

数据驱动下的网易严选 仓储供应链智能优化

邱似峰

网易严选高级技术经理

SPEAKER INTRODUCE

Qiu Sifeng Technical Manager

- 目前负责网易严选仓储物流供应链的技术研发工作。十年网易基础云计算平台经验，主持了分布式数据库、分布式文件系统、数据云存储平台和网易视频云的研发工作。

Email: giusf163@163.com



TABLE OF CONTENTS 大纲

- 仓储物流供应链产品体系架构
- 大数据和实际场景中的痛点
- 人工智能&算法的应用
- 数据驱动案例 — 商品库存布货
- 思考

仓储物流产品矩阵

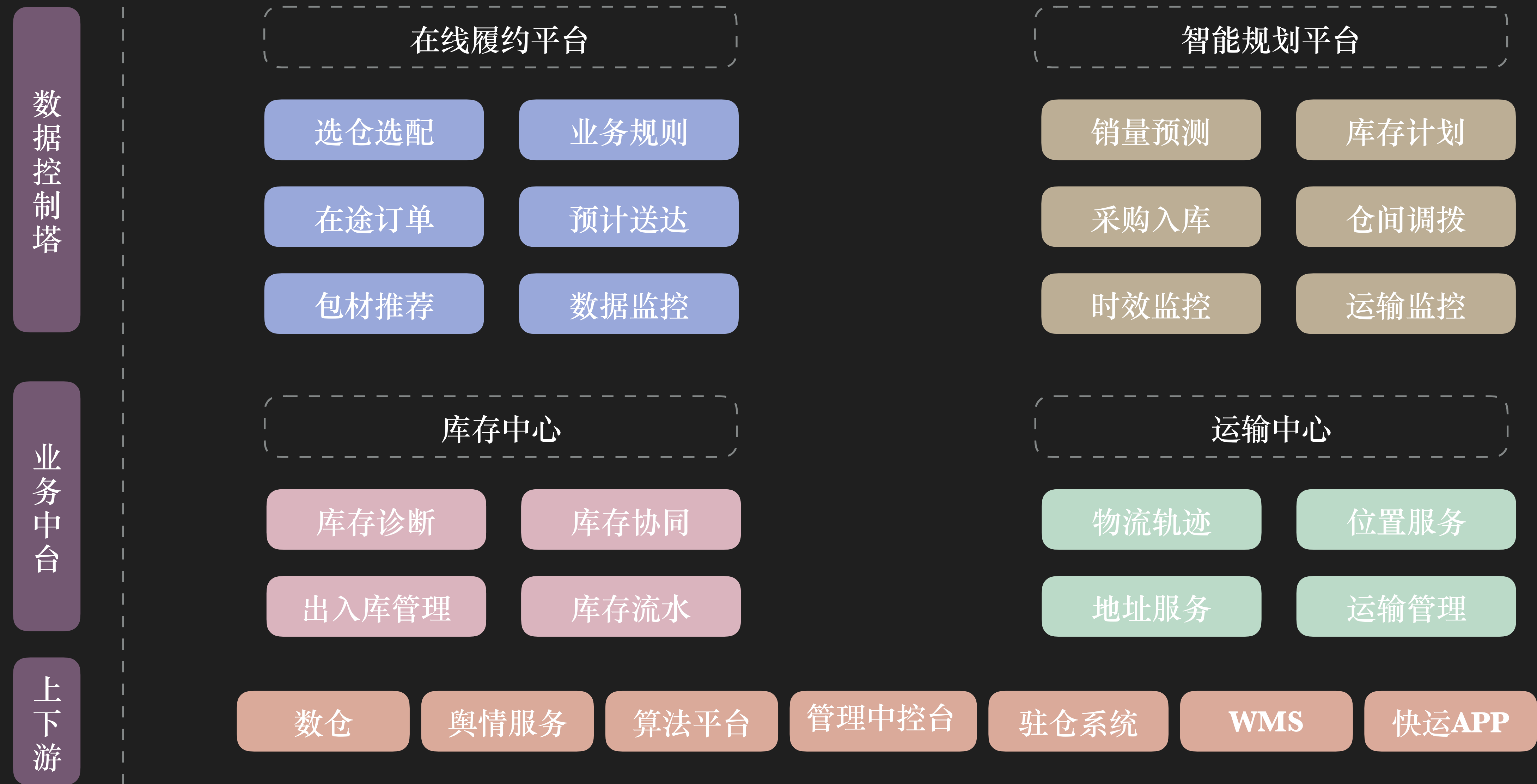


TABLE OF CONTENTS 大纲

- 仓储物流供应链产品体系架构
- 大数据和实际场景中的痛点
- 人工智能&算法的应用
- 数据驱动案例 — 商品库存布货
- 思考

大数据体系 — 数据分析



网易有数

仓配中控台

数据大屏

数据展现



Apache Flink

知乎 @攻城

数据处理



DDB 网易分布式数据库

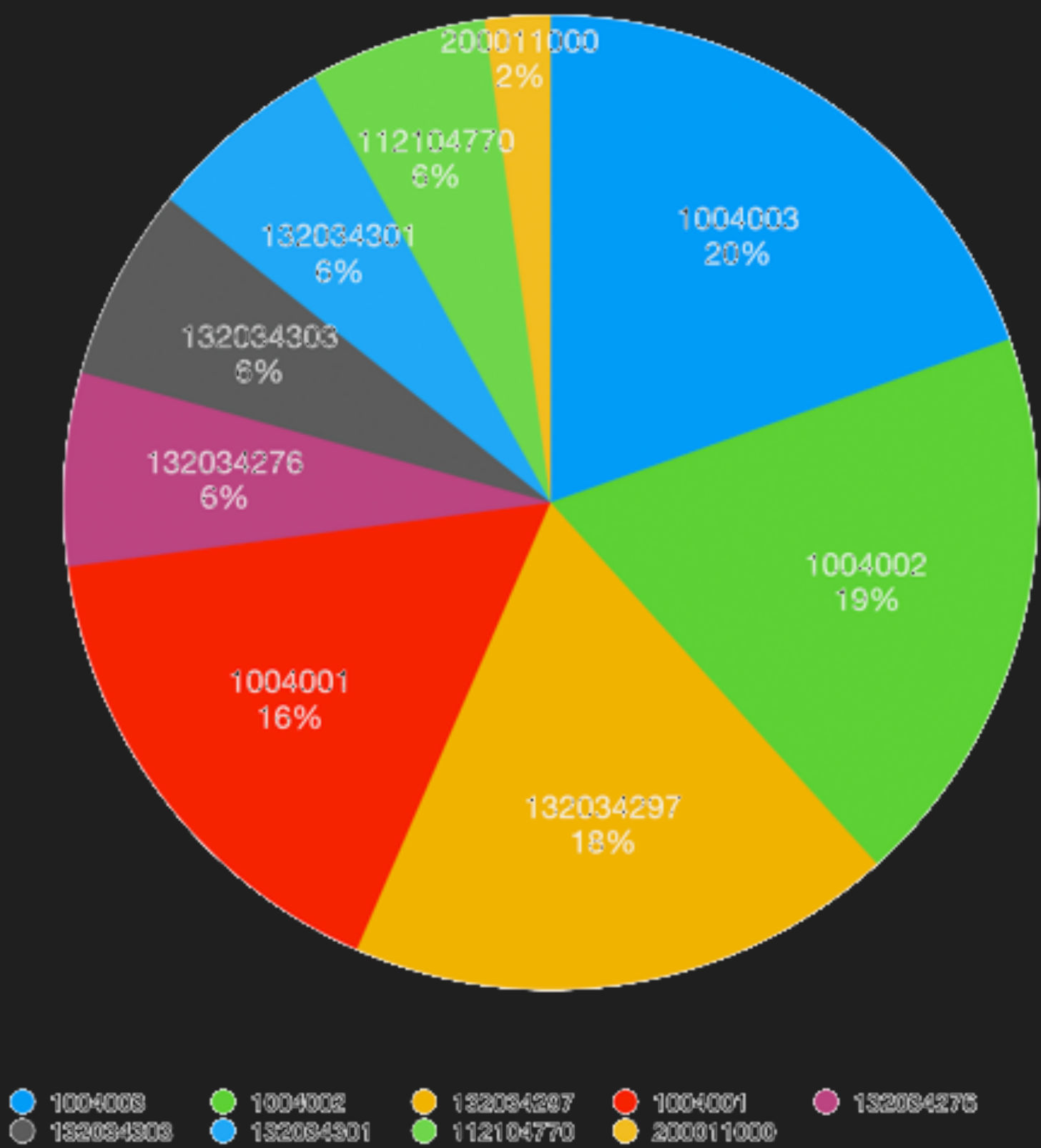
KIWI 严选数据存储格式

数据存储

应用场景

