[**Grupo Bimbo Inventory Demand**](https://www.kaggle.com/c/grupo-bimbo-inventory-demand)

宾堡集团的库存需求

# Maximize sales and minimize returns of bakery goods

最大限度地提高销售和最大限度地减少烘焙食品的退回

对烘焙食品最大限度地提高销售和最大限度地减少退回

Planning a celebration is a balancing act of preparing just enough food to go around without being stuck eating the same leftovers for the next week. The key is anticipating how many guests will come. Grupo Bimbo must weigh similar considerations as it strives to meet daily consumer demand for fresh bakery products on the shelves of over 1 million stores along its 45,000 routes across Mexico.

计划一场活动是一种平衡的行为，是为了准备足够的食物而不是下周吃同样的剩菜。关键是预计会有多少客人来了。Grupo Bimbo必须在超过100万家商店的货架上沿其45000年在墨西哥航线上，对满足日常消费者对新鲜烘焙产品的需求有着近似的估算。

Currently, daily inventory calculations are performed by direct delivery sales employees who must single-handedly predict the forces of supply, demand, and hunger based on their personal experiences with each store. With some breads carrying a one week shelf life, the acceptable margin for error is small.

目前，每天的库存计算是由直接交货的销售人员完成，他们必须在他们个人经验的基础上单方面对每家商店的供应量，需求和欲望作出预测。对于一些有一个星期保质期的面包，可接受的预测误差是很小的。

In this competition, Grupo Bimbo invites Kagglers to develop a model to accurately forecast inventory demand based on historical sales data. Doing so will make sure consumers of its over 100 bakery products aren’t staring at empty shelves, while also reducing the amount spent on refunds to store owners with surplus product unfit for sale.

在这场竞争中，宾堡集团邀请kagglers开发一个模型根据历史销售数据来准确预测库存需求。这样做将确保消费者的100多个面包产品没有盯着空的货架上，同时也减少了花在退款的金额，以储存有剩余产品不适合出售的业主。

In this competition, you will forecast the demand of a product for a given week, at a particular store. The dataset you are given consists of 9 weeks of sales transactions in Mexico. Every week, there are delivery trucks that deliver products to the vendors. Each transaction consists of sales and returns. Returns are the products that are unsold and expired. The demand for a product in a certain week is defined as the sales this week subtracted by the return next week.

在这个竞赛中，你会预测一个产品的需求，在一个给定的一周，在一个特定的商店。您所提供的数据集包括9个星期在墨西哥的销售交易。每周都有送货车运送产品给供应商。每一笔交易都包括销售和。退回是未售出的和过期的产品。在某个星期的某个产品的需求被定义为本周的销售减去下一周的退回。

The train and test dataset are split based on time, as well as the public and private leaderboard dataset split.

训练和测试数据集在时间的基础上被分割，以及公共和私人的排行榜(积分榜)数据分割。

**Things to note:**

**注意：**

* There may be products in the test set that don't exist in the train set. This is the expected behavior of inventory data, since there are new products being sold all the time. Your model should be able to accommodate this.

有可能是产品在测试集里，而不存在在测试集里。这是库存数据的预期行为，因为有新的产品 随时备生产出来出售。你的模型应该能够适应这些。

* There are duplicate Cliente\_ID's in cliente\_tabla, which means one Cliente\_ID may have multiple NombreCliente that are very similar. This is due to the NombreCliente being noisy and not standardized in the raw data, so it is up to you to decide how to clean up and use this information.

有重复的客户ID在客户数据表中，这意味着一客户ID可能有多个客户名非常相似。这是由于这nombrecliente原始数据中的噪音和不规范，所以它是由你来决定如何清理和使用这些信息。

* The adjusted demand (Demanda\_uni\_equil) is always >= 0 since demand should be either 0 or a positive value. The reason that Venta\_uni\_hoy - Dev\_uni\_proxima sometimes has negative values is that the returns records sometimes carry over a few weeks.

调整后的需求（demanda\_uni\_equil）总是> = 0由于需求应该是0或者一个正数。venta\_uni\_hoy - dev\_uni\_proxima数据有时为负值的原因是，有时退回记录会延续几周。



