

DiscountCrawler

A Crawler for fetching discount data

Requirement

- python 2.7
 - scrapy 0.24
 - scrapyd
-

Usage

1. Make sure you have successfully installed `python2.7` , `scrapy` and `scrapyd`
2. In your terminal, use the command `scrapy crawl <spider_name>` to crawl data
3. Available spider_name: `jd` (for data on `jd.com`), `smzdm` (for data on `smzdm.com`)
4. To ensure the data is up-to-date, you may use the `crontab` command in your Unix-like system to run the command periodically.

For example, you can create a file named `period_task` , and write `0 */2 * * * scrapy crawl jd` in it, then type `sudo crontab period_task` in your terminal to activate it. You can refer to the manual page of `crontab` to get more details.

5. The crawled data will be created under the project folder as a json file.