշյալ ընտրանքի ստորին սահմանը և ապահովելով ավելի բարձր նշանակալիություն, միաժամանակ լինելով նաև ներկայացուցչական ռեսպոնդենտների ավելի մեծ թվի պարագայում⁵⁶: Հաշվարկի հիմքում դրվել է ընտրանքի մեծության հաշվարկի համար կիրառվող Wald մեթոդը:

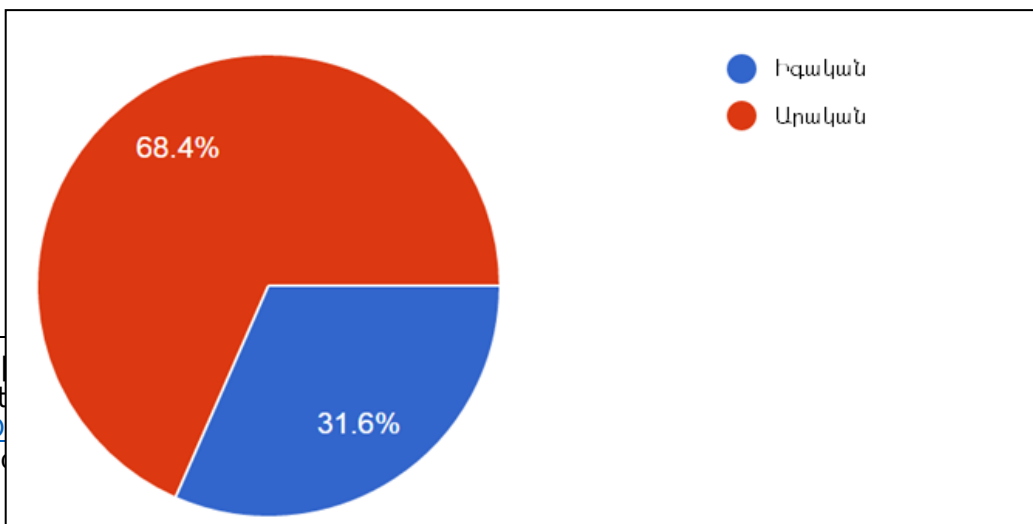
Տվյալագիտության ոլորտում իրականացվող հիմնարար գիտական հետազոտությունների և դասագրքերի հրատարակմամբ զբաղվող O'Reilly հրատարակչությունն արդեն 3 տարի անընդմեջ իրականացնում է տվյալագետների համաշխարհային հարցում՝ ներկայացնելով ոլորտում տիրող իրավիճակը, օգտագործվող հիմնական ծրագրային փաթեթները և տվյալագետների ժողովրդագրական կազմը⁵⁷: Մեր կողմից իրականացվելիք հարցման արդյունքները տվյալագետների համաշխարհային հարցման չափորոշիչներին համահունչ դարձնելու նպատակով O'Reilly կազմակերպությունից ստացել ենք իրենց կազմած հարցաթերթը և այն համապատասխանեցրել ՀՀ իրավիճակին՝ կրճատելով նաև որոշ նեղ մասնագիտական հարցեր, որոնք սույն հետազոտության խնդրո առարկան չէին հանդիսանում:

Հարցման մասնակիցները հարցաթերթում լրացրել են ժողովրդագրական բնույթի 3 դասական հարց, այն է՝ տարիքը, սեռը և աշխատանքային փորձը: Գծապատկեր 15-ում ցուցադրված է հարցման մասնակիցների սեռային բաշխվածությունը:

Գծապատկեր 15.

Տվյալագետների սեռային բաշխվածությունը ՀՀ-ում

**կառուցվել է մեր կողմից*



⁵⁶ 76 հոգի
անհրաժեշտ
⁵⁷ Տե՛ս [O'Reilly](#)
data-scienc

ամար
L մ:
ports/2015-

Ինչպես ներկայացված է գծապատկերում, տվյալ ազետների ճնշող մեծամասնությունը (68,4%-ը) հանդիսանում է արական սեռի ներկայացուցիչ, ինչն ավելի քան երկու անգամ գերազանցում է իգական սեռի ներկայացուցիչների թիվը: Տվյալ ագիտության ոլորտում իգական սեռի նման ցածր ներգրավվածությունը կարող է բացատրվել երկու պատճառով.

- Մեծ տվյալները, ինչպես հայտնի է, հնարավոր չէ վերլուծել դասական /ավանդական/ եղանակով. այս աշխատանքի իրականացման համար անհրաժեշտ են որոշակի, անհամեմատ ավելի խորը համակարգչային գիտելիքներ: Գծապատկերում ներկայացվածը լիովին համահունչ է SS ոլորտի մասնագետների սեռային բաշխվածության վիճակագրությանը, ինչը փաստում է /համաձայն «Ձեռնարկությունների ինկուբատոր» հիմնադրամի 2015 թվականի հաշվետվության/, որ ոլորտում ներգրավված մասնագետների 68%-ը արական սեռի ներկայացուցիչներ են⁵⁸:
- ՀՀ տնտեսության ոլորտների մեծ մասում, անկախ գենդերային հորիզոնական խորականության հանգամանքից, իգական սեռի զբաղվածության աստիճանը զգալիորեն ցածր է արական սեռի զբաղվածության աստիճանից⁵⁹:

Այնուամենայնիվ, այս ցուցանիշը ՀՀ-ում բնութագրում է մասնագետների սեռային բաշխվածության ավելի մեծ համասեռություն, քան համաշխարհային միջին ցուցանիշը: Համաձայն O'Reilly-ի հարցման արդյունքների՝ արական և իգական սեռերի հարաբերակցության համաշխարհային միջին ցուցանիշը հավասար է 4:1-ի:

Մասնագետների թիվը միակ փոփոխականը չէ, որն արական սեռի շրջանում իգականի համեմատ ավելի բարձր ցուցանիշ է: Ընտրանքային հարցման մասնակիցները պատասխանել են նաև այն

⁵⁸ Տե՛ս [Ձեռնարկությունների ինկուբատոր հիմնադրամ, «ՀՀ SS ոլորտի հաշվետվություն»](#), 2015 թ, էջ 28

⁵⁹ Տե՛ս [ՀՀ ԱԿԾ, «Աշխատանքի շուկան Հայաստանի Հանրապետությունում»](#), 2016 թ, էջ 55

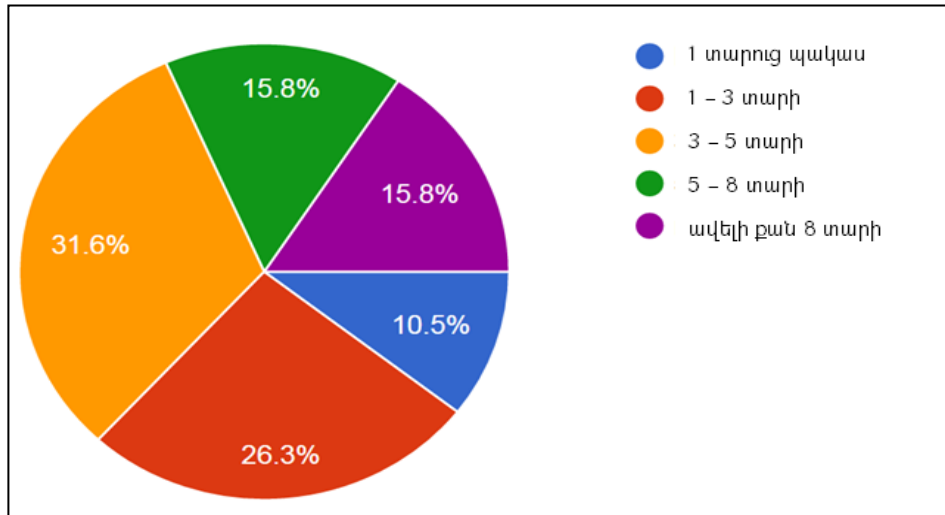
հարցին, թե ներկայիս աշխատանքը լքելու պարագայում որքանով հեշտ կլինի նոր աշխատանք գտնելը: Արական սեռի տվյալագետների շրջանում 1-5 սանդղակով այդ ցուցանիշը միջինում կազմում է 4,23, ինչը նշանակում է, որ տղամարդկանց կարծիքով իրենց աշխատատեղը փոխելը բավականին հեշտ կլինի: Պատկերն այլ է իգական սեռի շրջանում, որտեղ հեշտորեն աշխատանք գտնելու միջին ցուցանիշը 2,83 է: Թերևս դա պայմանավորված է տարիքով և աշխատանքային փորձով, քանզի արական սեռի հարցվածների միջին տարիքը (31) 5 տարով գերազանցել է իգական սեռի հարցվածների միջին տարիքը (26): Ինչ վերաբերվում է տարիքային ընդհանուր բաշխվածությանը, ապա հարցվողների ավելի քան կեսը (52%-ը) 24-28 տարեկան մասնագետներ են, ինչը դարձյալ վկայում է ոլորտի կադրերի երիտասարդության մասին: Աշխատանք գտնելու հեշտության մասին տվյալները O'Reilly-ի հաշվետվությունում չեն ներկայացվում ըստ սեռատարիքային բաշխման: Նշվում է միայն ընդհանուր միջին ցուցանիշը՝ 3,63, որը փոքր-ինչ գիջում է 33 աշխատառուկայում հաշվարկված միջինին՝ 3,78:

Մասնագիտական փորձառության մասին տրված հարցին մասնակիցների 15,8%-ը պատասխանել է, որ այդ ոլորտում ունի 8 տարուց ավելի աշխատանքային փորձ: Ելնելով այն հանգամանքից, որ տվյալագիտությունը նոր մասնագիտություն է, և 8 տարի առաջ դեռևս ձևավորված չէր որպես առանձին ոլորտ, ենթադրվում է, որ հարցվողներն այդ թվում ներառել են նաև տվյալների ավանդական վերլուծության իրենց մասնագիտական փորձը: Հատկանշական է, որ արական սեռի ներկայացուցիչների ավելի քան կեսն (52%) ունի 5 և ավելի տարիների փորձ, մինչդեռ իգական սեռի բոլոր ներկայացուցիչներն ունեն դրանից ավելի քիչ փորձառություն:

Գծապատկեր 16.

Տվյալագետների բաշխվածությունը՝ ըստ մասնագիտական փորձի

**Կառուցվել է մեր կողմից*



Չարմանալիորեն այս ցուցանիշը չի ընդգրկվել վերոնշյալ միջազգային հարցման հարցաթերթում: Այդ իսկ պատճառով, համեմատության համար հիմք է վերցվել Մասաչուսեթսի Տեղեկատվական տեխնոլոգիաների համալսարանի «Կառավարման գիտություններ» բաժանմունքի 2015թ. տարեկան հաշվետվությունը⁶⁰: Համաձայն այդ հաշվետվության՝ տվյալ ազետների աշխատանքային միջին փորձառությունը 5 տարին է:

Տվյալ ազետների աշխատավարձի մեծությունը զգալիորեն տարբեր է ՀՀ-ում միջին աշխատավարձի ցուցանիշից: ՀՀ-ում միջին ամսական անվանական աշխատավարձը կազմում է 187850 ՀՀ դրամ, որը ենթադրում է մոտ 130000 ՀՀ դրամ զուտ աշխատավարձ⁶¹: Ամենաբարձր միջին աշխատավարձերը գրանցվել են ֆինանսական և ապահովագրական գործուներության ոլորտում՝ մոտ 280,70 հազար դրամ և տեղեկատվության ոլկապի ոլորտում՝ մոտ 264,25 հազար դրամ: Հարցման մասնակիցները հնարավորություն են ունեցել ընտրել եկամտային տարբեր միջակայքեր, և արդյունքում պարզվել է, որ հարցված բոլոր տվյալ ազետներն ավելի շատ են վաստակում, քան միջին աշխատավարձն է: Ոլորտի բարձր արտադրողականության մասին վկայում է այն հանգամանքը, որ մասնազետների 78%-ն ավելի բարձր աշխատավարձ են ստանում, քան ֆինանսական և ապահովագրական գործուներության, տեղեկատվության ոլկապի

⁶⁰ Տե՛ս «Minding the Analytics Gap,” MIT Sloan Management Review, volume 56, no. 3 (Spring 2015), pp. 63-68

⁶¹ Տե՛ս ՀՀ ԱՎԾ, «Հայաստանի Հանրապետության սոցիալ-տնտեսական վիճակը 2016թ. հունվարի դրությամբ», 2016թ. էջ 53

ոլորտներում զբաղվածների միջին աշխատավարձն է: Հատկանշական է, որ հարցման մասնակիցների 47%-ը զբաղված է բարձր տեխնոլոգիաների բնագավառում գործող կազմակերպություններում, որոնք ներառված են տեղեկատվության կապի ոլորտում:

Այնուամենայնիվ, չնայած նման բարձր ցուցանիշներին, ՀՀ աշխատաշուկայում տվյալ ագեռներին միջին աշխատավարձն անհամեմատ ցածր է ԱՄՆ-ում ձևավորված համապատասխան աշխատավարձի մակարդակից. վերջինս, համաձայն O'Reilly-ի հաշվետվության, կազմում է ամսական 8,660 հազար ԱՄՆ դոլար: Այս երկու ցուցանիշները միմյանց հետ համադրելի դարձնելու նպատակով հաշվի են առնվել երկու երկրների սպառողական գների ինդեքսները և արդյունքում ստացվել է, որ ԱՄՆ-ում ամսական 8,660 հազար դոլար աշխատավարձը համապատասխանում է ՀՀ-ում ամսական 3,870 հազար դոլար աշխատավարձին: Այս գումարն ակնհայտորեն գերազանցում է տեղական շուկայի համար հաշվարկված միջին ցուցանիշը, ինչը նշանակում է, որ ՀՀ-ը միջազգային շուկայում կարող է գնային շոշափելի մրցակցային առավելություն ունենալ տվյալ ագիտական ծառայությունների մատուցման ասպարեզում:

Ելնելով այն հանգամանքից, որ ՄՏ-ի հավաքագրումն ու վերլուծությունը պահանջում է աշխատանքային, նյութական ու ֆինանսական ռեսուրսների զգալի ներդրումներ, ինչպես նաև հաշվի առնելով տվյալ ագեռների բարձր միջին աշխատավարձը, կարելի է ենթադրել, որ ՀՀ-ում տվյալ ագեռների պահանջարկ կներկայացնեն հիմնականում խոշոր կազմակերպությունները:

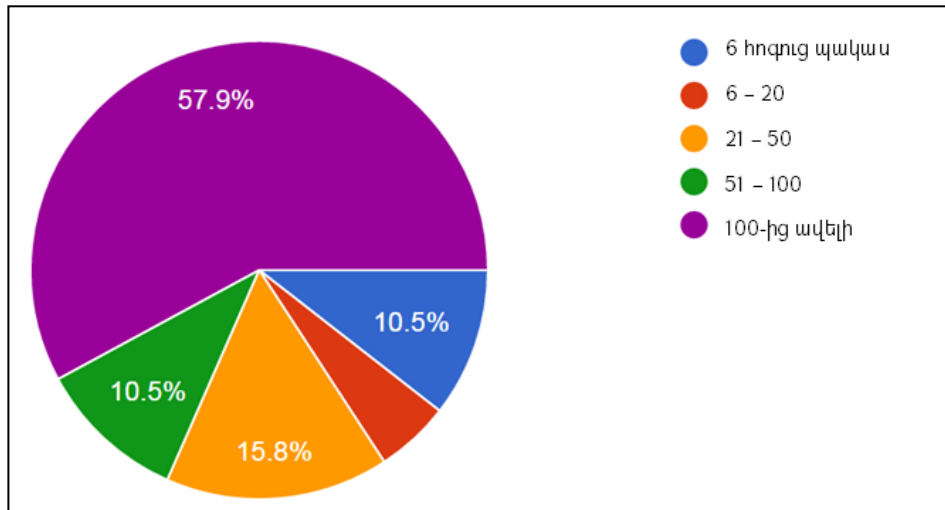
Հարցման արդյունքում պարզվել է, որ տվյալ ագեռների 57,9%-ը զբաղված են 100 և ավելի, իսկ 84,2%-ը՝ 21 և ավելի աշխատակից ունեցող ընկերություններում: 15,7%-ը ուսանող են կամ նորավարտ և դեռևս արակտիկական փորձառջան են անցնում: Հետաքրքրական է, որ հարցվողների միայն 15,8%-ի հաստիքի անվանումն է «Տվյալ ագետ», և բոլորն էլ աշխատում են 100-ից ավելի աշխատակից ունեցող կազմակերպություններում: Մնացած մասնակիցների հաստիքները կրում են ավելի նեղ մասնագիտական անվանումներ,

ինչպես, օրինակ՝ ֆինանսական տվյալների վերլուծաբան, միկրոբիոլոգիական տեխնոլոգիական ինժեներ և այլն:

Գծապատկեր 17.

Աշխատակիցների քանակը հարցմանը մասնակցած մասնագետների ընկերություններում

**կառուցվել է մեր կողմից*



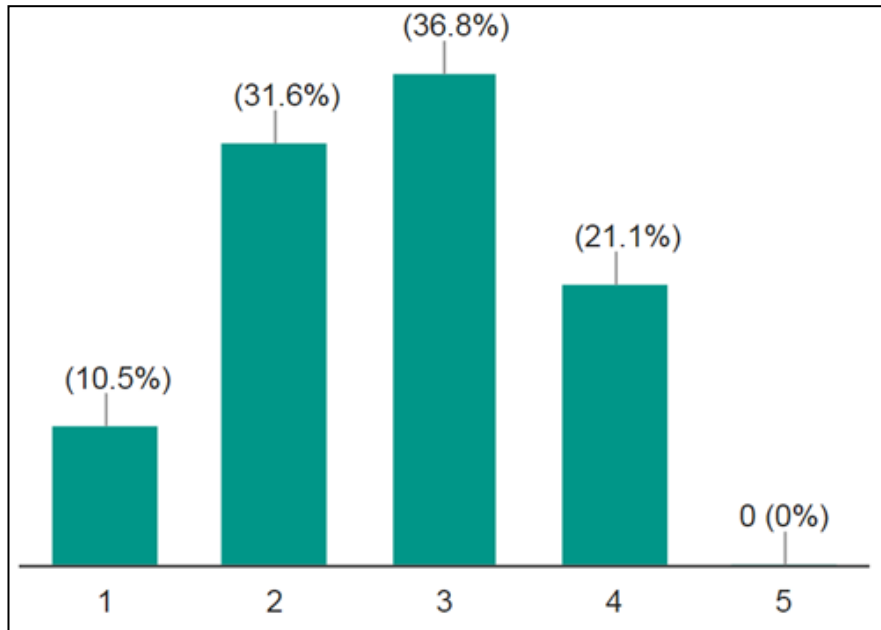
Հարցմանը մասնակցած բոլոր տվյալագետներն ունեն բարձրագույն կրթություն, իսկ նրանց 89,4%-ը՝ մագիստրոսի կամ գիտությունների թեկնածուի աստիճան: Բուհական մասնագիտացման տեսանկյունից առաջին տեղում են գտնվում տնտեսագետ-վիճակագիրները, ովքեր հարցվածների 36,8%-ն են: Երկրորդ տեղում ծրագրավորողներն ու մաթեմատիկոսներն են, որոնք կազմում են հարցման մասնակիցների 21,1%-ը: Այս փաստը ևս մեկ անգամ ընդգծում է տվյալագիտության զբաղվելու համար անհրաժեշտ երեք կարևոր՝ վիճակագրության, համակարգչագիտության և տնտեսագիտության ոլորտներում, ձեռք բերվող գիտելիքների զուգակշռված համակցման կարևորությունը:

O'Reilly-ի կողմից իրականացվող հարցման հարցաթերթին ավելացվել է ևս մեկ հարց: Տվյալագետները 1-5 սանդղակով գնահատել են իրենց իսկ հմտությունները: Հարցմանը մասնակցած տվյալագետների միջինացված ցուցանիշը 2,68 է: Այս թիվը հավասարապես հանդիսանում է միջին ցուցանիշ և՛ արական, և՛ իգական սեռի ներկայացուցիչների շրջանում:

Գծապատկեր 18.

Տվյալագետների հմտության սեփական գնահատականը

**Կառուցվել է մեր կողմից*



Ինչպես երևում է գծապատկեր 18-ից՝ տվյալագետների 10,5%-ն իրենց գիտելիքները գնահատել են ցածրագույն՝ 1 նիշով: Այս անձինք արակտիկաններ են, ովքեր, թերևս, նոր են ընտելանում ոլորտին, քանզի ներկայումս բուհական մասնագիտական կրթությունը դեռևս չի պատրաստում այդ ոլորտի մասնագետներ: Վերջինիս ապացույցն է նաև հանդիսանում այն փաստը, որ հարցման մասնակիցների 78,9%-ն իրենց ներկայիս հմտությունները ձեռք են բերել ինքնուսուցման եղանակով (մասնագիտական գրականություն, տարատեսակ «online» ուսումնական ծրագրեր և այլն):

Չարցաթերթի մեջ ընդգրկվել է նաև հարցառօրյայում US-ի վերլուծության նպատակով հարցվողի կողմից կիրառվող գործիքակազմի մասին: Գրանցվել են 27 տարբեր ծրագրային փաթեթների անվանումներ: Չարցվողների մեծամասնությունը՝ 52,4%-ը, տիրապետում է Python և R անվանումներով վիճակագրական ծրագրավորման լեզուներին, կամ՝ երկուսին միաժամանակ: Այս ցուցանիշը նույնպես համահունչ է O'Reilly-ի հաշվետվությանը: Վերջինս հավաստում է, որ վերոնշյալ ծրագրավորման լեզուները ոչ միայն ամենաօգտագործվողն են տվյալագիտության ոլորտում, այլև՝ ամենապահանջվածն են աշխատառուկայում: Ըստ էության սույն փաստը վկայում է տեղական աշխատառուկայում US-ի վերլուծության ամբողջական մասնագետների՝ համաշխարհային

զարգացման միտումներին տեղեկացվածության և դրանց հետևելու մասին:

Հարցաթերթում ներառված վերջին հարցը վերաբերում է տվյալ ազետների կողմից առօրյայում վերլուծվող տվյալների ծավալին: Վերջինս, վերոնշյալ այլ փոփոխականների հետ միասին, օգտագործվել է տվյալ ազետների եկամտային խմբի վրա ազդող փոփոխականները բացահայտելու նպատակով:

Տվյալ ազետների եկամտային խմբի վրա ազդող փոփոխականների ռեգրեսիոն վերլուծություն

Սույն ընտրանքային հարցման տվյալների հիման վրա իրականացվել է տվյալ ազետների եկամուտը բացատրող փոփոխականների ուսումնասիրություն⁶²: Ուսումնասիրության համար օգտագործվել են հարցաթերթում ընդգրկված բոլոր 12 հարցերից ստացված տվյալները, որոնցից 11-ը հանդիսանում են անկախ փոփոխականներ, իսկ 12-րդը՝ եկամտային խումբը, կախյալ փոփոխական: Մինչ ռեգրեսիոն մոդելին անցում կատարելը իրականացվել է տվյալների մշակում, որոնց վրա հնարավոր կլինի կիրառել մոդելը: Նախ և առաջ՝ ձևափոխվել են տեքստային արժեք ընդունող փոփոխականները և յուրաքանչյուր կատեգորիա ստացել է անվանական գնահատական (օրինակ՝ արական սեռի ներկայացուցիչներին վերագրվել է 1 արժեքը, իսկ իգական սեռին՝ 0): Այնուհետև իրականացվել են հետևյալ փոփոխությունները.

Այլ ուսակ 5.

