Management and Organization Review

Management and Organization Review 14:2, June 2018, 433–437
doi: 10.1017/mor.2018.19

Dialogue, Debate, and Discussion

Show Us Your Data: Connect the Dots, Improve Science

Sheen S. Levine
The University of Texas, Dallas

'The truth is under attack', I wrote earlier this decade (Levine, 2012). As the replication crisis became apparent, the alarm was timely. But now, a counter-attack is raging. In its arsenal are replications, open data, shared instruments, pre-registration of hypotheses and now – data visualizations.

'You can observe a lot by watching', reflected Yogi Berra, and scientists knew that all along (or at least, should have known). More than 40 years ago, writing in *The American Statistician*, F. J. Anscombe (1973) warned us that numbers are not substitutes for figures. 'Graphs are essential to good statistical analysis', he concluded, because graphs validate the assumptions underlying the analysis.

Quite naturally, Anscombe illustrated theory with a graph. He devised four data sets. To most regression methods, they would appear identical: the same means for X's and Y's, same standard deviations, same correlations, same measures of explained variance (R^2), and same regression line (3+0.5x).

Limit yourself to regression, and the data are ostensibly identical. Add a graph, and their vast differences become apparent (Figure 1). One is exactly what you'd imagine from the regression line, but the three others tell a radically different story, empirically and hence theoretically. But without a graph, you would never know that.

The Anscombe Four have recently been expanded to an almost-comical effect. The data could be visibly clustered around certain values, seemingly randomly distributed, or even be shaped like a beast – but you could never tell by the summary statistics (Figure 2).

Graphs can test assumptions. They help diagnose data. Their utility in research, from beginning to end, has been known for decades (Healy & Moody, 2014). So why do they remain rare? I can speculate: Perhaps we have become captive by a sociological institution telling us that 'proper' science precludes graphs.[1] Or maybe we are confined by what game theoreticians call a coordination game, in which anyone who diverges from convention is penalized, even when correct.

© 2018 The International Association for Chinese Management Research

https://doi.org/10.1017/mor.2018.19 Published online by Cambridge University Press

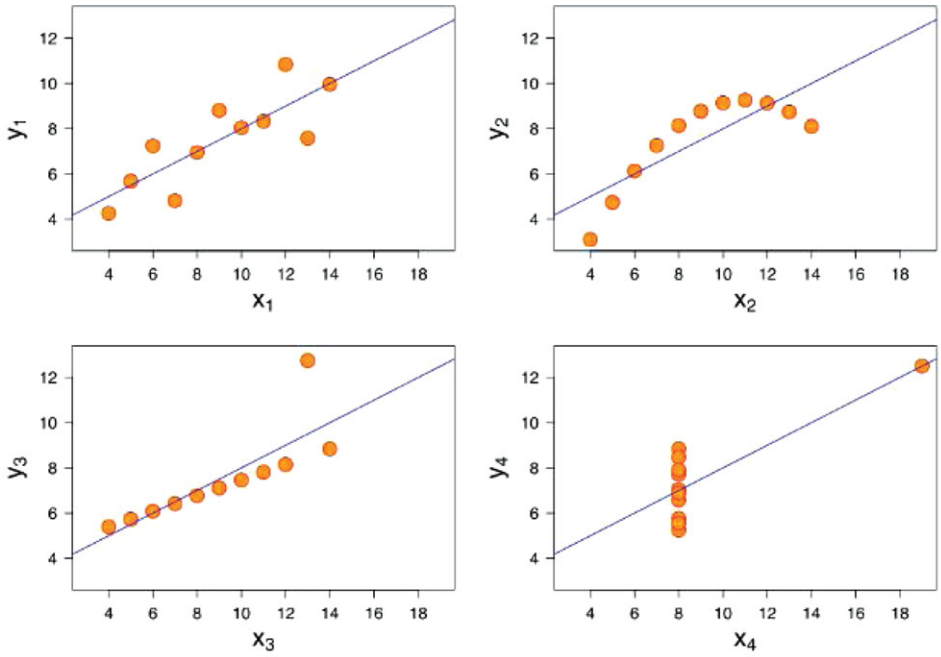


Figure 1. (Color online) Anscombe's (1973) quartet. The four datasets are visibly different but possess identical summary statistics.

