



JD.COM 京东

全球运筹优化挑战赛

Global Optimization Challenge

——赛题深度解析

01

城市物流运输车辆智能调度



概览

- **A榜赛题：城市物流运输车辆智能调度**
- **B榜赛题：按照赛制时间待公布**

问题一：城市物流运输车辆智能调度

业务背景

赛题概况

技术要点

解题思路

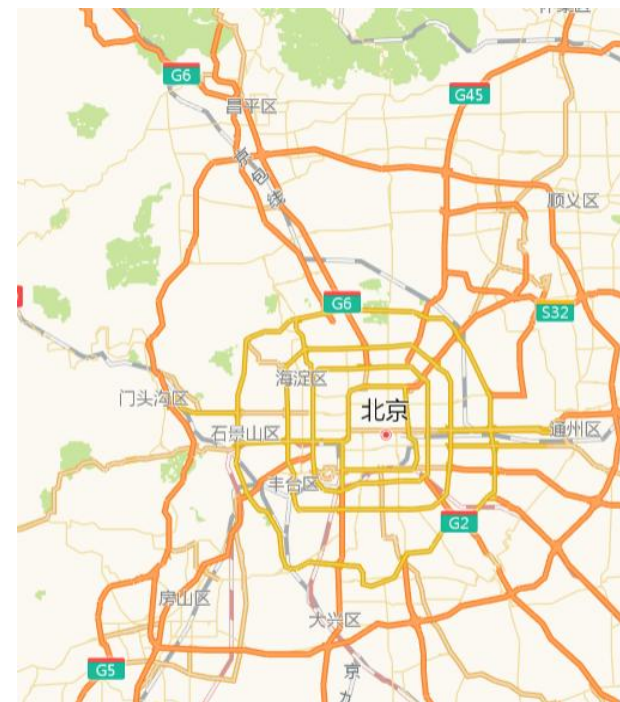
数据讲解

提交要求

常见问题

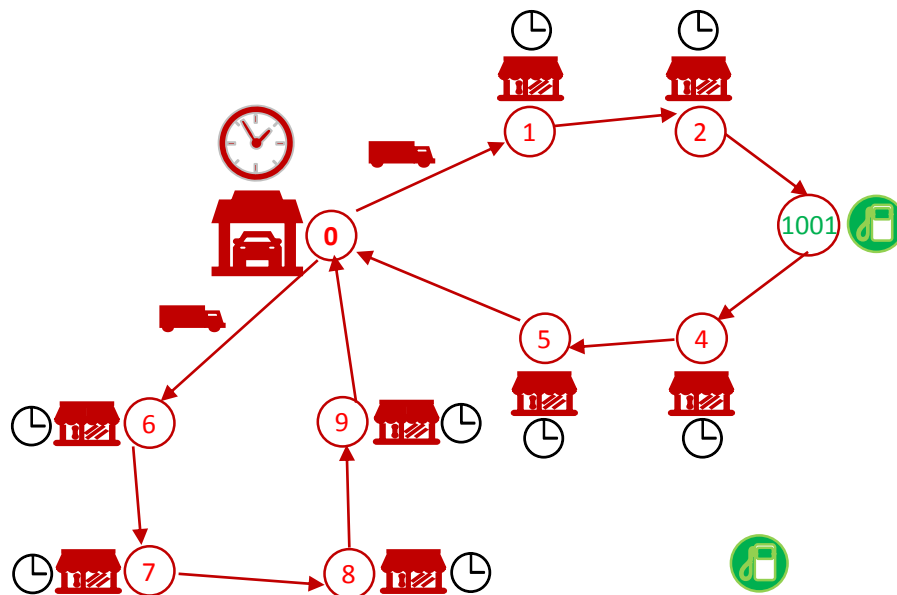
大规模城市物流城配车辆智能调度问题

- 以京东物流为例，城市A的城配物流中心B目前平均每天为分布在本城区的1000余个客户提供城市配送服务，服务对象为B2B或者大宗商品客户,每个服务对象有时效(时间窗)要求，每个车辆有容积和载重限制，车辆还需要充电，可往返配送中心，每天8点从城市配送中心出发，24点前返回本配送中心等。



