

# 某教育机构财务周报分析

某教育公司主营成人教育培训及教育硬件等方面的产品，公司领导希望每周了解各主要产品的收入、成本及利润等情况。

1、产品类别：统考培训、会计培训、职业培训、职业资格、教学硬件

2、分析维度：

- 收入：每个产品类别的总收入、每周收入增量
- 成本：采购成本、返款成本、税金6%
- 利润：收入减去成本

3、需要的数据：订单数据

任务目标：一份反映各产品类别的收入、成本及利润等情况的表格

# 互联网产品用户留存率分析

在互联网行业中，经常要用到留存率这个指标，留存，是指产品的新用户，经过一段时间后，仍然继续使用该产品的用户，被称为留存用户，通俗地说，就是有多少新用户留了下来，反映了用户对于产品的忠诚度。

留存用户人数占当时新增用户的比例，就是用户留存率。

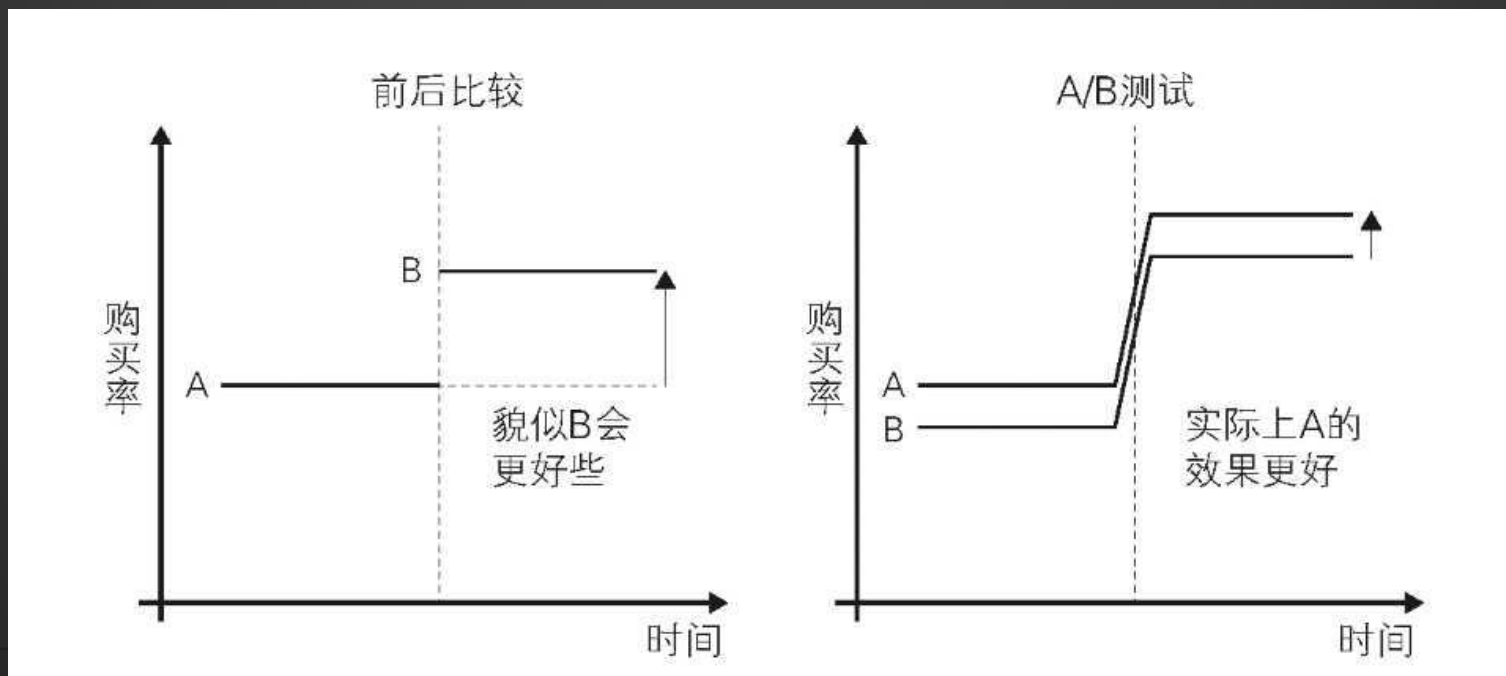
用户留存率可以按照天、周、月等时间维度来统计，常用的留存率有次日留存率、3日留存率、7日留存率、月留存率等。

