

frilly / December 10, 2019 04:45PM

[深入淺出，教你一步步實現數據分析入門、進階](#)

1. [初級階段：懂基本操作](#)
2. [中級階段：深入技術，做全方位的資料分析](#)
3. [高級階段：資料採擷](#)
4. [終極階段：從資料分析到商業決策](#)

最近有很多人在問資料分析的一些問題。關於資料分析到底應該怎麼學？如何快速入門，以及技術和業務之間的瓶頸如何突破？

其實，要學資料分析之前，至少要瞭解一下[資料分析師](#)

，到底需要哪些技能。有的同學看到資料分析幾個字，就馬上開始Python函數+控制語句、R語言和ggplot庫...上來一頓操作，還沒入門就放棄了。這就是需求不明確導致的，當然學習方式也值得商榷，那到底資料分析需要什麼樣的技能呢？

其實企業對資料分析師的基礎技能需求差別不大，可總結如下：

分析工具：一般要求R、Python、FineBI等分析工具至少會一種，會兩種以上加分，有的企業因內部需求，會指定的一種。

資料庫：絕大會要求會SQL，部分要求SQL、NoSQL會一種，高級的分析師或者大型企業要求能夠處理大數據，需要Hive（較少的需要Hadoop/Spark）。

統計學：若無相關專業背景，需要具備相應的統計學、概率論等基礎知識。

資料採擷：少部分要求會建模，瞭解基本的演算法模型，能夠做資料預測，即便不要求，演算法也是加分項。

結果輸出：Excel、PPT、Tableau、FineReport等專業BI工具。Excel和PPT要求的比較多，主要用作常規的資料呈現，與業務部門溝通等，[FineReport](#)

、Tableau等一般作為視覺化或者分析工具的加分項或者要求之一。

業務/思維：對某個領域（如電商、金融等）相關業務的瞭解或具有產品、運營方向的分析經驗，有自己的資料分析的方法論和專案經驗。

看上去很簡單，但其實你把每個技能拆分開來，都是一個不小的知識體系。如果我們按照資料分析的流程來細分的話，每個部分應該掌握的技能，大概是這樣的：

那對於這個技能體系，應該如何進行技能的訓練呢？先後順序是什麼？哪些地方可能出現困難和瓶頸？

這可能是初學者最困惑的地方，我們可以看看資料分析的不同層次：

1. 初級階段：懂基本操作

這個要求真的不高，如果是用Excel，也就是畫個分佈圖、用用場函數和Vlookup的事情，當然對於專業資料分析師而言，僅會Excel是遠遠不夠的，特別是對於擁有海量資料的巨頭。

用Python來做的話，也就是numpy、pandas和matplotlib的一些基本操作，簡單來說，你需要掌握基本的資料分析處理和[資料視覺化](#)，進行探索性的資料分析，觀察資料分佈、計算各種統計量，得出一些基本的結論。

當然你還需要補補基本的統計知識。

所以開始上手的話，就兩個部分，Python庫和統計學知識，當然你可以找一些資料集來進行這些操作。

在這個時候你已經可以做一些基本的資料清洗、資料分析工作了，像資料專員這樣的職位都可以嘗試，如果懂一些業務，初級資料分析的工作也可以著手了。

這個階段的工作主要進行基本資料處理，對於明確的分析任務能夠進行技術上的基本實現，能夠正常推進具體的分析計畫落地實施。

而處於這個階段，欠缺的能力在於技術的熟練度、全面性以及整體業務分析的規劃。

一方面是對於實際的專案缺乏多重技術的對比，對於複雜專案缺乏有效實施手段。；另一方面是對於整體的業務缺乏清晰的拆解方向，對於用什麼指標，需要構建什麼模型等問題的缺乏系統性思考。

2. 中級階段：深入技術，做全方位的資料分析

能夠做基本的操作之後，接下來你應該可以更深入下去了，掌握更多的技巧，以滿足不同的資料和場景。

所以這應該是一個連續的、循序漸進的過程。當然這個時候你更多的是要透過具體的資料集，去進行真實場景的分析，盡可能多的從不同的角度進行探索，得出更多的結論。

考慮到更多的場景，你可能還需要資料庫 ([SQL](#) or NoSQL) 的技能，至少能夠順利提取、查詢公司的資料庫，能夠通過資料庫的管理實現基本的資料清洗和分析。

還有就是爬蟲的技術，如果你有獲取網路資料的需求，你需要掌握另外的Python包 (比如BeautifulSoup、Requests 等)。當然這些可能並不是當下必須的技術，可以等到有需求的時候再去學習。

總的來說，這個部分，是資料分析工具深入和全面實戰的部分。確定你能夠熟練完成項目並輸出結果後，就完全可以去嘗試初級資料分析師相關的職位了。當然在此之前，你需要去大量進行專案練習，一來為熟悉流程和工具，二來培養分析思維，這些項目經驗也是求職的優質資源。

3. 高級階段：資料採擷

別慌，並不是要去真正地系統學習資料採擷。只是在做資料分析的時候，通常會涉及到預測型的資料分析，比如預測廣告的點擊、出行需要的時間、未來某個時間的銷售額.....這些對未來情況的預測是做決策的重要依據。

那其實就是要掌握一些資料採擷的方法，比如邏輯回歸、樸素貝葉斯方法、決策樹等等。當然你需要用到Python庫scikit-learn，scikit-learn裡都可以直接調用這些演算法，知道怎麼調用就可以了。

但隨著深入，你會有瞭解演算法原理的需求，那可以幫助你去調參，提升演算法效果。當然特徵選擇也是提升預測效果不可忽視的部分，如果演算法應用沒問題了，可以嘗試做特徵工程提升效果。

資料採擷技能 (或者說演算法能力) 並不是分析師的必修項目，但你去看招聘JD就會發現，這是一個很好的加分項，優秀的分析師應該具有一定的構建演算法模型的能力。

4. 終極階段：從資料分析到商業決策

老實說，這個步驟並不是最後來訓練的，而是貫穿在整個資料分析的過程中。因為資料分析的本質，或者說終極追求是為商業目的服務的，比如提供使用者增長的策略，為銷售提升提供解決辦法，通過資料分析提供更人性化的用戶體驗...

方法、技巧皆為工具，你將會總結出自己的一些方法論，如何透過一個完整的分析得出一個高價值的商業報告，甚至是一套自動化分析決策的商業系統（比如推薦系統，使用者畫像系統，風險模型）。

這些才是更高層次的資料分析工作應該去做的事情，在任何專案推進之前明確目標，並系統地解決資料獲取、資料存儲、資料建模和視覺化呈現的問題，推動整個專案的進度。

我們的終極目的不是去獲取幾個資料，也不是呈現幾張炫酷的圖表，而是去探索一個問題的解決方案，尋找一個市場的突破口，以技術的方式去降低策略制定的風險。

所以當你不再聚焦於具體的技術點、工作流程，而是重於不同業務指標的選取，解決問題的模型的構建，高效和低成本的決策輔助等時候，你就真正上升到比較高的層次了。

另外，如果你準備進入[大數據分析](#)領域，我強烈建議您使用這個資料分析報表工具[FineReport](#)。你可以點擊以下連結免費下載使用。初學者不用擔心，點擊這個學習計畫表頁面，它提供了從基礎到進階詳細的[學習計畫](#)，來幫助你快速入門！

[免費下載FineReport10.0](#)

相關文章：

[如何學習數據分析？十年數據分析經驗告訴你，看這一篇就足夠了](#)

[數據分析師需要掌握哪方面的電腦技能？——從基礎到進階](#)

[從事數據分析工作，如何找到適合的好企業？](#)
