

Adgang til mikrodata i Danmarks Statistik - Datasikkerhed

Forskningservice



Datasikkerhed og Datafortrolighed

Baggrund

- Danmarks Statistiks giver adgang til mikrodata via Forskerordningen fordi det giver værdi til samfundet
- Data er fortrolige og der er derfor nogle regler for adgang og håndtering af data som forskerne skal sætte sig godt ind i, før man får adgang til data
- Kravene til sikkerhed og fortrolighed er de samme som før, men vi har skærpet opmærksomheden på at reglerne overholdes
 - Øget fokus på forskernes hjemsendte output
 - Opstramning af sanktioner for miljøer der bryder reglerne
 - Lukker ned for institutioner der hjemsender mikrodata
 - Øget reaktion på hjemsendelse af detaljerede tabeller



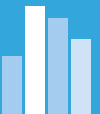
Formål med kurset

- At sætter jer ind i datasikkerhedsreglerne under forskerordningen
 - Kender regler for hvad man må hjemtage
 - Ved hvad man ikke må hjemtage
 - Ved hvad der menes med mikrodata
 - Kender "fælderne" ved hjemtagelse af resultater fra forskermaskinen
 - Ved hvad i skal gøre hvis i ved en fejl kommer til at sende mikrodata hjem



Datasikkerhed – generelle regler

- Individ-data udleveres *aldrig*. De bliver bag DST's mure, adskilt fra DST's produktionssystem
- Kun adgang til data via sikker krypteret forbindelse
- Kun adgang via autoriserede forskningsmiljøer
- Formelle aftaler underskrives
- Data stilles til rådighed efter "need to know"-princippet
- Data afidentificeres med unik nøgle på hvert projekt
- Password er strengt personlige (og må ikke lånes ud til andre)



Datasikkerhed – generelle regler

- Alt arbejde med data foregår på serverne i DST. Forskere kan følge deres arbejde på skærmen, men kan kun tage print hjem via en mail-server, hvor alt output logges
- **Det er tilladt at hjemsende resultater fra analyser, men ikke individdata** (mikrodata)
- FSE tager stikprøver af output hver dag og checker, at reglerne om datafortrolighed er overholdt



Hvad *må* I sende hjem?

- Alle former for:
 - Analyseresultater
 - aggregerede tabeller
 - figurer mv.

hvor man *ikke* kan identificere nogen enkeltindivider eller -virksomheder

- Alle celler i tabeller mv. bør indeholde mindst 3 observationer
 - I er velkomne til selv at skærpe dette



Hvad må I *ikke* sende hjem?

1. Individdata

- Individdata må aldrig sendes hjem. De skal blive på DST.
- Hvad er individdata?
 - Specifikke oplysninger enkeltpersoner/virksomheder. Det vil sige, *enhver form for individuelle "records"* eller datasæt med oplysninger på individniveau
 - Nøglevariable som personnumre, CVR-numre, adressekoder mv. opfattes *altid* som individdata og må *aldrig sendes hjem*, selvom de er afidentificerede. Heller ikke selvom de står alene

Hvad må I *ikke* sende hjem?

- Hjemsendelse af individdata betragtes som en **alvorlig** overtrædelse af datafortroligheds-reglerne. Hvis det sker, lukkes al dataadgang for pågældende forsker og alle den projektejende institutions projekter
- Pas på med programfiler og logfiler
 - Der er logfiler, der lister et udsnit af mikrodata. I skal være klar over hvad de programmer I anvender udskriver.
 - Pas på med at programmere på individniveau i programmer



Hvad må I *ikke* sende hjem – fiktive eksempler - Mikrodata!

PNR	KØN	INDKOMST	BOPÆL	ALDER	PNRB
1	M	1500000	København	40	15
2	M	1500	Aarhus	45	22
3	K	1000	Bornholm	50	37
4	M	500	Bornholm	50	65
4	M	250	København	40	87
5	M	150	Aarhus	20	
6	K	150	Odense	25	
7	M	750	Herning	35	19
8	K	500	København	60	74
9	M	250	Odense	65	74
10	K	150	Aarhus	20	
11	K	50	Bornholm	15	
12	M	0	København	10	



Hvad må I *ikke* sende hjem – fiktive eksempler - Mikrodata!

