

目录

- 一、背景介绍..... 2
  - 1、行业背景 .....2
  - 2、分析目标 .....2
- 二、数据预处理和统计 ..... 3
  - 1、数据预处理.....3
  - 2、数据统计 .....4
- 三、数据可视化..... 5
  - 1、生鲜类商品和一般类商品每日销售分析 .....5
  - 2、各大类商品销售占比分析.....6
  - 3、促销商品和促销商品周环比增长分析 .....10
- 四、用户画像与促销策略 ..... 12
  - 1、累计消费前十的用户的画像 .....12
  - 2、大类商品销售情况分析 .....16
  - 3、促销对商品销售影响分析..... 18

# 一、背景介绍

## 1、行业背景

随着当下科学技术的不断发展，新零售业态蓬勃发展，而传统的线下零售、电商零售等业态渐显势微。因此，借鉴新零售行业的销售特点，对超市的经营数据进行分析，可以提高传统超市的竞争力，为调整、提高传统超市的运营模式、战略提出建议。

## 2、分析目标

本次分析利用某超市 2015 年 1 月 1 日至 4 月 30 日的经营数据，对其进行以下三步的分析：

- 1) 对销售数据统计分析，并进行可视化展示。
- 2) 分析消费者的消费行为。
- 3) 研究促销对于消费行为的影响。

## 二、数据预处理和统计

### 1、数据预处理

在进行正式的统计处理之前，我们先依次对数据进行去重、去除缺失值、去除异常值。其中异常值包括销售数量、商品单价为负值或 0，销售金额为负值，或者销售日期为“20150229”的数据。经过数据预处理以后，我们得到正常的数据共计 **42721** 条。具体数据保存至附件“task1\_1.csv”

原附件数据有 42816 行,17 列,使用 Python 进行去除重复值,剩下 42813 行数据。

查看数据时还发现有些行数据内为空值,这可能使得统计分析时出错,使用 python 的 dropna 函数进行去除有缺失值的行。剩下 42723 行数据。

在将数值型数据转成时间类型时发现报错,查看数据发现有不合理的日期,即“2015 年 2 月 29 日”这种不合理的数据,使用程序将之删除。剩下 42721 行数据,保存到“task1\_1.csv”文件中。

## 2、数据统计

使用 groupby 函数对数据进行分组聚合，得到各大类商品和的销售总额数据。结果保存在 “task1\_2.csv” 中。groupby 对数据分组聚合，得到中类商品的促销销售金额以及非促销销售金额数据，保存至附件 “task1\_3.csv” 中。新增加一列标注对应日期为第几周，使用 python 分组聚合得到生鲜类产品和一般产品的每周销售金额。保存至附件 “task1\_4.csv”。使用 python 分组聚合，得到每位顾客每月消费额以及消费天数统计已保存至附件 “task1\_5.csv” 中。

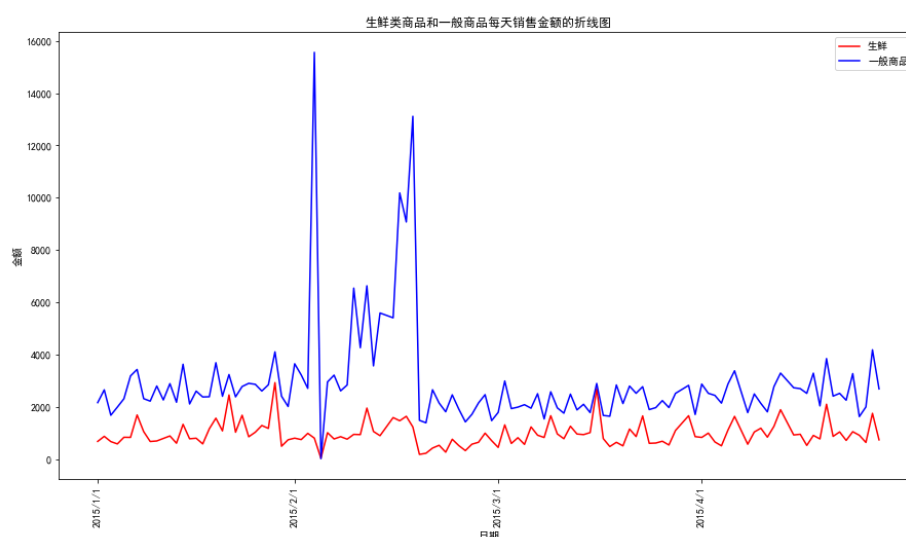
|   | 1月份消费总额 | 1月份消费天数 | 2月份消费总额 | 2月份消费天数 | 3月份消费总额 | 3月份消费天数 | 4月份消费总额 | 4月份消费天数 |
|---|---------|---------|---------|---------|---------|---------|---------|---------|
| 0 | 11.05   | 1.0     | 0.00    | 0.0     | 0.00    | 0.0     | 13.60   | 1.0     |
| 1 | 12.30   | 1.0     | 30.30   | 1.0     | 0.00    | 0.0     | 0.00    | 0.0     |
| 2 | 48.70   | 2.0     | 30.00   | 1.0     | 0.00    | 0.0     | 0.00    | 0.0     |
| 3 | 67.39   | 3.0     | 360.97  | 4.0     | 68.87   | 7.0     | 49.28   | 4.0     |
| 4 | 27.90   | 1.0     | 0.00    | 0.0     | 0.00    | 0.0     | 154.94  | 2.0     |
| 5 | 57.60   | 2.0     | 0.00    | 0.0     | 33.50   | 1.0     | 48.34   | 2.0     |
| 6 | 23.99   | 1.0     | 65.67   | 1.0     | 11.50   | 1.0     | 49.14   | 3.0     |
| 7 | 138.37  | 2.0     | 0.00    | 0.0     | 61.49   | 1.0     | 52.07   | 2.0     |
| 8 | 190.49  | 4.0     | 0.00    | 0.0     | 111.85  | 4.0     | 74.72   | 1.0     |
| 9 | 80.40   | 2.0     | 0.00    | 0.0     | 0.00    | 0.0     | 0.00    | 0.0     |

