**文件介紹**

# 環境處理

## html檔

該網站是教學如何備份、還原環境

## Eisai.yaml

該檔案為我做實驗使用的環境檔，可使用該環境檔進行環境的還原

# 天氣資料

## [CWB爬蟲資料.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%A4%A9%E6%B0%A3%E8%B3%87%E6%96%99/CWB%E7%88%AC%E8%9F%B2%E8%B3%87%E6%96%99.ipynb)

這隻程式會以爬蟲的方式抓取觀測站的歷史資料，如果未來做實驗有需要抓取觀測站資料可以使用

## [氣溫差.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%A4%A9%E6%B0%A3%E8%B3%87%E6%96%99/%E6%B0%A3%E6%BA%AB%E5%B7%AE.ipynb)

這隻程式會整理爬蟲回來的觀測站資料，並計算出每一天最高溫、最低溫、最大溫差、每一週的平均溫度差、每一週的異常溫度差的天數



# 成大醫院物料預測

## 2019~2021庫存

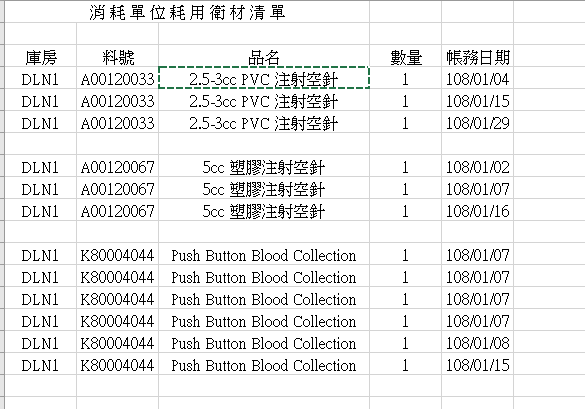
### 出入庫報表

該資料夾裡面存放著兩種資料，首先是出入庫報表，存放的是各個病房他們從大庫房領到味才的紀錄檔，最重要的欄位是受料量，受料量代表著該病房確實拿到的數量，會需要出入庫報表的原因是因為，有些衛材並不能確實記錄使用了多少，像是棉花，不可能一次用了一整包的棉花，只能以時間來估算可能一個月、2個禮包拜需要多少包的棉花



### 病房衛材帳務日期

該資料夾裡面存放的為各個病房確實使用了衛材的紀錄檔，舉個例子可能有一筆A00120033(2.5-3cc PVC 注射空針)再108/01/04使用了1支上面就會記錄一筆，該檔案室主要用來預測的檔案，因為他代表了這個病房衛材真實的用量



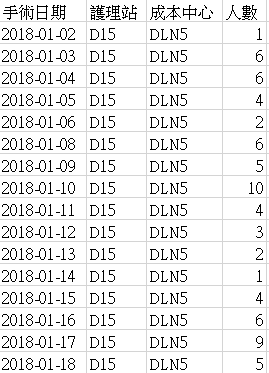
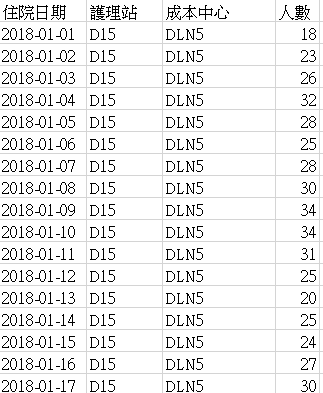
## 2022庫存

該資料夾裡面的資料為2022年的資料，由於案子再2022年就已執行，因此這是2023年後才提供的資料，並且該資料並沒經過人工處理，是由資料庫直接匯出，各個檔案內容需要再進行確認，目前暫定認為Charge開頭的檔案為病房衛材帳務日期的資料

# 醫院提供病房人數資料

## 病房人數

該excel檔為醫院提供的各個病房住院人數、開刀人數的原始資料



## [分病房計算人數.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E9%86%AB%E9%99%A2%E6%8F%90%E4%BE%9B%E7%97%85%E6%88%BF%E4%BA%BA%E6%95%B8%E8%B3%87%E6%96%99/%E5%88%86%E7%97%85%E6%88%BF%E8%A8%88%E7%AE%97%E4%BA%BA%E6%95%B8.ipynb)

