



Rigsarkivet

Danmarks
hukommelse

Brugervejledning til ASTA version 2.1.1

Brugervejledning til programmet ASTA (Aflevering af Statistikfiler Til Arkiv), der kan anvendes i forbindelse med produktion og test af en afleveringspakke med forskningsdata til Rigsarkivet, der overholder krav i bilag 9 til Bekendtgørelse om arkiveringsversioner nr. 128.

Version 2.1.1





Indhold

0. Læsevejledning til ASTA brugervejledningen	2
A. Vejledningens målgruppe og anvendelse	2
B. Henvi sning til øvrig vejledning	2
C. Lovgivning og retsforskrifter	3
D. Definitioner	3
1. ASTA – program til Aflevering af Statistikfiler Til Arkiv	3
A. Hvad er en afleveringspakke	3
B. Hvor finder jeg ASTA?	4
C. Systemtekniske krav til din computer, når du anvender ASTA	4
Tjekliste: Vigtige trin før brug af ASTA	6
D. Gennemgang af brugergrænsefladen i ASTA	9
E. Skab afleveringspakke	10
F. Test afleveringspakke	19
G. Kontroller at udtræk fra ASTA er tabsfrit	23
H. Rediger afleveringspakke	24
I. ASTA support i Rigsarkivet	26
Bilag 1	26
A. SPSS – forberedelse af statistikfilen før udtræk	27
B. Stata – forberedelse af statistikfilen før udtræk	34
C. SAS – forberedelse af statistikfilen før udtræk	39
D. R/RStudio – forberedelse af statistikfilen før udtræk	49





0. Læsevejledning til ASTA brugervejledningen

Offentlige myndigheder, herunder forskningsinstitutioner, er forpligtet til at aflevere en kopi af data og dokumenter af bevaringsværdige forskningsdata. Hvis de forskningsdata, der skal afleveres stammer fra statistikfiler eller tilsvarende (fx regneark), kaldes afleveringsformatet en afleveringspakke. Statslige myndigheder skal aflevere til Rigsarkivet. Kommuner og regioner kan vælge, om de vil aflevere til Rigsarkivet eller oprette deres eget arkiv.

Rigsarkivet har fastsat en række bestemmelser for en afleveringspakke af hensyn til bevaring og fremtidig brug af data, som alle myndigheder skal overholde ved aflevering. Disse bestemmelser er beskrevet i Rigsarkivets bekendtgørelse om arkiveringsversioner, bilag 9: Afleveringspakke for visse typer af forskningsdata.

Rigsarkivet har udviklet værktøjet ASTA (Aflevering af Statistikdata Til Arkiv), der kan anvendes til produktion og test af en afleveringspakke. ASTA kan downloades fra [Rigsarkivets hjemmeside](#).

Med ASTA kan du udtrække data- og metadatafiler automatisk fra statistikprogrammerne SAS, Stata, SPSS og R/RStudio. I bilag 1 findes vejledning i, hvordan du i de forskellige statistikprogrammer SAS, Stata, SPSS og R/RStudio sikrer dig, at dine statistikfiler overholder krav i bilag 9 før udtræk med ASTA.

ASTA brugervejledningen beskriver hvordan du anvender programmet, men også hvad du skal have klar før du begynder en produktion af en afleveringspakke med ASTA.

A. Vejledningens målgruppe og anvendelse

ASTA brugervejledningen henvender sig til dem, som producerer afleveringspakker med data fra statistikfiler i formaterne SAS, Stata, SPSS og R/RStudio, der afleveres til Rigsarkivet. Det kan fx være myndighedens egen datamanager, it-afdeling, forskeren selv eller en ekstern it-leverandør.

B. Henvisning til øvrig vejledning

Foruden ASTA brugervejledningen har Rigsarkivet udarbejdet andre vejledninger, der har betydning for produktion og aflevering af afleveringspakker:

- Quickguide – til produktion og test af en afleveringspakke med ASTA
- Vejledning til bilag 9 om afleveringspakker i bekendtgørelse om arkiveringsversioner nr. 128
- Vejledning til produktion af afleveringspakke med data fra regneark eller csv-filer
- Vejledning til Skab archiveIndex
- Vejledning til Skab contextDocumentationIndex
- Vejledning om konvertering af dokumenter til TIFF
- Vejledning om UTF-8
- Eksempelafleveringspakke med statistikdata FD.18005





Rigsarkivet

Alt vejledningsmateriale samt programmer kan tilgås fra [Rigsarkivets hjemmeside](#).

C. Lovgivning og retsforskrifter

Information om lovgivning m.v. findes på [Rigsarkivets hjemmeside](#).

D. Definitioner

Afleveringspakker med data fra statistikfiler består overordnet set af kontekstdokumenter, der skal afleveres i Rigsarkivets arkivformater, udtræk af data og metadata fra de statistikfiler, som skal afleveres, samt to indeksfiler i xml-format, der indeholder overordnet metadata om de afleverede data og kontekstdokumenterne.

Arkivformater: Rigsarkivet benytter 6 arkivformater: TIFF, JPEG2000, MP3, WAV, MPEG2 og MPEG4.

1. ASTA – program til Aflevering af Statistikfiler Til Arkiv

A. Hvad er en afleveringspakke

Som en del af Rigsarkivets krav til aflevering af statistiske data, skal data fra de oprindelige statistikfiler udtrækkes til en afleveringspakke og testes inden aflevering til arkivet. ASTA er et værktøj, der er udviklet til at hjælpe med at skabe og teste en afleveringspakke.

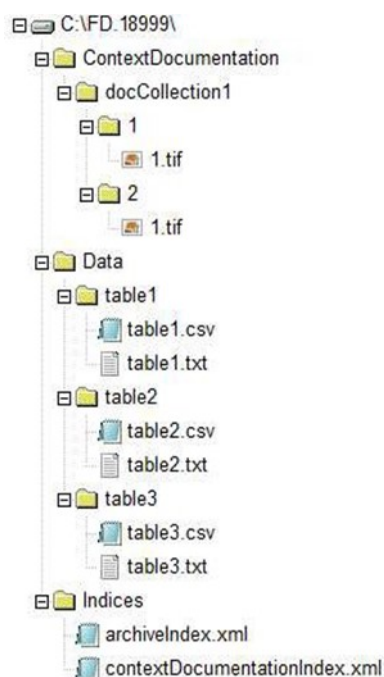
Afleveringspakken skal følge nogle specifikke krav vedrørende mappestruktur, navngivning af mapper, filer, datatyper, variabelnavne, variabellabels, formatnavne, manglende værdier osv. Alle krav til en afleveringspakke findes under Bilag 9 i Bekendtgørelse om arkiveringsversioner.

En afleveringspakke kan for eksempel se sådan ud:





Rigsarkivet



Skærbillede 1

B. Hvor finder jeg ASTA?

Versioner af ASTA til Windows og Mac kan downloades fra Rigsarkivets hjemmeside.

C. Systemtekniske krav til din computer, når du anvender ASTA

Nedenfor kan du se hvilke styresystemer, internet browsere og versioner af statistikprogrammer, som ASTA er testet til at kunne fungere på.

Anvender du andre styresystemer, internetbrowsere eller statistikprogramversioner kan det ikke garanteres, at ASTA virker fejlfrit.

Styresystemer:

ASTA er testet på følgende styresystemer:

- Windows 7 og 10
- macOS High Sierra

Internetbrowsere:

Testloggen i ASTA med visning af vejledning under fejl-ID'er virker i følgende browsere:

- Google Chrome, Version 77.0.3865.90 (Officiel version) (64-bit)





Rigsarkivet

- Mozilla Firefox, version 69.0.2 (64-bit)
- Microsoft Internet explorer 11 [Husk evt. at tillade blocked content for at se fejl-ID'erne]
- Microsoft Edge, version 17.17134
- Safari version 11.1.2 (13605.3.8)

Statistikprogrammer

Export scripts til udtræk af data og metadata fra statistikfilerne kan eksekveres i følgende versioner af statistikprogrammerne

- SAS v. 9.4
- SPSS v. 24
- Stata v. 14.2
- R v. 3.6.2 /Rstudio v 1.20.5033

Øvrige systemtekniske krav

Optimal skærmopløsning er 1920 x 1080 eller 1920 x 1200. Anvendes en mindre opløsning kan nogle knapper være skjulte i ASTA.

Da ASTA installeres på din computer, skal du have administrative rettigheder på din computer for at kunne køre programmet.





Tjekliste: Vigtige trin før brug af ASTA

Følgende trin bør tjekkes eller udføres, før du skaber en afleveringspakke med ASTA. Bilag 1 i slutningen af dette dokument giver en grundig gennemgang af, hvordan du i de forskellige statistikprogrammer SPSS, Stata, SAS og R/RStudio sikrer dig, at tjeklistens punkter 1-8 er opfyldt.

NR.	OPGAVE	UDFØRT
GENERELT		
1. Tjek/installer statistikprogram	Tjek at du har et statistikprogram, der kan åbne den fil, du skal lave udtræk fra med ASTA, installeret på den PC, du anvender ASTA på.	
2. Ny mappe med kopi af materiale	Opret en ny mappe på computeren med en kopi af alle data og dokumenter, der skal afleveres. Dvs. statistikfilerne, der skal afleveres,	
STATISTIK DATAFILER		
3. Datafiler	Datafilerne er enten i SPSS-, SAS- Stata-, og R/RStudio format. Datafilerne skal være indkodet som i UTF-8.	
4. Variable i datafiler	Variabelnavne må ikke være længere end 128 tegn, starte med et tal eller inkludere blanke mellemrum eller specialtegn, såsom %. Alle variable i filen skal have variabel labels, der beskriver deres indhold.	





Rigsarkivet

NR.	OPGAVE	UDFØRT
5. Variable type/format	Alle variable skal have tildelt type/format, der er accepteret af Rigsarkivet, jf. bilag 9, figur 9.3 i bekendtgørelse om arkiveringsversioner. De må ikke være standard formateret.	
6. Value labels	<p>Alle værdier i value labels skal være unikke og specificeret.</p> <p>I tilfælde af at nogle værdier ikke er specificeret (for eksempel i en Likert skala hvor kun den første og sidste værdi har specificeret labels), skal dette beskrives i beskrivelsen til variablen (variable label), fx med teksten "Skala anvendt. Ikke alle koder har kodebeskrivelse."</p>	
7. Missing values	<p>Manglende værdier må kun anvendes på numeriske og kategoriske variable og dette skal være konsekvent.</p> <p>Kontakt Rigsarkivets datamanager for forskningsdata, hvis du anvender koder for manglende værdier i variable med typerne tekst, tidspunkt eller dato.</p> <p>Alle koder for manglende værdier skal også tildeles en værdi i value labels.</p>	
8. Referencer (ved aflevering af flere datasæt	Når der afleveres mere end et datasæt i den samme afleveringspakke med en reference mellem disse statistiske filer (en fletningsnøgle) skal flettevariable have samme type/format og længde.	
KONTEKSTDOKUMENTATION		
9. Godkendelse fra arkivet	Kontekstdokumentationsfilen contextDocumentationIndex.xml skal godkendes af arkivet inden den indgår i afleveringspakken. Filen specificerer alle de ekstra dokumenter, der er inkluderet i afleveringspakken. Se 'Vejledning til Skab contextDocumentationindex'.	





Rigsarkivet

NR.	OPGAVE	UDFØRT
10. TIFF-dokumenter	Hvert dokument registreret i kontekstdokumentationsfilen skal konverteres til TIFF eller et andet arkivformat. Se 'Vejledning i konvertering af dokumenter til TIFF'.	
ARKIVBESKRIVELSESFIL		
11. Godkendelse fra arkivet	Arkivbeskrivelsesfilen skal godkendes af arkivet, inden den indgår i afleveringspakken. Filen indeholder overordnede metadata om de afleverede data. Se 'Vejledning til Skab archiveIndex'.	