- 全文索引用于日志检索，定位问题
- 图表化展示，做简单的数据挖掘

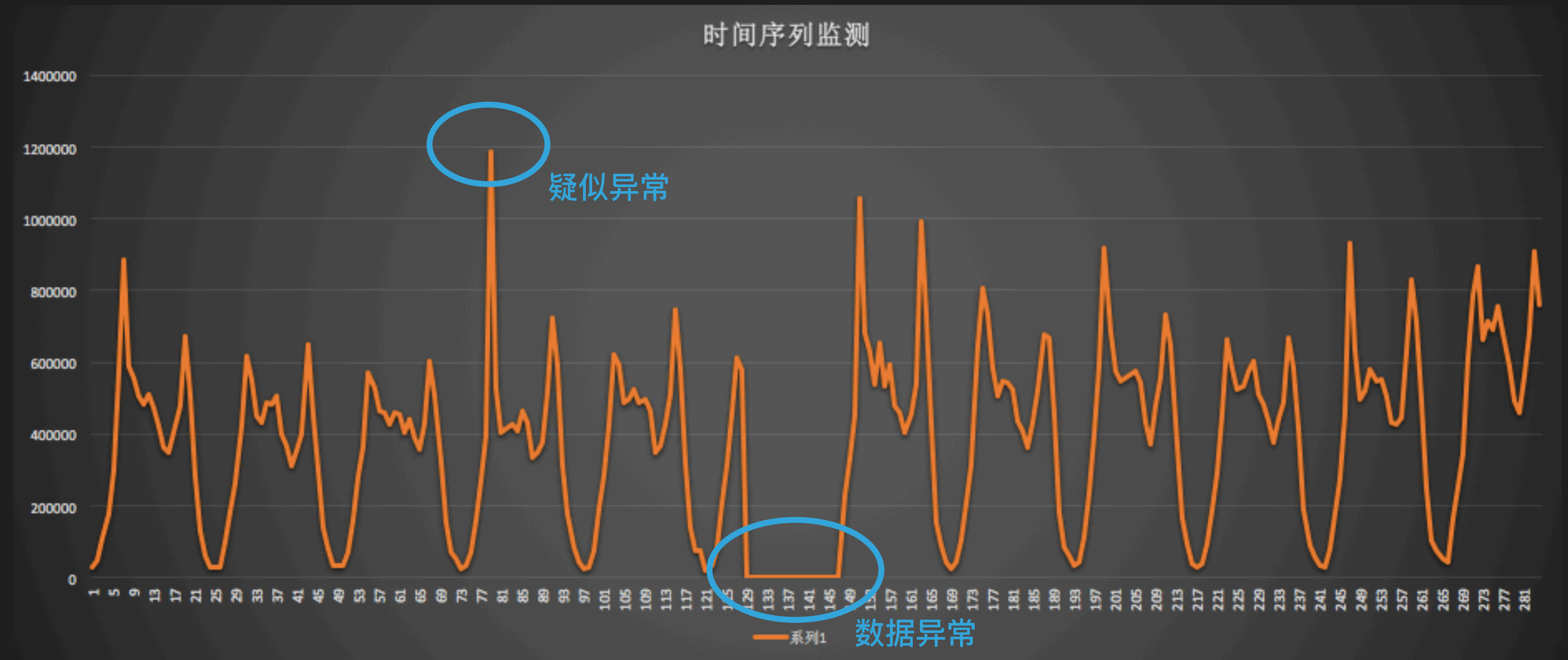
定位：适合研发和产品团队内部使用



订单选仓比例

大数据上的业务痛点

业务痛点1：300+张报表，如何快速发现当前的问题点？



- ▶ 时间序列数据智能判断
- ▶ 对逻辑上的异常数据进行自动追踪报警
- ▶ 集成在报警平台中

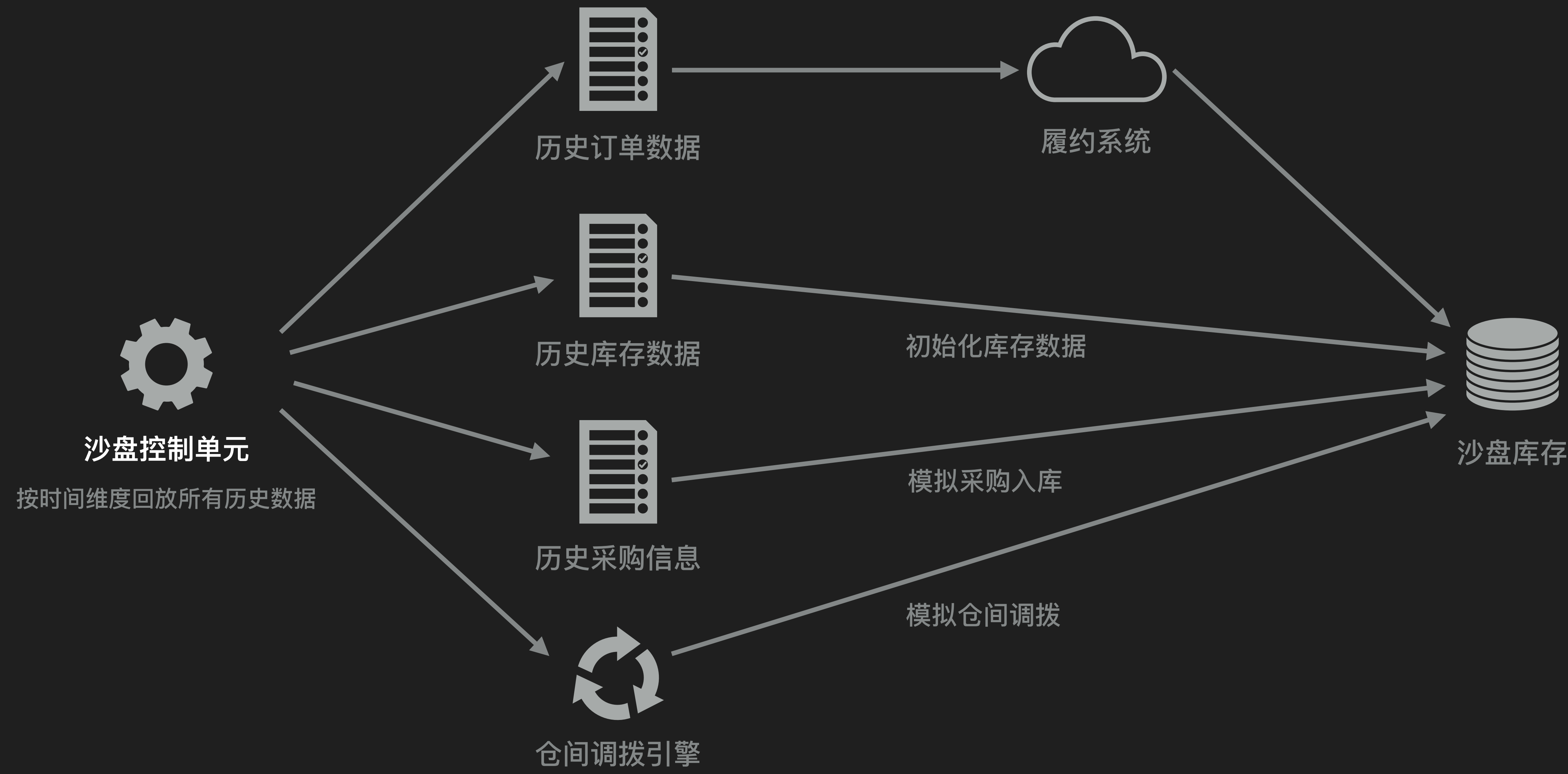
- ▶ 严重偏离的异常实时短信报警
- ▶ 监控报表页面对疑似异常进行排序



疑似异常汇总

业务痛点2：业务场景变化快，无法快速试错，如何控制好风险和成本？

模拟沙盘



准确率

- ▶ 与线上代码保持同一套代码
- ▶ 由同一个团队维护
- ▶ 线上报表进行准确率监控



性能

- ▶ 作为产品有交互上的时间要求
- ▶ 作为自动模型调优平台的基础有性能要求
- ▶ 决策与执行分离
- ▶ 进行一定程度的性能优化：增加缓存、减少数据库开销

沙盘应用（模拟类产品）

仓间调拨模拟

- ▶ 仓间调拨参数（例如安全库存）
- ▶ 输出关键指标（默认仓满足率、拆单率）
- ▶ 各场景对比

波次模拟

- ▶ 设置各仓库波次
- ▶ 输出当日达、次日达

