

# Interne notater

STATISTISK SENTRALBYRÅ

89/6

7. februar 1989

OPPBYGGINGEN AV FØDELANDSFILEN

av

Kåre Vassenden

## Innhold

	Side
1. Innledning .....	3
2. Generelt om fødelandsfilen .....	3
2.1. Målet med arbeidet .....	3
2.1.1. Massen .....	3
2.1.2. Kjennemerker .....	4
2.2. Datakilder .....	5
2.3. Retting av fødselsnummer .....	6
2.4. Koder for land og statsborgerskap .....	7
3. Arbeidets forløp .....	8
4. Dokumentasjon av fødelandsfilen 1988 .....	10
4.1. Ulike måter å dokumentere på .....	10
4.2. Systematisk dokumentasjon .....	10
4.2.1. Oversikt .....	10
4.2.2. Programnavn .....	13
4.2.3. A-delen: Innflyttere - statsborgerskap og fraflyttingsland .....	13
4.2.4. B-delen: Fob70 - fødeland .....	14
4.2.5. C-delen: Innflyttere - fødeland .....	15
4.2.6. D-delen: Fødte - fødeland Norge .....	16
4.2.7. E-delen: Sammensyningen til fødelandsfil .....	18
4.2.8. F-delen: Fødelandssituasjonen .....	18
4.2.9. G-delen: Kortversjoner .....	18
4.2.10 H-delen: Situasjonsfiler .....	19
4.2.11. J-delen: Statsborgerskapsendringer .....	19
4.2.12. K-delen: Tilleggsposter .....	20
4.2.13. L-delen: Linkfilen .....	20
4.2.14. M-delen: Mors, fars og ektefelles fødeland ..	21
4.3. Kjennemerkedokumentasjon .....	21
4.3.1. Kjennemerkene på fødelandsfilen .....	22
4.3.2. Eksempler på tilleggskjennemerker på føde- landssituasjonen .....	27
5. Kvalitetsvurdering av fødelandsfilen .....	28
5.1. Innledning .....	28
5.2. Anvendte kilder for fødelandskodene .....	28
5.3. Kodetyper .....	29
5.4. Sammenligninger .....	29
5.4.1. Sammenligningsmulighetene .....	29
5.4.2. Sammenligningsfeltene .....	30
5.4.3. Sammenligningsoversikt .....	30
5.5. Kvalitetsdiskusjon .....	32
5.5.1. Kvaliteten på fødeland fra fødefilene (D) ..	32
5.5.2. Kvaliteten på fødeland fra innvadrings- meldingene (C) .....	32
5.5.3. Kvaliteten på fødeland fra Fob 70 (B) .....	33
5.5.4. Kvaliteten på statsborgerskapsfødeland (A) ..	34
5.5.5. Kvaliteten på fraflyttingslandfødeland (A) ..	34
5.5.6. Utrekning og oppsummering av feil .....	35
6. Oppdateringer i framtida .....	36
Vedlegg:	
1. Kodetøkler/lister for land/statsborgerskap .....	37
2. Fordelingen på 'kodekvalitet' .....	43
3. Aktuelle filidenter .....	47

## 1. INNLEDNING

Dette notatet skal forklare og delvis dokumentere dannelsen av kjennemerket 'fødeland' som ble foretatt på Kontor for befolkningsstatistikk høsten 1987 og videreført i 1988. Det meste av dokumentasjonen befinner seg ellers i programmene. Notatet vil gi et innblikk i hvilke vurderinger som ble gjort i arbeidet.

Situasjonen før arbeidet tok til var slik: 'Fødeland' var et kjenne-merke i Folke- og bolig telling 1970, men for tida etter 1970 fantes det ingen opplysninger om befolkningens fødeland, bare dens statsborgerskap. Stadig flere personer med utenlandsk bakgrunn hadde norsk statsborgerskap, og behovet for en utvidet utlendingsstatistikk ble mer påtrengende. Kommunal- og arbeidsdepartementet stilte midler til disposisjon for å få laget en fil med opplysninger om fødeland. Detaljert prosjektskriv var utarbeidet (Ski 20/7-84), og det la til en viss grad opp rammene for prosjektet.

Tittelen på dette notatet er "Oppbyggingen av fødelandsfilen". "Oppbygging" er her brukt i begge betydninger av ordet, altså både "byggeprosess" og "konstruksjon". Den beste måten å forklare resultatet på er å forklare prosessen som ledet fram til resultatet. Dessuten er det sannsynlig at den beste måte å oppdatere fødelandsfilen på i framtida vil være å "bygge den opp" på nytt hvert år.

Navnet "fødelandsfilen" vil etterhvert bli gitt innhold. Selve fødelandsfilen er forresten ikke sluttproduktet i prosessen, bare nest siste trinn. Det er først når fødelandsfilen kobles til situasjonsfiler eller endringsfiler at opplysningene blir interessante. Det sluttproduktet som oppstår ved kobling til siste situasjonsfil blir her kalt "fødelandssituasjonen".

Det er viktig allerede nå å gjøre oppmerksom på at fødelandsprosjektet bygger på at 'første registrerte statsborgerskap' indikerer fødeland.

## 2. GENERELT OM FØDELANDSFILEN

### 2.1. Målet med arbeidet

#### 2.1.1. Massen

Det ble tidlig klart at resultatet av arbeidet måtte bli en fil som inneholdt alle som var bosatt og/eller hadde vært bosatt i Norge siden et visst tidspunkt. Folketellingsdatoen 1.11.70 pekte seg naturlig ut som et slikt starttidspunkt. Særlig gode data før 1970 kunne en dessuten ikke vente å finne. (At det er tatt med noen innflyttere og fødte fra de nærmeste årene før 1970 forandrer ikke på det at 1.11.70 er første dato vi i prinsippet har hele massen registrert).

Denne fødelandsfilen måtte så kunne kobles til situasjonsfiler for å få fram situasjonen pr. en bestemt dato. En slik kobling betegner vi "fødelandssituasjonen".

### 2.1.2. Kjennemerker

Bortsett fra kjennemerkene fødselsnummer og fødeland så måtte det på en fødelandsfil også finnes en del tilleggsopplysninger. For det første måtte registreringsdatoer for de ulike opplysninger tas med. Slike datoer representerer ingen tekniske problemer som er verdt å nevne.

Dernest var det ønskelig med kjennemerker som kunne si noe om kvaliteten på produktet, særlig med tanke på den tilnærmingen til fødeland som måtte foretas ved hjelp av statsborgerskap og fraflyttingsland. Disse kvalitetskjennemerkene var først relativt enkle, men ble i prosessens løp mer spesifisert.

To litt spesielle kjennemerker som kom med på fødelandsfilen er 'fødekommune' og 'antall opphold i utlandet'.

Til sist er de opprinnelige kildene for fødelandskoden tatt med som egne kjennemerker, bl.a. for å lette eventuelle kontroller.

Kjennemerkene nedenfor kom med på den endelige fødelandsfilen. Nærmere forklaring på kjennemerkene vil komme inn ulike steder i teksten (særlig i kap. 4.3).

FØDSELSNUMMER	
FØDELAND	
FØDELANDSDATO.	Registreringsdato for den i fødelandsopplysningen som er brukt 'fødeland'.
KODEKVALITET.	Består av kjennemerkene 'KILDE', 'KODETYPE' og 'SAMMENLIGNINGSFELTER'
KOBLINGSTYPE.	Viser kombinasjoner av kilder.
FØRSTE REGISTRERTE	
STATSBORGERSKAP	
STATSBORGERSKAPSDATO.	
FØRSTE OPPHOLDSDATO.	Dato for første opphold i Norge (= tidligste registrerte dato i våre kilder) (år-mnd-dag).
ANTALL OPPHOLD I	
UTLANDET.	I følge våre kilder.
FØDEKOMMUNE	
KOMMUNEÅRGANG.	Brukes sammen med FØDEKOMMUNE ved kommunegrenseendringer.
FØDELAND I INN-	
VANDRINGSMELDINGENE	
FØDELAND I FOB70	
FRAFLYTTINGSLAND	

Tidsbestemte og forandelige kjennemerker som bostedskommune, ekte-skapelig status, statsborgerskap m.v. må naturlig nok bli tatt fra de filer som fødelandsfilen kobles mot. Alder må regnes ut i hvert i enkelt tilfelle. Kjønn utledes samtidig på grunnlag av fødselsnummeret.

## 2.2. Datakilder

Flere datakilder ble vurdert, men det pekte seg tidlig ut tre hovedkilder:

- Filen fra Folke- og bolig telling 1970.
- Fødtefilene med alle fødte i Norge siden 1970.
- Innvandringsmeldingene med personer innvandret siden 1970.

Fob70-filen ville bidra med den opplysningen om 'fødeland' som folk hadde oppgitt i folketellingen 1970.

Fødtefilene ville peke ut en hel del som sikre norsk-fødte.

Innvandringsmeldingene ville gi informasjon om innvandrere. Den informasjonen omfattet dessverre ikke fødeland, men statsborgerskap og fraflyttingsland skulle tjene som substitutt. Ved oppdateringen i 1988 kunne også fødeland (registrert f.o.m. 1.7.87) tas fra innvandringsmeldingene og brukes som kilde.

Med ordet innvandringsmeldinger tenkes ikke her på selve journalmeldingene, men på statistikkfiler som inntil 1984 ble laget med innvandringsmeldinger (og utvandringsmeldinger) som kilde. Det er bare statistikkfilene som er lett tilgjengelig. I teorien skulle det ikke være noen forskjell mellom meldingene og statistikkfilene, men i praksis har ikke de seneste etterslepsmeldingene (gjelder både fødsler og innflyttinger) kommet med på statistikkfilene. Dette ble dessverre ikke klarlagt før langt ut på høsten 1987.

I 1985 ble befolkningsstatistikksystemet lagt om, og problemet med etterslepsmeldinger er i vår sammenheng opphørt fra da av. Kilden for 1985 og senere er de såkalte meldingsfilene. Meldingsfiler er mellomledet mellom journalfilene (som jevnlig mottas fra DSP) og statistikkfilene (årgangsfilerne).

En sak for seg er at ordet "innvandrings"-melding egentlig kan være villedende, fordi det alltid er snakk om innflyttinger til Norge av både "innvandrere" (som blir boende her), "besøkende" (kortidsinnflyttinger) og tilbakeflyttere til Norge. Vi velger å beholde det hevdvundne ordet "innvandringsmeldinger" som navn på de datafiler vi har nevnt, og så bruke ordet innflytter mer generelt.

Det ble vurdert å bruke også Fob60-filen som datakilde. Hensikten var for det første å skille ut de som hadde identiske fødeland i Fob60 og -70, fra de som ikke hadde det, for på den måten å få et begrep om kvaliteten. For det andre kunne noen innvandrere på 70-80-tallet som ikke var med i Fob70 bli gjenfunnet i Fob60, og dermed få en reell fødelandskode.

Fob60 ble droppet som datakilde da det ble klarlagt at en slik kobling krevde melding til datatilsynet, fordi opplysningene var samlet inn med en folketellingshjemmel. I Fob70 hadde imidlertid noen opplysninger blitt samlet inn med hjemmel i Folkeregistreringsloven. Dette gjelder alle de opplysninger som Folkeregistreringsloven hjemler, deriblant "fødested". Følgelig kunne 'fødeland' tas rett fra Fob70-filen uten nærmere formaliteter.

En annen grunn til å unngå Fob60 var at utbyttet ikke ville stå i forhold til innsatsen. Bare som hjelpemiddel i vurderingen av kvaliteten på fødeland i Fob70 ville Fob60-fødeland komme til sin rett. Denne kvalitetsvurderingen har skjedd utenfor fødelandsfilen (se kap. 5.4.4).

### 2.3. Retting av fødselsnummer

Fra Statens Datasentral ble anskaffet en såkalt linkfil, oppdatert t.o.m. september 1987. Den omfattet alle fødselsnummer som var endret siden fødselsnummer ble innført i 1964. Filen inneholder gammelt fødselsnummer, nytt fødselsnummer, og en kode som forteller om årsaken til endringen. Ved å koble en linkfil til en datafil kan en få byttet ut alle utgåtte fødselsnummer med de som gjelder på oppdateringsdatoen. En slik kobling forutsetter at datafilen er sortert på fødselsnummer. Med ett unntak (situasjonsfilen pr. 1.1.87) ble alle filer utsatt for fødselsnummerretting. I flere tilfeller medførte det at fødselsnummer som før var ulike ble like, altså at det oppsto (flere) fødselsnummerdubletter på filene. Disse ble etterhvert luket ut.

Før datakjøringene i '88 måtte linkfilen oppdateres til 8.04.88 (se L-delen).

Det er ikke alle som tenker på å rette gamle fødselsnummer når de kobler filer fra ulike tidspunkter. Holdningen er at nytten ikke står i forhold til innsatsen. Når slik retting ble gjennomført konsekvent i dette prosjektet så er begrunnelsen følgende:

1. Tallet på fødselsnummerendringer siden 1964 er faktisk ikke så lite (ca. 190 000 endringer).
2. Det er spesielt mange fødselsnummerendringer blant utlendinger - nettopp den kategorien som fødelandsprosjektet i første rekke omhandler.
3. Fordi 'fødeland' skulle lages på en måte som ville komme til å gi en del feil, var det viktig å eliminere flest mulig av de feilkildene en hadde rådighet over.
4. Ikke rettede fødselsnummer på fødelandsfilen ville bety at for en hel del personer ville opplysningene på fødelandsfilen ikke kunne nyttiggjøres. Ved koblinger med andre filer (situasjons- eller endringsfiler) er en avhengig av at fødselsnumrene på tilkoblingsfilen finner sin makker på fødelandsfilen. Det skjer ikke for alle fødselsnumrene hvis de ikke er oppdatert til samme dato på de to filene.

Noen eksempler på betydningen av retting:

Pga. en feilopplysning hadde ikke K70 fått rettet fødselsnumrene til 80-standard. Den rettingen som ble foretatt på K70 i 1987 berørte derfor fødselsnummer fra hele perioden 1.11.70 til 30.9.87. Det viste seg at hele 13 022 fødselsnummer ble endret. Ikke overraskende kom det fram en del dubletter etter rettingen (151 stk.).

I løpet av de neste 6 månedene (fram til 8.4.88) ble 97 av fødselsnumrene på Fob70-filen endret.

