

國立中山大學

應用數學系

一星二陽之 Top 4 股票技術分析

The Top 4 Stock Technical Analysis of  
One Star and Two Suns



B042040021 謝綺珊

B042040033 陳泓翰

B042040049 康雅滄

指導教授：羅夢娜教授

中華民國一百零七年六月二十六日

# 中文摘要

運用 R 語言套件蒐集所有台股的历史股價，運用 R 語言建立一星二陽策略程式對台股進行回測，檢驗此策略在台股的成效，並分析策略勝率最高的前 4 檔股票，觀察其歷史股價的表現，進行多方面的股價技術分析，如平均趨向指標 (ADX)、指數平滑異同移動平均線 (MACD) 等技術指標，並將成果視覺化，用以方便觀察趨勢或調整進出場時機。

關鍵詞：R 語言、一星二陽、勝率、回測、技術分析



# Abstract

Using the R Language Suite to collect the historical stock prices of Taiwanese stocks, and using the One Star Eryang strategy to back test Taiwan stocks. Test the effectiveness of this strategy in Taiwan stocks, and analyze the top 4 stocks with the highest strategy winning percentage to observe the performance of their historical stock prices. Conduct a variety of stock price technical analysis, such as the average directional indicator (ADX), moving average convergence / divergence (MACD) and other technical indicators. And visualizing the results to observe the trends or adjust the opportunities of entry and exit point.

Keywords and phrases : R language, one star and two suns, winning rate, back test, technical analysis



# 目 錄

摘 要 .....	I
Abstract .....	II
目 錄 .....	III
圖目錄 .....	IV
1. 前 言 .....	1
2. 研究方法 .....	2
3. 結果與分析 .....	4
4. 結論與後續工作 .....	22
參考文獻 .....	23

# 圖目錄

Figure1	2458.TW 義隆電子股票分析 . . . . .	5
Figure2	4952.TW 凌通科技股票分析 . . . . .	5
Figure3	4976.TW 佳凌科技股票分析 . . . . .	6
Figure4	2458.TW 可寧衛股票分析 . . . . .	6
Figure5	各自總結 . . . . .	7
Figure6	2458.TW 義隆電子股價波動率 . . . . .	9
Figure7	2458.TW 義隆電子新聞 . . . . .	9
Figure8	4952.TW 凌通科技股價波動率 . . . . .	10
Figure9	4952.TW 凌通科技新聞 . . . . .	10
Figure10	4976.TW 佳凌科技股價波動率 . . . . .	11
Figure11	4976.TW 佳凌科技新聞 . . . . .	11
Figure12	8422.TW 可寧衛股價波動率 . . . . .	12
Figure13	8422.TW 可寧衛新聞 . . . . .	12
Figure14	股票調整後股價之散佈圖 . . . . .	13
Figure15	調整後股價之散佈圖 (混合) . . . . .	14
Figure16	股票盒鬚圖 . . . . .	16
Figure16	道瓊指數盒鬚圖 . . . . .	16
Figure17	判斷相關係數與概率 . . . . .	17
Figure18	股票間的相關性 . . . . .	18
Figure19	總結 . . . . .	20
Figure20	總結 without 2458 . . . . .	20
Figure21	Shapiro-Wilk Normality Test . . . . .	21
Figure22	Score Test For Non-Constant Error Variance . . . . .	21

# 第 1 章 前言

隨著科技越來越進步，電腦的運算能力以指數般的速率提升，以往許多人類想像不到或計算量過於龐大的事情，都漸漸被電腦給實現。1956 年，麻省理工學院的約翰·麥卡錫在達特茅斯會議上為人工智慧 AI (Artificial Intelligence) 首次提出定義：讓機器的行為看起來像有智慧的人類一樣，時至今日，AI 的發展已成為當前市場上最火熱的話題。而在財金應用方面，可以運用電腦的高速運算能力來分析市場當前的情況，將所有可能影響市場的因素量化後輸入到程式中，透過機器學習的技術，讓模型自我修正，使得最後預測結果愈來愈符合市場實際情況。

由於之前上課有教過運用 R 語言查詢股票價格，搭配資料探勘的講義內容，以及郭美惠教授的演講，讓我們決定朝量化投資的方向進行研究。另一方面，由於我們組別成員對財金領域皆沒有太多涉略，因此投資標的決定以股票買賣為主，一方面是對股票這個名詞比較熟悉，另一方面是現今網路上的量化交易參考文獻主要也是以股票為主，資料來源相對較豐富。

## 第 2 章 研究方法

我們的研究方向主要是股票分析，由於之前對股票也沒有任何研究，因此決定從生活中較有可能接觸到的台股著手。分析的部分，經過搜尋網路上的資料，決定從判斷條件容易觀察的“一星二陽”作為回測的入場點。在兩根紅 K 線之間夾帶一根變盤線（十字線、電阻線），第三根紅 K 線的低點與第一根紅 K 線高點相近或較高，而此變盤線位置在兩根紅 K 的高點和低點之間，即為所謂的一星二陽型態。在投資領域中，有些投資者認為技術分析線圖上若出現一星二陽的 K 棒組合，代表這支股票未來會有一波多頭的走勢。由於一星二陽發生時一般是用肉眼觀察股價技術圖形，因此在轉換成函式時會有定義上的問題。這裡要特別感謝中山大學財務管理所王昭文教授的程式範例，替我們定義了許多艱深的技術圖形函式。而出場點的部分，則是收盤價跌破 20 日移動平均線時，以跌破當日的收盤價出場。20 日移動平均線的計算，是以前 20 個交易日收盤價之平均形成一個點，依次連接則形成了 20 日移動平均線。當股價運行到 20 日均線之下時，即代表這隻股票遇到阻力，需要考慮脫手。透過以上兩個策略的組合，我們可以對台股進行基本的買賣操作，並記錄交易明細，對結果進行分析。

而要進行股票分析的第一步就是要選擇適合的程式，這方面的選擇主要有 R 以及 Python，這兩者都是目前資料分析領域主流的程式。從定位角度來看，R 致力於提供良好的，對用戶友善的環境進行數據分析、統計分析和繪圖模型；而 Python 則強調生產效率和代碼的可讀性，適合進行工程開發的環境。我們的研究主要以 R 為主，一方面因為其較適合進行數據分析工作，另一方面是因為上課時已經提供許多 R 的使用教學，搭配自行上網查詢資料，可以減少學習新程式的麻煩。

要進行股票分析，我們需要的資料主要是台股的歷史股價。由於我們要從技術面著手，因此我們不需要研究公司的背景、財務狀況等，只單純從股票價格、成交量等量化的數據分析，並觀察現象。資料蒐集的方式是用 R 的 quantmod 套件直接取得，其資料來源為 YAHOO 財經每日更新的公開資料。由於這些公開資料是免費的，偶爾會有一些缺失的值，因此我們常常需要對數據做資料清洗，有時也

