



*Stein Bakke, Christoffer Berge, Ole Christian Erikstad
og Ole-Petter Kordahl*

**Tiltak for å forbedre dialogen
mellom fagseksjonene og
Kvartalsvis nasjonalregnskap**

Notater

Sammendrag

Kvartalsvis nasjonalregnskap (KNR) skal gi en helhetlig oversikt over den kortsiktige utviklingen i norsk økonomi, og er et viktig verktøy for overvåking og analyse av konjunkturforløpet. KNR lages av seksjon for Nasjonalregnskap (S930), og korttidsstatistikker fra flere av SSBs fagseksjoner utgjør en viktig del av datagrunnlaget.

Bakgrunnen for dette prosjektet har vært et ønske om å forbedre dialogen mellom fagseksjonene og KNR. En god dialog avhenger av en felles forståelse av datagrunnlaget, og prosjektgruppen har sett nærmere på hvordan man i større grad kan oppnå dette.

Tidlig i prosjektet kom vi fram til at det trengs bedre rutiner for utveksling av metadata, det vil si *data om data*. Metadata gir verdifull informasjon til S930 om korttidsstatistikkene som skal innarbeides i KNR-systemet. Prosjektgruppen har derfor jobbet med å kartlegge de mest relevante metadata, samt utarbeidet et forslag til hvordan metadata på en enkel måte kan formidles mellom fagseksjonene og S930.

Vi har på bakgrunn av et utvalg av de korttidsstatistikkene som inngår i KNR, egne erfaringer og samtaler med nøkkelpersoner hos KNR, foreslått følgende typer metadata: opplysninger om statistikken, opplysninger om dataleveransen, opplysninger om mottaker av dataleveransen, opplysninger om feilkilder og usikkerhet, samt opplysninger om endringer som kan påvirke beregningene i KNR.

Med opplysninger om statistikken mener vi for eksempel statistikkens navn, navn og telefonnr til statistikkansvarlig ved fagseksjonen, samt variabelnavn og variabeltype. Når det gjelder opplysninger om dataleveransen, så er dette for eksempel dato og klokkeslett for leveransen. Opplysninger om mottaker av dataleveransen vil være navn, telefonnummer og e-postadresse til aktuelle personer ved S930.

Når det gjelder feilkilder og usikkerhet, så har vi skilt mellom utvalgsbaserte statistikker og registerbaserte statistikker. For utvalgsbaserte statistikker er dekningsgrad, standardavvik, frafallsfeil og prediksjonsfeil viktige indikatorer for å si noe om kvaliteten på datamaterialet. For registerbaserte statistikker foreslår prosjektgruppa blant annet indikatorer som beskriver endringer i meldingsgangen til registrene, samt uoppbyggingsandeler for variablene som inngår i dataleveransen.

Prosjektgruppa foreslår at det etableres et elektronisk system for utveksling av metadata. Dette skal fungere som et viktig supplement til dagens arbeidsmøter og kontakt via telefon og e-post. I et slikt system skal det være mulig å registrere og ta ut informasjon om de metadata som er nevnt over. Vi har laget en skisse til et slikt system, og vi har lagt vekt på at systemet ikke skal være for ressurskrevende å lage, at det skal være brukervennlig og at det skal være lett å administrere. Derfor foreslår vi at systemet

bygges på samme måte som applikasjonen for elektronisk registrering av nye frivningsdatoer i Statistikkalenderen.

Prosjektet er gjennomført som en del av Byråskolens utviklingsprogram i Prosjekt- og teamarbeid 2007/2008. Prosjektdeltakere har vært Ole Christian Erikstad (prosjektleder), Ole-Petter Kordahl, Stein Bakke og Christoffer Berge.

Innhold

1 Innledning	5
1.1 Bakgrunn for prosjektet	5
1.2 Formål og avgrensninger	5
1.3 Prosjektorganisering	6
1.4 Notatets struktur	7
2 Bakgrunn	8
2.1 Kvartalsvis nasjonalregnskap (KNR)	8
2.2 Korttidsstatistikk	8
2.3 Overføring av korttidsstatistikk til KNR	9
2.4 Arbeidsmøter	9
2.5 utfordringer	9
2.6 Tidligere arbeider	9
3 Korttidsstatistikker som input i KNR	11
3.1 Bakgrunn for valg av korttidsstatistikker	11
3.2 Beskrivelse av de fire korttidsstatistikkene	12
3.2.1 Detaljomsetningsindeksen	12
3.2.2 Kvartalsvis investeringsstatistikk for industri, bergverksdrift og kraftforsyning	13
3.2.3 Kvartalsvis omsetningsindeks for samferdsel og reiseliv	13
3.2.4 Sysselsatte og arbeidsledige på korttidsopphold i Norge	14
4 Kvalitet, feilkilder og usikkerhet	15
4.1 Hva er kvalitet?	15
4.1.1 Relevans	16
4.1.2 Nøyaktighet	16
4.1.3 Aktualitet og punktlighet	16
4.1.4 Tilgjengelighet og klarhet	16
4.1.5 Sammenlignbarhet med andre statistikker	16
4.1.6 Sammenheng mellom statistikker	16
4.2 Feilkilder og usikkerhet	16
4.2.1 Utvalgsfeil	17
4.2.2 Registerfeil (dekningsfeil)	17
4.2.3 Målefeil	18
4.2.4 Frafallsfeil	18
4.2.5 Bearbeidingsfeil	19
4.2.6 Modellfeil	19
5 Metadata som grunnlag for dialog og samarbeid	20
5.1 Utgangspunkt for arbeidet	20
5.2 Kartlegging av relevante metadata	20
5.2.1 Opplysninger om statistikken	20
5.2.2 Opplysninger om dataleveransen	21
5.2.3 Opplysninger om mottager av dataleveransen	22
5.2.4 Opplysninger om feilkilder og usikkerhet	22
5.2.5 Opplysninger om endringer som kan påvirke beregningene i KNR	24
6 Forslag til system for utveksling av metadata mellom fagseksjonene og Seksjon for nasjonalregnskap	25
6.1 Behov for et informasjonssystem	25
6.2 Skisse til system for elektronisk lagring og formidling av metadata	26

6.2.1	Overordnet systembeskrivelse.....	26
6.2.2	Kriterier for valg av løsning.....	26
6.2.3	Prosjektgruppas forslag.....	26
6.2.4	Formidling av opplysningene.....	29
6.2.4.1	Formidling av opplysninger via e-post.....	29
6.2.4.2	Direkte uttrekk fra databasen.....	30
6.2.5	Gevinster ved å etablere en informasjonsbank.....	32
7	Anbefalinger for videre arbeid.....	33
7.1	Anbefalinger som krever lite ressurser.....	33
7.2	Anbefalinger som er mer ressurskrevende.....	33
7.3	ISEE – Integrert system for editering og estimering.....	34
	Begrepsforklaringer.....	35
	Referanser.....	37
	Vedlegg A. Prosjektskrivet.....	39
	Vedlegg B: Fagseksjoner som leverer data til KNR.....	42
	Vedlegg C: Korttidsstatistikker som er sentrale i KNR.....	43
	Vedlegg D. Eksempel på tidsplan for KNR.....	44

1 Innledning

1.1 Bakgrunn for prosjektet

Flere av fagseksjonene i Statistisk sentralbyrå (SSB) leverer hvert kvartal dataserier til Seksjon for nasjonalregnskap (S930) til bruk i produksjonen av Kvartalsvis nasjonalregnskap (KNR). Disse dataseriene er hovedsakelig selvstendige statistikker, som blir publisert av den enkelte fagseksjon.

I forbindelse med dataleveransene fra fagseksjonene til S930, oppstår det en del utfordringer i forhold til intern kommunikasjon, det være seg felles forståelse av datagrunnlaget, dokumentasjon av avvik i forhold til tidligere publiseringer, samt at de opprinnelige dataseriene fra fagseksjonene kan få en noe ulik utvikling når de har blitt innarbeidet i KNR-systemet.

På bakgrunn av dette foreslo seksjonssjef Leiv Solheim, ved Seksjon for samferdsels- og reiselivsstatistikk, at denne problemstillingen kunne inngå som et prosjekt i forbindelse med Byråskolens utviklingsprogram i prosjekt- og teamarbeid (Proteam), som gjennomføres hvert år. Dette prosjektet er derfor gjennomført som ett av seks prosjekt i Proteam 2007/08.

1.2 Formål og avgrensninger

Som Proteam-prosjekt, er prosjektet underlagt en del faste rammer. Prosjektgruppen skal blant annet kun bestå av fire personer og prosjektgruppen rapporterer til en styringsgruppe bestående av prosjektmedlemmenes respektive seksjonssjefer.

I det opprinnelige prosjektskrivet, som ble tildelt prosjektgruppen på den første samlingen i oktober 2007, bestod en vesentlig del av prosjektet i å sammenlikne fagseksjonenes dataserier før og etter at de har blitt innarbeidet i KNR. Ut fra dette skulle vi måle eventuelle avvik som oppstår, samt årsaksforklare disse. I tillegg skulle vi se på kvaliteten til de enkelte fagstatistikkene.

Dette prosjektskrivet forutsatte at prosjektgruppen skulle inkludere en person fra Seksjon for nasjonalregnskap. Da så ikke ble tilfelle, ble prosjektskrivet revidert. Måling av avvik ble noe nedprioritert, samtidig som de nye resultatmålene ble mer knyttet mot interaksjonen mellom fagseksjonene og KNR. Vi mente også at de nye resultatmålene på en bedre måte reflekterte prosjektets tittel, *Tiltak for å forbedre dialogen mellom fagseksjonene og KNR*. I det reviderte prosjektskrivet tolker vi ordet *dialog* ikke bare som kommunikasjon i arbeidsmøter og direkte kontakt mellom fagseksjonene og KNR, men også som formidling av metadata, det vil si data om data.

Det nye prosjektskrivet (se vedlegg A) ble godkjent av styringsgruppa i desember 2007 og har følgende målformuleringer.

Effekt mål:

- En bedre synliggjøring av kvaliteten på korttidsstatistikken.
- Å skape en felles forståelse av datagrunnlaget mellom fagseksjonene og KNR.
- En bedre og sikrere tolkning av avvik mellom fagstatistikkene og KNR-tallene.

Resultat mål:

- Utarbeide et opplegg for kvalitetsbeskrivelse av et utvalg korttidsstatistikker som inngår i datagrunnlaget til KNR.
- Avdekke mulige kilder til avvik ved innarbeiding av korttidsstatistikk i KNR-systemet, samt vurdere behovet for intern formidling av slik informasjon.
- Skrive en rapport i serien Notater.

Det første resultatmålet innebærer en kartlegging av metadata som er av interesse for fagseksjonene og KNR ved dataleveranser. I tillegg må indikatorer tilknyttet nøyaktigheten til data som overleveres KNR identifiseres. I forbindelse med den nevnte omarbeidingen av prosjektskrivet, ble det satt spørsmålsteget med hvorvidt prosjektgruppen ville ha ressurser til å kunne måle kilder til avvik ved innarbeiding i KNR, slik som nevnt i det andre resultatmålet. Dette punktet ble likevel opprettholdt, dog noe nedprioritert i forhold til det opprinnelige prosjektskrivet.

Grunnet tidspress ble det i etterkant, og i samråd med styringsgruppa, bestemt at vi *ikke* skulle arbeide med kilder til avvik.¹ I tillegg forutsetter måling av avvik at man også sammenlikner i forhold til det årlige nasjonalregnskapet samt strukturstatistikkene. Prosjektgruppen hadde allerede i startfasen av prosjektet bestemt at man kun skulle se på korttidsstatistikker.

