

# Hva skal fylles inn i de ulike feltene i Vardok? – En kort veiledning

Hvis det er noe du lurer på, ta kontakt med Anne Gro Hustoft (faglig veileder for Vardok).

Fig. 1 Vinduet som brukes for å legge inn en ny variabel. De gule feltene er obligatoriske å fylle ut.

\* Betyr i det følgende at feltene er obligatoriske å fylle ut for å få lagret variabelen i Vardok.  
*Det anbefales imidlertid at en fyller ut så mange felter som mulig, da dette vil øke nytten av variabelen i seinere bruk.*

## **Id:**

Dette er et nummer tildelt av VarDok.

## **Navn\*:**

Her legger du inn navnet på variabelen. Dette skal være et navn som er forståelig, ikke en forkortelse slik en f.eks. bruker i Datadok; dvs. i Vardok skal variabelnavnet være "Landbakgrunn" og ikke "innv\_landbak" som i Datadok. En skal heller ikke bruke generelle betegnelser som f.eks. "Spørsmål 11". Hvis spørsmål 11 f.eks. er "Hva er din årslønn?", skal tilhørende variabel være "Årslønn".

Prøv å unngå lange navn og legg ikke hele definisjonen i variabelnavnet. Hvis variablene i utgangspunktet har lange navn, bør du vurdere om navnet kan kortes ned, og noe av beskrivelsen heller legges i definisjonen. Ikke knytt variabelen til en bestemt undersøkelse eller datasett dersom den er egnet til gjenbruk, f.eks. ikke spesifiser FoB i navnet når det er snakk om Bydel (selv om denne variabelen også brukes i FoB).

#### **Kortnavn:**

Dette er en forkortet versjon av Navn, noen ganger brukt på grunn av begrensninger i maksimum antall karakter i andre systemer (også kjent som navn i SAS eller variabelnavn i DataDok). Dersom variabelen har samme variabelnavn i alle filbeskrivelser i Datadok (noe som oftest ikke er tilfellet), kan det på sikt være hensiktsmessig å sette inn dette som kortnavn. Hvis ikke, anbefaler vi at du lar være å fylle ut dette feltet.

#### **Definisjon\*:**

Her skal en sette inn definisjonen av variabelen. Det er viktig at en ikke bare gjentar variabelnavnet – brukeren skal bli klokere etter å ha sett definisjonen. Prøv å lage en kort og konsis definisjon. Denne skal kunne forstås av ikke-ekspertbrukere (f.eks. brukere av Statistikkbanken). Det er derfor viktig at en vurderer ordbruk (er dette et fremmedord som bare innvidde forstår?), og at en ikke bruker referanser til poster i regnskap etc. (som også de færreste forbinder noe med). Begrens deg til selve definisjonen, det finnes andre felter som passer for mer utfyllende opplysninger (Interne og eksterne kommentarer og Beregnings-feltet). Husk at dette er en overordnet definisjon som skal kunne brukes uavhengig av datasett. Dette betyr at du generelt ikke skal ha med opplysninger som er knyttet til spesifikke datasett. Slik informasjon må eventuelt ivaretas i Datadok eller Om statistikken som gir informasjon om spesifikke datasett/statistikker.

- Dette betyr bl.a. at opplysninger om måleenheter ikke skal med; f.eks. når en definerer Inntekt som ” Yrkesinntekter, kapitalinntekter, skattepliktige og skattefrie overføringer, i løpet av kalenderåret. Utlignet skatt og negative overføringer er trukket i fra.” skal en ikke ha med ”målt i kroner”. I noen datasett måles dette i kroner, i andre kanskje i 1000 kroner, men selve definisjonen av inntekt skal være uavhengig av måleenhet.
- Av samme grunn tar en heller ikke med populasjon i definisjonen. Definisjonen av inntekt er den samme enten en måler den for f.eks ”alle personer mellom 16 og 74 år” eller ”alle kvinner ansatt i undervisningssektoren”.

