Instructions for the use of run_script.exe for web scraping charities with NDIS registration.

By Jonathan Markey | ABN: 77810429163

Email: jonmarkey1232@gmail.com

STEP 1 – DOWNLOADS:

- To download the bot navigate to the GitHub link here: https://github.com/Jmarkey11/Unprofitable-NDIS-Charity-Web-Scraping-And-Analysis and download the dist.rar file which contains all the necessary files. Once extracted the folder contains several files as listed below
- Your anti-virus software may flag the run_scripts.exe file as malware when downloading or extracting. This is a false positive. You may be able to restore the file through your anti-virus software.

bot.py	17/03/2025 12:25 PM	Python Source File	18 KB
generate_results.py	17/03/2025 2:41 AM	Python Source File	33 KB
Reg_Prvdrs_15_March_2025.csv	16/03/2025 1:59 AM	Microsoft Excel C...	22,053 KB
requirements.txt	16/03/2025 3:51 PM	Text Document	1 KB
run_scripts.exe	17/03/2025 12:29 PM	Application	6,916 KB
run_scripts.py	17/03/2025 12:27 PM	Python Source File	2 KB
setup.py	16/03/2025 4:20 PM	Python Source File	1 KB

- The **Reg_Prvdrs_15_March_2025.csv** can be updated to a more current list of NDIS providers by downloading the .csv file from the NDIS registered provider website and replacing this file to a more current version (ensure only a **single Reg_Prvdrs.exe file** exists in the folder): https://www.ndis.gov.au/participants/working-providers/find-registered-provider/provider-finder
- **Do not remove** any of the files out of the folder as this can result in the .exe file not running.

STEP 2 – RUNNING THE EXE FILE:

- To run the web scaping right click on the **run_scripts.exe** file and click on **run as administrator**. This will open a black window called the terminal and will begin executing the code. In the terminal text will appear showing the progress of the bot. **Do not close the terminal window** while running the file.
- While running, text will flash up in the terminal window. This is normal and outlines the progress of the bot. The process can take some time to complete depending on the speed of your hardware and internet. As such it may appear that the terminal is frozen but it is just processing. Your computer may slow down when the file is running.
- When the .exe file is complete the terminal will automatically close and several new files will be in the folder.

STEP 3 – LOCATING YOUR FILES:

- The .exe file outputs 4 new files into your folder.

🛡 financial_report.html	17/03/2025 5:34 PM	Brave HTML Docu...	729 KB
📊 NDIS_CHARITIES_UNPROFITABLE_COND...	17/03/2025 5:34 PM	Microsoft Excel C...	435 KB
📊 NDIS_CHARITIES.csv	17/03/2025 5:33 PM	Microsoft Excel C...	2,121 KB
📁 charity_reports	17/03/2025 5:30 PM	File folder	

- The first is the **NDIS_CHARITY.csv** file which contains the output from the full web scraping for all the charities.
- The **NDIS_CHARITY_UNPROFITABLE_CONDENSED.csv** file is a more condensed version of the previous file and contains the information on the charities that operate at < 4% surplus.
- The **financial_report.html** file is a report that can be opened in your browser. This file contains a more interactable version of the previous .csv file. For instructions on the use refer to the next step.
- The **charity_reports** folder contains all the individual automatically generated charity reports that can be accessed using the previous .html file. For the financial_report.html file to display these reports it requires that these two files always remain together in the same folder. Do not put the financial_report.html file in the charity reports folder.

STEP 4 – HOW TO USE FINANCIAL_REPORT.HTML:

- The financial_report.html file is an automatically generated file and provides an overview of the results of the web scraping. The file contains two main sections.
- The first section of the file contains an executive summary that provides a short summary. The Key findings provides a count of the number of charities that are operating at <= 4% surplus of total revenue to expenses. This is then further broken down into counts by state and by revenue classification of small, medium and large.
- The second section of the report is the interactable table and window for the charities. The table contains the information from the NDIS_CHARITY_UNPROFITABLE_CONDENSED.csv file but with some actions. You can **filter** the table using the **Filter by State** and **Filter by** Classification drop down boxes. You can **Search** for information in any of the columns using the search bar on the top right of the table. The columns can be **sorted** by clicking on the up and down arrows next to their titles.
- For each charity an individual company report has been automatically generated and can be viewed by clicking on their named **link** under the **Charity_Legal_Name** column in the table. This will open the report in the window below the table. Each report contains information relating to the charity. Links to their ACNC website and most recent financial reports can be accessed by clicking on the links in the **ACNC Profile:** and **Latest Financial Report:** sections. A table listing who the charity helps is also available.
- A financial breakdown is available which firstly shows two interactable pie charts. Hover over the sections to see what part they correspond to. Below this is the income statement and balance sheet.
- Lastly a map is available showing the location of the where the charity is headquartered. You can zoom in and out using the +/- buttons or the middle mouse button.
- Note: As these reports are automatically generated some links may not work and some of the information may be missing from sections or incorrect.

Potential problems:

- Files removed from the dist folder will result in the run_scripts.exe file from not running.
- The web scraping bot uses chrome to run. Ensure that It is installed on your pc before executing the .exe file.
- Running the .exe file multiple times in a single day may overwhelm the database connection or cause a rate limit error to occur. This can be resolved by waiting and retrying 24 hours later. If this error persists the database connection may be paused and requires restarting. Contact the email for this to be resolved.
- If the .exe file is unable to run, you can execute the files manually using your terminal and Python. The required libraries to run the .py files are in the requirements.txt file and can be automatically installed into your Python environment using setup.py. Following this you can execute the run_scripts.py file using the command "python run_scripts.py".
- This program has been tested only using a Windows 11 system. The file may not be able to execute in alternative operating systems such as linux or macOS.

Thank you very much for using this program. For any further information please contact me using the email address at the beginning of this file.