

网易游戏 2016 校园实习生招聘

商业智能研究员笔试试题

注意事项：

1. 考试时间：3 小时。请在规定时间内完成答题并上传答卷至系统，超时上传的答卷将作 0 分处理。
2. 考试过程中可以合理使用网络、工具辅助答题，手写或手绘的部分，可以拍照后粘贴在答卷中，但请尽量使用电子作答方式。
3. 答卷无须复制题目内容，但请标明题号。
4. 请在答题开头部分填写姓名、学校、专业、学历

以下是试题正文内容：

1. 逻辑题（8 分）

有 10 名士兵居住在无法与外界通信的荒岛，有一天需要传递一份重要的情报到外界。假设从荒岛到外界需要 18 天，但是每人最多只能携带 12 天的食物，请问怎样用做少的食物可以将情报送到外界（不能饿死人）

2. 概率题（8 分）

据统计男性有 5%是患色盲的，女性有 0.25%的是患色盲的，今从男女人数相等的人群中随机地挑选一人，恰好是色盲患者，问此人是男性的概率是多少？

3. 简答题（8 分）

k 均值和 k 中心点算法都可以进行有效的聚类。概述 k 均指和 k 中心点算法的优缺点。并概述这 2 种方法与层次聚类方法相比有何优缺点。

4. SQL（10 分）

phone_call_log 是一张记录了某地运营商的用户通话情况的 log 表，表字段为：拨出用户 (call_id bigint)， 操作时间(the_time datetime 精确到秒)，操作类型 (type int, 1 代表通话开始，2 代表通话结束)。假设一个用户在同一时间只能有一个通话，每次通话的最小时长为 1 秒（即通话开始时间和通话结束时间不相等）。

- 1) 请用 sql 查询出这批用户的总通话时长。
- 2) 该运营商目前采用每分钟收费 1.5 毛的方式收费，不满 1 分钟按 1 分钟计算。如果该运营商改为按每秒 0.025 毛的方式收费，请用 sql 查询出这批用户给该运营商带来的收入会减少多少。

注意：创建索引或创建临时表等同样需要在答题中写出 sql

5. 分析题（12 分）

已知某月 App Store 中国区畅销榜排名前 30 的游戏应用分成后收入如下表。同时，按照当时了解到的统计数据，iOS 游戏收入约占整个手游市场收入的 80%。现在请你根据这些信息，估算当月中国手游市场营收规模，并给出你的估算过程。（注：苹果公司将 App Store 应用收入按 3：7 与开发者分成，开发者占 7 成，苹果占 3 成。）

应用排名	分成后收入（美元）	应用排名	分成后收入（美元）	应用排名	分成后收入（美元）
1	890842	11	330025	21	129156
2	672299	12	282393	22	117978
3	668511	13	272348	23	108261
4	578944	14	252101	24	96726
5	541721	15	248184	25	92803
6	516986	16	238135	26	92560
7	488672	17	231509	27	79885
8	369946	18	216444	28	77169
9	367326	19	203193	29	76739
10	355305	20	197583	30	76667

6. 程序优化（10 分）

如下是小 D 实现的一个计算 n 很大时， $1-2+3-4+5-6+7-8+\dots+n$ 的函数

```
long calc( long n){
    long temp=0;
    int flag=1;
    if (n<=0){
        printf("error");//输出
        exit(1);//退出
    }

    for (int i=1;i<=n;i++){
        temp=temp+flag*i;
        flag=flag*(-1);
    }
    return temp;
}
```

请分析 for 循环部分逻辑，从实现及效率的角度，剖析该逻辑可能存在的问题（可能有多个），并深入阐述导致问题的计算机原理；
如该 for 循环部分实现逻辑有问题，请给出你的优化建议（注意不能直接替换掉 for 循环，不能直接给出公式计算结果）

7. 算法设计与实现（10 分）

设计算法求以 a 为底 b 的对数， $a>1, b>0$;不能直接使用任何求对数的系统函数，精度控制在小数点后 7 位即可。要求写代码实现该逻辑（可在提交笔试结果时附加代码文件），如你觉得有必要，添加必要的文字描述，说明你的算法原理

