

National Asset Project Vejledning 5.0

Dette projekt har til formål at indsamle og analysere økonomiske regnskaber fra danske virksomheder i perioden 2014-nu. Projektet har en primær fokus på vurdering af, hvordan virksomheder forvalter deres aktiver. Gennem en omfattende dataindsamling og analyse vil vi generere grafer og visuelle repræsentationer for at belyse forskellige aspekter af årsregnskaberne.

All data hentes fra Virk.dk

Nyeste version af alle filer findes under på Remote Desktop under *National Asset Project 2023* og ekstra dokumenter kan finde på DriveHQ: \\Reliaset\Projekter\Nationale aktiver.

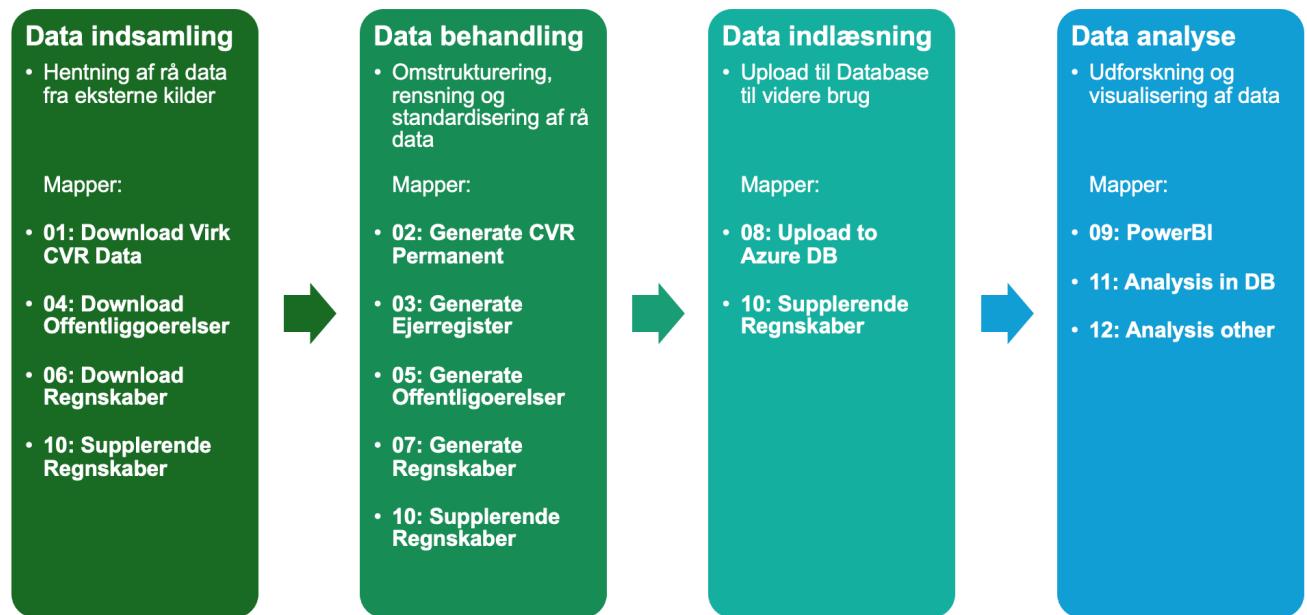
Bemærk, at versioner og årstal i dokumenterne kan være forældede og ikke nødvendigvis afspejler den aktuelle version.

Indholdsfortegnelse

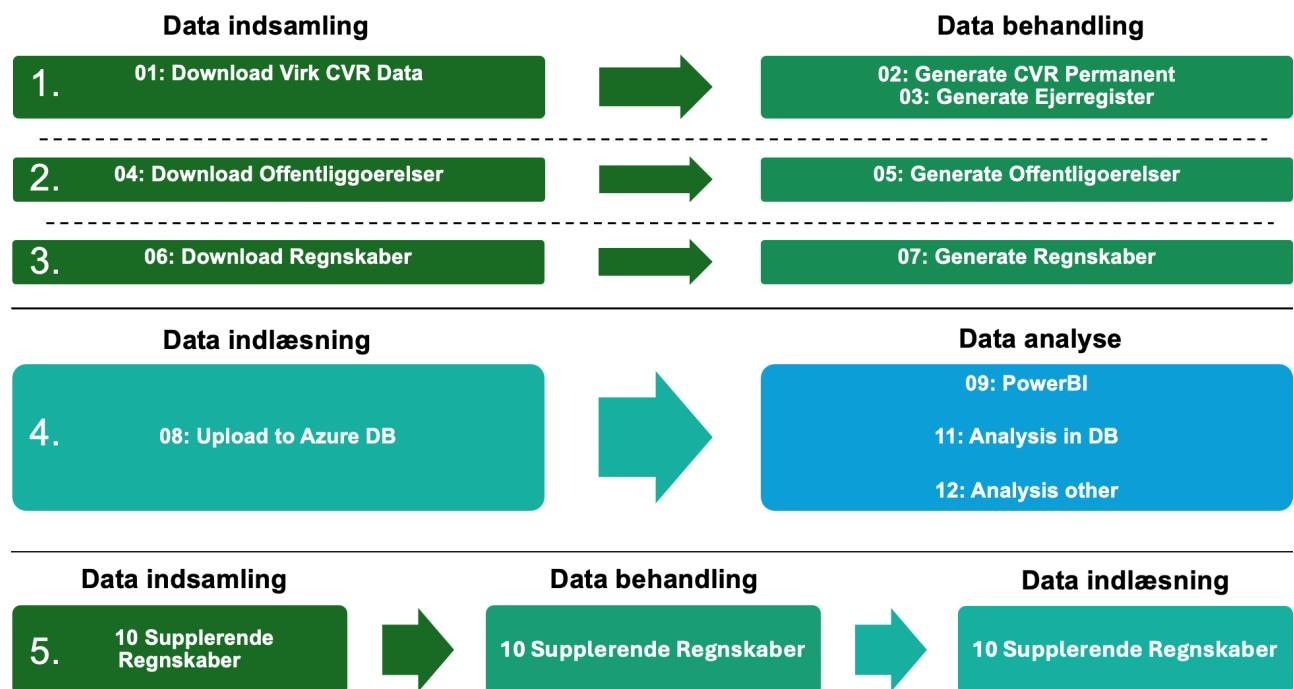
01 - Download Virk CVR data	3
02 - Generate CVR-Permanent	5
03 - Generate Ejerregister	6
04 - Download Offentliggørelser	7
05 - Generate Offentliggoerelser.....	8
06 - Download regnskaber	9
07 - Generate Regnskaber.....	11
08 - Upload to Azure DB	14
09 - Power BI	15
10 - Supplerende Data	16

National Assets Projekt struktur

'Mapper' refererer til Reliaset's remote desktop hvor mapperne kan findes lokalt.



