國立台北科技大學

資訊與財金管理系

股票市場最佳進場時機之相關性分析

指導教授：吳建文　　教授

學　　生：洪思平、莊凱捷

、賴奎皓

中 華 民 國 108 年 1 月

目錄

第一章　研究介紹

第二章　文獻探討

第三章　研究步驟

第一節　資料簡述

第二節　步驟

第四章　研究結果

第一節　前言

第二節　研究結果

第三節　研究結果說明

第五章　結論與建議

參考文獻

1. 研究介紹
2. 問題定義：

1.研究金融市場的人用盡各種方式就是找不到百分之百絕對獲   
 取正報酬的方法。

2.影響走勢的因素有無限多個，至今尚未有人找出最佳的參考  
 組合。

1. 問題的重要性：
2. 分析方法沒有侷限

　從股票成為一種交易標的以來，有非常多人預測與研究股票走勢，儘管現在投資人分析行情的方式成千論萬，至今尚未有人能精準預測股票市場的走向並且能保證獲利。

1. 隨機漫步理論(Random Walk Theory)

　顯示了股票的價格如同行人般，下一步走向哪邊沒有人能知道，若我們能從這些數據中發現高機率獲利的原則，便有機會打破此理論。

1. 現有方式的弱點：

1. 無法全面兼顧

　研究股票市場考量的因素有無限多種，包括緊急發生的事件，而這類型的因素往往導致重大變動，這會導致分析出來的結果和實際市場的走向是相差甚遠的。

2.無法數值化政治等因素

　影響股市走勢的因素有許多種，而有許多因素是沒辦法透過數值方法去研究，例如政治性因素、投資者心理因素等情況。

1. 數據已不具參考價值

　科技日新月異，下一波走勢主導的概念股及影響因素無人知曉，如果研究數據過於老舊，可能已經沒辦法符合現在時代的趨勢。

1. 我們的貢獻：
   1. 分析資料庫

　實際分析一個大資料庫，以真實交易數據做為研究基礎，相較其他研究都是以總體經濟去研究探討有更貼近真實交易市場性質。

* 1. 研究資料庫

　其他研究分析大多以單一或極少項因素影響股市走勢為探討依據，而我們彙整許多相關論文的資料，合併分析市場上普遍探討的因素合併研究股票走勢概況，並將所有的相關數據整理為大型資料庫，該資料庫也可供後續他人作為參考依據。

* 1. 找出最佳投資策略

　用分析出來找出最好的影響股市因素找到能在股市中賺取最高報酬的方法。

1. 我們將進行的工作：
   1. 過濾、整理、分析原始資料庫

　利用X公司的兩年完整交易資料作為原始資料，再以SQL Sever將資料作過濾及整理分析，並且找出連續兩年獲利最佳的交易人。

* 1. 查閱大量論文

　利用查閱大量論文，了解普遍對於研究股市的人想了解影響的因素有那些，進而去分析大量論文考量其對股票市場之因素，並將這些因素當作參考資料去了解在許多因素當作參考下的最佳影響因素會是哪些。

* 1. 查閱找出的因素對股市影響

　查找相關因素在國際市場上一整年的變動，研究其對最佳交易人的交易行為是否有其相關性，並把每一項因素的影響跟其他因素做比較。

1. 文獻探討

　　本章主要說明我們研究因素，以及回顧國內有關股票市場最佳進場時機的相關性分析，針對各項可能和股票報酬有關之文獻，最後說明與我們探討之議題異同之處。

　　我們的研究所使用之因素，有一般人參考之因素，以及以下文獻之探討因素。

　　賴芊卉(2017)以社會網路拓樸學來分析跨市場股票及金融市場資訊流動，先估計每檔股票或金融商品(石油、黃金、白銀、外匯)在跨市場訊息傳遞的網絡重要性，再將資料動態視覺化後，並探討變化背後之原因，最後實證分析跨市場網絡結構參數與報酬率特性之間的關係。

