

frilly / July 30, 2019 04:07PM

[數據分析工具怎麼選？十大諫言你值得一看！](#)

1. [從工具屬性和分析師需求來劃分](#)
2. [1.0.1 業務 or 技術](#)
3. [1.0.2 分析類工具](#)
4. [1.0.3 程式碼類工具](#)
5. [2. 從企業資料應用架構來劃分](#)
6. [2.0.1 資料存儲層](#)
7. [2.0.2 報表/BI層](#)
8. [2.0.3 資料分析層](#)
9. [2.0.4 表現層](#)
10. [2.0.5 關於視覺化](#)
11. [2.0.6 關於一些視覺化特效](#)
12. [3. 4大工具盤點](#)
13. [3.0.1 >>>>Excel](#)
14. [3.0.2 >>>>R](#)
15. [3.0.3 >>>>Python](#)
16. [3.0.4 >>>>BI](#)

對於[資料分析](#)

，我一直強調核心是業務，透過業務的分析邏輯影射到資料分析的處理邏輯，而資料分析工具則是說明我們實現結果的手段。

但是，你說工具不重要吧，他又很重要，就像什麼樣的路選擇什麼樣的交通工具，合適的工具能幫我們更快的達到終點。對應資料分析的不同環節，也要選擇不同的工具，甚至選擇更容易上手。今天這篇文章，就是來掃盲工具的。

估計網路上沒有比這個更全面的了

1. 從工具屬性和分析師需求來劃分

在企業中，資料分析師往往分為業務和技術兩類，兩者能力和工作內容有較大區別，對於工具的要求也各有側重。

1. 業務 or 技術

業務類分析師，往往在營運部，市場部，銷售部等，根據服務的業務部門的不同，可能叫資料營運，經營分析，會員分析，商業分析師等名字。因為各個業務線具體考慮的問題不同，分析思路與體系均有不同，所以會有這種區別。日常的工作更多是整理業務報表，針對特定業務做專題分析，圍繞業務增長做需要用到資料的測算、規劃、方案等。

技術類分析師，往往在IT部、資料中心。根據從事的工作環節不同，被分成資料庫工程師，ETL工程師，爬蟲工程師，演算法工程師等角色。在中小企業，往往一個技術小哥通吃這些流程。在大企業，一個標準的資料中心，一般都有數倉、專題分析、建模分析等組來完成資料開發工作，再大的公司，還有專門負責資料治理的小組。之所以有這個區分，是因為生產資料，需要一個多層次的複雜的資料系統。一個資料系統，需要資料獲取、資料內建、資料庫管理、資料演算法開發、報表設計幾個環節組合。這樣才能把分散在各處的一點一滴的資料集中起來，計算成常用的指標，展示成各種炫酷的圖表。這裡每一個環節都需要對應的技術支援和人員工作，因此有了不同的崗位。

大家在找資料分析崗時，一定要區分是技術還是業務，和自己的職業傾向是否匹配。

分析師有技術和業務之分，那對應工具也有這樣的屬性側重。

2. 分析類工具

對於初級[資料分析師](#)

，玩轉Excel是必須的，樞紐分析表和公式使用必須熟練，VBA是加分。另外，還要學會一個統計分析工具，SPSS作為入門是比較好的。

對於高級資料分析師，使用分析工具是核心能力，VBA基本必備，SPSS/SAS/R至少要熟練使用其中之一，其他分析工具（如Matlab）視情況而定。

對於資料採擷工程師……嗯，R和Python必備，要靠寫程式碼來解決。

3. 程式碼類工具

對於初級資料分析師，會寫SQL查詢，有需要的話寫寫Hadoop和Hive查詢，基本就OK了。

對於高級資料分析師，除了SQL以外，學習Python是很有必要的，用來獲取和處理資料都是事半功倍。當然其他程式設計語言也是可以的。

對於資料採擷工程師，Hadoop得熟悉，Python/Java/C++至少得熟悉一門，Shell得會用……總之程式設計語言絕對是資料採擷工程師的最核心能力。

一圖說明問題：

2. 從企業資料應用架構來劃分

工具的使用還要看企業的需求和環境。為什麼小企業招[數據分析師](#)

其實就是Excel做報表，大企業找資料分析是卻是把玩高大上的Python、R？這就要看企業的資料架構。

站在IT的角度，實際應用中可以把資料工具分為兩個維度：

第一維度：資料存儲層——資料包表層——資料分析層——資料展現層

第二維度：用戶級——部門級——企業級——BI級

1. 資料存儲層

資料存儲設計到資料庫的概念和[資料庫](#)

語言，這方面不一定要深鑽研，畢竟有專業的DBA。但至少理解資料的存儲方式，資料的基本結構和資料類型。SQL查詢語言必不可少，精通最好。可從常用的selece查詢，update修改，delete刪除，insert插入的基本結構和讀取入手。

Access這是最基本的個人資料庫，經常用於個人或部分基本的資料存儲；MySQL資料庫，這個對於部門級或者互聯網的資料庫應用是必要的，這個時候關鍵掌握資料庫的庫結構和SQL語言的資料查詢能力。SQL

