

# 目錄

圖目錄.....	1
表目錄.....	1
第一章 分析資料說明.....	2
第一節 研究背景與動機.....	2
第二節 研究目的.....	2
第三節 變數介紹.....	3
第二章 資料初步分析.....	6
第一節 變數觀察(敘述性統計量).....	6
第二節 簡單線性迴歸.....	9
一、參數估計值.....	10
二、變異數分析.....	11
三、殘差分析.....	12
四、Lack of Fit 分析.....	16
五、Box-Cox 變數轉換分析.....	16
六、Loess 無母數迴歸 .....	17
七、小節.....	19

## 圖目錄

圖 1-1 POKEMON—PIKACHU(皮卡丘).....	3
圖 2-1-2 TOTAL(神奇寶貝的能力得分)直方圖與 BOX PLOT.....	6
圖 2-1-3 HP(神奇寶貝生命值)直方圖與 BOX PLOT.....	7
圖 2-1-4 ATTACK(神奇寶貝攻擊力)直方圖與 BOX PLOT.....	7
圖 2-1-5 DEFENSE(神奇寶貝防禦力)直方圖與 BOX PLOT .....	7
圖 2-1-6 SP ATK(神奇寶貝特殊攻擊力)直方圖與 BOX PLOT .....	8
圖 2-1-7 SP DEF(神奇寶貝特殊防禦力)直方圖與 BOX PLOT .....	8
圖 2-1-8 SPEED(神奇寶貝攻擊速度)直方圖與 BOX PLOT.....	8
圖 2-2-1 應變數 TOTAL 之 QQ PLOT .....	9
圖 2-2-2 解釋變數 ATTACK 之 QQ PLOT.....	9
圖 2-2-3 SAS 線性迴歸結果.....	10
圖 2-2-4 個別預測值之信賴區間.....	11
圖 2-2-5 SAS 線性迴歸配適圖.....	12
圖 2-2-6 SAS 線性迴歸配適診斷.....	13
圖 2-2-7 殘差對 ATTACK 作散佈圖 .....	14
圖 2-2-8 誤差項之常態性相關檢定 .....	14
圖 2-2-9 SAS : TOTAL 之殘差 SHAPIRO-WILK 常態檢定 .....	15
圖 2-2-11 BROWN-FORSYTHE 檢定結果.....	16
圖 2-2-12 LACK OF FIT 檢定結果.....	16
圖 2-2-13 BOX-COX 轉換結果 .....	17
圖 2-2-14 LOWESS 配適(依平滑係數不同).....	17
圖 2-2-15 LOWESS 曲線及 SCATTER PLOT(依平滑係數不同) .....	18
圖 2-2-16 LOWESS 平滑曲線及信賴帶(平滑係數=0.85).....	18

## 表目錄

表 1-1 資料集變數說明 .....	3
表 2-1 連續型變數敘述性統計量 .....	6

## 第一章 分析資料說明

### 第一節 研究背景與動機

Pokemon GO！(寶可夢，或譯為神奇寶貝、口袋怪獸)是最近最流行的手機 APP 遊戲，由於 Pokemon(寶可夢、神奇寶貝)的卡通與相關遊戲已經發展一、二十年(官方網站：<http://www.pokemon.com/>)，故目前具有龐大的各式各樣神奇寶貝(第一世代至第六世代共 721 種，第七世代持續增加中)。

Pokemon(寶可夢、神奇寶貝)有豐富的屬性值，常常令人眼花撩亂，難以記憶為數眾多的 Pokemon(寶可夢、神奇寶貝)，故在 Pokemon(寶可夢、神奇寶貝)的卡通或遊戲中甚至有神奇寶貝圖鑑/百科(類似生物物種圖鑑)可供玩家參考(可參考中文版：<https://wiki.52poke.com/wiki/>)。

從神奇寶貝圖鑑可了解每種 Pokemon(神奇寶貝)的各種屬性值(類似生物物種的特徵值)以及其分類。如果可透過其重要屬性值即可判斷它會是哪一類型的神奇寶貝，找出它的規則；或是是否能透過重要屬性值，判斷其神奇寶貝實力值，會是十分有趣的事情。

### 第二節 研究目的

Pokemon(神奇寶貝)的類型包含：一般(Normal)、格鬥(Fighting)、飛行(Flying)、毒(Poison)、地面(Ground)、岩石(Rock)、蟲(Bug)、幽靈(Ghost)、鋼系(Steel)、火系(Fire)、水系(Water)、草系(Grass)、電氣(Electric)、超能力(Psychic)、冰系(Ice)、龍系(Dragon)、惡(Dark)、妖精(Fairy)等不同 18 種類別。

例如最有名的神奇寶貝，皮卡丘(Pikachu)即屬於電氣系(Electric)的神奇寶貝，妙蛙種子屬於草系(Grass)的神奇寶貝，傑尼龜屬於水系(Water)的神奇寶貝。



圖 1-1 Pokemon—Pikachu(皮卡丘)

本報告採用之分析資料是來自於知名資料分析競賽網站 Kaggle (<https://www.kaggle.com/>) 釋出之資料，此資料集為目前最紅的 Pokemon(神奇寶貝)種類資料之資料，共 721 種神奇寶貝 (神奇寶貝第一世代至第六世代)，800 筆資料(有幾種神奇寶貝有特別版)，它包含了各種 Pokemon(神奇寶貝)的屬性值資料。Kaggle 此次釋放的資料，主要是希望可建模找出是否可由特定數個屬性值(2~3 個)變數可預測是屬於哪一種類別 Pokemon(神奇寶貝)，找出其規則(pattern)？

然而在本次分析報告中，主要與 Kaggle 釋出此資料之分析目的較不同，本分析報告著重於建立神奇寶貝屬性值與神奇寶貝能力得分之模型，期望可用神奇寶貝屬性值即可判斷神奇寶貝能力得分。

### 第三節 變數介紹

本資料集之資料變數包含神奇寶貝名稱，屬於哪一種類型的神奇寶貝，該種神奇寶貝的實力值(Strong)，以及其六大屬性值：HP、攻擊、防禦、特攻、特防、速度，整理如下表說明之。

表 1-1 資料集變數說明

變數名稱	變數類型	變數說明
ID	連續	神奇寶貝編號(觀察值編號，流水號型式)

Name	類別變數	神奇寶貝名稱(Name of each pokemon)
Type 1	類別變數 (名目)	<p>神奇寶貝第一種類別，每個神奇寶貝都會有一個基本類別(Each pokemon has a type, this determines weakness/resistance to attacks)</p> <p>此變數值共有 18 種：Bug、Dark、Dragon、Electric、Fairy、Fighting、Fire、Flying、Ghost、Grass、Ground、Ice、Normal、Poison、Psychic、Rock、Steel、Water</p>
Type 2	類別變數 (名目)	<p>神奇寶貝第二種類別，有些神奇寶貝會有第二類別的歸屬，其能力會同時混合兩種類別的性質(Some pokemon are dual type and have 2)</p> <p>此變數值共有 18 種：Bug、Dark、Dragon、Electric、Fairy、Fighting、Fire、Flying、Ghost、Grass、Ground、Ice、Normal、Poison、Psychic、Rock、Steel、Water</p>
Total	連續變數	神奇寶貝的能力得分，可判斷神奇寶貝實力，有多強壯(sum of all stats that come after this, a general guide to how strong a pokemon is)
HP	連續變數	神奇寶貝生命值(或俗稱的血量)，可判斷神奇寶貝還能戰鬥多久( hit points, or health, defines how much damage a pokemon can withstand before fainting)
Attack	連續變數	神奇寶貝攻擊力(the base modifier for normal attacks (eg. Scratch, Punch))
Defense	連續變數	神奇寶貝防禦力(the base damage resistance against normal attacks)
SP Atk	連續變數	神奇寶貝特殊攻擊力(special attack, the base modifier for special attacks (e.g. fire blast, bubble beam))
SP Def	連續變數	神奇寶貝特殊防禦力(the base damage resistance against special attacks)

