

# Forløpsdatabasen-Trygd

*Forløpsdatabasen-Trygd (FD-Trygd) skal samle statistiske data på individnivå til velferdsforskningen. Trygdesystemet har en sentral plass i vårt velferdssystem og opplysninger om personers trygdestatus og trygdeforhold utgjør en stor del av dataene i basen. Videre finnes opplysninger om sysselsetting, arbeidssøking, økonomisk sosialhjelp, utdanning med videre, noe som gjør FD-Trygd til en unik kilde.*

## En kilde for velferdsforskning

Opplysningene i FD-Trygd er fra og med 1992 og bygger på administrative data hentet inn fra Rikstrygdeverket, Aetat og Statistisk sentralbyrå. Opplysningene omfatter hele befolkningen. Temaområdene eller ordningene som det er opplysninger for i basen, kan grovt inndeles slik:

- forskjellige typer folketrygdpensjon (alders-, uføre- og etterlattepensjon)
- avtalefestet pensjon
- pensjon og medlemskap i Statens Pensjonskasse
- foreløpig uførestønad
- avslag på uførepensjon
- enslig-forsørger-stønader
- grunn- og hjelpestønad
- medisinsk rehabilitering
- yrkesmessig attføring
- fødsels- og sykepenger
- arbeidssøking
- utdanning
- sysselsetting
- økonomisk sosialhjelp
- demografi (blant annet opplysninger om sivilstand, bosted, landbakgrunn og flytting)
- inntekt og formue.

For mange av disse temaområdene er dataene ferdigstilt til og med årgangene 1999/2000, mens det for andre områder ikke finnes data for senere år enn 1997/1998. Arbeidet med å ajourholde og legge inn data i FD-Trygd skjer fortløpende.

Dataene i basen er organisert på en måte som gjør det mulig å anvende forskjellige forskningsdesign ved analyser som gjennomføres på grunnlag av data fra basen. Det er tatt spesielt hensyn til at de statistiske opplysningene i basen skal kunne nyttes til forløpsanalyser, det vil si til analyser av for eksempel hvor lenge en person befinner seg i en bestemt trygdestatus før overgang til en ny og endret status, og hvilke faktorer som påvirker at personen går fra en status til en annen. Slike studier forutsetter data på individnivå, samt opplysninger over tidsperioder for hver person (forløpsdata eller hendelsesorienterte data).

## Datadesign

For statistiske data på individnivå er de tre mest vanlige datadesign skissert i figur 1.

---

*Grete Dahl*

---

*Grete Dahl* er rådgiver i Statistisk sentralbyrå, Seksjon for leveårsstatistikk ([grete.dahl@ssb.no](mailto:grete.dahl@ssb.no)).

Disse datadesignene er henholdsvis tverrsnittsdata, paneldata og forløpsdata. For å se forskjellen mellom disse datadesignene, er det vist i figuren hvordan en persons sivilstand blir registrert ved en statistisk undersøkelse avhengig av type datadesign som nyttes ved datainnsamlingen. Variabelen sivilstand er tilfeldig valgt.

