

統計等データの提供等に関するユーザーからの要望・提案募集
(令和2年12月22日から令和3年3月31日受付分) 検討結果一覧

- | | |
|---------------------|-----|
| 1. 具体的な統計等データの提供要望 | p.1 |
| 2. 二次利用の手続等への要望 | p.2 |
| 3. データ収集・処理に当たっての要望 | p.3 |
| 4. 調査対象者の負担軽減の要望 | p.3 |

1. 具体的な統計等データの提供要望

No	要望・提案名	具体的な要望・提案の内容とその理由	検討結果の内容
1	税務統計の活用	所得税、消費税にかかる個表のデータを集計したものを公表してほしい。今時のコロナの影響を見る際に法人統計だけではカバー率が低く飲食業などへの影響が明確に捉えられない。また12月に公表される法人標準調査は欠損企業の比率がわかる貴重なデータだが欠損企業についての詳細(業種、資本金別など属性別のデータ 継続年に関するデータ)くつかつて税調で個表を再集計したものが公表された)も合わせ公表してほしい	所得税及び消費税の申告書等のデータを集計・公表しているものとしましては、国税庁が公表しております「国税庁統計年報」におきまして、それぞれ統計表(申告所得税、源泉所得税、消費税)として公表しております。 また、例年5月頃に公表しております「会社標準調査」におきましても、「統計表第3表 繰越欠損金」におきまして、「資本金階級別」「業種別」に欠損法人に係る各種計数を公表し、継続年に関するデータにつきましては、長期時系列データとして公表しておりますので、ご参照ください。
2	意図的な情報の非開示の中止	現在の新型コロナウイルスの患者の情報がオープンでない、外国の方々を日本の公費で治療されているとか、日本人には生活保護のハードル高くて、韓国の人たちは働いているのに生活保護あげてるとか??正確な情報開示をお願いします	現在、厚生労働省では新型コロナウイルス感染症に関して、ホームページ(https://covid19.mhlw.go.jp/)において、データを公開しています。上記ページでは、性別・年代ごとの陽性者数等についても公開しています。
3	訪日外国人観光客アンケート調査の個票データについて	観光庁が行っている訪日外国人観光客に対するアンケート調査(消費動向調査など)について、個票データの利用を私立大学の研究者にも認めていただきたいと思います。国別など、観光客の基本的な属性別での分析結果は公表されていますが、訪日観光客においても個人旅行の比率が高まっており、国別などの外形的属性だけではニーズを的確に把握することができません。観光行動や消費行動のスタイルごとに分析するなどが必要ですが、これについては個票データを利用し、さまざまな統計的分析をやってみて、有効なものを採り出す必要があります。現在、国立大学を含む公的機関であれば利用できるのですが、観光関係の研究者は私立大学の比重が高く、私立大学の研究者が個票データを利用することで、今後観光客の積極的誘致が再び可能になった際に役立つ有効な研究が促進できると思われます。	統計法に規定する基幹統計及び一般統計につきましては、同種の統計調査の実施を抑制することができ、その結果として被調査者の負担軽減に資すること、社会の価値観やニーズが多様化する中、調査実施者があらかじめ想定していなかった統計の作成や統計的分析が公益に資する場合があることから、同法及び同法施行規則に基づき調査票情報の提供を行っているところです。 ご指摘の「官学(国立大学法人)」については、統計法第33条第1項に基づく提供に該当しますが、国立大学法人が組織として調査票情報の提供を受け、利用するような場合を念頭に提供を行っており、大学に属する個々の研究者が自らの研究のために利用しようとする場合については、私学と同様の扱いとなります。 また、「私学(私立大学)」については、学術研究の発展に資する統計の作成等又は調査票情報の提供を受け行うことについて相当の公益性を有する統計の作成等を目的とする者に対し、オンライン施設(※)を利用する形で、統計法第33条の2(令和元年5月改正にて新規追加)に基づく調査票情報の提供を行うことができます。 (※)独立行政法人統計センターと連携する大学や行政機関等に設置された情報セキュリティが確保された環境で、許可を受けた研究者が調査票情報を用いて、独自の集計・分析を行うことができる施設／専用室(手数料の支払いを要する。)
