

SAS formater i Danmarks Statistik

| | | |
|----|--|----|
| 1. | Indledning..... | 1 |
| 2. | Hvor findes formater og øvrige datafiler?..... | 2 |
| 3. | Hvordan bruges formater i SAS-programmet?..... | 2 |
| 4. | Formaterne findes i flere varianter | 4 |
| 5. | Hvordan finder man de relevante formater?..... | 5 |
| | • Metadatabanken | 7 |
| | • Uddannelserne..... | 8 |
| | • Brancherne..... | 10 |
| | • DISCO (Arbejdsfunktioner)..... | 11 |
| | • Sundhed..... | 12 |
| | • Geokoder | 13 |
| | • DREAM databasen..... | 14 |

1. Indledning

Vejledningen beskriver, hvordan man udnytter de SAS formater, som Danmarks Statistik stiller til rådighed for forskerne. Målgruppen er de forskere, der arbejder med registerdata under Danmarks Statistiks forskerordning. Der forudsættes et vist kendskab til SAS formater for at få udbytte af vejledningen.

På Danmarks Statistiks hjemmeside findes dokumentation af variable og værdisæt i Danmarks Statistiks registre, hvis man ser under *Dokumentation* og *Variabler* og *Klassifikationer*. SAS formaterne adskiller sig fra denne dokumentation på følgende måde:

- Formater kan bruges direkte i programmeringen, når registerdata behandles.
- Man kan nemt få overblik over de variable/værdisæt som er relevante.
- Man kan nemt danne en række af standardgrupperingerne fx uddannelsesgrupper og branchegrupper.

Formaterne findes ligeledes tilgængeligt for forskerne i form af datasæt til brug i SPSS og STATA, som SAS-datasæt og som tekstfiler. Hvis databehandlingen sker i SAS, så udmærker formater sig ved, at de nemt kan bruges i datastep til at danne nye variable, og i procedurerne kan de bruges til at give rapporter sigende tekster og grupperinger. Meget tid til programmering kan spares, hvis man på forhånd har adgang til de nødvendige SAS formater.

For at få et overblik over omfanget af formater og selve indholdet i det enkelte format findes der en række HTML-sider, jf. afsnit 5 "Hvordan finder man de relevante formater?"

Kommentarer til systemet modtages gerne. Send en mail til Ole Schnor osc@dst.dk

2. Hvor findes formater og øvrige datafiler?

Alle SAS-formater og de øvrige datafiler er samlet i ét bibliotek med denne mappestruktur:

| Navn |
|--|
| FORMATKATALOG |
| HTML_oversigter |
| PDF_lister |
| SAS_datsaet |
| SAS-FORMATER I DANMARKS STATISTIK-filer |
| SPSS_datsaet |
| STATA_datsaet |
| Tidligere versioner |
| TXT_filer |
| Vejledning mv |
| SAS formater i Danmarks Statistik - Genvej |
| SAS-FORMATER I DANMARKS STATISTIK.html |

Når man er på forskermaskinen, så er stien til biblioteket denne:

\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik

Fra de hostede forskermaskiner, så er stien til biblioteket denne:

\\srvfsenas5\formater\SAS formater i Danmarks Statistik

3. Hvordan bruges formater i SAS-programmet?

Permanente SAS formater lagres i såkaldte kataloger. Via katalogerne kan formaterne bruges direkte i programmeringen, og for at få adgang til katalogerne bruges nogle få linjers kode.

```
libname fmt '\\srvfsenas1\data\formater\SAS formater i Danmarks  
Statistik\FORMATKATALOG';
```

```
options fmtsearch=(fmt.metadatabank fmt.statistikbank fmt.brancher  
fmt.disced fmt.disco fmt.sundhed fmt.geokoder fmt.dream  
fmt.times_personstatistik fmt.times_bbr fmt.times_motor  
fmt.times_erhvervsstatistik fmt.uddannelser);
```

Hvis man arbejder på en hosted forskermaskine, så angives i stedet:

```
libname fmt "\\srvfsenas5\formater\SAS formater i Danmarks  
Statistik\FORMATKATALOG";
```

```
options fmtsearch=(fmt.metadatabank fmt.statistikbank fmt.brancher  
fmt.disced fmt.disco fmt.sundhed fmt.geokoder fmt.dream  
fmt.times_personstatistik fmt.times_bbr fmt.times_motor  
fmt.times_erhvervsstatistik fmt.uddannelser);
```

Hvis disse linjer eksekveres i starten af SAS-programmet, så vil SAS automatisk kigge efter formaterne, når man henviser til dem. På den måde er det nemt og effektivt at bruge katalogerne direkte i procedurerne. I nedenstående eksempel

dannes en statistik med befolkningen pr. 1. januar 2011 fordelt på køn og oprindelsesland. Via format-sætningen hentes der sigende tekster for køn og oprindelsesland.

```
proc tabulate data=bef.Bef201012 missing;
format koen KOEN_KT. OPR_LAND OPR_LAND_VERDEL_SB.;
class koen OPR_LAND;
table OPR_LAND all, koen*n*f=commax12. all*n*f=commax12.;
keylabel n='Antal' all='I alt';
title 'Befolkningen pr. 1. januar 2011 fordelt på verdensdel og køn';
run;
```

Resultatet er dette:

Befolkningen pr. 1. januar 2011 fordelt på verdensdel og køn

| | Køn | | I alt |
|-------------------------|-----------|-----------|-----------|
| | 1 Mand | 2 Kvinde | |
| | Antal | Antal | Antal |
| Oprindelsesland | | | |
| 10 Uoplyst | 330 | 275 | 605 |
| 01 Danmark | 2.478.984 | 2.519.127 | 4.998.111 |
| 09 Statsløse | 430 | 287 | 717 |
| 02 EU-28 | 67.359 | 69.830 | 137.189 |
| 03 Europa udenfor EU-28 | 80.728 | 81.863 | 162.591 |
| 04 Afrika | 26.079 | 23.664 | 49.743 |
| 08 Oceanien | 1.349 | 1.077 | 2.426 |
| 06 Syd- og Mellemerika | 4.317 | 6.411 | 10.728 |
| 05 Nordamerika | 5.131 | 4.866 | 9.997 |
| 07 Asien | 91.875 | 96.646 | 188.521 |
| I alt | 2.756.582 | 2.804.046 | 5.560.628 |

