

## 論自學-以行銷資料科學為例

上大學後，如果將時間劃分為3等分，

1. 1/3的時間，我們在睡覺；
2. 1/3的時間我們在打屁聊天、參與不同的大學活動；
3. 最後1/3的時間我們則是在學習產、銷、人、發、財及通識課程(企管系為例)；

如此的時間分配，造成大學生普遍「樣樣皆通，但樣樣不精」之窘況，更遑論自動自發真正的學習一件有用的新事物。

然而，反觀比爾蓋茲、祖克伯、唐鳳等名人，皆是自學成材之典範，那我們到底如何跳脫時間的桎梏，讓自學速達速成？如何越過重重學習的藩籬，徹底將我們生鏽的頭腦磨成光亮的寶劍？

以下便由小弟從自學0到1、自學入門、自學訣竅及進階自學，分享幾則自學心法

### 一、自學0到1-針對自身主修科目，加以擴展基礎，多方面加強自身科系

還記得我在2012年於澳洲Middle Ridge小學當文化志工的時候，看到他們的校訓就是「Where seeds of learning become forests of opportunity.」這句話的意思就是「有地方可讓種子成長茁壯(學習)，就有機會可以變成壯大的森林」，他們用這句話表示學習的重要性。

這同樣可類比到大學學習，千萬不要指望有老師可以「扶著我們的手寫字」，完完整整的將他畢生所學傳授給我們，大學的課程僅是帶你入這學門知識的一個入門而已，剩下的，便要由我們自己去發揮。舉我自己來說，企管系的孟彥老師引領我進入行銷這學門(Seeds)，而透過與不同老師及學長姐的交談下，我覺知到如果能將「資料科學」+「行銷」，衍伸為「行銷資料科學」，方能在身處大資料時代的浪頭上，擁一立足之地。再進一步去查詢台灣現階段有這兩種領域交集的人才、學者實在不多，如自身能擴大兩種領域的交集，便有創造自身未來機會(Forests)之可能性。

但是千不熟、萬不熟資料科學的我，該如何自學資料科學？

### 二、自學入門-領航員到底在哪裡？

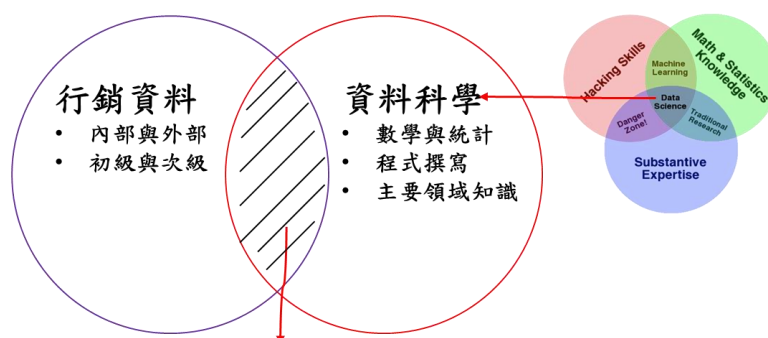
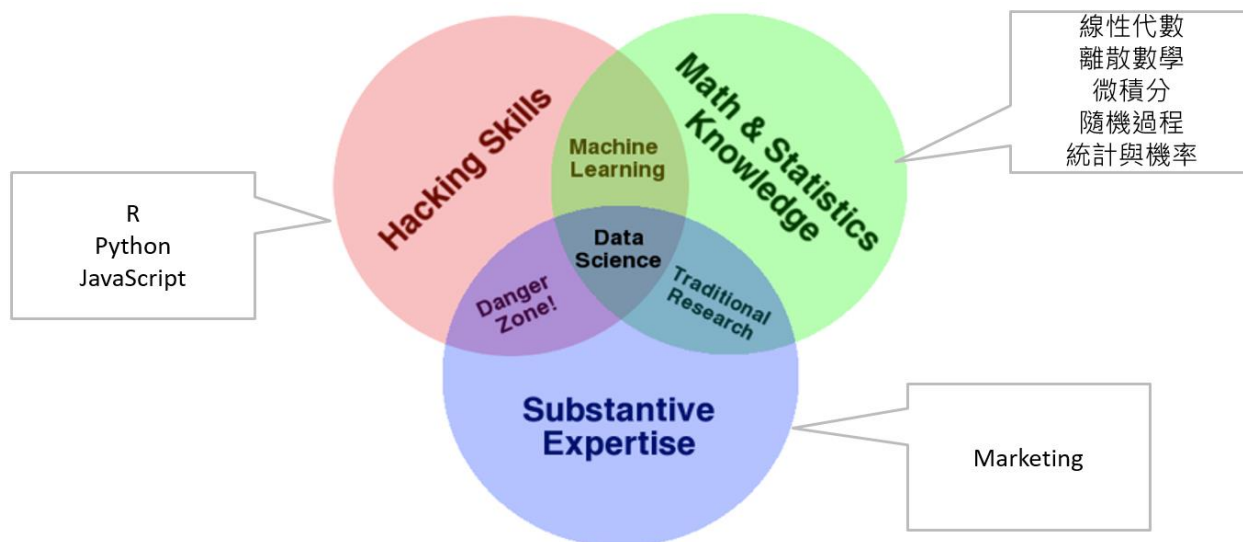
自學的第一步，就是要找到要學習科目的領航員，它可以是我們的老師、朋友，但有一種「遠在千里，卻近在咫尺」的領航員，更能讓我們方便找尋到自學入門的關鍵，那就是-「網路」。

