

11 | 数据的台子搭完了，但你还得想好戏该怎么唱

2020-04-29 郭忆

数据中台实战课

[进入课程 >](#)



讲述：郭忆

时长 22:22 大小 20.49M

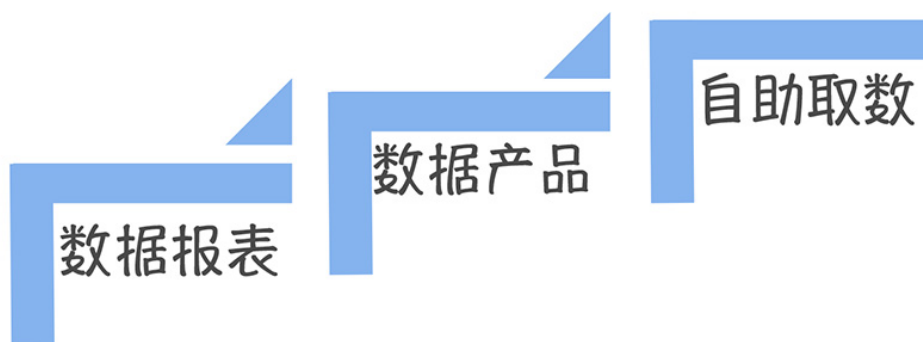


你好，我是郭忆。

从第 4 节元数据管理开始，到第 10 节数据安全，我已经详细讲了如何建成快、准、省和安全的数据中台。现在，数据中台的台子已经全部搭完了，接下来，好戏就可以上演了，也就是说，我们要在数据中台的基础上，构建企业数据应用体系，用好数据中台的数据。

对企业来说，用好数据非常关键，从我多年的数据建设经验来看，我把数据在企业的应用划分成三个阶段。





初级阶段。一般企业的数据应用都是从数据报表开始的，分析师会为业务部门的负责人、运营制作一些 BI 报表，把数据通过可视化的方式呈现出来，这是数据应用的初始阶段。

发展阶段。只是可视化的展现数据已经不能满足业务的需求，业务需要根据数据持续监控业务过程，发现问题、诊断分析，并给出决策建议，最后需要一键执行，形成完成的业务过程闭环，**这个时候就要借助数据产品来实现**，网易也是在 2018 年才开始大规模构建数据产品体系。

