

投稿類別：資訊類

篇名：

Big Data：由財務數據評估股票的投資價值

作者：

黃惟。國立新竹女子高級中學。高二 12 班

指導老師：

古佳怡老師

鐘培碩老師

壹、前言

一、研究動機

每年農曆春節從長輩手裡領到的紅包，父母讓我自由選購想要的玩具，剩下的便存進我的銀行戶頭裡，幾年下來也累積了一筆不小的數字。購物的過程，父母會教我分辨「價格與價值」、「想要和需要」的差異，除了物質帶來的幸福感，同時也建立正確的消費觀與儲蓄觀。

小學高年級暑假，媽媽送我一本有趣的書「富爸爸，窮爸爸」（Kiyosaki，1997），從中我學習到資產與負債的觀念，也發現自己存在銀行的錢，微薄的定存利率難以對抗通膨，實質購買力有不斷貶值的風險。就算今天吃得起雞腿飯，老了以後可能連滷肉飯都負擔不了。於是，我對大人們口中的投資與股票燃起興趣，拜託父母帶我去開了一個證券帳戶，但這時還是不懂該買什麼股票才能讓一流的企業家天天幫我賺錢。

媽媽建議我從日常的食衣住行觀察，尋找投資標的。我發現兩位阿嬤（奶奶與外婆）的冰箱裡總是有豆腐，而且常來自「中華食品」這個品牌。這家公司有幾個特點：豆腐營養美味，人人吃得起，需求不受景氣影響；保存期限短，不易遭受進口產品攻擊；中華食品是台灣市場的領導品牌。但，我還是不知道中華食品的股價多少才合理？

後來我找到另一本書「圖解新制財報選好股」（羅澤鈺，2013），學習如何從公司財報觀察企業的營運績效，以及隱藏的財務風險。經過計算，我決定買進中華食品的股票，幾年下來，不只配息足以幫家裡吃的豆腐買單，也能同享公司獲利成長的果實，配股、配息加上股價上漲，獲利可觀（圖一、圖二），我對投資的興趣也更濃厚了。

接下來的難題是，台灣的內需消費市場多由中小型企業供應，其中許多不是上市櫃公司，僅從日常生活觀察，能找到的投資標的有限。但要從公開證券市場挑選股票（超過 1600 家），宛如大海撈針，踩到地雷、血本無歸的風險不小。

於是我開始思考，是否能運用程式代替人力，快速地「閱讀」所有公司的財報數據，並設計數學公式作為篩選機制。如此一來，接下來的質性分析工作就會輕鬆不少。

二、研究目的

利用程式下載、處理台灣上市櫃公司的歷年財務報表、月營收數據及股價資訊。接著藉由資料分析與建立計算模型，預測企業未來獲利並評估股票價值。最後，套用股票價值預測模型於股市歷史資料，進行回測與驗證。

下單	明細	股名	種類	股數	成交均價	投資成本	現價	帳面收入	損益	損益率	損益平衡價
下單	明細	中華食	現股	3,330	39.46	131,586	96	318,267	186,681	141.87%	39.7
下單	明細									29.77%	16.11
下單	明細									25.39%	23.25
下單	明細									71.56%	80.94

