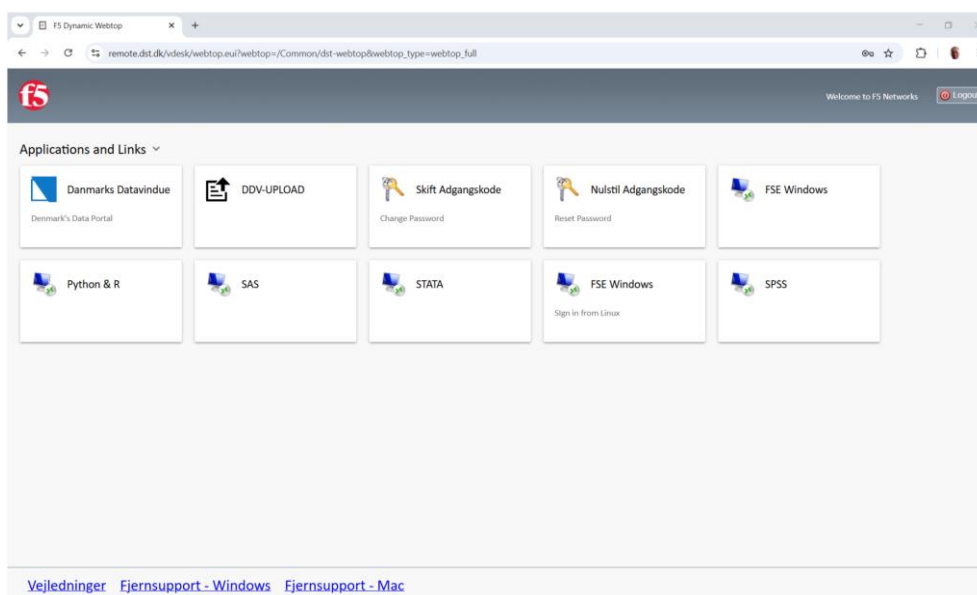


Arbejde på Forskermaskinen

Som bruger kan du enten være logget på Danmarks Statistiks fælles miljø ("Forskermaskinen") eller på en hosted server. Denne vejledning beskriver det setup du møder, når du er logget på DST's fælles miljø.

Frontend-server Fra remote.dst.dk kan du enten logge på en specifik applikationsserver, fx SAS eller STATA, eller på frontend-serveren (FSE Windows).

På frontend-serveren kan du ikke arbejde direkte med data. Denne server giver kun adgang til applikationsserverne, hvor der er installeret analysesoftware.



Du kan bl.a. arbejde med SAS, STATA, SPSS, R og PYTHON på vores servere. Se den komplette liste "Hardware og software på Danmarks Statistiks forskermaskiner" under:

<https://www.dst.dk/da/TilSalg/data-til-forskning/analyseplatform/forskermaskinen>

SAS- og STATA- og Python/R-serverne kører med Microsofts teknologi "Remote Desktop Connection Broker".

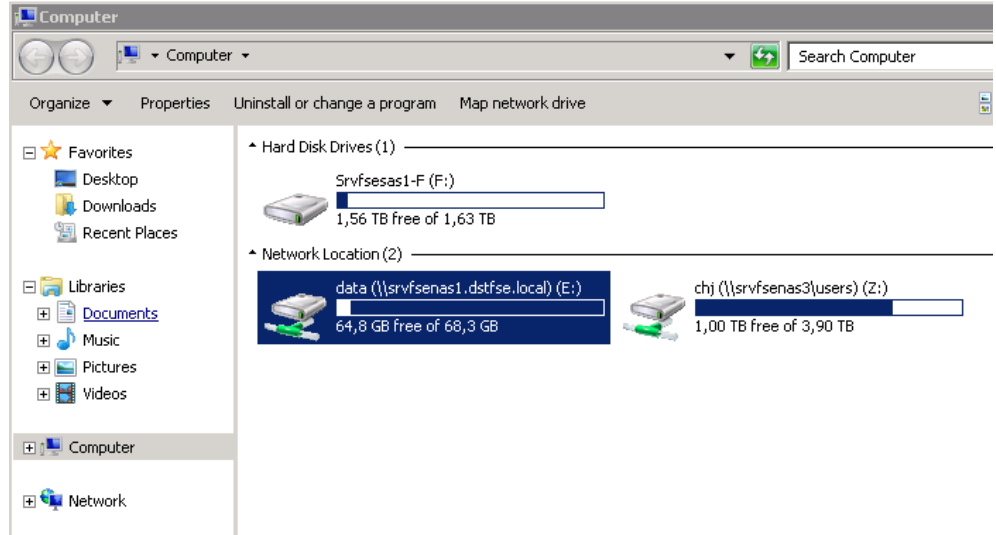
Dvs. at der kun er ét ikon for henholdsvis "SAS" og "STATA" og "Python & R", som tilslutter dig den mindst belastede server.