我透過網路，找尋到如下資源：

1. 人人都可成為資料科學大師！一整年的網路自學清單就在這了  
<https://buzzorange.com/techorange/2016/02/02/plan-to-be-a-data-scientist-in-new-year/>
2. 如何自學資料科學？21個優質課程資源幫你入門

...

等資源都可以協助我做進一步的延伸，如果信不過這些網路資源，便去帶著這些資源去詢問這領域的老師，經過幾次的詢問後，我們應該可以得出類似下面的交集圖



## 【行銷資料科學】

是一門：

「藉由資料科學之新興工具、設備進行大量、及時且多樣的行銷數據蒐集與分析，並依不同資料屬性建立預測性模型，增益行銷決策精準度的科學」

如此，我們於自學上有了頭緒，在程式學習上，我應該要精通R與python；數學知識上，我要懂統計機率、隨機過程(馬可夫鏈)、微積分、離散數學、線性代數；主領域知識上，應該具備主要科目，行銷學。

對該領域有了初步的了解後，我們便要開始進行真正的自學流程。

### 三、自學訣竅-海盜精神

自學的訣竅在於「資訊的流通」+「線上結合線下學習」+「熱忱」+「接案」+「不怕挫折的精神」

### (一)資訊的流通-資訊不對稱，乃是利益之所在

相信很多人都不知道Udemy、Udacity、Hahow、Coursera等MOOC平台，這就是會自學與不會自學的人之差別，如果連自學管道皆不知曉，便遑論能真正「結構化的自學」。所謂結構化自學便是透過有系統的方式，快速自學(通常在1-2禮拜內便要能知其然，4-6禮拜知其所以然)。有幸我們今生在網路無遠弗屆的時代，如果懂得利用這些網路線上結構化課程資源，便能在短時間內，學習所需知識。雖然許多線上課程均須付費，但因為線課競爭激烈，所以造成價格低、品質佳的狀況，這正是大學生可以多多利用資源阿~!

以下列出「邁向資料科學家」的線上學習資源(Udemy課程全部優惠300-450元台幣)

1. 書籍\_程式撰寫能力
  - [01\\_巨量資料蒐集\\_Text Mining with R](#)
  - [02\\_巨量資料蒐集\\_mining the social web](#)
2. 線上課程\_統計與數學
  - [01\\_MIT\\_線性代數](#)
  - [02\\_清大\\_微積分一](#)
  - [03\\_清大\\_微積分二](#)
  - [04\\_清大\\_統計](#)
3. 線上課程\_基本程式撰寫能力+統計與數學
  - [01\\_Udemy\\_Machine Learning A-Z™](#)
  - [02\\_Udemy\\_Data Science and Machine Learning Bootcamp with R](#)
  - [03\\_Udemy\\_Python for Data Science and Machine Learning Bootcamp](#)
4. 線上課程\_中階程式撰寫能力+統計與數學
  - [01\\_Udemy\\_線性迴歸](#)
  - [02\\_Udemy\\_邏輯思迴歸](#)
  - [03\\_Udemy\\_深度學習](#)
5. 線上課程\_高階程式撰寫能力+統計與數學
  - [01\\_Udemy\\_CNN 圖像化應用](#)
  - [02\\_Udemy\\_非監督式學習](#)
  - [03\\_Udemy\\_RNN 語意應用](#)
  - [04\\_Udemy\\_NLP](#)
6. 爬蟲\_程式撰寫能力
  - [01\\_hahow\\_python爬蟲](#)

這時候，另一個問題又來了，該如何找到這些資源? 這是從二、自學入門的諸多網站中，挑選線上課程平台，在挑選課程內容的結果。如是說，自學的訣竅並不是最重要的，最猶要反而是「一、自學0到1」，真正找到自己的方向後，才能往第二步第三步前行，更直白來說，沒有seeds，哪來forests?

