



Методолошке смјернице за израду извјештаја о квалитету



САДРЖАЈ

УВОД.....	5
1 УВОД У СТАТИСТИЧКИ ПРОЦЕС И ПРОИЗВОД.....	8
1.1 Намјена истраживања.....	8
1.2 Правни основ и одговорност статистичких институција.....	8
1.3 Коришћене класификације.....	8
1.4 Извјештајна јединица.....	8
1.5 Статистичка јединица посматрања.....	8
1.6 Покривеност и обухват.....	8
1.7 Статистички концепти и дефиниције.....	8
2 РЕЛЕВАНТНОСТ, ПРОЦЈЕНА ПОТРЕБА И ПЕРЦЕПЦИЈА КОРИСНИКА.....	9
2.1 Корисници података статистичког истраживања.....	9
2.1.1 Кључни корисници података из статистичког истраживања.....	9
2.1.2 Процјена корисничких потреба.....	9
2.1.3 Мјерење перцепције и задовољства корисника.....	9
2.2 Комплетност података.....	9
2.2.1 Индикатор квалитета и учинка - Стопа комплетности података (R1).....	9
3 ТАЧНОСТ И ПОУЗДАНОСТ.....	10
3.1 Узорачка грешка.....	11
3.1.1 Индикатор квалитета и учинка – Узорачка грешка (A1).....	11
3.1.2 Активности за смањење узорачких грешака.....	12
3.2 Неузорачке грешке.....	12
3.2.1 Неузорачке грешке - Грешке обухвата.....	12
3.2.1.1 Индикатор квалитета и учинка - Стопа прекомјерног обухвата (A2).....	12
3.2.1.2 Индикатор квалитета и учинка – Удио заједничких јединица (A3).....	14
3.2.1.3 Грешка недовољног обухвата.....	14
3.2.1.4 Мјере за смањење грешака обухвата.....	15
3.2.2 Неузорачке грешке - Грешке мјерења.....	15
3.2.2.1 Разлози за настанак грешака мјерења.....	15
3.2.2.2 Мјере за смањење броја грешака мјерења.....	15
3.2.3 Неузорачке грешке - Грешке неодговора.....	15
3.2.3.1 Индикатор квалитета и учинка - Стопа неодговора јединица (A4).....	16
3.2.3.2 Индикатор квалитета и учинка - Стопа неодговора варијабле (A5).....	18
3.2.3.3 Поступци у случају неодговора.....	19
3.2.3.4 Поступци за смањење стопе неодговора.....	19
3.2.4 Ревизије.....	19
3.2.4.1 Индикатор квалитета и учинка - Просјечна величина ревизије података (A6).....	19
3.2.5 Импутација.....	22
3.2.5.1 Индикатор квалитета и учинка - Стопа импутираних података (A7).....	22
4 ПРАВОВРЕМЕНОСТ И ТАЧНОСТ ОБЈАВЉИВАЊА.....	23
4.1 Правовременост објављивања.....	23
4.1.1 Индикатор квалитета и учинка - Правовременост првих резултата (TP1).....	23
4.1.2 Индикатор квалитета и учинка - Правовременост коначних резултата (TP2).....	24
4.2 Тачност објављивања.....	25
4.2.1 Индикатор квалитета и учинка – Тачност објављивања (TP3).....	25
4.3 Разлози за већа кашњења и мјере за побољшање правовремености и тачности.....	26
5 УСКЛАЂЕНОСТ И УПОРЕДИВОСТ.....	26
5.1 Усклађеност.....	26
5.1.1 Индикатор квалитета и учинка - Усклађеност између различитих извора података (SN1).....	26
5.1.2 Разлози за већа одступања.....	27
5.2 Упоредивост.....	27
5.2.1 Индикатор квалитета и учинка – Неподударност упоредивих статистика (CC1).....	27
5.2.2 Индикатор квалитета и учинка - Дужина упоредивих временских серија (CC2).....	28

5.2.3	Прекиди у временским серијама	30
5.3	Географска упоредивост	30
5.3.1	Упоредивост с чланицама Европског статистичког система	30
6	ДОСТУПНОСТ И РАЗУМЉИВОСТ, ФОРМАТ ДИСЕМИНАЦИЈЕ	31
6.1	Саопштења у којима се објављују подаци	31
6.2	Публикације у којима се објављују подаци	31
6.3	Онлајн база података	31
6.4	Пристап микроподацима	31
6.5	Доступност методолошке документације	31
6.6	Мјере за побољшање разумљивости дисеминираних резултата	31
6.7	Индикатор квалитета и учинка – Коришћење (консултовање) сетова података (АС1)	31
6.8	Индикатор квалитета и учинка – Коришћење (консултовање) метаподатака (АС2)	32
6.9	Индикатор квалитета и учинка - Стопа комплетности метаподатака (АС3)	32
7	ТРОШКОВИ ИСТРАЖИВАЊА И ОПТЕРЕЋЕНОСТ ДАВАЛАЦА ПОДАТАКА	33
7.1	Трошкови провођења статистичког истраживања	33
7.2	Оптерећеност давалаца података	33
7.3	Мјере за смањивање трошкова и оптерећености	33
8	ПОВЈЕРЉИВОСТ	34
8.1	Повјерљивост - политика	34
8.2	Повјерљивост – поступање са подацима	34
9	СТАТИСТИЧКА ОБРАДА	34
9.1	Извор података	34
9.2	Учесталост прикупљања података	34
9.3	Прикупљање података	35
9.4	Валидација података	35
9.5	Компилација података	35
9.6	Прилагођавања	35
9.6.1	Сезонско прилагођавање	35
	ПРЕГЛЕД ИНДИКАТОРА КВАЛИТЕТА И УЧИНКА	36
	ЛИТЕРАТУРА	37

УВОД

Квалитет у контексту производње и дисеминације статистике подразумијева обавезу произвођача званичне статистике да извјештавају о квалитету статистичких процеса и производа. На основу Приручника за извјештаје о квалитету Статистичке институције Европске уније, Еуростата, из 2009. године (*ESS Handbook for quality reports – EHQR, 2009*), те Стандарда за извјештаје о квалитету у Европском статистичком систему (*ESS Standard for quality reports – ESQR, 2009*) три статистичке институције у БиХ, у 2012. години, усагласиле су прве Смјернице за израду извјештаја о квалитету.

Еуростат је у 2014. години објавио ново издање Приручника за извјештаје о квалитету (*ESS Handbook for quality reports, 2014*). У циљу хармонизације са стандардима Европског статистичког система, три статистичке институције у БиХ су усагласиле садржај смјерница за извјештаје о квалитету са новим приручником Еуростата. Као резултат ових активности, Републички завод за статистику Републике Српске прилагодио је и припремио нове Методолошке смјернице за израду извјештаја о квалитету.

Општи циљ Приручника за извјештаје о квалитету Европског статистичког система (ЕСС) је да пружи препоруке за припрему свеобухватних извјештаја о квалитету за цијели низ статистичких процеса и производа. Главни циљеви Смјерница за израду извјештаја о квалитету је да обезбиједи препоруке за:

- извјештавање о квалитету хармонизовано са препорукама Европског статистичког система, како би се омогућила упоредивост процеса и производа са осталим чланицама ЕУ;
- израду извјештаја о квалитету који ће обухватати све информације неопходне за оцјену квалитета статистичких процеса и производа и идентификацију могућих унапређења.

Стандард Европског статистичког система за извјештај о квалитету дефинише шест типова статистичких процеса:

1. Истраживања на бази узорка

Ова истраживања се најчешће заснивају на пробабилистичкој процедури бирања узорка, а подаци се прикупљају директно од испитаника. За ову врсту истраживања успостављена је теорија тачности која омогућава извјештавање о прецизно дефинисаним компонентама тачности (узорачке и неузорачке грешке).

2. Попис

Попис је статистички процес у којем су истраживањем обухваћене све јединице из оквира (пописне јединице).

3. Статистички процеси који користе административни извор(е)

У овим статистичким процесима користе се подаци прикупљени за друге сврхе, а не за директну производњу статистике. Примјер коришћења административног извора је статистичко табелирање регистрованих пословних субјеката произведено из административне база података Регистра пословних субјеката, чије вођење је у надлежности Агенције за посредничке, информатичке и финансијске услуге.

4. Статистички процеси који укључују вишеструке изворе података

У многим статистичким областима производња статистика је таква да, у различитим фазама статистичког процеса, захтијева различите приступе бирању узорка и мјерењу. На примјер, у истраживањима из области структурних пословних статистика, у којима су основни економски подаци о пословању (производња, финансије, итд.) агрегирани, за различите сегменте истраживања могу да се користе различите јединице, упитници, узорачке шеме и/или друге процедуре истраживања.

5. Цијене и други економски индекси

Ово су посебне врсте статистичких процеса, често са непробабилистичким дизајном, за које: (а) постоји специјализована економска теорија која дефинише циљне концепте економских индекса; (б) структура грешака индекса укључује специјализоване концепте попут прилагођавања квалитета, замјене и поновног узорковања (*re-sampling*); (в) истраживања на узорку користе се у неколико димензија (пондери, производи, продајна мјеста), са комплексним комбиновањем пробабилистичких и непробабилистичких метода и (г) постоји више ових индекса који имају кључну улогу у националном статистичком систему или ЕСС .

6. Статистичка компилација

Овај статистички процес обједињује низ различитих примарних извора, укључујући све раније наведене, како би се добио агрегат од посебне концептуалне важности. Углавном је овдје ријеч о економским агрегатима као што су национални рачуни и платни биланс.

Структура извјештаја о квалитету

Структура извјештаја о квалитету се састоји из сљедећих поглавља:

- (1) Увод у статистички процес и производ;
- (2) Релевантност, процјена потреба и перцепција корисника;
- (3) Тачност и поузданост;
- (4) Правовременост и тачност објављивања;
- (5) Усклађеност и упоредивост;
- (6) Доступност и разумљивост, формат дисеминације;
- (7) Трошкови истраживања и оптерећеност давалаца података;
- (8) Повјерљивост и
- (9) Статистичка обрада.

Листа индикатора квалитета и учинка је урађена на основу ЕСС Приручника за извјештаје о квалитету (*ESS Handbook for quality reports, 2014*), ЕСС Смјерница за имплементацију индикатора квалитета и учинка (*ESS Guidelines for the Implementation of the ESS Quality and Performance Indicators, 2014*) и Техничког приручника о јединствено интегрисаној структури метаподатака (*Technical Manual of the Single Integrated Metadata Structure – SIMS, 2014*).