　　陳立鈞(2015)以1998年7月21日至2013年12月31日，以日資料採ADF單根檢定、VAR 向量自我迴歸模型迴歸評估…等，來研究臺灣股票、期貨市場、國際黃金價格及美元匯率判定四者之間是否存在關聯性，實證結果發現股價、期貨、黃金、匯率具有高度關聯性；期貨、股價、匯率具有關聯性；黃金價格與匯率和期貨具有關聯性；匯率對黃金和期貨有關聯性。

　　宋筧玲(2005)以台灣產業股價指數及西德州中級原油現貨價格之日資料，利用向量自我迴歸模型(VAR)，探討石油價格與台灣股票市場之報酬率是否存在領先-落後關係。實證結果顯示台灣股票市場中各種產業類股受到石油價格變動會因類股的不同而有所差異。

　　鍾蓓函(2017)以2012年1月至2015年12月止，運用向量自我迴歸(VAR)、ADF單根檢定、Granger因果關係檢定…等分析方式，實證後結果發現匯率較具影響力、黃金最容易受影響，油價與匯率彼此存在雙向回饋關係，台幣兌美元匯率會單向影響黃金價格。

　　楊士漢(2008)以動態條件相關模型 (DCC, Dynamic Conditional Correlation)，採美國NASDAQ綜合指數的隔日報酬對於台灣股市是否具有報酬和波動性的外溢效果，實證結果發現具有統計的顯著報酬外溢效果，其中對於台灣市場的隔日報酬影響最為顯著。

根據第一節文獻所探討因素之外，我們也加上一般人也會探討的因素，例如：大盤、三大法人、道瓊…等，我們與上述文獻相同之處，為探討股票市場與其他指數或金融商品之相關分析；而不同之處在於過去研究大多探討一項或兩項，至多四項因素與股票市場之間相關分析，而我們研究則合併了以上各個因素做相關分析，過去至今都曾為有人做過多項因素之分析。

1. 研究步驟
2. 資料簡述

　　本研究利用X公司之2013年及2014年完整交易資料，本資料含有22個欄位，根據下列圖示，以我們會使用到的資料作說明：

|  |  |
| --- | --- |
| BROKER\_NO3 | 交易員編號 |
| S\_ACNO | 顧客編號 |
| STOCK\_NO | 股票代號 |
| MYH\_DATE | 交易日期 |
| ORD\_NO | 訂單編號 |
| MTH\_PRICE | 股價 |
| TRADE\_KIND2 | 交易種類 |
| HAA09 |  |
| MTH\_QTY | 數量 |
| MTH\_AMT | 交易金額 |
| FEE | 手續費 |
| TAX | 交易稅 |
| IO\_AMT | 交易總額 |
| TRADE\_KIND1 |  |
| HAA20 |  |
| EX\_CTRADE-FLAG |  |
| BS\_STATUS | 買賣 |
| WEB\_ID |  |
| EX\_TRADE\_FLAG |  |
| EX\_MAKET\_FLAG |  |
| EX\_CORP\_FLAG |  |
| EX\_CREDIT\_FLAG |  |

　　先以X公司之資料，找出最佳獲利者，並研究該筆交易資料，從研究資料中，發現兩年度中最佳獲利者為同一交易人，因此我們會以此交易人之交易行為為研究出發點。

1. 步驟

　　由於數據資料過於龐大，已經超出了Excel工作表(1,048,576 列乘以 16,384 欄)能使用的範圍，因此我們利用SQL Server做為初步篩選資料的工具，下圖是查詢該交易紀錄的所有項目。

一、運用LIKE及INTO語法

我們使用 SQL語法中的LIKE語法，目的在於找尋交易記錄中日期欄位開頭為2013以過濾出2013年度的交易資料(這邊我們選擇以2013做分析，因在X公司的資料中2013是最為其全的)，由圖可見，查詢已過濾為2013開頭的資料，並藉由INTO語法使該查詢另外產生一個名為Trade2013的資料庫。

SELECT \*

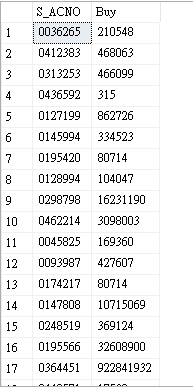
INTO　Trade2013

From [ODS].[dbo].[ST\_MATCH\_D]

