

## Metodenotat angående statistik fra Ungedatabasen

I dette notat beskrives nøgletal og dimensioner for statistikken baseret på Ungedatabasen. Der bemærkes en række forhold i relation til analyser på baggrund af statistikken, samt der gøres rede for datakilder, problematikker vedrørende datakvalitet i de data, der indberettes til Ungedatabasen samt metoden for opgørelse af data fra Ungedatabasen.

### Beskrivelse af nøgletal og dimensioner

Følgende nøgletal og fordeling på dimensioner dannes ud fra Ungedatabasen. Et nøgletal er den enhed, der tælles på, eller beregnede enheder. I denne statistik er der kun et nøgletal: antal personer. En dimension er en gruppering af nøgletal, her f.eks. alder.

<b>Nøgletalsgruppe: Status</b>	
Status angiver personers aktivitetsmæssige status på en given dato. En enkelt person kan kun have én aktivitet på en given dato, så hvis der foreligger flere samtidige aktiviteter for en person, vises den, der er højest prioriteret. Uddannelse prioriteres altid frem for andre aktiviteter, mens beskæftigelse prioriteres lavest. (jf. side 8)	
<b>Nøgletal</b>	<b>Beskrivelse</b>
Antal personer	Antal personer opgøres som det antal personer, der er i pågældende aktivitet på den pågældende dato. Hver person tæller kun én gang på opgørelsesdatoen. Det samlede antal personer på en dato er derfor lig populationen den pågældende dato.

<b>Dimensioner</b>	
<b>Dimension</b>	<b>Beskrivelse</b>
Person	Oplysninger for de personer, opgørelsen vedrører, som de ser ud på opgørelsesdatoen. Det er ikke muligt at se enkeltpersoner, men i stedet kan der vælges og fordeles på forskellige attributter som bopæl, UU-center, køn og alder.
Dato	Måned og år for opgørelsen. Status opgøres pr. den sidste dag i hver måned.
Placering (Aktivitetshierarkiet)	Den aktivitet (placering) personen har på opgørelsesdatoen. Aktiviteten kan vælges og fordeles efter et aktivitetshierarki udviklet af Ministeriet for Børn og Undervisning. Aktivitetshierarkiet tager udgangspunkt i UNI-C's Uddannelseshierarki, der anvendes i UNI-C's øvrige uddannelsesstatistik. Dette hierarki er til brug for statistik fra Ungedatabasen blevet udvidet og tilpasset således, at det klassificerer aktiviteter efter deres relevans for opfyldelsen af målsætningerne i Ungepakke II (Akti-

	<p>vitetshierarkiet er nærmere beskrevet nedenfor). Aktivitetshierarkiet og UNI-C's Uddannelseshierarki er dermed ikke identiske og kan ikke sammenlignes direkte.</p>
--	--

## De enkelte dimensioner

<b>Person</b>	
<b>Attribut</b>	<b>Beskrivelse</b>
Højest Fuldførte Uddannelse	Personens højest fuldførte uddannelse. Mulige værdier er 'Ingen', 'Ungdomsuddannelse' og 'Videregående uddannelse'.
Alder	Personens alder (i år).
Aldersgruppe	Aldersgruppering i intervallerne '15-17 år', '18-24 år' og '25-29 år' <sup>1</sup> .
Køn	Personens køn.
Kommune	Personens bopælskommune.
Landsdel	Landsdelen, hvori personens bopælskommune ligger.
Region	Regionen, hvori personens bopælskommune ligger.
<b>Hierarki</b>	<b>Beskrivelse</b>
Alder	Personens aldersgruppe og alder.
Bopæl - Region - Kommune	Region og kommunen for personens bopælskommune.
Bopæl - Region - Landsdel - Kommune	Region, landsdel og kommunen for personens bopælskommune.

<b>Dato</b>	
<b>Attribut</b>	<b>Beskrivelse</b>
Måned	Måned og år på formen 'Januar 2011'. Dvs. status den sidste dag i januar 2011.
Kalenderår	Årstal.
<b>Hierarki</b>	<b>Beskrivelse</b>
Kalenderår-måned	År og måned.