KØN	INDKOMST	BOPÆL	ALDER
M	1500000	København	40
M	1500	Aarhus	45
K	1000	Bornholm	50
M	500	Bornholm	50
M	250	København	40
M	150	Aarhus	20
K	150	Odense	25
M	750	Herning	35
K	500	København	60
M	250	Odense	65
K	150	Aarhus	20
K	50	Bornholm	15
M	0	København	10

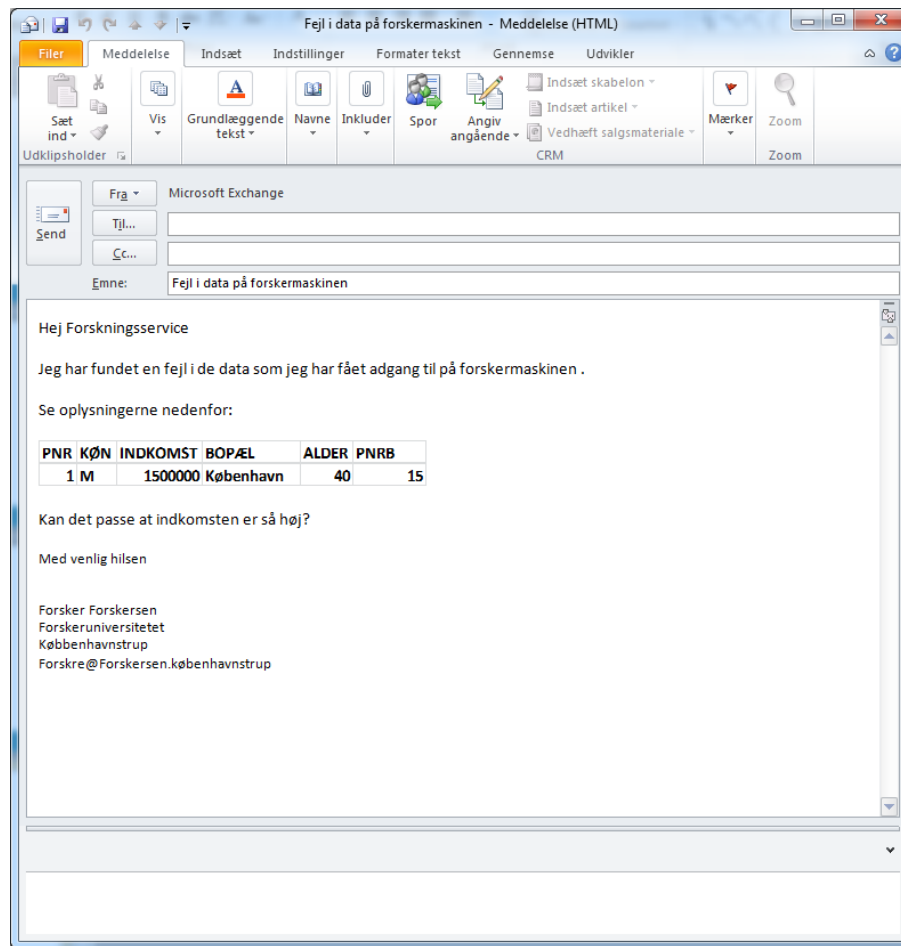


Hvad må I *ikke* sende hjem – fiktive eksempler - Mikrodata!

PNR
1
2
3
4
4
5
6
7
8
9
10
11
12



Hvad må I *ikke* sende hjem – Fiktive eksempler - mails, der indeholder mikrodata !- heller ikke til fejlsøgning!



Hvad må I *ikke* sende hjem – Fiktive eksempler - programmer der indeholder mikrodata !

SRV12 WIN 6.2.9200 Server

SAS initialization used:
real time
cpu time

Velkommen
Hvis du får

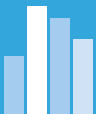
SAS 9.4 ve
opdateres
SAS du ska

Du kan også
automatisk
Det kan sp

AUTOEXEC

AUTOEXEC

```
Editor - Untitled1 *  
  
Data a;  
  set b;  
  if pnr=1 then delete;  
  if pnr=12 then delete;  
  if alder<20 then delete;  
run;  
  
Proc sort .....  
Run;
```



Hvad *bør* I ikke sende hjem?