如檔明所示，該支程式為計算各個病房每天的住院人數，最後統整出，該週總人數、該週最多人數、前1週人數、前1週最多人、前1週平均人數。



## [分病房計算開刀人數.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E9%86%AB%E9%99%A2%E6%8F%90%E4%BE%9B%E7%97%85%E6%88%BF%E4%BA%BA%E6%95%B8%E8%B3%87%E6%96%99/%E5%88%86%E7%97%85%E6%88%BF%E8%A8%88%E7%AE%97%E9%96%8B%E5%88%80%E4%BA%BA%E6%95%B8.ipynb)

如檔明所示，該支程式為計算各個病房每天的住院人數，最後統整出，該週開刀總人數、該週開刀最多人數、前1週開刀人數、前1週開刀最多人、前1週平均開刀人數。

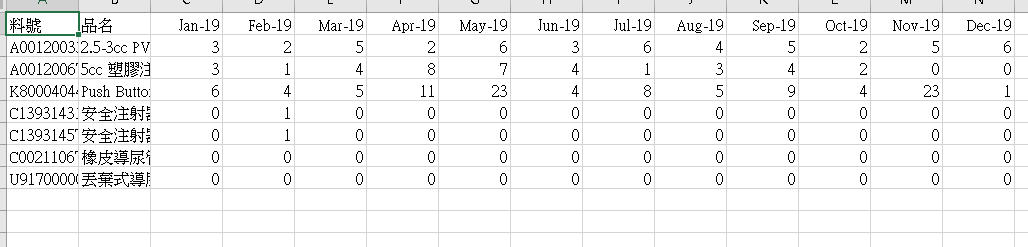


# 使用量計算

(已週、月為單位為當初想要查看各個病房衛材的資料情況，而撰寫來分析所使用)

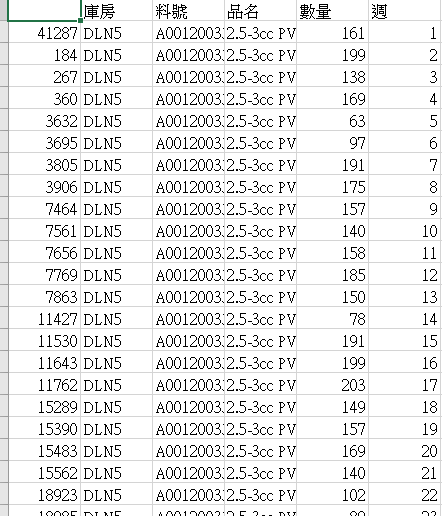
## [以月為單位\_使用量計算.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E8%A8%88%E7%AE%97/%E4%BB%A5%E6%9C%88%E7%82%BA%E5%96%AE%E4%BD%8D_%E4%BD%BF%E7%94%A8%E9%87%8F%E8%A8%88%E7%AE%97.ipynb)

該支程式為計算各個病房的月統計量，舉例來說該支程式會跑出一個病房所有衛材的使用量，紀錄單位為月。



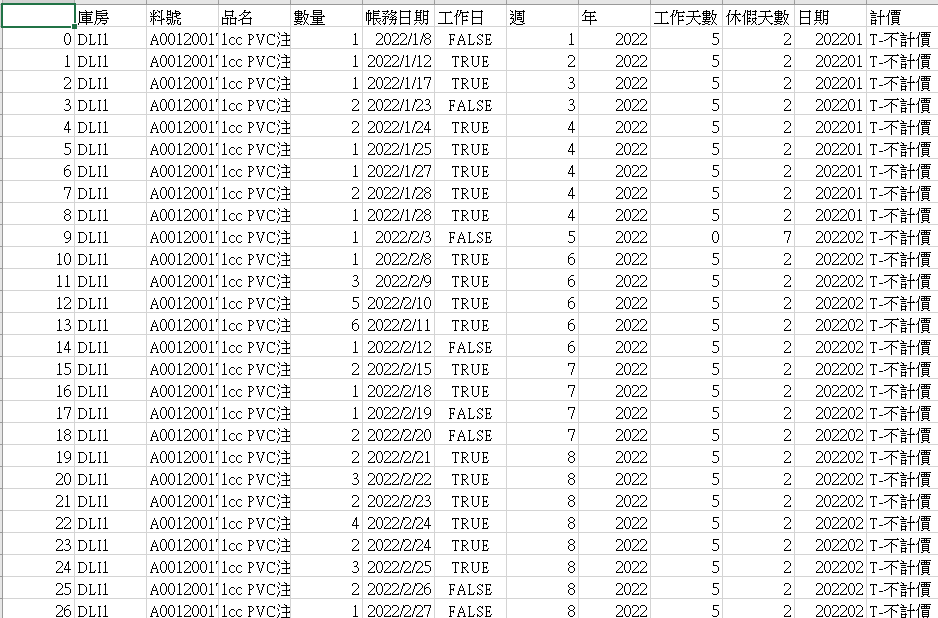
## [以週為單位\_使用量計算.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E8%A8%88%E7%AE%97/%E4%BB%A5%E6%9C%88%E7%82%BA%E5%96%AE%E4%BD%8D_%E4%BD%BF%E7%94%A8%E9%87%8F%E8%A8%88%E7%AE%97.ipynb)