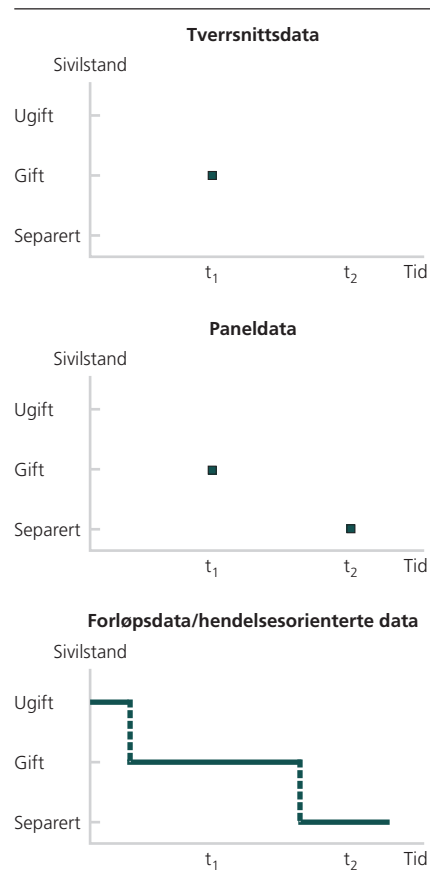
I figur 1 (øverst) er det avmerket at den personen som det er innhentet sivilstandsopplysninger for, er registrert som *gift* i tidspunkt  $t_1$ .  $t_1$  kan for eksempel være 1. oktober 1994. Men dette gir ingen informasjon om personens sivilstand før 1. oktober 1994 og heller ikke om, og eventuelt når, personen får endret sin sivilstand etter denne datoen. Ved statistiske undersøkelser blir det innhentet opplysninger ikke bare for én person. Det er vanlig å innhente data for et utvalg av personer eller for en hel populasjon, og dersom undersøkelsen er gjennomført slik at den gir sivilstandsopplysninger for et sett av personer pr. 1. oktober 1994, er dette *tverrsnittsdata* som kan nyttes til å gi statistisk informasjon om personenes sivilstand, for eksempel om fordelingen av personene på tilstandene *ugift*, *gift* eller *separert* på dette tidspunktet, (jf. den lodrette akse i figuren som angir mulige tilstander for en persons sivilstand). *Paneldata* impliserer at det er registrert opplysninger for de samme personene på to eller flere forskjellige tidspunkter, og midt i figuren vises slike registreringer for en person. Denne personen er registrert som *gift* i tidspunkt  $t_1$ , og som *separert* i tidspunkt  $t_2$ .  $t_2$  kan for eksempel være 15. mars 1998, og paneldataene gir ikke informasjon om personens sivilstand på andre tidspunkter enn 1. oktober 1994 og 15. mars 1998. Bare *forløpsdata* eller hendelsesorienterte data gir opplysninger om personenes sivilstand mer eller mindre kontinuerlig over tid. For den personen som det er avmerket opplysninger for nederst i figuren, inneholder forløpsdataene opplysninger om hvilke tidspunkter personen har skiftet sivilstand fra henholdsvis *ugift* til *gift* og fra *gift* til *separert*, samt informasjon om varigheten av giftermålet. Som det framgår av figuren, har personen vært *gift* en tid både før og etter tidspunktet  $t_1$ , og også *separert* en tid før tidspunktet  $t_2$ . (Når separasjonen eventuelt tar slutt, vises ikke i figuren.)

De aller fleste opplysningene i FD-Trygd er registrert som forløpsdata eller hendelsesorienterte data. Opplysningene består av registreringer av tilstandsendringer eller hendelser i den enkelte persons livsløp, og av målinger av varigheten mellom suksessive eller påfølgende hendelser. En hendelse for en person kan for eksempel være at en persons trygdestatus endres, (jf. figur 1 hvor vi ved et eksempel har valgt å se på endringer i en persons sivilstand). Både tidligere trygdestatus, det vil si den status som personen skifter fra, og personens nye trygdestatus er da registrert, samt den tiden personen var i sin tidligere trygdestatus før overgang til ny status. I FD-Trygd er det også registrert hendelsesorienterte data som kan nyttes dersom målet for en analyse er å fokusere på hvilke faktorer som synes å forklare varigheten i hver tilstand, eller som er bestemmende for de tilstandsendringene som skjer.

### Eksempler på bruk av dataene til forskning

Dataene i FD-Trygd byr på et mangfold av analysemuligheter. (Vi viser også til en artikkel om *Uførepensjon* i dette nummer av Samfunnspeilet. Artikke-

Figur 1. Typer datadesign



### FD-Trygd

FD-Trygd er en stor forløpsdatabase som i vel tre år har vært en vesentlig datakilde for velferdsforskningen i Norge.

