**一、问题**

这次分析主要想解决两个问题：

（1） 用户点击量和展示次数是否成正相关？

（2） 从性别、兴趣、年龄三个方面来看，哪些广告更有投入价值？

（3） 哪次广告活动更有价值

**二、数据来源**

数据来自kaggle，此项目中使用的数据来自匿名组织的社交媒体广告。文件conversion\_data.csv包含11个变量中的1143个观测值。以下是变量的描述。

1.）ad\_id：每个广告的唯一ID。

2.）xyz\_campaign\_id：与XYZ公司的每个广告活动相关联的ID。

3.）fb\_campaign\_id：与Facebook如何跟踪每个广告系列相关联的ID。

4.）age：广告所针对的人的年龄。

5.）gender：显示添加的人的性别

6.）interest：指定该人的利益所属类别的代码（利益在该人的Facebook公开简介中提及）。

7.）impressions：展示广告的次数。

8.）clicks：该广告的点击次数。

9.）spent：公司xyz向Facebook支付的金额，以显示该广告。

10.）Total\_Conversion：看到广告后询问产品的总人数。

11.）Approved\_Conversion：看到广告后购买产品的总人数。

**三、数据清洗**

（1）数据整理

因为本次分析从xyz公司的角度出发，故删除第3列和Facebook相关的广告ID。同时为了方便观察，将原有的英文的列重新用中文命名。

（2）引入点击率（CTR）、转化率(CR)、广告投资回报率（ROAS）

考虑到原始数据中只给出了销售量，没有销售额，投资回报率这个指标又比较重要，故假设每一次成功的销售能带来100美元的销售额。

指标的计算方式如下：

点击率=点击次数/展示次数

转化率=购买人数/点击次数

广告投资回报率=购买人数\*100/广告花费

此时发现转化率存在大于1的情况，这可能是因为Facebook提供的时间点和xyz公司统计的时间点存在区别，也可能是系统错误等原因，因此将此类数据剔除。同时剔除点击次数为0（无需支付费用）的情况，后还剩935条数据。

**四、数据分析**

（1） 用展示次数和用户点击量是否成正相关？

在散点图中我们可以发现，展示次数和点击次数成正相关，相关性没有那么强，存在部分广告展示次数高但是点击次数低。这说明部分广告的投放效果没有那么理想，可能是部分广告的目标受众不明确。

（2） 从性别、年龄、兴趣三个方面来看，哪些广告更有投入价值？

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **行标签** | **平均值项:点击率(CTR)** | **平均值项:转化率（CR）** | **平均值项:广告花费** | **平均值项:广告投资回报率（ROAS）** |
| F | 0.000230502 | 0.09031975 | 74.67755403 | 6.768662224 |
| M | 0.000170904 | 0.123307764 | 51.16830866 | 9.941962154 |
| **总计** | **0.000200352** | **0.107007804** | **62.78464167** | **8.373978659** |

从性别角度看，女性用户的点击意愿较高，但是转化率较低。是因为商品无法达到女性用户的预期？或者是销售话术有待改进？而男性用户正好相反，点击广告后购买意愿较强，可以考虑优化广告展示质量或者是广告投放的精准度。总的来说公司在对女性用户广告投入较大，投资回报率却没有那么理想，可以适当增加对男性用户的广告投放力度。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **行标签** | **平均值项:点击率(CTR)** | **平均值项:转化率（CR）** | **平均值项:广告花费** | **平均值项:广告投资回报率（ROAS）** |
| 30-34 | 0.000160683 | 0.148367073 | 49.83924832 | 10.72807505 |
| 35-39 | 0.000191065 | 0.12215294 | 52.665545 | 8.983983247 |
| 40-44 | 0.000225096 | 0.074049225 | 61.97716567 | 5.438192046 |
| 45-49 | 0.000241355 | 0.065067089 | 89.82974025 | 7.074964076 |
| **总计** | **0.000200352** | **0.107007804** | **62.78464167** | **8.373978659** |

从年龄角度看，对30-34岁年龄段受众投放的广告投资回报率较高，更适合投放广告。而随着受众年龄增大，广告点击率出现了上升，但是转化率没有增加，这可能因为他们对商品的要求较高或者付费意愿不强，是否是商品定价过高或者是功能上满足不了他们的需求。可以考虑优化产品或者是把更多的资源倾向30-34岁的受众。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **行标签** | **平均值项:点击率(CTR)** | **平均值项:转化率（CR）** | **平均值项:广告花费** | **平均值项:广告投资回报率（ROAS）** |
| 63 | 0.000206369 | 0.089420336 | 65.37000006 | 18.7716095 |
| 31 | 0.000204968 | 0.231480662 | 17.59411759 | 17.64185619 |
| 104 | 0.000160316 | 0.214354067 | 77.58399985 | 14.39874961 |
| 101 | 0.000142856 | 0.209134356 | 108.2414282 | 14.03581179 |
| 15 | 0.000171841 | 0.175369485 | 66.59641013 | 13.54077612 |

从兴趣分类看，上述兴趣标签的人更愿意购买公司的产品。兴趣编号为63号的转化率明显偏低，需要具体分析原因。

（3） 哪次广告活动更有价值

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **行标签** | **平均值项:点击率(CTR)** | **平均值项:转化率（CR）** | **平均值项:广告花费** | **平均值项:广告投资回报率（ROAS）** |
| 916 | 0.000277941 | 0.257142857 | 4.27742859 | 19.42862583 |
| 936 | 0.000265574 | 0.205873896 | 10.0758885 | 17.91693061 |
| 1178 | 0.000165387 | 0.052147616 | 90.80285474 | 3.274892427 |
| **总计** | **0.000200352** | **0.107007804** | **62.78464167** | **8.373978659** |

从广告活动角度看，1178活动的广告投资回报远低于前面两次，在点击率降低的情况下广告费用反而出现了大量的增加，是否因为广告投放数量的增加而忽视了投放的质量，投放广告的渠道是否出现了问题。而转化率很低，可以考虑是因为广告投放和目标用户的不匹配或是此次活动的广告设计存在问题。