<b>Placering (Aktivitetshierarkiet)</b>	
<b>Attribut</b>	<b>Beskrivelse</b>
Aktivitetssområde	1. niveau i aktivitetshierarkiet.
Hovedaktivitet	2. niveau i aktivitetshierarkiet.
Mellemaktivitet	3. niveau i aktivitetshierarkiet.
Underaktivitet	4. niveau i aktivitetshierarkiet.

<sup>1</sup> På nuværende tidspunkt er det udelukkende muligt at se statistikken for aldersgruppen '15-17 år'. På et senere tidspunkt vil det ligeledes være muligt at se statistikken for de to øvrige aldersgrupper.

Hierarki	Beskrivelse
Aktivitet	Aktivitetshierarki, som klassificerer aktiviteter efter Uddannelseshierarkiet og deres relevans for opfyldelsen af målsætningerne i Ungepakke II. (Se bemærkninger til dette nedenfor).

<b>Kommune med vejledningsforpligtelse</b>	
Attribut	Beskrivelse
Kommune	Kommune med vejledningsforpligtelsen over for personen.
Landsdel	Landsdelen, hvori kommunen med vejledningsforpligtelsen ligger.
Region	Regionen, hvori kommunen med vejledningsforpligtelsen ligger.
UU-center	Det UU-center, hvor kommunen med vejledningsforpligtelse hører under.
Kommune nr.	Kommunennummeret for kommunen med vejledningsforpligtelsen.
Landsdelskode	Landsdelskoden for landsdelen, hvori kommunen med vejledningsforpligtelsen ligger.
Regionskode	Regionskoden for regionen, hvori kommunen med vejledningsforpligtelsen ligger.
UU-centerkode	Koden for det UU-center, hvor kommunen med vejledningsforpligtelse hører under.
Hierarki	Beskrivelse
Vejledningsforpligtelse - Region - Kommune	Region og kommunen for kommunen med vejledningsforpligtelse.
Vejledningsforpligtelse - Region - Landsdel - Kommune	Region, landsdel og kommunen for kommunen med vejledningsforpligtelse.
Vejledningsforpligtelse - UU-center - Kommune	Det UU-center, hvor kommunen med vejledningsforpligtelse hører under og kommunen.

### **Bemærkninger til Aktivitetshierarkiet**

I forbindelse med placering af aktiviteter i Aktivitetshierarkiet anvendes en række forretningsregler på specifikke aktiviteter. De beskrives nedenfor.

#### **Beskæftigelse**

Når en persons aktivitet er beskæftigelse, afhænger placeringen af beskæftigelsesgraden. Beskæftigelsesgraden er beregnet på baggrund af månedlige indkomstop-

lysninger fra e-indkomstregisteret, anslåede timesatser<sup>2</sup> for aldersgrupperne og antal timer på en måned.

Beskæftigelse er bekræftet ud fra et ugentligt timetal i sammenhæng med personens alder. For personer under 18 år er aktiviteterne 'Arbejde - under 18 t/uge' og 'Arbejde - over 18 t/uge', mens den for personer på 18 år og derover er 'Arbejde - under 30 t/uge' og 'Arbejde - over 30 t/uge'.

### ***Uden aktivitet og ingen offentlig forsørgelse***

Når en person ikke er i nogen aktivitet og ikke modtager offentlig forsørgelse, er personerne placeret under hovedaktiviteten 'Uden aktivitet og ingen offentlig forsørgelse'. Personerne underinddeles i mellemaktiviteterne 'Under 1 måned', '1 til 2 måneder', '2 til 3 måneder' eller 'Over 3 måneder', afhængigt af hvor længe personen har været uden aktivitet og offentlig forsørgelse.

### ***Fejlagtige aktiviteter***

I visse tilfælde er personer indberettet til Ungedatabasen med en aktivitet, der ikke er logisk mulig. Det gælder eksempelvis aktiviteter, der er lovbestemte og afhængige af personens alder. I disse tilfælde placeres personen under aktiviteten 'Uoplyst/ukendt'.

### ***Uklassificerede aktiviteter***

Aktiviteter, hvor der ikke kunne oversættes fra kildesystemernes kodesystemer, placeres under aktiviteten 'Uklassificeret aktivitet'. Det kan eksempelvis være nye uddannelser, der endnu ikke er indplaceret i Uddannelseshierarkiet. Disse aktiviteter søges løbende indplaceret i Aktivitetshierarkiet.

## **Statistikken**

Statistikken fra Ungedatabasen adskiller sig fra anden statistik fra ministeriet på en række måder, som beskrives nedenfor.

