

Python⁺

51CTO学院微职位

2019 数据分析工程师

学习
模块

重点
项目

经典
案例

必备
技能

掌握最黑技术 占领更大市场

跟随知名企业优秀工程师、行业知名专家，系统全面的学习并掌握工作方法和技能技巧，获得全新的技能提升！

目录

CONTENTS

01

微职位简介

02

核心优势课程特色

03

课程介绍

04

应用领域

05

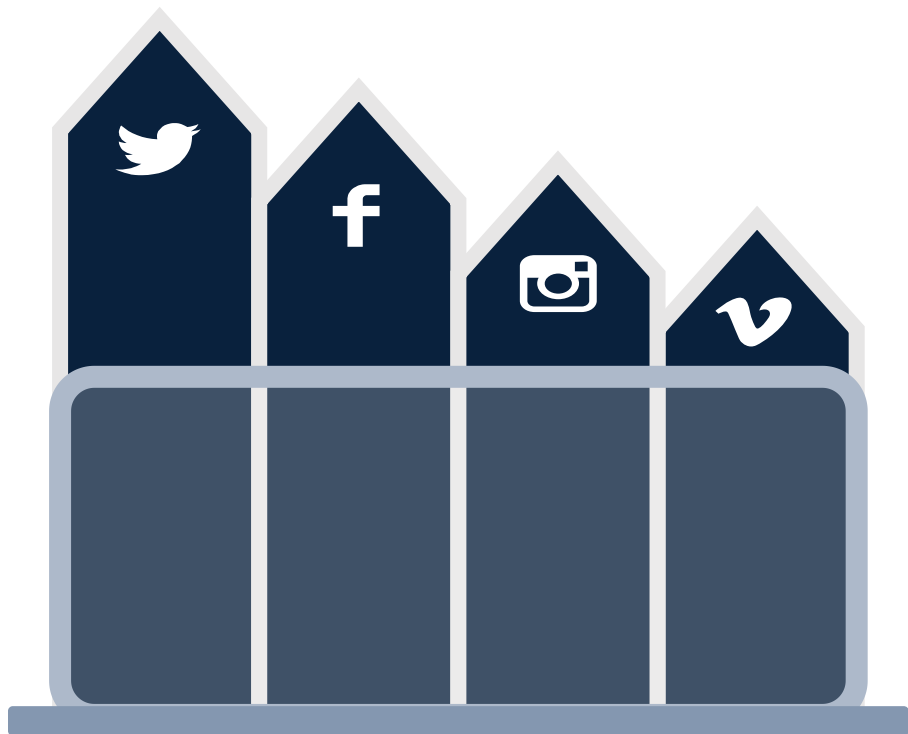
职业发展

06

教学任务



微职位简介



跟随知名企业优秀工程师、行业知名专家，系统全面的学习并掌握工作方法和技能技巧，获得全新的技能提升！

本微职位课程基于Python3.6版本，囊括Python基础+数据库+爬虫开发+数据分析+数学统计+数据模式+数据可视化+实战案例八大重点内容，全面掌握Python爬虫与数据分析工程师的实战技能。



核心优势课程特色



Python数据分析工程师课程专门为零基础学员定制，通过本课程学习您将掌握具备以下能力：

2. 通过Python爬虫学习与练习具备爬虫开发能力；

4. 理论强化：掌握统计学，数据分析等基础理论知识；

1. 掌握Python基本开发能力，为今后发展打下坚实基础；

3. 掌握常用的数据清洗，整理与可视化，具备一定数据分析能力；

5. 掌握常用的数据分析模型，通过实际项目掌握备数据分析相关知识点，具备应对数据分析工作能力；

通过以上学习，可以提升自己，满足自己工作或者换工作需求，为职业生涯发展打下良好基础。



课程名称	章	课程知识点	预期达到的目标
第一模块： 了解课程和学习计 算机基础	课程目标	1. 你为什么学编程	掌握技能： 1. 了解课程体系，确认学 习目标 2. 理解软件开发思想 3. 掌握常见编程工具的使 用 4. 掌握开发必备的计算机 基础技能
		2. 本课程的目标	
	课程简介与学习指导	1.课程的特点	
		2.anaconda安装问题解决方案	
		3.课程教学体系	
		4.学习最佳实践	
	计算机基础	1.计算机基础1	
		2.计算机基础2	
		3.计算机基础3	
	打造编程环境	1. 快速搭建Python环境	
		2.【选修】打造编程环境	
		3.【选修】打造编程环境4-vscode与python-1	
		4.【选修】打造编程环境4-vscode与python-2	
		5.打造编程环境5-精讲jupyter	
		6.【选修】在Linux安装和试用jupyter	