Da en Fob80-fil ble rettet til 8.4.88-nivå ga det 3 045 endrede fødselsnummer og 54 dubletter.

På en fødtefil som bestod av årgangene 67-84 ble 2 789 nummer endret. Etter rettingen var det 658 dubletter på filen. Noen færre er altså født her i landet enn det summen av fødselstallene for hvert år skulle tilsi. Men dette oppveies av at en del fødte ikke er med på statistikkfilene p.g.a. etterslep.

En test med linkfilen viser at 119 843 fødselsnummer på situasjonsfilen 1.1.88 er endrede nummer. De aller fleste av dem gjelder personer som var bosatt i Norge allerede 1.10.64. Av de resterende 15 192 personene ligger andelen med utenlandsfødte relativt høyt (32 %). Blant personer født eller kommet til Norge etter 1970 og som har fått endret sitt fødselsnummer, er andelen utlendinger ca. 60 %.

#### 2.4. Koder for land og statsborgerskap

Før programmeringen kunne begynne måtte det gjøres et arbeid med å bringe klarhet i de ulike kodene for land og statsborgerskap som befant seg på filene, og sammenhengen mellom dem ("kodenøkler"). Resultatet av dette arbeidet er gjengitt i vedlegg 1.

Den nå gjeldende kodelista (datert 1981, gjaldt i praksis fra 1975) ble tatt som utgangspunkt. Prinsippet i den er at hvert land har sin egen kode. Kodelistene for 1974 og tidligere (det gjelder 4 delvis ulike lister) er klart mindre detaljert enn 81-lista. Disse tidligere listene var basert på at bare de viktigste landene fikk egen kode, og så var de resterende landene samlet i sekkekoder som f.eks. "Resten av Europa". Tendensen til å slå sammen land (som naturlig hørte sammen) var også større før. Dette gjelder også for folketellingene.

Koder fra de eldre listene fikk en kode etter 81-standard. I flere tilfeller ble det nødvendig å lage nye koder for å få innpasset de gamle kodene. Tre typer koder ble opprinnelig laget: samlekoder, restkoder og uoppgittkoder. Samlekoder viser til land som nå er delt, mens restkoder viser til land som ikke hadde egen kode, men ble samlet i restkategorier. Uoppgittkoder eksisterte fra før, men det ble laget en til. Selv med innføringen av slike ekstrakoder ble

ikke kodenøkene fri for svakheter. Den viktigste av disse er at 101 ble stående som kode både for 'Danmark uten Færøyene og Grønland', 'Danmark uten Færøyene', 'Danmark med Færøyene og Grønland', og Danmark uspesifisert'. Se kodelista i vedlegg 1 for flere eksempler.

Situasjonen er altså at kildene inneholder mange "mindreverdige" koder, altså koder som ikke er så spesifisert som de nå gjeldende kodene er. I alle ledd i prosessen fram mot fødelandsfilen blir de "mindreverdige" kodene forsøkt erstattet med mer spesifiserte koder.

Eksemplet med Danmark demonstrerer hvordan arbeidet med fødelandsfilen var preget av skjønnsmessige avgjørelser. Resultatet har da også blitt deretter - nemlig fødelandsopplysninger som må tas med et visst slingringsmonn og kunnskap om de situasjonen for de enkelte land. Mer om det i kap. 5.

Erfaringene fra den første tabellproduksjonen var at samlekoder ikke bare var unødvendige - de var også til mye bryderi. Det viste seg nemlig at det var svært få som kom fra hhv. DDR, Bangladesh og Nord-Korea, men mange fra Forbundsrepublikken Tyskland, det nåværende Pakistan og Sør-Korea (de tre viktigste tilfellene). Dermed var det åpenbart hvor de fleste personene fra Tyskland, det gamle Pakistan og Korea hørte hjemme. Samlekodene ble da også avskaffet i 88-oppdateringen.

### 3. ARBEIDETS FORLØP

Den aller første delen av arbeidet innebar en utredning om hvilke kilder som forelå, hvor kildene var, hvilke kilder og hvilken framgangsmåte som var mest hensiktsmessig. Kildene (altså filene) måtte utprøves. Dokumentasjonen var tildels svært dårlig, og det dukket opp mange overraskelser når en studerte filene nærmere. Også i løpet av prosessen har det foregått en løpende vurdering av framgangsmåter, kjennemerker og kjennemerkeverdier.

Når framgangsmåten så var valgt bestod hovedproblemet i systemarbeidet for det første i å få statsborgerskap og fraflyttingsland til å bli fødeland, og dernest i å sy sammen opplysningene fra de tre kildene innvandringsmeldinger, Fob70 og fødtefiler.

Dette ble etterhvert gjort (på en måte som ligger nær opp til det som vises for 88-oppdateringen), og til slutt ble den dannede fødelandsfilen koblet med situasjonen pr. 1.1.87. Da viste det seg at 5 877 poster (poster = records) på situasjonsfilen ikke fantes på fødelandsfilen. Våre kilder hadde tydeligvis ikke vært uttømmende nok når det gjaldt å få med alle som hadde vært bosatt i Norge siden 70-tellinga. Forklaringen på manglene er (som forklart i kap. 2.2.) at årsfilene for innflytting og fødte hadde blitt avsluttet i rimelig tid etter årets utgang, og at etterslepen ikke hadde kommet med.



Blant annet p.g.a. de 5877 manglende personer på fødelandsfilen måtte sammensyningen av den foretas på nytt. For om mulig å få ned tallet på manglende personene ble statsborgerskapsendringer tatt i bruk som kilde, forøvrig uten særlig virkning. De som til slutt gjensto fikk sitt fødeland bestemt av 'nåværende statsborgerskap' på situasjonsfilen. Dette var en løsning som kunne brukes der situasjonsfilen hadde et statsborgerskapskjennemerke, men som ved koblinger med andre filer ville skape problemer. Dette problemet fikk en tilfredsstillende løsning først i 88-utgaven.

Erfaringene fra arbeidet med 87-utgaven av fødelandsfilen og tabellproduksjonen på grunnlag av den var følgende:

1. Samlekodene var unødvendige og brysomme.
2. Det ble for mange kodekvalitetsverdier.
3. Det var ikke brukbart at fødelandsfilen ikke inneholdt alle som var med på situasjonsfilen.
4. Statsborgerskapsendringene var tatt i bruk på en måte som var på grensa av det en kunne forsvare.
5. Filidentene var til dels usystematiske og ulogiske.
6. En del av arbeidet, særlig i begynnelsen, var preget av manglende programmeringserfaring og manglende kunnskap om innholdet på filene.
7. Mange faktorer (jfr. punkt 2-6) førte til at det ble et for stort avvik mellom det som burde vært gjort, og det som faktisk var gjort. Dette avviket skapte problemer for dokumentasjonen (den kom aldri over det elementære stadiet).
8. Det var upraktisk at 'fødekommune' og 'fødeland' var atskilte kjennemerker.

Dette skjedde i 88-oppdateringen som svar på erfaringene:

1. Samlekodene ble avskaffet (se kap. 2.5).
2. For mye programmeringsarbeid var lagt ned i kodekvalitetskjennemerket til at det var noe poeng å endre det. Men faktisk ble det vesentlig enklere p.g.a. bortfallet av samlekoder og nåværende statsborgerskap som kilde.
3. De manglende poster ble hentet fra situasjonsfilen og integrert i hhv. innflytterfilen og fødtefilen. Fordelingen mellom innflyttere og fødte er imidlertid usikker. I DSP-basen finnes de nødvendige kjennemerkene til sikker fordeling, men disse stod ikke med på vår situasjonsfil.
4. Statsborgerskapsendringene ble tatt i bruk til å forbedre mindreverdige koder også i 88, men forbedringen skjedde tidligere i prosessen og på en mer kvalifisert måte.
5. Filidentene ble mer systematiske (se kap. 4.1.).
6. Arbeidet i 88 kunne bygge på erfaringer og kunnskaper fra høsten 87, noe som gjorde arbeidet lettere og ryddigere.
7. Arbeidet ble hele tida gjort med vissheten om at det skulle kunne dokumenteres i detalj.
8. 'Fødekommune' og 'fødeland' ble kombinert, slik at hvert utland står fram som "kommune" (med førstesiffer 3).
9. Programmene ble tilpasset den nye kilden fødeland fra innvandringsmeldingene.

#### 4. DOKUMENTASJON AV FØDELANDSFILEN 1988

##### 4.1. Ulike måter å dokumentere på

Som tidligere forklart skulle fødelandsfilen bestå av de som var med i Fob70 pluss fødte og innflyttere etter den tid. Oppbyggingen av fødelandsfilen foregår ved at de delene gjøres klar hver for seg, og så blir de "sydd sammen", eller m.a.o. koblet. Først kobles innflytterdelen med fødtedelen, og så blir koblingsproduktet koblet med Fob70-delen. Resultatet er da fødelandsfilen, som så gjerne kobles til situasjons- og endringsfiler.

Når uttrykket "sy sammen" er brukt, så er det for å fortelle at koblingsprosessene foretas ved programmer det er lagt en viss omtanke i.

Når en skal dokumentere/presentere et slikt arbeid som arbeidet med fødelandsfilen har vært, kan en gjøre dette på flere måter:

1. Faktisk kronologisk (hva som skjedde dag for dag).
2. Ideelt kronologisk (hvordan det burde blitt gjort steg for steg).
3. Følge kjennemerke for kjennemerke gjennom prosessen.
4. Systematisk (etter et eller annetsystem).

Av disse beskrivelsesmåtene velger vi den systematiske, men også den ideelle kronologien vil bli beskrevet. Når det gjelder det å følge kjennemerker i systemet er det antydninger til det i teksten.

For at det skal være sagt: I programmene er det mye dokumentasjon som ikke kommer med i dette notatet.

##### 4.2. Systematisk dokumentasjon

###### 4.2.1. Oversikt

I standard filidenter for SSB består tredje leddet av et datasettnummer, et redigeringsnummer og et utgavenummer.

Datasettnummeret for Fødeland 1987 er I698, for 1988 er det I755 (nytt nummer ble valgt av praktiske grunner). I 88-oppdateringen ble redigeringsnummeret laget med utgangspunkt i bokstavkodene nedenfor, slik at f.eks. en fil som hadde med fødte å gjøre (D-delen) kunne få et 3. ledd som "I755D1A2". Disse bokstavkodene gjør det lettere å holde tråden, og blir brukt utover både i omtalen av filer, program og ledd i prosessen. Vi velger å følge dette systemet strengt alfabetisk, selv om vi skal se at den reelle prosessen var springende.

- A - Innflyttere: statsborgerskap og fraflyttingsland
- B - Fob70: fødeland
- C - Innflyttere: fødeland (f.o.m. 1.7.88)
- D - Fødte: fødeland Norge
- E - Koblinger mellom A, B, C og D. Resultatet blir Fødelandsfilen
- F - Fødelandssituasjonen 1.1.88 (evt. andre tidspunkter)
- G - Kortversjoner av fødelandssituasjoner
- H - Situasjonsfiler (bearbeidet for vårt formål)
- J - Forbedring av visse gamle koder
- K - Tilleggsposter
- L - Linkfilen
- M - Mor, fars og ektefelles fødeland
- N - Koblinger av fødelandsfilen mot fødtefiler
- O osv. - Eventuelle framtidige koblinger med fødelandsfilen

Om G og H er det å si at disse leddene er med bare av praktiske grunner. Bokstaven I brukes ikke fordi den kan forveksles med l.

M er et tillegg til fødelandssituasjonen, og tas derfor med i dokumentasjonen.

N (og utover) er anvendelser av fødelandsfilen og dokumenteres ikke nærmere her.

Rekkefølgen av A-, B-, C- og D-delen er ikke tilfeldig. Prinsippet er at delene er ordnet i stigende kvalitet. I de programmene som definerer fødeland blir opplysningene om statsborgerskap betraktet som mer pålitelige enn fraflyttingsland (fra et fødelandssynspunkt). Tilsvarende går Fob70-fødeland normalt foran statsborgerskap, innflytter-fødeland går foran Fob70-fødeland, og fødtefilene går foran alt. Sikrere informasjon om fødeland enn fødtefilene kan en ikke finne (med noen unntak, avhengig av hvordan "fødeland" defineres).

Noen vil kanskje si at fødeland i Fob70 minst må være like bra som fødeland på innvandringsmeldingene. Imidlertid er det bare ca. 2 000 tilfeller hvor en kan velge mellom disse to kildene, og i så godt som alle disse tilfellene er fødeland det samme begge steder. Rangeringen har dermed liten betydning.

I det følgende vil prosessen i 88-oppdateringen beskrives. Det er viktig å være klar over at denne oppdateringen av to grunner er mer komplisert enn senere oppdateringer må være. Det er ikke tatt stilling til hvordan senere oppdateringer skal foregå, så beskrivelsen her gjelder utelukkende det som til nå har skjedd.

Den første kompliserende grunnen er at det befinner seg 5 628 personer på situasjonsfilen 1.1.88 som ikke finnes i våre kilder. Disse personene er det nå funnet fram til en gang for alle (ledd K).

Den andre grunnen er det ble gjort en ekstra innsats for å skifte ut mindreverdige koder med fullverdige koder. Også dette kan sies å være gjort en gang for alle (men kan også gjentas senere hvis ønskelig).

Oppdateringen i '88 kunne i stor grad bygge på arbeid som var gjort i '87 (som vist i kap. 3) når det gjaldt programmer. Men bearbeidingen av filene ble gjort praktisk talt fra begynnelsen av.

Dokumenteringen av de 87-filene som 88-oppdateringen bygde på inngår ikke i samme systemet som selve 88-oppdateringen. I de etterfølgende kapitler vil de aktuelle 87-filene bli beskrevet under merkelappen "Utg.pkt."

