

Data Quality Benchmarking

WAVE 1 | H1 2025



Table of Contents

About	3
Methodology.....	4
Key Findings	5
Progress.....	5
Global Quality Removal Benchmarks.....	6
Global Quality Removal Benchmarks by Study Type.....	7
Research Agency Quality Removal Benchmarks	8
Supplier Quality Removal Benchmarks.....	9
US Quality Removal Benchmarks.....	10
US Quality Removal Benchmarks by Study Type.....	11
Global Quality Benchmarks.....	12
Global Quality Benchmarks by Study Type	12
Research Agency Quality Benchmarks	13
Supplier Quality Benchmarks.....	13
US Quality Benchmarks.....	14
US Quality Benchmarks by Study Type	14
How to use the Data.....	15
Next Steps.....	16

Appendix

Data Mapping for Benchmarks.....	17
Definitions.....	18
We Made Some Changes	19
Approach to Weighting.....	20
Countries and World Bank Regions	21
Additional Reports.....	22

About the GDQ Data Quality Benchmarking Project

The Global Data Quality (GDQ) Data Quality Benchmarking Project is a global research-on-research initiative to understand and improve data quality. It seeks to define and track a set of standards or norms to be used to assess data quality at both the industry and individual company levels. With these norms, this initiative aims to contribute to the development of globally accepted data quality buying signals and fuel innovation.

How it works

This initiative can only be done through industry collaboration, with the participating companies in the research, insights and analytics industry (including brands, research agencies, panel providers, tech platforms, digital analytics and consultancies) contributing random sets of data from online survey projects. Metrics are aggregated in various ways to assess our ecosystem.

Methodology

Data collected from January - June 2025

Participating Companies were asked to provide a random selection of n=10,000 pseudonymized data cases

This wave doubled the companies and doubled the data contributed from beta Wave 0, and became global

51
Companies

78
Countries

~2M
Records

Results in this report are shown by the below data cuts:

Weighted Research Agency

By global total,
key regions, and US

Weighted Supplier

By global total,
key regions, and US

Unweighted Global Study

Type (B2C, B2B, Healthcare
patient, Healthcare provider)

Details on the countries that make up the global picture as well as key regions can be found in the appendix.

US data has been included as it comprises a ~60% market share of the global insights and analytics industry, and represents ~50% of the global data in this wave 1 benchmark.

Where sample allows, data for other countries including UK, CA, AU, and JP can be downloaded through links at the end of the report.

Key Findings

Globally, Research Agencies report lower fraud removal than Suppliers. This is expected in an ecosystem where, upstream, Suppliers employ methods to clean points of entry.

Removals appear higher in North America than most other global regions, which is likely the result of a combination of factors from the varied data to the size of the market. Anecdotal evidence (from a small sample size) suggests disproportionately higher issues in some countries in East Asia.

Removals appear higher for B2B and Healthcare studies, which may be more susceptible to fraud as a result of generally higher incentives. Length of Interview (LOI) matters, as both Abandon Rate and Removals increase with longer surveys due to respondent fatigue.

Globally, across all study types, our industry appears to consistently under-estimate Incidence Rate (IR). Research Agencies report actual IR at ~7% below estimated, while Suppliers see studies at ~10% lower than estimated. This is most pronounced in Healthcare Provider work where actual IR is ~24% below estimated.

Progress

Wave 0: Process Findings

Key areas for improvement in engaging firms to contribute more data for B2B and healthcare online research, as well as gathering data from brands.

Wave 0: Benchmarks

Areas of immediate focus

Reducing the need for data removal through the reduction of fraud.

Increasing the use of technology to secure panels and surveys.

Decreasing abandon rates and data removal through improved survey experiences.

Wave 1: Process Updates

Collation of B2B and Healthcare data is much improved, with opportunity to shore up in key markets for robustness. Gathering data from brands remains an area of focus for future waves.

