


IDA – en integreret database for arbejdsmarkedsforskning

Hovedrapport

**IDA –
an integrated data base for
labour market research**

Main report

DANMARKS
STATISTIK

Da. 15
024
~~2x. 10~~


**IDA –
en integreret database for arbejds-
markedsforskning
Hovedrapport**

Udgivet af Danmarks Statistik

Januar 1991

ISBN 87-501-0792-5

Pris: 105,00 kr. inkl. 22% moms

Oplag: 500

Danmarks Statistiks trykkeri, København

21 FEB. 1991

DANMARKS STATISTIK
BIBLIOTEKET

© **Danmarks Statistik**
1991

Enhver form for hel eller delvis gengivelse eller mangfoldiggørelse af denne publikation, uden skriftligt samtykke fra Danmarks Statistik, er forbudt efter gældende lov om ophavsret.

Undtaget herfra er citatretten, der giver ret til at citere, med angivelse af denne publikation som kilde, i overensstemmelse med god skik og i det omfang, som betinges af formålet.

Forord

Denne publikation er hovedrapporten om IDA-databasen og projektet, som har ligget til grund for udviklingen af databasen. Publikationen giver en redegørelse for, hvilke oplysninger der findes i IDA-databasen og for dens grundlæggende opbygning.

Målgruppen er især brugere af databasen, for hvem rapporten kan give et overblik over indhold og muligheder. Men samtidig har formålet også været at give den almindeligt interesserede et indtryk af IDA-databasen. De første 3 kapitler - herunder især kapitel 1 - henvender sig således til mere generelt interesserede, mens kapitel 4 - om klassifikation af virksomhedernes identitet over tid - i højere grad henvender sig til specielt interesserede.

I tilknytning til denne hovedrapport er der udarbejdet en række særlige rapporter, hvori der gives en nærmere redegørelse for forskellige delområder, fx vedrørende personoplysninger og vedrørende fastlæggelsen af virksomheders identitet over tid.

Endvidere foreligger der et omfattende dokumentationsmateriale. Det drejer sig bl.a. om databeskrivelser, dvs. oversigter over koder og deres betydning for hver enkelt variabel - samt tekniske beskrivelser af edb-systemet i form af procedurebeskrivelser (programmeringsoplæg) mv. Disse dokumentationsmaterialer findes på pc-disketter.

Initiativet til opbygningen af IDA-databasen blev taget af afdelingschef *Poul Jensen*, Danmarks Statistik, på baggrund af det projekt, der i 1983-84 blev gennemført i samarbejde mellem Socialforskningsinstituttet og Danmarks Statistik (jf. Per Vejrup Hansen: "Arbejdskraftbevægelser og beskæftigelse", Socialforskningsinstituttets publikation 149, København 1985).

Opbygningen af IDA-databasen er blevet finansieret af Statens Samfundsvidenskabelige Forskningsråd (SSF) og af Danmarks Statistik. SSF har bidraget med lønmidler til projektgruppen og en basisbevilling til edb-udgifter. Det er samtidig et udtryk for, at der primært har været tale om et forskningsmæssigt sigte med IDA-databasen, dvs. at skabe et

datagrundlag for arbejdsmarkedsforskning mv. Databasen er dog ikke reserveret til forskningsmæssige anvendelser alene, men kan udnyttes af alle interesserede.

IDA-databasen er blevet udviklet i årene 1988-90, dvs. over en 3-års periode. IDA-projektet er blevet udført af en projektgruppe bestående af følgende:

Ruth Emerek, cand.stat., lektor ved Aalborg Universitetscenter.

Per Vejrup Hansen, cand.polit., lektor ved Handelshøjskolen i København.

Søren Leth-Sørensen, mag.scient.soc., fuldmægtig, med orlov fra Danmarks Statistik.

Aalborg Universitetscenter, henholdsvis Handelshøjskolen i København, har velvilligt givet "delvis tjenestefritagelse" fra undervisningsforpligtelser for de to førstnævnte i projektperioden.

Projektgruppen har sideløbende med opbygningen af databasen deltaget i en række seminarer mv. med oplæg om IDA-databasen, herunder i nordisk sammenhæng og i OECD-regi. I en tidlig fase af projektet aflagde gruppen en række besøg hos forskellige forskningsmiljøer inden for arbejdsmarkedsområdet for at få et indtryk af disses ønsker og behov.

IDA-projektet er blevet fulgt af en *referencegruppe* med repræsentanter udpeget af SSF og Danmarks Statistik. Det har drejet sig om følgende personer:

Professor Reinhard Lund, sekretariatschef Jørgen Søndergaard, professor Hans E. Zeuthen (udtrådte august 1988 med udnævnelsen til rigsstatistiker), lektor Peter Maskell (fra efteråret 1988) - alle udpeget af SSF; og konsulent Otto Andersen (sekretær for referencegruppen), afdelingschef Poul Jensen (til marts 1989, formand), kontorchef Vøgg Løwe Nielsen (til marts 1989), kontorchef Søren Hostrup-Pedersen (fra april 1989) og afdelingschef Lars Thygesen (formand fra april 1989) - alle fra Danmarks Statistik.

Referencegruppens status har været rådgivende, idet det er projektgruppen, der har ansvaret for de trufne valg i forbindelse med databasens principielle udformning og indhold.

IDA-projektet er blevet udført i *Danmarks Statistik*, da grundlaget for databasen er Danmarks Statistiks person- og erhvervsregistersystemer. Forudsætningen for IDA-databasen er den opbygning af nye statistik-registre, som i løbet af en årrække er blevet foretaget i Danmarks

Statistik. Udviklingen har bl.a. muliggjort en sammenkædning eller "integration" af oplysninger om personer og virksomheder over tid, hvilket er kernen i IDA-databasen.

Projektet er blevet udført i Danmarks Statistiks kontor for Forsknings- og metodespørgsmål, som ledes af konsulent *Otto Andersen*. Otto Andersen har udarbejdet bilag 3: "Hvordan får man adgang til IDA?". Overassistent *Bente Henriksen* og assistent *Jesper S. Brønnum* har udført sekretærarbejde for projektet.

Det omfattende programmeringsarbejde er blevet udført af programmør *Søren Nisted* og programmør *Birger Hansen*. Systemplanlægningen og udarbejdelsen af programmeringsoplæg er blevet udført af projektgruppen. Fuldmægtig *Leif Sølling*, Danmarks Statistik, har bidraget med rådgivning i forbindelse med anvendelsen af Danmarks Statistiks nye dokumentations-system TIMES. Under projektet forløb har der været kontakt med mange medarbejdere i Danmarks Statistik, der har bidraget positivt med råd og vejledning.

Ved udarbejdelsen af denne rapport har referencegruppen og lektor *Morten Lassen*, AUC, bidraget med konstruktive kommentater. Lay-out af manuskriptet er foretaget af sekretær *Vibeke Bo Jensen*, AUC. Rapporten er udarbejdet af projektgruppen.

Danmarks Statistik, december 1990

Hans E. Zeuthen

/

Ruth Emerek

Per Vejrup Hansen

Søren Leth-Sørensen

Indholdsfortegnelse

Oversigt over figurer mv.	9
Oversigt over Rapporter fra IDA-projektet	10
Læsevejledning	11
1. En introduktion til IDA	13
1.1. IDA-databasens grundstruktur	14
1.2. Centrale analytiske muligheder og anvendelsesområder	16
1.3. Grundlæggende forudsætninger og begreber	18
1.4. Virksomheders identitet over tid	19
2. Oplysninger om personer, ansættelser og virksomheder	23
2.1. Typer af oplysninger	23
2.2. Oplysninger om personer	26
2.3. Oplysninger om ansættelser/beskæftigelse	31
2.4. Oplysninger om virksomheder: Arbejdssteder og firmaer	34
3. Database-organiseringen af IDA - oversigt over datasættene	39
3.1. Databasestrukturen	40
3.2. De enkelte datasæt: Population, identifikationsnøgle, variabelindhold og anvendelsesområder	43
4. Klassifikationen af arbejdssteders identitet over tid	57
4.1. To delpopulationer af arbejdssteder	57
4.2. Klassifikation af arbejdssteder med ansatte pr. november	60
4.3. Klassifikation af arbejdssteder uden ansatte pr. november	66
4.4. Endeligt nedlagte og oprettede arbejdssteder i forhold til den samlede population af arbejdssteder	68

Bilag 1: Oversigt over variable i IDA	73
Bilag 2: Oversigt over EDB-systemet til dannelse af IDA-databasen	97
Bilag 3: Hvordan får man adgang til IDA?	107

Oversigt over figurer mv.

Figur 1.	IDA-databasens grundstruktur	14
Figur 2.	Databasestruktur. Datasæt pr. år	41
Figur 3.	Den samlede population af arbejdssteder samt selvstændige, inkl. bevægelser til og fra arbejdssteder med ansatte pr. november	59
Oversigt 1.	Klassifikation af arbejdssteders identitet set fremad i tid (bevaret/nedlagt), hhv. set tilbage i tid (bevaret/oprettet). Arbejdssteder med ansatte pr. november i det pågældende år	61
Figur 4.	Forskellige typer af bevarede, nedlagte og oprettede arbejdssteder	62
Oversigt 2.	Klassifikation af arbejdssteder uden ansatte pr. november i året: Identiteten set fremad i tid (samme kategori/afgang), hhv. set tilbage i tid (samme kategori/tilgang)	67
Bilagstabel 1	(til kap. 4). Identiteten af arbejdssteder pr. november 1983 i det følgende år pr. november 1984 (identiteten set fremad i tid: bevaret/nedlagt)	70
Bilagstabel 2	(til kap. 4). Identiteten af arbejdssteder pr. november 1984 set i forhold til året før pr. november 1983 (identiteten set tilbage i tid: bevaret/oprettet)	71

I bilag 2 vedrørende edb-systemet:

Figur 5.	Data-flow-diagram over IDA-systemet	98
Figur 6.	Modul 0: Afstemning af oplysninger for personer og arbejdssteder	103
Figur 7.	Modul 1: Fastlæggelse af arbejdssteders identitet over tid	104
Figur 8.	Modul 2: Dannelse af personoplysninger	105
Figur 9.	Modul 3: Dannelse af ansættelses- og virksomhedsdatasæt	106

Oversigt over Rapporter fra IDA-projektet

Kilder til personoplysninger i IDA om uddannelse, arbejdsmarkedstilknytning og indkomst.
Arbejdsnotat nr. 27, Danmarks Statistik, januar 1991.

Kilder til personoplysninger i IDA vedrørende demografiske forhold, samliv og familie.
Arbejdsnotat nr. 28, Danmarks Statistik januar 1991.

Virksomheders identitet over tid. Begrebsdannelse og operationalisering.
Arbejdsnotat nr. 29, Danmarks Statistik, januar 1991.

Anvendelse af oplysningssedlerne i IDA.
Arbejdsnotat nr. 30, Danmarks Statistik, januar 1991.

Timeløn og arbejdsomfang i den enkelte ansættelse.
Arbejdsnotat nr. 31, Danmarks Statistik, januar 1991.

Læsevejledning

Rapporten er opbygget, så den - kapitel for kapitel - gradvis bliver mere detaljeret.

I *kapitel 1* gives der en generel orientering om indholdet i IDA-databasen og om mulige anvendelser. De grundlæggende enheder, forudsætninger og begreber omtales kort. Dette kapitel henvender sig til alle, der er interesserede i hvad IDA er, hvad IDA indeholder, og hvordan IDA er opbygget.

I *kapitel 2* omtales oplysninger og variable i IDA nærmere. Formålet med dette kapitel er at give brugeren, også den potentielle bruger, indsigt i de mangfoldige oplysninger IDA rummer om personer, ansættelser og virksomheder.

Kapitel 3 henvender sig især til de personer, der har besluttet sig for at anvende IDA i deres forsknings- eller udredningsprojekter. Her omtales, hvordan IDA er opbygget som database, og hvad de enkelte datasæt indeholder. Endelig skitseres nogle generelle anvendelsesmuligheder for de enkelte datasæt.

Det sidste kapitel i rapporten, *kapitel 4*, henvender sig til de brugere, der ønsker en nærmere beskrivelse af, hvordan arbejdsstedernes identitet over tid er klassificeret i forskellige typer af bevarede, nedlagte og oprettede arbejdssteder.

I *bilag 1* findes en oversigt over variablene i de enkelte datasæt i IDA.

Som *bilag 2* bringes en oversigt over EDB-systemet til dannelse af IDA-databasen.

I *bilag 3* omtales, hvordan man får adgang til IDA.

1. En introduktion til IDA

IDA står for "Integreret Database for Arbejdsmarkedsforskning". IDA er en "integreret" database i den forstand, at der indgår både personer og virksomheder.

Et væsentligt kendetegn ved IDA er, at data om arbejdsmarkedet foreligger med udgangspunkt i et *virksomhedsniveau*. Det betyder, at den enkelte beskæftigede person sammenkædes med den enkelte virksomhed. En gang om året, ultimo november, fastlægges hvert enkelt ansættelsesforhold, dvs. på hvilken virksomhed den enkelte person er ansat.

Den grundlæggende virksomhedsenhed i IDA-databasen er det enkelte arbejdssted, der kan være en del af et firma med flere arbejdssteder¹.

Et andet kendetegn ved IDA er, at den er *longitudinal*, dvs. at databasen indeholder data over tid. Konkret er der tale om årlige data fra 1980 og frem til det seneste år, for hvilket der foreligger oplysninger i alle de tilgrundliggende statistikregistre². Personer og virksomheder kan således følges over tid, fra år til år.

Endelig er det et kendetegn, at IDA er *totaldækkende*, dvs. at IDA omfatter den samlede befolkning i Danmark, herunder den totale arbejdsstyrke, og samtlige virksomheder i såvel den offentlige som den private sektor.

Det specifikke eller unikke ved IDA er at alle tre kendetegn optræder samtidig. IDA er således en *totaldækkende, longitudinal database over personer og virksomheder*. Den totale population af henholdsvis personer og virksomheder sammenkædes og følges over en årrække fra 1980.

¹ Denne definition er begrundet i datagrundlaget for IDA, jf. afsnit 1.3.

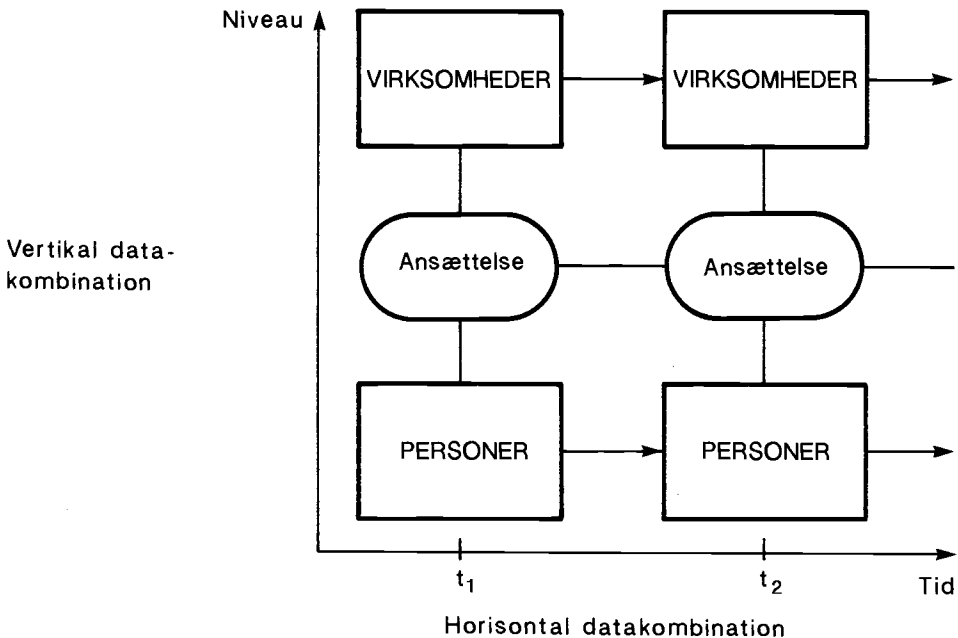
² Databasen omfatter ved etableringen oplysninger fra årene 1980-1987. Ved fremkomsten af data for efterfølgende år (1988 osv.) kan der successivt foretages en opdatering af databasen.

Datagrundlaget for IDA udgøres af statistikregistre i Danmarks Statistik. Disse statistikregistre bygger på centrale administrative registre såsom det Centrale Personregister, skattevæsenets registre, registeret vedrørende arbejdsløshed (CRAM), mv. IDA-databasens indhold af data er således bestemt af dataindholdet i statistikregistre, som igen er bestemt af indholdet i de administrative registre. Derudover indeholder IDA en omfattende bearbejdning og kombination af disse registerdata.

1.1. IDA-databasens grundstruktur

En illustration af IDA-databasens grundlæggende struktur fremgår af *figur 1*.

Figur 1. IDA-databasens grundstruktur



Af illustrationen fremgår, at der er tale om datakombination i *to dimensioner*, både vertikalt (på et givet tidspunkt) og horisontalt (over tid). Det er et udtryk for, at der ikke blot indgår én enhed, således som det er tilfældet i typiske longitudinale datasæt - men at der er tale om to enheder, personer og virksomheder, samt deres indbyrdes relation, ansættelser.

Grundstrukturen genfindes i database-organiseringen af IDA, dvs. den måde de forskellige datasæt i IDA er organiseret på, jf. kapitel 3. Heri er der tre hovedkategorier af datasæt, nemlig for henholdsvis personer, ansættelser og virksomheder.

I datasættet vedrørende personer findes de oplysninger, der karakteriserer personerne som personer, dvs. køn, alder, uddannelse mv. På tilsvarende måde indeholder datasæt vedrørende virksomheder de baggrundsoplysninger, der karakteriserer de enkelte virksomheder.

Oplysninger om en persons ansættelsesforhold - stilling, løn mv. - er placeret i datasættet ansættelser. Det understreger, at et ansættelsesforhold er et resultat af et samspil mellem en person- og en virksomhedsside - og ikke kun en egenskab ved personen som person. For "ansættelserne" findes en dobbelt identifikationsnøgle for både personen (den ansatte) og virksomheden, nemlig henholdsvis personnummeret og et virksomhedsnummer.

Indholdet i de tre datasæt fremgår af følgende *oversigt over eksempler på oplysninger*:

Personer	Ansættelser	Virksomheder
Nøgle: Personnr.	Nøgle: Personnr./ virksomhedsnr.	Nøgle: Virksomhedsnr.
- Køn, alder mv.	- Stilling, heltid/deltid	- Oprettelsesår
- Familie og samliv	- Timeløn	- Branche, beliggenhed mv.
- Uddannelse	- Anciennitet	- Ansatte og lønniveau
- Beskæftigelse og erhvervs erfaring	- Ændring i ansættelse: Tilgang/afgang	- Identitet over tid: Bevaret, nedlagt, oprettet
- Arbejdsløshed		
- Indkomst		

Kapitel 2 indeholder en nærmere omtale af de forskellige typer af oplysninger.

Udover de oplysninger, der allerede foreligger i databasen, kan der i forbindelse med konkrete forskningsprojekter ad hoc tilknyttes yderligere oplysninger fra andre registre. Endvidere kan der foretages yderligere aggregeringer eller summeringer over personer (ansættelser) pr. virksomhed eller pr. virksomhedskategori. Blandt andet kan der ved hjælp af personoplysningerne dannes supplerende oplysninger til karakteristik af den enkelte virksomhed som fx antallet af lærlinge på virksomheden.

1.2. Centrale analytiske muligheder og anvendelsesområder

De særlige analytiske muligheder ved IDA-databasen udspringer af, at der indgår to enheder, personer og virksomheder, samt deres indbyrdes relation, ansættelserne.

Det betyder, at man i analyser af udbudssiden på arbejdsmarkedet, dvs. af personers placering og bevægelser, kan inddrage data for de virksomheder, hvor personerne er ansat. Fx kan betydningen af ændringer i det totale antal ansatte for beskæftigelseschancer/afskedigelsesrisiko blandt forskellige persongrupper analyseres.

Tilsvarende kan man ved analyser af efterspørgselssiden på arbejdsmarkedet, fx ved analyser af virksomheders forløb, specielt af oprettelser og nedlæggelser, inddrage oplysninger om de involverede personer (hvilke kategorier af personer er ansat, hvem opretter nye virksomheder, osv.).

For arbejdsmarkedsforskningen betyder det forhold, at der foreligger et totaldækkende grundlag for opgørelser af arbejdskraftens mobilitet med udgangspunkt i et virksomhedsniveau, at man kan inddrage bevægelser af personer mellem virksomheder - og ikke blot bevægelser mellem hovedtilstandene beskæftigelse, ledighed og udenfor arbejdsstyrken. Denne egenskab indebærer, at man kan løse nogle målingsproblemer i forbindelse med de såkaldte "bruttostrømsanalyser".

Der kan peges på følgende generelle anvendelsesområder for IDA-databasen:

Personbevægelser på arbejdsmarkedet og ansættelsesforhold

På grundlag af databasen kan der foretages analyser af en række former for mobilitet: Arbejdspladsmobilitet (afgang/tilgang på virksomhedsniveau), erhvervs mobilitet, arbejdsstyrkemobilitet (mellem beskæftigelse/ledighed/udenfor arbejdsstyrken), mv.

Ansættelsesvilkår kan undersøges med hensyn til stilling, arbejdstid og lønniveau.

Under en teoretisk synsvinkel kan man pege på muligheden af at afprøve nyere arbejdsmarkedsteorier, fx segmenteringsteorier.

Desuden kan man undersøge betydningen af konjunkturudviklingen for bevægelserne på arbejdsmarkedet, både med hensyn til mobilitetens omfang og dens "retning" (eksempelvis: Vender personer uddannet i bygge- og anlægsbranchen tilbage til denne branche under en opgangskonjunktur?).

Der kan endvidere på grundlag af IDA etableres fremskrivnings- og konsekvensberegningmodeller for udbudssiden på baggrund af skøn over overgange mellem forskellige

tilstande i forhold til arbejdsmarkedet - afhængigt af fx køn, uddannelse, virksomhedsstørrelse mv.

Beskæftigelses- og lønudvikling på virksomhedsniveau

Hidtil har vi kun haft mulighed for at belyse beskæftigelses- og lønforhold på branche- eller makroniveau. Med et virksomhedsniveau kan vi fx beskrive, hvor store udsving der er i beskæftigelsesomfanget fra år til år på den enkelte virksomhed. Eller vi kan undersøge, hvordan sammensætningen af de ansatte på fx uddannelseskategorier ændres over tid i givne typer af enkeltvirksomheder. Sådanne data kan have betydning for arbejdskraftprognoser og for studier af substitution. Herunder kan der også peges på undersøgelser af, hvilke forskelle der er mellem nye og gamle virksomheder med hensyn til beskæftigelsesvækst og sammensætningen af de ansatte. Endelig kan vi fx undersøge forskelle i lønniveau og lønudvikling mellem enkeltvirksomheder, hvilket kan give indblik i løndannelsesprocesser.

På samme måde kan strukturelle træk og tendenser i forskellige regioner analyseres. Hvilke konsekvenser har det på længere sigt, at der fx er en forholdsvis stor andel ikke-faglærte blandt de beskæftigede i en region?

Virksomheders livsforløb

Med hensyn til analyser af virksomheders livsforløb er der tale om et helt nyt datagrundlag, specielt når det tages i betragtning, at der også indgår oplysninger om de ansatte personer og ejere.

Først og fremmest kan det beskrives, hvor stor mobiliteten af virksomheder er, dvs. hvor stort det reelle omfang af nedlæggelser og nyoprettelser af virksomheder er. Eller det kan undersøges, om der findes bestemte størrelsesmæssige barrierer for virksomheders vækst.

