第一题：

a,b两题答案已合并写在代码中

第二题：

首先需要对数据本身进行处理，检查是否有空值、错误的数据等。

鉴于商品属性不一样，有些商品属于快消品消费者回在一段时间内多次回购，有些商品属于耐用品短时间内消费者不会重复购买，可以用商品维度下半年内所有用户浏览、收藏、推荐，和购买数量的总数量进行KMeans聚类分析，看是否能通过消费者的消费习惯对商品进行分类。

想法一：

运用KMeans对用户进行聚类分析，使用每种商品下的用户根据半年内累计的浏览、收藏、推荐，和购买数量这四个数据，将该商品下分析结果中表现为对该商品感兴趣的用户推荐给拥有该商品的商家。由于题目中未提及是否有商品属性等信息，若有相关信息，可以对商品进行聚类分析，将同类商品下的分析出来的用户推送给经营该类商品的所有商家。

想法二：

运用Logistic Regression,以每个商品每个用户半年内累计浏览、收藏、推荐，和购买数量这四个数据为变量，回归的结果y放入由用户和商品构成的矩阵中，回归结果y值可以看作是每个用户对每个商品的感兴趣程度，越接近于1表明该用户对该商品越感兴趣，将结果高于threshold的那部分用户推荐给拥有该商品或该类商品的商家。

以上两个方法都可以得到每个商品下推荐给商家的用户名单，可以根据结果匹配出消费者购买行为发生的日期，对于有重复购买行为的消费者可以计算其平均复购间隔期限，由于数据为半年内的数据，时间上较为有限，若能发现消费者重复购买某一商品的时间规律，则可以利用该规律及该消费者最后一次购买日期，预测出下一次购买可能发生的时间，每月向商家重点推送最可能进行下一次购买的用户。

根据题目内容，已知数据集中不含有数据的标签等信息，可以根据商家反馈以及下一个月用户行为数据结果对模型进行改进，例如修改不同变量的权重，或者增加新的变量（例如：已有数据的平均数）等方法对于模型进行训练。