

# Wafer 良率之預測

諮詢者: HAMADA HEAVY INDUSTRIES LTD.

執行研究生: 吳漢銘

日期: 民國八十八年十一月 26 日

- 一 · 研究動機及目的
- 二 · 資料背景
- 三 · 統計分析方法
- 四 · 結論及建議

## 一． 研究動機及目的

企業主從兩家客戶進貨 8 吋晶圓，再把這些晶圓銷出。進貨之晶圓經過磨蝕，會有損耗不能使用的情況發生，爲了控制進銷存貨，供應客戶所需的晶圓片數，企業主必需求得出貨的晶圓片數，使不致於出貨不足或出貨過多的情況發生。因此本報告的目的便是想從所要求的晶圓成品(Products)片數中，去估計進貨(Input)需要多少片。

## 二． 資料背景

企業主從兩家客戶進貨 8 吋晶圓，資料紀錄從 1997 二月到 1999 七月的每月晶圓 Input 和 Products 的片數，將 missing data 去除後，有效的資料如下：

	第一家客戶	第一家客戶
1997	二月~四月，六月~十二月共 10 筆	二月~十二月共 11 筆
1998	一月~十二月共 12 筆	一月~十二月共 12 筆
1999	一月~七月共 7 筆	一月~七月共 7 筆
變數	Input 進貨片數 Products 可用成品片數	

### 三·統計分析方法

首先觀察 data (Input, Products)的時間序列圖(圖 1、圖 2)，基本上 Input 和 Products 呈現正相關的消漲關係，對時間上變化並沒有特別的 pattern 存在。因此要由 Products 來預測 Input 需多少，可利用 Regression 的方法來分析，故可把 Input 當成 response 變數，products 為解釋變數。

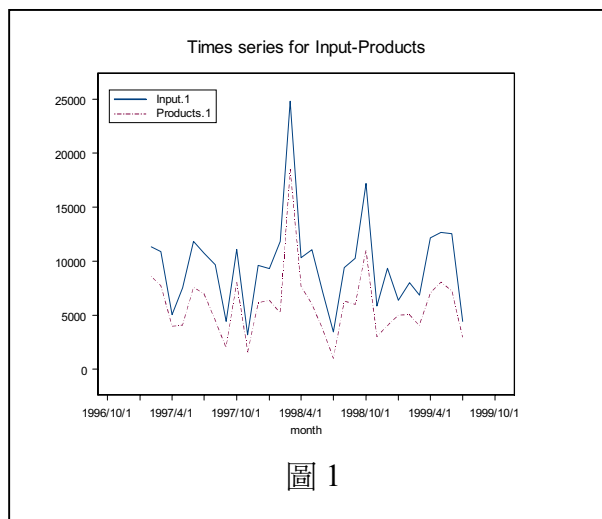


圖 1

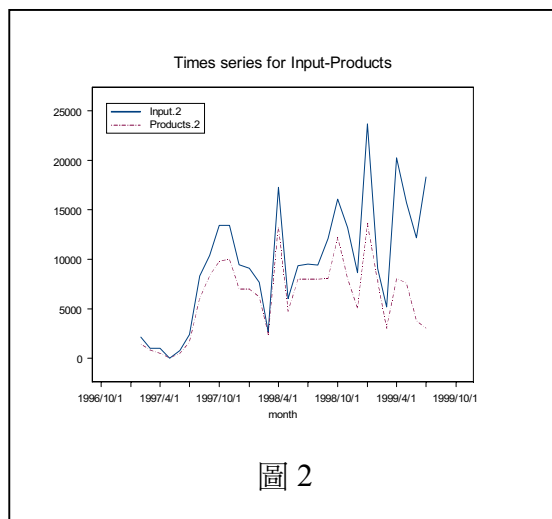


圖 2

#### Case 1: 兩家客戶分別做迴歸分析

圖 3 及圖 4 為兩家客戶的 scatterplot，實線為迴歸線，虛線為 95%的信賴區間。明顯地第一家的 Line Fit 比第二家要好的多。現考慮晶圓的來源不因客戶的不同而不同，所以把兩家資料合併做迴歸分析，圖 5 及表 1 為這個資料 fit 的結果。