Tilgang til vejledning



01 - Download Virk CVR data (30-35 min)

Remote Desktop: E:\National Asset Project 2023\01 - Download Virk CVR data

Download de informationsdata for alle virksomheder der findes inde på Virk fra og med år 2012.

Kør curl scriptet:

“curl_query_cvr.bat”

Curl scriptet henter alt data inde fra virk med hjælp fra Query: “curl_query_cvr.json”, som sikre at vi kun henter fra 2012 og op:

```
{
  "query": {
    "bool": {
      "filter": [
        {
          "bool": {
            "must": {
              "exists": {
                "field": "Vrvirksomhed.cvrNummer"
              }
            }
          }
        },
        {
          "bool": {
            "should": [
              {
                "range": {
                  "Vrvirksomhed.livsforloeb.periode.gyldigTil": {"gte": "2012-01-01"}
                }
              },
              {
                "bool": {
                  "must_not": {
                    "exists": {
                      "field": "Vrvirksomhed.livsforloeb.periode.gyldigTil"
                    }
                  }
                }
              }
            ],
            "minimum_should_match": 1
          }
        }
      ]
    },
    "_source": [ "Vrvirksomhed" ],
    "size": 3000
  }
}
```

Figur 1: curl_query_cvr.json

PROBLEMER: Tidligere har vi haft problemer med syntaks når scriptet “curl_query_cvr.bat” køres:

```
echo off
title Henter Virk CVR-data ud fra curl_query_cvr.json
cd "E:\National Asset Project 2023\01 - Download Virk CVR data"
del VirkAllCVRdata-*.json

curl -XPOST -H "Content-Type: application/json" -o VirkCVRdata-001.json -u Reliasset_CVR_I_SKYEN:522e52b7-32b4-475a-b941-74ef4562b315
http://distribution.virk.dk/cvr-permanent/_search?scroll=5m -d @"curl_query_cvr.json"
```

Screenshot fra curl_query_cvr.bat fil

Her skal anførelsestegn huskes når -d query tilføjes som ses på billedet.

Indsæt output filerne (som har navnet ”VirkCVRdata-<SUFFIX>.json”) i mappen ’E:\National Asset Project 2023\Data’ hvorefter alle filernes gemmes under ny mappe: ”VirkCVRData”.

Mappen VirkCVRData kan nu oprettes som database i BaseX (Grundet størrelsen af mappen kan det være nødvendigt at splitte database op i 2: VirkCVRData-1 og VirkCVRData-2)

HVIS: Der eksisterer en database i BaseX med samme navn skal den gamle database droppes.

PROBLEMER: BaseX skal åbnes igennem bat filen: “**basexgui.bat - Shortcut**” som findes i mappen: Remote Desktop: E:\National Asset Project 2023. Hvis ikke BaseX åbnes igennem scriptet kan følgende fejl opstå:



1) Åben BaseX igennem bat filen: "basexgui.bat - Shortcut"

2) Database/Open and Manage

Drop databases: VirkCVRdata-1 **and** VirkCVRdata-2

3) Database/New

Input file or directory: E:/National Asset Project 2023/Data/VirkCVRData-1/

Name of database: VirkCVRdata-1

Input format: JSON

Parse files in archives: x

Gøres for begge mapper:

Database/New

Input file or directory: E:/National Asset Project 2023/Data/VirkCVRData-2/

Name of database: VirkCVRdata-2

Input format: JSON

Parse files in archives: x

02 - Generate CVR-Permanent

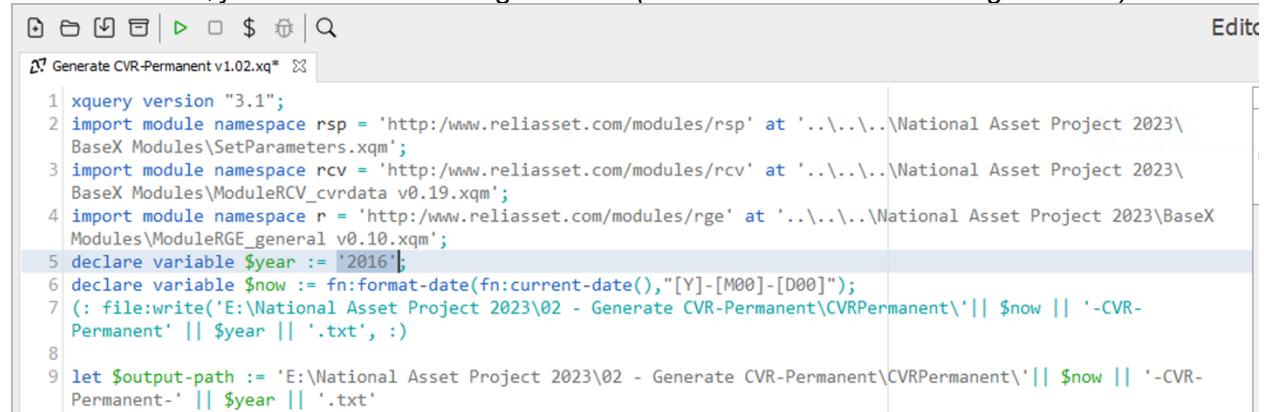
Remote Desktop: E:\National Asset Project 2023\02 - Generate CVR-Permanent

Baseret på den database der er blevet oprettet i **01** kan vi nu generere txt filer der indeholder basal CVR-information for alle virksomheder pr. år.

