

优达学城数据分析师纳米学位

A/B 测试项目

说明：[点击此处下载此文档的英文版本](#)。

模板格式

此模板可用于组织你最终项目的答案。应从你的答案中复制到小测试中的项用蓝色显示。

试验设计

指标选择

列出你将在项目中使用的不变指标和评估指标。（这些应与你在“选择不变指标”和“选择评估指标”小测试中使用的指标一样）

对于每个指标，解释你为什么使用或不使用它作为不变指标或评估指标。此外，说明你期望从评估指标中获得什么样的试验结果。

不变指标：

1. **COOKIE 的数量：**即访问课程概述页面的唯一 **COOKIE** 的数量。（D 最小=3000）
2. **点击次数：**即点击“开始免费试学”按钮的唯一 **COOKIE** 的数量（在免费试学筛选器触发前发生）。（D 最小 =240）
3. **点进概率：**即点击“开始免费试学”按钮的唯一 **COOKIE** 的数量除以查看课程概述 页的唯一 **COOKIE** 的数量所得的比率（D 最小=0.01）

实验研究为学生设定明确的期望，是否会减少因时间不够而受挫取消离开免费试学的学生数量。并不会在很大程度上减少继续通过免费试学并最终完成课程的学生数量。

实验通过添加学生点击免费试学时弹出选择学习时长以给学生提供明确学习期望，并根据时长的输入结果指向不同的用户下一步行为。

访问课程概述页面行为发生在‘弹框’之前，在实验组和对照组中访问课程概述页面的 **COOKIE** 数量应无显著差异。

点击‘开始免费试学’发生在‘弹框’前后都有，但唯一 **COOKIE** 应该‘弹框’前的数量较多，唯一 **COOKIE** 大致等于点击‘弹框’前的唯一 **COOKIE** 量，这个量在实验中应不受扰动。

点击‘开始免费试学’量按 **COOKIE** 分可以避免累计用户在‘弹框’前后的两次点击行为，访问课程概述页面 **COOKIE** 数量应无差异，点进概率也应无太大差异。

评估指标：

1. **留存率：**即在 14 天的期限过后仍参加课程（因此至少进行了一次付费）的用户 ID 数量除以完成登录的用户 ID 的数量。（D 最小 =0.01）
2. **净转换率：**即在 14 天的期限后仍参与课程的用户 ID 的数量（因此至少进行了一次付费）除以点击了“开始免费试学”按钮的唯一 **COOKIE** 的数量所得的比率。（D 最小

=0.0075)

3. 总转化率：即完成登录并参加免费试学的用户 ID 的数量除以点击“开始免费试学”按钮的唯一 COOKIE 的数量所得的比率。(D 最小 =0.01)

实验是为了防止学生受挫并且不减少原来的通过体验课并付费的学生数量，应当看到原来因时间少而受挫离开学生将不太可能进入免费试学环节，而最终付费学生数量却无明显减少。因此留存率应该增加，净转化率应该‘不降低’，即置信区间大于等于 0。

对于 ID，可能参与免费试学的学生会减少，导致出现差异，但由于实验对照组中 COOKIE 量可能不同，所以参与实验的 ID 数量差异可能是由于 COOKIE 不同导致的。

由于点击‘开始免费试学’的学员可能在被询问学习时间后不再打算参加学习，所以总转化率可能降低。

测量标准偏差

列出你的每个评估指标的标准偏差。(这些应是来自“计算标准偏差”小测试中的答案。)

1. RETENTION

0.0549

2. NET CONVERSION

0.0156

3. CROSS CONVERSION

0.0202

对于每个评估指标，说明你是否认为分析估计与经验变异是类似还是不同（如果不同，在时间允许的情况下将有必要进行经验估计）。简要说明每个情况的理由。

RETENTION 分析单位为 ID,分组单位为 COOKIE,分析估计与经验变异不相似。

NET CONVERSION 分析单位与分组单位一致，分析估计与经验变异类似。

CROSS CONVERSION 分析单位与分组单位一致，分析估计与经验变异类似。

规模

样本数量和功效

说明你是否会在分析阶段使用 Bonferroni 校正，并给出实验正确设计所需的页面浏览量。（这些应是来自“计算页面浏览量”小测试中的答案。）

不使用 BONFERRONI 校正，

留存率所需页面浏览量为 4741213，

净转化率所需页面浏览量为 685325，

总转化率所需页面浏览量为 645875。

持续时间和曝光

说明你会将多少百分比的页面流量转入此试验，以及鉴于此条件，你需要多少天来运行试验。（这些应是来自“选择持续时间和曝光”小测试中的答案。）

使用 NET CONVERSION 作为度量，比例 0.685325，时间 25 天。净转化率置信区间应大于等于 0。

说明你选择所转移流量部分的原因。你认为此试验对优达学城来说有多大风险？

- 1.实验只是询问学习时间，不属于敏感数据，不影响受试者身心健康。

- 2.即使学生每周学不到五小时，他们只是被页面的变更提醒引导到了另外的一个页面，如果今后有需要学生仍然可以进入免费试学、登陆并可能完成继续课程的，不会因此影响用户使用网站的习惯。
- 3.该试验没有关于数据库及后台的改变，不用担心数据的丢失及由于后台的失误导致网页奔溃用户无法访问网页等大问题。
- 4.此试验也不会对用户的个人信息安全造成风险，因为不论网页是否增加了提醒，用户在确认参加免费试学时都得输入信用卡信息，而很明显系统一定会保护用户的个人信息。
- 5.考虑实验周期不宜过长，选择了 30 天以下，即流量 0.685325，周期 25 天。

试验分析

合理性检查

对于每个不变指标，对你在 95%置信区间下期望观察到的值、实际观察的值及指标是否通过合理性检查给出结论。（这些应是来自“合理性检查”小测试中的答案）

Sanity checks

For each metric that you chose as an invariant metric, compute a 95% confidence interval for the value you expect to observe. Enter the upper and lower bounds, and the observed value, all to 4 decimal places. Check the box if the metric passes your sanity check.