WHERE [MTH\_DATE] LIKE ‘2013%’;

二、計算各個交易人之買入金額

為了計算出各個交易人的獲利(賣的交易總額減去買的交易總額)因此我們需要先整理出買賣的總額，因此我們選擇了[交易代號]及總合所有的[交易總額](並且重新命名為Buy欄位)，而篩選的條件設為[買賣狀態]為買(‘B’)，然後利用GROUP BY語法來將相同的[交易代號]結合(因此不會有重覆資料)之後存成名為Buy2013的資料庫，如圖為查詢結果。



SELECT [S\_ACNO],

SUM(IO\_AMT) as Buy

INTO Buy2013

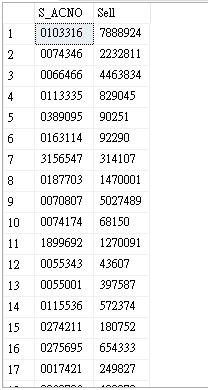
FROM [ODS].[dbo].[Trade2013]

WHERE [BS\_STATUS]=’B’

GROUP BY [S\_ACNO]

三、計算各個交易人之賣出金額

同理，可以找出[交易狀態]為賣(‘S’)的同樣找出各交易人賣的總和(取名為Sell)且存成名為Sell2013的資料庫。



SELECT [S\_ACNO],

SUM(IO\_AMT) as Sell

INTO Sell2013

FROM [ODS].[dbo].[Trade2013]

WHERE [BS\_STATUS]=’S’

GROUP BY [S\_ACNO]

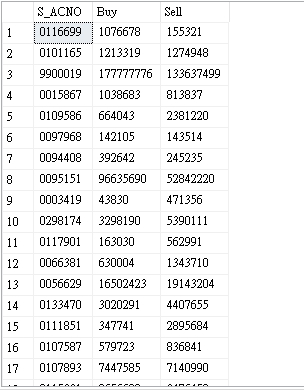
四、整併各個交易人買入即賣出金額

為了方便計算，我們分別將資料庫Buy2013的[交易代碼](與Sell2013的欄位相同)，Buy2013的Buy欄位與Sell2013的Sell欄位，條件在兩個資料庫[交易代碼]相同時將Buy、Sell做合併，下圖為輸出結果。

SELECT Buy2013.S\_ACNO,Buy2013.Buy,Sell2013.Sell

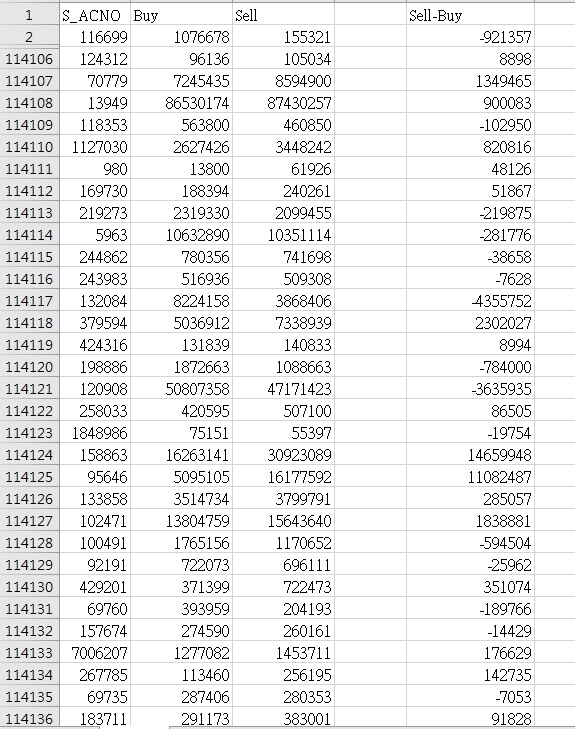
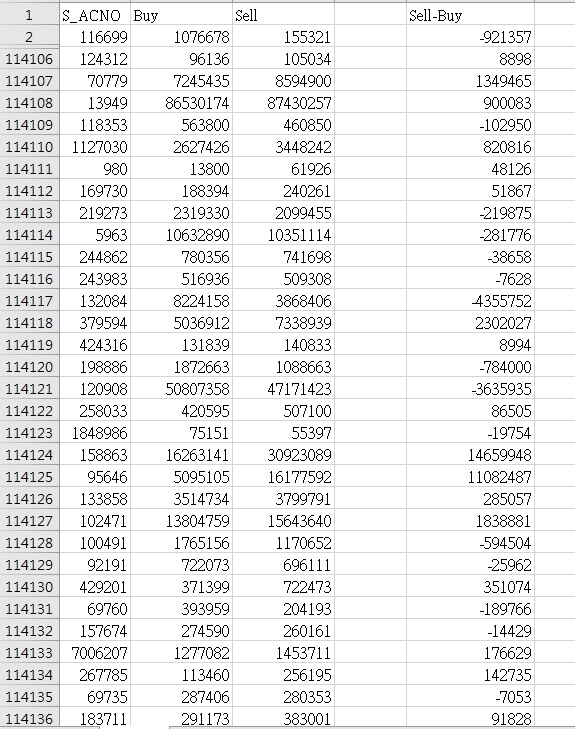
FROM Buy,Sell2013

