

*Brugervejledning til  
Høj kvalitetsdokumentationen på  
Danmarks Statistiks hjemmeside*

## **Forord**

Denne vejledning beskriver baggrunden for begreber og sammenhænge i Danmarks Statistiks dokumentationssystem på Internettet.

De variabler Danmarks Statistik har i sine registre er for en stor dels vedkommende beskrevet og dokumenteret i softwaresystemet Times. Samtidig med udviklingen af fjerde version, Times4, er visse variabler blevet høj kvalitetsdokumenteret, dvs. at der er koblet yderligere oplysninger til variabelen end de som er beskrevet for alle variabler.

Høj kvalitetsdokumentationen er udarbejdet i samarbejde mellem Danmarks Statistik og Det Koordinerende Organ for Registerforskning. Kvalitetskontrollen er udført af eksterne bedømmere udpeget af Det Koordinerende Organ for Registerforskning.

I denne vejledning gennemgås de forskellige komponenter af dokumentationssystemet.

Danmarks Statistik  
Marts 2009

## **Indhold**

<b>Forord .....</b>	<b>2</b>
<b>Indledning.....</b>	<b>4</b>
<b>Dokumentationssystemet.....</b>	<b>4</b>

## Indledning

Høj kvalitetsvariable er skabt som følge af en forespørgsel fra forskere, der havde behov for en detaljeret dokumentation af de variable der er mest brugt. Projektet blev startet i januar 2006, og de første høj kvalitetsvariable blev præsenteret på [www.dst.dk](http://www.dst.dk) i april 2006.

## Dokumentationssystemet

<i>Vejen til den enkelte variabel</i>	<p>På Danmarks Statistiks hjemmeside er der et link til "Forskningservice". Herfra kan der klikkes videre til link med "Høj kvalitetsdokumentation".</p> <p>Under høj kvalitetsdokumentation åbnes en liste med Høj kvalitetssemnegrupper. Ved at klikke videre på et emne fremkommer en liste med de høj kvalitetsvariable der hører til emnet inddelt i statistikområder. Der kan klikkes på variabelnavnene og dette åbner siden, der beskriver den enkelte variabel. Der kan ligeledes klikkes på statistikområdet og dette åbner en boks, der indeholder en beskrivende tekst om det pågældende statistikområde.</p> <p>Der er tre kolonner på siden, hvoraf kun den i midten indeholder dokumentationen. Kolonnerne til højre og venstre indeholder generelle oplysninger på Danmarks Statistiks hjemmeside. I denne beskrivelse gennemgås kun elementerne i den midterste kolonne.</p>
<i>To slags emnegrupper</i>	<p>Variabler inddeles traditionelt efter deres registertilhørsforhold i Danmarks Statistik. Denne emneopdeling omfatter samtlige variable i Times4 (Emnegruppe DST). Emneopdelingen kan dog virke ulogisk på udenforstående og derfor er der opbygget en alternativ emnegruppering (høj kvalitetssemnegrupper), der tager udgangspunkt i, hvad den enkelte variabel beskriver. Denne emneopdeling er kun implementeret på høj kvalitetsvariable.</p>
<i>Statistikområde</i>	<p>Statistikområdet inddeler variable i mindre portioner indenfor hver høj kvalitetssemnegruppe. Der vil således typisk være flere statistikområder indenfor hver høj kvalitetssemnegruppe. Statistikområderne er dannet på baggrund af de grupperinger der forekommer i statistikbanken indenfor DST emnegrupperne. Dette er valgt for, at der er en sammenhæng imellem præsentationen af statistik og dokumentation på Danmarks Statistiks hjemmeside.</p>
<i>Beskrivelse af den enkelte variabel</i>	<p>En variabel er beskrevet på en internetside efter en fast skabelon. I det følgende gennemgås de elementer skabelonen består af.</p>
<i>Den blå bjælke</i>	<p>I den blå bjælke står teksten "høj kvalitetsdokumentation" efterfulgt af variabelnavnet efterfulgt af "variabelbeskrivelse".</p>
<i>Variabel</i>	<p>I denne boks angives variabelens navn, samt beskrivende navn (label).</p>
<i>Emnegrupper</i>	<p>I boksen angives den emnegruppe som variabelen tilhører ifølge Danmarks Statistiks interne emneopdeling, samt den høj kvalitetssemnegruppe som variabelen tilhører. Ved at klikke på høj kvalitetssemnegruppen, fås en oversigt over alle variable i emnegruppen.</p>
<i>Gyldighed</i>	<p>Her angives i hvilken periode variabelen er i anvendelse.</p>
<i>Databrud</i>	<p>For en høj kvalitetsdokumenteret variabel er der angivet, hvorvidt der findes databrud indenfor den pågældende variabel eller databrud på tværs af variable.</p> <p>Databrud indenfor en variabel vil oftest bestå i ændringer af værdisættet (kode-tekst) for variabelen. Nye koder kommer til og tidligere koder forsvinder. Dette kan ske ved</p>