	Lower bound	Upper bound	Observed	Passes
• Number of cookies	<input type="text" value="0.4988"/>	<input type="text" value="0.5012"/>	<input type="text" value="0.5006"/>	<input checked="" type="checkbox"/>
• Number of user-ids	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>
• Number of clicks on "Start free trial"	<input type="text" value="0.4959"/>	<input type="text" value="0.5041"/>	<input type="text" value="0.5005"/>	<input checked="" type="checkbox"/>
• Click-through-probability on "Start free trial"	<input type="text" value="0.0812"/>	<input type="text" value="0.0830"/>	<input type="text" value="0.0821"/>	<input checked="" type="checkbox"/>
• Gross conversion	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>
• Retention	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>
• Net conversion	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>

对于任何未通过的合理性检查，根据每日数据解释你觉得最有可能的原因。在所有合理性检查通过前，不要开始其他分析工作。

结果分析

效应大小检验

对于每个评估指标，对试验和对照组之间的差异给出 95% 置信区间。说明每个指标是否具有统计和实际显著性。（这些应是来自“效应大小检验”小测试的答案。）

Effect Size Tests

For each of your evaluation metrics, compute a confidence interval around the difference.

Did you use the Bonferroni correction? ☐ Yes ☒ No

	Lower bound	Upper bound	Statistical significance	Practical significance
• Number of cookies	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Number of user-ids	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Number of clicks on "Start free trial"	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Click-through-probability on "Start free trial"	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Gross conversion	-0.0291	-0.0120	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
• Retention	0.0081	0.0541	<input checked="" type="checkbox"/>	<input type="checkbox"/>
• Net conversion	-0.0116	0.0019	<input type="checkbox"/>	<input type="checkbox"/>

留存率由于实验要求浏览量过多，实验间不能采集够样本量，所以可以不使用它作为评估指标。

符号检验

对于每个评估指标，使用每日数据进行符号检验，然后报告符号检验的 p 值以及结果是否具有统计显著性。（这些应是“符号检验”小测试中的答案。）

Sign Tests

Run a sign test on each of your evaluation metrics using the day-by-data data. Enter each p-value, and indicate whether each result is statistically significant.

Did you use the Bonferroni correction? ☐ Yes ☒ No

	p-value	Statistical significance
• Number of cookies	<input type="text"/>	<input type="checkbox"/>
• Number of user-ids	<input type="text"/>	<input type="checkbox"/>
• Number of clicks on "Start free trial"	<input type="text"/>	<input type="checkbox"/>
• Click-through-probability on "Start free trial"	<input type="text"/>	<input type="checkbox"/>
• Gross conversion	0.0026	<input checked="" type="checkbox"/>
• Retention	0.6776	<input type="checkbox"/>
• Net conversion	0.6776	<input type="checkbox"/>

汇总

说明你是否使用了 Bonferroni 校正，并解释原因。若效应大小假设检验和符号检验之间存在任何差异，描述差异并说明你认为导致差异的原因是什么。

BONFERRONI 校正不考虑实验度量间是否独立，它对非独立度量的显著性校正过于保守。总转化率和净转化率间相互关联，因此不建议使用 **BONFERRONI** 校正。

建议

提供建议并简要说明你的理由。

不建议进行更改，总转化率达到了统计和实际显著性，符号检验也具有显著性，且是我们希望看到的结果。但是净转化率置信区间包含负数，也就是说净转化率很大概率可能减小，并且有一定概率达到实际显著性 **0.0075**，所以无法说明‘不会在很大程度上减少继续通过免费试学和最终完成课程的学生数量’，所以不建议启动更改。

后续试验

对你会开展的后续试验进行概括说明，你的假设会是什么，你将测量哪些指标，你的转移单位将是什么，以及做出这些选择的理由。

进行对参加免费试学的学生打卡记分的实验，在免费试学期间每周完成课程量达到讲师设定的不同级别的学习目标后可以获得不同的积分，攒够一定积分可以兑换免费试学课程的 VIP，不同等级的 VIP 在信用卡扣费时有不同的打折力度，免费试学期相应课程的积分自免费试学结束起只保留一个月，过时清 0。

假设学生为了获得付费时的 VIP 拥有的打折力度而会在免费试学时努力达到目标，因而腾出时间去学习，于是‘提前终止’的行为会减少。

通过观察净转化率作为指标，点击的 COOKIE 不太会变化，期望最后付费的 ID 会增加。所以净转化率会增加。

转移单位为 ID,以保证用户前后的一致体验。

参考

[优达城论坛-A/B 测试 review01](#)

优达学城

2017 年 10 月