Speed	連續變數	神奇寶貝攻擊速度，每次神奇寶貝進攻時的速度(determines which pokemon attacks first each round)
Generation	類別變數	神奇寶貝世代目前共有七個世代，此資料集僅第一至第六世代(第七世代為最新世代，正在推出增加中)，第一世代至第六世代共 721 種。 此變數值共有六種，表示第一世代至第六世代：1、2、3、4、5、6
Legendary	類別變數	神奇寶貝是否為傳說中神奇寶貝(卡通及遊戲中的特殊隱藏神奇寶貝，不一定會出現)。 此變數值為 True/False 之二元變數值。

## 第二章 資料初步分析

### 第一節 變數觀察(敘述性統計量)

首先，我們選用了本資料集的連續型變數，包含：Total(神奇寶貝的能力得分)、HP(神奇寶貝生命值)、Attack(神奇寶貝攻擊力)、Defense(神奇寶貝防禦力)、SP Atk(神奇寶貝特殊攻擊力)、SP Def(神奇寶貝特殊防禦力)、Speed(神奇寶貝攻擊速度)等，共 7 種變數，其敘述性統計量如下表：

表 2-1-1 連續型變數敘述性統計量

變數	平均值	標準差	標準誤差	變異數	最小值	最大值	中位數	N
Total	434.81875	120.61364	4.2643362	14547.65	46	780	450	800
HP	69.2725	25.525089	0.9024482	651.53016	1	255	65	800
Attack	78.98	32.477347	1.1482476	1054.78	5	190	75	800
Defense	73.8425	31.183501	1.1025032	972.41071	5	230	70	800
Sp. Atk	72.82	32.722294	1.1569078	1070.75	10	194	65	800
Sp. Def	71.915	27.816811	0.9834728	773.77499	20	230	70	800
Speed	68.21625	29.150543	1.0306273	849.75418	1	180	65	800

可以看到變數中標準差最大為變數 Total，其觀察值界於 46~780 之間。變數 Sp. Atk 之平均值與中位數落差最大。每一項變數之直方圖與 Box plot 如下圖 2-1-3~圖 2-1-8。