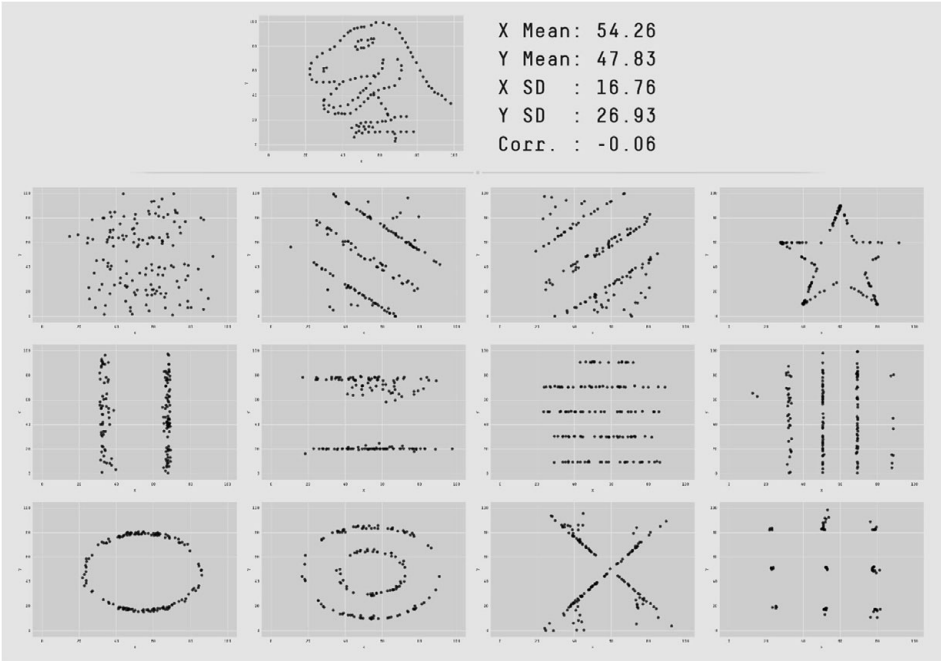


Figure 2. The Datasaurus Dozen (Matejka & Fitzmaurice, 2017), an extension of Anscombe's quartet. All of them have identical descriptive statistics (yes, even the Tyrannosaurus).

© 2018 The International Association for Chinese Management Research

https://doi.org/10.1017/mor.2018.19 Published online by Cambridge University Press

Whatever the cause, the way forward is clear: disrupt the institution, change the game. And this is what researchers, editors, and most recently journals have been doing about other damaging conventions. Take, for instance, statistical significance or hypothesis testing: Statisticians have been warning for years that p-value is not a measure of hypothesis veracity (e.g., Gelman & Stern, 2006; Wasserstein & Lazar, 2016), even if some mistakenly believe so (McShane & Gal, 2015). A turning point came when two studies showed how popular research decisions can, intentionally or not, cause anything to become statistically significant (John, Loewenstein, & Prelec, 2012; Simmons, Nelson, & Simonsohn, 2011). Psychologists heeded the problematic evidence by bolstering editorial policies (e.g., Eich, 2014), adopting 'new statistics' (Cumming, 2012, 2014), and commencing a vast and ongoing effort to replicate findings (e.g., Open Science Collaboration, 2015).[2] Curatescience.org was established to coordinate replications and document their results. Institutions began reforming.

Management scholars, already wary (e.g., Bettis, 2012), took notice. The *Strategic Management Journal* (SMJ) moved to discourage talk of 'statistically significant' findings and encourage replication (Bettis, Ethiraj, Gambardella, Helfat, & Mitchell, 2016; Ethiraj, Gambardella, & Helfat, 2016). Shortly thereafter, *Management and Organization Review* (MOR) extended these guidelines by adding incentives: Researchers who share data and instruments deserve recognition, and they receive it in the form of badges (Lewin et al., 2016). At MOR, recognition also goes to those who preregister their hypotheses, signaling that a counterintuitive finding was genuine, not the result of hypothesizing after the results were known (HARKing; Kerr, 1998). Soon thereafter, SMJ also began awarding badges for open data (full disclosure: this commentator's paper received such a badge, Levine, Bernard, & Nagel, 2017).

