

亞洲部分國家地區來台人次分析

林語萱	戴聆	陳惟中	程思嘉
資訊管理學系二年級	資訊管理學系二年級	資訊管理學系二年級	資訊管理學系二年級
國立台灣大學	國立台灣大學	國立台灣大學	國立台灣大學
B06705026@ntu.edu.tw	B06705008@ntu.edu.tw	B06705048@ntu.edu.tw	B06705038@ntu.edu.tw
朱家儀	應卓廷	王佩琳 [†]	
資訊管理學系二年級	資訊管理學系二年級	資訊管理學系二年級	
國立台灣大學	國立台灣大學	國立台灣大學	
B06705018@ntu.edu.tw	B04705030@ntu.edu.tw	B06705048@ntu.edu.tw	

一、分析主題

對近年來亞洲部分國家地區來台人次做時間序列分析，預測未來一年(民國 108 年 5 月到 109 年 4 月)每月來台人次並分析季節性指數。

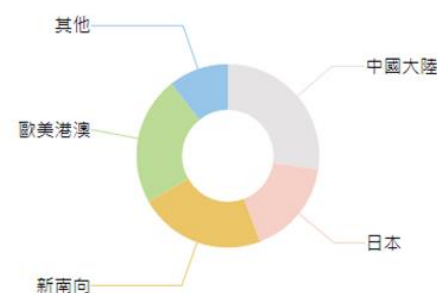
二、分析動機

在過去的十幾年中全球拚觀光的風潮下，我國政府陸續地推出「觀光客倍增計畫」、「開放陸客來台」、「亞洲旅遊重要目的地」等政策，希望能有效增加他國來台人次。根據交通部觀光局統計指出，2008 年到 2017 十年間，來台外籍旅客從每年不到 400 萬人次，增加到每年 1 千多萬人次，增幅高達 1.8 倍。

根據交通部觀光局網站統計資料顯示，108 年 1 月到 4 月來台人次以中國大陸最多，次之為日本，兩者就佔了約 44%，接著是新南向國家，包括馬來西亞、新加坡等國，以及歐美港澳和其他國家。

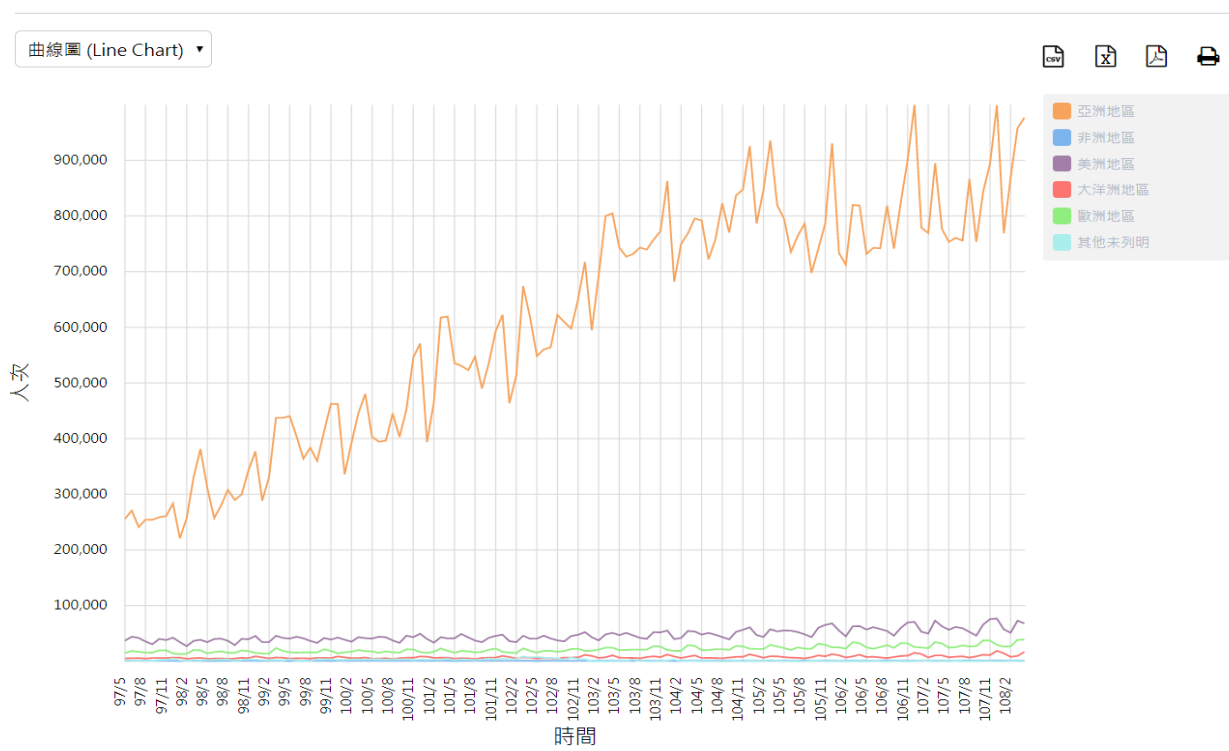
而我們實際到該網站上查詢從 97 年 5 月到 108 年 4 月的歷史來台資料，發現歷年來亞洲地區的來

108年1-4月來臺遊客

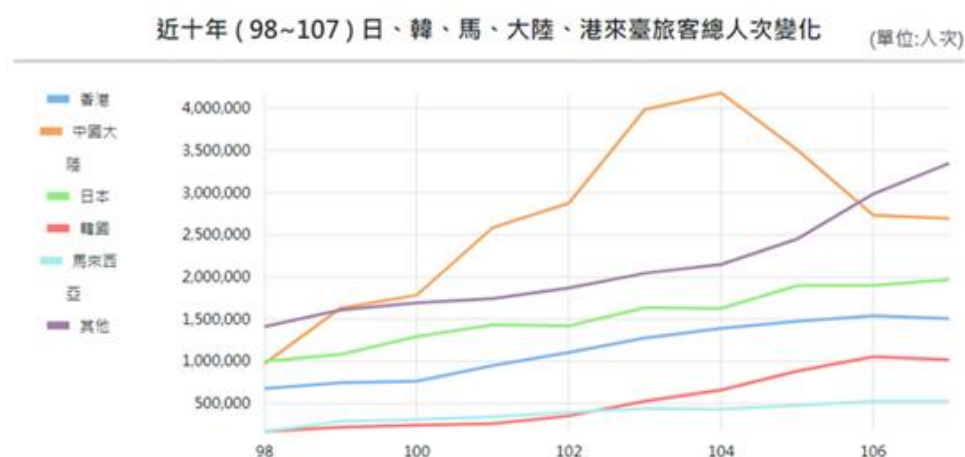


中國大陸	1,099,922	27%
日本	673,280	17%
新南向	895,862	22%
歐美港澳	910,886	23%
其他	424,854	11%

台人次都遠遠高於其他地區，因此我們想主要針對亞洲國家地區之來台人次進行時間序列分析，並根據分析預測結果提出建議。



而針對亞洲地區，該網站資料也顯示了近十年來日本、韓國、馬來西亞、大陸、和香港的來台人次變化。



除了中國大陸之外，其餘國家地區來台人次都呈現上升趨勢。由於中國大陸來台人次較容易受到政治或其他因素之影響，較難預測，因此我們並未對中國大陸來台人數進行分析預測。而我們最終選定以日本、馬來西亞、新加坡、香港、澳門和印度這六個國家地區來進行未來一年的來台人次預測。

三、資料描述

- 資料來源：交通部觀光局觀光統計資料庫

在交通部觀光局網站的資料庫中，選取六個國家地區的來台人次資料，時間範圍包含 97 年 5 月到 108 年 4 月。

網址：<https://stat.taiwan.net.tw/inboundSearch>

- 資料描述：

逾十年來每月日本、香港、澳門、馬來西亞、新加坡、印度來台觀光人次，主要欄位包含：Year、Month、Japan、Hong Kong、Macao、Malaysia、Singapore、India（單位：當月來台人次(人)）。

- 範例資料：

年別 (Year)	月別 (Month)	日本 (Japan)	香港 (Hong Kong)	澳門 (Macao)	馬來西亞 (Malaysia)
97	5	84,643	53,851	1,789	14,434
97	6	86,022	62,472	3,851	10,733
97	7	80,201	54,509	2,764	7,704
97	8	84,168	59,785	4,397	8,334
97	9	90,716	42,758	4,153	15,002

