Annals of Science and Inquiry
Vol. 2 No. 6 (2026)

A Rising Yes Share Can Hide Crossing Community-Membership Distributions

Derek Knight[1] · Wesley Stewart[2] · Ryan King[3]

[1] Sweden · [2] Denmark · [3] Norway

Published: June 17, 2026 | Article 855

Abstract

Among the most frequent Stack Overflow participants, stronger endorsement of community membership coexists with a larger fraction choosing the most negative response. We compare every ordinal cutpoint across six participation groups in 44,718 public survey records. Three adjacent comparisons show upward ordering in their observed distributions; only two meet a simultaneous working-interval criterion. The final two comparisons cross, although their combined Yes shares increase. Allowing every observed Not sure response an adverse category placement preserves those crossings. A single favorable percentage therefore gives an incomplete account of this engagement gradient. Keywords: ordinal response; community membership; participation frequency; distribution comparison; public feedback

A Rising Yes Share Can Hide Crossing Community-Membership Distributions

Abstract

Among the most frequent Stack Overflow participants, stronger endorsement of community membership coexists with a larger fraction choosing the most negative response. We compare every ordinal cutpoint across six participation groups in 44,718 public survey records. Three adjacent comparisons show upward ordering in their observed distributions; only two meet a simultaneous working-interval criterion. The final two comparisons cross, although their combined Yes shares increase. Allowing every observed Not sure response an adverse category placement preserves those crossings. A single favorable percentage therefore gives an incomplete account of this engagement gradient.

Keywords: ordinal response; community membership; participation frequency; distribution comparison; public feedback

1. What does an increasing membership percentage establish?

The fraction answering Yes is an accessible summary of a community survey. Its meaning becomes less straightforward when the source offers two Yes responses, a neutral response, and two No responses. An increase in the combined Yes fraction may accompany a decrease in neutral answers, a concentration at the strongest endorsement, or a simultaneous increase at both extremes. These configurations describe different experiences. Reporting only the favorable share leaves their distinction unresolved, even if the percentage has been calculated correctly.

The 2024 Stack Overflow Developer Survey records Q&A participation frequency and asks whether respondents consider themselves members of the Stack Overflow community (Stack Exchange Inc., 2024). Its membership question provides an opportunity to examine that distinction without constructing a new attitude scale. Our research question is whether adjacent participation groups are ordered in the same direction at every membership-response cutpoint, and whether that conclusion survives the unknown position of the separate Not sure response. This is a question about the distribution of recorded self-reports, rather than about whether participation produces a psychological sense of belonging.

Community theory makes the distinction between activity and membership worth preserving. McMillan and Chavis (1986) develop a definition and theory of sense of community, whereas Blanchard and Markus (2004) investigate the experience of a virtual community. Those broader constructs cannot be reconstructed from one public survey item. Here the item's own wording defines the observation: an answer about considering oneself a community member. Construct validity would require evidence beyond a convenient label or an orderly numerical pattern, as the foundational discussion by Cronbach and Meehl (1955) makes clear.

Anonymous-feedback analysis provides a related motivation for examining more than an overall average. The deposited abstract of Yuan et al. (2026) describes uncertainty-aware group differences and scale validation within a privacy-constrained framework for a private learning community. We accessed that abstract rather than its underlying private records or full manuscript. The present study takes a narrower public-data question: whether a group comparison remains true across the cutpoints already supplied by one questionnaire. It does not reproduce their longitudinal framework or claim to validate that framework's procedures.

2. Participation supplies a comparison, not a causal explanation

Research on lurking shows why low public activity should not automatically be interpreted as absence from a community. Preece et al. (2004) identify several reasons for not posting, including reasons unrelated to unwillingness to benefit others. Nonnecke et al. (2006) compare attitudes among public and nonpublic participants. These studies concern different communities and participation definitions, but they motivate keeping an activity measure separate from an attitudinal response. They do not establish the direction or size of the relationships in the developer records used here.

Knowledge-contribution research also treats participation as a substantive process. Wasko and Faraj (2005) examine social capital and contributions in electronic networks of practice. Joyce and Kraut (2006) study whether newcomers continue participating after their first newsgroup contribution. Neither setting supplies a causal effect for our survey. In particular, a cross-sectional respondent classified as a frequent participant is not the same observation as an initially infrequent participant followed after an intervention. Selection into frequent participation and the possibility that feeling like a member encourages participation both remain open.