該支程式為計算各個病房的週統計量，舉例來說該支程式會跑出一個病房所有衛材的使用量，紀錄單位為週。



## [工作日計算一次多個.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E8%A8%88%E7%AE%97/%E5%B7%A5%E4%BD%9C%E6%97%A5%E8%A8%88%E7%AE%97%E4%B8%80%E6%AC%A1%E5%A4%9A%E5%80%8B.ipynb)

該支程式算是較為重要的一支程式，這支程式事在整理每一個衛材從2019~2022年每一週的使用量，歷史用量為預測中模型會用到的重要資料，所有的衛材預測都必須先靠這支程式進行資料計算整理，進行病房的設定後，該支程式會跑出設定病房中的所有衛材每週的使用量資料



## 全部資料合併

該資料夾裡面會依照年分、病房去合併每一個病房每一個衛材的資料，合併後的資料情況如下圖

左邊依照年分、病房區分，右邊則是各個檔中的內容

# 使用頻率計算

[使用量排行+全部衛材品名.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E3%80%81%E9%A0%BB%E7%8E%87%E6%8E%92%E8%A1%8C/%E4%BD%BF%E7%94%A8%E9%87%8F%E6%8E%92%E8%A1%8C%2B%E5%85%A8%E9%83%A8%E8%A1%9B%E6%9D%90%E5%93%81%E5%90%8D.ipynb)

[全部衛材使用量以週為單位.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E3%80%81%E9%A0%BB%E7%8E%87%E6%8E%92%E8%A1%8C/%E5%85%A8%E9%83%A8%E8%A1%9B%E6%9D%90%E4%BD%BF%E7%94%A8%E9%87%8F%E4%BB%A5%E9%80%B1%E7%82%BA%E5%96%AE%E4%BD%8D.ipynb)

[各個衛材使用量加總和頻率合併.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E3%80%81%E9%A0%BB%E7%8E%87%E6%8E%92%E8%A1%8C/%E5%90%84%E5%80%8B%E8%A1%9B%E6%9D%90%E4%BD%BF%E7%94%A8%E9%87%8F%E5%8A%A0%E7%B8%BD%E5%92%8C%E9%A0%BB%E7%8E%87%E5%90%88%E4%BD%B5.ipynb)

由於中低頻率的衛材比較是用統計方法ARIMA，高頻率的衛材使用機器學習的方法，因此事先計算每個衛材的使用頻率。

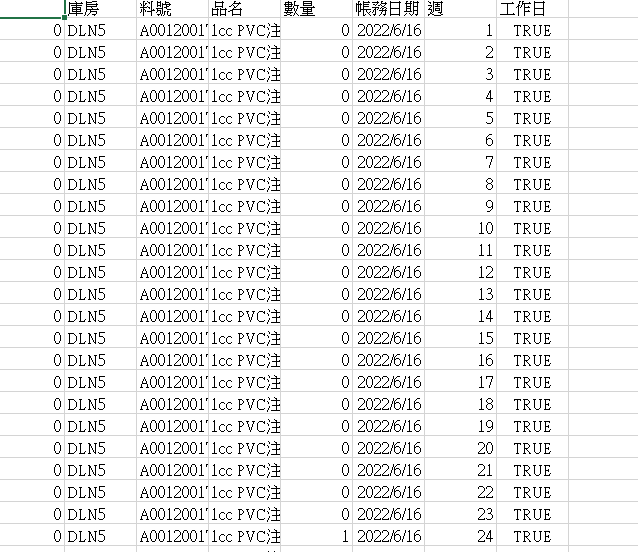
[使用量排行+全部衛材品名.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E3%80%81%E9%A0%BB%E7%8E%87%E6%8E%92%E8%A1%8C/%E4%BD%BF%E7%94%A8%E9%87%8F%E6%8E%92%E8%A1%8C%2B%E5%85%A8%E9%83%A8%E8%A1%9B%E6%9D%90%E5%93%81%E5%90%8D.ipynb)

