

# Kommunedataarkivet: Et resume.

## Databeskrivelse.

Data er organiseret i 8 arkiver dækkende forskellige perioder, forskellige variable og forskelligt aggregeringsniveau.

For alle arkivers vedkommende er enheden kommunen, der er den mindste enhed der er meningsfuldt og fyldestgørende beskrevet i officiel statistik. Da afgrænsningen af kommuner har ændret sig siden begyndelsen af 1960 ved frivillige sammenlægninger og gennem tvungne sammenlægninger efter 1970 (den kommunale inddelingsreform af 1970) har det været nødvendigt at aggregere bagud i tid hvad der forårsager tab af information.

Vi vil for hvert enkelt arkiv beskrive følgende:

1. Kommuneidentifikation
2. Variable
3. Aggregeringsniveau
4. Arkiver og program<sup>mer</sup> involveret i generering af arkivet
5. Filbeskrivelse

De tre oprindelige arkiver.

Det er nødvendigt at beskæftige sig med disse arkiver for at forstå de beslutninger med hensyn til opbygningen af de efterfølgende arkiver, der er taget.

De tre arkiver er:

1. JB-ARKIVET. Data fra folketingsvalg, folketællinger og udvalgt socialforsøgs, indkomst og formuestatistik. 1920-45. I det oprindelige arkiv defineredes en ny kommuneenhed ved den mindste ændring af kommunegrænser for alle involverede kommuner. Dette medførte et frafald på 18% over perioden. I denne version anvendes samme princip som i KB-ARKIVET hvor mindst 5% af befolkningen skulle være involveret i ændringen, før der redefineredes. Dette medfører kun et tab på 7% over perioden.

Variabel 1. Kommune-id fra KB-ARKivet (se dette)

Variabel 2. Kommune-id. position 1-2 regions og amtskode

3-4 kommunenummer

5 udgave

Hvis en kommune har været involveret i en grænseændring hvor den bevarede sit oprindelige navn, er den "nye" kommune identificeret med det samme amts og kommunenummer men med et højere udgavenummer (se 2 side 7).

Variabel 3. Kommunens navn

Variabel 4. Kommunens navn (fortsat)

Variabel 5. Kommunens navn (fortsat)

Variabel 6. Dato for kommunens skabelse ( lavest mulige værdi er 200921 dato for arkivets begyndelse. Feltbredde er 6pos. yymmdd)

Variabel 7. Dato for kommune afslutning (som var. 6 højst mulige dato er 471101 )

Fil beskrivelse

Antal variable: 269

Antal cases: 1500

Fortran format:(2I7,2X,3R10,3X,11I7,I4(/,18I7),/,I7)

Sorterings orden:stigende på variabel 2

Manglende dat: 0 §

§: En variabel vil få værdien 0 hvis en af følgende betingelser er opfyldt: 1) Enheden eksisterer ikke på det givne tidspunkt  
2) Variablen er ikke blevet kodet  
3) Variablen har den gyldige værdi 0 (nul)

## OB -ARKIVET

Dette kan betragtes som en videreføring af JE-ARKIVET med data fra folketings-valg og folketællinger for perioden 1947-68. Arkivet adskiller sig på to punkter fra JE-ARKIVET.

1. For hele perioden anvendes den kommuneopdeling der eksisterede per 22. november, 1966. Dette betyder at et antal aggregeringer bagud i tid og nogle opsplittings fremad i tid er nødvendige. Opsplittingsne er foretaget med 1966 stemmetallene som vægtfaktor under antagelse om at de oprindelige kommuner havde samme fordeling af vælgere over partier som de nye sammenlagte kommuner i 1968.

~~Identifikationsvariable.~~

~~Position 1-2 Amtsnr. som JE-ARKIVET~~

~~Position 3-4 Kommunenummer som JE-ARKIVET~~

2. For kommunerne København og Frederiksberg er der sket en opsplitting på 20 valgkredse. Dette har medført nogle ubetydelige overlapninger mellem tællings og valg data.

