**家乐福销售需求计划项目部署文档**

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

# 系统环境要求

本项目的运行需要如下系统环境

|  |  |
| --- | --- |
| 项目 | 版本号 |
| Airflow | 1.10.3 或更高版本 |
| Jupyter Lab | 0.35.4或更高版本 |
| Spark | 2.4.0 |
| openpyxl | 2.6.3或更高版本 |
|  |  |

# 安装Airflow

参考下列airflow官方文档进行airflow的安装

<http://airflow.apache.org/installation.html>

# 初始化文件夹

建立下列文件夹。确保Airflow对下列文件夹拥有读写权限

|  |  |
| --- | --- |
| 文件夹路径 | 说明 |
| /NFSdata/output/datachecking | 运行监测结果输出文件夹 |
| /NFSdata/output/weekly\_forecast | 销售预测输出文件夹 |
| /NFSdata/output/monitoring | 运行监控结果输出文件夹 |
| /NFSdata/output/daily\_order | 每日订单输出文件夹 |
| /data/jupyter/Carrefour-China-Supply-Chain-Forecast/workflow\_integration | 需求计划程序根文件夹 |
| /data/jupyter/Carrefour-China-Supply-Chain-Forecast/output | 程序输出存档文件夹 |
| /data/jupyter/Carrefour-China-Supply-Chain-Forecast/logs | 日志文件夹 |

完成创建后将需求计划程序文件复制到需求计划程序根文件夹

## 需求计划程序根文件夹结构

需求计划程序根文件夹包含下列文件

|  |  |
| --- | --- |
| 文件名 | 说明 |
| forecast\_orderflow.py | 自动化运行需求计划项目的脚本 |
| forecast\_simulation.py | 自动化需求计划项目模拟订单运行的脚本 |

需求计划程序根文件夹下有三个子文件夹

* jar 文件夹，包含系统运行所需scala jar包。文件夹包含下列文件。

|  |  |
| --- | --- |
| 文件名 | 说明 |
| carrefour\_forecast\_2.11-0.1.jar | 生成常规订单订货量的软件 |

* py\_script文件夹，包含系统运行所需python程序。文件夹包含下列文件。

|  |  |
| --- | --- |
| 文件名 | 说明 |
| forecast\_dm\_order.py | 生成DM订单订货量的软件 |
| forecast\_dm\_simulation.py | 模拟DM订单订货量的软件 |
| forecast\_data\_loading.py | 运行抓取数据语句的软件 |
| forecast\_order\_file.py | 生成Excel格式每日订单文件的软件 |
| forecast\_check\_missing\_orders.py | 检查每日订单生成情况的软件 |
| generate\_supplier\_forecast.py | 生成分享给供应商的预测文件的软件 |

* sql文件夹，包含系统运行所需SQL语句。文件夹包含下列文件。

|  |  |
| --- | --- |
| 文件名 | 说明 |
| forecast\_p4cm\_store\_item.sql | 从数据池抓取所需数据的查询语句 |
| forecast\_nsa\_dm\_extract\_log.sql | 从数据池抓取所需数据的查询语句 |
| forecast\_lfms\_daily\_dcstock.sql | 从数据池抓取所需数据的查询语句 |
| forecast\_item\_code\_id\_stock.sql | 从数据池抓取所需数据的查询语句 |
| forecast\_dc\_latest\_sales.sql | 从数据池抓取所需数据的查询语句 |
| forecast\_service\_level\_by\_item.sql | 计算单品服务率的语句 |
| forecast\_service\_level.sql | 计算服务率统计的语句 |

# 初始化Airflow环境

## 修改Airflow 配置

在Airflow安装目录找到airflow.cfg 文件，编辑文件内容。将dags\_folder 指向需求计划程序根文件夹。 将base\_log\_folder 指向日志文件夹。保存修改内容。

完成修改后，参考下列官方文档启动Airflow

<http://airflow.apache.org/start.html>

## 导入Airflow参数

打开浏览器，进入Airflow UI管控界面。点击Admin。在显示的下拉列表中选择Variable。

A screenshot of a cell phone

Description automatically generated

在Variables界面中点击Choose File。 选中Airflow 参数配置文件airflow\_variables.json。 点击Import Variables完成参数配置导入

