安徽财经大学字

**本科毕业设计**

|  |  |
| --- | --- |
| **题 目** | **疫情下旅游行业生存状况可视化分析** |
| **学 院** | **管理科学与工程学院** |
| **专 业** | **计算机科学与技术** |
| **班 级** | **20计科四班** |
| **学 号** | **20202229** |
| **姓 名** | **贾凯** |
| **指导老师** |  |

**2023 年 3 月**

安徽财经大学管理科学与工程学院

本科生毕业论文（设计）诚信承诺书

本人承诺：

1.所呈交的毕业论文（设计）《 毕业设计过程管理系统》，是在认真学习理解《安徽财经大学学位论文作假行为处理办法》和《管理科学与工程学院本科毕业论文（设计）工作管理办法》后，保质保量独立完成的，没有弄虚作假，没有抄袭别人的内容；

2.毕业论文（设计）所使用的相关资料、数据、观点等均真实可靠，文中所有引用的他人观点、材料、数据、图表均已注释说明来源；

3.毕业论文（设计）中无抄袭、剽窃或不正当引用他人学术观点、思想和学术成果，伪造、篡改数据的情况；

4.本人已被告知并清楚：学院对毕业论文（设计）中的抄袭、剽窃、弄虚作假等违反学术规范的行为将严肃处理，并可能导致毕业论文（设计）成绩不合格，无法正常毕业、取消学士学位资格或注销并追回已发放的毕业证书、学士学位证书等严重后果；

5.若在省教育厅、学校、学院组织的毕业论文（设计）检查中，被发现有抄袭、剽窃、弄虚作假等违反学术规范的行为，本人愿意接受学院按有关规定给予的处理，并承担相应责任。

学生（签名）：

年 月 日

指导老师（签名）：

年 月 日

**中文标题**

**（宋体 二号 加粗 1.5倍行间距 段前段后0.5行 居中）**

**摘 要**

正文：小四 宋体 1.5倍行距 段落首行缩进2字符

**关键词：xx**；xx

**Title**

**（times new roman 二号 加粗 1.5倍行间距 段前段后0.5行 居中）**

**Abstract**

正文：小四 times new roman 1.5倍行距 首行缩进2字符

**Keywords： xx; xx**

# 目 录

**一级标题三号加粗 行间距1.5倍 段前段后0.5行或6磅**

**中文宋体，英文和字符 times new roman**

# 1 引言

主要写开发该系统的背景、目的和意义（比摘要要详细和全面）、软件的定义、软件的开发工具，阐述课题设计的创新点、实用价值等。

(一级标题 宋体 三号 加粗 1.5倍行间距 段前段后各0.5行 左对齐)

正文：小四 宋体 1.5倍行距 首行缩进2字符

2019年年底新冠肺炎疫情暴发以来，感染新冠肺炎人数的激增使得98%旅游目的地实施边境关闭、出示疫苗接种和核酸检测证明等旅行限制。世界旅游理事会发布年度经济影响报告指出，2019年全球旅游行业创造了9.2万亿美元GDP降至2020年的4.7万亿美元，直接损失4.5万亿美元；2019年对全球经济贡献占比10.4%下降至2020年的5.5%;2020年旅游业失去6160多万个工作岗位，降幅达18.5%,全球旅游业备受打击。 在此背景下，文化和旅游部与国家卫生健康委多次出台政策对景区最大承载量进行动态调整。旅游业要恢复到新冠肺炎爆发前仍有较大的差距。面对前所未有的冲击，学术界涌现出大量关于旅游业与疫情相关的研究成果，本程序主要使用Python语言实现，主要包括数据的爬取，处理和可视化三部分。首先通过网络爬虫采集数据，其次对爬取的数据进行清洗处理并进行可视化分析，将数据呈现为柱状图、饼图和词云等，并进行词云的制作。

# 2 可行性研究

概述段落

## 2.1 需求概述

粗略分析系统操作流程以及功能和特色。

(二级标题 宋体 小三号 加粗 1.5倍行间距 段前段后各0.5行 左对齐)

结合系统流图或业务流程图概述系统需求及物理组成。

疫情下旅游行业生存状况可视化分析的设计原则是通过对大量疫情期间旅游数据的收集和整理，并按照目标进行有规律和有针对性的分析，从而对应各模块具体分析生成相关数据。我们可以从中对疫情下旅游行业生存状况有基本的认识和了解，最后的目标是形成一个可视化的分析结果。为更科学地进行评估，我们对北京、上海、成都、武汉等旅游热门景点评论数据进行数据爬取。

