**[壽險經營管理實務](https://ceiba.ntu.edu.tw/1081Fin5060_" \t "_blank)報告**

**——投資組合的報酬與風險歸因分析**

**11-7 詹凱翔 尹雨佳 錢云帆 黃容**

**一、專案起源**

近一世紀來金融機構增資的流動性大幅提升，資產與負債管理的重要性也隨之增長，投資機構在主動尋求投資標的時，才能穩定風險。投資標的不斷擴展，出現了多種投資組合，調整其中標的的權重可以幫助我們更近一步提升報酬率，同時降低風險。為此，我們的主題，是進行對投資組合的風險歸因，分析投資組合的報酬，比較投資組合與不同投資類別的「波動」（黃金、匯率、美元、原油、股票型基金、債券型基金），找出投資組合的風險歸因。

以基金為例，有股票型、債券型、平衡型，亦有分境外和境內的基金組合，但不能單以名稱當做基金的類型，而是應該研究他的波動來源。如股票型基金中若持有塑化股，其報酬就會受石油價格影響，而非像是單純的股票，而在債券型基金中，若持有可轉債，債券上所附加的買權使得資產有了股票的性質，並非單純的債券。當我們了解了投資組合的波動來源，就能分析、調整所希望的因子的權重，來滿足所希望達成的風險與報酬率。

為了貼近企業積極，且瞬息萬變的步伐，我們選擇以科技為核心，結合計量經濟知識，建構投資組合歸因的模型，以因應未來去紙本化的資料存取、迅速分析，更趁著機會，向明淇導師學習關於科技在金融、企業的應用。

經由明淇導師的指引，我們在製作這份專案時，將自己視為開發專案小組的一員。我們建立一個組內共通的資料庫(MariaDB)，擴充資料同時也共享資料；資料和文件均存放雲端(GitHub)，方便隨時存取與編輯；co-work一份投資組合歸因的程式碼，依功能分區編寫，方便未來更動因子與比重。實行的過程中，我們一邊學習程式語言、模型建構，更一邊學習如何討論與合作產生一份「投資組合的報酬與歸因」專案。明淇導師不斷提醒我們：「天花板是自己所定義」，而這份專案中，猶如他帶著我們進入職場合作的模式，除了學術上、技術上的成果，我們也因為合作、切磋，獲得校園內少有的實務學習。

以下的章節，將是我們分別對各個階段分工所進行的作業過程，以及成果展現。

**二、資料來源與處理**

在此一小節，我們共會有三類的資料檔案，依序為台灣股價資料、因子價格資料，以及想要測試的投資組合，以下依序做細部解釋：

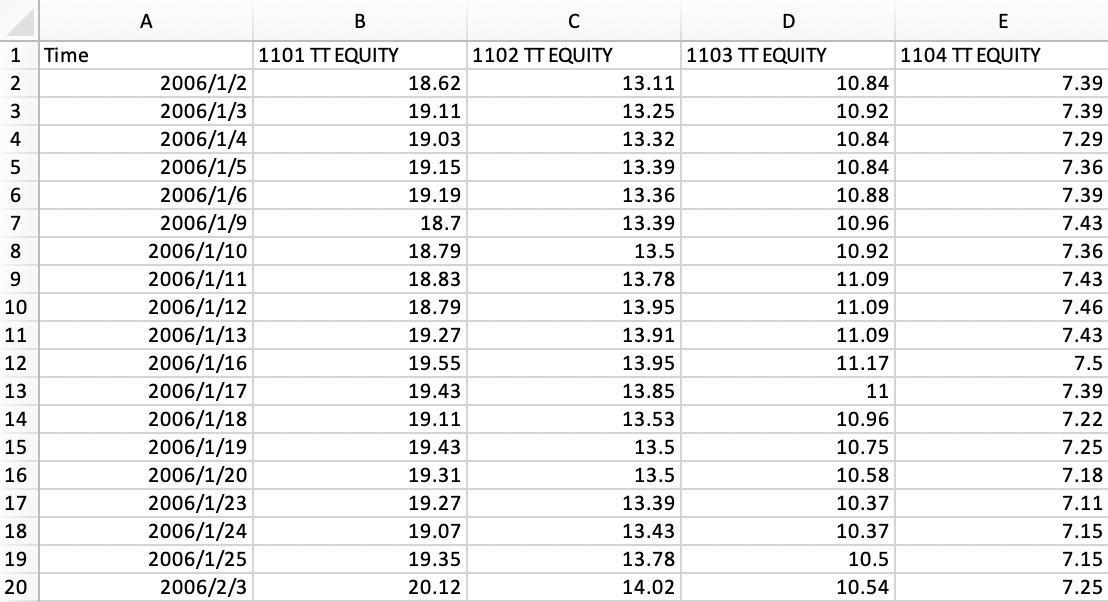
資料處理的部份，可以視為應變數與自變數兩區塊。其中應變數為要預測的投資組合報酬，單純只是計算給定一投資組合，隨者時間改變的報酬率 ; 自變數為用來預測的因子，則因為了滿足回歸分析的模型性質，或以利解釋結果，需對資料做較多的加工，而每一個加工的流程，我們都定義為一個函數。