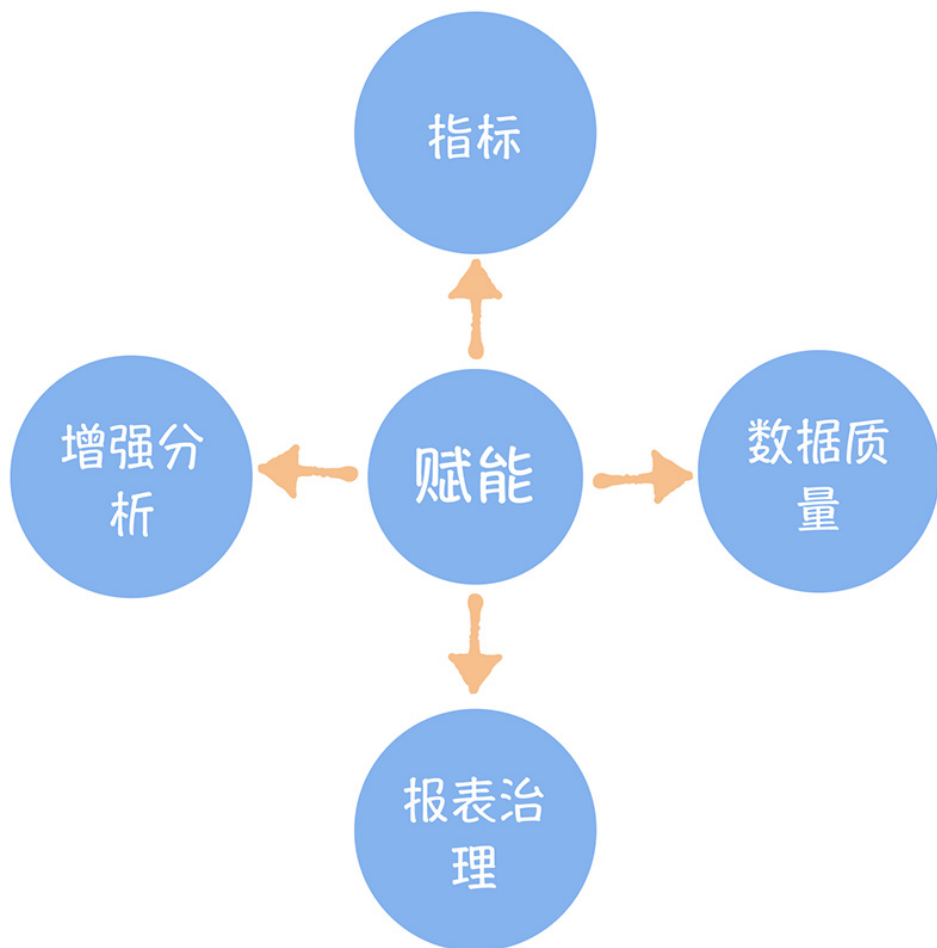
高级阶段。无论是数据报表、还是数据产品，它们呈现的都是固化的分析思路，只能解决知道的业务问题，但是日常工作还有很多未知的业务问题，比如销售额指标突然下降了，需要基于数据进行探索分析。这个时候，如果都依赖分析师，肯定不现实，**那么就要实现自助取数**，让每个人都能基于数据去做分析和决策，实现普惠大数据。我认为这是数据应用的最高级阶段，网易在 2019 年开始开放越来越多的中台数据，让更多的非技术人员去使用数据。

那么今天这节课，我们就从这三个阶段，谈一谈如何用好数据中台的数据。

数据中台该如何赋能 BI 工具

很多人对数据的了解，都是从 BI 工具做的报表开始的。关于 BI 工具的产品本身，不是我想说的重点，我主要想和你讨论的是数据中台时代，如何让数据中台帮助 BI 工具更强大。

我会从四个方面带你了解这部分内容。



第一，统一报表指标业务口径。

数据报表上会存在指标口径不一致的问题，相同指标名称，两个报表里的数据却相差很大，这会让数据使用者对数据失去信任。

而数据中台的所有的指标都是由指标系统统一管理的，如果能在数据报表上直接看到指标系统中，指标的口径定义，就可以让看报表的人准确理解数据的含义，也可以避免不同报表之间指标口径不一致的问题。

同时，如果我们在指标系统上修改了指标的口径定义，也可以同步到所有的呈现该指标的数据报表中。

第二，掌握任务影响了哪些数据报表。

当某个任务异常，影响了下游多个任务时，我们往往要根据任务的影响范围，决定任务恢复的优先级。如果任务影响了老板每天看的一张报表，而你却不知道，没有优先修复它，那你

就等着被批吧。

那我们要怎么知道一个任务影响了哪些数据报表呢？

在网易，数据报表在保存时，BI 工具可以把报表和数据的链路关系，推送给数据中台的元数据中心。当数据中台的任何一个任务出现异常，我们通过数据血缘，就可以快速找到这个任务影响了哪些数据报表，尤其是在故障恢复的时候，根据报表的优先级，我们可以优先恢复高优先级的报表。

第三，治理低价值的数据报表。

根据数据中台的全链路数据血缘，我们可以计算每一个报表上游所有的数据加工成本，然后得到这个报表的成本。然后根据报表的访问量和访问人群，我们可以计算报表的 ROI（投入产出比），下线低价值的数据报表。