會因為不乾淨的資料導致程式出現錯誤。

取得歷史股價後，我們需要依照一星二陽策略實際進行回測。回測的概念是，假設自己回到過去的某個時間點，依照所選策略進行投資一段時間，並將結果記錄下來以檢測該策略的績效。在我們的研究中，設定的時間點最早是從十年前開始，而上市未滿十年的股票則從上市日開始進行回測；結束的時間點皆設定到2018年6月。這裡運用到的程式是R的基本函式建立，從歷史股價中，將符合一星二陽策略的日期篩選出來，並記錄買進賣出價格、計算報酬率，建立出一個完整的交易紀錄表。

得到股票的交易紀錄表後，我們要對交易紀錄進行績效評估。績效評估主要以以下七個用來衡量投資策略的指標：平均報酬率、勝率、交易次數、報酬率標準差、最大報酬率、最小報酬率以及平均持有日數。透過這些數據，我們可以得知這支股票在一星二陽投資策略下，獲利的效益及所承受的風險。由於我們是要對所有台股進行分析，且台股有一千多檔股票，所以我們需要用迴圈進行分析，並將結果記錄成一張表格，方便比較其中的數據。這裡會遇到的一個問題是，要對這些數據進行排序時，會因為資料型態的不同導致其變成字典序，而非我們熟悉的數值大小順序，所以我們需要將表格重新輸出成數值型態的方式儲存。

之後，我們從表格中挑選出勝率較高且交易次數較多的前四檔股票進行技術分析，觀察其績效較其他股票優異的原因。技術分析的部分，我們從常見的技術指標中挑選出幾個跟一星二陽方向相符的進行分析。使用到的技術指標有：K線圖、布林線指標(BBands)、指數平滑異同移動平均線(MACD)、成交量(Vo)，並將結果視覺化，以便觀察趨勢。另外，我們找出4檔股票的歷史股價中漲幅較大的時間點，上網搜尋當時是否有相關的新聞報導，試圖了解其股價暴漲的原因。最後是將挑選出的四檔股票的股價搭配”元大五十(0050.TW)”及”道瓊工業指數”進行統計分析，如相關性分析、回歸分析，將結果視覺化，並畫出股價的散佈圖、盒鬚圖逕行觀察。



## 第 3 章 結果與分析

股價分析最重要且最常見的工具便是 K 線圖，每一根 K 棒同時傳達了當天的開盤價、收盤價、當日最高價、當日最低價及漲跌等技術分析最主要用到的資訊。藉由 R 的 quantmod 套件，我們可以取得歷史股價並直接畫出 K 線圖。

### 布林線指標 (BBands)

一般而言，股價的運動總是圍繞某一價值中樞（如均線、成本線等）在一定的範圍內變動，布林線指標正是在上述條件的基礎上，引進了“股價通道”的概念，其認為股價通道的寬窄隨著股價波動幅度的大小而變化，二期股價通道又具有變異性，它會隨著股價的變化而自動調整。正是由於它的靈活性、直觀性和趨勢性的特點，BOLL（布林）指標漸漸成為投資者廣為應用的市場上熱門指標。

### 指數平滑異同移動平均線 (MACD)

這是一個常用的震盪指標，由 Gerald Appel 發明。用序列的快速移動平均線減去慢速移動平均線，可用來識別市場趨勢。

### 成交量 (Volume)

觀察該股當天交易是否熱絡

從 K 線圖表中，可以觀察出許多現象。首先是當股價向下衝出布林通道時，通常會伴隨著一股回彈的力量將股價拉回至通道中，而一星二陽則常發生在這種現象之後，類似於觸底反彈的概念。而相較於其他指標，指數平滑異同移動平均線在視覺上可以良好反應股價的趨勢，方便投資者判斷當前股票處於上升或下降的哪個階段。另外成交量大的時候常代表股市有劇烈動盪，有大量資金在該股票上流通。