Du vil altid blive tilsluttet den samme server, hvis du i forvejen bare har en "disconnectet" session (klikket på krydset og ikke logget af).

Hvis du har logget af en server, vil du derimod blive tilsluttet den mindst belastede server.

Dataområder

- Adgang til samme filer Vores servere benytter en fælles filserver. Det betyder, at alle data er tilgængelige fra alle servere.
- E-drevet Data er placeret på E-drevet. Det er et fælles diskområde, som blandt andet indeholder mapperne Rawdata, Workdata og Formater.



Data er organiseret efter projektnumre.
Projektnummer består af 6 cifre, og starter med 7.

Rawdata

- Datavejledning Her placerer Danmarks Datavindue data for det givne projekt.
E:\Rawdata\<projektnummer>

Data vil normalt være placeret i tre undermapper afhængigt af typen.

-  Eksterne data
-  Grunddata
-  Population

Data er skrivebeskyttet og kan kun rettes af Danmarks Datavindue. Derved kan vi skelne mellem leverede datasæt og eventuelle filer dannet senere i analyseprocessen.

Workdata

- Arbejdsområde Workdata er arbejdsområdet på projektet.
E:\Workdata\<projektnummer>

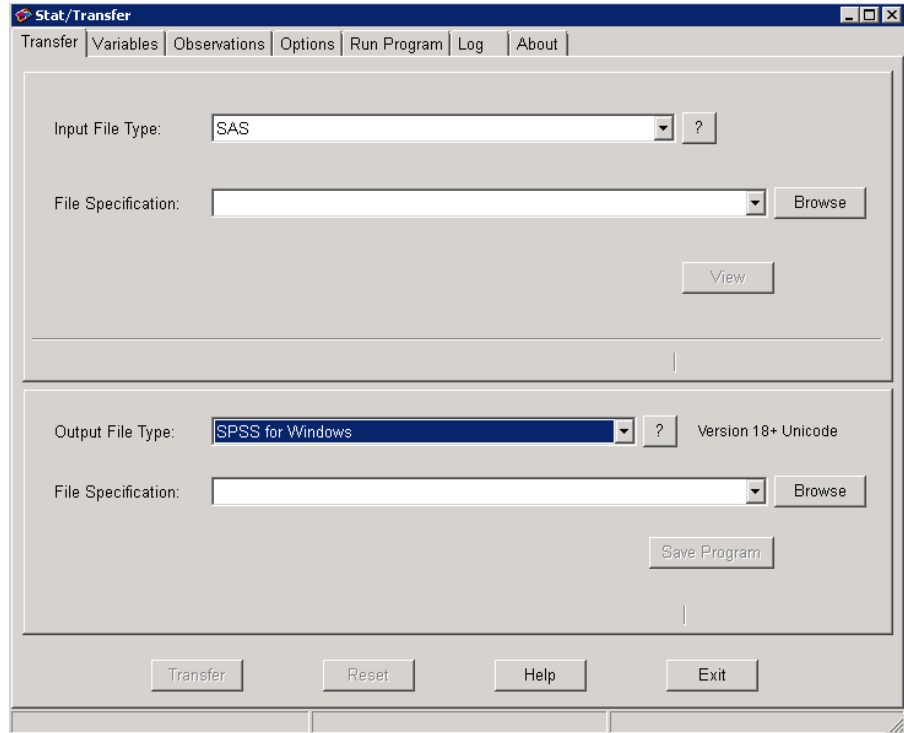
Her kan alle brugere, der er tilknyttet projektet læse og skrive data. Her skal både programmer, afledte data og færdige resultater lagres.

På det enkelte projekt bestemmer man selv en struktur på mapper og data i workdata.

