

НАЦИОНАЛНО ПРЕДСТАВИТЕЛНО КОЛИЧЕСТВЕНО ПРОУЧВАНЕ НА ПАЗАРИТЕ НА ТЕЛЕВИЗИЯ, ИНТЕРНЕТ И ПАКЕТНИ ЕЛЕКТРОННИ СЪОБЩИТЕЛНИ УСЛУГИ В БЪЛГАРИЯ

ЦЕЛИ НА ИЗСЛЕДВАНЕТО:

Според заданието на настоящото проучване Алфа Рисърч ще:

1. Проучи проникването сред домакинствата¹ в България на следните електронни съобщителни услуги:
 - Платена телевизия на база сключен договор с оператор и месечна абонаментна такса - кабелна телевизия (DVB-C или аналогова), спътникова (сателитна DVB-S) телевизия, IPTV с предоставено от оператор устройство (IPTV STB) или приложение за SMART телевизионен приемник, мобилна телевизия с предоставено от оператор приложение;
 - Платена телевизия без сключен договор с оператор, но със заплащане на месечна такса - телевизия от мобилно устройство с инсталирано приложение (например Da Paye), WEB TV (например VOYO, PLAY DIEMA XTRA);
 - Безплатна телевизия - спътникова (сателитна) телевизия, ефирна цифрова телевизия, телевизия от мобилно устройство с инсталирано приложение, IPTV с еднократно закупено устройство (IPTV STB) или приложение за SMART телевизионен приемник, WEB TV (безплатно гледане на телевизия през общодостъпен Интернет портал, например <https://btvplus.bg/live/>);
 - Интернет достъп чрез фиксирани мрежи;
 - Пакетни електронни съобщителни услуги, включващи приемане на радио- и телевизионни програми (кабелна, сателитна, IPTV или мобилна телевизия), достъп до интернет (фиксиран или мобилен).
2. Проучи пазарните дялове и направи естимационната оценка на брой абонати /абонаменти /потребители на доставчиците на изброените телекомуникационни услуги.
 - Платена телевизия на база сключен договор с оператор и месечна абонаментна такса - кабелна телевизия (DVB-C или аналогова), спътникова (сателитна DVB-S) телевизия, IPTV с предоставено от оператор устройство (IPTV STB) или приложение за SMART телевизионен приемник, мобилна телевизия с предоставено от оператор приложение;

¹ "Домакинство" - две или повече лица, които живеят заедно в едно жилище или в част от жилище, имат общ бюджет и се хранят заедно, независимо от това, че някои от тях може да нямат родствени връзки помежду си. Домакинство е и едно лице, което живее в самостоятелно жилище, в стая или в част от нея към дадено жилище, има самостоятелен бюджет по отношение на разходите за хранене и задоволяване на други потребности.

- Платена телевизия без сключен договор с оператор, но със заплащане на месечна такса - телевизия от мобилно устройство с инсталирано приложение (например Da Payer), WEB TV (например VOYO, PLAY DIEMA XTRA);
- Безплатна телевизия - спътникова (сателитна) телевизия, ефирна цифрова телевизия, телевизия от мобилно устройство с инсталирано приложение, IPTV с еднократно закупено устройство (IPTV STB) или приложение за SMART телевизионен приемник, WEB TV (безплатно гледане на телевизия през общодостъпен Интернет портал, например <https://btvplus.bg/live/>);
- Интернет достъп чрез фиксирани мрежи;
- Пакетни електронни съобщителни услуги, включващи приемане на радио- и телевизионни програми (кабелна, сателитна, IPTV или мобилна телевизия), достъп до интернет (фиксиран или мобилен).

ОБХВАТ И ПРЕДСТАВИТЕЛНОСТ НА ИЗВАДКАТА

Алфа Рисърч прилага изключително строга и прецизна система за изработване на извадки, в резултат на което постига висока степен на точност в реализираните от агенцията проучвания.

Предвид спецификата и целите на настоящото проучване ще бъде разработена представителна национална извадка за домакинствата в България.

ИЗВАДКОВ МЕТОД - многостепенна стратифицирана случайна извадка. Този метод осигурява максимално добро национално покритие и същевременно - обективен подбор на единиците на изследване. Извадката ще бъде изготвена на базата на система от критерии:

- Национално покритие на ниво област: 28 административни области (население: 7 000 039 души; данни НСИ https://www.nsi.bg/sites/default/files/files/data/timeseries/Pop_6.1.1_Pop_DR.xls) и окрупнено в 6-те района за планиране /NUTS/;
- Национално покритие по тип на населено място (столица, областен град, малък град, село);
- Национално покритие на ниво домакинство (общ брой домакинства 2 708 000; данни Евростат: https://ec.europa.eu/eurostat/statistics-explained/images/5/57/Private_households_by_household_composition%2C_2008_-_2018_%28number_of_households_in_1_000_and_%25_of_household_types%29.png)

В контекста на домакинския сегмент на пазарите на телевизия и интернет услуги (проникване сред домакинствата и пазарни дялове), изследването е представително за пазара, дефиниран като: използване на телевизионни и интернет услуги, мобилно използване.