Identifikationsvariable.

Position 1-2 Amtsnr. som JE-ARKIVET

Position 3-4 Kommunenummer do

Fil beskrivelse

Antal variable: 110

Antal cases: 1120

Fortran format: (6(18I7,/),2I7)

Sorteringsorden: Stigende på variabel 001

## KB-ARKIVET

Den oprindelige hensigt med dette arkiv var at foretage en klassifikation af lister som stillede op ved kommunevalgene, samt at undersøge de politiske partiers overtagelse af de lokale valg, for perioden 1909-66.

Da langt de fleste af de grupperede lister, ialt ca. 90 (se 1 app. B) ikke er repræsenterede ved et kommunevalg, ville en fast position medføre en enorm mængde manglende data, derfor valgtes et to-cifret identifikation af listen efterfulgt af stemmetallet på listen.

Da de fleste analyseprogrammer ikke kan håndtere en sådan datarepræsentation var det nødvendigt at passere filen gennem et konverteringsprogram, hvor brugeren kan slå listenavne sammen til et mere anvendeligt antal og hvor data er i faste variabelpositioner.

Variabelbeskrivelse (for fyldigere gennemgang se 1 side 30)

## Variabel 1. Kommuneidentifikation

position 1-2    Amtsnummer  
 position    3    Administrative kommune type  
                  1: Sognekommuner  
                  2: Købstæder, flækker og  
                                  handelspladser  
                  3: Hovedstadskommuner  
 position 4-6    Kommunenummer

## Variabel 2. Valg-type

Værdier    1: valg med to eller flere lister  
                  opstillet . Valg afholdt  
                  2: valg med en liste opstillet  
                                  valg afholdt  
                  3: valg med en liste opstillet  
                                  valg ikke afholdt  
                  4: manglende data

## Fil beskrivelse

Antal variable: 796

Antal cases: 1655

Fortran format: (44(10I7,/),4I7)

Manglende data: som note § side 42+ 4) data for enheden var ikke tilstede

Integrationen af JE,OB og KB arkiverne: IG-ARKIVET.

Da det er en fordel om alle oplysningerne i et arkiv om en enhed findes registreret i arkivet under enheden, har man oprettet et arkiv med oplysninger fra de tre oprindelige arkiver.

For at kunne sammenlægge geografisk identiske kommuner har man oprettet et ID-ARKIV (se 1 app. E) som indeholder identifikation fra de tre oprindelige arkiver. Denne information er input til et integrationsprogram sammen med de tre arkiver og out-put fra programmet er IG-ARKIVET.

Da ikke alle kommuner eksisterer i de tre arkiver er identifikationsvariable 001,002 og 003 i ID-ARKIVET udfyldt med tallet 9 i hele identifikationsfeltet for de ikke eksisterende kommuner. Når integrations-programmet støder på en række 9-taller i en identifikationsvariabel i ID-ARKIVET fylder det alle variable med nuller så de genererede enheder får samme længde.

Programmet har et antal optioner så brugeren kan skabe forskellige versioner af IG-ARKIVET:

1. Ved at omsortere ID-ARKIVET før program exekvering kan sekvensrækkefølgen af enheder ændres.
2. Brugeren kan bestemme antallet og rækkefølgen af de oprindelige arkiver som skal medtages.
3. Ved at anvende specielle formattsætninger er det muligt at udtrække et subsæt af variable fra et givent arkiv.
4. Ved at generere et subsæt af ID-ARKIVET kan skabes et subsæt af det mulige IG-ARKIV.
5. Liner fra ID-arkivet kan skrives på den tilsvarende enhed i IG-arkivet.

For hver enhed i IG-ARKIVET skrives første linie fra enhedens record i ID-ARKIVET, derefter følger JE,OB og KB arkiverne.

Filbeskrivelse.