#### ■ 變數 Total：平均值 434，中位數 450，標準差 120

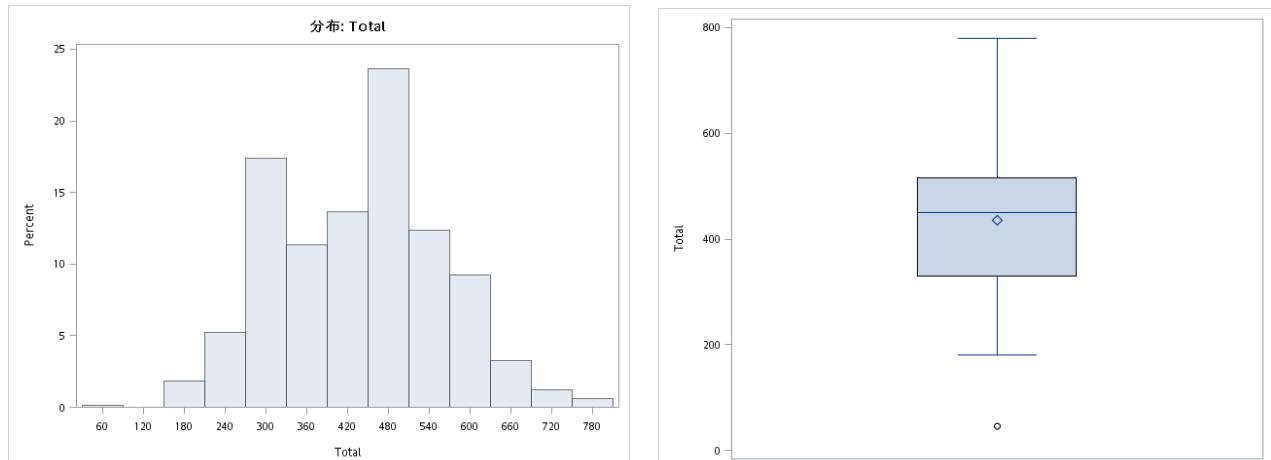


圖 2-1-2 Total(神奇寶貝的能力得分)直方圖與 Box Plot

■ 變數 HP：平均值 69，中位數 65，標準差 25

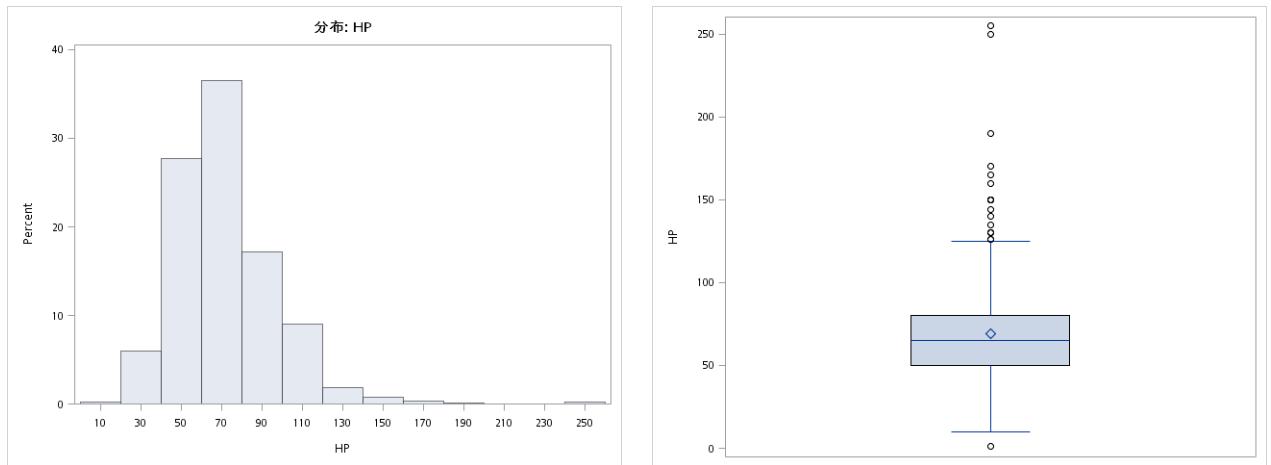


圖 2-1-3 HP(神奇寶貝生命值)直方圖與 Box Plot

■ 變數 Attack：平均值 79，中位數 75，標準差 32

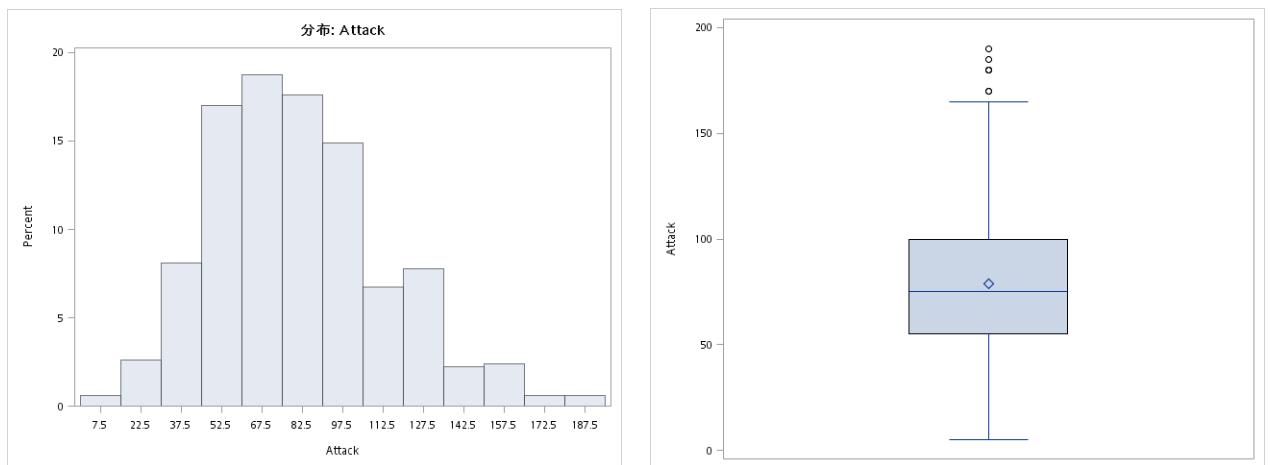


圖 2-1-4 Attack(神奇寶貝攻擊力)直方圖與 Box Plot

■ 變數 Defense：平均值 74，中位數 70，標準差 31

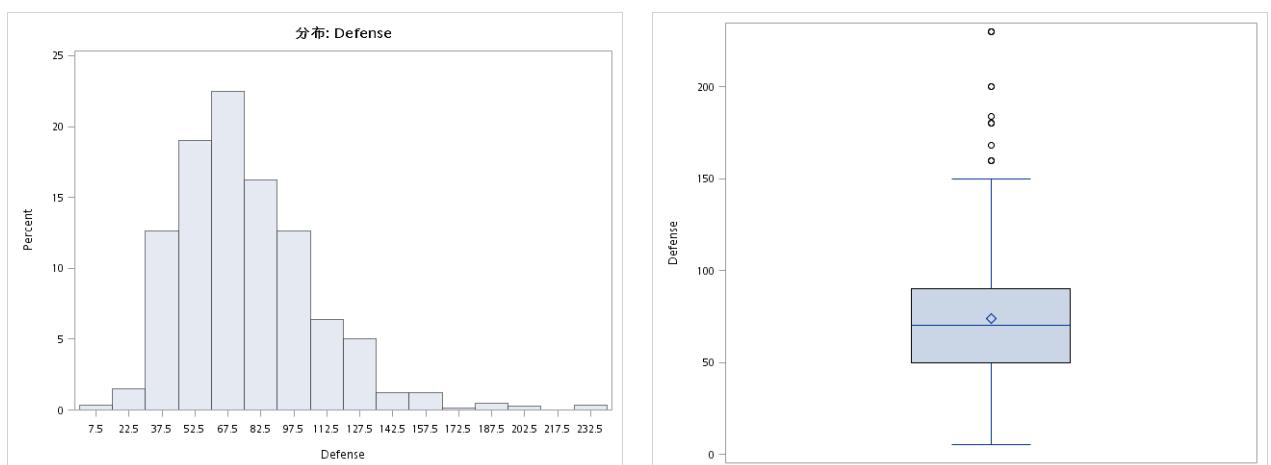


圖 2-1-5 Defense(神奇寶貝防禦力)直方圖與 Box Plot

■ 變數 SP Atk：平均值 73，中位數 65，標準差 32

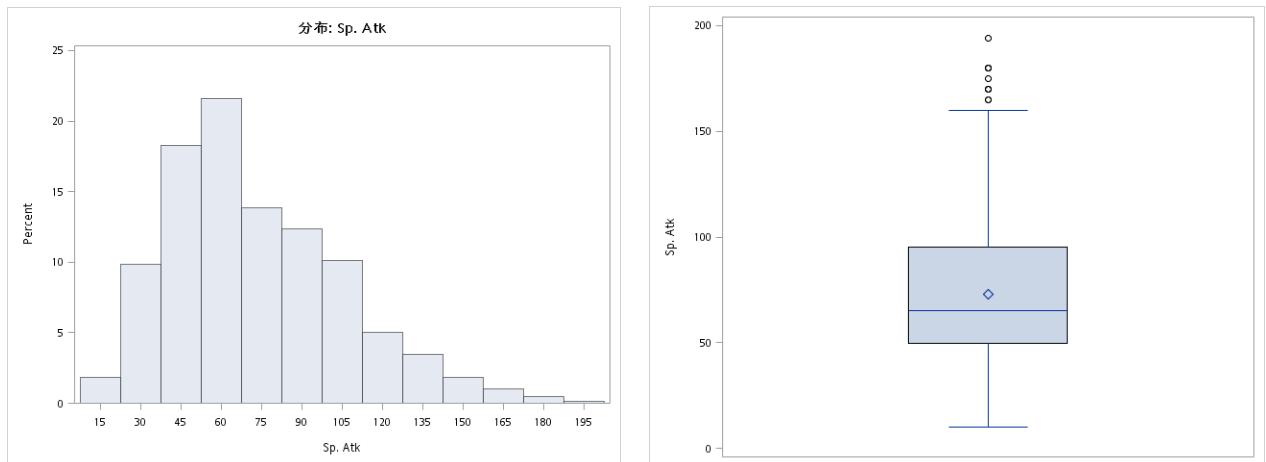


圖 2-1-6 SP Atk(神奇寶貝特殊攻擊力)直方圖與 Box Plot

■ 變數 SP Def：平均值 72，中位數 70，標準差 28

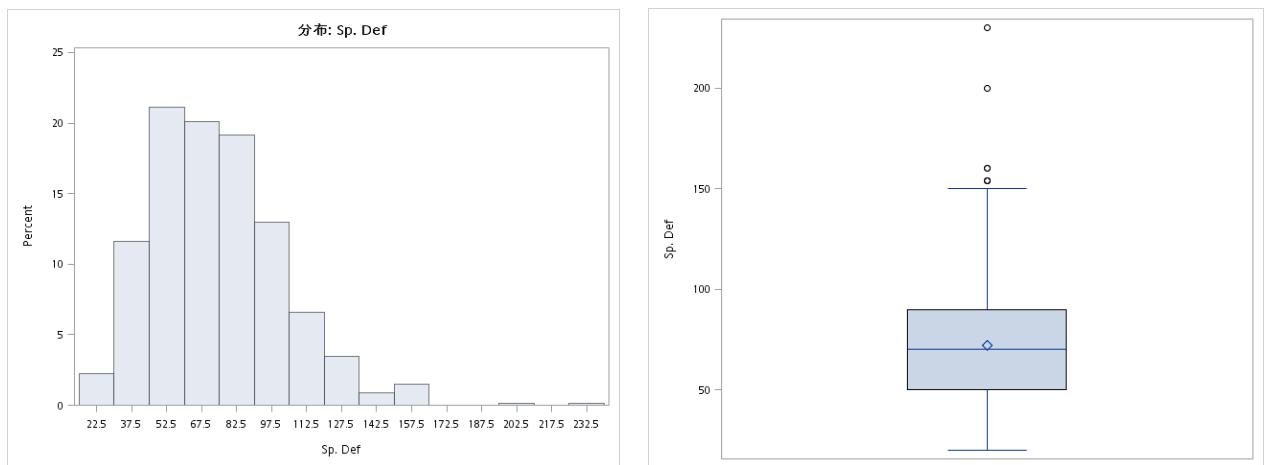


圖 2-1-7 SP Def(神奇寶貝特殊防禦力)直方圖與 Box Plot

■ 變數 Speed：平均值 68，中位數 65，標準差 29

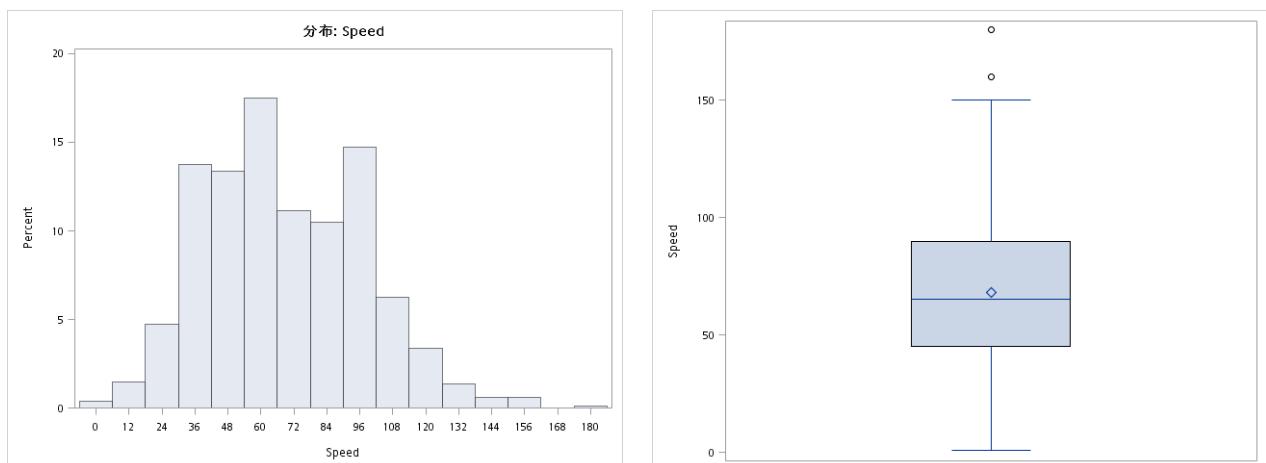


圖 2-1-8 Speed(神奇寶貝攻擊速度)直方圖與 Box Plot

## 第二節 簡單線性迴歸

在眾多屬性值中，我們最好奇 Attack(神奇寶貝攻擊力)是不是與 Total(神奇寶貝的能力得分)具線性相關性？故我們先試驗這兩個變數值，令：

應變數(Y)：Total(神奇寶貝的能力得分)

解釋變數(X)：Attack(神奇寶貝攻擊力)

先觀察這兩個變數的 QQ Plot，如圖 2-2-1、2-2-2，解釋變數 Attack 較接近常態分配，而應變數 Total 則較不似常態分配。

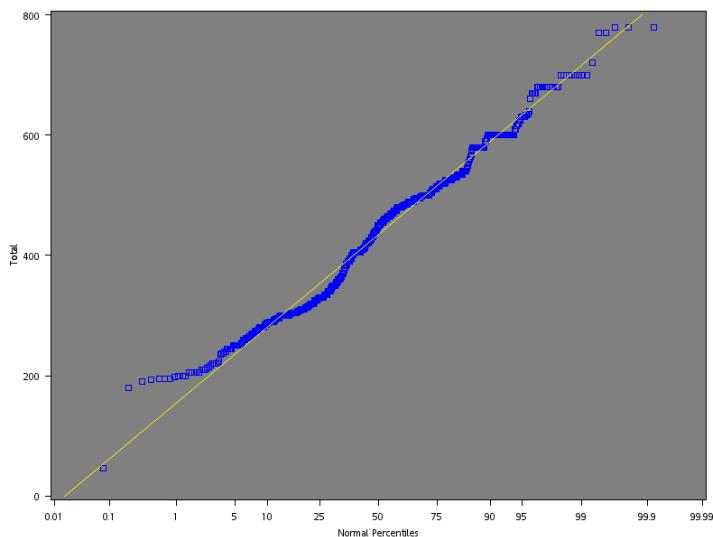


圖 2-2-1 應變數 Total 之 QQ Plot

