

---

学号	10212818125
----	-------------

武汉华夏理工学院

# 课 程 设 计

课程名称 python 程序设计课程设计

题 目	<u>节假日长假景点人满为患的数据分析</u>
专 业	<u>软件工程</u>
班 级	<u>软件 1181</u>
姓 名	<u>伍泽蒙</u>
成 绩	<u></u>
指 导 老 师	<u>徐勇</u>

2021 年 1 月 11 日至 2021 年 1 月 22 日

---

# 武汉华夏理工学院信息工程学院

## 课程设计任务书

课程名称: python 程序设计课程设计

指导教师: 徐勇

班级名称: 软件 1181-1182

开课院、系: 计算机与网络工程系

### 一、课程设计目的与任务

“python 程序设计课程设计”是一个综合性的学习实践型实验教学环节,将在“python 程序设计”课程的授课基础上,对 python 的基础语法、python 对文件的操作、python 对数据的操作、数据处理基础、数据可视化等若干个知识点进行综合运用。

python 是一种具有天然开源基因的编程语言,了解开源社区和广泛的使用开源工具,也是 Python 学习实践的重要环节。因此在本次课程设计中,要求掌握 jupyter notebook、Git 等常用工具,以及 github 等重要开源社区的使用。

### 二、课程设计的内容与基本要求

数据分析的基本技术和方法在“python 程序设计”课程中已经进行完整的讲授,由于数据分析本身是基于业务场景的,因此本次课程设计环节更加偏重于实际的业务场景的实践。通过对近期互联网热点的调查,准备了 15 个具体的业务场景,用于本次课程设计的具体任务场景。具体包括如下:

- 1) 北上广深租房状况分析;
- 2) 蔡某坤粉丝数及转发数据真假状况分析;
- 3) 地震的数据分析;
- 4) 英文名字的数据分析;
- 5) 外籍英文老师收入虚高情况数据分析;
- 6) 我国城市空气污染和烟花燃放的关系分析;
- 7) 针对 996 工作,程序员群体的看法的分析;
- 8) 吴某凡微博热点的分析;
- 9) 节假日长假景点人满为患的数据分析;
- 10) 针对荔枝的品种、销售地等维度,进行价格数据分析;
- 11) 分析芒果 TV《我是大侦探》的观众评论数据;
- 12) 针对当前儿科医生的缺乏,对相关数据进行分析;
- 13) 著名网游《绝地求生》的数据分析;
- 14) 实习岗位状况的数据分析;
- 15) 电影《流浪地球》的观众评价的数据分析。

每三个学生组建一个课程设计小组,最后的任务输出包括代码每组一份、课程设计

报告每人一份、答辩 ppt 每小组一份,并进行课程设计成果答辩。小组成员均参与前述工作,但是每个人的侧重点不同。

每个课程设计小组可以从上述 15 个场景中选取一个作为课程设计的选题,选题中提供了待分析的数据,和现有的分析方法。各小组,通过学习和实践现有的分析方法,理解实战分析的思维过程并锻炼实际动手能力,再此基础上可以扩展更多维度的分析和数据展现形式。

每个班每个选题最多只能被两个小组选中,先选先得。同时,如果各小组发现更有意思的场景,并能够获取到相关待分析的数据,也可以申请作为选题方向。

本次课程设计的目标是培养学生的团队协作能力、对 python 知识点的综合运用、对实际场景的理解和适应能力、针对答辩的表达能力等。注重过程,期待成果,但不强求结果的尽善尽美。

### 三、学时分配进度安排

序号	设计内容	所用时间
1	下发任务书,学生查阅资料	1 天
2	组建团队,并进行选题和团队匹配	1 天
3	各小组对自己的选题开始进行研究	2 天
4	各小组开始准备数据并形成初步处理意见	1 天
5	代码和实现分析的动手实践	2 天
6	形成初步的报告书和 ppt	1 天
7	答辩并完成报告书	2 天
合 计		2 周

### 四、课程设计考核及评分标准

#### 1. 设计报告要求

课程设计报告要求逻辑清晰、层次分明、书写整洁。课程设计报告为每人一份,同一个小组的各成员的整体报告内容结构一致,但个人侧重点不同,个人着重撰写自己的工作内容,其他人的内容只要体现文档结构的完整性即可。

课程设计考核将综合考虑学生考勤和参与度、团队协作能力,过程管理能力、成果达成情况等。

#### 2. 过程要求

整个过程要求通过 github 来进行过程化发布,即阶段性地在 github 上提交过程结果。

#### 3. 评分标准

评分依据	评分成绩
1. 团队协作能力	25 分
2. python 综合运用能力	25 分
3. 态度认真、刻苦钻研、遵守纪律	10 分
4. 过程完成、对工具的使用、对 github 的运用	20 分

