

# Chapter 4 - Exercise 4: Weather Data

Bộ dữ liệu Weather Data bao gồm thông tin về thời tiết của tất cả các ngày trong tháng 7 năm 2017

Bộ dữ liệu chứa trong thư mục data gồm các file dữ liệu:

- 1\_TempHighF: Nhiệt độ cao nhất (độ F)
  - 2\_TempLowF: Nhiệt độ thấp nhất (độ F)
  - 3\_HumidityHighPercent: Độ ẩm cao nhất (%)
  - 4\_HumidityLowPercent: Độ ẩm thấp nhất (%)
  - 5\_VisibilityHighMiles: Tầm nhìn xa (miles)
  - 6\_VisibilityLowMiles: Tầm nhìn gần (miles)
  - 7\_WindMPH: Sức gió (miles per hour)
  - 8\_PrecipitationSumInches: Lượng mưa (inches)
- Link thư mục data: [https://drive.google.com/drive/folders/1614aFbGoWimoWYVi\\_x-o5F9Tbc\\_bNnPc?usp=sharing](https://drive.google.com/drive/folders/1614aFbGoWimoWYVi_x-o5F9Tbc_bNnPc?usp=sharing) ([https://drive.google.com/drive/folders/1614aFbGoWimoWYVi\\_x-o5F9Tbc\\_bNnPc?usp=sharing](https://drive.google.com/drive/folders/1614aFbGoWimoWYVi_x-o5F9Tbc_bNnPc?usp=sharing))

## \* Yêu cầu

1. Đọc dữ liệu từ các tập tin và lưu trữ vào list data với mỗi phần tử là dữ liệu được đọc từ 1 tập tin tương ứng trong thư mục data (có kiểu numpy array). In kết quả list data

2. Từ các numpy array trong list data, tạo dataframe dfWeather gồm các cột: - 'TempHighF', 'TempLowF', 'HumidityHighPercent', 'HumidityLowPercent', 'VisibilityHighMiles', 'VisibilityLowMiles', 'WindMPH', 'PrecipitationSumInches'

- Nhiệt độ cao nhất, nhiệt độ thấp nhất, độ ẩm cao nhất, độ ẩm thấp nhất, tầm nhìn xa, tầm nhìn gần, sức gió và lượng mưa

- Xem 5 dòng đầu tiên trong dfWeather

3. Tạo dataframe dfWeather2 lần lượt với các cột sau đây: 'TempHighC', 'TempLowC', 'HumidityHighPercent', 'HumidityLowPercent', 'VisibilityHighKMs', 'VisibilityLowKMs', 'WindKMPH', 'PrecipitationSumMM'

3.1 Chuyển đổi cột nhiệt độ cao nhất từ độ F sang độ C và lưu vào cột TempHighC trong dfWeather2 (làm tròn 2 số lẻ), hiển thị 5 giá trị đầu của TempHighC

- công thức chuyển từ độ F sang C:  $C = (F - 32) \times 5/9$
- ví dụ chuyển từ 80F sang C:  $C = (80 - 32) \times 5/9 = 26.67C$

3.2 Chuyển đổi cột nhiệt độ thấp nhất từ độ F sang độ C và lưu vào cột TempLowC trong dfWeather2

3.3 Tính nhiệt độ trung bình TempMediumC trong dfWeather2

3.4 Tạo mới các cột độ ẩm HumidityHighPercent và HumidityLowPercent trong dfWeather2 bằng các cột HumidityHighPercent và HumidityLowPercent trong dfWeather

3.5 Chuyển đổi cột tầm nhìn xa từ miles sang kilomet và lưu vào cột VisibilityHighKMs trong dfWeather2

- công thức chuyển từ Mile sang Kilomet:  $km = mile \times 1.609344$
- ví dụ chuyển từ 10 mile sang km:  $km = 10 \times 1.609344 = 16.09344 km$



3.6 Chuyển đổi cột tầm nhìn gần từ miles sang kilomet và lưu vào cột VisibilityLowKMs trong dfWeather2

3.7 Chuyển đổi cột sức gió từ miles/hour (WindMPH) sang kilomets/hour và lưu vào cột WindKMPH trong dfWeather2

3.8 Chuyển đổi cột lượng mưa từ inches sang milimet và lưu vào cột PrecipitationSumMMs trong dfWeather2

- công thức chuyển từ Inches sang Milimet:  $\text{mm} = \text{inches} * 25.4$
- ví dụ chuyển từ 10 inches sang milimet:  $\text{mm} = 10 * 25.4 = 254 \text{ mm}$

4. Xem thông tin thời tiết của 5 ngày đầu tiên trong dfWeather2

5. Từ dfWeather2, liệt kê thông tin thời tiết của các ngày có nhiệt độ trung bình cao nhất

6. Từ dfWeather2, liệt kê thông tin thời tiết của các ngày có nhiệt độ trung bình thấp nhất

7. Từ dfWeather2, liệt kê thông tin thời tiết của các ngày có mưa, sắp giảm theo lượng mưa

8. Từ dfWeather2, liệt kê thông tin thời tiết của các ngày không có mưa và có độ ẩm trung bình cao nhất

9. Tạo series ser chứa hai giá trị trung bình cột TempMediumC của các ngày có mưa và các ngày không mưa trong tháng

```
In [1]: import numpy as np
import os
```

```
In [ ]:
```