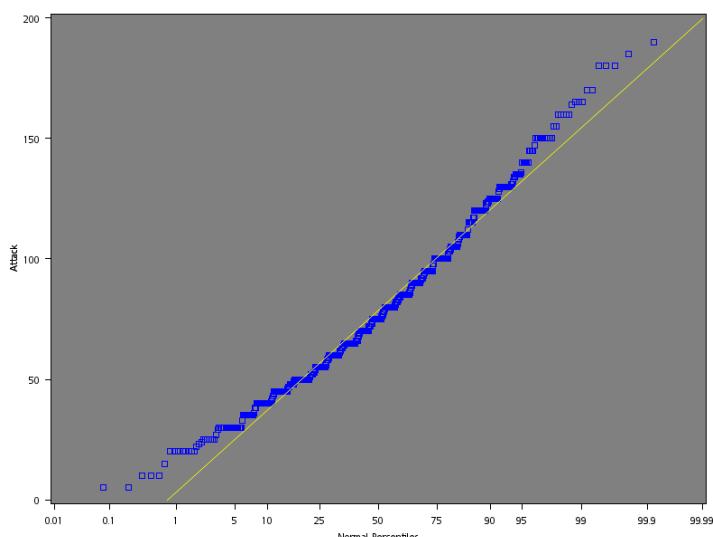


圖 2-2-2 解釋變數 Attack 之 QQ Plot

接著進行迴歸分析，由 SAS 計算後之可知下列線性迴歸結果：

線性迴歸結果					
REG 程序					
模型: Linear_Regression_Model					
應變數: Total					
讀取的觀測值數目					800
使用的觀測值數目					800
變異數分析					
來源	自由度	平方和	平均值 平方	F 值	Pr > F
模型	1	6287832	6287832	940.39	<.0001
誤差	798	5335741	6686.39210		
配適不足	109	803919	7375.40225	1.12	0.2030
純誤差	689	4531822	6577.39049		
已校正的總計	799	11623573			
根 MSE		81.77036	R 平方	0.5410	
應變平均值		434.81875	調整 R 平方	0.5404	
變異係數		18.80562			
參數估計值					
變數	自由度	參數 估計值	標準 誤差	t 值	Pr >  t
Intercept	1	219.08721	7.60579	28.81	<.0001
Attack	1	2.73147	0.08907	30.67	<.0001
				0.73550	2.55663
					2.90631
95% 信賴界限					

圖 2-2-3 SAS 線性迴歸結果

## 一、 參數估計值

$$\hat{\beta}_0 = 219.08721 ; \hat{\beta}_1 = 2.73147$$

其線性關係之迴歸函數為  $Y = 219.08721 + 2.73147X$

若我們將 Pikachu(皮卡丘)的 Attack 數值 55 代入此估計式，可得

Total = 219.08721 + 150.23085 = 369.31806，比實際 Pikachu(皮卡丘)的 Total(神奇寶貝的能力得分)數值 320 要高。

如下圖 2-2-4 為個別預測值之信賴區間，Pikachu(皮卡丘)為第 31 個觀測值，可看到  $X_h=55$  時， $E\{Y_h\}$  的 95% 信賴區間為  $362.2623 < E\{Y_h\} < 376.3738$  當  $X_h=55$  時， $\hat{Y}_h$  的 95% 預測區間為  $208.6527 < \hat{Y}_{h(\text{new})} < 529.9835$ 。