Wave 1: Benchmark Updates

Fraud removals are at similar levels, but there appears to be a shift to pre-survey versus in-survey removal.

Link encryption has significantly improved, but there is still a need to strengthen this – particularly for Suppliers.

Data removal rates and abandon rates have stabilized, with length of interview down and increasing use of mobile devices.

Global Quality Removal Benchmarks

Pre-Survey + In-Survey Fraud Removal + In-Survey Behavior Removal



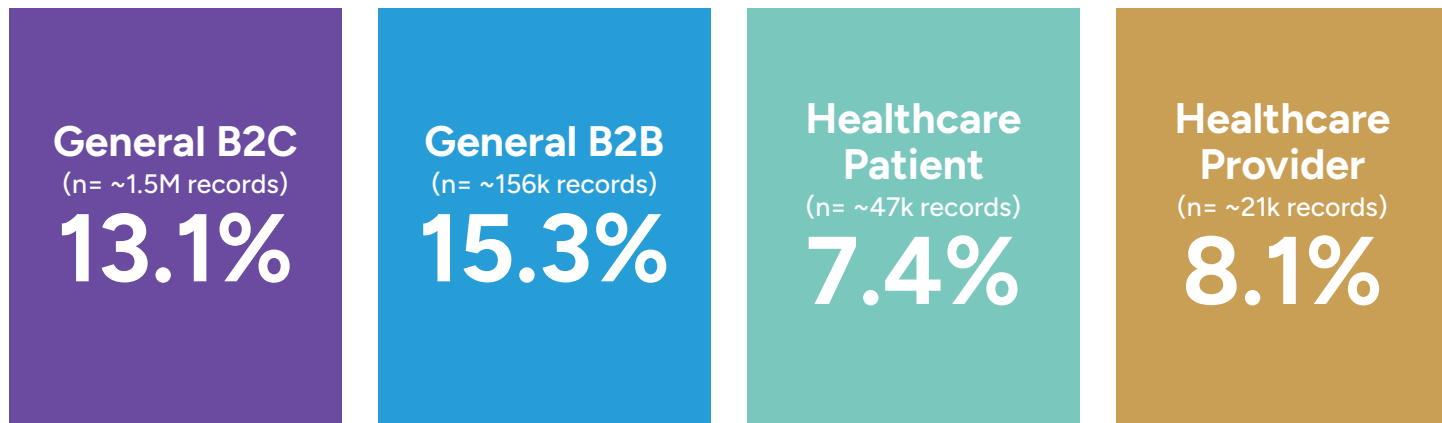
Key Findings

Suppliers show higher removals than Research Agencies, likely due to being further upstream in the process. Globally, Suppliers removed 7.4% of respondents pre-survey.

Global Benchmark	Research Agency (n= ~1.15M records)	Supplier (n= ~825k records)
Pre-Survey removal rates Defined as: Pre-survey removal - Quality Termination/ Block	2.8%	7.4%
In-Survey Fraud Removal Defined as: Terminated for Identity/ Device/ Tech/ Fraud	2.4%	1.3%
In-Survey Behavior Removal Defined as: Terminated for In-Survey Response Behaviors	4.2%	5.0%
Among Qualified Completes	Research Agency (n= ~8200 records)	Supplier (n= ~7900 records)
Post-Survey cleanout rate Defined as: Post-survey removal - Inattention/ Quality Removal	6.6%	7.4%

