

SQL Server® 2012  
商業智慧2.0

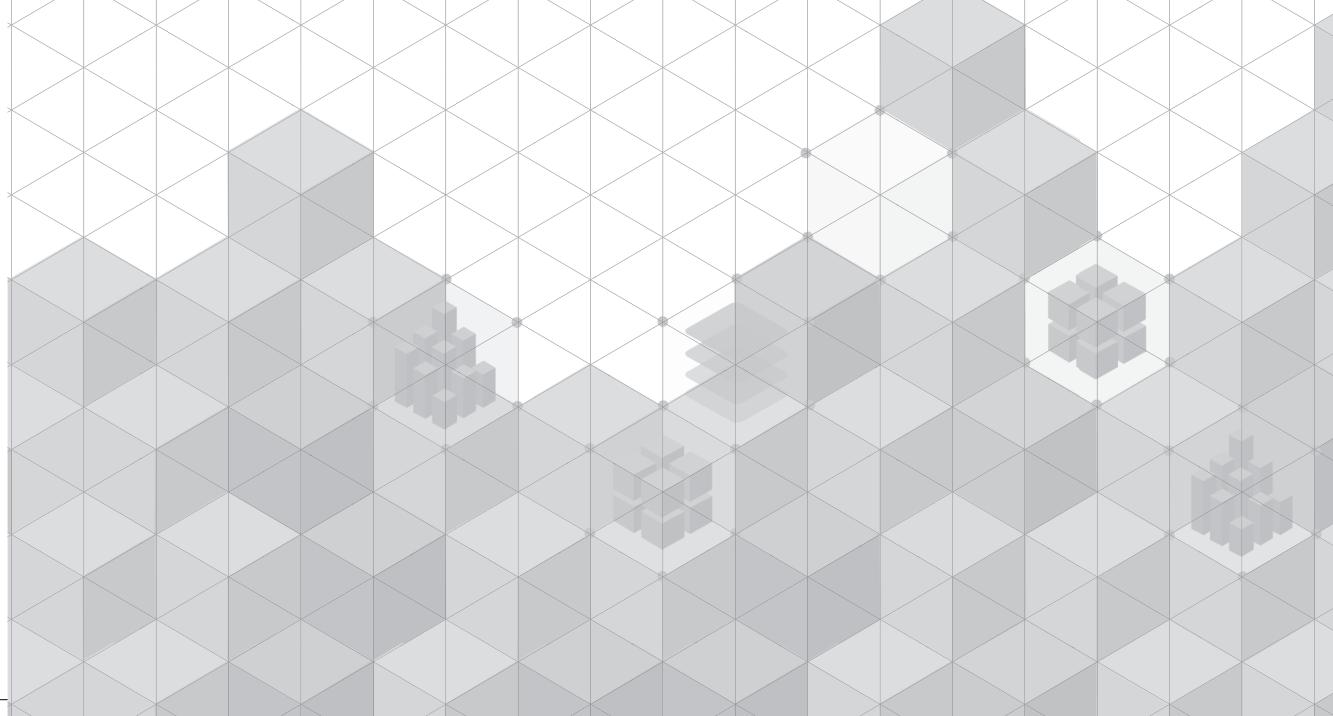
## 第15章地理資訊分析

地理資訊基本觀念

Reporting Services 地理資訊報表

表格式模型 地理資訊設計與應用

GeoFlow 地圖



隨著企業國際化的腳步以及行動應用的興起，地理資訊在商業智慧中所扮演的角色也越來越重要。相較於其他的商業智慧廠商，微軟很有可能是最後一個推出地理資訊分析的廠商。筆者在之前的書中就曾提到過，我認為這並非代表微軟在地理資訊商業智慧分析的漠視或是落後，而是微軟以更慎重的方式來推出地理資訊商業智慧分析。

事實證明筆者的推斷沒錯，自從Reporting Services 2008 R2首度推出地圖報表能後，在SQL Server 2012中，微軟更是發揮了它在地理資訊的強項，將資料庫、圖資、視覺化乃至於搜尋引擎的整合都發揮地淋漓盡致。接著，我們將以專章的方式來整合介紹微軟商業智慧在地理資訊上的發展。

全形

## 15-01 地理資訊基本觀念

### 15-1-1 地理資訊型別基本觀念

在SQL Server 2008之後，共提供了Geometry與Geography資料類型。Geometry(平面幾何)資料型別主要是以符合開放式地理空間協會（Open Geospatial Consortium，OGC）的SQL簡單特徵規格1.1.0版，來記錄地理資訊。至於Geography（球面測地）會儲存橢圓體（圓形地球）資料，例如，GPS 經緯度座標。這兩種資料型別有何差別呢？如果以台灣的大小來看，其實用哪一種並不會有太大的差異。因為就以台灣的大小而言，把地球看成平面或球面所得到的計算結果，是不會差異太大的。但如果是台北到紐約的距離，兩者計算的個別結果就會差很多，Geometry會是直穿地心，而Geography則是沿地球表面。

Geometry和Geography資料類型支援11種空間資料物件或執行個體類型。聽到這邊，各位可別被嚇到了，其實最常見的應用主要還是集中於3種物件：

- 點（Point）
- 線（LineString）
- 多邊形（Polygon）

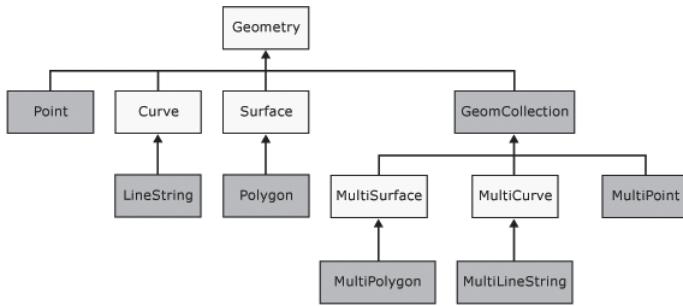


圖15-1：空間資料物件架構

「點」是最基本的地理資訊物件，它代表在空間中的一個位置。因此，無論是客戶的所在位置、分店的所在位置都可以利用「點」的方式來記錄。而地圖報表就可以將這些點標註在地圖上。

位置的描述當然是透過經緯度來代表。以亞洲資採辦公室所在位置來看，是落在東經121.57475 北緯25.035975。若要在SQL Server中建立一個可供分析計算的點物件（Geometry型別），請利用STGeomFromText函數（從文字轉換為平面幾何的地理資訊物件）。舉例來說，語法如下：

