

## Indholdsfortegnelse

Opdateringer af registre i Forskningservice .....	1
Dannelse af nye BOL og BBR registre .....	2
Revision af Lønstatistikken for 2009-2012.....	3
Kvartalsvis A-indkomst.....	4
Ejendomsstatistik registre .....	5
Hjemsendelse af resultater fra forskermaskinen.....	6
Ændringer i hardware og software på DST's forskermaskiner.....	7
<u>Personaletilgang i Forskningservice.....</u>	<u>9</u>

## Opdateringer af registre i Forskningservice

I Forskningservice 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 Forskningservice har fået via en ekstern leverandør, f.eks. Lægemedeldata. Opdateringen af disse registre forventes i 2012 at blive som følger:

- IDA-databasen for 2012: Sommer 2014
- LPR-data for 2013: Efterår 2014
- LMDB for 2. halvår 2013: april 2014

## Dannelse af nye BOL og BBR registre

*Rettelse af gamle adresser,  
før kommunalreformen*

Der er i efteråret 2013 dannet følgende nye registre:

- BOL – Boligtælling for årene 2005-2013
- BBRE - BBR data på enhedsniveau for årene 2005-2013
- BBRB - BBR data på bygningsniveau for årene 2005-2013
- BBRP - Nøgle mellem PNR/ADRESSE.

Den ”nye” BOL opbygges i forlængelse af BOL 1981-2004 således, at en anvendelig tidsserie, som er konsistent over årene, dannes. Den nye Bol vil erstatte den nuværende BBR.

Forskellen på den ”nye” BOL 2005-2013 og den ”gamle” BBR 2005-2013 er primært navnene på variable og hvilke variable, der er med. Indholdsmæssigt er der ikke nogen større forskel.

### Population i BOL

Grundlaget i den ”nye” BOL er den samme som i den eksisterende BOL. Det betyder, at det fortsat er adressen ved KOM BOPIKOM, der er enheden/grundlaget i den nye BOL. Antallet af adresser er ca. 3. millioner, hvorimod der er ca. 2,5 million adresser i hustandsdatasættene.

*Beboede og ubeboede  
adresser i BOL*

Dette skyldes populationsafgrænsningen i BOL, som omfatter både beboede og ubeboede adresser. Sidstnævnte er boliger og fritidshuse uden en cpr-tilmelding pr. 1. januar.

*CPR\_BBR\_MATCHKODE\_GL*

Populationen afgrænses ved hjælp af variabelen **CPR\_BBR\_MATCHKODE\_GL**. Variablen dannes i forbindelse med samkørsel mellem BBR-data (BED (Bolig & Ejendomsdatabasen og Geo-data)) og CPR for årene 2005-2013.

### Population i BBR

Grundlaget i de ”nye” BBR-registre er ligeledes adressen ved KOM BOPIKOM. De nye registre indeholder al information om bygningerne, dvs. oplysninger på adresser der ikke nødvendigvis kan genfindes i befolkningsregistrene. Forskellen mellem BOL og BBR er således, at BBR indeholder informationer om bygningerne, hvor BOL indeholder informationer om hvor den enkelte person bor - med diverse oplysninger om denne adresse og enheden derpå.

For yderligere spørgsmål henvises til Charlotte Leolnar Reif 3917 3232.

## Revision af Lønstatistikken for 2009-2012

Lønstatistikken er blevet revideret for årene 2009-2012.

### *Nyt lønsystem i DSB*

Baggrunden for revisionen er indfasning af et nyt lønsystem i DSB i 2008, som ændrede karakteren af indberetningen af løndata fra DSB til Danmarks Statistik. Ændringen vedrører de forudlønnede medarbejdere i DSB og betyder, at stigningen i lønniveauet fra 2007 til 2008 har været for høj i lønstrukturstatistikken. Lønniveauet har for disse medarbejdere derfor været overvurderet siden 2008.