图表 1 0-10 号顾客 1-4 月消费天数和消费金额统计表

### 三、数据可视化

#### 1、生鲜类商品和一般类商品每日销售分析

使用 python 读取 “task1\_1.csv” 文件，分组聚合操作后得到生鲜类商品和一般商品的每天销售金额数据，使用 matplotlib 进行数据可视化，得到折线图如下。

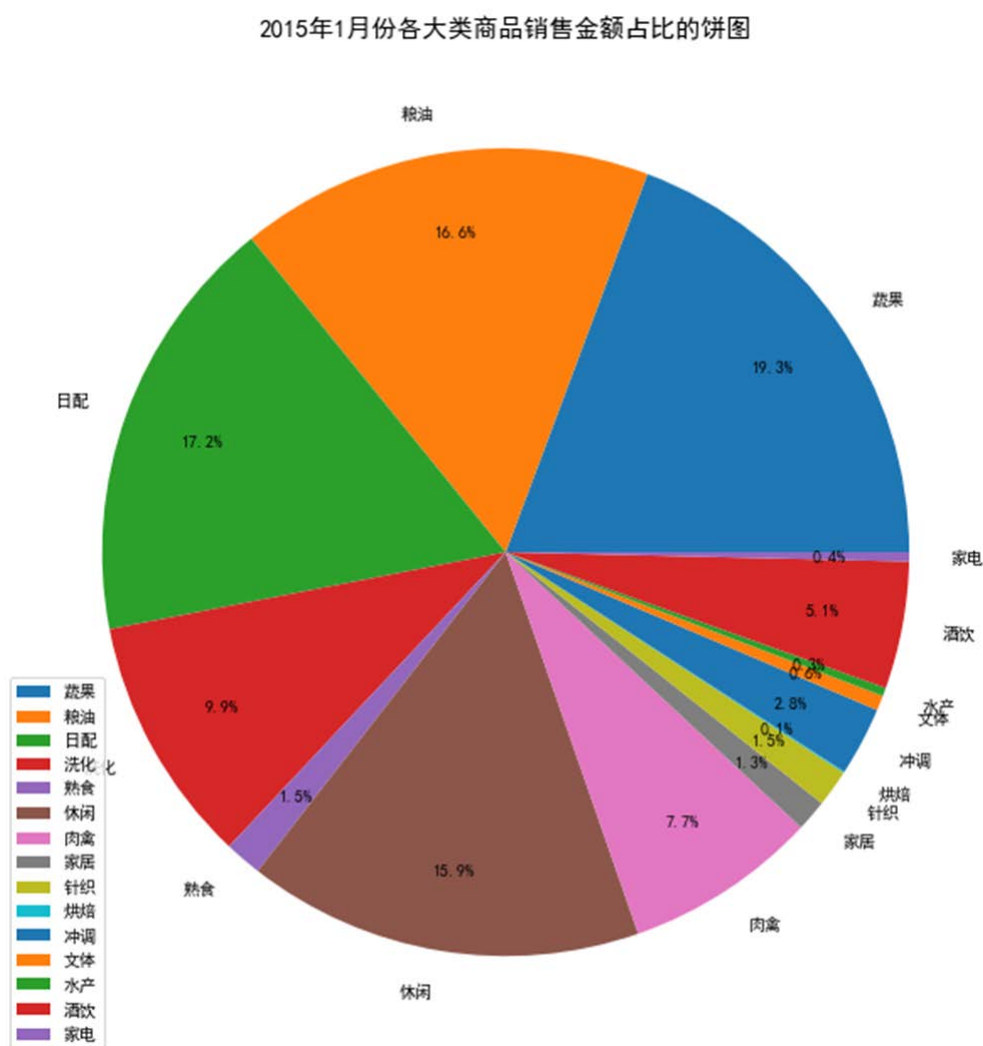


图表 2 生鲜类和一般类商品每日销售量折线图

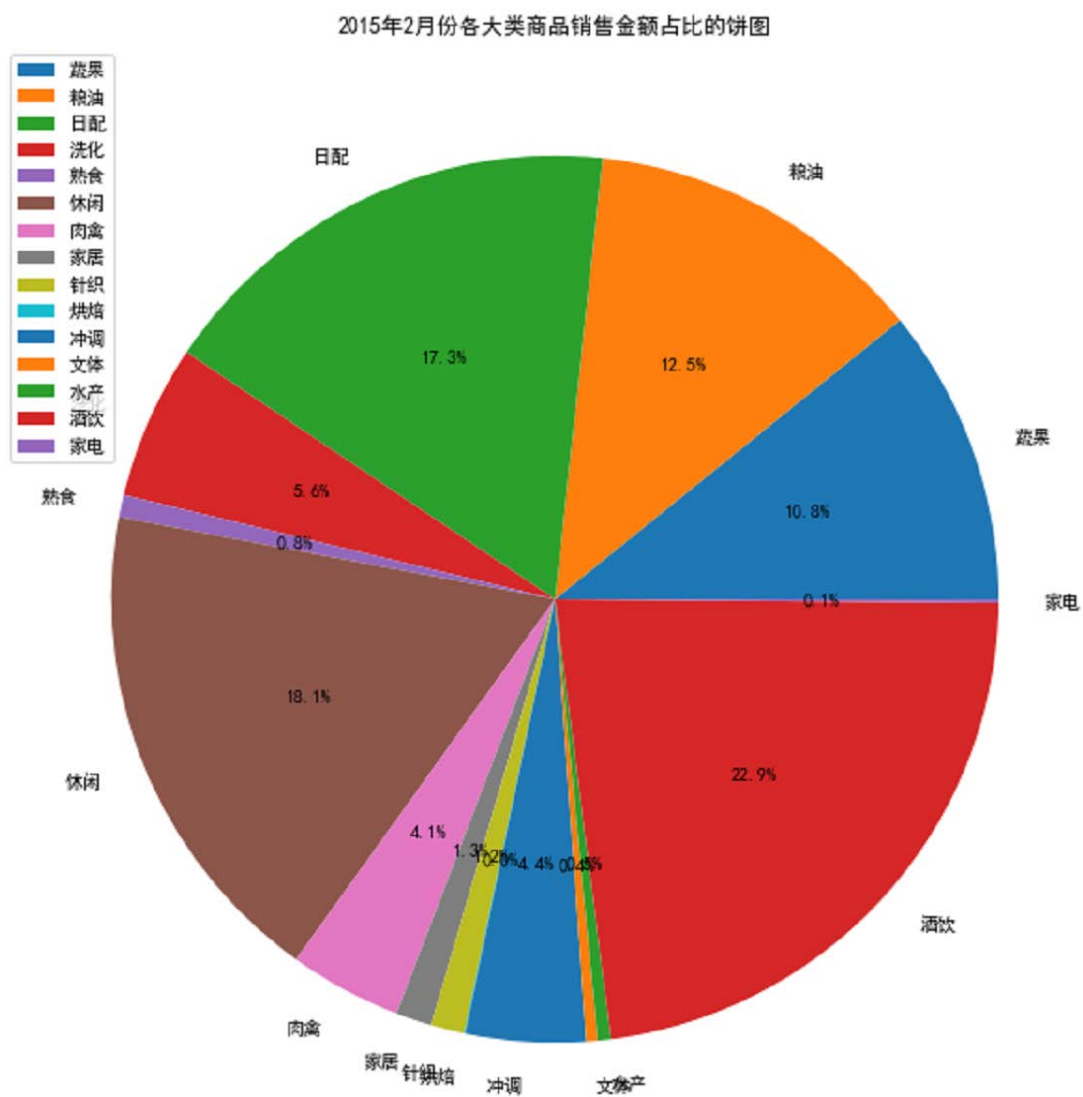
从折线图可以看出，春节期间一般商品销售量出现了强烈波动。而生鲜类商品仍然保持稳定。除假期影响外，总体上，两类商品的销售量保持稳定，一般商品的销售量普遍比生鲜类商品销售量更高。在 2015 年，假期的影响对一般商品比生鲜商品大。这与现实情况是不符的，在假期期间，一般商品的销售量也应该呈现较大的波动。我们初步猜测，由于生鲜类新零售行业的兴起，人们的选择发生了变化。针对这一异常，超市更应该调整假期中针对生鲜商品的销售策略。

## 2、各大类商品销售占比分析

分别计算 1 到 4 月的各大类的商品的销售金额，然后根据得到的每个大类的名称和对应的销售金额画得饼状图.

