

## 分母錯誤

資料來源：《為什麼數據會說謊》，彼得·施萊弗斯著，

張 羿譯，中信出版集團，2023

城市的交通擁堵嚴不嚴重，衡量的指標是什麼？《城市交通報告》是美國最被廣泛引用的關於城市擁堵及其相關成本的報告，該報告用來衡量的指標是「行程時間指數(TTI)」。TTI 是一個簡單的比值，即在高峰時段的出行時間，與在無擁堵條件下的出行時間之比。例如，在交通順暢情況下，通勤需要 1 小時，高峰時間需要 1.5 小時，TTI 就是 1.5，在這種情況下，每天「損失的時間」是 1 個小時(單程半小時)。

TTI 公式乍看之下相當合理，如果通常需要 1 小時，但在尖峰時間需要 1.5 個小時，當然比在尖峰時間只需 1 小時 10 分鐘要糟糕。但是我們把思路擴展到整個城市，數以百萬計通勤者的總出行時間，如果使用這樣的指標來評估，TTI 為 1.2 的 A 市是不是就比 TTI 為 1.3 的 B 市好？如果是的，那是因為 A 市的人開車上班的時間更長，還是因為 A 市上班的通勤時間本來就比較長，所以花在交通擁堵的時間比例比 B 市少？如果不知道每個城市的平均通勤時間是多少，這樣的指標是沒有意義的。Cortright 深入的研究後，對 TTI 提出了嚴厲的批評，指出在搜集數據、確定平均車速和油耗的模型方面都存在諸多的缺陷，他指出了許多用 TTI 衡量指標，扭曲了現實通勤情況的實際例子。例如，1982~2007 年波特蘭市由於城市規劃和交通政策的改善，從每天 54 分下降到 43 分鐘，這本來是個好現象，可是 TTI 卻從 1.07 上升到 1.29，擁堵率從 7% 上升到 29%，擁堵排名更難看。而真正的主要原因，是市民平均的通勤距離從 19.6 哩降到了 16 哩，完全足以說明這種指標的扭曲和荒謬。進一步考察不同 TTI 得分高低的 cities，並比較它們的平均通勤時間時，就會發現 TTI 似乎完全顛倒了城市通勤狀況的好壞。

作為衡量指標，TTI 是失敗的，因為它誤用了分母。我們的目標應該是讓通勤花費更少的時間，而不是比一些理想情況花費更少的時間，這就是使用非擁擠的時段來衡量額外的時間，反而忽略了花在通勤上的大部分時間，所產生的扭曲。TTI 關注的是快速移動，而不是關注你要走多遠或多久，這種測量追求的是達到最大速度，而不是減少通勤時間，於是就變成長途通勤要比短途通勤更好。其錯誤在於使用一個高峰時段的通勤時間，與在通暢交通中的通勤時間的比率，並假設減少擁堵會改善測量結果，忽略了另一種降低比率的方法，也就是增大分母，本質上等於增加出行距離。實際上影響通勤時間的因素，除了擁堵以外，行駛速度、停等紅綠燈都是。或許更好的衡量指標是：只報告每座城市的平均通勤時間再加上標準差。

由此我們得到一個幾乎可以用於任何衡量指標的教訓，那就是：在衡量指標中正確使用了「每」了嗎？如果用了，很可能就會變成，提高了「分數」實際上反而會惡化了狀況。

請思考一下：美國千禧一代比 X 一代的人少開車嗎？汽車擁有量在下降嗎？應該如何衡量這些問題？

根據 2015 年在《彭博商業》和《大西洋月刊》上發表的幾篇文章，這二個問題的答案都是「否」。這二家雜誌都表示，2014 年「千禧一代」購買了 370 萬輛汽車，而 X 一代只購買了 330 萬輛。很明顯的，千禧一代比他們老一輩購買了更多的汽車，對吧？所以汽車擁有量在年輕一代是增加的，這種比較顯現的問題很簡單。可是，實際上千禧一代要比 X 一代人口多得多：7800 萬比 4900 萬，而且還取決於劃分這些群體的方法。文章中定義的千禧一代是出生於 1977~1994 年共有 17 年，而 X 一代則是出生於 1965~1976 年只有 11 年，所以千禧一代的群體更大只是因為它包括了更大的年齡段。根據數據，千禧一代每千人購買 47.5 輛汽車，而 X 一代每千人購買了 67.1 輛汽車。《城市觀察》發現了這個明顯的錯誤，並發表了一篇反駁文章，指出他們完全忘記了除法。