5. 课程设计答辩逻辑清晰，内容正确	10 分
6. 课程设计期间的课堂考勤、创新能力	10 分
总分	100 分

注：按上述六项分别记分后求和，根据小组成员贡献率综合评定，记载个人最后成绩。

成绩等级：优（90分—100分）、良（80分—89分）、中（70分—79分）、及格（60分—69分）、60分以下为不及格。

## 五、指导地点与时间

本课程设计将安排在第 20-21 周，采用腾讯会议和 QQ 群的方式，以在线形式进行。具体安排如下：

周次	星期一	星期二	星期三	星期四	星期五
第 20 周	第 5-8 节	第 5-8 节		第 5-8 节	
第 21 周		第 5-8 节		第 5-8 节	

执笔:徐 勇      日期: 2021-1-8

审阅:钱小红      日期: 2021-1-8

---

# 目录

1. 绪言 .....	1
1.1 研究背景.....	1
1.2 研究对象.....	1
2. 编程环境与相关技术.....	1
2.1 硬件平台.....	1
2.2 软件平台.....	1
2.2.1 项目托管平台-GitHub 网站.....	1
2.2.2 jupyter 运行平台.....	2
2.2.3 编程语言-python.....	2
3. 程序设计.....	3
3.1 数据分析库.....	3
3.1.1 数据分析库图导入.....	3
3.1.2 数据分析库注解.....	3
3.2 数据处理.....	4
3.2.1 数据读取.....	4
3.2.2 字段保留.....	4
3.2.3 去重.....	5
3.2.4 类型转换.....	5
4. 问题分析.....	5
4.1 哪些城市的旅游选择最多? .....	5
4.2 哪些省份的旅游选择最多? .....	6
4.3 哪些城市最受游客青睐? .....	7
4.4 最热门的景点? .....	7
4.5 各省份的好评热门景点.....	8
4.6 很热门但是评分不好景点? .....	9
4.7 打折力度最大的景点? .....	9
4.8 各级别景区的门票? .....	10
5. 总结 .....	11
参考文献.....	12

---

# 1. 绪言

## 1.1 研究背景

随着国内经济的发展，近年来在假日旅游中散客旅游上升趋势明显，散客尤其是“家庭型”、“亲朋型”上升态势更为明显，占整个出游人数的 60%以上。人民在追求物质生活的基础上增加了对精神文化生活的追求。科学技术的进步使移动端用户在线订票和选择旅游地变得更加方便，国内旅游市场稳步上升。在当今社会下旅游成为大多数人们缓解工作、生活的压力的一种很普遍的方式。旅游地及景点的多样化使人们多了很多选择。

## 1.2 研究对象

本课程设计是对 2019 年各个城市/省份的五一节日长假人满为患的数据进行分析、统计分析热门景点、打折力度最大、哪些城市最受青睐、各级别景区的门票价格，观察这一年其最受青睐的城市，整理出来的数据，进行对比选出相对来说比较划算且有趣的旅游景点。数据的主要来源是网上公布的数据。

# 2. 编程环境与相关技术

## 2.1 硬件平台

一台 64 位 Windows 系统的计算机。

## 2.2 软件平台

### 2.2.1 项目托管平台-GitHub 网站

GitHub 是最大的开源代码托管平台，旨在促进在一个共同项目上工作的个人之间的代码托管、版本控制和协作。通过该平台，无论何时何地，都可以对项目进行操作（托管和审查代码，管理项目和与世界各地的其他开发者共同开发软件），GitHub 平台为开源项目和私人项目都提供了项目处理功能。

GitHub 于 2008 年 4 月 10 日正式上线，除了 Git 代码仓库托管及基本的 Web 管理界面以外，还提供了订阅、讨论组、文本渲染、在线文件编辑器、协作图谱（报表）、代码片段分享（Gist）等功能。目前，其注册用户已经超过 350 万，托管版本数量也是非常之多，其

---

中不乏知名开源项目 Ruby on Rails、jQuery、python 等。

GitHub 可以托管各种 git 库，并提供一个 web 界面，但它与外国的 SourceForge、Google Code 或中国的 coding 的服务不同，GitHub 的独特卖点在于从另外一个项目进行分支的简易性。为一个项目贡献代码非常简单：首先点击项目站点的“fork”的按钮，然后将代码检出并将修改加入到刚才分出的代码库中，最后通过内建的“pull request”机制向项目负责人申请代码合并。已经有人将 GitHub 称为代码玩家的 MySpace。