课程名称	章	课程知识点	预期达到的目标
第二模块： Python基础：从第一行 代码到掌握Python	成长型项目：51备忘录 v0.21	1. 导学-Python基础	掌握技能： 1. 初建项目概念，了解一个项目从需求到立项，到代码落地的过程 2. 掌握Python的基本数据类型和常见操作 3. 掌握Python的控制语句并灵活运用 4. 掌握函数的相关概念和用法，能用函数熟练编写简单项目 5. 掌握正则的基础知识和各种匹配的技巧，从数据中提取目标资源 6. 掌握面向对象思想，能够用面向对象思维构建项目
		2. 为什么叫成长型项目	
		3. Python程序的运行过程	
		4. Python脚本的常用结构	
		5. 伴你终身的代码规范	
		5. 基本数据类型：int、str、bool	
	字符串的增删改查	6. 基本数据结构list	
		1. 字符串的增删改查	
		2. 判断是否数字或字符串	
		3. 案例: 打印乘法表	
	基本数据结构dict、set tuple	4. 案例：token生成器	
		1. 案例: 翻译小程序	
		2. 案例: 单位转换器	
	控制程序的每个角落	3. 项目：51备忘录v0.24	
		1. 检查代码的真与假	
		2. 用if控制代码分支	
		3. 用for进行代码循环	
		4. 用while进行代码循环	
		5. 案例：21点小游戏	
	了解函数的来龙去脉	6. 项目：51备忘录v0.25	
		1. 为什么需要函数	
		2. KV编程论：解释函数	
		3. 几种参数的用法	
		4. 你了解作用域吗？	
		5. 程序出了异常怎么办？	
		6. 案例：登录验证命令行版	
		7. 案例：密码生成器	
		8. 项目：51备忘录v0.26	



学习安排——第3周

课程名称	章	课程知识点	预期达到目标
第二模块： Python基础：从第一 行代码到掌握Python	强大的正则表达式	1. KV编程论：各种模式匹配都只是代数	<p>掌握技能：</p> <ol style="list-style-type: none">1. 初建项目概念，了解一个项目从需求到立项，到代码落地的过程2. 掌握Python的基本数据类型和常见操作3. 掌握Python的控制语句并灵活运用4. 掌握函数的相关概念和用法，能用函数熟练编写简单项目5. 掌握正则的基础知识和各种匹配的技巧，从数据中提取目标资源6. 掌握面向对象思想，能够用面向对象思维构建项目
		2. 匹配一次还是多次	
		3. 如果想分组匹配怎么办	
		4. 匹配要不要贪心一点	
		5. 正则匹配之内容替换	
		6. 案例：找找她的联系方式	
		7. 案例：登录验证正则版	
		8. 项目：51备忘录v0.27	
	面向对象大法好	1. KV编程论：类与对象	
		2. 如何定义一个类	
		3. 方法中的self是怎么回事	
		4. 类里面的特殊方法	
		5. 面向对象的三大特征	
		6. 解析元编程	
		7. 7种可调用对象	
		8. 反射都能用来干什么？	
		9. 案例：汽车工厂	
		10. 案例：简易聊天机器人	
		11. 项目：51备忘录v0.28	



课程名称	章	课程知识点	预期达到的目标
第三模块： Python进阶：让工作 开始自动化	管理目录与文件	1. 探索整个目录树	掌握技能： 能够熟练使用Python编写自动化工具， 解决工作和生活问题，包括： 1. 管理目录与文件 2. 处理任何格式的文本 3. 熟练操作各种办公软件 4. 用日志记录各种操作 5. 管理时间常见技巧 6. 学会使用API和JSON获取数据 7. 用Python发邮件 8. 处理各种图像 9. 编写爬虫 10. 使用各种高级函数和特性优雅的解 决问题
		2. 文件的命名管理	
		3. 管理目录，文件属性与后缀名	
		4. 怎么压缩备份多个文件	
		5. 案例：代码备份机	
	处理任意格式的文本文件	1. KV编程论：解析文本文件	
		2. 案例：日志分析器	
		3. 项目：51备忘录v0.32	
	熟练操作各种办公文件	1. word文档快速编辑	
		2. excel文件读写套路	
		3. 用Python读写pdf文件	
		4. 案例：用代码写word情书	
		5. 案例：表格与文本转换器	
		6. 项目：51备忘录v0.33	
	给程序加点日志	1. 要日志有什么用？	
		2. 日志有哪些级别	
		3. 日志的输出控制	
		4. 案例：交通信号灯	
		5. 项目：51备忘录v0.34	