Անկախ փոփոխականների անվանումները և նկարագրությունը

Փոփոխական	Նկարագրություն
ՏԱՐԻՔ	տվյալ ազետի տարիքը
ՍԵՌ	տվյալ ազետի սեռը
ՖՐՄ	ստեղծվել է ՖՐՄ անունով փոփոխական, որը ցույց է տալիս, արդյոք տվյալ ազետը իր գիտելիքների մի մասը ձեռք է բերել ֆորմալ կրթության միջոցով, թե ոչ
ԹԴ	ստեղծվել է ԹԴ անունով փոփոխական, որը ցույց է

⁶² Դավթյան, Հ., Տվյալ ազետների եկամտային խմբին ազդող փոփոխականների ռեգրեսիոն վերլուծություն: «Այլ ընտրանք» գիտական հանդես, №1, 2017 թ., էջեր 269-276

	տալիս, արդյոք տվյալ ազգետը ունի գիտություն ունեցող թեկնածուի կամ դոկտորի աստիճան (ցանկացած բնագավառում), թե ոչ
ՏԾ	այս փոփոխականը ստեղծվել է պարզելու, արդյոք տնտեսագիտության և/կամ ծրագրավորման ոլորտում կրթությունը ազդեցություն ունի աշխատավարձային խմբի վրա, թե ոչ
ԾԱՎ	ցույց է տալիս տվյալ ազգետի կողմից առօրյայում վերլուծվող տվյալների ծավալային խումբը
ՅՍՏ	Մեծ տվյալների հետ աշխատելու սեփական հմտություն ունեցող գնահատականը
ԱԾԽ	ընկերությունում աշխատակիցների քանակը
ՅԱՍՏ	այս փոփոխականը ստեղծվել է հաստիքների անվան հիման վրա, ցույց է տալիս, արդյոք տվյալ ազգետի հաստիքը պարունակում է «տվյալ» բառը, թե ոչ
ՖԻՆՏՏ	տվյալ ազգետի գործունեության ոլորտը նկարագրվում է ՖԻՆՏՏ փոփոխականով, որն ընդունում է 1 արժեք, եթե նա զբաղված է ֆինանսների կամ ՏՏ ոլորտում, և 0՝ այլ ասպես
ՓՈՐՁ	տվյալ ազգետի մասնագիտական փորձառությունը
ՅԵՇՏ	տվյալ ազգետի կարծիքով նոր աշխատանք գտնելու հեշտությունը

Նախնական ենթադրություններ այս փոփոխականներից միայն «ՅԵՇՏ» փոփոխականի գործակիցն է ակնկալվում ստանալ բացասական թիվ: Ենթադրության հիմքում ընկած է այն վարկածը, որ բարձր աշխատավարձով նոր աշխատանք գտնելն այդքան էլ հեշտ է: Մնացած բոլոր փոփոխականների գործակիցներից ակնկալվում է դրական նշան՝ ենթադրելով որ նրանց գոյությունն ու աճը դրականորեն կազդի տվյալ ազգետների աշխատավարձի վրա:

Մինչ ռեգրեսիոն վերլուծությանը անց կատարելը դիտարկվել են անկախ փոփոխականների միջև կախվածությունները՝ ստեղծելով կորելացիոն մատրից: Մատրիցը պատկերված է աղյուսակ 6-ում, որտեղ բարձր դրական կորելացիոն կախվածությունը ներկված է կարմիր, իսկ ուժեղ բացասական կախվածությունը՝ կապույտ գույնով:

Ինչպես երևում է աղյուսակից՝ անկախ փոփոխականների միջև անկախության անհրաժեշտ ենթադրությունը խախտված է, քանի որ ամենաուժեղ առկա բացասական կորելացիան ընդամենը $-0,48$ է, դրականը՝ $0,67$: Ավելին, այս ցուցանիշները առկա են միայն մի քանի փոփոխականի դեպքում, իսկ մնացյալ փոփոխականների պարագայում կորելացիայի գործակցի բացարձակ արժեքը բավականին մոտ է 0-ին:

Ինչպես նշվեց վերևում՝ ամենաբարձր դրական կորելացիան գրանցել է $0,67$ արժեք և գործում է ՓՈՐՁ և ՏԱՐԻՔ փոփոխականների միջև: Իհարկե, նման արդյունքը կանխատեսելի էր, քանզի տարիքի աճի հետ մարդիկ ավելի երկար փորձ են ձեռք բերում: Նույնը, թերևս, կարելի է ասել ՏԱՐԻՔ և ԹԴ փոփոխականների միջև գոյություն ունեցող կորելացիայի մասին ($0,61$), քանզի գիտությունների թեկնածուի և/կամ դոկտորի աստիճան ձեռք բերելու համար հետազոտական փորձ է անհրաժեշտ, որը ձեռք է բերվում տարիների ընթացքում:

Աղյուսակ 6.

Անկախփոփոխականների միջև կորելացիան

*կազմվել է մեր կողմից

	ՔԿՄՔ	ՍԵՌ	ՖՐՄ	ԹԲ	ՏԾ	ՃԱԿ	ՀԱՏ	ԱՇԽ	ՀԱՍ	ՖԻԼՏՏ	ՓՈՐՁ	ՀԵՇՏ
ՏԱՐԻՔ	1.00	0.30	-0.05	0.61	-0.01	0.00	0.06	-0.31	-0.21	-0.05	0.67	0.23
ՍԵՌ	0.30	1.00	0.05	0.22	-0.04	0.27	0.01	-0.25	0.41	0.26	0.37	0.51
ՖՐՄ	-0.05	0.05	1.00	-0.05	0.15	-0.31	-0.33	-0.48	0.04	-0.37	0.09	-0.04
ԹԲ	0.61	0.22	-0.05	1.00	-0.42	0.07	-0.14	-0.08	-0.15	-0.26	0.47	0.11
ՏԾ	-0.01	-0.04	0.15	-0.42	1.00	-0.38	-0.02	-0.15	-0.09	0.48	0.17	0.40
ՃԱԿ	0.00	0.27	-0.31	0.07	-0.38	1.00	0.62	0.03	0.49	0.06	-0.13	0.16
ՀԱՏ	0.06	0.01	-0.33	-0.14	-0.02	0.62	1.00	-0.04	0.46	0.36	-0.24	0.30
ԱՇԽ	-0.31	-0.25	-0.48	-0.08	-0.15	0.03	-0.04	1.00	-0.09	-0.08	-0.13	-0.18
ՀԱՍ	-0.21	0.41	0.04	-0.15	-0.09	0.49	0.46	-0.09	1.00	0.33	-0.39	0.38
ՖԻԼՏՏ	-0.05	0.26	-0.37	-0.26	0.48	0.06	0.36	-0.08	0.33	1.00	-0.09	0.58
ՓՈՐՁ	0.67	0.37	0.09	0.47	0.17	-0.13	-0.24	-0.13	-0.39	-0.09	1.00	0.20
ՀԵՇՏ	0.23	0.51	-0.04	0.11	0.40	0.16	0.30	-0.18	0.38	0.58	0.20	1.00

Աղյուսակ 6-ից պարզ է դառնում նաև, որ ֆինանսների և/կամ ՏՏ ոլորտի մասնագետները իրենց հնարավորությունները նոր աշխատանք գտնելու հարցում բարձր են գնահատում: Այս կորելացիան ամենայն հավանականությամբ այդ երկու ոլորտների զարգացման և հաստիքների քանակի աճի հետևանք է:

Հետաքրքրաշարժ է բացասական կորելացիան ՖՐՄ և ԱՇԽ փոփոխականների միջև: Այն ցույց է տալիս, որ գիտելիքները ֆորմալ կրթության միջոցով ձեռք բերած անձիք ավելի փոքր կազմակերպություններում են աշխատում քան ոչ ֆորմալ եղանակով ձեռք բերածները: Այս փաստը ամենայն հավանականությամբ հանդիսանում է մեծ կազմակերպությունների կողմից կազմակերպվող արակտիկաների և ուսուցման ծրագրերի հետևանք, որոնք մեր տվյալներում ներկայացվում է որպես ոչ ֆորմալ կրթություն: Ելնելով այն հանգամանքից, որ կախյալ փոփոխական հանդիսացող աշխատավարձային խմբերը օրդինալ փոփոխականներ են,

այն է՝ աշխատավարձը ներկայացված է ոչ թե բացարձակ արժեքով, այլ անվանական խմբերով, որոնցից ավելի մեծ բալ ունեցող խումբը վկայում է ավելի բարձր աշխատավարձի մասին՝ կիրառվել է հարակից լոջիթ մոդելը /ordinal logit/⁶³:

Նախնական ռեգրեսիոն մոդելում ներգրավվել են բոլոր փոփոխականները: Այնուհետև մեկ առ մեկ ռեգրեսիոն մոդելից դուրս են հանվել ոչ նշանակալի փոփոխականները՝ գնահատելով արդյունքում ստեղծված մոդելների փոփոխականների նշանակալիությունը: Ռեգրեսիոն մոդելից փոփոխականները դուրս են հանվել հետևյալ հերթականությամբ.

- | | | |
|----------|---------|---------|
| 1 ՅՏՄ, | 2 ՓՈՐՁ, | 3 ՅԱՍՏ, |
| 4 ՏԱՐԻՔ, | 5 ՍԵՌ, | 6 ԱՃԽ, |
| 7 ՖՐՄ: | | |

Ռեգրեսիոն մոդելին նման մոտեցումը հիմնավորված է այն հանգամանքով, որ երբեմն ոչ նշանակալի փոփոխականները «իրենց հետևից տանում են» նաև նշանակալի, բայց իրենց հետ բարձր կորելացիա ունեցող փոփոխականներին: Այս հետազոտության պարագայում ՅԵԾՏ փոփոխականը նույնպես վիճակագրորեն նշանակալի չէր սկզբնական մոդելում: Սակայն ՍԵՌ փոփոխականը (որոնց միջև կորելացիան 0,51 է) մոդելից հանելուց հետո այն ապահովեց վիճակագրական նշանակալիություն՝ 94% վստահության միջակայքում: Վերոնշյալ 7 փոփոխականները մոդելից հանելուց հետո վերջնական տարբերակը ներառում է մնացյալ 5 անկախ փոփոխականները: Այդ 5 անկախ փոփոխականների հիման վրա իրականացված ռեգրեսիոն վերլուծության արդյունքները ներկայացված են աղյուսակ 7-ում:

Աղյուսակ 7-ում ներկայացված փոփոխականներով իրականացված ռեգրեսիոն մոդելի դետերմինացիայի R^2 գործակիցը ունի 0,812 արժեք: Վերջինս բավականին բարձր ցուցանիշ է և ցույց է տալիս, թե որքանով են մոդելում ընդգրկված անկախ փոփոխականները բացատրում եկամտային խմբի վարիացիան: Այսպիսով, մեր

⁶³ Հարակից լոջիթ մոդելի փոխարեն կարելի էր նաև կիրառել հարակից փրոբիթ մոդելը: Երկու մոդելների դեպքում արդյունքները պետք է մոտոստացվեն, քանզի տարբերությունը կայանում է համախառն բաշխման ֆունկցիայի մեջ:

վերջնական մոդելը կախյալ փոփոխականի վարիացիան բացատրում է 81,2%-ով:

Աղյուսակ 7.

Փոխականների նշանակալիությունը՝ վերջնական մոդելում

**կազմվել է մեր կողմից*

	Գործակից 64	Ստանդարտ սխալ	P- արժեք	Նշանակալիություն 65
Հաստատուն 1	-4,0058	2,2864	0,0798	92% ՎՄ
Հաստատուն 2	-9,1171	3,2827	0,0055	99% ՎՄ
Հաստատուն 3	-10,3271	3,7077	0,0053	99% ՎՄ
Հաստատուն 4	-13,085	4,4449	0,0032	99% ՎՄ
ԹԴ	6,0541	2,371	0,0107	99% ՎՄ
ՏԾ	3,4025	2,0491	0,0968	90% ՎՄ
ԾԱԿ	2,4459	0,9346	0,0089	99% ՎՄ
ՀԵՇՏ	1,4734	0,7836	0,0601	94% ՎՄ
ՖԻՆՏՏ	-7,8627	2,8494	0,0058	99% ՎՄ

Ինչպես երևում է աղյուսակ 7-ից՝ մոդելում գոյություն ունեն 4 հաստատուն: Սա պայմանավորված է այն հանգամանքով, որ կախյալ փոփոխականը՝ եկամտային խումբը (ԵԿԽ), բաղկացած է 5 խմբից, որոնց բաժանման համար անհրաժեշտ է 4 հատման կետ (այն է՝ 4 հաստատունները): Նրանց նշանակալիությունը հիմնավորում է ԵԿԽ փոփոխականի 5 դասերի բաժանումը:

Ելնելով այն հանգամանքից, որ լոջիստիկ մոդելը օգտագործում է էքսպոնենցիալ բաշխում՝ «Գործակից» սյան մեջ ներկայացված արժեքները հանդիսանում են իրական գործակիցների լոգարիթմական արտահայտությունը: Այնուամենայնիվ, բնական հիմքով լոգարիթմը հանդիսանում է մոնոտոն ձևափոխություն և, հետևաբար, կարելի է օգտագործել աղյուսակ 7-ի «Գործակից» սյան արժեքները՝ ԵԿԽ փոփոխականի վրա անկախ փոփոխականների ազդեցության ուղղությունը գնահատելու համար: Ինչպես երևում է աղյուսակից՝ մոդելում ընդգրկված 5 փոփոխականներից 4-ը

⁶⁴ Այս գործակիցը դեռևս չի հանդիսանում վերջնական մոդելում ընդգրկվող գործակիցը, քանզի լոգիստիկ \$ ունկցիան ենթադրում է էքսպոնենցիալ բաշխվածություն:

⁶⁵ ՎՄ՝ վստահելիություն միջակայք

դրականորեն են ազդում տվյալ ազետների եկամտային խմբի պատկանելիության վրա: Չարմանալիորեն բացասական արժեք է ստացել \$ԻՆՏՏ փոփոխականի գործակիցը, որը բավականին բարձր վիճակագրական նշանակալիություն ունի (99% վստահելիության միջակայքում): Վերջինից բխում է, որ \$ինանսների և /կամ ՏՏ ոլորտում աշխատանքը բացասաբար է ազդում տվյալ ազետների եկամուտների վրա: Թերևս այս արդյունքը բացատրվում է \$ինանսների և ՏՏ ոլորտում առկա աշխատանքների մեծաքանակությամբ և, առաջարկի օրենքի համաձայն, բերում է ցածր գների, տվյալ դեպքում՝ եկամուտների: Սխալ էր նախնական ենթադրությունը ՅԵԾՏ փոփոխականի մասին, որը, ինչպես երևում է աղյուսակ 7-ից, դրական ազդեցություն ունի եկամտային խմբի վրա:

Այս հետազոտության արդյունքները որոշ չափով տարբերվում են O'Reilly կողմից իրականացված ռեգրեսիոն վերլուծության արդյունքներից⁶⁶: Վերջինս ցույց է տալիս, որ տարիքը, սեռը և կազմակերպության անձնակազմի մեծությունը նշանակալի ազդեցություն ունեն եկամուտների վրա: Ըստ հաշվետվության՝ իգական սեռը և փոքրաքանակ անձնակազմ ունեցող կազմակերպությունում աշխատելը բացասական են ազդում եկամուտների վրա, իսկ տարիքը և մեծ կազմակերպությունում աշխատելը՝ դրական: Մյուս կողմից, երկու վերլուծության արդյունքում էլ եկամտի վրա կարևորագույն դրական ազդեցություններից մեկը ունի ԹԴ փոփոխականը՝ տվյալ ազետի գիտական աստիճանը: Արդյունքները համապատասխան են նաև \$ԻՆՏՏ փոփոխականի նշանակալիության և ազդեցության ուղղության հետ կապված:

Մեկ այլ տարբերություն, որ առկա է այս երկու հետազոտությունների արդյունքների միջև հաստիքի անվանումն է: Ի տարբերություն մեր վերլուծության՝ ԱՄՆ շուկայում հաստիքի անվանման մեջ տվյալ ազետ բառի առկայությունը զգալիորեն (դրականորեն) ազդում է եկամուտների վրա:

Նման տարբերությունները, թերևս, կարող են հետևանք հանդիսանալ հետևյալ 4 գործոնների հետևանք.

⁶⁶ O'Reilly, "2015 Data Science Salary Survey", 2015, էջ . 6

1. **Մոդել ու մ ընդգրկված փոփոխականներ** – սույն հետազոտության ընդ որպես համեմատության հիմք վերցրել է O'Reilly-ի կողմից իրականացված հետազոտությանը, սակայն վերջինիս ավելացրել է որոշ փոփոխականներ, որոնք, ինչպես ցուցադրված է աղյուսակ 11-ում, նշանակալի ազդեցություն ունեն ռեգրեսիոն վերլուծության արդյունքների վրա:
2. **Կիրառված մոդել** – Մեր հետազոտությանն ու մ որպես կախյալ փոփոխական հանդես է գալիս եկամտային խումբը և որպես գնահատման մեթոդ ընտրվել է հարակից լոջիթ մոդելը: Իսկ ահա O'Reilly-ի կողմից իրականացված հետազոտության մեջ կախյալ փոփոխականը հենց եկամուտներն են, իսկ օգտագործվող մոդելը՝ գծային ռեգրեսիան իր փոքրագույն քառակուսիների գնահատականով:
3. **Բացարժակ վարիացիա** – սույն հետազոտության արդյունքում ստացվել է, ինչպես նշվեց վերևում, 81,2% դետերմինացիայի գործակից, այն ինչ O'Reilly կողմից կիրառված մոդելով բացարժակ է վարիացիայի ընդամենը 38,2%-ը:
4. **Աշխատող ու կայ ական տարբերություններ** – ՀՀ շուկան միանշանակ տարբերվում է ԱՄՆ և/կամ այլ զարգացած երկրների շուկաներից, ինչը երևում է նաև ՀՀ տվյալ ազետների սեռային, եկամտային և այլ փոփոխականների բաշխվածության տարբերությունը:

Տվյալ ազետների նշանակալի որոնք բարձր աշխատավարձը կազմակերպության համար ծախսատար է դարձնում ՄՏ-ի վերլուծության կիրառությունը: Այդ առումով գործարար կառավարման տեսանկյունից կարևոր է պարզել, թե արդյոք իրենց ընկերության խնդիրների լուծման համար նպատակահարմար է կիրառել ավանդական տվյալների վերլուծություն, թե վարձել ավելի բարձր վարձատրության մեջ տվյալ ազետի: Կազմակերպությանը պատրաստ կլինի հավելյալ վճարել տվյալ ազետի ծառայությունների դիմաց միայն այն դեպքում, երբ

տվյալ ագիտական վերլուծությունն ապահովի ավելի բարձր ճշգրտության արդյունքներ, քան ավանդական վերլուծությունն է: Այս նմիշարքայլ գործարարական հարցերի պատասխան տալու համար մշակվել են ՄՏ-ի օգտագործման գործիքակազմ, որը ներկայացված է հաջորդ գլխում:

**ԳԼՈՒԽՅ – ԿԱՌԱՎԱՐՄԱՆ ՈԼՈՐՏՈՒՄ ՄԵՃ ՏՎՅԱԼՆԵՐԻ
ՕԳՏԱԳՈՐԾՄԱՆ ԳՈՐԾԻՔԱԿԱԶՄԸ ԵՎ
ԱՐԴՅՈՒՆԱՎԵՏՈՒԹՅԱՆ ԲԱՐՁՐԱՑՄԱՆ
ՅԻՄՆԱԽԱԴԻՐՆԵՐԸ**

3.1 Տվյալների վերլուծության դասական և տվյալագիտական մեթոդների համեմատական արդյունավետությունը գործարար կառավարման ոլորտում

(վարկային տվյալների օրինակով)

Գործարար կառավարման ոլորտում տվյալագիտության կիրառության անհրաժեշտությունը, ներուժն ու արդյունավետությունը ընդգծելու, ինչպես նաև նպատակով հետազոտության սույն հատվածում ներկայացվում է տվյալների տնտեսագիտական վերլուծության դասական և տվյալագիտական վերլուծության համեմատական արդյունքները՝ միևնույն տվյալների օգտագործմամբ կանխատեսում անելու պարագայում: Վերլուծության հիմքում ընկած են 2-րդ գլխում նշված առանձնահատկություններն ու հիմնական տարբերությունները, որոնք սպասվում է, որ կհանգեցնեն արդյունավետության տարբեր գույանիչների:

Ելնելով այն հանգամանքից, որ տվյալագիտությունը (և ույն ինքը՝ Մեծ տվյալների վերլուծությունը) իր կիրառությունն է գտել առավելապես բանկային համակարգում և հեռահաղորդակցման ոլորտում՝ սույն հետազոտությունն օգտագործում է վարկային տվյալներ՝ տվյալների դասական և տվյալագիտական վերլուծությունների համեմատության համար^{67,68}: Օգտագործելով Յամբոլրգի համալսարանի տվյալների շտեմարանում առկա German Credit Data տվյալների բազան՝ դասական տնտեսաչափական և տվյալագիտական մեթոդներով փորձ է արվում դասակարգել վարկառուներին⁶⁹:

⁶⁷ Դավթյան, Յ., Տվյալների վերլուծության կիրառությունը ֆինանսական շուկայում: «Եվրոպական ակադեմիա» գիտական հոդվածների ժողովածու, №7, Երևան, 2016 թ., էջ 231-237

⁶⁸ [Ռազմավարական վերլուծության գնահատականներ, IBM ընկերության տարեկան հաշվետվություն, 2015](https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data))

⁶⁹ [https://archive.ics.uci.edu/ml/datasets/Statlog+\(German+Credit+Data\)](https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data))

Տվյալների բազայի նկարագրությունը

German Credit Data-ն պարունակում է 20 անկախ փոփոխականի մասին դիտարկումներ, որոնք վերաբերում են 500 նախկին վարկառուներին: Ամեն վարկառու գնահատվել է որպես «լավ» (357 անձ) և «վատ» (143 անձ): Նոր վարկառուները նույնպես կարող են գնահատվել այս 20 անկախ փոփոխականներով: Հետազոտության նպատակն է զարգացնել վարկերի վարկանշավորման կանոն, որը կարող է օգտագործվել նոր վարկառուների վարկունակությունը գնահատելու համար: Արդյունքների ստուգման համար գոյություն ունի երկրորդային տվյալների բազա, նույնպես տեղադրված տվյալների շտեմարանում: Վերոնշյալ 20 անկախ փոփոխականները նկարագրված են ստորև՝ աղյուսակ 8-ում:

Աղյուսակ 8.