圖一：研究者的部分投資明細
（圖一資料來源：研究者的證券帳戶截圖）



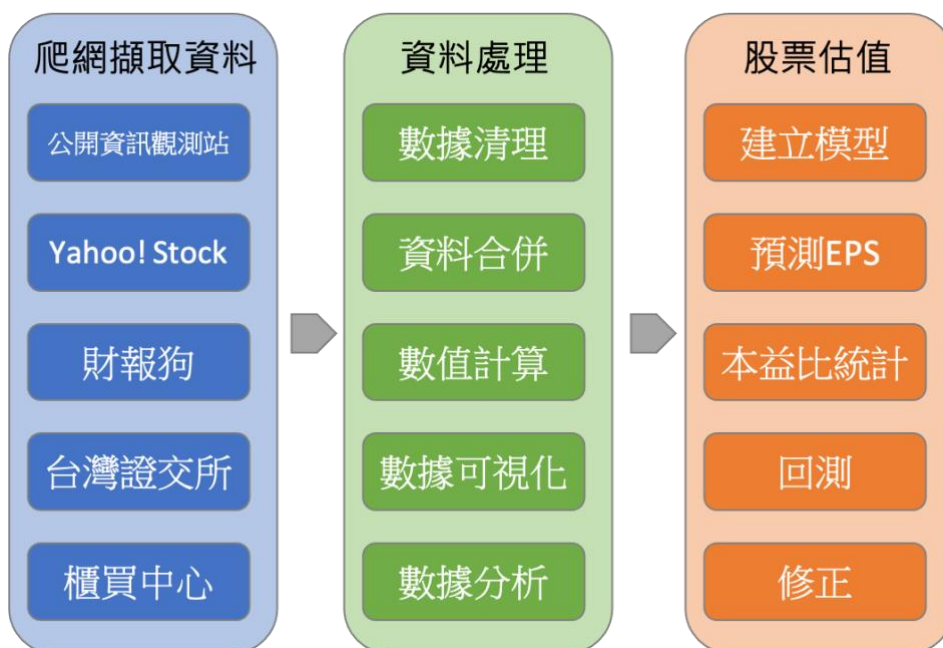
圖二：中華食品（股票代號：4205）的歷史股價。
（圖二資料來源：Yahoo! 股市，<https://tw.stock.yahoo.com/>）

三、研究方法

本論文中以 Python 語言為實作基礎，搭配幾個開源套件：Requests、NumPy、Pandas、Matplotlib。以程式爬網擷取資料、進行數據分析，再設計公式計算股票價值。

- (1) Requests：爬網擷取資料。
- (2) NumPy：處理數學計算。
- (3) Pandas：資料清理、格式化、分析及運算。
- (4) Matplotlib：資料的可視化呈現。

四、研究架構



圖三：研究架構
（圖三資料來源：研究者繪製）

貳、正文

一、價值型投資概述

股神巴菲特終身實踐其恩師班傑明葛拉漢傳授的價值型投資，這也是最能說服我的投資（而非投機）理論。因此，本研究將以價值型投資心法為基礎來估算股票價值。

價值型投資的秘訣，簡單卻重要：用便宜的價格買進好公司的股票並長期持有，享受時間帶來的複利效應（Graham, 1949）。因為買進的是穩定成長的好公司並且長期持有，身為股東可以透過每年的配息配股，共享公司經營的成果。當公司的價值獲得市場認可時，股價便會隨之上漲。此時可以繼續持有，也可以賣出股票收割資本利得。但是，困難的是，怎麼判斷一家公司好或不好？什麼樣的價格才叫便宜？

本研究將從財務報表出發，檢視各項財務指標與股價的相關性，進一步預測公司的獲利能力，據此估算股票價值。

財務報表包含四大表：「綜合損益表」、「資產負債表」、「現金流量表」、「股東權益變動表」。分析綜合損益表可以幫助我們瞭解一家公司的獲利能力與營運績效；分析資產負債表和現金流量表，則能檢查其財務體質是否健全，避免買進地雷股；股東權益變動表則可以幫助我們瞭解股東整體資金的獲利效率。

要判斷一家公司的獲利能力，不外乎檢查下列四個重要的財務指標，本研究將全部採納考慮。

（1）營收：企業因銷售產品或提供服務而取得的收入。

（2）毛利：營收扣除銷貨成本（直接成本）。可觀察企業生產成本的變化，也可檢查公司是否遭逢惡性競爭。

（3）營業利益：毛利扣除營業費用（間接成本），即本業帶來的利益。營業費用包含廣告、薪資、研發等費用。當營業費用增加，營業利益會下滑。每間公司的營運費用佔比不盡相同，因此營業利益率（營業利益除以營收）的走勢比數值大小重要。

