广西大学行健文理学院

实践周作业

|  |  |
| --- | --- |
| 题目： | 计科岗位 |
|  | 就业形势数据分析 |

|  |  |
| --- | --- |
| 学 部： | 理工学部 |
| 专 业： | 计算机科学 |
| 班 级： | 182 |
| 学 号： | 1738940341 |
| 学生姓名： | 杨启玢 |
| 指导教师： | 罗泉 |

二〇二〇年十二月

摘要

2020年的疫情爆发，让大学生的就业面临了严峻的挑战。疫情期间，很多公司，商店被迫关门。很多学长学姐已经本已经找到了企业，却因为疫情被告知公司已经倒闭。随着很多公司的倒闭，社会提供的岗位也越来越少。所以，今年的实践周，老师特意安排我们用爬虫工具爬取招聘网站上的职位信息，并写下这篇论文。

但是目前的疫情只是得到了控制，但是疫情还没有完全过去，偶尔还是会反弹一下。但是总体已经趋于可控范围内。

这篇论文主要就java,linux,web前端,游戏服务端，还有算法这5个领域，从薪资,学历要求，工作经验要求,这三个方向去分析疫情稳定后，我们国家上海的计算机相关行业的就业形式。

我使用Python，beautifulsoup4从boss直聘上爬取了近1000条数据，数据量较少，由于boss直聘有反爬手段。每一个关键字爬取20页就已经是极限了，20页往后直接是不让访问的。正因为每个关键字只能爬20页，所以我一共爬了5个关键自。分别是java,linux.web前端,游戏服务端,算法。

关键词：就业形式，计算机

目录

[第一章 引言](#_Toc33)

[1.1 汉德公式产生的理论前提](#_Toc2406)

[第二章 研究进展](#_Toc28833)

[2.1 环境中黑炭的主要来源](#_Toc7859)

[第四章 图表公式示例](#_Toc2574)

[4.1 图的要求及示例](#_Toc28153)

[4.2 表的要求及示例](#_Toc31801)

[4.3公式的要求及示例](#_Toc27971)

[第五章 结论及展望](#_Toc24200)

[参考文献](#_Toc13164)

[附录](#_Toc22801)

[致谢](#_Toc24542)

# 相关领域简介

## 1.1 java

java主要是指使用java语言进行开发的领域，目前java技术主要用在网站的服务器，以及安卓系统的开发上。由于其良好的规范，以及长期以来技术生态的发展，备受程序员的青睐。java语言拥有非常优秀的设计规范，完全为企业打造，所以社会对java程序员的需求量非常大。但是也由于太受欢迎，导致很多大学，和培训机构都在做相关的培训。java程序员的工资并不是特别的高，但也侧面的说明了java语言的受欢迎程度

1.2 Linux

Linux系统也许大家十分陌生，但是在服务器市场，它的地位是举足轻重的。Linux诞生于一个天才大学生的电脑上，他写完第一版的Linux内核后，将其源代码公布在了网络上。在其公布源代码后，立刻受到了很多从事计算机行业的人的关注。他们免费为Linux提供代码，后来在服务器市场上，因为其免费且可靠。所以在服务器市场备受青睐。

1.3 web前端

Web前端是一个大类，它包含的语言是和技术是比较多的。基础的开发语言包括javascript,html,css,jsp,asp。框架包括bootstrap,vue,elementPlus,等一系列的前端框架。web前端就是我们经常看到的网页。

1.4 游戏后端

游戏后端是一个有门槛的领域，因为很少有相关机构去专门做游戏后端相关的开发。游戏后端使用的语言比较局限，一般是用c++，lua，erlang，java，go。其实这个在网上是不好调查的，我查了百度，一般都是推荐c++。但是据我了解，每个公司自有自己的一套体系，这跟项目领导的风格是息息相关的。例如我目前实习的公司就是用c++做的底层通信，lua做的上层逻辑。并没有全用c++，甚至c++只占很小的一部分。

1.5 算法

算法可以说是计算机领域的一块非常难啃的骨头，算法是计算机的核心。所以专门从事算法领域的门槛是最高的。算法非常的抽象难以理解，涉及的知识非常的多。因为这方面的人才比较缺乏，所以一般offer拿的也比较的高。

# 第二章 相关数据简介

## 2.1数据分析的手段

### 2.1.1平均数

平均数是指在一组数据中所有数据之和再除以数据的个数 平均数是统计中的一个重要概念.平均数一般是指算术平均数,也就是一组数据的和除以这组数据的个数所得的商.在统计中算术平均数常用于表示统计对象的一般水平,它是描述数据集中程度的一个统计量.既可以用它来反映一组数据的一般情况,也可以用它进行不同组数据的比较,以看出组与组之间的差别.用平均数表示一组数据的情况,有直观、简明的特点,所以在日常生活中经常用到,如平均速度、平均身高、平均产量、平均成绩等等.

使用平均数这个工具，可以让我们知道在上海，我们研究的计算机领域的平均工资。帮助我们从平均数的角度去了解整个数据的离散情况。我们可以通过研究平均数，来了解我们毕业后，从事相关行业所很可能拿到的offer。

### 2.1.2 中位数

中数是按顺序排列的一组数据中居于中间位置的数，即在这组数据中，有一半的数据比他大，有一半的数据比他小。同样，通过研究中位数，可以帮助我们了解我们毕业后，从事相关行业很有可能拿到的offer。

### 2.2.3 众数

众数的定义：一组数据中，出现次数最多的数就叫这组数据的众数。

众数的个数：一组数据中可以只有一个众数，也可以很多个，也可以没有众数。

众数的特点：用众数代表一组数据，可靠性较差。不过，众数不受极端数据的影响，并且求法简便。在一组数据中，如果个别数据有很大的变动，选择中位数表示这组数据的“集中趋势”就比较适合，选择众数不一定适合，选择平均数肯定不适合。

使用众数可以帮助我们了解大多数相关行业从事者的薪资水平，相较于平均数与中位数，它不受极端数据的影响，所以可能更具有代表性。

### 2.1.3 图表呈现

使用图表去分析数据，已经是一个历史悠久的数据分析手段了。特别是随着计算机软件的发展，已经有了excel这样的可以方便的根据数据画出图表的软件了。更高级的，还有用python对数据进行清洗和绘制图形。

# 第三章 数据分析

## 3.1 java岗位分析

# 致谢

本论文是在xx老师的悉心指导下完成的。xx老师作为一名优秀的、经验丰富的教师，具有丰富的xx知识和xx经验，在整个论文实验和论文写作过程中，对我进行了耐心的指导和帮助，提出严格要求，引导我不断开阔思路，为我答疑解惑，鼓励我大胆创新，使我在这一段宝贵的时光中，既增长了知识、开阔了视野、锻炼了心态，又培养了良好的实验习惯和科研精神。在此，我向我的指导老师表示最诚挚的谢意！