German Credit Data տվյալների բազայում առկա անկախ փոփոխականների նկարագրությունը

Account.Balance	կարգային փոփոխական՝ կազմված 4 կարգից, ցույց է տալիս հայտատուի բանկային հաշվի բալանսը՝ դիմումներ կայացման օրվա դրությամբ
Credit.duration	վարկի տևողությունը՝ ամիսներով
Credit.history	կարգային փոփոխական՝ կազմված 5 կարգից, նկարագրում է վարկառուի վարկային պատմությունը
Purpose	անվանական փոփոխական՝ կազմված 10 խմբից, նշում է վարկ վերցնելու նպատակը (մեքենա, կահույք, կրթություն և այլն)
Credit.amount	վարկի մեծությունը
Savings.account	կարգային փոփոխական՝ կազմված 5 կարգից, ցույց է տալիս հայտատուի խնայողական հաշվի բալանսը՝ դիմումներ կայացման օրվա դրությամբ
Employment.length	կարգային փոփոխական՝ կազմված 5 կարգից, ցույց է տալիս վարկառուի

	աշխատանքային փորձը
Instalment	վարկառուի ամսական տնօրինելի եկամուտը:
Sex.marital.status	անվանական փոփոխական՝ կազմված 5 խմբից, ցույց է տալիս վարկառուի սեռը և ամուսնական կարգավիճակը
Guarantors	անվանական փոփոխական՝ կազմված 3 խմբից, ցույց է տալիս երաշխավորների և/կամ համավարկառուների առկայությունը
Present.residence.duration	ներկա բնակավայրում բնակվելու տարիները
Property	անվանական փոփոխական՝ կազմված 4 խմբից, ցույց է տալիս անշարժ գույքի առկայությունը որպես գրավ
Age	վարկառուի տարիքը
Concurrent.Credits	անվանական փոփոխական է՝ կազմված 3 խմբից, ցույց է տալիս, վարկառուի գուգահեռ ունեցած վարկերի առկայությունը
Housing	անվանական փոփոխական՝ կազմված 3 խմբից, ցույց է տալիս վարկառուի տաննկառմամբ սեփականության իրավունքը
No.of.Credits.at.this.Bank	վարկառուի՝ տվյալ բանկում ունեցած վարկերի քանակը
Occupation	կարգային փոփոխական՝ կազմված 4 կարգից, ցույց է տալիս վարկառուի աշխատանքային կարգավիճակը
No.of.people.provide	մարդկանց քանակը, ովքեր պարտավոր են ապահովել տեխնիկական սպասարկում
Telephone	կեղծ փոփոխական, ցույց է տալիս, արդյոք վարկառուն հեռախոս ունի, թե՛ ոչ

Foreign.worker	կեղծ փոփոխական, ցույց է տալիս արդյո՞ք վարկառուն օտարերկրացի աշխատող է, թե՛ ոչ
-----------------------	---

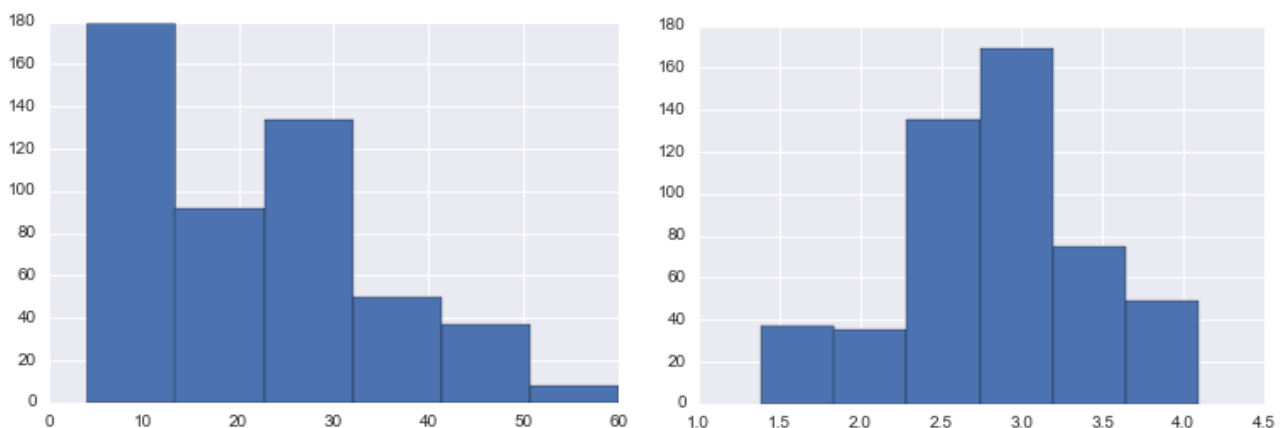
Աղյուսակ 8-ից երևում է, որ տվյալների բազայում առկա են մի շարք անվանական փոփոխականներ: Ի տարբերություն կարգային փոփոխականների՝ անվանական փոփոխականների դեպքում խմբերի միջև գոյություն ունեն կարգային կախվածություններ: Այդիսկ պատճառով անկախ կիրառվող մոդելից անվանական փոփոխականները պետք է նախօրոք վերածվեն կեղծ փոփոխականների: Բազմակուլի ներարկությունից խուսափելու նպատակով յուրաքանչյուր անկախ փոփոխականի համար, որն ունի N թվով խումբ կազմվել են N-1 քանակի կեղծ փոփոխականներ՝ որոնք ընդունում են 0 և 1 արժեքներ: N-րդ խումբը դիտարկվում է որպես 0-ական դեպք, այն է՝ գործ ունենք N-րդ խմբի հետ, երբ որ նորաստեղծ բոլոր N-1 կեղծ փոփոխականները ունեն 0 արժեք: Վերոնշյալ մոտեցումը կիրառելուց հետո տվյալների բազայում ստացել ենք 20 նոր կեղծ փոփոխական և ազատվել այն 8 անվանական փոփոխականներից, որոնց միջոցով այդ 20 կեղծ փոփոխականները ստացվել են:

Շարունակական արժեքներ ընդունող փոփոխականների համար ստուգվել է արդյո՞ք բաշխվածությունը Գաոսյան նորմալ է, թե ոչ: Անկախ փոփոխականները պարտադիր չէ, որ ունենան նորմալ բաշխում, սակայն արժեքների կենտրոնացվածությունը բաշխման աջ կամ ձախ ծայրերում կարող է հանգեցնել ծայրահեղ արժեքների ազդեցության անցանկալի մեծացման:

Գծապատկեր 19.

Credit.duration փոփոխականի բաշխվածությունը և ոգարիթմական ձևախոսքից առաջ (ձախ) և հետո (աջ)

**կառուցվել է մեր կողմից*



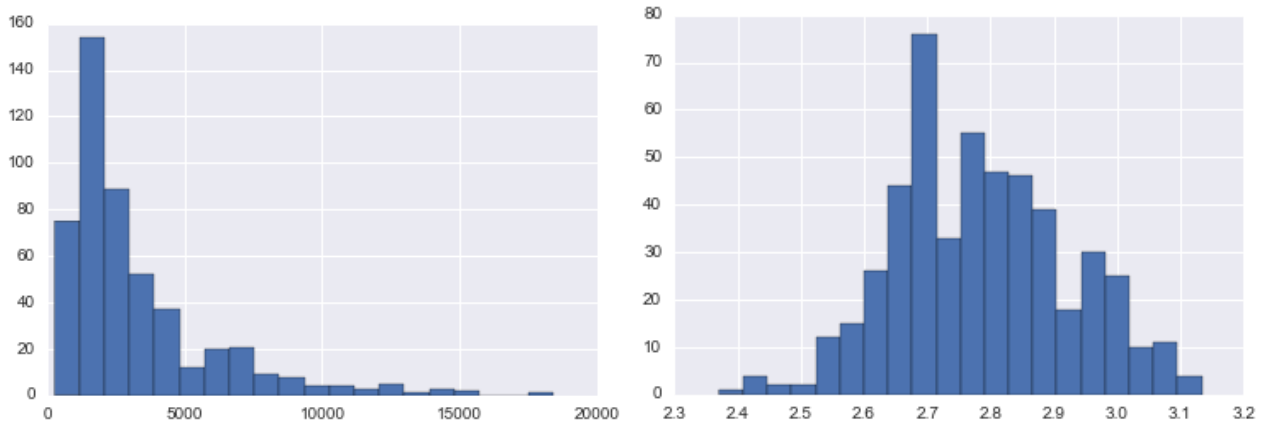
Ինչպես երևում է գծապատկեր 19-ից՝ Credit.duration փոփոխականը ունի ձախ հատվածում կենտրոնացված արժեքներ: Բաշխումը նորմալիզացնելու և ծայրահեղ արժեքների ազդեցությունը փոքրացնելու նպատակով կիրառվել է բնական հիմքով լոգարիթմական ձևափոխություն, որի արդյունքում ստացվել է նորմալ ինմոտբաշխում:

Գծապատկեր 20-ում և 21-ում ցուցադրված է նույն (բնական հիմքով լոգարիթմական) ձևափոխության դեպքում բաշխման փոփոխությունը համապատասխանաբար Credit.amount և Age փոփոխականների դեպքում:

Գծապատկեր 20.

Credit.amount փոփոխականի բաշխվածությունը լոգարիթմական ձևափոխությունից առաջ (ձախ) և հետո (աջ)

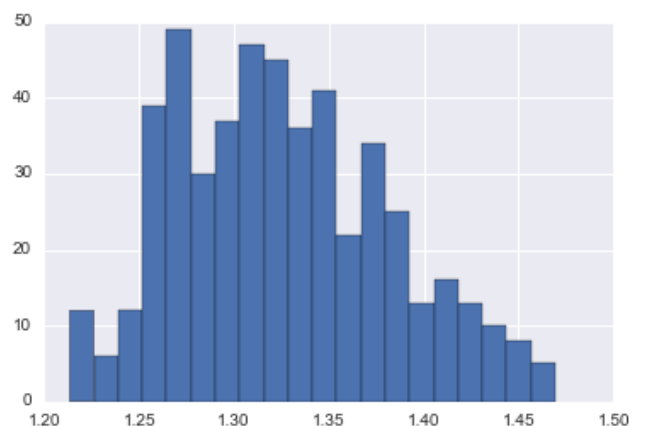
**կառուցվել է մեր կողմից*

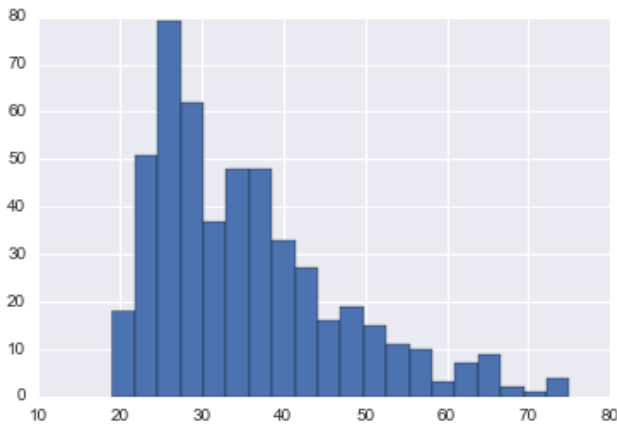


Գծապատկեր 21.

Age փոփոխականի բաշխվածությունը լոգարիթմական ձևափոխությունից առաջ (ձախ) և հետո (աջ)

**կառուցվել է մեր կողմից*





Բնական հիմքով և ոգարիթմական ձևափոխությունը բացի կիրառական նշանակությունից կրում է նաև տնտեսագիտական բովանդակություն: Այն ցույց է տալիս տվյալ փոփոխականի հարաբերական արժեքի

փոփոխությունը, որը շատ ժամանակ առավել կարևոր նշանակություն է ունենում, քան պարզապես բացարձակ մեծության փոփոխության ուսումնասիրությունը:

Մոդելի կիրառությանն անցնելուց առաջ տվյալների մշակման վերջին քայլն է անկախ փոփոխականների միջև կորելացիայի գոյություն ստուգումը: Ինչպես երևում է աղյուսակ 9-ում՝ առավել բարձր կորելացիան առկա է Credit.Amount և Credit.Duration փոփոխականների միջև՝ 57%: Այդ իսկ պատճառով, վերոնշյալ փոփոխականների համար հաշվարկվել է վարիացիայի մեծության գործակիցը՝ ստուգելու համար նրանց օգտագործման թույլ տվությունը մեր մոդելում: Արդյունքում պարզվել է, որ վարիացիայի մեծության գործակիցը ավելի փոքր է քան 5-ը, ինչը նշանակում է, որ 2 փոփոխականներն էլ կիրառելի են մոդելում:

Աղյուսակ 9.

Անընդհատարժեքները ուղուևող փոփոխականների կորելացիոն մարտիցը

**կազմվել է մեր կողմից*

	Credit.Duration	Credit.Amount	Age
Credit.Duration	100	57	-6
Credit.Amount	57	100	5
Age	-6	5	100

Գլուխ 2.1-ում ներկայացված դասական մոտեցման և տվյալ ագիտության հիմնական տարբերություններից մեկը բացակայող արժեքների նկատմամբ իրականացվող գործողություններն են: Ինչպես նշվեց՝ ավանդականորեն բացակայող արժեքները դիտվում են որպես խնդրնդոտ և ազատվում

են նրանցից: Սույն հետազոտության դասական վերլուծության պարագայում բացակայող արժեքները նույնպես դուրս կհանվեն տվյալների բազայից:

Տվյալների վերլուծությանը՝ դասական մոտեցմամբ

Վերոնշյալ փոփոխությունների արդյունքում կան 32 անկախ և մեկ կախյալ փոփոխական: Վերջինս ընդունում է միայն երկու արժեք՝ 1, եթե վարկառուն «լավ» է, և 2՝ հակառակ դեպքում: Տնտեսաչափական մոդելները, որոնք կիրառվում են ընդամենը երկու արժեք ընդունող կախյալ փոփոխականի դեպքում կոչվում են բինար ընտրության մոդելներ: Սույն աշխատությաննում կիրառվում է թերևս ամենահայտնի և լայն կիրառություն ունեցող բինար ընտրության մոդելը՝ Լոգիստիկ ռեգրեսիան: Այսպիսով, հետազոտվող մոդելը կունենա հետևյալ տեսքը.

$$\text{logit}(P_{ij}) = \log\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{32} X_{32} + \varepsilon,$$

որտեղ՝

$X_1 - X_{32}$ – մոդելի 32 անկախ փոփոխականներ,

$\beta_1 - \beta_{32}$ – անկախ փոփոխականների գործակիցներ, որոնք ենթակա են գնահատման,

P_{ij} – հայտատուի «վատ» վարկառու լինելու

հավանականությունը:

Հիշեցնենք, որ տվյալ հետազոտության նպատակն է, դասակարգել «լավ» և «վատ» հայտատուներին, և ոչ թե բացահայտել տարատեսակ փոփոխականների ազդեցության ուղղությունը, մեծությունը և նշանակալիությունը կախյալ փոփոխականի վրա: Այդ իսկ պատճառով, անկախ փոփոխականների հնարավոր ազդեցության մասին նախնական ենթադրություններ չեն արվի: Այնուամենայնիվ, կստուգվեն բոլոր փոփոխականների նշանակալիությունը, որպեսզի վերջնական մոդելում ընդգրկվեն միայն նշանակալի մոդելները:

Փոփոխականների գործակիցների գնահատման, ինչպես նաև նրանց ազդեցության նշանակալիության համար օգտագործվել է առավելագույն ճշմարտանմանության մեթոդը:

Գնահատման ընթացքում ընդհանուր մոդելի համար ստացվել է, որ $P=0.000$, ինչը նշանակում է, որ մոդելը նշանակալի է ($p_{արժեք} = 0.000 < \alpha = 0.05$): Մոդելի անկախ փոփոխականներից յուրաքանչյուրի նշանակալիությունը ներկայացված է ստորև՝ աղյուսակ 10-ում:

Անկախ փոփոխականների նշանակալիությունը

*կազմվել է մեր կողմից

Փոփոխական	P արժեք
Բացառություն	
Credit.Duration	$p_{value} = 0.121 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, վարկի տևողությունը նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Credit.Amount	$p_{value} = 0.031 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակցի արժեքը գնահատվել է 0.0469, որը նշանակում է, որ այս փոփոխականը վարկի հաստատման վրա դրական ձևով է ազդում:	
Instalment	$p_{value} = 0.016 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակցի արժեքը գնահատվել է 0.728, որը նույնպես դրական ձևով է ազդում կախյալի վրա:	
Guarrantors	$p_{value} = 0.687 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, երաշխավորների առկայությունը նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Present.residence.duration	$p_{value} = 0.730 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, այս փոփոխականը նույնպես նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Age	$p_{value} = 0.276 > \alpha = 0.05$

այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, վարկառուների տարիքը նույնպես նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Concurrent.Credits	
Այս փոփոխականը ընդունում է միայն մեկ արժեք, հետևաբար, այն կարելի է հանել մեր մոդելից:	
No.of.Credits.at.this.Bank	$p_{value} = 0.232 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, այս փոփոխականը նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Occupation	
Այս փոփոխականը ընդունում է միայն մեկ արժեք, հետևաբար, այն կարելի է հանել մեր մոդելից:	
No.of.people.provide	$p_{value} = 0.122 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, այս փոփոխականը նույնպես նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Telephone	$p_{value} = 0.715 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, հեռախոսի առկայությունը նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	
Foreign.worker	$p_{value} = 0.364 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն ոչ նշանակալի է: Այսպիսով, այս փոփոխականը նշանակալի ազդեցություն չի ունենում վարկի հաստատման որոշման մեջ:	

Ինչպես նշվեց վերևում՝ առկա անվանական փոփոխականները ներկայացվել են կեղծ փոփոխականների տեսքով: Առյուսակ 11-ում ներկայացված են կեղծ անկախ փոփոխականները, յուրաքանչյուրի նշանակալիությունը և բացատրությունը:

Առյուսակ 11.

Կեղծ փոփոխականների նշանակալիությունը

**Կազմվել է մեր կողմից*

Սկզբնական փոփոխական	
Կեղծ Փոփոխական	Նշանակալիություն
Բացատրություն	
Account.Balance	
accbal1	$P_{value} = 0.000 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 0.241 արժեք:	
accbal2	$P_{value} = 0.000 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 0.318 արժեք:	
Credit.History	
credhist1	$P_{value} = 0.000 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 0.149 արժեք:	
credhist2	$P_{value} = 0.010 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 0.414 արժեք:	
Purpose	
purpose1	$P_{value} = 0.000 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 6.513 արժեք:	
purpose2	$P_{value} = 0.025 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 2.254 արժեք:	
purpose3	$P_{value} = 0.078 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի չէ:	

Savings.Account	
savacc1	$P_{value} = 0.008 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 0.377 արժեք:	
savacc2	$P_{value} = 0.672 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի չէ:	
savacc3	$P_{value1} = 0.105 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի չէ:	
Employment.Length	
այս փոփոխականը լիովին վիճակագրորեն ոչ նշանակալի է:	
Sex.marital.status	
այս փոփոխականը լիովին վիճակագրորեն ոչ նշանակալի է:	
Property	
property1	$P_{value} = 0.024 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 4.006 արժեք:	
property2	$P_{value} = 0.255 > \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի չէ:	
property3	$P_{value} = 0.036 < \alpha = 0.05$
այս փոփոխականը վիճակագրորեն նշանակալի է: Նրա գործակիցն ընդունում է 3.391 արժեք:	
Housing	
այս փոփոխականը լիովին վիճակագրորեն ոչ նշանակալի է:	

Առանձնացնելով նշանակալի ազդեցությունն ունեցող փոփոխականները ոչ նշանակալի ազդեցությունն ունեցողներից՝ վերջնական մոդելում մնում են 2 անկախ փոփոխական, 2 ամբողջական անվանական փոփոխական (ներկայացված մի քանի կեղծ փոփոխականների տեսքով) և 3 մասնակի անվանական փոփոխական (անվանական փոփոխականներ, որոնցից ստացված կեղծ փոփոխականներից մեկը կամ մի քանիսը ոչ նշանակալի են և չեն ընդգրկվել վերջնական մոդելում): Մոդելի տեսքը հետևյալն է.

$$\begin{aligned} \text{logit}(P_{\text{վառ}}) &= \log\left(\frac{P_{\text{վառ}}}{1 - P_{\text{վառ}}}\right) \\ &= 0.0469 * \text{Credit.Amount} + 0.728 * \text{Instalment} + 0.241 * \text{accbal1} + 0.318 \\ &\quad * \text{accbal2} + 0.149 * \text{credhist1} + 0.414 * \text{credhist2} + 6.513 * \text{purpose1} \\ &\quad + 2.254 * \text{purpose2} + 0.377 * \text{savacc1} + 4.006 * \text{property1} + 3.391 \\ &\quad * \text{property3} \end{aligned}$$

Օգտագործելով վերոնշյալ մոդելը կարելի է դասակարգել ամևտրային վարկի նոր հայտատուի «լավ» կամ «վատ» լինելը: Օգտագործելով միևնույն շտեմարանում ամկա նոր վարկառուների մասին տվյալները՝ կատարվել է նրանց դասակարգումն ըստ մոդելի և համեմատվել իրական արդյունքի հետ: **Մեր կողմից իրականացված վերլուծության արդյունքում պարզվել է, որ տվյալ մոդելը կատարում է կանխատեսում 74.8% ճշգրտությամբ, այն է՝ սխալ դասակարգում է տեղի ունեցել մոտ 25.2% դեպքերում:**

Տվյալների վերլուծության ունը՝ տվյալ ագիտական մոտեցմամբ

Որպեսզի հնարավոր լինի համեմատել դասական և տվյալ ագիտական կանխատեսման արդյունքները միևնույն խնդրի լուծման պարագայում՝ հետազոտության սույն հատվածում կօգտագործվեն միևնույն վարկային տվյալները՝ դասակարգելով վարկառուներին տվյալ ագիտական մեթոդաբանությամբ:

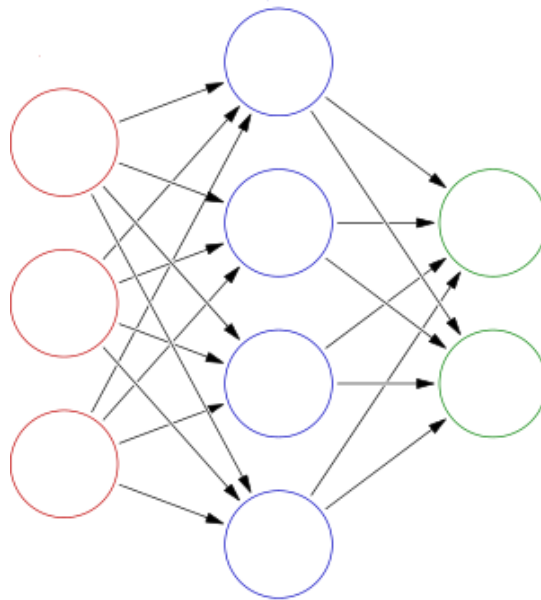
ՄՏ-ի վերլուծության համար կիրառվող մոդելները բազմազան ու բազմատեսակ են: Թերևս տվյալ ագիտական ամենատարածված մոդելը, որը հենց ընկած է խորը ուսուցման հիմքում, Նեյրոնային ցանցերի մոդելն է, որն էլ հենց կիրառվել է դասական եղանակով կատարված վերլուծության արդյունքների հետ համեմատության նպատակով:

Ինչպես արդեն նշվել է՝ տվյալ ագիտական հետազոտության պարագայում բացակայող արժեքներին յուրահատուկ ուշադրություն է դարձվում: Այդ տրամաբանությամբ առաջնորդվելով տվյալներում ամկա անվանական փոփոխականների բացակայող արժեքները վերագրվել են իրենց իսկ մոդայով, կարգային փոփոխականներինը՝ մեդիանայով: Անընդհատ արժեքները նդունող փոփոխականների դեպքում բացակայող արժեքներին տրվել է միջին արժեքը:

Չնայած այն հանգամանքին, որ բացակայող արժեքների նկատմամբ առանձնահատուկ վերաբերմունքը մեկ քայլով ավելացնում է տվյալ ագիտական վերլուծությանը, հարկավոր է նշել, որ մնացած տեսանկյուններից տվյալ ագիտական հետազոտության ավելի քիչ աշխատատար է: Տվյալ հետազոտության օրինակով, բացարձակապես պարտադիր չէ անկախ փոփոխականների բաշխվածությունը ստուգել և տրանսֆորմացնել նրանց՝ նորմալ բաշխում ստանալու նպատակով: Ավելին, մոդելի կիրառության ընթացքում չեն ստուգվում անկախ փոփոխականների նշանակալիությունը և նրանց ազդեցության մեծությունն ու ուղղությունը կախյալ փոփոխականի վրա:

Գծապատկեր 22.

