

專案名稱

104網站爬蟲並匹配相關工作技能及同義詞

目的

依關鍵字爬取104網站所有職缺內容，並提供客製化輸入相關技能及同義詞，結果以Excel逐筆呈現，可查看相關技能匹配結果(1=有，0=無)。

操作說明

※ 以下使用步驟:

1. 執行 exe.py 前安裝所需 requirements.txt 或執行 dist/exe.exe，執行成功開啟網頁 <http://127.0.0.1:8880>
2. 於網頁輸入職缺關鍵字、選擇工作地點(最多三個，可不選)、輸入爬蟲的網頁數，上傳已編輯好的技能同義詞字典synonym.json(僅接受json檔，若無可跳過)
3. 請參考synonym.json格式，synonym_dic中填寫欲搜尋的技能及其同義詞 (例如AI同義詞有artificial intelligence=人工智慧，則寫成"AI": ["artificial intelligence","人工智慧"])，英文不分大小寫，不同技能需換行輸入。
4. 程式出現 Processes all done. 代表資料抓取完畢。輸出檔案於 output資料夾: output_104.csv、output_104.xlsx，並於網頁呈現結果。
5. 爬取結果；包含公司名稱、職缺名稱、工作地點、待遇、工作內容、工作經驗、學歷要求、福利、各技能比對(1=有，0=無)、技能配對總數(match_skills)等詳細資料。

網頁輸入範例

搜尋104職缺

請輸入關鍵字:

數據分析

請選擇工作地點(最多三個):

台北市, 新北市, 桃園市

請輸入搜尋頁數(1-50):

2

技能同義詞json檔案上傳:

選擇檔案

synonym.json

搜尋

1.搜尋技能同義詞範例

```
{
  "synonym_dic": {
    "Python": ["python"],
    "Linux": ["linux","unix"],
    "MySQL": ["mysql","sql","sql server","關聯式資料庫"],
    "MongoDB": ["mongodb","非關聯式資料庫"],
    "R": ["r","r language","r語言"],
    "ETL": ["etl","elt","爬蟲"],
    "Data Analysis": ["data analysis","data mining","datamining","ab testing","a/b testing","資料分析","數據分析","大數據","大數據分析","資料探勘","資料工程","資料科學","智慧製造"],
    "Machine Learning": ["machine learning","deep learning","機器學習","深度學習","人臉辨識","臉孔辨識","影像辨識","圖形辨識","圖片辨識"],
    "視覺化": ["Tableau","powerbi","power bi","matplotlib","視覺化"],
    "AI": ["ai","artificial intelligence","人工智能","人工智慧"],
    "Cloud": ["cloud","aws","gcp","azure","cloud service","cloud server","雲端服務","雲端運算","雲服務"],
    "Kafka": ["kafka"],
    "Docker": ["docker","容器化"]
  }
}
```