Definisjonen skal heller ikke inneholde definisjon av andre tilknyttede variabler, disse bør heller lenkes, f.eks. hvis definisjonen av Variabel A er at den er summen av Variabel B og Variabel C, skal ikke definisjonen av B og C tas med i definisjonen av A, men en lenker variabel A til B og C (se Lenket i Vardok? nedenfor).

#### **Eid av og Kontakt:**

Fylles ut automatisk av VarDok basert på informasjonen du oppgir når du logger deg inn. Kontakt kan endres etter behov.

#### **Sensitivitet:**

Her er det flere valg i en meny, men i praksis bruker vi mest Sensitiv og Ordinær. Alle metadata er i utgangspunktet ordinære, men dette feltet skal markere om dataene tilhørende en slik variabel er sensitive. Dette betyr at selve definisjonen av variabelen ”Dødsårsak” ikke er sensitiv, men data vedrørende dødsårsak vil være det. Denne variabelen vil derfor ha Sensitiv i dette feltet.

#### **Gyldig fra/til:**

En variabel kan endre definisjon over tid. Disse feltene markerer når den gitte definisjonen er gyldig fra, og hvis den ikke lenger er gyldig, når den var gyldig til. Åpent Gyldig til-felt markerer at den gitte definisjonen fortsatt gjelder.

**Statistisk enhet\*:**

Statistisk enhet er den enheten variabelen skal fortelle noe om, f.eks. person, husholdning eller foretak. Statistisk enhet velges fra en meny. Dersom du ikke finner den enheten du er på jakt etter, må du ta kontakt med faglig veileder for Vardok for å få lagt den inn i menyen. Samme variabel (f.eks. Inntekt) kan ha ulike definisjoner alt etter om den statistiske enheten er person eller foretak.

I mange tilfeller vil en imidlertid bruke samme definisjon av en variabel både for f.eks. bedrift og foretak. Da er det ikke hensiktsmessig å legge inn samme variabel to ganger, bare med forskjellig statistisk enhet. Legg heller inn variabelen en gang (vi anbefaler å legge den inn på laveste nivå – altså med bedrift som statistisk enhet), og skriv at ”Definisjonen er den samme både for foretak og bedrift” i Eksterne kommentarer-feltet.

**Intern referanse:**

Her kan en legge ei lenke til dokumentasjon som gir viktig informasjon om variabelen, og som kun skal vises internt i SSB (f.eks. et dokument på Q/Dok eller ei side på Intranett). Dersom du vil henvise til flere interne dokumenter, må dette gjøres i Interne kommentarer, men her kan de ikke legges inn som hyperkobling. Søk-knappen brukes for å søke etter, og legge inn, relevante interne dokumenter.

**Ekstern referanse:**

Her har du mulighet til å legge inn ei lenke til noe som kan vises utenfor SSB (f.eks. et SSB-dokument eller ei side på [www.ssb.no](http://www.ssb.no)). Dersom du vil henvise til flere dokumenter, må dette gjøres i Eksterne kommentarer, men her kan de ikke legges inn som hyperkoblinger. Ekstern referanse brukes hvis det finnes dokumenter som gir viktig informasjon om variabelen, og som du gjerne vil henvise eksterne brukere til. Dette kan f.eks. gjelde kompliserte variabler som indikatorer der det ofte finnes egne dokumentasjonsrapporter.

**Beregning:**

Dersom en variabel bygger på andre variabler (Var A=varB+varC, ev. var A= post 100+post 120), kan en legge inn en enkel beregning her. For mer kompliserte beregninger bør en heller benytte Ekstern referanse til å henvise til et dokumentasjonsnotat (jamf. Ekstern referanse).

**Interne kommentarer:**

Tilleggsinformasjon for interne brukere. Dette kan f.eks. være oversikt over interne publikasjoner som er nyttige i forbindelse med denne variabelen, eller annen informasjon som vi ikke vil vise utenfor SSB.

