

出國報告（出國類別：研究）

# 金融業客戶行為分析 之研究

服務機關：臺灣土地銀行

姓名職稱：鄭欣如高級辦事員

派赴國家：美國

出國期間：105年9月10日至105年9月30日

報告日期：105年12月15日

# 摘 要

巨量資料已是近年來最火紅技術議題，分散式的計算架構、非結構化資料分析…等，彈性及即時的特點使握有資料的一方能快速掌握即時回應，而社群、物聯網等新興議題與巨量資料分析應用的交互影響下，金融服務不再是銀行的獨門生意，金融業生態正在面臨巨大的轉變，金融業者需要更多對於自家客戶、整個市場有更多的了解，發現需求並提供服務；如何運用分析去了解客戶的行為、正確的解讀其中的意義，進而回應，或是提供有需求的商品，都考驗著各家利用資料分析的能力。

本次赴美國從事「金融業客戶行為分析研究」，藉由參訪美國 **Teradata**、**UNISYS** 及 **Microsoft** 公司等知名企業所提供之簡報及研討會，瞭解國外各產業於巨量資料之應用及下一階段趨勢，作為本行發展與應用客戶資料及行為分析應用之參考。

# 目 錄

壹、	研究目的 .....	3
貳、	研習內容 .....	4
一、	Teradata Partners 2016.....	4
二、	UNISYS (優利) .....	5
三、	Strata + Hadoop World 2016.....	5
參、	研習心得 .....	7
一、	客戶行為研究的起點--“資料” .....	7
(一)	資料形式&特性.....	7
(二)	金融業客戶資料蒐集的範圍.....	7
二、	資料分析方式及運用 .....	9
(一)	分析方法 .....	9
(二)	資料的運用.....	12
三、	分析平台的建立 .....	17
(一)	分析的再進化 .....	17
(二)	奠定分析基礎—制定計劃 .....	18
(三)	建立分析平台 .....	19
(四)	金融業分析應用案例.....	21
四、	隱私權問題 .....	23
肆、	本行現況及建議事項 .....	26
一、	結論.....	26
(一)	先有研究/分析的目的，如何研究、使用技術是其次 .....	26
(二)	資料品質及時效性依然重要 .....	26
(三)	許多界線逐漸模糊 .....	26
(四)	分析文化的建立與態度影響成效.....	26
二、	本行現況 .....	27
三、	建議.....	27
(一)	資料架構的整理與單一客戶視圖的建立 .....	27
(二)	探索平台的建立.....	28
(三)	分析文化的建立，包含團隊、及資料使用的推廣.....	29
伍、	參考文獻 .....	30

## 壹、 研究目的

金融業的起點--銀行業是一項古老的行業，起源可追溯至 600 年前義大利文藝復興時期，其經營卻在這幾年裡因著行動網路、數位行動裝置、巨量資料處理分析平台、雲端、互聯網、物聯網等技術，面臨新的機會與挑戰，提出銀行 Bank 1.0 到 Bank3.0 進化論的 Brett King 說：銀行不再是地方，而是一個行為（Why Banking Is No Longer Somewhere You Go, But Something You Do.）。

金融業者不再是銷售產品，而是需要知道客戶真正的需求是什麼，以及能怎麼滿足其需求，巨量資料（或稱大數據）的處理技術的出現，使得過去礙於技術限制的資料處理能力，有了過去令人難以想像的突破，企業能夠更完整的去瞭解客戶的面貌，尤其以目前的消費、生活型態而言，客戶對於需求的探索/實現多是從網路上搜尋資訊開始，至分行多僅是下決定後最終的一個步驟，因此我們有必要瞭解從客戶的行為模式、取得服務的完整路徑，也就是從有興趣到最後決定的互動過程，而不只是測量最後一個步驟（分行、交易）。

本次研習報告目的在探討與客戶行為有關的資料，像是客戶個人資料、交易資料、通路行為等，歸納出目前國內外專業機構團體或知名企業提出的資料分析及應用的方式和觀點。

## 貳、 研習內容

本次赴美從事「金融業客戶行為分析之研究」，研習內容為參訪 Unisys 公司，以及參加由 Teradata 公司舉辦之 Teradata Partners 2016、Microsoft 舉辦之 Strata + Hadoop World 2016 等巨量資料分析及應用研討會，議程涵蓋了「資料分析」、「物聯網應用」、「用戶體驗與視覺化應用」、「巨量資料分析環境/架構」等。

### 一、 Teradata Partners 2016

Teradata(天睿) 公司以大數據分析、企業資料倉儲和整合行銷管理解決方案為主要業務，為目前市面上提供大型資料庫解決方案的龍頭廠商之一，總公司位於美國，在全世界 24 個國家都有分支據點，客戶則遍及全球 77 個國家。

天睿公司最為知名的大型資料倉儲產品使用的技術被稱作「無共享 (shared nothing)」架構，各個伺服器之間擁有獨立記憶體和處理能力，增加伺服器與節點即可擴充資料儲存空間，並由資料庫軟體集中管理承載負荷量。天睿公司也銷售可處理不同資料類型的應用服務與軟體，並於 2010 年加入文字分析功能，可用於追蹤非結構化資料（如文字檔案）或半結構性資料（如試算表），應用於商業分析，本行目前使用的資料倉儲系統係為天睿公司生產並負責系統維護。

Teradata PARTNERS 則是每年一度由 Teradata 籌辦，邀請許多知名企業的研究及應用實務經驗分享，包含了亞馬遜(Amazon)、富國銀行(Wells Fargo)、Ebay 等，分享內容包含如何運用資料處理企業的難題，或是如何在企業中構建資料分析的環境，主題皆著重於資料處理分析的領域。今年的大會主題為” Data Changes Everything”，十分呼應近一年來隨著數據處理平台、夾雜著雲端服務、物聯網從實驗階段進入成熟階段的發展，如何處理資料、運用資料的能力都是各企業努力加強的議題。

## 二、 UNISYS (優利)

Unisys (優利) 總公司位於美國賓州，是一家全球性的資訊技術服務和解決方案公司，將諮詢、系統整合、委外服務、基礎建設服務和銷售伺服器集於一身，客戶遍及全球 100 多個國家，當中包含政府單位、金融、航太製造、服務業、交通運輸、教育機構等各領域。

優利公司除了生產大型主機、票據分類機等常見的硬體產品之外，也提供業務分析規劃、多種系統網路應用整合、網路安全解決方案、生物辨識系統、委外服務及教育訓練等服務。至今已發展成為一家橫跨美、歐、亞、非四大洲的跨國企業。台灣優利系統股份有限公司在台灣提供資訊服務已有多年的歷史，其主要客戶包括：台灣銀行、土地銀行、彰化銀行、遠東銀行、福特六和汽車等。本行目前使用的中心作業主機係為優利公司所生產，並負責系統維護事宜。