亞洲資採辦公室的所在位置：

```
geometry::STGeomFromText('POINT (121.57475 25.035975)',0)
```



圖15-2：點的查詢結果

在SQL Server SSMS上面執行此查詢語法後，可以看到編碼過後的地理資訊結果，此外，查詢結果還會多出一個「空間結果」分頁，以圖形化的方式來呈現查詢結果（不過一個點實在很難看分明，所以此分頁留待介紹線時，再做介紹）。

至於「線」則是由兩點構成一直線，在地理資訊的意義上表示兩個點間的最短距離。當然，也可以透過很多段長短方向不一的線構成折線、鋸齒狀或者是曲線。

其實語法與之前的點很接近，只不過用逗點分隔構成線的每個點的座標。以亞洲資採（121.57475 25.035975）到台灣微軟（121.568663 25.038862）一直到台北101（121.564085 25.034128）來說，可利用以下的SQL語法來構成這條路徑。

從亞洲資採到台灣微軟再到台北101：

```
geometry::STGeomFromText('LINESTRING (
121.57475 25.035975,
121.568663 25.038862,
121.564085 25.034128)', 0)
```

切換到空間結果分頁，在平面投影會畫出以下這條路徑。

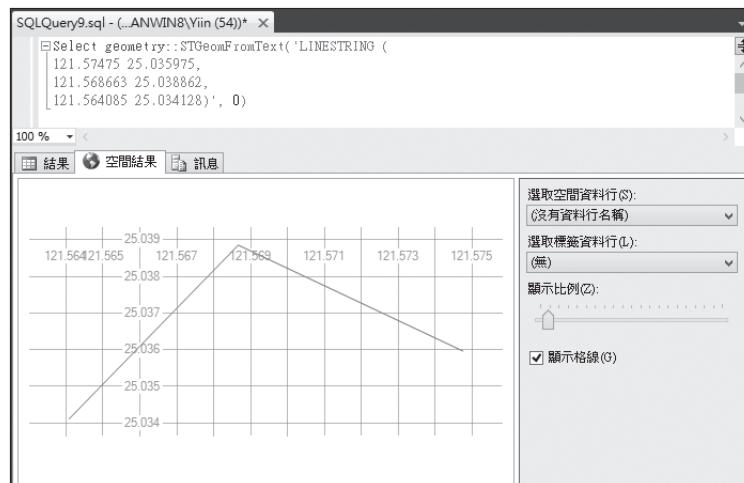


圖15-3：線的查詢結果

「多邊形」則是用來呈現在空間中不規則的任意色塊，包括了行政區、業務責任區等都可以透過多邊形的方式來呈現。在座標的表示上，多邊形本身是一個封閉

的曲線所構成，因此，它必須至少要有4個點。而且第一個點必須與最後一個點相同。

以下是假設某個公司在信義區的信義分店的業務責任區，就可以利用多個點的方式來構成責任區多邊形。當然，我們目前的範例點並沒有很多，因此這個責任區看起來會比較有稜有角。當點的數量越多，這個多邊形就會看起來越平滑。

信義分店業務責任區：

```
geometry::STPolyFromText('POLYGON ((  
121.559953 25.033174,  
121.564666 25.040993,  
121.576933 25.040639,  
121.57659 25.039295,  
121.576944 25.038761,  
121.577042 25.03849,  
121.577079 25.037894,  
121.577982 25.036067,  
121.5747 25.033588,  
121.572714 25.032729,  
121.559953 25.033174  
)',0)
```

查詢的結果如下，空間結果也會繪製出此空間的多邊形。

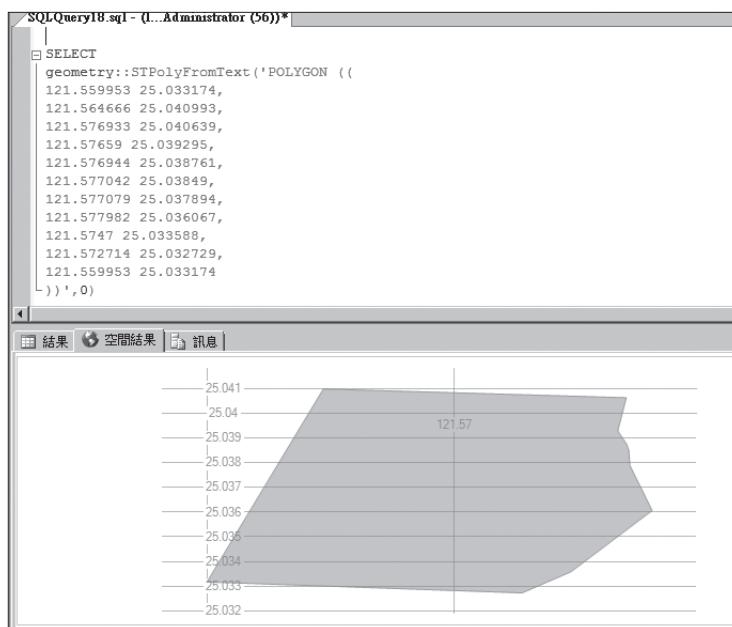


圖15-4：多邊形的查詢結果

透過空間物件可完整表達所有的空間觀念，並可以利用T-SQL的語法來表達。如圖15-5中，就透過了T-SQL語法再現世界名畫維納斯的誕生。各位可以從範例了解到，空間查詢不僅能用於地理資訊，還可以用來做任何物件形狀、位置的描述。

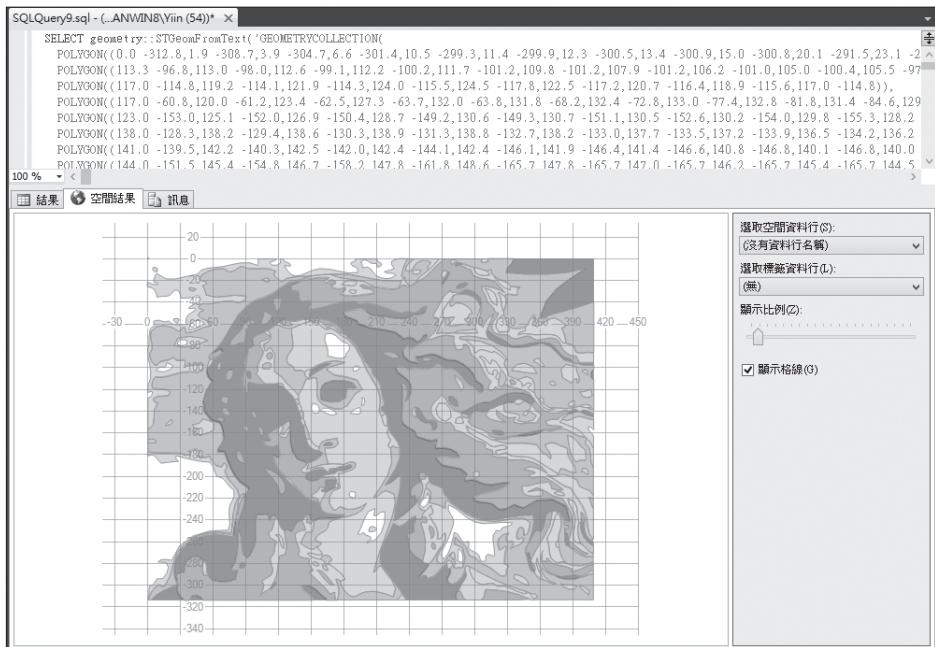


圖15-5：維納斯的誕生

有了空間物件之後，SQL Server其實還提供了一系列與空間有關的計算函數，可讓原先困難重重的地理資訊分析需求變得容易（例如，將篩出的名單按照最近的分店分配），而計算的結果自然還能夠透過商業智慧平台，來進行資料視覺化。

