1803 爬虫作业

１．完成全书网数据爬取

２．完成数据库表的设计

３．将数据存储进数据库

４．使用retsful搭建小说列表接口

　　，小说章节列表接口，小说详情

接口

1. 复习mongodb的聚合操作命令

1.首先新建数据库，数据库名称为：jobs

2.在jobs中新建一个collection名称为 jobdesc

3.在jobdesc中添加下列数据

jobname workyears tags needpople age

python 3 [python,scrapy,mongodb] 3 20

iOS 3 [iOS,mysql,mongodb] 2 25

java 2 [java, mysql] 1 23

python2 5 [python,scrapy] 5 24

python 2 [scrapy,mongodb] 2 30

1. 使用$group 根据jobname 进行分组，统计每种职位的总数

2. 使用$group 根据jobname 进行分组，统计每种职位需要的总人数

3. 使用$group将集合中的所有文档分为一组，统计职位总数，算出平均年纪age

4. 使用$project修改文档的结构，只展示jobname和needpople 两个文档字段

5. 使用$match函数筛选年纪在20～23岁的职位数量包含20和23，并且只返回jobname和age

6. 使用$match函数筛选年纪在20～23岁的职位数量包含20和23，并且只返回jobname和age,做降序排列

7. 使用$match函数筛选年纪在20～23岁的职位数量包含20和23后，通过jobname分组，并且显示每种职位的平均年龄

8. 使用limit和skip结合使用，返回集合中第二条开始的后2条数据

9. 使用$ｍａｔｃｈ筛选出集合中职位年龄要求在23～30的职位，并根据职位需要的人数做降序排列

10. 使用$unwind 将文档中的tags字段拆分成多条后，返回拆分后的

各职位的总数量。

１１．创建一个超级用户名为“mmm1803”,密码‘１２３４５６’

１２．创建一个普通用户‘beicai1803’,密码‘bc123’,只有操作class1804数据库的权限

以上作业，完成，周一下午３点发送到我邮箱