本次參訪行程包含金融領域的全通路解決方案、生物辨識應用、應用服務等，以下為參訪議程：

- Trends in banking and FinTech and Unisys Point of View
- Omnichannel and digital banking
- Advanced Analytics
- Biometrics in banking
- Brief tour of Facility
- Application Services

## 三、 Strata + Hadoop World 2016

由 Microsoft(微軟)協力承辦的「Strata + Hadoop World」是目前有關巨量資料領域裡最大型的研討會之一，集結了各項有關巨量資料技術分析及處理的原創者及企業領域的領導使用者，研討會中有許多特別的議題或應用領域經驗分享，包含巴拿馬文件的解密、F1 賽車上的運用及白宮裡有關資料科學領域的分享

等，可以看到資料分析及運用已影響並改變所有領域的工作。

在今年度的研討會上，可以看到資料分析與物聯網的結合應用、分析平台的架構、人工智慧、企業架構資料分析平台等都是熱門議題，而從今年大會議程安排上可觀察到「**Streaming data**」將會是未來接下來的熱門議題。

另外也在許多議題研討當中可看出巨量資料的應用在每個產業都有更進階的應用，除了金融業外，也有醫療、娛樂、網路應用服務...等，這些被邀請來的講者幾乎都是產業的前幾大甚至是獨佔市場的公司，除去這些公司產業的分類，在本質上其實都等同於一個大型的資料收集處理、分析應用公司，資料運用能力的重要性已隨著「資料分析」由過去的支援性質角色轉變為支撐公司營運的角色，成為影響公司營運的重要因素。

## 參、 研習心得

### 一、 客戶行為研究的起點--“資料”

#### (一) 資料形式&特性

##### 結構化資料 vs. 非結構化資料

在這個行動裝置普及的數位時代，金融行業進行數據分析時，不能再只注重傳統在進行交易時所獲得的結構化資料，應以包含結構化與非結構化的資料做分析應用的基礎，以下簡單說明何謂結構化與非結構化資料：

##### 結構化資料

將資料擺放整齊，在儲存時就已受到了精確定義，且拒絕例外。特色是清晰、明瞭、且容易被預期，也意味著資料能輕鬆就被後續程式取用分析，也很方便被整合到某些特定的框架之內。

##### 非結構化資料

簡單的說法是，無明確結構、讓人無從整理起的資料。如一支影片、一串缺乏欄位概念的純文字等都是，在形態上無法簡單整理存放的資料。

傳統的結構化資料因對格式的要求而無法納入某些不完備的資料，結構化的過程中會因清理資料而丟失以及拒絕了某些資料。就像是為了把家裡整理得乾乾淨淨，總是難免會要丟掉一些東西，是易用性和全面性之間的取捨問題。

#### (二) 金融業客戶資料蒐集的範圍

以目前台灣的市場現況來說，絕大多數的客戶都同時與多家金融機構往來，銀行若只靠著蒐集交易資料及合法外購的資料，對建立完整的客戶視圖已不足



夠，必須藉著擴增更多與客戶互動時產生的資料及一些外部公開資料來輔助分析的進行。

金融機構目前可取得的與客戶互動相關的資料，大致上可分為三類：1. 客戶互動資料，包含客服中心與客戶的互動紀錄及員工在各種接觸客戶過程中留下的互動紀錄或軌跡 2. 數位通路上客戶的各項互動軌跡，如網站及行動手機 app 的使用行為 3. 銀行本身所能掌握的社群媒體討論資訊，如自營的討論區或臉書粉絲頁等，代表了客戶意圖的表達。

歸納這些資料的內容，包含：

1. 內部業務資料：金融機構內部各種結構化的資料，包含客戶資料及交易資料等。
2. 外部/外購資料：如政府資料開放平台、合法取得的加值資料供應商所提供的資料等。
3. 互動記錄：整合金融機構所能掌握的客戶互動資料，多以非結構化文字型態存在，例如客服中心的客戶專員與客戶的互動紀錄、理財專員與客戶的互動記錄、或是逐漸普及的線上文字客服互動內容等。
4. 網頁點擊資訊：客戶在金融機構網頁的點擊資訊，往往代表其當下意圖，若能蒐集網頁使用行為，能提升對客戶的瞭解程度。
5. 社群媒體資訊：客戶在社群媒體的發言，可蒐集以作為商品的滿意度、話題性熱度或商機的產生。

## 二、 資料分析方式及運用

### (一) 分析方法

在此說明幾項巨量資料分析處理上常常談論到的幾項主要技術：

#### 1. 機器學習

機器學習為人工智慧研究主題之一，是為了以電腦實現人類自然進行學習能力的技術與手法。先集結了一定數量的抽樣資料為對象並進行分析，再從這些資料萃取出有用的法則、知識表達與判斷標準。

機器學習可以從資料或經驗當中，萃取出感興趣的部份，並構造模型 (Model)，於其中定義許多參數 (Parameters)，而學習 (Learning) 就是讓這個模型以程式的方式執行，利用測試資料 (Training Data) 來調整參數，等到訓練樣本到一定的程度，參數的定義也更成熟，模型也就更接近真實世界，而能進行預測或是產品功能提升等應用。

#### 2. 資料探勘

應用機器學習的方法處理大量資料庫的數據則稱為「資料探勘」(Data Mining)，資料探勘就是從巨量資料中提取出未知的、有價值的潛在資訊。

資料探勘的實際工作是對大規模資料進行自動或半自動的分析，以提取過去未知的、有價值的潛在資訊，例如資料的分組 (可利用「聚類分析」進行)、資料的異常記錄 (可利用「異常檢測」進行) 和資料之間的關係 (利用「關聯式規則挖掘」進行)。

### 3. 資料分群

所謂的資料分群，指的是將資料中相似的個體聚集在一起並分類成數個群組，是一種將資料分類成群的方法，其主要目的在於找出資料中較相似的幾個群聚（clusters），並找出各個群聚的代表點。

應用在推薦系統上，可以事先將興趣、嗜好相似的用戶分群，藉由不同的推薦來有效地對特定群組行銷。

### 4. 類神經網路

類神經網路是一種由軟體和硬體所組成之計算系統，使用許多相連的人工神經元來模仿生物神經網路。透過人工神經元從外界環境或其它神經元取得所需資訊，經過運算之後，將結果輸出到外界環境或其它神經元。

以重複學習的方法，將一連串的例子交予學習，歸納出一個足以區分的樣式。若面對新的例證數據，類神經網路即可根據其過去學習的成果歸納，推導出新的結果，屬於機器學習的一種。資料挖掘的相關問題也可採類神經學習的方式，其學習效果佳並可做預測功能。