### ***Dynamisk statistik***

Data i Ungedatabasen og dermed også statistikken fra Ungedatabasen er dynamisk og vil derfor løbende vil ændre sig. Det vil altid være de unges helt aktuelle status (månedsvist), der vises i statistikken, og der kan forekomme ændringer med tilbagevirkende kraft, hvorfor statistikken også vil kunne ændre sig bagudrettet.

### ***Baseret på den aktuelle population i aldersgruppe (Ikke årgangsbaseret)***

Data er ikke baseret på årgange af unge, men på den aktuelle population i en aldersgruppe eksempelvis 15-17-årige. Dette har en væsentlig betydning for, hvordan

---

<sup>2</sup> Følgende timesatser er grundlaget for beregningen af beskæftigelsesgraden: Unge under 18 år: 67 kr. pr. time. Unge 18-24-årige: 125 kr. pr. time. Timesatserne er udtryk for den direkte løn, dvs. uden løn under ferie, personalegoder mv. Kilde: DA's StrukturStatistik 2010.

der analyseres på udviklingen i tallene. Eksempelvis skal en stigning i antallet af unge 15-17-årige, der går i grundskolen i perioden fra januar til juni ikke forstås som en tilgang af elever til grundskolen i løbet af skoleåret, men at flere unge fylder 15 år i løbet af skoleåret og dermed indgår i statistikken. Tilsvarende vil nogle unge fylde 18 år og dermed ikke længere indgå i statistikken, hvilket eksempelvis vil være tilfældet med antallet af elever på gymnasiale uddannelser. Antallet af personer, der bliver 15 år og dermed tilgår gruppen af 15-17-årige er i ca. 6300 pr. måned og antallet af personer, der bliver 18 år og dermed fragår gruppen af 15-17-årige er i ca. 6400 pr. måned. Den samlede population ændrer sig kun minimalt, mens fordelingen på de forskellige aktiviteter er påvirket i forskellig grad af selve udviklingen i de unges alder.

### ***Udsving i sommerferieperiode***

Der ses generelt udsving i sommerferieperioden. Dette kan hænge sammen med, at status for en række unge vil være, at de er mellem uddannelser, samt at registreringen på uddannelsesstederne er påvirket (forsinket) af sommerferieafholdelse, hvilket afspejler sig i statusregistreringerne i Ungedatabasen og dermed også statistikken.

### ***Usikkerheder ved datakvalitet***

Ungedatabasen danner datagrundlaget for denne statistik. Datagrundlaget er mere usikkert end det normalt er tilfældet for de statistikker, som findes på ministeriets hjemmeside, idet data ikke har gennemgået den samme omfattende kvalitetssikring, jf. nedenfor for en uddybning om datakvaliteten.

## **Datakilder**

Ungedatabasen er baseret på fortløbende indberetninger om 15-29-åriges uddannelses- og beskæftigelsesaktiviteter fra administrative systemer i kommuner, gymnasier, erhvervsskoler, videregående uddannelser, UU-centrene, e-indkomstregisteret, m.fl.

Der er tre overordnede kilder, der leverer data til Ungedatabasen:

- Studieadministrative systemer, hvor data overføres maskinelt. Inkluderer webindberetninger fra uddannelsesinstitutioner.
- Kommunalbestyrelser (UU-centre).
- Statslige myndigheder, herunder Arbejdsmarkedsstyrelsen, Politiet, Forsvaret, Beredskabsstyrelsen og SKAT (e-indkomstregister).

## **Datakvalitet**

Ungedatabasen er baseret på aktivitetshændelser, der indberettes fra en række administrative systemer, jf. beskrivelsen af datakilder ovenfor. Disse hændelser (i gang med uddannelse, afslutning eller afbrud af uddannelse) definerer de unges aktivitet på et givet tidspunkt.

Indberetninger af data til Ungedatabasen sker altovervejende maskinelt. Kvalitetssikringen af data påhviler institutionerne og er afhængig af deres anvendelse af de respektive administrative systemer. Data kan kun kvalitetssikres ved kilden. Ved samling af data fra de forskellige kilder i Ungedatabasen er det kun i meget begrænset omfang muligt at validere data. Det valideres udelukkende, at personen har et validt cpr-nummer og er i den relevante aldersgruppe, samt at aktiviteten er mulig i forhold til aldersgruppen.