### 15-1-2 地理圖資的來源

當有了空間物件後，我們還需要將這些物件與真實世界的地理概念疊合，才能產生決策上的意義，而這些重現真實的資訊源稱之為地理圖資。目前常見的地理圖資來源包括了：

- ESRI形狀檔
- 網頁地圖圖資（Bing Map, Google Map…）

ESRI（Environmental Systems Research Institute, Inc.）是目前全球最大的地理資訊機構，目前Reporting Services支援了ESRI的地理資訊檔案規格。ESRI形狀檔是一組檔案，包含了：

- .shp 檔案：主要是記錄了地理資訊的點、線、多邊形資訊。
- .dbf 檔案：則是將以上地理資訊圖案連結至相關的欄位屬性。

在使用ESRI形狀檔時，這兩個檔案必須放在同一個資料夾中。在本書範例光碟中提供了台灣地圖的ESRI形狀檔。在設計地理資訊報表時，地圖報表會內嵌此ESRI形狀檔，以提供相關的圖資。

以外，SSRS中也提供了預設的地圖庫，不過，當然是以美國及世界地圖為主，因此，除了外商外，實用意義不大。另外，在<http://mapgallery.codeplex.com>中，也有人募集全世界國家的地理資訊圖庫以作為SSRS預設地圖庫，有興趣的人可以到上面尋找相關資訊。

至於網頁地圖圖資部份，我們介紹的當然不會是Google Map啦！微軟的地理資訊商業智慧預設使用的圖資自然是Bing Map。談到Bing Map，實在讓人對它又愛又恨，愛的是它的技術創新遠超過其他網路地圖，像是微軟特有的鳥瞰圖，如圖15-6，與Google透過實景車來拍攝街景不同，鳥瞰圖是透過衛星地圖加上3D塑模所構成的虛擬世界，因此，地圖中不會出現人類的身影，並能以任意角度與高度（這與街景車只能從地表的視角不同）切換視角，日後用於擴增實境等應用來說，實是非常強大的圖資。



圖15-6：Bing Map鳥瞰圖

至於恨呢？自然是中文化的問題。雖然到目前為止，Bing Map上面都還是英文圖資，所以在商業智慧分析時，都還只呈現英文的內容。不過，目前看來已有轉機，基於Windows Phone與Nokia的緊密合作，Nokia也宣布與Bing Map團隊互相合作，由Bing Map負責地理資訊應用技術研發，而由Nokia負責提供全世界本地化的圖資，所以眼尖的讀者可能會發現目前Windows 8及Windows Phone 8中的地圖，已經默默地變成中文了。因此，商業智慧可以使用中文圖資的那天，應該是指日可待了。

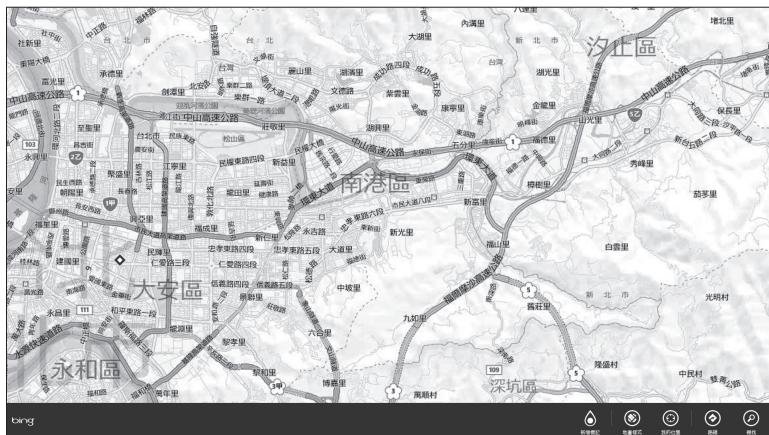


圖15-7：Windows 8地圖App

這樣與Nokia的分工合作果然讓Bing Map產生更多的創意火花，稍後我們將介紹表格式模型與Bing Map的整合，以及全新的Excel增益集GeoFlow，相信都會讓各位感受到其中的創新。

## 15-02 Reporting Services地理資訊報表

雖然本書定位為介紹新商業智慧技術，但由於涉及地理資訊的整合介紹，所以筆者還是留了些篇幅來回顧Reporting Services的地圖報表。

目前SSRS的圖層共分為以下4種：

- 圖格圖層（Tile Layer）：主要是用於呈現完整地理資訊圖資，目前僅支援Bing Map的圖資。



圖15-8：新增地圖圖層

請將選擇空間資料的來源切換為「ESRI形狀檔」，並選取加入光碟所附的台灣形狀檔，如圖15-9，將可預覽到形狀檔內所記錄的台灣地圖。



圖15-9：預覽形狀檔

點選「下一步」，就會進入到「選擇空間資料及地圖檢視選項」畫面。在此您可以利用左方的縮放卷軸進行地圖的縮放，同時利用位向指示器將地圖進行平移。利用這些功能便能進行地圖的裁切。由於使用ESRI形狀檔會將地理資訊內嵌於報表中，因而也可以在畫面右方指定地圖解析度。



圖15-10：裁切地圖



圖15-11：地段圖

除了地段圖之外，也可以選取「空照圖」，便會呈現衛星空照之結果。



圖15-12：空照圖

「下一步」

進入下一步，即可選取要設計的地圖圖表類型。總共包含了：

- **基本地圖**：以顏色深淺的方式呈現數據。
- **色彩分析地圖**：以不同色彩來呈現數據。
- **泡泡地圖**：多加了一個點圖層，可利用這些點的大小、色彩來呈現數據。



圖15-13：選取視覺效果

若選擇的是「色彩分析地圖」，接著則可選取包含地理資訊的資料集。在此範例提供的資料集中，並未包含之前提到的Geography或者是Geometry資料型別，而是提供一個欄位與ESRI形狀檔內的地理資訊欄位屬性的內容相符。



