



CRIC and MRS GDQ Guidance

CHECKLIST

10 Key Components of
the Survey
Data Cleaning and
Verification Workflow



The table on the next page summarizes the key steps for survey data cleaning and verifications.

Refer to the CRIC and MRS GDQ Guidance Survey Data Quality: Controls, Detection, Validation, and Removal for a comprehensive review of the data cleaning processing including pre-survey, in-survey and post-survey.

Step	Purpose	Key Activities/Checks	Action/Decision Point
1. Intake	Prepare raw data for review	Compile survey data files; remove or mask personal information unless necessary for field validation	Dataset ready for automated/manual review
2. Remove Duplicates	Ensure each participant is unique	Check for duplicate IDs, emails, IPs; analyze response similarity and timing for likely duplicates	Remove clear duplicates; flag probable cases for review
3. Automated Quality Flags	Identify obvious quality risks early	Automated flags: speeders, straight- liners, bots, red herring fails, open-end nonsense/gibberish	Flag cases for further review; rate, clear cases for removal
4. Technical & Metadata Review	Detect technical abuse or fraud	Review: IP/device/browser, geo- location inconsistency, unusual patterns	Flag for manual review; remove confirmed technical fraud
5. Review Response Patterns	Assess engagement and authenticity	Speeding (< 1/3 median LOI), patterned patterns, logic/consistency issues, response pattern anomalies	Flag cases; manual or semi-automated review
6. Open-Ended Quality Review	Evaluate open-ended responses	Check for gibberish, spam, profanity, excessive or “bot-like” perfection, copy-paste, irrelevant answers	Flag “soft” issues; accumulate for possible removal
7. Create a Composite Quality Score	Classify by severity of issues	<p>Tier 1: Clear excludes (e.g., multiple independent issues, egregious fraud) ▶ Remove</p> <p>Tier 2: Borderline (1-2 flags, context needed) ▶ Review</p> <p>Tier 3: Clear includes (no significant issues) ▶ Keep</p>	Remove only Tier 1; escalate Tier 2 for manual review
8. Documentation	Maintain transparency and enable auditing	Log all cleaning rules, thresholds, exclusions, and rationale; track percent of sample excluded/lost	Prepare detailed cleaning report
9. Representativeness & Sensitivity Analysis	Avoid bias post-cleaning	Compare sample demographics pre-/ post-cleaning; test sensitivity of results to level of cleaning/removal	Adjust weighting or flag findings as needed
10. Final Dataset Preparation & Reporting	Prepare dataset for analysis/ reporting and disclose cleaning methods in reporting	Finalize clean file, recode variables, check formatting; document all changes. Include sample sizes, removal criteria, impact on key results in reports	Dataset ready for analysis; ensure transparency for all audiences

