



**IZVJEŠTAJ O KVALITETU STATISTIČKIH ISTRAŽIVANJA**  
**- Smjernice za izradu -**

**Sarajevo, 2016**



## SADRŽAJ

UVOD	5
1 STATISTIČKI PROCES I STATISTIČKI PROIZVOD	8
2 RELEVANTNOST	9
2.1 Korisnici podataka statističkog istraživanja	9
2.1.1 Ključni korisnici podataka iz statističkog istraživanja	9
2.1.2 Procjena korisničkih potreba	9
2.1.3 Mjerenje percepcije i zadovoljstva korisnika	9
2.2 Potpurnost podataka	9
2.2.1 <u>Indikator kvaliteta i učinka - Stopa kompletnosti podataka (R1)</u>	9
3 TAČNOST I PRECIZNOST	10
3.1 Greška uzorkovanja	11
3.1.1 <u>Indikator kvaliteta i učinka – Greška uzorkovanja (A1)</u>	11
3.1.2 Aktivnosti za smanjenje grešaka uzorkovanja	12
3.2 Neuzoračke greške	12
3.2.1 Neuzoračke greške - Greške obuhvata	12
3.2.1.1 <u>Indikator kvaliteta i učinka - Stopa prekomjernog obuhvata (A2)</u>	13
3.2.1.2 <u>Indikator kvaliteta i učinka – Udio zajedničkih jedinica (A3)</u>	14
3.2.1.3 Greška nedovoljnog obuhvata	15
3.2.1.4 Mjere za smanjenje grešaka obuhvata	15
3.2.2 Neuzoračke greške - Greške mjerenja	15
3.2.2.1 Razlozi za nastanak grešaka mjerenja	16
3.2.2.2 Mjere za smanjenje broja grešaka mjerenja	16
3.2.3 Neuzoračke greške - Greške neodgovora	16
3.2.3.1 <u>Indikator kvaliteta i učinka-Stopa neodgovora jedinice(A4)</u>	16
3.2.3.2 <u>Indikator kvaliteta i učinka - Stopa neodgovora varijable (A5)</u>	19
3.2.3.3 Postupci u slučaju neodgovora	21
3.2.3.4 Postupci za smanjenje stope neodgovora	21
3.3 Revizije	21
3.3.1 <u>Indikator kvaliteta i učinka - Prosječna veličina revizije podataka (A6)</u>	21
3.4 Imputacija	24
3.4.1 <u>Indikator kvaliteta i učinka - Stopa imputiranih podataka (A7)</u>	24
4 PRAVOVREMENOST I TAČNOST OBJAVLJIVANJA	25
4.1 Pravovremenost objavljivanja	25
4.1.1 <u>Indikator kvaliteta i učinka - Pravovremenost prvih rezultata (TP1)</u>	25
4.1.2 <u>Indikator kvaliteta i učinka - Pravovremenost konačnih rezultata (TP2)</u>	26
4.2 Tačnost objavljivanja	27
4.2.1 <u>Indikator kvaliteta i učinka – Tačnost objavljivanja (TP3)</u>	27
4.3 Razlozi za veća kašnjenja i mjere za poboljšanje pravovremenosti i tačnosti	28

5	USKLAĐENOST I UPOREDIVOST	28
5.1	Usklađenost	28
5.1.1	<u>Indikator kvaliteta i učinka - Skladnost sa rezultatima iz referentnog izvora (CH1)</u>	28
5.1.2	Razlozi za veća odstupanja	29
5.2	Uporedivost	30
5.2.1	<u>Indikator kvaliteta i učinka – Nepodudarnost uporedivih statistika (CC1)</u>	30
5.2.2.	<u>Indikator kvaliteta i učinka - Dužina uporedivih vremenskih serija (CC2)</u>	31
5.2.3	Prekidi u vremenskim serijama	32
5.3	Geografska uporedivost	32
5.3.1	Uporedivost s ostalim članicama evropskog statističkog sistema	32
6	DOSTUPNOST I RAZUMLJIVOST, DISEMINACIJSKI FORMAT	32
6.1	Saopćenja u kojima se objavljuju podaci	33
6.2	Publikacije u kojima se objavljuju podaci	33
6.3	On – line baza podataka	33
6.4	Pristup mikropodacima	33
6.5	Dostupnost metodološke dokumentacije	33
6.6	Mjere za poboljšanje razumljivosti diseminiranih rezultata	33
6.7	<u>Indikator kvaliteta i učinka - Konsultovanje setova podataka (AC1)</u>	33
6.8	<u>Indikator kvaliteta i učinka – Konsultovanje meta podataka (AC2)</u>	33
6.9	<u>Indikator kvaliteta i učinka – Stopa kompletnosti meta podataka (AC3)</u>	34
7	TROŠKOVI ISTRAŽIVANJA I OPTEREĆENOST DAVALACA PODATAKA	34
7.1	Troškovi provođenja statističkog istraživanja	34
7.2	Opterećenost davalaca podataka	35
7.3	Mjere za smanjivanje troškova i opterećenosti	35
8	POVJERLJIVOST	35
8.1	Povjerljivost – politika	35
8.2	Povjerljivost – postupanje sa podacima	36
9	STATISTIČKA OBRADA	36
9.1	Izvor podataka	36
9.2	Učestalost prikupljanja podataka	36
9.3	Prikupljanje podataka	36
9.4	Validacija podataka	36
9.5	Kompilacija podataka	36
9.6	Prilagođavanja	37
9.6.1	Sezonsko prilagođavanje	37
	PREGLED INDIKATORA KVALITETA I UČINKA	38
	LITERATURA	39

## UVOD

Opći cilj Priručnika za izvještaje o kvalitetu Evropskog statističkog sistema (*ESS Handbook for Quality Reports*) je pružiti preporuke za pripremu sveobuhvatnih izvještaja o kvalitetu za cijeli niz statističkih procesa i statističkih proizvoda.

ESS standard za izvještaj o kvalitetu se odnosi na sljedeće vrste statističkih procesa:

### 1. Istraživanja na bazi uzorka

Ovdje je riječ o anketama temeljenim na najčešće probabilističkoj proceduri uzorkovanja, koja uključuje prikupljanje podataka direktno od ispitanika. Za ovakvu vrstu istraživanja utemeljena je teorija tačnosti koja dopušta izvještavanje o precizno definiranim komponentama tačnosti (greške uzorkovanja i greške koje nisu uvjetovane uzorkovanjem).

### 2. Popis

Ovdje je riječ o statističkom procesu u kojem su u istraživanju obuhvaćene sve jedinice iz okvira (popisne jedinice).

### 3. Statistički procesi koji koriste administrativni izvor(e)

Proces korištenja podataka prikupljenih za druge svrhe a ne direktnu proizvodnju statistike. Primjer je recimo statističko tabeliranje proizvedeno iz administrativnih baza podataka koje održava Centralna izborna komisija.

### 4. Statistički procesi koji uključuju višestruke izvore podataka

U mnogim statističkim oblastima sama proizvodnja statistika je takva da zahtjeva različite pristupe u različitim fazama statističkog procesa. Na primjer, u istraživanjima iz oblasti biznis statistika, u kojima su agregirani osnovni ekonomski podaci (proizvodnja, finansije itd.) o biznisu, različite jedinice, upitnici, uzoračke šeme i/ili druge procedure istraživanja mogu biti korištene za različite segmente istraživanja.

### 5. Cijene i ostali ekonomski indeksi

Uključivanje složenih anketnih istraživanja, često sa neprobabilističkim dizajnom, kao što su ekonomski indeksi kao posebne vrste statističkog procesa: (i) postoji specijalizirana ekonomska teorija koja definira ciljne koncepte ekonomskih indeksa; (ii) struktura grešaka indeksa uključuje specijalizirane koncepte poput prilagođavanja kvaliteta, zamjene i ponovnog uzorkovanja; (iii) uzoračke ankete koriste se u nekoliko dimenzija (težina, proizvodi, prodajna mjesta).

### 6. Statistička kompilacija

Ovaj statistički proces objedinjuje niz različitih primarnih izvora, uključujući gore navedene, kako bi stvorio agregat izvora te je od posebne konceptualne važnosti. Uglavnom je ovdje riječ o ekonomskim agregatima, kao što su nacionalni računi i bilans plaćanja.

Struktura izvještaja o kvalitetu se sastoji iz slijedećih poglavlja:

- (1) Uvod u statistički proces i statistički proizvod;
- (2) Relevantnost, procjena potreba i percepcija korisnika;
- (3) Tačnost i preciznost;
- (4) Pravovremenost i tačnost objavljivanja;
- (5) Usklađenost i uporedivost;
- (6) Dostupnost i razumljivost, diseminacijski format;
- (7) Troškovi istraživanja i opterećenost davalaca podataka;
- (8) Povjerljivost i
- (9) Statistička obrada

Popis indikatora je urađen na temelju slijedećih dokumenata: *ESS priručnik za izvještaje o kvalitetu* (ESS handbook for quality reports, 2014) i *ESS smjernice za implementaciju indikatora kvaliteta i učinka* (ESS Quidelines for the Implementation of the ESS Quality and Performance Indicators, 2014) i *Tehničkog priručnika o jedinstveno integriranoj strukturi meta podataka* (Technical Manuel of the Single Integrated Metadata Structure – SIMS, 2014).

U ovom dokumentu pružamo smjernice za pripremu izvještaja o kvalitetu u standardiziranom formatu. Tako se za svaku komponentu još jednom daje kratka definicija, za podkomponente i kratko uputstvo. U slučajevima, kada je podkomponenta vezana za izračunavanje određenog indikatora, dato je uputstvo za izračunavanje istih. Za sve indikatore kvaliteta i učinka se daju i primjeri za izračunavanje istih.

Glavni cilj izvještaja o kvalitetu je pružiti proizvođačima raznih statistika analitički uvid u cjelokupni statistički proces, a s druge strane korisnicima statističkih rezultata pružiti dodatne informacije za ispravno korištenje i tumačenje rezultata.

Glavna namjena indikatora kvaliteta je ponuditi provodiocu statističkog istraživanja i korisniku statističkih rezultata uvid u kvalitet statističkih rezultata, kvalitet procesa preko kojeg su ti rezultati dobijeni, te u određenoj mjeri uvid u kvalitet cjelokupnog institucionalnog okruženja, u kojem se istraživanje provodi.

Iako su ove smjernice orijentirane na proizvođače statistike, izvještaj o kvalitetu izrađen uz pomoć ovih smjernica uključuje takođe i sve informacije potrebne za izradu izvještaja orijentiranih na korisnika.

Indikatori se po svojoj definiciji mogu odnositi na tri različita predmeta u okviru provođenja istraživanja:

- na istraživanje kao cjelinu,
- na varijablu i
- na statistiku (statistički rezultat).

U prvom slučaju, pri provođenju konkretnog istraživanja nema dileme na koju vrijednost se indikator odnosi. U ostala dva slučaja to ne vrijedi, jer uobičajeno u jednom istraživanju mjerimo više varijabli i izračunavamo (procjenjujemo) više statistika.