（4）每股盈餘 (Earnings per Share, EPS)：扣除成本、費用、業外損益、營業稅後，每單位資本額的獲利能力。通常 EPS 成長率高的公司，享有的本益比倍數也越高。

$$\text{本益比} = \frac{\text{股價}}{\text{EPS}}$$

二、研究步驟

（一）爬網擷取上市櫃公司股票收盤價格

資料來源：Yahoo! 股市 (<https://tw.stock.yahoo.com/>)

程式碼：P1_get_Yahoo_price.py（程式碼如圖四所示，後續的程式碼礙於篇幅，只附上 GitHub 連結）。本程式碼的 GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P1_get_Yahoo_price.py

利用 requests 套件爬取網頁資料，回傳的每一筆資料的各欄位數值分別對應不同數據，{"t":日期, "o":開盤價格, "h":高點, "l":低點, "c":收盤價格, "v":成交量}，可以透過自己寫的 html parser 取出需要的資料，建立每月收盤價格表。

原始資料範例：{"t":20210129,"o":99.0,"h":99.0,"l":94.9,"c":95.7,"v":501}

```

1 import requests
2
3 def get_Yahoo_stock_price(ticker):
4     """抓取Yahoo Stock的個股股價
5     #== ticker: 股票代碼
6     """
7     url = "https://tw.quote.finance.yahoo.net/quote/"
8     url += f"q?type=ta&perd=m"
9     url += f"&mkt=10&sym={ticker}"
10    url += "&callback=jQuery111306117842047895292_1535895637574&_=1535895637575"
11
12    r = requests.get(url)
13    html = r.text
14    print(html)
15    data = html.split('{ "t":')
16    price = {}
17    for i in range(1, len(data)):
18        #== 找出月份與對應的股價
19        date = data[i].split(', "o":')[0][:-2]
20        closing_price = float(data[i].split(', "c":')[1].split(', "v":')[0])
21        price[date] = closing_price
22    return price

```

圖四：研究者的程式碼。P1_get_Yahoo_price.py
(圖四資料來源：研究者的電腦截圖)

(二) 爬網擷取上市上櫃公司每一季的財務報表

資料來源：公開資訊觀測站 (<https://mops.twse.com.tw/mops/web/index>)

程式碼：P2_get_financial_report.py，GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P2_get_financial_report.py

由於公開資訊觀測站上的彙總報表（綜合損益表、資產負債表）是以表格形式呈現在網頁，可以直接用 pandas.read_html 將表格逐季抓取下來，把上市及上櫃公司的同期資料合併在一起，再利用 pandas.to_csv 轉存成 csv 格式，逐季存檔，方便後續使用。

(三) 萃取個股的財報歷史資料，同時計算各項財務比例

程式碼：P3_cal_fin_idx.py，GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P3_cal_fin_idx.py

接續步驟（二），從各季度的財報匯總檔案中，找出對應指定股票代號的資料，順便計算出各種損益率，再合併成單一檔案，即為個股的歷史財報資料。表一以中華食品（股票代號：4205）為例，呈現該公司的部分歷史財報資料表。

	20194	20201	20202	20203
營業收入（千元）	413,320	409,278	389,138	426,234
營業毛利（千元）	160,794	166,124	164,273	181,876
營業費用（千元）	67,399	68,166	66,103	67,655
營業利益（千元）	93,395	97,958	98,170	114,221

營業外收入及支出（千元）	748	1307	788	10,888
稅前淨利（千元）	94,143	99,265	98,958	125,109
所得稅費用（千元）	18,504	19,853	19,892	25,022
本期淨利（千元）	75,639	79,412	79,066	100,087
基本每股盈餘（元）	0.94	0.98	0.98	1.24

表一：中華食品的歷史財報資料表
（表一資料來源：研究者的程式執行結果）

（四）分析財務指標與股價的相關性

程式碼：P4_cal_corr.py，GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P4_cal_corr.py

我們以三家公司：中華食品（股票代號：4205）、台積電（股票代號：2330）和聯發科（股票代號：2454）為例，採用從 2018 年第 1 季到 2020 年第 3 季的財報資料，經由程式碼 P4 計算出各項財務指標與股價的相關係數。結果整理於表二。