课程名称	章	课程知识点	预期达到的目标
第三模块： Python进阶：让工作开始自动化	自由管理时间	1. 工程中那些地方需要时间？	掌握技能： 能够熟练使用Python编写自动化工具，解决工作和生活问题，包括： 1. 管理目录与文件 2. 处理任何格式的文本 3. 熟练操作各种办公软件 4. 用日志记录各种操作 5. 管理时间常见技巧 6. 学会使用API和JSON获取数据 7. 用Python发邮件 8. 处理各种图像 9. 编写爬虫 10. 使用各种高级函数和特性优雅的解决问题
		2. 常用时间模块有哪些？	
		3. 案例：2B闹钟	
		4. 案例：时间转换器	
		5. 项目：51备忘录v0.35	
	JSON和API会是你的好伙伴	1.JSON其实还是文本	
		2. API会对你造成多大影响？	
		3. 案例：定制天气预报	
		4. 案例：博客词频统计	
		5. 项目：51备忘录v0.36	
	怎么用Python发邮件	1. 邮件的几个要素	
		2. 来吧！给我发个邮件	
		3. 案例：会员注册提醒	
		4. 项目：51备忘录v0.37	
	用Python来PS照片？	1. 图像的基础知识	
		2. 用Python玩转图像	
		3. 案例：mini照片管理系统	
		4. 项目：51备忘录v0.38	



课程名称	章	课程知识点	预期达到的目标
第三模块： Python进阶：让工作开始自动化	是时候做个小爬虫了	1. 8哥警告你，要守法！	掌握技能： 能够熟练使用Python编写自动化工具，解决工作和生活问题，包括： 1. 管理目录与文件 2. 处理任何格式的文本 3. 熟练操作各种办公软件 4. 用日志记录各种操作 5. 管理时间常见技巧 6. 学会使用API和JSON获取数据 7. 用Python发邮件 8. 处理各种图像 9. 编写爬虫 10. 使用各种高级函数和特性优雅的解决问题
		2. 训练爬虫-1-html-js简介	
		3. 训练爬虫-2-css-http	
		4. 训练爬虫-3-爬虫训练手册	
		5. 爬虫有哪些姿势-静态和动态网站抓取	
		6. 案例：电商网站差评分析	
	不自己造轮子之高阶函数，迭代器和生成器	1. map，filter，reduce的应用	
		2. Iterator和Generator的关系	
		3. lambda表达式很好玩	
		4. 这些代码可以一行搞定	
		5. collections里面的几个有用方法	
		6. 项目：51备忘录v0.310	
	装饰器的概念与应用	1. 装饰器的概念	
		2. 装饰器的多种用途	
		3. 案例：验证登录状态	
		4. 项目：51备忘录v0.311	



学习安排——选修（1周）

课程名称	章	课程知识点	预期达到的目标
第四模块：数据库详解（选修）	学前准备 数据库详解	1. mysql数据库操作详解	掌握数据库基本操作： 增删改查及对应Python模块使用
		2: pymysql模块使用	
		3: mongodb数据库详解	
		4. pymongo模块详解	