У овом документу дате су смјернице за припрему извјештаја о квалитету у стандардизованом формату. За сваку компоненту дата је кратка дефиниција, а за поткомпоненте и кратко упутство. У случајевима када је поткомпонента везана за израчунавање одређеног индикатора, дато је упутство за израчунавање. За све индикаторе квалитета и учинка дају се и примјери за њихово израчунавање.

Главни циљ извјештаја о квалитету је да произвођачима разних статистика пружи аналитички увид у цјелокупни статистички процес, а корисницима статистичких резултата обезбиједи додатне информације за исправно коришћење и тумачење резултата.

Главна намјена индикатора квалитета је да произвођачу статистике и кориснику статистичких резултата пружи увид у квалитет статистичких резултата, квалитет процеса путем којег су ти резултати добијени, те у одређеној мјери пружи увид и у квалитет цјелокупног институционалног окружења у којем се истраживање проводи.

Иако су ове смјернице усмјерене на израду извјештаја о квалитету намијењених произвођачима статистике, оне укључују и све информације потребне за израду извјештаја оријентисаних ка корисницима.

Индикатори се, по својој дефиницији, могу односити на три различита предмета у оквиру провођења истраживања:

- на истраживање као цјелину,
- на варијаблу и
- на статистику (статистички резултат).

У првом случају, при провођењу конкретног истраживања нема дилеме на коју вриједност се индикатор односи. У остала два случаја то не вриједи, јер обично у једном истраживању мјеримо више варијабли и израчунавамо (процјењујемо) више статистика.

Индикатори који се односе на варијабле и статистике, у извјештајима о квалитету се исказују само за кључне варијабле односно статистике.

Иако се у извјештајима о квалитету индикатори исказују на годишњем (или вишегодишњем) нивоу, код истраживања чија је периодика краћа од године (нпр. мјесечна и тромјесечна истраживања) индикатори се израчунавају за свако проведено истраживање. Код ових истраживања, у извјештају о квалитету наводе се све вриједности у оквиру посматране године, а уједно се даје и годишњи просјек.

Тамо гдје је примјењиво, поред основне вриједности индикатора која се односи на цијелу посматрану популацију, израчунавају се вриједности индикатора за неке важније домене. Тако се, рецимо, у неком истраживању поред основне вриједности стопе неодговора за цијели узорак, приказује и стопа неодговора по важнијим доменима за које се праве излазне табеле.

У неким случајевима, вриједност индикатора је потребно приказати и графички. Прије свега – код вишегодишњих истраживања гдје се (временске) вриједности индикатора исказују у облику линијских дијаграма, помоћу којих приказујемо кретање вриједности индикатора у различитим периодима.

Извјештај о квалитету израђују статистичари одговорни за одређено статистичко истраживање, у сарадњи са колегама задуженим за узорак, анализу и информационе технологије.

Извјештај о квалитету би требао да има стандардну структуру, са тачно рашчлањеним садржајем. Структура извјештаја је припремљена са намјером „покривања“ што ширег распона различитих истраживања, а одговорност аутора извјештаја је да просуди који дијелови извјештаја су за дато истраживање релевантни.

У случају да нека тачка стандардног извјештаја за одређено истраживање није релевантна, то је потребно навести, али не и испустити цијелу тачку.

У извјештају треба да се наведе веб адреса која корисника води до детаљнијих информација о описаној теми. Ово се прије свега односи на поглавље о доступности и разумљивости (упитници, методологије, публикације и сл. објављени на интернету).

Према смјерницама у овом документу, учесталост израде извјештаја о квалитету варира зависно од потреба, а ажурирање извјештаја се препоручује последице већих измјена структуре података или основних пословних процеса. Годишњи извјештаји о квалитету су стандард, тиме лица задужена за састављање извјештаја не би била преоптерећена с обзиром на то да би се, ако се не догоде неке значајније промјене, материјали из једне године једноставно могли копирати у другу, а једини нови материјал били би ажурирани подаци о индикаторима квалитета и учинка.

1 УВОД У СТАТИСТИЧКИ ПРОЦЕС И ПРОИЗВОД

У уводном дијелу извјештаја о квалитету наводи се кратак опис статистичког процеса и производа, прије свега зашто и како се проводи статистичко истраживање.

Потребно је дати кратак преглед сљедећих информација:

1.1 Намјена истраживања

Наводи се кратак опис сврхе, циља и предмета статистичког истраживања.

1.2 Правни основ и одговорност статистичких институција

Експлицитно се наводи постојећи правни основ (закон, програм, план, регулатива, итд.) на основу којег се врши провођење статистичког истраживања. Затим, наводи се одговорност статистичке институције за провођење истраживања.

1.3 Коришћене класификације

Наводи се списак класификација и номенклатура које су коришћене приликом провођења истраживања.

1.4 Извјештајна јединица

Извјештајна јединица је јединица која даје податке искључиво у статистичке сврхе, у садржају и роковима одређеним планом провођења статистичких истраживања. То могу да буду правна лица (пословни субјекти, предузећа) и њихови дијелови, физичка лица, занатлије, слободне професије, домаћинства, тијела јавне управе и јединице локалне самоуправе и сва друга лица које дају податке.

1.5 Статистичка јединица посматрања

Статистичка јединица посматрања је основна јединица на коју се подаци односе, за коју се подаци прикупљају или изводе на други начин (физичко лице, правно лице и његов дио, занатлија, слободна професија, домаћинство, тијело јавне управе, јединица локалне самоуправе и друге основне јединице посматрања на које се подаци односе). Статистичка јединица је предмет статистичких истраживања и носилац статистичког обиљежја.

1.6 Покривеност и обухват

Наводе се сљедеће информације:

- која метода је била коришћена при избору јединице посматрања (случајни узорак, попис – пуни обухват);
- број јединица посматрања, у случају узорка – број јединица у оквиру, те подаци о могућој стратификацији.

1.7 Статистички концепти и дефиниције

Наводе се дефиниције кључних варијабли.

2 РЕЛЕВАНТНОСТ, ПРОЦЈЕНА ПОТРЕБА И ПЕРЦЕПЦИЈА КОРИСНИКА

Релевантност је степен до којег статистички производ задовољава садашње и потенцијалне потребе корисника, односно у којој мјери произведена статистика и коришћени концепти (дефиниције, варијабле, класификације, итд) одражавају потребе корисника.

2.1 Корисници података статистичког истраживања

Почетна фаза дизајнирања и управљања статистичким процесима је дефинисање корисничких потреба. Процјена потреба корисника уопштено укључује класификацију корисника, дефинисање намјене за резултате који се желе добити, дефинисање приоритета у задовољавању корисничких потреба и начина на који се добијају повратне информације од корисника (нпр. преко савјетодавних одбора, редовних сазива корисничких група, повратних информација/примједби корисника, анкетним истраживањем, итд).

2.1.1 Кључни корисници података из статистичког истраживања

Наводе се кључни корисници по главним групама сегментације: јавни сектор; пословни субјекти; наука, истраживање и образовање; општа јавност; медији; страни корисници; интерни корисници. Кључни корисници се одређују на основу информација о захтјевима за подацима и публикацијама, на основу чланства/претплате/пријаве за примање публикација и других евиденција о корисницима података из истраживања.

2.1.2 Процјена корисничких потреба

Ако не постоје ЕСС регулативе и препоруке за одређено статистичко истраживање требало би навести (уколико су познати) главне захтјеве корисника за подацима из статистичког истраживања за које се припрема извјештај и описати у које сврхе се подаци користе, као и то да ли постоје документовани/познати захтјеви за подацима које није било могуће произвести истраживањем.

2.1.3 Мјерење перцепције и задовољства корисника

Задовољство корисника је приоритет првог реда. Најдјелотворнија метода мјерења је провођење анкете о задовољству корисника у складу са најбољом праксом анкетних истраживања, коришћењем репрезентативног узорка корисника из одговарајућег оквира. Мјерење перцепције и задовољства корисника се утврђује вриједностима скале задовољства корисника. Индексом задовољства корисника можемо мјерити степен задовољства корисника услугама и производима које му нудимо. Остале, трошковно мање оптерећене, методе процјене укључују анализу продаје или дистрибуције објављених публикација, коментаре корисника, анализу примљених захтјева и примједби, посјета интернетској страници, итд.

2.2 Комплетност података

2.2.1 Индикатор квалитета и учинка - Стопа комплетности података (R1)

Дефиниција индикатора

Ову стопу израчунавамо као однос између броја статистика (података) које објављујемо у истраживању и броја података који су прописани (захтијевани) одговарајућим уредбама и регулативима. Вриједност индикатора се односи на цјелокупно истраживање.

Напомена: Овај индикатор је примјењив само ако постоји ЕСС регулатива или смјерница о захтијеваним излазним подацима (статистикама).

Поступак израчунавања

$$R1_{PDR} = \frac{\#A_D^{rqd}}{\#D^{rqd}}$$

гдје је:

$\#D^{rqd}$... број елемената у сету захтијеваних статистика (података) према ЕУ регулативама

$\#A_D^{rqd}$... број елемената у подсету расположивих статистика (података)

Објашњења и рачунски примјери

Нпр, ако одређена регулатива прописује да требамо објавити **5** статистика за ниво Републике Српске, онда је број прописаних (захтјеваних) статистика **5**.

Претпоставимо да су за одређено истраживање објављене **4** статистике за ниво Републике Српске, вриједност индикатора је:

$$R1_U = \frac{4}{5} = 0,80 = 80\%$$

3 ТАЧНОСТ И ПОУЗДАНОСТ

Тачност и поузданост процјене се дефинише као степен слагања/блискости између процијењених/израчунатих вриједности (које добијамо на крају статистичке обраде) и стварних, али непознатих вриједности популације.

Најчешће је изражавамо у апсолутном облику (стандардна грешка), релативном облику (коэффициент варијације) или у смислу поузданости као интервал повјерења. Величина узорачке грешке зависи од коришћене процјене и коришћеног дизајна узорка.

Вриједност индикатора се односи на кључну статистику и кључне варијабле.

Узорачке грешке треба дати у табели. Потребно је навести за које статистике и/или варијабле се приказују узорачке грешке. Код мјесечног или тромјесечног истраживања, дају се вриједности грешака за сваки мјесец односно тромјесечје посебно, као и просјечна годишња вриједност узорачке грешке. У случају да постоји више различитих статистика и нивоа објављивања, наводе се грешке само за главне групе, остали резултати се дају у прилогу.

У *анкетним истраживањима* гдје је јединица посматрања домаћинство, често се грешке у случајном узорку приказују не само у форми коэффицијента варијације, него и преко интервала повјерења.