這個問題與 TTI 相反，TTI 是忘了在衡量指標中使用分母。人們試圖回答的問題不是千禧一代購買的汽車總數是否超過 X 一代，而是他們每人購買的汽車是否更多。忘記了除法，只簡單的以為是用「每」來衡量的情況，其實是比比皆是，不足為奇。這提示了我們是要測量強度問題，還是原始規模問題，常常會被混淆。我們經常看到一些測量方法，聲稱的是一件事，但實際上說的是另一件事，這些情況隨處可見。這類「分母忽略」或「分母膨脹」的錯誤，作者稱它為「每」的問題，或稱之為「中國謬誤」。政治人物、評論家、記者經常喜歡把事物拿來和中國比較，例如：你想證明你的國家污染不嚴重，就把排放總量和中國比較；你想證明你的國家哪些項目沒有足夠的投資，就和中國的投入相比，只提總數，而不提中國的龐大人口。

註：中國前總理溫加寶有句名言：

「任何小事，乘以 13 億，都會變成天大的事；任何大事，除以 13 億，也會變得非常渺小。」

衡量強度或效率的指標應該使用分母，在許多情況下，最好的分母是人數，即人均，但經常被有意的忽視。在城市之間的比較最常看到的是基於總體的衡量，如暴力犯罪數量、公園數量。要判斷哪個國家的人最富有、最聰明，就需要看人均。但是，不只忽略分母會導致問題，使用錯誤的分母同樣也會產生誤導。

對美國的行人來說，紐約似乎是相當危險的地方，因為平均每三天就有 1 個行人命喪街頭。根據美國國家公路交通安全管理局的數據顯示，2012 年，在美國 50 萬以上人口的城市中，紐約有 127 名行人死亡，超過了洛杉磯 99 人、芝加哥 47 人、舊金山 14 人等城市，以及更依賴汽車的休斯頓 46 人或鳳凰城 39 人。在全美所有的交通死亡事故中，14%的受害者是行人，而紐約，這一比例是 47%。衡量行人安全是一個強度問題，我們應該使用人均衡量指標來反映這個問題，在 2012 年時，紐約有 830 萬人口，拿來與人口為 150 萬的鳳凰城相比顯然並不恰當。但是，如果以紐約每年每 10 萬人中有 1.52 名行人喪生來看，還是要比全美的平均水準 1.51 人略差。然而，事實上紐約是美國行人最安全的地方之一，這又怎麼說呢？因為紐約約有 10%的人步行上班，如果再加上乘坐公共交通工具的人（因為在起點和終點還是需要步行），這個比率就達到 65%，躍居全美第一。所以使用人均指標，有時也不能正確反映試圖衡量的強度，以行人安全的角度，用步行人數的比例來衡量可能更好些。

另一個因計算方式被誤解的指標與疾病有關，導致扭曲的不是分母而是分子。通常衡量疾病的影響有三個指標：流行率、發病率和死亡率。流行率是指特定人群中患某種疾病的人數，如每 10 萬人中患流感的有 10 人；發病率是指在給定的時間段內，特定人群中有多少人患這種疾病，如一年內每 10 萬人中有 50 人患瘧疾；死亡率是指特定人群中死於該疾病的人數，如每 10 萬人中有 100 人死於癌症。

我們當然希望患病的人少一些、帶病生活的人少一些、死於疾病的人少一些，但是問題在於其中一項測量的積極變化，會使另一項測量惡化。例如，如果發現了一種可以延長瘧疾患者壽命的新療法，如果發病率保持不變，那新療法就會提高了瘧疾的流行率。相反，一個很快就能致死的疾病，流行率就會很低。因此，一種疾病的流行率下降，並不是因為感染的人減少了，而可能是因為死於這種疾病的速度加快了，所以未必就是個好消息。

分母不僅要針對被測量的特定目的或目標，還要有意義，它必須與某一事物的目的、影響或作用有合理的聯繫。如果，我們通過「每」本教科書的成本來衡量教育系統的成本，不是很愚蠢嗎？不幸的是，在很多情況下，「每」都是沒有意義的。以「人均」溫室氣體的排放量來看，加拿大高居世界第二，但是以每平方公里衡量溫室氣體的排放強度，加拿大就是全世界最低的。人口越多土地越大，交通運輸的需求就越多，當然溫室氣體的排放量就越大，可是交通運輸排放的增加，也可能是由於交通工具的選擇。所以，不能不明就理的就任意責怪自己，或盲目的相信別人的推論。

我們可以用二種方法檢測一個衡量指標是否合理。第一種方法是分析可以改進衡量指標的方法，並批判這些方法是否合理、可實現或者適得其反。第二種方法是找出測量結果有所改進但實際效果不佳的例子。當我們從理論上闡述一個指標可以得到改進，但結果卻惡化的方法時，那麼這個衡量指標就是有缺陷的，我們也會更好地理解衡量指標背後的真正含意。然而，當衡量指標只關注到了複雜系統的一方面，而忽略其他方面時，即使使用了正確的分母，所做的努力仍然可能導致適得其反的效果！