第三章 研究步驟

第一節 資料簡述

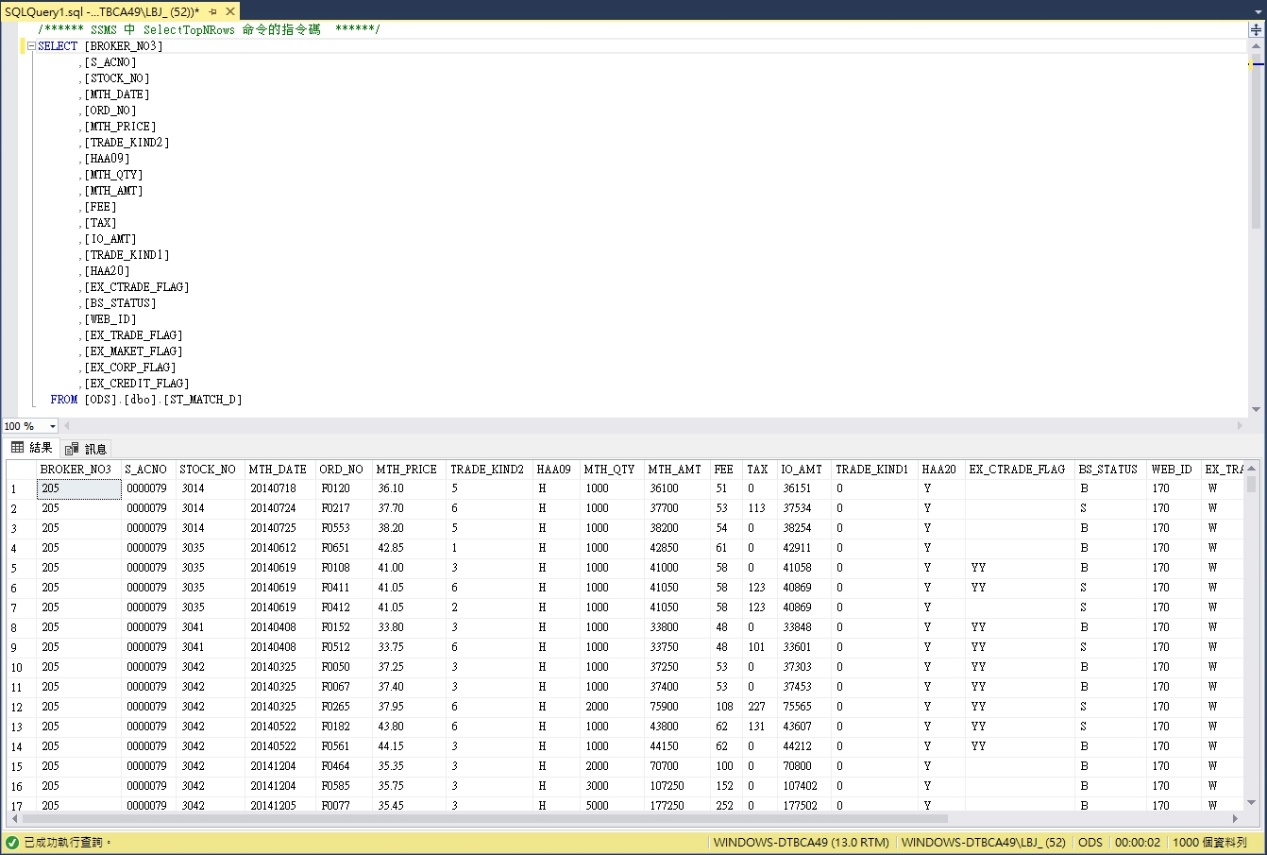
本研究利用X公司之2013年及2014年完整交易資料，本資料含有22個欄位，根據下列圖示，以我們會使用到的資料作說明：