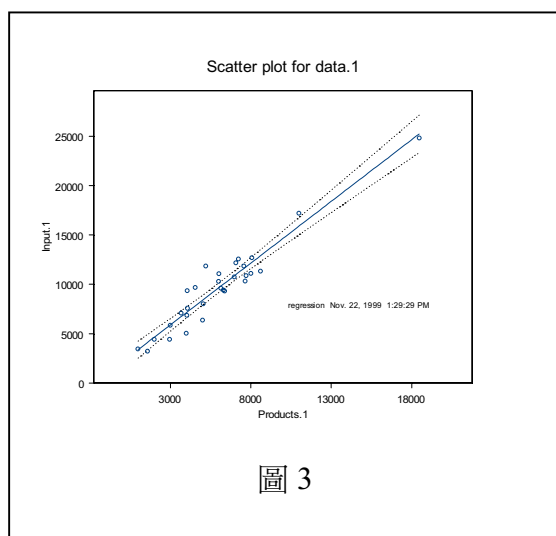


圖 3

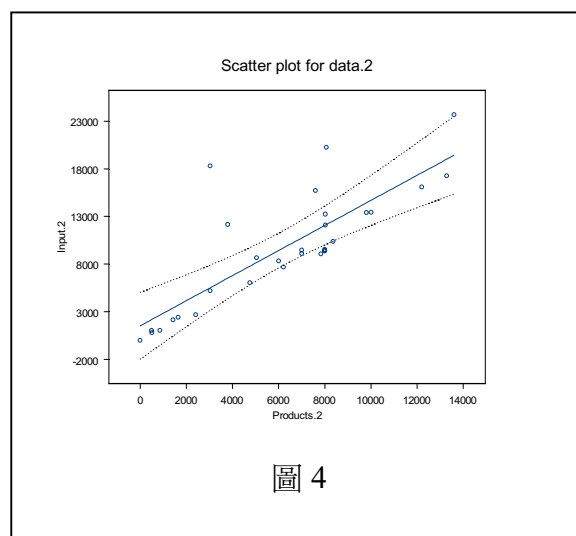


圖 4

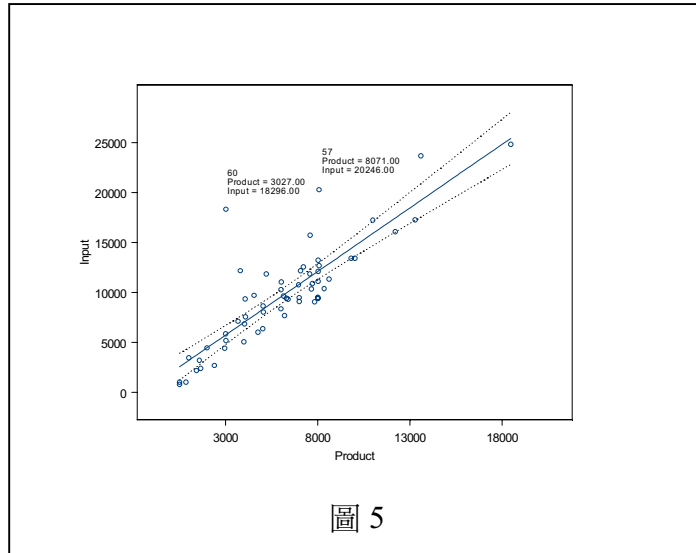


圖 5

表 1

```

*** Linear Model *** ALL
Call: lm(formula = Input ~ Product, data = all,
na.action = na.exclude)
Residuals:
  Min     1Q   Median     3Q    Max
-2835 -1566  -684.1   496.3 12596

Coefficients:
              Value Std. Error  t value Pr(>|t|)