GitHub 项目本身自然而然的也在 GitHub 上进行托管，只不过在一个私有的，公共视图不可见的库中。开源项目可以免费托管，但私有库则并不如此。Chris Wanstrath，GitHub 的开发者之一，肯定了通过付费的私有库来在财务上支持免费库的托管这一计划。

## 2.2.2 jupyter 运行平台

Jupyter Notebook 是一款开源的 web 应用，支持 markdown 编辑，可以使用 Jupyter Notebook 来创建一个技术分享的文档。Jupyter Notebook 最大的特点是：可以插入 Python 代码，并且还能运行代码，直接将运行结果展示在文档中。

Jupyter Notebook 是 IPython 项目的一个衍生项目，该项目以前曾经拥有 IPython Notebook 项目。Jupyter 的名称来自它支持的核心支持的编程语言：Julia，Python 和 R。Jupyter 随 IPython 内核一起提供，它允许您使用 Python 编写程序，但是目前有 100 多个其他内核也可以使用。

即 Jupyter Notebook 中所有交互计算、编写说明文档、数学公式、图片以及其他富媒体形式的输入和输出，都是以文档的形式体现的。

这些文档是保存为后缀名为 .ipynb 的 JSON 格式文件，不仅便于版本控制，也方便与他人共享。此外，文档还可以导出为：HTML、LaTeX、PDF 等格式。

## 2.2.3 编程语言-python

python 是一种脚本语言，写好了就可以直接运行，省去了编译链接的麻烦，对于需要多动手实践的初学者而言，也就是少了出错的机会。而且 python 还有一种交互的方式，如果是一段简单的小程序，连编辑器都可以省了，直接敲进去就能运行。

python 是一种清晰的语言，用缩进来表示程序的嵌套关系可谓是一种创举，把过去软性的编程风格升级为硬性的语法规则。再不需要在不同的风格间选择、再不需要为不同的风格争执。python 中没有各种隐晦的缩写，不需要去强记各种奇怪的符号的含义。python 写的程序很容易懂，这是不少人的共识。

python 是一种面向对象的语言，但它的面向对象却不象 C++ 那样强调概念，而是更注重实用。它用最简单的方法让编程者能够享受到面向对象带来的好处，这正是 Python 能像 Java、

---

C#那样吸引众多支持者的原因之一。

python 是一种功能丰富的语言，它拥有一个强大的基本类库和数量众多的第三方扩展。python 为程序员提供了丰富的基本功能使得人们写程序时用不着一切最底层做起。

人们通常会用一种担心：脚本语言通常很慢。但 python 的速度却比人们想象得快很多。虽然 python 是一种脚本语言，但实际上也可以对它进行编译，省去了对程序文本的分析解释，速度自然提升很多。

## 3. 程序设计

### 3.1 数据分析库

#### 3.1.1 数据分析库图导入

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from pandas.io.json import json_normalize
import pymysql
```

图 3-1

#### 3.1.2 数据分析库注解

**Pandas** 是 python 的一个数据分析包，最初由 AQR Capital Management 于 2008 年 4 月开发，并于 2009 年底开源出来，目前由专注于 Python 数据包开发的 PyData 开发 team 继续开发和维护，属于 PyData 项目的一部分。**Pandas** 最初被作为金融数据分析工具而开发出来，因此，pandas 为时间序列分析提供了很好的支持。**Pandas** 的名称来自于面板数据（panel data）和 python 数据分析（data analysis）。panel data 是经济学中关于多维数据集的一个术语，在 **Pandas** 中也提供了 panel 的数据类型。

**NumPy** 一个用 python 实现的科学计算，包括：1、一个强大的 N 维数组对象 Array；2、比较成熟的（广播）函数库；3、用于整合 C/C++ 和 Fortran 代码的工具包；4、实用的线性代数、傅里叶变换和随机数生成函数。numpy 和稀疏矩阵运算包 scipy 配合使用更加方便。**NumPy**（Numeric Python）提供了许多高级的数值编程工具，如：矩阵数据类型、矢量处理，以及精密的运算库。专为进行严格的数字处理而产生。多为很多大



型金融公司使用，以及核心的科学计算组织如：Lawrence Livermore，NASA 用其处理一些本来使用 C++，Fortran 或 Matlab 等所做的任务。NumPy 的前身为 Numeric，最早由 Jim Hugunin 与其它协作者共同开发，2005 年，Travis Oliphant 在 Numeric 中结合了另一个同性质的程序库 Numarray 的特色，并加入了其它扩展而开发了 NumPy。

NumPy 为开放源代码并且由许多协作者共同维护开发。

**Matplotlib** 是一个 Python 的 2D 绘图库，它以各种硬拷贝格式和跨平台的交互式环境生成出版质量级别的图形。通过 Matplotlib，开发者可以仅需要几行代码，便可以生成绘图，直方图，功率谱，条形图，错误图，散点图等。

使用 `%matplotlib` 命令可以将 matplotlib 的图表直接嵌入到 Notebook 之中，或者使用指定的界面库显示图表，它有一个参数指定 matplotlib 图表的显示方式。inline 表示将图表嵌入到 Notebook 中。

`%matplotlib inline` Seaborn 是一种基于 matplotlib 的图形可视化 python library。它提供了一种高度交互式界面，便于用户能够做出各种有吸引力的统计图表。

**Seaborn** 是基于 matplotlib 的图形可视化 python 包。它提供了一种高度交互式界面，便于用户能够做出各种有吸引力的统计图表。

Seaborn 是在 matplotlib 的基础上进行了更高级的 API 封装，从而使得作图更加容易，在大多数情况下使用 seaborn 能做出很具有吸引力的图，而使用 matplotlib 就能制作具有更多特色的图。应该把 Seaborn 视为 matplotlib 的补充，而不是替代物。同时它能高度兼容 numpy 与 pandas 数据结构以及 scipy 与 statsmodels 等统计模式。

## 3.2 数据处理

### 3.2.1 数据读取

