|  |  |
| --- | --- |
| 学号 | **10214518415** |



课 程 设 计

课程名称 python程序设计课程设计

|  |  |
| --- | --- |
| 题 目 | 数据分析的学习与实践-PUBG |
| 专 业 | \_\_ 软件工程\_\_\_\_ \_\_ |
| 班 级 | \_\_\_\_\_\_\_ 软件1182\_\_\_\_\_\_\_\_\_\_\_ |
| 姓 名 | \_\_\_\_\_\_\_\_ 夏可壮\_\_\_\_\_\_\_\_\_\_\_\_ |
| 成 绩 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| 指 导 老 师 | \_\_\_\_\_\_\_\_\_\_\_徐勇\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |

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小组一份、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 需求分析 6](#_Toc62227389)

[2.开发环境 6](#_Toc62227390)

[3.分析流程 6](#_Toc62227391)

[4.分析实施 6](#_Toc62227392)

[4.1数据获取 6](#_Toc62227393)

[4.2数据预处理 7](#_Toc62227394)

[4.3数据分析 7](#_Toc62227395)

[4.2.1：开局应该猥琐拾物资还是应该刚枪舔包？ 7](#_Toc62227396)

[4.2.2武器该怎么选择，什么武器更容易杀人？ 8](#_Toc62227397)

[4.2.3远处狙击，适合的武器 10](#_Toc62227398)

[4.2.5近距离作战适合用什么武器 12](#_Toc62227399)

[5.自我评价与总结 15](#_Toc62227400)

[参考文献 16](#_Toc62227401)

1 需求分析

绝地求生是前段时间大火的游戏，相信很多人都有接触过。同时腾讯公司也出了一款类似的游戏“刺激战场”，可见其当时的热度。但是，既然是游戏，就肯定有输有赢。怎样才能做到驰骋沙场，不败金身呢？

分析了70多万场绝地求生的比赛数据，总结出独家吃鸡攻略

分析了总共20多G、70多万场的绝地求生比赛数据，总结出各种地图哪里跳的人最多、近战最佳武器、狙击最佳武器等。

2.开发环境

操作系统：windows 10

开发工具：jupyter notebook

3.分析流程

数据获取:找出需要用作数据分析的数据并进行获取。

数据预处理:对获取数据进行去重、缺失值处理等一系列数据清洗操作。

数据可视化:文字不如数据，数据不如图片，可视化是对于数据最好的展现方式，让人一目了然。

深入分析:对于数据深入分析，寻找不同因素之间与薪资之间的关系。

4.分析实施

## 4.1数据获取

数据老师已提供

## 4.2数据预处理

首先进行的是数据清洗。然后进行重复项删除，最后只留下唯一的数据。

agg1 = pd.read\_csv('D:\py/agg\_match\_stats\_1.csv')*# 丢弃重复数据*

agg1.drop\_duplicates(inplace=True)

## 4.3数据分析

### 4.2.1：开局应该猥琐拾物资还是应该刚枪舔包？

争对这个问题，我们分析了在游戏里面，击杀人数和吃鸡的对应关系：

一局游戏总人数共100人，我们以40人为临界值，进行筛选。并画出柱状图：

通过图像很明显的可以看到，随着击杀人数的提升，吃鸡概率在不断提高，直到单局击杀37人时，吃鸡概率近乎为1。由此说明这个游戏想要获胜，猥琐，不刚枪杀人，获胜的几率反而不大。

# 添加是否成功吃鸡列

agg1['won'] = agg1['team\_placement'] == 1

agg1.loc[agg1['player\_kills']<40,['player\_kills','won']].groupby('player\_kills').won.mean().plot.bar(figsize=(15,6), rot=0)

plt.xlabel('击杀人数', fontsize=14)

plt.ylabel("吃鸡概率", fontsize=14)

plt.title('击杀人数与吃鸡概率的关系', fontsize=14)

Text(0.5, 1.0, '击杀人数与吃鸡概率的关系')

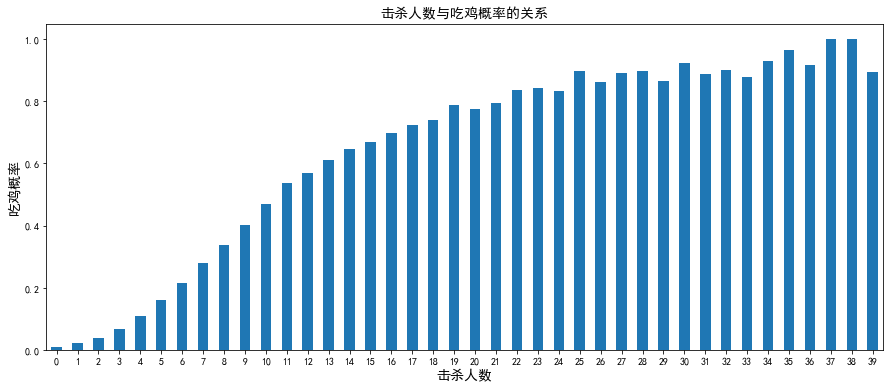


图 4.2.1 击杀与吃鸡概率图

各模式下的平均击杀数：

agg1.groupby('party\_size').player\_kills.mean()