(Intercept) 1799.5523  691.1942   2.6035  0.0117
  Product    1.2886    0.0989  13.0300  0.0000

Residual standard error: 2672 on 58 degrees of freedom
Multiple R-Squared:  0.7454
F-statistic: 169.8 on 1 and 58 degrees of freedom, the p-value is 0

Analysis of Variance Table

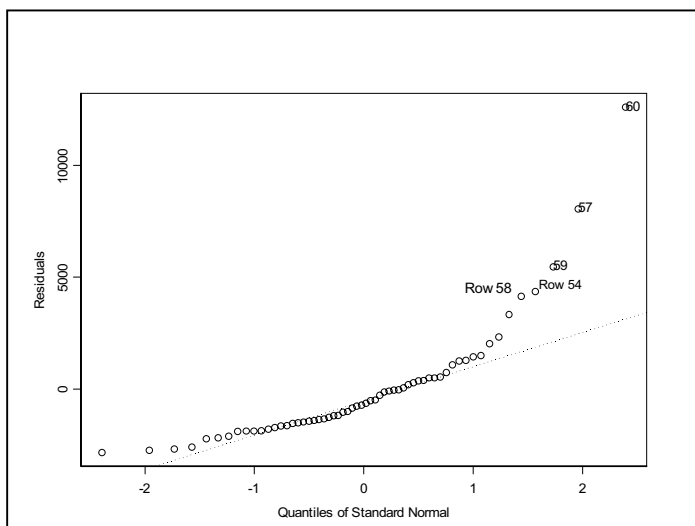
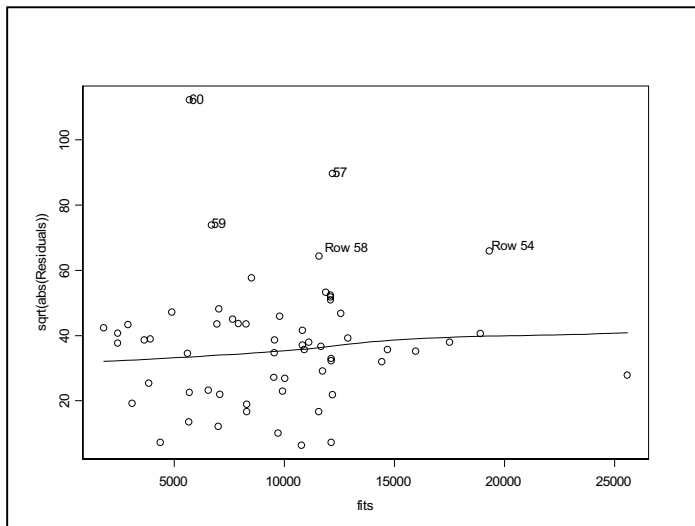