Med hensyn til nye virksomheder kan det (for enkeltmandsejede virksomheder) kortlægges, hvem der opretter dem, dvs. hvilke personkategorier.

Endvidere kan man belyse i hvor høj grad - og på hvilken måde - der er tale om "spin-offs" fra bestående virksomheder. Hvilke personkategorier ansættes på nye virksomheder, og hvor kommer de typisk fra (andre virksomheder, fra ledighed eller fra uddannelses-systemet)?

1.3. Grundlæggende forudsætninger og begreber

De grundlæggende begreber og enheder i IDA-databasen er bestemt af det statistiske grundlag for sammenkædningen af personer og virksomheder på et givet tidspunkt. Dette grundlag udgøres af *nye registerdata* vedrørende arbejdssteder, som foreligger fra og med året 1980. Det drejer sig om et register hos skattevæsenet over "*oplysningssedler*" for *samlige ansættelser*³.

En oplysningsseddel er en virksomheds (arbejdsgivers) årlige oplysning til skattevæsenet om den enkelte ansatte persons løn, indeholdte kildeskat, ATP-beløb og ansættelsesperiode mv.

Heraf fremgår et arbejdsgivernummer og den ansattes personnummer. I tilfælde af, at en arbejdsgiver har flere lokale arbejdssteder, indberettes der til Danmarks Statistik oplysning om, på hvilket arbejdssted den enkelte person er ansat. Det foregår via en "arbejdsstedskode".

Virksomhedsenheden

Ovennævnte registermæssige udgangspunkt betyder, at den grundlæggende virksomhedsenhed i IDA-databasen er det lokale arbejdssted, som eventuelt er en del af et firma med flere arbejdssteder. Et arbejdssted er - jf. ovenfor - en lokal enhed med ansatte i løbet af et år, for hvilke der er indeholdt kildeskat. Firmaenheden er identisk med *kildeskatteenheden*, også kaldet "SE-enheden".

Kildeskatteenheden er ikke nødvendigvis sammenfaldende med moms-enheden, og den kan desuden afvige fra den virksomhedsenhed, der anvendes i industristatistikken (SAF-enheden). En konsekvens af disse forskelle er, at data fra momsstatistikken og industristatistikken ikke umiddelbart kan inddrages i IDA-databasen. På dette punkt er der dog udsigt til, at der på et tidspunkt vil blive etableret en fælles virksomhedsenhed.

I databasen indgår også firmaenheden. Derimod er det ikke muligt at registrere en eventuel koncern, dvs. den større enhed, som omfatter flere firmaer.

³ For så vidt angår en diskussion af anvendelsen af disse oplysningssedler, som er centrale for hele IDA's opbygning, findes en selvstændigt rapport.: "Anvendelse af oplysningssedlerne i IDA", Arbejdsnotat nr. 30.

Tidsdimensionen

Sammenkædningen af personer og arbejdssteder i det enkelte år via oplysningsseddelregistret mv. refererer til ét bestemt tidspunkt, nemlig pr. *ultimo november måned*. Dette tidspunkt er valgt således, at det ligger nogenlunde tæt på slutningen af året, samtidig med, at registreringen ikke bliver forstyrret af sæsonmæssige forhold (lukning af virksomheder omkring Julen). Det betyder, at opgørelser af ansættelser i den enkelte virksomhed refererer til situationen pr. ultimo november.

Baggrundsoplysninger vedrørende de enkelte personer refererer derimod hovedsageligt til ultimo året. En undtagelse herfra er oplysninger om personers uddannelsesmæssige status, der refererer til oktober måned i det pågældende år. En nærmere omtale af de øvrige oplysningstyper findes i kapitel 2.

1.4. Virksomheders identitet over tid

En grundlæggende forudsætning for IDA-databasen er, at den enkelte virksomhed kan følges over tid. En sådan bestemmelse af virksomheders identitet over tid er ikke fastlagt på forhånd, således som det er tilfældet med personer, hvor det én-éntydige personnummer (CPR-nummeret) definerer samme person på forskellige tidspunkter. IDA-databasen bygger på en specifik begrebsdannelse vedrørende identiteten af virksomheder over tid, og en særlig operationalisering heraf i forbindelse med den konkrete løsning.

For det første foreligger der et *grundlæggende begrebsmæssigt problem*, når man vil fastlægge, hvad der er "samme" virksomhed på forskellige tidspunkter. Det skyldes, at en virksomhed er en abstrakt og kompleks enhed, som er kendetegnet ved flere forskellige karakteristika. Det drejer sig om sådanne egenskaber som ejerskab, geografisk beliggenhed, branche og arbejdsstyrke (som indgår i IDA), samt om fx produktkategori, produktionsteknologi, arten af produktmarkedet osv. Det principielle spørgsmål er da, hvilke og hvor store ændringer i én eller flere af disse karakteristika der bør bevirke, at en virksomheds identitet ændres (dvs. at den nedlægges eller må betragtes som en ny enhed).

For det andet er der tale om et *"teknisk" problem*, som består i, at identifikationsnummeret for den enkelte virksomhed (arbejdssted) kan skifte fra et tidspunkt til et andet. I tilfælde af et ejerskifte vil der således fremkomme et skift i arbejdsgivernummeret, og arbejdsstedskoden kan blive ændret af administrative grunde. Hvis virksomheden iøvrigt er den samme, ville det da være fejlagtigt at betragte den som henholdsvis nedlagt og

nyoprettet, idet det korrekte ville være at betragte virksomheden som bevaret (den samme) på trods af et eventuelt ejerskifte. Den operationelle løsning heraf i IDA-databasen kan foretages i kraft af, at der indgår data om de ansatte personer, således at "samme arbejdsstyrke" kan anvendes som kriterium for at arbejdsstedet er det samme.

Den *begrebsmæssige definition* af en virksomhed over tid - hvornår er der tale om "samme" virksomhed på forskellige tidspunkter - er fastlagt med *udgangspunkt i det analytiske formål med IDA-databasen*. Formålet er for det første at kunne foretage analyser på grundlag af arbejdskraftens bevægelser ind i og ud af virksomheder, dvs. af arbejdspladsmobiliteten, og for det andet at kunne analysere virksomheders livsforløb. Begge disse formål gør det nødvendigt, at der er tale om "samme" virksomhed på trods af ejerskifte, og at "samme" virksomhed skal kunne flytte til en anden lokalitet. Sidstnævnte krav om at muliggøre flytning til en anden beliggenhed indebærer, at den geografiske eller fysiske placering tillægges en underordnet betydning for identiteten af en virksomhed.

Dette udgangspunkt medfører, at det afgørende kriterium for, at virksomheden er den samme i to på hinanden følgende år, består i at enten ejeren eller arbejdsstyrken er den samme. Det betyder, at der er tale om et *personorienteret virksomhedsbegreb*. Det kommer til udtryk i det "regelsæt" vedrørende samme arbejdssted over tid, som er fastlagt. Det er kun for arbejdsstederne, at identiteten over tid bliver bestemt. Ifølge regelsættet foreligger samme arbejdssted, hvis én af følgende regler er opfyldt⁴:

1. Samme *ejer* og samme *branche*.
2. Samme *ejer* og samme *arbejdsstyrke*.
3. Samme *arbejdsstyrke* og samme *adresse eller branche*.

Det er et fælles træk, at der i alle regler indgår krav om enten samme ejer eller samme arbejdsstyrke. Kriteriet om samme ejer eller samme arbejdsstyrke står dog ikke alene, idet det også skal sandsynliggøres, at der er tale om samme type af erhvervsaktivitet (jf. fx kravet om samme branche i regel 1). Sammenfattende gælder det, at samme virksom-

⁴ Med samme "*ejer*" menes samme identifikationsnummer for arbejdsstedet, dvs. samme arbejdsgivernummer og samme arbejdsstedskode. Med samme *branche* menes samme ISIC-nummer på 5-ciffer niveau. Samme *adresse* defineres som samme kommunekode og vejkode. Samme *arbejdsstyrke* specificeres forskelligt i regel 2 og regel 3. I regel 2 kræves, at antal gangangere skal udgøre mindst 30 pct. af det samlede antal ansatte i det ene af årene - mens det i regel 3, hvor der er tale om et ejerskifte, kræves at gangangerne udgør mindst 30 pct. af de ansatte i begge år (dvs. et mere restriktivt krav).

hed foreligger, når en ejer eller de ansatte personer fortsætter den samme type af erhvervsaktivitet. Det primære er, at det er personerne i en virksomhed, der konstituerer identiteten over tid - og ikke fx den fysiske lokalitet, jf. ovenfor. Det betyder eksempelvis, at en given virksomhed (defineret ved ejeren og/eller de ansatte) kan overtage en anden virksomheds tidligere lokaler, og alligevel være bevaret. Hvis virksomheden blev opfattet som værende identisk med den tidligere virksomhed på den pågældende lokalitet, ville konsekvensen være, at personaleomsætningen kunne blive på 100 procent, og at man mistede forbindelsen til virksomhedens "egen" fortid.

En nærmere redegørelse for denne begrebsdannelse og for den operationelle løsning findes i rapporten: "Virksomheders identitet over tid", Arbejdsnotat nr. 29. Heri indgår også en empirisk belysning af de forskellige kriterier i regelsættet, herunder med hensyn til forekomsten af et eventuelt tautologi-problem⁵.

Selve klassifikationen af ændringstyper fremgår af kapitel 4. I klassifikationen indgår forskellige typer af bevarede, nedlagte eller oprettede arbejdssteder, idet der bl.a. foretages en registrering af opdelinger eller sammenlægninger af arbejdssteder. Det sker via oplysninger om personkredsen af ansatte på forskellige tidspunkter.

⁵ Det mulige tautologi- eller målingsproblem beror på, at omfanget af arbejdskraftens bevægelser har indflydelse på fastlæggelsen af, hvilke arbejdssteder der er bevarede - samtidig med, at disse arbejdssteder udgør grundlaget for opgørelser af arbejdskraftens bevægelser. Det kunne betyde, at arbejdssteder med en meget stor personaleomsætning ikke kommer til at indgå heri. Resultatet af undersøgelsen er, at problemets faktiske omfang må vurderes som værende ubetydeligt.

2. Oplysninger om personer, ansættelser og virksomheder

Formålet med dette kapitel er dels at præsentere de forskellige typer af oplysninger i IDA, dels at give en generel oversigt over de oplysninger, der indgår om de tre enheder: personer, ansættelser og virksomheder. Gennemgangen af oplysningerne vedrører de mere generelle forhold, og der omtales især de aspekter, der er specielle for IDA-databasen. I bilag 1 findes en samlet oversigt over variablene i IDA; her er variablene ordnet svarende til den måde, som data er struktureret på i IDA, og som fremgår af det følgende kapitel (dvs. pr. datasæt). Dokumentation af indholdet i de enkelte variable er på grund af omfanget ikke medtaget i denne publikation. Disse databeskrivelser findes på diskette og kan i givet fald også fås udskrevet på papir.

2.1. Typer af oplysninger

Hovedparten af oplysningerne i IDA refererer til et bestemt tidspunkt i året - dvs. at der i IDA primært er tale om *årlige statusoplysninger*. Det indebærer, at hændelser og bevægelser som hovedregel registreres som *ændringer i status* fra det ene år til det næste. Hvor det har været muligt og er skønnet hensigtsmæssigt, er der dog også inddraget datoen for den seneste hændelse eller bevægelse i året. Det drejer sig om de demografiske hændelser: fødsler, dødsfald, vielser, skilsmisser og flytninger¹.

På nogle områder indgår dog oplysninger om ændringer eller bevægelser i løbet af året. Dette gælder bl.a. i forbindelse med personernes arbejdsmarkedstilknytning. Bevægelserne/ændringerne i tilknytningen er bestemt på flere måder med hensyn til omfang og

¹ Det fremgår dog ikke, om en person fx har skiftet adresse mere end en gang i løbet af året, men det giver alligevel en mulighed for at sætte arbejdsplads- og geografisk mobilitet i relation til hinanden.

intensitet - fx i form af hvor mange ansættelser, der har været i løbet af året, antallet af ledighedsperioder, og hvilken ledighedsgrad der har været i de pågældende uger.

Endelig findes en række oplysninger, der knytter sig til hele året under ét (om fx indkomsten i året, den gennemsnitlige ledighed).

På trods af, at IDA er en database, hvor bevægelserne på arbejdsmarkedet først og fremmest registreres årligt, indgår altså også oplysninger om variation og fluktuation i løbet af året.

Generelt kan oplysningerne for de tre typer af enheder - personer, ansættelser og virksomheder - i IDA opdeles i følgende kategorier:

Faste eller relativt faste baggrundsuplysninger:

- faste historiske baggrundsuplysninger (fx oprettelsesår for arbejdsstederne og fødested for personer).
- baggrundsuplysninger, der kan ændres (fx statsborgerskab, adresse).

Statusoplysninger og oplysninger om ændringer i status:

- årlige statusoplysninger (fx stilling, forsikringskategori i A-kasse).
Hovedparten af oplysningerne er hentet fra Danmarks Statistiks registre og resten er dannet på grundlag heraf. For virksomhedernes vedkommende også via aggregering over de ansatte personer.
- oplysninger om ændringer fra en én statusopgørelse til den næste (fx afslutning af erhvervsuddannelse).

Årsoplysninger samt oplysninger om og indikatorer for hændelser:

- årsoplysninger i form af års-totaler eller gennemsnit over året (fx personers årsledighed, den gennemsnitlige timeløn i ansættelsen pr. november i det pågældende år).
- oplysninger om bevægelser i løbet af året (fx antal ledighedsperioder, antal ansættelser).
- oplysninger om datoen for visse hændelser i løbet af året (fx dato for seneste skift af adresse).

Ved anvendelse af IDA skal man være opmærksom på, hvilken type af opgørelse, der ligger til grund for oplysningen, dvs. om det fx er status- eller årsoplysninger.

Endvidere bør man være opmærksom på, at statutidspunktet ikke er det samme for alle oplysninger. Statusoplysningerne for personer refererer overvejende til ultimo året (31. december). Undtagelser herfra er bl.a. uddannelsesoplysningerne, der vedrører personernes uddannelsesmæssige status primo oktober, og personernes ansættelsesoplysninger, der knytter sig til situationen ultimo november. Samtlige statusoplysninger referer dog således til det samme kvartal.

Ved analyser af sammenhængen mellem ændringer i demografiske forhold, uddannelsesmæssig status og ansættelsesforhold er det vigtigt at tage højde for disse forskellige tidsreferencer. Her kan der drages nytte af, at datoen for de demografiske hændelser foreligger.

2.2. Oplysninger om personer

Personoplysninger er opdelt i følgende hovedgrupper: nøgle, baggrundoplysninger, familie og samliv, uddannelse, beskæftigelse, arbejdsløshed og indkomst.

Nøgle

- personnummer

Som nøgle for personerne anvendes personnummeret.

Baggrundoplysninger

- herunder:
køn, alder, statsborgerskab, bopælsadresse samt ind- og udvandring

For personer, der er ind- og/eller udvandret i løbet af året findes oplysninger om datoen for seneste ind- og/eller udvandring i året, samt oplysning om hvilket land personen er indvandret fra eller udvandret til².

For alle personer er anført bopælsadressen ultimo året. Derudover er angivet datoen for, hvornår personen (evt.) er flyttet til denne adresse.

Familie og samliv³

- opvækst-familie (mors og fars personnummer)
- egen familie (civilstand, samliv, ægtefælles personnummer)
- B-familie (type, antal børn)
- IDA-familie (type, antal børn)
- husstand (type)

For hovedparten af personerne født efter 1955 findes mors og fars personnummer. For gifte personer findes ligeledes ægtefællens personnummer, og tilsvarende gør sig gælden-

² Disse oplysninger findes også i form af en sammensat variabel, der angiver personens bevægelser i løbet af året. Denne variabel omtales nærmere i "Kilder til personoplysningerne i IDA vedr. demografiske forhold, samliv og familie", Arbejdsnotat nr. 28.

³ Oplysningerne vedr. familie og samliv er mere udførligt beskrevet i Arbejdsnotat nr. 28.

de for samboer/samlever. Sammen med personens civilstand ved årsskiftet findes datoen for sidste ændring af civilstand.

I IDA indgår som supplement til den juridiske civilstand en oplysning om det faktiske samlivsforhold⁴. Hensigten hermed er at give et mere nuanceret billede af personernes evt. forsørgermæssige forpligtigelser over for andre.

Herunder er der dannet en speciel "IDA-familie", der især med hensyn til samlivsbegrebet afviger fra den såkaldte "B-familie", der i perioden 1980-90 overvejende har været anvendt i Danmarks Statistik. Af hensyn til sammenligneligheden med tidligere undersøgelser er oplysninger om B-familien dog også inddraget.

For den enkelte person findes familietilhørsforhold i form af et familienummer (for både IDA-familien og B-familien). I både IDA-familien og i B-familien er familienummeret fastlagt ud fra primært kvindens personnummer.

I IDA er familieoplysningerne knyttet til de enkelte personer. Det betyder, at såvel oplysninger om personens status i familien som oplysninger om familiens sammensætning med hensyn til antal voksne og antallet af børn i forskellige aldersgrupper findes for samtlige personer i familien.

Udover de årlige statusoplysninger findes en oplysning, om hvorvidt der er født et barn i familien i løbet af året og datoen herfor, idet det yngste barns fødselsdato indgår.

Personer med samme bopælsadresse udgør en husstand. Disse er i IDA-basen klassificeret efter type samt antallet af familier og personer i husstanden.

⁴ I IDA indgår følgende former for samliv (for voksne):

- enlige (herunder ikke-samboende ægtepar)
- samlevende med ægtefælle
- samlevende ikke-gifte med fælles hjemmeboende børn
- samboende ikke-gifte (for to personer af forskelligt køn med samme bopælsadresse og med en aldersforskel på 15 år eller derunder, dog ikke søskende).

Uddannelse

- igangværende uddannelse (adresse for uddannelsesinstitution)
- afsluttet almen uddannelse
- afsluttet erhvervsuddannelse
- deltagelse i AMU-kurser

Til belysning af uddannelsesforhold findes først og fremmest oplysninger om igangværende uddannelse, og om afsluttet skole- og erhvervsuddannelse på et meget detaljeret niveau med hensyn til retning og niveau⁵. Oplysninger om afsluttet/ophørt uddannelse findes for personer, der er født efter 1920⁶.

Med henblik på at kunne differentiere gruppen af personer uden erhvervsuddannelse, er uddannelsesoplysningerne endvidere suppleret med oplysninger om deltagelse i AMU-kurser. Heri indgår branche/type for det seneste kursus i året og antallet af kurser gennemført i året. Disse oplysninger findes fra og med 1984.

Beskæftigelse

- status i november
- andre ansættelser i november (antal, lønsom)
- ansættelser i året
- årsoplysninger vedr. beskæftigelse (DFK-kode, hel-/deltid)
- historiske oplysninger (erhvervs erfaring)

Oplysningerne om status i forhold til arbejdsmarkedet (i beskæftigelse, arbejdsløs, uden for arbejdsstyrken) referer til situationen i den sidste uge i november⁷. Oplysninger om beskæftigelsen i november måned fremgår af oplysninger vedr. ansættelser/beskæftigelse (jf. afsnit 2.3.).

⁵ For en nærmere omtale af uddannelsesoplysningerne henvises til "Kilder til personoplysninger i IDA om uddannelse, arbejdsmarkedstilknytning og indkomst", Arbejdsnotat nr. 27.

⁶ Oplysningerne for personer, der har afsluttet deres uddannelse før 1970, bygger på oplysninger fra Folke- og boligtællingen i 1970, og herfra er medtaget oplysninger for personer født efter 1920. Oplysningerne for de ældre generationers vedkommende blev anset for at være for usikre.

⁷ For en nærmere redegørelse for opdelingen af befolkningen efter deres tilknytning til arbejdsmarkedet henvises til Arbejdsnotat nr. 27.

For personer med en anden ansættelse i løbet af året, hvor ansættelsesperioden således ikke overlapper referenceugen i slutningen af november, indgår endvidere oplysninger om den væsentligste ansættelse. Det drejer sig om oplysninger om identifikationsnummer for arbejdsstedet, løn, ATP og ansættelsesperiode. Antallet af øvrige ansættelser af denne type er ligeledes opgjort sammen med den samlede lønsum⁸.

Endelig indgår et skøn over erhvervsancienniteten som lønmodtager⁹ og om start og ophør på arbejdsmarkedet.

Arbejdsløshed

- statusoplysninger vedr. ledighed (A-kasse, forsikringskategori)
- ledighedsgrader (i året, kvartalsvis, i ledighedsperioder)
- ledighedsperioder
- ledighedsskift
- antal ledighedsuger (uger med delvis/fuld ledighed, opdelt på årsag)
- historiske oplysninger (sum af ledighedsgrader fra 1980)

IDA indeholder en række detaljerede oplysninger om ledighed. Der indgår for det første oplysninger, der umiddelbart har foreligget i Danmarks Statistik. Det gælder bl.a. oplysning om A-kasse og arbejdsløshedsforsikringskategori ultimo året samt den gennemsnitlige ledighedsgrad i året.

For det andet er der til brug i IDA-databasen dannet en række særlige variable, der kan give et mere differentieret billede af arbejdsløshedens omfang og fordeling. Bl.a. er antallet af uger med fuld eller delvis ledighed opgjort efter ledighedsårsag. Desuden er antallet af ledighedsperioder opgjort, og skift mellem perioder uden ledighed og perioder med hhv. delvis og fuld ledighed er beregnet. Hvor det er relevant, er oplysninger opgjort under hensyntagen til problemer omkring den såkaldte "AF-ferie"¹⁰.

⁸ For en nærmere redegørelse om disse oplysninger henvises til Arbejdsnotat nr. 27.

⁹ Skønnet er lavet på baggrund af oplysninger om det samlede indbetalte ATP-beløb pr. person frem til 1987 og de årlige ATP-indbetalinger fra 1980-1987. Se nærmere i Arbejdsnotat nr. 27, bilag 2.

¹⁰ For en detaljeret oversigt over ledighedsoplysningerne henvises til den samlede oversigt i bilag 1. I Arbejdsnotat nr. 27, findes en diskussion af ledighedsoplysningerne.

Indkomst

- samlet indkomst (brutto- og skattepligtig indkomst)
 - primærindkomst (løn, overskud af virksomhed)
 - overførselsindkomster (sygedagpenge, arbejdsløshedsunderstøttelse, bistandshjælp)
 - formueoplysninger
-

Oplysning om overskud vedr. ejerbolig indgår som en indikator for, om personen er indehaver af en ejerbolig.

2.3. Oplysninger om ansættelser/beskæftigelse

Formålet med oplysningerne om ansættelser/beskæftigelse er at give en karakteristik af personernes ansættelse på et arbejdssted og at registrere eventuelle ændringer i ansættelsesrelationen (tilgang til eller afgang fra et arbejdssted, dvs. arbejdspladsmobilitet).

Omfanget af disse oplysninger afhænger af, hvilken kategori af ansatte/beskæftigede, der er tale om. Det største antal oplysninger findes for hovedbeskæftigede lønmodtagere og arbejdsgivere - mens oplysningerne for bibeskæftigede lønmodtagere, selvstændige og medhjælpende ægtefæller er færre¹¹.