Global Quality Removal Benchmarks by Study Type

Pre-Survey + In-Survey Fraud Removal + In-Survey Behavior Removal



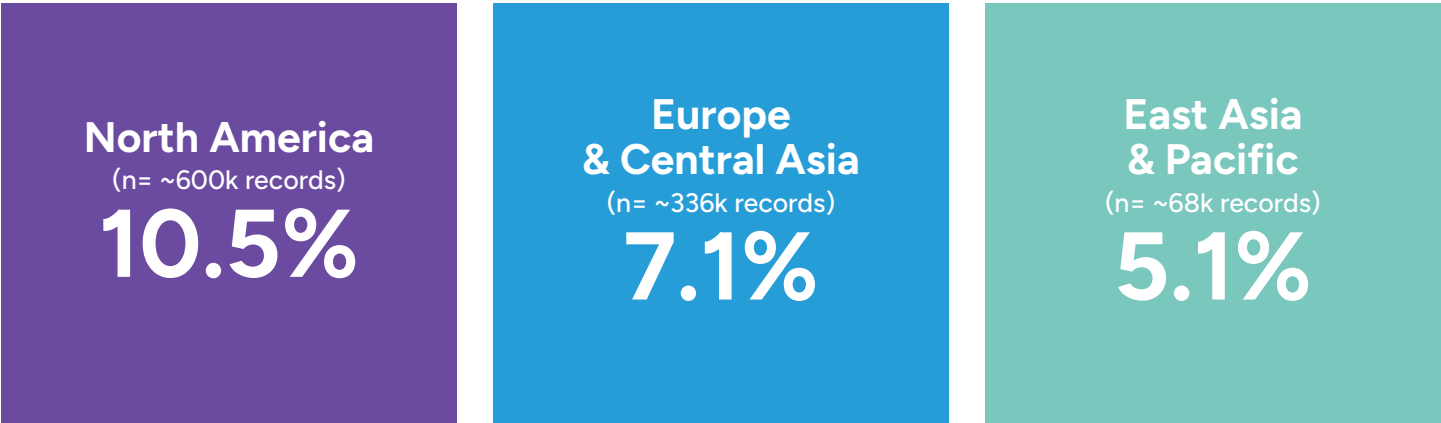
Key Findings

General B2B had the highest removal rate pre- and in-survey at 15.3%. Post-survey cleaning of that audience was considerably higher than for others, at 18.9%

Global Benchmarks by Study Type	General B2C (n= ~1.5M records)	General B2B (n= ~156k records)	Healthcare Patient (n= ~47k records)	Healthcare Provider (n= ~21k records)
Pre-Survey removal rates Defined as: Pre-survey removal - Quality Termination/ Block	4.6%	7.5%	2.1%	2.7%
In-Survey Fraud Removal Defined as: Terminated for Identity/ Device/ Tech/ Fraud	2.1%	2.0%	1.6%	1.6%
In-Survey Behavior Removal Defined as: Terminated for In-Survey Response Behaviors	6.4%	5.8%	3.7%	3.8%
Among Qualified Completes	General B2C (n= ~430k records)	General B2B (n= ~23k records)	Healthcare Patient (n= ~82k records)	Healthcare Provider (n= ~79k records)
Post-Survey cleanout rate Defined as: Post-survey removal - Inattention/ Quality Removal	5.6%	18.9%	9.8%	4.3%

Research Agency Quality Removal Benchmarks

Pre-Survey + In-Survey Fraud Removal + In-Survey Behavior Removal



Key Findings

Research agencies report higher pre- and in-survey removals from North American samples, due in part to a larger share of B2B records. Post-survey removals are generally consistent across regions.

Research Agency Benchmark by Region	North America (n= ~600k records)	Europe & Central Asia (n= ~336k records)	East Asia & Pacific (n= ~68k records)
Pre-Survey removal rates Defined as: Pre-survey removal - Quality Termination/ Block	3.1%	2.8%	1.4%
In-Survey Fraud Removal Defined as: Terminated for Identity/ Device/ Tech/ Fraud	3.3%	1.2%	0.8%
In-Survey Behavior Removal Defined as: Terminated for In-Survey Response Behaviors	4.1%	3.1%	2.9%
Among Qualified Completes	North America (n= ~200k records)	Europe & Central Asia (n= ~170k records)	East Asia & Pacific (n= ~26k records)
Post-Survey cleanout rate Defined as: Post-survey removal - Inattention/ Quality Removal	6.7%	5.9%	7.4%

Supplier Quality Removal Benchmarks

Pre-Survey + In-Survey Fraud Removal + In-Survey Behavior Removal



Key Findings

Suppliers report similar pre- and in-survey removal rates across regions. North American samples had more behavior removals, and East Asia & Pacific had the bulk of their removals pre-survey. Post-survey cleaning was highest in East Asia & Pacific.