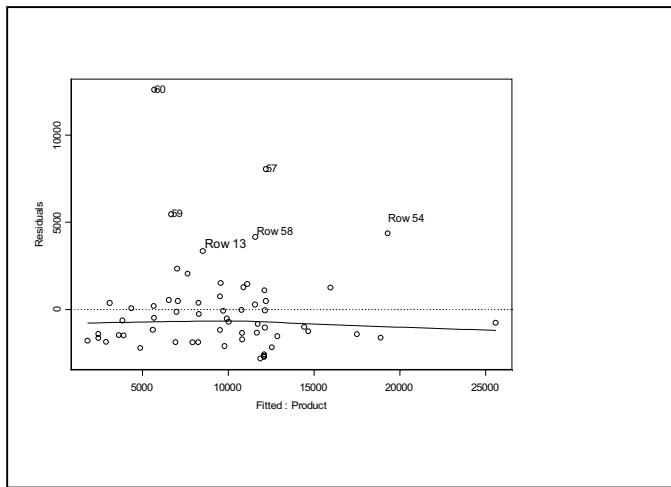
Response: Input

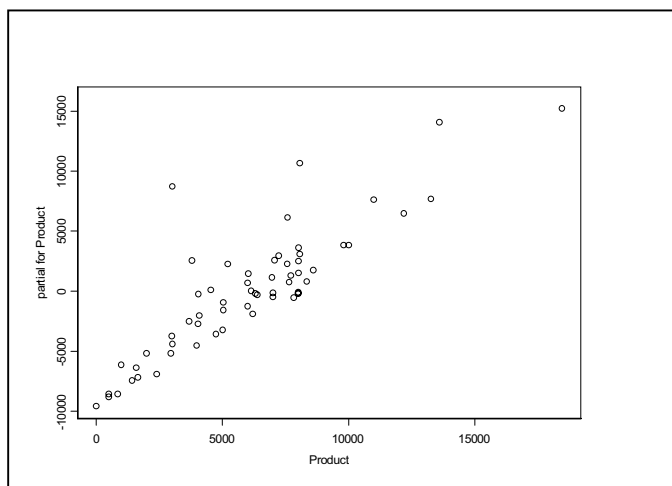
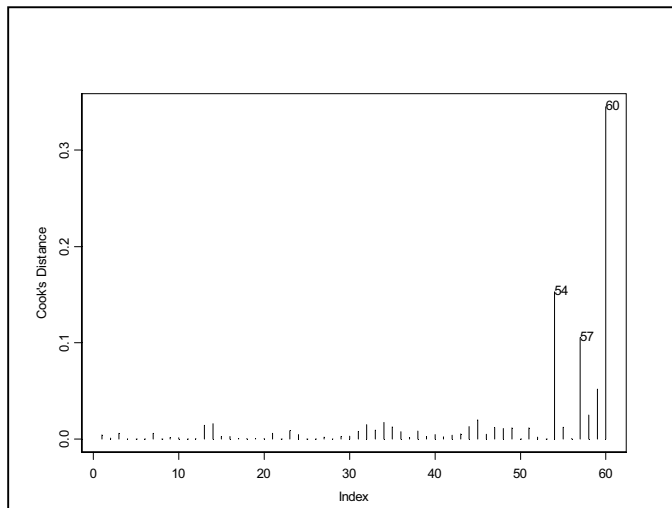
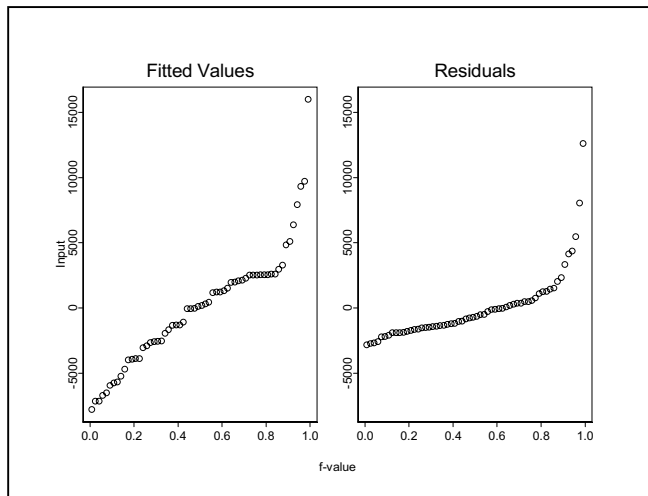
Terms added sequentially (first to last)
      Df Sum of Sq  Mean Sq  F Value Pr(F)
Product 1 1211810394 1211810394  169.781    0
Residuals 58 413974579  7137493

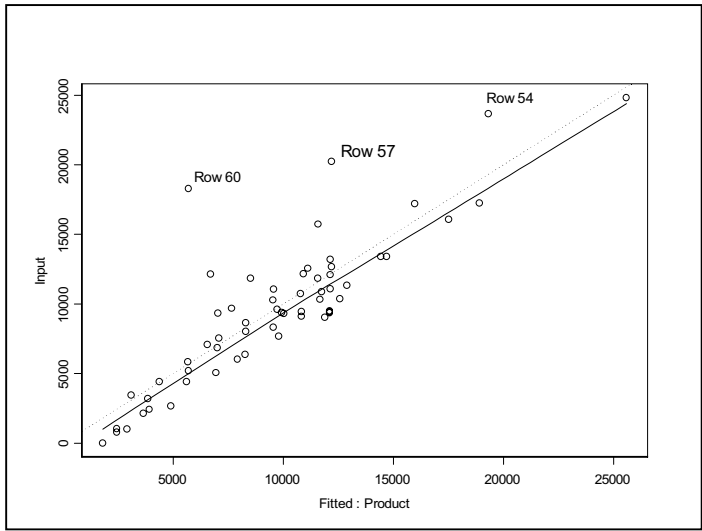
```