1.3 Prosjektorganisering

Prosjektgruppa har bestått av:

- Ole Christian Erikstad, Seksjon for bygg- og tjenestestatistikk, S460 (prosjektleder)
- Ole-Petter Kordahl, Seksjon for samferdsels- og reiselivsstatistikk, S440
- Stein Bakke, Seksjon for industristatistikk, S470
- Christoffer Berge, Seksjon for arbeidsmarkedsstatistikk, S260

¹ Styringsgruppemøte 07.05.08.

Kvalitetslos har vært:

- Harald Utne, Seksjon for befolkningsstatistikk, S320

Styringsgruppa har bestått av seksjonssjefene ved de involverte seksjonene:

- Leiv Solheim, Seksjon for samferdsels- og reiselivsstatistikk, S440 (leder av styringsgruppa)
- Roger Jensen, Seksjon for bygg- og tjenestestatistikk, S460
- Lasse Sandberg, Seksjon for prisstatistikk, S240² (fram til 1. februar 2008)
- Lise Dalen Mc Mahon, Seksjon for industristatistikk, S470 (fra og med 1. februar 2008)
- Helge N. Næsheim, Seksjon for arbeidsmarkedsstatistikk, S260

Prosjektet har benyttet seg av følgende referansepersoner:

- Ann Lisbet Brathaug, Seksjonssjef, Seksjon for nasjonalregnskap, S930
- Øystein Berge, Seksjon for nasjonalregnskap, S930
- Jennifer Holgate, Seksjon for nasjonalregnskap, S930
- Jan Henrik Wang, Seksjon for industristatistikk, S470

1.4 Notatets struktur

I kapittel 2 beskrives Det kvartalsvise nasjonalregnskapet (KNR) og informasjonsutvekslingen mellom fagseksjonene og KNR. Kapittel 3 beskriver utvalgte statistikker (Kvartalsvis investeringsstatistikk, Detaljomsetningsindeksen, Omsetningsindeks for samferdsel og reiseliv og Sysselsatte og arbeidsledige på korttidsopphold). Kvalitet, feilkilder og usikkerhet er temaet i kapittel 4, mens kapittel 5 forklarer hvordan metadata kan bidra til å bedre dialogen mellom fagseksjonene og Seksjon for nasjonalregnskap. I kapittel 6 beskrives et forslag til informasjonsutveksling mellom fagseksjonene og KNR. Forslag til videre arbeid er lagt til kapittel 7.

² Stein Bakke var opprinnelig ansatt ved seksjon 240 med Lasse Sandberg som seksjonssjef. Etter omorganiseringer i SSB fra årsskiftet, ble han overført til seksjon 470. Bakkes nye sjef, Lise Dalen Mc Mahon, overtok derfor Lasse Sandbergs plass i styringsgruppen fra og med 1. februar 2008.

2 Bakgrunn

Dette kapitlet beskriver KNR og arbeidsprosessene mellom KNR og fagseksjonene. For en detaljert beskrivelse av KNR henvises det til Evensen (2004) og *Om statistikken* på SSB sine nettsider.

2.1 Kvartalsvis nasjonalregnskap (KNR)

KNR utarbeides av Seksjon for nasjonalregnskap (S930) og frigis 50 dager etter kvartalets utløp. Statistikken gir en helhetlig oversikt over den kortsiktige utviklingen i norsk økonomi og er et viktig verktøy for overvåking og analyse av konjunkturforløpet. Konsistente kvartalstall foreligger tilbake til 1978.

KNR omfatter alle størrelser fra den delen av nasjonalregnskapet som representerer realøkonomiske forhold, det vil si vare- og tjenestestrømmer, samt innsats og avlønning av arbeid og kapital. KNR-størrelsene presenteres i tabeller der verdiene oppgis i løpende og faste priser. Bruttonasjonalprodukt (BNP) og bruttonasjonalprodukt for Fastlands-Norge er viktige makroøkonomiske størrelser som estimeres i statistikken.

KNR bygger på to hovedpilarer; endelig nasjonalregnskap og korttidsstatistikk. Siste årlige nasjonalregnskapet omtales som basisåret for KNR og brukes som nivåmessig grunnlag samt kilde til strukturinformasjon om økonomien. Korttidsstatistikkene benyttes som indikatorer på utviklingen i nivåtallene og utgjør dermed grunnlaget for fremskriving av disse.

Arbeidet med KNR er fordelt på i overkant av 20 personer ved S930.

- En person fungerer som driftsansvarlig. Denne rollen innebærer administrasjon av prosessen med å lage KNR samt faglig ansvar for resultater og publiseringer.
- To personer er teknisk ansvarlige. Av disse har en hovedansvar for innarbeiding av data i KNR-systemet, mens den andre tar seg av endringer i bakenforliggende programmer, modeller etc.
- De øvrige involverte er næringsberegnerne. Disse har som oppgave å tilrettelegge korttidsstatistikk for innarbeiding i KNR-sytemet.

2.2 Korttidsstatistikk

Med korttidsstatistikk menes statistiske undersøkelser med en hyppighet på 3 måneder eller kortere (kvartal, termin, måned uke). Ved fagseksjonene i SSB produseres et stort antall korttidsstatistikker, og

mange inngår i KNR-beregningene.³ Statistikkansvarlig ved fagseksjonene fungerer som kontaktperson mot S930.

2.3 Overføring av korttidsstatistikk til KNR

En del av korttidsstatistikkene som inngår i KNR samles inn ved direkte tilgang til databaser administrert av SSB sine fagseksjoner. Alternative metoder er elektronisk overføring av data fra fagseksjonene til databaser administrert av S930 eller overføring av datafiler via e-post.

2.4 Arbeidsmøter

Ved hvert kvartal utdeler S930 en tidsplan⁴ for KNR. Denne angir datoer for beregninger og dato for arbeidsmøter med fagseksjonene. Hensikten med arbeidsmøtene er at fagstatistikerne skal få mulighet til å komme med supplerende informasjon. S930 skriver referat fra arbeidsmøtene.

2.5 Utfordringer

Fravær, utskifting av medarbeidere eller omfordeling av ansvarsområder på seksjonene kompliserer dialogen på tvers av seksjonene. For å forebygge at kunnskap går tapt, er det viktig å dokumentere hva som blir gjort og hvorfor man gjør det. Læringsprosessen går da raskere for de som tar over, og det blir mindre rom for misforståelser. Informasjonsomfanget i KNR-prosessen kan gjøre det vanskelig å få full oversikt. Dette er noe av bakgrunnen for at S930 skriver og distribuerer referater fra arbeidsmøtene med fagseksjonene.

2.6 Tidligere arbeider

I 2003 gjennomførte SSB prosjektet *Kvalitetsarbeid knyttet til Kvartalsvis nasjonalregnskap*. Arbeidet ble ledet av Trude Nygård Evensen og er beskrevet i Notat 2004/42. Målsetningene i prosjektet var:

1. Beskriving av delprosesser i systemet, blant annet ved hjelp av prosess- og organisasjonskart.
2. Evaluering av delprosessene.
3. Utarbeiding av overordnet plan for videre arbeider med vekt på kvalitet og kvalitetssikring.
4. Gjennomføre enkelte kvalitetsfremmende endringer som er enkle å implementere som en del av prosjektet.

³ Navnet på de fagseksjoner som leverer tall til KNR, samt de mest sentrale korttidsstatistikkene er listet opp i henholdsvis vedlegg B og C.

⁴ I vedlegg D ligger det et eksempel på en tidsplan for KNR.

I kjølvannet av dette prosjektet er det gjennomført flere kvalitetsfremmende arbeider, blant annet for olje- og gassnæringen (Notat 2006/8, Notat 2006/9) og for investeringer i kraftforsyning (Notat 2007/25). Formålet med disse arbeidene har vært å forbedre konsistensen i tallene for produksjon og anvendelse (eksport) av olje og gass, samt å få bedre konsistens mellom korttidsstatistikk og strukturstatistikk for investeringer i kraftforsyning.

3 Korttidsstatistikker som input i KNR

I prosjektet har vi valgt ut fire av korttidsstatistikkene som inngår KNR, og sett nærmere på hvilke problemstillinger disse står overfor i dialogen med seksjon 930. Først begrunner vi valget av de fire statistikkene. Videre følger en kort beskrivelse av disse, samt hvordan de brukes i KNR-beregningene. For en mer detaljert innføring henvises det til Evensen (2004) og *Om statistikken* på SSB sine nettsider.

3.1 Bakgrunn for valg av korttidsstatistikker

Korttidsstatistikk fra fagseksjonene spiller en sentral rolle i dette prosjektet, da de sammen med det årlige Nasjonalregnskapet utgjør det viktigste grunnlaget for KNR. Gitt prosjektets tidsbegrensning valgte vi å ikke se nærmere på forholdet mellom KNR og det årlige nasjonalregnskapet. I tillegg valgte vi å kun se på fire av korttidsstatistikkene, basert på følgende kriterier:

1. De må inngå i KNR-beregningene.
2. Det bør være forskjellige typer korttidsstatistikk.
3. Det bør ikke være for tidkrevende å sette seg inn i dem.

Bakgrunnen for det andre kriteriet var at det ville være lite hensiktsmessig for prosjektet hvis man tok utgangspunkt i nærmest identiske undersøkelser. Jo større bredde som de fire statistikkene til sammen representerer, desto mer relevant blir prosjektets konklusjoner i forhold til dialogen med KNR. Det tredje kriteriet er en direkte følge av at prosjektet hadde begrenset med ressurser. Det ville derfor være best å kunne bruke statistikker som prosjektets medlemmer allerede kjenner til. Følgene korttidsstatistikker ble valgt ut:

- Detaljomsetningsindeksen (DOI)
- Kvartalsvis investeringsstatistikk for industri, bergverksdrift og kraftforsyning (KIS)
- Kvartalsvis omsetningsindeks for samferdsel og reiseliv (SROI)
- Sysselsatte og arbeidsledige på korttidsopphold i Norge (KORTSYS)

Det var noe diskusjon om hvorvidt undersøkelsen *Sysselsatte og arbeidsledige på korttidsopphold i Norge* kunne regnes som en korttidsstatistikk, da dette i utgangspunktet er en årlig undersøkelse. Av mangel på en tilsvarende kvartalsvis statistikk, og siden det arbeides med å lage et opplegg for å produsere kvartalsvise tall, valgte vi likevel å ta denne med.

Tabell 3-1 oppsummerer de utvalgte korttidsstatistikkene:

Tabell 3-1 Oversikt over viktige egenskaper ved de utvalgte statistikkene

	Innsamlingsmetode	Hypighet	Variabel	Variabeltype	Enhet
DOI	utvalgsbasert	måned	omsetning	verdi (volum)	indeks
KIS	utvalgsbasert	kvartal	investering	verdi	nivå
SROI	registerbasert	kvartal	omsetning	verdi	indeks
KORTSYS	registerbasert	år	sysselsetting	volum	nivå

3.2 Beskrivelse av de fire korttidsstatistikkene

Nedenfor følger en kort beskrivelse av de fire korttidsstatistikkene. For hver statistikk har vi i tillegg beskrevet hvordan disse brukes i KNR. Et konkret forslag fra prosjektgruppen er at det, for de statistikker dette gjelder, burde legges en link i *Om statistikken* hetende *Bruk i KNR*. I henhold til malen for *Om statistikken* vil det være naturlig å legge en slik link under avsnitt 6.2 *Sammenheng med annen statistikk*. Hensikten er at brukerne enkelt skal kunne finne mer informasjon om hvordan statistikkene inngår i KNR.

