



独立行政法人

統計センター

資料5

オンデマンドによる 統計作成機能・方策の研究 (中間報告)

平成26年3月

独立行政法人統計センター

オンデマンドによる統計作成機能・方策の研究

- ◆ 公的機関や学術研究などの利用において、利用者が調査項目を選択するだけで統計結果を自動的に出力する、新しい形の統計サービスを研究中
- ◆ これにより、既存の結果表にない任意の多重クロス集計が出力可能になり、学術研究をはじめとする多様なニーズに対応

従来の
形態



政府統計の総合窓口
あらかじめ作成した
統計表を提供

必要な統計表を探して
ダウンロード



統計利用者

利用者が自らのニーズに合わせ
希望する項目を組合せ

オンデマ
ンド集計



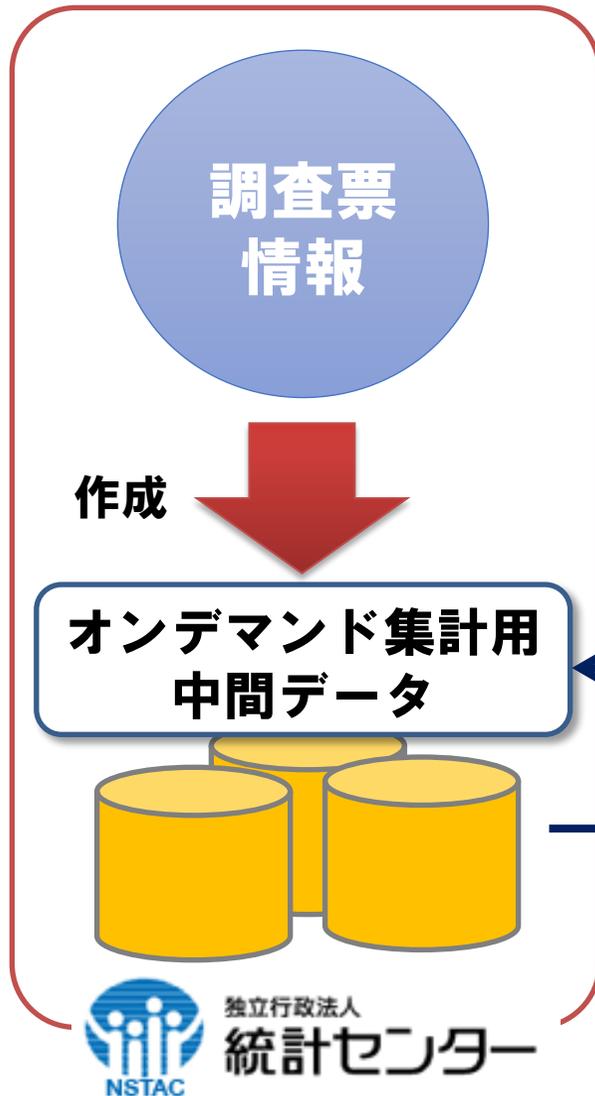
①利用者が調査項目を選択

階級(任意以上)	A 農業				B 畜産				C 漁業				D 製造業				E 建設業				F 卸売業・小売業				G 情報・通信業				H 金融業・保険業				I 学術・研究・開発業				J 教育・文化・娯楽業				K 健康・福祉業				L 業種不明																																											
	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不	総数	自派	...	下分派不																																																				
15-19歳	1260195	282123	435711	578886	977285	222123	375711	378431	282107	686876	157523	235536	209623	322652	127623	225586	178623	121164	48884	59259	220767	58488	100784	55346	185628	49488	90784	45346	35139	244893	88632	62862	85593	214807	78632	52862	85593	29786	802319	124500	180775	162863	448593	84908	150175	198830	158726	228003	22245	91360	91532	170137	12645	81300	81532	52866	178185	84636	37478	77731	141845	46836	27478	67731	38340																							
20-24歳	191131	45709	51337	127611	35709	41337	50065	69520	1185020	245179	381871	178115	902184	220179	361871	315134	282836	552344	144815	184867	178115	479697	126415	676687	168115	72747	232239	91348	80092	54887	196427	81348	70092	44887	28812	198228	30820	63136	10098	164054	20820	53136	90098	34174	131777	44247	61739	43030	119016	34247	51739	33030	12761	63176	95764	158904	422687	83764	168904	147019	200809	234761	78719	97074	24329	170122	68719	87074	14329	64639	228486	23145	103213	58445	154803	13145	93213	48445	71883	626331	182330	111462	227750	470542	152330	105462	217750	150789