4	後期高齢者医療に関する市町村別データの開示	後期高齢者医療に関する市町村別データの開示が不十分と思われる。国で収集されていると思われるので、比較検討がなされるよう、国民に広く開示いただきたい。	後期高齢者医療制度に関する市区町村別データにつきましては、現在、厚生労働省又はe-statのホームページ上で、 ①後期高齢者医療制度被保険者実態調査報告(https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00450388&tstat=00001044907)において、市町村別・均等割軽減状況別の被保険者数、均等割軽減額を、 ②医療保険データベース(https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/database/index.html)において、地域差分析を公表しており、その中で後期高齢者医療制度の市町村別にみた一人当たり医療費等を お示しており、市町村別データの開示は一定程度行っているものと認識しているところです。
5	「所得再分配調査」のマイクロデータの公開	我が国の再分配の実態を把握するためには、横軸に「再分配前の所得」、縦軸に「再分配後の所得」ととり、所得再分配調査のマイクロデータをプロットし、国民に開示することが重要に思います。どのプロット点が誰のものか特定できないように開示することは可能であり、個人情報の保護を図りつつ、これが可能となるデータをエクセル形式で公開をお願いします。	所得再分配調査では、階級等の一定の集団で分析することを目的としており、ご指摘のマイクロデータをプロットするような公表については今後も予定していません。 また、統計法(平成19年法律第53号)に基づく統計調査の調査票情報については、統計調査に対する国民の信頼確保等の観点から、同法第40条において、特別の定めがある場合を除き、目的外利用が禁止されています。 一方、この特別の定めの一環として、同法第33条において、公的機関が行う統計の作成等と同等の公益性を有する統計の作成等を行う者に調査票情報を提供することができるとされており、提供条件に合致すれば、一定の手続きを経た上で、調査票情報を利用することが可能となっております。
6	発電事業者の時間別・事業所別の二酸化炭素排出量に関するパネルデータの整備・提供依頼	2015年にいわゆる「パリ協定」が採択され、我が国でも、菅首相が所信表明演説(2020年10月)において「2050年までに温室効果ガスの排出量を実質ゼロ」とする公約を掲げ、パリ協定の目標達成に向けた取組みが本格化するものと期待されておりますが、目標をより効果的・経済効率的に達成するための具体的な施策に関して、国内における経済学的な研究成果・知見は、極めて脆弱と言わざるを得ません。そのような研究の最大のネックとなっているのは、欧米において公的に提供されている環境・経済データと比べ、明らかに精度・粒度が劣っていることです。例えば、米国EPAのAir Market Data(amp.epa.gov)では、発電を行っている事業所の発電量および汚染物質・二酸化炭素排出量が「事業者毎・日毎・時間毎に」簡単にDLできるようになっているだけでなく、事業所自体の識別データが提供されているため、EIAによって提供されて経済データやUSGSによって提供されているGISデータとの統合が容易に可能となっています。このようなデータは研究者にとって有益なだけでなく、日本の炭素マネジメントの実効性・経済効率性を高めるためにも極めて有益です。例えば、何らかの形でカーボン・プライシングが導入された場合、どの時間帯にどの事業者から電力を購入することが最もCO2排出量を削減することができ、企業にとって最も低コストの調達となるのかを明確に把握することができます。また、消費者にとっても、例えば、電気自動車を利用する際に、どの時間帯・地域での充電が最も経済的なのか等が把握できます。更に、米国では既に、このようなデータにAIを活用することで、炭素マーケット・ビジネスの事業化(例、 www.watttime.org)が進められており、AI/IO分野と環境分野の共進化に寄与しています。具体的には、以下のようなデータの整備・ご提供をお願い致します： ①各発電事業所の事業所レベルの(他の経済/GISデータとの統合が可能となるような)識別情報・各事業所の「日毎かつ時間毎の」Load ②「日毎かつ時間毎の」利用された発電源の詳細 ③「日毎かつ時間毎の」汚染物質・二酸化炭素の排出量・これらの長期間のパネルデータ・各事業所の親会社情報・雇用人数・保有電源・生産要素投入量などの経済データ(こちらは年ベースで可)	発電事業者の一覧は資源エネルギー庁のHP(https://www.enecho.meti.go.jp/category/electricity_and_gas/electricity_measures/004/list/)で公表しておりますが、「日毎かつ時間毎の」二酸化炭素の排出量、これらの長期間のパネルデータ、各事業所の親会社情報・雇用人数、「日毎かつ時間毎の」利用された発電源についてはデータを保有しておりません。 