3.2.1 Detaljomsetningsindeksen

Formålet med detaljomsetningsindeksen (DOI) er å beskrive verdi- og volumutviklingen i detaljomsetningen, unntatt motorkjøretøyer og bensin. Detaljomsetningen er den viktigste komponenten i beregningen av konsum i husholdningene. I henhold til standard for næringsgruppering⁵ (SN2002) omfatter populasjonen næring 52.1-52.6. Det vil si *Detaljhandel, unntatt handel med motorkjøretøyer og motorsykler, reparasjon av husholdningsvarer og varer til personlig bruk*.

Data innhentes ved hjelp av postalt skjema og undersøkelsen er underlagt svarplikt. I tillegg benyttes Skattedirektoratets oppgaveregister, samt opplysninger fra SSBs eget Bedrifts- og foretaksregister. Publisering skjer senest 30 dager etter månedens utløp.

DOI i KNR

Detaljomsetningsindeksen brukes som indikator til å beregne varekonsumet. Først og fremst er den input til varekonsumindeksen (VKI), som publiseres av Seksjon for nasjonalregnskap, og som legges direkte inn i det kvartalsvise nasjonalregnskapet.

Varekonsumindeksen er noe mer omfattende enn DOI da den, i tillegg til detaljomsetning, inkluderer kjøp av blant annet bensin fra bensinstasjoner, kjøp av førstegangsregistrerte kjøretøy, samt tall for forbruk av elektrisitet og brensel. Dette kan medføre en noe ulik utvikling i de to indeksene.

Siden DOI publiseres månedlig, og etter 30 dager, leveres tallene i god tid i forhold til KNR-publisering.

3.2.2 Kvartalsvis investeringsstatistikk for industri, bergverksdrift og kraftforsyning

Kvartalsvis investeringsstatistikk (KIS) er en utvalgsbasert undersøkelse som beregner nivå- og endringstall for investeringer innen industri, bergverksdrift og kraftforsyning. I henhold til SN2002 omfatter populasjonen næringene 10, 13 og 14, som er *bergverksdrift*, næringene 15-37, som er *industri*, samt næring 40 som er *kraftforsyning*. I undersøkelsen defineres investeringer som *Alle anskaffelser av nye varige driftsmidler som har en brukstid på 1 år eller mer; gjelder også maskiner mv. som fremstilles til eget bruk samt større ombyggingsarbeider.*

Investeringene skal rapporteres i det kvartalet utstyret er mottatt, selv om det ikke er betalt eller aktivert i regnskapet, og fordeles på kategoriene maskiner, biler og andre transportmidler samt bygg og anlegg. Undersøkelsen er underlagt svarplikt og offentliggjøres om lag 8 uker etter det aktuelle kvartalets utløp.

KIS i KNR

I KNR-beregningene benyttes 12-måneders-endringen i KIS til å fremskrive nivå-tallene for utførte investeringer innen industri, bergverksdrift og kraftforsyning fra det siste tilgjengelige nasjonalregnskapet. Tidlig i prosessen er det endringstall for antatte investeringer som inngår i beregningene, men fra og med første overføring til KNR benyttes endringstall for utførte investeringer. Disse tidsseriene består av innrapporterte data for utførte investeringer i det aktuelle kvartalet, samt *imputerte* verdier for skjemaoppgaver som enda ikke er mottatt.

Hovedhensikten med å etablere KIS var at undersøkelsen skulle bidra med nødvendige data til KNR. Som en følge av dette er det god sammenheng mellom KNR sitt behov og hva som mottas av data.

3.2.3 Kvartalsvis omsetningsindeks for samferdsel og reiseliv

Formålet med omsetningsindeksen for samferdsels- og reiselivsnæringene (SROI) er å måle endring i omsetning for bedrifter i samferdsels- og reiselivsnæringene. I henhold til standard for næringsgruppering (SN2002), omfatter statistikken alle bedrifter i næringshovedområde H, *Hotell- og restaurantvirksomhet*, næringshovedområde I, *Transport og kommunikasjon*, samt næring 92.2 *Radio og fjernsyn*. Indeksen beregnes hovedsakelig på bakgrunn av data fra Skattedirektoratets oppgaveregister.

SROI i KNR

I KNR brukes SROI som indikator for transport og kommunikasjonsnæringene. SROI publiseres 60 dager etter kvartalets utløp og kan derfor ikke brukes som indikator for det gjeldende kvartal. Indeksen benyttes

⁵ Mer om standard for næringsgruppering, se <http://www.ssb.no/nace>. I dette notatet refereres det alltid til versjon SN2002.

i stedet til å oppdatere tidligere kvartaler. Indekstall legges på regneark på Q-området. SROI er en verdiindeks og brukes i sammenheng med volum- og prisindikatorer for samferdsels- og reiselivsnæringene.

3.2.4 Sysselsatte og arbeidsledige på korttidsopphold i Norge

Formålet med statistikken er å beskrive situasjonen for sysselsetting blant personer som ikke er registrert bosatt i Norge, altså personer som bare er ment å være her for en kortere periode (kriteriet er at man forventes å oppholde seg mindre enn seks måneder i landet). Statistikken er årlig og gjelder per 4. kvartal. Statistikken er et supplement til de ordinære sysselsettingsstatistikkene (Arbeidskraftundersøkelsen og den ordinære registerbaserte sysselsettingsstatistikken) som kun dekker personer som er registrert bosatt i Norge.

Foreløpig omfatter ikke statistikken selvstendig næringsdrivende og sjøfolk som er bosatt utenfor EØS-land. Statistikken er basert på informasjon fra flere ulike registre. De viktigste er NAVs Arbeidstakerregister, Lønns- og trekkoppgaverregisteret og registre over personer/arbeidsforhold fra Sentralskattekontoret for utenlandssaker.

KORTSYSS i KNR

Det er mange sysselsettingsstatistikker som inngår i KNR. Sysselsetting (og lønn) blir beregnet i et eget delsystem, det såkalte Arbeidskraftregnskapet (AR). Sysselsettingstall beregnes i sammenheng med lønnskostnader og lønn i næringene. Tre kilder benyttes: Strukturstatistikker, Arbeidskraftundersøkelsen (AKU) og registerdata (SSBs registerbaserte sysselsettingsstatistikk, SST og PAI).

De vanlige sysselsettingsstatistikkene (AKU og den ordinære registerbaserte sysselsettingsstatistikken) dekker ikke hele KNR-populasjonen.⁶ Sysselsatte på korttidsopphold brukes for å supplere disse statistikkene, der målet er å komme nærmere populasjonen i KNR. Siden statistikken i dag kun gjelder per 4. kvartal, må KNR gjøre vurderinger/beregninger på hvor mange som er i Norge på korttidsopphold. 4. kvartals-tallene danner en basis for disse vurderingene. Videre benyttes antall personer med D-nummer fra NAVs arbeidstakerregister. Dette registeret oppdateres løpende og KNR mottar disse tallene hvert kvartal. I tillegg gjøres det en vurdering av sysselsettingsveksten og Utlendingsdirektoratets tall for gyldige arbeidstillatelser benyttes.

⁶ De vanlige sysselsettingsstatistikkene omfatter personer som er registrert bosatt i Det sentrale folkeregisteret. KNR dekker derimot all sysselsetting i innenlandske bedrifter, dvs. bedrifter som forventes å drive produksjonsvirksomhet i 12 måneder eller mer i Norge. Avgrensning i KNR skjer uavhengig om man er registrert bosatt eller ikke. KNR skal dermed ha med flere sysselsatte enn de vanlige sysselsettingsstatistikkene.

4 Kvalitet, feilkilder og usikkerhet

Kapittelet gir en kort innføring i sentrale begreper knyttet til kvalitet, feilkilder og usikkerhet. Dette er begreper som det vil bli henvist til senere i notatet. Lesere med god innsikt i dette kan gå videre til kapittel 5.

4.1 Hva er kvalitet?

Kvalitet kan defineres som de karakteristiske trekk ved et produkt eller en tjeneste som har noe å gjøre med produktet/tjenestenes evne til å tilfredsstille uttalte og implisitte behov.

EU har utarbeidet retningslinjer for europeisk statistikk (*Code of Practice*), som blant annet er ment for Eurostat og medlemslandenes statistikkbyråer. Retningslinjene består av ulike krav gruppert etter institusjonelle forhold (strukturkvalitet), krav til de statistiske prosessene (prosesskvalitet) og de statistiske produktene (produktkvalitet).

Strukturkvalitet henviser til de rammer som må ligge til grunn for at statistikkproduksjonen skal være gjennomførbar. Det inkluderer faglig uavhengighet, mandat til å samle inn data, tilstrekkelige ressurser, kvalitetsbevissthet, statistisk konfidensialitet og upartiskhet og objektivitet.

Ved organisering, innsamling, produksjon og formidling av statistikk skal internasjonale standarder, retningslinjer og god praksis benyttes. Viktige momenter når det gjelder prosesskvaliteten er gode metoder, egnede statistiske prosedyrer, rimelige krav til oppgavegiverne og kostnadseffektivitet.

Kravene til de statistiske produktene er de samme som SSBs krav til kvalitet. De viktigste faktorene er:

- Relevans
- Nøyaktighet
- Aktualitet og punktlighet
- Informasjon om tilgjengelighet og klarhet
- Sammenlignbarhet med andre statistikker
- Sammenheng mellom statistikker

Nedenfor er disse faktorene beskrevet nærmere.

4.1.1 Relevans

En undersøkelse er relevant hvis brukernes behov blir ivaretatt. Det er derfor nødvendig å identifisere hvem brukerne er og hvilke behov de har.

4.1.2 Nøyaktighet

Statistikk skal avspeile virkeligheten på en nøyaktig og pålitelig måte. For å kunne si noe om nøyaktigheten til statistikken, er analyse av mulige feilkilder viktig. Det er vanlig å skille mellom utvalgsfeil og ikke-utvalgsfeil. Dette er nærmere beskrevet i kapittel 4.2.

4.1.3 Aktualitet og punktlighet

Publisering av statistikken skal være aktuell og punktlig. Det innebærer blant annet faste publiseringstidspunkt (SSB publiserer for eksempel fast kl 10 hver dag), at periodeinndelingen tar hensyn til brukerne i størst mulig grad og at alle avvik fra fastsatt publiseringsdag i statistikkalenderen gjøres kjent og forklares på forhånd.

4.1.4 Tilgjengelighet og klarhet

Statistikken skal publiseres klart, lettfattelig og på en egnet og praktisk måte. Dataene må være enkle å tolke og sammenligningene skal være meningsfulle. Mikrodata skal også tilrettelegges for forskere og metadata dokumenteres.

4.1.5 Sammenlignbarhet med andre statistikker

Statistikken skal kunne sammenlignes over tid og mellom regioner og land. Det er derfor viktig at endringer og virkningen av disse dokumenteres. Det er også viktig at statistikk fra forskjellige kilder sammenlignes og samordnes.

4.1.6 Sammenheng mellom statistikker

Sammenheng mellom ulike statistikker er rettet mot sammenligning av ulike statistikker, der disse statistikkene ofte blir produsert på ulik måte og til forskjellig bruk. For å få sammenheng mellom ulike statistikker må disse være basert på felles definisjoner, klassifikasjoner og metodiske standarder.

4.2 Feilkilder og usikkerhet

Som nevnt i kapittel 4.1.2 skal statistikken avspeiler virkeligheten nøyaktig og pålitelig. Imidlertid er det knyttet en viss usikkerhet til enhver statistikk.

