

KOMMISSIONENS GENNEMFØRELSESFORORDNING (EU) 2017/881**af 23. maj 2017****om gennemførelse af Europa-Parlamentets og Rådets forordning (EF) nr. 763/2008 om folke- og boligtællinger for så vidt angår kvalitetsrapporternes metoder og struktur og det tekniske format for indberetning af data og om ændring af forordning (EU) nr. 1151/2010****(EØS-relevant tekst)**

EUROPA-KOMMISSIONEN HAR —

under henvisning til traktaten om Den Europæiske Unions funktionsmåde,

under henvisning til Europa-Parlamentets og Rådets forordning (EF) nr. 763/2008 af 9. juli 2008 om folke- og boligtællinger ⁽¹⁾, særlig artikel 5, stk. 5, og artikel 6, stk. 3, og

ud fra følgende betragtninger:

- (1) I forordning (EF) nr. 763/2008 fastsættes der fælles regler for levering af fuldstændige oplysninger om befolknings- og boligforhold hvert tiende år.
- (2) I henhold til Kommissionens forordning (EU) 2017/712 ⁽²⁾ bør den næste folke- og boligtælling vedrøre referenceåret 2021.
- (3) I Kommissionens forordning (EU) nr. 1151/2010 ⁽³⁾ fastsættes kvalitetsrapporternes metoder og struktur og det tekniske format for indberetning af data om folke- og boligtællinger for referenceåret 2011.
- (4) Med henblik på den næste folke- og boligtælling i 2021 og for at vurdere kvaliteten af de data, der indberettes til Kommissionen (Eurostat) er det nødvendigt at fastsætte nye metoder og en ny struktur for kvalitetsrapporterne og det tekniske format for indberetning af data.
- (5) I henhold til artikel 5, stk. 5, i forordning (EF) nr. 763/2008 skal medlemsstaterne fremsende deres validerede data og metadata elektronisk i et egnet teknisk format, som skal vedtages af Kommissionen. Initiativet Statistical Data and Metadata Exchange (SDMX) vedrørende statistiske og tekniske standarder for udveksling og deling af data og metadata, som Census Hub er baseret på, blev iværksat af Den Internationale Betalingsbank, Den Europæiske Centralbank, Kommissionen (Eurostat), Den Internationale Valutafond, Organisationen for Økonomisk Samarbejde og Udvikling (OECD), FN og Verdensbanken. SDMX og Census Hub fastsætter statistiske og tekniske standarder samt indberetningsstandarder for udveksling af officielle statistikker. Der bør derfor indføres et teknisk format i overensstemmelse med disse standarder.
- (6) I henhold til forordning (EU) nr. 1151/2010 skal medlemsstaterne opbevare dataene vedrørende folke- og boligtællingen for referenceåret 2011 indtil den 1. januar 2025. For at brugerne skal kunne sammenligne de to tællinger, bør data fra 2011-tællingen være tilgængelige indtil den 1. januar 2035 sideløbende med 2021-dataene.
- (7) Forordning (EU) nr. 1151/2010 bør derfor ændres.
- (8) Foranstaltningerne i denne forordning er i overensstemmelse med udtalelse fra Udvalget for det Europæiske Statistiske System —

⁽¹⁾ EUT L 218 af 13.8.2008, s. 14.⁽²⁾ Kommissionens forordning (EU) 2017/712 af 20. april 2017 om fastsættelse af referenceåret og programmet for de statistiske data og metadata vedrørende folke- og boligtællinger, der er foreskrevet i Europa-Parlamentets og Rådets forordning (EF) nr. 763/2008 (EUT L 105 af 21.4.2017, s. 1).⁽³⁾ Kommissionens forordning (EU) nr. 1151/2010 af 8. december 2010 om gennemførelse af Europa-Parlamentets og Rådets forordning (EF) nr. 763/2008 om folke- og boligtællinger for så vidt angår kvalitetsrapporternes metoder og struktur og det tekniske format for indberetning af data (EUT L 324 af 9.12.2010, s. 1).