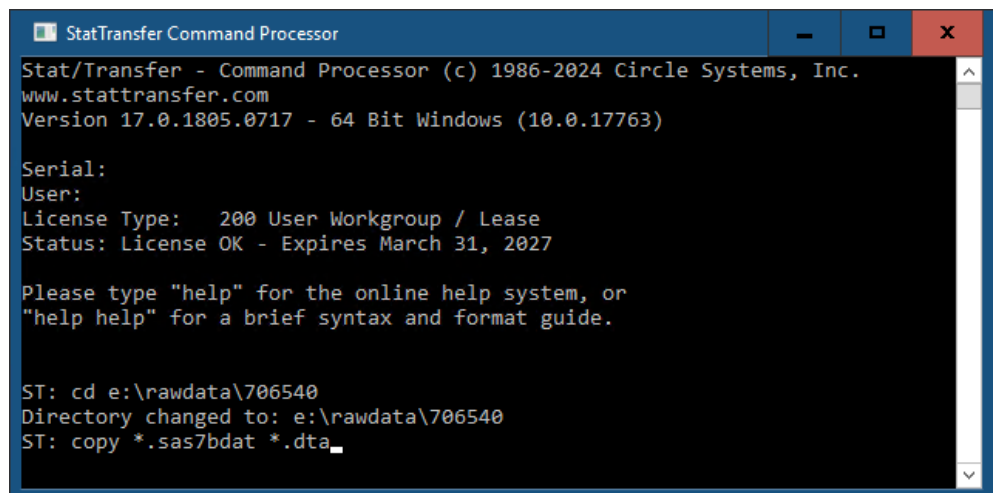
Konvertering af data

Det er muligt at konvertere data fra en type til en anden og skrive dem til work-data. En del analysesoftware kan foretage denne konvertering fx fra SAS til SPSS.

Det er også muligt at anvende Stat Transfer, der kan foretage en konvertering af langt de fleste datatyper.



Stat Transfer Command Processor kan anvendes til at konvertere flere filer samtidig – her konverteres fx alle SAS-filer i mappen til STATA-filer:



Formater

Formater Forskningservice stiller en række formater og kode-teksttabeller til rådighed for alle brugere.

Ved at klikke på genvejen ”DST-Formater” på skrivebordet får du adgang til en oversigt over Danmarks Statistiks formater.



Hjemtagelse af færdige resultater

Det er muligt at hjemtage de færdige analyseresultater.

Se vejledning under:

<https://www.dst.dk/da/TilSalg/data-til-forskning/regler-og-datasikkerhed/regler-for-hjemtagelse-af-analyseresultater>

Backup af data

Backup	Der skrives fuld backup af alle servere, som gemmes i op til 5 år.
Fælles diskområde	Det fælles diskområde (E-drevet) benytter snapshots, som dannes hver dag.
Restore	Du kan finde ”Vejledning til hurtig genskabelse af filer” her: https://www.dst.dk/da/TilSalg/data-til-forskning/analyseplatform/forskermaskinen

De hostede servere benytter båndbackups. Så her skal du henvende dig til ServiceDesk@dst.dk, hvis du ønsker noget genskabt.

Hostede servere

Det generelle setup på hostede servere ligner i udgangspunktet det fælles miljø. Der kan dog være forbehold, som denne vejledning ikke dækker. Læs mere på:

<https://www.dst.dk/da/TilSalg/data-til-forskning/analyseplatform/hosted-server>

Genstart af servere

Alle serverne bliver genstartet tidlig mandag morgen. Derfor bliver alle brugere automatisk logget af en gang om ugen. Har du et program der skal køre længe, så sæt det derfor på i starten af ugen, så det kan nå at køre færdigt.

Styring af RAM-forbrug på serverne

På applikationsserverne i det fælles miljø vil du automatisk blive logget af, hvis dit forbrug af RAM (hukommelse) overstiger 250 Gb. Hvis en server samlet har under 10 % ledig RAM, vil den bruger med det største forbrug blive logget af, selvom forbruget er under 250 GB.

Det er derfor meget vigtigt, at du **gemmer din kode inden du afvikler dine analyser**, da der ikke gemmes, hvis du automatisk bliver logget af pga. et stort forbrug af ram.

For tips til hvordan du kan reducere dit RAM-forbrug, se ”Reduktion af RAM-forbrug i det fælles miljø” under:

<https://www.dst.dk/da/TilSalg/data-til-forskning/analyseplatform/forskermaskinen>