Indikatori koji se odnose na varijable i statistike u izvještajima o kvalitetu se iskazuju samo za ključne varijable odnosno statistike.

Iako se indikatori u izvještajima o kvalitetu iskazuju na godišnjem (ili višegodišnjem) nivou, isti se kod istraživanja čija je periodika kraća od godine (npr. mjesečna i kvartalna istraživanja) izračunavaju za svako provedeno istraživanje. Kod ovih istraživanja se u izvještaju o kvalitetu navode sve vrijednosti u okviru posmatrane godine, a ujedno se daje i godišnji prosjek.

Tamo gdje je primjenljivo, pored osnovne vrijednosti indikatora, koja se odnosi na cijelu posmatranu populaciju, izračunavaju se vrijednosti indikatora za neke važnije domene. Tako ćemo recimo, u nekom istraživanju pored osnovne vrijednosti stope neodgovora za cijeli uzorak, prikazati takođe stopu neodgovora po važnijim domenima po kojima se prave izlazne tabele.

U nekim slučajevima, vrijednost indikatora je potrebno prikazati i grafički. Prije svega kod višegodišnjih istraživanja, gdje se (vremenske) vrijednosti indikatora iskazuju u obliku linijskih dijagrama pomoću kojih prikazujemo kretanje vrijednosti indikatora u različitim razdobljima

Izvještaj o kvalitetu izrađuju statističari odgovorni za određeno statističko istraživanje u saradnji sa kolegama zaduženim za uzorak, analizu i IT.

Izvještaj o kvalitetu bi trebao imati standardnu strukturu sa tačno raščlanjenim sadržajem. Struktura izvještaja je pripremljena sa namjerom »pokrivanja« što šireg raspona različitih istraživanja a odgovornost autora izvještaja je - da prosudi, koji dijelovi izvještaja su za dato istraživanje relevantni.

**U slučaju, da neka tačka standardnog izvještaja za određeno istraživanje nije relevantna, to je potrebno navesti, ali ne ispustiti cijelu tačku.**

U izvještaju treba navesti web adresu, koja korisnika vodi do detaljnijih informacija o prezentiranoj temi. Prije svega to važi za poglavlja o dostupnosti i razumljivosti (internetske objave).

Učestalost izrade izvještaja o kvalitetu prema smjernicama u ovom dokumentu varira ovisno od potreba, a ažuriranje istog se preporučuje poslije većih izmjena strukture podataka ili temeljnih poslovnih procesa. Godišnji izvještaji o kvalitetu su standard, time osobe zadužene za sastavljanje izvještaja ne bi bile preopterećene budući da bi se, ako se ne dogode neke značajnije promjene, materijali iz jedne godine jednostavno mogli prekopirati u drugu, a jedini novi materijal bili bi ažurirani podaci o indikatorima kvaliteta i učinka.

## 1 STATISTIČKI PROCES I STATISTIČKI PROIZVOD

U ovom uvodnom dijelu navodi se kratak opis statističkog procesa i proizvoda, prije svega - zašto i kako se provodi statističko istraživanje.

Potrebno je dati kratak pregled sljedećih informacija:

### 1.1 Namjena istraživanja

Navodi se kratak opis svrhe, cilja i predmeta statističkog istraživanja.

### 1.2 Pravni osnov i odgovornost statističkih institucija

Navodi se eksplicitno postojeći pravni osnov (zakon, program, plan, regulativa, itd.) na temelju kojeg se poduzima provođenje statističkog istraživanja. Zatim, navodi se decidno odgovornost statističkih institucija pri provođenju samog istraživanja.

### 1.3 Korištene klasifikacije

Navodi se spisak klasifikacija, nomenklatura koje su korištene prilikom provođenja istraživanja.

### 1.4 Izvještajna jedinica

Jedinica koja daje podatke isključivo u statističke svrhe, u sadržaju i rokovima određenim planom provođenja statističkih istraživanja. To mogu biti pravne osobe (preduzeća) i njeni dijelovi, fizičke osobe, obrtnici, slobodne profesije, domaćinstva, državna tijela te tijela jedinica lokalne samouprave i sve druge osobe koje daju podatke.

### 1.5 Statistička jedinica posmatranja

Osnovna jedinica na koju se podaci odnose, ultimativno prikupljaju ili izvode na drugi način (fizička osoba, pravna osoba i njezin dio, obrtnik, slobodna profesija, domaćinstvo, državno tijelo, jedinica lokalne i područne (regionalne) samouprave i druge osnovne jedinice posmatranja na koje se podaci odnose. Statistička jedinica je predmet statističkih istraživanja i nosilac statističkog obilježja.

### 1.6 Pokrivenost i obuhvat

Navode se sljedeće informacije:

- koja metoda (slučajni uzorak, popis - potpun obuhvat) je bila korištena pri izboru jedinice posmatranja;
- navodi se broj jedinica posmatranja, u slučaju uzorka, broj jedinica u okviru, te podaci o mogućoj stratifikaciji.

### 1.7 Statistički koncepti i definicije

Navode se definicije ključnih varijabli.



## 2 RELEVANTNOST

Relevantnost je stepen do kojeg statistički proizvod zadovoljava sadašnje i potencijalne potrebe korisnika. To ovisi o tome, u kojoj mjeri statistika koja je proizvedena kao i korišteni koncepti (definicije, varijable, klasifikacije, itd) odražavaju potrebe korisnika.

### 2.1 *Korisnici podataka statističkog istraživanja*

Početna faza dizajniranja i upravljanja statističkim procesima je definiranje korisničkih potreba. Procjena potreba korisnika općenito uključuje klasifikaciju korisnika, naznaku namjene za rezultate koji se žele dobiti, prioritete u zadovoljavanju njihovih potreba, načina kako su ove informacije dobivene od korisnika, npr. preko savjetodavnih odbora, redovitih saziva korisničkih grupa, povratnih informacija/pritužbi korisnika, anketnim istraživanjem.

#### 2.1.1 *Ključni korisnici podataka iz statističkog istraživanja*

Navode se ključni korisnici po glavnim grupama segmentacije: javni sektor, privredni subjekti, pravosuđe, nauka, istraživanje i obrazovanje, opća javnost, mediji, strani korisnici. Ključni korisnici se određuju na osnovi podataka o naručenim publikacijama, na osnovi članstva/pretplata, odnosno na osnovi ostalih evidencija o korisnicima podataka iz istraživanja.

#### 2.1.2 *Procjena korisničkih potreba*

Ako ne postoje ESS regulative i preporuke za određeno statističko istraživanje trebalo bi navesti (ukoliko su poznati) glavne zahtjeve korisnika za podacima iz statističkog istraživanja za koje se priprema izvještaj i opisati u koje svrhe se koriste, kao i to da li postoje dokumentovani/poznati zahtjevi za podacima koje nije bilo moguće proizvesti istraživanjem.

#### 2.1.3 *Mjerenje percepcije i zadovoljstva korisnika*

Zadovoljstvo korisnika je prioritet prvog reda. Najučinkovitija metoda mjerenja je provođenje ankete o zadovoljstvu korisnika u skladu sa najboljom praksom anketnih istraživanja koristeći reprezentativni uzorak korisnika iz odgovarajućeg okvira. Mjerenje percepcije i zadovoljstva korisnika se utvrđuje vrijednostima skale zadovoljstva korisnika. Indeksom zadovoljstva korisnika možemo mjeriti stepen zadovoljstva korisnika uslugama i proizvodima koje mu nudimo. Ostale, troškovno manje opterećene metode procjene uključuju analizu prodaje objavljenih publikacija, komentare korisnika, zaprimljene zahtjeve i pritužbe, posjete internetske stranice, itd.

## 2.2 Kompletnost podataka

### 2.2.1 *Indikator kvaliteta i učinka - Stopa kompletnosti podataka (R1)*

#### *Definicija indikatora*

Ovu stopu izračunavamo kao odnos između broja statistika (podataka) koje objavljujemo u istraživanju i broja podataka koji su propisani (zahtjevani) u odgovarajućim uredbama i regulativima.

Napomena: Ovaj indikator je primjenjiv samo ako postoji ESS propis ili smjernica o zahtjevanim izlaznim podacima (statistikama).

Vrijednost indikatora se odnosi na cjelokupno istraživanje.

#### Postupak izračunavanja

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

gdje je:

$\#D^{rqd}$  ... broj elemenata u setu zahtjevanih statistika (podataka) prema EU propisima

$\#A_D^{rqd}$  ... broj elemenata u podsetu raspoloživih statistika (podataka)

#### Objašnjenja i računski primjeri

Npr. ako određena regulativa propisuje da trebamo objaviti **5** statistika za nivo cijele BiH i za nivo statističkih regija (NUTS 2), onda je broj propisanih (zahtjevanih) statistika **5 + (5x3) = 20**.

Pretpostavimo da smo za određeno istraživanje objavili **4** statistike za nivo BiH, i **4** statistike objavljene za **2** statističke regije BiH, onda je **4 + (4x2) = 12**.

Vrijednost indikatora je:

$$R1_U = \frac{12}{20} = 0,60 = 60\%$$

### 3 TAČNOST I PRECIZNOST

Tačnost i pouzdanost procjene se definira kao stepen slaganja/bliskosti između procjenjenih/izračunatih vrijednosti (koje dobijamo na kraju statističke obrade) i ispravnih, ali nepoznatih populacijskih vrijednosti.

Najčešće je izražavamo u apsolutnom obliku (standardna greška), relativnom obliku (koeficijent varijacije) ili u smislu pouzdanosti kao interval pouzdanosti. Veličina greške uzorkovanja ovisi od korištene procjene i korištenog dizajna uzorka.

Vrijednost indikatora se odnosi na ključnu statistiku i ključne varijable.

Greške uzorkovanja treba dati u tabeli. Potrebno je navesti za koje statistike i/ili varijabe se prikazuju greške uzorkovanja. Kod mjesečnog istraživanja, daju se vrijednosti grešaka za svaki mjesec (ili kvartal) posebno, kao i prosječna godišnja vrijednost greške uzorkovanja. U slučaju, da imamo više različitih statistika te nivoa objavljivanja, navode se samo greške za glavne grupe, ostali rezultati se daju u prilogu.

U *anketnim istraživanjima* gdje je jedinica posmatranja domaćinstvo, često se greške u slučajnom uzorku prikazuju ne samo u formi koeficijenta koeficijenta varijacije, nego i preko intervala pouzdanosti.

Kod *poslovnih istraživanja*, posebno gdje npr. promet može biti i negativan (proizvodnje, prihoda, izvoza itd) najbolji način da izrazimo grešku uzorkovanja je koeficijent varijacije.

U slučaju mjesečnih istraživanja preporučujemo, da se »kretanje« grešaka uzorkovanja prikaže i u grafičkom obliku.