### 5. 迴歸分析

迴歸分析主要目標在於在變數間（自變數、因變數）找出其變動或受影響而變動之關係。建立輸入之自變數與輸出之因變數間的適切關係模型。例如：在自動駕駛的技術中，方向盤轉動的角度為輸出之因變數，各種偵測裝置得到的道路環境資料為輸入之自變數。

## 6. 決策樹

決策樹是進行預測或分類的分析手法，以樹狀圖來表示決策或行動前的條件。對於一項條件，分別會敘述著 YES 和 NO 兩種情況下的處理方式。

## 7. 關聯分析

也稱為購物籃分析，找出像是“購買產品 A 也有購買產品 B 的傾向”這樣的關聯資訊，最主要的目的在於找出什麼樣的東西應該放在一起？應用在藉由顧客的購買行為來瞭解是什麼樣的顧客以及這些顧客為什麼買這些產品，找出相關的聯想（association）規則。舉例來說，零售店可藉由此分析改變置物架上的商品排列或是設計吸引客戶的商業套餐等。

## 8. 自然語言處理

是以電腦程式對人們在日常溝通上使用的語言進行分析與運用，讓機器理解人類的語言。自然語言處理的應用相當廣泛，如大量的文字資料萃取出有用資訊的文字探勘、利用文字探勘分析社群媒體上產品與服務的評價都是。

## 9. 語意搜尋

從搜尋文章內詞類間的關聯性來分析“詞彙的意義”，並提升準確度的搜尋技術，系統會判斷輸入的關鍵字的意義，目標在於提升搜尋結果的精準度與參考價值，而不僅是列出包含有關鍵字、但卻無參考意義的搜尋結果。

## 10. 連結探勘

對社群網路或網頁間的連結、郵件收發信關係與文章的引用等，各種網路上

的聯繫進行分析的探勘技術。在社群網站如 Facebook 上常看到的運用，如：“你可能認識的人”或是具有影響力的“社群影響者”的層面。

## 11. A/B 測試

指的是為了實現網站的優化，同時推出幾個網頁版本（A 與 B 等），來進行哪個版本較受歡迎的測試。在不同版本中呈現不同網頁的內容、設計、版面與文字說明的變更，根據實際的效果來決定最後採用哪個版本。

### （二） 資料的運用

#### 1. 巨量資料下各種資料的分析方式

##### (1) 交易資料的分析方式

傳統的關聯性資料庫下所存放的客戶交易相關資料，均屬結構化的資料，一般是以資料倉儲存放，而常應用包括資料探勘、統計迴歸分析、隨機查詢(Ad hoc)、線上分析(OLAP)等方式，進行客戶分群，以瞭解客戶的樣貌。

##### (2) 互動記錄（非結構化文字）分析

透過自然語言處理，來理解文章中隱含的含義，包含了理解語言中微妙的含義、同意詞、縮寫等的的能力。文字探勘的軟體則用於輔助非結構化文字的分析，能讓金融機構迅速地從隱藏在非結構化的資訊中獲得洞察，像是利用公司郵件、社群媒體、演講文件、法律合約等等，理解含義並識別潛在的機會或威脅，有新的洞察等。

##### (3)數位通路資料（網站互動記錄）分析

透過網頁程式中標籤(Tag)或是使用「數位通路行為蒐集」軟體，目標在於較

能掌握客戶在數位通路上的行為，進而推測需求，甚至透過客戶使用瀏覽器所產生的視窗瀏覽 Cookie ID，目前已能將其與登入網站的使用者串連起來，識別出一特定客戶，可進行更進一步的統計或分析。

數位通路行為蒐集經營目的通常是強化兩大主題：「理解訪客特質」和「管理網站內容與功能分析」。蒐集到的數位通路行為資料本身能產生如頁面點擊、網頁瀏覽、使用者來源及環境等資訊，能夠提供網頁功能使用的分析重要依據，而在識別出特定客戶後，則可透過企業內已有的結構化客戶資料，更進一步的分析或行銷運用。

以下某一北美前三大跨國銀行應用實例：

<p><b>② 遭遇問題</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> 高流量網站並無法代表帶來高收入</li> <li><input type="checkbox"/> 無法從中得知客戶實際行為</li> <li><input type="checkbox"/> 無法辨別客戶從何通路進來 – 無法確認各數位通路的投資收益</li> </ul>	<p><b>③ 解決方法</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> 找出高放棄率的產品、內容及功能</li> <li><input type="checkbox"/> 辨別各線上服務是否被客戶所需要</li> <li><input type="checkbox"/> 改善流程以減少客戶放棄率 – 如精簡複雜的內容(如法律條文)</li> </ul>
<p><b>④ 解決目標</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> 希望能夠深入洞察客戶行為</li> <li><input type="checkbox"/> 希望能夠預測客戶期待那些新產品及新功能</li> </ul>	<p><b>⑤ 使用效益</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> 提高29%線上帳戶申請達成率</li> <li><input type="checkbox"/> 實現客戶追蹤及定位，增加線上客戶並強化交叉銷售能力</li> <li><input type="checkbox"/> 客戶滿意度增加，並透過不斷的改善以提供無縫的線上銀行體驗</li> </ul>

#### (4) 社群資料分析

社群資料來源包含各項社群媒體（如臉書、Instagram、Twitter 等），這些也都是使用者真實生活及想法的呈現，可以幫助企業量化瞭解自己的評價、討論議題的熱度、評量行銷方案的實際效果…等。社群討論資料屬於非結構化文字，分析方法與文字探勘類似。

