

ふだん何気なく使っているリサーチ用語がグローバルには通用しない和製英語だったり、ただの社内用語だったことはありませんか？

それを知らずに使ってしまうと、思わぬ混乱や間違いを招いてしまうこともあります。

第5回は、「パネル・モニター」と「サンプル追加の呼び方」、「データクリーニング工程の呼び方」、「母数」について焦点を当てました。ここで書いた解説は探偵団による暫定的な交通整理にすぎませんので、別のご意見がありましたら編集部までお寄せください。……………『マーケティング・リサーチャー』編集部

データクリーニングの工程に関する呼称



調査実施後のデータクリーニングの工程はどのように呼び分けられているでしょうか？

調査実施後にはデータクリーニング(データクレンジング)を行います。この工程を指して「エディティング」「バリデーション」「インスペクション」「検票」といった用語が使用されますが、どのように呼び分けられているでしょうか。

調査票を用いた定量調査では、調査票回収後にまず回答の点検作業を行います。これを一般的に「検票」と呼んでいます。続いて、検票の結果として誤りが発見された回答に対して実施する修正作業が「エディティング」です。また、調査員が介する他記式調査において、調査員が適切に調査を実施したかどうかをチェックする工程が「インスペクション」です。

現在では、こういったデータクリーニングの工程全般を「バリデーション」と呼ぶリサーチャーがいるようです。もともとバリデーションとはコーディングされたスクリプトや入力データが適切かどうかの検証を意味していましたが、デジタル記録されたリサーチ結果のデータクリーニングが一般化すると、データクリーニングを広い意味で「バリデーション」と呼ぶようになったと編集部では考えています。

参考文献

ESOMAR “ESOMAR/GRBN Online Research Guideline” 2011, 2015
JMRA “JMRA マーケティングリサーチ品質基準” 2001, 2002
日経パソコン “デジタル・IT用語辞典” 日経BP社, 2012

「母数」＝「分母の数字」？



集計担当者と話していたところ、「集計ベース」＝「分母」の意味で「母数」といっていましたが……。

「母数」は、もともと「Parameter」の訳語で、母集団のある特性を表す定数のことを指し、具体的には母平均や母分散などをあげることができます。

私達がリサーチを行う時、ほとんどが悉皆調査ではなく標本調査の形で行います。ある母集団を調査対象として定義して、そこから標本を抽出して、その標本の値をもとに母集団の値を推定する、というのがそもそも標本調査です。その場合の母集団の何らかの値が、「母数」「Parameter」です。一部の統計ソフトでは、例えば重回帰分析の出力でも、偏回帰係数を「Parameter estimates」と表現していますが、これも分析に用いた標本をもとに母集団の係数を推定したらこの値になった、ということの意味しています。

「母数」という言葉から「分母の数字」を連想して、例えば調査結果から1,000名中400名が「そう思う」と答えていて回答率40%だった、などの場合、この1,000名を「母数」といってしまいがちですが、本来の母数は母集団の特性値ですから、この用法は誤りということになります。この場合には、「集計ベース」などの表現が適切だと思われます。

参考文献

JIS Z8101-1:2015