Via formater kan man meget effektivt danne nye variable i et SAS-datasæt. Det sker ved at benytte PUT-funktionen, hvilket fremgår af nedenstående eksempel, hvor der dannes to nye variable.

```
data ras;
set ras.ras2009;
* Udvælg de primære beskæftigede;
if i_befolkningen_kode=1 and PRIMAER_STATUS_KODE=1 and (110 le
SOC_STATUS_KODE le 136);
* Dan branchegruppe;
stdgrp21=put (ARB_HOVED_BRA_DB07,DB07_L1L4_KT.);
* Køn som kode/tekst;
kon=put (koen,KOEN_KT.);
run;

proc tabulate data=ras missing;
class stdgrp21 kon;
table all stdgrp21='Branchegruppe', kon='Køn'*n*f=commax12.;
keylabel n='Antal' all='I alt';
title 'Primær beskæftigede i RAS 2009 på branchegrupe og køn';
run;
```

Tabellen køres direkte på de to nye variable og resultatet ses nedenfor.

Primær beskæftigede i RAS 2009 på branchegrube og køn

| | Køn | |
|---|-----------|-----------|
| | 1 Mand | 2 Kvinde |
| | Antal | Antal |
| I alt | 1.410.785 | 1.314.939 |
| Branchegrube | | |
| A Landbrug, skovbrug og fiskeri | 59.129 | 15.285 |
| B Råstofindvinding | 3.863 | 693 |
| C Industri | 212.921 | 96.735 |
| D Energiforsyning | 8.746 | 3.165 |
| E Vandforsyning og renovation | 9.593 | 1.968 |
| F Bygge og anlæg | 144.546 | 15.612 |
| G Handel | 236.635 | 183.613 |
| H Transport | 103.455 | 30.638 |
| I Hoteller og restauranter | 39.425 | 44.371 |
| J Information og kommunikation | 67.793 | 30.752 |
| K Finansiering og forsikring | 41.134 | 42.628 |
| L Ejendomshandel og udlejning | 26.144 | 14.540 |
| M Vidensservice | 85.415 | 67.437 |
| N Rejsebureauer, rengøring og anden operationel service | 64.976 | 61.404 |
| O Offentlig administration, forsvar og politi | 73.878 | 70.988 |

4. Formaterne findes i flere varianter

Der dannes i hvert enkelt tilfælde både karakterformater og numeriske formater (hvor det er muligt), fordi det ikke altid er helt entydigt, hvordan en variabel optræder i Danmarks Statistiks registre. Variablen KØN/KOEN/KON kan fx optræde både som en numerisk variabel og som en karaktervariabel alt efter hvilket register man benytter. I hvert enkelt tilfælde må man derfor undersøge hvordan den enkelte variabel er defineret. SAS kommer dog med en fejlmeddelelse, hvis ikke man bruger det korrekte formatnavn, og så er det nemt at rette til det korrekte formatnavn. Der skelnes mellem de to typer formater ved, at karakterformater starter med '\$'

I arbejdet med en variabel kan man have behov for 3 varianter af output fra et format:

- a. Koden
- b. Den forklarende tekst
- c. Koden og den forklarende tekst

Hvis man fx har kommunekoden i sit datasæt og vil danne regionen, så er der 3 muligheder for output (med Københavns kommune=101 som eksempel):

- a. 84
- b. Hovedstaden
- c. 84 Hovedstaden

Med baggrund i ovenstående findes der op til 6 forskellige varianter for et givent formatnavn:

1. **\$formatnavn_K** = Karakterformat som giver koden
2. **\$formatnavn_T** = Karakterformat som giver den forklarende tekst
3. **\$formatnavn_KT** = Karakterformat som giver kode + den forklarende tekst
4. **formatnavn_K** = Numerisk format som giver koden
5. **formatnavn_T** = Numerisk format som giver den forklarende tekst
6. **formatnavn_KT** = Numerisk format som giver kode + den forklarende tekst

5. Hvordan finder man de relevante formater?

Danmarks Statistik har udarbejdet en række HTML-sider, hvorfra man kan få en oversigt over formaterne, og det er ligeledes muligt at se, hvad de enkelte formater indeholder. På forskermaskinerne er der adgang til HTML-siderne fra denne genvej på skrivebordet:



Formater til SAS, STATA og SPSS

På de hostede forskermaskiner skal man bruge denne genvej for at få oversigten:

\\srvfsenas3\formater\ SAS formater i Danmarks Statistik\SAS-FORMATER I DANMARKS STATISTIK.html



SAS-formater til forskere og kundeservice

Koder

[Metadatabanken](#)

[Personstatistik fra TIMES \(afsluttet\)](#)

[Erhvervsstatistik fra TIMES \(afsluttet\)](#)

[Bygning og Bolig \(BBR\) fra TIMES \(afsluttet\)](#)

[Motorregistret fra TIMES \(afsluttet\)](#)

Klassifikationer og grupperinger

[Grupperinger i Statistikbanken](#)

[Uddannelser \(DISCED og ISCED2011\)](#)

[Uddannelser \(DISCED - Hovedområder med historik\)](#)

[Uddannelser \(Forspalter\)](#)

[Uddannelser \(ISCED1997\)](#)

[Uddannelser \(DUN\)](#)

[Brancher og standardgrupperinger](#)

[Brancher og NACE](#)

[DISCO](#)

[Sundhed](#)

[Dream databasen](#)

Geokoder

[Kommuner, Regioner/Amter og Landsdele](#)

Vejledning, allokering og andre filtyper

[Vejledning](#)

[Allokering af formatbibliotek \(forskermaskiner\)](#)

[Allokering af formatbibliotek \(hostede forskermaskiner\)](#)

[Allokering af formatbibliotek \(Lovmodel\)](#)

[Allokering af formatbibliotek \(Dst. medarbejdere\)](#)

[Her findes SAS formater for Dst. medarbejdere](#)

[Indlæs formater i WPS \(hostede forskermaskiner\)](#)

[SPSS](#)

[STATA](#)

[PDF](#)

[TXT](#)

Kodeformaterne hentes via Danmarks Statistiks **Metadatabank**, der opdateres løbende. Kodeformater baseret på **TIMES** er afsluttet og opdateres ikke, men formatkatalogerne findes fortsat, så hvis man ønsker af benytte dem. Se hvordan, i næste afsnit.