The developer setting adds further reasons to avoid treating all activity as a single psychological variable. Mamykina et al. (2011) examine the design of Stack Overflow, and Anderson et al. (2012) investigate value derived from activity on that focused Q&A site. The survey's participation question combines asking, answering, voting and commenting. We do not observe which component primarily produced a respondent's frequency answer. A group may consequently include people who engage through different roles, with different reasons for calling themselves members.

The analysis has three linked propositions. P1 retains the five ordered membership categories and compares their four exceedance probabilities. P2 asks whether the higher participation group has an upward distributional ordering at all four cutpoints; a positive Yes contrast alone is insufficient. P3 asks whether assigning the unordered Not sure answers to any ordered category can change that conclusion. We leave P2's direction open. Greater participation may accompany greater endorsement overall while the most negative tail behaves differently. Figure 1 connects these propositions to their actual measurements and calculations.

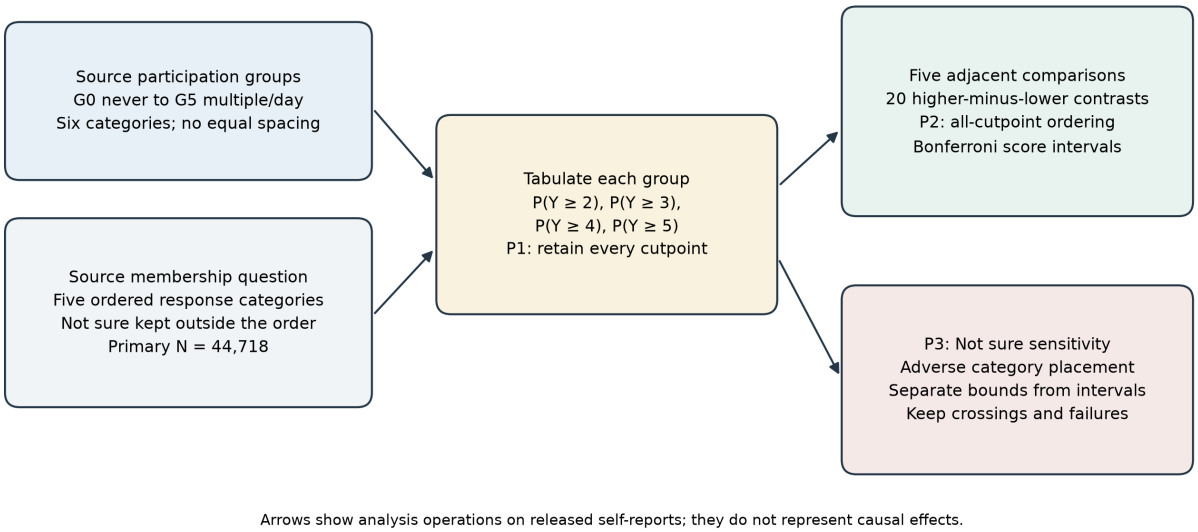

Figure 1. Analysis model for P1-P3. Arrows indicate tabulation and comparison operations, not causal paths. No latent psychological scale or participant-level behavior model is fitted.

3. Source records, categories and the analysis population

We reused the official public 2024 response CSV, schema and questionnaire, downloaded on 16 November 2026 and identified by their saved SHA-256 hashes. The archive attributes the database to Stack Exchange Inc. and supplies ODbL 1.0 terms, with DbCL 1.0 terms for individual contents. Reuse conditions and attribution remain attached to the local derivative. The analysis reads only participation and membership fields; an identifier is checked for source integrity and replaced by a local row number. It uses no free text, geographic combinations, employer fields or identity resolution.

The actual questionnaire on pages 29-30 establishes the category wording. Participation runs from never participating in Q&A through monthly-or-less, monthly-or-weekly, a few times per week, daily-or-almost-daily, and multiple times per day. These are ordered descriptions with uneven and partly combined frequency intervals. We label them G0 through G5 for compact reporting. We do not turn them into counts of contributions, hours spent on the site, or equally spaced numerical exposure values.

