

Kłamstwo dla początkujących

Przyjrzyjmy się teraz podstawom posługiwania się danymi w mylący sposób. Pewnie niektóre opisane tu sposoby są dość oczywiste, ale nie zaszkodzi przypomnieć, co może pójść nie tak, gdy ludzie posługują się danymi na poparcie swoich argumentów.

ZACIERANIE GRANIC MIĘDZY KORELACJĄ A ZWIĄZKIEM PRZYCZYNOWYM

Jeżeli dwa zdarzenia (nazwijmy je A i B) są ze sobą skorelowane, to znaczy, że z czasem zmiany ich wartości zachodzą w tym samym kierunku. Jeżeli rośnie A , rośnie też B , a jeżeli A maleje, maleje też B . Istnieją cztery możliwe wyjaśnienia tej prawidłowości:

1. A powoduje B .
2. B powoduje A .
3. Trzeci czynnik, C , powoduje A i B .
4. Między A i B nie występuje żadna zależność przyczynowa.

Tak naprawdę to nie jest wyczerpująca lista, istnieją również bardziej skomplikowane możliwości, takie jak „A powoduje C, a C powoduje B”. Na potrzeby naszego przykładu wystarczą jednak cztery najprostsze kategorie.

Czwarte wyjaśnienie występuje na przykład w sytuacji, w której obserwacja nastąpiła w krótkim okresie i korelacja A i B wynika ze zwykłego przypadku. Pierwszy wniosek jest zatem taki, że jeśli analizujemy dane z krótkiego okresu, w ogóle nie ma sensu zaczynać szukać związków przyczynowych.

Zawsze warto natomiast zastanowić się nad trzecim wyjaśnieniem. Weźmy na przykład stwierdzenie, że łysienie powoduje choroby naczyniowe. Te dwa zjawiska mogą być ze sobą skorelowane, ale istnieje znacznie większe prawdopodobieństwo, że oba są powodowane przez trzeci czynnik, czyli starość, która znacząco zwiększa ryzyko występowania obu.

Ogólnie wygląda to tak, że gdy stwierdzimy występowanie korelacji między zdarzeniami A i B , jesteśmy skłonni zakładać, że mamy do czynienia z wyjaśnieniem pierwszym lub drugim. *Spurious Correlations*, znakomita książka i strona internetowa Tylera Vigena, opisuje zabawne przypadki par zjawisk, które nie mogą być powiązane przyczynowo, a mimo to są skorelowane. Na przykład w latach 1999–2008 wiek pań wybieranych jako Miss America wykazywał dość bliską korelację z morderstwami za pomocą pary wodnej, oparów i gorących przedmiotów. W podobnym okresie amerykański import ropy naftowej z Norwegii wykazywał korelację z liczbą kierowców, którzy zginęli w wypadkach z udziałem pociągów.

W tego rodzaju przypadkach na pierwszy rzut oka widać, że korelacja ma charakter przypadkowy, ponieważ trudno wyobrazić sobie jakikolwiek związek między tymi zdarzeniami. Mózg człowieka jest naturalnie zaprogramowany na poszukiwanie prawdziwości,

więc gdy mamy do czynienia z dwoma zjawiskami, w przypadku których można wyobrazić sobie zależność przyczynową, bardzo chętnie zakładamy jej występowanie.

W jaki sposób kłamcy mogą wykorzystywać tę naturalną dla człowieka skłonność? Weźmy na przykład słynną reklamę papierosów Camel z 1946 roku, w której zapewniano, że „Camele to najpopularniejsze papierosy wśród lekarzy”. Były to lata, kiedy zaczynaliśmy rozumieć, że palenie może być dla nas niezbyt korzystne, a zatem autorzy tej reklamy postanowili wykorzystać powszechne przekonanie, że lekarze wiedzą, co jest zdrowe, więc palą najzdrowsze papierosy. Tym samym reklama sugerowała występowanie zależności przyczynowej między „wiedzą o tym, co jest zdrowe” a „paleniem papierosów Camel”.

Zacznijmy od tego, że nie ma żadnych powodów, by sądzić, że lekarze jako eksperci od zdrowia sami dokonują zdrowych wyborów. Być może prawdą jest, że lekarze częściej sięgają po amfetaminę, niż wynosi średnia dla społeczeństwa, ale jeśli rzeczywiście tak jest, najprawdopodobniej wynika to z faktu, że ich praca jest obciążająca i stresująca, a nie z przekonania, że amfetamina jest zdrowa.

Tak się składa, że w tym przypadku podstęp miał charakter dwuwymiarowy. Badania, na których opierało się to zapewnienie, zostały przeprowadzone przez William Esty & Co., agencję reklamową zatrudnioną przez koncern tytoniowy R.J. Reynolds. Zastosowana metodologia nie należała do najrzetelniejszych: pracownicy agencji rozmawiali z lekarzami na konferencjach medycznych albo w ich gabinetach, a każdy wywiad zaczynali od wręczenia im kartonu Cameli.

Gdyby nawet jednak stwierdzenie opierało się na rzetelnych badaniach, nie byłoby żadnych powodów, żeby sądzić, że w tym przypadku korelacja oznacza związek przyczynowy.