A screenshot of a cell phone

Description automatically generated

# 初始化订单参数

将下列文件放置于同一文件夹

|  |  |
| --- | --- |
| 文件名 | 说明 |
| load\_store\_parameters.ipynb | 载入或更新门店订单参数的软件 |
| load\_dc\_parameters.ipynb | 载入或更新大仓订单参数的软件 |
| prepare\_calendar\_table.ipynb | 载入或更新订单日历的软件 |
| East Parameter 20190906.xlsx | 最新门店订单参数 |
| East 3 Supps DC Item list 20190830.xlsx | 最新大仓订单参数 |
| store\_delivery\_time.xlsx | 门店订单到货时间 |

在Jupyter Lab中打开上面的 ipynb 文件。点击Run，在下拉列表中选择Run All Cells 以执行文件中的全部语句。

请确保如下表在数据库中可见

|  |  |
| --- | --- |
| 表名 | 说明 |
| 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 |

# 初始化数据库表

打开脚本文件create table.sql，将其中的SQL语句拷贝至Impala运行环境。运行语句完成表的创建。

请确保如下表在数据库中可见

|  |  |
| --- | --- |
| 表名 | 说明 |
| vartefact.forecast\_script\_runs | 系统运行记录 |
| 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.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.t\_forecast\_daily\_sales\_prediction | 销售预测 |
| vartefact.forecast\_store\_daily\_order\_files | 门店订单文件存档 |
| vartefact.forecast\_dc\_daily\_order\_files | 大仓定单文件存档 |
| 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 |
| 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.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 |
| 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 |
| vartefact.v\_forecast\_weekly\_dc\_order\_forecast | 存库单品的供应商订货量预测 |
| vartefact.v\_forecast\_weekly\_xdock\_order\_forecast | 越库单品的供应商订货量预测 |

# 首次运行

系统自动运行需要有至少一次成功运行。如首次运行为2019年9月9日，则需有一次2019年9月8日的成功运行。

使用如下指令手动启动2019年9月8日的运行.

airflow backfill forecast\_orderflow -s 20190906 -e 2090907

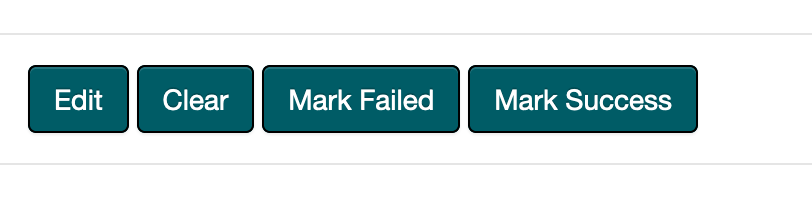
-s 日期为运行日期贱两日

命令执行后进入Airflow UI管理界面，进入相应DAG。 在Tree View界面中应可见如下图示，代表正在执行的运行。

A drawing of a face

Description automatically generated

点击绿色圆形图标，在弹出窗口中选择Mark Success



在新窗口中选择 OK！，确认状态更高

图示应转化为如下式样，代表此次运行以标记为成功运行。

A drawing of a person

Description automatically generated

# 开启自动化执行

可以通过两种方式开启系统自动化执行。

## 通过Airflow UI管理界面

Airflow UI管理界面，选择的DAGs

A screenshot of a social media post

Description automatically generated

DAG 名称左侧图式代表自动化执行是否开始。 On 代表自动化执行启动。 Off 代表自动化执行停止。单击图标可进行状态切换。

A screenshot of a cell phone

Description automatically generated

## 通过命令行

在Linux运行环境中执行下列命令

|  |  |  |
| --- | --- | --- |
| 命令· | 说明 | 示例 |
| airflow pause {DAG名称} | 开启自动化执行 | airflow pause forecast\_orderflow |
| airflow unpause {DAG名称} | 关闭自动化执行 | airflow unpause forecast\_orderflow |