```
In [4]: #直接从本地读取csv数据
df=pd.read_csv(r'D:\data.csv')
```

图 3-2

### 3.2.2 字段保留

```
#保留的字段data.columns = ['城市', '评论', '折扣价格', '产品特点', 'ID',
#                             '评分', '纬度', '经度', '原价',
#                             '城市', '年销量', '月销数字', '特征',
#                             '标题', '类型']
col_keep = ['city', 'fields.comment', 'fields.discountPrice', 'fields.features', 'fields.itemId',
            'fields.itemTotalScore', 'fields.latitude', 'fields.longitude', 'fields.price',
            'fields.shortInfo', 'fields.sold365', 'fields.soldRecentNum', 'fields.tagList',
            'fields.title', 'trip_main_business_type']
df = df[col_keep]
```

图 3-3

### 3.2.3 去重

如图这里通过 `pandas.DataFrame.drop_duplicates` 函数实现将字段” `fields.itemId`” 重复的去除。

```
df.drop_duplicates(subset='fields.itemId', inplace=True)
#代码中subset对应的值是列名, 表示只考虑这一列, 将这两列对应值相同的行进行去重。
#默认值为subset=None表示考虑所有列。
#inplace=True表示直接在原来的DataFrame上删除重复项, 而默认值False表示生成一个副本。
```

图 3-4

### 3.2.4 类型转换

如图先筛选出字段” `fields.comment`” 为 `np.nan` 的数据, 然后将此字段转换为 `float` 类型, 然后将字段” `fields.discountPrice`” 和字段” `fields.price`” 转换为 `float` 类型”, 将” `fields.itemId`” 和” `fields.soldRecentNum`” 转换为 `int` 类型

```
# 由于np.nan是float, 所以这里整体转换为float
df.loc[df['fields.comment'].isnull(), 'fields.comment'] = np.nan
df['fields.comment'] = df['fields.comment'].astype(float)

df['fields.discountPrice'] = df['fields.discountPrice'].astype(float)
df['fields.itemId'] = df['fields.itemId'].astype(int)
df['fields.price'] = df['fields.price'].astype(float)
df['fields.soldRecentNum'] = df['fields.soldRecentNum'].astype(int)
```

图 3-5

## 4. 问题分析

### 4.1 哪些城市的旅游选择最多?

如下图, 根据数据的分析, 生成对比图可以看出排行首位的城市是三亚, 其次是杭州、北京、桂林、上海、广州、厦门、苏州、香港、成都、重庆、清远。我想选择三亚的人最多可能是因为三亚景点众多还四季如春的原因吧, 其他城市基本不是经济较为发达就是风景比较宜人的城市, 这些城市自然而然就成了旅选择最多的了。

