

ВЪЗМОЖНОСТИ ЗА ПОВИШАВАНЕ КАЧЕСТВОТО НА ДАННИТЕ ОТ ОНЛАЙН АНКЕТИ

Гл. ас. д-р Радослав Йорданов

Резюме

С появата и развитието на интернет и свързани с него приложения се промениха много области на обществения живот. Промениха се и начините за намиране на маркетингова информация. Все по-рядко анкетите се отпечатват на хартия и все по-често се провеждат в уебсреда. Причините са прости – онлайн анкетите се провеждат по-бързо и по-евтино, лесно се разпространяват, предлагат полезни функционалности, каквито класическите анкети не притежават. Съществено предизвикателство пред маркетинговите изследователи е осигуряването на представителност на онлайн анкетите. Основната цел на това изследване е да провери доколко структурата на извадката, получена чрез онлайн анкета, съответства на структурата на изучаваната съвкупност. Второстепенна цел на изследването е да се провери доколко респондентите реагират на стимули, предложени за участие в онлайн анкета. В настоящото изследване са представени резултатите от три онлайн анкети. Всяка от тях е адресирана към проучването на специфична съвкупност: фенове на Facebook страница; посетители на сайт; население на България на възраст от 13 до 54 г. Две анкети са реализирани посредством Facebook реклама, а една – сред посетителите на уебсайт. За оценка на съответствието между извадка и генерална съвкупност са използвани два инструмента – χ^2 критерий за една извадка и интегрален коефициент на Кирил Гатев. При допитването до феновете на Facebook страница е получена извадка, която плътно се доближава до изучаваната съвкупност по променливите „пол“ и „местоживееене“. При валидирането на извадките са използвани следните източници на информация за изучаваните съвкупности: Национален статистически институт; раздел „Анализи“ на страницата във Facebook; отчети на Google Analytics за профила на посетителите на уебсайт. При двете допитвания, реализирани чрез Facebook реклама, е постигната много ниска цена за една попълнена анкета. При допитването до посетителите на сайт цената за една попълнена анкета е на практика нулева. За да се провери как влияе материалното стимулиране върху активността на респондентите, са предложени два вида награди – USB Flash памет (3 бр.) и външен хард диск (2 бр.). В конкретния случай стимулирането на целевите респонденти не повишава възвръщаемостта на анкетите в степен, оправдаваща разходите за закупуване и доставяне на наградите.

Ключови думи: онлайн анкета, представителност на извадка, възвръщаемост на анкета, фактори за повишаване на възвръщаемостта, Facebook реклама.

JEL: M31, M37.

OPPORTUNITIES FOR INCREASING THE QUALITY OF DATA FROM ONLINE SURVEYS

Head Assist. Prof. Radoslav Yordanov, PhD

Abstract

The emergence and evolution of Internet and related applications have changed many areas of public life. The methods of gathering marketing information have also changed. Less and less often surveys are printed on paper and are conducted in a web environment. The reasons are simple – online surveys are conducted faster and cheaper, they are easy to spread, they offer useful functionalities. A major challenge for marketing researchers is to ensure that online surveys are representative. The main purpose of this study is to examine whether the structure of a sample, drawn from an online survey, corresponds to the structure of the studied aggregation. Another objective of the study is to examine how respondents react to incentives offered for participation in an online survey. This study presents the results of three online surveys. Each survey examines specific aggregation: Facebook page fans; visitors to a web site; population of Bulgaria aged 13-54. Two surveys were conducted through Facebook advertising. One survey was conducted among the visitors of a web site. To compare the sample and general aggregation, we have used χ^2 criterion for one sample and Kiril Gatev's integral coefficient. In the case of studying Facebook page fans, the sample matches the set of variables "sex" and "residence". In order to validate the three samples, the following information sources were used: The National Statistical Institute; Facebook page Insights; Google Analytics reports. In both surveys made via Facebook advertising, a very low cost was obtained for a response. In the survey of site visitors, the cost of a response is practically zero. To test how incentives affect respondents' activity, two types of prizes are offered – USB Flash memory stick (3 pcs.) and External hard disk drive (2 pcs.). In this study, we have found that the incentives offered do not increase respondents' responsiveness to a degree justifying the costs of purchasing and delivering the prizes.

Keywords: Online survey, sample representativeness, survey (response) return rate, factors that increase the survey return rate, Facebook ad.

JEL: M31, M37.

Увод

Появата и разпространението на интернет доведоха до силно технологизиране на маркетинговите изследвания. Новите технологии притежават редица предимства спрямо конвенционалните си аналози. Така например, в сравнение с класическата анкета, онлайн анкетата осигурява по-бързо и по-евтино набиране на (възможно) по-качествени данни.

Онлайн анкетата предоставя възможност за пълноцветно оформление, адаптиране на въпросника съобразно предходни отговори на респондента, включване на аудио- или видеосъдържание, подобряване разбирането на задаваните въпроси, верифициране на отговорите в момента на попълване на анкетата, лесен достъп до изследваните единици. Интерактивният характер на онлайн анкетите повишава ангажираността на респондентите и склонността им за участие.

Изброените предимства обясняват ръста на индустрията на онлайн изследванията, в частност на онлайн анкетирането. Приходите от маркетингови изследвания (общо офлайн и онлайн), реализирани в света като цяло за периода 2009 – 2013 г., са: 2009 г. – \$33,2 млрд., 2010 г. – \$34,5 млрд., 2011 г. – \$39,5 млрд., 2012 г. – \$39,7 млрд., 2013 г. – \$40,3 млрд. (ESOMAR, 2014, р. 99). Любопитен факт е, че за 2013 г. (спрямо 2012 г.) България е третата най-бързо развиваща се европейска страна, що се отнася до показателя „приходи от маркетингови изследвания“ (ESOMAR, 2014, р. 7). Делът на онлайн маркетинговите изследвания (в рамките на маркетинговите изследвания като цяло) е различен в отделните страни. През 2013 г. от обхванатите 54 страни в доклад на ESOMAR се констатира среден дял на онлайн изследванията (в рамките на маркетинговите изследвания като цяло) в размер на 28%. За България стойността на този показател е 47%. Това нарежда страната на първо място (заедно с Япония) по показателя „дял на онлайн маркетинговите изследвания“ (ESOMAR, 2014, р. 222-223).

Онлайн анкетите се радват на бързо разпространение не само в бизнес средите. Нараства използването им и за академични цели. Това е видно от търсенето по ключови думи „Web survey“, „Internet survey“, „Online survey“, „Telephone survey“ и „Mail survey“ в Google Scholar за всяка година от периода 1994 – 2017 (по идея на Фийлдинг, Лий и Бланк (Fielding, et al., 2008, р. 10)). Резултатът от търсенето е брой на публикациите, съдържащи избраните ключови думи (в случая словосъчетания). Констатира се, че в индексиранията от Google Scholar англоезични публикации през последните години устойчиво намалява използването на словосъчетанията Telephone survey“ и „Mail survey“ и нараства употребата на словосъчетанието „Online survey“. Важна причина за това са развитието на информационните технологии и произтичащите от тях предимства на онлайн анкетите спрямо класическите форми на допитване.

Обект на изследване в тази разработка са онлайн анкетите, а неин **предмет** – възможностите за подобряване качеството на данните, набирани чрез онлайн анкети. В случая, макар че е възможно да се открият различия, приемаме понятията „онлайн анкета“, „интернет анкета“ и „уебанкета“ за взаимозаменяеми. Под качество на данните разбираме степента, в която получената извадка съответства на изучаваната генерална съвкупност по определени признаци (по друг начин казано – доколко извадката представлява умалено копие на генералната съвкупност).

Съществува разбиране, че провежданите в интернет допитвания не могат да са представителни или че не може да се провери доколко са представителни. В това изследване **се поддържа тезата**, че при определени обстоятелства (естество на изучаваната съвкупност), при използване на подходящи канали за достигане до респондентите и при съблюдаване на определени правила за провеждане на онлайн анкети, събраните данни могат да са представителни или поне силно да се доближават до свойството представителност.

В съответствие с така дефинираната теза са и **основните цели на изследването**:¹

1. Да се провери доколко онлайн извадките могат да се доближават до генералната съвкупност по променливите „пол“, „възраст“ и „местоживеене“ (най-често използваните признаци, по които се търси представителност на извадковите изследвания) при реализиране на анкети в различен контекст;
2. Да се провери дали предлагането на стимул за участие в анкетата провокира активността на респондентите до степен, оправдаваща разходите за стимулиране.

Ключова предпоставка за получаване на представителни резултати е високата възвръщаемост на анкетите. Под възвръщаемост се разбира отношението между брой получени отговори към брой достигнати адресати. По-високата възвръщаемост е предпоставка за по-плътното припокриване между структура на извадка и структура на съвкупност.² Затова предпо-

¹ Цели с второстепенно значение и стоящи извън обхвата на настоящото изследване са: (а) Да се подпомогнат административните органи на Стопанската академия във вземането на решения, касаещи приема и обучението на студенти в бакалавърска и магистърска степен. (б) Да се получи обратна връзка от завършилите Стопанската академия студенти – за удовлетвореността от обучението като цяло и от отделни негови компоненти, както и за професионалната им реализация. Данните от проведените три анкети са предоставени на ръководството на Стопанска академия „Д. А. Ценов“.