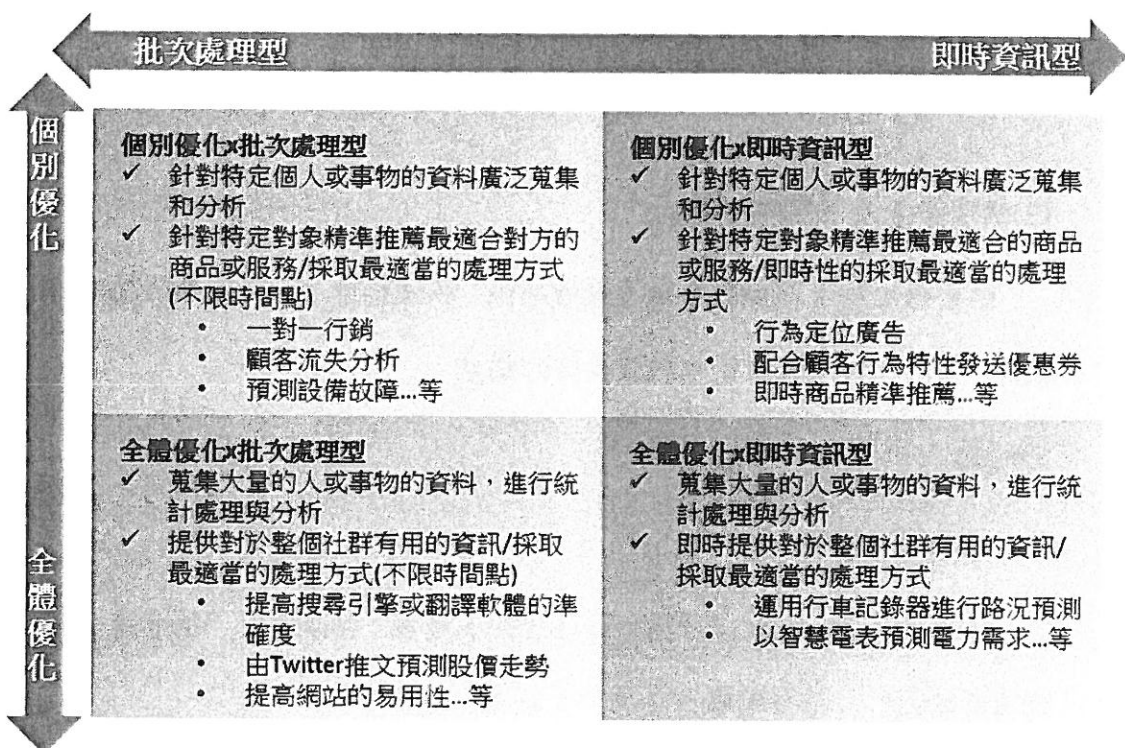
## 2. 巨量資料分析的四個運用模式

在資料處理模式上一般分為即時線上分析或批次分析處理，而資料分析運用的角度可分為總體概觀或對單一個體之探討分析。

在“大數據的獲利模式”一書中，提出 2x2 矩陣區分出四種應用模式，橫軸為『即時資訊型/批次處理型』，縱軸為『個別優化/全體優化』：

橫軸的『即時資訊型』指的是資料分析結果將依據受益者所處的狀況提供立即回饋；『批次處理型』不著重在反饋分析結果的時效性，但能處理較大量的資料。縱軸的『個別優化』指的是分析結果的受益對象為特定的個人或事物，針對其提供最適合的資訊服務或施以最適當的處理；『全體優化』指的則是分析結果的受益者並非特定的個人或事物，而是個人或事物所屬的群集。

用兩個指標，提出四種運用資料思考方式：



資料來源：野村總合研究所

## (1) 個別優化 x 批次處理型

針對特定個人或事物蒐集其相關資料，然後針對此對象精準推薦最適合的商品或服務，或採取最適當處理方式，但不限時間點。

### <案例> 新型態的車險定價方式

在美國目前已出現許多車險公司的定價方式，是依據行車記錄決定實際保險費的方案。保險公司提供依簽約車主的駕駛習慣，能享有不同程度的保險費折扣。

簽約車主將保險公司提供的專用設備安裝在車上後，設備會蒐集該車主的駕駛行為記錄，包含駕駛頻率、速度、行駛距離、緊急煞車次數等資料，以無線傳輸方式將資料回傳保險公司，保險公司則依據這些資料來決定保費的折扣率，這些資料也能提供給簽約車主檢視自己的駕駛習慣，修正駕駛方式。

依據不同車主各自駕駛習慣的詳細資料，相當於是“個別優化”模式，而決定出適當的保險費，是長期將行車記錄分析所產生的，屬於“批次處理型”。

## (2) 個別優化 x 即時資訊型

針對特定個人或事物蒐集其相關資料，然後對此人精準推薦最適合的商品或服務；或針對該事務採取最適當處理的模式。

### <案例>

#### ■ 線上訂房網站

針對已經進入到飯店訂房網頁頁面的瀏覽者，會彈跳出“目前有\_\_人正在瀏覽這家飯店”，促使瀏覽者能及早下訂單。

■ 美國特殊保險公司 **Assurant Solutions** 開發出一套客服機制，把每位顧客和客服中心人員定義出親和性分數，當客戶打電話進來時，不單只是將客戶轉接到當時有空的客服人員那裡，而是把客戶轉接到當下個性最適合應對該客



戶的客服人員。而據說該公司除了營收在六年內增加了 190%，顧客的解約防止率增加 117%，客服人員離職率也減少了 25%。

### (3) 全體優化 x 批次處理型

蒐集並儲存來自多數人或事物的資料，將儲存的資料進行整批統計處理與分析，提出對該個人或事物所屬的整個群集有用的統計資訊，或是以非即時的方式，優化相關業務。

#### <案例>

- 倫敦的德溫特資本(Derwent Capital)基金公司，針對數百萬則 Twitter 推文中的 10%進行隨機分析，將市場情緒分為“警戒”、“冷靜”、“熱衷”等類型，用以預測市場走向。
- 分析龐大的信用卡使用記錄，針對每位顧客找出顯示其卡片可能遭到盜用的模式，建立警示盜用的模型。從此便能在線上監控是否有盜用情形，或判斷可否核准該筆交易。
- 電子商務網站分析使用者點擊串流資料，找出有問題的頁面，用 A/B 測試頁面的方式，進行網站使用者體驗改善。