|  |  |
| --- | --- |
| 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。

第二節

首先，由於數據資料過於龐大，以經超出了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。



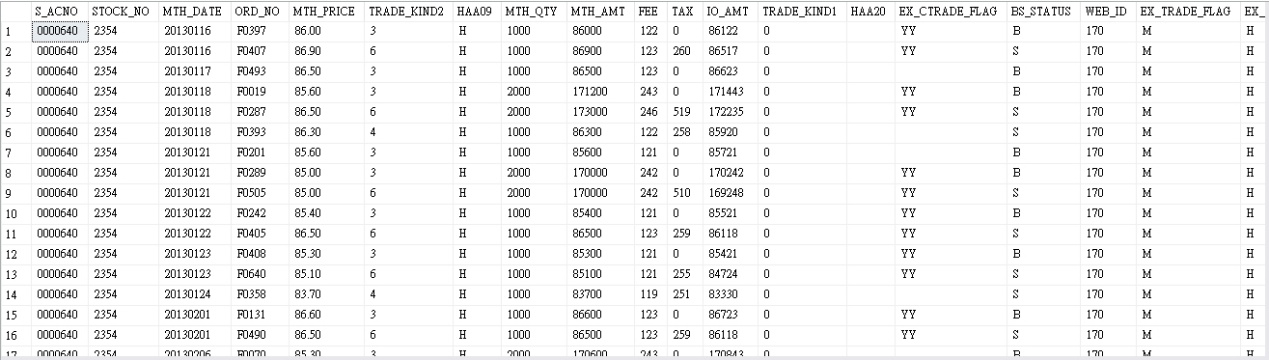
2.我們使用 SQL語法中的LIKE語法，目的在於找尋交易記錄中日期欄位開頭為2013以過濾出2013年度的交易資料(這邊我們選擇以2013做分析，因在X公司的資料中2013是最為其全的)，由圖可見，查詢已過濾為2013開頭的資料，並藉由INTO語法使該查詢另外產生一個名為Trade2013的資料庫。

SELECT \*

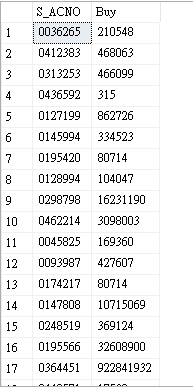
INTO　Trade2013

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

WHERE [MTH\_DATE] LIKE ‘2013%’;



3.再來，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]

4.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],

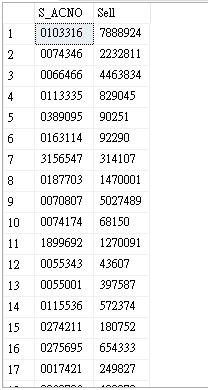
SUM(IO\_AMT) as Sell

INTO Sell2013

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

WHERE [BS\_STATUS]=’S’

GROUP BY [S\_ACNO]

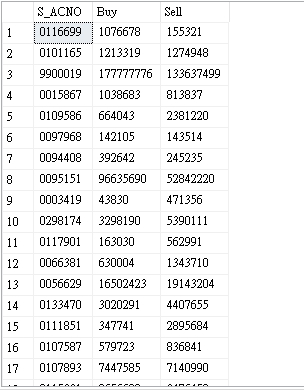


5.為了方便計算，我們分別將資料庫Buy2013的[交易代碼](與Sell2013的欄位相同)，Buy2013的Buy欄位與Sell2013的Sell欄位，條件在兩個資料庫[交易代碼]相同時將Buy、Sell做合併，下圖為輸出結果。