### (二)線上結合線下學習-台科大位處台灣實體學習資源最豐富之帶

身處台灣大學聯盟的台科大，方圓五公里內，擁有全台灣最好的教育資源(而且旁聽全部都是免費)，如果不好好利用，未免也太枉費台科大學生的身分了。

以下列出「邁向資料科學家」的實體學習資源

1. 實體課程\_主領域知識

[01\\_台大\\_大數據與商業分析](#)

[02\\_台大\\_大數據行銷](#)

2. 實體課程\_高階程式撰寫能力+統計與數學

[02\\_台大\\_機器學習](#)

[02\\_台大\\_深度學習](#)

其中不少課程居然有錄影(如:[機器學習](#))。

這時有人又會問，都有網路線上課程了，為什麼要上實體課程?!

第一，我們要好好利用在校身身分

第二，實體課程可扎基礎，增加學理知識

第三，最重要的，線上結合線下實體課程；因為實體課程大多以學理為重，實作時間極少，如果能將線上的實作課程結合學理知識，便能付諸實行所謂「理論應用到實務」的概念。

### (三)接案+熱忱-一把槍的威力

當「一把槍指著我們的頭」，我們是否都會依照拿槍者的指令去執行任何事項? 接案便是能夠發揮這種效用的自學絕招，因為當我們接案後，必須要交出可交付成果時，我們都必須要長出三頭六臂，生出成果(可以回想大學硬課要我們做的報告XD，對! 就是那種感覺!)。這時候的學習效率、速度、成效都是任何一堂大學課程與線上課程難以披靡的(不過壓力真的會很大就是)。

但是，如何接案? 最簡單的方法便是找「親戚」、「友人」或「大學老師」等關係人，來執行他們與自己自學領域相關的專案。我便是找大學老師(孟彥、凱揚老師)一同做案子，增廣經歷。而我們始接案之時，千萬謹記，對方沒有開價，就不要索價，說白了，先免費幫關係人做案子，因為剛開始時，我們或許對自學領域上不熟悉，所以就算做砸了，也不會有太大的人情壓力，待我們對該自學領域開始熟習時、關係人對我們有所信任時，他不給我們一些「補貼」也難過意得去。

最後，這一把槍的威力，可以帶領我們幫身旁的關係人進行專案外，慢慢的，他便可以為我們接下不少其他經費更多的案子，其中滋味，等著大家親自去體會。

## 四、進階自學-無招勝有招

當我們對自學領域的實力緩緩上升到實體課程或者線上課程的內容已經無法添益的地步時，我們便要想辦法當那「無招勝有招」的創新學習者，也就是透過網路社群快速的獲得即時新知，當然，還沒有到達本地步的同學，同樣也可以採用進階自學，讓自己成長的更快! 而且透過擁有這些進階自學資源，能夠變相的再擁有更多自學資訊!

以下分享我「邁向資料科學家」的進階自學網路資源

## 1. 國內外相關社群平台(FB)的資訊

[01 國內平台 Data Man 的資料視覺化筆記](#)

[02 國內平台 台灣資料科學年會 \(爆推，多參加他的講座\)](#)

[03 國內平台 大鼻觀點-統計與資料科學 \(爆推，犀利統計觀點\)](#)

[04 國內平台 Python Taiwan \(爆推\)](#)

[05 國外平台 R bloggers \(爆推，資訊即時\)](#)

[06 國外平台 R Computing](#)

[07 國外平台 Learning Python Programming](#)

## 2. Kaggle

[Kaggle](#)裡面有許多Datasets及competitions，幾乎都具有Kernels，也就是資料科學家比賽的codes都會公布在上面，只要能將他們幾個厲害的資料科學家的kernels看懂，並且充分利用他們的codes，便能解決現實中資料科學的許多問題。

## 五、結語

Joachim Baan 曾說:Experts are made, not born，當我們再羨慕唐鳳等成功人士時，細觀他們成功的背後，都充滿著自學的熱忱與不畏挫折的勇氣。所以我們在自學時，每每遇挫折與失敗時，我們都應從失敗中爬起來，不斷進步。最終，如同R blogger這篇文章所說[The one thing you need to master data science is practice](#)，唯有不斷的「Practice」方能成就那真正自學的我們。

加油! 皓軒 筆