8. 算法设计与实现（10 分）

某人饲养了一头 5 岁的母袋鼠，假设袋鼠 4 岁后 (≥ 4 岁) 方可生小袋鼠；而且之后每年必定只生一头小袋鼠，假设所有生的袋鼠必定是母袋鼠，而且袋鼠 10 岁会死亡且当年不可生育，请问 N 年后，共有多少袋鼠？利用程序实现，并给出 N=2000 的返回值。

9. 作图题（9 分）

下表为一次游戏玩家调研得到的某游戏人口属性数据，请你设计一个“信、达、雅”的可视化方法，将表格中的数据以图形方式展现出来。

性别		学历	
男	60%	初中以下	12.50%
女	40%	高中中专	35.40%
		大专本科	48.50%
		硕士以上	3.60%
年龄		收入	
18 岁以下	8%	0-499	10.50%
19-23 岁	36%	500-999	9.60%
24-29 岁	42%	1000-1999	17.60%
30-35 岁	11%	2000-2999	18.30%
35 岁以上	3%	3000-4999	23.60%
手游玩龄		职业	
不到半年	40%	个体老板	14.50%
半年及以上~1 年以下	30%	上班族	33.40%
1 年及以上~2 年以下	18%	打工族	23.50%
2 年以上~3 年以下	20%	中小学生	7.60%
3 年以上	11%	大学生	16.50%

10. 依据题目内容详细作答（15 分）

徐经理是某游戏公司 N 公司的一名营销经理，负责 N 公司旗下自主研发的手机游戏 X 产品的渠道推广工作。由于 X 产品是 N 公司第一款手游产品，在推广方面尚没有实际经验和数据积累，因此 N 公司管理层决定先在各类型渠道中挑选典型渠道进行测试。徐经理在 A、B、C 这 3 类渠道中分别选取了 a、b、c 这 3 个渠道进行推广测试，相关数据如下：

一、3 种类型推广渠道市场状况如下：

渠道类型	现有渠道数量	Top5 市场份额总和（%）
A 类渠道	300+	75%

B 类渠道	1000+	42%
C 类渠道	86	80%

二、通过估算得到的 X 产品在 a、b、c 渠道测试阶段前 10 天所投入的推广费用

a 渠道	b 渠道	c 渠道
230 万	10 万	1.5 万

三、a、b、c 渠道测试阶段前 10 天核心数据表现

	a 渠道			b 渠道			c 渠道		
日期	新增量	次日留存	30 日 ARPU	新增量	次日留存	30 日 ARPU	新增量	次日留存	30 日 ARPU
2014/7/1	34,004	23.8%	20.8	4,026	30.00%	15.7	91	42.90%	7.2
2014/7/2	31,255	27.7%	3.74	2,338	28.20%	7.39	82	30.50%	133.0
2014/7/3	26,518	26.6%	5.72	2,127	30.60%	10.8	71	40.80%	47.6
2014/7/4	39,916	2.4%	0.81	3,506	28.90%	10.4	66	31.80%	2.0
2014/7/5	44,307	1.8%	0.47	3,515	28.40%	6.51	94	29.80%	180.0
2014/7/6	38,631	4.3%	4.22	2,718	27.80%	8.71	210	39.50%	23.3
2014/7/7	31,675	28.2%	7.99	2,188	29.30%	6.61	204	40.20%	7.1
2014/7/8	31,665	22.8%	17.4	1,753	28.90%	8.29	209	42.10%	36.1
2014/7/9	33,604	17.9%	77.2	1,573	29.40%	7.51	245	34.70%	17.0
2014/7/10	26,933	10.9%	3.84	2,038	30.10%	3.44	186	28.50%	26.6

指标解释：

1. 次日留存：第二天登录了游戏的用户占新增用户的比例
2. 30 日 ARPU：新增用户在新增之后 30 日内的人均消费额

四、a、b、c 渠道测试阶段前 10 天数据汇总

	a 渠道	b 渠道	c 渠道
累计新增量	338,508	25,782	1,458
累计 30 日消费	4693744	234295	54986

现在渠道测试工作基本结束，公司管理层要求徐经理汇报推广测试结果。你是徐经理的一名智囊团成员，将在 15 分钟后出席推广策略研讨会，并准备在会议上提出自己的见解。请整理你在数据中的发现。