I forbindelse med etableringen af Ungedatabasen er der konstateret en række forhold ved kvaliteten af indberetningerne til Ungedatabasen, der medfører en række usikkerheder ved statistikker på baggrund af Ungedatabasen. Der er igangsat et datakvalitetsprojekt, der skal arbejde med at forbedre datakvaliteten i Ungedatabasen og imødekomme nogle af disse usikkerheder.

Den større usikkerhed ved data betyder, at analyser af statistikken bør tage højde for dette ved eksempelvis ikke at følge og sammenligne tallene på alt for detaljeret niveau, men i højere grad fokusere på nogle mere overordnede tendenser.

Nedenfor er beskrevet nogle af de større kendte fejl, som søges løst, såfremt det er muligt.

### ***Manglende og sent indberettede hændelser***

Det er vigtigt, at alle aktivitetshændelser bliver indberettet korrekt til Ungedatabasen. Det er desværre konstateret, at en lille del af disse hændelser på trods af indberetning fra de studieadministrative systemer ikke når Ungedatabasen, og dermed at de unges aktivitet ikke er korrekt beskrevet. Da effekten af en sådan manglende hændelse først bliver korrigeret, når der kommer en efterfølgende indberetning vedrørende samme aktivitet, kan fejlen godt blive bibeholdt i flere år. Hvis f.eks. en 'I gang'-indberetning forsvinder, vil den unge i Ungedatabasen stå som uden aktivitet. Tilsvarende vil en forsvundet 'Afbrudt' eller 'Afsluttet'-indberetning medføre, at den unge i Ungedatabasen vil stå som værende i aktiviteten i al fremtid. Problemet er dermed potentielt kumulativt.

En lille stikprøve på indberetningen af erhvervsuddannelser har identificeret, at problemets omfang udgør mellem en halv til en procent af alle indberetninger af denne type til Ungedatabasen. Såfremt problemet er generelt for indberetning af alle typer ungdomsuddannelser, som der er ca. 90.000 af årligt for 15-17-årige, vil det betyde, at der potentielt mangler 450-900 indberetninger. Problemets omfang er ikke anslået for de øvrige aldersgrupper. Disse manglende indberetninger kan have betydning for størrelsen af gruppen, som ikke er i uddannelse, beskæftigelse m.v.

Tilsvarende problemer opstår ved manglende indberetning og til dels også ved sen indberetning. Sen indberetning rettes der dog op på tilbage i tiden, dvs. det må forventes at statistik vedrørende de seneste måneder er mere usikker end opgørelse af de tidligere måneders. Andelen af manglende og sene indberetninger er det pt. ikke muligt at opgøre.

### **Usikkerhed om den totale population**

Ved beregning af størrelsen af gruppen af unge, der er uden aktivitet og ikke modtager offentlig forsørgelse, benyttes den samlede danske befolkning i de givne aldersgrupper, da gruppen netop er defineret ved den del af befolkningen, der ikke er i aktivitet. Derfor er befolkningens størrelse af betydning for denne gruppes størrelse.

Befolkningens reelle størrelse kan ikke aflæses af folkeregisteret, da fraflytninger fra Danmark ikke altid registreres. Det gælder specielt for personer på au-pair ophold, udvekslingsstuderende og udlændinge, der arbejder i Danmark i en periode. Disse personer er udlændinge, som får tildelt et CPR-nummer ved indrejse til landet og dermed optræder som en del af befolkningen. Desværre melder en del af disse personer ikke udrejse ved opholdets afslutning og dermed ser befolkningen ud til at være større, end den rent faktisk er. Disse 'ekstra' personer vil for nuværende, i statistikken baseret på Ungedatabasen, typisk være registreret som værende 'uden aktivitet og ingen offentlig forsørgelse', da de ikke har nogen registreret aktivitet.

Denne usikkerhed vil formodentlig have størst betydning for aldersgrupper over 18 år, men det vides ikke med sikkerhed.

Reduktion af denne usikkerhed indgår i datakvalitetsprojektet.

