

	状态	举办方	赛季2	奖金	参赛队伍
阿里音乐流行趋势预测大赛	已结束	 阿里音乐	2016-07-15	¥ 380000	5475

- 赛制
- 赛题与数据
- 排行榜 
- 论坛
- FAQ

Q: 如何提交结果?

A: 初赛，通过天池网站提交本地结果，临近评测时，赛题详情页，左侧导航有个“提交结果”的入口。复赛，在本队伍所在项目空间，按规定的命名和表结构，产出结果表即可。

Q: 初赛提交的结果CSV文件格式约定?

A: 1、结果文件中，字段与字段之间以半角逗号分隔，记录与记录之间以LF（换行符）或者CRLF（回车换行）分隔。注意不要有多余的引号或者其他特殊字符。正确示例：1234,5678 错误示例：“1234”,“5678” 1234, 5678 2、结果文件中，一定要添加表头行，格式遵循上述1.中的约定。注意字母大小写和拼写准确。 3、结果文件以UTF-8编码，有无bom头均可兼容。请选手提交结果文件前先确认，在文本编辑器中以UTF-8打开时没有乱码。

Q: 为什么gender变量有三个属性值?

A: 男、女、乐队

Q: 发布日期20160207，为什么2015年会有播放记录?

A: 发布时间是整张专辑发布才算，有些歌曲先前已经存在，所以系统中存在少量歌曲视听时间会先于发布时间。有的是先发的单曲，然后后面整专辑的时候把单曲收录了。

Q: 用户播放时间是指什么?

A: 用户播放时间，是指用户播放该歌曲的UNIX时间，参见这个函数 from_unixtime()
https://help.aliyun.com/document_detail/odps/SQL/builtin_function.html?spm=5176.product8314999_odps.6.145.nJyaLI

Q: 用户行为表（mars_tianchi_user_actions）中，gmt_create这个时间戳所在时区是什么?

A: 北京时区。<https://bbs.aliyun.com/read/274901.html>

Q: 如果用户下载歌曲后，离线在本地播放，还会产生记录么?

A: 离线播放也会有播放记录。

Q: 预测的是daily plays，也就是说不包括downloads和collects?

A: 最后预测的是所有用户60天里每天的播放量，预测艺人下面所有歌曲的播放数的和。

Q:要预测的艺人当天的播放数据是基于user表中已有用户的播放，还是预测网站中所有用户的播放?

A:需要预测网站中所有用户的播放。注：初赛是针对用户做了一次随机抽样，后续的真实结果也做了处理。所以原则上，只需要根据已有的用户规模做合理预测即可。

Q: language中各个数字分别代表什么意思?

A: 为了保证数据安全, 字段都经过编码, language中各个字符串分别代表中不同的语言, 如英语、中文、韩文等。

Q: 实际预测结果集中有没有为0的情况?

A: 没有

Q: 是否有对选手恶意的预测做过限定?

A: 有限定

Q: 为什么提交的成绩是负数?

A: 根据公式如果预测出很大的负值是有可能的, 可能预测的太离谱, 需要注意数据的处理和分析

Q: 为什么歌曲初始播放量为0, 但是用户之前却听过?

A: 播放量只是一个参考, 3月1号之前的统计。

Q: 歌曲发布前为什么有人听过?

A: 这个发布是包含歌曲的专辑发布时间, 不是单曲发布的时间

Q: 记录收集日 (分区) 是什么意思

A: 就是当天的行为包括下载、收藏和播放

【关于赛制】

Q: 实名认证是必要条件吗? 会影响到晋级复赛吗?

A: 队伍内所有成员完成支付宝实名认证, 方可进入第一赛季最后阶段, 否则队伍将被淘汰; 如遇到队伍内有成员为海外选手等无法进行支付宝实名认证的情况, 请在[个人中心-认证]页面添加ID card并按照提示上传照片。如仍有问题, 请联系组委会, 官方邮箱: tianchi_bigdata@member.alibaba.com。如队伍内有成员无法继续参赛而未进行实名认证, 请及时清理, 遇到问题及时联系组委会, 官方邮箱: tianchi_bigdata@member.alibaba.com。

【关于计算平台的疑问】

Q: 数据平台里面, 除了写sql之外, 可以写shell吗, 因为我想在sql里加循环, 所以想把sql写在shell里?

A: 御膳房不支持shell脚本, 可以通过开发环境调度形式进行按日期循环操作。

Q: 使用PAI算法平台, 创建“读ODPS数据”节点时, 表名为“tianchi_data.mars_tianchi_user_actions”, 提示“Acesses denied...You doesn't exists in project tianchi_data”, 是什么原因?

A: 算法平台不能跨项目引用表, 需要用数据开发工作台通过sql的方式copy到自己project里, 就能在算法工作台的节点里访问了。

Q: 如何在御膳房上开发map/reduce任务?

A: 为了保证数据平台使用的安全性, 御膳房的MR任务需要基于maven开发.具体配置见文档:
<http://setting.tenant.yushanfang.com/portal/help/doc.html?file=MrUdfLocalDev>

Q: 如果使用机器学习算法来求解问题？

A: 有两种方式: 1. 使用PAI算法平台, 其中提供了很多分类、回归的算法, 包括gbdt, xgboost, LR等; 2. 也可以自己开发, 通过map/reduce接口实现新算法的实现。御膳房的MR任务需要基于maven开发,具体配置见文档: <http://setting.tenant.yushanfang.com/portal/help/doc.html?spm=5176.bbsr274574.0.0.BpnbEf&file=MrUdfLocalDev>

[关于我们](#) [法务协议](#)

[了解更多, 请关注天池微信](#)

解决方案

[成熟解决方案](#) | [专业架构师咨询](#) | [成功客户案例分享](#) | [专属定制服务](#)




行业架构师咨询

热门产品	云服务器ECS 云计算	云数据库RDS	云存储OSS	NAT网关	负载均衡	域名注册	网站建设	大数据
用户热搜	网站备案 云安全	网安法	CDN加速	API网关	企业邮箱	whois查询	视频直播	视频转码
更多推荐	全民云计算 合规安全解决方案	免费套餐	学生机	IT论坛	数据可视化	云虚拟机	com域名	cn域名

[关于我们](#) [法律声明及隐私权政策](#) [廉正举报](#) [联系我们](#) [加入阿里云](#)

[阿里巴巴集团](#) [淘宝网](#) [天猫](#) [聚划算](#) [全球速卖通](#) [阿里巴巴国际交易市场](#) [1688](#) [阿里妈妈](#) [飞猪](#) [阿里云计算](#) [YunOS](#) [阿里通信](#) [万网](#) [高德](#) [UC](#) [友盟](#) [虾米](#) [阿里星球](#) [来往](#) [钉钉](#) [支付宝](#)

© 2009-2019 Aliyun.com 版权所有 ICP证: 浙B2-20080101

   浙公网安备 33010002000099号