第四，全维度钻取。

在制作报表时，分析师只能依靠经验去判断一个指标有哪些可分析维度。如果 BI 工具能根据元数据中心提供的所有指标可分析维度，自动根据指标在各个维度下的取值，找出指标波动的原因，那这就是全维度钻取了，它是目前业界最为热门的研究领域，增强分析的一个方向。

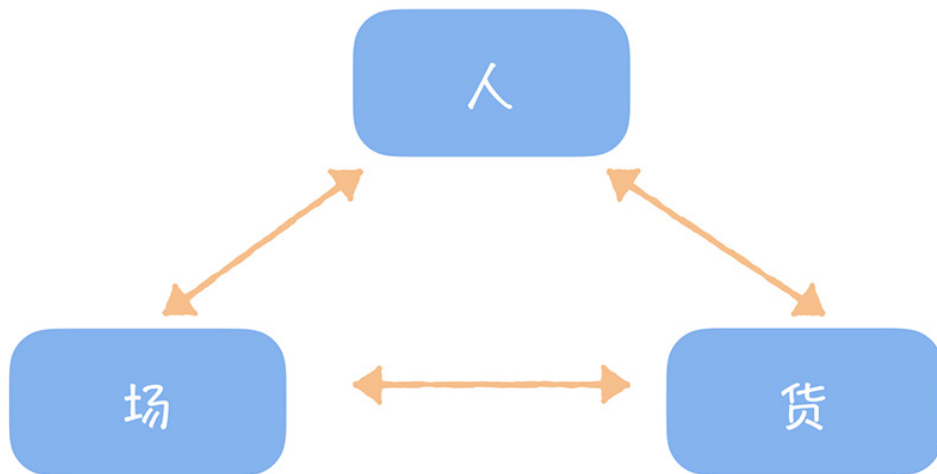
比如，有一个单车租赁公司，发现 8 月份的营业额下降了，系统通过根据各个维度的数据对比和分析发现，8 月份营业额下降，是因为那个月雨天的天数增多导致的。如果分析师不知道用天气的维度去分析营业额，很可能就不知道原因。但是全维度钻取，可以基于数据中台营业额的所有可分析维度，包括天气，自动计算出雨天的销售额相比晴天的销售额低，同时进行交叉分析，发现 8 月份的雨天数量比其他月份多，最后找到问题的原因。

你看，数据中台是不是很大程度上增强了 BI 工具的产品能力？在 BI 工具的基础上制作数据报表，这才是数据应用的初级阶段，接下来，咱们继续看一下，基于数据中台，我们能做出什么数据产品，提升业务的运营效率。

打造零售行业精益数据运营体系

零售行业是目前我见过的所有行业中，对数据使用程度最深的行业，所以我会以零售行业为例，带你了解如何借助数据实现精益运营。

假如你是“贾天真连锁奶茶店”的老板，你的目标是把更多的奶茶卖给更多的人，赚更多的钱。那你要时刻谨记零售行业一个很经典的理论，那就是：人、货、场，在正确的地点，把正确的商品，卖给正确的人。



让更多的人，买更多的奶茶

让更多的人，买更多的奶茶，你必须要解决客户拉新和促活的问题。那如何拉新呢？

获得新用户的方式，一般就是做广告，但是做广告也有很多渠道：

微信公众号；

抖音；

快手短视频；

小区电梯；

.....

可这么多的广告渠道，到底哪个渠道的广告效果最好，性价比最高呢？数据说了算！

我们一般用新消用户数、单个新消用户的平均消费金额（新消 ARPU）、新消单客成本来衡量各个渠道的广告投放效果。你可以参考这几项，选择最优的广告投放渠道。例如，微信

公众号相比快手短视频，每日新消用户数更多、单个新消的平均消费金额更多、新消客成本更低，那你就应该果断选择微信公众号。

当然，广告中选择的奶茶种类也会在很大程度上影响广告拉新效果。比如高档小区投放广告时，应该选择价格高、健康的饮品；普通小区的话，更加亲民的奶茶才能吸引更多的客户。那如何来选择奶茶的种类呢？还是数据说了算！