圖15-14：選取資料集



圖15-15：資料集比對

最後，則是設定視覺化選項。在「要視覺化的欄位」中選擇色彩變化的基礎欄位，在此選擇了「SALES\_AMT」欄位，同時也可在此選取預設的色彩規則。若希望能標示出每個行政區的名稱，則需要勾選「顯示標籤」，並於下拉式選單中選取代表縣市別名稱的欄位。在此是選取「CITY\_NAME」，如此即完成基礎的設定。



圖15-16：設定視覺化選項

若要做進一步的設定，可以點選地圖，則右方會顯示目前的地圖僅有單層的多邊形圖層，點選右鍵，即可選取其他的設定選項。

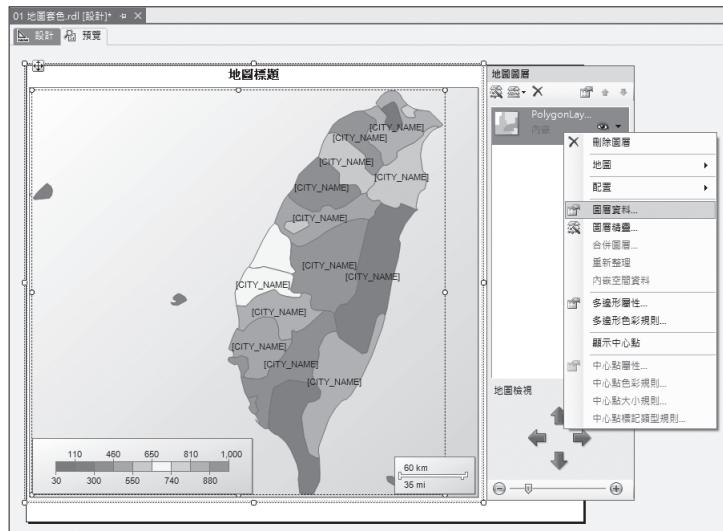


圖15-17：設定圖層

09  
10  
11  
12  
13  
14

15-02

Reporting Services地理資訊報表

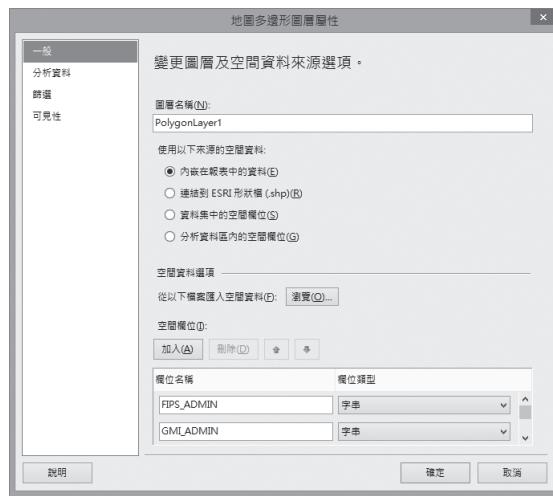


圖15-18：設定圖層

至於點選「多邊形色彩規則」，在此可用來指定每個多邊形的色彩呈現原則，如圖15-19。



圖15-19：多邊形色彩規則一般頁籤

至於在多邊形色彩規則散發頁籤中（散發是個翻譯錯誤，在此散發（distribution）指的是分布（distribution）），可用來指定「子範圍的數目」，子範圍的數目越多，則畫面的漸層效果越好。筆者在此將子範圍的數目數量修改為10。



圖15-20：多邊形色彩規則散發頁籤

完成所有的設定後，預覽的地圖就會如圖15-21所示。

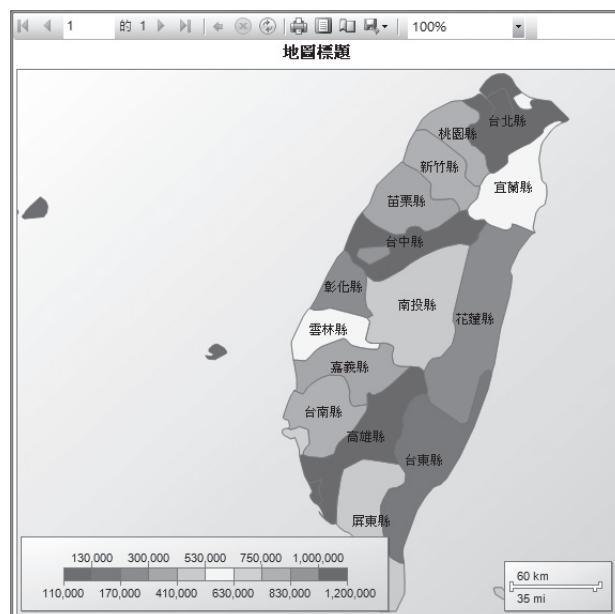


圖15-21：最終設定結果

09  
10  
11  
12  
13  
14

15-02

Reporting Services地理資訊報表



圖15-22：最終設定結果

若要改變泡泡的顏色，就要利用「中心點色彩規則」來指定。在此將泡泡的色彩連結至毛利率欄位（`MARGIN_RATE`）。

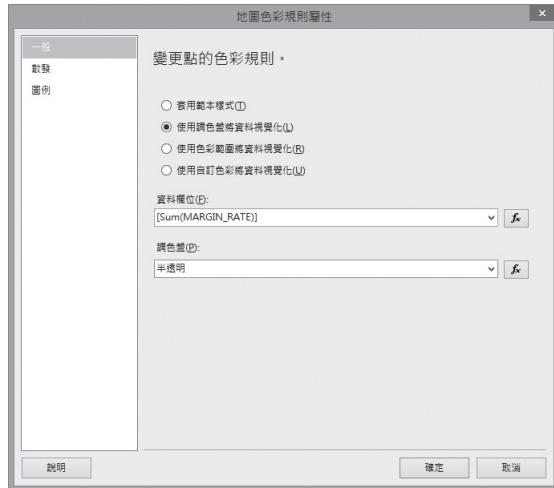


圖15-23：設定泡泡色彩

最後，由於形狀檔圖層系統預設是以「隨機」套色，會讓畫面看起來花花綠綠的不容易閱讀，因此，我們於「多邊形色彩規則」將變更多邊形色彩規則，切換為「套用範本樣式」。設定完成後，預覽報表結果如圖15-24。

