

統計センターにおける 二次的利用拡大等に向けた取組

独立行政法人 統計センター
樁 広計

1

内容

- はじめに: 国勢調査と統計センターの最近の活動
- 公的統計マイクロデータの研究利用
 - ニーズ: 人間・社会研究における公的統計活用
 - 進行形の活動: 公的統計マイクロデータ活用のためのネットワーク基盤
- 公的統計を含むマイクロデータ分析の課題

2

統計におけるオープンデータの高度化

所属組織
独立行政法人
統計センターの
最近の活動
「統計を活かす」

JSTAT MAPは、
事業所に
多様な近傍を
定義し、性別年齢
別人口分布を
すぐに表示

ある研修：
浦安市の介護
施設配置の
あるべき政策？

統計データの提供方法を高度化し、新たな付加価値を創造するサービスや革新的な事業の創出などを支援する取り組みを、総務省統計局と連携し行っており、政府が取り組んでいるオープンデータの推進を先導。

API機能による統計データの提供

統計データを機械判読可能な形式で提供するAPI機能（Application Programming Interface）を提供中

利用登録者数 **3162人**
統計データへのリクエスト件数 **約2272万件**
(平成27年3月26日現在)

活用例1：利用者の情報システムにe-Statのデータを自動的に反映

活用例2：ユーザー保有やインターネット上のデータ等と連動させた高度な統計データ分析

開発支援情報も提供中

2014.10.31から運用開始

地図による小地域分析 (JSTAT MAP)

任意に指定したエリアによる集計や利用者が保有するデータの取り込み集計する機能などを提供

利用登録者数 **2334人**
ログイン件数 **約23000件**
(平成27年3月26日現在)

活用例1：任意に指定したエリアによる集計や、利用者が保有するデータと統計データを組み合わせ、集計結果を地図上で視覚的に把握可能

活用例2：選択したエリアの年齢構成等の基本的な分析結果のレポート作成

タブレット用アプリも提供中

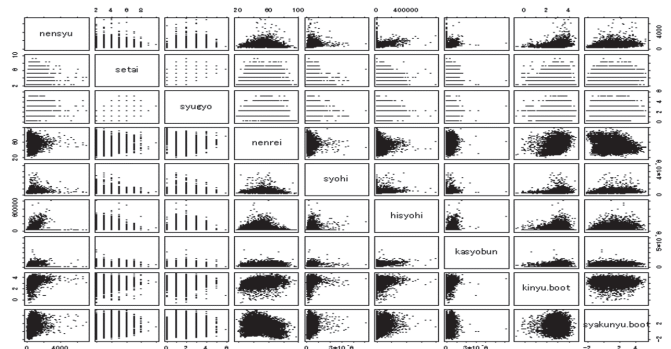
2015.1.20から運用開始

Needs: 人間・社会研究における 公的統計活用

人間・社会研究における公的統計活用

一例たりとも実データが存在しない
家計調査+貯蓄動向調査のイミテーションデータ
重回帰分析位ならば再現する(椿:1998)

- 国際競争力を低下させる日本の社会科学的研究
 - 1980年代欧米で大きな変化
 - 自国公的統計マイクロ(個票)データの研究利用が可能に
 - 経済・社会研究の中心が集計データからマイクロデータ分析に
 - 日本の経済学者の危機感
 - 科研費特定領域:1996-1999
 - 統計情報活用のフロンティアの拡大の総括的研究:
 - 松田芳郎(一橋大)
 - 13研究班の要請を集約:5省庁17調査「目的外使用申請」
 - データ処理センター活動→現行の**オーダーメイド集計制度**
 - リサンプリングデータの作成→現行の**匿名データ提供制度**
 - イミテーションデータの作成→現行の**教育用擬似データ制度**
 - データ処理結果および処理要求をデータの秘密保持を保ったうえで各研究班相互を計算機ネットワーク化して計算結果等の情報を流通させるシステムを開発
 - 今後の**リモートアクセス制度**
 - 日本評論社:マイクロデータ分析シリーズにより啓発
 - 新統計法下のマイクロデータ研究利用の先駆け研究