party\_size

1 0.897479

2 0.897744

4 0.876073

Name: player\_kills, dtype: float64

### 4.2.2武器该怎么选择，什么武器更容易杀人？

通过枪械，我们筛选出了不同枪械击杀的人数，如下图所示：

death1 = pd.read\_csv('D:\py/kill\_match\_stats\_final\_1.csv')

erg\_died\_of['killed\_by'].value\_counts()[:10].plot.barh(figsize=(10,5))

plt.xlabel("被击杀人数", fontsize=14)

plt.ylabel("击杀的武器", fontsize=14)

plt.title('武器跟击杀人数的统计(绝地海岛艾伦格)', fontsize=14)

plt.yticks(fontsize=12)

(array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9]),

[Text(0, 0, 'M416'),

Text(0, 1, 'M16A4'),

Text(0, 2, 'SCAR-L'),

Text(0, 3, 'AKM'),

Text(0, 4, 'UMP9'),

Text(0, 5, 'S1897'),

Text(0, 6, 'Mini 14'),

Text(0, 7, 'Kar98k'),

Text(0, 8, 'S686'),

Text(0, 9, 'S12K')])

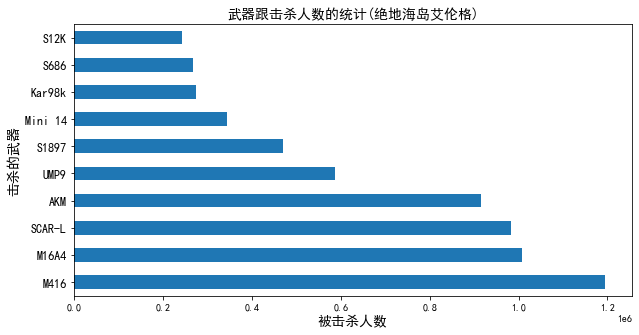


图 4.2.2.1海岛武器击杀图

mrm\_died\_of['killed\_by'].value\_counts()[:10].plot.barh(figsize=(10,5))

plt.xlabel("被击杀人数", fontsize=14)

plt.ylabel("击杀的武器", fontsize=14)

plt.title('武器跟击杀人数的统计(热情沙漠米拉玛)', fontsize=14)

plt.yticks(fontsize=12)

(array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9]),

[Text(0, 0, 'M416'),

Text(0, 1, 'SCAR-L'),

Text(0, 2, 'M16A4'),

Text(0, 3, 'AKM'),

Text(0, 4, 'UMP9'),

Text(0, 5, 'S1897'),

Text(0, 6, 'Kar98k'),

Text(0, 7, 'SKS'),

Text(0, 8, 'Mini 14'),

Text(0, 9, 'S686')])

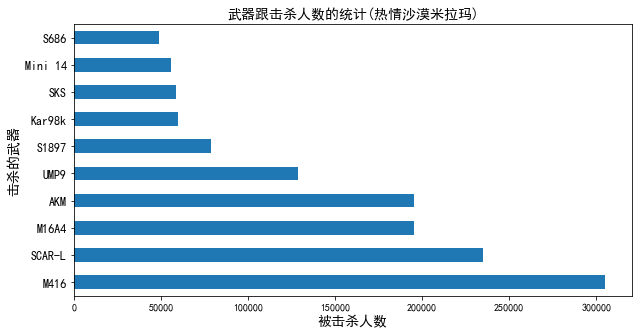


图 4.2.2.2 沙漠武器击杀图

通过这张图都可以很清晰的看出来，这个游戏M416是不二之选，喷子（S686,S12k）终究是错付了。

### 4.2.3远处狙击，适合的武器

在这里，狙击范围我们采用的时八百米到一千五百米，因为距离问题，可能与实际情况有些许区别。

erg\_died\_of.loc[(erg\_distance>800)&(erg\_distance<1500), 'killed\_by'].value\_counts()[:10].plot.bar(rot=30)

plt.xlabel("狙击的武器", fontsize=14)

plt.ylabel("被狙击的人数", fontsize=14)

plt.title('狙击武器跟击杀人数的统计(绝地海岛艾伦格)', fontsize=14)

plt.yticks(fontsize=12)

(array([ 0., 100., 200., 300., 400., 500., 600., 700., 800., 900.]),

[Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, '')])

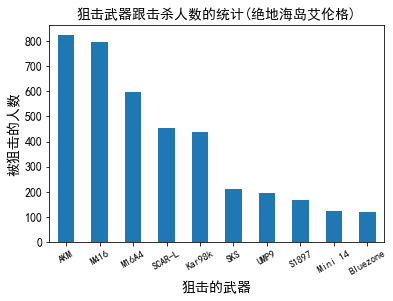


图 4.2.3.1海岛狙击武器击杀图

mrm\_died\_of.loc[(mrm\_distance>800)&(mrm\_distance<1000), 'killed\_by'].value\_counts()[:10].plot.bar(rot=30)

plt.xlabel("狙击的武器", fontsize=14)

plt.ylabel("被狙击的人数", fontsize=14)

plt.title('狙击武器跟击杀人数的统计(热情沙漠米拉玛)', fontsize=14)

plt.yticks(fontsize=12)

