

# 互联网人必备思维方式

AB测试|数据分析师晋升篇

CDA数据分析师|唐辰

走进AB测试

---

拓展AB测试

---

实战AB测试

---





# 走进AB测试

---

初步认识

熟悉核心

应用场景

实施步骤

推进节奏

注意事项

# AB测试来源

## 双盲测试

双盲测试中病人被随机分成两组，在不知情的情况下分别给予安慰剂和测试用药，经过一段时间的实验后再来比较这两组病人的表现是否具有显著的差异，从而决定测试用药是否有效。

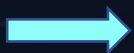


## AB测试

将Web或App界面或流程的两个或多个版本，在同一时间维度，分别让两个或多个属性或组成成分相同（相似）的访客群组访问，收集各群组的用户体验数据和业务数据，最后分析评估出最好版本正式采用。

# AB测试要点

两个关键



同一时间维度

成分相同（相似）的访客群组

一个任务



收集各群组的用户体验数据和业务数据

一个目标



分析评估出最好版本正式采用



# AB测试应用场景

×

注册

用QQ注册

用微信注册

用电子邮箱注册

名字

Paul Jensen

邮箱

paul@invisionapp.com

密码

.....

性别

☒ 帅哥 ☐ 美女

生日

1983年09月12日

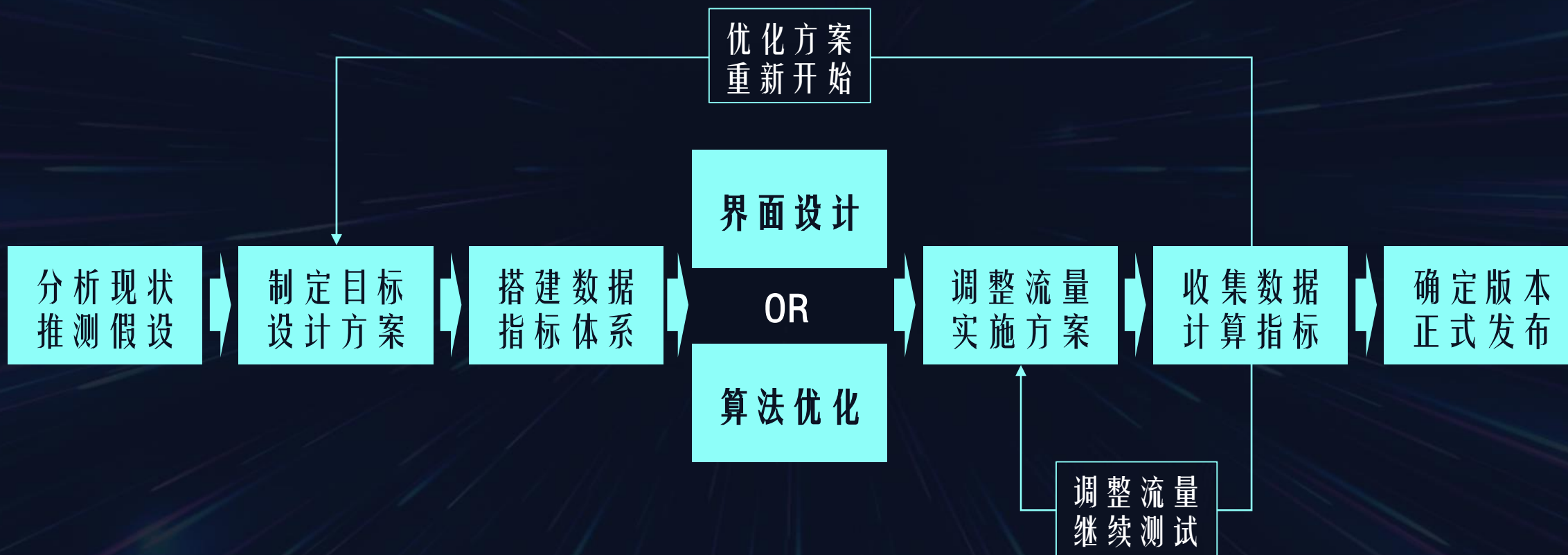
提交

界面设计

算法优化



# AB测试实施步骤



# 测试流程

- 1、分析实际业务场景，明确改进关键点，制定优化建议。
- 2、与项目经理确定优化目标，设计优化方案。
- 3、根据优化目标及方案搭建数据分析指标体系，明确数据口径。
- 4、由UI设计师或算法工程师根据方案完成样例开发，并设定数据埋点。
- 5、调整流量池，在同一时间内对目标用户开始测试。
- 6、根据时间进度及时收集数据，并根据数据情况调整流量池或重新进行测试。
- 7、经过多次试验取得成功后，确定新版本功能，正式发布。



