INSIGHTS · GENERATING INSIGHT

# The Contribution of Outlier Cases and Edge Examples to Novel Insights

Why outliers, anomalies and extreme cases are a valuable source of new insight, how lead users and positive deviants reveal opportunities, and how analysts can tell meaningful outliers from errors.

ModQu research article · 984 words · 10 references (APA 7th edition)

Analysis usually focuses on the typical case: the average customer, the most common complaint, the central trend. Outliers, the cases that do not fit, are often treated as noise to be removed. Yet some of the most valuable insights come from exactly these cases. An unusual customer may be using a product in a way that points to a new market. A store that performs far better than similar stores may have discovered a practice others could copy. A data point that contradicts the expected pattern may reveal that the pattern itself is wrong. This article explains why outliers matter for insight and how to examine them productively.

## Anomalies as the start of new understanding

Kuhn (1962) argued that major changes in scientific understanding begin with anomalies, observations that do not fit the prevailing framework. Dunbar (1995) studied scientists at work in molecular biology laboratories and found that unexpected findings were common, and that how scientists responded to them was a key difference between laboratories. Productive laboratories took surprising results seriously, discussed them as a group and used analogies to related findings to explain them.

Klein (2013) studied over a hundred cases of insight in business, science and other fields and found that one common path to insight was noticing a contradiction: something that did not fit with what a person expected. People who gained insight did not dismiss the inconsistency but investigated it.

## Lead users and extreme customers

Some customers are outliers because they are ahead of the market. Von Hippel (1986) described lead users: customers who face needs months or years before most of the market does and who expect to benefit significantly from solving them. Lead users often modify existing products or build their own. Lilien et al. (2002) studied the use of lead-user methods at 3M and found that projects based on lead-user ideas had substantially higher projected sales than projects developed by conventional methods.

Christensen (1997) observed that established firms often focus on their mainstream customers and miss the needs of customers at the edge of the market, where disruptive innovations tend to begin. Customers who use a product in unusual ways, or who are poorly served by it, may reveal opportunities that average customers do not.

## Positive deviants

Another kind of outlier is the individual or group that achieves much better results than others facing the same conditions. Pascale et al. (2010) described the positive deviance approach, in which communities and organisations identify such cases and study what they do differently. Because the successful practice already exists within the same constraints, it is more likely to be adopted by others than a solution imported from outside.

For insight teams, this suggests looking deliberately for the best-performing stores, sales teams, regions or customers and asking what they do that others do not.

