

ZORG SOCIAL

دليل اختبارات A/B

للإعلانات المرئية المدعومة بالذكاء الاصطناعي

هيكلية اختبارات مقارنة الأساليب بدلالة إحصائية
أحجام العينات ومدة الاختبار وإطار التحليل والتحسين بالذكاء الاصطناعي

لماذا اختبار A/B للإعلانات المرئية؟

الحالة للتجريب المنظم في حملات الفيديو

لا يوجد أسلوب إعلاني واحد متفوق عالميًا. ما يؤثر في جمهور الرعاية الصحية في السعودية قد يفشل مع مستخدمي التكنولوجيا المالية في دبي. اختبار A/B يزيل التخمين بترك البيانات تحدد أي أسلوب إبداعي ونمط إنتاج ونبرة رسالة واستراتيجية منصة تحقق أفضل النتائج.

١.١ ما الذي يمكن اختباره في الإعلانات المرئية

متغير الاختبار	النسخة A (الضابطة)	النسخة B (المعالجة)
أسلوب الإعلان	عاطفي — قصة عائلية مؤثرة	إقناعي — عرض عائد مبني على بيانات
نمط الإنتاج	تصوير حي مع ممثلين	رسوم متحركة ثنائية
نبرة الرسالة	مهني وموثوق	دافئ وودي
الخطاف / الافتتاحية	البدء بسؤال مثير	البدء بإحصائية مفاجئة
مدة الفيديو	١٥ ثانية للخلاصات	٣٠ ثانية مع سرد أعمق
الدعوة للإجراء	«ابدا تجربة مجانية»	«احجز عرضًا توضيحيًا»
الصورة المصغرة	شخص ينظر للكاميرا مع نص	لقطة منتج مع بيانات النتائج
الموسيقى	موسيقى آلية مبهجة	تعليق صوتي بدون موسيقى
تنسيق المنصة	مربع 1:1 لفيديو	عمودي 9:16 لريلز فيسبوك
نمط الترجمة	شريط ترجمة سفلي	رسوم كلمة بكلمة متحركة

⚡ اختبار متغيرًا واحدًا في كل مرة. تغيير الأسلوب والموسيقى معًا يجعل من المستحيل تحديد سبب الفرق في الأداء. اعزل المتغير لعزل الرؤية.

١.٢ عائد الاختبار

- زيادة التحويل: اختبار A/B المنهجي يحقق معدلات تحويل أعلى بـ ٢٠-٤٩٪ مقارنة بالحملات غير المختبرة.
- كفاءة التكلفة: إلغاء النسخ ضعيفة الأداء ميكراً يوفر ١٥-٣٠٪ من الإنفاق الإعلاني.
- مكاسب مركبة: كل دورة اختبار تبني معرفة مؤسسية. عبر ١٢ شهرًا ترى العلامات تحسناً بـ ٣-٥ أضعاف.

القسم ٢ الأسس الإحصائية

الرياضيات وراء اتخاذ القرارات بثقة

فهم الدلالة الإحصائية يضمن أنك تتخذ قرارات بناءً على فروق أداء حقيقية — وليس ضوضاء عشوائية.

٢.١ المفاهيم الأساسية

المفهوم	ماذا يعني لاختبارك
الدلالة الإحصائية	احتمال أن الفرق الملاحظ حقيقي وليس صدفة. المعيار: ٥٪ فقط احتمال أن تكون النتيجة عشوائية ($p < 0.05$).
مستوى الثقة	كم تريد أن تكون متأكدًا قبل إعلان فائز. ٩٥٪ للتسويق العام. ٩٩٪ للقرارات الكبرى.
قيمة P	احتمال رؤية نتيجةك إذا لم يكن هناك فرق فعلي. $p < 0.05$ = دال إحصائيًا. $p < 0.01$ = دال جدًا.
أقل تأثير قابل للكشف (MDE)	أصغر تحسن تهتم باكتشافه. تحسن ١٪ يحتاج عينة ضخمة؛ ٢٠٪ يحتاج أقل بكثير.
القوة الإحصائية	احتمال أن يكتشف اختبارك فرقًا حقيقيًا إن وُجد. المعيار: ٨٠٪. أقل = خطر تفويت فائز حقيقي.