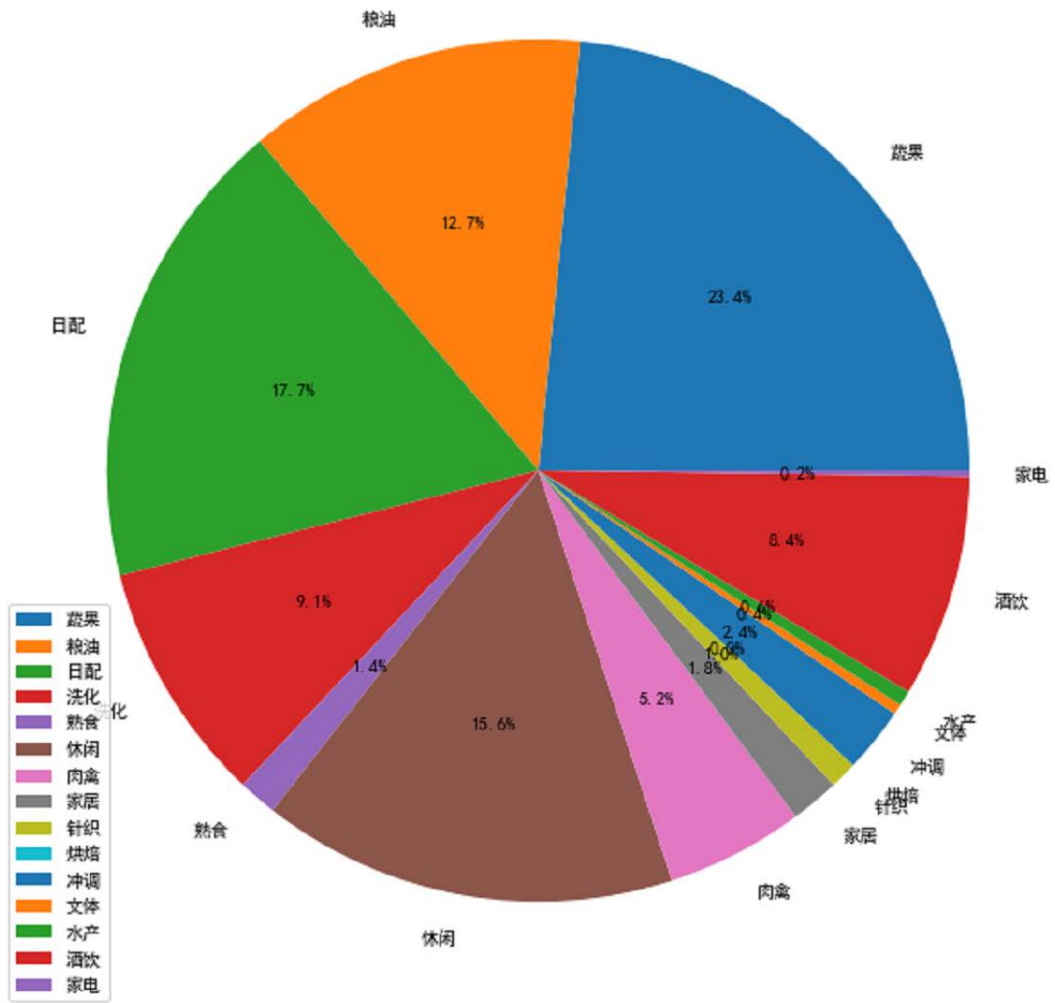


图表 3 2015 年 1 月各大类商品销售金额占比饼状图



图表 4 2015 年 2 月各大类商品销售金额占比饼状图

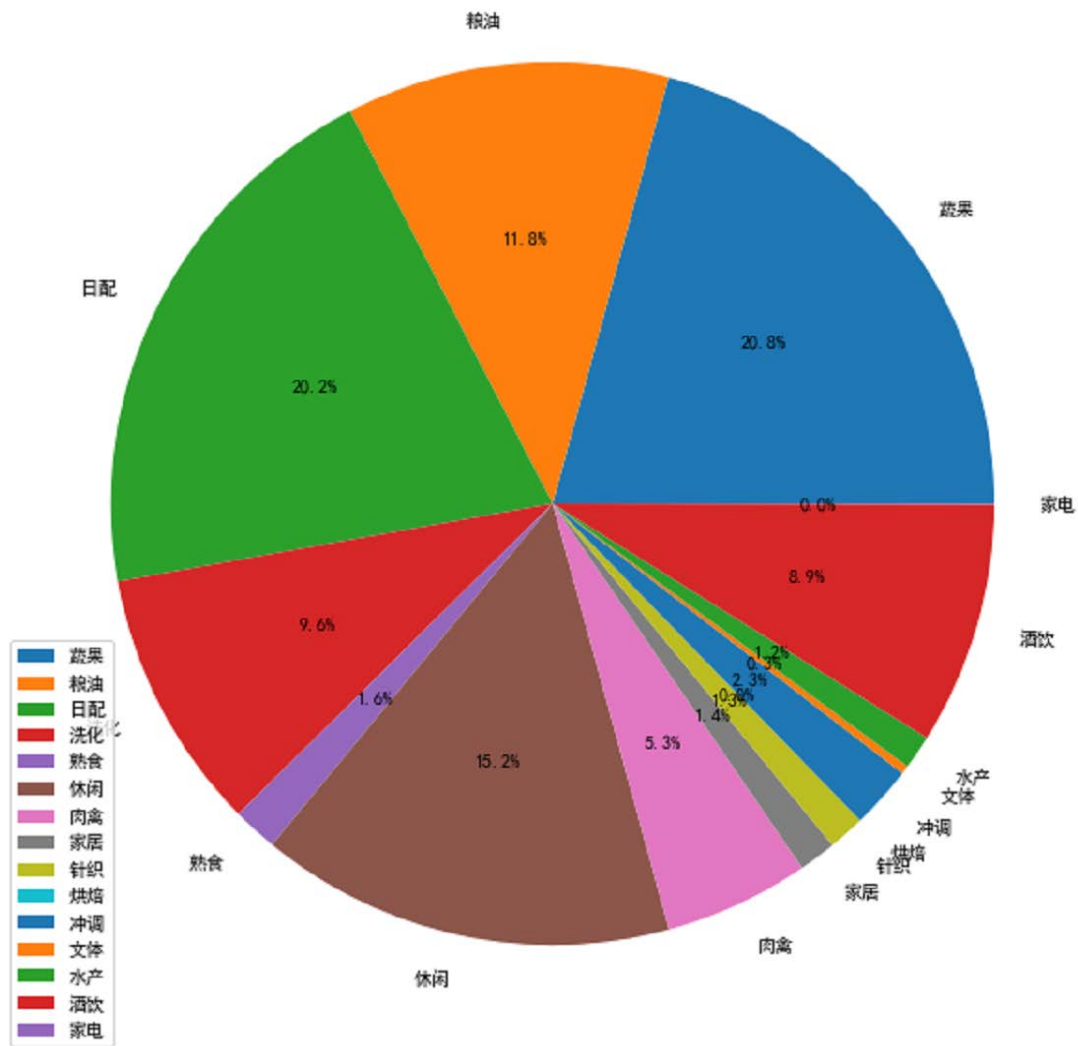
2015年3月份各大类商品销售金额占比的饼图



图表 5 2015 年 3 月各大类商品销售金额占比饼状图



2015年4月份各大类商品销售金额占比的饼图



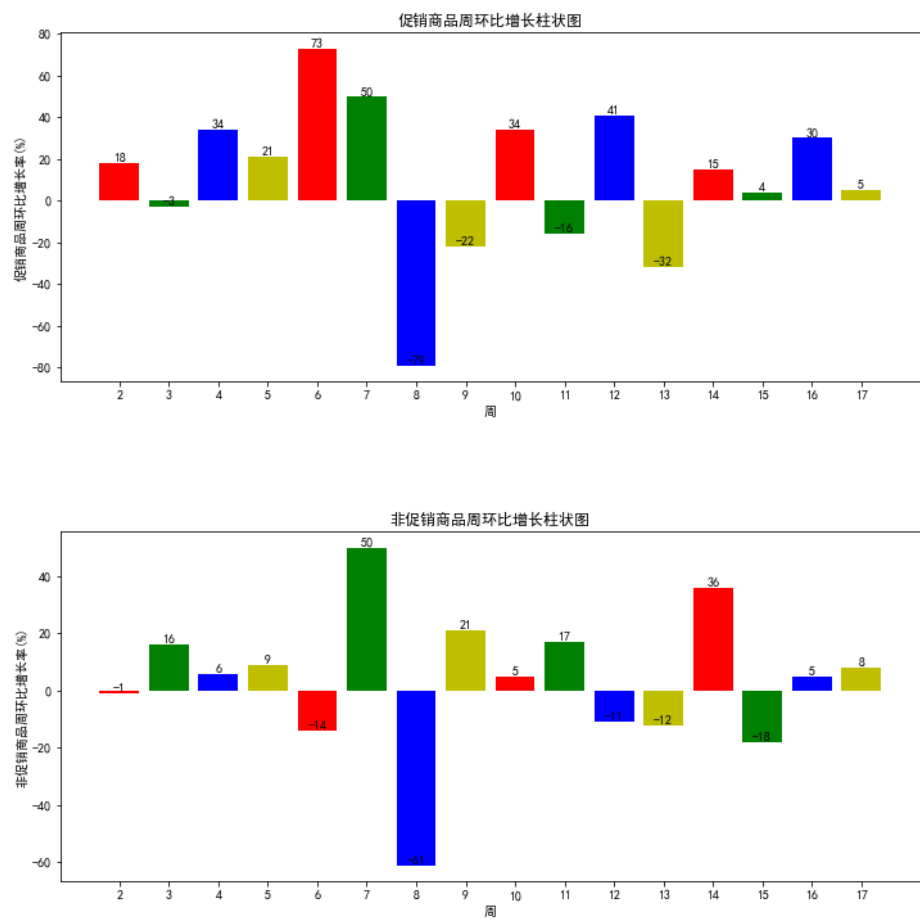
图表 6 2015 年 4 月各大类商品销售金额占比饼状图

总体上，日配，粮油，蔬菜，休闲各大类商品销售稳定占据每月销售总额 15%–20%，每月占较大比重。肉禽、酒饮、洗化次之，每月销售总额 5%–10%，其他大类商品基本占据 5% 以下，占比很小。这与现实中超市实际情况是较相符的。值得注意在较于一月份及三月份，二月份酒水占比增大。在此情况下，日配的占比稳定，其他商品占比均有所下降。我们进行初步猜测，这与 2015 年 2 月份正值春节期间，人们设宴请客，添置新物是有所相关的。在此期间，

超市需要及时预测客人对酒水、日配需求增大的情况，调整商品的补货计划。

