# Python数据分析

**计划工程文档**

版本2.0

IMG_256

江苏润和软件股份有限公司

南京润云科技有限公司

2019年7月

履历

|  |  |  |  |
| --- | --- | --- | --- |
| 版本 | 更新日期 | 更新描述 | 作者 |
| 1.0 | 2019-07-09 | 初稿 | 晁圣博、黄优、赵韩兵 |
| 2.0 | 2019-07-10 | 修改了内容格式等问题，详见第一次计划工程评审报告 | 晁圣博、黄优、赵韩兵 |
| 3.0 | 2019-07.15 | 精简了项目概述 | 晁圣博、黄优、赵韩兵 |
|  |  |  |  |

目 录

[1．项目概述 1](#_Toc14594)

[1.1项目背景 1](#_Toc2659)

[1.2项目目标 1](#_Toc19891)

[1.3项目规模 2](#_Toc20313)

[2．项目组织结构 3](#_Toc20951)

[2.1人力资源配置 3](#_Toc13689)

[2.2人员沟通计划 4](#_Toc1266)

[3．项目管理流程 4](#_Toc13402)

[3.1估算项目 4](#_Toc26068)

[3.2划分项目开发阶段 4](#_Toc1728)

[3.3项目阶段时间进度 4](#_Toc26274)

[3.4计划项目的风险管理 5](#_Toc5883)

[4．项目技术流程 6](#_Toc22594)

[4.1开发案例 6](#_Toc8719)

[4.2项目开发中使用的方法、工具和技术 6](#_Toc28957)

[4.3产品验收计划 7](#_Toc16067)

[5．项目支持流程 8](#_Toc23866)

[5.1文档编写计划 8](#_Toc11705)

[5.2评审计划 8](#_Toc19235)

[5.3测试计划 9](#_Toc19222)

## 1.项目概述

### 1.1项目背景

大数据开启了一次重大的时代转型，它的核心就是挖掘出庞大的数据库独有的价值。数据分析要引入统计学的思想，大数据背景下，传统的抽样分析已经不适合大数据分析，如果在大数据时代下进行抽样分析就像在汽车时代骑马一样实际情况是“样本=总体”，通过对数据网络之间的联系进行分析。

随着数据挖掘技术的发展，数据的获取自然会越来越容易，但统计学作为从数据中读取信息的科学，应该永远和获取数据的学问相伴相生。大数据的数据量巨大，不能代表我们能观测到总体。有时候，总体是可测的，但更多时候，总体从理论上说是不可观测的。所以应该引入统计学思想，帮我从数据里还原出背后的真实。

大数据的科学价值和社会价值正体现在这里。一方面，对大数据的掌握程度可以转化为经济价值的来源；另一方面，大数据已经撼动了世界的方方面面，从商业科技到医疗、政府、商业、经济、人文以及社会的其他各个领域。数据分析可以帮助制定聪明的决策----他是把数据转化为信息的过程，分析信息以得到见解，并制定可以影响商业绩效的策略和计划。在大数据时代，需要冲扥挖掘数据的隐藏价值，对数据加强分析。

### 1.2项目目标

数据分析，无论是现在的互联网企业，还是传统型企业，都需要数据分析。

本项目的目的是对部门人员信息进行数据的分析处理，其目的如下：

1.现状分析：通过分析部门的整体人员情况，明确目前部门的人员分布流动状况。

2.原因分析：通过第一点，了解部门存在某种隐患，则需要去分析该隐患。了解隐患为什么会存在？是如何产生的？了解了原因才能更好的针对事件进行分析从而达到我们要的目的。

