## 淘宝用户行为数据分析

### 一、项目背景

本数据集共有大约1200万条数据，数据为淘宝APP2014年11月18日至2014年12月18日的用户行为数据。

### 二、问题描述

不同时间维度下用户活跃度如何变化？

用户的留存情况如何（复购率及漏斗流失情况）？

用户价值情况？

### 三、数据描述

本数据集共有大约1200万条数据，数据为淘宝APP2014年11月18日至2014年12月18日的用户行为数据，共6列字段，列字段分别是：

* + user\_id：用户身份，脱敏
  + item\_id：商品ID，脱敏
  + behavior\_type：用户行为类型（包含点击、收藏、加购物车、支付四种行为，分别用数字1、2、3、4表示）
  + user\_geohash：地理位置
  + item\_category：品类ID（商品所属的品类）
  + time：用户行为发生的时间

### 四、数据来源

<https://tianchi.aliyun.com/dataset/dataDetail?dataId=46&userId=1>

### 五、分析思路

**1.数据获取与读取**

**2.数据清洗**

**3.数据探索性分析（EDA）**

**4.用户行为分析**

用户行为定义和分类（如浏览、点击、购买等）。

用户活跃度分析。

用户路径分析。

转化率分析。

**5.用户群体分析**

用户分群（如新用户、活跃用户、沉默用户等）。

用户特征分析（如地域、设备、时间等）。

**6.模型构建与预测**

**7.结果展示与报告**