觀測	應變數	預測值	輸出統計值					
			標準誤差	95% CL 平均值	95% CL 預測值	殘差		
20	495.0000	628.8078	6.9552	615.1551	642.4605	467.7178	789.8978	-133.8078
21	251.0000	342.0034	4.1855	333.7874	350.2193	181.2828	502.7239	-91.0034
22	349.0000	382.9754	3.3490	376.4015	389.5494	222.3305	543.6204	-33.9754
23	479.0000	437.6048	2.8924	431.9271	443.2826	276.9941	598.2156	41.3952
24	579.0000	437.6048	2.8924	431.9271	443.2826	276.9941	598.2156	141.3952
25	253.0000	372.0496	3.5423	365.0963	379.0028	211.3886	532.7105	-119.0496
26	413.0000	440.3363	2.8966	434.6504	446.0222	279.7252	600.9474	-27.3363
27	262.0000	382.9754	3.3490	376.4015	389.5494	222.3305	543.6204	-120.9754
28	442.0000	464.9196	3.0531	458.9265	470.9126	304.2973	625.5418	-22.9196
29	288.0000	382.9754	3.3490	376.4015	389.5494	222.3305	543.6204	-94.9754
30	438.0000	451.2622	2.9403	445.4905	457.0339	290.6481	611.8764	-13.2622
31	320.0000	369.3181	3.5945	362.2623	376.3738	208.6527	529.9835	-49.3181
32	485.0000	464.9196	3.0531	458.9265	470.9126	304.2973	625.5418	20.0804
33	300.0000	423.9475	2.9127	418.2301	429.6649	263.3353	584.5597	-123.9475
34	450.0000	492.2343	3.4443	485.4732	498.9953	331.5815	652.8870	-42.2343
35	275.0000	347.4663	4.0586	339.4996	355.4331	186.7583	508.1743	-72.4663

圖 2-2-4 個別預測值之信賴區間

承前述參數估計值之結果，可知若  $\alpha=0.05$  時， $2.55663 < \beta_1 < 2.90631$ ；因此，在 95% 的信賴係數下，我們估計 Attack 每增加 1 單位，Total 增加量約介於 2.55663 到 2.90631 之間。

並且由於我們假設  $H_0: \beta_1 = 0$ ， $H_1: \beta_1 \neq 0$ ，故若  $\alpha=0.05$  時，觀察 p-value  $< 0.0001$ ，拒絕  $H_0$ ，代表我們有足夠的證據顯示  $\beta_1 \neq 0$ ，表示 Total(神奇寶貝的能力得分)與 Attack(神奇寶貝攻擊力)存在著線性關係。

## 二、 變異數分析

針對變異數分析結果，可觀察到  $F=940.39$ , p-value  $< 0.0001$ ，故拒絕  $H_0$ ，即 Total(神奇寶貝的能力得分)與 Attack(神奇寶貝攻擊力)具線性關係。另外，判定係數  $R^2$  為 0.5410，故若考慮 Attack(神奇寶貝攻擊力)後，Total(神奇寶貝的能力得分)的變異削減了 54%。

迴歸函數與資料點之配適圖如下圖 2-2-5，為 Total(神奇寶貝的能力得分)與 Attack(神奇寶貝攻擊力)之散佈圖，繪製估計出來的簡單線性迴歸與其 95% 期望值估計信賴區間 95% 預測值信賴區間。

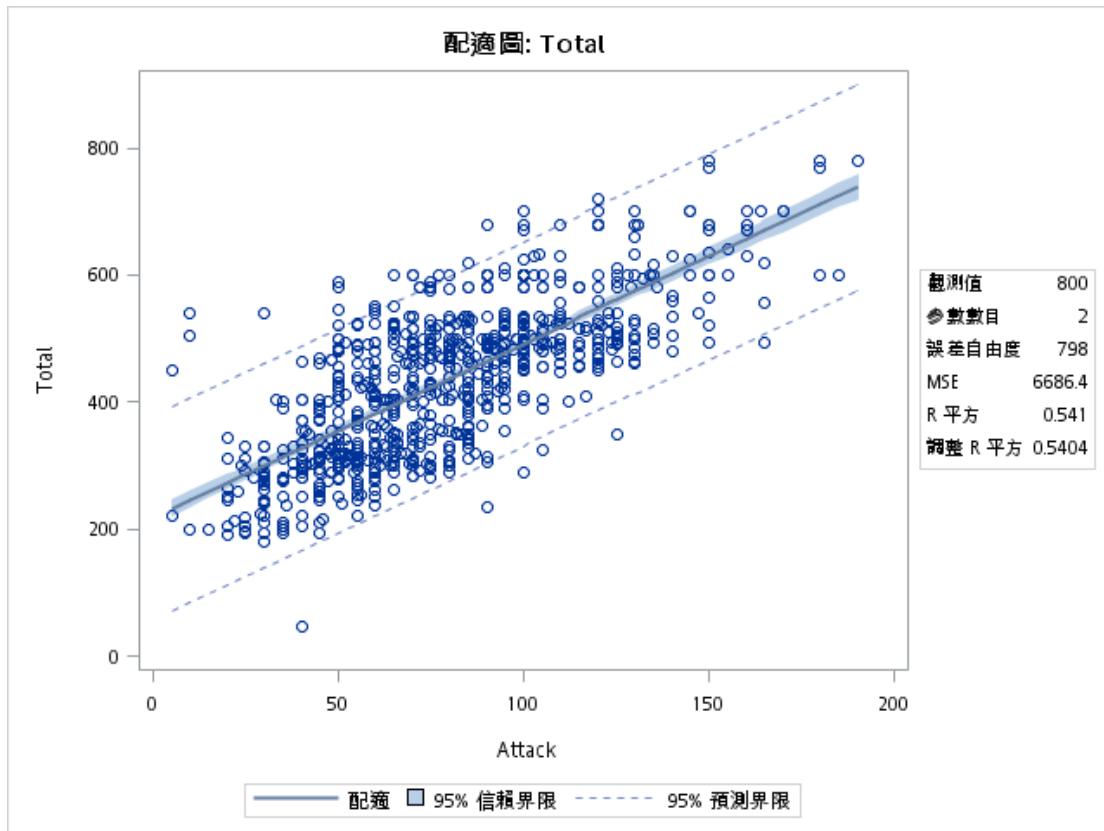


圖 2-2-5 SAS 線性迴歸配適圖

### 三、殘差分析

#### (一) 初步診斷

圖 2-2-6 為為殘差對應變數 Total 作的各種圖形分析，由最左方第一個圖，及最左方第二個圖觀察，可知二項重要事項：

- (1). 殘差分布的情形無顯著存在某趨勢性，其均勻散佈在此圖當中，故殘差與應變數 Total 應無相關，可視殘差與應變數無關。
- (2). 殘差之常態機率圖(QQ-Plot)貼近 45 度角的斜直線，各殘差值都在其附近，看起來殘差似乎應符合常態分配之假設。

