### CDNow网站用户消费数据分析

#### 一、背景和目的

**数据：**CDNow网站的消费数据，共69659条，4个字段，它们是ID、日期、数量和总金额。

**目的：**计算并分析主要运营指标，了解用户消费现状。

#### 二、**分析**思路

**分析内容：**

1. 客单价
2. 用户平均消费额
3. 复购率
4. 用户消费周期

**工具说明：**使用Python的pandas库清洗和分析数据。

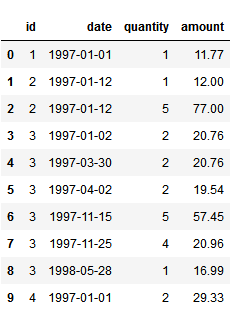
**代码地址**：

#### 三、**正文**

**1、数据表**

图 1是经简单清洗后的数据表，id代表用户ID，date代表消费日期，quantity代表购买CD的数量，amout代表订单金额。

图 1 数据表前10行

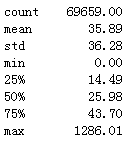


**2、客单价**

客单价概念来源于市场营销，指一位消费者每一次到场消费的平均金额。在互联网中，客单价指每一笔订单的平均金额，计算方法是总收入/订单数。

从数据中计算出的总收入为2500315.63美元，订单数为69659，客单价为35.89美元。订单金额的统计描述，如图 2所示。

图 2 订单金额统计描述



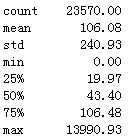
从图 2中发现，订单金额的中位数为25.98美元，小于客单价，这说明大部分订单金额小于客单价，少部分大于客单价，这符合电商行业的实际情况。另外，数据中存在订单金额为0但购买数量不为0的情况，原因可能是平台针对新用户实施了首单免费之类的优惠活动，这可以算作拉新成本。

**3、用户平均消费金额：**

用户的平均消费金额可以用来衡量用户的消费能力，针对不同消费能力的用户采取不同的营销策略，实现精细化营销。用户平均消费金额的计算方法是总收入/用户数。

从数据中计算出总收入为2500315.63美元，用户数为69659，用户平均消费金额为106.08美元。用户消费金额的统计描述，如图 3所示。

图 3 用户消费金额的统计描述



从图中可以看出，用户平均消费金额的第3分位数为106.47美元，略低于平均数，这说明75%的用户消费金额小于106.08美元。我们再计算贡献率，消费金额最高的25%的用户贡献了收入的 73.42%，这符合二八定律。这个25%的用户属于高价值用户，应该被重点关注。

**4、复购率**

复购率是指某时间窗口内消费两次及以上的用户在总消费用户中占比。复购率针对消费用户，更高的复购率意味着更高的收入。这里设时间窗口为1个月，复购率如图 4所示。

图 4 复购率

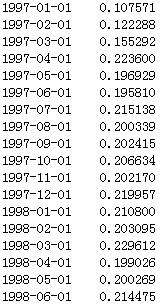
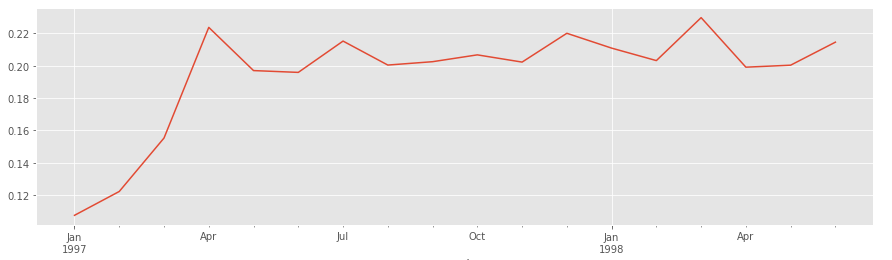


图 5 复购率折线图

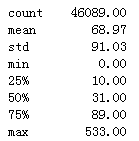


从图 5可以看出，前期复购率较低为10%左右，可能是因为有大量的新客产生，后期趋于稳定在20%左右。

**5、用户消费周期**

用户消费周期指两次消费间隔的时间，它表明了用户消费的时间规律。掌握用户大概的消费时间规律有助于制定更好的运营策略，引导用户更多地消费。用户消费周期的统计描述，如所示。

图 6 用户消费周期的统计描述



从图 6可以看出，用户的平均消费间隔时间是68天，可以在用户消费60天左右的时间后，通过短信邮件等方式，对用户进行营销。

#### 四、**小结**

本文根据用户消费数据，计算并分析了主要运营指标，对形成运营策略、促进增长具有重要意义。