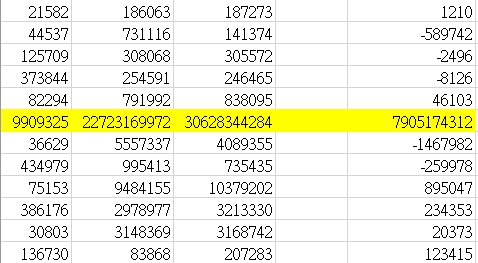
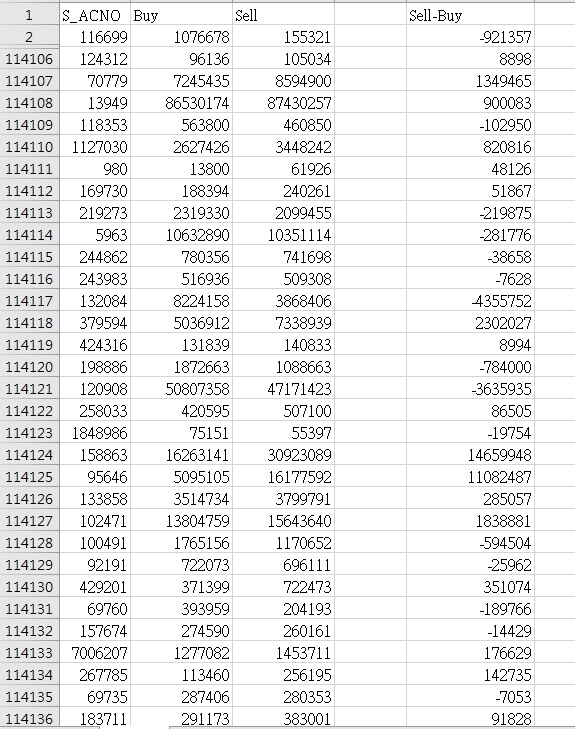
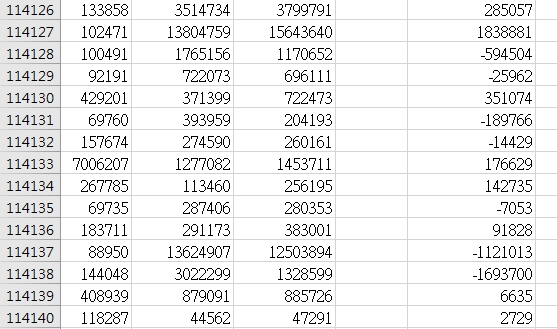
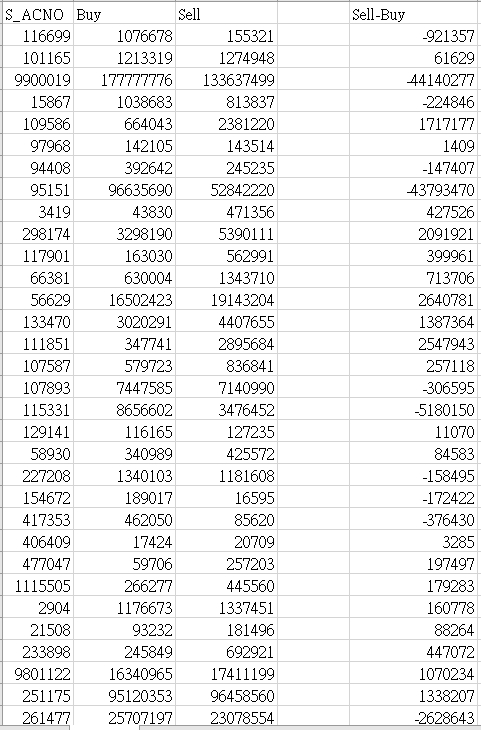
SELECT Buy2013.S\_ACNO,Buy2013.Buy,Sell2013.Sell

FROM Buy,Sell2013

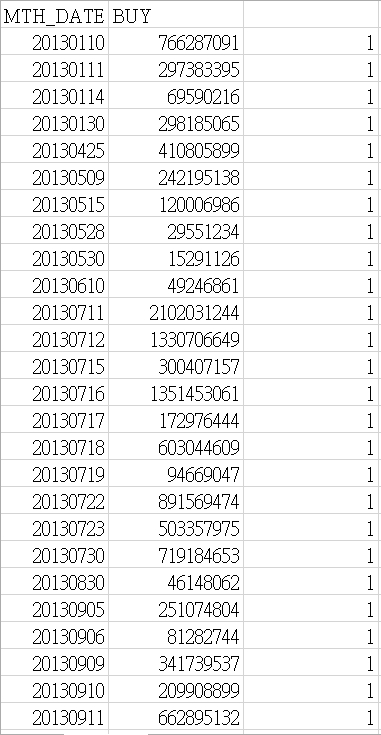
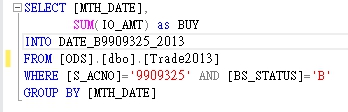
WHERE Buy2013,S\_AC，NO=Sell2013.S\_ACNO