Նեյրոնային ցանցի օրինակ



Գծապատկեր 22-ում ներկայացված է, թե ինչպես է աշխատում նեյրոնային ցանցը: Առաջին ուղղահայաց շարքում գտնվող երեք կարմիր շրջանները մոդել ներմուծվող արժեքներն են՝ անկախ փոփոխականները (տվյալ հետազոտության պարագայում դրանց թիվը հավասար է 32-ի): Երկրորդ շարքում գտնվող կապույտ շրջանները ցույց են տալիս այն տարբեր մոդելները, որոնք իրականացնում են գնահատում՝ անկախ փոփոխականների տարբեր կշռված արժեքների պարագայում: Վերջապես երրորդ ուղղահայաց սյան մեջ գտնվում են

կանաչ շրջանները, որոնք վերադարձնում են կախյալ փոփոխականի գնահատականը:

Մեր կողմից իրականացված վերլուծության արդյունքում պարզվել է, որ կիրառելով այդ մոդելը վարկային տվյալների բազայի վրաստանում ենք կանխատեսման 82.4% ճշգրտություն, ինչը նշանակում է, որ սխալ են գնահատվել հայտատուների միայն 17.6%:

Ինչպես երևում է այս օրինակից՝ նեյրոնային ցանցերի կիրառումը թույլ է տալիս բարձրացնել մոդելի ճշգրտությունը ամբողջ 7.6%-ով: Սա, իհարկե, վերջնական ցուցանիշ չէ, քանի որ մոդելում կիրառվող տարբեր պարամետրերի փոփոխության արդյունքում հնարավոր է ճշգրտության ցուցանիշների բարելավում⁷⁰: Ավելին, գոյություն ունեն բազմազան և բազմատեսակ մոդելներ, որոնք հնարավոր է ավելի բարձր ճշգրտություն ապահովեն: Այնուամենայնիվ, այս հետազոտության նպատակն է ձեռնարկել ուշադրությունը սևեռել տվյալ ազիտական մեթոդաբանական առանձնահատկություններին և ցույց տալ, վարկային տվյալների օրինակով, որ տվյալ ազիտական մոտեցումը կարող է հեշտ և ավելի արդյունավետ լինել, քան դասական տնտեսաչափական մոտեցումն է: ճշգրտության ցուցանիշի 7.6%-ով բարելավումը մրացակցային առեւտրի առավելություն կարող է հանդիսանալ գործարար ոլորտում՝ մեծացնելով ընկերության հասույթը կամ փոքրացնելով ծախսերը:

Չնայած վերոնշյալ մեթոդաբանական հեշտությանը՝ տվյալ ազիտական մոտեցումը հիմնականում կիրառելի չէ դասական տվյալների վերլուծության փաթեթների միջոցով: Խորը ուսուցման մոդելների և մեթոդների կիրառման համար ծրագրավորման իմացությունը անհրաժեշտ պայման է: Այսպիսով, կարելի է եզրակացնել, որ տվյալ ազիտության մեջ հաջողությամբ զբաղվելու համար անհրաժեշտ պայման են հանդիսանում ծրագրավորման և վիճակագրության, ինչպես նաև տնտեսագիտության և գործարարության գիտելիքները: Վերջինս հնարավորություն կտա մասնագետին լիովին տիրապետել այն գործարար խնդրի առանձնահատկություններին, որի լուծման համար էլ հենց

⁷⁰ Օրինակ, հնարավոր է ստանալ ճշգրտության այլ արդյունք՝ փոփոխելով երկրորդ ուղղահայաց սյան մեջ գտնվող մոդելների քանակը:

կիրառվում է տվյալ ազդեցությունը: Թերևս վերոնշյալ անհրաժեշտ պայմաններն էլ հենց պայմանավորել են տվյալ ազդեցության սակավ առաջարկն ու բարձր միջին աշխատավարձը, որի մասին փաստում են ենթազույգ 2.2-ում ներկայացված վերլուծության արդյունքները:

Այսպիսով, ՄՏ-ի վերլուծությանը՝ ավելի բարձր ճշգրտության ցուցանիշ ապահովելու միջոցով, կարող է մրցակցային առավելության/ շահույթի մեծացման հիմք հանդիսանալ: Այնուամենայնիվ, տվյալ ազդեցության բարձր աշխատավարձը վերջինիս կիրառությունը ծախսատար է դարձնում ընկերության համար: Այդ իսկ պատճառով, հիմնվելով ՄՏ-ի վերլուծության միջոցով բարձր կանխատեսողական ճշգրտության ներուժի վրա, հետազոտության հաջորդ հատվածում ներկայացվում է գործիքակազմ, որը հնարավորություն կտա ընկերությանն անհրաժեշտ գնահատել ՄՏ-ի օգտագործման արդյունավետությունը և բացահայտել վերջինիս բարձրացման համար անհրաժեշտ ներդրումների ծավալը: Գործիքակազմը կիրառվել է ՀՀ 19 կազմակերպությունների համար՝ օգտագործելով որակական հարցման միջոցով հավաքագրված տվյալները:

3.2 Մեծ տվյալների օգտագործման արդյունավետության մադիցը. ՀՀ կազմակերպությունների կողմից Մեծ տվյալների օգտագործման համեմատական արդյունավետության գնահատում

Մեծ տվյալների օգտագործման արդյունավետության մադիցը Տվյալների մեծություն գնահատական (SUF)

Սույն հետազոտության առաջին գլխում ներկայացվեց ՄՏ-ի երկրորդը՝ ուսումնասիրության հիմքում դնելով Դուգլաս Լանեյի կողմից տրված երեքի V-երի դասական սահմանումը: Այնուհետև ներկայացվեց ՄՏ-ի և «ավանդական» տվյալների միջև տարբերությունը: Հետազոտության սկզբնական շրջանում պարզ դարձավ, որ մի շարք մենեջերներ չեն կարողանում հստակ տարանջատել, արդյոք իրենց կազմակերպության կողմից հավաքագրվող տվյալները կարող են դասվել ՄՏ-ի շարքին, թե ոչ: Խնդիրն ուսումնասիրելուց հետո բացահայտվեց, որ մինչ այժմ

գոյությունն չունի ՄՏ-ի «մեծություն» չափման քանակական գնահատական, որը հնարավորություն կտա տարանջատել ՄՏ-ին՝ դասականից: Հետազոտության այս հատվածում ներկայացվում է տվյալների մեծության գնահատականը (ՏՄԳ), որը թույլ է տալիս չափել «մեծությունը»՝ հիմնվելով ՄՏ-ի դասական սահմանման վրա:

Համաձայն Դուգլաս Լանեյի՝ ՄՏ-ն նկարագրվում են մեծ ծավալով, բարձր շարժունություն և խիստ բազմազանություն: ՏՄԳ-ն հանդիսանում է այդ 3 ցուցանիշների թվաքանակական միջինը՝ հավասար կշիռներ տալով 3 բնութագրիչներին էլ: Պարզեցման համար յուրաքանչյուր բաղադրիչ չափվում է 0-ից 4 սանդղակով, որից բխում է, որ ՏՄԳ-ն կարող է ընդունել արժեքներ 0-ից 4 փակ միջակայքում:

Ճավալ

Տվյալների ծավալի մեծության չափման համար օգտագործվում է բայթերի վրա հիմնված համակարգը, որը Ստանդարտացման միջազգային կազմակերպության կողմից ճանաչված է որպես տվյալների ծավալի մեծության չափման միջազգային համակարգ (ISO/IEC 8000)⁷¹:

Աղյուսակ 12-ում ներկայացված է յուրաքանչյուր ծավալային խմբին վերագրվող բալը, որն աճում է ծավալի չափման միավորի աճին գուցե ընթաց: Աղյուսակի երկրորդ սյունը տալիս է տվյալ դասի վերլուծական բարդության կարճ բացատրությունը:

Աղյուսակ 12.

Մեծ տվյալների ծավալի չափման բալային համակարգ

**կազմվել է մեր կողմից*

Ծավալ	Բացատրություն	Բալ
Մեզաբայթ	Ազատորեն կարող են վերլուծվել դասական ծրագրային փաթեթների կողմից	1
Գիգաբայթ	Դժվարություն բայց կարող են վերլուծվել դասական ծրագրային փաթեթների կողմից	2

⁷¹ [ISO 8000-1 Quantities and units. Part 1: General \(1st ed.\). Switzerland: ISO \(the International Organization for Standardization\). 2009-11-15](#)

Տեռաբայթ	Վերլուծելու, ինչպես նաև պահպանելու համար անհրաժեշտ է օգտվել ամպային լուծումներից ⁷²	3
Պետաբայթ (և ավելի)	Վերլուծելու, ինչպես նաև պահպանելու համար անհրաժեշտ է օգտվել լայնածավալ բաշխողական համակարգերից ⁷³	4

Շարժունություն

V2 բնութագրիչը՝ տվյալների շարժունությունը, չունի ստանդարտացված դասակարգման համակարգ: Սակայն պարզ են շարժունականության երկու բևեռները՝ ստատիկ և իրական ժամանակահատվածում հավաքագրվող տվյալներ: Աղյուսակ 2-ում վերջիններիս վերագրված է համապատասխանաբար 1 և 4 բալ: Մյուս երկու միջանկյալ դեպքերը դասակարգվել են հեղինակի կողմից որպես դիսկրետ ժամանակային շարք (օր.՝ Ազգային վիճակագրական ծառայության կողմից համայնքներից ամենամսյա կտրվածքով հավաքագրվող տվյալներ) և գրեթե իրական ժամանակահատվածում հավաքագրվող տվյալներ (օր.՝ մասնաճյուղեր ունեցող կազմակերպության տարբեր մասնաճյուղերից ամենօրյա գլխամաս ուղարկվող հաշվետվություններ):

Աղյուսակ 13.

Մեծ տվյալների շարժունության չափման բալային համակարգ

**կազմվել է մեր կողմից*

Շարժունություն	Բացատրություն	Բալ
Ստատիկ	Կոնկրետ ֆիքսված (մեկ) ժամանակաշրջանի մասին տեղեկատվություն պարունակող տվյալներ	1

⁷² Mell P., Grance T., “The NIST Definition of Cloud Computing (Technical report)”. National Institute of Standards and Technology: U.S. Department of Commerce, September 2011

⁷³ Cappello F., Djilali S., Fedak G., Herault T., Magniette F., Nérib V., Lodygensky O. , “Computing on large-scale distributed systems: XtremWeb architecture, programming models, security, tests and convergence with grid”, Elsevier March 2005

Դիսկրետ ժամանակային շարք	Կոնկրետ ֆիքսված ժամանակահատվածի մասին տվյալներ՝ ներկայացված ըստ տարիների, ամիսների, օրերի կամ մեկ այլ հաճախականության	2
Գրեթե իրական ժամանակահատված	Իրական ժամանակահատվածի գործողությունների մասին որոշակի ժամանակային լագով հավաքագրվող տվյալներ	3
Իրական ժամանակահատված	Իրական ժամանակահատվածում ակնթարթորեն հավաքագրվող տվյալներ	4

Բազմազանություն

Վերջին բնութագրիչը, ինչպես շարժունությունը, նույնպես չունի ստանդարտացված դասակարգում: Աղյուսակ 14-ում ներկայացված դասերի և յուրաքանչյուր դասին վերագրված բալային համակարգի համար հիմք են հանդիսացել տվյալների աղբյուրների հանդիպման հաճախականությունը և վերլուծման բարդության աստիճանը: Առավել հաճախաբար մացվող, անընդհատ շարժի մեջ գտնվող, տարատեսակ գործարար որոշումների կայացման համար վերլուծական բարձր արժեք ունեցող, սակայն առավել դժվար վերլուծվող տվյալներին (օր.՝ սոցիալական ցանցերի անընդհատ թարմացվող հրապարակումներ, մեկնաբանություններ և այլն) տրվել է առավել ագույն գնահատականը՝ 4 բալ:

Աղյուսակ 14.

Մեծ տվյալների բազմազանության չափման բալային համակարգ

**կազմվել է մեր կողմից*

Բազմազանություն	Բացատրություն	Բալ
Աղյուսակային	Աղյուսակային տվյալներ՝ պարունակող դիտարկումներ մի քանի փոփոխականի մասին	1
Տվյալների	Բազմափոփոխական տվյալների	2

շ տե մ ար ան	շ տե մ ար ան ն եր	
Նկար, ձայնագրող թյուրև, կայքէջ	Վեբկայքի այցելությունների մասին, նկարների պիքսելներից և ձայնագրող թյուրևների ձայնային հաճախականությունից ստացված տվյալներ	3
Տեքստային, տեսահոլովակային, շարժական տեղանքային	Սոցիալական կայքերում արված մեկնաբանություններից, հրապարակված տեքստերից, աշխարհագրական դիրքի մասին շարժական հեռախոսակապից, տեսահոլովակից պիքսելային և ձայնային հաճախականության տվյալներ	4

Հարկավոր է նշել, որ US-ի 3 բնութագրիչները միմյանց հետ փոխկապակցված են: Երբ տվյալները բազմազան են և/կամ ունեն բարձր շարժունություն, նրանց պահպանելու համար անհրաժեշտ ծավալը ավելի մեծ է լինում: Եվ հակառակը, ավելի մեծ ծավալով տվյալները *սովորաբար* պարունակում են ավելի մեծ բազմազանություն և շարժունություն: Այդ իսկ պատճառով, տվյալների մեծության գնահատականը կընդունի 0 արժեք միայն տվյալների բացակայության դեպքում, իսկ առկայության դեպքում ընդունվող փոքրագույն արժեքը կլինի 1-ը: Հասկանալու համար եկեք կատարենք հաշվարկը «փոքրագույն» տվյալների համար: Եթե տվյալներն առկա են, ապա նրանք ինչ-որ ծավալ են զբաղեցնում, որը փոքրագույն դեպքում կարելի է գնահատել 1 բալով: Նմանապես, ամենապարզ դեպքում տվյալները կարելի է արտապատկերել աղյուսակի միջոցով (1 բալ): Իսկ շարժունության բացակայությունը նկարագրվում է ստատիկ վիճակով, որին նույնպես համարժեք է 1 բալը: Այդ դեպքում.

SUQ

$$= \frac{\text{Մավալի գնահատական} + \text{Շարժունության գնահատական} + \text{Բազմազանության}}{3}$$
$$= \frac{1 + 1 + 1}{3} = 1$$

Ըստ US-ի դասական սահմանման՝ վերջինս պետք է ունենա մեծ ծավալ, բարձր շարժունություն և մեծ բազմազանություն: Յետևաբար, որպեսզի տվյալները բավարարեն US-ի սահմանմանը, բնութագրիչներից յուրաքանչյուրը պետք է ընդունի 3 կամ 4 բավարժեք, այն է՝ մեծ 2-ից: Ինչպես գիտենք, 2-ից մեծ երեք ցուցանիշի միջին թվաբանականը նույնպես մեծ է 2-ից: Արդյունքում, Մեծ տվյալ կարող են հանդիսանալ միայն այն տվյալները, որոնց համար $SUQ > 2$:

SUQ մոտեցումը թույլ է տալիս քանակապես գնահատել տվյալների մեծությունը և հասկանալ արդյոք ընկերությունը գործունի US-ի հետ, թե դասական տվյալների: Գործարար կառավարման տեսանկյունից նման տեղեկատվությունը յուրահատուկ կարևորություն է կրում: Նախ՝ պետք է նշել, որ US-ն պահպանելու, նրանց մշակելու համար պահանջվող տեխնոլոգիաները տարբերվում են դասական տեխնոլոգիաներից (և ավելի բարձր արժեք ունեն): Այդիսկ պատճառով, ձեռներեցը, օգտագործելով SUQ-ն, զերծ կմնա հավելյալ ծախսերից՝ մինչև ունեցած տվյալները անցնեն $SUQ = 2$ շեմը (այն է՝ դասակարգվեն որպես Մեծ տվյալներ): Մյուս կողմից, US-ի հետ աշխատող մասնագետները նույնպես ունեն կրթության և գիտելիքների տարբեր մակարդակ և հետևաբար՝ տարբեր (ավելի բարձր) վարձատրություն՝ քան ավանդական տվյալների վերլուծաբանները: SUQ-ն թույլ կտա որոշել, թե որ մասնագետն է ավելի նպատակահարմար աշխատանքի ընդունել՝ օպտիմալ ացնելով աշխատավարձային վճարները: Եվ վերջում, ձեռներեցը ժամանակի ինչ-որ պահի կարող է ցանկություն ունենալ վճարել և ձեռք բերել տվյալներ, որոնք կկարողանա օգտագործել որոշումների կայացման գործընթացում: Երկու այլընտրանքների միջև որոշում կայացնելու լավ գործիք է SUQ-ն: Ձեռներեցը, համասեռ գնի պարագայում, պետք է ցանկանա ձեռք բերել այն տվյալները, որոնք ունեն ավելի բարձր SUQ, քանզի դրանք ավելի

շատ տեղեկատվություն կարող են պարունակել, քան ավելի ցածր SUԳ ունեցող տվյալները:

Այնուամենայնիվ, US-ի միայն գոյությունը կազմակերպության համար հաջողության աղբյուր չի կարող հանդիսանալ: Վերջինս միայն արդյունավետ կառավարման համար անհրաժեշտ պայման է, քանզի առանց տվյալների հնարավոր չէ չափել ի որոշումներ կայացնել: Կառավարման արդյունավետության բարձրացման համար պետք է նաև գոյություն ունեցող US-ն արդյունավետ օգտագործել:

Տվյալների կիրառության գնահատական (ՏԿԳ)

Մինչ օրս գոյություն ունի չունի ընդունված մեթոդ, որն օգտագործելով ձեռներեցը կարող է պարզել, արդյոք իր ձեռնարկությունում US-ն արդյունավետ են օգտագործվել, թե ոչ: Յետագոտության այս հատվածում ներկայացվում է US-ի օգտագործման արդյունավետության մատրիցը (USOU), որը հիմնված է SUԳ-ի և Տվյալների կիրառության գնահատականի (ՏԿԳ) վրա: Այն թույլ է տալիս գնահատել կազմակերպության կողմից US-ի օգտագործման արդյունավետությունը և պլանավորել վերջինիս հետևյալ հետազաներդրումային քայլերը:

Յետագոտության առաջին գլխում ներկայացվեց US-ի կիրառության դասակարգումն ըստ 3 տարբեր մոտեցումների: ՏԿԳ հաշվարկման համար հիմք է հանդիսանում Ցուցանիշների հաշվարկման համակարգի (ՑՀՀ) 4 առանցքներից յուրաքանչյուրում US-ի կիրառությունը: Այս մոտեցումն ընտրվել է այն պարզ պատճառով, որ կիրառության դասակարգման մյուս երկու մոտեցումները (ըստ գործունեության ոլորտի և ըստ տվյալների հավաքագրման աղբյուրների) կախված են կազմակերպության յուրահատուկ գործունեության, մինչդեռ ՑՀՀ մոտեցումը կիրառելի է ցանկացած ոլորտում գործող և ցանկացած տեսակի աղբյուրից տվյալներ հավաքագրող ընկերության համար:

Այսպիսով, ՏԿԳ-ն հանդիսանում է 4 առանցքներից յուրաքանչյուրում US-ի կիրառության գնահատականների միջին թվաբանականը: ՑՀՀ առանցքներից յուրաքանչյուրում, ինչպես նաև SUԳ դեպքում, գնահատումն իրականացվում է 0-4 սանդղակում: Օբալը

ցույց է տալիս կիրառության բացարձակ բացակայությունը: 1 և 4 բալերը համապատասխանում են մինիմալ և ամբողջական կիրառությանը: Առանցքներից յուրաքանչյուրում կիրառությունը պատկերելու համար օգտագործվել է հիերարխիկ բուրգի մոտեցումը, որը ցույց է տալիս, որ յուրաքանչյուր հաջորդ մակարդակին հասնելը հնարավոր է միայն նրան նախորդող փուլերն անցնելու միջոցով:

Յաճախորդներ

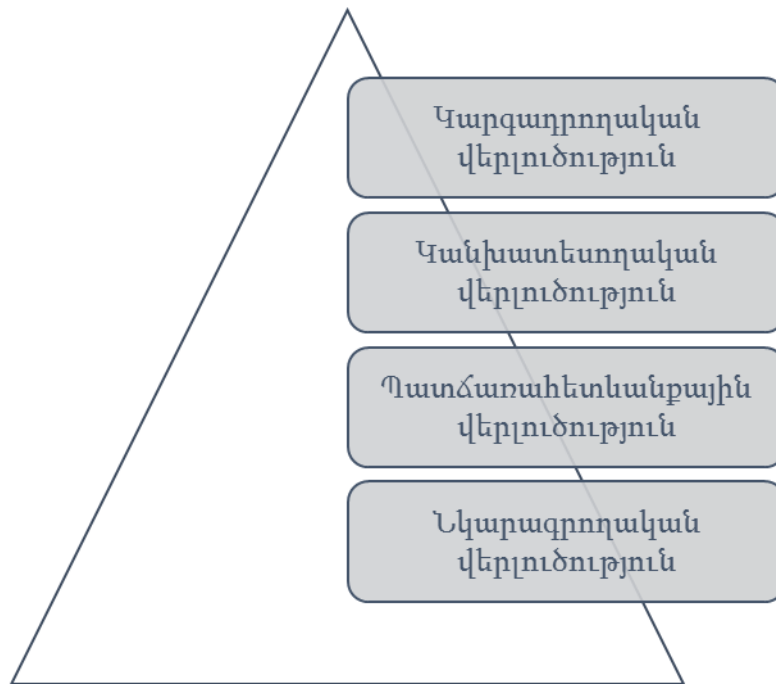
Գծապատկեր 23-ում ներկայացված է հաճախորդների հետկապված որոշումների ընդունման գործընթացում տվյալների կիրառության հիերարխիկ բուրգը: Բուրգի ստորին աստիճանում է գտնվում նկարագրողական վերլուծությունը, որը թույլ է տալիս խմբավորել և դասակարգել հաճախորդներին անհատական կամ խմբային բնույթի բնութագրիչների հիման վրա: Վերջինս անհրաժեշտ պայման է հանդիսանում մյուս փուլի՝ պատճառահետևանքային վերլուծության համար, քանզի պատճառականության բացահայտումն անհնար է առանց նկարագրողական հետազոտության առկայության:

Բուրգում ավելի բարձր աստիճան է զբաղեցնում կանխատեսողական վերլուծությունը, որը թույլ է տալիս նկարագրողական հետազոտության արդյունքների և բացահայտված պատճառահետևանքային կապերի հիման վրա կատարել կանխատեսումներ հաճախորդների ապագավարքի վերաբերյալ:

Գծապատկեր 23.

