

2010 年国勢調査小地域統計のオープン・アトラスの作成・公開

佐谷岳穂・矢野桂司・中谷友樹・Alex Singleton・Christopher Brunsdon

Construction and Web Publication of Open Social Atlas based on the small area statistics of the 2010 Population Census of Japan

Takeho SATANI, Keiji YANO, Tomoki NAKAYA, Alex SINGLETON and Christopher BRUNSDON

Abstract: This paper aims to create *cho-cho*-based social atlas, based on the latest 2010 Population Census of Japan, which provides various variables about neighborhood attributes by *cho-cho* unit as the smallest area unit of the population census. Japan has 211,548 *cho-cho* units, besides 6,099 hidden areas, and the average *cho-cho* has 605 persons and 245 households. In order to automatically create the social atlas that includes a large number of census maps, we improved the R program developed for the UK 2011 Census Open Atlas Project (Singleton, 2013). As a result, we could create over seven hundred maps as PDF files for about 1,600 municipalities and open the maps to the public through the web.

Keywords: 2010 年国勢調査小地域統計 (the 2010 Population census of Japan), 社会地図 (social atlas), センサス・マッピング (census mapping), R, 日本 (Japan)

1. はじめに

1980 年代後半における欧米での GIS 革命以降、国勢調査などの官庁統計を用いたジオデモグラフィクス研究は、地理情報科学の主要な研究分野の 1 つを形成してきた (Harris, et al., 2005; Longley et al., 2011)。

1920 年代の米国シカゴ大学都市社会学者による社会調査基礎地図を嚆矢とする都市の居住構造に関する研究は、国勢調査データの整備に伴って、社会地区分析そして因子生態研究へと発展した (森川, 1975)。

GIS の普及と国勢調査の小地域レベルでの GIS データの公開により、様々な社会地図 (センサス・マップ) を容易に作成できるようになった。例えば、アメリカ合衆国では、米国センサス局が 1970 年センサスから導入した、DIME (Dual Independent Map Encoding), さらにそれを発展させた TIGER (Topologically Integrated Geographic Encoding and Referencing) により、ブロック単位でのセンサス・マッピングが可能である。そして、英国では、2001 年センサスから、英国全体で最小

空間単位である Output Area でのセンサス・マッピングが可能となっている。

日本では、1970 年国勢調査から国土数値情報と合わせて、地域メッシュ統計が整備された。しかし、その空間単位は、日本測地系の経緯度に基づいて区分された 3 次メッシュ (1km 基準メッシュ) で、人口集中地区に対して、4 次メッシュ (500m 分割メッシュ) での社会地図の作成が可能となった。例えば、倉沢進編『東京の社会地図』(倉沢, 1980) は、4 次メッシュを用いて、東京 23 区を対象とした、日本で初めての本格的な社会地図を作成している。また、倉沢・浅川 (2004) のまえがきには、東京の社会地図を作成するのに 10 年を超す大変な作業であったことが記されている。

1980 年代後半の GIS 革命を通して、Stan Openshaw (Abrahart et al., 2000) が GeoComputation を提唱し、コンピュータのメモリや処理速度の問題を克服し、地理情報科学の新たな展開を示唆したように、現在のコンピュータ環境であれば、小地域レベルの膨大な社会地図を容易に作成することが可能となった。

日本においても、1990 年国勢調査から基本単位

区（街区レベル）でのデータ提供を始めたが、本格的な小地域レベルでのデータの提供は、1995年国勢調査の町丁・字等集計からである。そして、10年ごとの大規模調査での小地域統計は、2000年国勢調査が初めてで、本研究が扱う2010年国勢調査小地域統計は、国勢調査としては4回目、大規模調査としては2回目のものである。

本研究の目的は、最新の2010年国勢調査小地域統計（町丁・字等集計）を用いて、日本全国の社会地図を作成することにある。しかし、小地域統計での社会地図に関して、膨大なGISデータを整備して、多くの地図を描くことは、大変な労力を要する。ここでは、日本の2010年国勢調査を対象とした、社会地図の自動作成によるオープン・アトラスを構築する。

2. 手法の概略

2.1 英国でのモデルケースについて

2013年2月にイングランド、ウェールズの2011年英国センサスのOutput Areaの統計表と境域データがOffice for National Statistics (2013)のWebサイトで公開されたことに伴い、著者の一人である英国リバプール大学のSingleton (2013a)が中心となって348のLocal Authority（日本の市町村にあたる行政区）に対して約380変数分の地図を生成する「2011 Census Open Atlas Project」を実施し、オープン・アトラスのシステムを開発した。

このプロジェクトの最大の特徴はオープン・データである英国センサスと、オープンソース・ソフトウェアであるRを利用することによって金銭的コストを最低限に留めるとともに、プログラム化することによって地図生成の過程を自動化し、時間的コストも大幅に削減している点にある。