### (4) 全體優化 x 即時資訊型

蒐集儲存來自多數人或事物的資料，將儲存的資料整批先進行彙總性統計處理與分析後，配合所處的情境，即時性地對該群體提出有用的統計資訊，或進一步讓相關業務優化。

#### <案例>

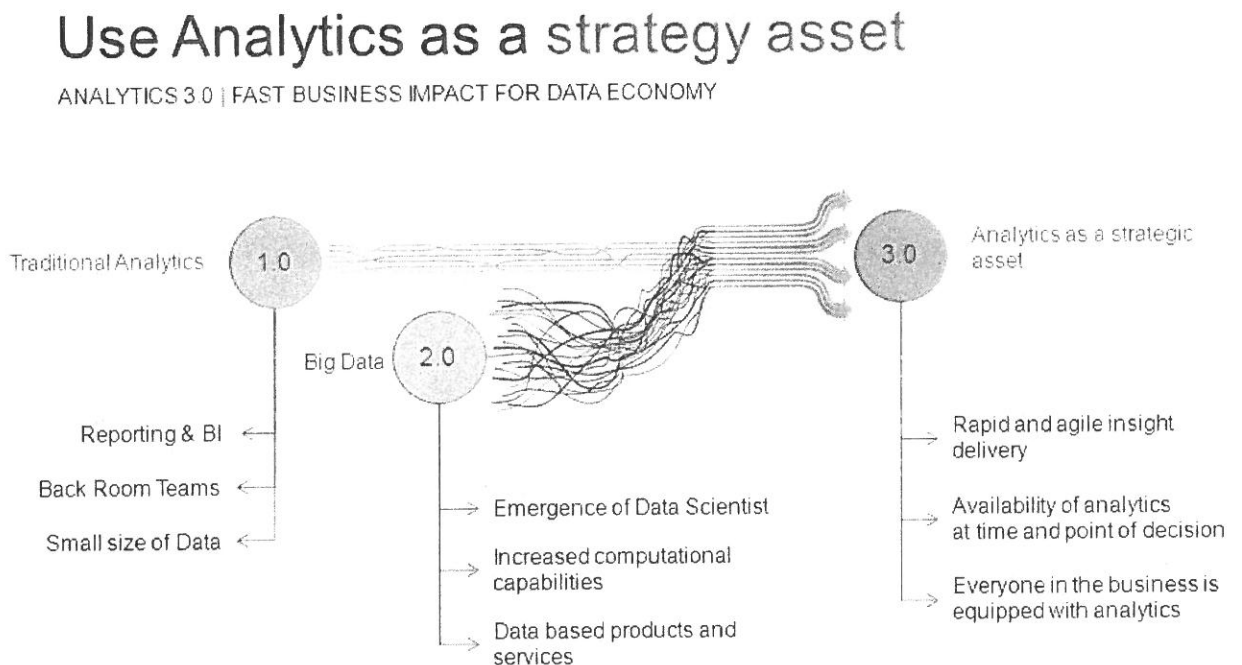
- 路況預測

以特定車輛作為探測車(Probe Vehicle)，蒐集交通參數資料，作為其他車輛之旅行時間預估，一般利用其既設之定位模組 (ex:GPS 定位)，回傳通過該特定道路上之旅行時間與車速資訊，藉此進行道路之旅行時間及交通資訊的蒐集。

### 三、 分析平台的建立

#### (一) 分析的再進化

“大數據的另一章”一書中提到，資料的分析，已由過去的傳統分析--以描述性及分析性報表為主的階段，進展成為可支援營運、即時、容納巨量資料的分析，在國際分析研究所 (International Institute of Analytics, IIA) 所提出的分析 3.0 框架中 (如下圖)，分析成為企業戰略的核心組成部分。



#### ■ 目標是分析，而不是產品

大量的產品正在收集資料，很多的情況下針對資料執行的分析是主要目標。

也就是說，當今的產品除了提供服務外，更是資料收集的機制，甚至在某些情況下，產品對客戶的價值就是產品蒐集資料後所提供的分析，例如手機的記帳軟體、睡眠品質監控軟體等。

以提供免費電子郵件的公司為例，目的不是單純的提供免費的電子郵件服務，而是當人們開始在使用郵件服務時，服務供應商可以獲得許多公司想知道的事情。像是當朋友之間密集頻繁地使用郵件與討論旅行時，服務供應商便將郵件瀏覽頁面的廣告欄位推播相關的商品或服務，服務供應商也能將這樣的訊息在除去個人訊息後賣給相關資源的公司。

#### ■ 正在打破的產業界限

隨著物聯網不斷被應用在許多商品上，產品紛紛嵌入了感測器、收集資料並為其客戶提供分析，產業界線模糊了。這不僅是新競爭的需要，也是客戶選擇產品的原因，因為能依據裝置所收集的資料和分析作為判斷，是更可靠的依據。

『Nike』是許多人再熟悉不過的品牌名稱，絕大多數的人對這個品牌的印象會是運動產品、設備製造商，然而不只是這樣，2012年時 Nike 發表了一款像手錶的裝置 FuelBand（後來與 Apple 合作，即現在的 Apple Watch NIKE+），可以測量每日行走的步數及關於睡眠模式等幾個指標資訊。

購買產品的客戶需要在電腦或手機下載軟體以搭配使用，使 Nike 開始涉足了軟體產業的領域，而裝置收集的各類資料，又使 Nike 進入資料收集領域甚至可以進一步分析資料進入健康業務領域。

## （二） 奠定分析基礎—制定計劃

為了完成分析的需要，企業需有人力、工具、技術等多方面的投資，但企業要取得內部一致的共識及批准並不容易，因此在規劃建立分析平台時，制定計劃是相當重要的基礎。

## 1. 設定優先順序，制定計劃

在資料收集及分析之前，應先設定好分析的目的是為了什麼，為了解決什麼業務問題，應先將業務待解決的問題釐清，目標不應是為了取得新資料或技術，並將問題依其重要性、急迫性排定解決的優先次序。

## 2. 選擇合適的決策標準

選擇計劃時，決策標準是重要的，哪些因素應最大化，哪些應最小化，有一明確的標準才能正確的鎖定目標並理解選擇帶來的影響。例如：

- 哪個選項下，企業能更有效率的搭建、部署及測試新的營運分析流程？
- 利用新技術進行實驗性的探索有多容易？

選擇決策的標準時除了可見的設備成本、軟體成本外，也應將人力成本、維護成本等與易用性或企業內部實際條件有關因素列入考量。

### （三） 建立分析平台

為了使分析成為營運的重要支援，尤其在巨量資料的時代，不同的資料類型及不同的分析需求，使企業需建立一套統一的平台，在『大數據的另一章』一書中，作者提出建立統一的分析環境的三大支柱：

#### 1. 關聯式資料庫

關聯式資料庫是現今每一家企業都會使用的，在具備著平行處理的關聯式資料庫中，不僅在資料量和處理上具備大規模計算能力，資源管理能力也是強大的一環，通常是部署分析的最佳處。

#### 2. 非關聯式資料庫

能夠處理常見的非結構或半結構化資料，但在資料格式上的靈活，也相對喪失分析使用的靈活性，非關聯式資料庫為企業提供了一個能存放低價值資料的同時還能在需要時找得到的方式。

### 3. 探索平台

探索不是一個新的概念，但卻因為分析工具的使用方式而有了新的改變，一個探索平台能支援關聯式資料處理，也能支援非關聯式資料處理，能簡化以往的資料探索過程，沒有太多的約束和限制，目前市面上的產品如 Aster、Greenplum 不僅能提供自己的分析演算法，也能支援如 SAS、SPSS 或 R 這些常見的分析工具。對於探索平台而言，處理速度和可擴充性都不如靈活性和易用性來得重要。