由表 1，得知迴歸模型的 R-square 值為 0.7454，且對此 Data 的配適性也是顯著的。但由圖 5 可看出某些觀察值可能是 Outliers，因此茲對此迴歸模型做一些圖形上的迴歸假設檢測，並找出可能的 Outliers。

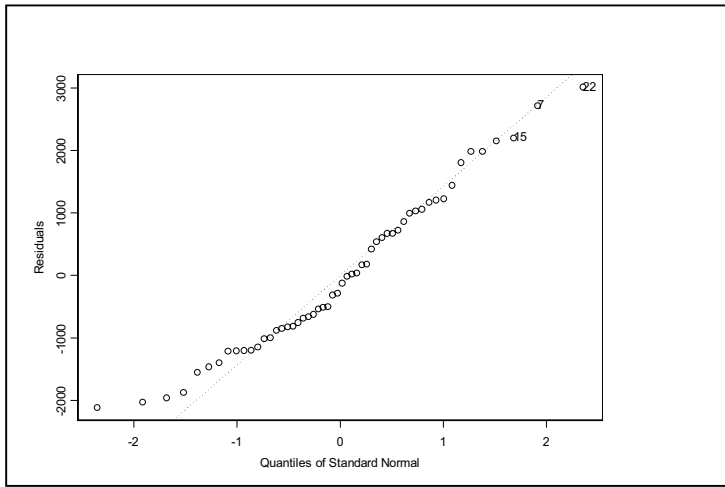
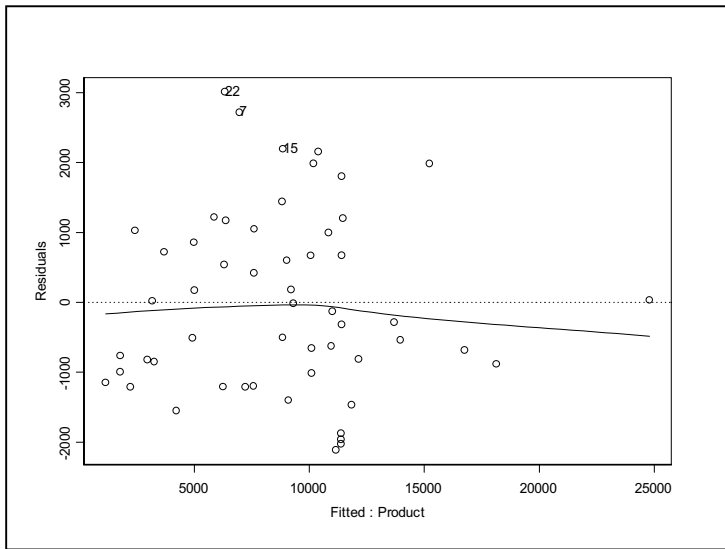
# Diagnostics