四、資料敘述統計

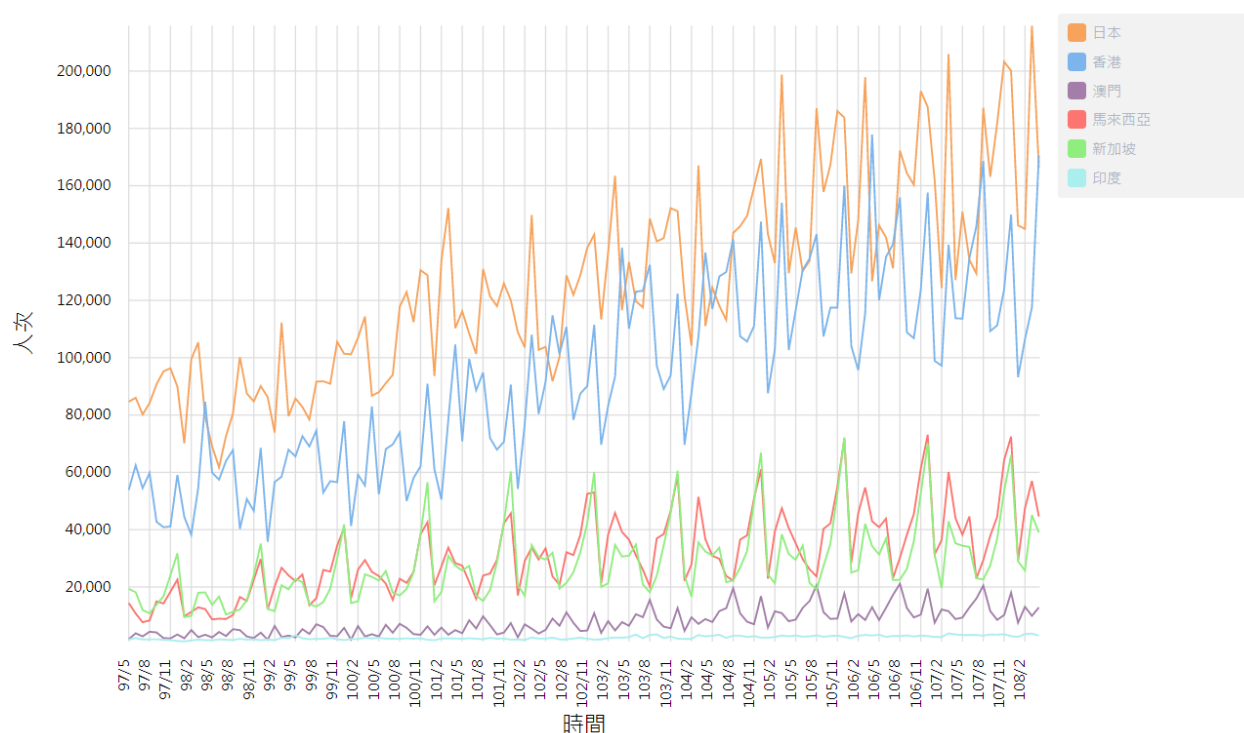
	Mean	Median	Std. dev	Max	Min	Correlation coef
Japan	139434.165	133793.0	29661.2081	215809	86743	0.6918252
HK	106577.897	107358.0	28952.3728	177852	49986	0.6943874
Macao	8345.198	7647.5	4551.7565	21141	1606	0.7186604

Malaysia	33384.455	30204.0	13940.5351	73122	8694	0.6422593
Singapore	32019.381	29570.0	13025.0068	72216	15044	0.2915491
India	2543.289	2556.0	592.8169	3822	1445	0.7736049

從每月來台人次的平均值中，我們可以看到日本、香港每月來台人次的平均超過 10 萬人，遠高於其他四個國家地區，其最多人次來台的月份甚至分別超過二十萬與十五萬人次。馬來西亞和新加坡的平均每月來台人次為三萬多人次，澳門八千多人次，印度最少，為兩千多人次。而標準差的部分則跟該國家地區的每月來台人次平均值呈現正比關係。

而除新加坡外，每月來台人次和時間的相關係數都逾 0.6，呈現趨近高度相關的現象，印度和澳門雖來台人次較低，其每月來台人次和時間的相關程度卻是六者中較高者。

五、資料視覺化



我們以每月為單位繪出六個國家地區來台人次折線圖，由上圖我們觀察到六個國家地區來台的人次可能存在長期正向趨勢，且呈現規率性的變動，即存在季節變動 (Seasonal Effect) 的因素，因此我們將採用時間序列分析，並將季節變動因素考慮進去。

六、時間序列分析

我們採用以下三種方法來進行時間序列分析：

1. Smoothing by Centered Moving Average
2. Smoothing by Linear Regression Model
3. Regression Model by Indicator Variables

使用第一、二種方法，我們能得到各國季節性指標，反應季節規律的變動性，其公式為：

$$\text{Seasonal index} = \frac{\text{Seasonal averaged ratio} \times (\text{number of seasons})}{\text{Sum of averaged ratios}}$$

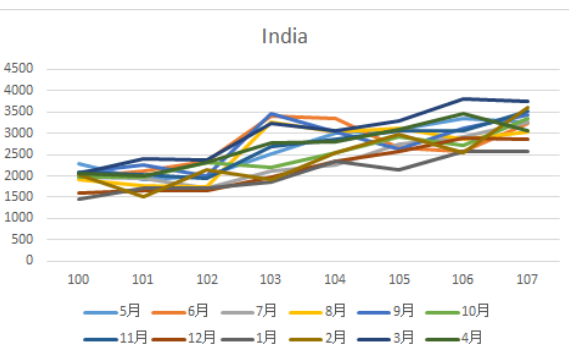
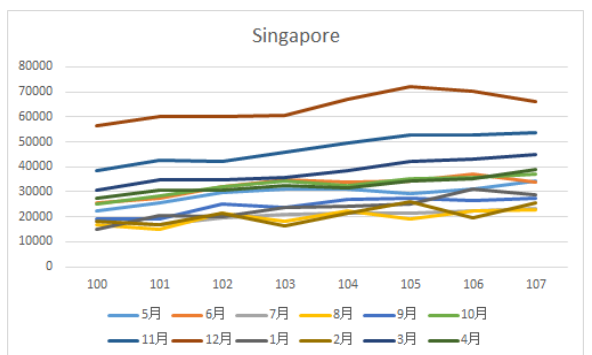
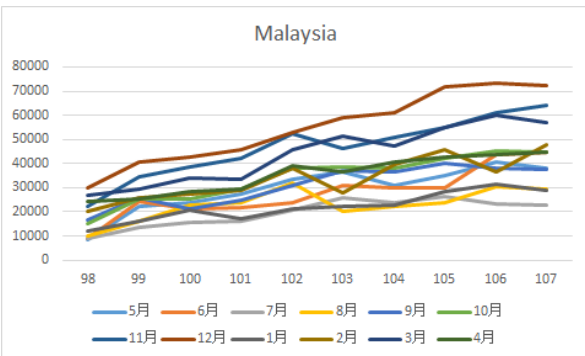
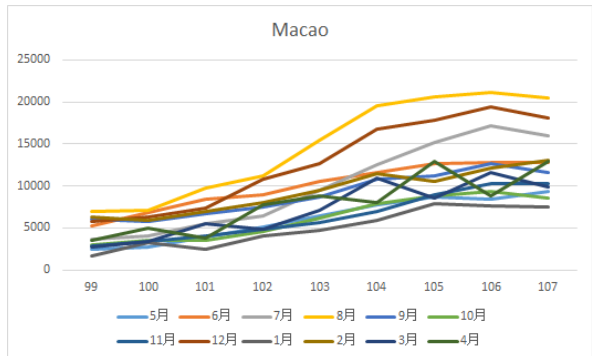
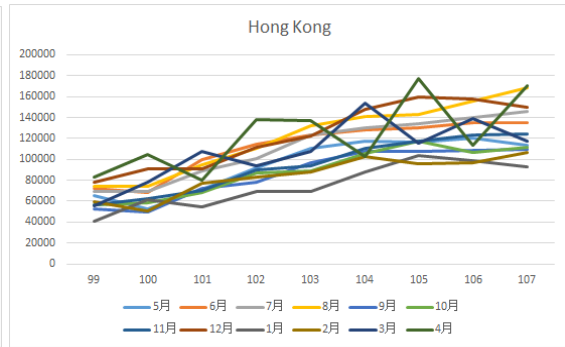
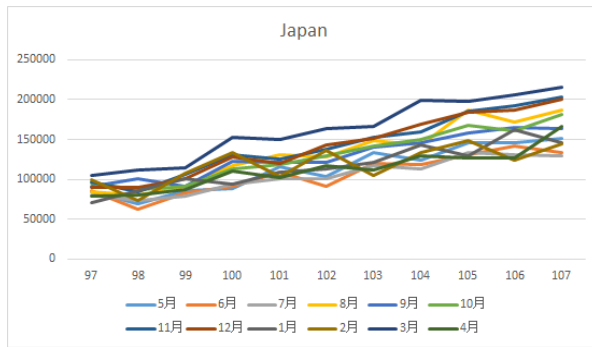
季節性指標代表意義為某個季節的時間序列值與整體季節性平均值之間的比率。得到季節性指標後，我們能用來消除時間序列的季節影響，並建立模型：

$$\text{Deseasonalized_Visitors} = \beta_0 + \beta_1 \times t + e$$

where Deseasonalized_Visitors = numbers of inbound visitors which have been deseasonalized, t = time in chronological order

最後再對模型的預測結果根據季節進行調整。

第三種方式使用 Regression Model by Indicator Variables，由於資料顯示個國家地區來台人次以十二個月做為一個規律循環，並且根據以下折線圖，我們認為這六個國家地區的十二個月分，來台人次都與時間呈正向線性關係：



因此我們首先建立模型，並設了十一個 Dummy Variables：

$$\text{Visitors} = \beta_0 + \beta_1 * t + \beta_2 * Q1 + \beta_3 * Q2 + \beta_4 * Q3 + \beta_5 * Q4 + \beta_6 * Q5 + \beta_7 * Q6 + \beta_8 * Q7 + \beta_9 * Q8 + \beta_{10} * Q9 + \beta_{11} * Q10 + \beta_{12} * Q11 + e$$

where Visitors = number of inbound visitors, t = time in chronological order, and Q_i = indicator variable (0, 1).

接著我們將分別對日本、香港、澳門、馬來西亞、新加坡、印度進行以上三種時間序列的預測分析之建模，並選擇最適模型。

(一) 針對各國選擇分析模型

我們針對六國家地區資料進行上述三種時間序列分析建模，先檢查模型之 **Residual Analysis**，並排除掉 **Residual Analysis** 沒有通過之方法，接著 **assess the fitted model**，若還有兩個以上的方法適合用來預測該國家地區的來台人次，則透過誤差比較從剩餘方法中選擇誤差較小者作為我們的分析模型。以下我們分別針對各國家地區進行上述步驟，並且事先將 **Residual Analysis** 之結果整理成表格 (原始輸出結果可參考附檔 **html** 檔，並且 **Normality test** 部分均使用 **shapiro test** 來做檢定)：

- 日本：

Residual analysis:

1. Normality test

H_0 : Errors are normally distributed.

H_1 : Errors are not normally distributed.

2. Homoscedasticity & Heteroscedasticity

H_0 : The variance of e is the same for all values of x .

H_1 : The variance of e is not the same for all values of x .

3. Dependence of the Error Variable

H_0 : Randomness exists.