Nøgle

- personnummer og arbejdsstedets løbenummer

Ansættelsens karakter

- herunder:
stillingsgruppe, arbejdsomfang, tilknytning til arbejdssted, jobtype, ansættelsesår, pendling

Som de vigtigste oplysninger til karakteristik af ansættelsesforholdet i hoved- og bi-beskæftigelsen er anført statusoplysninger om stillingsgruppe og arbejdsomfang. Som supplement til stillingsoplysningerne indgår endvidere heri oplysninger om A-kasse, uddannelseskategori, deltagelse i AMU-kurser og DFK-koden (fagkode).

Endvidere er der for hovedbeskæftigede lønmodtagere udviklet en variabel, der angiver graden af tilknytning til virksomheden i løbet af året. Variablen angiver om personen er kontinuert eller periodevis tilknyttet virksomheden, herunder med hvor mange arbejdstimer i gennemsnit hele året eller i en del af året¹².

¹¹ Dette afspejler sig også i, at der er et datasæt for hver af disse grupper/populationer (jf. næste kapitel). For en fuldstændig oversigt over oplysningerne vedrørende de forskellige grupper må henvises til bilag 1, 2A-G, - i dette afsnit gives alene nogle hovedtræk.

¹² For en nærmere redegørelse vedr. denne tilknytningsvariabel se "Timeløn og arbejdsomfang i den enkelte ansættelse", Arbejdsnotat nr. 31.

Endelig er typen af ansættelse eller job for de hoved- og bibeskæftigede lønmodtageres med hensyn til ugentligt arbejdstid beskrevet¹³.

For hovedbeskæftigede lønmodtagere indgår ansættelsesåret på virksomheden. Dette gælder dog kun for personer, der er ansat efter 1980. For personer, der er ansat tidligere er ansættelsesåret anført som 1980. Ud fra ansættelsesåret kan arbejdsstedsancienniteten beregnes.

Afstanden i kilometer mellem arbejdsstedskommune og bopælskommune er angivet for både hoved- og bibeskæftigede.

Løn og ATP

- herunder:
løn og ATP i ansættelse, timeløn

Ud fra oplysningssedlerne er løn og ATP angivet for såvel hoved- som bibeskæftigede. Desuden er den gennemsnitlige timeløn i ansættelsen estimeret. Estimatet er lavet på grundlag af lønnen i året i ansættelsen og et skøn over det samlede timetal baseret på ATP-indbetalingen og ansættelsesperiodens længde, suppleret med oplysninger om evt. ledighed og sygdom samt et skøn over timetallet i anden beskæftigelse. Kvaliteten af skønnet over timelønnen er angivet i form af den relative usikkerhed på timelønnen bestemt som forskellen mellem den højeste og laveste timeløn, der kan estimeres på grundlag af det indbetalte ATP-beløb, i relation til den estimerede timeløn¹⁴.

Ændringer i ansættelsen

- arbejdspladsmobilitet

Ændringer i ansættelsesforholdet beskrives - dels retrospektivt, dvs. om personen er uændret ansat eller tilgået til arbejdsstedet i forhold til november året før - dels prospek-

¹³ Klassifikationen har følgende kategorier:

- fuldtidsjob (over 30 timer)
- deltidsjob (15-30 timer)
- bijob (under 15 timer).

¹⁴ For en nærmere beskrivelse af bestemmelsen af timelønnen henvises til Arbejdsnotat nr. 31.

tivt, dvs. om personen er uændret ansat eller afgang fra arbejdsstedet i løbet af den følgende årsperiode. Disse skift kan ske enten mellem arbejdssteder inden for et firma eller mellem forskellige firmaer.

For ansatte i den kommunale sektor vil et skift af arbejdssted inden for en given kommune også udgøre et skift af "firma" (arbejdsgiverenhed). Dette skyldes, at hvert fysisk adskilt arbejdssted i en kommune betragtes som et firma. Tilsvarende vil ansatte i staten, der skifter ansættelse fra én styrelse til en anden, også skifte "firma" eller arbejdsgiver.

En ændring i "ansættelses"-relationen for en arbejdsgiver skal forstås som en ændring i tilknytningen til arbejdsstedet via tilgang eller afgang i forbindelse med et ejerskifte for et givet arbejdssted.

For hovedbeskæftigede lønmodtagere og arbejdsgivere er endvidere anført arten af mobilitet¹⁵. For bibeskæftigede er desuden anført, om der er sket et skift mellem hovedbeskæftigelse og bibeskæftigelse på samme arbejdssted.

For selvstændige findes tilsvarende oplysninger med hensyn til ændring i status - henholdsvis retrospektivt (uændret selvstændig, tilgang til gruppen af selvstændige) og prospektivt (uændret, afgang). Herunder indgår også oplysninger om tidligere eller efterfølgende placering på/uden for arbejdsmarkedet).

Tilsvarende oplysninger findes for medhjælpende ægtefæller.

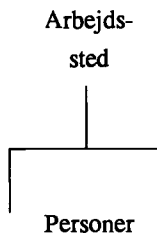
¹⁵ Der indgår følgende kategorier:

- fra/til lønmodtager/ikke-lønmodtager, på et andet arbejdssted i samme firma eller i et andet firma
- overgang fra/til arbejdsløshed eller uden for arbejdsstyrken.

2.4. Oplysninger om virksomheder: Arbejdssteder og firmaer

Nogle af variablene for arbejdssteder er dannet via aggregering over de ansatte personer pr. arbejdssted. Det gælder målene for stillingsstruktur, personaleomsætning, længennemsnit samt ledighed i årets løb blandt de ansatte.

Aggregering over personer:



Det skal her understreges, at dannelse af tilsvarende variable via aggregering kan foretages ud fra individuelle ønsker. Det betyder, at man i udtræk af data til bestemte analyser kan danne andre variable via yderligere aggregeringer over *grupper* af personer. Det kan fx dreje sig om aldersgrupper, uddannelseskategorier, detaljerede stillingsgrupper osv., herunder også i form af kombinationer af forskellige personkarakteristika. Eksempelvis kan man danne variable vedrørende antallet af lærlinge pr. arbejdssted, herunder lærlinge, der starter eller afslutter lærlinguddannelsen. Sådanne aggregeringer behøver ikke at blive foretaget for det enkelte arbejdssted, men kan også foretages for kategorier af arbejdssteder, fx størrelses- eller branchegrupper. De aggregerede variable, som allerede er inddraget i datasættet vedr. arbejdsstederne, er af summarisk karakter og har først og fremmest til formål af give en supplerende, generel karakteristik af arbejdsstederne.

Arbejdssteder

Nøgle

- arbejdsstedets løbenummer¹⁶

Baggrundsoplysninger

- herunder:
ejerform, branche, adresse, oprettelsesår

I IDA-databasen indgår desuden en oplysning, der angiver om arbejdsstedet er et enkeltarbejdsstedsfirma eller en filial inden for et firma med flere arbejdssteder.

For arbejdssteder oprettet efter 1980 (dvs. efter startåret for IDA) findes oprettelsesåret. I tilfælde af at arbejdsstedet fx er oprettet via udskilning fra et andet arbejdssted findes en henvisning til "donor"-arbejdsstedet i det foregående år. Disse oplysninger er del-resultater af fastlæggelsen af arbejdsstedernes identitet over tid, jf. nedenfor.

Antal ansatte og ændring heri

- herunder:
ansatte i alt, bibeskæftigede

For hvert arbejdssted angives desuden antallet af "ansættelsesforhold" i alt i løbet af året¹⁷. Den procentvise ændring i antal *hovedbeskæftigede* lønmodtagere er angivet, både set i relation til november året før og til november året efter. Med udgangspunkt heri findes også en variabel, der beskriver ændringen i beskæftigelsesniveauet på arbejdsstedet (hhv. faldende, uændret, eller stigende beskæftigelsesomfang).

¹⁶ Arbejdsstedets løbenummer er et 10-cifret nummer, hvor de 4 første cifre er lig med oprettelsesåret og de 6 øvrige er et tilfældigt nummer. Oprettelsesåret er året for den novemberopgørelse, hvor arbejdsstedet findes første gang.

¹⁷ Heri indgår *alle* ansættelsesforhold, for hvilke der foreligger en oplysningsseddel.

Personalesammensætning og -omsætning

- andel ikke-faglærte/faglærte arbejdere, lavere/højere funktionærer
- andel kvinder i stillingsgrupper
- ledighed blandt ansatte
- af- og tilgangsrater

De ovennævnte oplysninger om andelen af beskæftigede udgør summariske mål for personalesammensætningen ultimo november. De er dannet via en aggregering over de ansatte. Andelen af kvinder inden for disse fire grupper er ligeledes beregnet.

Endvidere findes oplysninger om det gennemsnitlige ledighedsomfang for den del af gruppen af arbejdere, der også var ansat på arbejdsstedet pr. ultimo november året før. Det drejer sig om den gennemsnitlige ledighedsgrad og det gennemsnitlige antal ledighedsperioder i året (dvs. en indikator for omfanget af midlertidige hjemsendelser).

Endelig findes oplysninger om tilgangsraten (fra november året før) og afgangsraten (til november året efter) blandt ansatte med hovedbeskæftigelse. Der er tale om mål for omfanget af "udskiftning" blandt de ansatte i løbet af en årsperiode - og ikke om et mål for den løbende personaleomsætning ("fra dag til dag")¹⁸.

Lønsum og lønniveau

- herunder:
gennemsnitlig timeløn

Den samlede lønsum for hele året for alle ansatte indgår i IDA. Desuden foreligger et skøn over den gennemsnitlige timeløn for ansatte med hovedbeskæftigelse, dels under ét og dels for hver af grupperne ikke-faglærte, faglærte, lavere og højere funktionærer.

Ændring i arbejdsstedets identitet over tid mv.

- herunder:
retrospektiv/prospektiv klassifikation af identitet

¹⁸ En nærmere redegørelse herfor findes i Per Vejrup Hansen: Arbejdskraftbevægelser og beskæftigelse. Socialforskningsinstituttets publ. 149, København 1985, p. 42-44.

Ændringen i arbejdsstedets identitet fra et år til det næste beskrives både:

- retrospektivt, dvs. set i forhold til året før (bevaret, oprettet), og
- prospektivt, dvs. set fremad til året efter (bevaret, nedlagt).

Endvidere findes en detailklassifikation bestående af forskellige typer af bevarede, oprettede og nedlagte arbejdssteder, jf. kapitel 4. Der indgår en særlig klassifikation vedrørende identitetsændringen for arbejdssteder, der ikke har ansatte pr. november.

Samtidig foreligger der en registrering af, hvilken "regel" der ligger til grund for, at et arbejdssted eventuelt er klassificeret som "bevaret"¹⁹. Herunder indgår oplysninger om ejerskifte, adresseændring og brancheændring for arbejdsstedet.

Endelig findes særlige variable vedrørende oprettelser af arbejdssteder "via udskilning" fra et arbejdssted, henholdsvis nedlæggelser "via opsugning" i et andet arbejdssted. For oprettelser via udskilning drejer det sig bl.a. om identifikationsnummeret for det arbejdssted, som arbejdsstedet er udskilt fra, og om en angivelse af, hvorvidt oprettelsen sker inden for det samme firma eller i form af et nyt firma.

¹⁹ Jf. gennemgang af regelsættet i kapitel 1.

Firmaer

Et firma består af ét eller flere arbejdssteder. Hvis et firma kun består af ét arbejdssted svarer firmaenheden til arbejdsstedsenheden. Ellers er firma-oplysningerne udelukkende dannet via aggregering over de arbejdssteder, der indgår i det pågældende firma i det givne år.

Nøgle

- arbejdsgivernummer

Beskæftigede og arbejdssteder

- herunder:
ansatte i alt, bibeskæftigede, antal arbejdssteder, største arbejdssted

Til karakteristik af firmaet indgår antallet af arbejdssteder (herunder antal bevarede arbejdssteder set i forhold til året før²⁰). Endvidere findes det samlede antal ansatte i firmaet - opdelt på antal hovedbeskæftigede og antal bibeskæftigede. Antallet af ansatte, for hvem det er uoplyst hvilket arbejdssted, de er ansat på, er ligeledes opgjort.

Til karakteristik af firmaet indgår endvidere kendetegn for det største arbejdssted i firmaet, med angivelse af arbejdsstedsløbenummer, ændring i identitet set tilbage i tid, ændring i arbejdsgivernummer i forhold til året før, samt arbejdsstedets branche.

²⁰ Dette kan evt. indgå i en indirekte fastlæggelse af identitet over tid for firmaer.

3. Database-organiseringen af IDA - oversigt over datasættene

IDA er ikke blot en database af navn, men også af gavn.

Databaseorganiseringen kommer konkret til udtryk i, at dataene er delt op i et antal *forskellige datasæt*. For det første er der tre hovedgrupper af datasæt, svarende til de tre enheder i IDA: Personer, ansættelser og virksomheder (jf. figur 1 i kapitel 1). For det andet findes der forskellige datasæt for hver af disse enheder. Specielt for enheden ansættelser/beskæftigede findes der et større antal datasæt, idet der her bl.a. er tale om forskellige kategorier af ansatte/beskæftigede, som fx lønmodtagere versus arbejdsgivere. På dette punkt bevirker opdelingen i forskellige datasæt, at man sikrer, at der ikke foretages sammenblandinger af forskellige kategorier. Det bevirker også en forenkling af analysearbejdet, idet man til mange formål kan tænkes kun at interessere sig for lønmodtagere (ansatte). I så fald er det hensigtsmæssigt, at der foreligger et datasæt, som udelukkende omfatter denne bestemte population.

I forbindelse med givne analyser skal man selvfølgelig kunne forbinde data (variable) fra *flere* datasæt. Det muliggøres ved hjælp af "*identifikations-nøgler*" i datasættene, dvs. ensartede identifikationsnumre for de enkelte enheder (personer, arbejdssteder mv.). I de fleste analyser vil der typisk være tale om at inddrage data fra flere datasæt, fx for forskellige enheder (personer og arbejdssteder) eller for forskellige år. Databaseorganiseringen kan i denne forbindelse bidrage til at specificere den *population*, som skal være udgangspunkt for analysen - dvs. fx en bestemt gruppe af ansatte eller en bestemt kategori af arbejdssteder. Et centralt formål med opdelingen i en række datasæt består i at fastlægge bestemte populationer af enheder.

I virkeligheden er databaseorganiseringens formål således at gøre dataindholdet mere overskueligt, idet organiseringen understreger, at der findes forskellige typer af enheder og kategorier. I denne forstand er databaseorganiseringen et udtryk for en begrebsmæssig model af "virkeligheden".

Inden vi i det følgende kaster os ud i, hvordan IDA ser ud som database, skal vi kort berøre spørgsmålet om *IDA-brugerens nytte af at kende databaseorganiseringen*. Vores budskab er, at det ikke er et "must", at brugeren skal kende databasedesignet - men at det kan være særdeles nyttigt. Sidstnævnte gælder især i de tilfælde, hvor brugeren ikke har en helt klar opfattelse af opgavens karakter og omfang, dvs. når der ikke er tale om specifikke, afgrænsede analyser. Et kendskab til databaseorganiseringen og strukturen af datasættene giver et grundlag for at forstå *mulighedsområdet* for analyser. Det er selvfølgelig Danmarks Statistiks opgave - som leverandør af IDA-data - at give de tekniske løsninger og svar på givne analysebehov, men det er på den anden side kun brugeren, der har den fulde fornemmelse af den analytiske problemstilling, og derfor er det en stor fordel, hvis brugeren selv har et indtryk af analysemulighederne.

I afsnit 3.2 om de enkelte datasæt vil der da også blive peget på nogle generelle analysemuligheder med udgangspunkt i de enkelte datasæt.

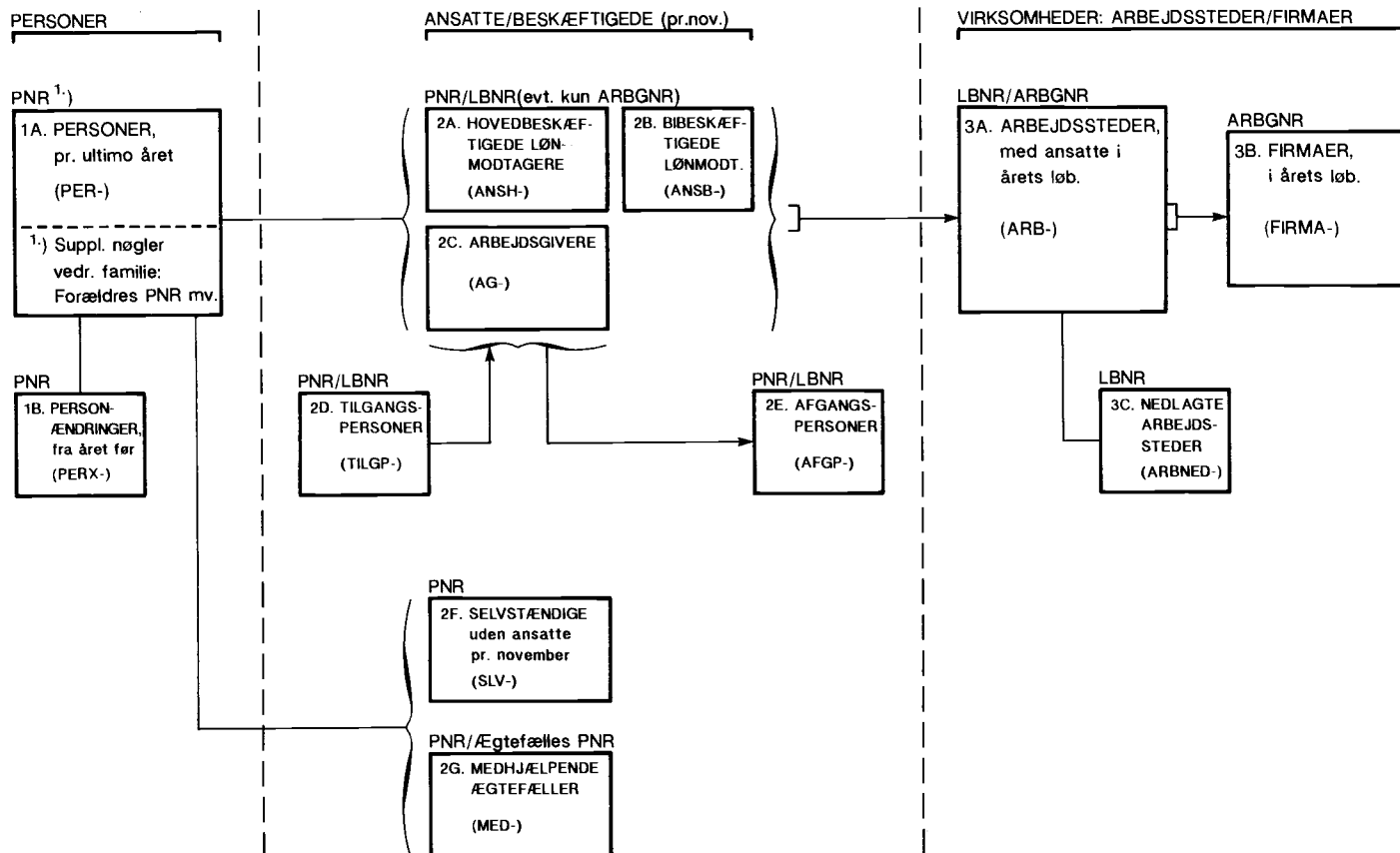
3.1. Databasestrukturen

Databaseorganiseringen fremgår af *figur 2*.

Der findes i alt 12 datasæt *pr. år*, dvs. for henholdsvis 1980, 1981, 1982 etc. frem til og med det seneste år, som indgår i IDA, dvs. det seneste kalenderår, for hvilket de tilgrundsiggende statistikregistre foreligger. For *hver* kategori af datasæt foreligger der altså et antal fysiske datasæt svarende til hvert af de år, som på et givet tidspunkt indgår i IDA. Der er dog ét datasæt, som kun findes i én udgave, nemlig "3C. NEDLAGTE ARBEJDSSTEDER", idet dette datasæt indeholder en successiv kumulering af endeligt nedlagte arbejdssteder i alle årene efter 1980.

I venstre side af figuren vises datasættene vedrørende *personer*, dvs. med data om hver enkelt person i befolkningen. Det drejer sig om 2 datasæt, 1A og 1B, hvoraf 1B vedrører ændringer i personernes baggrundsdata mellem to år. I højre side fremgår datasættene vedrørende *virksomheder*, henholdsvis arbejdssteder og firmaer, som omfatter i alt 3 datasæt, 3A, 3B og 3C. I midten af figuren findes endelig det resterende antal datasæt, i alt 7 (2A - 2G), vedrørende *ansatte/beskæftigede* pr. november i året.

Figur 2. Databasestruktur. Datasæt pr. år.



Enhederne (personerne) i de tre øverste af disse datasæt vedrørende ansatte mv. (2A - 2C) kan forbindes med de enkelte enheder på virksomhedssiden, idet disse personer er tilknyttet et arbejdssted eller et firma qua ansatte (lønmotagere) eller qua arbejdsgivere. For disse personer findes ud over personnummeret PNR også identifikationsnummeret LBNR for det arbejdssted, hvor den enkelte person er ansat eller er arbejdsgiver. (LBNR står for det arbejdssteds-løbenummer, der genereres i IDA-systemet, og som er konstant over tid for givne arbejdssteder). Generelt er identifikationsnumrene angivet ovenpå kassen vedrørende det enkelte datasæt.

For de selvstændiges vedkommende (datasæt 2F) er det af begrebsmæssige og tekniske grunde ikke muligt at sammenkæde personen med et arbejdssted¹.

Det samme gælder de medhjælpende ægtefæller (datasæt 2G) i de tilfælde, hvor ægtefællen er selvstændig. Er ægtefællen derimod arbejdsgiver, er det muligt indirekte at forbinde den medhjælpende ægtefælle med et arbejdssted (eller firma) via arbejdsgiver-ægtefællen.

Blandt datasæt vedrørende ansatte mv. findes der endelig to særlige datasæt for de personer blandt hovedbeskæftigede lønmotagere og arbejdsgivere på et arbejdssted pr. november i det givne år, som er tilgået arbejdsstedet i løbet af den forudgående årsperiode (datasættet 2D), eller som afgang fra arbejdsstedet i løbet af den følgende årsperiode (datasættet 2E). Disse særlige datasæt kan betragtes som "analytiske" datasæt. Heri ligger, at de i princippet kun indeholder oplysninger, som også findes i datasættene over alle ansatte mv. (dvs. i 3A - 3C). De to datasæt over henholdsvis "tilgangspersoner" og "afgangspersoner" indeholder således analytisk orienterede eller bearbejdede oplysninger om bl.a. personernes placering i henholdsvis året før og året efter.

På tilsvarende måde er datasættet 1B vedrørende "personændringer" og datasættet 3C vedrørende nedlagte arbejdssteder analytisk orienterede datasæt. I figur 2 over database-

¹ Det skyldes på den ene side, at arbejdssteder defineres som enheder med lønnet personale, hvilket betyder, at der for selvstændige, som overhovedet ikke har ansatte i hele det pågældende år ikke findes en tilsvarende virksomhedsenhed. På den anden side er der tale om tekniske årsager til, at man heller ikke for selvstændige med ansatte i løbet af året (uden for november) kan sammenkæde den selvstændige som person med den tilsvarende arbejdsstedsenhed (her med ansatte i årets løb uden for november). Inden for denne kategori af arbejdssteder kan der således være tale om en overlapning med individer blandt selvstændige.

strukturen findes disse fire analytisk orienterede datasæt aftegnet som de fire små kasser på stribe under de store kasser for de grundlæggende enheder.