Grupperinger i Statistikbanken indeholder nogle af de centrale grupperinger i Statistikbanken. De øvrige **Klassifikationer** og **Geokoder** indeholder hierarkier/standardgrupperinger for de mest anvendte variable, og svarer til de grupperinger der anvendes af Danmarks Statistik. Her er tale om standard grupperinger for Uddannelser, Brancher, Disco og Sundhed. **Geokoder** indeholder grupperinger af Kommuner, Regioner, Amter og Landsdele.

Metadatabanken

Oversigterne er inddelt efter Grunddatasæt i Forskningservice og variabel.

Koder i Metadatabanken for variable i Forskningservice - opdateret 18jan23
SAS-formatkatalog: Metadatabank

| FSE Grunddatasæt | Navn på grunddatasæt | Variabel Navn | Label for variabel | Karakter - tekst | Karakter - kode og tekst |
|------------------|--|--------------------------|---|-------------------------|--------------------------|
| AEFB | Ældreokumentation forebyggende hjemmebesøg | IMP_MARK | Markering for imputering af borgere | SIMP_MARK_T | SIMP_MARK_KT |
| AEFV | Ældreokumentation visiteret hjemmehjælp frt valg | LEV_TYPE_VIS | Leverandør type for visiteret hjemmehjælp frt valg | SLEV_TYPE_VIS_T | SLEV_TYPE_VIS_KT |
| AEHJSP | Ældreokumentation - hjemmesygepleje i eget hjem | IMP_MARK | Markering for imputering af borgere | SIMP_MARK_T | SIMP_MARK_KT |
| AEIH | Ældreokumentation leveret hjemmehjælp | IMP_MARK | Markering for imputering af borgere | SIMP_MARK_T | SIMP_MARK_KT |
| | | LEV_TYPE_LEV | Leverandørtipe for leveret hjemmehjælp frt valg | SLEV_TYPE_LEV_T | SLEV_TYPE_LEV_KT |
| AEPB | Ældreokumentation visiteret hjemmehjælp plejebolig | IMP_MARK | Markering for imputering af borgere | SIMP_MARK_T | SIMP_MARK_KT |
| AEPI | Ældreokumentation borgere på plejehjem bolig (imputeret) | IMP_MARK | Markering for imputering af borgere | SIMP_MARK_T | SIMP_MARK_KT |
| AEREHAB | Ældreokumentation - Rehabilitering | GOODKEND_AAR | Godkendt år | SGOODKEND_AAR_T | SGOODKEND_AAR_KT |
| AETR | Ældreokumentation genoptræning vedligeholdelsestræning | TRAEN_ART | Træningsydelse (genoptræning/vedligeholdelsestræning) | STRAEN_ART_T | STRAEN_ART_KT |
| AKAS | A-kasse medlemmer | AKASSE_ID | Entydig identifikation af a-kasse/tilshørsforhold | SAKASSE_NAVN_T | SAKASSE_NAVN_KT |
| | | EF_MARK | EF_MARK | SMODUL_AKASSE_EF_MARK_T | SMODUL_AKASSE_EF_MARK_KT |
| | | FORSIKRINGSKATEGORI_KODE | Kode for forsikringskategorien | SHKD_MODULE_FORSKAT_T | SHKD_MODULE_FORSKAT_KT |
| AKM | Arbejdsklassifikationsmodul | BESKST | Beskæftigelsesstatus 1980 til 2001 (Indkomst/AKM) | SBESKST_T | SBESKST_KT |
| | | BESKST02 | Beskæftigelsesstatus fra 2002 (Indkomst/AKM) | SBESKST02_T | SBESKST02_KT |
| | | BESKST13 | Kode for personens væsentligste indkomstkilde (INDKOMST/AKM) | SBESKST02_T | SBESKST02_KT |
| | | BRANCHE_77 | Dansk branchekode 1977 til 1993 (AKM) | SDST_77_T | SDST_77_KT |
| | | DISCO08_ALLE_INDK_13 | Fagkode for væsentligste beskæftigelse i året- fra 2010 (AKM) | SAKM_DISCO_T | SAKM_DISCO_KT |
| | | DISCO08_LOEN_INDK | Fagkode for væsentligste lønmodtagerbeskæftigelse i året. (AKM) | SAKM_DISCO_T | SAKM_DISCO_KT |

Hvis der klikkes på de enkelte formatnavne, så får man adgang til en stump SAS-kode der angiver formatet via PROC FORMAT. Det giver en information om formatets indhold, og hvis man ønsker at gemme/ændre formatet, så kan man klippe koden ud og gemme i sit SAS-program. I nedenstående eksempel er karakterformatet **\$HUSTYPE_T** hentet ind i SAS. Når dette format benyttes, så er output den sigende tekst uden kode. Hvis man ikke skal ændre i et eksisterende format, så er det ikke nødvendigt at hente koden ind i SAS-programmet. Man kan i stedet pege direkte ned i SAS-kataloget jf. afsnit 3.

```

Proc Format;
value $HUSTYPE_T
'1'='Enlig mand'
'2'='Enlig kvinde'
'3'='Egtepar'
'4'='Par i øvrigt'
'5'='Ikke hjemmeboende børn (under 18 år)'
'6'='Andre husstande bestående af flere familier'
;
run;