城配范围

- 1个城配中心
- 1000个商家
- 100个充电站



真实脱敏数据

2018年某时期数据

- 商家基础数据及配送数据
- 配送中心数据及运力数据
- 充电站数据

假设条件 & 问题定位

假设条件

- 1) 车辆从配送中心出发，服务完客户后需返回配送中心，车辆可多次往返循环配送，发车时间为早晨8:00后(含)，回配送中心最晚时间为当日24:00；
- 2) 车辆到达客户时间必须在客户要求的最晚时间前到达(含)，同时，先于客户要求的最早到达时间则有等待成本；配送中心首次发车不计等待成本，其余等待均计算成本；车辆如果未被使用(即未访问客户)则不花费任何成本，车辆数不限；
- 3) 配送中心出发的车辆(包括多次往返配送中心场景)为满电状态，不计充电费用和充电时间，但对于多次往返配送中心场景，第二次起从配送中心出发需要在配送中心等待1h,计算等待成本；
- 4) 配送中心、充电站、客户之间的距离(m)和车辆行驶时间(min)以距离时间表形式体现；
- 5) 每辆货车在行驶中的车载质量和装货体积不超过该车型的质量和体积限制；
- 6) 客户都要被服务，且每个客户一天只能被一辆车服务，卸货时间恒定为0.5h, 装车时间不计；
- 7) 充电站充电费用为100元/h；
- 8) 充电站的充电桩无限制，车辆需要在可持续里程到达前在充电站充电，车辆一次充满，每次充电时间恒定为0.5h；
- 9) 上述等待成本的成本系数为24元/h。

问题定位

VRP(H,TW,G,C,T,M), NP-Hard

评测方式

总成本(TC)=车辆运输成本(DC)+车辆等待成本(WC)+车辆充电成本(CC)+车辆使用成本(FC)

其中：

DC为所有使用的车辆运输成本；

WC为所有使用车辆的等待成本；

CC为所有充电车辆的充电成本；

FC为车辆每天使用成本。

算法评分**主要依据TC**(越小越好)，即总成本最低，同等条件下，A榜以有效提交时间为依据。

解题思路

问题甄别：

最大容量约束、复合时间窗口约束、混合车辆约束、多循环约束、带充电站的电车里程约束。

建模与工具方法选择：

Metaheuristic , heuristic , etc.

宽进严出

宽进：可以考虑1000辆车，每辆车服务一个客户，需要充电则充电，否则返回配送中心。

严出：大规模NP-Hard 问题是理论难题，三甲难。



数据讲解

车型运力信息(参考input_vehicle_type.xlsx)

Col_name	Data_type	Is not null	Comment
vehicle_type_ID	VARCHAR	yes	主键,车型编号
vehicle_type_name	VARCHAR	yes	车型名称
max_volume	DOUBLE	yes	最高装载容积(m³)
max_weight	DOUBLE	yes	核定载重(t)
vehicle_cnt	INT	yes	车辆数量(台)
driving_range	INT	yes	持续里程(m)
charge_tm	DOUBLE	yes	充电时间(h)
unit_trans_cost	DOUBLE	yes	每公里运输成本(元)
vehicle_cost	INT	yes	车辆使用成本(元/天)

商家派送数据信息(参考input_node.xlsx)

Col_name	Data_type	Is not null	Comment
ID	INT	yes	主键,序号
type	TINYINT(2)	yes	类型,1代表配送中心, 2代表商家, 3代表充电站
lng	DOUBLE	yes	经度
lat	DOUBLE	yes	纬度
pack_total_weight	DOUBLE	yes	包裹总重量(t)
pack_total_volume	DOUBLE	yes	包裹总体积(m³)
first_receive_tm	TIME	yes	商家最早收货时间,HH:MM
last_receive_tm	TIME	yes	商家最晚收货时间, HH:MM



数据讲解

距离时间信息(参考input_distance-time.txt & 距离时间数据说明.docx)

Col_name	Data_type	Is not null	Comment
ID	INT	yes	主键,序号
from_node	INT	yes	起点,0配送中心, 1-1000商家, 1001-1100充电站
to_node	INT	yes	终点,0配送中心, 1-1000商家, 1001-1100充电站
distance	INT	yes	两点间距离, 单位米, 整数。
spend_tm	INT	yes	两点间行走时间, 单位分钟, 整数

更详细数据示例解析（赛题与数据中2.1&3.2）

附：A榜赛题数据列表

序号	文件名	文件内容	MD5验证	数据大小 (kb)	记录条数 (条, 不含表头)
1	input_node.xlsx	商家配送数据	654eac873b9e88109335bae2944b7dcb	78	1,101
2	input_vehicle_type.xlsx	车型运力数据	b97b2aa5d125e79e01e5394cf175045f	11	2
3	input_distance-time.txt	距离时间表	6c43262a41d9338f9e0911548436213c	29600	1,211,100
4	距离时间数据说明.docx	距离时间数据说明	e891f8c2c23a073cd5827c7054236bfb	15	5