包材模拟

- ▶ 设置包材规格和种类
- ▶ 输出包材体积利用率和使用量

辅助业务日常运营工作

应用场景

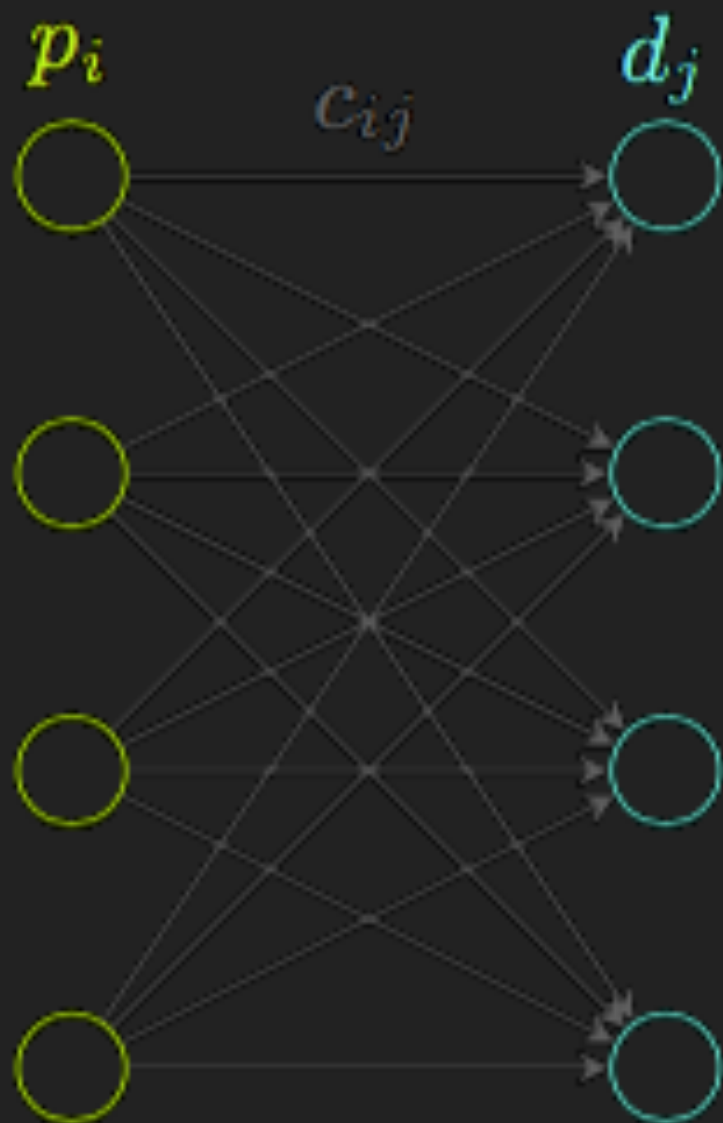
- ▶ 从静态配置表更改为动态选仓选配策略
- ▶ 计算仓间调拨的库存分布
- ▶ 在途策略从最近发货改为妥投时间最优发货
- ▶ 多级库存模型效果评估
- ▶ 履约策略：大订单拆分

降低重大策略更新在系统中上线的风险

TABLE OF CONTENTS 大纲

- 仓储物流供应链产品体系架构
- 大数据和实际场景中的痛点
- 人工智能&算法的应用
- 数据驱动案例 — 商品库存布货
- 思考

运筹学 (线性规划)



仓间调拨

- 供应仓库: i
- 需求仓库: j

下标集(Index Sets)

- ▷ $i \in \{1, \dots, m\}$: 供应仓库
- ▷ $j \in \{1, \dots, n\}$: 需求仓库

参数(Parameters)

- ▷ p_i : 仓库 i 供应的数量
- ▷ d_j : 仓库 j 需求的数量
- ▷ c_{ij} : 单位商品从 i 到 j 的运费

决策变量(Decision Variables)

- ▷ x_{ij} : i 调拨到 j 的商品数量

目标(Objective)

