



Anbefalte kvalitetsindikatorer i offisiell statistikk

TALL

SOM FORTELLER

NOTATER / DOCUMENTS

2024/5

Aslaug Hurlen Foss, Per Ola Haugen

I serien Notater publiseres dokumentasjon, metodebeskrivelser, modellbeskrivelser og standarder.

© Statistisk sentralbyrå
Publisert: 28. januar 2024

ISBN 978-82-587-1899-1 (elektronisk)
ISSN 2535-7271 (elektronisk)

Standardtegn i tabeller	Symbol
Ikke mulig å oppgi tall Tall finnes ikke på dette tidspunktet fordi kategorien ikke var i bruk da tallene ble samlet inn.	.
Tallgrunnlag mangler Tall er ikke kommet inn i våre databaser eller er for usikre til å publiseres.	..
Vises ikke av konfidensialitetshensyn Tall publiseres ikke for å unngå å identifisere personer eller virksomheter.	:
Desimaltegn	,

Forord

Kvalitetsindikatorer er viktig for å måle etterlevelsen av kvalitetskravene som er definert i «Statistikkloven» (Lovdata, 2019) og «Retningslinjer for europeisk statistikk» (SSB, 2017). «Rapport om kvalitet i offisiell statistikk, 2023» (SSB, 2023), viser at bruk av kvalitetsindikatorer for å måle kvalitet i produksjon av offisiell statistikk er varierende og generelt sett lite utbredt. Å ta i bruk kvalitetsindikatorer er en av hovedanbefalingene i rapporten, og produsentene av offisiell statistikk anbefales derfor å starte et arbeid med å vurdere relevante kvalitetsindikatorer for utvalgte statistikker. Kvalitetsindikatorer gir objektiv informasjon for vurdering av kvalitet i statistiske prosesser eller på statistiske produkter, og vil gi mulighet til å måle og følge et kvalitetsaspekt over tid, og å dokumentere om forbedringstiltak har effekt.

SSB har igangsatt et arbeid med anbefalte kvalitetsindikatorer, og ønsker å dele anbefalinger og erfaringer med alle produsenter av offisiell statistikk. Notatet har sin bakgrunn i arbeid med kvalitetsindikatorer på Seksjon for metoder i SSB, og lista over anbefalte kvalitetsindikatorer er forankret i SSBs standardutvalg¹. Ved å publisere dette notatet på ssb.no, tilgjengeliggjøres innholdet for alle produsenter av offisiell statistikk og andre interessenter.

Statistisk sentralbyrå, 12.01.2024

Arvid Olav Lysø

¹ Standardutvalget er et internt utvalg i SSB, med representanter fra alle avdelingene, som har ansvar for SSBs sentrale standardiseringsarbeid. Utvalget skal bidra til økt bruk av eksisterende standarder, samt utvikling av nye standarder.

Innhold

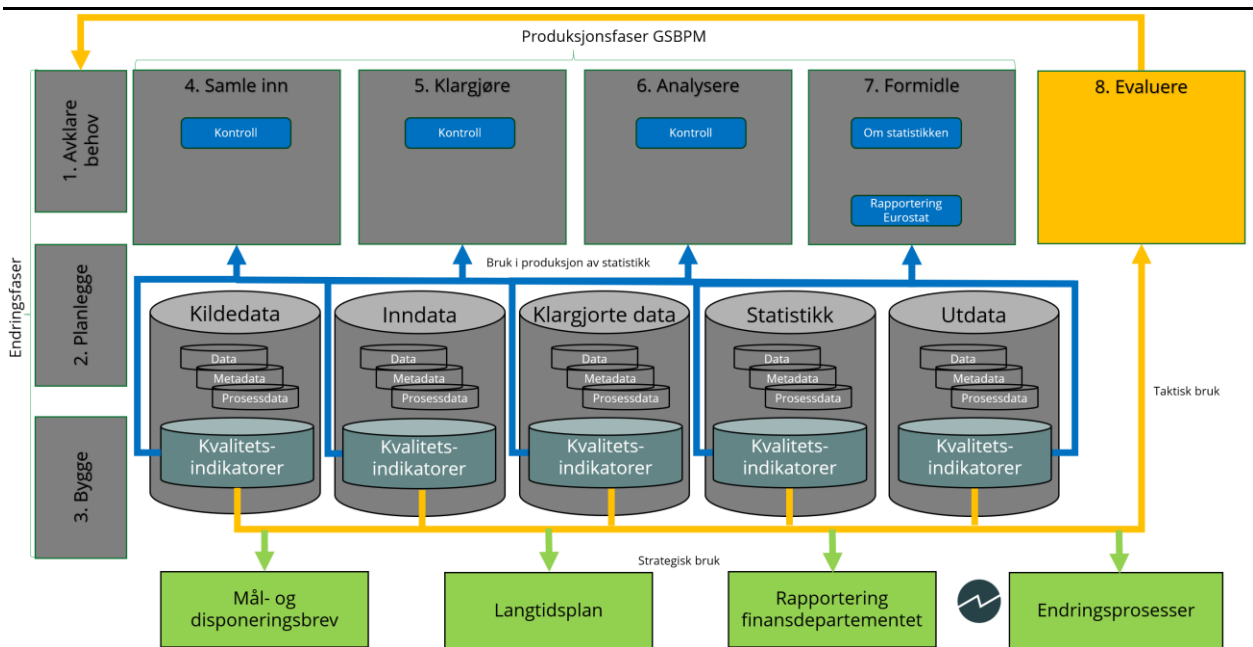
Forord	3
1. Om kvalitetsindikatorer	5
1.1. Hvorfor kvalitetsindikatorer?.....	5
1.2. Om anbefalingene	5
1.3. Produksjon av kvalitetsindikatorer	6
1.4. Internasjonale rammeverk for kvalitetsindikatorer	7
1.5. Leseveiledning.....	8
2. Anbefalte kvalitetsindikatorer	9
2.1. Editeringsandel.....	9
2.2. Kontrollutslagsandel.....	10
2.3. Koblingsandel	11
2.4. Effekten av editering.....	12
2.5. Usikkerhet	13
2.6. Gjennomsnittlig revisjon	14
2.7. Speilstrømmer	15
2.8. Lengde på tidsserie.....	16
2.9. Overdekningsandel.....	17
2.10. Tidspunkt for oppdatering av Om statistikken	18
2.11. Feil rettet etter publisering	18
2.12. Antall oppslag av statistikk.....	19
2.13. Antall oppslag av metadata	19
2.14. Kompletthetsandel metadata	20
2.15. Aktualitet foreløpige tall	21
2.16. Aktualitet endelige tall	22
2.17. Punktlighet	23
2.18. Avvik data mottatt	24
2.19. Andel ugyldige data	25
2.20. Enhetsfracfall.....	26
2.21. Partielt fracfall	27
2.22. Datakompletthetsandel	28
2.23. Tidsbruk skjemautfylling	29
2.24. Potensiell aktualitet	30
2.25. Andelen dubletter i datasettet	31
2.26. Treffsikkerhet.....	31
Referanser	32
Vedlegg A: Dokumentasjon av kvalitetsrapporter i ISEE	33

1. Om kvalitetsindikatorer

1.1. Hvorfor kvalitetsindikatorer?

Ved å gi en anbefaling om valg av kvalitetsindikatorer, skal statistikkprodusenter få hjelp til å definere hvilke kvalitetsindikatorer som er viktige. Det er også en målsetning at flere skal inspireres til å bruke indikatorer som et verktøy for å måle og overvåke kvaliteten i statistikken. Anbefalingene som gis i dette notatet gjelder kvantitative kvalitetsindikatorer, altså numeriske størrelser som kan måles og beregnes i et produksjonsløp. Kvalitetsindikatorerne skal bidra til at man kan oppdage svakheter og feil i produksjonen, og at det deretter settes inn tiltak for å rette opp disse. Noen indikatorer retter seg mot og gir mål på drift av statistikken i en produksjonssyklus, mens andre indikatorer gir informasjon som gjør at man kan sette inn tiltak for langsiktig forbedring av kvaliteten i statistikken. Ved å sammenstille eller aggregere indikatorer på tvers av statistikker kan indikatorerne også brukes på strategisk nivå, til rapportering og oppfølging, eller som input til endringsprosesser på avdelings- eller institusjonsnivå.

Figur 1.1 Bruksområder for kvalitetsindikatorer



Kilde: Statistisk sentralbyrå

1.2. Om anbefalingene

Det finnes svært mange ulike kvalitetsindikatorer som kan brukes til å måle og overvåke kvalitet i statistikkproduksjonen. Det er derfor utarbeidet et sett med anbefalte kvalitetsindikatorer som er forankret i Standardutvalget, der saken er behandlet flere ganger. Det er gjennomført to workshoper med statistikkavdelingene i SSB for å få en grundig og mer detaljert tilbakemelding på hvilke indikatorer som bør være anbefalte. Dette notatet presenterer de anbefalte kvalitetsindikatorerne som fikk endelig tilslutning i Standardutvalget i 2023.

Hver statistikk bør definere sin egen målsetning for kvalitet og velge indikatorer som støtter opp om denne. Valget av indikatorer bør også ta hensyn til hvilke kvalitetsaspekter brukerne av statistikken er opptatt av. Det finnes ofte flere brukergrupper som kan ha ulike behov, og disse må balanseres og prioriteres.

Når en skal velge kvalitetsindikatorer, er det viktig å vurdere hvilke risikoer man står overfor i statistikkproduksjonen. Hvilke feil er mulige, når kan de oppstå og hvilke konsekvenser kan feilene få for prosesser, produkter og metoder?

Lista over anbefalte kvalitetsindikatorer er ingen fasit. Ikke alle de anbefalte indikatorene er relevante for alle statistikker, og hvilke indikatorer som gir best effekt vil variere fra statistikk til statistikk. De anbefalte indikatorene må derfor ses på som et hjelpemiddel for å velge ut og implementere relevante indikatorer for det enkelte produksjonsløp.