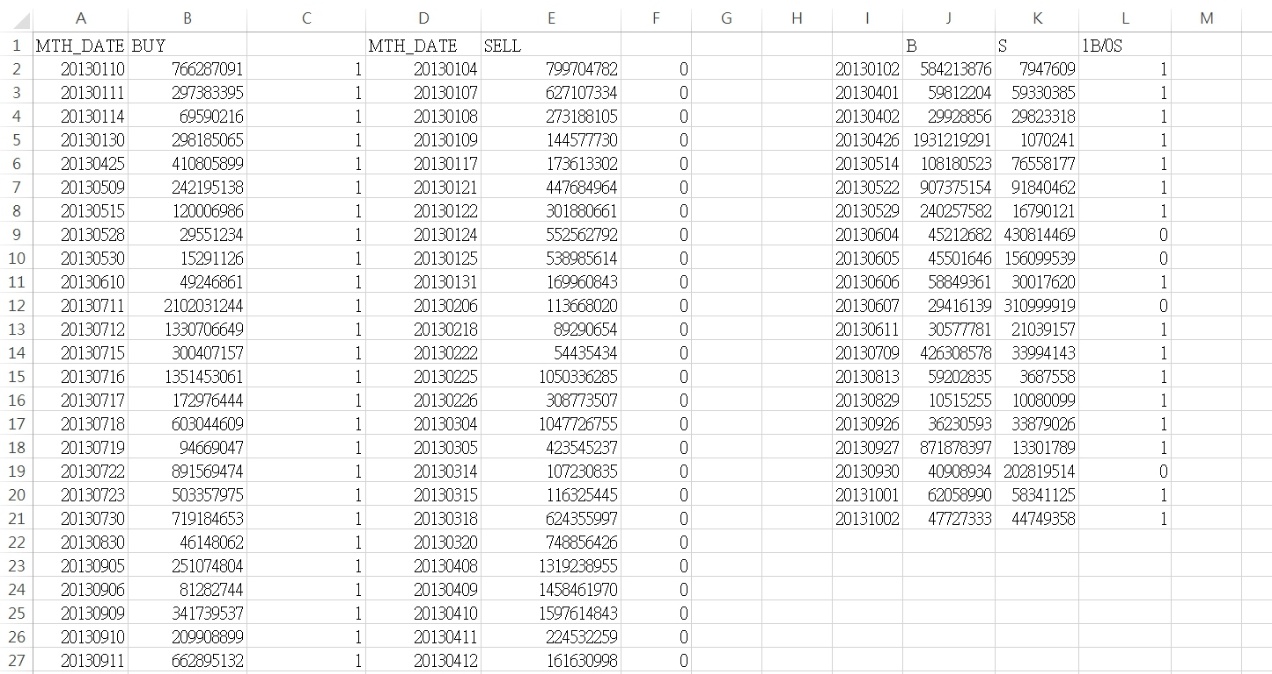


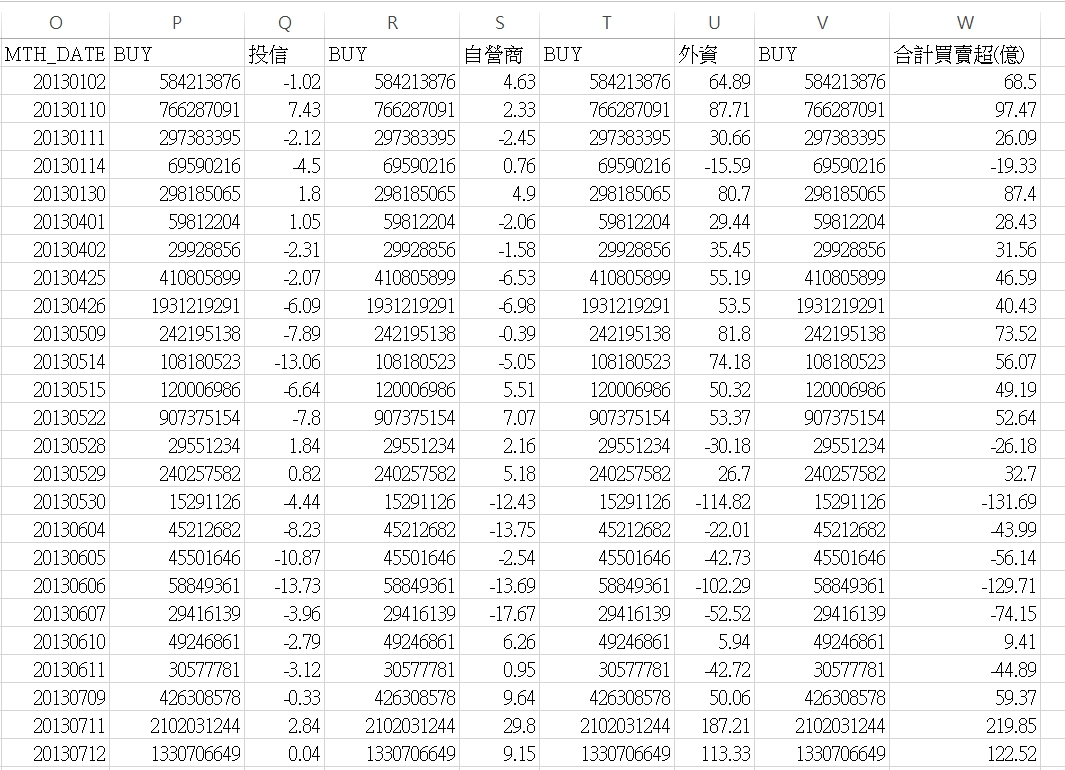
6.為了找出最賺錢的交易人，且我們已經將龐大的數據庫精簡為我們所需的資料，因此在這邊可以放入Excel做數據分析，下圖我們計算出各交易人的獲利(賣-買)且利用MAX語法找出獲利的最大值。如圖，找出最大值7905174312，之後以簡易查詢(Ctrl+F)反求找出獲利為該值的交易人，得到的結果為[交易代號]9909325，在2014同樣方法的驗證中，也同樣顯示最賺錢的為相同代號的交易人，因此我們合理推測，該交易人的成功並非偶然，而是有一套規則。