H_1 : Randomness does not exist.

日本	normality test	Heteroscedasticity	Dependence of the Error Variable (run test)
CMV	0.6509	no sign	2.741969e-05
SLR	0.8631	no sign	0.0001208429
Dummy	0.9987	no sign	0.1621229

我們可以看到在殘差分析中，CMV 和 SLR 的建模過程中，使用 **run test** 檢定誤差項的獨立性(Dependence of the Error Variable)時，其結果為拒絕假設檢定($p\text{-value} < 0.05$)，即偏向對立假設(alternative hypothesis)：樣本不是隨機選擇的。因此，我們不使用這兩個方法去預測日本來台人次。

我們使用 Regression model by indicator variables 並進一步 assess the fitted model :

Residual standard error: 11700 on 119 degrees of freedom
Multiple R-squared: 0.8952, Adjusted R-squared: 0.8846
F-statistic: 84.67 on 12 and 119 DF, p-value: < 2.2e-16

1. The Standard Error of Estimate

The standard error of estimate (11700) is small compared to mean_y (126028).
We can conclude that the model fit the data.

2. The Coefficient of Determination

$r^2 = 89.52\%$ of the variation in the fund value is explained by the variation in Des_x. The rest (10.48%) remains unexplained by this model.

3. The *F*-test of ANOVA

$H_0: \beta_{a_1} = \beta_{a_2} = \dots = \beta_{a_{11}} = 0$
 $H_1: \text{At least one } \beta_{a_i} \text{ is not equal to zero.}$

Analysis of Variance Table

Response: Visitors

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Data.Visitors	12	1.3917e+11	1.1598e+10	84.667	< 2.2e-16 ***
Residuals	119	1.6301e+10	1.3698e+08		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Since $p\text{-value} < 2.2e-16 < 0.05$, there is sufficient evidence to reject the null hypothesis. At least one of the β_{a_i} is not equal to zero. Thus, at least one independent variable is related to y . This regression model is valid.

4. Testing of the Coefficients (t test)

$H_0: \beta_{a_i} = 0$
 $H_1: \beta_{a_i} \neq 0$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	92983.23	3960.76	23.476	< 2e-16	***
Data.Visitors	735.67	26.84	27.405	< 2e-16	***
Data.VisitorsQ1	-27066.76	4990.60	-5.424	3.11e-07	***
Data.VisitorsQ2	-24762.07	4990.81	-4.962	2.35e-06	***
Data.VisitorsQ3	17594.53	4991.18	3.525	0.000602	***
Data.VisitorsQ4	-32854.05	4991.68	-6.582	1.31e-09	***
Data.VisitorsQ5	-23681.02	4994.06	-4.742	5.93e-06	***
Data.VisitorsQ6	-31858.69	4993.12	-6.381	3.52e-09	***
Data.VisitorsQ7	-33875.27	4992.33	-6.785	4.78e-10	***
Data.VisitorsQ8	-5481.86	4991.68	-1.098	0.274335	
Data.VisitorsQ9	-10914.89	4991.18	-2.187	0.030710	*
Data.VisitorsQ10	-10479.84	4990.81	-2.100	0.037857	*
Data.VisitorsQ11	1678.40	4990.60	0.336	0.737228	

Since all the p_value except for that of Q8 and Q11 are < 0.05 , there is overwhelming evidence to infer that t and all the Q_i except for Q8 and Q11 affects the number of visitors. That is, we do not have enough evidence to infer that Q8 and Q11 affects the number of visitors.

因此在後續的分析中，我們可以採用 **Regression model by indicator variables** 對日本來台人次作分析，並且我們有足夠證據證明在此模型中時間對於日本來台人次是重要變數。

- 香港：

Residual analysis:

1. Normality test :

H_0 : Errors are normally distributed.

H_1 : Errors are not normally distributed.

2. Homoscedasticity & Heteroscedasticity

H_0 : The variance of e is the same for all values of x.

H_1 : The variance of e is not the same for all values of x.

3. Dependence of the Error Variable

H_0 : Randomness exists.

H_1 : Randomness does not exist.

	normality test	Heteroscedasticity	Dependence of the Error Variable (run test)
--	----------------	--------------------	---

CMV	0.3321	no sign	0.1219004
SLR	0.1219004	no sign	0.02032634
Dummy	0.06569	no sign	0.006789827

在殘差分析中，SLR 和 Dummy 的建模過程中，使用 run test 檢定誤差項的獨立性時，其結果為拒絕假設檢定(p-value < 0.05)，即偏向對立假設：樣本不是隨機選擇的。因此，我們不使用這兩個方法去預測香港來台人次。

```
Residual standard error: 11240 on 106 degrees of freedom
Multiple R-squared:  0.8166,    Adjusted R-squared:  0.8148
F-statistic: 471.8 on 1 and 106 DF,  p-value: < 2.2e-16
```

我們使用 CMV 進一步的 Assess the fitted model：

1. The Standard Error of Estimate：

The standard error of estimate (11240) is small compared to mean_y (102116).
We can conclude that the model fit the data.

2. The Coefficient of Determination：

$r^2 = 81.66\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(18.34%) remains unexplained by this model.

3. Testing of the Coefficients (t test)：

Testing the coefficients beta1.

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  61801.34    2148.09   28.77  <2e-16 ***
t              753.55     34.69   21.72  <2e-16 ***
```

Since p-value < 2e-16, we are 95% confident to reject H_0 . There is overwhelming evidence that t affects Des_x.

因此在後續的分析中，我們可以採用 **CMV** 對香港來台人次作分析，並且我們有足夠證據證明在此模型中，時間對於香港來台人次是重要變數。

- **澳門：**

Residual analysis :

1. Normality test :

H_0 : Errors are normally distributed.

H_1 : Errors are not normally distributed.

2. Homoscedasticity & Heteroscedasticity

H_0 : The variance of e is the same for all values of x .

H_1 : The variance of e is not the same for all values of x .

3. Dependence of the Error Variable

H_0 : Randomness exists.

H_1 : Randomness does not exist.

	normality test	Heteroscedasticity	Dependence of the Error Variable
CMV	0.2454	no sign	0.6989698
SLR	0.2504	no sign	0.6989698
Dummy	0.04282	no sign	0.08182361

我們可以看到在殘差分析中，**Dummy** 的建模過程中，使用 **shapiro test** 檢定誤差項的常態分布時，其結果為拒絕假設檢定($p\text{-value} < 0.05$)，即偏向對立假設：誤差不是常態分布的。因此，我們不使用這個方法去預測澳門來台人次。

我們使用 **CMV** 和 **SLR** 進一步的 Assess the fitted model :

CMV

Residual standard error: 1284 on 106 degrees of freedom
Multiple R-squared: 0.8683, Adjusted R-squared: 0.8671
F-statistic: 698.9 on 1 and 106 DF, p-value: < 2.2e-16

1. The Standard Error of Estimate

The standard error of estimate (1284) is small compared to mean_y (8773).
We can conclude that the model fit the data.

2. The Coefficient of Determination

$r^2 = 86.83\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(13.17%) remains unexplained by this model.

3. Testing of the Coefficients (t test) :

Testing the coefficients beta1.

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3168.808	245.332	12.92	<2e-16 ***
t	104.746	3.962	26.44	<2e-16 ***

Since p-value < 2e-16 , we are 95% confident to reject H_0 . There is overwhelming evidence that t affects Des_x.

SLR

Residual standard error: 1299 on 106 degrees of freedom
Multiple R-squared: 0.8659, Adjusted R-squared: 0.8646
F-statistic: 684.5 on 1 and 106 DF, p-value: < 2.2e-16

1. The Standard Error of Estimate :

The standard error of estimate (1299) is small compared to mean_y (8775).
We can conclude that the model fit the data.

2. The Coefficient of Determination :

$r^2 = 86.59\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(13.41%) remains unexplained by this model.

3. Testing of the Coefficients (t test) :

$H_0: \beta_1 = 0$

$$H_1 : \beta_{a_i} \neq 0$$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3163.558	248.214	12.74	<2e-16 ***
t	104.878	4.009	26.16	<2e-16 ***

Since the $p_value < 0.05$, there is overwhelming evidence to infer that t affects the # of visitors.

因為我們可以採用 **CMV** 和 **SLR** 來對澳門來台人次作分析預測，故我們比較兩者的 **error** 來選擇最適的模型。

	CMV	SLR
MAD	9.535612e+02	9.616170e+02
MSE	1.623418e+06	1.683360e+06
MAPE	1.180574e+01	1.177422e+01

由於 **CMV** 的 **MAD**、**MSE** 分別比 **SLR** 的小了 8、59942，並且它的 **MAPE** 只比 **SLR** 略大了 0.03，因此我們選擇 **CMV** 作為預測澳門來台人次的模型，並且我們有足夠證據證明在此 **CMV** 模型中，時間對於澳門來台人次是重要變數。

- 馬來西亞：

Residual analysis:

1. Normality test :

H_0 : Errors are normally distributed.

H_1 : Errors are not normally distributed.

2. Homoscedasticity & Heteroscedasticity

H_0 : The variance of e is the same for all values of x .

H_1 : The variance of e is not the same for all values of x .

3. Dependence of the Error Variable

H_0 : Randomness exists.

H_1 : Randomness does not exist.

	normality test	Heteroscedasticity	Dependence of the Error Variable
CMV	0.5239	no sign	0.1424388
SLR	0.4103	no sign	0.1424388
Dummy	0.7282	no sign	0.06673393

我們可以看到在殘差分析中，在建模過程中使用 normality test 和 run test 檢定誤差項的常態分佈與獨立性時，沒有結果為拒絕假設檢定($p\text{-value} < 0.05$)。

因此我們使用全部三種方法進一步的 **Assess the fitted model** :

CMV