5

2009年統計法全面改訂

- 旧統計法(1947年)法の目的
 - 第一条 この法律は、**統計の真実性**を確保し、統計調査の重複を除き、統計の体系を整備し、及び統計制度の改善発達を図ることを目的とする。
- 新統計法第一条
 - この法律は、**公的統計が国民にとって合理的な意思決定を行うための基盤となる重要な情報**であることにかんがみ、公的統計の作成及び提供に関し基本となる事項を定めることにより、公的統計の体系的かつ効率的な整備及びその有用性の確保を図り、もって**国民経済の健全な発展及び国民生活の向上に寄与**することを目的とする。
 - オーダーメイド集計や匿名データの作成・研究目的などでの提供を可能とする**統計データの二次的利用の制度**設立

6

統計データの二次的利用促進に関する研究会

- 総務省政策統括官室：2007年スタート
 - 廣松毅(東京大学, 現情報セキュリティ大学院大学)座長
 - 内閣府統計委員会へのInput
 - 統計法改正に向けた二次的利用のスタイルと展開
 - **オーダーメイド集計・匿名データ・疑似マイクロデータ**の提供
 - 利用目的としての公益性
- 統計法公布後の活動
 - 提供方法：各国制度比較と日本の特異性
 - 目的外申請：個人情報・法人情報が付随したデータをセキュアな環境を持たない研究者が管理する可能性
 - オンサイト拠点
 - 目的外申請で得た個票をセキュアな監視環境下で分析
 - 探索的なモデリングを可能とする
 - しかし、各拠点ごとの人員・設備整備にかなりなコスト
 - リモートアクセスによるネットワーク形成：応用統計学会からの学会会議マスタープラン提案
 - 各拠点にはデータは置かない：監視は中央で一括
 - 統計データ・アーカイブの整備

7

匿名データの提供

一般からの申出を受け、利用要件を満たした申出者に対し、特定の個人又は団体等が識別できないように加工して作成した調査票情報の利用を一定期間認める制度。匿名データを利用することで、行政機関が作成していない統計表の作成のみならず、多変量解析など**マイクロデータ**に基づく実証分析を行うことが可能。

【利用要件】

- 統計の作成または統計的研究にのみ利用されること
- 学術研究目的または高等教育目的の用に供することを直接の目的とすること
- 学術研究の成果または高等教育の内容が公表され、社会に還元されること
- 匿名データが適切に管理されること

8

オーダーメイド集計

一般からの委託を受けて、利用要件を満たした申出者に対し、調査票情報を用いて集計を行い、その結果の提供を行う制度。オーダーメイド集計を利用することで、行政機関等が作成していない統計表に基づいた分析が可能。

【利用要件】

- 統計の作成または統計的研究にのみ利用されること
- 学術研究目的または高等教育目的の用に供することを直接の目的とすること
- 学術研究の成果または高等教育の内容が公表され、社会に還元されること

9

擬似マイクロデータの試行的提供

擬似マイクロデータは、我が国の行政機関が実施した統計調査の集計表から作成したマイクロデータ形式の擬似的なデータで、**大学等の教育機関における授業や演習**及び公的統計の二次的利用の際のテストデータなどの利用が可能であり、マイクロデータ利用者の裾野を広げ、公的統計の二次的利用の拡大を図るため無償で提供している。

【利用要件】

- 申請者及び利用者以外の者に利用させないこと
- 集計表から擬似的に作成したデータであるので、分析結果は実証研究の結果と見なすことができないことを理解すること
- 利用者アンケートを提出すること

10

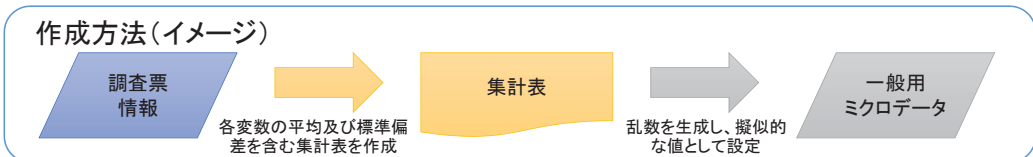
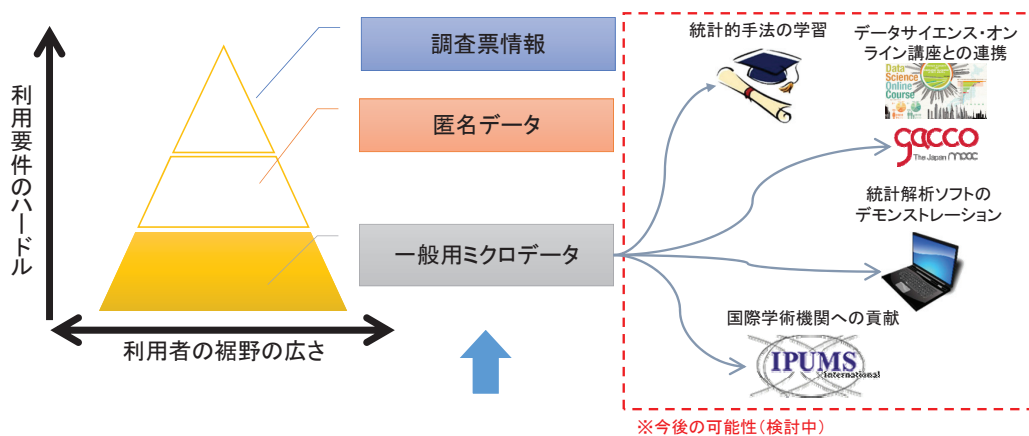
試行提供している擬似マイクロデータ (平成27年5月末現在)