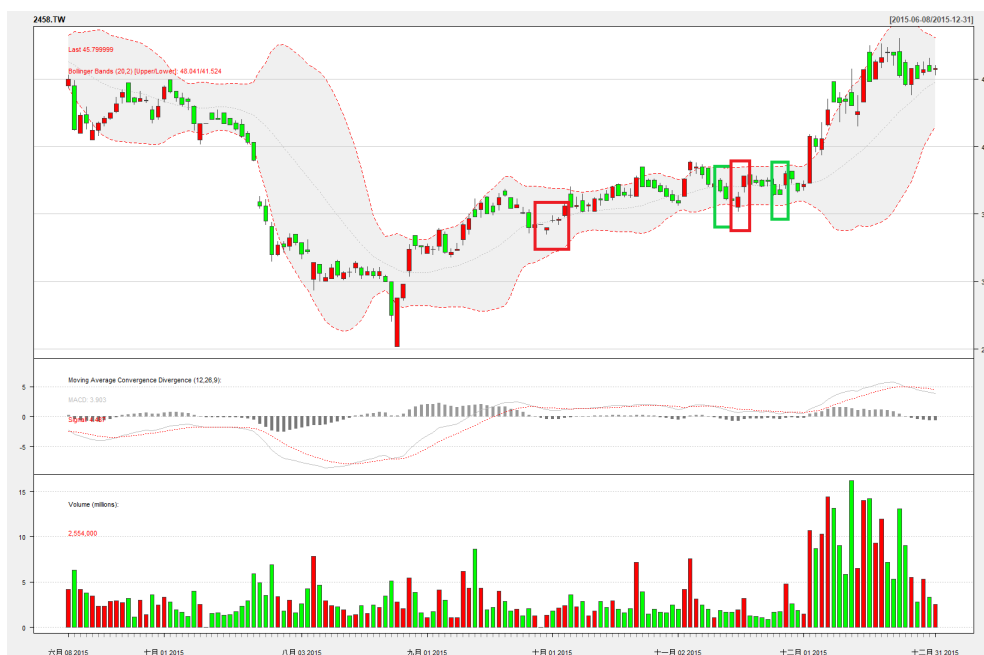


Figure 1: 2458.TW 義隆電子股票分析



Figure 2: 4952.TW 凌通科技股票分析

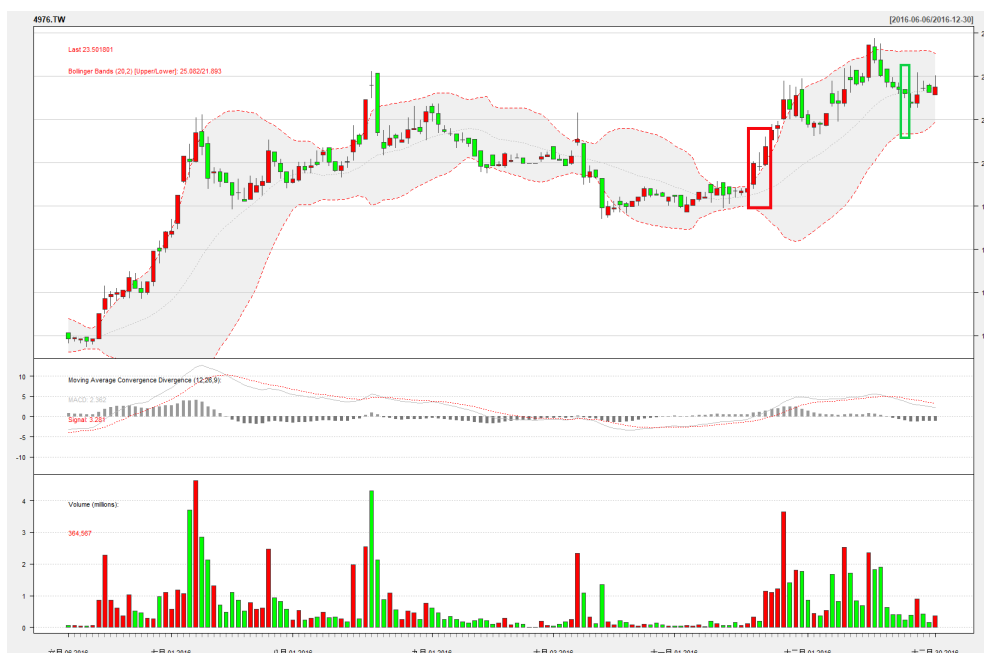


Figure 3: 4976.TW 佳凌科技股票分析

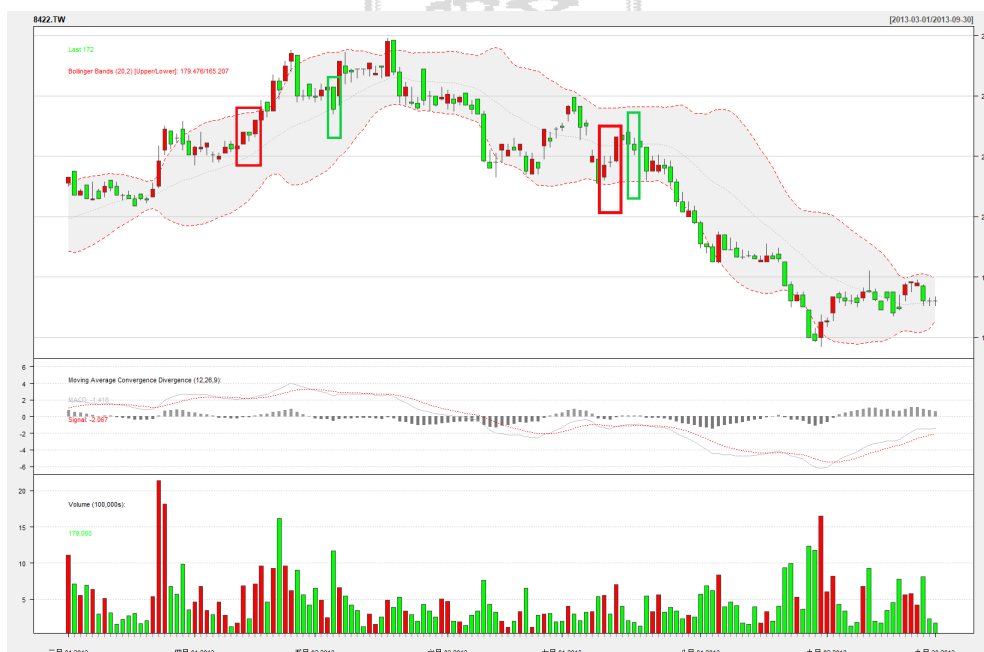


Figure 4: 2458.TW 可寧衛股票分析

```

> summary(`8422.Tw`)
      Index      8422.Tw.Open      8422.Tw.High      8422.Tw.Low      8422.Tw.Close      8422.Tw.Volume      8422.Tw.Adjusted
Min.   :2011-10-05 Min.   :126.5 Min.   :129.5 Min.   :126.5 Min.   :126.5 Min.   : 0 Min.   :106.9
1st Qu.:2013-06-10 1st Qu.:169.5 1st Qu.:171.0 1st Qu.:168.0 1st Qu.:169.1 1st Qu.: 163000 1st Qu.:144.8
Median :2015-02-05 Median :176.5 Median :178.0 Median :175.0 Median :176.5 Median : 255500 Median :154.6
Mean   :2015-02-10 Mean   :180.9 Mean   :182.6 Mean   :179.0 Mean   :180.6 Mean   : 359452 Mean   :153.7
3rd Qu.:2016-10-13 3rd Qu.:192.0 3rd Qu.:193.4 3rd Qu.:190.0 3rd Qu.:191.4 3rd Qu.: 427794 3rd Qu.:162.8
Max.   :2018-06-22 Max.   :258.5 Max.   :259.0 Max.   :248.0 Max.   :258.0 Max.   :3397000 Max.   :197.5
NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6

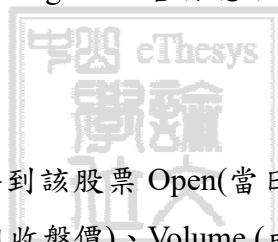
> summary(`2458.Tw`)
      Index      2458.Tw.Open      2458.Tw.High      2458.Tw.Low      2458.Tw.Close      2458.Tw.Volume      2458.Tw.Adjusted
Min.   :2007-01-02 Min.   :14.80 Min.   :15.84 Min.   :14.41 Min.   :15.35 Min.   :0.000e+00 Min.   :12.75
1st Qu.:2009-11-11 1st Qu.:37.40 1st Qu.:38.00 1st Qu.:36.80 1st Qu.:37.38 1st Qu.:3.066e+06 1st Qu.:34.30
Median :2012-09-17 Median :44.90 Median :45.54 Median :44.30 Median :44.80 Median :6.544e+06 Median :40.86
Mean   :2012-09-23 Mean   :44.91 Mean   :45.62 Mean   :44.14 Mean   :44.76 Mean   :4.034e+07 Mean   :40.23
3rd Qu.:2015-08-05 3rd Qu.:51.10 3rd Qu.:52.00 3rd Qu.:50.20 3rd Qu.:50.90 3rd Qu.:1.402e+07 3rd Qu.:46.05
Max.   :2018-06-22 Max.   :95.84 Max.   :96.63 Max.   :88.32 Max.   :94.95 Max.   :5.045e+10 Max.   :75.60
NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6

> summary(`4976.Tw`)
      Index      4976.Tw.Open      4976.Tw.High      4976.Tw.Low      4976.Tw.Close      4976.Tw.Volume      4976.Tw.Adjusted
Min.   :2012-10-03 Min.   :11.26 Min.   :11.51 Min.   :11.21 Min.   :11.21 Min.   : 0 Min.   :11.21
1st Qu.:2014-03-16 1st Qu.:20.17 1st Qu.:20.47 1st Qu.:19.88 1st Qu.:20.09 1st Qu.: 59484 1st Qu.:19.91
Median :2015-08-13 Median :29.35 Median :29.63 Median :28.98 Median :29.38 Median : 160981 Median :28.50
Mean   :2015-08-13 Mean   :28.96 Mean   :29.42 Mean   :28.49 Mean   :28.88 Mean   :1462277 Mean   :28.16
3rd Qu.:2017-01-10 3rd Qu.:36.09 3rd Qu.:36.62 3rd Qu.:35.82 3rd Qu.:36.19 3rd Qu.:1035239 3rd Qu.:33.53
Max.   :2018-06-22 Max.   :73.50 Max.   :74.00 Max.   :67.50 Max.   :70.90 Max.   :46614491 Max.   :70.90
NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6

> summary(`4952.Tw`)
      Index      4952.Tw.Open      4952.Tw.High      4952.Tw.Low      4952.Tw.Close      4952.Tw.Volume      4952.Tw.Adjusted
Min.   :2011-11-01 Min.   :17.10 Min.   :17.25 Min.   :17.00 Min.   :17.10 Min.   : 0 Min.   :12.03
1st Qu.:2013-07-01 1st Qu.:28.99 1st Qu.:29.44 1st Qu.:28.40 1st Qu.:28.80 1st Qu.: 289000 1st Qu.:21.04
Median :2015-02-28 Median :37.15 Median :37.55 Median :36.60 Median :37.10 Median : 604000 Median :29.23
Mean   :2015-02-24 Mean   :35.23 Mean   :35.68 Mean   :34.72 Mean   :35.12 Mean   :1260734 Mean   :29.50
3rd Qu.:2016-10-19 3rd Qu.:41.10 3rd Qu.:41.50 3rd Qu.:40.60 3rd Qu.:41.00 3rd Qu.:1318250 3rd Qu.:36.21
Max.   :2018-06-22 Max.   :68.00 Max.   :71.20 Max.   :65.00 Max.   :67.00 Max.   :29269563 Max.   :67.00
NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6 NA's   :6