Det er vanlig å klassifisere feil i utvalgsfeil og ikke-utvalgsfeil. Utvalgsfeil er feil som skyldes at kun en del av, og ikke hele, målpopulasjonen brukes i undersøkelsen. Denne typen feil er omtalt i delkapittel

4.2.1. Ikke-utvalgsfeil er feil som er knyttet til selve datainnsamlingen, og kan oppstå for både utvalgsundersøkelser og registerbasert statistikk (fulltelling). Ikke-utvalgsfeil kan videre deles inn i:

- Registerfeil (dekningsfeil)
- Målefeil
- Bearbeidingsfeil
- Frafallsfeil
- Modellfeil

Disse feilene beskrives i delkapitlene 4.2.2-4.2.6.

4.2.1 Utvalgsfeil

Denne typen feil er forårsaket av at det ved innhenting av datagrunnlag for en statistikk benyttes et utvalg i stedet for en totaltelling, og er mer presist en stokastisk usikkerhet enn en egentlig feil.

Det skilles mellom usikkerhet i trekkingen av utvalget (designerusikkerhet) og usikkerhet fra en statistisk modell (usikkerhet i de estimerte verdiene). Det er også vanlig å skille mellom feil som oppstår i sannsynlighetsutvalg (der det trekkes et utvalg) eller ikke-sannsynlighetsutvalg (der det utpekes et utvalg).

Vanlige mål på utvalgsfeil er utvalgsprosent, dekningsgrad, varians og standardavvik. Utvalgsprosenten angir hvor stor del av populasjonen som er med i utvalget, mens dekningsgraden forteller hvor stor andel av målpopulasjonen som inngår i utvalget. Variansen for en variabel viser spredningen omkring gjennomsnittet for den empiriske verdien. Standardavviket er kvadratroten av variansen.

Etablering av usikkerhetsmål for en utvalgsundersøkelse gir mulighet for en fortløpende vurdering av om kvaliteten på estimerte verdier er god nok til at disse kan videreformidles. Samtidig kan man se om usikkerheten tiltar eller avtar over tid, og bruke denne informasjonen til iverksette nødvendige tiltak.

Ulempen med standardavvik og varians er at de er relativt ressurskrevende å lage når ikke-sannsynlighetsutvalg benyttes.

4.2.2 Registerfeil (dekningsfeil)

Med registerfeil eller dekningsfeil menes feil i registre som brukes som populasjonsregistre. Dette kan være feil som skyldes feilklassifisering, overdekning eller underdekning.

Feilklassifisering betyr at enhetene som tilhører målpopulasjonen ikke blir klassifisert eller blir feilaktig plassert i en klasse, for eksempel med hensyn til næring eller størrelse.

Overdekning skjer når enheter som ikke hører til i populasjonen eller enheter som ikke finnes, er feilaktig medregnet i populasjonen. Manglende opphørsmeldinger kan være en årsak til denne typen feil.

Underdekning forekommer når (nye) enheter ikke er tatt med i registeret, enten på grunn av nyetablering eller fusjon, manglende innmelding av for eksempel personer/bedrifter, forsinkelser i meldingsgangen til registeret eller fordi enheter er klassifisert feil.

Konsekvensen av underdekning kan være systematiske feil i estimatene. Når det gjelder overdekning, vil dette som regel føre til skjevheter i estimatene, det vil si avvik mellom den forventede verdien og den sanne verdien.⁷ Det er vanskelig å kvantifisere denne type feil.

4.2.3 Målefeil

Feil som oppstår i datainnsamlingen kalles målefeil. Kilder til målefeil kan være at respondenten ikke har gitt riktig svar (respondentfeil), intervjueren ikke har registrert svaret korrekt (intervjufeil) eller uklare spørsmålsformuleringer i spørreskjemaet. Målefeil kan også oppstå dersom referanseperioden for statistikken ikke er den samme som den perioden dataene er samlet inn for.

Ulike teknikker kan brukes for å vurdere målefeil, som for eksempel gjentatt intervju og estimering av intervju-effektene, men kostnadene vil ofte være høye. Imidlertid vil den økte bruken av elektronisk innrapportering gjøre det noe enklere å gjennomføre ulike kontroller.

4.2.4 Frafallsfeil

Det finnes to typer frafall: enhetsfracfall og partielt frafall. Enhetsfracfall (eller totalfracfall) inntreffer når respondenten unnlater å svare, mens partielt frafall (eller svarfracfall) inntreffer når respondenten unnlater å svare på et eller flere spørsmål.

Enhetsfracfall er ofte en viktig kilde til feil i undersøkelser og kan føre til skjevheter i estimatene. En enkel indikator for frafall er svarprosenten. Lav svarprosent er en indikasjon på at det knyttet usikkerhet til tallene. Vektet svarprosent er ofte en enda bedre indikator da dette gir et mer korrekt bilde av omfanget ved at det tar hensyn til om kritiske enheter mangler.

⁷ *Forventet verdi* er det hypotetiske gjennomsnittet over alle tenkte utvalg fra en undersøkelse. *Sann verdi* er den faktiske verdien vi får hvis vi ikke gjør noen feil, verken ved innsamling av data eller ved beregninger.

4.2.5 Bearbeidingsfeil

Med bearbeidingsfeil menes feil som oppstår i datainnsamlingsprosessen ved koding, redigering og fremstilling av tabeller. Vanlige bearbeidingsfeil er feil under inntasting (punchefeil), feiltolkning ved optisk lesning og feilkoding (for eksempel korrekte tall vurderes som uriktige og feilaktig blir rettet). Det er imidlertid vanskelig å tallfeste denne typen feil.

4.2.6 Modellfeil

Modellfeil kan oppstå hvis en feilaktig eller mindre god modell velges, for eksempel ved sesongjustering av tidsserier. Slike feil vil kunne skyldes avvik mellom de betingelsene som ligger til grunn for metoden som benyttes for å sesongjustere. Bevegelige helligdager vil kunne føre til problemer for modellen som benyttes, men disse problemene vil ofte være større for månedsstatistikker enn kvartalsstatistikker.

5 Metadata som grunnlag for dialog og samarbeid

Kapittel 5 inneholder en kartlegging av *metadata*, det vil si data om data, som bidrar med relevant informasjon ved dataleveranser fra fagseksjonene til Seksjon for nasjonalregnskap. Dette er viktig da tilgang på slik informasjon er viktig for konstruktiv dialog og samarbeid.

5.1 Utgangspunkt for arbeidet

I Evensen (2004) side 99, er det utarbeidet en tiltaksliste for videre arbeid knyttet til kvalitet i KNR. Et av punktene på listen lyder som følger: "*Seksjon for nasjonalregnskap må i samarbeid med fagseksjonene spesifisere krav til metadata for den enkelte statistikk. Metadata bør overleveres fra fagseksjonen sammen med oppdatert korttidsstatistikk*".

Tilgang på relevant informasjon er en viktig forutsetning for konstruktiv dialog og samarbeid. Dette gjelder også i forholdet mellom fagseksjonene, som leverandører av data, og Seksjon for nasjonalregnskap, som mottaker av data. Med utgangspunkt i tiltakslisten fra Evensen (2004) valgte prosjektgruppa å kartlegge relevante metadata ved leveranser av korttidsstatistikk fra fagseksjonene til Seksjon for nasjonalregnskap. Hvilke korttidsstatistikker som inngår i dette arbeidet samt bakgrunnen for at nettopp disse ble valgt er beskrevet i kapittel 3. Nøkkelpersoner ved fagseksjonene og Seksjon for nasjonalregnskap ble konsultert for å kvalitetssikre resultatet av kartleggingen.

5.2 Kartlegging av relevante metadata

Etter noe diskusjon konkluderte prosjektgruppa med at følgende typer metadata er relevante ved interne leveranser av korttidsstatistikk fra fagseksjonene til Seksjon for nasjonalregnskap:

1. Opplysninger om statistikken
2. Opplysninger om dataleveransen
3. Opplysninger om mottager av dataleveransen
4. Opplysninger om feilkilder og usikkerhet
5. Opplysninger om endringer som kan påvirke beregningene i KNR

5.2.1 Opplysninger om statistikken

Denne typen informasjon vil vanligvis være uendret fra periode til periode. Derfor kan det oppstå problemer/forvirring ved endringer. Prosjektgruppa kom frem til at det kan være nyttig å ha tilgang til følgende opplysninger:

- Navn på statistikken
- Navn på fagseksjonen som er ansvarlig for å lage statistikken
- Navn på statistikkansvarlig
- Telefonnummer til statistikkansvarlig
- E-postadresse til statistikkansvarlig
- Næringer som inngår i statistikken (NACE-kode, eller eventuell annen type klassifisering)
- Om næringer oppdateres fortløpende, eller for eksempel bare en gang hvert år
- Variabelnavn (for eksempel *utførte investeringer*)
- Variabeltype (verditall eller volumtall)
- Datainnsamlingsmetode (utvalgsundersøkelse eller registerundersøkelse)
- Hyppighet (ukentlig, månedlig, terminvis, kvartalsvis)

Opplysninger om navn på ansvarlig seksjon og navn på statistikkansvarlig er viktig da dette sier noe om hvor man skal henvende seg ved behov for informasjon. Telefonnummer og e-postadresse er tilgjengelig på intranettet, men det er praktisk om slik informasjonen også står oppført sammen med navnet på personen. Opplysninger om hvilke næringer som inngår i statistikken forteller hvorvidt statistikken er dekkende (relevans). Informasjon om variabelnavn og variabeltype benyttes ved innarbeiding av datamaterialet i KNR-systemet. Når det gjelder informasjon om datainnsamlingsmetode, sier dette noe om resultatene er beheftet med usikkerhet på grunn av utvalgsfeil, mens hyppighet oppgir frekvens for frigivning av statistikken.

5.2.2 Opplysninger om dataleveransen

Denne typen informasjon vil endre seg ved overgang til ny periode og/eller ved nye dataleveranser fra fagseksjonen. Prosjektgruppa kom frem til at det kan være nyttig å ha tilgang til følgende opplysninger:

- Statistikkperiode (for eksempel 1. kvartal 2008)
- Dato for dataleveransen
- Klokkeslett for dataleveransen
- Fritekstfelt

Variabelen statistikkperiode forteller om dataleveransen oppfyller gjeldende krav til aktualitet, mens dato og klokkeslett forteller når tallene er overført. Det siste er nødvendig ved absolutte tidsfrister. I fritekstfeltet kan man legge informasjon om spesielle forhold som berører akkurat denne overføringen.

5.2.3 Opplysninger om mottager av dataleveransen

Slik informasjonen vil vanligvis være uendret fra periode til periode. Prosjektgruppa kom frem til at det kan være nyttig for avsender å ha tilgang til følgende opplysninger:

- Navn på mottager av data (næringsberegner i KNR)
- Telefonnummer til mottager av data
- E-postadresse til mottager av data

Opplysninger om navn og e-postadresse er viktig da dette forteller hvem fagseksjonene skal sende informasjonen til. Telefonnummer er nyttig hvis det er noe som haster.

5.2.4 Opplysninger om feilkilder og usikkerhet

Opplysninger om feilkilder og usikkerhet er viktig da dette sier noe om i hvilken grad man bør legge vekt på resultatene fra statistikken. I tillegg gir det en mulighet til å vurdere ulike informasjonskilder opp mot hverandre. Mulige indikatorer på feilkilder og usikkerhet i korttidsstatistikken er identifisert ved å ta utgangspunkt i avsnitt 4.2 og knytte dette opp mot undersøkelsene som er beskrevet i kapittel 3.