$$\min \sum_{i,j} c_{ij} x_{ij}$$

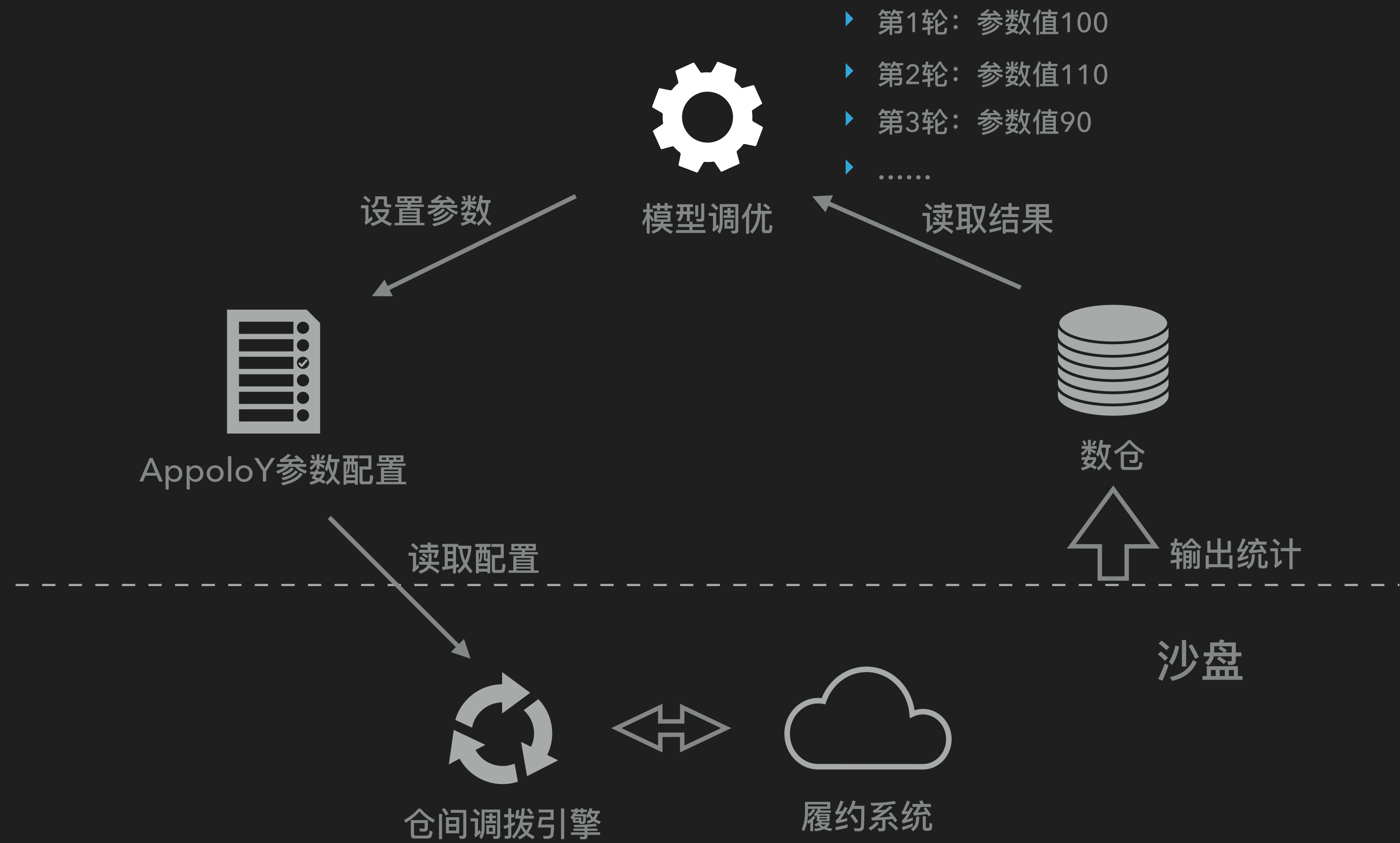
约束(Constraints)

- ▷ 仓库 i 最多供应 p_i : $\sum_j x_{ij} \leq p_i$
- ▷ 仓库 j 最少需要 d_j : $\sum_i x_{ij} \geq d_j$
- ▷ 调拨量非负: $x_{ij} \geq 0$