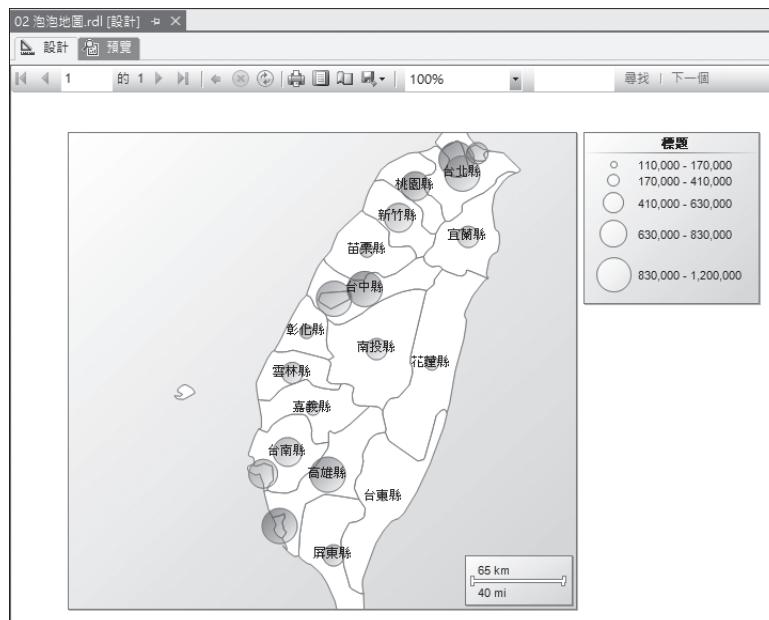


圖15-24：泡泡地圖結果

## 15-2-2 使用SQL Server地理資訊資料設計地圖報表

除了使用ESRI形狀檔外，接下來，我們將介紹如何利用SQL Server空間資料型別來產生更彈性的地理資訊商業智慧分析。在以下的範例中，我加入了一些地理資訊資料型別運算的範例。

台北各分店銷售狀況：

```
Select
    [StoreName]
    ,Geometry::STGeomFromText('POINT ('+Cast([STX] as varchar)+' '+Cast([STY] as
Varchar)+')',0) as Location
    ,TotalSales
    ,MarginRate
    from [ASIAMINER_BI20].[dbo].[Stores] a
    Left Join
    (
        Select [StoreId]
        ,Sum([SalesAmount]) as TotalSales
        ,(Sum([SalesAmount])-Sum([ProductCost]))/Cast(Sum([SalesAmount])) as
        Numeric(18,2)) as MarginRate
        from [ASIAMINER_BI20].[dbo].[TrxDetails]
        Group by [StoreId]
    ) b
    on a.[StoreId]=b.[StoreId]
    Where [AdministrativeAreaName]='台北市'
```

以上的語法其實就是計算各分店的銷售金額與毛利，同時串接至分店主檔，將該資料表中的經緯度轉換為Geomerty型別。接著，我們將以上語法置入Reporting Services中，以建立共用資料集。

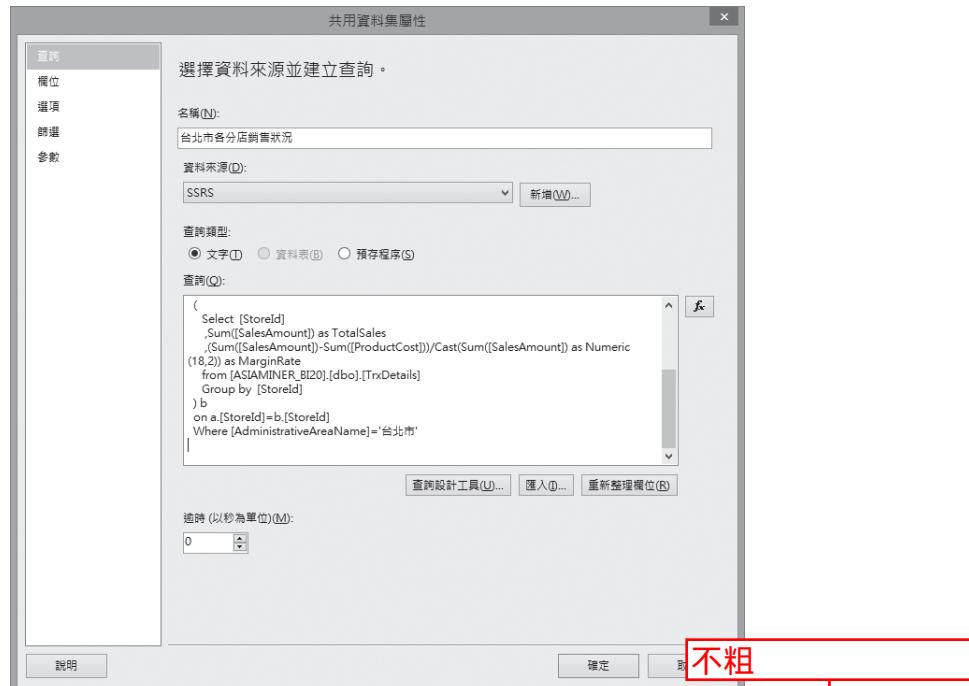


圖15-25：建立包含空間資料之資料集

接著新增報表，將剛才的共用資料集加入報表中，並拖放一個新的地圖物件，再將「選取空間資料的來源」切換為「SQL Server空間查詢」（請參考《圖15-8：新增地圖圖層》）。如圖15-26，請選取剛才建立包含空間查詢的資料集。



圖15-26：使用包含空間資料之資料集



圖15-27：設定空間資料

在接下來的設定畫面中，可以進一步的設定泡泡圖的呈現邏輯。由於我們的空間查詢中包含銷售金額（泡泡大小）、毛利率（泡泡顏色）以及分店名稱（泡泡標籤），各位可參考圖15-28的方式來進行設定。



圖15-28：設定泡泡大小與色彩

完成設定後，點選預覽報表，即可得到如圖15-29透過泡泡圖的方式來顯示各分店位置，以及檢視其銷售金額的相對大小。若是您對於泡泡大小或顏色有任何想法，都可以點選「Point圖層」進行細部設定修正。