### 3.1 Greška uzorkovanja

Greška uzorkovanja se javlja kod istraživanja koja se temelje na slučajnom uzorku, i posljedica je činjenice da u istraživanju ne posmatramo cijelu populaciju, nego samo uzorak. U tom slučaju za ključne statistike je potrebno procijeniti greške uzorkovanja. Grešku uzorkovanja možemo prikazati na više načina.

#### 3.1.1 Indikator kvaliteta i učinka – Greška uzorkovanja (A1)

Postupak izračunavanja je određen dizajnom uzorka te korištenim estimatorom, zato nije moguće napisati opću formulu. Ovdje dajemo samo osnovne formule za slučaj procjene prosječne vrijednosti populacije kod jednostavnog slučajnog uzorka bez ponavljanja.

Pretpostavimo, da smo iz populacije od  $N$  elemenata izabrali jednostavan slučajan uzorak veličine  $n$  i želimo da na temelju uzorka procijenimo prosječnu varijablu  $Y$ .

#### Objašnjenja i računski primjeri

Vrednovanje greške uzorkovanja statističkih procjena je kompleksno, kako sa teoretskog tako i praktičnog aspekta. Jednostavne formule, važe samo pri jednostavnom slučajnom uzorkovanju, što se u praksi rijetko koristi. Kod korištenja kompleksnog dizajna uzorka ili pri upotrebi nelinearnih estimatora teoretski rezultati su mnogo zahtjevniji i ponekad neiskazivi u tačnom analitičkom obliku.

Primjer: Recimo da nakon provođenja statističkog istraživanja (kod kojeg na temelju jednostavnog slučajnog uzorkovanja bez ponavljanja procjenjujemo vrijednost aritmetičke sredine varijable  $Y$ ) imamo, podatke kakvi su prikazani u sljedećoj tabeli.

Veličina populacije ( $N$ )	Veličina uzorka ( $n$ )	Procjenjena vrijednost aritmetičke sredine populacije ( $\hat{Y}$ )	Procjena varijanse u populaciji ( $s^2$ )
10.000	500	800	50.000

Zatim izračunavamo:

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

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

**Standardna greška procjene** je jednaka kvadratnom korjenu greške uzorkovanja, tj:

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

$$se(\hat{Y}) = \sqrt{V95} = 9,75$$

**Koeficijent varijacije** se definira kao odnos između standardne devijacije i aritmetičke sredine i obično se iskazuje u procentima:

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

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

Donje i gornje granice **intervala pouzdanosti** pri 95% stepenu pouzdanosti se izračunavaju na slijedeći način:

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

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

Interval pouzdanosti je dakle: **(780,9 ; 819,1)**

Iz dobijenih rezultata zaključujemo da je prosječno odstupanje posmatrane varijable od njene prosječne vrijednosti 9,75 ili u relativnom iznosu 1,22%.

Ujedno, zbog poboljšanja preciznosti procjena izračunava se i interval pouzdanosti navedene varijable kojim se uz 95% pouzdanosti procjenjuje da se vrijednost varijable nalazi u intervalu između 780,9 i 819,1.

### 3.1.2 Aktivnosti za smanjenje grešaka uzorkovanja

Navodi se komentar za greške uzorkovanja, prije svega navesti uzroke za greške uzorkovanja, koji prelaze unaprijed određene, standardizirane granice prihvatljive greške uzorkovanja.

U slučaju, da intervali povjerljivosti nisu objavljeni, daje se uputstvo, kako se mogu izračunati iz objavljenih grešaka uzorkovanja.

U slučaju, da su procijenjene greške uzorkovanja po korištenim kriterijima previsoke, potrebno je opisati akcije za njihovo smanjenje u budućnosti.

## 3.2 Neuzoračke greške

### 3.2.1 Neuzoračke greške - Greške obuhvata

Greška obuhvata (ili greška okvira) nastaje zbog razlika između populacije koja je obuhvaćena okvirom i ciljne populacije.

Razlikujemo tri vrste grešaka obuhvata: prekomjerni obuhvat, nedovoljni obuhvat (podobuhvat), i višestruki listing (dupliciranje).

### 3.2.1.1 *Indikator kvaliteta i učinka - Stopa prekomjernog obuhvata (A2)*

#### *Definicija indikatora*

Udio jedinica (nerelevantnih) dostupnih u okviru koje ne pripadaju ciljnoj populaciji. Pod nerelevantnim jedinicama se misli na jedinice, koje smo (obično zbog pogreške ili zastarjelih podataka) uključili u okvir ili u samo istraživanje, iako iste nisu dio ciljne populacije.

Ako istraživanje provodimo na uzorku, stopa prekomjernog obuhvata procjenjuje se na temelju podataka prikupljenih na uzorku i u tom slučaju može se računati ponderisana i neponderisana stopa prekomjernog obuhvata. Ako se istraživanje zasniva na uzorku te svaka jedinica iz uzorka predstavlja određen broj jedinica u ciljnoj populaciji, logično je računati ponderisanu stopu prekomjernog obuhvata.

U slučaju da sve jedinice iz uzorka imaju jednak ponder, ponder se u postupku izračunavanja zanemaruje i tada govorimo o izračunavanju neponderisane stope prekomjernog obuhvata.

U slučaju, da provodimo periodično istraživanje (npr. mjesečno, kvartalno), potrebno je navesti vrijednosti indikatora za svaki period (npr. mjesec, kvartal) kao i prosječnu godišnju vrijednost indikatora. Vrijednosti indikatora treba dati u tabeli. Kod periodičnih istraživanja možemo dati takođe i grafični prikaz (npr. linijski dijagram) kretanja vrijednosti indikatora kroz vrijeme. Treba dati takođe objašnjenja o mogućim ekstremnim vrijednostima indikatora.

Vrijednost indikatora se odnosi na jedinice u uzorku.

#### *Postupak izračunavanja*

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

*gdje je:*

*O*... broj jedinica izvan obuhvata (ne pripadaju ciljnoj populaciji)

*E*... broj jedinica unutar obuhvata (pripadaju ciljnoj populaciji)

*Q*... broj jedinica nepoznate vrijednosti

*w<sub>j</sub>*... uzorački ponder jedinice

*α*...procjenjeni udio jedinica nepoznate relevantnosti koje su ustvari relevantne (u praksi se uglavnom pretpostavlja da iznose 1: brojnik u tom slučaju obuhvata samo jedinice izvan obuhvata)

#### *Objašnjenja i računski primjeri*

*Primjer 1:* U uzorak za Anketu o potrošnji domaćinstava izabrano je 9 400 adresa (stanova), od čega je 815 bilo praznih stanova. Na 205 adresa, stambena jedinica nije bila dostupna ili je nije bilo moguće pronaći. Primjer prikazuje postupak izračuna neponderisane stope prekomjernog obuhvata (tj. stope za adrese) pod pretpostavkom da sve jedinice imaju isti ponder.

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

Indikator računamo stavljanjem u omjer jedinica koje nisu relevantne (dakle praznih stanova) uključujući jedinice koje nisu bile dostupne, sa svim jedinicama koje su izabrane u uzorak.

Zaključujemo kako 10,85% jedinica iz uzorka ne pripada ciljanoj populaciji.

U *Primjeru2* je prikazano izračunavanje ponderisane stope prekomjernog obuhvata pod pretpostavkom da sve jedinice imaju svoj ponder, i to prema oznaci relevantnosti jedinice.

Status relevantnosti jedinice:

- 1 – relevantna jedinica koja pripada ciljanoj populaciji
- 2 - jedinica koja ne pripada ciljanoj populaciji (izvan obuhvata) – nerelevantne jedinice
- 3 – jedinice nepoznate vrijednosti (nedostupne i i jedinice koje nije moguće pronaći).

Oznaka relevantnosti jedinice	Ponder jedinice
1	6,3
3	19,0
2	3,7
1	112,3
2	115,5
3	31,2
1	8,8

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

U brojniku se nalazi zbir pondera svih jedinica koje iz određenih razloga nisu relevantne ili s njima nije bilo moguće ostvariti kontakt. Dakle, to su ponderi jedinica koje u primjeru imaju oznake 2 i 3. Nazivnik sadrži sume pondera svih jedinica iz uzorka.

Zaključujemo kako 57,1% svih jedinica iz uzorka ne pripada ciljanoj populaciji.

### 3.2.1.2 Indikator kvaliteta i učinka – Udio zajedničkih jedinica (A3)

#### Definicija indikatora

Udio jedinica obuhvaćenih iz dva različita izvora (statističko istraživanje i administrativni izvor) u odnosu na ukupan broj jedinica u istraživanju.

Indikator se koristi kada se administrativni podaci kombinuju sa podacima iz istraživanja, odnosno kada su podaci na nivou jedinica dobiveni iz oba izvora (odnosno, neke varijable dolaze iz istraživanja a druge varijable iz administrativnih podataka), ili kada podaci za dio jedinica dolaze iz istraživanja a za drugi dio jedinica iz jednog ili više administrativnih izvora.

Indikator nam pruža informaciju kompletnosti/obuhvata izvora – tj. u kojoj mjeri jedinice (varijable) egzistiraju u oba izvora (statističko istraživanje i administrativni izvor).

#### *Postupak izračunavanja*

$$Ad = \frac{\text{Broj zajedničkih jedinica iz statističkog istraživanja i administrativnog izvora}}{\text{Broj jedinstvenih jedinica u statističkom istraživanju}}$$

Zajedničke jedinice odnose se na one jedinice koje su uključene u oba izvora u podacima koji dolaze iz istraživanja i administrativnog izvora.

Indikator "jedinstvene jedinice u statističkom istraživanju" u nazivniku znači da ako jedinica egzistira u više izvora, onda se računa samo jednom.

#### *3.2.1.3 Greška nedovoljnog obuhvata*

Problem nedovoljnog obuhvata je u poređenju s prekomjernim obuhvatom teže izmjeriti, i javlja se u slučajevima, kada jedinicu nismo uključili u okvir (samim tim nismo ni u uzorak), iako po definiciji tamo pripada. O ovim jedinicama nemamo nikakvih neposrednih informacija. Zbog toga ne definiramo posebni indikator, te je potrebno dati, bilo kakvu (posrednu) informaciju, koja barem sugerira na opseg greške.

#### *3.2.1.4 Mjere za smanjenje grešaka obuhvata*

Potrebno je opisati sve postupke, koji se izvode s ciljem smanjivanja grešaka obuhvata. Pri tome je potrebno opisati kako trenutne tako i planirane aktivnosti. Problemi zbog pogrešnog obuhvata se recimo mogu riješiti na sljedeći način: (1) u posmatranoj godini zabilježimo odgovore jedinica, koje dobijemo s terena, i dodijelimo im status izvještajne jedinice, a zatim kod istraživanja za narednu godinu prihvatimo te nove statuse; (2) prije slanja upitnika na teren, pripremimo statistički spisak (adresar) npr. visokoškolskih ustanova i uporedimo isti sa spiskom koji održava Ministarstvo obrazovanja. Ako se javi nesklad između ovih adresara, statistički adresar popravimo (dopunimo).