事業所レベルの識別情報や保有電源のデータなどは、競争情報であることや公表目的が漏れ出されていないといった課題もあり、ご提供することは困難です。 「汚染物質データ」については、大気汚染防止法において、工場及び事業場における事業活動等による大気汚染に關し、国民の健康を保護するとともに生活環境を保全する目的で、同法に定める「ばい煙発生施設」を規制対象とし、また、当該施設から排出される大気汚染物質の排出状況を把握するために大気汚染物質排出量総合調査を実施しています。当該調査は、統計法に基づく統計調査のため、調査によって得た情報を、誰がどのような回答をしたか特定できる形で公表することができます。 なお、大気汚染物質排出量総合調査の結果については、以下のサイトで公表されており、データの入手は可能です。 https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00650202&tstat=000001014479

2. 二次利用の手続等への要望

No		要望・提案名	具体的な要望・提案の内容とその理由	検討結果の内容
1		個票データ二次利用手続きの簡素化	各種統計の個票データを二次利用する場合、例えばIPUMSなどと比べて、いまだに煩雑な手続きが必要です。こうした利用のしにくさは、日本の定量的分析を大きく遅らせる要因となり、政策分析を滞らせることになると考えられます。せめて手続きはすべてオンラインで完結するようにして、1%程度の標本抽出を行い匿名化処理をしたデータは、教育研究目的であれば自由にいつでも使えるようにしていただきたいと思います。	統計法においては、調査票情報を匿名処理した匿名データについて、学術研究の発展に資すると認められる場合その他の一定程度の公益性が認められる統計の作成又は統計的研究を行う場合に、一般からの求めに応じて提供することができるとされております。これは、匿名データが、秘密保護のための措置が施されているものではあっても、集計された統計ではないことから、統計制度に対する国民の不安を招かないようにするために一定の要件を科しているものであり、ご理解をいただけますと幸いです。 なお、統計センターでは、総務省統計局から委託を受け、統計演習など教育用に利用可能な一般用マイクロデータの無償提供を行っていますので、適宜ご利用ください。 https://www.nstac.go.jp/services/ppan-microdata.html 二次の利用に係る申出等の手続については、署名・押印の見直しを図ることにより、電子メールによることを可能とするなど利便性向上に順次取り組んでいるところです。 また、統計データの二次的利用につきましては、（磁気媒体の貸し渡しではなく）より探索的な研究が可能となるオンサイト利用であれば、事前に提出する申出書に添付する詳細な集計表様式は不要とし、手続の簡素化を図っております。マイクロデータ利用ポータルサイトにおいて手続をご紹介しておりますので、是非ご利用ください。 https://www.e-stat.go.jp/microdata/
2		統計等データの提供のための手続に関すること	データ担当部署の方々には日ごろからご多忙の中、大変な作業をして頂いており、心から感謝申し上げます。そのうえで、以下の点、限られた研究時間を有効に使うために手続等を簡素化して頂けるよう要望する。1)転写様式(申請した統計において必要な変数(項目)に丸をつける)と集計様式(分析の結果を表示する図や表の作成)を現状作成し、提出しているが、転写様式に丸がついていて、集計様式に掲載されていない変数は「転写様式でなぜ丸がついているのか」などの問い合わせがある。データに触る前に分析の詳細が完全にイメージできていることは少なく、分析をしている段階で方向性が変わる可能性は十分にあり得る。そのため、基本的には申請した統計の全変数を申請できるものとし、転写様式と集計様式の提出の双方を廃止してほしい。(研究の内容を示した研究計画書は別に提出しているため。)2)研究室のレイアウトの提出も廃止してほしい。