- 平成16年全国消費実態調査に関する試行版の擬似マイクロデータ
 - 大規模データ (CSV形式のみ)
 - 約200項目: 世帯情報、収入、実支出(消費支出、非消費支出)、実支出以外、繰越金
 - 約3万2千レコード
 - 簡易データ (CSV形式 及び Excel形式)
 - 25項目: 世帯情報、消費支出
 - 約8千レコード

11

今後の予定

擬似マイクロデータの提供(無償)は、平成27年度末で終了する。
平成28年度からは、一般用マイクロデータの提供に移行する。
(利用制限を設けないマイクロデータを開発・提供する。有償か無償かは未定。)



12

学術研究機関等との連携協力の取組

統計センターでは、平成21年4月施行の統計法において創設された公的統計の二次的利用制度の充実と学術研究の発展を図っていくため、学術研究機関等と連携した取組を展開。公的統計の利用拡大に係る取組に賛同する法人と連携協力協定を締結し、公的統計の二次的利用に関する研究・開発、普及・啓発を推進するほか、これらの法人に**統計データアーカイブのサテライト機関の役割を担ってもらい、研究者等に向けた二次的利用のサービスの充実を図っている。**

法人の要件

1. 国立大学法人法に基づき設置された国立大学法人及び大学共同利用機関法人
2. 私立学校法に基づく学校法人により設置された私立大学
3. 独立行政法人通則法及び個別法の定めるところにより設立された独立行政法人
4. 法人税法別表第1に掲げる公共法人
5. 公益社団法人及び公益財団法人の認定等に関する法律により公益性の認定を受けた公益法人（特例民法法人を含む。）

連携協力事項

- 公的統計に関するデータアーカイブの運営に関すること（**施設基準に適合**）
 - ・匿名データの提供
 - ・オンサイト利用環境の提供など
- 公的統計の二次的利用に関する研究・開発
- 公的統計の二次的利用に関する普及・啓発
- 人材交流
- その他協定の目的を達成するために必要な事項