#### *3.2.2 Neuzoračke greške - Greške mjerenja*

Greške mjerenja su greške koje se javljaju prilikom prikupljanja podataka i dovode do toga da se zabilježene vrijednosti varijabli razlikuju od stvarnih vrijednosti.

Uzroci za ove razlike su sljedeći:

- (1) kod instrumenata istraživanja: obrazac koji se koristi za prikupljanje podataka može dovesti do bilježenja pogrešnih vrijednosti;
- (2) kod ispitanika/izvještajne jedinice: ispitanici mogu svjesno ili nesvjesno davati pogrešne podatke;
- (3) kod anketara: anketar može utjecati na odgovore koje daje ispitanik/izvještajna jedinica.

### 3.2.2.1 Razlozi za nastanak grešaka mjerenja

Potrebno je navesti glavne razloge, koji prouzrokuju nastanak grešaka u mjerenju. Ako postoji povratna informacija od strane izvještajnih jedinica, potrebno je prije svega navesti, koji su glavni razlozi (po njihovom mišljenju) za greške u mjerenju. U nekim istraživanjima (posebno koja se provode u domaćinstvima) treba navesti napomene anketara (ako postoje).

Najčešći razlozi za nastanak grešaka u mjerenju su:

- osoba, koja ispunjava upitnik, nije dovoljno stručno osposobljena za ispunjavanje;
- osoba, koja ispunjava upitnik, nije dovoljno pažljivo pročitala upute za ispunjavanje;
- nedovoljna pažnja osobe koja unosi podatke u tabele;
- nedostatak evidencija kod izvještajnih jedinica;
- preobimna ili oskudna metodološka uputstva za ispunjavanje upitnika i
- upitnik neprecizan i uputstvo nepotpuno.

### 3.2.2.2 Mjere za smanjenje broja grešaka mjerenja

Postupak uređivanja/editovanja podataka identificira nedosljednosti. One su obično posljedica grešaka u originalnim podacima, ali mogu takođe biti rezultat procesnih grešaka kod šifriranja i unosa podataka. Potrebno je ukratko opisati postupak u slučaju otkrivanja grešaka. Navesti prije svega, da li jedinicu ponovo anketiramo ili se podatak ispravlja ručno ili pak koristimo postupak automatskog uređivanja. Potrebno je navesti sve aktivnosti, koje se poduzimaju s ciljem smanjenja grešaka mjerenja. Najučinkovitiji način za smanjenje broja grešaka u mjerenju su precizna i razumljiva metodološka objašnjenja, koja ne smiju biti preobimna i svakako uspostaviti direktan kontakt s izvještajnim jedinicama

### 3.2.3 Neuzoračke greške - Greške neodgovora

Razlika između statistika izračunatih iz prikupljenih podataka i onih koje bi bile izračunate da ne postoje nedostajuće vrijednosti predstavlja grešku neodgovora.

Postoje dvije vrste neodgovora:

- (1) neodgovor izvještajne jedinice, koji se javlja kada nisu prikupljeni podaci o izvještajnoj jedinici dizajniranoj za prikupljanje podataka, i
- (2) neodgovor za određene varijable, koji se javlja kada se prikupe podaci samo za neke - a ne za sve varijable koje su zahtjevane određenim istraživanjem.

#### 3.2.3.1 Indikator kvaliteta i učinka - Stopa neodgovora jedinica (A4)

#### Definicija indikatora

Odnos između broja jedinica posmatranja za koje nismo prikupili podatke (barem za neke varijable) prema ukupnom broju jedinica (dizajniranih - određenih za prikupljanje podataka). Nekontaktirane jedinice posmatranja, za koje ne znamo, kakav je njihov status (da li su relevantne ili nerelevantne), takođe smatramo neodgovorom.



Ovisno o tome računamo li pokazatelje za uzorak ili cijelu populaciju, može se izračunati ponderisana i neponderisana stopa neodgovora jedinica.

U slučaju, da obavljamo periodično istraživanje (npr. mjesečno, kvartalno), potrebno je navesti vrijednosti indikatora za svaki period (npr. mjesec, kvartal) kao i prosječnu godišnju vrijednost indikatora.

Vrijednost indikatora se odnosi na cijelo istraživanje. Vrijednosti indikatora treba dati u tabeli.

### *Postupak izračunavanja*

Po završenom prikupljanju podataka jedinice uzorka možemo podijeliti na četiri grupe:

- Nerelevantne jedinice (NJ). Jedinice koje su inače bile izabrane u uzorak, ali se u fazi prikupljanja podataka pokazalo da nisu više dio ciljne populacije koju posmatramo. Najčešći razlog što su sve ove jedinice dio uzorka je neažurnost i nepotpunost izvora koje koristimo za utvrđivanje okvira uzorka.
- Neodgovori (NR). Jedinice koje su relevantne za istraživanje, ali nismo od njih uspjeli dobiti željene podatke. U ovoj grupi su najčešće jedinice koje odbijaju sudjelovanje u istraživanju.
- Jedinice sa nepoznatom valjanošću (Q). Jedinice za koje nismo uspjeli dobiti željene informacije i ne znamo, da li su relevantne za istraživanje ili ne. Ovo važi uglavnom za jedinice sa kojima nije moguće uspostaviti kontakt.
- Odgovori (R). Jedinice za koje smo uspjeli dobiti sve željene informacije. Pri tom je potrebno odrediti (za pojedino istraživanje) koja je najmanja količina potrebnih informacija koje jedinica mora dati da bismo je mogli uvrstiti u odgovore.

### Neponderisana stopa

U odnosu na definirane kategorije neponderiranu stopu neodgovora jedinice izračunavamo po formuli:

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

gdje je:

*R*... broj relevantnih jedinica koje su odgovorile

*NR*... broj relevantnih jedinica koje nisu odgovorile

*Q*... broj odabranih jedinica nepoznate relevantnosti

*α*... (procjenjen) udio jedinica nepoznate relevantnosti a koje su zapravo relevantne.

Ako nema jako utemeljenih pretpostavki za procjenu parametra **α**, uzimamo da je **α = 1**, što znači da sve jedinice sa nepoznatom relevantnošću smatramo neodgovorom.

### Ponderisana stopa

Ponderisanu stopu izračunavamo prema formuli:

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

gdje je:

R... broj relevantnih jedinica koje su odgovorile

NR... broj relevantnih jedinica koje nisu odgovorile

Q... broj odabranih jedinica nepoznate relevantnosti

w<sub>j</sub>... uzorački ponder jedinice

α... procjenjeni udio jedinica nepoznate relevantnosti koje su zapravo relevantne (ako nema jako utemeljenih pretpostavki za procjenu parametra uglavnom se pretpostavlja da iznose 1)

Ovaj pristup bi trebao biti korišten u anketama osoba i domaćinstava.

U poslovnim i poljoprivrednim istraživanjima je njihov uticaj na konačne rezultate mnogo važniji od drugih, tako da u tom slučaju kao pomoćnu varijablu moramo uzeti u obzir varijablu koja određuje veličinu jedinice (npr. broj zaposlenih u preduzeću).

### Objašnjenja i računski primjeri

Navodimo hipotetički primjer istraživanja u kojem smo nakon konačnog postupka prikupljanja podataka dobili slijedeće podatke:

	Broj jedinica posmatranja	Zbir uzoračkih pondera	Zbir vrijednosti pomoćne varijable jedinica posmatranja	Ponderisani zbir pomoćne varijable
Odgovori (R)	700	5 500	15 600	62 500
Neodgovori (NR)	250	1 520	2 900	6 200
Nepoznata relevantnost (Q)	80	550	1 400	2 850
Uzorak	1 030	7 570	19 900	71 550

Ako uzmemo da je pretpostavljena vrijednost udjela valjanih jedinica među jedinicama nepoznate relevantnosti jednaka α = 0,90 - vrijednost indikatora je:

### Neponderisana stopa neodgovora jedinica

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

*Ponderisana stopa neodgovora jedinice*

- sa ponderima izbora uzorka  $NRr_w = 1 - \frac{5500}{5500+1520+0,90 \cdot 550} = 1 - 0,73 = 0,27 = 27\%$

- samo pomoćna varijabla  $NRr_w = 1 - \frac{15600}{15600+2900+0,90 \cdot 1400} = 1 - 0,79 = 0,21 = 21\%$

- sa korekcijom pondera za pomoćnu varijablu

$$NRr_w = 1 - \frac{62500}{62500+6200+0,90 \cdot 2850} = 1 - 0,88 = 0,12 = 12\%$$

**3.2.3.2 Indikator kvaliteta i učinka - Stopa neodgovora varijable (A5)***Opis indikatora:*

Odnos između broja jedinica posmatranja za koje nismo uspjeli dobiti podatak za datu varijablu i broja svih jedinica koje su trebale dati podatak za posmatranu varijablu. Ovaj odnos se izračunava samo u okviru jedinica posmatranja, koje su relevantne za posmatranu varijablu.

Vrijednost indikatora se odnosi isključivo na ključnu varijablu.

*Postupak izračunavanja*

Kao i kod stope neodgovora jedinica, može se računati neponderisana stopa, ponderisana stopa sa ponderima izbora uzorka ili ponderisana stopa sa ispravljenim ponderima sa vrijednostima pomoćne varijable.

$$NR_{Yr_w} = 1 - \frac{\sum_{R_Y} w_j}{\sum_{R_Y} w_j + \sum_{NR_Y} w_j}$$

gdje je:

$R_Y$ ... broj relevantnih jedinica koje su odgovorile na varijablu Y

$NR_Y$  ...broj relevantnih jedinica koje nisu odgovorile na varijablu Y iako se odgovor za istu zahtjeva

$w_j$ ... ponder jedinice

Tri osnovna načina izračunavanja stope su:

- Neponderisana stopa:  $w_j = 1$
- Stopa ponderisana ponderima dizajna uzorka:  $w_j = d_j$
- Stopa ponderisana značajnošću jedinica:  $w_j = d_j x_j$ , gdje je  $x_j$  vrijednost pomoćne varijable X

### Objašnjenja i računski primjeri

Za izabranu varijablu relevantne jedinice su one jedinice za koje bismo morali dobiti podatak za tu varijablu. Nije potrebno, da su sve relevantne jedinice relevantne takođe i za izabranu varijablu.

Razlika između broja relevantnih jedinica i broja jedinica, relevantnih takođe za izabranu varijablu je najviše povezana sa tzv. preskocima u upitniku. Prema tome, jedinice kod kojih dolazi do preskoka pitanja ne ulaze u izračunavanje stope neodgovora na određenu varijablu koju smo preskočili.

#### Primjer:

Ako nakon pitanja »Da li vaše domaćinstvo ima pristup internetu?« slijedi pitanje, »Da li domaćinstvo pristupa internetu preko PC«, za odgovaranje na drugo pitanje relevantne su samo jedinice, kod kojih je odgovor na prvo pitanje bilo »DA«.