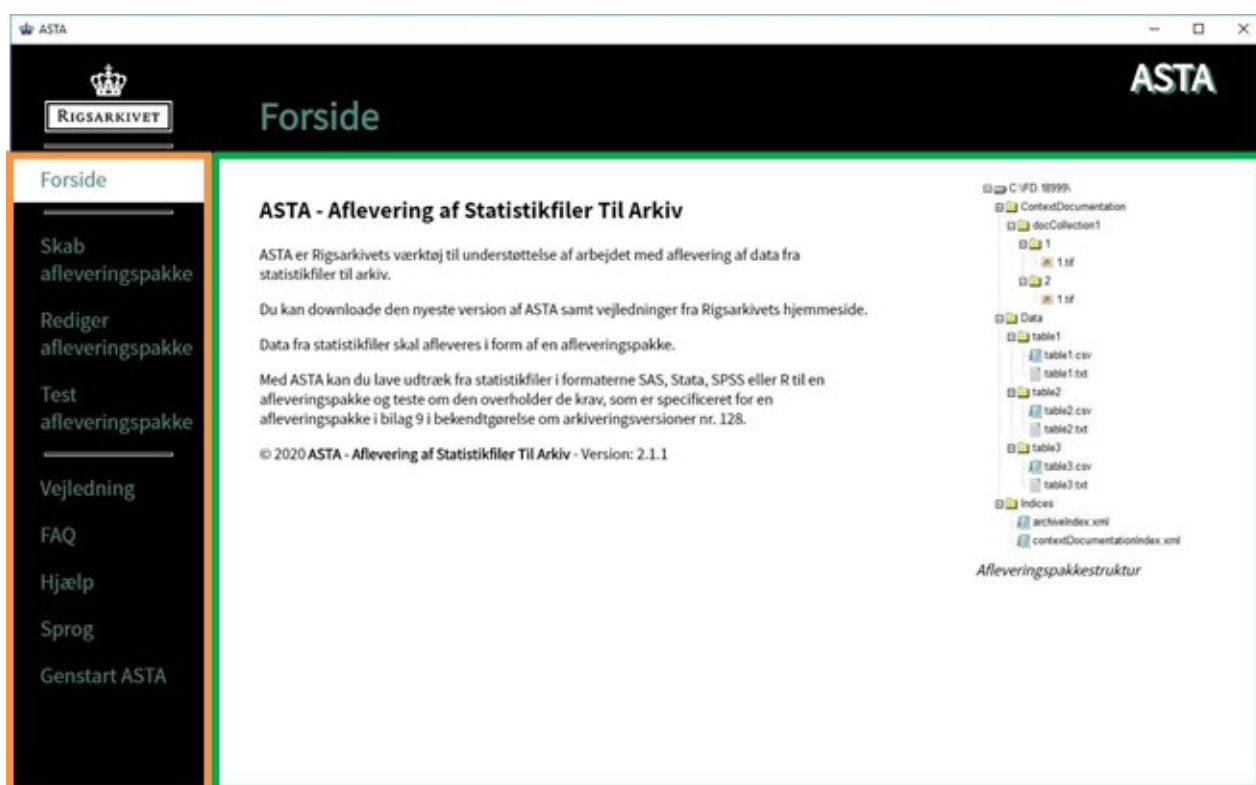




D. Gennemgang af brugergrænsefladen i ASTA

Elementer i hovedmenuen (Boks omgivet af orange streg i figur 1)

På venstre side i ASTA under Rigsarkivets logo findes ASTA's hovedmenu. Den første del af menuen er en forside, der introducerer ASTA og viser, hvordan struktur og indhold af en afleveringspakke ser ud.



Figur 1: Forsiden i ASTA

Orange: Hovedmenuboks med 7 forskellige arbejdsniveauer.

Grøn: Detaljeret underniveau af det valgte hovedmenuniveau.

Efter forsiden er der følgende tre hovedfunktioner i programmet.

- Skab afleveringspakke - Ved at klikke på dette menupunkt kan du straks starte processen med at skabe en afleveringspakke, som består af 5 trin:
 1. Oprette afleveringspakkens mappestruktur
 2. Udtrække data og metadata fra statistikfilen/-erne
 3. Indtaste informationer om statistikfilen/-erne
 4. Placere indeksfilerne
 5. Placere kontekstdokumenterne





Rigsarkivet

- Rediger afleveringspakke – Under dette menupunkt kan du uploade eventuelt nye eller manglende indeksfiler og kontekstdokumenter til en eksisterende afleveringspakke.
- Test afleveringspakke – Dette menupunkt skal du anvende til test af afleveringspakken, når afleveringspakken er skabt. Dette er et vigtigt trin før aflevering til arkiv, fordi funktionen validerer, om afleveringspakkens struktur og indhold overholder de krav, der er specificeret i "Bilag 9 Afleveringspakke for visse typer af forskningsdata" til bekendtgørelse om arkiveringsversioner. Når Rigsarkivet modtager afleveringspakken, anvender Rigsarkivet samme værktøj til test, og hvis kravene ikke er overholdt, bliver arkivskaber bedt om at justere afleveringspakken og indsende den igen. Derfor anbefales det, at køre testen, for at rette eventuelle fejl og mangler vist i testloggen. Dette skal gøres, indtil testen kører uden fejl.

Den tredje del af hovedmenuen indeholder supportinformation om ASTA og anden relevant information til aflevering af data fra statistikfiler til Rigsarkivet.

E. Skab afleveringspakke

Forberedelse før anvendelse af funktionen Skab afleveringspakke

Jævnfør tjeklisten i punkt D bør du bl.a. have følgende klar inden du påbegynder skabelsen af afleveringspakken med ASTA:

- Afleveringspakkens løbenummer, som blev oplyst i arkivets afleveringsbestemmelse.
- En ny mappe med kopier af al materiale, som skal indgå i afleveringspakken inkl. de kopier af de statistikfiler, som skal afleveres til arkivet
- Du skal sikre dig at statistikfilerne overholder krav i bilag 9, jf. 3-8 i afsnit 1 punkt D.
- Godkendte versioner af indeksfilerne `archiveIndex.xml` og `contextDocumentationIndex.xml`.
- Alle kontekstdokumenter registreret i filen `contextDocumentationIndex.xml` konverteret til et af arkivformaterne TIFF, JPEG-2000, MP3-, WAVE- eller MPEG.
- En beskrivelse af datasættet.
- Oplysninger om nøglevariablen i filen (hvis den findes).
- I tilfælde af at flere datafiler i afleveringspakken kan sammenflettes, skal du have oplysninger parat om hvilke datasæt, der kan sammenflettes samt over, hvilke variable denne sammenfletning sker.

Mappestruktur

Når du klikker på menupunktet 'Skab afleveringspakke' åbner siden vist i figur 2, hvor du kan skabe afleveringspakken mappestruktur.

1. Indtast løbenummeret på afleveringspakken i overensstemmelse med det løbenummer, der er tildelt af arkivet og som fremgår af din afleveringsbestemmelse. Hvis det tildelte afleveringspakkenavn er FD.12345, skal du skrive 12345 i indtastningsfeltet.





Rigsarkivet

2. Vælg destinationen, hvor afleveringspakken skal gemmes ved at klikke på knappen 'Browse' og finde den rigtige mappe.
3. Klik på knappen 'Skab mappestruktur' for at oprette afleveringsmappestrukturen. Bemærk, at du i bunden af den side, der fremkommer ved klik på denne knap, kan klikke på linket til den mappestruktur, der blev genereret (se figur 3).

ASTA

RIGSARKIVET

Forside

Skab afleveringspakke

Rediger afleveringspakke

Test afleveringspakke

Vejledning

FAQ

Hjælp

Sprog

Genstart ASTA

Skab afleveringspakke

Mappestruktur

Du kan her skabe en tom mappestruktur til afleveringspakken

Indtast løbenummer for afleveringspakken ①

12345

Vælg destination for afleveringspakken ②

C:\Users\lokalAdmin\Desktop\12345 Browse

Skab mappestruktur ③

Checkliste

Du bør have følgende klar, før du skaber en afleveringspakke:

1. **Løbenummer for afleveringspakken** - løbenummeret er udleveret af arkivet og fremgår af afleveringsbestemmelsen
2. **Statistikfilerne** der skal indgå i afleveringspakken og afleveres til arkivet - husk at kvalitetssikre statistikfilerne før udtræk
Advarsel! Anvend altid en KOPI af din originale statistikfil!
3. **Indeksfilerne** archiveIndex.xml og contextDocumentationIndex.xml
4. **Kontekstdokumenterne** der skal indgå i afleveringspakken. Kontekstdokumenterne skal konverteres til TIFF, JPEG-2000, MP3, WAVE eller MPEG inden de afleveres til arkivet
5. **Beskrivelse af statistikfilernes indhold**
6. **Afklar nøglevariable** - hvilke unikke variable i statistikfilerne skal angives som nøglevariable?
7. **Afklar referencer** - over hvilke variable dannes referencer mellem hvilke statistikfiler i afleveringspakken?

Se hvordan du kvalitetssikrer statistikfilen og gør øvrige dokumenter klar for udtræk med ASTA under menupunktet [Vejledning](#).

Figur 2: Siden "Mappestruktur" under "Skab afleveringspakke" i ASTA

Udtræk data fra statistikfilen

Når du klikker på knappen 'Skab mappestruktur' (se forrige afsnit) fremkommer siden med funktion til at udtrække data og metadata fra statistikfilen vist i figur 3.

Udtræk af data og metadata fra statistikfilen består af tre trin.

I TRIN 1 i figur 3 vælger du den statistikfil, som udtrækket skal laves fra:

1. Vælg den kopi af din statistikfil du vil lave udtræk fra med 'Browse' knappen
2. Klik på 'Næste' for at gå til TRIN 2

TRIN 2: Kør eksport script





Rigsarkivet

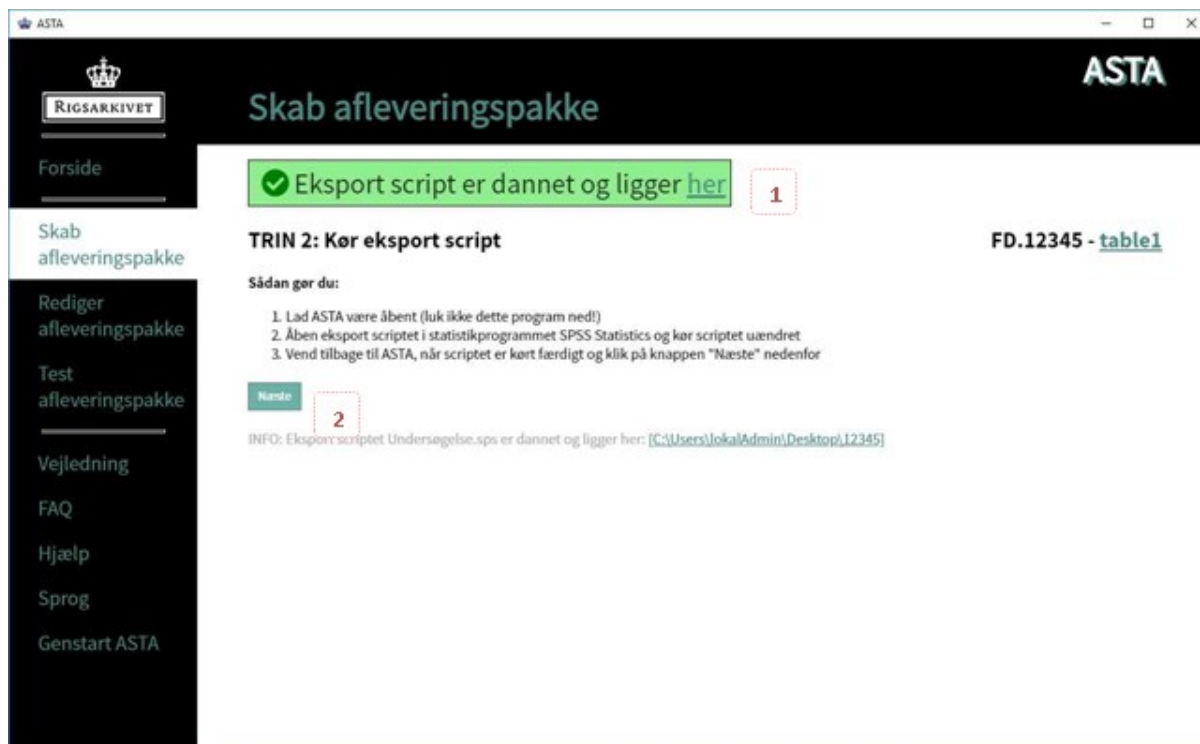
Data og metadata udtrækkes fra statistikfilen til en datafil (fx table1.csv) og en metadatafil (fx table1.txt), der overholder krav til struktur i bilag 9, ved at du kører en eksport syntaks fra et script, som ASTA automatisk har dannet på baggrund af den valgte statistikfil. I trin 2 i figur 4 skal du køre dette eksport script.

Sådan gør du:

1. Luk ikke programmet ASTA mens du kører eksport scriptet, som du finder ved at klikke på 'her' i den grønne boks med teksten "Eksport script er dannet og ligger her". Ved klik på 'her' åbnes en stifinder med præcis placering for, hvor eksport scriptet er placeret. Eksportskriptet er navngivet med statistikfilens navn samt en script ekstension for henholdsvis SPSS (.sps), Stata (.do), SAS (.sas) og R/RStudio (.rds, .RDS, .RData). Bemærk, at der i bunden af skærbilledet er angivet scriptets præcise navn (generationsundersøgelsen.sps). Dobbeltklik på scriptfilen så den åbnes op i det tilsvarende statistikprogram, som du selv skal have installeret på computeren. Kør syntaksen i eksport scriptet.
2. Når scriptet er kørt færdigt, skal du vende tilbage til ASTA og klikke på knappen "Næste" for at gå til TRIN 3.

Figur 3: Siden "Udtræk data fra statistikfilen" under "Skab afleveringspakke" i ASTA





Figur 4: Siden "TRIN 2: Kør eksport script" under "Skab afleveringspakke" i ASTA

TRIN 3: Indtast informationer om statistikfilen

Figur 5: Siden "TRIN 3: Indtast informationer om statistikfilen" under "Skab afleveringspakke" i ASTA



Rigsarkivet

I trin 3 i figur 5 skal du indtaste informationer om den valgte statistikfil.

