

国勢調査データの再編集について

★ 国勢調査の地域区画と地図上の地域区画の違いによるアンマッチの処理

国から公表される国勢調査の町丁別地域区画は、例えば市販のGIS用の地図で用いる町丁別地域区画とは必ずしも一致しません。その原因は、主に下記によります。

- ・ 国勢調査データは、5年に1度実施される調査結果に基づくものであり、最新の地図とは調査時点にズレがあること
- ・ 国勢調査における町丁別データの住所体系は、一般に用いられる住所体系(財団法人国土地理協会設定の住所マスター等)とは異なること

※GISで用いる地図の多くは、上記公益財団法人国土地理協会設定の住所マスターをベースとしています。

具体的にみていくと、例えば前述の公益財団法人国土地理協会の住所マスター「全国町・字ファイル」掲載の町丁数(大字町丁目数)が約**18万件**であるのに対し、2015年国勢調査の町丁数は約**22万件**あります。

※国勢調査における町丁には、一部の小字や行政区等が混在していることによります。

この町丁数の差に加え、さらに地名の違い等もあるため、単純に地名で両者をマッチングすると、**国勢調査データのうち6万件弱がアンマッチ町丁**として残ります。

当社では、このような国勢調査データにおけるアンマッチ町丁を、地域の実情に近づけるよう、独自のロジックにより地図上の町丁との対応づけを行い、集約・配分等の処理を行ったうえでデータベース化しています。

★ 過去の国勢調査データの最新の地域区画へのマッチング処理

単年次の国勢調査データだけでは、その地域の変遷を把握することができません。例えば、一戸建て住宅が散在していた地域に高層住宅が建設された場合、その地域の市場性は量的にも質的にも変化していることが考えられます。このような変化を正確に把握しておくことによって、例えば自社の売上動向の変化要因等の分析が可能となります。

国から公表された町丁別の国勢調査データについて、2010年調査結果と2015年調査結果の町丁をマッチングすると、**約22万件の町丁のうち8千件がアンマッチ**となり、そのままでは比較することが困難です。

当社では、国勢調査データの時系列比較を可能にするため、過去の調査結果データについても最新の町丁との対応づけを行い、集約・配分等の処理を行ったうえでデータベース化しています。

★ 国勢調査における欠測地域のデータ推計処理

国から公表される町丁別の国勢調査データには、実際には居住者がいるにもかかわらず、秘匿等により欠測となっている地域があります。中には、数千人規模の町丁のデータが欠測となっているケースもあり、これらを考慮せずにマーケティング分析を行うと、大きな過ちを犯す可能性があります。

最新の2015年国勢調査の場合、**約22万件の町丁のうち6千件以上が欠測**となっています。この欠測の多さは地域によって異なりますが、例えば京都府では約400件、北海道では約1400件みられます。

当社では、このような国勢調査における欠測地域のデータについても、市区町村別の集計結果や他の統計資料等を基に推計処理を行ったうえでデータベース化しています。