U tabeli su date pretpostavljene vrijednosti za izabranu ključnu varijablu:

	Broj jedinica posmatranja	Zbir konačnih pondera	Zbir vrijednosti pomoćne varijable posmatranih jedinica	Ponderisani zbir pomoćne varijable
Odgovori jedinice	700	5 500	15 600	70 000
Relevantne jedinice od kojih se očekuje odgovor na ključnu varijablu	450	3 800	11 500	55 000
Odgovori relevantnih jedinica na ključnu varijablu	370	3 000	10 500	52 000

#### Neponderisana stopa neodgovora na ključnu varijablu

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

#### Ponderisana stopa neodgovora na ključnu varijablu

- sa ponderima dizajna uzorka  $NR_{Yr_w} = 1 - \frac{3000}{3800} = 0,211 = 21,1\%$

- samo pomoćna varijabla  $NR_{Yr_w} = 1 - \frac{10500}{11500} = 0,087 = 8,7\%$

- sa korekcijom pondera za pomoćnu varijablu

$$NR_{Yr_w} = 1 - \frac{52000}{55000} = 0,055 = 5,5\%$$

U slučaju da obavljamo periodično istraživanje (npr. mjesečno, kvartalno), potrebno je navesti vrijednosti indikatora za svaki period (npr. mjesec, kvartal) kao i prosječnu godišnju vrijednost indikatora. Vrijednosti indikatora treba dati u tabeli. Možemo takođe dati grafički prikaz (npr. linijski dijagram) kretanja vrijednosti indikatora u vremenu.

### 3.2.3.3 Postupci u slučaju neodgovora

Potrebno je opisati sve postupke (npr. ponderisanje, imputacije), koje smo koristili zbog neodgovora, kako za slučaj neodgovora jedinice posmatranja tako i za slučaj neodgovora varijable. Ako se koristi postupak ponderisanja, potrebno je navesti formulu za izračunavanje pondera.

Na primjer, ako nedostaju podaci o određenim varijablama, uspostavlja se telefonska veza s izvještajnom jedinicom i po potrebi dopunjavaju nedostajuće vrijednosti. Izuzetno, ako nije uspostavljen kontakt sa izvještajnom jedinicom, vrijednost varijable se može procijeniti.

### 3.2.3.4 Postupci za smanjenje stope neodgovora

Potrebno je opisati sve postupke, koje izvodimo s ciljem smanjivanja stope neodgovora jedinice posmatranja i varijable. Kod uputstva npr. za izračunavanje ekvivalenta punog radnog vremena, dodamo i neke primjere za rješavanje najčešćih kombinacija pri izračunavanju ovog indikatora. Takođe, potrebno je poslati izvještajnim jedinicama obavijest da nam se obrate telefonski ili elektronski, ako pri ispunjavanju upitnika imaju određenih poteškoća i nejasnoća.

## 3.3 Revizije

Revizija može biti planirana i neplanirana. Neplaniranu reviziju uzrokuje uglavnom otkrivanje grešaka u publikovanim rezultatima. ESS Code of Practice zahtjeva da planirane revizije slijede standardne, dobro uspostavljene i transparentne procedure. To znači, da su unaprijed najavljene revizije poželjne i da treba navesti razloge za poduzimanje i prirodu same revizije (novi izvor podataka, nove metode, itd...).

### 3.3.1 Indikator kvaliteta i učinka - Prosječna veličina revizije podataka (A6)

#### Opis indikatora:

Revizija se definira kao razlika između kasnije i ranije procjene ključne varijable. Prosječna veličina revizije podataka je prosjek revizija ključnih varijabli tokom određenog perioda.

	Posmatrani period				
	1	...	$t$	...	$n$
Objavlivanje					
1 objava	$X_{11}$	...	$X_{1t}$	...	$X_{1n}$
...	...	...	...	...	...
$k$ -ta objava	$X_{k1}$	...	$X_{kt}$	...	$X_{kn}$
...	...	...	...	...	...
$K$ -ta i konačna objava	$X_{K1}$	...	$X_{Kt}$	...	$X_{Kn}$

### Postupak izračunavanja

Obzirom na dvodimenzionalnu situaciju koja je opisana u definiciji, postoji više strategija za izračunavanje indikatora. Predlaže se uzimanje prosjeka za određena objavljivanja za  $n$  posmatranih perioda, odnosno MAR (prosječna apsolutna revizija).

MAR (prosječna apsolutna revizija):

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

gdje je:

$X_{Lt}$  - „kasnije“ objavljivanje,

$X_{Pt}$  - „prijašnje“ objavljivanje

$n$  - broj revizija

Prosječna apsolutna revizija uglavnom se primjenjuje kod indeksa, proporcija i ostalih relativnih podataka. Indikator nije preporučljiv za godišnje procjene.

AR (apsolutna revizija):

Reviziju podataka možemo računati i kao razliku prvog i posljednjeg objavljivanja.

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

gdje je:

$X_{Lt}$  - „kasnije“ objavljivanje,

$X_{Pt}$  - „prijašnje“ objavljivanje

Broj revizija ( $n$ ) se računa kao broj razlika u objavljivanjima u vremenskim serijama. U principu, to je broj objavljivanja umanjen za 1, tj. ako imamo četiri objavljivanja  $n = 4 - 1 = 3$ .

Prosječna relativna revizija (RMAR):

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

Prosječna relativna revizija se primjenjuje za podatke u apsolutnim iznosima (npr. broj zaposlenih).

### Objašnjenja i računski primjeri

Podaci koji se objavljuju, a odnose se na određeni posmatrani period  $t$ , tokom vremena se mogu mijenjati. Broj posmatranih perioda za koje se podaci objavljuju označava se sa  $n$ , a broj objava tokom vremena sa  $K$ . Revizija podataka nastaje između pojedinih objava vezano uz određeni posmatrani period, te se može definirati kao razlika između kasnije i ranije procjene.

Primjer1: Provodi se istraživanje u kojem se mjesečni indeks u posmatranom periodu revidira u šest uzastopnih mjeseci te je izračunata revizija podataka. Kako je riječ o indeksima izračunava se prosječna apsolutna revizija (MAR).

Ako imamo slijedeće vrijednosti mjesečnog indeksa ukupnog broja zaposlenih (od prvog do šestog objavljivanja):

	1. objav.	2. objav.	3. objav.	4. objav.	5. objav.	6. objav.
Vrijednost indeksa	98,20	98,40	98,20	98,30	98,10	98,00
Razlika	-	0,20	0,20	0,10	0,20	0,10

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

Zaključak je da se mjesečni indeks ukupnog broja zaposlenih u ovih šest objavljivanja u prosjeku promijenio za 0,16.

**Primjer2:** Provodi se istraživanje u kojem se ukupan mjesečni broj zaposlenih u posmatranom periodu revidira u dvanaest uzastopnih mjeseci. Kako je indikator izražen u apsolutnim vrijednostima računa se prosječna relativna revizija (RMAR). Indikator se prvo izračuna za svaki mjesec - a zatim i za cijelu godinu (kao prosjek svih prosječnih relativnih revizija po mjesecima).

Mjesec	Prvi rezultati	Konačni rezultati	Prosječna relativna revizija (%)
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
<b>2014.</b>	-	-	<b>0,20</b>

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

Zaključak je da se ukupan broj zaposlenih u ovih dvanaest objavljivanja u godišnjem prosjeku promijenio za 0,20%.

### 3.4 Imputacija

Imputacija je odgovor na nedostatke u dobivenim podacima. U statističkim istraživanjima na bazi uzorka ili popisa razlog za imputaciju može biti neodziv (uobičajeno neodziv varijable), dok recimo u obradi indeksa cijena, do imputacije može doći iz razloga nedostajućih cijena.

#### 3.4.1 Indikator kvaliteta i učinka - Stopa imputiranih podataka (A7)

##### Opis indikatora:

Indikator se izračunava za ključne varijable i definira kao *odnos između broja jedinica, za koje smo imputirali podatke za posmatrane ključne varijable (da li zbog nedostajućih ili zbog neadekvatnih vrijednosti) te broja svih jedinica za koje imamo bilo koji podatak*.

Stopa imputiranih podataka je indikator, koji se izračunava u slučaju, kada dio vrijednosti za ključne varijable procjenimo (imputiramo) s jednom od uobičajenih metoda imputacije. Podatak možemo imputirati zbog nedostajućih vrijednosti ili zbog neadekvatnih vrijednosti, do kojih smo došli u toku uređivanja podataka.

Ovaj pokazatelj je pod uticajem kako neodgovora varijable tako i procesa editovanja, i mjeri kako relativni iznos imputirane vrijednosti tako i relativni uticaj na konačne procjene iz postupaka imputacije.

Takođe, ovdje se može izračunati kako ponderisana - tako i neponderisana vrijednost indikatora.

Vrijednost indikatora se odnosi na ključnu varijablu. Vrijednosti indikatora treba dati u tabeli.

U slučaju, da provodimo periodično istraživanje (npr. mjesečno, kvartalno), potrebno je navesti vrijednosti indikatora za svaki period (npr. mjesec, kvartal) kao i prosječnu godišnju vrijednost indikatora. Kod periodičnih istraživanja možemo dati takođe i grafični prikaz (npr. linijski dijagram) kretanja vrijednosti indikatora kroz vrijeme.

##### Postupak izračunavanja

Neponderisanu vrijednost indikatora izračunavamo na slijedeći način:

$$IR_Yr = \frac{\sum I_Y}{\sum I_Y + \sum K_Y}$$

gdje je:

$I_Y$ ... broj jedinica za koje je varijabla Y imputirana,

$K_Y$ ... broj jedinica za koje je vrijednost varijable Y ostala nepromijenjena

Ako je  $w_j$  finalni ponder, ponderisanu vrijednost izračunavamo na slijedeći način:

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



### Objašnjenja i računski primjeri

Kod izračunavanja stope imputiranih podataka moramo uzeti u obzir samo one vrijednosti, koje su bile imputirane sa relevantnom statističkom (imputacijskom) metodom.

Jedinice, koje su bile u postupku uređenja podataka popravljene sa ponovnim provjeravanjem kod izvještajnih jedinica, ne smatramo jedinicama sa imputiranim podacima.

U tabeli su dati podaci za naš primjer:

	Broj jedinica posmatranja	Ponderisani zbir varijable Y
Podaci za varijablu Y	500	12500
Imputirani podaci	150	1750

Neponderisana vrijednost indikatora  $IR_Y r = \frac{150}{500} = 0,30$

Ponderisana vrijednost indikatora  $IR_Y r_W = \frac{1750}{12500} = 0,14$

## 4 PRAVOVREMENOST I TAČNOST OBJAVLJIVANJA

Pravovremenost objavljivanja podataka odnosi se na vremenski razmak između posljednjeg dana referentnog perioda na koji se podaci i datuma objavljivanja.