```
Residual standard error: 3894 on 118 degrees of freedom
Multiple R-squared:  0.8233,    Adjusted R-squared:  0.8218
F-statistic: 549.7 on 1 and 118 DF,  p-value: < 2.2e-16
```

1. The Standard Error of Estimate :

The standard error of estimate (3894) is small compared to mean_y (33423).
We can conclude that the model fit the data.

2. The Coefficient of Determination :

$r^2 = 82.33\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(17.67%) remains unexplained by this model.

3. Testing of the Coefficients (t test) :

Testing the coefficients beta1.

H0: beta1 = 0

H1: beta1 != 0

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	19109.24	706.45	27.05	<2e-16 ***
t	240.58	10.26	23.45	<2e-16 ***

Since $p\text{-value} < 2e-16$, we are 95% confident to reject H0. There is overwhelming evidence that t affects Des_x.

SLR

Residual standard error: 3922 on 118 degrees of freedom
Multiple R-squared: 0.8206, Adjusted R-squared: 0.819
F-statistic: 539.6 on 1 and 118 DF, p-value: < 2.2e-16

1. The Standard Error of Estimate

The standard error of estimate (3922) is small compared to mean_y (33452).
We can conclude that the model fit the data.

2. The Coefficient of Determination

$r^2 = 82.06\%$ of the variation in the fund value is explained by the variation in Des_x. The rest (17.94%) remains unexplained by this model.

3. Testing of the Coefficients (t test)

Testing the coefficients β_{t_1} .

$H_0: \beta_{t_1} = 0$

$H_1: \beta_{t_1} \neq 0$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	19166.11	711.63	26.93	<2e-16 ***
t	240.10	10.34	23.23	<2e-16 ***

Since p-value <2e-16, we are 95% confident to reject H_0 . There is overwhelming evidence that t affects Des_x.

Regression model by indicator variables

Residual standard error: 4566 on 107 degrees of freedom
Multiple R-squared: 0.9025, Adjusted R-squared: 0.8915
F-statistic: 82.5 on 12 and 107 DF, p-value: < 2.2e-16

1. The Standard Error of Estimate

The standard error of estimate (4566) is small compared to mean_y (33560).
We can conclude that the model fit the data.

2. The Coefficient of Determination

$r^2 = 90.25\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(9.75%) remains unexplained by this model.

3. The F-test of ANOVA

$H_0: \beta_{t_1} = \beta_{t_2} = \dots = \beta_{t_{11}} = 0$

$H_1: \text{At least one } \beta_{t_i} \text{ is not equal to zero.}$

Analysis of Variance Table

Response: Visitors

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Data.Visitors	12	2.0642e+10	1720155928	82.504	< 2.2e-16 ***
Residuals	107	2.2309e+09	20849299		

Since $p\text{-value} < 2.2e-16 < 0.05$, there is sufficient evidence to reject the null hypothesis. At least one of the β_{α_i} is not equal to zero. Thus, at least one independent variable is related to y . This regression model is valid.

4. Testing of the Coefficients (t test) :

$$H_0 : \beta_{\alpha_i} = 0$$

$$H_1 : \beta_{\alpha_i} \neq 0$$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	40156.86	1621.46	24.766	< 2e-16 ***
Data.Visitorst	241.83	12.09	19.997	< 2e-16 ***
Data.VisitorsQ1	-33016.73	2042.06	-16.168	< 2e-16 ***
Data.VisitorsQ2	-21626.45	2042.17	-10.590	< 2e-16 ***
Data.VisitorsQ3	-11611.88	2042.35	-5.686	1.14e-07 ***
Data.VisitorsQ4	-20363.90	2042.60	-9.970	< 2e-16 ***
Data.VisitorsQ5	-23487.12	2043.78	-11.492	< 2e-16 ***
Data.VisitorsQ6	-25547.65	2043.31	-12.503	< 2e-16 ***
Data.VisitorsQ7	-34017.77	2042.92	-16.652	< 2e-16 ***
Data.VisitorsQ8	-30835.10	2042.60	-15.096	< 2e-16 ***
Data.VisitorsQ9	-23226.72	2042.35	-11.373	< 2e-16 ***
Data.VisitorsQ10	-20220.25	2042.17	-9.901	< 2e-16 ***
Data.VisitorsQ11	-7870.97	2042.06	-3.854	0.000198 ***

Since the $p\text{-value} < 0.05$, there is overwhelming evidence to infer that t and all the dummy variables affect the # of visitors.

因為三種方法都可對馬來西亞來台人次作分析預測，故我們比較三者的 **error** 來選擇最好的模型。

	CMV	SLR	Dummy
MAD	2.729081e+03	2.723617e+03	3.404047e+03
MSE	1.197448e+07	1.196451e+07	1.859062e+07
MAPE	1.035718e+01	1.029726e+01	1.245622e+01

由於 SLR 的 **MAD**、**MSE**、**MAPE** 均為三種方法中最小，因此我們選擇 SLR 作為預測馬來西亞來台人次的模型，並且我們有足夠證據證明在此 SLR 模型中，時間對於馬來西亞來台人次是重要變數。

● 新加坡：

Residual analysis:

1. Normality test :

H_0 : Errors are normally distributed.

H_1 : Errors are not normally distributed.

2. Homoscedasticity & Heteroscedasticity

H_0 : The variance of e is the same for all values of x.

H_1 : The variance of e is not the same for all values of x.

3. Dependence of the Error Variable

H_0 : Randomness exists.

H_1 : Randomness does not exist.

	normality test	Heteroscedasticity	Dependence of the Error Variable
CMV	0.02356	no sign	0.02399074
SLR	0.02958	no sign	0.06476695
Dummy	0.8984	no sign	0.06476695

我們可以看到在殘差分析中，CMV 和 SLR 的建模過程中，使用 Shapiro-Wilk normality test，其結果為拒絕假設檢定($p\text{-value} < 0.05$)，即偏向對立假設：樣本不是常態分佈的。因此，我們不使用這兩個方法去預測新加坡來台人次。

我們使用 Regression model by indicator variables 進一步的 Assess the fitted model :

Regression model by indicator variables

Residual standard error: 2151 on 83 degrees of freedom

Multiple R-squared: 0.9763, Adjusted R-squared: 0.9729

F-statistic: 285.2 on 12 and 83 DF, p-value: $< 2.2e-16$

1. The Standard Error of Estimate :

The standard error of estimate (2151) is small compared to mean_y (32108).
We can conclude that the model fit the data.

2. The Coefficient of Determination

$r^2 = 97.63\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(2.37%) remains unexplained by this model.

3. The *F*-test of ANOVA

$H_0: \beta_{a_1} = \beta_{a_2} = \dots = \beta_{a_{11}} = 0$

H_1 : At least one β_{a_i} is not equal to zero.

Analysis of Variance Table

Response: Visitors

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Data.Visitors	12	1.5830e+10	1319167172	285.15	< 2.2e-16 ***
Residuals	83	3.8398e+08	4626218		

Since $p\text{-value} < 2.2e-16 < 0.05$, there is sufficient evidence to reject the null hypothesis. At least one of the β_{a_i} is not equal to zero. Thus, at least one independent variable is related to y . This regression model is valid.

4. Testing of the Coefficients (t test) :

$H_0: \beta_{a_i} = 0$

$H_1: \beta_{a_i} \neq 0$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	58115.747	855.175	67.96	<2e-16 ***
Data.Visitorst	122.406	7.984	15.33	<2e-16 ***
Data.VisitorsQ1	-40575.281	1075.462	-37.73	<2e-16 ***
Data.VisitorsQ2	-43667.936	1075.551	-40.60	<2e-16 ***
Data.VisitorsQ3	-26455.467	1075.699	-24.59	<2e-16 ***
Data.VisitorsQ4	-31986.998	1075.906	-29.73	<2e-16 ***
Data.VisitorsQ5	-33913.410	1076.883	-31.49	<2e-16 ***
Data.VisitorsQ6	-31027.066	1076.499	-28.82	<2e-16 ***
Data.VisitorsQ7	-43006.597	1076.173	-39.96	<2e-16 ***
Data.VisitorsQ8	-43834.252	1075.906	-40.74	<2e-16 ***
Data.VisitorsQ9	-39339.158	1075.699	-36.57	<2e-16 ***
Data.VisitorsQ10	-31266.689	1075.551	-29.07	<2e-16 ***
Data.VisitorsQ11	-16795.969	1075.462	-15.62	<2e-16 ***

Since all the $p\text{-value} < 0.05$, there is overwhelming evidence to infer that t and all the dummy variables affect the # of visitors.

因此在後續的分析中，我們可以採用 Regression model by indicator variables 對新加坡來台人次作分析，並且我們有足夠證據證明在此模型中，時間對於新加坡來台人次是重要變數。

● 印度：

Residual analysis:

1. Normality test :

H_0 : Errors are normally distributed.

H_1 : Errors are not normally distributed.

2. Homoscedasticity & Heteroscedasticity

H_0 : The variance of e is the same for all values of x .

H_1 : The variance of e is not the same for all values of x .

3. Dependence of the Error Variable

H_0 : Randomness exists.

H_1 : Randomness does not exist.

	normality test	Heteroscedasticity	Dependence of the Error Variable
CMV	0.08338	no sign	0.8374104
SLR	0.1295	no sign	0.8374104
Dummy	0.03138	no sign	0.8374104

在 Dummy 的建模過程中，使用 Shapiro-Wilk normality test，其結果為拒絕假設檢定($p\text{-value} < 0.05$)，即偏向對立假設：樣本不是常態分佈的。因此，我們不使用這個方法去預測印度來台人次。

我們使用另外兩個方法進一步的 Assess the fitted model：

CMV

Residual standard error: 271.8 on 94 degrees of freedom
Multiple R-squared: 0.7527, Adjusted R-squared: 0.7501
F-statistic: 286.2 on 1 and 94 DF, p-value: < 2.2e-16

1. The Standard Error of Estimate :

The standard error of estimate (272) is small compared to mean_y (2551). We can conclude that the model fit the data.

2. The Coefficient of Determination :

$r^2 = 75.27\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(24.73%) remains unexplained by this model.

3. Testing of the Coefficients (t test) :

$$H_0 : \beta_{a1} = 0$$

$$H_1 : \beta_{a1} \neq 0$$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1746.691	55.059	31.72	<2e-16 ***
t	16.937	1.001	16.92	<2e-16 ***

Since the p_value < 0.05, there is overwhelming evidence to infer that t affects the # of visitors.

SLR