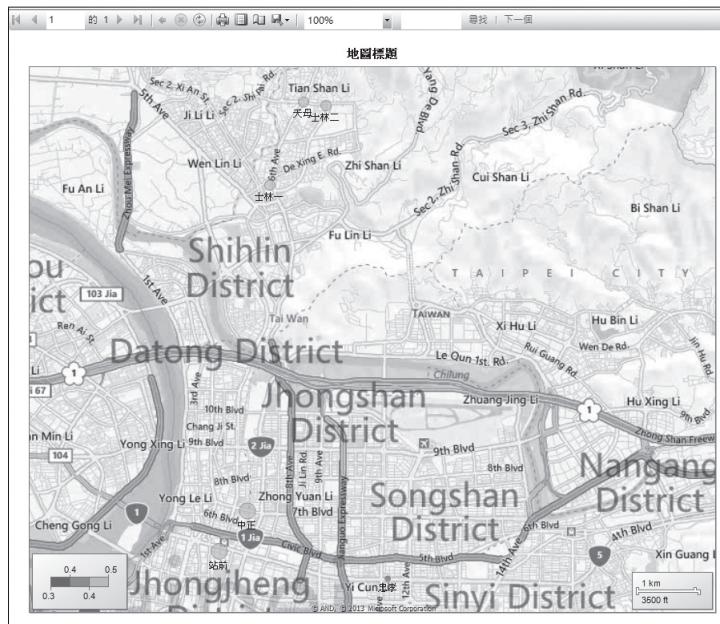


圖15-29：檢視泡泡地圖

## 15-03 表格式模型地理資訊設計與應用

### 15-3-1 地理資訊標記

大多數的商業智慧平台中的地理資訊分析功能都是透過之前介紹過的形狀檔或經緯度的方式來實作。但對於絕大部分的企業來說，無論是哪一種，進入門檻都過高，因此也造成地理資訊分析不夠普及的問題。

為了能夠讓企業更廣泛地應用地理資訊分析，微軟在SQL Server 2012 SP1的表格式模型以及Excel 2013版的PowerPivot中，新增加了資料類別的概念，用來進行特殊邏輯資料的標示。而這些標示可讓前端視覺化工具執行特殊處理與應用。而其中地理資訊就是資料類別中非常重要的一部分。

09  
10  
11  
12  
13  
14

15-03

表  
格  
式  
模  
型  
地  
理  
資  
訊  
設  
計  
與  
應  
用



圖15-30：在表格式模型中設定資料類別

不過要注意的是，由於「資料類別」是SQL Server 2012 SP1後才有的功能，因此設定時，必須確認模型檔（Model.bim）的相容層級是否已切換至「SQL Server 2012 SP1(1103)」，否則是不會出現此選項。



圖15-31：專案相容層級

在Excel 2013 PowerPivot（舊版是沒有的）中，同樣提供了資料類別的功能，只需要點選欲指定之資料行，切換至「進階」索引標籤即可進行設定。



圖15-32：在PowerPivot中設定資料類別

根據筆者實際操作的結果發現，PowerPivot相較於表格式模型，多了「自動辨識」的功能。也就是說，當筆者把分店別資料匯入模型時，表格式模型必須自己手動設定各資料行的資料類別，但PowerPivot一部份的資料行卻能夠自動辨識資料類別。經筆者深入追查後發現，PowerPivot設計工具會自動透過Windows Azure平台的推薦機制（<https://recommendation.azure.com/rest>）來執行自動辨識。有了自動辨識機制，對於一般非專業資訊人員來說非常便利，也可省下不少設定資料模型的時間。

目前與地理資訊相關的資料類別如下：

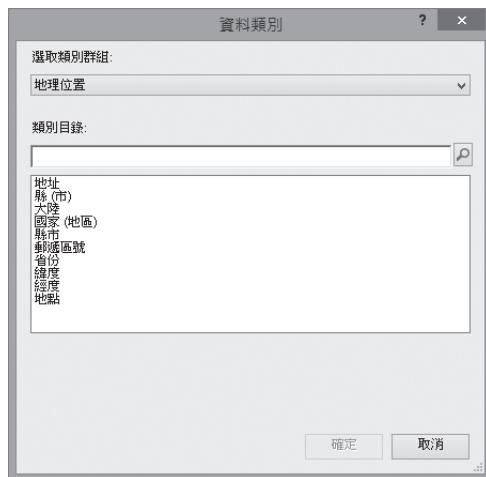


圖15-33：地理資訊相關資料類別

地理資訊資料類別本身只是一種標記性質，要讓資料類別發生作用，必須由視覺化工具來達成。以目前來說，僅有Power View才支援資料類別的效果。在執行地圖報表展示時，Power View會將有標示為地理資訊之欄位去重複值送至VirtualEarth（Bing Map的地圖引擎）進行地理位置查詢，然後，Power View就會自動將資料標記於回傳之地理位置之上。經筆者實際測試，雖然Bing Map地圖目前還是英文，但Power View地理資訊資料行是支援中文內容查詢的。目前各項地理資訊資料類別之定義如下：

- 地址：指向該地址之精確位置，例如：台北市松德路171號。
- 縣(市)：指向縣市或直轄市之中心位置，例如：台北市。
- 大陸：指向該大陸（洲）之中心位置，例如：亞洲。但筆者測試過，如果您希望指向美洲，中文名稱務必標明北美洲或是南美洲，否則會無法判定。
- 國家（地區）：指向該國家地區之中心位置，例如：台灣。
- 縣市：指向行政區或都會之中心位置，例如：板橋區。
- 郵遞區號：指向該由地地區號之中心位置，目前僅支援美國郵遞區號。
- 省分：指向該省份之中心位置，例如：加州（沒錯，用中文也能指向California）。
- 緯度：套用此資料類型，此欄位值會被視為緯度。
- 經度：套用此資料類型，此欄位值會被視為經度。
- 地點：指向特殊景點或知名地標之中心位置，例如：台北101。

稍後，我們將介紹透過Power View，來展示這些地理資訊資料類別之效果。

### 15-3-2 Power View地理資訊報表

在SQL Server 2012 SP1表格式模型及Excel 2013 PowerPivot，正式加入了資料類別的觀念。在對應版本的SQL Server 2012 SP1 Power View以及Excel 2013 Power View中，也針對這些資料類別做了對應的視覺化展示。其中，針對地理資訊類別資料行，就能夠自動以地圖的方式呈現。在Power View欄位清單中，若是資料模型標示為地理資訊類別的資料行，就會顯示地球的小圖示，見圖15-34。

如



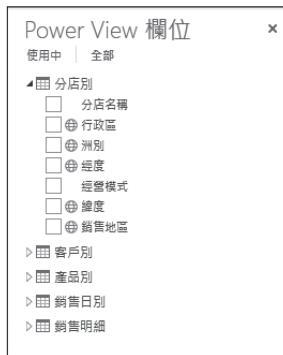


圖15-34：Power View欄位清單中的地理資訊資料行

與其他廠牌商業智慧工具中的地理資訊分析最大的差別在於，企業無需自行準備圖資或自行查詢位置之經緯度。當透過Power View的地圖控制項來展示地理資訊資料行時，Power View會自動將資料行內的值送至VirtualEarth查詢，並自動將資料標示在回傳的地理位置之上。舉例來說，被筆者標示為「大陸」的「洲別（資料庫中為RegionGroupName）」資料行，Power View就會自動將裡面的值「亞洲」，透過以下Url查詢其中心位置，其中「includeEntityTypes=Continent」表示該查詢資料類別為「大陸」，至於「亞洲」這個查詢關鍵字，則被UrlEncode後，會變成%e4%ba%9e%e6%b4%b2。