1.3. Produksjon av kvalitetsindikatorer

Vi anbefaler å tegne opp et prosessflytkart over statistikkproduksjonen for å gi en visuell oversikt over produksjonsprosessen. En slik visualisering kan ta utgangspunkt i den internasjonale prosessmodellen for statistikk «GSBPM» (SSB, 2019). Statistikkproduksjon kan være kompleks, og den kan bestå av mange datakilder som brukes i et nettverk av prosesser. For hver hovedprosess bør det lages kvalitetsindikatorer.

En forutsetning for å skape mange kvalitetsindikatorer er at endringer og hendelser i dataene logges, med informasjon om:

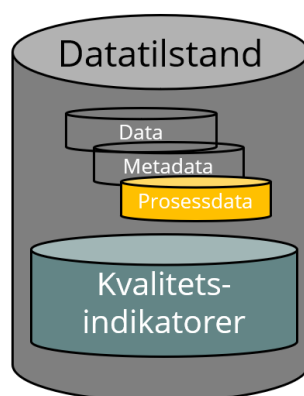
- Tidspunkt for hendelse eller endring, f.eks. tidspunkt for innlasting av data
- Hvem som trigget en manuell hendelse eller endring
- Hvilken type hendelse: ny verdi, endret verdi, slettet verdi
- Gammel og ny verdi
- Hvordan hendelsene eller endringen ble trigget: manuelt eller automatisk

Det er også god praksis å dokumentere hvorfor en manuell endring er gjort, det er derfor viktig å tilrettelegge for dokumentasjon.

Den informasjonen som oppstår i produksjonsprosessen, kaller vi prosessdata. Andre eksempler på prosessdata er opptelling av antall rader i et datasett som er lest inn, antall celler behandlet eller tiden en del av en prosess har tatt. Prosessdata er ofte nyttige i seg selv og bør derfor lagres. For makrostatistikk blir ofte hele tidsserier revidert bakover i tid. For disse statistikkene vil mengden av revisjon være en kvalitetsindikator. De er det viktig at alle versjoner av en tidsserie blir tatt vare på slik at det blir enkelt å beregne forskjell mellom første og siste versjon. En annen forutsetning for å skape indikatorer er at det finnes metadata som beskriver datasett og variabler. Eksempler på slike metadata er informasjon om variabelnavn, type variabel og knytning til klassifikasjon eller kodeverk.

Kvalitetsindikatorer skapes ofte ved å sammenstille data, metadata og prosessdata. Ved å standardisere lagringsstrukturer for datasett, metadata og prosessdata muliggjør man at kvalitetsindikatorer kan skapes effektivt og på en ensartet måte på tvers av statistikker. Dashboard for visualisering av kvalitetsindikatorer kan da brukes uten at man trenger å tilpasse disse for enkeltstatistikker.

Kvalitetsindikatorer bør lagres for å kunne sammenligne perioder og også ulike statistikker. Noen kvalitetsindikatorer inngår også i ekstern rapportering, for eksempel til Finansdepartementet eller Eurostat.

Figur 1.3. Prosessdata

Kilde: Statistisk sentralbyrå

1.4. Internasjonale rammeverk for kvalitetsindikatorer

Retningslinjer for europeisk statistikk

«Retningslinjer for europeisk statistikk» (SSB, 2017) er et felles kvalitetsrammeverk for det europeiske statistikkssystemet. Det er basert på 16 prinsipper som omfatter institusjonelle forhold, statistiske prosesser og statistiske produkter. For hvert prinsipp finnes det et sett av indikatorer for god praksis og standarder som gir veiledning og referanser for evaluering av gjennomføringen av retningslinjene. Hver kvalitetsindikator som anbefales i dette notatet vil knyttes opp mot prinsippene i disse retningslinjene. Prinsippene for produktkvalitet er: «relevans», "nøyaktighet og pålitelighet», «aktualitet og punktlighet», «sammenheng og sammenlignbarhet» og "tilgjengelighet og klarhet». Prinsippene for prosesskvalitet er: "gode metoder», «egne statistiske prosedyrer», "rimelige krav til oppgavegiverne» og "kostnadseffektivitet".

Kvalitetsindikatorer for statistisk prosessmodellen, GSBPM

UNECE² har laget et rammeverk for kvalitetsindikatorer basert på prosessmodellen for statistikkproduksjon, «Generic Statistical Business Process Model (GSBPM)» (SSB, 2019). For alle faser og prosesser er det foreslått kvalitetsindikatorer, til sammen 237 kvalitative og kvantitative kvalitetsindikatorer.³ Det er foreslått kvalitetsindikatorer som både passer for utvalgsstatistikk og for statistikk basert på administrative datakilder. Hvis en er interessert i hvilke kvalitetsindikatorer som gjelder for en bestemt prosess, kan man bruke en klikkbar GSBPM-ressurs⁴. Denne gir en beskrivelse av (del)prosessen og hvilke kvalitetsindikatorer som er anbefalt for denne prosessen. Alle kvalitetsindikatorer er knyttet opp til produksjonsfasene i denne modellen som er «samle inn», «klargjøre», «analysere» og «formidle».

Single Integrated Metadata Structure, SIMS

Det europeiske statistikkssystemet (ESS) sitt rammeverk for organisering og rapportering av referansemetadata er SIMS, «Single Integrated Metadata Structure» (Eurostat, 2023). Referansemetadata, i form av SIMS, er dokumentasjon av statistikk og beskriver hvordan prinsippene i Retningslinjer for europeisk statistikk oppfylles for den enkelte statistikk. SIMS er en overbygning av to standarder som er i ferd med å bli faset ut til fordel for SIMS: ESMS⁵ som er

² The United Nations Economic Commission for Europe

³ Kvalitetsindikatorer etter prosessmodellen, GSBPM: <https://statswiki.unece.org/display/GSBPM/Quality+Indicators>

⁴ Klikkbar prosessmodell, GSBPM, med kvalitet: <https://statswiki.unece.org/display/GSBPM/Clickable+GSBPM+v5.1>

⁵ Euro-SDMX Metadata Structure

brukerorientert og ESQRS⁶ som er produsentorientert. SIMS består av 19 kapitler og 80 delkapitler, herunder 16 kvantitative kvalitetsindikatorer og mange kvalitative indikatorer. En anbefaling om å ta i bruk SIMS fra Eurostat til medlemslandene kom i 2023. SSBs «Om statistikken» inneholder referansemetadata, men følger ikke SIMS fullt ut. Alle kvantitative kvalitetsindikatorer fra SIMS er tatt med i de kvalitetsindikatorerne som anbefales i dette notatet.

1.5. Leseveiledning

Hvert delkapittel i dette notatet presenterer én anbefalt kvalitetsindikator og følger malen som er gitt i tabellen under.

ID	Unik ID for kvalitetsindikatoren
Navn	Navn på indikatoren
Beskrivelse	Detaljert beskrivelse
Anvendelse	Hvilke statistiske prosesser og hvilken type statistiske produkter indikatoren kan anvendes på
Formel	Formel for utledning av kvalitetsindikator
Målverdi	Hvilken verdi indikatoren bør være i nærheten av eller bevege seg mot over tid. Jo nærmere målet verdien er jo bedre kvalitet indikeres.
Aggregeringsnivåer og prinsipper	Beskrivelse av hvilke typer aggregering som er relevant for indikatoren
Tolkning	Beskrivelse av hvordan resultatet av en indikator skal tolkes
Kvalitetsprinsipp	Hvilket eller hvilke kvalitetsprinsipp i Retningslinjer for europeisk statistikk kan indikatoren relateres til
Fase i prosessmodellen	Hvilken eller hvilke faser/hovedprosesser i prosessmodellen (GSBPM) kan indikatoren relateres til
Datatilstand	Hvilken eller hvilke datatilstander er indikatoren relevant for? Med datatilstander mener vi ulike tilstander data har i en statistikkproduksjon. Datatilstandene som en kvalitetsindikator kan relateres til er kildedata, inndata, klargjorte data, statistikk og utdata. De tre første tilstandene, er i hovedsak mikrodata som gir informasjon om enkeltenheter, mens statistikk og utdata i hovedsak er aggregerte data. Datatilstander er nærmere beskrevet i «Datatilstander SSB» (SSB, 2023).
SIMS	Er indikatoren en del av SIMS? Ja – hvilken / nei
Bruksnivå	På hvilket nivå brukes indikatoren: <ul style="list-style-type: none"> • Strategisk: Rapportering til Finansdepartementet eller som input til endringsprosesser eller andre aktiviteter på avdelings- eller organisasjonsnivå • Taktisk: Som input til endrings- og forbedringsprosesser på statistikk- eller statistikkområdenivå. F.eks. som grunnlag for fasen Evaluere i GSBPM. Denne typen kvalitetsindikatorer er utmerkede for å vurdere om igangsatte tiltak, f.eks. knyttet til data- eller prosesskvalitet, har hatt effekt. • Produksjon: Indikator som brukes direkte i produksjon av statistikken
Nivå	Er indikatoren på enhetsnivå (representasjon), variabelnivå (mål) eller statistikknivå?
Kodereferanser	Lenke til kode i Git eller andre versjonshåndteringssystemer der kode for generering av indikatoren er implementert
Annen dokumentasjon	Lenke til annen relevant dokumentasjon, f.eks. internasjonalt