```

Mange af formater findes både i Metadatabanken og i TIMES. SAS 'kigger' i katalogerne i den rækkefølge de er angivet i FMTSEARCH. I denne kode tages formatet fra Metadatabanken, hvis det findes der.

```

options fmtsearch=(fmt.metadatabank fmt.statistikbank fmt.brancher
fmt.disced fmt.disco fmt.sundhed fmt.geokoder fmt.dream
fmt.times_personstatistik fmt.times_bbr fmt.times_motor
fmt.times_erhvervsstatistik fmt.uddannelser);

```

Hvis man i stedet vil benytte formatet fra TIMES fremfor Metadatabanken, så kræver det en ændring i koden med FMTSEARCH, hvor 'fmt.metadatabank' blot skal sættes til sidst.

```
options fmtsearch=(fmt.statistikbank fmt.brancher fmt.disced
fmt.disco fmt.sundhed fmt.geokoder fmt.dream
fmt.times_personstatistik fmt.times_bbr fmt.times_motor
fmt.times_erhvervsstatistik fmt.uddannelser fmt.metadatabank);
```

Uddannelserne

Fra og med 2015 bruger Uddannelsesstatistikken klassifikationen DISCED-15. Denne klassifikation erstatter de tidligere nomenklaturer 'Forspalte', 'ISCED1997' og 'DUN'.

Bemærk, at formaterne til 'Forspalter', 'ISCED1997' og 'DUN' ligger i SAS-formatkataloget '**Uddannelser**' mens DISCED-15 formaterne ligger i SAS-formatkataloget '**DISCED**'.

Forspalte var den foretrukne danske klassifikation før 2015. Dansk Uddannelsesnomenklatur (DUN) er en gruppering af uddannelseskompetencer efter deres faglige indhold. DUN bruges primært af virksomheder til en uddannelsesgruppering af deres ansatte.

Hovedområde-dimensionen i DISCED-15 afspejler en fordeling af uddannelserne på de lovregulerede hovedområder for de videregående uddannelser og på indgange for erhvervsuddannelserne. Dermed følger Hovedområderne strukturen i det danske uddannelsessystem.

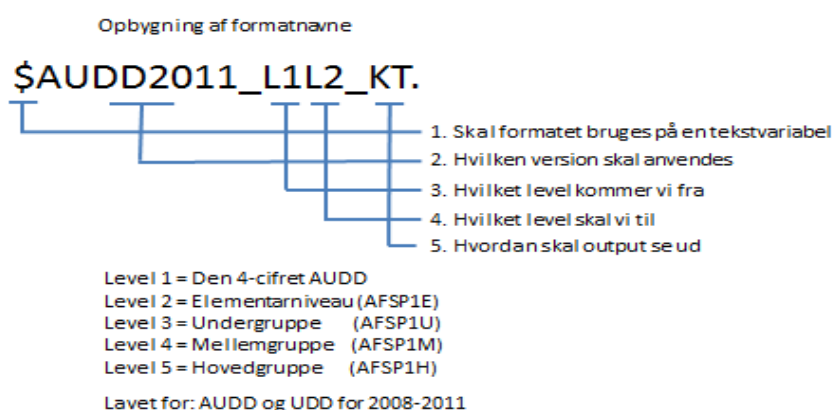
Uddannelsesstypen i DISCED-15 gør det muligt at differentiere uddannelserne i det danske uddannelsessystem efter type af uddannelse, uafhængigt af uddannelsernes niveau, fag- eller hovedområde. Uddannelsesstypen tillader en underopdeling af klassifikationens øvrige dimensioner på fx ordinære og voksenuddannelser.

Uddannelsesniveau i DISCED-15 er en dimension, som sikrer overensstemmelse imellem niveauklassificeringen i det danske uddannelsessystem og den internationale uddannelsesklassifikation, ISCED2011. Uddannelsesniveautet i DISCED-15 er opbygget efter samme grundlæggende principper som Level i ISCED2011, hvormed der sikres en entydig sammenhæng.

Fagområde-dimensionen i DISCED-15 sikrer en fuldstændig sammenlignelighed imellem national og international uddannelsesklassifikation. Klassificeringen af uddannelserne indenfor de enkelte fagområder sker uafhængigt af uddannelsernes niveau. Det grundlæggende princip i opbygningen af fagområderne følger ideen om, hvilket beskæftigelsesfunktions-/brancheområde uddannelsen orienterer sig mod med henblik på senere beskæftigelse.

I Elevregistret (**KOTRE**) skelnes mellem kode for igangværende/afbrudt uddannelse (UDD) og for afsluttet uddannelse (AUDD). Bemærk derfor, at der findes uddannelsesformater til begge typer af kode.

Navngivningen af formaterne følger denne systematik:



Man kan derved komme fra det mest detaljerede niveau (den 4 cifrede uddannelseskode) som typisk er den variabel man har i sit grunddatasæt, og til alle andre niveauer.

Her ses et eksempel med højeste fuldførte uddannelse (**UDDA**), hvor formatet **AUDD_HOVED_L1L5_KT** omsætter den 4 cifrede uddannelseskode (HFAUDD) til Hovedområde overgruppeniveau.

```
proc tabulate data=udda.udda201809 missing order=formatted;
format hfaudd AUDD_HOVED_L1L5_KT.;
class HFAUDD;
table HFAUDD all, n*f=commax12.;
keylabel n='Antal' all='I alt';
title 'Befolkningens højeste fuldførte uddannelse 1. oktober 2018';
run;
```

Giver dette resultat:

Befolkningens højeste fuldførte uddannelse 1. oktober 2018

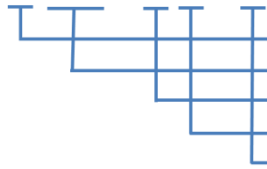
| | Antal |
|---|-----------|
| Højste fuldførte uddannelse | |
| . | 883.563 |
| 10 Grundskole | 1.478.863 |
| 20 Gymnasiale uddannelser | 441.941 |
| 30 Erhvervsfaglige uddannelser | 1.489.271 |
| 35 Adgangsgivende uddannelsesforløb | 4.028 |
| 40 Korte videregående uddannelser, KVU | 224.953 |
| 50 Mellemlange videregående uddannelser, MVU | 707.085 |
| 60 Bacheloruddannelser, BACH | 98.776 |
| 70 Lange videregående uddannelser, LVU | 439.787 |
| 80 Ph.d. og forskeruddannelser | 37.748 |
| I alt | 5.806.015 |

Da Uddannelsesstatistikken opdateres hvert år med de seneste koder også bagud i tid, så man vil som regel kun have brug for den seneste version af formaterne.