Код *пословних истраживања*, посебно гдје процијењене вриједности могу да буду и негативне (нпр, профит, нето вриједности извоза/увоза, итд), најбољи начин да изразимо грешку узорковања је коэффициент варијације.

У случају мјесечних истраживања препоручујемо да се „кретање“ узорачких грешака прикаже и у графичком облику.

3.1 Узорачка грешка

Узорачка грешка се јавља код истраживања која се заснивају на случајном узорку и последица је чињенице да у истраживању не посматрамо цијелу популацију, него само узорак. У том случају, за кључне статистике је потребно процијенити узорачке грешке. Узорачку грешку можемо приказати на више начина.

3.1.1 Индикатор квалитета и учинка – Узорачка грешка (A1)

Поступак израчунавања је одређен дизајном узорка те коришћеним естиматором, зато није могуће написати општу формулу. Овдје дајемо само основне формуле за случај процјене просјечне вриједности популације код једноставног случајног узорка без понављања.

Претпоставимо да смо из популације од N елемената изабрали једноставан случајан узорак величине n и да желимо да на основу узорка процијенимо просјечну варијаблу Y .

Објашњења и рачунски примјери

Вредновање узорачке грешке статистичких процјена је комплексно, како са теоретског тако и са практичног аспекта. Једноставне формуле важе само при једноставном случајном узорковању, што се у пракси ријетко користи. Код коришћења комплексног дизајна узорка или при употреби нелинеарних естиматора теоретски резултати су много захтјевнији и понекад неисказиви у тачном аналитичком облику.

Примјер: Рецимо да након провођења статистичког истраживања (код којег на основу једноставног случајног узорковања без понављања процјењујемо вриједност аритметичке средине варијабле Y) имамо податке приказане у сљедећој табели.

Величина популације (N)	Величина узорка (n)	Процјењена вриједност аритметичке средине популације (\hat{Y})	Процјена варијансе у популацији (s^2)
10 000	500	800	50 000

Затим израчунавамо:

$$\text{Var}(\hat{Y}) = \left(1 - \frac{n}{N}\right) \cdot \frac{s^2}{n}$$

$$\text{Var}(\hat{Y}) = \left(1 - \frac{500}{10000}\right) \cdot \frac{50000}{500} = 0,95 \cdot 100 = 95$$

Стандардна грешка процјене је једнака квадратном коријену узорачке грешке, тј:

$$se(\hat{Y}) = \sqrt{\text{Var}(\hat{Y})}$$

$$se(\hat{Y}) = \sqrt{95} = 9,75$$

Коефицијент варијације се дефинише као однос између стандардне девијације и аритметичке средине и обично се исказује у процентима:

$$CV(\hat{Y}) = \frac{se(\hat{Y})}{\hat{Y}}$$

$$CV(\hat{Y}) = \frac{9,75}{800} = 1,22\%$$

Доње и горње границе *интервала повјерења* при 95% степену поузданости се израчунавају на сљедећи начин:

$$P_{DG} = \hat{Y} - 1,96 \cdot se(\hat{Y}); I_{DG} = 800 - 1,96 \cdot 9,75 = 780,9$$

$$P_{GG} = \hat{Y} + 1,96 \cdot se(\hat{Y}); I_{GG} = 800 + 1,96 \cdot 9,75 = 819,1$$

Интервал повјерења је дакле: **(780,9 ; 819,1)**

Из добијених резултата закључујемо да је просјечно одступање посматране варијабле од њене просјечне вриједности 9,75 или у релативном износу 1,22%.

Уједно, због побољшања прецизности процјена израчунава се и интервал повјерења наведене варијабле којим се, уз 95% поузданости, процјењује да се вриједност варијабле налази у интервалу између 780,9 и 819,1.

3.1.2 Активности за смањење узорачких грешака

Наводи се коментар за узорачке грешке, прије свега – узроци грешака које прелазе унапријед одређене, стандардизоване границе прихватљиве узорачке грешке.

У случају да интервали повјерења нису објављени, даје се упутство како се могу израчунати из објављених узорачких грешака.

У случају да су процијењене узорачке грешке по коришћеним критеријумима превисоке, потребно је описати акције за њихово смањење у будућности.

3.2 Неузорачке грешке

3.2.1 Неузорачке грешке - Грешке обухвата

Грешка обухвата (или грешка оквира) настаје због разлика између популације која је обухваћена оквиром и циљне популације.

Разликујемо три врсте грешака обухвата:

- прекомјерни обухват (надобухват),
- недовољни обухват (подобухват) и
- вишеструки листинг (дуплирање).

3.2.1.1 Индикатор квалитета и учинка - Стопа прекомјерног обухвата (A2)

Дефиниција индикатора

Удио (нерелевантних) јединица доступних у оквиру које не припадају циљној популацији. Нерелевантним јединицама сматрају се јединице које су (обично због грешке или застарјелих података) укључене у оквир или у само истраживање иако нису дио циљне популације.

Ако се истраживање проводи на узорку, стопа прекомјерног обухвата процјењује се на основу података прикупљених на узорку и у том случају може се рачунати пондерисана и непондерисана стопа прекомјерног обухвата. Ако се истраживање заснива на узорку те свака јединица из узорка представља одређен број јединица у циљној популацији, логично је рачунати пондерисану стопу прекомјерног обухвата.

У случају да све јединице из узорка имају једнак пондер, пондер се у поступку израчунавања занемарује и тада говоримо о израчунавању непондерисане стопе прекомјерног обухвата.

У случају да се проводи периодично истраживање (нпр. мјесечно, тромјесечно), потребно је навести вриједности индикатора за сваки период (нпр. мјесец, тромјесечје) као и просјечну годишњу вриједност индикатора. Вриједности индикатора треба дати у табели. Код периодичних истраживања може се дати и графички приказ кретања вриједности индикатора кроз вријеме (нпр. линијски дијаграм). Такође, треба дати објашњења о могућим екстремним вриједностима индикатора.

Вриједност индикатора се односи на јединице у узорку.

Поступак израчунавања

$$OCr_w = \frac{\sum_o w_j + (1-\alpha)\sum_Q w_j}{\sum_o w_j + \sum_E w_j + \sum_Q w_j}$$

гдје је:

O... број јединица изван обухвата (не припадају циљној популацији)

E... број јединица унутар обухвата (припадају циљној популацији)

Q... број јединица непознате вриједности

W_j... узорачки пондер јединице

α...процијењени удио јединица непознате релевантности које су у ствари релевантне (у пракси се углавном претпоставља да износи 1, бројник у том случају обухвата само јединице изван обухвата)

Објашњења и рачунски примјери

Примјер 1: У узорак за Анкету о потрошњи домаћинстава изабрано је 9 400 адреса (станова), од чега је празних станова било 815. На 205 адреса, стамбена јединица није била доступна или је није било могуће пронаћи. Примјер приказује поступак израчунавања непондерисане стопе прекомјерног обухвата (тј. стопе за адресе) под претпоставком да све јединице имају исти пондер.

$$OCr_w = \frac{815+205}{9400} = 0,1085 = 10,85\%$$

Индикатор се рачуна стављањем у омјер јединица које нису релевантне (дакле, празних станова) укључујући јединице које нису биле доступне, са свим јединицама које су изабране у узорак. Закључујемо како 10,85% јединица из узорка не припада циљној популацији.

У **Примјеру 2** је приказано израчунавање пондерисане стопе прекомјерног обухвата под претпоставком да све јединице имају свој пондер, и то према ознаци релевантности јединице.

Статус релевантности јединице:

- 1 – релевантна јединица која припада циљној популацији,
- 2 – јединица која не припада циљној популацији (изван обухвата) – нерелевантне јединице,
- 3 – јединице непознате вриједности (недоступне и и јединице које није могуће пронаћи).

Ознака релевантности јединице	Пондер јединице
1	6,3
3	19,0
2	3,7
1	112,3
2	115,5
3	31,2
1	8,8

$$OCr_w = \frac{(3,7+115,5)+(19,0+31,2)}{296,8} = \frac{169,4}{296,8} = 0,571\% = 57,1\%$$

У бројнику се налази збир пондера свих јединица које из одређених разлога нису релевантне или с њима није било могуће остварити контакт. Дакле, то су пондери јединица које у примјеру имају ознаке 2 и 3. Називник садржи суме пондера свих јединица из узорка.

Закључујемо како 57,1% свих јединица из узорка не припада циљној популацији.

3.2.1.2 Индикатор квалитета и учинка – Удио заједничких јединица (A3)

Дефиниција индикатора

Индикатор показује удио јединица обухваћених из два различита извора (статистичко истраживање и административни извор) у односу на укупан број јединица у истраживању.

Индикатор се користи када се административни подаци комбинују са подацима из истраживања, односно када су подаци на нивоу јединица добијени из оба извора (неке варијабле долазе из истраживања, а друге варијабле из административних података), или када подаци за дио јединица долазе из истраживања, а за други дио јединица – из једног или више административних извора.

Индикатор нам пружа информацију комплетности/обухвата извора, тј. у којој мјери јединице (варијабле) постоје у оба извора (статистичко истраживање и административни извор).

Поступак израчунавања

$$Ad = \frac{\text{Број заједничких јединица из статистичког истраживања и административног извора}}{\text{Број јединствених јединица у статистичком истраживању}}$$

Заједничке јединице односе се на оне јединице које су укључене у оба извора у подацима који долазе из истраживања и административног извора.

Индикатор "јединствене јединице у статистичком истраживању" у називнику значи да ако јединица постоји у више извора, онда се рачуна само једном.

3.2.1.3 Грешка недовољног обухвата

Проблем недовољног обухвата је у поређењу с прекомјерним обухватом теже измјерити, а јавља се у случајевима када јединицу нисмо укључили у оквир (самим тим нисмо ни у узорак), иако по дефиницији тамо припада. О овим јединицама немамо никаквих непосредних информација. Због тога не дефинишемо посебан индикатор те је потребно дати било какву (посредну) информацију која бар сугерише на опсег грешке.

3.2.1.4 Мјере за смањење грешака обухвата

Потребно је описати све поступке који се изводе с циљем смањивања грешака обухвата. При томе треба описати како тренутне тако и планиране активности. Проблеми због погрешног обухвата се, рецимо, могу ријешити на сљедећи начин: (1) у посматраној години се забиљеже одговори јединица који се добију с терена, додијели им се статус јединице посматрања, а затим се код истраживања за наредну годину прихвате ти нови статуси; (2) прије слања упитника на терен, припреми се статистички списак (адресар) нпр. високошколских установа, који се упореди са списком који одржава Министарство образовања, ако се јави несклад између ових адресара, статистички адресар се поправи (допуни).