Power View自動查詢「亞洲」的中心位置

```
https://dev.virtualearth.net/REST/v1/Locations?key=Am9jDNwXCG3tUgleAtYKexNZKrTTzKgSiNPIgUP0YnWC1Tr5n1iwkvde4MHwGSBZ&c=zh-TW&includeEntityTypes=Continent&maxRes=20&q=%e4%ba%9e%e6%b4%b2
```

舉例來說，當我們透過Power View欄位清單勾選分店別的「銷售地區」以及量值「總銷售金額」後，原本是以表格的方式呈現，如圖15-35左側。但當我們切換此資料表為地圖後，除了會以地圖方式呈現外，Power View還會自動將各個國名換成對應的地理中心位置進行標示，如此一來，就可以很簡單地做到泡泡地圖了。比起之前我們用Reporting Services設計的過程相較之下，是否顯得簡單許多？

09  
10  
11  
12  
13  
14

15-03

表  
格  
式  
模  
型  
地  
理  
資  
訊  
設  
計  
與  
應  
用



圖15-35：Power View泡泡地圖

此外，我們還可以將其它的類別資料行加入地圖的「色彩」區域，這樣，就會顯示如圖15-36的圓餅泡泡地圖。



圖15-36：Power View圓餅泡泡地圖

在Power View中，也允許切換為不同的底圖，以便利分析者展示其資料。

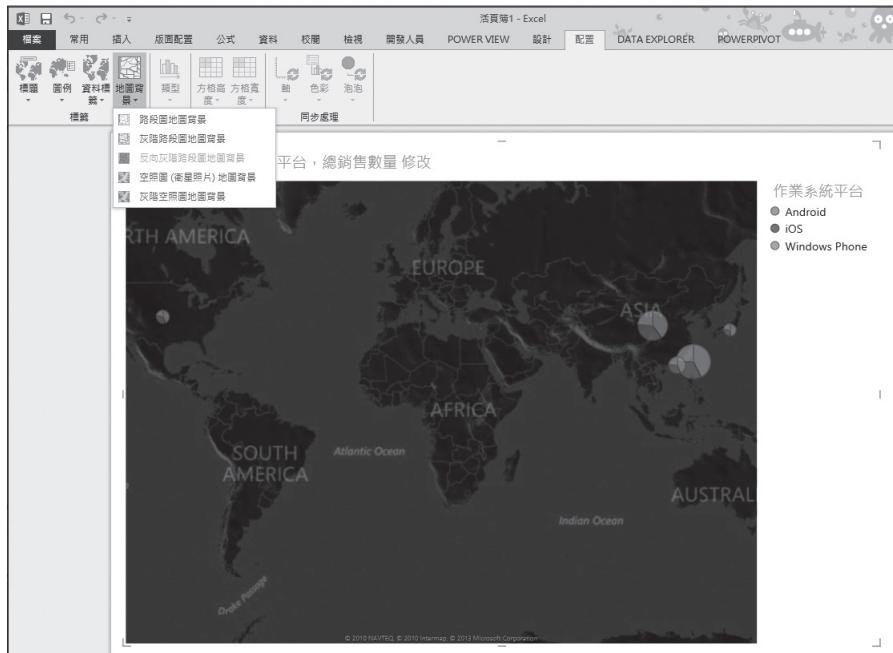


圖15-37：切換底圖

或者您也可以如圖15-38，搭配交叉分析篩選器。如此一來，就可以點選到指定地區，以顯示對應地區之細部銷售狀況分布。



圖15-38：地圖搭配交叉分析篩選器

## 15-04 GeoFlow地圖

從Power View的地圖功能會發現，微軟其實無縫整合了商業智慧表格式模型、Bing Map圖資以及其搜尋功能，以發展出更多地理分析的可能性。事實上，微軟目前針對地理資訊還有尚未發表的秘密武器「GeoFlow」，更是將地理資訊的能力發揮到極致。

GeoFlow目前還是Beta 1（尚未公開技術預覽），由於正式版操作介面可能仍會有大變動，因此，在後續的篇幅中，將不深入介紹操作介面，主要希望能讓各位讀者先一睹為快。GeoFlow是基於Excel 2013的增益集，其技術架構與Power View類似，都是基於PowerPivot資料模型的資料視覺化工具。當設計完資料模型後，可以直接點選「插入」索引標籤中的「3D Map」啟動Geo Flow地圖。

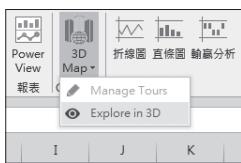


圖15-39：啟動GeoFlow

筆者在此使用的是，我們公司協助銀行進行不動產整批評價的資料庫來做範例，首先，映入眼簾的就是一個立體的地球，這也是為何稱為3D Map的緣故。

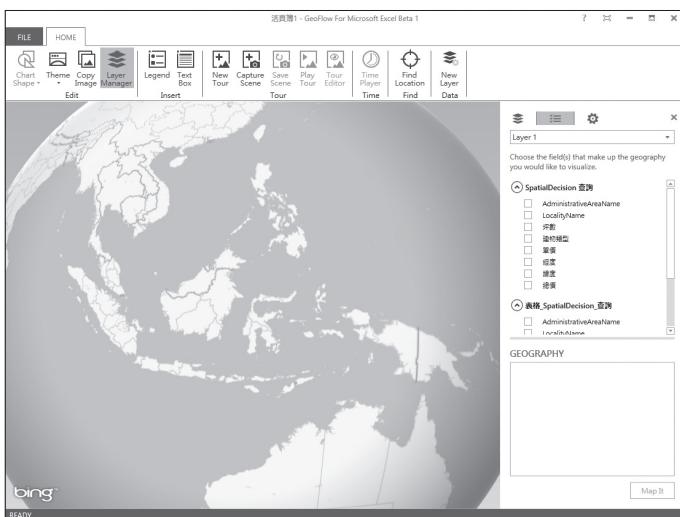


圖15-40：GeoFlow設計畫面

GeoFlow是以圖層為基礎的視覺化工具，目前提供以下數種圖層：

- 地圖底圖：每個報表可以選定一個地圖底圖，如圖15-41。
- 柱狀圖：數值以立體柱狀圖（或堆疊柱狀圖）的方式標註在地理位置中心。
- 圓餅圖：數值以圓餅圖的方式標註在地理位置中心。
- 熱溫圖：當資料位置為零散區域時，它會自動根據數值計算類似溫度分布的熱溫圖。