疫情下旅游行业生存状况可视化分析可以帮助旅游行业提供数据分析，辅助运营决策依据。旅游行业运营者则可以根据数据走势预测结果，辅助决策。对疫情下旅游行业生存状况进行详细分析，数据可视化在疫情稳定后对旅游业的复苏起到了关键作用。

## 2.2 系统可行性

### **2.2.1经济可行性**

(三级标题 宋体 四号 加粗 1.5倍行间距 段前段后各0.5行 左对齐)

**（1）工作量估算**

在软件开发阶段需要使用到的人力工作量百分比如下表2-1所示。

表2-1毕业设计过程管理系统各个开发阶段的人力百分比

（表格标题： 中文黑体 英文times new roman 五号 不加粗 表格上方 居中对齐）

|  |  |
| --- | --- |
| 任务 | 人力（%） |
| 可行性研究 | 5 |
| 需求分析 | 10 |
| 概要设计和详细设计 | 25 |
| 编码和测试 | 60 |
| 总计 | 100 |

**（2）成本估算**

在软件开发阶段需要其他一次性支出如下表2-2所示。

表2-2软件开发过程中各个开发阶段的一次性支出

|  |  |
| --- | --- |
| 项目 | 费用（元） |
| 系统前期需求研究 | 500 |
| 开发计划与测试基准研究 | 500 |
| 数据库的建立与数据字典 | 1500 |
| 检查费用和管理性费用 | 1500 |
| 培训费及软件开发人员所需的一次性支出 | 1000 |
| 总计 | 5000 |

表2-3软件开发过程中成本估算

|  |  |  |  |
| --- | --- | --- | --- |
| 项目 | 单价 | 数量 | 费用 |
| PC机 | 5000 | 2 | 10000 |
| 打印机 | 1000 | 1 | 1000 |
| 管理员工资 | 2000元/月 | 12 | 24000 |
| 不可预知费用 | \*\*\* | \*\*\* | 10000 |
| 一次性支出 | \*\*\* | \*\*\* | 5000 |
| 总计 | | 50000 | |

**（3）效益**

表2-4软件开发过程中各个开发阶段的效益

|  |  |
| --- | --- |
| 项目 | 收益（元/年） |
| 一次性收益 | 无 |
| 经常性收益（下载文献） | 80000 |
| 不可定量收益 | 无 |
| 企业定制服务 | 20000 |
| 总计 | 100000 |

**（4）收益/投资比**

一次性支出：26000元

经常性支出：24000元/年

收益：100000元/年

收益/投资比：100000\*5/(26000+24000\*5)=3.424657

**（5）货币的时间价值**

五年预计收益：100000\*5-（26000+24000\*5）=354000（元）

文献管理系统的前期软件开发阶段，需要耗费巨大的人力，投资的成本和未来获得的效益会随着软件每年的货币时间价值有所改变。文献管理系统五年收益为354000元。实际收益需要通过货币时间价值来进行核算。用利率的形式表示货币的时间价值。假设年利率为i，如果现在存入P元，则n年后可以得到的钱数为：

H:\temp\ksohtml14132\wps1.jpg (1)

这也就是P元钱在n年后的价值。反之，如果n年后能收入F元钱，那么这些钱的现在价值是：

H:\temp\ksohtml14132\wps2.jpg (2)

假定年利率为12%，利用上面计算货币现在价值的公式可以算出系统5年预计收益的现在价值，如下表2-5所示。

表2-5 将来收入折算成现在值

|  |  |  |
| --- | --- | --- |
| 将来值（元） | （1+i)^n | 现在值（元） |
| 354000 | 1.7623 | 200873 |

**（6）投资回收期估算**

文献管理系统基础建设成本约为26000元，第一年内收入为100000元，软件投资的回收期约为

 （3)

因此软件的投资回收期约为0.26年，本系统开发成本较低，维修成本适中，软件投资回报期较短可以较快获得利润，用户群体丰富且需求量大，经济利益客观值得投资。

### **2.2.2技术可行性**

技术可行性是指对系统的性能、可维护性、可靠性以及生产率等方面的信息进行评价。本系统所使用的技术包含：

（1）网络爬虫是一个从Web上自动下载网页的计算机程序。爬虫技术是一个可以连接数据和解析数据，并将这些数据进行分析，最后将结果可视化的工具，Python具有丰富的网络爬虫模块，且Python语言较为简单易懂易于学习。