Det samlede befolkningstal pr. aldersgruppe i Ungedatabasen er baseret på CPR-registeret. Befolkningstallene stemmer ikke fuldstændig overens med Danmarks Statistiks befolkningstal. Det skyldes, at Danmarks Statistik tager hensyn til det forhold, at hændelser indberettes til CPR med en vis forsinkelse. Ved opgørelsen af den samlede befolkning f.eks. pr. 1. januar er udgangspunktet de personer, der på denne dato står tilmeldt folkeregistret (CPR). Danmarks Statistik korrigerer denne befolkning for fødsler, dødsfald, flytninger, vielser og skilsmisser, som er sket før 1. januar, men som indberettes til CPR senest 30 dage efter.

### **Beskæftigelsesgrad**

Ifølge bestemmelserne i Ungepakke II har unge under 18 år pligt til at være i uddannelse, beskæftigelse eller anden aktivitet. Pligten er opfyldt, hvis den unge arbejder mindst 18 timer ugentligt. Den unges beskæftigelsesgrad kan derfor afgøre, om han/hun opfylder pligten eller tilhører gruppen, som ikke er i uddannelse, beskæftigelse m.v..

Beskæftigelsesgraden beregnes som tidligere beskrevet på baggrund af månedlige indkomstoplysninger fra e-indkomstregisteret, anslåede timesatser for aldersgrupperne og antal timer på en måned. Anvendelsen af denne fremgangsmåde kan potentielt give anledning til forkerte beregnede beskæftigelsesgrader, hvis de unges løn afviger fra gennemsnitslønnen.

Der er pt. ikke noget skøn over effekten af dette kvalitetsproblem.

### **Metode ved konfliktende indberetninger**

I Ungedatabasen er registreret en række aktiviteter for unge mellem 15 og 30 år. Disse aktiviteter er kendetegnet ved en indberetning af en startdato og senere evt. indberetning af en slutdato, hvis forløbet er afsluttet eller afbrudt.

Den samme aktivitet kan være indberettet flere gange for samme person, men med forskellig startdato. Det er ved udarbejdelse af data til statistik defineret, at hvis startdato i to indberetninger af samme aktivitet på samme institution for samme person er adskilt med mindre end 30 dage, så anses det som værende et forløb, hvor den tidligste af de to startdatoer sættes til startdato til statistikbrug.

En person kan være i gang med flere aktiviteter på samme tid. Dette vil i nogle tilfælde være korrekt, i andre tilfælde skyldes det fejl eller forsinkelse i indberetning fra f.eks. uddannelsesinstitutionerne. Da hver person kun kan optræde med en aktivitet for hver månedsopgørelse, så er der opbygget et aktivitetshierarki, således at den aktivitet, der er højest prioriteret på listen, vælges. Første niveau i dette hierarki er opgjort i tabel 1.

**Tabel 1 Aktivitetshierarki – første niveau**

Aktivetsområde	Prioritet
Videregående uddannelse	1
Ungdomsuddannelse	2
Grundskole eller specifikke forberedende aktiviteter mv. (jf. note 1)	3
Specifikke forberedende aktiviteter mv. (jf. note 2)	4
Specifikke forberedende aktiviteter mv. (jf. note 3) eller Ikke i uddannelse, beskæftigelse m.v.	5

**Note 1:** Gælder følgende forberedende aktiviteter mv.: Ungdomsskole, efterskole, husholdnings- eller håndarbejdsskole, maritimt forberedelseskursus, introduktionsforløb, musikalsk grundkursus, øvrige uddannelser og kurser, almen voksenuddannelse (AVU), GSK, HF-enkeltfag, HF-fagpakker, daghøjskoler, folke- og ungdomshøjskoler, ordblindeundervisning, forberedende voksenundervisning (FVU), danskundervisning for udlændinge, voksenspecialundervisning, AMU, TAMU, produktionsskoler, andre udviklende og forberedende aktiviteter aftalt med UU, ophold i udlandet, praktik i ungevejledningen, vejledning og opkvalificering efter LAB-loven, virksomhedspraktik efter LAB-loven.

**Note 2:** Gælder følgende forberedende aktiviteter mv.: Værnepligt, foranstaltning på fuld tid efter Serviceloven, fritagelse for uddannelsespligt iht. Vejledningsloven, sygdom, orlov mv., arbejde - over 18 t/uge (for 15-17 årige), arbejde - over 30 t/uge (for 18-25 årige), selvstændig erhvervsdrivende.

**Note 3:** Gælder følgende forberedende aktiviteter mv.: Offentlig forsørgelse.