Antal variable: 1187

Antal cases: 1697

Fortran format: (I6,I6,I5,1X,3R10,I2,2I3,2I7,/,  
 2I7,2X,3R10,3X,11I7,14(/,18I7),/,/,  
 6(18I7,/),2I7,/,  
 44(18I7,/),4I7)

Som det ses af format-sætningen har hver enhed 69 linier

linie 1: 1 enhed fra ID-ARKIVET

linie 2 til 17: 1 enhed fra JE-ARKIVET

linie 18 til 24: 1 enhed fra OB-ARKIVET

linie 25 til 69: 1 enhed fra KB-ARKIVET

Sorteringsorden: Efter linie-orden i ID-ARKIVET.

Kommunerne i ID-ARKIVET er sorteret amts-vis  
 og indenfor amtet i provinsbyer og dernæst  
 sognekommuner, og indenfor disse grupper  
 alfabetisk.

Manglende data: 0

Udover at anvendes i de tilfælde som er nævnt  
 for de enkelte arkiver, vil alle variable i et  
 eller to arkiver have værdien 0 i det integrerede  
 arkiv, hvis enheden ikke eksisterer i det/de  
 nævnte arkiver.

Dette er i ID-ARKIVET repræsenteret med 999999  
 for KB-ARKIVET, 99999 for JE-ARKIVET og 9999  
 for OB-ARKIVET (id-variable 1,2 og 3)

# Aggregering af IG-ARKIVET: G6-ARKIVET

På grund af grænse-ændringer er det ikke muligt at opnå en fuldstændig integration af de tre oprindelige arkiver.

Således eksisterer kun 61 % af 1909 kommunerne på samme form i 1966 i arkivet. ( se 1 app. A). Ydermere er frafaldet ikke tilfældigt ( se 5).

For at løse disse problemer er foretaget en aggregering således at alle kommuner i det skabte arkiv har samme udstrækning som i november 1966.

Programmet som genererer det aggregerede arkiv tager højde for fire forskellige situationer:

1. Kommunen er identisk med 1966 kommunen
2. Kommunen er indeholdt i 1966 kommunen
3. Kommunen er indeholdt i to 1966 kommuner.
4. Kommunen er indeholdt i tre 1966 kommuner.

Programmet giver brugeren mulighed for at specificere hvor stor afvigelse mellem de aggregerede kommuner og de i 1966 faktisk forekommende kommuner han vil akceptere. F. eks. vil en 10% afvigelse i befolkningens størrelse mellem de aggregerede og virkelige kommuner i 1966 medføre 64 yderligere aggregeringer så antallet af kommuner i arkivet daler fra 1102 til 1038 ( se 1 s.76).

Filbeskrivelse:

?

Antal variable: 525

Antal enheder: 1034

Portran format: I7,2X,3R103X,I7,28(/,18I7),/,16I7)

Sorteringsorden: Stigende på variabel 1 på første linie

Manglende data: 0 se kommentar på side A-2

NK-ARKIVET: Valg og folketællings-statistik fra 1970 og 1975.  
Arkivet er en fortsættelse af OB-ARKIVET, men grundet kommunesammenlægningen i 1970 er antallet af kommuner dalet fra 1107 i 1966 til 278 i 1970.

Variabelbeskrivelse. Alle variable er heltal

Variabel 1. Kommune id. -  
Pos. 1-2 Amt  
Pos. 3-4 Kommunenummer

Variabel 2. Udgave: Værdi = 0

Variabel 3. Kredsnummer  
Pos. 1-2 Kommune som variabel 1 pos. 1-2  
Pos. 3-4 Kredsnummer. (se 1 s. 01)

Filbeskrivelse.

Antal variable: 126

Fortran format: (6(18I7,/),18I7)

Antal enheder: 278.

GN-ARKIVET: Integration af IG-ARKIVET og NK-ARKIVET.

På samme måde som for G6-ARKIVET, men på grundlag af 1970 kommunestrukturen (oplysninger herom for en enhed findes på linie 3 af enhedens information i ID-ARKIVET) aggregeres IG-ARKIVET til G7-ARKIVET.