3.2. De enkelte datasæt: Population, identifikationsnøgle, variabelindhold og anvendelsesområder

I dette afsnit gives en nærmere redegørelse for hver enkelt af de 12 datasæt i databasen, jf. figur 2. I redegørelsen vil der indgå fire hovedpunkter i karakteristikken af det enkelte datasæt, nemlig:

- *populationen*, dvs. definitionen og afgrænsningen af de pågældende enheder, samt eventuelt en omtale af relevante del-populationer.
- *identifikationsnøglen/-nøglerne*, dvs. det én-entydige nummer, der definerer den enkelte enhed (person, ansættelse, arbejdssted).
- *variabelindholdet*, hvori indgår en generel, kortfattet karakteristik af variabelindholdet.
- *anvendelsesområder*, hvorunder det for nogle af datasættene kort skitseres, hvilke analyseområder der især kan være relevante på basis af - eller med udgangspunkt i - det pågældende datasæt.

Datasæt 1A. Personer

Datasætnavn: PER(år), fx PER80.

Population: Datasættet omfatter den samlede "aktive" befolkning i Danmark pr. ultimo året (den 31. december), dvs. personer med bopæl i Danmark. Heri indgår også personer, som er midlertidigt fraværende fx på grund af ferie, forretningsrejse eller lignende.

Identifikationsnøgle: Personnummer (PNR).

Variable: Datasættet, som er meget omfattende, indeholder historiske baggrundsvariable, statusvariable og årsvariable for personerne.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 1A.

Anvendelsesområder: I forbindelse med analyser af ansættelses- og beskæftigelsesforhold udgør disse variable et grundlag for udvælgelse af bestemte personkategorier, herunder fx i relation til familieforhold.

Datasæt 1B. Personændringer

Datasætnavn: PERX(år)(år), fx PERX8081.

Population: Datasættet indeholder både den samlede befolkning i Danmark ved årsskiftet det ene år og ved årsskiftet det følgende år (dvs. foreningsmængden af de to befolkninger, svarende til befolkningen ved årsskiftet det følgende år med tillæg af udvandrede og døde i årets løb).

Identifikationsnøgle: Personnummer (PNR).

Variable: Datasættet omfatter variable, der beskriver *ændringer* i forskellige status- og baggrundsvARIABLE for personerne fra det ene år til det næste.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 1B.

Anvendelsesområder: Analyser af sammenhænge mellem ændringer i baggrundsvARIABLE mv. og ansættelses- og beskæftigelsesforhold (fx beskæftigelsessituation efter afslutning af uddannelse, eller som følge af barnefødsel). Endvidere analyser af arbejdsstyrkemobilitet, herunder især vedrørende tilgang til arbejdsstyrken.

Datasæt 2A. Hovedbeskæftigede lønmodtagere

Datasætnavn: ANSH(år), for seneste år dog ANSTIH(år), fx ANSTIH87.

Population: Datasættet omfatter alle personer i den "aktive" befolkning ved årsskiftet, der i den vigtigste tilknytning til arbejdsmarkedet pr. ultimo november var lønmodtagere. I denne population indgår også skoleelever, studerende mv. med en varierende grad af tilknytning til arbejdsmarkedet, og hvoraf den vigtigste pr. ultimo november bestod i at være lønmodtagere. Det skal derfor understreges, at de hovedbeskæftigede lønmodtagere er personer, hvis *vigtigste tilknytning til arbejdsmarkedet* er at være lønmodtagere. En relevant *del*population kan således bestå af personer, hvis "hovedaktivitet" består i at være lønmodtagere (dvs. at der ses bort fra skoleelever o.l. i arbejdsstyrken). Personer, der er fuldt ledige i registreringsugen (den sidste uge af november), klassificeres som arbejdsløse og *indgår ikke* i dette datasæt, selvom de fx har haft hovedbeskæftigelse som lønmodtager resten af året².

Identifikationsnøgler: Personnummer (PNR) og Arbejdssteds-løbenummer (LBNR).

Endvidere indgår nummeret for firmaet/arbejdsgiverenheden (ARBGNR) af hensyn til de tilfælde, hvor det er uoplyst, på hvilket arbejdssted en ansat er beskæftiget.

Variable: Formålet med variablene er at karakterisere det pågældende ansættelsesforhold pr. ultimo november, samt at registrere og karakterisere eventuelle ændringer i ansættelsesforholdet (tilgang til/afgang fra arbejdsstedet).

Udover deciderede ansættelsesvariable (om ansættelsens karakter, løn mv.) indeholder datasættet af praktiske grunde nogle få personoplysninger som fx køn, alder (pr. ultimo november) og uddannelseskategori.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2A.

Anvendelsesområder: Analyser vedrørende ansættelsesvilkår samt skift af arbejdssted, skift af stilling på samme arbejdssted, mv. Endvidere udgør dette datasæt hovedgrundlaget for aggregeringer (optællinger mv.) over de ansatte personer pr. arbejdssted. Via sådanne aggregeringer kan man foretage analyser af omfanget af arbejdspladsmobilitet, samt fx af midlertidige hjemsendelser inden for forskellige erhverv/brancher.

² For en definition af personens vigtigste tilknytning til arbejdsmarkedet se Arbejdsnotat nr. 27. En nærmere omtale af hovedprincipperne findes i den årlige artikel i Statistiske Efterretninger vedrørende den Registerbaserede arbejdsstyrkestatistik (RAS).

Datasæt 2B. Bibeskæftigede lønmodtagere

Datasætnavn: ANSB(år), for seneste år dog ANSTIB(år).

Population: Datasættet omfatter personer, der som næstvigtigste tilknytning til arbejdsmarkedet pr. ultimo november er lønmodtagere. Der er altså tale om et bijob ved siden af en vigtigste tilknytning som enten lønmodtager, arbejdsgiver, selvstændig eller medhjælpende ægtefælle (dvs. at personen tillige findes i et af de øvrige datasæt vedrørende ansættelser/beskæftigede, 2A, 2C, 2F eller 2G). Populationen omfatter dog også efterlønsmodtagere med (begrænset) erhvervsmæssig beskæftigelse.

Identifikationsnøgler: Personnummer (PNR) og arbejdsstedsløbenummer (LBNR), samt firmanummeret (ARBGNR).

Variable: Variablene karakteriserer ansættelsesforholdet samt ændringer i ansættelsen som bibeskæftiget på arbejdsstedet, herunder skift mellem hovedbeskæftigelse og bibeskæftigelse på samme arbejdssted.

Af praktiske grunde indgår nogle få personvariable som køn og alder pr. ultimo november.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2B.

Anvendelsesområder: Analyser af omfang og karakter af lønmodtager-bibeskæftigelse blandt personer i arbejdsstyrken, herunder blandt fx hovedbeskæftigede lønmodtagere eller selvstændige. Via aggregering over bibeskæftigede lønmodtagere pr. arbejdssted eller pr. branche kan omfanget og karakteren af bibeskæftigelser set fra virksomhedssiden belyses. Ved opgørelser af den samlede beskæftigelse pr. arbejdssted mv. kan summen af hoved- og bibeskæftigede inddrages.

Datasæt 2C. Arbejdsgivere

Datasætnavn: AG(år), for seneste år dog AGTI(år).

Population: Datasættet omfatter de personer, der ejer (eller er medejer af) en virksomhed, hvor der er mindst én lønmodtager ansat (som hoved- eller bibeskæftiget) pr. ultimo november. Denne placering som arbejdsgiver betragtes som personens vigtigste tilknytning til arbejdsmarkedet³.

Identifikationsnøgler: Personnummer (PNR) og arbejdsstedsløbenummer (LBNR).

Endvidere indgår firmanummeret (ARBGNR), idet identifikationsnummeret LBNR for arbejdsstedet kan være uoplyst, hvis arbejdsgiveren har flere arbejdssteder. Da kan arbejdsgiveren kun forbindes med firmaet.

Variable: I dette datasæt indgår de samme variable som i datasættet 2A vedrørende hovedbeskæftigede lønmodtagere⁴.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2C.

Anvendelsesområder: Udgangspunkt for analyser af omfanget af mobilitet til/fra gruppen af arbejdsgivere (sammen med datasættene vedr. tilgangs- og afgangspersoner, 2D og 2E). Uddannelsesbaggrund mv. blandt arbejdsgivere. Kan inddrages som en del af det samlede beskæftigelsesomfang på arbejdsstedet.

³ Antallet af arbejdsgivere i IDA svarer ikke præcist til de offentliggjorte oplysninger fra den Registerbaserede arbejdsstyrkestatistik (RAS), idet tallet er reduceret med nogle få hundrede.

Det skyldes, at der findes arbejdssteder uden ansatte i november, hvor der ifølge RAS samtidig findes en person, der er klassificeret som "arbejdsgiver" på det pågældende arbejdssted. I disse tilfælde er personen i IDA omklassificeret til "selvstændig". Det medfører selvsagt, at gruppen af selvstændige vokser tilsvarende.

⁴ Variablene til karakteristik af ansættelsesforholdet har dog en begrænset relevans. Løn-oplysningen er således irrelevant, og stillingskoden er altid lig med værdien for arbejdsgiver.

Datasæt 2D. Tilgangspersoner

Datasætnavn: TILGP(år).

Population: Dette datasæt omfatter den *del* af de hovedbeskæftigede lønmodtagere og arbejdsgivere, der er "nyansatte" på et arbejdssted pr. november i det pågældende år, dvs. som ikke var ansat på arbejdsstedet ultimo november året før.

Populationen er således en delpopulation af populationerne i datasættene 2A og 2C.

Identifikationsnøgler: Personnummer (PNR) og arbejdsstedsløbenummer (LBNR) for tilgangs-arbejdsstedet.

Variable: Variablene beskriver primært personens placering pr. november *året før*, dvs. hvor personen kommer fra. For tidligere beskæftigede beskrives fx, om personen har skiftet erhverv og/eller branche - og et eventuelt skift af bopælsadresse i løbet af året registreres. I forbindelse med arbejdssteder, der er oprettede "via udskilning" fra et andet arbejdssted, angives om tilgangspersonen kommer fra "donorarbejdsstedet" eller ej.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2D.

Anvendelsesområder: For personer, der var beskæftiget ultimo november året før, og som tilgår et arbejdssted som lønmodtagere/arbejdsgivere, kan der foretages analyser af bl.a. erhvervsmobilitet og stillingsskift. Generelt kan man undersøge, hvor de nyansatte kommer fra.

Datasæt 2E. Afgangspersoner

Datasætnavn: AFGP(år), idet det seneste år, for hvilket datasættet kan optræde, er det næstsidste år i IDA.

Population: Datasættet omfatter (analogt til datasættet om tilgangspersoner) den del af de hovedbeskæftigede lønmodtagere og arbejdsgivere, der i løbet af den følgende årsperiode afgår fra det arbejdssted, hvor de er ansat pr. ultimo november i det pågældende år - dvs. at de ikke genfindes på arbejdsstedet ultimo november året efter.

Populationen er således en delpopulation af populationerne i datasættene 2A og 2C.

Identifikationsnøgler: Personnummer (PNR) og arbejdsstedsløbenummer (LBNR) for afgangsarbejdsstedet.

Variable: Variablene beskriver primært personens placering pr. november *året efter*, dvs. hvortil personen bevæger sig. For beskæftigede året efter beskrives fx, om personen foretager et skift af erhverv og/eller branche, og et eventuelt skift af bopælsadresse i løbet af året registreres. For arbejdssteder, der er nedlagt "via opsugning" i et andet arbejdssted, angives om personen overgår til det modtagende (opsugende) arbejdssted.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2E.

Anvendelsesområder: For fortsat beskæftigede personer, der afgår fra et arbejdssted, kan der foretages analyser af bl.a. erhvervsmobilitet og stillingsskift. Generelt kan man undersøge, hvortil afgangspersoner bevæger sig.

Datasæt 2F. Selvstændige

Datasætnavn: SLV(år), for seneste år dog SLVTI(år).

Population: Datasættet omfatter personer med selvstændig virksomhed, som ikke har ansatte pr. ultimo november i det pågældende år, og hvor der er tale om personens vigtigste tilknytning til arbejdsmarkedet⁵.

Der kan forekomme beskæftigelse af lønnet personale i løbet af året uden for registreringsugen i november.

Identifikationsnøgle: Personnummer (PNR).

Variable: Datasættet indeholder relativt få variable, der karakteriserer den selvstændige som person og arten af virksomheden. Desuden registreres ændringer i status som selvstændig.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2F.

Anvendelsesområder: Beskrivelse af den population af virksomheder, der indehaves af selvstændige. Denne virksomhedspopulation kan således supplere populationen af arbejdssteder (datasættet 3A), omend det af tekniske grunde ikke er muligt eksakt at identificere den overlapning, der forekommer mellem selvstændige virksomheder og arbejdssteder med ansatte i løbet af året (uden for november).

Beskrivelse af de selvstændiges uddannelsesbaggrund mv., samt mobiliteten til/fra gruppen af selvstændige.

⁵ En person klassificeres som selvstændig, hvis personen enten er momsbetaler, arbejdsløshedsforsikret som selvstændig, eller har sin væsentligste indkomst fra selvstændig virksomhed.

Datasæt 2G. Medhjælpende ægtefæller

Datasætnavn: MED(år), for seneste år dog MEDTI(år).

Population: Datasættet omfatter personer, der har deres væsentligste indkomst i løbet af året som medhjælpende ægtefælle, og som er gift med en person, der er arbejdsgiver eller selvstændig⁶.

Identifikationsnøgler: Personnummer (PNR) og ægtefælles personnummer (AGTEPNR).

Variabelkategorier: Datasættet indeholder relativt få variable, herunder baggrundsdata for personen, arten af virksomheden og ændringer i status som medhjælpende ægtefælle. En fuldstændig oversigt over variablene findes i bilag 1, afsnit 2G.

Anvendelsesområder: Beskrivelse af gruppen af medhjælpende ægtefæller, samt mobiliteten til/fra gruppen. Via ægtefælles personnummer kan personer i gruppen forbindes med gruppen af selvstændige eller gruppen af arbejdsgivere, og dermed i sidstnævnte tilfælde også med arbejdsstederne.

⁶ Mere præcist skal personen blot indgå i *bruttopopulationen* af selvstændige i RAS. Det betyder, at der i ganske få tilfælde vil forekomme personer klassificeret som "medhjælpende ægtefæller", hvor ægtefællen ikke indgår som arbejdsgiver eller selvstændig.

Datasæt 3A. Arbejdssteder

Datasætnavn: ARB(år), for seneste år dog ARBTI(år).

Population: Datasættet omfatter de erhvervsenheder, der i årets løb har haft lønnet personale beskæftiget, og for hvilke der er indbetalt kildeskat. Arbejdsstedet er den enkelte (lokale) erhvervsenhed, som organisatorisk, geografisk og branchemæssigt er sammenhængende.

Der er tale om en "bruttopopulation" af arbejdssteder, som også omfatter arbejdssteder, der ikke har ansatte pr. ultimo november - fx på grund af sæsonpræget beskæftigelse, eller fordi arbejdsstedet kun eksisterer i en del af året (dvs. at det nedlægges før, eller oprettes efter, ultimo november)⁷.

En relevant *del*population er således arbejdssteder *med* ansatte pr. ultimo november.

En anden relevant *del*population kan være arbejdssteder med *hoved*beskæftigede pr. ultimo november - idet det kan forekomme, at der alene er tale om *bibeskæftigede* på arbejdsstedet.

Endelig kan man anvende forskellige *del*populationer af bevarede, nedlagte eller oprettede arbejdssteder (jf. kapitel 4).

Identifikationsnøgler: Arbejdsstedsløbenummer (LBNR) og firmanummer (ARBGNR).

Endvidere indgår den originale arbejdsstedskode af hensyn til eventuelle ad hoc samkøring med supplerende statistikregistre.

Variable: Datasættet omfatter baggrundsdata for arbejdsstedet samt ændringer i arbejdsstedets identitet over tid (bevaret, nedlagt, oprettet). Der indgår endvidere nogle summariske mål vedrørende arbejdsstyrken og ændringer i arbejdsstyrken, som kan udgøre en karakteristik af arbejdsstedet (med hensyn til fx sammensætningen af personalet, lønniveauet og omfanget af til- og afgang blandt de ansatte).

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 3A.

⁷ En nærmere beskrivelse af bruttopopulation og klassifikationer fremgår af kapitel 4.

I IDA indgår ikke de "arbejdssteder" i arbejdsstedsregistret, som består af ansatte med uoplyst arbejdsstedskode inden for et firma. Det drejer sig om ca. 2000 arbejdssteder, der ikke medregnes i IDA.

Lokale enheder inden for forsvaret indgår ikke i arbejdssteder i årene 1980-1986 (inkl.). Disse inddrages i 1987 som oprettede af typen O1, dvs. "nyoprettet".

Anvendelsesområder: Forløb af bevarede arbejdssteder over tid, fx vedrørende beskæftigelses- og lønudvikling. Omfang og type af nedlæggelser og oprettelser af arbejdssteder. Som udgangspunkt for analyser af personers mobilitet mv. kan der foretages grupperinger af arbejdssteder med hensyn til branche, størrelsesgruppe, vækst eller nedgang i beskæftigelsesomfanget, nedlagte eller oprettede enheder osv. I forbindelse med visse analyser kan det være relevant kun at tage udgangspunkt i bevarede arbejdssteder, herunder evt. kun i bevarede og identiske enheder (af typen B1); sidstnævnte kan fx gælde ved analyser af omfanget af arbejdspladsmobilitet, jf. også kapitel 4.

Datasæt 3B. Firmaer

Datasætnavn: FIRMA(år).

Population: Ved et firma - en arbejdsgiverenhed - forstås i den private sektor den enhed, der indeholder og afregner kildeskat for ansatte (SE-enheden). Et firma kan være en del af en "koncern", bestående af flere firmaer. I den statslige sektor er firmaenheden identisk med en styrelse, defineret ved hovedkonto i finansloven. I den kommunale sektor udgør hvert fysisk adskilt arbejdssted et firma.

Et firma består af ét eller flere arbejdssteder. I tilfælde af, at et firma kun omfatter ét arbejdssted, er firmaenheden identisk med den tilsvarende arbejdsstedsenhed (i datasættet 3A vedrørende arbejdssteder)⁸.

Identifikationsnøgle: De enkelte enheder er defineret ved arbejdsgivernummeret (ARBGNR) - men der foreligger *ikke* en fastlæggelse af identiteten over tid for firmaer. Identifikationsnøglen ARBGNR er ikke konstant over tid for givne, bevarede enheder, idet nummeret fx kan ændres i forbindelse med ejerskifte⁹.

Variable: Variablene er dannet via aggregering over de arbejdssteder, der indgår i det pågældende firma i det givne år. Det drejer sig om antallet af arbejdssteder og det samlede antal ansatte i firmaet. Til karakteristik af firmaet indgår endvidere oplysninger om det største arbejdssted i firmaet.

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 3B.

Anvendelsesområder: Analyser af firmaniveau, herunder filialstruktur, diversifikation mv.

⁸ Firmaer, der ikke er "repræsenteret" ved arbejdssteder - herunder forsvaret i årene 1980-86 samt enkelte andre tilfælde - indgår ikke i datasættet firmaer.

⁹ Dog har nogle af variablene til formål at give en delvis indikation af et firmas identitet over tid (antal arbejdssteder, det største arbejdssteds identitet tilbage i tid, skift i firmaets arbejdsgivernummer).

Datasæt 3C. Nedlagte arbejdssteder

Datasætnavn: ARBNED(år), hvor år er lig med det seneste år, som indgår i IDA.

Population: Datasættet omfatter "endeligt" nedlagte arbejdssteder i perioden efter 1980 og indtil ultimo det seneste år i databasen¹⁰. Der er således tale om en kumulering af samtlige "endeligt" nedlagte arbejdssteder i hele perioden siden 1980.

Endeligt nedlagte arbejdssteder udgøres af afgangene fra *bruttopopulationen* af arbejdssteder, dvs. eksklusive fx enheder, der overgår fra arbejdssteder med ansatte pr. november til arbejdssteder uden ansatte pr. november¹¹.

Identifikationsnøgle: Arbejdsstedsløbenummer (LBNR).

Variable: Datasættet omfatter årstallet for det sidste år, hvor arbejdsstedet eksisterer, samt identiteten set fremad for det sidste år (dvs. "arten" af nedlæggelse).

En fuldstændig oversigt over variablene findes i bilag 1, afsnit 3C.

Anvendelsesområder: Overlevelsestavler for arbejdssteder, herunder nedlæggeshypotheder afhængigt af arbejdsstedets alder mv.

¹⁰ Når et arbejdssted nedlægges "endeligt", ophører det pågældende arbejdsstedsløbenr. LBNR med at eksistere i populationen af "aktive" arbejdssteder i datasættet 3A.

¹¹ For en præcis definition af endeligt nedlagte arbejdssteder henvises til kapitel 4.

4. Klassifikationen af arbejdssteders identitet over tid

I dette kapitel gives en nærmere redegørelse for klassifikationen af arbejdssteders identitet over tid, dvs. af de forskellige kategorier af ændringstyper (bevarede, nedlagte og oprettede arbejdssteder mv.)

Til bestemte analyseformål er der hensigtsmæssigt at tage udgangspunkt i bestemte kategorier af ændringstyper. Det gælder især i forbindelse med analyser af oprettelser og nedlæggelser af arbejdssteder.

4.1. To delpopulationer af arbejdssteder

I forbindelse med klassifikationen opereres der med *to forskellige delpopulationer* af arbejdssteder.

Arbejdsstederne som helhed omfatter enheder, der har ansatte *i løbet af* et givet kalenderår. Men en del af arbejdsstederne har *kun beskæftigede i en del af året*, fx kun i sommer-sæsonen. Sådanne arbejdssteder indgår ikke i den centrale sammenkædning af personer og arbejdssteder, idet sammenkædningen "kun" vedrører et bestemt tidspunkt i året, nemlig pr. ultimo november måned. Det betyder, at antallet af ansatte personer og sammensætningen af disse ikke kan registreres for arbejdssteder, der ikke har ansatte pr. ultimo november.

Hertil kommer, at gruppen af arbejdssteder uden ansatte pr. november er *heterogent* sammensat. Kategorien omfatter ikke blot arbejdssteder med sæsonpræget beskæftigelse, men også enheder, der er ophørt i løbet af året (dvs. før ultimo november) eller som er nystartede i slutningen af året (efter ultimo november). Hvorvidt der er tale om det ene eller det andet fremgår ikke direkte af registret, idet der ikke indgår oplysninger om start eller ophør af arbejdssteder.

Hovedinteressen må følgelig samle sig om de *arbejdssteder, der har ansatte pr. ultimo november*.

Konsekvensen af denne forekomst af to forskellige delpopulationer er, at der er udarbejdet *særskilte klassifikationer for hver af populationerne*. Det betyder, at fx klassifikationen af bevægelser ind i og ud af delpopulationen af arbejdssteder *med* ansatte pr. november både omfatter skift fra eller til den anden population af arbejdssteder *uden* ansatte pr. november (men dermed inden for "bruttopopulationen" af arbejdssteder), *og* bevægelser ind i eller ud af bruttopopulationen af samtlige arbejdssteder. Dette er illustreret i *figur 3*, som vi skal vende tilbage til i det følgende.