Prosjektgruppa har lagt vekt på at indikatorene bør være enkle å forstå samt at det ikke må være for ressurskrevende for statistikkansvarlig å fremskaffe den nødvendige informasjonen.

5.2.4.1 Utvalgsbasert statistikk

For utvalgsbasert statistikk er det vanlig å skille mellom utvalgsfeil og ikke-utvalgsfeil (se avsnitt 4.2).

Indikatorer på utvalgsfeil:

- **Utvalgets dekningsgrad**

Utvalgets dekningsgrad oppgir hvor stor andel av målpopulasjonen som inngår i datagrunnlaget. Dersom statistikken har lav dekningsgrad, er dette en indikasjon på at det er knyttet usikkerhet til estimatene.

- **Varians/standardavvik**

Variansen viser spredningen omkring gjennomsnittet for den empiriske verdien. Jo høyere varians, jo større usikkerhet. Standardavviket beregnes som kvadratroten av variansen og gir en indikasjon på i hvilken grad man risikerer at det beregnede estimatet avviker fra den verdien man ønsker å estimere. Varians og standardavvik er relativt krevende å beregne, men mindre man har tilgang til et statistisk verktøy som er i stand til å utføre jobben.

Indikatorer på ikke-utvalgsfeil:

- **Frafallsfeil**

Svarprosenten gir en indikasjon på usikkerhet i statistikken da manglende data må erstattes gjennom logisk redigering eller ved bruk av statistiske metoder. I statistikker hvor enkeltobservasjoner har stor innvirkning på det estimerte resultatet kan det være bedre å bruke vektet svarprosent. Dette innebærer at man tar høyde for at noen enheter er viktigere enn andre.

- **Prediksjonsfeil**

Da KNR er underlagt strenge krav til aktualitet, forekommer det at foreløpige tall, fremskrevne tall eller prognoser inngår i datagrunnlaget fra fagseksjonene. En visualisering av historisk avvik mellom disse og endelige tall, gir en indikasjon på usikkerhet i datamaterialet. En forutsetning for denne typen kontroll er at det etableres rutiner for lagring historiske data. Slik informasjon kan også benyttes til å bedre nøyaktigheten på fremskrivingene.

I avsnitt 4.2 omtales også andre kilder til usikkerhet i utvalgsstatistikk, men prosjektgruppa har ikke klart identifisere gode indikatorer for disse da de enten er kompliserte å beregne eller anses som mindre relevante i forhold til KNR. Registerfeil som kilde til usikkerhet gjennomgås i kommende avsnitt.

5.2.4.2 Registerbasert statistikk

Mange av SSB sine korttidsstatistikker er registerbaserte. Dette innebærer at datainnsamlingen skjer direkte fra sentrale registre. Et eksempel på et slikt register er Skattedirektoratets oppgaveregister. Registerbasert statistikk baserer seg på fulltelling og vil dermed ikke være beheftet med utvalgsfeil. I stedet kan registerfeil være en kilde til usikkerhet.

Prosjektgruppa er av den oppfatning at det har vært mye vanskeligere å finne gode indikatorer på usikkerhet i registerundersøkelser enn gode indikatorer på usikkerhet i utvalgsundersøkelser. Blant annet skyldes dette at registergrunnlaget kan være forskjellig for ulike statistikker. Derfor er det vanskelig å være konkret med henblikk på indikatorer. Nedfor følger noen generelle forslag til indikatorer på usikkerhet i registerbasert statistikk.

Registerfeil (dekningsfeil):

- Har det vært endringer i meldingsgangen til ett eller flere registre som benyttes/ligger til grunn for statistikken?
 - Hva slags register?
 - Hva slags endringer?

- Hvor mange personer/husholdninger/bedrifter/foretak er berørt?

Dette vil gi en indikasjon på kvaliteten i registeret. Jo flere personer/bedrifter etc. som er berørt, jo større grunn er det til å stille spørsmålstegn ved kvaliteten.

- Har det vært administrative endringer i ett eller flere av registrene som benyttes/ligger til grunn for statistikken?
 - Hva slags endringer?

For eksempel endringer i IT-systemet, nytt IT-system, endringer i administrative rutiner, osv. hos registreier (som ofte ikke er SSB).

- Indikatorer på variabelnivå:
 - Andel uoppgitt

Eventuelle store endringer i uoppgittandelen vil kunne si noe om kvaliteten på variabelen.

- Avvik mellom foreløpige og endelige tall (prediksjonsfeil)

På samme måte som ved utvalgsbasert statistikk, kan man se på historisk avvik mellom foreløpige og endelige tall.

I tillegg til registerindikatorerne nevnt over, er *imputeringsmengde* en viktig indikator for kvalitet. Blanke felter i registeret kan tolkes som frafall. Imputeringsprosessen kan tolkes som en slags vektning eller oppblåsning på samme måte som ved utvalgsundersøkelser. Ut fra dette skal det være mulig å regne ut standardavvik og andre usikkerhetsmål.

5.2.5 Opplysninger om endringer som kan påvirke beregningene i KNR

Det er viktig at Seksjon for nasjonalregnskap mottar informasjon om endringer som kan påvirke beregningen av KNR. Eksempler på slike opplysninger kan være:

- Revisjoner bakover i tid
- Næringsendringer for viktige observasjoner
- Nye definisjon av variabler
- Endret hyppighet
- Lovendringer eller nye reguleringer

6 Forslag til system for utveksling av metadata mellom fagseksjonene og Seksjon for nasjonalregnskap

6.1 Behov for et informasjonssystem

Samhandlingen mellom fagseksjonene og KNR foregår primært ved dialog mellom fagansvarlig og næringsberegner. Det er disse som besitter detaljkunnskap om de spesifikke indikatorene. Dagens rutiner fungerer bra så lenge de involverte kjenner hverandre og har god kjennskap til sine respektive fagområder. Det kan imidlertid oppstå problemer når nye personer skal sette seg inn i disse rutinene. Dette spesielt med tanke på at verken e-post eller telefonsamtaler blir lagret systematisk. For nye medarbeidere, eller i forbindelse med sykdom og ferieavvikling, kan derfor denne informasjonen oppfattes som utilgjengelig.

Det er prosjektgruppens mening at en mer strukturert form for informasjonsutveksling kan gi en bedre dialog mellom fagseksjonene og KNR. Med informasjonsutveksling tenker vi først og fremst på metadata, slik det er beskrevet i kapittel 5. Når det gjelder hvilket format de enkelte tidsserier sendes til KNR på, det være seg FAME, Excel eller på papir, så er ikke det et tema vi går inn på i dette notatet. Vi holder oss til å si noe om utvekslingen av metadata, det vil si *data om data*.

Per i dag er metadata tilgjengelig fra ulike kilder. Først og fremst er *Om statistikken* et nyttig oppslagsverktøy. Likevel er det heller ikke her nevnt eksplisitt hvordan de forskjellige korttidsstatistikkene inngår i KNR. I kapittel 3 har vi kommet med forslag til et slik punkt, som bør innarbeides i malen til *Om statistikken*. I tillegg til *Om statistikken* har det de siste årene blitt utarbeidet flere andre metadatasystemer. STABAS er ett eksempel, der man på en enkel måte kan finne fram til de ulike typer klassifiseringer som brukes i SSB. I tillegg har man utviklet VARDOK, der man på en enkel måte skal kunne finne definisjoner av de ulike variablene som inngår i statistikkene. Mer om disse informasjonssystemene finner man på <http://www.ssb.no/metadata>.

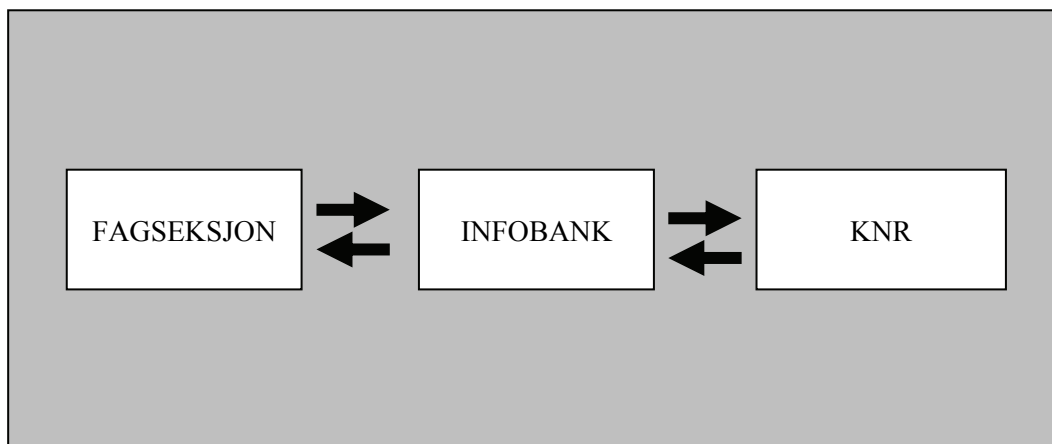
Alle de eksisterende metadatasystemene er nyttige informasjonskilder for KNR. Likevel inneholder disse informasjon som stort sett er konstant over tid. Derfor ønsker vi oss et system som er spesialdesignet for informasjonsutvekslingen mellom fagseksjonene og KNR, der vi samtidig kan få informasjon som endrer seg fra kvartal til kvartal. Videre i kapittelet skal vi presentere prosjektgruppas forslag til hvordan et slik system kan se ut. Viktige stikkord er relevans, samt at systemet skal være enkelt å bruke.

6.2 Skisse til system for elektronisk lagring og formidling av metadata

6.2.1 Overordnet systembeskrivelse

Prosjektgruppen foreslår en informasjonsbank med toveis kommunikasjon mellom fagseksjonene og KNR. Denne skal fungere som et viktig supplement til dagens arbeidsmøter, samt kontakt via telefon og e-post. Systemet skal holde oversikt over alle statistikker som inngår i KNR. Det skal være mulig å registrere overføringer, få oversikt over administrative opplysninger, oversende kvalitetsindikatorer, samt få tilgang til dokumentasjon og historiske data.

Figur 6-1 System for å registrere overføring av metadata



6.2.2 Kriterier for valg av løsning

Prosjektgruppen kom fram til følgende kriterier for valg av løsning:

1. Systemet må ikke være for ressurskrevende å lage
2. Systemet må være brukervennlig
3. Systemet må være lett å administrere

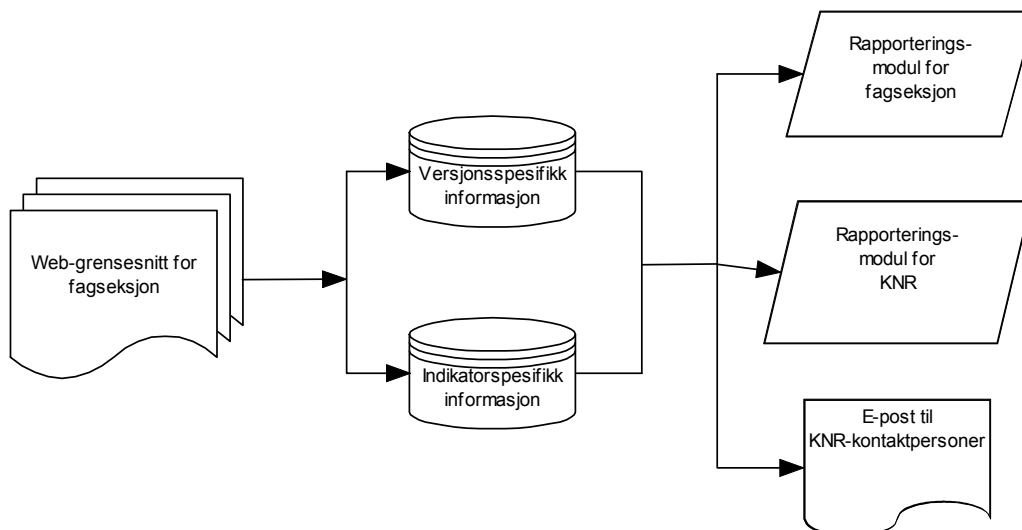
Da ingen av prosjektdeltagerne har spisskompetanse innenfor IT, valgte man å søke veiledning hos ressurspersoner internt i organisasjonen. Denne informasjonen ble så brukt til å lage et løsningsforslag i tråd med gjeldende kriterier.