3.预测分析：分析了现状，也分析了原因，接下来就需要预测。通过现在掌握的数据，和之前的事件的经验来预测接下来部门的人员流动发展趋势等。

其实三种作用，总结来看，可以分析：过去、现在和未来。掌握过去部门整体人员分布情况，分析现在部门可能存在的隐患，预测未来部门的人员流动发展趋势。

以上就是我们根据三种数据分析来进行的判断，从而能做到事件有条不紊的进行，这就是我们来做数据分析的目的和更深层次的结果。

### 1.3项目规模

本项目是针对润云内的人员信息做的一个对数据进行处理的项目，由于是新人演习项目，缺乏项目开发流程经验，所以可以初步定位为一个小型的演练项目，内容比较简单。初步界定为开发小组6人，工期7.8-7.30。

## 2．项目组织结构

### 2.1.人力资源配置

|  |  |
| --- | --- |
| 计划项 | 人员安排 |
| 1.计划安排 2.SVN部署 | 1.蒋安全  2.All |
| 1.RA 2.PP 3.评审 | 1.蒋安全，郑能锦，陈国华  2.赵韩兵，黄优，晁圣博  3.All |
| 1.BD 2.FD 3.SD 4.评审 | 1.郑能锦，晁圣博  2.蒋安全，陈国华  3.赵韩兵，黄优  4.All |
| 1.DD | All |
|  |  |
| 1.DD 2.评审报告总结 | All |
| 1.Git部署 2.编码实现及单元测试 | All |
|  |  |
| 1.编码实现 2.单元测试 | All |
| 模块测试 | All |
| 集成测试 | All |
| 系统测试 | All |
|  |  |
| 完善验收 | All |

### 2.2.人员沟通计划

1.组员间每日针对项目交流讨论，各抒己见，吸收可用的建议

2.组员之间互相了解各自对工作的看法，以作改进

3.针对已完成的阶段进行组内评审，交流意见，完善工作

## 3．项目管理流程

### 3.1.估算项目

时间估算：周期为24天

人力估算：6人

### 3.2.划分项目开发阶段

|  |  |
| --- | --- |
| 7.8 | SVN局域网部署，项目计划 |
| 7.9--7.15 | 文档编写 |
| 7.16--7.23 | 源代码编写 |
| 7.24--7.26 | 模块测试、集成测试、系统测试 |
| 7.29--7.30 | 完善验收 |

### 3.3.项目阶段时间进度

|  |  |  |
| --- | --- | --- |
| 员工数据分析网站计划进度表 | | |
| 阶段 | | 主要完成工作 |
| 第一阶段 | 7/8 | 安排总体计划、部署SVN |
|
| 7/9 | 完成PA、PP文档撰写和评审 |
| 7/10 | 完成BD、FD、SD文档撰写和评审 |
| 7/11 |
| 7/12 | 完成DD文档撰写、评审和报告总结 |
| 7/15 |
| 第二阶段 | 7/16 | 部署Git  搭建开发环境  编码实现  单元测试 |
| 7/17 |
| 7/18 |
| 7/19 |
| 7/22 |
| 7/23 |
| 第三阶段 | 7/24 | 模块测试 |
| 7/25 | 集成测试 |
| 7/26 | 系统测试 |
| 第四阶段 | 7/29 | 完善验收 |
| 7/30 |
| 7/31 | 项目演示 |