```

Figure 5: 各自總結



藉由 Summary, 我們可以得到該股票 Open(當日開盤價)、High(當日最高價)、Low(當日最低價)、Close(當日收盤價)、Volume(成交量)、Adjust(調整後股價)的最大最小值, 以及四分位數值的資料。

挑出這些股票暴漲的時間點（例如開盤價或收盤價比前一天漲跌幅度超過 9%），通過搜尋引擎尋找是什麼原因導致這些暴漲暴跌，觀察資料，看就暴漲暴跌事件是否有可以利用的買賣規律以及和一星二陽發生的時間點有無關聯。而我們也確實找到了一些新聞，當中釋出一些利多的資訊，最終反映到股價上。如義隆公司 (2458.TW) 在打贏與蘋果公司的專利訴訟後，股價確實一度達到漲停板。而其他公司也因為宣布接單量增加或是擴大合作等消息而造成後續股價大幅上漲。

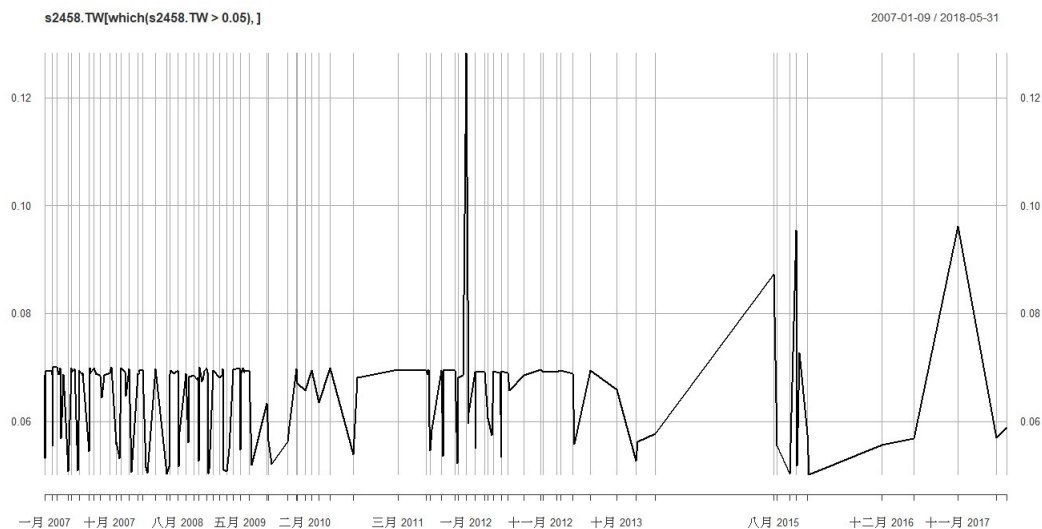


Figure 6: 2458.TW 義隆電子股價波動率

義隆、蘋果互告侵權 和解收場



2012-01-06

〔記者洪友芳 / 新竹報導〕IC設計義隆（2458）與美商蘋果（Apple）互告專利侵權案，經過2年多的訴訟，雙方以交互授權達成和解，蘋果並支付500萬美元（約1.5億台幣）給義隆，這樁和解在義隆近期即將公布的財報中呈現。

Figure 7: 2458.TW 義隆電子新聞

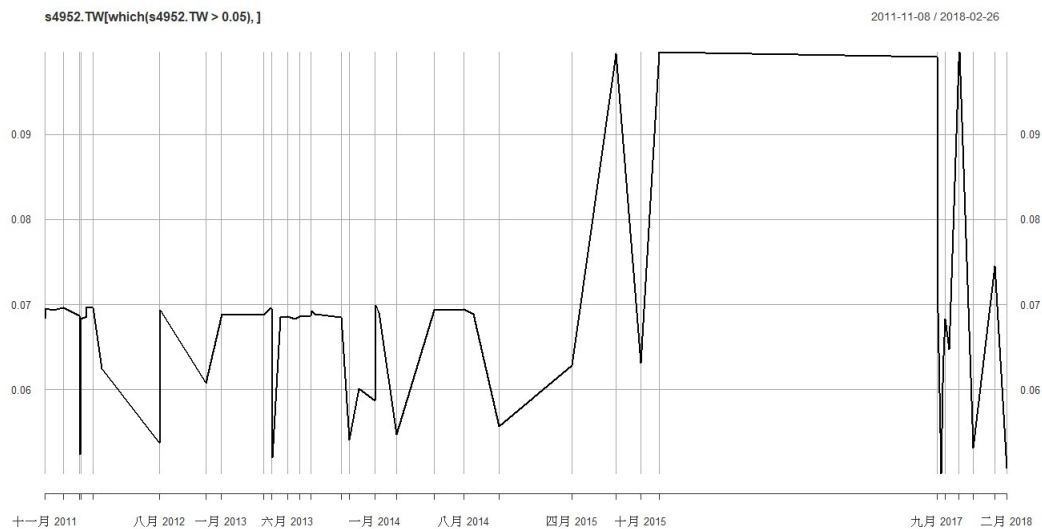


Figure 8: 4952.TW 凌通科技股價波動率



2016年03月14日 04:09 [工商時報](#) 記者楊曉芳 / 台北報導

IC設計凌陽（2401）集團全力朝向車用市場發展，包括集團旗下小金雞MCU廠凌通（4952）也不例外，凌通藉由在影像處理晶片累計超過5年的經驗，去年積極推出行車紀錄器、車用碰撞警示等車用影像相關晶片，去年獲利改寫上市以來新高後，今年將跟隨集團步伐在大陸車用後裝市場發展、再戰新高。