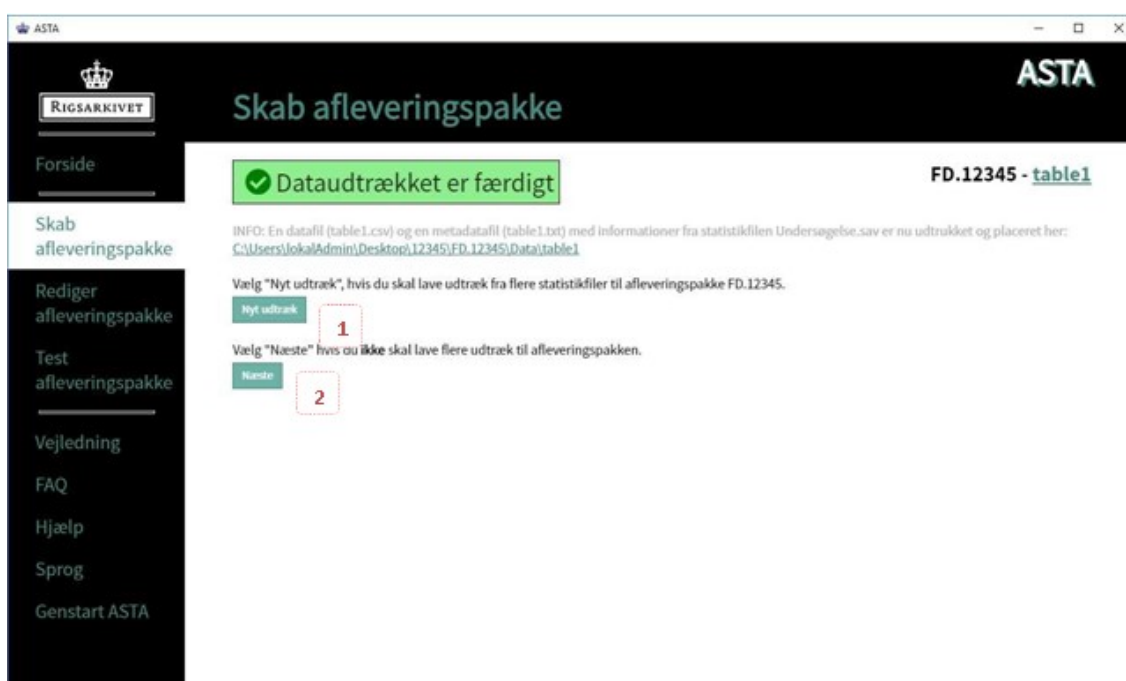
1. Angiv datasættets nav, som de benævnes i brugssammenhæng (obligatorisk)
2. Angiv en beskrivelse af datasættets indhold, maksimalt 4096 tegn (obligatorisk)
Hvis du har en unik nøglevariabel i datafilen, skal du følge trin 3 og 4. Ellers gå til trin 5
3. Klik på den lille pil og vælg nøglevariabelen fra dropdown menuen, der fremkommer
4. Klik på 'Tilføj variabel' for at inkludere den valgte variabel
Hvis din nøglevariabel er sammensat af flere variabel, fortsæt med trin 3 og 4, indtil alle variable er valgt. Den valgte nøglevariabel kan slettes igen ved at klikke på knappen 'Fortryd nøglevariabel'.
5. Klik på 'Næste' for at gå til næste side

Dataudtrækket er færdigt

Ved klik på "Næste" i forrige TRIN 3, fremkommer siden vist i figur 6 med en bekræftelse på, at dataudtrækket er færdigt for den valgte fil. Bemærk at der under INFO vises links til de generede data- og metadatafiler, fx table1.csv og table1.txt.

Du kan nu vælge at lave udtræk fra andre statistikfiler, der skal afleveres i samme afleveringspakke.

1. Klik på knappen "Nyt udtræk", hvis du ønsker at lave udtræk fra flere statistikfiler.
Dette trin vil tage en tilbage til TRIN 1 på siden "Udtræk data fra statistikfilen", hvor du vælger den statistikfil, der skal laves udtræk fra.
2. Klik på knappen "Næste", hvis afleveringspakken ikke skal indeholde udtræk fra flere statistikfiler.



Figur 6: Siden "Dataudtrækket er færdigt" under "Skab afleveringspakke" i ASTA





Rigsarkivet

Kopiering af filer til kontrol af udtræk

Du kan lave en kopi af filer som kan anvendes til at kontrollere din udtræk med ASTA ved at klikke på "Næste". Filer bliver kopierede i mappen ASTA_kontrolfiler_FD.XXXX, som kan afleveres til arkivet sammen med afleveringspakken.

ASTA

RIGSARKIVET

Skab afleveringspakke

Kopiering af filer til kontrol af udtræk **FD.12345**

Du kan her lave en kopi af filer, der kan anvendes til din kontrol af om udtræk med ASTA er tabsfrit samt understøtte arkivets arbejde med test og godkendelse af afleveringspakken. Filerne kopieres til mappen ASTA_kontrolfiler_FD.XXXXX, som placeres ved siden af den skabte afleveringspakke.

Du kan lave en kopi af følgende filer:

- a. Din originale statistikfil
- b. Eksport-script anvendt til udtræk
- c. Eventuel outputfil fra udtræk (kun SPSS)
- d. Filen statistikfilnavn_Exportscriptlog.spv (kun SPSS). Filen indeholder resultatet af sammenligningen (compare) af den valgte originale statistikfil og outputfil fra udtræk i SPSS.

Information! Mappen ASTA_kontrolfiler_FD.XXXXX kan med fordel afleveres til arkivet sammen med afleveringspakken

Information! Kopiering af meget store filer kan tage lang tid!

Vælg filer	Fileres originale placering	Placering af udtræk	Størrelse (Bytes)
<input checked="" type="checkbox"/> Undersøgelse.sav	C:\Users\lokalAdmin\Desktop\12345	table1	123169
<input checked="" type="checkbox"/> Undersøgelse.sps			
<input checked="" type="checkbox"/> Undersøgelse_Exportscriptlog.spv			
<input checked="" type="checkbox"/> Undersøgelse_output.sav			

1

Figur 7: Siden "Kopiering af filer til kontrol af udtræk" under "Skab afleveringspakke" i ASTA



Rigsarkivet

Placering af indeksfiler i afleveringspakken

Du skal nu placere indeksfilerne i afleveringspakken, se figur 8.

Figur 8: Siden "Placer indeksfiler i afleveringspakken" under "Skab afleveringspakke" i ASTA

1. Vælg filen *archiveIndex.xml* ved brug af 'Browse' knappen.
2. Vælg filen *contextDocumentationIndex.xml* ved brug af 'Browse' knappen.
3. Klik på knappen 'Næste' for at uploade de valgte filer.





Rigsarkivet

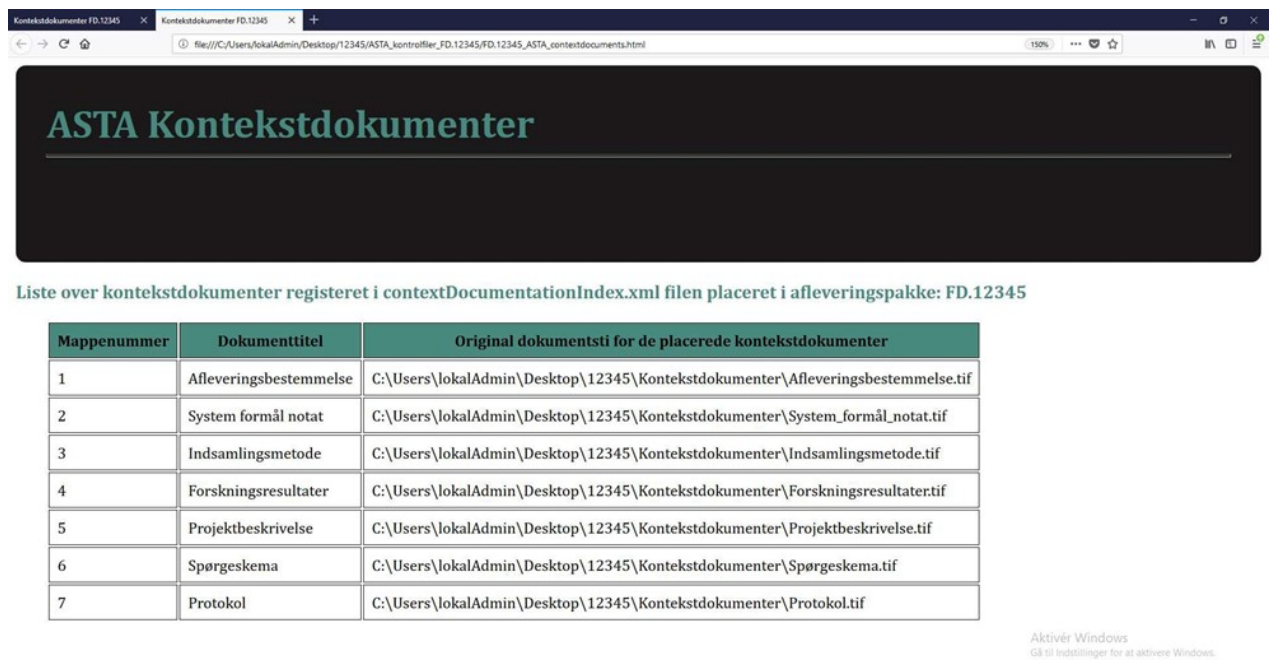
Placering af kontekstdokumenter i afleveringspakke

Du skal nu placere kontekstdokumenterne, der er registreret i contextDocumentationIndex.xml i afleveringspakken, se figur 9.

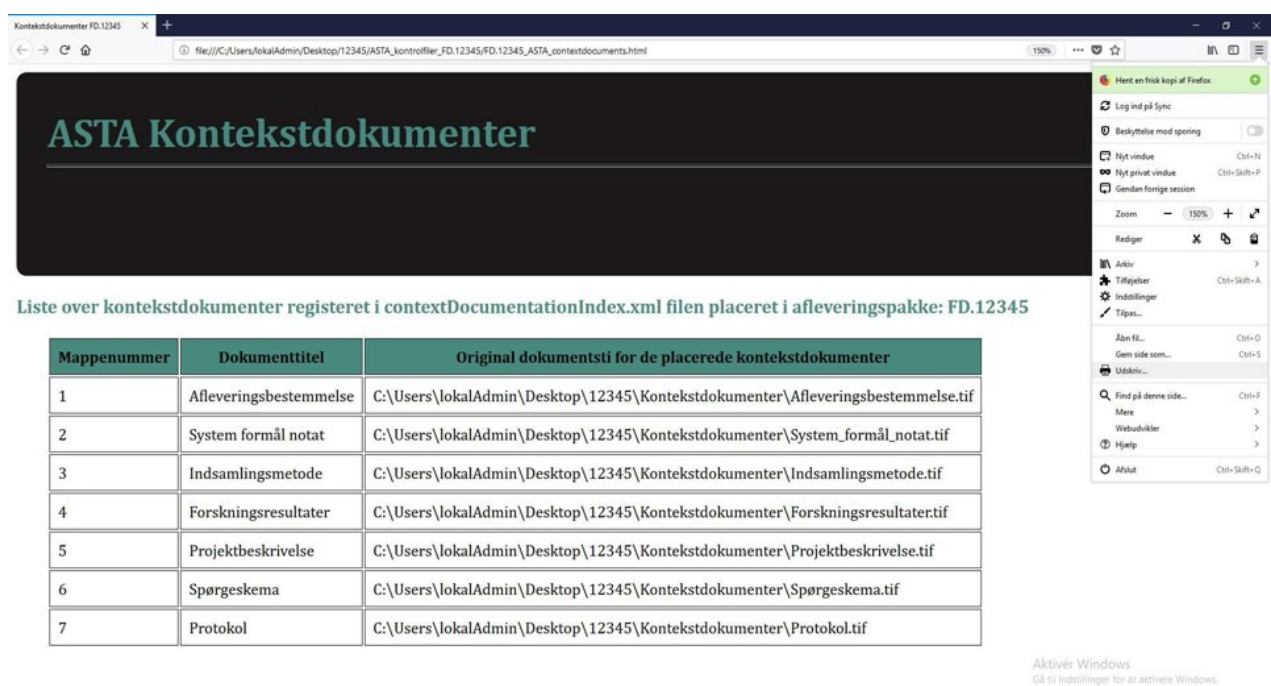
Mappe nummer	Dokumenttitel	Vælg dokument
1	Afleveringsbestemmelse	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\Afleringsbestemmelse.tif Browse
2	System formål notat	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\System_formål_notat.tif Browse
3	Indsamlingemetode	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\Indsamlingemetode.tif Browse
4	Forskningsresultater	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\Forskningsresultater.tif Browse
5	Projektbeskrivelse	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\Projektbeskrivelse.tif Browse
6	Spørgeskema	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\Spørgeskema.tif Browse
7	Protokol	C:\Users\JokulAdmin\Desktop\12345\Kontekstdokumenter\Protokol.tif Browse

Figur 9: Siden "Placer kontekstdokumenter i afleveringspakken" under "Skab afleveringspakke" i ASTA

1. Alle kontekstdokumenter registreret i contextDocumentationIndex.xml vises i en dokumentliste. Klik på knappen 'Browse' ud for hvert dokument og vælg det kontekstdokument, der svarer til dokumenttitlen i dokumentlisten. Hvis du ikke har alle kontekstdokumenter klar, kan du tilføje dem senere under menupunktet 'Rediger afleveringspakke'.
2. Det anbefales, at du udskriver dokumentlisten på knappen 'Udskriv dokumentliste' og kontrollerer, at indholdet af de valgte dokumenter svarer til dokumenttitlen. Når du klikker på knappen "Udskriv dokumentliste" vises dokumentlisten i din standardbrowser, fx Edge (se figur 10). Anvend browserens printfunktion til at udskrive listen (se figur 11).
3. Klik på knappen 'Næste' for at uploade de valgte dokumenter.