### 3、促销商品和促销商品周环比增长分析

在 task1-1.csv 数据的基础上增加一列周并为每个商品标记具体为哪周然后命名为 task2-3.csv 文件，然后通过 loc 函数选出促销商品来画出周环比增长率图，非促销商品类似可得。



图表 4 促销商品和非促销商品周环比增长柱状图

从图中我们可以看出，促销商品在 6 月、7 月、8 月期间，非促销产品在 7、8 月间环比增长均超过 50%，销售金额波动明显。

而促销产品的周环比增长比非促销产品周同比增长更为波动。由于正值春节假期期间,我们猜测假期对于超市的销售金额影响是很明显的,且该超市的促销活动是具有明显的聚客、引流的作用的。所以该超市应该继续制定针对性的促销活动。

# 四、用户画像与促销策略

## 1、累计消费前十的用户的画像

首先算出累计消费前十的顾客，然后分别选出 task1-1 中每个顾客买的商品所属小类名称依次存在 txt 文件中。根据这个 t 文件画出顾客的画像。

经过处理，我们得出累计消费前十的用户分别为编号为 1177，52，986，1385，108，210，12，395，74，1594 的用户。其累计消费分别为 13597.06,3589.15,2611.24,2475.4,2349.22,1989.73,1838.19,1820.20,1678.37,1563.9。以下分别这十位用户的画像。