Յաճախորդների հետկապված որոշումների ընդունման գործընթացում տվյալների կիրառության հիերարխիկ բուրգը

**կառուցվել է մեր կողմից*



Գագաթում է գտնվում կարգադրողական վերլուծությունը: Վերջինս հանդիսանում է կանխատեսողական տարբեր սցենարների ամբողջականություն: Ի տարբերություն միայն կանխատեսողական վերլուծության, որտեղ ձեռներեցը օգտագործում է պատմական և ներկայիս իրավիճակը ապագայի մասին տեղեկատվություն քաղելու նպատակով՝ կարգադրողական հետազոտությունը բացահայտում է այն գործողությունները, որոնք ձեռներեցը պետք է իրականացնի հստակ սահմանված նպատակին հասնելու համար:

Անձնակազմ

Նմանօրինակ մոտեցում է կիրառվում նաև անձնակազմի կառավարման հետ կապված որոշումների գործընթացում տվյալների կիրառման դասակարգման համար: Այստեղ հիմք է հանդիսացել նաև Պենսիլվանիայի համալսարանում գործող Ուարթոն /Wharton/ բիզնես դպրոցի կողմից հետազոտական աշխատանքներում և ուսուցման ծրագրերում կիրառվող մոտեցումը⁷⁴: Համաձայն այդ մոտեցման հիերարխիկ բուրգի 4 մակարդակներից առաջինը (1 բալ) աշխատակիցների արտադրողականության գնահատումն է, որը ներկայացված է Գծապատկեր 24-ում: Միայն ներկա աշխատակիցների արտադրողականության գնահատման միջոցով է հնարավոր իրականացնել ճիշտ կողմնորոշում ունեցող կարգային

⁷⁴ Wharton People Analytics, Wharton, University of Pennsylvania, (15/01/2017), <https://wpa.wharton.upenn.edu/>

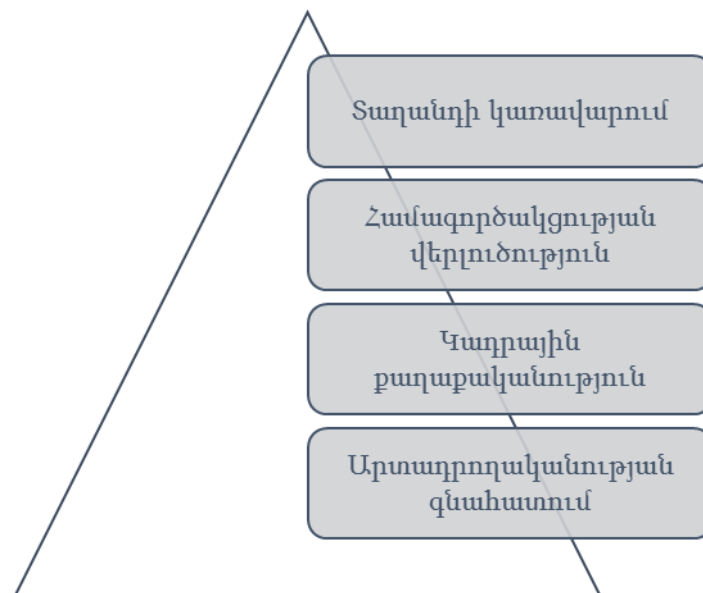
քաղաքականություն: Այդ պատճառով, հիերարխիկ բուրգում վերջինս զբաղեցնում է 2-րդ հորիզոնականը:

Գծապատկեր 24.

Անձնակազմի հետկապված որոշումների ընդունման գործընթացում տվյալների կիրառության հիերարխիկ բուրգը

**կառուցվել է մեր կողմից*

3-րդ հորիզոնականում է գտնվում համագործակցության վերլուծությունը, որը թույլ է տալիս կազմակերպության ներսում իրականացնել պահանջարկի պլանավորում: Ելնելով այն հանգամանքից, որ համագործակցության վերլուծության համար շատ ժամանակ անհրաժեշտ են տվյալներ, որոնք կարելի է հավաքագրել միայն կազմակերպության ներսում (օրինակ՝ կազմակերպության



ներսում աշխատակիցների տեղաշարժի մասին տվյալներ)՝ այն ավելի բարձր է գտնվում կադրային քաղաքականության հից:

Բուրգի գագաթն է զբաղեցնում Տաղանդի կառավարումը, որը համեմատաբար նոր ուղղություն է և աջակից է հանդիսանում բուրգի առաջին 3 մակարդակներին՝ հնարավորություն է տալով բացահայտել յուրաքանչյուր առանձին աշխատակցի տաղանդները, հմտությունները և վերջիններիս կիրառման արդյունավետ եղանակները:

Գործառնություններ

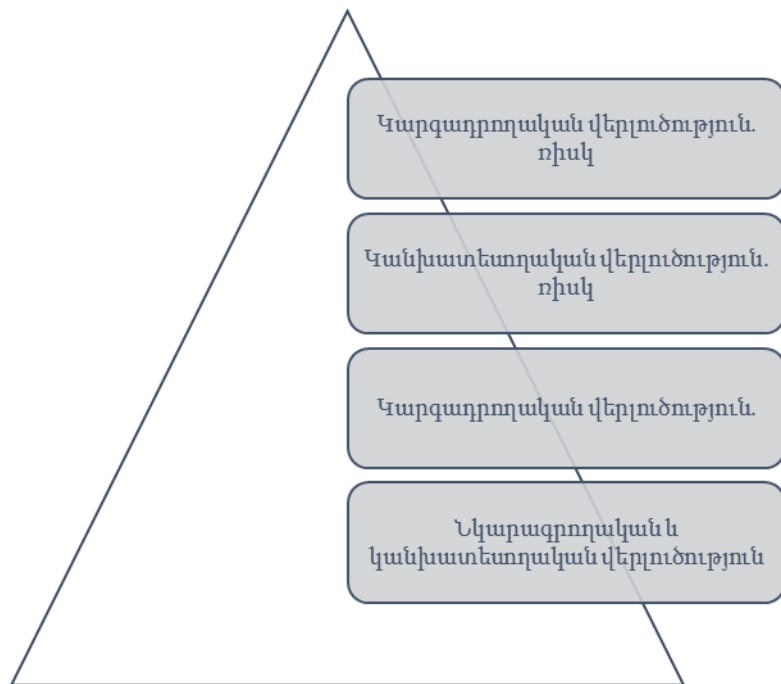
Գործառնությունների հետազոտման և վերջիններիս հետ կապված որոշումների ընդունման գործընթացում տվյալների տվյալների

վերլուծության համար կիրառվող մոդելներն ու մեթոդները բազմազան ու բազմատեսակ են: Այդ մոդելների և մեթոդների վերլուծության արդյունքում առանձնացվել են 4 հիմնական ձեռներեցական խնդիրների հիերարխիկ խումբը, որոնք ներկայացված են գծապատկեր 25-ում:

Գծապատկեր 25.

Գործառույթների հետկապված որոշումների ընդունման գործընթացում տվյալների կիրառության հիերարխիկ բուրգը

**կառուցվել է մեր կողմից*



Չափի առնելով այն հանգամանքը, որ աշխատանքի ավտոմատացման պահանջներից ելնելով գործառույթների մասին տվյալները, թերևս, ամենաշատն են հավաքագրվում, ի տարբերություն հաճախորդների հետ կապված որոշումների՝ այստեղ առանձնացվում են կանխատեսողական և կարգադրողական վերլուծության 2 տարբեր մակարդակ:

Կարգադրողական վերլուծությունը այդ հանգամանքի վերոնշյալ պատճառներով գտնվում է ավելի բարձր հորիզոնականի վրա քան կանխատեսողականը: Այնուամենայնիվ, ցածր ռիսկայնությամբ կարգադրողական հետազոտությանը գերազանցում է իր բարդությամբ բարձր ռիսկայնությամբ կանխատեսողական վերլուծությունը: Վերջինս ներառում է նաև ռիսկային վնասների

և շահերի տարատեսակ սցենարներում վերլուծությունն ու որոշումների կայացումը:

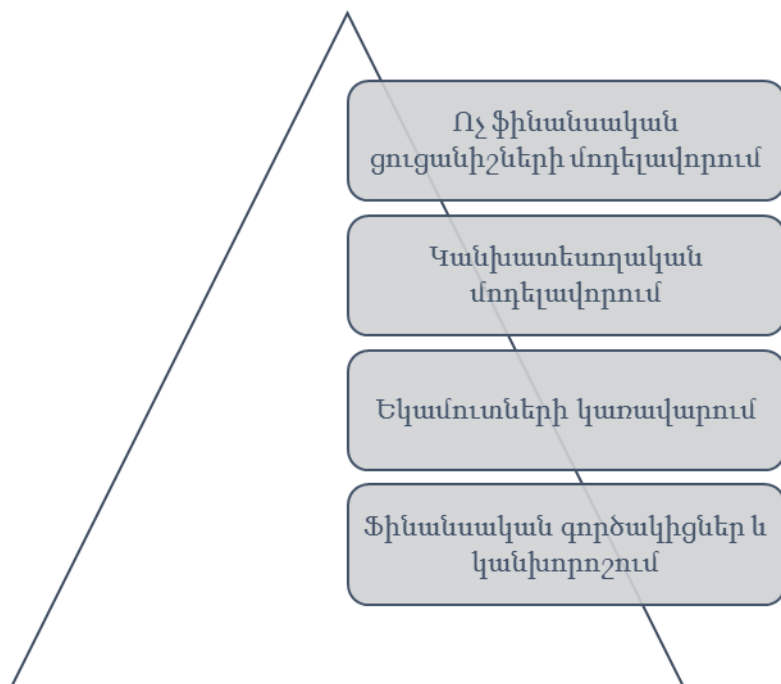
Ֆինանսներ

Ֆինանսների հետ կապված որոշումների ընդունման գործընթացում տվյալների կիրառության գնահատման համար օգտագործվող հիերարխիկ բուրգը ներկայացված է գծապատկեր 26-ում: Այնտեղ գազաթում է գտնվում ոչ ֆինանսական ցուցանիշների մոդելավորումը, որը ենթադրում է ոչ ֆինանսական տեսակի տվյալների կիրառությունը ֆինանսական որոշումների կայացման գործընթացում:

Գծապատկեր 26.

Ֆինանսների հետկապված որոշումների ընդունման գործընթացում տվյալների կիրառության հիերարխիկ բուրգը

**կազմվել է մեր կողմից*



Կանխատեսողական մոդելավորումը, ինչպես հաճախորդների հետ կապված որոշումների դեպքում, նույնպես զբաղեցնում է 3-րդ հորիզոնականը: Նրան նախորդում են ֆինանսների կառավարման 2 ավելի ցածր մակարդակները՝ ֆինանսական գործակիցներն ու եկամուտների կառավարումը:

Ուշադրության է արժանի այն փաստը, որ կազմակերպությունը կարող է բավականին ակտիվորեն կիրառել ՄՏ-ի վերլուծությունն 4 առանցքներից մեկում կամ մի քանիսում, և գրեթե չկիրառել այն

մյուսներում: Ելնելով այդ հանգամանքից՝ կարելի է եզրակացնել, որ այս 4 գործոնների միջև փոխկապակցվածությունը բավականաչափ ուժեղ չէ միայն ագրեգացված ցուցանիշին ուշադրություն դարձնելու համար: Վերոնշյալ ձևակերպումը պարզ դարձնելու նպատակով եկեք դիտարկենք հետևյալ օրինակը: Ենթադրենք «Ա» կազմակերպության կողմից հաճախորդների և ֆինանսների հետ կապված գործարար որոշումների կայացման գործընթացում ՄՏ-ի կիրառությունը գնահատվում է 4, իսկ աշխատակիցների և գործառույթների հետ կապված որոշումների կայացման գործընթացում՝ համապատասխանաբար 2: Արդյունքում, «Ա» կազմակերպության ՏԿԳ-ն կհաշվարկվի հետևյալ կերպ.

$$\begin{aligned} \text{ՏԿԳ} &= \frac{\text{Հաճախորդների գնահատական} + \text{Աշխատակիցների գնահատական} + \\ &\quad + \text{Գործառույթների գնահատական} + \text{Ֆինանսների գնահատական}}{4} \\ &= \frac{4 + 2 + 2 + 4}{4} = 3 \end{aligned}$$

Ենթադրենք մեկ այլ՝ «Բ», կազմակերպության համար այդ ցուցանիշը ունի հետևյալ տեսքը.

$$\text{ՏԿԳ} = \frac{3 + 3 + 3 + 3}{4} = 3$$

Արդյունքում ստացվում է, որ երկու կազմակերպություններն էլ ունեն ՏԿԳ 3 ցուցանիշ, սակայն 4 առանցքներում նրանց կողմից տվյալների կիրառությունը խտորեն տարբերվում է: Այդ իսկ պատճառով, որպես ՄՏ-ի կիրառության բարձր մակարդակ կարելի է սահմանել այն դեպքը, երբ $\text{ՏԿԳ} > 2$, սակայն հաճախ անհրաժեշտություն կարող է առաջանալ այն մասնատել և տեղեկատվություն ստանալ նրա մասը կազմող յուրաքանչյուր առանցքում ընդունվող որոշումներում կիրառման մակարդակի մասին:

ՏԿԳ մոտեցումը թույլ է տալիս քանակապես գնահատել տվյալների կիրառությունը որոշումների կայացման գործընթացում և հասկանալ, արդյոք կարելի է այն բարելավել թե ոչ: Այս տեղեկատվության կարևորությունը յուրահատուկ նշանակություն ունի գործարար կառավարման ոլորտում, քանզի կիրառության մակարդակի փոփոխությունը անհրաժեշտություն է ծնում նորանոր ներդրումների: Գործարար կառավարման 4

առանցքներից յուրաքանչյուրում ՄՏ-ի կիրառության մակարդակը բարձրացնելը, թերևս, դրսևորվում է նոր ձեռներեցական խնդիրներին ՄՏ-ի միջոցով լուծում տալու և/կամ նախկինում եղած խնդիրներին ավելի արդյունավետ լուծում տալու տեսքով: Սակայն նոր խնդրի լուծումը կամ հին խնդրի առավել արդյունավետ լուծումը, հավանաբար, պահանջելու է նոր աշխատակիցների ներգրավում, որը կազմակերպության կողմից հավելյալ ներդրում է պահանջում: ՏԿԳ-ի միջոցով ձեռներեցը կարող է պարզել, թե որ մակարդակում է գտնվում և հասկանալ արդյոք խելամիտ է այդ պարագայում նոր տվյալ ագետների աշխատանքի ընդունելը:

Մեծ տվյալների օգտագործման արդյունավետության մատրից (USOU մատրից)

Վերոնշյալ երկու գնահատականները հիմք են հանդիսանում ՄՏ-ի օգտագործման արդյունավետության մատրիցի կառուցման համար: Վերջինս իրենից ներկայացնում է երկչափ կոորդինատային համակարգ, որի հորիզոնական առանցքն է ՏԿԳ-ն, իսկ ուղղահայացը՝ ՏՄԳ-ն:

Գծապատկեր 27-ում ցուցադրվում է ՄՏ-ի օգտագործման արդյունավետության մատրիցը, որը կազմակերպություններին բաժանում է 4 խմբի՝ հիմնվելով նրանց ՏՄԳ և ՏԿԳ մակարդակների բարձր կամ ցածր լինելու վրա: 4-րդ խումբը համարվում է ամենաարդյունավետը, քանզի ոչ միայն ունի ՏՄԳ բարձր մակարդակ, այլև ակտիվորեն կիրառում է հասանելի Մեծ տվյալները՝ որոշումների կայացման գործընթացում: Ենթադրվում է, որ մյուս 3 խմբերից յուրաքանչյուրում գործող կազմակերպությունները պետք է ցանկանան հայտնվել 4-րդ դիրքում: Սակայն թռիչքային անցում ապահովելը հնարավոր և հավանական չի դիտվում:

ՄՏՕԱ մաարիցը

**Կառուցվել է մեր կողմից*



Այսպես, եթե կազմակերպությունը գտնվում է առաջին խմբում, այն է՝ ունի ավանդական տվյալներ և բավականին ցածր է վերջիններիս կիրառությունը որոշումների կայացման գործընթացում, ապա տրամաբանական չէ ենթադրել, որ այդ կազմակերպությունը կկարողանա ներդրումների շնորհիվ հայտնվել 4-րդ դիրքում: Այդ դիրքում հայտնվելու համար խորհուրդ տրվող ճանապարհն այսպիսին է.

1. Նախ և առաջ, ձեռներեցը պետք է կարողանա արդյունավետորեն կիրառել իրեն հասանելի տվյալները՝ 1-ին մակարդակից 3-րդ խումբ տեղափոխվելով:
2. Այնուհետև, ելնելով այն հանգամանքից, որ հասանելի ռեսուրսները (այս դեպքում՝ տվյալները) արդյունավետ են օգտագործվում՝ ձեռներեցը սանձ կառնի Մեծ տվյալների հավաքագրման գործընթացին:
3. 2-րդ փուլի գործողությունները ամենայն հավանականությամբ կհասցնեն ոչ թե 4-րդ այլ 2-րդ խումբ,

քանզի Մեծ տվյալների հավաքագրումն արդեն իսկ ներդրումներ է պահանջում: Իսկ 4-րդ փուլը հասնելու համար անհրաժեշտ են նաև ներդրումներ վերջինիս վերլուծության գործընթացում:

Չնայած այն հանգամանքին, որ դեպի 4-րդ մակարդակ ճանապարհը անցնում է սկզբից 3-րդ, և միայն հետո 2-րդ մակարդակներով՝ խմբերի այս աստիճանավորումը հիմնավորված է ձեռներեցի գործունեության արդյունավետության ամբ: Այլ կերպ ասած, 3-րդ խմբի 2-ից բարձր դասակարգվելու պատճառը կայանում է այն հանգամանքում, որ այդ մակարդակում գտնվող կազմակերպությունը արդյունավետ է օգտագործում իրեն հասանելի տվյալները, մինչդեռ 2-րդ խմբում, չնայած տվյալների մեծությանը, վերջիններս արդյունավետ օգտագործման չեն ենթարկվում:

Յաճախ ընկերությունները հավասար չեն գնահատում 4 առանցքներում ՄՏ-ի օգտագործման կարևորությունը: Այդ իսկ պատճառով տրամաբանական է ենթադրել, որ հնարավոր է դեպի 4-րդ խումբ անցման ճանապարհին իրականացնել ոչ համասեռ ներդրումներ՝ գործարար կառավարման 4 առանցքներից յուրաքանչյուրում ՏՄԳ և ՏԿԳ մեծացման համար: Նման պարագայում ավելի մանրակրկիտ ուսումնասիրության անհրաժեշտություն է առաջանում, որը հնարավորություն է տալիս հասկանալ, թե 4 առանցքներից որ մեկում է կազմակերպությունը ավելի «առաջնային» դիրք գրավում, և որում՝ ոչ:

Աղյուսակ 15.

ՄՏՕԱ մարտիցի վերլուծություն

**կազմվել է մեր կողմից*

	Ճավալ	Շարժ	Բազմազանություն	ՏՄԳ
Յաճախորդներ	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$(x_{1,1}+x_{1,2}+x_{1,3}) * 1/3$
Աշխատակիցներ	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$(x_{2,1}+x_{2,2}+x_{2,3}) * 1/3$
Գործառույթներ	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$(x_{3,1}+x_{3,2}+$

				$x_{3,3}$ *1/3
Ֆինանսներ	$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$(x_{4,1}+x_{4,2}+x_{4,3})$ *1/3
Բնութագրիչի գնահատական	$(x_{1,1}+x_{2,1}+x_{3,1}+x_{4,1})$ *1/4	$(x_{1,2}+x_{2,2}+x_{3,2}+x_{4,2})$ *1/4	$(x_{1,3}+x_{2,3}+x_{3,3}+x_{4,3})$ *1/4	

Վերևում պատկերված աղյուսակը հանդիսանում է USOU մատրիցի մանրակրկիտ վերլուծության գործիքը: Ինչպես տեսնում ենք աղյուսակ 15-ից՝ յուրաքանչյուր տող իրենից ներկայացնում է գործարար կառավարման 4 առանցքերից մեկը, իսկ սյուները՝ ՄՏ-ի 3 բնութագրիչներն են: Աղյուսակի 15-ի վանդակները ցույց են տալիս տվյալ տողին համապատասխանող առանցքում տվյալ սյանը համապատասխանող բնութագրիչի գնահատականը: Օրինակ՝ $x_{1,2}$ -ը ցույց է տալիս հաճախորդներին վերաբերող տվյալների շարժունության գնահատականը: Յետևաբար, յուրաքանչյուր տողի 3 վանդակների արժեքների միջին թվաբանականը ցույց կտա տվյալ առանցքի վերաբերյալ հավաքագրվող տվյալների մեծությունը, այն է՝ ՏՄԳ-ն առանցքային մակարդակով: Յարկավոր է նշել, որ աղյուսակ 3-ը կարող է նաև ստուգիչ գործիք հանդիսանալ ՏՄԳ-ի համար. գումարելով միևնույն սյան գնահատականները բոլոր 4 առանցքներում և հաշվելով նրանց միջին թվաբանականը՝ պետք է ստանալ տվյալ բնութագրիչի գնահատականը (օրինակ՝ ծավալ): Եթե յուրաքանչյուր սյան ստորին վանդակում ստացված արժեքը չի համապատասխանում տվյալ բնութագրիչի գնահատականին, ապա վերջինս խոսում է հաշվարկներում կատարված սխալի մասին:

Յետագոտության հաջորդ հատվածում ներկայացվում է ՀՀ շուկայում գործող 19 կազմակերպությունների հետ անցկացված որակական (խորքային) հարցման արդյունքները, որոնք երկրորդային հետազոտության հետ միասին ոչ միայն հիմք են հանդիսացել USOU մատրիցի և նրա հիմքում ընկած ՏՄԳ-ի և ՏԿԳ-ի ստացման համար, այլև օգտագործվում են ՀՀ Մեծ տվյալների հավաքագրմամբ և գործարար որոշումների գործընթացում

վերջինիս վերլուծության կիրառմամբ գբաղվող ընկերությունների վերլուծության համար:

ՀՀ կազմակերպությունների կողմից Մեծ տվյալների օգտագործման համեմատական արդյունավետության գնահատում

ՀՀ աշխատաշուկայում Մեծ տվյալների վերլուծության կիրառությունը, արդյունավետությունը և տվյալագետների նկատմամբ եղած պահանջակն ուսումնասիրելու նպատակով անց է կացվել որակական հարցում⁷⁵: Ներկայումս շուկայում գործող կազմակերպություններից բացահայտվել են ՄՏ-ի վերլուծությամբ գբաղվելու հավանականություն ունեցող 19 կազմակերպություն: Ընկերությունների փոքր թվի և ՄՏ-ի վերլուծություն կիրառող ընկերությունների մասին հստակ և վստահելի տեղեկատվության բացակայության պատճառով քանակական հարցում անցկացնելը ներկայումս հնարավոր չէ:

Որակական հարցումը նպատակ էր հետապնդում պարզել ՀՀ-ում գործող կազմակերպությունների կողմից տվյալագիտության կիրառությունը գործարարական տարբեր որոշումների կայացման գործընթացում, ապագա պլանները այդ հարցի վերաբերյալ, առօրյայում վերլուծվող տվյալների ծավալը, ինչպես նաև տեղեկատվություն ստանալ ընկերության ֆինանսական ցուցանիշների վրա ՄՏ-ի վերլուծության գրանցված կամ սպասված ազդեցության մասին: Սակայն հարցվող ընկերությունների ներկայացուցիչների ֆինանսական տվյալների մասին տեղեկատվություն տալու մասին հիմնականում բացասական որոշման հետևանքով այդ հարցերը հանվեցին հարցաշարից: Արդյունքում վերջնական հարցաթերթը ներառում է 6 հարցի մասին պատասխաններ 19 ընկերության ՄՏ-ին հարող հաստիք ունեցող մասնագետներից: Ընկերությունների ցանկը կազմվել է քանակական հարցման արդյունքում ստացված տվյալների հիման վրա:

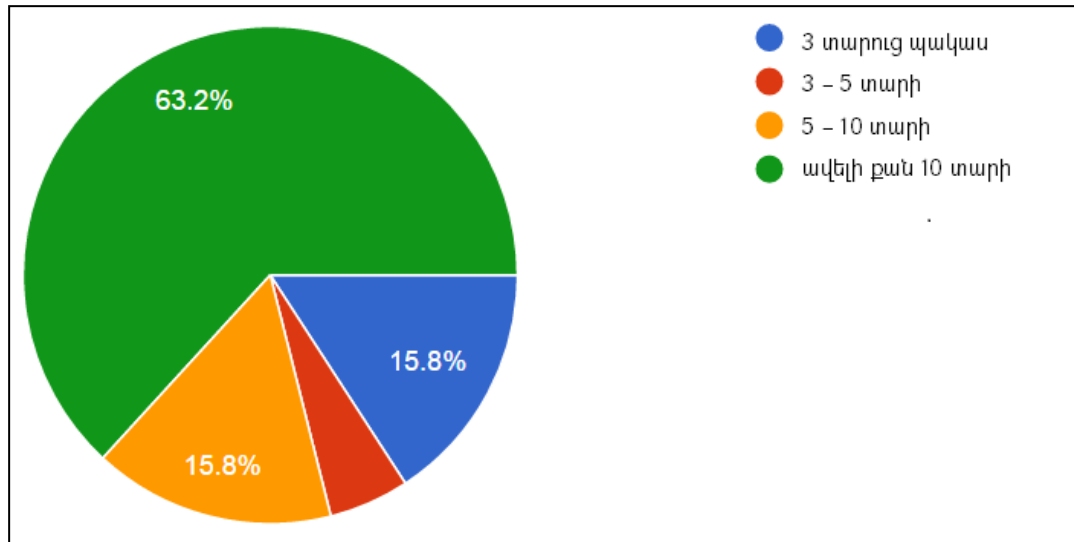
Հարցված կազմակերպություններից ավելի քան ¾-ը՝ 80%, գործում են ավելի քան 5 տարի: Մյուս կողմից, ՄՏ-ի

⁷⁵ Դավթյան, Հ., ՀՀ կազմակերպությունների կողմից մեծ տվյալների օգտագործման համեմատական արդյունավետության գնահատում: «Ֆինանսներ և Էկոնոմիկա» հանդես, 3-4 (199-200), 2017 թ.

