

13、大数据的例子

1、网站系统各功能访问的信息

#ip信息 请求资源 状态码 访问的时间
192.168.1.100 /user/login.do 200 2017-9-1 15:12:25
192.168.1.105 /user/login.do 200 2017-9-1 15:12:25
192.168.1.101 /user/login.do 200 2017-9-1 15:12:27
192.168.1.101 /user/register.do 500 2017-9-1 15:12:35
192.168.1.101 /user/register.do 501 2017-9-1 15:12:35
192.168.1.102 /user/register.do 200 2017-9-1 15:12:55
192.168.1.102 /user/register31231.do 404 2017-9-1 15:13:25

需求：

- a) 找到所有5xx的的请求的所有请求资源
- b) 找到所有ip, 请求总次数, 状态码200的次数的列表
- c) 找到最近7天里每一天的请求总量, 并按时间排序

...

2、网站系统用户行为信息收集

- x) 查看的行为
- x) 搜索的行为

...

用户名 查看的商品 查看时间
张伟 豆浆机 2017-9-1 15:13:25
张伟 豆浆机 2017-9-1 15:13:25
柳岩 电视 2017-9-1 15:13:15
柳岩 电视 2017-9-1 15:13:35
柳岩 电视 2017-9-1 15:13:45
柳岩 衣服 2017-9-1 15:23:15
张伟 耳机 2017-9-1 15:33:25
柳岩 耳机 2017-9-1 15:13:15
张伟 豆浆机 2017-9-1 15:14:25
柳岩 电视 2017-9-1 15:15:25

需求：找到每个用户最想购买的商品（及查看次数最多的）

3、监测系统的数据分析

- x) 汽车轨迹数据分析
- x) 空气信息检测数据分析
- x) 气温检测数据分析

...

日期 时:分:秒 出租车编号 经纬度 不用管
2015-04-01 00:00:11 114568 121.5304 31.3254 0 0 0.0 02
2015-04-01 00:00:12 18068 121.5075 31.7209 1 2 0.0 02
2015-04-01 00:00:11 116369 121.3286 31.4825 0 6 2.0 02
2015-04-01 00:00:11 114763 121.1479 31.3107 0 1 0.0 02
2015-04-01 00:00:10 12264 121.3907 31.2743 1 5 7.0 02

需求：找出相同的出租车编号，然后按照日期和时:分:秒升序排序。

输出格式例如：

2015-04-01 00:00:11 114568 121.5304 31.3254
2015-04-01 00:00:32 114568 121.5075 31.7609
2015-04-01 00:00:51 114568 121.3286 31.4825