

第一屆 Python 解難大賽題目

注意事項：

1. 請使用 `mock_server.bat` 啟動伺服器，然後可以通過 `http://localhost:8000` 訪問各網站。
2. 所有題目各自有其資料夾。而每個資料夾內，有已建立但未完成的 `python` 檔案。請不要更改這些檔案結構及名稱。
3. 另外，每個資料夾內均有一個 `inputs` 及一個 `outputs` 資料夾。輸入的資料均放在題目的 `inputs` 資料夾，而輸出的檔案則放在 `outputs` 資料夾內。
4. 評分只會在每隊的一部電腦上進行，另一部電腦不會用作任何評分。
5. 所有 `YYYY-MM-DD` 為日期的表述，其中月份及日子如為單數，請補 `0`，即三月為 `03`。
6. 所有輸出的圖表均需要有標題、圖例、`X` 軸標題、`Y` 軸標題，如題目沒有要求，請自行撰寫。
7. 所有代碼請適當注釋，以保持可讀性及易維護性。
8. 評審時 `localhost:8000` 上的動態資料將有所替換，但已說明為靜態的內容不會改變。

題目：七日天氣

輸入：天文台網址，內含多個 XML 檔案。

輸出：未來七日天氣純文字檔

描述：

現有一個天文台網站，網址為：

http://localhost:8000/weather/forecast_7_days.xml 及

http://localhost:8000/weather/weather_status.html

當中：

- `forecast_7_days.xml` 為動態七日天氣
- `weather_status.html` 包括天氣狀態對應表，這個狀態表不會改變。

請修改 `macao_weather.py` 檔案，並從網站中，下載相關 XML 動態信息，並以以下格式輸出至純文字檔，命名為 `forecast-YYYY-MM-DD.txt`。檔案名中的 `YYYY-MM-DD` 為今日日期。

檔案內應有七行，每行格式如下：

YYYY-MM-DD (WEATHER) LOWEST 至 HIGHEST 度

例如：

2021-10-01 (天晴) 23 至 28 度

2021-10-02 (天晴) 23 至 28 度

2021-10-03 (天晴) 23 至 28 度

2021-10-04 (天晴) 23 至 28 度

2021-10-05 (天晴) 23 至 28 度

2021-10-06 (天晴) 23 至 28 度

2021-10-07 (天晴) 23 至 28 度

題目：匯率

輸入：包括匯率的網站

輸出：一個寫好的函數，放於 `outputs/macao_rates.py` 檔案內。

描述：

現時本澳居民常於大灣區生活及消費。現希望取得一個銀行網站所發佈的最新人民幣兌澳門元匯價。

已知網站位置為: `http://localhost:8000/rates/rates.html`

考慮到獲取匯價是一個在不同程序及場景中都會用到的代碼。所以本次我們希望你可以把代碼製作成一個供匯入的檔案。

作為測試，我們已製作一個名為 `test_rates.py` 的測試用代碼，以匯入及調用你的檔案的代碼。

`test_rates.py` 檔案內容

```
from macao_rates import latest_cny_mop
rate = latest_cny_mop()
print(100*rate)
```

請修改 `output` 資料夾內的 `macao_rates.py`，當中必須有一個命為 `latest_cny_mop` 的函數，回傳值為一個浮點數 (`float`)。

注意：這個銀行網站的匯價列表，不一定全部記錄每次齊全，有時會有一些外幣是從缺的，但人民幣兌澳門元則從不從缺。

評審

先替換 `rates.html`

有提交檔案 成功取得資料並正確。

題目：幸運博彩毛收入

輸入：網址

輸出：

1. 11 個 XML 檔案
2. 一個名為 `dicj_summary.txt` 檔案
3. 一個名為 `dicj_average.xlsx` 檔案

描述：

博彩監察協調局每月都會公佈每月幸運博彩毛收入，其網址為：

- http://localhost:8000/dicj/2021/report_cn.xml
- http://localhost:8000/dicj/2020/report_cn.xml
- http://localhost:8000/dicj/2019/report_cn.xml
- http://localhost:8000/dicj/2018/report_cn.xml
- http://localhost:8000/dicj/2017/report_cn.xml
- http://localhost:8000/dicj/2016/report_cn.xml
- http://localhost:8000/dicj/2015/report_cn.xml
- http://localhost:8000/dicj/2014/report_cn.xml
- http://localhost:8000/dicj/2013/report_cn.xml
- http://localhost:8000/dicj/2012/report_cn.xml
- http://localhost:8000/dicj/2011/report_cn.xml
- http://localhost:8000/dicj/2010/report_cn.xml

11 個 XML 檔案備份

現在希望下載上述的 11 個 XML 檔案作備份，請修改 `download_report.py`，把上述網址下載成為以 `report_cn_YYYY.xml` 格式命名的檔案，並儲存到 `outputs` 資料夾。

`dicj_summary.txt`

請修改 `dicj_summary.py` 檔案，以找出 2019 年至 2021 年中，單月幸運博彩毛收入的最低值、最高值，及發生的年份月份，並生成為 `dicj_summary.txt` 的檔案。

