



LHD og SCK, 30.10.2024

Til stede: Roy Nielsen, Forskningsinstituttens fellesarena (FFA)
Lars Dommermuth, SSB/Forskningsavdelingen
Kjersti Misje Østbakken, Forskningsinstituttens fellesarena (FFA)
Anneli Pellerud, Folkehelseinstituttet (FHI)
Guri Kvesetberget, Lin Hege Austad Dalen og Kenneth Schnelle SSB/Seksjon for mikrodata

Kopi: Ann-Kristin Brændvang SSB/Avdeling for person- og sosialstatistikk

Referat fra møtet i Brukerutvalget for mikrodatatjenesten i SSB 15. oktober 2024

Sak 1. Status for SSB sin mikrodatatjeneste v/Kenneth Schnelle (SSB)

- Det ble gitt en statusoppdatering for både microdata.no og tradisjonell tilgang til mikrodata.
- Interesse for testingen av Kreftregisteret i microdata.no. Det gjenstår fremdeles noen avklaringer / testing før data kan gjøres tilgjengelig for andre.
- Neste år preges av et stramt budsjett, hovedsakelig på grunn av flytting til skyplattform og flytting til Jernbaneplassen.
- Diskusjon rundt samarbeid med helseregistrene. Det ble stilt spørsmål om det vil være mulig å koordinere søknad til SSB og søknad til helseregistrene, slik at prosjektene får tilgang til de ulike registrene på samme tid. Det ble tatt opp et eksempel der SSB har levert for halvt år siden, der helseregistrene først er klare til å gi tilgang nå. Resultatet er mismatch i årganger fra de ulike registreierne. Dette resulterer i endringssøknad til SSB.
- Det ble forespurt om SSB foretar en kontrollsjekk av innkommende søknader og om det gis tilbakemelding til prosjektene fortløpende om manglende dokumentasjon. SSB informerte om at dette erfaringsmessig ikke er hensiktsmessig, da det ikke lar seg gjøre å gi en fullgod tilbakemelding uten å sette seg godt inn i prosjektene. Praksis er derfor at vi gir en samlet tilbakemelding, både angående manglende dokumentasjon og manglende samsvar i mottatt dokumentasjon, først etter at saksbehandlingen er påbegynt.
- SSB informerte om at det fra 2025 ikke lenger vil være mulig med forhåndsfakturering. Etter møtet er det kommet ny informasjon i saken, og endelig avklaring er enda ikke på plass.
- Nettside for brukerutvalg er opprettet på [ssb.no](https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/brukerutvalg-for-statistisk-sentralbyras-mikrodatatjeneste). <https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/brukerutvalg-for-statistisk-sentralbyras-mikrodatatjeneste>
- Gjennomgang av endring i variabellister, spesielt med tanke på merknader til førstelinje. Viktig at siste publiserte versjon benyttes ved søknad om tilgang til mikrodata. Dette vil redusere saksbehandlingstiden.
- Informasjon om oppfølging av slettefrister. Før sommeren sendte SSB ut samlet informasjon om aktive prosjekter til alle institusjoner. 96 av totalt 498 aktive prosjekter hadde ikke slettet data i tråd med vilkår i avtale / vedtak fra SSB.
- Det vil kunne komme ytterligere krav fra SSB om krav til lagring. Det ses på løsninger for bruk av sikre analyserom, evt. selvdeklarasjon fra institusjonene. Det ble fremhevet at det ikke burde være SSB sitt ansvar å vurdere sikkerheten ved institusjonenes serverer ei heller at vi burde purre på slettefrister.
- Filformater:
Parquetfiler er det valgte formatet på den nye dataplattformen.



Det vil være mer effektivt om dette benyttes som valgt format fremover. SSB vil forslå dette som nytt standard format i leveranser og legge ut informasjon om innlesing av formatet på våre nettsider.

- Inntektsdata blir tidligst tilgjengelig i februar, på grunn av utsatt publisering.
- PIAAC – voksenes ferdigheter:

Den nye undersøkelsen (2022-2023) har innhentet samtykke for kobling mot registeropplysninger. Det vil være begrensninger når det kommer til hvilke opplysninger PIAAC kan kobles mot og det vil komme faste datasett som kan kobles innenfor samtykket. Det vil bli utarbeidet en egen variabelliste for denne undersøkelsen.

Publisering av rapporten fra undersøkelsen er satt til 10.12.2024. Etter dette kan det søkes om tilgang. SSB vil publisere inngangsside og variabelliste så fort som mulig.

Sak 2. Kobling på navn og kobling på geografi

Informasjon om grunnprinsippet og avanserte metoder som benyttes ved navnematching. SSB oppfordrer prosjektene selv til å innhente org. ved kobling av bedrifter/foretak og fødselsnummer der det er mulig.

Det vil komme informasjon om standardisering / krav ved navnematching.

Hva er minimumsinformasjon for å få gode nok treff? Navn alene vil ikke være tilstrekkelig, annet enn for de med særegne navn. Opplysninger som kjønn, fødselsdato og bosted bedrer matchingsprosenten.

Det ble stilt spørsmålsteget ved at SSB påtar seg å gjøre slike koblinger. Prosjektene burde innhente fnr. i de tilfellene hvor man vet at det skal kobles på registerdata. Dette er imidlertid ikke mulig i alle tilfeller. Patentdata ble trukket frem som et eksempel der navnematching må gjøres, da patenter kun registreres på navn.

Geodata:

Naturskadeerstatning og støydata er typiske eksempler på informasjon som ofte nødvendiggjør geokoding.

To løsninger ved påkobling av slike opplysninger:

- 1) SSB kan oversende koordinat på midlertidig løpenummer til prosjektet som kobler på ønskede opplysninger og sender disse tilbake til SSB
- 2) Opplysningene kan oversendes SSB som foretar omkodingen.

Det vil komme en spesifisering på nettsidene om hvilke formater som må benyttes for at SSB skal kunne få treff, dvs. for at kobling skal kunne gjøres.

Det er også ønsket å få laget et standard datasett for eksempelvis reiseavstander til jobb, skole, sykehus med mer, samt avstander som vi ser går igjen i bestillinger.

Sak 3. Eventuelt

Spørsmål rundt variabellister og om det planlegges flere: De opplysningene som er mest brukt er dekket av variabellistene som ligger ute på nettsidene. Det jobbes kontinuerlig med å



Statistisk sentralbyrå
Statistics Norway

utarbeide nye variabellister. Foreligger konkrete planer for FoBene bakover i tid, konkurs og historiske befolkningsregister.

SSB tar gjerne imot innspill til tema til neste møte. Send forslag til SSB v/Kenneth Schnelle på sck@ssb.no.