٢.٢ حساب حجم العينة

حجم العينة المطلوب يعتمد على: معدل التحويل الأساسي، وأقل تأثير قابل للكشف، ومستوى الثقة المطلوب.

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

معادلة حجم العينة (لكل نسخة)
حيث $Z_{\alpha/2} = 1.96$ (ثقة ٩٥٪)، $Z_{\beta} = 0.84$ (قوة ٨٠٪)، p_1 = المعدل الضابط، p_2 = المعدل المتوقع

مرجع سريع: أحجام العينات المطلوبة لكل نسخة

المعدل الأساسي	٥٠٪ رفع	١٠٪ رفع	٢٠٪ رفع	٥٠٪+ رفع
١٪ CTR	31,234	7,660	1,901	306
٢٪ CTR	15,366	3,784	949	157
٣٪ CTR	10,073	2,493	631	108
٥٪ تحويل	5,892	1,470	377	68
١٠٪ تحويل	2,758	698	183	36
٢٠٪ تحويل	1,230	317	86	19
٥٠٪ إكمال	384	104	31	9

بناءً على مستوى ثقة ٩٥٪، قوة ٨٠٪، اختبار ثنائي الطرف. محرك اختبار A/B في ZorgSocial بحسب المتطلبات الدقيقة بناءً على بيانات حملتك. المصطلح "Click-Through Rate : CTR" معدل النقر إلى الظهور

القسم ٣

إطار تصميم الاختبار

كيف تبني اختبار A/B صارمة للإعلانات المرئية

٣.١ عملية تصميم الاختبار في ٧ خطوات

#	الخطوة	الإجراءات والتفاصيل
١	حدد الفرضية	اكتب عبارة واضحة قابلة للاختبار: «تغيير الأسلوب من العاطفي إلى الإقناعي سيزيد طلبات العرض بـ ١٥٪ على الأقل»
٢	اختر المتغير	اختر متغيرًا واحدًا. الاختبارات الأولى الشائعة: الأسلوب، نمط الخطاف، مدة الفيديو، نص الدعوة
٣	حدد مؤشر الأداء الرئيسي	اختر مقياسًا واحدًا يحدد الفائز. المقاييس الثانوية للسباق فقط
٤	احسب حجم العينة	استخدم الجدول المرجعي أو حاسبة ZorgSocial. أدخل: المعدل الأساسي، MDE، ثقة ٩٥٪، قوة ٨٠٪
٥	حدد مدة الاختبار	الحد الأدنى ٧ أيام لمراعاة تباين أيام الأسبوع. الحد الأقصى ٢٨ يومًا
٦	التوزيع والإطلاق	قسّم الجمهور ٥٠/٥٠ عشوائيًا. النسختان تعملان متزامنتين بنفس الجداول والميزات
٧	التحليل والقرار	انتظر حتى يتم الوصول لحجم العينة. تحقق من الدلالة. وثّق النتائج. نفذ الفائز

٣.٢ إرشادات مدة الاختبار

الملاحظات اليومية (لكل نسخة)	تأثير صغير (٥٪)	تأثير متوسط (١٥٪)	تأثير كبير (٣٠٪)	التوصية
١,٠٠٠-٥٠٠	+٦٠ يوم	٣٥-٢١ يوم	١٤-١٠ يوم	اختبر التأثيرات الكبيرة فقط أو زد الميزانية
٥,٠٠٠-١٠,٠٠٠	٤٥-٣٠ يوم	٢١-١٠ يوم	٧-٥ أيام	مناسب لمعظم حملات الشركات الصغيرة
٢٠,٠٠٠-٥٠,٠٠٠	٢١-١٤ يوم	١٠-٧ أيام	٥-٣ أيام	سرعة اختبار مثالية
١٠٠,٠٠٠-٢٠٠,٠٠٠	١٠-٧ أيام	٥-٣ أيام	٢-١ يوم	سرعة مؤسسية
+١٠٠,٠٠٠	٥-٣ أيام	٢-١ يوم	> ٢٤ ساعة	تكرار سريع ممكن

⚡ لا تنتقد النتائج وتوقف مبكرًا عندما تبدو إحدى النسخ متقدمة. هذا يسمى «تحيز الاطلاع» ويضخم معدلات الإيجابية الكاذبة حتى ٣٠٪. حدد حجم العينة مسبقًا وانتظر.