2.2 統計データについて

本研究が用いるデータは、2010年10月1日に実施された大規模調査の国勢調査小地域統計（町丁・字等集計）である。

小地域統計の空間単位は町丁・字等によって定

義され、日本全土（竹島及び北方領土を除く）は228,286の区域に分割されている。また小地域統計で表章されている変数は669存在し、全人口128,057,352人、全世帯51,950,504世帯について人口、世帯・家族構成、住居、労働・就業、学歴、住所の移動、通勤・通学といった分野から調査・集計されている。特に、学歴と人口移動に関する変数は、大規模調査時のみ調査される。また、小地域統計のデータは2010年国勢調査から、全ての表章変数がWebサイトを通して無償で提供されるようになった。そこでは、小地域統計のデータは、第1表から第20表が都道府県ごとにダウンロードできる（第1表は基本単位区集計）。

これら小地域統計の町丁・字等にかかる境域データと一部の統計データはともに「e-Stat」と呼ばれる政府統計のポータルサイトより入手することが可能であり、本研究においても使用した。しかし、境域データは、市区町村ごとダウンロードし、そのデータを結合する必要がある、膨大な作業となる。

2.3 Rプログラムについて

「2011 Census Open Atlas Project」のRスクリプトは大きく分けて2部構成になっている。前半は変数グループごとにファイルが分かれている2011年英国センサスのOutput Areaの集計データをWebサイト上からダウンロードし、それらを1ファイルに統合する。後半は地図データを読み込み、集計データの関連付けを行い、RのGISライブラリを使用してLocal Authorityごとのセンサスマップの描画を行う。

3. 適用例

3.1 適用方針

「2011 Census Open Atlas Project」の特徴であったオープン・データ、オープンソース・ソフトウェアを利用した金銭的コストの削減と地図生成過程のプログラム化による時間的コストの削減は本研究でも方針として引き継いだ。しかし境域

データの整備については、一部、Microsoft Office 社の Excel や Esri 社の ArcGIS といった有償パッケージ・ソフトウェアを使用した。

英国版からの改善点として、①スケールバーの挿入、②組版ソフトウェア LaTeX を導入することによる、スケールバー、凡例、タイトルといった要素のレイアウト調整の簡易化、③同じく LaTeX による PDF ファイルの目次ページから対応する地図ページへのジャンプ機能の追加、が挙げられる。

3.2 集計データ、境域データの整備

「e-Stat」から入手した日本全国の小地域の集計データについてまず、「市区町村コード」「町丁字コード」「地域識別番号」を文字列結合した「GEO_CODE」を作成する。また集計データには町丁・字等単位、市区町村単位の合計値レコードと秘匿地域レコードが含まれているため、合計値レコードに関しては削除、秘匿地域レコードに関しては「GEO_CODE」を合算対象地域の「GEO_CODE」に統合する。

境域データに関しては行政界を基本としているため、海岸線や湖沼面といった水面上にもポリゴンが描画されてしまっている。そのため、データ上「水面調査区」と認識されるポリゴンを削除し、さらに国土数値情報の Web サイトより入手可能な「湖沼」のデータにより、内陸の水面部分にあたるポリゴンを削除する。また、「e-Stat」より入手できるポリゴンデータは「KEY_CODE」と呼ばれる集計データとの結びつきの役割を持つコードが付与されているが、この「KEY_CODE」が一意でないため、「KEY_CODE」が重複しているポリゴンを一つのポリゴンに統合し、「KEY_CODE」の付与されていないポリゴンを削除する。

上記で整備した集計データの「GEO_CODE」と、境域データの「KEY_CODE」とを互いに結びつけることで GIS 上でのデータ表現が可能となる。

3.3 R スクリプトの整備

まず分子・分母となる変数が明示された 2010 年国勢調査の全変数のリストを作成し、R スクリプトが読み込み可能な形式にする。R 上で、①集計データと変数リストの読み込み→各変数のパーセンテージ計算→異常値（分母 0 など）処理、②境域データの読み込み→代表地点名の挿入→市区町村ごとに各変数についての地図の描画（「classInt」パッケージの「classIntervals」メソッド内の「fisher」オプションによる 5 段階の自然分類法）→スケールバー、凡例、タイトルの挿入、③描画した地図の PDF 化、が実行され地図が自動生成される。

3.4 自動生成された地図

本研究で作成された地図は、東京都 23 区と、19 の政令指定都市（札幌市、仙台市、さいたま市、千葉市、横浜市、川崎市、相模原市、新潟市、静岡市、浜松市、名古屋市、京都市、大阪市、堺市、神戸市、広島市、岡山市、北九州市、福岡市）、および 1,610 の市町村（政令指定都市の区部と一部の市町村を除く）についての 485 変数分であり、枚数にして約 650,000 枚、データサイズは約 68.1GB である（図 1、2）。

これら自動生成された地図データは現在、Web サイトの「Census Atlas Japan」（Singleton, 2013b）において公開されており、PDF ファイルとして、自由に閲覧・印刷が可能である。

4. おわりに

本研究では、オープン・データである 2010 年国勢調査小地域統計の町丁・字等集計と、オープンソース・ソフトウェア R による社会地図自動描画システムを用いて、市町村別に町丁字等単位での社会地図を作成し、Web 上に公開した。小地域レベルの社会地図はジオデモグラフィクス研究の基礎となるもので、今後、様々な活用が期待される。