VISUALIZING THE FUTURE

Henrich Greve counts the challenges we face – statistically significant but weak findings, vast unexplained variation, and proprietary data that blocks anyone from interpreting the findings, except the original authors.[3] He shows the risk – static science – and offers an easy countermeasure. Graphs, while not resolving all these challenges, are a necessary addition to our shared grammar. And graphing is much easier than it was.

Whereas Anscombe (1973: 20) had to contend with a 'line printer or typewriter terminal', the contemporary, open-source statistical package R offers anything from a simple scatterplot to complex creations in two and three dimensions. Users have been adding to it through open collaboration: offering detailed examples, sharing statistical code, providing support (Levine & Prietula, 2014). With such expansive possibilities, graphing is within anyone's reach.

Now, the challenge is not statistical or technological, but institutional: Who would dare to doodle? Greve's journal, *Administration Science Quarterly* (ASQ), is set

© 2018 The International Association for Chinese Management Research

https://doi.org/10.1017/mor.2018.19 Published online by Cambridge University Press

to make visuals legit again. Each of us can contribute, at ASQ and anywhere else:

1. Open your eyes to the wealth of graphing possibilities (Healy & Moody, 2014) and understand how to construct, characterize, and interpret graphs (Cook & Weisberg, 1999).
2. As an author, include in your manuscript graphs such as scatterplot matrixes, residual plots, and Cook's distance (Cook, 1977). Explain to your readers how the graphs aid exploration, improve evidence, encourage extensions, and promote collaboration and accumulation of evidence. As a reviewer, communicate why you expect to see graphs.
3. Train peers and students in graphing and interpretation thereof, regardless of method or approach (Healy, Forthcoming).[4] Graphs have an important place in empirical papers, but they can also play a useful role in visualizing theory (e.g., Levine & Prietula, 2012; Levinthal & Posen, 2007).

There is a light at the end of the tunnel. I can see it.

NOTES

For advice, I am thankful to Paul Alison of the Sociology Department and Paul Rosenbaum of the Statistics Department, both the University of Pennsylvania, and to Kieran Healy of the Sociology Department at Duke University. Conversations with Rosemarie Nagel of the Barcelona Graduate School of Economics inspired the heart of this commentary.

[1] As Andrew Gelman (2009) sarcastically instructed, 'if you've found something great and you want to share it with the world, sure, make a pretty graph and put it on a blog. But please, please, keep these abominations out of our scientific journals'.

[2] In economics, similar replication projects showed that many findings failed to replicate (e.g., Camerer et al., 2016; Chang & Li, 2017).

[3] Kieran Healy and James Moody (2014: 108) see a link between the paucity of graphs and the dominance of proprietary data.

[4] For a draft, see http://socviz.co/

REFERENCES

Anscombe, F. J. 1973. Graphs in statistical analysis. ***American Statistician***, 27(1): 17–21. doi: 10.1080/00031305.1973.10478966

Bettis, R. A. 2012. The search for asterisks: Compromised statistical tests and flawed theories. ***Strategic Management Journal***, 33(1): 108–113. doi: 10.1002/smj.975

Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. 2016. Creating repeatable cumulative knowledge in strategic management. ***Strategic Management Journal***, 37(2): 257–261. doi: 10.1002/smj.2477

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., . . . Wu, H. 2016. Evaluating replicability of laboratory experiments in economics. ***Science***, 351(6280): 1433–1436. doi: 10.1126/science.aaf0918

Chang, A. C., & Li, P. 2017. A preanalysis plan to replicate sixty economics research papers that worked half of the time. ***American Economic Review***, 107(5): 60–64. doi: 10.1257/aer.p20171034

Cook, R. D. 1977. Detection of influential observation in linear regression. ***Technometrics***, 19(1): 15–18. doi: 10.1080/00401706.1977.10489493

© 2018 The International Association for Chinese Management Research

https://doi.org/10.1017/mor.2018.19 Published online by Cambridge University Press

Cook, R. D., & Weisberg, S. 1999. Graphs in statistical analysis: Is the medium the message? ***American Statistician***, 53(1): 29–37. doi: 10.1080/00031305.1999.10474426

Cumming, G. 2012. ***Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis***. New York: Routledge.