Oppdateringen i '88 i hovedtrekk:

0. Utg.pkt.: tre filer med innflyttinger 67-70, 71-74 og 75-84. Fob70-fil, fil med fødte 1966-84, linkfil pr. 30.09.87.
1. L-delen: Linkfilen oppdateres til kjøredato for situasjonsfilen 1.1.88, altså 8.4.88. L-delen må gjennomføres først fordi de andre delene er avhengige av en fungerende linkfil.
2. A, B og D-filer bearbeides så langt at de får det riktige antall poster (dvs. at poster t.o.m. 1987 tas med, og at dubletter lukes ut). (C-filen er ikke nevnt her, da det ikke finnes noen poster på den som ikke også finnes på A-filen.
3. K-delen: A, B og D-filen kobles mot situasjonsfilen for å finne fram til de som mangler. Disse blir igjen fordelt på "innflyttere" og "fødte".
4. J-delen: J-filen lages av statsborgerskapsendringer 1977-87.
5. A- og B-filene forbedres av J-filens opplysninger om tidligere statsborgerskap ved koblinger der mindreverdige koder byttes ut med fullverdige.
6. A- og B-filene forbedres av situasjonsfilens opplysninger om nåværende statsborgerskap ved koblinger der mindreverdige koder byttes ut med fullverdige.
7. "Innflyttertillaget" og "fødtetillaget" fra K-delen blir lagt til hhv. A-filen og D-filen.
8. E-delen: A-, B-, C- og D-filen blir sydd sammen til en E-fil, som altså er det som kalles "Fødelandsfilen".
9. F-delen: Situasjonsfilen blir koblet med Fødelandsfilen og resultatet blir "Fødelandssituasjonen pr. 1.1.88".
10. M-delen: Mors, fars og ektefelles fødselsnummer trekkes ut av situasjonsfilen. Etter tre koblinger med fødelandsfilen er deres fødselsnummer byttet ut med fødeland.
11. F-delen på nytt: M-delen kobles til fødelandssituasjonen, og vi får utgave nr. 2 av den.

En kortere måte å beskrive gangen i det hele er følgende: (L H) A B D K J A B C D E F M F. Senere kan det gjøres så enkelt: (L H) A B C D E F (M F). Når L og H er satt i parentes, så er det fordi disse leddene strengt tatt ikke er en del av selve fødelandsprosjektet. L(inkfilen) kan og bør brukes også til andre prosjekter. Nødvendigheten av H er et praktisk spørsmål som har med bl.a. diskplass å gjøre. (M F) er bare et tillegg.

#### 4.2.2. Programnavn

De første to posisjoner i et programnavn er lik redigeringsnummeret til den filen som programmet lager (eks.: "D1"). Dernext kommer et siffer som angir sifferet i utgavenummeret (slik at en får f.eks. "D12"). At bare ett siffer er avsatt til utgavenummeret forutsetter at bokstaven i utgavenummeret alltid er A. De fire neste posisjonene er identiske med generasjon. Noen programnavn har et tillegg som S eller R. De brukes der programmet ikke endrer noe filnavn (evt. at bare versjonsnummeret endres). S brukes ved sortering, og R ved "reparasjoner". Oppsummert så får vi at programmet D126687 laget filen P-22.S0130.I755D1A2.G6687.V00. Lagringskoden framgår ikke av programnavnet. (Kontorkoden 22 er etter at kjøringene foregikk blitt forandret til 214).

#### 4.2.3. A-delen: Innflyttere - statsborgerskap og fraflyttingsland.

Innvandringsmeldingsfilene kan naturlig deles i 4: 67-70, 71-74, 75-84 og 85-87. Kodene for statsborgerskap og land er ikke de samme i 6770 som i 7174, som igjen er forskjellig fra kodene som er brukt i 75-filen og senere filer. 85-87-filene er spesielle ved at de inneholder flere opplysninger enn de tidligere filer. Alt dette innebærer at filene må omkodes før de kan slås sammen. Omkodingen er gjort etter en kodenøkkel som vises i vedlegg 1.

6770-filen måtte tas med for å dekke innvandringen i de to månedene mellom Fob70 og 1. januar 1971. Når vi likevel måtte befatte oss med filen var det greit å ta med hele perioden 67-70. De fleste av disse ville en vente skulle finnes igjen i Fob70-filen, med de muligheter for sammenligning det innebar.

Framgangsmåte:

Utg.pkt - Tre filer (67-70, 71-74, 75-84) var laget høsten 87 ved å trekke ut innflyttere (posttype 2) fra innvandringsmeldingene. Bare relevante kjennemerker var tatt med. For perioden 75-84 hadde innvandringsmeldingene foreligget som årsfiler, men årsuttrekkene hadde blitt konkatenerert til en fil 75-84.

De tre filene hadde hver for seg fått rettet fødselsnummer (og blitt sortert på nytt), men dette var også den eneste form for bearbeiding. Statsborgerskaps- og landkoder hadde ikke blitt omkodet på disse filene.

All6770 - Trekker ut fødelandsrelevante opplysninger (fødselsnummer statsborgerskap, fraflyttingsland og registreringsdato/-hendelsesdato) om innflyttere til Norge 1967-70 (fra en fil laget høsten '87) og omkoder de gamle kodene for statsborgerskap og fraflyttingsland. Dessuten blir registreringsdatoen snudd til år-mnd-dag. Antall utposter: 63753.

- A117174 - Som A116770, men statsborgerskaps- og landkodene som skal omkodes er fra en annen kodeliste. Innflyttere 1971-74. Antall utposter: 74277.
- A117584 - Trekker ut opplysninger om innflyttere til Norge 1975-84, og retter/omkoder noen få statsborgerskaps- og landkoder. Antall utposter: 193258.
- A118500 - Trekker ut opplysninger fra meldingsfilen for innvandringer 1985. Lager dessuten 'første oppholdsdato'.
- A118600 - Som over, men for 1986.
- A118700 - Som over, men for 1987. Merk at 'fødeland', som fra 1.7.87 er med på meldingsfilen for innvandringer ikke tas med her. Antall utposter: 37139.
- A116787 - Konkatenering av All-filene til en stor fil med alle innflyttinger 1967-87.
- A116787S - Filen sorteres.
- A126787 - Kobler mot linkfilen for å bytte ut endrede fødselsnummer. 660 rettinger.
- A126787R - Retting av et kjennemerke som hadde blitt feil i A12.
- A126787S - Filen sorteres.
- A136787 - Luker ut postdubletter. Antall utposter: 414643. Denne filen er altså en fil med innflyttinger til Norge 1967-87 (ikke medregnet "innflyttertillaget" A31).
- A146787 - Fødselsnummerdubletter lukes ut på en måte som gjør at den "beste" posten i hvert dublettpar blir tatt med. Dette er poster med koder som ikke er restkoder eller uoppgitte koder. Samtidig blir tidligste registreringsdato registrert under navnet 'første oppholdsdato'. Antall poster: 367627. Denne filen og de senere A-filene er altså filer med innflyttere til Norge.
- A156787 - Kobler mot fil med statsborgerskapsendringer (se J) for å bytte ut mindreverdige statsborgerskapskoder med fullverdige.
- A166787 - Som A15, men nå kobles det mot situasjonsfilen 1.1.88.
- A216787 - Teller opp fnr-dubletter og legger antallet på egen fil.
- A310000 - Fra K-filen tas ut de manglende poster som skal defineres som innflyttere, og legger dem på en A-merket fil (kalles heretter "innflyttertillaget"). Det er personer med 'registreringsstatusdato' forskjellig fra fødselsdato som defineres som innflyttere. Antall utposter: 4618.
- A316787 - Kobler innflytterfilen (A16) med opptellingsfilen (A21)
- A326787 - Konkatenerer A316787 og innflyttertillaget A310000. Antall utposter: 372245.
- A326787S - Filen sorteres
- A416787 - Lager en "fødelandsfil" basert bare på innflyttere. Statsborgerskap og fraflyttingsland sammenlignes, og en av de to blir erklært som fødeland. Dessuten blir 'kodekvalitet' laget som en del av sammenligningsprosessen.

#### 4.2.4. B-delen: Fob70 - fødeland

Kildefilen er en 70-fil (kalt K70) der flest mulig av kjennemerkene er omkodet til 80-standard. 'Fødeland' er et av disse omkodede kjennemerkene. I prinsippet skulle en derfor kunne ta 'fødeland' direkte. Imidlertid var noen koder blitt feil, slik at det likevel var nødvendig å foreta noen omkodinger. Dessuten var det noen spesielle koder som ikke samsvarte med den kodenøkkel som var blitt utarbeidet for fødelandsprosjektet.

## Framgangsmåte:

- Utg.pkt - En Fob70-fil hadde blitt laget høsten '87 ved å trekke 'fødselsnummer' og 'fødeland' fra Folketellingskontorets 70-fil. Denne filen ble så utsatt for fødselsnummerretting med påfølgende sortering. Deretter hadde dubletter blitt luket ut.
- B117000 - Kobler mot linkfilen for å bytte ut endrede fødselsnummer. Samtidig omkodes/rettes noen fødelandskoder.
- B117000S - Filen sorteres.
- B127000 - Luker ut eventuelle fødselsnummerdubletter. Antall utposter: 3873975.
- B137000 - Kobler mot filen med statsborgerskapsendringer (J-filen) for å bytte ut mindreverdige fødelandskoder med fullverdige statsborgerskapskoder.
- B217000 - Som B13, men nå kobles det mot situasjonsfilen 1.1.88. I de tilfeller der personens registreringsstatus-dato på situasjonsfilen er mindre enn 701101, blir B-filens hendelsesdato satt lik personens reg.status-dato (dog ikke lavere enn 641001).

4.2.5. C-delen: innflyttere - fødeland

Denne delen er ny i 88-oppdateringen, og er dessverre preget av det. Situasjonen da C-delen ble laget var at meldingsfilen for 1987 var feil oppdatert mht. de nye fødelandsopplysningene. Forsøket på å finne hva som var feil lyktes bare delvis, bl.a. på grunn av feilopplysninger fra systemkontor/personregister. Letingen etter feilen truet med å stoppe hele prosjektet, så til slutt ble meldingsfilen brukt som den var, til tross for at fødeland manglet for alle familieinnflyttere og til tross for at det åpenbart var for mange utenlandske statsborgere som oppga at de var født i Norge.

Først i ettertid er feilen helt oppklart. Egentlig er feilen todelt:

1. Meldingsfilen med innvandring fram t.o.m. mars 1988 har tatt med fødeland bare for enkeltinnflyttere, ikke familieflyttere. Denne mangelen har ikke hatt andre konsekvenser enn at C-delen ikke har blitt utnyttet i så stor grad som det var grunnlag for. Mangelen kan bli utbedret til neste oppdatering.

2. Fødeland ble ikke registrert før i juli 1987, ikke fra 1. juni som det ble oppgitt flere ganger. Alle innflyttere registrert i juni hadde blitt registrert med blankt fødeland. Blank ble stående som '000' på meldingsfilen, og fødeland Norge ble dermed overrepresentert.

Hvilke betydning denne feilen har tas opp i kap. 5. Ved neste oppdatering vil feilen ikke bli gjentatt.

## Framgangsmåte:

- C118700 - Trekker ut fødelandsrelevante opplysninger (inkludert 'fødeland' fra en meldingsfil med innflyttinger registrert i perioden 1.1.87 - 31.05.88. Poster med reg.-dato før 870601 og etter 880408 tas ikke med, og heller ikke dem med fødelandskode blank eller '990'. (Merk at 870601 viste seg å være feil dato).
- C118700S - Filen sorteres.
- C128700 - Kobler mot linkfilen for å bytte ut endrede fødselsnummer.
- C128700S - Filen sorteres.
- C138700 - Fødselsnummerdubletter lukes ut. Antall utposter: 17769.

4.2.6. D-delen: Fødte - fødeland Norge

Fødtefilene byr på små utfordringer sett fra et fødelandssynspunkt: alle får fødeland Norge.

Riktignok finnes det f.o.m. 1.7.87 et kjennemerke 'fødeland' på fødte-meldingsfilene, men dette kjennemerket blir ikke brukt. Det forteller om faktisk fødeland for personer som kommer til verden mens mora (som er registrert bosatt i Norge) er en kortere eller lengre tur utenlands. Noen av disse tilfellene skyldes nok at mora ikke har husket på å levere melding om utflytting til folkeregisteret. Bakgrunnen for at 'faktisk fødeland' ikke blir brukt er både av praktisk art (systemet ville måtte omarbeides fullstendig), og av prinsipiell art. Det prinsipielle består i at 'fødekommune' defineres som 'mors bostedskommune ved fødselsen', noe som er en forbedring fra tidligere tider da "fødestuekommunen" ble registrert. Det ville være riktigst å definere fødeland på en parallell måte til fødekommune. Dessuten ligger metoden med bruk av statsborgerskap som fødelandssubstitutt nærmest den definisjonen.

## Fødekommune

'Fødekommune' var mer problematisk enn fødeland. De kommunenummer som stod på filene var selvsagt de nummer som gjaldt på fødselstidspunktet. Det er altså kommunenummer fra og med 1966 vi taler om. Noen av disse kommunenumrene eksisterer ikke lenger, og noen av dem står for noe annet nå enn de gjorde tidligere. Av den grunn måtte kommunenumrene omkodes i størst mulig grad til de som gjaldt pr. 1.1.88.

I denne sammenhengen kan vi snakke om 4 typer kommunenummerendringer:

1. Sammenslåing av kommuner.
2. Endring av kommunenummer (at samme område skifter nummer).
3. Deling av kommuner
4. Overføring av et område fra en kommune til en annen.

Endringstype 1 og 2 byr ikke på problemer: numrene omkodes direkte. Endringstype 3 - tidligere kommuner som nå er delt - er det ikke noe å gjøre med.



Endringstype 4 kan variere i betydning. Ideelt sett burde en se på både ubebodde og bebodde overførte områder, for områder som på overføringstidspunktet er ubebodd kan ha vært bebodd tidligere.

(Ubeboede områder som senere blir bebodd er uinteressante når det gjelder fødsler). I praksis har imidlertid ubebodde områder liten betydning - det er overføring av bebodde områder som skaper problemer. Men de fleste overføringer har heldigvis karakter av ubetydelige justeringer.

Løsningen på problemene er å gi en tilleggskode (kalt 'kommuneårgang') som forteller hvilket siste år den aktuelle kommunen eksisterte i sin daværende form.

Alle kommunekoder som står for kommuner som nå er delt har fått en årgangskode (altså type 3).

Når det gjelder overføring av personer (type 4) har vi ikke brydd oss om alle småtilfeller. Nedre grense for tildeling av årgangskode ble satt der tallet på overførte personer utgjorde 1 % av folketallet i den kommunen som personene ble overført fra. Sagt på en annen måte: for en person med en årgangskodeløs fødekommune er det 99 % sikkert at han/hun ble født i den oppgitte kommunen. (For kommuner som ikke har avgitt områder er sikkerheten 100 %, feilkoding ikke tatt i betraktning).