# AB测试推进节奏

## 灰度测试

由于调整版本对企业影响极大，设定目标中的两个版本可能都不是有效版本，故不能将所有用户直接一分为二全部进行AB测试，对此可以采用灰度发布的方式，即只选择其中极小部分用户进行AB测试，并逐步推广到所有用户。

推进进度：

灰度 ▶ 5% ▶ 10% ▶ 20% ▶ 50% ▶ 80% ▶ 100%

# AB测试注意事项（一）

## AB测试进行中，个别用户结果被质疑

AB测试需要分析参与测试用户的整体分布情况，而不仅仅是个别用户的情况。

如出现了个别异常值，重点是需要保证整体分布没有大的异常即可继续进行AB测试，对于异常值，可以重点关注其产生的根本原因，可能是其他可优化点。

# AB测试注意事项（二）

## AB测试中，数据的有效性

正常情况下，在推进AB测试时，前三天的数据波动会比较大，不建议直接使用；一般进行AB测试需要大约7-14天为周期进行推进，切不可心急。

针对不同行业应用的产品，也需要根据用户使用周期进行调整，一定要完整覆盖整个用户周期。如电商平台在进行AB测试时，就需要注意工作日与节假日的影响不同，需要将测试周期定在一周以上。



# AB测试注意事项（三）

## AB测试开始后，用户反馈反方向发展

在这种情况下，先缩小测试用的流量池，减少测试的用户数量；另一方面，分析目前的用户反馈情况，跟进反馈背后问题的真实原因，优化调整AB测试方案。

千万不要着急直接停止AB测试，认真分析问题产生的根源，待情况明确，新方案制定后，继续进行AB测试。

# 拓展AB测试

---

ABN测试

多维测试

ANB测试

任务与职责

# ABN测试

在AB测试设计初期，可能提出了多种优化版本，对此并不一定只能要求其中一种版本进行AB测试，可同时选择类型相同的版本同时进行AB测试，此类方法称为ABN测试。

如下针对注册按钮，同时测验用户对四种颜色的反映情况，以判断能否提升注册率。

注册/登录

注册/登录

注册/登录

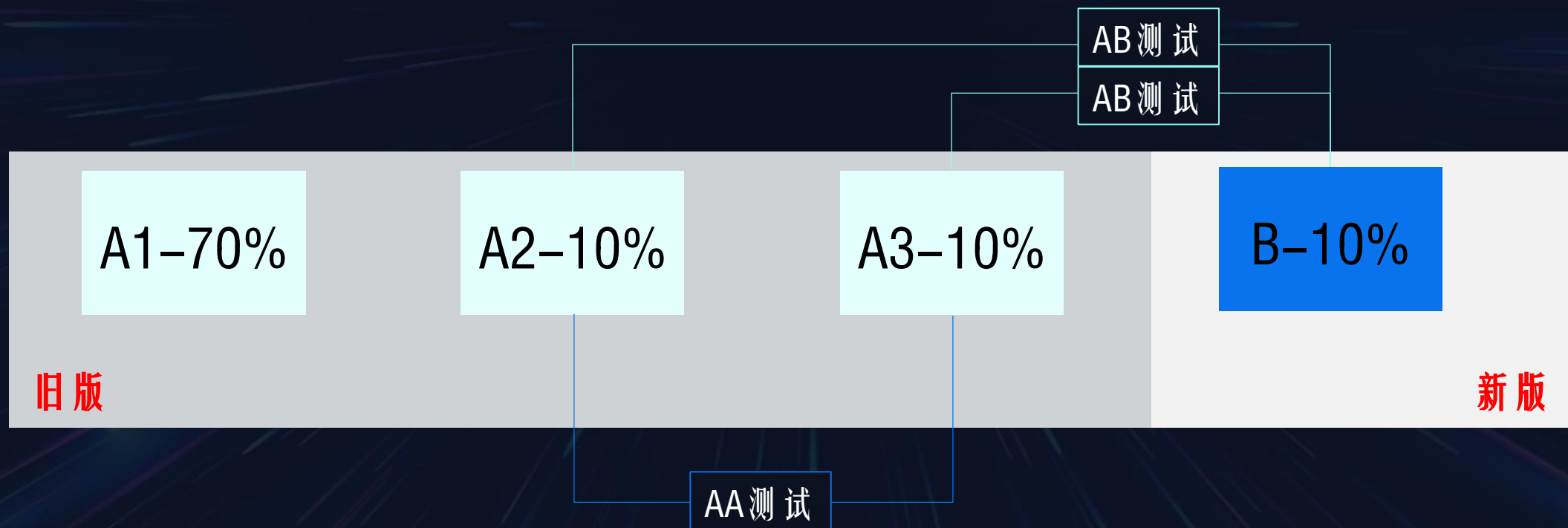
注册/登录

如登录用的LOGIN的按钮颜色，



# ANB测试

在进AB测试过程中，可能会有个别特殊因素对测试结果有异常明确的影响，如遇到有特殊用户刷量等操作。为了避免此类情况的发生对AB测试产生影响，可以考虑进行ANB测试（AA测试的加强版），具体原理如下图：



