

从一顿晚饭，到一份临时礼物

社区零售 AI补货实验

把日常购物观察，转化为
需求预测、库存策略与IP商品数据设计。

个人项目 2026.06 - 2026.09

职业间歇期研究 · 本次完成公开数据离线回测与方案整理



非美团或其他品牌委托项目。使用真实公开零售数据做方法验证；业务场景设计、插画与决策界面均为原创示意。

晚餐要准时，礼物要有货，库存不能都一样

观察起点是即时到家的便利性。模型边界收敛为“自营社区零售库存”，不混入餐饮商家的食材库存。

晚餐补货

牛奶、生鲜与日用品承担高频需求。供应不足影响体验，过量又可能损耗。需要按保质期、时段和到货节奏制定策略。

临时礼赠

礼物和主题商品有节日、内容热度及上新波动。高热度未必能持续，需要区分日常基线与事件增量。

同一决策入口

先预测需求，再校验库存与供应能力；解释为什么补、补多少、何时到，以及什么条件下不补。

现实参照：小象超市的自营服务站模式

美团2023年公开介绍：社区服务站集存储、分拣与配送于一体，定位“30分钟快送超市”。本项目借鉴这种业务边界，不使用其内部交易或库存数据，也不宣称验证其履约能力。

[M1] 美团官方品牌升级说明。插画为原创小熊，不使用袋鼠、小象或品牌官方Logo。

先把能验证的部分，和不能验证的部分分开

公开数据是一家英国线上礼品零售商的交易记录；它可检验订单预测方法，但不能直接代表中国即时零售。

541,909

原始交易行

5,268

完全重复行

476,637

过滤后英国正向交易行

保留的事实

UCI Online Retail，2010.12.01 - 2011.12.09。字段包括发票、SKU、描述、数量、时间、单价、客户和国家；货币为GBP。
去重后排除取消、非正数量和非正价格；仅英国，且商品码匹配五位数字及可选字母。

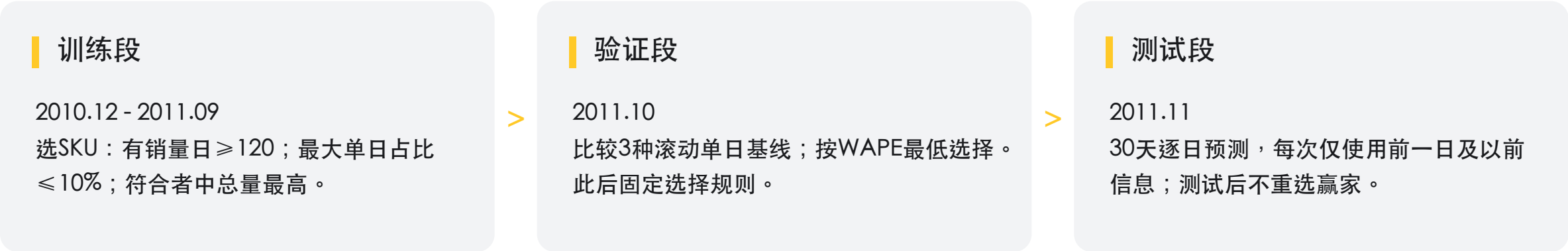
不能偷换的口径

分析目标是正向订单数量，不是净销量或真实消耗。正向订单可能之后被取消，本次未配对冲销。
无库存、缺货、交期、天气、IP授权与本地仓字段；缺失日期补0仅表示没有记录，不证明需求为0。