6.2.3 Prosjektgruppas forslag

Prosjektgruppas skisse til løsning er basert på en eksisterende applikasjon for elektronisk registrering av nye frigivningsdatoer i Statistikkalenderen (Statistikkalender-applikasjonen), og vår løsning vil videre bli forklart ut i fra denne. Kort fortalt, består systemet av et nettbasert brukergrensesnitt som er en overbygging av en orakeldatabase med to tabeller. Tanken er at man skal kunne logge seg på systemet via

Byrånettet og legge inn metadata for de statistikkene man er ansvarlig for. Informasjon formidles via e-post eller hentes ut i form av rapporter. Dette er illustrert i figur 6-2, som forklares under.

Figur 6-2 Flytdiagram for informasjonsbanken



Pålogging og registrering:

Ved pålogging i Statistikkalender-applikasjonen kommer man til et administrasjonsbilde hvor tilgjengelige statistikker er seksjonsbestemt, slik vi ser av figur 6-3. Her velger man hvilken statistikk det skal registreres opplysninger for. Applikasjonen gir mulighet for å registrere ny informasjon samt oppdatere eksisterende informasjon.

Figur 6-3 Registreringsbilde

Statistikkalenderen

Ny frigivning

Logget inn som: **Stein Bakke** Du jobber i **Seksjon for industristatistikk (470)** så du kan registrere frigivning på følgende statistikker:

- Bergverk, industri og kraftforsyning - investeringsstatistikk.
- Forskning og utvikling, næringslivet
- Industri, lagerstatistikk
- Industri, ordrestatistikk
- Industri, produksjonsindeks
- Industri, strukturtall
- Industri, utenlandsk eierskap
- Industri, varetall
- Innovasjoner i norsk næringsliv
- Konjunkturbarometeret
- Olje- og gassutvinning, industri, bergverksdrift og kraftforsyning - omsetningsstatistikk

Ny frigivning Vis registrerte

Vis/endre kontaktpersoner

Vis/endre produktregister info

I prosjektgruppas forslag til system, der applikasjonen er tilpasset for dataoverføringer mellom fagseksjonene og KNR, skal det i stedet stå:

- Logget inn som: <navn på person>
- Du jobber i seksjon for <navn på seksjon> og har tilgang til følgende statistikker: <oppgis i tekstboksen>

Statistikkansvarlig markerer hvilken statistikk han/hun ønsker å legge inn metadata for og klikker på knappen <Registrer opplysninger>. Dette tilsvarer knappen *Ny frigivning* i den eksisterende applikasjonen.

Oracle-tabeller:

Prosjektgruppas løsningsforslag krever at det etableres to orakeltabeller. Den ene vil være reservert for versjonsspesifikk informasjon, mens den andre vil være reservert for statistikkspesifikk informasjon. Dette forklarer vi nærmere under. I Statistikkalender-applikasjonen blir slike tabeller tilgjengelige ved å trykke på knappen *Ny frigivning*. Da kommer man til følgende side:

Figur 6-4 Opplysninger fra Oracle-tabeller

Statistikkalenderen	
Ny frigivning	
Om statistikken:	
Kortnavn:	kis
Statistikk navn:	Bergverk, industri og kraftforsyning - investeringsstatistikk.
Emnekode:	08.03.30
Kontaktpersoner:	jan.henrik.wang@ssb.no 21 09 47 20 Stein.Bakke@ssb.no 21 09 47 45
Om frigivningen:	
Norsk tittel:	Investeringsstatistikk. Industri, bergverksdrift og kraftfo
Engelsk tittel:	Quarterly Investment Statistics. Manufacturing, mining &
Målform:	<input checked="" type="radio"/> Bokmål <input type="radio"/> Nynorsk
Frigivning statistikkbanken	<input type="radio"/> Alt <input type="radio"/> Delvis <input checked="" type="radio"/> Ikke (denne publiseringen)
Tidspunkt:	Uke nr: <input type="text"/> (aaaa-uu)  Dato: <input type="text"/> (aaaa-mm-dd) kalender
Neste >	

Versjonsspesifikk informasjon:

Med versjonsspesifikk informasjon, menes opplysninger knyttet til en bestemt dataleveranse. I prosjektgruppas skisse til system for elektronisk lagring og formidling av metadata, vil versjonsspesifikk informasjon være registrerte opplysninger om hver enkelt dataleveranse fra fagseksjonene til Seksjon for nasjonalregnskap. I kapittel 5, der vi har redegjort for de mest relevante metadata, er avsnittene 5.2.2, 5.2.4 og 5.2.5 eksempler på versjonsspesifikk informasjon. Det sistnevnte avsnittet bør legges i et fritekstfelt.

Statistikkspesifikk informasjon:

Med statistikkspesifikk informasjon menes metadata knyttet til hver enkelt statistikk. Slike opplysninger oppdateres fortløpende ved endringer. I Statistikkalender-applikasjonen foretas endringer ved å registre nye opplysninger i bildet som kommer til syne, ved å trykke på knappene *Vis/endre kontaktpersoner* eller *Vis/endre produktregister info* (se figur 6-3).

Figur 6-5 Oppdateringer



Epost	Telefon	
jan.henrik.wang@ssb.no	21 09 47 20	fjern
Stein.Bakke@ssb.no	21 09 47 45	fjern

- [Legg til ny kontaktperson](#)

I prosjektgruppas skisse til system for elektronisk lagring og formidling av metadata, vil statistikkspesifikk informasjon tilsvare standardiserte opplysninger om de ulike korttidsstatistikkene som inngår i KNR. Se avsnitt 5.2.1 og 5.2.3 for eksempler på slik informasjon.

6.2.4 Formidling av opplysningene

Når det gjelder formidling av opplysninger fra databasen, ser prosjektgruppa for seg en løsning der dette skjer via to forskjellige kanaler:

1. Formidling av opplysninger via e-post
2. Direkte uttrekk fra databasen

6.2.4.1 Formidling av opplysninger via e-post

Alternativ 1 innebærer at applikasjonen implementerer funksjonalitet som automatisk genererer en e-post når statistikksansvarlig trykker på en knapp med tittelen *Send*. Denne bør ligge i administrasjonsbildet. E-posten vil inneholde både statistikkspesifikk- og versjonsspesifikk informasjon, men mengden opplysninger bør begrenses. Mottagere av e-posten vil være kontaktpersoner ved Seksjon for

nasjonalregnskap for den aktuelle statistikken. E-post adressene deres hentes fra orakeltabellen med statistikkspesifikk informasjon.

Om man tar utgangspunkt i kapittel 5, vil et mulig forslag til innhold i en slik e-post være:

Statistikkansvarlig:	Gjeldende periode:	Variabel:
Navn	Kvartal	Navn
Telefonnummer	Måned	Type (verditall, volumtall)
E-post	Etc.	
Usikkerhet:	Revisjoner bakover i tid:	Kommentarer:
Dekningsgrad	Ja/nei	Fritekstfelt
Svarprosent	Revidert periode	
Etc.		

6.2.4.2 Direkte uttrekk fra databasen

Alternativ 2 innebærer at det etableres to rapporteringsmoduler som gir mulighet for direkte uttrekk av opplysninger fra tabellene i orakeldatabasen. Når det gjelder Statistikkalender-applikasjonen, så har ikke denne en slik rapporteringsmodul, men har noe som likner:

Figur 6-6 Uttrekk fra databasen

kis

Klikk på tittelen til frigjøringen for å se detaljer

Følgende frigivninger er sendt til DS redaksjonen, og avventer godkjenning for å bli lagt i den offentlige kalenderen. Du kan slette frigivninger herfra ved å klikke *fjern*.

Ingen frigivninger

Kommende frigivninger meldt i offentlig kalender, sortert stigende etter frigivningsdato. For å endre frigivninger som allerede er meldt i den offentlige kalenderen kontakt [Oddveig Dystland](#)

Frigis	Tittel
02.06	Investeringsstatistikk. Industri, bergverksdrift og kraftforsyning, 2. kv. 2008

Detaljer for frigivning

Frigivningsdato: 02.06.2008 10:00	SB: alt	MF: Bokmål
Tittel: Investeringsstatistikk. Industri, bergverksdrift og kraftforsyning, 2. kv. 2008		
Engelsk tittel: Quarterly Investment Statistics. Manufacturing, mining and quarrying and electricity supply, 2nd quarter 2008		

Dette bildet er laget slik at statistikkansvarlig kan slå opp registrerte opplysninger for en bestemt frigivning ved å klikke på lenken med den aktuelle frigivningsdatoen. Disse kommer til syne i feltene under *Detaljer for frigivning*.

Prosjektgruppas skisse til system for elektronisk lagring og formidling av metadata, krever at det etableres to rapporteringsmoduler. Den ene vil være tilgjengelig for statistikkansvarlige, mens den andre vil være tilgjengelig for kontaktpersoner på Seksjon for nasjonalregnskap. Hensikten er å gi mulighet for å lage skreddersydde rapporter. Ulike typer skreddersydde rapporter er:

- Metadata for siste beregning av en bestemt statistikk i en bestemt periode.
- Metadata for flere beregninger av en bestemt statistikk i en bestemt periode.
- Metadata for siste beregning av en bestemt statistikk over flere perioder.
- Metadata for flere beregninger av en bestemt statistikk over flere perioder.
- Metadata for siste beregning av flere statistikker i en bestemt periode.
- Metadata for flere beregninger av flere statistikker i en bestemt periode.
- Metadata for siste beregning av flere statistikker over flere perioder.
- Metadata for flere beregninger av flere statistikker over flere perioder.

Med dette systemet skulle det være mulig å selektere på tvers av statistikker, perioder og versjoner ved hjelp av informasjon fra tabellene. Et mulig eksempel skisseres i tabell 6-1.

Tabell 6-1: Eksempel på uttrekk av versjonsspesifikk informasjon

Statistikk	Periode	Versjon	Metadata 1	Metadata 2
KIS	200801	1	X	X
KIS	200801	2	X	X
KIS	200802	1	X	X
PI	200801	1	X	X
PI	200801	2	X	X
PI	200802	1	X	X

Denne typen rapporter gir en hel rekke muligheter. Noen av de mest opplagte er:

- Det vil bli mulig å vurdere kvalitetsmessige egenskaper for en bestemt statistikk i en bestemt periode mot resultatene fra tidligere perioder. Dermed blir det lettere å identifisere potensielle avvik fra "normalen".
- Det vil bli mulig å vurdere kvalitetsmessige egenskaper for en bestemt statistikk mot kvalitetsmessige egenskaper for tilsvarende undersøkelser. Dermed blir det lettere å se i hvilken grad man skiller seg ut.
- I tilfeller hvor statistikken beregnes på nytt etter hvert som det kommer inn mer data, vil man enkelt kunne synliggjøre i hvilken grad ny informasjon bidrar til å redusere usikkerheten. Slik informasjon vil blant annet fortelle hvor hensiktsmessig det er å foreta nye beregninger. Dermed blir det lettere å vite når man kan sette sluttstrek.