温馨提示：如需下载赛题数据，请先登录并完成报名参赛。

下载赛题数据

由京东万象提供下载服务

点击下载赛题数据

提交要求

- 选手每日可以提交2次结果进行评测；
- 考虑自主知识产权，决赛不建议使用商业优化软件，如CPLEX等；
- 考虑大赛的公平性及实际价值，暴力求解方法(如枚举等)原则上不进入前100名；
- 严格按照作品要求提交文件(如列名称，数据精度等)，不得缺失或提交虚假内容，经沟通仍未在规定时间内提交的，将进行诚信记录，影响参赛资格及奖励；
- 作品包括，**结果文件和代码文件**，结果文件为CSV格式(表输出数据格式)，代码文件包含源代码、使用安装说明、最终结果分析说明(含硬件配置、CPU计算时间，过程结果)。



提交要求

更详细解析（参考赛题与数据中04文字描述&示例数据）

fixed_use_cost	INT	yes	固定成本(元)
total_cost	DOUBLE	yes	总成本(元)
charge_cnt	INT	yes	充电次数

[作品要求详情下载](#)

点击下载赛题数据

派车调度结果样例信息(参考Result.csv)

Trans_code	Vehicle_type	Dist_seq	Distribute_lea_tm	Distribute_arr_tm	Distance	Trans_cost	Charge_cost	Wait_cost	Fixed_use_cost	Total_cost	Charge_cnt
DP0001	1	0;2;100;3;10;12;0	8:00	13:00	63489	761.87	0	0	200	961.87	0
...
DP1000	2	0;9;1006;0;106;23;0	9:24	13:30	3458	48.41	50	0	300	398.41	1

备注：每一行记录代表一辆车一天内的往返记录，其中，里程(m)、运输成本(元)、充电成本(元)、等待成本(元)、固定成本(元)、总成本(元)、充电次数依据当前节点计算，其中运输成本、等待成本、总成本精确到小数点后两位，里程精确到m。



常见问题

1. 问：车辆是否单程不返回？ 答：车辆从配送中心出发最终须回到配送中心。
2. 问：1000个商家数据是否分布在几个区域？ 答：请选手自己分析数据。
3. 问：车辆中途回到配送中心是否必要, 因为还要计算等待费用？ 答：赛题源于实际场景。
4. 问：车辆充电是否需要排队等？ 答：充电站内充电桩不做限制。
5. 问：是否可以用求解器？ 答：决赛环节不提倡使用，鼓励自主开发。
6. 问：可以用启发式算法吗？ 答：可以，算法不限。
7. 问：近似解可以吗？ 答：可以。
8. 问：评测代码 是否会公开？编程语言是否有限制。 答：评测代码不会公开，赛制规定每天可以提交两次，参赛者可以利用这些机会自检答案。编程语言不限。
9. 问：关于“第二次起从配送中心出发需要在配送中心等待1h, 计算等待成本”，是否可以理解为，同一辆车执行多次配送，则连续的两次中间【至少】需要等待1h，那么如果我让车等待2h，那么等待成本是按照1h计算，还是2h计算？ 答：按照2h计算，比如路径0;1;2;0;3;4;5;0;6;7;0，中间两次重新从配送中心出发，则等待时间2h，等待成本按照2h计算。第二次开始每次返回配送中心再出发都要在配送中心等待1小时。
10. 问：车辆每次在充电站充电的0.5h是否算作是等待；车辆每次的卸货时间0.5h是否算作是等待。 答：充电时间和卸货时间都不计入等待时间。
11. 问：车辆一趟是可以送多个客户吗？ 答：可以送多个客户。
12. 问：算法是否需要考虑鲁棒性？ 答：优秀的算法不仅仅需要考虑鲁棒性。
13. 问：提交结果时提示各种问题？ 答：请选手仔细阅读“作品要求”，严格自查。