除了根据数据选择奶茶种类之外，广告的投放也要讲究策略，就拿微信公众号这个渠道来说，年纪大的客户群体，注重健康饮品；年轻的客户群体，注重价格亲民，口感，样式。所以，必须要基于人群画像（年龄、地区、学历等），决定推送哪些人哪些商品。至于人群画像，需要基于日常的顾客交易数据计算而来。

不过，光拉新用户，但是如果留不住用户也不行。那么如何让老用户，增加消费奶茶的频率呢？

我相信你肯定也见过一些套路，比如经常收到一些短信、App 站内消息、小程序、微信公众号推送的打折信息，然后没忍住，就“剁手”了。那你有没有想过，这些商家是怎么抓住你的，怎么就知道你喜欢这一款？

我曾经做过 2 年的推荐算法，这个算法有一个很经典的论述：大数据可以做到让机器比你更了解自己。所以，如果你曾经购买过奶茶，那系统就可以交易行为数据计算出你喜欢的奶茶口味、品类，你平时喜欢在哪家店购买，然后定向把这些店对应的奶茶优惠信息推送给你，这样你大概率会中招！

你可以看到，店家总是有各种各样的套路促进你消费。

店家在数据的基础上，一方面可以让新客源源不断；又可以增加老客复购的频率，这时整个奶茶生意的销售额就实现了最大化。

保障奶茶不要断货

作为老板的你，要让更多的奶茶，卖给更多的人，那前提必须要保障奶茶的充足供应，这就涉及到供应链管理的问题。

因为奶茶本质上属于生鲜品，如果门店囤货太多，鲜果就会烂掉。但如果缺货，又会影响门店的销售，所以如何在保证不缺货的前提下，尽量减少门店的囤货，这是你必须要解决的问题。

而供应链涉及到销售、补货、到货和库存四个环节。如果有一款数据产品，可以根据奶茶的实际销售情况和销售计划、结合门店库存的安全水位、采购时间周期，自动计算需要补货的原材料，然后推送给采购系统进行补货，那你是不是会觉得很省心？