⁶ ESS Standard for Quality Reports Structure

2. Anbefalte kvalitetsindikatorer

2.1. Editeringsandel

ID	1
Navn	Editeringsandel
Beskrivelse	Andel verdier editert i forhold til antall mulige verdier. Denne indikatoren beskriver omfanget av endringer i datasettet. Indikatoren kjøres på ferdig klargjorte data, dvs. etter at editering og imputering er gjort. Det kan beregnes andel editerte verdier per variabel og per enhet. Hvis det både blir utført automatisk og manuell editering, kan indikatoren også skille på dette. Indikatoren kan bidra både til å forbedre den statistiske nøyaktigheten til dataene og effektivisere produksjonsprosessen.
Anvendelse	Indikatoren er relevant for alle statistikker som samler inn og editerer data.
Formel	Forutsetninger: Alle data må ha tilhørende prosessdata om hvordan hver enkelt datacelle i datasettet er endret. Dette kan bli gjort med en loggfil på alle endringer som blir gjort eller at denne prosessdata blir inkorporert i datasettet ved hjelp av attributtvariabler. Ideelt sett burde alle dataceller ha informasjon om behandling med følgende koder: <ul style="list-style-type: none"> • U (eller blank) = Uendret • A = Endret som følge av en automatisk prosess • M = Endret som følge av en manuell prosess <p>Som minimum må det være dokumentasjon av endringer på radnivå i et datasett.</p> <p>Kalkulasjonsformler, for hver variabel eller hver nøkkelvariabel:</p> <ul style="list-style-type: none"> • Andel med U = Antall dataceller med U (eller blank) / Antall dataceller totalt • Andel med A = Antall dataceller med A / Antall dataceller totalt • Andel med M = Antall dataceller med M / Antall dataceller totalt
Målverdi	Her er det flere målverdier. For det første et mål at nøyaktigheten på inndata skal være så god som mulig. Slik sett bør andelen med verdier som ikke er editert over tid nærme seg 1, forutsatt at årsaken til at den nærmer seg 1 er at nøyaktigheten på inndata over tid forbedrer seg. På den annen side er det et mål at mest mulig av klargjøringen skal skje automatisk. Andel med manuelt editerte verdier bør derfor over tid nærme seg 0.
Aggregeringsnivåer og prinsipper	Indikatoren kan beregnes på variabelnivå og enhetsnivå, men kan også beregnes på statistikknivå. Videre kan indikatoren aggregeres på tvers av statistikker for å få tall på statistikkområde, eller på avdelings- eller organisasjonsnivå.
Tolkning	Indikatoren viser hvilke variabler eller enheter som editeres mye. En høy editeringsandel kan være en indikasjon på at nøyaktigheten på inndata er for dårlig, men det kan også si noe om hva som er de viktigste variablene i statistikken. En høy verdi på indikatoren kan gi grunnlag for tiltak for å forbedre datakvaliteten. Å følge indikatoren over tid er en fin pekepinn på om igangsatte tiltak har effekt.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet, Prinsipp 10: Kostnadseffektivitet
Fase i prosessmodellen	Klargjøre
Datatilstand	Ferdig klargjorte data
SIMS	A7. Imputation rate
Bruksnivå	I utgangspunktet taktisk, men det kan også være interessant å følge denne indikatoren på strategisk nivå, f.eks. i form av andel manuell editering på avdelings eller organisasjonsnivå over tid.
Nivå	Variabel- og statistikknivå for hver statistikkproduksjon
Kodereferanser	Annen dokumentasjon
	Dokumentasjon av kvalitetsrapporter i ISEE (system for editering i SSB), rapport-id 01 – 06. Se vedlegg A. "Statistics Canada Quality Guidelines, Sixth edition" (Statistics Canada, 2019), side 12.

2.2. Kontrollutslagsandel

ID	2
Navn	Kontrollutslagsandel
Beskrivelse	Andel utslag av kontroller i forhold til antall enheter kontrollert
Anvendelse	<p>Indikatoren er relevant for alle statistikker som samler inn og editerer data. Formålet med indikatoren er å vurdere nøyaktigheten på data og hvor treffsikre kontrollene er. Kontrollene som er relevante her er kontroller som gir utslag og kan logges på enhetsnivå.</p> <p>Kontroller av data kan settes opp i hele produksjonsprosessen fra fasen samle inn til formidling, men hovedtyngden er i klargjøringsfasen. Her kjøres kontroller på Klargjorte data-tilstanden <i>etter</i> at eventuelle automatiske korreksjoner og imputeringer er kjørt, men før eventuell manuell editering utføres.</p> <p>Tilsvarende kontrollutslagsandel er også relevante i Samle-inn-prosessen, f.eks. i forbindelse med kontroller i web-skjemaer.</p>
Formel	<p>Forutsetninger / avhengigheter: Alle kontrollutslag må logges med informasjon om undersøkelse, periode, enhet og kontroll-id.</p> <p>Kalkulasjonsformel, for hver kontroll:</p> <ul style="list-style-type: none"> Verdi = Antall utslag av kontrollen / antall dataceller som kontrollen er kjørt på.
Målverdi	<p>Høyt utslag av kontroller kan indikere to ting:</p> <ul style="list-style-type: none"> Dårlig datakvalitet At kontrollene ikke er treffsikre, f.eks. at grenseverdier i kontroller er satt feil. <p>Verdien bør over tid nærme seg 0. Det indikerer både at nøyaktigheten på inndata er god, automatiske korreksjoner er effektive og at kontrollen er treffsikker.</p>
Aggregeringsnivåer og prinsipper	
Tolkning	<p>Kontroller kjøres for å identifisere mulige feil i datagrunnlaget. Målet er at kontrollene skal være så treffsikre som mulig. Man må unngå utslag av kontroller som ikke fører til editeringer som forbedrer den statistiske nøyaktigheten i dataene.</p> <p>Kontrollutslagsandel vurderes i prosess 8 – evaluere. Aktuelle tiltak i tilknytning til dette vil være:</p> <ul style="list-style-type: none"> Fjerne eller endre kontroller som er lite treffsikre. F.eks. kontroller med kontrollutslagsandel > 0,1. Vurdere relevante tiltak i Samle inn for kontroller med høy kontrollutslagsrate, f.eks. å legge inn relevante kontroller i web-skjema. Legge inn automatiske korreksjoner av data. <p>Se også indikator «Treffsikkerhet» for analyse av kontrollene.</p>
Kvalitetsprinsipp	<p>Prinsipp 12: Nøyaktighet og pålitelighet</p> <p>Prinsipp 10: Kostnadseffektivitet</p>
Fase i prosessmodellen	Klargjøre i utgangspunktet, men er også relevant for fase Samle inn for skjemaundersøkelser.
Datatilstand	I utgangspunktet Klargjorte data, men indikatoren kan også være relevant på kilde-data for skjemaundersøkelser.
SIMS	Nei
Bruksnivå	Taktisk
Nivå	Variabelnivå for hver statistikkproduksjon
Kodereferanser	
Annen dokumentasjon	Dokumentasjon av kvalitetsrapporter i ISEE, rapport-id 07. Se vedlegg A.

2.3. Koblingsandel

ID	3
Navn	Koblingsandel
Beskrivelse	Andel koblede og ikke-koblede data ved kobling av to datasett
Anvendelse	Denne kvalitetsindikatoren bør anvendes når statistikken er basert på integrasjon av flere datakilder.
Formel	<p>Kalkulasjonsformel, for hver kobling av to datasett (A & B):</p> <ul style="list-style-type: none"> • Andel kun A = Antall kun i A / Totalt antall med unik koblings-id i A & B • Andel kun B = Antall kun i B / Totalt antall med unik koblings-id i A & B • Andel koblede = Antall både i A&B / Totalt antall med unik koblings-id i A & B <p>Andel kun A + andel kun B + andel koblede skal summere seg til 1.</p>
Målværdi	Kan variere fra statistikk til statistikk, men i utgangspunktet er det ønskelig med en verdi så nærme 1 som mulig for Andel koblede.
Aggregeringsnivåer og prinsipper	
Tolkning	En lav verdi indikerer at det kan være problemer (dubletter eller feil/manglende verdier i id-variabler) med datasettene som skal kobles. Det bør da vurderes å sette i gang tiltak for å øke datakvaliteten og man kan bruke denne indikatoren for å måle om igangsatte tiltak har ønsket effekt.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	Klargjøre.
Datatilstand	Klargjorte data
SIMS	A3. Common units proportion
Bruksnivå	Produksjon
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	"European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition" (Eurostat, 2021), kapittel: S.13.3.1.2.

2.4. Effekten av editering

ID	4
Navn	Effekt av editering
Beskrivelse	<p>Indikatoren gir en oversikt over prosentvis endring på nøkkelvariable (aggregerte og eventuelt vektet) som følge av klargjøre-prosessen som omfatter både automatisk korreksjon av data og eventuell manuell korreksjon.</p> <p>En måte å lage denne indikatoren på er å først kjøre gjennom hele produksjonsløpet til statistikk eller utdata uten noen form for editering eller imputering. Da har man et datasett i statistikk- eller utdata-tilstanden som er aggregert og vektet, men helt <i>uten</i> nøyaktighetsforbedrende tiltak. Deretter, etter man er ferdig med editering og imputering, både automatiske og manuelle prosesser, og produksjonsløpet er kjørt så har man et tilsvarende datasett <i>med</i> nøyaktighetsforbedrende tiltak. Da kan man sammenligne variablene i disse to datasettene ut ifra formelen nedenfor.</p> <p>En annen måte å lage indikatoren på, som også viser produksjonstid som en dimensjon, er å kjøre ut denne indikatoren daglig (automatisert). Da vil man kunne lage en grafisk fremstilling med tidsdimensjonen på x-aksen og forbedring av nøyaktigheten på y-aksen. Man vil da av en slik kurve se når i produksjonsprosessen editering har størst effekt og når man kan avslutte editering uten at det går utover den statistiske nøyaktigheten.</p>
Anvendelse	
Formel	<p>Kalkulasjonsformel, for nøkkelvariable:</p> <ul style="list-style-type: none"> Verdi = verdi ikke-klargjorte data / verdi klargjorte data × 100.
Målverdi	
Aggregeringsnivåer og prinsipper	Denne indikatoren kan kjøres på ulike aggregerte nivåer. Man kan tenke seg en type drille-funksjonalitet der man enkelt kan gå mellom ulike aggregerte nivåer.
Tolkning	<p>Indikatoren må ses i sammenheng med bl.a. ressursbruk på editering.</p> <p>Ved høy ressursbruk og/eller lang produksjonstid og liten effekt av klargjøre-prosessen (lav prosentvis endring) bør man vurdere tiltak for å få ned ressursbruk og produksjonstid.</p> <p>Høy prosentvis endring indikerer problemer med nøyaktigheten på inndata. Tiltak for å forbedre nøyaktigheten på inndata bør vurderes.</p>
Kvalitetsprinsipp	<p>Prinsipp 12: Nøyaktighet og pålitelighet</p> <p>Prinsipp 10: Kostnadseffektivitet</p> <p>Prinsipp 13: Aktualitet og punktlighet</p>
Fase i prosessmodellen	Klargjøre.
Datatilstand	Klargjorte data
SIMS	Nei
Bruksnivå	Taktisk
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	

