الارتباط مقابل السببية

الارتباط يعني أن شيئين يميلان إلى التحرك معا، مثل وقت المذاكرة ودرجات الامتحان. أما السببية فتعني أن تغير أحدهما ينتج فعلا تغيرا في الآخر. وكثيرا ما يخلط بينهما، وقد يقود هذا الخلط إلى قرارات مكلفة. ونتابع دينا، المحللة التسويقية في شركة بيع بالتجزئة، وعليها أن تشرح ما تظهره بياناتها حقا.

1. ما يظهره الارتباط

يصف الارتباط كيف يتحرك مقياسان معا. في الارتباط الموجب يرتفعان معا، وفي الارتباط السالب يرتفع أحدهما بينما ينخفض الآخر، ومع غياب الارتباط لا يوجد نمط. وكثيرا ما يلخص الإحصائيون القوة برقم بين سالب واحد وموجب واحد، فالقيم القريبة من الطرفين تعني صلة قوية، والقيم القريبة من الصفر تعني صلة ضعيفة. وهذا الرقم يصف نمطا في البيانات، ولا يفسر سبب وجوده.

مثال: يبيع مقهى حساء ساخنا أقل كلما ارتفعت درجة الحرارة في الخارج: كلما زاد أحدهما نقص الآخر، وهذا ارتباط سالب. والنمط واضح، ومن السهل فهم سببه.

في العمل: تجد دينا أن الأسابيع ذات الإنفاق الإعلاني الأكبر تشهد مبيعات أعلى أيضا، بارتباط يبلغ نحو ثمانية أعشار. وهذه صلة موجبة قوية، لكنها لا تعرف بعد لماذا يتحركان معا.

2. لماذا لا يكفي الارتباط

هناك ثلاثة أسباب شائعة لأن الارتباط قد لا يعني السببية. الأول هو الصدفة: فمع بيانات كثيرة تتحرك أشياء غير متصلة معا بالمصادفة. والثاني هو عامل ثالث خفي، يسمى المتغير المشوش، يؤثر في الشيئين معا. والثالث هو الاتجاه العكسي، حيث يكون الأثر هو ما يسبب السبب المزعوم. ويمكن للارتباط القوي أن يحمل أيا من هذه التفسيرات، لذلك هو نقطة بداية للأسئلة فقط.

مثال: ترتفع مبيعات المثلجات وحالات الغرق معا في الصيف، لكن المثلجات لا تسبب الغرق. فالطقس الحار هو العامل الثالث الخفي الذي يزيد الاثنين.

في العمل: تلحظ دينا أن العملاء الذين يفتحون رسائل كثيرة ينفقون أكثر أيضا، فتتساءل هل تسبب الرسائل الإنفاق الزائد. وقد يكون العامل الثالث الخفي هو الولاء: فالعملاء المخلصون يفتحون الرسائل ويشترون أكثر. والارتباط وحده لا يستطيع التمييز بين الاحتمالين.

حقوق الطبع والنشر محمية © - أكاديمية أساسيات (شركة أساسيات للخدمات الاستشارية)

3. كيف نختبر السبب

أقوى طريقة لاختبار السببية هي التجربة. نقسم الناس عشوائيا إلى مجموعتين، ونطبق التغيير على إحداهما، ونقارن النتائج. والتوزيع العشوائي يجعل المجموعتين متشابهتين في كل ما عدا ذلك، فيكون الفرق أقرب إلى أن ينتج عن التغيير. وعندما لا تمكن التجربة، تعطي المقارنات الدقيقة والتوقيت وآلية معقولة ونتائج متكررة أدلة أضعف لكنها مفيدة.

مثال: لاختبار فاعلية سماد جديد، يعالج بستاني نصف النباتات المختارة عشوائيا ويترك النصف الآخر دون معالجة. فإذا نما النصف المعالج أفضل في الظروف نفسها، يكون السماد سببا محتملا.

في العمل: تجري دينا اختبار A/B: يتلقى نصف عشوائي من العملاء الرسالة ولا يتلقاها النصف الآخر. وتشتري المجموعة التي وصلتها الرسالة أكثر من المجموعة التي لم تصلها. وبما أن المجموعتين اختيرتا عشوائيا، تستطيع القول إن الرسالة ربما ساعدت.

4. التقرير بحذر

الصياغة الحذرة تحمي القرارات. نستطيع القول إن شيئين مرتبطان أو بينهما صلة، وينبغي أن نحتفظ بكلمات مثل «يسبب» أو «يؤدي إلى» للحالات ذات الأدلة القوية. ويذكر التقرير الجيد أيضا العوامل الخفية المحتملة وحدود البيانات. وأخيرا، يقترح الخطوة التالية، مثل اختبار، قبل أن يعتمد المال أو السياسة على النتيجة.

مثال: يقول عنوان إن من يتناولون الفطور يتفوقون في المدرسة. وتقول الصياغة الحذرة إن الفطور ونتائج المدرسة بينهما صلة، وإن عادات الأسرة قد يكون لها دور.

في العمل: تكتب دينا في تقريرها أن من يفتحون الرسائل ينفقون أكثر، وأن هذه صلة لا دليل. وتذكر الولاء عاملا خفيا محتملا وتعرض نتيجة الاختبار. فيمول المدير خطة الرسائل بثقة أكبر.

الخاتمة

وخلاصة القول، الارتباط يظهر أن شيئين يتحركان معا، أما السببية فتعني أن أحدهما ينتج الآخر. ويطبق ذلك في أربع خطوات: رؤية النمط، والتشكيك فيه، واختبار السبب، والتقرير بحذر. وبهذا تقود البيانات إلى قرارات سليمة بدل التخمين المكلف.

المصادر

• أوستن برادفورد هيل (1965): The Environment and Disease: Association or Causation؟ مجلة Proceedings of the Royal Society of Medicine – https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/
• Correlation does not imply causation: نظرة عامة على الصدفة والعوامل المشوشة والاتجاه العكسي (ويكيبيديا) – https://en.wikipedia.org/wiki/Correlation_does_not_imply_causation
• Confounding: كيف يصنع عامل ثالث خفي صلة زائفة (ويكيبيديا) – https://en.wikipedia.org/wiki/Confounding

حقوق الطبع والنشر محمية © - أكاديمية أساسيات (شركة أساسيات للخدمات الاستشارية)