

frilly / September 15, 2015 09:49PM

[資料分析的五大思維方式](#)

今天要講[url=http://www.finereport.com/tw/]資料分析[/url]的五大思維方式。

首先，我們要知道，什麼叫資料分析。其實從資料到信息的這個過程，就是資料分析。資料本身並沒有什麼價值，有價值的是我們從資料中提取出來的信息。

然而，我們還要搞清楚資料分析的目的是什麼？

目的是解決我們現實中的某個問題或者滿足現實中的某個需求。

那麼，在這個從資料到信息的過程中，肯定是有的一些固定的思路，或者稱之為思維方式。下面零一給你一一介紹。（本文用到的指標和維度是同一個意思）

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujvfenxi1.png[/img]

第一大思維【對照】

【對照】俗稱對比，單獨看一個資料是不會有感覺的，必需跟另一個資料做對比才会有感覺。比如下面的圖a和圖b。

圖a毫無感覺

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujvfenxi2.png[/img]

圖b經過跟昨天的成交量對比，就會發現，今天跟昨天實則差了一大截。

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujvfenxi3.png[/img]

這是最基本的思路，也是最重要的思路。在現實中的應用非常廣，比如選款測款、監控店鋪資料等，這些過程就是在做【對照】，分析人員拿到資料後，如果資料是獨立的，無法進行對比的話，就無法判斷，等於無法從資料中讀取有用的信息。

第二大思維【拆分】

分析這個詞從字面上來理解，就是拆分和解析。因此可見，拆分在資料分析中的重要性。在派代上面也隨處可見「拆分」一詞，很多作者都會用這樣的口吻：經過拆分後，我們就清晰了……。不過，我相信有很多朋友並沒有弄清楚，拆分是怎麼用的。

我們回到第一個思維【對比】上面來，當某個維度可以對比的時候，我們選擇對比。再對比後發現問題需要找出原因的時候？或者根本就沒有得對比。這個時候，【拆分】就閃亮登場了。

大家看下面一個場景。

運營小美，經過對比店鋪的資料，發現今天的銷售額只有昨天的50%，這個時候，我們再怎麼對比銷售額這個維度，已經沒有意義了。這時需要對銷售額這個維度做分解，拆分指標。

銷售額=成交用戶數*客單價，成交用戶數又等於訪客數*轉化率。

詳見圖c和圖d

圖c是一個指標公式的拆解

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujvfenxi4.png[/img]

圖b是對流量的組成成分做的簡單分解（還可以分很細很全）

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujvfenxi5.png[/img]

拆分後的結果，相對於拆分前會清晰許多，便於分析，找細節。可見，拆分是分析人員必備的思維之一。

第三大思維【降維】

是否有面對一大堆維度的資料卻促手無策的經歷？當資料維度太多的時候，我們不可能每個維度都拿來分析，有一些有關聯的指標，是可以從中篩選出代表的維度即可。如下表

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujvfenxi6.png

[/img]

這麼多的維度，其實不必每個都分析。我們知道成交用戶數/訪客數=轉化率，當存在這種維度，是可以通過其他兩個維度通過計算轉化出來的時候，我們就可以【降維】。

成交用戶數、訪客數和轉化率，只要三選二即可。另外，成交用戶數*客單價=銷售額，這三個也可以三擇二。

另外，我們一般只關心對我們有用的資料，當有某些維度的資料跟我們的分析無關時，我們就可以篩選掉，達到【降維】的目的。

第四大思維【增維】

增維和降維是對應的，有降必有增。當我們當前的維度不能很好地解釋我們的問題時，我們就需要對資料做一個運算，增加多一個指標。請看下圖。

[img]http://www.finereport.com/tw/wp-content/themes/BusinessNews/images/shujufenxi7.png[/img]

我們發現一個搜索指數和一個寶貝數，這兩個指標一個代表需求，一個代表競爭，有很多人把搜索指數/寶貝數=倍數，用倍數來代表一個詞的競爭度（僅供參考）。這種做法，就是在增維。增加的維度有一種叫法稱之為【輔助列】。

【增維】和【降維】是必需對資料的意義有充分的了解後，為了方便我們進行分析，有目的的對資料進行轉換運算。

第五大思維【假說】

當我們拿不準未來的時候，或者說是迷茫的時候。我們可以應用【假說】，假說是統計學的專業名詞吧，俗稱假設。

當我們不知道結果，或者有幾種選擇的時候，那麼我們就召喚【假說】，我們先假設有結果，然後運用逆向思維。

從結果到原因，要有怎麼樣的因，才能產生這種結果。這有點尋根的味道。那麼，我們可以知道，現在滿足了多少因，還需要多少因。如果是多選的情況下，我們就可以通過這種方法來找到最佳路徑（決策）

當然，【假說】的威力不僅僅如此。【假說】可是一匹天馬（行空），除了結果可以假設，過程也是可以假設的。

我們回到資料分析的目的，我們就會知道只有明確了問題和需求，我們才能選擇分析的方法。

順帶給大家講講三巨量資料類型。這個屬於偷換概念，其實就是時間序列的細分，不是真正意義上的資料類型，但這個卻是在處理店鋪資料時經常會碰到的事情。資料放在坐標軸上面分【過去】、【現在】和【未來】

第一 巨量資料類型【過去】

【過去】的資料指歷史資料，已經發生過的資料。

作用：用於總結、對照和提煉知識

如：歷史店鋪運營資料，退款資料，訂單資料

第二 巨量資料類型【現在】

【現在】的概念比較模糊，當天，當月，今年這些都可以是現在的資料，看我們的時間單位而定。如果我們是以天作為單位，那麼，今天的資料，就是現在的資料。現在的資料和過去的資料做比較，才可以知道現在自己是在哪個位置，單有現在的資料，是沒什麼用處的。

作用：用於了解現況，發現問題

如：當天的店鋪資料

第三 巨量資料類型【未來】

【未來】的資料指未發生的資料，通過預測得到。比如我們做得規劃，預算等，這些就是在時間點上還沒有到，但是卻已經有了資料。這個資料是作為參考的資料，預測沒有100%，總是有點兒出入的。

作用：用於預測

如：店鋪規劃，銷售計劃

三種資料是單向流動的，未來終究會變成現在，直到變成過去。

他人我不知道，但我自己非常喜歡把資料往坐標軸上面放，按時間段一划分，每個資料的作用就非常清晰。

4500+企業選擇FineReport報表與 [url=http://www.finereport.com/tw/]BI[url]

[url=http://www.finereport.com/tw/]商業智慧[url]工具【免費下載】

opensource開發，類excel設計，全方位異質資料庫整合，資料填報、Flash列印、權限控制、行動應用、客制化、交互分析、報表協同作業管理系統。

分享自：愛數據