02

仓储网络库存管理

- **初赛问题一：区域销售预测**
- **初赛问题二：仓储网络库存调拨**
- **复赛问题 = 初赛问题一 + 初赛问题二**

问题一：区域销售预测

业务背景

赛题概况

技术要点

解题思路

数据讲解

提交格式

京东如何做到618海量商品精准极速配送？

高效靠谱的物流体系

多层级仓储网络结构

准确的区域销量预测



销量预测范围

- 6个销售区域
- 1000商品 (SKU)
- 2018年1月1日-2018年1月31日

真实历史数据 2016年初-2017年底

- 商品销量数据
- 商品属性信息
- 商品价格与促销信息

均值预测 vs. 分位点预测

赛题中要求进行分位点预测

满足销量均值不是京东履约的目标

需要满足绝大多数的销量情况

能够应对销量突增等较小概率事件

对高分为点目标进行预测

对销量分为点 $q\%$ 进行预测，即 $q\%$ 概率情况下当日销量将不会超出预测值

评测公式

$$L = \sum_{ij} L_{ij}(q_i)$$

$$L_{ij}(q_i) = \frac{\sum_t \left[q_i (y_{ijt} - \hat{y}_{ijt}^{q_i})^+ + (1 - q_i) (\hat{y}_{ijt}^{q_i} - y_{ijt})^+ \right]}{T \sum_t y_{ijt}}$$

i 代表商品， j 代表仓库， t 代表时间

$L_{ij}(q_i)$ 代表商品 i 在销量区域 j 在事件 t 的 q_i 分位点的预测误差

每个商品的 q_i 事先给定

$\frac{1}{\sum_t y_{ijt}}$ 为评价权重，为使销量较低的商品在评价时权重升高

分位点预测的含义

如果预测偏高，会受到 $1-q_i$ 的惩罚

如果预测偏低，会收到 q_i 的惩罚

赛题中 q_i 的取值为0.85-0.99

所以预测偏高惩罚少，预测偏低惩罚多

$$\sum_t \left[(1 - q_i) (y_{ijt} - \hat{y}_{ijt}^{q_i})^+ + q_i (\hat{y}_{ijt}^{q_i} - y_{ijt})^+ \right]$$

如何提升成绩

理解分为点评测标准，预测 q_i 分为点的销量

利用历史数据，特别是时间上相近与往年同期的销量数据

探索强特征

注意选取训练和验证数据集的方式

学习利用python现成的quantile regression的分析包框架



数据讲解

商品基本信息

Column_name	Description
item_sku_id	SKU唯一识别代码
item_first_cate_cd	一级品类
item_second_cate_cd	二级品类
item_third_cate_cd	三级品类
brand_code	品牌

商品属性信息

一个SKU可对应多个属性

一个属性可以对应多个值

Column_name	Description
item_sku_id	SKU唯一识别代码
attr_cd	属性代码
attr_value_cd	属性值

机身颜色	黑色
机身长度 (mm)	141.2
机身宽度 (mm)	71.45
机身厚度 (mm)	7.92
机身重量 (g)	153
输入方式	触控
运营商标志或内容 ?	无
机身材质分类	金属边框 ; 玻璃后盖
屏占比 ?	74%

某手机属性表

某属性上多个属性值
attr_value

多种属性
attr

销售信息

Column_name	Sample_data	Description
item_sku_id	36	SKU唯一识别代码
dc_id	1	仓库代码
date	2017/2/13	日期
quantity	1	销量
vendibility	0	当日结束库存状态
original_price	0.0373797	当日价格(0-1)
discount	10	销售折扣(0-10)

促销信息

Column_name	Description
item_sku_id	SKU唯一识别代码
item_third_cate_cd	三级品类
date	日期
promotion_type	促销形式代码

更详细数据解析EDA下载（FAQ中第一个问题）

赛制介绍

赛题与数据

FAQ

1. Q：哪里可以看到数据说明与提交要求。

A：赛题与数据页面中点击“下载查看完整赛题”可下载赛题详情PDF。赛题说明与数据说明会根据选手的反馈进行必要的更新，请各位参赛者保持关注比赛页面，下载最新的赛题与数据说明。我们提供了一些[初步分析](#)。