Brancherne

Formaterne for brancher kan bruges til gruppering af data til både den internationale nomenklatur NACE og til Danmarks Statistiks standardgrupperinger. For at gøre navngivningen af formaterne uafhængig af versionen af branchenomenklaturen og antal grupper i standardgrupperingerne, er der lavet systematiske navne til formaterne:

\$DB07_L1L2_KT.



1. Skal formatet bruges på en tekstvariabel
2. Hvilken version skal anvendes
3. Hvilket level kommer vi fra
4. Hvilket level skal vi til
5. Hvordan skal output se ud

Det fremgår af tabellen, hvilken sammenhæng der er mellem niveau og Danmarks Statistiks standardgrupperinger for de 4 nomenklaturer.

| Nomenklatur | Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
|-------------|------------------|---------------|---------------|---------------|---------------|
| DB07 | 6-cifret br.kode | 127 grp. | 38 grp. | 21 grp. | 10 grp. |
| DB03 | 6-cifret br.kode | 111 grp. | 53 grp. | 27 grp. | 9 grp. |
| DB93 | 6-cifret br.kode | 111 grp. | 53 grp. | 27 grp. | 9 grp. |
| DSE77 | 5-cifret br.kode | 4-cifret grp. | 3-cifret grp. | 2-cifret grp. | 1-cifret grp. |

HTML-oversigten er inddelt efter niveau og version af branchenomenklaturen:

| Niveau | Version af branche-nomenklaturen | Fra - Til | Karakter - kode | Karakter - tekst | Karakter - kode og tekst | Numerisk - kode | Numerisk - tekst | Numerisk - kode og tekst |
|--------|----------------------------------|--------------------------------|-----------------|------------------|--------------------------|-----------------|------------------|--------------------------|
| L1L1 | DB07 | DB07 til DB07 | Ikke relevant | \$DB07_L1L1_T | \$DB07_L1L1_KT | Ikke relevant | DB07_L1L1_T | DB07_L1L1_KT |
| | DB03 | DB03 til DB03 | Ikke relevant | \$DB03_L1L1_T | \$DB03_L1L1_KT | Ikke relevant | DB03_L1L1_T | DB03_L1L1_KT |
| | DB93 | DB93 til DB93 | Ikke relevant | \$DB93_L1L1_T | \$DB93_L1L1_KT | Ikke relevant | DB93_L1L1_T | DB93_L1L1_KT |
| | DSE77 | DSE77 til DSE77 | Ikke relevant | \$DSE77_L1L1_T | \$DSE77_L1L1_KT | Ikke relevant | DSE77_L1L1_T | DSE77_L1L1_KT |
| L1L2 | DB07 | DB07 til STD 127-GRP | \$DB07_L1L2_K | \$DB07_L1L2_T | \$DB07_L1L2_KT | DB07_L1L2_K | DB07_L1L2_T | DB07_L1L2_KT |
| | DB03 | DB03 til STD 111-GRP | \$DB03_L1L2_K | \$DB03_L1L2_T | \$DB03_L1L2_KT | DB03_L1L2_K | DB03_L1L2_T | DB03_L1L2_KT |
| | DB93 | DB93 til STD 111-GRP | \$DB93_L1L2_K | \$DB93_L1L2_T | \$DB93_L1L2_KT | DB93_L1L2_K | DB93_L1L2_T | DB93_L1L2_KT |
| | DSE77 | DSE77 til 4-cifret undergruppe | \$DSE77_L1L2_K | \$DSE77_L1L2_T | \$DSE77_L1L2_KT | DSE77_L1L2_K | DSE77_L1L2_T | DSE77_L1L2_KT |
| L1L3 | DB07 | DB07 til STD 38- | \$DB07_L1L3_K | \$DB07_L1L3_T | \$DB07_L1L3_KT | DB07_L1L3_K | DB07_L1L3_T | DB07_L1L3_KT |

For den internationale nomenklatur NACE findes kun grupperingen i forhold til DB07. Formaterne findes med både *danske* og *engelske* tekster.

| Nomenklatur | Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
|-------------|------------------|----------------------|------------------|----------------------|------------------------|
| DB07 | 6-cifret br.kode | 4-cifret Undergruppe | 3- cifret Gruppe | 2-cifret Hovedgruppe | 1-cifret Hovedafdeling |

DISCO (Arbejdsfunktioner)

DISCO er en klassifikation over arbejdsfunktioner og således et redskab til at organisere jobs eller stillinger inden for en organisation, en branche eller et område, i nogle klart definerede grupper, i forhold til de opgaver der udføres i jobbet eller stillingen. Anvendelsen af DISCO gør det muligt at sammenligne personer med samme arbejdsfunktion - uanset medarbejderens formelle titel og uddannelse.