```
Residual standard error: 268.7 on 94 degrees of
freedom
Multiple R-squared:  0.7581,    Adjusted R-squared:
0.7556
F-statistic: 294.6 on 1 and 94 DF,  p-value: < 2.2e-
16
```

1. The Standard Error of Estimate :

The standard error of estimate (268.7) is small compared to mean_y (2551). We can conclude that the model fit the data.

2. The Coefficient of Determination :

$r^2 = 75.81\%$ of the variation in the fund value is explained by the variation in Des_x. The rest(24.19%) remains unexplained by this model.

3. Testing of the Coefficients (t test) :

$$H_0 : \beta_{a1} = 0$$

$$H_1 : \beta_{a1} \neq 0$$

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1744.0242	54.4274	32.04	<2e-16 ***
t	16.9885	0.9897	17.16	<2e-16 ***

Since the p_value < 2e-16 < 0.05, there is overwhelming evidence to infer that t affects the # of visitors.

因為 CMV 和 SLR 方法都可對印度來台人次作分析預測，故我們比較兩者的 error 來選擇最適的模型。

	CMV	SLR
MAD	212.605414	208.551551
MSE	75985.384251	74639.590962
MAPE	8.505054	8.371494

由於 SLR 的的 MAD、MSE、MAPE 均為較小，因此我們選擇 SLR 作為預測印度來台人次的模型，並且我們有足夠證據證明在此 SLR 模型中，時間對於印度來台人次是重要變數。

整理以上六國家地區之模型選擇結果，日本、新加坡採用 Regression model by indicator variables 作為時間序列分析之模型，香港、澳門採用 CMV 作為時間序列分析之模型，馬來西亞、印度採用 SLR 作為時間序列分析之模型。

(二) 建模預測之結果

1. 日本 (Regression model by indicator variables)

```

Call:
lm(formula = Visitors ~ Data.Visitors)

Residuals:
    Min       1Q   Median       3Q      Max
-30005   -7843    167    7453   32467

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    92983.23    3960.76   23.476 < 2e-16 ***
Data.Visitorst     735.67     26.84   27.405 < 2e-16 ***
Data.VisitorsQ1  -27066.76    4990.60   -5.424 3.11e-07 ***
Data.VisitorsQ2  -24762.07    4990.81   -4.962 2.35e-06 ***
Data.VisitorsQ3   17594.53    4991.18    3.525 0.000602 ***
Data.VisitorsQ4  -32854.05    4991.68   -6.582 1.31e-09 ***
Data.VisitorsQ5  -23681.02    4994.06   -4.742 5.93e-06 ***
Data.VisitorsQ6  -31858.69    4993.12   -6.381 3.52e-09 ***
Data.VisitorsQ7  -33875.27    4992.33   -6.785 4.78e-10 ***
Data.VisitorsQ8   -5481.86    4991.68   -1.098 0.274335
Data.VisitorsQ9  -10914.89    4991.18   -2.187 0.030710 *
Data.VisitorsQ10 -10479.84    4990.81   -2.100 0.037857 *
Data.VisitorsQ11   1678.40    4990.60    0.336 0.737228

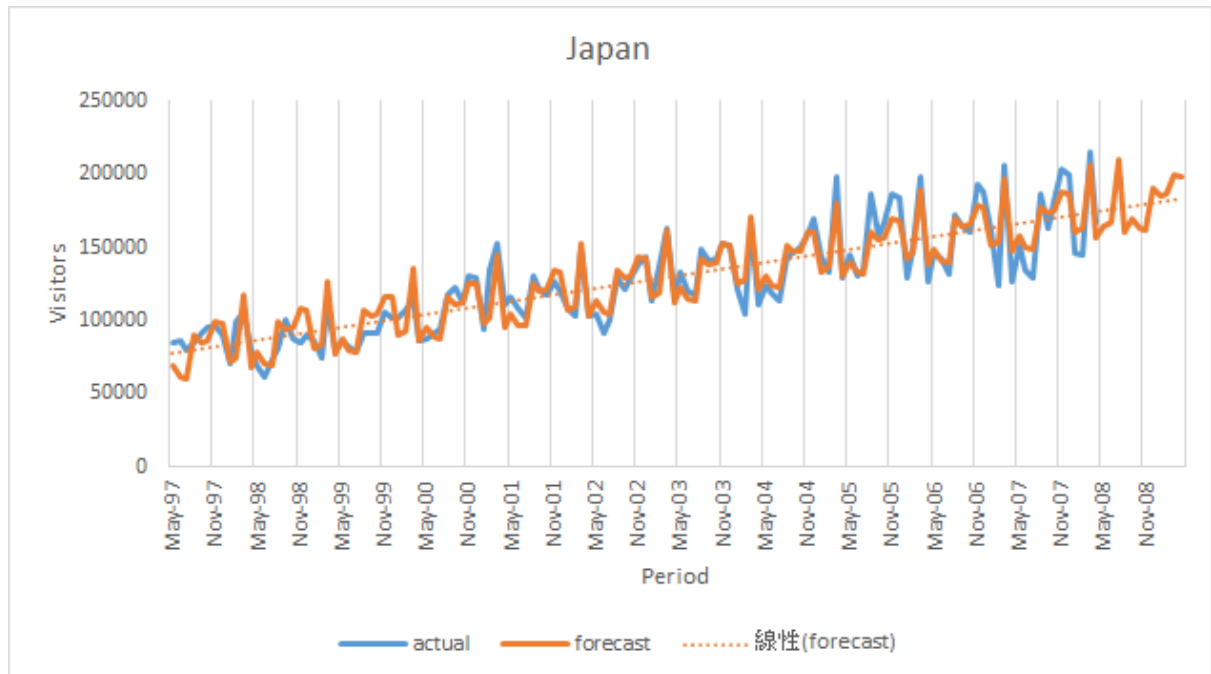
```

- 趨勢線:

$$\text{Visitors} = 92983.23 + 735.67 * t - 27066.76 * Q1 - 24762.07 * Q2 + 17594.53 * Q3 - 32854.05 * Q4 - 23681.02 * Q5 - 31858.69 * Q6 - 33875.27 * Q7 - 5481.86 * Q8 - 10914.89 * Q9 - 10479.84 * Q10 + 1678.4 * Q11$$

根據趨勢線我們可以看到的是，來台人次時間序列的線性趨勢，趨勢方程式的斜率為 735.67，表示來台人次的平均成長約每月 736 人次。而和同年十二月相比，除了三月和十一月的係數為正外，其他月份的平均來台人次均較十二月低。整體而言，以三月最高、七月最低。

- 未來一年來台人次預測(108 年五月~109 年四月):



By Regression model by indicator variables, we forecast that the number of visitors from Japan for the next 12 months are 155143.78, 158832.10, 220215.00, 147621.57, 163669.24, 149796.74, 146693.08, 194293.32, 180323.86, 210573.68, 202387.93, 201723.35 respectively.

其中我們可以看到，108年7月、109年2、3、4月的預測人次會超過20萬人次。108年8、10、11月預測人次低於15萬人次。

2. 香港(CMV)：

```
### Linear Model ###
Call:
lm(formula = Des_x ~ t)

Residuals:
    Min       1Q   Median       3Q      Max
-35270  -6700    354   7274  33608

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  61801.34   2148.09   28.77  <2e-16 ***
t             753.55    34.69   21.72  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

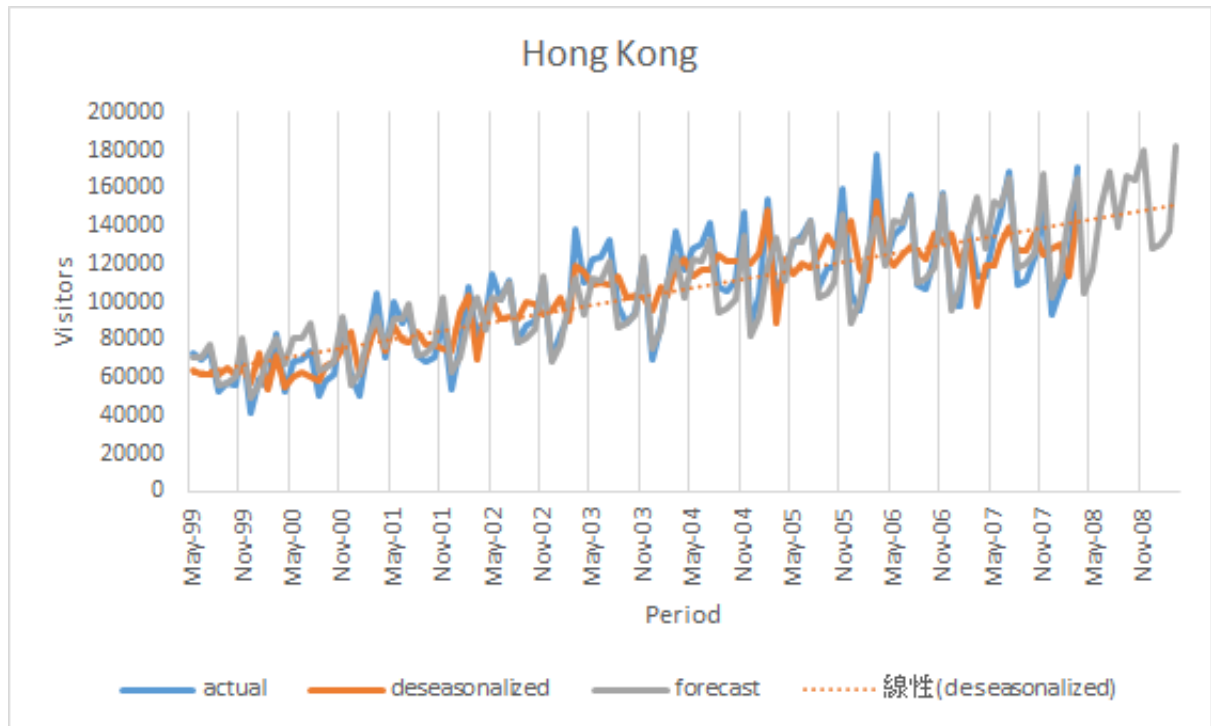
Residual standard error: 11240 on 106 degrees of freedom
Multiple R-squared:  0.8166,    Adjusted R-squared:  0.8148
F-statistic: 471.8 on 1 and 106 DF,  p-value: < 2.2e-16
```

- 趨勢線:

$$\text{Visitors} = 61801.34 + 753.55 * t$$

根據趨勢線我們可以看到的是，來台人次時間序列的線性趨勢，趨勢方程式的斜率為 753.55，表示去掉季節影響後，來台人次的平均成長約每月 754 人次。

- 未來一年預測(108 年五月~109 年四月):



By CMV, we forecast that the number of visitors from Hong Kong for the next 12 months will be 103529.51, 112451.53, 150877.93, 162155.38, 139664.27, 164917.41, 162528.71, 177895.17, 128127.75, 146923.38, 135307.21, and 178065.39 respectively.

其中我們可以看到，108 年 12 月、109 年 4 月的預測人次會超過 17 萬人次。108 年 5、6 月、109 年 1 月預測人次低於 13 萬人次。

3. 澳門(CMV)


