

LIST EXTRACTION INSTRUCTIONS

01

File name

- Ensure the file name **EXACTLY** matches the file number

02

File type

- Ensure the file is saved in **CSV in UTF-8 format**

03

File accessibility

- Ensure you have the necessary **permissions** to access the file.
- **Verify** that the file is not corrupted or encrypted.

04

Follow file instructions

- **Read and understand** the instructions or guidelines provided for the file.
- **Clarify** any ambiguous instructions by messaging Arvin though Upwork

05

Header

- Include column **headings** in the first row of the file, such as Name, Company Name, etc., which must match the list from the url. The goal is to recreate the exact list from the URL in the file.

06

Columns

- **Collect all fields** the from the list on the url provided, all data for each individual or company from the list must be extracted, including hyperlinks.

SWIPE NEXT

LIST EXTRACTION INSTRUCTIONS

07

Rows

- **Collect all rows** of data from the url, there are no exceptions. All data for each individual or company must be in one row, and the total number of records you extract must match the list on the date you collect the data. Contact Arvin if you are unable to capture all records/rows.

08

Hyperlinks

- **Add extra** columns to capture hyperlinks. If a word is hyperlinked, extract both the word and the corresponding URL, using two columns.

09

Code

- You must **send us the python code** you used to scrape the data.
- Zip the codes when you submit them.
- Ensure the name **EXACTY** matches the file number.

10

Count lines in CSV file

- Report the **number of lines** or data you scraped by using this tracker:
<https://docs.google.com/spreadsheets/d/1raMaZKFaY-3mvcFFFxkqomBioHs30br3FUzhYZvBTh0/edit#gid=0>
- Make sure to indicate the **date** when you scraped the data.

FILES SUBMISSION INSTRUCTIONS

01

Uploading Link

- Once finished, upload files to this site: <https://amazing.sunset.photos.click/>
- You will need to create an account to use the link

02

Bulk Submissions

Zip all CSV files together and zip all Python code together separately. If you are not using Python, you don't have to send us the code. Do not create a folder for each dataset. Do not create any subfolders.

03

CSV and PY CODES ONLY

Ensure that there are no other files aside from the CSV and Python code files.

WHEN YOU ENCOUNTERED AN ERROR/ISSUE

01

If you are dealing with large datasets

- Do not separate the data into different sheet tabs. Put all the data under only one tab.
- We understand that it may take some time to complete, but we expect you to work on other smaller datasets simultaneously.

02

If you encounter datasets that are challenging to scrape or if you encounter an error

- Use this tracker:
<https://docs.google.com/spreadsheets/d/1raMaZKFaY-3mvcFFFxkqomBioHs30br3FUzhYZvBTh0/edit#gid=0>

03

CSV and PY CODES ONLY

Ensure that there are no other files aside from the CSV and Python code files.