各种留存率只是统计时间维度不同，求法大同小异，下面演示一下月留存率的求法。

## A/B测试：广告效果分析

A/B测试，简单来说，就是为同一目标制定A、B两种方案，并行测试，选择更优的一个方案。

例如，研究广告方案对购买率的影响，可以设计两套广告方案：广告A和广告B，如果采用前后比较法，则先执行广告A一段时间，然后执行广告B一段时间，然后比较前后两个时间段的广告方案的效果，如下面左边这个图所示。



## A/B测试：广告效果分析

进行A/B测试的时候，需要将用户分到A组或B组中，然后对A组用户投放广告A，对B组用户投放广告B，然后比较这两组用户的购买率。

将用户分成A、B两组的时候，需要保证随机性（**随机分配原则**），避免因为个体的差异对结果产生影响。

假设现在通过抽样的方式得到了两组用户，称为两个样本，现在需要判断这两组样本是否存在差异。

通过假设检验中的**卡方检验**来判断这种差异性。

## 案例：两种广告效果比较

某公司投放了两种广告计划（方案），想通过比较点击率知道哪种广告计划的效果更好？

现在已知两种广告计划的曝光次数和点击次数，如下表所示。

广告曝光日期	广告计划	用户ID	事务ID
2013/10/1	B	36703	25622
2013/10/1	A	44339	25623
2013/10/1	B	32087	25624
2013/10/1	B	10160	25625
2013/10/1	B	46113	25626
2013/10/1	A	6605	25627
2013/10/1	A	346	25628
2013/10/1	A	42710	25629
2013/10/1	A	37194	25630

广告点击日期	广告计划	用户ID	事务ID
2013/10/1	B	15021	25638
2013/10/1	B	351	25704
2013/10/1	B	8276	25739
2013/10/1	B	1230	25742
2013/10/1	B	17471	25743
2013/10/1	B	48728	25746
2013/10/1	B	16929	25769
2013/10/1	B	30111	25780
2013/10/1	B	30328	25791

## 案例：两种广告效果比较

主要操作步骤：

1. 根据事务ID合并数据集：将广告曝光次数表中的用户是否点击广告进行标记
2. 统计出广告计划A和B的点击次数、未点击次数和点击比例
3. 对广告计划A和B的点击率进行卡方检验（请参考统计学假设检验：卡方检验）

# 客户价值分析：RFM模型

在用户运营过程中，需要根据用户的属性对用户进行归类，以便于在运营过程中获得更大的收益。但是，用户有很多属性，究竟选择哪些属性进行分析呢？

根据美国数据库营销研究所Arthur Hughes的研究，客户数据库中有三个神奇的要素，这三个要素构成了客户分类最好的指标：

- **R, Recency, 最近一次消费距离现在的时长**：R越大，客户价值越低，如果太大，则视为流失客户。
- **F, Frequency, 消费频率**：F越大，客户价值越高，反之，则越低。
- **M, Monetary, 消费金额**：M越大，客户价值越高，反之，则越低。

