

Quiz9

- 題目說明

(一)

1. 請取出"owid-covid-data.csv" ('location', 'date', 'new_cases', 'new_deaths') 的欄位並取出 location 為 Thailand 的資料存成新的 dataframe，然而將缺失值補 0。
2. 求出 2021 年每個月平均新增案例(monthly_avg_case)以及每個月平均新增死亡案例(monthly_avg_death)，並代入提供的繪圖程式碼，如範例提供 Thailand Monthly Average Cases, Thailand Monthly Average Deaths 的圖。

(二)

1. 將 1-1 取出的資料複製成一個新的 dataframe 名為 df_new，須將 df_new 的索引設置為 date 欄的值，將 df_new 中的 new_cases 和 new_deaths 欄位的資料進行正規化。
2. 將 df_new 中的'new_cases'和'new_deaths' 分別依照日期存成 train(2020/1/3-2022-06-30)和 test(2022-07-01~ 2023/4/26) 的 dataframe，並將結果代入提供的 ARIMA 模型進行 2023-05-01 到 2024-01-01 的案例以及死亡結果預測。如範例 Daily New Cases ,Daily New Deaths 所示。
(y 軸數值可能不同)

- 正規化公式(Z-score normalization)(*可用快速方法)

公式： $Z = (x - \mu) / \sigma$

*公式中的變數:

Z 是正規化後的數值

x 是正規化前的數值

μ 是該批資料的算術平均數

σ 是該批資料的標準差

- 題目要求

1. 最後繳交之檔案，請繳交 Jupyter notebook 格式 (學號.ipynb)，否則一律以零分計算。

● 題目範例