Figur 10: "Dokumentliste" som fremkommer ved klik på "Udskriv dokumentliste" i ASTA



Figur 11: Print af "Dokumentliste" via browserens printfunktion



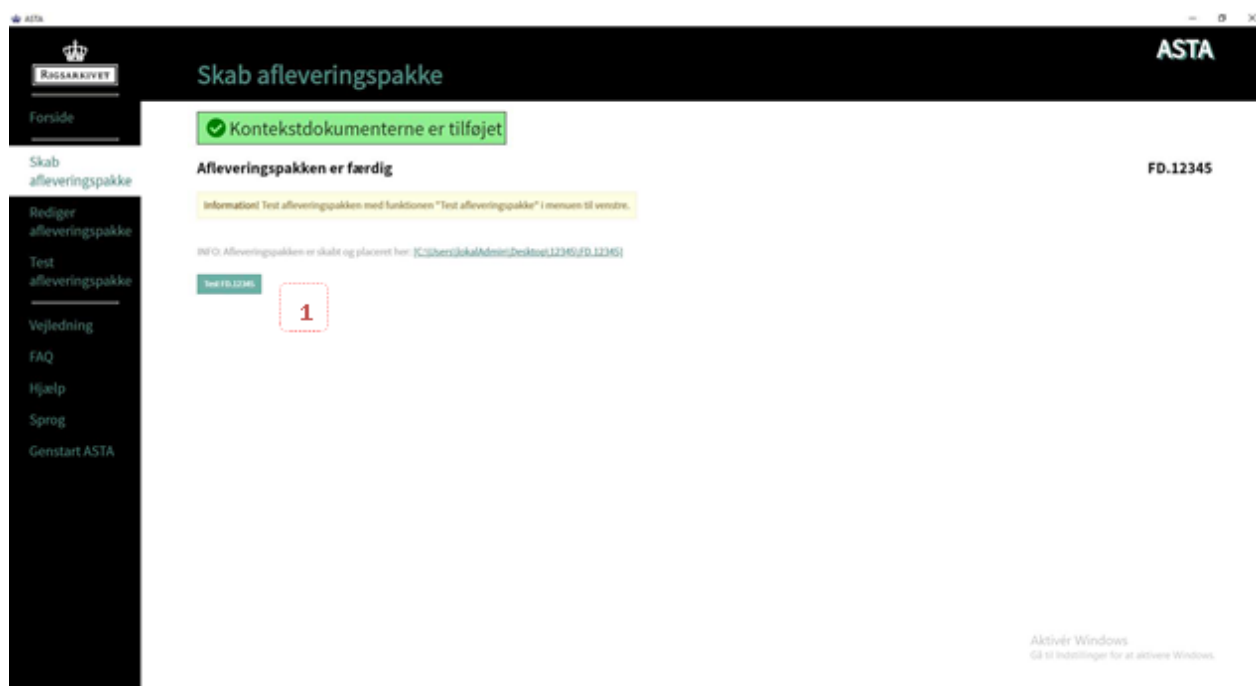


Rigsarkivet

Afleveringspakken er færdig

Afleveringspakken er nu færdig og skal testes.

Klik på knappen '**Test FD.XXXXX**' for en genvej til test af den netop skabte afleveringspakke. Du kan også teste afleveringspakken under menupunktet 'Test afleveringspakke'.



Figur 12: Siden "Afleveringspakken er færdig" under "Skab afleveringspakke" i ASTA

F. Test afleveringspakke

Test af afleveringspakke

Det er dit ansvar at sikre, at afleveringspakken overholder alle krav i bilag 9 i Bekendtgørelse om arkiveringsversioner. Derfor er det vigtigt, at du tester afleveringspakken og retter eventuelle fejl, inden du afleverer afleveringspakken til arkivet.

Du kan teste afleveringspakken under 'Test afleveringspakke', se figur 13.

1. Vælg den afleveringspakke du vil teste (fx FD.12345) ved at klikke på knappen 'Browse'
2. Klik på knappen 'Start test'
3. Testen er nu påbegyndt, og teststatus vises i testloggen, se figur 13.





Rigsarkivet

Teststatus og rettelse af fejl i testlog

1. Gennemgå nøje de genererede fejlmeddelelser markeret med rødt kryds i testloggen (4 i figur 13) og ret altid disse inden aflevering til arkivet. Hints vises som en advarsel med gult udråbstegn. Hints er ikke altid fejl og skal kun rettes, hvis det faktisk er en fejl eller mangel i den pågældende afleveringspakke.
2. Ud over den testlog der vises på skærbilledet genereres også en html-version af testloggen, som også gemmes automatisk på samme placering, som den valgte afleveringspakke efter endt test (se figur 14). Hvis du klikker på linket til denne testlog (5 i figur 13), åbnes den i din standard browser og kan printes ved hjælp af browserens printfunktion (se figur 11). Html-testloggen indeholder også vejledning til, hvordan fejlbeskeder i testloggen forstås og rettes (se figur 15).
3. Når testen er kørt færdig vises en af følgende tre statusmeddelelser:

Afleveringspakken er gennemtestet uden fejl og kan nu afleveres til arkivet

Dette indikerer at afleveringspakken er klar til at blive afleveret til arkivet. BEMÆRK: Da ASTA ikke tester for alle krav i bilag 9 kan arkivet vende tilbage med yderligere fejl eller mangler, der skal rettes.

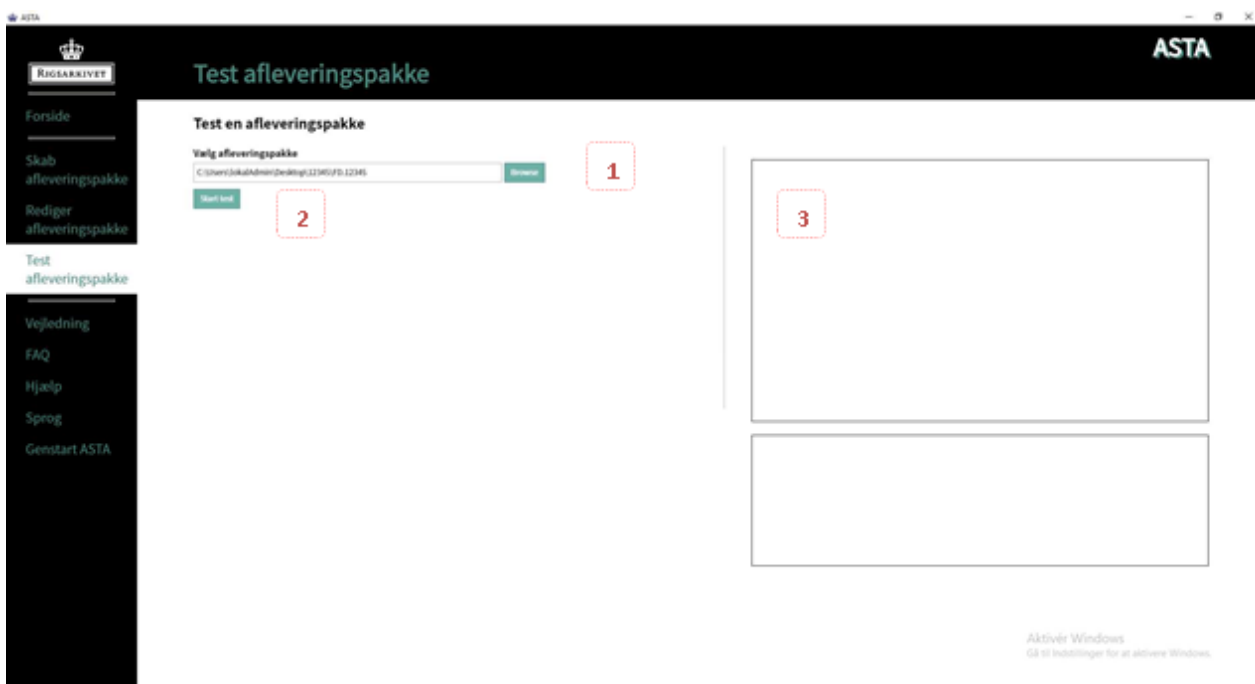
Afleveringspakken er fuldt gennemtestet. Fejl vist i loggen skal rettes før aflevering til arkiv

Dette indikerer, at testen er afsluttet, men der stadig findes fejl, der skal rettes. Når du har rettet de viste fejl i testloggen og kører en ny test fremkommer ikke flere nye fejl.

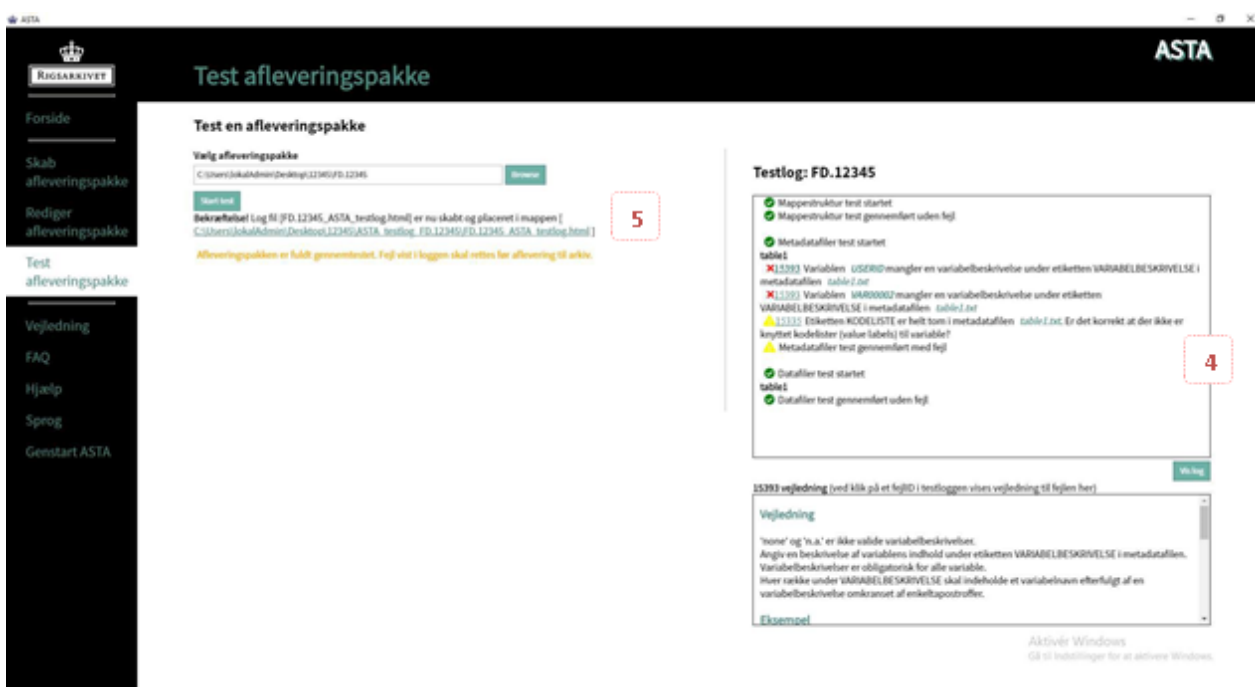
Testen er afbrudt, og afleveringspakken er ikke testet færdig. Fejl vist i loggen skal rettes før testen kan fortsætte

Dette indikerer, at der er fejl, der skal rettes, inden testen kan fortsætte. Når du har rettet de viste fejl i testloggen og kører en ny test, kan der fremkomme flere nye fejl. Ved flere tests vises max 40 af samme type fejl i testloggen, hvilket betyder, at når du har rettet de viste 40 fejl og kører en ny test, kan der forekomme flere af samme type.



