**家乐福销售需求计划项目设计文档**

|  |  |  |
| --- | --- | --- |
| **版本** | **日期** | **更新内容** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

目录

[1 需求 3](#_Toc19540441)

[2 现有环境 3](#_Toc19540442)

[3 概要设计 3](#_Toc19540443)

[3.1 系统架构 3](#_Toc19540444)

[3.2 系统模块 4](#_Toc19540445)

[3.3 模块结构 4](#_Toc19540446)

[3.3.1 订单生成 5](#_Toc19540447)

[3.3.2 系统监控 5](#_Toc19540448)

[4 详细设计 6](#_Toc19540449)

[4.1 数据库设计 6](#_Toc19540450)

[4.1.1 参数表 6](#_Toc19540451)

[4.1.2 数据输入表 6](#_Toc19540452)

[4.1.3 数据输出表 6](#_Toc19540453)

[4.1.4 KPI监控表 7](#_Toc19540454)

[4.1.5 模拟运行表 7](#_Toc19540455)

[4.2 DM订单模块设计 8](#_Toc19540456)

[4.3 常规订单模块设计 9](#_Toc19540457)

[5 附录 10](#_Toc19540458)

[5.1 门店订单文件模板 10](#_Toc19540459)

[5.2 大仓订单文件模板 10](#_Toc19540460)

[5.3 预测文件模板 10](#_Toc19540461)

# 需求

开发销售需求计划系统。结合供应商，门店，和大仓订单参数，把销售预测转化成大仓和门店的订单。此项目模型为概念验证型模型。

系统应在给定范围内进行稳定的自动化运行，无需手动输入数据或者干预。

系统应提供如下产出

* 按日生成的门店和大仓订单文件。
* 按周生成的需求计划预测文件
* 所需的可测量KPI的测量结果

# 现有环境

系统可以使用家乐福数据池获取和储存文件。数据池为Cloudera 大数据解决方案。可用工具包括Spark 运行环境，Impala运行环境。数据池中的数据以Hive表和Kudu表形式储存。数据池每日清晨从各个系统导入数据。

系统可以使用家乐福提供的服务器。服务器提供JupyterLab 开发环境。可以使用服务器进行运算。也可以通过Spark 运行环境和Impala运行环境=使用数据池集群，进行运算。

家乐福提供Airflow工作引擎用以实现自动化运行。

# 概要设计

## 系统架构

系统每次运行需要产生百万条以上的数据，逻辑较为复杂。系统需在每日8点30分前生成运算结果。数据池中数据最早在每日5点30分可以准备完毕。系统可用运行时间较少。鉴于以上两点和现有环境，系统基于Spark 批处理技术进行搭建。Spark由加州大学伯克利分校AMP实验室开发，是一个围绕速度、易用性和复杂分析构建的大数据处理框架。十分适合用来构建大型的、低延迟的数据应用程序。

系统包含以下几个组成

* 数据载入部分

Spark 运行环境对Kudu表的支持并不理想。家福乐数据池中的Kudu表多以全量备份形式更新，难以获取过去某一时间点的数据状态。数据载入部分使用Impala运行环境获取数据池表当前快照，并写入需求计划自用表中，以便以后使用。

* 逻辑部分

此部分为系统核心。在数据载入后此部分运行订单逻辑，生成订单信息，生成需求计划。生成的信息会写入数据池表中。此部分有个分为两个子类，DM订单逻辑和日常订单逻辑。

DM订单逻辑较为简单，逻辑分支较少，使用Pyspark技术基于Spark运行环境开发。日常订单逻辑复杂，分支众多，使用Scala技术基于Spark运行环境开发。

* 输出部分

此部分读取订单逻辑部分输出，将其按照规定的格式生成订单文件和需求计划文件，并放置于指定位置。此部分基于Python技术开发。使用openpyxl包以生成规定格式文件

* 监控部分

此部分用以计算所需的可测量KPI。此部分使用Jupyter Notebook技术为基础进行开发。使用运行环境获取所需数据。

* 运行部分

此部分用以实现系统的自动化运行。使用Airflow技术进行开发。

## 系统模块

系统包含如下模块

|  |  |  |
| --- | --- | --- |
| 模块归类 | 模块名 | 用途 |
| 运行部分 | show\_dag\_args | 记录当次运行输入参数，如运行日期 |
| 输入载入部分 | load\_data | 获取数据池表的快照 |
| 逻辑部分 | calculate\_service\_level | 计算服务水平 |
| 逻辑部分 | run\_dm\_order | 计算DM订单订量 |
| 逻辑部分 | run\_onstock\_store\_order | 计算存库单品门店订单订量 |
| 逻辑部分 | run\_xdock\_order | 计算越库单品门店订单订量 |
| 逻辑部分 | run\_dc\_order | 计算存库单品大仓订单订量 |
| 输出部分 | generate\_store\_order\_file | 生成门店订单文件 |
| 输出部分 | generate\_dc\_order\_file | 生成大仓订单文件 |
| 监控部分 | check\_store\_order | 检查门店订单生成情况 |
| 监控部分 | check\_dc\_order | 检查大仓订单生成情况 |
| 输出部分 | generate\_forecast\_file | 生成需求计划文件 |
| 监控部分 | run\_monitoring | 以计算所需的可测量KPI |
| 输出部分 | copy\_output | 将KPI测量结果放至指定位置 |