連携協力協定を締結している大学等

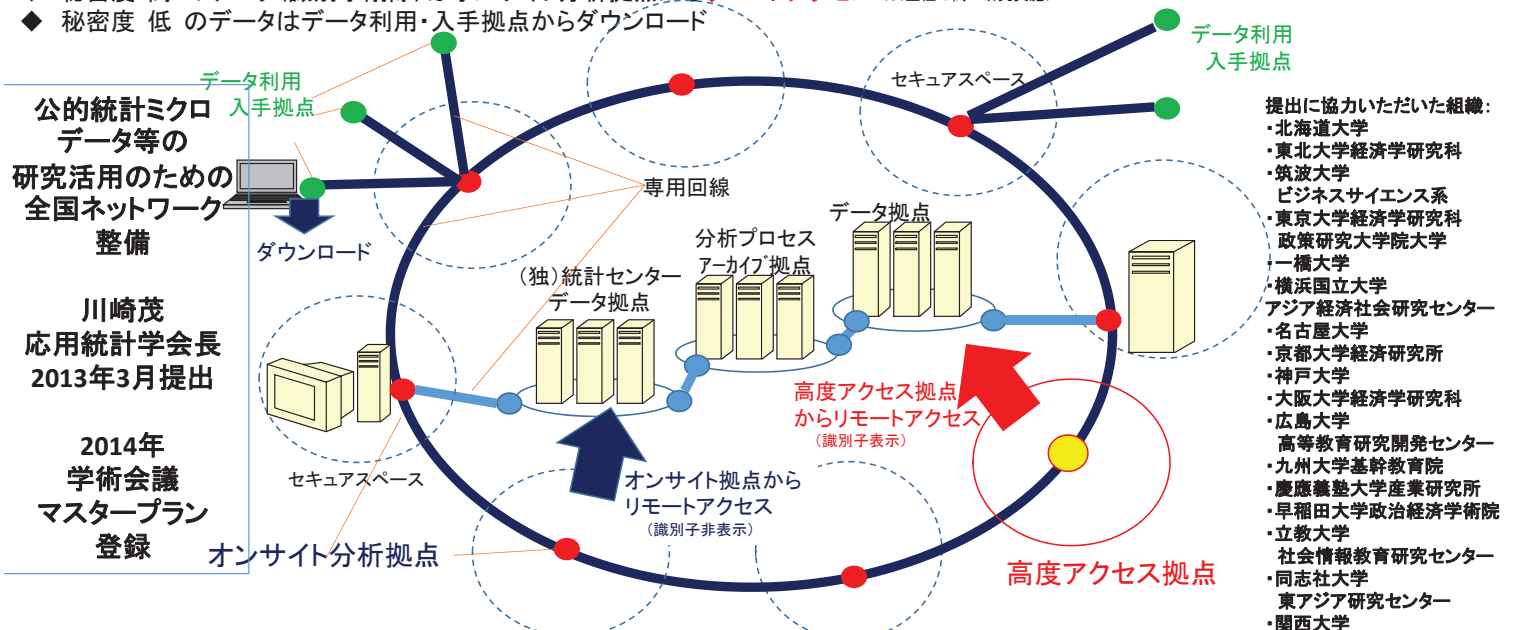
| サテライト機関名 | 匿名データの提供 | オンサイト利用環境 | 二次的利用に係るURL |
|---|----------|-----------|---|
| 国立大学法人 一橋大学 経済研究所附属社会科学統計情報研究センター | ○ | ○ | http://rciss.ier.hit-u.ac.jp/Japanese/micro/index.html |
| 国立大学法人 神戸大学大学院 経済学研究科 | ○ | — | http://www.econ.kobe-u.ac.jp/kuma/satellite/index.html |
| 法政大学 日本統計研究所 | ○ | — | http://www.hosei.ac.jp/toukei/micro/index.html |
| 大学共同利用機関法人 情報・システム研究機構 新領域融合研究センター（統計数理研究所） | ○ | ○ | http://www.rois.ac.jp/tric/tokumei/tokumei.html |

進行形の活動： 公的統計マイクロデータ 活用のためのネットワーク基盤

総務省「統計の二次的利用研究会」
川崎茂応用統計学会長(当時)「学術会議マスタープラン」提出
内閣府統計委員会の「公的統計の整備に関する基本計画
基本計画2014年3月25日閣議決定

15

- ◆ 秘密度超高のデータ(識別子付与)はごく一部の**高度アクセス拠点**からリモートアクセス<根幹を成す高度検証等に参照>
- ◆ 秘密度 高 のデータ(識別子削除)はオンサイト分析拠点から**リモートアクセス**<公益性の高い研究実施>
- ◆ 秘密度 低 のデータはデータ利用・入手拠点からダウンロード



※ データ拠点：2拠点 分析プロセスアーカイブ拠点：1拠点 高度アクセス拠点：4拠点
オンサイト分析拠点：18拠点 データ利用・入手拠点47拠点

16

公的統計の整備に関する基本計画への組み込み

- 内閣府統計委員会
 - 2013年10月30日総務大臣より諮問, 2014年1月31日答申
- 2014年3月25日閣議決定
- 調査票情報等の提供及び活用については、
 - セキュリティレベルや調査票情報等の匿名性の程度に応じた利用形態ごとの特性
 - 諸外国における取組状況等を総合的に勘案した上
- 法制度上の整理を含め、以下の取組を行う
 - オーダーメイド集計における利用条件の緩和に向けた検討
 - 調査票情報の提供におけるリモートアクセスを含むオンサイト利用やプログラム送付型集計・分析の実現に向けた整理・検討
 - 匿名データの作成及び提供における提供対象統計調査の種類や年次の追加等によるサービスの充実
- その際、効率性及び利便性の観点から、政府一体として一元的な取組を推進する。
 - 主担当総務省政策統括官, 府省横断の中核機関:(独)統計センター
 - 二次的利用研究会での検討⇔リモートアクセス拠点形成⇔大学などの二次的利用フォーラム形成(今年度中)
 - 総務省関連概算要求