Den som vil nøye seg med en lavere sikkerhet kan fjerne årgangskoder på en eller flere av følgende kommuner:

Kommune/ årgang	Prosent overført
0134 67	1,6
1815 70	1,7
1941 81	3,6
1569 75	4,9
0806 67	7,4
2018 83	9,9
1724 67	8,8
1941 71	31,5

#### Framgangsmåte:

Utg.pkt - En fødtefil for årene 1966-84 hadde blitt laget høsten '87 ved å trekke levendefødte fra fødtefilene, med kjennemerkene 'fødselsnummer', 'fødekommune' og 'statsborgerskap'. De aller første årene t.o.m. 1971 måtte behandles særskilt, da innholdet på filene ikke stemte helt med det som var oppgitt. Utdragene for hvert år ble konkatenerert til en fil. Denne filen ble sortert, postdubletter ble luket ut, fødselsnummer ble rettet, og filen ble sortert på nytt.

D118600 - Trekker ut felter (fødselsnummer, fødekommune, statsborgerskap) fra meldingsfil for fødte 1986. Dette året er strengt tatt tatt med i fødtefilen fra høsten 87, men pga. usikkerhet omkring det som ble gjort da tas 86-fødte med for sikkerhets skyld.

- D118700 - Trekker ut fødelandsrelevante opplysninger fra en meldingsfil med fødte registrert i perioden 1.1.87 - 31.05.88. Poster med reg.dato etter 880408 tas ikke med.
- D116687 - Konkatenerer alle fødtefilene.
- D116687S - Filen sorteres.
- D126687 - Kobler mot linkfilen for å bytte ut endrede fødselsnummer.
- D126687S - Filen sorteres.
- D130000 - Fra K-filen. Tar ut de manglende poster som skal defineres som fødte, og legger dem på en D-merket fil (kalles heretter "fødtetillegget"). Det er personer med registreringstatusdato lik fødselsdato som defineres som fødte i Norge. Antall utposter: 1010.
- D136687 - Dubletter lukes ut fra D12. Antall utposter: 1271696.
- D146687 - Konkatenerer fødtefilen med fødtetillegget. Antall utposter: 1272706.
- D146687S - Filen sorteres.
- D216687 - Omkoder/retter statsborgerskap, og lager et nytt kjennermerke 'fødselsdato (år-mnd-dag)'.

#### 4.2.7. E-delen: Sammensyningen til fødelandsfil

Prinsippet i programmene er - som tidligere nevnt - at fødeland fra D går foran C, C foran B og B foran A. I A-delen går statsborgerskap foran fraflyttingsland. Unntak fra denne hovedregelen skjer når en kode fra en høyere prioritert kilde er av dårligere kvalitet (dvs. er restkode eller uoppgittkode) enn en kode fra en lavere prioritert kilde. Dessuten brukes fraflyttingsland framfor statsborgerskap i 3 spesielle tilfeller der statsborgerskapskoden, selv om den er fullverdig, ikke blir vurdert å gi uttrykk for riktig fødeland.

Den viktigste grunnen til at disse programmene blir relativt lange er kodingen av 'kodekvalitet'.

- E11AB88 - Kobler A-delen og B-delen. 'Fødeland' definert i A-delen må ved kobling normalt vike for fødeland fra Fob70. Antall utposter: 4147972.
- E12AC88 - Kobler AB-filen med C-delen. 'Fødeland' i AB-filen må ved kobling vike for fødeland fra C-delen.
- E21AD88 - Kobler ABC-filen med D-delen. 'Fødeland' i ABC-filen må ved kobling vike for fødeland Norge fra D-delen. Med denne siste koblingen er Fødelandsfilen etablert. Antall utposter: 5 092 281.

#### 4.2.8. F-delen: Fødelandssituasjonen

- F118800 - Kobler Fødelandsfilen med en bearbeidet versjon av situasjonsfilen pr. 1.1.88. Resultatet blir "Fødelandssituasjonen 1988".
- F218800 - Kobler F11 med M71 slik at fødelandssituasjonen får et lite tillegg (mors, fars og ektefelles fødeland).

#### 4.2.9. G-delen: Kortversjoner

- G-delen er av rent praktisk karakter.
- G118800 - Trekker ut utlendinger (utenlandsfødte og/eller utenlandske statsborgere) fra fødelandssituasjonen utg.1.
- G218800 - Trekker ut personer med utenlandsk bakgrunn fra fødelandssituasjonen utg. 2.

4.2.10. H-delen: Situasjonsfiler

H-delen er hovedsaklig av rent praktisk karakter.  
H318800 - Trekker enkelte felter fra situasjonsfilen.

4.2.11. J-delen: Statsborgerskapsendringer

En fil med statsborgerskapsendringer siden 1977 var laget høsten '87. Den inneholdt skifte fra utenlandsk til norsk statsborgerskap for bl.a. personer som hadde flyttet til Norge flere år før, f.eks. i årene rundt 1970. På bakgrunn av at innvandringsmeldingene for den tida (og også Fob70) ble registrert med mange mindreverdige koder, var det fristende å lage en form for forbedringsfil som kunne forbedre disse kodene. De mindreverdige kodene er så problematiske å hanskles med både i programmer og i tabellproduksjon, at ekstraarbeidet med å redusere antallet av dem vil være verdt innsatsen.

J-filen, som denne rettefilen kalles, ble brukt både i A- og B-delen (s.d.). Programmene bytter ut mindreverdige koder med fullverdige koder fra J-filen hvis J-filens kode er inneholdt i den mindreverdige koden på hhv. A- og B-filen. På A-filen ble statsborgerskapskoder forbedret på denne måten (ikke fraflyttingsland), og på B-filen ble fødeland forbedret. Uoppgitt statsborgerskap var inkludert i forbedringen av A-filen, men uoppgitt fødeland ble ikke tatt med i forbedringen av B-filen.

Utg.pkt. - Som nevnt over hadde det høsten '87 blitt laget en fil med endringer av statsborgerskap for perioden 1977-84. Endringer fra utenlandsk statsborgerskap (oppgitt med gyldig kode utenom uoppgittkoder) ble trukket ut fra årsfiler. Uttrekkene ble konkatenerert til en fil, som igjen ble sortert, fødselsnummerrettet, sortert på nytt og til sist utsatt for fødselsnummerdublett-utlukning. I dublettutlukningen ble første post i dublettparet beholdt, slik at bare en persons første statsborgerskapsendring er kommet med.

- J118600 - Trekker ut 'tidligere statsborgerskap' og hendelsesdato fra meldingsfilen for statsborgerskapsendringer i 1986. Poster med ugyldige koder, uoppgittkoder og med norsk 'tidligere statsborgerskap' tas ikke med.
- J118700 - Tilsvarende som over, men for 1987 og 1988 t.o.m. mai.
- J117787 - Konkaterer de tre statsborgerskapsendringsfilene (inkludert filen fra høsten 87).
- J117787S - Filen sorteres.
- J127787 - Kobler mot linkfilen for å bytte ut endrede fødselsnummer.
- J127787S - Filen sorteres.
- J137787 - Fødselsnummerdubletter lukes ut. Den første post i hvert dublettpar beholdes. Antall utposter: 27538.

4.2.12. K-delen: Tilleggsposter

K110000 - Kobler den foreløpige A-delen, B-delen, den foreløpige D-delen og situasjonsfilen. Skriver ut bare de poster på situasjonsfilen som ikke er med på de andre filene. Antall utposter: 5 628.

4.2.13. L-delen: Linkfilen

Oppdatering av linkfilen skjedde dessverre ikke på samme systematiske måten som for de andre delene, uten at det skal ha hatt noen praktisk betydning. Linkfiloppdateringen ble nemlig gjort så tidlig at dokumentasjonssystemet ikke var kommet skikkelig i gjenge.

Etter alle kjøringer ble det oppdaget at det var 15 (post-)dubletter på linkfilen. Pga. av et uhell måtte linkfilen oppdateres på nytt, og denne oppdateringen ble gjort på en litt annen måte enn ved første gangs oppdatering for å bli kvitt dublettene og for å dokumentere alle ledd.

13 av dublettene skrev seg fra linkfilen vi hadde anskaffet fra Statens datsentral. De 2 resterende skyldtes at 2 poster fra journalfilen for oktober 1987 også befant seg på linkfilen som bare gikk fram t.o.m. september.

Framgangsmåte slik det skjedde i 88-oppdateringen (ledd i parentes er dårligere dokumentert enn ønskelig):

Utg.pkt - Linkfil anskaffet fra SDS som inneholdt alle fødselsnummerendringer siden 1964 t.o.m. september 1987. Antall poster: 1 921 105.

L118788 - Trekker ut felter for fødselsnummerendring fra en journalfil. Redigerer feltene. Tar bare endringer som har skjedd 880408 eller tidligere. Dette programmet gjentas for månedene oktober 87 t.o.m. april 88. Konkatenering skjer i editor.

L218788 - Trekker ut 'utgått fødselsnummer' og 'siste fødselsnummer' fra L11.

L228788 - Sorterer filen etter 'siste fødselsnummer'.

L238788 - Kobler L21 og L22, som er den samme filen men sortert på forskjellig måte. Der det kobler blir det "aller siste fødselsnummer" valgt, på bekostning av "nest siste fødselsnummer".

(L238709) - Fra linkfilen pr. september 87 trekkes endringer der utgått fødselsnummers registreringsstatus er '5' (utgått D-nummer), '6' (dublett) eller '9' (retting av fødselsdato eller kjønn). Status '7' (adopsjon) blir ikke tatt med fordi utgått fødselsnummer ikke har link til siste fødselsnummer. Å få inn blanke fødselsnummer på våre filer har ingen hensikt. Status '8' (dødfødt) er det ingen vits å ta med heller. Den nye filen har bare med kjennemerkene 'utgått fødselsnummer' og 'siste fødselsnummer'. Antall utposter: 188 805.

(L238709S)

- Filen sorteres etter utgått fødselsnummer.
- L238804 - Kobler L238788 og den tidligere linkfilen pr. 1.10.87 (L238709). Der sistefnr på 8709-linkfilen matcher med utgåttfnr på 8788-tillegget, blir sistefnr fra 8788-tillegget skrevet ut i stedet for sistefnr fra 8709-linkfilen. Utgåttfnr fra 8709-linkfilen beholdes altså, men nå med link til aller siste fnr. Det gamle sistefnr, som nå kan betegnes "nest siste fnr", blir på den oppdaterte linkfilen et utgåttfnr med link til "aller siste fnr". Antall utposter: 704.

Tillegg ved den nye framgangsmåten: L318709 - Som L238709

- L318709S - Filen sorteres etter utgått fødselsnummer
- L328709S - Postdubletter lukes ut.
- L338804 - Som L238804
- L348804 - Postdubletter lukes ut igjen for å få med seg dem som oppstod pga. matchingen. Antall utposter: 189494.

#### 4.2.14. M-delen: Mors, fars og ektefelles fødeland

- M118800 - Trekker ut eget fødselsnummer og mors fødselsnummer fra situasjonsfilen. Antall utposter: 2351901.
- M118800S - Sorterer etter mors fødselsnummer.
- M218800 - Trekker ut eget fødselsnummer og fars fødselsnummer fra situasjonsfilen. Antall utposter: 2286495.
- M218800S - Sorterer etter fars fødselsnummer.
- M318800 - Trekker ut eget fødselsnummer og ektefelles fødselsnummer fra situasjonsfilen. Antall utposter: 1816035.
- M318800S - Sorterer etter ektefelles fødselsnummer.
- M418800 - Kobler mot fødelandsfilen og bytter ut mors fødselsnummer med fødeland (som altså blir mors fødeland). 2347198 koblinger. Antall utposter: 2351901.
- M418800S - Sorterer etter eget fødselsnummer.
- M518800 - Kobler mot fødelandsfilen og bytter ut fars fødselsnummer med fødeland (som altså blir fars fødeland). 2268189 koblinger. Antall utposter: 2286495.
- M518800S - Sorterer etter eget fødselsnummer.
- M618800 - Kobler mot fødelandsfilen og bytter ut ektefelles fødselsnummer med fødeland (som altså blir ektefelles fødeland). 1815884 koblinger. Antall utposter: 1816035.
- M618800S - Sorterer etter eget fødselsnummer.
- M718800 - Kobler M41, M51 og M61 sammen til en fil. Den består da av eget fødselsnummer, mors, fars og ektefelles fødeland.

#### 4.3. Kjennemerkedokumentasjon

I dette kapittelet skal det sies noe om kjennemerkenes innhold, bruk, verdier, feilkilder m.v.

#### 4.3.1. Kjennemerkene på fødelandsfilen.

FØDSELSNUMMER. 11 pos. Gjeldende pr. 8.04.88.

FØDELAND. 3 pos. Den begrepsmessige definisjonen på fødeland er "mors bostedsland ved fødselen". Den operasjonaliserte definisjonen av fødeland er temmelig vid: Oppgitt fødeland eller registrert fødeland Norge eller tidligst registrerte statsborgerskap eller fraflyttingsland for innflyttere til Norge. Bak "oppgitt fødeland" skjuler det seg det faktum at det har vært opp til personene selv å avgjøre hvor de er født. Om de har valgt sitt fødeland i samsvar med vår definisjon er selvsagt umulig å vite.

For koder og forklaringer, se kodeliste i vedlegg 1.

Noen koder som en bør være oppmerksom på:

#### 000 Norge

- Et spørsmål som naturlig dukker opp er hvilket fødeland barn født av nordmenn i utlandet får. De barn som var kommet hjem til Norge til folketellingen i 1970 har fått det fødeland som foreldrene oppga i tellingen. Barn kommet til Norge senere vil få fødeland Norge hvis de har norsk statsborgerskap (noe de skal ha hvis foreldrene er norske). En kunne selvsagt brukt fraflyttingsland som kilde til fødeland i de tilfellene det er åpenbart at barnet må være født i utlandet, noe det er hvis det ikke står på fødtefilen. En slik test inngår imidlertid ikke i systemet.

Fra og med 1.7.87 er det igjen opp til foreldrene å fastsette det innvandrede barnets fødeland. En eventuell tilbøyelighet til å oppgi Norge som fødeland for barnet skulle være mindre når de nettopp har kommet til landet enn den kan ha vært i Fob70, da utenlandsoppholdet kunne ligge mange år tilbake.