både nedlægning og nyopretning af koder og ved sammenlægninger eller opsplittning af tidligere koder. Desuden kan der forekomme ændringer i den population som variabelen beskriver.

Databrud på tværs af variabler forekommer når variabler udskiftes, sammenlægges eller opsplittes over tid.

Hvis der er databrud for en variabel vil de enkelte databrud være beskrevet i den detaljerede beskrivelse.

<i>Kvalitetssikring</i>	Her angives det om variabelen er kvalitetssikret af Danmarks Statistik, eller om den også har været igennem en ekstern bedømmer udpeget af KOR.
<i>Generel beskrivelse</i>	En beskrivelse af variabelen uden længere tekniske udredninger eller detaljer. Teksten skal give et umiddelbart billede af, hvad variabelen dækker over.
<i>Detaljeret beskrivelse</i>	En mere dybdegående beskrivelse af variabelen. Her skal detaljer om variabelen beskrives. Det kan være sammenhæng til andre variabler, ændringer gennem tiden, specielle oplysninger der er gode at kende i forbindelse med statistisk arbejde med variabelen. Hvis der har været angivet databrud for variabelen, skal dette være beskrevet, evt. suppleret med en forklaring på, hvordan det kan afhjælpes. Der kan også for visse variabler været vedhæftet bilag som belyser databruddet yderligere.
<i>Bilag - Tabel og graf</i>	For en stor del af de højkvalitetsdokumenterede variabler er der vedhæftet en tabel (html-fil) samt en graf (png-fil), der giver en kort beskrivelse af den enkelte variabel gennem tiden, for den population der er beskrevet (se afsnittet "population"). Formålet med tabellen og grafen er at give et indtryk af variabelen og dens udvikling gennem tiden. Der er ikke tale om udarbejdelse af egentlig statistik med umiddelbar samfundsfaglig relevans. Teksterne i tabellen kan være skåret delvist af hvis den reelle tekst er meget lang. Den fulde tekst kan om nødvendigt aflæses i værdisættet, se afsnit om dette. I tabellen er der desuden ofte angivet en række "mellemtotaler til graf". Dette er summeringer af de enkelte kategorier, der ofte er angivet længere nede i tabellen. Det fremgår hvilke kategorier der er summeret til mellemtotaler, enten i teksten for mellemtotalen, eller med angivelse af de kodeværdier der indgår.
<i>Andre bilag</i>	Der kan desuden være vedhæftet andre bilag i de tilfælde, hvor det er fundet nyttigt til at forklare særlige egenskaber ved variabelen.
<i>Population</i>	Angiver den eller de populationer som variabelen beskriver. For hver population er angivet et navn samt en kort beskrivelse.
<i>Værdisæt</i>	For de variabler, der er opbygget af koder angiver værdisættet sammenhængen mellem de enkelte koder og en forklarende tekst. For koder, hvor den forklarende tekst er udskiftet i løbet af variabelens levetid vil der være angivet gyldighedsperioder. Disse angiver i hvilken periode den enkelte tekst er gældende.