(少なくとも勤務先に変更がなければ廃止してほしい。)以下は、申請後のデータ提供方法などに関する要望である。3)データは省庁により(例えば経済産業研究所など)はセキアUSBで提供されるが、セキアUSBで使える機能を増やしてほしい。セキアUSBは基本的に一切のコピー処理ができないよう設定されているため、統計分析ソフトから結果をアウトプットし、それをエクセルなどに貼り付けるといふ単純処理ができないなど、分析上、不都合な問題が発生している。(一方で研究所に所属する研究員はこうした処理が可能と聞いている。)4)各個票を可能な限りバネル化した状態で提供してもらえると助かる。申請の度にバネル化作業を一から各研究者が行い、プロジェクトが終了するとそのデータは消去、返却する。こうした作業を政府が行うことで、研究の効率が格段上がると思われる。5)統計法が改正されたことにより、データの利用可能期間が短くなった模様。(例えば、経済産業研究所では研究プロジェクト期間終了日の1か月前を目途に終了することの指示があった。)申請から承認、データ入手までかなりの時間を要するため、研究者が実質利用可能な期間が短くならないような措置を希望したい。	1) 統計データの保護の観点から、統計的研究に必要な最低限の調査票情報を提供することとしており、そのためには提供の審査に必要な書類を提出いただく必要がありますので、その点をご理解いただければと存じます。 2) 統計データの利用者は、提供を受けた統計データの適正管理措置として、統計データを取り扱う区域を特定し、特定された区域の立入りを制限するための措置を講ずることが必要となります。そのため審査に必要な情報を提供いただくことについてはご理解いただきたいと思います。 3) 各府省における個別の事情については承知していませんが、一般論としては、統計データを電磁的記録媒体により提供する場合、情報漏えい防止の観点から当該統計データの内容等に応じた暗号化やパスワードの付与など、必要な措置を講ずることとしております。 4) 二次的利用にあたり、研究に必要な処理等については、利用者において実施していただくようお願いいたします。 5) データの利用可能期間については、調査票情報の提供に関するガイドライン等において定めており、適正な運営に努めて参りたいと考えます。 統計データの二次的利用につきましては、（磁気媒体の貸し渡しではなく）より探索的な研究が可能となるオンサイト利用であれば、事前に提出する申出書に添付する詳細な集計表様式は不要とし、手続の簡素化を図っております。マイクロデータ利用ポータルサイトにおいて手続をご紹介しておりますので、是非ご利用ください。 https://www.e-stat.go.jp/microdata/
3		二次利用手続きの際の利用項目書き出しの省略	要望：二次利用手続きに際しては、その統計に含まれる提供可能な項目をすべて提供するようにし、利用する項目をいちいち書き出す必要をなくしてほしい理由：個票データの二次利用申請の際には利用する質問項目をすべて書き出したうえで研究計画書にその内容を盛り込むことが要求されるが、調査によっては毎年微妙に質問項目が異なるなど、この作業にかなりの時間を取られる。また、往々にして研究の初期段階では最終的にどの変数が必要になるのかははっきりとは分からないので、なるべく多くの変数をカバーするような研究計画を作成しなければならず、こちらも相当な手間がかかる。審査する側も、利用する質問項目がすべて研究計画に盛り込まれているかの確認作業にかなりの労力を割いているだろうと予想される。一方で、二次利用の審査に通るような研究者に対して、実際の分析では使わない項目を提供することによって生じる不利益はほとんどない。匿名性の担保に関わるような変数以外は原則すべて提供することによって、特段の問題を生じることなく申請者と審査担当双方の負担を減らすことができると思う。	統計データの保護の観点から、統計的研究に必要な最低限の調査票情報を提供することとしており、そのためには提供の審査に必要な書類を提出いただく必要がありますので、その点をご理解いただければと存じます。統計データの二次的利用につきましては、（磁気媒体の貸し渡しではなく）より探索的な研究が可能となるオンサイト利用であれば、事前に提出する申出書に添付する詳細な集計表様式は不要とし、手続の簡素化を図っております。マイクロデータ利用ポータルサイトにおいて手続をご紹介しておりますので、是非ご利用ください。 https://www.e-stat.go.jp/microdata/