Tačnost objavljivanja je vremenski razmak između stvarnog datuma objavljivanja podataka i najavljenog datuma koji je naznačen u zvaničnom kalendaru publikovanja.

### 4.1 Pravovremenost objavljivanja

Pojedine statistike se objavljuju u nekoliko verzija (npr. preliminarna, revidirana i konačna). U tom slučaju svako izdanje ima svoj profil pravovremenosti.

Potrebno je posebno objasniti i navesti razloge za moguće kašnjenje u publiciranju, kao i navesti napore koji su poduzeti za poboljšanje situacije.

#### 4.1.1 Indikator kvaliteta i učinka - Pravovremenost prvih rezultata (TP1)

##### Opis indikatora

Pravovremenost prvih rezultata je vremenski razmak između kraja perioda za koji su podaci izračunati i tačnog datuma prvog objavljivanja rezultata.

Podatke o pravovremenosti prvog objavljivanja rezultata treba prikazati u tabeli.

##### Postupak izračunavanja

Ako datum koji označava posljednji dan (datum) perioda posmatranja statističkih podataka označimo sa  $d_{refp}$ , a datum objavljivanja prvih rezultata sa  $d_{frst}$ , vrijednost indikatora je:

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

### Objašnjenja i računski primjeri

Vrijednost indikatora kod mjesečnih, tromjesečnih i polugodišnjih istraživanja se iskazuje u ukupnom broju dana, dok se kod godišnjih i višegodišnjih istraživanja vrijednost indikatora iskazuje u broju mjeseci. U svakom slučaju kod iskazivanja vrijednosti indikatora treba jasno navesti korištenu vremensku jedinicu.

**Primjer 1:** Ako su rezultati mjesečnog istraživanja koji se odnose na januar 2014. godine prvi put objavljeni 17.03.2014., vrijednost indikatora je jednaka:  $T_1 = (17.3.2014) - (31.1.2014) = T+45$ .

**Primjer 2:** Ako su rezultati APD (Anketa o potrošnji domaćinstava) koji se odnose na 2014. godinu prvi put objavljeni 20.11.2015., vrijednost indikatora:  $T_1 = (20.11.2015) - (31.12.2014) = T+11$ .

U izvještaju o kvalitetu treba navesti: period posmatranja; datum publikovanja prvih rezultata; vremenski razmak u formatu  $T+x$ , gdje je  $T$  kraj perioda posmatranja, a  $x$  broj dana (ili mjeseci).

#### 4.1.2 Indikator kvaliteta i učinka - Pravovremenost konačnih rezultata (TP2)

##### Opis indikatora

Pravovremenost konačnih statističkih rezultata je vremenski razmak između kraja perioda posmatranja na koji se objavljeni rezultati odnose i datuma objave konačnih rezultata.

Podatke o pravovremenosti konačnih rezultata treba prikazati u tabelarnom formatu.

##### Postupak izračunavanja

Ako datum koji označava posljednji dan (datum) perioda posmatranja statističkih podataka označimo sa  $d_{refp}$ , a datum objavljivanja konačnih rezultata sa  $d_{finl}$ , vrijednost indikatora je:

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

##### Objašnjenja i računski primjeri

Vrijednost indikatora kod mjesečnih, tromjesečnih i polugodišnjih istraživanja se iskazuje u ukupnom broju dana, dok se kod godišnjih i višegodišnjih istraživanja se kao jedinica iskazuje u broju mjeseci. U svakom slučaju treba kod iskazivanja vrijednosti indikatora treba jasno navesti korištenu vremensku jedinicu.

**Primjer 1:** Ako su konačni rezultati mjesečnog istraživanja koji se odnose na januar 2014. godine objavljeni 25.03.2015., vrijednost indikatora je jednaka:  $T_2 = (25.03.2015) - (31.1.2014) = T+53$ .

**Primjer 2:** Ako su konačni rezultati APD (Anketa o potrošnji domaćinstava) koje se odnosi na 2014. godinu objavljeni 15.02.2016., vrijednost indikatora:  $T_1 = (15.02.2016) - (31.12.2014) = T+13$ .

U izvještaju o kvalitetu treba navesti: period posmatranja; datum publikovanja konačnih rezultata; vremenski razmak u formatu  $T+x$ , gdje je  $T$  kraj perioda posmatranja, a  $x$  broj dana (ili mjeseci).

**Ako se objavljuju samo jedni (konačni) rezultati istraživanja, indikator TP1 se ne izračunava, te se pod ovom tačkom daje samo napomena o tome.**

## 4.2 Tačnost objavljivanja

### 4.2.1 Indikator kvaliteta i učinka – Tačnost objavljivanja (TP3)

#### Opis indikatora

Tačnost objavljivanja je vremenski razmak (broj dana) između datuma objavljivanja podataka i najavljenog/planiranog datuma za objavljivanje podataka. Najavljeni datum se odnosi na planirani datum u godišnjem Kalendaru publikovanja.

Indikator se računa na dva načina.

#### a) Postupak izračunavanja indikatora za proizvođače:

Ako je predviđeni datum objavljivanja statistika jednak  $d_{sch}$ , a stvarni datum objavljivanja je  $d_{act}$ , vrijednost indikatora je:  $P_3 = d_{act} - d_{sch}$

#### Objašnjenja i računski primjeri

Vrijednost indikatora se izračunava i iskazuje u broju dana. Ako su rezultati za određenu statistiku stvarno objavljeni i prije predviđenog datuma objavljivanja rezultata, vrijednost indikatora može takođe biti i negativna.

**Primjer 1:** Ako je predviđeni datum objavljivanja 16.3.2015. godine, a rezultati su bili objavljeni 20.3.2015., vrijednost indikatora jednaka je  $P_3 = (20.3.2015) - (16.3.2015) = 4$ .

**Primjer 2:** Ako je predviđeni datum objavljivanja 16.3.2015. godine, a rezultati su bili objavljeni 13.3.2015., vrijednost indikatora jednaka je  $P_3 = (13.3.2015) - (16.3.2015) = -3$ .

U izvještaju o kvalitetu treba navesti: najavljeni/planirani datum objavljivanja rezultata; stvarni datum objavljivanja rezultata; vremensko kašnjenje u formatu  $T+x$ , gdje je  $T$  najavljeni datum objavljivanja, a  $x$  broj dana.

#### a) Postupak izračunavanja indikatora za korisnike:

Izračunava se Stopa tačnosti objavljivanja podataka ( $P_{3R}$ ) relevantnih za grupu statistika/rezultata.

Ako broj statistika/rezultata objavljenih na predviđeni datum objavljivanja u kalendaru ili objavljenih ranije označimo sa  $m_{pc}$ , a broj statistika/rezultata koji nisu objavljeni do datuma najavljenig u kalendaru objavljivanja sa  $m_{up}$ , vrijednost indikatora je:

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

#### Objašnjenja i računski primjeri

Vrijednost indikatora se izračunava i iskazuje u postotku.

*Primjer:* Ako je 8 statistika/rezultata publikovano na/ili prije predviđenog datuma objavljivanja, a prema Kalendaru publikovanja planirana je objava ukupno 12 statistika/rezultata, vrijednost indikatora jednaka je:

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

#### 4.3 Razlozi za veća kašnjenja i mjere za poboljšanje pravovremenosti i tačnosti

Potrebno je navesti razloge i dati dodatna objašnjenja u slučaju velikih vrijednosti vremenskih pomaka za vrijednosti indikatora TP1, TP2 i TP3. Ako isti postoje, treba navesti mjere preduzete s ciljem poboljšanja pravovremenosti i tačnosti objavljivanja vremenskih rokova iz kalendara publikovanja.

## 5 USKLAĐENOST I UPOREDIVOST

*Usklađenost* dva ili više statističkih (podataka) proizvoda odnosi se na stepen do kojeg su statistički procesi po kojem su isti proizvedeni koristili iste koncepte, klasifikacije, definicije i ciljnu populaciju i harmonizirane metode.

*Uporedivost* se odnosi na potrebu da dobiveni podaci i informacije budu uporedive u različitim periodima posmatranja, između različitih domena te između različitih geografskih područja.

### 5.1 Usklađenost

#### 5.1.1 Indikator kvaliteta i učinka - *Skladnost sa rezultatima iz referentnog izvora (CH1)*

##### *Opis indikatora:*

Skladnost sa rezultatima iz referentnog izvora pokazuje odnos između statističkih rezultata koji su proizvedeni istraživanjima i statističkih rezultata iz drugih izvora (npr. administrativnih izvora, kratkoročnih i strukturnih poslovnih pokazatelja, nacionalnih računa, itd).

Ako je riječ o apsolutnim iznosima, skladnost se prikazuje u relativnom obliku, a ako je riječ o relativnim podacima (tj. indeksi, postoci), skladnost se izražava u apsolutnom obliku.

### Postupak izračunavanja

Vrijednost pokazatelja se za ključne varijable, gdje je to moguće (ako imamo uporedivu varijablu iz drugog izvora), izračunava po formuli (može se iskazati u relativnom ili apsolutnom obliku, u zavisnosti od vrste podataka):

$$\text{u relativnom obliku: } CH1 = \frac{X_{iref} - X_i}{X_i}$$

$$\text{u apsolutnom obliku: } CH1 = X_{iref} - X_i$$

gdje je:

$X_{iref}$ ... vrijednost varijable u posmatranom statističkom istraživanju

$X_i$  ... vrijednost varijable iz drugog izvora

$i$  ... broj izvora podataka

### Objašnjenja i računski primjeri

#### Primjer1: Skladnost između različitih izvora podataka – relativna razlika

Ako je broj zaposlenih dobijen kao rezultat u anketi o radnoj snazi iznosio 816036, a podatak za istu ključnu varijablu u drugom referentnom istraživanju iznosio 693941, vrijednost indikatora je:

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

Zaključak je da je broj zaposlenih koji je rezultat anketnog istraživanja veći za 18% od broja zaposlenih iz drugog izvora.

#### Primjer2: Skladnost između različitih izvora podataka – apsolutna razlika

Ako je kao rezultat u anketi o radnoj snazi, postotak (udio) osoba sa završenom srednjom iznosio 44,5 %, a podatak za istu ključnu varijablu iz drugog izvora (u ovom slučaju, ankete o potrošnji domaćinstava) iznosio 41,6 %, vrijednost indikatora je:

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

#### 5.1.2 Razlozi za veća odstupanja

U slučaju većih odstupanja u rezultatima istraživanja, potrebno je navesti razloge.

## 5.2 Uporedivost

### 5.2.1 Indikator kvaliteta i učinka – Nepodudarnost uporedivih statistika (CC1)

#### Opis indikatora:

U domenima gdje postoji uporediva statistika moguće je procijeniti geografsku uporedivost mjerenjem neslaganja između odlaznih i ulaznih tokova za odabrane parove zemalja. Klasičan primjer ove statistike je vanjska trgovina. Načelno, izvoz zemlje A u zemlju B kroz određeni period treba biti jednak uvozu zemlje B iz zemlje A. Pored vanjske trgovine, primjer za ovu statistiku može biti i statistika migracija, turizma, FATS, bilans plaćanja, itd).