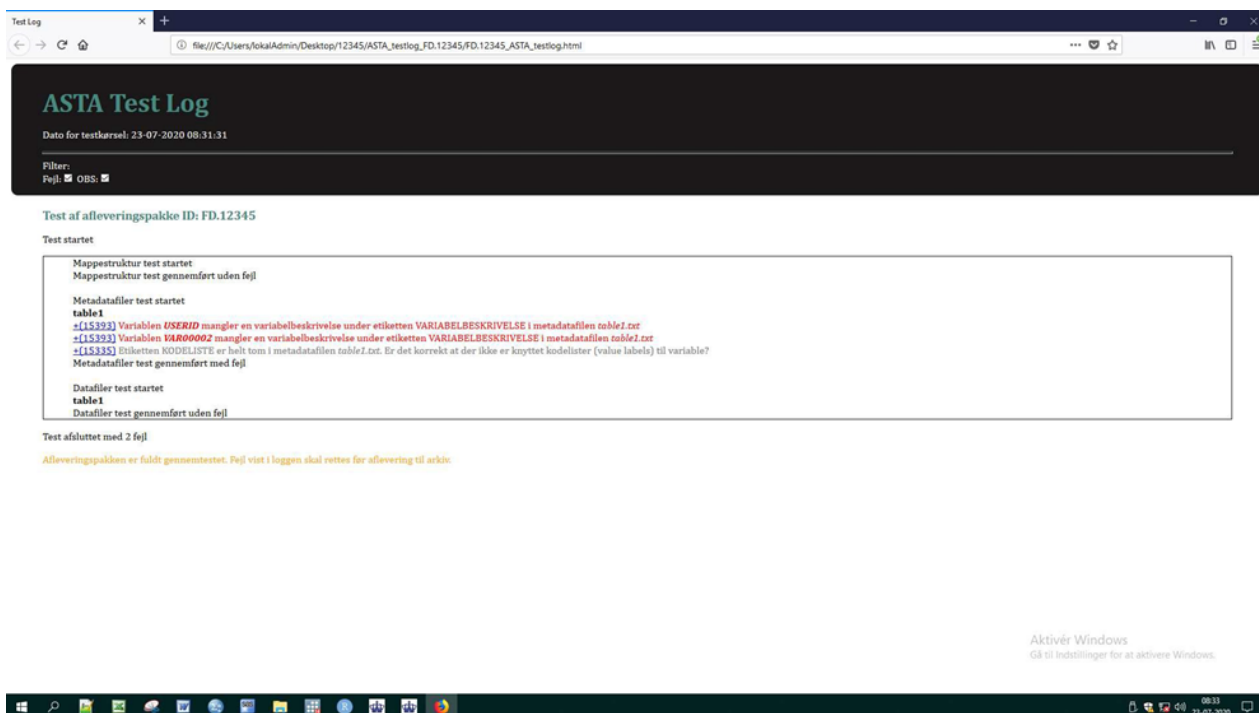


Figur 13: Siden "Test afleveringspakke" i ASTA før testen påbegyndes



Figur 14: Siden "Test afleveringspakke" i ASTA når testen er afsluttet



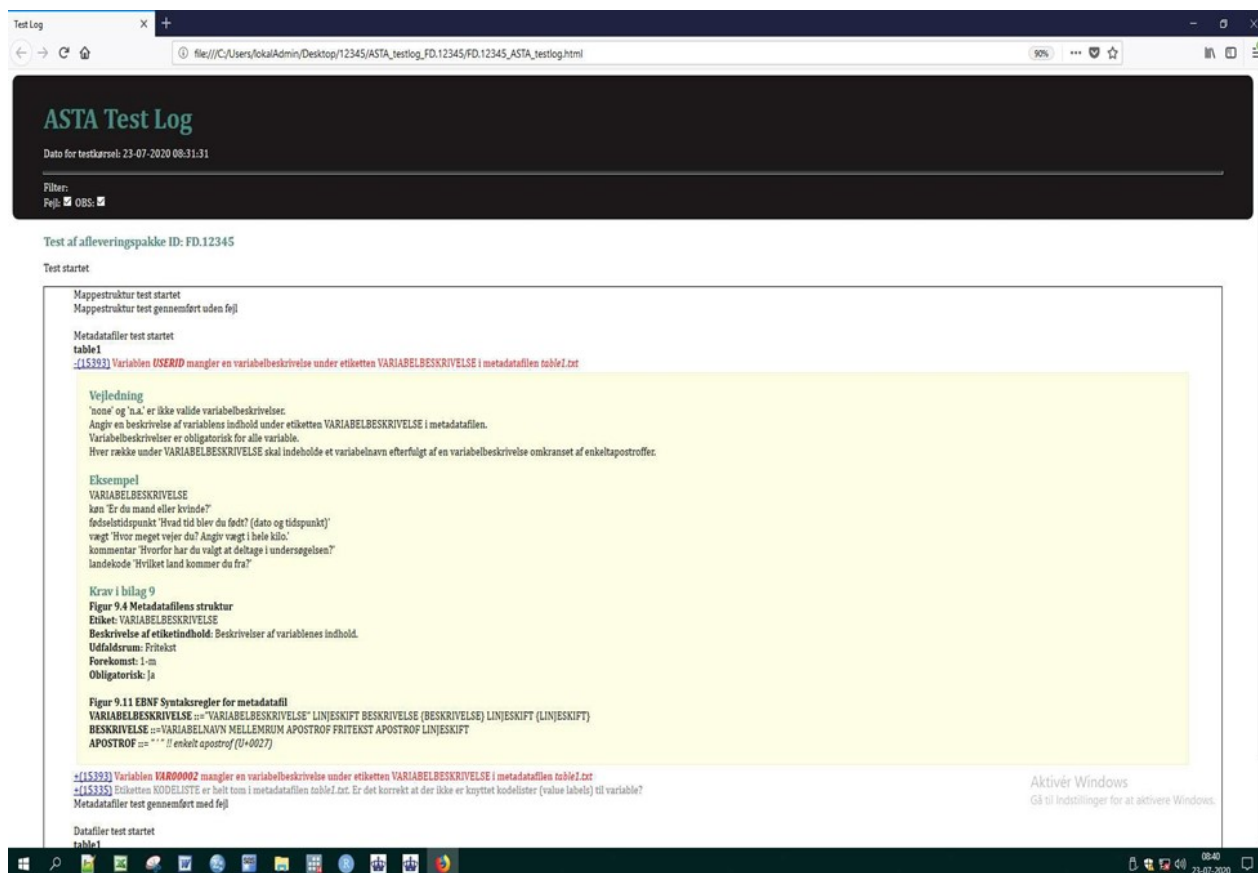


Figur 15: HTML-Versionen af testloggen fra "Test afleveringspakke" i ASTA

HTML-Version af testlog med vejledning

I html-testloggen som fremkommer, når du klikker på linket fra ASTA til denne (se 5 i figur 14) vises et fejl-ID (fx 15393) ud fra hver opdaget fejl. Ved klik på dette fejl-ID fremkommer en vejledningstekst til, hvordan fejlen skal forstås og rettes. Der vises også et eksempel på data og metadata uden fejl samt angivelse af den konkrete ordlyd af kravene, som de står skrevet i bilag 9 i Bekendtgørelse om arkiveringsversioner (se figur 16)





Figur 16: Html-testlog med visning af vejledning ved klik på fejl-ID 15393

G. Kontroller at udtræk fra ASTA er tabsfrit

Det er vigtigt at de data der afleveres til Rigsarkivet er autentiske og svarer til indholdet i de originale statistikfiler. Det er den afleverende myndigheds ansvar at sikre, at det udtræk der laves med ASTA eller andre værktøjer er tabsfrit. Du skal derfor foretage en visuel kontrol af data før og efter udtræk med ASTA, dvs. foretage en sammenligning af data i den originale statistikfil med data i datafilen (fx table1.csv) i afleveringspakken. Du skal også kontrollere om metadata i metadatafilen (fx table1.txt) i afleveringspakken er udtrukket korrekt og tabsfrit fra den originale statistikfil.

Kopiering af kontrolfiler

Under udtræk med ASTA får du mulighed for at kopiere relevante filer der kan understøtte kvalitetskontrollen af udtrækket til mappen ASTA_kontrolfiler_FD.XXXXX. Herunder den valgte (originale) statistikfil som udtrækket laves fra, samt eventuelle outputfiler og logs, som ASTA skaber under udtræk.





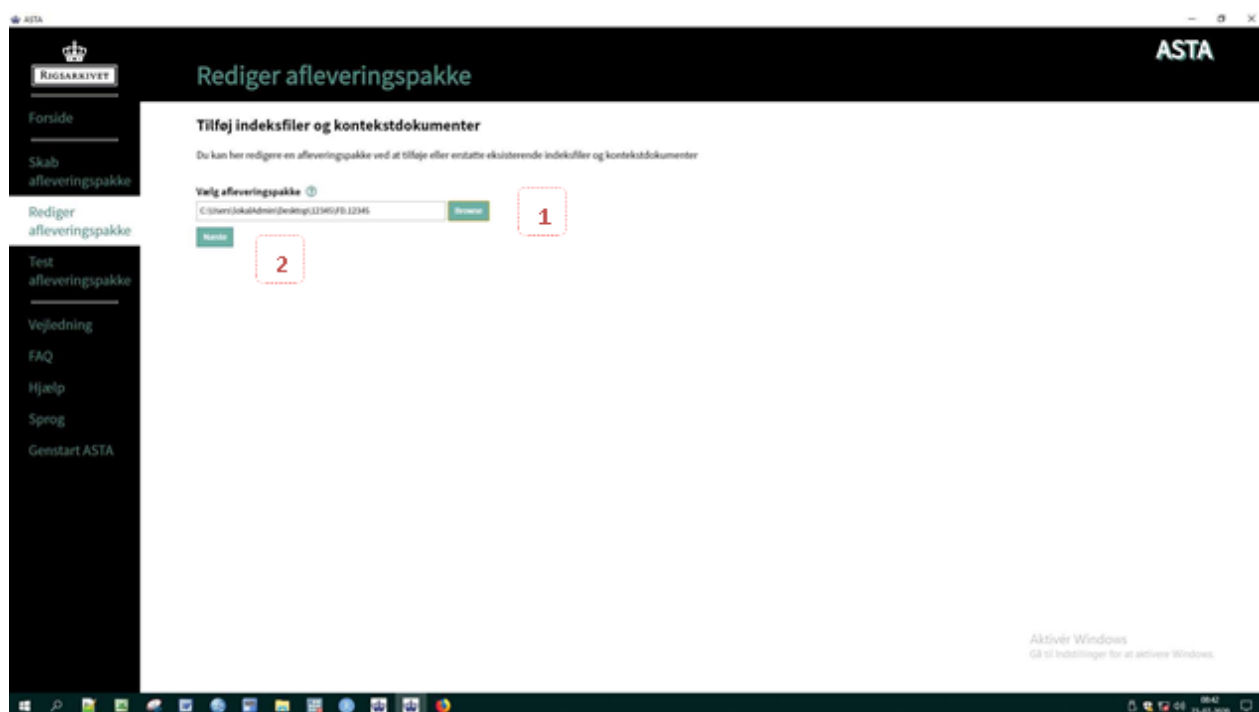
Rigsarkivet

Kontrol af placering af kontekstdokumenter

Det anbefales, at du under udtræk med ASTA udskriver dokumentlisten over tilføjede kontekstdokumenter (se figur 11) og anvender denne til kontrol af, om alle kontekstdokumenter er placeret i de korrekte dokumentmapper.

H. Rediger afleveringspakke

Under menupunktet 'Rediger afleveringspakke' kan du redigere afleveringspakken med hensyn til placering af indeksfiler og kontekstdokumenter, se figur 17.



Figur 17: Siden "Tilføj indeksfiler og kontekstdokumenter" under "Rediger afleveringspakke" i ASTA

1. Vælg den afleveringspakke (fx FD.12345), der skal redigeres, ved brug af 'Browse' knappen
2. Klik på knappen 'Næste' for at gå til næste side

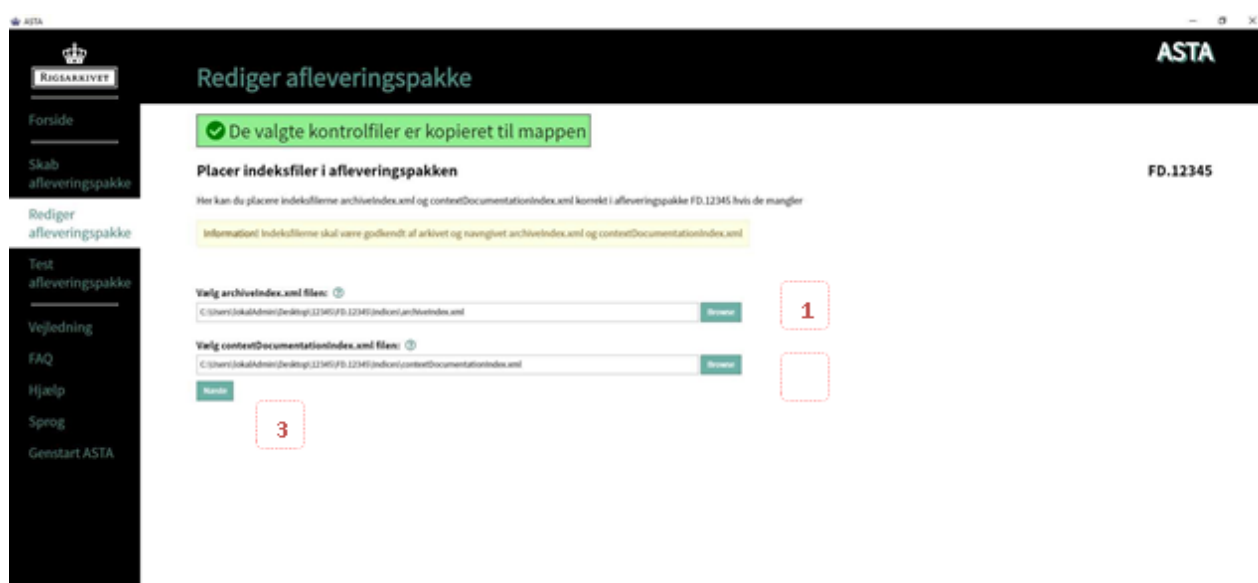
Placér indeksfiler i afleveringspakken

Ved klik på 'Næste' på forrige side, fremkommer siden vist i figur 18. Du kan her placere indeksfilerne *archiveIndex.xml* og *contextDocumentationIndex.xml* i afleveringspakken, hvis de helt mangler, eller erstatte de eksisterende med nyere opdaterede versioner (1 og 2).