準總統蔡英文上周赴竹科參訪5家IC設計公司技術展示，其中由凌陽集團董事長黃洲杰親自展示的車用ADAS相關系統技術中，即有多項影像MCU技術是來自自家IC設計公司凌通。

Figure 9: 4952.TW 凌通科技新聞

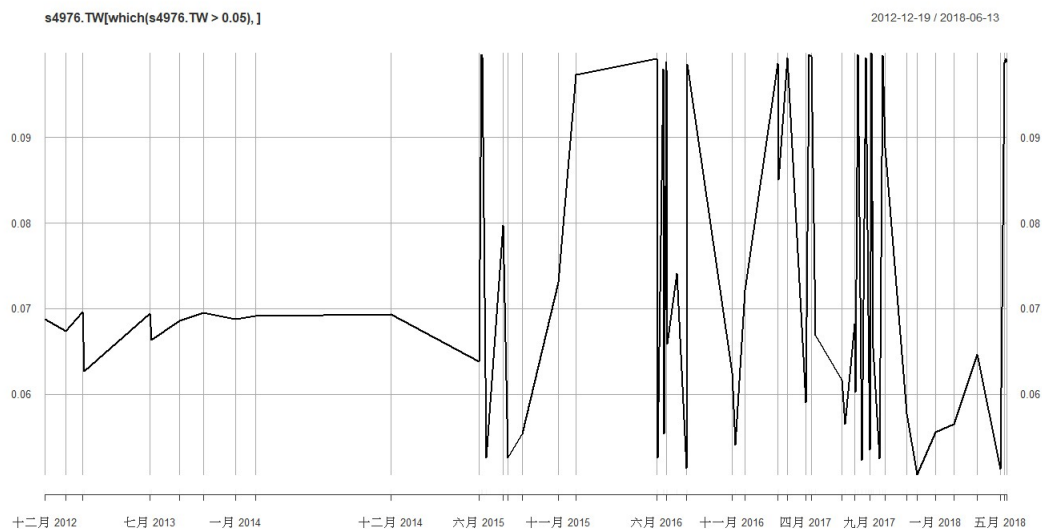


Figure 10: 4976.TW 佳凌科技股價波動率

中國 eThesys 學術論壇

## 《光電股》車載鏡頭出貨放量，佳凌今年業績看俏

財經

20180624 14:26  
清大研發聰明AI呵護  
晶圓製造 恆溫又節能

20180624 14:23  
孔劉歐巴二度為  
ZenFone代言 今現身  
與千名粉絲互動

20180624 12:55  
國內汽、柴油價格明  
(25)日起各調降0.2元

20180624 08:30  
《財訊雙周刊》政府

Facebook Google+ Twitter Weibo

2018年02月23日 13:24 時報資訊 【時報記者張漢源台北報導】

佳凌(4976)專注車載及安控等非手機鏡頭領域，逐漸走出自己的一片天，佳凌預估，今年車載鏡頭出貨量可望較去年成長50%，上看300萬顆，在車載鏡頭出貨持續放量下，今年車載鏡頭佔營收比重可望突破20%，受惠於日系投影機客戶推出新機，以及車載鏡頭穩定成長，佳凌預估，第1季合併營收有機會較去年同期成長20%到30%，今年業績也會比去年好。

Figure 11: 4976.TW 佳凌科技新聞

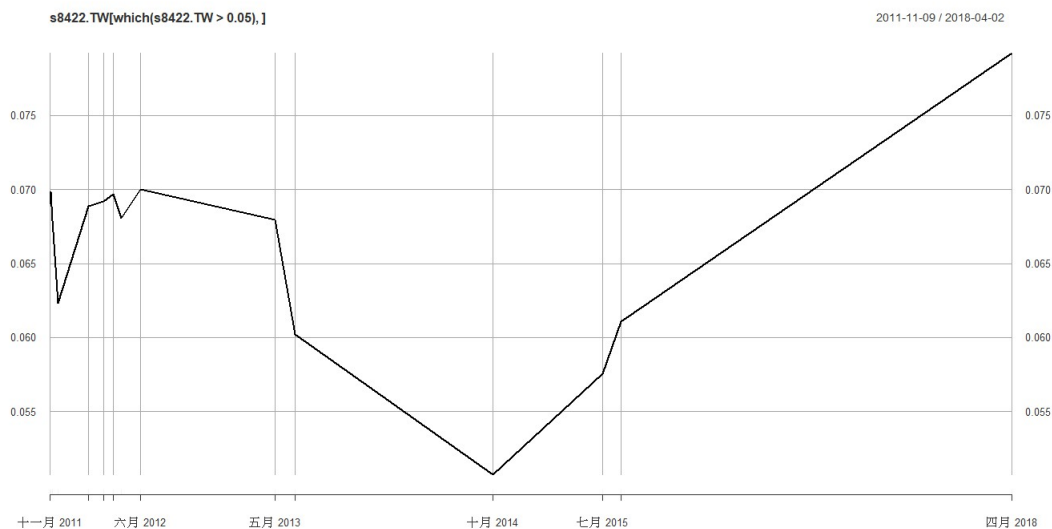


Figure 12: 8422.TW 可寧衛股價波動率



**可寧衛攜手全球環保大廠，正式跨入焚化領域**  
 2018/04/09 12:05 回應(0) 人氣(2762) 收藏(0)

MoneyDJ新聞 2018-04-09 12:05:39 記者 林詩茵 報導



廢棄物處理大廠**可寧衛**(8422)日前公告與台灣昇達國際廢料公司共同取得榮工大發廠**民營化**案最優投資人團資格，三方將共同設立新項目公司，可寧衛、台灣昇達將持股各29%、而榮工持股42%，此案對可寧衛而言意義重大，將是公司正式進入焚化處理領域的開始，在有相關實績及經驗後，有利後續至中國大陸或東協等地參與焚化廠興建、營運，尤其，昇達為全球第二大環保廠商，未來是否在海外合作更引起法人圈注目。

可寧衛為國內最大事業廢棄物處理廠商，業務從清運、**固化**到掩埋一條龍，為擴大營運觸角，並降低有限的掩埋空間影響營運動能，公司近幾年積極參與爭取其他廢棄處理領域業務，如今，終於獲得榮工大發廠焚化爐的營運案，此為可寧衛首次跨入焚化業務，盼藉此經驗拓展更多元的營運模式。