Da brugeren kan afgøre hvor stor en afvigelse i befolkningsstørrelse mellem faktisk eksisterende og aggregeret kommune han vil acceptere er det nødvendigt at foretage den samme aggregeringsprocedure på NK-ARKIVET, hvorved der fremkommer et arkiv med samme struktur som G7-ARKIVET, men med data fra efter 1970.

Det i forløbs-diagrammet nævnte IN-ARKIV er en speciel version af ID-ARKIVET som indeholder oprindelige og ændrede kommune-id og navne.

Slutteligen integreres G7-ARKIV og N7-ARKIV til GN-ARKIV.

Filbeskrivelse:

Antal variable: 649

Antal enheder: 235

Fortran format: (I7,2X,3R10,3X,2I7,28(/,18I7),/,16I7,6(/,18I7),/,15I7)  
note: linie 1 indeholder 6 id-variable VAR04 - VAR06

linie 2 til linie 30 indeholder 520 udvalgte variable  
fra IG-ARKIVET : VAR007 - VAR526  
linie 31 til 37 indeholder 123 udvalgte variable  
fra NK-ARKIVET: VAR527 - VAR649

Manglende data : 0 se note side **A-2**

Litteraturliste:

1. Bentzon, K.H., Borre, O. , Elklit, J.: Kommune-data-arkivet 1909-68  
En foreløbig oversigt over det  
økologiske data-arkiv ved  
institut for statskundskab  
Århus 1972
2. Bentzon, K.H.: The danish statistical commune data archive  
Forskningsrapport 1975/5  
Institut for samfundsfag  
Københavns universitet
3. Statistisk Departement: Folketal, areal og klima 1909-60.  
Statistiske undersøgelser  
nr. 10, 1964
4. Statistiske meddelelser fra hele perioden , se variabelbeskrivelser  
i 2.
5. Bentzon, K.H.: The process of Politicization of the danish  
local elections, 1909-58, pp. 180-184  
International journal of Politics, 1974, 1-2

Appendix CData Files in the Archive.C 1. Basic Data Files.

The following files contain the basic information for all file manipulations described in Appendix D.

IDARKIV

The file contains information about every commune unit in the archive which can be used for case selection, file integration and file aggregation procedures. IDARKIV is reproduced in Appendix E.

JEORIG

The file has information for the creation of JEARKIV.

OBARKIV

The file is used directly as archive file OBARKIV.

KBSTEM

The file contains information about voters and votes for the local elections 1909-66.

KBLIST

The file contains information about list types (see Appendix B.) and vote cast on the list types for the local elections 1909-66. KBSTEM and KBLIST are used for the generation of KBARKIV.

NKARKIV

The file has data in and after 1970 and is used directly in the archive as NKARKIV.

C 2. Basic Archive Files.JEARKIV

Created from JEOIG. See C 1.

KBARKIV

Created from KBSTEM and KBLIST. See C 1.

Other Basic Data Files used as Basic Archive Files are: IDARKIV, OBARKIV and NKARKIV

C 3. Integrated Archive Files (IDARKIV, JEARKIV, OBARKIV, KBARKIV).IGARKIV

The file contains information from IDARKIV and from one or more of the files: JEARKIV, OBARKIV, and KBARKIV. See Section 5.

INARKIV

The file has data from IDARKIV and NKARKIV. See Section 7.

C 4. Aggregated Archive Files.

G6ARKIV

The file is a 1966 aggregation of a version of IGARKIV. See Section 6.

G7ARKIV

The file is a 1970 aggregation of a version of IGARKIV. See Section 7.