1. Åben BaseX igennem bat filen: "basegui.bat - Shortcut"

2. Åben "Generate CVR-Permanent v1.02.xq"

3. Erstat året i \$year til året som skal genereres (*Alle år fra 2012-2023 skal genereres*).

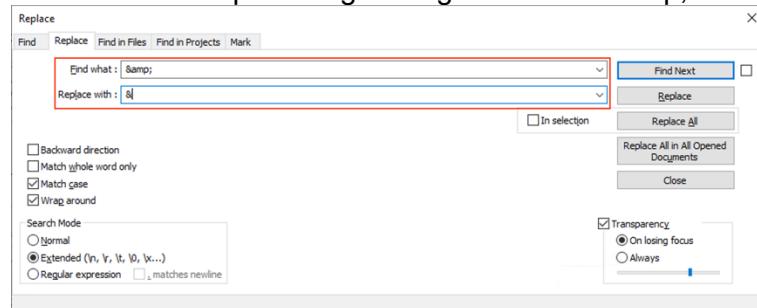


```
xquery version "3.1";
import module namespace rsp = 'http://www.reliasset.com/modules/rsp' at '../../National Asset Project 2023/BaseX Modules/SetParameters.xqm';
import module namespace rcv = 'http://www.reliasset.com/modules/rcv' at '../../National Asset Project 2023/BaseX Modules/ModuleRCV_cvrdatas v0.19.xqm';
import module namespace r = 'http://www.reliasset.com/modules/rge' at '../../National Asset Project 2023/BaseX Modules\ModuleRGE_general v0.10.xqm';
declare variable $year := "2016";
declare variable $now := fn:format-date(fn:current-date(),"[Y]-[M00]-[D00]");
(: file:write('E:\National Asset Project 2023\02 - Generate CVR-Permanent\CVRPermanent'|| $now || '-CVR-Permanent' || $year || '.txt', :)
let $output-path := 'E:\National Asset Project 2023\02 - Generate CVR-Permanent\CVRPermanent'|| $now || '-CVR-Permanent-' || $year || '.txt'
```

4. Kør BaseX query

5. Output fil findes I *Remote Desktop: E:\National Asset Project 2023\02 - Generate CVR-Permanent\CVRPermanent*
gemt med navnet: <dato>-CVR-Permanent-<år>.txt

Åben txt fil i Notepad++ og find og erstat alle '&' til '&'



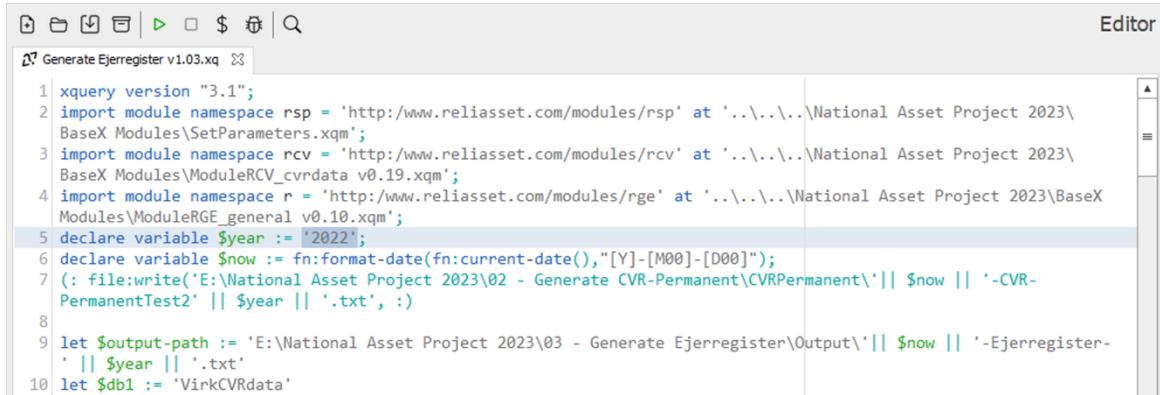
6. Gentag for alle år.

03 - Generate Ejerregister

Remote Desktop: E:\National Asset Project 2023\03 - Generate Ejerregister

Baseret på den database der er blevet oprettet i **01** kan vi nu generere txt filer der indeholder Ejerregister information for alle virksomheder pr. år.

1. Åben BaseX igennem bat filen: "basexgui.bat - Shortcut"
2. Åben "Generate Ejerregister v1.03.xq"
3. Erstat året i \$year til året som skal genereres.

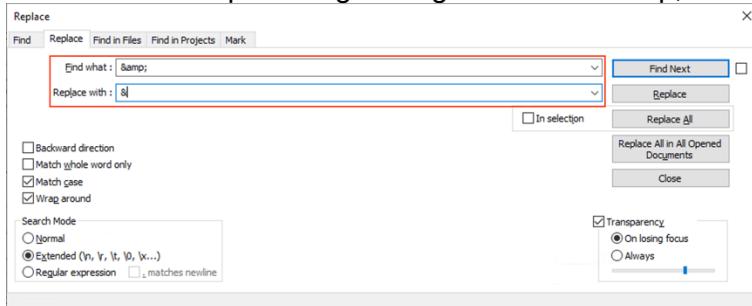


```
xquery version "3.1";
import module namespace rsp = 'http://www.reliasset.com/modules/rsp' at '../../National Asset Project 2023/BaseX Modules/SetParameters.xqm';
import module namespace rcv = 'http://www.reliasset.com/modules/rcv' at '../../National Asset Project 2023/BaseX Modules/ModuleRCV_cvrdta v0.19.xqm';
import module namespace r = 'http://www.reliasset.com/modules/rge' at '../../National Asset Project 2023/BaseX Modules/ModuleRGE_general v0.10.xqm';
declare variable $year := '2022';
declare variable $now := fn:format-date(fn:current-date(),"[Y]-[M00]-[D00]");
(: file:write('E:\National Asset Project 2023\02 - Generate CVR-Permanent\CVRPermanent'|| $now || '-CVR-PermanentTest2' || $year || '.txt', :)
let $output-path := 'E:\National Asset Project 2023\03 - Generate Ejerregister\Output'|| $now || '-Ejerregister-' || $year || '.txt'
let $db1 := 'VirkCVRdata'
```

4. Kør BaseX query

5. Output fil findes I Remote Desktop: E:\National Asset Project 2023\02 - Generate CVR-Permanent\CVRPermanent
gemt med navnet: <dato>-Ejerregister-<år>.txt

Åben txt fil i Notepad++ og find og erstat alle '&' til '&



6. Gentag for alle år.

04 - Download Offentliggørelser

Remote Desktop: E:\National Asset Project 2023\04 - Download Offentliggoerelser

Vi downloader nu de offentliggørelser for alle virksomheder der findes inde på Virk fra og med år 2012.