图表 5 消费额第一名 1177 用户画像



禽蛋 类乳 麦片 新鲜 豆奶粉  
酸 乳 麦片 花生油  
纯味 花生油  
普通 综合 冷藏 白酒





图表 10 消费额第六名 210 用户画像



图表 11 消费额第七名 12 用户画像



图表 12 消费额第八名 395 用户画像



图表 13 消费额第九名 74 用户画像



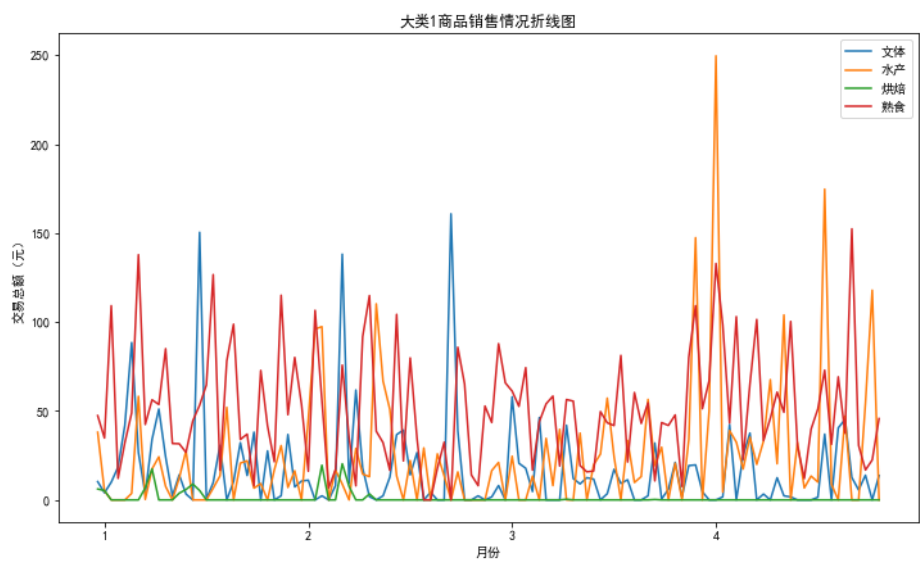
图表 14 消费额第十名 1594 用户画像

根据用户画像，我们了解大部分消费排在前十的消费者在超市会日常购置蔬果类、肉类、零食类、饮品类的商品。有两位消费者会进行少量、单价较高的物品采购。是符合商品消费中的二八规律的。我们建议超市针对消费者消费的预期性的特点，展开促销活动，首先应该让消费者提前预知促销活动的内容。可以通过结合线上和线下方式提前宣传促销活动。线上方式应该具有针对性，根据个人消费特征通过网络、收集 APP 信息推送，而线下宣传则可以根据主要的消费种类进行促销，通过张贴海报、宣传册等方式进行宣传。