## 模块结构

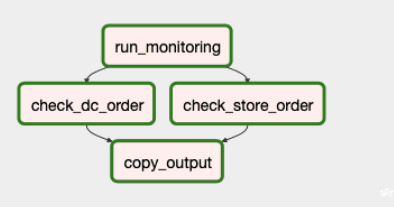
系统模块依据功能分为订单生成和系统监控两部分。两者无依靠关系，独立运行。两者无依靠关系，独立运行。

### 订单生成

图片包含 屏幕截图

描述已自动生成

### 系统监控



# 详细设计

## 数据库设计

### 参数表

下列表为需求计划系统参数表。储存程序运行基本参数。表内容更新不频繁。

|  |  |
| --- | --- |
| vartefact.forecast\_store\_item\_details | 门店单品信息 |
| vartefact.forecast\_onstock\_order\_delivery\_mapping | 存库单品送货信息 |
| vartefact.forecast\_xdock\_order\_mapping | 越库单品送货信息 |
| vartefact.forecast\_dc\_item\_details | 大仓单品信息 |
| vartefact.forecast\_calendar | 系统日 |
| vartefact.forecast\_dc\_order\_delivery\_mapping | 大仓订货日 |
| vartefact.forecast\_stores\_delv\_time | 门店送货到店时间 |
| vartefact.v\_forecast\_inscope\_store\_item\_details | 门店可订货单品信息 |
| vartefact.v\_forecast\_inscope\_dc\_item\_details | 大仓可订货单品信息 |
| vartefact.v\_forecast\_daily\_onstock\_order\_items | 当日应在门店下订单的存库单品信息 |
| vartefact.v\_forecast\_daily\_xdock\_order\_items | 当日应在门店下订单的越库单品信息 |

### 数据输入表

下列表为需求计划系统数据输入表。表中信息按日更新。

|  |  |
| --- | --- |
| vartefact.forecast\_nsa\_dm\_extract\_log | nsa.dm\_extract\_log 表的快照 |
| vartefact.forecast\_item\_code\_id\_stock | 门店每日库存 |
| vartefact.forecast\_p4cm\_store\_item | p4cm.store\_item表的每日快照 |
| vartefact.forecast\_lfms\_daily\_dcstock | 大仓每日库存 |
| vartefact.forecast\_dc\_latest\_sales | 大仓最新平均销量 |
| vartefact.t\_forecast\_daily\_sales\_prediction | 销售预测 |

### 数据输出表

下列表为需求计划数据输出表，用以接收每次运行的产出。

|  |  |
| --- | --- |
| vartefact.forecast\_onstock\_orders | 存库单品门店订单 |
| vartefact.forecast\_xdock\_orders | 越库单品门店订单 |
| vartefact.forecast\_dc\_orders | 存库单品大仓订单 |
| vartefact.forecast\_onstock\_orders\_hist | 存库单品门店订单历史输出记录 |
| vartefact.forecast\_xdock\_orders\_hist | 越库单品门店订单历史输出记录 |
| vartefact.forecast\_dc\_orders\_hist | 存库单品大仓订单历史输出记录 |
| vartefact.forecast\_dm\_orders | DM单品门店订单 |
| vartefact.forecast\_dm\_dc\_orders | DM单品大仓订单 |
| vartefact.forecast\_store\_daily\_order\_files | 门店订单文件存档 |
| vartefact.forecast\_dc\_daily\_order\_files | 大仓订单文件存档 |

### KPI监控表

下列表储存所需的可测量KPI的测量结果。表中信息按日更新。

|  |  |
| --- | --- |
| vartefact.foreacst\_store\_monitor | 门店库存状况监控 |
| vartefact.foreacst\_dc\_monitor | 大仓库存状况监控 |
| vartefact.forecast\_monitor\_store\_order\_diff | 门店订单执行情况监控 |
| vartefact.forecast\_monitor\_dc\_order\_diff | 大仓订单执行情况监控 |

### 模拟运行表

下列表为模拟运行所需表。日常运行时这些表不会被使用。

|  |  |
| --- | --- |
| vartefact.forecast\_simulation\_dm\_orders | 模拟运行产生的DM单品门店订单 |
| vartefact.forecast\_simulation\_dm\_dc\_orders | 模拟运行产生的DM单品大仓订单 |
| vartefact.forecast\_simulation\_orders\_hist | 模拟运行产生的订单历史输出记录 |
| vartefact.forecast\_simulation\_result | 模拟运行历史输出记录 |
| vartefact.v\_forecast\_simulation\_lastest\_result | 模拟运行输出 |
| vartefact.v\_forecast\_simulation\_stock | 模拟运行产生的每日预计库存 |
| vartefact.v\_forecast\_simulation\_orders | 模拟运行产生的订单 |

## DM订单模块设计

开始

载入运行参数

读取DM相关信息

剔除已处理过的DM

保留距运行日最近的DM

计算DM门店订单日

读取DM销售预测

读取DM前后日常销售预测

计算DM门店订单

储存DM门店订单

计算DM大仓订单

储存DM大仓订单

计算DM大仓订单日

结束

## 常规订单模块设计

开始

载入运行参数

读取当日为订单日的的单品

读取单品信息和未来订单日

剔除不可订货单品

读取库存

读取已生成的DM订单信息

读取在途订单信息

读取过去生成的订单计划

是否为大仓订单

读取未来门店向大仓订货计划

读取销售预测

否

是

计算订单量

储存订单

结束

# 附录

## 门店订单文件模板



## 大仓订单文件模板



## 预测文件模板