Rigsarkivet

1. Vælg filen *archiveIndex.xml* ved brug af 'Browse' knappen, hvis den mangler, eller du ønsker at erstatte den eksisterende med en ny opdateret version.
2. Vælg filen *contextDocumentationIndex.xml* ved brug af 'Browse' knappen, hvis den mangler, eller du ønsker at erstatte den eksisterende med en ny opdateret version.
3. Klik på knappen 'Næste' for at uploade de valgte filer, eller hvis du fortsat skal anvende de allerede eksisterende indeksfiler.



Figur 18: Siden "Placer indeksfiler i afleveringspakken" under "Rediger afleveringspakke" i ASTA



Placer kontekstdokumenter i afleveringspakken

Ved klik på 'Næste' på forrige side fremkommer siden vist i figur 19.

Du kan her placere kontekstdokumenter i afleveringspakken eller erstatte eksisterende kontekstdokumenter med nye.

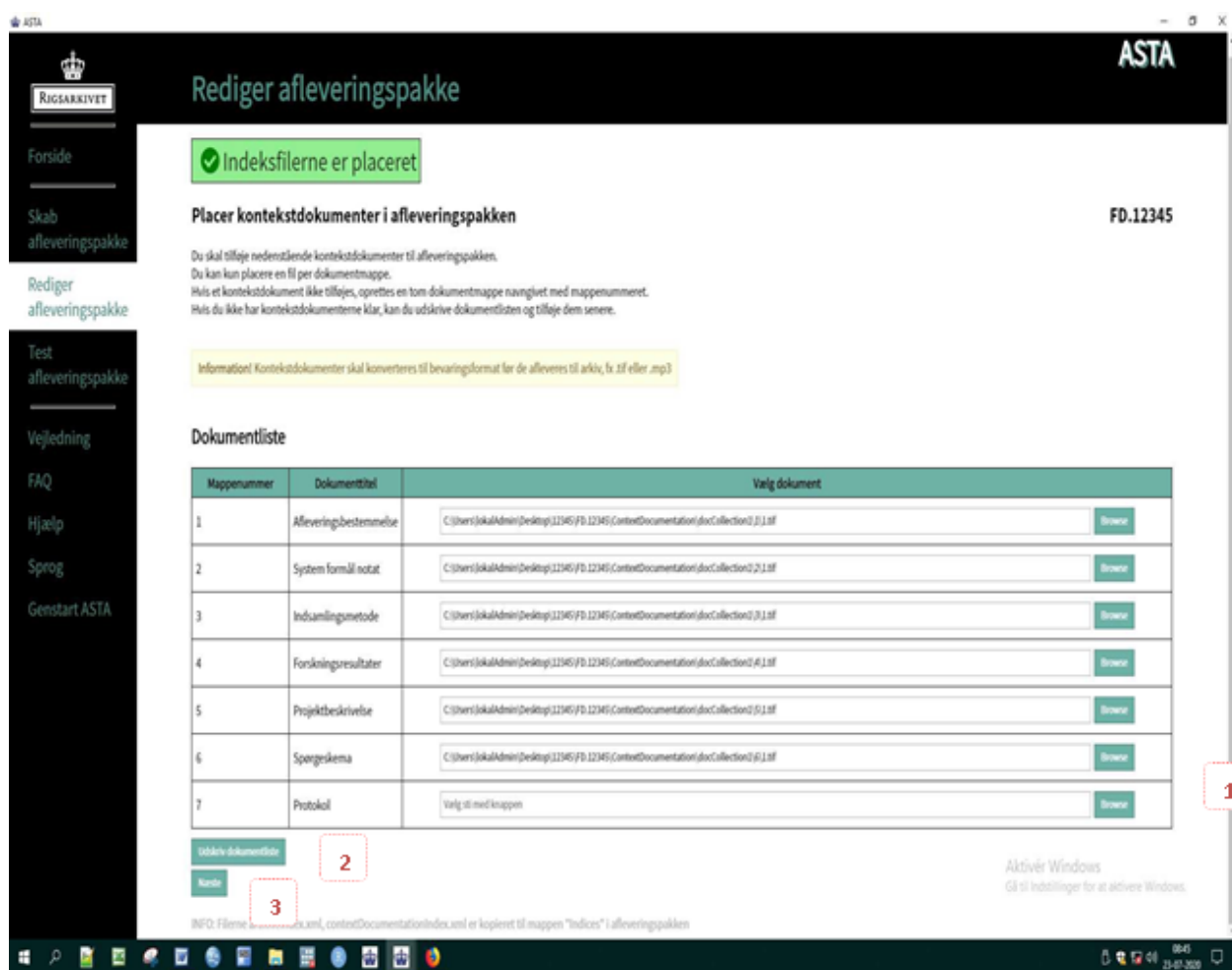
1. Alle kontekstdokumenter registreret i *contextDocumentationIndex.xml* vises i en dokumentliste. Klik på knappen 'Browse' ud for hvert dokument og vælg det kontekstdokument, der svarer til dokumenttitlen i dokumentlisten. Hvis der findes et kontekstdokument i mappen allerede, vises stien til denne. Klik på 'Browse' knappen hvis du ønsker at erstatte det eksisterende kontekstdokument i mappen med et nyt.
2. Når du er færdig med at placere alle kontekstdokumenterne, anbefales det, at du udskriver dokumentlisten på knappen 'Udskriv dokumentliste' og kontrollerer, at indholdet af de valgte dokumenter svarer til dokumenttitlen. Når du klikker på knappen 'Udskriv dokumentliste' vises dokumentlisten i din standardbrowser, fx Edge (se figur 10). Anvend browserens printfunktion til at udskrive listen (se figur 11). Dokumentlisten gemmes også automatisk i en



Rigsarkivet

html-version samme sted, hvor afleveringspakken er placeret, når du klikker på knappen 'Udskriv dokumentliste'.

3. Klik på knappen 'Næste' for at uploade de valgte dokumenter. Skærbilledet i figur 9 med bekræftelse, af at afleveringspakken er færdig, fremkommer. Klik på knappen 'Test FD.XXXXX' for at teste afleveringspakken (som vist i figur 12).



Mappe nummer	Dokumenttitel	Vælg dokument
1	Afleveringsbestemmelse	C:\Users\jokaldm\Desktp\12345\FD.12345\ContextDocumentation\docCollection2\1.1.tif Størrelse
2	System formål notat	C:\Users\jokaldm\Desktp\12345\FD.12345\ContextDocumentation\docCollection2\2.1.tif Størrelse
3	Indsamlingsmetode	C:\Users\jokaldm\Desktp\12345\FD.12345\ContextDocumentation\docCollection2\3.1.tif Størrelse
4	Forskningsresultater	C:\Users\jokaldm\Desktp\12345\FD.12345\ContextDocumentation\docCollection2\4.1.tif Størrelse
5	Projektbeskrivelse	C:\Users\jokaldm\Desktp\12345\FD.12345\ContextDocumentation\docCollection2\5.1.tif Størrelse
6	Spørgeskema	C:\Users\jokaldm\Desktp\12345\FD.12345\ContextDocumentation\docCollection2\6.1.tif Størrelse
7	Protokol	Vælg til med knappen Størrelse

Figur 19: Siden "Placer kontekstdokumenter i afleveringspakken" under "Rediger afleveringspakke" i ASTA

I. ASTA support i Rigsarkivet

Bilag 1

Dette bilag giver en grundig gennemgang af, hvordan du i de forskellige statistikprogrammer SPSS, SAS Stata og R/RStudio sikrer dig, at tjeklistens punkter 1-8 er opfyldt.



**A. SPSS – forberedelse af statistikfilen før udtræk**

SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
GENERELT			
1. Tjek/installer statistikprogram	Tjek at du har statistikprogrammet SPSS installeret på den PC, du anvender ASTA på.		Du kan ikke udtrække data fra en SPSS-fil med ASTA, hvis statistikprogrammet SPSS ikke er installeret på PC'en
2. Ny mappe med kopi af materiale	<p>Opret en ny mappe på computeren med en kopi af statistikfilerne, der skal afleveres, kontekstdokumenterne, som fremgår af afleveringsbestemmelsen, og de to indeksfiler.</p> <p>Navngiv evt. mappen med afleveringspakkens løbenummer, der fremgår af arkivets afleveringsbestemmelse.</p>	Da forberedelsen af afleveringen involverer ændringer i variablene, justering af formater og konvertering af filer, er det vigtigt at bevare integriteten af de originale filer ved at separere dem fra afleveringsmaterialet.	De originale data kan blive overskrevet eller slettet.
SPSS DATAFILER			





SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
3.Datafiler	Tjek at SPSS filerne har ekstension .sav	Data skal være i et dataformat, som ASTA kan lave udtræk fra. For SPSS-filer skal ekstension .sav anvendes.	ASTA kan kun udtrække data fra formaterne SPSS, SAS, Stata og R
	Tjek at datafiler er indkodet til UTF-8	<p>Alle datafiler skal være kodet til UTF-8.</p> <p>Tip: Nederst i venstre side af SPSS-vinduet kan man se "Unicode: ON" eller "Unicode: OFF". Juster indkodningen så Unicode (som i SPSS betyder UTF-8) er "ON".</p> <p>Ydermere skal du undersøge, om dit datasæt viser alle tegnene korrekt specielt æ, ø og å.</p> <p>Se vejledning om UTF-8 på Rigsarkivets hjemmeside.</p>	<p>Hvis data ikke er indkodet til UTF-8, vil nogle tegn, såsom æ, ø og å ikke blive vist korrekt.</p> <p>Efter aflevering til arkiv vil arkivet også teste afleveringen. Hvis data ikke er indkodet som UTF-8, kan det medføre andre fejl i afleveringen, som kun kan rettes ved at lave helt nye udtræk fra de originale statistikfiler.</p> <p>Så spar tid ved at kontrollere det inden udtræk.</p>





SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
4.Variable i datafiler	Tjek at alle variabelnavne overholder regler for navngivning af variable.	Variabelnavne må ikke være længere end 128 tegn, starte med et tal eller inkludere blanktegn (mellemrum) eller andre tegn end bogstaver og tal. Brug "Variable View"-fanen for at se alle dine variabelnavne, labels og typer, bredde og decimaler. Juster enhver variabel, der ikke følger Rigsarkivets regler om aflevering.	ASTA vil vise fejl under testningen af afleveringspakken.
	Tjek at alle variable har variable labels	Alle variable skal have variabel labels, der beskriver deres indhold. Brug "Variable View"-fanen og kolonnen "Label" for at sikre, at alle variable har variabel labels	ASTA vil vise fejl under testningen af afleveringspakken.





SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
5.Variable type/format	Tjek at alle variable har angivet en type/format, som er accepteret af Rigsarkivet og ikke er angivet med "default formats"	Alle variable skal have angivet type og format, som er accepteret af Rigsarkivet og ikke "default formats". Brug "Variable View"-fanen og kolonnen "Type" for at sikre dig, at alle variable har den accepterede type/format.	ASTA vil vise fejl under testningen af afleveringspakken.





SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
6.Value labels	Tjek at alle value labels er unikke og specificeret.	<p>Alle værdier i value labels skal være unikke og specificeret.</p> <p>I tilfælde af at nogle værdier ikke er specificeret (fx i en Likert skala hvor kun den første og sidste værdi har specificeret labels), skal dette beskrives i variabelens label fx med teksten "Skala anvendt. Ikke alle koder har kodebeskrivelse."</p> <p>Variable labels ændres i "Variable View"-fanen i kolonnen "Label" i SPSS.</p>	Efter aflevering til arkiv vil arkivet teste, om alle value labels findes.





SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
7.Missing values	Kontrollér at det kun er numeriske og kategoriske variable, der har manglende koder i "Missing" kolonnen.	Manglende værdier må kun anvendes på numeriske og kategoriske variable. Kontakt Rigsarkivets datamanager for forskningsdata, hvis du anvender koder for manglende værdier i variable med typerne tekst, tidspunkt eller dato.	ASTA vil vise fejl under testningen af afleveringspakken.
	Alle tildelte koder for manglende værdier skal også tildeles en værdi i value labels knyttet til den variabel, hvor koden anvendes	Forståelse af den tildelte kode for en missing value er vigtig for at forstå data. Derfor skal alle anvendte koder for manglende værdier (missing values) også tildeles en værdi i value labels knyttet til den variabel, hvor koden anvendes. Fx 9 = Uoplyst 10 = Irrelevant 11 = Deltager ikke	ASTA vil vise fejl under testningen af afleveringspakken.





SPSS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
8.Referencer (ved aflevering af flere datasæt)	Tjek at eventuelle flettevariable har samme type/format og længde	<p>Flettevariable angives i afleveringspakken for at beskrive mulig sammenkobling mellem flere datasæt.</p> <p>Når der findes flere datasæt i en afleveringspakke og referencer mellem disse (en fletningsnøgle), skal der være fuldstændig overensstemmelse mellem flettevariablenes type og længde.</p> <p>Tjek at formatet og typen af disse flettevariable er identiske i "Variabel view"- fanen for hver af de afleverede filer.</p>	<p>Manglende overensstemmelse mellem type og længde i flettevariable kan påvirke muligheden for at sammenkoble de afleverede datasæt.</p> <p>Efter aflevering til arkiv vil arkivet kontrollere, om denne regel er overholdt. Spar tid ved at gøre dette på forhånd.</p>



**B. Stata – forberedelse af statistikfilen før udtræk**

Stata			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
GENERELT			
1. Tjek/installer statistikprogram	Tjek at du har statistikprogrammet Stata installeret på den PC, du anvender ASTA på.		Du kan ikke udtrække data fra en Stata-fil med ASTA, hvis statistikprogrammet Stata ikke er installeret på PC'en
2. Ny mappe med kopi af materiale	<p>Opret en ny mappe på computeren med en kopi af statistikfilerne der skal afleveres, kontekstdokumenterne, som fremgår af afleveringsbestemmelsen, og de to indeksfiler.</p> <p>Navngiv evt. mappen med afleveringspakkens løbenummer, der fremgår af arkivets afleveringsbestemmelse.</p>	Da forberedelsen af afleveringen involverer ændringer i variablene, justering af formater og konvertering af filer, er det vigtigt at bevare integriteten af de originale filer ved at separere dem fra afleveringsmaterialet.	De originale data kan blive overskrevet eller slettet
STATA DATAFILER			
3. Datafiler	Tjek at STATA-filerne har ekstension .dta	Data skal være i et dataformat, som ASTA kan lave udtræk fra. For STATA filer skal ekstension <i>.dta</i> anvendes.	ASTA kan kun udtrække data fra formaterne SPSS, SAS , Stata og R.





Stata			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
	Tjek at datafiler er kodet til UTF-8	<p>Alle datafiler skal være kodet til UTF-8.</p> <p>Ydermere skal du undersøge, om dit datasæt viser alle tegnene korrekt specielt æ, ø og å.</p> <p>Se vejledning om UTF-8 på Rigsarkivets hjemmeside.</p>	<p>Hvis data ikke er indkodet til UTF-8, vil nogle tegn, såsom æ, ø og å ikke blive vist korrekt.</p> <p>Efter aflevering til arkiv vil arkivet også teste afleveringen. Hvis data ikke er indkodet som UTF-8 kan det medføre andre fejl i afleveringen, som kun kan rettes ved at lave helt nye udtræk fra de originale statistikfiler. Så spar tid ved at kontrollere det inden udtræk.</p>
4.Variable i datafiler	Tjek at alle variabelnavne overholder regler for navngivning af variable	<p>Variabelnavne må ikke være længere end 128 tegn, starte med et tal eller inkludere blanktegn (mellemrum) eller andre tegn end bogstaver og tal.</p> <p>Brug "Property" vinduet (nederst til højre i hovedvinduet i Stata) for at se dine variable, deres navne, labels og formater. Juster enhver variabel, der ikke overholder regler for</p>	ASTA vil vise fejl under testningen af afleveringspakken





Stata			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		navngivning af variable.	
	Tjek at alle variable har variable labels	Alle variable skal have variabel labels, der beskriver deres indhold. Se på variablene for at sikre dig at alle variable har variabel labels (beskrivelser). Brug "Variables Manager" under menuen "Data" for at ændre eller tilføje variabel labels.	ASTA vil vise fejl under testningen af afleveringspakken
5.Variable type/format	Tjek at alle variable har angivet en type/format som er accepteret af Rigsarkivet og ikke er angivet med "default formats"	Alle variable skal have angivet type og format som er accepteret af Rigsarkivet og ikke "default formats". Brug "Properties" vinduet for at sikre dig, at alle variable har de accepterede typer/formater. Hvis ikke, brug "Variables Manager" under menuen "Data" for at ændre dine variabel labels.	ASTA vil vise fejl under testningen af afleveringspakken.
6.Value labels	Tjek at alle value labels er unikke og specificeret	Alle værdier i value labels skal være unikke og specificeret.	Efter aflevering til arkiv vil arkivet teste om alle value labels findes.





Stata			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>I tilfælde af at nogle værdier ikke er specificeret (fx i en Likert skala hvor kun den første og sidste værdi har specificeret labels), skal dette beskrives i variabelens variable label fx med teksten "Skala anvendt. Ikke alle koder har kodebeskrivelse."</p> <p>Variable labels ændres i "Variables Manager" under menuen 'Data' i Stata.</p>	
7.Missing values	Kontrollér at det kun er numeriske og kategoriske variable, der anvender koder for missing values.	<p>Manglende værdier må kun anvendes på numeriske og kategoriske variable.</p> <p>Kontrollér at det kun er numeriske og kategoriske variable, der anvender koder for missing values i "value labels" kolonnen i "Variables Manager" vinduet under menuen "Data" i Stata.</p> <p>Koder for manglende værdier i Stata (specialkoder) skal</p>	ASTA vil vise fejl under testningen af afleveringspakken





Stata			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		ligge inden for udfaldsrummet .a-.z Kontakt Rigsarkivets datamanager for forskningsdata, hvis du anvender koder for manglende værdier i variable med typerne tekst, tidspunkt eller dato.	
	Alle tildelte koder for manglende værdier skal også tildeles en værdi i value labels knyttet til den variabel, hvor koden anvendes	Forståelse af den tildelte kode for en missing value er vigtig for at forstå data. Derfor skal alle anvendte koder for manglende værdier (missing values) også tildeles en værdi i value labels knyttet til den variabel, hvor koden anvendes. Fx .u = Uoplyst .i = Irrelevant .d = Deltager ikke Du kan tilføje value labels for disse koder for missing values i	ASTA vil vise fejl under testningen af afleveringspakken





Stata			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		"Variables Manager" i Stata.	
8.Referencer (ved aflevering af flere datasæt)	Tjek at eventuelle flettevariable har samme type/format og længde	<p>Flettevariable angives i afleveringspakken for at beskrive mulig sammenkobling mellem flere datasæt.</p> <p>Når der findes flere datasæt i en afleveringspakke og referencer mellem disse (en fletningsnøgle), skal der være fuldstændig overensstemmelse mellem flettevariablenes type og længde.</p> <p>Tjek at formatet og typen af disse flettevariable er identiske i "Properties"-vinduet i Stata for hver af de afleverede filer.</p>	<p>Manglende overensstemmelse mellem type og længde i flettevariable kan påvirke muligheden for at sammenkoble de afleverede datasæt.</p> <p>Efter aflevering til arkiv vil arkivet kontrollere, om denne regel er overholdt. Spar tid ved at gøre dette på forhånd.</p>

C. SAS – forberedelse af statistikfilen før udtræk

SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
GENERELT			





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
1. Tjek/installer statistikprogram	Tjek at du har statistikprogrammet SAS installeret på den PC du anvender ASTA på.		Du kan ikke udtrække data fra en SAS-fil med ASTA, hvis statistikprogrammet SAS ikke er installeret på PC'en
2. Ny mappe med kopi af materiale	<p>Opret en ny mappe på computeren med en kopi af statistikfilerne der skal afleveres, kontekstdokumenterne, som fremgår af afleveringsbestemmelsen, og de to indeksfiler.</p> <p>Navngiv evt. mappen med afleveringspakkens løbenummer, der fremgår af arkivets afleveringsbestemmelse.</p>	Da forberedelsen af afleveringen involverer ændringer i variablene, justering af formater og konvertering af filer, er det vigtigt at bevare integriteten af de originale filer ved at separere dem fra afleveringsmaterialet.	De originale data kan blive overskrevet eller slettet
SAS DATAFILER			





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
3.Datafiler	Tjek at SAS filerne har extension <i>.sas7bdat</i>	<p>Data skal være i et dataformat, som ASTA kan lave udtræk fra.</p> <p>For SAS filer skal ekstension <i>.sas7bdat</i> anvendes til datafilen og katalogfilen (formatfilen) skal have extension <i>.sas7bcats</i></p> <p>Bemærk at både datafilen og katalogfilen skal have præcis samme navn, før udtræk af formater fra katalogfilen med ASTA er mulig.</p>	ASTA kan kun udtrække data fra formaterne SPSS, SAS, Stata og R/RStudio.
	Datafiler er kodet til UTF-8	<p>Alle datafiler skal være kodet til UTF-8.</p> <p>Ydermere skal du undersøge om dit datasæt viser alle tegnene korrekt specielt æ, ø og å også.</p> <p>Se vejledning om UTF-8 på Rigsarkivets hjemmeside.</p>	<p>Hvis data ikke er indkodet til UTF-8, vil nogle tegn såsom æ, ø og å ikke blive vist korrekt.</p> <p>Efter aflevering til arkiv vil arkivet også teste afleveringen. Hvis data ikke er indkodet som UTF-8 kan det medføre andre fejl i afleveringen, som kun kan rettes ved at lave helt nye</p>





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
			udtræk fra de originale statistikfiler. Så spar tid ved at kontrollere det inden udtræk.
4.Variable i datafiler	Tjek at alle variabelnavne overholder regler for navngivning af variable	Brug 'Proc Contents' i SAS for at se alle dine variable, deres navne, labels og formater. Juster enhver variabel, der ikke overholder regler for navngivning af variable.	ASTA vil vise fejl under testningen af afleveringspakken
	Tjek at alle variable har variable labels	Alle variable skal have variabel labels, der beskriver deres indhold. Brug 'Proc Contents' i SAS til at kontrollere at alle variable har labels. Hvis en variabel mangler et "Label Statement" kan du tilføje dette med denne syntaks: 'LABEL variable = "label" ... ' Fx 'LABEL koen = "Respondentens køn" alder = "Respondentens alder";'	ASTA vil vise fejl under testningen af afleveringspakken





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
5.Variable type/format	Tjek at alle variable har angivet en type/format, som er accepteret af Rigsarkivet, og ikke er angivet med "default formats"	<p>Alle variable skal have angivet type og format, som er accepteret af Rigsarkivet og ikke "default formats".</p> <p>Denne del er særligt vigtig for SAS-brugere, da standarden for SAS er, at datasæt ikke har tildelt formater. Derfor anvendes typisk standard indstillingen (default formatet), med mindre brugeren eksplicit har angivet format.</p> <p>For at tjekke formaterne kør en 'Proc Contents' i SAS og kontrollér at outputformatet i kolonnen "Format" ikke er tom. Kolonnen skal enten vise formater fra et formatkatalog (via 'Proc Format' proceduren) eller et format angivet af dig via et 'Format statement', som dette:</p> <pre>'FORMAT variable-1 <. . . variable-n> <format>;'</pre>	ASTA vil vise fejl under testningen af afleveringspakken





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		Fx 'FORMAT koen koen. alder f3.,'	





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
6.Value labels	Tjek at alle value labels er unikke og specificeret	<p>Alle værdier i value labels skal være unikke og specificeret.</p> <p>I tilfælde af at nogle værdier ikke er specificeret (fx i en Likert skala hvor kun den første og sidste værdi har specificeret labels), skal dette beskrives i variabelens variable label fx med teksten "Skala anvendt. Ikke alle koder har kodebeskrivelse."</p> <p>Kør nedenstående syntaks for at tjekke, om alle værdier har fået angivet value labels i datasættet og kontrollér, at alle niveauer og missing values er specificeret i value labels. Hvis dette ikke er tilfældet, tilføjes de manglende value labels.</p> <p>'Proc format library=xxx.yyy select x;'</p> <p>Fx: Hvis du vil se formatet for variablen "alder" i katalogfilen "Formats" i biblioteket</p>	Efter aflevering til arkiv vil arkivets tester teste om alle value labels findes