القسم ٤

اختبارات مقارنة الأساليب

خطط اختبار جاهزة لمقارنة الأساليب العشرة

مقارنة الأساليب هي أعلى اختبار تأثيرًا لأنها تغير الاستراتيجية الإبداعية الأساسية. فيما يلي خطط اختبار جاهزة لأكثر المقارنات شيوعًا.

٤.١ هدف الوعي: العاطفي مقابل السرد

الفرضية	السرد سيغفوق على العاطفي في تذكر العلامة لأن البنية السردية تخلق ترميزًا أقوى في الذاكرة
النسخة A (الضابطة)	عاطفي: فيديو ٣٠ ثانية تصوير حي لحظة عائلية، موسيقى ناعمة، كشف لطيف عن العلامة
النسخة B (المعالجة)	سرد: فيديو ٣٠ ثانية مع بطل قابل للتعاطف، صراع، حل عبر العلامة. نفس الموسيقى والممثلين
مؤشر رئيسي	رفع تذكر العلامة التجارية
مؤشرات ثانوية	معدل الإكمال، المشاركات، الإعلام المكتسب، تحليل المشاعر
حجم العينة	٨,٠٠٠-١٢,٠٠٠ مشاهدة لكل نسخة
المدة	١٤ يومًا كحد أدنى
المنصات	يوتيوب (16:9)، إنستغرام ريلز (9:16)، فيسبوك (1:1)
الصناعات	غير ربحية، رعاية صحية، ضيافة، العلامة الوظيفية

٤.٢ توليد العملاء: الإقناعي مقابل العرض التوضيحي

الفرضية	العرض التوضيحي سينتج تسجيلات تجريبية أعلى لأن «الرؤية إيمان»، بينما الإقناعي سيولد طلبات عرض أكثر
النسخة A	إقناعي: رسوم متحركة مع بيانات ROI وإحصائيات. ينتهي بـ «أحجز عرضًا»
النسخة B	عرض توضيحي: تسجيل شاشة للميزات الرئيسية مع نتائج حقيقية. ينتهي بـ «ابدأ تجربة مجانية»
مؤشر رئيسي	معدل التحويل (طلبات عرض أو تسجيلات تجريبية)
حجم العينة	٥,٠٠٠-٨,٠٠٠ مشاهدة لكل نسخة
المنصات	لينكدان (16:9)، يوتيوب (16:9)، فيسبوك (1:1)
الصناعات	SaaS، خدمات مالية، تكنولوجيا، خدمات مهنية



إطار التحليل

كيف تقرأ النتائج وتعلن الفائزين وتستخلص الروى

٥.١ مصفوفة القرار

النتيجة	الدلالة	الإجراء
فوز B	$p < 0.05$	نفذ B كضابط جديد. وثق الأسلوب الفائز. وسع الإنفاق على B. أرشف A
فوز B	$p > 0.05$	النتيجة غير حاسمة. مدد الاختبار أو اختبر متغيرًا آخر
فوز A	$p < 0.05$	A يبقى الضابط — التغيير لم يحسن. سجل الرؤية. اختبر متغيرًا آخر
لا فرق	$p > 0.05$	المتغير المختبر لا يهم لهذا الجمهور. انتقل لمتغير أعلى تأثيرًا

٥.٢ الأخطاء الشائعة وكيفية تجنبها

الخطأ	ماذا يحدث	كيف تتجنب
تحيز الاطلاع	فحص النتائج يوميًا والتوقف عندما تبدو A جيدة يضحك الإيجابية الكاذبة إلى ٢٥-٣٠٪	التزم مسبقًا بحجم العينة وتاريخ الانتهاء. استخدم قفل ZorgSocial التلقائي
الاختبارات المتعددة	تشغيل ٥ اختبارات متزامنة على نفس الجمهور يجد «فائزين» بالصدفة	طبق تصحيح Bonferroni أو شغل الاختبارات تتابعيًا على شرائح منفصلة
تأثير الجدة	B تفوز مبدئيًا لأنها جديدة وليس لأنها أفضل. النتائج تتراجع بعد أسبوعين	شغل الاختبارات لـ ١٤-٢٨ يومًا. تحقق من انخفاض الأداء في الأسبوع الأخير
اختبارات ضعيفة القوة	اختبار ٣ أيام بـ ٥٠٠ مشاهدة. تحسن حقيقي بـ ٢٠٪ لا يُكتشف	احسب حجم العينة قبل الإطلاق. إن لم تستطع الوصول، زد MDE أو الميزانية

