拟据











## 腾讯微保实习二面(数据开发工程师)

动动小手,点个赞呗 腾讯微保二面 (35分钟+5分钟提问)压力面 1 自我介绍 2 项目介绍

3 大数据相关 (很多我忘记了)

spark sql读取文件,内存不够使用,如何处理 spark sql与DataFrame的使用 数据倾斜问题,Flink问了一点点 sql如何实现数据倾斜中加盐操作

示例: 给 表中的某个字段插入随机数数据, 取值500~2000

update 表名 set 字段名 = floor(500+rand()\*1500);

UPDATE `表名` SET `字段名`=ceiling(rand()\*500000+500000) WHERE (条件);

floor: 函数只返回整数部分, 小数部分舍弃, 即向上取整

rand: 函数用于产生0 (包含) 到1 (不包含) 的随机数

Spark做实时流处理如何防止乱序现象,Structured Streaming 全都是生产上常见的问题。。。

4 机器学习,数据挖掘相关(太多问题)

问了比赛相关的,主要机器学习的问题使用的什么方法,XGBoost与LightGBM的应用场景逻辑回归原理,如何防止过拟合,如何评价模型,如何进行特征工程如何选择最优的特征,数据预处理,spark的机器学习库,底层源码问题CNN常见模型,一些原理,适用场景。

5写了一个sql,不能额外使用其他表

姓名,课程,课目成绩这三个字段转换成 姓名 课程1,课程2,课程3(对应成绩)

姓名	课程	分数					
张三	 语文	74	 			-结果	
张三	数学	83					
张三	物理	93	姓名	语文	数学		物理
李四	语文	74					
李四	数学	84	李四	74	84	94	
李四	物理	94 学 华客@MarshalJS	张三	74	83	93	毬 华客@MarshalJS

## 方法一

select 姓名,

max(case 课程 when '语文' then 分数 else 0 end)语文, max(case 课程 when '数学'then 分数 else 0 end)数学, max(case 课程 when '物理'then 分数 else 0 end)物理

## 全站热榜

- 关于计算机专业; 网?
- 腾讯软件后台开划 面 40分钟 挂
- 3 美团面试 (已OC
- 4 0419阿里笔试代
- 5 2023春招实习面 (已接阿里offer
- 阿里笔试419 45i 解
- 7 请教如何刷题、i
- Java-2022暑期到 2023春招部分面
- 9 4月19日 0offer
- 京东JD Young实 (后端开发)

## 正在热议

#2023届毁约公司名

3613次浏览

#滴! 实习打卡#

3047次浏览

#五一调休情绪发泄的 2531次浏览

#阿里巴巴信息集散
2689次浏览

#通信硬件2023笔面 3713次浏览

#现在前端的就业环境

7674次浏览

#如何拒绝/ [ **6**]

074次浏览

#没有实习经

9715次浏览

#要钱还是要

**F**4

#我的成功项...... 7058次浏览

#你收到了团子



5993次浏览



2

☆ 20

#23届找工作求助阵:
122270次浏览

#京东信息集散地#
1631次浏览

#国企/银行/研究所2
12643次浏览

#Ooffer是寒冬太冷辽
58847次浏览

#为什么那么多公司指
6452次浏览

#工作丧失热情的瞬间
8142次浏览

#机械制造2023笔面

#你觉得今年春招回

33160次浏览



4









☆ 425 월 …

**心** 174 **⊡** 143

☆ 20 回 14 ピ 分享

<u>((</u>

6

4