2.5. Usikkerhet

ID	5
Navn	Usikkerhet
Beskrivelse	Usikkerhet er en indikator som måler hvor usikre totaler og aggregater er. Ofte er denne indikatoren begrenset til utvalgsusikkerheten, som er et mål for den feilkilden som oppstår når man kun har informasjon fra en delmengde av populasjonen. Et vanlig mål for usikkerhet er variasjonskoeffisienten, som måler hvor stor prosentandel standardfeilen utgjør av den beregnede totalen eller snittet. Konfidensintervall kan også bli brukt.
Anvendelse	Dette er mest aktuelt for utvalgsbaserte statistikker og statistikker som imputerer celleverdier ved hjelp av en modell
Formel	Variasjonskoeffisient= standardfeil / total eller gjennomsnitt. Indikatoren oppgis vanligvis i prosent. Beregning av standardfeilen vil avhenge av antagelser gjort om dataene og hvilken metode som er brukt for imputering eller vektning. Usikkerhetsberegninger kan være kompliserte og det bør brukes anerkjente funksjoner for dette. I tillegg må en være oppmerksom på hvordan ekstremverdier er håndtert, siden dette kan påvirke beregningene
Målverdi	Målet er at denne indikatoren skal være så liten som mulig. Hovedtall bør ha en variasjonskoeffisient som ligger under 3 prosent. Ved variasjonskoeffisient på nærmere 20 prosent må en vurdere nytten av å publisere disse tallene
Aggregeringsnivåer og prinsipper	Variasjonskoeffisienten må beregnes for hvert aggregat og kan ikke summeres.
Tolkning	Variasjonskoeffisient er et mål på den relative usikkerheten til en statistikk.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet Prinsipp 15: Tilgjengelighet og klarhet
Fase i prosessmodellen	Klargjøre
Datatilstand	Klargjorte data
SIMS	A1. Sampling error – indicators
Bruksnivå	Produksjon
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.13.2.

2.6. Gjennomsnittlig revisjon

ID	6
Navn	Gjennomsnittlig revisjon
Beskrivelse	<p>Statistikk kan bli gjenstand for planlagt revisjon, det vil si endring av publiserte verdier. Dette kan blant annet skyldes nye og bedre kildedata, justeringer av tall og endringer av metoder, klassifikasjoner og definisjoner. Når det er revisjon av tidsserier, da kan gjennomsnittlig revisjon beregnes for tidsserien. Ved sesongjustering av tidsserier blir de sesongjusterte seriene revidert bakover i tid. Gjennomsnittlig revisjon bør da også bli beregnet på disse seriene.</p> <p>Gjennomsnittlig revisjon kan også beregnes mellom foreløpige og endelige tall, men da for snittet over aggregater og ikke tidsserien.</p>
Anvendelse	Dette gjelder spesielt økonomisk korttidsstatistikk, slik som Nasjonalregnskapet. Men det kan beregnes for alle statistikker som publiserer reviderte tall.
Formel	<p>Ulike indikatorer kan utledes ved forskjellige måter å beregne gjennomsnittet av revisjonene for en tidsserie. Revisjoner kan beregnes som gjennomsnitt av absolutt verdi eller ikke, indikatoren kan være absolutt eller relativ.</p> <p>Gjennomsnitt absolutt revisjon (MAR):</p> $MAR = \frac{1}{n} \sum_{t=1}^n L_t - P_t $ <ul style="list-style-type: none"> t = 1, ..., n antall referansetidspunkt for analyseperioden. L_t: Siste estimat for referanse perioden t P_t: Tidlig estimat for referanse perioden t
Måleverdi	Måleverdien bør være så liten som mulig.
Aggregeringsnivåer og prinsipper	Indikatoren må beregnes for hvert aggregat
Tolkning	Hvor stor gjennomsnittsendringen er ved revisjon av statistikk
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	Analysere
Datatilstand	Statistikk
SIMS	A6. Data revision – average size
Bruksnivå	Strategisk
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	<p>“Interpreting the results of Revision Analyses: Recommended Summary Statistics” (McKenzie R. & Gamba M., 2008)</p> <p>“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.17.2.</p> <p>https://www.ssb.no/omssb/kvalitet-i-offisiell-statistikk/prinsipper-for-kommunikasjon-og-formidling/prinsipper-for-revisjon-i-ssb</p>

2.7. Speilstrømmer

ID	7
Navn	Speilstrømmer
Beskrivelse	<p>Avvik mellom data knyttet til flyter som har sammenheng og som skal speile hverandre. Eksempler:</p> <ul style="list-style-type: none"> • Antall som innvandrer fra Sverige til Norge bør speile statistikk i Sverige over antall som utvandrer fra Sverige til Norge • Antall skattemeldinger bør tilsvare antall skatteoppgjør, da alle som har en skattemelding også skal ha et skatteoppgjør
Anvendelse	Statistikk over strømmer som handel, migrasjon, turismestatistikk, betalingsbalanse.
Formel	<p>Bilateral speilstatistikk. For hvert par av land, anta:</p> <ul style="list-style-type: none"> • A – Land A • B – Land B $CC2A_B = \frac{OF_{AB} - mIF_{AB}}{\frac{OF_{AB} + mIF_{AB}}{2}}$ $CC2B_A = \frac{OF_{BA} - mIF_{BA}}{\frac{OF_{BA} + mIF_{BA}}{2}}$ <ul style="list-style-type: none"> • OF_{AB} – utgående flyt som går fra land A til land B • mIF_{AB} – speil inngående flyt • mIF_{BA} – speil inngående flyt til land B fra land A • mOF_{AB} – speil utgående strømning <p>Et felles kvalitetsmål kan oppnås fra de to forskjellene i forhold til en gjennomsnittlig strømning (flere muligheter, en er gitt nedenfor):</p> $CC2_{AB} = \frac{\frac{ OF_{AB} - mIF_{AB} }{\frac{OF_{AB} + mIF_{AB}}{2}} + \frac{ OF_{BA} - mIF_{BA} }{\frac{OF_{BA} + mIF_{BA}}{2}}}{2}$ <p>Multilateral speilstatistikk:</p> <ul style="list-style-type: none"> • OF_{AiOj} – utgående strøm som går fra land A_i til et hvilket som helst annet land O_j • mIF_{AiOj} – speil inngående flyt • A_i – land A_i • O_j – Et annet land O_j • K – antall land land A_i kan ha kontakter med • C – gruppe av land EU + EFTA $CC2_C = \frac{\sum_{i=1}^C \sum_{j=1}^K OF_{AiOj} - mIF_{AiOj} }{\sum_{i=1}^C \sum_{j=1}^K \frac{OF_{AiOj} + mIF_{AiOj}}{2}}$
Målværdi	Verdien av denne indikatoren bør være så nær null som mulig, siden verdien av inngående og utgående strømmer mellom par av land bør matche.
Aggregeringsnivåer og prinsipper	Denne indikatoren utarbeides på aggregert nasjonalt nivå eller mellom statistikker fra forskjellige land.
Tolkning	
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet Prinsipp 14: Sammenheng og sammenlignbarhet
Fase i prosessmodellen	Analysere
Datatilstand	Statistikk
SIMS	CC1. Data revision – asymmetry for mirror flow statistics – coefficient
Bruksnivå	Strategisk
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.15.1.1.

2.8. Lengde på tidsserie

ID	8
Navn	Lengde på tidsserie
Beskrivelse	Antall perioder siden forrige brudd i tidsserien
Anvendelse	Alle statistikker som publiserer tall regelmessig på ssb.no
Formel	Verdi = Siste periode publisert – første periode med sammenlignbar statistikk + 1
Målverdi	En lang sammenlignbar tidsserie er i utgangspunktet ønskelig, men å gjøre endringer i statistikken som fører til brudd i tidsserien kan være riktig, f.eks. for å tilpasse seg nytt datagrunnlag, forbedrede estimeringsmetoder eller for å oppnå sammenlignbarhet med annen statistikk.
Aggregeringsnivåer og prinsipper	
Tolkning	Hvis det ikke har vært noen brudd i tidsserien, er verdien lik antall perioder som statistikken er publisert.
Kvalitetsprinsipp	Prinsipp 14: Sammenheng og sammenlignbarhet
Fase i prosessmodellen	Formidle
Datatilstand	Statistikk
SIMS	CC2. Length of comparable time series.
Bruksnivå	Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.15.2

2.9. Overdekningsandel

ID	9
Navn	Overdekningsandel
Beskrivelse	Andel enheter som er fjernet fordi enhetene ikke er innenfor kriteriene til populasjonen som det skal lages statistikk for. Overdekning kan skyldes feil i enhetsidentifikasjon, klassifisering og koding etc.
Anvendelse	Overdekningsgraden gjelder både for utvalgsundersøkelser og for statistikk basert på bruk av administrative kilder. Dersom undersøkelsen har mer enn én enhetstype, kan det beregnes en indikator for hver type.
Formel	Overdekningsandel (Ocr): $Ocr = \frac{\sum_O w_j}{\sum_O w_j + \sum_E w_j}$ <ul style="list-style-type: none"> • O: sett med enheter som ikke tilhører målgruppen • E: sett med enheter som tilhører målgruppen • w_j: vekt av enhet j, beskrevet nedenfor. <p>De tre hovedtilfellene er:</p> <ul style="list-style-type: none"> • Uvektet: $w_j = 1$ • Designvektet: $w_j = d_j$ hvor i utgangspunktet $d_j = 1/\pi_j$, noe som betyr at designvekten er den inverse av seleksjonssannsynligheten. • Størrelsesvektet: $w_j = d_j x_j$ hvor x_j er verdien av en variabel X
Målværdi	Indikatoren bør være så nærme 0 som mulig.
Aggregeringsnivåer og prinsipper	
Tolkning	Indikatoren kan gi en helhetlig idé om kvaliteten på populasjonen. <ul style="list-style-type: none"> • Den uveide overdekningsraten gir antall enheter som har blitt funnet å ikke tilhøre målpopulasjonen i forhold til det totale antallet observerte enheter. • Den designvektede overdekningsgraden er et estimat for overdekningsraten basert på informasjon tilgjengelig i utvalgsundersøkelser. • Den størrelsesvektede overdekningsgraden uttrykkes i form av valgt størrelsesvariabel, f.eks. omsetning i næringsstatistikk.
Kvalitetsprinsipp	12. Nøyaktighet og pålitelighet
Fase i prosessmodellen	Samle inn
Datatilstand	Inndata
SIMS	A2. Over-coverage – rate
Bruksnivå	Produksjon, taktisk
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.13.3.1.

