《数据可视化》教学大纲

课程编号：071183B

课程类型：□通识教育必修课 □通识教育选修课

□学科基础课 ☑专业核心课

□专业提升课 □专业拓展课

总 学 时：48 讲课学时：16 实验（上机）学时：32

学　　分：2

适用对象：数据科学与大数据技术（信息技术）专业

□是 ☑否 适合作为其他专业学生的个性化选修课

先修课程：计算机应用基础、

计算机图形学、概率论与数理统计

一、教学目标

首要目标为立德树人，通过课程思政的方式，培养社会主义事业的建设者。本课程是信息学院信息管理与信息系统专业（大数据方向）学生的专业选修课。本课程是为适应信息时代对大数据和商业智能人才的要求而开设的，既重视学生相关理论的系统学习，又强调培养学生发现问题、分析问题和解决问题的实践应用能力。本课程是系统设计、数据挖掘、软件开发、决策支持等领域人才培养方面重要组成部分，是专业人才培养目标得以实现的重要保证。

通过本课程的学习，培养学生的信息数据处理能力、信息分析和应用能力，信息表达能力。使学生掌握数据可视化的一般原理和处理方法，能使用数据可视化工具对数据进行可视化处理。能够针对不同业务情景下的海量数据，洞察和分析隐藏在数据背后的重要信息，并可视化展示为相关决策者提供辅助决策支持。该课程能够帮助学生养成良好的职业素质，培养动手能力、创新能力和独立解决实际问题的能力。

二、教学内容及其与毕业要求的对应关系

（一）教学内容

本课程主要学习数据可视化的基础理论和概念，针对实际应用中遇到的不同类型的数据，学习相应的可视化方法，并学习可视化综合应用及实用系统。

教学内容包括：数据可视化的基础理论和概念，视觉感知和视觉通道，数据处理和分析方法，不同类型数据的可视化方法，可视化交互与评估，以及可视化软件与工具等。其中，有关数据可视化的基础理论和数据处理和分析方法应细讲和精讲，而视觉感知和视觉通道则粗讲，重点是不同类型数据的可视化方法。课程的难点在于数据变换方法和可视化编码方法，拟通过实验例子讲解与学生自己动手相结合方式进行讲授。

（二）教学方法和教学手段

根据教学目标，拟采用的教学方法有：课堂讲解基本理论、概念和核心知识，讲授和讨论相结合领会知识要点，实验教学训练独立解决问题的能力，最后借助可视化软件工具让学生进行上机操作和具体实践。

（三）实践教学环节的要求

根据教学进度和要求布置相应的课后作业，通过上机实践。每一章节都有对应的思考题和上机内容，如可视化软件工具包的开发环境配置、可视化的流程设计、各章节提供的用于可视化设计和展示的例子等。

（四）学习要求

为有效学习本课程，要求学生首先具备计算机应用基础知识、计算机图形学，以及概率论与数理统计的基本理论等方面的基本知识，熟练掌握数据处理和数据分析等方面的完整专业知识。自学时应该熟读大纲，提纲挈领地掌握数据可视化的流程和方法，随后按照大纲熟读教材，并通过课后思考题和上机实践进行多角度和多层次地反复学习。

（五）与毕业要求的关系

数据可视化是大数据产业方向的重要组成部分，也是未来发展的必然趋势。环顾全球，许多国家已经把数字化、信息化、智能化作为国家战略的关键主题，信息基础设施建设成为后金融危机时代振兴经济的重要手段。特别是大数据时代下，云计算、物联网技术的逐步成熟，使得基于数据分析结果的数据可视化正在成为各个领域内的战略决策。因此，学习和掌握数据可视化的相关知识、方法和技术，不仅是当前大数据产业发展的需求，还是适应未来各种互联网商业战略决策的基本要求，也是信息管理与信息系统专业培养的基本要求。