վերլուծությամբ զբաղվող հարցված ընկերությունների 15,8%-ը գործում են ավելի քիչ քան 3 տարի: Սավկայում է այն փաստի մասին, որ ՀՀ-ում կան ՄՏ-ի վերլուծության կարևորությունը գիտակցող երիտասարդ կազմակերպություններ: Ընկերությունների գործունեության ժամկետը նկարագրող ընդհանրական պատկերը ցուցադրված է գծապատկեր 28-ում:

**Մեծ տվյալների վերլուծության կիրառության ամբողջական գնահատումը
ռեսպոնդենտներին վերաբերող ոլորտների գործունեության ամփոփումները**

**կառուցվել է մեր կողմից*



Բոլոր 19 կազմակերպություններին տրվել է այն հարցը, թե գործունեության 4 առանցքներից յուրաքանչյուրում (հաճախորդներ, աշխատակիցներ, գործառույթներ և ֆինանսներ) որոշումներ կայացնելիս որքանով են օգտագործվում ՄՏ-ն: Որպեսզի հարցման մասնակիցների պատասխանների միջև ապահովվի համասեռությունը՝ վերջիններին ներկայացվել է, թե որոնք են ՏԿԳ 5 մակարդակները, ներառյալ՝ 0-ական մակարդակը): Հարցման ընթացքում 0-ական մակարդակը կողմորվել է որպես 1 միավոր, իսկ հիերարխիկ բուրգի 4 մակարդակները՝ 2 – 5 միավորով: Այս մոտեցումը կիրառվել է գործարար կառավարման բոլոր 4 առանցքներում որոշում կայացնելիս Մեծ տվյալների կիրառման մասին դիտարկումների համար:

Այդ ուսական 16-ը ցույց է տալիս, որ ներկայումս հարցվող ընկերություններում գործարար որոշումների կայացման գործընթացում ՄՏ-ի կիրառությունը միջինում բավականին ցածր մակարդակի վրա է՝ անկախ գործունեության առանցքից: Ընդհանուր առմամբ միջինից շեղումները շատ չեն, ինչը հիմնավորվում է 18,2%-ից 25,2% միջակայքում ընկած ստանդարտ շեղման գնահատականով:

**Հարցվող կազմակերպությունների կողմից գործունեության 4
առանցքներում ընդունվող որոշումներում Մեծ տվյալների
կիրառությունը**

**կազմվել է մեր կողմից*

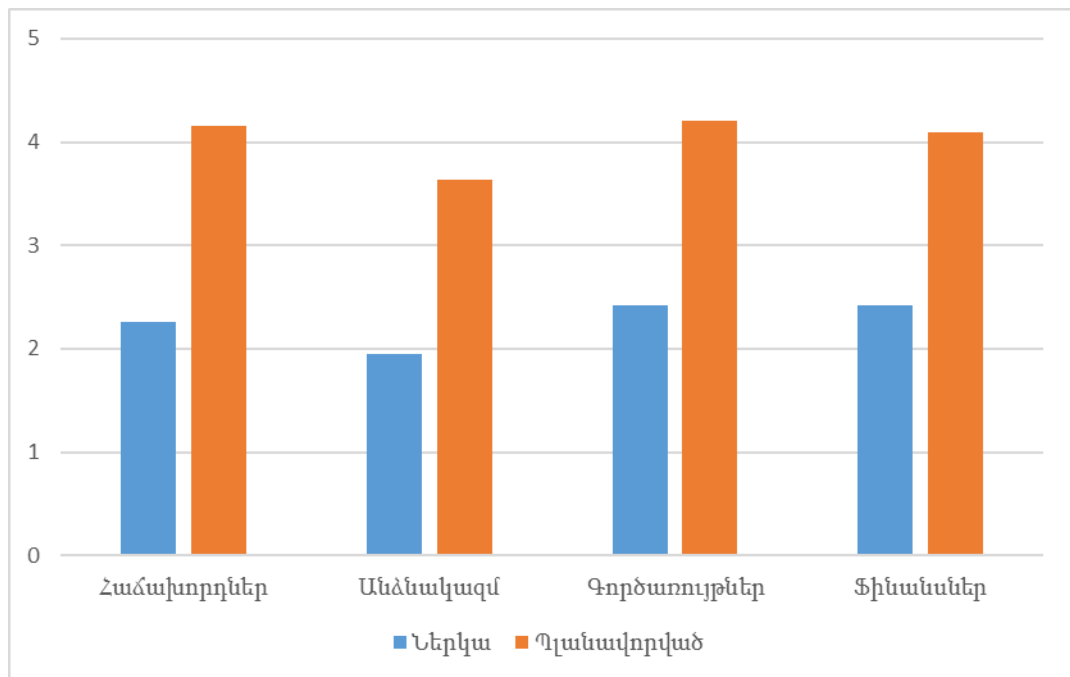
	Հաճախորդներ	Անձնակազմ	Գործառույթներ	Ֆինանսներ
Միջին արժեք	2,26	1,95	2,42	2,42
Ստանդարտ շեղում	21,8%	18,2%	25,2%	25,2%

Չնայած ներկայիս նման ցածր կիրառելիությանը՝ հարցված կազմակերպությունները բավականին կարևորում են ՄՏ-ի վերլուծության դերը բոլոր 4 առանցքներում: Հաճախորդների, գործառույթների և ֆինանսների հետ կապված որոշումների կայացման գործընթացում ՄՏ-ի վերլուծության կիրառության դերը հարցված կազմակերպությունները միջինում բավականին բարձր են գնահատում և պլանավորում են կիրառել մոտակա ապագայում: ՅՀՀ 4 առանցքներում ապագա կիրառության գնահատականների ստանդարտ շեղումը այս անգամ նույնպես մեծ է:

Գծապատկեր 29.

**Մեծ տվյալների վերլուծության ներկան պլանավորված
կիրառության համեմատությունը գործարար գործունեության 4
առանցքներում**

**կառուցվել է մեր կողմից*



Վերոնշյալ առանցքներում ՄՏ-ի ներկայիս կիրառության և ապագա պլանների մասին տվյալները համեմատությունը ներկայացված է գծապատկեր 29-ում: Ինչպես երևում է գծապատկերից՝ մի փոքր ցածր է գնահատվում աշխատակիցների մասին որոշումների ընդունման գործընթացում վերջինիս կիրառությունը՝ ընդունելով միջինում 3,63 արժեք: Ամենայն հավանականությամբ այս ցուցանիշի ցածր լինելն էլ հենց պայմանավորում է ներկայիս ցածր կիրառության աստիճանը՝ գործարար գործունեության մնացյալ 3 առանցքների հետ համեմատ:

Պատասխանելով այն հարցին, թե ինչու է ՄՏ-ի վերլուծության մակարդակը ներկայումս ցածր պլանավորվածից՝ ընկերությունների մեծ մասը ընդգծեց համապատասխան գիտելիքներ և փորձ ունեցող մասնագետների՝ տվյալ ազետների սակավությունը: Հետաքրքրաշարժ է, որ վերապատրաստման ծրագրերի առկայությունը ռեսպոնդենտ ընկերությունների կողմից որպես խնդիր չի դիտվել: Այս փաստը վկայում է ՀՀ աշխատանքային տվյալ ազետների պահանջարկի մասին թերի տեղեկացվածության, այդ ոլորտի հանրությանը դեռևս ոչ այդքան հայտնի լինելու մասին: Հատկանշական է այն հանգամանքը, որ «Ֆինանսական բրենդ» հանրահայտ միջազգային ֆորումի հարցման

համաձայն այս խոչընդոտը զբաղեցնում է 2-րդ տեղը՝ զիջելով միայն US-ի վերլուծության համար պահանջվող ժամանակին⁷⁶:

Որպես US-ի գործարար որոշումների ընդունման գործընթացում US-ի վերլուծության կիրառման 2-րդ խոչորագույն խոչընդոտ հավասարապես նշվում են տվյալների հասանելիության ու որակը (հատկապես՝ ազգային վիճակագրության և/կամ հանրապետության մակարդակով պետական մարմինների կողմից հավաքագրվող այլ աղբյուրների տվյալների դեպքում): Անմիջապես ընկերության կողմից անհարժեշտ տվյալների հավաքագրումը միջինում գնահատվել է 3,16՝ կրկին 5 բալանի սանդղակում:

Չնայած այն հանգամանքին, որ ներկայումս US-ի վերլուծությունը գործունեության 4 առանցքերից յուրաքանչյուրում գտնվում է բավականին ցածր մակարդակում՝ որոշ առանձին խնդիրների լուծման համար վերջնիս կիրառությունը միջինում շատ ավելի բարձր ցուցանիշներ է գրանցում: Այդ առումով առանձնանում են նոր արտադրանքի կամ ծառայության վերաբերյալ որոշումները, որոնց US-ի վրահիմնված լինելը միջինում գնահատվում է 2,89:

Վերջին հարցը, որին պատասխանել են հարցվող կազմակերպությունները, կապված է նրանց կողմից հավաքագրվող և/կամ վերլուծվող տվյալների մեծության հետ: Ըստ հարցման մասնակիցների՝ իրենց կազմակերպությունում օգտագործվող տվյալների շտեմարանները ունեն գծապատկեր 30-ում ցուցադրված ծավալները: Ինչպես երևում է գծապատկերից՝ հարցմանը մասնակցած կազմակերպությունների գրեթե կեսը՝ 47,4%-ը, վերլուծում են բարձր մեգաբայթ կամ ցածր գիգաբայթ ծավալ զբաղեցնող տվյալներ: Սա, իհարկե, անհամեմատելի ցածր ցուցանիշ է զարգացած երկրների համեմատ:

Գծապատկեր 30.

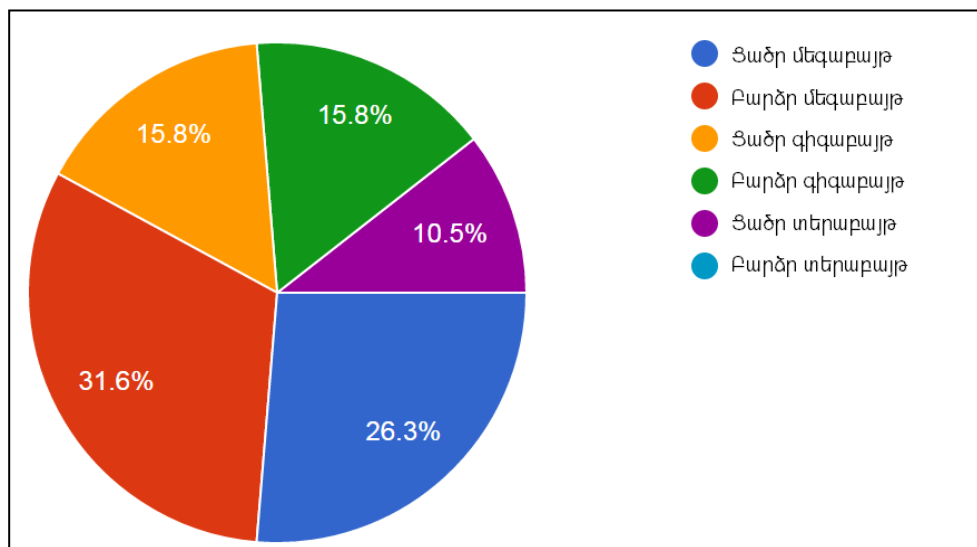
Հարցվող ընկերությունների կողմից օգտագործվող տվյալների շտեմարանների ծավալը

**կառուցվել է մեր կողմից*

⁷⁶ [The Financial Brand, "Big Data: Profitability, Potential and Problems in Banking", 2015, https://thefinancialbrand.com/38801/big-data-profitability-strategy-analytics-banking/](https://thefinancialbrand.com/38801/big-data-profitability-strategy-analytics-banking/)

Հարցման (խորքային հարցազրույցների) արդյունքում հավաքագրված տվյալների հիման վրա հաշվարկվել է ռեսպոնդենտ 19 կազմակերպությունների տվյալների մեծության և տվյալների կիրառության գնահատականները: Վերջիններս օգտագործվել են US-ի օգտագործման արդյունավետության մատրիցի ստեղծման համար:

Ստորև ներկայացված գծապատկերում կապույտ շրջանները ցույց են տալիս ռեսպոնդենտ ընկերությունների ներկայիս դիրքը մատրիցում⁷⁷: Ինչպես երևում է ոչ մի ընկերություն և, համաձայն իրենց իսկ կողմից տրված տվյալների, չի համապատասխանում 4-րդ մակարդակին՝ չնայած այն հանգամանքին, որ կան 4-րդ խմբի սահմանին գտնվող կազմակերպություններ: Քանակով առաջին տեղն են զբաղեցնում 1-ին խմբերում ներգրավված կազմակերպությունները: Միայն 1-ին մակարդակում ընդգրկված ընկերությունների թիվը (ներառյալ՝ սահմանայինները) հավասար է 14-ի: Ինչպես երևում է գծապատկերից՝ միայն 3 ընկերություն են



լիովին համապատասխանում 3-րդ խմբին, և ընդամենը 2-ը՝ USOU մատրիցի 2-րդ մակարդակին:

Գծապատկեր 31.

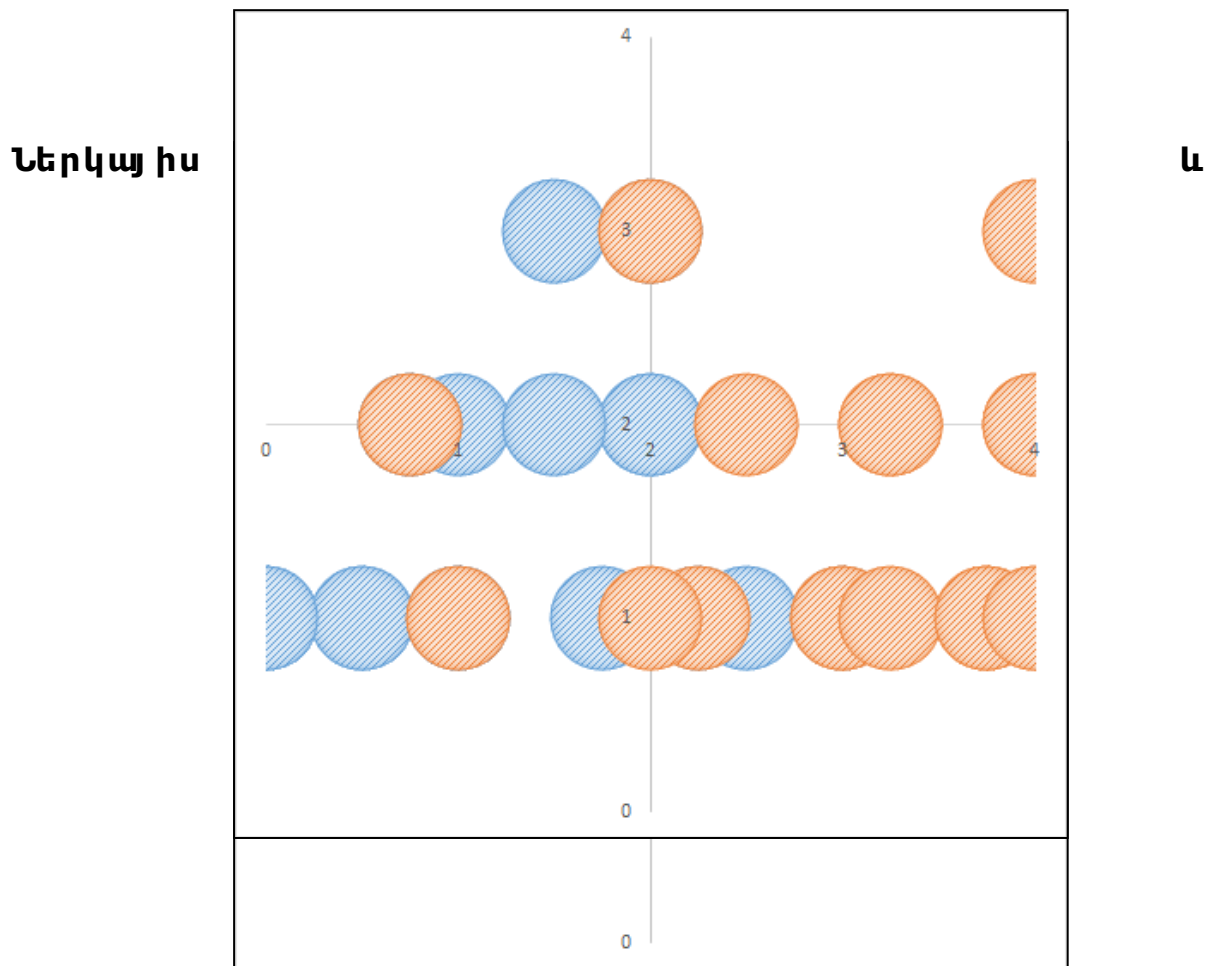
Ռեսպոնդենտ կազմակերպությունների դիրքը USOU մատրիցում

**կառուցվել է մեր կողմից*

⁷⁷ Գծապատկերում առկա կապույտ շրջանների քանակը գիշում է հարցված ընկերությունների քանակին: Պատճառն այն է, որ որոշ կազմակերպությունների արդյունքները համընկել են և արդյունքում մեկ ձեռնարկության շրջանը գծապատկերում ծածկել է մյուսին:

Օգտվելով այն հանգամանքից, որ հավաքագրվել են տվյալ կազմակերպությունների կողմից որոշումների կայացման գործընթացում ՄՏ-ի կիրառման ոչ միայն ներկա իրավիճակի մասին, այլև մոտակա ապագայում պլանավորված կիրառության մասին տվյալներ՝ հնարավորություն է ստեղծվել հաշվարկել նաև պլանավորված ՏԿԳ-ն: Վերջինս նույնպես օգտագործվել է ՄՏՈՒ մատրիցի ստեղծման գործընթացում: Արդյունքում, գծապատկեր 32-ում պատկերված մատրիցը տեղեկատվություն է հաղորդում ոչ միայն ռեսպոնդենտ ընկերությունների ներկայիս դիրքի, այլև մոտակա ապագայում պլանավորված զարգացման մասին:

Գծապատկեր 32.



պլանավորված ապագայի ՄՏՈՒ մատրիցը

**կառուցվել է մեր կողմից*

Այս գծապատկերում (32) ներկայացված կարմիր շրջանները ցույց են տալիս ՄՏՈՒ մատրիցում ռեսպոնդենտ 19

կազմակերպությունների դիրքը՝ համաձայն այն անավորված ՏԿԳ-ի: Ինչպես տեսնում ենք ընկերությունների ճնշող մեծամասնությունը դեպի աջ շարժ է գրանցել: Այս հանգամանքը ցույց է տալիս գործարար կառավարման բոլոր 4 առանցքներում որոշումների ընդունման գործընթացում ՄՏ-ի վերլուծության դերի գիտակցումը և այդ ուղղությամբ այն անների գոյությունը: Վերջինիս վկայությունն է նաև գծապատկեր 29-ում ներկայացված համեմատությունը:

Այնուամենայնիվ հարկավոր է նշել, որ ոչ բոլոր կազմակերպություններն են նախատեսում զարգացում գրանցել մատրիցի հորիզոնական առանցքով: Հարցված ընկերություններից 3-ը բավարարվում են իրենց ներկայիս վիճակով և չեն այն անավորում ներդրումներ կատարել ՄՏ-ի վերլուծության կիրառության ուղղությամբ, իսկ 1-ը՝ նախատեսում է ՏԿԳ-ի 0.25 բալով իջեցում մոտակա ապագայում: Մանրակրկիտ վերլուծությունը ցույց է տալիս, որ այս կազմակերպությունը այն անավորում է ապագայում բավարարվել հաճախորդների նկարագրողական և պատճառահետևանքային վերլուծությամբ՝ չնայած այն հանգամանքին, որ ներկայումս կիրառվում է նաև կանխատեսողական վերլուծություն:

Ցավոք սրտի հնարավոր չէ պարզել ռեսպոնդենտ կազմակերպությունների դիրքը USOU մատրիցում կախված ՏԿԳ-ի ապագա արժեքից: Պատճառն այն է, որ հարցմանը մասնակցող կազմակերպությունների ներկայացուցիչները դժվարացան մեկնաբանել թե՛ ՄՏ-ի 3 բնութագրիչների հետ կապված ինչ փոփոխություններ են սպասվում մոտագա ապագայում: Նույն պատճառով հնարավոր չէ իրականացնել USOU մատրիցի վերլուծություն՝ առյուսակ 15-ի կիրառմամբ, քանզի վերջինս ցույց է տալիս ՏԿԳ մեծացման համար անհրաժեշտ ուղղությունները, ինչի համեմատության համար տվյալները հասանելի չեն: Վերջինս թերևս մեկ այլ հետազոտության խնդրո առարկա է:

Այսպիսով, հետազոտության սույն հատվածում ներկայացված գործիքակազմը կարող է կազմակերպությունների համար «Մեծ

տվյալների դարաշրջանում» հանդիսանալ օգտագործման արդյունավետության գնահատման հիմնական միջոց, որը թույլ կտա կազմակերպության ռազմավարական նպատակներին հասնելու համար իրականացնել տվյալ ոլորտի ներդրումային ավանավորում: Վերջինս կարող է նաև միջոց հանդիսանալ մակրոտնտեսական կիրառության տեսանկյունից՝ մատրիցում ընկերությունների փոխարեն պատկերելով երկրների կամ մարզերի համեմատական արդյունավետությունը տվյալ ոլորտում:

Յետագոտության հաջորդ և վերջին ենթագլխում ներկայացվում է USOU մատրիցի ներկայիս դիրքից ավանավորված դիրքին անցում կատարելու համար անհրաժեշտ ներդրումների կառուցվածքը և դիտարկվում է ՀՀ-ում Մեծ տվյալների ոլորտի զարգացման հեռանկարները:

3.3 Մեծ տվյալների օգտագործման արդյունավետության բարձրացման հիմնախնդիրները և նախադրյալները ՀՀ-ում

Մեծ տվյալների օգտագործման արդյունավետության բարձրացման հիմնախնդիրները

Մեծ տվյալների վերլուծության արդյունավետությունը խթանում է կազմակերպությունների կողմից այդ ոլորտում իրականացվող ներդրումները⁷⁸: Վերոնշյալ USOU մատրիցը հնարավորություն է տալիս գնահատել ՄՏ-ի օգտագործման արդյունավետությունը և հասկանալ այն ուղղությունները, որոնցում անհրաժեշտ է իրականացնել ներդրում՝ գործարար կառավարման ասպարեզում կայացվող որոշումների արդյունավետության բարձրացման նպատակով: Այնուամենայնիվ, մատրիցը դեռևս չի պատասխանում այն հարցին, թե որքան ներդրումներ են անհրաժեշտ արդյունավետության որևէ թիրախային մակարդակի հասնելու համար⁷⁹: Յետագոտության այս հատվածում ամբողջականացվում է ՄՏ-ի օգտագործման

⁷⁸ Տե՛ս [Gartner, "Gartner Symposium, survey summary" 2016,
http://www.gartner.com/newsroom/id/3130817](http://www.gartner.com/newsroom/id/3130817)

⁷⁹ Դավթյան, Յ., Մեծ տվյալների ոլորտում ներդրումների իրականացման արդյունավետությունը: «Այլ ընտրանք» գիտական հանդես, №4, 2016 թ., էջեր 228-232

գործիքակազմը՝ ներկայացնելով US-ի ոլորտում իրականացնող ներդրումների կառուցցածքային բաղադրիչները և վերջիններիս ծավալի գնահատման համար անհրաժեշտ միջին վիճակագրական մեծությունները:

Համաձայն USOU մատրիցի՝ US-ի օգտագործման արդյունավետությանը բարձրանում է SU9 և S49 արժեքների մեծացման պարագայում: Այդ տեսանկյունից այս ոլորտում իրականացվող ներդրումները կարելի է բաժանել երեք խմբի՝ տվյալների մեծությանն ու տվյալների կիրառությանն ուղղված ներդրումներ և US-ի կառավարման ուղղված ներդրումներ, որոնք կամուրջ են ստեղծում նախորդ երկուսի միջև:

SU9 բարձրացմանն ուղղված ներդրումներ

Մեծ տվյալներ ձեռք բերելու համար անհրաժեշտ է հետևյալ երկու պայմանների բավարարումը.