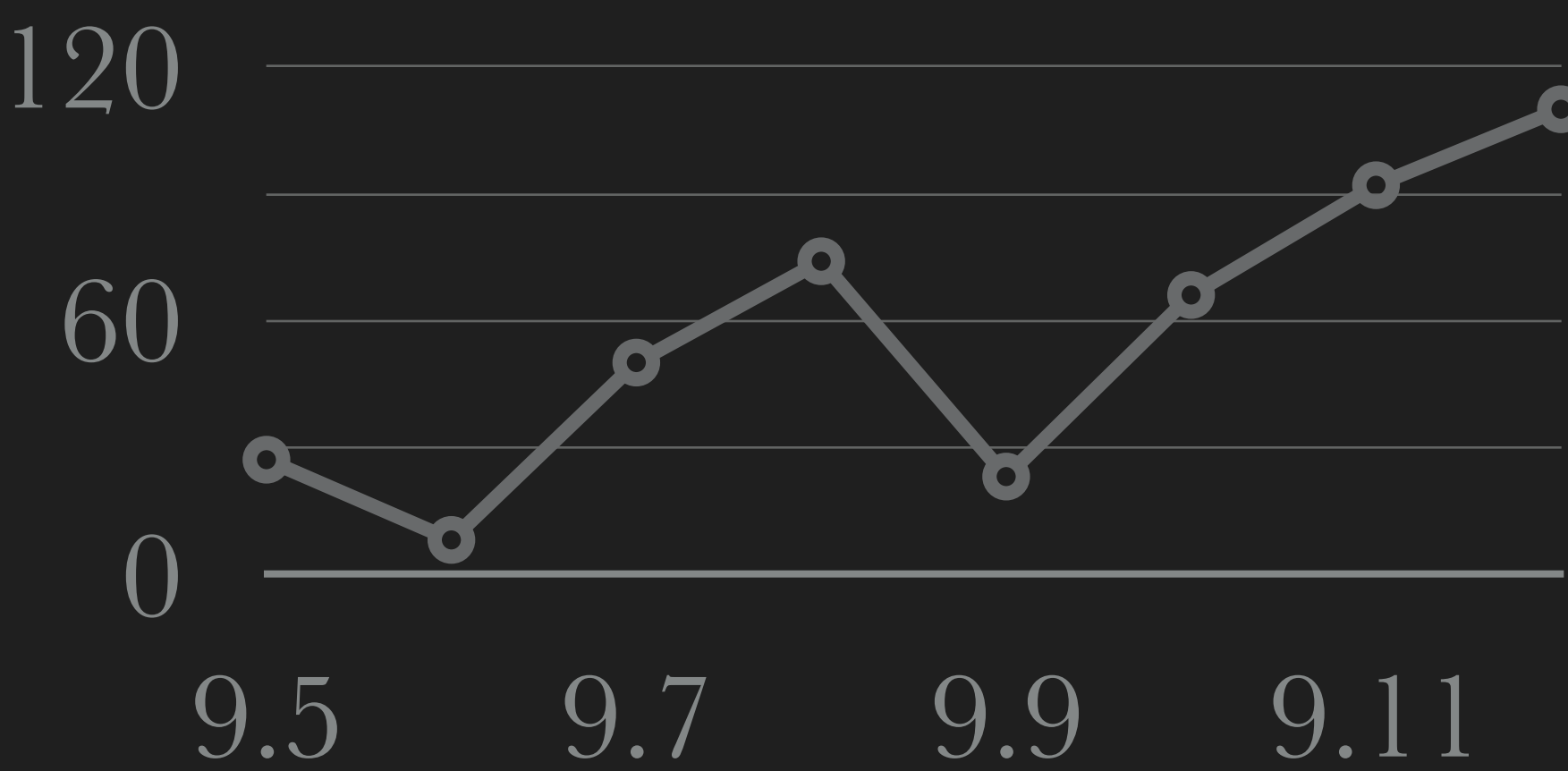
业务痛点：建模速度跟不上业务节奏

- ▶ 渠道的限制每天都变
- ▶ 新开仓影响现有库存模型
- ▶ 仓库产能根据实际情况经常调整
- ▶ 快递和干线运输承运商规则易变

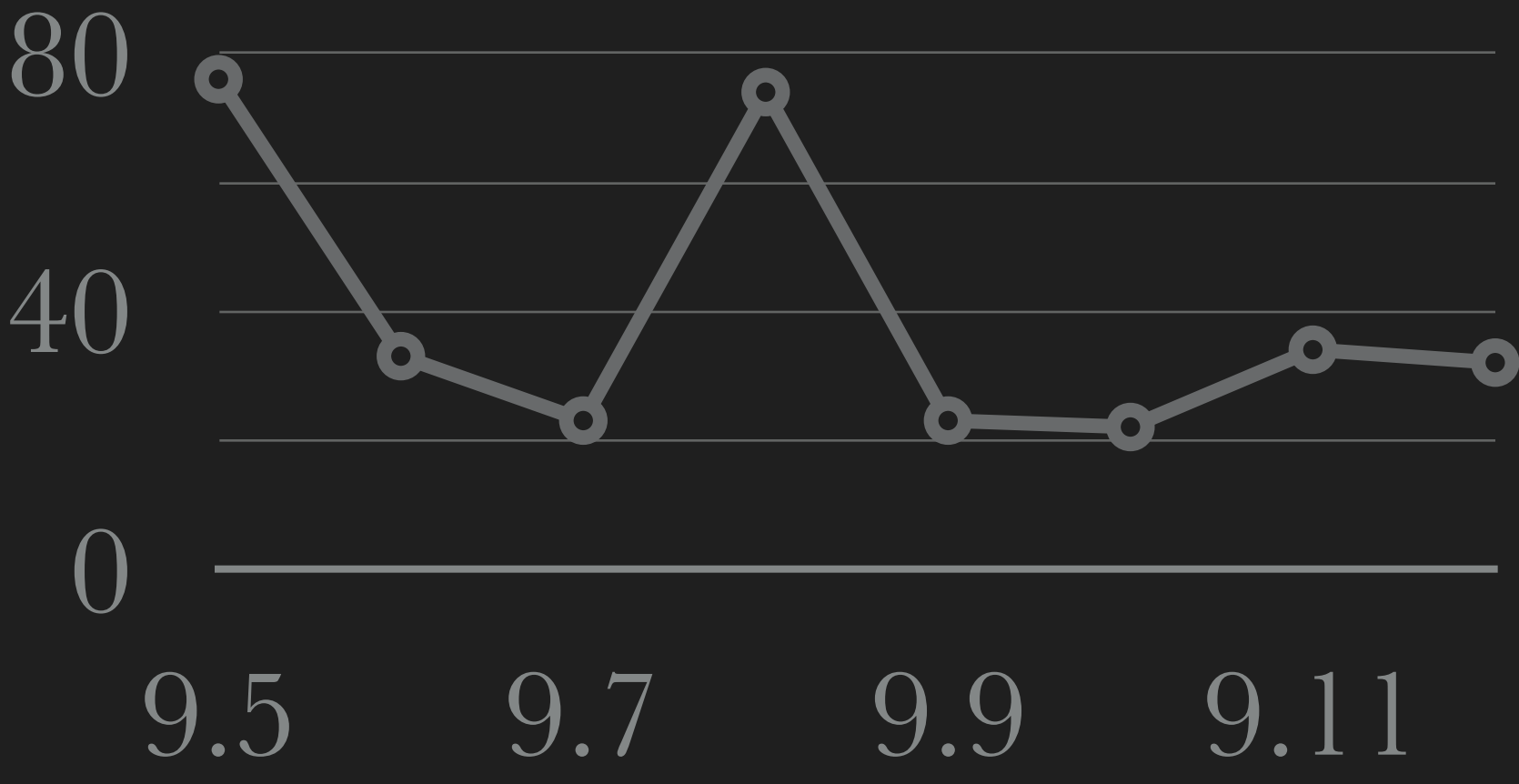
模型智能分析



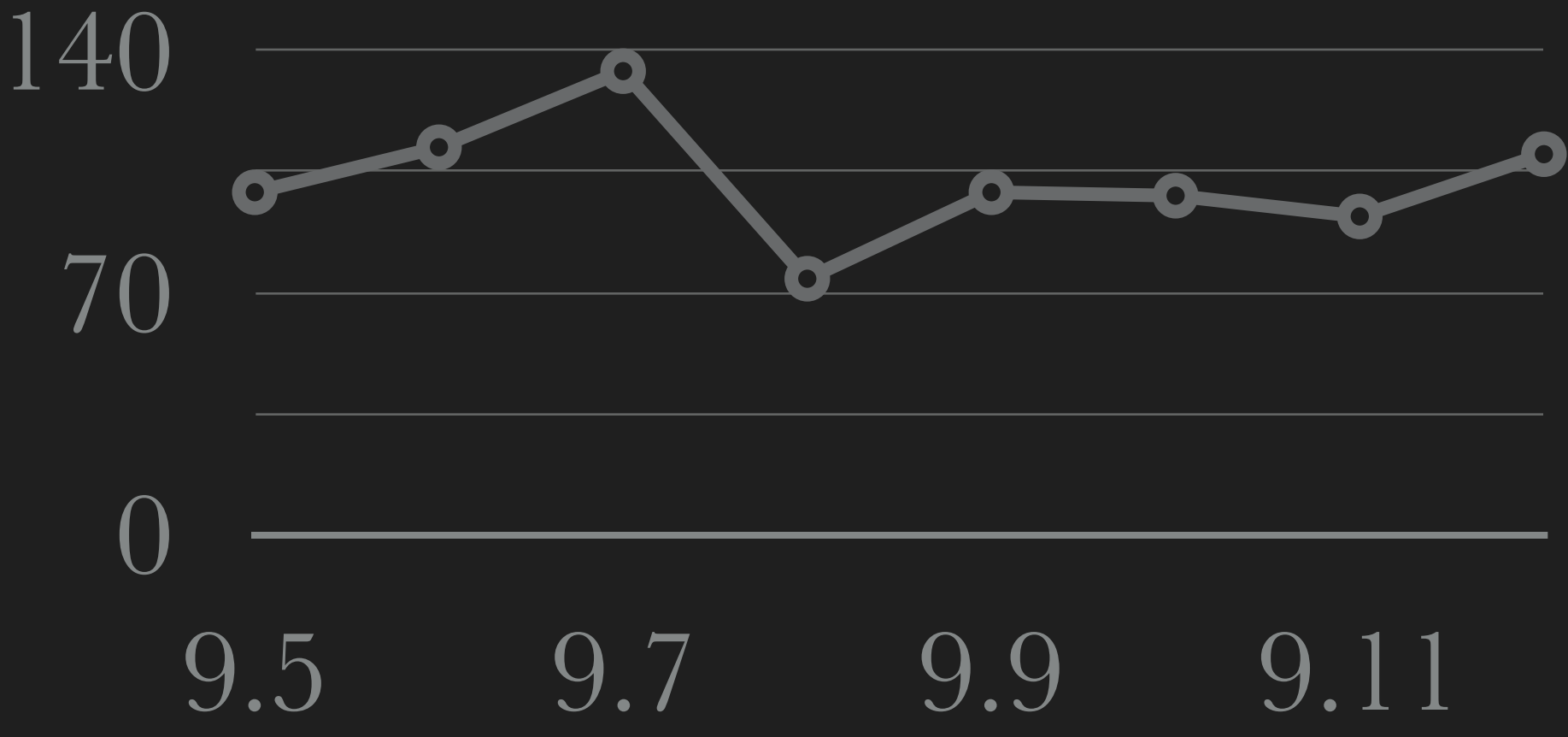
应用场景 — 设置安全库存



杭州仓销量



西南仓销量



华北仓销量

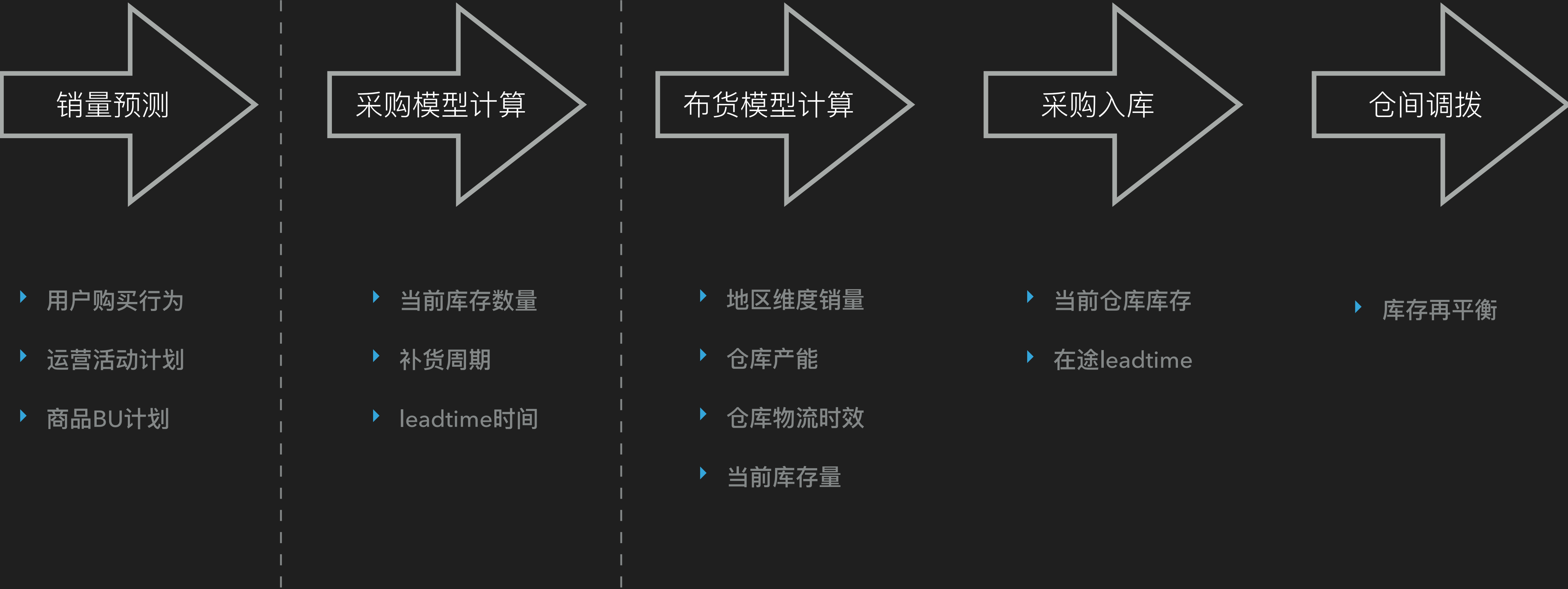
目标

- ▶ 避免缺货，默认仓满足率最大
- ▶ 避免来回调拨

TABLE OF CONTENTS 大纲

- 仓储物流供应链产品体系架构
- 大数据和实际场景中的痛点
- 人工智能&算法的应用
- 数据驱动案例 — 商品库存布货
- 思考

数据驱动案例 — 供应链全生命周期



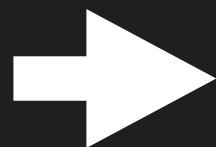
补货模型计算 — 线上订单履约流程