**Eksterne kommentarer:**

Tilleggsinformasjon som også er aktuell for eksterne brukere. Dette kan f.eks. være opplysninger om spesiell bruk av variabelen, endringer som har skjedd i variabelen etc.

**Klassifikasjon:**

Dersom en variabel beskrives vha. en klassifikasjon, kan en bruke Stabas-knappen, og gå inn og velge den rette klassifikasjonen fra Stabas (f.eks. henter variabelen Landbakgrunn sine verdier fra klassifikasjonen Standard for land og statsborgerskap i personstatistikk). Vi legger fortsatt klassifikasjoner inn i Stabas (klassifikasjonsdatabasen) så den er ennå ikke fullstendig. Hvis du ikke finner klassifikasjonen du er på jakt etter, anbefaler vi at du skriver hvilken klassifikasjon som brukes, i feltet for Eksterne kommentarer. Når en klassifikasjon er valgt, kan du dobbelsjekke klassifikasjonen ved å dobbelklikke på navnet i klassifikasjonsfeltet. Du kan enten lenke variabelen på klassifikasjonsnivå (f.eks. Standard for næringsgruppering) eller på versjonsnivå (f.eks. SN2002). Det første gjøres når en har en variabel der variabeldefinisjonen sjelden endrer seg, f.eks. skolekommune, men der klassifikasjonen variabelen er knyttet til, endrer seg oftere. Når en lenker til Standard for kommuneinndeling, vil brukeren få tilgang til de ulike versjonene av kommuneinndelingen. Dersom en imidlertid hadde hatt variabler knyttet til næringsvirksomhet i Folke- og boligtellingsen i 2001, ville en ha lenket disse variablene til den spesifikke versjonen SN94 av Standard for næringsgruppering siden det bare var den som var aktuell i 2001. Du kan også knytte variabelen til et bestemt nivå i en versjon dersom disse kategoriene er de eneste som brukes for den aktuelle variabelen.

**Kodeliste:**

*Dette feltet bør ikke brukes lenger.* Det er en overlevning fra Vardoks tidligere dager da det ikke var noen forbindelse til Datadok, og kodelistene derfor måtte legges inn i Vardok. Dersom en variabel har ei kodeliste, skal denne dokumenteres i Datadok, og så bør det opprettes ei lenke fra Vardok til kodelista i Datadok. Dette gjør at en forbeholder oppdatering av kodelister for Datadok (se lenger nede ; Filer i Datadok) . Vi skiller mellom standard klassifikasjoner, som ligger i Stabas, og mer datasettspesifikke kodelister (f.eks. 1=Ja, 2=Nei, 3=Vet ikke) som skal dokumenteres i filbeskrivelsene i Datadok.

**Statistikkemne\*:**

Her velger du den kategorien fra SSBs emneinndeling som variabelen tilhører. Valget foretas fra en meny, og du kan velge hvor detaljert nivå du synes det er relevant å legge inn. Se også Statistikk.

**Statistikk:**

Når du har valgt statistikkemne, får du ei liste med statistikkene som tilhører det valgte emnet, bak Statistikk-knappen. Dersom variabelen din er sterkt knyttet til en statistikk, kan du legge inn denne statistikken i dette feltet. Dersom den er en variabel som brukes i mange ulike statistikker, f.eks. alder, lar du bare statistikkfeltet stå åpent.

**SSB kilde:**

SSB kilde velges fra menyen bak SSB kilde-knappen, og beskriver det ”stedet” i SSB der variabelen først ble definert (hvis den er definert av SSB). Dette feltet har altså ikke noe med datakilde å gjøre, det gjelder kilden for selve variabeldefinisjonen. Kilden vil som oftest være et register eller en statistikk. Dersom du ikke finner din SSB-kilde i menyen, eller du ser noe som er feil, gi beskjed til faglig veileder for Vardok.