Bilateralna uporediva statistika predstavlja razliku izlaznih i ulaznih tokova između dvije zemlje podijeljena prosjekom ove dvije vrijednosti.

Primjer za zemlje A i B

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

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

Zajednička mjera se može dobiti iz dvije razlike u odnosu na prosjek tokova (jedna od mogućnosti je prikazana u slijedećoj formuli):

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

gdje je:

$OF_{AB}$  ... odlazni tok (izvoz) koji ide iz zemlje A u zemlju B

$mIF_{AB}$ ... uporedivi dolazni tok za zemlju A

$OF_{BA}$  ... odlazni tok (izvoz) koji ide iz zemlje B u zemlju A

$mIF_{BA}$ ... uporedivi dolazni tok za zemlju B

#### Objašnjenja i računski primjeri

Primjer: Uzmimo podatke vanjskotrgovinske razmjene između BIH i HRV za 2014. godinu. Zemlja A je Bosna i Hercegovina, a zemlja B je Hrvatska.

*Podaci o vanjskotrgovinskoj razmjeni sa HRV (objavljeni u statistici BIH) su slijedeći:*

IZVOZ ( $OF_{AB}$ ) - 488.308 (u 000 €)

UVOZ ( $mIF_{BA}$ ) - 946.756 (u 000 €)

Podaci o vanjskotrgovinskoj razmjeni sa BIH (objavljeni u statistici HRV) su slijedeći:

IZVOZ ( $OF_{BA}$ ) - 1.223.664 (u 000 €)

UVOZ ( $mIF_{AB}$ ) - 460.275 (u 000 €)

Na osnovu datih podataka izračunavamo nepodudarnost između ovih uporedivih statistika:

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

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

Kada bi koeficijent bio jednak nuli postojala bi perfektna simetrija (podudarnost) podataka (odnosno odlazni tokovi bi u tom slučaju bili jednaki uporedivim dolaznim tokovima).

Međutim u našem slučaju vidimo da je koeficijent asimetrije (nepodudarnosti) 0,20 – odnosno zaključujemo da nepodudarnost ovih statistika iznosi 20%.

### 5.2.2. Indikator kvaliteta i učinka - Dužina uporedivih vremenskih serija (CC2)

#### Opis indikatora

Dužina uporedive vremenske serije od zadnjeg prekida vremenske serije odnosno broj posmatranih perioda u vremenskoj seriji od zadnjeg prekida.

Vrijednost indikatora se odnosi na ključne statistike istraživanja.

#### Postupak izračunavanja

Ako je  $J_{last}$  redni broj zadnjeg perioda posmatranja za koji su objavljeni rezultati (zadnja tačka u vremenskoj seriji),  $J_{first}$  je redni broj prvog perioda posmatranja sa već uporedivim statističkim rezultatima (prva vremenska tačka prema mogućem prekidu), vrijednost indikatora je:

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

#### Objašnjenja i računski primjeri

Vrijednost indikatora pokazuje broj perioda u vremenskoj seriji; to znači, da je mjerna jedinica ovisna od periodike istraživanja. Kod mjesečnog istraživanja vrijednost indikatora se izražava kao broj mjeseci, kod kvartalnog kao broj kvartala itd.

Vrijednost indikatora je u velikoj mjeri određena time, kako definiramo prekid u vremenskoj seriji. Činjenica je, da je nemoguće postaviti posve tačna mjerila s kojima bi opazili prekid, što znači, da obično treba – barem djelomično – koristiti subjektivnu procjenu. Uopšteno bi se moglo reći da prekid u vremenskoj seriji nastaje onda kada dođe do takve promjene jednog ili više aspekata istraživanja (npr. u izvoru podataka ili korištenoj metodologiji), da konačni rezultati među sobom nisu više razumno uporedivi.

Ako vremenska serija nema prekida, vrijednost indikatora pokazuje dužinu cijele vremenske serije (vidjeti primjer1 i primjer2).

*Primjer 1:* Pretpostavimo, da početak uporedive vremenske serije mjesečnog istraživanja predstavlja posmatrani period januar '2009. U tom slučaju, indikator ima vrijednost za posmatrani period juni '2016.  $CC2 = (\text{juni '16} - \text{januar '09}) = 90$  (mjeseci).

*Primjer 2:* Ako je kod godišnjeg istraživanja godina prvog uporedivog objavljivanja rezultata 2005., onda je po objavljivanju rezultata za godinu 2016. vrijednost indikatora jednaka  $P1 = 12$  (godina).

Ako je početkom 2011. izmjenjena KD, a na osnovu toga promijenjen uzorak jedinica posmatranja i nije moguće uspostaviti odgovarajuću vezu sa starim podacima, te je došlo do prekida u vremenskoj seriji. Kako vremenska serija ima prekid u određenom dijelu unutar serije, a podaci su se i dalje nastavili objavljivati moguće je izračunati dvije vrijednosti pokazatelja (vidjeti primjer3).

*Primjer 3:* Ako je kod kvartalnog istraživanja *indeksa prometa u trgovini na veliko* (prvi posmatrani period je Q1 2007. godine a zadnji kraj Q4 2014. godine) došlo do prekida serije u 2011. godini (zbog izmjene KD), dužina vremenske serije iznosi:

$$CC2 = J_{last} - J_{first} + 1 = 16 - 1 + 1 = 16$$

$$CC2 = J_{last} - J_{first} + 1 = 15 - 1 + 1 = 15$$

### 5.2.3 Prekidi u vremenskim serijama

Opisuju se svi prekidi u vremenskim serijama te razlozi za nastanak prekida. Opisuju se svi faktori (koji doduše nisu prouzrokovali prekid vremenskih serija), ali su ipak na određeni način uticali na smanjenu uporedivost rezultata u različitim vremenskim tačkama

## 5.3 Geografska uporedivost

### 5.3.1 Uporedivost s ostalim članicama evropskog statističkog sistema

Navode se faktori, koji bi mogli uticati na to, da rezultati određenog istraživanja nisu sasvim uporedivi s rezultatima sličnih istraživanja u Evropskom statističkom sistemu. Navodi se takođe, da li postoji regulativa, koja uređuje harmoniziranu upotrebu (nekih dijelova) metodologije.

## 6 DOSTUPNOST I RAZUMLJIVOST, DISEMINACIJSKI FORMAT

Definicija komponente:

Dostupnost i razumljivost statističkih proizvoda/rezultata podrazumijevaju jednostavan i lak način na koji korisnici mogu pristupiti statističkim podacima koristeći jednostavne i lake procedure. Dostupnost statističkih rezultata/proizvoda se odnosi na konkretne fizičke okolnosti, u kojima su podaci dostupni korisniku: gdje se podaci fizički nalaze, kakve su mogućnosti korištenja, raspored objavljivanja, jasna politika plaćanja, dostupnost mikro i makropodataka, različiti formati i mediji (npr. saopćenja, publikacije, on-line baze podataka, dokumentacija o metodologiji i kvalitetu).



### 6.1 Saopćenja u kojima se objavljuju podaci

Navesti nazive redovnih i vanrednih saopćenja sa listom objavljenog seta podataka i on-line linkom na saopćenje.

### 6.2 Publikacije u kojima se objavljuju podaci

Navesti nazive redovnih i vanrednih publikacija sa listom objavljenog seta podataka i on-line linkom na publikaciju.

### 6.3 On – line baza podataka

Navesti informaciju o dostupnoj on-line bazi podataka kojoj se može pristupiti diseminiranim podacima sa linkom na istu.

### 6.4 Pristup mikropodacima

Navesti informaciju da li su dostupni mikropodaci, te ako jesu kratko opisati pravila anonimiziranja mikropodataka.

### 6.5 Dostupnost metodološke dokumentacije

Navesti informaciju o dostupnosti datoteka referentnih metapodataka, važnijih metodoloških priručnika, itd .... (naslov, izdavač, godina i link na on-line dokument).

### 6.6 Mjere za poboljšanje razumljivosti diseminiranih rezultata

Opisati bilo koje aktivnosti, koje su planirane u vezi s poboljšanjem razumljivosti printanih, internetskih objava, kao i diseminacionih baza podataka.

### 6.7 Indikator kvaliteta i učinka – Konsultovanje (korištenje) setova podataka (AC1)

*Opis indikatora:*

Broj konsultacija setova podataka unutar statističke oblasti u određenom periodu. Napomena: interno pregledavanje stranica se isključuje.

*Postupak izračunavanja, primjer i objašnjenje*

AC1 = #CONS

Broj korisničkih konsultacija u vezi statistike *Indeksa potrošačkih cijena* u 2014. godini iznosi 1.255 (web pages hits).

### 6.8 Indikator kvaliteta i učinka – Konsultovanje (korištenje) meta podataka (AC2)

*Opis indikatora:*

Broj konsultacija referentnih meta podataka (ESMS) unutar statističke oblasti u određenom periodu (*povratna informacija na uvid ili download po predmetnoj statističkoj oblasti – temi*)

*Postupak izračunavanja, primjer i objašnjenje*

AC2 = # ESMS

Broj konsultacija referentnih meta podataka (ESMS) u vezi statistike *Indeksa potrošačkih cijena* u 2014. godini iznosi 1.120 (web pages hits).

**6.9 Indikator kvaliteta i učinka - Stopa kompletnosti meta podataka (AC3)***Opis indikatora:*

Odnos broja raspoloživih (kompletiranih) metadata elemenata prema ukupnom broju preporučenih metadata elemenata.

*Postupak izračunavanja, primjer i objašnjenje*

Od ukupnog broja (60) metadata elemenata sadržanih u ESMS v.2.0 (*EURO-SDMX Metadata Structure*) za statistiku *Indeksa potrošačkih cijena* osigurano je 55.

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

Što znači - stopa kompletnosti meta podataka za statistiku *Indeksa potrošačkih cijena* je 92%.

**7 TROŠKOVI ISTRAŽIVANJA I OPTEREĆENOST DAVALACA PODATAKA***Definicija komponente:*

Troškovi i opterećenost davalaca podataka nisu stvarne dimenzije kvaliteta. Međutim, treba razmotriti kompromise između troška i opterećenja davalaca podataka ili, drugim riječima, troškovi i opterećenje davalaca podataka ograničavaju kvalitet statističkih podataka.

Mogućnost izračunavanja troškova nužna je za efikasno upravljanje, a posebno za procjenu kvaliteta i učinka. Analiza *koristi* i *troškova* potrebna je kako bi se odredili odgovarajući kompromisi između troškova i koristi (u smislu kvaliteta statističkih podataka). Na isti način, na sudjelovanje davalaca podataka mora se gledati kao na *trošak* (za davalaca podataka) koji se mora uravnotežiti sa *koristima* od podataka koji se daju.