The membership responses are ordered as No, not at all; No, not really; Neutral; Yes, somewhat; and Yes, definitely. Category numbers one through five index this order, without specifying distances between the answers. Not sure is a separate response whose location cannot be read from the instrument. Survey-response research emphasizes the processes connecting a question to its recorded answer (Tourangeau et al., 2000). Our coding preserves the distinction between neutrality and uncertainty about how to answer rather than assuming those processes produce the same response.

Of 65,437 released records, 20,200 lack a recorded participation category. Among the 45,237 with participation observed, 156 lack membership responses and 363 choose Not sure. The resulting primary cohort contains 44,718 ordered pairs. Table 1 reports every retained response category and the separately retained Not sure counts. These exclusions define the target collection; they are not evidence that excluded respondents resemble included respondents. Missing-data analysis requires assumptions that the presence of a large public file cannot supply (Little and Rubin, 2019).

Table 1. Source-category counts in the declared participation groups. Short group labels abbreviate the original wording; counts preserve the original categories.

Group	Participation frequency	Ordered N	Not sure N	No, not at all	No, not really	Neutral	Yes, somewhat	Yes, definitely
G0	Never	10928	165	1972	4020	2600	1781	555
G1	Monthly or less	23942	147	1942	7190	5676	6948	2186
G2	Monthly or weekly	6221	35	104	703	1100	2717	1597
G3	A few times/week	2264	5	23	172	283	973	813
G4	Daily or almost daily	927	6	12	37	72	315	491
G5	Multiple times/day	436	5	15	13	21	100	287

4. Comparing an ordinal distribution without choosing its spacing

The distinction between order and metric spacing has a long history in measurement theory (Stevens, 1946). More recent demonstrations show that treating ordinal observations as metric can change statistical conclusions, including effect directions (Liddell and Kruschke, 2018). Svensson (2001) discusses evaluation of rating-scale data, and Agresti (2010) presents methods organized around ordinal categorical responses. These sources support making the category interpretation explicit before deciding which summary to calculate.

For each group and each cutpoint k from two through five, we calculate the fraction selecting category k or a higher category. The four fractions successively concern avoiding the most negative response, choosing at least Neutral, choosing one of the Yes responses, and choosing Yes, definitely. For each adjacent participation pair, we subtract the lower-frequency group's fraction from the higher-frequency group's fraction. The fixed family therefore contains twenty differences in proportions. We express them in percentage points so that their units remain the same at every cutpoint.

This family answers a stronger question than one favorable-share comparison. If all four differences are nonnegative, with at least one positive, the observed higher-frequency distribution has upward first-order ordinal dominance. For any strictly increasing category scores, the difference between the two mean scores equals a sum of these four differences, each multiplied by the corresponding positive increase between neighboring scores. Consequently the same direction survives every increasing recoding. When differences have opposing signs, an analyst can emphasize one region of the scale enough to obtain a different mean-score ordering. This identity follows directly from summing the five category probabilities.

The design does not require a cumulative regression model or a proportional-odds assumption. Bürkner and Vuorre (2019) describe ordinal regression as a way to respect ordered outcomes; our question is resolved earlier, by showing every threshold contrast in the released groups. There is no adjusted slope to interpret and no assumption that the relationship has a common odds ratio across cutpoints. This choice also prevents a fitted one-number gradient from obscuring an observed crossing that the research question specifically asks us to retain.

For working uncertainty, we calculate Newcombe's hybrid-score intervals for differences between independent proportions, using Wilson component limits and no continuity correction (Wilson, 1927; Newcombe, 1998). We apply a Bonferroni allocation over the twenty prespecified contrasts: each interval has nominal coverage 99.75%, giving nominal joint coverage of at least 95% if the component coverage conditions hold. Score coverage is approximate, so this is not an exact finite-sample guarantee. The general simultaneous-comparison problem discussed by Dunn (1961) motivates declaring the family before inspecting favorable results; we do not apply a comparison-of-means model from that paper.

Different cutpoints reuse respondents, and adjacent comparisons share groups. Bonferroni adjustment does not require those twenty comparisons to be independent. The underlying two-group score intervals nevertheless use an independent-respondent binomial working model. We report upward working statistical support only when all four adjusted lower limits exceed zero. Failure of this criterion means that the joint positive claim is unsupported by these intervals; it neither establishes equality nor changes an observed positive difference into a negative one.