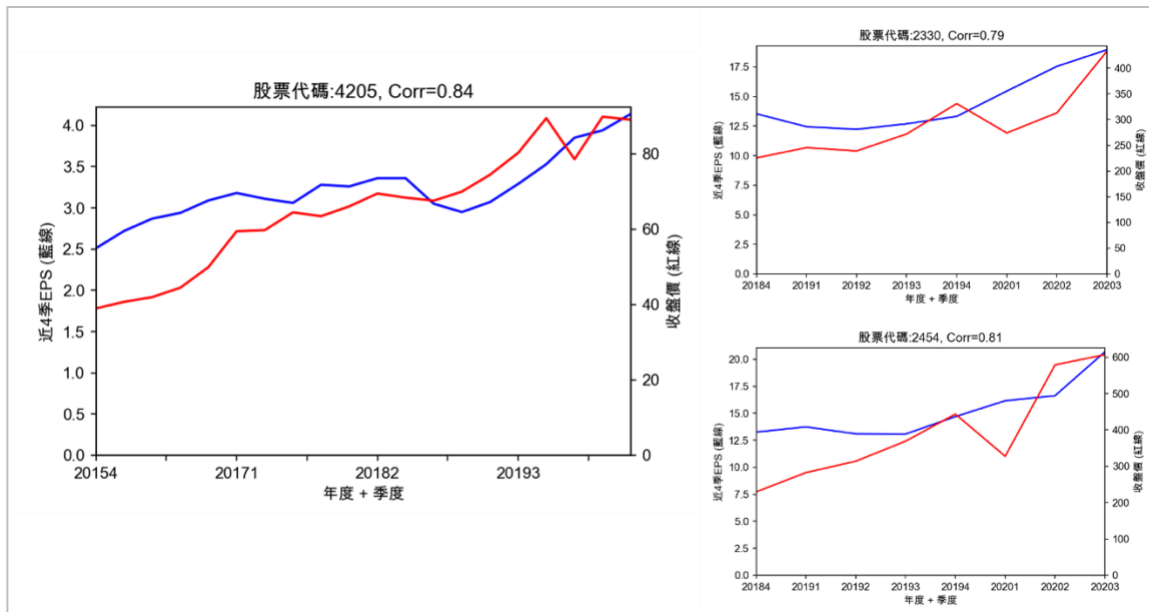
排名	4205 中華食品		2330 台積電		2454 聯發科	
	會計科目	相關係數	會計科目	相關係數	會計科目	相關係數
1	所得稅費用（利益）	0.86	營業費用	0.92	營業利益率	0.85
2	營業收入	0.86	營業毛利（毛損）	0.79	毛利率	0.84
3	營業毛利（毛損）	0.86	營業收入	0.79	本期淨利（淨損）	0.83
4	營業利益（損失）	0.85	近 4 季 EPS	0.79	稅後淨利率	0.83
5	近 4 季 EPS	0.84	稅前淨利（淨損）	0.77	基本每股盈餘（元）	0.82
6	稅前淨利（淨損）	0.81	營業利益（損失）	0.77	營業毛利（毛損）	0.82
7	本期淨利（淨損）	0.78	基本每股盈餘（元）	0.76	營業利益（損失）	0.82
8	營業利益率	0.76	本期淨利（淨損）	0.76	稅前淨利（淨損）	0.82
9	稅前淨利率	0.70	毛利率	0.70	營業費用	0.81
10	營業費用	0.69	營業外收入及支出	0.70	近 4 季 EPS	0.81
11	毛利率	0.68	所得稅費用（利益）	0.68	稅前淨利率	0.78
12	基本每股盈餘（元）	0.67	稅前淨利率	0.66	營業收入	0.77
13	稅後淨利率	0.63	營業利益率	0.65	所得稅費用（利益）	0.68
14	所得稅率	0.61	稅後淨利率	0.64	營業費用佔營收比	-0.61
15	營業費用佔營收比	-0.31	業外收支佔稅前淨利比	-0.46	業外收支佔稅前淨利比	-0.43
16	營業外收入及支出	0.18	營業費用佔營收比	0.20	營業外收入及支出	0.09
17	業外收支佔稅前淨利比	0.10	所得稅率	-0.15	所得稅率	-0.03

