【项目08】 电商打折套路解析

作业要求

1. 从现有数据中，分析出“各个品牌都有多少商品参加了双十一活动？”

要求：

① 计算得到：商品总数、品牌总数

② 双十一当天在售的商品占比情况

（思考：是不是只有双十一当天在售的商品是“参与双十一活动的商品？”）

1. 未参与双十一当天活动的商品，在双十一之后的去向如何？
2. 真正参与双十一活动的品牌有哪些？

其各个品牌参与双十一活动的商品数量分布是怎样的？

\* 用bokeh绘制柱状图表示

提示：

① 数据的“id”字段为商品的实际唯一标识，“title”字段则为商品在网页上显示的名称

\* 仔细看数据可以发现，同一个id的title不一定一样（双十一前后）

② 数据的“店名”字段为品牌的唯一标识

③ 按照商品销售节奏分类，我们可以将商品分为7类

A. 11.11前后及当天都在售 → 一直在售

B. 11.11之后停止销售 → 双十一后停止销售

C. 11.11开始销售并当天不停止 → 双十一当天上架并持续在售

D. 11.11开始销售且当天停止 → 仅双十一当天有售

E. 11.5 - 11.10 → 双十一前停止销售

F. 仅11.11当天停止销售 → 仅双十一当天停止销售

G. 11.12开始销售 → 双十一后上架

④ 未参与双十一当天活动的商品，可能有四种情况：

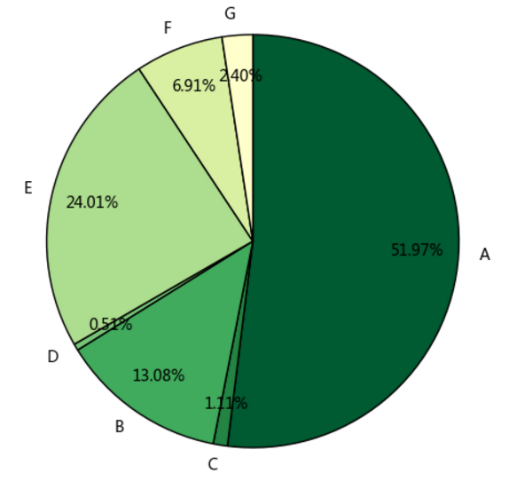
con1 → 暂时下架（F）

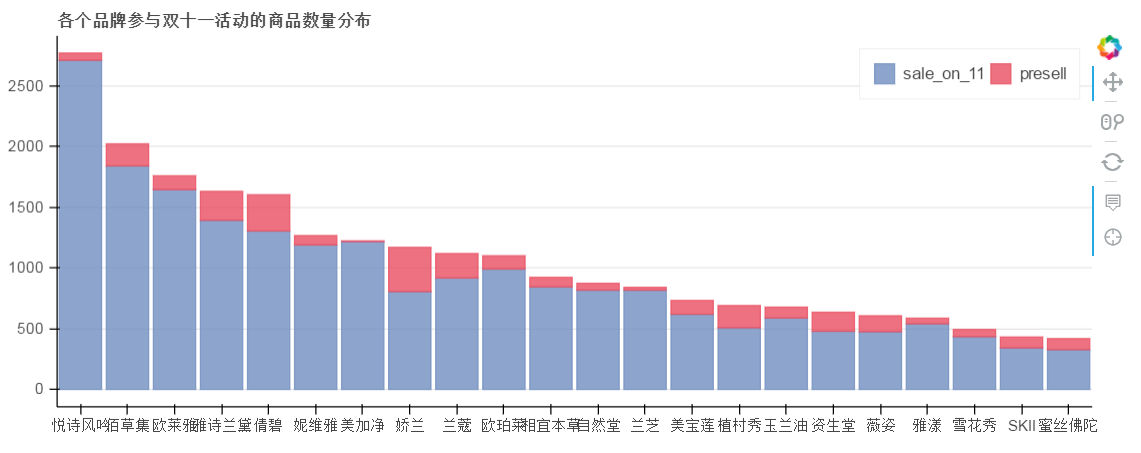
con2 → 重新上架（E中部分数据，数据中同一个id可能有不同title，“换个马甲重新上架”），字符串查找特定字符 dataframe.str.contains('预售')