網頁輸出範例

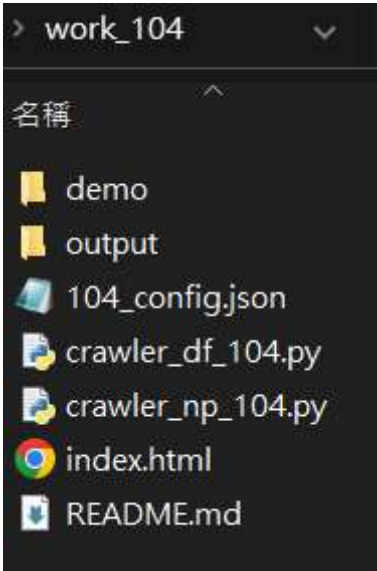
Output Table

Downloaded the output file at output/output_104.xlsx & output_104.csv

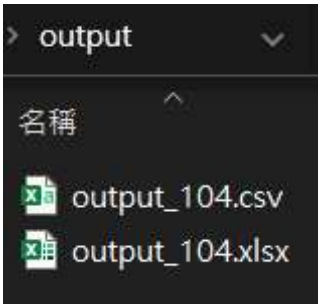
company	job_name	job_area	job_salary	job_content	job_exp	job_require_major	job_welfare	job_contact	URL	Python	Linux	MySQL	MongoDB	R
麥絲瑪拉台灣零售股份有限公司	CRM 數據分析主任副理	台北市中山區	待遇面議	- 會員資料庫統計及進行成效分析及提出優化建議 (會員活動成效、行為分析)- 監控CRM數據，瞭解使用者需求，運用會員資料庫統計數據分析，探索消費者旅程，擬定CRM分眾行銷策略，進行成效追蹤與自動化再行銷- 品牌會員行銷或相關專案之規劃與執行 (滿額禮、點數、滿意度調查、異業合作洽...等)- 彙整公司各部門意見、優化會員活動或經營機制- 協助營運分析	工作經驗5年以上		1.專職同仁排變形工時排班，依假期天數依辦法令規定 2.每月/每季依權點績效發放業績獎金及業績達成獎金 3.優於勞基法之年休假制度 4.完整休假制度(給薪病假/婚假/陪产假/家庭照顧假/女性同仁生理假等) 5.年終獎金/開工紅包 6.年終尾牙餐會與摸彩 7.團體保險(壽險/重大疾病險/意外醫療險/癌症醫療保險/農業災害保險) 8.年度免費健康檢查。(與國	聯絡人 HR	https://www.104.com.tw/job/7wpm5	0	0	0	0	0