表二：財務指標與股價的相關係數
（表二資料來源：研究者的程式執行結果）

由表二觀察，財務指標與股價的相關係數排名，因公司而異，缺少具有普遍代表性的單一指標。可能原因如下：個股的營業收入與營業毛利會受淡旺季週期影響，營業費用與產業屬性有關，稅前淨利易受業外收益左右，稅後淨利可能與政府的研發獎勵或補助款有關，基本每股盈餘（EPS）則是所有營運績效的綜和結果，但受淡旺季週期影響。

為了排除淡旺季週期影響，我們加入一個自行計算的衍生性指標「近 4 季 EPS」，其為本季與過去三季的每股盈餘總和，發現近 4 季 EPS 與股價的趨勢一致（圖五），其

與股價的相關係數，不分企業，均排名在前（表二）。因此，本論文選擇近 4 季 EPS 作為預測的對象，應可排除單一財務指標造成的變異。



圖五：三家公司的近 4 季 EPS 與股價的趨勢對照圖
（圖五資料來源：研究者的程式執行結果）

（五）股票價值預測

程式碼：P5A_get_mth_revenue.py，GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P5A_get_mth_revenue.py

程式碼：P5B_cal_EPS.py，GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P5B_fcst_EPS.py

程式碼：P5C_get_SDog_PER.py，GitHub 連結如下：

https://github.com/YuiHuang/Report_2021_03/blob/main/P5C_get_SDog_PER.py

財務報表呈現的是企業的實際營運績效，而股票價格反應的是投資人對企業的期待。本論文利用已經發佈的企業財報，結合最近 3 個月的月營收數字（EPS 屬未公開資訊），預測下一個週期的近四季 EPS（例：2020 年 5 月時，預估 2019 年第 3 季到 2020 年第 2 季的 EPS 總和），並據此推估前瞻股票價值。

在此我們假設企業的營運（包括其產品佈局、製造成本及營銷費用、所得稅率、同業競爭、其它上下游關係）不會在短時間內出現巨大變化，因此可以運用下列自行設計的公式來預測 EPS 變化。

$$EPS_{new} = \frac{\text{近 4 季 EPS} \times (1 + \text{最近 3 個月營收 YoY}) \times [1 + \min(GM \text{ YoY}, OPM \text{ YoY})]}{(1 + \text{最近 4 季平均股本 YoY})}$$

（註）YoY = year-on-year change，比之去年同期的增減百分比

在產品的銷售毛利及企業的營運費用維持不變的情況下，營收（Revenue）的增減會直接影響 EPS。本論文利用程式碼 P5A 爬網下載月營收資訊，接著計算以下公式。

$$\text{最近 3 個月營收 YoY} = \frac{\sum_{i=1}^3 Rev_{m-i} - \sum_{i=1}^3 Rev_{m-12-i}}{\sum_{i=1}^3 Rev_{m-12-i}}, Rev = \text{月營收}, m = \text{本月}$$

成本結構或外部競爭改變時，會影響毛利率（Gross Margin，GM）。管理、銷售、研發等費用的變化則會影響營業利益率（Operating Margin，OPM），考量毛利率和營業利益率的短期變化通常不大，本論文採用簡化的方式，以前三季毛利率的平均值來預估下一季的毛利率，營業利益率也以相同方式計算，再取毛利率和營業利益率的最小值對 EPS 進行預測，避免高估。如此可保有預測失準時的安全邊際（margin of safety）。

$$GM \text{ YoY} (\%) = \frac{\frac{\sum_{i=1}^3 GM_{s-i}}{3} - GM_{s-4}}{GM_{s-4}}, s = \text{本季}$$