據可寧衛公告，公司與台灣昇達共同取得此民營化案最優投資人團資格，未來可寧衛、台灣昇達和榮工等三方將合組新項目公司，公司取得29%，總投資金額為6.5億元。

Figure 13: 8422.TW 可寧衛新聞



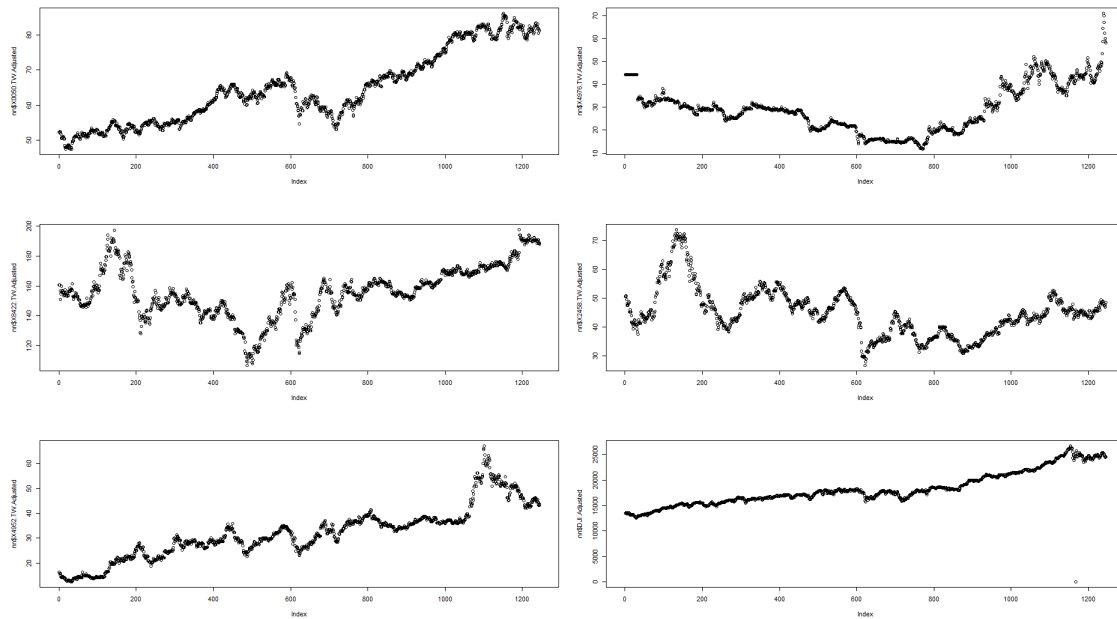


Figure 14: 股票調整後股價之散佈圖

由散佈圖可以觀察出元大 50 和道瓊工業指數的趨勢非常相似，但當道瓊出現一點波動時，元大 50 的波動會更加劇烈，其充分顯現出台灣作為一出口貿易為主國家，且易受到美國的波動影響。

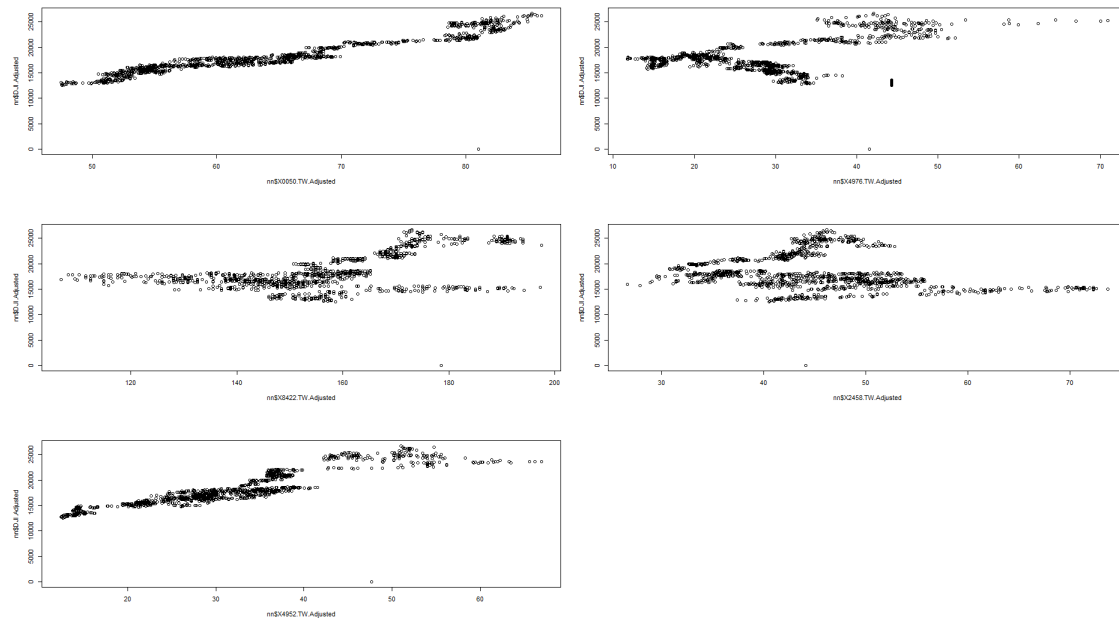
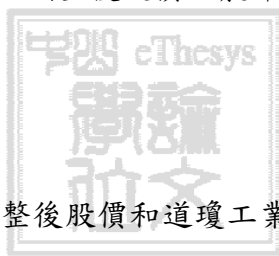


Figure 15: 調整後股價之散佈圖 (混合)