### 3.4.计划项目的风险管理

（1）风险管理的主要内容

1.识别风险：在项目进行中确认有可能会影响项目进展的风险，并记录每个风险所具有的特点

2.量化风险：评估风险和风险之间的相互作用，以便评定项目可能产出结果的范围

3.提出风险对策：确定对机会进行选择及对危险做出应对的步骤

4.实施控制风险对策：对项目进程中风险所产生的变化作出反应

（2）项目实施中可能存在的风险

1.进度风险

因时间安排不合理造成项目未能按时完成，在开发过程中项目负责人应对开发进度实时监控，若认为安排不合理或者开发人员提出进度建议，需及时作出判断和调整。

2.资源风险

因时间或个人能力不足，未能按时完成开发阶段中的具体工作造成项目延期完成，若存在资源风险则需要提高工作效率，必要时延长工作时间完成任务。

3.技术风险

因缺乏开发经验、开发技能不足造成项目未能按时完成，项目负责人需考虑开发人员的个人技能掌握情况和开发经验来安排工作任务。

## 4．项目技术流程

### 4.1.开发案例

(1)体育赛事预测

世界杯期间，谷歌、百度、微软和高盛等公司都推出了比赛结果预测平台。百度预测结果最为亮眼，预测全程64场比赛，准确率为67%，进入淘汰赛后准确率为94%。现在互联网公司取代章鱼保罗试水赛事预测也意味着未来的体育赛事会被大数据预测所掌控。

(2)交通行为预测

基于用户和车辆的LBS定位数据，分析人车出行的个体和群体特征，进行交通行为的预测。交通部门可预测不同时点不同道路的车流量进行智能的车辆调度，或应用潮汐车道；用户则可以根据预测结果选择拥堵几率更低的道路。

### 4.2.项目开发中使用的方法、工具和技术

1.方法：web前后端分离

2.工具

开发语言：Python、JavaScript

开发环境：PyCharm、VSCode

开发数据模块库：pandas

3.技术

Css、React 、Echarts

### 4.3.产品验收计划

(1)验收标准

1.文档审查

检验系统建设文档是否齐全、完整、规范。

2.功能模块审查

审查系统的各功能模块是否按照规划完成。

3.性能审查

审查python数据分析项目小组提供的《测试报告》

4.用户可用性审查

审查python数据分析项目小组提供的《用户使用说明书》

(2)验收小组及职责

1.验收小组成员

组长：蒋安全

组员：郑能锦，陈国华，赵韩兵，黄优，晁圣博

1. 验收小组职责

按照验收流程组织验收会议，协调相关人员确保验收工作按计划完成，对照需求说 明书审核是否已完成所有建设任务。审查所有项目文档。

3.验收小组成员分工

Python数据分析项目系统功能确认：对照需求分析报告，检查Python数据分析项目各功能是否可用，郑能锦，晁圣博负责。

Python数据分析项目系统文档确认：检验系统建设文档是否齐全、完整、规范，蒋安全，陈国华负责。

Python数据分析项目系统性能及可用性确认：审查python数据分析项目小组提供的《测试报告》和《用户使用说明书》，赵韩兵，黄优负责。

## 5．项目支持流程

### 5.1文档编写计划

|  |  |  |  |
| --- | --- | --- | --- |
| 文档名称 | 内容 | 完成时间 | 负责人 |
| 需求分析 | 项目功能需求、项目性能需求 | 2019.7.9 | 蒋安全，郑能锦，陈国华 |
| 计划工程 | 项目概述、项目组织结构、项目管理流程项目技术流程、项目支持流程 | 2019.7.9 | 赵韩兵，黄优，晁圣博 |
| 基本设计 | 项目的结构，采用的技术 | 2019.7.10 | 郑能锦，晁圣博 |
| 功能设计 | 每个功能模块的详细功能设计 | 2019.7.10 | 蒋安全，陈国华 |
| 结构设计 | 每个功能的逻辑设计 | 2019.7.11 | 赵韩兵，黄优 |
| 详细设计 | 每个功能的详细设计 | 2019.7.11 | 全员 |
| 评审文档 | 文档存在的问题 | 2019.7.15 | 全员 |
| 测试结果 | 代码测试记录 | 2019.7.26 | 全员 |

### 5.2评审计划

(1)评审时间：

文档和代码完成之后

(2)评审人员**：**

小组全部成员

(3)评审结果

评审报告

### 5.3测试计划

1. 单元测试

代码开发过成中每个功能模块完成后进行白盒测试，前端使用mock测试，赵韩兵，黄优负责。后端使用uinttest和postman测试。郑能锦，陈国华，晁圣博，蒋安全负责。

1. 集成测试

所有的功能全部完成后进行集成测试。