² Нека си представим следната ситуация: Изследовател желае да установи какво е отношението на потребителите към продукт X. Съставен е проект на случайна извадка и е определено кои единици попадат в нея. Търсената представителност е по променливите „пол“ и „възраст“. Започва полевата работа по набиране на данни. Оказва се, че една значителна част от целевите респонденти отказват участие в анкетата и се налага тяхното заместване с други единици. Изследователят няма как да установи дали отказалите участие респонденти са единодушни в своето негативно отношение към изучавания продукт. Да предположим, че ситуацията е точно такава. На мястото на отказалите участие респонденти с негативно отношение се избират други, които поради положителната си оценка за продукта X по-скоро са склонни да участват в анкетата. В случая ниската степен на възвръщаемост, породена от невъзможността да се провокира участието на всички групи респонденти, би довела до нереалистично завишаване на оценката за продукта.

ставките за получаване на по-висока възвръщаемост са от специален интерес в настоящото изследване.

Постановка на проблема

Наред с преимуществата, използването на онлайн анкети е съпроводено и с някои предизвикателства, най-важното от които е **осигуряването на представителност на получените резултати**. Под представителност се разбира, способността на извадката да отразява изучаваната съвкупност по ключови за изследването променливи (ESOMAR, 2015). За да е представителна една извадка, е необходимо да бъде излъчена по случаен начин и да обхваща достатъчен брой единици в зависимост от разнородността на изучаваната съвкупност (Желев, 1995, с. 194). Причините за проблемната представителност на онлайн анкетите са основно три:

а) Не всички единици от изследваните съвкупности ползват интернет (*coverage bias*). За съставяне на случайна извадка е необходимо, изследователят да разполага с основа на извадката (списък на единиците от съвкупността). Когато обект на изследване е цялото население, онлайн изследванията не могат да претендират за представителност. Най-малкото, защото не всички имат достъп до интернет. Освен това, различна е интензивността и честотата на ползване на интернет. Поради това някои единици от съвкупността биха имали по-големи, други – по-малки, а трети – нулеви шансове за попадане в извадката. Към края на 2017 г. в България 30,3% от населението на възраст от 16 до 74 години никога не са използвали интернет (НСИ, 2017). Дори и всички единици от съвкупността да имат достъп до интернет, липсата на данни за осъществяване на онлайн контакт възпрепятства достъпа до респондентите.

Проблемът с достъпа до интернет може да се преодолее, като онлайн анкетата бъде допълнена от офлайн анкета, покриваща онлайн недостъпните единици. Друг подход е възприел операторът на German Internet Panel (GIP). За провеждане на лонгитюдни изследвания, които са представителни за населението на Германия на възраст от 16 до 75 години, компанията провежда предварителни срещи с представители на случайно подбрани домакинства. При тези срещи се установява кои единици са склонни да участват в онлайн панел, както и наличието на компютър и интернет в домакинството. За да постигнат представителност на онлайн панела за цялата съвкупност, администраторите на GIP инсталират компютър и/или осигуряват достъп до интернет на лицата, които не разполагат с такива, и предоставят техническа поддръжка и обучение. В резултат на този подход се получава извадка, която е представителна за изучаваната съвкупност (Blom, 2015).

Проблемът с достъпа до интернет отпада, когато обект на изследване са съвкупности, формирани в резултат или с помощта на интернет. Такива съвкупности са например членовете на сдружения, феновете на

Facebook страници, регистър на туроператорите и туристическите агенции, клиентите на онлайн магазин и др. Това са затворени общности, състоящи се от известен брой достижими по електронен път единици.

б) Ниска степен на възвръщаемост, водеща до неизпълнение на планираните обем и структура на извадката (*non-response bias*). Дори когато извадката по своя замисъл е планирана като случайна, респондентите не могат да бъдат заставени да участват. На практика при ниска степен на отговаряне извадката от случайна се превръща в неслучайна, в извадка на отзовалите се. Високата степен на неотговаряне е силна предпоставка за непредставителност на резултатите. В такъв случай е необходимо да се предвиди по-голям брой покани за участие в анкетата. Така например, ако целта е да се анкетира посетителите на сайт, който има приблизително 50 000 уникални посетители дневно, за съставяне на извадка от 1 000 респонденти е необходимо, всеки петдесети посетител да попадне в извадката. Ако на базата на минал опит е установено, че само 10% от виделите поканата за участие се включват в анкетата, то покана за участие следва да се показва на всеки пети посетител на сайта. По-долу са коментирани някои ключови предпоставки за стимулиране участието на респондентите, за повишаване на ангажираността им и в крайна сметка за повишаване степента на възвръщаемост:

- **Дизайн на въпросника.** Онлайн анкетите позволяват пълноцветно оформление и добавяне на текст, графични обекти, аудио, видео, линкове към интернет ресурси. В рамките на въпросника респондентът може да изтегля или да качва файлове. Въпросите могат да бъдат обособени в тематични раздели. Всички тези възможности позволяват на изследователя да конструира въпросник, който да съответства на целите на изследването и на езиковото и образователно равнище на респондентите, да предразполага към участие, да улеснява разбирането и отговарянето на поставените въпроси. Добре е да се има предвид, че претрупването на страниците с текст е предпоставка, респондентът да откаже участие или да маркира отговори, без да е изчел указанията, въпросите и предложените опции за отговор. Необходим е баланс между малкото текст и риска да останем неразбрани, от една страна, и многото текст и риска да откажем респондента от участие, от друга страна. Онлайн анкетата позволява разделяне на въпросника на секции, показвани на отделни страници. При подобно структуриране респондентът преминава към следваща секция от въпроси едва след попълване на въпросите от предходната секция. Такова едно дозиране на съдържанието създава усещане за лекота на анкетата и не позволява на респондента да види накуп всички въпроси, а го заставя да маркира своите отговори в заложен от изследователя порядък.

- **Продължителност на попълване на анкетата.** Колкото повече въпроси и отговори съдържа една анкета, колкото повече усилия и време изисква нейното попълване, толкова по-малко респонденти ще са склонни да участват. Всички излишни обяснения и въпроси следва да се спестяват. Съблюдаването на принципа „кратко и ясно“ е предпоставка за по-малко откази за участие. Коректно е в поканата за участие и в самото начало на анкетата (във въведението) да се уточни очакваната продължителност на попълване. Различните респонденти попълват една и съща анкета за различно време. Тогава изглежда логично, продължителността на анкетата да бъде обявена като интервал – например от 4 до 6 минути.
- **Използване на индикатор за степента на завършеност на анкетата.** Това е графичен обект, най-често хоризонтално поле, което постепенно се запълва с преминаването на респондента към следваща секция с въпроси от анкетата. Смесът на този индикатор е да ориентира респондента, докде е стигнал и какво все още му предстои с попълването на анкетата.
- **По-голям контрол върху анкетата.** Уеббазираните анкети предлагат редица технически възможности, чието уместно използване позволява събирането на търсените от изследователя данни. Някои от тези технически възможности са: = Ако един въпрос е избран така, че да позволява маркиране само на един отговор, то респондентът и да иска, няма възможност за отбелязване на повече отговори. При хартиената анкета, въпреки направените указания, това няма как да се контролира и често респондентите маркират не един, а няколко отговора. = Ако изследователят желае, може да направи даден въпрос задължителен за отговаряне (респондентът не може да продължи по-напред в анкетата, ако не отговори на такъв въпрос). Това обикновено са ключовите, основните въпроси в една анкета. Ако, обаче, респондентът не желае да отговори на даден въпрос, той най-вероятно няма да го стори, независимо че въпросът е задължителен. В такъв случай е по-вероятно той да прекъсне попълването на анкетата като цяло. = Верифициране на отговора на респондента. Така например, ако сумата от отговорите към даден въпрос трябва да е 100%, то системата автоматично проверява спазването на това изискване и ако не е спазено, подканва респондента да направи съответни корекции. Ако респондентът трябва собственооръчно да запише своята възраст и в съответното поле е попълнил четирицифрено число, системата може да го подкани да въведе коректен отговор. = Определяне на точен времеви интервал, в рамките

на който онлайн анкетата да е отворена за попълване. Преди започването и след изтичането на указания период достъпът на респонденти до анкетата е невъзможен.

