

A/B Testing

جهان به مانند یک آزمایشگاه

گروه دایچه . dayche.com



□ همبستگی یا علیت؟

مثال 1: همبستگی بالا بین مصرف بستنی و آفتاب سوختگی

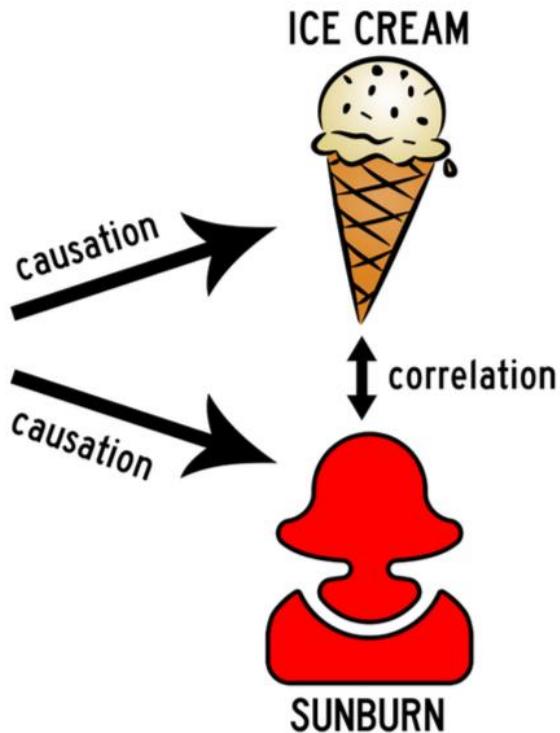
علت اصلی: هوای خشک و آفتابی در تابستان

مثال 2: همبستگی معنادار بین سلامتی و مصرف کنترل شده الکل

علت اصلی: وضعیت روانی سالم در خانواده و دوستان



DRY, HOT AND SUNNY
SUMMER WEATHER



سوگیری حذف
عامل اصلی

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه

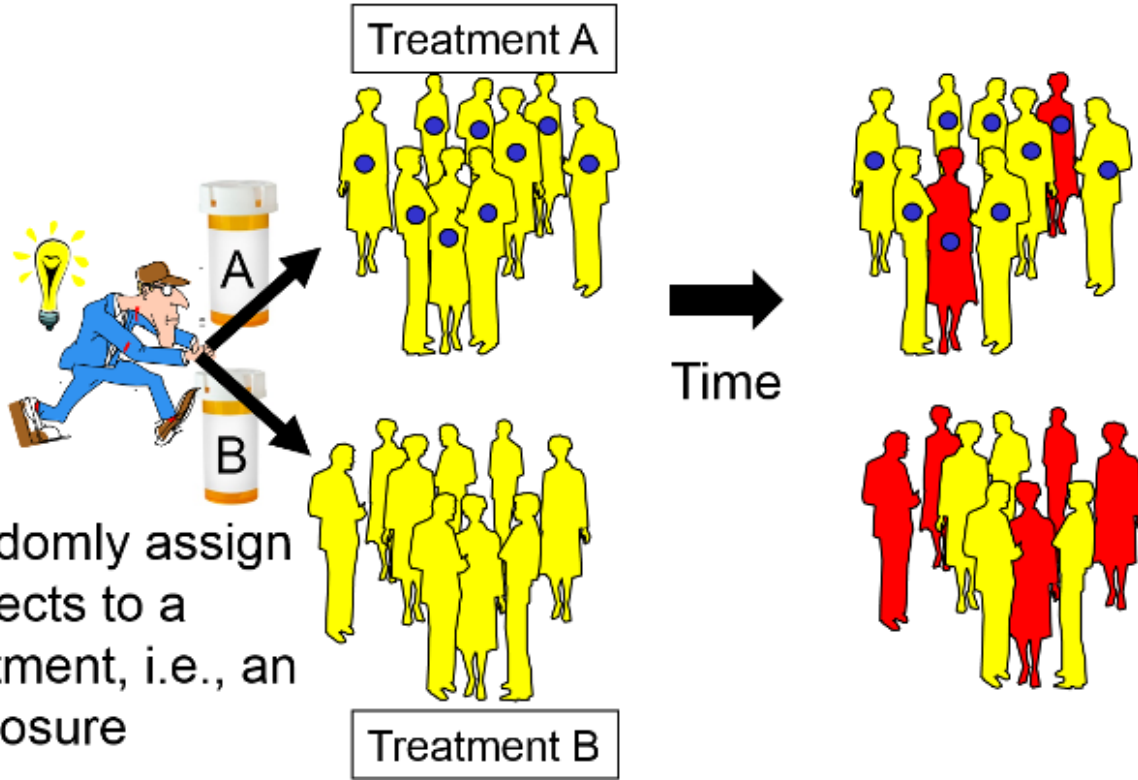


□ آزمایش تصادفی کنترل شده

راه حل طلایی:

- تقسیم تصادفی به دو گروه مستقل
- اجرای دو روش متفاوت به تفکیک هر گروه
- مقایسه نتایج دو گروه
- ارزیابی تفاوت عملکرد بر اساس تفاوت دو روش

Compare
Incidence



Randomly assign subjects to a treatment, i.e., an "exposure"

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه



□ استفاده از آزمون A/B در علوم اجتماعی

مثال: شیوه های ارتقای آموزش در مدارس هند (استر دوفلو، دانشگاه MIT)

مسئله: سطح سواد بسیار پایین دانش آموزان مقطع راهنمایی در روستاهای هند

متغیر مورد بررسی: میزان غیبت معلمین در مدارس

طراحی آزمایش:

- انتخاب تصادفی مدارس در دو گروه مجزا
- پرداخت حق الزحمه جداگانه به مدرسین گروه آزمایش بابت حضور در هر جلسه کلاس

نتیجه: افزایش معنادار موفقیت دانش آموزان در گروه آزمایش



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه



□ چالش های اجرایی آزمون A/B

○ هزینه های بالای اجرایی (نیاز به منابع مالی و انسانی)

○ زمان بر بودن اجرای طرح و پیگیری مداوم نتایج

○ اجرای سریع آزمایش ها

○ ارزان بودن اجرای آزمایش ها

○ در دسترس بودن نمونه های
آزمایش و نتایج آزمایش ها

دنیای دیجیتال



**جهان به مانند
یک آزمایشگاه**





□ کاربردهای آزمون A/B در بستر اینترنت

- انجام بیش از 100 هزار تست A/B بصورت سالانه در شرکت های گوگل و مایکروسافت
- انجام روزانه بیش از 1000 تست A/B در فیسبوک



برخی از مسائل رایج در طراحی آزمون A/B

- طراحی رابط کاربری و بهبود تجربه کاربر
- طراحی محتوای بازاریابی ایمیلی برای افزایش نرخ پاسخ
- قیمت گذاری سرویس ها و محصولات
- افزایش نرخ مشارکت در فعالیت های خیریه، سیاسی، اجتماعی

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه

□ کاربردهای آزمون A/B در بستر اینترنت

مثال: استفاده از آزمون A/B در طراحی کمپین تبلیغاتی باراک اوباما جهت جذب افراد بیشتر



- تغییر تصویر اوباما با تصویر خانواده
- تغییر گزینه SIGN UP با LEARN MORE


نتیجه:

- افزایش 40% در تعداد عضویت
- افزایش 60 میلیون دلار کمک مالی بیشتر

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



□ درک شهودی در مقابل انجام آزمون


مثال: انتخاب تیتر در سایت خبری بوستون گلوب بر اساس آزمون A/B

تیتر B	تیتر A
آیا این پهپاد می تواند نهنگ ها را نجات دهد؟	آیا پهپاد اسنات بات می تواند نهنگ ها را نجات دهد؟
این جستجوی برتر گوگل در ماساچوست واقعا خجالت آور است.	عجیب نیست که "توپ های کم باد" برترین عبارت جستجو شده در ماساچوست باشد.
زنی که از کارت های بیسبال نایاب 179000 دلار به دست آورد.	زنی که از کارت های بیسبال نایاب پول کلانی به جیب زد.

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 



□ درک شهودی در مقابل انجام آزمون

برنده	تیترا B	تیترا A
53% کلیک بیشتر روی A	آیا این پهپاد می تواند نهنگ ها را نجات دهد؟	آیا پهپاد اسنات بات می تواند نهنگ ها را نجات دهد؟
986% کلیک بیشتر روی B	این جستجوی برتر گوگل در ماساچوست واقعا خجالت آور است.	عجیب نیست که "توپ های کم باد" برترین عبارت جستجو شده در ماساچوست باشد.
38% کلیک بیشتر روی A	زنی که از کارت های بیسبال نایاب 179000 دلار به دست آورد.	زنی که از کارت های بیسبال نایاب پول کلانی به جیب زد.

مردم پیش بینی ناپذیرند!