計算某一個病房所有衛材的使用量頻率



[全部衛材使用量以週為單位.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E3%80%81%E9%A0%BB%E7%8E%87%E6%8E%92%E8%A1%8C/%E5%85%A8%E9%83%A8%E8%A1%9B%E6%9D%90%E4%BD%BF%E7%94%A8%E9%87%8F%E4%BB%A5%E9%80%B1%E7%82%BA%E5%96%AE%E4%BD%8D.ipynb)

該檔案可以直接計算出某一病房的每一個衛材每一週的用量



[各個衛材使用量加總和頻率合併.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E4%BD%BF%E7%94%A8%E9%87%8F%E3%80%81%E9%A0%BB%E7%8E%87%E6%8E%92%E8%A1%8C/%E5%90%84%E5%80%8B%E8%A1%9B%E6%9D%90%E4%BD%BF%E7%94%A8%E9%87%8F%E5%8A%A0%E7%B8%BD%E5%92%8C%E9%A0%BB%E7%8E%87%E5%90%88%E4%BD%B5.ipynb)

單純把品名跟頻率檔案做合併

# 模型預測

## [base\_line.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/base_line.ipynb)

## [base\_line\_twoweek.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/base_line_twoweek.ipynb)

## [base\_line\_onemonth.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/base_line_onemonth.ipynb)

這3個檔案為計算基線的檔案，基線就是我們透過衛材上一週的歷史用量當作預測值，直接計算出如果這樣做的績效為何，因為假設模型預測出的結果連基線都贏不了的話，那該模型就沒有意義了，之所以分為3個檔案的原因是因為每個衛材的使用頻率都不同，有些衛材可能每週都有使用，就需要每週預測，有些可能使用頻率是2週1次，因此2週才預測1次，總共分為3類每週預測、2週預測1次、每月預測1次

## [k-shape.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/k-shape.ipynb)

由於衛材跟衛材之間可能有關連，比如說空針可能會搭配針頭的使用，因此空針跟針頭的使用量曲線可能會很相同，而k-shape是個分群的演算法，他會透過演算法計把時間序列相似的曲線分為同一群，因此就可以找出哪些衛材的使用量曲線是類似的，再把他們一起當作訓練資料共同進行訓練的動作，相當於變相的增加資料集，並且讓模型去學習這些資料資間的關聯，該檔案只負責計算分群的部分。

## [k-shape+XGB.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/k-shape%2BXGB.ipynb)

透過該檔(k-shape.ipynb)知道哪些衛材為同一組之後，把同一組的衛材，利用此檔進行訓練的動作，這個檔案會將同組的衛材一起進行訓練後，跑出績效。

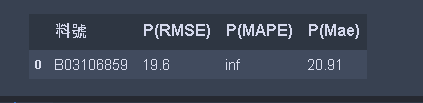


## [XGB.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/XGB.ipynb)

## [XGB預測多個衛材.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/XGB%E9%A0%90%E6%B8%AC%E5%A4%9A%E5%80%8B%E8%A1%9B%E6%9D%90.ipynb)

這兩支則為使用頻率為每週的衛材使用的預測檔案，差別在於一支一次只能預測一個衛材，另一支一次可以預測多個衛材(但和k-shape不同並沒有一起預測，只是單純個別預測)

[XGB.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/XGB.ipynb)



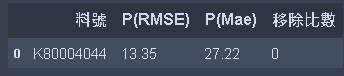
[XGB預測多個衛材.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/XGB%E9%A0%90%E6%B8%AC%E5%A4%9A%E5%80%8B%E8%A1%9B%E6%9D%90.ipynb)