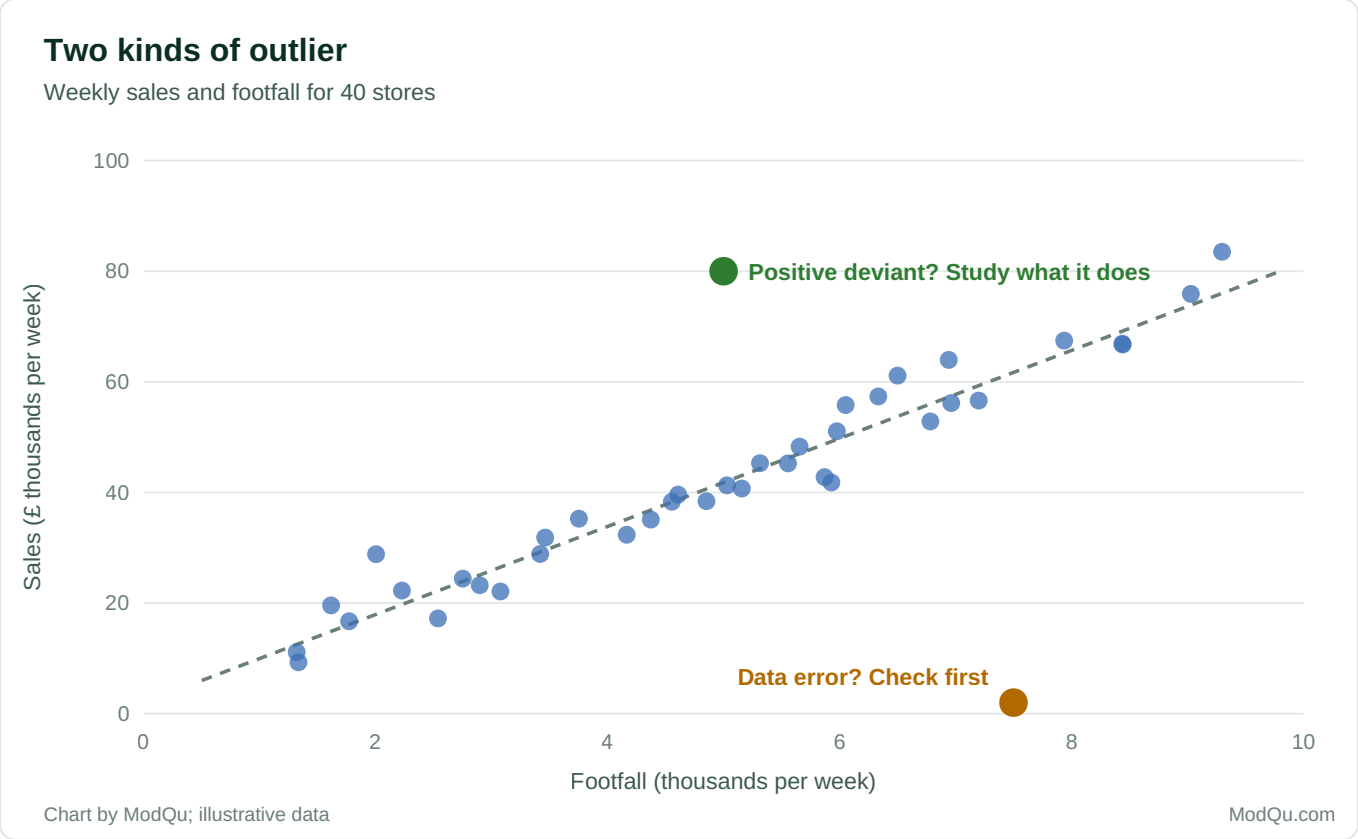



Figure 1. Two kinds of outlier: one worth studying, one to check. Chart by ModQu; illustrative data.

## Telling signals from errors

Not every outlier is meaningful. Many are errors in measurement, recording or processing. Hodge and Austin (2004) and Chandola et al. (2009) reviewed statistical methods for detecting outliers and anomalies, which can identify unusual cases but cannot say whether they matter. That requires judgement and further investigation.

A practical sequence helps. First, check whether the case could be an error: an impossible value, a duplicate record or a change in how data were collected. Second, describe the case in detail, including its context. Third, ask what it would mean if it were real: what would have to be true for this case to make sense? Fourth, look for other cases like it. Finally, if the outlier suggests a new explanation, test it.

Flyvbjerg (2006) argued that extreme or deviant cases are often more informative than typical ones, because they reveal more about the mechanisms at work. A single well-chosen case can challenge a general claim, much as one black swan disproves the claim that all swans are white.

Outliers also matter in qualitative research. When most interview participants describe a product in similar terms, the one who describes it very differently deserves attention rather than dismissal. That person may belong to a segment the study did not expect, may use the product in a different situation or may have noticed a problem others have not yet encountered. Recording and revisiting such cases, instead of averaging them away in a summary of common themes, keeps open the possibility that they point to something new. In quantitative work, a useful habit is to keep a short log of every outlier removed from an analysis, with the reason for its removal, so that the decision can be reviewed later if the pattern changes.

## An example

A bank analysing mortgage applications noticed that one branch had an approval rate in line with others but a default rate far lower. The first assumption was an error in the data, but the records were correct. Interviews revealed that the branch manager had a habit of calling applicants a week after approval to explain the first statement and payment schedule. Customers who understood their repayments were less likely to fall behind. The bank tested the practice in several branches and later adopted it more widely.

## Common mistakes

Common mistakes include deleting outliers automatically before analysis, which removes the cases most likely to reveal something new; treating every outlier as a discovery without checking for errors; and generalising from a single unusual case without looking for others. Another is focusing only on negative outliers, such as failing stores, and ignoring positive ones that may hold practical solutions.

## Conclusion

Outliers are where established patterns break down, which makes them a natural starting point for new understanding (Kuhn, 1962; Klein, 2013). Lead users, positive deviants and unexpected data points can reveal needs, practices and mechanisms that averages hide. Insight teams that check outliers carefully, rather than discarding them, give themselves more chances to discover something new.

## References

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), Article 15. https://doi.org/10.1145/1541880.1541882

Christensen, C. M. (1997). *The innovator's dilemma: When new technologies cause great firms to fail*. Harvard Business School Press.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365–395). MIT Press.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry, 12*(2), 219–245. https://doi.org/10.1177/1077800405284363

Hodge, V. J., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review, 22*(2), 85–126. https://doi.org/10.1007/s10462-004-4304-y

Klein, G. (2013). *Seeing what others don't: The remarkable ways we gain insights*. PublicAffairs.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Lilien, G. L., Morrison, P. D., Searls, K., Sonnack, M., & von Hippel, E. (2002). Performance assessment of the lead user idea-generation process for new product development. *Management Science, 48*(8), 1042–1059. https://doi.org/10.1287/mnsc.48.8.1042.171

Pascale, R., Sternin, J., & Sternin, M. (2010). *The power of positive deviance: How unlikely innovators solve the world's toughest problems*. Harvard Business Press.

von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science, 32*(7), 791–805. https://doi.org/10.1287/mnsc.32.7.791