DSICO08

Fra og med 2010 data er det nomenklaturen for DISCO08 der gælder. Der er tale om disse levels i navngivningen af formaterne:

| Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
|--------------------|----------------------|-------------------------|---------------------|----------------------|
| 6-cifret discokode | 4-cifret Undergruppe | 3- cifret Mellem-gruppe | 2-cifret Overgruppe | 1-cifret Hovedgruppe |

De tidligere DISCO koder

Før 2010 er alle koder i nomenklaturerne DISCO88, DISCO96 og DISCOLØN04 samlet i ét fælles format. Da koderne før 2010 ikke nødvendigvis er 6 cifrede, så giver det ikke mening at danne formater for de forskellige niveauer/levels. Bemærk, at der i nogle registre før 2010 kan der optræde ikke valide koder og som derved ikke fanges af formatet. Formatet findes nederst på oversigten.

| Version | Niveau | Fra - Til | Karakter - kode | Karakter - tekst | Karakter - kode og tekst | Numerisk - kode | Numerisk - tekst | Numerisk - kode og tekst |
|---------------------------------|--------|---|----------------------------------|----------------------------------|-----------------------------------|--------------------------------|--------------------------------|---------------------------------|
| DISCO08 | L1L1 | Disco08 til Disco08 | Ikke relevant | \$DISCO08 L1L1 T | \$DISCO08 L1L1 KT | Ikke relevant | DISCO08 L1L1 T | DISCO08 L1L1 KT |
| | L1L2 | Disco08 til 4-cifret undergruppe | \$DISCO08 L1L2 K | \$DISCO08 L1L2 T | \$DISCO08 L1L2 KT | DISCO08 L1L2 K | DISCO08 L1L2 T | DISCO08 L1L2 KT |
| | L1L3 | Disco08 til 3-cifret mellemgruppe | \$DISCO08 L1L3 K | \$DISCO08 L1L3 T | \$DISCO08 L1L3 KT | DISCO08 L1L3 K | DISCO08 L1L3 T | DISCO08 L1L3 KT |
| | L1L4 | Disco08 til 2-cifret overgruppe | \$DISCO08 L1L4 K | \$DISCO08 L1L4 T | \$DISCO08 L1L4 KT | DISCO08 L1L4 K | DISCO08 L1L4 T | DISCO08 L1L4 KT |
| | L1L5 | Disco08 til 1-cifret hovedgruppe | \$DISCO08 L1L5 K | \$DISCO08 L1L5 T | \$DISCO08 L1L5 KT | DISCO08 L1L5 K | DISCO08 L1L5 T | DISCO08 L1L5 KT |
| | L2L2 | 4-cifret undergruppe til 4-cifret undergruppe | Ikke relevant | \$DISCO08 L2L2 T | \$DISCO08 L2L2 KT | Ikke relevant | DISCO08 L2L2 T | DISCO08 L2L2 KT |
| | L2L3 | 4-cifret undergruppe til 3-cifret mellemgruppe | \$DISCO08 L2L3 K | \$DISCO08 L2L3 T | \$DISCO08 L2L3 KT | DISCO08 L2L3 K | DISCO08 L2L3 T | DISCO08 L2L3 KT |
| | L2L4 | 4-cifret undergruppe til 2-cifret overgruppe | \$DISCO08 L2L4 K | \$DISCO08 L2L4 T | \$DISCO08 L2L4 KT | DISCO08 L2L4 K | DISCO08 L2L4 T | DISCO08 L2L4 KT |
| | L2L5 | 4-cifret undergruppe til 1-cifret hovedgruppe | \$DISCO08 L2L5 K | \$DISCO08 L2L5 T | \$DISCO08 L2L5 KT | DISCO08 L2L5 K | DISCO08 L2L5 T | DISCO08 L2L5 KT |
| | L3L3 | 3-cifret mellemgruppe til 3-cifret mellemgruppe | Ikke relevant | \$DISCO08 L3L3 T | \$DISCO08 L3L3 KT | Ikke relevant | DISCO08 L3L3 T | DISCO08 L3L3 KT |
| | L3L4 | 3-cifret mellemgruppe til 2-cifret overgruppe | \$DISCO08 L3L4 K | \$DISCO08 L3L4 T | \$DISCO08 L3L4 KT | DISCO08 L3L4 K | DISCO08 L3L4 T | DISCO08 L3L4 KT |
| | L3L5 | 3-cifret mellemgruppe til 1-cifret hovedgruppe | \$DISCO08 L3L5 K | \$DISCO08 L3L5 T | \$DISCO08 L3L5 KT | DISCO08 L3L5 K | DISCO08 L3L5 T | DISCO08 L3L5 KT |
| | L4L4 | 2-cifret overgruppe til 2-cifret overgruppe | Ikke relevant | \$DISCO08 L4L4 T | \$DISCO08 L4L4 KT | Ikke relevant | DISCO08 L4L4 T | DISCO08 L4L4 KT |
| | L4L5 | 2-cifret overgruppe til 1-cifret hovedgruppe | \$DISCO08 L4L5 K | \$DISCO08 L4L5 T | \$DISCO08 L4L5 KT | DISCO08 L4L5 K | DISCO08 L4L5 T | DISCO08 L4L5 KT |
| | L5L5 | 1-cifret hovedgruppe til 1-cifret hovedgruppe | Ikke relevant | \$DISCO08 L5L5 T | \$DISCO08 L5L5 KT | Ikke relevant | DISCO08 L5L5 T | DISCO08 L5L5 KT |
| Indkomster, RAS og lønstatistik | Alle | Format indeholder alle niveauer | Ikke relevant | \$DISCO08 T | \$DISCO08 KT | Ikke relevant | Ikke relevant | Ikke relevant |

Sundhed

Disse formater bygger på helt overvejende på Sundhedsvæsenets Klassifikationssystem (SKS). Værdisæt i dette system opdateres én gang i kvartalet. SAS- formaterne opdateres i takt med, at der kommer nye versioner.