```

### Linear Model ###
Call:
lm(formula = Des_x ~ t)

Residuals:
    Min       1Q   Median       3Q      Max
-3101.3  -736.1   -84.4    883.4   3767.3

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3168.808    245.332   12.92  <2e-16 ***
t             104.746      3.962   26.44  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1284 on 106 degrees of freedom
Multiple R-squared:  0.8683,    Adjusted R-squared:  0.8671
F-statistic: 698.9 on 1 and 106 DF,  p-value: < 2.2e-16

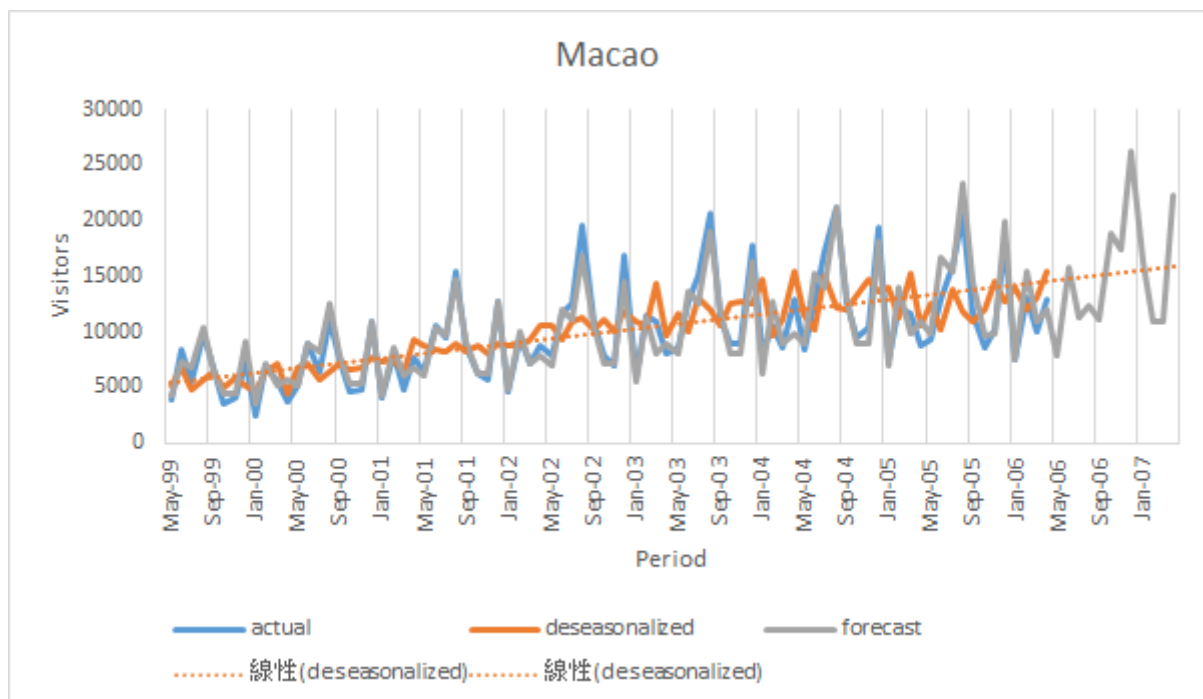
```

- 趨勢線:

$$\text{Visitors} = 3168.808 + 104.746 * t$$

根據趨勢線我們可以看到的是，來台人次時間序列的線性趨勢，趨勢方程式的斜率為 104.746，表示去掉季節影響後，來台人次的平均成長約每月 105 人次。

- 未來一年預測(108 年五月~109 年四月):



By CMV, we forecast that the number of visitors from Macao for the next 12 months will be 7800.82, 15848.05, 11201.14, 12363.07, 11068.55, 18862.17, 17387.61, 26277.88, 16236.04, 10926.63, 10940.54, 22297.09 respectively.

其中我們可以看到，108 年 12 月、109 年 4 月的預測人次會超過 2 萬人次。108 年 5、7、9 月、109 年 2、3 月預測人次低於 1.3 萬人次。

4. 馬來西亞(SLR)：

```
### Linear Model ###
Call:
lm(formula = Des_x ~ t)

Residuals:
    Min       1Q   Median       3Q      Max
-9693.9 -2386.3 -173.5  2155.6 13626.0

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 19166.11     711.63   26.93  <2e-16 ***
t           240.10       10.34   23.23  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

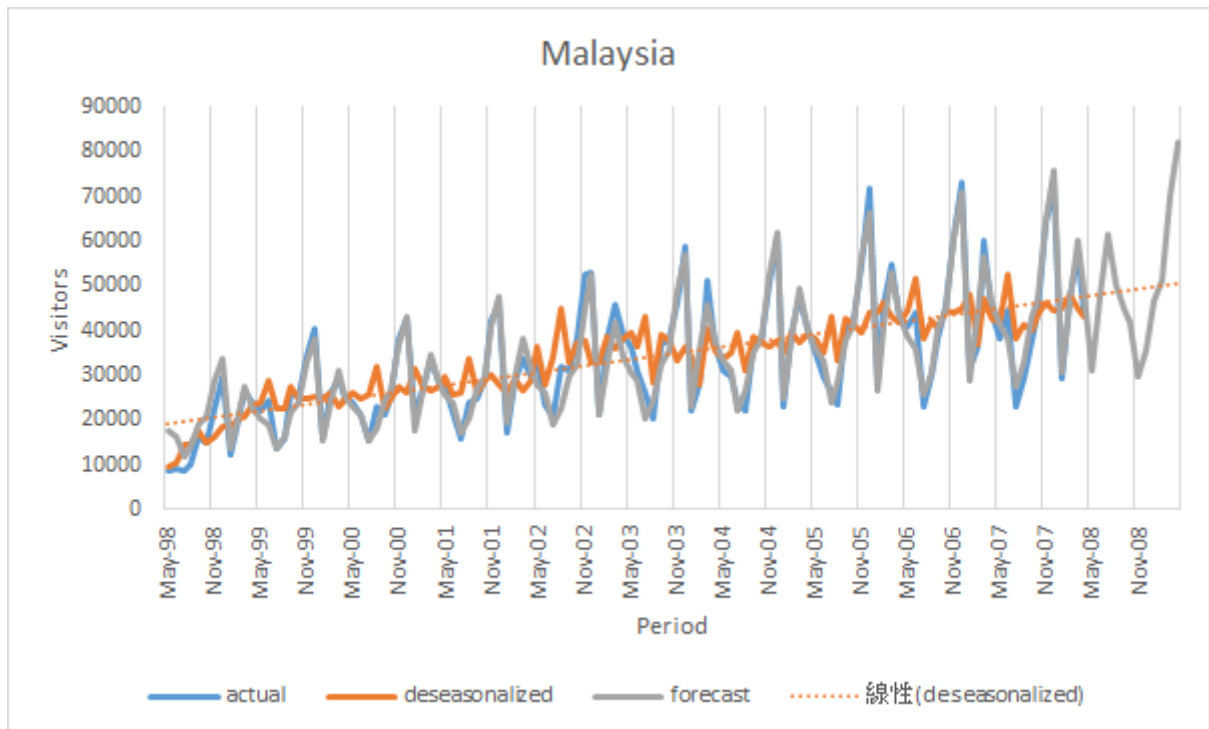
Residual standard error: 3922 on 118 degrees of freedom
Multiple R-squared:  0.8206,    Adjusted R-squared:  0.819
F-statistic: 539.6 on 1 and 118 DF,  p-value: < 2.2e-16
```

- 趨勢線:

$$\text{Visitors} = 19166.11 + 240.10 * t$$

根據趨勢線我們可以看到的是，來台人次時間序列的線性趨勢，趨勢方程式的斜率為 240.10，表示去掉季節影響後，來台人次的平均成長約每月 240 人次。

- 未來一年預測(108 年五月~109 年四月):



By SLR, we forecast that the number of visitors from Malaysia for the next 12 months will be 31189.49, 47780.60, 61563.53, 50174.54, 44918.51, 41675.48, 29818.83, 35430.62, 46955.87, 51187.61, 70204.90, 82041.74 respectively.

其中我們可以看到，109 年 3、4 月的預測人數會超過 7 萬人次。108 年 5、11、12 月預測人數低於 4 萬人次。

5. 新加坡(Regression model by indicator variables)：

```

Call:
lm(formula = Visitors ~ Data.Visitors)