- **Адаптивност на анкетата.** Понякога се налага, на една група респонденти да се поставят едни въпроси, а на друга – съвсем различни. Примерно, ако на по-ранен етап се пита дали респондентът пие кафе, то на следващ етап за отношение към конкретни марки е логично да се питат само потребителите на кафе. При хартиената анкета това се реализира с указание от типа „Ако отговорът Ви на този въпрос е X, моля преминете към въпрос Y!“ . Онлайн анкетите позволяват автоматично пренасочване на респондента към един или друг набор от въпроси в зависимост от предходните му отговори. Така респондентът вижда само въпроси, които го касаят.
- **Защита на анкетата от достъп на лица, различни от целевата аудитория.** Една сравнително мека защита срещу достъп на странични лица е да се забрани индексването на анкетата от търсещите машини. Така тя не би се показвала в резултатите от търсенето, ако някой реализира търсене по ключови думи, налични и в самата анкета. Други варианти за защита включват: да се забрани повторното попълване на анкетата от потребители с един и същ IP адрес; да се забрани повторното попълване на анкетата от потребители, ползващи един и същ браузър от едно и също устройство; оторизиране на достъпа до анкетата с парола; позволяване на достъп до анкетата за потребители, идващи от точно определен сайт; генериране на уникален линк за всеки отделен респондент, позволяващ еднократно попълване и изпращане на анкетата; насочване на анкетата само към респонденти, притежаващи определени признаци – демографски, социални, поведенчески и др. (примерно чрез рекламния алгоритъм на Facebook). Допускането до участие само на представители на изучаваната съвкупност е важна предпоставка за набирането на данни, подходящи за постигане целите на изследването.
- **Видимост на анкетата в рамките на екрана.** Потребителите имат добре изграден навик за ползване скролбара на браузъра за вертикално придвижване в рамките на уебстраницата. Същото не може да бъде казано за използването на хоризонталния скролбар. Ако анкетата не се побира по своята широчина в рамките на екрана, твърде вероятно е, потребителят изобщо да не погледне какво е скритото съдържание вдясно на текущия екран. Препоръчително е, съдържанието на анкетната карта да

е така подбрано и разположено, че да не се налага, потребителят да използва хоризонталния скролбар.

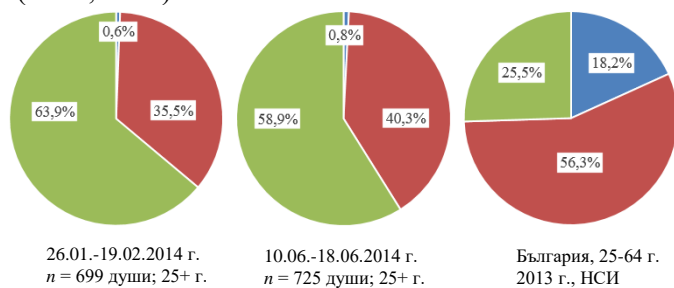
- **Предварително тестване на анкетата.** То помага на изследвателя да открие пропуски, логически и правописни грешки в анкетната карта, отстраняването на които да доведе до постигане по-висока степен на възвръщаемост при същинското изследване.
- **Възможност за попълване на анкетата на части.** Понякога се налага, респондентът да прекъсне попълването на анкетата. В такива случаи системата запомня вече попълнените отговори и позволява по-късно продължаване на анкетата от мястото, докдето респондентът е стигнал. Липсата на такава техническа възможност би означавала загуба на респонденти, които иначе са склонни да участват.
- **Легитимиране на изследвателя.** Възвръщаемостта на анкетата е по-висока, когато изследвателят представи себе си и целите на изследването. Колкото по-популярна и авторитетна е организацията, стояща зад изследването, толкова по-склонни са респондентите да участват в него.
- **Гарантиране на анонимност.** Използването на въпроси или технологии, които дават възможност за разкриване самоличността на респондентите, е предпоставка за влошаване степента на възвръщаемост на анкетата. Хората ценят правото си на анонимност. Все пак предлагането на достатъчно сериозен стимул е в състояние да неутрализира стремежа към анонимност и да накара респондентите с охота да споделят свои лични данни, например име, e-mail-адрес, телефон и др.
- **Стимулиране на респондентите.** Стимулирането може да има (а) материален характер (възможност за спечелване на предметна награда, директно заплащане за участие в анкетата, ваучер за ценова отстъпка при пазаруване или др.) или (б) нематериален характер (например, предлагане на обобщените резултати след края на изследването). И в двата случая върху поведението на респондентите се влияе чрез предоставяне или изтъкване на някаква полза, която ангажира респондентите с изследването и в крайна сметка повишава степента на възвръщаемост. Съществува мнение, че материалното стимулиране може да се окаже по-скоро вредно. Така например често се случва, респондентите да разгадаят целта на изследването и в стремежа да „заслужат“ обещаното възнаграждение, да се стремят към даване на „правилните“ отговори. Освен това възможно е да се развият т.нар. „професионални респонденти“ (ловци

на награди) – хора, които се стремят да участват в колкото може повече изследвания, само заради материалните стимули.

- **Съответствие на тема на анкетата с интересите на респондентите.** Колкото по-профилирана е извадката и колкото по-тясна е връзката между темата на изследването и интересите на респондентите, толкова по-ангажирани ще са те и по-висока ще е степента на възвръщаемост. Така например в две реални изследвания са постигнати степени на възвръщаемост 2,58% и 16,36%. В първия случай генералната съвкупност е голяма и разнородна, а темата на изследването не е от първостепенен интерес за респондентите. Във втория случай генералната съвкупност е сравнително малка и темата на изследването се припокрива с професионалните интереси на респондентите.

в) Някои респонденти са в по-голяма степен склонни да участват в онлайн анкети с конкретна тематика (*self-selection bias*). Изкривяване на резултатите може да възникне вследствие на това, че в онлайн анкети са по-склонни да участват респонденти, които (Йорданов, 2016, с. 433):

- **Имат по-висока образователна степен.** В периодите 26.01.–19.02.2014 г. и 10–18.06.2014 г. бяха проведени две вълни на онлайн анкета, целяща да оцени туристическия потенциал на региона Свищов–Белене–Ценово. За достигане до респондентите бяха използвани Facebook-групи, в които бе публикуван линк към анкетата (типичен пример за извадка на отзовалите се). На Фигура 1 може да се сравнят структурите на двете извадки и структурата на генералната съвкупност (по данни на НСИ за 2013 г.) по променливата „образование“ (НСИ, 2014):



Фигура 1. Структура по променливата „образование“ (основно, средно, висше)

И двете вълни са произвели извадки, които не съответстват на изучаваната съвкупност – респондентите с висше образование са свръх-представени за сметка на подпредставяне на групите с основно и средно

образование. До същия извод за свръхпредставяне на респонденти с висше образование стигат Фенър и колектив (Fenner et al., 2012).

- **Имат интерес към темата на изследването.** Хората с вече изградено отношение към темата на изследването е по-вероятно да вземат участие в него (Utts & Neckard, 2007). По-малка е вероятността, в една анкета да участват хора, които не са достатъчно информирани, не са заинтересовани или нямат твърдо мнение по въпросите, включени в анкетата.
- **Са от женски пол.** Жените са по-отзивчиви и по-често от мъжете вземат участие в онлайн анкети. За цитираното по-горе изследване на региона Свищов-Белене-Ценово бяха регистрирани следните дялове: Първа вълна ($n = 902$) – 33,1% мъже и 66,9% жени. Втора вълна ($n = 973$) – 31,3% мъже и 68,7% жени. По данни на НСИ за България към края на 2014 г. делът на мъжете е 48,6%, а на жените – 51,4% (НСИ, 2014).
- **Боравят по-интензивно с информационни и комуникационни технологии.** От една страна, за по-опитните потребители на ИКТ е по-лесно да разбират и попълват онлайн въпросници. От друга страна, по-вероятно е анкетата да бъде показана на екрана на потребител, който ползва интернет по-често и по-интензивно.

Когато е позната структурата на изучаваната съвкупност по важни за изследването променливи, отклоненията може да се коригират чрез претегляне на случаите в извадката. С тази процедура се цели постигане на съответствие между структурата на съвкупността и структурата на извадката по отношение на ключови променливи. Няма никаква гаранция, обаче, че подпредставените случаи, които следва да се умножат с коефициент по-голям от единица, са типични и представителни за съответната подсъвкупност.

В различни изследвания се докладват добри резултати за получаване на представителни данни на приемлива цена чрез извадки, излъчени с помощта на Facebook-реклама (Ramo & Prochaska, 2012). В подобни случаи следва да се има предвид още една възможност за изкривяване на резултатите – декларирането на фалшиви лични данни от страна на респондентите. При регистриране на Facebook-профил се изисква потребителят да отбележи своите пол, възраст, местоживее, образование, интереси и др. Ако един потребител попълни неверни данни за себе си, ще попадне в група, към която реално не принадлежи. Показването на реклама-покана за участие в анкета на лица, които не представляват интерес за изследването, е предпоставка за съществено изкривяване на получените резултати. По оценки на Facebook, фалшивите профили в социалната мрежа са между 5,5% и 11,2% от всички активни потребителски профили. Делът на фалшивите профили е по-нисък в развитите страни (примерно САЩ и

Великобритания) и по-висок в развиващите се страни (примерно Индия и Турция) (Facebook, Inc., 2014). Съгласно условията за ползване на Facebook, право на личен профил имат лица на 13 и повече години. По данни на MinorMonitor от април 2012 г. 38% от децата във Facebook са на възраст под 13 години (4% са на възраст до 6 г., 6% са от 7 до 8 години, 9% са от 9 до 10 години и 19% са от 11 до 12 години) (MinorMonitor, 2012). Следователно голяма е вероятността, изследвателят да попадне на респонденти, които във Facebook-профила си са обявили неверни данни.

