泰迪科技数据分析项目实训

供应链销售数据分析

**培训解决方案**

**广东泰迪智能科技股份有限公司 版权所有**

地址：广州市经济技术开发区开泰大道36号1栋212

网址：http://www.tipdm.com

邮箱：services@tipdm.com

邮编：510663

联系人：

电话：

目录

[1 项目介绍 3](#_Toc35352239)

[1.1 项目背景 3](#_Toc35352240)

[1.2 项目目标 5](#_Toc35352241)

[1.3 项目数据 5](#_Toc35352242)

[1.4 项目周期 5](#_Toc35352243)

[1.5 项目难度 5](#_Toc35352244)

[2 项目任务 5](#_Toc35352245)

[3 项目流程 5](#_Toc35352246)

[4 项目核心 6](#_Toc35352247)

[5 实现工具 6](#_Toc35352248)

[6 实训对象 6](#_Toc35352249)

[7 前置知识 6](#_Toc35352250)

[8 实训对应的就业岗位 6](#_Toc35352251)

[8.1 就业岗位 6](#_Toc35352252)

[8.2 岗位分析 6](#_Toc35352253)

[9 实训收益 7](#_Toc35352254)

[10 证书 7](#_Toc35352255)

[11 交付成果 8](#_Toc35352256)

[12 附件一 前置课程课表 9](#_Toc35352257)

# 项目介绍

## 项目背景

“Python供应链商品销售数数据分析”是泰迪科技专门为高校在校学员设计的一套实训项目。本项目综合了Python数据分析、数据可视化等技术，主要使用pandas、matplotlib、seaborn库可视化分析数据。

本项目首先探索数据总体的分布情况，检测缺失值和异常值；然后根据业务问题从不同维度分析数据，然后使用matplotlib等库进行可视化分析；

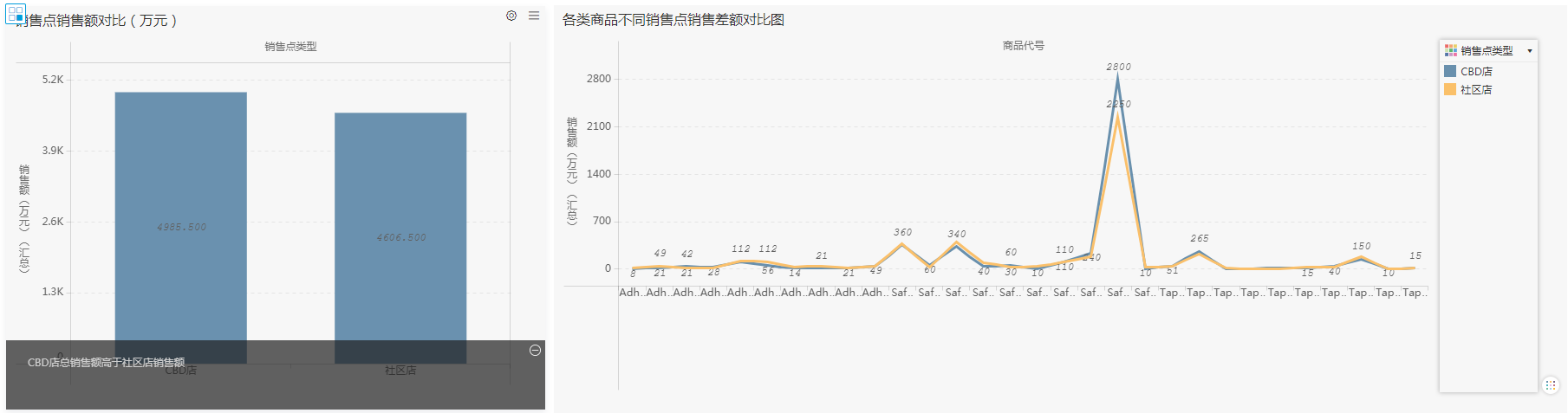
通过相对较短的时间内完成一个数据挖掘中最基础最常见的任务，向学生展示数据分析应用中的分析流程，使学生对数据分析在现实中的应用有一个整体的了解和掌握。通过本项目的学习，学员不仅能够掌握主流的数据挖掘方法，同时还为从事数据挖掘相关工作累积了数据分析项目的方法、流程和经验。