# 多维测试

按照正常情况，完成一个AB测试的整个周期约为15天-1个月，按照这种情况，一年能完成的AB测试不会超过24个，为了保证高速迭代速度，可以在多个维度内同时推进AB测试，但是一定要保证任意两个维度内互不影响。

$A \alpha - 25W$

$A \beta - 25W$

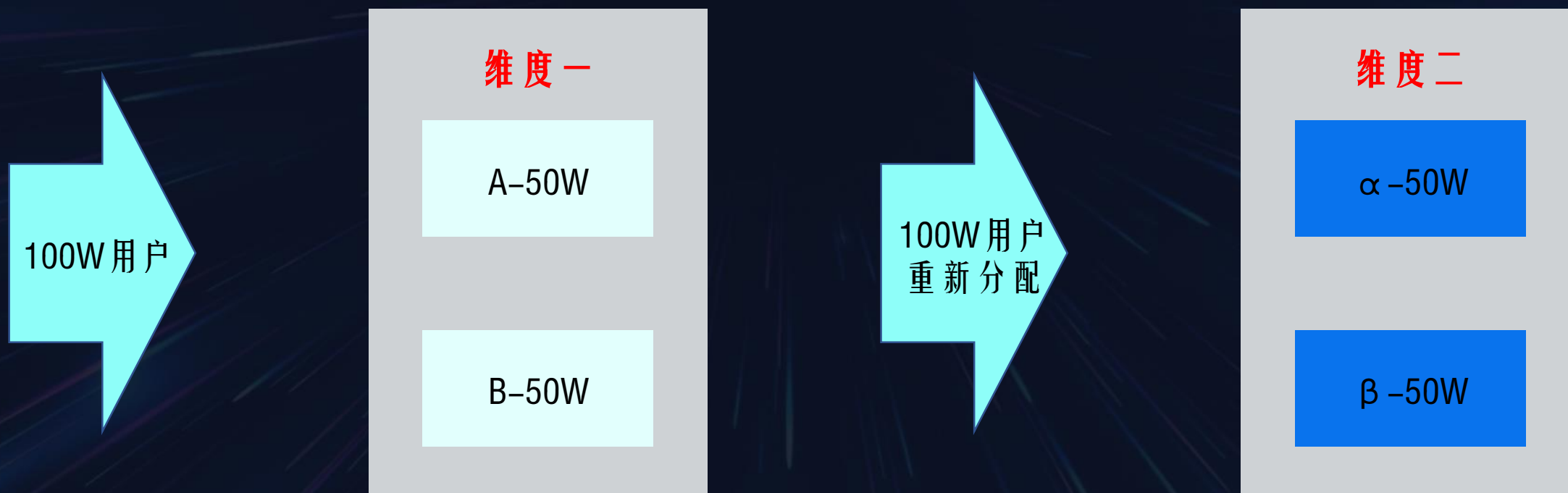
$B \alpha - 25W$

$B \beta - 25W$

100W 用户

# 多维测试优化

在只有有限的用户流量池时，可以将目标范围用户根据测试维度的不同，分别使用两次切分方式进行测试。如下图，目标用户在完成维度一的测试之后，在重新进行分配，分别在维度二的测试，这样既保证了用户数量，也能确定确保各个维度之间互不干扰。





# 任务与职责

## 数据分析师需要扮演的角色：

- 1、针对业务场景，制定专题分析，明确分析目标。
- 2、AB测试初期，分析业务场景，提出优化方向。
- 3、与产品经理沟通设计解决方案，制定数据分析体系，明确数据口径。
- 4、与测试人员明确流量分配情况，确保流量分配方式有效。
- 5、及时收集数据，分析并对测试过程不断优化。
- 6、跟进测试过程，了解测试用户的反馈情况，分析其他优化点。
- 7、完成测试后，撰写专题报告，整理测试方案、测试数据、测试报告等完成文件归档。

# 任务与职责-专题分析

**专题分析**是指针对某一专门问题所作的不定期的专项分析，这种分析内容集中，目标明确，遇有情况发生或问题发现，可随时进行开展，以便采取有效措施，改进工作。

作为数据分析师，每一个AB测试，都需要将其作为专题分析对待。



# 任务与职责-优化方向

在进行AB测试时，多数情况属于通过日常数据发现了用户的某些指标波动明显，或是为了提升注册量等因素，而本次优化不论是基于UI方面，或者是算法方面，都需要和相关产品经理进行沟通，明确本次测试的具体优化方向。