The separate sensitivity calculation addresses a different uncertainty. Within each participation group, we place every observed Not sure answer below or above a given cutpoint to obtain exact lower and upper fractions. For each difference, adverse placements in its two groups yield the most negative and most

positive possible contrast. This bounds an unresolved response location instead of choosing a single convenient code, following the broader logic of partial identification (Manski, 2003). These bounds are not confidence intervals, and they do not account for respondents whose membership answer or participation frequency is wholly missing.

5. A broad gradient with local distributional crossings

Table 2 shows an increase in the combined Yes share at every adjacent step, although the final increase has a working interval that includes zero. The first increase is 16.774 percentage points and the second is 31.195 points. More substantially for P2, both comparisons have positive differences at every cutpoint, and their smallest simultaneous lower limits are 3.184 and 5.672 points. These two comparisons therefore meet both the observed-ordering criterion and the stronger working-interval criterion.

Table 2. Adjacent-group results. Yes combines Yes, somewhat and Yes, definitely. Each working interval belongs to the fixed twenty-contrast family. The model is an independent-binomial group-contrast model with Newcombe/Wilson score limits; no covariates or fitted regression coefficients are used. All effects and limits are percentage points.

Comparison	Higher N / lower N	Yes difference, pp	Working simultaneous interval, pp	All-cutpoint point range, pp	Smallest simultaneous lower limit, pp
G1 - G0	23,942 / 10,928	16.774	[15.240, 18.277]	[4.052, 16.774]	3.184
G2 - G1	6,221 / 23,942	31.195	[29.163, 33.173]	[6.440, 31.195]	5.672
G3 - G2	2,264 / 6,221	9.541	[6.325, 12.598]	[0.656, 10.239]	-0.306
G4 - G3	927 / 2,264	8.060	[3.602, 12.092]	[-0.279, 17.057]	-2.032
G5 - G4	436 / 927	1.814	[-4.349, 7.124]	[-2.146, 12.859]	-6.008

The third comparison shows why observed ordering and working statistical support must be reported separately. A few-times-per-week respondents have positive point differences relative to the monthly-or-weekly group at all four cutpoints. Their smallest increase is only 0.656 points, at the cutpoint that excludes No, not at all. Its simultaneous interval runs from -0.306 to 1.392 points. The Yes contrast is much larger, at 9.541 points, but its favorable interval cannot substitute for the unresolved lower-tail comparison. This pair satisfies observed upward ordering while failing the declared joint support criterion.

The next two pairs fail the observed ordering itself. Daily-or-almost-daily respondents have an 8.060-point larger Yes share than the a-few-times-per-week group, yet a 0.279-point smaller fraction avoiding No, not at all. Multiple-times-per-day respondents have a 1.814-point larger Yes share than daily respondents, but their fractions reaching at least No, not really and at least Neutral are lower by 2.146 and 1.136 points. Meanwhile their strongest endorsement is 12.859 points more common. A greater Yes share and a heavier negative tail coexist in the same released distributions.

Figure 2 makes the cutpoint dependence visible on a common percentage-point axis. Its points and intervals are displayed at every threshold rather than using color to announce a successful comparison. The choice of aligned positions is consistent with the graphical-perception problem examined by Cleveland and McGill (1984). The figure also shows that the negative tail estimates have working intervals spanning zero. Their observed crossings are therefore descriptive features of this collection, not evidence that corresponding population tail differences have been established as negative.