Cumming, G. 2014. The new statistics: Why and how. ***Psychological Science***, 25(1): 7–29. doi: 10.1177/0956797613504966

Eich, E. 2014. Business not as usual. ***Psychological Science***, 25(1): 3–6. doi: 10.1177/0956797613512465

Ethiraj, S. K., Gambardella, A., & Helfat, C. E. 2016. Replication in strategic management. ***Strategic Management Journal***, 37(11): 2191–2192. doi: 10.1002/smj.2581

Gelman, A. 2009. Why tables are really much better than graphs. Retrieved from http://andrewgelman.com/2009/04/01/why_tables_are/

Gelman, A., & Stern, H. 2006. The difference between 'significant' and 'not significant' is not itself statistically significant. ***American Statistician***, 60(4): 328–331. doi: 10.1198/000313006X152649

Healy, K. Forthcoming. ***Data visualization: A practical introduction***. Princeton, NJ: Princeton University Press.

Healy, K., & Moody, J. 2014. Data visualization in sociology. ***Annual Review of Sociology***, 40(1): 105–128. doi: 10.1146/annurev-soc-071312-145551

John, L. K., Loewenstein, G., & Prelec, D. 2012. Measuring the prevalence of questionable research practices with incentives for truth-telling. ***Psychological Science***, 23(5): 524–532.

Kerr, N. L. 1998. Harking: Hypothesizing after the results are known. ***Personality and Social Psychology Review***, 2(3): 196–217. doi: 10.1207/s15327957pspr0203_4

Levine, S. S. 2012. Walter R. Nord and Ann F. Connell: Rethinking the knowledge controversy in organization studies: A generative uncertainty perspective. ***Administrative Science Quarterly***, 57(3): 537–540. doi: 10.1177/0001839212462542

Levine, S. S., Bernard, M., & Nagel, R. 2017. Strategic intelligence: The cognitive capability to anticipate competitor behavior. ***Strategic Management Journal***, 38(12): 2390–2423. doi: 10.1002/smj.2660

Levine, S. S., & Prietula, M. J. 2012. How knowledge transfer impacts performance: A multilevel model of benefits and liabilities. ***Organization Science***, 23(6): 1748–1766. doi: 10.1287/orsc.1110.0697

Levine, S. S., & Prietula, M. J. 2014. Open collaboration for innovation: Principles and performance. ***Organization Science***, 25(5): 1414–1433. doi: 10.1287/orsc.2013.0872

Levinthal, D., & Posen, H. E. 2007. Myopia of selection: Does organizational adaptation limit the efficacy of population selection? ***Administrative Science Quarterly***, 52(4): 586–620.

Lewin, A. Y., Chiu, C.-Y., Fey, C. F., Levine, S. S., McDermott, G., Murmann, J. P., & Tsang, E. 2016. The critique of empirical social science: New policies at management and organization review. ***Management and Organization Review***, 12(4): 649–658.

Matejka, J., & Fitzmaurice, G. 2017. *Same stats, different graphs: Generating datasets with varied appearance and identical statistics through simulated annealing*. Paper presented at the CHI 2017, Denver, CO, USA.

McShane, B. B., & Gal, D. 2015. Blinding us to the obvious? The effect of statistical training on the evaluation of evidence. ***Management Science***. doi:10.1287/mnsc.2015.2212

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. ***Science***, 349(6251): aac4716. doi: 10.1126/science.aac4716

Simmons, J. P., Nelson, L. D., & Simonsohn, U. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. ***Psychological Science***, 22(11): 1359–1366.

Wasserstein, R. L., & Lazar, N. A. 2016. The American Statistical Association statement on p-values: Context, process, and purpose. ***American Statistician***, 70(2): 129–133. doi: 10.1080/00031305.2016.1154108

© 2018 The International Association for Chinese Management Research

https://doi.org/10.1017/mor.2018.19 Published online by Cambridge University Press