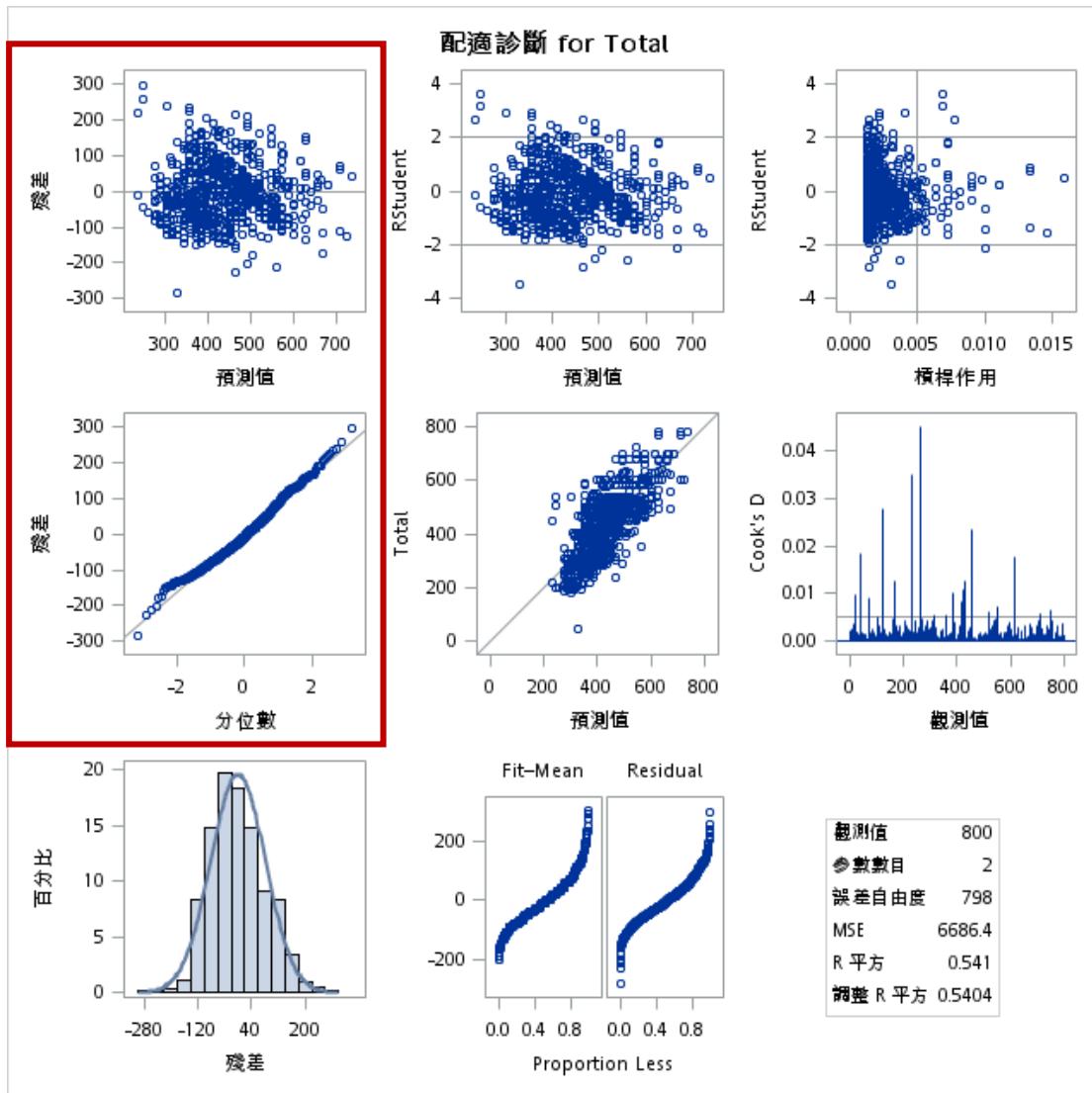


圖 2-2-6 SAS 線性迴歸配適診斷

另外，我們可再多觀察殘差對 Attack 作的散佈圖，如圖 2-12。目的是為了觀察殘差與解釋變數 Attack 之間是否存在某種相關？因解釋變數不為一隨機變數，理論上應該與殘差並無相關。

如圖 2-2-7 中，殘差分布的情形無顯著存在某趨勢性，我們可視殘差與解釋變數 Attack 無關。

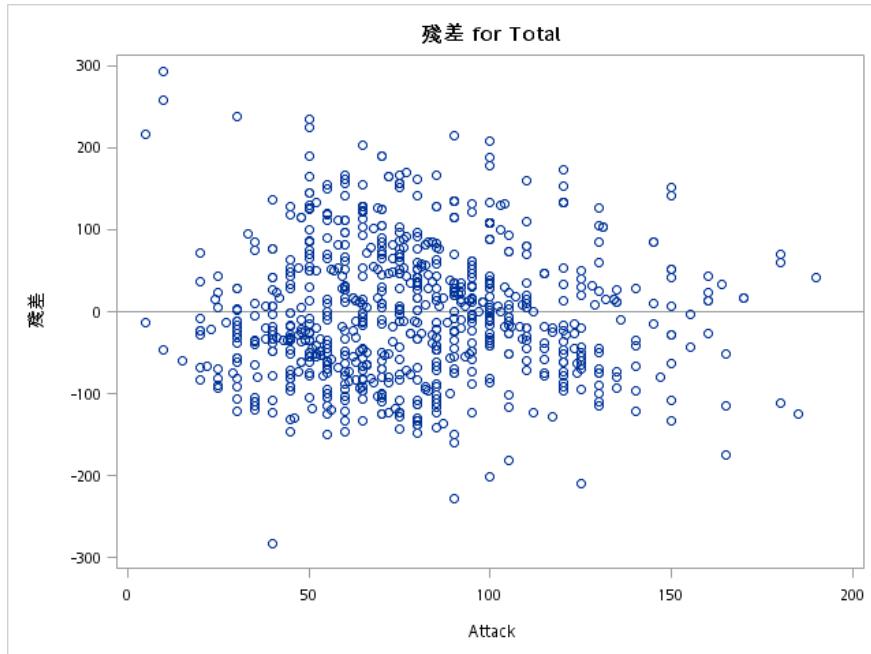


圖 2-2-7 殘差對 Attack 作散佈圖

## (二)、 殘差常態檢定

由前述應變數 Total 之殘差分析，初步診斷看起來似乎符合常態。先以 Looney 與 Gullledge 的方式進行檢定，根據常態誤差下的順序殘差與常態期望值之相關係數臨界值表， $n=100$  時(本研究樣本數為 800，故遠大於表中之  $n=100$ )，若  $\alpha=0.05$ ，則臨界值為 0.987。可以看到圖 2-2-8 中，不論是 Pearson 或 Spearman 的相關係數皆已大於 0.987，故可知誤差項並未嚴重偏離常態分配。

Pearson 相關係數, N = 800 Prob >  r  (位於 H0 底下): Rho=0		
	residual_Total	expec
residual_Total	1.00000	0.99427
殘差		<.0001
expec	0.99427	1.00000
		<.0001

Spearman 相關係數, N = 800 Prob >  r  (位於 H0 底下): Rho=0		
	residual_Total	expec
residual_Total	1.00000	1.00000
殘差		<.0001
expec	1.00000	1.00000
		<.0001

圖 2-2-8 誤差項之常態性相關檢定

但若以常態檢定實際進行檢定後，如圖 2-2-9，可發現 Shapiro-Wilk 與 Kolmogorov-Smirnov 檢定在  $\alpha=0.05$  下，拒絕  $H_0$  之假設，亦即拒絕殘差為常態分配( $H_0$ ：服從常態分配;  $H_1$ ：不服從常態分配)。但若  $\alpha=0.01$ ，則 Kolmogorov-Smirnov 檢定勉強不拒絕  $H_0$  之假設，亦即不拒絕殘差為常態分配。

Tests for Normality				
Test	Statistic	p Value		
Shapiro-Wilk	W	0.988890	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.047931	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.446210	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	2.793528	Pr > A-Sq	<0.0050

