資料源: 內政部不動產時價登錄網 (http://plvr.land.moi.gov.tw/DownloadOpenData)

程式語言: Python

- 題目: 1. 撰寫網路爬蟲程式, 抓取【內政部不動產時價登錄網】中位於【臺北市/新北市/桃園市/臺中市/高雄市】的【不動產買賣】資料。請選擇【非本期下載】【CSV格式】, 【資料內容: 發布日期 108 年第 2 季】。
 - 2. 使用檔名【 a_lvr_land_a 】【 b_lvr_land_a 】【 e_lvr_land_a 】【 f_lvr_land_a 】 【 h_lvr_land_a 】五份資料集。
 - 3. 【設計/建立 RESTful API】供查詢資料及的下列資訊: 【以 JSON 格式回傳,請自訂 Schema】
 - 【鄉鎮市區】
 - 輸入【阿拉伯數字】查詢【總樓層數】
 - 【建物型態】
 - 4. 使用 Spark 合併資料集,以下列條件從篩選出結果
 - -【主要用途】為【住家用】
 - -【建物型態】為【住宅大樓】
 - -【總樓層數】需【大於等於十三層】
 - 5. 使用 Spark 將步驟 4 的篩選結果,轉換成 Json 格式,產生【 result-part1.json 】和【 result-part2.json 】(五個城市隨機放入兩個檔案),格式如下:
 - time_slots 需根據 date 作 desc 排序(越近期的放前面)
 - 每行各自是valid Json (結果範例如下頁)



結果範例

```
("city": "新北市", "time_slots": [("date": "2018-12-22", "events": [("type": "其他", "district": "清水區")]), ("date": "2018-12-21", "events": [("type": "其他", "district": "清水區")]), ("date": "2018-12-24", "events": [("type": "其他", "district": "龜山區")]), ("date": "2018-12-24", "events": [("type": "公寓(5樓含以下)
("city": "臺中市", "time_slots": [("date": "2018-12-24", "events": [("type": "其他", "district": "瑞芳區"), ("type": "住宅大樓(11層含以上有電梯)", "district": "新店區")]
("city": "台北市", "time_slots": [("date": "2018-12-25", "events": [("type": "公寓(5樓含以下無電梯)", "district": "信義區")]), ("date": "2018-12-22", "events": [("type": "株園市", "time_slots": [("date": "2018-12-25", "events": [("type": "其他", "district": "小港區")]), ("date": "2018-12-24", "events": [("type": "賽房(1房1廳1衛)
```

繳交內容與方式

- 請以簡報方式呈現以下內容:
 - 1. 實作處理過程說明
 - 2. 過程中遇到的問題與解決方法
- 繳交**包含註解的.py 程式碼**至個人 github, 並提供連結。