课程名称	章	课程知识点	预期达到的目标
第五模块： 爬虫开发详解	第一章： 爬虫简介与工具使用	1. 爬虫简介	掌握技能： 1. 熟悉http协议，掌握爬虫常用开发工具 2. 熟练掌握Python常用爬虫模块，数据提取与存储方式，根据不同项目选择合适模块。 3. 爬虫系统的设计与性能优化，有一定的分布式爬虫经验； 4. 掌握爬取过程中数据加解密问题处理
		2. 爬虫基础之http协议详解	
		3. 爬虫开发必备之常用工具使用	
	第二章： Urllib模块详解与开发实战	1. urllib模块详解与第一次页面抓取	
		2. Url构建与自定义请求头信息	
		3~6: 第一个项目，项目关键点实现，代码设计与类实现，代码实现与调试	
		4. 项目练习：下载豆瓣电影top250海报	
		7~8. BeautifulSoup模块详解	
		9. 作业：爬虫基本练习	
		10. 项目练习：请求行为与提取信息分析	
		11~13. 项目练习：实现思路及爬虫类定义，关键数据提取与数据整理，数据与异常处理	
		14. 项目练习：数据存储及存在问题分析	
		15.代码模块化设计	
		16~19. 文本, csv,mysql, mogodb存储类实现，	
		20. 作业：抓取电影信息并保存	
	第三章： post请求与登录实现	1. post请求详解	
		2. post登录oschina分析	
		3. 登录github需求分析	
		4. 登录github实现	
		5. 登录失败分析与urlopen详解	
		6. cookie处理，保存与使用cookie登录	
		4. 登录失败分析与urlopen详解	



课程名称	章	课程知识点	预期达到的目标
第五模块： 爬虫开发详解	第四章： 反爬机制与高并发爬虫设计	1. robots详解	解决问题： 1. 分析项目需求，理解数据需求，根据项目需求制定爬取策略与实现，完成数据提取，清洗，入库，维护。 2. 高并发爬虫设计 3. 具有一定破解加密与反扒机制处理。
		2. 常见反爬机制与应对方式-伪装浏览器	
		3. 代理使用	
		4. 获取免费代理	
		5. 通过API获取代理	
		6. 代理去重与验证	
		7. 高并发爬虫设计与实现	
		8.课程总结	
	第五章 requests模块详解	1. requests模块简介与get详解	
		2. ajax分析与360图片抓取	
		3. 练习：使用关键字搜索下载图片	
		4. Post请求：表单，数据，文件	
		5. 练习：使用Post请求对人脸图片进行标记	
		6. lxml模块与xpath语法详解	
		7. 练习：抓取微博头条实战	
		8. 构建多线程并发爬虫-生产者和消费者模型构建并发爬虫	
		9. 作业：抓取NBA指定赛季数据	





课程名称	章	课程知识点	预期达到的目标
第五模块： 爬虫开发详解	第六章： scrapy框架	1-2. scrapy基础详解	市场价值： 掌握基本爬虫开发技能
		3. scrapy get与post请求	
		4. scrapy shell操作与信息提取	
		5. 使用scrapy抓取爬虫职位招聘信息	
		6. item与数据保存	
		7. scrapy框架，数据流，去重详解	
		8-9. scrapy中pipeline实现	
		10-11. scrapy中中间件处理	
		12. 练习：抓取爬虫职位详细信息	
		13. 数据清洗提取	
		14.自定义itemLoader	
		15-16. 文件下载，小说下载器实现	
		17. 构建分布式爬虫前提	
		18. 基于scrapy-redis构建分布式爬虫	
		19. scrapy-redis Request与去重详解	
		20. scrapy-redis 存储数据分析	



课程名称	章	课程知识点	预期达到的目标
第五模块： 爬虫开发详解	第七章： 爬虫加密与破解	1-4. 常见加密算法详解： md5,hash,Base64,RSA, AES	（薪资范围： 8K~12K）
		5. 爬取评论信息分析	
		6. JS断点调试与JS代码分析	
		7. Python代码实现JS行为	
		8-9. 爬取评论信息代码实现	
		10. selenium开发环境搭建与测试	
		11. selenium详解	
		12. 使用selenium构建微博自动评论器	
		13. 作业	





课程名称	章	课程知识点	预期达到的目标
第六模块： 数据分析可视化	第一章： numpy详解与实战	1. 认识numpy	掌握技能： 1. 掌握常用数据分析模块使用，具有一定数据分析能力 2. 掌握elasticsearch使用 解决问题： 1. 数据关联，分析力能。 2. 站内搜索引擎建立。 市场价值： 具有一定的数据分析能力,薪资范围： (9K~12K+)
		2.order与array属性	
		3.array访问方式	
		4.numpy生成array方法详解	
		5.numpy数据类型	
		6~7.numpy自定义数据类型	
		8.numpy基本运算	
		9.numpy比较运算与Boolean索引	
		10.numpy统计相关方法与axis详解	
	第二章： 数据可视化分析 _matplotlib详解	1.matplotlib简介与plot方法	
		2.pyplot方法	
		3.matplotlib添加子图和多图	
		4~6.综合练习：北京空气质量分析（数据抓取，清洗，整理）	
		7.bar状图详解	
		8.练习：统计分析每个月天气质量情况	
		9.练习：饼状图详解与应用	