Възможността за прецизно насочване на Facebook реклама към целева аудитория с конкретен профил е особено интересна при планирането на онлайн извадки. За да се провери доколко рекламният модул на Facebook достига до желаните респонденти, са проведени две онлайн анкети, целящи проучване на различни генерални съвкупности. Проведена е и трета анкета, реализирана сред посетителите на сайт в рамките на потребителската им сесия.

В следващата методологична част са описани подготовката и реализацията на онлайн анкетите, а в аналитичната част на изследването са докладвани получените резултати.

Методология на изследването

За постигане на дефинираните цели на изследването са проведени три отделни емпирични проучвания. Използвана е дескриптивна изследователска стратегия. Методът за набиране на данни е допитване, в частност онлайн анкета. Трите анкети са реализирани с помощта на платформата QuestionPro. При подготовка на анкетите са отчетени коментиранията по-горе препоръки за стимулиране интереса и ангажираността на респондентите и за повишаване на възвръщаемостта. При проектиране дизайн на въпросниците е отчетено как те ще се виждат на екрана на десктоп-машина и лаптопи, таблети и мобилни телефони. Това е важно съображение, тъй като значителна част от респондентите (понякога и преобладаваща) зареждат и попълват онлайн анкети през мобилните си телефони.

Анкета 1 – Проучване мнението на студенти, завършили Стопанска академия „Д. А. Ценов“

Целта на анкетата е да се получат оценките на вече завършилите студенти за качеството на обучение във висшето училище, както и да се получат данни за професионалната им реализация. Генералната съвкупност са лица, харесали страницата „Стопанска академия „Д. А. Ценов“. Към началото на кампанията (05.07.2018) групата е с големина 3 533 души. Информация за структурата на изучаваната съвкупност по променливите „пол“, „възраст“ и „местоживеене“ е използван раздел „Анализ“ към Facebook страницата на Стопанската академия. За достигане до целевите респонденти е използван рекламният алгоритъм на Facebook. От възможните варианти е избрана рекламна кампания от типа „Трафик“. При

него рекламата се оптимизира за максимизиране броя кликове върху рекламния банер. Рекламата се показва на лица, склонни да кликнат върху банера (преценено на база минало поведение). При този вариант на оптимизиране се получават възможно повече кликове на възможно най-ниска цена, каквито са обичайните цели на един рекламодадел. Не е ясно доколко този подход на насочване на рекламата е удачен при рекрутирането на респонденти. Това, че Facebook показва рекламата приоритетно на лица, склонни да взаимодействат с рекламния банер, е вече вид преднамерена намеса при съставяне на извадката и не всички потенциални респонденти се ползват с равен шанс за попадане в извадката. Кампанията стартира на 05.07.2018 г. и приключи на 09.07.2018 г. Не е подбиран конкретен график за показване на рекламата (показва се през целия период). Върху рекламния банер е изписан следният текст: „Искаме 7 минути от Вашето време. Мнението Ви е важно за нас. Включете се в анкета на Стопанска академия „Д. А. Ценов“, сега!“ . Изразходваният общ бюджет за кампанията е 21,76€. Целевата аудитория е дефинирана по следния начин: лица, харесали Facebook страницата на Стопанската академия; местоположение – България; пол – мъже и жени; възраст – 23-65+ години. Възрастовото ограничение е възприето в съответствие с целта на анкетата – да се проучи професионалната реализация на завършилите Стопанска академия „Д. А. Ценов“. Завършилите бакалавърска степен може да са най-малко на 23 години. Това логично ограничение на целевата аудитория обаче не бе съобразено с възрастовите интервали, които се докладват за феновете на Facebook страниците – 13-17 г., 18-24 г., 25-34 г., 45-54 г., 55-64 г., 65+ г. На практика в изучаваната Facebook страница има фенове на възраст от 13 до 22 г., които не са достигнати от рекламата и съответно не са имали шанс да попаднат в извадката. Поради тази причина в нея е подпредставен възрастовият интервал 18-24 г. При публикуване на рекламата се таргетира само избраната генерална съвкупност, но не може да се ограничат виделите рекламата на свой ред да споделят линк към анкетата. Тогавашанс да попълнят анкетата ще имат и странични лица, попадащи извън обсега на целевата генерална съвкупност. В края на кампанията рекламният банер е харесан от 60 души, а 19 са го споделили в своите профили. Не можем да определим колко потребители извън генералната съвкупност са видели рекламния банер, колко са кликнали върху него и колко са попълнили анкетата като резултат от тези споделяния. Големият дял на външни за генералната съвкупност респонденти е предпоставка, структурата на извадката да се различава значително от структурата на изучаваната съвкупност.

Анкета 2 – Проучване мнението на кандидат-студенти в магистърска степен

Втората анкета се състои от две последователни вълни. При първата вълна не се предлага награда за участие, а при втората вълна респондентите могат да спечелят едно от три устройства за съхраняване на инфор-

мация (флаш памет с вместимост 32GB). Целта е да се провери как целевите респонденти ще реагират на предложени стимул за участие в изследването. Впоследствие извадките от двете вълни се обединяват и се проверява доколко получената извадка съответства на генералната съвкупност по признаците „пол“, „възраст“ и „местоживеене“. Генералната съвкупност е дефинирана така – всички посетители на раздел „Магистър“ от сайта на Стопанска академия „Д. А. Ценов“. Идеята на това тематично насочване е да се проучи мнението по различни въпроси на кандидат-студентите в магистърска степен. Конкретната насоченост на анкетата е към установяване оценката на кандидат-студентите в магистърска степен за качеството на обучението в Стопанската академия, както и към получаване на обратна връзка за пригодността на сайта на висшето училище. Поканата за участие в анкетата се показва на всеки посетител на някоя от седемте страници от раздел „Магистър“. Три секунди след зареждане на съответната страница отдясно се появява поле, което закрива $\frac{1}{4}$ от екрана. Поканата за участие се показва само веднъж на всеки уникален посетител. Търсеният ефект от това ограничение е поканата да не се вижда повторно от вече участвалите в анкетата. Нежеланият ефект е, че само един контакт се реализира и посетители, които не са откликнали на първото показване на поканата. Хипотетично, при повторен или следващ контакт, вероятно част от тези посетители биха се включили в анкетата. От анализа са изключени посетителите, които са напуснали съответната страница в рамките на 5 секунди. Аргументът за това редуциране е следният: Поканата се показва 3 секунди след зареждане на страницата, а изчитането на текста на поканата не може да стане за по-малко от 2 секунди. На практика информирано решение да се включи в анкетата или да отклони поканата, може да вземе посетител, който се е задържал на страница от раздел „Магистър“ повече от 5 секунди. Информация за структурата на генералната съвкупност по променливите „пол“, „възраст“ и „местоживеене“ се получава от отчета на Google Analytics за профила на посетителите на сайта в раздел „Магистър“ в рамките на наблюдавания период. За прецизността на тази информация косвено може да се съди по дела на успешно класифицираните посетители. Така например по променливата „местоживеене“ успешно определени са всички посетители, докато по променливите „пол“ и „възраст“ – съответно 58,36% и 53,7%. **Първата вълна** продължава от 3 до 17 септември 2018 г. Поканата за участие съдържа следния текст: „Здравейте! Искаме 7 минути от Вашето време. Мнението Ви е важно за нас. Включете се в анкетата на Стопанска академия „Д. А. Ценов“, сега!“. **Втората вълна** започва на 18 септември и завършва на 17 октомври 2018 г. Текстът на поканата е: „Здравейте! Искаме 7 минути от Вашето време. Мнението Ви е важно за нас. Участвайте в тази анкета на Стопанската академия и спечелете **USB Flash памет 32GB!**“. Преди да се пристъпи към обработка, от събраните данни са изчистени случаите с повтарящи се IP-адреси, зад които е ясно, че стои един и същ потребител.

Анкета 3 – Проучване мотивацията за кандидатстване в икономически висши училища

Целта на тази анкета е да се установят съображенията на българите при избора на висше училище и да се определят нагласите към икономическото образование. Генералната съвкупност е дефинирана като населението на България на възраст от 13 до 54 години. За достигане до респондентите е използвана Facebook реклама. Теоретично погледнато, генералната съвкупност няма как да бъде проучена по надежден начин, ако до респондентите се достига единствено с Facebook реклама. Причината е, че не всички хора на възраст от 13 до 54 години в България ползват Facebook (единиците от съвкупността нямат равен шанс да попаднат в извадката). Въпреки това, решихме да проверим доколко структурата на извадката, излъчена чрез Facebook, съответства на структурата на генералната съвкупност по ключовите променливи „пол“, „възраст“ и „местоживеене“. В крайна сметка, ако броят на отказите е голям, така или иначе извадката, първоначално замислената като случайна, би се превърнала в извадка на отзовалите се. Една неслучайна извадка също може да е представителна, но тази възможност ще се потвърди или отхвърли само ако е известна структурата на генералната съвкупност по променливите, представляващи интерес за изследването. За разпространението на линка към анкетата е избрана рекламна кампания от типа „Трафик“. Както вече бе казано, тя е насочена към осигуряване на повече кликове върху линка, препрещащ към платформата с онлайн анкетата. Въз основа на минало поведение, рекламният алгоритъм на Facebook оценява кои лица са по-склонни да кликнат върху рекламния банер. Именно на тези лица се показва рекламата. За параметрите на генералната съвкупност по променливите „пол“, „възраст“ и „местоживеене“ ползваме данни от „Статистически справочник 2018“ – официално издание на Националния статистически институт на Република България (НСИ, 2018, с. 21 и 29). Данните от справочника са актуални към 31.12.2017 г.