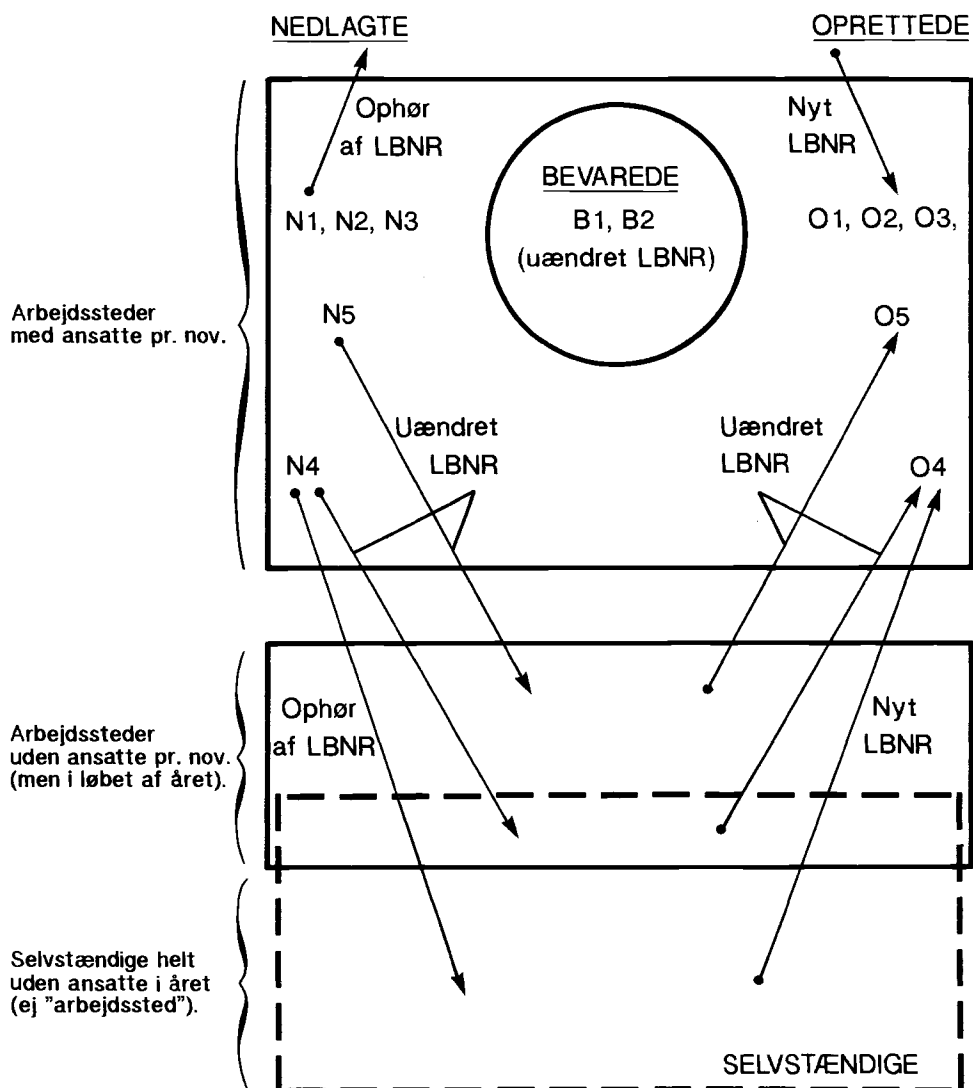
En følge af sondringen mellem de to delpopulationer er således, at en afgang af et arbejdssted fra en af populationerne ikke nødvendigvis er ensbetydende med en afgang fra *bruttopopulationen* af arbejdssteder - og dermed ikke en "endelig" nedlæggelse" - idet der kan ske overgang til den anden delpopulation. Alligevel rubriceres sådanne tilfælde af "overgang" som en "nedlæggelse" (eller afgang) i klassifikationen af den enkelte population - omend det sker i form af særlige kategorier, der angiver, at det sker i form af en overgang til/fra den anden population af arbejdssteder. Man kan sige, at der på denne måde skelnes skarpt mellem tilstandene 0 ansat og 1 ansat. Begrundelsen herfor er imidlertid, at et skift til/fra nul ansatte *kan* være ensbetydende med et ophør eller en nyoprettelse. På den anden side medfører et skift fra 1 til 0 ansat altså ikke, at der foreligger en "endelig" nedlæggelse, eller at der foreligger en nyoprettelse ved skift fra 0 til 1 ansat.

Fastlæggelsen af, hvornår der er tale om en "endelig nedlæggelse", omtales i det sidste afsnit 4.4. Heri indgår som et centralt aspekt en omtale af relationen mellem populationen af arbejdssteder og gruppen af selvstændige (jf. også figur 3).

I de følgende to afsnit skal klassifikationen af ændringer over tid behandles for henholdsvis arbejdssteder med ansatte pr. november og arbejdssteder uden ansatte pr. november.

Klassifikationen er et slutresultat af begrebsdannelsen og operationaliseringen vedrørende identiteten over tid. Dette grundlag fremgår kortfattet af kapitel 1, afsnit 1.4, og detaljeret af rapporten om virksomheders identitet over tid (Arbejdsnotat nr. 29).

Figur 3. Den samlede population af arbejdssteder samt selvstændige, incl. bevægelser til og fra arbejdssteder med ansatte pr. nov.



4.2. Klassifikation af arbejdssteder med ansatte pr. november

Ændringen af identiteten over tid kan dels dreje sig om ændringen set "*fremad i tid*", hvor et arbejdssted kan være bevaret eller nedlagt - og dels om ændringen set "*tilbage i tid*", hvor der enten kan være tale om en bevaret eller en oprettet enhed. Tidsdimensionen er årsperioden mellem ultimo november i et år og ultimo november i året efter eller i året før. Arbejdsstederne, der indgår i kategorien af bevarede enheder, er de samme, uanset om man ser fremad i tid eller tilbage i tid mellem to november-opgørelser.

En oversigt over klassifikationen fremgår af *oversigt 1*, som både indeholder kategorierne vedrørende fremad i tid og tilbage i tid. Der er tre hovedkategorier: Bevaret, nedlagt og oprettet. Desuden findes i *figur 4* en illustration af nogle af kategorierne. Et indtryk af størrelsesforholdene kan fås af *bilagstabellerne 1 og 2* bagest i kapitlet. Disse vedrører 1983-84.

Bevarede arbejdssteder

Det centrale aspekt er her sondringen mellem "*bevaret og identisk*" (B1) og "*bevaret, men ikke-identisk*" (B2).

Ikke-identisk vil sige, at der foreligger en "organisatorisk" ændring i den forstand, at der forekommer en udskilning af ansatte personer til et andet arbejdssted, der (som hovedregel) er oprettet - eller en "opsugning" af ansatte fra et andet arbejdssted, der (som hovedregel) er nedlagt. Der er med andre ord tale om henholdsvis en opdeling eller en sammenlægning, hvori den bevarede (men ikke-identiske) enhed udgør den betydeligste del¹.

For kategorien "bevaret, ikke-identisk" er der således tale om, at afgang eller tilgang af ansatte - og dermed antallet af ansatte - er påvirket af en opdeling eller en sammenlægning. Det skal man være opmærksom på i forbindelse med analyser af arbejdskraftbevægelser og ændringer i beskæftigelsesomfanget for disse arbejdssteder.

¹ Defineret som den enhed, der i forbindelse med en opdeling eller sammenlægning har en uændret ejer-identifikation (arbejdsgivernummer og arbejdsstedskode), eller som er størst målt ved antal ansatte. Den anden enhed i forbindelse med en opdeling vil oftest være oprettet via "udskilning" (typen O2), mens den anden enhed i forbindelse med en sammenlægning typisk vil være nedlagt via "opsugning" (typen N2), jf. nedenfor.

Oversigt 1 Klassifikation af arbejdssteders identitet set fremad i tid (bevaret/nedlagt), hhv. set tilbage i tid (bevaret/oprettet). Arbejdssteder med ansatte pr. november i det pågældende år.

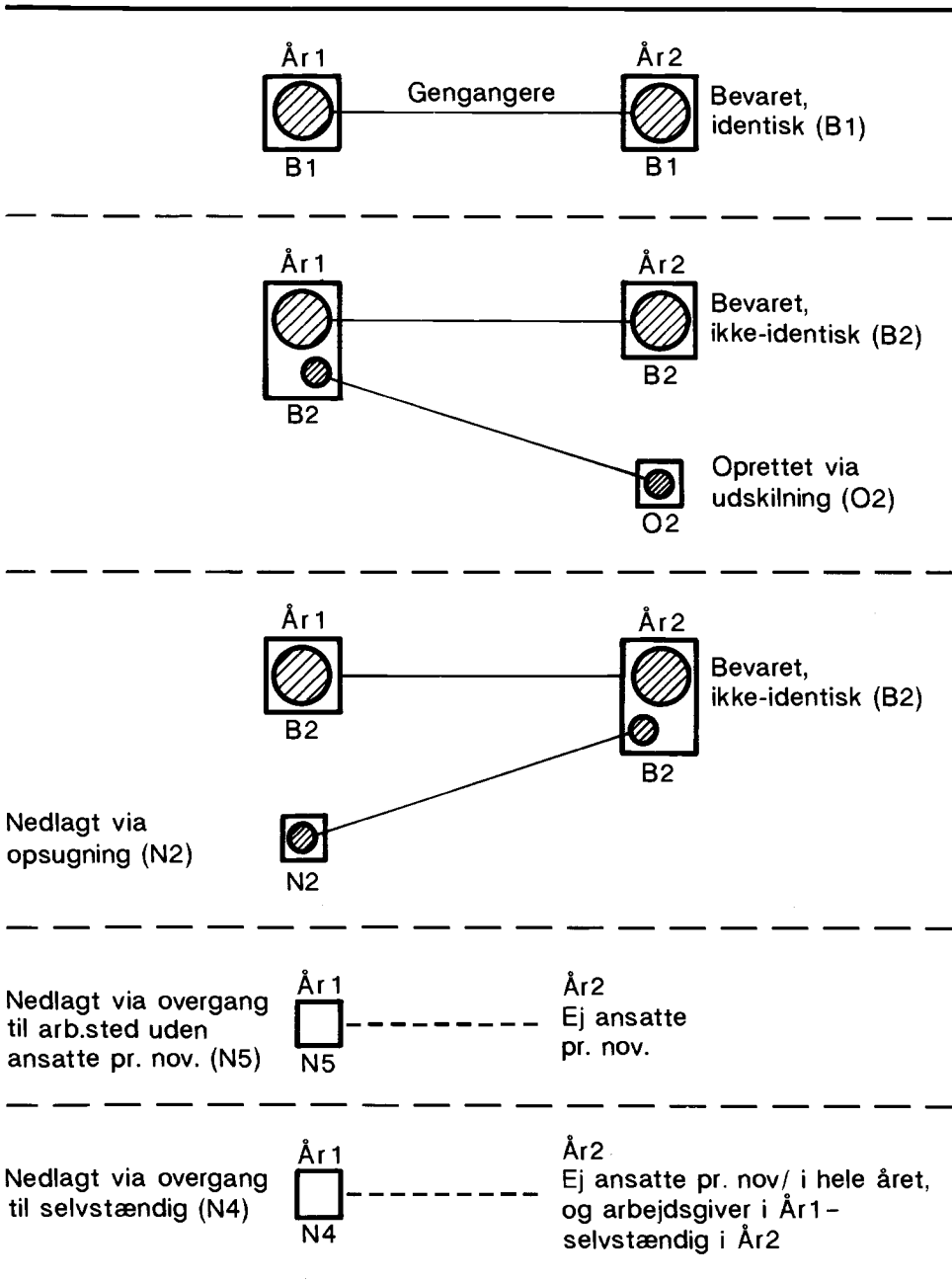
1. Fremad i tid, til november året efter (bevaret/nedlagt):

B1	Bevaret, identisk.
B2	Bevaret, ikke-identisk. Dvs. et arbejdssted hvorfra der er udskilt personer til et andet (oprettet) arbejdssted - eller hvortil der er "opsuget" personer fra et andet (nedlagt) arbejdssted. Antal udskilte/opsugede personer skal udgøre mindst 2 personer (ved adresseskift).
N1	<i>Nedlagt.</i>
N2	Nedlagt via <i>opsugning</i> i et andet arbejdssted. Hvis mindst 30 pct. af de ansatte overgår til det andet arbejdssted, og mindst 2 personer (ved adresseskift).
N3	Nedlagt via " <i>overdragelse</i> " af lokaler/bygninger (til et nyt arbejdssted i samme branche).
N4	Nedlagt via overgang <i>til selvstændig</i> uden ansatte pr. november (i samme branche).
N5	Nedlagt via overgang <i>til arbejdssted uden ansatte</i> pr. november (i samme branche).

2. Tilbage i tid, fra november året før (bevaret/oprettet):

B1	Bevaret, identisk.
B2	Bevaret, ikke-identisk (jf. ovenfor under pkt. 1).
O1	<i>Nyoprettet.</i>
O2	Oprettet via <i>udskilning</i> fra et arbejdssted. Hvis mindst 30 pct. af de ansatte kommer fra samme arbejdssted, og mindst 2 personer (ved adresseskift).
O3	Oprettet via " <i>overtagelse</i> " af lokaler/bygninger (fra et nedlagt arbejdssted i samme branche).
O4	Oprettet via overgang <i>fra selvstændig</i> uden ansatte pr. november (i samme branche).
O5	Oprettet via overgang <i>fra arbejdssted uden ansatte</i> pr. november (i samme branche).

Figur 4. Forskellige typer af bevarede, nedlagte og oprettede arbejdssteder



Til gengæld kan man umiddelbart foretage analyser af personaleomsætning og beskæftigelsesændringer med udgangspunkt i kategorien af bevarede, identiske arbejdssteder (B1). Hvis man herunder kun er interesseret i at se på gruppen af hovedbeskæftigede lønmodtagere, kan man endvidere se bort fra arbejdssteder, der kun har bibeskæftigede blandt de ansatte i det ene eller i begge af årene. Arbejdssteder med ansatte afgrænses som arbejdssteder med ansatte overhovedet pr. november, dvs. at der udelukkende kan være tale om ansatte med bibeskæftigelse på arbejdsstedet.

For begge kategorier af bevarede arbejdssteder gælder det, at regelsættet vedrørende "samme" arbejdssted skal være opfyldt (jf. kapitel 1, afsnit 1.4).

Nedlagte arbejdssteder

Der indgår fem forskellige kategorier af nedlagte arbejdssteder. Med "nedlagte" menes arbejdssteder, der afgår fra populationen af arbejdssteder med ansatte pr. november (jf. afsnit 4.1).

Kategorien "*nedlagt af typen N1*" består af "rene" nedlæggelser. Det kan være arbejdssteder, som ikke findes året efter på identifikationsnummeret (ejeren), eller som ikke opfylder arbejdsstyrkekriteriet med hensyn til antal gengangere blandt de ansatte. Sidstnævnte vil konkret sige, at mindre end 30 procent af de ansatte genfindes på et fælles arbejdssted året efter - svarende til, at de ansatte er blevet spredt "for alle vinde".

Hvis en vis mindste del af de ansatte (mindst 30 pct.) derimod overgår samlet til et andet, bestående arbejdssted, vil der være tale om en *nedlæggelse "via opsugning"*, dvs. typen *N2*. Her er der dog tillige et krav om, at det mindst skal dreje sig om to (hovedbeskæftigede) personer, som overgår til det andet arbejdssted. I modsat fald kan der være tale om almindelig mobilitet blandt enkeltpersoner².

² Sidestykket til en nedlæggelse via opsugning er et arbejdssted af typen bevaret, ikke-identisk (B2). De udgør under ét en "sammenlægning". Når vi opererer med begrebet sammenlægninger (og analogt opdelinger) af arbejdssteder, beror det således udelukkende på *bevægelser af "klumper" af personer* (hvor en "klump" af personer kan bestå af blot to personer). Vi kan strengt taget ikke vide, om der foreligger organisationsmæssige beslutninger, herunder af typen opkøb. Det er grunden til, at vi bruger udtrykket nedlagt via "opsugning", dvs. opsugning af personer (og analogt oprettet via "udskilning"). Når det drejer sig om en oprettelse af et nyt arbejdssted via udskilning, må det dog antages, at man er på mere sikker grund. Her er der rent faktisk tale om en ny enhed.

Vi skal herefter springe til kategorien N5: *Nedlagt via overgang til et arbejdssted uden ansatte* pr. november, dvs. i november i det følgende år. Her er der tale om et arbejdssted, der opfylder regelsættet vedrørende "samme" arbejdssted (nemlig regel 1, jf. kap. 1) - men der findes ingen ansatte pr. november i det følgende år. Enheden vil således fortsat optræde i bruttopopulationen af arbejdssteder, så i en vis forstand er arbejdsstedet bevaret, og man kan vælge at betragte det som sådan i givne analyser.

Når vi alligevel klassificerer tilfældet som en "nedlæggelse", er det ikke kun på grund af, at der er tale om en afgang fra delpopulationen af arbejdssteder med ansatte pr. november. Det skyldes også, at der meget vel kan være tale om en faktisk nedlæggelse, dvs. et ophør af erhvervsaktiviteten, nemlig i løbet af det følgende år *før* ultimo november (jf. også afsnit 4.1). Men der kan også være tale om en "dvaletilstand" eller passiv tilstand i relation til forekomst af ansatte (pr. november), evt. i form af en overgang til udelukkende sæsonmæssig beskæftigelse. Hvilken af disse muligheder, der er tilfældet, vil først vise sig endnu et år frem, dvs. i det tredje år.

Kategorien N4, *"nedlagt via overgang til selvstændig"*, er en variant heraf. Der er tale om en supplerende klassifikation, der fremkommer via en omklassificering af kategorierne N5 og N1. Et nedlagt arbejdssted af typen N5 eller N1 kan således blive ændret til typen N4, overgang til selvstændig i det følgende år. En forudsætning herfor er dog, at det drejer sig om fx enkeltmandsejede arbejdssteder, hvor indehaveren overgår fra "arbejdsgiver" til "selvstændig".

Herunder kan der være tale om to forskellige kategorier. Det kan dels dreje sig om en selvstændig med ansatte i løbet af det følgende år uden for november, dvs. et arbejdssted i delpopulationen af arbejdssteder uden ansatte pr. november (disse opstår via N5)³. Og dels kan det dreje sig om selvstændige, der overhovedet ikke har ansatte i det følgende år (disse opstår via N1). I sidstnævnte tilfælde er der ikke tale om et "arbejdssted" året efter; enheden udgår dermed af bruttopopulationen af arbejdssteder, og arbejdsstedet er "endeligt nedlagt". Antalsmæssigt er der tale om betydelige størrelser i forbindelse med denne omklassificering, jf. bilagstabel 1 (og bilagstabel 2 for så vidt angår oprettelse via overgang fra selvstændig).

³ I denne forbindelse kan kategorien N5 også blive ændret til N1, dvs. at det afgøres allerede her, at der faktisk er tale om en nedlæggelse af arbejdsstedet i løbet af det følgende år. Det sker, hvis den tidligere arbejdsgiver ikke er selvstændig pr. november året efter, men fx lønmodtager eller pensionist.

Endelig findes kategorien *N3*, nedlagt via "overdragelse" af lokaler/bygninger til et oprettet arbejdssted i samme branche. Dette skal indikere, at et nedlagt arbejdssted (af typen *N1* eller *N5*) kan have overdraget lokaler mv. til et oprettet arbejdssted. Det oprettede arbejdssted på vedkommende lokalitet vil samtidig blive klassificeret som oprettet via overtagelse af lokaler, dvs. typen *O3*. Et eksempel herpå kan være et generationsskifte af en landbrugsejendom. Det er et temmelig begrænset antal tilfælde, som optræder i disse kategorier, jf. bilagstabellerne.

Oprettede arbejdssteder

Som for nedlæggelsernes vedkommende findes der også fem forskellige kategorier af oprettelser, jf. den nederste del af *oversigt 1* vedrørende identiteten set tilbage i tid.

Indholdet i de fem kategorier er *analogt til kategorierne af nedlæggelser*. Der skal således kun knyttes nogle enkelte bemærkninger til den "omvendte" situation med oprettelser.

En klassifikation som "*nyoprettet*" (*O1*) vil betyde, at de ansatte er kommet fra mange forskellige "steder" (fra forskellige arbejdssteder eller fra arbejdsløshed mv.).

For så vidt angår "*oprettelse via udskilning*" fra et arbejdssted (*O2*) skal det nævnes, at det bl.a. kan dreje sig om såkaldte "spin-offs" fra større arbejdssteder. Herunder gælder det, at personer, der overgår fra lønmodtager (på det tidligere arbejdssted) til arbejdsgiver (indehaver) på det udskilte, nye arbejdssted, medregnes i den gruppe af personer, der overgår til det nye arbejdssted. Herved medregnes de tilfælde, hvor det drejer sig om en overgang af to personer, hvoraf den ene bliver indehaver af det udskilte arbejdssted.

Med hensyn til en *oprettelse via overgang fra selvstændig* (*O4*) gælder det analogt til nedlæggelse, at der dels kan være tale om en overgang fra en selvstændig med ansatte i løbet af det foregående år, og dels om en overgang fra en selvstændig helt uden ansatte. Sidstnævnte betyder, at der foreligger et nyt "arbejdssted", dvs. en tilgang til bruttopopulationen af arbejdssteder.

4.3. Klassifikation af arbejdssteder uden ansatte pr. november

Populationen udgør omkring 20 pct. af det samlede antal arbejdssteder, jf. bilagstabel 1 og 2.

Klassifikationen af identitetsændringen fremgår af *oversigt 2*. Kategorierne fremkommer som noget, der afledes fra fastlæggelsen af identiteten for arbejdssteder med ansatte.

For det første er der tale om en kategori af arbejdssteder, der forbliver som bevarede inden for populationen af arbejdssteder uden ansatte pr. november, benævnt "SM" (for "samme"). De udgør over en trediedel af denne delpopulation, og kan opfattes som den gruppe af arbejdssteder, der har sæsonpræget beskæftigelse (i det mindste i to på hinanden følgende år).

For det andet findes tre kategorier af *afgang* fra arbejdssteder uden ansatte pr. november (A1-A3). For den første kategori drejer det sig om en afgang fra hele populationen af arbejdssteder, dvs. en endelig nedlæggelse. I den anden kategori er der tale om en afgang (overgang) som følge af, at arbejdsstedet får ansatte pr. november i det følgende år. Herunder kan der være tale om, at arbejdsstedet oprettes i slutningen af året, efter ultimo november. Denne kategori A2 er modstykket til kategorien O5 blandt arbejdssteder med ansatte pr. november⁴. Endelig er der en noget speciel kategori A3, som er et "bi-produkt" fra fastlæggelsen af identiteten for arbejdssteder med ansatte. I forbindelse med ejerskifte af allerede bestående (og bevarede) enheder får nogle arbejdssteder uden ansatte pr. november en "fiktiv" karakter, og de betragtes som nedlagte⁵.

På tilsvarende måde findes tre kategorier af tilgang til populationen af arbejdssteder uden ansatte pr. november, jf. nederst i *oversigt 2*.

⁴ Af tekniske grunde er der også i kategorien A2 inkluderet arbejdssteder, der i lighed med den første kategori A1 bliver endelig nedlagt, jf. noten i *oversigten*.

⁵ Når ejerskiftet sker i slutningen af året (efter ultimo november), optræder der et arbejdssted - et nummer - uden ansatte pr. november. Dette er en "fiktiv" enhed, idet den reelle enhed er det allerede bestående arbejdssted med ansatte pr. november (med et andet nummer). Der er tale om et formelt ejerskifte, som bliver negligeret, jf. nærmere i rapporten "Virksomheders identitet over tid", afsnit 3.3.2.

Oversigt 2 Klassifikation af arbejdssteder uden ansatte pr. november i året: Identiteten set fremad i tid (samme kategori/afgang), hhv. set tilbage i tid (samme kategori/tilgang).

