**课程名称：运动健康与数据科学 （备用名：数据分析与R语言应用/高级体育统计学与R语言应用）**

**英文名称**：**Data Science in Exercise and Health (Data Analysis with R/Advanced Statistics in Exercise Science with R)**

**领衔教师（姓名、工资号）：**

张展嘉、2006191077

**上课老师：**

张展嘉（30学时），校（院）外专家 （2学时）

**学分：2 周学时：2 总学时：32**

**申报学期（指第一或第二学期开课）：**

第二学期

**课程简介**

大数据、数据科学的发展正在影响着社会的各行各业，本课程将主要介绍多种运动与健康领域常见的模型和数据分析方法，主要涵盖：一般线性模型、广义线性模型、多层模型、结构方程模型、有监督和无监督机器学习等。本课程将以R语言的实操应用为重点，结合现实中的真实数据库，培养学生数据获取、数据清洗、数据可视化、数据分析与结果解读的能力。

**授课对象**

此课程主要面向体育专业硕士研究生，同时也欢迎其余院系的研究生选课。

**考核形式**

作业+期末报告与展示：此课程会安排三次的平时作业，作业由每位同学单独完成，三次作业共占总成绩的45%；期末数据分析项目占总成绩的40%，要求学生根据各自的研究问题选取特定的数据库进行分析，提交报告并进行课堂展示，此项目由2人一组完成；考勤与课堂表现占总成绩的15%。

**课程大纲及学时安排：**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 课次 | 内容 | 任课教师 | 课时 | 地点 |
| 1 | 数据的理解；R语言基础 | 张展嘉 | 2 | 教室 |
| 2 | 数据的准备与预处理 | 张展嘉 | 2 | 教室 |
| 3 | 数据可视化基础 | 张展嘉 | 2 | 教室 |
| 4 | 统计推论基础 | 张展嘉 | 2 | 教室 |
| 5 | 方差分析；协方差分析 | 张展嘉 | 2 | 教室 |
| 6-7 | 一般线性模型；广义线性模型 | 张展嘉 | 4 | 教室 |
| 8-9 | 纵向数据分析；多层回归 | 张展嘉 | 4 | 教室 |
| 10 | 运动表现分析 | 校外专家 | 2 | 教室 |
| 11-12 | 结构方程模型 | 张展嘉 | 4 | 教室 |
| 13 | 数据报告的撰写与展示 | 张展嘉 | 2 | 教室 |
| 14-15 | 机器学习 | 张展嘉 | 4 | 教室 |
| 16 | 期末报告与展示 | 张展嘉 | 2 | 教室 |
|  |  | 合计 | 32 |  |

**参考书：**

1. Rob Kabacoff - R in Action: Data Analysis and Graphics with R (2015, Manning)
2. Marc A. Scott, Jeffrey S. Simonoff, Brian D. Marx - The SAGE Handbook of Multilevel Modeling (2013, Sage)
3. Rex B. Kline-Principles and practice of structural equation modeling (2015, Guilford publications)