VEDTAGET DENNE FORORDNING:

Artikel 1

Genstand

I denne forordning fastsættes metoderne og strukturen for medlemsstaternes rapporter om kvaliteten af de befolknings- og boligdata, som de indberetter til Kommissionen (Eurostat) for referenceåret 2021, og det tekniske format for dataindberetningen.

Artikel 2

Definitioner

Definitionerne og de tekniske specifikationer i forordning (EF) nr. 763/2008, Kommissionens gennemførelsesforordning (EU) 2017/543 ⁽¹⁾ og forordning (EU) 2017/712 anvendes i nærværende forordning.

Herudover forstås ved:

1. »statistisk enhed«: basisobservationsenhed, dvs. en fysisk person, en husstand, en familie eller en (konventionel) bolig
2. »individuel optælling«: at oplysninger om de enkelte statistiske enheder indsamles på en sådan måde, at deres karakteristika kan registreres særskilt og krydsklassificeres med andre karakteristika
3. »samtidighed«: at de oplysninger, der indhentes i en tælling, referer til samme tidspunkt (referencedato)
4. »universalitet inden for et givent område«: at der leveres data for alle statistiske enheder inden for et præcist defineret område. Hvis de statistiske enheder er personer, forstås ved »universalitet inden for et givent område«, at de data, der leveres, er baseret på oplysninger for alle personer, der har deres sædvanlige bopæl i det givne område (samlet befolkning)
5. »tilgængelighed af data for mindre områder«: hvorvidt der foreligger data for små geografiske områder og for små grupper af statistiske enheder
6. »bestemt hyppighed«: regelmæssig gennemførelse af folke- og boligtællinger i begyndelsen af hvert årti, herunder vedligeholdelse af registre
7. »datakilde«: de dataposter for statistiske enheder og/eller begivenheder i tilknytning til statistiske enheder, som udgør det direkte grundlag for udarbejdelse af folke- og boligtællingsdata for en eller flere specifikke variabler for en bestemt målpopulation
8. »målpopulation«: alle statistiske enheder inden for et bestemt geografisk område, som på referencedatoen opfylder betingelserne for indberetning af en eller flere specifikke variabler. I målpopulationen indgår hver statistisk enhed én gang
9. »estimeret målpopulation«: den bedst tilgængelige approksimation af målpopulationen. Den består af folke- og boligtællingspopulationen plus underdækning minus overdækning
10. »folke- og boligtællingspopulation«: de statistiske enheder, som faktisk er repræsenteret i folke- og boligtællingsresultaterne for en eller flere af de variabler, der er fastsat for en bestemt målpopulation. Folke- og boligtællingspopulationens dataposter er de dataposter, der er omfattet af datakilden for den fastsatte målpopulation, inklusive alle imputerede dataposter og eksklusive alle slettede dataposter. Hvis en datakilde på grund af et metodologisk princip udelukkende omfatter dataposter for en stikprøve af de statistiske enheder i den estimerede målpopulation, omfatter folke- og boligtællingspopulationen ud over de statistiske enheder i stikprøven, det supplerende sæt statistiske enheder
11. »supplerende sæt statistiske enheder«: statistiske enheder, der tilhører en estimeret målpopulation, men som der ikke findes nogen dataposter om i datakilden som følge af den anvendte stikprøvemetode

⁽¹⁾ Kommissionens gennemførelsesforordning (EU) 2017/543 af 22. marts 2017 om fastlæggelse af regler for anvendelsen af Europa-Parlamentets og Rådets forordning (EF) nr. 763/2008 om folke- og boligtællinger for så vidt angår variabelernes tekniske specifikationer samt deres opdelinger (EUT L 78 af 23.3.2017, s. 13).