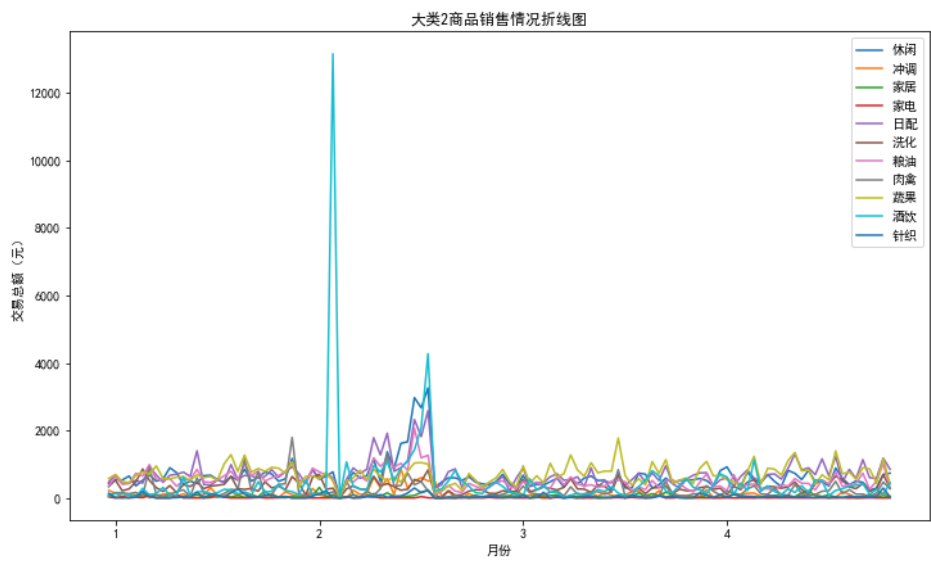
## 2、大类商品销售情况分析

计算出每个大类的商品每天的销售额，画出每个大类每天销售额的折线图，分析规律。

由大类的商品销售饼状图，我们发现文体、烘焙、水产和熟食类商品销售份额很小，故我们根据销售量分开画折线图。



图表 15 文体、水产、烘焙、熟食类商品销售折线图

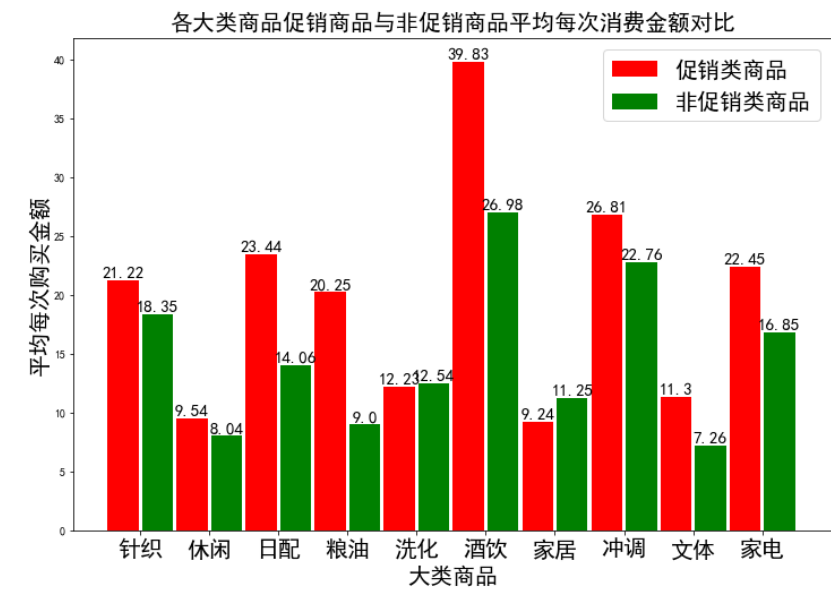




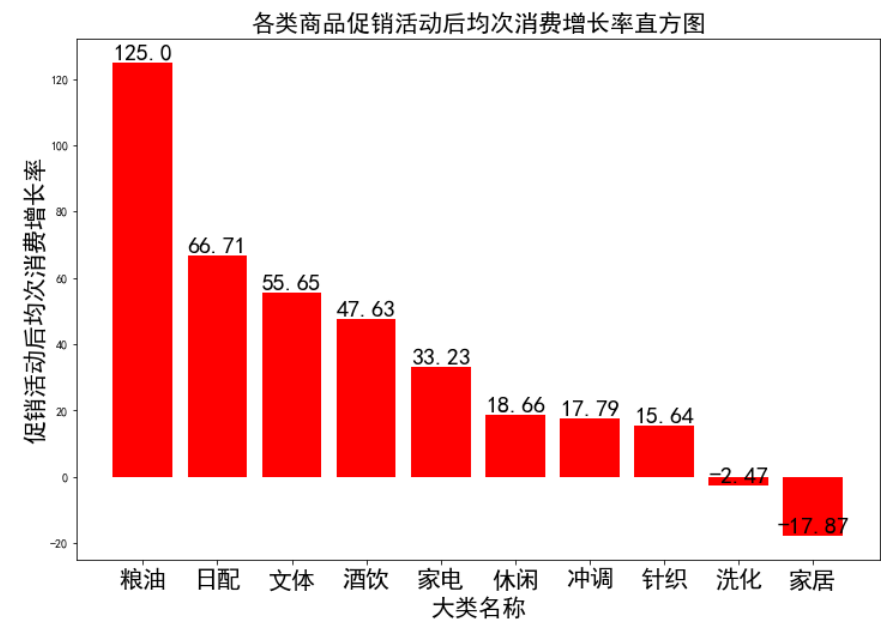
图表 16 酒水、日配、粮油共 11 种商品销售折线图

从十五种大类商品的折线图，我们可以看出在周末、假期，各类商品的销售量都会出现不同程度的变化。其中酒饮的销售额受假期影响最大。粮油、日配次之。图中可以看出在假期前期，酒水类商品的销售量增加可以作为日配、粮油销售量的风向标。在日常中，各类商品的销售量都较为稳定。各类商品之间的销售额也有较大区别。其中日配、蔬果、粮油、酒饮均为比较畅销的商品。而烘焙、文体、熟食、水产类商品销售少，销售慢。针对以上所述的销售规律我们建议超市对淡季、滞销类的商品，把握节假日所带来的人流量，举办对应的促销活动。并以酒水类商品的销售作为风向标，对酒水、蔬果、水产等畅销商品及时补货，避免缺货的情况而导致利润减少。提高超市促销时期的宣传力度。另外在节假日进行有意义，符合主题的活动。总而言之，以特定节假日为契机，从消费者的实际利益出发，结合超市内部大类商品的各自消费特征，加强线上线下的宣传力度，吸引消费者并且提高消费者的体验。