3.2.2 Неузорачке грешке - Грешке мјерења

Грешке мјерења су грешке које се јављају приликом прикупљања података и доводе до тога да се забиљежене вриједности варијабли разликују од стварних вриједности.

Узроци за ове разлике су сљедећи:

- (1) код инструмената истраживања: образац који се користи за прикупљање података може да доведе до биљежења погрешних вриједности;
- (2) код давалаца података: даваоци података могу свјесно или несвјесно давати погрешне податке;
- (3) код анкетара: анкетар може да утиче на одговоре давалаца података.

3.2.2.1 Разлози за настанак грешака мјерења

Потребно је навести главне разлоге који проузрокују настанак грешака у мјерењу. Ако постоји повратна информација од стране извјештајних јединица потребно је, прије свега, навести који су главни разлози (по њиховом мишљењу) за грешке у мјерењу. У неким истраживањима (посебно оним која се проводе у домаћинствима) треба да се наведу напомене анкетара (ако постоје).

Најчешћи разлози за настанак грешака у мјерењу су:

- лице које попуњава упитник није довољно стручно оспособљено за попуњавање;
- лице које попуњава упитник није довољно пажљиво прочитало упутство за попуњавање;
- недовољна пажња лица које уноси податке у табеле;
- недостатак евиденција код извјештајних јединица;
- преобимна или оскудна методолошка упутства за попуњавање упитника и
- упитник непрецизан и упутство непотпуно.

3.2.2.2 Мјере за смањење броја грешака мјерења

Поступак уређивања/едитовања података идентификује недоследности. Оне су обично последица грешака у оригиналним подацима, али могу такође да буду резултат процесних грешака код шифрирања и уноса података. Потребно је укратко описати поступак у случају откривања грешака. Навести, прије свега, да ли се врши поновно анкетање јединице или се податак исправља ручно или пак, користимо поступак аутоматског уређивања. Потребно је навести све активности које се предузимају с циљем смањења грешака мјерења. Најдјелотворнији начин за смањење броја грешака у мјерењу су прецизна и разумљива методолошка објашњења и упутства која не смију да буду преобимна, као и успостављање директног контакта с извјештајним јединицама.

3.2.3 Неузорачке грешке - Грешке неодговора

Грешка неодговора представља разлику између статистика израчунатих из прикупљених података и оних које би биле израчунате да не постоје недостајуће вриједности.

Постоје двије врсте неодговора:

- (1) неодговор јединице, који се јавља када уопште нису прикупљени подаци о јединици популације одређеној за прикупљање података, и
- (2) неодговор за одређене варијабле, који се јавља када се прикупе подаци само за неке, али не и за све варијабле о одређеној јединици популације.

3.2.3.1 Индикатор квалитета и учинка - Стопа неодговора јединица (A4)

Дефиниција индикатора

Однос између броја јединица посматрања за које нисмо прикупили податке (барем за неке варијабле) према укупном броју јединица (дизајнираних – одређених за прикупљање података). Неконтактиране јединице посматрања за које не знамо какав је њихов статус (да ли су релевантне или нерелевантне) такође сматрамо неодговором.

У зависности од тога да ли рачунамо показатеље за узорак или цијелу популацију, може се израчунати пондерисана и непондерисана стопа неодговора јединица.

У случају када се ради о периодичном истраживању (нпр. мјесечно, тромјесечно), потребно је да се наведу вриједности индикатора за сваки период (нпр. мјесец, тромјесечје) као и просјечна годишња вриједност индикатора.

Вриједност индикатора се односи на цијело истраживање. Вриједности индикатора треба приказати у табели.

Поступак израчунавања

По завршеном прикупљању података јединице можемо подијелити на четири групе:

- **Нерелевантне јединице (NJ)**. Јединице које су иначе биле изабране у узорак, али се у фази прикупљања података показало да нису више дио циљне популације коју посматрамо. Најчешћи разлог што су све ове јединице дио узорка је неажурност и непотпуност извора које користимо за утврђивање оквира узорка.
- **Неодговори (NR)**. Јединице које су релевантне за истраживање, али нисмо од њих успјели добити жељене податке. У овој групи су најчешће јединице које одбијају да учествују у истраживању.
- **Јединице са непознатом ваљаношћу (Q)**. Јединице за које нисмо успјели добити жељене информације и не знамо да ли су релевантне за истраживање или не. Ово важи углавном за јединице са којима није могуће успоставити контакт.
- **Одговори (R)**. Јединице за које смо успјели добити све жељене информације. При том је потребно одредити (за поједино истраживање) која је најмања количина потребних информација које јединица мора дати да би се могла уврстити у одговоре.

Непондерисана стопа

У односу на дефинисане категорије, непондерисану стопу неодговора јединице израчунавамо по формули:

$$NR_r = 1 - \frac{R}{R + NR + \alpha \cdot Q}$$

гдје је:

R ... број релевантних јединица које су одговориле

NR ... број релевантних јединица које нису одговориле

Q ... број одабраних јединица непознате релевантности

α ... (процијењен) удио јединица непознате релевантности, а које су заправо релевантне (ако нема добро утемељених претпоставки за процјену параметра α , узимамо да је $\alpha = 1$, што значи да све јединице са непознатом релевантношћу сматрамо неодговором).

Пондерисана стопа

Пондерисану стопу израчунавамо према формули:

$$NRr_w = 1 - \frac{\sum_R w_j}{\sum_R w_j + \sum_{NR} w_j + \alpha \cdot \sum_Q w_j}$$

гдје је:

R... број релевантних јединица које су одговориле

NR... број релевантних јединица које нису одговориле

Q... број одабраних јединица непознате релевантности

w_j... узорачки пондер јединице

α... процијењени удио јединица непознате релевантности које су заправо релевантне (ако нема добро утемељених претпоставки за процјену параметра, углавном се претпоставља да износи 1)

Овај приступ би требао бити коришћен у анкетама гдје су јединице лица и домаћинства.

У пословним и пољопривредним истраживањима је удио јединица, с обзиром на њихов утицај на коначне резултате, много важнији од других, тако да у том случају као помоћну варијаблу морамо узети у обзир варијаблу која одређује величину јединице (нпр. број запослених у предузећу).

Објашњења и рачунски примјери

Наводимо хипотетички примјер истраживања у којем смо након коначног поступка прикупљања података добили сљедеће податке:

	Број јединица посматрања	Збир узорачких пондера	Збир вриједности помоћне варијабле јединица посматрања	Пондерисани збир помоћне варијабле
Одговори (<i>R</i>)	700	5 500	15 600	62 500
Неодговори (<i>NR</i>)	250	1 520	2 900	6 200
Непозната релевантност (<i>Q</i>)	80	550	1 400	2 850
Узорак	1 030	7 570	19 900	71 550

Ако узмемо да је претпостављена вриједност удјела релевантних јединица међу јединицама непознате релевантности једнака $\alpha = 0,90$, вриједност индикатора је:

Непондерисана стопа неодговора јединица

$$- \quad NRr = 1 - \frac{700}{700+250+0,90 \cdot 80} = 1 - 0,68 = 0,32 = 32\%$$

Пондерисана стопа неодговора јединице

$$- \quad \text{са пондерима избора узорка} \quad NRr_w = 1 - \frac{5500}{5500+1520+0,90 \cdot 550} = 1 - 0,73 = 0,27 = 27\%$$

$$- \quad \text{само помоћна варијабла} \quad NRr_w = 1 - \frac{15600}{15600+2900+0,90 \cdot 1400} = 1 - 0,79 = 0,21 = 21\%$$

$$- \quad \text{са корекцијом пондера за помоћну варијаблу} \quad NRr_w = 1 - \frac{62500}{62500+6200+0,90 \cdot 2850} = 1 - 0,88 = 0,12 = 12\%$$

3.2.3.2 Индикатор квалитета и учинка - Стопа неодговора варијабле (A5)

Опис индикатора

Однос између броја јединица посматрања за које нисмо успјели добити податак за дату варијаблу и броја свих јединица које су требале да дају податак за посматрану варијаблу. Овај однос се израчунава само у оквиру јединица посматрања које су релевантне за посматрану варијаблу.

Вриједност индикатора се односи искључиво на кључну варијаблу.

Поступак израчунавања

Као и код стопе неодговора јединица, може се рачунати непондерисана стопа, пондерисана стопа са пондерима избора узорка или пондерисана стопа са исправљеним пондерима са вриједностима помоћне варијабле.

$$NR_{YTW} = 1 - \frac{\sum R_Y w_j}{\sum R_Y w_j + \sum NR_Y w_j}$$

гдје је:

R_Y ... број релевантних јединица које су одговориле на варијаблу Y

NR_Y ... број релевантних јединица које нису одговориле на варијаблу Y иако је тражен одговор

w_j ... пондер јединице

Три основна начина израчунавања стопе су:

- Непондерисана стопа: $w_j = 1$
- Стопа пондерисана пондерима дизајна узорка: $w_j = d_j$
- Стопа пондерисана значајношћу јединица: $w_j = d_j x_j$, гдје је x_j вриједност помоћне варијабле X

Објашњења и рачунски примјери

За изабрану варијаблу релевантне су оне јединице за које бисмо морали да добијемо податак за ту варијаблу. Све јединице релевантне за истраживање не морају да буду релевантне и за изабрану варијаблу.

Разлика између укупног броја релевантних јединица и броја јединица које су релевантне и за изабрану варијаблу је највише повезана са тзв. скоковима у упитнику. Према томе, јединице код којих долази до прескакања одређених питања не улазе у израчунавање стопе неодговора на одређену варијаблу коју смо прескочили.

Примјер:

Ако након питања „Да ли ваше домаћинство има приступ интернету?“ слиједи питање „Да ли домаћинство приступа интернету преко персоналног рачунара?“, за одговор на друго питање релевантне су само јединице, код којих је одговор на прво питање било „ДА“.