```
# pyecharts v1版
from pyecharts.charts import Funnel
from pyecharts.faker import Faker
from pyecharts import options as opts
funnel=(Funnel(init_opts=opts.InitOpts(page_title="旅游选择最多的前12个城市"))
        .set_global_opts(title_opts=opts.TitleOpts(title="旅游选择最多的前12个城市"))
        .add("",[list(z) for z in zip(count_top12.index.tolist(),count_top12.values.tolist())])
        )
funnel.render("旅游选择最多的前12个城市.html")
funnel.render_notebook()
```

图 4-1

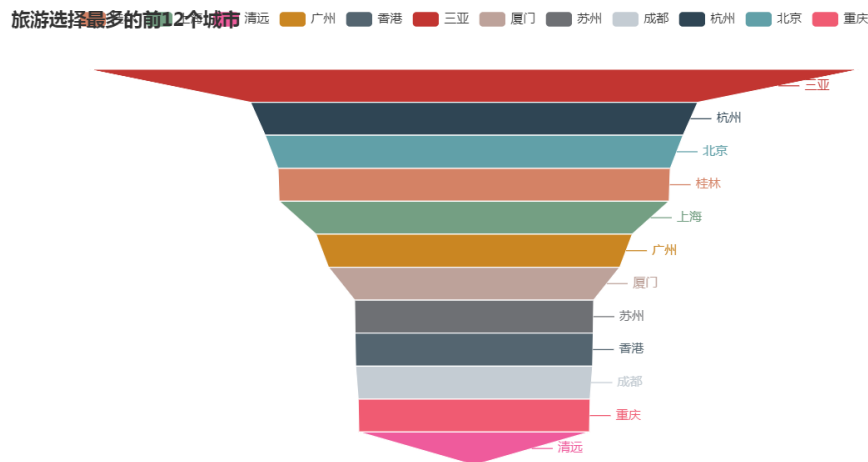


图 4-2

## 4.2 哪些省份的旅游选择最多？