- **Մեծ տվյալների հավաքագրում:** Ի տարբերություն «ավանդական» տվյալների՝ US-ն հնարավոր է հավաքագրել հարցման միջոցով: Պատճառն այն է, որ սահմանման համաձայն US-ն պետք է ունենան բարձր շարժունություն, որը հարցումն ապահովել չի կարող: Այդ իսկ պատճառով վերջինիս հավաքագրումը ենթադրում է ժամանակակից S3S օգտագործում: Ներկայումս տարածված լուծում է Կիրառական ծրագրավորման ինտերֆեյսների (ԿԾԻ, անգլերեն՝ API) օգտագործումը, որը թույլ է տալիս հավաքագրել տարատեսակ աղբյուրներից ստացվող տվյալները: Վերջինիս վառ օրինակ է Amazon կամ eBay առցանց կայքերի ԿԾԻ-ն^{80,81}: Այն թույլ է տալիս իրական ժամանակահատվածում ստանալ տվյալներ այդ կայքերում տարատեսակ ապրանքների դիտումների, գների, քանակների և գնումների վերաբերյալ: Վերջին տարիներին US-ի ոլորտի բուռն աճը այնքան է զարգացրել ԿԾԻ օգտագործումը, որ վերջինս դիտվում է առանձին ոլորտ՝ հանդիսանալով և՛ գիտական, և՛ գործնական հետազոտությունների առարկա⁸²: Հետազոտության ընթացքում դիտարկված ԿԾԻ-ները տարատեսակ և բազմազան են և ընկերություններին առաջարկում են տարբեր բիզնես փաթեթներ:

⁸⁰ Տե՛ս <https://aws.amazon.com/api-gateway/>, (24/02/2017)

⁸¹ Տե՛ս <https://go.developer.ebay.com/>, (24/02/2017)

⁸² Տե՛ս <http://nordicapis.com/tracking-the-growth-of-the-api-economy/>

Մեր կողմից հաշվարկված միջին ԿՃԻ ծախսը տարբեր տվյալների աղբյուրների դեպքում ընկած է ամսական 400-ից 700 ԱՄՆ դոլար միջակայքում: Այնուամենայնիվ ԿՃԻ-ն հանդիսանում է տվյալների հավաքագրման անուղղակի աղբյուր՝ ընկերությանը կախված դարձնելով տվյալների անմիջական հավաքագրմամբ զբաղվող կազմակերպության հից: Ընկերության նախապատվություններին համապատասխան տվյալներ հավաքագրելու նպատակով շատ ժամանակ անհրաժեշտություն է առաջանում ուղղակի ներդրումների իրականացման օրինակ՝ Ռադիոհաճախականային նույնականացման չիպեր (անգլերեն՝ RFID) ձեռք բերելու նպատակով: Այս չիպերը կարող են հավաքագրել տվյալներ բազմատեսակ գործողությունների վերաբերյալ՝ ներառյալ հետազոտության առարկայի տեղակայումը, տեղաշարժը և այլն: **Նման 1000 հատ չիպի միջին արժեքը գնահատվում է 300 ԱՄՆ դոլար,** որը հանդիսանում է միանգամյա ներդրում, քանզի չիպերը բազմամյա օգտագործման են ենթակա:

▪ **Մեծ տվյալների պահպանում:** Վերոնշյալ ուղիներով հավաքագրված տվյալներն անհրաժեշտ է պահպանել՝ վերլուծության նպատակով: ՄՏ-ի ծավալն ու բազմազանությունը անհնարին է դարձնում համակարգչի/կոշտ սկավառակի վրա վերջինիս պահպանումը: Միաժամանակ կարևորվում է այդ տվյալների պահպանումը այնպիսի վայրում, որը կապահովի ՄՏ-ի իրական ժամանակահատվածում հասանելիություն: Այդ խնդրի լուծումն են տալիս վերջին տարիներին զարգացում ապրող ամպային տեխնոլոգիաները (անգլերեն՝ cloud technologies), որոնք հնարավոր են դարձնում ՄՏ-ի պահպանումը համացանցում: Չնայած նրան, որ տվյալները պահպանվում են համացանցում՝ նրանց գաղտնիությունն ապահովվում է տարատեսակ նույնականացման մեխանիզմների միջոցով: Ամպային տեխնոլոգիաներ են առաջարկում ինտերնետային գիգանտներ Amazon-ը, Google-ը և մի շարք այլ ընկերություններ: **Ամպային տեխնոլոգիաներից օգտվելու միջին տարեկան ծախսը 1 տեռաբայթ ծավալով տվյալների պահպանման համար մեր կողմից գնահատվել է 282 ԱՄՆ դոլար՝** հիմքում ընդունելով վերոնշյալ կազմակերպության կողմից առաջարկվող

տարատեսակ փաթեթների գները: Այդ լուծումների վրա 1000-ից ավել աշխատակից ունեցող ընկերությունները ԱՄՆ-ում տարեկան ծախսում են միջինում ավելի քան 2 մլն ԱՄՆ դոլար⁸³: Այս ցուցանիշը ևս մեկ անգամ ընդգծում է ՄՏ-ի առեւտլի ծավալը. հիմնվելով այս ցուցանիշի վրա կարելի է եզրակացնել, որ միջինում 1000-ից ավել աշխատակից ունեցող ընկերության տարեկան տվյալները զբաղեցնում են 7000 տեռաբայթ տարածք:

ՏԿԳ բարձրացման ուղղված ներդրումներ

Վերոնշյալ ներդրումները կապահովեն կազմակերպության դրական տեղաշարժը USOU մատրիցի օրդինատների առանցքով: Սակայն ՄՏ-ի միայն ձեռքբերումն ու պահպանումը արդյունավետության համար բավարար պայման չէ: Ներդրումային արդյունավետության համար անհրաժեշտ է նաև ապահովել այդ տվյալների կիրառությունը: ՄՏ-ի կիրառության համար անհրաժեշտ ներդրումները բաժանվում են երկու խմբի.

- **Ծրագրային ապահովվածություն:** Մեծ տվյալների և ավանդական տվյալների հիմնական տարբերակիչ առանձնահատկություններից մեկը վերլուծության համար օգտագործվող ծրագրային փաթեթներն են: Ելնելով ՄՏ-ի բազմազանությունից, առեւտլի ծավալից և շարժունությունից՝ դասական տվյալների վերլուծության ծրագրային փաթեթերը անհամատեղելի են դառնում: Այդ պարագայում համապատասխան ծրագրային ապահովվածություն ունենալու համար ներդրումների իրականացման անհրաժեշտություն է առաջանում: Բարեբախտաբար ՄՏ-ի վերլուծության համար օգտագործվող հիմնական գործիքները՝ Python և R ծրագրավորման լեզուները, ամբողջությամբ անվճար են և հասանելի ցանկացած օգտագատիրոջ: Համեմատության համար պետք է նշել, որ ավանդական տվյալների վերլուծության համար օգտագործվող SPSS և Stata ծրագրային փաթեթների արժեքներն են համապատասխանաբար 1200 ԱՄՆ դոլար (տարեկան)⁸⁴ և 2495 ԱՄՆ

⁸³ Տե՛ս IDG, “2016 Cloud computing summary”, https://www.idgenterprise.com/resource/research/2016-idg-enterprise-cloud-computing-survey/?utm_campaign=Cloud%20Computing%20Survey%202016&utm_medium=Press%20Release&utm_source=Press%20Release

⁸⁴ <https://www.ibm.com/us-en/marketplace/spss-statistics/purchase#product-header-top>, (02/12/16)

(մշտական) դոկար⁸⁵: Այս տեսանկյունից US-ի վերլուծությունը բերում է նաև ծրագրային ապահովվածության վրա կատարվող ներդրումների խնայողություն: Այնուամենայնիվ պետք է նշել, որ բացի վերոնշյալ երկու ծրագրավորման լեզուներից կան նաև միջարքայլ փաթեթներ, որոնք կիրառվում են US-ի վերլուծության նպատակով և որոնց տարեկան վճարը կազմում է 500-ից 9000 ԱՄՆ դոկար⁸⁶:

▪ **Կարգային ապահովվածություն:** Մեծ տվյալների և վերջինիս վերլուծության համար անհրաժեշտ ծրագրային փաթեթների հասանելիության պարագայում ընկերության հաջորդ քայլն է համապատասխան մասնագետների՝ տվյալագետների աշխատանքի ընդունելը: Ինչպես նշվեց 2-րդ գլխում՝ տվյալագետների աշխատավարձը թե՛ ԶԶ-ում, թե՛ ԱՄՆ կամ այլ երկրներում բավականին բարձր է և մի քանի անգամ գերազանցում է միջին աշխատավարձը⁸⁷: Եթե նախորդող ներդրումները հիմնականում համացանցային լուծումներ էին և աշխարհագրական տեսանկյունից տարբերություններ չէին ծնվում, ապա տվյալագետների պարագայում կա բավականին մեծ տարբերություն՝ կախված կազմակերպության գործունեության երկրից: Ելնելով այն հանգամանքից, որ ԱՄՆ-ում մեկ տվյալագետի միջին ամսական գումար աշխատավարձը գերազանցում է 8000 ԱՄՆ դոկարը՝ տվյալագիտական ծառայությունների արտահանումը կարող է ռազմավարական նշանակություն ունենալ ԶԶ ընկերությունների համար:

Մեծ տվյալների կառավարմանն ուղղված ներդրումներ

Ամպային համակարգում պահպանվող US-ն տվյալագետին վերլուծության համար հասանելի դարձնելու նպատակով գոյություն ունեն տեխնոլոգիաներ, որոնք կարելի է ասել կամուրջ են ծառայում US-ի և տվյալագետների միջև: Այդպիսի տեխնոլոգիաներից են Hadoop և Spark համակարգերը, որոնք հնարավորություն են տալիս փոքր մասերի բաժանել US-ն, վերլուծությունը կիրառել մասերից յուրաքանչյուրի վրա

⁸⁵ <http://www.quantec.co.za/software/stata/>, (02/12/16)

⁸⁶ <https://www.sas.com/store/index.ep>, (02/12/16)

⁸⁷ Տե՛ս O'REILLY, "2015 Data Science Salary Survey", էջ 6.

<https://duu86o6n09pv.cloudfront.net/reports/2015-data-science-salary-survey.pdf>

միաժամանակ՝ առցանց միջավայրում, և այնուհետև միացնել վերլուծված մասնիկները մեկ ամբողջական բազայի մեջ: Վերոնշյալ համակարգերն իրականացնում են այս գործընթացն ավտոմատաբար, որն էլ նրանց նկատմամբ մեծ պահանջարկի հիմքն է հանդիսանում: Հատկանշական է, որ չնայած այդ համակարգերի առավել ություններին և նրանց նկատմամբ եղած պահանջարկին՝ և՛ Hadoop-ը, և՛ Spark-ը լիովին անվճար են: Այսպիսով, ՄՏ-ի կառավարման համար անհրաժեշտ համակարգերի ներդրումը նույնպես ծախսատար չէ:

2արգացման նախադրյալները ՀՀ-ում

Մեծ տվյալների ոլորտը նոր զարգացում ապրող ոլորտ է, որն օրեցօր ավելի է ընդարձակվում: Սույն հետազոտությունը ներկայացրեց ՄՏ-ի տնտեսագիտական վերլուծության անհրաժեշտությունը, առանձնահատկությունները և գործիքակազմը՝ այդ ամենը դիտարկելով որպես գործարար որոշումների կայացման հիմնախնդիր: Ելնելով խնդրո առարկայի արդիականությունից և բնական ռեսուրսների սահմանափակությունից՝ ՄՏ-ի ոլորտը կարող է ՀՀ տնտեսության համար խթան հանդիսանալ և դառնալ առաջնային թիրախ: Ոլորտի զարգացման համար անհրաժեշտ են որոշակի նախապայմաններ, որոնք փոխկապակցված են վերոնշյալ ներդրումային ուղղությունների հետ: Այդ նախապայմանների բավարարումը կարող է ՀՀ-ում ՄՏ-ի էկոհամակարգի ձևավորվման հիմք հանդիսանալ:

▪ **մասնագիտական ուսումնական հաստատություններում և ուսումնական կենտրոններում տվյալագիտության և առնչվող կրթական ծրագրերի ներդրում և ոլորտին հարող առկա ծրագրերի բարելավում:** Լինելով նոր ձևավորվող գիտություն՝ տվյալագիտությանը հարող ուսումնական ծրագրեր դեռևս գրեթե չկան: Հարվարդի, Սթենֆորդի և աշխարհի լավագույն այլ համալսարանների կողմից տվյալագիտության մագիստրոսական ծրագրեր միայն վերջերս են սկսվել առաջարկվել⁸⁸: Հատկանշական է, որ ելնելով տվյալագիտության բազմաբնագաբառությունից այդ կրթական ծրագրերը առաջարկվում են համալսարանների տարբեր

⁸⁸ Տե՛ս <http://harvardmagazine.com/2017/03/data-science-institute>, (21/03/2017)

Ֆակուլտետների կողմից՝ սկսված գործարարության ֆակուլտետից և վերջացված առողջապահության կամ վիճակագրության ֆակուլտետով:

ՀՀ-ում դեռևս որպես այդպիսին տվյալագիտության ֆորմալ ուսումնական ծրագիր գոյություն չունի: Ելնելով այն հանգամանքից, որ տվյալագետները պետք է ունենան գործարարական/տնտեսագիտական գիտելիքներ և տիրապետեն համակարգչային հմտությունների՝ առաջակվում է ՄՏ-ի վերլուծության ուսումնական ծրագրեր առաջարկել տնտեսագիտության և/կամ կառավարման բակալավրի ծրագրերում, որտեղ ուսանողները կկարողանան ձեռք բերել անհրաժեշտ հմտությունները: Մեկ կամ երկու ուսումնական ծրագրի առաջարկումը կարող է հիմք հանդիսանալ մոտապագայում առավել մասնագիտացված մագիստրոսական ծրագրի ստեղծման համար: Միաժամանակ առաջակվում է փոփոխություն մտցնել տվյալագիտությանը հարող և ներկայումս դասավանդվող առարկաների ուսուցման գործիքակազմում: Ավելի կոնկրետ, առաջակվում է տնտեսաչափություն և/կամ վիճակագրություն և առարկա դասավանդելիս միայն «ավանդական» տվյալների վերլուծությանը համապատասխան SPSS, Stata, Eviews կամ այլ ծրագրային փաթեթների փոխարեն դասավանդման գործընթացում կիրառել Python, R կամ այլ փաթեթներ, որոնք համապատասխան են և՛ «ավանդական», և՛ Մեծ տվյալների վերլուծության պարագայում:

▪ **Մեծ տվյալների պահպանման ամպային համակարգերի տեղակայում:** Մեծ տվյալների պահպանման համար նախատեսված ամպային համակարգերի հիմքում իրականում ընկած են «տվյալների կենտրոնները» (անգլերեն՝ data center), որոնք որոշակի տարածքում կոնտրոնոցված և համացանցից անընդհատ և անխափան կապ ունոցող սարքավորումների ամբողջություն են: ՀՀ-ում տվյալների կենտրոնի ստեղծումը հնարավորություն կտա առցանց ամպային ծառայություններ արտահանել, որոնք, ինչպես նշվեց վերևում, բավականին պահանջված են: Նման կենտրոնում 1000-ից ավել աշխատակից ունեցող քսան ամերիկյան կազմակերպության տվյալներ պահպանելը կապահովի տարեկան 40 մլն ԱՄՆ դոլարի

ներհոսք⁸⁹: 2016թ. դրույթամբ ՀՀ ընթացիկ հաշվի պակասուրդը կազմել է 238.1 մլն ԱՄՆ դոլար⁹⁰: Ստացվում է, որ մեկ տվյալների կենտրոնի կողմից ընդամենը քսան հաճախորդի սպասարկումը կարող է հանգեցնել միջինում ՀՀ ընթացիկ հաշվի պակասուրդի գրեթե 1/6-ով (16.8%) կրճատմանը՝ մեկ քայլ մոտեցնելով ակտիվ վճարային հաշվեկշռի ապահովմանը:

Տվյալների կենտրոնի ստեղծմանն ուղղված ներդրումները նաև կարող են գրավիչ լինել մասնավոր կազմակերպությունների համար՝ իրենց բարձր եկամտաբերությամբ: ՀՀ-ում նման կենտրոնի շահութաբերության ներքին նորման գնահատվում է 24%⁹¹: Այս թվի հիմքում ընկած են տարածքի գնման/վարձակալման, պահպանման, աշխատուժի և այլ գործոնների մրցակցային արժեքը ՀՀ-ում: Թերևս սա էլ հենց պատճառ է հանդիսացել այն հանգամանքի, որ ամպային լուծումներ առաջարկող առաջատար ընկերություններից մեկը՝ Cloudflare-ն, 2017թ. մարտի 20-ին իր կայքում պաշտոնապես հայտարարեց երևանում նոր տվյալների կենտրոնի բացման պլանների մասին⁹²: Վերջինս մեկ քայլ ավելի է մոտեցնում ՀՀ-ում ՄՏ-ի էկոհամակարգի ձևավորմանը:

▪ **Մեծ տվյալների ոլորտում գիտական հետազոտությունների իրականացում:** Հանդիսանալով երիտասարդ ոլորտ՝ տվյալագիտության ուսումնասիրության աստիճանը դեռևս բավականին ցածր մակարդակի վրա է: Այնուամենայնիվ, ցանկացած ոլորտի զարգացման գործընթացում կրթական և գիտական հաստատությունները կարող են որոշիչ դեր խաղալ: Հարկավոր է նշել, որ առաջատար համալսարանները միայն նոր են սկսում ներդրումներ կատարել ՄՏ-ի ոլորտի գիտական ուսումնասիրության վրա: Այս տեսանկյունից հատկապես առանձնանում է Հարվարդի համալսարանը, որը 2017թ. մարտի 27-ին հայտարարեց

⁸⁹ 20 x 2 = 40 մլն ԱՄՆ դոլար, համաձայն IDG տվյալների (տե՛ս [IDG, “2016 Cloud computing summary”](#))

⁹⁰ Տե՛ս ԿԲ, ՀՀ վճարային հաշվեկշիռ, 12.06.2017, <https://www.cba.am/am/SitePages/statexternalsector.aspx>

⁹¹ <http://www.investmentprojects.am/en/67/>, (21/03/2017)

⁹² Տե՛ս <https://blog.cloudflare.com/yerevan-armenia-cloudflare-data-center-103/>, (21/03/2017)

«Տվյալագիտությունը նախաձեռնությունների կենտրոնի» (անգլերեն՝ Data Science Initiative) բացման մասին⁹³:

ՀՀ-ում Մեծ տվյալների էկոհամակարգի ստեղծման կարևորագույն նախապայմաններից է տվյալագիտական ուղղվածությունը գիտահետազոտական կառույցների ստեղծումը կամ գոյություն ունեցող կառույցներում տվյալագիտական ուղղվածության առանձնացումը:

• Մեծ տվյալների հավաքագրման և հասանելիության մեխանիզմների ներդրում և օգտագործման խրախուսում: Վերոնշյալ բոլոր նախապայմանները կարևոր դեր են խաղում ՀՀ-ում Մեծ տվյալների ոլորտի զարգացման համար: Այնուամենայնիվ, ոլորտը չի կարող զարգանալ եթե բացակայի թերևս ամենակենտրոնական տարրը՝ ՄՏ-ն: Ինչպես նշվեց վերևում՝ ՄՏ-ի հավաքագրման միանշանակ և համասեռ եղանակ գոյություն ունի: Տարբեր կազմակերպություններ կարող են տարբեր մոտեցումներ կիրառելով հավաքագրել ՄՏ-ն և ինքնուրույն որոշել, թե այդ տվյալների որ մասն է հասանելի օգտատերերի համար: Սակայն կառավարությունը կարող է հստակ քայլեր ձեռնարկել հանրապետության մակարդակով ՄՏ-ի հավաքագրման և հասանելիության ապահովման նպատակով:

Ներկայումս Ազգային վիճակագրական ծառայությունը (ԱՎԾ), Կենտրոնական բանկը (ԿԲ), Արտակարգ իրավիճակների նախարարությունը (ԱԻՆ), Պետական եկամուտների կոմիտեն (ՊԵԿ) և մի շարք այլ պետական մարմիններ ունեն իրական ժամանակատվածում մեծ ծավալով տվյալների հավաքագրման և հասանելիության ապահովման ներուժ: ԱՎԾ առանձին հարցումների տվյալների, ի հարկե, Մեծ տվյալներ համարվել չեն կարող, քանզի չեն բավարարում ՄՏ-ի սահմանման 3 բնութագրիչներին: Այնուամենայնիվ, արխիվում առկա բոլոր տվյալները միասին արդեն կարող են դասակարգվել որպես Մեծ տվյալ: ԱՎԾ կայքում տվյալները հասանելի են դարձել վերլուծության համար մատչելի ֆորմատով (ներկա դրությամբ՝ որպես Excel ծրագրի փաստաթուղթ) միայն վերջերս: Առաջարկվում է այդ տվյալները պահպանել մեկ

⁹³ Տե՛ս <http://news.harvard.edu/gazette/story/2017/03/harvard-launches-sweeping-data-science-initiative/>

համընդհանուր կենտրոնում և ապահովել ԿԾԻ հասանելիություն, որը ԱՎԾ տվյալների կիրառությունը ավելի հեշտ կոարձնի վերլուծաբանների համար (վերջինս ռեսպոնդենտ կազմակերպությունների կողմից նշված հիմնական խոչընդոտներից մեկն էր):

ԿԾԻ ստեղծման մոտեցումը արդեն իսկ որդեգրվել է մի շարք զարգացած երկրների կողմից: ԱՄՆ կառավարությունը ՄՏ-ի իրական ժամանակահատվածում վերլուծման տարատեսակ վահանակներ է կիրառում՝ տնտեսության տարբեր ոլորտներում որոշումների կայացման նպատակով⁹⁴: Առաջարկվում է ԿԾԻ ստեղծման մոտեցումը կիրառել ՀՀ-ում վերոնշյալ պետական կառույցների կողմից: Վերջինս ոչ միայն հնարավորություն է կտա ՀՀ-ում էկոհամակարգի ձևավորմանը, այլև կարող է հիմք հանդիսանալ առավել արդյունավետ մակրոտնտեսական քաղաքականության մշակման հարցում:

⁹⁴ Տե՛ս <https://learn.arcgis.com/en/arcgis-book/chapter9/>, (04/03/2017)

ԵՐԱԿԱՑՈՒ ԹՅՈՒՆՆԵՐ

Հետազոտության արդյունքում հանգել ենք եզրակացությունների և նրանցից բխող առաջարկությունների, որոնք ներկայացված են սույն հատվածում: Հիմնական եզրահանգումները դասակարգվել են 3 խմբի՝ ելնելով բնույթից և ուրոտային կիրառելիությունից: Ստորև ներկայացված են յուրաքանչյուր առանձին խմբին պատկանող ստացված հիմնական արդյունքները, դրանցից բխող եզրահանգումներն ու առաջարկությունները:

Հետազոտական եզրահանգումներ

Աշխատության սկզբնամասում նշվեց, որ ՄՏ-ի ուրոտը բավականին երիտասարդ է և իր զարգացումն է ապրում ոչ ավանդական եղանակով: Այդ ամենը վկայում է տվյալ ուրոտում տնտեսագիտական հետազոտությունն իրականացնելու անհրաժեշտության մասին: Սույն ուսումնասիրությունը մեզ հայտնի առաջին աշխատանքն է, որը փորձ է կատարում համապարփակ ձևով վերլուծել և ներկայացնել ՄՏ-ի դերը, կարևորությունը, կիրառությունները, արդյունավետությունը և գործիքակազմը՝ գործարարորոշումները ընդունման տեսանկյունից: Հետազոտության արդյունքում մշակվել և կազմվել է ցուցանիշների համակարգ, որը թույլ է տալիս գնահատել ՄՏ-ի օգտագործումը ընկերության կողմից: Վերջինս կարող է հիմք հանդիսանալ ապագա հետազոտությունների, որոնք կարող են.

