

## 赛道 B：电商零售商家需求预测及库存优化问题

电商平台存在着上千个商家，他们会将商品货物放在电商配套的仓库，电商平台会对这些货物进行统一管理。通过科学的管理手段和智能决策，大数据智能驱动的供应链可以显著降低库存成本，同时保证商品的按时履约。一般来说，以上供应链优化问题会包含以下方面：

### 需求预测

预测往往是智能供应链的决策基础，它可以让管理者提前预知各地的需求，从而将库存提前放在靠近需求的仓库中，此时的预测任务为：根据历史一段时间的需求量，预测各仓库中各商品未来需求，“预测维度”即为不同商家在各仓库中存放的各种商品每天的数量。一般来说，企业会首先根据数据的历史情况，分析出需求量序列的数理特征，对相似的需求量序列进行归类，并根据分类结果做到更加精准的预测。预测准确率有很多评价指标，较常用的指标为  $1-wmape$ ，定义如下：

$$1 - wmape = 1 - \frac{\sum |y_i - \hat{y}_i|}{\sum y_i}$$

其中  $y_i$  为第  $i$  个序列（商家在各仓库中存放的各种商品每天的数量）的真实需求量， $\hat{y}_i$  为第  $i$  个序列的预测需求量。

然而，在实际的电商供应链预测任务中，常常会出现多种不同的问题。例如，部分商品的销售时间过短、仓库存在新增或切换等情况，导致该预测维度下历史数据过少；另外，部分大型促销期间货量的陡增并由此带来的不规律性，也给需求量的精准预测带来了不小的难度。此时便需要通过算法得到历史一般规律，找出相似的历史情况（如相似的仓或商品），从而实现精准预测。

### 库存优化

为了有效管理库存水平，企业通常会制定各种库存策略来控制商品的补货频次和补货量。定期盘点库存策略(s, S)是一种常见的库存策略，其中

$s$  表示该种策略下的库存下限， $S$  为库存上限。在这种策略中，企业会每隔一段固定时间查看库存水平，盘点之间的间隔设为盘点周期（NRT）。如果在盘点时库存水平  $T$  低于最小库存水平  $s$ ，则会将库存补充至  $S$ ，此时的补货量  $Q = S - T$ ，如果在盘点时库存水平高于  $s$ ，则不进行补充，直到下一次库存盘点。同时需要考虑到，在每次发起补货后，补充的货物会经过一定时间后才能到达当前仓库，这个时间间隔为提前期（LT）。

使用以上的库存策略，管理者需要对两个方面进行权衡取舍。一方面，为了及时满足用户订单，公司需要保有大量库存以防止缺货；另一方面，库存水位升高带来了过高的库存成本。库存优化的目标主要为：

1) 降低库存总持有成本，可由单位持有成本  $h$  得到（一件商品存储一天产生的成本）；

2) 满足服务水平：确保有足够的产品可以及时满足客户需求，降低总缺货成本，可由缺货成本得到（每天每缺一件商品产生的成本）；

3) 降低库存周转天数：库存周转天数越小，表示库存的流动性越好，企业的库存管理效率越高。

服务水平及库存周转天数计算公式为：

$$\text{服务水平} = \frac{\text{被满足的商品件数}}{\text{实际需求总件数}} \times 100\%$$

$$\text{库存周转天数} = \frac{(\text{期初库存数量} + \text{期末库存数量})}{2} \times \frac{\text{计算时间段内天数}}{\text{计算时间段内总销量}}$$

现有一张电商零售商家的历史出货量表（附件 1），给出了历史 6 个月各商家存放在电商不同仓库的商品每天的出货量。假设该出货量即为历史各商品在各仓库的需求量。同时，还可以取到各商品、商家、仓库的信息（附件 2-4），例如分类、品牌、生效日期等，这些信息的选择和引入会帮助更好的预测并管理供应链中的库存。

**初赛问题：**

**问题一：**使用附件 1-4 中的数据，预测出各商家在各仓库的商品 2023-05-16 至 2023-05-30 的需求量，请将预测结果填写在结果表 1 并上传

至竞赛平台，并对你们模型的预测性能进行评价。另外请讨论：根据数据分析及建模过程，这些由商家、仓库、商品形成的时间序列如何分类，使同一类别在需求上的特征最为相似？

**问题二：**现有一些新出现的商家+仓库+商品维度（附件 5），导致这种情况出现的原因可能是新上市的商品，或是改变了某些商品所存放的仓库。请讨论这些新出现的预测维度如何通过历史附件 1 中的数据进行参考，找到相似序列并完成这些维度在 2023-05-16 至 2023-05-30 的预测值。请把预测结果填写在结果表 2，并上传至竞赛平台。

**问题三：**每年 6 月会出现规律性的大型促销，为需求量的精准预测以及履约带来了很大的挑战。附件 6 给出了附件 1 对应的商家+仓库+商品维度在去年双十一期间的需求量数据，请参考这些数据，给出 2023-06-01 至 2023-06-20 的预测值。请把预测结果填写在结果表 3，并上传至竞赛平台。

注：结果表 1-3 打包一个压缩包，“结果表.zip”，提交 zip 文件到竞赛系统中。

#### 附件 1-4 的字段说明：

附件 1：商家历史出货量表

字段名	字段类型	描述
seller_no	String	商家编码
product_no	String	商品编码
warehouse_no	String	仓库编码
date	String	日期
qty	Float	出货量（可看做需求量）

附件 2：商品信息表

字段名	字段类型	描述
product_no	String	商品编码
category1	String	商品一级分类

category2	String	商品二级分类
category3	String	商品三级分类

附件 3：商家信息表

字段名	字段类型	描述
seller_no	String	商家编码
seller_category	String	商家分类
inventory_category	String	库存分类
seller_level	String	商家规模

附件 4：仓库信息表

字段名	字段类型	描述
warehouse_no	String	仓库编码
warehouse_category	String	仓库类别
warehouse_region	String	仓库区域

各参赛队员可自行查找相关文献，以下参考文献仅供参考：

- [1] 智能供应链:预测算法理论与实战[M]，北京：电子工业出版社，2023.
- [2] Makridakis S , Spiliotis E , Assimakopoulos V .The M5 Accuracy competition: Results, findings and conclusions[J].International Journal of Forecasting, 2020, 36(1):224-227.
- [3] Makridakis S , Spiliotis E , Assimakopoulos V ,et al.The M4 Competition: 100,000 time series and 61 forecasting methods[J].International Journal of Forecasting, 2020, 36(1):54-74..