

Q1 分析报告

(一) 问题描述

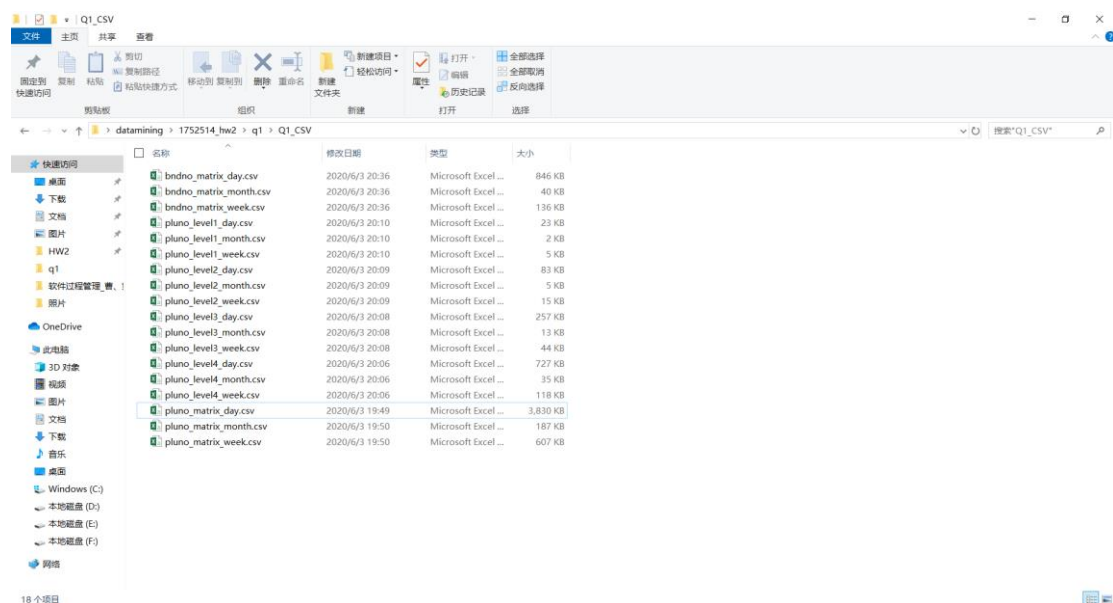
数据预处理：首先要求对每个商品编号/商品类别结构/品牌编号等按照购买时间对购买数量进行汇总求和，包括按天、周、月为周期分别形成该商品编号/商品类别结构/品牌编号对应购买数量 汇总的每天、每周、每月 3 个时间序列数据，该时间序列数据按照商品编号分组对应的时间信息进行排序。

(二) 解决思路

我的预处理生成的结果是一系列 csv 文件，分别是针对不同的区别商品的字段（根据商品编号、商品四级类别结构、品牌编号）和销量汇总的时间单位（天、周、月），每个 csv 文件中包含每一个不同商品从最早的日期（2016-02-01）到最晚的日期（2016-07-31），每一个时间单位的销量记录，没有购买记录的日期补 0。最后呈现的就是一个以时间跨度为横轴、区别商品字段为纵轴的矩阵。