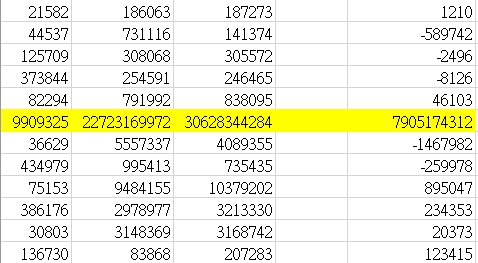
WHERE Buy2013,S\_ACNO=Sell2013.S\_ACNO



* 1. 利用Excel計算獲利

為了找出最賺錢的交易人，且我們已經將龐大的數據庫精簡為我們所需的資料，因此在這邊可以放入Excel做數據分析，下圖我們計算出各交易人的獲利(賣-買)且利用MAX語法找出獲利的最大值。如圖，找出最大值7905174312，之後以簡易查詢(Ctrl+F)反求找出獲利為該值的交易人，得到的結果為[交易代號]9909325，在2014同樣方法的驗證中，也同樣顯示最賺錢的為相同代號的交易人，因此我們合理推測，該交易人的成功並非偶然，而是有一套規則。





1. 找出最賺錢的人之交易紀錄

已知最賺錢的[交易代碼]為’9909325’而我們的目的在於分析該交易人的買點，因此我們在所有2013年度交易資料的資料庫(Trade2013)以交易代碼為’9909325’及[交易狀態]為買(‘B’)做為查詢，找出2013年該交易人的所有買點，並且以日期做合併，存成資料表DATE\_B9909325\_2013，如下圖。

SELECT [MTH\_DATE],

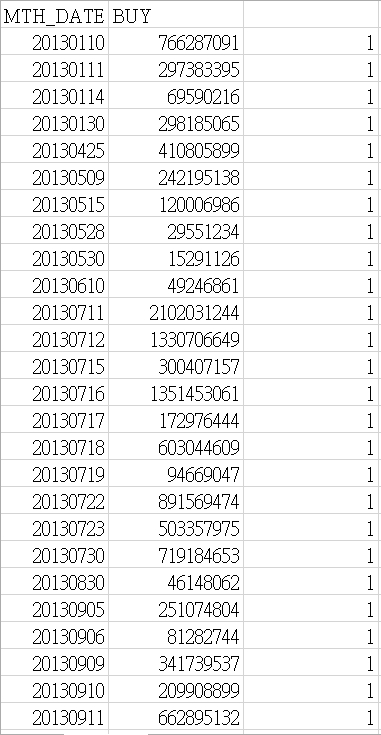
SUM(IO\_AMT) as BUY

INTO DATE\_B9909325\_2013

FROM [ODS].[dbo].[Trade2013]

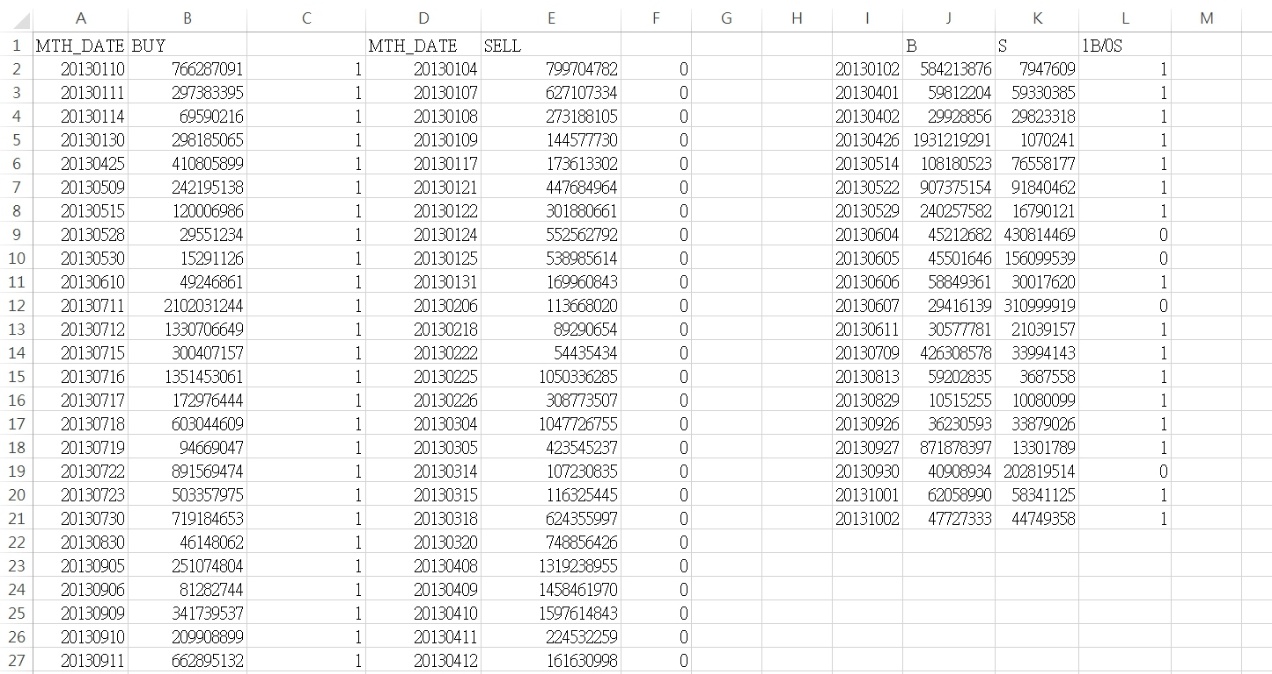
WHERE [S\_ACNO]=’9909325’ AND [BS\_STATUS]=’B’

GROUP BY [MTH\_DATE]



1. 列出最賺錢的人之交易紀錄

至此由於我們已經找出最賺錢的人，後面我們開始尋找資料並開始進行各因素與該交易人所有交易天數的比對(下圖為此人的買賣的金額與日期)

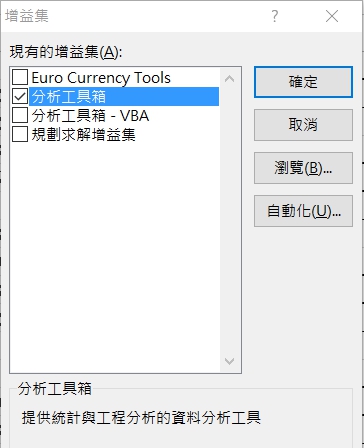


1. 以三大法人作為因素並且進行分析

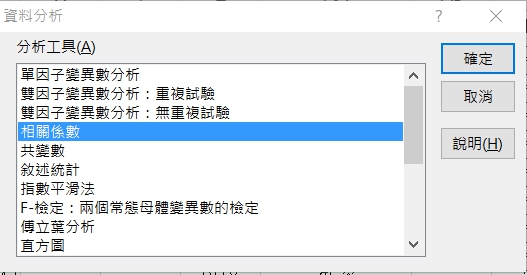
以三大法人的各項交易金額為例，同樣以日期做區分進行分析，透過Excel內部的分析工具箱中的資料分析，在使用相關性分析。

