**Scripts**

**project\_data\_extractor.py** is made to be much more adaptable to new changes. Including switching to be run on either single-process or multi processes, getting rows of unique urls from the csv file, handling connection errors by waiting for connection to be stable before extracting, handling captcha when detected, having correct data inputs, resumes previously last scraped row etc  
**project\_data\_extractor(chhabra).py** prints less of the errors to screen. Mainly for firing runs.

**db\_to\_csv.py** converts the database output into a csv file

**Project\_url\_extractor.py(csv)** is designed to extract URLs from CSV files found on the webrobots [website](https://webrobots.io/kickstarter-datasets/). It has been modified to get all information about the campaigns from URLs alone and can be configured to use either single or multiple processes to handle each URL.

**Project\_url\_extractor.py(json)**  is designed to extract URLs from JSON files found on the webrobots website. The script has been modified to extract all information about campaigns only from URLs that are live, successful, or failed (i.e., already launched). It is capable of using either single or multiple processes to handle each URL, improving efficiency. Additionally, the script retrieves the HTML content (soup) only once per URL to reduce the likelihood of detection and efficiently gathers all necessary information from the single soup extraction.