FD-Trygd er utviklet i Statistisk sentralbyrå, og basen inneholder statistiske data fra og med 1992 – blant annet om personers trygdeforhold, sysselsetting, arbeidssøking, utdanning, inntekt med videre. Basen ble åpnet for drift 1. februar 1999, og til nå er det levert data fra basen både til forskningsformål og til offentlig utredningsarbeid.

Web-adressen er <http://www.ssb.no/emner/03/fd-trygd/>

len er basert på data fra FD-Trygd). Her skal vi bare gi noen få eksempler på de mulighetene som finnes:

- kartlegge og studere ulikheter i trygdekarrierer for grupper av personer
- belyse overganger fra trygd til arbeid, eller vice versa, for grupper av personer
- framskaffe forskningsbasert kunnskap om hvordan velferdssystemet fungerer for karakteristiske grupper som for eksempel innvandrere, vanskeligstilte ungdomsgrupper eller marginaliserte grupper
- identifisere persongrupper som på samme tid eller i samme tidsperiode er brukere av to eller flere velferdsordninger
- beskrive persongrupper som går ut i tidligpensjonering
- påvise hvilke persongrupper som kombinerer arbeid og trygd eller utdanning og trygd
- evaluere effektene av regelverksendringer eller strategitiltak knyttet til velferdsordningene og studere om regelverksendringene eller tiltakene har hatt tilsiktet virkning.

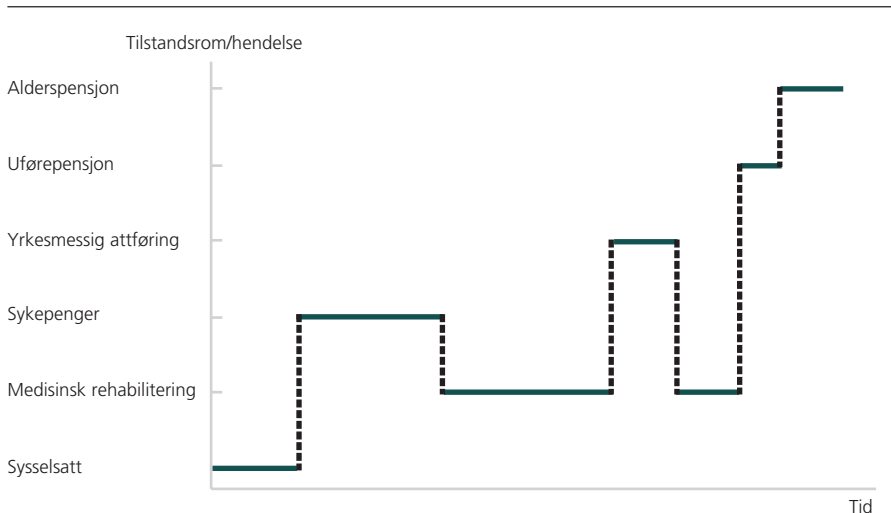
### Dataorganisering og datering av tilstandsendringer

Opplysningene i FD-Trygd er organisert som "tabeller" i en database. Det er en tabell for hvert temaområde eller hver ordning; for eksempel en tabell for uførepensjon, en for igangværende utdanning, en for overgangsstønad til enslig forsørger, etc. Opplysningene i hver tabell gir informasjon om når en person kommer inn i ordningen (tilgang), og når en person går ut av ordningen (avgang). I tillegg er det i hver tabell innført en del variabler som måler om det skjer endringer i personens forhold til ordningen og eventuelt når slike endringer skjer. Innen ordningen uførepensjon kan for eksempel en endring for en person bestå i en økning i uføregraden fra 50 prosent til 100 prosent. Dette registreres ved å endre verdien for variabelen uføregrad fra 50 til 100 for denne personen og med angivelse av når endringen skjer.

