

---

## 美团点评 2020 面试题 – 数据工程师

---

1. 缓慢变化维处理
2. 数据流（实时数据流-离线数据流）
3. Mapreduce 的流程
4. 数据仓库分层
5. hive 的 map/reduce 数怎么确定？
6. hive 数据倾斜怎么办？
7. 二叉树种类？完全二叉树知道么？有什么应用？
8. 维度建模过程？
9. 支架表？拉链表？统计带状维度如何更新？
10. 元数据怎么管理？
11. kafka 了解吗？干什么用的？
12. 数据库引擎结构、区别
13. 索引底层结构，插入数据索引结构的变化等
14. 索引的优化方法，以及为什么
15. 索引使用（优化器如何界定）
16. 数据库几种日志，以及如何刷新，区别
17. 数据库三个线程用处
18. HDFS 数据如何保证不会丢失？
19. MR 详细过程
20. JVM 虚拟机回收机制

---

21. 分布式存储系统和分布式计算框架区别

22. Spark MR 区别

23. MR 和 HDFS 架构

24. YARN 容错机制

25. HDFS 数据本地性

## 业务题

1. 陌生城市，想通过美团 app 选择到店就餐。若周围 5km 内，店铺数量过少，用户无选择空间；若店铺太多，其实也没意义，不会花太多时间去逐一研究。  
如何确定门店密度？（假设横轴为门店密度，纵轴为满意度）
2. 每个渠道的佣金上涨了，但平均佣金率下降的问题
3. 估计整个餐饮市场团购的规模
4. 怎么提高我们的业务占有等。
5. 百团大战期间商家端还是用户端更重要
6. 项目中所用的模型及应用成果
7. 项目/实习中和他人的沟通情况

## SQL 题

1. 给定 uid,login\_time,logout\_time 求每分钟最大的在线人数.两个 time 时间都是标准时间 datetime(2019-01-01 12:00:00)这种