

## Indholdsfortegnelse

|   |           |
|---|-----------|
| Status på datasikkerhed .....   | 1         |
| Opdateringer af registre i Forskningservice .....   | 4         |
| NYT datasæt for Indkomst (IND) .....  | 5         |
| Nye datasæt for RAS (Registerbaseret Arbejdsstyrkestatistik) .....  | 7         |
| Data fra Danske Almen Medicinsk Database (DAMD) .....   | 8         |
| Studerendes adgang til registerforskningsprojekter i forbindelse med udarbejdelse af bachelor og speciale projekter ..... | 8         |
| Ændringer i hardware og software på DST's forskermaskiner.....  | 9         |
| <u>Telefontider i Forskningservice.....</u>   | <u>10</u> |

## Status på datasikkerhed

### Ny vejledning i hjemsendelse af analyseresultater:

*Vejledning på hjemmesiden*

Forskningservice har nu en vejledning klar med retningslinjer for hjemsendelse af analyseresultater.

Denne vejledning ligger på vores hjemmeside under linket:

<http://www.dst.dk/da/TilSalg/Forskningservice/Vejledninger>.

Her er retningslinjerne for hjemsendelse beskrevet detaljeret.

I afsnittet: **Hjemsendelse af analyseresultater**, se side 2, følger et lille uddrag af de vigtigste regler for hjemsendelse af analyseresultater, og hvad man skal gøre, hvis man ved et uheld har sendt mikrodata hjem.

## Nyt hjemsendelsessystem

*Begrænsning af forskernes adgang til hjemsendelse*

Der har fra de autoriserede institutioners side været et ønske om at kunne begrænse forskernes adgang til at hjemsende output, så det fx kun er udvalgte personer i et forskningsmiljø, der har denne rettighed. Denne facilitet er nu til rådighed i det nye hjemsendelsessystem, der har afløst det gamle hjemsendelsessystem på forskermaskinerne og de hostede maskiner.

I afsnittet om ”**Det nye hjemsendelsessystem**”, se side 4, gives et overblik over de vigtigste ændringer i forhold til det gamle hjemsendelsessystem, og hvad man skal gøre, hvis man ønsker at lukke for forskernes adgang til at hjemsende analyseresultater fra forskermaskinen.

## Kursus i datasikkerhed

*Planlægger datasikkerhedskursus*

I løbet af foråret og hen over sommeren, har vi holdt flere foredrag om vores datasikkerhedsregler for autoriserede miljøer. Vi har fået meget god og konstruktiv feedback på disse foredrag, og med den erfaring planlægger vi at tilbyde et egentligt datasikkerhedskursus i Danmarks Statistik. Kurset bliver meldt ud i løbet af efteråret, når det har fundet sin endelige form.

## Hjemsendelse af analyseresultater

Nedenfor følger et kort uddrag af de vigtigste regler for hjemsendelse af analyseresultater og hvad man skal gøre, hvis man ved et uheld har sendt mikrodata hjem.

### Hvad må hjemsendes:

*Ej identificerbare resultater*

Som hovedregel må der kun hjemsendes analyseresultater, aggregerede tabeller eller figurer, hvor det ikke er muligt at identificere enkelte enheder, dvs. hverken personer, husstande, familier, virksomheder og andre enheder med af-identificerede løbenumre. Dette gælder også selvom det af-identificerede løbenummer er fjernet. **Det må med andre ord ikke være muligt at genkende nogen enkeltpersoner eller -virksomheder i det hjemsendte materiale.**

Generelt gælder nedenstående tommelfingerregler for aggregerede data.

|   |
|---|
| Hjemsendte resultater bør som udgangspunkt have sådan en form, at de umiddelbart kan anvendes i en publikation. |
|---|

*Viderebearbejdning er tilladt*

Men det er tilladt at arbejde videre med det hjemsendte aggregerede materiale, fx til dannelse af figurer eller videre analyser i statistikprogrammer.