圖 2-2-9 SAS : Total 之殘差 Shapiro-Wilk 常態檢定

### (三)、 Brown-Forsythe 檢定

由於 Brown-Forsythe Test 不需考慮是否為常態分析，故先以 BF 檢定誤差項是否為常數變異數。BF 檢定之假設如下：

$H_0$ ：誤差項變異數為常數

$H_1$ ：誤差項變異數不是常數

group	N	平均值	標準差	標準誤差	最小值	最大值
1	409	69.0914	54.1691	2.6785	0	311.8
2	391	61.3153	46.6287	2.3581	0	233.0
Diff (1-2)		7.7761	50.6244	3.5806		

group	方法	平均值	95% CL 平均值	標準差	95% CL 標準差
1		69.0914	63.8260	74.3568	54.1691
2		61.3153	56.6791	65.9515	46.6287
Diff (1-2)	集區	7.7761	0.7476	14.8046	50.6244
Diff (1-2)	Satterthwaite	7.7761	0.7710	14.7812	

方法	變異數	自由度	t 值	Pr >  t
集區	均等	798	2.17	0.0302
Satterthwaite	不均等	789.43	2.18	0.0296

變異數相等性				
方法	分子自由度	分母自由度	F 值	Pr > F
Folded F	408	390	1.35	0.0029

圖 2-2-10 Brown-Forsythe 檢定結果

取解釋變數 Attack 之中位數將資料分成兩群(中位數=75)。由圖 2-2-10 檢定結果可知，變異數不均等，故其  $p\text{-value}=0.0029 < 0.05$ ，故可知拒

絕  $H_0$  之假設，亦即誤差項變異數不是常數。

#### (四)、 Breusch-Pagan 檢定

由於此資料集的樣本較大(樣本數為 800)，故接著再繼續透過檢力更強，適合用於大樣本的 Breusch-Pagan Test，檢定誤差項是否為常數變異數。BP 檢定之假設如下：

$H_0$ ：誤差項變異數為常數

$H_1$ ：誤差項變異數不是常數

Obs	ssrs	sse	nobs	tests	pv
1	375975638.89	5335740.89	800	4.22591	0.96019

圖 2-2-11 Brown-Forsythe 檢定結果

由圖 2-2-11 可知，不拒絕  $H_0$  之假設，亦即誤差項變異數為常數。

#### 四、 Lack of Fit 分析

F Test of Lack of Fit 的假設如下：

$H_0$ ：regression function is linear ( $E(Y)=\beta_0 + \beta_1 X$ )

$H_1$ ：regression function is not linear ( $E(Y)\neq\beta_0 + \beta_1 X$ )

由圖 2-2-12 可知，配適不足的 p-value = 0.203，不拒絕  $H_0$  之假設，故表示此迴歸函數為線性關係。

變異數分析					
來源	自由度	平方和	平均值 平方	F 值	Pr > F
模型	1	6287832	6287832	940.39	<.0001
誤差	798	5335741	6686.39210		
配適不足	109	803919	7375.40225	1.12	0.2030
純誤差	689	4531822	6577.39049		
已校正的總計	799	11623573			

圖 2-2-12 Lack of Fit 檢定結果

#### 五、 Box-Cox 變數轉換分析

由前述之各項殘差分析中顯示，有些檢定結果不服從常態，故再以 Box-Cox 轉換進行評估解釋變數 Attack 是否可再做轉換？由圖 2-2-13 可知， $\lambda = 1$  (no transformation)，表示不需轉換。

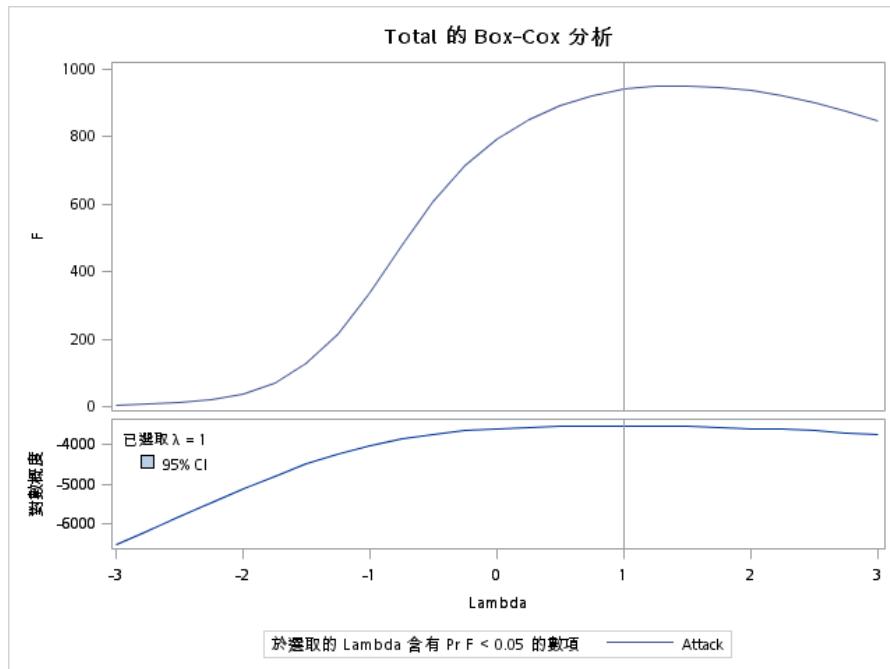


圖 2-2-13 Box-COX 轉換結果

## 六、Loess 無母數迴歸

利用 Lowess 無母數迴歸進行配適，平滑係數取 0.55, 0.65, 0.75, 0.85 等，如圖 2-2-14，發現平滑係數 0.85 之配適效果較佳。

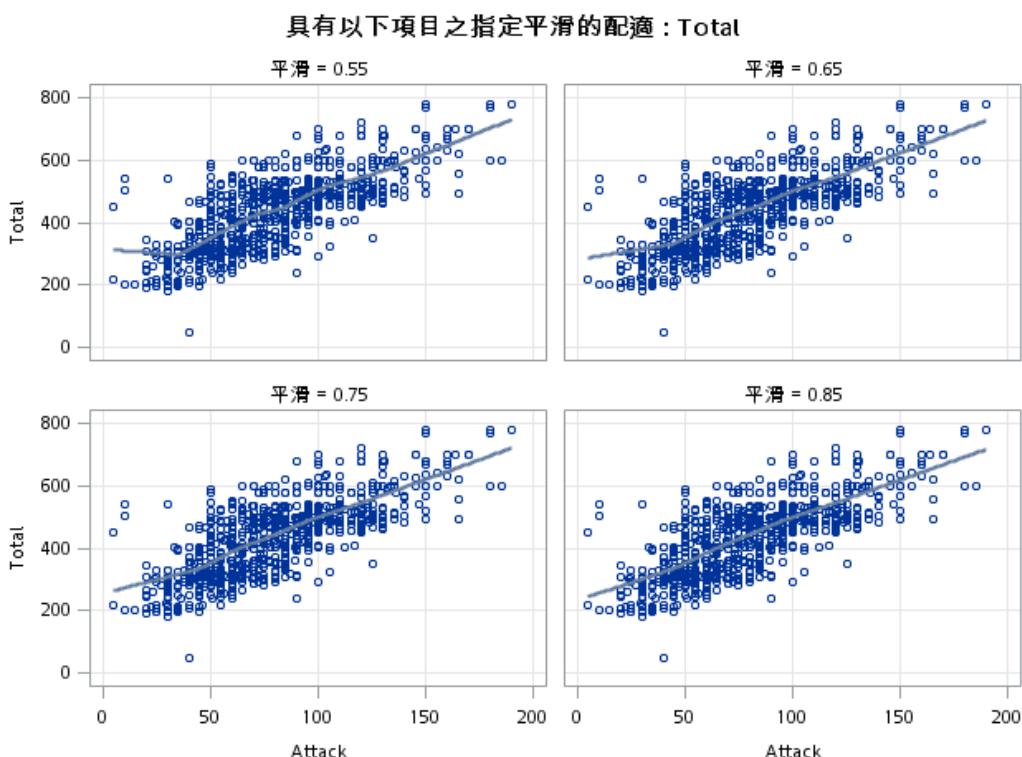


圖 2-2-14 Lowess 配適(依平滑係數不同)