（2）数据采集。由于携程的网页没有反爬虫策略，通过对网址进行分析，发现无论哪个城市的景点，或者用户的那一页进行评论，URL都是一样的。我们对Request Payload进行分析，发现每一页的参数是有变化的，可以断定是通过改变这些参数来改变接口返回的结果。

（3）数据整理。对于获取的数据，需要进一步处理，以满足分析分析的需求，我们主要做的数据处理步骤：删除重复项、删除不需要的列、处理不规范的数据格式。然后使用pyecharts，对热门城市景点的数据进行数据可视化分析。

综上所述，本系统涉及的技术难度较低，且有学习途径、性能优良、兼容性强、可靠性好、效率高、资源的要求低、维护方便。

### **2.2.3 操作可行性**

操作可行性是对开发系统在一个特定的工作环境中能否运行或运行的好坏程度的衡量。该系统采用前端web页面显示可视化结果，界面友好，操作简单，用户无需掌握相关技术，即可进行相关操作，故该系统具有操作可行性。

### **2.2.4 法律可行性**

可视化分析的代码和数据都是自己编写或取自网络开放数据，并不存在侵权等问题，且分析结果可以帮助旅游行业提供参考，具有很大的实际意义。因此这个系统的开发和设计，在法律上完全具有可行性。

## 2.3 可行性研究结论

由上述分析可知，该系统具有开发成本较低、技术难度一般、用户操作简便以及经济效益可观等特点，因此，对疫情下旅游行业生存状况可视化分析在技术、经济、操作和法律上都具有很高的可行性，开发此程序是可行的。

# 3 需求分析

概述段落

## 3.1 系统综合需求

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

精神文明日益发展的今天, 出行旅游成为人们的主要休闲方式和社会经济活动, 旅游业不仅能推动多个产业的发展, 调节产业结构, 同时也满足了人们日益增长的文化需要。其越来越成为国民经济的重要支撑。相关数据显示, 疫情平稳期间的旅游人数仍然稳中有增。而与之相对应的是海量的无序化旅游数据日益激增。针对有效旅游数据进行准确分析、 合理预测有助于制定与旅游发展相关的规划, 同时能有效地帮助后续在旅游设施分配和资源调度等方面上制定明智的决策, 发挥旅游信息价值的最大化。由此可知, 旅游经济的发展与对数据的合理分析是密不可分的。

## 3.2 数据流图分析

### **3.2.1 顶层数据流图**



图3-1 软件顶层数据流图

系统的源点和终点均为用户，用户输入需要分析的景点数据，经过系统爬取数据并可视化分析处理后，反馈给用户旅游信息的可视化分析结果。

### **3.2.2 数据流图细化与分解**

**（1）0层数据流图**

管理员具有以下功能：

1. 用户信息管理功能：对用户基本信息进行检索、增删改和权限修改；
2. 旅游数据系统中所有信息进行查询和增删改管理；
3. 快捷操作和帮助模块：一些系统的辅助操作，便于用户使用系统；

普通用户则可以通过登录该系统，实现旅游数据可视化结果的查询。



**图3-2软件0层数据流图**

**（2）1层数据流图**

在顶层数据流图和0层数据流图的基础上，将继续细化用户数据流图中各个处理过程，让读者更加清晰的了解到系统的处理过程，列出系统的1层数据流图如下：

用户登录管理加工包括用户登录、注册、找回密码、注销四个部分，旅游数据分析则包括旅游信息爬取、旅游数据分析、数据可视化三个部分，具体细节如下图所示。



**图3-3 软件1-1层数据流图**



**图3-4 软件1-2层数据流图**

**（3）2层数据流图**

在1层数据流图的基础上，对旅游数据分析和数据可视化加工进行进一步的分解，旅游数据分析可分为XX、XX、XX，数据可视化加工则可分解为原始结果展示、词云展示以及图表展示，具体细节如下图所示：