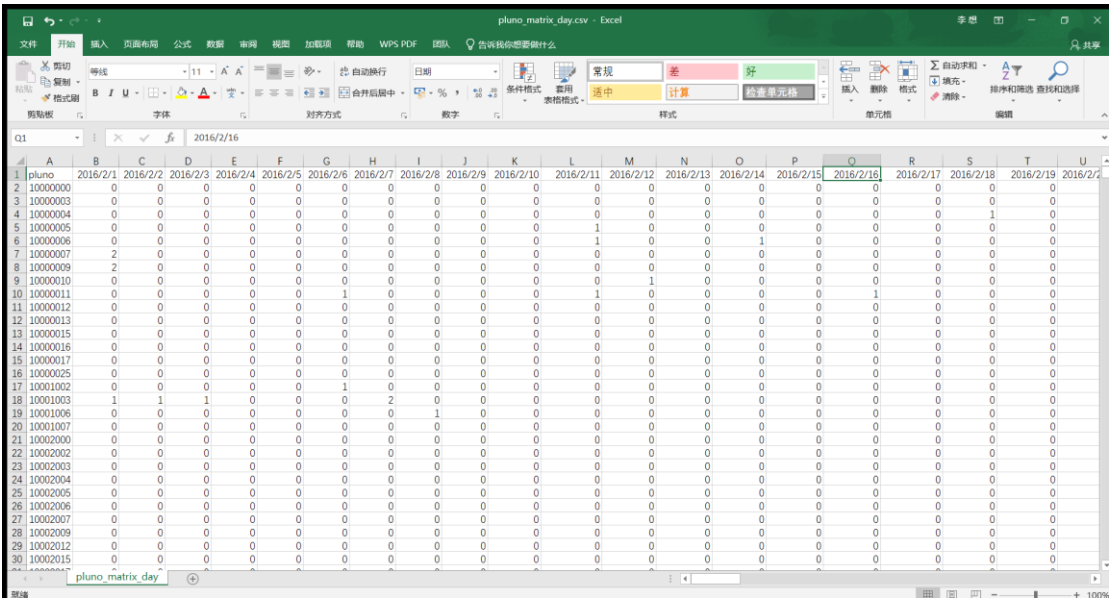
(三) 预处理结果

下面是所有的 csv 文件的截图，所有 csv 文件都会附在文件夹中：

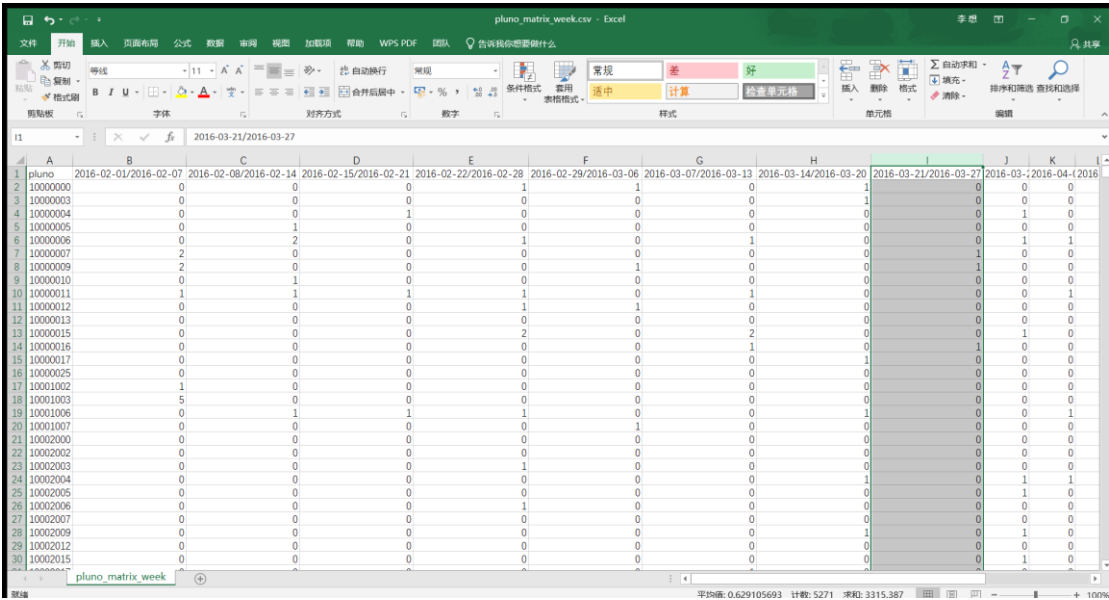


下面是截取了根据商品编号（Pluno）区分不同商品，同时分别以天、周、月为时间单位的部分 csv 数据：

Pluno/Day:



Pluno/Week:



Pluno/Month:

