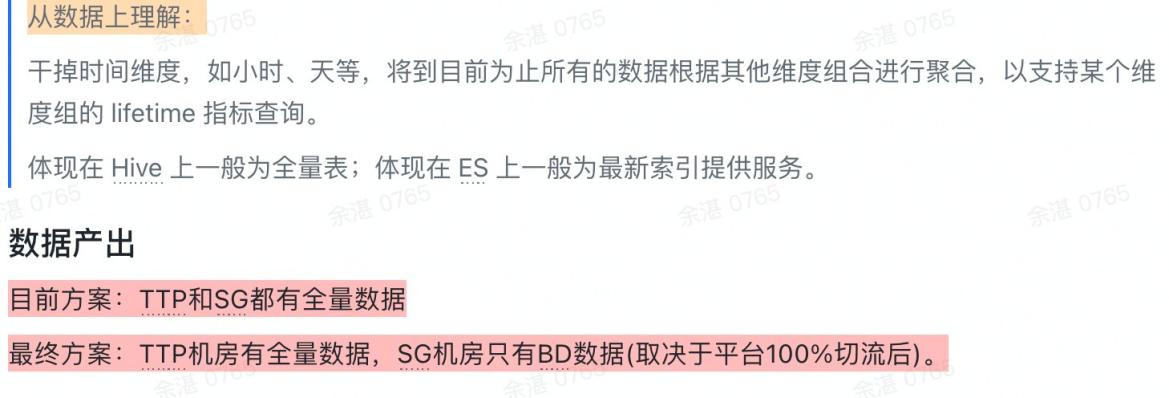
# 字节

1. 非中国区广告的大数据, 数据来源有各种 app，主要是 tiktok，然后我们把这些数据处理好在一个看板上展示给广告主。看板有一些事件归因(广告主看我们的点击率 转化率啥的)，自助挑战赛，

看板上还有受众预估的功能，广告平台在创建广告计划的时候会在平台去筛选，比如投 在 tiktok，男，兴趣 tag 为打篮球，筛选后返回一个 1 亿这个数作为受众预估数

1. 数据模块化管理：我们的数据是好多链路，数据分离线数据和流式数据，流式数据我没怎么接触，主要是 Flink Streaming 任务。离线数据我接触到的主要是一些 hive sql 任务， 这里面就是用 hive sql 对数据进行聚合或者过滤啥的，会分比如小时级任务 天级任务， 然后一般会写到 es 或者 druid 来给下游提供查询服务。还有一些跨机房的同步任务(因为去年美国要限制美国数据的流出，也就是不能把所有数据都给我们，所以我们在美国新弄了一个机房，我的一些工作也是和美国机房数据服务部署以及和原有的新加坡机房建立联系相关的)
2. 写数据服务

在 hive sql 的离线任务里用中间件资源(也就是用 java 写的数据查询服务，里面会去 es 等一些地方取数，这个 es 是一个天级 hive2es 任务写入映射数据的)拿到两个字段(国家和大区)的映射关系来传到下游，这样广告主就可以在数据大盘上分国家和大区看到数据

请求独立访客数据，本来是请求本地 es 或者 druid，现在要跨域去美国机房取数

技术上：

用 SQL 来写一些数据流任务

用 java，python 写一些数据查询相关的服务，比如读取 es，druid 等

业务上：

1. 在大数据开发平台里创建任务或者根据需求更改任务 sql，并且配置一些任务的参数， 比如监控报警，超时重试时间等，还有对任务回溯数据
2. 在 es 查询服务里加上字段映射并配置路由，这样别人就可以通过查询服务找对应的接口用配置完的指标和维度名来请求数据。这个查询服务通过 python 来写，主要是去 es 里查数来返回
3. 搭建测试环境，建一些测试用的任务和 es 集群和索引

# 商汤

1. 企业微信用户填写员工工位信息的功能
2. 企业滴滴员工订单信息管理
3. 访客以及车辆预约系统应用，用预约单的形式允许访客新增/修改/取消预约，预约单里可以添加多个访客信息以及多个车牌信息。

访客在一楼时候可以手机号获取验证码来进行验证，这个验证码用 redis 缓存，设置缓存时间等

预约车位的时候可能会有两个人同时预约到同一个剩余的车位， 所以预约的时候用synchronize 锁来保证，(也可以用数据库的行锁锁定某一行)

select \* from 数据库表 where id=5 for update