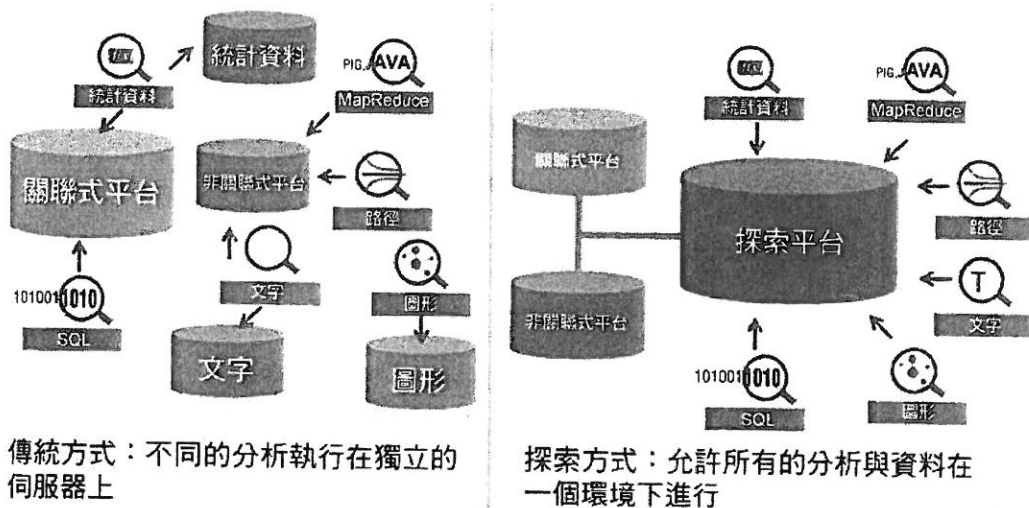


圖 5-4 探索平台如何簡化分析

建立一個統一的分析環境的終極目標是能透過「任何分析方法」在「任何時間」分析「任何類型和規模的資料」。然而企業的分析需求量在不同階段是會變化的，企業依據成本收益考量決定是否需要增加工具元件（或支柱），或發展現有工具，亦或是尋找其他替代方案，也可以定期的方式審查評估，一旦決定需求量已足以增加構建新的元件，企業應立即行動並開始引進新元件的前置作業，因為引進會帶來更多的複雜性及功能要求。

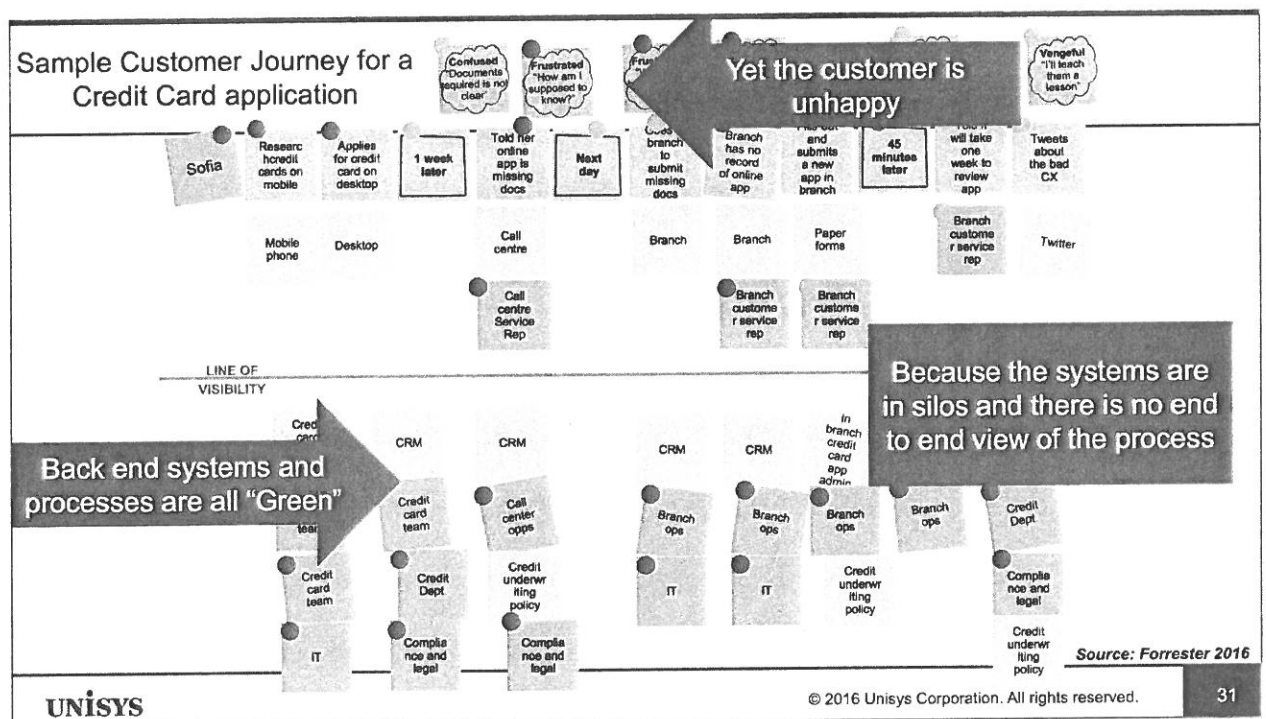
#### (四) 金融業分析應用案例

以下提供實際上應用客戶行為資料、交易資料等分析應用的場景：

##### 1. 客戶行為體驗改善

此種應用類似前面四種模式的個別優化 x 批次處理型，銀行透過資料的分析及整合，目標是能在所有的通路、所有服務項目，提供客戶整合的、一致的資訊，並提供客戶少阻力（frictionless）的用戶體驗，各個業務資料、客戶行為資料、交易資料的整合能使銀行完整的認識一個客戶，甚至在任何一個通路的接觸機會都能精準地提供客戶需要的資訊及功能。

下圖是以客戶申請信用卡的歷程為例，主要在說明即使從企業現有服務檢視的角度來看，該做的都做了，但當銀行內部的資料是分散在各個系統中、服務亦未經整合時，常會發生客戶在過程有不好的服務體驗，而企業卻未能意識到。



##### 2. 精準行銷

對客戶群進行精準定位，分析出潛在金融服務的需求，並運用客戶行為分析

結果，發展鎖定特定客群行銷推廣，有效提升經營績效。

#### (1) 數位通路行為的分析

在這個已經很少上分行的時代，數位通路行為的觀測變成不可或缺的一項重要工具，而透過數位通路行為的分析，可以針對符合事先定義好條件的使用者進行行銷活動的推播，例如某 B 客戶連續兩天透過網路銀行或手機 App 觀看基金資訊超過半小時，這項訊息透露客戶可能有購買基金的需求，因此接下來銀行可以透過網頁或 email 推薦，既準確也能使成本的效益提高，減少過去在缺乏參考資訊的情況下，毫無章法的推銷所帶來的支出。

#### (2) 分析客戶需求，提供客製化服務

採用分析系統觀察與分析顧客消費行為，並對其作出適當建議，而於虛擬通路上，結合巨量資料平台，幫助銀行掌握客戶的動向與需求。

### 3. 詐欺行為分析

在有效預防與管理金融犯罪上，巨量資料分析所得的客戶消費行為模式，藉由即時詐欺交易識別與分析，一旦發現不正常消費，可立即得知並採取行動，在極短時間內找出盜用者，及時遏阻銀行因盜用事件所可能蒙受之損失。

### 4. 加值服務、優化服務

透過巨量資料分析，金融業可監控各種市場推廣通路運作狀況，將客戶行為轉化為可分析的資訊，從中了解客戶的個性特徵、風險偏好，瞭解客戶的金融往來習慣及使用行為，將精準行銷擴展至服務的創新與優化。