База за изготвяне и претегляне на извадката – официалните данни на Националния Статистически Институт за структурата на населението в България от последното преброяване през 2011 година (с коригирана актуализация, спрямо ежегодни публикации на Националния Статистически Институт). Наличните данни са на национално, областно и местно /общинско/ равнище и отразяват всички основни социално демографски признаци. Това позволява максимално прецизното им отчитане при формиране структурата на извадката с оглед постигането на пълно съответствие със социално-демографската структура на целевата група. Алфа Рисърч разполага с пълната база данни, закупена от НСИ.

МЕТОДИКА ЗА ОПРЕДЕЛЯНЕ ОБЕМА И РЕГИОНАЛНОТО РАЗПРЕДЕЛЕНИЕ НА ИЗВАДКАТА

Предвижданият от нас обем на извадката е 5 700 ефективни интервюта. По този начин ще бъдат спазени всички статистически изисквания за надеждност и ще се гарантира достатъчен брой интервюта във всички сегменти на целевата група (т.е. по географско покритие, типове населени места, брой членове на домакинство, пол, етнос, възрастови под-групи, равнище на образование и т.н.).

Методиката, по която е определен размерът на извадката, е следната:

Обема на извадката първоначално се изчислява на ниво район за планиране. Тя трябва да съдържа достатъчен брой единици, така че да се постигне малка случайна грешка.

Оттук, формулата за обем на извадката за район с около 450 000 домакинства се получава въз основа на формулата за максимална допустима грешка:

$$\text{обем на извадката } SS = \frac{N * Z^2 * p * (1-p)}{(C^2 * (N-1) + Z^2 * p * (1-p))}, \text{ където}$$

N= общ брой домакинства за съответния район за планиране

Z = приема стойност 1.96 при 95% доверителен интервал

p = очакван пропорция на попадналите в дадената група (50%)

C= доверителния интервал, изразен с десетично число.

Прилагането на горната формула гарантира, че при 95% гаранционна вероятност, максималното стандартно отклонение на най-малката извадка ще бъде до +/- 5% при 50-процентен дял.

По този начин ще бъдат спазени всички статистически изисквания за надеждност и ще се гарантира валидни, надеждни и представителни резултати.

КАЛКУЛАТОР ЗА ИЗЧИСЛЯВАНЕ НА ОБЕМ НА ИЗВАДКА



Домакинства на най-малката съвкупност (Североизточен район)	371228
Доверителен интервал (стандартно използван 95%)	95 %
Марж на доверителен интервал (+/- 5%)	5 %
Очаквана пропорция в дадена група (т.е. p=50%)	50 %
НЕОБХОДИМ ОБЕМ НА ИЗВАДКАТА	382

От така направените изчисления приемаме, че най-малката допустима подизвадка (на ниво район за планиране) ще бъде с обем минимум 400 интервюирани домакинства. За да се гарантира по-висока степен на регионална представителност на областите, влизаци в съответните райони за планиране, най-малката възможна под-извадка е увеличена на 720 души. Стартирайки от нея, чрез прилагане на статистически процедури, определяме размера на подизвадките във всички 6 района за планиране, които са представени пропорционално на тяхната тежест на национално равнище.

Таблица №1 илюстрира предварителното разпределение на анкетираните домакинства по район за планиране и тип населено място.

Район за планиране	%	Общо		В градовете		В селата	
		Анкетни	Гнезда брой	%	Анкетни	%	Анкетни
ОБЩО	100.0%	5700	950	68.3%	3940	31.7%	1760
Северозападен	13.4%	760	127	7.6%	440	5.8%	320
Североизточен	12.7%	720	120	8.8%	500	3.9%	220
Северен централен	12.7%	720	120	8.0%	450	4.8%	270
Югозападен	27.3%	1560	260	21.9%	1270	5.4%	290
Югоизточен	14.6%	830	138	9.8%	570	4.8%	260
Южен централен	19.3%	1110	185	12.2%	710	7.1%	400

Формулата за определяна на максималната статистическа грешка за цялата извадка е следната:

$$sample_error = \frac{z\sqrt{p\%(100 - p\%)}}{\sqrt{s}}$$

s = големина на извадката

z = приема стойност 1.96 при 95% доверителен интервал

p = очакван пропорция на попадналите в дадената група (50%)

На база статистическите изчисления и общият размер на извадката сред целевата група от 5 700 анкетирани домакинства получаваме максимално допустимо стандартно отклонение на национално ниво от +/-1,30% при 50-процентен дял.