17

(参考) 調査票情報のオンサイト利用 (現在検討中)

- オンサイト利用については、長年にわたる課題であったが、IT技術の進展等を踏まえ、平成28年度を目途に、リモートアクセスを活用したオンサイト利用の試行を開始する。今後も、総務省政策統括官(統計基準担当)は制度面の検討・関係府省と調整するとともに、総務省統計局と(独)統計センターが連携して技術面を検討する。
- 具体的な試行については、統計局において、学界や各府省の協力を得つつ、統計局の統計調査の調査票情報を主たる対象に、(独)統計センターを政府共通の基盤である中央データ管理施設の管理者とし、オンサイト施設を整備する。



18

公的統計を含む マイクロデータ分析の課題

19

政府統計部局の役割 研究の障害除去：データがない、使いにくい

- 提供可能な公的統計マイクロデータの範囲拡大
 - 匿名データ・オーダーメイド集計の範囲拡大
 - 事業所・企業統計などの匿名化困難
- 適切なメタデータの付与
 - 現状：統計局：CSVあるいはXMLでの提供
 - 研究者の負担
 - 研究に用いる複数データセットの変数結合 * 国内外の類似データとの結合
 - 適切なメタデータの付与：LOD (Linked Open Data) 形式によるデータ提供
 - LOD：関連するデータ同士が相互に結びついているデータ
 - 統計オープンデータモデル事業
 - 福井県・統計局・統計センター協働事業
 - 統計局提供：国勢調査 ¥ 社会・人口統計体系等の統計データ
 - 福井県提供各種行政データ

20

公的統計部局の課題: 提供データ範囲の拡大:

統計センターにおける匿名データ及びオーダーメイド集計の提供

提供データ一覧(厚生労働省が別途国民生活基礎調査提供)

| 調査名(年次等) | 匿名 | オーダー | 調査名(年次等) | 匿名 | オーダー | 調査名(年次等) | 匿名 | オーダー | 調査名(年次等) | 匿名 | オーダー |
|-----------------|----|------|------------------|----|------|---------------------------------------|----|------|------------------------|----|------|
| 国勢調査 | | | 住宅・土地統計調査 | | | 社会生活基本調査[調査票A] 生活時間編・生活行動編 | | | 家計調査 | | |
| 昭和55年 | | ○ | 昭和53年 | | ○ | 昭和56年 | | ○ | 昭和56年1月～ 平成26年12月 | | ● |
| 昭和60年 | | ○ | 昭和58年 | | ○ | 昭和61年 | | ○ | 家計消費状況調査 | | |
| 平成2年 | | ○ | 昭和63年 | | ○ | 平成3年 | | ○ | 平成14年1月～26年12月 | | ○ |
| 平成7年 | | ○ | 平成5年 | | ○ | 平成8年 | | ○ | 消費動向調査 | | |
| 平成12年 | ○ | ○ | 平成10年 | | ○ | 平成13年 | | ○ | 平成16年4月～27年3月 | | ○ |
| 平成17年 | ○ | ○ | 平成15年 | | ○ | 平成18年 | | ○ | 企業行動に関するアンケート調査 | | |
| 平成22年 | | ○ | 平成20年 | | ○ | 平成23年 | | ○ | 平成18年度～26年度 | | ○ |
| 就業構造基本調査 | | | 平成25年 | | ● | 社会生活基本調査[調査票B] 生活時間編 | | | 賃金構造基本統計調査 | | |
| 昭和54年 | | ○ | 全国消費実態調査 | | | 平成13年 | | ○ | 平成18年～26年 | | ○ |
| 昭和57年 | | ○ | 平成元年 | | ○ | 平成18年 | | ○ | 学校基本調査[高等教育機関編] | | |
| 昭和62年 | | ○ | 平成6年 | | ○ | 労働力調査[基礎調査票] | | | 平成20年度～26年度 | | ● |
| 平成4年 | ○ | ○ | 平成11年 | | ○ | 昭和55年1月～63年12月 | | ○ | 学校基本調査[初等教育機関編] | | |
| 平成9年 | ○ | ○ | 平成16年 | | ○ | 平成元年1月～23年12月 | | ● | 平成20年度～22年度 | | ○ |
| 平成14年 | ○ | ○ | 平成21年 | | ○ | 平成24年1月～26年12月 | | ○ | 建設工統計調査 | | |
| 平成19年 | | ○ | | | | 労働力調査[特定調査票] | | | 平成21年4月～27年3月 | | ○ |
| 平成24年 | | ○ | | | | 平成14年1月～26年12月 | | ○ | | | |