1. Kør "curl_query_offentliggoerelser_all.bat"
2. Indtast offentliggørelsесår
3. Vent på at filerne er downloadet
4. Flyt alle Offentliggoerelser<ÅR>-<SUFFIX>.json til:
"Remote Desktop: E:\National Asset Project 2023\Data\Offentliggoerelser<ÅR>"
(opret mappen "Ofentliggoerelser<ÅR>" hvis den ikke findes)
5. Åben BaseX og opret ny database:
Database/New
Input file or directory: E:/National Asset Project 2023/Data/Offentliggoerelser<ÅR>/
Name of database: Offentliggoerelser<ÅR>
Input format: JSON
Parse files in archives: x
6. Gentag for alle år

05 - Generate Offentliggoerelser

Remote Desktop: E:\National Asset Project 2023\05 - Generate Offentliggoerelser

Baseret på den database der er blevet oprettet i **04** kan vi nu generere txt filer der indeholder offentliggoerelser information for alle virksomheder pr. år.

1. Åben BaseX igennem bat filen: "basexgui.bat - Shortcut"

2. Åben "Generate Offentliggoerelser v1.03.xq"

3. Erstat året i \$year til året som skal genereres



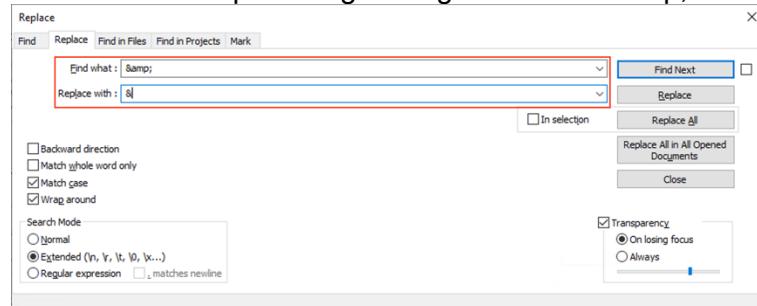
```
Editor
D:\ Generate Offentliggoerelser v1.03.xq
1 xquery version "3.1";
2
3 import module namespace rsp = 'http://www.reliasset.com/modules/rsp' at
4 '../../..\\National Asset Project 2023\\BaseX Modules\\SetParameters.xqm';
5 import module namespace r = 'http://www.reliasset.com/modules/rge' at '../../..\\..\\..\\National Asset Project 2023\\BaseX
Modules\\ModulerGE_general v0.10.xqm';
6
7 declare variable $year := '2023';
8
9 let $now := fn:format-date(fn:current-date(), "[Y]-[M00]-[D00]")
10 let $output-path :=
11   'E:\\National Asset Project 2023\\06 - Generate Offentliggoerelser\\Offentliggoerelser\\' || $now || '-
Offentliggoerelse' || $year || '.txt'
```

4. Kør BaseX query

5. Output fil kan findes:

Remote Desktop: E:\National Asset Project 2023\06 - Generate Offentliggoerelser\Offentliggoerelser
gemt med navnet: <DATO>-Offentliggoerelse<ÅR>.txt

Åben txt fil i Notepad++ og find og erstat alle '&' til '&



6. Gentag for alle år

06 - Download regnskaber

Remote Desktop: E:\National Asset Project 2023\06 - Download regnskaber

1. Kør BaseX filen "QueryUrlForDownload.xq" for at hente regnskabs url.
2. Database der skal åbnes (oprettet i 04) ændres til år ønsket ('Offentliggoerelser<ÅR>')
3. Filen gemmes manuelt som "<DATO>-UrlForDownloadXml<ÅR>.txt"

The screenshot shows the BaseX Query Editor interface. On the left, the file tree displays the 'BaseX' directory structure. In the center, the 'Query UrlForDownload.xq' editor pane contains the following XQuery code:

```
for $x in db:open("Offentliggoerelser2023")//hits/hits/_/source
let $dokumentUrl := $x/dokumenter/_/dokumentMimeType = "application/xml"/dokumentUrl
return data($dokumentUrl)
```

A red arrow points from the text "Ændres til år der ønskes" to the line 'let \$dokumentUrl := \$x/dokumenter/_/dokumentMimeType = "application/xml"/dokumentUrl'. Below the editor is the 'Result' pane, which shows the output of the query. A red arrow points to the 'After output: gem som txt fil' button. The result pane displays a large list of URLs, starting with:

```
http://regnskaber.virk.dk/9670001/
amlsb3VkcsovLzA1zR1L2JhLz1Mz10LzBmNDItNDU4MS04M02LTQyNjkxYzA1IjQ3Yw.xml
http://regnskaber.virk.dk/6680001/
amlsb3VkcsovLzA1zE0LzN1L20LzdkLzBjMwQtHDkMy04YTZhTk1NTFjMDg5MGUSzg.xml
http://regnskaber.virk.dk/3231001/
amlsb3VkcsovLzA1z2Lz2KzL2MzLzN1NDItNDQ8M11NGImLT1INGU3NT1kTEyWw.xml
http://regnskaber.virk.dk/5671001/
amlsb3VkcsovLzA1z2LzYLz1L2Jl2RhLzLyIDgtNDM4y1hNWJjlNR1YwJ1MjySMj1kYg.xml
http://regnskaber.virk.dk/20230001/
amlsb3VkcsovLzA1z2LzYLz1L2Jl2RhLzLyIDgtNDM4y1hNWJjlNR1YwJ1MjySMj1kYg.xml
http://regnskaber.virk.dk/56830001/
amlsb3VkcsovLzA1z2Lz1g5LzM1L2R1L21ZGmtNDY90C05ZjIxLTQvMTvhK0W5NTE1Yw.xml
http://regnskaber.virk.dk/02040001/
amlsb3VkcsovLzA1z2Lz1YzLzR1L2ExLzQ4NTEtNDM0MS04ZmFnLw30Dj3jNGYzZjcvNh.xml
http://regnskaber.virk.dk/38150001/
amlsb3VkcsovLzA1zI4LzR1LzgvLzE0LzJ3YItNDE1ZC1iy2Y2LTE1ZTmzM2HzWljZg.xml
http://regnskaber.virk.dk/53500001/
```

The bottom right corner of the result pane shows a progress bar indicating the total time taken for the query execution.