例如國內某家銀行其每一家分行服務櫃台之擺設、理財專區、座位的設計，係依照資料庫中分行所在地的人口特性、年齡及交易量複雜度等數據，客製化該分行的業務洽辦區位。

### 5. 因應潛在風險，貸款違約處理

風險控管是金融業信託貸款管理業務中的重要組成部分，對於目前訊息和資料在網路上不斷地膨脹，包含社群平台、文章討論等，利用巨量資料的發產所衍生出來分析應用的如機器學習，提供金融業者有更全面衡量個別客戶風險的機會。

透過機器學習，可以將更多的資訊加入作為評估借款人的還款能力等各種風險評估作業，而目前在美國已有部分銀行在客戶開戶時，要求提供臉書帳號，列為授信評分其中一個，透過客戶的虛擬社群，更加了解客戶，成為客戶取得好的貸款條件之一。

而多面向的客戶資料，透過演算法及模型計算，銀行也能更早預測可能潛在違約風險的客戶，能有更積極的因應作法，像是安排人員關心或提供小額的信貸協助。貸前、貸中及貸後都能夠藉由巨量資料的技術運用獲得許多有價值的資訊。

#### 四、 隱私權問題

##### ■ 隱私與創新的兩難

在企業追求服務貼近使用者的需求的想法下，隨著許多關於行為資訊的蒐集工具越發的普遍，物聯網更是將資訊的利用推進了一個新的層次，而另一方面最終將越來越接近個資或隱私權等敏感性資訊，是無法被逃避掉的議題，尤其若期望能透過其創造出巨大的商業價值。

##### ■ 歐盟與美國的做法

巨量資料技術發展與隱私權保護的價值爭議由來已久，而相關的法令政策在具體上則為美國與歐盟兩種取向為代表。美國在 2014 年由總統執行辦公室所發表的「大數據白皮書」可看出，美國政府看重巨量資料所為經濟社會發展帶來的創新動力，對於與隱私權產生的衝突，以解決問題的態度、希望以改良的政策框架與法律規則來處理。

而相對於美國，歐盟對於侵犯個人數據的行為處罰措施十分嚴格，與商業創



新發展的考量相比，個人資料被認為更具保護價值，於 2009 年修法將利用 Cookie 蒐集使用者行為紀錄等定為應採“選擇性參與”，也就是儲存在使用者終端設備裡的資訊，應在已明確且全面性地將使用目的告知使用者、並取得其事先同意的情況下，始准許使用；並於 2016 年 4 月通過延宕許久的「一般性個人資料保護規則(GDPR)」，正式使該法全面性的直接適用於歐盟會員國內，並將「被遺忘權」具體列於條文中。（「被遺忘權」維基百科解釋：人們有權利要求移除自己負面或過時的個人身分資訊搜尋結果）

## ■ 基礎原則

觀察以美國、歐盟分別為代表的隱私政策，可以歸納出共通的基礎原則：

- 無論是什麼通路或是產品，有關於隱私的政策說明，使用者應要容易取得，也就是說企業應面對並負起告知義務之最基本的要求。
- 使用者有權知道究竟有哪些資料被收集，及蒐集後將如何被使用，企業需能配合消費者此權利的行使。
- 有權要求刪除或修改資料：前提及的歐盟，及臺灣的個人資料保護法均涵蓋這一塊，個人有權利要求刪除過時、失真的資料。

## ■ 與使用者的對話、互惠

有關「透明度」、「使用者參與決策的機會」等具體的實踐原則，其實並不存在標準答案。但面對此問題的關鍵其實在於需要「與使用者對話」。

試想使用者在自己不知情的情況下，被蒐集關於自己的資訊，然後就在上網瀏覽時，出現看似依據自己的興趣或曾經搜尋的主題相關的廣告，這樣另類的“被跟蹤”，總是會帶來不舒服的感覺。但若反過來，例如能事先控制傳送的廣告內容，一方面使用者有被告知、受尊重的感覺，另一方面也能降低使用者因頻繁看到不感興趣的廣告內容的不悅。

舉例來說，某些網站提供讓使用者選擇能看到廣告的類型及內容，使用者只

會看到自己感興趣的廣告內容，另一方面網站業者也能直接鎖定對產品可能有興趣的使用者，可說是雙贏。

像這樣透過與使用者對話，決定傳送內容，也就能在某種程度上達到「確保透明度」、「確保使用者參與決策的機會」等條件。如果只是對業者有單方面的好處，勢必難以獲得使用者的理解，而如果能對使用者帶來好處，取得同意的難度就不會太高了。

## 肆、 本行現況及建議事項

### 一、 結論

#### (一) 先有研究/分析的目的，如何研究、使用技術是其次

無論是哪一種行為研究，執行的目標在於了解客戶的行為、趨勢、喜好、生活，但分析結果的下一步呢？仍然需要回歸到清楚的目標策略或任務，也才能針對研究獲得的結果進行下一個因應的準備動作。

#### (二) 資料品質及時效性依然重要

資料品質及時效性對分析一直是很關鍵的，當分析隨著技術突破過去效能的限制時，可以預期分析的工作會越來越多種類，進而成為一種營運的決策支援，此時實際上沒有機會修正資料錯誤，資料需要做到及時和準確。

#### (三) 許多界線逐漸模糊

產品界線的模糊—產品不再只是產品，更像是一種通路、或是手段，透過提供服務、收集資料紀錄，企業得以進行改善或研發。

產業界線的模糊—如同 Brat King 所說的，不再是“Bank”，而是“Banking”，面對“Banking”的需求，不一定得開銀行，而是要能提供市場需要的服務。產業經驗並不見得會構成太大的限制，換個方向思考，將現有的金融服務延伸至各種生活的需求，對不同產業的企業來說機會是不會相差太大的。

#### (四) 分析文化的建立與態度影響成效

分析、資料使用的觀念建立，第一步始於高層決策者，但要能使分析的效果達到最佳，取決於是否組織中每一個角色知道資料的價值。須確保參與的所有人能改變行為習慣，例如依照推薦的結果或方式進行，若忽視建議內容依照原本的

習慣進行，分析的價值就無法產生。若分析過程有問題，則應該修正這些問題，但不應讓員工自行任意改變推薦的建議。

分析的結果或者功能目的不應在於指責，而是在於幫助員工將事情做得更好，這種態度必須成為企業文化的一部分，否則分析文化的推行勢必有更多的阻力。

## 二、 本行現況

隨著金融商品及服務的多元化，且眾多銀行皆提供類似或有稍微差異的產品或服務，銀行的客戶常常有在同一銀行有多個帳戶、多張不同的信用卡作為不同用途的使用，而客戶的資料散落在各個業務系統中。