②統計結果を自動的に出力

オンデマンド集計の仕組み

- ① 統計センターにおいて、調査票情報からオンデマンド集計用中間データ（データキューブ）を作成



- ② 統計利用者が集計上の分類項目（表頭・表側）を指定



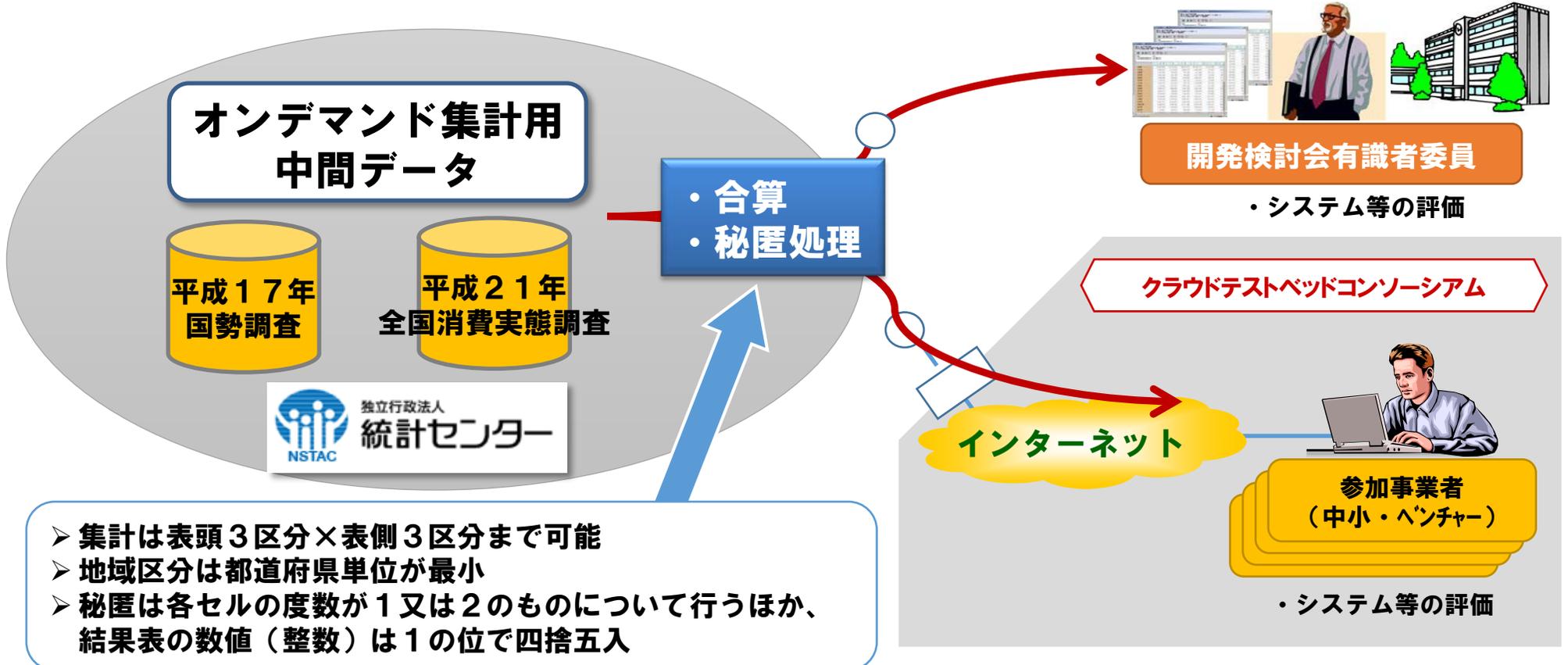
統計利用者

- ③ データキューブから必要な情報を合算し、秘匿処理を施した集計結果を出力

実証実験の概要

- ◆ 平成17年国勢調査及び平成21年全国消費実態調査の調査票情報からオンデマンド集計用中間データ（データキューブ）を実際に作成し、オンデマンドによる統計作成機能を開発
- ◆ 一般利用者（商用利用を含む）を想定し、クラウドテストベッドコンソーシアム※参加メンバー及び次世代統計利用システム開発検討会有識者委員らがシステム等を評価

※ 総務省施策「中小ベンチャー企業向け先進的クラウドサービス開発支援事業」に基づき運営されている団体。中小企業やベンチャー企業向けに仮想マシン等の提供を行っている。



委員・メンバーからの主な意見

使用感について

- ◆ 指定した集計表は特段の問題なく、正しく出力された。
- ◆ 国勢調査のデータキューブは大規模であるため演算に時間を要することもあったが（最大で30分程度）、全国消費実態調査についてはストレスなく結果が出力された（最大で3分程度）。