```
In [39]: # pyecharts v1版
from pyecharts.charts import Map#地图
map=(Map(init_opts=opts.InitOpts(page_title="全国各省份旅游选择数量图"))
      .add("", [list(z) for z in zip(province, count_province_top10.values.tolist())], "china")
      .set_global_opts(title_opts=opts.TitleOpts(title="全国各省份旅游选择数量图"),visualmap_opts=opts.VisualMapOpts(max_=5000))
      )
map.render("全国各省份旅游选择数量图.html")
map.render_notebook()
```

Out[39]: 全国各省份旅游选择数量图

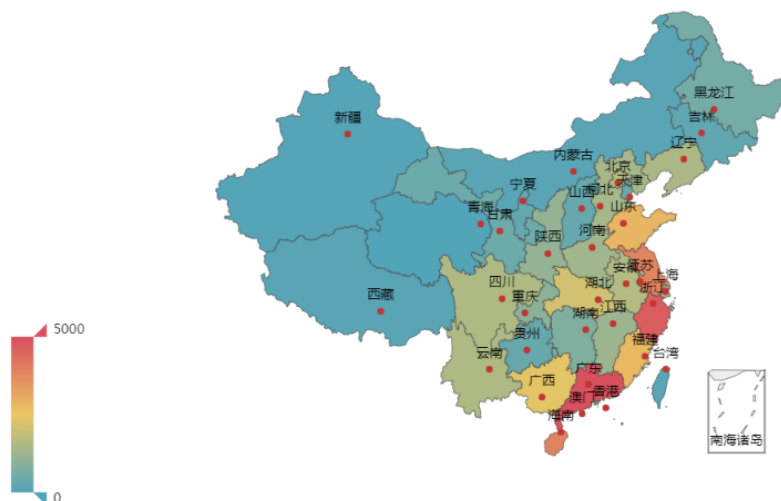


图 4-3

### 4.3 哪些城市最受游客青睐？

通过下图词云图显示出平均评分最高的 20 个城市。词海中的字越大评价越好，可以看出旅游的人们对武汉，海口的评价颇高。通过结合“选择旅游最多的城市”及“评分最高的城市”发现杭州、澳门、苏州等地在游客的选择和评价来说都是较高的。



图 4-4

### 4.4 最热门的景点？

通过城市最近一个月售出门票数量和城市景点的平均分这两个指标来确定可见杭州灵隐飞来峰和上海迪士尼遥遥领先。这可能也是杭州和上海城市排行位居前面的原因吧。

最热门的10个景点

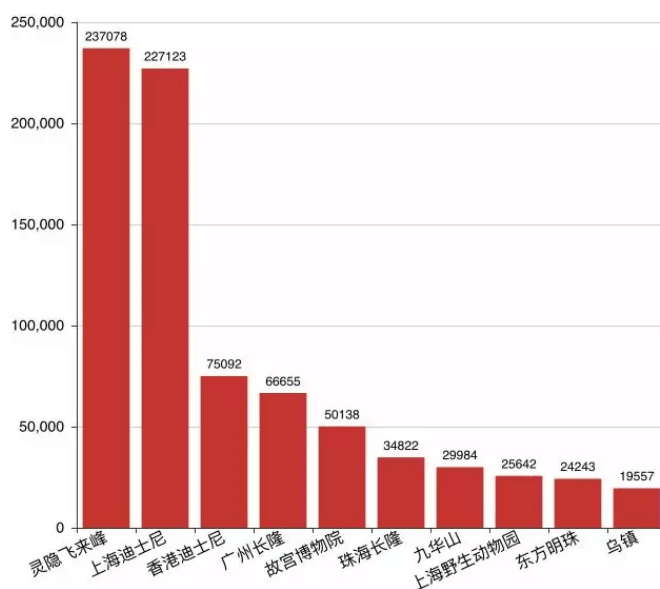


图 4-5

## 4.5 各省份的好评热门景点

	省份	项目名	销售数量	评分
0	上海市	✔上海迪士尼乐园门票 上海迪士尼乐园门票1日票 迪斯尼	9022	5.0
1	云南省	昆明青青湾水上乐园门票 青青湾【现买现用·秒出票】一票畅玩	543	5.0
2	内蒙古自治区	【超值体验】内蒙古旅游鄂尔多斯草原演绎套票	16	5.0
3	北京市	【官方直营】北京古北水镇国际旅游度假区全天门票-散客全票	4100	5.0
4	台湾	雄狮旅游 台湾台北国立故宫博物院门票二维码电子票儿童免票	123	5.0
5	吉林省	I【自动出票】吉林长春伪满皇宫博物院门票 博物馆门票	125	5.0
6	四川省	✔【即买即用】成都熊猫基地门票 大熊猫繁育研究基地大门票	5491	5.0
7	天津市	【官方门票】天津方特欢乐世界门票 刷身份证入园	481	5.0
8	宁夏回族自治区	I【当天可订】宁夏银川镇北堡西部影视城大门票 华夏西部影城	351	5.0
9	安徽省	【当天可订】宏村门票 安徽黄山宏村优惠票 宏村团队 黄山旅游	3070	5.0
10	山东省	【免取票电子票】山东德州济南泉城海洋极地世界旅游景区景点门票	3613	5.0
11	山西省	大同云冈石窟大门票成人票	599	5.0
12	广东省	五一特惠 当天 广州长隆国际大马戏动物马戏团门票成人女士一等座	2919	5.0
13	广西壮族自治区	【官方门票】南宁方特东盟神画乐园门票 1票畅玩东盟十国方特门票	759	5.0
14	新疆维吾尔自治区	【当天可订】新疆 乌鲁木齐 天山大峡谷 新疆旅游 天山大峡谷L	13	5.0
15	江苏省	【当天可订 自动出票】南京总统府景区门票南京总统府门票电子票	6924	5.0
16	江西省	江西 南昌 滕王阁门票 滕王阁门票+电子导览	630	5.0
17	河北省	清西陵景区 套票	216	5.0
18	河南省	【官方旗舰店】河南开封清明上河园/白天门票/大门票	3083	5.0
19	浙江省	【当天可订 官方品质】杭州西湖灵隐飞来峰（不含灵隐寺）门票	117729	5.0
20	海南省	免抱佛脚随时出票三亚南山寺门票+电瓶车南海观音南山海南旅游	4080	5.0
21	湖北省	【官方旗舰店】武汉欢乐谷 成人/儿童/青少年/大学生票 全日票	1796	5.0
22	湖南省	【官方门票当日可订】株洲方特梦幻王国门票身份证入园极速出票	590	5.0
23	澳门	扫码即用澳门新濠影汇酒店8字摩天轮八字摩天轮电子门票旅游之星	1090	5.0
24	甘肃省	【官方】敦煌鸣沙山月牙泉门票 敦煌旅游景区成人电子门票	504	5.0
25	福建省	【官方旗舰店】菽庄花园门票成人票 厦门鼓浪屿景区景点门票飞猪	4576	5.0
26	西藏自治区	西藏拉萨旅游文成公主大型实景剧演出门票表演票市区游可选接送	279	5.0