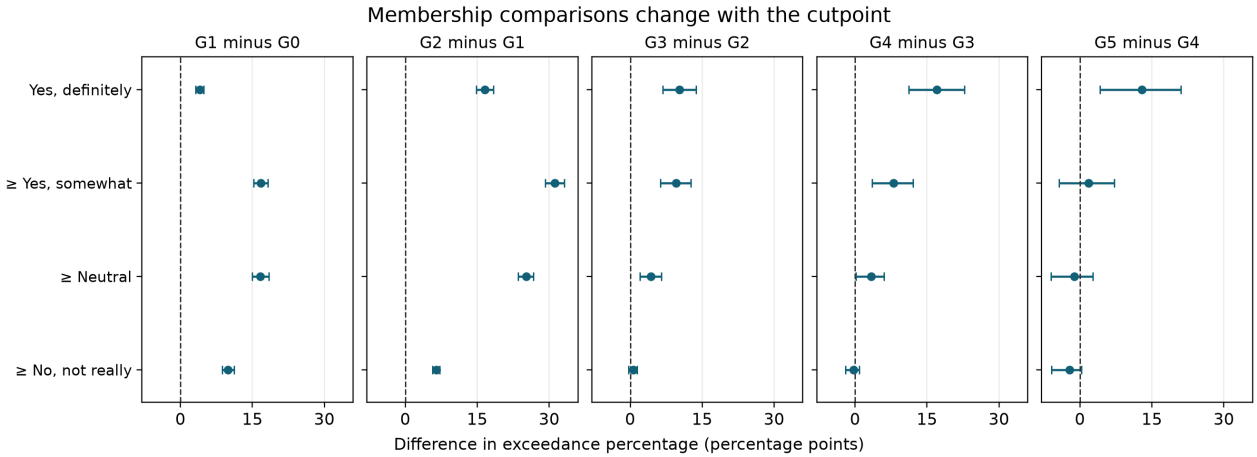

Figure 2. Points are observed higher-minus-lower exceedance differences. Bars are two-sided Newcombe intervals using Bonferroni allocation over twenty contrasts, with nominal joint 95% working coverage. Shared axes preserve the size of each contrast. G0-G5 follow Table 1.

Table 3 retains the full result family and the Not sure sensitivity. For the first three adjacent pairs, every adverse-placement lower bound remains positive. For the two crossing pairs, some sensitivity intervals remain entirely negative while other intervals remain positive. Thus arbitrary placement of the observed Not sure answers does not eliminate either crossing. For example, the final pair's contrast at cutpoint two remains between -3.249 and -1.472 points, while its strongest-endorsement contrast remains between 11.810 and 13.587 points. This sensitivity strengthens a claim about category-placement robustness, not a claim about sampling precision.

Table 3. Complete cutpoint family and separate placement sensitivity. The latter bounds include recorded Not sure responses within each participation group. Ordered sample sizes are given in Tables 1-2. Bounds concern unresolved response location; working intervals concern the declared binomial model.

Contrast	Cutpoint	Difference, pp	Working simultaneous interval, pp	Not sure placement bounds, pp
G1 - G0	2	9.934	[8.718, 11.185]	[9.105, 11.203]
G1 - G0	3	16.689	[14.960, 18.407]	[15.496, 17.594]
G1 - G0	4	16.774	[15.240, 18.277]	[15.372, 17.470]
G1 - G0	5	4.052	[3.184, 4.884]	[2.584, 4.682]
G2 - G1	2	6.440	[5.672, 7.134]	[5.840, 7.010]
G2 - G1	3	25.170	[23.529, 26.729]	[24.450, 25.620]
G2 - G1	4	31.195	[29.163, 33.173]	[30.430, 31.600]
G2 - G1	5	16.541	[14.804, 18.335]	[15.843, 17.012]
G3 - G2	2	0.656	[-0.306, 1.392]	[0.428, 1.208]
G3 - G2	3	4.359	[2.049, 6.463]	[4.085, 4.865]
G3 - G2	4	9.541	[6.325, 12.598]	[9.196, 9.976]
G3 - G2	5	10.239	[6.798, 13.744]	[9.744, 10.524]
G4 - G3	2	-0.279	[-2.032, 0.857]	[-0.916, -0.052]
G4 - G3	3	3.327	[0.187, 5.995]	[2.699, 3.563]
G4 - G3	4	8.060	[3.602, 12.092]	[7.455, 8.318]
G4 - G3	5	17.057	[11.207, 22.797]	[16.575, 17.438]
G5 - G4	2	-2.146	[-5.955, 0.341]	[-3.249, -1.472]
G5 - G4	3	-1.136	[-6.008, 2.697]	[-2.231, -0.454]
G5 - G4	4	1.814	[-4.349, 7.124]	[0.724, 2.501]
G5 - G4	5	12.859	[4.207, 21.020]	[11.810, 13.587]

6. What a feedback report can conclude

The answer to the research question depends on which adjacent comparison is being described. Infrequent and moderately frequent participation groups display a consistent upward ordering in their recorded membership responses. That ordering weakens at one lower-tail threshold before crossing at the two highest-frequency comparisons. A report stating that greater participation always corresponds to a more favorable distribution would therefore overstate the data, even though the conventional Yes fraction increases at every step. The full threshold vector identifies precisely where that broad statement stops being supported.