### *Feriebetaling i private sektor*

Der er ligeledes foretaget en revision af lønkomponenten ”feriebetalinger pr. præsteret time” for ansatte i den private sektor for året 2011. Lønkomponenten har fejlagtigt indeholdt den særlige feriegodtgørelse. Da den særlige feriegodtgørelse ligeledes optræder som et selvstændigt lønelement, er den fejlagtigt indgået to gange i statistikken. Fejlen påvirker ikke den overordnede løn, dvs. fortjenesten pr. præsteret time.

I følgende link til Nyt fra Danmarks Statistik, som udkom d. 10. oktober 2013, uddybes ovenstående :

<http://www.dst.dk/pukora/epub/Nyt/2013/NR548.pdf>

Revisionen berører alle lønvariable dvs. alle ”prae”-variable og alle ”stand”-variable.

Lønstatistikkontoret har udarbejdet et ”Question and answer”-dokument:

#### *Spørgsmål 1* **Hvorfor er fejlen opstået?**

*Svar 1* DSB indfasede i 2008 et nyt lønsystem, som betød at indberetningen til DSt ændrede karakter. Det betød at de forudlønnede i DSB fejlagtigt fik en meget stor lønstigning mellem 2007 og 2008. Men da DSB, ved særskilt møde mellem DSB og DSt i starten af 2009 oplyste, at visse personalegrupper i DSB i 2008 havde fået nogle store lønstigninger, lyste alarmklokkerne desværre ikke tilstrækkelig voldsomt, til at vi blev urolige ved de alt for store lønstigninger i 2008.

#### *Spørgsmål 2* **Har fejlen implikationer ud over lønstatistikken?**

*Svar 2* Nej

#### *Spørgsmål 3* **Hvilke personalegrupper er berørt – og hvor store er justeringerne?**

*Svar 3* Da DSB er en stor virksomhed, er der mange personalegrupper, som er berørt. De grupper, hvor DSBs forudlønnede fylder meget, er naturligvis ekstraordinært meget berørt. Det drejer sig primært om følgende arbejdsfunktionsgrupperne (DISCO-grupper), Lokomotivførerarbejde (831100), Kontrolarbejde under rejser (511200), Tograngering (831200), Rejsebureauarbejde (422100) og arbejde med kontrol af arbejdsmiljø og sikkerhed (3257..). Månedslønnen for 2011 på disse grupper korrigeres ned med henholdsvis; 6.500 kr., 6.300 kr., 4.100 kr., 3.500 kr. og 3.200 kr. For andre grupper drejer det sig om mindre end 600 kr. pr. måned.

#### *Spørgsmål 4* **Hvor mange år tilbage har I revideret tallene?**

*Svar 4* Tilbage til 2008

**Spørgsmål 5** Hvorfor har I først opdaget fejlen nu?

**Svar 5** I forbindelse med produktion af en statistik som den årlige lønstrukturstatistik, er der mange forskellige datamæssige problemer som skal løses. Problemerne skifter ofte karakter fra år til år, og giver på hver deres måde anledning til at se både data og produktionssystemer efter på forskellig måde. I forbindelse med valideringen af data for 2011 har der været anledning til at se data og systemer efter på en måde, som ledte os frem til problemet med de forudlønnede medarbejdere i DSB.

**Spørgsmål 6** Bliver Statens lønniveau påvirket samlet set? Og med hvor meget?

**Svar 6** Ja. Alle år bliver påvirket. De to seneste år dvs. 2010 og 2011 bliver påvirket med henholdsvis 1,5 kr. pr. time (207 kr. på månedslønnen) og 1,30 kr. pr. time (177 kr. på månedslønnen).