За оценка степента на съответствие между извадка и генерална съвкупност са използвани два инструмента:

- **χ^2 тест за една извадка** – непараметрична статистическа процедура, с която може да се сравняват относителни дялове в извадката с някакви хипотетични стойности или със стойности, валидни за относителните дялове в генералната съвкупност (Кръстевич, Смокова, 2010, с. 296). Тестът е подходящ за проверка на съответствието между две съвкупности въз основа на променливи, имащи две или повече значения (изследваните единици се разпределят в две или повече групи). Процедурата по проверка на съответствието започва с формулиране на две хипотези – нулева (H_0) и алтернативна (H_1). Нулевата хипотеза съдържа твърдение, че разликата между параметъра на генералната съвкупност и неговата оценка от извадката е случайна, не-

закономерна, статистически незначима. Алтернативната хипотеза съдържа противоположното твърдение – разликата между параметъра на генералната съвкупност и неговата извадкова оценка не е случайна, а закономерна и статистически значима. Нулевата хипотеза се отхвърля, ако наблюдаваното равнище на значимост ($\alpha_{\text{емпирично}}$) е по-малко от предварително дефинираното критично равнище на значимост ($\alpha_{\text{критично}}$). Обикновено в социално-икономическите изследвания се работи с критично равнище на значимост $\alpha=0,05$, но ако е необходима по-висока степен на сигурност на заключенията, се използват и по-ниски стойности за α (Манов, 2000, с. 129). В настоящото изследване χ^2 тестът се използва за сравняване на извадка и генерална съвкупност по променливите „пол“, „възраст“ и „местоживееене“. Прилагането му е коректно, когато теоретичните (очакваните) честоти във всяка група са не по-малки от 5. Ако това условие не е изпълнено, може да се премине към округняване на групите (обединяване на отделни значения на изследваната променлива – примерно, интервалите „под 17 г.“ и „от 18 до 24 г.“ да се обединят в един възрастов интервал, интервалите „от 25 до 34 г.“ и „от 35 до 44 г.“ – във втори възрастов интервал и т.н.). В настоящото изследване такова округняване на групите е направено за променливата „местоживееене“ – оригинално събраните данни са по области, но впоследствие са обединени по райони – североизточен, северен централен, северозападен, югоизточен, южен централен, югозападен и чужбина. Следва да се има предвид, че χ^2 тестът е приложим при случайни извадки. Строго погледнато, извадките и при трите анкети не отговарят на изискването за случайност. По тази причина ще използваме и представения по-долу интегриран коефициент на структурните различия на Кирил Гатев, изводите от прилагането на който ще сравним със заключенията, получени на база на χ^2 теста.

- **Интегриран коефициент на структурни различия на Гатев** (Гатев, 1989, с. 258-260) – използва се за сравняване степента на съответствие между структурите на различни съвкупности. В случая коефициентът се използва за измерване съответствието на извадка и генерална съвкупност по променливите пол, възраст и местоживееене. Коефициентът се изчислява по следната формула:

$$K_S = \sqrt{\frac{\sum_{i=1}^k (v_{i2} - v_{i1})^2}{\sum_{i=1}^k (v_{i1})^2 + \sum_{i=1}^k (v_{i2})^2}}, \text{ където:} \quad (1)$$

K_S – интегриран коефициент на структурни различия на Гатев;
 v_{i1} – дял на група i в извадката;
 v_{i2} – дял на група i в генералната съвкупност;
 k – брой на значенията на изследваната променлива;
 $1, 2$ – индекси за обозначаване на извадката и на генералната съвкупност.

За интерпретиране стойностите на коефициента се използва следната скала (Шопова, 2018, с. 221):

Стойност на коефициента	Сила на структурните различия
От 0,00 до 0,03	Липсват структурни различия
Над 0,03 до 0,07	Незначителни структурни различия
Над 0,07 до 0,15	Слаби структурни различия
Над 0,15 до 0,30	Умерени структурни различия
Над 0,30 до 0,50	Значителни структурни различия
Над 0,50 до 0,70	Силни структурни различия
Над 0,70 до 0,90	Много силни структурни различия
Над 0,90	Пълни структурни различия

Три основни достойнства се сочат за интегрирания коефициент на Гатев, а именно (Шопова, 2018, с. 64):

- (1) Коефициентът е нормиран в теоретични граници от 0 до 1, като ниските (високите) стойности свидетелстват за наличие на малки (големи) разлики между структурите на сравняваните съвкупности.
- (2) При изчисляването му се отчита както разликата между относителните дялове в двете съвкупности, така и фактическото равнище на самите относителни дялове. Така например, между 80% и 90%, от една страна, и между 10% и 20%, от друга, разликата е 10 процентни пункта, но във втория случай тя е много по-значима от първия.
- (3) Коефициентът е чувствителен спрямо структурните различия – при едно и също абсолютно изменение по-голямо влияние върху стойността на коефициента има разликата между по-малките относителни дялове, отколкото между по-големите.

При работа с интегрирания коефициент на Гатев следва да се има предвид, че не е коректно сравняването на стойности, изчислени за структури с различен брой елементи (броят на значенията на отделните променливи, респ. броят на групите, на които се подразделят съвкупностите). Така например, не е редно да се сравняват стойности на коефициента, изчислени на база променливите „пол“ и „възраст“, ако едната променлива има 2, а другата – 7 значения. Сравнения са възможни, но ако предварително броят на значенията на двете променливи се уеднакви – примерно, като възрастовите интервали се округлят и се приведат към два на брой.

Извадките и в трите анкети са неслучайни, от типа на „отзовалите се“. Това не дава основание за прилагане на статистически инструментариум за определяне обема на извадката и оценка на нейната представителност. По презумпция, за да е представителна, една извадка трябва да е излъчена на случаен принцип. За целта е необходим изчерпателен списък на всички единици от изучаваната генерална съвкупност (основа на извадката). При изследване 3 (национална анкета чрез Facebook реклама) е невъзможно осигуряването на такава основа. При изследване 2 (анкетирание посетителите на сайт) е възможно, всички посетители за определен период да се разглеждат като основа за излъчване на извадка. Ако, примерно, на всеки десети посетител на сайта се показва покана за участие в анкетата, би се получила систематична случайна извадка. При използване на голяма стъпка неголям брой посетители на сайта и сравнително ниска степен на възвръщаемост на онлайн анкетите, съществува риск от получаване на твърде малка извадка. По тази причина е взето решение, анкетата да се показва на всеки посетител, което е в съответствие с принципа за формиране на систематична извадка, но със стъпка, равна на единица. При изследване 1 (допитване до феновете на Facebook страница) е възможно да се състави основа за излъчване на случайна извадка (списък на всички, харесали страницата). Съществуват технологични възможности, поканата за участие в анкетата да се изпрати само до избраните по случаен принцип фенове на Facebook страницата. В резултат на отказите за участие, макар и замислена като случайна, получената извадка би била от типа на доброволно отзовалите се. Поради тази причина е предпочетено, покана за участие да се показва на всички фенове на страницата, като се използва рекламният алгоритъм на Facebook.

Анализ на резултатите от изследването

В тази секция са анализирани и коментирани резултатите, получени при провеждането на анкети 1, 2 и 3. Направени са сравнения в две посоки:

1. Сравняване на данните за активност и възвръщаемост от отделните извадки (ако в рамките на съответната анкета са реализирани повече от една извадки). Целта е да се установи дали и доколко обявената награда влияе върху активността на респондентите.
2. Сравняване структурата на генералната съвкупност със структурата на извадката (или обединените извадки, ако в рамките на съответната анкета са реализирани повече от една извадки). Целта е да се провери доколко получената извадка се доближава до параметрите на генералната съвкупност по променливите пол, възраст и местоживееене.

Анализ на резултатите от анкета 1 – Допитване до феновете на Facebook страница

Данни за активността на респондентите:

Период на провеждане	Импресии	Достигнати лица	Кликове в/у банера	Отворени анкети	Започнати анкети	Завършени анкети	Започнати, но незавършени анкети
05-09.07.2018	18 759	2 437	266	347	260	182	68

Изразходена сума	Средна продължителност на попълване	% отзовали се ¹	Степен на възвръщаемост ²	Степен на завършване ³	Степен на отпадане ⁴
21,76€	4 минути	14,24%	7,47%	70,00%	26,15%

¹ Процент отзовали се на поканата = (Брой респонденти, заредили началния екран на анкетата / Брой адресати, видели поканата за участие)*100.

Високите стойности показват, че поканата привлича вниманието на адресатите, провокира интереса им към изследването, кара ги да кликнат върху линка и да заредят началния екран на анкетата.