此 TXT 檔案要求格式如下，最低值：數字，於 X 年 Y 月發生。最高值：數字，於 X 年 Y 月發生。

例子：

最低值：12345，於 2010 年 4 月發生。

最高值：54321，於 2011 年 5 月發生。

dicj_average.xlsx

另外，我們亦想統計每年的平均值。請修改 `dicj_average.py` 檔案，找出每年的平均單月博彩毛收入，精準之個位數。

注：2021 年取截至十月的數值作平均數。

並輸出成一個 `dicj_average.xlsx` Excel 檔案。當中有 2 欄 11 列，分別是年份及平均值，共 11 年。沒有標題列。

CONFIDENTIAL

基礎 Markdown 轉 DOCX

輸入：不同的檔案，包括 .txt 及 .md 檔案

輸出：轉換後的 .docx

描述：

我們希望製作一個文字生成 DOCX 的簡單轉換器。在這個版本中，我們不希望實現所有 Markdown 功能。我們只要求能轉換大標題、文字段落、及分頁。

這是一種輕量化純文字格式，例如當一行起始為 # 時，則表示為標題。當一行是 --- 或 ---- 時，則表示分頁符。其他文字每行則為段落。

而段落有個比較特別的規則，就是一個跳行不當為段落，而且一個跳行不起任何作用，即在輸出的 Word 中不會跳行。而兩個跳行的（即有空行的）才計算為段落。

以下為一個示例解說：

標題

段落一的文字，段落一的文字，
仍然是段落一的文字，且在同一行。

段落二開始了。
是段落二的文字。
仍然是段落二的文字。

段落三了。

這是標題，及第二頁開始

第二頁第一段文字。

第二頁第二段文字。

所有輸入的文件均在 inputs 資料夾內。請修改 markdown_to_docx.py 檔案，按輸入資料夾中所有 .txt 文字檔案或 .md 文字檔案，逐個輸出成相同名稱的 .DOCX 檔案，並儲存到 outputs 資料夾內。

例如 sample.txt 輸出為 sample.docx、report.md 輸出為 report.docx 等。

題目：人口數據

輸入：countries.csv

輸出：

1. 全球總人口數：global-population.xlsx
2. 全球總人口數折線圖：global-population.png
3. 各年全球頭十位：top-10-population.xlsx

描述：

現有一人口數據 CSV，名為 countries.csv。這個 CSV 檔案內包括各國分別 12 個年份的人口數目。

12 個年份分別為：

- 1952 年
- 1957 年
- 1962 年
- 1967 年
- 1972 年
- 1977 年
- 1982 年
- 1987 年
- 1992 年
- 1997 年
- 2002 年
- 2007 年

現在希望計算全球總人口數從 1952 年至 2007 年的趨勢折線圖，及各年全球頭十位排名。

請修改 population.py 檔案，並得出以下輸出。

全球總人口數 Excel

請輸出一個 Excel 名為 global-population.xlsx

當中有 12 個年份的人口總數。

格式為 2 欄 12 列，沒有標題列。第一欄為年份，如「1952 年」第二欄為人口數總數。

趨勢折線圖

請輸出一幅名為 `global-population.png` 的折線圖。

X 軸為各年份，Y 軸為人口數，以百萬為單位。Y 軸標題亦請注明以百萬計 折線圖需要為實線及有每點標記。

全球頭十位 Excel

希望計算各年的全球頭十位，並以一個名為 `top-10-population.xlsx` 的 Excel 表輸出，這個 Excel 表內有 13 個分頁 (Sheet)。

第一個分頁是介紹頁 (命名為 `Summary`)，內裡只有一行靜態信息：「這個檔案列出各年全球人口頭十位的國家。」

緊接便是 12 個分頁。每年一個分頁，以年份作為分頁名稱，如「1952 年」等。由遠到近。於其中每個分頁，分別有 2 欄 11 列，第一列為標題列，當中第一欄為「國家」，第二欄為「人口」

然後 10 列則是頭 10 位的國家名稱及人口數字。

題目：圖像檔案分類成 Word 文檔

輸入：多個圖像檔案

輸出：多個 Word DOCX 檔案

描述：

我們的一位兼職設計師為我們的數個活動設計了社交媒體用廣告，每個廣告配有數張有順序的圖片檔案。現在希望分別將這些廣告放到 Word DOCX 檔案裏，作為存檔之用。

由於只作存檔之用，圖片只要正常順序放到 DOCX 便可以。沒有特定尺寸要求，但不能辨認不到圖片。

例如，若有以下圖片檔案，目標是有 Word DOCX 檔，分類包括各圖片檔案。

- christmas-1.png
- christmas-2.jpeg
- xmas-fb-1.jpeg
- xmas-fb-2.png
- lunar-new-year-1.png
- lunar-new-year-2.jpg
- lunar-new-year-3.jpg

這些檔案都按特定名稱命名，例如 `xmas-fb-1` 是項目名稱 `xmas-fb` 的第一張圖等。

請修改 `campaign.py` 檔案，按上述需求，輸出 DOCX 檔案，以項目名稱作為檔案名。例如，上述的輸入檔案將生成以下的輸出檔案。

- christmas.docx
- xmas-fb.docx
- lunar-new-year.docx