Таблица №2 Размери на максимална стохастична грешка при описания размер на извадката:

Относителен дял	Максимална грешка	Доверителен интервал	
0.5%	±0.18	0.32	0.68
1%	±0.26	0.74	1.26
2%	±0.36	1.64	2.36
3%	±0.44	2.56	3.44
5%	±0.57	4.43	5.57
10%	±0.78	9.22	10.78
15%	±0.93	14.07	15.93
20%	±1.04	18.96	21.04
25%	±1.13	23.87	26.13
30%	±1.19	28.81	31.19
35%	±1.24	33.76	36.24
40%	±1.27	38.73	41.27
45%	±1.29	43.71	46.29
50%	±1.30	48.70	51.30

CONFEDENCE LEVEL 95%

Прилагайки формулата за изчисляване на максимална статистическа грешка при 50 процентен дял в най-малките райони за планиране (Североизточен и Северен централен) с извадка от по 720 домакинства, получаваме максимално допустимо стандартно отклонение +/- 3,65, а в най-големия район за планиране (Югозападен) с 1 560 домакинства е +/- 2,48.

При така описаната методология изследването ще може да оцени доставчици на платена телевизия, интернет и/или пакетни услуги с 2300+ абоната на национално ниво.

ЕТАПИ ЗА ИЗГОТВЯНЕ НА ИЗВАДКАТА

Първа стъпка - населението на страната в 28-те административни области ще бъде стратифицирано по район и тип населено място. Във всеки регион се избират случайни населени места, така че да са репрезентативни за съответната административна област. Анкети ще бъдат проведени във всички административни области (пропорционално на тяхната тежест Таблица 1) и във всички типове населени места в тях – големи градове, малки градове и села.

По данни на НСИ, както се вижда и от разпределението в Таблица 3, общият брой населени места в страната е 5256. От тях 164 са без население, а 1900 с население под 50 души. Те ще бъдат изключени от списъка на населени места, обхващащи генералната съвкупност, тъй като не дават възможност за надеждна естимация на резултатите от тях. Така, при 3192 населени места с население над 50 души, обхванатите от извадката 450, ще представляват приблизително 15% от населените места с реално живущо население, което осигурява висока надеждност на резултатите както на национално равнище, така и за отчитане на регионалните специфики.

Таблица №3

	Население	% от Населението	Брой населени места
0 жители		0,00%	164
Под 150 души	112 067	1,60%	2146
От 150 до 250 души	122 626	1,75%	619
От 251 до 400 души	199 333	2,85%	620
От 401 до 550 души	175 221	2,50%	374
От 551 до 1000 души	447 186	6,39%	608
От 1001 до 2999 души	845 376	12,08%	521
От 3000 до 17999 души	1 018 533	14,55%	158
От 18000 до 99999 души	1 673 573	23,91%	40
Повече от 100000 души	1 164 449	16,63%	5
София	1 241 675	17,74%	1
ОБЩО	7000039	100,00%	5256

Втора стъпка - подбираме необходимия общ брой гнезда (950 бр.) на случаен принцип в избраните от стъпка едно населени места. Във всяко гнездо ще бъдат анкетирани по 6 респондента с цел избягване на вътрешно-гнездовата корелация и минимизиране на систематичните отклонения. Съответно колкото по-голяма е населеното място, толкова по-голям брой гнезда могат попадат в него.

Трета стъпка - на случаен принцип ще бъдат подбрани началните точки за стартиране на анкетирането във всяко гнездо и ще бъде определена стъпка за систематичен подбор на домакинствата по следната формула:

$$s_i = \frac{T_i}{U_i}$$

s_i - стъпка за определяна на подбора на конкретно домакинство

T_i - общ брой домакинства в i -тото гнездо

U_i - брой на домакинствата попаднали в извадката от i -тото гнездо

Четвърта стъпка – От всяко домакинство интервюираме главата на домакинството или човекът, който отговаря за плащането на сметките/избора на комунални услуги на възраст 18+ години.

За да се избегне редуцирането на броя на извадката поради евентуални откази, отсъствие на респондентите и пр., ще се генерират 10% по-големи масиви с адреси, отколкото е реално заложената бройка. При отсъствие на респондента, всяко домакинство ще бъде посетено още два пъти в друг часови диапазон (делничен ден след 17:30 часа или уикенд след 10:00 часа). Отказалите и неоткрити на адреса домакинства ще бъдат заместени по описаната в трета стъпка методология.