Residuals:
    Min       1Q   Median       3Q      Max
-4759.7 -1638.0   -71.3   1537.0  5899.1

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   58115.747    855.175   67.96  <2e-16 ***
Data.Visitorst  122.406      7.984   15.33  <2e-16 ***
Data.VisitorsQ1 -40575.281   1075.462  -37.73  <2e-16 ***
Data.VisitorsQ2 -43667.936   1075.551  -40.60  <2e-16 ***
Data.VisitorsQ3 -26455.467   1075.699  -24.59  <2e-16 ***
Data.VisitorsQ4 -31986.998   1075.906  -29.73  <2e-16 ***
Data.VisitorsQ5 -33913.410   1076.883  -31.49  <2e-16 ***
Data.VisitorsQ6 -31027.066   1076.499  -28.82  <2e-16 ***
Data.VisitorsQ7 -43006.597   1076.173  -39.96  <2e-16 ***
Data.VisitorsQ8 -43834.252   1075.906  -40.74  <2e-16 ***
Data.VisitorsQ9 -39339.158   1075.699  -36.57  <2e-16 ***
Data.VisitorsQ10 -31266.689   1075.551  -29.07  <2e-16 ***
Data.VisitorsQ11 -16795.969   1075.462  -15.62  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2151 on 83 degrees of freedom
Multiple R-squared:  0.9763,    Adjusted R-squared:  0.9729
F-statistic: 285.2 on 12 and 83 DF,  p-value: < 2.2e-16

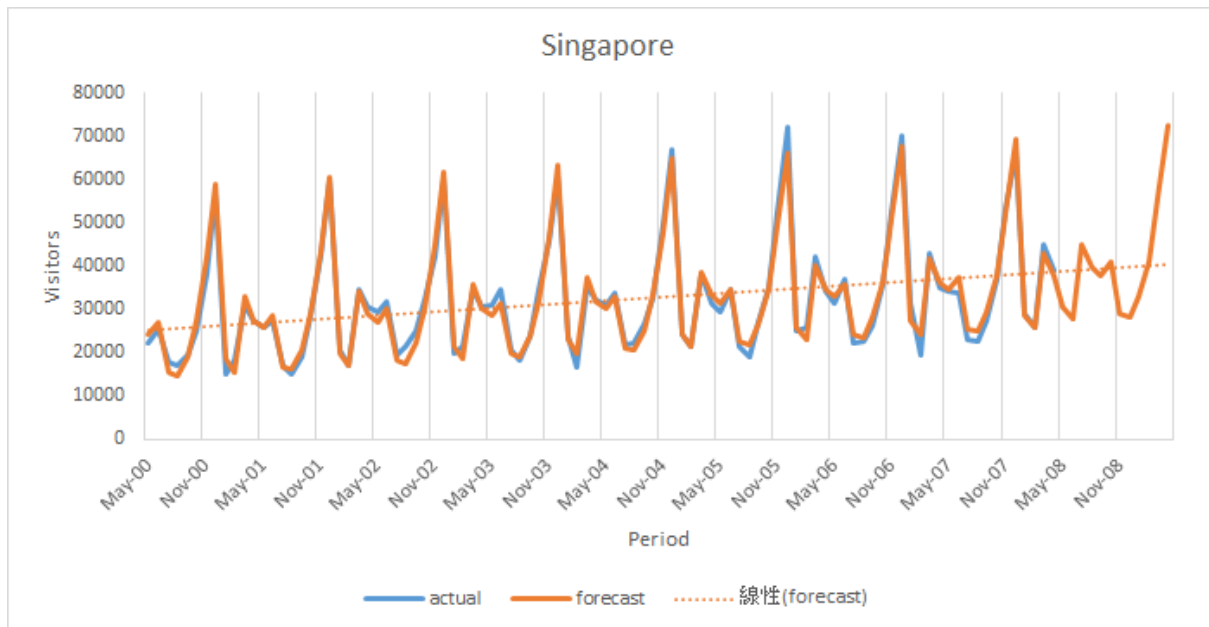
```

- 趨勢線:

$$\text{Visitors} = 58115.747 + 122.406 * t - 40575.281 * Q1 - 43667.936 * Q2 - 26455.467 * Q3 - 31986.998 * Q4 - 33913.41 * Q5 - 31027.066 * Q6 - 43006.597 * Q7 - 43834.252 * Q8 - 39339.158 * Q9 - 31266.689 * Q10 - 16795.969 * Q11$$

根據趨勢線我們可以看到的是，來台人次時間序列的線性趨勢，趨勢方程式的斜率為 122.406，表示來台人次的平均成長約每月 122 人次。而和同年十二月相比，所有係數均為負數，表示其他月份的平均來台人次均較十二月低。

- 未來一年預測(108 年五月~109 年四月):



By Regression model by indicator variables, we forecast that the number of visitors from Singapore for the next 12 months will be 30760.28, 27790.03, 45124.90, 39715.78, 37911.77, 40920.52, 29063.40, 28358.15, 32975.65, 41170.52, 55763.65, and 72682.02. respectively.

其中我們可以看到，109 年 4 月的預測人數會超過 7 萬人次。108 年 5、6、11、12 月、109 年 1 月預測人數低於 3.5 萬人次。

6. 印度(SLR)：

```
### Linear Model ###
Call:
lm(formula = Des_x ~ t)

Residuals:
    Min       1Q   Median       3Q      Max
-652.53 -184.84    6.85   138.13   766.88

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1744.0242    54.4274   32.04  <2e-16 ***
t           16.9885     0.9897   17.16  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