本项目采用“技术顾问”+“项目经理”+“学员”的团队组织模式。以完全企业化的方式与学生进行交流。学员在项目进行中，能熟悉企业的工作环境，在规定时间内完成项目需求、提升专业技术、锻炼团队协作能力与沟通能力。

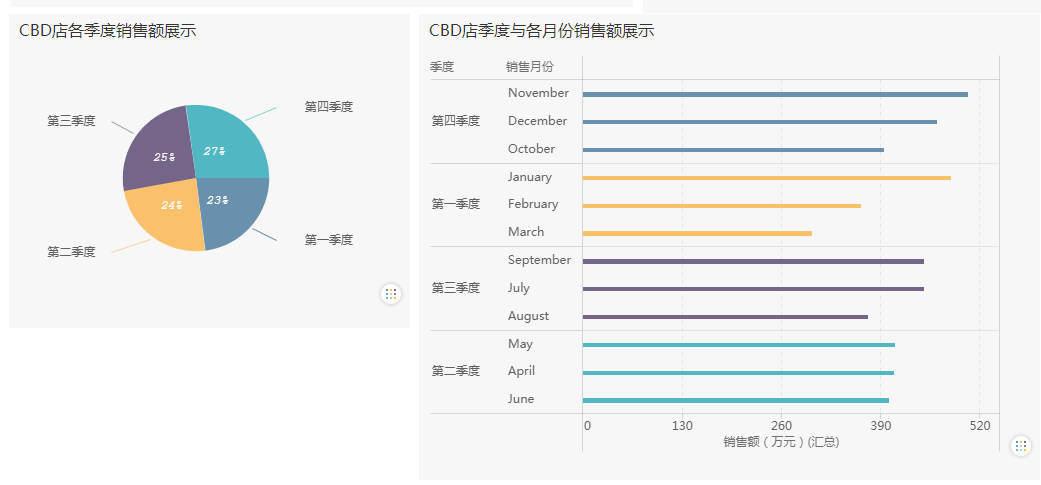
A市客户有意加盟380开设连锁门店，请你根据A市已有销售点的销售数据分析，给予该客户铺货支持和经营策略建议。