**7.1 Troškovi provođenja statističkog istraživanja**

Kod izračunavanja troškova statističke institucije pri provođenju istraživanja treba uzeti u obzir materijalne troškove, te troškove koji proizilaze iz utroška radnog vremena statističkog osoblja.

Izračunavanje godišnjih operativnih troškova prema glavnim troškovnim komponentama prikazati u standardnoj tabeli:

Broj utrošenih sati rada	
Materijalni troškovi (štampanje i slanje obrazaca na teren)	
Godišnji broj obrazaca koji se dostavljaju izvještajnim jedinicama	

## 7.2 Opterećenost davalaca podataka

Kao indikator opterećenosti davalaca podataka, ocjenjuje se vrijeme, koje je potrebno za izvještajne jedinice za popunjavanje upitnika u toku jedne godine. Podaci se prikazuju u standardnoj tabeli:

Izračunavanje godišnjeg opterećenja davalaca podataka prikazati u standardnoj tabeli:

Broj davalaca podataka koji su popunili obrazac	
Vrijeme potrebno za popunjavanje jednog obrasca (sati)	
Ukupno upotrebjeno vrijeme (sati)	

## 7.3 Mjere za smanjivanje troškova i opterećenosti

Navesti moguće mjere za smanjenje troškova i opterećenosti davalaca podataka (npr. smanjenje veličine i koordinacija uzorka; smanjenje frekvencije prikupljanja podataka; smanjenje broja traženih podataka u istraživanju; smanjenje broja kontakata sa izvještajnom jedinicom; korištenje administrativnih podataka (može i dijela istih); testiranje upitnika sa izvještajnim jedinicama u cilju boljeg razumijevanja i popune istih; u kalendaru istraživanja koji je dostupan izvještajnoj jedinici navesti zahtjev za podacima; ukoliko treba pomoć pri popunjavanju upitnika izvještajna jedinica vas može kontaktirati; itd ...).

## 8 POVJERLJIVOST

U dokumentu *Statistics Code of Practice*, u načelima se navodi - da tajnost davalaca podataka (osoba, domaćinstava, preduzeća) i povjerljivost informacija koje pružaju mora biti apsolutno zagarantirana, a njihovi podaci mogu biti korišteni samo za statističke svrhe (načelo 5); statističke institucije moraju producirati i diseminirati statistiku (respektirajući stručnu neovisnost) na objektivan, stručan i transparentan način, u kojem će svi korisnici biti tretirani na isti način (načelo 6).

### 8.1 Povjerljivost - politika

Uobičajeno - povjerljivost statističkih podataka je zahtjevana zakonom a osoblje koje provodi statističko istraživanje ima po istom pravnom osnovu obavezu zaštite povjerljivosti. Zakoni o statistici određuju poverljivost kao jedan od glavnih principa. U zakonima je posebno poglavlje posvećeno ovom pitanju – poglavlje sadrži definiciju povjerljivih podataka, predviđa upotrebu poverljivih podataka samo za statističke potrebe, određuje obaveze proizvođača podataka koji treba precizno da definiraju mjere i procedure zaštite povjerljivosti podataka. Zakoni o statistici su dostupni na internet stranicama statističkih institucija.

Dakle, Izvještaj o kvalitetu treba potvrditi ovaj aranžman, ili izvještaj o bilo kojoj iznimci od ovog pravila.

## 8.2 Povjerljivost – postupanje sa podacima

Takođe, potrebno je u glavnim crtama navesti procedure za osiguranje povjerljivosti za vrijeme prikupljanja, obrade i diseminacije – što uključuje, protokole za zaštitu individualnih podataka kojima se ima pristup, pravila za definiranje povjerljivih ćelija u izlaznim tabelama i procedure za otkrivanje i prevenciju naknadnih razotkrivanja. Isto tako, treba navesti - ako spoljni korisnici imaju pristup mikropodacima za istraživačke namjene, da li se primjenjuju odredbe o povjerljivosti. U izvještaju treba opisati odredbe koje osiguravaju: zaštitu i integritet kompletnog upitnika, mikro i makro baze podataka i rezultata.

## 9 STATISTIČKA OBRADA

Statistička obrada se odnosi na operacije koje se izvode na podacima u cilju dobivanja novih informacija prema određenom setu pravila. Ovaj koncept se dalje dijeli na:

### 9.1 Izvor podataka

Navesti na kojem izvoru se temelji set podataka: na istraživanju, na administrativnim izvorima podataka, na miksu više izvora podataka ili podataka iz drugih statističkih aktivnosti. Ukoliko je korišten uzorak, treba dati neke karakteristike uzorka (npr. veličinu populacije, bruto i neto veličinu uzorka, vrsta dizajna uzorkovanja, itd.). Ako se koriste administrativni registri, treba dati opis registara (izvor, primarna svrha, itd).

### 9.2 Učestalost prikupljanja podataka

Navesti učestalost (periodiku) prikupljanja podataka (npr. mjesečno, tromjesečno, godišnje, kontinuirano).

### 9.3 Prikupljanje podataka

Navesti način prikupljanja podataka, npr.: iz administrativnih izvora; putem snimatelja cijena na upitniku (na terenu ili putem telefona); putem upitnika za određeno statističko istraživanje koji se kreiraju u statističkoj instituciji). Opisati korištene metode prikupljanja podataka za jedinice posmatranja (npr. uzorkom, putem terenske ankete, CAPI, on-line istraživanja, itd). Neke dodatne informacije o dizajnu i testiranju upitnika, obuci anketara, metode koje se koriste za praćenje neodgovora i sl. Priložiti link korištenih upitnika.

### 9.4 Validacija podataka

Opisati postupke za provjeru i vrednovanje izvora i izlaznih podataka i kako se rezultati ovih validacija prate i koriste. Aktivnosti validacije mogu uključivati: provjeru obuhvata populacije i po potrebi stope odziva; poređenje statistike sa prethodnim ciklusima (ako je primjenjivo); poređenje statistike sa drugim relevantnim podacima (internim i eksternim); istraživanje nedosljednosti u statistici; izvedbu uređivanja mikro i makro podataka; otkrivanje outlier-a.

### 9.5 Kompilacija podataka

Kompilacija podataka podrazumijeva postupke na podacima koje služe za izvođenje novih informacija u skladu sa datim setom pravila.

Opisati proces kompilacije podataka (npr. imputacija, ponderisanje, podešavanje za neodgovor, kalibracija i sl). Za imputacije: informacija u kojoj mjeri se koristi imputacija i razlozi; kratak opis korištenog metoda i njegov uticaj na procjenu. Ponderisanje: svaki korak ponderisanja opisati odvojeno; izračunavanje pondera. Podešavanje neodziva: kako je korigiran inicijalni ponder, uzimajući u obzir razlike u stopama odgovora; Kalibracije: varijable koje se koriste u podešavanju, primijenjene metode. Izračunavanje finalnih pondera.

## 9.6 Prilagođavanja

Set procedura za modifikaciju statističkih podataka u cilju osiguranja kompatibilnosti sa nacionalnim ili međunarodnim standardima, ili procedure za rješavanje razlika u kvalitetu podataka ako kompiliramo specifične setove podataka. Opisati vremenske serije koje su prilagođene i statističke postupke koje su korištene za podešavanje serije (metode desezoniranja npr. TRAMOSEATS, ARIMA ili druge slične metode). U slučaju prilagođavanja, navesti tip prilagođavanja (npr. sezonsko, kalendarsko, trend-ciklus). Ako je otkriven outlier i izvršena zamjena, navesti koje vrste outlier-a (impuls, prolazne promjene, smjene nivo) su otkrivene. Izvjestiti o softveru za podešavanje.

### 9.6.1 Sezonsko prilagođavanje

Statistička tehnika koja se koristi za otklanjanje efekata uticaja sezonskog kalendara na seriji podataka. Navesti kratak opis korištenog metoda.

## PREGLED INDIKATORA KVALITETA I UČINKA

Komponenta kvaliteta i učinka	Oznaka	Naziv indikatora kvaliteta i učinka
<b>Relevatnost</b>	R1	Stopa kompletnosti podataka
<b>Tačnost i preciznost</b>	A1	Uzoračka greška - Koeficijent varijacije
	A2	Stopa prekomjernog obuhvata
	A3	Udio zajedničkih jedinica
	A4	Stopa neodziva/neodgovora jedinice posmatranja
	A5	Stopa neodziva/neodgovora varijable
	A6	Prosječna veličina revizije
	A7	Stopa imputiranih podataka
<b>Pravovremenost i tačnost objavljivanja</b>	TP1	Pravovremenost prvih rezultata
	TP2	Pravovremenost konačnih rezultata
	TP3	Tačnost objavljivanja
<b>Dostupnost i razumljivost</b>	AC1	Konsultacije (korištenje) setova podataka
	AC2	Konsultacije (korištenje) meta podataka
	AC3	Stopa kompletnosti meta podataka
<b>Usklađenost i uporedivost</b>	CH1	Skladnost sa rezultatima iz referentnog izvora*
	CC1	Nepodudarnost između uporedivih statistika
	CC2	Dužina uporedivih vremenskih serija
<b>Troškovi istraživanja i opterećenost izvještajnih jedinica</b>	-	Troškovi provođenja statističkog istraživanja
	-	Opterećenje davalaca podataka

\* - Izračunavanje ovih indikatora nije sadržano u preporukama za izradu izvještaja o kvalitetu, ali su zbog značaja sastavni dio smjernica.

## LITERATURA

- Methodological documents, Standard Report: Working Group “*Assessment of quality in statistics*”, Sixth meeting, Luxembourg 2-3 October 2003
- Statistics Canada (2003): *Statistics Canada Quality Guidelines*, 4th edition, October 2003.
- Eurostat (2007): “*Handbook on Data Quality Assessment Methods and Tools (DatQAM)*”, Handbook written by Körner, T., Bergdahl, M., Elvers, E., Lohaus, P., Nimmergut, A., Šcbr, H.V., Szép, K., Timm, U., and Zilhão M.J.
- Eurostat (2009): “*ESS Standard for Quality Reports*”, Methodologies and Working Papers, Eurostat, Luxembourg
- Statistični ured Republike Slovenije (2011) – „*Kazalniki kakovosti*“, Ljubljana, 2011
- Office for National Statistics (2013): *Guidelines for Measuring Statistical Output Quality, - Version 4.1*, London 2013
- Eurostat (2014): “*ESS Handbook for Quality Reports*”, Manuals and guidelines, Eurostat, Luxembourg
- EU – India Seminar on Quality Matters in National Accounts (2014) – „*Quality Reporting, Quality Indicators and Metadata*“ Mr. Kari Djerf, 25-26 Nov 2014, New Delhi, India
- Državni zavod za statistiku Republike Hrvatske (2015): „*Priručnik za izračun pokazatelja kvalitete*“, Zagreb, 2015