After outlier deleted

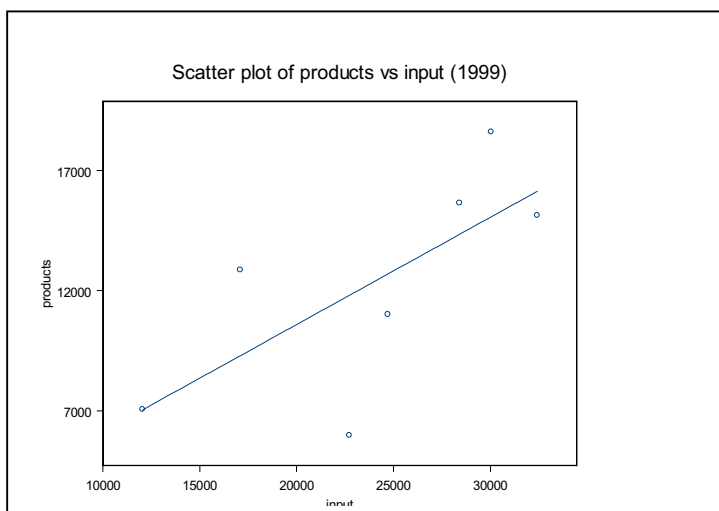
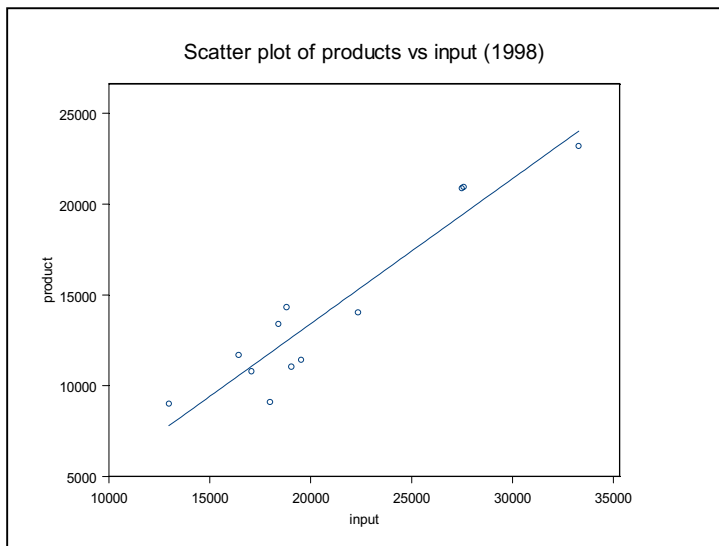
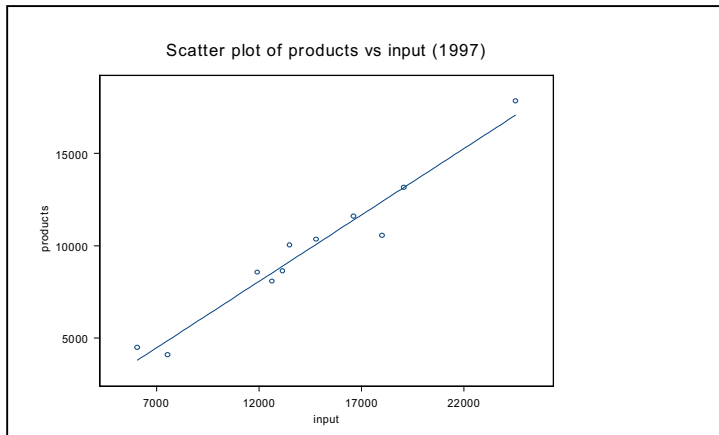