The crossing pattern does not show that intensive participation harms membership feelings. Frequent contributors may have different roles, histories, expectations or experiences; none is measured here. The concentration at both strong endorsement and negative answers could reflect several such mixtures. Explaining it requires additional evidence rather than relabeling the present distributional comparison as a

mechanism. The contribution is diagnostic: it distinguishes an increase in one favorable percentage from a claim that survives every increasing treatment of the response labels.

Generalization is further limited by recruitment and missingness. The source methodology says that recruitment largely used Stack Overflow channels and that highly engaged users were more likely to notice invitations. Bethlehem (2010) explains the selection problems of web surveys, while Elliott and Valliant (2017) distinguish assumptions needed for inference from nonprobability samples. Our nominal score intervals cannot repair that selection. Groves and Peytcheva (2008) also connect nonresponse to bias rather than treating a response rate as a complete quality measure. Accordingly, neither the retained cohort size nor the narrow early-comparison intervals identify a representative developer-population gradient.

The sensitivity analysis has a similarly specific boundary. It shows what remains true if the observed Not sure responses had any location in the five-category order. It does not establish that these respondents possess an unreported ordered attitude, nor does it resolve the wholly missing participation and membership records. A dashboard could display their counts alongside the full ordered distribution and allow readers to inspect each cutpoint. Such a display would preserve useful comparisons without presenting a single category collapse as a comprehensive community measure.

7. Computational record and conclusion

All estimates were executed locally from the hashed public CSV. The primary calculation used NumPy arrays for category counts and arithmetic, with SciPy's normal quantile for the fixed interval critical value (Harris et al., 2020; Virtanen et al., 2020). Figures were drawn locally with Matplotlib (Hunter, 2007). A separate implementation reread the raw CSV using counters and the standard-library normal distribution, reconstructing all counts, twenty effects, interval endpoints and placement bounds. The largest absolute numerical discrepancy was below 2.4×10^{-16}, with no check failures. This is computational agreement between two implementations, not an independent human evaluation of the scientific claims.

Scientific Agent Skills informed evidence recording, statistical reporting and figure review, and is cited as actual procedural software used in the work (Kassis et al., 2026). Source-specific reading limits, the frozen protocol, source hashes, code and machine-readable results accompany the draft. Accountable-author approval, declarations and human source verification remain pending. No ethics exemption, institutional approval, participant consent beyond the source's documented process, funding status or publication authorization is inferred from public availability.

An ordinal survey gradient can be informative while remaining incomplete. In this release, the first two adjacent comparisons support a positive all-cutpoint claim under the declared working intervals; the next comparison remains uncertain at one threshold; and the final two show observed crossings that survive all placements of recorded Not sure answers. Reporting those distinctions answers the question more precisely than an uninterrupted sequence of increasing Yes percentages. The practical implication is to show the distribution that supports a comparison and identify the strength of the ordering actually established.

References

1. Liddell, Torrin M.; Kruschke, John K.. (2018). Analyzing ordinal data with metric models: What could possibly go wrong?. Journal of Experimental Social Psychology. https://doi.org/10.1016/j.jesp.2018.08.009
2. Wasko, Molly McLure; Faraj, Samer. (2005). Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. MIS Quarterly. https://doi.org/10.2307/25148667
3. Cronbach, Lee J.; Meehl, Paul E.. (1955). Construct validity in psychological tests.. Psychological Bulletin. https://doi.org/10.1037/h0040957

4. Tourangeau, Roger; Rips, Lance J.; Rasinski, Kenneth. (2000). The Psychology of Survey Response. Cambridge University Press. https://doi.org/10.1017/cbo9780511819322

5. Nonnecke, Blair; Andrews, Dorine; Preece, Jenny. (2006). Non-public and public online community participation: Needs, attitudes and behavior. Electronic Commerce Research. https://doi.org/10.1007/s10660-006-5985-x