[U1] Daqing Chen (2015)，CC BY 4.0。过滤后行数覆盖原文件全部日期；训练与测试时间在下一页单独限定。

先锁规则，再看留出的30天

不为了让结果好看而挑选测试期。选择条件仅使用2011年9月30日及以前的数据。



选中SKU 84077

WORLD WAR 2 GLIDERS ASSTD DESIGNS

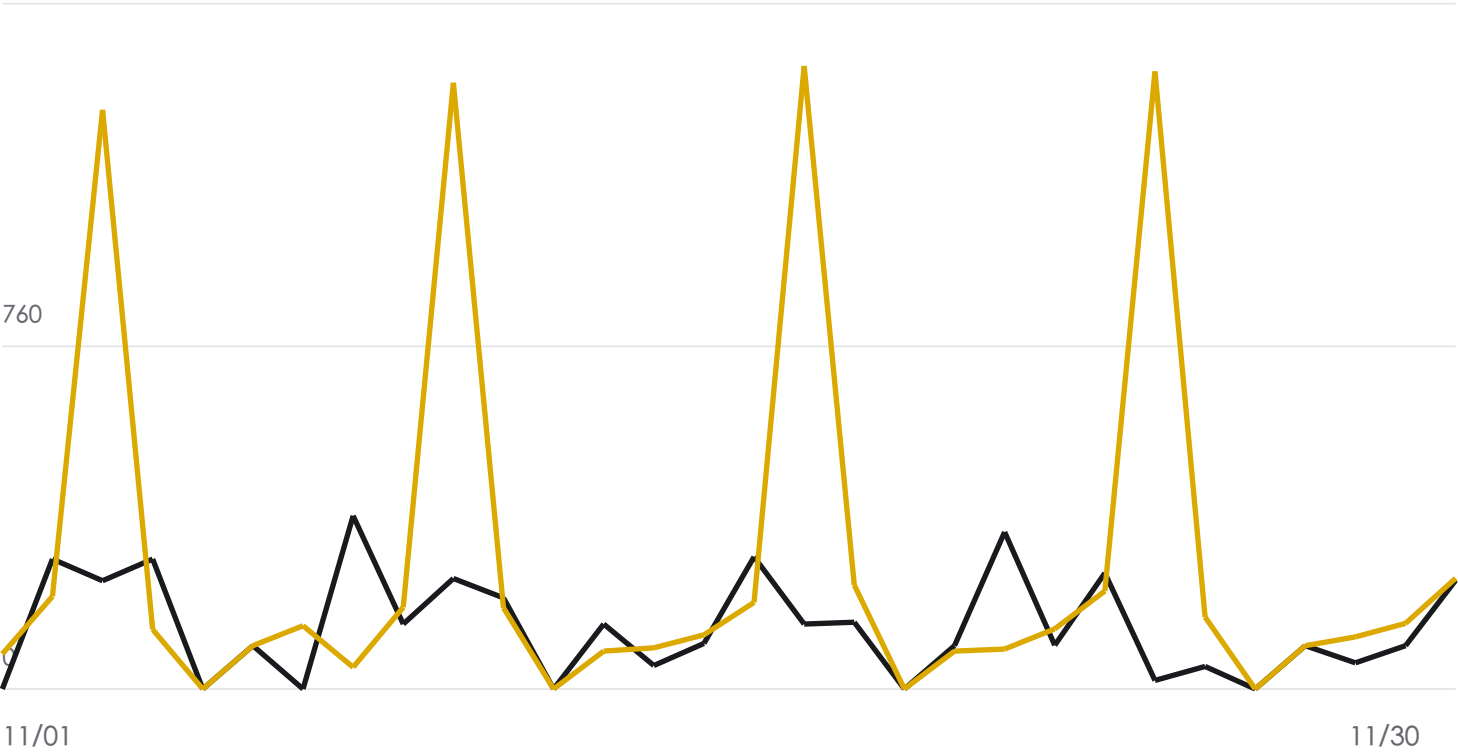
该商品为玩具滑翔机，属于礼赠类测试样本，不等于授权IP产品。训练段共35,996件。保留批发订单波动，并披露预测失败；不能将本样本外推成多门店准确率。

基线：上周同日；过去4个同星期日均值；过去28天日均值。末尾不完整的2011年12月不参与回测。[U1 / F1]

真实结果：模型还不具备自动补货条件

2011年11月留出期。选中模型为“过去4个同星期日均值”；图中预测为滚动一步预测。

单位：件 / 日
1519



黑：实际订单 黄：预测订单

156.3%

WAPE / 越低越好

+104.6%

预测偏差 / 显著高估

11月实际4,127件。模型受前期波动影响；销量稀疏与批发大单，使简单周期规律难以直接迁移。

WAPE = $\sum | \text{预测} - \text{实际} | / \sum \text{实际}$ ；偏差 = $\sum (\text{预测} - \text{实际}) / \sum \text{实际}$ 。WAPE可能超过100%，不等于“负准确率”。