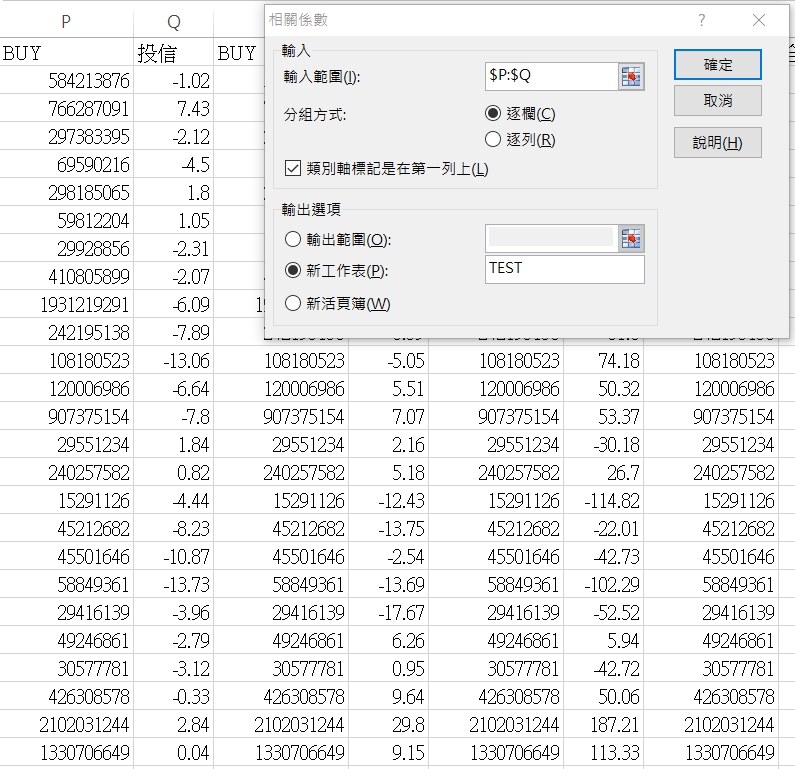
以下介紹分析工具箱的使用：

1. 首先先由檔案->選項->增益集->執行，開起分析工具箱如下圖。



1. 開啟後至資料->資料分析開起如下圖。



1. 選擇相關係數並分別將主要項目(僅能兩項)框起來做為輸入範圍(在此以該交易人買的金額與投信金額做為範例)，並輸出至新工作表(TEST)。
2. 相關係數分析後獲得的結果如下，其他因素結果的獲得方式以同樣方法取得。



1. 研究結果
2. 前言

　　我們透過相關性分析來了解X公司最賺錢的交易人與其他條件交易的相關性，這個基礎是建立在假設該交易人為絕對正確的前提下(所有交易皆是正收益)，並以普遍交易大眾的認知做為結果好壞的判斷；若結果為非顯著，也視為是不好的分析結果。

1. 結果

分析方式：Excel資料分析-相關分析

主要樣本: 大盤買賣超、三大法人(合計)買賣超、那斯達克、黃金價格、石油價格、美元匯率

分別計算:

大盤買賣超：



三大法人(合計)買賣超：



那斯達克指數：



黃金價格：



石油價格：



美元匯率：



第三節　研究結果說明

|  |  |  |
| --- | --- | --- |
|  | 結果正常 | 結果反常 |
| 正相關 | 投信　　　　　0.1169  自營商　　　　0.3505  大盤　　　　　0.3549  外資　　　　　0.5310  三大合計　　　0.5276  台積電(外資)　0.3176  台積電(合計)　0.3037 | 黃金　　　　　 0.2232 |
| 負相關 | 美金匯率　　 -0.3327 | 台積電(投信) 　-0.1012  　台積電(自營)　 -0.1465 |
| 無相關(相關係數過低) |  | 石油價格　　　　0.0940  那斯達克指數　　0.0480 |