Formaterne er baseret på SKS-klassifikationer fra SSI pr. 17062019
 DOSO og INDO er baseret på data fra SSI pr. 05jan2016
 SPECIALE i Sygesikring er hentet i Danmarks Statistik (TIMES)
 Formater for øvrige sygesikringsvariable findes under TIMES variable/Personstatistik

SAS-formatkatalog: Sundhed

| Område | Klassifikation (forkortelse) | Niveau | Fra - Til | Karakter - kode | Karakter - tekst | Karakter - kode og tekst |
|---------------------------|------------------------------|--------|--|--------------------------------|--------------------------------------|---------------------------------------|
| 01. Sygdomme | ICD10 | L1L1 | Sygdomme mv., hovedgruppe D | Ikke relevant | \$ICD10_L1L1_T | \$ICD10_L1L1_KT |
| | | L1L2 | ICD10 til 99 gruppering | \$ICD10_L1L2_K | \$ICD10_L1L2_T | \$ICD10_L1L2_KT |
| | | L1L3 | ICD10 til 23 gruppering | \$ICD10_L1L3_K | \$ICD10_L1L3_T | \$ICD10_L1L3_KT |
| | ICD8 | L2L2 | 99 Gruppering | Ikke relevant | \$ICD10_L2L2_T | \$ICD10_L2L2_KT |
| | | L3L3 | 23 Gruppering | Ikke relevant | \$ICD10_L3L3_T | \$ICD10_L3L3_KT |
| | | L1L1 | ICD8 | Ikke relevant | \$ICD8_L1L1_T | \$ICD8_L1L1_KT |
| 02. Operationer | OPR | L1L1 | ICD8 til 99 gruppering | \$ICD8_L1L2_K | \$ICD8_L1L2_T | \$ICD8_L1L2_KT |
| | | L1L3 | ICD8 til 23 gruppering | \$ICD8_L1L3_K | \$ICD8_L1L3_T | \$ICD8_L1L3_KT |
| | | L1L1 | Operationer, hovedgruppe K | Ikke relevant | \$OPR_L1L1_T | \$OPR_L1L1_KT |
| 03. Lægemidler | OPR95 | L1L1 | Operation og behandling (1971-1995) | Ikke relevant | \$OPR95_L1L1_T | \$OPR95_L1L1_KT |
| | | L1L1 | Lægemidler, hovedgruppe M | Ikke relevant | \$ATC_L1L1_T | \$ATC_L1L1_KT |
| | | L1L1 | Dosering | Ikke relevant | \$DOSO_L1L1_T | \$DOSO_L1L1_KT |
| 04. Ulykker | INDO | L1L1 | Indikation | Ikke relevant | \$INDO_L1L1_T | \$INDO_L1L1_KT |
| | | L1L1 | Ulykker, hovedgruppe E | Ikke relevant | \$ULY_L1L1_T | \$ULY_L1L1_KT |
| 05. Undersøgelser | UND | L1L1 | Undersøgelser, hovedgruppe U | Ikke relevant | \$UND_L1L1_T | \$UND_L1L1_KT |
| 06. Behandlinger | BEH | L1L1 | Behandlinger og pleje, hovedgruppe B | Ikke relevant | \$BEH_L1L1_T | \$BEH_L1L1_KT |
| 07. Anæstesi | ANE | L1L1 | Anæstesi, hovedgruppe N | Ikke relevant | \$ANE_L1L1_T | \$ANE_L1L1_KT |
| 08. Funktionsevne mv. | ICF | L1L1 | Funktionsevne og helbred, hovedgruppe F | Ikke relevant | \$ICF_L1L1_T | \$ICF_L1L1_KT |
| 09. Klinisk fysiologi mv. | KFN | L1L1 | Klinisk fysiologi og nuklearmedicin, hovedgruppe W | Ikke relevant | \$KFN_L1L1_T | \$KFN_L1L1_KT |
| 10. Sygesikring | SPECIALE | L1L1 | 6 cifret specialekode | Ikke relevant | \$SPECIALE_L1L1_T | \$SPECIALE_L1L1_KT |
| | | L2L2 | 4 cifret specialekode | Ikke relevant | \$SPECIALE_L2L2_T | \$SPECIALE_L2L2_KT |
| | | L3L3 | 3 cifret specialekode | Ikke relevant | \$SPECIALE_L3L3_T | \$SPECIALE_L3L3_KT |
| | | L4L4 | 2 cifret specialekode | Ikke relevant | \$SPECIALE_L4L4_T | \$SPECIALE_L4L4_KT |
| 11. Hospitaler | LPRSPECIALE | L1L1 | Speciale | Ikke relevant | \$LPRSPECIALE_L1L1_T | \$LPRSPECIALE_L1L1_KT |
| 12. Øvrige | ADM | L1L1 | Administrative forhold, hovedgruppe A | Ikke relevant | \$ADM_L1L1_T | \$ADM_L1L1_KT |
| | | L1L1 | Tillægskoder, hovedgruppe Z | Ikke relevant | \$TIL_L1L1_T | \$TIL_L1L1_KT |

For sygdomme og operationer findes også formater for tidligere nomenklaturer (ICD8 og Operationer 1995). Men bortset fra det, så indeholder formaterne ikke historik. Så hvis en værdikode får ny betydning, vil formatet kun indeholde den seneste betydning.

SPECIALE i Sygesikringsstatistikken er undtagelsesvis hentet i Danmarks Statistiks dokumentationssystemer. Men er medtaget under 'Sundhed' da variabelen kan grupperes i efter både 2, 3, 4 og 6 cifre.

Geokoder

Med formaterne for geokoder kan man ud fra kommunekoder danne regioner/amter og landsdele. Med kommunesammenlægningerne siden 2003 og kommunalreformen i 2007 har det været nødvendigt at danne 4 versioner af formaterne (V1-V4). Navngivningen af formaterne følger denne systematik:

\$KOMREG_V4_KT.