4		リモートアクセスの導入	政府はオンサイトの拡充を図っているが、実はオンサイト施設の端末から分析サーバにリモートアクセスする形になっている。端末を置く場所はオンサイト施設に限定される必要はなく、セキュリティ環境が担保されていることを条件に、行政機関や大学研究室などからのリモートアクセスも許すことが望ましい。あるいは行政機関の特定の個室や大学の個別研究室をオンサイト施設として認定することを積極的に進めるべき。	オンサイト施設につきましては、今後も利用可能な施設を順次拡充していく予定としているところ、入退室管理や利用者の利用状況のチェック、成果物の持ち出し審査など、データの取扱いのみならず物理的な適正管理要件も満たすべきものであるため、指定された場所での利用に限定するものとなっておりますことをご理解いただければと存じます。

3. データ収集・処理に当たっての要望			
No	要望・提案名	具体的な要望・提案の内容とその理由	検討結果の内容
1	機械判別可能な統計データの作成方法について	ニュージーランドにおけるの機械判読可能なデータ(csvファイル)を参考に、機械判別可能な統計データの作成をお願い致します。 https://www.stats.govt.nz/large-datasets/csv-files-for-download/	お問合せのありました件については、各国の提供データも参考に、全省庁共通の方針を策定しており(一部の内容については決定し、各府省合意済み(※))、今後も、機械判読可能なデータとなるよう検討してまいります。 ※統計表における機械判読可能なデータ作成に関する表記方法 (https://www.soumu.go.jp/menu_news/s-news/01toukatsu01_02000186.html)

4. 調査対象者の負担軽減の要望			
No	要望・提案名	具体的な要望・提案の内容とその理由	検討結果の内容
1	シームレスなデータ活用について	統計データ収集は、利活用を具体的にすることから行わないと、政府統計オンライン調査(https://www.e-survey.go.jp/)のような省庁縦割りの無秩序且つ大量の調査依頼になり、受け手の企業は答えたくなくなる。統計法に基づく調査依頼で義務だからという記載があるが、答えるにも時間がかかり、スタッフの貴重な時間＝コストをかけたとは思えない。方やデータ収集は、正確かつ低コストでなければならず、政府の託業者含めて企業が人手をかけてデータを収集提供する今のスタイルは、タイムリーでもなく正確でもなく、高コストである(コロナ禍にも関わらず、大量の封書が会社へ届く)。各企業の決算期も異なり当社もそうだが上場している企業であれば、開示できるタイミングもある(委託業者スタッフが閲覧できるとなると報告したくない)。殆どの調査内容は、企業が利用する会計ソフトなどから吸い上げて済む内容も多い。吸い上げられる機能を追加し(クラウドならなお簡単にできる)、データを吸い上げれば事足りる。また、企業情報(住所や担当者情報などの個人情報含む)も封書で送られてきており、テレワーク中に盗難による漏洩リスクもあることを考えると、本当にデータ収集の為に企業情報の提供が必要なのかを再度検討いただきたいと思います。景況感などは、過去の膨大なデータからAIによる未来予測と確率論で業界毎にトップ数社の企業の景況感をヒアリングしAIとの予測との比較を示せばよいと考えます。紙による調査をWebで回答できるようになっていても、多くの人的のパワーを使っている時点で、デジタル化には程遠く、ムダな作業であり、誰も幸せになっていないと考えております。是非とも改善をお願いします。	公的統計は、統計法の中で「国民の合理的な意思決定の基盤となる重要な情報」と位置付けられており、政府全体として、その体系的・効率的な整備と有用性の確保を図っております。 負担の軽減については、統計調査を新たに実施したり調査内容の変更を行う前に総務省が審査を行い、他の統計調査との重複が合理的な範囲を超えていないか、調査事項を代替できる行政情報がないかなどについて確認を行っております。 統計作成における行政情報やビッグデータ等の利活用については、業務の効率化や負担の軽減に資するため、政府として積極的に取り組んでいます。具体的には、総務省においてビッグデータに関する産官学の連携会議を開催し、ビッグデータが有する偏りやノイズなどの課題への対応や官民の優良事例を共有するなど、利活用の推進を図っております。 引き続き、回答される皆様の負担の軽減を図るとともに、品質の高い公的統計を提供できるよう努めてまいります。ご理解のほど、よろしく申し上げます。