课程名称	章	课程知识点	预期达到的目标
第六模块： 数据分析可视化	第三章： pandas/seaborn模块 详解与应用	1~2.pandas数据结构Series与DataFrame详解	<p>掌握技能：</p> <p>1. 掌握常用数据分析模块使用，具有一定数据分析能力</p> <p>2. 掌握elasticsearch使用</p> <p>解决问题：</p> <p>1. 数据关联，分析力能。</p> <p>2. 站内搜索引擎建立。</p> <p>市场价值：</p> <p>具有一定的数据分析能力,薪资范围： (9K~12K+)</p>
		3.pandas基本计算	
		4.缺失数据处理	
		5~6.pandas多级索引处理	
		7.pandas索引转换	
		8~9.pandas数据合并处理	
		10.pandas数据分组处理	
		11.分组数据统计计算与MovlieLens数据集分析	
		12.区间分组与pandas数据可视化	
		13.pandas时间处理	
		14.时间索引	
		15~16.pandas练习：数据可视化处理与数据整理	
		17.seaborn基本设置	
		18.seaborn数据集与第一个图标	
		19.直方图与KDE图	
		20.joinplot方法与皮尔逊系数	
		21.seaborn分类统计_柱状图	
		22~23.seaborn数据分布统计_箱状图与四分位数	
		24.多维度数据中不同维度相关性	
		25.作业：职位需求分析与可视化	



课程名称	章	课程知识点	预期达到的目标
第六模块： 数据分析可视化	第四章 数据可视化利器 _pyecharts详解与应用	1.pyecharts安装与第一个图标	掌握技能： 1. 掌握常用数据分析模块使用，具有一定数据分析能力 2. 掌握elasticsearch使用 解决问题： 1. 数据关联，分析力能。 2. 站内搜索引擎建立。 市场价值： 具有一定的数据分析能力, 薪资范围： (9K~12K+)
		2.根据需求快速掌握pyecharts相应图标	
		3~6.NBA球队数据与球员数据分析	
		7.基于位置与地图可视化	
		8.关系图与K线图	
	第五章 项目实战	9.制作词云	
		1~6.抓取北京所有二手房信息，并对二手房数据做分析	
		7~13.抓取某影评网站1980~2018所有电影数据并做相应分析	
		14~16.抓取某音乐软件某首歌曲评论及评论用户信息，对评论信息做相应分析	
		17.作业：出租房源分析	



课程名称	章	课程知识点	预期达到的目标
第七模块：数据分析常用模型与实战	第一章 统计学基础	1.数据集中趋势	掌握技能： 1. 掌握常用数据分析模块使用，具有一定数据分析能力 2. 掌握elasticsearch使用 解决问题： 1. 数据关联，分析力能。 2. 站内搜索引擎建立。 市场价值： 具有一定的数据分析能力,薪资范围： (9K~12K+)
		2.数据离散程度	
		3.概率	
		4.随机变量数据特征	
		5.线性变换	
		6.独立事件	
		7.几何分布详解	
		8.二项式分步详解与运用	
		9.泊松分布详解	
		10.分布计算与可视化	
		11.数学相关基础知识	





课程名称	章	课程知识点	预期达到的目标
第七模块：数据分析常用模型与实战	第二章 常用机器学习模型分析与应用	1.机器学习简介	掌握技能： 1. 掌握常用数据分析模块使用，具有一定数据分析能力 2. 掌握elasticsearch使用 解决问题： 1. 数据关联，分析力能。 2. 站内搜索引擎建立。 市场价值： 具有一定的数据分析能力,薪资范围： (9K~12K+)
		2.线性回归	
		3.实现线性代码回归实现	
		4.sklearn模块详解	
		5.波士顿房价预测	
		6.逻辑回归详解	
		7.分类算法之Knn	
		8.分类算法值贝叶斯分类器	
		9.分类算法治svm分类器	
		10.分类算法之决策树	
		11.聚类算法之层次聚类	
		12.kmeans聚类	
		13.pca降维与主成分析	
		14.案例分析	