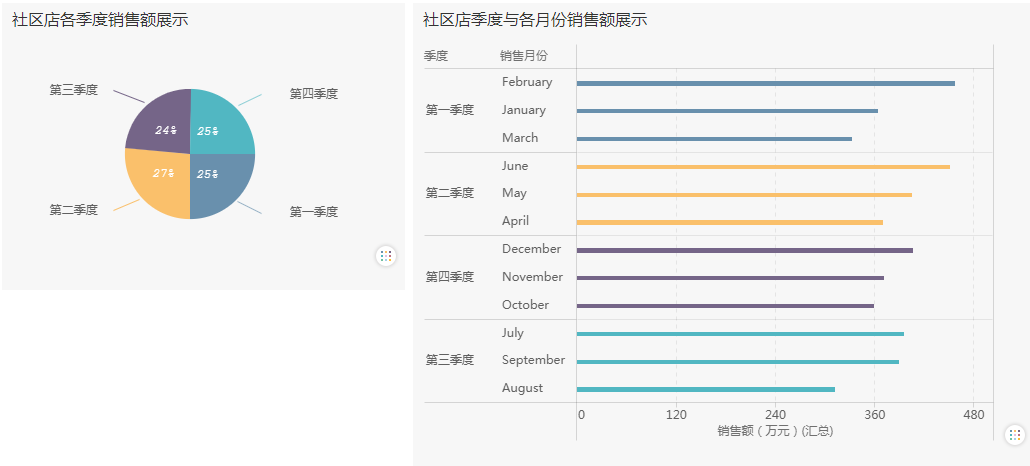
1. 提交结果

柱形图查看不同店的总销售额高低，折线图可查看不同商品在CBD店和社区店内的销售额对比状况，可参考以此分析。

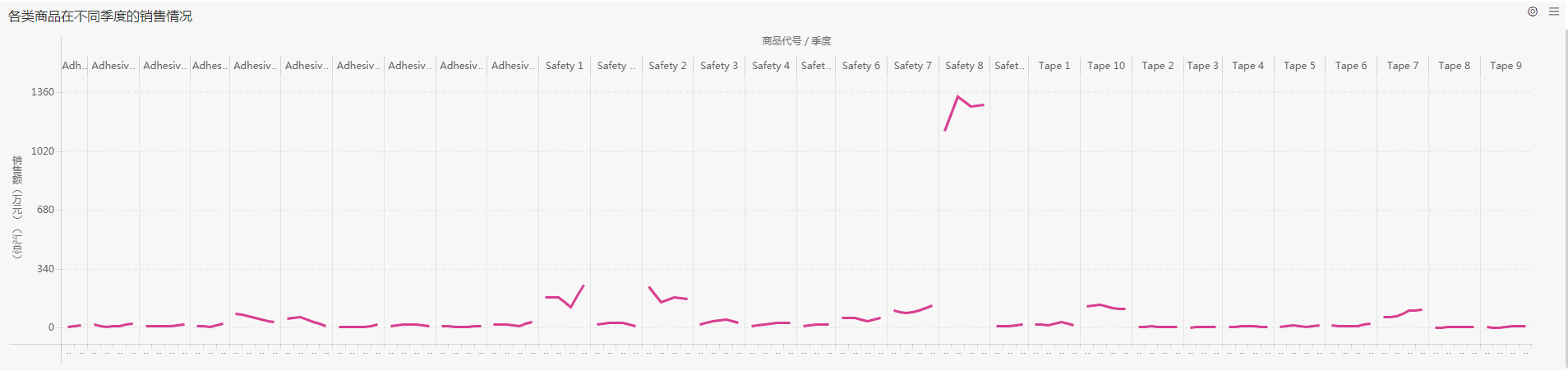


月份新建分组生成季度字段，可用多种图形分析各季度或者各月份销售额高低以分析休假等的任务，可参考以此分析。





使用分组好的季度字段，并使用折线图绘制不同商品在各个季度的销售趋势，折线幅度大的即为季节性强的产品，可参考以此分析。



2、主流技术

本项目主要涉及技术为利用pandas库进行数据预处理，利用matplotlib、seaborn进行数据可视化。

## 项目目标

A市客户有意加盟380开设连锁门店，请你根据A市已有销售点的销售数据分析，给予该客户铺货支持和经营策略建议。通过对部门的运营情况、财务状况、物流管理等不同维度的分析，评估该部门健康状况和发展趋势，指导平台发现问题并进行优化。

## 项目数据

见附件

## 项目周期

2周

## 项目难度

★★★

# 项目任务

任务1 明确项目需求与目标

任务2 环境准备：numpy、pandas、matplotlib、seaborn等

任务3 分析新开拓销售点的选址，是CBD店还是社区店更有利？（作业代码）

任务4 可挑选哪些产品做为新销售店的主要铺货产品，以帮助新店顺利经营。（作业代码）

任务5 分析销售季节性特点，据以安排店员年休假或培训时间，以及订货量控制。（作业代码）

任务6 假设进货商品均余半年保质期，分析社区店有哪些季节性很强的产品，必须在哪个时点前促销出清，否则会滞销过期。（作业代码）

任务7 完成项目报告，根据项目背景和数据分析结果，给客户提供店铺运营的个性化定制方案（作业）

# 项目流程

1. 明确项目目标：项目经理对项目进行介绍展示，明确项目交付内容。
2. 学习前置知识：提供项目所需知识的云课堂课程，快速掌握项目前置知识。
3. 项目实践：在企业工程师带领下获取项目需求，动手做项目。
4. 项目验收：对项目成果进行验收，指导改进项目成果。
5. 信息入库：参训学生信息录入泰迪人才库，为优秀结业生引荐合作公司。