6.2.5 Gevinster ved å etablere en informasjonsbank

Prosjektgruppa ser for seg følgende gevinster ved å ta i bruk en informasjonsbank, slik den er skissert over:

- det er enkelt å ta ut rapporter.
- alle får lett tilgang på dokumentasjon siden dokumentasjonen blir en del av produksjonsopplegget
- det blir enkelt å sammenligne data (med historiske tall og andre datakilder).
- KNR-prosessen blir mindre personavhengig.
- mens andre metadatasystemer i SSB inneholder metadata som er konstante over tid, vil man med en informasjonsbank få ut metadata som endrer seg fra kvartal til kvartal.

I avsnitt 6.2.2 satte vi opp tre kriterier for vårt valg av løsning, og etter vår mening oppfyller forslaget alle kriteriene. Gjenbruk av elementer fra allerede eksisterende systemer i SSB vil bidra til å dempe ressursbehovet. Ut fra egen erfaring vet vi at en slik løsning er svært brukervennlig. Systemet er også enkelt å administrere da endringer i statistikkspesifikk informasjon kan oppdateres direkte i databasen. Tilfeller der man vil være avhengig av IT-ressurser vil stort sett kun være i forbindelse med organisasjonsmessige endringer eller hvis statistikker skifter navn o.l.

Prosjektgruppa mener at applikasjon for elektronisk registrering av nye frigivningsdatoer i Statistikkalenderen og applikasjon for elektronisk lagring og formidling av metadata ligger så nær hverandre, at det kanskje kan være hensiktsmessig å utvide den eksisterende løsningen med ny funksjonalitet. På den måten begrenses kostnadene ytterligere. Funksjonalitet for lagring og formidling av metadata bør kun være synlig for fagseksjoner som har leveranser til Seksjon for nasjonalregnskap.

7 anbefalinger for videre arbeid

I dette notatet har vi kommet med forslag til hvordan dialogen mellom fagseksjonene og seksjon for Nasjonalregnskap kan forbedres. Vår oppfatning er at god dialog avhenger av kunnskap. Vi har vist hvordan en informasjonsbank gjør denne kunnskapen lettere tilgjengelig. I dette avsnittet gir vi forslag til tiltak som kan iverksettes i kjølvannet av dette prosjektet. Noen anbefalinger er lite ressurskrevende å iverksette. Andre vil kreve mer av både fagseksjonene og S930.

7.1 anbefalinger som krever lite ressurser

- Det opprettes en arbeidsgruppe som utarbeider en endelig spesifisering av krav til metadata.
- Informasjonsbanken utvikles i samarbeid mellom S930, fagseksjoner og seksjon for IT-utvikling.
- Det jobbes videre med å pleie den personlige kontakten mellom seksjonene. I desember 2007 gjennomførte seksjon for Nasjonalregnskap en KNR-presentasjon i Kongsvinger. Temaer var blant annet teoretisk forankring og praktiske eksempler. Initiativet ble meget godt mottatt i Kongsvinger og vi oppmuntrer S930 til å fortsette med dette. Vi tror ønsket om personlig kontakt er gjennomgående for de involverte i KNR-arbeidet. Vi anbefaler derfor at det årlig avholdes en konferanse hvor de involverte får en anledning til å møtes. Dette tror vi kan bidra til økt fellesskapsfølelse.
- For de statistikker som leverer tall til KNR burde det legges en link i *Om statistikken* hetende **Bruk i KNR**. I henhold til malen for *Om statistikken* vil det være naturlig å legge en slik link under avsnitt 6.2 *Sammenheng med annen statistikk*. Tanken er at spesielt interesserte brukere enkelt skal kunne finne mer informasjon om hvordan statistikkene inngår i KNR.

7.2 anbefalinger som er mer ressurskrevende

Fagseksjonene har for lite innblikk i hvordan tallene fra KNR beregnes. Fagansvarlige kan oppleve å få spørsmål som er vanskelig å svare på. Vi anbefaler derfor følgende:

- Fagseksjonene utvikler egne indikatorer som brukes til sammenligning med KNR-tallene.
- Byråskolen tilbyr et kurs rettet mot praktisk arbeid med KNR

På lengre sikt anbefaler vi i tillegg at følgende tiltak vurderes:

- En fullstendig kartlegging av alle indikatorer som brukes i KNR. Vi anbefaler at våre forslag til metadata brukes i dette arbeidet.
- En sammenligning av korttidsstatistikkene og deres respektive strukturstatistikker. Det bør være et økt fokus på samstemming mellom korttidsstatistikker og strukturstatistikker.

7.3 ISEE – Integrert system for editering og estimering

Vi mener det er mye funksjonalitet ved ISEE som er relevant for vårt prosjekt. Dette gjelder primært muligheten til å kunne ta ut rapporter fra ISEE, som på en enkel måte kan brukes som input til informasjonsbanken vi har presentert i avsnitt 6.2. Per i dag er det ikke alle metdata vi har omtalt i dette prosjektet, som kan listes ut i ISEE. Et innspill fra prosjektgruppen til de som videreutvikler ISEE er at det skal kunne være mulig å kjøre ut rapporter på de metadata som er nevnt i kapittel 5. Videre etableres det et script som overfører metadataene fra ISEE-applikasjonene til den foreslåtte informasjonsbanken.

I tillegg er det vår oppfatning at applikasjonene PRIS og STRUKTUR først og fremst er egnet for produksjon av utvalgsbaserte statistikker. Spesielt i STRUKTUR, burde det utvikles funksjoner som også er nyttig for registerbasert statistikk.

Begrepsforklaringer

Aktualitet

I teksten refereres det ofte til en statistikkens aktualitet. Med aktualitet mener vi tiden fra statistikkperiodens utløpt til publisering. For eksempel vil detaljomsetningsindeksens aktualitet være 30 dager, når detaljomsetningen for mai måned publiseres den 30. juni. Når vi snakker om høy aktualitet, betyr dette at statistikken publiseres relativt raskt, og motsatt når vi snakker om lav aktualitet.

D-nummer

D-nummer er bygd opp som et fødselsnummer, men har 40 lagt til personens fødselsdato. Alle som kommer til Norge for å arbeide vil få utdelt et fødselsnummer eller et såkalt D-nummer. De fleste som arbeider i Norge på korttidsopphold har et D-nummer.

Estimering

Det vil si å beregne verdien for en ukjent størrelse med data fra et utvalg.

Indikator

En indikator er et mål som er avledet av data og/eller statistikk og indikerer status eller utvikling på et angitt område. For eksempel kan svarprosenten i en utvalgsundersøkelse være en indikator for kvalitet.

Imputering

Imputering vil si at man erstatter en manglende observasjon gjennom en logisk redigering eller ved bruk av statistiske metoder. Imputering brukes både når respondenten unnlater å svare (enhetsfracfall/totalfracfall) og når det er fracfall på et eller flere spørsmål (partielt fracfall).

ISEE – Integrert system for editering og estimering

ISSE omfatter applikasjonene DynaRev (editering), PRIS (beregning av prisindekser) og STRUKTUR (utregning av populasjonstotaler) og er en viktig del av SSBs kvalitetsarbeid. Mens man tidligere skreddersydde produksjonsopplegg for hver enkelt statistikk, er ISEE en felles løsning for revisjon og editering. Tanken er at felles løsninger og gjenbruk av arbeidsprosesser bidrar til å forenkle statistikkproduksjonen, samt gjør det lettere å kvalitetssikre de enkelte statistikkene.

KNR – Kvartalsvis Nasjonalregnskap

Kvartalsvis nasjonalregnskap (KNR) er et system for beregning av kvartalsvise tall og foreløpige årstall i nasjonalregnskapet. KNR er omtalt i kapittel 2.

Kortidsstatistikk

Med korttidsstatistikk menes statistiske undersøkelser med en hyppighet på 3 måneder eller kortere (kvartal, termin, måned uke).

Metadata

Metadata er data som beskriver data.

Register

Et register består av fortegninger, lister, oversikter, m.v. der informasjonen er lagret systematisk slik at opplysninger enkelt kan finnes igjen.

Utvalg

Det er ofte dyrt og komplisert å undersøke hele populasjonen. Ved å trekke en del av populasjonen, det vil si et utvalg, vil man som regel oppnå tilfredsstillende resultater.

Variabel

En variabel er en egenskap ved en statistisk enhet, og den kan derfor anta mer enn én verdi, i form av en numerisk verdi (kvantitativ variabel) eller en kategori fra en klassifisering (kategorisk variabel).

Eksempler på variabler: inntekt, alder, yrke, næring.

Referanser

Bakke, Stein og Todsen, Steinar (2007): Elektrisitetsstatistikk - dokumentasjon og vurdering av kilder, beregningsopplegg for investeringer. Notater 2007/25. Statistisk sentralbyrå.

Evensen, Trude Nygård (2004): Kvalitetsarbeid knyttet til kvartalsvis nasjonalregnskap (KNR). Rapport fra prosjektgruppen. Notater 2004/22. Statistisk sentralbyrå.

Evensen, Trude Nygård (2006): Oljenæringen - Dokumentasjon av kilder, beregninger mv. (SKA-prosjekt). Notater 2006/8. Statistisk sentralbyrå.

Hustoft, Sæbø (2006): Nokre sentrale omgrep knytte til metadata – til bruk i dei felles metadatasystema i SSB. Interne dokumenter 2006/3. Statistisk sentralbyrå.

SSB 2004: Prosjekthåndboka – slik gjør vi det i SSB. Statistisk sentralbyrå.

Statistisk sentralbyrå: Detaljomsetningsindeksen (<http://www.ssb.no/doi/om.html>).

Statistisk sentralbyrå: Kvartalsvis nasjonalregnskap, KNR (<http://www.ssb.no/knr/om.html>).

Statistisk sentralbyrå: Investeringsstatistikk (<http://www.ssb.no/kis/om.html>).

Statistisk sentralbyrå: Omsetningsindeks for samferdsel og reiseliv (<http://www.ssb.no/sroi/om.html>).

Statistisk sentralbyrå: Sysselsatte på korttidsopphold (<http://www.ssb.no/kortsys/om.html>).

Statistisk sentralbyrå: Retningslinjer for europeisk statistikk (<http://www.ssb.no/omssb/retningslinjer.html>).

Stålnacke, Hustoft, Solheim (1999): Vurdering av kvalitet i statistikk. En oversettelse av notater fra Eurostat om kvalitetsrapportering. Notater 99/42. Statistisk sentralbyrå.

Todsen, Steinar (1999): Kvartalsvis nasjonalregnskap. Dokumentasjon av beregningsopplegget. Rapporter 99/25. Statistisk sentralbyrå.

Walle, Mari Aasgaard mfl.(2006): Oljenæringen - Dokumentasjon av kilder, beregninger mv. (SKA-prosjekt). Notater 2006/9. Statistisk sentralbyrå.

Zhang, Faldmo og Lien (2007): ISEE, Integrert system for editering og estimering. Innlegg i forbindelse med Nordisk statistikermøte i Reykjavik, 25. - 28. juni 2007.