## [最佳組合.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/%E6%9C%80%E4%BD%B3%E7%B5%84%E5%90%88.ipynb)

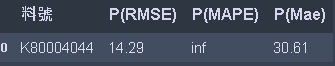
由於整理後的資訊，能使用的特徵可能有10幾20幾個，但每一種衛才可能有自己適用的特徵組合，應此這支檔案為實踐包裝方法來找出每一個衛材的最佳組合，最後的結果會跑出該衛材的最佳組合以及他的績效。





## [滑動arima-訓練資料長度一樣.ipynb](http://140.125.84.107:8888/notebooks/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E6%A8%A1%E5%9E%8B%E9%A0%90%E6%B8%AC/%E6%BB%91%E5%8B%95arima-%E8%A8%93%E7%B7%B4%E8%B3%87%E6%96%99%E9%95%B7%E5%BA%A6%E4%B8%80%E6%A8%A3.ipynb)

這支則為使用頻率為2週預測1次、每月預測1次的衛材使用的預測檔案，透過ARIMA統計模型來進行預測，比較特別的點是，不會一次呈現1年的預測值，而是預測一次後把預測週的真實用量加入訓練資料後，重複該動作26、13次，最後才計算績效



# 差值(該資料夾為系統用的程式)

實務上會以下列的使用週期進行區分

* 大於40週以上每周都進行預測 (SVR模型)
* 介於39~23週2週進行預測一次(ARIMA模型)
* 介於13~22週1個月進行預測一次(ARIMA模型)

不管哪一種預測週期，補足的都週期會以差值的方式補上，舉例，44週少了8週的使用量，讓模型學習0並沒有太大的意義，並且如果是0的話也預測不出來，因此以線性差值的方式將沒有紀錄的週數給補上。

## [run.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/run.py)

該檔案為和系統的接口(API)，裡面包含了

功能有:

1. 整個病房衛材預測
2. 整個病房使用頻率>40以上的衛材更新模型
3. 指定病房、料號，返回預測值

先透過抓取計算完整個病房每一個衛材使用頻率的資料，再依序將各個衛材分入要執行預測的檔案

## [predict.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/predict.py)

## [predict\_single.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/predict_single.py)

兩支檔案皆為預測衛材的程式，要使用甚麼模型進行預測的動作，會依照使用頻率的多寡進行預測的動作，40以上用SVR，40以下都使用ARIＭＡ，但預測後除了要返回預測值之外，還需要討論需要返回甚麼值

[predict\_single.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/predict_single.py)

這支為選擇特定的病房特定一個衛材進行預測的動作

[predict.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/predict.py)

這支為選擇其中一個病房隊全部的衛材進行預測的動作

## [renew.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/renew.py)

由於衛材的使用量可能過幾個月就會有改變，因此希望透過該支程式，來更新某個病房所有高使用率衛材的機器學習特徵，自動從計算一次每個衛材的最佳組合

## [last\_useage.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/last_useage.py)

該支程式為抓取衛材的上一周使用量為多少，由於有些衛材可能資料量不夠，或是各種原因導致報錯的關係，目前方法暫定報錯的衛材暫時直接以上一週的使用量作為代替

[every\_week.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/every_week.py)

[two\_week.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/two_week.py)

[four\_week.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/four_week.py)

我將衛材的預測依照頻率分為３種情況

* 每周都使用每周都預測[every\_week.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/every_week.py)
* ２個禮拜有一筆紀錄，２個禮拜預測一次[two\_week.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/two_week.py)
* １個月有一筆紀錄，１個月預測一次[four\_week.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E4%BA%A4%E6%8E%A5%E6%AA%94/%E5%B7%AE%E5%80%BC/four_week.py)

因此[predict.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/predict.py)　[predict\_single.py](http://140.125.84.107:8888/edit/%E6%88%90%E5%A4%A7%E9%86%AB%E9%99%A2/%E5%B7%AE%E5%80%BC/predict_single.py)這兩支程式都會先抓取衛材頻率的表，判斷要預測的衛材頻率為多少再決定要進入上面３支程式的哪一支程式進行預測的動作