利用R、F、M这三个指标对客户进行分类的模型被称为**RFM模型**，也叫作客户关系模型。

例如，客户A，R：180天，F：1次/月，M：100元/月；客户B，R：3天，F：10次/月，M：1000元/月。很明显，客户B价值远高于客户A

# 客户价值分析：RFM模型

在RFM模型中，将R、F、M这三个指标分别与其均值比较，将客户划分为8个类别。

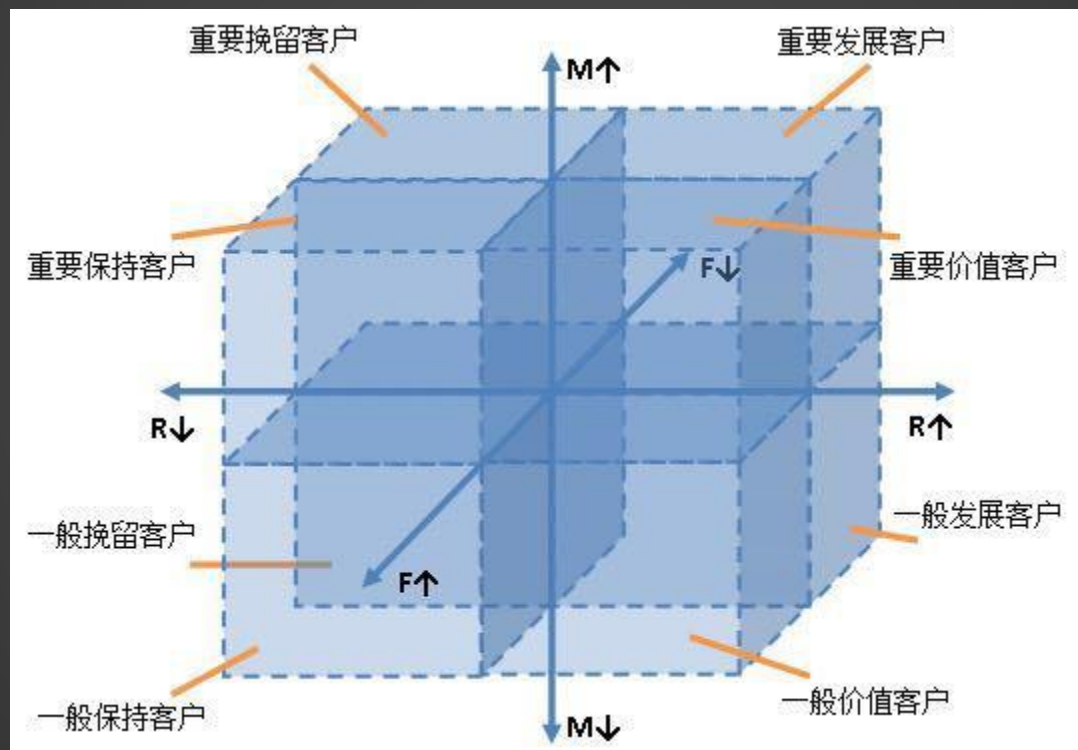
R	F	M	分析	客户类别
+	+	+	R, F, M都很大, 属于高价值客户	重要价值客户
+	+	-	最近有消费, 消费频率高, 但消费金额低	一般价值客户
+	-	+	最近有消费, 消费金额高, 但消费频率低	重要发展客户
+	-	-	最近有消费, 但消费频率、消费金额较低	一般发展客户
-	+	+	最近没有消费, 但消费频率、消费金额较高	重要保持客户
-	+	-	最近没有消费, 消费频率高, 但消费金额低	一般保持客户
-	-	+	最近没有消费, 消费频率低, 但消费金额高	重要挽留客户
-	-	-	最近没有消费, 且消费频率、消费金额都较低	一般挽留客户或流失客户

说明：R很高是指最近一次消费距离现在很近。



## 客户价值分析：RFM模型

将之前的客户分类表用一个三维坐标表示，R、F、M表示三个坐标轴，形成8个象限。



## 客户价值分析：RFM模型

构建RFM模型，需要构建R、F、M这三个指标，这三个指标又需要以下5个字段的数据：

1. 客户的编号：代表每一个客户
2. 客户的最近一次消费时间，可以计算出指标R：最近一次消费距离现在的时长
3. 客户的首次消费时间，可以计算出客户消费时长（天、月、季度等）
4. 客户的消费总次数， $\text{总消费次数} / \text{消费时长（月）} = \text{月均消费次数（日均消费次数等）}$
5. 客户的消费总金额， $\text{总消费金额} / \text{消费时长（月）} = \text{月均消费金额（日均消费金额等）}$

有了这三个指标之后，就可以按照前面所讲的思路来构建RFM模型。

## 案例：RFM客户价值分析模型

小李作为一家互联网公司的数据分析师，2016年7月20日，小李接到一个任务，公司希望进行用户精细化运营，也就是想知道现有用户中，哪些是高价值客户，哪些是一般价值客户，哪些是重要发展客户，哪些是流失客户等，然后对不同客户群体采取不同的运营策略。

实现步骤：

- 1、提数：提取用户相关的数据，包括5个字段：用户编码、首次消费时间、最近一次消费时间、消费总金额和消费总次数。
- 2、构建分析指标：利用数据源中的字段构建R、F、M三个指标，并进行标准化。
- 3、聚类分析：利用R、F、M这三个指标进行K均值聚类分析，将用户分成4个不同的类别
- 4、计算出这4个类别的R、F、M的均值，根据均值的高低标注用户类别。