订单关键信息

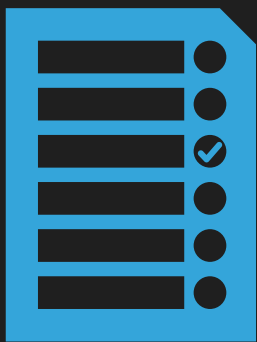
- ▶ 下单时间
- ▶ 投递地址
- ▶ sku+购买数量

+



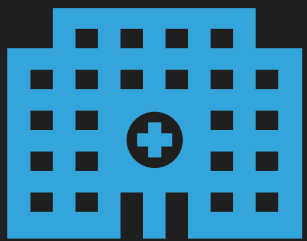
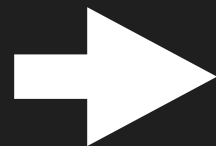
| | | | |
|-------|-------|-------|-------|
| 杭州顺丰仓 | 上午11点 | 浙江 | 当日达 |
| | | 江苏 | 当日达 |
| | | 江西 | 当日达 |
| | | 安徽 | 当日达 |
| | 下午6点 | 广东 | 次日达 |
| | | 广西 | 次日达 |
| | 下午8点 | 全国其他 | 隔日达 |
| 华北仓 | | | |

波次表

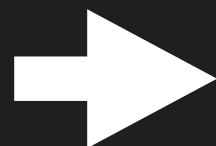


履约计划表

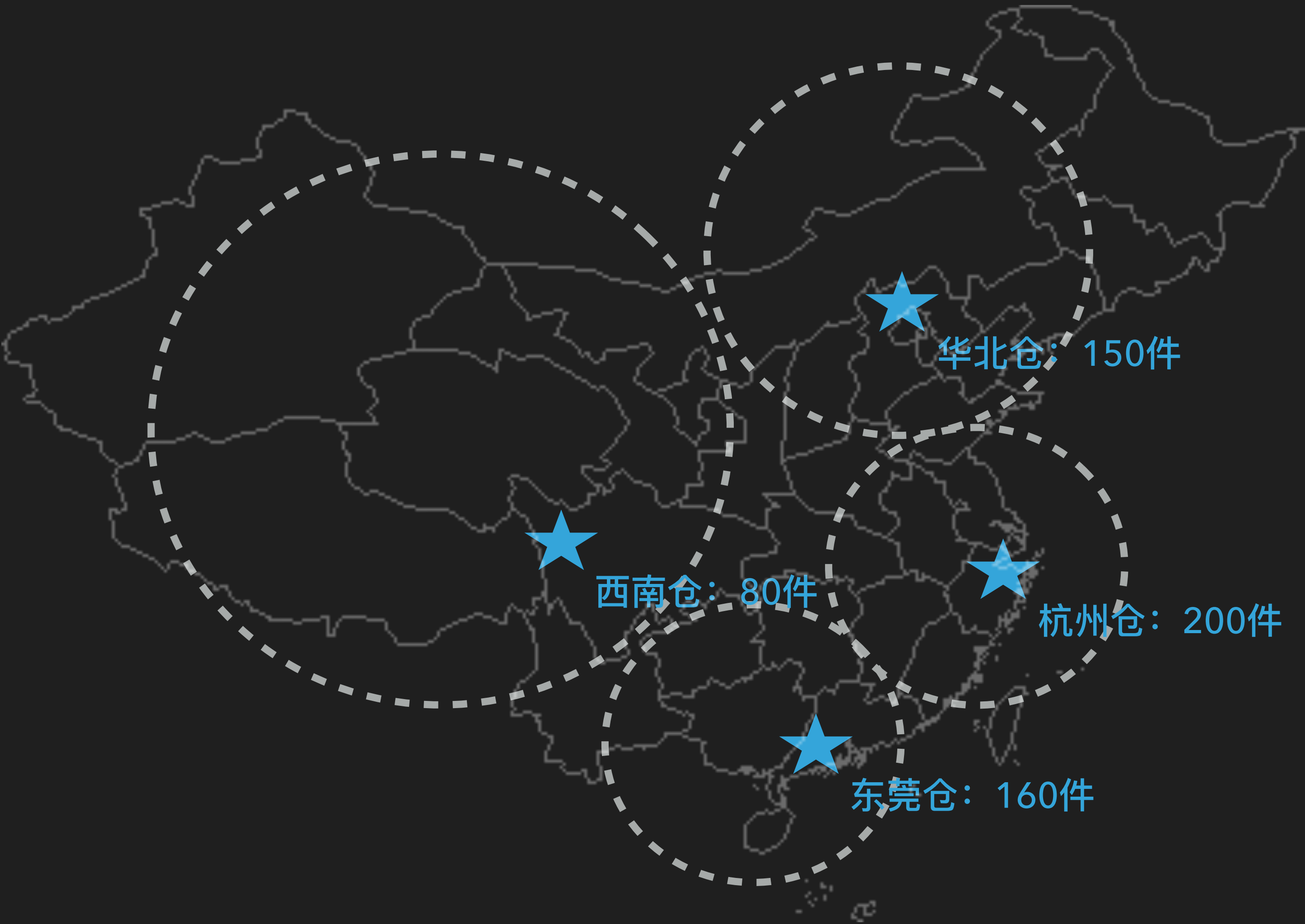
- ▶ 预计推仓时间
- ▶ 预计拣货时间
- ▶ 预计揽收时间
- ▶ 预计妥投时间
- ▶