Research Agency Benchmark by Region	North America (n= ~502k records)	Europe & Central Asia (n= ~120k records)	East Asia & Pacific (n= ~99k records)
Pre-Survey removal rates Defined as: Pre-survey removal - Quality Termination/ Block	6.7%	5.2%	9.6%
In-Survey Fraud Removal Defined as: Terminated for Identity/ Device/ Tech/ Fraud	0.8%	2.6%	1.0%
In-Survey Behavior Removal Defined as: Terminated for In-Survey Response Behaviors	6.6%	4.0%	2.1%
Among Qualified Completes	North America (n= ~93k records)	Europe & Central Asia (n= ~59k records)	East Asia & Pacific (n= ~22k records)
Post-Survey cleanout rate Defined as: Post-survey removal - Inattention/ Quality Removal	6.4%	7.3%	11.1%

US Quality Removal Benchmarks

Pre-Survey + In-Survey Fraud Removal + In-Survey Behavior Removal



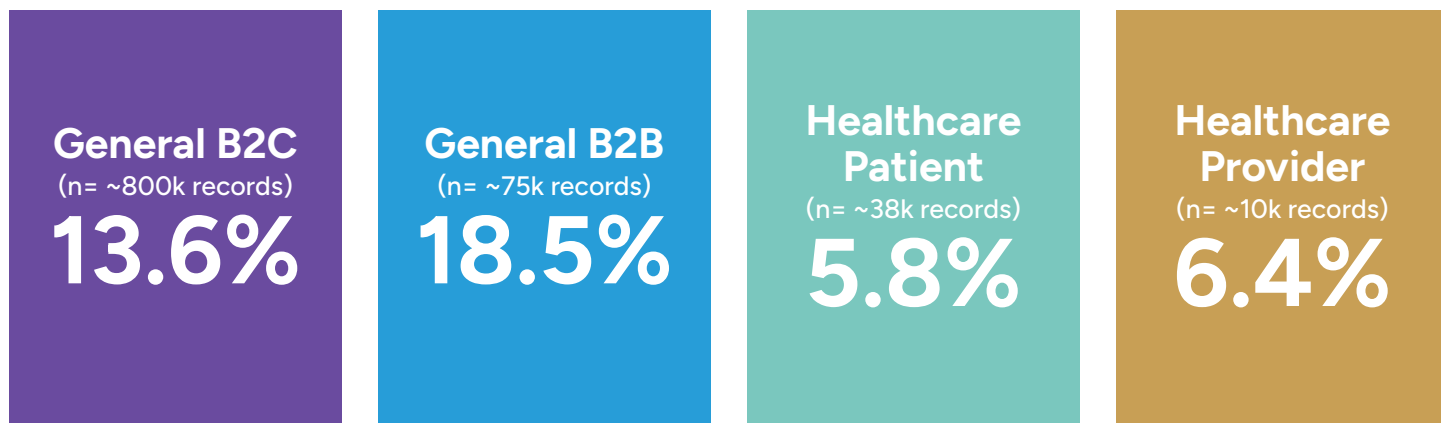
Key Findings

For US surveys, Research Agencies and Suppliers report fairly similar double-digit removal rates, with Suppliers removing more sample pre-survey

