

Ensuring integrity through comprehensive survey data cleaning

At Drive Research, we understand that survey data cleaning is not a one-size-fits-all solution.

Identifying fraudulent respondents requires a strategic combination of approaches encompassing question design, in-survey checks, and post-fieldwork review.



Our survey data cleaning process

We implement a multi-faceted approach to weed out unreliable responses:



ReCAPTCHA Integration

We start surveys with a reCAPTCHA to eliminate bot submissions



Comprehension Questions

Incorporating questions that gauge comprehension to help filter out unreliable responses



Specific Requests

Including questions that demand specific responses to add an additional layer of data quality



Red Herring Questions

Integration of [red herrings](#) aids in identifying respondents who may not be engaging genuinely



Articulation Questions

We ask an articulation question and review the open-ends (e.g., what is your favorite movie and why?)



Duplicate IP Addresses

Metadata analysis helps eliminate responses from the same location, reducing the risk of repeat submission



Time to Complete

Reviewing this helps identify and remove outliers, ensuring the integrity of response times



Straightlining Checks

We review grid questions to identify respondents who provide consistently repetitive responses



Consistency in Open-ends

All open-ended responses are reviewed for consistency, avoiding contradictions across question sets