Herefter kan vi via curl trække alle url ud separat:

4. Kør script:
"curl_download_all.bat"
 - a. Input fil gives som følgende:
"<DATO>-UrlForDownloadXml<ÅR>.txt"
 - b. Outputfolder skal oprettes for gældende år og gives som output folder i script:
Remote Desktop: E:\National Asset Project 2023\06 - Download
regnskaber\Output\<ÅR>

Herefter oprettes outputfolderen som en database i BaseX med navn: "XBRL<ÅR>":

5. Åben BaseX og opret ny database:

Database/New

Input file or directory: E:\National Asset Project 2023\06 - Download regnskaber\Output\<ÅR>

Name of database: XBRL<ÅR>

Input format: XML

Parse files in archives: x

PROBLEMER: Hvis mappen er for stor til at blive oprettet i databasen splittes mappen op i flere

mapper og gemmes på følgende vis:

[XBRL2022-1](#)

[XBRL2022-2](#)

[XBRL2022-3](#)

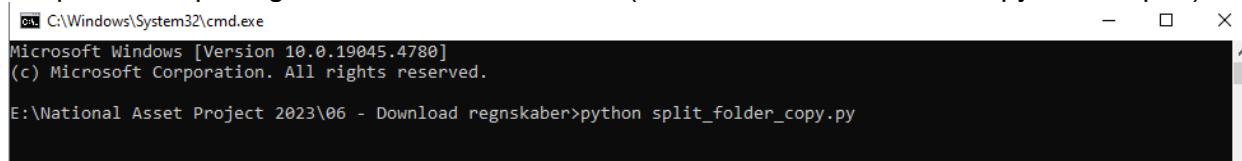
[XBRL2022-4](#)

For at splitte mappen op kan scriptet, 'split_folder_copy.py' bruges:

```
35 if __name__ == "__main__":
36     # Define source and destination paths
37     source_folder = r"E:\National Asset Project 2023\06 - Download regnskaber\Output\2023"
38     destination_folders = [
39         r"E:\National Asset Project 2023\Data\XBRL2023-1",
40         r"E:\National Asset Project 2023\Data\XBRL2023-2",
41         r"E:\National Asset Project 2023\Data\XBRL2023-3",
42         r"E:\National Asset Project 2023\Data\XBRL2023-4",
43         r"E:\National Asset Project 2023\Data\XBRL2023-5"
44     ]
```

Her ændres 'source_folder' og 'destination_folders' til ønsket.

Scriptet køres på følgende måde fra terminalen (terminal sat til lokationen af python scriptet):



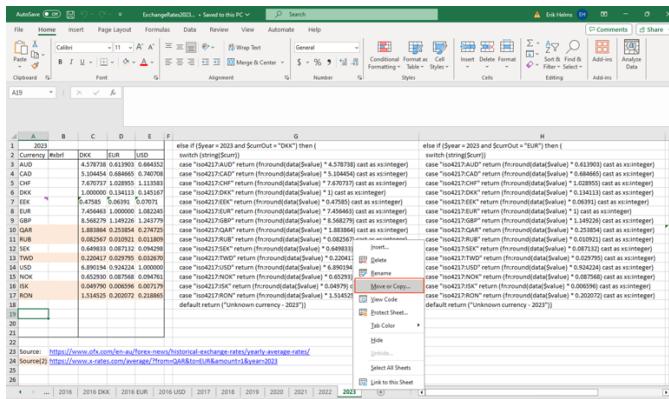
A screenshot of a Windows Command Prompt window. The title bar says 'cmd C:\Windows\System32\cmd.exe'. The window content shows the command 'E:\National Asset Project 2023\06 - Download regnskaber>python split_folder_copy.py' being typed into the prompt. The background of the window is black, and the text is white.

5. Gentag for alle år.

07 - Generate Regnskaber

Remote Desktop: E:\National Asset Project 2023\07 - Generate Regnskaber

1. Inden Regnskaber kan genereres skal **ExchangeRates2023.xlsx** dokumentet opdateres for gældende år. (Dokumentet er på nuværende tidspunkt opdateret til og med 2023 - 01/03 2024)



Kopier ark fra tidligere år og opdater alle valuta kurser for gældende år fra følgende sider:
<https://www.ofx.com/en-au/forex-news/historical-exchange-rates/yearly-average-rates/>
<https://www.x-rates.com/average/?from=QAR&to=EUR&amount=1&year=2023>

2. Efter at have opdateret valuta kurser kan følgende kolonner kopieres over i "ModuleREX_exchangerate v0.13.xqm" som findes:

Remote Desktop: E:\National Asset Project 2023\BaseX modules

<pre>else if (\$year = 2023 and \$currOut = "DKK") then (switch (string(\$curr)) case "iso4217:AUD" return fn:round(data[\$value] * 4.578738) cast as xs:integer case "iso4217:CAD" return fn:round(data[\$value] * 5.104454) cast as xs:integer case "iso4217:CHF" return fn:round(data[\$value] * 7.67073) cast as xs:integer case "iso4217:DKK" return fn:round(data[\$value] * 1) cast as xs:integer case "iso4217:EEK" return fn:round(data[\$value] * 0.47585) cast as xs:integer case "iso4217:EUR" return fn:round(data[\$value] * 7.456463) cast as xs:integer case "iso4217:GBP" return fn:round(data[\$value] * 8.568279) cast as xs:integer case "iso4217:QAR" return fn:round(data[\$value] * 1.883864) cast as xs:integer case "iso4217:RUB" return fn:round(data[\$value] * 0.082567) cast as xs:integer case "iso4217:SEK" return fn:round(data[\$value] * 0.649833) cast as xs:integer case "iso4217:TWD" return fn:round(data[\$value] * 0.220417) cast as xs:integer case "iso4217:NOK" return fn:round(data[\$value] * 0.65293) cast as xs:integer case "iso4217:ISK" return fn:round(data[\$value] * 0.04979) cast as xs:integer case "iso4217:RON" return fn:round(data[\$value] * 1.514525) cast as xs:integer default return ("Unknown currency - 2023")</pre>	<pre>else if (\$year = 2023 and \$currOut = "EUR") then (switch (string(\$curr)) case "iso4217:AUD" return fn:round(data[\$value] * 0.613903) cast as xs:integer case "iso4217:CAD" return fn:round(data[\$value] * 0.684665) cast as xs:integer case "iso4217:CHF" return fn:round(data[\$value] * 1.028955) cast as xs:integer case "iso4217:DKK" return fn:round(data[\$value] * 0.134113) cast as xs:integer case "iso4217:EEK" return fn:round(data[\$value] * 0.06391) cast as xs:integer case "iso4217:EUR" return fn:round(data[\$value] * 1) cast as xs:integer case "iso4217:GBP" return fn:round(data[\$value] * 1.149216) cast as xs:integer case "iso4217:QAR" return fn:round(data[\$value] * 0.253854) cast as xs:integer case "iso4217:RUB" return fn:round(data[\$value] * 0.010921) cast as xs:integer case "iso4217:SEK" return fn:round(data[\$value] * 0.087132) cast as xs:integer case "iso4217:TWD" return fn:round(data[\$value] * 0.029795) cast as xs:integer case "iso4217:NOK" return fn:round(data[\$value] * 0.087568) cast as xs:integer case "iso4217:ISK" return fn:round(data[\$value] * 0.006596) cast as xs:integer case "iso4217:RON" return fn:round(data[\$value] * 0.202072) cast as xs:integer default return ("Unknown currency - 2023")</pre>	<pre>else if (\$year = 2023 and \$currOut = "USD") then (switch (string(\$curr)) case "iso4217:AUD" return fn:round(data[\$value] * 0.664352) cast as xs:integer case "iso4217:CAD" return fn:round(data[\$value] * 0.740708) cast as xs:integer case "iso4217:CHF" return fn:round(data[\$value] * 1.113583) cast as xs:integer case "iso4217:DKK" return fn:round(data[\$value] * 0.145167) cast as xs:integer case "iso4217:EEK" return fn:round(data[\$value] * 0.07071) cast as xs:integer case "iso4217:EUR" return fn:round(data[\$value] * 1.082245) cast as xs:integer case "iso4217:GBP" return fn:round(data[\$value] * 1.243779) cast as xs:integer case "iso4217:QAR" return fn:round(data[\$value] * 0.094298) cast as xs:integer case "iso4217:RUB" return fn:round(data[\$value] * 0.011809) cast as xs:integer case "iso4217:SEK" return fn:round(data[\$value] * 0.094298) cast as xs:integer case "iso4217:TWD" return fn:round(data[\$value] * 0.03267) cast as xs:integer case "iso4217:USD" return fn:round(data[\$value] * 1) cast as xs:integer case "iso4217:NOK" return fn:round(data[\$value] * 0.094761) cast as xs:integer case "iso4217:ISK" return fn:round(data[\$value] * 0.007179) cast as xs:integer case "iso4217:RON" return fn:round(data[\$value] * 0.218865) cast as xs:integer default return ("Unknown currency - 2023")</pre>
---	--	---

ModuleREX_exchangerate v0.13.xqm:

```

553 case "iso4217:RON" return (fn:round(data($value) * 1.514525) cast as xs:integer)
554 default return ("Unknown currency - 2023"))
555 else if ($year = 2023 and $currOut = "EUR") then (
556 switch (string($curr))
557 case "iso4217:AUD" return (fn:round(data($value) * 0.613903) cast as xs:integer)
558 case "iso4217:CAD" return (fn:round(data($value) * 0.684665) cast as xs:integer)
559 case "iso4217:DKK" return (fn:round(data($value) * 0.134113) cast as xs:integer)
560 case "iso4217:EKK" return (fn:round(data($value) * 0.06391) cast as xs:integer)
561 case "iso4217:EUR" return (fn:round(data($value) * 1) cast as xs:integer)
562 case "iso4217:GBP" return (fn:round(data($value) * 1.149226) cast as xs:integer)
563 case "iso4217:QAR" return (fn:round(data($value) * 0.253854) cast as xs:integer)
564 case "iso4217:RUB" return (fn:round(data($value) * 0.010921) cast as xs:integer)
565 case "iso4217:SEK" return (fn:round(data($value) * 0.087132) cast as xs:integer)
566 case "iso4217:TWD" return (fn:round(data($value) * 0.029795) cast as xs:integer)
567 case "iso4217:USD" return (fn:round(data($value) * 0.924224) cast as xs:integer)
568 case "iso4217:NOK" return (fn:round(data($value) * 0.087568) cast as xs:integer)
569 case "iso4217:ISK" return (fn:round(data($value) * 0.006596) cast as xs:integer)
570 case "iso4217:RON" return (fn:round(data($value) * 0.202072) cast as xs:integer)
571 default return ("Unknown currency - 2023"))
572
573 else if ($year = 2023 and $currOut = "USD") then (
574 switch (string($curr))
575 case "iso4217:AUD" return (fn:round(data($value) * 0.664352) cast as xs:integer)
576 case "iso4217:CAD" return (fn:round(data($value) * 0.740708) cast as xs:integer)
577 case "iso4217:CHF" return (fn:round(data($value) * 1.113583) cast as xs:integer)
578 case "iso4217:DKK" return (fn:round(data($value) * 0.145167) cast as xs:integer)

```

Aa [...] * Find... Replace with... OK 594 : 3

Efter exchange rates er opdateret, kan følgende punkter begyndes på:

Baseret på den database der er blevet oprettet i **06** kan vi nu generere txt filer der indeholder regnskaber for alle virksomheder pr. år.

1. Åben BaseX igennem bat filen: *"basexgui.bat - Shortcut"*

2. Åben "Generate Regnskaber v1.7.xq"