6. Bethlehem, Jelke. (2010). Selection Bias in Web Surveys. International Statistical Review. https://doi.org/10.1111/j.1751-5823.2010.00112.x

7. Joyce, Elisabeth; Kraut, Robert E.. (2006). Predicting Continued Participation in Newsgroups. Journal of Computer-Mediated Communication. https://doi.org/10.1111/j.1083-6101.2006.00033.x

8. Dunn, Olive Jean. (1961). Multiple Comparisons among Means. Journal of the American Statistical Association. https://doi.org/10.1080/01621459.1961.10482090

9. Wilson, Edwin B.. (1927). Probable Inference, the Law of Succession, and Statistical Inference. Journal of the American Statistical Association. https://doi.org/10.1080/01621459.1927.10502953

10. Groves, R. M.; Peytcheva, E.. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. Public Opinion Quarterly. https://doi.org/10.1093/poq/nfn011

11. Stevens, S. S.. (1946). On the Theory of Scales of Measurement. Science. https://doi.org/10.1126/science.103.2684.677

12. Hunter, John D.. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering. https://doi.org/10.1109/mcse.2007.55

13. Blanchard, Anita L.; Markus, M. Lynne. (2004). The experienced "sense" of a virtual community: characteristics and processes. ACM SIGMIS Database: the DATABASE for Advances in Information Systems. https://doi.org/10.1145/968464.968470

14. McMillan, David W.; Chavis, David M.. (1986). Sense of community: A definition and theory. Journal of Community Psychology. https://doi.org/10.1002/1520-6629(198601)14:1<6::aid-jcop2290140103>3.0.co;2-i

15. Cleveland, William S.; McGill, Robert. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Journal of the American Statistical Association. https://doi.org/10.1080/01621459.1984.10478080

16. Manski, Charles F.. (2003). Partial Identification of Probability Distributions. Springer New York. https://doi.org/10.1007/b97478

17. Newcombe, Robert G.. (1998). Interval estimation for the difference between independent proportions: comparison of eleven methods. Statistics in Medicine. https://doi.org/10.1002/(sici)1097-0258(19980430)17:8<873::aid-sim779>3.0.co;2-i

18. Svensson, Elisabeth. (2001). GUIDELINES TO STATISTICAL EVALUATION OF DATA FROM RATING SCALES AND QUESTIONNAIRES. Journal of Rehabilitation Medicine. https://doi.org/10.1080/165019701300006542

19. Little, Roderick; Rubin, Donald. (2019). Statistical Analysis with Missing Data, Third Edition. Wiley Series in Probability and Statistics. https://doi.org/10.1002/9781119482260

20. Stack Exchange Inc.. (2024). 2024 Stack Overflow Developer Survey [public response data, schema and questionnaire]. Stack Overflow Developer Survey archive. https://survey.stackoverflow.co/2024/

21. Elliott, Michael R.; Valliant, Richard. (2017). Inference for Nonprobability Samples. Statistical Science. https://doi.org/10.1214/16-sts598

22. Harris, Charles R.; Millman, K. Jarrod; van der Walt, Stéfan J.; Gommers, Ralf; Virtanen, Pauli; Cournapeau, David; Wieser, Eric; Taylor, Julian; Berg, Sebastian; Smith, Nathaniel J.; Kern, Robert; Picus, Matti; Hoyer, Stephan; van Kerkwijk, Marten H.; Brett, Matthew; Haldane, Allan; del Río, Jaime Fernández; Wiebe, Mark; Peterson, Pearu; Gérard-Marchant, Pierre; Sheppard, Kevin; Reddy, Tyler; Weckesser, Warren; Abbasi, Hameer; Gohlke, Christoph; Oliphant, Travis E.. (2020). Array programming with NumPy. Nature. https://doi.org/10.1038/s41586-020-2649-2

23. Preece, Jenny; Nonnecke, Blair; Andrews, Dorine. (2004). The top five reasons for lurking: improving community experiences for everyone. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2003.10.015

24. Mamykina, Lena; Manoim, Bella; Mittal, Manas; Hripcsak, George; Hartmann, Björn. (2011). Design lessons from the fastest q&a site in the west. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/1978942.1979366