تولید محتوا: زهرا ذوالقدر

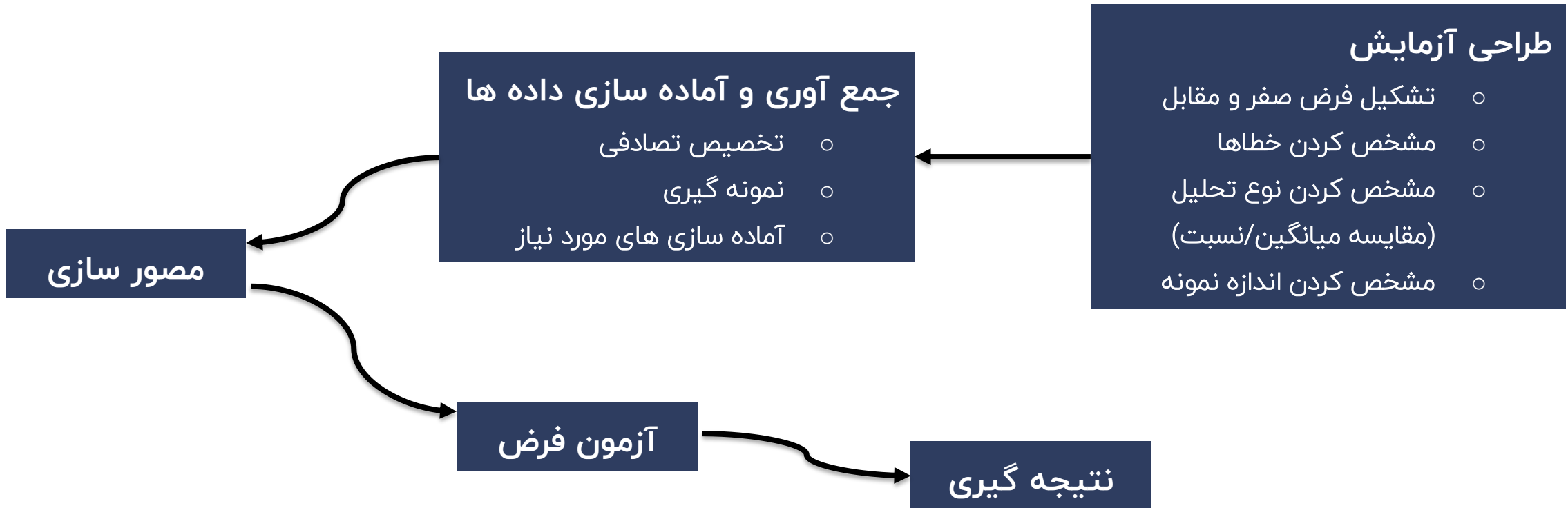
daychegroup

daychegroup

dayche.com | گروه دایکه



□ مراحل انجام آزمون A/B



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایکه



□ تعیین اندازه نمونه

نوع تحلیل: مقایسه نسبت گروه ها

Represents the **desired power** (typically .84 for 80% power).

A measure of **variability** (similar to standard deviation)

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2 [P_1(1-P_1) + P_2(1-P_2)]}{(P_1 - P_2)^2}$$

Represents the **desired level of statistical significance** (typically 1.96 for $\alpha = 0.05$).

Minimum meaningful difference or Effect Size

Sample size in each group (assumes equal sized groups)

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه



□ تعیین اندازه نمونه


نوع تحلیل: مقایسه میانگین گروه ها

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2 (2\sigma^2)}{(\mu_1 - \mu_2)^2}$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 



□ نمونه موردی

طرح مسئله: بیا بید تصور کنیم که شما در یک تیم تجارت الکترونیکی آنلاین با تیم محصول کار می کنید. طراح UX روی **نسخه جدید صفحه محصول** بسیار تلاش کرد، با این امید که منجر به **نرخ تبدیل (Conversion Rate)** بالاتری شود. مدیر محصول به شما گفته که نرخ تبدیل فعلی به طور متوسط حدود 13 درصد در طول سال است و تیم از افزایش 2 درصدی خوشحال خواهد شد، به این معنی که اگر طرح جدید منجر به نرخ تبدیل 15 درصدی شود، این طرح موفقیت آمیز تلقی می شود.

شما به عنوان تحلیلگر داده چگونه موفقیت طرح جدید را اندازه گیری می کنید؟

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

~
dayche.com | گروه دایکه



□ نمونه موردی

$$H_0: \rho_{Control} = \rho_{experiment}$$

$$H_a: \rho_{Control} \neq \rho_{experiment}$$

تشکیل فرض صفر و مقابل

راهکار: استفاده از آزمون A/B

نمایش همزمان طرح قدیم و جدید

برای دو گروه آزمایش و کنترل

$$\alpha = 0.05 , \beta = 0.2$$

مشخص کردن خطاها


آزمون استقلال یا مقایسه نسبت ها

مشخص کردن نوع تحلیل
(مقایسه میانگین/نسبت)

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



□ نمونه موردی

مشخص کردن اندازه نمونه

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2 [p_1(1-p_1) + p_2(1-p_2)]}{(p_1 - p_2)^2}$$

$$n = \frac{(1.96 + 0.84)^2 [0.13(1-0.13) + 0.15(1-0.15)]}{(0.13 - 0.15)^2} = 4714$$

در هر گروه

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 