**Spørgsmål 7** Kan samme fejl ikke være opstået for andre personalegrupper?

**Svar 7** Nej

**Spørgsmål 8** Hvordan vil I forhindre, at den slags sker igen?

**Svar 8** Være mere fokuseret i forbindelse med systemændringer- og vedligeholdelse.

## Kvartalsvis A-indkomst

Med den Kvartalsvise A-indkomst stilles oplysninger om personindkomster til rådighed allerede 90 dage efter udgangen af kvartalet. Registeret indeholder data for 1. kvartal 2010 og frem. Statistikken omfatter løn, skattemæssig værdi af frynsegoder, pensioner- både offentlig og privat, efterløn, dagpenge samt kontanthjælp og lign.

Indkomsten er summen af indkomsterne i løbet af kvartalet. For indkomstudbetalinger, der omfatter dage i mere end et kvartal, fordeles beløbet forholdsmæssigt efter antallet af dage.

Den endelige årsstatistik for personindkomster offentliggøres 350 dage efter udgangen af året og indeholder selvstændiges indkomst samt boligstøtte og børneydelser, som ikke indgår i kvartalsstatistikken. Det gør, at kvartalsstatistikken undervurderer den disponible personindkomstmasse med ca. 15 pct. og gennemsnittet med 12 pct.

For yderligere information se :

<http://www.dst.dk/da/Statistik/dokumentation/Times/a-indkomst-kvartal.aspx>

## Ejendomsstatistik registrene

*Statistikens 6 registre* Ejendomsstatistik registrene dækker nedenstående 6 registre i Forskningservices grunddata.

EJER - Ejere af ejendomme  
EJID - Ejendomsidentifikation  
EJMT - Ejendomsmatrikler  
EJSK - Ejendomsskatter  
EJVK - Ejendomsvurderinger  
EJSA - Ejendomme salgsoplysninger.

*Datakvalitet for statistikvariabler* Kun ganske få variabler fra ejendomsstatistik registrene anvendes til statistik, og det er derfor også kun få variable, der fejlsøges. Disse variabler kan udleveres under samme vilkår som for alle andre fejlsøgte variabler i Danmarks Statistiks registre.

*Øvrige variabler er ikke fejlsøgt* Alle øvrige variabler er ikke fejlsøgte, og Danmarks Statistik kender derfor ikke til kvalitet og indhold i variablerne. Forskningservice kan derfor heller ikke rådgive, fejlrette eller hjælpe med problemløsninger for variablerne.

*Data på eget ansvar* Ønskes ikke-fejlsøgte data på trods heraf, kan de udleveres på eget ansvar. Vi understreger i den forbindelse, at man skal være opmærksom på følgende:

1. Datakvaliteten på variablerne kan være ringe
2. Variablerne fejlsøges og fejlrettes ikke
3. Danmarks Statistik yder ikke rådgivnings- og dokumentationsservice på variabler, der ikke anvendes til statistik, og som ikke allerede er dokumenteret i vores dokumentationssystem.
4. Danmarks Statistik tager ikke ansvar for data

Som følge heraf anbefaler Forskningservice, at man kun gør brug af variabler der indgår i statistikkerne. Samtidig fraråder vi anvendelse af de øvrige ejendomsstatistikvariabler.

*Dokumentation af statistik variabler* Dokumentation af ejendomsstatistik variabler, der anvendes til statistik, stammer fra registrene EJSA, EJER og EJID, og kan findes under højkvalitetsdokumentationen på vores hjemmeside for henholdsvis:

*Ejendomssalg her:*

<http://www.dst.dk/da/TilSalg/Forskningservice/Dokumentation/hoekvalitetsvariable/ejendomssalg.aspx>

*og Ejendomme her:*

<http://www.dst.dk/da/TilSalg/Forskningservice/Dokumentation/hoekvalitetsvariable/ejendomme.aspx>

*Rettelse af dato variabler* Da der har været problemer med datovariablerne Ejerstartdato (EJER-STARTDTO) og vurdering (VURDAN) i henholdsvis EJER og EJVK registrene, har Forskningservice besluttet at tilbyde datovariablerne med 2-cifrede og/eller 4-cifrede årstal for hele tidsserien.