2.10. Tidspunkt for oppdatering av Om statistikken

ID	10
Navn	Tidspunkt for oppdatering av Om statistikken
Beskrivelse	Antall dager fra siste publisering til «Om statistikken» ble sist oppdatert. Denne kvalitetsindikatoren kjøres i forbindelse med publisering.
Anvendelse	Alle statistikker som publiserer statistikk på egen hjemmeside.
Formel	Verdi = Publiseringsdato - dato for sist oppdatering av Om statistikken
Målverdi	Så lav som mulig. Om statistikken bør vurderes for oppdatering ved hver publisering.
Aggregeringsnivåer og prinsipper	
Tolkning	Høy verdi viser at det er lenge siden Om statistikken sist er oppdatert, noe som indikerer at dokumentasjon av statistikken ikke er godt nok oppdatert.
Kvalitetsprinsipp	Prinsipp 15: Tilgjengelighet og klarhet
Fase i prosessmodellen	Formidle
Datatilstand	Statistikk, utdata
SIMS	Nei
Bruksnivå	Produksjon, Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	

2.11. Feil rettet etter publisering

ID	11
Navn	Feil rettet etter publisering
Beskrivelse	Antall ganger en statistikk er rettet etter publisering av endelige tall.
Anvendelse	Alle statistikker som publiserer tall på SSB.no
Formel	Verdi = Sum av antall ulike tidspunkt for innlasting av data til statistikkbanken etter første publisering av endelige tall.
Målverdi	Alle tidspunkt for lasting til statistikkbanken logges og kan hentes ut. Det må være mulig å skille mellom publisering av foreløpige og endelige tall. I utgangspunktet er målverdien 0. Det indikerer at statistikken er riktig og at det ikke er behov for å publisere på nytt. Noen statistikker har imidlertid planlagte revisjoner som publiseres.
Aggregeringsnivåer og prinsipper	
Tolkning	Et tall > 0 indikerer at nøyaktigheten på statistikken ikke var god nok ved opprinnelig publisering. Alvorlighetsgraden på eventuelle feil som oppdages må imidlertid vurderes i sammenheng med denne indikatoren. Det kan være et dilemma mellom kvalitetskravene om nøyaktighet, punktlighet og aktualitet. Hvilket eller hvilke av disse kravene som skal tillegges størst vekt vil variere fra statistikk til statistikk. Det vil være interessant å se denne indikatoren i sammenheng med indikatorene 15 - 17 om aktualitet og punktlighet, for å se om eventuelt hvor mye forbedret aktualitet påvirker nøyaktighet på de statistiske produktene.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet.
Fase i prosessmodellen	Formidle
Datatilstand	Statistikk
SIMS	Nei
Bruksnivå	Taktisk, strategisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	

2.12. Antall oppslag av statistikk

ID	13
Navn	Antall oppslag av statistikk
Beskrivelse	Antall treff på statistikkens hjemmeside på SSB.no innenfor en gitt tidsperiode, fortrinnsvis månedlig.
Anvendelse	Alle statistikker som publiserer tall på SSB.no
Formel	Forutsetter at antall treff telles og tilgjengeliggjøres. Opptelling av antall treff.
Målverdi	
Aggregeringsnivåer og prinsipper	Indikatoren kan aggregeres over tid, for eksempel fra måned til år, og fra statistikk til statistikkområde.
Tolkning	Indikatoren gir en pekepinn på hvor stor interesse det er for statistikken. Det er interessant å følge denne indikatoren over tid for å se hvordan interessen utvikler seg.
Kvalitetsprinsipp	Prinsipp 15: Tilgjengelighet og klarhet Prinsipp 11: Relevans
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	AC1. Data tables - consultations
Bruksnivå	Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	"European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition" (Eurostat, 2021), kapittel: S.10.3.1.

2.13. Antall oppslag av metadata

ID	13
Navn	Antall oppslag i metadata
Beskrivelse	Antall oppslag på statistikkens «Om statistikken».
	Alternativt telle oppslag på metadatafiler som publiseres hos Eurostat.
Anvendelse	Alle statistikker som publiserer tall på ssb.no.
Formel	Forutsetter at oppslag på «Om statistikken» logges. Opptelling av antall treff.
Målverdi	
Aggregeringsnivåer og prinsipper	
Tolkning	Indikatoren gir en pekepinn på brukernes interesse for statistikken og er interessant å følge over tid.
Kvalitetsprinsipp	Prinsipp 15: Tilgjengelighet og klarhet
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	AC2. Metadata tables - consultations
Bruksnivå	Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	"European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition" (Eurostat, 2021), kapittel: S.10.5.1.

2.14. Kompletthetsandel metadata

ID	14
Navn	Kompletthetsandel metadata
Beskrivelse	Andel metadataelementer i SIMS utfylt i forhold til antall metadataelementer i SIMS totalt. Med metadata på denne indikatoren menes referansemetadata, altså metadata man i dag kjenner fra «Om statistikken» eller metadatarapporter som sendes til Eurostat.
Anvendelse	Alle statistikker som publiserer tall på SSB.no.
Formel	For hvert element i SIMS: Verdi = Antall metadataelementer utfylt / antall metadataelementer totalt. Forutsetning at informasjon kan hentes ut fra Om statistikken eller metadatarapporter som sendes til Eurostat.
Målverdi	Så nærme 1 som mulig
Aggregeringsnivåer og prinsipper	
Tolkning	Indikatoren avdekker manglende metadata. Vær oppmerksom at denne indikatoren kun gir et bilde av om metadatainformasjon er utfylt. Den sier ikke noe om kvaliteten på metadata. Verdt å merke seg er også at alle metadataelementer ikke er relevante for alle statistikker. For eksempel så er det ikke relevant for årsstatistikker å dokumentere sesongjustering. Men i de tilfellene skal termen «Ikke relevant» benyttes fremfor å ha et tomt felt.
Kvalitetsprinsipp	Prinsipp 15: Tilgjengelighet og klarhet
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	AC3. Metadata completeness - rate
Bruksnivå	Produksjon, Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.10.6.1.

2.15. Aktualitet foreløpige tall

ID	15
Navn	Aktualitet foreløpige tall
Beskrivelse	Antall dager fra referanseperiodens utløp til publisering av foreløpige tall. Denne indikatoren rapporteres til finansdepartementet på SSB-nivå (aggregert for alle statistikker).
Anvendelse	Alle statistikker som publiserer foreløpige tall. Indikatoren er ikke relevant for statistikker som kun publiserer endelige tall.
Formel	Verdi = Tidspunkt (dato) for publisering av foreløpige tall – tidspunkt (dato) for referanseperiodens utløp. Tidspunkt (dato) for publisering av foreløpige tall refererer til faktisk tidspunkt, ikke forhåndsvarslet tidspunkt i statistikkalenderen. For statistikker som har flere publiseringer i løpet av et år, slik som måneds- og kvartalsstatistikk, beregnes et gjennomsnitt av denne indikatoren.
Målverdi	Jo lavere verdi, jo bedre aktualitet.
Aggregeringsnivåer og prinsipper	
Tolkning	Indikatoren kan brukes til å måle utviklingen av aktualitet over tid. Spesielt interessant er det å følge denne der det er igangsatt tiltak for å forbedre aktualitet. Indikatoren kan også brukes til å sammenligne aktualitet på tvers av statistikker.
Kvalitetsprinsipp	Prinsipp 13: Aktualitet og punktlighet
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	TP1. Time lag – first results
Bruksnivå	Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.14.1.1.

2.16. Aktualitet endelige tall

ID	16
Navn	Aktualitet endelige tall
Beskrivelse	Antall dager fra referanseperiodens utløp til publisering av endelige tall. Denne indikatoren rapporteres til finansdepartementet på SSB-nivå (aggregert for alle statistikker).
Anvendelse	Alle statistikker som publiserer tall på SSB.no
Formel	Verdi = Tidspunkt for publisering av endelige tall – tidspunkt (dato) for referanseperiodens utløp. Tidspunkt (dato) for publisering av endelige tall refererer til faktisk tidspunkt, ikke forhåndsvarslet tidspunkt i statistikkalenderen. For statistikker som har flere publiseringer i løpet av et år, slik som måneds- og kvartalsstatistikk, beregnes et gjennomsnitt av denne indikatoren.
Målverdi	Jo lavere verdi jo bedre aktualitet. På aggregert nivå er det satt måltall (styringsindikator) av finansdepartementet som det er viktig å oppnå.
Aggregeringsnivåer og prinsipper	Rapporteres til finansdepartementet på SSB-nivå
Tolkning	Indikatoren kan brukes til å måle utviklingen av aktualitet over tid. Spesielt interessant er det å følge denne der det er igangsatt tiltak for å forbedre aktualitet. Indikatoren kan også brukes til å sammenligne aktualitet på tvers av statistikker.
Kvalitetsprinsipp	Prinsipp 13: Aktualitet og punktlighet
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	TP2. Time lag – final results
Bruksnivå	Taktisk, strategisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.14.1.2.