For tabeller gælder nedenstående:

*Tabeller*

|   |
|---|
| Tabeller bør indeholde mindst 3 observationer pr. celle |
|---|

For tabeller er tommelfingerreglen, at disse bør indeholde mindst 3 observationer pr. celle, men aggregeringsniveauet skal altid afhænge af en konkret

vurdering af hvad tabellen indeholder. For yderligere information henvises til ”[Datafortrolighedspolitik i Danmarks Statistik](#)”, kapitel 3.

*Max, min og percentiler*

Max. og min. og percentiler, som fx median, må hjemsendes, men kun hvis der ikke er fare for identifikation af enkeltindivider. Ingen enkeltpersoner/virksomheder må kunne genkendes i det hjemsendte output.

Funktioner som min. og max. henviser ofte til enkeltindivider, men de kan hjemsendes, hvis det er sikret, at ingen personer kan genkendes i materialet.

*Figurer*

Figurer må hjemsendes, hvis de ikke indeholder identificerbar information. Man skal ved figurer være særligt opmærksomme på såkaldte outliers/ ekstreme værdier.

Det skal derudover sikres, at figurerne ikke indeholder indlejrede mikrodata. Sidstnævnte kan fx, sikres ved at hjemsende figuren i pdf-format.

### Hvad må ikke hjemsendes?

*Hjemsend aldrig mikrodata*

Ingen form for mikrodata må hjemsendes

Mikrodata defineres som data, der fx indeholder nedenstående:

- Datasæt – eller dele af datasæt - med oplysninger på individniveau - dvs. hvor der i datasættet findes en række værdier for et enkelt individ (enkeltekords). Det gælder også i de tilfælde, hvor man har fjernet det af-identificerede løbenummer fx personnummeret. Dvs. et datasæt på individniveau bestående af baggrundsoplysninger som fx indkomst, uddannelse og socioøkonomisk status, men hvor det af-identificerede personnummer for hver person er fjernet
- Af-identificerede nøglevariabler som personnumre, CVR-numre, arbejdsstedsnumre, adressekoder mv. betragtes *altid* som mikrodata, da de angiver et unikt nummer, der henviser direkte til en enkelt person eller virksomhed. Selv om der fx blot hjemsendes af-identificerede personnumre uden anden information, er dette stadig ikke tilladt. Dette gælder også eksterne data, dvs. data som man selv har indsendt til Danmarks Statistik og som skal indgå i et konkret projekt

### Hvad kan I gøre ved et sikkerhedsbrud?

*Orienter os straks ved regelbrud*

Hvis I konstaterer, at der er hjemsendt mikrodata fra forskerordningen, så er det meget vigtigt, at Forskningservice orienteres herom snarest.

Beskriv i en mail hvornår data er hjemsendt, hvilken eller hvilke filtyper der er hjemsendt samt hvor mange filer med mikrodata, I har sendt hjem. Vi betragter det som en betydelig formildende omstændighed, at I selv orienterer os om bruddet.

## Det nye hjemsendelsessystem

Den nye løsning til hjemsendelse af filer blev endeligt indført ultimo maj 2015 og erstatter det gamle hjemsendelsessystem.

*Mail sendes nu med det samme*

Det nye system er gjort mere brugervenligt og har fået et par væsentligt ændringer i forhold til det gamle system. En vigtig ændring er, at mailen sendes med det samme, så man ikke skal vente på sit output. Det betyder på den anden side også, at der ikke er nogen "fortrydelsesret" ved hjemsendelse, som der var i det gamle system, hvor mails kun blev afsendt hver 10. minut.

Nedenfor en oversigt over de vigtigste ændringer i det nye system:

- Man kan sende flere filer i samme E-post
- Filer sendes med det samme, når man klikker "Hjemtag"
- Der er nu begrænsning på hvilke filtyper, der kan hjemsendes
- Der er fortsat en grænse på 3 MB på filstørrelsen, men man vil fremover få direkte besked, hvis en fil er for stor. Tidligere blev sådanne filer overført til en undermappe (2Big2ToSend)
- Forskeradgang uden mulighed for at hjemsende analyseresultater

*ADVARSEL – filerne sendes med det samme*

Man skal passe ekstra på nu, når man sender filer hjem. Det er nemmere, og det går hurtigere – der vil ikke blive tid til at fortryde, når man har klikket "Hjemtag".