# 项目核心

* Python数据分析
* Python数据可视化

# 实现工具

numpy、pandas、matplotlib、seaborn等

# 实训对象

数学、统计学、计算机等相关专业学生。

# 前置知识

项目所需的所有前置知识，皆提供在线课程供学生免费学习。

第1模块：Python编程基础

第2模块：Python数据分析与应用

第3模块：Python数据可视化

# 实训对应的就业岗位

## 就业岗位

数据分析师

## 岗位分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 岗位 | 主要业务工作 | 所需技能 | 相应课程设置 |
| 1 | 数据分析师 | 数据处理、分析建模、撰写分析报告 | Python  数据分析  分析报告 | Python语言设计  Python数据分析与应用  数据库 |

# 实训收益

* 足不出校门即可获得实战技能。
* 了解数据分析岗位目前的就业形式和前景，了解需要掌握的技能。
* 掌握一定的挖掘技能和工具，体验一个实际项目的全过程。
* 获得企业实习证明，无需奔波解决实习需求。
* 获得CBDA大数据分析工程师证（初级），国际高水平认证。
* 参训学生信息录入泰迪人才库，为优秀结业生引荐合作公司。

# 证书





# 交付成果

交付内容包括项目报告书，详情请看附件

# 附件一 前置课程课表

|  |  |
| --- | --- |
| Python编程基础 | Python数据分析与应用 |
| 1 准备工作  1.1 认识Python  1.2 搭建Python环境  1.3 安装PyCharm并创建一个应声虫程序  2 Python基础知识  2.1 掌握Python固定语法  2.2 创建字符串变量并提取里面的数值  2.3 计算圆形的各参数  3 Python数据结构  3.1 创建一个列表（list）并进行增删改查操作  3.2 转换一个元组（tuple）并进行取值操作  3.3 创建一个字典（dict）并进行增删改查操作  3.4 将两个列表转换为集合（set）并进行集合运算  4程序流程控制语句  4.1实现考试成绩划分  4.2实现一组数的连加与连乘  4.3使用冒泡排序法排序  4.4实训（猜数字游戏）  5 函数  5.1自定义函数实现输出方差  5.2使用匿名函数添加列表元素  5.3存储并导入函数模块  6 面向对象  6.1 认识面向对象编程  6.2 创建Car类  6.3 创建Car对象  6.4 迭代Car对象  6.5 产生Land\_Rover对象（子类）  7 文件基础  7.1 认识文件  7.2 读取txt文件中的数据  7.3 保存数据为csv格式文件  7.4 认识os模块 | 1 Python数据分析概述  1.1认识数据分析  1.2熟悉Python数据分析的工具  1.3安装Python3的Anaconda发行版  1.4掌握Jupyter Notebook常用功能  2 NumPy数值计算基础  2.1认识NumPy数组对象ndarray  2.2认识NumPy矩阵与通用函数  2.3利用NumPy进行统计分析  3 Matplotlib数据可视化基础  3.1了解绘图基础语法与常用参数  3.2分析特征间的关系  3.3分析特征内部数据分布与分散状况  4 Pandas统计分析基础  4.1读写不同数据源的数据  4.2掌握DataFrame的常用操作  4.3转换与处理时间序列数据  4.4使用分组聚合进行组内计算  4.5创建透视表与交叉表  5 使用Pandas进行数据预处理  5.1合并数据  5.2清洗数据  5.3标准化数据  5.4转换数据  6使用scikit-learn构建模型  6.1使用sklearn转换器处理数据  6.2构建并评价聚类模型  6.3构建并评价分类模型  6.4构建并评价回归模型 |