12. »dækningsvurdering«: en undersøgelse af forskellen mellem en bestemt målpopulation og den tilsvarende folke- og boligtællingspopulation
13. »undersøgelse efter optællingen«: en undersøgelse, der gennemføres kort tid efter optællingen med henblik på vurdering af dækning og indhold
14. »underdækning«: alle de statistiske enheder, der tilhører en bestemt målpopulation, men som ikke indgår i den tilsvarende folke- og boligtællingspopulation
15. »overdækning«: alle de statistiske enheder, der indgår i en folke- og boligtællingspopulation, og som anvendes til indberetning om en bestemt målpopulation, men uden at tilhøre denne målpopulation
16. »imputering af datapost«: henførelse af en kunstig men plausibel datapost til nøjagtigt ét geografisk område på det mest detaljerede geografiske niveau, som folke- og boligtællingsdata udarbejdes på, og imputering af denne datapost i en datakilde
17. »slettelse af en datapost«: den handling, hvorved en datapost, der indgår i en datakilde for en bestemt målpopulation, men som ikke indeholder nogen valide oplysninger om de statistiske enheder i målpopulationen, slettes eller ignoreres/ikke tages i betragtning
18. »partiell imputering«: indsættelse af en kunstig men plausibel værdi om en bestemt variabel i en datapost, som allerede findes i en datakilde, men som enten ikke indeholder denne værdi eller indeholder en værdi, som anses for ikkeplausibel
19. »spørgeskemabaserede data«: data, der oprindeligt er indhentet fra respondenter ved hjælp af et spørgeskema i forbindelse med indsamling af statistiske data, som vedrører et bestemt tidspunkt
20. »sammenkædning af dataposter«: en proces, der anvendes til at slå oplysninger fra forskellige datakilder sammen, ved at sammenligne dataposterne for de enkelte statistiske enheder og slå oplysninger, der referer til den samme statistiske enhed, sammen
21. »entydig identifikator«: en variabel eller et sæt af variabler i dataposterne i en datakilde eller en liste over statistiske enheder, som anvendes til:
 - at kontrollere, at datakilden (eller listen over statistiske enheder) kun indeholder én datapost for hver statistisk enhed, og/eller
 - sammenkædning af dataposter
22. »register«: et lager med oplysninger om statistiske enheder, som opdateres direkte, når der indtræffer begivenheder, der berører de statistiske enheder
23. »registerbaserede data«: data i eller fra et register
24. »matchning af registre«: en sammenkædning af dataposter, hvor alle de matchede datakilder findes i registre
25. »dataudtræk«: en proces, der anvendes til at hente folke- og boligtællingsoplysninger blandt oplysninger, der findes i et register, og som vedrører de enkelte statistiske enheder
26. »kodning«: en proces, der anvendes til at konvertere oplysninger til koder, der repræsenterer klasser i et klassifikationssystem
27. »datafangst«: en proces, hvorved indsamlede data importeres i et format, der egner sig til yderlig behandling
28. »dataredigering«: en proces, der anvendes til at kontrollere og ændre dataposter, for at gøre dem plausible, samtidig med at store dele af disse dataposter bibeholdes
29. »generering af en husstand«: identificering af en privat husstand i overensstemmelse med husstandsbo-pælsbegrebet som defineret i bilaget til gennemførelsesforordning (EU) 2017/543 under variabelen »husstandsstatus«
30. »generering af en familie«: identificering af en familie på grundlag af oplysninger om, hvorvidt personerne bor i samme husstand, men uden eller med ufuldstændige oplysninger om deres indbyrdes familierelationer. Begrebet »familie« defineres som »kernefamilie« i bilaget til gennemførelsesforordning (EU) 2017/543 under variabelen »familiestatus«
31. »enhedsbortfald«: manglende indsamling af data fra en statistisk enhed, som indgår i folke- og boligtællingspopulationen
32. »kontrol med statistisk fortrolighed«: metoder og processer med henblik på at minimere risikoen for, at der videregives oplysninger om individuelle statistiske enheder, når der frigives statistiske oplysninger

33. »estimering«: beregning af statistiske estimater ved hjælp af en matematisk formel og/eller algoritme, der anvendes på de tilgængelige data
34. »definition af datastruktur«: strukturelle metadata, der er knyttet til et datasæt, og som indeholder oplysninger om, hvordan variabler knyttes til en hyperkubus' mål, dimensioner og attributter, samt opdelinger, oplysninger om repræsentation af data og tilhørende deskriptive metadata.

Artikel 3

Metadata og kvalitetsrapportering

Medlemsstaterne forelægger senest den 31. marts 2024 Kommissionen (Eurostat) de baggrundsoplysninger samt de kvalitetsrelaterede data og metadata, der er fastsat i bilaget til denne forordning, vedrørende deres folke- og boligtællinger for referenceåret 2021 og de data og metadata, der indberettes til Kommissionen (Eurostat) i henhold til forordning (EU) 2017/712.