点击下载
赛题数据分析



提交格式

- 预测量为 $6*1000*31=186000$ 行
- 选手每日可以提交2次结果进行评测
- 注意提交文件需要有列名称行，销量数字只取小数后两位

date	dc_id	item_sku_id	quantity



概览

问题二：仓储网络库存调拨

业务背景

赛题概况

仿真程序

评测方法

解题思路

数据解析

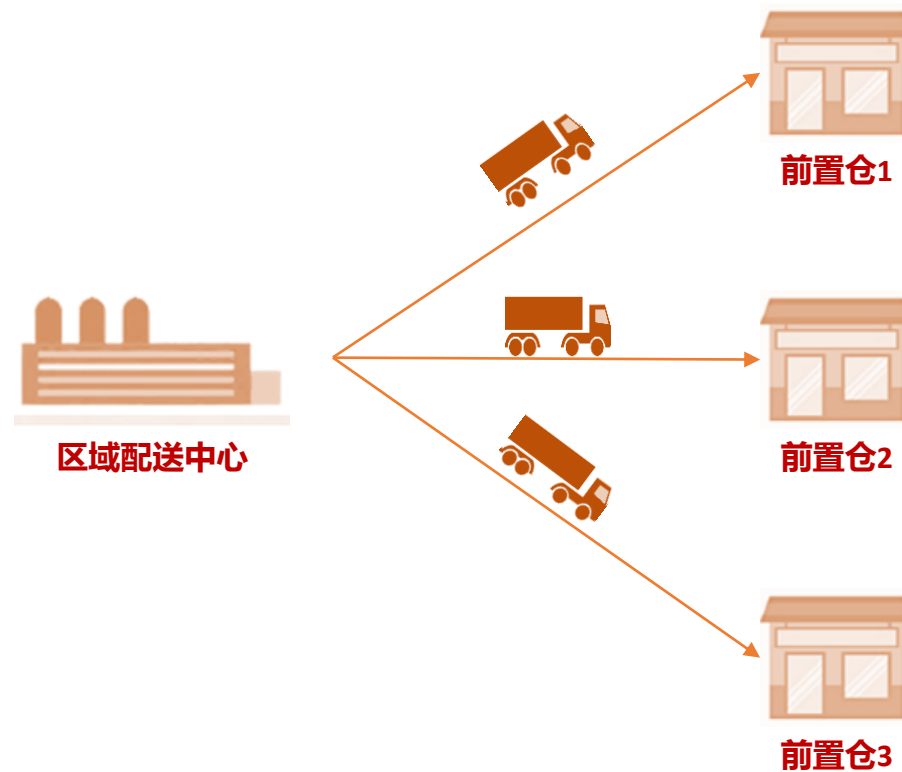
提交格式

京东如何做到618海量商品精准极速配送

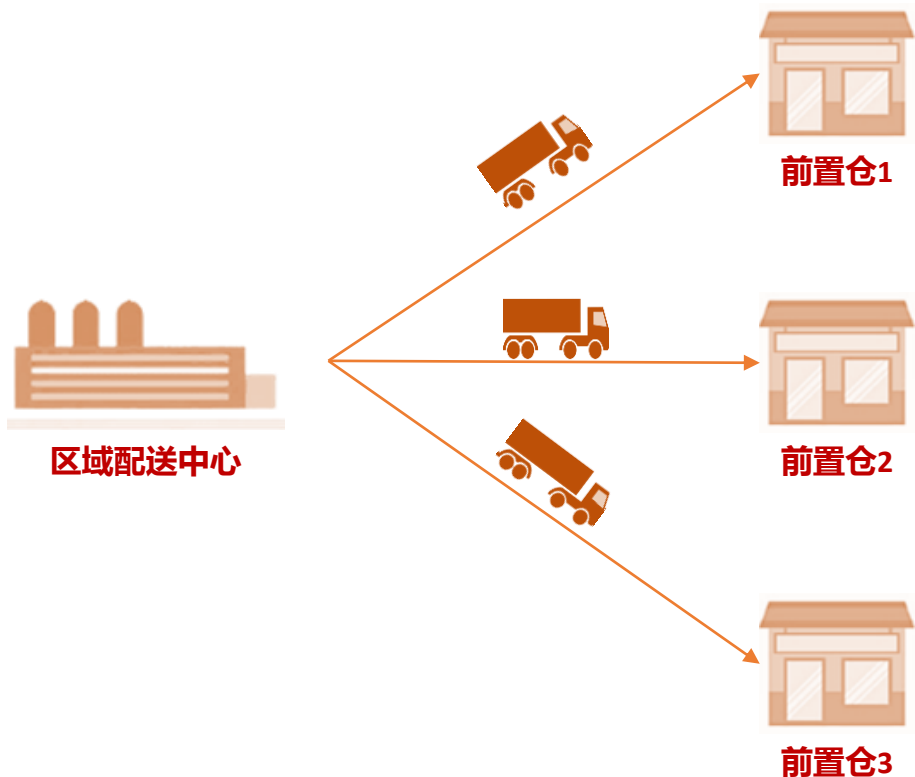
- 高效靠谱的物流体系
- 多层级仓储网络结构
- 准确的区域销量预测

二级仓储配送网络结构

区域配送中心（RDC）和前置仓（FDC）



赛题概览



- **RDC接受供应商补货，FDC接受RDC调拨**

- **调拨是从RDC向FDC补充商品储备的过程**

- **调拨决策每日进行（早晨调拨，夜间达到）**

- **调拨决策需要对每一个FDC每一个SKU进行**

- **赛题调拨范围：**

- 1个RDC，5个FDC

- 1000个SKU

- 30天

- **调拨运力限制**

- SKU调拨种类上限

- 总调拨数量

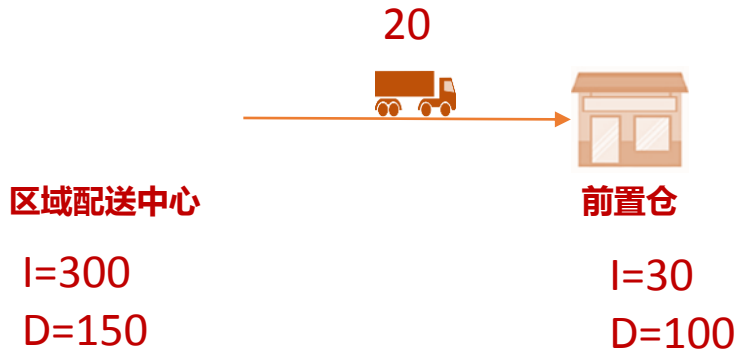
每天依序进行

- 进行调拨决策（参赛队伍程序输入）
- 验证调拨决策正确性，观察当日顾客需求，进行履约，记录成本
- 接收当日到达 RDC 和 FDC 的补货与调拨量

仿真评测程序simulation.py

- 每组销量模拟数据仿真30天
- 共100组模拟销量数据
- 取平均值作为最终成绩

评测方法



成本计算方法（案例说明）

初始库存：FDC库存30，在途库存20，RDC库存300

当日需求：FDC需求100，RDC需求150

- 首先FDC库存满足本地需求30
- 本地需求70不能满足
- 因时效原因造成10%直接销量损失， $70 \times 0.1 = 7$ （公式一）
- 其余不能满足部分 $70 - 7 = 63$ 首先利用在途库存20满足
- 不能满足部分 $63 - 20 = 43$ 向RDC申请支援
- RDC首先考虑满足自身需求150，剩余库存 $300 - 150 = 150$
- RDC支援FDC需求43，产生额外履约成本（公式三）