目前行內的客戶主檔及交易相關行為資料呈現分散的狀態，如中心帳務主機系統、財富管理系統、基金系統等，而資料倉儲系統則是目前統一收集各業務系統資料的平台，行內於今年（2016年）年中亦啟動了巨量資料處理平台專案，提供非結構化資料的存放空間，並於該專案中，購置了數位通路行為資料分析工具及視覺化分析工具。

而有關個人金融方面的業務，分行則扮演主要的銷售窗口，客戶可從個人網路銀行直接購買基金等理財商品、或安裝軟體透過一般電腦上網購買證券商品，亦可透過手機下載 App 進行購買。

## 三、 建議

### （一） 資料架構的整理與單一客戶視圖的建立

隨著金融商品及服務的多元化，銀行的客戶常常在同一銀行有多種帳戶、多張不同的信用卡作為不同用途的使用，客戶的資料散落在各個業務系統中，資料倉儲系統則是各業務系統欲使用其他系統資料源時的主要管道之一，配合業務系統將指定的資料收納後提供使用。

目前資料倉儲扮演的角色主要在於存放資料並提供一致格式的資料，對於許多不同來源系統的資料放入資料倉儲後，尚未有一個較具系統性的、有架構的整理收納，業務系統於使用資料時多是自行將資料取回後，再進行統計、彙整及後續的分析應用。當各類資料處於分散的狀態時，要獲得一個完整的客戶視圖——每個客戶與銀行的所有往來互動紀錄，包含所有產品交易諮詢紀錄和通路的行為資訊，是不容易完成的。

因此若能將資料進行有系統性、有規則可依循的整理，例如透過引進“資料模型”架構產品，將資料依業務、主題區分並進行基礎的切割、合併等處理，將有助於提升各系統使用資料的能力，而面對一些共通性資料的需求，如基礎統計、彙總性資訊等，也可直接在資料倉儲內完成基本的整理，將能縮短業務系統的作業流程；前面提及的單一客戶視圖的建立也能在這樣的基礎上完成，結合其他行為蒐集或統計分析資料，對於客戶的掌握能力將大大提升，有助於行銷及業務推展，未來產品的制定或其他重要業務行銷決策，也能夠提供參考資料。

無論是資料架構的整理或是單一客戶視圖的建立，目的都是為了提升資料的使用，在架構設計上需考量其擴充性及使用性，資訊單位也需與業務單位進行詳實的訪談並取得共識，業務單位也可一併重新審視系統可進行的應用或調整。

## （二） 探索平台的建立

行內現在已具備了資料倉儲，刻正建置巨量資料處理平台，處理結構化及非結構化的資料存放，於安裝完工具後搭配程式的撰寫可完成資料分析查詢使用。

可預期的是未來兩類資料結合使用的需求會越來越多元，前述的探索平台建立，目標在於能同時查詢結構化及非結構化的資料、以統一且簡單的資料存取方式，且降低使用上的技術門檻，使分析的需求單位能夠較容易取得資料並進行資料探索，易用性及靈活性會是重要的評估方式，能使得資料探索的過程變得容易且有效率。

### (三) 分析文化的建立，包含團隊、及資料使用的推廣

#### 1. 跨部門團隊的建立

以本行現行分工架構，在資訊單位對於業務功能需求缺乏業務知識時，以及業務單位未提出足夠明確需求情況下常會有理解上的誤差，導致開發測試過程延遲。而當要進入資料的分析進行時，業務單位及技術單位對於結果有一定程度的共識與否，事關到分析結果是否有真正的價值，因此，跨部門團隊的成立目標在於能夠所有成員使資料的使用、資料分析的目的及結果能有相同的期待，使分析作業能產生有價值的成果。

許多分析尤其是在探索的過程時，都不會產生期望的結果，但關鍵還是在於如何管理失敗，讓失敗可以掌控或預期，並且知道如何處理失敗。

#### 2. 從總行部處開始，作為先導使用者，再逐步漸進式的推廣。

資料分析的使用要實施在企業中，由於是新的嘗試仍在摸索的階段，可以從業務單位中挑選出其需求較急迫或是依其他考量選出先導使用單位，待機制、流程能運作順暢時，再逐漸往更多業務單位推廣。

#### 3. 分行行員也需懂得使用資料

在數位金融時代的衝擊下，分行將面臨減少的命運，行員扮演的角色可能將會是更重要的顧問諮詢，需要具備多元、豐富的專業服務技能，因此解讀資料、運用資料的能力亦將能提升行員在提供諮詢時的服務專業度，為一項重要的優勢。

## 伍、 參考文獻

1. 金融研訓院，「金融數位力」，2016 年版。
2. 結構、半結構、非結構式資料是啥意思？-- Kevin 的 MongoDB 工作筆記  
<https://kevinwang.gitbooks.io/bigdata/content/general/structured-data.html>
3. 什麼是機器學習(Machine Learning)?  
[http://tzanfeng.pixnet.net/blog/post/25614373-%E4%BB%80%E9%BA%BC%E6%98%AF%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92\(machine-learning\)%3F](http://tzanfeng.pixnet.net/blog/post/25614373-%E4%BB%80%E9%BA%BC%E6%98%AF%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92(machine-learning)%3F)
4. 資料探勘，維基百科。  
[https://zh.wikipedia.org/wiki/%E6%95%B0%E6%8D%AE%E6%8C%96%E6%8E%98#cite\\_note-acm-1](https://zh.wikipedia.org/wiki/%E6%95%B0%E6%8D%AE%E6%8C%96%E6%8E%98#cite_note-acm-1)
5. Data Mining 的十種分析方法。  
<http://dasanlin888.pixnet.net/blog/post/34467788-data-mining%E7%9A%84%E5%8D%81%E7%A8%AE%E5%88%86%E6%9E%90%E6%96%B9%E6%B3%95>
6. 一文讀懂機器學習 <https://read01.com/az5gje.html>
7. 朱啟恆，大數據於金融業之應用，財金資訊季刊 / No.84 / 2015.10  
<https://www.fisc.com.tw/Upload/8eaaa580-8592-4511-bd3d-a95d0d5ccdc8/TC/8402.pdf>
8. IBM 智慧型領導白皮書：以六項簡單步驟完成智慧型治理。  
<ftp://public.dhe.ibm.com/software/tw/data/db2/six-steps-smart-governance.pdf>
9. 博誌出版社，大數據的另一章！，
10. Omnichannel banking，優利公司簡報。
11. BRETT KING 著，孫一仕譯，「BANK 3.0-銀行轉型未來式」，2013 年 10 月。

12. 城田真琴著，鐘慧真、梁世英譯，「大數據的獲利模式」，2013年8月。
13. Bill Franks 著，張建輝、車皓陽、劉靜如、范歡動譯，「大數據的另一章！」，  
2016年6月。