

Python 数据分析与可视化实战培训

【课程目标】

本课程为基础课程，主要讲解如何利用 Python 进行数据分析，以及数据可视化。假定学员已经基本掌握 Python 语言的使用。

通过本课程的学习，达到如下目的：

- 1、掌握数据分析的基本步骤和过程（数据分析六步曲）
- 2、掌握搭建数据分析框架的基本思想（数据分析框架）
- 3、熟悉 Pandas 常用数据结构，掌握用 Python 访问、操作数据集
- 4、掌握 Pandas 常用的统计功能（函数和方法）
- 5、理解统计分析原理，掌握统计分析常用的分析方法
- 6、熟练掌握 matplotlib 模块，熟练画图函数
- 7、学会解读图形，形成业务结论和业务策略。

【授课时间】

2 天时间

(要根据学员的实际情况调整重点内容及时间)

【授课对象】

业务支持部、IT 系统部、系统开发部、网络运维部等相关技术人员。

【学员要求】

- 1、 每个学员自备一台便携机(必须)。
- 2、 便携机中事先安装好 Python 3.9 版本及以上。

注：讲师现场提供开源的安装程序、扩展库，以及现场分析的数据源。

【授课方式】

分析步骤 + 分析框架 + 分析方法 + 可视化呈现 + 案例实战

采用互动式教学，围绕业务问题，展开数据分析过程，全过程演练操作，

让学员在分析、分享、讲授、总结、自我实践过程中获得能力提升。

【课程大纲】

第一部分：数据分析基础

目的：掌握数据分析基本步骤和过程，学会如何构造数据分析框架

- 1、 数据分析 VS 数据挖掘

2、数据分析的六步曲

- 步骤 1：明确目的--理清思路
- 步骤 2：数据收集—理清思路
- 步骤 3：数据预处理—寻找答案
- 步骤 4：数据分析--寻找答案
- 步骤 5：数据展示--观点表达
- 步骤 6：报表撰写--观点表达

3、搭建精准营销分析框架

演练：如何用大数据来支撑手机精准营销项目

第二部分：数据操作基础

1、简化的 Python 操作过程

2、常用扩展包

- Numpy 数组处理支持
- Pandas 数据分析和探索工具
- Matplotlib 可视化工具库

3、数据集读写

- 读取文件 (CSV 文件、Excel 文件)

- 数据集保存 (CSV、Excel)

4、数据集结构

- 数据集基本属性

- Index : 位置索引、标签索引

- Series : 一维结构

- DataFrame : 二维结构

5、数据集基本操作

- 数据访问 : 行访问/列访问/值访问

- 字段类型

 - ◇ 类型检查

 - ◇ 类型转换

 - ◇ 定义有序类别变量

- 排序

 - ◇ 按值排序

 - ◇ 按索引排序

- 数据筛选

- 数据修改

- 数据删除

第三部分：统计分析方法篇

1、统计分析基础

- 统计分析的关键要素

- 统计分析三个步骤

2、六种统计操作

- 描述统计 describe

- 分类计数 value_counts

- 分段计数/分箱计数 value_counts(bins)

- 分类汇总(groupby, count/sum/mean/...)

- 透视表（多维统计分析）pivot_table

- 按日期汇总 resample/to_period

案例实战：掌握常用的 Python 统计函数/方法

3、五种统计分析方法

- 对比分析法(不同用户的消费水平差异)

- 结构分析法(用户的学历结构、收入结构分析、动态结构分析)
- 分布分析法(用户的年龄分布、用户消费层次)
- 交叉分析法(产品偏好分析)
- 趋势分析法(销售淡旺季节、用户活跃时间)

案例实战：掌握常用的统计分析方法

第四部分：数据可视化

目的：掌握作图扩展库，实现数据可视化

- 1、中文显示的问题解决
- 2、了解图形元素及其函数
 - 标题、坐标轴、刻度
 - 数据标签、文本、注释
 - 图例、网格线、边框
- 3、简单图形的画法
 - 柱状图（简单/复式/堆积/堆积百分比柱状图）
 - 直方图（分布分析，查看分布特征）
 - 箱图（判断离群值）

- 饼图（结构分析）

- 折线图（趋势分析）

4、复杂图形的画法

- 多子图

- 多坐标系作图

- 多区域作图

5、图形保存

第五部分：数据预处理

1、数据预处理四大任务

- 数据清洗

- 数据集成

- 样本处理

- 变量处理

2、数据集成

- 数据追回

- 变量合并

- 拼接

演练：样本追加与变量合并

3、数据清洗

- 四大异常数据
- 重复值检查与处理
- 无效值检查与处理
- 离群值检查与处理
- 缺失值检查与处理

演练：异常值查找、删除、填充

4、样本处理

5、变量处理

第六部分：实战篇(上述知识点都融入下面分析实战中)

1、零售商用户消费行为分析

- 用户行为分析框架：5W2H
- 用户的典型特征
- 用户的消费能力
- 用户的消费水平

2、运营商用户购买行为分析

➤ 用户维度

- ◇ 用户地域分布
- ◇ 用户学历结构
- ◇ 用户消费能力/消费层次
- ◇ 用户流量分布/层次
- ◇ 用户流失分析

➤ 产品维度

- ◇ 套餐销量分析
- ◇ 套餐贡献分析
- ◇ 服务满意度分析
- ◇ 套餐偏好分析

➤ 时间维度

- ◇ 产品淡旺季分析
- ◇ 用户活跃度分析
- ◇ 重购周期分析

➤ 金额维度

- ◇ 收入结构（用户、产品、区域）
- ◇ 价格偏好分析
- ◇ 成本/利润分析

3、金融风险数据分析

- 用户维度
 - ◇ 违约用户的典型特征
 - ◇ 违约用户的消费水平
- 违约的影响因素分析
 - ◇ 违约与学历/岗位的关系
 - ◇ 违约与行业/职业的关系

注：会根据学员所在行业选择合适的实战案例。

结束：课程总结与问题答疑。