3. Erstat året i \$year til året som skal genereres

```

10 declare default element namespace "http://www.xbrl.org/2003/instance";
11 declare namespace cvr = "http://cvr.dk/report";
12 declare namespace cmn = "http://xbrl.dcca.dk/cmn";
13 declare namespace fsa = "http://xbrl.dcca.dk/fsa";
14 declare namespace gsd = "http://xbrl.dcca.dk/gsd";
15 declare namespace xbrldi = "http://xbrl.org/2006/xbrldi";
16 declare namespace xbrli = "http://www.xbrl.org/2003/instance";
17 declare namespace xlink = "http://www.w3.org/1999/xlink";
18 declare namespace link = "http://www.xbrl.org/2003/linkbase";
19 declare namespace ifrs = "http://xbrl.ifrs.org/taxonomy/2011-03-25/ifrs";
20 declare namespace ifrs-dk = "http://xbrl.dcca.dk/ifrs-dk-cor_2014-12-20";
21 (: declare namespace ifrs-full = "http://xbrl.ifrs.org/taxonomy/2014-03-05;ifrs-full"; :)
22 (: declare namespace ifrs-full="http://xbrl.ifrs.org/taxonomy/2021-03-24;ifrs-full"; :) (:NEW NAMESPACE:)
23
24
25
26 declare variable $year := 2021;
27 declare variable $currOut := 'DKK';
28
29 (: for $year in 2020 to 2022 :)
30 let $now := fn:format-date(fn:current-date(), "[Y]-[M00]-[D00]")
31 let $output-path :=
32   'E:\National Asset Project 2023\07 - Generate Regnskaber\Regnskaber\' || $now || '-Regnskab' || $year || '.txt'
33
34 let $result := (
35 let $xbr10BPart := concat('XBRL', $year)
36 let $ext := ('','-'1','-'2','-'3','-'4','-'5','-'6','-'7','-'8','-'9')

```

4. Kør BaseX query

Eksternt script forståelse

BaseX-queryen **Generate Regnskaber v1.7.xq** benytter et eksternt script til at hente CVR-numre for virksomheder, der anvender deres LEI-nummer som ID i deres regnskaber. Dette sikrer, at alle relevante CVR-numre bliver korrekt matchet og inkluderet i analysen.

Et snippet af hvor det eksterne script køres kan ses for neden:

```
let $LEI :=
  if (string-length($IdentificationNumberCvrOfReportingEntity) > 8)
    then $IdentificationNumberCvrOfReportingEntity
    else ()

let $CVR :=
  if (string-length($IdentificationNumberCvrOfReportingEntity) <= 8)
    then $IdentificationNumberCvrOfReportingEntity
    else (
      let $input := string($LEI)
      let $batchScript := "E:\National Asset Project 2023\07 - Generate Regnskaber\get_cvr_from_lei.bat"
      (: Construct the command as a single string :)
      let $command := string-join(($batchScript, $input), ' ')
      (: Run the batch script using the proc:system function :)
      let $result := proc:system($command)
      (: Save the output as a variable or process it as needed :)
      return fn:normalize-space($result)
    )
)
```

Screenshot fra Generate Regnskaber v1.7.xq

Queryen kalder først batchfilen **get_cvr_from_lei.bat**, som derefter eksekverer Python-scriptet **get_cvr_from_lei.py**.

Vejledning til Python-scriptet

For detaljeret information om, hvordan Python-scriptet fungerer, henvises til den separate vejledning:

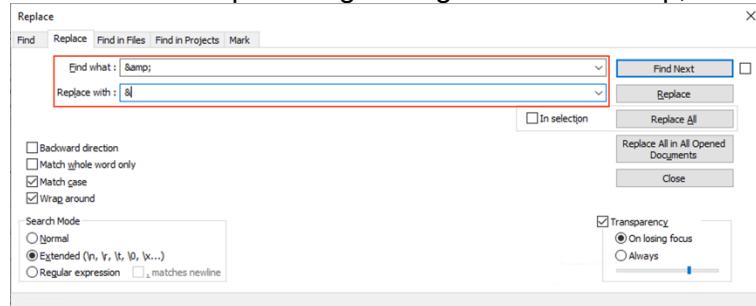
'Vejledning til brug af det eksterne Python-script.docx'

5. Output fil kan findes:

Remote Desktop: E:\National Asset Project 2023\07 - Generate Regnskaber\Regnskaber med navn:

<DATO>-Regnskab<ÅR>.txt

Åben txt fil i Notepad++ og find og erstat alle '&' til '&



6. Gentag for alle år

08 - Upload to Azure DB

Remote Desktop: E:\National Asset Project 2023\08 - Upload to Azure DB

1. Åben SQL Server Management Studio Management Studio 20

2. Log ind på Azure Databasen

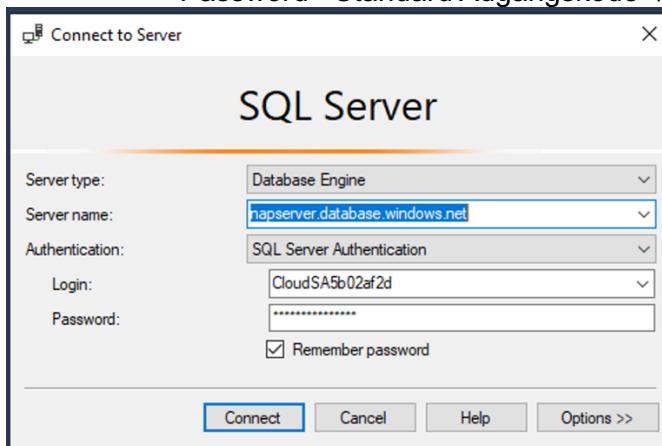
 Server type: Database Engine

 Server name: napserver.database.windows.net

 Authentication: SQL Server Authentication

 Login: CloudSA5b02af2d

 Password <Standard Adgangskode>#1



3. Klik "Connect"

4. Udvid: Databases

5. Højre klik på NationalAssetProject -> Tasks -> Import flat file

6. Find enten "CVR-PermanentYYYY", "OffentliggoerelseYYYY" eller "RegnskabYYYY" (YYYY er år)

7. New table name: fjern dato fra navn

(f.eks. <YYYY-MM-DD>-CVR-Permanent-<ÅR> -> CVR-Permanent-<ÅR>)

8. Table Schema: dbo

9. Modify Columns:

'Data Type' varierer af kolonnens størrelse og type. Undgå at bruge nvarchar(MAX).

PROBLEM: Der kan være et potentielt problem med at have mange af de ens felter fra forskellige tabeller i forskellige formater såsom CVR. Dette kan f.eks. ændres ved følgende query:

ALTER TABLE dbo.Regnskab

ALTER COLUMN CVR INT;

Dog skal man være sikker på at alle felter kan formateres til valge format ellers vil der blive slettet rækker fra databasen.

