

## Anvendelse af eksterne data

### Anvendelse af eksterne data

Til forskning og analysearbejde er det ofte nødvendigt at koble oplysninger fra flere forskellige datakilder. Danmarks Statistiks data kan samkøres med forskerens eget datamateriale – eksterne data - fx surveys eller udtræk fra registre, som ikke findes i Danmarks Statistik.

Forskningservice stiller krav til de eksterne data som indsendes til brug i forskningsprojekter. Kravene skal sikre, at der kun indsendes data som skal bruges i projektet og for at sikre, at Forskningservice får de oplysninger som er nødvendige for at kunne overføre data til det konkrete projekt på forskermaskinen.

### Krav til indsendelse af eksterne data.

#### Krav til eksterne data

Data må generelt kun indeholde oplysninger der skal bruges til det konkrete projekt. Variabler der ikke skal anvendes, skal derfor fjernes før data uploades/indsendes.

#### Kun de nøglevariable der skal bruges:

- Kun de nøglevariabler – fx personnummeret - der skal bruges til at koble data med de øvrige data på projektet på forskermaskinen må medtages. Øvrige nøglevariabler fx fremmednøgler, som ikke skal bruges, skal fjernes før data uploades/indsendes til Forskningservice

#### Kun numeriske variable og kategoriserede tekstvariable

- Kun de numeriske og kategoriserede variabler som skal bruges til projektet må medtages. Variabler der ikke forventes brugt skal slettes, inden data indsendes til Forskningservice

#### Ingen navne og besvarelser med fritekst

- Data skal overholde de samme krav om diskretion som Danmarks Statistik i øvrigt benytter. Dette betyder at oplysninger som tilnærmelsesvis identificerer personer eller virksomheder ikke må indsendes – fx personnavne og besvarelser med fritekst i surveys. Sådanne variabler skal fjernes eller kategoriseres før data indsendes til Forskningservice

#### Fil-formater

#### Fil-formater

Eksterne filer må indsendes i følgende formater:

1. ASC
2. CSV
3. DTA
4. ODS
5. SAS7BDAT
6. SAV
7. XLS
8. XLSX

#### Dokumentation af eksterne data

#### Dokumentation

Eksterne data skal dokumenteres og indeholde nedenstående:

1. projektnummer data skal overføres til
2. navn eller navne på variable der skal afidentificeres, dvs. de variable som skal bruges til at koble data med øvrige data på forskermaskinen.
3. Hvis de eksterne data indeholder flere nøgler der skal bruges for at koble fx flere eksterne filer sammen, så skal disse variable også fremhæves
4. Når de eksterne data er uploadet og dokumentation indsendt, så anbefales det at advisere projektets kontaktperson i Forskningservice pr. mail

Send dokumentationen til projektets kontaktperson i Forskningservice.

Data kan leveres til Danmarks Statistik på nedenstående måder

*Anvendelse af Upload-løsning*

1. Forskere med projektadgang kan anvende af vores Upload-løsning. En vejledning heri findes på vores hjemmeside under linket: <http://www.dst.dk/da/TilSalg/Forskningservice/Vejledninger>

*Anvendelse af sikker e-mail*

2. Sende data via sikker e-mail:  
Data kan sendes via sikker e-mail til [forskerpost@dst.dk](mailto:forskerpost@dst.dk) på en af to følgende måder:
  - a) Institutionen henter Danmarks Statistiks certifikat på [www.dst.dk/certifikater](http://www.dst.dk/certifikater). Vælg [forskerpost@dst.dk](mailto:forskerpost@dst.dk). Vcf-formatet anbefales til Outlook.  
Data skal herefter sendes til [forskerpost@dst.dk](mailto:forskerpost@dst.dk). Angiv altid projektnummer og projektansvarlig i Forskningservice i emnefeltet på e-mailen og orienter den projektansvarlige i Forskningservice direkte, når data er afsendt.  
Danmarks Statistik yder ikke support på kryptering og Digital Signatur. Der henvises i stedet til [www.nemid.nu/digital\\_signatur](http://www.nemid.nu/digital_signatur) - herunder særligt [https://www.nemid.nu/dk-da/kom\\_i\\_gang\\_med\\_nemid/sikker\\_e-mail/send\\_og\\_modtag\\_sikker\\_e-mail/](https://www.nemid.nu/dk-da/kom_i_gang_med_nemid/sikker_e-mail/send_og_modtag_sikker_e-mail/)
  - b) Institutionen kan anvende tunnelkryptering (SEPO). Opsætningen er individuel for forskellige institutioner og forstås af institutionens IT-afdeling, som bør inddrages før data indsendes. Angiv altid projektnummer og orienter om at mailen er sendt med tunnelkryptering (SEPO). Anvendes tunnelkryptering skal data ligeledes sendes til [forskerpost@dst.dk](mailto:forskerpost@dst.dk).

*Anbefalet brev / personlig aflevering*

Af datasikkerhedsmæssige grunde foretrækkes det at en af de to ovenstående muligheder anvendes, men hvis det ikke er muligt (fx hvis filerne er meget store), så kan passwordbeskyttede eksterne data også leveres direkte til Danmarks Statistik på nedenstående adresse. Projektets kontaktperson i Forskningservice (FSE) samt projektnummer skal altid fremgå af det indleverede materiale:

Danmarks Statistik, Servicedesk  
Sejrøgade 11  
2100 København Ø

Att. Projektets kontaktperson i Forskningservice  
Vedr. projektnummer: 70XXXX

Data kan enten leveres personligt i Danmarks Statistiks reception eller blive sendt anbefalet til Danmarks Statistik på DVD, CD-rom eller USB-stik (leveres ikke tilbage). I begge tilfælde skal data være passwordbeskyttet. Når data er modtaget, så skal password sendes på e-mail til projektets kontaktperson i Forskningsservice.

*Send IKKE data på almindelig mail* Eksterne data med nøglevariable der henviser til personer eller virksomheder (fx CPR-nummer, CVR-nummer, løbenummer mv) må ikke fremsendes med almindelig e-mail eller post til Forskningsservice.

*Andre dataleverandører* Eksterne data kan også leveres fra andre dataleverandører direkte til Danmarks Statistik på projektforskerens bestilling og efter aftale med Forskningsservice og dataleverandøren. Levering af data skal ske på en af ovenstående måder, men I skal sikre, at data opfylder betingelserne ovenfor før data indsendes til Forskningsservice.

*Afidentificering* Når de eksterne data er modtaget i Forskningsservice, så bliver de afidentificeret på samme måde som de øvrige data der hører til projektet. Dette sker ved at nøglevariable afidentificeres. Derefter lægges data tilgængeligt på forskermaskine sammen med de øvrige data. Leveringstiden fra modtagelse af data til data er klar på forskermaskinen er maksimalt 2 måneder.

*Pris* Danmarks Statistik Forskningsservice tager betaling for de timer, der er forbundet med håndteringen af de eksterne data efter gældende timetakst.