课程名称	章	课程知识点	预期达到的目标
第七模块：数据分析常用模型与实战	第三章 推荐算法与应用	1.数据准备	掌握技能： 1. 掌握常用数据分析模块使用，具有一定数据分析能力 2. 掌握elasticsearch使用 解决问题： 1. 数据关联，分析力能。 2. 站内搜索引擎建立。 市场价值： 具有一定的数据分析能力,薪资范围： (9K~12K+)
		2.用户画像	
		3.用户行为分析	
		4.用户可视化分析	
		5.产品信息与知识图谱	
		6.推荐算法详解与应用	
		7.使用爬虫抓取音乐及用户信息，根据用户行为构建推荐系统	
		8.项目实战	





应用领域

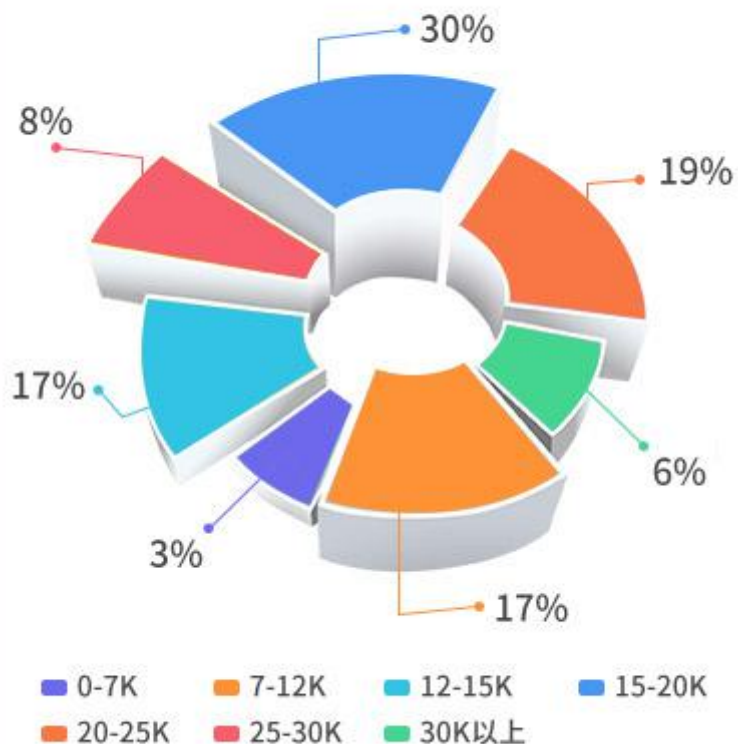


无处不在的数据分析师，企业的核心职位



职业发展

北京薪资水平分布占比



I 数据分析：

是指用适当的统计分析方法对收集来的大量数据进行分析，提取有用信息和形成结论而对数据加以详细研究和概括总结的过程。

I 就业前景：

大数据分析已广泛应用于各个领域，无论是国家政府部门、企事业单位，数据分析都是进行决策和制作决定的重要环节。



教学任务



教学方案

我们紧跟国际IT技术潮流，四大名企经验名师打造
为学员打造行业所需的教学大纲



实战项目

100多家企业真实的项目实战，10大行业大型项目
学完即可解决工作中实际应用问题



学习模式

报名就可以立马开始学习，无需再等开班，专享个人方案
全程录播学习加直播答疑



助学模式

授课名师，辅导老师，班主任，就业导师4大讲师角色
4位名师全程指导，一站式服务



考核模式

每个课程都设置关卡，通关后进入下个模块，督学促学
夯实每个模块的基础，高效学习



就业模式

超过1000家就业合作企业，遍及全国80多个城市
优秀学员推荐名企就业

Python⁺

51CTO学院微职位

2019 数据分析工程师

学习
模块

重点
项目

经典
案例

必备
技能

掌握最黑技术 占领更大市场

跟随知名企业优秀工程师、行业知名专家，系统全面的学习并掌握工作方法和技能技巧，获得全新的技能提升！

2019

感谢您的观看