² Степен на възвръщаемост = (Брой респонденти, попълнили анкетата / Брой адресати, видели поканата за участие)*100.

Измерва каква част от виделите поканата са завършили анкетата. Показва каква част от планираната извадка е успешно реализирана и, евентуално, каква част предстои да се осигури. Свидетелства за способността на анкетата да привлече, но и да задържи интереса на респондента до самия ѝ край.

³ Степен на завършване на анкетата = (Брой респонденти, попълнили анкетата / Брой респонденти, стартирали попълването на анкетата)*100. Показва доколко анкетата е в състояние да задържи интереса на респондента до самия ѝ край.

⁴ Степен на отпадане = (Брой респонденти, стартирали, но незавършили анкетата / Брой респонденти, стартирали попълването на анкетата)*100.

Показва доколко анкетата не успява да задържи интереса на респондента до самия ѝ край.

Границите на изменение на тези показатели са от 0 до 1 (когато са изчислени като коефициент) или от 0 до 100 (ако са изчислени като процент).

Отгук могат да се изчислят и други показатели, полезни при обясняване на резултатите (защо са такива, каквито са) и тяхната себестойност (колко са стрували на изследователя в парично изражение): Цена за достигане до един потенциален респондент = 21,76€ / 2 476 = 0,0088€; Цена за получаване на една попълнена анкета = 21,76€ / 182 = 0,1196€; Средна честота на контакт с банера = 18 759 / 2 437 = 7,7.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност по променливата „пол“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
Мъж	26,39%	3533	27,32%	183
Жена	73,13%		72,68%	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
0,061	1	0,805	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,009

Заключение на база χ^2 теста: $\alpha_{\text{емпирично}} > \alpha_{\text{критично}}$ ($0,805 > 0,05$) – следва да се приеме нулевата хипотеза – различията между разпределенията в извадката и генералната съвкупност по променливата „пол“ са несъществени и не са статистически значими.

Заключение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката липсват структурни различия различия по критерия „пол“.

Оценяване съответствието между структура на извадката и структура на генералната съвкупност по променливата „възраст“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
От 18 до 24 г.	27,73	3533	21,86	183
От 25 до 34 г.	34,45		48,63	
От 35 до 44 г.	16,18		19,13	
От 45 до 54 г.	11,45		7,1	
От 55 до 64 г.	7,19		3,28	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
20,066	4	0,000	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,223

Заключение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,000 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в извадката и генералната съвкупност по променливата „възраст“ са налице съществени, статистически значими различия.

Заключение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице умерени структурни различия по критерия „възраст“.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност по променливата „местоживееене“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
Североизточен район	7,32	3533	8,2	183
Северен центр. район	30,19		25,14	
Северозападен район	16,97		22,95	
Югоизточен район	7,79		7,65	
Южен центр. район	5,43		6,56	
Югозападен район	30,90		28,42	
Чужбина	1,39		1,09	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
7,054	6	0,316	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,125

Заключение на база χ^2 теста: $\alpha_{\text{емпирично}} > \alpha_{\text{критично}}$ ($0,316 > 0,05$) – следва да се приеме нулевата хипотеза – различията между разпределенията в извадката и генералната съвкупност по променливата „пол“ са не-съществени и не са статистически значими.

Заключение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице слаби структурни различия по критерия „местоживееене“.

Някои възможни обяснения за различията между извадката и генералната съвкупност:

- В анкетата се пита за местоживееене през последната година, което може да е различно от информацията за местоживееене, попълнена от респондента във Facebook профила.
- Първият докладван от Facebook възрастов интервал е от 18 до 24 г. Анкетата бе насочена към респонденти на 23 и повече години (в съответствие с целта да се проучи професионалната реализация на възпитаниците на Стопанска академия „Д. А. Ценов“). При това положение обаче се стига до сравняване между (а) дела в генералната съвкупност на лицата на възраст от 18 до 24 години (27,73%) и (б) дела в извадката на респондентите на възраст от 23 до 24 години (21,68%). Тази неточност при проектирането на извадката води до изкривено представяне и на другите възрастови интервали.

Анализ на резултатите от анкета 2 – Допитване до посетителите на сайт

Големината на генералната съвкупност при анализа по отделни променливи е различна, защото алгоритъмът на Google Analytics не успява да определи всички респонденти по пол, възраст и местоположение. При променливата „пол“ успешно класифицирани са 4230 посетители, при променливата „възраст“ – 3892 посетители, а при променливата „местоживее-не“ – 7248 посетители на сайта. Въпреки че реалният брой на посетителите за периода от 3 септември до 17 октомври 2018 г. е по-голям, за генерална съвкупност приемаме само онези единици, които са успешно класифицирани. В противен случай не е възможно да се структурира генералната съвкупност по променливите, въз основа на които се проверява за съответствие с получената извадка.

Сравняване активността на респондентите от двете вълни:

Извадка 1 (без награда)

Период на провеждане	Респонденти, прочели поканата	Брой отворени анкети	Брой стартирани анкети	Брой завършени анкети	Брой започнати, но незавършени анкети
03-17.09.2018	3 748	375	233	161	72

Средна продължителност на попълване	% отзовали се	Степен на възвръщаемост	Степен на завършване	Степен на отпадане
7 минути	10,01%	4,3%	69,1%	30,9%

Извадка 2 (с награда)

Период на провеждане	Респонденти, прочели поканата	Брой отворени анкети	Брой стартирани анкети	Брой завършени анкети	Брой започнати, но незавършени анкети
18.09-17.10.2018	4 168	390	259	179	80

Средна продължителност на попълване	% отзовали се	Степен на възвръщаемост	Степен на завършване	Степен на отпадане
6 минути	9,36%	4,3%	69,11%	30,89%

Първата вълна от анкетата е с продължителност две седмици – от 3 до 17 септември. Първоначалният замисъл бе, втората вълна да има същата продължителност. Като се има предвид, че при втората вълна е добавена

на награда, не сме допускали, че отговорите ще се съберат толкова трудно. За да постигнем приблизително същия брой попълнени анкети като при първата вълна, се наложи да оставим анкетата цял месец – от 18 септември до 17 октомври. Основната причина за бавното набиране на респонденти при втората вълна е следната: в секция „Магистър“ на сайта на Стопанска академия влизат основно две категории – кандидат-студенти и вече записали се студенти – едните търсят информация за кандидатстване, другите търсят учебния график за предстоящата учебна година; към края на месец септември и началото на октомври и двете групи посетители на сайта вече са наясно с интересувашата ги информация и посещенията на страниците в този раздел чувствително намаляват. Оттук може да се направи заключение, че при провеждане на тестове с две или повече извадки е редно, те да се реализират паралелно, а не последователно във времето. Редица обстоятелства може да са различни за отделните периоди и това да се отрази върху резултатите, констатирани при различните вълни на допитване.

От сравнението на показателите „процент отзовали се“, „степен на възвръщаемост“, „степен на завършване“ и „степен на отпадане“ за двете вълни става ясно, че предлагането на стимул за участниците не повишава тяхната активност. Възможно обяснение – наградите се възприемат като твърде малки и не провокират интерес у целевите респонденти. Остава хипотезата, че включването на по-голяма награда би действало в по-голяма степен за мобилизиране на респондентите. Тази хипотеза бе подложена на проверка с провеждането на анкета 3.

В допълнение, видно е, че попълването на анкетата при първата вълна (без награда) е отнело средно по 7 минути, а при втората вълна (с награда) – средно по 6 минути. Възможно обяснение – някои респонденти са участвали във втората вълна по-скоро заради наградата и не са се задълбочавали в четенето и разбирането на поставените въпроси. Ако хипотезата е вярна, то добавянето на награда води до влошаване качеството на данните от анкетата. В настоящата разработка тази хипотеза не е подложена на проверка.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност въз основа на променливата „пол“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
Мъж	24,85%	4230	19,24%	343
Жена	75,15%		80,76%	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
6,065	1	0,014	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,069

Заключение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,014 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в

извадката и генералната съвкупност по променливата „пол“ са налице съществени, статистически значими различия.

Заклучение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице **незначителни** структурни различия по критерия „пол“.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност въз основа на променливата „възраст“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
От 18 до 24 г.	14,77	3892	29,45	343
От 25 до 34 г.	45,68		34,40	
От 35 до 44 г.	26,72		27,41	
От 45 до 54 г.	12,82		8,75	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
64,063	3	0,000	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,254

Заклучение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,000 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в извадката и генералната съвкупност по променливата „възраст“ са налице съществени, статистически значими различия.

Заклучение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице **умерени** структурни различия по критерия „възраст“.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност въз основа на променливата „местоживееене“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
Североизточен район	5,28	7248	8,75	343
Северен центр. район	15,29		20,99	
Северозападен район	11,49		29,15	
Югоизточен район	7,64		10,20	
Южен центр. район	5,71		6,41	
Югозападен район	50,33		23,03	
Чужбина	4,25		1,46	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
168,825	6	0,000	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,469

Заключение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,000 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в извадката и генералната съвкупност по променливата „местоживеене“ са налице съществени, статистически значими различия.

Заключение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице **значителни** структурни различия по критерия „местоживеене“.

