**某电商平台用户和投放广告数据分析项目**

* 1. **数据理解**

**.advertisement\_user.xlsx**

包含字段:

'用户id',

'时间戳-Bigint',:就是一个从1970-01-01 08:00:00到记录的时间的相隔的秒数

'广告id',

'资源位',:广告类型

'是否点击',投给的这个用户是否点击了广告

'群id',用户加入的不同群组

'性别',

'年龄层次',:含义如下:

0:18岁及以下

1:19-25岁

2:26-30岁

3:31-35岁

4:36-40岁

5:41-50岁

6:51岁及以上

'消费档次',

'用户层次',

'是否学生',

'城市层级',

'商品类别id',投放的广告对应的商品id

**. product\_price.xlsx**

包含字段:

['商品类别id',:与advertisement\_user.xlsx数据里的商品类别id一一对应

'商品品牌',

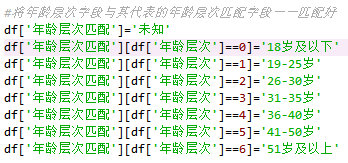
'商品价格']

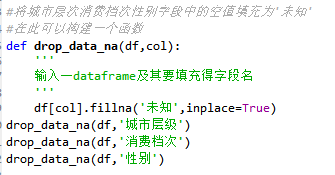
* 1. **数据清洗**

先读取数据



新建一个字段[年龄层次匹配],将advertisement\_user数据中的用户年龄层次数字与年龄层次含义匹配起来,并把年龄层次中的空值匹配为未知.





将两个表连接成为新表

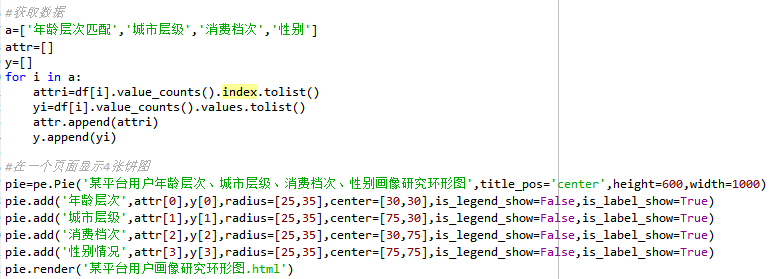


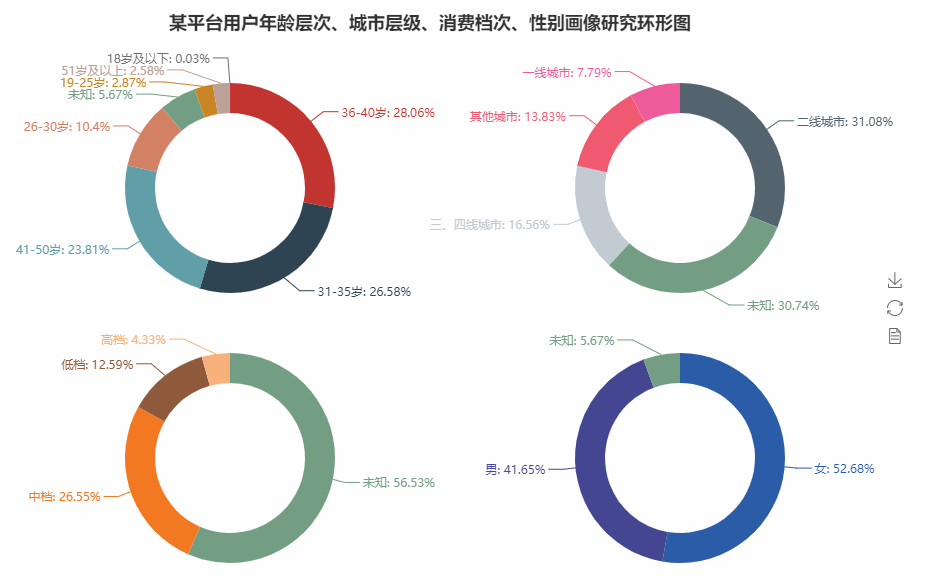
将新表中的时间戳-Bigint转化成时间序列



* 1. **数据可视化**

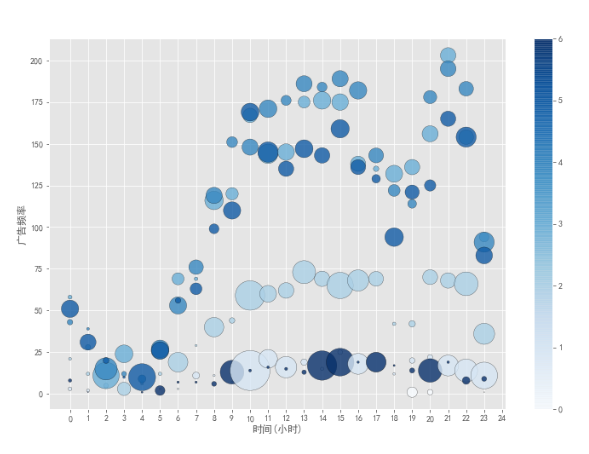
**.通过制作饼图来研究'年龄层次匹配','城市层级','消费档次','性别'四个字段的用户画像**





**.通过制作气泡图来研究该平台广告投放上的规律**





* 1. **分析与结论**

**对于该电商平台用户的用户画像研究**

1. 该平台的主要用户是30-40岁人群，其中以女性用户为主，猜测该平台的人群购买力好。

2. 在已知的城市层级和消费档次用户中，大部分人群来自二线城市，其次是三四线城市，一线城市用户最少，且其中的大部分人群的消费档次为中档，其次为底档，高档消费占的比例最低，因此猜测该平台的整体消费水平中等偏低。

**对于该电商平台投放广告的规律研究**

1. 位于图上方的气泡集中在10-15，20-22点，因此可以判断一天当中该平台投放广告的高频时间集中在10-15点以及20-22点。2. 偏大气泡和偏小气泡都在分布在图的下方，而中等的气泡在图的上方，因此可以判断该平台对于价格很高或很低的商品，投放频率都比较低频，中等价格区间的商品,投放频率高。3. 偏深色气泡和偏浅色气泡分布在图的下方且气泡大的居多，因此可以判断对于年龄小于18岁和年龄在51岁以上的用户，平台的策略倾向于低频高价。4. 对于年龄在19-25岁的用户，气泡大小适中偏大，频率也适中，因此平台的策略倾向于中频中高价格。