Vedlegg A. Prosjektskrivet

Prosjektskriv

Prosjektnavn	Tiltak for å forbedre dialogen mellom fagseksjonene og Kvartalsvis nasjonalregnskap		
Kortnavn	KNR-dialog		
Planlagt startdato	24.10.2007	Planlagt sluttdato	06.06.2008
Oppdragsgiver	Leiv Solheim		
Ansvarlig enhet	Seksjon for samferdsels-og reiselivsstatistikk w/ seksjonssjef Leiv Solheim		
Er prosjektet del av et større hovedprosjekt?	<input type="checkbox"/> Ja	<input checked="" type="checkbox"/> Nei	
Hvis ja, oppgi navn og prosjektnr på hovedprosjekt			
Prosjektleder	Ole Christian Erikstad		
Kvalitetslos	Harald Utne		
Produktnummer	1771 KNR realregnskap		
Godkjent dato	13.12.2007	Av	Leiv Solheim

Bakgrunn for prosjektet - Problembeskrivelse

Mange av korttidsstatistikkene som produseres av SSB sine fagseksjoner inngår i datagrunnlaget til Kvartalsvis nasjonalregnskap (KNR). Dette stiller strenge krav til aktualitet og kvalitet, men foreløpig er det ikke gjort nok for å synliggjøre styrke og svakheter ved statistikkene. Slik informasjon vil være nyttig både for KNR og fagseksjonene og omfattes av fremtidige internasjonale krav.

I noen tilfeller oppstår det avik mellom inn-data fra korttidsstatistikkene og ut-data som presenteres i KNR. Dette kan være en kilde til usikkerhet hos de ansvarlige ved fagseksjonene. For å sikre at SSB taler med en stemme, er det behov for å belyse kilder til endringer ved innarbeiding av korttidsstatistikker i KNR-systemet, samt etablere rutiner for intern formidling av denne typen informasjon. I denne sammenheng vil det også være behov for å se nærmere på, i hvilken grad, det er samsvar mellom KNR sine krav til dataleveransene og hva som faktisk leveres.

Det vil være nyttig å ta utgangspunkt i **Notat 2004/42 - Kvalitetsarbeid knyttet til kvartalsvis nasjonalregnskap**. Målsetningene for prosjektet er formulert slik at de i størst mulig grad skal harmonere med FOSS-tankegangen.

Effekt mål

Prosjektet skal bidra til:

- en bedre synliggjøring av kvaliteten på korttidsstatistikkene
- å skape en felles forståelse av datagrunnlaget mellom fagseksjonene og KNR
- en bedre og sikrere tolkning av avik mellom fagstatistikkene og KNR-tallene

Målgruppen for prosjektet er produsenter av korttidsstatistikk som inngår i datagrunnlaget til KNR samt Seksjon for nasjonalregnskap (S930).

Resultat mål

- 1: Utarbeide opplegg for kvalitetsbeskrivelse av et utvalg korttidsstatistikker som inngår i datagrunnlaget til KNR.
- 2: Avdekke mulige kilder til avik ved innarbeiding av korttidsstatistikk i KNR-systemet, samt vurdere behovet for intern formidling av slik informasjon.
- 3: Skrive en rapport i serien notater.

Faglige problemstillinger

Til resultatmål 1:

En kvalitetsbeskrivelse innebærer at man kartlegger viktige metadata som størrelsen på populasjonene og eventuelt utvalgene, estimerings- og beregningsmetoder, sammenliknbarhet med andre statistikker samt feilkilder og usikkerhet. DESAP, ISEE samt håndbook i risikoanalyse er eksempler på eksisterende verktøy som muligens kan benyttes i dette arbeidet.

Det vil være nyttig å ta utgangspunkt i "Om statistikken" for de aktuelle korttidsstatistikkene.

Til resultatmål 2:

Prosjektgruppen bør ta utgangspunkt i hvorvidt KNR mottar det tallmaterialet de faktisk etterspør. Samtidig bør det kartlegges eventuelle forbedringspotensialer av fagstatistikkene, sett fra KNR sitt ståsted. I tillegg bør prosjektgruppen se på hvilke tilbakemeldingsrutiner man i dag har mellom KNR og fagseksjonene, og hvordan disse eventuelt kan forbedres.

Det må på et tidlig tidspunkt i prosjektet avklares hvor mye ressurser KNR stiller med til gjennomføringen av resultatmål 2.

Det flere måter KNR kan involveres i prosjektet på. Et alternativ er å bestille en rapport/notat fra KNR. Et annet og kanskje mer effektivt alternativ er å intervju nøkkelpersoner hos KNR, der prosjektgruppa selv sammenfatter en delrapport.

Grunnlagsdokumentasjon, aktuelle linker og andre prosjekter

- Evensen (2004): Kvalitetsarbeid knyttet til kvartalsvis nasjonalregnskap (KNR). Rapport fra prosjektgruppen. Notater 2004/4
- Bakke og Todsen (2007): Elektrisitetsstatistikk - dokumentasjon og vurdering av kilder, beregningsopplegg for investeringer. Rapport fra prosjektgruppen. Notater 2007/25
- Eurostat (2001): Methodology of industrial short-term statistics. Rules and Recommendations.
- Johannessen, Randi (1999): Kvalitetssikring av korttidsstatistikk. Notater 1999/60.
- Statistiska centralbyrån (2001): Resultater fra SCBs satsing på systematisk kvalitetsarbeid. Utredning enligt oppdrag i Reguleringsbrev før SVB år 2000 (datert 2002-03-01).
- Steinar Todsen (1999): Kvartalsvis nasjonalregnskap. Dokumentasjon av beregningsopplegget. Rapporter 99/25.
- Pia E. Tønjum (1999): Teknisk dokumentasjon av beregningsopplegget for kvartalsvis nasjonalregnskap (KNR). Notater 99/19.
- Pia E. Tønjum (2003): Tilbakemelding/dokumentasjon av prosjektet: Avstemning av KNR mot nye årstall ifølge tallrevisjonen. Notater 2003/65.
- Om statistikken for de utvalgte fagstatistikkene

Vedlegg B: Fagseksjoner som leverer data til KNR

S215 - Seksjon for energistatistikk

S225 – Seksjon for inntekts- og lønnsstatistikk

S260 - Seksjon for arbeidsmarkedsstatistikk

S270 - Seksjon for utenrikshandel

S430 - Seksjon for primærnæringsstatistikk

S440 - Seksjon for samferdsel og reiselivsstatistikk

S460 - Seksjon for bygg- og tjenestestatistikk

S470 - Seksjon for industristatistikk

S940 - Seksjon for kredittmarkedsstatistikk

S950 - Seksjon for offentlige finanser

Vedlegg C: Korttidsstatistikker som er sentrale i KNR

- Detaljomsetningsindeksen (sentral for beregning av konsum i husholdninger og - indirekte - produksjon i næringen varehandel)
- Omsetningsindeksen for eiendomsdrift, forretningsmessig tjenesteyting og utleievirksomhet (sentral i beregningen av produksjon i næringen forretningsmessig tjenesteyting)
- Kvartalsvis omsetningsindeks for samferdsel og reiseliv (produksjon av transport og samferdselstjenester, hotell- og restaurantvirksomhet)
- Produksjonsindeks for industrien (produksjon industri og kraftforsyning, produksjon av olje)
- Produksjonsindeksen for bygge- og anleggsvirksomhet (produksjon i bygg- og anleggsnæringen, sentral i beregningen av investeringer i bygg og anlegg)
- Byggearealstatistikken (sentral i beregningen av investeringer i bygg og anlegg)
- Investeringsstatistikken for olje- og gassvirksomheten (investeringer i olje- og gassvirksomhet)
- Investeringsstatistikken for industri, bergverksdrift og kraftforsyning
- KOSTRA (produksjon, sysselsetting, konsum og investeringer i kommunal forvaltning)
- Statsregnskapets inntekter og utgifter (produksjon, sysselsetting, konsum og investeringer i statlig forvaltning, eksklusive sykehusvirksomheten)
- Kvartalsvis regnskapsinformasjon fra helseforetakene (produksjon, sysselsetting, konsum og investeringer i den statlige helsenæringen)
- Utenrikshandelen med varer. Volum- og prisindekser (eksport og import av varer)
- Konsumprisindeksen, KPI (priser i husholdningenes konsum, priser for produksjonen i enkelte tjenestenæringer)
- Vareprisindeksen for industrien (industriens produksjonspriser for leveranser til hjemmemarkedet)
- Arbeidskraftundersøkelsen (sentral i beregningen av sysselsatte personer fordelt på yrkesstatus (lønnstakere eller selvstendig næringsdrivende) og for sysselsatte personer totalt)
- Registerbasert sysselsettingsstatistikk (sentral i beregningen av sysselsatte personer fordelt på næring)
- Lønnstatistikk, lønnsindeks, kvartalsvis (sentral i beregningen av lønn per normalårsverk for de fleste næringer utenom offentlig forvaltning)
- I tillegg benyttes diverse indikatorer som gjestedøgn ved norske hoteller, avreiste passasjerer fra norske flyplasser, antall personkilometer og tonnkilometer transportert med jernbanen m.fl.

Vedlegg D. Eksempel på tidsplan for KNR

øyb 31.3.08

TIDSPLAN KNR 1. KV. 08

INNARBEIDET STATISTIKK

MANDAG	21. APR	TESTKJØRING S930	
MANDAG	5. MAI	1. KJØRING	Overnattingsstatistikk feb, VKI mars, KPI mars, VPPI mars, Omsetn. ind feb, Samferdsel og reiseliv 4. kv. 07, UHS 3. kv. 07, Prodind. ind feb, BAR feb, AR/AKU, oljeinvest. 3. kv
ONSDAG	7. MAI	2. KJØRING (tabeller til arb.møter)	Prodind. Ind. mars
ONSDAG	7. MAI	ARBEIDSMØTE S930/Berørte fagseksjoner S460 KL. 12.30-13.30, rom DU01/2M6	
TORS DAG	8. MAI	ARBEIDSMØTE S930/Berørte fagseksjoner S440 KL. 12.30-13.30, rom DU01/2M6	
FREDAG	9. MAI	ARBEIDSMØTE S930/Berørte fagseksjoner S215/470/270 KL 10.00-11.30, rom K11 - 714 S950. KL 13.00-14.00, rom K11 - 714	
TIRSDAG	13. MAI	ARBEIDSMØTE S930/Berørte fagseksjoner S260/950/970/530 kl.10.00-11.30 rom B214/2M2	
TIRSDAG	13. MAI	3. KJØRING	INV Industri (ca 80%).Registerbasert syss/AKU, STAT/KOMMUNE 4. kv. 07
ONSDAG	14. MAI	KT-MØTE KL 14.00 – 15.30, 7M3	
TORS DAG	15. MAI	DS-UTKAST SENDES UT	
FREDAG	16. MAI	GODKJENNINGSMØTE KL.14.30-16.00, rom 3M1.	
FREDAG	16. MAI	4. KJØRING (LÅSING AV TALL)	Prodind BoA, INV industri (ca 90 %)
ONSDAG	21. MAI	KNR vedlegg + tabeller leveres til ØA	
ONSDAG	21. MAI	PUBLISERING KNR	

KVARTALSREGNSKAP 2007 - 2008

AR/AKU	= Utdrag fra arbeidstakerregisteret (avstemt mot AKU på totalen)
BAR	= Byggearealstatistikk
DOI	= Detaljomsetningsindeksen
KOSTRA	= Regnskap for kommuneforvaltningen
KPI	= Konsumprisindeksen
LØNN	= Kvartalsvis lønnstatistikk, enkelte hovednæringer
Prodind. bygg og anlegg (selvforklarende)	
Prodind. ind	= Produksjonsindeks for industri
STAT	= Statsregnskapet
Sykehusforetak	= De 5 regionale statlige sykehusforetakene
UHS	= Utenrikshandel med varer; volum og prisindekser
UR	= Utenriksregnskapet
VPPI	= Utdrag av Vareprisindeksen; industriens priser på leveranser til hjemmemarkedet