Някои възможни обяснения за различията между извадката и генералната съвкупност:

- В анкетата се пита за местоживеене през последната година, а Google Analytics засича мястото, на което се намира респондентът в момента на попълване на анкетата.
- Не е ясна технологията и прецизността, с която Google Analytics класифицира респондентите по пол и възраст. Повод за разсъждаване в тази посока е фактът, че от 7 248 уникални посетители успешно (според Google Analytics) класифицирани са 4 230 респонденти по пол и 3 892 респонденти по възраст.

Анализ на резултатите от анкета 3

Сравняване активността на респондентите от двете паралелни извадки:

Извадка 1 (без награда)

Период на провеждане	Импресии	Достигнати лица	Кликове в/у банера	Отворени анкети	Започнати анкети	Завършени анкети	Започнати, но незавършени анкети
18-28.10.2018	99 810	45 679	1 742	1 837	761	241	520

Изразходена сума	Средна продължителност на попълване	% отзовали се ¹	Степен на възвръщаемост ²	Степен на завършване ³	Степен на отпадане ⁴
57,36€	4 минути	4,02%	0,53%	31,67%	68,33%

Допълнителни показатели: Цена за достигане до един потенциален респондент = $57,36\text{€} / 45\,679 = 0,0013\text{€}$; Цена за получаване на една попълнена анкета = $57,36\text{€} / 241 = 0,238\text{€}$; Средна честота на контакт с банера = $99\,810 / 45\,679 = 2,19$.

Извадка 2 (с награда)

Период на провеждане	Импресии	Достигнати лица	Кликове в/у банера	Отворени анкети	Започнати анкети	Завършени анкети	Започнати, но незавършени анкети
18-28.10.2018	97 486	45 928	1 752	1 864	797	297	500

Изразходена сума	Средна продължителност на попълване	% отзова-ли се ¹	Степен на възвръщаемост ²	Степен на завършване ³	Степен на отпадане ⁴
57,39€	5 минути	4,06%	0,65%	37,26%	62,74%

Допълнителни показатели: Цена за достигане до един потенциален респондент = $57,39€ / 45\,928 = 0,0013€$; Цена за получаване на една попълнена анкета = $57,39€ / 297 = 0,1932€$; Средна честота на контакт с банера = $97\,486 / 45\,928 = 2,12$.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност въз основа на променливата „пол“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
Мъж	51,37%	3 723 994	19,8%	491
Жена	48,63%		80,2%	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
195,544	1	0,000	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,411

Заключение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,000 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в извадката и генералната съвкупност по променливата „пол“ са налице съществени, статистически значими различия.

Заключение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице **значителни** структурни различия по критерия „пол“.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност въз основа на променливата „възраст“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
От 13 до 17 г.	8,56	3 723 994	6,92	490*
От 18 до 24 г.	12,09		21,38	
От 25 до 34 г.	24,84		24,85	
От 35 до 44 г.	28,16		29,33	
От 45 до 54 г.	26,35		17,52	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
51,403	4	0,000	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,191

* От получените общо 538 отговора са изключени респондентите на 55 и повече години, тъй като попадат извън обхвата на целевата съвкупност на конкретното допитване.

Заклучение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,000 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в извадката и генералната съвкупност по променливата „възраст“ са налице съществени, статистически значими различия.

Заклучение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице **умерени** структурни различия по критерия „възраст“.

Оценяване съответствието между структурата на извадката и структурата на генералната съвкупност въз основа на променливата „местоживееене“

Значения на променливата	Дял в генералната съвкупност	Големина на генералната съвкупност (N)	Дял в извадката	Големина на извадката (n)
Североизточен район	13,24	3 723 994	9,69	485
Северен центр. район	11,28		23,09	
Северозападен район	10,72		22,06	
Югоизточен район	14,75		14,43	
Южен центр. район	20,11		9,90	
Югозападен район	29,91		20,82	

χ^2	Степени на свобода	$\alpha_{\text{емпирично}}$	$\alpha_{\text{критично}}$	H ₀	H ₁	K _S
161,401	5	0,000	0,05	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0,351

Заклучение на база χ^2 теста: $\alpha_{\text{емпирично}} < \alpha_{\text{критично}}$ ($0,000 < 0,05$) – следва да се приеме алтернативната хипотеза – между разпределенията в извадката и генералната съвкупност по променливата „местоживееене“ са налице съществени, статистически значими различия.

Заклучение на база интегралния коефициент на Гатев: Между генералната съвкупност и извадката са налице **значителни** структурни различия по критерия „местоживееене“.

Някои възможни обяснения за различията между извадката и генералната съвкупност:

- В анкетата са свръхпредставени респонденти от Северен централен и Северозападен район. Това са районите, от които традиционно Стопанската академия привлича най-много студенти. В случая колкото по-близо до Свищов живеят респондентите, толкова по-склонни са да участват в анкетата. Съответно това води до подпредставяне на респондентите от по-отдалечените региони.
- Извадката представлява твърде малък процент от големината на генералната съвкупност. Една по-продължителна кампания и по-голяма извадка могат да осигурят значително по-добри данни.

В случая на анкета 3 допълнителните взаимодействия с рекламния банер, като харесване, коментар и споделяне, са желан резултат и допринасят за по-голямата видимост сред потребителите на Facebook, съответно осигуряват по-широкия обхват на кампанията. За периода от 18 до 28 октомври 2018 г. рекламният банер без обявена награда е харесан 91 пъти, има един коментар и е споделен 14 пъти. Банерът с обявена награда е харесан 85 пъти и е споделен 19 пъти. Размерът на вирусния ефект е измерен чрез съпоставяне данните за броя на кликовете върху рекламния банер (от Facebook) и данните за броя зареждания на началния екран на анкетата (от bit.ly).

Кликове върху рекламния банер (по данни от Facebook)		Брой зареждания на началния екран на анкетата (по данни от bit.ly)	
С награда	Без награда	С награда	Без награда
1 752	1 742	1 864	1 837

Допълнителните взаимодействия с рекламния банер са увеличили видимостта му (по вирусен път) и са осигурили 207 допълнителни зареждания на анкетата.

При констатиране на значителни разлики между структурата на извадката и структурата на изучаваната съвкупност изследователят може да:

- Претегли данните от анкетата, за да се подсили присъствието на подпредставените единици и да се осигури необходимата структура на извадката по ключовите променливи;
- Продължи рекламната кампания във Facebook, но вече насочена само към подпредставените единици. Например насочване на рекламата само към мъже на възраст между 35 и 55 г., ако тази група не е представена в извадката с необходимия относителен дял. Подобен подход е лесно приложим при съставянето на квотна извадка, която е и една от най-често прилаганите неслучайни извадки (Желев, 1995, с. 187).

С уговорката, че проведените анкети са само частни случаи и не дават основание за извеждане на категорични и общовалидни заключения, в настоящото изследване се констатира следните по-важни резултати:

- Онлайн анкетите могат да са източник за бързо и евтино набиране на данни, които в много висока степен да отразяват структурата на генералната съвкупност.
- По-склонни да участват в онлайн анкетите са респонденти, живеещи по-близо до Стопанска академия „Д. А. Ценов“. И обратно – лица, които живеят по-далеч и са по-незапознати със Свищов и Стопанската академия, са по-слабо представени в получените извадки. Това е една от причините за чувствителните несъответствия между извадка и съвкупност по променливата „местоживееене“ от анкети 2 и 3.
- За пореден път се констатира, че в онлайн анкети по-склонни към участие са жените.
- Колкото по-близка е темата на изследването до целевите респонденти, толкова по-висока е възвръщаемостта на анкетата и по-силно е съответствието между извадка и изучавана съвкупност. Най-лоши в това отношение са резултатите при анкета 3, а най-добри са при анкета 1.
- Онлайн анкетите, проведени чрез реклама във Facebook, осигуряват сравнително добро съответствие между извадка и генерална съвкупност, когато обект на изследване са неголеми затворени общности (в случая – феновете на Facebook страница);
- Допитването до посетителите на сайт може да осигури извадка, която сравнително добре описва изучаваната съвкупност (в случая – посетителите на сайта в рамките на определен период).
- Тестваните стимули не подобряват активността на респондентите в степен, оправдаваща разходите за закупуване и доставяне на наградите.
- Обявяването на награда за участие в онлайн анкетите смъква чувствителността на респондентите по отношение на аноним-

ността им и те с готовност обявяват свои лични данни – име, телефонен номер или e-mail адрес. Във варианта на анкета 2 с обявена награда 67,04% от респондентите са предоставили свои лични данни. При анкета 3 стойността на този показател е 88,22%. За значително по-високия процент при анкета 3 очевидно има значение големината на обявената награда (флаш памет 32 GB срещу външен хард диск 1 TB).