- RDC最终库存为 $150 - 43 = 107$ ，不产生RDC销量损失
- 如果RDC库存不足以支持FDC，则产生RDC销量损失（公式二）

假设SKU价值为 p_i ，额外履约成本为 q_i ，则当日由该SKU造成的总成本为

$$7p_i + 43q_i$$

详细成本计算方法请查看完整赛题描述

解题思路

理解销量分布（外部参数）

- G代表伽马分布（波动相对较小）
- N代表负二项分布（波动相对较大）

Column_name	Description
item_sku_id	SKU唯一识别代码
dc_id	仓库代码
dist_type	销量分布名称
para1	第一系数
para1	第二系数

思考：库存调拨的核心是什么？

- 调拨过少，造成直接销量损失和额外履约成本
- 调拨过多，造成RDC库存减少，之后可能无货可调，影响其他FDC库存



下载仿真程序simulation.py，进行线下测试

数据解析

FDC_ID	sku_Limit	capacity_Limit
1	200	3200
2	200	1600
3	200	1200
4	200	3600
5	200	1600

FDC_ID	abandon_rate
1	1%
2	7%
3	10%
4	9%
5	8%

● 运力限制

思考：运力不够时如何决定调拨次序？

● 补货时间表

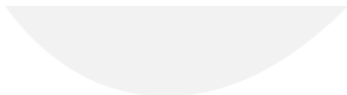
思考：补货即将到来时调拨决策是否应该变化？

● SKU价值差异

● FDC本地销量损失比例差异

● 更详细数据解析EDA下载（FAQ中第一个问题）

- 选手需要提交simulation.py
- 仿真程序simulation.py每天调用submission.py中
- UserPolicy类函数
- 选手具体修改UserPolicy即可
- 选手提交代码需要在要求时间内运行完成
- 选手每日可以提交2次代码



```
.
├── simulation.py
├── submission.py
├── initial_inventory.csv
├── inventory_replenishment.csv
├── sku_demand_distribution.csv
└── sku_cost.csv
```



GOC选手交流QQ群：
729162237



GOC大赛选手群

扫一扫二维码，加入群聊。