DSCP Final Project

Name: 洪維駿

Student ID: 40923205L

1. Target
   I'm Wei-Chun Hung, a geography student. I'm interested in spatial information technology and climatology. In geography text books in junior high and high school, there are many beautiful climate graphs with temperature and precipitation data. Therefore, I decided to write a program which could help users getting the climate data from Central Weather Bureau, and visualizing the data as an average climate graph in month scale.

2. Operation
   First, users have to import the packages and the station data, that is to say, the station data(.csv) and the program file must be in the same path.

匯入函式庫與讀取測站資料

In [1]:
```
import csv
import datetime
import urllib.parse
import requests
from bs4 import BeautifulSoup
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline

files = open("./station.csv", "r", encoding="utf-8-sig")
data = files.readlines()
files.close()
```

Second, users can input the station ID or station name(the station data can be seen in station.csv file) that he or she wants to look up. After inputting, the program will check whether the input is valid, and then get the temperature and precipitation data with web crawler.

Note: Some stations have the same name, to input the station ID is a good solvation

選取測站並爬取資料

In [*]:
```
print("請輸入測站站號或站名：")
station = input()
flag = 0

print("")
stationList = []
for i in data:
    list = i.strip().split(",")
    #print(list[0], list[1], station)

    if station in str(list[0]) or station in str(list[1]):
        stationList = [list[0], list[1], list[2], urllib.parse.quote(urllib.parse.quote(list[1]))]

if len(stationList) != 0:
    tonow = datetime.datetime.now()
    today = [int(tonow.year), int(tonow.month), int(tonow.day)]
    print("選取測站：" + stationList[1])

    output = [["日期", "溫度", "降水量"]]
    for i in range(today[0] - 11, today[0] + 1):
        print("----", i, "----")
        for j in range(1, 13):
            if i == today[0] and j == today[1]:
                break
            if i == today[0] - 11 and j < today[1]:
                continue
            url = "https://e-service.cwb.gov.tw/HistoryDataQuery/MonthDataController.do?command=viewMain&station=" + stationList[
            page = requests.get(url)
            soup = BeautifulSoup(page.text,"html.parser")
            sel = soup.select("div.CSSTableGenerator td")
            for k in range(0, len(sel), 35):
                date = str(i) + "-" + str(j).zfill(2) + "-" + str(sel[k].text)
                print(date, sel[k + 7].text.rstrip(), sel[k + 21].text.rstrip())
                output.append([date, sel[k + 7].text.rstrip(), sel[k + 21].text.rstrip()])

else:
    print("無法讀取，請確認輸入是否有誤")
```

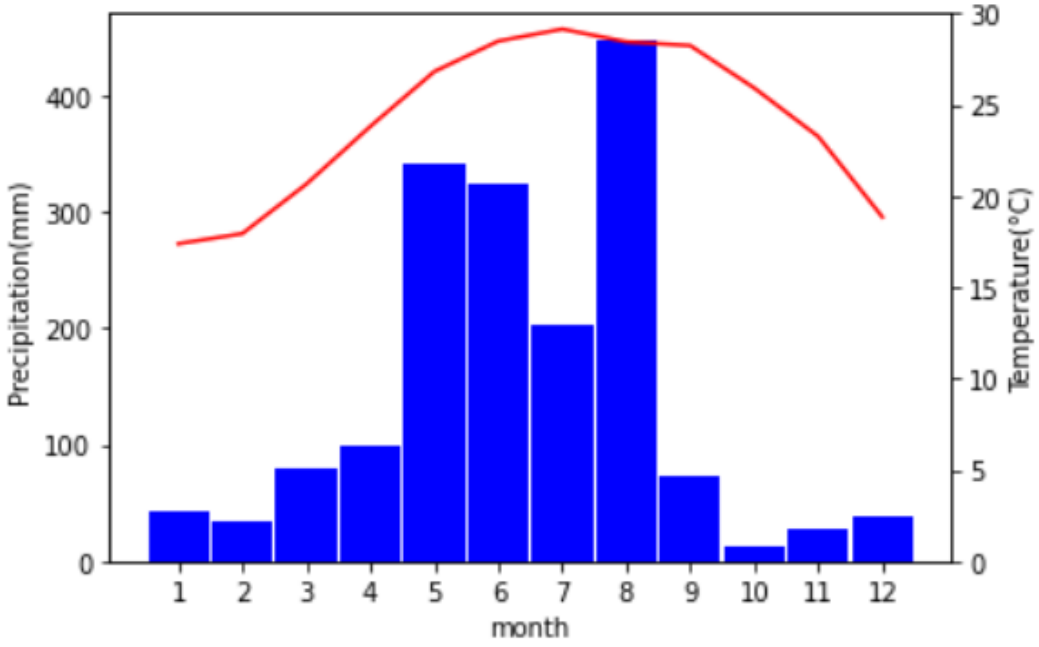
請輸入測站站號或站名：

```
請輸入測站站號或站名：
臺中

選取測站：臺中
---- 2011 ----
2011-12-01 20.1 0.0
2011-12-02 15.9 T
2011-12-03 17.2 0.0
2011-12-04 19.2 0.0
2011-12-05 18.8 30.9
2011-12-06 21.8 0.5
2011-12-07 22.0 0.0
2011-12-08 20.1 0.0
2011-12-09 15.5 0.8
2011-12-10 16.2 0.0
2011-12-11 16.0 2.2
2011-12-12 18.7 0.0
2011-12-13 19.8 0.2
2011-12-14 20.6 1.0
```

After that, users can create a climate graph like the below one.

Note: There is no temperature data in the station ID with C1 header.



In the end, users can choose whether to output the csv file or not. There are two file, 氣候資料 is the every-day data, 逐月平均氣候資料 is the average data each month.

輸出csv

```
with open("./create_data/" + stationList[1] + "氣候資料" + '.csv', 'w', newline='') as csvfile:
    writer = csv.writer(csvfile)
    writer.writerows(output)

with open("./create_data/" + stationList[1] + "逐月平均氣候資料" + '.csv', 'w', newline='') as csvfile:
    writer = csv.writer(csvfile)
    writer.writerow(["月份", "溫度", "降水量"])
    for i in range(1, 13):
        writer.writerow([i, daySumT[i-1], daySumP[i-1]])
```

3. Appendix

   Source web: https://e-service.cwb.gov.tw/HistoryDataQuery/

   Github: https://github.com/Wirra487/DSCP