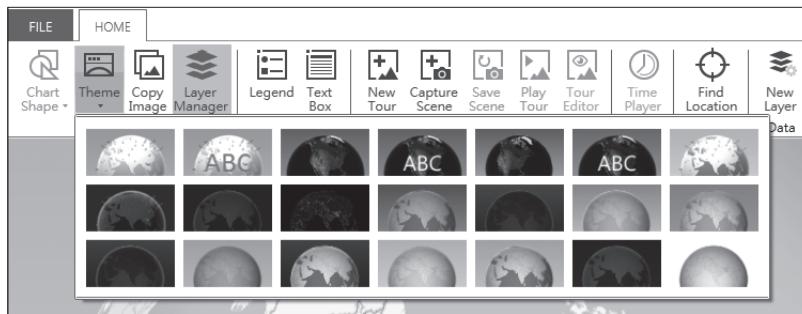


圖15-41：地圖底圖

GeoFlow設計地理圖層的概念與Power View很類似，您只需要指定地理資訊資料行即可完成設定。例如，筆者先指定「LocalityName(行政區)」資料行，並將其資料類別指定為「City」後，再執行「Map It」，此時，GeoFlow就會開始協助我們自動標示位置了。

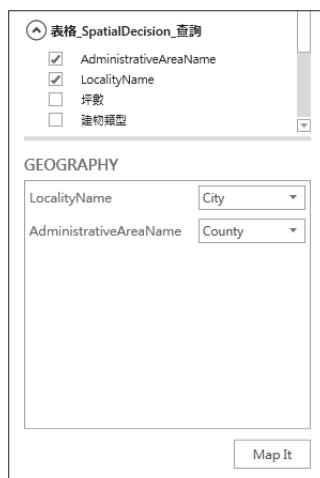


圖15-42：指定資料行之資料類別

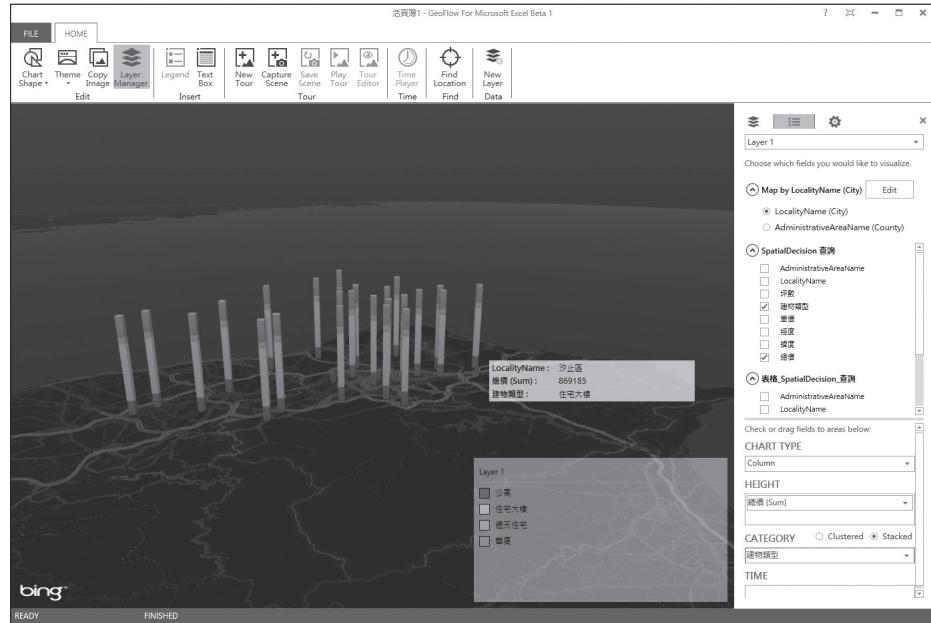


圖15-43：顯示3D柱狀圖

由於筆者使用的資料庫有每一筆成交行情詳細經緯度，因此，也可以加入新的圖層使用經緯度來做為定位基準。接著，將塗層設定為熱溫圖（Heat），以房屋平均單價做為顯示量值（Value），圖形就會顯示如圖15-44這般在既有地圖上又疊加一層熱溫圖。透過色溫的變化，可以很清楚地看到那些是高房價的熱區（紅色），這是非常有用的視覺化技術。

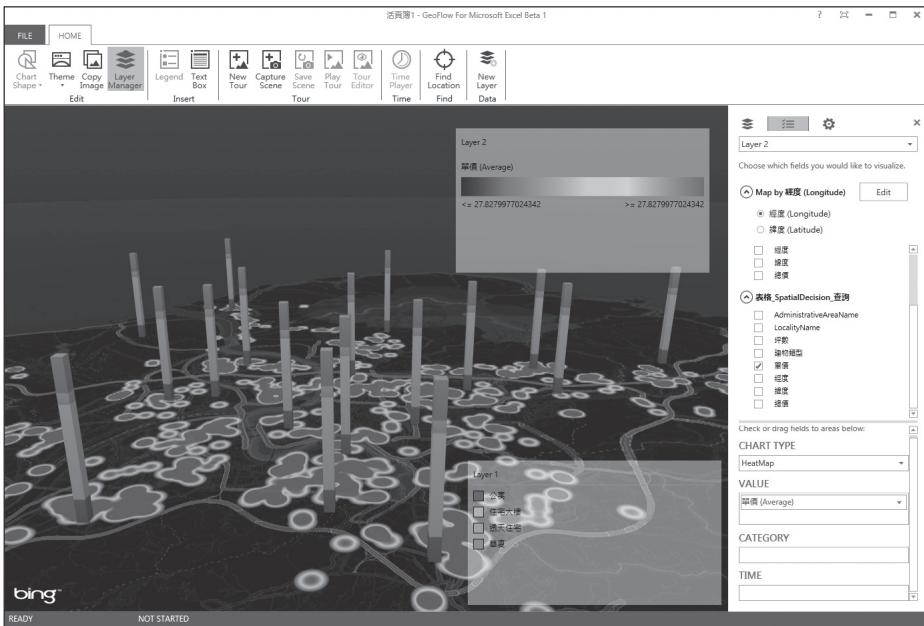


圖15-44：顯示熱溫圖

從以上的幾個展示，就能看到微軟在地理資訊商業智慧分析上的企圖心以及技術上的創新。將地圖與搜尋技術整合的觀念非常新穎，GeoFlow可說開創了一條地理資訊分析的新方向。至於GeoFlow未來會推出的時間點，微軟並沒有明確的時間表，但筆者認為，它無論是背後的資料模型以及目前的操作介面，幾乎都與Power View如出一轍，因此，未來整併至Power View的可能性應該不小，就讓我們靜觀其變吧！◆◆