Det fremgår af tabellen, hvilken sammenhæng der er mellem version og historik.

| Version | Historik |
|---------|--|
| V1 | Gældende før 1. jan. 2003 |
| V2 | Gældende fra 1. jan. 2003 til 31. dec. 2005. Kommunerne på Bornholm er slået sammen til én kommune |
| V3 | Gældende fra 1. jan. 2006 til 31. dec. 2006. Marstal og Ærøskøbing sammenlægges til én kommune |
| V4 | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner. Amter nedlægges og regioner oprettes. |

HTML-oversigten er inddelt efter geogrupping og version af formatet:

| Niveau | Fra - Til | Karakter - kode | Karakter - tekst | Karakter - kode og tekst | Numerisk - kode | Numerisk - tekst | Numerisk - kode og tekst |
|--------------------|--|-----------------|------------------|--------------------------|-----------------|------------------|--------------------------|
| Amt | Gældende før 1. jan. 2003 | Ikke relevant | \$AMT_V1_T | \$AMT_V1_KT | Ikke relevant | AMT_V1_T | AMT_V1_KT |
| | Gældende fra 1. jan. 2003 til 31. dec. 2005. Kommunerne på Bornholm er slået sammen til én kommune | Ikke relevant | \$AMT_V2_T | \$AMT_V2_KT | Ikke relevant | AMT_V2_T | AMT_V2_KT |
| | Gældende fra 1. jan. 2006 til 31. dec. 2006. Marstal og Ærøskøbing sammenlægges til én kommune | Ikke relevant | \$AMT_V3_T | \$AMT_V3_KT | Ikke relevant | AMT_V3_T | AMT_V3_KT |
| Kommune | Gældende før 1. jan. 2003 | Ikke relevant | \$KOM_V1_T | \$KOM_V1_KT | Ikke relevant | KOM_V1_T | KOM_V1_KT |
| | Gældende fra 1. jan. 2003 til 31. dec. 2005. Kommunerne på Bornholm er slået sammen til én kommune | Ikke relevant | \$KOM_V2_T | \$KOM_V2_KT | Ikke relevant | KOM_V2_T | KOM_V2_KT |
| | Gældende fra 1. jan. 2006 til 31. dec. 2006. Marstal og Ærøskøbing sammenlægges til én kommune | Ikke relevant | \$KOM_V3_T | \$KOM_V3_KT | Ikke relevant | KOM_V3_T | KOM_V3_KT |
| Kommune til amt | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner | Ikke relevant | \$KOM_V4_T | \$KOM_V4_KT | Ikke relevant | KOM_V4_T | KOM_V4_KT |
| | Gældende før 1. jan. 2003 | \$KOMAMT_V1_K | \$KOMAMT_V1_T | \$KOMAMT_V1_KT | KOMAMT_V1_K | KOMAMT_V1_T | KOMAMT_V1_KT |
| | Gældende fra 1. jan. 2003 til 31. dec. 2005. Kommunerne på Bornholm er slået sammen til én kommune | \$KOMAMT_V2_K | \$KOMAMT_V2_T | \$KOMAMT_V2_KT | KOMAMT_V2_K | KOMAMT_V2_T | KOMAMT_V2_KT |
| Kommune til region | Gældende fra 1. jan. 2006 til 31. dec. 2006. Marstal og Ærøskøbing sammenlægges til én kommune | \$KOMAMT_V3_K | \$KOMAMT_V3_T | \$KOMAMT_V3_KT | KOMAMT_V3_K | KOMAMT_V3_T | KOMAMT_V3_KT |
| | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner | \$KOMLAN_V4_K | \$KOMLAN_V4_T | \$KOMLAN_V4_KT | KOMLAN_V4_K | KOMLAN_V4_T | KOMLAN_V4_KT |
| | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner | \$KOMREG_V4_K | \$KOMREG_V4_T | \$KOMREG_V4_KT | KOMREG_V4_K | KOMREG_V4_T | KOMREG_V4_KT |
| Kommunegruppe | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner | \$KOMGRP_V4_K | \$KOMGRP_V4_T | \$KOMGRP_V4_KT | KOMGRP_V4_K | KOMGRP_V4_T | KOMGRP_V4_KT |
| | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner | Ikke relevant | \$LAN_V4_T | \$LAN_V4_KT | Ikke relevant | LAN_V4_T | LAN_V4_KT |
| Region | Gældende fra 1. jan. 2007. Kommunesammenlægninger til 98 kommuner | Ikke relevant | \$REG_V4_T | \$REG_V4_KT | Ikke relevant | REG_V4_T | REG_V4_KT |

DREAM databasen

Beskæftigelsesministeriets forløbsdatabase DREAM, administreres af Styrelsen for Arbejdsmarked og Rekruttering (STAR). Databasen indeholder beskæftigelsesoplysninger og andre grundlæggende personoplysninger. Formaterne er modtaget fra STAR. Dokumentation af data og koder kan findes på denne side:

https://www.dst.dk/da/TilSalg/Forskningservice/Data/Andre_Styrelser

Bemærk, at STAR har værdisæt for fx Køn og Herkomst som ikke flugter Danmarks Statistiks værdisæt. Derfor findes disse som særskilte DREAM formater.

Variable og formater i DREAM databasen - opdateret 06FEB23 SAS-formatkatalog: 'Dream'

| Emne | Variabel | Label for variabel | Karakter | Numerisk |
|-----------------|------------------|----------------------|-------------------------------------|-----------------------------------|
| A-kasse | AKyyyy | A-kassekode | \$DREAM_AKASSE_KT | |
| | | | \$DREAM_AKASSE_I | |
| Herkomst | HERKOMST | Herkomst | \$DREAM_HERKOMST_KT | |
| | | | \$DREAM_HERKOMST_I | |
| Køn | KON | Køn | \$DREAM_KON_KT | |
| | | | \$DREAM_KON_I | |
| Matchkategori | MATCH_xKVyy | Matchkatgegori | \$DREAM_MATCH_KT | |
| | | | \$DREAM_MATCH_I | |
| Statsborgerskab | STATSBORGERSKAB | Statsborgerskab,kort | \$DREAM_STATKORT_KT | |
| | | | \$DREAM_STATKORT_I | |
| Virksomhedsform | VIRKFORM_yyyy_mm | Virksomhedsform | | DREAM_VIRKFORM_KT |
| | | | | DREAM_VIRKFORM_I |
| Visitation | VISIT_xKVyy | Visitationskode | \$DREAM_VISIT_KT | |
| | | | \$DREAM_VISIT_I | |
| Ydelse | Y_yyuu | Ydelsesgruppering | \$DREAM_YDELSE_GRP | |
| | | Ydelseskode | \$DREAM_YDELSE_KT | |
| | | | \$DREAM_YDELSE_I | |