图3-5 软件2-1层数据流图



图3-6 软件2-2层数据流图

## 3.3 数据字典

### **3.3.1数据流条目定义**

实例：

表3-X 验证用户信息数据字典定义

根据3.2节数据流图分析情况可知，系统主要涉及的数据项包括：xx, xx, xx, xx和xx分别定义如图3-X~3-Y所示。

名称：验证用户信息

别名：无

描述：用户登录信息的输入

数据流组成：验证用户信息=用户名+密码+用户权限

数据流来源：用户

数据流去向：用户登录

（每个类型的数据字典定义前应有一段总结性的开场白。）

### **3.3.2数据存储条目定义**

实例：

表3-X 用户信息数据字典定义

名字：用户信息表

别名：用户信息

描述：唯一标识用户

定义：用户信息表=用户账号+用户密码+用户权限

位置：输入到电脑

### **3.3.3数据项条目定义**

实例：

表3-X 用户账号数据字典定义

名称：用户账号

别名：用户编号

含义：唯一确定用户身份编号

类型：数值

长度：11

取值范围：10000000000~19999999999

## 3.4实体联系分析

### **3.4.1 实体提取及实体图**

表3-18 实体提取结果

|  |  |  |
| --- | --- | --- |
| 实体 | 主码 | 其他属性 |
| 用户 | 用户编号 | 用户编号，用户密码，用户权限，用户类型 |
| 房屋产权人 | 姓名 | 姓名，联系方式 |
| 住房 | 地址 | 地址，价格，户型，面积 |

图 3-2 住房数据分析系统-用户属性图



图 3-3 住房数据分析系统-房屋产权人属性图



图 3-4 住房数据分析系统-住房属性图



### **3.4.2 实体联系图**



# 4软件设计

## 4.1 功能简述

根据前期需求分析可以得到数据分析系统主要有3大功能模块，用户登录管理功能模块、住房数据分析功能模块、使用帮助功能模块，每个子模块再调用更下层的模块从而完成程序的子功能，最下层的模块完成最基本的功能。



图4-1 住房分析系统的一级功能结构图

用户登录管理功能模块包括用户登录、用户注册、找回密码、用户注销四个部分。

1. 管理员可以对用户的基本信息进行检索、增删改和权限修改。
2. 用户则可以通过登录该系统，实现查询电影信息的服务。

（3） 用户注册时，系统将引导用户进行注册，不能使用已存在的用户名，注册成功后返回用户登录界面进行登录。当用户输入账号密码后，系统会与用户信息表中的信息进行比对，比对成功则登录成功。

（4） 登陆成功后，用户可以进行注销以及修改密码等操作，对比失败可以通过找回密码功能再次登录。



图4-2用户登录管理功能模块结构图

电影数据分析功能模块共分为电影信息爬取、电影信息存储、电影数据分析及数据可视化四个部分。

（1）电影信息爬取通过使用Python爬虫工具在各网站爬取大量相关数据，筛选出有效的数据，并进行信息存储。

（2）电影信息存储通过构建数据库来实现信息的安全存储，管理员可以对其进行管理。

（3）电影数据分析将存储在数据库内的数据进行分类整理和分析，将数据划分为许多个以不同主题小的数据集，为可视化提供数据来源。

（4）数据可视化将数据以各种形象的方式显示在用户面前，便于用户了解该电影的相关内容。



图4-3 电影数据分析功能模块结构图

电影数据分析可以分成电影信息分类、影评关键词提取、分析数据存储及同类电影特点分析四个部分。将整理好的电影信息进行不同主题的分类，并提取影评中的关键词，将这些信息存储整理并分析还可以对同类电影的特点进行总结分析，并将分析好的数据提供给数据可视化模块。



图4-4 电影数据分析结构图

电影信息可视化可以分成原始结果展示、图表展示及词云展示。将数据分析后的电影信息通过原始分析数据进行展示，原始分析数据使用更加形象的图表进行展示，也可以使用词云展示可视化的信息。



图4-5 数据可视化结构图

综上所述，将基于Python爬虫的电影数据分析系统进行模块化结构表示，并对每一个子模块在进行模块表示，构建出整体系统结构图，并实现完整的子功能，用户使用电影数据分析系统时结构更加清晰明了，并使用可视化进行图表词云等数据表示，更加方便用户使用该电影数据分析系统。

图4-6 基于Python爬虫的电影数据分析系统整体系统结构图

## 4.2 基于数据流分析构建软件结构

### **4.2.1面向数据流分析**

面向数据流的设计方法把信息流映射成软件结构，信息流的类型决定了映射的方法。

（1）参看图3-2，信息沿输入通路进入系统，同时由外部形式变换为内部形式，进入系统的信息通过变换中心，经加工处理以后再沿输出通路变换成外部形式离开软件系统，可见这些信息流具有交换流的总特征。

（2）确定输入流和输出流的边界，从而孤立出变换中心。

（3）完成各级分解，得到未经精化的软件结构图。



图4-7未经精化的输入结构



图4-8未经精化的变换结构



图4-9未经精化的输出结构

图4-10 基于Python爬虫的电影数据分析系统未经精化的软件结构图

### **4.2.2结构调整与优化**

由图3-6可知，电影信息可视化有明显的事务中心，可以将原本的输出结构调整为事务流。

由事务流映射成的软件结构包括一个接收分支和一个发送分支，映射出接收分支结构的方法和变换分析映射出输入结构的方法很相像，即从事务中心的边界开始，把沿着接收流通路的处理映射成模块。发送分支的结构包含一个调度模块，它控制下层的所有活动模块，然后把数据流图中的每个活动流通路映射成与他的流特征相对应的结构。