仓库



快递商



理想库存分布模型

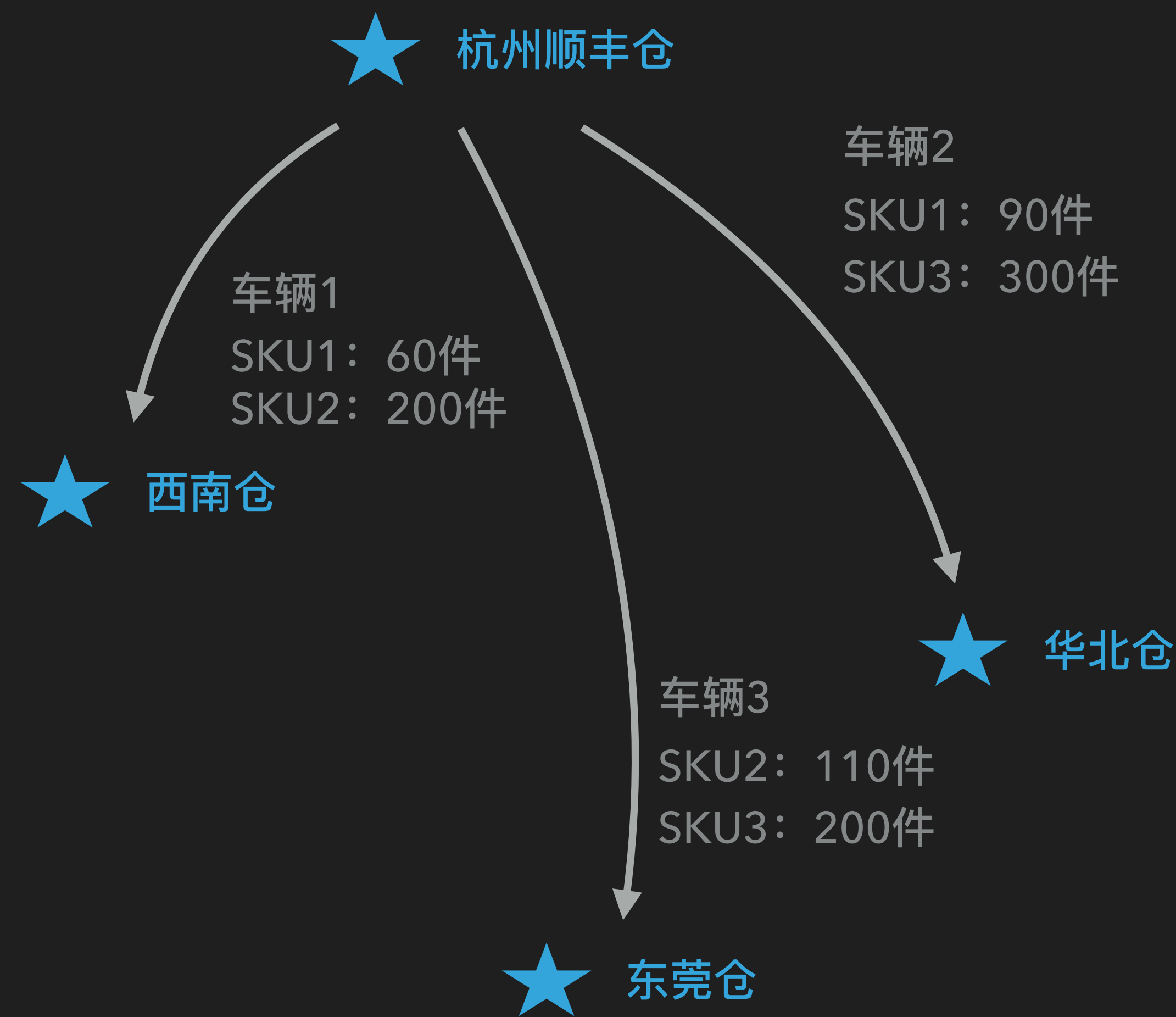
如何计算

- ▶ 获取分地区的销量预测
- ▶ 结合沙盘计算分仓比
- ▶ 考虑不同商品类目波动设置安全库存

限制条件

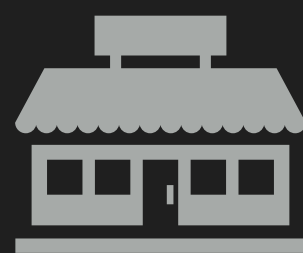
- ▶ 仓库容量
- ▶ 仓库吞吐量产能
- ▶

仓间调拨策略



| | | | |
|-------|-------|------------|----------|
| 杭州顺丰仓 | 车型1 | 40000 (kg) | 110 (m³) |
| | 车型2 | 30000 (kg) | 110 (m³) |
| | 车型3 | 25000 (kg) | 70 (m³) |
| | 车型4 | 8000 (kg) | 35 (m³) |
| 西南仓 | 车型1 | 40000 (kg) | 110 (m³) |
| | 车型2 | 30000 (kg) | 110 (m³) |
| | 车型3 | 25000 (kg) | 70 (m³) |
| 华北仓 | | | |

目标：成本最优（线性规划问题）



仓库出库

- ▶ 拣货时间预期
- ▶ 出库生产预期
- ▶ 揽收预期



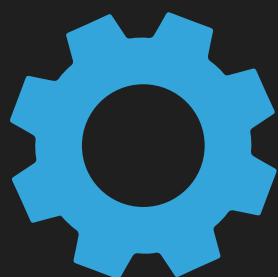
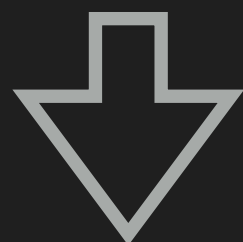
仓间运输