3、促销对商品销售影响分析



图表 17 促销商品与非促销商品销售额直方图



图表 18 各类商品促销后消费增长率直方图

由以上直方图可知，日配类商品、粮油类商品、酒饮类商品、文体类商品、家电类商品的销售额对促销活动较为敏感，当有促销活动时，这几大类的商品均次消费金额比非促销时有较大提升。在展开促销活动时候，由于其价格优势，将引导消费者产生消费行为。同时，促销活动对于家电类商品比其他三种商品的影响稍微小。这与实际情况是相符的。由于消费者对于粮油、日配、文体该类日常消耗的商品而言，消费者对其的品牌要求相对较低，故消费者不容易因为是不常用品牌而放弃价格的优势，而对于家电等耐用商品而言，其质量会对寿命等产生直接的影响，故消费者对于其价格优势，会对多个品牌综合比较再做出选择。

我们建议商家，首先在选择促销商品的种类方面，加大销售额对促销活动敏感的商品销售范围，或者首先优先对这几类商品进行促销，而洗化类商品的销售额对促销活动不敏感，故建议减少相关的促销活动，避免人力财力的浪费。此外，家居类商品对促销活动的敏感度为负，建议家居类商品不进行促销活动。结合上文的分析，第二我们建议商家选择合适的降价时机。由于节假日会带来销售高峰，商家需要把握好这一时机，最大限度提高促销产品在节假日期间的销售量。另外我们建议商家提高商场环境、服务等，提高消费者对于商场开展的促销品牌活动的探求欲望，从而提升降价促销策略的有效性，切实发挥促销活动对提高商品销售的作用。