檔案輸出範例

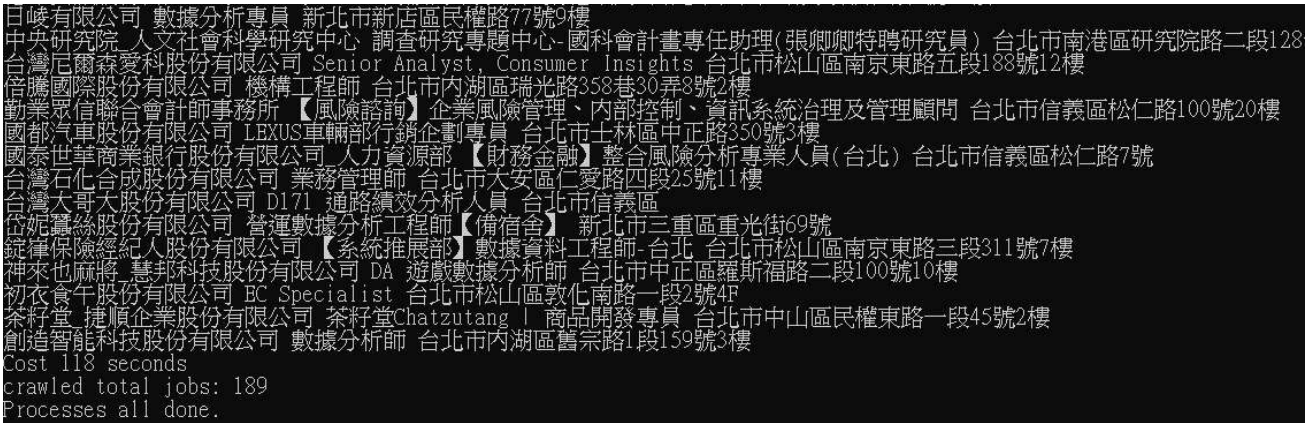
1.資料夾畫面



2.output資料夾畫面



3.程式執行畫面



4.Excel畫面

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	company	job_name	job_area	job_salary	job_content	job_exp	job_require_major	job_welfare	ob_contac	URL	Python	Linux	MySQL	MongoDB
2	香港商蒙卓有限公司台灣	數據分析師	台北市松	月薪35,000	2.有遊戲產業經驗	工作經驗:3年以上			聯絡人:De	https://www	1	0	1	0
3	杏豐實業股份有限公司	大數據分析	新北市	三待遇面議		工作經驗:1年以上		7.定期福委會	聯絡人:陳	https://www	1	0	0	0
4	網路中文資訊股份有限公司	數據分析師	台北市中	月薪30,000	工作內容	工作經驗:不拘			聯絡人:劉	https://www	1	0	1	0
5	台灣壽司郎股份有限公司	【日商 壽司郎】台中北屯松	台中市北	時薪190~22	計畫拓展全台灣	工作經驗:不拘			聯絡人:人	https://www	0	0	0	0
6	麥絲瑪拉台灣零售股份有	CRM 數據分析主任/副理	台北市中	待遇面議		工作經驗:5年以上			聯絡人:HF	https://www	0	0	0	0
7	昇恒昌股份有限公司	數據分析師(內湖總公司)	台北市內	月薪40,000	3.運用機器學習、	工作經驗:2年以上	一般商業學類、資		聯絡人:En	https://www	1	0	1	0
8	傲群科技有限公司	數據分析師	台中市西	月薪50,000		工作經驗:2年以上		全勤獎金、	聯絡人:楊	https://www	0	0	1	0
9	遊戲橘子數位科技股份有限公司	數據分析師_0167	台北市內	待遇面議	3.透過資料觀察推	工作經驗:3年以上	應用數學相關、數	更多Gamania	聯絡人:HF	https://www	0	0	1	0
10	東昇休閒有限公司	數據分析師	台中市北	月薪40,000		工作經驗:不拘	統計學相關、數理	目前公司健康	聯絡人:林	https://www	1	0	1	0
11	AIR SPACE_新地國際股份有	商品數據分析	台北市士	待遇面議	7.熟悉Excel函數或	工作經驗:2年以上			聯絡人:HF	https://www	0	0	0	0
12	影響力數據顧問股份有限	【Quickseek】數據分析實習	台北市松	時薪176~18	5.每週必須可排班	工作經驗:不拘	政治學相關、社會	勞保、健保、	聯絡人:To	https://www	0	0	0	0
13	日峻有限公司	數據分析專員	新北市新	待遇面議	公司介紹 https://a	工作經驗:不拘		4.年終獎金/	聯絡人:吳	https://www	0	0	1	0
14	群益金鼎證券股份有限公	【資訊】大數據分析人員	台北市松	待遇面議	4.例行資料整理及	工作經驗:2年以上	統計學相關、數理	03.優惠電影	聯絡人:粘	https://www	1	0	1	0
15	全曜財經資訊股份有限公	營運數據分析師	新北市板	待遇面議		工作經驗:不拘	其他商業及管理相	官方網站:h	聯絡人:HF	https://www	1	0	1	0
16	傳奇網路遊戲股份有限公	遊戲數據分析師	台北市信	月薪40,000	4.針對進行中的各	工作經驗:1年以上		兼職、工讀、	聯絡人:劉	https://www	1	1	0	0
17	新澄企業有限公司	數據分析師	台中市西	月薪50,000	具備GA / GA4證照	工作經驗:不拘		★飲品點心	聯絡人:張	https://www	1	0	1	0
18	英屬開曼群島商雲數位智	數據分析師 - Google Analytics	台北市松	月薪38,000		工作經驗:2年以上		12.運動健身、	聯絡人:HF	https://www	1	0	1	0
19	財團法人中衛發展中心	G8-長照及健康數據分析應用	台北市中	月薪35,000		工作經驗:2年以上	公共衛生相關、醫	4.陪病假3	聯絡人:HF	https://www	0	0	0	0
20	松崗數位文創股份有限公	數據分析產品規劃師	台北市中	待遇面議		工作經驗:2年以上	資訊工程相關、數	結婚、生子	聯絡人:歐	https://www	1	0	0	0
21	新普科技股份有限公司	儲能數據分析工程師	新竹縣湖	待遇面議		工作經驗:3年以上	應用數學相關、統	(駐外職缺的	聯絡人:人	https://www	1	0	1	1