25. Bürkner, Paul-Christian; Vuorre, Matti. (2019). Ordinal Regression Models in Psychology: A Tutorial. Advances in Methods and Practices in Psychological Science. https://doi.org/10.1177/2515245918823199

26. Anderson, Ashton; Huttenlocher, Daniel; Kleinberg, Jon; Leskovec, Jure. (2012). Discovering value from community activity on focused question answering sites: a case study of stack overflow. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. https://doi.org/10.1145/2339530.2339665

27. Kassis, Timothy; Agarwal, Vinayak; He, Yuhuan; Patel, Darshil; Brueckner, Aubrey M.. (2026). Scientific Agent Skills: A Library of Procedural Knowledge for Research Agents. arXiv preprint. https://arxiv.org/abs/2609.00065

28. Yuan, Yuewei; Huang, Jialei; Yin, Jiayang; Xu, Tianyi. (2026). Visual Verification of user Experience Differences under Privacy Constraints in Anonymous Feedback Data. SSRN working paper 7393298. https://doi.org/10.2139/ssrn.7393298.

29. Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; Haberland, Matt; Reddy, Tyler; Cournapeau, David; Burovski, Evgeni; Peterson, Pearu; Weckesser, Warren; Bright, Jonathan; van der Walt, Stéfan J.; Brett, Matthew; Wilson, Joshua; Millman, K. Jarrod; Mayorov, Nikolay; Nelson, Andrew R. J.; Jones, Eric; Kern, Robert; Larson, Eric; Carey, C J; Polat, İlhan; Feng, Yu; Moore, Eric W.; VanderPlas, Jake; Laxalde, Denis; Perktold, Josef; Cimrman, Robert; Henriksen, Ian; Quintero, E. A.; Harris, Charles R.; Archibald, Anne M.; Ribeiro, Antônio H.; Pedregosa, Fabian; van Mulbregt, Paul; SciPy 1.0 Contributors; Vijaykumar, Aditya; Bardelli, Alessandro Pietro; Rothberg, Alex; Hilboll, Andreas; Kloeckner, Andreas; Scopatz, Anthony; Lee, Antony; Rokem, Ariel; Woods, C. Nathan; Fulton, Chad; Masson, Charles; Häggström, Christian; Fitzgerald, Clark; Nicholson, David A.; Hagen, David R.; Pasechnik, Dmitrii V.; Olivetti, Emanuele; Martin, Eric; Wieser, Eric; Silva, Fabrice; Lenders, Felix; Wilhelm, Florian; Young, G.; Price, Gavin A.; Ingold, Gert-Ludwig; Allen, Gregory E.; Lee, Gregory R.; Audren, Hervé; Probst, Irvin; Dietrich, Jörg P.; Silterra, Jacob; Webber, James T; Slavič, Janko; Nothman, Joel; Buchner, Johannes; Kulick, Johannes; Schönberger, Johannes L.; de Miranda Cardoso, José Vinícius; Reimer, Joscha; Harrington, Joseph; Rodríguez, Juan Luis Cano; Nunez-Iglesias, Juan; Kuczynski, Justin; Tritz, Kevin; Thoma, Martin; Newville, Matthew; Kümmerer, Matthias; Bolingbroke, Maximilian; Tartre, Michael; Pak, Mikhail; Smith, Nathaniel J.; Nowaczyk, Nikolai; Shebanov, Nikolay; Pavlyk, Oleksandr; Brodtkorb, Per A.; Lee, Perry; McGibbon, Robert T.; Feldbauer, Roman; Lewis, Sam; Tygier, Sam; Sievert, Scott; Vigna, Sebastiano; Peterson, Stefan; More, Surhud; Pudlik, Tadeusz; Oshima, Takuya; Pingel, Thomas J.; Robitaille, Thomas P.; Spura, Thomas; Jones, Thouis R.; Cera, Tim; Leslie, Tim; Zito, Tiziano; Krauss, Tom; Upadhyay, Utkarsh; Halchenko, Yaroslav O.; Vázquez-Baeza, Yoshiki. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. https://doi.org/10.1038/s41592-019-0686-2

30. Agresti, Alan. (2010). Analysis of Ordinal Categorical Data. Wiley Series in Probability and Statistics. https://doi.org/10.1002/9780470594001