МОДЕЛНА ИЗВАДКАТА	
Обем на извадката:	5 700 ефективни интервюта
Брой населени места:	450 населени места; в по-големите населени места ще има пропорционално на големината им по-голям брой гнезда в отделните микрорайони и квартали
Общ брой гнезда:	950 гнезда, макс. 6 респондента в гнездо
Представителност:	За домакинствата на страната, разпределено във всички шест района за планиране.
Метод на формирането на извадката:	Многостепенна стратифицирана случайна извадка на ниво <ul style="list-style-type: none"> ✓ район за планиране ✓ област ✓ тип на населеното място ✓ домакинство
Метод на набиране на информацията:	Пряко стандартизирано интервю (face-to-face) с таблети по домовете на интервюираните лица

КОНТРОЛ И ВАЛИДИРАНЕ НА ДАННИТЕ

Алфа Рисърч е възприела политика на постоянен контрол над качеството на работа, включващ следните процедури, които се прилагат на всеки етап от изследователския проект:

- ✓ Специализиран инструктаж за всички интервюери и регионални супервайзери за разясняване на целите на изследването и начина на неговото провеждане, рекрутирането на респондентите, процедурите на интервюиране и начина на приложение на въпросника
- ✓ Индивидуален инструктаж на местните анкетъорски екипи, извършван от регионалните супервайзери
- ✓ Тестване на въпросниците/пилотно изследване

- ✓ Ежедневен контрол над изпълнението на извадката
- ✓ 20% контрол на реализираните интервюта, извършван от независим екип от сътрудници, следящи качеството на теренната работа, като се покрива не по-малко от 15% от работата на всеки отделен анкетьор
- ✓ Логически оглед и контрол на реализираните интервюта карти

ПРЕТЕГЛЯНЕ НА РЕЗУЛТАТИТЕ:

Вследствие влиянието на различни фактори (нестохастични влияния, изместване от неотговаряне и др.) получените резултати на изследваните лица от извадковите проучвания е възможно да се различават от признаците на генералната съвкупност. Признаците, по които се различават могат да бъдат най-различни: пол, възраст, местоживееене и пр. За да не се получат изместени оценки от изследването, е възможно да се подходи по следните два начина:

- Претегляне на резултатите с признака, при който има изместване (например по признака брой членове на домакинство). Този подход е възможен само, когато има разминаване между данните от извадката и тези за цялата страна само по един признак.
- Калибриране на резултатите. При този подход е възможно да се калибрират данните от извадките с тези за страната едновременно по няколко на брой признака. Този подход се препоръчва от Евростат за всички извадкови изследвания.

Притеглящите променливи се базират на официалните данни на Националния Статистически Институт за структурата на населението в България от последното преброяване през 2011 година (с коригирана актуализация, спрямо ежегодни публикации на НСИ)

ЕКСТРАПОЛАЦИЯ НА ОСНОВНАТА БАЗА ДАННИ

За целите на анализа на пазарните дялове и обем на пазарите (абонаменти) се генерират 3 допълнителни екстраполации на основната база данни:

- Екстраполация „ТВ абонаменти“ – базата данни включва всички ТВ абонаменти (първи и втори ТВ абонамент в основното жилище) като отделни случаи.

- **Екстраполация „Интернет абонаменти“** – базата включва всички интернет абонаменти (първи и втори интернет абонамент в основното жилище) като отделни случаи.
- **Екстраполация „Пакетни абонаменти“** – базата включва всички абонаменти за пакетни услуги за телевизия и интернет.

ЕСТИМАЦИОННИ ОЦЕНКИ НА БРОЙ АБОНАМЕНТИ/АБОНАТИ

Всички оценки на абонаменти/абонати на домакинска база се осъществяват на два етапа:

Етап 1: проекция от извадката към генералната съвкупност чрез проекционен коефициент 475,085 (при извадка с обем $SS= 5\ 700$ и генерална съвкупност 2 708 000, зад едно интервюирано домакинство „стоят“ 475,08 реални домакинства)

Етап 2: претегляне на естимираните бройки според броя домакинства, ползващи един абонамент. Целта е да се избегне естимиране два пъти на един и същи абонамент, ако е ползван от повече от едно домакинство.

- Когато едно домакинство има два отделни абонамента, коефициентът за претегляне е 2.
- Когато едно домакинство има един абонамент и не го дели с друго домакинство, коефициентът е равен на 1 (най-масовия случай).
- Когато две домакинства ползват един и същи абонамент, претеглящият коефициент е 0,5
- Когато три домакинства ползват един и същи абонамент, претеглящият коефициент е 0,33, и т.н.