Server2005或更高版本，對中小企業，一些大型企業也可以採用SQL

Server資料庫，其實這個時候本身除了資料存儲，也包括了資料包表和資料分析了。

DB2，Oracle資料庫都是大型資料庫，主要是企業級，特別是大型企業或者對資料海量存儲需求的就是必須的了，一般大型資料庫公司都提供非常好的資料整合應用平臺。

BI級別，實際上這個不是資料庫，而是建立在前面資料庫基礎上的，企業級應用的數倉。Data

Warehouse，建立在DW機上的資料存儲基本上都是商業智慧平臺，整合了各種資料分析，報表、分析和展現。

2. 報表/BI層

企業存儲了資料需要讀取，需要展現，[報表工具](#)

則是最普遍應用的工具，尤其是在國內。過去傳統報表大多解決的是展現問題，如今衍生了一些分析型報表工具，也會和其他應用交叉，做資料分析報表，透過介面開放功能、填報、決策報表功能，能夠做到打通資料的進出，涵蓋了早期商業智慧的功能。

像Tableau、PowerBI、FineReport、FineBI、Qlikview這類BI (商業智慧) 工具，涵蓋了報表、資料分析、視覺化等多層。底層還可於數倉銜接，構建OLAP分析模型。

3. 資料分析層

這個層其實有很多分析工具，當然我們最常用的就是Excel。

Excel軟體，首先版本越高越好用這是肯定的。當然對excel來講很多人只是掌握了5%Excel功能，Excel功能非常強大，甚至可以完成所有的統計分析工作！但是我也常說，有能力把Excel玩成統計工具不如專門學會統計軟體。

SPSS軟體：當前版本是18，名字也改成了PASW Statistics；我從3.0開始Dos環境下程式設計分析，到現在版本的變遷也可以看出SPSS社會科學統計套裝軟體的變化，從重視醫學、化學等開始越來越重視商業分析，現在已經成為了預測分析軟體。

SAS軟體：SAS相對SPSS其實功能更強大，SAS是平臺化的，EM挖掘模組平臺整合，相對來講，SAS比較難學些，但如果掌握了SAS會更有價值，比如離散選擇模型，抽樣問題，正交實驗設計等還是SAS比較好用，另外，SAS的學習材料比較多。

其他還有Python和R，後面還會詳細講。

4. 表現層

表現層也叫[資料視覺化](#)

，以上每種工具都幾乎提供了一點展現功能。但要說企業級最常應用的還是BI，做分析做報告。而我們的[FineReport](#)本身是一個通用的報表工具和資料視覺化工具。就好比Excel，小到可以存儲統計資料、製作各式各樣的資料圖表、dashboard，大到製作[財務報表](#)、開發進銷存系統。

- 它可以連結ERP、CRM、OA、MIS在內的各種業務系統資料，作為一個中間資料管理平臺。
- 它可以快速製作報表，搭建統一的資料分析和視覺化平臺。

就因為其強大的資料整合能力，再結合10多年來成熟的視覺化元件，finereport可以製作各類[資料視覺化](#)戰情室。

5. 關於視覺化

FineReport，擁有60多種圖表樣式，基本上涵蓋是市面上所有基礎的圖表，都是自主研發的HTML5圖表，具有優秀的動態效果和強大的交互體驗。使用時能夠根據需求設定各類特性，並且在行動端，LED大屏也能自我調整展示。

6. 關於一些視覺化特效

FineReport目前具有自動刷新、圖表切換、資料高亮顯示等特效，都是根據使用者的實際需求開發，除此之外，帆軟 (finereport的母公司) 背後有一批愛好的開發者，會利用視覺化開源庫，設計開發視覺化外掛程式，專門為finereport客製，目前這一生態已十分成熟。

需要說明的是，這樣的分類並不是區分軟體，只是想說明軟體的應用。有時候我們把資料庫就用來進行報表分析，有時候報表就是分析，有時候分析就是展現；當然有時候展現就是分析，分析也是報表，報表就是資料存儲了！

3. 4大工具盤點

以上囉嗦了那麼多，具體講講Excel、R、Python、BI吧。

1. >>>>Excel

適用場景：

1. 一般的辦公需求下的資料處理工作；
2. 中小公司資料管理，存儲（很多國有企業都用）；
3. 學校學生，老師做簡單的統計分析（如方差分析，回歸分析）；
4. 結合Word，PowerPoint製作資料分析報告；
5. 資料分析師的主力分析工具（部分資料分析師的輔助工具）；
6. 部分商業雜誌，報刊圖表製作（資料視覺化）；

優點：

1. 容易上手；
2. 學習資源十分豐富；
3. 可以用Excel做很多事情，建模，視覺化，報表，動態圖表；
4. 幫助你在進一步學習其它工具之前（比如Python，R），理解很多操作的含義；

缺點：

1. 深入學習需要掌握VBA，難度還是很高；
2. 當資料量較大時，會出現卡頓的情況；
3. 到Excel2016版，在不借助其它工具的情況下，Excel資料檔案本身能夠容納的資料僅有108萬行，不適合處理大規模資料集；
4. 內置統計分析種類太簡單，實用價值不大；
5. 不像Python，R語言等開源軟體，正版Excel需要付費，比如我用office365.每年需要支付300多塊錢（不過也值了）

