

# 一份来自清华的「数据分析」学习资料，拿走不谢！

学习python的正确姿势 今天

之前发过很多关于数据分析的文章，收到大家不少的好评，不过也有一些困惑：

- 入门数据分析该学哪些知识点？该看哪些书？
- 是从Python入手还是R语言？常用的算法有哪些？
- 可以练手做项目的数据库去哪里找？好用的爬虫工具又有哪些？
- 网上看了很多文章，但感觉没有个系统，号主有系统性的资料推荐吗？

的确，我当时学数据分析也有同样的苦恼，网上关于数据分析的学习资料非常多，但缺乏系统性，没有老师时时给你解答问题，很多时候都感觉自己要放弃了。

但，我不得不说一句：**学数据分析绝对没错，坚持下去你会看到一个完全不一样的自己。**

之前在百度的曹政举了他身边的例子，令人印象深刻，蛮多看上去并不优秀的人却都能靠着**数据分析C位出道**。

十来年前我在百度招聘过一个大本科应届的小伙子邓明生，从学历背景看在百度并不占优势，当时开始跟我做数据分析，写程序分析百度的业务数据，后来慢慢独挡一面，因为对百度所有业务线的数据都清晰，后来百度出现一些人事危机的时候开始成为救火队长，连续在不同业务部门担纲重要职位，一路升到联盟事业部总经理，今年离职出来创办御势资本，青出于蓝而胜于蓝，人家现在比我厉害很多了。

还是十来年前，又有一个应届生吴海生，从百度产品部门申请内部调动去做数据分析，开始经验不足，写报告还被我嘲讽过的那种。好多年不见，最近看新闻才知道，已经某新近上市的金融公司CEO，妥妥的C位出道有没有，真是让人刮目相看。

那么，数据分析到底该怎么学呢？

其实真的没那么复杂，你只需要做好这三件事：

- 1.找到一个实力与经验俱佳的“教练”，**从思维、工具、实战带你“即学即用”**
- 2.制定一份**正确的学习计划与路径**，你真正需要的是好方法而不是蛮力
- 3.**有效工具的运用**会让你事半功倍

在这里，不得不给大家介绍下蛮不多的数据分析学习资料 —— 极客时间的**《数据分析实战45讲》**。作者是清华大学计算机系博士陈旸，最近刚刚更新完毕，好评度也超高。

在这个专栏中，陈旸清晰地把数据分析拆解成下面三个组成部分：**数据采集、数据可视化和数据挖掘**。在后面文章里，我会给大家分享这三部分所需要掌握的知识，让你有个更深入了解。

专栏里一直秉承着“**MAS学习法**“，即 Multi-Dimension（多维度认识）、Ask（提问）和Sharing（分享），**从“思维”到“工具”再到“实践”**，学以致用，更高效上手数据分析。而且老师**还会直接提供项目数据，让你上手练习，可以在简历上完善项目经历，顺利找到工作。**（**有兴趣的可以直接拉到文章末尾，享受粉丝专属优惠**）

练手的数据项目

好，下面接着给大家分享上图中**数据采集、数据可视化和数据挖掘**需要掌握的知识。

## 数据采集

当你入门Python后，接下来就算是正式进入数据处理阶段。「数据分析」涵盖两部分：**数据是基础，分析是过程**，所以数据的前

期准备工作也很重要。第一步，就是采集数据。

你可以用Python自动采集数据，也可以使用第三方平台，比如用八爪鱼来采集数据。《数据分析实战45讲》中，作者陈咏用了两个实战案例来讲解如何用Python和八爪鱼来采集数据，讲的非常细致，看完你可以掌握这两种常用方法。

老师制作的「Python爬虫总结图」

详细地，你可以去直接看《数据分析实战45讲》专栏里这两篇文章：

- 第9讲 | 如何用八爪鱼采集微博上的“D&G”评论？
- 第10讲 | 如何用Python自动化下载王祖贤海报？

## 数据可视化

大多数人都很容易被数据可视化吸引。试想一下，用各种酷炫的图片将数据的规律直观地呈现给大家，是一件特别有成就感的事情。应用也很广泛，比如天猫双十一的数据大屏等。我们可以用各种工具、编程语言做数据可视化，比如DataV、Tableau、Python或者R语言。

在《数据分析实战45讲》中，主要用Python的 Matplotlib 工具来做数据可视化。Matplotlib 是Python的可视化基础库，非常适合入门学习。学完专栏，下面的这几张图我也可以做出来，非常抢眼。