Artikel 4

Datakilder

Medlemsstaterne indberetter alle datakilder, der anvendes til indsamlingen af de oplysninger, der er nødvendige for at opfylde kravene i forordning (EF) nr. 763/2008; datakilderne skal navnlig:

- a) være i overensstemmelse med de væsentlige kendetegn, der er nævnt i artikel 2, litra i), i forordning (EF) nr. 763/2008
- b) være repræsentative for målpopulationen
- c) overholde de relevante tekniske specifikationer, der er fastsat i gennemførelsesforordning (EU) 2017/543 og
- d) bidrage til leveringen af data til programmet for statistiske data, jf. forordning (EU) 2017/712.

Artikel 5

Adgang til relevante oplysninger

1. På anmodning fra Kommissionen (Eurostat) giver medlemsstaterne Kommissionen (Eurostat) adgang til alle oplysninger, der er relevante for vurderingen af kvaliteten af de indberettede data og metadata i henhold til forordning (EU) 2017/712.
2. I forbindelse med overholdelsen af stk. 1 er medlemsstaterne ikke forpligtet til at levere mikrodata eller fortrolige data til Kommissionen (Eurostat).

Artikel 6

Teknisk format for indberetning af data

Det tekniske format, der skal anvendes til indberetning af data og metadata for referenceåret 2021, er formatet Statistical Data and Metadata eXchange (SDMX) som gennemført via Census Hub. Medlemsstaterne indberetter de krævede data i overensstemmelse med Kommissionens (Eurostat) datastrukturdefinitioner og tilhørende tekniske specifikationer. Medlemsstaterne opbevarer de krævede data og metadata indtil den 1. januar 2035, i tilfælde af at Kommissionen (Eurostat) på et senere tidspunkt anmoder om, at de indberettes.

Artikel 7

Ændring af forordning (EU) nr. 1151/2010

Artikel 6, tredje punktum, i forordning (EU) nr. 1151/2010 affattes således:

»Medlemsstaterne opbevarer dataene og metadataene for referenceåret 2011 indtil den 1. januar 2035. Medlemsstaterne er ikke forpligtet til at ændre eller revidere disse data efter den 1. januar 2025. Medlemsstater, der vælger at gøre dette, underretter Kommissionen (Eurostat) om disse ændringer og revisioner, inden de gennemføres.«

*Artikel 8***Ikrafttræden**

Denne forordning træder i kraft på tyvendedagen efter offentliggørelsen i *Den Europæiske Unions Tidende*.

Denne forordning er bindende i alle enkeltheder og gælder umiddelbart i hver medlemsstat.

Udfærdiget i Bruxelles, den 23. maj 2017.

På Kommissionens vegne

Jean-Claude JUNKER

Formand

BILAG

Indhold og struktur af kvalitetsrapporterne for indberetning af data

De **tekstmæssige** og **kvantitative** metadataoplysninger vedrørende de folke- og boligtællinger, der gennemføres i medlemsstaterne for referenceåret 2021, omfatter følgende afsnit:

1. OVERSIGT

- 1.1. Retsgrundlag
- 1.2. Ansvarlige organer

2. DATAKILDER

- 2.1. Klassifikation af datakilder i overensstemmelse med artikel 4, stk. 1, i forordning (EF) nr. 763/2008
- 2.2. Liste over datakilder, der anvendes til 2021-tællingen
- 2.3. Matrix »datakilder x variable«
- 2.4. Datakilders egnethed: I hvilket omfang datakilderne er i overensstemmelse med de væsentlige kendetegn i artikel 4, stk. 4, i forordning (EF) nr. 763/2008
 - 2.4.1. *Individuel optælling*
 - 2.4.2. *Samtidighed*
 - 2.4.3. *Universalitet inden for et givent område*
 - 2.4.4. *Tilgængelighed af data for mindre områder*
 - 2.4.5. *Bestemt hyppighed.*