三、各教学环节学时分配

**教学课时分配**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 章节内容 | 讲课 | 实验 | 其他 | 合计 |
| 1 | 数据可视化简介 | 2 |  |  | 2 |
| 2 | 视觉感知与视觉通道 |  | 2 |  | 2 |
| 3 | 数据处理与分析 | 2 | 2 |  | 4 |
| 4 | 数据可视化基础 | 2 | 2 |  | 2 |
| 5 | 时空数据可视化 | 2 | 4 |  | 6 |
| 6 | 地理空间数据可视化 | 2 | 4 |  | 6 |
| 7 | 高维非空间数据可视化 | 2 | 4 |  | 6 |
| 8 | 层次和网络数据可视化 | 2 | 4 |  | 6 |
| 9 | 跨媒体数据可视化 | 2 | 4 |  | 6 |
| 10 | 可视化交互与评估 |  | 2 |  | 2 |
| 11 | 可视化软件与工具 |  | 4 |  | 4 |
| 合计 |  | 16 | 32 |  | 48 |

四、教学内容

**第一章 数据可视化简介**

课程思政元素设计：通过中国对疫情有效控制案例，阐述课程价值。

教学内容：概要介绍可视化简史，可视化的目标、作用和意义。对数据可视化分类，并阐述与其他学科方向的关系。

重点和难点：数据可视化的目标和作用。

考核要求：使学生了解可视化简史，深入理解可视化的目标、作用和意义。

思考题一：举例说明数据可视化的三类通用目标。

**第二章 视觉感知和视觉通道**

教学内容：介绍感知、认知、格式塔理论、视觉感知相对性。学习视觉通道类型和视觉通道特性。

重点和难点：数据通道类型。

考核要求：使学生掌握视觉感知与认知，了解视觉通道。

实验一：通过图像对比，理解视觉感知的相对性。

实验二：用编程语言实现2个常用的颜色空间相互转换算法。

实验三：熟悉几种在线颜色选择工具，如colorbrewer和adobe kuler等。

**第三章 数据处理与分析**

课程思政元素设计：疫情数据质量提升，有助于中国疫情防控。

教学内容：介绍数据的属性和特征，数据预处理的方法，数据的存储和分析方法。

重点和难点：数据预处理，数据分析。

考核要求：使学生掌握数据特征，掌握数据预处理和数据分析方法。

实验四：对给定的数据集，进行数据预处理，清洗数据噪声并填补缺失数据。

实验五：对给定的数据集，在数据预处理后，对数据进行线性回归分析。

**第四章 数据可视化基础**

教学内容：着重讲授数据变换方法，数据映射和视图的交互设计。

重点和难点：数据变换。

考核要求：使学生掌握可视化流程，掌握数据处理和数据变换方法，理解可视化编码方法，掌握可视化设计原理。

思考题二：讨论不同元素在可视化双变量数据集时可能产生的可视化效果和遇到的问题。

实验六：选择一个多维数据集，设计一个可视化隐喻方案。

**第五章 时空数据可视化**

课程思政元素设计：准确的时间序列预测有助于疫情防控。

教学内容：着重讲授低维度数据的可视化方法，多变量的空间数据可视化，以及时序数据可视化方法。

重点和难点：时序数据可视化方法。

考核要求：使学生掌握低维度数据的可视化方法，理解三维标量数据可视化方法，掌握时间序列数据可视化方法，掌握多变量空间数据可视化方法。

实验七：用不同颜色对二维温度场进行可视化，并观察效果。

实验八：搜索中美历史单行最高领导人姓名、任期以及相互访问的时间、地点。设计可视化直观有效地展示研究成果。

**第六章 地理空间数据可视化**

教学内容：介绍地图投影，点形数据、线形数据和区域数据的可视化方法，以及地理信息可视化方法