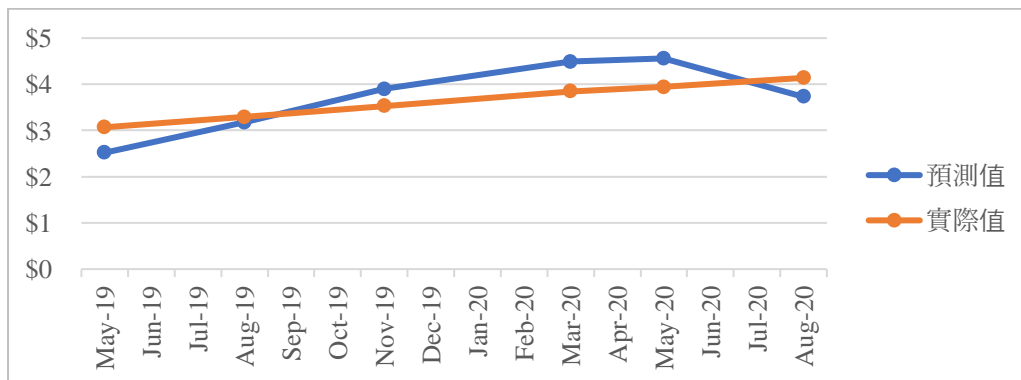
$$OPM \text{ YoY} (\%) = \frac{\frac{\sum_{i=1}^3 OPM_{s-i}}{3} - OPM_{s-4}}{OPM_{s-4}}, s = \text{本季}$$

此外，股本的變化（比如增資或減資）會直接稀釋或放大 EPS。本論文在預測 EPS 時，也一併將之列入考慮。

表三以筆者持有的中華食品（股票代號：4205）為例，依上述公式，執行程式碼 P5B。基於財報揭露的績效，搭配最新的月營收資料，預測其下一個週期的近 4 季 EPS。估算出的結果與事後官方公佈的數值趨勢吻合，如圖六所示。

公司	預測日期	近 3 月 營收和 YoY (%)	毛利率 YoY (%)	營業 利益率 YoY (%)	股本 YoY (%)	EPS 預測值	下季公告 EPS 實際值
中華食品 (4205)	2019-05-15	6.6	-7.9	-14.5	0	\$ 2.52	\$ 3.07
	2019-08-14	7.1	-3.2	-2.6	0	3.18	3.29
	2019-11-14	5.6	11.9	23.2	0	3.90	3.53
	2020-03-31	7.4	18.5	37.3	0	4.49	3.85
	2020-05-15	4.8	3.2	3.4	0	4.56	3.94
	2020-08-14	3.5	-3.3	-5.6	0	3.73	4.14

表三：中華食品近 4 季 EPS 預測值與幾個月後公佈的實際值比較
（表三資料來源：研究者的程式執行結果）



圖六：中華食品近 4 季 EPS 預測值與幾個月後公佈的實際值比較
（圖六資料來源：研究者的程式執行結果）

投資人對一家公司的評價（喜愛或厭惡的程度）反應在本益比（P/E ratio, price-to-earnings ratio），本益比不是一個固定數值，其與產業別有關，也受企業形象影響。

因此，我們預測一家企業的近 4 季 EPS 後，還要決定合理的本益比，才能因時因地制宜地評估市場如何看待一檔股票的價值。圖七以中華食品為例，以程式碼 P5C 下載並畫出其過去五年的本益比走勢圖，可以看出本益比隨著時間常有劇烈的波動。

在此引入統計學的觀念來描述本益比的特性，避免被離群值（特別高或特別低的數值）誤導，並採用第 1、2、3 四分位數當作決策標準。本益比的統計結果整理於表四，股票價值估算結果整理於表五。由於 EPS 的預測可能不完全準確，如果用比表五中建議的買進價格還低的價位購入股票，就會有更大的安全邊際來守護資產。



圖七：中華食品的五年本益比趨勢

（圖七資料來源：財報狗，<https://statementdog.com/>）

	第 1 四分位數	第 2 四分位數	第 3 四分位數
自 2020 年回溯五年	17.3	20.2	22.8
自 2016 年回溯五年	16.5	20.0	21.8