(array([ 0., 20., 40., 60., 80., 100., 120., 140., 160.]),

[Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

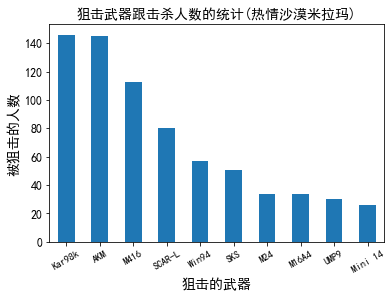
Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, '')])



4.2.3.2 沙漠狙击武器击杀图

海岛地图可以看到，数据领先的是AKM,M416,是步枪，而沙漠确实98K的数据高于他们，可能是范围选取过小，还有可能是枪械刷新率的问题。

### 4.2.5近距离作战适合用什么武器

erg\_died\_of.loc[erg\_distance<10, 'killed\_by'].value\_counts()[:10].plot.bar(rot=30)

plt.xlabel("近战武器", fontsize=14)

plt.ylabel("被击杀的人数", fontsize=14)

plt.title('近战武器跟击杀人数的统计(绝地海岛艾伦格)', fontsize=14)

plt.yticks(fontsize=12)

(array([ 0., 50000., 100000., 150000., 200000., 250000., 300000.,

350000., 400000., 450000.]),

[Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

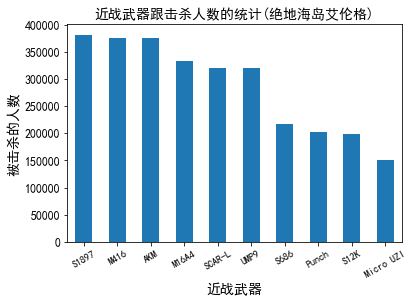
Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, '')])



mrm\_died\_of.loc[mrm\_distance<10, 'killed\_by'].value\_counts()[:10].plot.bar(rot=30)

plt.xlabel("近战武器", fontsize=14)

plt.ylabel("被击杀的人数", fontsize=14)

plt.title('近战武器武器跟击杀人数的统计(热情沙漠米拉玛)', fontsize=14)

plt.yticks(fontsize=12)

(array([ 0., 20000., 40000., 60000., 80000., 100000., 120000.]),

[Text(0, 0, ''),

Text(0, 0, ''),

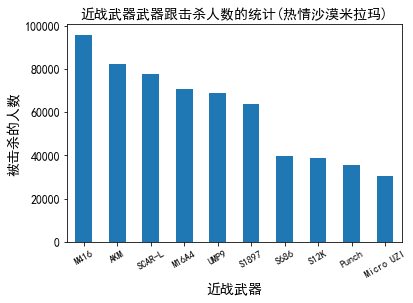
Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, ''),

Text(0, 0, '')])



5.自我评价与总结

经过这次的学习，我们收获很多，不仅仅是知识上的。这一个学期的学习让我们第一次能看到自己码出来的代码可以显示在电脑上，充满了成就感。虽然中间经历了很多困难，程序当中有很多我知识薄弱的点，特别是数据采集和清洗这块。但是老师会不厌其烦，一遍又一遍的帮你梳理你的思路，强调语句的用法。最终还是在我们的不懈努力，他有了最初的雏形。设计这个系统，我们复习巩固了上学期学到的基础知识，进一步了解掌握。知道怎么利用所学知识把理论和实际结合起来，利用资源，采用模块化的结构，学会使用模仿、修改、自主设计相结合的方法。通过pubg的数据芬妮下，掌握结构化的程序设计思想，实现系统的正常运行，并且提高学生综合分析问题的素质和解决实际问题的编程能力。培养学生在项目开发中的求职精神、团队合作精神、创新意识以及实战能力。

作为python结课作业，验证自身的学习效果，同时希望利用所学的基本知识和技能，进行简单github的基本操作。在这两个星期的作业中，也让我看到了一些自身方面薄弱的地方，在接下来的时间里我要好好巩固这方面的知识，争取学好python这方面的知识熟练掌握。

参考文献

[1]刘宇宙.Python 3.5从零开始学[M].北京:清华大学出版社，2017.

[2]董付国.Python可以这样学[M].北京:清华大学出版社，2017.

[3]刘凌霞，郝宁波，吴海涛，21天学通Python[M].北京:电子工业出版社，2016.

**课程设计成绩评定表**

|  |  |  |  |
| --- | --- | --- | --- |
| 课程设计题目 | PUBG数据分析的学习与实践 | | |
| 课程设计学生答辩或质疑记录：  Python里面match()和search()的区别？  答：re模块中match(pattern,string[,flags]),检查string的开头是否与pattern匹配。  re模块中research(pattern,string[,flags]),在string搜索pattern的第一个匹配值。  如何在一个function里面设置一个全局的变量？  答：解决方法是在function的开始插入一个global声明： | | | |
| **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 |  |
| 最终评定等级为：  指导老师签字：  2021 年 1 月 15日 | | | |