القسم ٦

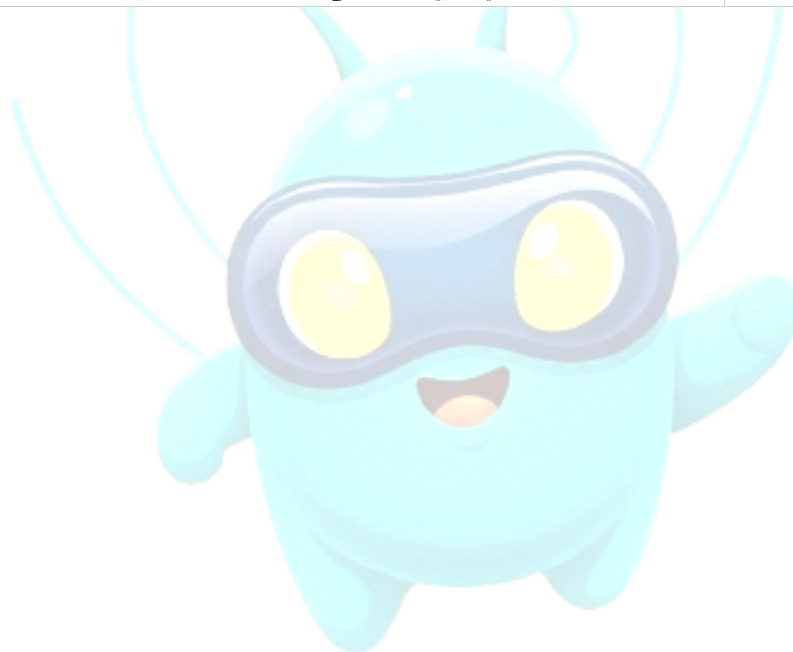
الاختبار المدعوم بالذكاء الاصطناعي مع ZorgSocial

كيف توثقت المنصة وتحسن سير عمل اختبار A/B

محرك اختبار A/B في ZorgSocial يتكامل مباشرة في سير عمل الحملة، مما يؤتمت العمل الإحصائي الثقيل ويسرع المسار من الفرضية إلى الرؤية.

٦.١ ميزات اختبار A/B في ZorgSocial

الميزة	كيف يعمل
توليد النسخ التلقائي	أدخل موجز الحملة مرة واحدة. الذكاء الاصطناعي يولد النسخة B تلقائيًا بتطبيق أسلوب بديل مع الحفاظ على باقي العناصر
حاسبة العينة الذكية	أدخل مقاييسك الأساسية وMDE. المحرك يحسب حجم العينة ويقدر الوقت ويوصي بتوزيع الميزانية
لوحة محمية من الاطلاع	النتائج مغلقة حتى الوصول لحجم العينة. لا تسرب بيانات وسيطة لتحفيز القرارات
وضع القاطع متعدد الأذرع	للحملات التي تهم فيها السرعة: الذكاء الاصطناعي يحول الحركة تدريجيًا للنسخة الأفضل أثناء الاختبار
الاختبار التتابعي	للحملات منخفضة الحركة: يسمح بالتوقف المبكر الصالح إذا كان التأثير كبيرًا جدًا
توليد التقارير التلقائي	تقرير تحليلي موحد بنقرة واحدة. قابل للتصدير ك PDF أو مشاركة داخل فريق العمل
سجل الاختبارات ومركز التعلم	جميع الاختبارات المكتملة تُسجل تلقائيًا. الذكاء الاصطناعي يكشف الأنماط عبر الاختبارات
الاختبار التنبؤي المسبق	قبل الإطلاق، الذكاء الاصطناعي يحاكي الأداء المتوقع بناءً على بيانات تاريخية مشابهة



الجزء ٧ خارطة طريق الاختبار لـ ١٢ شهرًا

إيقاع اختبار منظم التحسين المستمر

خارطة الطريق التالية توفر تسلسل اختبار موصى به للعلامات الجديدة على اختبار A/B المنظم. كل ربع يبني على تعلم الربع السابق.