- ▶ 风险监控
- ▶ 预计送达时间



仓库入库

- ▶ 到达时间预期
- ▶ 入库时间预期
- ▶ 上架时间预期



决策中心

- ▶ 修改入库/调拨计划
- ▶ 修改在途订单计划
- ▶ 修改用户预期送达时间

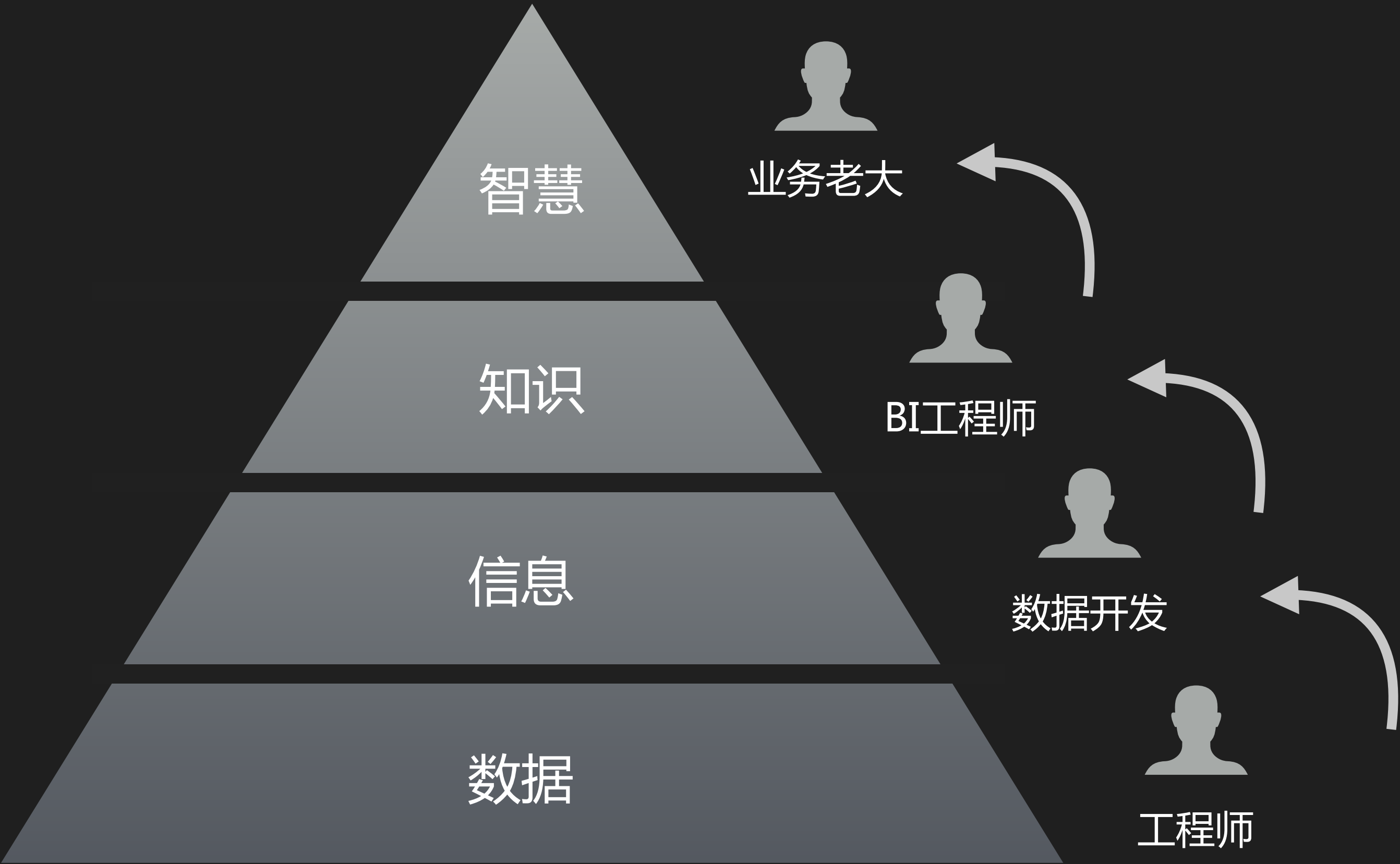
数据驱动案例总结



TABLE OF CONTENTS 大纲

- 仓储物流供应链产品体系架构
- 大数据和实际场景中的痛点
- 人工智能&算法的应用
- 数据驱动案例 — 商品库存布货
- 思考

数据驱动 — 数据金字塔



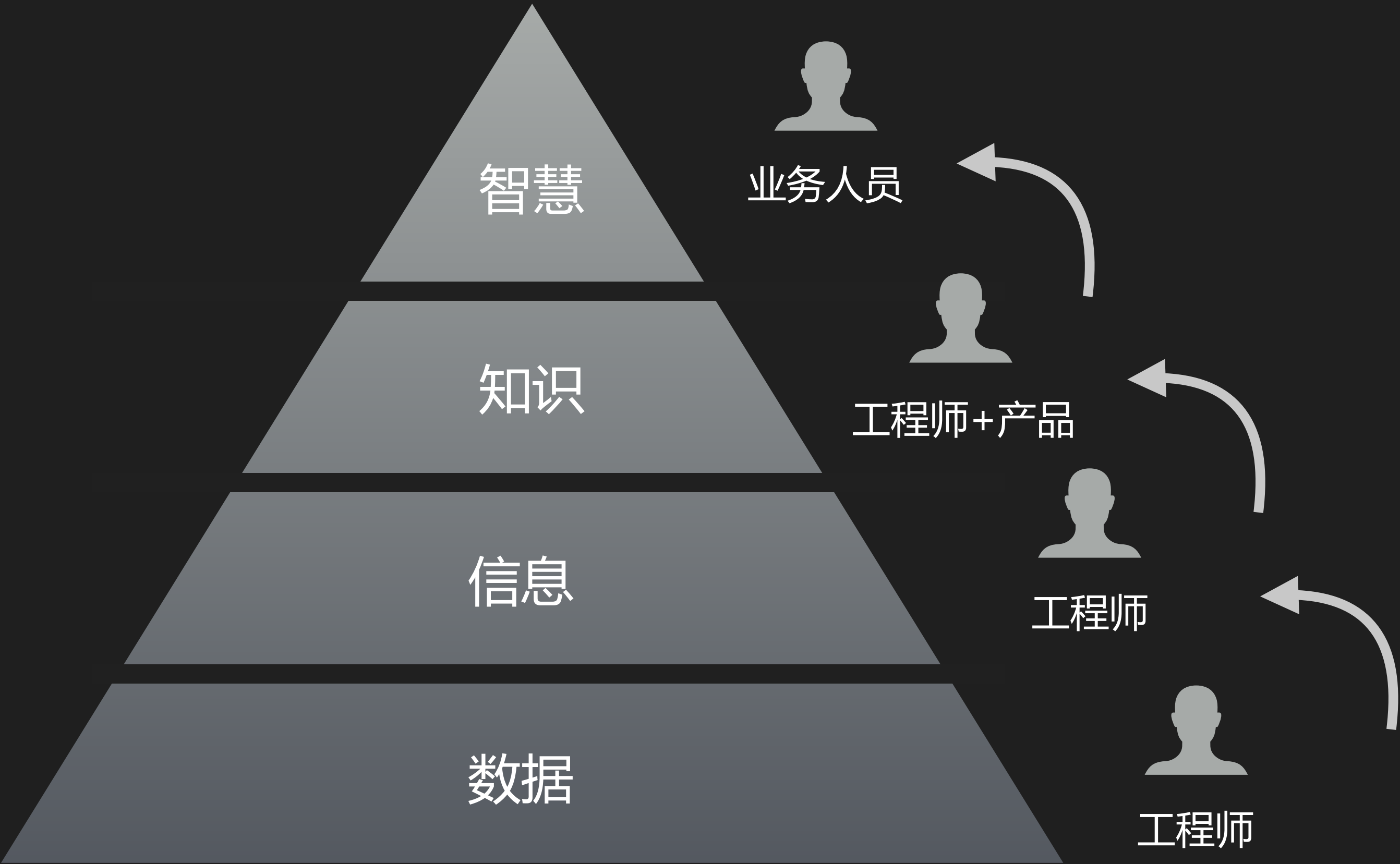
优势

- ▶ 组织分工明确
- ▶ 各职责专业化程度高
- ▶ 适用于开发全局类报表，帮助业务发现问题

劣势

- ▶ 沟通成本高
- ▶ 数据+工程+算法的综合优化问题，没有人全局把控

数据驱动 — 扁平化架构



改变

- ▶ 工程师兼BI和数据开发
- ▶ 算法由算法工程师完成
- ▶ 架构设计时把决策系统与执行系统分开
- ▶ 一个人能够了解整个事情
- ▶ 适用于优化类专业任务

劣势

- ▶ 对工程师综合要求较高
- ▶ 容易把工程师陷入事务性工作，需要产品辅助

THANKS