2.17. Punktlighet

ID	17
Navn	Punktlighet
Beskrivelse	Avvik mellom forhåndskunngjort publiseringstidspunkt og faktisk publisering
Anvendelse	Alle statistikker som publiserer tall på SSB.no
Formel	<p>Indikatoren er todelt:</p> <ul style="list-style-type: none"> 17.1 Antall dager mellom faktisk publisering og forhåndskunngjort publiseringstidspunkt. Formel: Verdi = dato for faktisk publisering – dato for forhåndskunngjort publiseringstidspunkt i statistikkalenderen. Denne er rettet mot produsenter. 17.2 Andel statistikker publisert <i>på</i> annonsert dato i statistikkalenderen. Verdi = Statistikker publisert i henhold til statistikkalenderen / Totalt antall statistikker publisert. Denne er rettet mot brukere og er på aggregert nivå. <p>Forutsetning er at informasjon om forhåndskunngjort publiseringstidspunkt (statistikkalenderen) og tidspunkt for faktisk publisering er tilgjengelig.</p>
Målverdi	<ul style="list-style-type: none"> For 17.1: Målverdi = 0, noe som indikerer full punktlighet For 17.2: Målverdi = 1, noe som indikerer at alle statistikker publiseres i henhold til statistikkalenderen. <p>Verdt å merke seg er også at publisering <i>før</i> forhåndskunngjort dato vil gi negativ verdi, noe som også vil være et avvik.</p>
Aggregeringsnivåer og prinsipper	
Tolkning	<p>Indikator 17.2 ble til og med 2020 rapportert til finansdepartementet på aggregert nivå. Selv om indikatoren ikke lengre rapporteres til finansdepartementet er det viktig at SSB ser hvordan denne indikatoren utvikler seg.</p> <p>Indikatoren kan ses i sammenheng med indikator 12 – Feil rettet etter publisering. Det vil kunne være en avveining mellom punktlighet og nøyaktighet.</p>
Kvalitetsprinsipp	Prinsipp 13: Aktualitet og punktlighet
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	TP3. Punctuality
Bruksnivå	Taktisk, strategisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.14.2.1.

2.18. Avvik data mottatt

ID	18
Navn	Avvik data mottatt
Beskrivelse	Avvik i antall mottatte enheter sammenlignet med tidligere perioder; Forrige periode og tilsvarende periode forrige år for korttidsstatistikker.
Anvendelse	Alle statistikker som har egen datainnsamling. Aktuell for både skjema og registerbaserte statistikker.
Formel	Indikatoren er todelt: <ul style="list-style-type: none"> • 18.1 Verdi = Antall mottatte enheter / antall mottatte enheter forrige periode. • 18.2 for korttidsstatistikk: Verdi = Antall mottatte enheter / antall mottatte enheter samme periode forrige år.
Målverdi	Ingen spesifikk målverdi, men hvis avviket er stort er det en indikasjon på at mottatte data ikke er riktig.
Aggregeringsnivåer og prinsipper	
Tolkning	Indikatoren brukes til å avdekke mangler i datagrunnlaget som er mottatt ved å sammenligne med antall enheter mottatt i tidligere perioder. Indikatoren brukes typisk i produksjon av statistikken.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	Samle inn
Datatilstand	Kildedata
SIMS	Nei
Bruksnivå	Produksjon
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	

2.19. Andel ugyldige data

ID	19
Navn	Andel ugyldige data
Beskrivelse	Andel data pr variabel som ikke matcher databeskrivelsen. Databeskrivelsen kan typisk være en XSD ⁷ -fil.
Anvendelse	Alle statistikker som samler inn data
Formel	For hver variabel: <ul style="list-style-type: none"> Numeriske variable: Verdi = antall enheter med ugyldig numerisk verdi / antall enheter. Ugyldig verdi er for eksempel ikke-numeriske tegn som ikke er desimal eller 1000-skilletegn. Datovariabel: Verdi = antall enheter med ugyldig datoverdi / antall enheter. Verdiene må matche angitt formatmaske og datoene må være gyldige. Boolske variable: Verdi = Antall enheter med ugyldig boolsk verdi / antall enheter
Målverdi	En verdi så nærme 0 som mulig indikerer at data rent teknisk sett er riktige. NB! Denne indikatoren sier ikke noe om nøyaktighet i variablene.
Aggregeringsnivåer og prinsipper	
Tolkning	Indikatoren brukes i Samle inn-fasen for å avdekke om mottatte data stemmer overens med beskrivelsen av data. Indikatoren vil avdekke omfanget av data som rent teknisk sett ikke er riktig i forhold til angitt datatype. <p>Feil av denne typen må håndteres:</p> <ul style="list-style-type: none"> ved at denne typen ugyldige data utelates fra datagrunnlaget og slik sett ikke blir en del av klargjorte data ved at man i dialog med oppgavegiver får tilsendt nye data ved at man korrigerer feilene i klargjøre-prosessen, fortrinnsvis automatisk.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	I utgangspunktet fase 4. Samle inn, men av praktiske årsaker kan man også gjøre dette som en del av fase 5. Klargjøre.
Datatilstand	Inndata
SIMS	Nei
Bruksnivå	Produksjon, eventuelt også taktisk
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	I ISEE er dette implementert i form av såkalte datatype-kontroller

⁷ XSD: XML Schema Definition. Filbeskrivelse av en XML-fil.

2.20. Enhetsfrafall

ID	20
Navn	Enhetsfrafall
Beskrivelse	Andel enheter ikke rapportert i forhold til antall enheter i utvalget som skal rapportere. Rapporteres fra SSB til Finansdepartementet på aggregert nivå.
Anvendelse	Utvalgsbaserte undersøkelser. Statistikknivå: Verdi = 1 - Antall enheter mottatt / Antall enheter i utvalg.
Formel	For rapportering til Finansdepartementet er formelen antall enheter mottatt / antall enheter i utvalget * 100 på SSB-nivå. Statistikknivå: Verdien bør ligge så nærme 0 som mulig.
Målverdi	Finansdepartementet fastsetter måltall for denne indikatoren. Måltallet gjelder aggregert nivå fordelt på oppgavepliktige og frivillige undersøkelser. Det er viktig at SSB når disse måltallene.
Aggregeringsnivåer og prinsipper	Rapporteres til Finansdepartementet på SSB-nivå. Det skilles mellom undersøkelser med opplysningsplikt og frivillige undersøkelser.
Tolkning	Et lavt enhetsfrafall reduserer utvalgsusikkerhet
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	Samle inn
Datatilstand	Inndata
SIMS	A4. Unit non-response - rate
Bruksnivå	Produksjon, taktisk, strategisk
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	"European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition" (Eurostat, 2021), kapittel: S.13.3.

2.21. Partielt frafall

ID	21
Navn	Partielt frafall
Beskrivelse	For hver variabel: Andel med blank verdi i forhold til antall enheter rapportert og som skal fylle ut variabelen.
Anvendelse	Indikatoren kjøres på inndata-tilstanden, det vil si før editering og imputering utføres. I utgangspunktet på utvalgsbaserte skjemaundersøkelser, men indikatoren kan også være relevant for statistikk basert på administrative data.
Formel	For hver (nøkkel)variabel: Verdi = $1 - \frac{\text{Antall enheter med verdi}}{\text{Antall enheter i utvalget}}$.
Målverdi	Verdien bør ligge så nærme 0 som mulig.
Aggregeringsnivåer og prinsipper	
Tolkning	En høy verdi indikerer at det vanskelig å få tak i informasjonen fra oppgavegiverne. For eksempel ved at formuleringer i spørreskjemaer er uklare, eller at informasjon ikke er tilgjengelig fra regnskapssystemer i økonomisk statistikk. En høy verdi på indikatoren kan også være en pekepinn på mulig skjevhet. Når man skal tolke denne indikatoren må man være bevisst på at ikke alle variabler er obligatoriske eller skal fylles ut. Dette gjelder f.eks. ved sporvalg på et Altinn-skjema der et sett med felt kan skjules ut ifra hva man svarer på et spørsmål i skjemaet. Det er derfor viktig med god kunnskap om innsamlingsinstrumentet når man skal tolke denne indikatoren.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	Samle inn
Datatilstand	Inndata
SIMS	A5. Item non-response - rate
Bruksnivå	Produksjon, taktisk
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	"European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition" (Eurostat, 2021), kapittel: S.13.3.

2.22. Datakomplethetsandel

ID	22
Navn	Datakomplethetsandel
Beskrivelse	<p>For hver variabel: Andel med verdi (ikke blank) i forhold til antall enheter i datasettet.</p> <p>Denne indikatoren ligner på indikator 21 – partielt frafall. Forskjellen er at denne kjøres på datatilstandene fra Klargjorte data og utover, altså <i>etter</i> editering og imputering. Mens indikatoren for partielt frafall kjøres på inndata-tilstanden, altså <i>før</i> editering og imputering.</p> <p>Indikatoren kan beregnes for nøkkelvariable, eller på alle variabler i et datasett. Indikatoren har to varianter:</p> <ul style="list-style-type: none"> • 22.1 Produsentorientert: Tar ikke hensyn til dataceller som er prikket eller endret som følge av konfidensialitet • 22.2 Brukerorientert: Tar hensyn til dataceller som er prikket eller endret som følge av konfidensialitet. Relevant for statistikk og utdata.
Anvendelse	På klargjorte data og statistikk.
Formel	<p>Formel, for hver variabel:</p> <ul style="list-style-type: none"> • $\text{Verdi} = \text{Antall med verdi} / \text{Antall enheter i datasett}$. <p>Inputen her er datasett med ferdig klargjort data eller statistikk.</p>
Målverdi	Verdien bør ligge så nærme 1 som mulig.
Aggregeringsnivåer og prinsipper	Verdien er i utgangspunktet på variabelnivå. Kan aggregeres opp på statistikknivå og avdelings- eller organisasjonsnivå.
Tolkning	For nøkkelvariable (klassifikasjonsvariable og statistikkvariable) er det et mål at andelen med verdi er så nærme 1 som mulig, spesielt på den produsentorienterte varianten. Det indikerer at det innsamlede datamaterialet har god kompletthet og/eller at imputering for frafall er utført.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	5. Klargjøre 6. Analysere
Datatilstand	Ferdig klargjorte data Statistikk Utdata
SIMS	R1. Data completeness - rate
Bruksnivå	Produksjon, taktisk
Nivå	Variabel
Kodereferanser	
Annen dokumentasjon	<p>Dokumentasjon av kvalitetsrapporter i ISEE, rapport-id 09. Se vedlegg A.</p> <p>“European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition” (Eurostat, 2021), kapittel: S.12.3.</p>