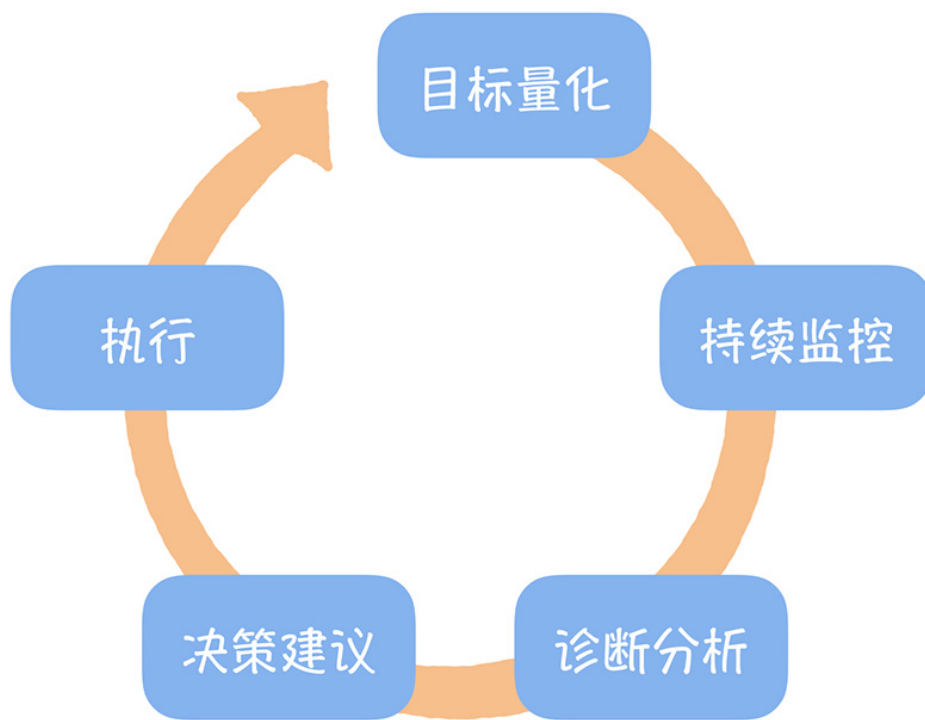
实现门店的利润最大化

当然了，奶茶卖的多不多，还和门店有很大的关系。如果你的店员，可以根据数据，及时发现滞销的奶茶，然后在客户结账的时候，主动推荐这些奶茶，那你就可以获得更高的门店收益。我们一般使用“效坪（每天每平米门店的营业额）”来衡量单个门店的经营状况。

通过这几点，其实你可以看到，零售行业有很多赚钱的窍门。接下来，我带你了解一下如何基于数据产品，轻松地使用这些窍门。

构建数据产品，实现数据驱动下的精益运营

数据产品与 BI 报表最大的不同，在于它们不仅可以实现数据的可视化展示，更为重要的是，可以基于数据，对业务过程进行持续的监控，及时发现问题，进行诊断，并形成决策建议，付诸执行。



数据产品，首先要实现对业务目标的量化。对于卖奶茶来说，你要关注的重点是研发出更多的网红款的奶茶，确保圈住更多的“奶茶粉儿”，同时降低库存周转的压力，因为有越多的滞销奶茶，就会导致积压更多的货物，产生更多的成本。

为了实现这个目标，你可以用动销率来评估目标的达成。

动销率：销售商品的品类数量占库存的商品品类数量的比例。

为了提高动销率，数据产品必须对每个奶茶品类进行销售的跟踪，及时发现零动销的奶茶。

所以，你可能会经常收到“xxx 款奶茶零动销”“xxx 款奶茶慢动销”的预警信息，然后接下来你就要对这款奶茶出现零动销进行分析了：数据产品会通过不同季节横向对比这款奶茶的销售情况，也会通过顾客消费问卷去分析这款奶茶的口感，最终找到这款奶茶滞销的原因。

接下来，你就要根据原因产生决策建议了。比如如果是因为奶茶口感的因素，应该及时下架这款奶茶，否则会影响口碑。数据产品可以推送给运营进行审核，然后运营确认后，一键下线商品，此后各个奶茶店的菜单中，不会再出现该款奶茶。

当然了，我只是拿零售行业举了个例子，因为很多问题都是共通的，用奶茶店，我总结了一些方法论，你可以结合自己所在的行业去应用：

找到业务问题、量化业务目标，比如，我们找到提高奶茶周转的关键，在于及时发现滞销奶茶品类，那么我们用动销率来衡量业务目标；

然后要对业务目标持续监控，及时发现问题，比如，我们监控各个品类奶茶的销售情况，及时发现零动销奶茶；

紧接着，要对问题进行诊断，比如，我们要发现奶茶滞销是因为口感太差；

当然，还要根据原因形成决策，比如下线这款奶茶；

最后付诸执行，比如通过一键，在所有门店菜单中去掉了该品类奶茶。

你看，数据产品实现了从监控问题、发现问题、解决问题的完整闭环。可数据产品毕竟还是按照固化的分析思路进行分析和产生决策建议，在日常运营中，还会有很多数据产品或者数据报表无法解释的问题，这个时候就必须依赖探索式的数据分析来解决，而探索分析的门槛主要在于获取数据，接下来，咱们就来聊聊自助取数的问题。