○印は提供中、●印は(複数年存在するものについては最新年(度)分を)平成27年度中に提供予定

～ 利用可能な統計調査は、今後、さらに拡大予定 ～

公的統計調査は情報という形の税徴収

- 国勢調査など基幹統計調査に協力しないとどのようなことになるか？
 - 罰則規定がある
 - 国民の義務の一つ
- 租庸調+「報」
 - 税を金銭で支払う
 - 税を労働で支払う(兵役:戦前は義務)
 - 税を物品で支払う(今は無い)
 - **税を情報で支払う:報告(情報には明らかな経済的価値がある)**
 - 統計以外に、届け出義務を課し徴集する情報も多数ある
 - **国をEvidenceに基づいて運営するのに必要な個人情報, 法人情報**
 - 個人情報保護の枠外
 - 目的外の利用を公益性のある統計研究等に制限
- 「報」を適切に政策利用できないのは逆に国家の怠慢

(参考) 匿名データの利用実績例

| 利用目的 | 調査名 | 研究の名称 |
|--------|----------------------------------|---|
| 学術研究目的 | 社会生活基本調査 | 正規雇用者における平日の労働時間と休息時間 —「社会生活基本調査」マイクロデータによる分析— 生活行動からみる高齢者の行動特性について —社会生活基本調査の匿名データを用いて— 子供のいる世帯における夫と妻の2次活動時間の差異について —社会生活基本調査の匿名データを用いて— 趣味・娯楽活動の時間について 個人・世帯属性と行楽・観光旅行行動の関係 |
| | 全国消費実態調査 | 等価尺度の推計と比較—消費上の尺度・制度的尺度・OECD尺度— 『季刊社会保障研究』Vol.48 Spring 2013 No.4 所得格差変動の年齢階級別要因分解 :全国消費実態調査マイクロデータを用いて |
| | 全国消費実態調査 住宅・土地統計調査 | 持家取得における既婚女性の就業の役割 |
| | 就業構造基本調査 | 転職経験および転職理由と転職希望意識との関連について —就業構造基本調査匿名データによる統計分析— 若年者就業率における賃金弾力性の推定 女性事務職の賃金と就業行動 —男女雇用機会均等法施行後の三時点比較— 税負担と労働供給 —「日本労働研究雑誌」No.605 2010.12 — 若者の有業・無業状態における属性の考察 女性の働き方と少子化に関する考察 |
| 高等教育目的 | 就業構造基本調査 全国消費実態調査 社会生活基本調査 | 一橋大学大学院経済学研究科「演習」(労働経済学Ⅰ) |

論文等: <http://www.nstac.go.jp/services/jisseki.html>

23

(参考) オーダーメイド集計の利用実績例

| 調査名 | 研究の名称 |
|----------|---|
| 国勢調査 | 岡山大学大学院社会文化科学研究科紀要第35号(2013.3) ・2005年国勢調査にみる在日外国人女性の結婚と仕事・住居 ・2005年国勢調査にみる外国人の教育 ・2005年国勢調査にみる在日外国人の仕事 Access to childcare and the employment of women with preschool-aged children in Tokyo 看護人材の就業率の推移 —再検討した潜在者数推計方法による結果から— 地域別経済指標に基づくSDモデルの開発 在日外国人の仕事 —2000年国勢調査データの分析から— 家族・ジェンダーからみる在日外国人 —2000年国勢調査データの分析から— 在学率と通学率から見る在日外国人青少年の教育 —2000年国勢調査データの分析から— エリア・サンプリングにおける問題点 |
| 全国消費実態調査 | 新・家計消費論 —高齢層が支える都市部消費— |
| 就業構造基本調査 | 近年における都道府県別貧困率の推移について —ワーキングプアを中心に— |

論文等: <http://www.nstac.go.jp/services/jisseki.html>

24

参考資料

- 匿名データの利用に関するFAQ:
 - <http://www.nstac.go.jp/services/faq-anonymity.html>
- 擬似マイクロデータの利用に関するFAQ
 - <http://www.nstac.go.jp/services/faq-gijimicro.html>