US Benchmark	Research Agency (n= ~540k records)	Supplier (n= ~550k records)
Pre-Survey removal rates Defined as: Pre-survey removal - Quality Termination/ Block	3.2%	6.2%
In-Survey Fraud Removal Defined as: Terminated for Identity/ Device/ Tech/ Fraud	3.8%	0.5%
In-Survey Behavior Removal Defined as: Terminated for In-Survey Response Behaviors	4.1%	6.7%
Among Qualified Completes	Research Agency (n= ~180k records)	Supplier (n= ~87k records)
Post-Survey cleanout rate Defined as: Post-survey removal - Inattention/ Quality Removal	7.4%	6.2%

US Quality Removal Benchmarks by Study Type

Pre-Survey + In-Survey Fraud Removal + In-Survey Behavior Removal



Key Findings

Among US samples, General B2B had considerably higher removal rates than other audiences. Despite strong front-end quality measures, considerable data loss occurred during post-survey cleaning.

US Benchmark by Study Type	General B2C (n= ~800k records)	General B2B (n= ~75k records)	Healthcare Patient (n= ~38k records)	Healthcare Provider (n= ~10k records)
Pre-Survey removal rates Defined as: Pre-survey removal - Quality Termination/ Block	3.5%	9.2%	1.3%	1.9%
In-Survey Fraud Removal Defined as: Terminated for Identity/ Device/ Tech/ Fraud	2.7%	3.3%	0.0%	0.1%
In-Survey Behavior Removal Defined as: Terminated for In-Survey Response Behaviors	7.4%	6.0%	4.5%	4.4%
Among Qualified Completes	General B2C (n= ~230k records)	General B2B (n= ~12k records)	Healthcare Patient (n= ~7k records)	Healthcare Provider (n= ~4k records)
Post-Survey cleanout rate Defined as: Post-survey removal - Inattention/ Quality Removal	5.4%	11.8%	9.3%	2.1%

Global Quality Benchmarks

Link encryption is widely employed by Research Agencies to prevent fraud and is nearly universal in Healthcare Patient research, where data privacy and security standards are paramount.

	Research Agency (n= ~1.15M records)	Supplier (n= ~825k records)
Incidence rate Defined as: Mean incidence provided	Sold 61.6% Actual 55.2%	Sold 50.8% Actual 41.1%
Length of interview Defined as: Median LOI for qualified completes only	10 minutes	13 minutes
Abandon rate Defined as: Dropped Out	12.6%	13.9%
Device type	67.6% mobile	51.4% mobile
Use of Link encryption Defined as: Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	91.5%	75.2%

Global Quality Benchmarks by Study Type

Estimating Incidence in Healthcare studies appears challenging. Abandon rates are generally consistent across study types, with the exception of Healthcare Patient research where respondents have a stronger personal stake in participation.

	General B2C (n= ~1.5M records)	General B2B (n= ~156k records)	Healthcare Patient (n= ~47k records)	Healthcare Provider (n= ~21k records)
Incidence rate Defined as: Mean incidence provided	Sold 52.5% Actual 44.1%	Sold 35.4% Actual 31.3%	Sold 39.2% Actual 21.9%	Sold 65.7% Actual 42.5%
Length of interview Defined as: Median LOI for qualified completes only	12 minutes	17 minutes	15 minutes	34 minutes
Abandon rate Defined as: Dropped Out	17.2%	16.5%	7.1%	16.5%
Device type	61.5% mobile	48.9% mobile	52.6% mobile	26.4% mobile
Use of Link encryption Defined as: Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	76%	81.3%	98.8%	78%

Research Agency Quality Benchmarks

Interestingly, Research Agencies tend to overestimate incidence rates (IR) in North American studies, while significantly underestimating them in East Asia and the Pacific. Use of link encryption lags somewhat with North American samples.