让技术人员不再是数据的搬运工，释放取数效能

对于传统行业来说，BI 部门一般有两项职责，一个是做报表，一个是取数。而取数的工作量远远多于报表的工作量。

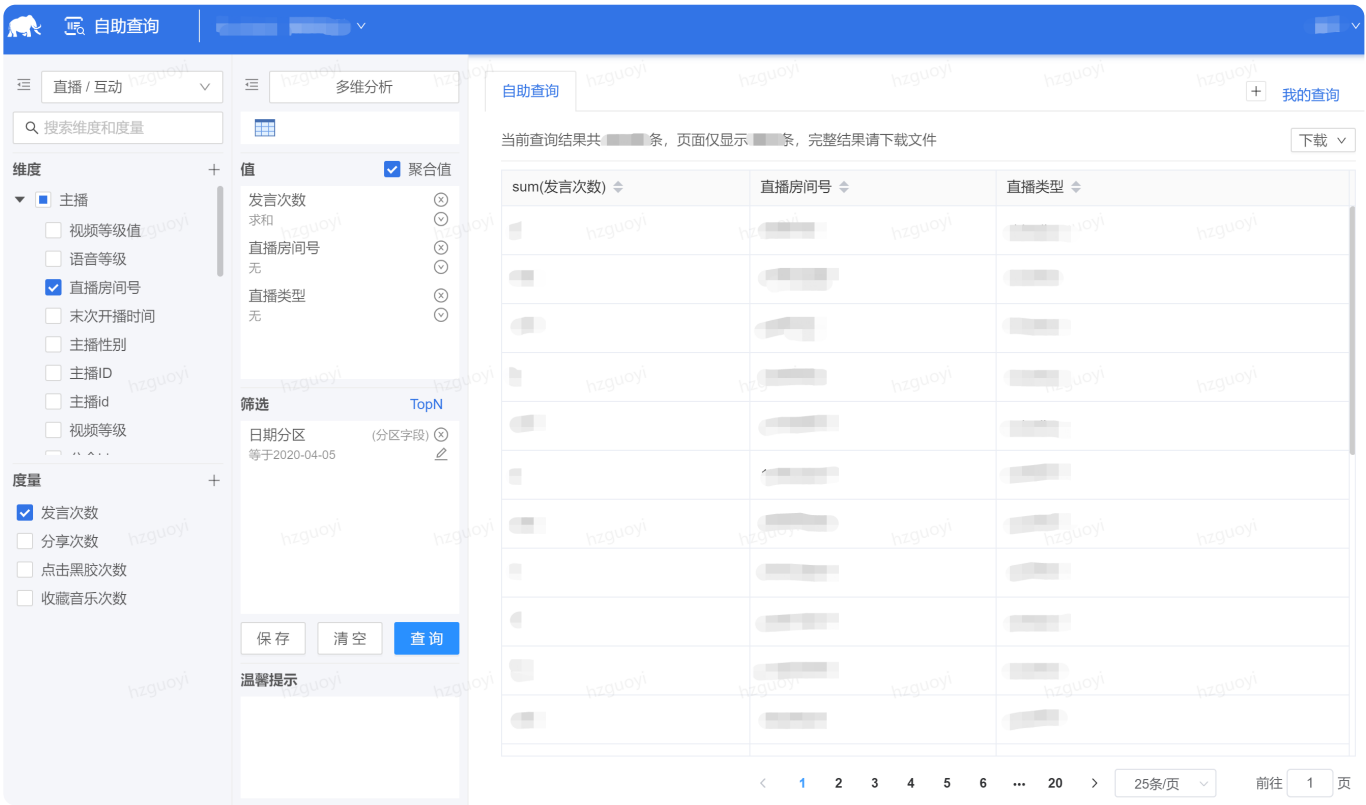
一年中做的报表可能就几百张，但是取数，一年可能要取几千次，或者上万次。而大部分传统企业的取数会依赖技术人员，因为他们离数据更近，取数还涉及写代码，所以，如果你是非技术人员，根本不可能基于数据去做探索式的分析。