Hvis foreldrene har vært registrert bosatt i Norge på fødselstidspunktet, og barnet er født i 1966 eller seinere, skal barnet i prinsippet stå på fødtefilen og dermed bli erklært som norskfødt (se mer om det i kap. 4.2.5.).

#### 101 Danmark

- Omfatter også Færøyene og Grønland i de fleste tilfeller, men ikke alltid. Færøyene er forsøkt skilt ut f.o.m. 1975 da det begynte å opptre som fraflyttingsland. I tilfeller med dansk statsborgerskap og fraflyttingsland Færøyene ble 'fødeland' satt til Færøyene. Det er klart at noen dansker har blitt færøyinger på den måten, men sannsynligvis er det mange flere færøyinger som står som danskfødt (særlig p.g.a. at Fob70 ikke skilte ut Færøyene som eget land).

Det ble vurdert som for dristig å gjøre det samme med Grønland, da det ville være for stor sannsynlighet for at en person innflyttet fra Grønland ville være danskfødt. De 3 grønlenderne som opptrer på Fødelandssituasjonen 1988 er alle fra C-filen.

- 102 Grønland  
- se Danmark.
- 104 Færøyene  
- se Danmark.
- 117 Frankrike  
- omfatter delvis også Monaco og Andorra.
- 123 Italia  
- omfatter delvis også Vatikanstaten og San Marino.
- 132 Portugal  
- omfatter delvis også Azorene og Madeira.
- 137 Spania  
- omfatter delvis også Balearene, Kanariøyene og Gibraltar.
- 118 Gibraltar  
- inngår i Spania eller Storbritannia.
- 141 Sveits  
- omfatter delvis også Lichtenstein og Schaffhausen.
- 144 Forbundsrepublikken Tyskland  
- omfatter også DDR før 1975.
- 151 Tyske Demokratiske Republikk  
- inngår i Forbundsrepublikken før 1975.
- 216 Burundi  
- i 329 Rwanda i L70. (L70 = kodeliste Fob70).
- 283 Liberia  
- i 339 Sierra Leone i L70.
- 296 Malawi  
- i 326 Zimbabwe i L70.
- 398 Resten av Afrika L70  
- Angola, Botswana, Britiske områder i Vest-Afrika, Britisk Indiske Hav, Komorene, Ekvatorial-Guinea, Djibouti, Gabon, Gambia, Guinea, Kapp Verde, Kongo, Lesotho, Libya, Madagaskar, Mali, Mauretania, Mauritius, Namibia, Niger, Mayotte, Reunion, Sao Tome og Principe, Den sentralafrikanske republikken, Seychellene, Somalia, Spanske områder i Afrika, Swaziland, Tchad, Togo.
- 410 Bangladesh  
- i 534 Pakistan i L70 og LS71.
- 432 Taiwan  
- eksisterer ikke som statsborgerskapskode. Statsborgerskap Kina (Folkerepublikken) og fraflyttingsland Taiwan gir fødeland Taiwan.
- 436 Hongkong  
- personer med fraflyttingsland Hong Kong og statsborgerskap Storbritannia, Hongkong, Folkerepublikken Kina eller statsløs, har fått fødeland Hongkong.
- 488 Nord-Korea  
- del av Sør-Korea før 1975, men i praksis betyr det lite.
- 534 Pakistan  
- inkludert Bangladesh i L70 og LS71, men i praksis betyr det lite.
- 598 Resten av Asia L70  
- Omfatter følgende land: Bahrain, Bhutan, Brunei, Sri Lanka, De forente arabiske emiratene, Taiwan (uklart), Irak, Jordan, Kuwait, Kypros, Laos, Libanon, Macao, Malaysia, Maldivene, Mongolia, Oman, Nepal, Timor, Qatar, Saudi-Arabia, Sikkim, Singapore, Yemen, Syria, Nord-Yemen.

698 Resten av Nord- og Mellom-Amerika L70

- Omfatter følgende land: Jomfruøyene (am.), Barbados, Antigua, Belize, Bahamas, Bermuda, Jomfruøyene (br.), Caymanøyene, Cuba, Dominica, Den dominikanske republikken, Franske Antiller, Grenada, Guadeloupe, Haiti, Jamaica, Martinique, Montserrat, Nederlandske Antiller, Panamakanalsonen, El Salvador, Saint-Pierre og Miquelon, St. Kitts - Nevis-Anguilla, St. Lucia, St. Vincent, Trinidad og Tobago, Turks og Caicosøyene.

798 Resten av Sør-Amerika L70

- Omfatter følgende land: Guyana, Falklandsøyene, Fransk Guyana, Suriname.

FØDELANDSDATO. 6 pos. Registreringsdato for den fødelandsopplysningen som er brukt i FØDELAND (år-mnd-dag). Datoen kan være fødselsdato, 701101, eller den innvandringsdatoen som fødelandsopplysningen er fra (fortrinnsvis etter 1974). 000000 = Dato mangler

KODEKVALITET. Består av kjennemerkene 'kilde', 'kodetype' og 'sammenligningsfelter'. 7 pos.

KILDE. 1 pos. Kilde for opplysningene i FØDELAND

- 1 = Fødeland fra fødtetiler
- 2 = Fødeland fra innvandringsmeldingene
- 3 = Fødeland fra Fob70
- 4 = Statsborgerskap fra innvandringsmeldingene
- 5 = Fraflyttingsland fra innvandringsmeldingene
- 6 = Nåværende statsborgerskap

Verdi 6 er en reserveverdi for det tilfellet at fødelandsfilen ikke er uttømmende.

KODETYPE. 1 pos. Kjennemerket trengs til tolkningen av sammenligningsfeltene.

- 1 = Enkeltlandskode utlandet
- 2 = Enkeltlandskode 000 (Norge)
- 3 = Restkode
- 9 = Uoppgittkode (980, 989 eller 990)

SAMMENLIGNINGSFELTER (5 pos.) Sammenligningsfeltene viser først og fremst likhet/ulikhet mellom den koden som er valgt til å stå for fødeland og de andre kodene. Dernest vises likhet/ulikhet disse andre kodene i mellom.

Sammen med 'kilde' og 'kodetype' utgjør sammenligningsfeltene kjennemerket 'kodekvalitet'. Tolkningen av verdiene på dette kjennemerket er ikke lett å ta på sparket. I kap. 5 og vedlegg 3 vises eksempler på hvordan sammenligningsfeltene brukes.

- felt 1: Fødeland fra fødtetilene
- felt 2: Fødeland fra innvandringsmeldingene
- felt 3: Fødeland fra Fob70
- felt 4: Statsborgerskap
- felt 5: Fraflyttingsland



- X = Sammenligningsfeltet tilhører koden
- 0 = Ikke kobling, evt. ikke forsøkt kobling, eller kobling uten at sammenligning er foretatt.
- 1 = FØDELANDskoden er lik sammenligningskoden (tall for tall, ellers:
- 3 = FØDELANDskoden er inneholdt i sammenligningskoden. ellers:
- 4 = Sammenligningskoden er 000 (Norge). (Koden er ikke 000) ellers:
- 5 = Sammenligningskoden er irrelevant/ikke til hjelp. Gjelder når sammenligningskoden er uoppgitt (980, 989, 990). ellers:
- 6 = FØDELANDskoden er ulik sammenligningskoden. Ved flere sammenligningskoder: disse er innbyrdes forskjellige. ellers:
- 7 = FØDELANDskoden er ulik sammenligningskodene, men disse står sannsynligvis for det samme (de er "like", om enn ikke tall for tall). ellers:
- 8 = FØDELANDskoden er ulik sammenligningskodene, og disse er innbyrdes like. ellers:
- 9 = d.s.s 8. Brukes når det er 2 "8-par".

KOBLINGSTYPE. 2 pos. Strengt tatt går det fram av sammenligningsfeltene hvilke filer som er koblet, slik at KOBLINGSTYPE kan synes overflødig. Imidlertid letter kjennemerket kontroller, og er dessuten mer oversiktlig enn sammenligningsfeltene.

De koblingstyper som ikke skal eksistere er tilfeller der C er med i koblingen uten at A også er det. (Ved senere oppdateringer kan det bli annerledes.)

- |                                      |                                     |
|--------------------------------------|-------------------------------------|
| 01 = ABCD                            | A: Innflytterfilen med statsborger- |
| 02 = ABC                             | skap og fraflyttingsland            |
| 03 = AB                              | B: Fob70-filen                      |
| 04 = A                               | C: Innflytterfilen med fødeland     |
| 05 = A CD                            | D: Fødtefilen                       |
| 06 = A C                             |                                     |
| 07 = AB D                            |                                     |
| 08 = A D                             |                                     |
| (09 = BCD)                           | Skal ikke eksistere                 |
| (10 = BC )                           | Skal ikke eksistere                 |
| 11 = B                               |                                     |
| 12 = B D                             |                                     |
| (13 = CD)                            | Skal ikke eksistere                 |
| (14 = C )                            | Skal ikke eksistere                 |
| 15 = D                               |                                     |
| 99 = Ikke kobling med fødelandsfilen |                                     |

FØRSTE REGISTRERTE STATSBOGERSKAP. 3 pos. Statsborgerskap ved fødsel eller ved første innflytting til Norge fra og med 1966 eller gammelt statsborgerskap ved skifte til norsk statsborgerskap eller annen tilsvarende kilde.

STATSBORGERSKAPSDATO. 6 pos. Registreringsdato for den statsborgerskapsopplysningen som er brukt i 'første registrerte statsborgerskap'. År-mnd-dag.

000000 = Dato mangler

For fødte gjøres det ikke noe for å bytte ut verdien '990' (som fødte før 1985 har) med en mer brukbar kode (det ville vært mye bry for lite). En må bare leve med at mange norskfødte står med '990' som første registrerte statsborgerskap.

FØRSTE OPPHOLDSDATO. 6 pos. Dato for første opphold i Norge (= tidligste registrerte dato i våre kilder). År-mnd-dag.

000000 = Dato mangler

Merk at personer som bare står på B-filen (Fob70) kan ha lavere første oppholdsdato enn 701101. Det skyldes at 'personens registreringsstatus-dato' er hentet fra situasjonsfilen og brukt som første oppholdsdato (der regstatus-dato var mindre enn 701101). Bare 46189 personer på fødelandssituasjonen har første oppholdsdato lik 701101.

ANTALL OPPHOLD I UTLANDET. 2 pos. Egentlig telles antall registrerte opphold i utlandet f.o.m. 1967, og før 1967 for utenlandsfødte på Fob70-filen (men da bare med ett opphold). Det utenlandsoppholdet som innvandrere har hatt før de kom til Norge for første gang, telles altså med på lik linje med andre utenlandsopphold.

I ettertid er det lett å se at 70-filen ikke skulle vært med i dette kjennemerket, for da hadde det vært lett å si at det omfattet "alle utenlandsopphold f.o.m. 1967". Det får bli gjort i senere oppdateringer.

Innflyttere fra innflyttertillaget skulle stå med 00 i dette kjennemerket for å markere manglende opplysninger om antall opphold i utlandet og usikkerhet om de virkelig er innflyttere. Dette bør nok endres til neste gang, slik at de får 01 som verdi. En annen sak er at en programmeringsfeil har gjort at 00 ble til '0 ', og at siste siffer i første oppholdsdato alltid er 0.

FØDEKommUNE. 4 pos. Mors bostedskommune ved fødselen.

XXXX = Uoppgitt FØDELAND (dvs. 990)

0000 = Født i Norge iflg. FØDELAND, men med uoppgitt fødekommune

0101-2030 = Kommunekoder fra ulike tidspunkter

3101- = Født i utlandet iflg. FØDELAND

Dette kjennemerket fungerer for alle som er født i 1966 eller seinere, men også for disse er det rundt 500 personer årlig som har fødekommune '0000'.

Kodene fra 3101 og oppover består av 3 + fødelandskode, slik at hvert land utenfor Norge blir en "fødekommune".

KOMMUNEÅRGANG. 2 pos. Kjennemerket brukes bare for personer som er født i delte kommuner eller kommuner som hadde videre grenser på fødselstidspunktet enn nå. Se kap. 4.2.5.

blank = Område (dvs. kommune, Norge/utlandet) med nåværende grenser. (For land er det ikke oppgitt årgang).

XX = Uppgitt FØDELAND (dvs. kode 990)

71 = Kommune med 71-grenser

osv.

FØDELAND FRA INNVANDRINGSMELDINGENE. 3 pos.

FØDELAND FRA FOB70. 3 pos.

FRAFLYTTINGSLAND. 3 pos.

#### 4.3.2. Eksempler på tilleggskjennemerker på fødelandssituasjonen

PERSONENS REGISTRERINGSSTATUSDATO. 6 pos. Gjelder registreringsstatusdato pr. 1.1.88.

##### 1 KONTINUITET

1 = Personen har ikke vært utenlands siden første oppholds- dato (personens reg.status-dato er ikke større enn første oppholdsdato)

2 = Personen har vært utenlands siden første oppholdsdato (personens reg.status-dato er større enn første oppholdsdato)

KJØNN. 1 pos.

ALDER. 3 pos. (alder = 1987 - fødselsdato)

NÅVÆRENDE STATSBERGERSKAP. 3 pos.

BOSTEDSKOMMUNE. Kommuneinndeling som 1.1.88

KRETS. Står tomt foreløpig.

EKTESKAPELIG STATUS

ANTALL LEVENDEFØDTE BARN

ANTALL DØDFØDTE BARN

MORS FØDELAND

FARS FØDELAND

EKTEFELLES FØDELAND

## 5. KVALITETSVURDERING AV FØDELANDSFILEN

### 5.1. Innledning

Overskrifta antyder at det er fødelandsfilen som skal vurderes. I praksis er det imidlertid lettest å forstå kvaliteten på fødelandsfilen ved å se på hvordan den framtrer på fødelandssituasjonen.

Spørsmålet om kvaliteten på en slik fil som fødelandsfilen har mange aspekter. I dette kapitlet er vi imidlertid mest opptatt av kvaliteten på hovedkjennemerket, nemlig 'fødeland', på bakgrunn av den metoden som blir brukt for å lage dette kjennemerket.

Som tidligere nevnt er prinsippet i tilordningen av fødeland at koden hentes fra den datakilden som er tilgjengelig. Hvis flere kilder er tilgjengelig, hentes fødelandskoden fra den av de tilgjengelige kildene som regnes å være mest pålitelig og som har en fullverdig kode (ikke rest- eller uoppgittkode).