秘匿処理について

- ◆ 地域区分は都道府県までとなっているため、市町村や小地域での分析が行えるとなお良い。
- ◆ 集計結果数値の丸め（四捨五入）について、緩和されると良い。

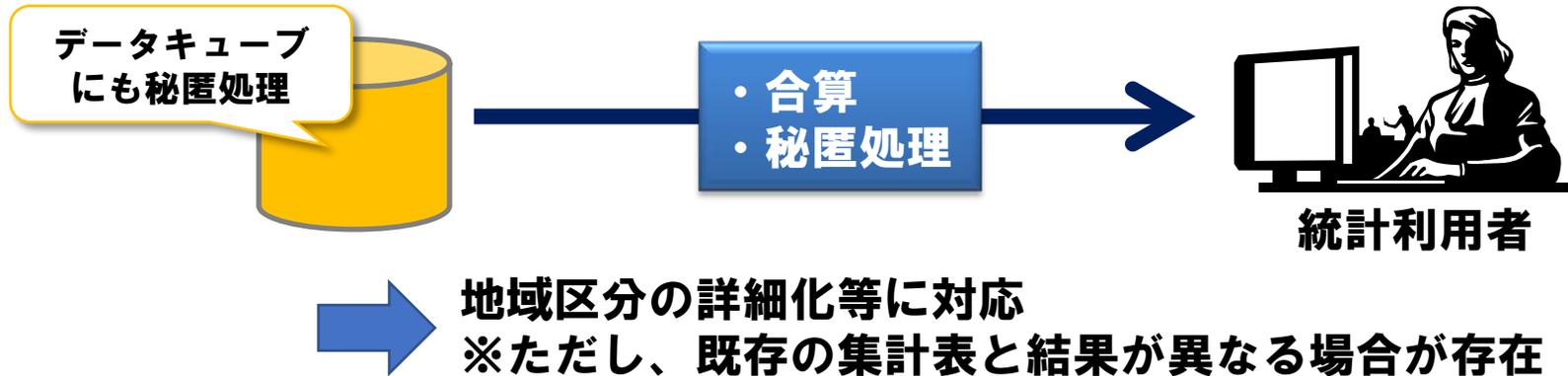
【参考】実証実験における評価を踏まえて改善した機能等

- 表示画面の大きさに応じた画面サイズの変更
- キーワードによる分類事項検索機能の追加
- 集計処理の終了を通知する機能の追加
- 集計処理時間の改善
- ダウンロードファイルの形式を追加（従来のCSV形式に加え、XML形式での出力も可能に）

今後の取組内容（案）

新しい秘匿手法の検討

有用性を保持しつつ秘匿性を高めるため、集計結果に対する秘匿処理だけでなく、データキューブに対する秘匿処理についても検討



利用者のニーズに応じた秘匿方法の検討

利用目的等によって、集計の細かさや秘匿に対するニーズが異なると考えられることから、適切な秘匿方法を検討

e-Statとの連携

利用者が既存の集計表と同内容の集計を行うよう指示を出した場合に、e-Statから該当する統計表を参照する機能の開発について検討

【参考】オンデマンド集計用中間データ（データキューブ）

調査票情報から、調査事項（集計事項）における回答パターンの組合せを集積し、個別識別性を除去した粒度の細かなミクロ集計データ

調査票情報

	調査事項 A	調査事項 B	調査事項 C	...	調査事項 E	乗率
被調査者1	1	1	1		2	25
被調査者2	2	2	1		1	30
被調査者3	1	1	2		5	40
被調査者4	1	1	1		1	15
被調査者5	1	1	2		1	15
被調査者6	2	1	2		1	15
被調査者7	2	2	1		2	10
被調査者8	1	1	2		2	5
被調査者9	1	1	1		3	20
被調査者10	3	1	2		4	30
被調査者11	3	1	2		4	30
被調査者12	1	1	1		2	20
被調査者13	2	1	2		5	20
被調査者14	4	1	2		5	15
被調査者15	4	1	2		2	15

加工・集計

オンデマンド集計用中間データ（データキューブ）

集計区分A	集計区分B								
	総数	集計区分C		集計区分C		集計区分C		総数	
		1	2	1	2	1	2		
総数	305	120	185	265	80	185	40	40	-
1	140	80	60	140	80	60	-	-	-
2	75	40	35	35	-	35	40	40	-
3	60	-	60	60	-	60	-	-	-
4	30	-	30	30	-	30	-	-	-

クロス統計のデータ保持構造

	調査事項 A	調査事項 B	調査事項 C	集計値
セルレコード1	1	1	1	80
セルレコード2	2	2	1	40
セルレコード3	1	1	2	60
セルレコード4	2	1	2	35
セルレコード5	3	1	2	60
セルレコード6	4	1	2	30

統計（集計結果）

【参考】海外におけるオンデマンド集計の事例

ベースとなるデータ	概要	採用例
事前集計済み結果表	事前に利用が想定できる分類事項の組合せで集計を行い、あらかじめ用意した組合せによる集計表のみを提供	オランダ統計局「StatLine」 香港統計局「Interactive Data Dissemination System」
超高次元クロス集計表	すべての分類事項について多次元クロス集計した単表を、超高次元クロス集計表としてシステムに格納	豪州統計局「Census CData Online」
マイクロデータ	マイクロデータファイルから検索・作表	米国センサス局「Advanced Query System」 豪州統計局「Census Table Builder」