所以，大量的取数工作就落在了懂技术的数据开发的头上。

靠别人取数，会存在大量的沟通和协作的成本，同时因为公共集市层数据不完善，导致无法基于现有的数据，直接完成取数，需要数据开发加工新的数据，所以耗时会非常的长，一般需要一周时间。高昂的取数成本，压制了取数的需求，也导致探索式的数据分析，根本不可能大规模的使用。

对于数据开发来说，他们更希望自己的工作重心放在建设公共集市层的数据上，因为公共集市层越完善，取数的成本就越低，不需要额外的开发。但是他们忙于临时的取数需求，根本就没有时间和精力去做这些工作。最后就形成了不良循环，越是集市层数据不完善，取数的工作量就会越大（要开发新的模型），越多的时间去临时取数，集市层越没人建设。

这个问题该如何破解呢？ 我们研发了一个自助取数平台，叫 EasyFetch（意为简单取数）。



这个平台主要有这样几个优点：

- 用图形化的方式，替代了写 SQL 的方式；
- 提供了对业务人员比较友好的业务过程、指标、维度的概念，替换了表、字段；
- 每个指标的业务口径都能够直接显示；
- 用户通过选取一些指标和维度，添加一些筛选值，就可以完成取数过程；
- 界面非常简洁，使用门槛非常低。

在实现层面，我们在数据中台里，加工了多个面向不同业务过程的集市层的表，取数平台会自动根据用户选择的度量和维度，去对应的表中关联多张表进行查询，SQL 会自动根据查询进行优化，避免非技术人员调试 SQL 以及写的 SQL 质量非常差的问题。

通过自助取数平台，原先我们数据开发 50% 的时间都在临时取数，而现在只有 10% 的时间，在自助取数平台无法满足（需要加工集市层模型）的情况下，帮助用户取数。

同时，这部分的工作也会对集市层模型的不断优化产生促进作用。对于取数效率来说，原先 10 个数据开发，一周做 100 个取数需求，已经是濒临极限。而现在，我们一周有 1000 多次有效取数的需求在自助取数平台完成，取数效率提升了 10 倍以上。