	North America (n= ~600k records)	Europe & Central Asia (n= ~336k records)	East Asia & Pacific (n= ~68k records)
Incidence rate Defined as: Mean incidence provided	Sold 62% Actual 50%	Sold 66% Actual 65%	Sold 33% Actual 55%
Length of interview Defined as: Median LOI for qualified completes only	10 minutes	10 minutes	14 minutes
Abandon rate Defined as: Dropped Out	11.5%	12.7%	18.6%
Device type	63.8% mobile	74.6% mobile	57.2% mobile
Use of Link encryption Defined as: Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	86%	99.2%	98%

Supplier Quality Benchmarks

Supplier metrics are generally consistent across regions; however, link encryption adoption appears less prevalent in East Asia & Pacific samples.

	North America (n= ~502k records)	Europe & Central Asia (n= ~120k records)	East Asia & Pacific (n= ~99k records)
Incidence rate Defined as: Mean incidence provided	Sold 46% Actual 36%	Sold 60% Actual 47%	Sold 47% Actual 40%
Length of interview Defined as: Median LOI for qualified completes only	13 minutes	15 minutes	10 minutes
Abandon rate Defined as: Dropped Out	14.3%	14.7%	10%
Device type	50.6% mobile	47.6% mobile	50.1% mobile
Use of Link encryption Defined as: Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	76.8%	85.6%	50.5%

US Quality Benchmarks

US metrics are generally similar to Global metrics (as they comprise about half the sample), with the exception of lower link encryption use among Research Agencies.

	Research Agency (n= ~540k records)	Supplier (n= ~550k records)
Incidence rate Defined as: Mean incidence provided	Sold 64.7% Actual 49.4%	Sold 46.2% Actual 36.3%
Length of interview Defined as: Median LOI for qualified completes only	10 minutes	12 minutes
Abandon rate Defined as: Dropped Out	11.7%	14.4%
Device type	65.2% mobile	52.6% mobile
Use of Link encryption Defined as: Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	84%	75.4%

US Benchmarks by Study Type

Reported LOI for Healthcare Provider studies is nearly twice that of other survey types. Despite the longer duration, abandon rates are lower than in General B2C and B2B research, likely due to the higher incentives.

	General B2C (n= ~800k records)	General B2B (n= ~75k records)	Healthcare Patient (n= ~38k records)	Healthcare Provider (n= ~10k records)
Incidence rate Defined as: Mean incidence provided	Sold 51% Actual 39.4%	Sold 35.3% Actual 28.5%	Sold 40% Actual 24.4%	Sold 60.7% Actual 47%
Length of interview Defined as: Median LOI for qualified completes only	14 minutes	16 minutes	15 minutes	28 minutes
Abandon rate Defined as: Dropped Out	16%	15.4%	4.6%	12.9%
Device type	58.1% mobile	45.5% mobile	62.9% mobile	35.2% mobile
Use of Link encryption Defined as: Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	68.8%	85.3%	98.6%	68.4%

How to use the Data

1.

Benchmark, don't judge

Use these metrics to understand how your data compares with industry norms, not as an absolute measure of success or failure

2.

Segment comparisons

Compare your study outcomes against relevant benchmarks (e.g., by supplier type, research agency, country, or study type)

3.

Identify gaps

Look for areas where your results diverge significantly from the benchmarks; these may indicate opportunities for process improvements or quality innovations

4.

Track progress

Apply these benchmarks across time to measure your organization's progress wave to wave

5.

Collaborate & contribute

Your participation not only helps your organization – it refines future benchmarks and ensures the data reflects the broadest possible view of industry standards

Next Steps

Wave 1: Process Improvements

Key areas for improvement in engaging firms to contribute more global data for healthcare online research, and a focus on gathering data from brands.

Wave 1: Benchmarks

Areas of immediate focus:

- Reducing the need for data removal through the reduction of fraud.
- Link encryption has improved but there is still work to be done to secure panels and surveys.
- Increasing the use of technology to identify data removals earlier in the survey process.

Wave 2

To be completed in Q1 of 2026. More information forthcoming.