У табели су дате претпостављене вриједности за изабрану кључну варијаблу:

	Број јединица посматрања	Збир коначних пондера	Збир вриједности помоћне варијабле посматраних јединица	Пондерисани збир помоћне варијабле
Одговори јединице	700	5 500	15 600	70 000
Релевантне јединице од којих се очекује одговор на кључну варијаблу	450	3 800	11 500	55 000
Одговори релевантних јединица на кључну варијаблу	370	3 000	10 500	52 000

Непондерисана стопа неодговора на кључну варијаблу

$$- NR_{Yr} = 1 - \frac{370}{450} = 0,18 = 18\%$$

Пондерисана стопа неодговора на кључну варијаблу

$$- \text{са пондерима дизајна узорка} \quad NR_{Yr_w} = 1 - \frac{3000}{3800} = 0,211 = 21,1\%$$

$$- \text{само помоћна варијабла} \quad NR_{Yr_w} = 1 - \frac{10500}{11500} = 0,087 = 8,7\%$$

$$- \text{са корекцијом пондера за помоћну варијаблу} \quad NR_{Yr_w} = 1 - \frac{52000}{55000} = 0,055 = 5,5\%$$

Код периодичних истраживања (нпр. мјесечних, тромјесечних), потребно је навести вриједности индикатора за сваки период (нпр. мјесец, тромјесечје) као и просјечну годишњу вриједност индикатора. Вриједности индикатора треба дати у табели. Можемо такође дати графички приказ (нпр. линијски дијаграм) кретања вриједности индикатора у времену.

3.2.3.3 Поступци у случају неодговора

Потребно је описати све поступке (нпр. пондерисање, импутације), које су коришћене због неодговора, како за случај неодговора јединице посматрања тако и за случај неодговора варијабле. Ако се користи поступак пондерисања, потребно је навести формулу за израчунавање пондера. На примјер, ако недостају подаци о одређеним варијаблама, успоставља се телефонска веза с извјештајном јединицом и по потреби допуњавају недостајуће вриједности. Изузетно, ако није успостављен контакт са извјештајном јединицом, вриједност варијабле се може процијенити.

3.2.3.4 Поступци за смањење стопе неодговора

У овом дијелу извјештаја о квалитету описују се сви поступци који се изводе у циљу смањивања стопе неодговора јединице посматрања и варијабле. Код упутства нпр. за израчунавање еквивалента пуног радног времена, додајемо и неке примјере за рјешавање најчешћих комбинација при израчунавању овог индикатора. Такође, потребно да се извјештајним јединицама пошаље обавјештење у којем се позивају да се у случају потешкоћа и нејасноћа при попуњавању упитника обратe Заводу, телефонски или путем електронске поште.

3.2.4 Ревизије

Ревизија може бити планирана и непланирана. Узрок непланиране ревизије углавном је откривање грешака у објављеним резултатима. Кодекс праксе Европске статистике захтијева да планиране ревизије слиједе стандардне, добро успостављене и транспарентне процедуре. То значи да су унапријед најављене ревизије пожељне и да треба навести разлоге за предузимање и природу саме ревизије (нови извори података, нове методе, итд...).

3.2.4.1 Индикатор квалитета и учинка - Просјечна величина ревизије података (А6)**Опис индикатора**

Ревизија се дефинише као разлика између касније и раније процјене кључне варијабле. Просјечна величина ревизије података је просјек ревизија кључних варијабли током одређеног периода.

Објављивање	Посматрани период				
	1	...	t	...	n
1. објављивање	X_{11}	...	X_{1t}	...	X_{1n}
...
k-to објављивање	X_{k1}	...	X_{kt}	...	X_{kn}
...
K-to и коначно објављивање	X_{K1}	...	X_{Kt}	...	X_{Kn}

Поступак израчунавања

С обзиром на дводимензионалну ситуацију описану у дефиницији, постоји више стратегија за израчунавање индикатора. Предлаже се узимање просјека за одређена објављивања за n посматраних периода, односно MAR (просјечна апсолутна ревизија – *Mean Absolute Revision*).

MAR (просјечна апсолутна ревизија):

$$MAR = \frac{1}{n} \sum_{t=1}^n |X_{Lt} - X_{Pt}|$$

гдје је:

X_{Lt} - „касније“ објављивање (L -то објављивање у посматраном периоду t),

X_{Pt} - „пријашње“ објављивање, (P -то објављивање у посматраном периоду t),

n - број ревизија

Индикатор није препоручљив за годишње процјене.

Просјечна апсолутна ревизија углавном се примјењује код индекса, пропорција и осталих релативних података.

AR (апсолутна ревизија):

Ревизију података можемо рачунати и као разлику између први и посљедњи пут објављених података.

$$AR = |X_{Lt} - X_{Pt}|$$

гдје је:

X_{Lt} - „каснија“ процјена, L -то објављивање у посматраном периоду t ,

X_{Pt} - „пријашња“ процјена, P -то објављивање у посматраном периоду t ,

Просјечна релативна ревизија (RMAR):

$$RMAR = \frac{1}{n} \sum_{t=1}^n \left| \frac{X_{Lt} - X_{Pt}}{X_{Pt}} \right|$$

Просјечна релативна ревизија се примјењује за податке у апсолутним износима (нпр. број запослених).

гдје је:

X_{Lt} - „каснија“ процјена, L -то објављивање у посматраном периоду t ,

X_{Pt} - „пријашња“ процјена, P -то објављивање у посматраном периоду t ,

Број ревизија (n) се рачуна као број разлика у објављивањима у временским серијама. У принципу, то је број објављивања умањен за 1, тј. ако су подаци објављени четири пута $n = 4 - 1 = 3$.

Објашњења и рачунски примјери

Подаци који се објављују, а односе се на одређени посматрани период t , током времена се могу мијењати. Број посматраних периода за које се подаци објављују означава се са n , а број објављивања током времена са K . Ревизија података настаје између појединих објављивања везаних за одређени посматрани период и може се дефинисати као разлика између касније и раније процјене.

Примјер 1: Проводи се истраживање у којем се мјесечни индекс у посматраном периоду ревидира у току шест узастопних мјесеци па је израчуната величина ревизије података. Како је ријеч о индексима, израчунава се просјечна апсолутна ревизија (MAR).

Ако имамо сљедеће вриједности мјесечног индекса укупног броја запослених за јануар 2016. (од првог до шестог објављивања):

	1. објав.	2. објав.	3. објав.	4. објав.	5. објав.	6. објав.
Вриједност индекса (јан. 2016, $\emptyset 2015=100$)	98,20	98,40	98,20	98,30	98,10	98,00
Разлика	-	0,20	0,20	0,10	0,20	0,10

величина *просјечне апсолутне ревизије* рачунаће се на сљедећи начин:

$$MAR = \frac{0,20+0,20+0,10+0,20+0,10}{5} = \frac{0,80}{5} = 0,16$$

Закључак је да се мјесечни индекс укупног броја запослених у ових шест објављивања у просјеку мијењао за 0,16.

Примјер 2: Проводи се истраживање у којем се укупан мјесечни број запослених у посматраном периоду ревидира у дванаест узастопних мјесеци. Како је индикатор изражен у апсолутним вриједностима рачуна се просјечна релативна ревизија (RMAR). Индикатор се прво израчуна за сваки мјесец, а затим и за цијелу годину (као просјек свих просјечних релативних ревизија по мјесецима).

Мјесец	Први резултати	Коначни резултати	Просјечна релативна ревизија (%)
I	695 250	696 870	0,23
II	696 200	697 480	0,18
III	697 100	699 200	0,30
IV	696 350	697 800	0,21
V	698 210	698 210	0,00
VI	698 480	699 400	0,13
VII	697 100	699 100	0,29
VIII	701 200	702 850	0,24
IX	700 150	701 800	0,24
X	701 200	702 750	0,22
XI	701 330	702 990	0,24
XII	702 350	702 950	0,09
2014.	-	-	0,20

$$RMAR = (0,23+0,18+0,30+0,21+0,00+0,13+0,29+0,24+0,24+0,22+0,24+0,09)/12 = 0,20$$

Закључак је да се укупан број запослених у ових дванаест објављивања у годишњем просјеку промијенио за 0,20%.

3.2.5 Импутација

Импутација је одговор на недостатке у добијеним подацима. У статистичким истраживањима на бази узорка или пописа разлог за импутацију може да буде неодговор (обично неодговор варијабле), док рецимо у обради индекса цијена, до импутације може доћи због недостајућих цијена.

3.2.5.1 Индикатор квалитета и учинка - Стопа импутираних података (A7)

Опис индикатора

Индикатор се израчунава за кључне варијабле и дефинише као *однос између броја јединица, за које су импутирани подаци за посматране кључне варијабле (због недостајућих или неадекватних вриједности) те броја свих јединица за које имамо било који податак.*

Стопа импутираних података је индикатор који се израчунава у случају када дио вриједности за кључне варијабле процијенимо (импутирамо) једном од уобичајених метода импутације. Податак се може импутирати због недостајућих вриједности или због неадекватних вриједности до којих смо дошли у току уређивања података.

Овај индикатор је под утицајем како неодговора варијабле тако и процеса едитовања (уређивања података) и мјери и релативни износ импутиране вриједности и релативни утицај на коначне процјене из поступака импутације.

Такође, овдје се може израчунати и пондерисана и непондерисана вриједност индикатора. Вриједност индикатора се односи на кључну варијаблу. Вриједности индикатора треба дати у табели. У случају да проводимо периодично истраживање (нпр. мјесечно, тромјесечно), потребно је навести вриједности индикатора за сваки период (нпр. мјесец, тромјесечје) као и просјечну годишњу вриједност индикатора. Код периодичних истраживања можемо дати и графички приказ кретања вриједности индикатора кроз вријеме (нпр. линијски дијаграм).

Поступак израчунавања

Непондерисану вриједност индикатора израчунавамо на сљедећи начин:

$$IR_Y r = \frac{\sum I_Y}{\sum I_Y + \sum K_Y}$$

гдје је:

I_Y ... број јединица за које је варијабла Y импутирана,

K_Y ... број јединица за које је вриједност варијабле Y остала непромијењена

Ако је w_j финални пондер, пондерисану вриједност израчунавамо на сљедећи начин:

$$IR_Y r_W = \frac{\sum_{I_Y} w_j y_j}{\sum_{I_Y} w_j y_j + \sum_{K_Y} w_j y_j}$$

Објашњења и рачунски примјери

Код израчунавања стопе импутираних података морамо узети у обзир само оне вриједности, које су биле импутиране са релевантном статистичком (импутацијском) методом.

Јединице за које су подаци у поступку уређења података кориговани поновним провјеравањем код извјештајних јединица, не сматрамо јединицама са импутираним подацима.

У табели су дати подаци за наш примјер:

	Број јединица посматрања	Пондерисани збир варијабле Y
Подаци за варијаблу Y	500	12 500
Импутирани подаци	150	1 750

Непондерисана вриједност индикатора $IR_Y r = \frac{150}{500} = 0,30$

Пондерисана вриједност индикатора $IR_Y r_W = \frac{1\,750}{12\,500} = 0,14$

4 ПРАВОВРЕМЕНОСТ И ТАЧНОСТ ОБЈАВЉИВАЊА

Правовременост објављивања података односи се на временски размак између посљедњег дана периода посматрања на који се подаци односе и датума објављивања.