*Kontakt omgående  
Forskningsservice ved  
mistanke om fejl*

Hvis man mener, at man har begået en fejl, skal man omgående henvende sig til institutionens kontaktperson i Forskningsservice.

### Hvordan begrænser man hjemsendelsesrettigheden?

*Begrænsning i  
hjemsendelsesrettighed*

Forskningsservice kan med det nye system fjerne og tildele en given forsker rettigheden til at hjemsende filer.

*Hvad institutionen skal  
gøre*

Hvis institutionen ønsker, at nogle af deres forskere IKKE skal have adgang til at hjemsende, skal den person, der er ansvarlig for institutionens aftale med Danmarks Statistik, sende en liste med forskernes projekt-identer til institutionens kontaktperson i Forskningsservice. Denne liste kan for eksempel være:

- Alle forskere på et givent projekt.
- Alle projekter for en given forsker.
- Udvalgte projekter for udvalgte forskere.

### Opdateringer af registre i Forskningsservice

I Forskningsservice har vi adgang til data fra en række forskellige statistikkontorer i Danmarks Statistik. Opdatering af vores data sker løbende umiddelbart efter offentliggørelse af data i statistikkontorerne.

Information om udgivelsestider i Danmarks Statistik kan findes via følgende link:

<http://www.dst.dk/da/Statistik/offentliggoerelser.aspx>

Visse data vil ikke være at finde på listen, idet det enten er data, der ikke offentliggøres statistik på, f.eks. IDA-data, eller er data, som Forskningsservice har fået via en ekstern leverandør, f.eks. Lægemedeldata. Opdateringen af disse registre forventes i 2015 at blive som følger:

- IDA-databasen for 2013: Sommer 2015
- LPR-data for 2014: Efterår 2015
- LMDB for 1. halvår 2015: august 2015

## NYT datasæt for Indkomst (IND)

### *IND – et nyt indkomstregister*

Der er dannet et nyt indkomstregister – IND - Indkomst - 2013 definition.

I dette er indeholdt variabler, der alle er dannet i forbindelse med revisionen af indkomststatistikken, som fandt sted med udgivelsen af indkomsterne for 2013 16. december 2014.

Revisionen af indkomststatistikken er delvist foretaget som følge af revisionen af nationalregnskabet i september 2014. Der er foretaget en række ændringer i statistikken.

### *Kvalitetsdeklaration*

Kvalitetsdeklaration for Indkomststatistikken findes via dette link: <http://www.dst.dk/da/Statistik/dokumentation/kvalitetsdeklarationer/indkomststatistik>

### *Notat om revision af indkomststatistikken*

Yderligere kan der læses om selve revisionen af indkomststatistikken 2013 via dette link;

<http://www.dst.dk/da/Statistik/emner/indkomster/personindkomster?tab=dok>

### *Periode*

Registeret er dannet for perioden 2000-2013 og vil fremover blive opdateret årligt. På sigt vil data blive kørt tilbage til og med 1987.

### *Variabler*

Der er tale om variabler, der tidligere eksisterede, men som er blevet erstattet af nye beregninger og dermed nye variabelnavne. Registeret indeholder følgende variabler:

- ALDER\_ULT\_INK (Personens alder på referencetidspunktet, 31 .december
- det pågældende år)
- CPR-tjek
- CPR-type
- DAGPENGE\_KONTANT\_13 (Dagpenge og kontanthjælp mv. i alt)
- DISPON\_13 (Indkomst efter skat og renter (samlet indkomst fratrukket renteudgifter)
- ERHVERVSINDK\_13 (Erhvervsindkomst, løn og nettooverskud af selvstændig virksomhed inkl.)
- FOLKEFORTID\_13
- KONTANTHJ\_13 (Kontanthjælp (underhold/forsørgelse))
- LOENMV\_13 (Lønindkomst i alt)
- NETOVSKUD\_13 (Nettooverskud af selvstændig virksomhed inkl.