I tilfellet med flere tilgjengelige kilder gir 'sammenligningsfeltene' muligheter for å vurdere samsvaret mellom de ulike kildene, og dermed kvaliteten på kildene ut fra et fødelandssynspunkt. Dette gir til slutt muligheter for å vurdere kvaliteten på fødelandskjennemerket slik det framtrer på fødelandsfilen og dens avledninger.

Først vises hvilke kilder fødelandskodene kommer fra, dernest hvilke sammenligningsmuligheter som har eksistert, og til sist tas resultatet av sammenligningene opp.

### 5.2. Hvilke kilder fødelandskodene kommer fra

Etter alle kodeforbedringer og koblinger ble fødeland til syvende og sist hentet fra de kildene som tabell 5.1 viser. Det er to store kilder: Fob70 og fødtefilene. Mest usikkerhet knytter det seg til de 131 704 personene som har fått sitt fødeland fra statsborgerskap på innvandringsmeldingene. Den kilden er forøvrig klart den viktigste i tilordningen av utenlandsk fødeland, og fortjener derfor særlig oppmerksomhet. Fraflyttingsland er også usikkert, men der er antallet heller lite.

Tabell 5.2. Hvilke kilder fødelandskodene kommer fra.  
Fødelandssituasjonen

	I alt	Fødeland Norge	Utenlandsk fødeland
I alt .....	4 198 289	4 036 664	161 625
1 = Fødeland fra fødtefiler .....	1 226 331	1 226 331	-
2 = Fødeland fra innvandrings- meldingene .....	12 869	2 751	10 118
3 = Fødeland fra Fob70 .....	2 826 636	2 773 778	52 858
4 = Statsborgerskap fra innv.meld. ..	131 704	33 804	97 900
5 = Fraflyttingsland fra innv.meld. .	749	-	749

### 5.3. Kodetyper

På fødelandssituasjonen er det i alt 769 mindreverdige koder, fordelt på 604 restkoder og 165 uoppgittkoder.

I praksis kan uoppgitt settes lik fødeland Norge, og dermed slipper en det bryderiet uoppgitt medfører. Relativt mange med uoppgitt fødeland har norsk statsborgerskap, noe som indikerer at de virkelig kan være norskfødte.

Et argument mot en slik enkel løsning er at brukerne har krav på å få vite hvor mye uoppgitt et tallmateriale inneholder, for på den måten å kunne danne seg et inntrykk av kvaliteten. Men som det skulle framgå av denne dokumentasjonen er tallet 165 for uoppgitte intetsigende som kvalitetsmål.

Både uoppgitt og mindreverdige koder vil etterhvert bli mindre av i seinere oppdateringer.

### 5.4. Sammenligninger

#### 5.4.1. Sammenligningsmuligheter

Fordelingen på 'koblingstype' gir oss svar på hvilke muligheter som foreligger for å sammenligne opplysninger fra flere kilder.

Tabell 5.4.1. Befolkningen etter koblingstype. 1.1.88

I alt	4198289	
11 = B	2752402	A: Innflytterfilen med statsborgerskap og fraflyttingsland
15 = D	904592	B: Fob70-filen
12 = B D	303910	C: Innflytterfilen med fødeland
04 = A	132419	D: Fødtefilen
03 = AB	74268	
08 = A D	12203	
06 = A C	11330	
07 = AB D	5219	
02 = ABC	1539	
01 = ABCD	263	
05 = A CD	144	

For 90,3 % av befolkningen finnes det ingen sammenligningsmuligheter - det finnes bare én kilde. For resten har vi kunnet sammenholde opplysninger fra 2 kilder, og for noen svært få fra 3 eller 4 kilder.

#### 5.4.2. Sammenligningsfeltene

Først når en tar i bruk sammenligningsfeltene får en vite hvordan sammenholdingen av opplysninger fra flere kilder har falt ut. Sammenligningsfeltet viser først og fremst likhet/ulikhet mellom den koden som er valgt til å stå for fødeland og de andre kodene. Dernest vises likhet/ulikhet disse andre kodene i mellom. Om de "andre" kodene er forskjellige fra den valgte koden, men innbyrdes like, så antyder det større sannsynlighet for at den valgte koden er gal enn når opplysningene spriker i alle retninger (dette er avhengig av hvilket land en er opptatt av).

Verdiene på kodekvalitetskjennermerket er vist i vedlegg 2.

Et foreløpig inntrykk fra vedlegg 2 er at kvaliteten på fødelandskjennermerket synes å være god for de fleste. De kodekvalitetsverdier som antyder usikkerhet eller feil er det relativt få personer som har.

#### Utslaget av en datafeil

Imidlertid er det viktige unntak fra denne tendensen til god kvalitet, og det gjelder verdi 220X088 og noen andre lignende verdier som i alt omfatter ca. 1050 personer. Så mange som 1050 personer har altså fått seg tildelt fødeland Norge på grunnlag av opplysninger på innvandringsmeldingene fra andre halvpart av 1987 og de første måneder av 1988, til tross for at de har utenlandsk statsborgerskap og fraflyttingslandet er likt statsborgerskapet. Bortsett fra en håndfull av dem er dette en feil. Årsaken til feilen ble forklart i kap. 4.2.4. om C-delen.

Tabell 5.4.2. Oversikt over det tillegget i antall personer med utenlandsk fødeland som ville fått hvis datafeilen med innvandringsmeldingene ikke var gjort. Sortert etter størrelsen på tillegget i prosent. Bare land der tillegget utgjør mer enn 1 prosent av antallet på fødelandssituasjonen

Fødeland	Utenlands- fødte 1.1.88	Tillegg	Tillegget i prosent
Togo .....	13	5	38,4
Den Dominikanske Rep. .	22	2	9,0
Resten av Europa S67 ..	13	1	7,6
Uganda .....	80	5	6,2
Zimbabwe .....	24	1	4,1
Madagaskar .....	26	1	3,8
Malta .....	27	1	3,7
Chile .....	3 062	110	3,5
Kypros .....	31	1	3,2
Ghana .....	369	11	2,9
Libanon .....	268	8	2,9
Malaysia .....	122	3	2,4
Bolivia .....	86	2	2,3
Sri Lanka .....	3 281	73	2,2
Honduras .....	49	1	2,0
Sudan .....	50	1	2,0
Iran .....	2 738	50	1,8
Etiopia .....	873	16	1,8
Folkerepublikken Kina .	989	16	1,6
Algerie .....	250	4	1,6
Somalia .....	381	6	1,5
Filippinene .....	2 535	34	1,3
Guatemala .....	170	2	1,1
Hellas .....	424	5	1,1
Australia .....	662	7	1,0
Syria .....	92	1	1,0
Sierra Leone .....	95	1	1,0
Afghanistan .....	96	1	1,0
Romania .....	190	2	1,0

Tabellen viser at for de største landene gir ikke feilen særlig store utslag, med unntak av Chile, Sri Lanka og Iran.

#### 5.4.3. Sammenligningsoversikt

En oversikt over hvordan selve sammenligningene falt ut vises i tabellen nedenfor. Den viser utfallet når B sammenlignes med A. Derneft sammenlignes C med B og med A, og til sist sammenlignes D med C, med B og med A.

Tabell 5.4.3. Resultatet av sammenligningene mellom de ulike kildene, der sammenligning har vært mulig. (C sin sammenligning er vist to ganger, først slik den faktisk er på fødelandsituasjonen og dernest slik den kunne ha blitt om ikke feil nr. 2 omtalt i kap. 5.4.1 var gjort.)

	I alt	Lik	Ulik	Andel ulik
B sammenlignet med A-statsb. ....	74 129	66 653	7 476	10,1
C faktisk sammenl. med andre i alt.	3 627	2 108	1 519	
" " med B .....	1 539	1 505	34	
" " med A-statsb. ....	2 088	603	1 485	
C tenkt sammenl. med andre i alt .	3 627	3 158	469	12,9
" " med B .....	1 539	1 505	34	2,2
" " med A-statsb. ....	2 088	1 653	435	20,8
D sammenlignet med andre i alt ....	17 828	16 350	1 478	8,2
" " med C .....	407	398	9	2,2
" " med B .....	5 413	5 182	231	4,3
" " med A-statsb. ....	12 203	10 770	1 433	11,7

Tabellen viser bl.a. at statsborgerskap fra innvandringsmeldingene relativt ofte (10 - 20 % av tilfellene) er forskjellig fra fødeland, enten fødeland er hentet fra Fob70, fra innvandringsmeldinger eller fødtefiler.

Det er en positiv side ved dette, og det er at opplysningene fra B, C og D åpenbart har rettet opp en del "statsborgerskapsfødeland", noe som også var hensikten. Den negative sida ved dette er de "statsborgerskapsfødeland" som ikke er blitt rettet opp.

## 5.5. Kvalitetsdiskusjon

### 5.5.1. Kvaliteten på fødeland fra fødtefilene (D)

Kvaliteten burde i prinsippet være 100 %, men det er likevel mulig å få avvik mellom opplysningene på denne filen og de to kildene for 'oppgitt fødeland' i de tilfellene der personer oppgir faktisk fødeland i stedet for mors bostedsland ved fødselen.

### 5.5.2. Kvaliteten på fødeland fra innvandringsmeldingene (C)

Som forklart over er det en viss feil på filen i forbindelse med med C-delen, men den feilen skyldes ikke at folk har gitt gale opplysninger.



Tester på journalfiler etter juni 1987 viser at det bare er 10 - 30 personer hver måned som en kan mistenke for feilaktig å oppgi fødeland Norge (eller ev. at de ble utsatt for andres feilregistrering). Ved nærmere gransking viser det seg at de mange av disse godt kan tenkes å være født i Norge likevel. Det står likevel igjen en rest på kanskje opp mot 100 innflyttere i året der det oppgitte fødelandet (Norge) virker lite sannsynlig (når 60-årige vietnamesiske statsborgere står med fødeland Norge er det ikke troverdig).

### 5.5.3. Kvaliteten på fødeland i Fob70 (B)

På en koblet fil mellom Fob60 og -70 har det blitt gjort en test på samsvaret mellom fødelandsopplysningene i disse to tellingene. Bare 6329 eller 0,20 % av de som deltok i begge tellingene hadde oppgitt forskjellig fødeland på de to tidspunktene. Av dem som stod som utenlandsfødt i minst en av tellingene (og deltok i begge), utgjør de 6329 personene 13,6 %.

Nær halvparten av fødelandsskifterne var norsk i den ene tellingen og utenlandsk i den andre. Litt flere gikk fra utenlandsk til norsk (2575) enn den andre veien (2247). Denne tendensen hadde blitt tydeligere om ikke nesten alle bulgarere i 1970 hadde stått som norskfødt i 1960 (476 personer)! De land som var utsatt for overgang til/fra Norge var først og fremst Sverige, Danmark, USA og andre nærtstående land. Overganger mellom andre fødeland enn Norge skjedde som oftest mellom landpar som f.eks. Sverige/Danmark, Polen/Tyskland, USA/Storbritannia, Irland/Storbritannia.

En konklusjon av dette er at det riktige tallet på utenlandsfødte i 1970 kan ha vært opp mot 2575 høyere enn det Fob70-filen viser. De manglende utenlandsfødte i Fob70 er stort sett personer født i Sverige, Danmark, USA, Storbritannia og Tyskland. Fob70-tallene for disse fødelandene kan i det minste sies å være beheftet med en viss usikkerhet.

Et intressant moment å ta med seg er at selv om det er avvik mellom fødelandsopplysninger i 60 og 70, så er avvikene små målt i geografisk og kulturell avstand mellom landene.

I tillegg til usikkerheten knyttet til dem som deltok både i 60 og 70, vil det være en tilsvarende usikkerhet knyttet til dem som var med bare i 70. På den annen side reduseres problemet ved at bare en del av dem som står på Fob70-filen fortsatt er med i 1988.

Tabell 5.4.3. viser at det er 4,3 % uoverensstemmelse mellom D og B. Dette er tilfeller der utenlandsfødte blir "rettet" av D til å bli norskfødt. Dette prosenttallet kan imidlertid ikke brukes på de andre utenlandsfødte fra B (de som ikke er koblet mot D), for nettopp det at de ikke er gjenfunnet på D er et tegn på kvalitet. Avviksprosenten 4,3 er derfor relativt intetsigende i denne sammenheng.

#### 5.5.4. Kvaliteten på statsborgerskapsfødeland (A)

Utgangspunktet for fødelandsprosjektet var en idé om at statsborgerskap ved første innvandring skulle kunne brukes som indikator på fødeland. I hvilken grad har dette slått til? I alt har 131704 personer fått påført fødeland ved hjelp av statsborgerskap. For innflyttinger etter juni 1987 har vi muligheten for å kontrollere statsborgerskaps- og fraflyttingslandsopplysningene mot en fasit, nemlig fødeland. Det er fire aktuelle kombinasjoner mellom statsborgerskap og fraflyttingsland:

Kombinasjonstype 1 er det tilfellet at statsborgerskapet er norsk. Fraflyttingslandet ser en da bort fra og fødeland blir Norge. Kontrollen viser at dette er feil i 14,4 % av tilfellene. De viktigste nasjonaliteter som feilaktig blir norsk på denne måten er fra Sverige, USA, Storbritannia og Danmark.

Kombinasjonstype 2 er det når statsborgerskap og fraflyttingsland er likt. Bare i 3,7 % av disse tilfellene blir statsborgerskap feil indikator for fødeland. Den dominerende type 2-feilen er at norskfødte feilaktig blir klassifisert som utenlandsfødt.

Kombinasjonstype 3 er det når statsborgerskap og fraflyttingsland er forskjellig. Da blir statsborgerskap feil indikator for fødeland i hele 25,5 % av tilfellene. Blant annet dansk statsborgerskap kombinert med fraflyttingsland Færøyene og britisk eller kinesisk statsborgerskap kombinert med Hongkong, blir i programmene ført over til fraflyttingsland som kilde. Det fører til at feilprosenten blir redusert til 20,0 %.

Kombinasjonstype 4 brukes når fraflyttingslandet er uoppgitt (statsborgerskapet er utenlandsk). Feilprosenten er da 5,4.