你可以看看《数据分析实战45讲》专栏里这篇文章：

- 第15讲 | 如何用Python绘制10种常见的可视化视图？

## 数据挖掘

当你掌握了数据分析中基础的操作后，接下来就该正式处理数据了。为了进行数据挖掘任务，数据科学家们提出了各种算法，《数据分析实战45讲》中详细讲解了数据挖掘十大经典算法，根据用途，把它们分为四大类：

- 分类算法：C4.5，朴素贝叶斯（Naive Bayes），SVM，KNN，Adaboost，CART
- 聚类算法：K-Means，EM
- 关联分析：Apriori
- 连接分析：PageRank

值得一提的是，专栏里用了大篇幅内容、许多案例来讲解这十大算法，还会提供一些数据库让大家去实操，亲测有效。

此课程，享受极客时间 11.11 优惠，《数据分析实战45讲》超值参团只需 ¥79，原价 ¥99，立省 ¥20。想深入进阶 Python 的同学，别错过。

再加上我的**粉丝专属优惠口令**  
**支付时输入：handsomeb**  
还能**再减¥20**  
到手价**只要¥59（基本5折了）**

在留言区，你能看到大家各种各样的**解题思路**，其中有的你可能会意想不到，可以说，在留言区你也能学到很多。作者也会留言回复，**解答大家提出的问题，或者给予及时反馈。**

《数据分析实战45讲》超值特惠，优惠口令参团只需 **¥59**，原价 **¥99**，**基本半折入手**！数据分析能力必然是每个互联网人必须具备的，无论是运营、产品还是程序员，之后还可以往数据分析师、数据挖掘工程师等方向发展。

扫码订阅

粉丝专属福利

只需 **¥59**，拿下数据分析！

因为有了这个全场通用的**¥20元优惠口令**，买极客时间的课程真的很划算！！再给大家推荐几个Python相关的课程，口碑很不错！

我的**粉丝专属优惠口令**

**支付时输入：handsomeb**

全场还能**再减¥20**

## 《零基础学Python》

基于最新 Python 3，通过实际场景来学习，全面讲解标准库、第三方库和机器学习库。另有一个大模块专门讲解爬虫技能，是一个入门 Python 必看的视频课程。现已有8600+用户订阅，跟着这个课程学，从零基础到入门绝对没问题。

「**handsomeb**」口令后优惠价 **¥69**，原价 **¥199**

（很划算！！）

## 《Python核心技术与实战》

5 大模块，45 讲，Facebook 资深工程师景霄带你从工程角度，系统掌握 Python，进阶 Python 高手。已经有**近 1.1W 人**加入学习，评价都挺不错的。而且景霄**还会直接带你手把手完成一个用 Python 搭建的交易系统，让你上手练习，即学即用。**

除此之外，每篇文章后，作者都会留下一个**思考题**，帮助大家**更好吸收知识**。

举个例子，在「**第 4 篇 | 字典、集合，你真的了解吗？**」中，作者讲解了 Python 最常见的 2 种数据结构，留下了这道思考题：

「**handsomeb**」口令后优惠价 **¥59**，原价 **¥99**

## 《数据结构与算法之美》

订阅量Top1，60000+程序员的算法课堂，整个专栏会涵盖100 多个算法真实项目场景案例，更难得的是它跟市面上晦涩的算法书籍不同的是，还手绘了一些清晰易懂的详解图（总共有 300 多张），市面上的大多数的算法教程都看过，走心的说，这个专

栏是市面上唯一一门真正适用于工程师的专栏，作者是前Google工程师王争，相信会开启你的趣味学习算法之旅。

手绘图一出自《数据结构与算法之美》

「handsomeb」口令后优惠价 ¥59，原价 ¥99

《人工智能基础课》

王天一副教授从数学基础开始，层层深入，覆盖机器学习、人工神经网络、深度学习等最新领域，给你呈现学习人工智能的最佳路径。

「handsomeb」口令后优惠价 ¥35，原价 ¥68

以上课程都想要怎么办？开通极客时间VIP年卡会员，解锁全部课程。100+一线技术专家的实战秘籍，365天畅学。今日开卡立省¥1000，还有爆款程序员周边等加码福利等候。扫描下方二维码了解详情。

最后，大家别忘了  
满¥40减¥20优惠口令：  
handsomeb

80多门课程优惠，「handsomeb」再减20，戳此进入

阅读原文