3. TÆLLINGENS LIVSCYKLUS**3.1. Referencedato**

Medlemsstaterne skal oplyse Kommissionen (Eurostat) om referencedatoen i henhold til artikel 3 i Kommissionens forordning (EU) 2017/712

3.2. Forberedelse og gennemførelse af dataindsamlingen**3.2.1. Spørgeskemabaserede data**

- udformning og afprøvning af spørgeskemaer (med eksemplarer af alle endelige spørgeskemaer)
- udarbejdelse af eventuelle adresselister, forberedelse af feltarbejdet, kortlægning, offentlig omtale
- dataindsamling (herunder feltarbejde)
- retlig forpligtelse til at indsamle oplysninger, incitamenter til at levere sandfærdige oplysninger eller mulige grunde til at levere urigtige oplysninger.

3.2.2. Registerbaserede data

- oprettelse af nye registre fra og med 2011 (hvis det er relevant)
- ændring af eksisterende registre fra og med 2011 (herunder ændringer i registerindhold, tilpasning af folke- og boligtællingspopulationen, tilpasning af definitioner og/eller tekniske specifikationer) (hvis det er relevant)

- registrenes vedligeholdelse (for hvert register, der anvendes til 2021-tællingen), herunder
 - registerindhold (statistiske enheder og oplysninger herom, dataredigering og/eller partiel imputering i registret)
 - administrativt ansvar
 - retlig forpligtelse til at registrere oplysninger, incitamenter til at levere sandfærdige oplysninger og mulige grunde til at levere urigtige oplysninger
 - forsinkelser i indberetningen, navnlig legale/officielle forsinkelser, forsinkelser i dataregistreringen, for sen indberetning
 - evaluering af og justering for ikke-registrering, ikke-afregistrering, dobbeltregistrering
 - større registerrevisioner eller ajourføringer af dataposter, der berører 2021-tællingsdataene, hyppighed af registerrevisioner
 - anvendelse, herunder »statistisk anvendelse af registret til andre formål end tællingsformål« og »anvendelse af registret til andre formål end statistiske formål (f.eks. administrative formål)«
- matchning og sammenkædning af registre (herunder entydig(e) identifikator(er), som anvendes til sammenkædning af data)
- dataudtræk.

3.2.3. Data indsamlet ved hjælp af en stikprøve

For variabler, for hvilke der er indsamlet oplysninger ved hjælp af en stikprøve, beskriver metadataene tillige:

- stikprøveplanen
- metodologier anvendt til estimeringer, modeller eller imputeringer
- mulige skævheder i estimeringen som følge af de anvendte metodologier
- formler og algoritmer, der anvendes til at beregne standardafvigelsen.

3.2.4. Data indsamlet ved hjælp af kombinerede metoder (data baseret på mere end én type datakilde)

For variabler, for hvilke der er indsamlet oplysninger ved hjælp af kombinerede metoder, skal metadataene tillige indeholde:

- en beskrivelse af de metoder (typer anvendte datakilder, og hvordan oplysninger fra forskellige datakilder blev kombineret, hvordan de forskellige anvendte kilder og metoder supplerer og støtter hinanden, og, hvis det er relevant, hvilke dele af populationen der var dækket af hvilke kilder)
- andre kvalitetsspørgsmål i forbindelse med anvendelsen af kombinerede metoder.

3.3. Bearbejdning og evaluering

3.3.1. Databehandling (herunder datafangst, kodning, identifikationsvariabel/-variabler, dataredigering, imputering af dataposter, slettelse af dataposter, estimering, sammenkædning af dataposter, herunder identifikationsvariabler til brug for sammenkædningen, generering af husstande og familier, foranstaltninger til identificering eller begrænsning af enhedsbortfald)

3.3.2. Aktiviteter i forbindelse med vurdering af dækning, metodologi til behandling af bortfald, undersøgelse(r) efter optælling (hvis det er relevant), endelig validering af data: metode til vurdering af under- og overdækning, herunder oplysninger om kvaliteten af estimerterne for under- og overdækning.