- Ուսումնասիրել այս աշխատանքում մշակված գնահատականների և կազմակերպության \$ինանսական ցուցանիշների մեջ եղած կորելացիոն և/կամ պատճառահետևանքային կապերը:
- Գնահատել ՄՏ-ի օգտագործման արդյունավետության մատրիցի մեկ մակարդակից մյուսն անցնելիս կազմակերպության \$ինանսական, արտադրողականության և այլ ցուցանիշների փոփոխությունը:
- Իրականացնել մատրիցի մեկ մակարդակից մյուսն անցնելիս պահանջվող ներդրումների արդյունավետության գնահատում՝ ընդգծելով

մակարդակներից յուրաքանչյուրի հարաբերական արդյունավետություները:

Հետազոտության երկրորդ գլխում ներկայացվում է «ավանդական» տվյալների վերլուծության և տվյալագիտության հիմնական առանձնահատկությունները գործարար կառավարման տեսանկյունից: Տվյալագիտության հնարավոր համեմատական արդյունավետությունն ընդգծելու նպատակով իրականացվել է վարկային տվյալների վերլուծություն: Ապագա հետազոտությունները կարող են ընդլայնել այս աշխատանքի ուսումնասիրությունը և:

- իրականացնել նմանատիպ վերլուծություն տնտեսության այլ ոլորտների համար,
- ուսումնասիրել գործարար կառավարման այն որոշումները, որոնց ընդունման գործընթացում ավելի նպատակահարմար է օգտագործել տվյալագիտական մոտեցումները և այն որոշումները, որոնց դեպքում «ավանդական» տվյալների վերլուծության մեթոդներն ավելի կիրառելի են,
- ուսումնասիրել խնդիրը տնտեսամաթեմատիկական մոդելավորման տեսանկյունից:

Իրական հատվածին հարող եզրահանգումներ

Սույն հետազոտության արդյունքների մեծամասնությունը կիրառելի են հենց իրական հատվածում, քանզի ՄՏ-ի դիտարկումը անց է կացվել հենց գործարար կառավարման տեսանկյունից: Նախ և առաջ՝ պետք է նշել, որ տվյալագետները, ինչպես ներկայացված էր առաջին գլխում, Համաշխարհային տնտեսագիտական \$որոումի կողմից 2 տարի անընդմեջ ճանաչվել են աշխարհում ամենապահանջված մասնագետները: Հետազոտության արդյունքում պարզվեց, որ ՀՀ-ում տվյալագիտությամբ զբաղվող մասնագետների միջին աշխատավարձը նշանակալիորեն գերազանցում է մեր տնտեսության բոլոր ոլորտներում գրանցված պաշտոնական վիճակագրական միջին ցուցանիշին: Այդուհանդերձ, վերջինս զգալիորեն ցածր է զարգացած երկրներում (օրինակ՝ ԱՄՆ) նույն մասնագետների միջին

աշխատավարձին՝ նույնիսկ գնաճի ճշգրտումից հետո: Ելնելով այս հանգամանքից՝ մշակվել են հետևյալ առաջարկությունները.

- Ընկերություններին առաջարկվում է պարզել, թե գործարար որոշումների կայացման գործընթացում իրենց հատուկ ձեռներեցական խնդիրների լուծման համար որ վերլուծությունն է ավելի նպատակահարմար («ավանդական» կամ տվյալագիտական): Այս հարցի պատասխանը թույլ կտա հասկանալ, թե որ մասնագետին է ավելի նպատակահարմար աշխատանքի ընդունել:
- Չափի առնելով այն հանգամանքը, որ տվյալագետների միջին վարձատրությունը, նույնիսկ սպառողական զամբյուղի արժեքը հաշվի առնելուց հետո, զգալիորեն գիջում է ԱՄՆ-ում տվյալագետների միջին վարձատրությանը՝ առաջարկվում է կենտրոնանալ արտասահմանյան կազմակերպություններին տվյալագիտական ծառայություններ մատուցելու/արտահանելու վրա, որը ՉՉ ընկերություններին համար կարող է նոր շահույթի աղբյուր հանդիսանալ:

Աշխատության երրորդ գլխում մշակված գործիքակազմը առաջարկվում է օգտագործել հետևյալ նպատակներով.

- Տվյալագետների և «ավանդական» տվյալների վերլուծաբանների միջև ընտրություն կայացնելու հարցին տվյալահենք պատասխան տալու համար առաջարկվում է կիրառել ՏՄԳ-ն, որը հնարավորություն կտա պարզել տվյալների մեծությունը և հասկանալ տվյալագետի աշխատանքի ընդունելու անհրաժեշտությունը: ՏՄԳ բարձր մակարդակը հուշում է վերջինիս անհրաժեշտության մասին, իսկ ցածր մակարդակի դեպքում ձեռներեցը պետք է որոշում կայացնի՝ հիմնվելով սեփական կազմակերպության գործարարական խնդիրների, հասանելի ռեսուրսների և տվյալների հավաքագրման հետ կապված ապագալանների վրա:

- Վերլուծաբանների ռեսուրսներն արդյունավետ օգտագործելու և գործարար որոշումները քանակապես չափելի դարձնելու նպատակով առաջարկվում է կիրառել S49-ն, որը հնարավորություն կտա ձեռներեցին հասկանալ, թե գործարար կառավարման առանցքներից յուրաքանչյուրում որոշում կայացնելիս որքանով են կիրառվում տվյալները:
- Տվյալների հավաքագրման և կիրառման համեմատական արդյունավետության գնահատման համար առաջարկվում է կիրառել USOU մատրիցը, որը հնարավորություն կտա ոչ միայն պարզել ընկերության ներկայիս դիրքը, գնահատել հասանելի տվյալների օգտագործման արդյունավետությունը, այլև իրականացնել ապագա ներդրումների պլանավորում:

Մակրոտնտեսական/մեզոտնտեսական եզրահանգումներ

Յեռագոտության արդյունքում պարզվել են US-ի վերլուծության կիրառության հետկապված ՅՅ-ում առկա հիմնական խոչընդոտները և զարգացման նախադրյալները: Ելնելով այդ խոչընդոտների բնույթից, ինչպես նաև տվյալ ազետների և US-ի նկատմամբ մեծ պահանջարկից՝ առաջարկվում է:

- Մշակել տվյալ ոլորտի մասնագետների պատրաստող կրթական ծրագրեր, որոնք հնարավորություն կտան մեծացնել տվյալ ազիտական ռեսուրսը ՅՅ-ում և օգտագործել գնային առավելությունը US-ի վերլուծության ծառայություններ արտահանելու նպատակով:
- Ստեղծել US-ի վերլուծության վիճակագրական և աշխատանքային հարթակ, որը թույլ կտա պարզել այն կազմակերպություններին և անձանց, որոնք ներկայումս զբաղվում են US-ի վերլուծությամբ:
- Ստեղծել ազգային վիճակագրական տվյալների համապարփակ հարթակ, որը հնարավորություն կտա իրական ժամանակահատվածում հասանելի դարձնել ներկայումս հավաքագրվող տվյալները:

- Ներդնել ՄՏ-ի պահպանման ամպային համակարգեր, որոնք կկրճատեն ՄՏ-ի վերլուծությամբ զբաղվող տեղական կազմակերպությունների ծախքերը և միաժամանակ հնարավորություն կտան ՄՏ-ի պահպանման ծառայություններ արտահանել:

Այսպիսով՝ հաշվի առնելով ՄՏ ոլորտի աճն ու ՀՀ-ում առկա նախադրյալները՝ պետք է իրականացնել տվյալ ագիտության էկոհամակարգի ձևավորմանն ու զարգացմանն ուղղված քաղաքականություն:

ՕԳՏԱԳՈՐԾՎԱՆՑ ԳՐԱԿԱՆՈՒ ԹՅԱՆ ՑԱՆԿ

Հոդվածներ, գրքեր և հրատարակումներ

1. A History of Accountancy, New York State Society of CPAs, November 2003 – [31]
2. Arrow, K. J. (1965). "Aspects of the Theory of Risk Bearing" – [40]
3. Awadallah A., "How Hadoop Revolutionized Business Intelligence and Advanced Data Analytics", November 2011 – [33], [34]
4. Barua A., Manis D., Mukherjee R., Measuring the Business Impacts of Effective Data, University of Texas, Austin 2015 – [24], [25]
5. Blackwell D., "Equivalent Comparisons of Experiments". The Annals of Mathematical Statistics, 1953, էջ Եր 265-272 – [21]
6. Brynjolfsson E., Heekyung K., Hitt L., Strength in Numbers: How Does Data-Driven Decision-making Affect Firm Performance?", 2011 – [26]
7. Cappello F., Djilali S., Fedak G., Herault T., Magniettea F., Nérib V., Lodygenskyc O. , "Computing on large-scale distributed systems: XtremWeb architecture, programming models, security, tests and convergence with grid", Elsevier March 2005 – [73]
8. Denning P., "Saving All The Bits", American Scientist, Volume 78, Issue 5, Oct 1990, էջ Եր 402-405 – [18]
9. Gartner Business Intelligence & Information Management Summit 2013 – [27]
10. Gartner, "Gartner Symposium, survey summary" 2016 – [41], [78]
11. Global Digital Payment Transactions Set to Reach 10 Percent Growth for First Time Finds World Payments Report 2016 – [39]
12. GlobalWebIndex, Q3 Report, 2016 – [35]
13. Harvard Business Review, "Cost of bad data", 2016 – [11]
14. IDG, "2016 Cloud computing summary" – [83], [89]
15. International Data Corporation (IDC), "Worldwide Semiannual Big Data and Analytics Spending Guide" 2015 – [42]
16. ISO 80000-1 Quantities and units. Part 1: General (1st ed.). Switzerland: ISO (the International Organization for Standardization). 2009-11-15 – [71]
17. John R. Mashey. "Big Data ... and the Next Wave of InfraStress", Usenix, April 1998 – [4]
18. Kaplan R., Norton D., "The Balanced Scorecard: Translating Strategy into Action", Harvard Business School Press, August 1996, էջ Եր 43–46 – [30]

19. KPMG, "CEO Outlook 2016", 2016, p. 20 – [10]
20. Laney, Douglas. "3D Data Management: Controlling Data Volume, Velocity and Variety". Gartner. February 2001 – [8]
21. M. Hashem Pesaran, "Econometrics, The New Palgrave: A Dictionary of Economics", v. 2, 1987, էջ 8-22 – [47]
22. Matthew A., Stanley E., "Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management", Journal of Business Logistics, 2013, 34(2), էջ 77–84 – [28]
23. McCulloch, Warren; Walter Pitts (1943). "A Logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathematical Biophysics – [51]
24. McKinsey & Co, "Big data: The next frontier for innovation, competition, and productivity" – [44]
25. McKinsey Global Institute (MGI), Big Data: The next frontier for innovation, competition, and productivity, Report, June, 2012, էջ 11 – [6]
26. Mell P., Grance T., "The NIST Definition of Cloud Computing (Technical report)". National Institute of Standards and Technology: U.S. Department of Commerce, September 2011 – [72]
27. MIT Sloan Management Review, "Minding the Analytics Gap" volume 56, no. 3, Spring 2015, pp. 63-68 – [60]
28. Morgan Stanley Research Center, "Mobile Internet Report 2016" – [38]
29. O'REILLY, "2015 Data Science Salary Survey" – [57], [66], [87]
30. Paul Zikopoulos, "Understanding Big Data", 2012 – [13]
31. Petter Bae Brandtzæg, "Big Data, for better or worse: 90% of world's data generated over last two years." ScienceDaily. ScienceDaily, May 2013 [1], [2]
32. Reinsel, D. Gantz, J. Extracting value from chaos. 2011, էջ 6 – [7]
33. Sinan Aral, Erik Brynjolfsson, and Marshall W. Van Alstyne, "Information, Technology, and Information Worker Productivity" Information Systems Research, March, 2012 – [23]
34. Tarabanis K. Kalampokis E., Tambouris E., "Understanding the Predictive Power of Social Media", Internet Research, Vol.23, September 2013 – [14]
35. The Financial Brand, "Big Data: Profitability, Potential and Problems in Banking", 2015 – [76]
36. Wharton People Analytics, Wharton, University of Pennsylvania – [74]

37. Ваганян ГА, “Машинная графика в управлении”, Монография, “Айастан”, Ереван, 1985 – [19]
38. Ваганян ГА, Ваганян ОГ, “Виртуальные технологии менеджмента (системотехника менеджмента)”, Монография, “Нжар”, Ереван, 2005 [20]
39. Դավթյան, Յ., Տվյալների վերլուծության կիրառությունը ֆինանսական շուկայում: «Եվրոպական ակադեմիա» գիտական հոդվածների ժողովածու, №7, Երևան, 2016 թ., էջ 231-237 – [67]
40. Դավթյան, Յ., Big data in different sectors of the economy: application scenarios. «Ակունք» ուսումնամեթոդական հոդվածների ժողովածու, №2 (14), Երևան, 2016 թ., էջ 122-125 – [29]
41. Դավթյան, Յ., Մեծ տվյալների ոլորտում ներդրումների իրականացման արդյունավետությունը: «Այլընտրանք» գիտական հանդես, №4, Երևան, 2016 թ., էջ 228-232 – [79]
42. Դավթյան, Յ., Մեծ Տվյալների տնտեսագիտական վերլուծության առանձնահատկությունները գործարար կառավարման սպարեզում: «Ֆինանսներևէկոնոմիկա» հանդես, 1-2 (197-198), Երևան, 2017 թ., էջ 12-15 – [48]
43. Դավթյան, Յ., ՀՀ կազմակերպությունների կողմից մեծ տվյալների օգտագործման համեմատական արդյունավետության գնահատում: «Ֆինանսներևէկոնոմիկա» հանդես, 3-4 (199-200), Երևան, 2017 թ., էջ 26-28 – [75]
44. Դավթյան, Յ., Տվյալագետների եկամտային խմբին ազդող փոփոխականների ռեգրեսիոն վերլուծություն: «Այլընտրանք» գիտական հանդես, №1, Երևան, 2017 թ., էջ 269-276 – [62]
45. Մարգարյան Յ., Դավթյան Յ., Տվյալագետների աշխատուժի առաջարկը ՀՀ աշխատաշուկայում. ընտրանքային հարցման արդյունքների վերլուծություն: «Ակունք» ուսումնամեթոդական հոդվածների ժողովածու, №1 (15), Երևան, 2017 թ., էջ 240-246 – [54]
46. ՀՀ ԱԿԾ, «Աշխատանքի շուկան Հայաստանի Հանրապետությունում», 2016 թ., էջ 55 – [59]
47. ՀՀ ԱԿԾ, «Հայաստանի Հանրապետության սոցիալ-տնտեսական վիճակը 2016թ. հունվարի դրությամբ», 2016թ. էջ 53 – [61]

48. Ձեռնարկ ու թյ ու ն ն երի ի ն կ ու բ առ ուր հ ի մ ն ադր ա մ , « Զ Զ Տ Տ ու լ ուր տի հ աշ վ ե տվ ու թյ ու ն », 2015 թ, Է ջ 28 – [58]
49. Ռ ազ մ ա վ ա ր ա կ ա ն լ ու ծ ու մ ն երի գ ն ա հ ա տ ա կ ա ն ն եր , IBM ը ն կ եր ու թյ ա ն տ ա ր ե կ ա ն հ աշ վ ե տվ ու թյ ու ն , 2015 – [32], [68]

Կայքեր

50. <http://ebusiness.mit.edu/erik/ITandBusinessvalue.html> – [22]
51. <http://harvardmagazine.com/2017/03/data-science-institute> – [88]
52. <http://news.harvard.edu/gazette/story/2017/03/harvard-launches-sweeping-data-science-initiative/> – [93]
53. <http://nordicapis.com/tracking-the-growth-of-the-api-economy/> – [82]
54. <http://www.ibmbigdatahub.com/> – [9]
55. <http://www.investmentprojects.am/en/67> – [91]
56. <https://www.cba.am/am/SitePages/statexternalsector.aspx> – [90]
57. <http://www.quantec.co.za/software/stata/> – [85]
58. <https://archive.ics.uci.edu/ml/datasets/> – [69]
59. <https://aws.amazon.com/api-gateway/> – [80]
60. <https://blog.cloudflare.com/yerevan-armenia-cloudflare-data-center-103/> – [92]
61. <https://go.developer.ebay.com/> – [81]
62. <https://learn.arcgis.com/en/arcgis-book/chapter9/> – [94]
63. <https://trends.google.com/> – [5], [49], [52], [53]
64. <https://www.google.com/insidesearch/howsearchworks/thestory/index.html> – [12]
65. <https://www.ibm.com/us-en/marketplace/spss-statistics/purchase#product-header-top> – [84]
66. <https://www.sas.com/store/index.ep> – [86]
67. <https://www.statista.com/> – [36], [37]
68. <https://www.weforum.org/> – [43], [46]

ՀԱՊԱՎՈՒՄՆԵՐԻ ՑԱՆԿ

ԳՏՅ	Գլ ոբալ տեղորոշման համակարգ
ԱԻՆ	Արտակարգ իրավիճակների նախարարություն
ԱՄՆ	Ամերիկայի Միացյալ Նահանգներ
ԱՎԾ	Ազգային վիճակագրական ծառայություն
ԵԿՏԱ	Եվրոպական կրթական տարածաշրջանային ակադեմիա
ԿԲ	Կենտրոնական բանկ
ԿԾԻ	Կիրառական ծրագրավորման ինտերֆեյս
ՀԱՅ	Հայաստանի ամերիկյան համալսարան
ՀՀ	Հայաստանի Հանրապետություն
ՀՆԱ	Համախառն ներքին արդյունք
ՄՏՕԱ	Մեծ տվյալների արդյունավետ օգտագործում
ՏԿԳ	Տվյալների կիրառություն գնահատական
ՏՀՏ	Տեղեկատվական և հաղորդակցություն տեխնոլոգիաներ
ՏՄԳ	Տվյալների մեծություն գնահատական
ՏՏ	Տեղեկատվական տեխնոլոգիաներ
ՎՄ	Վստահելիություն միջակայք
ՑՀՀ	Ցուցանիշների հաշվեկշռված համակարգ

ՀԱՎԵԼ ՎԱՃ 1. ՔԱՆԱԿԱԿԱՆ ՀԱՐՑՄԱՆ ՀԱՐՑԱԹԵՐԹ

1. Սեռը

1	Իգական
2	Արական

2. Տարիքը

3. Աշխատավարձային խումբը (ամսական գումարը տԱՄՆ դոլար)

1	ավելի քիչ քան 300
2	300 – 500
3	500 – 1000
4	1000 – 1500
5	1500 – 2000
6	ավելի քան 2000

4. Ներկայիս աշխատանքը լքելու պարագայում նոր աշխատանք գտնելու Ձեր հնարավորությունը ներքի գնահատականը:

1	2	3	4	5
Շատ հեշտ				Շատ բարդ

5. Ձեր մասնագիտական փորձը

1	1 տարուց պակաս
2	1 – 3
3	3 – 5
4	5 – 8
5	8 տարուց ավելի

6. Ձեր կազմակերպության աշխատակիցների քանակը

1	6 հոգուց պակաս
2	6 – 20

3	21 – 50
4	51 – 100
5	100-ից ավելի

7. Ձեր կազմակերպության գործունեության ոլորտը

1	Ֆինանսներ
2	SS
3	Այլ՝ նշել ստորև

8. Ձեր հաստիքը

9. Յետևյալ անդումներից, ընտրենք նրանք, որոնք համապատասխանում են Ձեզ

Ես ստացել եմ մագիստրոսական կրթություն
Ես ունեմ գիտությունների թեկնածուի/դոկտորի աստիճան
Իմ կրթությունը տնտեսագիտության/վիճակագրության ոլորտում է
Իմ կրթությունը մաթեմատիկայի/ ծրագրավորման ոլորտում է

10. Ինչպե՞ս եք ձեռք բերել ձեր մասնագիտական գիտելիքները:

1	Ֆորմալ կրթություն
2	Ոչ ֆորմալ կրթություն

11. Մեծ տվյալների հետաշխատելու Ձեր ներկա հմտությունների գնահատականը

1	2	3	4	5
Սկսնակ				Հմուտ

12.Նշել այն գործիքները, որոնք Ձեր կողմից առօրյայում կիրառվում են Մեծ տվյալների վերլուծության նպատակով

--

13.Առօրյայում Ձեր կողմից վերլուծվող տվյալների ծավալը

1	2	3	4	5
Փոքրածավալ				Մեծածավալ

ՀԱՎԵԼ ՎԱԾ 2. ՈՐԱԿԱԿԱՆ ՀԱՐՑՄԱՆ ՀԱՐՑԱԹԵՐԹ

1. Կազմակերպության տարիքը

1	մինչև 3 տարի
2	3 – 5 տարի
3	5 – 10 տարի
4	ավելի քան 10 տարի

2. Ձեր կազմակերպության կողմից գործարար կառավարման 4 առանցքներից յուրաքանչյուրում որոշում կայացնելիս Մեծ տվյալների կիրառությունը (յուրաքանչյուր սյան մեջ նշել մեկ վանդակ)

	Հաճախորդներ	Աշխատակիցներ	Գործառույթներ	Ֆինանսներ
1				
2				
3				
4				
5				

3. Ձեր կազմակերպության կողմից գործարար կառավարման 4 առանցքներից յուրաքանչյուրում որոշում կայացնելիս Մեծ տվյալների ապագա կիրառության պլանները (յուրաքանչյուր սյան մեջ նշել մեկ վանդակ)

	Հաճախորդներ	Աշխատակիցներ	Գործառույթներ	Ֆինանսներ
1				
2				
3				
4				
5				

4. Նշել որոշումներ, որոնց կայացման գործընթացում կիրառվում է Մեծ տվյալների վերլուծություն:

1			
2			
3			
4			
5			

5. Նշել Մեծ տվյալների վերլուծության հետ կապված Ձեր կազմակերպությանը հանդիպող հիմնական խոչընդոտները:

1			
2			
3			
4			
5			

6. Ձեր կազմակերպության կողմից հավաքագրվող տվյալների Մեծությունը:

	Ճավալ	Շարժունություն	Բազմազանություն
0			
1			
2			
3			
4			