还有一个有趣的现象，我也想分享给你，就是我们发现，在周末，也有很多人在使用取数平台，经过调研，我们发现很多人在基于数据写周报，这是之前完全无法想象的事情。

最后，我建议你在设计取数平台时，一定要注重简洁、对用户的引导、降低用户的使用门槛。因为我们面临的是非技术人员，我们要拿出做 C 端产品的姿态去做取数产品。

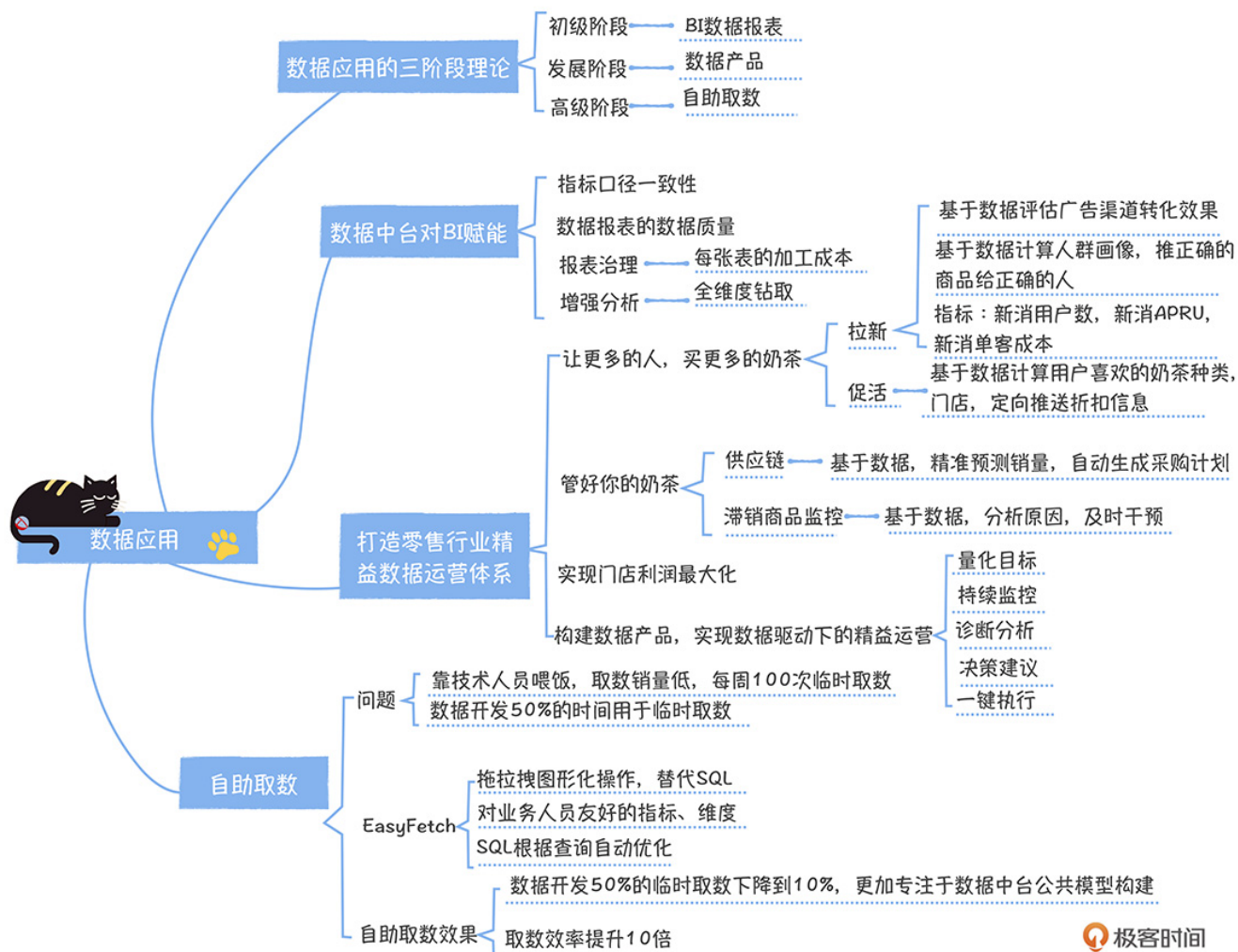
课堂小结

这就是今天我要讲的全部内容了，你可以看到，数据中台之上，可以有这么多的数据应用场景，数据可以帮助我们实现这么多原先不可能做到的事情。在课程的最后，我想再强调几个重点：

数据中台对 BI 赋能体现在指标口径的一致、任务影响分析、数据报表的成本以及基于数据中台的元数据之上的全维度钻取；

数据产品实现了从目标量化，持续跟踪，异常诊断，决策建议，最后到执行的完整数据驱动业务目标达成的闭环；

通过实现面向非技术人员友好的自助取数平台，让数据开发专注于集市模型的构建，可以释放取数的效能，大幅度促进数据的应用范围和深度。



思考时间

今天我主要介绍的都是零售行业数据应用的场景，在其他的行业，比如农业、物流、金融、教育、制造业等等，来谈谈你所在的行业有哪些数据应用的场景，如何实现业务目标的数据驱动？欢迎在留言区与我互动。

最后，感谢你的阅读，如果这节课让你有所收获，也欢迎你将它分享给更多的朋友。

点击参与 

和郭忆一起，落地数据中台



扫一扫参与小程序话题



新版升级：点击「 请朋友读」，20位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 10 | 怎么一劳永逸地解决数据安全问题？

精选留言

 写留言

由作者筛选后的优质留言将会公开显示，欢迎踊跃留言。