3.4. **Formidling** (formidlingskanaler, der sikrer statistisk fortrolighed, herunder kontrol med den statistiske fortrolighed)

3.5. Foranstaltninger til sikring af omkostningseffektivitet

4. VURDERING AF DATAKVALITET

4.1. Sammenlignelighed

For hver variabel skal medlemsstaterne indberette alle afvigelser fra de krævede begreber eller definitioner eller alle fremgangsmåder i medlemsstaten, som kan forringe sammenligneligheden af data på EU-plan.

For så vidt angår variabelen »nuværende erhvervssituation« skal medlemsstaterne indberette alle de estimeringsmetoder, der anvendes til at justere data for i højere grad at opfylde definitionen i bilaget til Kommissionens gennemførelsesforordning (EU) 2017/543. Medlemsstaterne skal indberette, i hvilket omfang datakilderne og de anvendte estimeringsmetoder resulterer i afvigelser fra definitionen af »nuværende erhvervssituation«, som er fastsat i nævnte forordning.

4.2 Aktualitet og punktlighed

Følgende oplysninger skal leveres for det nationale niveau:

- dato(er) for indberetning af data til Kommissionen (Eurostat) opdelt efter hyperkubusser
- dato(er) for større revision(er) af de indberettede data opdelt efter hyperkubusser
- dato(er) for indberetning af metadataene.

I tilfælde af større revisioner fra og med den 1. april 2024 skal medlemsstaterne indberette de(n) respektive dato(er) separat til Kommissionen (Eurostat) inden for senest en uge.

4.3. Kohærens

Medlemsstaterne skal indberette væsentlige uoverensstemmelser mellem data, der indberettes i de forskellige datasæt, der er defineret i Kommissionens forordning (EU) 2017/712.

4.4. Dækning og nøjagtighed

For at angive dækning skal følgende **absolutte værdier** leveres for persontællinger på nationalt plan og opdeles efter køn og brede aldersgrupper som defineret i Kommissionens gennemførelsesforordning (EU) 2017/543:

- a) folke- og boligtællingspopulation
- b) samlet antal imputerede dataposter
- c) samlet antal slettede dataposter
- d) underdækning (estimeret)
- e) overdækning (estimeret)
- f) estimeret målpopulation.

For at vurdere nøjagtighed skal følgende **absolutte værdier** leveres for persontællinger på nationalt plan og opdeles efter køn og brede aldersgrupper som defineret i Kommissionens gennemførelsesforordning (EU) 2017/543:

- a) folke- og boligtællingspopulation
- b) antal observerede dataposter vedrørende variabelen afledt af traditionelle tællinger
- c) antal observerede dataposter vedrørende variabelen afledt af administrative registre
- d) antal observerede dataposter vedrørende variabelen afledt af stikprøveundersøgelser
- e) antal observerede dataposter vedrørende variabelen afledt af flere forskellige datakilder
- f) supplerende sæt statistiske enheder vedrørende variabelen (for stikprøver)
- g) antal imputerede observationer vedrørende variabelen
- h) antal dataposter med manglende oplysninger vedrørende variabelen.

Ovenstående absolutte værdier til vurdering af nøjagtighed skal leveres for følgende tællingsvariabler:

- a) civilstand (LMS)
- b) familiestatus (FST)
- c) husstandsstatus (HST)
- d) nuværende erhvervssituation (CAS)
- e) beskæftigelse (OCC)
- f) erhvervsgren (IND)
- g) jobniveau (SIE)
- h) arbejdspladsens beliggenhed (LPW)
- i) uddannelsesniveau (EDU)
- j) fødeland/fødested (POB)
- k) statsborgerskabsland (COC)
- l) år for ankomst til landet 2010 eller senere (YAT)
- m) år for ankomst til landet 1980 eller senere (YAE)
- n) sædvanlig bopæl et år før tællingen (ROY)
- o) boligforhold (HAR).

4.5. Fuldstændighed

Medlemsstaterne skal indberette dataenes grad af fuldstændighed med hensyn til kravene i forordning (EF) nr. 763/2008. De skal give nærmere oplysninger om tællingsvariabler eller tilknyttede opdelinger, for hvilke der ikke leveres data.

4.6. Relevans

Oplysninger om følgende skal leveres på EU-plan:

- a) foranstaltninger, der er truffet for at identificere og opfylde brugerbehov
 - b) overvågning af omfanget af dataudtræk.
-