27	贵州省	【官方预定】铜仁梵净山门票+含往返环保车+往返索道 贵州大门票	540	5.0
28	辽宁省	【四馆套票】辽宁大连圣亚海洋世界公园极地馆旅游景区景点门票	4049	5.0
29	重庆市	【赠10元餐券】重庆黑山谷门票 重庆万盛区黑山谷景区大门票HP	1108	5.0
30	陕西省	西安旅游 西安兵马俑博物馆门票 兵马俑门票兵马俑大门票秦始皇陵	5307	5.0
31	青海省	青海景点门票青海高原野生动物园西宁野生动物园西藏动物园门票	52	5.0
32	香港	【当天可订】香港迪士尼三合一餐券午餐+晚餐+小食香港迪士尼门票	1605	5.0
33	黑龙江省	【官方直营】哈尔滨极地馆大门票（跟着企鹅看表演）	159	5.0

图 4-6

## 4.6 很热门但是评分不好景点？

	fields.shortInfo	fields.title	fields.soldRecentNum	fields.itemTotalScore
2545	大连	【领券立减】大连二〇三樱花园大门票旅顺203樱花成人票扫码入园	118	4.3
7124	无锡	【当天可定】无锡善卷洞门票善卷洞游船联票大门票成人电子票	760	4.2
8317	杭州	浙江杭州临安风之谷景区门票 户外昌化风之谷门票 杭州接送跟车	199	4.3
22782	武汉	【汉口粤汉码头登船】武汉两江游览船票B线（不含餐）夜游长江游	109	4.3
26939	深圳	ZY【当天可定即定即用】深圳观澜山水田园娱乐套票A 含景区大门票	155	3.8
29856	佛山	顺德长鹿农庄门票 休博园 大门票+游乐金+动物园 单人票/双人票	1001	4.4
29877	佛山	长鹿农庄门票休博园大门票 套票动物园2大2小（不含矿洞海洋馆）	210	4.3
40353	西安	陕西 西安 西安城墙景区 大门票+电子导览	221	4.4
42576	北京	北京昌平第七届北京农业嘉年华大门票 农业嘉年华门票 成人票	112	4.4
46919	香港	Dm香港迪士尼乐园门票1日迪斯尼香港旅游自由行景点电子票不刷单	261	4.3

图 4-7

## 4.7 打折力度最大的景点？

门票折扣价最低的是 2.7 折，是位于广东 东莞的“南社明清古村落”，折后仅需 13.5 元就能够购买门票。优惠力度最大的是珠海的“海泉湾海洋温泉”，打折前为 1980.0 元，折后为 589.0 元，推荐有意向去的旅客可以考虑购买。

	fields.discountPrice	fields.price	discount	fields.title	fields.shortInfo
28979	13.5	50.00	2.7	ZM【当天可定】广东 东莞南社明清古村落 成人儿童老人观光大门票	东莞
22713	180.0	630.00	2.9	【大学生3人特惠】武汉海昌极地海洋公园门票 海洋世界萌宠乐园	武汉
29456	598.0	1980.00	3.0	【598元/10次卡】超值钜惠 珠海海泉湾海洋温泉10次卡 可4人使用	珠海
37355	18.0	53.10	3.4	【双人特价 电子票】成都大熊猫基地门票成都大熊猫繁育研究基地	成都
27530	34.5	100.00	3.4	ZY【提前1小时下单】广东 清远田野绿世界 成人儿童老人大门票	清远
30308	36.0	100.00	3.6	ZM【当天可定】广东 阳江闸坡大角湾景区 成人1日景区大门票	阳江
26931	60.0	166.50	3.6	即买即用 索道+缆车】深圳东部华侨城大侠谷门票 大峡谷亲子家庭	深圳
1004	49.9	135.00	3.7	当日可订 河北 秦皇岛 沙雕海洋乐园门票 沙雕大世界大门票	秦皇岛
30549	20.0	52.99	3.8	【即买即用】南宁园博园门票 第十二届园博会园博园门票	南宁
30541	21.4	55.00	3.9	【即买即用】南宁园博园门票 第十二届园博会园博园门票	南宁
30426	40.0	100.00	4.0	ZM【当天可定】提前1小时下单 广东茂名浪漫海岸成人儿童情侣门票	茂名
28929	24.0	60.00	4.0	【当天可定 即定即用】广东 东莞梦幻百花洲 成人儿童老人大门票	东莞
42044	138.0	336.00	4.1	银川天山海世界·黄河明珠五一特惠双人票	银川
6922	35.0	83.00	4.2	【当天可订】江苏无锡鼋头渚门票+客渡船门票（含园内公交车）	无锡
313	69.9	168.00	4.2	【官方旗舰店】石家庄天山海世界69.9元五一特惠票	石家庄

图 4-8

## 4.8 各级别景区的门票？