1. Fremad i tid, til november året efter (samme/afgang):

SM	<i>Uændret, uden ansatte pr. november også i det følgende år (bevaret).</i>
A1	Afgang fra bruttopopulationen af arbejdssteder (ingen forbindelse til et arbejdssted året efter).
A2	Afgang via overgang til arbejdssted med ansatte pr. nov. (i samme branche) ⁶ .
A3	Afgang af en "fiktiv" enhed (jf. teksten).

2. Tilbage i tid, fra november året før (samme/tilgang):

SM	<i>Uændret, uden ansatte pr. november også i året før (bevaret).</i>
T1	Tilgang til bruttopopulationen af arbejdssteder (ingen forbindelse til et arbejdssted året før).
T2	Tilgang via overgang fra arbejdssted med ansatte pr. nov. (i samme branche) ⁷ .
T3	Tilgang af en "fiktiv" enhed (jf. teksten).

⁶ Inkluderer af tekniske grunde overgange til oprettede arbejdssteder af typen O1, evt. O3. Dvs. at arbejdssteds-løbenummeret i disse tilfælde ophører (nedlægges endeligt).

⁷ Inkluderer af tekniske grunde overgange fra nedlagte arbejdssteder af typen N1, evt. N3. I så fald er der tale om et nyt arbejdssted (med et nyt arbejdssteds-løbenummer).

4.4. Endeligt nedlagte og oprettede arbejdssteder i forhold til den samlede population af arbejdssteder

Formålet med dette afsnit er at præcisere bevægelserne af arbejdssteder ind i og ud af populationen af arbejdssteder *med ansatte*.

For kategorierne N1, N2 og N3 er det klart, at der er tale om en afgang fra hele populationen af arbejdssteder - og for kategorien N5, at der er tale om en overgang til delpopulationen af arbejdssteder uden ansatte pr. november (se figur 3). I sidstnævnte tilfælde er der ikke tale om en "endelig" nedlæggelse af et arbejdssted. Det kommer også til udtryk i, at arbejdssteds-løbenummeret (LBNR) bliver bevaret for et arbejdssted af typen N5.

Med hensyn til kategorien N4, dvs. en overgang til selvstændig, foreligger der imidlertid af registertekniske grunde en komplikation. Det skyldes, at selvstændige helt uden ansatte ikke indgår i "arbejdssteder" (og de har ikke noget arbejdsgivernummer i de anvendte registre).

Af denne grund må de tilfælde, hvor der sker overgang til selvstændige helt uden ansatte, klassificeres som *endeligt nedlagte*. Det fremgår af en særlig variabel. Samtidig ophører det pågældende arbejdssteds-løbenummer med at eksistere.

På tilsvarende måde må en overgang *fra* selvstændig helt uden ansatte til et arbejdssted med ansatte pr. november (dvs. en del af kategorien O4), opfattes som en tilgang af et *nyt arbejdssted*. Til forskel fra de øvrige tilfælde i kategorien O4 gælder det her, at oprettelsesåret da er lig med det pågældende eller aktuelle år (oprettelsesåret udgør de 4 første cifre af arbejdssteds-løbenummeret).

Men strengt taget er der ikke tale om en nedlæggelse i betydningen nedlæggelse af en "erhvervsenhed". Det gælder kun i betydningen et "arbejdssted". På den anden side foreligger der faktisk en angivelse af, at der fortsat findes (eller i forvejen var) tale om en erhvervsenhed i bred forstand. Det kan fx udnyttes i analyser af, hvilken rolle henholdsvis bestående og nye virksomheder spiller for ændringer i den samlede beskæftigelse. Her kan overgange til og fra selvstændige helt uden ansatte betragtes som bevarede og eksisterende "erhvervsenheder".

Der foreligger yderligere et forhold i forbindelse med gruppen af selvstændige. Af registreringsmæssige årsager findes der heller ikke et arbejdsgivernummer for selvstændige

med ansatte i løbet af året. Det betyder, at det ikke er muligt at registrere eksakt, hvilken del af arbejdsstederne uden ansatte pr. november der svarer til bestemte personer inden for gruppen af selvstændige⁸.

Det betyder, at man ikke kan opgøre den *totale* population af "erhvervsenheder", dvs. summen af arbejdssteder og selvstændige helt uden ansatte.

Derimod gælder det for arbejdssteder med ansatte pr. november-tidspunktet, at størrelsen af den totale population kendes, og at bevægelserne af arbejdssteder (til- og afgang) samtidig inkluderer skift fra eller til selvstændige helt uden ansatte.

⁸ Af samme grund kan man ikke registrere overgange mellem selvstændige helt uden ansatte og populationen af arbejdssteder uden ansatte pr. november.

Bilagstabel 1 Identiteten af arbejdssteder pr. november 1983 i det følgende år pr. november 1984 (identiteten set fremad i tid: bevaret/nedlagt).

1983 → ?		Antal	Pct. af arb. steder med ansatte	Pct. af alle arb. steder
I. Arbejdssteder med ansatte pr. nov. 83:				
			--- pct. ---	
B1.	Bevaret, identisk	156.294	84,8	69,1
B2.	Bevaret, ikke-identisk	3.502	1,9	1,5
<hr/>				
N1.	Nedlagt	9.897	5,4	4,4
N2.	Nedlagt via opslugning	1.777	1,0	0,8
N3.	Nedlagt via "overdragelse"	106	0,1	0,0
N4.	Nedlagt/overgang til selvstændig ⁹	7.514	4,1	3,3
N5.	Nedlagt/overgang til ej ansatte	5.307	2,9	2,3
<hr/>				
I alt, arbejdssteder med ansatte		184.397	100,0	81,5
<hr/>				
II. Arbejdssteder uden ansatte pr. nov. 83:				
SM.	Uændret ej ansatte (bevaret)	15.129	•	6,7
<hr/>				
A1.	Afgang (nedlagt)	17.675	•	7,8
A2.	Afgang til arb. sted med ansatte ¹⁰	8.966	•	4,0
A3.	Afgang ("fiktiv" enhed)	113	•	0,0
<hr/>				
I alt, arbejdssteder uden ansatte		41.883	•	18,5
<hr/>				
Arbejdssteder i alt (I + II)		226.280	•	100,0

⁹ Herunder overgang til selvstændig uden ansatte i hele det følgende år 1984.

¹⁰ Herunder til oprettede arbejdssteder af typen O1, evt. O3 (via procedure 61-3/60-4 i modul 1).

Bilagstabel 2 Identiteten af arbejdssteder pr. november 1984 set i forhold til året før pr. november 1983 (identiteten set tilbage i tid: bevaret/oprettet).

? —> 1984		Antal	Pct. af arb. steder med ansatte	Pct. af alle arb. steder
I. Arbejdssteder med ansatte pr. nov. 84:				
				---- pct. ----
B1.	Bevaret, identisk	156.294	83,0	67,9
B2.	Bevaret, ikke-identisk	3.502	1,9	1,5

O1.	Nyoprettet	14.200	7,5	6,2
O2.	Oprettet via udskilning	3.410	1,8	1,5
O3.	Oprettet via "overtagelse"	106	0,1	0,0
O4.	Oprettet/overgang fra selvstændig ¹¹	7.895	4,2	3,4
O5.	Oprettet/overgang fra ej ansatte	2.930	1,6	1,3

I alt, arbejdssteder med ansatte		188.337	100,0	81,8
II. Arbejdssteder uden ansatte pr. nov. 84:				
SM.	Uændret ej ansatte (bevaret)	15.129	•	6,6

T1.	Tilgang (oprettet)	10.635	•	4,6
T2.	Tilgang fra arb. sted med ansatte ¹²	15.073	•	6,6
T3.	Tilgang ("fiktiv" enhed)	903	•	0,4

I alt, arbejdssteder uden ansatte		41.740	•	18,2

Arbejdssteder i alt (I + II)		230.077	•	100,0

¹¹ Herunder overgang fra selvstændig uden ansatte i hele det forudgående år 1983.

¹² Herunder fra nedlagte arbejdssteder af typen N1, evt. N3 (via procedure 60-3/60-4 i modul 1).

Bilag 1:

Oversigt over variable i IDA

På de følgende sider findes en oversigt over variabel-indholdet i de enkelte datasæt i IDA-databasen.

Oversigten er ordnet efter hver af de 12 kategorier af datasæt i IDA-databasen, svarende til opdelingen i kapitel 3 og i figur 2. Inden for hvert datasæt er oplysningerne ordnet emnemæssigt svarende til omtalen i kapitel 2. Dette indebærer imidlertid, at der i oversigten i enkelte datasæt indgår variable, der er irrelevante, fordi de ikke er defineret eller ikke har mening for de pågældende enheder.

I oversigten findes først en kort betegnelse for den enkelte variabls indhold og dernæst den forkortelse, der er anvendt for variablen i IDA-sammenhæng. Selve databeskrivelserne med de enkelte variables kodeværdier og deres betydning bringes af pladsmæssige grunde ikke her, men findes tilgængelig på diskette, og kan i givet fald leveres som papirudskrift. Det skal bemærkes, at variable, der indgår i flere datasæt, kan have forskelligt indhold såvel som forskellige koder. Tilsvarende skal man være opmærksom på, at der i enkelte tilfælde kan være sket ændringer i variables definition eller kodeværdier i løbet af perioden 1980-1987.

1A. Personer

Baggrundoplysninger

Nøgle

Personnummer PNR

Baggrundvariable

Køn KON2
 Alder (pr. november) ALDER2
 Alder (pr. ultimo året) ALDERP
 Fødselsstedskode FODREG
 Statsborgerskabskode STATKOD
 Statusvar. for bevægelse i løbet af året ARSSTAT
 Fra/til befolkning STATX

Bopæl

Bopælsadressekode pr. ultimo året BOPLADR2
 GEO-kode 1 for bopæl BOPGEOK1
 GEO-kode 2 for bopæl BOPGEOK2

Familie og samliv

Opvækst-familie

Mors personnummer PNRM
 Fars personnummer PNRF

Egen familie

Civilstand CIVST2
 Ægtefælles personnummer AGTEPNR
 Samlivsindikator SAMLIV
 Fødselsindikator BABY
 Børne-indikator BORN

IDA-familie

Samboerens personnummer SAMPNR
 Samboers alder SAMALD
 Samboers statsborgerskab SAMSTAT
 Samboers fødested SAMFOD
 IDA-familienummer FAMNIDA

1A. Personer fortsat

IDA-familietype	FAMTIDA
IDA-familiestatus	FAMSIDA
Antal spædbørn (0-2 år) i IDA-familien	ANI02
Antal småbørn (3-6 år) i IDA-familien	ANI36
Antal skolebørn (7-9 år) i IDA-familien	ANI79
Antal skolebørn (10-14 år) i IDA-familien	ANI1014
Antal børn 15-17 år i IDA-familien	ANI1517
Antal børn 18-19 år i IDA-familien	ANI1819
Antal børn 20-25 år i IDA-familien	ANI2025
Antal børn 0-17 år i IDA-familien	ANI017
Antal børn 0-25 år i IDA-familien	ANI025
Antal voksne i IDA-familie	VOKSI
Antal personer ialt i IDA-familie	PERSI
Yngste barns alder i IDA-familien	YNGSTI
Ældste barn i IDA-familien	ALDSTI
Yngste barns fødselsdato i IDA-familie	YFDATOI
Flerfødsel i IDA-familie	FLERFI

B-familie

B-familienummer	FAMNB
B-familietype	FAMTB
B-familiestatus	FAMSB
Antal spædbørn (0-2 år) i B-familien	ANB02
Antal småbørn (3-6 år) i B-familien	ANB36
Antal skolebørn 7-14 år i B-familien	ANB714
Antal børn 15-17 år i B-familien	ANB1517
Antal børn 18-19 år i B-familien	ANB1819
Antal børn 20-25 år i B-familien	ANB2025
Antal børn 0-17 år i B-familien	ANB017
Antal børn 0-25 år i B-familien	ANB025
Antal voksne i B-familien	VOKSB
Antal personer ialt i B-familien	PERSB
Yngste barns alder i B-familien	YNGSTB
Ældste barns alder i B-familien	ALDSTB

Husstand

Husstandstype (baseret på B-familien)	HUSTT
---------------------------------------	-------

Uddannelse**Igangværende uddannelse**

Under uddannelse	UDDSOEG2
Igangværende uddannelse	IGUDD2

1A. Personer *fortsat*

Igangværende uddannelse omkodet, 1 ciffer	IG1UDD
Løn under uddannelse	UDDLOEN2
Heltid/deltids uddannelse	UDDHDTI2
Helårs-elev	UDDHELA2
Uddannelses-institutionsadresse	UDDADR2
Uddannelsesadresse GEO-kode 1	UDDGEOK1
Uddannelsesadresse GEO-kode 2	UDDGEOK2
Pendlingsafstand (KM) til uddannelsested	UAFST

Almen uddannelse

Senest afsluttede almenuddannelse	ALMENUMD2
-----------------------------------	-----------

Erhvervsuddannelse

Højeste fuldførte erhvervsuddannelse	ERHVUDD2
Højeste fuldførte erhvervsuddannelse, 1 ciffer	FLDF1UD
AMU-kursus som højeste erhvervsuddannelse	KURSUS
Højeste uddannelse (inkl. AMU-kurser)	HOJUDD

AMU-kurser

AMU specialarbejderuddannelser retning	AMUSPUD
AMU specialarbejderuddannelser antal	AMUSPAN
AMU kurser for faglærte retning	AMUFAUD
AMU kurser for faglærte antal	AMUFAAN

Historiske oplysninger vedr. uddannelse

År ophør almenuddannelse	ALMUDAF
Erhvervsuddannelse start (år)	ERHUDST
Erhvervsuddannelse ophør/afslutning (år)	ERHUDAF

Beskæftigelse**Status i november**

Branchekode primær arbejdssted	PBRANCH2
Ejerforholdskode for sekundær arbejdssted	SEJERKO2
Branchekode for sekundær arbejdssted	SBRANCH2
Sekundær arbejdsstedes adressekode	SARBADR2
Aktuelt arbejdsomfang	NOOMFAN2
Pensionsmarkering (ja/nej)	PENSMAR2

1A. Personer fortsat

Andre ansættelser i november

Supplerende november-ansættelser (antal)	NSUP
Supplerende løn i november-ansættelse (sum)	NSUPSUM

Ansættelser i året

Ej-november arbejdsgiversnummer	ARBGNR3
Ej-november arbejdsstedskode	ARBSTK3
Dage ansat i ej-november-ansættelse	ANDAGE3
Ej-november formålskode	FORMAL3
Antal supplerende ej-november ansættelser	EJNOV
Ej-novemberansættelser lønsum	EJNOVSUM

Årsoplysninger vedr. beskæftigelse

Stillingskode/DFK	DFKSTIL2
Heltid-deltid på arbejdsmarkedet	HDTID2
Stilling som kommuneansat	KOMMST
Stilling som statsansat	STATST
År-DSejer (lønmotagere i året)	AEJERKO2
År-Branche	ABRANCH2
Socio-økonomisk gruppe (i året)	ASOCIO2
Samlet løn fra oplysningssedler	SLON
ATP-beløb i året	ATPXX

Historiske oplysninger vedr. beskæftigelse

Første år på arbejdsmarkedet	STARTAR
Seneste start på arbejdsmarkedet	SENSTAR
Seneste afslutning på arbejdsmarkedet	SENAFAR
Samlet ATP ultimo 1979	SATP79
Erhvervs erfaring ultimo 1979	ERHVER79
Erhvervs erfaring fra 1980 i 1000	ERHVER
Antal år som lønmotager	ATPAR
Alder ved pensionering (offentlig)	PENSALD

Arbejdsløshed

Statusoplysninger vedr. ledighed

A-kasse	AKASKOD
Forsikringskategori i CRAM	FORSKAT2
Ledighedsgrad i uge 48 (ultimo november)	ULEDGRA2
Ledighedsårsag i uge 48	LEDARSA2

1A. Personer *fortsat***Ledighedsgrader**

Ledighed i uge 1	LEDUG1
Ledighed i uge 52	LEDUG52
Årsledighedsgrad	ARLEDGR
Ledighedsgrad i arbejdsår (ekskl. ferie)	ARBLED
Ledighedsgrad i ledighedsperioder	GRADPER
Ledighedsgrad i 1. kvartal	LED1GR
Ledighedsgrad i 2. kvartal	LED2GR
Ledighedsgrad i 3. kvartal	LED3GR
Ledighedsgrad i 4. kvartal	LED4GR

Ledighedsperioder

(Numre angiver beregning uden følgende ledighedårsag:

4=AF-ferie som ansat, 5=AF-ferie som ledig)

Ledighedsperioder i året (4,5)	LEDPERI
Perioder fuld ledighed (4,5)	LEDPF
Perioder delvis ledighed (4,5)	LEDPD
Ledighedsperioder i året (5)	LEDPERA
Perioder fuld ledighed (5)	LEDPFA
Perioder delvis ledighed (5)	LEDPDA

Ledighedsskift

Skift af ledighedstilstand (4,5)	LEDSKIB
Skift af ledighedstilstand (4)	LEDSKIF

Antal ledighedsuger

Uger fuld ledighed	LEDFULD
Uger fuldt ledig vejrlig	LEFUVE
Uger fuldt ledig arbejdsfordeling	LEFUARB
Uger fuldt ledig AF-ferie ansat	LEFUANS
Uger fuldt ledig AF-ferie ledig	LEFULED
Uger fuldt ledig AF-ferie og andet	LEFUAND
Uger fuldt ledig bistandsydelse	LEFUBIS
Uger fuldt ledig uoplyst	LEFUUOP
Uger delvis ledig	LEDEL
Uger delvis ledig vejrlig	LEDEVE
Uger delvis ledig arbejdsfordeling	LEDEARB
Uger delvis ledig AF-ferie ansat	LEDEANS
Uger delvis ledig AF-ferie ledig	LEDELED
Uger delvis ledig AF-ferie og andet	LEDEAND
Uger delvis ledig bistandsydelse	LEDEBIS

1A. Personer fortsat

Uger delvis ledig uoplyst	LEDEUOP
Uger uden ledighed	UGENUL

Historiske oplysninger vedr. ledighed

Seneste startår A-kasse medlem	AKASST
Seneste år som medlem af A-kasse medlem	AKASAF
Første ledighed (årstal)	LEDAR
Forsikret/ledig antal år (fra 1980)	POTLEAR
Sum af ledighedsgrader (fra 1980)	SUMGRAD

Indkomst**Samlet indkomst**

Bruttoindkomst	BRINDK2
Skattepligtig indkomst	SKPLIND2

Primærindkomster

Løn	LONIND
Arbejdsgiveradministrerede pensionsindbetalinger	PENSORD
Overskud af virksomhed	OVSKVI
Overskud af fast ejendom	OVSKEJ
Medarbejdende ægtefælle indkomst/fradrag	FRAMARB

Overførselsindkomster

SU-stipendium	STIP
Arbejdsløshedsunderstøttelse	ARBLP
A-sygedagpenge	ADAGP
B-sygedagpenge	BDAGP
Social bistand	SOCBIST
Efterløn	EFTLON
Offentlig pension	OFFPENS
Pension fra ATP	ATPPENS

Formueoplysninger

Aktiver	AKTIV
Passiver	PASSIV
Formue	FORMUE

Ansættelse/beskæftigelse i november**Ansættelsens art og karakter**

Primær arbejdsstilling	PSTILL2
------------------------	---------

1A. Personer *fortsat*

Sekundær arbejdsstilling
Tilknytningen til prim. arbejdssted

SSTILL2
TILKNYT

Løn og ATP

Ej-november ATP
Løn i ej-november ansættelser

ATP3
LON3

1B. Personændringer

Baggrundsoplysninger

Nøgle

Personnummer

PNR

Baggrundvariable

Vandringer i løbet af året

VANDRE

Dødsdato

DATODOD

Indvandringsdato

DATOIND

Indvandringslandskode

INDLAND

Udvandringsdato

DATOUD

Udvandringslandskode

UDLAND

Fra/til befolkning

STATX

Bopæl

Dato for tilflytning til adresse

ADRDTO

Ændring af bopælsadresse

BOPLX

Afstand i kilometer bopælskommune år 1 og år 2

FLYTAFS

Familie og samliv

Født i Danmark i løbet af året

FODT

Mors personnummer

PNRM

Civilstandsændringsdato

CIVDTO

Uddannelse

Status i november

Ændring af uddannelsessituation

UDDX

Beskæftigelse

Status i november

Statusændring i forhold til arbejdsmarked

SOCIOX

Heltid-deltidsændring

HDTIDX

Til (og fra) pensionering

PENSX

Arbejdsløshed

Statusoplysninger vedr. ledighed

Ændring af A-kasse

AKASSEX

1B. Personændringer *fortsat*

Fra og til ledighed

LEDIGX

Ansættelse/beskæftigelse i november**Ansættelsens art og karakter**

Primær arbejdsstilling - år 1

PSTILL1

Primær arbejdsstilling - år 2

PSTILL2

Sekundær arbejdsstilling - år 1

SSTILL1

Sekundær arbejdsstilling - år 2

SSTILL2

2A. Hovedbeskæftigede lønmodtagere

Baggrundsoplysninger

Nøgle

Personnummer PNR

Baggrundsvariable

Køn	KON
Alder i år (pr. november)	ALDER
Igangværende uddannelse omkodet, 1 ciffer	IG1UDD
Højeste fuldførte erhvervuddannelse, 1 ciffer	FLDF1UD
AMU-kursus som højeste erhvervsuddannelse	KURSUS
"Personlig" branche	PERSBRCH
"Personlig" arbejdsstedsadresse	PERSADR
Stillingskode/DFK	DFKSTIL
A-kasse	AKASSE
Ledighedsgrad i arbejdsår (ekskl. ferie)	ARBLED
Ledighedsperioder i året (4,5)	LEDPERI

Ansættelse/beskæftigelse i november

Ansættelsens art og karakter

Antal dage ansat	ANSDAGE
Lønmodtagere eller arbejdsgivere	TYPE
Arbejdsstilling	STILL
Skiftende arbejdssted (kun primær)	SKIFTARB
Arbejdsomfang for lønmodtagere	ARBOMF
Kvalitet af oplysninger om persons arbejdssted	ARBSTOPL
Formålskode for kommunalt ansatte	FORMAL
Jobtype primær ansættelse	PJOB
Afstand fra bopæl til arbejdssted	KMAFST
Tilknytningen til primær arbejdssted	TILKNYT

Løn og ATP

Timeløn i ansættelsen	TIMELON
Kvalitet af timelønsskøn	TLONKVAL
Løn i november ansættelse	NOVLON
ATP i november ansættelse	NOVATP

2A. Hovedbeskæftigede lønmodtagere *fortsat***Ændring i ansættelse/historiske oplysninger**

Lønmodtager året efter

LONMFREM

Lønmodtager året før

LONMTILB

Ansættelsesændring frem til året efter

ANSXFREM

Ansættelsesændring i forh. til året før

ANSXTILB

År for ansættelse på arbejdssted

ANSAAR

Arbejdssteder/firmaer**Nøgle**

Arbejdssteds-løbenummer

LBNR

År-parameter

AARPRM

Arbejdsgivernr.