De originale datovariabler er leveret til Danmarks Statistik med 2-cifrede årstal, og for at omsætte disse dato variabler til SAS datoformat er de 2-cifrede årstal blevet konverteret til 4-cifrede årstal i Forskningservice. I praksis betyder det dog, at et årstal kan komme til at tilhøre et forkert århundrede efter konverteringen.

I EJER registeret er 4-cifrede årstal erstattet med de originale 2-cifrede årstal frem til 2010. I de senere registerår (fra 2010 og frem) har vi også modtaget data med 4-cifrede årstal (selvom disse stadig er baseret på 2-cifrede årstal). Fra 2010 har vi derfor valgt at tage begge dato-variable med.

I EJVK registeret findes de 2-cifrede datovariabler kun frem til og med 2006. Herefter modtager vi kun datovariabler med 4-cifrede årstal, og kan derfor kun tilbyde dette.

Variablen med 2-cifrede årstal er et tekstfelt med dato angivet som YYMMDD.

## Hjemsendelse af resultater fra forskermaskinen

I Forskningservice kan man som forsker hjemsende resultater af analyser på e-mail. Det er til stor glæde for vores forskere, der på en let og enkel måde kan arbejde videre med de resultater, de har beregnet på forskerserverne. Mange af de data, som stilles til rådighed for forskerne, er oplysninger, som i teorien kan anvendes til identifikation af personer eller virksomheder. Derfor udfører Danmarks Statistik kontrol med de hjemsendte filer for at se, om de overholder datasikkerhedskravene. Dette er i forskernes interesse, da det bidrager til at bevare en ordening med en stor frihedsgrad til, hvilke registeroplysninger forskerne kan få stillet til rådighed og lave analyser på.

Vi har ændret på den måde, vi foretager kontrol af hjemsendte filer på. I den forbindelse er vi blevet opmærksomme på flere ting, som vi gerne vil dele med jer:

*Systemfiler mv.* Ordningen er tiltænkt til hjemsendelse af analyseresultater. Vi ser indimellem hjemsendelse af konfigurations og systemfiler mv., som vi derved også skal bruge ressourcer på at lave kontrol af. Muligvis skyldes dette at forskeren (ukritisk) vælger at hjemsende alle filer i en given mappe. Der må kun hjemsendes resultater, som I har brug for.

*Filtyper og extension*s Alle filer skal fremover have en filtype ekstension fx .xlsx (Excel). Og der skal være sammenhæng mellem filtypen og indholdet i filen, så det er muligt for Danmarks Statistik at åbne filen med det program, den er associeret med.

*Komprimerede filer* Der er en del forskere som vælger at pakke deres hjemsendte filer til zip-filer med Win-Zip. Det oprindelige formål har sandsynligvis været at kunne hjemsende filer, der har været større end mailsystemet kunne håndtere. Efter en gennemgang af disse filer ændrer vi på tre vigtige punkter:

- Vi hæver den tilladte størrelse på filer fra 2 MB til 3MB
- Det er kun tilladt at komprimere én enkelt fil i hver ZIP-fil.
- Det er ikke tilladt at opdele en enkelt meget stor fil i flere ZIPPEDE-filer fx z01, z02

## 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/Forskningservice/Vejledninger.aspx>

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

Bemærk, at SPAD 7.3.12 ikke vil blive installeret på den nye DIVERSE server.