**Ekstern kilde:**

Ekstern kilde velges fra menyen bak Ekstern kilde-knappen, og beskriver ”stedet” utenfor SSB (f.eks. Skattedirektoratet) der variabelen først ble definert (hvis den er definert utenfor SSB). Dette feltet har altså ikke noe med datakilde å gjøre, det gjelder kilden for selve variabeldefinisjonen. Kilden vil som oftest være en organisasjon eller et register. Dersom du ikke finner din SSB-kilde i menyen, eller du ser noe som er feil, gi beskjed til faglig veileder for Vardok.

I noen tilfeller der variabelen er eksternt definert, har brukerne også valgt å fylle ut SSB-kilde for å markere i hvilken sammenheng i SSB variabelen etableres. Dette er OK, men vi anbefaler at i alle fall en av kildene fylles ut (selv om dette ikke er gjort obligatorisk).

**Filer i Datadok:**

Denne knappen brukes når du vil knytte variabeldefinisjonen i Vardok til en variabel i en filbeskrivelse i Datadok. Gjennom denne lenkingen oppnår du to ting. 1) Når du er inne i Datadok, kan du for enhver lenket variabel se den tilhørende Vardok-definisjonen. Dersom du skriver ut en filbeskrivelse f.eks. for å la den følge et datasett til en bruker, vil du også få med Vardok-definisjonen for variablene. 2) Dette er også måten du skal knytte ei kodeliste til en variabel på. Når du lenker en Vardok-variabel til en Datadok-variabel med kodeliste, vil du også kunne se denne kodelista fra Vardok. (Det betyr at kodelista kun ligger i, og oppdateres i, Datadok, men er tilgjengelig også for brukerne i Vardok). Dersom en variabel endrer kodeliste fra et år til et annet (f.eks. ulike grupperinger av alder da skolealder endret seg fra 7 til 6 år), knytter en bare Vardok-variabelen opp til samme variabel i begge disse to årgangene av filbeskrivelser, og får dermed med begge kodelistene som er knyttet til variabelen. En kan lenke opp til alle filbeskrivelser der en synes det er nyttig å ha med Vardok-definisjonen. I tilfeller der informasjon fra Datadok publiseres eksternt (f.eks. i Mikrodata for

forskere), brukes Vardok-tilknytningen til å bestemme hvilke Datadok-variabler som kan vises eksternt (bare de som har en definisjon fra Vardok knyttet til seg, vises). Bak Datadok-knappen vil du kunne se alle Datadok-filbeskrivelser som den aktuelle variabelen er lenket til; både de som tilhører egen seksjon, og de som tilhører andre seksjoner som bruker din variabeldefinisjon. Samme variabeldefinisjon kan også knyttes til flere variabler i samme filbeskrivelse. Dersom en f.eks. har lagt inn en generell definisjonen av ”Månedlig sosialhjelp” i Vardok , kan denne variabelen knyttes til Datadokvariablene Jan, Febr etc. som inneholder sosialhjelp for januar, februar etc. De beskriver alle månedlig sosialhjelp – spesifisert på måned.

#### **Tabeller i Statbank:**

Dette oppdateres automatisk. Dersom denne variabelen er brukt i noen tabeller i Statistikkbanken, vil radioknappen ha verdien Ja. Da kan du trykke på knappen ved siden av, og få se i hvilke Statbanktabeller variabelen brukes.

#### **Brukt i MetaDB:**

Dette oppdateres automatisk. Dersom denne variabelen er brukt i MetaDB (metadatabasen til FD-Trygd og Norsk utdanningsdatabase), vil radioknappen ha verdien Ja. Da kan du trykke på knappen ved siden av , og få se i hvilke MetaDB-variabler den aktuelle Vardok-variabelen brukes.