Why Participate

Contribute to the development of global data quality benchmarks that can be used to:

- Gain a quantitative understanding of what “good” data quality looks like
- Compare your studies’ performance against the industry overall, by country, company type (e.g., sample provider, research agency) and study type (e.g., B2C, healthcare provider)
- Identify white spaces against which your company might innovate

Appendix

Data mapping for benchmarks

Metric	Defined as...	Based on...
Pre-Survey removal rates	Pre-survey removal - Quality Termination/ Block	Total
In-Survey cleanout rate – Fraud removals	Terminated for Identity/ Device/ Tech/ Fraud	Total
In-Survey cleanout rate - Behavior removals	Terminated for In-Survey Response Behaviors	Total
Post-Survey cleanout rate	Post-survey removal - Inattention/ Quality Removal (Terminated)	Qualified completes
Incidence rate - Sold	Study Sold Incidence Rate, Mean incidence provided	Total
Incidence rate - Actual	Study Actual Incidence Rate, Mean incidence provided	Total
Length of interview (minutes)	Median LOI	Qualified completes
Abandon rate	In-Survey Participation Status - Dropped Out	Total
Device type - mobile		Total
Use of Link encryption	Uses Server-to-server, Link Encryption, Formula, or Secure Mobile	Total

Subgroup	Defined as...
Research Agency	Self-defined Research Agency
Supplier	Self-defined Sample Provider Company, Quality Tech Company, Platform Company, or Other

Definitions

Country of Fieldwork	The category indicating the country the fieldwork was conducted in. ISO country codes should be used (Alpha-2 code).
Study type	The category indicating the type of study being conducted.
Study Sold Incidence Rate	The expected incidence rate of the study as sold to the client.
Study Actual Incidence Rate	The actual incidence rate observed during the study.
Device Type of Respondent	The type of device used by the respondent to participate in the survey.
User Agent (full string)	This string contains information about the browser, operating system, and device of the respondent. For non-web surveys, please include information on operating system (e.g., iOS, Android), or format (e.g., SMS).
Level of Survey Link Encryption	The category indicating the level of encryption used for the survey link.
• Server-to-server	Survey data is transmitted securely from server to server.
• Link Encryption	The survey link is encrypted to protect data transmission.
• Formula	A formula-based encryption is used for the survey link.
• Secure Mobile	Survey data is transmitted via secure mobile messaging (App or SMS for examples)
• None	No encryption is used for the survey link.
Pre-Survey Attempt Status	The status indicating the initial phase of a respondent's interaction with a survey before they officially start it.
• Successfully Started	Indicates that the respondent successfully began the survey.
• Duplicate Fingerprint	A respondent identified as a duplicate through fingerprinting technology.
• Quality Termination/Block	Survey termination due to failing quality criteria OR indicating that the respondent has been blocked from participating in the survey due to previous actions or criteria not met.
In-Survey Participation Status	General category indicating the various statuses a respondent can have during a survey participation process.
• Qualified Complete	A respondent who meets the necessary criteria to be included in the survey.
• Dropped Out	A respondent who started the survey but did not complete it.
• Terminated for In-Survey Response Behaviors	Survey termination due to failing quality criteria.
• Terminated for Identity/Device/Tech/Fraud	A respondent terminated due to the participant does the following deliberately misrepresents their identity, profiling information or responses, including organizations that use bots to impersonate participants.
• Duplicate	A respondent identified as having attempted the survey more than once.
• Quota Full	Survey termination due to failing quality criteria.
• Screened Out Did Not Qualify	A respondent who was disqualified based on specific screening criteria.
Post-Survey Removal	The reason why a completed survey was removed after data collection.
• Inattention/Quality Removal (Terminated)	The respondent was removed due to failing post-field quality checks.
Validation: Duration	A calculated field and subtracting the Start time from the End time, providing overall Length of interview to assist with validation. Interviews shorter than 5 seconds will highlighted in light red, while any interviews with a negative duration - i.e., where the End time is prior to the Start time - will be highlighted in dark red.

We Made Some Changes...

