实验目录：

1. **基础语言学习实验（4个子模块，42个实验）**

**1.1 R语言编程基础实验**

* 实验一：R基本环境
* 实验二：R包
* 实验三：查看帮助
* 实验四：定义变量
* 实验五：变量赋值
* 实验六：定义数组
* 实验七：定义向量
* 实验八：定义数据框
* 实验九：定义函数
* 实验十：条件判断语句
* 实验十一：循环语句
* 实验十二：可视化语句
* 作业1

**1.2 R语言统计建模与分析基础实验**

* 实验一：数据源连接
* 实验二：数据查看
* 实验三：数据切片
* 实验四：数据编辑
* 实验五：描述性统计
* 实验六：点图和条形图
* 实验七：直方图和核密度图
* 实验八：箱线图
* 实验九：Hexbin图
* 实验十：散点图矩阵
* 作业2

**1.3 Python语言编程基础实验**

* 实验一：打印Hello World
* 实验二：查看帮助
* 实验三：定义变量
* 实验四：变量赋值
* 实验五：定义数组
* 实验六：定义向量
* 实验七：定义函数
* 实验八：条件判断语句
* 实验九：循环语句
* 实验十：可视化语句
* 作业3

**1.4 Python语言统计建模与分析基础实验**

* 实验一：数据库连接
* 实验二：数据导入
* 实验三：数据切片
* 实验四：数据导出
* 实验五：描述性统计
* 实验六：可视化处理
* 实验七：Numpy库操作
* 实验八：Scipy库操作
* 实验九：Pandas库操作
* 实验十：Sciki-learn库操作
* 作业4

1. **数据分析与挖掘实验（4个子模块，28个实验）**

**2.1 R数据探索与预处理实验**

* 实验一：数据质量分析实验
* 实验二：数据特征分析实验
* 实验三：统计特征函数实验
* 实验四：统计作图函数实验
* 实验五：数据清洗实验
* 实验六：数据集成实验
* 实验七：数据变换实验
* 实验八：数据规约实验
* 作业5
* 作业6

**2.2 R数据挖掘实验**

* 实验一：R语言分类预测基本流程实验
* 实验二：R语言数据预处理实验
* 实验三：R语言决策树分类方法实验
* 实验四：R语言高级分类方法实验
* 实验五：R语言聚类分析实验
* 实验六：R语言关联分析实验
* 实验七：R语言回归分析实验
* 作业7
* 作业8

**2.3 Python数据探索与预处理实验**

* 实验一：数据质量分析实验
* 实验二：数据特征分析实验
* 实验三：统计特征函数实验
* 实验四：统计作图函数实验
* 实验五：数据清洗实验
* 实验六：数据集成实验
* 实验七：数据变换实验
* 实验八：数据规约实验
* 作业9
* 作业10

**2.4 Python数据分析实验**

* 实验一：Python数据降维实验
* 实验二：Python聚类分析实验
* 实验三：Python K近邻分类实验
* 实验四：Python回归分析实验
* 实验五：Python决策树分类实验
* 作业11
* 作业12

1. **大数据基础实验（4个子模块，13个实验）**

**3.1 Hadoop基础实验**

* 实验一：Hadoop安装与配置实验
* 实验二：HDFS基本操作实验
* 实验三：HDFS进阶实验
* 实验四：Hadoop Streaming实验
* 作业13
* 作业14

**3.2 MapReduce基础实验**

* 实验一：MapReduce基本操作实验
* 实验二：MapReduce高级特性实验
* 实验三：迭代式MapReduce实验
* 作业15
* 作业16

**3.3 HBase基本实验**

* 实验一：HBase安装与配置实验
* 实验二：HBase基本操作实验
* 实验三：HBase进阶实验
* 作业17

**3.4 Hive基础实验**

* 实验一：Hive安装与配置实验
* 实验二：Hive基本操作实验
* 实验三：Hive进阶实验
* 作业18

1. **大数据挖掘实验（2个子模块，12个实验）**

**4.1 R大数据挖掘实验**

* 实验一：KNN分类实验
* 实验二：支持向量机分类实验
* 实验三：决策树分类实验
* 实验四：随机森林分类实验
* 实验五：神经网络分类实验
* 实验六：不同算法的比较
* 作业19
* 作业20

**4.2 Python大数据挖掘实验**

* 实验一：KNN分类实验
* 实验二：支持向量机分类实验
* 实验三：决策树分类实验
* 实验四：随机森林分类实验
* 实验五：神经网络分类实验
* 实验六：不同算法的比较
* 作业21
* 作业22

1. **案例分析实验（3个子模块，18）**

**5.1案例一：客户流失与分析**

* 实验一：数据读取
* 实验二：数据探索
* 实验三：数据预处理
* 实验四：分析建模
* 实验五：模型评价与优化
* 实验六：可视化输出
* 作业23

**5.2案例二：人流分析与预警**

* 实验一：数据读取
* 实验二：数据探索
* 实验三：数据预处理
* 实验四：分析建模
* 实验五：模型评价与优化
* 实验六：可视化输出
* 作业24

**5.3案例三：门店销量预测**

* 实验一：数据读取
* 实验二：数据探索
* 实验三：数据预处理
* 实验四：分析建模
* 实验五：模型评价与优化
* 实验六：可视化输出
* 作业25

具体实验内容可以参见：https://tjxlab.gitbooks.io/bigdata/content/

学生账户：

1352741 张子豪

1451724 黄靖凯

1454009 娄勇进

1459351 羅偉成

1351675 陈志

1451206 郑煜铭

1452169 刘义

1452485 向丹阳

1452901 吴竞舟

1459350 歐宣妤

1552397 张北

1552633 韦尧

1552678 王思源

1553594 张世宇

1559039 喻支俊

1552809 陈怡坤

1454096 周峰

1552219 余淼

1551235 郑航

1451933 王雨磊

教师账户：

09145 苏艺

08155 王珣

07158 范斌

10145 张翔