- visse honorarer, før kap)
- OFF\_PENS\_EFTERLON\_13 (Offentlige pensioner, inkl. Efterløn)
- OFF\_OVERFORSEK\_13 (Offentlige overførsler)
- OMFANG (Omfang af skattepligt)
- OVRIG\_DAGPENGE\_AKAS\_13
- OVRIG\_KONTANTHJALP\_13 (Øvrige kontanthjælpsydelse)
- OVRIG\_OVERFORSEL\_13 (Andre overførsler, SU, boligstøtte, grøn check og børnetilskud)
- PERINDKIALT\_13 (Personindkomst i alt ekskl. beregnet lejeværdi af egen bolig og før fr)
- PEROEVRIGFORMUE\_13 (Aktieindkomst, overskud/underskud af skibsanpart og lejeindtægt sommer)
- Personnummer
- PRIVAT\_PENSION\_13 (Udbetalte private pensioner.)
- RESUINK\_13 (Restindkomst inkl. Børnebidrag)
- SKATMVIALT\_13 (Skat, arbejdsmarkedsbidrag og særlig pension i alt. ekskl. Kirkeskat)
- SKATTOT\_13 (Samlet personlig slutskat for en skatteyder inkl. udenlandsk betalt sk)

### Øvrige indkomstregistre

De øvrige indkomstregistre i Forskningservice grunddata, INDK og INDH, eksisterer fortsat.

*INDK– nye variabler* I INDK er der kommet nye variabler for årene 2010-2013:

#### Kun 2010-2013

|                      |                                     |
|----------------------|-------------------------------------|
| EJD_SKAT_EJER_BOLIG  | Ejendomsskatter ejer (grundskyld)   |
| EJD_SKAT_LEJER_BOLIG | ejendomsskatter lejer               |
| GRON_CHECK           | Grøn check                          |
| SYG_BARSEL           | Udbetalte sygedagpenge fra kommuner |

#### Kun 2013

|                   |                                 |
|-------------------|---------------------------------|
| FLEXLN_TILSKUD    | Tilskud til fleksløn            |
| KOM_RES_FORLOB    | kommunalt ressourceforløb       |
| RNT_UDG_OFF_GAELD | Renteudgifter på offentlig gæld |
| UDD_YD            | uddannelses ydelse              |

*TIMES-dokumentation* Alle variabler i IND og INDK kan findes i TIMES-dokumentationen på følgende link;  
<http://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst>

## Nye datasæt for RAS (Registerbaseret Arbejdsstyrkestatistik)

*Detaljerede  
arbejdsmarkedsdata  
Nyt datagrundlag*

I maj 2015 er RAS, den registerbaserede arbejdsstyrkestatistik, blevet opdateret for årene 2008-2013. Den væsentligste ændring er, at datagrundlaget nu udgøres af Danmarks Statistiks nye arbejdsmarkedsregnskab (AMR).

AMR er et befolkningsregnskab, hvor arbejdsmarkedsrelaterede aktiviteter vægtes højest. Befolkningens tilknytning til arbejdsmarkedet fordeles på socio-økonomiske grupper, dvs. beskæftigede, studerende, ledige og andre offentligt forsørgede, børn og unge samt øvrige personer uden for arbejdsstyrken. I første omgang afvejes registrene imod hinanden, uden at der sker en nedskrivning af arbejdstimer ud over normal arbejdstid. Herved dannes et AMR uden normering; dette er det nye datagrundlag for den registerbaserede arbejdsstyrkestatistik (RAS).

*Nye data for perioden  
2008-2013*

Det er ikke muligt at føre et AMR-baseret RAS tilbage til før 2008, da AMR bygger på E-indkomstregisteret, som starter i 2008. AMR vil på et senere tidspunkt blive tilgængelig som et selvstændigt produkt i Forskningservices grunddatabank.