重点和难点：点形数据、线形数据的可视化，地理信息可视化。

考核要求：使学生理解地图投影，掌握点形数据、线形数据的可视化，理解区域数据的可视化，了解基于地理位置的综合信息可视化。

实验九：通过网络搜索世界各国在过去20年间的GDP总量，通过专门工具生成GDP在20年的动态变化并可视化展示。

**第七章 高维非空间数据可视化**

课程思政元素设计：透过现象看本质，将复杂问题简单化可视化，有助于找到有效疫情防控的方法。

教学内容：介绍主成分分析法，多维尺度分析法，等距映射法和局部线性嵌入法，以及高维数据的可视化呈现和交互。

重点和难点：高维数据的可视化呈现。

考核要求：使学生理解地图投影，掌握点形数据、线形数据的可视化，理解区域数据的可视化，了解基于地理位置的综合信息可视化。

实验十：通过网络下载一个多维度数据，分别采用像素图、平行坐标和马赛克图等方法比较可视化结果。

**第八章 层次和网络数据可视化**

教学内容：介绍结点链接法，空间嵌套填充法，相邻矩阵布局，动态网络数据可视化，图可视化数据效果和图可视化交互。

重点和难点：结点链接法，空间嵌套填充法，相邻矩阵布局。

考核要求：使学生理解地图投影，掌握点形数据、线形数据的可视化，理解区域数据的可视化，了解基于地理位置的综合信息可视化。

实验十一：用缩进法和径向布局法为美国NBA设计赛程可视化，对比两者的区别和优势。

实验十二：用相邻矩阵布局可视化任意一个社交网络。

**第九章 跨媒体数据可视化**

课程思政元素设计：如何有效监控媒体，做好舆情分析，有助于社会稳定与疫情防控。

教学内容：介绍文本数据、文档数据、日志数据可视化，以及社交网络可视化。

重点和难点：日志数据可视化。

考核要求：使学生掌握文本与文档可视化方法，了解社交网络可视化，掌握日志数据可视化。

思考题三：找一段文本并计算出该文本的词频向量。

实验十三：对你的微博好友关注这一社交网按照结构型可视化方法中的结点链接图进行可视化，并对比不同用户的社交网络结构的异同。

**第十章 可视化交互与评估**

教学内容：介绍可视化交互，可视化价值和评估。

重点和难点：可视化交互。

考核要求：使学生掌握和了解可视化的价值和评估。

实验十四：选择一种可视化交互方法，设计细节调控的交互并实现。

实验十五：选择两种可视化方法显示二维流场数据，设计定量评估，并在小范围用户中完成实验，讨论评估结果。

**第十一章 可视化软件与工具**

课程思政元素设计：熟练掌握可视化工具，可以全面掌控疫情现状与发展。

教学内容：介绍可视化软件分类，科学可视化软件与工具的使用，信息可视化软件与工具的使用，可视化分析软件与开发工具的使用。

重点和难点：可视化软件与工具的使用。

考核要求：使学生了解可视化软件分类，掌握常规的可视化软件与工具使用。

实验十六：利用excel绘制数据视图。

实验十七：利用tableau绘制数据视图。

五、考核方式、成绩评定

本课程所采用平时、课程设计相结合的考核方式，尤其注重实践环节的考核。各部分所占比重为平时成绩占40%，期末课程设计成绩占60%。平时成绩包括考勤成绩10%和上机作业成绩30%。期末考核采用课程设计形式，根据具体业务场景的考核要求，进行相应的设计与实现，注重知识应用能力和解决问题能力的考核，需要学生提交自己设计的调查报告（调查问卷设计、调查统计、数据分析、数据可视化、综合效果。

六、主要参考书及其他内容

（一）推荐教材：

[1] 陈为著。数据可视化的基本原理与方法。科学出版社，2013.

（二）主要学习参考书

[1] 祝洪凯著。数据可视化之美。机械工业出版社，2011.

[2] 陈为著。数据可视化。电子工业出版社，2013.

执笔人：

教研室主任：

系教学主任审核签名：