N7ARKIV

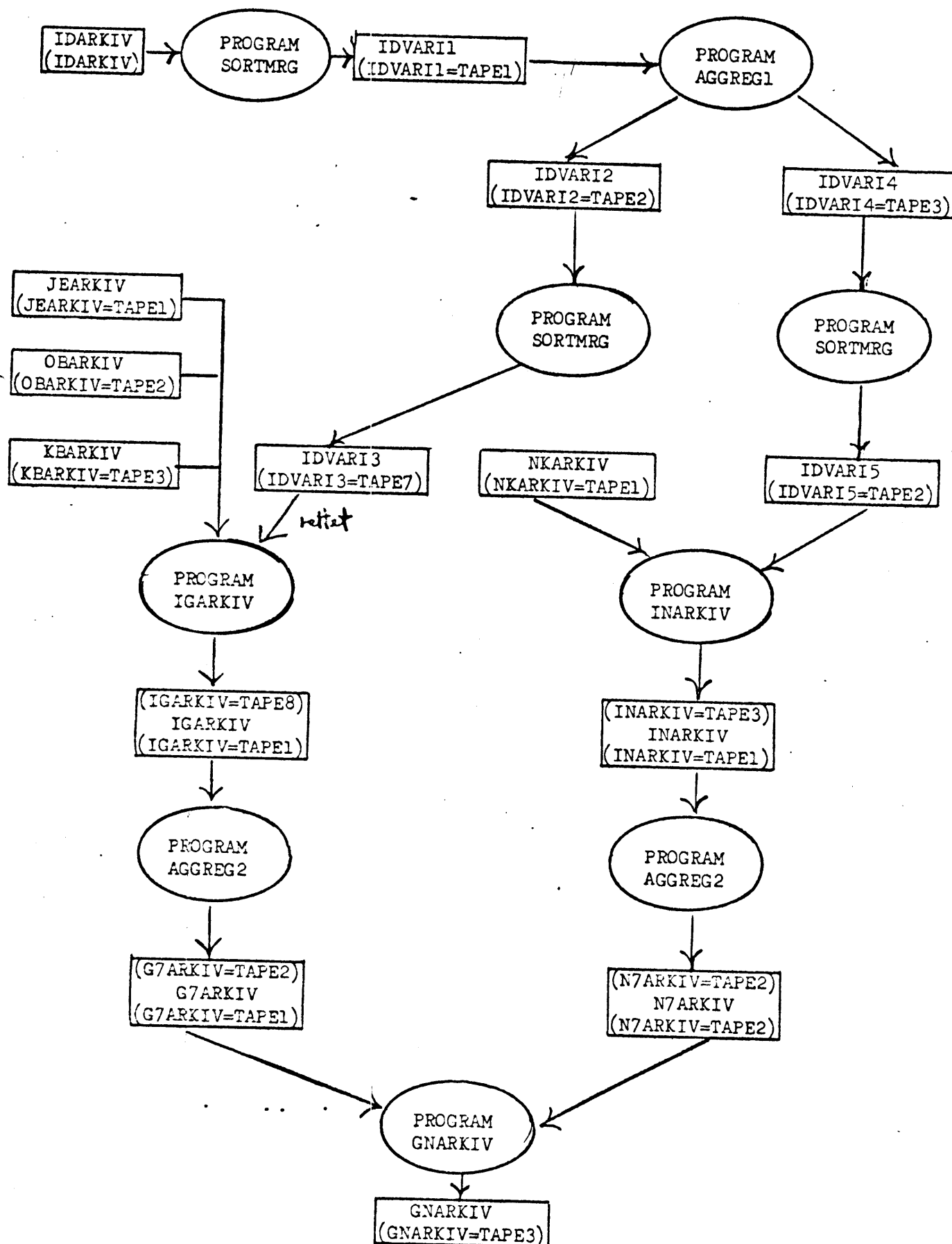
The file is a 1970 aggregation of INARKIV. See Section 7.

C 5. Integrated Archive Files (G7ARKIV, N7ARKIV).

GNARKIV

The file is a combined version of G7ARKIV and N7ARKIV. See Section 7.

Flow Chart of the Creation of the 'GN-ARCHIVE'.



Beskrivelse af kommunedataarkivet (fælles 13)

Indholdsfortegnelse:

Beskrivelse af de enkelte arkiver	A 1-9
Litteraturliste	B 1
Datafiler i arkivet	C 1-2
Forløbsdiagram over skabelsen af GN-ARKIVET	D 1