|  |  |
| --- | --- |
| Python数据可视化 |  |
| 1 准备工作环境  1.1准备介绍  1.2安装matplotlib、Numpy和Scipy库  1.3安装图像处理工具：Python图像库（PIL）  1.4配置matplotlib参数  2绘制并定制化图表  2.1简介  2.2定义图表类型—柱状图、线形图和堆积柱状图  2.3设置坐标轴长度和范围、线型、属性和格式化字符串  2.4添加图例和注解  3学习更多图表和定制化  3.1简介  3.2设置坐标轴标签的透明度和大小  3.3为图表线条添加阴影  3.4向图表添加数据表  3.5使用subplots(子区)  3.6定制化网格  4创建3D可视化图表  4.1简介  4.2创建3D柱状图  4.3创建3D直方图.  4.4在matplotlib中创建动画  5用图像和地图绘制图表  5.1用PIL做图像处理  5.2在具有其他图形的图表中显示图像  5.3使用Basemap在地图上绘制数据 |  |

# 作业提交时间安排

|  |  |  |  |
| --- | --- | --- | --- |
| 任务安排 | 内容 | 提交场所 | 提交时间 |
| 第一周 | 任务1 明确项目需求与目标 | 泰迪云课堂 | Day1 22:00 |
| 任务2 环境准备：numpy、pandas、matplotlib、seaborn、sklearn（5%） |
| 任务3 分析新开拓销售点的选址，是CBD店还是社区店更有利？（作业代码）（20%） | Day5 22:00 |
| 任务4 分析销售季节性特点，据以安排店员年休假或培训时间，以及订货量控制（作业代码）（50%） |
| 第二周 | 任务5 假设进货商品均余半年保质期，分析社区店有哪些季节性很强的产品，必须在哪个时点前促销出清，否则会滞销过期。（作业代码）（80%） | Day1 22:00 |
| 任务6 完成项目报告，根据项目背景和数据分析结果，给客户提供店铺运营的个性化定制方案（作业）（100%） | Day5 22:00 |
| 在线答疑 |  | 工作室微信群 |  |

# 评分标准

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 注意：项目总分=总计\*项目难度系数 项目结果有3次修改机会，以最终提交成果为依据 | | | | | | |
| **序号** | **评审项目** | **子项目** | **指 标** | **满分** | **分数** | **备注** |
| 1 | 项目进度 | 项目进度 | 在项目指定截止时刻的项目进度（如90%，70%） | 20 |  | 分数=20\*项目进度 |
| 2 | 报告文档 | 内容完整性 | 与模板进行对比，是否各个步骤都在文档中体现出来 | 10 |  |  |
| 排版 | 与模板进行对比，是否按标准格式进行排版 | 10 |  |  |
| 内容质量 | 方案合理性 思路是否清晰 模型结果是否优异 创新性 | 30 |  |  |
| 3 | 汇报PPT | 内容完整性 | 与模板进行对比，是否各个步骤都在文档中体现出来 | 5 |  |  |
| 排版 | 与模板进行对比，是否按标准格式进行排版 | 5 |  |  |
| 内容质量 | 方案合理性 思路是否清晰 模型结果是否优异 创新性 | 5 |  |  |
| 4 | 代码&中间数据 | 内容完整性 | 是否包含项目解决方案的各个环节对应的代码；是否有必要的中间数据。 | 10 |  |  |
| 代码规范与设计 | 代码编写是否简洁、规范、高效；脚本分割与命名是否与模板对应。 | 5 |  |  |
| 总计 | | | | |  |  |
| 总分 | | | | |  |  |

**工作室邀请函**

为了适应大数据与人工智能及发展的需求，顺应教育部提倡的深化校企合作的号召，更好的服务于广大数据分析爱好者，广东泰迪智能科技股份有限公司诚邀各高校相关专业老师、相关协会学会、俱乐部等组织合作成立“泰迪·智能工作室”，工作室以独立的模式运营，并以学生为中心成立，受泰迪科技监督且由其免费提供各种工作室所需资源的创新型数据智能工作室。



工作室旨在通过教育与产业之间的联动，实行“引进来，走出去”模式，引导学生学习数据科学与人工智能方法为导向，通过与企业的联系、合作、实践，激发学生的数据分析思维，全面推进数据分析与人工智能发展，提高大学生的数据分析素质，激发学生的创新创业精神，以实现创新型数据智能创业人才为培养目标。