Тачност објављивања је временски размак између стварног датума објављивања података и најављеног датума који је назначен у званичном календару публикавања.

4.1 Правовременост објављивања

Поједине статистике се објављују у неколико верзија (нпр. прва, претходна или прелиминарна, ревидирана и коначна). У том случају свако издање има свој профил правовремености.

Потребно је посебно објаснити и навести разлоге за могуће кашњење у публикавању, као и навести напоре који су предузети за побољшање ситуације.

4.1.1 Индикатор квалитета и учинка - Правовременост првих резултата (TP1)

Опис индикатора

Правовременост првих резултата је временски размак између краја периода за који су подаци израчунати и тачног датума првог објављивања резултата.

Податке о правовремености првог објављивања резултата треба приказати у табели.

Овај индикатор се рачуна за сва статистичка истраживања која објављују претходне или прелиминарне податке.

Поступак израчунавања

Ако датум који означава посљедњи дан (датум) периода посматрања статистичких података означимо са d_{refp} , а датум објављивања првих резултата са d_{first} , вриједност индикатора је:

$$T_1 = d_{first} - d_{refp}$$

Објашњења и рачунски примјери

Вриједност индикатора код мјесечних, тромјесечних и полугодишњих истраживања се исказује у укупном броју дана, док се код годишњих и вишегодишњих истраживања вриједност индикатора исказује у броју мјесеци. У сваком случају, код исказивања вриједности индикатора треба јасно навести коришћену временску јединицу.

Примјер 1: Ако су резултати мјесечног истраживања који се односе на јануар 2014. године први пут објављени 17.03.2014, вриједност индикатора је једнака:

$$T_1 = (17.3.2014) - (31.1.2014) = T+45.$$

Примјер 2: Ако су резултати АПД (Анкете о потрошњи домаћинстава) који се односе на 2014. годину први пут објављени 20.11.2015, вриједност индикатора је:

$$T_1 = (20.11.2015) - (31.12.2014) = T+11.$$

У извјештају о квалитету треба навести: период посматрања; датум публикавања првих (претходних, прелиминарних) резултата; временски размак у формату $T+x$, гдје је T крај периода посматрања, а x број дана (или мјесеци).

4.1.2 Индикатор квалитета и учинка - Правовременост коначних резултата (TP2)

Опис индикатора

Правовременост коначних статистичких резултата је временски размак између краја периода посматрања на који се објављени резултати односе и датума објављивања коначних резултата. Податке о правовремености коначних резултата треба приказати у табеларном формату.

Поступак израчунавања

Ако датум који означава посљедњи дан (датум) периода посматрања статистичких података означимо са d_{refp} , а датум објављивања коначних резултата са d_{finl} , вриједност индикатора је:

$$T_2 = d_{finl} - d_{refp}$$

Објашњења и рачунски примјери

Вриједност индикатора код мјесечних, тромјесечних и полугодишњих истраживања се исказује у укупном броју дана, док се код годишњих и вишегодишњих истраживања исказује у броју мјесеци. Код исказивања вриједности индикатора треба јасно навести коришћену временску јединицу.

Примјер 1: Ако су коначни резултати мјесечног истраживања који се односе на јануар 2014. године објављени 25.03.2015, вриједност индикатора је једнака:

$$T_2 = (25.03.2015) - (31.1.2014) = T+53.$$

Примјер 2: Ако су коначни резултати АПД (Анкете о потрошњи домаћинстава) за 2014. годину објављени 15.02.2016, вриједност индикатора је:

$$T_1 = (15.02.2016) - (31.12.2014) = T+13.$$

У извјештају о квалитету треба навести: период посматрања; датум публикавања коначних резултата; временски размак у формату $T+x$, гдје је T крај периода посматрања, а x број дана (или мјесеци).

Ако се објављују само једни (коначни) резултати истраживања, индикатор TP1 се не израчунава, те се под овом тачком даје само напомена о томе.

4.2 Тачност објављивања

4.2.1 Индикатор квалитета и учинка – Тачност објављивања (TP3)

Опис индикатора

Тачност објављивања је временски размак (број дана) између датума објављивања података и најављеног/планираног датума за објављивање података. Најављени датум се односи на планирани датум у годишњем Календару публикација.

Индикатор се рачуна на два начина.

а) Поступак израчунавања индикатора за произвођаче:

Ако је предвиђени датум објављивања (достављања) статистика једнак d_{sch} , а стварни датум објављивања је d_{act} , вриједност индикатора је: $P_3 = d_{act} - d_{sch}$

Објашњења и рачунски примјери

Вриједност индикатора се израчунава и исказује у броју дана. Ако су резултати за одређену статистику стварно објављени прије предвиђеног датума објављивања резултата, вриједност индикатора може да буде и негативна.

Примјер 1: Ако је предвиђени датум објављивања 16.3.2015. године, а резултати су били објављени 20.3.2015, вриједност индикатора једнака је $P_3 = (20.3.2015) - (16.3.2015) = 4$.

Примјер 2: Ако је предвиђени датум објављивања 16.3.2015. године, а резултати су били објављени 13.3.2015, вриједност индикатора једнака је $P_3 = (13.3.2015) - (16.3.2015) = -3$.

У извјештају о квалитету треба навести: најављени/планирани датум објављивања резултата; стварни датум објављивања резултата; временско кашњење у формату $T+x$, гдје је T најављени датум објављивања, а x број дана.

б) Поступак израчунавања индикатора за кориснике:

Израчунава се стопа тачности објављивања података (P_{3R}) релевантних за групу статистика/резултата.

Ако број статистика/резултата објављених на предвиђени датум објављивања у календару или раније означимо са m_{pc} , а број статистика/резултата који нису објављени до датума најављеног у календару објављивања са m_{up} , вриједност индикатора је:

$$P_{3R} = \frac{m_{pc}}{m_{pc} + m_{up}}$$

Објашњења и рачунски примјери

Вриједност индикатора се израчунава и исказује у процентима.

Примјер 1: Ако је 8 статистика/резултата објављено на предвиђени датума објављивања или раније, а према Календару публикација је планирано објављивање укупно 12 статистика/резултата, вриједност индикатора једнака је:

$$P_{3P} = \frac{m_{pc}}{m_{pc} + m_{up}} = \frac{8}{8+4} = 0,667 = 66,7\%$$

4.3 Разлози за већа кашњења и мјере за побољшање правремености и тачности

У овом дијелу се наводе разлози и додатна објашњења у случају великих временских одступања за вриједности TP1, TP2 и TP3. Ако постоје значајна одступања, треба навести мјере предузете с циљем побољшања правремености и тачности објављивања.

5 УСКЛАЂЕНОСТ И УПОРЕДИВОСТ

Усклађеност два или више статистичких производа (података) односи се на степен до којег су статистички процеси производње података користили исте концепте, класификације, дефиниције, циљну популацију и хармонизоване методе.

Упоредивост се односи на потребу да добијени подаци и информације буду упоредиве у времену (периодима посматрања), између различитих домена те различитих географских подручја.

5.1 Усклађеност

5.1.1 Индикатор квалитета и учинка - Усклађеност између различитих извора података (CH1)

Опис индикатора

Усклађеност између различитих извора података показује однос између статистичких резултата који су произведени истраживањима и сличних статистичких резултата из других референтних извора (нпр. административних извора, краткорочних и структурних пословних показатеља, националних рачуна, итд).

Ако је ријеч о апсолутним износима, усклађеност се приказује у релативном облику, а ако је ријеч о релативним подацима (тј. индекси, проценти), усклађеност се изражава у апсолутном облику.

Поступак израчунавања

Вриједност индикатора се рачуна за кључне варијабле, гдје је то могуће (ако имамо упоредиву варијаблу из референтног извора) и може се исказати у релативном или апсолутном облику, у зависности од врсте података. Индикатор се рачуна према формули:

$$\text{у релативном облику} \quad CH1 = \frac{X_{iref} - X_i}{X_i}$$

$$\text{у апсолутном облику} \quad CH1 = |X_{iref} - X_i|$$

гдје је:

X_{iref} ... вриједност варијабле у посматраном статистичком истраживању,

X_i ... вриједност варијабле из другог (референтног) извора,

i ... број извора података

*Објашњења и рачунски примјери***Примјер 1: Усклађеност између различитих извора података – релативна разлика**

Ако је број запослених добијен као резултат у Анкети о радној снази износио 816 036, а податак за исту кључну варијаблу у другом (референтном) истраживању износио 693 941, вриједност индикатора је:

$$CH1 = \frac{X_{iref} - X_i}{X_i} = \frac{816036 - 693941}{693941} = 0,18$$

Закључак је да је број запослених који је резултат анкетног истраживања већи за 18% од броја запослених из другог извора (у овом случају, референтног истраживања).

Примјер 2: Усклађеност између различитих извора података – апсолутна разлика

Ако је као резултат у Анкети о радној снази, проценат лица са завршеном средњом школом износио 44,5 %, а податак за исту кључну варијаблу из другог извора (у овом случају, Анкете о потрошњи домаћинства) износио 41,6 %, вриједност индикатора је:

$$CH1 = |X_{iref} - X_i| = 44,5 \% - 41,6 \% = 2,9 \%$$

Закључак је да је број лица са завршеном средњом школом (који је резултат Анкете о радној снази) већи за 2,9 процентних поена од броја лица са завршеном средњом школом из другог извора (у овом случају ријеч је о подацима из Анкете о потрошњи домаћинства).

5.1.2 Разлози за већа одступања

У случају већих одступања у резултатима истраживања, потребно је навести разлоге.

5.2 Упоредивост**5.2.1 Индикатор квалитета и учинка – Неподударност упоредивих статистика (CC1)***Опис индикатора*

У доменима гдје постоји упоредива статистика могуће је процијенити географску упоредивост мјерењем неслагања између одлазних и улазних токова за одабране парове земаља. Класичан примјер ове статистике је спољна трговина. Начелно, извоз земље А у земљу В, кроз одређени период треба да буде једнак увозу земље В из земље А. Поред спољне трговине, примјер за ову статистику може да буде и статистика миграција, туризма, FATS, платни биланс, итд).

Билатерална упоредива статистика представља разлику излазних и улазних токова између двије земље, подијељену просјеком ове двије вриједности.