الربع	مجال التركيز	الاختبارات	النتيجة المتوقعة	الرفع التراكمي
الربع ١	الأسلوب	عاطفي مقابل إقناعي عرض مقابل شهادات	تحديد الأسلوب الفائق لكل هدف	٪٢٥-١٥+
الربع ٢	العناصر الإبداعية	نمط الخطاب، نص CTA، المدة (١٥ مقابل ٣٠ ثانية)	تحسين داخل الأسلوب الفائق	٪٢٠-١٠+
الربع ٣	الإنتاج والمنصة	تصوير مقابل رسوم، مربع مقابل عمودي	النمط الأمثل لكل منصة	٪١٥-٨+
الربع ٤	التحسين المتقدم	تجزئة الجمهور، اختبار التوقيت، التكرار	الاستهداف الدقيق والتتابع	٪١٢-٥+

٧.١ التأثير المركب

اختبار A/B يحقق عوائد مركبة. إذا حقق الربع ١ تحسُّبًا بـ ٢٠٪، وأضاف الربع ٢ ١٥٪ فوق ذلك، والربع ٣ ١٢٪، والربع ٤ ٨٪، فالتحسن التراكمي ليس ٥٥٪ — بل ٧٧٪ مركبة.

معادلة التركيب

$$(1.20) \times (1.15) \times (1.12) \times (1.08) - 1 = 0.669 = 66.9\%$$

الرفع الإجمالي = (١ + ٢٠٪) × (١ + ١٥٪) × (١ + ١٢٪) × (١ + ٨٪) - ١

العلامات التي تلتزم ببرنامج اختبار منظم لمدة ١٢ شهرًا على ZorgSocial ترى عادةً تحسُّبًا بين ٥٠-٨٠٪ في مؤشرات أداء الفيديو الرئيسية.

القسم ٨

قوائم المراجعة والنماذج

أدوات جاهزة للتنفيذ الفوري

٨.١ قائمة ما قبل الإطلاق

☐	الفرضية مكتوبة كعبارة واضحة قابلة للاختبار مع الاتجاه والحجم المتوقع
☐	متغير واحد محدد — جميع العناصر الأخرى متطابقة بين النسختين
☐	مؤشر الأداء الرئيسي محدد ومتفق عليه من جميع الأطراف قبل الإطلاق
☐	حجم العينة محسوب (المعدل الأساسي، MDE، ثقة ٩٥٪، قوة ٨٠٪)
☐	مدة الاختبار محددة (الحد الأدنى ٧ أيام، الحد الأقصى ٢٨ يومًا)
☐	كلا الفيديوهين مُنتجان ومراجعان ومعتمدان بتكافؤ الجودة
☐	الاستهداف والميزانية والجدول وإعدادات المنصة متطابقة للنسختين
☐	التتبع والإسناد مُعد (UTM، بكسل التحويل، تحليلات ZorgSocial)
☐	الاختبار مسجل في سجل الاختبارات بمعايير نجاح محددة مسبقًا
☐	الأطراف المعنية على علم بأن النتائج لن تُشارك حتى اكتمال حجم العينة

٨.٢ قائمة ما بعد الاختبار

☐	حجم العينة المطلوب تم الوصول إليه لكلا النسختين
☐	الاختبار عمل للمدة الكاملة الملتزم بها (بدون توقف مبكر أو تعديلات)
☐	قيمة P محسوبة ومقارنة بعتبة 0.05
☐	النتائج مقسمة حسب المنصة وشريحة الجمهور ويوم الأسبوع ونوع الجهاز
☐	الفائز مُعلن بناءً على المؤشر الرئيسي فقط
☐	تقرير التحليل مُعد باستخدام النموذج الموحد
☐	الرؤية الرئيسية موثقة بلغة واضحة: «ماذا تعلمنا عن جمهورنا؟»
☐	الاختبار التالي محدد بناءً على هذه الرؤية
☐	سجل الاختبارات محدث: التاريخ، المتغير، النتيجة، حجم التأثير، الثقة، الرؤية، الإجراء التالي
☐	النسخة الفائزة موسعة كضابط جديد؛ الخاسرة مؤرشفة مع الدروس