再以平滑係數 0.85 進行平滑曲線描繪，如圖 2-2-15 是 Lowess 曲線及

Scatter Plot。接著再看圖 2-2-16，可以看出來配適良好，整條 Lowess 曲線，大致皆落在迴歸與 Lowess 曲線的信賴帶(Confidence Band)內，顯示應變數 Total 與解釋變數 Attack 之關係可能為具有些微彎度之曲線。

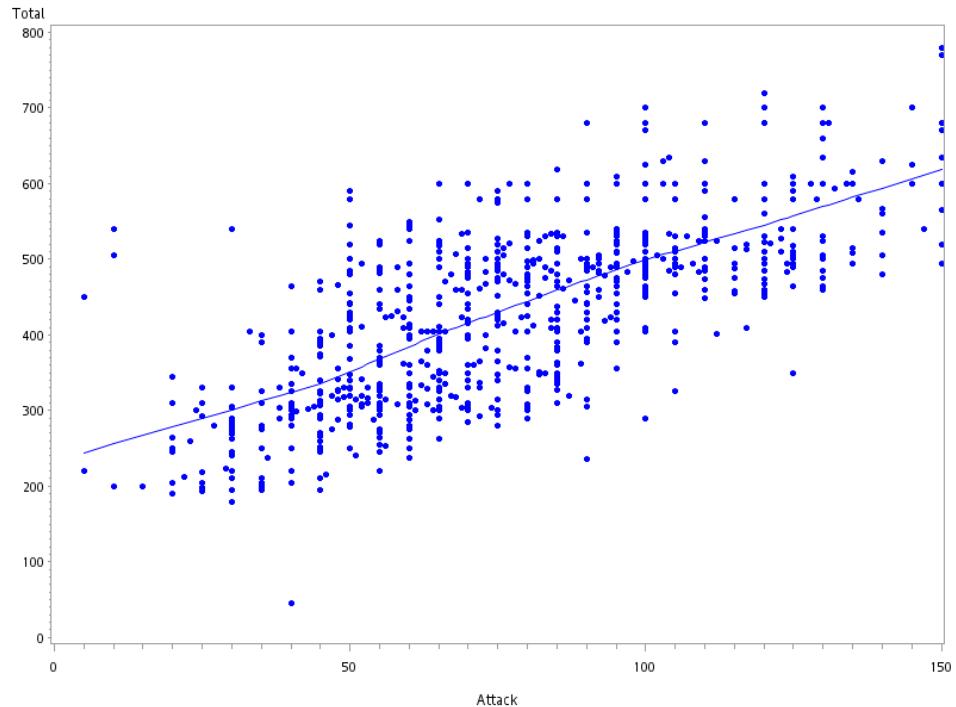


圖 2-2-15 Lowess 曲線及 Scatter Plot(依平滑係數不同)

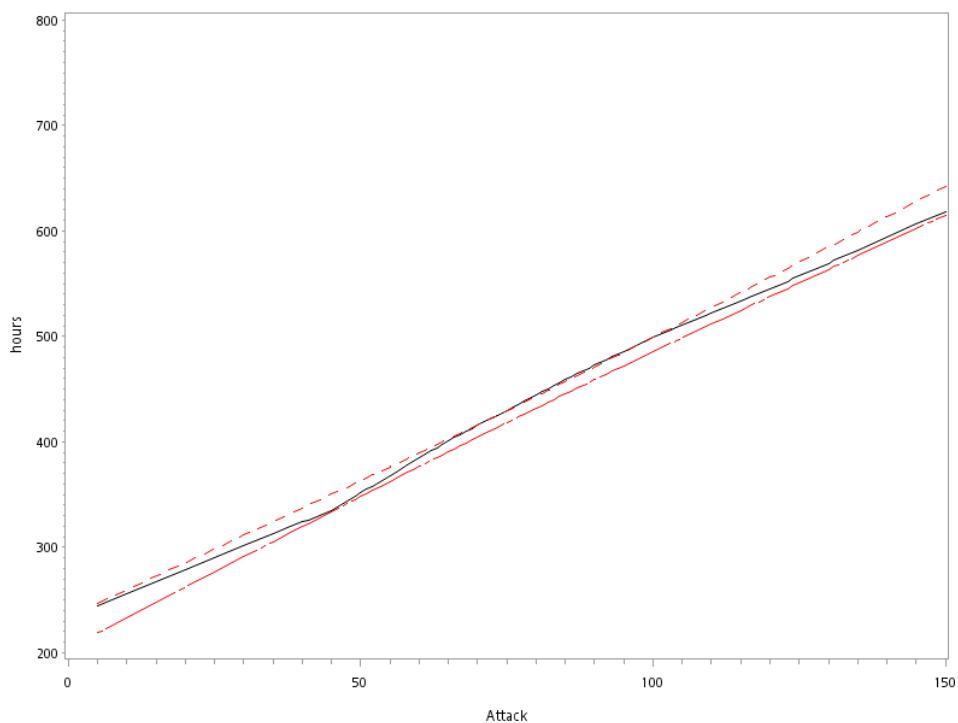


圖 2-2-16 Lowess 平滑曲線及信賴帶(平滑係數=0.85)

## 七、 小節

綜合前述分析，進行應變數 Total 與解釋變數 Attack 之迴歸分析，結果摘要如下：

線性關係之迴歸函數為  $Y = 219.08721 + 2.73147X$ ，判定係數  $R^2$  為 0.5410。

- ✓ 變異數分析： $F=940.39$ ,  $p\text{-value} < 0.0001$ ，故拒絕  $H_0$ ，即 Total(神奇寶貝的能力得分)與 Attack(神奇寶貝攻擊力)具線性關係
- ✓ 殘差常態檢定：殘差之常態機率圖(QQ-Plot)貼近 45 度角的斜直線，各殘差值都在其附近，看起來殘差似乎應符合常態分配之假設。以常態性相關檢定進行檢定時，亦發現誤差項並未嚴重偏離常態分配。故殘差大致符合常態分配。
- ✓ Brown-Forsythe Test：拒絕  $H_0$  之假設，即誤差項變異數不是常數
- ✓ Breusch-Pagan Test：不拒絕  $H_0$  之假設，即誤差項變異數為常數
- ✓ Lack of Fit 分析：配適不足的  $p\text{-value} = 0.203$ ，不拒絕  $H_0$  之假設，故表示此迴歸函數為線性關係。
- ✓ Box-Cox 轉換：評估解釋變數 Attack 是否可再做轉換，分析結果為  $\lambda = 1$  (no transformation)，表示不需轉換。
- ✓ Lowess 無母數迴歸分析：進行配適後可看出來配適良好，整條 Lowess 曲線，大致皆落在迴歸與 Lowess 曲線的信賴帶(Confidence Band)內，顯示應變數 Total 與解釋變數 Attack 之關係可能為具有些微彎度之曲線。

由上述結果知，此迴歸模型之殘差大致為常態分布，整體迴歸函數亦大致為線性關係(變異性分析與 Lack of Fit 分析)。雖然在判別係數上不高(僅 0.54)，且若嚴格的常態性檢定判斷殘差，亦僅為勉強接受，而 BF 檢定與 BP 檢定亦未無全通過；但經過 Box-Cox 轉換分析，亦無適合之變數轉換方式。故可知整體說來，此迴歸函數配適尚可，但非最佳。