con3 → 预售（E中部分数据，预售商品的title中包含“预售”二字）

con4 → 彻底下架（E中部分数据），可忽略

⑤ 真正参加活动的商品 = 双十一当天在售的商品 + 预售商品 （可以尝试结果去重）

 通过上述几个指标计算，研究出哪些是真正参与双十一活动的品牌，且其商品数量是多



2、哪些商品真的在打折呢？

要求：

① 针对每个商品，评估其打折的情况

② 针对在打折的商品，其折扣率是多少

\* 用bokeh绘制折线图：x轴为折扣率，y轴为商品数量

③ 按照品牌分析，不同品牌的打折力度

\* 用bokeh绘制浮动散点图，y坐标为品牌类型，x坐标为折扣力度

提示：

① 打折情况评估方法：

真打折：商品的价格在10天内有波动、双11价格为10天内最低价、不存在涨价现象

不打折：商品价格无变化

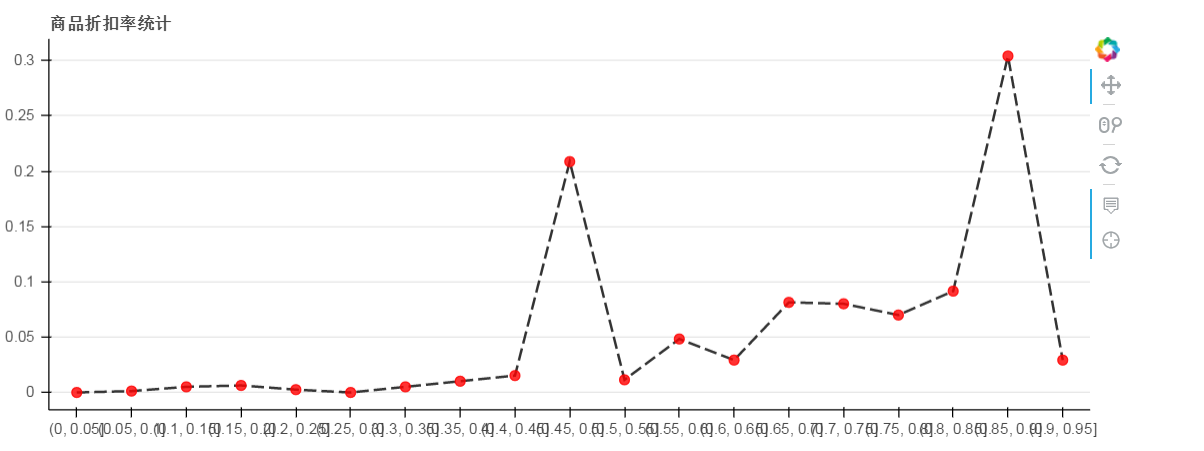
② 针对每个商品做price字段的value值统计，查看价格是否有波动，可以先用pd.cut将date分为不同周期：'双十一前','双十一当天','双十一后'，得到period字段

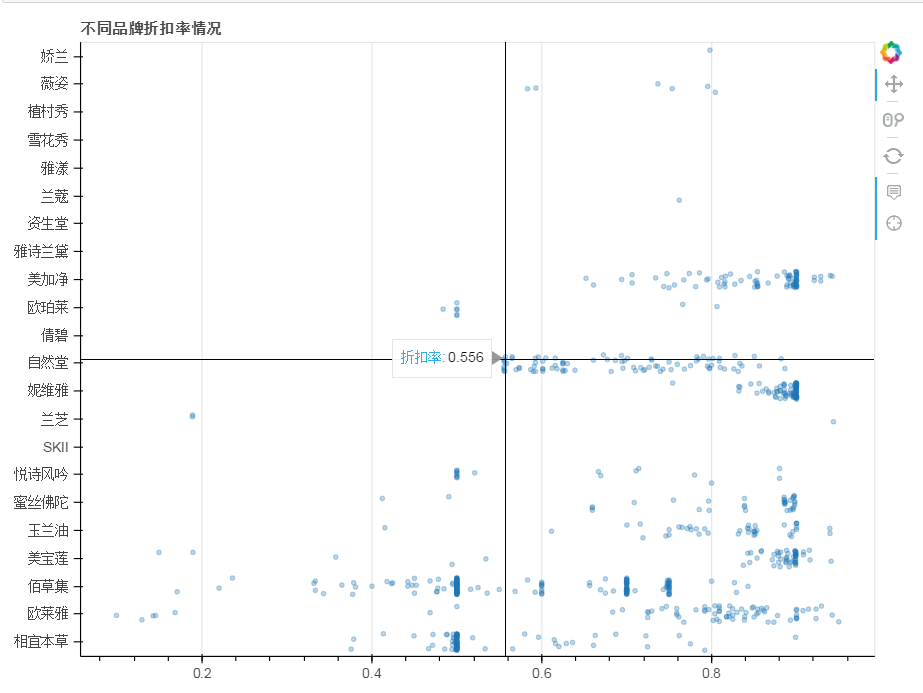
data[['id','price','date']].groupby(['id','price']).min()

针对统计出来的结果，如果按照id和price分组仍只有一个唯一值，则说明价格未变，没打折；否则为打折

③ 折扣率 = 双十一当天价格 / 双十一之前价格

④ 作图过程中，清除掉折扣率大于95%的数据





3、商家营销套路挖掘？

要求：

① 解析出不同品牌的参与打折商品比例及折扣力度，并做散点图，总结打折套路

\* 用bokeh绘制散点图，x轴为参与打折商品比例，y轴为折扣力度，点的大小代表该品牌参加双11活动的商品总数

提示：

① 折扣力度为该品牌所有打折商品的折扣均值，这里去掉品牌中不打折的数据

② 绘制散点图后，可以将x、y轴绘制均值辅助线，将绘图空间分为四个象限，基于该象限来总结套路