觀察 (Figure15) 五隻台股調整後股價和道瓊工業指數，發現元大 50(0050) 與凌通 (4952) 和道瓊工業指數間存在較強烈的線性關係。

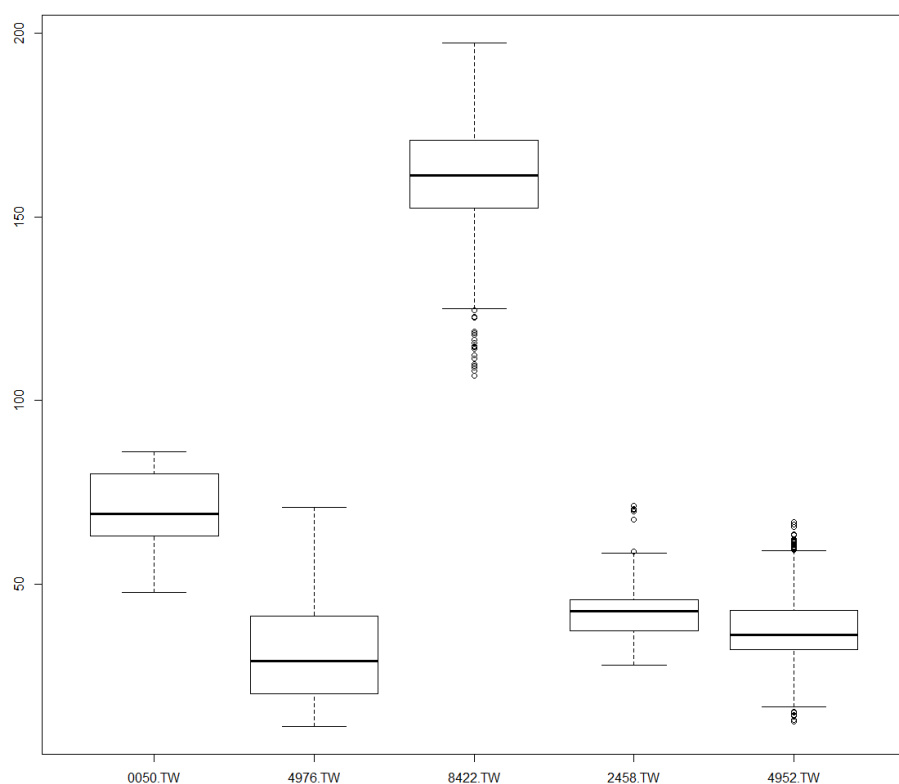


Figure 16: 股票盒鬚圖

由 (Figure16) 可發現，佳凌 (4976) 義隆 (2458) 凌通 (4952) 的平均股價大致相同，且其皆為科技公司；而可寧衛 (8422) 的調整後股價平均比其他都高，其為台灣事業廢棄物處理大廠，因其資本額較其他公司高，所以股價也較高。還有元大 50(0050)，其涵蓋臺灣證券市場中市值前五十大上市公司，可以將其視為台灣主要股市經濟趨勢的風向球。

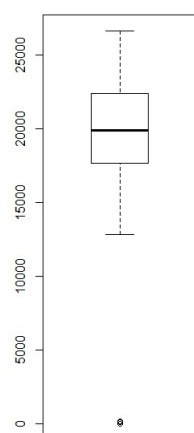


Figure 17: 道瓊指數盒鬚圖

道瓊工業指數，我們將它視為美國股市的主要股市經濟指標。



```

> corr.test(N)
Call:corr.test(x = N)
Correlation matrix
      x4976.Tw.Adjusted x8422.Tw.Adjusted x2458.Tw.Adjusted
x4976.Tw.Adjusted      1.00              0.71              0.49
x8422.Tw.Adjusted      0.71              1.00              0.24
x2458.Tw.Adjusted      0.49              0.24              1.00
x4952.Tw.Adjusted      0.62              0.67              0.14
DJI.Adjusted           0.75              0.73              0.16
x0050.Tw.Adjusted      0.80              0.72              0.18
      x4952.Tw.Adjusted DJI.Adjusted x0050.Tw.Adjusted
x4976.Tw.Adjusted      0.62              0.75              0.80
x8422.Tw.Adjusted      0.67              0.73              0.72
x2458.Tw.Adjusted      0.14              0.16              0.18
x4952.Tw.Adjusted      1.00              0.84              0.83
DJI.Adjusted           0.84              1.00              0.91
x0050.Tw.Adjusted      0.83              0.91              1.00
sample size
[1] 700
Probability values (Entries above the diagonal are adjusted for multiple tests.)
      x4976.Tw.Adjusted x8422.Tw.Adjusted x2458.Tw.Adjusted
x4976.Tw.Adjusted      0                  0                  0
x8422.Tw.Adjusted      0                  0                  0
x2458.Tw.Adjusted      0                  0                  0
x4952.Tw.Adjusted      0                  0                  0
DJI.Adjusted           0                  0                  0
x0050.Tw.Adjusted      0                  0                  0
      x4952.Tw.Adjusted DJI.Adjusted x0050.Tw.Adjusted
x4976.Tw.Adjusted      0                  0                  0
x8422.Tw.Adjusted      0                  0                  0
x2458.Tw.Adjusted      0                  0                  0
x4952.Tw.Adjusted      0                  0                  0
DJI.Adjusted           0                  0                  0
x0050.Tw.Adjusted      0                  0                  0

To see confidence intervals of the correlations, print with the short=FALSE option