#### 5.5.5. Kvaliteten på fraflyttingsland-fødeland (A)

Det sier seg selv at fraflyttingsland er en høyst usikker indikator for fødeland, så det er brukt bare i de tilfeller at fraflyttingsland tross alt er den beste kilden. Av de 745 personene med fødeland hentet fra fraflyttingslandet har 472 flyttet fra Hongkong (og har kinesisk eller britisk statsborgerskap, eller er statsløs). Fra Færøyene har 135 flyttet (og har dansk statsborgerskap). 54 har flyttet fra Taiwan (kinesisk statsborgerskap) og 21 fra Syria (statsløs). De statsløse fra Syria er sannsynligvis palestinerne, som kan tenkes å være født i flere land i Midt-Østen.

Pga. små tall er det vanskelig å foreta en kontroll med fødeland på innvandringsmeldingene, men en enkel foreløpig test tyder på at omkring 3/4 av Hongkong-tilfellene virkelig er født i Hongkong (resten stort sett i Kina). For Færøyene er tilsvarende andel 2/3.

En konklusjon er at fraflyttingsland gir et usikkert fødeland, men at det er klart bedre enn en kunne frykte. Om ikke annet blir verdensdel nokså riktig.

#### 5.5.6. Utregning og oppsummering av feilkilder

Opplysningene i D er definert som 100 % korrekt.

Feilen fra C er allerede utregnet til 1050. Alle disse er utenlandsfødte som feilaktig har blitt definert som norskfødte. Ved neste oppdatering vil feilmengden nok bli atskillig mindre (en ørliten prosent vil sannsynligvis ikke stemme med D).

Sammenligningen mellom Fob60 og -70 avslørte en liten uoverensstemmelse. Men fordi uoverensstemmelsene til dels er vanskelige å tolke må feilmengden anslås bare løselig. Om en antar at de fleste er utenlandsfødte som i minst en telling oppgir utenlandsk fødeland, så kan en gi et anslag på 5000 feil fra B-delen. Kanskje er det 1 500 for mange norskfødte. For "statsborgerskapsfødeland" er det funnet feilprosenter som kan brukes på de omlag 131 704 med statsborgerskapsdefinert fødeland. Da tar en i bruk fordelingen på kodekvalitet og finner ut hvilken kombinasjonstype de viser til:

	Komb.- type	Antall
41000X1	2	87 521
41000X5	4	510
41000X6	3	9 782
41005X1	2	11
41005X6	3	3
42000X0	1	33 787
42005X0	1	16
43000X1	2	50
43000X6	3	23
49000X1	-	1

Dette gir grunnlag for følgende regnestykker:

Komb.- An- type tall	Feil fødeland	Er egentlig utenlands- født	Er egentlig norskfødt
1 33803	x 14,4% = 4868	... x 14,4% = 4868	
2 87582	x 3,7% = 3241		... x 1,4% = 1226
3 9808	x 20,0% = 1962		... x 2,9% = 284
4 510	x 5,4% = 28		... x 1,5% = 8
Sum	10099	4868	1518

En kan altså anta at 10 099 personer på fødelandssituasjonen har feil fødeland pga. metoden med bruk av statsborgerskapskoder. Av disse skal 4868 utenlandsfødte feilaktig ha fått fødeland Norge, og motsatt skal 1518 norskfødte feilaktig ha blitt definert som utenlandsfødte. Det er altså 3350 for mange norskfødte på filen.

Resten av feilene (3713 feil) er tilfeller der utenlandsfødte har fått feil utenlandsk fødeland.

For fraflyttingslandfødeland er det regnet med en samlet feilprosent på ca. 30.

Tabell 5.5.6. Anslått antall personer på fødelandssituasjonen med feil fødeland, etter kilde. Avrundede tall

	Brutto- feil i alt	For mange norskfødte (netto)
Fødeland fra fødtefiler .....	-	-
Fødeland fra innvandringsmeldingene ...	1 050	1 050
Fødeland fra Fob70 .....	5 000	1 500
Statsborgerskap fra innv.meld. ....	10 100	3 350
Fraflyttingsland fra innv.meld. ....	250	-
I alt .....	16 400	5 900

Tabellen får tale for seg selv.

## 6. OPPDATERINGER SENERE

Fødelandsfilen måtte opprettes på én måte, men kan i prinsippet oppdateres på flere måter. Hvis en vil opprettholde og helst forbedre kvaliteten, og samtidig ikke måtte lage nye programmer, så er eneste løsning å oppdatere fødelandsfilen ved å opprette den på nytt hver gang. Imidlertid kan oppdateringen muligens forenkles etterhvert fordi opplysningene om fødeland på innvandringsmeldingene og om fødte på fødtemeldingene fra nå av vil være nok til å fange opp nykommerne. Nye opplysninger om statsborgerskap og fraflyttingsland vil bare ha interesse som sammenligningsgrunnlag (men det kan være viktig nok de aller nærmeste årene), og nye Fob70-opplysninger vil selvsagt ikke komme.

I årene framover vil fødeland fra innvandringsmeldingene bli en stadig viktigere kilde. Foreløpig har vi ikke så mye erfaring med denne kilden. Såpass vet vi imidlertid at personer som nå bare har "statsborgerskapsfødeland" vil kunne få endret fødelandet etter et utenlandsopphold. På en måte er dette bra - det betyr nemlig at våre opplysninger stadig blir forbedret. Baksida av medaljen er at en persons fødeland, som i prinsippet skulle være en fast verdi, likevel ikke blir fast. Hvor stor betydning dette har vil vi få mer oversikt etterhvert.

## Kodenøkler/lister for land/statsborgerskap

## KODENØKLER/-LISTER FOR LAND/STATSBORGERSKAP

Kodenøkkelen er satt opp med en god del skjønn og med praktiske behov i tankene. Enkelte løsninger kan nok derfor diskuteres.

- L60 = Kodeliste for fødeland i Fob60.  
 S67 = Kodeliste for statsborgerskap for 1967 (seinest) - 31.12.70.  
 L68 = Kodeliste for fraflyttingsland ???.68 - 31.12.70. Inkluderer håndskrevne koder f.o.m. en gang i 1968 (1/1 ?).  
 L70 = Kodeliste for fødeland i Fob70.  
 LS71 = Kodeliste for statsborgerskap og fraflyttingsland 1.01.71 - 31.12.74.  
 LS81 = Landkodeliste som gjelder nå, med ulike tilleggskoder.  
 S81 = Viser statsborgerskap for land uten eget statsborgerskap.

- X, XX = Landet er uttrykkelig nevnt et annet sted.  
 ' = Tilleggskoder til L81 som det har vært nødvendig å opprette.  
 L = Landkoder uten eget statsborgerskap. For flere av disse er det usikkert om de er blitt regnet som eget land (og blitt lagt under "Resten av ..." -koder), eller om de er betraktet som del av moderlandet.  
 L- = som L, men bør behandles spesielt.  
 s = Koden omfatter også et annet land på en eller flere kodelister.  
 x = Landet inngår i et annet land på en eller flere kodelister.  
 R = Restkoder  
 U = Uoppgittkode

L S L L LS S LS  
 60 67 68 70 71 81 81

```

=====
      0      01 00      000  Norge
51  X XX XX XX      101  Danmark uten Færøyene
XX  1 01 XX 01      101  Danmark (uspesifisert)
XX  X XX 02 XX      101  Danmark m/Færøyene og Grønland
      XX      101 102 L Grønland
52  2 02 07 02      103  Finland
53      XX      101 104 L-Færøyene
54  3 03 08 03      105  Island
55  4 04 03 04      106  Sverige
56      09      111  Albania
57  05 10 05      112  Belgia
58  06 11 06      113  Bulgaria
      XX      114  Andorra
59  6 08 XX 08      117  Frankrike (uspesifisert)
XX  X XX 12 XX      117  Frankrike, Monaco, Andorra
      XX      139 118 L Gibraltar
60  09 13 09      119  Hellas
61  07 14 07      121  Irland
  
```

62	10	XX	10	123	Italia (uspesifisert)
XX	X	XX	15	XX	123 Italia, Vatikanstaten, San Marino
63	11	16	11	125	Jugoslavia
		18		126	Malta
65	5	12	19	12	127 Nederland
		XX		128	Liechtenstein
64		17		129	Luxembourg
		XX		130	Monaco
66	13	20	13	131	Polen
67	14	XX	14	132	Portugal (uspesifisert)
XX	X	XX	21	XX	132 Portugal m/Azorene og Madeira
68	15	22	15	133	Romania
		XX		134	San Marino
69	17	23	17	135	Sovjetunionen
70	18	XX	18	137	Spania (uspesifisert)
XX	X	XX	XX	XX	137 Spania m/Balearene og Kanariøyene
XX	X	XX	24	XX	137 Spania m/Balearene og Kanariøyene, Gibraltar
71	7	19	04	19	139 Storbritannia og Nord-Irland
72	16	XX	16	141	Sveits (uspesifisert)
XX	X	XX	25	XX	141 Sveits, Lichtenstein, Schaffhausen
73	20	26	20	142	Tsjekkoslovakia
XX	X	XX	XX	24	143 Tyrkia
		58	27	87	143 Tyrkia
74	8	21	05	21	144 s Forbundsrepublikken Tyskland (og DDR før 1975)
XX	X	XX	XX	XX	151 x Den tyske demokratiske republikken (i 144 før 1975)
75	22	28	22	152	Ungarn
76	23	29	23	153	Østerrike
		XX		154	Vatikanstaten

9				196	R'Resten av Europa S67
	29	29		199	R'Resten av Europa L68 og LS71
		30	50	203	Algerie
				204	Angola
				205	Botswana
				139	209 L Britiske områder i Vest-Afrika
				139	213 L Britisk Indiske Hav
		XX		216	x Burundi (i 329 Rwanda i L70)
				220	Komorene
		32		229	Benin
				235	Ekvatorial-Guinea
		33		239	Elfenbeinskysten
	41	34	51	246	Etiopia
77	40	35	52	249	Egypt
				250	Djibouti
				254	Gabon
				256	Gambia
		36	53	260	Ghana
			54	264	Guinea
		XX		266	Guinea-Bissau
XX	X	XX	37	XX	266 Guinea-Bissau, Guinea Port L70
			38	270	Kamerun
				273	Kapp Verde
	44	39	55	276	Kenya
				278	Kongo
	42	40	56	279	Zaire
				281	Lesotho
		XX	57	283	x Liberia (i 339 Sierra Leone i L70)
			58	286	Libya
				289	Madagaskar

XX		296	x	Malawi	(i 326 Zimbabwe i L70)
		299		Mali	
	42	59		303	Marokko
				306	Mauretania
				307	Mauritius
				308	Namibia
				309	Niger
	45	44	60	313	Nigeria
		43		319	Mosambik
			117	322	L Mayotte
			117	323	L Reunion
	46	45	61	326	s Zimbabwe (inkludert Malawi i L68)
		31		329	s Rwanda (inkludert Burundi i L70)
				333	Sao Tome og Principe
		46	62	336	Senegal
				337	Den sentralafrikanske republikken
				338	Seychellene
		41		339	s Sierra Leone (inkludert Liberia i L70)
				346	Somalia
			137	349	L Spanske områder i Afrika
		47	63	356	Sudan
				357	Swaziland
78	43	48	64	359	Sør-Afrika
	47	49	65	369	Tanzania
				373	Tchad
				376	Togo
		50	66	379	Tunisia
	48	51	67	386	Uganda
		52	68	389	Zambia
		53		393	Burkina Faso
-----					
79				395	R'Resten av Afrika L60
	49			397	R'Resten av Afrika L68
		54		398	R'Resten av Afrika L70
			69	399	R'Resten av Afrika LS71
=====					
		74	70	404	Afghanistan
				409	Bahrain
	XX	XX		410	x Bangladesh (i 534 Pakistan i L70 og LS71)
				412	Bhutan
			139	416	L Brunei
		75	71	420	Burma
				424	Sri Lanka
				426	De forente arabiske emiratene
	50	76	72	428	Filippinene
			484	432	L-Taiwan
	51	77	73	div 436	L-Hongkong (St.b.skap 139, 436, 484 el. 980)
87	52	78	74	444	India
88	53	79	75	448	Indonesia
			76	452	Irak
	54	80	77	456	Iran
	55	81	78	460	Israel
89	56	82	79	464	Japan
			80	476	Jordan
		83		478	Kampuchea
86	57	84	81	484	Folkerepublikken Kina
	XX	XX		488	x Nord-Korea (i Sør-Korea i L70 og LS71)
		85	82	492	s Sør-Korea (inkl. Nord-Korea i L70 og LS71)
				496	Kuwait
				500	Kypros
				504	Laos

		83		508	Libanon
			132	510	L Macao
				512	Malaysia
				513	Maldivene
				516	Mongolia
				520	Oman
		84		528	Nepal
		86	85	534	s Pakistan (inkl. Bangladesh i L70 og LS71)
				537	Timor
				540	Qatar
				544	Saudi-Arabia
				546	Sikkim
				548	Singapore
				556	Yemen
				564	Syria
		87	86	568	Thailand
XX	X	XX	XX	XX	572 (Nord-Vietnam) nå: 575
		88	88	575	Vietnam ((572 + 576))
XX	X	XX	XX	XX	576 (Sør-Vietnam) nå: 575
				578	Nord-Yemen
-----					
		90		595	R'Resten av Asia L60 (muligens inkludert Tyrkia)
			59	597	R'Resten av Asia L68
			89	598	R'Resten av Asia L70
			89	599	R'Resten av Asia LS71
=====					
				684	601 L Jomfruøyene (am.)
					602 Barbados
			139	603	L Antigua og Barbuda
				604	Belize
				605	Bahamas
			139	606	L Bermuda
			139	608	L Jomfruøyene (br.)
80	30	55	30	612	Canada
			139	613	L Caymanøyene
		56		616	Costa Rica
			33	620	Cuba
				622	Dominica
				624	Den dominikanske republikken
			117	628	L Franske Antiller
				629	Grenada
			117	631	L Guadeloupe
		57		632	Guatemala
			34	636	Haiti
		58		644	Honduras
			35	648	Jamaica
			117	650	L Martinique
84	31	59	31	652	Mexico
			139	654	L Montserrat
			127	656	L Nederlandske Antiller
		60		664	Nicaragua
		61	36	668	Panama
			684	669	L Panamakanalsonen
				672	El Salvador
			117	676	L Saint-Pierre og Miquelon
			139	677	L St. Kitts, Nevis-Anguilla
				678	St. Lucia
				679	St. Vincent
				680	Trinidad og Tobago
			139	681	L Turks- og Caicosøyene
81	-	32	06	32	684 USA