与产品经理沟通优化方向时，切记自己手中的数据情况，以及本次测试的优化目标，一定要将**目标与现实实际情况相结合**，针对性制定优化方向。





# 任务与职责-分析体系

设计分析体系时，需要注意以下几点：

- 1、围绕核心目标，先行制定出最重要的指标；
- 2、优化并不是一个指标就能反映真实情况的，而且也无法追述原因；
- 3、指标体系需要由核心指标进行拆解，逐步到可落地实施；
- 4、在制定指标时，不能为了制定指标而制定指标，需要结合实际业务场景；
- 5、在明确指标的同时，也需要明确取数的口径以及推算要求。



# 任务与职责-流量分配

在进行测试时，流量分配重要是由技术测试人员进行设计和调整，但是作为数据分析师，需要及时与相关同事进行沟通，以了解具体的流量分配方式是否与此次测试要求相匹配。

在测试过程中，如发现流量分配方式需要调整，需要及时沟通，以确保测试能够稳定推进。



# 任务与职责-测试优化

从此阶段开始，数据分析师将开始面临整个测试过程中最重要的工作。

首先，需要加强注意数据的收集与整理，判断收集来的数据是否完整，能否支持分析；

其次，对数据进行整理与分析，确定数据情况能否反映真实业务，与预期情况的匹配程度；

再次，通过数据结果，判断是否需要调整流量分配情况进行调整，是否需要加大测试量；

再次，如一段时间的数据情况并不理想，判断真实原因，以及方案调整情况；

最后，注意收集测试期间用户的反馈，为下一次测试提前做好准备。





# 任务与职责-分析总结

测试进行到此阶段时，基本将接近尾声，但是仍需要小心谨慎。

首先，对本次测试情况进行整理汇总，分析测试的每一个环节；

其次，需要针对收集到的数据完成数据整理与分析，完成数据指标体系的计算及分析；

再次，完成整个测试的情况梳理，将所有重要情况整理为汇报材料；

最后，向上级汇报本次测试情况，复盘及梳理测试中的不足与改进点。



# 任务与职责-文件归档

到了此环节，表明本次测试基本告一段落，但是切记，需要收集整理本次测试过程中的所有文件，并完成编号归档，根据公司情况的不同，一般需要归档的材料包括但不限于以下几类：

- 测试启动会议记录
- 测试方案研讨会会议记录
- 测试实施方案及进度甘特图
- 测试目标及数据指标体系
- 测试过程数据集和数据口径
- 测试最终总结报告



# 实战AB测试

---

商业理解

数据理解

数据整理

数据分析

形成结论

报告呈现



# 商业理解

本次实战的数据集是从网络获取的，具体情况展示如下：

## 商业背景

我们网站开通了，然后我们给网站投放广告的时候，一开始给大家看到的落地页是访问课程资料，现在我们又推出了开始免费试学这个落地页，然后想着以后换上新的落地页，这时候就要检测两个落地页的转化效果了，下面是测试的24天，看数据会有24天的。

## 目的

确定是使用新落地页，还是使用旧落地页，或是延长测试时间。

数据来源：<https://www.jianshu.com/p/1c0c4c3ea889>

# 数据理解

user\_id



用户的id

timestamp



用户访问页面的时间

group



把新的落地页分到treatment组、把旧的落地页分到了control组

landing\_page



落地页

converted



是否转化，1表示转化，0表示未转化

# 数据整理

- 1、观察数据集，了解数据整体情况。
- 2、查看并去除缺失值，保留有效数据。
- 3、landing\_page 字段和converted 字段中：  
treatment 和 new\_page 相对应；control 和 old\_page 相对应。
- 4、去除重复数据，保证数据唯一性。



# 数据分析

1、确定用户数量、计算落地页的转化率及对比情况。

2、设计原假设和备择假设，分析数据结果。

原假设  $H_0: P_{\text{new}} - P_{\text{old}} \leq 0$

备择假设  $H_1: P_{\text{new}} - P_{\text{old}} > 0$

3、计算显著性水平Z值。

$$Z = \frac{P_b - P_a}{\sqrt{\frac{P_b(1 - P_b)}{n_b} + \frac{P_a(1 - P_a)}{n_a}}}$$

# 形成结论

## 数据

- 1、实际参与AB测试的用户共计290584名。
- 2、用户转化率：  
整体（11.96%），新页面（11.88%），旧页面（12.04%）
- 3、用户收到新旧页面的概率：50%
- 4、显著性检验Z值为1.31，落在95%的置信区间中。

## 结论

建议延长测试时间，继续观测情况。

# 欢 迎 关 注

数据分析师晋升系列文章

CDA数据分析师|唐辰