**数据分析项目设计**

**一 引言  
1.1 业务概述:**

**概述：挖掘各大招聘网站Python招聘数据并且对数据进行数据分析和可视化.**

**1.2设计语言 : Python3.**

**1.2.1 产品适用范围**

**1.3现状分析：**

**当前Python学习和其他语言转行到Python行业的人数众多,但是对于全国或者成都各个公司对Python工程师(爬虫工程师,全栈工程师,Python工程师等等)的招聘的情况知之甚少,所以我们团队制定做一个Python招聘的数据分析和解读.**

**二 业务描述:**

**21世纪的年轻人越来越喜欢Python这一门语言,学习者众多,同时也包含转行者.因为Python行业发展前景比较开阔,薪资待遇优.其中就包括Python全栈工程师,机器学习,深度学习,人工智能,大数据,爬虫工程师,网络安全工程师Python web开发工程师等等**

**但是同时由于学习者人数众多,多多少少有盲目跟随者不知道Python行业的就业宽度和薪资待遇,为了众多辛勤的Python学习者,我们就Python就业的未来一一分析,给Python学习者一个清晰的方向.**

**根据现状分析的数据,我们给Python学习者提供一个可以参考的Python就业方向的选择 和薪资范畴的了解,以便大家对Python行业和各个公司对Python招聘的需求有一定的了解. 本数据方案仅做参考使用,最终解释权归团队所有,转发请说明转发源地址.**

**2.1番外篇:  
 数据分析的过程如同烧一顿饭，先要数据采集（买菜），然后数据建模（搭配菜）、数据清洗（清洗洗菜,把废菜扔掉）、数据分析（煮菜菜）、数据可视化（摆盘上菜）**

**2.2 用户特点分析  
<根据实际用户情况，指出软件应该如何满足用户需要。分析使用该产品的不同用户类并描述他们相关的特征。特别应注意将该产品的重要用户类与那些不太重要的用户类区分开>  
  
  
2.3 假定约束分析  
2.3.1 直接约束  
<指出为达到所需要的功能、性能，并满足各种约束条件，所需要满足的条件。包括需要的软件产品支持，需要的硬件以及其它支持等>  
  
  
  
  
2.4 功能分析  
<概括地总结产品所具有的主要功能。例如用列表的方法给出。也可以用图形表示主要的需求分组以及他们之间的联系>  
  
  
  
三 团队成员介绍**

1. **团队成员:曾强, 肖祺怀, 张金灿, 谢国毅**

**2、团队成员分配:**

**a.爬虫-----🡪曾强和张金灿**

**b.数据分析----🡪谢国毅和肖祺怀**

**c.数据可视化---🡪谢国毅和肖祺怀**

**d. web展示---🡪4个人一起做**

**四 用户界面  
<陈述所需要的用户界面统一规范风格。这可能包括但不限于：所有界面统一的布局、页面元素、功能；将要采用的图形用户界面（GUI）标准或产品系列的风格；屏幕布局或解决方案的限制；快捷键；错误信息显示标准>  
  
3.2 硬件接口  
<描述系统中软件和硬件每一接口的特征。这种描述可能包括支持的硬件类型、软硬件之间交流的数据和控制信息的性质>  
  
3.3 软件接口  
<描述该产品与其他外部组件（由名字和版本识别）的连接，包括数据库、操作系 统、工具、库和集成的商业组件。明确并描述在软件组件之间交换数据或消息的目的。描述所需要的服务以及内部组件通信的性质。确定将在组件之间共享的数据。 如果必须用一种特殊的方法来实现数据共享机制，例如在多任务操作系统中的一个全局数据区，那么就必须把它定义为一种实现上的限制>  
  
3.4 通信接口  
<描述与产品所使用的通信功能相关的需求，包括电子邮件、Web浏览器、网络通信标准或协议及电子表格等等。定义了相关的消息格式。规定通信安全或加密问题、数据传输速率和同步通信机制>  
  
五 详细设计：**

**1、技术运用:**

**a、数据挖掘/爬虫:**

**scrapy, requests, xpath, bs4, 正则**

**b、数据分析:**

**numpy, pandas , SPSS ,机器学习**

**c、数据可视化:**

**matplotlib fineBIOS (软件)**

**d、web展示:**

**Django drf**

**2、数据库设计：**

**Redis mysql MongoDB 三种数据库配合使用**

**3、添加薪资选项(自己设计):**

**4 数据分析结论(项目完成补充)**

**5 非功能需求  
5.1 性能需求  
<阐述需求中要求的性能对应于软件的性能，并将之定量化。应解释性能需求的原理 以帮助开发人员做出合理的设计选择。如果性能需求很不完善，应将其补充。完整的性能需求应包括但不限于相互合作的用户数或者所支持的操作、响应时间以及与 实时系统的时间关系；容量需求，例如存储器和磁盘空间的需求或者存储在数据库中表中的最大行数；各种功能的响应时间；内存、计算资源、存储资源的消耗。可 能需要针对每个功能需求或特性分别陈述其性能需求>  
  
  
5.2 安全需求  
5.2.1 系统安全需求  
<详尽陈述与产品使用过程中可能发生的损失、破坏或危害相关的需求。定义必须采取的安全保护或动作，还有那些预防的潜在的危险动作。明确产品必须遵从的安全标准、策略或规则>  
  
5.2.2 数据安全需求  
<关于数据完整性或与私人问题相关的需求也应在此处描述。这些问题将会影响到产品的使用和产品所创建或使用的数据的保护。定义用户身份确认或授权需求。明确产品必须满足的安全性或保密性策略。你可能更喜欢通过称为完整性的质量属性来阐述这些需求>  
  
5.3 其它软件质量属性  
<详尽陈述与客户或开发人员至关重要的其他产品质量特性。这些特性必须是确定、定量的并在可能时是可验证的。至少应指明不通属性的相对侧重点。例如易用程度优于易学程度，或者可移植性优于有效性>  
  
  
  
6 其他需求  
<定义在软件需求规格说明的其他部分未出现的需求，例如国际化需求或法律上的需 求。还可以增加有关操作、管理和维护部分来完善产品安装、配置、启动和关闭、修复和容错，以及登录和监控操作等方面的需求。在模板中加入与你相关的新部 分。如果你不需要增加其它需求，就省略这一部分>  
  
  
  
7 尚未解决的问题**