現時点では、階級区分や色パターンの変更ができないが、階級区分図のパラメータを任意に変え

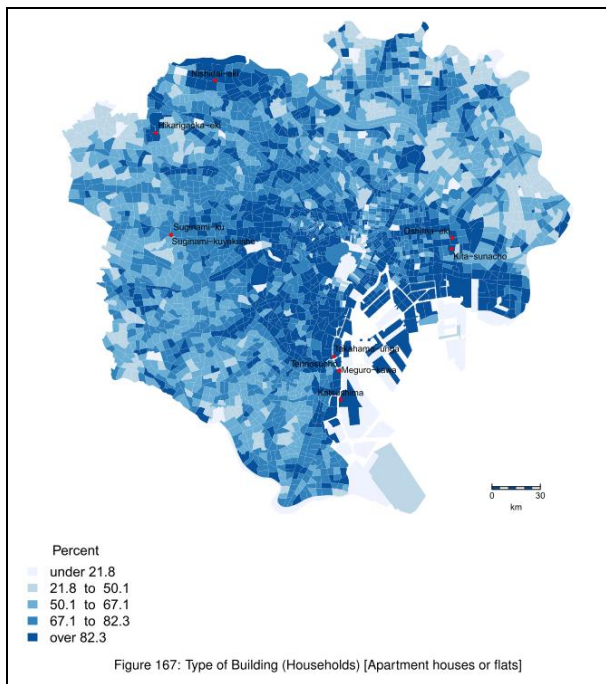


図1 東京都23区におけるマンション・アパートに住んでいる世帯の割合

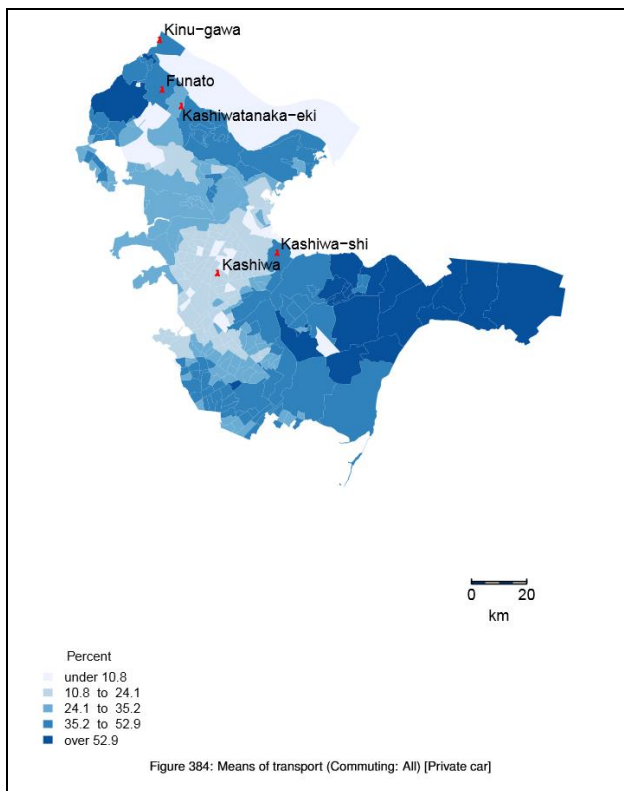


図2 柏市における自家用車で通勤している世帯の割合

る仕組みを導入することも可能である。さらに、この社会地図のGISデータを他のGISソフトに取

り込んで分析可能となるように、今後はPDFだけではなく、市区町村ごとでのShapeファイルや、Google Earthへの取り込みが可能なKMLファイルの自動生成と提供も検討したい。

謝辞

本研究は、文部科学省科学研究費助成金基盤研究(B)「GISベースの日本版センサス地理学の確立とその応用に関する研究」(代表者: 矢野桂司)によって行われた。

参考文献

- 倉沢進編(1980)『東京の社会地図』, 東京大学出版会
- 倉沢進, 浅川達人編(2004)『新編 東京圏の社会地図 1975-90』, 東京大学出版会
- 国土数値情報(2006)湖沼データ, <<http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-W09.html>>
- 森川洋(1975)都市社会地理研究の進展--社会地区分析から因子生態研究へ,「人文地理」, 27(6), 638-666.
- Abrahart, R. J., Openshaw, S. and SeeHarris, L. M. (2004) *GeoComputation*, CRC Press
- Harris, R. J., Sleight, P. and Webber, R. (2005) *Geodemographics, GIS and Neighbourhood Targeting*, John Wiley
- Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (2011) *Geographic Information Systems and Science (3rd Ed.)*, John Wiley
- Office for National Statistics (2013) Census 2011, <<http://www.ons.gov.uk/ons/guide-method/census/2011/index.html>>
- Singleton, A. D. (2013a) 2011 Census Open Atlas Project, <<http://www.alex-singleton.com/2011-census-open-atlas-project>>
- Singleton, A. D. (2013b) Census Atlas Japan, <<http://www.alex-singleton.com/census-atlas-japan/>>