10. Efter at have indsat ønsket år kombineres alle årene til en enkelt tabel, via følgende query: (hvis alle årene allerede er i samme tabel og der ønskes at tilføje et ekstra år skal samlet tabel kombineres med det enkelte år):

Eksempel på Offentliggoerelse<ÅR> kombination:

```
-- Create the new table 'Offentligoerelser' if it doesn't exist
IF NOT EXISTS (SELECT * FROM INFORMATION_SCHEMA.TABLES WHERE TABLE_NAME = 'Offentligoerelser')
BEGIN
    CREATE TABLE Offentligoerelser (
        cvrNummer INT,
        filnavn_xml NVARCHAR(200),
        url_xml NVARCHAR(200),
        filnavn_pdf NVARCHAR(200),
        url_pdf NVARCHAR(100),
        slutDato DATE,
        [year] INT,
        PRIMARY KEY (cvrNummer, slutDato) -- Define composite primary key
    );
END;

-- Insert data from each source table into the target table
INSERT INTO Offentligoerelser (cvrNummer, filnavn_xml, url_xml, filnavn_pdf, url_pdf, slutDato, [year])
SELECT cvrNummer, filnavn_xml, url_xml, filnavn_pdf, url_pdf, slutDato, [year]
FROM Offentliggoerelse2012
UNION ALL
SELECT cvrNummer, filnavn_xml, url_xml, filnavn_pdf, url_pdf, slutDato, [year]
FROM Offentliggoerelse<ÅR>
UNION ALL
```

Hvor <ÅR> er det ønskede år

09 - Power BI

10 - Supplerende Data

Der findes offentlige regnskaber som ikke kan hentes via Virk. Det gælder følgende sektorer:

- 10.1 'Only PDF' regnskaber
- 10.2 Kommuner & Regioner
- 10.3 Statslige virksomheder

10.1 'OnlyPDF' regnskaber hentes fra Proff via webscraping script.

'OnlyPDF' regnskaber er de regnskaber som kun har en PDF tilgængelig og derfor er deres regnskaber ikke hentet via tidligere metoder.

Hjemmesiden Proff.dk har indskrevet meget PDF data på deres hjemmeside hvilket betyder at vi kan webscrape data'en fra deres side via et CVR nummer.

1. Kørsel af webscraping-script

1. **Hent:** CSV-numre med manglende data fra databasen.
(Find '**Find-CVR-for-OnlyPDFs.sql**' query enten fra Remote Desktop (*Remote Desktop: E:\National Asset Project 2023\10 - Supplerende Regnskaber*)
eller DriveHQ (*\Reliasset\Projekter\Nationale aktiver>Status 2024*))
2. **Input:** Opret en CSV-fil med CVR-numrene (f.eks. OnlyPDFsCVR.csv).'

A	B
1	10003210
2	10003466
3	10004055
4	10004500
5	10005558
6	10006422
7	10007151
8	10007267
9	10010179
10	10012848
11	10013321
12	10014522
13	10016665
14	10016762
15	10019923
16	10020166
17	10020328
18	10021464
19	10022851
20	10026156
21	10026512

(CVR numre indsættes i første kolonne uden header)

3. **Kør scriptet:** Åben terminal i korrekt mappe og brug følgende kommando:
`python Get-data-from-Proff-Upgraded-v0.4 OnlyPDFsCVR.csv outputfilnavn.csv`

Bemærk: Detaljeret vejledning til brug og struktur af script kan findes i følgende mappe:

DRIVEHQ: /Nationale aktiver/Status 2024/Vejledning-getproffdata.pdf

Da scriptet er semi-statisk opbygget, kan det stoppe med at fungere korrekt, hvis strukturen på Proff.dk ændres. I sådanne tilfælde skal koden opdateres, så den matcher den nye HTML-struktur.

4. **Output:** Scriptet genererer en CSV-fil med regnskabsdata.

2. Rensning af data

- Åbn den genererede CSV-fil i Excel.
- Kontrollér, at data er korrekt formateret, og udfyld evt. manglende værdier.
- Gem den rensede fil, som skal bruges til upload.

3. Upload til databasen

1. Åbn **SQL Server Management Studio** og forbind til Azure-databasen.
2. Importér CSV-filen:
 - o Højreklik på databasen > Tasks > Import Flat File.
 - o Vælg combined_financial_data.csv.
3. Kontrollér, at data er uploadet korrekt, og at der ikke er fejl. Eventuelt vend tilbage til 08 - *Upload to Azure DB* for at se hvordan man kombinere tabeller i databasen.

10.2 Kommuner, Regioner hentes manuelt via Danmarks statistik.

1. Dataindsamling for Kommuner og Regioner

For at hente data om kommuner og regioner bruges følgende tabeller fra **Danmarks Statistik**:

- **REGK4:** Kommunernes balancer efter funktion, område og tid.
- **REGK11:** Kommunernes regnskaber på hovedkonti efter prisenhed, hovedkonto, tid, område og art.

Fremgangsmåde:

1. Gå til [Danmarks Statistik](#).
2. Find og download de relevante tabeller (**REGK4** og **REGK11**) i Excel- eller CSV-format.
3. Gem filerne i følgende mappe:

Remote Desktop: E:\National Asset Project 2023\10 - Supplerende Regnskaber\Kommuner og Regioner data

2. Databehandling

Efter dataindsamling skal der udføres rensning og beregninger for at gøre data klar til analyse.

Rensning:

1. Åbn de hentede filer i Excel.
2. Kontrollér, at alle kolonner er korrekt navngivet og struktureret.
3. Fjern unødvendige kolonner eller værdier, der ikke er relevante for analysen.

Beregninger:

1. Udfør nødvendige beregninger for at summere, om kategorisere eller omregne data efter projektets behov.

3. Upload til databasen

Når data er renset og klar:

1. Gem de behandlede filer i en konsolideret CSV-fil.
2. Følg processen beskrevet i trin **3. Upload til databasen** fra tidligere sektion.

10.3 Data indsamling af statslige virksomheder

1. Hentning af data

Data for statslige virksomheder findes på følgende side:

<https://oes-cs.dk/aarsrapporter/>

Fremgangsmåde:

1. Besøg hjemmesiden ovenfor.
2. Vælg det ønskede år i menuen og download årsrapporterne for alle relevante statslige virksomheder som PDF-filer.
3. Gem de downloadede filer i følgende mappe, afhængigt af året:
*Remote Desktop: E:\National Asset Project 2023\10 - Supplerende
Regnskaber\Ministerie data\<ÅR>*

2. Databehandling

Efter at have hentet årsrapporterne, skal data udtrækkes og behandles manuelt.