Den primære arbejdsmarkedstilknytning opgøres som hidtil pr. ultimo november, men det sker for befolkningen på det tidspunkt og ikke for befolkningen ultimo december i året. Navngivningen er samtidig ændret, så det nyeste datasæt hedder RAS 2013 og er opgjort november 2013.

*Øvrige ændringer*

Blandt øvrige ændringer kan følgende noteres:

- Fra og med 2013 eksisterer funktionskoden ikke længere. I stedet anvendes ARB\_SEKTORKODE.
- Kode-tekst-tabellen for SOC\_STATUS\_KODE kan fås ved henvendelse til Forskningservice. Selektion af populationen ultimo november (svarende til NOVPRIO=1) sker ved: I\_BEFOLKNINGEN\_KODE=1 og PRIMAER\_STATUS\_KODE=1. De beskæftigede selekteres ved: I\_BEFOLKNINGEN\_KODE=1 og PRIMAER\_STATUS\_KODE=1 og SOC\_STATUS\_KODE<=136.
- FIKTIV findes ikke mere, men de fiktive kan udskilles via FIKTIV\_ARB\_KODE, der har et andet værdisæt. 2 og 3 er fiktive arbejdssteder.
- DISCO\_MATCHPRIO\_KODE angiver, hvor "godt" DISCO'en er matchet på jobbet. Den svarer nogenlunde til den gamle DISCO\_KILDE\_KODE.
- FULDTID er afløst af TILSTAND\_GRAD\_AMR\_AAR. 1 svarer til 37 timer.
- SEKTORKODE hedder ARB\_SEKTORKODE.
- TILSTAND\_RAS er nu indeholdt i SOC\_STATUS og TILSTAND\_KODE.

Der henvises i øvrigt til dokumentationen i TIMES:

<http://dst.dk/da/Statistik/dokumentation/Times/moduldata-for-arbejdsmarked>

samt til kvalitetsdeklarationen:

<http://www.dst.dk/da/Statistik/dokumentation/kvalitetsdeklarationer/registerbaseret-arbejdsstyrkestatistik.aspx>

*Databrud og leverance af nye data*

Hvis man benytter tidligere udgivne data for 2008-2012 sammen med nye data for 2013, opstår der et databrud mellem 2012 og 2013. For at undgå dette kan man få leveret nye data for 2008-2012; dette sker til gældende takster. Du kan rette henvendelse til din FSE-kontaktperson for nærmere oplysninger herom.

## Data fra Danske Almen Medicinsk Database (DAMD)

*Ulovligt indsamlede DAMD data*

Det har, efter møde mellem Region Syddanmark, Datatilsynet og Sundhedsministeriet, vist sig, at der ikke er lovhjemmel i Persondataloven til at opbevare samtlige data fra Dansk Almen Medicinsk Database (DAMD) - en database med oplysninger om patientbehandling i almen praksis, idet der kun er lovhjemmel for de fire sygdomsspecifikke områder diabetes, KOL, depression og hjertesvigt. Det er den enkelte forskers ansvar, at ulovligt indsamlede data, som pågældende har fået lagt ind på Danmarks Statistiks forskningsservere, bliver slettet.

*Danmarks Statistik har kontaktet forskerne*

I forlængelse af debatten om lovligheden af DAMD-databasen har Danmarks Statistik kontaktet relevante forskere og understreget deres ansvar for at slette ulovligt indsamlede data. Danmarks Statistik udsendte derfor mails til alle forskere, der aktivt benytter – eller har benyttet - vores forskningsservere.

I henvendelsen opfordrede vi forskerne til straks at kontakte Danmarks Statistik, såfremt de havde data fra DAMD-databasen kombineret med data fra Danmarks Statistik på forskningsserveren, så det kan afklares, hvilke data der skal slettes.

## Studerendes adgang til registerforskningsprojekter i forbindelse med udarbejdelse af bachelor- og speciale projekter

*Betingelser for adgang for studerende ved universiteter*

I tilfælde hvor en autoriseret institution har et ønske om at tilvejebringe ny viden af et højt fagligt niveau på *eksisterende* projekter, kan studerende ved universiteter få adgang til disse data under forskerordningen med henblik på at udarbejde et bachelor- eller speciale projekt. Det er en forudsætning, at visse formelle krav er mødt.