2. >>>>R

使用場景：

透過擴展的協力廠商R包，R能夠做的事情幾乎涵蓋了任何需要資料的領域。就我們一般的資料分析或者學術資料分析工作而言，R能做的事情包括但不限於如下方面：

1. 資料清洗與整理；
2. 網路爬蟲；
3. 數據視覺化；
4. 統計假設檢驗（t檢驗，方差分析，卡方檢驗等）；
5. 統計建模（線性回歸，邏輯回歸，樹模型，神經網路等）；
6. 資料分析報告輸出（Rmarkdown）；

R容易學嗎？

從我個人來看，想要入門R是非常簡單的，10天的集中學習，對於掌握R的基本使用，基本資料結構，資料導入匯出，簡單的資料視覺化，是完全沒有問題的。有了這些基礎，在遇到實際的問題時，去找到需要使用的R包，通過閱讀R的說明文檔，以及網路上的資料，就能夠相對快速的解決具體問題了。

3. >>>>Python

R語言和Python同為需要程式設計的資料分析工具，所不同的是，R專門用於資料分析領域，而科學計算與資料分析

只是Python的一個應用分支，Python還可以用來開發web頁面，開發遊戲，做系統的後端開發，以及運維工作。

現在的一個趨勢是，Python在資料分析領域正在追趕R，在某些方面已經超越了R，比如機器學習，文本挖掘等程式設計的領域，但R語言在偏統計的領域仍然保持優勢。Python在資料分析方面的發展，很多地方借鑒了R語言中的一些特色。所以，如果你現在還是一片空白，還沒開始學習，要做決定學習R還是Python的話，建議從Python入手。

Python和R都比較容易學習，但是如果你同時學習兩者，由於在很多地方它們非常相似，就會很容易混淆，所以建議不要同時學習它們。等其中一個掌握到一定的程度，再著手學習另外一個。

Python能做什麼？

1. 網路資料爬取，使用Python能夠很容易的編寫強大的爬蟲，抓取網路資料；
2. 數據清洗；
3. 數據建模；
4. 根據業務場景和實際問題構造資料分析演算法；
5. 資料視覺化（個人感覺不如R好用）；
6. 機器學習，文本挖掘等高級資料採擷與分析領域；

應該學習R還是Python？

如果因為時間有限，只能選擇其中的一種來學習的話，我建議使用Python。但我仍然建議兩者都瞭解一下，畢竟每個人都不一樣。可能你在某些地方聽說，Python在工作中更加常用，但是工作中，解決問題才是最重要的，如果你能夠用R高效的解決問題，那就用R。實際上，Python很多資料分析方面的特色，是模仿R來實現的，比如pandas的資料框，正在開發中的ggplot視覺化包模仿的是R語言中非常著名的ggplot2。

4. >>>>BI

多數分析師日常的工作就是做報表，而資料分析師更多用到的報表是BI。

BI全稱商業智慧，在傳統企業中，它是一套完整的解決方案。將企業的資料有效整合，快速**報表製作**以作出決策。涉及資料倉庫，ETL，OLAP，許可權控制等模組。

BI工具主要有兩種用途。一種是利用BI製作自動化報表，資料類工作每天都會接觸大量資料，並且需要整理匯總，這是一塊很大的工作量。這部分工作可以交給BI自動化完成，從資料規整、建模到下載。

另外一種是使用其視覺化功能進行分析，BI的優點在於它提供比Excel更豐富的視覺化功能，操作簡單上手，而且美觀，如果大家每天作圖需要兩小時，BI會縮短一半時間。

BI作為企業級應用，可以通過它連接公司資料庫，實現企業級報表的製作。這塊涉及資料架構，就不深入講了。

這裡我以2019年非常流行的BI工具FineReport為例。它有兩個主要用途：

一種是自動生成報告。資料分析師每天都要接觸大量的資料。資料需要整理和匯總，這是一個很大的工作量。這部分工作可以移交給FineReport。它自動地對資料進行整形、建模和下載。二是利用其視覺化功能進行分析。FineReport的優點是它提供了比Excel更豐富的視覺化功能。而且它很容易使用。如果你每天花兩個小時做報表，FineReport會把它縮短一半。就我個人而言，在學習資料分析的初始階段，BI工具無疑是最容易學習的。

如果您準備進入資料分析領域，我強烈建議您使用這個資料分析工具FineReport。你可以點擊這裡**免費下載**使用。它的官方網站也提供了教程來幫助你快速入門。

[免費試用FineReport >](#)

獲得帆軟最新動態：數據分析，報表實例，專業的人都在這裡！加入[FineReport臉書粉絲團](#)！

相關文章：

[零基礎快速自學SQL，1天足矣！](#)

[30個值得推薦的資料視覺化工具，趕緊收藏！](#)

[書單 | 從入門到精通，數據分析不得不看的10本「好書」！](#)