不隐藏失败，才能决定下一步该改什么

数值来自附在PDF内的results.json与daily_backtest.csv；可在支持附件的PDF阅读器中提取。

模型	10月验证WAPE	11月测试WAPE	11月偏差
上周同日	93.8%	179.7%	+111.2%
同星期4次均值（选中）	91.0%	156.3%	+104.6%
过去28天均值	98.7%	102.8%	+88.7%

先拆订单类型

批发与零售混在一起。下一版先按客户历史订单量分群，再分别建模；分群规则也须只用历史。

再补业务约束

接入真实缺货、促销、营业日、交期和可售库存。当前销售记录不足以还原未满足的需求。

保留独立复测

28天均值测试期更好，也不能回头宣布它是预先选中的赢家。迭代后需新的留出期检验。

结论仅适用于这一SKU与这一时间段；不发布虚构的降本、库存下降或“AI精度提升”指标。

库存策略先做沙盘，不让失准预测直接下单

以下是独立的常温商品情景：用于展示规则正确性，不使用上一页失准预测自动生成采购单。

常温补货演算

日均需求10件；日标准差3件

交期2天；每天复核一次

95%周期服务水平， $z = 1.645$

现货18，在途8，待交付4

库存位置 = $18 + 8 - 4 = 22$ 件

$SS \approx 9$ 件 $S = 39$ 件

$Q = 17$ 件 → 整箱后20件

鲜品：先看可售剩余寿命

按到期批次执行先到期先出。用到期前累计可售需求限制进货，扣除同批次可售量与到货；箱规若导致超限，转为不补或人工拆箱。

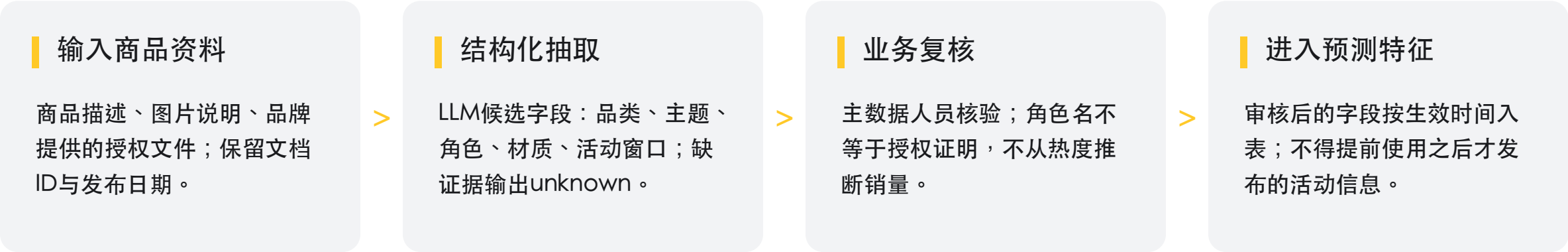
IP礼赠：先看上新与授权周期

日常基线与事件增量分别估计。增加角色/主题、上新日、活动窗口与授权截止字段；热度只作特征，不能直接按倍数放大采购量。

[1] $S = 10 \times (2+1) + \text{ceil}(1.645 \times 3 \times \sqrt{3}) = 39$ ； $Q = \max(0, 39 - 22) = 17$ ，箱规5件向上取整20。所有输入为假设。

AI把描述变成可检查的商品信息

已经完成的是AI辅助材料整理、代码编写和公开数据回测；以下IP抽取与工具编排是设计原型，尚未接入生产。



示例：公开样本84077的字段卡

品类：玩具滑翔机（由描述人工确认） 主题：二战（描述文本）
角色IP：unknown 授权方：unknown 上新日期：unknown
抽取依据：WORLD WAR 2 GLIDERS ASSTD DESIGNS
原则：未知字段保持未知；本次未执行批量LLM标注，也未虚构IP销量数据。