ARBGNR

2B. Bibeskæftigede lønmodtagere

Baggrundsoplysninger

Nøgle

Personnummer PNR

BaggrundsvARIABLE

Køn KON
Alder i år (pr. november) ALDER

Ansættelse/beskæftigelse i november

Ansættelsens art og karakter

Antal dage ansat ANSDAGE
Lønmodtagere eller arbejdsgivere TYPE
Primær arbejdsstilling PSTILL
Arbejdsomfang for lønmodtagere ARBOMF
Kvalitet af oplysninger om persons arbejdssted ARBSTOPL
Formålskode for kommunalt ansatte FORMAL
Afstand fra bopæl til arbejdssted KMAFST
Sekundær arbejdsstilling SSTILL
Jobtype sekundær ansættelse SJOB

Løn og ATP

Timeløn i ansættelsen TIMELON
Kvalitet af timelønsskøn TLONKVAL
Løn i november ansættelse NOVLOK
ATP i november ansættelse NOVATP

Ændring i ansættelse

Ændring i bibeskæftigelse til året efter BIXFRÉM
Ændring i bibeskæftigelse fra året før BIXTILB

Arbejdssteder/firmaer

Nøgle

Arbejdssteds-løbenummer LBNR
År-parameter AARPRM
Arbejdsgivernr. ARBGNR

2C. Arbejdsgivere

Variable mærket med *) er irrelevante - eventuelt ikke definerede - for arbejdsgivere. Oplysningerne er alligevel listet her, da dette datasæt har samme struktur som datasættet for de hovedbeskæftigede lønmodtagere (2A).

Baggrundsoplysninger

Nøgle

Personnummer PNR

Baggrundsvariable

Køn	KON
Alder i år (pr. november)	ALDER
Igangværende uddannelse omkodet, 1 ciffer	IGIUDD
Højeste fuldførte erhvervsuddannelse, 1 ciffer	FLDF1UD
AMU-kursus som højeste erhvervsuddannelse	KURSUS
"Personlig" branche	PERSBRCH
"Personlig" arbejdsstedsadresse	PERSADR
Stillingskode/DFK	DFKSTIL
A-kasse	AKASSE
Ledighedsgrad i arbejdsår (ekskl. ferie)	ARbled
Ledighedsperioder i året (4,5)	LEDPERI

Ansættelse/beskæftigelse

Ansættelsens art og karakter

Antal dage ansat *)	ANSDAGE
Lønmodtagere eller arbejdsgivere	TYPE
Arbejdsstilling *)	STILL
Skiftende arbejdssted (kun primær) *)	SKIFTARB
Arbejdsomfang for lønmodtagere *)	ARBOMF
Kvalitet af oplysninger om persons arbejdssted	ARBSTOPL
Formålkode for kommunalt ansatte *)	FORMAL
Jobtype primær ansættelse *)	PJOB
Afstand fra bopæl til arbejdssted	KMAFST
Tilknytningen til primær arbejdssted *)	TILKNYT

Løn og ATP

Timeløn i ansættelsen *)	TIMELON
Kvalitet af timelønsskøn *)	TLONKVAL

2C. Arbejdsgivere fortsat

Løn i november ansættelse *)
 ATP i november ansættelse *)

NOVLON
 NOVATP

Ændring i ansættelse/historiske oplysninger

Lønmodtager året efter
 Lønmodtager året før
 Ansættelses-ændring frem til året efter
 Ansættelses-ændring i forhold til året før
 År for ansættelse på arbejdssted

LONMFREM
 LONMTILB
 ANSXFREM
 ANSXTILB
 ANSAAR

Arbejdssteder/firmaer**Nøgle**

Arbejdssteds-løbenummer
 År-parameter
 Arbejdsgivernr.

LBNR
 AARPRM
 ARBGNR

2D. Tilgangspersoner

Baggrundsplysninger

Nøgle

Personnummer PNR

Bopæl

Ændring af bopælsadresse BOPLX

Ansættelse/beskæftigelse

Ansættelsens art og karakter

Lønmodtagere eller arbejdsgivere TYPE

Ændring i ansættelse

Tilgangspersoners stilling mv. året før	GLSTILL
Tilgangspersoners branchekode året før	GLBR3
Branchekode året før (5-cifret)	GLBR5
Brancheskift ved arbejdsstedsskift (3 cifre)	BRCH3X
Brancheskift ved arbejdsstedsskift (5 cifre)	BRCH5X
Moderarbejdssted for O2-kategori	MODERO2
Ansættelses-ændring i forhold til året før	ANSXTILB

Arbejdssteder/firmaer

Nøgle

Arbejdssteds-løbenummer	LBNR
År-parameter	AARPRM

2E. Afgangspersoner

Baggrundsoplysninger

Nøgle

Personnummer PNR

Bopæl

Ændring af bopælsadresse BOPLX

Ansættelse/beskæftigelse

Ansættelsens art og karakter

Lønmodtagere eller arbejdsgivere TYPE

Ændring i ansættelse

Ny stilling mv. for afgangspersoner	NYSTILL
Branchekode (3-cifret) året efter	NYBR3
Branchekode (5-cifret) året efter	NYBR5
Brancheskift ved arbejdsstedsskift (3 cifre)	BRCH3X
Brancheskift ved arbejdsstedsskift (5 cifre)	BRCH5X
Moderarbejdssted for N2-kategori	MODERN2
Ansættelses-ændring frem til året efter	ANSXFREM

Arbejdssteder/firmaer

Nøgle

Arbejdssteds-løbenummer	LBNR
År-parameter	AARPRM

2F. Selvstændige

Baggrundsoplysninger

Nøgle

Personnummer

PNR

Baggrundsvariable

Køn

KON

Alder i år (pr. november)

ALDER

Igangværende uddannelse omkodet, 1 ciffer

IG1UDD

Højeste fuldførte erhvervuddannelse, 1 ciffer

FLDF1UD

AMU-kursus som højeste erhvervsuddannelse

KURSUS

Beskæftigelse

Ansættelsens art og karakter

Status for selvstændige året efter

SLVXFREM

Status for selvstændige året før

SLVXTILB

Arbejdssteder/firmaer

Nøgle

År-parameter

AARPRM

Baggrundsoplysninger

Arbejdsstedsadresse

ARBADR

Ejerforholdskode

EJERKO

Branchekode

BRANCHE

2G. Medhjælpende ægtefæller

Baggrundsoplysninger

Nøgle

Personnummer

PNR

Baggrundsvariable

Køn

KON

Alder i år (pr. november)

ALDER

Ægtefælles personnummer

AGTEPNR

Igangværende uddannelse omkodet, 1 ciffer

IG1UDD

Højeste fuldførte erhvervuddannelse, 1 ciffer

FLDF1UD

AMU-kursus som højeste erhvervsuddannelse

KURSUS

Beskæftigelse

Status i november

Status for medhjælpende ægtefælle året efter

MEDXFREM

Status for medhjælpende ægtefælle året før

MEDXTILB

Arbejdssteder/firmaer

Nøgle

År-parameter

AARPRM

Baggrundsoplysninger

Arbejdsstedsadresse

ARBADR

Branchekode

BRANCHE

3A. Arbejdssteder

Nøgle

Arbejdssteds-løbenummer	LBNR
Arbejdsstedskode	ARBSTK
År-parameter	AARPRM
Arbejdsgivernr.	ARBGNR
Arbejdsgivernr. for arbejdssted året efter	ARBGNRFR
Arbejdsgivernr. for arbejdssted året før	ARBGNRTI
Arbejdsstedskode for arbejdssted året efter	ARBSTKFR
Arbejdsstedskode for arbejdssted året før	ARBSTKTI
LBNR for relateret arbejdssted året efter	LBNRFREM
LBNR for relateret arbejdssted året før	LBNRTILB

Baggrundsoplysninger

Arbejdssteds adressekode (fra ERHVERV)	ADRESSE
Arbejdsstedets GEO-kode 1	ADRGE01
Arbejdsstedets GEO-kode 2	ADRGE02
Ejerforholdskode	EJERKO
Branchekode	BRANCHE
Branchekode, 3-cifret (SFI) for arbejdssted	BR3
Adresseændring for arbejdssted året efter	ADRXFREM
Adresseændring for arbejdssted i året før	ADRXTILB
Brancheændringer for arbejdssted året efter (3 cif)	BR3XFR
Brancheændringer for arbejdssted året før (3 cif)	BR3XTI
Brancheændringer for arbejdssted året efter (5 cif)	BR5XFR
Brancheændringer for arbejdssted året før (5 cif)	BR5XTI

Antal ansatte og ændring heri

Antal ansatte pr. november	ANTNOV
Antal bibeskæftigede pr. november	ANTNOVBI
Antal årsværk for alle ansættelsesforhold	AARSVRK
Antal ansættelsesforhold i året	ANTAAR
Afgangsrate i efterfølgende årsperiode, pct.	AFGRATE
Andel af "opsugede" blandt ansatte	ANDELOPS
Andel af "udskilte" blandt ansatte	ANDELUDS
Ændring i antal ansatte i efterfølgende år	ANTALXFR
Ændring i antal ansatte året før	ANTALXTI
Ændring i beskæftigelsesomfang efterfølgende år	BESKXFR
Ændring i beskæftigelsesomfang året før	BESKXTI
Tilgangsrate i forudgående årsperiode, pct.	TILGRATE

3A. Arbejdssteder *fortsat*

Personalesammensætning og -omsætning

Andel højere funktionærer pr. arbejdssted	ANDHF
Andel lavere funktionærer pr. arbejdssted	ANDLF
Andel faglærte pr. arbejdssted	ANDFA
Andel ikke-faglærte pr. arbejdssted	ANDIF
Andel kvinder blandt højere funktionærer	KVANDHF
Andel kvinder blandt lavere funktionærer	KVANDLF
Andel kvinder blandt faglærte arbejdere	KVANDFA
Andel kvinder blandt ikke-faglærte arbejdere	KVANDIF
Andel af arbejdere med årsledighed, pct.	ANDLEDAR
Ledighedsperioder blandt ledige arb., gennemsnit	LEDPERG
Årsledighedsgrad blandt ledige arb., gennemsnit	LGRGNSAR

Lønsum og lønniveau

Timeløn for alle hovedbeskæftigelse, gennemsnit	LONGNS
Timeløn for højere funktionærer, gennemsnit	LONGNSHF
Timeløn for lavere funktionærer, gennemsnit	LONGNSLF
Timeløn for faglærte, gennemsnit	LONGNSFA
Timeløn for ufaglærte, gennemsnit	LONGNSIF
Lønsum i året for alle ansættelsesforhold	LONAAR

Ændring i arbejdsstedets identitet

Gengangere på arbejdsstedet pr. året efter	PCTFREM
Gengangere på arbejdsstedet fra året før	PCTILB
Firmaintern/ekstern nedlæggelse via opsugning	FRELFREM
Firmaintern/ekstern oprettelse via udskilning	FRELTILB
Identitet for arbejdssted fremad i tid	IDFREM
Identitet for arbejdssted tilbage i tid	IDTILB
Regel vedr. bevaret arbejdssted året efter	REGELFR
Regel vedr. bevaret arbejdssted året før	REGELTI
Endelig nedlæggelse af arbejdssted	ENDNEDL

Firma-oplysninger

Antal arbejdssteder i firmaet (max. 9)	FILIAL
--	--------

3B. Firmaer

Nøgle

År-parameter	AARPRM
Arbejdsgivernr.	ARBGNR

Firma-oplysninger

Antal arbejdssteder i firma	ANTALARB
Antal bevarede arbejdssteder i firma	ANTARBB
Antal hovedbeskæftigede i firma	FANSH
Antal bibeskæftigede i firma	FANSB
Hovedbeskæftigelse i firma med uoplyst arbejdssted	UOPLANSH
Bibeskæftigelse i firma med uoplyst arbejdssted	UOPLANSB
Branchekode største arbejdssted i firma	STORBRA
Løbenr. største arbejdssted i firma	SLBNR
Identitet tilbage største arbejdssted	SIDTILB
Nummerskift for største arbejdssted i firma	SARBGNRX

3C. Nedlagte arbejdssteder

Nøgle

Arbejdssteds-løbenummer

LBNR

År-parameter

AARPRM

Ændring i arbejdsstedets identitet

Identitet for arbejdssted fremad i tid

IDFREM

Bilag 2:

Oversigt over EDB-systemet til dannelse af IDA-databasen

1. Indledning

I dette afsnit er det hensigten at give et indtryk af det EDB-system, der udgør grundlaget for de endelige datasæt i IDA-databasen. En væsentlig del af arbejdet med IDA's etablering har bestået i at udarbejde dette system, der er meget omfattende. Der kan derfor kun gives en beskrivelse i meget grove træk, men dog forhåbentligt på en sådan måde, at der fås et indtryk af hovedtrækkene i systemet. Den detaljerede beskrivelse af EDB-systemet findes tilgængeligt i Danmarks Statistik, idet der til dette formål er anvendt TIMES¹, der er et dokumentationsværktøj, udarbejdet i Danmarks Statistik.

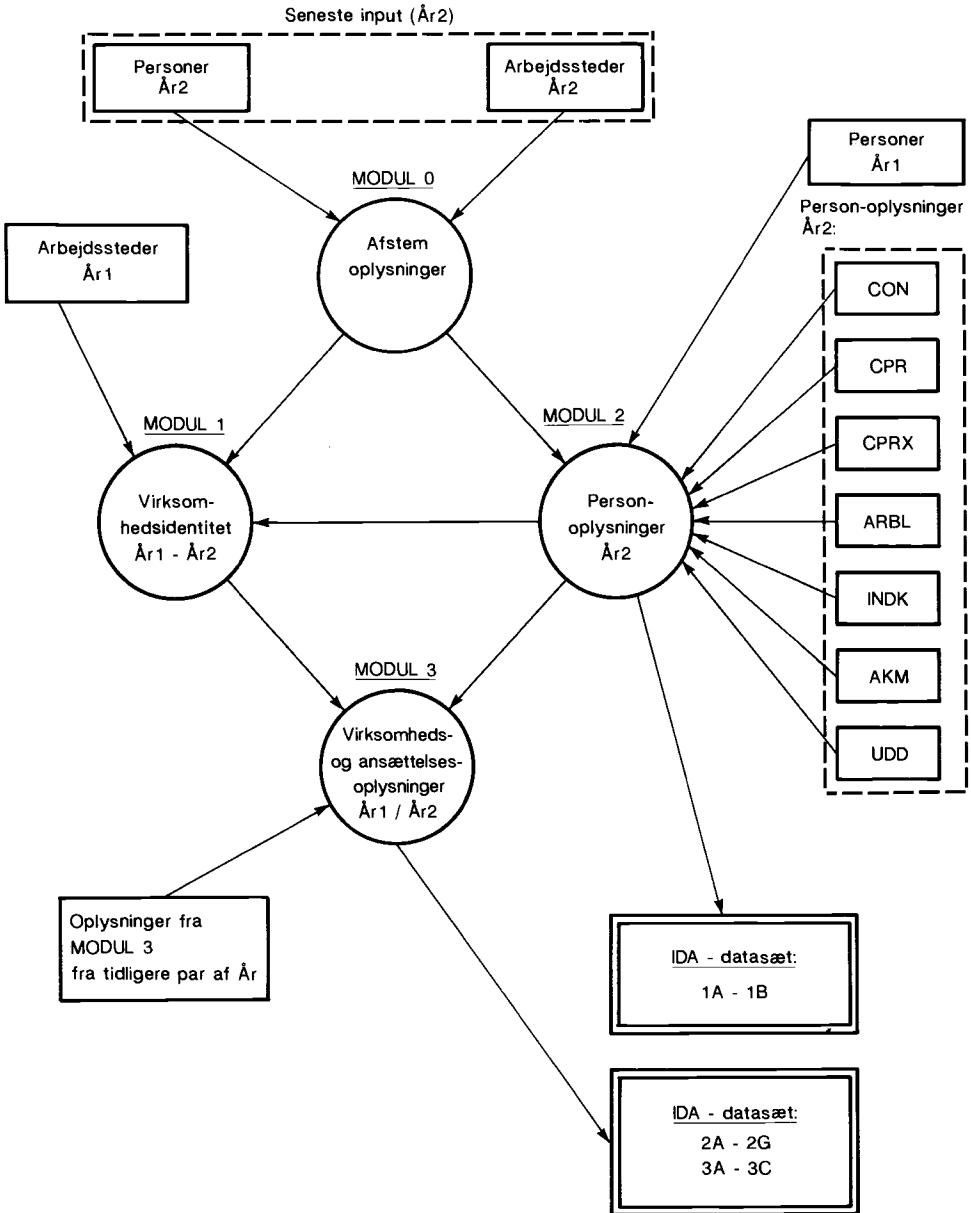
2. EDB-systemet

EDB-systemet til dannelse af IDA-databasen består af *fire hoved-moduler*, som er indbyrdes forbundne, og som har hvert sit formål. Modulerne benævnes med numrene 0, 1, 2 og 3, jf. figur 5. Bagest i bilaget findes en figur for hvert modul, der mere detaljeret viser indholdet i det enkelte modul.

I oversigten over modulerne og data-flowet opereres der med et "år-1" og et "år-2". Ved "år-2" forstås det seneste år, for hvilket der foreligger data (statistikregistre), og som således kan danne grundlag for en ajourføring af IDA-databasen. "År-2" benævnes i det følgende "det aktuelle år". "År-1" er "året før", som inddrages af hensyn til dannelse af data om bevægelser (ændringer) mellem det aktuelle år (år-2) og året før (år-1). Med hen-

¹ Dokumentationen ved hjælp af TIMES inkluderer data-flow-diagrammer, procedurebeskrivelser, record-layouts, databeskrivelser og formatbiblioteker.

Figur 5. Data-flow-diagram over IDA-systemet



Anm. År1: Året for
År2: Seneste År

syn til data for "året før" betyder fremkomsten af det nye, efterfølgende år-2 samtidig, at data om ændringer set "fremad i tid" fra år-1 nu kan dannes².

EDB-systemet bygger således på oplysninger for parvise år. I forbindelse med etableringen af IDA-databasen har der foreligget en serie af parvise år begyndende med året 1980, dvs. 1980-1981, 1981-1982, 1982-1983, osv. Det samlede EDB-system (som vist i oversigten) har således skullet gennemløbes for hvert par af år, i alt 7 gange, idet det seneste år, som indgår ved etableringen af databasen, er 1987. Der har endvidere været inddraget nogle "special-procedurer" vedrørende året 1980, dvs. startåret, for hvilket der ikke (som for de øvrige års vedkommende) har foreligget data om "året før".

De grundlæggende oplysninger i IDA, der knytter personer og virksomheder sammen, stammer fra den Registerbaserede arbejdsstyrkestatistik (RAS) og fra den Registerbaserede arbejdsstedsstatistik.

I den Registerbaserede arbejdsstyrkestatistik er enheden *personer*. Men der findes også for bl.a. lønmodtagerne oplysninger, der angiver hvilket firma og arbejdssted personerne er ansat i, samt oplysninger der nærmere karakteriserer ansættelsesforholdet som fx stilling.

I arbejdsstedsstatistikken er *virksomheder* (arbejdssteder) enheden og oplysningerne heri dannes delvist ud fra oplysninger i den Registerbaserede arbejdsstyrkestatistik.

Oplysninger fra disse to kilder udgør det centrale udgangspunkt for dannelsen af IDA. Herover inddrages andre kilder til beskrivelse af specielt personerne.

3. Formålet med de enkelte moduler

EDB-systemet består som nævnt af 4 moduler. I det følgende skal der gives en meget kortfattet beskrivelse af de enkelte moduler med hovedvægt på formålet med hvert af modulerne.

² Jf. grundstruktur illustreret i figur 1 i kap. 1.

Modul 0 er et mindre, indledende modul, hvor hovedformålet er at justere visse person- og virksomhedsdata, således at der er indbyrdes konsistens mellem disse³. Personoplysningerne anvendes i modul 2, mens virksomhedsoplysninger indgår i modul 1.

I *modul 1* dannes oplysningerne om arbejdsstedernes identitet over tid. Dette baseres dels på data om samtlige ansatte personer i det aktuelle år og året før (fra modul 2), hvori der indgår oplysninger om den enkelte persons arbejdssted i de to år. Og dels baseres det på data om den totale population af arbejdssteder i de to år. For arbejdsstederne i det første år fastlægges identiteten set fremad i tid (til det aktuelle år); for arbejdsstederne i det aktuelle år bestemmes identiteten set tilbage i tid (dvs. fra året før). Modul 1 beskrives detaljeret i rapporten: "Virksomheders identitet over tid. Begrebsdannelse og operationalisering", kap. 3, Arbejdsnotat nr. 29.

I *modul 2* dannes alle oplysningerne om personerne i det aktuelle år via bearbejdning af input fra en række personstatistikregistre (jf. afsnit 4). Ved at sammenholde oplysninger fra det aktuelle år med oplysninger fra året før dannes oplysningerne om ændringer i den enkelte persons baggrundsdata og de oplysninger, der fremkommer ved en summering af data fra de enkelte år. Fra modul 2 fremkommer således datasættene vedr. personer, dvs. datasættene 1A og 1B i IDA-databasen.

Endelig dannes i *modul 3* alle datasæt vedrørende ansættelser/beskæftigede og virksomheder (arbejdssteder og firmaer), dvs. datasættene 2A-2G og 3A-3G i databasen. Det sker via en samkøring af personoplysninger (fra modul 2) og data om arbejdsstedernes identitet over tid (fra modul 1). Heri indgår data for det aktuelle år og året før, idet data om ændringer i ansættelsesforhold og om ændringer i baggrundsdata for ansatte/beskæftigede personer og for arbejdssteder bliver dannet. Herunder inddrages de hidtidige datasæt i databasen vedrørende året før, med henblik på overførsel til det aktuelle år af de "historiske" oplysninger om arbejdssteds-løbenummer og oprettelsesåret (for bevarede arbejdssteder), samt af ansættelsesåret (for uændret eller fortsat ansatte på et arbejdssted). Samtidig dannes også de endelige og permanente datasæt for *året før*, idet data vedrørende ændringer set fremad i tid (dvs. frem til det aktuelle eller nye år) nu kan inddrages. Dette gælder datasættene 2A-2C og 2F-2G vedrørende ansættelser/beskæftigede og datasættet

³ Der overføres oplysninger fra virksomhederne til personerne om bl.a. branche og adressekode. Endvidere optælles antallet af beskæftigede på de enkelte virksomheder ud fra oplysninger i RAS. Endelig fravælges de "arbejdssteder", der består af ansatte med uoplyst arbejdsstedskode inden for et firma (ca. 2000 arbejdssteder).

3A vedrørende arbejdssteder. Det er et udtryk for, at disse datasæt for *det aktuelle år* ikke indeholder oplysninger om ændringer set fremad i tid, idet der (jo) er tale om det seneste år, for hvilket der foreligger oplysninger.

4. Data-flow i EDB-systemet

I det følgende skal kort omtales, på hvilken måde data anvendes i EDB-systemet.

I *modul 0* er input de tidligere omtalte oplysninger fra den Registerbaserede arbejdsstyrkestatistik (personer) og den Registerbaserede arbejdsstedsstatistik (arbejdssteder). Fra dette modul dannes et output vedr. arbejdsstederne, der går til modul 1. For personernes vedkommende går et output til modul 2, hvor personoplysningerne endeligt dannes.