## After outliers deleted

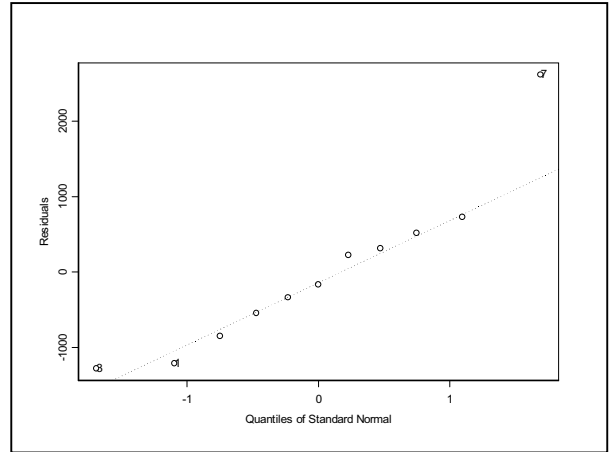
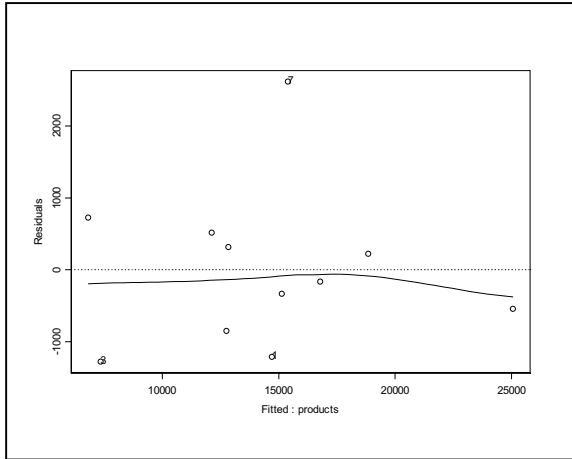
```
*** Linear Model ***  
Call: lm(formula = Input ~ Product, data =  
  nonoutliers, na.action =  
  na.exclude)  
Residuals:  
  Min      1Q  Median      3Q     Max  
-2117 -968.8 -209.3  960.6 3010  
Coefficients:  
                Value Std. Error  
(Intercept) 1147.9245  349.8446  
  Product      1.2801   0.0507  
                t value Pr(>|t|)  
(Intercept)   3.2812   0.0018  
  Product     25.2558   0.0000  
Residual standard error: 1293 on 52 degrees of freedom  
Multiple R-Squared: 0.9246  
F-statistic: 637.9 on 1 and 52 degrees of freedom, the p-value is 0  
Analysis of Variance Table  
Response: Input  
Terms added sequentially (first to last)  
      Df Sum of Sq  Mean Sq  F Value Pr(F)  
Product 1 1066937548 1066937548 637.854  0  
Residuals 52  86980332  1672699
```



分時間 1997,1998,1999 算，兩家公司合起來算



1997



```
*** Linear Model ***
Call: lm(formula = input ~ products, data = year97,
  na.action = na.exclude)
Residuals:
  Min     1Q  Median     3Q    Max
-1280 -699.1 -165.8  413.8  2616

Coefficients:
              Value Std. Error  t value
(Intercept) 1397.7714  996.0713   1.4033
  products    1.3271    0.0956  13.8807
              Pr(>|t|)
(Intercept)  0.1941
  products    0.0000

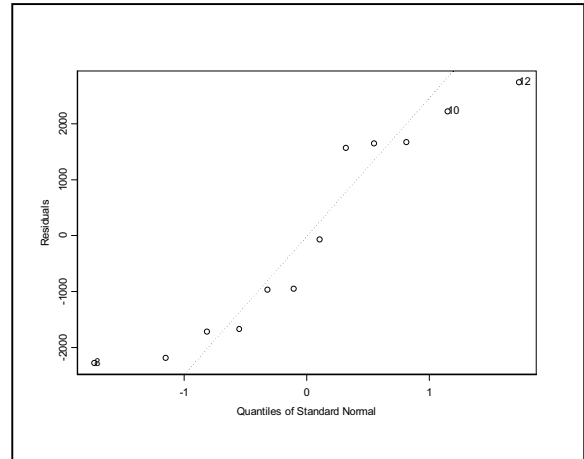
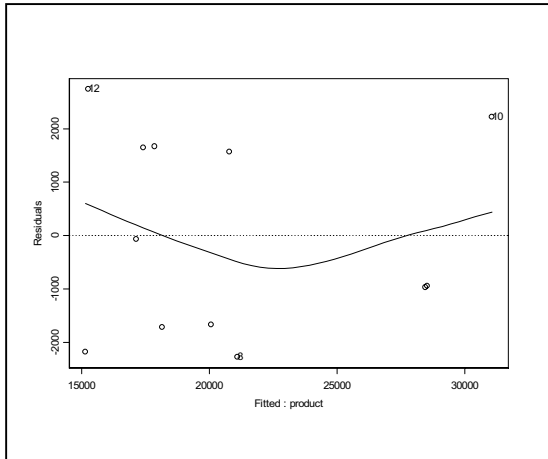
Residual standard error: 1158 on 9 degrees of freedom
Multiple R-Squared:  0.9554
F-statistic: 192.7 on 1 and 9 degrees of freedom, the p-value is 2.208e-007

Analysis of Variance Table

Response: input

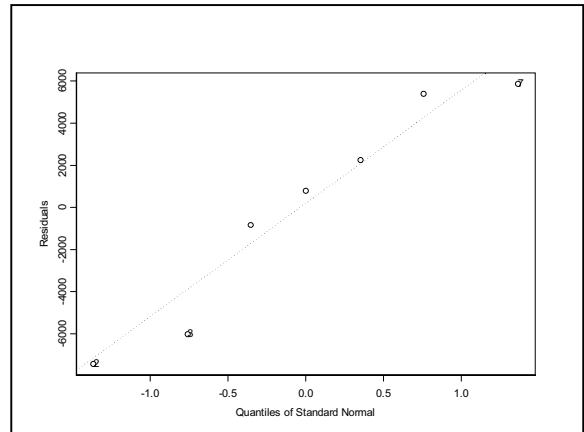
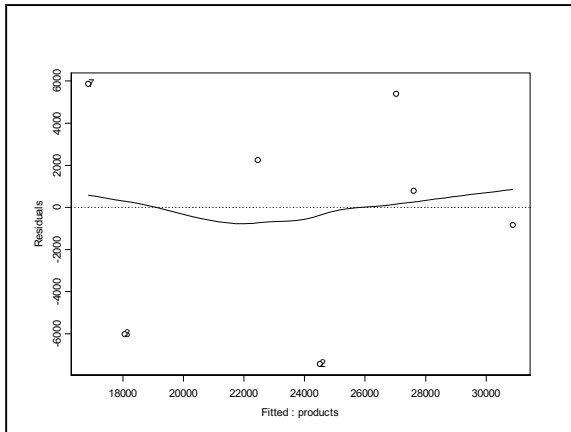
Terms added sequentially (first to last)
      Df Sum of Sq  Mean Sq F Value
products 1 258149421 258149421 192.6734
Residuals 9 12058464 1339829
              Pr(F)
products 2.208195e-007
Residuals
```