```

Figure 18: 判斷相關係數與概率

計算出四家公司股票的調整價格和元大 50 以及道瓊的相關係數和機率，利用 `corr.test()` 函數。

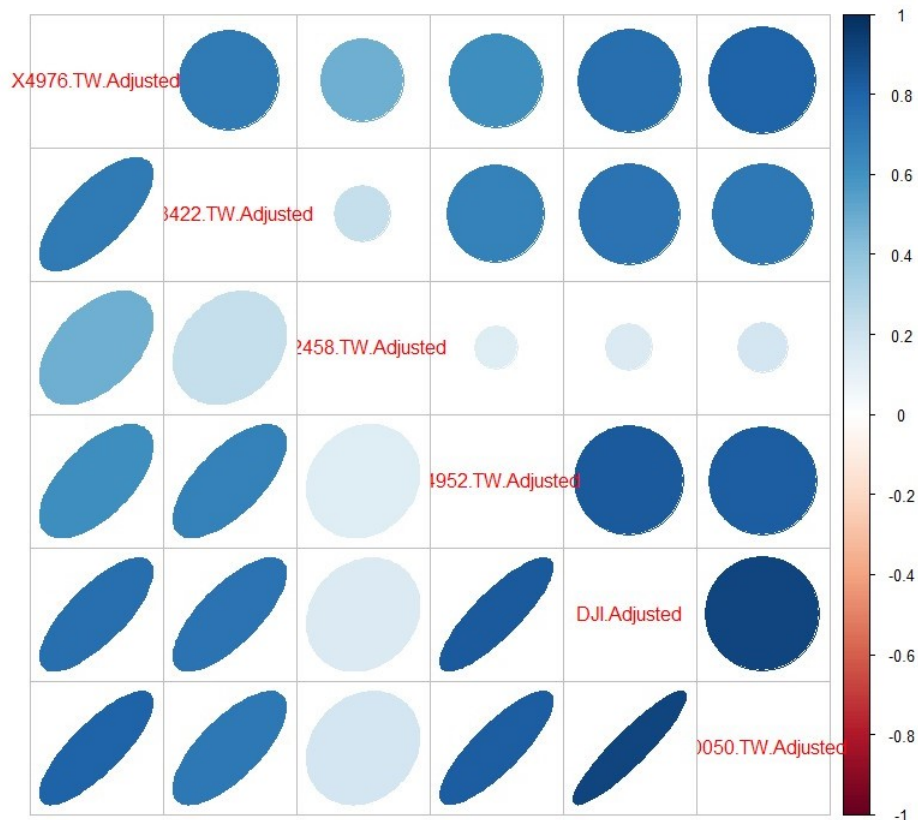


Figure 19: 股票間的相關性

由 (Figure18) 我們可以知道各個股票的調整後股價之間的相關性，他們全部都呈現藍色，所以我們可以知道，他們互相都呈正相關。最明顯的可以觀察到，除了 (2458) 之外都和元大 50(0050.TW) 與道瓊 (DJI) 之間的藍色都偏深，可以知道他們呈現高度相關的趨勢。抑或是藉由左下的圖形也可以觀察出，若圖形越趨向圓形，則相關性就越低 (趨向 0 相關)，反之若圖形較狹長，相關性就較強，圖形之斜率為正的趨勢，則為正相關，反之則為負相關。這是藉由 `corrplot.mixed` 繪製出的相關性圖。

```
> summary(m1m)

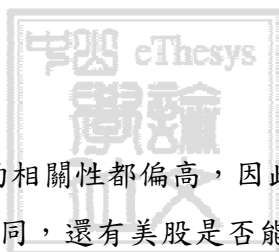
Call:
lm(formula = n$DJI.Adjusted ~ n$x0050.Tw.Adjusted + n$x4976.Tw.Adjusted +
    n$x8422.Tw.Adjusted + n$x2458.Tw.Adjusted + n$x4952.Tw.Adjusted,
    data = n)

Residuals:
    Min       1Q   Median       3Q      Max
-23542.5  -527.5    31.9    528.1   2083.7

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      324.704     816.033   0.398  0.69082
n$x0050.Tw.Adjusted 197.796     12.064  16.396 < 2e-16 ***
n$x4976.Tw.Adjusted  23.485       8.277   2.837  0.00468 **
n$x8422.Tw.Adjusted  18.287       4.063   4.501 7.94e-06 ***
n$x2458.Tw.Adjusted -22.946       8.561  -2.680  0.00753 **
n$x4952.Tw.Adjusted  83.195       8.803   9.450 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1179 on 694 degrees of freedom
Multiple R-squared:  0.865,    Adjusted R-squared:  0.864
F-statistic: 889.1 on 5 and 694 DF,  p-value: < 2.2e-16
```

Figure 20: 總結



由於他們除了 (2458) 之外的相關性都偏高，因此我們很好奇他們之間做迴歸分析的結果是否和我們的想法相同，還有美股是否能由台股去預測，因為台股容易被美股影響加上兩者之間有時差，如果能達到預測效果，可以對股市分析有很大的幫助。

首先，將所有的股票都加入做迴歸分析，得到  $R\text{-squared} = 0.865$ ，之後將 (2458) 這個股票拿掉，會得到  $R\text{-squared} = 0.8636$ ，雖然差異性不大，但是我們認為將一個相關性較低的變數拿掉之後，應該是會上升，但結果和我們的想法不一樣，反而是下降；還有反而是 (4976) 對道瓊工業指數較沒有明顯影響力，並非相關性較小的 (2458)。然而，在用 AIC 去檢測模型，也是發現使用所有的變數時的值較刪除 (2458) 低，即原本的模型較好。

```
> summary(m1m)

Call:
lm(formula = n$DJI.Adjusted ~ n$X0050.Tw.Adjusted + n$X4976.Tw.Adjusted +
    n$X8422.Tw.Adjusted + n$X4952.Tw.Adjusted, data = n)

Residuals:
    Min       1Q   Median       3Q      Max
-23552.3  -590.6    41.6    583.0   1998.0

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1042.621     639.741   -1.630    0.104
n$X0050.Tw.Adjusted  208.407     11.446   18.207 < 2e-16 ***
n$X4976.Tw.Adjusted   10.859       6.836    1.588    0.113
n$X8422.Tw.Adjusted   18.953       4.073    4.653 3.92e-06 ***
n$X4952.Tw.Adjusted   81.334       8.815    9.227 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1184 on 695 degrees of freedom
Multiple R-squared:  0.8636,    Adjusted R-squared:  0.8628
F-statistic: 1100 on 4 and 695 DF,  p-value: < 2.2e-16
```

Figure 21: 總結 without 2458

```
> shapiro.test(SH$name)

shapiro-wilk normality test

data:  SH$name
W = 0.55262, p-value < 2.2e-16
```

Figure 22: Shapiro-Wilk Normality Test

然而，當我們建立出一個線性回歸時，必須要確認其殘差 (residual) 是否符合下面三個假設：

1. 常態性 (Normality)
2. 獨立性 (Independence)
3. 變異數同質性 (Homogeneity of Variance)

由 Figure21，虛無假設  $H_0$ : 殘差服從常態分配，因為  $p\text{-value} < 0.05$ ，拒絕  $H_0$ ，所以其不服從常態分佈。

由 (Figure22)，虛無假設  $H_0$ : 殘差變異數具有同質性，因為  $p\text{-value} < 0.05$ ，拒絕



```
> ncvTest(mlm)
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 267.8818    Df = 1    p = 3.285189e-60
```

Figure 23: Score Test For Non-Constant Error Variance

H0，所以殘差變異數不具有同質性。

所以其實我們的線性模型無法使用，因為殘差不符合假設。



## 第 4 章 結論與後續工作

在做完一系列的股價研究後，我們了解了一些事情。首先是一星二陽的部分，由於一星二陽策略本身屬於一種順勢交易，因此進場點希望是越靠近最初上漲的時機越好。但一星二陽的發生即代表前三天已經有一大段漲幅，令我們離最佳入場點永遠有三天以上的距離，因此單靠一星二陽作為入場依據確實不夠有力，整理出的台股績效表也反映了這一點。績效表中的數據顯示，820 支數據完整的台股中，平均報酬率大於零的有 373 支，勝率超過 50% 的有 145 支，這樣的成績很明顯是無法賺錢的。透過新聞的分析，會發現股票的基本面對上漲的時機點有極大的影響，因此我們認為好的程式交易除了技術面外，還得搭配研究一家公司的基本面才能有較完善的交易策略。

而在研究 R 的過程中，我們也遇到許多困難。像是資料型態的問題，我們常常因為資料型態的誤用而導致程式碼產生 error，又因為我們對 R 的熟悉度不足導致這些 error 要處理很久。另一方面，我們一開始想要利用 R 連結 python 抓取台股股價資料，因為 python 比較適合進行網路爬蟲，可以抓取更多資訊；但中途一直遇到環境設定問題且一直無法解決，只好改回使用 R 內建的套件，中間浪費了許多時間。

還有股價的分析並沒有我們想像中的那麼容易，首先股價是一個隨時間變動的資料，所以若是使用即時股價去分析會非常困難。還有會影響股價的因素有太多了，我們不能就用幾支股票就去分析出來。雖然我們做出的線性模型不能使用，但我們從中也學習到許多資料預處理的方法，像是去視覺化資料，還有遺漏值的處理等等，這部分占了我們這份報告花費時間的最大部分，所以資料分析並非僅僅分析資料如此而已，這次的報告讓我們深深的瞭解到這件事。

## 參考文獻

- a. 蔡立崙 (2017)。金融科技實戰：R 語言與量化投資。臺灣: 博碩。
- b. 王昭文 (2017)。R 軟體財務投資應用：一星二陽 K 棒組合回測程式範例。檢自 [http://cm.nsysu.edu.tw/msrc/wp/rmarkdown/一星二陽 K 棒組合程式範例.html](http://cm.nsysu.edu.tw/msrc/wp/rmarkdown/一星二陽K棒組合程式範例.html)(2018 年 6 月 26 日)
- c. 台灣資料科學年會 (2017)。使用 R 語言建立自己的演算法交易事業。檢自 [https://www.slideshare.net/tw\\_dsconf/r-71023024](https://www.slideshare.net/tw_dsconf/r-71023024)(2018 年 6 月 26 日)