2.23. Tidsbruk skjema utfylling

ID	24
Navn	Tidsbruk skjema utfylling
Beskrivelse	Tidsbruk på utfylling av svarskjema
Anvendelse	Skjemabaserte undersøkelser
Formel	Verdi = minutter brukt på innhenting av informasjon pluss minutter brukt på utfylling av skjema. Beregne gjennomsnitt eller median for alle respondenter.
Målverdi	Ingen spesiell målverdi. Målverdi vil kunne variere fra skjema til skjema. Men over tid er det et mål at verdien på indikatoren reduseres eller holdes stabil.
Aggregeringsnivåer og prinsipper	
Tolkning	Det at et skjema tar lang tid å fylle ut for respondenten kan være en indikasjon på at skjemaet er vanskelig å forstå eller at det er vanskelig å fremskaffe de opplysningene det spørres etter. Det kan igjen medføre at nøyaktigheten på data som rapporteres blir dårlig ved at verdien estimeres etter skjønn. Lang tidsbruk påvirker også oppgavebyrden, som er en indikator som rapporteres på SSB-nivå til Finansdepartementet.
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet Prinsipp 9: Rimelige krav til oppgavegivere Prinsipp 8: Egnede statistiske prosedyrer
Fase i prosessmodellen	Samle inn
Datatilstand	Kildedata
SIMS	Nei
Bruksnivå	Taktisk
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	

2.24. Potensiell aktualitet

ID	24
Navn	Potensiell aktualitet
Beskrivelse	Antall dager fra statistikken er klar til publisering til den faktisk blir publisert. Formålet med indikatoren er å kunne gi en pekepinn på om det er rom for å forbedre aktualiteten på statistikken.
Anvendelse	Alle statistikker
Formel	Formel: <ul style="list-style-type: none"> Verdi = Tidspunkt for publisering minus tidspunkt for lasting til statistikkbanken. <p>Antall dager fra data blir lastet til statistikkbanken til faktisk publisering</p> <p>Forutsetter at statistikkbanken lagrer tidspunkt for innlasting av data og at denne informasjonen kan tilgjengeliggjøres for kodeverktøy.</p>
Målverdi	En høy verdi indikerer at det kan være potensiale til å forbedre aktualiteten ved at statistikken kan publiseres på et tidligere tidspunkt. Dette fordrer imidlertid at statistikkprodusentene laster til statistikkbanken så fort de anser at statistikken har god nok nøyaktighet til at den faktisk kan publiseres.
Aggregeringsnivåer og prinsipper	
Tolkning	Publiseringstidspunktet for en statistikk styres av statistikkalenderen der publiseringstidspunktet kunngjøres tre måneder før planlagt publisering. Hvis man over tid ser at denne indikatoren øker, eller er stabilt høy, så kan det gi grunnlag for å forbedre aktualiteten ved å fremskynde planlagt publisering ved neste produksjonssyklus.
	Indikatoren må ses i sammenheng med andre indikatorer, for eksempel indikator nr. 12 – Feil rettet etter publisering, for å se om og hvordan andre kvalitetskrav, påvirkes av forbedret aktualitet.
Kvalitetsprinsipp	Prinsipp 13: Aktualitet og punktlighet Prinsipp 10: Kostnadseffektivitet
Fase i prosessmodellen	Formidle
Datatilstand	Utdata
SIMS	
Bruksnivå	Taktisk
Nivå	Statistikk
Kodereferanser	
Annen dokumentasjon	

2.25. Andelen dubletter i datasettet

ID	25
Navn	Andelen dubletter i datasettet
Beskrivelse	<p>En dublett er når enheten i undersøkelsen har flere forekomster i innrapporteringen. Det finnes i hovedsak 3 typer dubletter:</p> <ul style="list-style-type: none"> • <i>Identisk</i> – alle verdiene som er rapportert inn er identiske like • <i>Revidert</i> – noen av verdiene som er rapportert inn er endret • <i>Partiell</i> – verdiene blir delt opp og rapportert i forskjellige omganger <p>Partiell rapportering kan oppstå når det blir rapportert inn forskjellige deler av et år, eller oppsplitting av rapportering av deler av skjema</p>
Anvendelse	<p>Er relevant både for skjemaundersøkelser og for administrative datakilder:</p> <ul style="list-style-type: none"> • Skjemaundersøkelser: Finne omfang av respondenter som rapporterer mer enn en gang • Administrative datakilder: Finne omfang av data der den unike id-en i datamaterialet finnes mer enn en gang.
Formel	Verdi = Antall linjer med ikke unik id / antall linjer med data mottatt
Målverdi	Målverdien er 0, noe som indikerer at det ikke finnes dubletter i datamaterialet.
Aggregeringsnivåer og prinsipper	
Tolkning	<p>Indikatoren brukes til å finne feil i datamaterialet.</p> <p>For skjemaundersøkelser vil et høyt tall bety at det er et stort omfang av flere rapporteringer pr respondent. Det kan indikere at det har vært vanskelig å fylle ut skjema eller at det har vært behov for rekontakt med respondent.</p> <p>For administrative data er høyt omfang av dubletter en indikasjon på at datamaterialet har for dårlig nøyaktighet eller er ufullstendig, eller at samme data er sendt flere ganger.</p>
Kvalitetsprinsipp	<p>Prinsipp 12: Nøyaktighet og pålitelighet</p> <p>Prinsipp 9: Rimelige krav til oppgavegivere</p> <p>Prinsipp 8: Egnede statistiske prosedyrer</p> <p>Prinsipp 7: Gode metoder</p>
Fase i prosessmodellen	Samle inn
Datatilstand	Kildedata, inndata
SIMS	Nei
Bruksnivå	Taktisk, produksjon
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	Dokumentasjon av kvalitetsrapporter i ISEE, rapport-id 08. Se vedlegg A.

2.26. Treffsikkerhet

ID	26
Navn	Treffsikkerhet
Beskrivelse	Andelen enheter som blir tatt ut i en kontroll som blir endret
Anvendelse	Denne indikatorer blir brukt til å analysere kontroller
Formel	Verdi = Antall endringer pga. en kontroll / antall utslag av en kontroll
Målverdi	Målverdien er 100, noe som indikerer at kontrollen finner potensielle feil og alle disse blir korrigert.
Aggregeringsnivåer og prinsipper	
Tolkning	<p>En lav treffsikkerhet for en kontroll, betyr at det er mange kontrollutslag som ikke fører til en endring. Dette er ikke effektivt.</p>
Kvalitetsprinsipp	Prinsipp 12: Nøyaktighet og pålitelighet
Fase i prosessmodellen	Klargjøre
Datatilstand	Klargjorte data
SIMS	Nei
Bruksnivå	Taktisk, produksjon
Nivå	Enhet
Kodereferanser	
Annen dokumentasjon	

Referanser

Eurostat (websiteside): ESS Reference Metadata Reporting Standards. SIMS, Single Integrated Metadata Structure, kan lastes ned fra denne siden. Lenke:

<https://ec.europa.eu/eurostat/documents/64157/4373903/SIMS-2-0-Revised-standards-November-2015-ESSC-final.pdf>

Eurostat (2021): European Statistical System (ESS) handbook for quality and metadata reports — 2021 re-edition. Lenke: <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gg-21-021>

Lovdata (2019): Lov om offisiell statistikk og Statistisk sentralbyrå (statistikkloven). Lenke til Lovdata: <https://lovdata.no/dokument/LTI/lov/2019-06-21-32>

SSB (2017): Retningslinjer for europeisk statistikk. 16. november 2017. Uoffisiell norsk oversettelse. Kan lastes ned fra <https://www.ssb.no/omssb/kvalitet-i-offisiell-statistikk/retningslinjer-for-europeisk-statistikk>

SSB (2019): Generic Statistical Business Process Model GSBPM. Notater 2019/43. Kan lastes ned fra: <https://www.ssb.no/befolkning/artikler-og-publikasjoner/generic-statistical-business-process-model-gsbpm>

SSB (2023): Datatilstander SSB. Interne dokumenter 2023/14.

SSB (2023): Rapport om kvalitet i offisiell statistikk, 2023. Planer og meldinger 2023/6. Kan lastes ned fra: <https://www.ssb.no/omssb/kvalitet-i-offisiell-statistikk/publikasjoner-om-kvalitet-i-offisiell-statistikk>

Statistics Canada (2019): Statistics Canada Quality Guidelines, Sixth edition. Lenke: <https://www150.statcan.gc.ca/n1/pub/12-539-x/12-539-x2019001-eng.htm>

McKenzie R. & Gamba M. (2008): Interpreting the results of Revision Analyses: Recommended Summary Statistics. Lenke: <https://web-archiv.oecd.org/2012-06-15/129056-40315546.pdf>

Vedlegg A: Dokumentasjon av kvalitetsrapporter i ISEE⁸

Tabell A1. Kvalitetsrapporter i ISEE

Id	Kvalitetsrapport	Beskrivelse
01	Andel editeringer i tidsserie	<p>Rapporten gir oversikt over hvor stor andel av enhetene som det har blitt gjort editeringer på i ISEE for en tidsserie av delregistre. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Periode: Hvilken periode informasjonen gjelder for. Format: Årstall, periodenr og periodetype • Delreg_nr: Delregisternummer • Skjema: Skjema-id • Antall med data: Antall enheter for aktuelt delreg_nr/skjema med data i ISEE • Antall editert: Antall enheter for aktuelt delreg_nr/skjema der data er endret i ISEE. Både manuelle og automatiske endringer telles med. • Antall enheter: Antall enheter for aktuelt delreg_nr / skjema i SFU (Antall enheter i utvalget). • Andel editert: Antall editert / Antall enheter <p>Rapporten er sortert synkende etter periode, og deretter stigende etter skjema-id.</p>
02	Andel editeringer, manuelle, i tidsserie	<p>Rapporten gir oversikt over hvor stor andel av enhetene som det har blitt gjort manuelle editeringer på i ISEE for en tidsserie av delregistre. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Periode: Hvilken periode informasjonen gjelder for. Format: Årstall, periodenr og periodetype • Delreg_nr: Delregisternummer • Skjema: Skjema-id • Antall med data: Antall enheter for aktuelt delreg_nr / skjema med data i ISEE • Antall editert: Antall enheter for aktuelt delreg_nr / skjema der data er endret i ISEE. Både manuelle og automatiske endringer telles med. • Antall enheter: Antall enheter for aktuelt delreg_nr / skjema i SFU (Antall enheter i utvalget). • Andel editert: Antall editert / Antall enheter <p>Rapporten er sortert synkende etter periode, og deretter stigende etter skjema-id.</p>
03	Andel editert, ikke-editert og automatisk editert	<p>Rapporten gir oversikt over antall andel manuelt editert, automatisk editert og uendret gruppert på delregister og skjema innenfor et angitt tidsintervall. Det vil si at alle data som er satt inn eller endret (automatisk eller manuelt) innenfor det angitte datointervallet telles. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Delreg: Navn på delregister • Periode: Hvilken periode informasjonen gjelder for. Format: Årstall, periodenr og periodetype • Delreg_nr • Skjema: Skjema-id • Automatisk: Antall automatisk prosesserte data • Manuell: Antall manuelt prosesserte data • Uendret: Antall uendrede data. Dette er altså data som er lastet inn uten å ha blitt endret i ISEE i datointervallet. • Total: Totalt antall prosesserte data • Andel automatisk: Antall automatisk / total • Andel manuell: Antall manuell / total • Andel uendret: Antall uendret / total <p>Rapporten er sortert stigende etter delregisternavn og periode</p> <p>NB! Denne rapporten sendes ut regelmessig med månedlig ISEE Nyhetsbrev med fra-dato og til-dato for foregående måned.</p>
04	Antall editeringer pr variabel	<p>Rapporten gir oversikt over antall editeringer pr. variabel for et gitt delregisternr. Input-parameter er Delregisternummer. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Skjema: Skjema-id • Felt_id: Navn på felt/variabel • Antall: For aktuelt felt/variabel - totalt antall forekomster av denne variabelen for alle enheter på aktuelt delreg/skjema i ISEE sin historikk-tabell, som logger innsetting (insert), endring(update) og sletting (delete). Summen av de tre neste kolonnene som viser hhv antall insert, update og delete. • Antall ins: Delmengde av foregående kolonne (Antall) og omhandler innsetting (insert)