2. Analyseresultater, aggregerede tabeller mv.

- Analyseresultater, aggregerede tabeller, figurer mv. hvor man *kan* identificere enkeltindivider eller virksomheder. Dvs. tabeller mv. hvor der optræder meget få (færre end 3) personer i cellerne.
- Da der er tale om aggregerede data, betragter vi sagen som mindre alvorlig, end hvis der er tale om individdata.
- Der kan være situationer, hvor det er OK, at enkelte celler indeholder mindre end 3 obs. (gns./max./min.)
- Brug jeres sunde fornuft her! Er der risiko for, at man kan identificere enkeltpersoner, så lad være med sende det hjem.



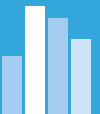
Hvad *bør* I ikke sende hjem?

- Vi reagerer på detaljerede tabeller fx hvis der er mange et-taller – men – vi lukker IKKE institutionen ned hvis vi vurderer at der har fundet en aggregering sted
- Her er normalt blot tale om en orienterende og vejledende henvendelse
- Særlig regler for erhvervsdata, jf. generelt datafortrolighedspolitik i Danmarks Statistik



Hvis vi opdager et brud på datafortroligheden (hjemsendelse af mikrodata)

- Vi lukker for alle forskere fra den institution, der ejer projektet hvor databrudet har fundet sted
 - - også hvis det er en tilknyttet forsker fra en anden institution der har begået bruddet
- Kontakter institutionen (forsker og ansvarlig)
 - Redegørelse for databrudet
 - Plan for hvordan institutionen fremadrettet vil undgå nye brud på datasikkerheden
- Danmarks Statistiks direktion orienteres og tager stilling til sagen
- Sanktion meldes ud til institutionen



Hvad kan institutionen gøre ved et regelbrud?

- Orienter altid snarest Danmarks Statistik ved et regelbrud
- Beskriv i mailen hvornår data er sendt hjem samt omfanget af hjemsendelsen. Adgangen til mikrodata vil fortsat blive lukket og der skal sendes en redegørelse og plan for forebyggelse
- Orienteres der hurtigt om et regelbrud betragtes det som en formildende omstændighed og kan reducere den periode man er lukket ned



Gode råd for forskeren

- For forskeren, der har adgang til mikrodata:
 - 1. Kontroller altid samtlige filer, der skal hjemsendes
 - Kontakt altid Forskningservice i tvivlstilfælde
 - 2. Begræns hjemsendelsen af filer til det mest nødvendige.
 - a. Er det fx nødvendigt at hjemsende programmer og logfiler? – de opbevares sikkert under forskerordningen
 - b. Er det nødvendigt at hjemsende foreløbige resultater?



Gode råd for den ansvarlige for autorisationen

- For den ansvarlige for autorisationen:
 - 1. Lav en plan der sikrer at alle forskere er orienteret om reglerne vedrørende datafortrolighed og datasikkerhed under forskerordningen
 - a. Herunder også at både den sidst tilkomne og den udenlandske forsker har tilstrækkelig viden om reglerne!
 - b. At der i miljøet findes en kontaktperson ved tvivlstilfælde
 - 2. At alle autoriserede forskere/analytikere du har ansvar for kender proceduren ved et regelbrud
 - 3. At alle autoriserede forskere er bekendt med de "fælder" der kan være med at arbejde med mikrodata – fx at output kan indeholde mikrodata
 - 4. Lav evt. kontrolforanstaltninger ved hjemsendelse af resultater
 - 5. Jævnligt at kontrollere at kun de relevante forskere har adgang til mikrodata til institutionens projekter og orientere Forskningservice hvis forskere skal have deaktiveret deres adgang



Nyt system til hjemtagelse af filer taget i brug

- Ikke alle filtyper kan sendes hjem fremover
- Vi kan lukke for enkeltforskeres adgang til at sende filer hjem
- Automatisk kontrol af om filerne indeholder mikrodata - får en advarsel før i sender output hjem
- I år forskermaskine med tekstbehandlingsprogrammer, så rapporten også kan skrives på forskermaskinen



Vejledning i hjemsendelse af resultater på vores hjemmeside

- Vejledning tilgængelig på vores hjemmeside under linket:
<http://www.dst.dk/da/TilSalg/Forskningservice/Vejledninger>
- Retningslinjer for hjemsendelse af analyseresultater
- Hjemsendelse af filer (hjemsendelsessystem)
- Hjemsendelse af filer - tilladte filtyper