表四：中華食品歷史本益比統計分析

（表四資料來源：研究者的程式執行結果）

公司	預測日期	EPS 預測值	買進 (17.3 倍)	持有 (20.2 倍)	賣出 (22.8 倍)	股票市價	評價
中華食品 (4205)	2020-03-31	\$ 4.49	\$ 77.68	\$ 90.70	\$ 102.37	\$ 78.2	持有
	2020-05-15	4.56	78.89	92.11	103.97	89.6	持有
	2020-08-14	3.73	64.53	75.35	85.04	89.7	賣出

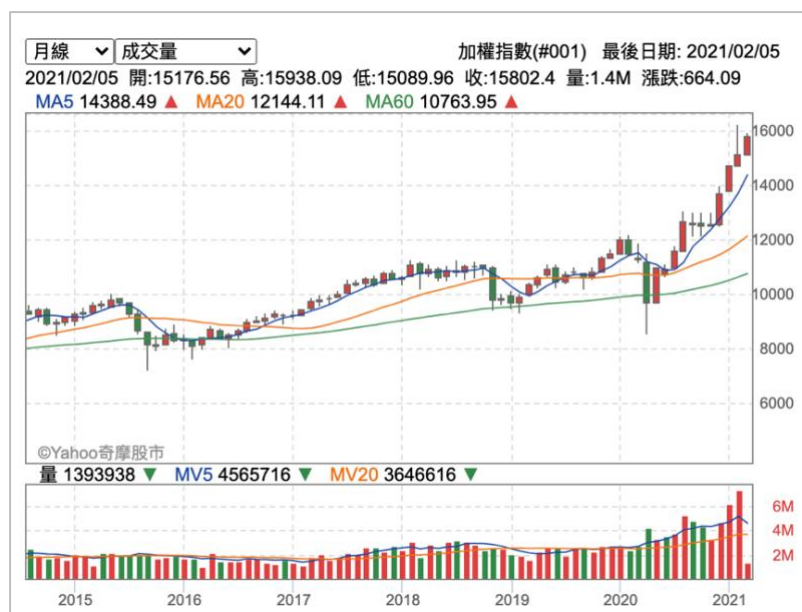
表五：中華食品股票價值評估

（表五資料來源：研究者的程式執行結果）

台股大盤指數自從 2017 年 5 月重登萬點後，至今維持多頭走勢（圖八），個股的本益比在投資人的追捧下，頻創新高。對於價值投資人而言，在行情過熱的情況下，找不到便宜的買點為正常現象，需耐心等待市場回復理性。

(六) 回測與驗證

觀察圖八的台股大盤加權指數後，筆者不禁思考，倘若時間能倒流回到 2016 年，本論文前述步驟的做法是否能夠幫助投資人獲利呢？再度以中華食品（股票代號：4205）為例，重新執行程式碼 P5B 與 P5C，並將計算結果整理於表六。



圖八：台股大盤加權指數歷史趨勢

（圖八資料來源：Yahoo! 股市，<https://tw.stock.yahoo.com/>）

公司	預測日期	EPS 預測值	買進 (16.5 倍)	持有 (20.0 倍)	賣出 (21.8 倍)	股票市價	評價
中華食品 (4205)	2016-05-15	\$ 2.79	\$ 46.04	\$ 55.80	\$ 60.82	\$ 41.15	買進
	2016-08-14	2.60	42.90	52.00	56.68	43.95	持有
	2016-11-14	3.35	55.28	67.00	73.03	48.80	買進

表六：中華食品股票價值估算

（表六資料來源：研究者的程式執行結果）

由表六可以看出，在 2016 年的 5 月、8 月、11 月，每季財報發佈後，立即用程式碼評估股票的價值，5 月和 11 月時都分別出現買進訊號（收盤價格低於模型預估的買進價格）。假設投資人在 2016 年 5 月，以每股 \$41.15 元購入 1,000 股中華食品的股票（成本 \$41,150），並持有至今（2021 年 2 月 5 日收盤價格 \$95.5），由於中華食品曾經在 2017 年，為了擴大產能，盈餘轉投資，每股配發股票股利 1.1 元，相當於投資人的持有股數增加為 1,110 股，期間還領到五次現金股利（共計 \$10,160 元），投資期間總報酬率 182%，年化報酬率 23%，細節彙整於表七。

