

数据科学实战训练营课程安排

讲课内容	内容安排	具体内容
数据科学素养 暨南大学 王斌会	第一讲 Python 基础入门 6月17日 14:30-16:30	1. 数据收集与整理 1.1 数据的类型 1.2 数据的收集 1.3 数据的管理 2. 数据分析软件介绍 2.1 数据分析软件简介 2.2 Python 语言介绍 2.3 Python 数据分析平台 2.4 Python 编程入门 操作及练习 参考教程:《Python 数据分析基础教程》 王斌会 王术编著
	第二讲 Python 数据处理 6月22日 14:30-16:30	3. Python 数据分析基础 3.1 Python 数据类型 3.2 数值分析库 numpy 3.3 数据分析库 pandas 3.4 Python 编程运算 4. 数据的探索性分析 4.1 数据的描述分析 4.2 数据的分类分析 操作及练习
	第三讲 Python 可视化 6月24日 14:30-16:30	5. 数据的可视化分析 5.1 基本绘图命令 5.2 特殊统计图绘制 5.3 Seaborn 统计绘图 5.4 plotnine 绘图系统 5.5 pyecharts 绘图 操作及练习
数据科学素养 (第四讲) 与应用实例 (第五、六讲)	第四讲 数据采集 6月29日 14:30-16:30	1. 爬虫原理简介 2. 网页数据抓取、解析 2.1 网页解析 2.2 使用 requests 库抓取网页数据 2.3 使用 Beautiful Soup 解析网页 2.4 数据清洗与组织 操作及练习

<p>华东理工大学 张晨峰</p>	<p>第五讲</p> <p>统计与数据科学</p> <p>7月1日 14:30-16:30</p>	<p>1. 数学建模</p> <p>2. 统计分析</p> <p>3. 常用统计分析方法</p> <p>3.1 常用统计分析方法</p> <p>3.2 回归、分类、聚类</p> <p>操作及练习</p>
	<p>第六讲</p> <p>Python 在计量经济学中的应用</p> <p>7月6日 14:30-16:30</p>	<p>1. 线性回归模型</p> <p>2. 面板数据方法</p> <p>3. 政策分析初步</p> <p>操作及练习</p>
<p>应用实例 ——图书情报 南京财经大学 李树青</p>	<p>第七讲</p> <p>读者借阅记录分析</p> <p>7月9日 14:30-16:30</p>	<p>利用读者借阅记录，结合 pandas 数据分析功能，并通过可视化呈现方法，多角度展示常见表格数据的分析方法，同时利用近 10 年读者借阅记录展示读者兴趣变化的特点。</p> <p>操作及练习</p>
	<p>第八讲</p> <p>作者相关度分析</p> <p>7月10日 14:30-16:30</p>	<p>利用作者发文信息，结合网络可视化和地图可视化等功能，展示作者相关度分析，并结合 VOSViewer 和 Gephi 介绍对比分析功能。</p> <p>操作及练习</p>
<p>应用实例 ——金融数据分析与挖掘 广东金融学院 侯昉</p>	<p>第九讲</p> <p>银行新产品电话营销重点客户筛选</p> <p>7月13日 14:30-16:30</p>	<p>1. 任务概述：根据客户个人信息、账户信息等属性，对客户是否接受新产品的电话营销推广进行预测，从而确定重点客户目标。</p> <p>2. 数据集统计描述</p> <p>3. 数据预处理与属性相关分析</p> <p>4. 算法实施</p> <p>5. 模型应用与总结分析</p> <p>操作及练习</p>
	<p>第十讲</p> <p>贷款客户违约预测</p> <p>7月15日 14:30-16:30</p>	<p>1. 任务概述：根据客户个人信息、账户信息、第三方信用评估信息、当前贷款业务（额度、用途、期限等）等信息，对用户违约可能性进行预测，对贷款发放与否进行辅助决策。</p> <p>2. 数据集统计描述</p> <p>3. 数据预处理与属性相关分析</p> <p>4. 算法实施</p> <p>5. 模型应用与总结分析</p> <p>操作及练习</p>

注:除了上述内容外,根据实际教学情况,拟增设“小班培训”作为补充,如 Python 零基础入门精讲、爬虫进阶、可视化进阶等专题,让学员们可以针对性地提高自身的学习效果。

附件 2：培训教师介绍

王斌会：



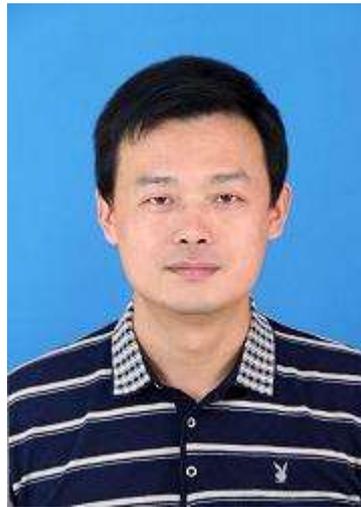
暨南大学管理学院教授、博士生导师，现为暨南大学经济管理国家实验教学中心主任，商业大数据分析中心主任。从事数学、统计学及经济管理教学和科研工作 30 多年，出版各类数据分析教程 10 余部，开发计算机软件 10 余套，研制多媒体课件多项。精通数据分析及 R 语言和 Python 编程与系统开发。

张晨峰：



经济学博士，现为华东理工大学商学院教师。1997 年、2003 年及 2012 年分别于南京大学、伦敦大学学院和复旦大学获得物理学学士、计算机硕士和经济学博士学位。主要研究领域是区域经济学、空间计量经济学以及机器学习和大数据在经济领域的应用。主持和参与多项省部级基金课题，在财经科学、南京社会科学等期刊发表多篇相关的学术论文。长期从事统计学和计量经济学相关的本科和研究生层次的教学工作，擅长基于 Python 和 R 的大数据及机器学习分析。

李树青：



南京财经大学信息工程学院副院长，教授，硕士生导师，南京大学信息管理系统博士，江苏省青蓝工程优秀青年骨干教师，江苏省科学技术情报学会理事，江苏省科技期刊学会科技评价专业委员会委员，美国纽约州立大学布法罗大学访问学者，国家自然科学基金委评审专家。教育部人文社会科学研究项目评审专家。《情报学报》、《现代图书情报技术》审稿专家。江苏省计算机学会计算机应用科技传播专家服务团成员。

侯昉：



广东金融学院互联网金融与信息工程学院副教授/系统分析师，华南理工大学计算机应用技术工学博士，2011年9月-2012年2月德国弗莱贝格工业大学访问学者。曾获《计算机学报》2011-2015年度优秀论文奖；完成科研项目五项；编写教材四本。主要研究方向为大数据存储与应用。

林晗川：



毕业于天津大学软件工程专业。北京百智享科技有限公司项目经理，主要负责公司项目管理、财经联盟运维等。