AI辅助写代码不等于AI预测模型。数值由可执行代码计算；IP语义特征需单独标注验证，不能用“智能化”替代证据。

一张能看懂、能拒绝、能追溯的补货卡

界面原型。Agent仅在授权范围内调用查询与计算工具；采购动作始终由业务审批。

社区仓 / 补货建议工作台

状态：阻断自动采购

原因：样本回测WAPE 156.3%，业务适配未经验证。

动作：保留人工计划，不创建采购单。

可用证据：回测CSV、数据清洗规则与模型版本。

待补信息：仓级销售、缺货、可售库存和供应交期。

查看证据 | 人工接管 | 保存分析

工具边界

query_snapshot只读；forecast返回误差与版本；simulate_policy计算情景；create_order仅接收有效审批ID。

恢复机制

时间戳过期即停止；超时可重试但不重复下单。工具返回不足时输出“无法判断”，不让语言模型猜库存。

[A1] 架构参考。示例阻断是本作品的设计决策，不是美团系统阈值；实际阈值由业务成本与服务目标共同确定。

从离线实验，到可上线方案还差哪些证据

研究期按个人安排标为2026.06 - 2026.09；本次回测与作品整理于2026.09完成，不伪造逐月实施记录。

状态	本次交付	下一次验证
已完成	公开数据清洗、单SKU时序回测	复现结果；扩大SKU并报告分层表现
已设计	补货沙盘、IP字段卡、Agent工具边界	人工标注集；供应约束与异常用例测试
未完成	仓级真实需求、库存与履约回测	经授权接入真实数据；开展影子运行
不宣称	美团内部经历、投产系统或降本收益	只有实际执行与对照证据齐备后才报告

影子运行只输出建议，不写采购；同时观察满足率、损耗、库存周转与资金占用。服务水平改善不能以不可接受的报损和库存增长换取。

可复用资产：选样规则、回测脚本、预测CSV、库存公式与审批机制。原始数据另由UCI公开下载，未嵌入整个数据集。

完整来源与复现入口

公开链接可点击；资料查询与回测日期：2026-09-28。没有引用无法追溯的行业营销数字。

M1

美团 · 美团买菜升级为“小象超市”

2023-12-01；仅用于自营社区服务站与30分钟快送定位的场景说明。
<https://www.meituan.com/news/NN231201067004457>

U1

UCI · Online Retail / Daqing Chen (2015)

541,909条交易；CC BY 4.0；数据期2010.12.01 - 2011.12.09。
<https://archive.ics.uci.edu/dataset/352/online%2Bretail>

F1

Forecasting: Principles and Practice (3rd ed)

第5.10节；训练、验证与滚动时序检验的参考方法。
<https://otexts.com/fpp3/tscv.html>

I1

NC State University · Periodic Review System

2011-01-28；周期复核库存策略参考，具体数值为本作品情景。
<https://scm.ncsu.edu/scm-articles/article/periodic-review-system-inventory-management-models-a-tutorial>

证据不止在参考文献里

本PDF附带分析脚本和计算结果，避免只有图表、没有复算入口。

A1 Anthropic · Building effective agents
2024-12-19；工具边界、工作流与人工检查。
<https://www.anthropic.com/engineering/building-effective-agents>

U2 UCI数据集永久标识
Daqing Chen，Online Retail，DOI：10.24432/C5BW33。
<https://doi.org/10.24432/C5BW33>

U3 原始Excel压缩包下载
原始下载包SHA-256见下；本作品使用Python、pandas与openpyxl读取。
<https://archive.ics.uci.edu/static/public/352/online+retail.zip>

附加文件：analyze.py（计算逻辑）、results.json（指标）、daily_backtest.csv（日序列）。
SHA-256：f5385cbb54bbebf7196389109c6b0621faab0c304e3702548165e71c84aede8b

复现：下载ZIP放入脚本同级data目录，安装pandas与openpyxl后运行analyze.py。插画及图表为矢量原创，无品牌授权暗示。