33			697 R'Resten av Nord- og Mellom-Amerika L68
	62		698 R'Resten av Nord- og Mellom-Amerika L70
		39	699 R'Resten av Nord- og Mellom-Amerika LS71
82	34	63 40	705 Argentina
		64 42	710 Bolivia
83	35	65 41	715 Brasil
			720 Guyana
		66 43	725 Chile
		67 44	730 Colombia
		68	735 Ecuador
		139	740 L Falklandsøyene
		117	745 L Fransk Guyana
		69 45	755 Paraguay
		70 46	760 Peru
			765 Suriname
		71 47	770 Uruguay
	36	72 48	775 Venezuela
85			795 R'Resten av Mellom- og Sør-Amerika L60
	39		797 R'Resten av Sør-Amerika L68
		73	798 R'Resten av Sør-Amerika L70
		49	799 R'Resten av Sør-Amerika L71
	60	90 90	684 802 L Amerikansk Samoa
			805 Australia
		139	806 L Salomonøyene
		805	807 L Christmasøya
		805	808 L Kokosøyene
		820	809 L Cookøyene
			811 Fiji
			812 Vanuatu
			813 Tonga
		117	814 L Fransk Polynesia
			815 Kiribati
			816 Tuvalu
		684	817 L Guam
			818 Nauru
		684	819 L Midwayøyene
	61	91 91	820 New Zealand
			820 821 L Niue
			805 822 L Norfolkøya
			805 823 L Ny-Guineaterritoriet
		684	826 L Pacificøyene
			827 Papua Ny-Guinea
		139	828 L Pitcairn
		820	829 L Tokelau
			830 Samoa
		684	831 L Wakeøyene
		117	832 L Wallis- og Fortunaøyene
		117	833 L Ny-Caledonia
91			895 R'Oseania L60
	69	92 92	899 R'Resten av Oseania L68, L70 og LS71
98		93	970 U Annet (på havet, i lufta)
		98	980 U Statsløs
		&	989 U Ikke-europeiske land, også Tyrkia, unntatt USA
99	70	99 99	990 U Uoppgitt/ukjent



## Fordelingen på 'kodekvalitet'

Poeng for kvalitet er gitt skjønsmessig. Antall pluss- og minustegn gjenspeiler både antall sammenligningsmuligheter og hvordan sammenligningene har falt ut. Vurderingen av kvaliteten er vanskelig uten nærmere undersøkelser. I poenggivingen er det ikke tatt hensyn til at fødeland fra fødtefilene i prinsippet er 100 % korrekt, slik at det er uinteressant om f.eks. Fob70 viser noe annet enn Norge. Det er heller ikke tatt hensyn til at motstridende opplysninger ikke alltid er tegn på dårlig kvalitet, men tvert i mot gjenspeiler virkelige fenomener som f.eks. at norskfødte drar utenlands og skaffer seg utenlandsk statsborgerskap, for så å komme tilbake til gamlelandet med et statsborgerskap og et fraflyttingsland som ikke "stemmer" med det de har oppgitt som fødeland i Fob70 eller på innvandringsmeldingen.

Tabell V5.4. Befolkningen etter 'kodekvalitet'. Sortert etter antallet. Fødelandssituasjonen.

Verdi	Antall	Poeng	Forklaring
	4198289		I alt
3200X00	2713108	+	Norge, fra Fob70.
12X0000	904592	+	Norge, fra fødtefilene.
12X0100	303696	++	Norge, fra fødtefilene. Norge også i Fob70.
41000X1	87521	++	Utenlandsk statsborgerskap fra innvandringsmeldingene. Fraflyttingslandet er likt.
3200X10	58909	++	Norge fra Fob70. Også norsk statsborgerskap på innvandringsmeldingene.
3100X00	38787	+	Utenlandsk fødeland fra Fob70.
42000X0	33787	+	Norsk statsborgerskap fra innvandringsmeldingene.
12X0010	10770	++	Norge fra fødtefilene. Også norsk statsborgerskap på innvandringsmeldingene.
41000X6	9782	-	Utenlandsk statsborgerskap fra innv.meld., men fraflyttingslandet er forskjellig.
210X011	9236	+++	Utenlandsk fødeland fra innv.meld.. Statsborgerskapet og fraflyttingslandet er det samme som fødelandet.
3100X11	6495	+++	Utenlandsk fødeland fra Fob70. Statsborgerskap og fraflyttingsland er det samme som fødelandet.
12X0110	4944	+++	Norge fra fødtefilene. Også norsk fødeland og statsborgerskap på innvandringsmeldingene.
3100X41	3223	++	Utenlandsk fødeland fra Fob70. Norsk statsborgerskap på innv.meld., men fraflyttingslandet er det samme som fødelandet.
3100X46	1472	-	Utenlandsk fødeland fra Fob70. Norsk statsborgerskap på innv.meld. og fraflyttingslandet er ikke det samme som fødelandet.
220X110	1423	+++	Fødeland Norge fra innv.meld. Også Norge i Fob70, og norsk statsborgerskap på innv.meld.
3200X88	1390	--	Norge fra Fob70. Statsborgerskap og fraflyttingsland på innv.meld. er utenlandsk og likt hverandre. Dette er muligens hjemvendte norskfødte.
12X0088	1240	-	Fødeland Norge fra fødtefilene. Statsborgerskap og fraflyttingsland på innv.meld. er utenlandsk og likt hverandre. Dette er

			hjemvendte norskfødte.
220X088	1012	---	Fødeland Norge fra innv.meld. Statsborgerskap og fraflyttingsland er utenlandsk og like. Dette kunne være hjemvendte norskfødte, men er dessverre stort sett utenlandsfødte som ved en feil har fått fødeland Norge. Se mer etter oversikten.
3100X16	948	+	Utenlandsk fødeland fra Fob70. Statsborgerskapet på innv.meld. er det samme, men fraflyttingsland er forskjellig.
510003X	739	+	Fraflyttingsland fra innv.meld. Statsborgerskapet er sannsynligvis det samme.
41000X5	510	+	Utenlandsk statsborgerskap fra innv.meld.
210X016	370	+	Fødeland fra innv.meld. Statsborgerskapet er det samme, men ikke fraflyttingslandet.
3100X43	352	+	Utenlandsk fødeland fra Fob70. Statsborgerskapet på innv.meld. er norsk, men fraflyttingslandet er muligens det samme som fødelandet.
3300X00	343	+	Fødeland fra Fob70 i form av restkode.
3100X88	297	--	Utenlandsk fødeland fra Fob70. Både statsborgerskap på innv.meld. og fraflyttingslandet er forskjellig fra fødelandskoden, og innbyrdes like.
3100X31	297	++	Utenlandsk fødeland fra Fob70. Fraflyttingslandet er likt, og muligens også statsborgerskapet.
12X1110	258	++++	Norge fra fødtefilene. Også fødeland fra innv.meld., Fob70 og statsborgerskap viser Norge.
220X010	227	++	Fødeland Norge fra innv.meld. Statsborgerskapet er også norsk.
3100X61	204	+	Utenlandsk fødeland fra Fob70. Fraflyttingslandet er det sammen, men ikke statsborgerskapet.
210X088	199	---	Utenlandsk fødeland fra innv.meld. Statsborgerskapet og fraflyttingslandet er ikke det samme som fødelandet, men innbyrdes like.
12X0600	194	-	Fødeland Norge fra fødtefilene. Fob70 har utenlandsk fødeland.
12X0066	187	--	Fødeland Norge fra fødtefilene. Har utenlandsk statsborgerskap på innv.meld.. Fraflyttingslandet er ikke det samme.
3200X66	185	--	Fødeland Norge fra Fob70. Det utenlandske statsborgerskapet på innv.meld. og fraflyttingslandet er ikke det samme.
3900X00	164	+	Uoppgittkode fra Fob70.
12X0188	161	++-	Fødeland Norge fra fødtefilene og fra Fob70. Statsborgerskapet er imidlertid utenlandsk, og det samme som fødelandet.
3100X66	119	--	
3100X36	106	+-	
210X041	97	+	
12X1010	76	+++	
12X0166	73	+-	
3300X43	68	+	
210X061	66	+	
3300X46	58	-	
220X066	51	--	
43000X1	50	+	
12X1088	46	+++	
3300X88	30	---	
210X110	27	+++	

210X046	27	-
210X066	26	--
3300X66	26	--
210X146	23	+
43000X6	23	-
210X141	21	++
3200X77	21	+
12X0500	20	+
42005X0	16	+
12X0818	15	+-
12X1066	14	+-
220X666	13	---
3100X77	12	+
12X0616	11	+-
41005X1	11	+
3100X63	10	+-
220X188	8	+-
220X888	7	----
210X056	7	-
12X0888	7	----
220X050	6	+
210X051	6	++
12X8018	6	+-
12X0077	6	-
510005X	5	+
3300X63	5	+-
3100X13	4	+
210X446	4	--
12X1188	4	+++
12X0177	4	++-
210X411	3	+++
510034X	3	+
41005X6	3	-
220X868	2	----
220X166	2	++-
210X441	2	+-
12X8088	2	---
210X646	1	---
210X611	1	++-
210X641	1	+-
210X188	1	+-
12X8816	1	---+
12X0878	1	----
12X0886	1	----
12X0510	1	+
12X0717	1	+-
510033X	1	+
530003X	1	+
49000X1	1	+
3200X55	1	+
3100X56	1	-
3100X33	1	+

---



## Noen aktuelle filidenter

Filidentene er ikke vist med kontorkode eller lagringskode, da slike ting kan endres. Det framgår heller ikke om filene ligger på IBM eller Honeywell Bull.

## =====

FILER BRUKT SOM KILDER:

## Filer med innvandringsmeldinger

0116.A009A1A3.G6770.V00	Statistikkfil
0116.A009A5A2.G7174.V00	
0116.A009A5A2.G7500.V00	
0116.A009A5A2.G7600.V00	
	,
	,
0116.A009A5A2.G8300.V00	
0116.A009A5A2.G8400.V00	
0108.I450A6A1.G8500.V00	Meldingsfil
0108.I450A6A1.G8600.V00	"
0108.I450A6A2.G8700.V00	"
0108.I450A6A2.G8600.V	Meldingsfil som vesamfil.

## -----

Fob70-filen

0101.I056A5A2.G7000.V00

## -----

Filer med fødtmeldinger

0111.A024A1A3.G6771.V00	
0111.A024A1B0.G7200.V00	
0111.A024A2A3.G6900.V00	
0111.A024A5A1.G6600.V00	
0111.A024A5A1.G6679.V00	
0111.A024A5A1.G6700.V00	
0111.A024A5A1.G7100.V01	
0111.A024A5A1.G7200.V00	
0111.A024A5A1.G7300.V00	
	,
	,
0111.A024A5A1.G8300.V00	
0111.A024A5A1.G8400.V00	
0108.I450A3A1.G8500.V00	
0108.I450A3A1.G8600.V00	
0108.I450A3A2.G8600.V	
0108.I450A3A2.G8600.V02	
0108.I450A3A2.G8700.V00	

## -----

Situasjonsfiler

0101.I265A6A1.g8000.V00	
0108.I459A8A5.G86MC.V00	
0108.I459A8A7.G87MC.V00	Filen som ble brukt

## Statsborgerskapsendringer

0119.B002A1A2.G7700.V00

0119.B002A1A2.G8400.V00

0108.I450AAA1.G8500.V00

0108.I450AAA1.G8600.V00

0108.I450AAA2.G8700.V00

0108.I450AAA2.G8600.V

## NOEN VIKTIGE FILER FRA FØDELANDSPROSJEKTET

0130.I698A1A1.G8709.V00	Linkfilen fra SDS
0130.I698A2A1.G6770.V01	Innvandringsmeldinger, ikke ankodet
0130.I698A2A1.G7174.V01	Innvandringsmeldinger, ikke ankodet
0130.I698A2A1.G7584.V01	
0130.I698A4A3.G7786.V00	Statsborgerskapsendringer
0130.I698B1A1.G6686.V01	
0130.I698B1A2.G6686.V00	
0130.I698C1A2.G7000.V00	
0130.I698D1A1.G6786.V01	
0130.I698D1A2.G6686.V00	Fødelandsfilen pr. 1.1.87
0130.I698D2A1.G86MC.V01	Fødelandssituasjonen pr. 1.1.87
0130.I698D2A3.G86MC.V01	Kortversjon av fødelandssituasjonen 1.1.87
0130.I755A1A2.G6787.V00	
0130.I755A1A3.G6787.V00	Innflyttinger 1967-87
0130.I755A1A4.G6787.V00	Innflyttere 1967-87 (ingen fmr.-dubletter)
0130.I755A1A6.G6787.V00	
0130.I755A2A1.G6787.V00	Fil med antall opphold i utlandet
0130.I755A3A1.G0000.V00	Innvandrer tillegget
0130.I755A3A1.G6787.V00	
0130.I755A3A2.G6787.V00	Innflyttere 1967-87, inkl. innvandrer tillegget
0130.I755A4A1.G6787.V00	Foreløpig "fødelandsfil" for innflyttere.
0130.I755B1A2.G7000.V00	
0130.I755B2A1.G7000.V00	Fob70-filen etter kodeforbedring.
0130.I755C1A1.G8700.V00	
0130.I755C1A3.G8700.V00	Innflyttere med fødeland, eksl. dubletter.
0130.I755D1A3.G0000.V00	Fødtetillegget
0130.I755D1A3.G6687.V00	Fødtefilen uten fødtetillegget (eksl. dubl.)
0130.I755D1A4.G6687.V00	Fødtefilen med fødtetillegget
0130.I755D2A1.G6687.V00	Den ferdige fødtefilen.
0130.I755E1A1.GAB88.V00	Koblingen mellom A og B
0130.I755E1A2.GAC88.V00	Koblingen mellom AB og C
0130.I755E2A1.GAD88.V00	Fødelandsfilen



0130.I755F1A1.G8800.V00  
0130.I755F2A1.G8800.V00 Fødelandssituasjonen med ektefelles og foreldres fødeland.  
0130.I755F4A1.G8000.V00 Fødelandssituasjonen pr. 1.11.80  
0130.I755G2A1.G8800.V00  
0130.I755J1A3.G7787.V00 Statsborgerskapsendringer  
0130.I755K1A1.G0000.V00 Tilleggspostene  
0130.I755L2A3.G8709.V00  
0130.I755L3A2.G8709.V00  
0130.I755L3A4.G8804.V00 Linkfilen oppdatert til 8.4.88