*Hvis du gjør endringer i navn, definisjon, statistisk enhet eller gyldighetsperiode i en variabel som er knyttet til Datadok, Statbank eller MetaDB, vil du få et spørsmål om du ønsker å informere dem som har knyttet seg til variabelen, om endringene. Dersom du svarer ja, vil opplysningen om hvilke endringer som er gjort, automatisk sendes til dem som har knyttet egne variabler til din Vardok-variabel. Vi anbefaler at du ikke sender slike meldinger dersom det bare dreier som om mindre ting som oppretting av skrivefeil etc.*

#### **Lenket i Vardok?**

Lenket- knappen gir deg mulighet til å lenke din variabel til andre Vardok-variabler. Dersom du f.eks. bruker andre variabler (f.eks. B) for å definere den aktuelle variabelen (A), trenger du ikke gjenta definisjonen av B i definisjonen av A, men kan lenke variabel B til A. Når noen senere tar opp definisjonen av A, vil de da samtidig få tilgang på definisjonen av B. Du kan også lenke variabler som ikke direkte brukes i definisjonen, men som har en sammenheng med den aktuelle variabelen, og som du ønsker å henlede oppmerksomheten på. Denne lenkingen er en manuell operasjon. Bak lenkeknappen finnes imidlertid også en automatisk funksjon. Der er en radioknapp Tid som er aktiv dersom det finnes flere versjoner av den aktuelle variabelen, dvs. dersom definisjonen har endret seg over tid. Her vil en da få tilgang på de ulike versjonene som utgjør et slags arkiv for denne variabelen.

#### **Godkj. Internt SSB:**

Når seksjonen mener at informasjonen om variabelen kan vises internt i SSB, hakes denne boksen av.

#### **Def. godkj. For Internett:**

Når seksjonen mener at informasjonen om variabelen kan vises eksternt (f.eks. i Statbank og Om statistikken), hakes denne boksen av.

Dersom ingen av boksene er haket av, vises ikke variabelen i variabeloversikten i Vardok til brukere med lesetilgang. Det er da bare de som går inn med seksjonens passord, som får tilgang til variabelen.

## Noen momenter til ettertanke når en skal velge hva som skal legges inn som variabel

### SAS

Det kan mange ganger være vanskelig å se hva som er en variabel og hva som er ei kodeliste, særlig fordi det lages kombinasjoner av disse i tabeller i SAS. En "SAS- variabel" vil ofte ikke være det samme som det vi kaller variabel i Vardok. Eksempel: Vi har variabelen Henvendelsesgrunn i Vardok, med tilhørende kodeliste av forskjellige henvendelsesgrunner (rusproblemer, familieproblemer, psykiske problemer osv.). I SAS kan dette bli oppdelt i flere SAS-variabler; Henvendelsesgrunn rusproblemer, Henvendelsesgrunn familieproblemer, Henvendelsesgrunn psykiske problemer osv. hver med en kodeliste 0/1/missing eller Ja/Nei/'blank' osv. Hver enkelt av disse er variabler i SAS-forstand, men ikke i Vardok der vi altså ser dette som én variabel (Henvendelsesgrunn) med tilhørende kodeliste/klassifikasjon over henvendelsesgrunner.

### Begreper

- Det kan være vanskelig å skille mellom variabler og begreper. Publiseres det tall for denne termen, f.eks. Produktinnsats, er det sannsynligvis en variabel. Et begrep publiseres det ikke tall for, men termen er viktig for å forklare noe som har med statistikken å gjøre ( f.eks. Basisår)

## Eksempler på variabeldefinisjoner og statistiske enheter

**Alder:** Personens alder ved utgangen av året  
Statistisk enhet: Person

**Toppskattegrunnlag:** Toppskattegrunnlaget omfatter personinntekt lønn, personinntekt primærnæringer, personinntekt annen næring og personinntekt pensjoner. Toppskatten til staten blir beregnet ut fra dette grunnlaget.  
Statistisk enhet: Person

**Egenvekt på kjøretøyet:** Egenvekt er vekten av kjøretøyet i full driftsferdig stand med permanent montert utstyr, smøreolje, vann og fulle drivstofftanker, samt eventuelle reservehjul, verktøy mv.  
Statistisk enhet: Kjøretøy