I *modul 2* kobles først personoplysningerne fra det seneste år sammen med de tilsvarende oplysninger året før. Herved sammenkædes oplysninger om arbejdsgivernummer og arbejdsstedskode i de to år for hver enkelt person. Disse personoplysninger overføres til modul 1. I modul 2 foretages derefter en overførsel af data fra en række andre personregistre. Der er tale om følgende statistikregistre:

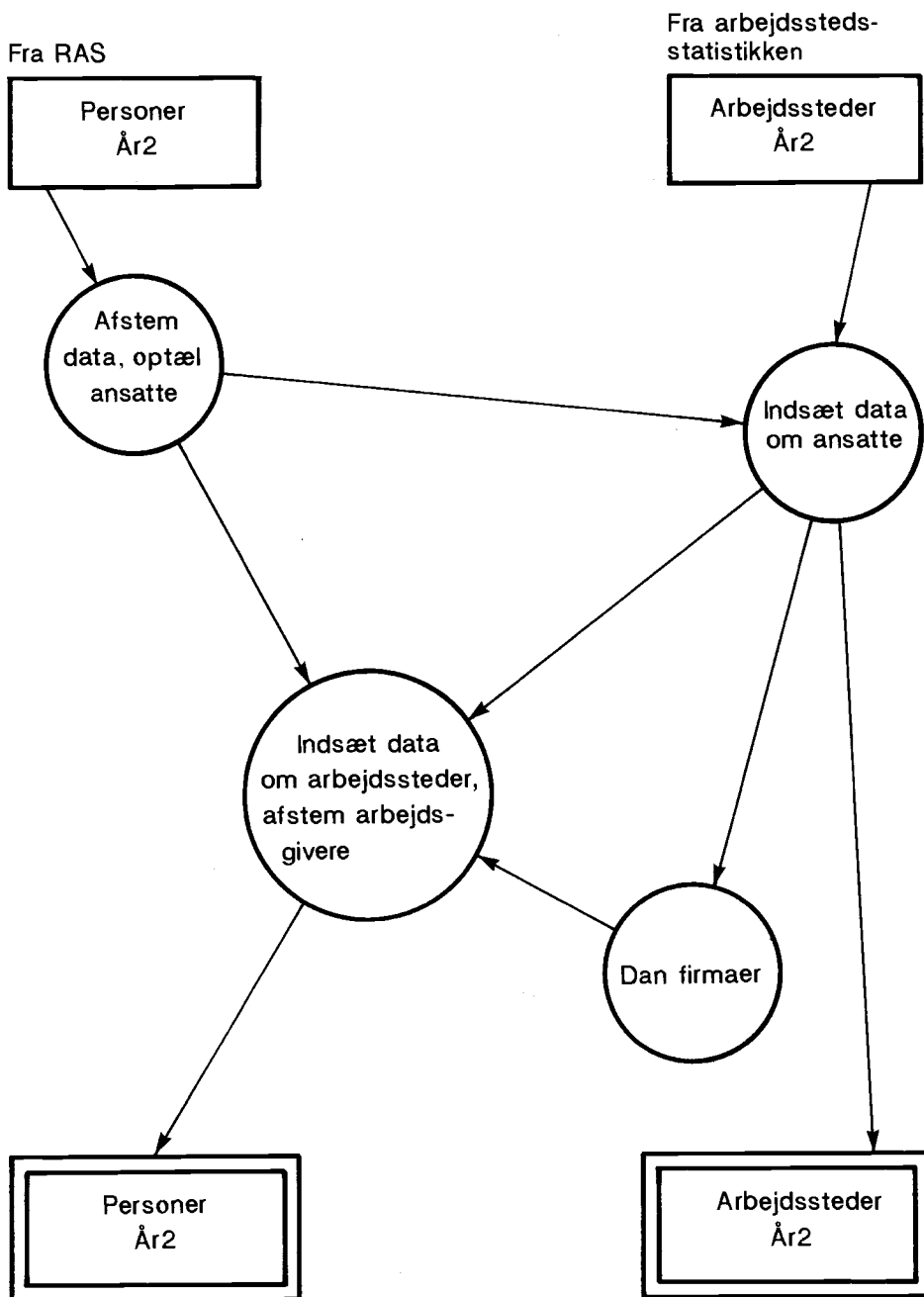
- CON: Arbejdspladsstatistikregisteret.
Løn, ansættelsesperiode mv.
- CPR: Befolkningsstatistikregisteret.
Demografiske oplysninger, familieforhold.
- CPRX: Befolkningsstatistikregisteret (ændringer i året).
Ind-og udvandring, fødsler og dødsfald.
- ARBL: Arbejdsløshedsstatistikregisteret.
Ledighedsoplysninger.
- INDK: Indkomststatistikregisteret.
Indtægts- og formueoplysninger.
- AKM: Arbejdsklassifikationsmodulet.
ATP-oplysninger mv.
- UDD: Uddannelsesstatistikregisteret.
Højeste uddannelse. Kurser for faglærte/specialarbejdere.

Ud fra disse oplysninger dannes en række nye variable til at karakterisere personerne.

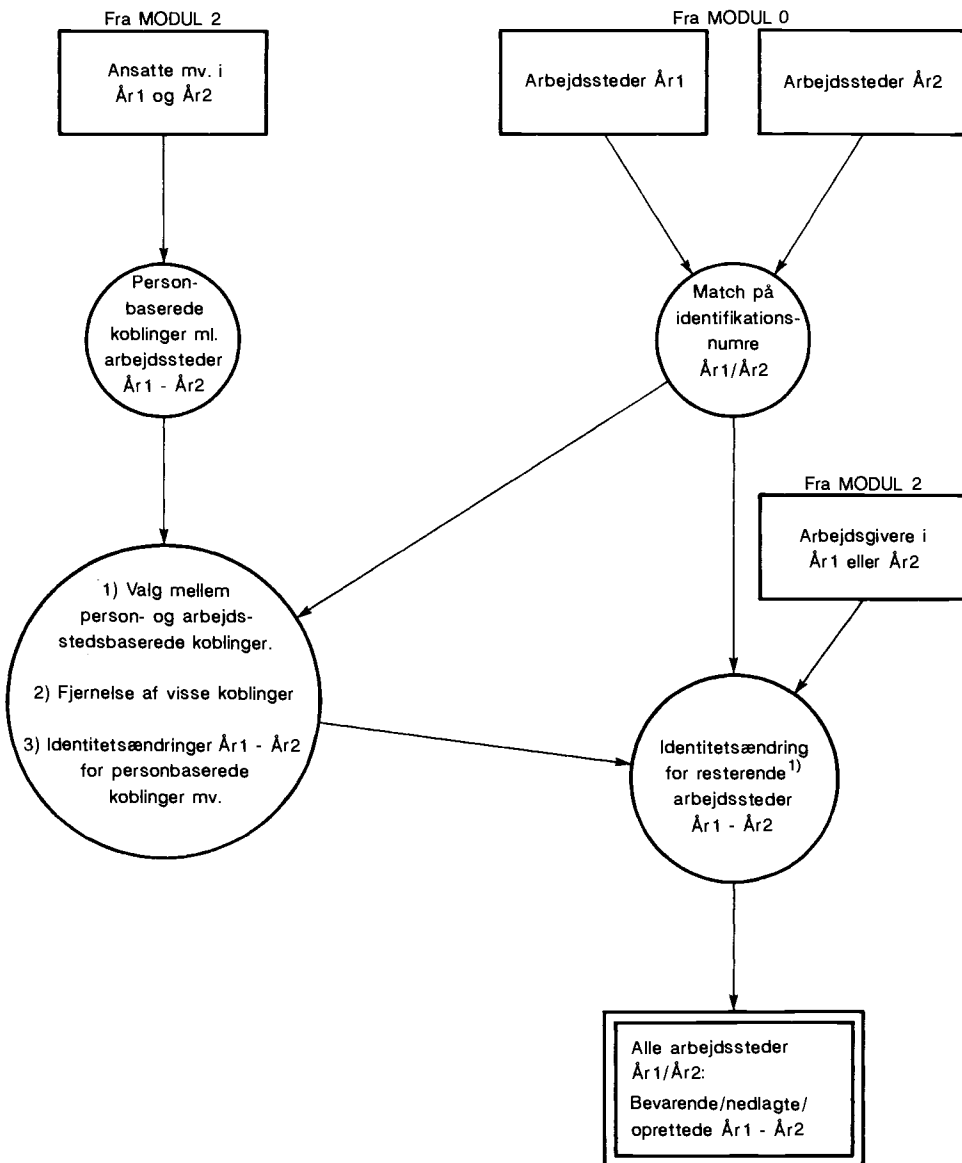
I *modul 1* er input som nævnt arbejdsstedsoplysninger (fra modul 0) og ansættelsesoplysninger for personerne fra de 2 seneste år (fra modul 2). Endvidere anvendes arbejdsstedsoplysningerne fra det foregående år. Disse oplysninger danner udgangspunkt for fastlæggelsen af ændringer i arbejdsstedernes identitet mellem de to år.

I *modul 3* kobles data om arbejdsstederne inkl. oplysninger om identiteten over tid (fra modul 1) sammen med person- og ansættelsesoplysninger (fra modul 2), samt oplysninger fra de tilsvarende datasæt året før. På grundlag heraf dannes oplysninger om ansættelser/beskæftigede og om virksomheder (arbejdssteder og firmaer).

Figur 6. MODUL 0: Afstemning af oplysninger for personer og arbejdssteder

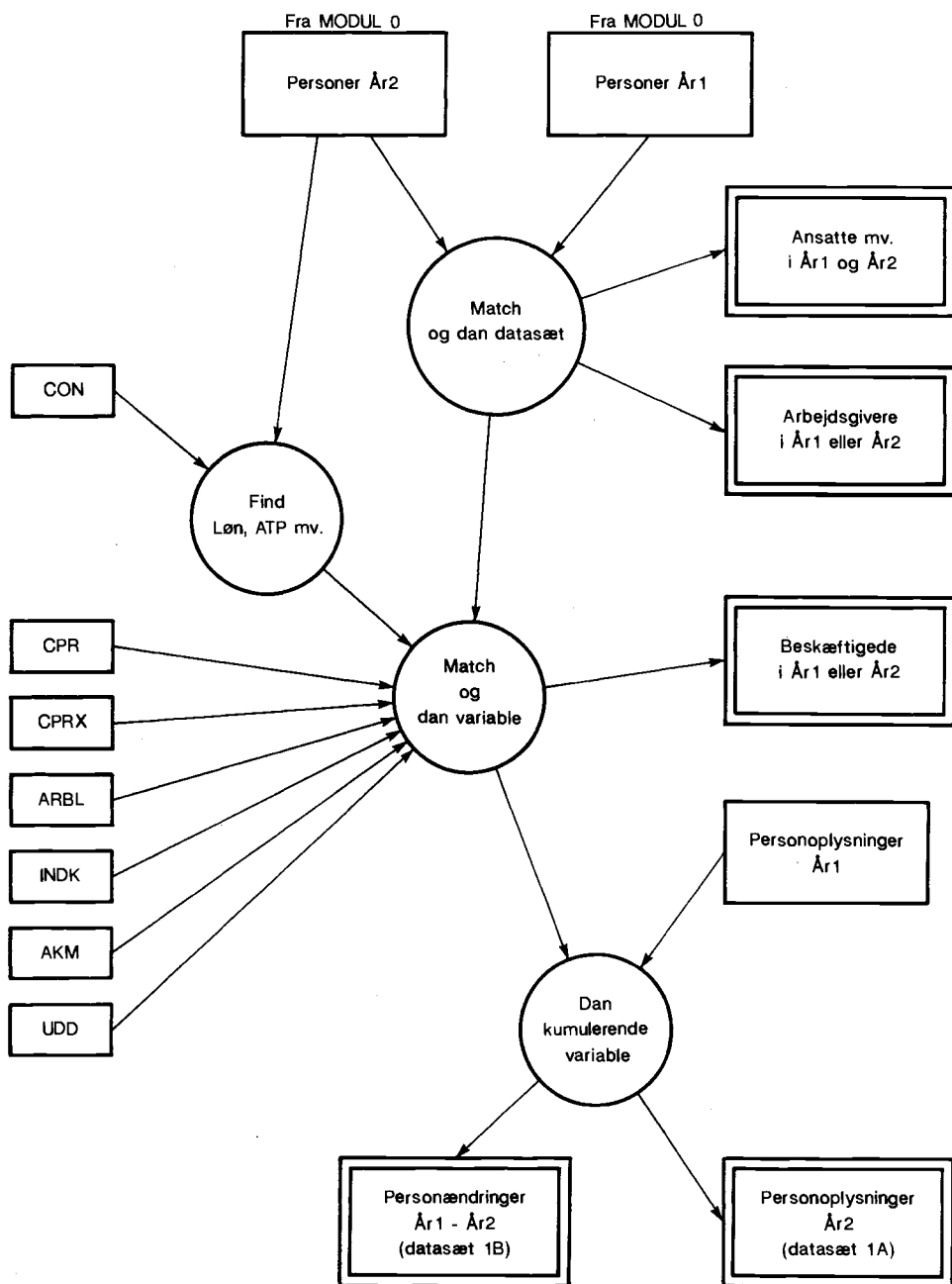


Figur 7. MODUL 1: Fastlæggelse af arbejdssteders identitet over tid (mellem År1 og År2)

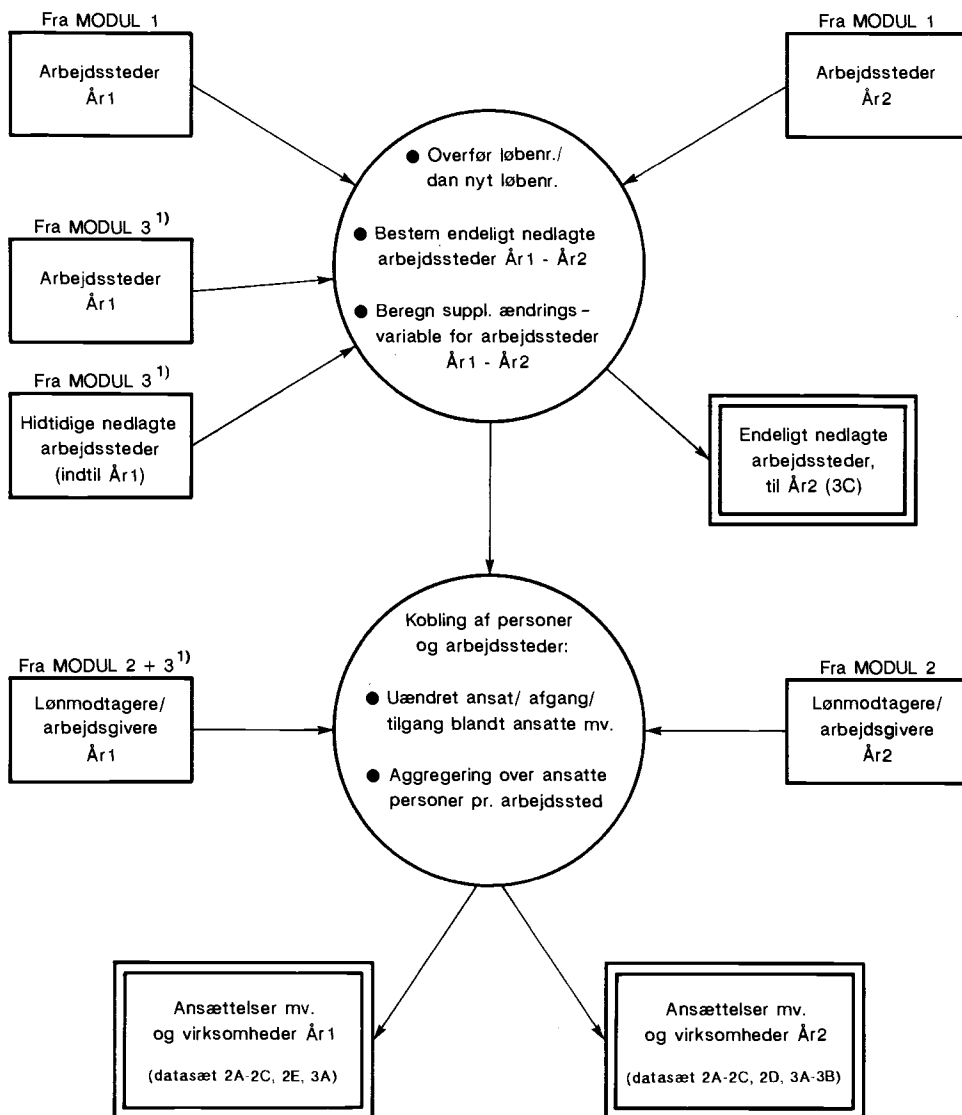


¹⁾ Arbejdssteder, som ikke indgår via de personbaserede koblinger, herunder arbejdssteder uden ansatte pr. november i det ene eller begge år.

Figur 8. MODUL 2: Dannelse af personoplysninger



Figur 9. MODUL 3: Dannelse af ansættelses- og virksomhedsdatasæt



Anm. Oversigten er ekskl. dannelse af datasæt vedrørende selvstændige og medhjælpende ægtefæller, der ikke indgår i koblingen mellem personer og arbejdssteder.

¹⁾ Fra MODUL 3 vedrørende det forudgående par af År.

Bilag 3:

Hvordan får man adgang til IDA?

Dataindholdet i IDA skal fortrinsvis anvendes som grundlag for forskningsprojekter indenfor arbejdsmarkedsforskningen. I dette afsnit beskrives de muligheder, forskerne har for at få rådighed over data. Beskrivelsen er generel, idet mulighederne helt afhænger af det enkelte projekts karakter.

Det må indledningsvis nævnes, at selv om der er tale om en database, er der *ikke* direkte (on line) adgang til data. Der er en række gode grunde hertil.

Data er organiseret som et register, idet datasættene indeholder person- og virksomhedsnumre, som skal anvendes som nøgle ved databehandlingen. Registerlovgivningen og Danmarks Statistiks egne diskretionsprincipper fastslår, at der *aldrig* må videregives oplysninger om personer og virksomheder, der medfører, at personen eller virksomheden kan identificeres.

Det er i den forbindelse ikke nok, at identifikationsnøgler fjernes, idet data kan være så detaljerede, at dette ikke betyder noget. Selvom man fx fjerner virksomhedsnummeret for en stor virksomhed, kan geografiske, økonomiske og personalemæssige oplysninger alligevel let føre til en afsløring af virksomhedens identitet.

Begrænsningerne fastlægges som nævnt i Registerloven, men derudover har Danmarks Statistik selv indført begrænsninger, idet institutionen har en række afgørende hensyn at tage. Det væsentligste er, at grunddata er indsamlet med den almindelige statistikproduktion for øje.

Der er også praktiske, tekniske hensyn, der gør, at direkte adgang er *umulig*. Datamængden er så omfattende, at der typisk skal dannes særlige udtræk af data, eventuelt fra flere datasæt, for at kunne anvendes i et bestemt forskningsprojekt. Denne opgave kan

kun hensigtsmæssigt ske i Danmarks Statistik med anvendelse af Danmarks Statistiks specielle edb-programmer.

Det typiske projektforsløb

Det enkelte projekt må beskrives overfor Danmarks Statistik, og der må foretages en afgrænsning af de data, der er nødvendige for projektets gennemførelse. Dette er en af de vanskeligste opgaver for den enkelte forsker, som naturligt nok vil kunne forestille sig mange interessante aspekter ved at medtage mange variable.

På grundlag af projektbeskrivelsen og dataafgrænsningen giver Danmarks Statistik et tilbud på, hvorledes data kan stilles til rådighed.

Der er *tre hovedmuligheder*, som imidlertid vil kunne kombineres. Mulighederne er tabelleverancer, levering af modeldata og anvendelse af Danmarks Statistiks forskerplaceringsordning.

1. Tabelleverance fra Danmarks Statistik

Den traditionelle tabelleverance fra Danmarks Statistik er en mulighed. Forskeren afgiver sin bestilling med angivelse af tabelhoveder og -forspalter. Tabelleverancer har den fordel, at det er rimeligt let at løse diskretionsproblemerne, idet Danmarks Statistik vil kunne udelade "tal, der er for små", men det må også erkendes, at tabellerne - uanset at de kan leveres på disketter - kun vanskeligt tilgodeser en moderne forskningsproces.

2. Modeldata

Danmarks Statistik har til *forskningsformål* mulighed for at levere såkaldte modeldata fra IDA-databasen. Modeldata består i princippet af en record pr. person/virksomhed uden formel identifikation dvs. uden personnr. eller virksomhedsnr. Hovedreglen er, at modeldata kun kan bearbejdes ved forskerplacering, jf. pkt. 3. med undtagelse af datasæt, hvor specifikationsgraden er lav, eller hvor der er tale om en meget lav udvalgsandel. I så fald kan forskeren få data udleveret mod at underskrive en aftale, som fastlægger betingelserne for anvendelsen. Disse betingelser vedrører bl.a.

Begrænsning af brugen. Data må kun bruges af bestemte personer (forskeren og dennes medarbejdere), evt. på bestemte edb-anlæg og under iagttagelse af bestemte forskrifter.

Begrænsning af videregivelse. Videregivelse af data fra forskeren til andre må normalt ikke finde sted, med mindre oplysningerne forinden er bearbejdet i en sådan grad, at de er publicerbare. Hvornår dette er tilfældet skal enten beskrives i aftalen eller afgøres konkret af Danmarks Statistik ved forhåndsgennemsyn af manuskripter inden videregivelse eller offentliggørelse.

Ret til efterkritik. Levering til en bestemt bruger af meget detaljerede oplysninger, som ikke er offentligt tilgængelige, kan gøre det vanskeligt for andre at gennemskue eller kritisere modtagerens brug af oplysningerne. Derfor må Danmarks Statistik forbeholde sig ret til at genbruge leverede data.

Krav om offentliggørelse. Danmarks Statistik kan kræve, at modtageren af data forpligter sig til at offentliggøre resultaterne af sin analyse.

Eksklusivaftaler, som giver brugeren eneret til anvendelse af et givet datamateriale, kan ikke indgås.

Uanset at data udleveres fra Danmarks Statistik, skal data fortsat betragtes som fortrolige, jf. Forvaltningslovens § 27, stk. 3 og Straffelovens § 152.

Modeldata vil kunne danne grundlag for analyser på større datamaterialer, idet forskeren på sin egen arbejdsplads kan udvikle edb-programmer, som kan videresendes til Danmarks Statistik til kørsel på større datasæt.

3. Forskerplacering i Danmarks Statistik

I de tilfælde, hvor leverancer af tabeller eller modeldata ikke kan anvendes, kan Danmarks Statistik tilbyde forskeren at få arbejdsplads i Danmarks Statistik en periode. Fordelene er, at forskeren kan få direkte adgang til større, detaljerede datasæt under sit arbejde med data. Forskeren kan selv gennemføre sine analysekørsler (evt. via en studentermedarbejder).

Der oprettes en aftale for forskerplaceringen, som beskriver betingelserne, der stort set er de samme som nævnt under modeldata. Resultater af bearbejdningen skal være af en sådan karakter, at personer og virksomheder ikke kan identificeres, og materialer kan kun medtages med Danmarks Statistiks samtykke.

Økonomi

Fordelen ved IDA-databasens etablering er, at der på forhånd er sket en bearbejdning af data, for hvilke den enkelte forsker ikke skal betale.

Det kan dog ikke undgås, at der er omkostninger forbundet med brugen af IDA.

Det er vanskeligt at angive omkostningerne konkret, idet det må formodes, at de enkelte projekter vil være af meget forskellig karakter og omfang.

Hovedprincippet er, at Danmarks Statistiks arbejdsindsats og CPU-tidsforbrug skal betales efter de generelle takster for sådanne opgaver.

Ved levering af modeldata vil Danmarks Statistiks tidsforbrug blive afgørende. Ved forskerplaceringsordninger vil omkostningerne til udtræk af data blive beregnet på tilsvarende måde, medens omkostninger ved forskerens arbejde i Danmarks Statistik beregnes ud fra forbruget af CPU-tid, konsulentbistand fra Danmarks Statistiks medarbejdere, samt et beløb pr. uge forskeren beslaglægger en arbejdsstation.