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>"MyLib" skrives:</p> <pre>'Proc format library=MyLib.Formats select alder;;'</pre> <p>Fx: Hvis du vil se formater på alle variable i katalogfilen "Formats" i biblioteket "Mylib" kan du skrive: 'Proc format library=MyLib.Formats select a-z;;'</p>	
7.Missing values	Kontrollér at det kun er numeriske og kategoriske variable, der anvender koder for missing values.	<p>Manglende værdier må kun anvendes på numeriske og kategoriske variable.</p> <p>Kontrollér at det kun er numeriske og kategoriske variable, der anvender koder for missing values i "value labels" ved at køre en Proc format syntaks:</p> <pre>'Proc format library=xxx.yyy select x;;'</pre> <p>Se eksempler under punkt 6.</p> <p>Koder for manglende værdier i SAS (specialkoder) skal ligge inden for udfaldsrummet A-Z.</p> <p>Kontakt Rigsarkivets</p>	ASTA vil vise fejl under testningen af afleveringspakken





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		datamanager for forskningsdata, hvis du anvender koder for manglende værdier i variable med typerne tekst, tidspunkt eller dato.	
	Alle tildelte koder for manglende værdier skal også tildeles en værdi i value labels knyttet til den variabel, hvor koden anvendes.	<p>Forståelse af den tildelte kode for en missing value er vigtig for at forstå data.</p> <p>Derfor skal alle anvendte koder for manglende værdier (missing values) også tildeles en værdi i value labels knyttet til den variabel, hvor koden anvendes.</p> <p>Fx</p> <p>U = Uoplyst</p> <p>I = Irrelevant</p> <p>D = Deltager ikke</p> <p>Du kan tilføje value labels for disse koder for missing values via en 'Proc Format' procedure i SAS.</p>	ASTA vil vise fejl under testningen af afleveringspakken
8.Referencer (ved aflevering af flere datasæt)	Tjek at eventuelle flettevariable har samme type/format og længde	Flettevariable angives i afleveringspakken for at beskrive mulig sammenkobling mellem flere datasæt.	Manglende overensstemmelse mellem type og længde i flettevariable kan påvirke muligheden





SAS			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>Når der findes flere datasæt i en afleveringspakke og referencer mellem disse (en fletningsnøgle), skal der være fuldstændig overensstemmelse mellem flettevariablenes type og længde.</p> <p>Tjek at formatet og typen af disse flettevariable er identiske. For at tjekke formaterne kørs en 'Proc Contents' i SAS og kontrollér type af outputformatet i kolonnen "Format" og længden i kolonne "Len".</p>	<p>for at sammenkoble de afleverede datasæt.</p> <p>Efter aflevering til arkiv vil arkivet kontrollere, om denne regel er overholdt. Spar tid ved at gøre dette på forhånd.</p>



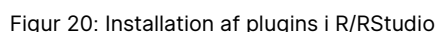


Du skal installere følgende plugin før du kan køre R eksport scriptet fra ASTA:

- Du skal installere følgende plugin før du kan konvertere tidspunkter og datoformater i R/RStudio til de tilladte dataformater for disse, jf. figur 9.3, i bilag 9 til bekendtgørelse om arkiveringsversioner nr. 128:

- Datoerne skal konverteres til tilladte formater inden du laver udtræk med R eksport scriptet fra ASTA.

For at installere disse plugins (packages) i R/RStudio skal du i console, skrive syntaksen `Install.packages("plugin-name")`, f.x. `Install.packages("haven")`, efter at de er downloadet fra Internet. Hvis computeren ikke er online, kan du downloade disse plugins fra <https://cran.Rstudio.com/src/contrib/PACKAGES>. Bagefter skal du klikke på 'Install Packages' på vinduet i højre side af RStudio, hente plugin og køre den ind i programmet med at angive syntaksen `Library(plugin-name)`, f.x. `Library(haven)`.





Rigsarkivet

Vi anbefaler at du kun udtrækker data fra R filer, som oprindeligt er skabt i R sprog med R eksport scriptet fra ASTA. Selvom R kan importere SAS, SPSS og Stata filer og gem dem i R format, sker der under import af data i R flere transformationer der ændrer data, fx bliver alle brugerdefinerede koder eller specialkoder for manglende værdier (missing values) erstattet med NA eller blanktegn. Generelt er der risiko for konverteringsfejl og tab af vigtig information, når andre statistikformater indlæses i R/RStudio.

R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
GENERELT			
1.Tjek/installer statistikprogram	Tjek at du har statistikprogrammet, enten R eller Rstudio, installeret på den PC du anvender ASTA på.		Du kan ikke udtrække data fra en R-fil med ASTA, hvis et af statistikprogrammene R eller RStudio ikke er installeret på PC'en
2.Ny mappe med kopi af materiale	Opret en ny mappe på computeren med en kopi af statistikfilerne der skal afleveres, kontekstdokumenterne, som fremgår af afleveringsbestemmelsen, og de to indeksfiler. Navngiv evt. mappen med afleveringspakkens løbenummer, der fremgår af arkivets afleveringsbestemmelse.	Da forberedelsen af afleveringen involverer ændringer i variablene, justering af formater og konvertering af filer, er det vigtigt at bevare integriteten af de originale filer ved at separere dem fra afleveringsmaterialet.	De originale data kan blive overskrevet eller slettet
R DATAFILER			
3.Datafiler	Tjek at R filerne har extension <i>.rds</i> .	Data skal være i et dataformat, som ASTA kan lave udtræk fra.	ASTA kan kun udtrække data fra





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		R/RStudio filer skal have ekstensions <i>.rds</i> .	SPSS, SAS, Stata og R/RStudio
	Tjek at datafiler er kodet til UTF-8	<p>Alle datafiler, der afleveres til arkivet, skal være kodet til UTF-8.</p> <p>Ydermere skal du undersøge om datafilen viser alle tegnene korrekt, specielt æ, ø og å.</p> <p>Se vejledning om UTF-8 på Rigsarkivets hjemmeside.</p>	<p>Hvis data ikke er indkodet til UTF-8, vil nogle tegn såsom æ, ø og å ikke blive vist korrekt.</p> <p>Efter aflevering til arkiv vil arkivet også teste afleveringen. Hvis data ikke er indkodet som UTF-8 kan det medføre andre fejl i afleveringen, som kun kan rettes ved at lave helt nye udtræk fra de originale statistikfiler. Du sparer således tid ved at kontrollere det inden udtræk.</p>
4.Variable i datafiler	Tjek at alle variabelnavne overholder regler for navngivning af variable	<p>Til at tjekke alle variabelnavne i et datasæt kan du i R/Rstudio bruge syntaksen:</p> <p><code>str(objekt,...)</code></p> <p>Eksempel, hvor datasættet hedder <code>mydata</code>:</p> <p><code>str(mydata)</code></p>	ASTA vil vise fejl under testningen af afleveringspakken





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>Denne syntaks viser alle variable; deres navne, labels, factor levels (koder og kodeforklaringer) og de første få værdier i hver variabel.</p> <p>Du kan ændre variabelnavne i R/RStudio, så de overholder regler for navngivning af variable med syntaksen:</p> <pre>names(datasetname) [variableposition] = "newvariablename"</pre> <p>Eksempel hvor datasættet hedder mydata og navn på første [1] variabel navngives til cpr_nr:</p> <pre>names(mydata) [1] = 'cpr_nr'</pre> <p>Husk at gemme de ændringer du laver i R/RStudio, fx i en ny fil. Du kan anvende følgende syntaks for at gemme det ændrede datasæt som en ny R datafil:</p> <pre>saveRDS(object=datasetname, file="path/datasetname.extension")</pre>	





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>Eksempel, hvor datasættet hedder mydata og gemmes med det nye datasætnavn mydata2:</p> <pre>saveRDS(object=mydata, file="C:/Users/lokalAdmin/Desktop/Testdata/mydata2.rds")</pre>	
	Tjek at alle variable har variabel labels (variabelbeskrivelser).	<p>Alle variable skal have variabel labels, der beskriver deres indhold. Det er et krav i Bilag 9.</p> <p>En statistikfil der oprindeligt er skabt i R har sandsynligvis ikke variabel labels. Derfor skal du selv angive disse variabelbeskrivelser.</p> <p>Du kan anvende følgende syntaks:</p> <pre>attributes(datasetname\$variablename)\$label <- "Description"</pre> <p>Eksempel, hvor variabelbeskrivelsen Respondentens cpr-nummer angives for variabelen cpr_nr i datafilen mydata:</p>	ASTA vil vise fejl under testningen af afleveringspakken





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<code>attributes(mydata\$cpr_nr)</code> <code>)\$label <-</code> <code>"Respondentens cpr-</code> <code>nummer"</code>	
5.Variable type/format	Tjek at alle variable har angivet en type/format (dataformatnotation), som er accepteret af Rigsarkivet, og ikke er angivet med "default formats"	<p>Alle variable skal have angivet en notation for dataformatet anvendt i datasættet, som er accepteret af Rigsarkivet. "Default formats" er ikke tilladt.</p> <p>For at tjekke alle variablenes type/format i datasættet, kan du anvende syntasken <code>str(object,...)</code> i R/Rstudio.</p> <p>Eksempel:</p> <p><code>str(mydata)</code> for at se alle attributter i filen 'mydata'</p> <p>Hvis en eller flere angivne dataformatnotationer ikke overholder Rigsarkivets krav til disse, jf. figur 9.3 i bilag 9 i bekendtgørelse nr. 128, skal du ændre variablenes dataformatnotationer.</p> <p>Alle datoer, tidsstempler og tidspunkter indlæses i R/RStudio som tekst (string). Dvs. der ikke nødvendigvis er en korrekt</p>	ASTA vil vise fejl under testningen af afleveringspakken





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>dataformatnotation på disse. Du skal således være ekstra opmærksom på at tjekke om disse datatyper har fået påsat korrekt</p> <p>dataformatnotation under udtræk med ASTA. Hvis de er angivet som string skal det ændres til en korrekt</p> <p>dataformatnotation (fx xml notationerne time, date eller datetime afhængigt af indholdet i datasættet). Hvis indholdet i datasættet ikke overholder den angivne tilladte dataformatnotation skal fx datoen konverteres til korrekt gyldigt datoformat i R- datafilen før udtræk. Se figur 9.3 i bilag 9 over tilladte datatyper i datafilen.</p> <p>Plugin 'chron' til R/RStudio kan anvendes til disse konverteringer af tid- og datoformater.</p>	
6.Value labels	Tjek at alle factor levels (value labels; koder og kodebeskrivelser) er unikke og specificeret.	Alle værdier i factor levels (value labels; koder og kodebeskrivelser) skal være unikke og specificeret.	Efter aflevering til arkiv vil arkivets tester teste om alle factor levels/value labels findes.





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		<p>Statistikfiler som oprindeligt er skabt R/RStudio har sandsynligvis ikke factor levels/value labels defineret. Derfor skal du selv angive factor levels til variablene som indeholder koder som skal forklares. Du kan anvende nedenstående syntaks:</p> <pre>x <- c(0,1)</pre> <pre>names(x) <- c("X", "Y")</pre> <pre>attributes(datasetname\$variablename)\$labels <- x</pre> <p>Eksempel, hvor datasættets navn er mydata og der påsættes faktor levels til variablen sex:</p> <pre>x <- c(1,2)</pre> <pre>names(x) <- c("Male", "Female")</pre> <pre>attributes(mydata\$sex)\$labels <- x</pre>	
7.Missing values	I R-filer er det ikke muligt at angive koder for manglende værdier (missing values)	Statistikfiler som oprindeligt er skabt i R/RStudio indeholder ikke koder for manglende værdier.	





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		Når en SPSS, Stata eller SAS fil indlæses i R erstattes de angivne koder for manglende værdier med NA eller blanktegn. Derfor anbefales det ikke at udtrække data fra disse filtyper med R eksport scriptet.	
8.Referencer (ved aflevering af flere datasæt)	Tjek at eventuelle flettevariable har samme type/format og længde	<p>Flettevariable angives i afleveringspakken for at beskrive mulig sammenkobling mellem flere datasæt.</p> <p>Når der findes flere datasæt i en afleveringspakke og referencer mellem disse (en fletningsnøgle), skal der være fuldstændig overensstemmelse mellem flettevariablenes type og længde.</p> <p>Tjek at formatet og typen af disse flettevariable er identiske.</p> <p>For at tjekke formaterne kørsyntaxen <code>str(objekt, ...)</code> i R/RStudio og kontrollér type af formatet på variablene listet i output.</p>	<p>Manglende overensstemmelse mellem type og længde i flettevariable kan påvirke muligheden for at sammenkoble de afleverede datasæt.</p> <p>Efter aflevering til arkiv vil arkivet kontrollere, om denne regel er overholdt. Du sparer således tid ved at gøre dette på forhånd.</p>





R/RStudio			
NR.	OPGAVE	VEJLEDNING	HVIS JEG IKKE GØR DET
		Eksempel: str(mydata) for at se alle attributterne, herunder formaterne, i filen mydata	