	2016	2017	2018	2019	2020
持股數	1,000	1,000	1,110	1,110	1,110
每股配息	\$ 2.0	\$ 1.5	\$ 3.0	\$ 2.8	\$ 3.0
每股配股	0	1.1	0	0	0
領取股息	\$ 2,000	\$ 1,500	\$ 3,330	\$ 3,108	\$ 3,330

表七：投資中華食品的獲利試算（2016 年 5 月～2021 年 2 月）

（表七資料來源：配息配股數字來自 Yahoo! 股市，<https://tw.stock.yahoo.com/>）

筆者也將模型套用在台積電（股票代號：2330），計算結果整理於表八。可以看出 2016 年 11 月曾出現買進訊號。假設投資人在 2016 年 11 月，以每股 \$180.5 元購入 1,000 股台積電股票（成本 \$180,500），並持有至今（2021 年 2 月 5 日收盤價格 \$632），期間配發四次現金股利（共計 \$35,000 元），投資期間總報酬率 270%，年化報酬率 30%，細節彙整於表九。

公司	預測日期	EPS 預測值	買進 (12.3 倍)	持有 (14.7 倍)	賣出 (15.8 倍)	股票市價	評價
台積電 (2330)	2016-05-15	\$ 10.40	\$ 127.92	\$ 152.88	\$ 164.32	\$ 144.5	持有
	2016-08-14	12.11	148.95	178.02	191.34	179.0	持有
	2016-11-14	15.41	189.54	226.53	243.48	180.5	買進

表八：台積電的股票價值估算
（表八資料來源：研究者的程式執行結果）

	2016	2017	2018	2019	2020
持股數	1,000	1,000	1,000	1,000	1,000
每股配息		\$ 7.0	\$ 8.0	\$ 10.0	\$ 10.0
領取股息		\$ 7,000	\$ 8,000	\$ 10,000	\$ 10,000

表九：投資台積電的獲利試算（2016 年 11 月～2021 年 2 月）
（表九資料來源：配息配股數字來自 Yahoo! 股市，<https://tw.stock.yahoo.com/>）

參、結論

本論文從價值型投資出發，透過程式發掘大數據的魅力與威力。透過觀察財務指標與股價的相關性，決定以衍生性指標「近 4 季 EPS」作為估算股票價值的基礎。接著設計一個「近 4 季 EPS」預測公式，並以歷史本益比的統計值來評估股票價值。以程式全程實作，並分別以 2020 年與 2016 年為例，進行回測驗證。有合理的數學計算為基礎，佐以電腦程式的效率，要理性、有紀律地實踐價值型投資，不隨群眾在股海中追漲殺跌，要抵抗人性的貪婪和恐慌也變得容易許多。

礙於時間及篇幅，很多值得研究的面向筆者力有未逮，希望未來能進一步拓展，讓程式自動閱讀分析所有企業的季報及月報，不僅節省時間，也能進行大規模的回測與檢討，進一步優化模型。評估股票價值重要，掃雷（排除財務體質不好的公司）的工作也不遑多讓，如何有效評估企業的財務體質是否健全，也可以作為下一階段的學習重點。此外，本論文中提出的公式，是由試誤法（trial and error）得來，未來研讀機器學習（machine learning）演算法後，或許能發掘更巧妙的方式，進一步提升股票價值評估的準確性和可靠度。

肆、引註資料

楊軍、楊明（譯）（2001）。**富爸爸，窮爸爸**（原作者：Robert T. Kiyosaki，Sharon L. Lechter）。台北市：高寶書版。（原著出版年：1997）

羅澤鈺（2013）。**圖解新制財報選好股**。台北市：商周出版。

俞濟群（譯）（1996）。**智慧型股票投資人**（原作者：Benjamin Graham）。台北市：寰宇出版。（原著出版年：1949）