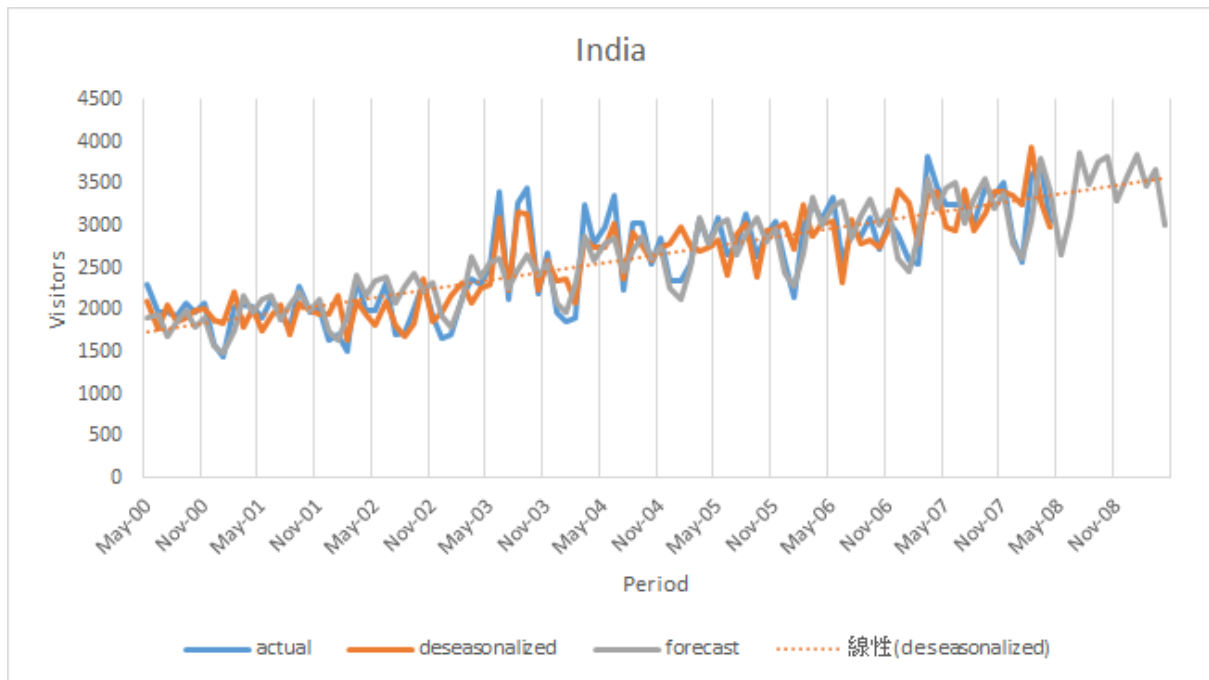
Residual standard error: 268.7 on 94 degrees of freedom
Multiple R-squared:  0.7581,    Adjusted R-squared:  0.7556
F-statistic: 294.6 on 1 and 94 DF,  p-value: < 2.2e-16
```

- 趨勢線:

$$\text{Visitors} = 1744.0242 + 16.9885 * t$$

根據趨勢線我們可以看到的是，來台人次時間序列的線性趨勢，趨勢方程式的斜率為 16.9885，表示去掉季節影響後，來台人次的平均成長約每月 17 人次。

- 未來一年預測(108年五月~109年四月):



By SLR, we forecast that the # of visitors from India for the next 12 months will be 2662.367, 3115.873, 3876.104, 3490.296, 3750.878, 3822.444, 3298.713, 3607.836, 3847.571, 3477.944, 3667.776, 3008.126 respectively.

其中我們可以看到，除了 108 年五月，其他月份的預測人數均超過 3 千人次。

(三) 季節性分析

- 香港、澳門、馬來西亞、印度之季節性指標

	香港(CMV)	澳門(CMV)	馬來西亞(SLR)	印度(SLR)
Jan	0.726234142	0.538677722	0.650069871	0.788867244
Feb	0.814102464	1.086512377	0.990912631	0.918618034
Mar	1.03953699	0.762453836	1.27042809	1.137053688
Apr	1.159604669	0.83558748	1.030299254	1.018799742
May	0.95532039	0.742835208	0.917845031	1.089459795
Jun	1.134033594	1.257046381	0.847420834	1.104794895
Jul	1.114922192	1.150742605	0.603384111	0.948763222
Aug	1.209227336	1.727143586	0.713472439	1.032626339
Sep	0.860024298	1.059835742	0.941008571	1.095913958
Oct	0.87248716	0.708410841	1.020901342	0.985861665
Nov	0.911517248	0.704528189	1.393514868	1.03468899
Dec	1.202989517	1.426226034	1.620742958	0.844552428

- 日本之 Dummy variable 係數與季節性指標

	Estimate(compae	Std. error	t value	Pr(> t)		CMV seasonal index	SLR seasonal index
Q1	-27066.76	4990.6	-5.424	3.11E-07 ***		0.921407	0.9065778
Q2	-24762.07	4990.81	-4.962	2.35E-06 ***		0.9582428	0.9431494
Q3	17594.53	4991.18	3.525	0.000602 ***		1.2596817	1.2489661
Q4	-32854.05	4991.68	-6.582	1.31E-09 ***		0.865704	0.8667965
Q5	-23681.02	4994.06	-4.742	5.93E-06 ***		0.922058	0.9311954
Q6	-31858.69	4993.12	-6.381	3.52E-09 ***		0.8520871	0.8695664
Q7	-33875.27	4992.33	-6.785	4.78E-10 ***		0.8459173	0.8563085
Q8	-5481.86	4991.68	-1.098	0.274335		1.0762823	1.0674273
Q9	-10914.89	4991.18	-2.187	0.03071 *		1.0432899	1.0415225
Q10	-10479.84	4990.81	-2.1	0.037857 *		1.0319428	1.0362937
Q11	1678.4	4990.6	0.336	0.737228		1.1174021	1.1239475
						1.105985	1.1082489

● 新加坡之 Dummy variable 係數與季節性指標

	Estimate(compae	Std. error	t value	Pr(> t)		CMV seasonal index	SLR seasonal index
Q1	-40575.281	1075.462	-37.73	<2e-16 ***		0.7143272	0.7226949
Q2	-43667.936	1075.551	-40.6	<2e-16 ***		0.6277514	0.6354943
Q3	-26455.467	1075.699	-24.59	<2e-16 ***		1.1584893	1.1619297
Q4	-31986.998	1075.906	-29.73	<2e-16 ***		0.9915091	0.9952488
Q5	-33913.41	1076.883	-31.49	<2e-16 ***		0.9445099	0.9364848
Q6	-31027.066	1076.499	-28.82	<2e-16 ***		1.0311287	1.0290398
Q7	-43006.597	1076.173	-39.96	<2e-16 ***		0.6414735	0.6496689
Q8	-43834.252	1075.906	-40.74	<2e-16 ***		0.6176663	0.62394
Q9	-39339.158	1075.699	-36.57	<2e-16 ***		0.7660893	0.7646471
Q10	-31266.689	1075.551	-29.07	<2e-16 ***		1.0243304	1.0183484
Q11	-16795.969	1075.462	-15.62	<2e-16 ***		1.4648784	1.4674456
						2.0178465	1.9950578

根據上表中六個國家地區的季節性指數和 Dummy variables 係數，我們發現結果並非完全符合我們的認知，例如我們原本以為六個國家地區的七、八月都會是來台旺季。以下我們將對於六國家地區的季節性指數進行分析，並猜測可能造成此現象的原因：

首先，我們可以發現每年一、二月普遍是來台旅遊的淡季。我們推測造成此現象的原因之一是因為寒假、春節假期是國人出國高峰期，在機位有限情況下，因而排擠到他國來台的機位，導致來台人次下滑。而歷年來一、二月來台旅客人次會受到春節連假有所變動，有「人次挪移」的觀光現象。接著我們個別分析各國家地區：

1. 日本

我們可以發現日本之 Dummy variables 係數高低趨勢與用 CMV 和 SLR 推算的季節性指標趨勢接近，並且根據 Dummy variables 係數，我們沒有足夠證據證明八月和十一月之來台人次與十二月有顯著差異。而根據日本來台人次的季節性指標，我們發現三、十一、十二月相對來說是日本人來台的旺季，其中三月

又最熱門，而四、六、七月則為淡季，其中又以七月最低。我們猜測，日本在三月有春假，造成出遊人次增加，四月上旬則常是各個單位的年度決算期，工作繁忙可能影響到外出旅遊，五月是日本黃金周，有計劃出遠門的家庭往往在五月出遊完後在後面月分減少出遊，加上台灣並非日本觀光首選以及六、七月是台灣赴日熱門的時段，可能會出現排擠現象，因此六、七月反而為日本來台淡季。

2. 香港、澳門

香港和澳門有著較接近的季節規律，一月都是季節性指標最低的月份，而來的熱門月份出現在六到八月以及十二月，其中八月和十二月為兩個高峰，我們推測假期可以解釋此現象：香港和澳門旅客常利用四月的復活節四天假期、六月到八月的暑假與十二月的耶誕假期到台灣旅遊。但是相較於香港，澳門各月份間的季節性指數落差較大。

3. 馬來西亞

馬來西亞各月份間的季節性指數落差較大，淡旺季明顯，其中十一、十二月季節性指數特別高，以十二月最高，一、六、七、八月較低，以七月最低，並不如我們所預想的暑假會是來台旺季。我們推測影響因素主要為氣候和假期。

氣候方面，馬來西亞屬於炎熱潮濕的赤道型氣候，受到季風性降雨影響，十月到次年二月是主要雨季，而三到五月則是最熱的時候。由於氣候影響，六、七、八月天氣相對乾燥而且不那麼炎熱，可以說是馬來西亞國內旅遊的黃金時期，因此造成這三個月的旅外人次減少，並且此現象反映在了六、七、八月明顯較低的季節性指標之上。

而十一、十二月季節性指數特別高，我們推測主要受到氣候與假期影響：十一、十二月馬來西亞正逢雨季，較少人會選擇留在國內旅遊，再加上馬來西亞學制為三個學期，十二月中之後學校就開始放假，又恰逢耶誕假期，正是出國旅遊的好時機，所以可以看到十二月的季節性指標特別的高。

4. 新加坡

我們可以發現新加坡之 *Dummy variables* 係數高低趨勢與用 CMV 和 SLR 推算的季節性指標趨勢接近，並且根據 *Dummy variables* 係數，我們有足夠證據證明一到十一月新加坡之來台人次都與十二月有顯著差異。而根據新加坡的季節性指標可以看出十二月份的來台人次為整年最高，十一月相對來說也算是比較多新加坡人來台的月份。一、二月和七、八、九月則為淡季，其中又以八月的季節性指標最低。

我們可以發現新加坡和馬來西亞的來台人次淡旺季大致接近，主要也是因為氣候與假期的因素。兩國在地理位置與氣候上相當接近，並且新加坡從 11 月中到 12 月底是為期六星期的學校假期，再加上十二月份的耶誕假期，可以解釋為何新加坡與馬來西亞的來台人次淡旺季接近，並且十二月為每年新加坡來台人數的最高峰。

5. 印度

印度各月的季節性指數差異並不大。一月季節性指數最低，十二月次低，我們推測或許也跟氣候有關：三月到六月是印度的夏季，而六月到十一月則為雨季，可能因此造成印度國人選擇利用十二月到二月進行國內旅遊，進而使旅外人次減少，並反映在季節性指標上。

而雖然印度來台比較熱門的月份出現在三到六月，也就是印度的夏季，並且在 3 月最高，但其季節指數都在 1.15 之下，代表僅有高於平均不到 15%，並沒有比較顯著集中來台的月份。

● 小結

以各別國家地區來看，日本三、十一、十二月相對來說是來台旺季，而四、六、七月則為淡季，我們推測主要跟假期有關。

香港與澳門的季節指標的起伏相當接近，一月都是季節性指標最低的月份，並且來台的熱門月份出現在六到八月以及十二月，主要原因推測為假期因素。而相較於香港，澳門各月份間的季節性指數落差更大。

馬來西亞與新加坡的來台人數走向相當接近，十一、十二為來台旺季，一月、六月到九月為來台淡季，推測主要是氣候與假期因素。

印度季節性指數在一月與十二月最低，推測和春節假期與氣候有關，但整體而言各月份間並沒有那麼大的來台人數差異。

總的來說，我們發現對於這六個國家地區，春節、該地氣候和假期大致上可以解釋它們來台人次的季節性指標與 *Dummy variables* 係數的變化。因此我們猜測這三項因素可能是影響這六個國家地區每月來台人次主要的原因。

七、結論與建議

(一) 結論

1. 根據趨勢圖，這六個國家地區的來台人次都存在長期正向趨勢，並且有季節性變化。

2. 根據這六個國家地區的時間序列模型，我們都有足夠證據證明時間對於來台人次是重要變數。
3. 對於這六個國家地區，春節、該國家地區的氣候和假期大致上可以解釋它們來台人次的季節性指標或是 **Dummy variables** 係數的變化，因此我們猜測這三項因素可能是影響它們每月來台人次主要的原因。

(二) 建議

1. 制定觀光政策

來台旅客的人次連年上升，同時帶來巨大的經濟效益。了解這六個國家地區旅客來台的淡旺季，依據差異制定不同的行銷策略吸引旅客來台觀光，以及分析預測未來的人流，有利加強台灣觀光。例如：針對日本旅客的喜好，推動鐵道及宣傳台灣特色景點與節慶活動；因應香港及新加坡自由行市場旅遊趨勢，深化自由行市場的宣傳，在當地投放廣告，或邀請當地名人赴台合作拍攝宣傳影片。分析結果也可以檢驗既有政策（例如：新南向政策、來台放寬簽證措施）的成效。

2. 旅遊業者推案參考

外國旅客來到台灣之後，實際的體驗更為關鍵。旅遊相關業者藉由來台人次的分析預測調整對應策略，例如旅館業者可以根據各月份的住宿需求合理的調整價格及人事、各類型活動及展覽可選在適合的時機舉辦以吸引更多人潮，既能提升自己的營運收益，也能維持觀光服務的品質。不僅在旺季時要有足夠的資源分配，在淡季時更應重視每個來訪的旅客，讓每位旅客不管何時來台都能留下美好的回憶，這才是推展台灣旅遊、提升來台人次最根本的作法。