```
In [69]: # pyecharts v1版
bar=(
    Bar(init_opts=opts.InitOpts(page_title="各级别景区的门票价格"))
    .set_global_opts(title_opts=opts.TitleOpts(title="各级别景区的门票价格"))
    .add_xaxis(level_data.index.tolist())
    .add_yaxis("", np.round(level_data.values,0).tolist())
)
bar.render("各级别景区的门票价格.html")
bar.render_notebook()
```

Out[69]: 各级别景区的门票价格

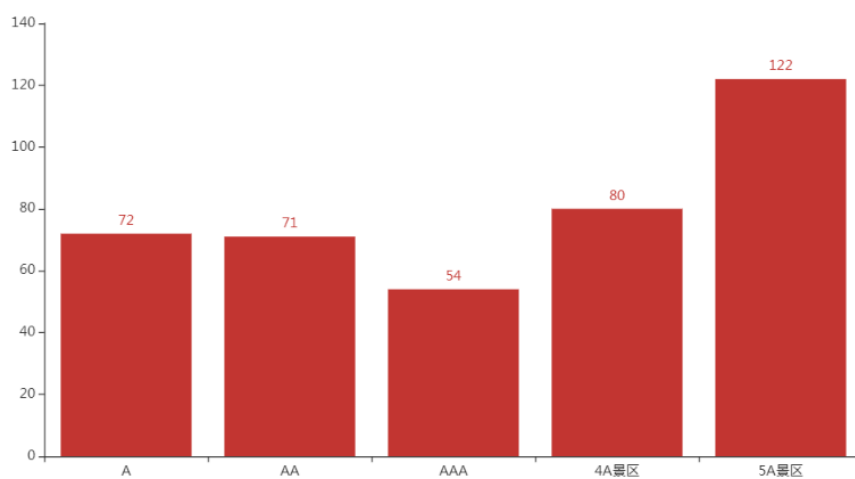


图 4-9

---

## 5. 总结

两周的课程设计结束了，在这次的课程设计中不仅检验了我所学习的知识，也培养了我如何去把握一件事情，如何去做一件事情，又如何完成一件事情。在设计过程中，与同学分工设计，和同学们相互探讨，相互学习，相互监督。学会了合作，这里的感谢队友对我的包涵，耐心给我讲解我不会的东西，还需要感谢老师每次准时上课为我们讲解，虽然是线上老师也依然陪我们到下课时间，成就了这次课设的圆满结束。

课程设计诚然是一门专业课，给我很多专业知识以及专业技能上的提升，通过这次课程设计，我掌握了 `GitHub` 网站的熟练应用，掌握了 `jupyter notebook` 的使用，和对 `Python` 语言语句的了解更加清楚，以前对数据清洗不是很明白通过这次的课设对数据清洗了解额更加透彻了。

回顾起此课程设计，至今我仍感慨颇多，从理论到实践，在这段日子里，可以说得是苦多于甜，但是可以学到很多很多的东西，同时不仅可以巩固了以前所学过的知识，而且学到了很多在书本上所没有学到过的知识。通过这次课程设计使我懂得了理论与实际相结合是很重要的，只有理论知识是远远不够的，只有把所学的理论知识与实践结合起来，从理论中得出结论，才能真正为社会服务，从而提高自己的实际动手能力和独立思考的能力。在设计的过程中遇到问题，可以说得是困难重重，但可喜的是最终都得到了解决。



---

## 参考文献

- [1] 丁辉. Python 基础与大数据应用. 人民邮电出版社
- [2] 刘宇宙. python3.5 从零开始学[M]. 北京: 清华大学出版社, 2017.
- [3] <https://mp.weixin.qq.com/s/iuCNreCuKzrggdXtvurpkQ>
- [4] <https://www.baidu.com>

## 课程设计成绩评定表

课程设计题目	节假日长假景点人满为患的数据分析	
<p>课程设计学生答辩或质疑记录：</p> <p>1.%matplotlib inline 的含义？</p> <p>答：用在 Jupyter notebook 中具体作用是当你调用 matplotlib.pyplot 的绘图函数 plot() 进行绘图的时候，或者生成一个 figure 画布的时候，可以直接在你的 python console 里面生成图像。</p> <p>2.简述 Seaborn 的功能。</p> <p>答：Seaborn 是基于 matplotlib 的图形可视化 python 包，其实是在 matplotlib 的基础上进行了更高级的 API 封装，从而使得作图更加容易，在大多数情况下使用 seaborn 就能做出很具有吸引力的图，而使用 matplotlib 就能制作具有更多特色的图。</p> <p>3.用的什么版本的 pyecharts？</p> <p>答：pyecharts 是用于生成一个 echart 图表的类库，需要用 1.9.0 的版本就能兼容，图就能正常出来了。</p>		
评 分 依 据	分 值	评分成绩
1. 团队协作能力	25 分	
2. python 综合运用能力	25 分	
3. 态度认真、刻苦钻研、创新能力	10 分	
4. 过程完成、对工具的使用、对 github 的运用	20 分	
5. 课程设计答辩逻辑清晰，内容正确	10 分	
6. 课程设计期间的课堂考勤、遵守纪律	10 分	
总 分	100 分	
<div style="display: flex; justify-content: space-between; align-items: flex-end;"> <div>最终评定等级为：</div> <div>             指导老师签字：_____             <div style="text-align: right; margin-top: 10px;">                 2021    年    1 月    15 日             </div> </div> </div>		