Примјер за земље А и В

$$CC1A_B = \frac{OF_{AB} - mIF_{AB}}{OF_{AB} + mIF_{AB}} \cdot \frac{1}{2}$$

$$CC1B_A = \frac{OF_{BA} - mIF_{BA}}{OF_{BA} + mIF_{BA}} \cdot \frac{1}{2}$$

Заједничка мјера се може добити из двије разлике у односу на просјек токова (једна од могућности је приказана у сљедећој формули):

$$CC1_{AB} = \frac{|OF_{AB} - mIF_{AB}| + |OF_{BA} - mIF_{BA}|}{\frac{OF_{AB} + mIF_{AB}}{2} + \frac{OF_{BA} + mIF_{BA}}{2}}$$

гдје је:

OF_{AB} ... одлазни ток (извоз) који иде из земље А у земљу В,

mIF_{AB} ... упоредиви долазни ток за земљу А,

OF_{BA} ... одлазни ток (извоз) који иде из земље В у земљу А,

mIF_{BA} ... упоредиви долазни ток за земљу В.

Објашњења и рачунски примјери

Примјер: Узмимо податке спољнотрговинске размјене између БиХ и Хрватске за 2014. годину. Земља А је Босна и Херцеговина, а земља В је Хрватска.

Подаци о спољнотрговинској размјени са Хрватском (објављени у статистици БиХ) су сљедећи:

ИЗВОЗ (OF_{AB}) = 488 308 (у хиљ. €)

УВОЗ (mIF_{BA}) = 946 756 (у хиљ. €)

Подаци о спољнотрговинској размјени са БиХ (објављени у статистици Хрватске) су сљедећи:

ИЗВОЗ (OF_{BA}) = 1 223 664 (у хиљ. €)

УВОЗ (mIF_{AB}) = 460 275 (у хиљ. €)

На основу датих података израчунавамо неподударност између ових упоредивих статистика:

$$CC1_{AB} = \frac{|OF_{AB} - mIF_{AB}| + |OF_{BA} - mIF_{BA}|}{\frac{OF_{AB} + mIF_{AB}}{2} + \frac{OF_{BA} + mIF_{BA}}{2}} = \frac{(488308 - 460275) + (1223664 - 946756)}{\frac{949083}{2} + \frac{2170420}{2}} =$$

$$= \frac{(28033) + (276908)}{474542 + 1085210} = \frac{304941}{1559752} = 0,20$$

Када би коефицијент био једнак нули постојала би савршена симетрија (подударност) података (односно одлазни токови би били једнаки упоредивим долазним токовима).

Међутим, у нашем случају видимо да коефицијент асиметрије (неподударности) износи 0,20 па закључујемо да неподударност ових статистика износи 20%.

5.2.2 Индикатор квалитета и учинка - Дужина упоредивих временских серија (CC2)

Опис индикатора

Дужина упоредиве временске серије од задњег прекида временске серије, односно број периода посматрања у временској серији од посљедњег прекида.

Вриједност индикатора се односи на кључне статистике истраживања.

Поступак израчунавања

Ако је J_{last} редни број посљедњег периода посматрања за који су објављени резултати (посљедња тачка у временској серији), а J_{first} редни број првог периода посматрања са већ упоредивим статистичким резултатима (прва временска тачка према могућем прекиду), вриједност индикатора се рачуна према формули:

$$CC2 = J_{last} - J_{first} + 1$$

Објашњења и рачунски примјери

Вриједност индикатора показује број периода у временској серији што значи да мјерна јединица зависи од периодике истраживања. Код мјесечног истраживања вриједност индикатора се изражава као број мјесеци, код тромјесечног као број тромјесечја, итд.

Вриједност индикатора је у великој мјери одређена тиме како дефинишемо прекид у временској серији. Како није могуће поставити посве тачна мјерила за идентификовање прекида, обично треба, бар дјелимично, користити субјективну процјену. Уопштено би се могло рећи да прекид у временској серији настаје онда када дође до такве промјене једног или више аспеката истраживања (нпр. у извору података, у дефиницији параметра који се процјењује или у коришћеној методологији) да коначни резултати међу собом нису више разумно упоредиви.

Ако временска серија нема прекида, вриједност индикатора показује дужину цијеле временске серије (видјети *Примјер 1* и *Примјер 2*).

Примјер 1: Претпоставимо да је период посматрања који представља почетак упоредиве временске серије мјесечног истраживања јануар 2009. године. Ако је посљедњи период посматрања јуни 2016. године, индикатор ће имати вриједност:

$$CC2 = (\text{јуни } 2016. - \text{јануар } 2009.) = 90 \text{ (мјесеци)}.$$

Примјер 2: Ако је код годишњег истраживања година првог упоредивог објављивања резултата 2005, онда је по објављивању резултата за 2016. годину вриједност индикатора једнака:

$$CC2 = (2016. \text{ година} - 2005. \text{ година}) = 12 \text{ (година)}.$$

Ако је, на примјер, почетком 2011. године измијењена Класификација дјелатности, а на основу тога промијењен и узорак јединица посматрања, није било могуће успоставити одговарајућу везу са старим подацима и долази до прекида у временској серији. Како временска серија има прекид у одређеном дијелу унутар серије, а подаци су се и даље наставили објављивати, могуће је израчунати двије вриједности показатеља (видјети *Примјер 3*).

Примјер 3: Ако је, нпр, код тромјесечног истраживања *индекса промета у трговини на велико* (први период посматрања је 1. тромјесечје 2007. године, а посљедњи је 4. тромјесечје 2014. године) дошло до прекида серије у 2011. години (због измјена у Класификацији дјелатности), дужина временске серије износи:

Тромјесечје	Индекси, 2007=100	Број периода посматрања
I 2007	80,4	1
II 2007	102,1	2
III 2007	118,2	3
IV 2007	11,3	4
I 2008	71,5	5
II 2008	97,2	6
III 2008	102,8	7
IV 2008	97,9	8
I 2009	77,9	9
II 2009	101,7	10
III 2009	114,2	11
IV 2009	107,9	12
I 2010	84,6	13
II 2010	116,8	14
III 2010	117,6	15
IV 2010	114,0	16

Тромјесечје	Индекси, 2007=100	Број периода посматрања
I 2011	прекид	прекид
II 2011	119,2	1
III 2011	123,2	2
IV 2011	122,2	3
I 2012	105,3	4
II 2012	148,4	5
III 2012	150,2	6
IV 2012	149,4	7
I 2013	99,8	8
II 2013	127,1	9
III 2013	145,7	10
IV 2013	146,2	11
I 2014	94,0	12
II 2014	132,2	13
III 2014	141,0	14
IV 2014	129,0	15

$$CC2_1 = J_{last} - J_{first} + 1 = 16 - 1 + 1 = 16$$

$$CC2_2 = J_{last} - J_{first} + 1 = 15 - 1 + 1 = 15$$

5.2.3 Прекиди у временским серијама

Описују се сви прекиди у временским серијама као и разлози за настанак прекида. Описују се и сви фактори који нису проузроковали прекид временских серија, али су на одређени начин утицали на смањену упоредивост резултата у различитим временским тачкама.

5.3 Географска упоредивост

5.3.1 Упоредивост с чланицама Европског статистичког система

Наводе се фактори који би могли утицати на то да резултати одређеног истраживања не буду сасвим упоредиви с резултатима сличних истраживања у Европском статистичком систему. Наводи се, такође, да ли постоји регулатива која уређује хармонизовану употребу (неких дијелова) методологије.

6 ДОСТУПНОСТ И РАЗУМЉИВОСТ, ФОРМАТ ДИСЕМИНАЦИЈЕ

Дефиниција компоненте

Доступност и разумљивост статистичких производа/резултата подразумевају једноставан и лак начин на који корисници могу да приступе статистичким подацима користећи једноставне и лаке процедуре. Доступност статистичких резултата/производа се односи на конкретне физичке околности, у којима су подаци доступни кориснику: гдје се подаци физички налазе, какве су могућности коришћења, распоред објављивања, јасна политика плаћања, доступност микро и макроподатака, различити формати и медији (нпр. саопштења, публикације, онлајн базе података, документација о методологији и квалитету).

6.1 Саопштења у којима се објављују подаци

Навести називе редовних и ванредних саопштења са листом објављених сетова података и линком на онлајн саопштење.

6.2 Публикације у којима се објављују подаци

Навести називе редовних и ванредних публикација са листом објављених сетова података и линком на онлајн публикацију.

6.3 Онлајн база података

Навести информацију о доступној онлајн бази података у којој се може приступити објављеним подацима, са линком на базу.

6.4 Приступ микроподацима

Навести информацију да ли су доступни микроподаци, те ако јесу, кратко описати правила анонимизације микроподатака.

6.5 Доступност методолошке документације

Навести информацију о доступности датотека референтних метаподатака¹, важнијих методолошких објашњења, упутстава, приручника, итд (наслов, издавач, година и линк на онлајн документ).

6.6 Мјере за побољшање разумљивости дисеминираних резултата

Описати било које активности које су планиране у вези с побољшањем разумљивости штампаних и интернетских објава, као и дисеминационих база података.

6.7 Индикатор квалитета и учинка – Коришћење (консултовање) сетова података (AC1)

Опис индикатора

Број консултација (прегледа) сетова података унутар статистичке области у одређеном периоду. Напомена: интерно прегледање страница се искључује. Прикупљање података о броју прегледа требало би се вршити на мјесечном нивоу, а ради разумљивијег прегледа, индикатор се може приказати и графички.

Поступак израчунавања, примјер и објашњење

$$AC1 = \#CONS$$

Број корисничких консултација за статистику *Индекса потрошачких цијена* у 2014. години износи 1 255 (*web pages hits*).

¹ Референтни метаподаци су метаподаци који описују садржај и квалитет статистичких података. (Statistical Data and Metadata Exchange, SDMX Glossary, version 1.0, 2016, p77)

6.8 Индикатор квалитета и учинка – Коришћење (консултовање) метаподатака (AC2)

Опис индикатора

Број консултација референтних метаподатака (ESMS – *Euro-SDMX Metadata Structure*) унутар статистичке области, у одређеном периоду (*прегледи или преузимања по предметној статистичкој области – теми*). Напомена: интерно прегледање страница се искључује. Ради разумљивијег прегледа, индикатор се може приказати и графички.

Поступак израчунавања, примјер и објашњење

AC2 = #ESMS

Број консултација референтних метаподатака (ESMS) за статистику *Индекса потрошачких цијена* у 2014. години износи 1 120 (*web pages hits*).

6.9 Индикатор квалитета и учинка - Стопа комплетности метаподатака (AC3)

Опис индикатора

Однос броја расположивих (комплетираних) елемената метаподатака према укупном броју препоручених елемената метаподатака.