　　相關係數的值介於-1與+1之間時，即-1 ≤ r ≤ +1。其性質如下：

當r > 0時，表示兩變數正相關，r < 0時，兩變數為負相關。

當|r| = 1時，表示兩變數為完全線性相關，即為函數關係。

當r = 0時，表示兩變數之間無線性相關關係。

當0 < |r| < 1時，表示兩變數存在一定程度的線性相關。且|r|越接近1，兩變數間線性關係越密切；|r|越接近於0，表示兩變數的線性相關越弱。

一般可按三級劃分：|r| < 0.4為低度線性相關；0.4 ≤ |r|< 0.7為顯著性相關；0.7 ≤ |r|< 1為高度線性相關。

　　以正常情況來講三大法人，包括投信、自營商、外資是主力，要與該投資法人的買賣超呈現正相關的關係，而大盤指數表示台灣上市公司加權指數，台積電(2330)是台灣股市權重最高的公司，擷取至2018年6月1日為止，占台灣上市公司市值17.8589%(資料取自公開資訊觀測站)，故大盤指數與台積電走勢也應與該法人投資人買賣呈正相關。

　　以基本面而言，台灣是以外銷為主的國家，尤其以大量外銷的企業而言，外國貨幣貶值(本國貨幣升值)，在財報上的結果會減少營收，所以勢必造成售出股票，股價下跌。

　　以籌碼面而言，預期美元持續升值(本國貨幣持續貶值)，對台灣股市外資佔一定比重而言，投資台股會損失匯差，所以進而賣掉股票造成股價下跌。

　　而本研究顯示美元對台幣匯率變化與該法人投資人當日買賣呈現負相關。相關係數為-0.337。

　　黃金是一種投資工具，正常情況下，全球政局動盪不安會造成金價上漲，資金流動勢必從股票市場流轉到黃金市場，而本研究顯示黃金價格的變化和該法人投資人買賣台股呈現正相關，相關係數為0.2232。

　CNNMONEY研究近年來，油價與股價同漲同跌，油價代表的是基本經濟物資，但油價影響的是一般投資人的信心。一般而言，油價與股市呈現正相關的關係，此研究顯示國際油價與該法人投資人當天買賣台股金額相關係數為0.094。

1. 結論與建議

　　在這研究中，我們取用的將近70項的因素以供做相關性分析，而不同於一般的研究，我們以交易人的交易記錄與大量的資料相比，以大量的數據得出結論，儘管最後我們的結果並沒有非常的符合期望(相關係數>0.7)，也有異常數據的產生，但研究總是由不斷的嘗試、累積，才能讓結果更加完美，而對於我們的研究數據，我們也建議相關研究者可以參考我們的數據以做進一步的分析(如:迴歸分析、假設檢定)，這些資料也可放在Python或R上做更完整的運算，期待此項研究能做為助力。

參考文獻

賴芊卉。跨市場股票、石油、黃金、白銀及外匯之報酬相關性網絡分析。檢自https://hdl.handle.net/11296/rj2htr(Jun.8,2018)

陳立鈞。臺灣股票及期貨市場與國際黃金價格和美元匯率之關聯性研究。檢自https://hdl.handle.net/11296/p54w6w(Jun.8,2018)

宋筧玲。國際石油價格波動對台灣股票市場影響之實證研究。檢自https://hdl.handle.net/11296/mqautd(Jun.8,2018)

鍾蓓函。世界新局勢下之美元，石油，和黃金的關聯性再檢視。https://hdl.handle.net/11296/v96p33(Jun.8,2018)

楊士漢。台灣股市與美國那斯達克的動態相關性分析。檢自https://hdl.handle.net/11296/sy26an(Jun.8,2018)

元大證券點精靈網路下單系統

公開資訊觀測站，台積電加權指數。檢自http://mops.twse.com.tw/mops/web/index(Jan.7,2018)

中華民國中央銀行全球資訊網，[新臺幣對美元銀行間成交之收盤匯率](https://www.cbc.gov.tw/np.asp?ctNode=862&mp=1)。檢自https://www.cbc.gov.tw/lp.asp?ctNode=863&CtUnit=506&BaseDSD=7&mp=1

Trading Economics,Crude Oil. Retrieved from https://tradingeconomics.com/commodity/crude-oil(Jan.7,2018)

Gold Price,Gold Price. Retrieved from https://goldprice.org/(Jan.10,2018)