1998



```
*** Linear Model ***  
Call: lm(formula = input ~ product, data = year98,  
          na.action = na.exclude)  
Residuals:  
   Min     1Q   Median     3Q    Max  
-2279 -1684 -509.5  1651  2740  
  
Coefficients:  
                Value Std. Error  t value  
(Intercept) 5063.0934 1797.4808    2.8168  
   product      1.1212    0.1207    9.2879  
             Pr(>|t|)  
(Intercept)    0.0183  
   product      0.0000  
  
Residual standard error: 1948 on 10 degrees of freedom  
Multiple R-Squared:  0.8961  
F-statistic: 86.26 on 1 and 10 degrees of freedom, the p-value is 3.115e-006  
  
Analysis of Variance Table  
  
Response: input  
  
Terms added sequentially (first to last)  
   Df Sum of Sq  Mean Sq F Value  
product 1 327195621 327195621 86.26466  
Residuals 10 37929277 3792928  
             Pr(F)  
   product 3.115078e-006  
Residuals
```

1999



```
*** Linear Model ***  
Call: lm(formula = input ~ products, data = year99,  
          na.action = na.exclude)  
Residuals:  
    1     2     3     4     5     6     7  
-845.7 -7435 -6012  5397  790.3  2247  5858  
Coefficients:  
                Value Std. Error  t value  
(Intercept) 10212.1272  6541.9318    1.5610  
  products      1.1108    0.5010    2.2173  
                Pr(>|t|)  
(Intercept)    0.1793  
  products      0.0774  
Residual standard error: 5679 on 5 degrees of freedom  
Multiple R-Squared:  0.4958  
F-statistic: 4.916 on 1 and 5 degrees of freedom, the p-value is 0.07739  
Analysis of Variance Table  
Response: input  
Terms added sequentially (first to last)  
      Df Sum of Sq  Mean Sq F Value  
products 1 158559987 158559987 4.916261  
Residuals 5 161260761 32252152  
      Pr(F)  
products 0.07739403  
Residuals
```