1. **應變數資料：**

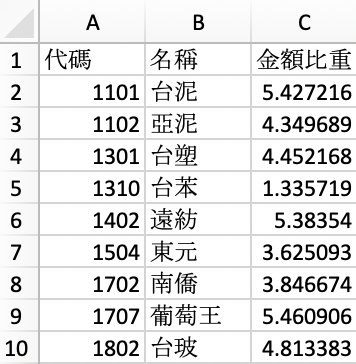
Step 1:

首先，我們在從Bloomberg抓下股價資料時，就先利用了excel的vlookup將股價時間對齊了（若當日無交易，則會以前一天收盤價自動填入，因為將vlookup第四個參數設為1，可以補入最相近者，而剛好是前一天收盤價）。

所以我們在用程式讀取時，預設股價檔案的長相如下：



我們也會有一組想要分析風險與報酬性質的投組，輸入格式如下：



Step 2:

我們計算報酬的思路為：

1. 我們會計算每一天的股票的“ 1 + 報酬率 ”，如：台積電連續三天的價格為100、110、121，那第一天的“ 1 + 報酬率 ”為110 / 100 =1.1 ; 第二天的“ 1 + 報酬率 ”為121 / 110 =1.1。
2. 想要得到投資組合的“ 1 + 報酬率 ”就將每一檔股票的“ 1 + 報酬率 ”乘上持有的權重。
3. 想要得到投資組合的累積報酬率，就將每一天的“ 1 + 報酬率 ”相乘。如果整個投組只有台積電，投組從第一天到第三天的報酬就會是 1.1 \* 1.1 – 1 = 0.21

Note: 我們丟進去的stock\_price為一個DataFrame，最後返回的portfolio\_return也是一個DataFrame。

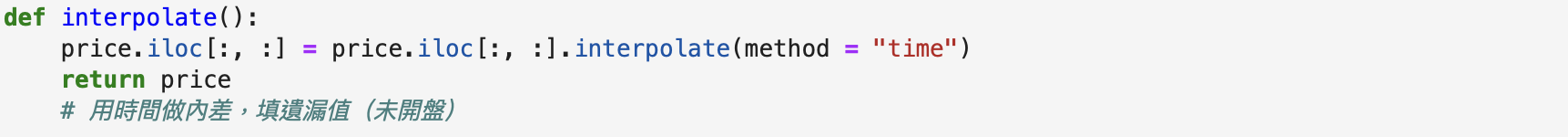
以下為此部分程式碼：



1. **自變數資料：**

函數1: 補遺漏值

我們拿取的資料可能因所屬國家不同，休市時間不一，因此我們利用時間長短來做內差法，補足空缺的資料。



函數2: 做標準化

因為不同的因子，其資料尺度不同，將資料做標準化，將來回歸係數的大小可以來拿比較對於自變數的影響大小。



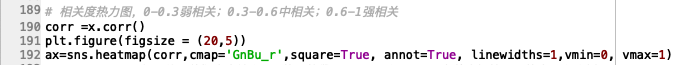
函數3: 做ICA

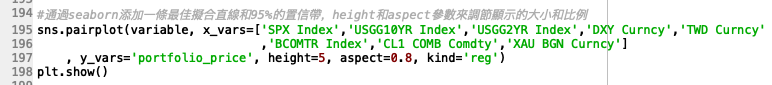
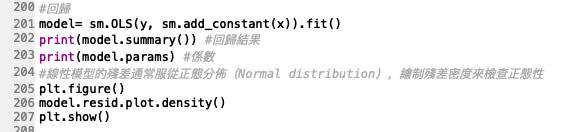
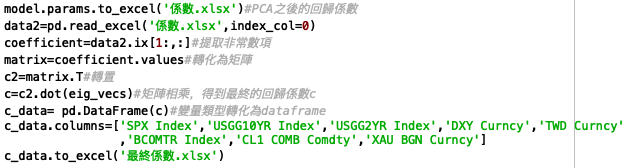
ICA 為將原本有一定程度相依的因子，轉化為獨立的因子的技術。如匯率與利率的關係數不等於0，而ICA套件所做的事就是計算出一組新的因子，其因子間的相關係數為0，而新產生的因子與原因子的相關性要最大。



**三、回歸分析**

1. **畫出相關度熱力圖，判斷變量之間的相關係數。代碼如下：**



1. **畫出每個解釋變量與因變量的散點圖，判斷其線型關係並畫出﻿添加一條最佳擬合曲線。代碼如下：**
2. **進行OLS回歸。代碼如下：**
3. **將3中的回歸係數轉化為矩陣，與PCA矩陣相乘得到最終的模型回歸係數。代碼如下：**