- Достигането до сравнително малки съвкупности чрез Facebook реклама дава добри резултати – и като структурно съответствие между извадка и съвкупност, и като цена за получаване на един отговор. Малката съвкупност дава възможност и за постигане на по-голяма честота на контакт с рекламния банер. При анкета 1 се констатира много висока степен на съответствие между извадка и съвкупност по наблюдаваните три признака. Цената за една попълнена анкета е 0,1196€. Средният брой контакти е 7,7. Големината на извадката представлява 5,18% от големината на съвкупността. Степента на възвръщаемост е 7,43%. За сравнение – при анкета 3 различията между извадка и изучавана съвкупност са от умерени до значителни. Обяснение за този резултат може да се потърсят в неблагоприятните стойности на показателите: Цена за една попълнена анкета = 0,2133€; Среден брой контакти = 2,15; Големината на извадката спрямо генералната съвкупност = 0,013%; Средна степен на възвръщаемост = 0,59%. Възможно е, възприетият подход за достигане до целевите респонденти да даде добри резултати и при анкета 3, ако рекламната кампания продължи по-дълго време. Това забавяне ще е съпроводено със значително повишаване цената за получаване на един отговор, което може да постави под въпрос ефективността на допитването.
- В случаите на структурно съответствие между извадка и генерална съвкупност, констатирано въз основа на χ^2 критерия, коефициентът на Гатев заема стойности в интервала от 0,00 до 0,150 (т.е. диапазоните на липсващи структурни различия, незначителни структурни различия и слаби структурни различия). В този смисъл за степента на припокриване между извадката и изучаваната съвкупност би могло да се състави заключение и само въз основа на интегралния коефициент на Гатев.
- Анализ на отпаданията – показва кои са критичните въпроси в една анкета, на които отпадат голям брой респонденти. Общото заключение е, че откритите въпроси и тежките въпроси от матричен тип отказват не малка част от респондентите, започнали попълването на анкетата. Понякога обаче не е целесъобразно, открит въпрос да бъде заместен със закрит (примерно,

когато се търси автентичното мнение на респондентите, а не да им се подсказват насоки за мислене и оценяване). Впечатление прави и сравнително големият брой респонденти, които отпадат на последния въпрос, непосредствено преди натискане на бутона „Изпрати“. Възможно решение на този проблем е към последния въпрос да се направи кратко уточнение „За да изпратите анкетата, моля натиснете бутона „Изпрати“!“

Заклучение

В хода на изследването възникнаха въпроси, на които може да се потърсят отговори в евентуални бъдещи изследвания, а именно:

- Дали обявяването на награда за участие се отразява върху степента на съответствие между генерална съвкупност и извадка (т.е. върху качеството на данните)?
- Включването на награда провокира ли участието на лица, които в противен случай не биха се отзовали? Ако това са лица със сходни признаци, отсъствието им от извадката би влошило нейната представителност. И обратно, активирането на тези лица би довело до по-силно съответствие на извадката с генералната съвкупност.
- Данните от анкети 2 и 3 показват, че при обявена награда за участие средната продължителност на попълване е по-малка. Изглежда логичен въпросът, дали някои респонденти са участвали само заради възможността да спечелят награда, без да се задълбочават в изчитането и разбирането на въпросите от анкетата.
- В анкета 2 са предложени три награди – флаш памет с капацитет 32GB и единична пазарна цена около 10 лв. към момента на провеждане на анкетата. Приемаме такава награда като малка. В случая тази малка награда се оказва недостатъчен стимул за повишаване активността на респондентите. В анкета 3 са предложени 2 външни твърди диска, всеки с капацитет 1TB и единична цена около 120 лв. към момента на провеждане на допитването. Определяме такава награда като сравнително голяма. Ефектът от тази награда е положителен – 297 попълнени анкети с обявена награда срещу 241 попълнени анкети без обявена награда. Този ефект обаче е недостатъчен, за да оправдае разходите за закупуване и доставяне на наградите до печелившите респонденти. Възниква въпросът дали обявяването на сравнително малка, но сигурна награда за всеки респондент би повишила интереса към участие.

- В настоящото изследване е тествана само реклама от типа „Трафик“. При нея алгоритъмът на Facebook оценява вероятността, един респондент от целевата аудитория да кликне върху показания му рекламен банер и доставя рекламата на най-подходящите лица. Подобна селекция противоречи на принципа за излъчване на случайна извадка. Това е една от възможните причини за констатираните несъответствия между извадка и генерална съвкупност. Измежду наличните към момента други типове реклама най-подходящ изглежда вариантът „Достигане“, при който Facebook оптимизира кампанията за достигане до максимален брой хора, без оглед на склонността им да реагират на показан им рекламен банер. Полезно би било да се установи дали рекламна кампания от типа „Достигане“ би осигурил по-добри резултати в сравнение с кампанията от типа „Трафик“.
- Както бе споменато, за да е представителна извадката, респондентите трябва да са подбрани по случаен начин и броят им да е достатъчно голям. В настоящото изследване акцентът е поставен върху структурните сравнения между извадка и съвкупност, но не е изследван въпросът за големината на извадката. Възможен подход за определяне големината на една неслучайна извадка е, като се следи дали с нарастване броя на респондентите се стабилизират стойностите на важните за изследването величини – средни стойности или относителни дялове.

След приключването на това изследване остават много открити въпроси. На други опитахме да дадем отговор, но с ясното съзнание, че тези отговори съвсем не са еднозначни. Те са такива, каквито са тук и сега. Възможно е, в друг контекст резултатите от подобно изследване да са коренно различни. Полезното, което остава, е предложеният подход за валидиране на неслучайна извадка с помощта на χ^2 критерия и интегралния коефициент на Гатев. Полезно е да се знае, че посредством рекламния алгоритъм на Facebook могат да се реализират бързи допитвания на сравнително ниска цена, които са в състояние да осигурят представителни или близки до представителни резултати, но най-вече, когато целевата съвкупност е сравнително малка и/или добре профилирана. В такива случаи възможността за прецизно насочване на рекламата, а с нея и поканата за участие в анкетата, може да осигури надеждни откъм представителност данни. Стремещът следва да бъде към формиране на случайна извадка, защото по-добре е по-малка случайна извадка, отколкото по-голяма неслучайна. Ако не е възможно съставяне на случайна извадка, за предпочитане са такива неслучайни извадки, които наподобяват случайни (примерно квотна извадка) или в рамките на неслучайните извадки да се прилагат процедури, характерни за случайните (Желев, 1995, с. 224).

Използвани източници

- Гатев, К., Спасов, А. & Радилов, Д. (1989). *Обща теория на статистиката и икономическа статистика*. Издателство „Наука и изкуство“, София.
- Желев, С. (1995). *Маркетингови изследвания*. Университетско издателство „Стопанство“, София.
- Йорданов, Р. (2016) Представителни ли са онлайн анкетите Научна конференция с международно участие Съвременни управленски практики IX „Управленска наука, икономика и бизнес практики – съвременни ракурси и предизвикателства“, Бургас, 17-18 юни, стр. 428-436.
- Кръстевич, Т. & Смокова, М., SAS и SPSS за начинаещи: подготовка, визуализация и анализ на данни. Свищов, 2010, стр. 296.
- Манов, А. (2000). *Статистика със SPSS*. Издателство „РИКОЛ - Б“, София.
- Шопова, М. (2018). *Статистически анализ на структури*, Академично издателство „Ценов“, Свищов.
- Blom, A., Gathmann, C. & Krieger, U. (2015). Setting Up an Online Panel Representative of the German Population: The German Internet Panel. *Field Methods*, November, vol. 27, 4, 391-408.
- Fenner, Y., Garland, S., Moore, E., Jayasinghe, Y., Fletcher, A., Tabrizi, S., Gunasekaran, B. & Wark, J. (2012). Web-based recruiting for health research using a social networking site: an exploratory study. *Journal of Medical Internet Research*, Vol. 14, No 1, Jan-Feb. Available at <http://www.jmir.org/2012/1/e20/#ref26> (Последен достъп: 04.11.2018)
- Fielding, N., Lee, R. & Blank, G. (2008). The Internet as a Research Medium: An Editorial Introduction to *The SAGE Handbook of Online Research Methods*. An Article in “The SAGE Handbook of Online Research Methods”, Edited by Nigel Fielding, Raymond Lee, Grant Blank.
- Utts, J. & Heckard, R. (2007). *Mind on Statistics*. Thomson Learning.
- НСИ. (2014). *Образование в Република България*. София. Докладът е наличен на адрес <http://bit.ly/2QeXTY5> (Последен достъп: 04.11.2018)
- НСИ. (2017). Доклад на НСИ „Лица, които никога не са използвали интернет“, наличен на адрес <http://bit.ly/1Np8jBd> (Последен достъп: 21.02.2018)
- НСИ. (2018). *Статистически справочник 2018*. София.
- ESOMAR. *Global Market Research 2014. An ESOMAR Industry Report*.
- ESOMAR and GLOBAL RESEARCH BUSINESS NETWORK. (2015). *Online Sample Quality Guideline*. February. Available at <http://bit.ly/297RfPT>, p. 6. (Последен достъп: 04.11.2018)
- Facebook, Inc. (2014). *Annual report for 2013*, Published on Jan 31. Available at <http://bit.ly/2JBZzZ3> (Последен достъп: 04.11.2018)
- MinorMonitor. (2012). *Report of Survey Results*. April 4. Available at <http://bit.ly/2qsZiz9> (Последен достъп: 04.11.2018)