*Krav til autoriseret institution*

Disse krav indbefatter, at projektet ikke alene oprettes med henblik på udarbejdelse af bachelor- eller speciale projekter. Endvidere skal den autoriserede institution tage det datasikkerhedsmæssige, kontraktlige og juridiske ansvar for den studerende, herunder underskriver forskeraftalen. Desuden er det et



krav, at der knyttes en faglig vejleder, ansat ved den autoriserede institution, til projektet.

## Ændringer i hardware og software på DST's forskermaskiner

Den fulde liste over hardware og software på DST's Forskermaskiner findes i dokumentet

[Hardware og Software på Danmarks Statistiks forskermaskiner](#)

som ligger under følgende link:

<http://www.dst.dk/da/TilSalg/Forskningsservice/Vejledninger.aspx>

Der er sket følgende ændringer i hardware og software på DST's Forskermaskiner siden seneste nyhedsbrev.

Ud over de faste program-pakker til GAUSS, SAS, WPS, SPSS og STATA har vi også en såkaldt DIVERSE-server, hvor vi som en særlig ydelse stiller forskellige programmer til rådighed for forskerne.

### DIVERSE

|                            |   |
|----------------------------|---|
| <i>Statistikprogrammer</i> | R, version 3.2.0 (64-bit)<br>Rstudio Version 0.98.1103  |
| <i>Microsoft</i>           | Microsoft ® Office Excel Viewer, Version 12.0.6717.5000 |

### GAUSS1

|                  |   |
|------------------|---|
| <i>Microsoft</i> | Microsoft ® Office Excel Viewer, Version 12.0.6717.5000 |
|------------------|---|

### SAS1

|                  |   |
|------------------|---|
| <i>Microsoft</i> | Microsoft ® Office Excel Viewer, Version 12.0.6717.5000 |
|------------------|---|

### SAS2

|                  |   |
|------------------|---|
| <i>Microsoft</i> | Microsoft ® Office Excel Viewer, Version 12.0.6717.5000 |
|------------------|---|

### SPSS1

|                  |   |
|------------------|---|
| <i>Microsoft</i> | Microsoft ® Office Excel Viewer, Version 12.0.6717.5000 |
|------------------|---|

### SPSS2

*Microsoft* Microsoft ® Office Excel Viewer, Version 12.0.6717.5000

### **STATA1**

*Tekstbehandling* Sublime Text 3, Build 3065

*Microsoft* Microsoft ® Office Excel Viewer, Version 12.0.6717.5000

### **STATA2**

*Tekstbehandling* Sublime Text 3, Build 3065

*Microsoft* Microsoft ® Office Excel Viewer, Version 12.0.6717.5000

### **WPS1**

*Hardware* Antal processorer: 1  
Processor frekvens: 2.4 GHz  
Antal kerner: 4  
Memory: 16 GB  
Temporært diskområde: 1,63 TB

*Statistikprogrammer* WPS Workbench, version 3.1.1.0.315

*Microsoft* Microsoft ® Office Excel Viewer, Version 12.0.6717.5000

## **Telefontider i Forskningservice**

*Telefontid fremover* Telefoniske henvendelser skal som hovedregel adresseres til din kontaktperson i Forskningservice. Alternativt er du velkommen til at kontakte os på følgende telefonnumre, der dagligt vil være åbne mandag til fredag kl 10-15. Uden for denne tidsperiode vil du i nogle tilfælde få en automatisk telefonbesked.

### **Administration Tlf. 39 17 31 32**

Spørgsmål om forskeraftaler, forskeroprettelser, password mv.

### **IT-support Tlf. 39 17 31 40**

Spørgsmål om log-on, servere, IT-problemer mv.

### **Forskerprojekter Tlf. 39 17 31 30**

Spørgsmål om data, leverancer, dokumentation mv

Mere information

Kontakt Forsknings-  
service på telefon  
+45 3917 3485  
forskningsser-  
vice@dst.dk