### DIVERSE (erstatte DIVERSE1)

<i>Hardware</i>	Antal processorer: 4 Processor frekvens: 3.19 GHz Antal kerner: 32 Memory: 512 GB
<i>Operativsystem</i>	Windows Server 2008 R2 Enterprise, Microsoft® Windows Version 6.1 (7601, Service pack 1) x64-based (64-bit)
<i>Statistikprogrammer</i>	aML, version 2.09 GAMS Integrated Development Environment, version 23.2.1 LatentGOLD® version 4.5.0.11031 Mplus Version 6.12 Demo SCD/DIGRAM, Version 2.14.0 SPSSToDigram R, version 3.0.2 (64-bit) Rstudio Version 0.98.501
<i>Tekstbehandling</i>	GNU Emacs, version 24.3.1 gVim - VIM - Vi IMproved, version 7.1 Tinn-R, version 3.0.3.5 MiKTeX TeXworks, version 0.4.3 r.857 (MiKTeX 2.9) MiKTeX Previewer, Yap - Yet Another Previewer 2.9.4206
<i>Filkonvertering</i>	Stat/Transfer (grafisk), Version 12.0.122.0123 Stat/Transfer (kommandobaseret), Version 12.0.121.0123
<i>Filkomprimeringsprogram</i>	WinZip®, version 11.0 (7313)
<i>Microsoft</i>	Microsoft ® Calculator, Version 6.1 (7601) Microsoft ® Wordpad, Version 6.1 (7601) Microsoft ® Paint, Version 6.1 (7601)

## **GAUSS1**

*Filkonvertering* Stat/Transfer (grafisk) er opdateret til version 11.2.2092.0301  
Stat/Transfer (kommandobaseret) er opdateret til version 11.2.2092.0301

## **SAS1**

*Filkonvertering* Stat/Transfer (grafisk) er opdateret til version 11.2.2092.0301  
Stat/Transfer (kommandobaseret) er opdateret til version 11.2.2092.0301

## **SAS2**

*Filkonvertering* Stat/Transfer (grafisk) er opdateret til version 11.2.2092.0301  
Stat/Transfer (kommandobaseret) er opdateret til version 11.2.2092.0301

## **SPSS1**

*Filkonvertering* Stat/Transfer (grafisk) er opdateret til version 11.2.2092.0301  
Stat/Transfer (kommandobaseret) er opdateret til version 11.2.2092.0301

## **STATA1**

*Filkonvertering* Stat/Transfer (grafisk) er opdateret til version 11.2.2092.0301  
Stat/Transfer (kommandobaseret) er opdateret til version 11.2.2092.0301

*Statistikprogrammer* STATA/MP 12.1 for Windows 64-bit x86-64 er fjernet  
STATA/MP 13.1 for Windows 64-bit x86-64 er tilføjet



## STATA2 (Ny server)

<i>Hardware</i>	Antal processorer: 4 Processor frekvens: 3.19 GHz Antal kerner: 32 Memory: 512 GB
<i>Operativsystem</i>	Windows Server® 2008 R2 Enterprise, Microsoft Windows Server, Version 6.1 (7601, Service Pack 1) x64-based (64-bit)
<i>Statistikprogrammer</i>	STATA/MP 13.1 for Windows 64-bit x86-64
<i>Tekstbehandling</i>	GNU Emacs, version 24.3.1
<i>Filkonvertering</i>	Stat/Transfer (grafisk), Version 12.0.122.0123 Stat/Transfer (kommandobaseret), Version 12.0.121.0123
<i>Filkomprimeringsprogram</i>	WinZip®, version 11.0 (7313)
<i>Microsoft</i>	Microsoft ® Calculator, Version 6.1 (7601) Microsoft ® Wordpad, Version 6.1 (7601) Microsoft ® Paint, Version 6.1 (7601)

## WPS1

<i>Filkonvertering</i>	Stat/Transfer (grafisk) er opdateret til version 11.2.2092.0301 Stat/Transfer (kommandobaseret) er opdateret til version 11.2.2092.0301
------------------------	--

## Personaletilgang i Forskningservice

Vi har siden sidste nyhedsbrev få to ny medarbejdere i Forskningservice:

Pr. 1. november 2013 tiltrådte David Michelsen og John Hertz Nielsen.

Mere information

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