图4-11精化后的输出结构

### **4.2.3软件结构图**

图4-12 基于Python爬虫的电影数据分析系统精化后的软件结构图

## 4.3 模块接口、算法设计

### **4.3.1 人机界面**

人机交互过程中人机交互界面处于视感知层面的主要操作区域，是用户与系统直接接触的部分，交互界面的好坏决定着用户体验的直接感受。在人机界面设计的过程中，我们主要考虑以下四个方面的问题：

（1）系统响应时间

系统响应时间是交互式系统中用户抱怨最多的问题。如果系统响应时间过长，用户就会感到紧张和沮丧；但是，当用户工作速度是由人机界面决定的时候，系统响应时间过短也不好，这会迫使用户加快操作节奏，从而可能会犯错误。除了响应时间的绝对长短外，用户对不同命令在响应时间上的差别亦很在意，若过于悬殊用户将难以建立起稳定的工作节奏，在敏感条件下将会难以接受。保证响应时间的稳定，可以大大提高系统的属性。

（2）用户帮助设施

人机交互界面的使用使得每个用户都需要帮助，当遇到复杂问题时甚至需要查看用户手册以寻找答案，我们在设计数据分析系统时也要考虑用户遇到问题时能否很好的解决，提供用户帮助设施，使得用户无需离开用户界面就能解决自己的问题。而在具体设计帮助设施时，我们必须考虑用户在使用帮助设施时出现的问题。用户与系统交互期间，我们可以通过提供部分功能的帮助信息和提供全部功能的帮助信息，以保证用户在任何时候都能获得关于系统任何功能的帮助信息。除此以外，我们也可以通过帮助菜单，特殊功能键或者是HELP命令，帮助用户快速找到帮助入口。显示帮助信息时，可以选择在屏幕固定位置显示简短提示。在用户返回到正常的交互方式方面，可以提供屏幕上的返回按钮和功能键，让用户了解帮助后可以快速回到原界面。用户求助机制宜采用集成式，也避免叠加式系统导致用户求助某项指南而不得不浏览大量无关信息。

（3）出错信息处理

一般来说，由于用户的不熟悉或是误触等一系列操作，使得交互界面发出出错信息和警告信息几乎是必然的。对于错误和警告信息的选择，必须选用用户明了、含义准确的术语描述，同时还应尽可能提供一些有关错误恢复的建议，还有指出错误可能导致哪些负面后果(例如，破坏数据文件)，以便用户检查是否出现了这些问题，并在确实出现问题时方便解决。伴随着听觉上或视觉上的提示，也可以给用户更为直观的感受。信息更不能带有指责色彩，给用户被责怪的体验。

（4）命令交互

作为用户和系统软件交互最常用的方式，命令行依旧是许多高级用户的偏爱。数据分析系统选择菜单与键盘命令并存，在交互界面上提供用户更多的选择，提高系统的实用性。选择“命令宏机制”可以用自己定义的名字代表一个常用的命令序列，保证用户使用命令序列时无须依次输入每个命令，只需输入命令宏的名字就可以顺序执行它所代表的全部命令。除此以外，还要考虑大多数应用软件命令的使用方法，将命令普遍化，保证在软件切换过程中达成一致。

# 5 编码与测试

概述段落

## 5.1 编码

### **5.1.1 编码规则简介**

### **5.1.2代表性模块示例**

## 5.2测试

概述段落

### **5.2.1 白盒测试**

#### 5.2.1.1 xx模块独立路径分析

#### 5.2.1.2 xx模块测试用例设计

#### 5.2.1.3 xx模块测试情况分析

### **5.2.2 黑盒测试**

#### 5.2.2.1 xx模块的等价类分析

#### 5.2.2.2 xx模块测试用例设计

#### 5.2.2.3 xx模块测试情况分析

# 6 系统使用说明

概述段落

## 6.1 系统运行环境和配置

## 6.2 系统操作说明（按照结构图或层次图的框架依次介绍）

### **6.2.1 XX1模块说明**

### **6.2.2 XX2模块说明**

### **6.2.3 XX3模块说明**

### **6.2.4 XX4模块说明**

# 7 总结

# 参考文献

按国家标准GB3469执行。

# 致谢