Newly Added for Wave 1

- Number of studies completed per Quarter
- Global scope!
- User Agent

Wave 0	Wave 1
Study type	
B2B	General B2B
B2C	General B2C
Healthcare Provider	Healthcare Provider
	Healthcare Patient
Unknown	Other
Pre-Survey Attempt Status	
Successfully Started	Successfully Started
Quality Termination	Quality Termination/ Block
Survey Unavailable	
Respondent Blocked Status	
Participation Limit Exceeded	
System Termination	
Duplicate Fingerprint	Duplicate Fingerprint
Quota Termination	
Pre-Screener Terminated	
Other	Other

Removed for Wave 1

- Collection of Respondent ID

Wave 0	Wave 1
In-Survey Participation Status	
Qualified	Qualified Complete
Dropped Out	Dropped Out
Terminated	Terminated for In-Survey Response Behaviors
Fraud Terminated	Terminated for Identity/Device/ Tech/Fraud
Duplicate	Duplicate
Quota Full	Quota Full
Pre-Screener Terminated	Screened Out/ Did Not Qualify
Post-Survey Participation Status (Reason removed post-field)	
Quality	Inattention/ Quality Removal (Terminated)
Quota	
Other	
Unknown	
N/A	Not applicable
Level of Survey Link Encryption	
Server-to-server	Server-to-server
Link Encryption	Link Encryption
Formula	Formula
	Secure Mobile
None	None
Unknown	Don't Know

Approach to Weighting

Using learnings from Wave 0, we have implemented a 2-cell, 3-tier weighting structure on Wave 1 data to ensure the sample is balanced and can be assessed consistently in future waves.

The weighting mechanism is based on company size, using the self-reported “number of completes” as the indicator. This approach will ensure larger companies have appropriate representation while preventing smaller companies from being weighted too far down.

While the data now includes 78 countries across multiple regions, sample in many countries is small. Therefore, cross-region and in-region representation will be addressed in future waves.

Est. # of completes per Quarter	Research Agencies (n=24)	Suppliers (n=22)
0-100,000	15%	15%
100,001-2,000,000	45%	45%
2,000,001+	40%	40%

Weighted data will be used for total and country-level analysis.
Unweighted data will be used for subgroup analyses.

Countries and World Bank Regions

East Asia and Pacific			
Australia	Indonesia	Philippines	Thailand
China	Japan	Singapore	Vietnam
Guam	Malaysia	South Korea	
Hong Kong	New Zealand	Taiwan	
South Asia			
Bangladesh	India	Pakistan	
Europe and Central Asia			
Austria	French Guiana	Luxembourg	Slovenia
Belgium	Germany	Monaco	Spain
Bulgaria	Greece	Netherlands	Sweden
Croatia	Guernsey	Norway	Switzerland
Czechia	Hungary	Poland	Turkiye
Denmark	Ireland	Portugal	Ukraine
Estonia	Italy	Romania	United Kingdom
Finland	Jersey	Serbia	
France	Latvia	Slovakia	
Middle East and North Africa			
Algeria	Israel	Saudi Arabia	
Egypt	Morocco	United Arab Emirates	
Sub-Saharan Africa			
Ghana	Nigeria	South Africa	
Kenya	Reunion		
North America			
Canada	United States		
Latin America and Caribbean			
Argentina	Ecuador	Mexico	Suriname
Brazil	El Salvador	Panama	United States Virgin Islands
Chile	Guatemala	Peru	
Colombia	Honduras	Puerto Rico	

Country-Specific Supplemental Reports

This report is comprised of data provided by 51 companies, representing surveys conducted in 78 countries – a total of close to 2 million records. Sample size was sufficient to produce separate reports for the UK, Canada, Australia, and Japan.

You may access these supplemental reports below. Thanks to the member associations of the Global Data Quality initiative for their help in spreading the word and making the latest wave of this report global in scope.

[UK](#)

[Canada](#)

[Australia](#)

[Japan](#)