7.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。

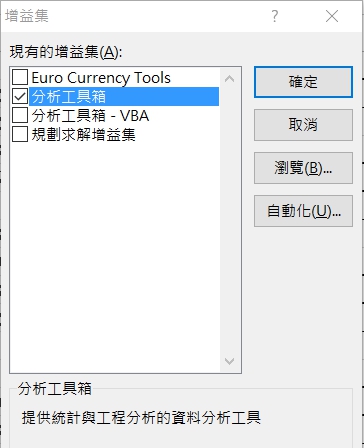


8.至此由於我們已經找出最賺錢的人，後面我們開始尋找資料並開始進行各因素與該交易人所有交易天數的比對(下圖為此人的買賣的金額與日期)

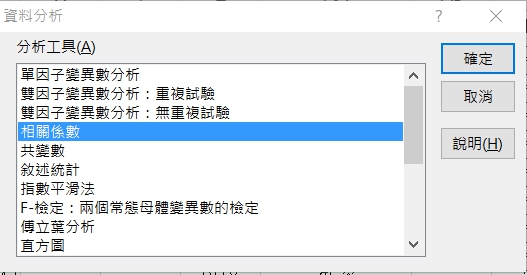
9.以三大法人的各項交易金額為例，同樣以日期做區分進行分析

10.下面介紹分析工具箱的使用

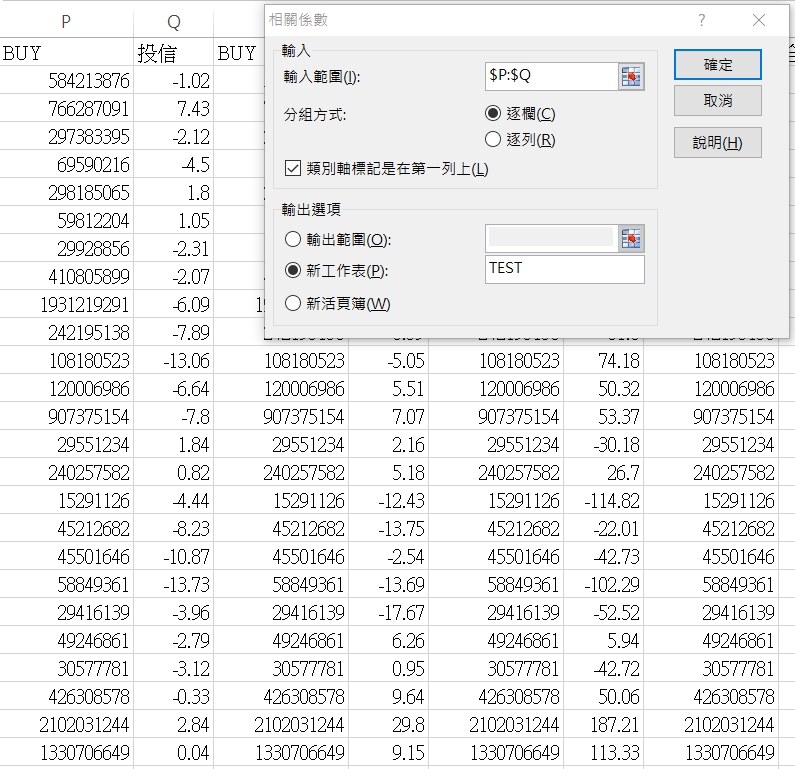
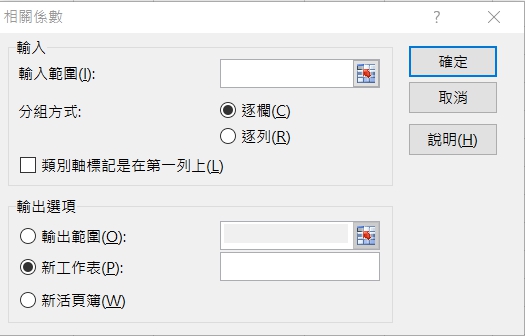
(1)首先先由檔案->選項->增益集->執行，開起分析工具箱如下圖。



(2)開啟後至資料->資料分析開起如下圖



(3)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)



10.相關係數分析後獲得的結果如下，其他因素結果的獲得同理。