⁸ Integrert System for Editering og Estimering, SSBs fellessystem for klargjøre-prosessen i GSBPM

Id	Kvalitetsrapport	Beskrivelse
		<ul style="list-style-type: none"> • Antall upd: Delmengde av foregående kolonne (Antall) og omhandler endring (update). Både manuelle og automatiske endringer telles. • Antall del: Delmengde av foregående kolonne (Antall) og omhandler sletting (delete) • Listefelt: Om den aktuelle variabelen er et listefelt (flerradet felt). Vil kun være Ja for skjemaer med dynamiske lister der antall rader varierer fra enhet til enhet. • Antall rapporter: Totalt antall innrapporteringer for aktuelt delregnr/skjema • Antall rapportlinjer: Totalt antall rader for listefelt for alle rapporter.
		Rapporten er sortert synkende etter Antall upd (endringer)
05	Antall editeringer pr variabel i tidsserie	<p>Rapporten gir oversikt over antall editeringer pr variabel for en tidsserie av delregistre. Input-parameter er Delregisternavn. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Delregnr: Delregisternummer • Skjema: Skjema-id • Felt_id: Navn på felt/variabel • Antall: For aktuelt felt/variabel - totalt antall forekomster av denne variabelen for alle enheter på aktuelt delreg/skjema i ISEE sin historikk-tabell, som logger innsetting (insert), endring(update) og sletting (delete). Summen av de tre neste kolonnene som viser hhv antall insert, update og delete. • Antall ins: Delmengde av foregående kolonne (Antall) og omhandler innsetting (insert) • Antall upd: Delmengde av foregående kolonne (Antall) og omhandler endring (update). Både manuelle og automatiske endringer telles. • Antall del: Delmengde av foregående kolonne (Antall) og omhandler sletting (delete) • Listefelt: Om den aktuelle variabelen er et listefelt (flerradet felt). Vil kun være Ja for skjemaer med dynamiske lister der antall rader varierer fra enhet til enhet. • Antall rapporter: Totalt antall innrapporteringer for aktuelt delregnr/skjema • Antall rapportlinjer: Totalt antall rader for listefelt for alle rapporter.
		Rapporten er sortert stigende etter delregisternummer, skjema og Antall upd, synkende)
06	Antall editeringer, manuelle, pr variabel	<p>Rapporten gir oversikt over antall editeringer pr variabel for et gitt delregisternr. Input-parameter er Delregisternummer. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Skjema: Skjema-id • Felt_id: Navn på felt/variabel • Antall: For aktuelt felt/variabel - totalt antall forekomster av denne variabelen for alle enheter på aktuelt delreg/skjema i ISEE sin historikk-tabell, som logger innsetting (insert), endring(update) og sletting (delete). Summen av de tre neste kolonnene som viser hhv antall insert, update og delete. • Antall ins: Delmengde av foregående kolonne (Antall) og gjelder antall ganger det manuelt er lagt inn en verdi på et felt der hvor verdi manglet. • Antall upd: Delmengde av foregående kolonne (Antall) og gjelder antall ganger en verdi på et felt er endret manuelt. • Antall del: Delmengde av foregående kolonne (Antall) og gjelder antall ganger en verdi på et felt er fjernet manuelt. • Listefelt: Om den aktuelle variabelen er et listefelt (flerradet felt). Vil kun være Ja for skjemaer med dynamiske lister der antall rader varierer fra enhet til enhet. • Antall rapporter: Totalt antall innrapporteringer for aktuelt delregnr/skjema • Antall rapportlinjer: Totalt antall rader for listefelt for alle rapporter.
		Rapporten er sortert synkende etter Antall upd (endringer)
07	Antall kontrollutslag	<p>Rapporten gir oversikt over antall feilkontrollutslag for et gitt delregisternr. Input-parameter er Delregisternummer. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Delreg_nr: Delregisternummer • Skjema: Skjema-id • Feiltekst: Tekst på aktuell kontroll, definert i Kontrollbygger • Felt_id: Navn på felt/variabel tilknyttet aktuell kontroll • Antall: Antall ganger denne kontrollen har slått ut totalt på alle enheter i delregisteret/skjemaet • Listefelt: Om den aktuelle variabelen tilknyttet kontrollen er et listefelt (flerradet felt). Vil kun være Ja for skjemaer med dynamiske lister der antall rader varierer fra enhet til enhet. • Antall rapporter: Totalt antall innrapporteringer for aktuelt delregnr/skjema • Antall rapportlinjer: Totalt antall rader for listefelt for alle rapporter.
		Rapporten er sortert etter Antall, fra størst til minst
08	Antall rapporteringer	<p>Rapporten gir oversikt over hvor mange ganger rapportørene har sendt inn skjema gruppert på antall skjemainnsendinger. Input-parameter er Delregisternummer. Rapporten gir følgende info:</p>

Id	Kvalitetsrapport	Beskrivelse
		<ul style="list-style-type: none"> • Delreg_nr: Delregisternummer • Skjema: Skjema-id • Antall innsendinger: Antall innsendinger pr rapportør • Forekomst: Antall forekomster av antall innsendinger • Totalt antall innsendinger: Totalt antall innsender for aktuelt delreg_nr og skjema
09	Fullstendighet av datasett	<p>Rapporten er sortert stigende etter antall innsendinger</p> <p>Rapporten gir oversikt over fullstendighet av variablene i datasettet. Alle variablene listes opp nedover og rapporten gir informasjon om hvor mange ganger denne variabelen har verdi og hva som er maks antall verdier (summen av utfylte og ikke-utfylte).</p> <ul style="list-style-type: none"> • Delreg_nr: Delregisternummer • Skjema: Skjema-id • Felt_id: Navn på felt/variabel • Listefelt: Om den aktuelle variabelen er et listefelt (flerradet felt). Vil kun være Ja for skjemaer med dynamiske lister der antall rader varierer fra enhet til enhet. • Felt_Type: Type felt - Numerisk (NUMBER eller DESIMAL), Tekst (CHAR eller KOMMENTAR), Kodeliste/klassifikasjon (LISTE eller CHECKBOX) • Antall_med_verdi: Antall forekomster med verdi for denne variabelen i dette delregisteret • Antall_totalt: Maks antall mulige forekomster denne variabelen i dette delregisteret
10	Oversikt datasett	<p>Rapporten er sortert etter skjema og antall_med_verdi synkende</p> <p>Rapporten gir oversikt over antall enheter, antall variable totalt antall verdier i datasettet fordelt på type variabel. Rapporten gir følgende info:</p> <ul style="list-style-type: none"> • Delreg_nr: Delregisternummer • Skjema: Skjema-id • Feltype: Hvilken feltype; CHAR, NUMBER, LISTE, DESIMAL, CHECKBOX • Antall felt: Antall felt av aktuell type • Antall listefelt. Av antallet i forrige kolonne, hvor mange av disse som er felt i en dynamisk liste • Antall statiske felt. Av antallet felt, hvor mange av disse som er felt på den statiske delen av skjemaet • Antall verdier: Antall felt med verdier for denne typen. • Antall enheter: Antall enheter med data i delregisteret

Data, metadata og prosessdata i ISEE ligger til grunn for kvalitetsrapportene:

- Alle hendelser på data i ISEE blir logget i en egen historikk/logg-tabell. Med hendelser menes innsetting (insert), oppdatering (update) og sletting (delete) av data. Logging skjer på variabelnivå og gjør at man har full oversikt over livssyklusen fra innlasting av data via editering til ferdig klaggjorte data som kan trekkes ut fra ISEE. Det som logges er bl.a. tidspunkt, hvem (initialer) som trigget hendelsen og type hendelse (insert, update, delete). Slik logginformasjon omtales som prosessdata.
- Metadata. Alle kontroller defineres og dokumenteres i form av metadata. Alle kontrollutslag logges, og alle endringer i kontroller (metadata) logges.
- Alle data og metadata i ISEE er lagret på en standardisert måte i felles data- og metadatastrukturer

Summen av de tre punktene over gjør det mulig å hente ut kvalitetsrapporter fra ISEE som gir oversikt over editeringsprosessen. Rapportene gir bilde av f.eks. hvilke variable det gjøres flest endringer på. Rapportene kan gi indikasjoner hvor det er problemer med datakvalitet, og være et viktig hjelpemiddel til å f.eks. kunne avdekke svakheter og problemer med Altinn-skjemaer. Rapportene anbefales brukt til evaluering av statistikkproduksjonsprosessen, jf. prosess 8 evaluere i prosessmodellen GSBPM.