Forløpsanalyse går ut på å analysere tidsintervall mellom suksessive tilstandsendringer eller hendelser med tiden som stokastisk variabel, det vil si som en størrelse som kan anta sine mulige verdier med en viss sannsynlighet. Både tilgang til ordninger, avgang fra ordninger og endringer i personenes forhold til ordningene kan ses på som tilstandsendringer for personene.

I FD-Trygd er disse tilstandsendringene registrert ved hjelp av datovariabler. Datovariablene reflekterer hvor hyppig tilstandsendringene er målt. For en stor del av tilstandsendringene inneholder datovariablen verdien *år*, *måned*, *dag*, det vil si at tilstandsendringene er målt med *dag* som fineste måleenhet for tid. Eksempler på dette er tilstandsendringer knyttet til sykepenger, fødselspenger, medisinsk rehabilitering, arbeidssøking og sysselsetting. For andre tilstandsendringer har datovariablen verdien *år*, *måned*, det vil si at *måned* er fineste måleenhet. Dette gjelder eksempelvis for sosialhjelp, utdanning og pensjoner i folketryg-

Figur 2. Eksempel på individuelt forløp



den. Forholdsvis få tilstandsendringer i FD-Trygd er målt sjeldnere enn måned. Figur 2 viser eksempel på et individuelt forløp, hvor en person over tid er forutsatt å "oppleve" forskjellige typer tilstandsendringer eller hendelser.

### En videreutvikling av Kirut

FD-Trygd kan ses på som en videreutvikling av prosjektet Kirut (Klientstrømmer inn, rundt, ut av trygdesystemet), som på 1990-tallet har vært en viktig datakilde for trygde- og sosialpolitisk forskning. Kirut inneholder opplysninger for årene 1989-1996, og Norsk samfunnsvitenskapelig datatjeneste i Bergen står for utlevering av data fra Kirut. Kirut har vært en vesentlig inspirasjonskilde i arbeidet med utviklingen av FD-Trygd, og det er i hovedsak de samme typer data som inngår i FD-Trygd som i Kirut. De faglig sett mange likhetspunktene mellom FD-Trygd og Kirut gjør det videre mulig å "etterbruke" Kirut-dataene sammen med opplysninger fra FD-Trygd. For å muliggjøre dette er det foretatt merking av recordene i FD-Trygd for de personene som har deltatt i Kirut.

Til tross for mange likhetspunkter mellom FD-Trygd og Kirut, er det flere forhold som skiller disse to prosjektene. Kirut omfatter bare 10 prosent av befolkningen mellom 16-67 år, mens FD-Trygd omfatter hele befolkningen. Til sammenligning med Kirut er også dataene i FD-Trygd organisert på en måte som sikrer en mer dynamisk datastruktur. Dette gjelder spesielt for opplysninger som er knyttet til demografiske forhold, utdanning og sosialhjelp hvor opplysningene i FD-Trygd fanger opp de fortløpende hendelsene, mens tilsvarende informasjon i Kirut bare er registrert en gang pr. år.

### Brukertjenesten

Forløpsdatabasen er plassert i Statistisk sentralbyrå, og databasen er ikke fysisk tilgjengelig for andre etater. Dette medfører at alle datauttakene fra FD-Trygd må foretas i og av Statistisk sentralbyrå.

Utover dette konkrete arbeidet med datauttakene hører flere viktige oppaver til brukertjenestens arbeidsområde. Eksempler på dette er kontakt med brukere, orientering og forklaring overfor brukere om hvilke data som finnes i FD-Trygd, tilrettelegging og anonymisering av datafiler etc. For å utføre disse arbeidsoppgavene er det lagt opp til et samarbeid mellom Statistisk sentralbyrå og Norsk samfunnsvitenskapelig datatjeneste, slik at Norsk samfunnsvitenskapelig datatjeneste har kontakt med forskere og studenter, mens Statistisk sentralbyrå har kontakt med offentlig forvaltning og andre brukere.

FD-Trygd ble åpnet for utlevering av data i februar 1999, og det er allerede levert mye data til forskjellige typer forskningsformål og offentlig utredningsarbeid.