Поступак израчунавања, примјер и објашњење

Од укупног броја (60) елемената метаподатака садржаних у ESMS v.2.0 (*Euro-SDMX Metadata Structure*) за статистику *Индекса потрошачких цијена*, осигурано је 55. Индикатор се рачуна према формули:

$$AC3 = \frac{\sum \#ML}{\sum \#L} = 55/60 = 0,92$$

гдје је:

#ML... укупан број обезбијеђених (комплетираних) елемената метаподатака,
#L... укупан број препоручених елемената метаподатака према ESMS v.2.0

У конкретном примјеру, стопа комплетности метаподатака за статистику *Индекса потрошачких цијена* је 92%.

7 ТРОШКОВИ ИСТРАЖИВАЊА И ОПТЕРЕЋЕНОСТ ДАВАЛАЦА ПОДАТАКА

Дефиниција компоненте

Трошкови и оптерећеност давалаца података нису стварне димензије квалитета. Међутим, треба размотрити компромисе између трошкова и оптерећења давалаца података или, другим ријечима, размотрити на који начин трошкови и оптерећење давалаца података ограничавају квалитет статистичких података.

Могућност израчунавања трошкова неопходна је за ефикасно управљање, а посебно за процјену квалитета и учинка. Анализа *трошкова и користи* потребна је како би се утврдили одговарајући компромиси између трошкова и користи у погледу квалитета статистичких података. На исти начин, на учешће давалаца података мора се гледати као на *трошак* (за даваоце података) који се мора уравнотежити са *користима* од података који се дају.

7.1 Трошкови провођења статистичког истраживања

Код израчунавања трошкова статистичке институције при провођењу истраживања треба узети у обзир материјалне трошкове те трошкове који произилазе из утрошка радног времена статистичког особља.

Израчунавање годишњих оперативних трошкова према главним трошковним компонентама приказује се у стандардној табели:

Број утрошених часова рада	
Материјални трошкови (штампање и слање образаца на терен)	
Годишњи број образаца који се достављају извјештајним јединицама	

7.2 Оптерећеност давалаца података

Као индикатор оптерећености давалаца података, оцјењује се вријеме које је потребно извјештајним јединицама за попуњавање упитника у току једне године. Подаци о годишњем оптерећењу давалаца података приказују се у стандардној табели:

Број давалаца података који су попунили образац	
Вријеме потребно за попуњавање једног обрасца (часова)	
Укупно утрошено вријеме (часова)	

7.3 Мјере за смањивање трошкова и оптерећености

У овом дијелу наводе се могуће мјере за смањење трошкова и оптерећености давалаца података (нпр. смањење величине узорка и координација узорка; смањење фреквенције прикупљања података; смањење броја тражених података у истраживању; смањење броја контаката са извјештајном јединицом; коришћење административних података (или дијела административних података); тестирање упитника са извјештајним јединицама у циљу бољег разумијевања и попуњавања; навођење захтјева за подацима у календару истраживања доступном извјештајној јединици; могућност остваривања контакта уколико је извјештајној јединици потребна помоћ при попуњавању упитника; итд ...).

8 ПОВЈЕРЉИВОСТ

У Кодекси праксе Европске статистике (*European Statistics Code of Practice*), у начелима се наводи да тајност давалаца података (лица, домаћинства, предузећа) и повјерљивост информација које пружају мора да буде апсолутно загарантована, а њихови подаци могу да се користе само за статистичке сврхе (начело 5); статистичке институције имају обавезу да развијају, производе и дисеминирају статистику, поштујући стручну независност, на објективан, стручан и транспарентан начин, тако да се сви корисници третирају једнако (начело 6).

8.1 Повјерљивост - политика

Повјерљивост статистичких података је најчешће гарантована законом, а запослени који проводе статистичка истраживања по истом правном основу имају обавезу заштите повјерљивости. Закони о статистици одређују поверљивост као једно од главних начела. У законима је посебно поглавље посвећено овом питању – поглавље садржи дефиницију повјерљивих података, предвиђа употребу поверљивих података само за статистичке потребе, одређује обавезе произвођача података који треба прецизно да дефинишу мјере и процедуре заштите повјерљивости података. Закони о статистици су доступни на интернет страницама статистичких институција.

У извјештају о квалитету треба да се наведу све информације везане за политику повјерљивости података, али и да се извјести о било којем изузетку од овог правила.

8.2 Повјерљивост – поступање са подацима

Такође, потребно је у главним цртама навести процедуре за осигурање повјерљивости за вријеме прикупљања, обраде и дисеминације података – што укључује протоколе за заштиту индивидуалних података којима се приступа, правила за дефинисање повјерљивих ћелија у излазним табелама и процедуре за откривање и превенцију накнадних разоткривања. Исто тако, треба да се наведе да ли се примјењују одредбе о повјерљивости ако спољни корисници имају приступ микроподацима за истраживачке сврхе. У извјештају треба описати одредбе које осигуравају заштиту и интегритет комплетног упитника, микро и макро базе података и резултата.

9 СТАТИСТИЧКА ОБРАДА

Статистичка обрада се односи на операције које се изводе на подацима у циљу добијања нових информација према одређеном скупу правила. Овај концепт се даље дијели на:

9.1 Извор података

Навести на којем извору се заснива скуп података: на истраживању, на административним изворима података, на комбинацији више извора података или података из других статистичких активности. Уколико је коришћен узорак, треба дати неке карактеристике узорка (нпр. величину популације, бруто и нето величину узорка, врсту дизајна узорковања, итд). Ако се користе административни регистри, треба дати опис регистара (извор, примарна сврха, итд).

9.2 Учесталост прикупљања података

Навести учесталост (периодичу) прикупљања података (нпр мјесечно, тромјесечно, годишње, континуирано).

9.3 Прикупљање података

Навести начин прикупљања података, нпр, из административних извора; путем сниматеља цијена који попуњавају упитник (на терену или путем телефона); путем упитника за одређено статистичко истраживање који се дизајнирају у статистичкој институцији. Описати коришћене методе прикупљања података за јединице посматрања (нпр. узорак, теренска анкета, САТИ, онлајн истраживања, итд). Ако постоје, треба навести и неке додатне информације о дизајну и тестирању упитника, обуци анкетара, методе које се користе за праћење неодговора и сл. У тексту се даје и линк на коришћене упитнике.

9.4 Валидација података

У овом дијелу описују се поступци за провјеру и вредновање извора и излазних података и описује се како се резултати ових валидација прате и користе. Активности валидације могу да укључују: провјеру обухвата популације и, по потреби, стопе одговора; поређење статистике са претходним циклусима (ако је примјењиво); поређење статистике са другим релевантним подацима (интерним и екстерним); истраживање недоследности у статистици; уређивање микро и макро података; откривање екстремних вриједности (*outliers*).

9.5 Компилација података

Компилација података подразумијева поступке на подацима који служе за извођење нових информација у складу са датим скупом правила.

Описати процес компилације података (нпр. импутација, пондерисање, подешавање за неодговор, калибрација и сл). За импутације: потребно је навести у којој мјери се користи импутација и навести разлоге за импутацију; дати кратак опис коришћене метода и њен утицај на процјену. Пондерисање: сваки корак пондерисања описати одвојено; описати израчунавање пондера. Подешавање неодговора: описати како је коригован иницијални пондер, узимајући у обзир разлике у стопама одговора. Калибрације: описати варијабле које се користе у подешавању, примијењене методе. Описати израчунавање финалних пондера.

9.6 Прилагођавања

Прилагођавања представљају скуп процедура за модификацију статистичких података у циљу осигуравања компатибилности са националним или међународним стандардима, или процедуре за рјешавање разлика у квалитету података ако се врши компилација специфичних скупова података. У овом дијелу описују се временске серије које су прилагођене и статистички поступци коришћени за подешавање серије (методе десезонирања, нпр. TRAMOSEATS, ARIMA или друге сличне методе). У случају прилагођавања, навести тип прилагођавања (нпр. сезонско, календарско, тренд-циклус). Ако је откривена екстремна вриједност (*outlier*) и извршена замјена, навести које врсте екстремних вриједности су откривене (импулс, пролазне промјене, помаци нивоа). Навести информације о софтверу за подешавање.

9.6.1 Сезонско прилагођавање

Статистичка техника која се користи за отклањање ефеката утицаја сезонског календара на серију података. Навести кратак опис коришћеног метода.

ПРЕГЛЕД ИНДИКАТОРА КВАЛИТЕТА И УЧИНКА

Компонента квалитета и учинка	Ознака	Назив индикатора квалитета и учинка
Релевантност	R1	Стопа комплетности података
Тачност и поузданост	A1	Узорачка грешка - коефицијент варијације
	A2	Стопа прекомјерног обухвата
	A3	Удио заједничких јединица
	A4	Стопа неодговора јединица посматрања
	A5	Стопа неодговора варијабле
	A6	Просјечна величина ревизије података
	A7	Стопа импутираних података
Правовременост и тачност објављивања	TP1	Правовременост првих резултата
	TP2	Правовременост коначних резултата
	TP3	Тачност објављивања
Доступност и разумљивост	AC1	Коришћење (консултовање) сетова података
	AC2	Коришћење (консултовање) метаподатака
	AC3	Стопа комплетности метаподатака
Усклађеност и упоредивост	CH1	Усклађеност између различитих извора података ¹⁾
	CC1	Неподударност упоредивих статистика
	CC2	Дужина упоредивих временских серија
Трошкови истраживања и оптерећеност давалаца података	-	Трошкови провођења статистичког истраживања ¹⁾
	-	Оптерећеност давалаца података ¹⁾

¹⁾ Израчунавање индикатора није садржано у ЕСС препорукама за израду извјештаја о квалитету из 2014. године, али су због значаја саставни дио смјерница.

ЛИТЕРАТУРА

- Агенција за статистику Босне и Херцеговине (2017): *Извјештај о квалитету статистичких истраживања – Смјернице за израду*
- Eurostat (2014): *“ESS Handbook for Quality Reports”, Manuals and guidelines*, Eurostat, Luxembourg
- Eurostat (2014): *“ESS Guidelines for the Implementation of the ESS Quality and Performance Indicators”, Eurostat, Luxembourg*
- Eurostat (2009): *“ESS Standard for Quality Reports”, Methodologies and Working Papers*, Eurostat, Luxembourg
- Државни завод за статистику Републике Хрватске (2015): *„Приручник за израчун показатеља квалитете“*, Загреб
- Statistical Data and Metadata Exchange (2016): *SDMX Glossary, version 1.0*